SNIA

Education

HIGH AVAILABILITY AND DISASTER RECOVERY FOR NAS DATA

Paul Massiglia
Chief Technology Strategist
agámi Systems, Inc.

Link to abstract

Link to SNIA legal notice

Data protection and high availability

- Data protection
  - A copy of critical data is intact after the disaster event
  - Restoring application and client access is treated as a separate problem
- High availability
  - A copy of critical data is intact after the disaster event
  - Procedures (usually automatic) are in place to restore application and client access

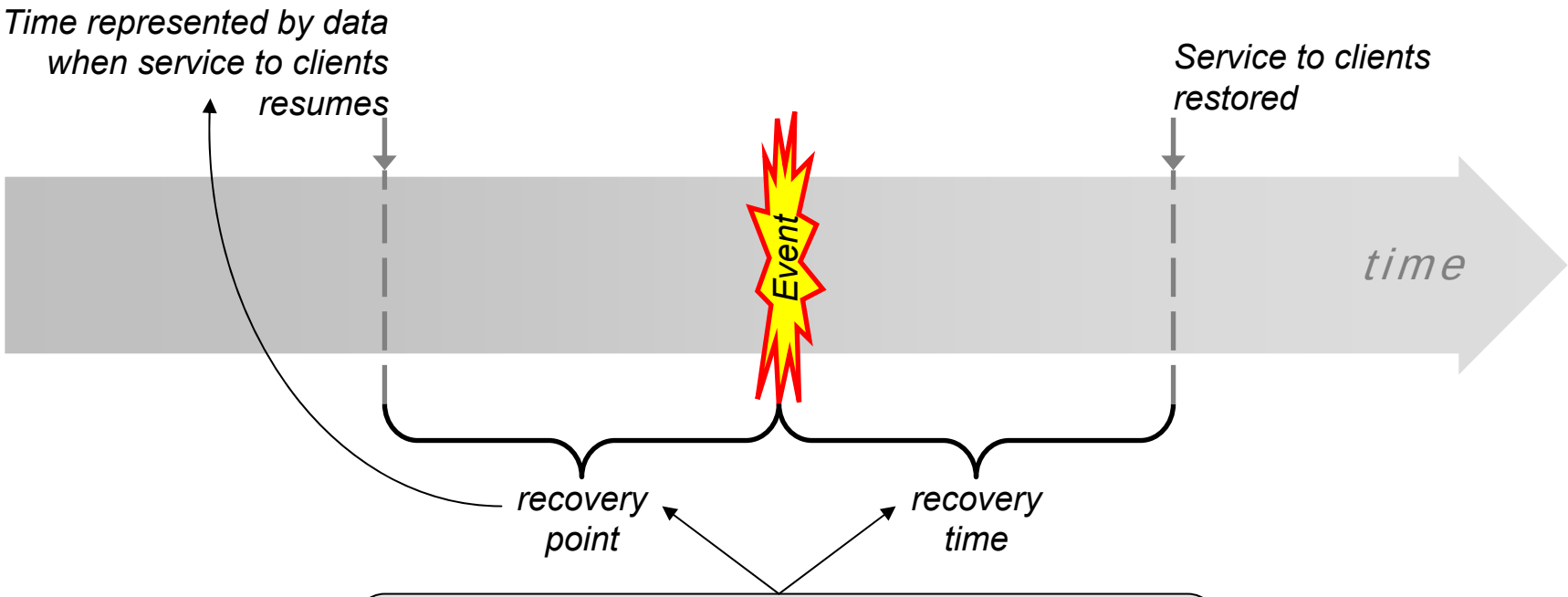
With high availability becoming more common, and continuous data protection (CDP), the distinction is becoming less clear

Measuring availability

Education SNIA

Link to RPO details

Link to RTO details



Business objectives for these measures determine appropriate data recovery and availability techniques

In general, ↓RTO-RPO↓ ⇒ ↑$$$↑

Disasters

- Logical (virtual?)
  - Equipment and facilities remain physically intact
  - Data has been destroyed or corrupted

- Physical
  - System failure
    The surrounding environment is intact
  - Data center loss
    The entire IT environment must be recreated

Logical disaster

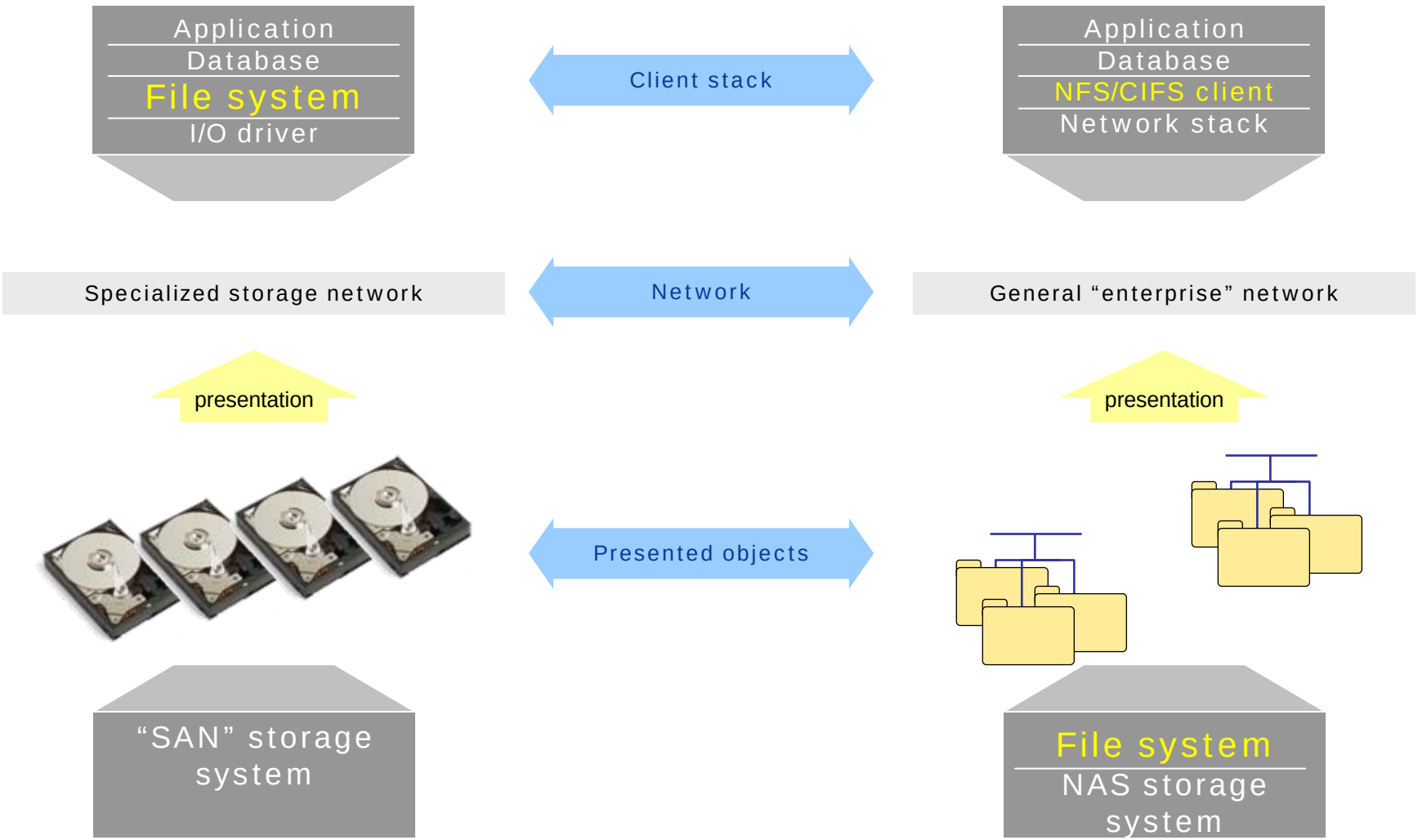
- Causes
  - Malice (malware or human action)
  - Innocent error
- General recovery strategy: “turn back to clock”
  - Restore a known good copy of data (i.e., that precedes the corrupting event)
  - Requires periodic copies stored in a safe place
  - Inevitability: recreate updates that occurred between recovery point and time of fault discovery

A “good copy” is…

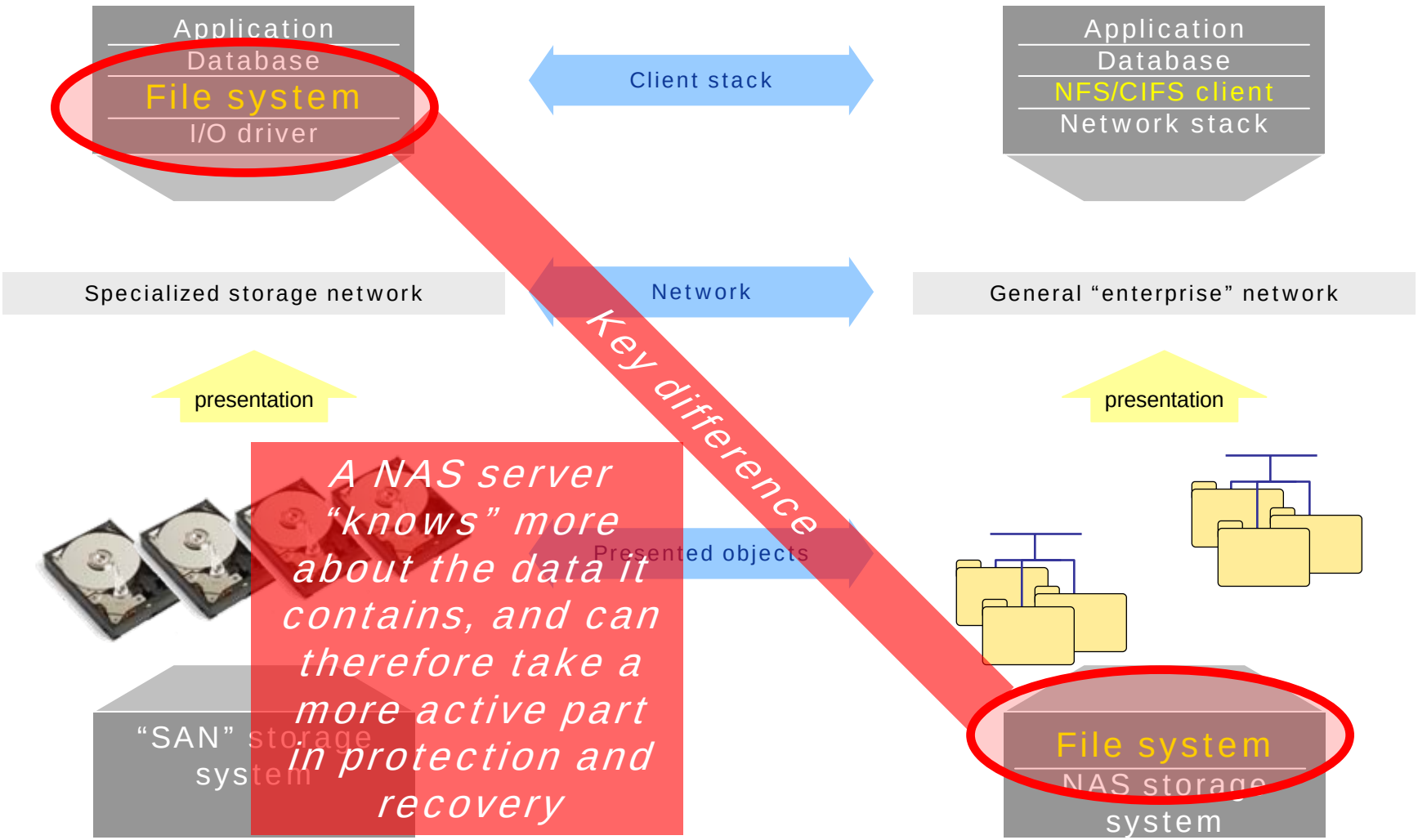
- Consistent
  - Represents a “snapshot” of the data set
- Modular
  - Is restorable en masse or file-by-file
- Unobtrusive
  - Creating it doesn’t bring applications to their knees
- Near-current
  - Isn’t missing much data when restored
- Durable
  - Is restorable after years or decades

snapshot **CONTEXT [Data Recovery] [Storage System]**
A fully usable copy of a defined collection of data that contains an image of the data as it appeared at the point in time at which the copy was initiated. A snapshot may be either a duplicate or a replicate of the data it represents.
http://www.snia.org/education/dictionary/s/

What's unique about NAS ?



What's unique about NAS ?

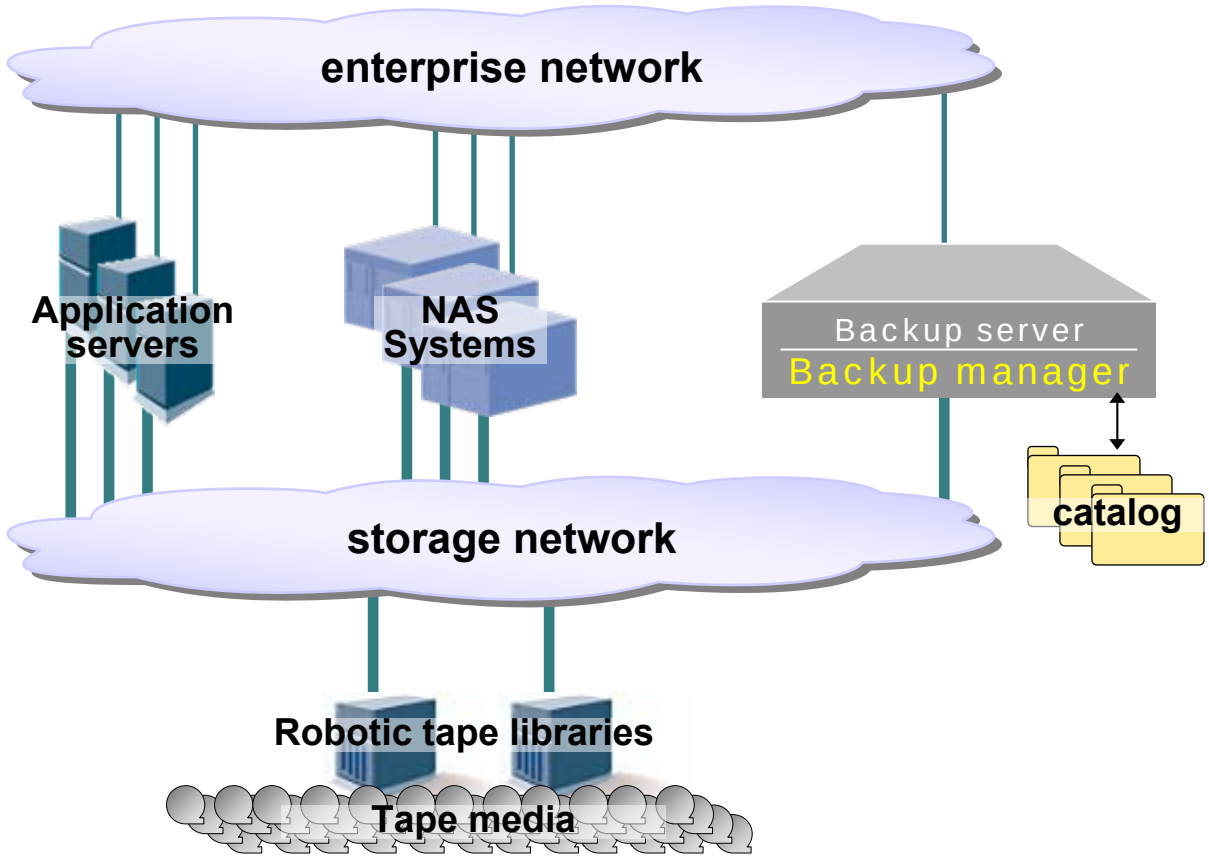


Techniques for backing up NAS data
`cp -r -x * /backup_location`

Advantages	Limitations
Nothing to buy	Bandwidth-intensive
No specific configuration or training	Platform-specific
Long-term stability of data formats and access protocols (NFS/CIFS)	Administration-intensive: data sets, backup devices, and media must be hand-selected and managed

Techniques for backing up NAS data
Why use backup management software ?

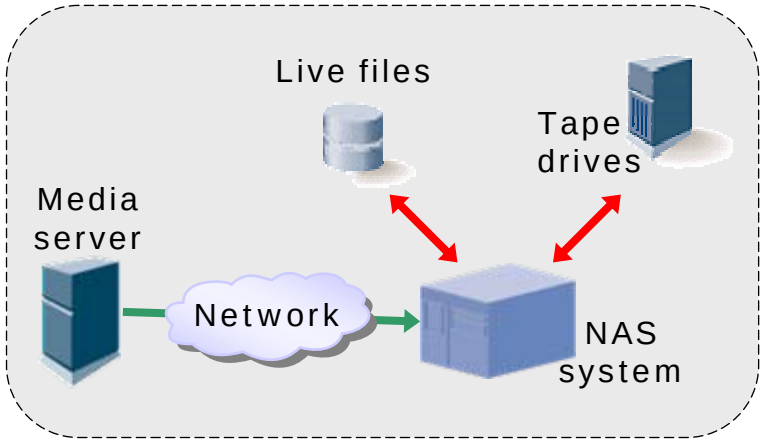
- In a word, management
  - Data
  - Schedule
  - Device
  - Data flow
  - Media

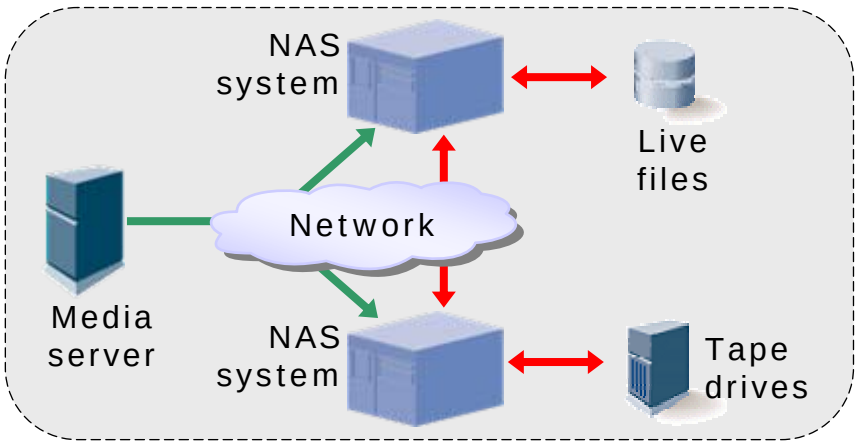


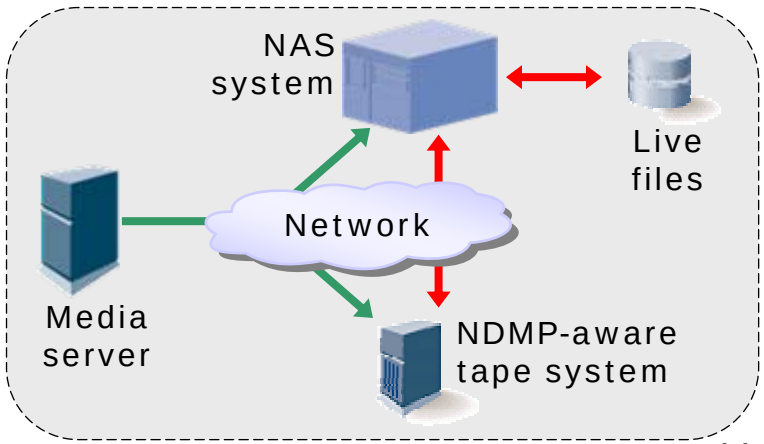
Link to +/— details

In another word, Network Data Management Protocol (NDMP)

- Augments backup management software
  - Minimizes network traffic
- NDMP roles
  - Data source
  - Storage destination
  - Control point







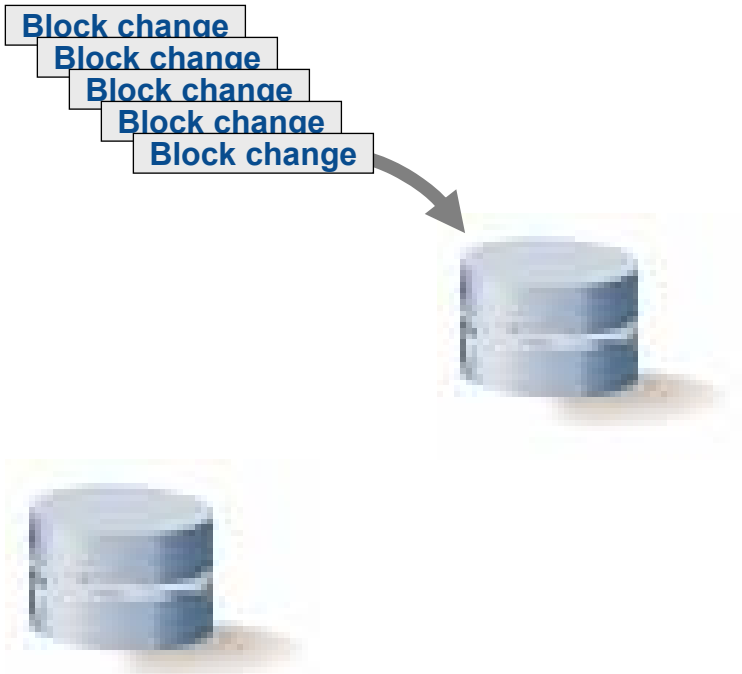
Backup to disk: presentation options

- Virtual tape library (VTL)
- Emulates
  - Tape drives
  - + Media
  - + Robotic loader

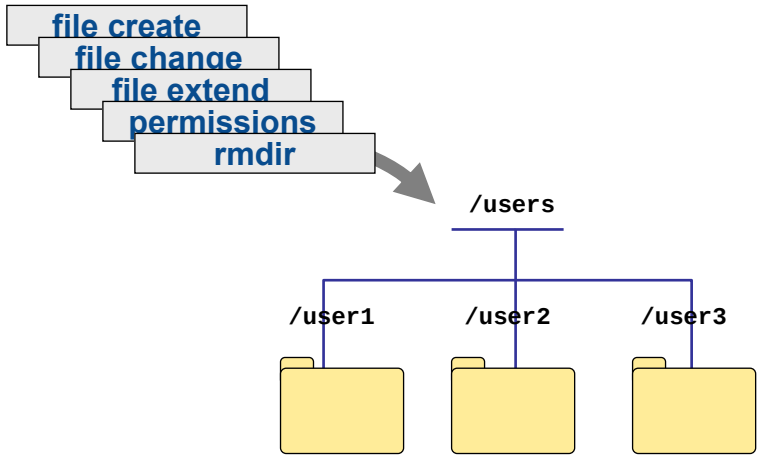
- On-line storage "containers"
- Managed by
  - backup management software

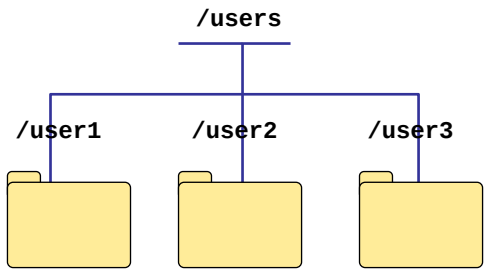
Link to VTL details

Snapshots: making data “stand still”



Frozen image of logical blocks

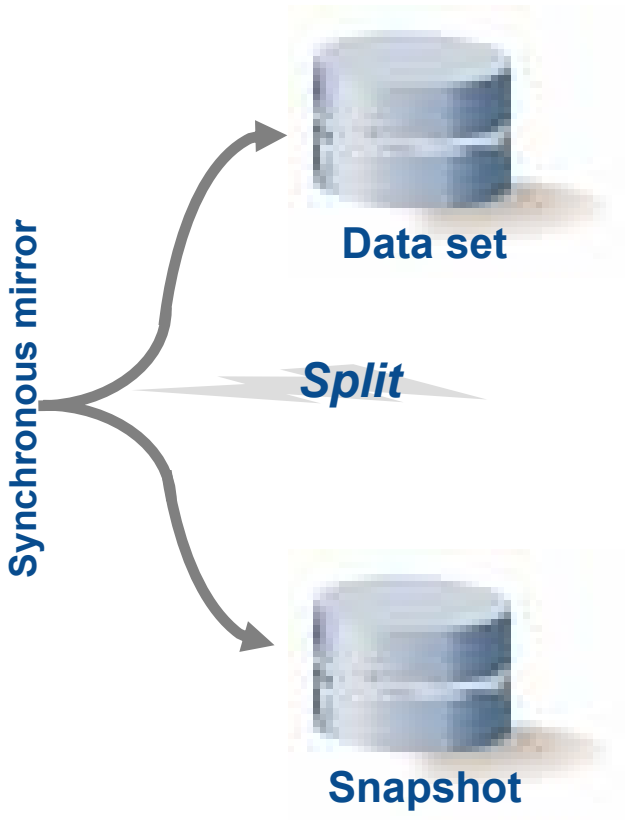




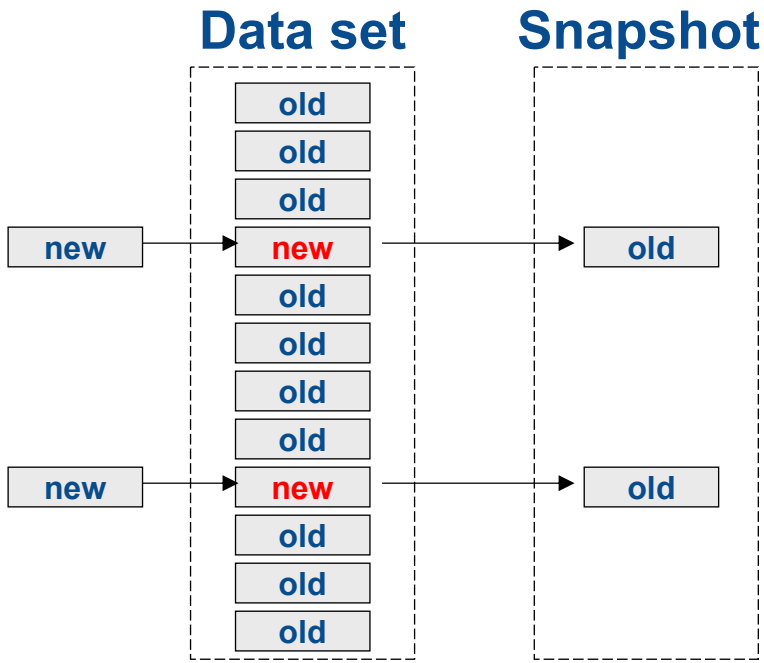
Frozen image of directory tree

Link to +/— details

Full-size and space-saving snapshots



Pretty much limited to block-level snapshots (which may contain file systems)



copy on write
CONTEXT [Storage System, Backup]
A technique for maintaining a point in time copy of a collection of data by copying only data which is modified after the instant of replicate initiation. The original source data is used to satisfy read requests for both the source data itself and for the unmodified portion of the point in time copy. *cf.* pointer remapping

http://www.snia.org/education/dictionary/c/

Link to +/— details

Which kind of snapshot is best ?

- It depends on your definition of “best”
- Full-copy (“split mirror”) snapshots
  - Very common with SAN (not NAS)
  - Zero-impact on production data
  - Greatest value is in off-hosting ability
- “Copy-on-write” snapshots
  - Make it feasible to maintain many active snapshots
  - Some impact on application write performance

Backup may not be so cheap

Conventional backup	Cost advantage	Replication
Tape drive, media, and library	←	Online storage capacity
Physical transportation	←	Replication link with adequate bandwidth and latency
Acquisition or activation of recovery site	←	Recovery facility premises and staff
Value of downtime to restore data (hours to days)	→	Seconds to minutes
Value of data lost due to recovery point (hours to days)	→	Zero to seconds

Disasters

- Logical (virtual?)
  - Equipment and facilities remain physically intact
  - Data has been destroyed or corrupted

- Physical
  - System failure
    The surrounding environment is intact
  - Data center loss
    The entire IT environment must be recreated

System failure

- Definition:
  - Failure that is beyond the reach of the classic recovery techniques:
    - RAID and mirroring
    - Active/active or active/passive “dual heads”
- Needed to recover:
  - An intact and accessible copy of live data
  - Connectivity to application servers

Data center loss

- Definition
  - Entire IT environment is incapacitated
    - Storage systems
    - App servers
    - Local clients
    - Connectivity to remote clients
- Needed to recover:
  - An intact recovery site outside the area impacted by disaster
  - IT staff including provisioning (transportation, safety, food, etc.)
  - Adequate hardware and connectivity to run critical applications
  - Reasonably up-to-date critical data accessible by recovery systems

Recovery from physical disasters

- Have or acquire adequate equipment for access to critical data and applications

- Have or recreate a copy of critical data at the recovery location

Recovery from physical disasters

- Have or acquire adequate equipment for access to critical data and applications
- Have or recreate a copy of critical data at the recovery location
- The difference between have and { acquire / recreate } is the distinction between high and "standard" availability

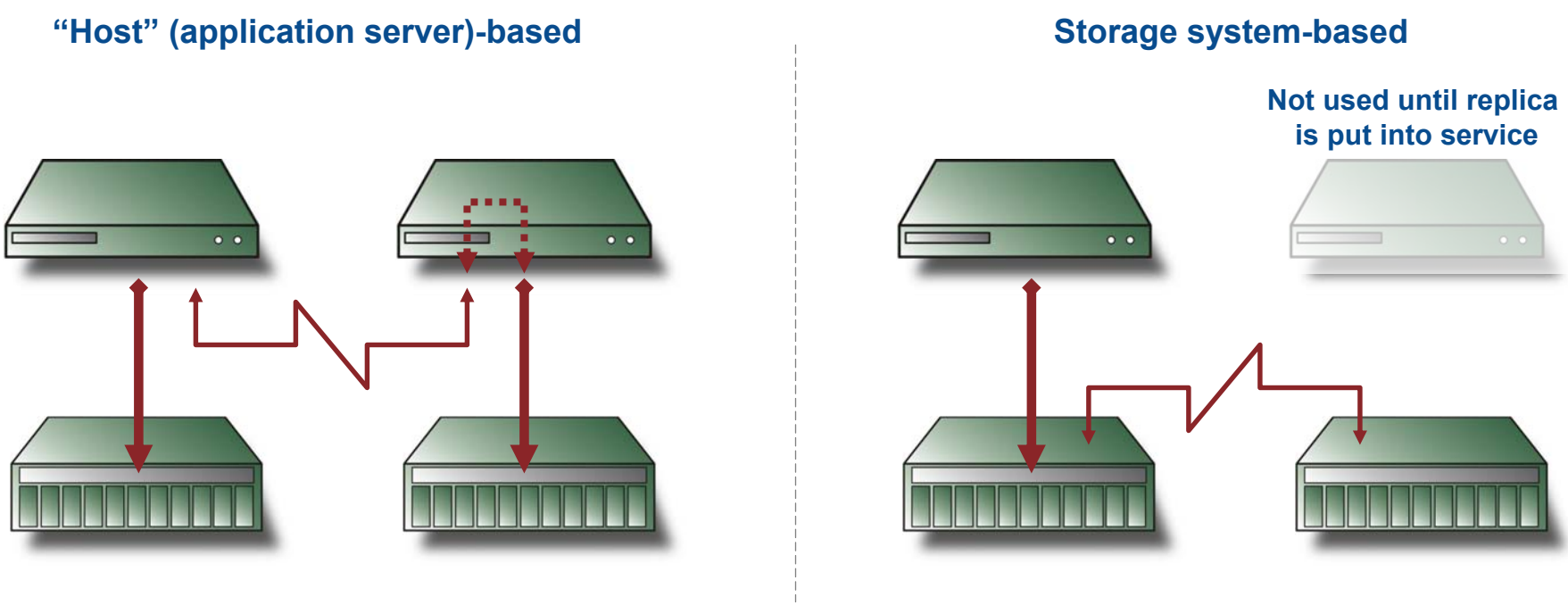
Replication

- Basic purpose: keep a (near-)current copy of a data set on separate storage resources attached to separate processing resources

- Use cases
  - Recovery from physical disaster
  - Data distribution/consolidation
  - Second data source for certain read-only applications

- Replication is not just mirroring
  - Primary-secondary relationship
  - Time ordered updates to a “consistency group” of devices
  - Asynchronous option

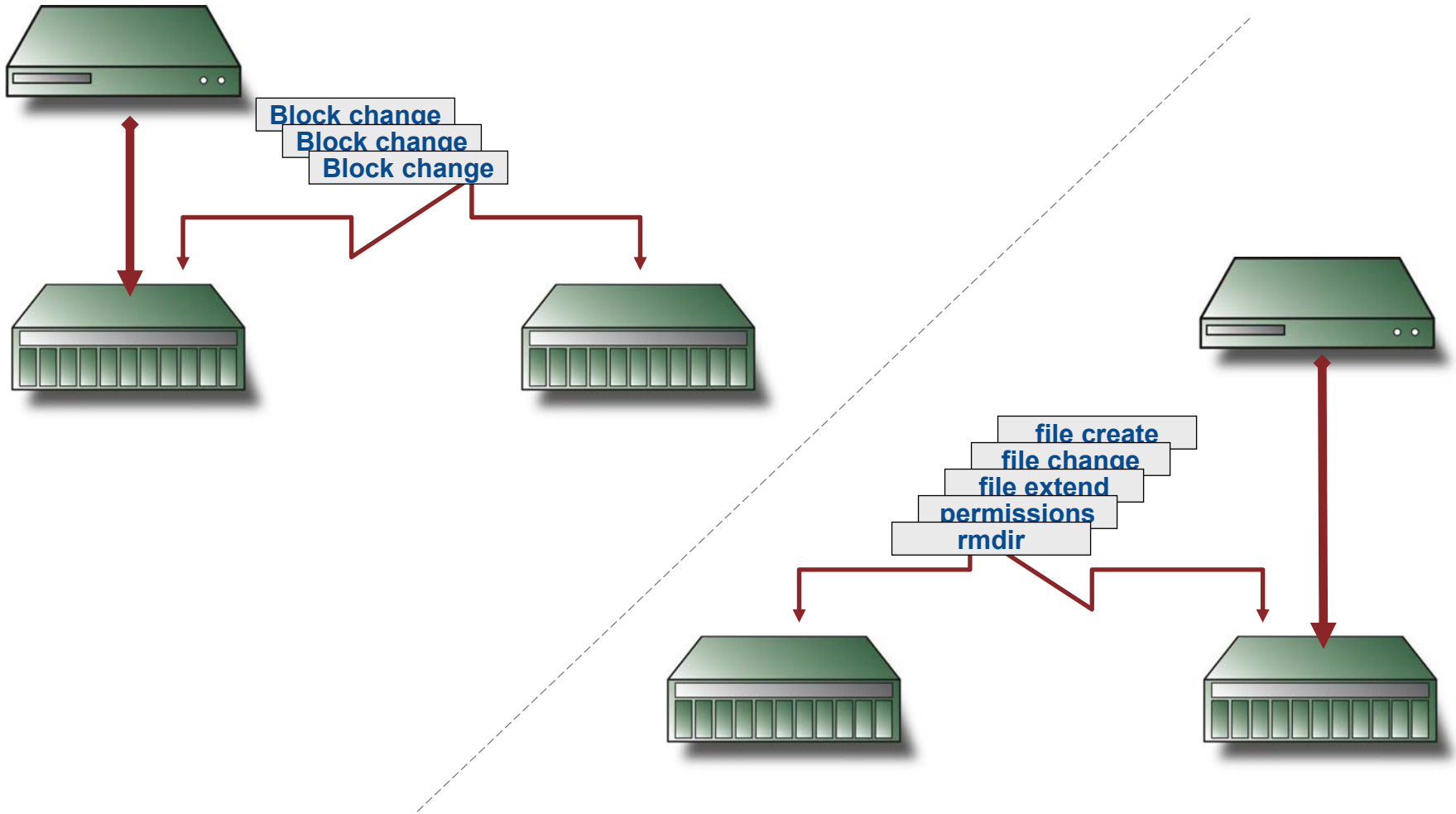
Host-based and off-host replication



- Other variations:
  - SAN switch, NAS aggregator, dedicated appliance
  - All are roughly equivalent to storage system-based replication

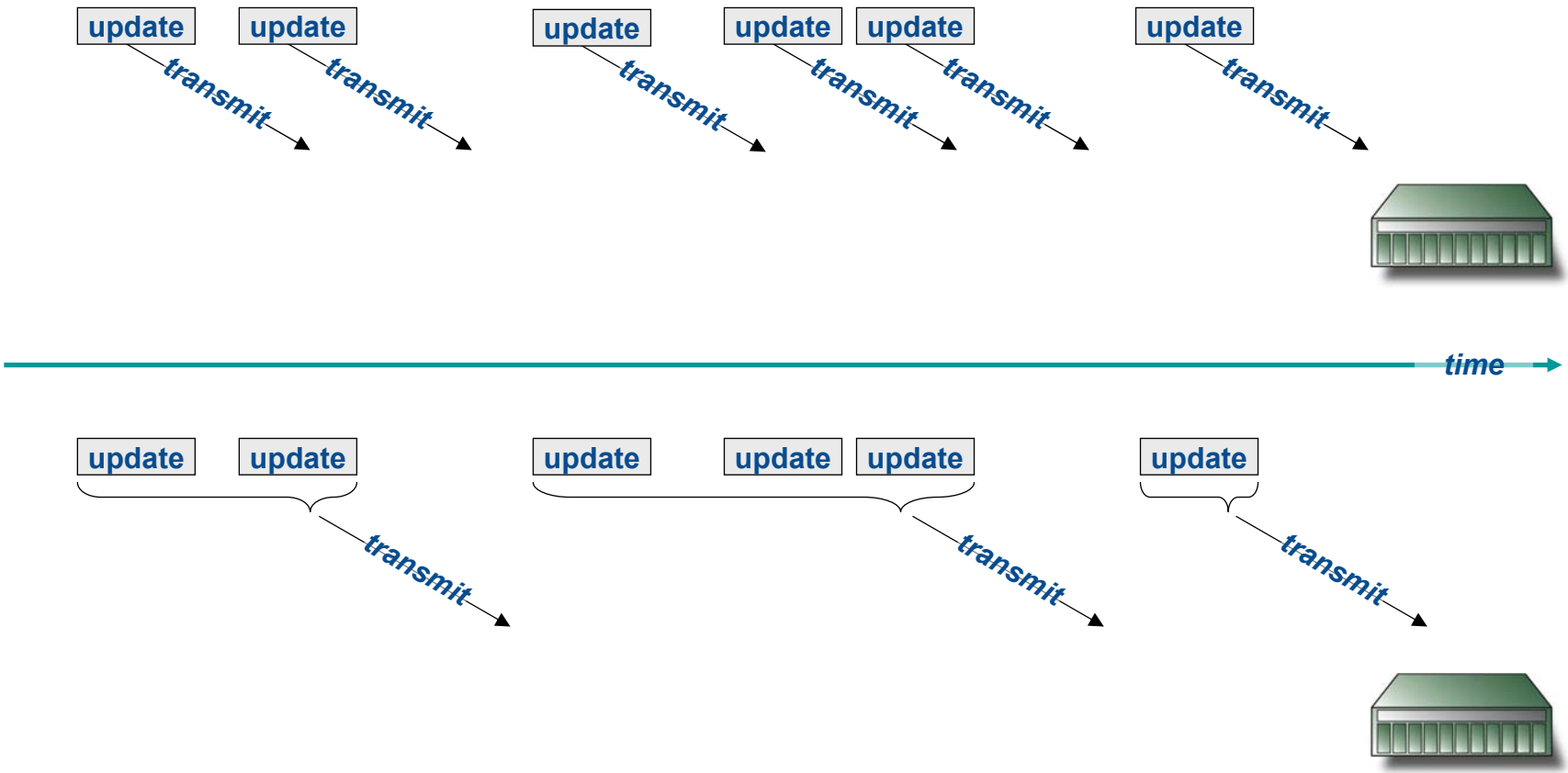
Link to +/— details

“Block” and file replication



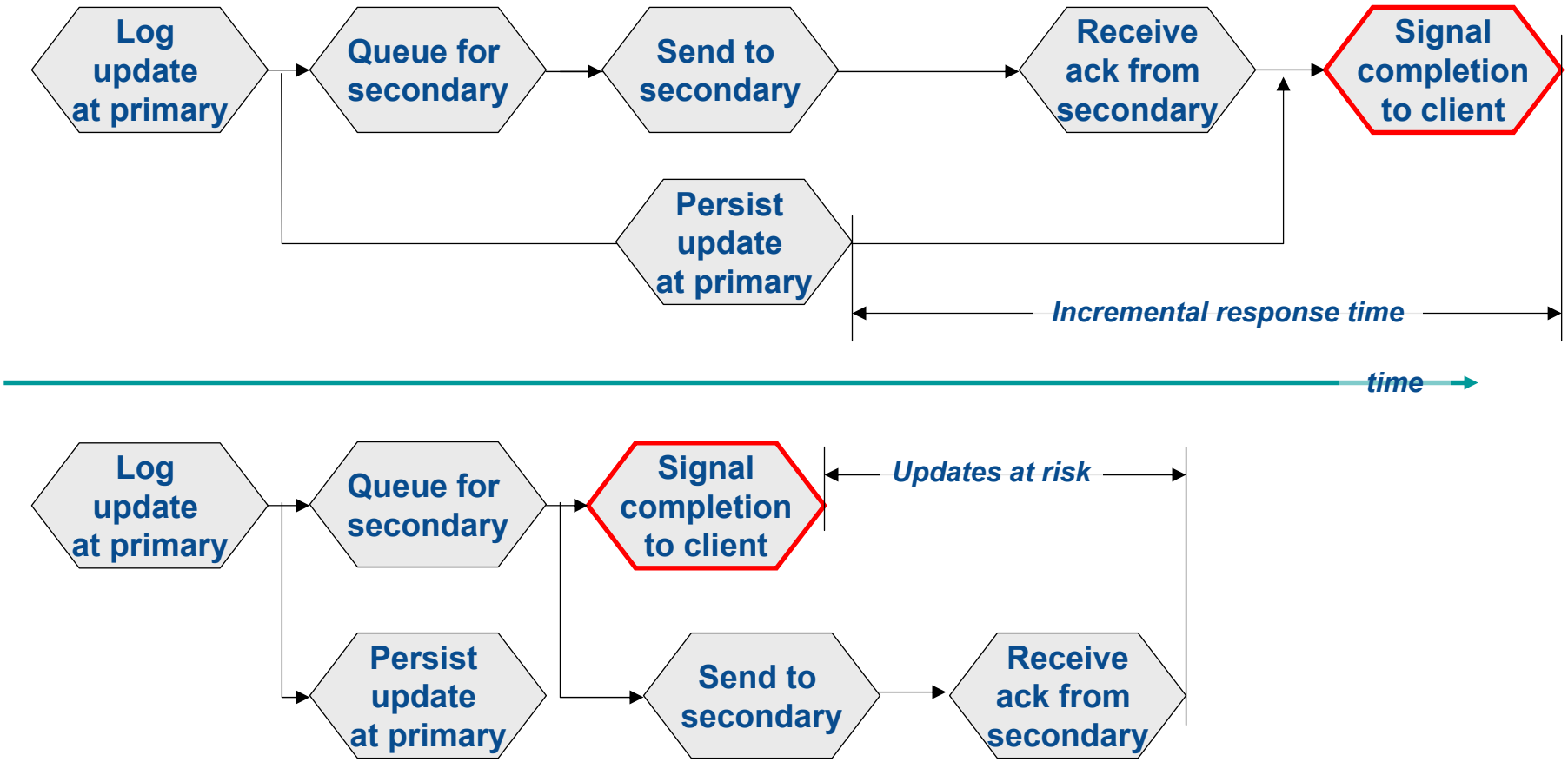
Link to +/— details

Real-time and batch replication



Link to +/— details

Synchronous and asynchronous replication



Link to +/— details

Which kind of replication is best ?

- It depends on your definition of “best”
- Block-level vs. file-level
  - Trade ubiquity against bandwidth and readability
- Synchronous vs. asynchronous
  - Trade replication link cost against client responsiveness
- Real-time vs batch update replication
  - Trade bandwidth against recovery point granularity

Beyond backup

- Basic premise
  - Record every transaction on a data set in a log
  - “Redo log for any kind of data”
- Compared to snapshots
  - Snapshot:….. images of data at fixed times
  - CDP:…….. Image of data recreated dynamically at recovery time
  - Defers specification of recovery point until restore time
- Challenges
  - Definition of “transaction” and “data set”
  - Storage consumption
  - Impact on application performance
  - Time to recreate a point-in-time data set image
  - Integration with applications and data managers
- Look before you leap

Summary

- Different threats; different data recovery techniques
- Recovery from logical disaster
  - "Turn back the clock" (recover from backup or snapshot)
  - Inherently some data loss
- Recovery from physical disaster
  - Substitute facilities (recovery site)
  - Have or recreate the best possible image of data
- Fundamentally a cost vs value decision

Q&A / Feedback

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

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

SNIA Education Committee

David Dale	Netapp
Rob Peglar	Xiotech
Warren Avery	storagenetworking.org
Norman Owens	storagenetworking.org

SNIA Legal Notice

- 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.

Abstract

As network attached storage has matured, users have entrusted increasingly critical data to it, creating requirements for protection against failures and disasters. This session will present a survey of techniques available for protecting data stored on NAS systems against loss or destruction by threats ranging from hardware and software component failure to accidental or deliberate corruption to disasters that completely incapacitate an entire data center.

Backup, RAID and mirroring, system fail-soft, snapshots, continuous and periodic replication, and NAS system clustering will all be discussed. For each technique, the threats against which it protects, the capital and operating costs, and the expected recovery time and recovery point objectives will be presented. The goal fo the session is to give students an appreciation for the high availability and disaster protection options available for their NAS-managed data, in order to better equip them to make well-informed decisions when purchasing or defining operating procedures.

Back

Availability objectives: RTO

Recovery time objective	Requirements
Days	Prepare and staff facilities and hardware Acquire backup copy of data and restore Replay logs, restore app environment and restart apps
Hours	Restore data from onsite incremental backup Replay logs, restore app environment and restart apps
Minutes	Replay logs against data replica and activate Restart apps Restore client connections
Seconds	Detect failure (avoiding false positives) Switch data replica to live mode Switch apps to live or full-service

Availability objectives: RPO

Recovery point objective	Recovery point is a consequence of the data preservation technique
Days	Time of newest backup copy
Hours	Time and location of newest incremental backup
Minutes	Amount of time by which data replica “lags” live data
Seconds	Zero

Back

Techniques for backing up NAS data
Backup management software

Advantages	Limitations
"Set-and-forget" automation Schedules Device and media management Data selection	Cost License & maintenance Training
Added-value features Incremental backup Multi-streaming & multiplexing	Requires data to "stand still" during backup
Application (e.g., database) integration	Inherently coarse-grained recovery times & recovery points

Snapshots:
what to snap …*"blocks" vs files*

Blocks	Files
Very fast creation (short time to usability)	Creation implies some metadata duplication
Block atomicity	File system operation atomicity
Unit of expansion: virtual volume	Unit of expansion: file system dependent

NAS developers have the luxury of choosing either

Back

Physical and logical snapshots

"Split mirror" (bit-for-bit copy)	"Copy-on-write" (snapshot = changes only)
Protects against device failures and media defects	**Requires separate physical protection mechanism**
Storage requirement: full data set size	**Storage requirement: ∝ change in data set size during snapshot life**
Overhead to maintain: zero	**Overhead to maintain: ≈ 2-3x for every "first write"**
Deletion cost: Resynchronization with data set	**Deletion cost: Space reclamation**

Back

Host and server-based replication

Host-based	Storage system-based
Uses app server processing resources	No processing impact on app server
Arbitrary consistency groups (e.g., volumes from multiple arrays)	Consistency group limited to storage system's scope

Back

Block and file-based replication

Block replication	File replication
Sends every block update over the replication link	Sends file system operations over the replication link (uses less network bandwidth)
Replicates file system operations I/O by I/O (i.e., is “bug-compatible”)	Performs source and target file system actions independently
No context for block updates: (Replica is not usable during replication)	File system operation-atomic (Replica can be used by read-only apps during replication)

Back

Real-time and batch replication

Real-time (op-by-op)	Batch
Sends repeated ops repeatedly	Uses network bandwidth more efficiently
Replicates every primary data set state (any app that can recover from local faults is recoverable)	May not represent all primary data set states in the replica (may affect recoverability)

Back

Synchronous and asynchronous replication

Synchronous	Asynchronous
Data update is at the replica before client I/O completes (No data lost in a disaster)	Data is queued for transmission to replica before client I/O completes
"Round trip" time to replica is added to client response time	Replication causes very little increment in client response time

Back

Outline

- High Availability and Disaster Recovery for NAS systems
- ---

- Learning objectives:
  - -deliver knowledge of the techniques available for protecting data stored on NAS systems against component failures, system failures, and wide-scope disasters
  - -improve students' ability to select and implement NAS HA/DR solutions that are appropriate to business value of data assets

- Outline

- I. The NAS HA/DR data protection problem space summary
  - -physical vs logical failures and recovery techniques
  - -protecting data vs. preserving accessibility
  - -component failures and protection/recovery techniques
  - -system failures and protection techniques
  - -data center disasters and recovery techniques

- II. Protecting data against component failures
  - -differences between SATA, SAS, and Fibre Channel disk drives
  - -what mirroring and RAID can and cannot protect against
  - -mirroring and RAID metrics--inline performance, recovery time, recovery impact
  - -matching the data protection technique to data requirements

- III. System failures and data replication
  - -replication vs mirroring
  - -forms of replication: synchronous vs. asynchronous
  - -forms of replication: continuous vs. episodic
  - -restarting applications using data replicas

- IV. Recovering NAS data from disasters
  - -what's unique about NAS data in a DR context (replica asynchrony and network addressing)
  - -NAS system failures (excepting disk drive failures)
  - -"dual-head" NAS clusters
  - -"shared nothing" NAS clusters
  - -combining HA and DR--inherent costs and properties of the dual-head and shared-nothing techniques