



Education

# How NAS Systems Participate in Data Protection

Paul Massiglia, Symantec Corporation

- 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.**

Network Attached Storage (NAS) systems are able to use their awareness of file structures to take a more proactive part in protecting data against faults and disasters than is possible with disk arrays. This session surveys common techniques for protecting data against hardware and software failures, accidental or deliberate corruption, and disasters that incapacitate entire data centers, and shows how NAS systems are able to participate actively in all forms of protection. Backup, snapshots, continuous and periodic replication, and continuous data protection will be discussed. For each technique, the threats it covers, the costs of using it, and expected recovery times and recovery points will be pointed out. The goal of the session is to give students an appreciation for the high availability and disaster protection options available for data stored in NAS systems, to better equip them to make informed decisions when purchasing equipment or defining operating procedures.

### ➤ Data protection

- ◆ An intact copy of critical data survives disaster events
- ◆ Restoring application and client access is treated as a separate problem

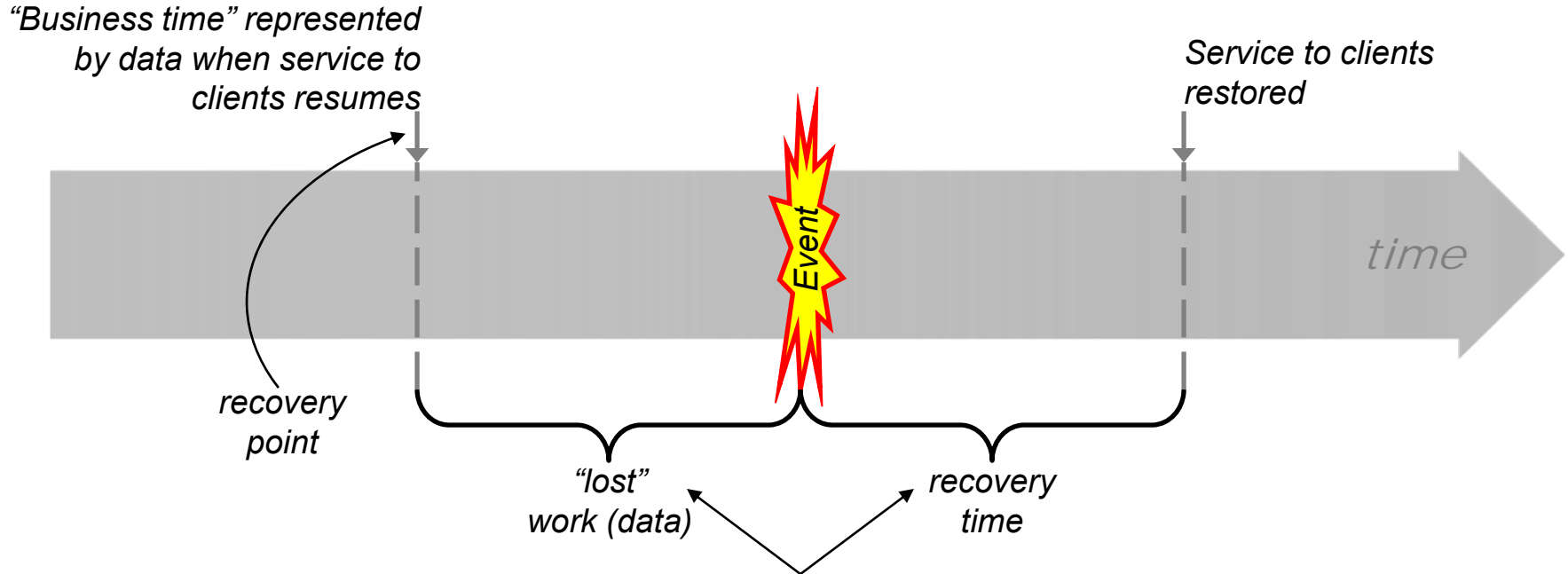
### ➤ High availability

- ◆ An intact copy of critical data survives disaster events
- ◆ Procedures (usually automatic) in place to restore service to applications and clients

# Background: Measuring availability

RPO details: slide 36

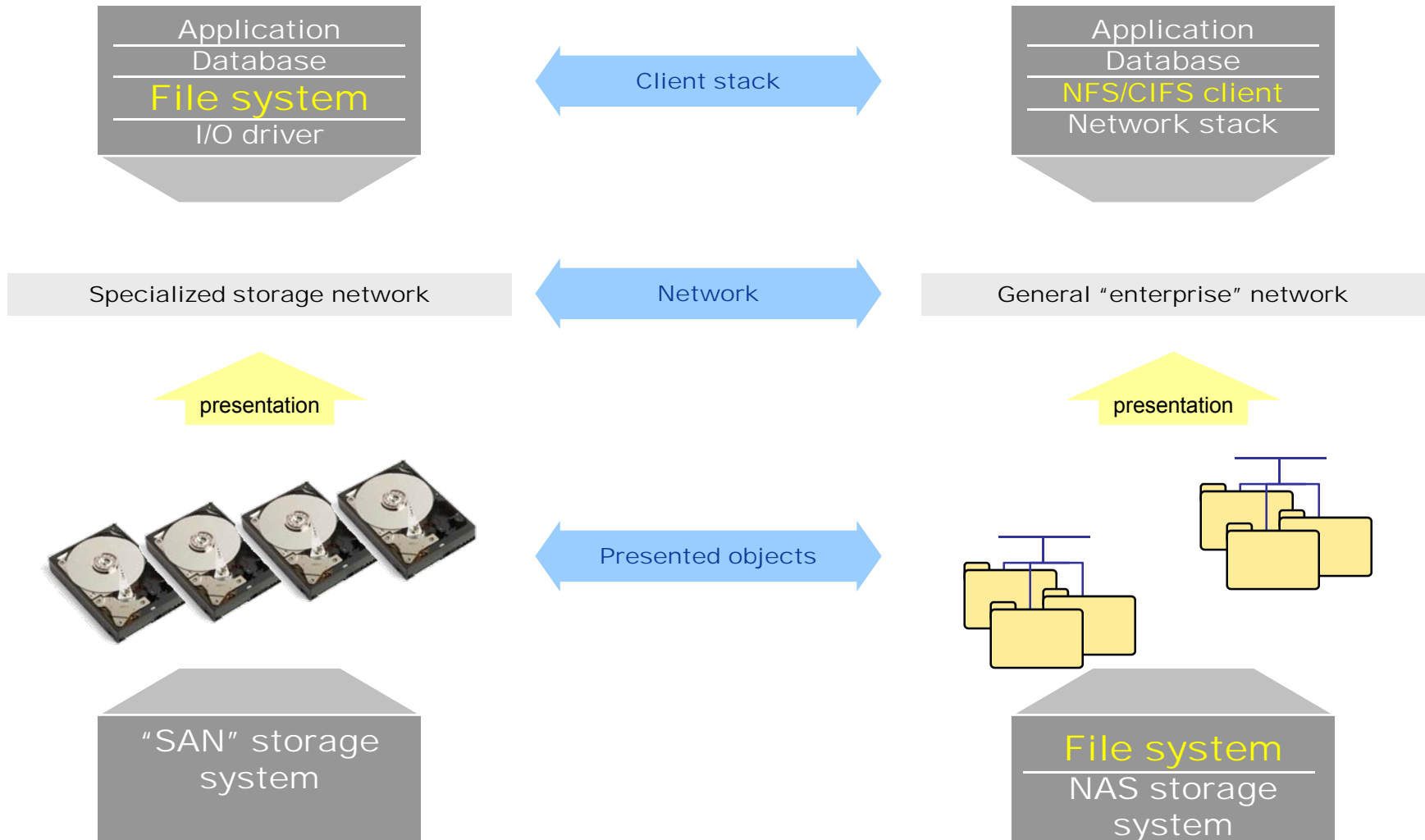
RTO details: slide 35



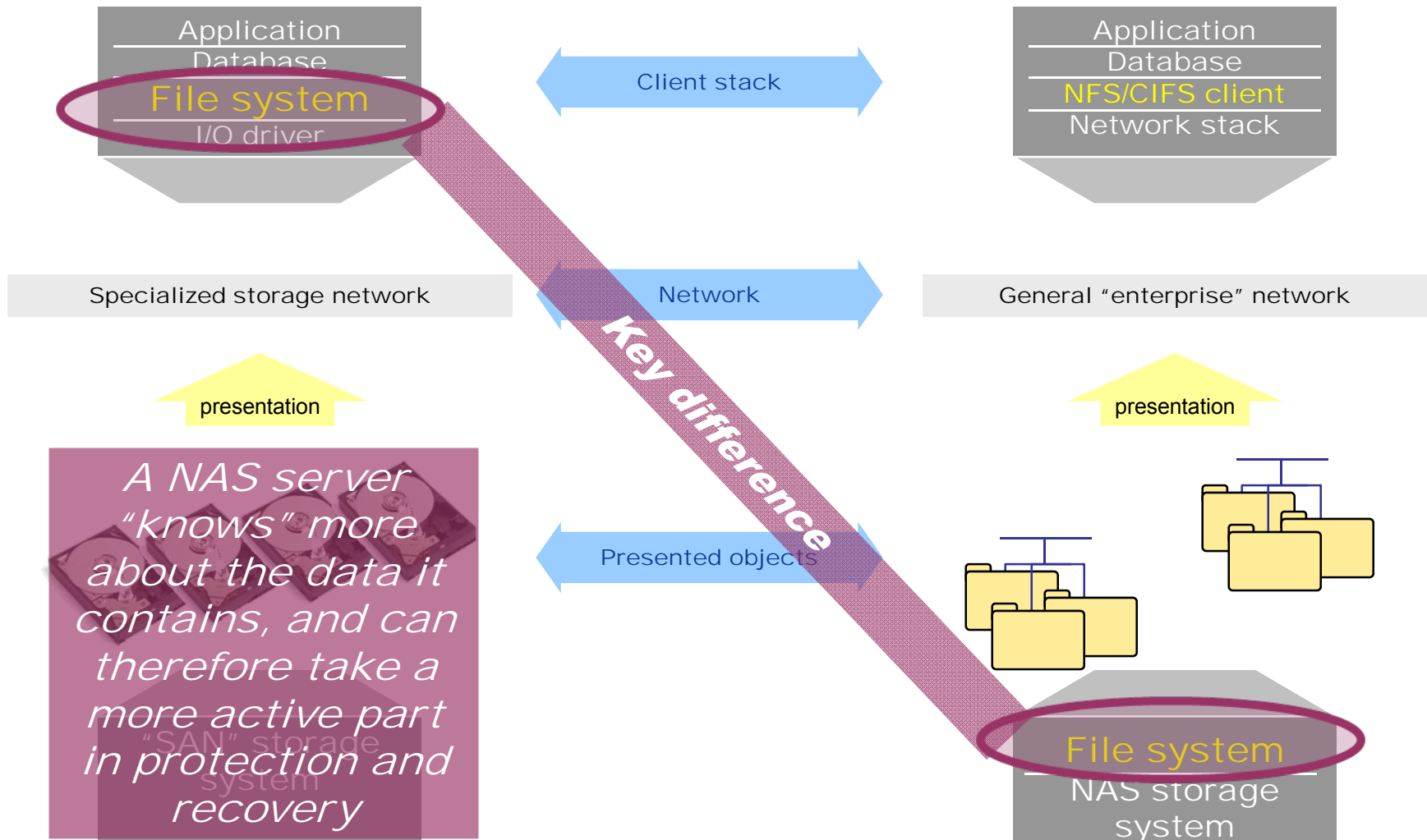
*Business objectives for these measures determine appropriate data recovery and availability techniques*

*In general,  $\downarrow RTO \& RPO \downarrow \Rightarrow \uparrow \$ \$ \$ \uparrow$*

# What's unique about NAS ?



# What's unique about NAS ?



## ➤ Logical

- ◆ Equipment and facilities remain physically intact
- ◆ Data has been destroyed or corrupted

## ➤ Physical

- ◆ System failure  
*Surrounding environment is intact*
- ◆ Data center outage  
*Entire IT environment must be restored or recreated*

## ➤ Causes

- ◆ Malice (malware or human action)
- ◆ Innocent human error
- ◆ Hardware or software fault

## ➤ General recovery strategy: “turn back the clock”

- ◆ Restore and revert to a “good” version of data

## ➤ Inherent consequences:

- ◆ Periodic copies of data sets stored in safe locations
- ◆ Updates that occur between latest recovery point and time of fault discovery are “lost”

# A “good” version of data is...

## ➤ Correct

- ◆ Is captured prior to the corrupting event

## ➤ Consistent

- ◆ Represents a “snapshot” of the data set

## ➤ Modular

- ◆ Is restorable en masse or file-by-file

## ➤ (Near-)current

- ◆ Represents a recent recovery point when restored

## ➤ Persistent

- ◆ Is restorable after years or decades

persistence

Synonym for non-volatility. Usually used to distinguish between data and [metadata](#) held in [DRAM](#), which is lost when electrical power is lost, and data held on [non-volatile](#) storage (disk, tape, battery-backed DRAM, etc.) that survives, or *persists* across power

<http://www.snia.org/education/dictionary/s/>

Technique	Properties	NAS Role
Copy files	<ul style="list-style-type: none"><li>▪ Low cost (nothing to buy)</li><li>▪ Labor intensive</li><li>▪ Fragile</li></ul>	<ul style="list-style-type: none"><li>▪ Low-cost storage for disk-based backup</li><li>▪ VTL or filer model</li></ul>
Backup management software	<ul style="list-style-type: none"><li>▪ Higher cost</li><li>▪ Robust</li><li>▪ Infrequent recovery points</li></ul>	<ul style="list-style-type: none"><li>▪ Low-cost storage for disk-based backup</li><li>▪ NDMP: direct copy from storage to backup device or system</li></ul>

**Backup details: slide 37**

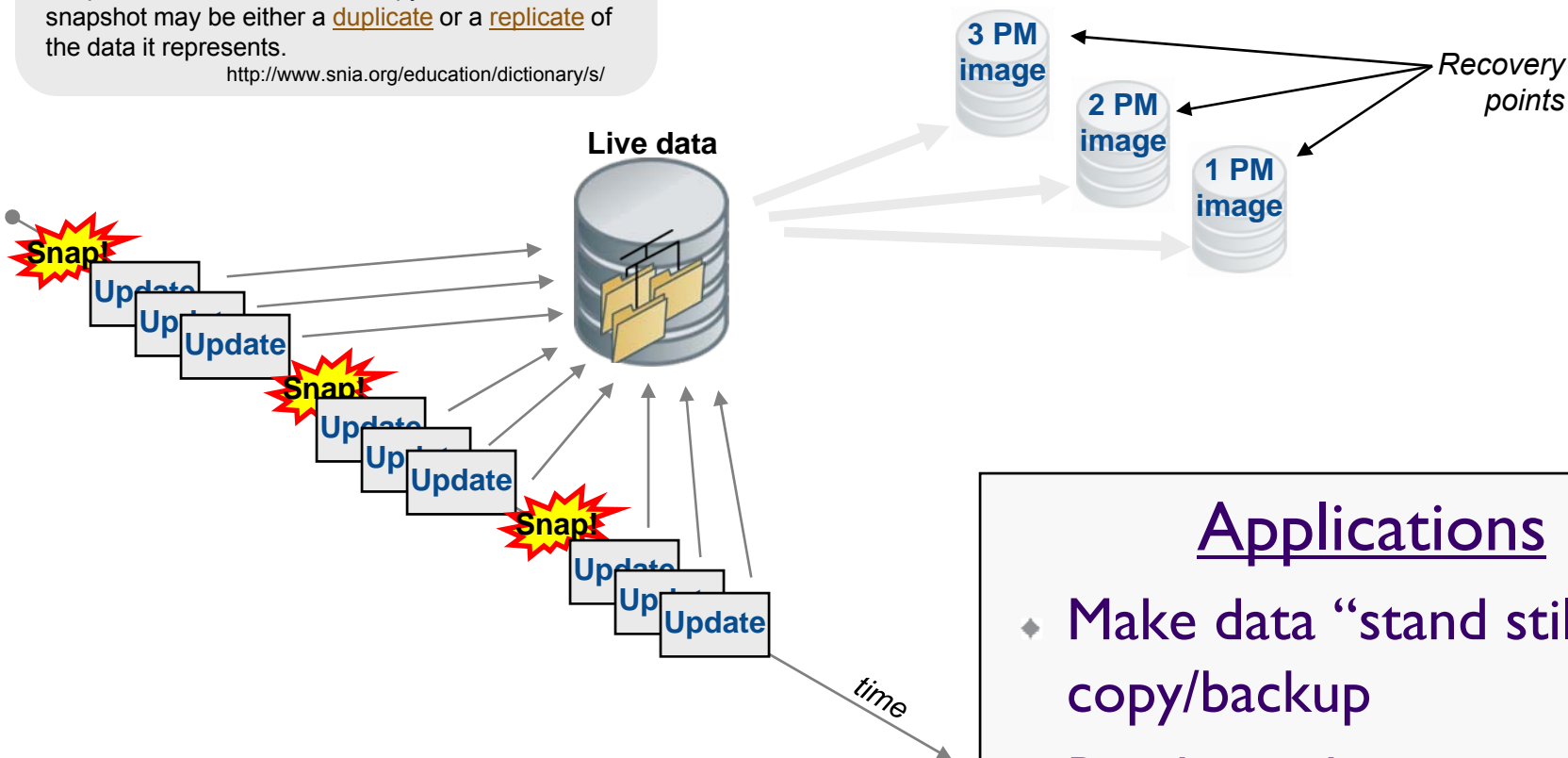
**VTL details: slide 42**

# Enabling technology: Snapshots

## 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.

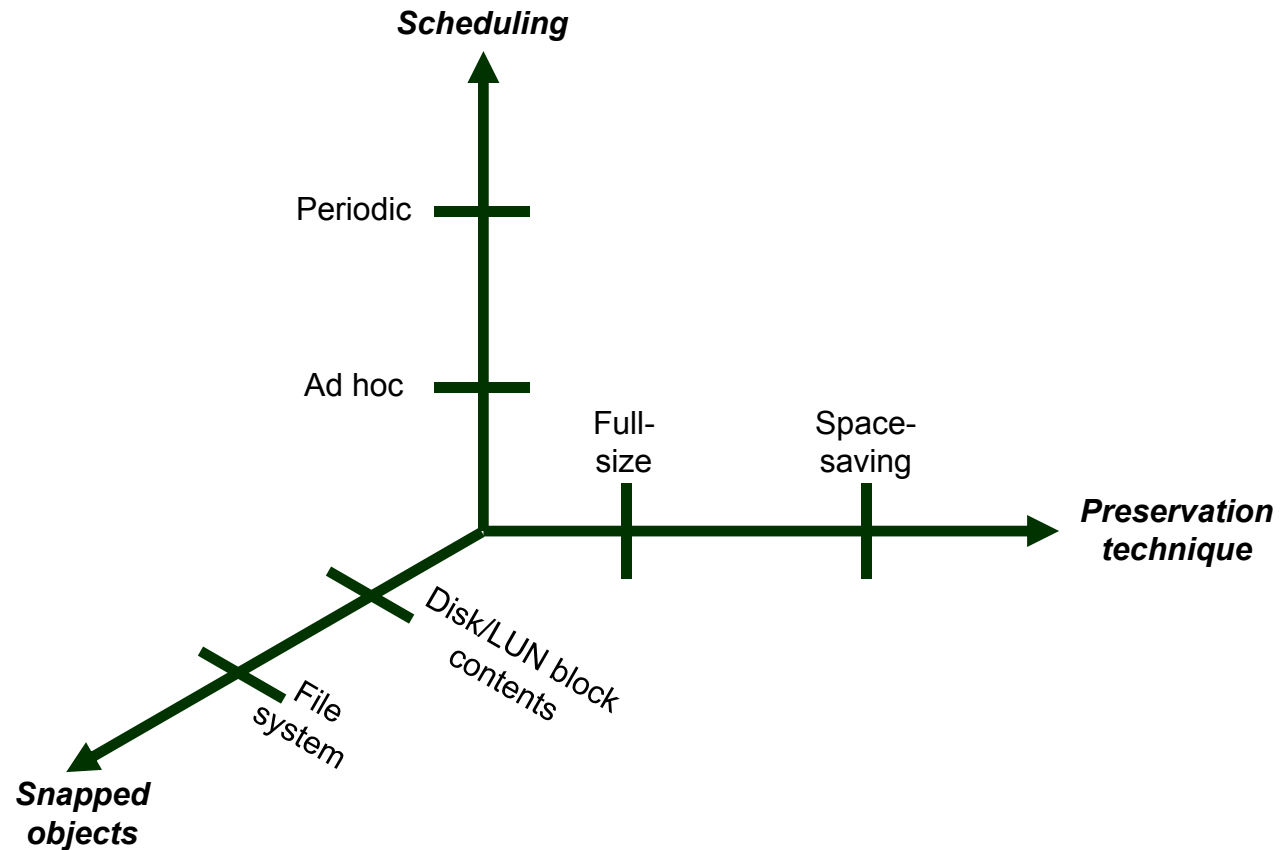
<http://www.snia.org/education/dictionary/s/>



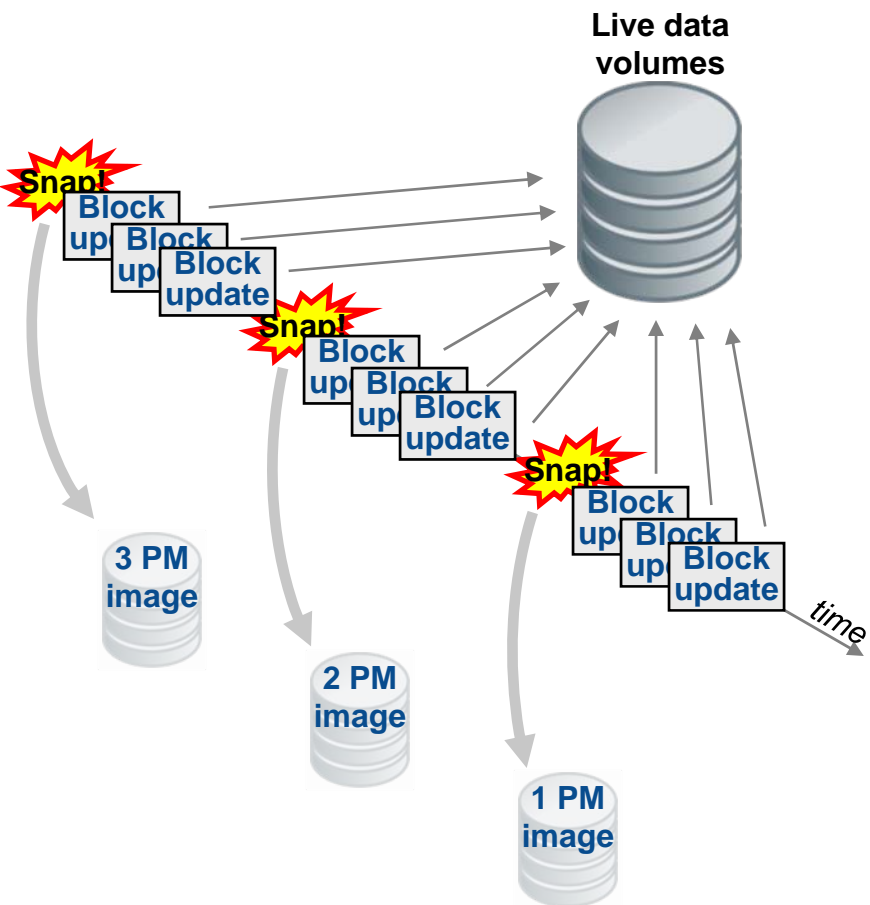
## Applications

- ◆ Make data “stand still” for copy/backup
- ◆ Ready-made recovery from logical corruption

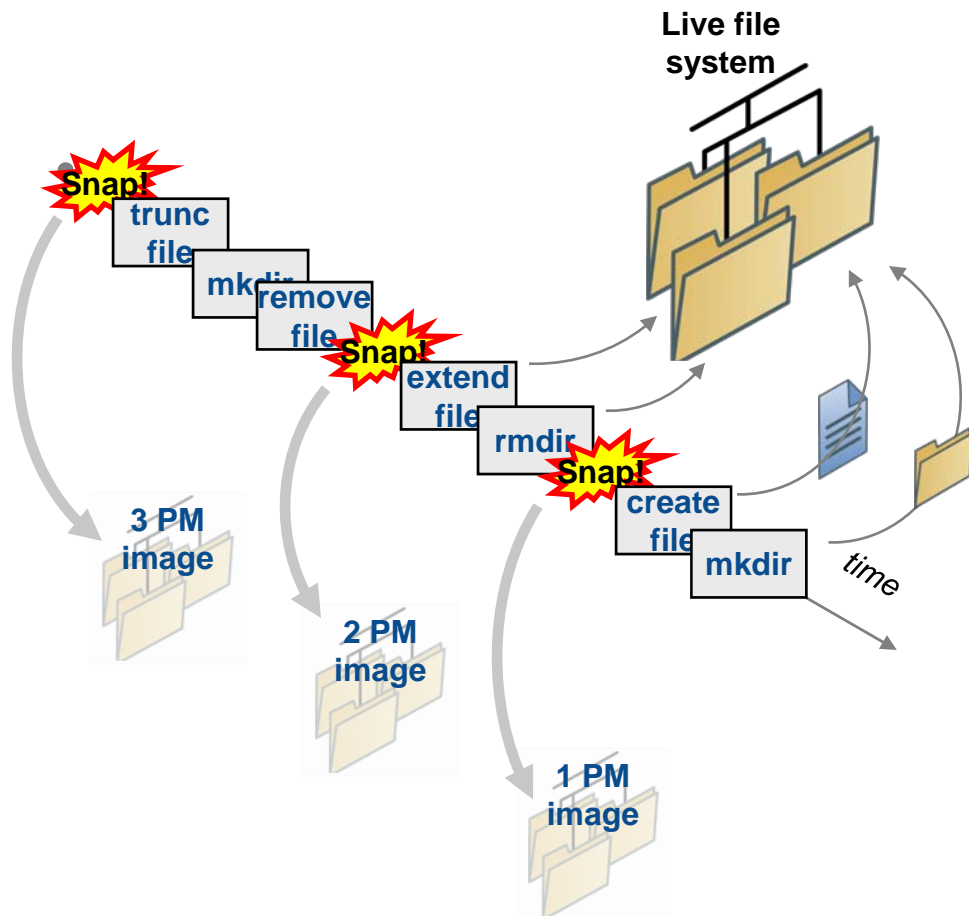
# The three basic snapshot choices



# Block vs. file system snapshots



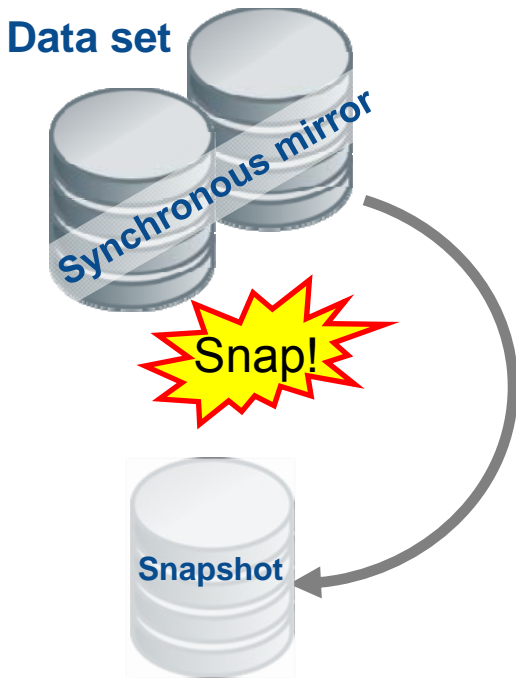
Point-in-time images of logical volume contents



Point-in-time images of directory tree contents

Compare properties: Slide

# Full-size vs. space-saving snapshots

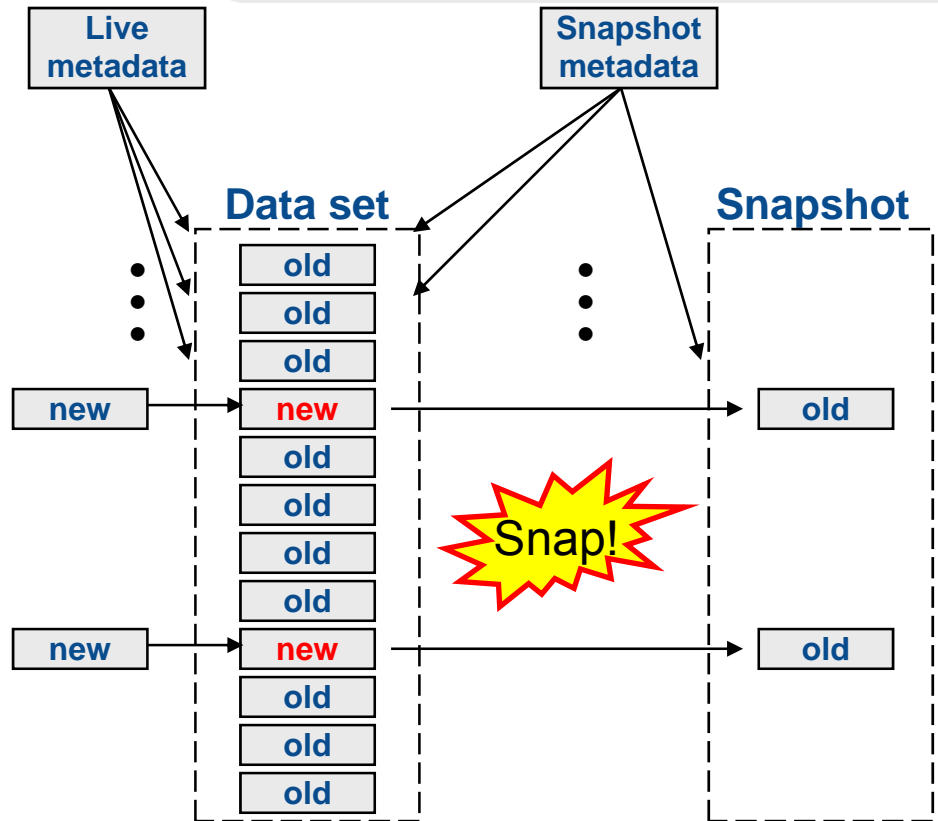


copy on write

### 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/)

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



# Which kind of snapshot is best ?

- It depends on your definition of “best”
  
- Full-size (“split mirror”) snapshots
  - 👍 Protect against physical destruction of data set
  - 👍 Enable “off-hosting”
  - 👍 Almost no impact on production data
  - 👎 High cost in storage and re-synchronization
  
- Space-saving snapshots
  - 👍 Space-saving (⇒ frequent recovery points)
  - 👎 Some (usually minor) impact on application write performance
  - 👎 Do not protect against physical destruction of data set

# Which kind of snapshot is best ?

- It depends on your definition of “best”
  
- File system snapshots
  - 👍 Ready-to-use as read-only file systems
  - 👎 File-system specific
  - 👎 Certain performance anomalies (e.g., bulk deletions)
  
- Disk (LUN) snapshots
  - 👍 Universal: snap any kind of data
  - 👍 Coordinated snapshots of multiple name spaces
  - 👎 Require external declaration of “sync points”

# Which kind of snapshot is best ?

- It depends on your definition of “best”
  
- Copy-on-write snapshots
  - 👍 Favor live file performance
  
- Pointer-mapped snapshots
  - 👍 Favor snapshot performance
  
- Why not a hybrid?
  - 👍 Why not, indeed?
  - 👍 Some vendors are at least experimenting with the concept

# The data protection game changer: *Continuous Data Protection (CDP)*

## ➤ Fundamental technique

- ◆ Log every transaction on a data set (a universal “redo log”)

## ➤ Fundamental benefit

- ◆ Defers specification of recovery point until restore time

## ➤ Compared to snapshots

- ◆ Snapshots:.....Pre-stored images of data taken at fixed times
- ◆ CDP:.....Image of recovery point data created dynamically at recovery time

## ➤ Challenges

- ◆ What’s in a “transaction” and a “data set” ?
- ◆ Resource consumption
- ◆ Impact on application performance
- ◆ Time to identify and present a point-in-time data set image
- ◆ Integration with applications and data managers

Continuous Data Protection – CDP

### **CONTEXT [Data Recovery]**

A data protection service that captures changes to data to a separate storage location. There are multiple methods for capturing the continuous changes involving different technologies that serve different needs. CDP-based solutions can provide fine granularities of restorable objects ranging from crash-consistent images to logical objects such as files, mail boxes, messages, etc.

<http://www.snia.org/education/dictionary/s/>

	<b>System failure</b>	<b>Data center loss</b>
<b>Definition</b>	<p>Failure beyond the reach of usual techniques like...</p> <ul style="list-style-type: none"> <li>▪ RAID/mirroring</li> <li>▪ NAS head “clustering”</li> <li>▪ Network redundancy</li> </ul>	<p>Entire IT environment incapacitated</p> <ul style="list-style-type: none"> <li>▪ Storage systems</li> <li>▪ App servers</li> <li>▪ Local clients</li> <li>▪ Connections to remote clients</li> </ul>
<b>Needed to recover</b>	<ul style="list-style-type: none"> <li>➤ Intact, accessible copy of critical production data</li> <li>➤ Connectivity to app servers</li> </ul>	<ul style="list-style-type: none"> <li>➤ Intact recovery site</li> <li>➤ Staff (including provisioning)</li> <li>➤ Hardware and connectivity for critical apps</li> <li>➤ Intact, accessible copy of critical production data</li> </ul>

# The case for data replication

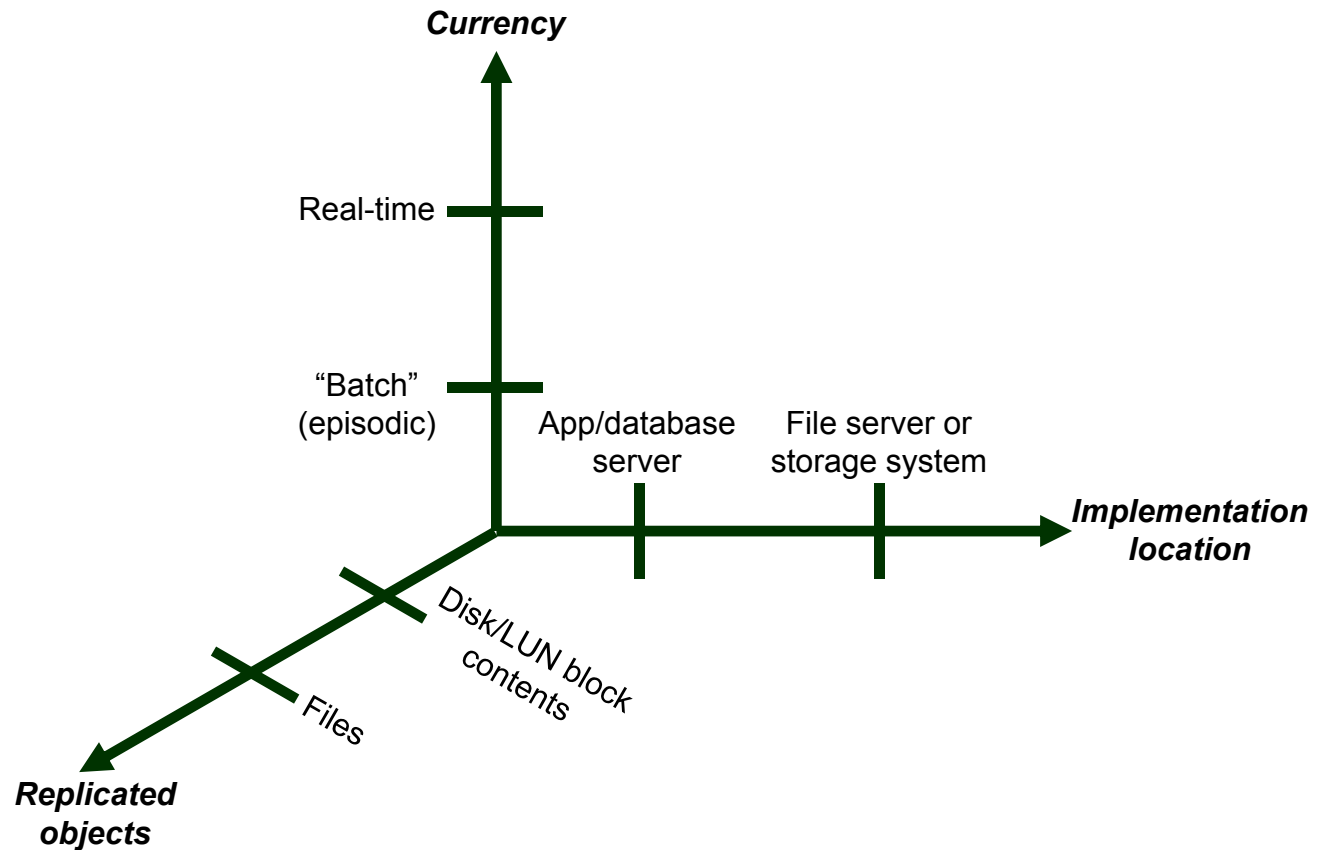
Backup-restore	Cost advantage	Replication
Tape drive, media, and library	←	Online storage capacity
Physical transportation	←	Replication link with adequate bandwidth and latency
Acquisition or activation of recovery site	←	Full-time recovery facility premises and staff
Cost of downtime to restore data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

- Have or acquire adequate equipment for access to critical data and applications
  
- Have or recreate a copy of critical data at the recovery location

- **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 “normal” availability

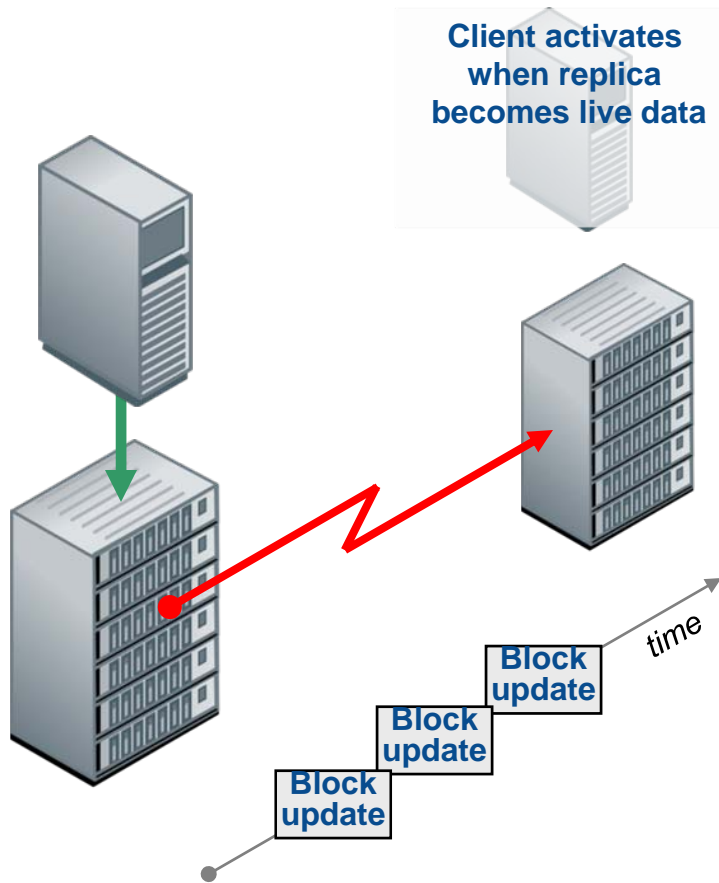
- Basic purpose: keep an up-to-date copy of a data set on separate storage resources  
*(usually attached to separate processing resources)*
  
- Different from mirroring
  - ◆ Primary-secondary relationship
  - ◆ Time ordered updates to a “consistency group” of devices
  - ◆ May be asynchronous
  
- Use cases
  - ◆ Recovery from physical disaster
  - ◆ Data distribution/consolidation
  - ◆ Second data source for read-only applications
  - ◆ Baseline for writable “clones” of data

# The three basic replication choices

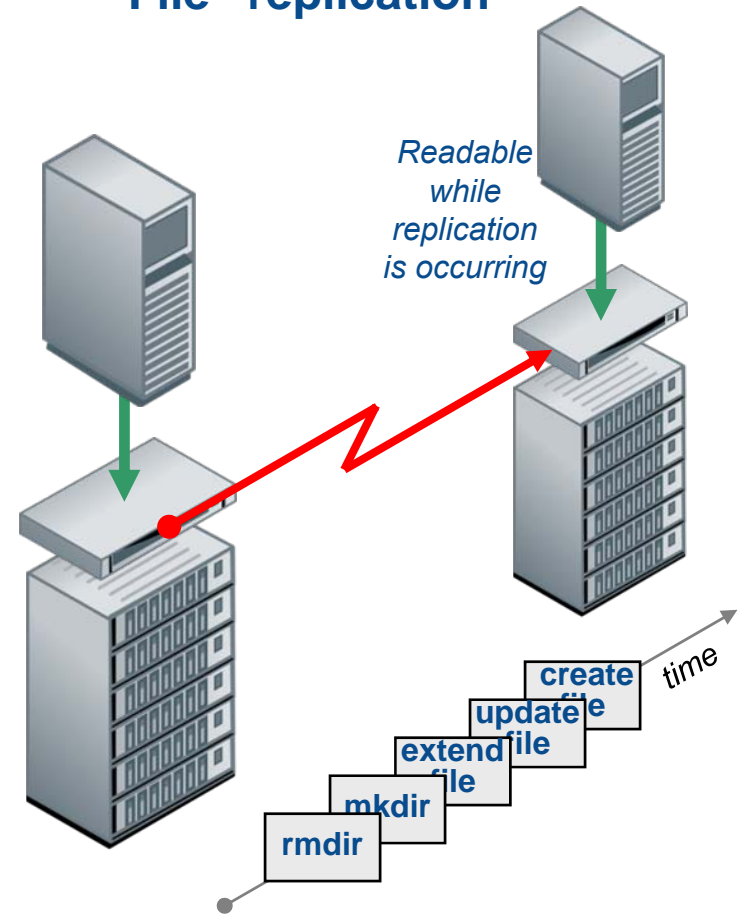


# “Block” vs. file-level replication

## “Block” replication



## “File” replication

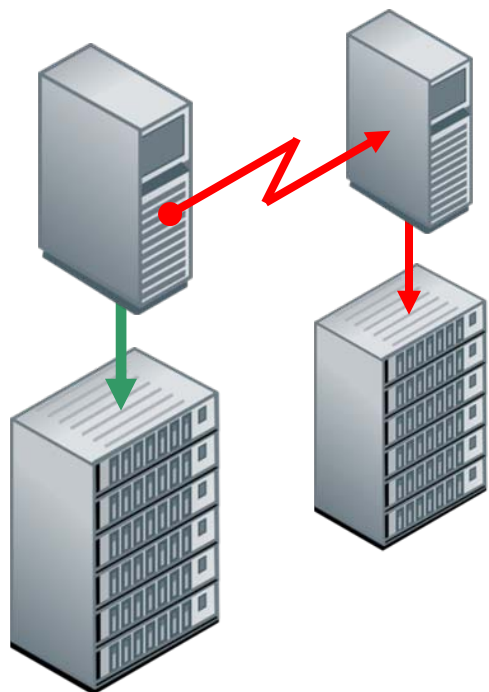


— Client read/write access  
— Replication traffic

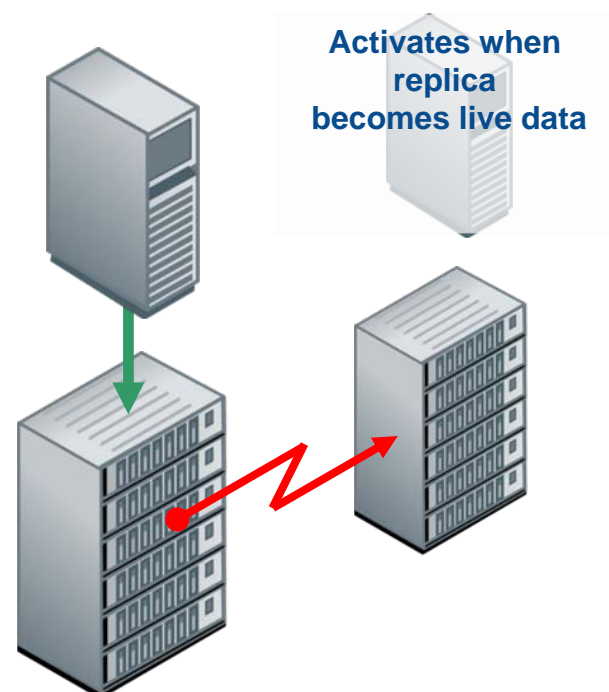
Compare properties: slide

# Host-based vs. storage system-based replication

## “Host” (application server)-based



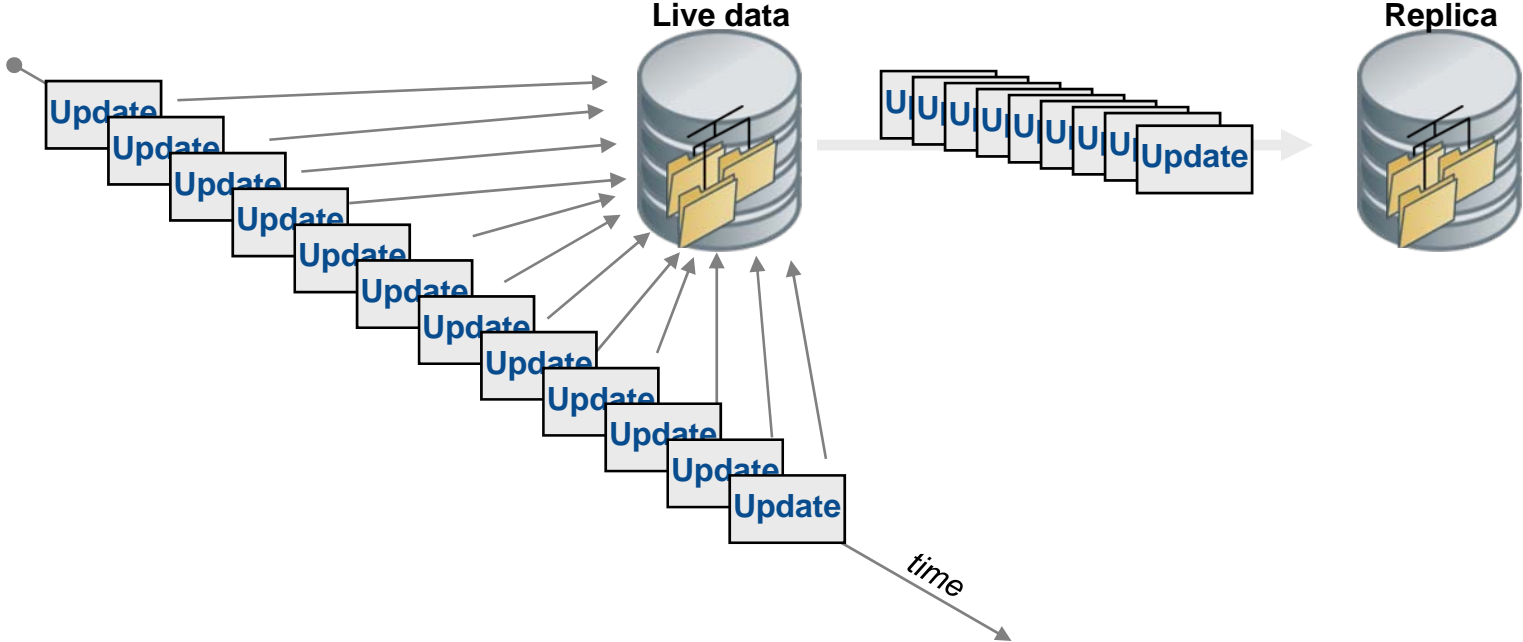
## Storage system-based



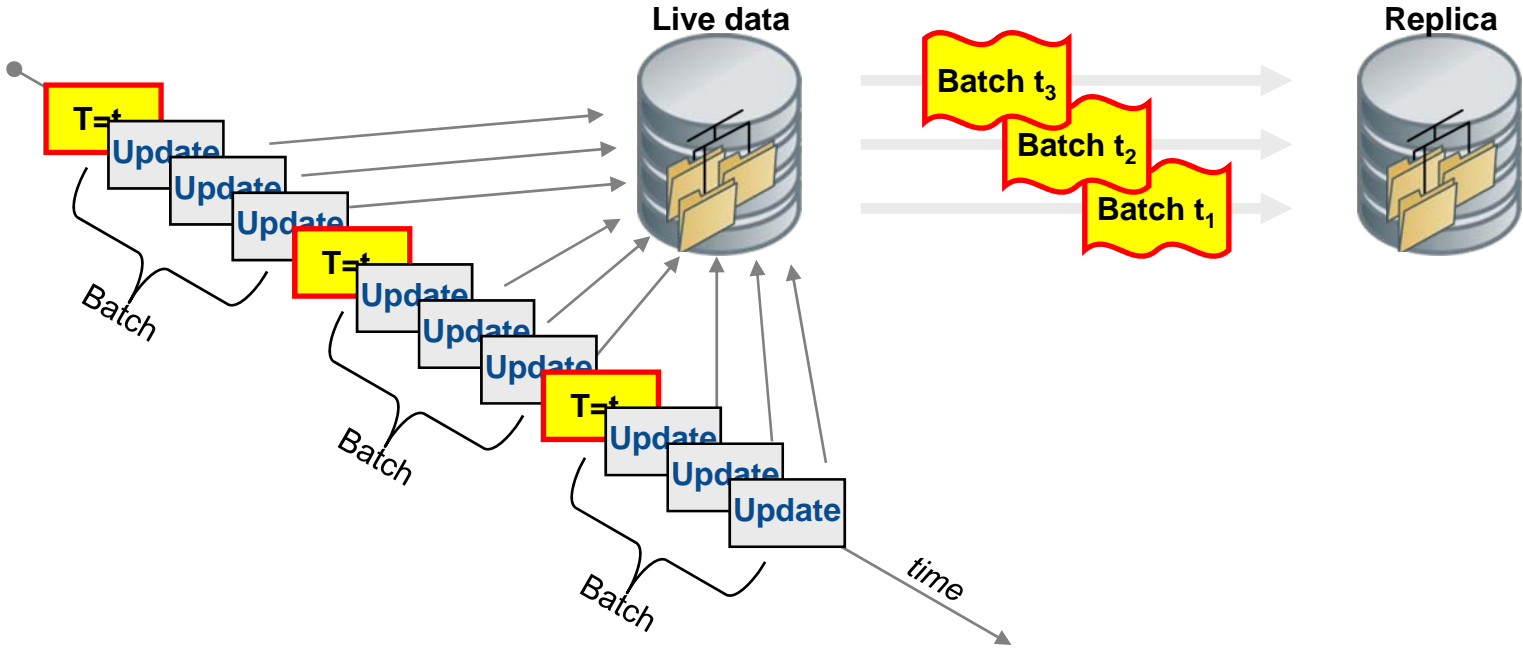
## ➤ Other variations:

- ◆ SAN switch, NAS aggregator, dedicated appliance
- ◆ Roughly equivalent to storage system-based replication

# Real-time vs. batch replication

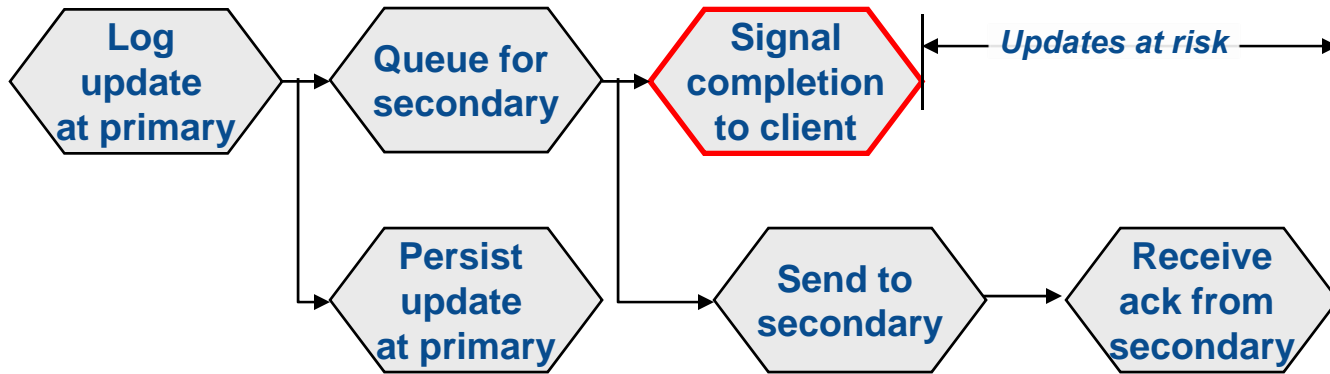
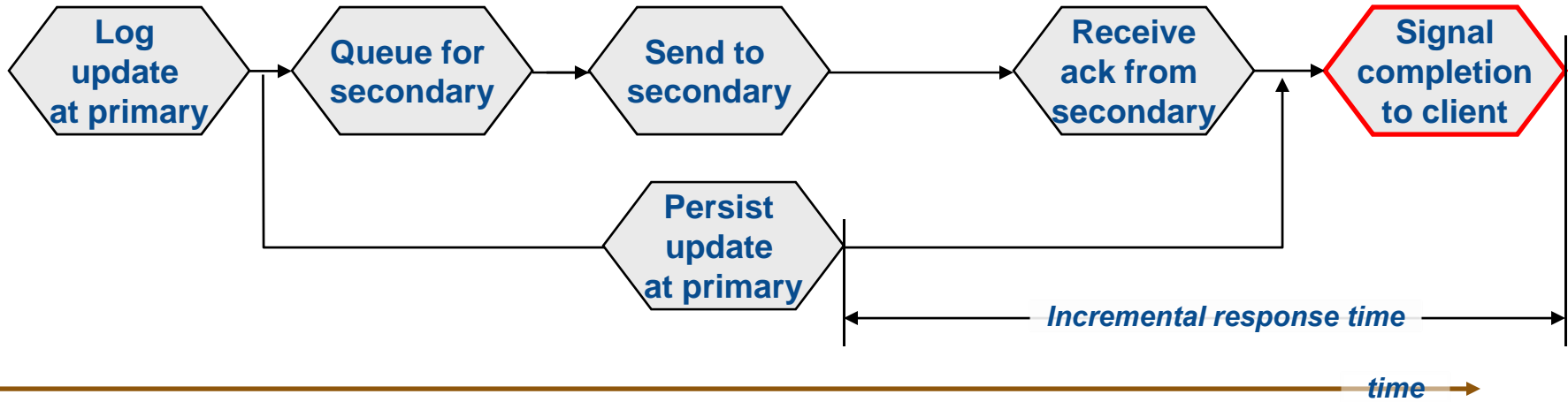


# Real-time vs. batch replication



Compare properties: slide

# Synchronous vs. asynchronous replication



› Variation on a theme:  
"limited asynchronous" replication

# Which kind of replication is best ?

- It depends on your definition of “best”
  
- Block-level vs. file-level
  - ◆ Universal availability vs. bandwidth and readability
  
- Synchronous vs. asynchronous
  - ◆ Potential for data loss vs. client responsiveness
  
- Real-time vs batch update
  - ◆ Bandwidth consumption vs. recovery point granularity

## ➤ Different threats→different recovery techniques

### **Logical disaster**

- ◆ Technique: “turn back the clock”
- ◆ Property: Inherently some data loss

### **Physical disaster**

- ◆ Technique: Recovery facility + replica of critical data
- ◆ Property: Cost vs. recovery point tradeoff

## ➤ Different data→different requirements

- ◆ Critical: RTO = seconds; RPO = now!
- ◆ Less-critical Cost becomes a factor

## ➤ NAS is unique because it “knows” about data structure

- ◆ Able to participate proactively in protection and recovery processes

- Please submit any questions or comments on this tutorial to the SNIA at [trackfilemgmt@snia.org](mailto:trackfilemgmt@snia.org)

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

*SNIA Education Committee*

**David Dale  
Rob Peglar  
Warren Avery  
Norman Owens**

**Netapp  
Xitech  
storagenetworking.org  
storagenetworking.org**

# **Background slides—for publication not to be presented**

# 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

# Techniques for backing up NAS data

## Backup management software

Advantages	Limitations
<p>“Set-and-forget” automation</p> <ul style="list-style-type: none"><li>Schedules</li><li>Device and media management</li><li>Data selection</li></ul>	<p>Cost</p> <ul style="list-style-type: none"><li>License &amp; maintenance</li><li>Training</li></ul>
<p>Added-value features</p> <ul style="list-style-type: none"><li>Incremental backup</li><li>Multi-streaming &amp; multiplexing</li></ul>	<p>Requires data to “stand still” during backup</p>
<p>Application (e.g., database) integration</p>	<p>Inherently coarse-grained recovery times &amp; recovery points</p>

# Snapshots: what to snap ... “blocks” vs files

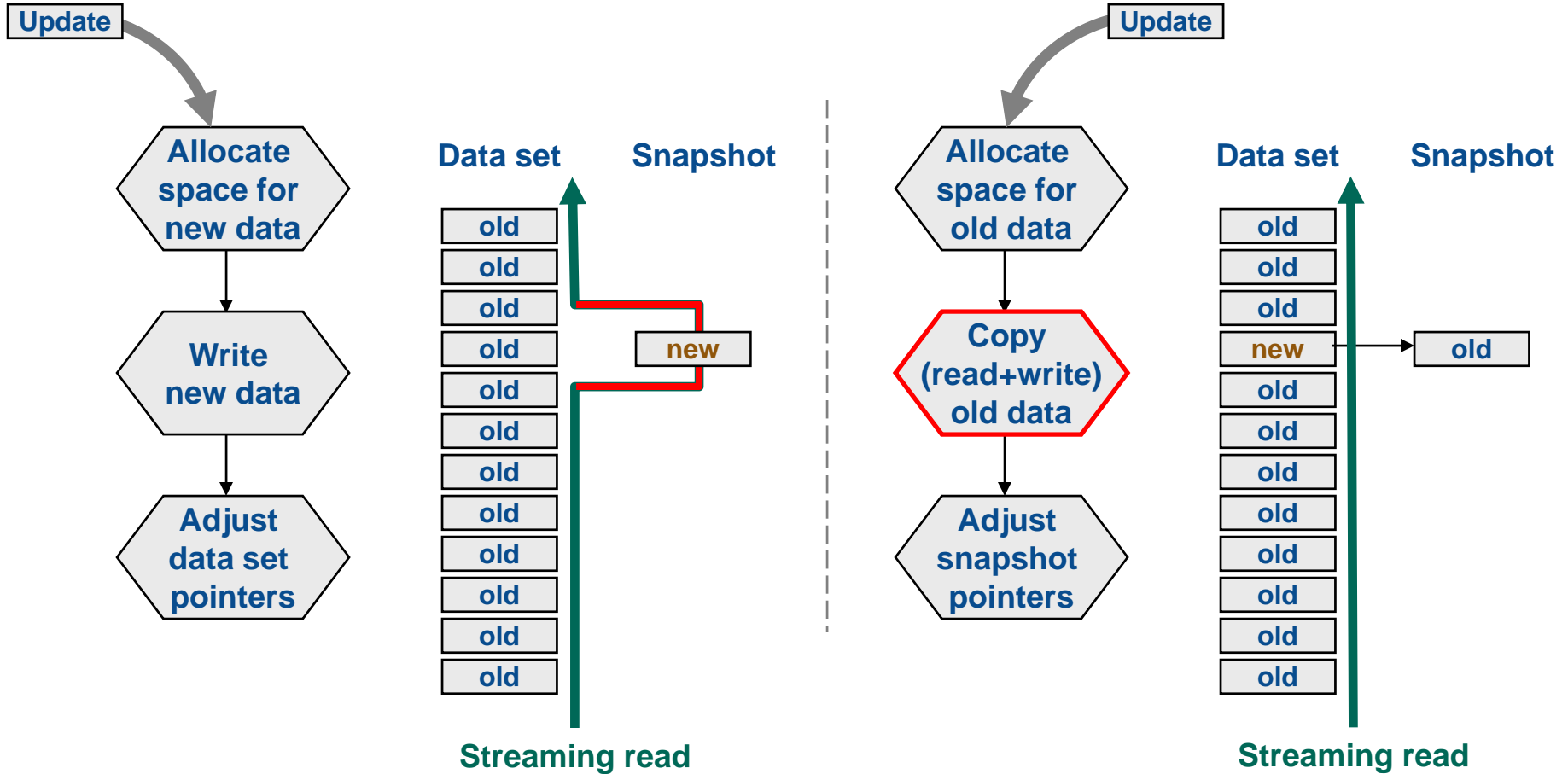
	Blocks	Files
Creation time	Fast	Some metadata creation required
Atomicity	Block update	File system operation
Impact on live data	Directly related to disk I/O operations	Some “surprises” possible (e.g., deletions)

*NAS developers have the luxury of choosing either*

# Full-size vs. space-saving snapshots

<b>“Split mirror”</b> (content = bit-for-bit copy)	<b>“Copy-on-write”</b> (content = changes only)
<b>Protects against device failures and media defects</b>	<b>Requires separate physical protection mechanism</b>
<b>Storage requirement: full data set size</b>	<b>Storage requirement: <math>\propto</math> change in data set size during snapshot life</b>
<b>Overhead to maintain: zero</b>	<b>Overhead to maintain: <math>\approx</math> 2-3x for every “first write”</b>
<b>Deletion cost: Resynchronization with data set</b>	<b>Deletion cost: Space reclamation</b>

# Copy-on-write vs. pointer-mapped snapshots



Performance impact —

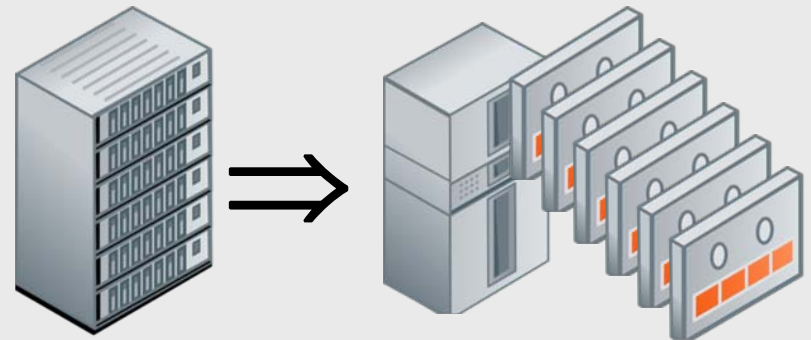
# Copy-on-write vs pointer-mapped snapshots

Pointer-mapped	Copy-on-write
<b>Faster initial update-in-place (allocate+write+pointer adjust)</b>	<b>More work for initial update- in-place (allocate+read+write+pointer adjust)</b>
<b>For small updates: Tends to fragment live data set</b>	<b>For small updates: no live data set fragmentation</b>
<b>For large updates: Both tend to preserve physical contiguity (for later streaming reads)</b>	

- Disk-based storage that emulates tape libraries
- Exploits existing backup infrastructure
  - ◆ Easy to implement
- Can use tiered storage to minimize cost per byte stored
  - ◆ e.g., “MAID”
  - ◆ e.g., compression & de-duplication
- Enhances NDMP capabilities
  - ◆ Mitigates impact of tape drive reservations
  - ◆ Mitigates impact of single-streaming

**Tape  
media**

- Behaves like
  - ◆ Tape drives
  - + Media
  - + Robotic loader



# Host-based vs. storage system-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

# 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)

Real-time (op-by-op)	Batch (periodic snapshots)
Transmits every operation (i.e., sends repeated ops repeatedly)	May consume less network bandwidth
Replicates every primary data set state (any primary data set recovery point can be recreated from the replica)	May not represent all primary data set states in the replica (recovery points at batch boundaries only)

# Synchronous vs 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  (committed updates may be lost in a disaster)
“Round trip” time is additive to client response time	Little if any increment in client response time

- Please use this icon to refer to other SNIA Tutorials where appropriate.



**Check out SNIA Tutorial:  
Enter Tutorial Title Here**