Redesign Operational Recovery for Critical Applications

Edwin Galang, Senior Architect, Data Protection, Hitachi Data Systems
Jonathan Howard, Product Manager, CommVault

July 23, 2014
Redesign Operational Recovery for Critical Applications

How are you protecting the business-critical data and applications on your Tier-1 storage systems? Traditional incremental and full backup paradigms don’t work when you can’t stop the system to back it up (backup window), can’t afford to lose any data (recovery point objective – RPO) and can’t afford downtime (recovery time objective – RTO). Join us to review new approaches to data protection and recovery that don’t require a rip-and-replace of your existing infrastructure.

By attending this webcast, you’ll learn how to:

- Eliminate the need for a backup window and reduce the stress on your application server
- Protect data more often, reducing the amount of data at risk
- Layer capabilities for operational recovery, disaster recovery, and continuous availability

WebTech Educational Series
The Challenge and the Opportunity

- DP spend (total cost) is equal to or more than primary capacity costs (33% of total IT budget)
  - For every dollar spent on primary capacity, another dollar is spend to protect/recover it (conservatively)
- Just looking at DP spend alone, ~30% is physical capacity (disk and tape targets)
  - Remainder is spent on infrastructure and effort
    - Data movement/copy, software, network, people
- **99+% of all restores are for logical operational recovery, not site failures**
  - DB or app cloning, file restore, logical corruption or human deletions, and so forth
- Most important data is already centrally located on enterprise platform
- Restore RPO/RTO is very high for critical applications, operationally difficult, and not always 100% reliable
- Sub-24-hour recovery point is nearly impossible for large systems such as databases

![Diagram showing the cost distribution between primary storage, data protection, people, infrastructure, and physical media/targets]

- **Primary Storage (Total: $100)**
- **Data Protection (Total: $100)**
- **Other (Total: $100)**
- **Physical Media/Targets (Total: $30)**
- **People, Infrastructure (Total: $70)**
Although the entire backup solution has a growing cost structure, a majority of the cost (70%) is in the copy/move process:

- Software, network, servers, people
Clone and Snapshot Use Cases by RTO

- Point-in-time snapshots for logical recovery (local and remote) <24 hours to improve backup and restore RTO/RPO and cost

- Eliminate the traditional backup window and have multiple point-in-times within a 24-hour window while not consuming space

- Refresh and run a non-production environment in a fraction of the space

- Not a solution for disaster recovery; works in conjunction with one at both sites
Hitachi ShadowImage Clones and Hitachi Thin Image Snapshots

**ShadowImage Replication software**
All data is saved from P-VOL to S-VOL

**Thin Image snapshot software**
Only changed data is saved from P-VOL to data pool; pool is shared by multiple snapshot images (V-VOL)

Consistent read and read/write access is available only in split states
What Is a ShadowImage Clone?

Clone technology that creates point-in-TIME copies of your data for LOGICAL PROTECTION

- Creates full clone copies of data for backup or application testing purposes
- Enables full point-in-time clones for disk-based backups
  - Speeds backups from hours to a few minutes
  - Virtually eliminates traditional backup windows
- Creates up to 9 copies of data per primary volume
  - Can be used for application testing
  - Multiple backups limit exposure to data loss
- Data restoration
  - Application- and OS-independent but can be integrated with application backup triggers
  - Can restore the S-VOL data to the P-VOL by copying the data (normal) or swapping the S-VOL and P-VOL (quick)
- Can be replicated with Hitachi TrueCopy or Hitachi Universal Replicator
ShadowImage: Quicksplit and QuickResync

**QuickResync**

- Nearly instant resync
- Differential data map copied immediately
- Actual data copied during update copy operations

**QuickSplit**

- Immediate read/write access to secondary volume
- Pending writes copied in the background
Data is copied from the S-VOL to the P-VOL.

- The restore process will copy over the P-VOL data. Once the restore process has started, it cannot be stopped.
- Unmount the P-VOL before starting the restore process.
- The amount of time required to copy the data is dependent on the amount of differential data, disk types, and copy pace.
  - Use the –c option to manage the copy pace.
- During the restore process, SI will be in RCPY status.
- When the restore process is complete, ShadowImage will be in PAIR mode.

pairresync –IM0 –g SI-group –restore –fq normal
The quick restore process will swap the S-VOL and P-VOLs
- Unmount the P-VOL before starting the restore process
- It is an instantaneous process
- After the quick restore process is complete, the production server will be using the HDP pool/parity groups that the S-VOL was using
  - Make sure there are sufficient resources available to meet the production performance requirements

pairresync –IM0 –g SI-group –restore –fq quick
What Is Hitachi Thin Image?

Snapshot technology that creates instant Point in TIME copies of your data for LOGICAL PROTECTION

- Creates instant copies of data for backup or application testing purposes
- Saves up to 90% or more disk space by storing only changed data blocks
- Enables fast and frequent point-in-time snapshots for disk-based backups
  - Speeds backups from hours to a few minutes
  - Virtually eliminates traditional backup windows
- Rapidly create up to 1,024 snapshots per volume, 32K per array
  - Great for application testing
  - Multiple backups limit exposure to data loss
- High-performance data restoration
  - Near-instant restore of critical data, without losing other point-in-times
  - Application- and OS-independent but can be integrated with application backup triggers

Fast, simple, and reliable snapshot software
Enhanced Performance, scalability, and capacity

- **70X improvement** in data capacity
  - Maximum source capacity increased to 2.1PB
    - Support up to 70 times more virtual machines
    - Protect up to 70 times larger data sets
    - Maximum pool size increases to 12.1PB

- **32X improvement** in snapshot copies
  - Maximum snapshot limit increased to 1,024
    - Snapshots can be maintained for longer periods
    - Snapshots can be taken more frequently

- **25X improvement** in performance
  - Asynchronous (copy after write) operation greatly improves response time to host
  - Enhanced metadata table structure for super-fast data recovery
(1) Host Write

Host → Data B

(2) Write Complete

Data B → Data A → P - VOL

- Data is copied to the snap pool after the write has been completed to cache of the array. Host does not need to wait for the array operations.
- Subsequent writes to the same block for the same snapshot don’t have to be moved.
- Single instance of data stored in HDP snap pool regardless of number of snaps.
- Restore is a quick restore. Primary volume is available for use immediately after restore operation is executed. Data is copied in background.
- Separate pools for data and snapshot ensure scalability, production performance, and capacity protection.

(3) Asynchronous Upstage to Cache (Read Miss)
How Thin Image Works: Pairing Snapshots

HTI Snapshots Are in Paired State
How Thin Image Works: Tracking Changes for 1 Snapshot

The Differential Data Between the P-VOL and V-VOL Are Written to the HTI Pool
How Thin Image Works:
Tracking Changes for 2 Snapshots

When the Bottom V-VOL Is Also Split, Both Snapshots Track Changes That Occur After the Split
How Thin Image Works: Tracking Changes for 2 Snapshots

When the Bottom V-VOL Is Also Split, Both Snapshots Track Changes That Occur After the Split

Update I/O

P-VOL

SPLIT

V-VOL

SPLIT

HTI Pool

V-VOL
How Thin Image Works: Restore Without Impacting Other Snaps

Host Can Access “Restored” Data Immediately After Operation Is Executed To Reduce RTO. Background Copy Is Performed
How Thin Image Works:
Restore Without Impacting Other Snaps

All Other Snapshots, Prior and Future to Restore Point, Are Retained and Maintained
HTI pool capacity sizing considerations

- Point-in-time changed data is stored in the HTI pool
- Writes to the V-VOL(s) are stored in the HTI pool
  - Total changed data for the life snap(s) + V-VOL writes = total HTI pool
Combining In-System Replication Configurations

Unparalleled performance, scalability and capacity
IntelliSnap Sequence

1. Simpana iDA requests application/database to be quiesced/unquiesced
2. Raid Manager software installed on host used to communicate to the command device for array communication
3. Snapshots (V-VOLs) are created from an existing HTI pool
4. IntelliSnap mounts the snapshot on the proxy and indexes the contents to enable granular recovery
It’s Easy Now! No, Really…

- From a CommVault perspective enable
  - Check “Create VVOLs for Thin Image”
  - Provide the TI pool
- From a Hitachi perspective add a command device and CCI software on the host (or with the Remote Snap MediaAgent)

Set up in minutes
IntelliSnap Sequence

1. Simpana iDA requests application / database to be quiesced / unquiesced
2. Raid Manager software installed on host used to communicate to the Command Device for array communication
3. Clones (S-VOLs) are created from an existing SI Pool which performs the initial copy of data; once completed the clone is split
4. IntelliSnap mounts the clone on the proxy and indexes the contents to enable granular recovery
5. Clones can be resynchronized to update the data from the primary volume as required
New Array Options for CCI Integration
Individual Client Configuration Options
Array Settings Can Be Overridden on a Per-Subclient Basis

**Note:** The modification in Snap configuration values are specific to this Subclient.
Remote Snap MediaAgent
Centralized Communication Option for IntelliSnap Hitachi Data Systems CCI Engine

No Remote Snap MediaAgent

- Each client will utilize CCI to communicate through the command device to the storage array
- In larger deployments this can require an increase in complexity and effort to deploy and manage the solution

With Remote Snap MediaAgent

- The Remote Snap MediaAgent centralizes the communication from multiple clients, requiring that the command device and CCI software only be installed on that host
- Each Remote Snap MediaAgent can manage up to 25 snap jobs per hour, and should be located on the same host as the proxy for secondary mounting operations
Cascaded Snapshots
Extending the Use of Clones and Snapshots

- Snapshots are well suited for their quick recovery; however, utilizing them for additional workloads will impact production resources.
- Clones are well suited for additional workloads; however, there is a heavy infrastructure cost to keep multiple copies around.
- Cascaded snapshots enable smarter tiering and retention inside of the array by:
  - Eliminating the load on the production volume for additional workloads.
  - Reducing the need and use cases for more than 1 clone to be required for multiple workload and retention points.
  - Eliminating lengthy and manual refresh cycles.

Use Case
Cascaded Snapshot Anatomy (Host)
Create Thin Image Snapshots from ShadowImage Clones

IntelliSnap Sequence

1. IntelliSnap Cascaded Snapshot requires that a ShadowImage relationship is pre-existing, or can be set up via IntelliSnap (separately)
2. During the IntelliSnap snapshot operation, the ShadowImage (S-VOL) clone will be re-synchronized with the LUN (P-VOL)
3. The ShadowImage (S-VOL) clone will be suspended/split from the LUN (P-VOL)
4. A Thin Image snapshot (V-VOLs) will be taken from the ShadowImage (S-VOL) clone and not the LUN (P-VOL)
5. IntelliSnap will perform all secondary operations (mount, index, stream) from the Thin Image snapshot
6. Snapshots can be mounted to alternate hosts for additional test/support operations
Simplified and Enhanced Troubleshooting

Use Case

- Enables additional logging from the Hitachi CCI engine
  - This does not increase the debug level on the Simpana software
- Information is logged under the C:\Program Files\CommVault\Simpana\Base\Temp directory
  - Logged in the formation of “script” and “resp”
  - Script includes the commands that IntelliSnap has generated to send to the CCI instance for an operation
  - Response is the array response
IntelliSnap – Anatomy (Host)
One Policy Orchestrating Distributed Components

Recovery copy

On storage array protection
- Native format
- Auto-discovery of storage
- Completely integrated (no scripting)
- Snapshot and/or clone support

Application integrated protection
- Log management
- Application consistent

On storage array recovery
- Revert volume
- Clone/mount – copy back data

Application recovery
- DB recovery RPO
- Log restore
- DB mount > object recovery tool

Indexed/management

Indexed live copies
- Searchable data from snapshots/clones
- Browsable indexes

Mining operations
- Granular mining for objects
- Database mounts

Retention copy (optional)

Indexed retention copies
- Protect entire dataset or only relevant data
- Full tiering options available (disk/cloud/tape)
- Align retention and recovery policies

Off-array retention
- Deduplication, compression, and encryption options
- Storage and location independent

Testing/development enhancements
- Automated creation of dev/test datasets

Retention copy

Retention Copy
Proxy (MediaAgent)
Mount/Index

Recovery Copy

Storage Array

IDM

Production LUN/Volum

Production Host

2 Hour Snapshots

Disk
Cloud
Tape

Testing/developers enhancements

- Automated creation of dev/test datasets

Full recovery options
- Agent dependent
- Cross server; full (BMR) down to granular options

Content indexing aware
- Agent/granular options
IntelliSnap – Anatomy (VMware)
VM Integrated Snapshots Tackle Scale, DR, and Access Needs

### IntelliSnap Sequence

1. **vCenter integration for auto-discovery of new virtual machines for protection**
2. **Virtual Server Agent (VSA) contacts vCenter and creates a consistent VMware snapshot for all VMs being protected**
3. **IntelliSnap communicates with the storage array to take snapshot(s)**
4. **Mount snapshot to ESXi Proxy for indexing and LiveBrowse operations; VSA performs indexing operations**
5. **Optionally data can be streamed off-array to create a long-term retention copy**
Improve Performance and Reduce Complexity
Combining the Storage and Database Layer

**Benefits**

- **Automated database integration for snap/clone** – use of integrated tools automates quiesce (hot backup mode, VSS, BRTools, and so forth)
- **No manual script generation**
- **Multiple PIT array copies** – allow for multiple recovery points on array
- **Storage array leverage** – large databases quickly protected
- **Low to zero production I/O impact**
  - Clones = no I/O impact
  - Snaps = minor I/O impact
- **Fast accessibility** – mount and browse or automate snap recovery through the standard “browse and recover” process
- **Multi-streaming for data movement**
- **Mixed mode protection** – logs may be protected more frequently than the entire database for log manipulation to exact PIT database restores with array snaps
- **Seamless recovery from multiple tiers** – automatic sequencing of log files and database files for direct PIT recoveries across storage tiers
- **Fastest recovery mechanism leveraged by default** – SAN copies are utilized for LAN free recoveries 1st

---

**2-Hour RPO**

**Database LUN/Volume**

**Two-Hour Snap**

**Managed Log Backups**

**30-Min. Interval**

**Mount**

**Retention Copy**

**Disk**

**Cloud**

**Tape**

**Contents ... /Vol1/DB/ /Vol2/Logs/**

---

**Disaster Recovery and Business Continuity**

**2-Hour RPO**

**Disk**

**Cloud**

**Tape**

---

© Hitachi Data Systems Corporation 2014. All rights reserved.
Database Awareness
Integrated Intelligent Recovery Process

Corruption Occurs at 6:42 P.M.

Storage Array

Recovery Copy

Database

4 P.M.
6 P.M.
8 P.M.

Managed Log Backups

30-Min. Interval

Retention Copy

Disk
Cloud
Tape

Log Contents ... (On Disk)

- /Logfiles-xxx.log - 4:30pm 2014/05/24
- /Logfiles-xxx.log - 5:00pm 2014/05/24
- /Logfiles-xxx.log - 5:30pm 2014/05/24
- /Logfiles-xxx.log - 6:00pm 2014/05/24
- /Logfiles-xxx.log - 6:30pm 2014/05/24
- /Logfiles-xxx.log - 7:00pm 2014/05/24
- /Logfiles-xxx.log - 8:00pm 2014/05/24
- /Logfiles-xxx.log - 8:30pm 2014/05/24

Retention Copy

Storage Array

Recovery Copy

Database

4 P.M.
6 P.M.
8 P.M.

Managed Log Backups

30-Min. Interval

Retention Copy

Disk
Cloud
Tape

Log Contents ... (On Disk)

- /Logfiles-xxx.log - 4:30pm 2014/05/24
- /Logfiles-xxx.log - 5:00pm 2014/05/24
- /Logfiles-xxx.log - 5:30pm 2014/05/24
- /Logfiles-xxx.log - 6:00pm 2014/05/24
- /Logfiles-xxx.log - 6:30pm 2014/05/24
- /Logfiles-xxx.log - 7:00pm 2014/05/24
- /Logfiles-xxx.log - 8:00pm 2014/05/24
- /Logfiles-xxx.log - 8:30pm 2014/05/24
Effectively use your dollar spend for operational recovery on both technology (hardware and software) and architectures that meet recovery RPO and RTO needs for the business.

Allow the best-of-breed solution to perform its designed functions to the maximum capability. Servers for processing transactions, storage for data copy and movement, and software for orchestration.

Implement an effective operational recovery solution that is not only application aware, but can actually provide flexibility for meeting restores under varying unforeseen scenarios.

Do not trade off production performance, scalability, and reliability for operational protection – have your cake and eat it too.

Orchestrate not only backups but also refreshes of non-production environments for a fraction of the capacity cost and time.

Integrate effective operational recovery to the complete lifecycle of data protection for your business.
Questions and Discussion
### Upcoming WebTechs

- **WebTechs**, 9 a.m. PT, 12 p.m. ET
  - *Secure, Efficient Workforce Mobility*, August 19
  - *Visibility and Control with Secure Private Cloud*, September 10
  - *Extend IT Beyond the Data Center*, September 24

- **Check** [www.hds.com/webtech](http://www.hds.com/webtech) for
  - Links to the recording, the presentation, and Q&A (available next week)
  - Schedule and registration for upcoming WebTech sessions
  - Questions will be posted in the HDS Community: [http://community.hds.com/groups/webtech](http://community.hds.com/groups/webtech)
Thank You