Hitachi Data Systems

Hitachi Data Migrator to Cloud Best Practices Guide

MK-92HNAS045-05

Hitachi, Ltd., reserves the right to make changes to this document at any time without notice and assumes no responsibility for its use. This document contains the most current information available at the time of publication. When new or revised information becomes available, this entire document will be updated and distributed to all registered users.

Some of the features described in this document might not be currently available. Refer to the most recent product announcement for information about feature and product availability, or contact Hitachi Data Systems Corporation at https://support.hds.com/en_us/contact-us.html.

Notice: Hitachi, Ltd., products and services can be ordered only under the terms and conditions of the applicable Hitachi Data Systems Corporation agreements. The use of Hitachi, Ltd., products is governed by the terms of your agreements with Hitachi Data Systems Corporation.

Hitachi Data Systems products and services can be ordered only under the terms and conditions of Hitachi Data Systems' applicable agreements. The use of Hitachi Data Systems products is governed by the terms of your agreements with Hitachi Data Systems.

Hitachi is a registered trademark of Hitachi, Ltd., in the United States and other countries. Hitachi Data Systems is a registered trademark and service mark of Hitachi, Ltd., in the United States and other countries.

Archivas, Dynamic Provisioning, Essential NAS Platform, HiCommand, Hi-Track, ShadowImage, Tagmaserve, Tagmasoft, Tagmasolve, Tagmastore, TrueCopy, Universal Star Network, and Universal Storage Platform are registered trademarks of Hitachi Data Systems Corporation.

VMware is a registered trademark or trademark of VMware, Inc. in the United States and/or other jurisdictions.

All other trademarks, service marks, and company names in this document or website are properties of their respective owners.

The product described in this guide may be protected by one or more U.S. patents, foreign patents, or pending applications.

Notice of Export Controls

Export of technical data contained in this document may require an export license from the United States government and/or the government of Japan. Contact the Hitachi Data Systems Legal Department for any export compliance questions.

Document Revision Level

Revision	*Date*	*Description*
MK-92HNAS045-00	May 2013	Initial Release
MK-92HNAS045-01	October 2014	Revision 1, replaces and supersedes MK-92HNAS045-00
MK-92HNAS045-02	May 2015	Revision 2, replaces and supersedes MK-92HNAS045-01
MK-92HNAS045-03	June 2015	Revision 3, replaces and supersedes MK-92HNAS045-02
MK-92HNAS045-04	June 2016	Revision 4, replaces and supersedes MK-92HNAS045-03
MK-92HNAS045-05	September 2016	Revision 5, replaces and supersedes MK-92HNAS045-04

Contact

Hitachi Data Systems
2845 Lafayette Street
Santa Clara, California 95050-2627
https://portal.hds.com

North America: 1-800-446-0744

Table of Contents

Intended audience 1

Overview 1

Data Migrator to Cloud 1

Data Migrator to Cloud architecture 2

Apcellerator and the CloudApp 2

Networking 3

Transitioning from the private network 4

Migrated data identifier on the cloud target. 5

Data Migrator to Cloud performance 5

Using Data Migrator to Cloud 5

Interoperability with other features 6

Classic Data Migrator 6

File Replication 7

Scenario 1 7

Scenario 2 7

Scenario 3 7

Object replication 8

Primary deduplication 8

NDMP backup 8

Character sets 9

Virtual Server Security 9

Multi-tenancy 9

Hitachi Content Platform view of XVLs 9

Intended audience

The intended audience for this guide is Hitachi Data Systems (HDS) customers, employees, and partners.

Overview

Data Migrator to Cloud, introduced in HNAS release 11.1, allows files hosted on the HNAS server to be transparently migrated to cloud storage, providing the benefits associated with both local and cloud storage. Data Migrator to Cloud combines External Cross-Volume Link (XVL) technology in HNAS with cloud storage targets such as the Hitachi Content Platform, Hitachi Cloud Services, Amazon S3, and Microsoft Azure.

A public cloud (such as Hitachi Cloud Services – Content Archive, Amazon S3, and Microsoft Azure) is provided by external entities hosting storage at their facility and paid for on a per-use basis. A private cloud is purchased and controlled by the end user. Data Migrator to Cloud supports both and the user can decide which model best suits their business needs. In all cases, it can transparently access the cloud storage directly to view and download data. In all scenarios, data is protected both in-flight and at-rest regardless of where the physical storage is hosted.

Data Migrator to Cloud

Data Migrator to Cloud was designed to leverage private and public clouds It also provides integration with the Hitachi Content Platform (HCP). Data Migrator to Cloud provides the following functionality:

- HCP Namespace verification: HNAS will validate best practice settings on HCP to ensure optimal configuration. HNAS will check that the default retention class is disabled and that the Data Access user permissions are sufficient. The required permissions are for all three levels—tenant, namespace, and namespace user—and should include Read, Write, Delete, Purge, and Search.
- Migration Performance to HCP: The migration engine is multi-threaded which aids in file system walking performance and migration job duration. In HNAS release 11.1, all threads connect to single HCP node. In HNAS release 11.2, HNAS connects to multiple HCP nodes which improves upload performance.
- Writing/modifying migrated data: In order for data migration to HCP and cloud targets to be transparent to users, the users need the ability to modify migrated files. With Data Migrator to Cloud, users can seamlessly modify files that have been migrated.
- Deletion of stub files: When the user deletes a file that has migrated to HCP or one of the other supported cloud targets, the delete is propagated to the target. With 11.2 and beyond, HNAS snapshots protect stubbed data. A file that has been stubbed will not be removed from the target until all XVL references, including snapshots references to the archive file, have been removed.
- HCP "no delete": As of HNAS Version 12.1 and higher, this feature adds an extra level of data protection by delaying deletion of archived data on HCP even if the primary XVL is deleted.

  If an XVL is accidently deleted that represents a file that has been uploaded to HCP by Data Migrator to Cloud and no instance of the XVL exists in an HNAS file system snapshot, HCP retains the object in its namespace for a user-defined period of time (based on the setting of the retention class named "HNAS") rather than issuing the purge. If the accidental deletion is recognized in time, the file can be retrieved manually from HCP back to the HNAS file system. The decision whether to purge or retain a file upon deletion of the XVL depends on the presence of a retention class in the HCP namespace. After the retention period expires, the HCP disposition service will automatically clean up these files from the namespace. Note however, that HNAS does not support the use of the default retention class. A default retention class should not be configured.

- SSL (HTTPS) Support: Support for encrypted communications to public and private clouds. As of HNAS release 11.3, HTTP is the default protocol used for communication with local HCP targets.
- Full CLI support: Data Migrator to Cloud can be fully managed via the HNAS CLI.
- Fully implemented in the server: Data Migrator to Cloud is implemented on the HNAS server itself. It has no dependencies on the System Management Unit (SMU) and NDMP.
- Data Migrator to Cloud can be fully configured via the SMU GUI
- Data Migrator to Cloud allows multiple file systems to share the same HCP namespace and directory by assigning each file system a unique identifier (UUID)
- Data Migrator to Cloud preserves the migrated files path and name for easier browsing on the cloud target.

There are also some differences when compared to Classic Data Migrator.

- Reverse migration of any cloud-migrated files can only be initiated through the CLI. DM2C does not support policy based recall.
- In HNAS releases prior to release 12.3, the HNAS node eth0 and eth1 1GbE network interfaces are used for communication to the cloud target. Starting in HNAS release 12.3, Data Migrator now supports the use of network aggregates using the front-end networking ports on 1GbE and 10Gbe network interfaces. Refer to the *Data Migrator Administration Guide* for more details.
- Since HNAS 12.5, VLAN support for port aggregates has been added.

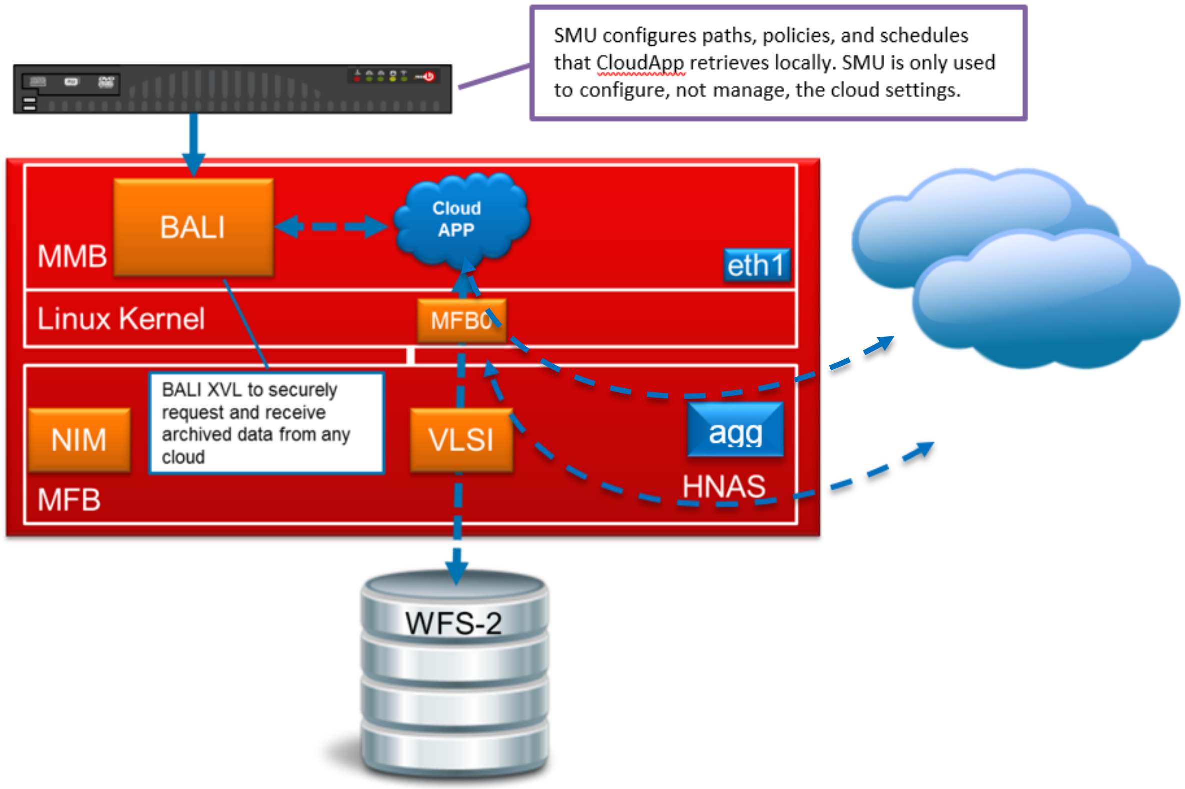
Data Migrator to Cloud architecture

Apcellerator and the CloudApp

Data Migrator to Cloud uses the HNAS Appcellerator infrastructure. This infrastructure allows storage management applications to run on the Linux Platform within HNAS. Appcellerator provides storage management applications on the HNAS Linux Platform access to HNAS WFS-2 file systems. Data Migrator to Cloud is an application that runs on the HNAS Linux Platform known as the CloudApp. The CloudApp consists of the tree walker, uploader, and a web server.

The tree walker scans the file system. The uploader sends data to the cloud target, and web server performs caching and read ahead.

A simplified diagram of this architecture is shown in in the following illustration.



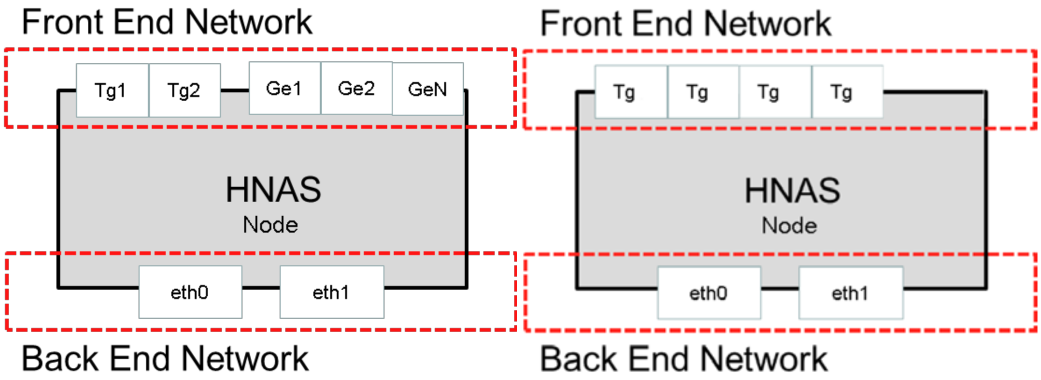
SMU configures paths, policies, and schedules that CloudApp retrieves locally. SMU is only used to configure, not manage, the cloud settings.

BALI XVL to securely request and receive archived data from any cloud

Networking

Data Migrator to Cloud uses the HNAS Linux Platform to encrypt (optional for local HCP targets starting in HNAS 11.3) to send and retrieve cloud data. In HNAS 12.3, Data Migrator supports migrating data through the 1GbE and 10GbE aggregate networks on the FPGA board. The use of aggregate networks is now the default for new data migration paths to the cloud. For non-encrypted targets, data is transferred directly from the SAN to VLSI to networks (not read into Linux).

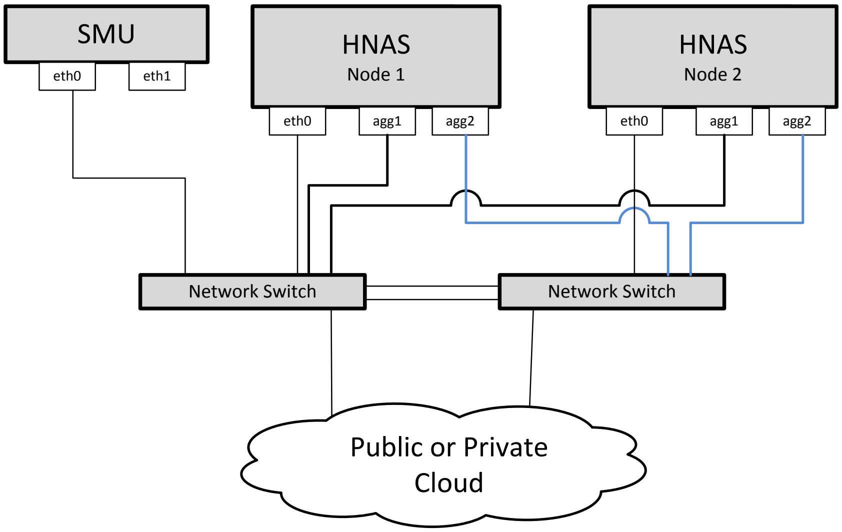
For encrypted targets, data is read to Linux, encrypted, sent back to VLSI then to aggregate networks. The use of front-end network aggregates provides Data Migrator to Cloud the ability to use faster networks, link aggregation, and zero copy to Linux for HCP. For HCP targets zero copy is now enabled by default. For public cloud targets, encrypt in Linux and send via aggregates is enabled by default. Data Migrator to Cloud chooses the aggregate based on BALI routing table that maps to the destination. Aggregates provide more resilient networking with seamless failover. Starting in HNAS 12.4, DNS resolution is no longer dependent on the HNAS management network ports. DNS resolution is now done over the aggregate networks.



To connect to a cloud target, HNAS must have DNS and network routes configured accordingly.

To successfully resolve the FQDN of the cloud target, the DNS query must be able to execute successfully as specified within the Global Security and EVS Context. All nodes in the cluster must have the ability to route to the cloud target, particularly those hosting the service and admin EVS.

For detailed instructions on configuring these interfaces, DNS, and routing, see the *Data Migrator Administration Guide*.



Note: It is recommended that both the service EVS and admin EVS have a valid route to the cloud target through the aggregate interface.

Transitioning from the private network

For information on updating existing Data Migrator to Cloud services to use aggregates, see the *Data Migrator Administration Guide*.

Migrated data identifier on the cloud target.

Data Migrator to Cloud uses a special universally unique identifier (UUID) for each file system that is migrated to a cloud target. This allows the user to migrate data from multiple source file systems to the same directory on a cloud target without have to worry about name collisions.

Though not required, you can create individual folders (Cloud Destinations) named after the source file system for easier identification of data on the cloud if several file systems are migrating data to the same target.

For example, when HCP is the cloud target, you can connect to the HCP namespace and browse the migrated data.

For chargeback, consider creating one namespace for each department or group and set the file system cloud migration to that namespace. This makes chargeback reporting easier.

Data Migrator to Cloud performance

Data Migrator to Cloud has been enhanced over Classic Data Migrator. Specifically, with the implementation of the tree walker, and the multi-threaded uploader mentioned previously in this document, upload performance has been improved. In HNAS release 12.5, further enhancements to the performance were done like multi-threaded recall of an individual file, multi-threaded reverse-migration, multiple reverse-migration processes are allowed simultaneously and other performance tunings.

Data set	*Single session upload*	*Concurrent sessions upload*	*Single Session reverse-migration*	*Single Client with 20 threads- recall*
2GBx50	~165 MB/s	~320 MB/s --> 2 sessions	~110 MB/s	~140 MB/s
10MBx5120	~140 MB/s	~450 MB/s --> 6 sessions	~100 MB/s	~135 MB/s

Cloud Data migrator performance has been improved in 12.5 when compared to 12.3 results. The data shown here was from a test was run on:

- HNAS (2 node cluster) model: HNAS 4100 software: 12.5.4038.03
- With HCP as a target: HCP500XL (10G)-4 node cluster OS 7.2.0.87

Using Data Migrator to Cloud

Data Migrator to Cloud supports multiple cloud providers. The table below lists each cloud provider and the required information you will need when adding a cloud account and destination. Please refer to the *Data Migrator Administration Guide* v.12.4 for detailed instructions on how to configure Data Migrator to Cloud for each provider.

Provider	*Server Name*	*User credentials*	*Server credentials*	*References*
HCP	Fully qualified domain name of the HCP namespace for the account credentials	User name of the Data Access account	The password of the Data Access account with read/write permission to the user account	

Provider	Server Name	User credentials	Server credentials	References
Hitachi Cloud Services	Fully qualified domain name of the Hitachi Cloud Services namespace for the account credentials	User name of the Data Access account	The password of the Data Access account with read/write permission to the user account	
Amazon S3	Auto-populates with aws-amazon.com	An IAM account	Security Credential key	https://console.aws.amazon.com/iam
Microsoft Azure	Auto-populates with azure.microsoft.com	Name of storage account	Primary or Secondary Access Key	https://azure.microsoft.com

Interoperability with other features

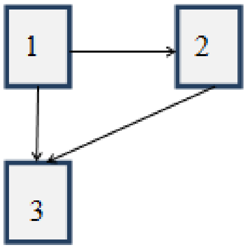
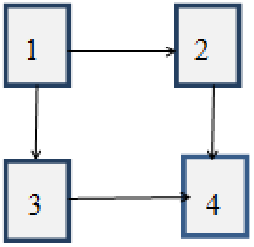
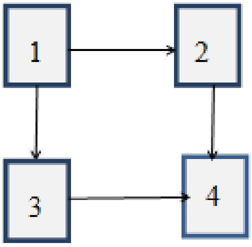
Classic Data Migrator

The environment variable `xvl-auto-recall-on-read` can be used with both Classic Data Migrator and Data Migrator to Cloud although setting `xvl-auto-recall-on-read` for Data Migrator to Cloud is not necessary.

Note: If `xvl-auto-recall-on-read` is set with Object Replication, the replication will trigger a reverse migration of the cloud migrated files.

File Replication

Hitachi NAS File Replication is the recommended method to replicate file systems that contain data migrated to the cloud.

Scenario 1	Scenario 2	Scenario 3
 1 = Primary HNAS 2 = Secondary HNAS 3 = HCP **or Amazon S3**	 1 = Primary HNAS 2 = Secondary HNAS 3 = Primary HCP 4 = replica HCP, will assume name of primary HCP	 1 = Primary HNAS 2 = Secondary HNAS 3 = Primary HCP 4 = replica HCP, will keep its unique name at all times

Scenario 1

Illustrates replicating file systems between HNAS clusters, both of which point to a single HCP or Amazon S3 target.

Warning! In this scenario, both HNAS clusters/entities map to the same HCP or Amazon S3 target. With HNAS file replication it is possible to access the secondary file system(s) at any time. It is strongly recommended to keep the destination file system syslocked to avoid unintentional deletion of data on the HCP system.

Scenario 2

Illustrates replicating file systems between HNAS clusters, where each cluster points to a HCP. The HCP replicates migrated data and also performs a DNS failover so that the secondary HCP maintains the same name resolution as the primary system.

Warning! In this scenario, HCP uses a DNS failover capability. Due to the way the HCP failover functionality operates, the secondary HNAS will also point to the primary HCP. With HNAS file replication it is possible to access the secondary file system(s) at any time. It is strongly recommended to keep the destination file system syslocked to avoid unintentional deletion of data on the HCP system.

Scenario 3

Illustrates replicating file systems between HNAS clusters, where each cluster points to a HCP. The HCP's replicate migrated data and maintain their own unique name resolution.

HCP replication is neither real time nor is it point-in-time. For this reason, you will be unable to guarantee a recovery point objective for any data migrated to HCP.

Important: DM2C only support File Replication when the configured Replication Policy/Rule is set to "recreate_link." DM2C does not support "remigrate." Migration-recreate-links-mode should always be set to "always-recreate-links."

For instructions on configuring HNAS in any of the above scenarios, see the *Data Migrator Administration Guide*, *HNAS Release Notes*, and *Replication and Disaster Recovery Guide*.

Object replication

Hitachi NAS object replication, when used to replicate a file system that contains data migrated to cloud, will follow the links and replicate all data to the destination file system. You cannot change this behavior. Use the options previously recommended for Hitachi NAS File Replication to properly handle the replication of XVLs.

Primary deduplication

Any file that has been fully or partially deduplicated on a Data Migrator to Cloud source file system will be rehydrated upon migration to the cloud target. If the file is recalled, HNAS will attempt to dedupe the data during the next deduplication job for that file system as it normally would for any newly written data.

Important: Certain HNAS 3080/3090/4040 deployment configurations may not have sufficient free MMB memory space to support both Data Migrator to Cloud and primary deduplication. Currently, see the HNAS product Release Notes for details on the memory requirements for Data Migrator to Cloud. In a future product release, this memory information will be located instead in the main product documentation.

NDMP backup

Hitachi NAS NDMP offers several variables to control how migrated (tiered) data is handled during backup and restore. These variables can typically be controlled through the backup application, and the way in which they are called is specific to each backup platform. For example, In NetBackup, environment variables can be set within the backup selections list by specifying one or more SET directives in a stanza. One should consult the documentation of the Backup application for specific guidance.

There are two main NDMP variables that control behavior of migrated files:

- NDMP_BLUEARC_EXCLUDE_MIGRATED: Controls how an NDMP backup interacts with CVL (files that have been migrated internally, for example, from SAS to NL-SAS). The Valid values are `y` or `n`. If set to `y`, the backup or copy will not include files whose data has been migrated to another volume. The default setting is `n` meaning that migrated files and their data will be backed up as normal files. The backup/copy retains the information that these files had originally been migrated.
- NDMP_BLUEARC_EXTERNAL_LINKS: Controls how an NDMP Backup interacts with XVLs (files that have been migrated to an external storage tier / cloud provider). The valid value are `remigrate`, `ignore` and `recreate_link`.
    - If set to `remigrate`, externally migrated files and their data will be backed up as normal files. On recovery the file will be restored and then an attempt will be made to remigrate the file to external storage again.

- If set to `ignore`, the backup or copy will not include files whose data has been migrated externally.
- If set to `recreate_link`, the backup or copy will include details of the link but none of the data contents. On recovery an attempt will be made to recreate the link to an existing file on the external storage system.

For platforms such as TSM that cannot directly manipulate NDMP variables, the CLI `ndmp-option command backup_ignore_external_links` option exists to allow the backup platform to ignore files migrate to external storage tiers.

For further details please consult the NDMP Backup Administrator Guide.

Note: if the `xvl-auto-recall-on-read` environment variable is enabled, an NDMP job will not cause the migrated files to be recalled.

Character sets

For versions **12.1 and below**, HNAS uses the Latin-1 as its default character set. To support special characters, change this to UTF-8 - otherwise files with special characters will not migrate. Note: If the character set is incorrect, the migration will fail and will not be able to progress further than the file with special characters. To set the correct character set, issue the following command:

```
protocol-character-set --all UTF-8
```

Virtual Server Security

The Virtual Secure Servers feature is compatible with Data Migrator to Cloud, provided the following requirements are met:

- A cloud target can be resolved in a DNS server configured in Global Context
- A route from the aggregate ports to the cloud provider server (HCS, HCP, AmazonS3, or Azure) exists on all nodes

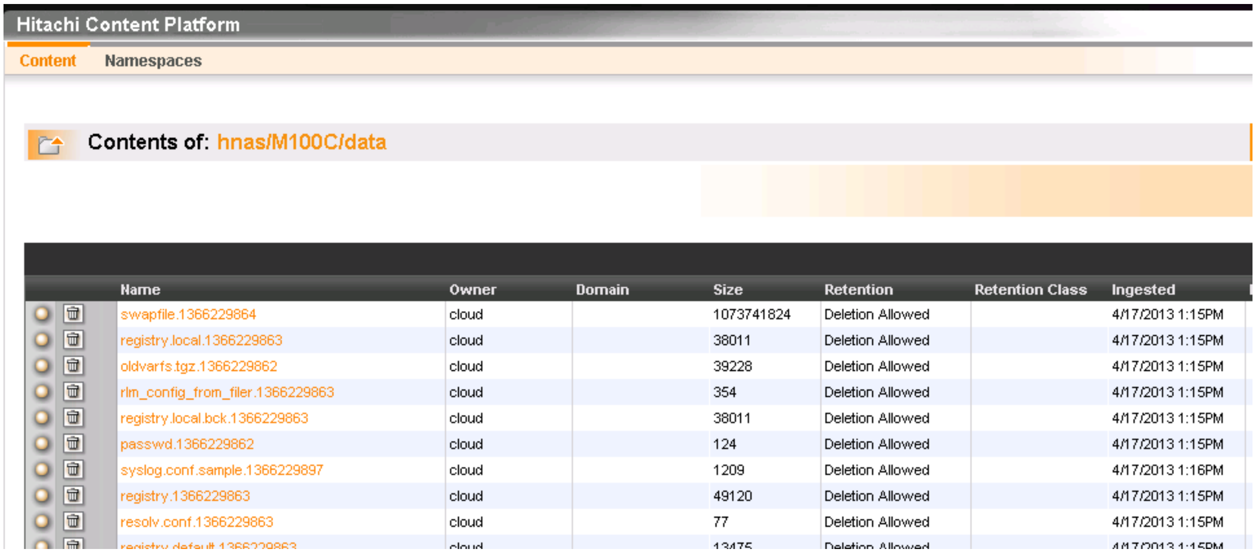
Multi-tenancy

Currently, multi-tenancy is not supported with the Data Migrator to Cloud.

Hitachi Content Platform view of XVLs

When using HCP as cloud target, the HCP administrator can optionally browse the namespace. To browse the namespace, login to the Namespace Console.

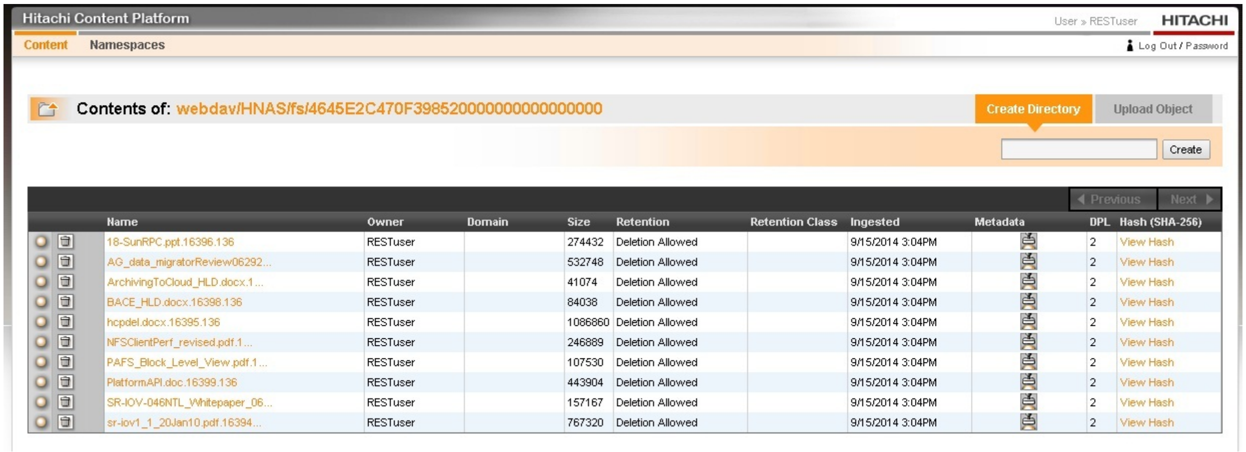
The first illustration shows the view in HCP when using Classic Data Migrator. The second screenshot shows the view of HCP when using Data Migrator to Cloud. With Data Migrator to Cloud, a directory with the UUID of the source file system is created which then holds any migrated data for that file system. This is shown in the second illustration, the directory starting with 743A2.



Contents of: hnas/M100C/data

		Name	Owner	Domain	Size	Retention	Retention Class	Ingested
●	🗑	swapfile.1366229864	cloud		1073741824	Deletion Allowed		4/17/2013 1:15PM
●	🗑	registry.local.1366229863	cloud		38011	Deletion Allowed		4/17/2013 1:15PM
●	🗑	oldvarfs.tgz.1366229862	cloud		39228	Deletion Allowed		4/17/2013 1:15PM
●	🗑	rlm_config_from_filer.1366229863	cloud		354	Deletion Allowed		4/17/2013 1:15PM
●	🗑	registry.local.bck.1366229863	cloud		38011	Deletion Allowed		4/17/2013 1:15PM
●	🗑	passwd.1366229862	cloud		124	Deletion Allowed		4/17/2013 1:15PM
●	🗑	syslog.conf.sample.1366229897	cloud		1209	Deletion Allowed		4/17/2013 1:16PM
●	🗑	registry.1366229863	cloud		49120	Deletion Allowed		4/17/2013 1:15PM
●	🗑	resolv.conf.1366229863	cloud		77	Deletion Allowed		4/17/2013 1:15PM
●	🗑	registry.default.1366229863	cloud		13475	Deletion Allowed		4/17/2013 1:15PM

The second screenshot shows the view of HCP when using Data Migrator to Cloud. With Data Migrator to Cloud, a directory with the UUID of the source file system is created which then holds any migrated data for that file system. In the illustration, it is the directory starting with 743A2.



Contents of: webdav/HNAS/fs/4645E2C470F398520000000000000000

		Name	Owner	Domain	Size	Retention	Retention Class	Ingested	Metadata	DPL	Hash (SHA-256)
●	🗑	18-SunRPC.ppt.16396.136	RESTuser		274432	Deletion Allowed		9/15/2014 3:04PM		2	View Hash
●	🗑	AG_data_migratorReview06292...	RESTuser		532748	Deletion Allowed		9/15/2014 3:04PM		2	View Hash
●	🗑	ArchivingToCloud_HLD.docx.1...	RESTuser		41074	Deletion Allowed		9/15/2014 3:04PM		2	View Hash
●	🗑	BACE_HLD.docx.16398.136	RESTuser		84038	Deletion Allowed		9/15/2014 3:04PM		2	View Hash
●	🗑	hcpdel.docx.16395.136	RESTuser		1086860	Deletion Allowed		9/15/2014 3:04PM		2	View Hash
●	🗑	NFSClientPerf_revised.pdf.1...	RESTuser		246889	Deletion Allowed		9/15/2014 3:04PM		2	View Hash
●	🗑	PAFS_Block_Level_View.pdf.1...	RESTuser		107530	Deletion Allowed		9/15/2014 3:04PM		2	View Hash
●	🗑	PlatformAPI.doc.16399.136	RESTuser		443904	Deletion Allowed		9/15/2014 3:04PM		2	View Hash
●	🗑	SR-IOV-046NTL_Whitepaper_06...	RESTuser		157167	Deletion Allowed		9/15/2014 3:04PM		2	View Hash
●	🗑	sr-iov1_1_20Jan10.pdf.16394...	RESTuser		767320	Deletion Allowed		9/15/2014 3:04PM		2	View Hash

Hitachi Data Systems

Corporate Headquarters
2845 Lafayette Street
Santa Clara, California 95050-2639
U.S.A.
www.hds.com

Regional Contact Information

Americas
+1 408 970 1000
info@hds.com

Europe, Middle East, and Africa
+44 (0) 1753 618000
info.emea@hds.com

Asia Pacific
+852 3189 7900
hds.marketing.apac@hds.com

MK-92HNAS045-05

Hitachi Data Systems