ABSTRACT
This white paper describes the Dell EMC Cloud Tiering Appliance (CTA) and how it can be used with Dell EMC Unity systems. CTA allows administrators to move block and file data from a Dell EMC Unity storage system to and from the cloud. CTA also facilitates cloud repository migration, stub aware file migration, and the usage of Dell EMC Unity as a destination.

November 2017
# TABLE OF CONTENTS

## EXECUTIVE SUMMARY
- Audience ................................................................. 6
- Terminology .............................................................. 6

## CLOUD TIERING APPLIANCE OVERVIEW
- Introduction .................................................................. 7
- Licensing ..................................................................... 7
- CTA Configuration ........................................................ 7
- Cloud Repositories ....................................................... 8
- Scheduler .................................................................... 8
- Policy .......................................................................... 8
- Simulation .................................................................... 9
- Compression ............................................................... 9
- Encryption .................................................................... 9
- Compatibility ............................................................. 9
- Common API Settings .................................................. 9
- CTA Backup and Restore ............................................... 10

## CTA FOR FILE
- Distributed Hierarchical Storage Management (DHSM) .......... 12
- File Policies .................................................................. 12
- Providing File Data to CTA ............................................ 14
- Recall Policy ............................................................... 15
- Recall Using CTA-HA ..................................................... 16
- CTA Database with File Data .......................................... 17
- Stub Scanner Jobs ......................................................... 17
- Orphans ........................................................................ 18
- File General Workflow .................................................. 18
- Reporting ...................................................................... 19

## CTA FOR BLOCK
- Providing Block Data to CTA ........................................... 20
- Block Archiving ............................................................ 20
- Block Restore Policy ...................................................... 21
- Block General Workflow ............................................... 21
FILE MIGRATION ........................................................................................................... 22
  Migration Source ........................................................................................................ 23
  Migration Targets ........................................................................................................ 23
  File Migration Workflow ............................................................................................. 25
  Stub Migration ............................................................................................................. 26

MISCELLANEOUS ........................................................................................................ 27
  Dell EMC Unity as a Destination .................................................................................. 27
  Troubleshooting ........................................................................................................... 27
  CTA Ports .................................................................................................................... 28

CONCLUSION ............................................................................................................... 28

REFERENCES ............................................................................................................... 28
EXECUTIVE SUMMARY
The Dell EMC Cloud Tiering Appliance (CTA) enables administrators to automatically move data from Dell EMC Unity to and from the cloud based on user-configured policies. This ability optimizes primary storage usage, dramatically improves storage efficiency, shortens the time required to back up data, and reduces overall TCO for primary storage.

Figure 1 shows a graphical representation of CTA actions for the Dell EMC Unity. CTA can be used to tier file data and archive block data to the cloud, as shown by the Tiering/Archiving diagram in Figure 1. Likewise, CTA can be used to recall file data and restore block data from cloud, as pictured in the Retrieval diagram in Figure 1. Lastly, CTA can be used to migrate repositories between clouds, as shown in the Repository Migration diagram in Figure 1. Figure 16 in the File Migration section provides a graphical representation for the File migration feature.

SCOPE
This white paper describes the following CTA features for Dell EMC Unity:

- File tiering
- File migration
- Block archiving

For file tiering, CTA acts as a policy engine by interacting with the Dell EMC Unity storage system and identifying files that fit an administrator-defined criteria. For these files, CTA initiates movement to a target cloud repository and places an 8 KB stub file in the original file location. When a client reads the stub, CTA recalls or passes the IO through the original file located in the cloud. To the user, the file appears to be in its original location on the Dell EMC Unity storage system. The stub utilizes only 8 KB of space, instead of the full size of the file. This solution makes file tiering seamless for the user who reads the archived file data.

For block archiving, CTA acts as a policy engine by interacting with the Dell EMC Unity storage system and identifying block snapshots that fit an administrator-defined criteria. For these snapshots, CTA initiates a backup to a target cloud repository and leaves the original snapshot untouched. After the snapshots are backed up to the cloud, they can be erased from the source system to free space. The user has the option to restore a snapshot to a new block storage resource located in the source system or in a new storage system. Block archiving is available only for Dell EMC Unity storage systems.

For file tiering and block archiving, the supported target cloud repositories are Virtustream, Dell EMC Elastic Cloud Storage (ECS), Microsoft Azure, Amazon S3, and IBM Cloud Object Storage (COS). CTA is supported with other storage platforms and has additional functionalities that are not applicable to Dell EMC Unity, for which reason, these functionalities are not discussed in this white paper. For a more in-depth look at CTA’s functionalities, refer to the Cloud Tiering Appliance Getting Started Guide found on Dell EMC Online Support.

For file migration, CTA acts as a policy engine by interacting with the source VNX or NetApp storage system and the target Dell EMC Unity storage system. CTA identifies files in the source system that fit an administrator-defined criteria and moves them from the source system to the target system. For source files that had been tiered to a cloud repository, the movement is stub aware, meaning that it maintains the stubs without recalling the files to the target.
**AUDIENCE**

This white paper is intended for Dell EMC customers, partners, and employees who are considering the use of Cloud Tiering Appliance (CTA) for the Dell EMC Unity family of storage systems. It assumes familiarity with Dell EMC Unity, Dell EMC’s management software and Cloud Tiering Appliance (CTA).

**TERMINOLOGY**

**Amazon S3** – Amazon Simple Storage Service (Amazon S3) is an object storage cloud platform from Amazon.

**Block archiving** – A primary CTA function that scans the block snapshots that meet an administrator-defined criteria, and takes a backup of the block resource to the cloud.

**DHSM (Distributed Hierarchical Storage Management)** – An API on the Dell EMC Unity storage system that enables the file stubbing and recall of archived files.

**ECS** – Elastic Cloud Storage (ECS) is an object storage platform from Dell EMC.

**CTA-HA** – Optional CTA software that is used to configure encryption. Also used for recalling files when the primary CTA is not available.

**File tiering** – A primary CTA function that scans a file system for files that meet an administrator-defined criteria and moves them to the cloud. CTA replaces the file on the file system with an 8 KB stub file that points to the full size file in the cloud repository.

**File migration** – The movement of files from a source location to a destination target.

**FileMover** – A service on the Data Mover that provides an API that CTA uses for file tiering and file migration on the VNX.

**IBM COS** – IBM Cloud Object Storage (COS) is an object storage platform from IBM. It might also be known as Cleversafe.

**LUN** – A block based storage resource that a user provisions. It represents a SCSI logical unit.

**NDMP (Network Data Management Protocol)** – An open standard network protocol that is designed for enterprise-wide backup and recovery of heterogeneous network-attached storage.

**Orphan file** – A file in a repository that does not have a stub file pointing to it. When a file is archived, a stub on the source NAS server points to the archived file. Deleting a stub does not automatically delete the archived file. Instead, the archived file becomes an orphan. To delete orphans, run an orphan delete job on the repository.

**Policy** – A combination of rule(s) and repositories for tiering and archiving. A policy might contain rules, for example, that would send files that have not been accessed in two years or files older than 3 weeks to a public Amazon S3 repository.

**Recall** – When a file has been tiered to a repository, and the user clicks on the stub file to quickly access the original file.

**Repository** – The target cloud location for the tiered file data and archived block data.

**Restore** – A method of transferring archived block data to a new storage resource.

**Snapshot** – A point-in-time view of a storage resource. When a snapshot is taken, the snapshot is an exact copy of the source storage resource, and shares all blocks of data with it. As data changes on the source, new blocks are allocated and data is written to them.

**Storage resource** – The top-level object that a user can provision and associated with a specific quantity of storage. All host access and data protection activities are performed at this level. In this document, storage resource refers specifically to those resources that support file tiering (NAS Servers and file systems) and block archiving (LUNs and Consistency Groups).

**Stub** – An 8 KB file that replaces the original file on the Dell EMC Unity file system. The stub file contains all the metadata associated with the archived file that is required for accessing the archived data on the cloud repository.

**Thin Clone** – A read-write copy of a Thin Block storage resource (LUN, Consistency Group, or VMware VMFS Datastore) that shares blocks with the parent resource. Thin Clones can have Snapshot Schedules assigned to them and can be replicated.

**Unisphere** – A web-based graphical user interface for managing the Dell EMC Unity storage system.
CLOUD TIERING APPLIANCE OVERVIEW

INTRODUCTION

CTA can help customers achieve many benefits, which include reducing capital expenses by reclaiming capacity on primary storage, lowering operating expenses by reducing the number of administrative tasks, and improving performance by reducing backup times. Figure 2 shows that by leveraging CTA to move data to the cloud, there is less data to be back up than without CTA.

![Diagram showing Reduced backup window with CTA.]

The three primary CTA functions that we will discuss are file tiering, file migration, and block archiving. File tiering moves file data to the cloud and leaves behind an 8 KB stub file. In this whitepaper, file tiering describes the functionality of leaving an 8 KB stub in place of the original file, but in other documentation tiering and archiving might be used interchangeably. In this whitepaper, archiving describes block archiving, which is the method of taking a backup of block data to the cloud. File migration is the action of moving files from a source array to Dell EMC Unity. This paper also discusses some additional aspects that should be considered to successfully configure CTA with a Dell EMC Unity storage system. In addition to archiving to the cloud, a Dell EMC Unity system can be used as the archiving destination for the VNX and NetApp storage systems. However, Dell EMC Unity does not support tiering to other Dell EMC storage systems or SMB/NFS shares.

The following sections give an overview of some of the CTA functionalities and components that apply to file tiering, file migration, and block archiving.

LICENSING

CTA is supported on all Dell EMC Unity systems and is included at no additional cost. See the Compatibility section for the minimum Dell EMC Unity OE requirements. The supported systems include the All Flash models, the Hybrid models, and the Dell EMC UnityVSA.

CTA CONFIGURATION

Two versions of the Cloud Tiering Appliance (CTA) are available – CTA and CTA/VE. The CTA is the deployment of the CTA software on a physical server, while the CTA/VE is deployed on top of an ESXi host. For the full details on what hardware is supported and the steps to set up a physical CTA server, see the sections in the Cloud Tiering Appliance Getting Started Guide that provide the qualified hardware and the appliance setup. The Installing the Virtual Edition section in the Cloud Tiering Appliance Getting Started Guide provides the requirements and steps to install the CTA/VE version on an ESXi host.

To use CTA with Dell EMC Unity, you must complete some configuration steps on the CTA. The Deploying CTA with Dell EMC Unity section in the Cloud Tiering Appliance Getting Started Guide, provides the full procedure to setup the CTA with the Dell EMC Unity storage system.
**CTA-HA AND CTA/VE-HA**

Both CTA editions, CTA and CTA/VE, include High Availability (HA) options for recall purposes. Both the CTA-HA and the CTA/VE-HA deliver solutions for redundancy, which ensure that clients can always access their data in case the primary appliance fails.

CTA-HA also performs keystore replication for encryption keys, which are required when the encryption feature is enabled during tiering and archiving to cloud platforms. Without CTA-HA, encryption will not work because the key generator cannot generate keys without the key replication daemon.

Refer to the *CTA and CTA/VE for High Availability* and *File Encryption* sections in the *Cloud Tiering Appliance Getting Started Guide* for full details to deploy and configure CTA-HA and encryption. The *Cloud Tiering Appliance Getting Started Guide* is available on Dell EMC Online Support.

**CLOUD REPOSITORIES**

The supported target cloud repositories from a Dell EMC Unity array include:

- Public clouds
  - Virtustream
  - Microsoft Azure
  - Amazon S3
  - IBM Cloud Object Storage (COS)

- Private clouds
  - Dell EMC Elastic Cloud Storage (ECS)

**CLOUD REPOSITORY MIGRATION**

CTA also supports the repository migration of one supported cloud to another supported cloud. The repository can be migrated when the repository is being used for file tiering or block archiving. Repository migration moves all archived data from one repository to another. After repository migration, the original repository can be removed. For file tiering, the stubs on the Dell EMC Unity primary storage system will point to the archived files on the new repository.

Repository migration with cloud storage retains the same encryption key when moving data. However, if an encrypted data that was archived without compression is migrated to cloud storage with compression, the destination data will be re-encrypted with the most recent encryption key.

Repository migrations are useful when you need to replace the cloud platform. The cloud target must have enough space to hold the data that is being moved, but the space does not need to be the same size or have the same layout as the source. Refer to the *Cloud Tiering Appliance Getting Started Guide* for more details on how to configure CTA for repository migration.

**SCHEDULER**

The CTA scheduler sets the task start time. For example, you can schedule an archive task to start at 2 a.m. on Saturday to scan share01 and evaluate the files with a policy for tiering to the cloud. You can also do capacity-based scheduling for file tiering.

Administrators usually schedule a task to run on a regular cadence: weekly, every other week, or monthly. The first archive task often tiers the most data and can require a substantial amount of time. Subsequent jobs move incremental amounts of data so they will run faster.

**POLICY**

You can use CTA to define a task to perform a series of actions, for example:

- Move any files that are larger than 10MB and have not been accessed in 30 days to the cloud
- Move any block snapshots that are more than 60 days old to the cloud

A policy is a set of one or more user-created rules. For each rule, you can associate an archive, recall, or restore action with the rule. When a policy is applied to a storage resource in CTA, the rules are applied to each share or snapshot under the storage resource. If a file or snapshot falls within the rule, the action is taken. The sections for file tiering and block archiving functions provide more details with regard to rules and actions that apply to each one.
SIMULATION
You can schedule a task with a policy, and then run the task in simulation mode. In simulation mode, CTA scans the source shares or snapshots, and applies the policy rules against each one, but CTA does not perform any tiering or archiving action. Instead, CTA tracks the number of files or snapshots and amount of data it would have tiered or archived. You can view a report at the end of the simulation by navigating to the History (Simulation) option for a task. Simulation is a good way to test the policy’s effectiveness and to edit the policy rules before you run a real job.

COMPRESSION
When moving data to a cloud repository, you can use CTA encryption and compression options. Encryption requires that you deploy and configure the CTA-HA or CTA/VE-HA.

The compression option is available when you create a policy to tier or archive to cloud repositories. If you select compression, it can be configured as either fast or strong. If you select the fast option, the compression process is fast, but the compressed data is not as small. If you select strong, the compressed data is smaller, but the compression process is slower. Compressed data in Dell EMC Unity will be uncompressed in memory and sent to the cloud repository. It is recommended to leverage the CTA’s compression to ensure that the space in the cloud repository is used efficiently.

ENCRYPTION
When moving data to a cloud repository, you can also use CTA encryption option. Encryption requires that you deploy and configure the CTA-HA or CTA/VE-HA. The encryption option is available when you create a policy to tier or archive to cloud repositories.

CTA stores the key in the keystore and replicates it to the CTA-HA. Every task that uses a policy with encryption enabled, uses the key. If a new key is generated, it will be used by any new tasks that uses encryption. The old keys will remain in the keystore and continue to apply to the data that was encrypted using the old keys.

Keystore replication is sufficient to sustain normal outages. However, Dell EMC recommends that you back up the CTA configuration to preserve the keystore after CTA generates a new key. Scheduled CTA backups can be configured as a Backup task from the CTA CLI and GUI. A one-time backup of the CTA configuration is possible from CLI using the fmbackup command. For more information about CTA backups, see the CTA Backup and Restore section.

Before you can enable encryption, you must meet the following prerequisites:

1. Generate an encryption key using the CTA GUI.
2. Install CTA-HA and run krdsetup to enable the encryption key and to start the keystore replication.

COMPATIBILITY
Table 1 below shows the Dell EMC Unity OE and CTA versions required per feature.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Dell EMC Unity OE</th>
<th>CTA</th>
<th>Source Storage System</th>
</tr>
</thead>
<tbody>
<tr>
<td>File tiering</td>
<td>Version 4.1 or later</td>
<td>Version 11 or later</td>
<td>N/A</td>
</tr>
<tr>
<td>File migration</td>
<td>Version 4.2 or later</td>
<td>Version 12 or later</td>
<td>NetApp: ONTAP 7.2 or later VNX: OE 7.0 or later</td>
</tr>
<tr>
<td>Block archiving</td>
<td>Version 4.2 or later</td>
<td>Version 12 or later</td>
<td>N/A</td>
</tr>
</tbody>
</table>

COMMON API SETTINGS
The Management and DHSM Credentials for the Dell EMC Unity are configured on CTA’s Common API Settings page, as seen in Figure 3. You need to populate the Management and DHSM credentials when configuring CTA with Dell EMC Unity. CTA will use the Management credentials when making any REST API calls to Dell EMC Unity. The DHSM credentials are used for stub management.
Note that the DHSM option can still be enabled manually on the NAS Server. DHSM is enabled per NAS Server, as seen in Figure 4.

CTA BACKUP AND RESTORE

BACKUP
To take a backup of the CTA configuration and database, you will need to set up the location of the backups in the **Backup and Recovery** page under **Configuration**, as shown Figure 5. Then, schedule the **Backup** task, as shown in Figure 6.
To restore from backup, navigate to the Backup and Recovery page under Configuration, as shown in Figure 5. Select a file in the Available Backup Files list under Restore Backup, and click Restore.

Refer to the Backing up the configuration section in the Cloud Tiering Appliance Getting Started Guide for more details on backing up and restoring.
CTA FOR FILE

The file tiering feature found on CTA allows the movement of file data to the cloud. The following sections of this paper will discuss:

- The benefits of file tiering
- How the file tiering works
- The general tasks that can be applied to file data

BENEFITS

Leveraging CTA file tiering you can use your Dell EMC Unity file shares more cost effectively and efficiently. For example, you can free up capacity on a Dell EMC Unity storage system by transferring your file data to the cloud and leaving an 8 KB stub in place of the files. File tiering can also decrease the time required to backup file data, because it eliminates the need to back up the full-size original files.

DISTRIBUTED HIERARCHICAL STORAGE MANAGEMENT (DHSM)

The Dell EMC Unity storage system has a native DHSM API, which moves a file from the storage system to the cloud. Moving the file leaves behind a small stub pointer in the file’s original location. This process of stubbing and relocating file data describes file tiering.

After the stubs files are in place, CTA constantly monitors them. At a basic level, DHSM intercepts client access to data and takes action before the client accesses the data. When a client tries to read or write to the file, CTA intercepts that access request and takes an action on the archived data using the stub’s information. See the Recall Policy section for more details with regard to the available recall actions.

REQUIREMENTS FOR THE SOURCE NAS SERVER

DHSM achieves the file tiering services. To identify stub files, DHSM reads the offline bit on the stubs. The SMB protocol supports offline bits, but NFS does not so Dell EMC Unity handles offline bits internally for NFS-only archival. CTA communicates with Dell EMC Unity using the DHSM API. Before archiving data from a storage array, the CTA must be configured with the details of the source array. CTA automatically creates the DHSM connections when they are needed by sending a REST API call to Dell EMC Unity. Users can also create the DHSM connection manually using Unisphere CLI (UEMCLI) on Dell EMC Unity.

Deleting the DHSM connection on a Dell EMC Unity using the UEMCLI can optionally trigger a recall of all stubbed data from the repository that is linked to the Dell EMC Unity file system that uses that connection. Before triggering a recall, ensure that enough space is available on the system for all of the recalled data. For more details on the UEMCLI, refer to the Unisphere Command Line Interface Guide found on Dell EMC Online Support.

When tiering from SMB shares, the source server must belong to a domain and the CTA configuration settings require a username and password from that domain. The username should be part of the SMB servers’ local Administrators group. By default, all members of the Active Directory Domain Admins group are members of the local Administrators group on the Dell EMC Unity SMB Server. The fully qualified domain name (FQDN) and IP address of the Domain Controller should be configured, in the CTA’s Fully Qualified Domain page, for Kerberos authentication. If the source includes NFS shares, these shares must have root and read/write permission for CTA and CTA-HA IP addresses.

FILE POLICIES

The following table, Table 2, provides the available tasks for file data:

<table>
<thead>
<tr>
<th>Task</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archive</td>
<td>Checks files on a source against policy rules and tiers the files to a repository</td>
</tr>
<tr>
<td>Multi Tier</td>
<td>Identifies and acts on files that match the multi-tier archive policy</td>
</tr>
<tr>
<td>Multi Tier Stub</td>
<td>Updates stubs files and relocates archived data per the multi-tier stub archive policy</td>
</tr>
<tr>
<td>Delete Orphans</td>
<td>Identifies orphans and deletes the ones that fit the delete orphan policy</td>
</tr>
<tr>
<td>Delete Stubs</td>
<td>Used to delete stubs</td>
</tr>
<tr>
<td>Stub Scan</td>
<td>Verifies the source for stub files</td>
</tr>
<tr>
<td>Recall</td>
<td>Used to recall and replace stubs with normal files on the source</td>
</tr>
</tbody>
</table>

Table 2. Available CTA tasks for file data.
A CTA policy for tiering consists of one or more rules and destinations. For example, a simple policy might read:

If this file has not been accessed in 6 months, send it to the ECS, and replace it with a stub.

The one-rule, one-destination policy is common, and many CTA users use this type of policy on their data. However, CTA rules are flexible. CTA supports the creation of more complex rules that archive data to multiple tiers. For example, a policy consisting of multiple rules might look like this:

If any file has not been accessed in more than 1 year and is larger than 1 MB in size, send it to Virtustream and leave a stub. However, if the file is a PDF, don’t archive it at all. Then, when these PDF files have not been accessed for two years, move them to Amazon S3, and update the stub file to point to the new location.

Policy rules are based on attributes such as access time, modify time, change time, file size, file name, or directory name. The archive policy action is either “archive” or “don’t archive.” A single expression or a combination of expressions defines the archive policy. See Figure 7 for the example of the attributes in the CTA GUI.

A policy is not associated to a share name hence a single policy can be used by multiple tasks to evaluate different shares. An Archive policy could move files to different cloud repositories by having multiple rules that point to different cloud repositories.

**MULTI-TIERED ARCHIVE POLICY**

CTA supports multi-tiered archiving, a feature used to specify how files of different ages, for example, can be stored to different types of repositories. A Multi-Tiered Archive policy applies to normal files as well as to files which are already stubbed.

Consider the following example of a multi-tiered archive:

Find files that have not been accessed in 1 year, and tier them to Virtustream public cloud storage. Find files that have not been accessed in 6 months, and send them to ECS private cloud.

By creating a multi-tiered policy type with several rules, each with a different repository, you can design data movement schemes to fit your needs. For the previous example, the multi-tiered archiving policy could have the following rules:

1. If access time > 1 year, archive to Virtustream public cloud
2. If access time > 6 month, archive to ECS private cloud
When files that were tiered to ECS are not accessed for 1 year, then they qualify for the first rule. Consequently, the files will be migrated from ECS to Virtustream with the stubs being updated accordingly.

Figure 8 provides an example of the multi-tiered policy in CTA.

The order of the rules is important because they are applied sequentially. When the first rule evaluates as true, CTA takes the action this rule specifies, either “archive” or “don’t archive.” CTA does not apply the subsequent rules, and the policy moves on to the next file.

In the example shown in Figure 8, reversing the order of Rule 1 and Rule 2 would produce an unintended result. The 6-month old rule would be applied first, and the 1-year old rule would never be applied because any file older than 1 year is also older than 6 months. All data older than 6 months would be archived to the private ECS cloud, and no files would be archived to the Virtustream. If there are multiple rules in the policy, CTA continues to apply the rules until a rule evaluates to “true.” It then takes the action associated with the rule (such as “archive” or “don’t archive”) and moves on to the next file.

PROVIDING FILE DATA TO CTA

File tiering is supported on the following storage resources: SMB and NFS Shares on Dell EMC Unity NAS Servers. Administrators usually direct a CTA archive policy to evaluate an SMB or NFS share. CTA scans the files and applies the policy rules to each file, one rule at a time.

A CTA policy can also be directed at specific files. Instead of directing CTA to scan an entire SMB or NFS share, you might instead import a list of filenames; then CTA will scan and apply the archive policy only to the files in that list. The Cloud Tiering Appliance Getting Started Guide describes the file ingest feature and is available from Dell EMC Online Support.

DELAYED STUBBING

When you use one of the following policy types: Archive, Multi Tier, or Multi Tier Stub, you can set a delayed stubbing option. Figure 9 shows the Define Policy Wizard when creating a multi-tier policy. As the figure shows, you can set the Delay Period property, for which you can provide the number of days when the files should be stubbed on the source. By default, the setting is set to 0 days, so there is no delayed stubbing.
RETENTION

You can also set a Retention Period and Stub Retention parameters for the Archive, Multi Tier, or Multi Tier Stub policies, as seen in Figure 9. The Retention Period states the period of time for which a file is to be retained in the repository. If Stub Retention is selected, the same retention period applies to the stubs of all the archived files. Before the retention period expires, the stubs that have retention enabled cannot be deleted by the users or by a delete task from CTA. Once the retention period expires, stubs can be deleted by the user from the share manually or by running a delete task.

RECALL POLICY

When a file has been tiered to a cloud repository and replaced with a stub on the source share, the stub should look and behave like the original file. When a file has been tiered, as seen in Figure 10, it will have an X icon and the size on disk attribute for the file will be 8 KB. File recall is the process by which the user clicks on the stub file and quickly accesses the original file.

The stub file contains information needed to find the actual file. The Dell EMC Unity sets the offline bit on the stub when the file is archived. When a user attempts to read a stub file, the Dell EMC Unity sends the recall request to CTA, which then executes the recall and passes the file to the Dell EMC Unity system.
When recalling the file, Dell EMC Unity performs one of the actions given in Table 3:

<table>
<thead>
<tr>
<th>Action</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passthrough recall (default)</td>
<td>Retrieves the data without recalling the data to NAS Server. The stubs stay in place.</td>
</tr>
<tr>
<td>Partial recall</td>
<td>Recalls only as much of the file’s contents as it needs to satisfy the client read request and saves the partial file on Dell EMC Unity.</td>
</tr>
<tr>
<td>Full recall</td>
<td>Writes the file back to its original location and deletes the stub.</td>
</tr>
<tr>
<td>None</td>
<td>Specifies no override of the method that is specified in the stub file.</td>
</tr>
</tbody>
</table>

In Dell EMC Unity, you can use the Unisphere CLI to set the recall style, which can be configured on each file system. The usage of the optional readPolicy flag that is available for the /net/nas/dhsmconn create and modify commands follows.

```
uemcli -d x.x.x.x -u username -p password /net/nas/dhsmconn -help
```

Manage DHSM connections between the specified primary file system of Unity and a secondary file system of CTA.

Actions:

- [Create/Modify]
- ...
- [Optional] -readPolicy { none | full | passthrough | partial }

Specify the migration method option used by the NAS Server in the connection level to override the migration method specified in the stub file.

**RECALL USING CTA-HA**

If an archive task job fails, no data is lost. Simply correct the problem and then rerun the job. For this reason, the complexity of a High Availability (HA) configuration for moving data to the cloud is not necessary or justified. However, recalls are mission-critical because they affect clients’ ability to access their data. Therefore, configurations where CTA is in the recall path require a CTA-HA to be configured.

CTA-HA is a recall-only version of CTA. The HA configuration pairs the CTA-HA physical appliance or the CTA/VE-HA virtual appliance with one or more CTA or CTA/VE systems. If the primary CTA cannot perform the recall, its associated CTA-HA partner can.

By creating a DNS hostname that maps to the IP addresses of both the CTA and CTA-HA appliances and is configured in the ACD DNS Name property for the Dell EMC Unity Fileserver in CTA, as shown in Figure 11 and Figure 12, either CTA or CTA-HA could be used when performing recalls for that Dell EMC Unity Fileserver in a round-robin fashion. This setup balances the recall load. If recalls fail with one of the appliances, the other appliance can perform recalls until the failed appliance returns to service. This configuration also allows maintenance of one appliance while the other continues to perform recalls.
CTA DATABASE WITH FILE DATA

When a file is archived to a cloud repository, the stub in the source NAS Server points to the file location in the cloud repository. However, the file in the cloud repository has no pointer back to the source, which means that repository files have no connection to the source. Archived files are not named the same way on the cloud repository as they are on the source.

The CTA database solves this problem. Each time a file is archived, an entry in the CTA database records the file's original location on the source and the file's location in the repository. The database includes entries for every archived file. CTA does not use the database for recalls because the stub on the source includes the information required to locate the file in the repository.

STUB SCANNER JOBS

For every scheduled archive job, CTA automatically schedules a monthly Stub Scan task. Stub Scan is a utility that reads the stubs in a share and compares them to the entries in the CTA database. If stubs move to different locations or if orphans appear, Stub Scan ensures that the CTA database is kept current.

Because a stub on the source has the information necessary to recall a file from the repository, CTA does not need to query for stub and repository file locations in the CTA database. However, CTA can more efficiently manage the repository storage if information in the database is synchronized with the system. You can run the Stub Scan task more frequently than the default 30 days, but doing so generally not necessary.
**ORPHANS**

Orphans are created when stub files are deleted from the source, as the actual files in the repository are not automatically deleted and become orphans. The files are not deleted automatically to prevent unwanted deletion of a file being used by a service or application. If CTA had deleted the archived files when the stubs were deleted, restoring the backup would restore stubs that point to nothing.

To delete orphan files and recover space on the repository, run the Delete Orphans task. Do not delete orphans until you are certain that you will not restore stubs that point to orphans. For example, if backups are kept for six months, define the orphan deletion job to delete files that have been orphans for at least that long.

The CTA database and the Stub Scan jobs play important roles in managing orphan files. Every time the Stub Scan sees a stub, CTA records a "last seen" time in its database. If the stub is deleted, the stub scanner identifies the file in the repository that was linked to the stub as an orphan. Because it records the "last seen" time in its database, CTA can calculate how long the file has been an orphan. CTA uses the orphan age to determine which orphans to delete. You can set a policy to delete stubs that fit a criteria that you set. For example, delete any orphans larger than 2 MB or older than 4 months.

**FILE GENERAL WORKFLOW**

The same workflow and procedures apply to both CTA and CTA/VE deployments.

**CONFIGURE FILE TIERING**

1) In CTA:
   a) Configure the Common API settings in CTA, as seen in Figure 3. CTA will use the management credentials to automatically enable DHSM on the NAS Server.

2) In Unisphere:
   a) Make sure that DHSM is enabled for the NAS Server for which file data will be tiered from. If not enabled, manually enable it, as seen in Figure 4.

3) In CTA:
   a) Add the cloud repository that will be used as a Cloud Server.

   **Note:** To use ECS and IBM COS as a destination repository, the AmazonS3 option can be used. See Figure 13 for an example of it in the CTA GUI. Refer to the Deploying CTA with ECS and Deploying CTA with IBM Cloud Object Storage (Cleversafe) sections in the CTA Getting Started Guide for the full details.
b) Configure Dell EMC Unity as a NAS File Server.

c) Create an Archive Policy Scheduled task.

   i) Select the share path from which files will be tiered from.

   ii) Select or add a policy for the files that will fit a predefined criteria, for example, any files greater than 2MB in size.

   iii) Select the schedule for the task.

d) Run the Archive task.

e) View Archived File List.

The Deploying CTA with Dell EMC Unity section in the Cloud Tiering Appliance Getting Started Guide, provides more details in order to configure file tiering.

REPORTING

CTA includes a robust reporting interface that provides valuable insight into the efficacy of file tiering policies. CTA generates reports on the files it tiers only. For tiered files, the reports display the size, number of files tiered, and breakdown by file types, but CTA does not give a detailed profile of the data in the file system. Figure 14 shows an example of a report for a share that had a variety of files, including PPT, XLS, PDF, and other files.
CTA FOR BLOCK

OVERVIEW

The CTA block archiving feature allows you to move block data to the cloud. The following sections of this paper will discuss:

- The benefits of block archiving
- How block archiving works
- The general tasks that apply to block data

BENEFITS

By leveraging CTA and its block archiving feature, you can more efficiently use your storage system. For example, you can free up space on the storage system by deleting source snapshots after they are backed up to the cloud. Block archiving can also fulfill any compliance need that you might have for your data, for example, when a requirement exists to keep snapshots for a longer period of time.

PROVIDING BLOCK DATA TO CTA

Block archiving is supported on the following storage resources:

- LUNs
- Consistency Groups
- Thin Clones

BLOCK ARCHIVING

CTA leverages the native snapshot differentials API in order to efficiently take backups of the block data to the cloud. Users can archive a full copy of a LUN, Consistency Group or Thin Clone to the cloud as well as all subsequent snapshots. By default, every thirtieth archived snapshot is a full-copy baseline snapshot. This can be customize using the Common Base Frequency attribute in the CTA management GUI. By setting this attribute to a custom value, CTA can establish a baseline for the restore operation that depends upon the set value.

Block snapshots that have been archived to the cloud are independent of the base resource on the source and are not altered by the archiving. Instead, they share blocks with the baseline snapshot established by CTA. After the snapshot is archived, the LUN and/or snapshot can be deleted in the source array to free space.

BLOCK ARCHIVE POLICY

You can use a Block Archive policy to analyze a Dell EMC Unity storage system for the block snapshots and archive them to the cloud. Figure 15 shows the snapshot attributes that are used by CTA to evaluate the snapshots. CTA scans the snapshots and applies the policy rules to each snapshot, one rule at a time. If there are multiple rules in the policy, CTA continues to apply the rules until a rule evaluates to “true.” It then takes the action associated with the rule (such as “archive” or “don’t archive”) and moves to the next snapshot. For a Block Archive policy, there is the option to set the retention period for the snapshot in the cloud. You can set the retention policy to a longer period of time, to establish compliance with policies that a company or government agency might have and to increase the resiliency. A policy is not associated to a block resource (LUN or CG). Hence a single policy can be used by multiple tasks to evaluate different block resources.
BLOCK RESTORE POLICY
After block resource and its snapshots have been archived to the cloud, they can be restored back to the storage array. The Block Restore policy can be used to restore snapshots to a new block resource. The new block resource can reside in the same Dell EMC Unity storage systems from which it was archived or in a new Dell EMC Unity system.

PRE-REQUESITES
In Unisphere:

- Ensure that you have iSCSI Interfaces. For physical systems, it is recommended that you have at least one iSCSI Interface per SP to ensure that multiple paths are available.

- Ensure that the CTA has host access to the destination LUN or CG

BLOCK GENERAL WORKFLOW
The same workflow and procedures apply to both physical CTA and CTA/VE deployments.

CONFIGURE BLOCK ARCHIVING
1. In CTA:
   a) Configure iSCSI IQN targets
   b) Add the cloud repository that will be used as a Cloud Server.

   **Note:** To use ECS and IBM COS as a destination repository, the AmazonS3 option can be used. See Figure 13 for an example of it in the CTA GUI. Refer to the *Deploying CTA with ECS* and *Deploying CTA with IBM Cloud Object Storage (Cleversafe)* sections in the CTA Getting Started Guide for the full details.

c) Configure Dell EMC Unity as a Block Server
d) Create a Block Archive Policy Scheduled task
   i) Create a Snapshot Attribute Expression. This determines the snapshots that will be archived with the task, for example, manually user-created snapshots and/or snapshots created by a snapshot schedule
   ii) Create a Block Archive Scheduled Task
(1) Select the block LUN or CG that snapshots will be archived from

(2) Select the Common base Snapshot. This will be snapshot that be the full-copy of the block resource.

(3) Select the Block Archive Policy that you previously created

e) Run the Block Archive task

f) View Archived Snapshot List

**BLOCK RESTORE**

1. In Unisphere, if not already created, create a LUN or CG in which CTA will restore the archived snapshot(s)

2. In CTA:
   a) Create a Block Restore Scheduled Task
      i) Select the source block LUN or CG that was already archived
      ii) Select the snapshot to restore from the cloud
      iii) Select a destination LUN or CG. The destination resource can be located in a different Dell EMC Unity system.

   b) Run the Block Restore Task

The *Deploying CTA with Dell EMC Unity* section in the *Cloud Tiering Appliance Getting Started Guide*, provides more details in order to configure block archiving.

**FILE MIGRATION**

**OVERVIEW**

With CTA version 12 SP1 and later, the CTA file migration feature is supported with Dell EMC Unity systems. File migration moves files from one location to another while preserving stubs and without needing to re-hydrate the data.

File migration is a task in CTA to move data from a source File System to a target File System. The target File System needs to be large enough to hold the data being moved, but it does not need to be of the same size as the source File System. Administrators can use CTA to move data from a legacy array to Dell EMC Unity with minimal disruption to the clients. CTA can perform multi-protocol, incremental, stub-aware, cross-vendor migrations that can greatly reduce the effort and complexity when replacing legacy arrays and implementing their NAS environment on a Dell EMC Unity.

CTA supports SMB (CIFS), NFS, and multi-protocol source File Systems. When having a Dell EMC Unity as a target, the supported source platforms are VNX and NetApp. Figure 16 provides a graphical representation for file migration. File Migration requires both source and destination file systems to be shared as well as exported. Hence when migrating to Dell EMC Unity, the NAS Server on which the destination file system will reside should be Multiprotocol Enabled.

For source SMB (CIFS) servers, CTA needs access permissions of a domain user that is part of the Backup Operators and the Local Administrators group on the source and target File Systems. For NFS servers, the NFS exports need to have root and read/write permissions for the CTA IPs.
MIGRATION SOURCE

CTA Migration is policy-based and uses NDMP as its file transfer engine when migrating from VNX or NetApp. Figure 17 shows an example for a rule used to migrate all PDF files excluding files older than 3 years. CTA uses the FileMover to create snapshots and migrate files from VNX using NDMPCOPY. CTA creates snapshots on the source File System by making API calls. Snapshots enable users to continue to access and write to the source share while the migration is taking place. Note: The CTA-HA is not used for CTA Migration.

Figure 17. CTA. Creating a rule.

To migrate all the data without filtering any files, requires a simple policy with a rule similar to the following:

1. If size >= 0 bytes, then Migrate

MIGRATION TARGETS

Before starting a migration task, the target File System in the Dell EMC Unity must already exist. CTA does not create the target File System automatically. Migrating files to the root level requires that the directories to be empty, except for system directories such as lost+found. However, files can be migrated to any empty directory in a share.
The protocol of the target Dell EMC Unity File System must be multi-protocol. Requiring that the target NAS Server must be multi-protocol. For more information on how to configure a multi-protocol NAS Server on a Dell EMC Unity system, review the Configuring Multiprotocol File Sharing guide on Dell EMC Online Support.

For NDMP-style migrations, NDMP must be enabled on the Dell EMC Unity target NAS Server from Unisphere, as seen in Figure 18. By default, the ndmp username is used. The server configuration for the target server on CTA requires the same NDMP password as the one set in Unisphere for the ndmp username. For the source server the NDMP username in the server configuration could be different.

For NDMP-style migrations, NDMP must be enabled on the Dell EMC Unity target NAS Server from Unisphere, as seen in Figure 18. By default, the ndmp username is used. The server configuration for the target server on CTA requires the same NDMP password as the one set in Unisphere for the ndmp username. For the source server the NDMP username in the server configuration could be different.

![Figure 18. Unisphere. NAS Server Properties. Protection & Events. NDMP Backup.](image)

**MIGRATION PROCESS**

Migrations run as scheduled batch jobs. Schedule the migration task on the CTA Schedule page or through the CLI, where you specify the following information:

- A source share
- A policy
- An empty target share or directory

When the job begins, CTA creates a snapshot on the source. The CTA copies the data that matches the policy to the target. The migration can begin anywhere in the source share (for example, at the top level or in a subdirectory), and can move files to any empty share or directory on the target.

After the first pass completes, the target has a copy of the data from the snapshot on the source. However, if users have continued updating the source during migration, the source share will probably have changed from the snapshot. You can run a second pass to pick up those incremental changes. CTA will create another snapshot, compare the second to the first, and only transfer the changes between the two snapshots such as new files, deleted files, metadata changes, and so forth.

You can continue running passes as needed to pick up changes. Prior to running the final pass, put the source in offline or in read-only mode to ensure that the target is identical to the source. To complete the migration, transfer the client mounts to the target. The target has copied all CIFS and/or NFS ACLs or permissions. After transferring the clients, delete the source share and recover its storage space on the source NAS system.

CTA can run migration tasks in a continuous mode. The migration task has the option of running incremental migrations until reaching a files-moved threshold. The threshold can be set when creating a File Migration task, by selecting Automatic Recursive File Migration and setting a File Threshold Limit in the Schedule step, as shown in Figure 19. For example, you might want to continue running incremental migrations until fewer than 1000 files are moved in one run. At that point, you would want the system to stop and notify you. Administrators typically perform incremental runs during scheduled off-hour periods over the course of several evenings, with the final locked cutover also performed during off hours.
If you wish to throttle CTA migration tasks, you can specify a maximum bandwidth rate when creating the task schedule. This way, the migration does not consume all the network bandwidth. You can also use the CTA SID translation tool to create a SID mapping of local to domain SIDs. CTA applies the mapping at migration time to ensure that the SIDs on the target will be correct. This guards against having permissions issues after the migration is complete. For more information on the SID translation tool see the SID translator section in the Cloud Tiering Appliance Getting Started Guide on Dell EMC Online Support.

In addition, CTA can migrate stub files. For a stub aware File Migration from VNX/NetApp to Dell EMC Unity, source file should be archived to a cloud repository. If the source and target both support CTA stubs and are properly configured, CTA migrates stubs without recalling the archived data and the stubs continue to function after migration. To confirm that the servers are configured properly, run a Verify from the Server page for the source and target servers.

**FILE MIGRATION WORKFLOW**

The same workflow and procedures apply to both CTA and CTA/VE deployments.

**CONFIGURE FILE MIGRATION**

1) In CTA:
   a) Under the Common API Settings page, configure the Dell EMC Unity Management and DHSM Credentials. For Dell EMC Unity, CTA will use the management credentials to automatically enable DHSM on the NAS Server. If migrating from a VNX source system also configure the VNX Management credentials, reference Figure 3.

2) In Unisphere:
   a) Create a new multiprotocol target NAS Server.
   b) Enable NDMP on the target NAS Server and set a password for the ndmp username. As shown in Figure 18.
   c) Create a multiprotocol File System and create a multiprotocol share on it.

3) In CTA:
   a) Add the source system as a NAS File Server. Making sure to select Yes in the Server as Source step. The option added will depend on the source type, either VNX or NetApp.
   b) Add the Dell EMC Unity target NAS Server as a NAS File Server. In the Dell EMC Unity as Source step select No since the Dell EMC Unity is a target.
   c) Create a Scheduled task for a File Migration Policy.
i) Select the Primary Server from which files will be migrated from. Figure 20 shows an example of the Primary Server step when creating a File Migration task.

![Figure 20. CTA. File Migration. Primary Server.](image)

ii) Select the Secondary Server to which files will be migrated to.

iii) Select or add a policy for the files that will fit a predefined criteria, for example, any files greater than 2MB in size.

iv) Select the schedule for the task.

d) Run the File Migration task.

e) The files that fit the predefined criteria are migrated to the target Dell EMC Unity system. To migrate all the data without filtering any files, the policy configured needs to have a rule set to size >= 0 bytes, then Migrate.

## STUB MIGRATION

As an alternative to the CTA file migration feature, you can migrate stubs in a VNX system to a Dell EMC Unity system without recalling the files, by using EMCopy. The EMCopy tool can be downloaded from Dell EMC Online Support. Table 4 provides the supported version for the Dell EMC Unity, VNX, and EMCopy.

<table>
<thead>
<tr>
<th>System/Tool</th>
<th>Supported version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dell EMC Unity</td>
<td>OE 4.1.2 or later</td>
</tr>
<tr>
<td>VNX</td>
<td>OE 7.0 or later</td>
</tr>
<tr>
<td>EMCopy</td>
<td>4.17 or later</td>
</tr>
</tbody>
</table>

## MIGRATE STUBS

1. In VNX, after the file data has been tiered, change the DHSM settings per file system to set the recall style to offline.

   a) Below is the usage of `backup` switch available for the VNX `fs_dhsm modify` command:

   ```bash
   [nasadmin@dr-vnx-cs0 ~]$ fs_dhsm
   -list
   | -info [<fs_name>|id=<fs_id>]``
b) From an SSH session to the VNX:

i) Get a list of all file systems with DHSM enabled by typing the following command: `fs_dhsm -list`

   **Example:**
   ```shell
   [nasadmin@dr-vnx-cs0 ~]$ fs_dhsm -list
   id      name
   220     cta_fs
   36      Test_FS1
   171     ctamigvnx
   ```

ii) Type the following command: `fs_dhsm -modify <file system name> -backup offline`.

   **Example:**
   ```shell
   [nasadmin@dr-vnx-cs0 ~]$ fs_dhsm -modify Test_FS1 -backup offline
   Test_FS1:
   state                = enabled
   offline attr         = on
   popup timeout        = 0
   backup               = offline
   ```

   Ensure that the backup method is changed to offline for all of the file systems for which stubs will be migrated.

2. In Dell EMC Unity, create the destination NAS Server with file systems that have enough space for the migration of the data.

3. In CTA, ensure that the target Dell EMC Unity NAS Server has been added as a NAS File Server.

4. For each file system that has stubs that need to be migrated, map the source and target shares in the host.

5. Open a command window in the folder where the EMCopy tool is located. Run the following EMCopy command for each share:

   ```cmd
   emcopy.exe [source] [destination]
   ```

   **Example:**
   ```cmd
   C:\Users\Administrator\Desktop\emcopy042002>emcopy.exe U:\source Y:\dest
   ```

   The stubs from the VNX source file system will be migrated to the target Dell EMC Unity file system without re-hydrating the data. The files can be opened and recalled from the Dell EMC Unity system. Reference the documentation on Dell EMC Online Support for EMCopy to see additional flags and options available.

### MISCELLANEOUS

**DELL EMC UNITY AS A DESTINATION**

Dell EMC Unity can be used as the archiving destination for the VNX and NetApp storage systems. For more details see the *CTA and CTA/VE 12.0 SP1 Interoperability Matrix* and the *Cloud Tiering Appliance Getting Started Guide* on Dell EMC Online Support.

**TROUBLESHOOTING**

When facing any issues configuring CTA with any of the features, review the *Cloud Tiering Appliance Getting Started Guide* and search for *CTA* on the Dell EMC Online Support site to find Knowledgebase articles.
The following list provides some suggestions for configuring CTA with a Dell EMC Unity system.

- Ensure that the time between the ESXi server, the Dell EMC Unity storage system, and CTA are synchronized. It is recommended to use the same NTP Server for all the components. Also ensure that the CTA has the correct time zone configured.

- When setting the Networking for the CTA, ensure that the Speed value for the Physical Interface is set to Auto.

- Ensure that your CTA instance has been added as a host to your environment’s DNS server. Ensure that the FQDN for the CTA does not have any dashes in it.

- Ensure that your DNS server has been added as a Fully Qualified Domain Name under the Configuration > Fully Qualified Domain page on CTA.

- Ensure that the File Management Daemon and Atmos Callback Daemon services are running. The Deploying CTA with Dell EMC Unity section in the Cloud Tiering Appliance Getting Started Guide, provides more details in order to configure Dell EMC Unity on CTA.

- To verify if you can authenticate with the current configuration given for a file server use: rffm authServer file_system_name, for example, rffm authServer MIG-SMB

**CTA PORTS**

The following provides a list of the ports that are used by CTA when used with Dell EMC Unity:

- SSH Access : port 22
- To archive data from a VNX and Dell EMC Unity - Common API: TCP port 5080
- To archive CIFS data, the CTA needs SMB over NetBIOS: TCP port 139
- To create the connection for a cloud storage server : port 9000
- Amazon S3 :
  - port 443 (XML API access)
  - port 80 (Management GUI over HTTP)
- ECS withAmazon S3 API : ports 9020, 9021
- Portmap v2 RPC service: TCP port 111
- Configuring SNMP alerts : UDP port (typically SNMP uses UDP port 161 for general SNMP messages and UDP port 162 for SNMP trap messages)

**CONCLUSION**

In this paper, we discussed the various features provided by the Cloud Tiering Appliance (CTA) that apply for the Dell EMC Unity. Configuring file tiering and block archiving dramatically improves storage usage efficiency. CTA seamlessly automates policy-based file tiering and block archiving with easy management and zero impact on the user’s data. CTA allows storage administrators to create scheduled policies that move inactive file and block data off the Dell EMC Unity system to public clouds Virtustream, Microsoft Azure, Amazon S3, and IBM COS, or to a Dell EMC ECS private cloud.

**REFERENCES**

The following references can be found on Dell EMC Online Support:

- Cloud Tiering Appliance Getting Started Guide
- CTA and CTA/VE 12.0 SP1 Interoperability Matrix
- Configuring Multiprotocol File Sharing