

# EMC Disk Library Automated Tape Caching Feature

*A Detailed Review*

---

**Abstract**

This white paper details the EMC<sup>®</sup> Disk Library configuration and best practices when using the EMC Disk Library Automated Tape Caching option.

February 2010

---

---

Copyright © 2010 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on [EMC.com](http://EMC.com)

All other trademarks used herein are the property of their respective owners.

Part Number h6888

---

## Table of Contents

<b>Executive summary .....</b>	<b>4</b>
<b>Introduction .....</b>	<b>4</b>
Audience .....	4
Terminology .....	4
<b>Automated Tape Caching.....</b>	<b>5</b>
Technical overview .....	5
Key benefits .....	5
Connecting a target tape library.....	6
Usage in failover configurations.....	7
Target tape library requirements.....	7
Discovering devices .....	7
<b>Configuring Automated Tape Caching .....</b>	<b>8</b>
Creating a VTL/target library relationship .....	8
Disk Library compression considerations .....	8
Creating or changing Automated Tape Caching policies .....	9
Disabling an Automated Tape Caching policy .....	11
Create a cache for the physical tapes .....	11
Creating non-cache-enabled virtual tapes .....	13
Manual migrations to target tape .....	13
Reclaim disk space manually.....	13
Renew cache for a direct link tape.....	13
<b>Recovering data using Automated Tape Caching .....</b>	<b>13</b>
<b>Tape caching behaviors .....</b>	<b>15</b>
Queued migration jobs.....	15
Managing direct access to target tapes .....	15
Managing writes to data that is on virtual tape and target tape.....	15
Ejecting virtual/target tapes.....	15
Out-of-sync behavior.....	16
Handling error conditions during migration .....	16
<b>Conclusion .....</b>	<b>16</b>
<b>Appendix A: Using Automated Tape Caching with a Data Domain system</b>	<b>17</b>
Connection from Disk Library to Data Domain .....	17
Hardware compression.....	17
Managing storage capacity .....	17
Drive types .....	18
Reclamation .....	18
Sizing .....	18

---

## Executive summary

Automatic Tape Caching is a licensable option of the EMC® Disk Library that provides transparent access to data regardless of its location. Tape migration policies automatically trigger the copying of data from a virtual tape to a corresponding tape in a target tape library. Tape reclamation policies automatically reclaim space on the EMC Disk Library occupied by virtual tapes whose data has been copied to tape. After reclamation has occurred, data is written directly to or read directly from the tape in the target library.

## Introduction

This white paper presents an overview of the EMC Disk Library with the Automated Tape Caching option.

## Audience

This white paper is intended for EMC customers, system engineers, and members of the EMC and partners professional services community who are interested in configuration and best practices information when using the Disk Library's Automated Tape Caching option.

## Terminology

- **Automated Tape Caching** – Licensable option that allows data to be temporarily stored on the Disk Library. That data is eventually written to back-end physical tape, allowing space to be freed up on the Disk Library.
- **Backup-to-disk (B2D)** – A backup solution where data is written to hard disk instead of tape.
- **Backup image** – A group of files or a file system that has been backed up on storage media.
- **Data Protection Advisor (DPA)** – An EMC data protection management solution that provides automatic and continuous data collection, conditional analysis that triggers alerts, and a single, consistent interface for reporting.
- **Engine** — A Disk Library or Data Domain® deduplication appliance server.
- **Flex port** — Fibre Channel (FC) ports on the Disk Library server that can be configured as either front-end (SAN client) ports or back-end (physical library) ports. Flex ports do not connect to the EMC storage arrays. See also *library port*.
- **Ingest** – Process of receiving backup data from the backup server.
- **Library port** — Fibre Channel (FC) ports on the Disk Library server(s) used to connect to a back-end physical library, another Disk Library, or a Data Domain appliance. These ports are also referred to as *initiator ports*.
- **Logical Unit Number (LUN)** – Identifying number of a SCSI object that processes SCSI commands.
- **Data Migration** - The process of data movement from the front-end Disk Library system to the back-end tape library device using Advanced Tape Caching.
- **Media Server** – System that writes data to an attached backup device.
- **Native format data** – Uncompressed backup data.
- **SAN client** — A backup server that connects through a FC SAN to a Disk Library.
- **SAN client port** — FC ports on the Disk Library server used to connect backup servers (clients of the Disk Library). These ports are also referred to as *target ports*.
- **Server** — A Disk Library appliance server. Also known as an *engine*.
- **TLU** — Tape library unit, sometimes referred to as a *physical library unit (PLU)*.
- **Virtual tape library (VTL)** – Software emulation of a physical tape library system.

---

## Automated Tape Caching

The Automated Tape Caching feature provides users of the Disk Library with the ability to migrate data from disk to tape based on user-defined policies, while providing transparent access to data regardless of its location. This feature is available in version 3.1 and later of the Disk Library. Automated Tape Caching is configured on a per-VTL basis.

The following sections describe the tape caching feature requirements.

### **Technical overview**

With Automated Tape Caching, data is first written to virtual tape on the Disk Library, and later copied from there to a tape located in a tape library attached to the back end of the Disk Library. Once on the tape, the space consumed by the data in the virtual tape can be freed up, creating more available space on the Disk Library. The data is always accessible via the Disk Library and the tapes are visible to the backup application, regardless of the data location.

In an Automated Tape Caching configuration, the Disk Library acts as a *transparent cache* to the target tape library. This feature delivers greater data management simplicity because the tape in the target tape library and virtual tape are seen as the *same entity* by the backup application, regardless of whether the data is on disk (virtual tape) or on tape. Because of the various disk space reclamation policies provided, it is possible that data can reside on any of the following at any time:

- Disk (virtual tape) only - All backup application access to the data will occur with the virtual tape.
- Disk (virtual tape) and tape - All backup application access to the data will occur with the virtual tape.
- Tape only - All backup application access to the data will occur with the tape.

With this feature, the backup software manages backups to the virtual tape library while the data copy to tape is handled by the Disk Library as a background process that does not impact production servers. And, since Automated Tape Caching provides *direct access* to data on the tape, the restore process is transparent from the perspective of the backup application.

To simplify deployment, the Automated Tape Caching feature detects the target tape library that has been attached to the Disk Library and automatically matches the VTL's configuration to it. This creates a 1:1 mapping, with matching barcodes between virtual tape and the tapes in the target library. This is an effortless approach for many backup administrators who prefer automation to manually setting up the initial target tape library emulation (as required by the Advanced Disk Library Configuration). Moreover, this high level of simplicity streamlines the backup process for organizations with multiple backup servers (each with their own target tape library), since Automated Tape Caching supports the mapping of multiple target tape libraries simultaneously.

### **Key benefits**

Automated Tape Caching does the following:

- Provides direct access to data regardless of whether it is on virtual tape or tape in an attached library
- Detects the target tape library that has been attached to the Disk Library and automatically matches the Disk Library's configuration to it
- Allows for easy tape management. Backup applications only see one tape, even if the Disk Library has migrated the data to the attached tape library
- Automatically migrates data to tape based on highly flexible policies, enabling backup administrators to tailor policies according to their ILM requirements
- Encrypts data on tapes, adding security so that data cannot be accessed by unauthorized persons if tapes are lost or misplaced
- Provides flexible space reclamation policies

- Duplicates data between the Disk Library and the attached tape library without impacting the backup server
- Allows Disk Library to handle data copy to tape as a background operation for offsite vaulting and disaster protection
- Provides data restore that is straightforward; the backup application identifies the tape containing data to restore irrespective of whether data is on virtual tape or on tape in a target library

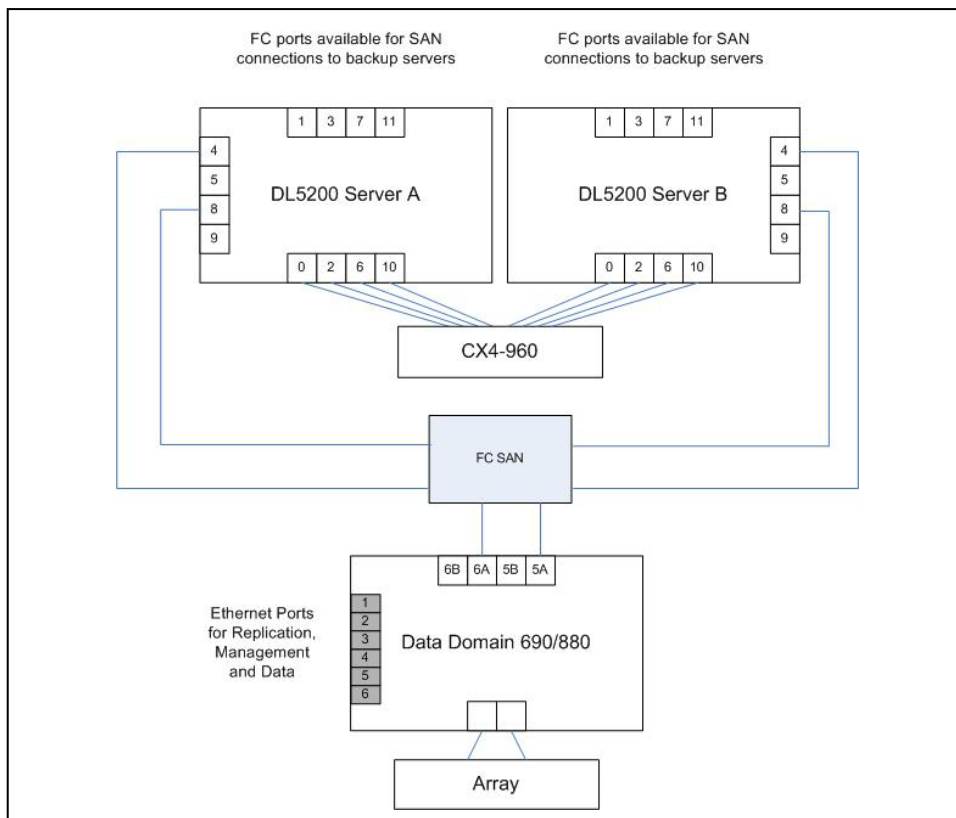
## Connecting a target tape library

The target tape library that can be attached to a Disk Library can be:

- A Data Domain deduplication system (see “Appendix A: Using Automated Tape Caching with a Data Domain system” for details)
- Another EMC Disk Library
- A physical tape library

The Disk Library has multiple Fibre Channel ports. Up to eight integrated Fibre Channel ports are front-end ports and are used to implement target-mode SCSI drivers (SCSI over Fibre Channel) that make the Disk Library look to systems on a SAN as a number of tape libraries with associated tape drives. Up to eight Fibre Channel ports are direct-attached back-end ports and are connected to the Disk Library storage.

Figure 1 illustrates the connection of a Data Domain system as the target tape library for the Automated Tape Caching feature.



**Figure 1. An example of a Data Domain VTL system's interconnections**

---

## **Usage in failover configurations**

For Disk Libraries configured for active engine failover (AEF), if an engine fails over while a data migration of data to the target tape library device is in progress, that migration will fail. Likewise if a failback occurs while a migration is in progress on the engine being failed back to, the migration will also fail. But, once a system has successfully failed over, if a migration is triggered for tapes associated with VTLs on the failed engine, the migration will execute properly.

When a migration fails, the Disk Library provides an alert both in attention required and in the event log.

There is no option to retry failed migrations automatically, but depending on the migration trigger and the reason for the failure it may be automatically retried.

For example if a time-based migration trigger policy specifies to migrate every day at 8 a.m., and the migration fails before completing, the migration should try to kick off again at 8 a.m. the next day assuming the failover has occurred successfully.

With policy-based migration such as End of Backup, the migration will not restart until the tape is mounted and unmounted or until a user makes some change to the migration trigger to initiate migration.

You can manually initiate migration to the target tape library by right-clicking on the tape that failed to migrate and selecting to migrate to the target tape library. This will restart the migration from the very beginning of the tape, not from the point where the migration failed.

## **Target tape library requirements**

Please refer to the *EMC Support Matrix* for a list of all supported tape library systems that can be used for Automated Tape Caching.

## **Discovering devices**

The Disk Library engine can have up to four fibre “initiator” ports available for attaching a target tape library through a Fibre Channel switch. If the target tape library has more than one tape drive and can have multiple fibre connections, you may connect to more than one Disk Library initiator fibre port as long as more than one is available. In general, you should use multiple connections to get the best performance. For example, for a library with four drives, the changer and two drives might be configured on one fibre connection and two other tape drives on a second connection.

Follow these steps to discover a target tape library attached to the Disk Library. Repeat steps 2 through 5 for each Disk Library HBA port connected to the target library.

1. Reboot the Disk Library from the DL Console.
2. Right-click an adapter in the **Physical Resources > Storage HBAs** tree.
3. Select **Rescan**.
4. Select **Discover New Devices**.
5. Click **OK**.

When the scan completes, click the **SCSI Devices** tab and observe the entries for the tape library devices: The library and the tape drives connected to the port should list.

If no expected tape devices are displayed:

- Verify that the target library is powered up.

- 
- Verify that the correct FC cable is connected to the tape library. You may check the port up/down status in the DL Console. Verify that the adapter in question does not have a red down-arrow icon when you expand the physical adapter's tree.
  - Because the connection to the target tape library is through a switch, the Disk Library's initiator port may not be logging in to the switch. Verify that the switch port is enabled and it is set to the proper type. (For example, the Disk Library will not log in to an E\_Port being used for interswitch communication.) Check the port up/down status in the DL Console. Verify that the adapter in question does not have a red down-arrow icon when you expand the physical adapters tree.
  - The switch may not be zoned properly. When a switch is zoned properly, the Disk Library will *see* the WWPNs of the tape library. (If the library contains native fibre drives, there will be one WWPN for the library and one for each drive.) Verify that the Disk Library sees the WWPNs by right-clicking the adapter(s) in question and selecting **Refresh SNS**. WWPNs should be listed in the **SNS Table** tab. If they do not appear, switch zoning must be corrected before continuing.

If *some* but not all of the expected tape devices display:

- Verify all cabling, including per-tape drive connections.
- Because the connection to the library is through a switch, verify that each device is connected to switch ports, that devices have logged in to the switch, and that the switch has a zone with the Disk Library port and the tape devices.
- Verify that the switch zone is set up correctly.

## Configuring Automated Tape Caching

The Automated Tape Caching feature makes use of two policies to manage virtual tape data migration and disk space reclamation. These two policies have been categorized as data migration trigger policies and data reclamation trigger policies and the policies are properties of virtual libraries, a set of tapes, or individual tapes. Trigger policies specify the events that determine “when” virtual tapes are to be migrated. Reclamation policies determine “how” disk space associated with already migrated data is reclaimed.

Migration will be supported only on virtual tapes that use barcodes mapped to target tapes. A virtual library that has its tape cache property enabled can have virtual tapes that do not use barcodes from target tapes. Tape caching operations will not be applicable to such virtual tapes.

In order to use Automated Tape Caching, you will need to do the following:

- Create the Automated Tape Caching policies for the virtual tape library
- Synchronize the tapes in the libraries

### ***Creating a VTL/target library relationship***

If the Automated Tape Caching license is installed and a target tape library is configured on the Disk Library, when creating a new VTL through the Console, the **Equivalent Library Creation** window appears with a list of assigned target tape libraries. Selecting a library creates an Automated Tape Caching-enabled VTL with the same library type, drive type, and serial number as the selected target tape library. EMC recommends creating a virtual library type and drive type that are identical to that of the target library connected to the system.

### ***Disk Library compression considerations***

Typically, tape library compression algorithms vary between virtual and target tape library devices and may affect the Disk Library's ability to reliably migrate data from the virtual tape library to the target tape library device using Automated Tape Caching. If Disk Library compression is absolutely necessary within any given environment, it is suggested that the tapes created within the Disk Library be 20 percent less in capacity than that of their corresponding tapes on the target tape library. (This can be achieved by using the Advanced Tape Creation process when creating virtual tapes.) This should ensure that when a data

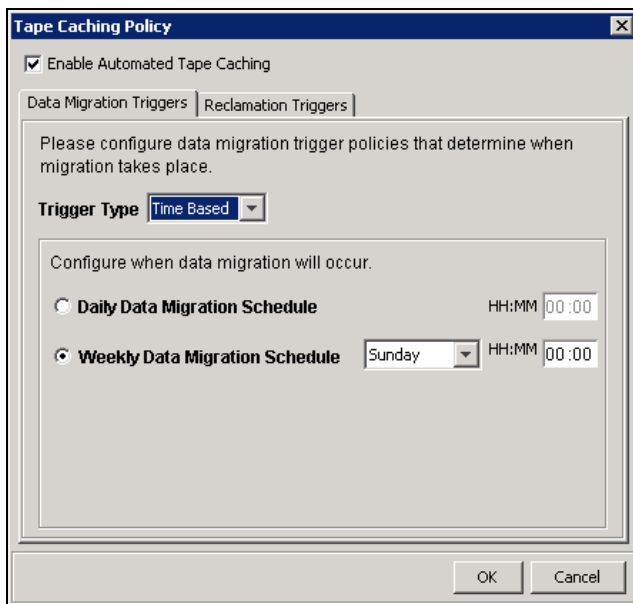
---

migration takes place, there is enough capacity on the target tape to accommodate the data being migrated from the virtual tape.

## ***Creating or changing Automated Tape Caching policies***

To create or change an Automated Tape Caching policy:

1. Right-click on a virtual tape library and click **Automated Tape Caching**.
2. Select the **Enable Automated Tape Caching** checkbox.
3. On the **Data Migration Triggers** tab, select the type of data migration trigger in the Trigger Type list box. Data migration triggers control when data in the virtual tape will be copied to target tape.

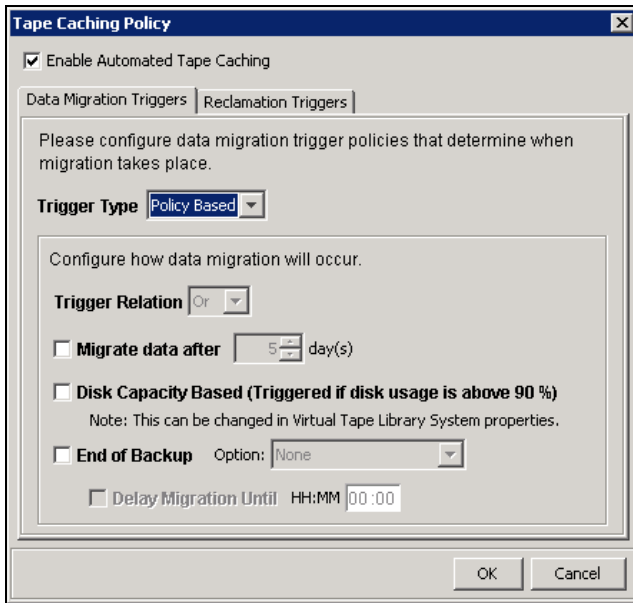


**Figure 2. Time Based data migration trigger selection**

For Time Based triggers, specify when data migration should actually occur with one of these two radio buttons:

- **Daily Data Migration Schedule** - Migration occurs at a specific time of day. Type the hour and minute (in 24-hour format) in the box. For example, if you want the migration to occur at 11:30 p.m., enter 23:30. Note that if the specified time has already elapsed when the policy is configured, the migration will occur when the trigger occurs at that time on the next day.
- **Weekly Data Migration Check Schedule** - Migration occurs on a specific day of the week. Specify the day of the week from the list and type the hour and minute (in 24-hour format) in the text box. Note that if the specified time has already elapsed when the migration is configured, the migration will occur at the next scheduled day and time.

For Policy Based triggers, select the criteria that will trigger data migration to the target tape.



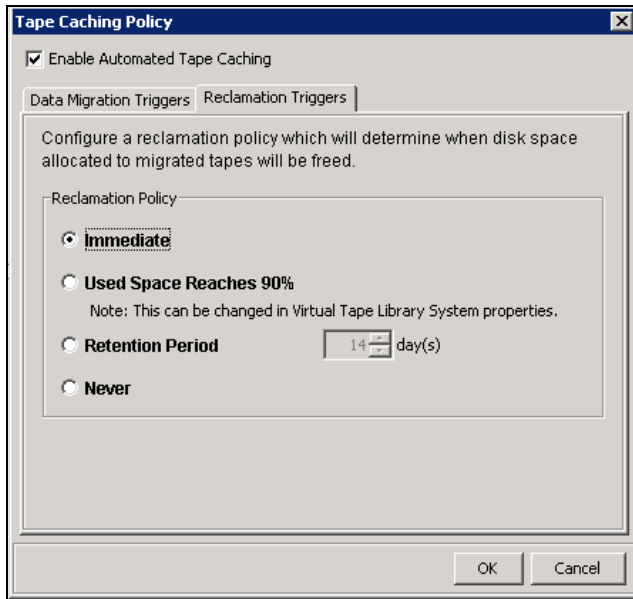
**Figure 3. Policy Based data migration trigger selection**

- Select **And** in the **Trigger Relation** field if all the selected criteria must be met to initiate the data migration, or select **Or** if meeting any one of them will initiate the data migration.  
For example, if you select **And** and both **Age Based** and **Disk Capacity Based**, data migration will occur only when both the specified number of days has elapsed and the specified disk capacity has been reached. If you select **Or** for **Trigger Relation**, the occurrence of either one of those events will cause data to migrate to the target tape.
- **Age Based** - Migration will occur when the data has been on the virtual tape for the specified number of days. Enter the desired number of days in the list box.
- **Disk Capacity Based** - Migration will occur when the used disk space exceeds the specified disk capacity.  
To specify the amount of used space that will trigger data migration, right-click **Virtual Tape Library System** in the tree, click **Properties**, and type the desired percentage in the **Tape Caching Policy Disk Capacity Threshold** box.

---

Note: The Tape Caching Policy Disk Capacity Threshold setting affects other capacity-based actions as well.

- **End of Backup** - Migration will occur when a backup has completed and the virtual tape has been moved out of the virtual drive. If you select **Only When Tape is Full**, migration will only occur if the tape is full.
  - **Delay Migration Until** - Migration will be delayed until the time you specify after one of the above policy events have occurred. You may want to select a time when system usage is very light. Enter the hour and minute (in 24-hour format) in the box.
4. Click the **Reclamation Triggers** tab and specify when the data that has been migrated to target tape can be deleted to free up disk space. The following radio buttons appear:



**Figure 4. Reclamation trigger policy selection**

- **Immediate** - Cache disk space is freed up as soon as the data migration is complete.
- **Used Space Reaches 90%** - Cache disk space is freed up when used space equals 90% of capacity. (As noted above: This can be changed in the Virtual Tape Library "Properties" section as well.)
- **Retention Period** - Disk space is freed up after a specified number of days have elapsed. Specify the number of days that the data should be retained in the adjacent list.
- **Never** - Disk space is never reclaimed

5. Click **OK**.

The policy takes effect immediately.

---

Note: When you move a tape from the virtual tape library to a vault, it retains the Automated Tape Caching policy associated with the original virtual tape library.

---

## ***Disabling an Automated Tape Caching policy***

To disable a Automated Tape Caching policy:

1. Right-click on a virtual tape library and click **Automated Tape Caching**.
2. Clear the **Enable Tape Caching Policy** checkbox.

All the options that you previously set are retained, but the Automated Tape Caching policies will not execute until you select this check box again.

3. Click **OK**.

## ***Create a cache for the physical tapes***

Tape caching-enabled tapes are identified by the letter **A** located in the upper left portion of the virtual tape icon. Tape caching-enabled tapes are created by synchronizing a VTL with a target library assigned to the back end of the Disk Library.

---

When *non-cache* enabled tapes have been previously created, you cannot make this a cached tape in the future.

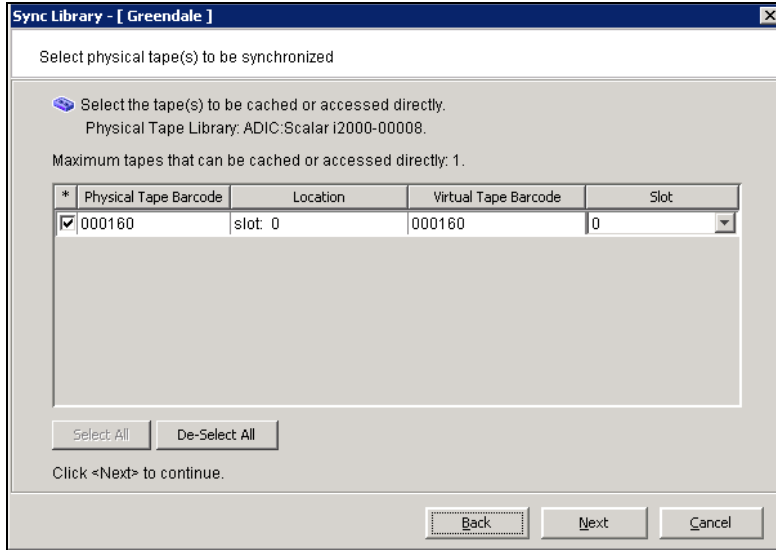
---

---

With Automated Tape Caching, the Disk Library stores data on disk before it migrates it to target tape. To migrate data to target tape, you must create a cache for each of your target tapes.

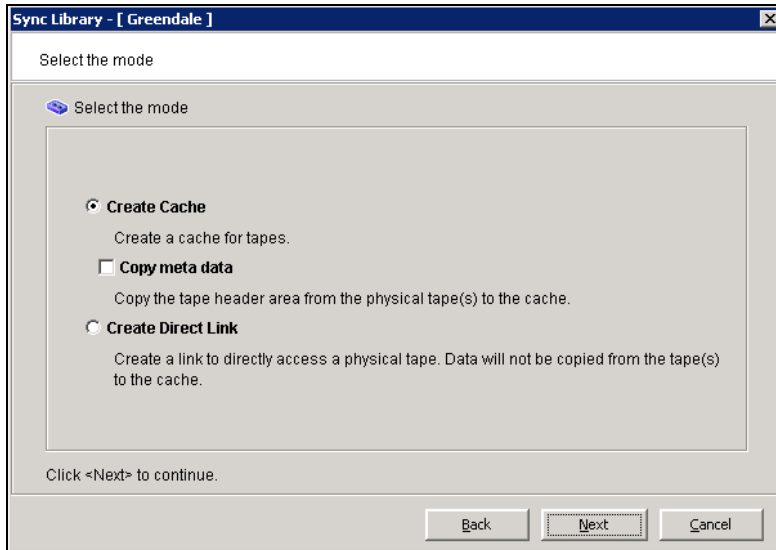
Do the following to synchronize the tapes in the libraries and create cache for the target tapes:

1. Right-click on the virtual tape library to synchronize with the target tape library and select **Sync Library**.
2. If you have multiple libraries, select the appropriate target library.
3. Select the target tape(s) for which you want to create a cache.



**Figure 5. Selecting tapes for synchronization**

4. Select **Create Cache**.



**Figure 6. Selecting the caching mode**

---

**Copy meta data** - Copies the tape header from the target tape to the cache. Select this option if your backup application requires a tape header to identify a tape.

5. Specify which target device should be used to create the cache.
6. Specify a prefix, size, and starting number for the cache.
7. Confirm all information and click **Finish**.

## ***Creating non-cache-enabled virtual tapes***

Even when using Automated Tape Caching, you can still create *non-cache-enabled* virtual tapes in a VTL that will not migrate to target tapes. This can be useful for a single backup that is not part of your normal backup routine. You can create one or more virtual tapes by right-clicking on a virtual tape library or on the **Tapes** object and selecting **New Tape(s)**.

---

Note: If you create virtual tapes in a virtual library with Automated Tape Caching enabled, they cannot match the barcodes of your physical tapes.

---

## ***Manual migrations to target tape***

You can manually cause data in a cache to migrate to target tape. To do this, right-click on a tape cache-enabled virtual tape and select **Migrate to Physical Tape**. Note that this will overwrite all data on the target tape.

## ***Reclaim disk space manually***

You can manually cause the data that has been migrated to target tape to be deleted to free up disk space. To do this for a single tape, right-click on a tape cache-enabled virtual tape and select **Reclaim Disk Space**. Note that this will delete all data in the cache for that tape.

To do this for multiple tapes, right-click on the **Virtual Tape Library System** object and select **Reclaim Disk Space**.

## ***Renew cache for a direct link tape***

When the data from a migrated tape is reclaimed and that space freed on the Disk Library, the virtual tape will become a *direct link* tape. A direct link tape is not an actual tape but a link to a target tape. If your backup application ever overwrites the header, the Disk Library will automatically create a cache for the target tape. This eliminates the direct link and creates a cache for the target tape.

You can also manually renew the cache for a direct link tape. To do this, right-click on a direct link tape and select **Renew Cache**.

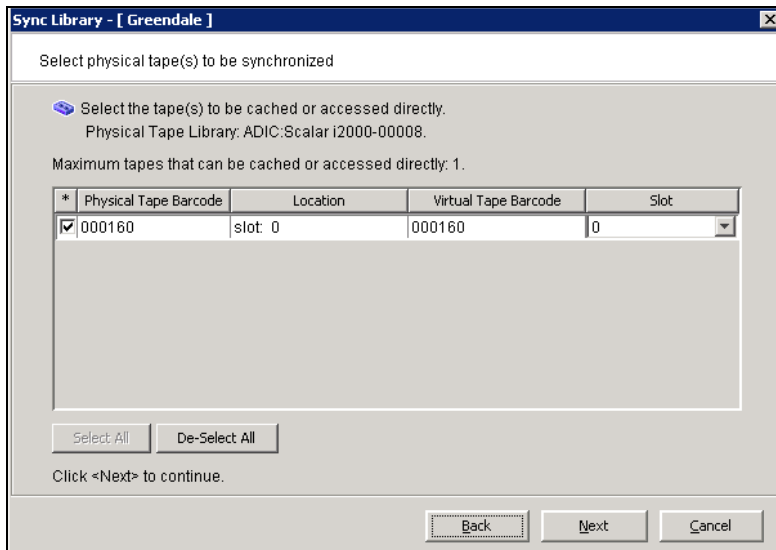
## **Recovering data using Automated Tape Caching**

With Automated Tape Caching, tapes are always visible to the backup application regardless of whether the data is actually on disk or on a target tape in the target tape library. When it comes time to restore data, your backup application will seamlessly read the data from disk (if it is still there) or from the target tape.

Once data has been migrated from virtual tape to target tape and the tape ejected out of the target library, it needs be imported back into the target library before it can be exposed to the backup application through the virtual libraries. This requires you to create a link to the target tape when it is returned to the library. To do this:

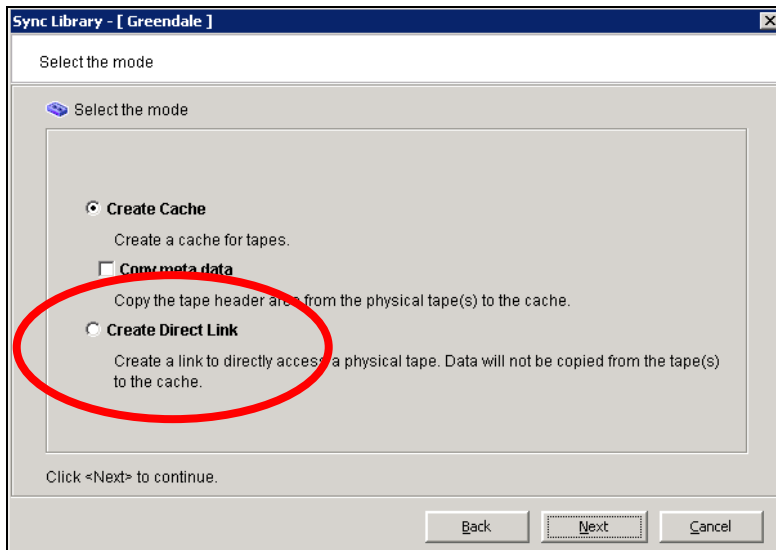
1. Right-click on your virtual tape library and select **Sync Library**.

2. If you have multiple libraries, select the appropriate target library.
3. Select the target tape from which you need to restore data.



**Figure 7. Selecting physical tapes for restore**

4. Select **Create Direct Link**.



**Figure 8. Creating a direct link to a physical tape for restore operations**

If the data was encrypted before being migrated, select the appropriate key to decrypt the data.

---

## Tape caching behaviors

This section discusses expected behavior for a number of commonly performed functions.

### ***Queued migration jobs***

All migration operations triggered by a policy that cannot be run immediately because of resource (target drives, tapes) availability will be placed in the job queue. Queued jobs can be managed the same way import/export jobs are managed.

Queued migration jobs will be cancelled automatically if any event that invalidates the pending operation takes place (such as overwriting/appending to virtual tape).

If a migration job is submitted for a tape that already had a migration job pending, the first job will be cancelled automatically since there is no need to run the migration job twice.

### ***Managing direct access to target tapes***

Since target tape drives are required in order to provide access to data on target tapes, it is recommended that the number of virtual tape drives match the number of target tape drives attached to the Disk Library.

If the number of virtual tape drives is larger than the available target tape drives, the following behavior should be expected.

- If a free target tape drive is available, it will be used.
- If no target tape drive is available, mount operations will wait for a certain period of (back-end-configurable) time, before failing the mount request. This mechanism should help with full library inventory conditions where a backup application is trying to read all tapes using all virtual tape drives. Failing mount requests can cause significant administrative issues for backup applications.

It should also be noted that a no-free tape drive condition could arise under the following conditions.

- Number of Virtual drives = Number of target drives and target drives are being used for migration
- Number of Virtual drives > Number of target drives and trying to load all target tapes at the same time

### ***Managing writes to data that is on virtual tape and target tape***

Any writes to a virtual tape (overwrite or append) will invalidate data on an already existing target tape. If a tape caching policy is enabled on the tape, data migration will be triggered again based on the migration policies specified for that tape.

If a migration operation is in progress when an overwrite operation takes place, the migration operation will be cancelled.

### ***Ejecting virtual/target tapes***

Tape ejection is handled as follows:

- **Ejecting a virtual tape on disk only (no physical copy):** Virtual tape is moved to vault.
- **Ejecting a virtual tape that has data that has been migrated:** Virtual tape is moved out of the library to vault. Target tape is left in the target library.
- **Ejecting a virtual tape that has data that has been reclaimed (stubbed):** The virtual tape stub is deleted and the corresponding target tape is moved to the IE slot in the target library.
- **Ejecting a target tape:** The tape moves to the IE slot of the target library, if an IE slot is available. If no IE slot is available, an eject job will be created. The job will remain in the queue until a free IE slot becomes available. The ejected tape will no longer be visible to the backup application.

---

## ***Out-of-sync behavior***

Out-of-sync conditions are handled as follows.

- **If the data is on disk and tape:** If the data has already migrated to target tape, reclamation will be allowed. If the data has not been migrated, reclamation will not be allowed.
- **If the data is on target tape only:** A mount operation for the tape will fail, which will cause the slot where the tape was located to be marked as empty by the Disk Library.

## ***Handling error conditions during migration***

If an error condition occurs during data migration operations, data migration will be considered “incomplete.” All “incomplete” data migration jobs will be left in the queue with a “failed” status. Users will have the option to restart failed operations after the necessary steps have been taken to ensure that a retry operation would succeed. Please note that unless a user restarts a “failed” job manually, no action will be taken by the software to retry the job. Users will have the option to delete “failed” jobs if they do not wish to restart them.

## **Conclusion**

The Disk Library’s Automated Tape Caching feature makes it easy for backup administrators to tailor Disk Library *data migration policies* according to their information lifecycle management (ILM) requirements. Flexible policies (as outlined in this paper) automatically trigger data export to physical tapes based on a variety of thresholds, such as the length of time data has been on virtual tape or when disk space available for virtual tapes is running low. Automated Tape Caching also provides very flexible *space reclamation policies* such as free space immediately upon migration, after a specified retention period, or when running out of space to ensure sufficient data storage is available on the Disk Library.

---

## Appendix A: Using Automated Tape Caching with a Data Domain system

Automated Tape Caching-enabled virtual tape libraries will migrate data to the Data Domain system for deduplication *and* subsequent replication to a target Data Domain system. The Data Domain system will appear as a physical library to the Disk Library and will present its library and tape drives to the Disk Library by assigning them via the DL console.

When using Automated Tape Caching when a Data Domain system is the target tape library for the Disk Library, it is important to adhere to these best practices in this interoperability configuration.

### ***Connection from Disk Library to Data Domain***

To connect the EMC Disk Library with the Data Domain system, initiator ports on the Disk Library need to connect to the target ports on the Data Domain system through a Fibre Channel SAN. In this configuration, EMC recommends the following:

- Connect two initiator ports on the Disk Library to two target ports on the Data Domain system to ensure best performance and availability.
- On the Data Domain system, use two Fibre Channel HBAs and split the connections between the two.
- Evenly distribute the Data Domain virtual tape drives between the two connections.

For example, port 4 (initiator port) on the Disk Library is zoned to port 5A (target port) on a Data Domain DD880 system and port 8 (initiator port) on the Disk Library is zoned to port 6A (target port) on the Data Domain DD880 system. A virtual library is created on the Data Domain system that contains 10 tape drives. Five of these drives are assigned to port 4 on the Disk Library and the other five drives to port 8.

---

Note: If a DD690 or DD880 system is configured with two Fibre Channel HBAs (four Fibre Channel ports) then the system will have two 1 GbE network ports for replication and management. There will not be a slot available to add in the 10 Gb Ethernet option.

---

### ***Hardware compression***

EMC recommends the following:

- Do not enable Disk Library hardware compression. When enabled, the data is uncompressed before it is migrated to the Data Domain system. There is a chance that the uncompressed data will be larger than the 800 GB Data Domain tape, which will cause the migration to fail.
- If hardware compression is disabled, create virtual tapes on the Data Domain appliance 5 percent larger than the intended size on the Disk Library.
- If the customer plans on using hardware compression, create virtual tapes on the Data Domain system that are larger than the tape size on the Disk Library by the expected/measured compression ratio for that data type.

### ***Managing storage capacity***

When a backup application expires a tape, there is no direct communication of the event to the Disk Library. As a result, the Disk Library and the Data Domain appliance still have the tape data and therefore the data still occupies storage space. For the Disk Library to reuse the tape, the backup application must also relabel the volume by writing a data block to its beginning. Expiring and relabeling Disk Library Automated Tape Caching-enabled tapes will allow the Disk Library to remove the tape stub and recover storage capacity in the Disk Library. The deduplicated data space associated with those tapes is recovered when new data written to the relabeled tape is migrated to the Data Domain appliance.

---

The Disk Library will not allow a relabeled tape to be migrated to the Data Domain appliance. New backup data must be written to the tape first before it can be migrated (either manually or with a policy).

---

When new data written to that tape is migrated to the Data Domain appliance, the previous deduplicated data associated with the expired version is marked as invalid and the space it occupies is eligible for recovery through the Data Domain appliance's cleaning process.

## ***Drive types***

When configuring virtual tape drives on the Data Domain and on the Disk Library, EMC recommends using the same drive types configured on both systems. Using the same drive types assures that size capacity and drive uniqueness is intact.

## ***Reclamation***

Perform reclamation only on full tapes in the Disk Library. Backup and restore performance is much better for data that has not yet been reclaimed and still resides on the Disk Library when compared to its deduplicated form on the Data Domain system. Therefore, keep data in native format as long as it is actively being used for backups (including appends), or is needed for quick restores. By performing reclamation only on full tapes, it assures that data accesses occur only with the tape on the Disk Library until it is reclaimed.

## ***Sizing***

In order to properly size the Disk Library when configured with the Data Domain system, retention requirements must be thoroughly reviewed and changed to accommodate the longer retention times possible with deduplication technology. Each system must be sized separately according to the retention scheme desired and anticipated data access needs to take full advantage of the features of each system. The sizing exercise may require backup policies to be re-evaluated. Storage capacity must be sized to adequately handle the amount of data expected to be retained in both native and deduplicated format.

---

Please contact your EMC representatives to properly size the environment in which this interoperability will be used.

---