Abstract

This white paper outlines the configuration and best practices for deploying Dell EMC™ ECS with Commvault® software to provide a flexible and scalable backup solution.

March 2019
Revisions

<table>
<thead>
<tr>
<th>Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>March 2019</td>
<td>Initial release</td>
</tr>
</tbody>
</table>

Acknowledgements

This paper was produced by members of the Dell EMC Unstructured Technical Marketing Engineering and Solution Architects team.

Please send comments, suggestions, or feedback to unstructured.tme.sa@emc.com.

Copyright © 2019 Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, Dell EMC and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners. [3/21/2019] [Configuration and Deployment] [H17576]
Table of contents

Revisions ........................................................................................................................................... 2
Acknowledgements ............................................................................................................................ 2
Table of contents ............................................................................................................................... 3
Executive summary ............................................................................................................................. 5
Audience ............................................................................................................................................. 5

1 Introduction ................................................................................................................................. 6
   1.1 Backup challenges with traditional storage .............................................................................. 6
       1.1.1 Compounding effect of backups ......................................................................................... 6
       1.1.2 Management overhead ...................................................................................................... 6
       1.1.3 Refresh cycle and data migration ......................................................................................... 6
       1.2 Backup solutions with ECS and Commvault ........................................................................... 6
       1.3 Terminology ........................................................................................................................... 7
       1.4 Solution components ............................................................................................................ 7

2 Reference architecture .................................................................................................................. 8
   2.1 Architectural overview .............................................................................................................. 8
       2.1.1 CommServe ......................................................................................................................... 8
       2.1.2 MediaAgent ........................................................................................................................ 9
       2.1.3 Client .................................................................................................................................. 9
       2.1.4 Agent .................................................................................................................................. 9
       2.1.5 CommCell Console ............................................................................................................ 9
       2.1.6 Storage policy .................................................................................................................... 9
       2.1.7 Subclient ............................................................................................................................. 9
       2.1.8 Cloud storage library ........................................................................................................ 9
       2.1.9 Load balancer ..................................................................................................................... 9

3 Backup/archive target configuration ............................................................................................... 11
   3.1 Cloud storage library connection ............................................................................................ 12
       3.1.1 Data access ports ............................................................................................................... 12
       3.1.2 ECS configuration .............................................................................................................. 12
       3.1.3 Amazon S3 ....................................................................................................................... 12
       3.1.4 EMC Atmos ...................................................................................................................... 14
       3.1.5 OpenStack ....................................................................................................................... 17
       3.1.6 Centera CAS ..................................................................................................................... 18
   3.2 Storage policy .......................................................................................................................... 20
       3.2.1 Stream randomization ...................................................................................................... 20
Table of contents

3.2.2 Alternate data paths for a shared library configuration ................................................................. 20
3.3 Subclient ........................................................................................................................................... 20
3.4 Configuration and tuning .................................................................................................................. 21
3.4.1 Cloud library recommendations .................................................................................................... 21
3.4.2 MediaAgent properties .................................................................................................................. 21
3.4.3 Subclient properties ...................................................................................................................... 24
3.4.4 Storage policy properties .............................................................................................................. 24
4 Advanced configuration options ......................................................................................................... 26
4.1 Commvault configuration best practices ......................................................................................... 26
4.2 Retention ........................................................................................................................................ 26
4.2.1 ECS bucket configuration ............................................................................................................ 26
4.2.2 Storage policy .............................................................................................................................. 26
4.2.3 Cloud storage library retention setting .......................................................................................... 27
4.2.4 MediaAgent retention property .................................................................................................... 27
4.3 Multi-stream restore operations ....................................................................................................... 28
5 Best practices and general information .............................................................................................. 29
5.1 Load balancing and connectivity ..................................................................................................... 29
5.2 Deduplication .................................................................................................................................. 29
5.3 Data pruning .................................................................................................................................... 29
5.4 Compression .................................................................................................................................... 30
5.5 Multi-site replication .......................................................................................................................... 30
6 Troubleshooting and known issues ...................................................................................................... 31
6.1 Data retention forecast and compliance report .................................................................................. 31
A Technical support and resources ......................................................................................................... 32
A.1 Related resources .............................................................................................................................. 32
Executive summary

This paper describes the best practices and solution-specific configuration steps for deploying the Commvault® software platform with Dell EMC™ ECS for both basic and advanced deployments. The combination of ECS and Commvault data and information management software provides a modern solution for data protection and preservation that provides a simple, flexible, cost-effective, scalable solution that lets enterprises address their needs from data protection to preservation while providing secure self-service access to data. ECS and Commvault provide installation flexibility with support for popular object interfaces including S3 and Swift as well as legacy Centera™ CAS and Atmos™ REST.

Audience

This document is intended for administrators who deploy and configure ECS with Commvault. This paper assumes a high level of technical knowledge for the devices and technologies described.
1 Introduction

Protecting and preserving data is a huge challenge for all enterprises because of the unprecedented and unrelenting growth of file-based unstructured data. This explosive growth is making it harder to meet backup SLAs and greatly increases the importance of having an effective strategy to move static data out of the backup stream and on to fast and cost-effective archive storage.

1.1 Backup challenges with traditional storage

Traditional storage solutions can present backup challenges such as the compounding effect of backups, the management overhead required, technology refresh cycles, and data migration.

1.1.1 Compounding effect of backups

Traditionally, many full backup copies are kept weekly, monthly, and yearly which can cause data growth to be compounded exponentially. For example, if the most recent weekly backups are kept for 8 weeks, monthly backups for 10 months (covering the remainder of the year), and yearly backups for 7 years, then one file would have 25 copies just using the backup algorithm. Using this example, just 40 TB of new data would need 1 PB of storage. To reduce some of the common data, there are mechanisms like compression, deduplication, and snapshots, but portions of the data will be unique or not easily reduced.

This compounding growth affects the cost of tape even more, since many of these data-reduction mechanisms are not available, and native tape drive compression is only supported by tape. In addition, there are many often-overlooked issues that need to be considered when using tape — the cost of secure off-site storage, the cost to periodically retrieve backups from storage to perform test restores as needed for compliance or business policy, and the risk of not being able to read older tapes if too many tape drive generations have passed.

1.1.2 Management overhead

Individually monitoring performance and free space on traditional RAID-based volumes or LUNs can become a huge burden. The constant juggling and adding of new volumes or LUNs when capacity or hardware limits are reached consumes more and more time. Each storage change often requires the application’s configuration to be updated as well.

1.1.3 Refresh cycle and data migration

The inevitable, and often overlooked, hardware refresh every 3–5 years will more than likely take up many nights and weekends, many data center resources, and much budget with traditional RAID-based storage systems. A technology refresh requires extensive planning and a complete re-evaluation of the performance and capacity requirements for 3 or more years.

1.2 Backup solutions with ECS and Commvault

The ECS object storage model removes the hurdles of multiple backup copies, management overhead, application reconfiguration, and data migration, supporting a more focused backup and archive strategy.

ECS is a multi-purpose, software-defined, cloud-scale object storage platform that delivers the ease-of-use and economics of the public cloud without the data sovereignty and security concerns. ECS features native multi-tenancy to deliver backup and archiving-as-a-service that can support multiple tenants and applications
Audience

on a single, easy-to-manage cloud-storage platform. ECS brings the cost profile, simplicity, and scale of public cloud services to everyone with the trust, reliability, and support expected from Dell EMC.

Commvault software is a comprehensive information protection and preservation platform that offers advanced features like OnePass™ which converges backup, archive, and reporting operations into a single job, maximizing efficiency while minimizing impact on production. Accessible through a single console, automated policies manage data in the ContentStore — the efficient, virtual repository that provides a flexible and efficient way to store, access, and understand data. Content-based retention through the Reference Copy feature takes advantage of ContentStore intelligence to store only data valuable to the business.

Together, ECS and Commvault provide a comprehensive cloud solution for any enterprise with demanding data protection and preservation needs. Commvault provides a proven software solution with intelligent policy-based data management, monitoring, and reporting and ECS provides geo-distributed cloud-scale object storage disk-based archive system with world-wide trusted support.

This document describes best practices and solution-specific configuration steps for deploying Commvault with Dell EMC ECS for both basic and advanced deployments. Basic best practices enable quick, easy, and straightforward deployments using the fewest settings to get started. The advanced best practices identify opportunities to configure the system for performance, scalability, or highly secure environments for a more optimized deployment model. This optimized deployment model requires a high level of knowledge, support, and time to plan the deployment of all the components in advance.

1.3 Terminology
The abbreviations used in this document are summarized in Table 1.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>CommServe server</td>
</tr>
<tr>
<td>MA</td>
<td>MediaAgent server</td>
</tr>
<tr>
<td>MP</td>
<td>Mount path (Cloud Storage Library)</td>
</tr>
<tr>
<td>DB</td>
<td>Database</td>
</tr>
<tr>
<td>DDB</td>
<td>Deduplication database</td>
</tr>
<tr>
<td>SQL</td>
<td>Microsoft® SQL Server</td>
</tr>
<tr>
<td>WORM</td>
<td>Write once read many</td>
</tr>
</tbody>
</table>

1.4 Solution components
The solution components described in this document include Commvault Complete Backup & Recovery 11 and Dell EMC ECS 3.2.2.
2 Reference architecture

Understanding the Commvault architecture first requires discussing the various components and their functions. The iDataAgent software is deployed on all servers, workstations, and laptops where backups are needed. Virtual Server Agents (collectively called clients) are used to support hypervisor protection. There are specific iDataAgents for each of the different file systems, applications, and databases supported by Commvault. The MediaAgent software is deployed on dedicated servers, maintains the deduplication databases, and manages the transfer of data between the clients and ECS. Multiple MediaAgents are used to distribute the load. ECS is set up as a Cloud storage library and is a backup target. The CommServe software is deployed on a dedicated server that communicates and coordinates all operations, backups, restores, copies, media management, and other functions. The CommServe creates a logical grouping of MediaAgents and clients called a CommCell. Depending on the size of the environment, or the network and physical boundaries, there can be multiple CommCells within an organization.

2.1 Architectural overview

A CommCell management group is the logical grouping of all software components described previously. This group protects, moves, stores, and manages the movement of data and information.

Figure 1 shows the logical architecture for a CommCell group with Dell EMC ECS storage defined as its CommCell Storage Library. With this configuration, the Storage Library settings never need to be updated as capacity is added. When capacity is added to the ECS environment, the CommServe and MediaAgents immediately see the additional capacity without additional configuration or manual intervention.

![Logical architecture overview]

2.1.1 CommServe

The CommServe host is the central management component of the CommCell group. It coordinates and executes all CommCell group operations, maintaining Microsoft SQL Server databases that contain all configuration, security, and operational history for the CommCell group. There can be only one CommServe
host in a CommCell group. The CommServe software can be installed in physical, virtual, and clustered environments but only on a Microsoft Windows® host. Commvault supports high availability by replicating data to a standby CommServe host.

2.1.2 MediaAgent
The MediaAgent is the data-transmission manager in the CommCell group. It provides high-performance data movement and manages the storage libraries. The CommServe server coordinates MediaAgent tasks. For scalability, there can be more than one MediaAgent in a CommCell. The MediaAgent software can be installed in physical, virtual, and clustered environments. A MediaAgent can be installed on a Windows or Unix host.

2.1.3 Client
A client is a logical grouping of the software agents that facilitate the protection, management, and movement of data associated with the client.

2.1.4 Agent
An agent is a software module that is installed on a client computer to protect a specific type of data. Different agent software is available to manage different data types on a client, such as Windows file-system data or Oracle® databases. Agent software can be installed in physical, virtual, and clustered environments, and may be installed either on the computer or on a proxy server. In addition, a Virtual Server Agent may be used for hypervisor support.

2.1.5 CommCell Console
The CommCell Console is the central management user interface for managing the CommCell group — monitoring and controlling active jobs, and viewing events related to all activities. The CommCell Console allows centralized and decentralized organizations to manage all data-movement activities through a single, common interface.

2.1.6 Storage policy
Storage policies act as a channel for backup and restore operations. They map data from its original location to physical media. A policy can be used for either data protection and archiving or disaster recovery. Retention and reduplication properties can also be defined here.

2.1.7 Subclient
Subclients are logical containers that define the specific production data (virtual machines, drives, folders, files, databases, mailboxes) to be protected on a client.

2.1.8 Cloud storage library
The Cloud storage library enables configuring and using online cloud storage devices such as Dell EMC ECS as a storage targets.

2.1.9 Load balancer
The Dell EMC Atmos, Amazon S3, and OpenStack Cloud Storage Libraries all expect a single ECS hostname or IP address. They do not support DNS round robin or multiple addresses. The best practice for distribution
of REST traffic among the ECS nodes requires an external IP load balancer. The ECS Best Practices Guide describes several approaches when configuring a load balancer with ECS.

**Note:** The ECS CAS protocol supports internal load balancing, so the Centera Access Node address should resolve directly to a single ECS node. No external IP load balancer should be used.
Backup/archive target configuration

Figure 2 shows an overview of the configuration steps for Commvault and ECS deployed as a backup target. The subsections that follow provide best practices or required settings for each step.

**Create the Cloud Storage Library**
- S3 | Swift | Atmos | CAS

**Create the storage policy**
- Configure for backup/archive
- Select the Cloud Storage Library
- Specify # streams / configure retention
- [Options] Enable deduplication
- Configure data path options on primary copy

**Create a Subclient**
- Choose data path(s) to back up/archive
- Select the storage policy
- Select SW compression options
- Select deduplication options
- Configure advanced options
- Choose scheduling options

**Apply recommended settings**
- MediaAgent properties
- Subclient properties
- Storage Library properties

Figure 2  Configuration workflow
3.1 Cloud storage library connection
Commvault supports direct connection to ECS using the Atmos, S3, Swift, or Centera CAS rest APIs. Security for each of these APIs must first be configured for the appropriate ECS user before they can be used within Commvault. The following sections review best practices for configuring a single Storage Library using each of the supported APIs to be shared with the MediaAgents, the CommCell policy configurations, and the necessary alerts for virtual deployments.

3.1.1 Data access ports
The ECS Object Services communicate on different ports depending on the protocol and data access API.

<table>
<thead>
<tr>
<th>Data access method</th>
<th>HTTP</th>
<th>HTTPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3</td>
<td>9020</td>
<td>9021</td>
</tr>
<tr>
<td>Atmos</td>
<td>9022</td>
<td>9023</td>
</tr>
<tr>
<td>Swift</td>
<td>9024</td>
<td>9025</td>
</tr>
</tbody>
</table>

3.1.2 ECS configuration
ECS requires a storage pool, virtual data center, replication group, namespace, and an object user with an S3 and Swift password, an Atmos sub-tenant and CAS password, bucket, and pea file depending on the API being used.

3.1.3 Amazon S3
For new installations of Commvault with ECS, S3 is the recommended API for data access. The Commvault implementation of Amazon S3 uses path-based addressing, so no DNS or ECS object-based URL modifications need to be made. MediaAgents can be configured to connect to ECS through its S3 API using the following method.

**Note:** A new cloud storage connector has been introduced in SP14 named S3 Compatible Storage. This new connector eliminates many of the fields that are irrelevant to ECS when using the Amazon S3 connector.

1. Expand **Storage Resources** and right-click **Libraries**. Choose **Add > Cloud Storage Library**...
2. If using SP13 or earlier, select **Amazon S3** as the Cloud Storage type; otherwise select **S3 Compatible Storage** if using SP14 and above. Enter the appropriate **Access Information** shown in Figure 3 and Figure 4 using information from the ECS user account in the web portal.

   - For the S3 cloud storage connector using SP13 or earlier, enter the following:

     **MediaAgent**: Choose the MediaAgent that will drive data to ECS.

     **Service Host**: ECS supports both HTTP (port 9020) and HTTPS (port 9021) for the S3 protocol. Enter the IP Address or DNS name of the load balancer that has been placed in front of the ECS nodes including the port. In the above example, a load balancer has been specified using the mapped port 80 for client/application-facing connections.

     **Access Key ID**: Enter the ECS object user name.

     **Secret Access Key**: Enter the ECS object user secret access key.

     **Bucket**: Enter the name of the ECS S3 bucket to use. If the bucket does not exist, it will be created automatically.

     **Storage Class**: Keep the default value for this field.

---

Figure 4  S3 cloud storage connector using SP13 or earlier
For the S3 cloud storage connector using SP14 or greater, enter the following:

**MediaAgent:** Choose the MediaAgent that will drive data to ECS.

**Service Host:** ECS supports both HTTP (port 9020) and HTTPS (port 9021) for the S3 protocol. Enter the IP Address or DNS name of the load balancer that has been placed in front of the ECS nodes including the port. In the above example, a load balancer has been specified using the mapped port 80 for client/application-facing connections.

**Access Key ID:** Enter the ECS object user name.

**Secret Access Key:** Enter the ECS object user secret access key.

**Bucket:** Use one of the following options to select the ECS bucket to use for backup/archive.

> Click the **Detect** button to retrieve the buckets owned by the ECS object user and use the drop-down menu to select the bucket to use.

> Enter the bucket name in the field. If it does not exist, Commvault will create it.

---

3.1.4 **EMC Atmos**

The EMC Atmos storage library uses the Atmos Namespace API to write objects to ECS. Configure MediaAgents to connect to ECS through its Atmos API using the following method.

An Atmos subtenant must be created before ECS can be accessed through its Atmos REST API. Atmos subtenants are created using the ECS CLI. Refer to the following article for more information on how to create an Atmos subtenant in ECS: [https://community.emc.com/docs/DOC-27822](https://community.emc.com/docs/DOC-27822)
1. Expand **Storage Resources** and right-click **Libraries**. Choose **Add > Cloud Storage Library**...

![Dell EMC Atmos Cloud Storage Library](image)

**Figure 6** Dell EMC Atmos Cloud Storage Library

2. Select **EMC Atmos** as the cloud storage type. Fill in the appropriate **Access Information** described as follows (also shown in Figure 6) using information from the ECS user account in the web portal.

- **MediaAgent**: Choose the Media Agent that will drive data to ECS.
- **Service Host**: ECS supports both HTTP (port 9022) and HTTPS (port 9023) for the Atmos protocol. Enter the IP Address or DNS name of the load balancer that has been placed in front of the ECS nodes including the port. In the above example, a load balancer has been specified using the mapped port 80 for client/application-facing connections.
- **Access Key ID**: Enter the ECS object user name.
- **Secret Access Key**: Enter the ECS object user secret access key.
- **Token ID**: Enter the token ID in standard Atmos format, using `<SubtenantID>/<ECS Object User Name>`.
- **Root Folder**: Enter the name of the ECS Atmos folder to use. If the folder does not exist, it will be created automatically.
Note: In SP14 and above, there is a Detect button next to the Root Folder field. Do not use this to detect the Atmos root folder because it is not supported.
3.1.5 OpenStack

Configure MediaAgents to connect to ECS through its OpenStack API using the following method.

1. Expand Storage Resources and right-click Libraries. Choose Add > Cloud Storage Library…

![OpenStack Cloud Storage Library](image)

**Figure 8** OpenStack Cloud Storage Library

2. Select OpenStack Object Storage as the cloud storage type. Fill in the appropriate Access Information shown in the following (also shown in Figure 9) using information from the ECS user account in the web portal.

   - **MediaAgent**: Choose the Media Agent that will drive data to ECS
   - **Service Host**: ECS supports both HTTP (port 9024), and HTTPS (port 9025) for the OpenStack protocol. Enter the IP Address or DNS name of the load balancer that has been placed in front of the ECS nodes including the port and authentication type (v1.0 or v2.0). In the above example, a load balancer has been specified using the mapped port 80 for client/application-facing connections and V2 authentication.
   - **Username**: Enter the Username in standard OpenStack/Swift format, using `<ECS Namespace>:<ECS Object User Name>`.
   - **API Key**: Enter the Swift password.
   - **Container**: Enter the name of the Swift Container (bucket) to use. If the bucket does not exist it will be created automatically.

The Use Internal URLs and Region fields can be ignored.
Note: In SP14 and above, there is a Detect button next to the container field which can be used to select the bucket to use if one was manually created.

3.1.6 Centera CAS

For more information about the Commvault Centera storage library, refer to the following URL: [http://documentation.commvault.com/commvault/v11/article?p=features/centera_clusters/centera_clusters_getting_started.htm](http://documentation.commvault.com/commvault/v11/article?p=features/centera_clusters/centera_clusters_getting_started.htm)

MediaAgents using the Windows operating system can be configured to connect to ECS as a Centera Cluster using the Centera SDK 3.2 P5 (includes support for application registration). For more information on how to configure ECS for CAS access and how to create a Centera PEA file, refer to the ECS administration guide.

Creating a Centera storage library is a two-step process:

**Step 1: Add ECS Centera cluster information**

1. Use the following steps to set up the environment variable for the ECS Centera PEA file location.
2. From the computer desktop, right-click My Computer and click Properties.
3. From the Advanced tab, click Environment Variables. In the Environment Variables dialog box, under System Variables, select New.
4. In the Variable name field, enter CENTERA_PEA_LOCATION.
5. In the Variable value field, add the full path to the PEA file including the file itself. For example, enter C:\PEA\Centera.pea. It is recommended that there are no spaces in either the folder location or the file name.
6. Click OK and click OK again. Reboot the computer to commit the variable.
7. On the ribbon in the CommCell Console, click the Storage tab, and click Expert Storage Configuration.
8. Under **Available MediaAgents**, select the MediaAgents that are attached to the library and on which the library needs to be configured, and click **Add >>**. Click **OK**.
9. To configure any shared library, make sure to select all MediaAgents that share the library.
10. If a device has already been configured for the MediaAgent, the device appears in the **Expert Storage Configuration** dialog box.
11. For the cluster, select the cluster server as the MediaAgent.
12. Configure the library:

   - If a library is not configured, a message appears that asks to configure the library. Click **OK**.
   - If a library is already configured, this message does not appear. Subsequently, the detected devices are displayed with the detection status detect success in the **Expert Storage Configuration** dialog box.

13. From the **Expert Storage Configuration** window, click the **Start** menu, point to *Centera*, and select **Cluster**.
14. From the **Centera Cluster List** dialog box, click **New Cluster**.
15. From the **Centera Cluster** dialog box, click **Add**.
16. From the **Add Access Node** dialog box, enter either the DNS name or IP address of one of the ECS nodes. Use port 3218 for the Centera cluster to access ECS data and click **OK**.

   **Note:** The ECS CAS protocol supports internal load balancing, so this address should resolve directly to a single ECS node. No load balancer should be used.

17. A message appears that asks to select a MediaAgent to detect the cluster information. Click **OK**.
18. Detect the Centera cluster using the following steps:

   a. In the **Access Nodes** box, select the DNS name or IP address recently added from the **Access Node** list.
   b. Select the MediaAgent to detect the Center cluster from the **Detect Information from MediaAgent** list. Click **Detect**.
   c. The ClusterID, Capacity, and Version information of the ECS Centera cluster are displayed. If an error occurs, verify and ensure the information accuracy in the **Add Access Node** dialog box. Click **OK**.
   d. The Centera cluster is displayed. Click **Close**.

**Step 2: Configure an ECS Centera cluster as a disk library**

The following procedure describes the steps involved in configuring an ECS Centera cluster as a disk library.

1. From the **Expert Storage Configuration** window, click the **Start** menu, point to *Centera*, and click **Add Centera Library**.
2. In the **Add Disk Library** dialog box, select the name of the MediaAgent that will control the library and click **OK**.
3. Select the ECS Centera cluster to configure from the **Centera Cluster** list and click **OK**.
4. The Centera cluster is added as the mount path to the Centera disk library.
5. The disk library appears in the **Expert Storage Configuration** window with the configured status.
3.2 Storage policy

3.2.1 Stream randomization

When a storage policy is configured to use more than one device stream, it is important that the device streams are equally used. If enabling stream randomization, the system chooses the device streams randomly. This should increase the rate of data transfer by copying data from different streams in parallel (see Figure 10).

![Figure 10 Storage policy stream randomization](image)

3.2.2 Alternate data paths for a shared library configuration

A data path is the combination of the MediaAgent, library, drive pool, and scratch pool that are used by the storage policy copy to perform a backup operation. Each storage policy copy has a default data path that is used to perform backup operations.

Alternate data paths can be defined in each of the storage policy copies to ensure the success of backup and other operations that use the storage policy. There are many benefits of this feature such as increased resiliency in the event of component failure using automatic switchover to an alternate data path, or load-balancing (round robin) between alternate data paths to increase synchronous operations.

Refer to Alternate Data Paths for more information.

3.3 Subclient

Subclients should be configured to use the storage policy created above with a full and incremental schedule policy that is appropriate for the client.
3.4 Configuration and tuning

The backup process is a resource-intensive operation, especially the compression and deduplication process, and the amount of spare CPU, RAM, and hard-drive performance on a client all contribute to its backup performance.

If a client requires the total backup time to be reduced, in addition to monitoring the CommServe and MediaAgent performance, the client performance should also be monitored during the backups. Additional CPU and RAM can help with the compression deduplication hashing algorithms and therefore reduce the backup time. Be sure to verify that disk performance and other workload or maintenance operations are not impacting the backup. If resources cannot be easily increased, test if offloading the compression, deduplication, or both processes from the client to the MediaAgents helps. Balance the performance gains of these adjustments against the increased network bandwidth consumption between the client and the MediaAgent.

Commvault recommends tuning several parameters when using Cloud Storage Libraries. The following section describes these settings and where to configure them. Refer to CommCell Performance Tuning for detailed information.

3.4.1 Cloud library recommendations

Follow these recommendations for optimized cloud library performance settings. The cloud library itself has tuning parameters as well as each of the mount paths.

<table>
<thead>
<tr>
<th>Table 3 Cloud library properties</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Location</strong></td>
</tr>
<tr>
<td>Mount paths</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

3.4.2 MediaAgent properties

The additional Commvault settings shown in Table 4 are used to make changes to the local Windows server registry settings. These settings generally apply to the CommServe and Windows MediaAgent systems.

In V11, the setting nCloudEnableHighPerformance is set to 1 by default. This setting enables a group of tunable parameters for optimal cloud library performance. Each parameter can also be set individually if required. These parameters are referenced in the following table in the Additional Settings tab section.
### MediaAgent properties

<table>
<thead>
<tr>
<th>Location</th>
<th>Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control tab</td>
<td>Data Transfer / Optimize for concurrent LAN backups: This option is enabled by default and is useful when the MediaAgent is used for concurrent data protection operations from a large number of clients. For example, if there are 25 clients concurrently using the MediaAgent for 50 or more data transfer streams in a specific operation window, it is recommended to enable this option, which helps optimize the operating system resources on the MediaAgent.</td>
</tr>
<tr>
<td>General tab</td>
<td>Maximum number of parallel data transfer operations / Restrict to: Select how many concurrent read/write operations need to be established to the MediaAgent. The default number of concurrent read/write operations to a MediaAgent is 25. This value can be set between 1 and 75. If the Optimize for Concurrent LAN Backup option is enabled in the MediaAgent Properties &gt; Control tab, the default value is set to 100 (this value can be set to a maximum of 500).</td>
</tr>
<tr>
<td>Additional</td>
<td>nCloudMaxSubFileSizeKB Description: Used to set the size of the chunks uploaded to the cloud servers.</td>
</tr>
<tr>
<td>Settings tab</td>
<td>Default value: 32768</td>
</tr>
<tr>
<td></td>
<td>Recommended value: 131072</td>
</tr>
<tr>
<td></td>
<td>Use: This parameter sets the chunk size of files to upload to ECS when not using deduplication.</td>
</tr>
<tr>
<td></td>
<td>Note: Increasing this value may consume an additional 12.8 GB of RAM on the MediaAgent when all threads are being utilized.</td>
</tr>
<tr>
<td></td>
<td>nCloudPageModeSubFileSizeKB Description: Used to set the size of the chunks to upload for the files in page mode.</td>
</tr>
<tr>
<td></td>
<td>Default value: 8192</td>
</tr>
<tr>
<td></td>
<td>Use: This parameter sets the chunk size (page mode) of files to upload to ECS when using deduplication.</td>
</tr>
<tr>
<td></td>
<td>Note: This parameter sets the deduplication chunk size. Caution should be taken when modifying this value because it has the potential to cause large amounts of un-reclaimable garbage in ECS.</td>
</tr>
</tbody>
</table>
## Backup/archive target configuration

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Default Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>nCloudNumOfReadAheadFiles</td>
<td>Sets the number of the read-ahead temp files to keep for the subfiles under the chunk folder on the cloud server. Should be set to nCloudNumOfDownloadThreads + 1.</td>
<td>4</td>
</tr>
<tr>
<td>nCloudNumOfReadAheadThreads</td>
<td>Sets the number of the read-ahead threads for the subfiles under the chunk folder on the cloud server. <strong>Note:</strong> This setting is being deprecated in SP14.</td>
<td>2</td>
</tr>
<tr>
<td>nCloudNumOfUploadThreads</td>
<td>Sets the number of concurrent upload threads for each file uploaded to the cloud servers. <strong>Note:</strong> This key will be deprecated starting from SP13 unless nCloudGlobalUploadThreadPoolMaxCount is set to 0.</td>
<td>3</td>
</tr>
<tr>
<td>nCloudSocketReceiveBufferBytes</td>
<td>Sets the number of receive bytes a connection before imposing flow control. <strong>Use:</strong> This is used for buffer control when utilizing a cloud library as the storage destination.</td>
<td>1024576</td>
</tr>
<tr>
<td>nCloudSocketSendBufferBytes</td>
<td>Sets the number of send bytes on a connection before imposing flow control. <strong>Use:</strong> This is used for buffer control when utilizing a cloud library as the storage destination.</td>
<td></td>
</tr>
<tr>
<td>nCloudUseTempFile</td>
<td>Used to enable or disable the usage of the temp file for uploading data to the cloud server. <strong>Default value:</strong> 0 (Disabled)</td>
<td></td>
</tr>
</tbody>
</table>
### Subclient properties

**Table 5** Subclient properties

<table>
<thead>
<tr>
<th>Location</th>
<th>Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced / Performance Tab</td>
<td>Number of Data Readers: Modifying the number of Data Readers made a significant difference with the performance of the storage libraries. Refer to the Commvault documentation <a href="#">Data Readers</a> regarding data streams for more information on optimizing.</td>
</tr>
<tr>
<td></td>
<td>Allow multiple data readers within a drive or mount point: Specifies whether multiple data reads are allowed for a single Windows physical drive or Unix mount point during backups on this subclient.</td>
</tr>
<tr>
<td></td>
<td>Resource Tuning / Network Agents: Specifies whether multiple data reads are allowed for a single Windows physical drive or Unix mount point during backups on this subclient.</td>
</tr>
</tbody>
</table>

### Storage policy properties

**Table 6** Storage policy properties

<table>
<thead>
<tr>
<th>Location</th>
<th>Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Paths tab / Media Agent properties</td>
<td>Chunk Size (MB): Specifies the sum of the size of files written to each stream’s folder.</td>
</tr>
<tr>
<td></td>
<td>• Use Application Setting: Specifies that the data path will use the chunk size configured for the agent in the Chunk Size tab of the Media Management Configuration dialog box available in the Control Panel.</td>
</tr>
<tr>
<td></td>
<td>• Use Specified Value: Specifies that the data path will use the specified value as the chunk size. This value will override the chunk size configured for a specific Agent, and the default for the disk. Use the space to type the new value.</td>
</tr>
<tr>
<td></td>
<td>Note: Commvault recommends using a 4 GB chunk size if cloud connectivity is through a high-speed network of about 1 GB per second or higher.</td>
</tr>
</tbody>
</table>
Backup/archive target configuration

<table>
<thead>
<tr>
<th>Block Size (KB): Specifies the size of each file written to each stream’s folder. For on-disk libraries to Centera Clusters, a block size of 32 KB is always used and user-defined block sizes are not applicable.</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Use Media Type Setting: Specifies that the data path will use a fixed block size of 64 KB to read/write to all media types.</td>
</tr>
<tr>
<td>• Use Specified Value: Specifies that the data path will use the specified value as the block size. There are several considerations associated with changing the block size.</td>
</tr>
</tbody>
</table>

Note: Commvault recommends using a 1024 KB block size if cloud connectivity is through a high-speed network of about 1 GB per second or higher.

Device Streams: Commvault recommends setting the number of device streams to 50 or above, when cloud connectivity is through a high-speed network of about 1 GB per second or higher.

Block-level Deduplication factor: Commvault recommends increasing the block level factor to 512 KB, when cloud connectivity is through a high-speed network of about 1 GB per second or higher.
4 Advanced configuration options

This section reviews advanced configuration and performance considerations when using ECS as the backup/archive target.

4.1 Commvault configuration best practices

Refer to the Commvault documentation for up-to-date Commvault specific details regarding performance and scaling best practices including Cloud Connection Performance tuning when using a high-speed network of 10 GbE or higher.

4.2 Retention

Commvault supports the ECS retention extension which enables the storage of objects using a WORM model. With retention policies, objects are prevented from being deleted or overwritten for a fixed amount of time or indefinitely. Retention enables you to meet regulatory requirements that require a layer of protection against object changes and deletion.

Refer to the ECS Administration Guide to learn more about bucket-level retention in ECS.

**Note:** This functionality is supported in Commvault 11 SP14 and above for the ECS S3, EMC Atmos, and OpenStack object services.

**Note:** Commvault does not support native deduplication when using retention. Commvault needs to do micro pruning from time to time which updates file sets, once metadata and index files are rewritten, you can’t update the content again until the retention period expires which may cause pruning issues.

4.2.1 ECS bucket configuration

The ECS bucket being used for retention must have retention set to an equal or lower period than the retention value set in Commvault. Otherwise, application errors may occur.

![Figure 11 ECS Bucket Retention Period](image)

4.2.2 Storage policy

The retention rule on the storage policy should be set higher or equal to the ECS bucket-level retention.
### 4.2.3 Cloud storage library retention setting

Enable WORM on the storage library.

<table>
<thead>
<tr>
<th>Location</th>
<th>Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>Mark Archive files as Read-Only: Configure the library to create archive/backup files as read-only to achieve WORM functionality on cloud storage. This enables the corresponding read-only lock mechanism on the cloud storage and saves the archive/backup files as read-only files. The expiration date for the read-only lock is set to match the data retention time (backup or archiver retention days, whichever is higher) established in the storage policy copies. Refer to the Commvault article Configuring Files as Read Only for more information.</td>
</tr>
</tbody>
</table>

#### Figure 12 Cloud Storage General Tab, enable read-only files

### 4.2.4 MediaAgent retention property

Enable ECS retention in the MediaAgent Advanced Settings section.

<table>
<thead>
<tr>
<th>Location</th>
<th>Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced Settings</td>
<td>nCloudDellECSRetention: This option enables Commvault to use the ECS retention extension. The default value is enabled [1]. Values: 0 – Disables retention 1 – Enables retention 2 – The current implementation is keying off a specific server response from ECS and uses an ECS extension that is not available in the standard S3 API. If a customer decides to override the server response via a load balancer or other method, then this value can be used to send the ECS retention extension regardless of the server response.</td>
</tr>
</tbody>
</table>
Multi-stream restore operations

By default, restore operations use a single stream. You can use multiple streams so that restore operations complete faster.

Reference the Commvault documentation for detailed information on multi-stream restores.

The following agents support multi-stream restore operations:

- Windows File System Agent
- Linux File System Agent
- AIX File System Agent
- FreeBSD File System Agent
- HP-UX File System Agent
- Solaris File System Agent
- Virtual Server Agent (supports up to 10 streams, no support for multiple nodes)
5  Best practices and general information

This section describes best practices and general information when using ECS as a Commvault backup/archive target.

5.1 Load balancing and connectivity

- The best practice for distribution of S3, Atmos, and OpenStack REST traffic among the ECS nodes is to use an external IP load balancer. This resource describes several approaches to configure a load balancer with ECS.
- The ECS CAS protocol supports internal load balancing, so the address should resolve directly to a single ECS node; no external load balancer should be used.
- Commvault provides several cloud tools which enable the ability to connect to the cloud library and explore its contents or test connectivity by uploading and downloading data.

5.2 Deduplication

- When using deduplication, the cloud storage report may not accurately report the amount of space used. The option Process volume size updates for cloud mount paths can be enabled in Media Management to reread the size on disk as pruning takes place.
- Refer to the Commvault Deduplication Best Practices document for recommended best practices.
- Deduplication in Commvault combines data from numerous clients into objects stored in ECS and uses pruning methods to remove data from chunks as needed based on the retention policy.

Combining backups and archives for applications/clients requiring the same retention policies to dedicated Commvault storage policies may ensure a higher likelihood that all data for a given object will eventually be a candidate for pruning and deleted to free up capacity.

5.3 Data pruning

- Micro pruning is a term that is specific to using the DDB (deduplicated database), and it removes specific references and data from disk storage based on retention criteria.
- Logical pruning occurs and decrements reference counts in the DDB once all references to an archive file (the way Commvault tracks data) become zero, and the data is then physically deleted.
  - This feature uploads data in 4 GB chunks (may be less depending on the size of the job) and breaks the chunk into 8 MB parts (default size: nCloudPageModeSubFileSizeKB) for upload to the cloud.
  - The entire part is deleted when all deduped records in it are prunable, then the 8 MB part is pruned (micro-pruning).
  - The 4 GB chunk will be deleted when all dedupe records for all part files are prunable.
- For cloud storage configured with deduplication, micro pruning is enabled by default to ensure that data will be pruned from cloud storage, as soon as the backup job meets the retention rules, without sealing the corresponding DDB (deduplication database).

Note: Due to the way Commvault prunes deduplicated data, there is potential that a large amount of unreclaimable garbage will remain in ECS until all dedupe records for all part files are pruned.
Macro pruning is the method used for non-deduplicated data. Once archive files associated to jobs meet the retention period then the data is immediately removed from disk or storage. Macro pruning also occurs when the DDB is sealed and all jobs tied to that DDB have met retention (the aging process is then like non-dedupe data because Commvault will no longer use the DDB process as noted in micro prune).

- The archive (chunk) maximum size is 4 GB (may be less depending on the size of the job) which is broken into 32 MB (default size - \( nCloudMaxSubFileSizeKB \)) parts for upload to the cloud.
- The entire archive is deleted when pruned.

It is highly recommended to monitor the capacity utilization of ECS. You can monitor the capacity utilization of storage pools and nodes as well as the entire VDC from the ECS Portal Monitor > Capacity Utilization page. The following reference document clarifies the various monitoring options. [https://www.emc.com/collateral/TechnicalDocument/docu88138.pdf](https://www.emc.com/collateral/TechnicalDocument/docu88138.pdf)

### 5.4 Compression

Software compression can be used to minimize the amount of data transferred over the wire and reduce the storage footprint. Software compression is available for all storage media. Note that for Auxiliary Copy operations, data is copied in the compressed state.

The following entities include options to compress data:

- **Client**: Client compression is useful if the client and MediaAgent reside on separate computers and the client must send its data using a network. Client compression reduces the network load since the data is compressed before it leaves the client.
- **MediaAgent**: MediaAgent compression can be useful if the MediaAgent software resides on a computer that is more powerful than the client computer. Using software to compress data can be processor intensive. Consequently, you may want to use MediaAgent compression for client computers with limited processing power.

**Note:** There is certain data that is compressed or deduplicated in origin such as SQL or Microsoft Exchange that may not require further processing.

### 5.5 Multi-site replication

In a multi-site deployment, more than one VDC is managed as a federation and/or geo-replicated. In a geo-replicated deployment, data can be read or written from any active site within the defined replication group.

Geo-replication provides enhanced protection against site failures by having multiple copies of the data, for example, a primary copy of the data at the original site and a secondary copy of the data at a remote site/VDC. Refer to the [ECS Architectural Guide](https://www.emc.com/collateral/TechnicalDocument/docu88138.pdf) for detailed information.

Managing application traffic both locally and globally can provide high availability (HA) and efficient use of ECS storage resources. HA is obtained by directing application traffic to known-to-be-available local or global storage resources. Optimal efficiency can be gained by balancing application load across local storage resources. An IP load balancer is required to direct traffic to an alternate site when the primary site is unavailable. This can be accomplished using an active/standby or fixed weighting configuration using GSLB.

Careful planning should be observed to evaluate what data should be replicated or kept local.
6 Troubleshooting and known issues

The following errors may occur when the Cloud storage library is incorrectly configured with ECS.

<table>
<thead>
<tr>
<th>API</th>
<th>Error</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atmos</td>
<td>‘[[Cloud] HTTP Error - forbidden.]]’</td>
<td>The object user was not appended to the token ID.</td>
</tr>
<tr>
<td></td>
<td>Occurs when configuring the Atmos cloud library.</td>
<td></td>
</tr>
<tr>
<td>Swift</td>
<td>‘[[Cloud] The argument is invalid.]’</td>
<td>The username is incorrect. Make sure that the username is in the format 'namespace:object user'.</td>
</tr>
<tr>
<td></td>
<td>Or</td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘[[Cloud] HTTP Error - bad request.]’</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Occurs when configuring the OpenStack Swift cloud library</td>
<td></td>
</tr>
<tr>
<td>Swift</td>
<td>‘[[Cloud] Access denied.]’</td>
<td>Make sure that '/v1.0' or '/v2.0' has been appended to the Service Host URL</td>
</tr>
<tr>
<td></td>
<td>Occurs when configuring the OpenStack Swift cloud library</td>
<td></td>
</tr>
</tbody>
</table>

6.1 Data retention forecast and compliance report

When using SP14 and retention is configured, the Data Retention Forecast and Compliance Report displays unknown error in the Delay Reason for physical space cleanup column. This is expected to be fixed in a future service pack.
A  Technical support and resources

Dell.com/support is focused on meeting customer needs with proven services and support.

Storage technical documents and videos provide expertise that helps to ensure customer success on Dell EMC storage platforms.

A.1  Related resources

Additional resources that may be helpful when configuring Commvault with Dell EMC ECS.

<table>
<thead>
<tr>
<th>Description</th>
<th>Detail / links</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commvault V11 Documentation</td>
<td>Commvault Version 11 documentation</td>
</tr>
<tr>
<td>Commvault Cloud Connection Performance Tuning</td>
<td>Commvault recommendations to get the maximum performance for cloud-based backup and restore for high-speed networks (1 GBps or higher).</td>
</tr>
<tr>
<td>Commvault Streams Overview</td>
<td>Overview of Commvault stream architecture</td>
</tr>
<tr>
<td>Commvault Cloud Storage Overview</td>
<td>Commvault Cloud Storage overview, Details regarding configuration, troubleshooting and best practices.</td>
</tr>
<tr>
<td>Dell EMC ECS Overview and Architecture White Paper</td>
<td>White paper providing an in-depth architecture overview of Dell EMC ECS</td>
</tr>
<tr>
<td>Commvault CommCell Sizing</td>
<td>Commvault sizing guidelines for CommServe and MediaAgents</td>
</tr>
<tr>
<td>Dell EMC ECS High Availability Design White Paper</td>
<td>ECS High Availability architectural details</td>
</tr>
<tr>
<td>Dell EMC ECS Best Practices</td>
<td>ECS Best Practices Guide</td>
</tr>
</tbody>
</table>