COHESITY

Version 2.1

October 2021

Protect Microsoft 365 Exchange Online with Cohesity

Reduce Complexity with Seamless Data Protection for Exchange Online

ABSTRACT

As more and more organizations transition to Microsoft 365 for their business-critical operations, the need to protect M365 data has grown more than ever. Cohesity delivers the features and capabilities organizations need to ensure their data is protected and their operations are uninterrupted. Now you can embrace Microsoft 365 without compromising organizational data-protection SLAs or worrying about compliance breaches.

Table of Contents

Figures

Tables

Introduction to Exchange Online Data Protection with Cohesity

Growing adoption of Microsoft 365®, Microsoft's business productivity tools suite delivered in a SaaS (Software as a Service) model (including Exchange Online OneDrive™ for Business, Sharepoint Online, and more), has sparked fresh conversations on the division of responsibilities between Microsoft, now a *service provider*, and the organization that owns the data that is created and managed in Microsoft 365. Before SaaS models, data protection was the responsibility of the organization, as both the data and the infrastructure to manage it was owned by the organization, and typically resided on-premises with the organization.

However, with the introduction of SaaS like Microsoft 365, where the infrastructure is distributed, data protection is now a shared responsibility between the customers, who create and use the data in cloud-based applications, and providers, who run the infrastructure for those applications.

As the SaaS provider, Microsoft protects your data in the event of:

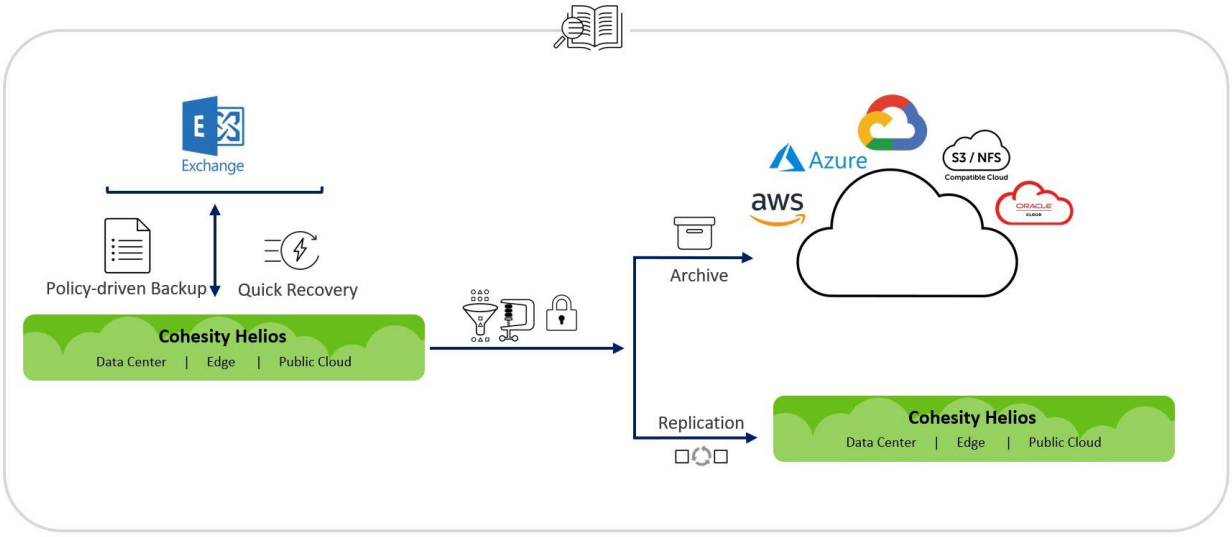
- Application failures
- Infrastructure failures
- Disaster scenarios

However, you, the customer, are responsible for protecting the data from a series of other threats, such as:

- Accidental data deletion
- Insider attack
- Security breaches
- Administrative error
- Ransomware/malware attack *(Email phishing is one of the major attack vectors for Ransomware attacks)*

Cohesity offers to help you mitigate the risk of data loss due to these reasons by providing a web-scale, space-efficient, highly available, and robust platform to back up your Exchange Online data. Cohesity also helps you to meet the compliance requirements around backup SLAs and long-term retention.

Figure 1: Backup Exchange Online using Cohesity



Features and Benefits of Cohesity Data Protection for Exchange Online

Cohesity Helios platform is a modern, software-defined platform for data management. Taking inspiration from its web-scale architecture and leveraging its unique distributed file systems (SpanFS), Cohesity offers high-scalability and reliability, all the while maintaining its simplicity and sophistication for all your Exchange Online data protection requirements.

Cohesity's flexible architecture allows easy expansion, further increasing operational simplicity, and improved TCO. It is a solution that works on-premises, on qualified Cisco, HPE, Dell, or Cohesity C Series platforms, in the public cloud, as well as remote and branch offices on hypervisors of your choice like VMware and Hyper-V. Hence, you can choose where you want to keep your backup for exchange online data.

Table 1: Features and Benefits of Cohesity

FEATURES	BENEFITS
Scheduled Automated Backups	Automated scheduled backup for Exchange Online enables you to have multiple points in time for the data.
Attribute-based Global Searches	Search globally through the different backups based on multiple mail attributes to find the relevant emails to be restored easily.
Granular Recovery	Restore data at various granularities such as mailbox, folder-level, and email-level.
Flexible Deployment Options	Cohesity can be deployed on-premises, in the cloud, and in a hybrid mode, giving customers full flexibility in deployment planning.
Data Lock and Legal Hold	Data Lock and Legal Hold functionality help meet regulatory compliance.
Forever Incremental	'Once full and forever incremental' backup helps organizations meet the backup SLAs.
Global Deduplication	Global deduplication ensures space-saving across different workloads.
Auto-Protect Feature	The auto-protect feature enables protection for any new user added to the domain.
Data Migration across Microsoft 365 domains	Allows migration of data from one Microsoft 365 domain to another.
Disaster Recovery for Microsoft 365 backup	Replication ensures you have a disaster recovery copy for the Microsoft 365 data.
CloudArchive Microsoft 365 backup	You can archive the Microsoft 365 data using CloudArchive features for long term retention.

Protect Your Exchange Online Data

Cohesity offers a policy-based, highly-scalable data protection infrastructure for your Exchange Online data. With a few steps, you can set up Cohesity Helios platform for your data protection requirements.

Understand the Backup Workflows

Cohesity employs two different workflows to back up and protect your Exchange Online data. Since Exchange Online is a SaaS offering and it throttles data download over time, Cohesity makes use of an incremental-forever approach which drastically reduces network traffic and improves backup times.

After the user initially configures Cohesity for Exchange Online, Cohesity starts off the data protection with a full backup of your Exchange Online data, which is followed by continuous incremental backups on regular intervals as per the schedule defined in the policy.

Before you start to set up the data protection and perform the first full backup, it is advisable to understand what goes under the hood in each backup.

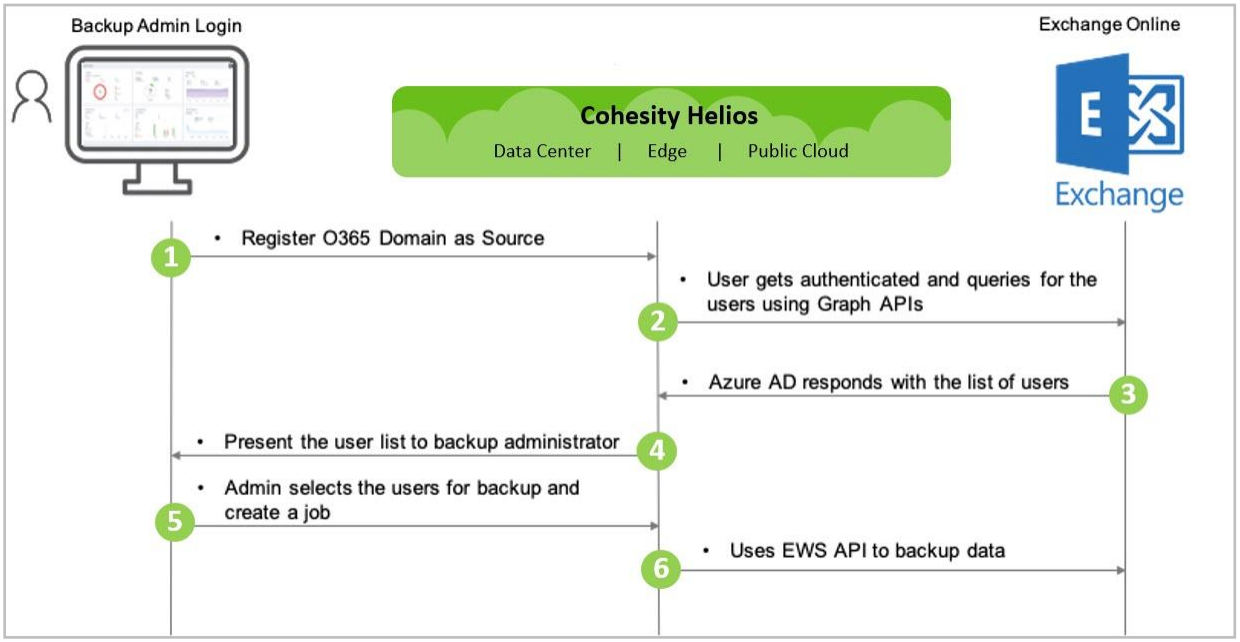
First Full Backup Workflow

The initial full backup of your Exchange Online data involves the following steps:

1. The administrator configures the Exchange Online data protection with Cohesity.
2. The administrator either schedules the first backup in the Protection Policy or manually triggers it.
3. Cohesity queries Azure by a combination of Graph APIs and Exchange Web Service (EWS) APIs for licensed domain users who have mailboxes assigned.
4. Once the list of users and related mailboxes is available, mailboxes are queued for backup on different nodes and are backed up in parallel.
5. Cohesity leverages EWS APIs to back up mailboxes. Mailboxes are backed up in parallel. Each mailbox is queried for a list of folders. One folder is backed up at a time for a mailbox.
6. Data is indexed as the backup for the mailboxes complete.

NOTE: Since Microsoft 365 domain is registered as a source on Cohesity using a user with the “Application Impersonation” role, the same user can be used to impersonate and backup other user mailboxes as well.

Figure 2: Cohesity Full Backup Workflow for Exchange Online



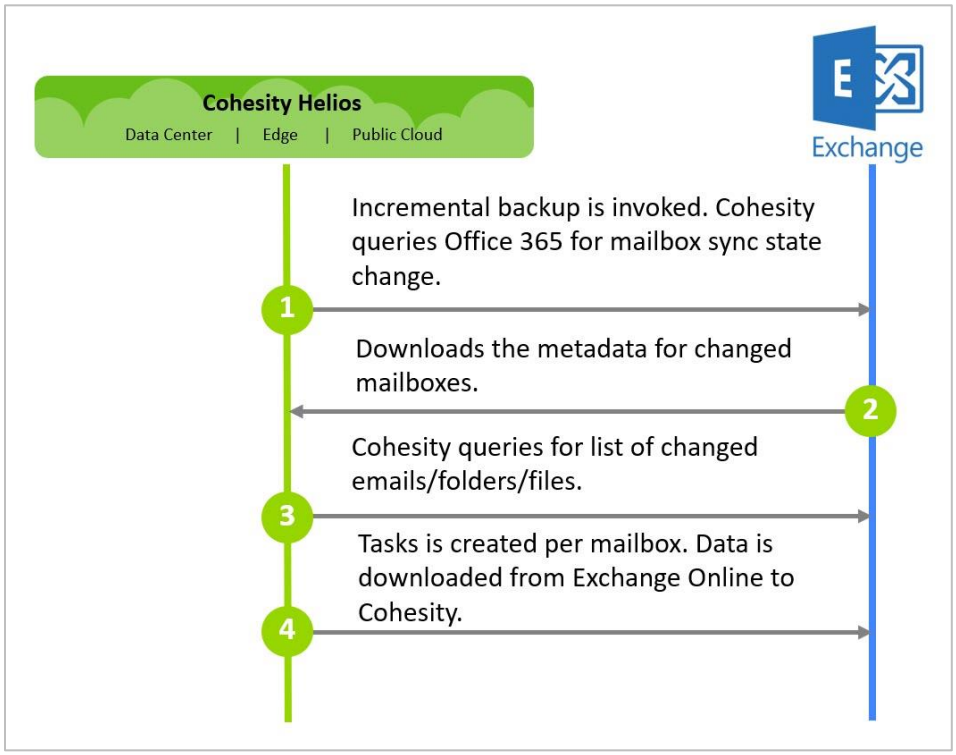
Incremental Backup Workflow

After the first full backup is complete, Cohesity runs the Protection Job at regular intervals as you specify, performing incremental backups to capture all data that has changed since the last backup.

The process of configuring and running incremental backups in Cohesity involves the following steps:

1. The administrator either schedules automatic incremental backups in the Protection Policy or manually triggers it when required.
2. Upon triggering, Cohesity queries Exchange Online for the sync state. If the current sync state is different from the previous state, it means there are changes to the Exchange Online.
3. Cohesity queries the Exchange Online to retrieve the list of changed folders in each mailbox.
4. A backup task is created for each mailbox that has changed folders.
5. Changed emails are backed up and indexed in batches.

Figure 3: Cohesity Incremental Backup Workflow for Exchange Online



Now that you have understood what happens under the hood during your Exchange Online data protection with Cohesity, it is time to configure Cohesity and initiate the first full backup of your data.

Configure Cohesity for Exchange Online Data Protection

To set up Cohesity for your Exchange Online data protection, you will do the following tasks in sequence:

1. Plan and Prepare
   a) Add Roles in Microsoft 365 User Account.
   b) Create an Azure Application with Minimum Permissions.
2. Register your Microsoft 365 domain as a source on Cohesity.
3. Protect Exchange Online Mailboxes
   a) Create or edit a Protection Policy. A Protection Policy defines how and when data objects, such as your Exchange Online data in Microsoft 365, are protected, replicated, and archived.
   b) Create a Protection Group. Protection Groups combine operational requirements with the business data protection SLAs that are defined in a Protection Policy. Using a Protection Group, you can combine mailboxes that require the same data protection SLAs and attach a Protection Policy to it.
4. Once set up, the full backup is triggered immediately unless a different **Start time** is selected under **Protection Group** > **Advanced Setting**.

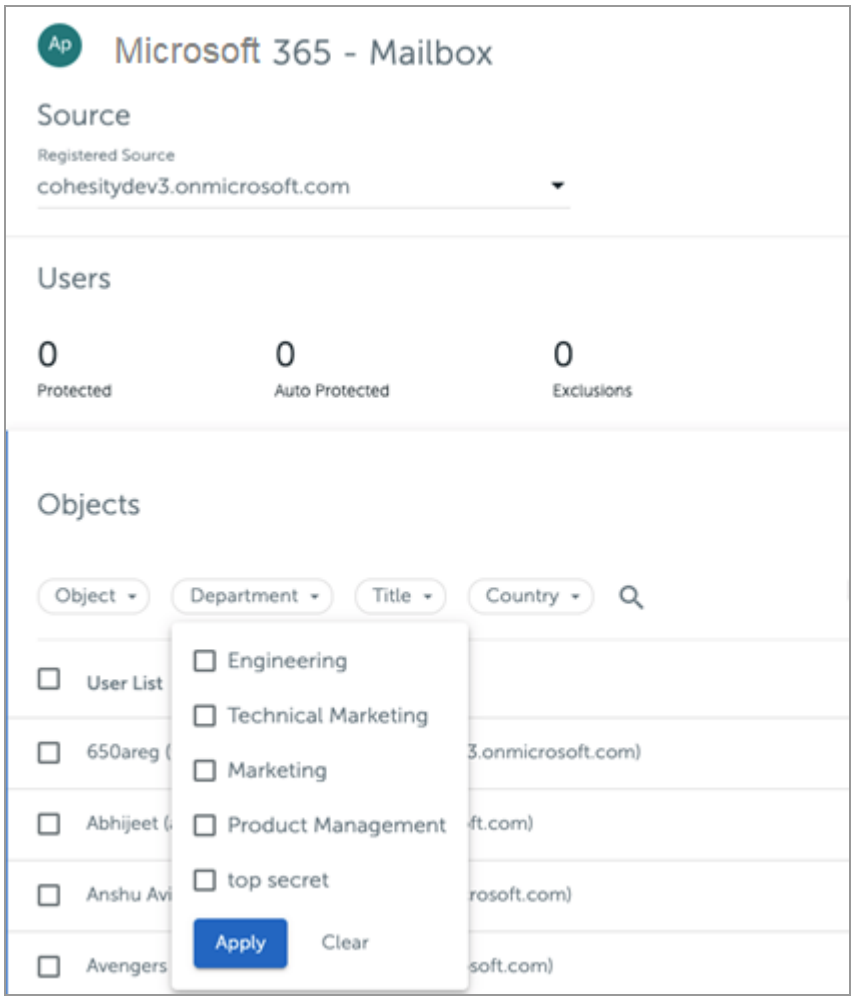
5. After the first backup, Cohesity runs the incremental backups periodically as you scheduled in the policy.

NOTE:

**If you are using Cohesity version:

1. **6.2**: Cohesity uses the EWS API to run the user mailbox query. (Doesn't require app ID or access key to register.)
2. **6.3 or later**: Cohesity uses a combination of the EWS API and Microsoft Graph API to query a list of users who have an Exchange Online license and mailbox assigned (Requires app ID and access key to register.) and EWS APIs for backup and restore.

**From Cohesity 6.5.1, you can now use Azure active directory user attributes under *Job Info (Title, Department, Manager, etc.) and Contact Info (Location, Address, etc.)* to filter out users to create a Protection Group.



Restore Exchange Online Data

Once you have backed up your Online Exchange data, Cohesity allows you to restore it whenever you require with a few simple and intuitive steps. Restore can be performed at various granularities, such as at the mailbox, folder, and email-levels. You can also restore in-place, out-of-place, and across Microsoft 365 domains. The workflow for restoring mailboxes varies slightly from the workflow for restoring emails and folders.

Restore Mailboxes

To start a mailbox restore, the administrator searches through the backups by username for the mailbox that needs to be restored and then launches the restore operation.

At that point, Cohesity:

1. Recreates the folder hierarchy from metadata in the backup.
2. Searches the internal metadata to locate the data blocks.
3. Collates the data and sends it to M365 using EWS Restore APIs.
4. Once M365 acknowledges the restore, the task status shows **Success**.

For more, see Restore Mailbox in the Online help.

Restore Emails and Folders

To start email/folders restore, the administrator searches through the backups using the filters listed in the table below.

Table 2: Mail Attributes as Search Filters to Narrow Your Searches

SEARCH FILTERS	
From	Mail sender email address
To	Mail recipient email address
Time Interval	Time interval within which the email existed
Subject	Subject line for the email
Folder Name	Folder name which needs to be restored
cc	The user email address that was cc'd in the email
bcc	The user email address that was bcc'd in the email
Within an M365 domain	Search for the email within the specified Microsoft 365 domain
Within a Mailbox	Search for the email within the specified user mailbox

SEARCH FILTERS	
Within a Protection Job	Search for the email within a particular Protection Job

Once the search yields the desired result, the administrator selects the desired mail or folder and launches the restore operation.

At that point, Cohesity:

1. Searches through the internal metadata to locate the data blocks containing the requested data.
2. Data is collated and sent for restore to M365 using EWS restore APIs.

For more, see Restore Emails and Folders in the Online help.

Maintain Business Continuity and Compliance

Cohesity provides two mechanisms for protecting your data from disruptions and disasters:

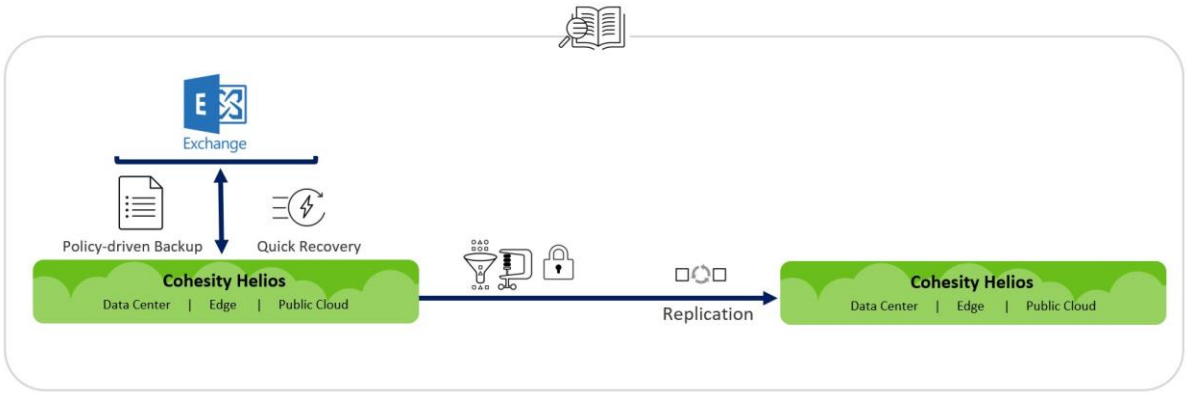
- **Replication** offers a simple way to store your valuable data as replicas in an alternate Cohesity cluster and later retrieve it in the event of major business disruptions such as natural disasters and IT failures.
- **Cloud Archive** provides you an easy and economical way to archive data to an external storage medium such as a cloud for long term retention and restore it to an alternate Cohesity cluster upon requirement.

Replicate Backups to Other Cohesity Clusters

Organizations can take advantage of Cohesity's replication for cost-effective disaster recovery (DR). Cohesity provides a policy-based data replication solution from the core to the cloud to the edge, from one cluster to another cluster in your DR site.

As part of replication, Cohesity always performs deduplication and compression first, and sends only the changed data over the network. In the event of the primary site becoming unavailable, application and backup admins can failover to the DR site for backup and recovery of their data.

Figure 4: Replicate Backups to Other Cohesity Clusters



To replicate Microsoft 365 data to another Cohesity cluster:

1. Configure replication to another Cohesity cluster. See, Replication and Remote Access Setup in the Online help.
2. Edit the Protection Policy that you created for backup (see, Create a Protection Policy in Protect OneDrive for Business with Cohesity) and configure the **Replication** settings to meet your requirements. See Step 12 in Create or Edit a Standard Policy in the Online help.

The replication settings will take effect in the next scheduled Protection Run.

Use Archive for Long Term Retention

Long term retention is often a business requirement predominantly due to compliance that regulates the business. With Cohesity CloudArchive, you can securely retain your data for a long time in a very cost-effective and air-gapped storage medium.

With CloudArchive, Cohesity provides a policy-based method to archive data to public clouds (AWS, Azure, GCP), to any S3-compatible storage, and/or to any NFS mount point. It offers a complete, self-contained copy of your backup, containing backup data, backup metadata, indexing data, and deduplication fingerprints, to pull through a jeopardizing disaster and maintain business continuity.

Cohesity upholds the storage and network efficiency of backups using strict deduplication, compression, and incremental-forever approach.

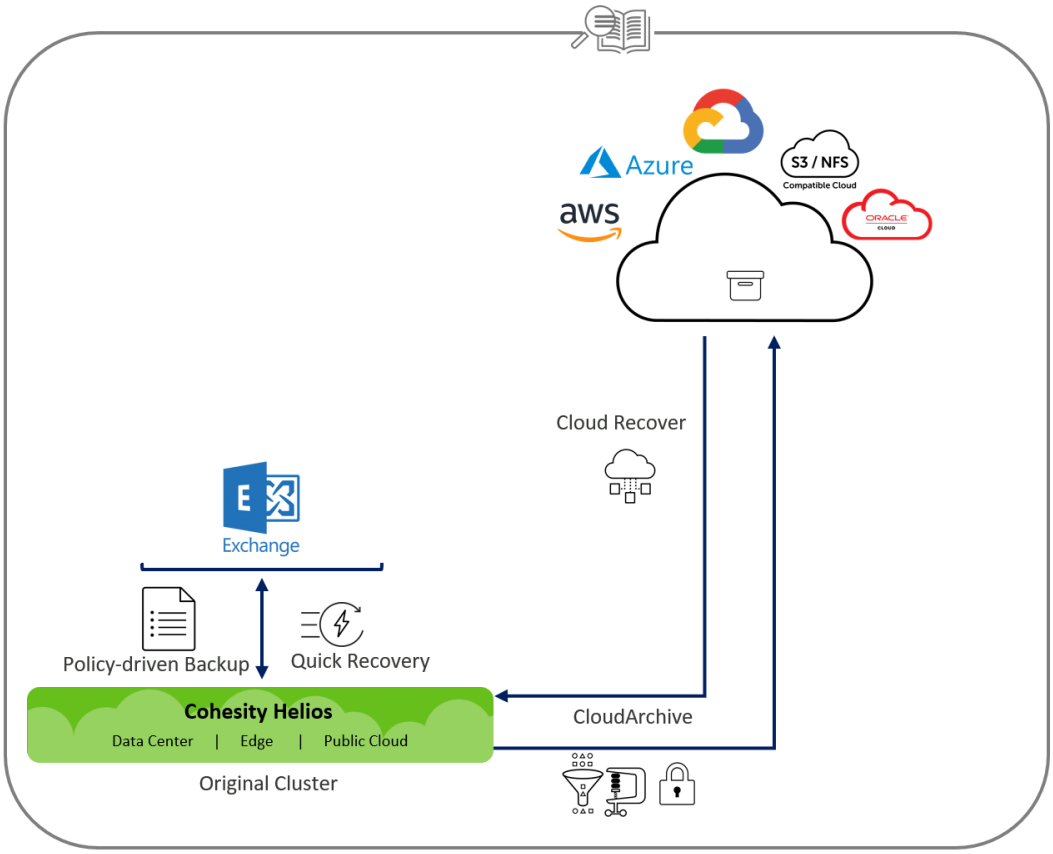
Figure 5: Leverage Public Cloud Infrastructure for Long-term Data Retention and Archival

CloudArchive is very flexible. You can use it with AWS, Azure, GCP, NAS, and S3-Compatible cloud object storage.

Access Your CloudArchived Data

Once your data is archived, Microsoft 365 administrators can also take advantage of Cloud Recover the archived snapshots to the source cluster and then restore it to Exchange Online.

Figure 6: Cloud Recover Workflow

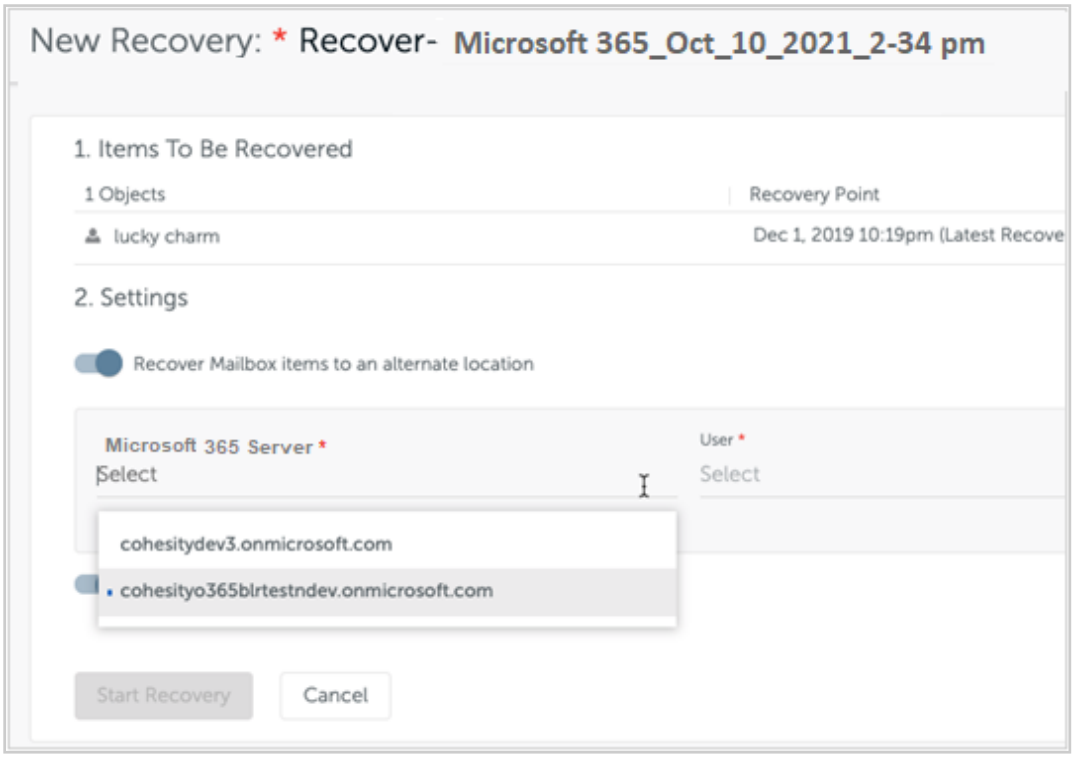


To learn more, see the *CloudArchive Deployment & Recovery Guide* for AWS, Azure, GCP, NAS, and S3-Compatible cloud object storage.

Migrate Data From One Microsoft 365 domain to Another

Due to merger and acquisition, rebranding, or a divestiture, sometimes enterprises are faced with a situation wherein they need to migrate Exchange Online data across Microsoft 365 domains. Cohesity enables you to tackle this challenge by offering to recover data to alternate locations, which includes multiple domains.

You can add multiple Microsoft 365 domains as a source and recover data across domains using the “Recover Mailbox items to an alternate location” option.



Best Practices

For best results using Cohesity to protect your Exchange Online data, Cohesity recommends:

- Set **Retry Options** to **1**, as the Microsoft 365 connector already has the retry mechanism handling failed requests. There is no need to increase this setting in the Cohesity Protection Policy.
- Make thorough use of the **Search** filters when setting up a granular recovery.
- Use **Auto-Protect** to avoid having to add new users to the Protection Job manually.
- To be able to recover at granular levels later, enable Indexing during Protection Job creation.
- To use the **Exclude Folder** option during Protection Job creation (for example, to exclude some custom folders in user Inboxes from backups), you must provide the *exact name* of the custom folders, and the folder names are case-sensitive.
- Avoid adding the same users in multiple Protection Jobs.

Your Feedback

Was this document helpful? Send us your feedback!

About the Authors

Saurabh Singh is a Staff Product Manager at Cohesity. In his role, he focuses on product management for Cohesity Data Management as a Service (DMaaS), Security, SaaS Backup, and Secure Multi-tenancy.

Other essential contributors included:

- Mayank Joshi, Product Line Manager

Document Version History

VERSION	DATE	DOCUMENT HISTORY
0.9	April 2019	First draft
1.0	July 2019	First release
2.0	Aug 2020	Second release
2.1	Oct 2021	Rebranding updates

ABOUT COHESITY

Cohesity radically simplifies data management. We make it easy to protect, manage, and derive value from data -- across the data center, edge, and cloud. We offer a full suite of services consolidated on one multicloud data platform: backup and recovery, disaster recovery, file and object services, dev/test, and data compliance, security, and analytics -- reducing complexity and eliminating mass data fragmentation. Cohesity can be delivered as a service, self-managed, or provided by a Cohesity-powered partner.

Visit our website and blog, follow us on Twitter and LinkedIn and like us on Facebook.