COHESITY

Version 1.1

July 2024

CloudArchive & CloudRetrieve Deployment & Recovery Guide for NAS

Use NAS to Store Your Protected Data for Long-Term Retention and Disaster Recovery

ABSTRACT

Long-term data and application retention is critical to organizations that seek to prevent data loss and meet security, legal, and compliance requirements. Cohesity Platform™ and DataProtect™ offer robust on-premises solutions for enterprise data protection and storage. Cohesity's CloudArchive™ and CloudRetrieve™ bring data protection and recovery together with any NAS.

Table of Contents

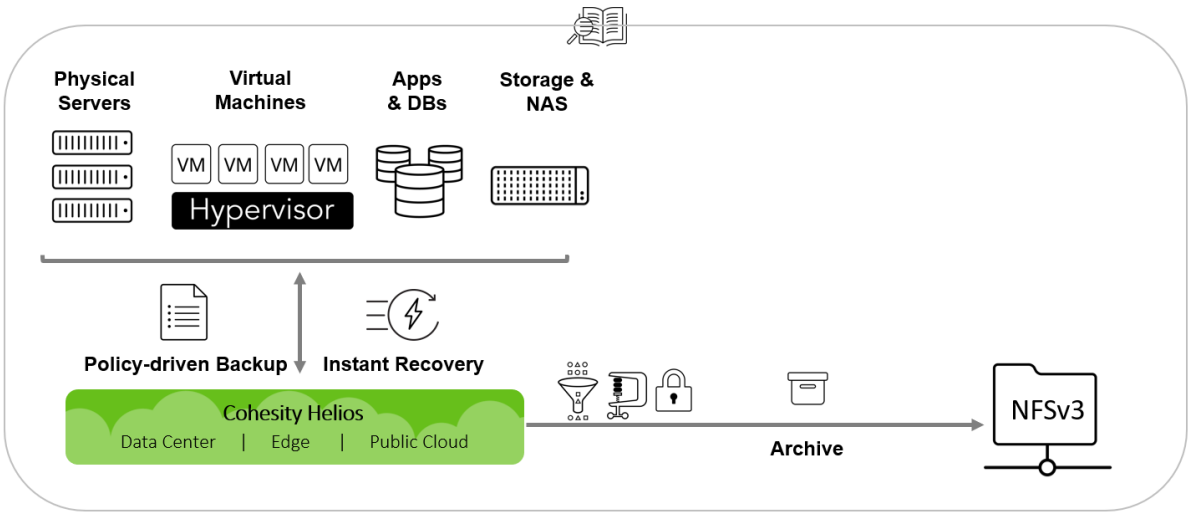
Figures

Tables

CloudArchive Connects NAS to Cohesity Platform

Long-term data and application retention is critical to organizations that seek to prevent data loss and meet security, legal, and compliance requirements. In most organizations, there are many underutilized NAS (Network-Attached Storage) appliances. With CloudArchive, customers can use them to store data for long-term data retention and disaster recovery.

Figure 1: CloudArchive Connects NAS to Cohesity Platform



IMPORTANT: Cohesity Platform supports only the NFSv3 protocol for connecting to NAS. However, you can use NAS appliances from any vendor, on premises or in the cloud, as long as they use NFSv3.

With Cohesity Platform, IT organizations save time by quickly archiving data to multiple targets—public clouds, private clouds, any S3-compatible device, NAS, and QStar managed tape libraries, eliminating the need for cloud gateways and point solutions to archive, while increasing operational efficiency and lowering total cost of ownership (TCO).

NOTE: This document covers only Cohesity Platform operations for archiving to NFSv3-compatible storage, and not tape or S3 targets.

For archiving to tape, see Long Term Retention to Tape with Cohesity DataProtect solution guide. For archiving to S3-compatible storage, see CloudArchive & CloudRetrieve Deployment & Recovery Guide for S3 . For archiving to public cloud vendors, see guides for AWS, Azure, and Google Cloud Platform.

CloudArchive Features and Benefits

Cohesity's CloudArchive provides many key features, each of which delivers several benefits to organizations and their IT administration staff. Specifically:

Table 1: CloudArchive Features and Benefits

FEATURES	BENEFITS
Policy-based cloud archival	• Easy to use. • Archive unique data differently by mapping Protection Policies to the required SLA. • Reduce bandwidth & storage costs.
Off-site copies	• Geo-redundancy • Disaster recovery
Deduplication and compression	• Efficient data transfer and storage.
Granular recovery	• Instantly locate VMs, files, and folders. • Recover just what you need.
Encryption	• Data is secure both in flight and at rest.

Classes of Supported Storage for CloudArchive

CloudArchive supports all of the leading object storage from cloud providers, any S3-compatible device, as well as NAS from storage vendors, including but not limited to:

Table 2: Supported Storage Classes

AWS	AZURE	GCP	ORACLE CLOUD	NAS-NFSv3	S3 COMPATIBLE
S3	Hot Blob	Regional Multi-Regional	Object Storage	• NetApp Cluster Mode • NetApp 7 Mode • EMC Isilon • EMC Unity • Pure FlashBlade • Generic NFS • Cohesity NAS	• IBM Cloud Object Storage • Cloudian • IIJ Object Store • Iron Mountain • EMC ECS • Pure FlashBlade • DDN • NetApp StorageGRID • Scality • Cohesity S3
S3-IA	Cool Blob	Nearline			
Glacier	Archive	Coldline	Archive Storage		
S3 to Glacier					

NOTE: This guide helps you set up archiving to and recovery from any NAS. The only requirement is that you use NFSv3. Cohesity Platform interacts with NFSv3 mount points in the same ways, regardless of vendor.

CloudArchive Terminology

There are several terms that are important to understand as you learn about the specific ways CloudArchive works.

Table 3: CloudArchive Terminology

TERM	DEFINITION	NOTES
Cohesity Platform	Cohesity Platform consolidates secondary data and applications, including backups, files, objects, test/ dev, and analytics on a single, software-defined platform. Inspired by web-scale architecture. Cohesity Platform is a scale-out solution based on a unique distributed file system, SpanFS®.	
Archive	Completely self-contained copy of the backup (data and metadata) that is stored outside the Cohesity cluster.	
Archive Chain	The set of a Full Archive and the Incremental Archives that depend on it and the preceding Incrementals.	If the Full Archive is lost for any reason, the entire archive chain becomes unusable. If an Incremental Archive is lost, the restore points that follow it are lost as well.
CloudRetrieve	The process of retrieving an archived Protection Job and its Job Run details from an External Target to a different cluster. Used for geo-redundancy and disaster recovery.	CloudRetrieve cannot be performed on the same cluster that performed the archive.
Cluster	An instance of Cohesity Platform.	
Deduplication Chain	The set of a Reference Archive and all the archive chains that depend on it for deduplication. This includes the Scheduled Full and Incremental Archives for each archive chain in the deduplication chain.	These dependencies determine when Cohesity Platform can retire and eventually delete Reference Archives.
External Target	Any storage to which data is sent outside the source Cohesity cluster.	Archive to Cloud, Tape, NFS, and replication targets are all External Targets in Cohesity Platform.

TERM	DEFINITION	NOTES
Full Archive	A full copy of the Protection Job that is archived.	
Incremental Archive	An archive that records just the changed data since the most recent archive.	
Protection Job	Defines operational requirements, such as which source objects to protect, the Protection Policy to use, and operational considerations like indexing, exclusions and inclusions, alerts, app consistency, and more.	Each Protection Job has a schedule of Job Runs, and each archive is a collection of those Job Runs.
Protection Policy	Reflects business needs of Recovery Point Objectives (RPOs) and Recovery Time Objectives (RTOs) by defining the frequency and retention requirements of backup, archival, and replication.	
Scheduled Full Archive	A Full Archive that runs at regular intervals (configurable, 90 days by default).	The Scheduled Full Archive does not send the same amount of data, as it is deduplicated against the Active Reference Archive. In those cases when there is no Active Reference Archive, the data sent for the Scheduled Full is deduplicated only with itself and not against any other archive. For example, if the Active Reference Archive size is 100GB and the Scheduled Full deduplication usage is 60%, then only 40GB is sent. If there is no Active Reference Archive, then the size of the Scheduled Full is 100GB.
Source-Side Deduplication	The process of eliminating redundant copies of data to reduce storage use before sending over the network. Redundant data blocks are replaced with a pointer to the unique data copy, which ensures that only unique instances of data are transferred over the network and retained on storage media.	Reduces storage as well as network bandwidth requirements and, in doing so, saves time and money.

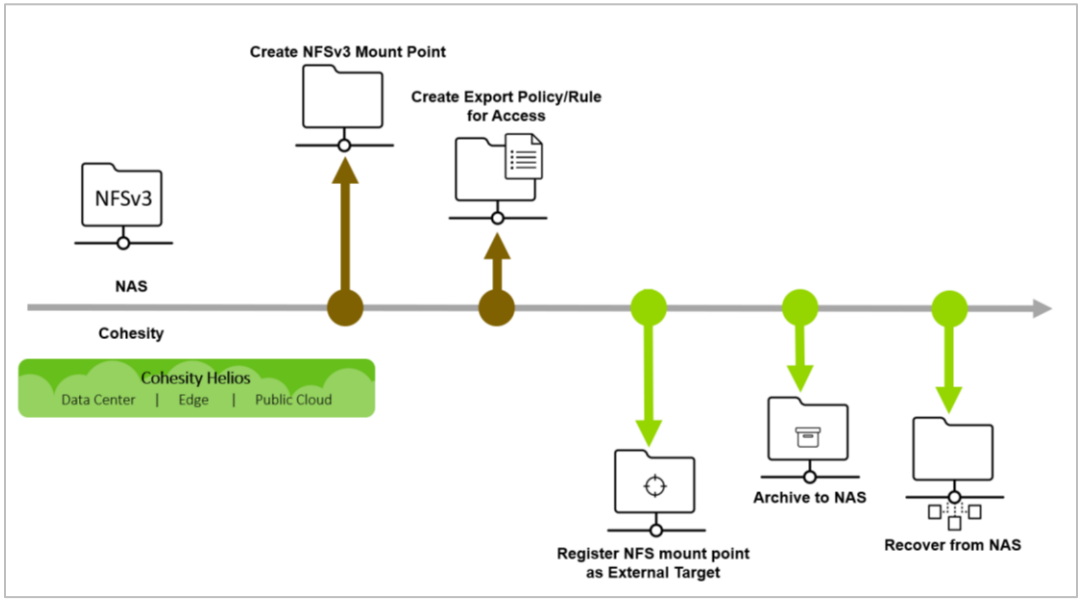
TERM	DEFINITION	NOTES
Recover	Retrieve an entire data object, such as a VM or database, or granularly recover files and folders from an External Target onto the original cluster.	
Reference Archive	The Full Archive against which all subsequent Incremental Archives (in the archive chain) and Scheduled Full Archives *as well as their* Incrementals are deduplicated.	All Reference Archives are full archives. A new Reference Archive is created when Cohesity Platform detects that deduplication with it is below 50%. **NOTE**: 50% is the default threshold. This is internally configurable, but changing this value only delays *when* (and not whether) the full data set is sent.
Retired Archive	A Reference Archive that is no longer used for deduplication.	

CloudArchive High-Level Workflow

At the highest level, leveraging CloudArchive for NAS involves several sequential tasks:

1. Create an NFS mount point on NAS.
2. Create an export policy (or rule) that grants access to the node IPs of your Cohesity cluster.
3. Register your NFS mount point to Cohesity Platform as an External Target.
4. Archive your data to NAS.
   a) Create a Cohesity Protection Policy.
   b) Create a Cohesity Protection Job.
5. Recover your data from the NFSv3 archive.

Figure 2: Leverage NAS with Cohesity



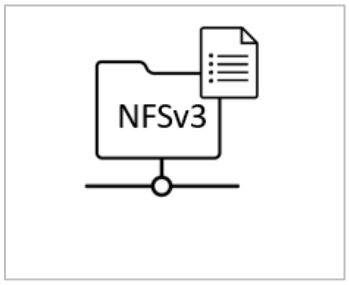
Create NFS Mount Point

The first thing you'll do is create an NFSv3 (Network File System Version 3) mount point on the NAS appliance from the vendor of your choice. Though the process is slightly different for each vendor, it always involves creating an NFS mount point on your NAS appliance and an export policy that provides the access for Cohesity Platform to read and write data.

Create Export Policy for Access

Next, follow your vendor's procedure to create an export policy (or rule) that grants access to your Cohesity Platform node IPs.

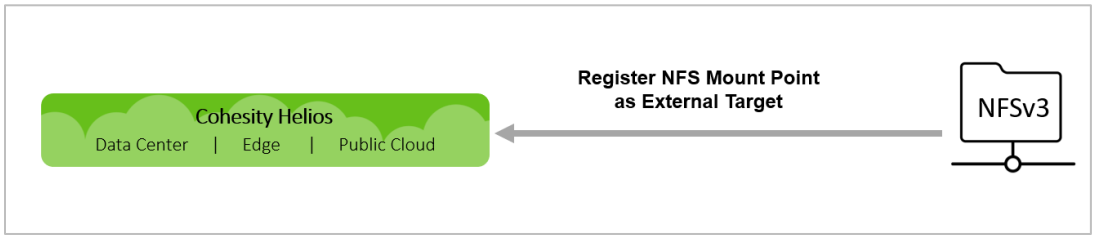
Figure 3: Create Your NFS Mount Point and a Policy for Access



Register NFS Mount Point

Next, you need to connect that new NFS mount point to Cohesity Platform by registering it as an External Target in Cohesity Platform. For this, you'll need the NAS Host (IP address or FQDN) and mount path.

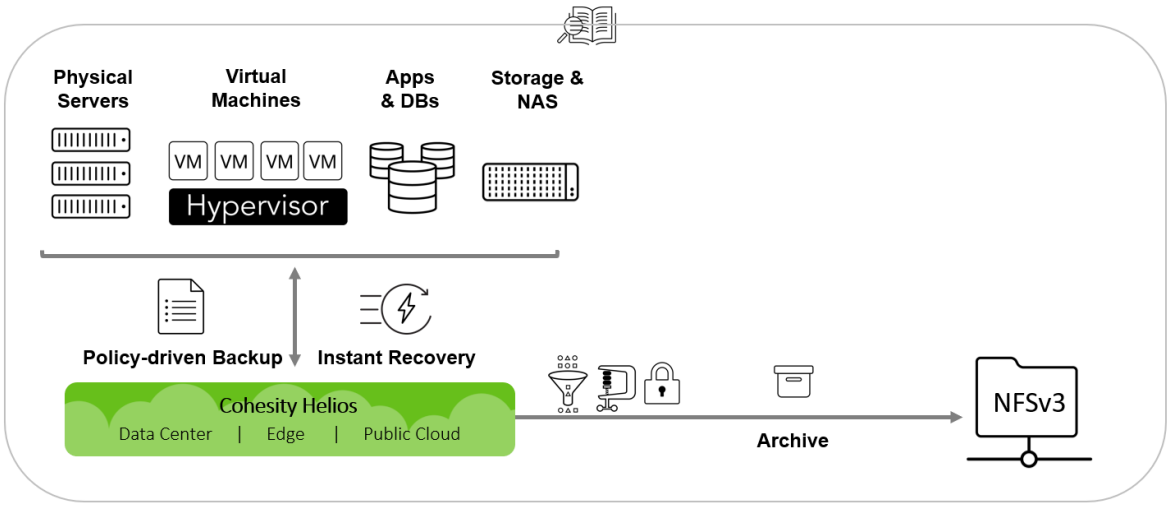
Figure 4: Register NFS Mount Point with Cohesity Platform



Archive Your Data to NAS

With your NFS mount point now registered with Cohesity Platform, the next step is to archive your data by creating a Protection Policy (which reflects your business needs, like frequency and archival retention requirements) and running a Protection Job (where you define operational requirements, such as which data objects to protect, the Protection Policy to use, indexing, alerts, and SLA requirements).

Figure 5: Archive Data to NAS



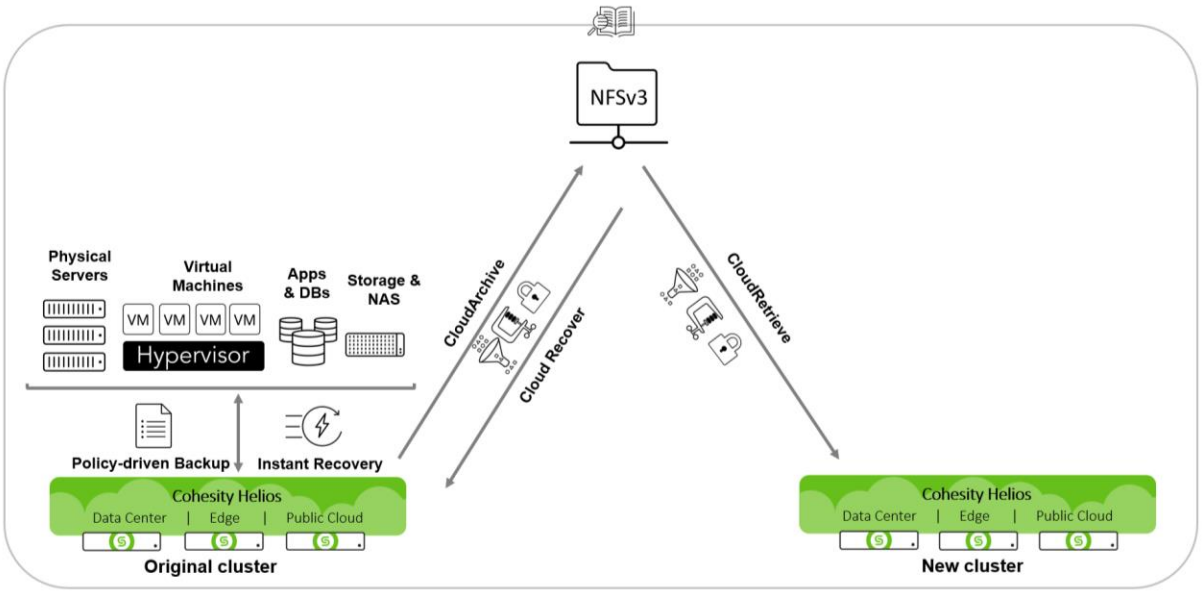
Recover Your Data from NAS

As with any archive, the challenge is to locate, identify, and recover it quickly and reliably. Cohesity includes an indexing engine that enables rapid search and recovery of files and virtual machines from archives stored both on-premises and in the cloud. As data is being backed up, Cohesity Platform's indexing engine indexes the data. This enables extremely fast, wild-card search results that are then used for granular recovery.

Once your data is archived with CloudArchive, when you need to access it again, you'll be able to get it back using Cloud Recover (to your original cluster) or CloudRetrieve (to a new cluster).

- **Cloud Recover to source cluster**: Recover entire objects (VMs, databases, NAS, etc.), or individual files and folders, to your original cluster.
- **CloudRetrieve to new cluster**: Retrieve your previously archived data onto an entirely new cluster, for disaster recovery and geo-redundancy.

Figure 6: Recover Data from NAS with Cloud Recover and CloudRetrieve



In the next chapter, we cover the individual steps that are involved in each of these tasks. Following that, we walk through the specific procedures for connecting your NAS to Cohesity Platform, archiving your data to NAS, and recovering and restoring your data from NAS.

Leverage Your NAS with Cohesity

This chapter provides a quick overview of the sequence of actions that you will be undertaking to set up your NFS mount point as an External Target in Cohesity Platform. In the chapters that follow, you'll find step-by-step instructions for each CloudArchive feature you can use with your NFS mount point.

Create and Register NFS Mount Point

Start with setting up your NAS by creating a new NFS mount point.

When you create the NFS mount point, be sure to:

- Allow read and write access.
- Disable NFS root squash.
- Whitelist the node IPs (and not the VIPs) of your Cohesity cluster.

> **TIP**: To see your cluster's node IPs, log in to Cohesity Platform and select **Platform** > **Cluster**. On the **Cluster** page, click the **Nodes** tab.

Once you have the NFS mount point, capture and save the NAS host IP or FQDN (fully qualified domain name) and mount path. Using the NAS host and mount path, you can then register your NFS mount point as an External Target in Cohesity Platform.

> **NOTE**: The same NFS mount point can be registered multiple times in the same cluster as different External Targets, as well as in multiple clusters.

Required NAS Fields

To register your NFS mount point as an External Target, Cohesity Platform requires the following fields:

- NAS Host
- Mount Path

Also, remember that the NFS mount point must have read and write access permissions, and the NFS export policy (rule) must whitelist the Cohesity cluster node IPs and disable NFS root squash.

Configure Your Policy-based Archive

Once Cohesity Platform registers your NFS mount point as an External Target, you will create a Protection Policy to define your business needs. The Protection Policy allows you to incorporate the NFS External Target that you created earlier as an archive target with a specific retention period.

In the Policy, you configure how virtual and physical servers, databases, and unstructured data are protected:

- Backup frequency and retention period.
- Whether to have your backups archived, how often, and how long to retain.

> **NOTE**: You can add more than one archival schedule to the same Policy, and you can use the same or a different External Target, with the same or different frequency and retention.

- Which External Target to use (in this case, your newly registered NFS mount point).

Protect Your Data

Protection Jobs combine operational requirements with the business requirements that are defined in a Protection Policy.

In the Job, you select the source, which data objects from that source to store, the Protection Policy and the storage domain (the named storage location) to use, and operational details such as Start Time, End Date, QoS Policy, Pre & Post Scripts, and more. See the all the advanced Protection Job settings in the Appendix.

Once you save a Protection Job, it will run on the schedule you define.

> **NOTE**: Multiple Protection Jobs can use the same Protection Policy, but each Job can have only one Policy.

Recover Data from your Archive

When the time comes to recover your archived data, Cohesity Platform gives you three options:

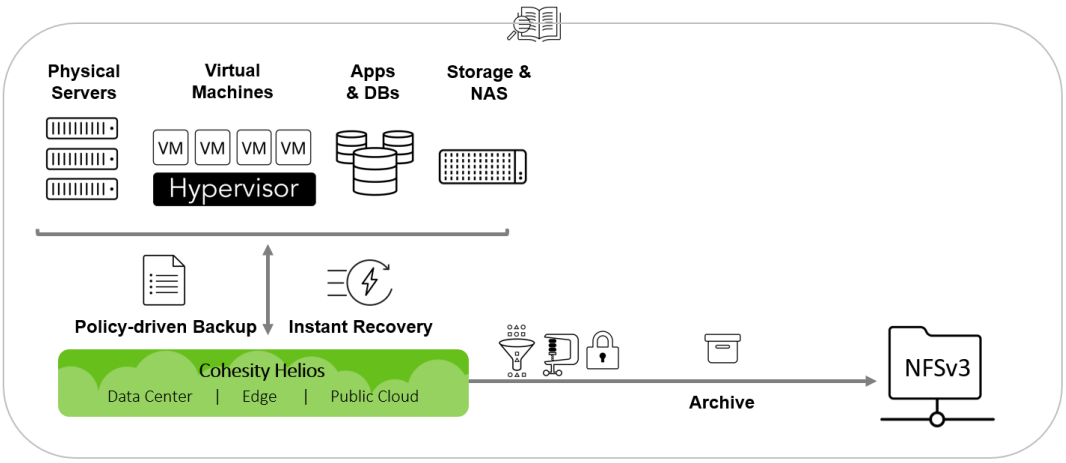
- Restore entire data objects (VMs, databases, NAS, etc.).
- Recover individual files and folders.
- Retrieve your data onto an entirely new Cohesity cluster (for disaster recovery, etc.).

For instructions, see Recover Data from CloudArchive below.

Connect NAS to Cohesity

Cohesity's CloudArchive enables customers to connect seamlessly to various classes of cloud and on-premises storage as an extension of the data center infrastructure. Customers are using CloudArchive to reduce their reliance on tape for cost-effective, long-term data retention, as well as a low-cost, disaster-recovery solution.

Figure 7: Cohesity CloudArchive with NAS



Let's get started!

Prepare Your NAS for CloudArchive

Cohesity Platform supports archiving to NAS using the NFSv3 protocol. First, you'll create an NFS mount point and write an export policy or rule that grants read and write access and whitelists the node IPs of your Cohesity cluster.

Create an NFS Mount Point

Create an NFS mount point using the procedure defined by your storage provider.

On your NAS, Cohesity recommends:

- **Disable snapshots**. If your NAS appliance supports periodic snapshots, it is better to disable them, to avoid trapping storage consumption due snapshot retention when the archive has deleted the data.
- **Disable encryption**. By default, any data that is sent and stored from Cohesity Platform is already encrypted unless you have disabled it. Because you already exchange some performance for data security by encrypting the data, there is no benefit to incurring this impact twice.
- **Disable deduplication and compression**. The archive data being written to the External Target is already deduplicated, compressed, and encrypted on the source side by your Cohesity cluster. Because deduplicating already deduplicated and encrypted data yields no gains, Cohesity recommends disabling both inline and post-process deduplication and compression in your NAS.

When you register your NFS mount point to Cohesity Platform, encryption, compression, and deduplication are already enabled by default, which Cohesity recommends keeping, as they reduce your network and storage utilization.

Get Access to Your NAS

Next, you need to add an NFS export policy (rule) that whitelists your Cohesity cluster's node IPs.

1. Using your NAS vendor's procedure, create an export policy or rule that:
   a) Grants read and write access to your NAS.
   b) Disables NFS root squash.
2. Whitelist the node IPs (and not the VIPs) of the Cohesity cluster. To get the list of node IPs, log in to Platform and select Platform > Cluster > Nodes.

TIP: For simplicity, you can whitelist the node IP subnet instead of individually adding each of them. This way, when you expand or reduce your cluster, you won't have to manage the new IP addresses separately.

Use NAS as External target in Cohesity Platform

Now that you have the NFS mount point that you need, you're ready to connect it to Cohesity Platform (whether your cluster is on-premises, Cloud Edition, or Virtual Edition).

Cohesity External Target Features

When you register your NFS mount point as an External Target in Cohesity Platform in the following section, you will be able to enable or disable:

Table 4: External Target Options

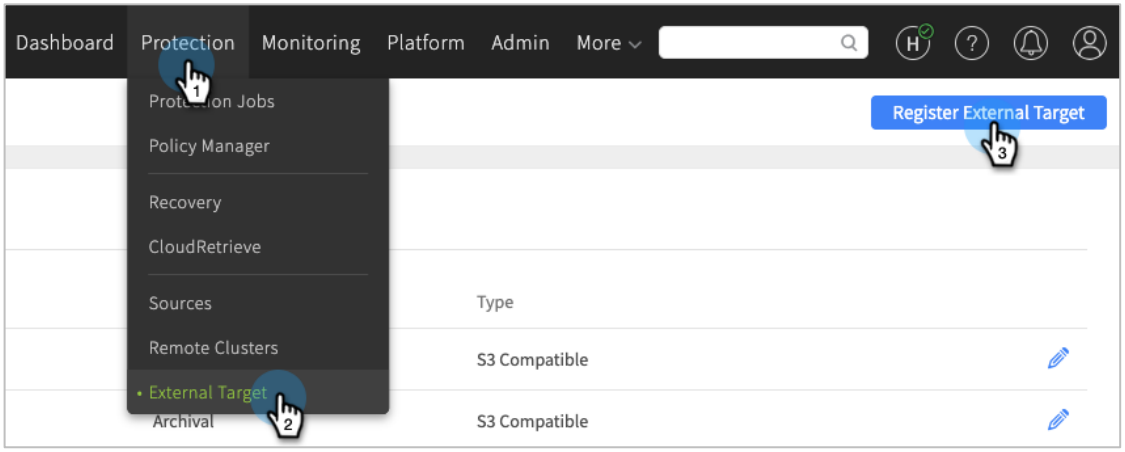
FEATURE	DESCRIPTION
Encryption	By default, Cohesity Platform writes the data into External Targets in encrypted format in real time. You can disable it, but Cohesity recommends you leave it enabled in almost all cases, except when the data is already encrypted. You can choose to keep your encryption key in the cloud with your archive, or, for additional security, to manage it manually. **NOTE**: If you choose the manual option, you will need to download the key after registering the External Target and store it outside the Cohesity cluster.
Compression	Reduces the impact on data transfers and data storage. Useful except when the data format doesn't compress well, such as with databases and large image files.

FEATURE	DESCRIPTION
Source-Side Deduplication	The process of eliminating redundant copies of data to reduce storage use. Redundant data blocks are replaced with a pointer to the unique data copy, which ensures that only unique instances of data are sent across the network and retained on storage media, and dramatically reduces the impact on bandwidth and storage utilization. Cohesity strongly recommends it in all cases.
Incremental Archival	An archive that records just the changed data since the most recent archive. This allows you to return to any restore point without having to create, transfer, and keep a backup copy of your whole dataset each time. Cohesity strongly recommends this setting in all cases. If this option is not enabled, it will send a full archive on every archive run.
Bandwidth Throttling	If needed, you can throttle the upload and download bandwidth that is consumed by network traffic between Cohesity Platform and an External Target. You can also limit bandwidth throttling to a specific time range, if there are particular days and times when it is needed. **NOTE**: You cannot set Bandwidth Throttling to lower than 1Mbps.

Register NFS Mount Point as External Target

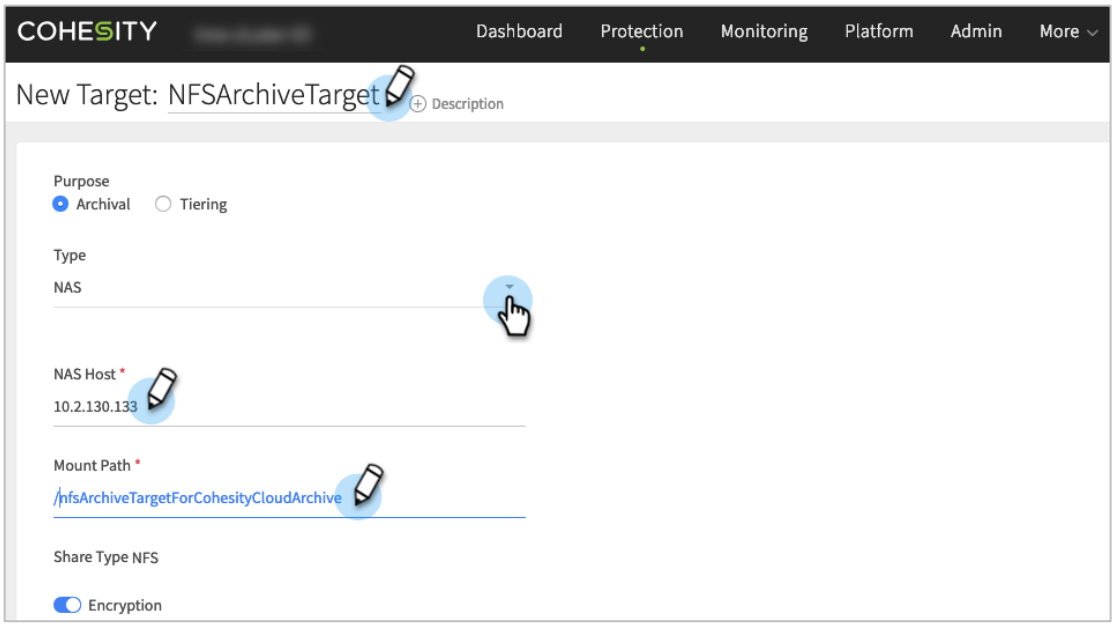
To register an External Target with your cluster:

1. Log in to Cohesity Platform.
2. Click **Protection** > **External Target**. Then click **Register External Target**.



3. In the form that opens:
   c) Enter a unique target Name.
   d) Select **Purpose** (Archival).
   e) Select **Type** (NAS).
   f) Enter the **NAS Host** IP address or FQDN.

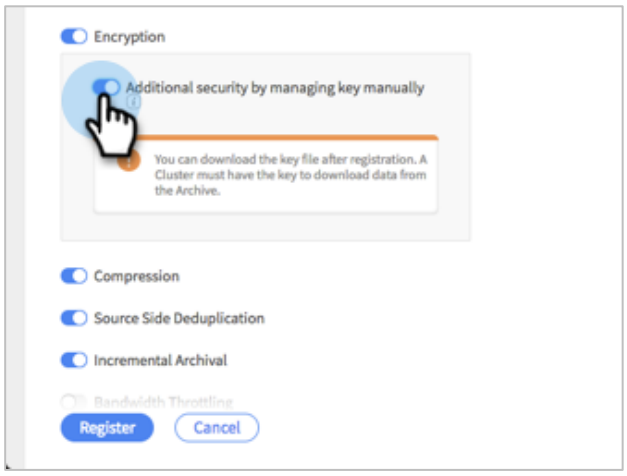
4. Enter the **Mount Path**.



NOTE: Do not include 'http' or 'https' in the mount path.

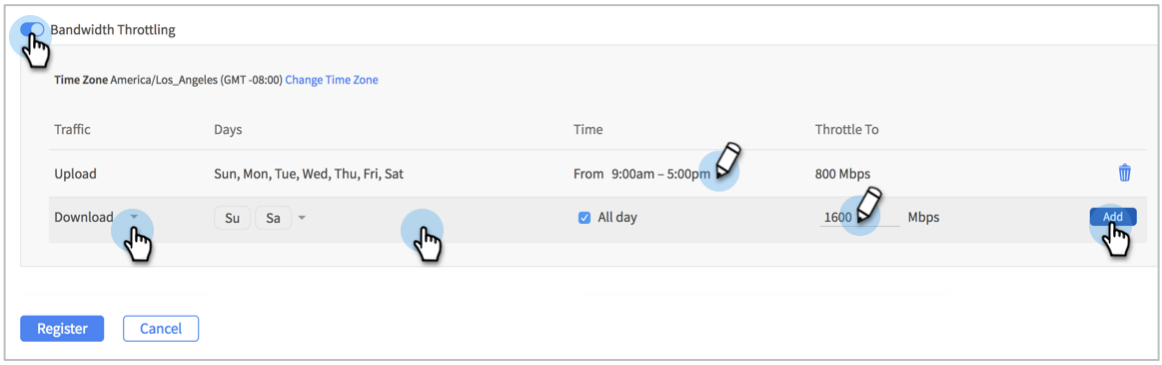
5. On the same screen, check your default settings. By default, **Encryption**, **Compression**, **Source Side Deduplication**, and **Incremental Archival** are enabled, while **Additional security by managing key manually** and **Bandwidth Throttling** are disabled.

a) If you want to enable manual key management for extra security, turn it on it here:



IMPORTANT: With this option on, a cluster must have the key to access data from the archive. You can download the key file (only once) after you register your NFS mount point. This key is required when you use CloudRetrieve. If you do not have it, you will still be able to recover data to its original cluster, but you will not be able to retrieve it onto a new cluster (in a disaster-recovery scenario, for example).

Enable Bandwidth Throttling if needed. You can throttle upload and download speeds separately, and apply throttling all the time or only specific days and times.



NOTE: For more on Encryption, Compression, Source Side Deduplication, Incremental Archival, and Bandwidth Throttling, see Create, Register Cloud Object Storage above.

6. Click **Register**.

Your NFS mount point is now an External Target in the Cohesity cluster, and is available to select when you create a Cohesity Protection Policy for use in Protection Jobs.

IMPORTANT: Customers should never manually edit, change, or delete Cohesity Platform archives directly in NAS.

Create a Protection Policy

In Cohesity, Protection Jobs use Protection Policies. Protection Policies reflect business needs, such as backup and archival frequency, retention requirements for each, as well as Recovery Point Objectives and Recovery Time Objectives, while a Protection Job defines operational requirements, such as which source objects to protect, the Protection Policy to use, and operational considerations like indexing, exclusions and inclusions, alerts, app consistency, and more. This process of combining business needs (Policy) with the objects to protect (source data) and the operational requirements (Job) provides rich flexibility to customers.

A Protection Policy defines:

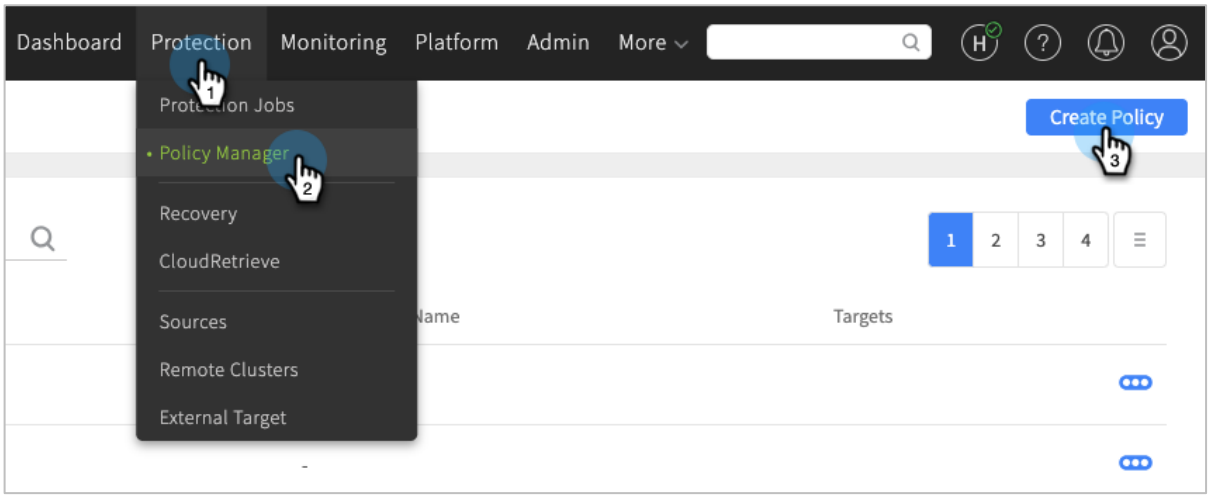
- How source data (like virtual and physical servers, databases, unstructured data, etc.) will be backed up and then archived.
- Where and how frequently they will be archived.
- How long the archives will be retained.

This list addresses parameters that affect CloudArchive operations. For the complete list of Protection Policy parameters, see Create or Edit a Policy in the online Help.

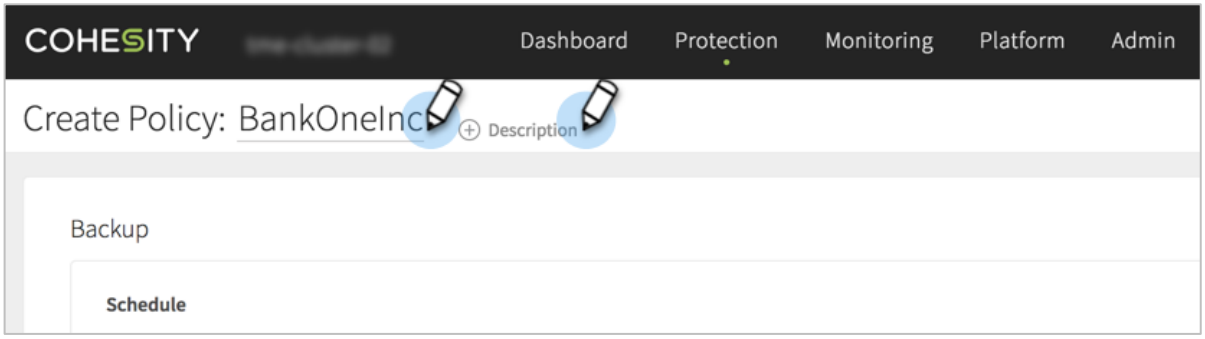
In the Protection Policy, you can select the cloud-based External Target you just created and registered as an External Target.

To create a Protection Policy:

1. Log in to Cohesity Platform.
2. Click **Protection** > **Policy Manager**. Then click **Create Policy**.

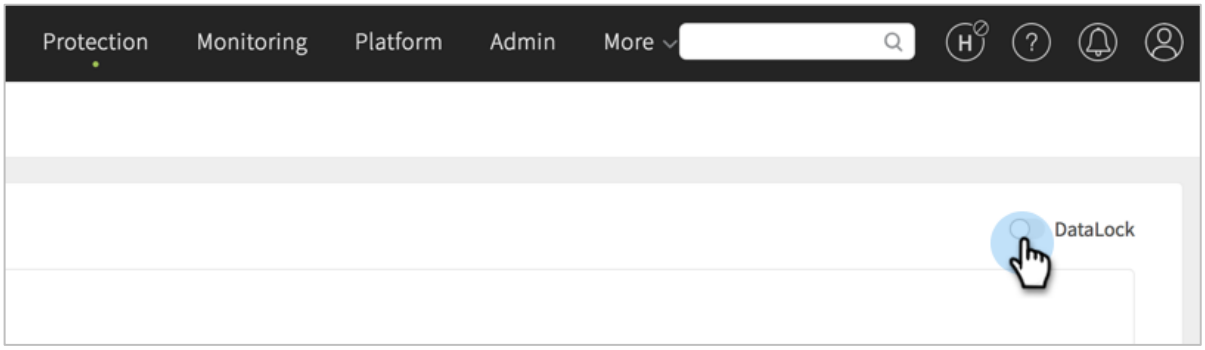


3. In the form that opens, complete the rest of these steps. Add a name, and optionally a **Description**.



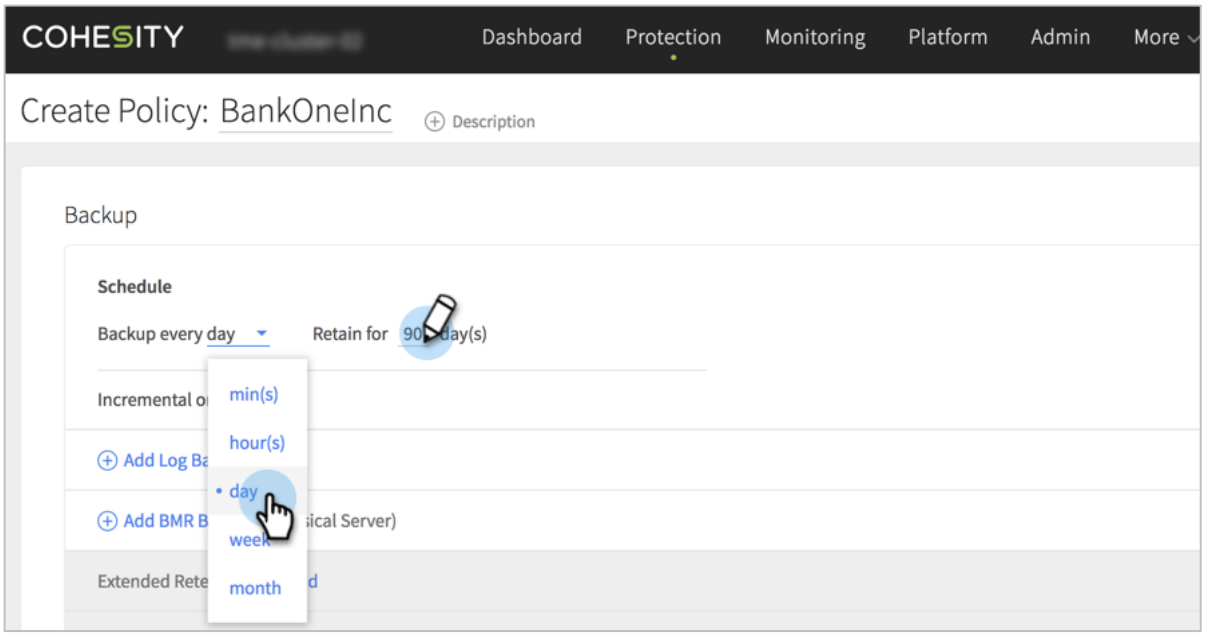
4. Add a **DataLock** for compliance and regulatory requirements, to ensure that your protected data, including local backups, archives, and replication, cannot be modified until the DataLock expiration.

   Once applied, a DataLocked Snapshot will be deleted only after its retention period expires. A DataLock prevents all users, including those who have the Data Security role in Cohesity Platform, from modifying or deleting any Snapshots that were generated by the Protection Jobs that use this policy. Only users with the Data Security role can add, modify, or remove a DataLock from a Policy. See online Help for more information.

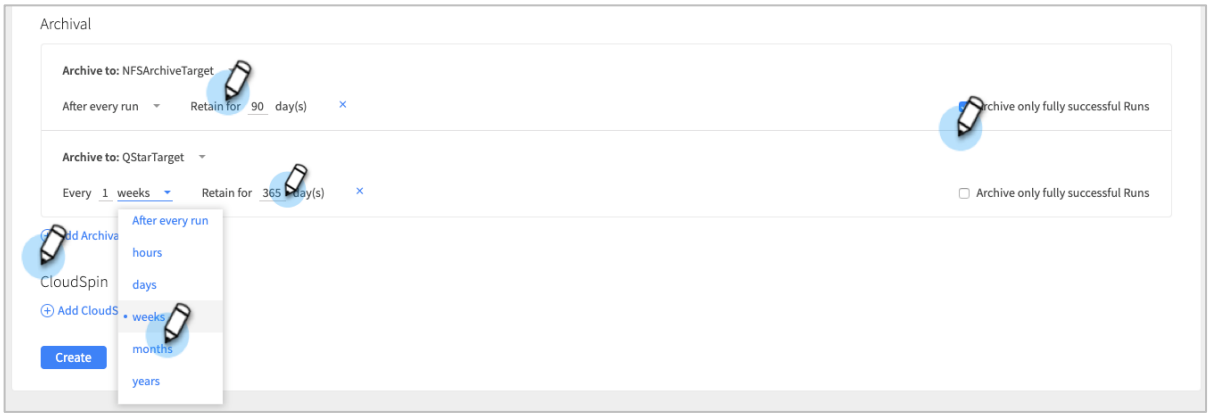


NOTE: You can also add a legal hold to a specific Protection Job run (a Snapshot) to preserve it for legal reasons. See Apply Legal Hold to Completed Job Run below.

5. Under **Backup**, set the Backup interval (every day, by default) and **Retain for** period (90 days, by default) for your local cluster.



6. Under **Archival**, click **Add Archival** and for **Archive to**, select the External Target you just created. Set the Archival interval (every day, by default) and **Retain for** period (90 days, by default). You can also enable **Archive only fully successful Runs** on the right.



Click **Add Archival** again if you need additional archival schedules.

NOTE: You can add multiple archival schedules that use the same or different External Targets, as well as the same or different intervals and retention periods, to a given Protection Policy. When you add more schedules and send them to the same External Target with different retention and schedule times, the schedules rationalize among themselves and only the necessary archive is sent, with the longest retention.

For example, if you add these three archival schedules to the same External Target:

- Once a day, retain for 90 days.
- Once every 7 days, retain for 180 days.
- Once every 30 days, retain for 365 days.

Then:

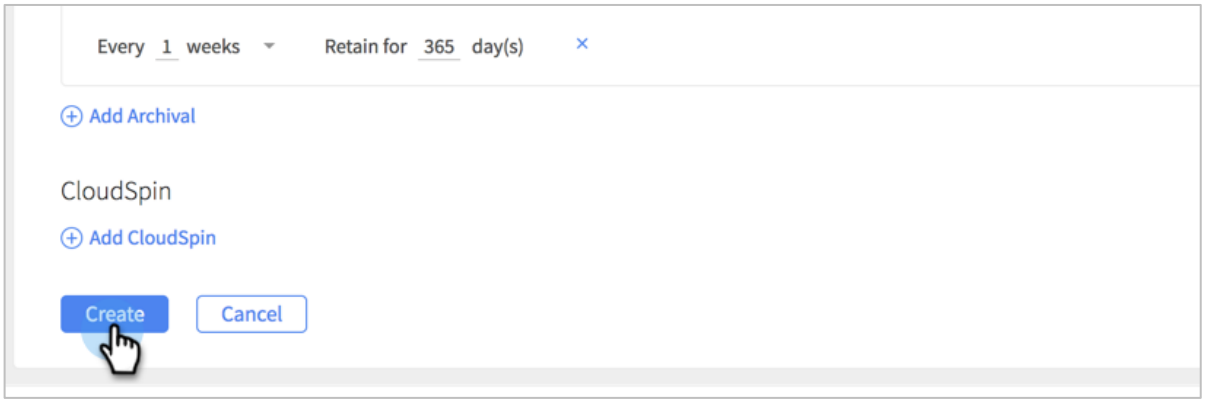
- On Day 7, only one archive is sent, meeting both Schedule 1 and Schedule 2 (and retained for 180 days, per Schedule 2, as it is the longer of the two).
- On Day 30, only one archive is sent, meeting both Schedule 1 and Schedule 3, but is retained for 365 days, to meet the Schedule 3 retention requirement.

By contrast, if you send the archives to different External Targets, then:

- On Day 7, per Schedule 1, the archive is sent to the first External Target and retained for 90 days. Per Schedule 2, the archive is also sent to the second External Target and retained for 180 days.
- On Day 30, per Schedule 1, the archive is sent to the first External Target and retained for 90 days. Per Schedule 3, the archive is also sent to the third External Target and retained for 365 days.

When you use multiple schedules with different External Targets, the schedules don't rationalize, and you accrue network and storage usage for each scheduled run.

7. Click **Create**.



Your new Policy can now be used in Protection Jobs. For the complete list of Protection Policy parameters, see online Help.

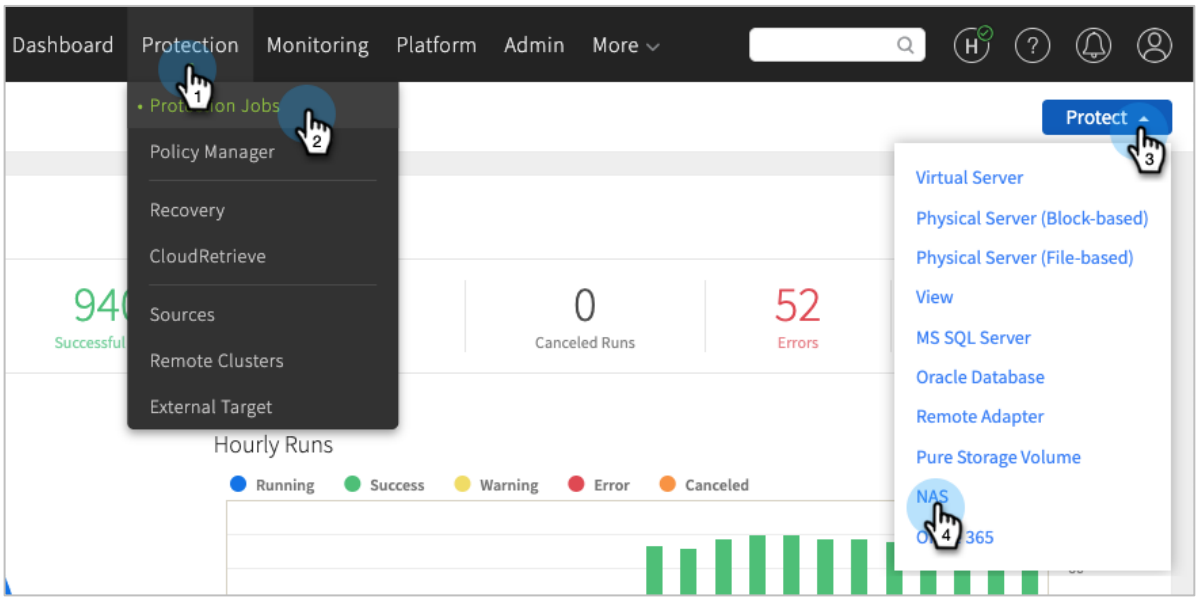
Create a Protection Job

Protection Jobs combine operational requirements with the business requirements that are defined in a Protection Policy. Multiple Protection Jobs can use the same Protection Policy, but each Job can have only one Policy. Protection Jobs protect specific source objects, such as virtual servers, physical servers, Views, SQL servers, Oracle databases, Remote adapters, Pure Storage Volumes, or network-attached storage (NAS).

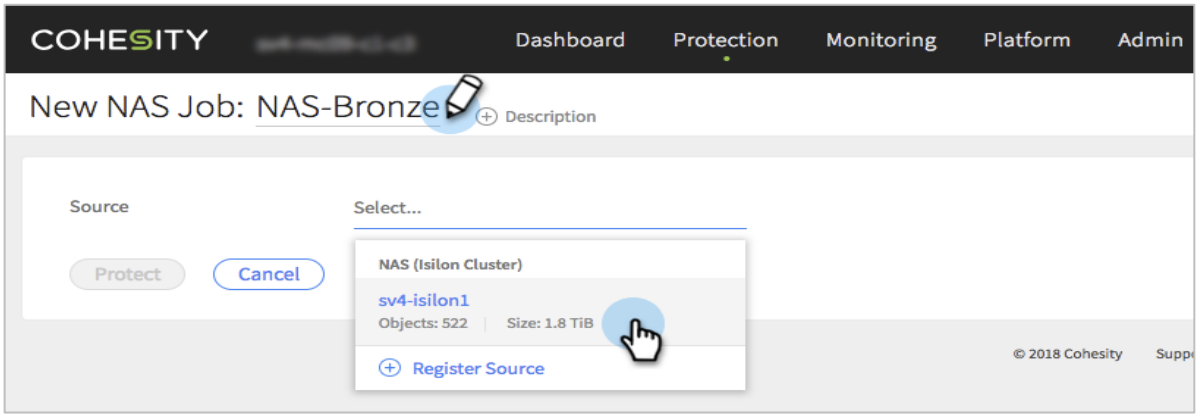
For this example, we look at the steps to create a Protection Job for NAS data, but the steps to protect other source objects are very similar.

To create a Protection Job:

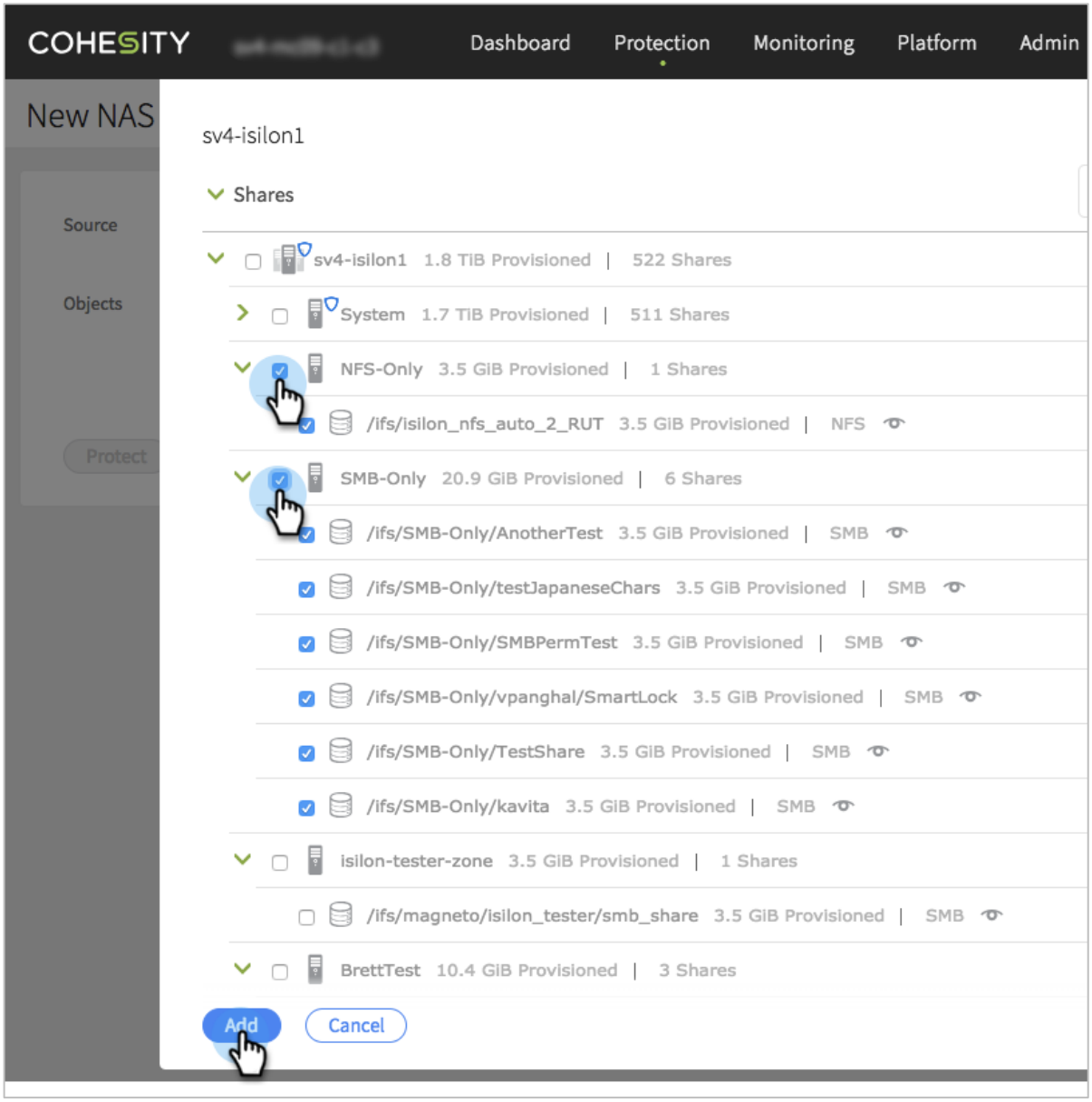
1. Log in to Cohesity Platform.
2. Click **Protection** > **Protection Jobs**. Then click **Protect** and choose the type of data to protect.



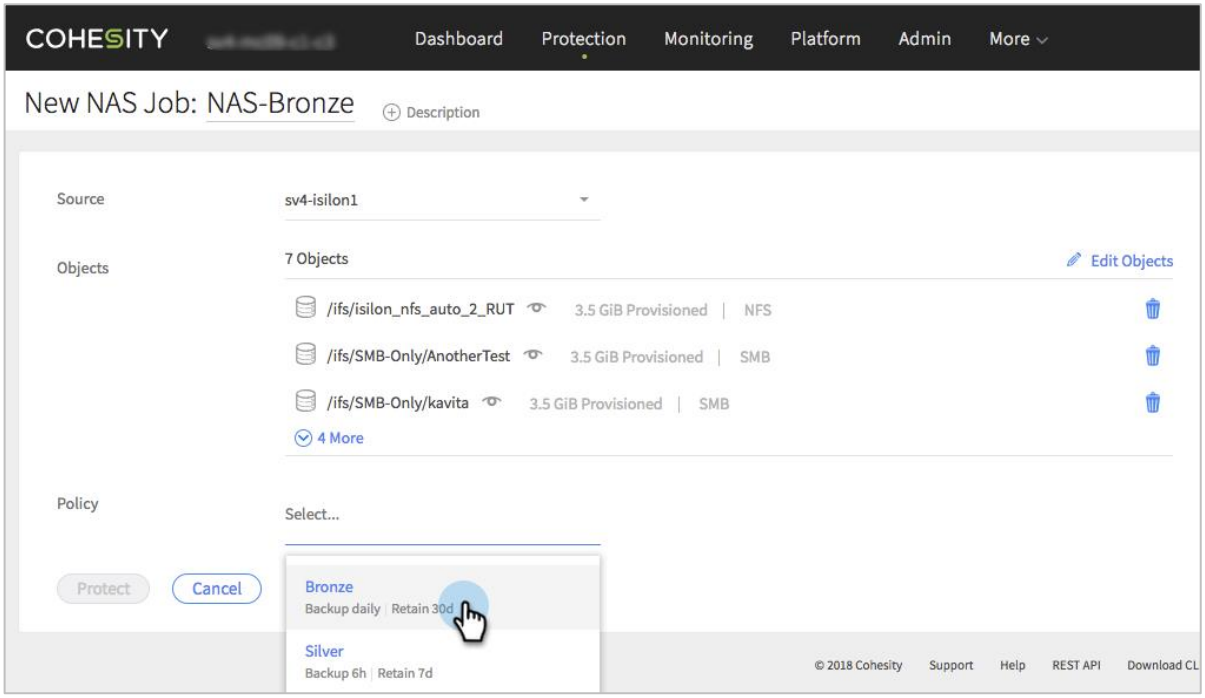
3. Name your Protection Job and select a **Source** to protect.



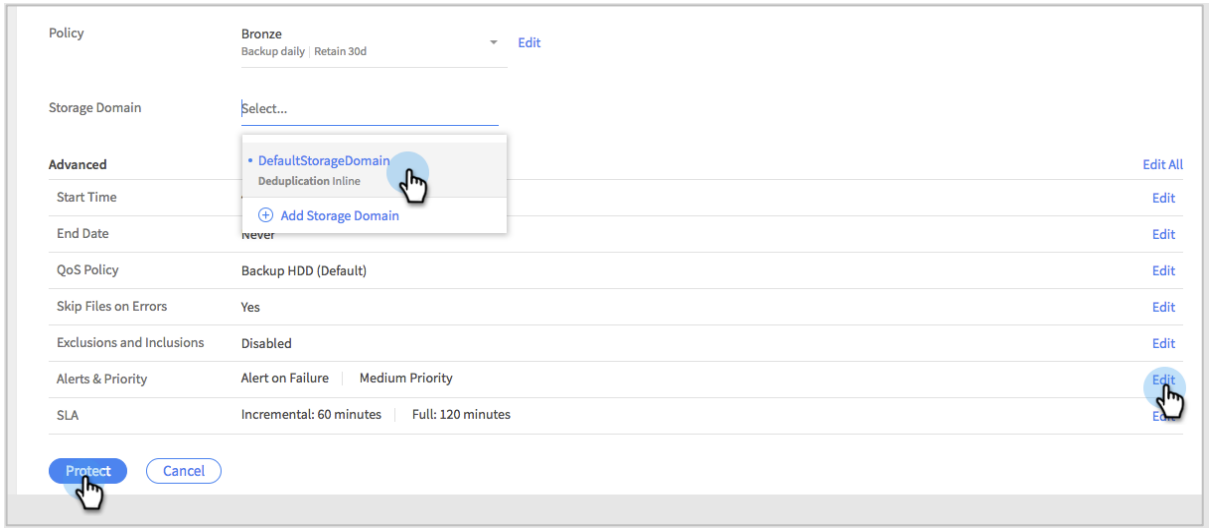
4. Select the specific objects you wish to protect with this Job, then click **Add**.



5. Select a **Policy**.

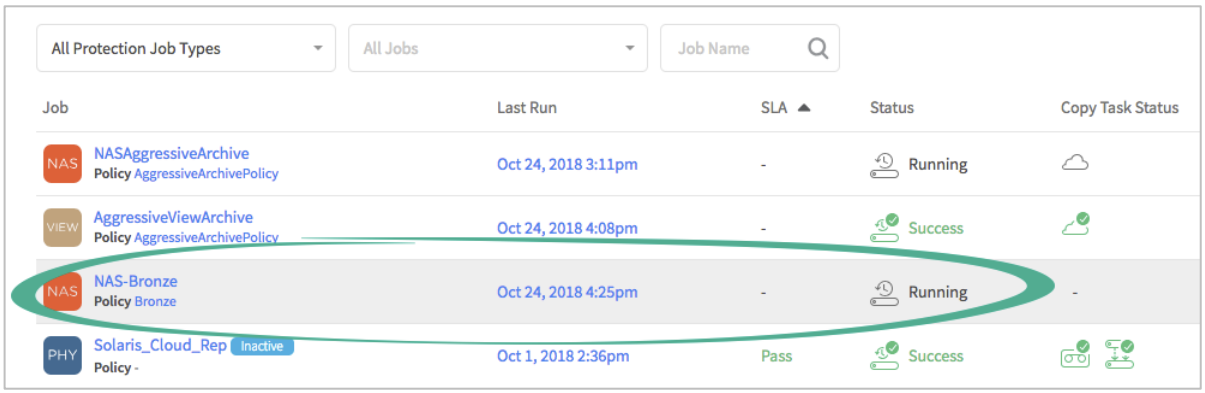


6. On the same screen, select a **Storage Domain**. If you need to change any of the **Advanced** settings, click **Edit** on the right. When you're done, click **Protect**.



NOTE: See the complete list of **Advanced** settings and the Job types that contain them in the Appendix.

7. In the top navigation, select **Protection** > **Protection Jobs** to verify that your new Job is in the list.



Your new Protection Job is now active and running. To manage Protection Jobs, see online Help.

Apply Legal Hold to Completed Job Run

Only users who are assigned the Data Security role can put a legal hold on existing Snapshots (Protection Job runs), to preserve them for legal purposes. Once a legal hold is applied, the retention period is ignored, and the Snapshot is preserved until the legal hold is removed. Legal hold Snapshots can only be deleted by a user with the Data Security role.

> **NOTE**: A legal hold can be added to both regular and DataLocked Snapshots.

You can add a legal hold to a Protection Job run or to individual objects in a Job run:

- If you add a legal hold to a Job Run, it applies to all the Snapshot objects that were backed up by that Job Run, and the legal hold is propagated to replicated and archived objects.
- If you add a legal hold only to selected objects in a Job Run, the legal hold is propagated to archived objects, but not to the replicated objects. You must manage the legal hold status on the remote replication cluster manually.

> **NOTE**: A legal hold prevents Snapshots from being deleted until the legal hold is removed. Using a legal hold for long periods of time can result in the cluster running out of space.

To add or remove a legal hold from a Protection Job Run, see Adding a Legal Hold to a Snapshot in the online Help.

The Difference between Legal Hold and DataLock

While both a legal hold and DataLock are features that empower the Data Security role in Cohesity Platform to prevent backed up and archived data from being deleted, the differ in purpose and function.

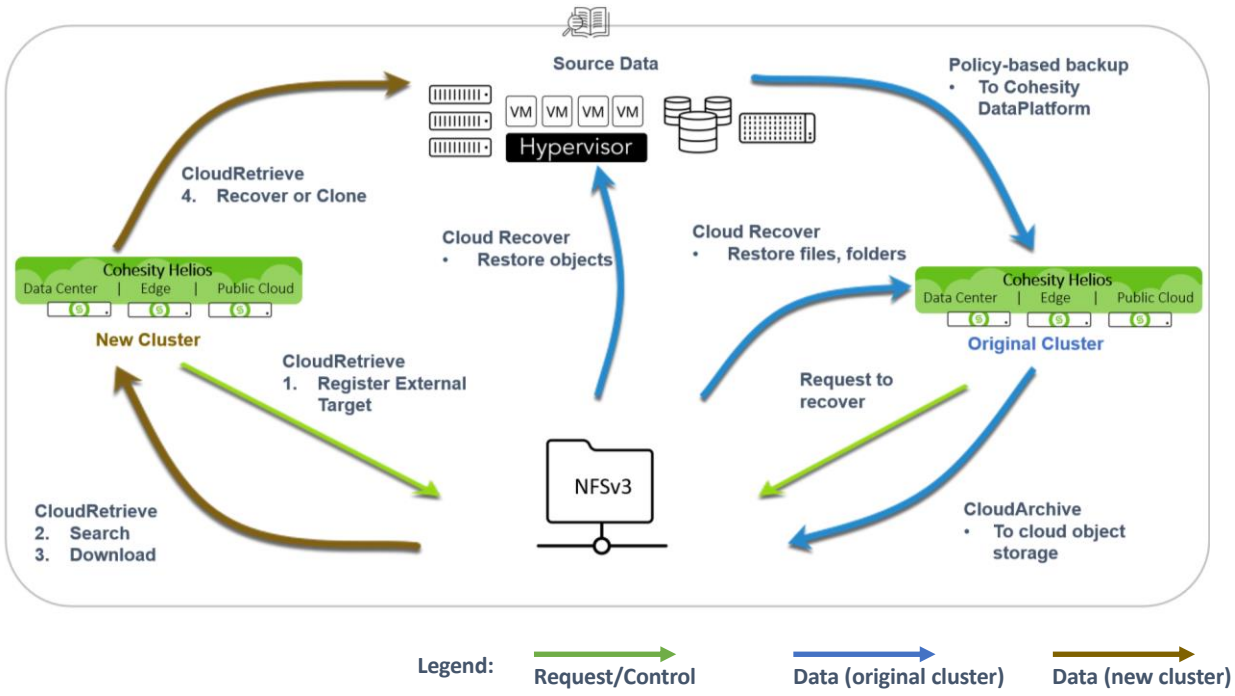
Table 5: The Difference between Legal Hold and DataLock

PURPOSE	LEGAL HOLD	DATALOCK
Business need	Reactive: Set on a specific Snapshot (i.e., Job Run), usually prompted by legal requirements.	Planned: Set on all Job Runs that use a Protection Policy with DataLock, usually for compliance.
Expiration period	No expiration. Removal managed by the user.	Defined in the Protection Policy.
Granularity	Set on individual Job Runs and at the Object Level.	Applies to all Job Runs of any Protection Jobs that use a Policy with DataLock.
Deletion	Can be deleted to recover storage space, but only by a user with the Data Security role.	Cannot be deleted before the DataLock expiration date, even by a user with the Data Security role.

Recover Data from CloudArchive

Cohesity Platform provides two ways to get your data back from storage: Cloud Recover and CloudRetrieve.

Figure 8: Cohesity Protection, CloudArchive, Cloud Recover, and CloudRetrieve



- Cloud Recover: Recover entire objects (such as VMs, databases, NAS, etc.) or individual files and folders back onto the Cohesity Platform that archived them.

  NOTE: When you recover a complete object (such as a VM or database), it is restored to its original location once it is downloaded to the Cohesity Platform from the cloud, and restored via the Instant Volume Mounting capability in Cohesity Platform.

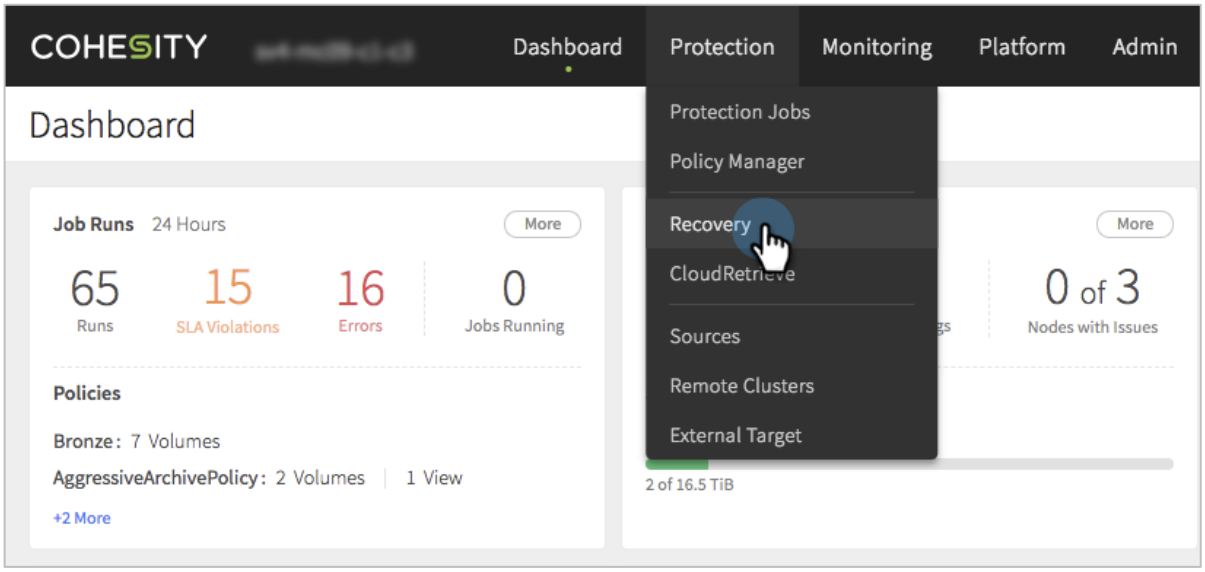
- CloudRetrieve: CloudRetrieve allows you to extract your Protection Job and its metadata, including Job Run details, from the archive in the cloud, so you can search it and recover the data you need onto a new or different cluster. This approach involves several steps:
  a) Register the External Target containing your archived data.
  b) Search the archive in the cloud.
  c) Select and download metadata for the archived Protection Jobs.
  d) Recover objects from the downloaded Protection Job Run.

But first, let's start with recovering data onto your original Cohesity cluster.

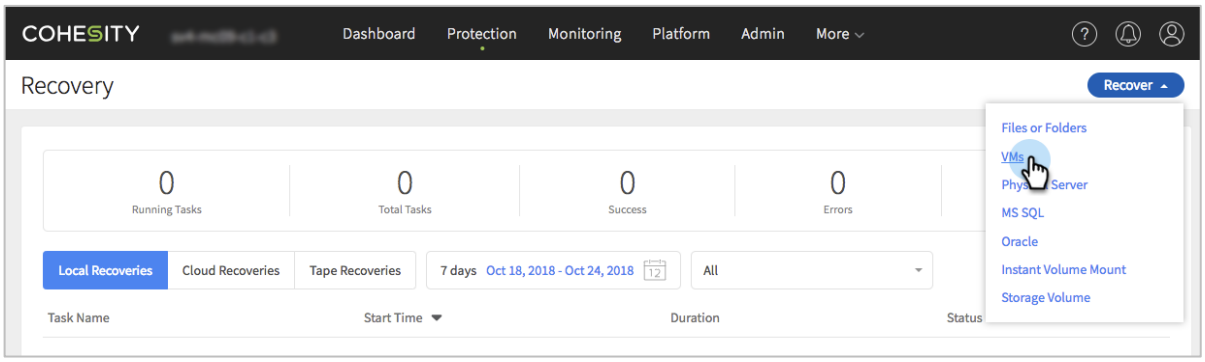
Recover Your Data to Original Cluster

To locate and recover a file, a folder, or an entire virtual machine to the original cluster:

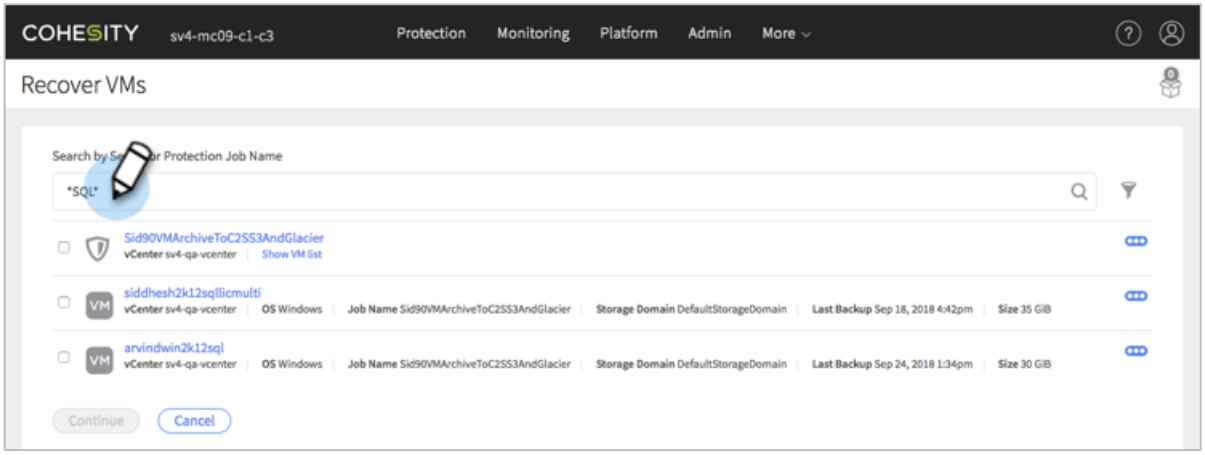
1. Log in to Cohesity Platform.
2. Select **Protection** > **Recovery**.



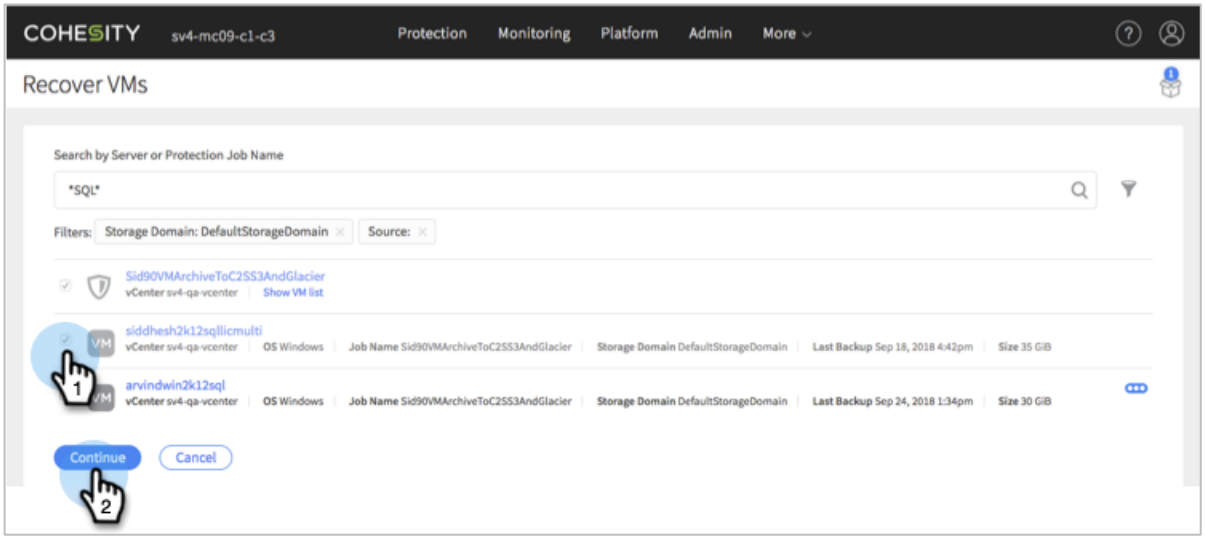
3. Click **Recover** and select the type of object you seek — a file or folder, VMs, physical server, and more.



4. To retrieve a list of virtual machines, for example, select VMs and enter part or all of the VM names:



5. Select the VMs you need, or select an entire Protection Job to recover all the VMs it archived, and then click **Continue**.



6. Edit the **Task Name** and **Recover as** fields, if necessary. To rename the VM, choose **Rename Recovered VMs** and add the appropriate **Prefix** and **Suffix**. Choose a **Recovery Location**, select **Networking Options** and **Additional Options**.

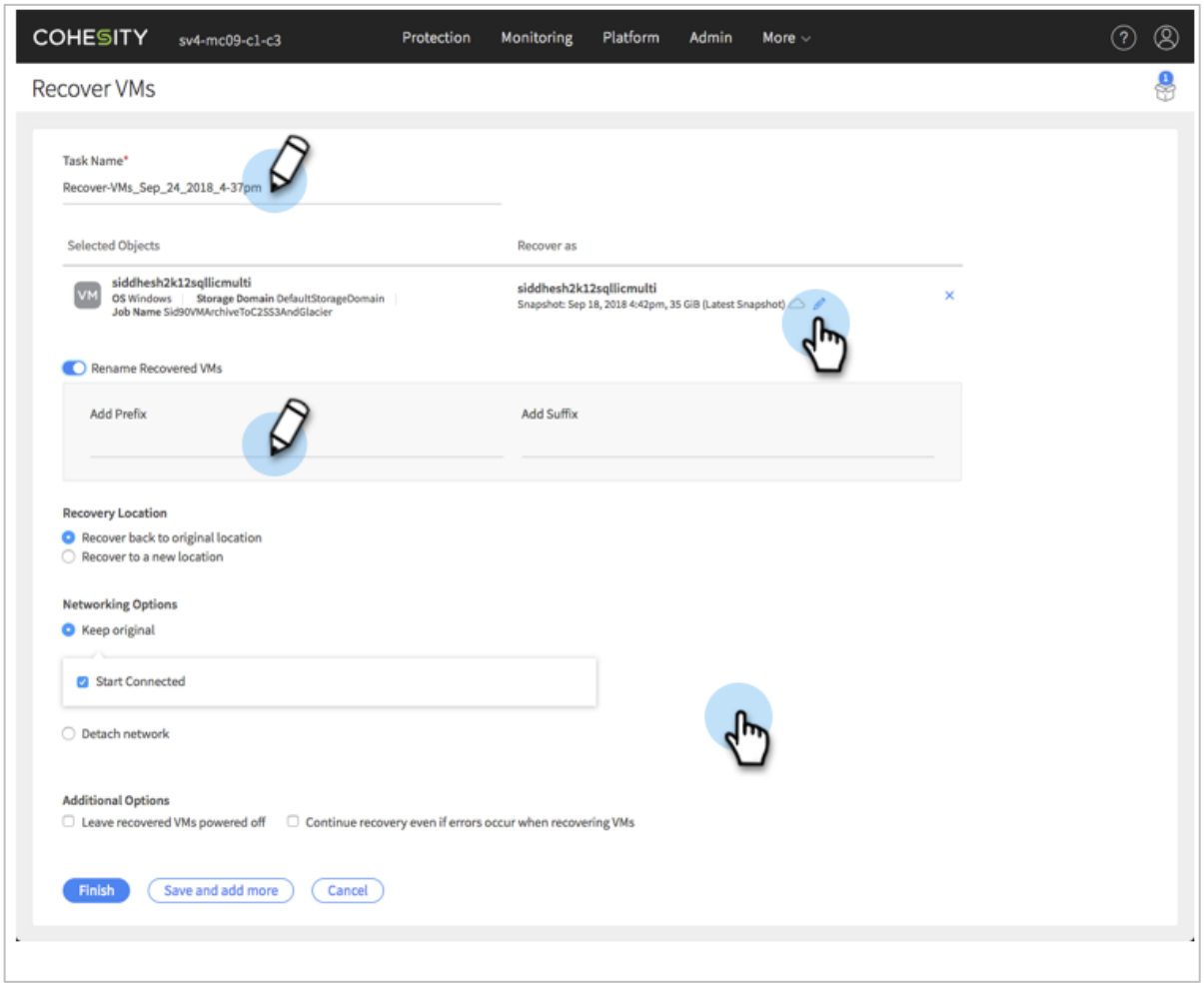


Table 6: Recover Task Options

RECOVERY OPTIONS	DETAILS
Recover to a new location	Specify this option to recover the VM files (such as the VMDK files) to their original datastores and create new instances of the VMs in the original location in the original source. For more, see Recover to Original Location in the online Help.
Keep original	For each recovered VM, keep the original virtual Network Interface Cards (vNICs) and attach them to the original network connections. **NOTE**: This option is only supported when VMs are recovered back to their original location.
Start Connected	For each recovered VM, connect to the original or new network when the VM reboots. **IMPORTANT**: If this option is not selected, the VMs are not connected to any network on reboot.
Detach network	For each recovered VM, the virtual Network Interface Card (vNIC) is removed from the VM.
Leave recovered VMs powered off	The recovered VMs remain powered off after they are created. **TIP**: Cohesity recommends this option if you are recovering from a storage domain that has CloudTier enabled.
Continue recovery even if errors occur when receiving VMs	With this option, if one of the VMs cannot be created, Cohesity will still attempt to create the other VMs.

NOTE: This example is for recovering a VM. The recovery options vary by Protection Job type.

7. Click **Finish** to start the recovery process.

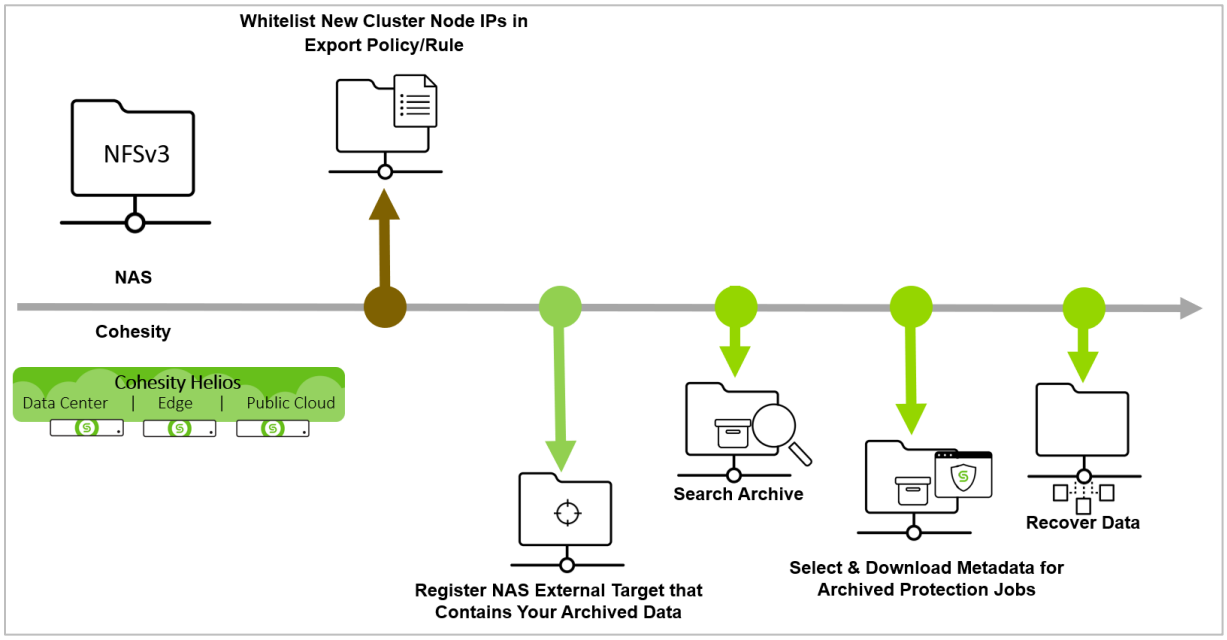
For more on the many capabilities and choices in our recovery process, see Recovery in the online Help.

CloudRetrieve Your Data to New Cluster

CloudRetrieve provides the ability to download data that was archived from a cluster to an alternate (non-original) cluster. In other words, you have Cluster A, which archives data to an External Target, but you need to download that archived data to Cluster B, for geo-redundancy or disaster recovery.

When you need to recover data from NAS to a different Cohesity cluster, there are several steps:

Figure 9: CloudRetrieve Workflow



The sections below describe the steps to:

1. Write a new export policy (rule) that whitelists the node IPs (or IP subnet) of your new Cohesity cluster.
2. Register the External Target containing your archived data to the new cluster.
3. Enter the retrieve parameters (cluster name, date range, Protection Job name) to search the archive in the cloud. (The search can take from minutes to several hours, depending the data-retrieval SLA for the class of storage you are using from your cloud provider. For example, some storage classes have a standard retrieval SLA of up to several hours.)

   NOTE: If your External Target is protected by a manually managed key, before you can search it, you will need to upload the External Target's access key.

4. From your search results, select and download the metadata (the Job Run details) for the archived Protection Jobs onto the new cluster, so that you can review Job Run details and choose just the specific you need to recover or clone.

   NOTE: In this step, you are prompted to select a date range, and if you know exactly which Job Run (Snapshot) you need, you can also choose to download it along with the metadata, to be able to recover your data objects as soon as it completes.

5. After the metadata download completes, select the necessary Job Run from the archived Protection Job to recover or clone your objects.

Register External Target Containing Archived Data

To register your NFS mount point as an External Target on the new cluster:

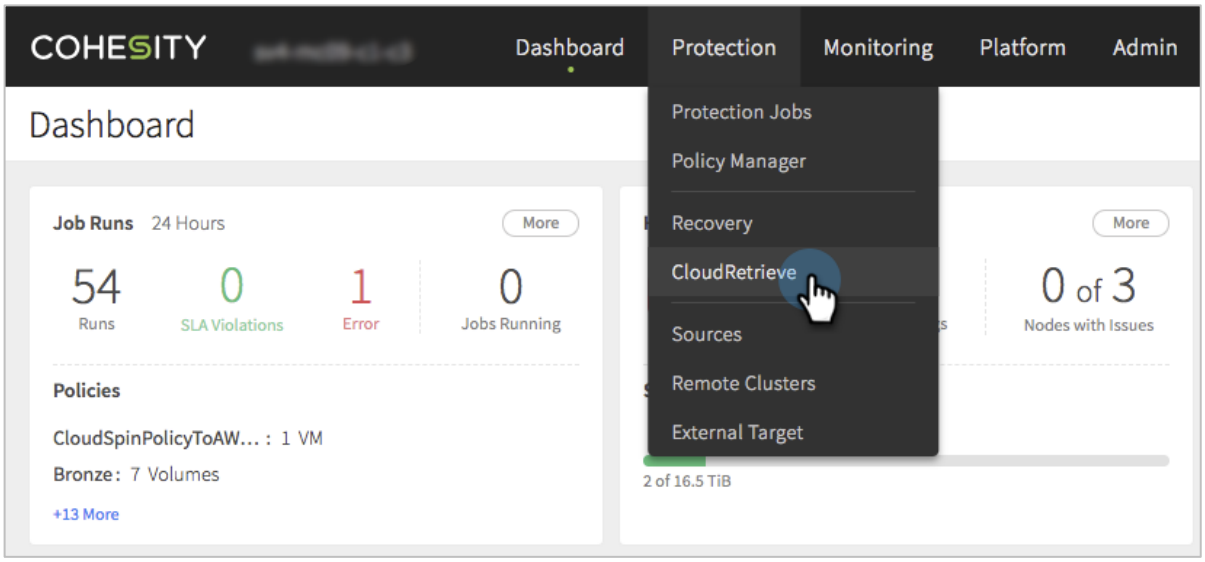
1. Log in to a cluster other than the cluster that archived your data, or stand up a new cluster.
2. Log in to Cohesity Platform on your new cluster.

3. Follow the steps in Register NFS Mount Point as External Target above to register your archived data from NAS to the new cluster.

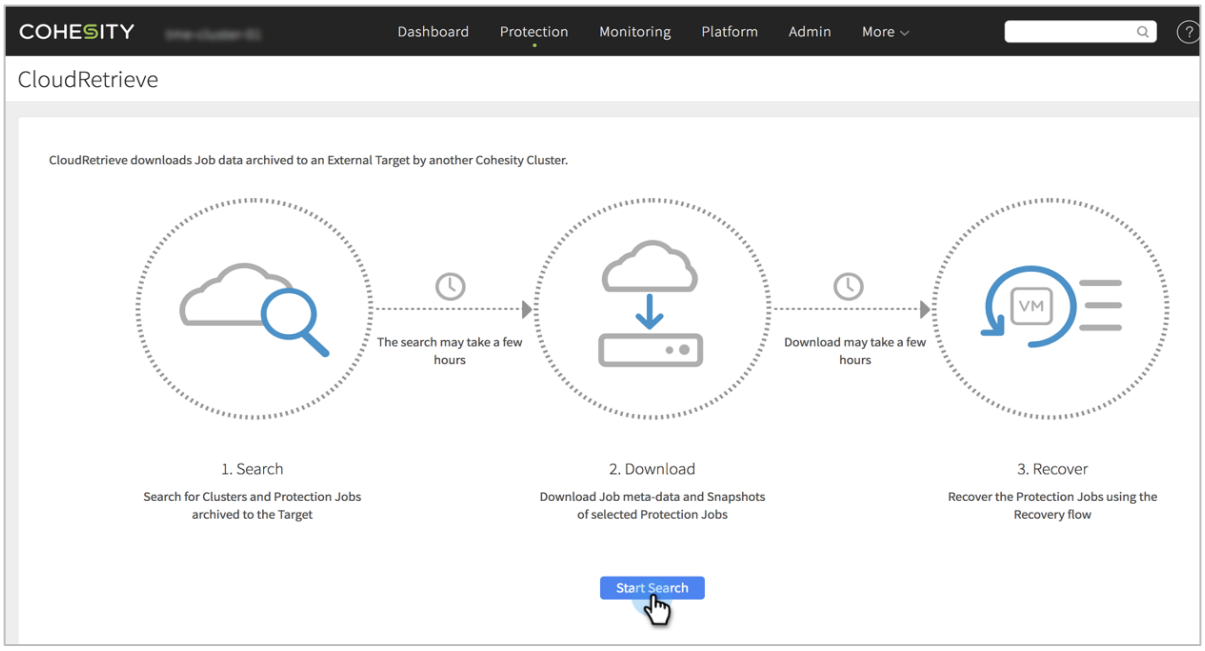
Search Archived Data in the Cloud

To submit a search request for a list of archived clusters and Protection Jobs:

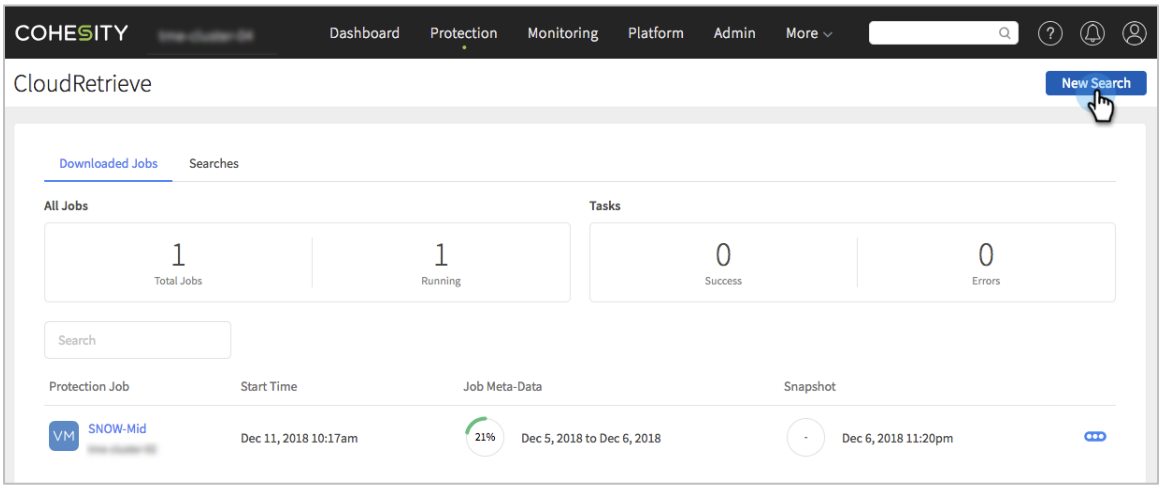
1. Log in to Cohesity Platform on the new cluster.
2. Select **Protection** > **CloudRetrieve**.



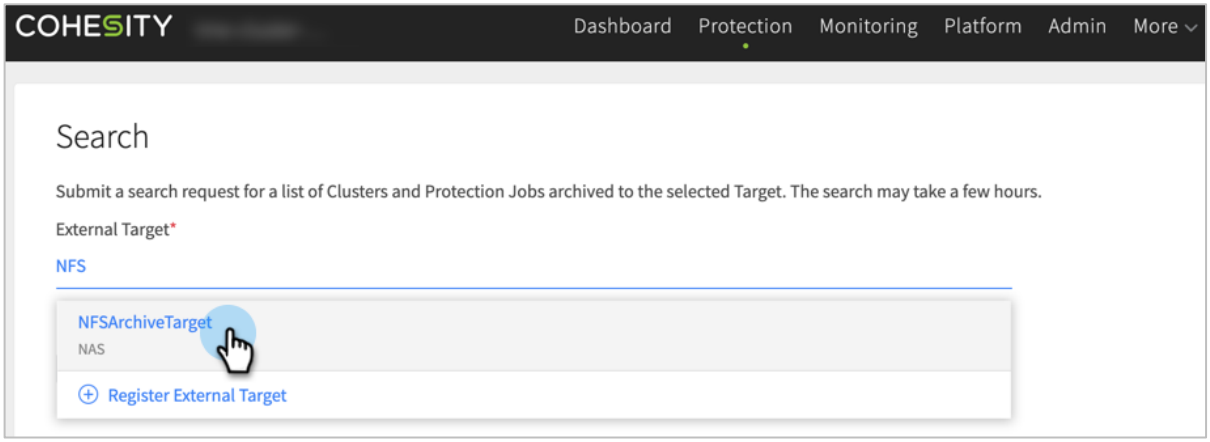
3. If this is the first time you have used CloudRetrieve, the CloudRetrieve summary screen appears. Click **Start Search**.



a) If this is not your first visit, the list of downloaded Jobs appears. In that case, click **New Search**.



4. Select your **External Target** from the drop-down list.



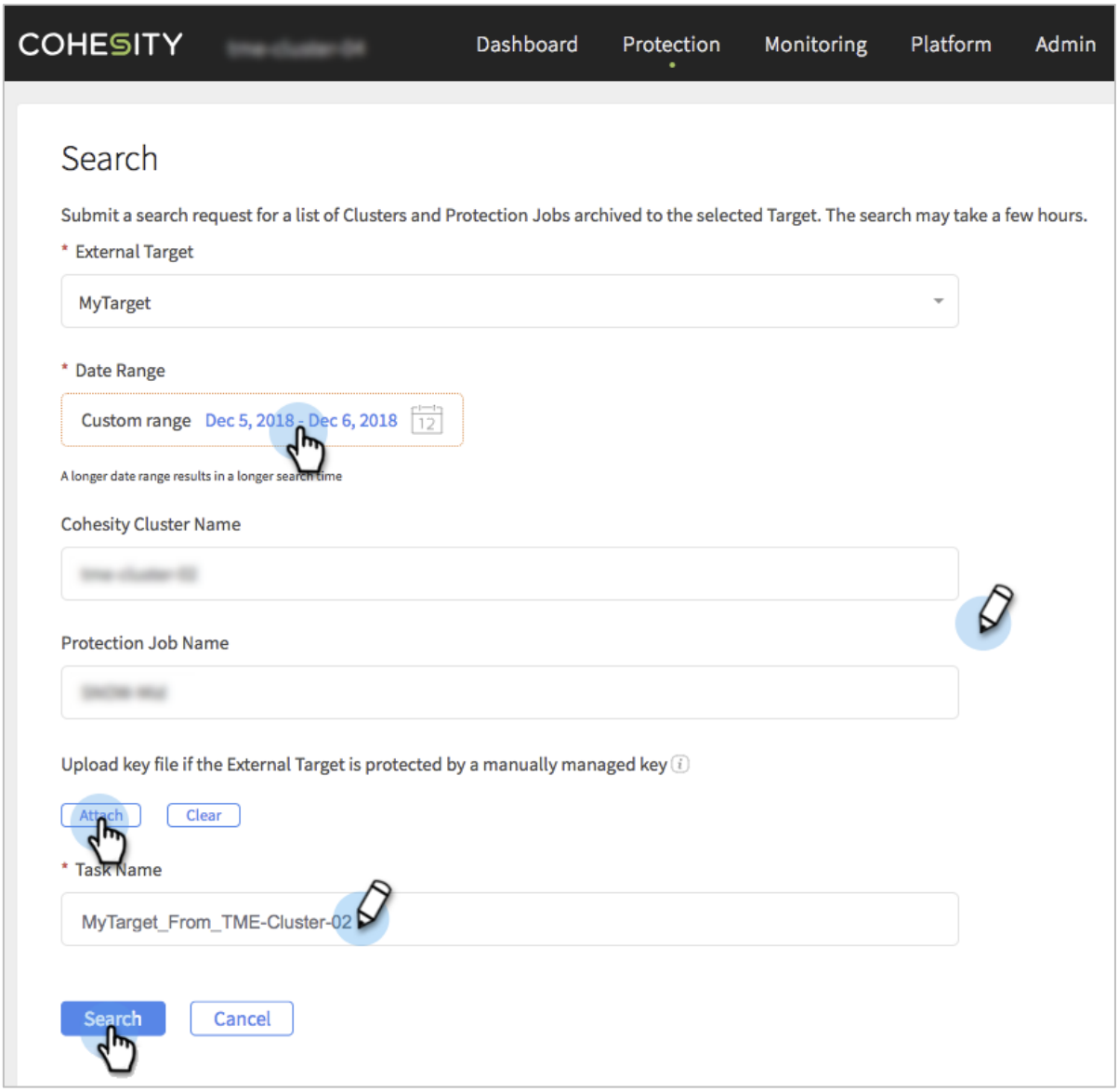
NOTE: If you skipped the first step and have not yet registered your External Target, you can register it here. To do so, click Register External Target from the drop-down menu and follow the steps in Register NFS Mount Point with Cohesity Platform above.

5. In the form that opens, enter:

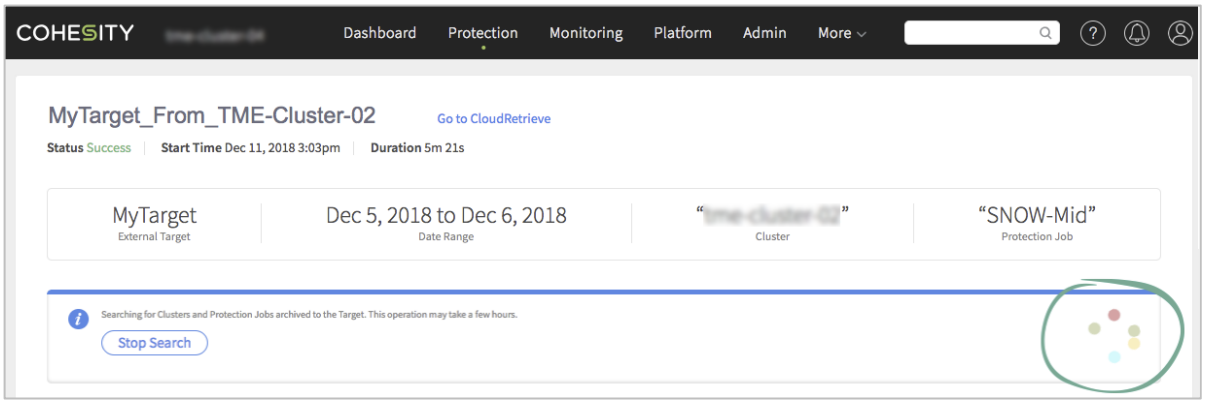
Table 7: CloudRetrieve Search Options

FIELD	DESCRIPTION	NOTES
Date Range (required)	Select a Date Range (past year by default) to limit the scope of your search.	
Cohesity Cluster Name (optional)	To narrow your search to a specific cluster, enter a cluster name. This is especially helpful if the same NFS mount point is used with more than one cluster. To broaden your search to match more than one cluster, use a partial name (for example, 'Acme' instead of 'Acme_Raleigh').	**IMPORTANT**: Wildcard characters (like '*****') are NOT supported. If you enter search terms for both **Cluster Name** *and* **Protection Job Name**, your search must find

FIELD	DESCRIPTION	NOTES
Protection Job Name (optional)	To narrow your search to a specific Protection Job, enter a Job name. This is especially helpful if the same NFS mount point is used for more than one Protection Job. To broaden your search to match more than one Protection Job, use a partial name (for example, 'NAS' instead of 'NAS-Bronze').	matches for the Protection Job *within* clusters that match. If your search is too narrow, try entering a search term for just **Cluster Name** *or* **Protection Job Name**, or leave one or both empty.
Upload key file (optional)	If your External Target is protected by a manually managed key, click **Attach**.	
Task Name (required)	By default, Cohesity uses the current timestamp to name the task automatically (for example, 'Cloud_search_***<CurrentTime>***'). Cohesity recommends you replace the automatic **Task Name** with terms that will make it easy to identify (for example, '***<ExternalTarget>***_From_***<SourceCluster>***_***<Purpose>***').	



6. Click **Search.**
7. Wait while the search runs.



> **NOTE**: The search can take from minutes to several hours, depending the data-retrieval SLA for the class of storage you are using from your cloud provider. For example, some storage classes have a standard retrieval SLA of up to several hours.

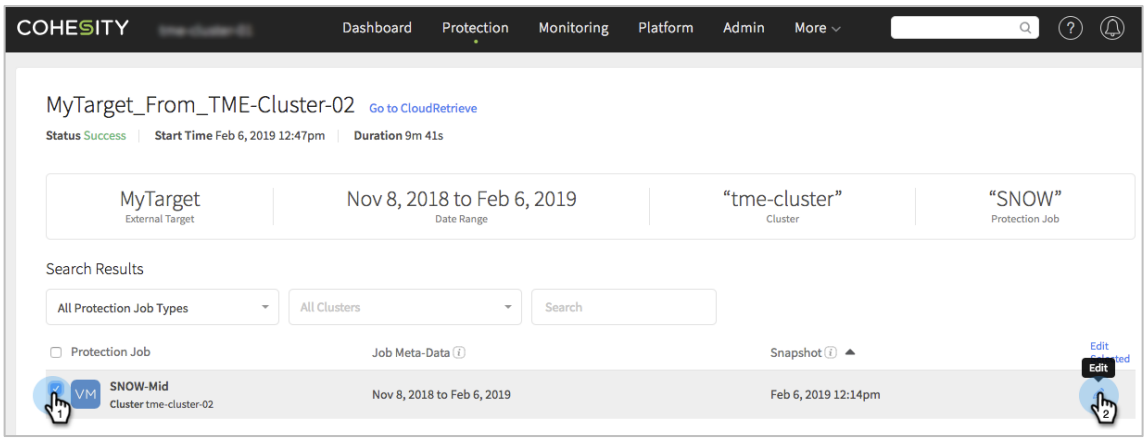
The success of a CloudRetrieve search does not guarantee that the search found any matches. It means only that the search operation completed successfully. If your search results came up empty, broaden your search with partial names for the cluster and/or Job, leave them blank, and/or extend the date range.

Select and Download Metadata for the Archived Protection Jobs

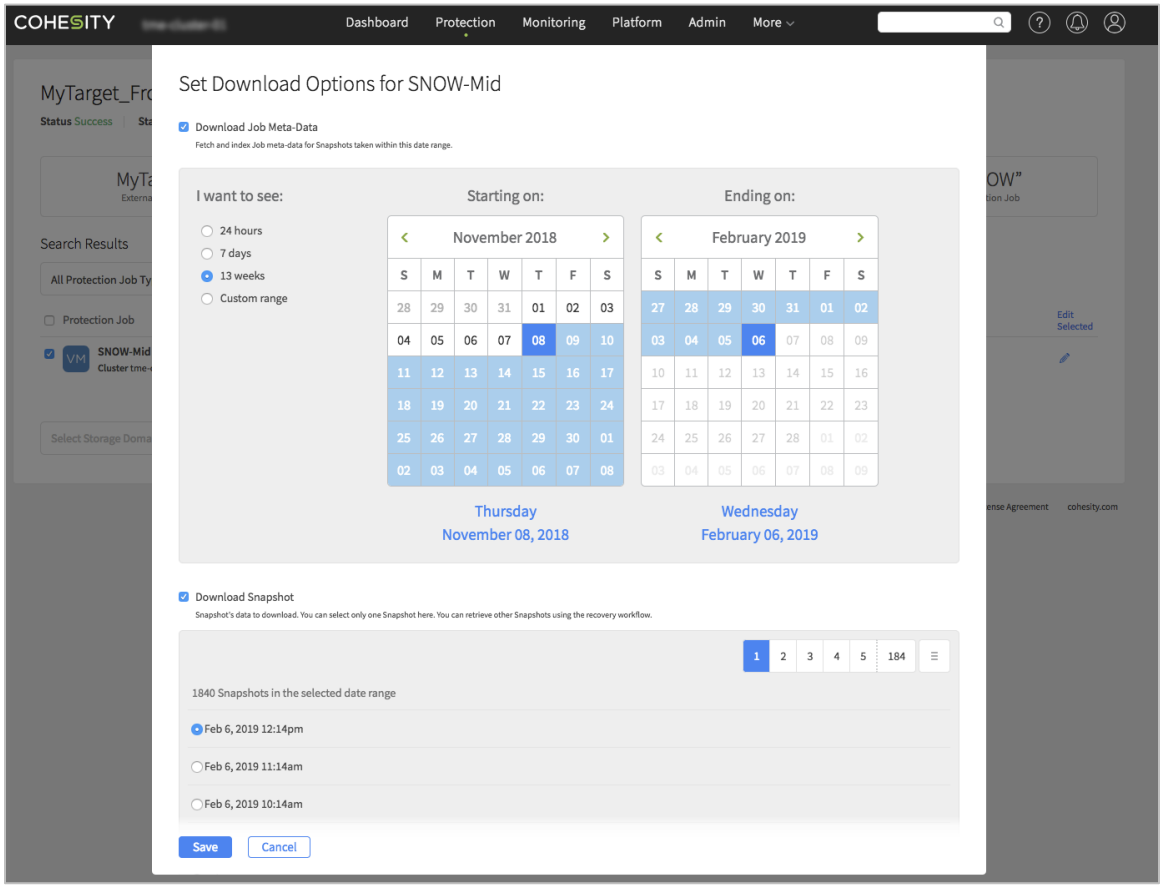
Once you have your search results, choose the Protection Jobs to download to your new cluster. After the download, you will be able recover your data from the downloaded archive. See Figure 8 above.

When your search completes:

1. Select the Protection Job(s) you wish to recover from the search results and click **Edit**.

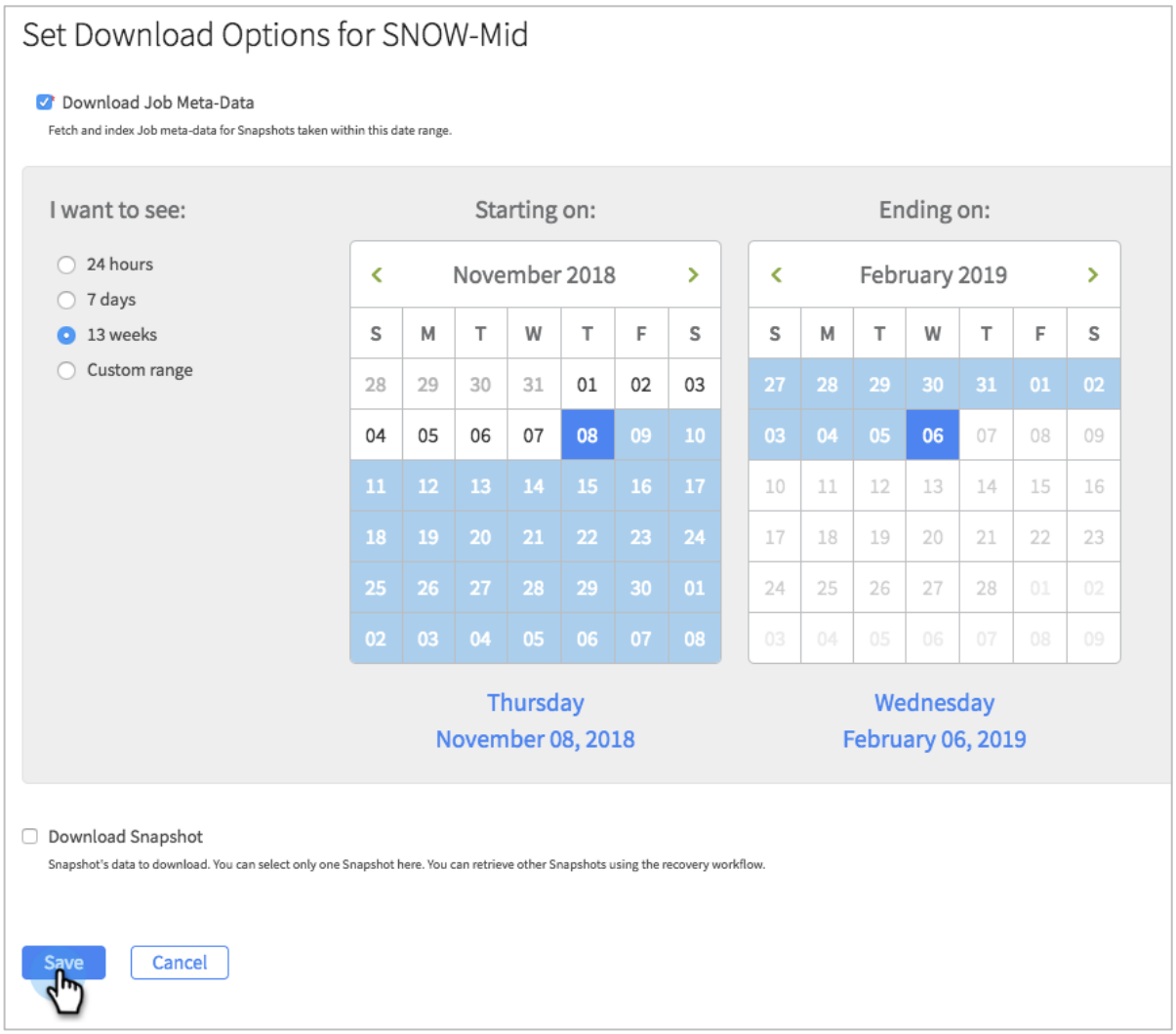


2. In the form that opens, you can choose to **Download Job Meta-Data** (that is, the details of each Job Run in the archived Protection Job), **Download Snapshot** (a specific Job Run), or both.

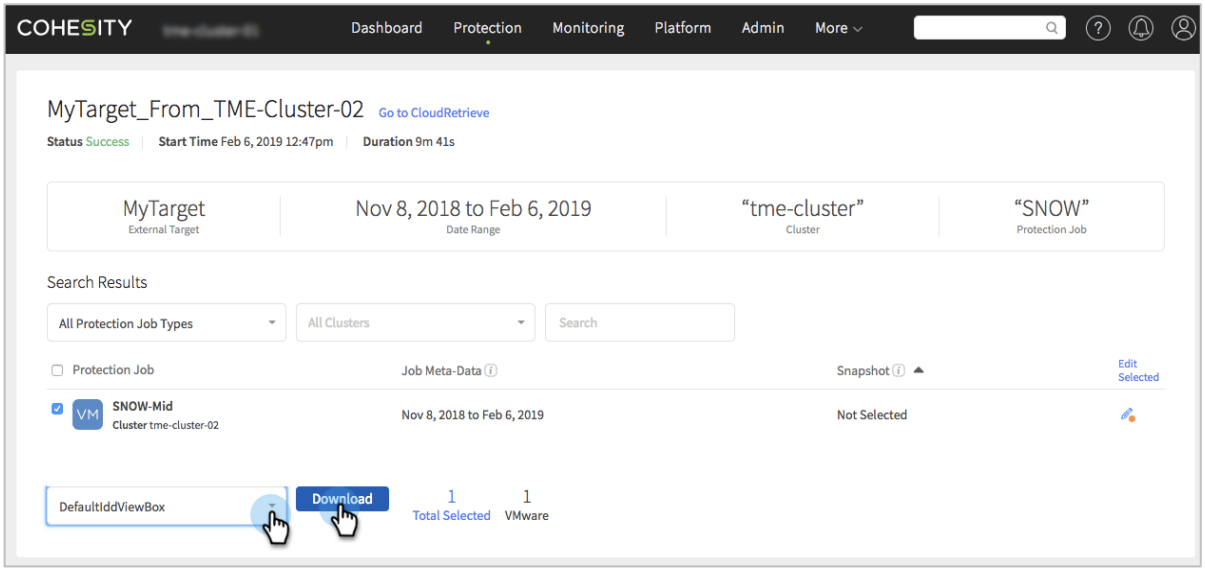


NOTE: If you are not certain which Snapshot contains the objects you need to restore, Cohesity recommends you deselect **Download Snapshot**. Once you have the Job metadata, you will be able to review the details of each Snapshot in the Protection Job, to help you narrow the download to just the specific data you need.

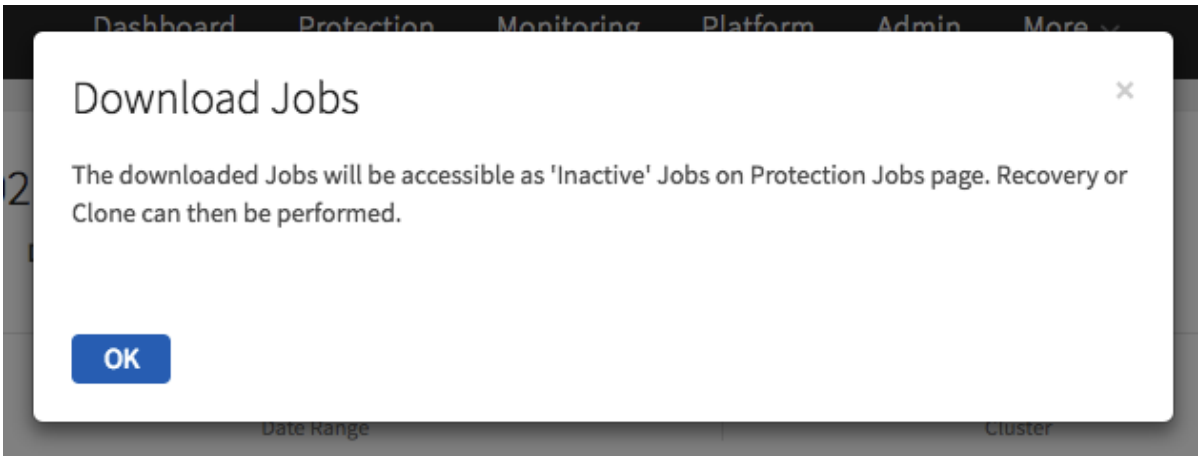
3. Make your choices and click **Save**.



4. Select the **Storage Domain** and click **Download**.

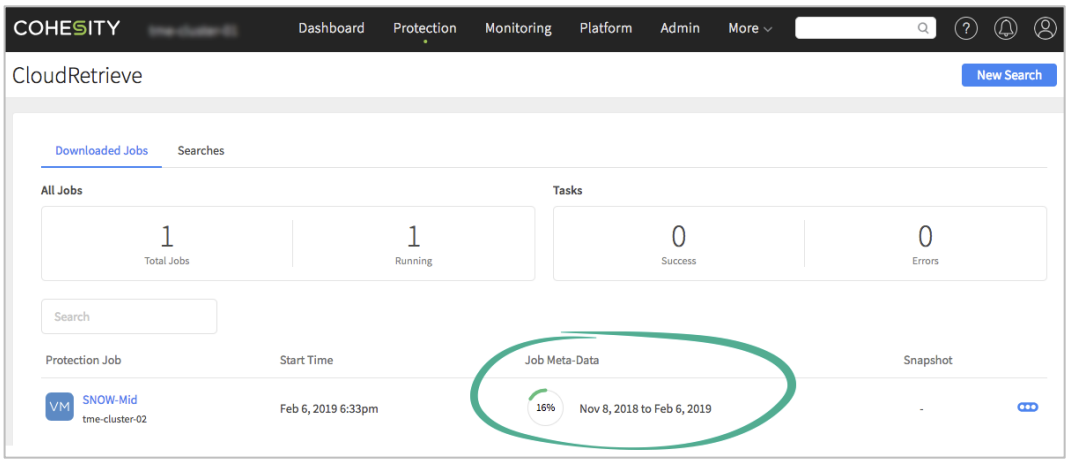


5. The downloaded Protection Job(s) will be accessible as **Inactive Jobs** under **Protection > Protection Jobs**.

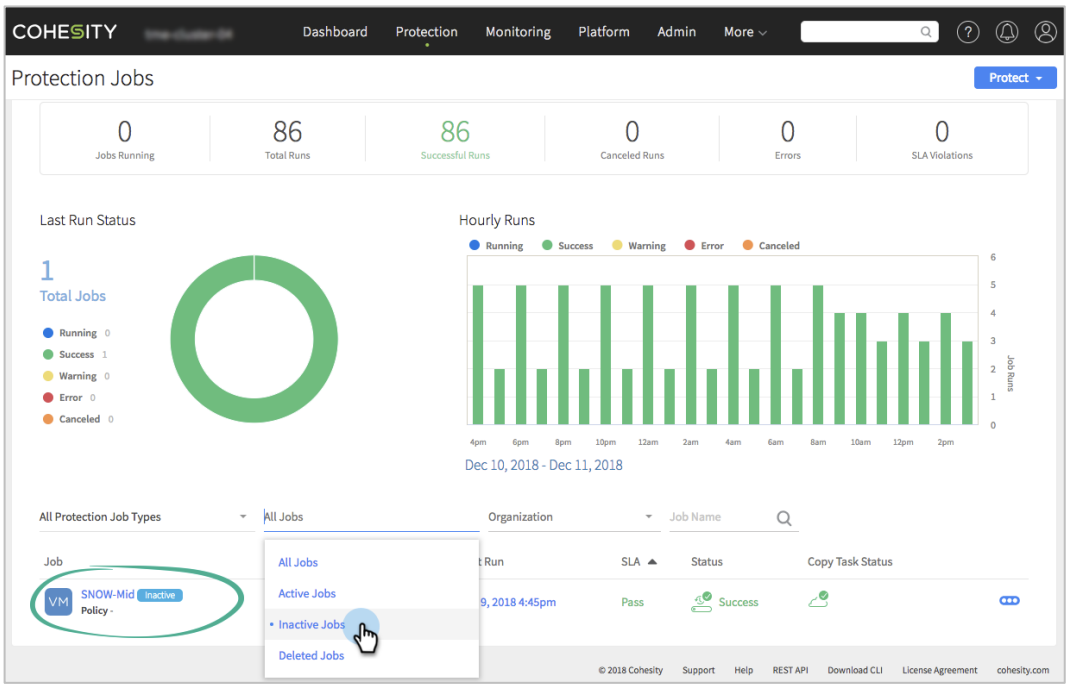


Wait for the download to complete.

6. Go to **Protection > CloudRetrieve** to monitor the progress of your download.



7. To confirm that the Protection Job is available on the new cluster, go to **Protection > Protection Jobs** and select **Inactive Jobs** from the **All Jobs** column.



The Protection Job is now available on your new Cohesity cluster, and can be used to recover your archived data.

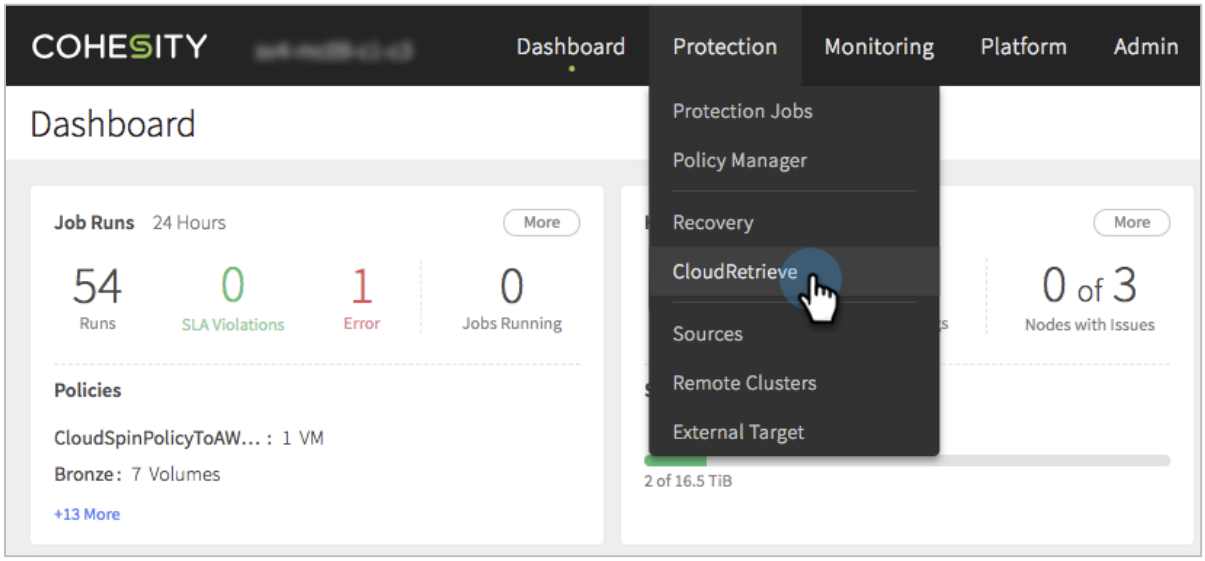
NOTE: CloudRetrieved Snapshots are not expired automatically by the new cluster. Once you have recovered the data you need, if you need to reduce your NAS size, you will have to delete the archived data from your NAS manually. Do NOT do this if the original cluster is still intact.

Recover Source Objects from Retrieved Archive on New Cluster

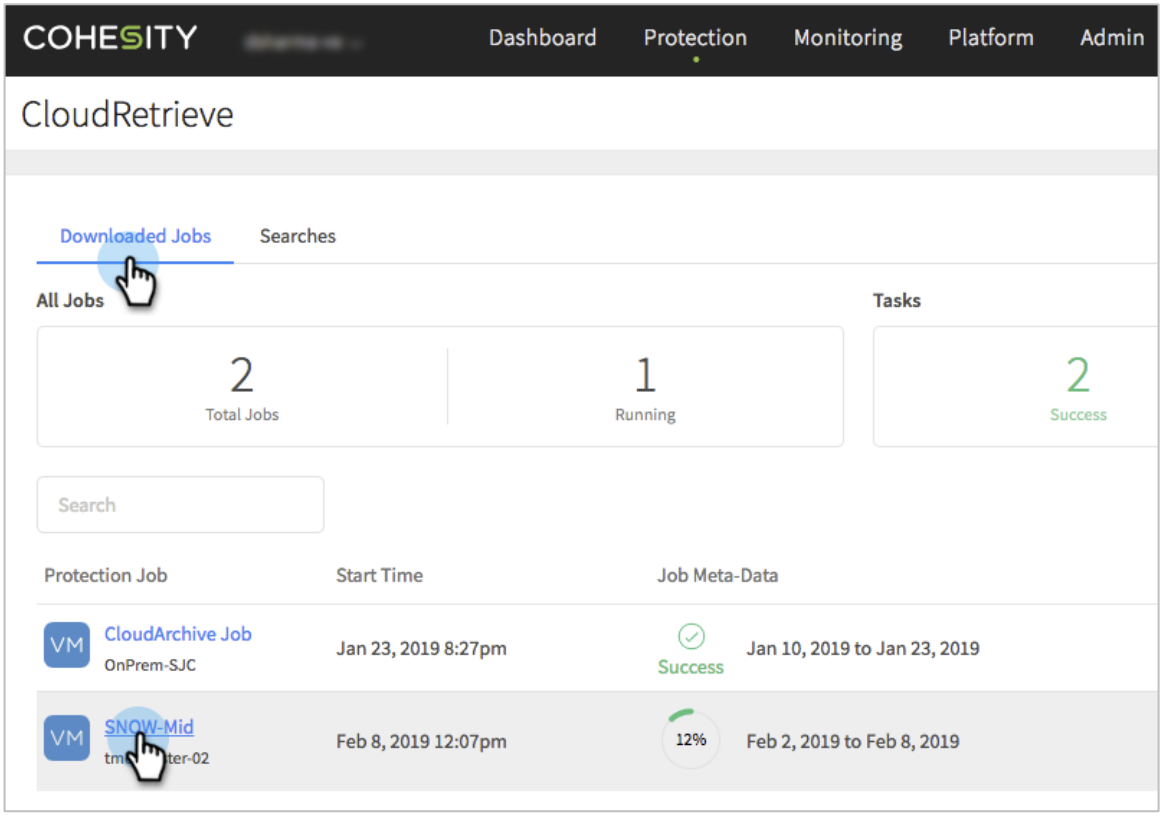
Now that you have downloaded the archived Job Runs metadata onto the new cluster, you can recover whole objects or individual files from the downloaded archive.

To recover an entire data object from a CloudRetrieved archive:

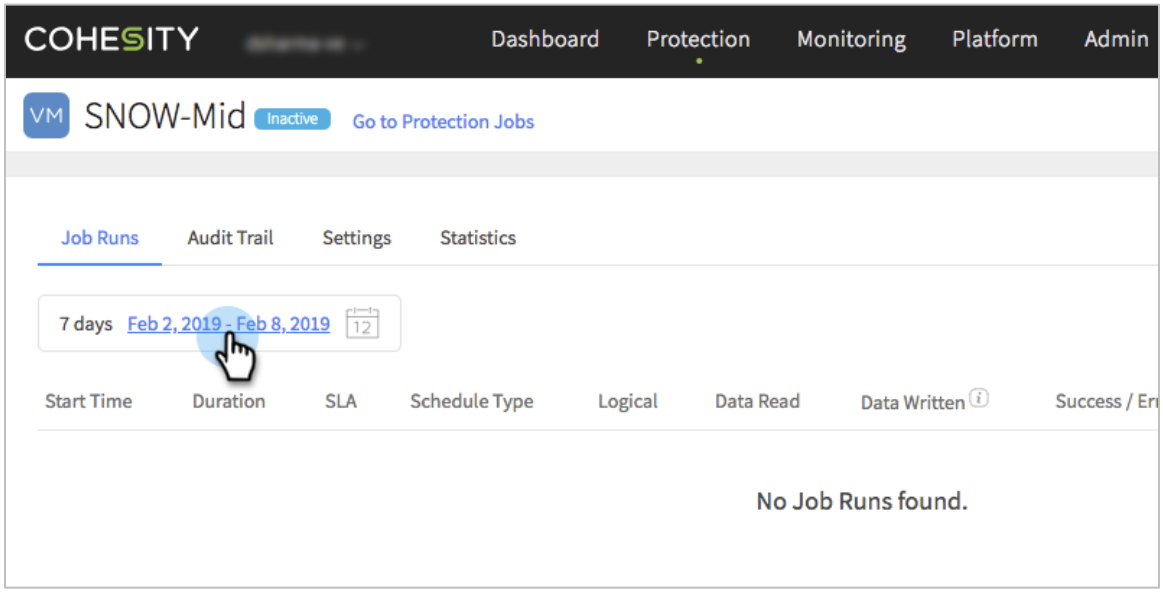
1. Log in to Cohesity Platform on the new cluster.
2. Select **Protection** > **CloudRetrieve**.



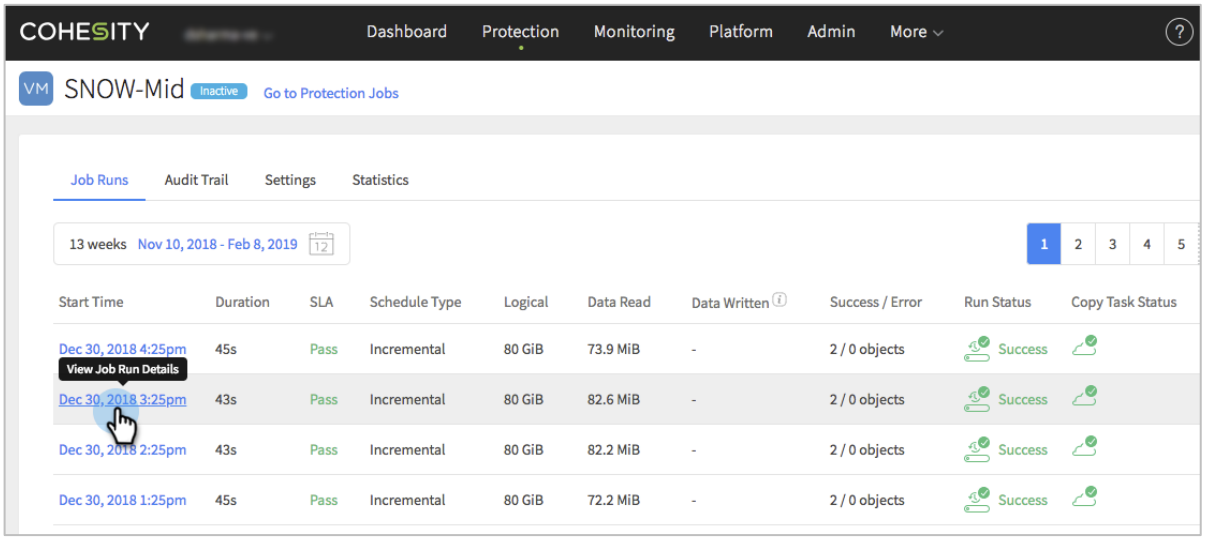
3. On the **Downloaded Jobs** tab, find the Protection Job you retrieved and click into it.



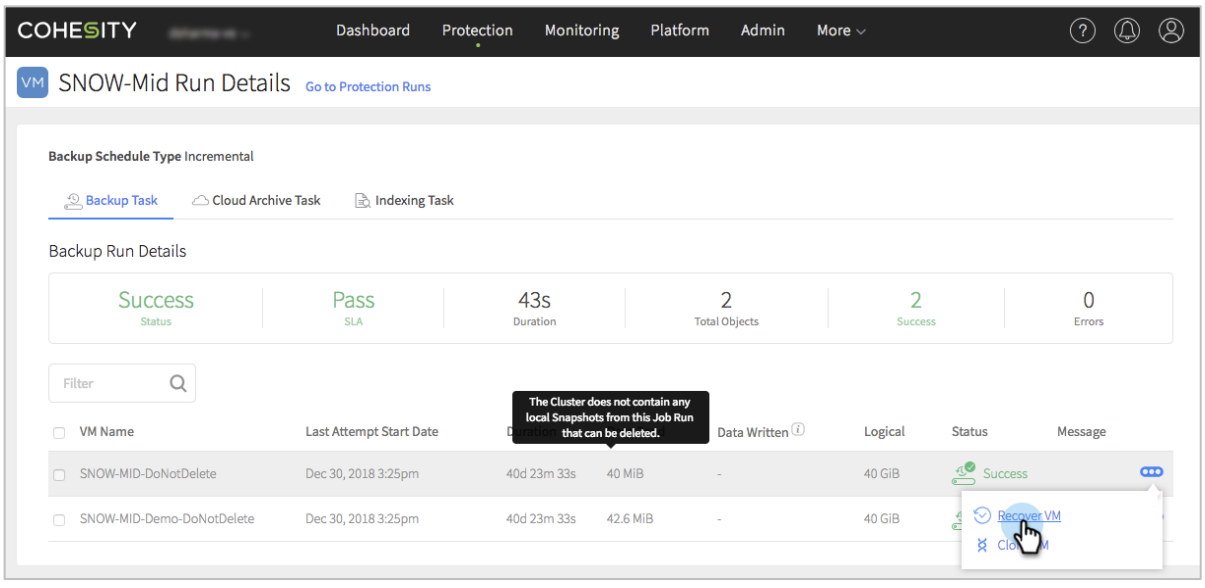
4. If the list of Job Runs is empty, click the date range and select a longer range.



5. When the list of Job Runs in the retrieved archive appears, inspect the details for each Run (**SLA**, **Schedule Type**, **Logical**, **Data Read**, **Success/Error**, and **Run Status**) and click the most appropriate Job Run.



6. In the list of data objects included in that Job Run, find the object you need to recover (for example, a particular VM), hover over the Action menu on the right, and select **Recover VM** or **Clone VM**.



7. Edit the **Task Name** and **Recover** as fields, if necessary, and then follow the rest of the standard procedure for recovery above to complete your recovery task.

See About CloudRetrieve in the online Help for more.

Appendix: Protection Job Advanced Settings

Protection Jobs combine operational requirements with the business requirements that are defined in a Protection Policy. See the all the advanced Protection Job settings, and the Job types that include them, in Table 8.

Table 8: Protection Job Advanced Settings

FIELD	DESCRIPTION	APPLICABLE JOB TYPE
Start Time	Available only if the selected Policy has a **Backup** frequency other than hourly. Indicates when the job should run. The current time is displayed by default, but you can change it.	All job types
End Date (optional)	Toggle on **End Date** and select the date on which the Protection Job stops capturing Snapshots. A Job Run that starts prior to this date will run until completion even if it completes after this date.	All job types
QoS Policy	Select **HDD** or **SSD**. **Backup HDD**: The Cohesity Cluster writes the data directly to an HDD drive for this Protection Job. **Backup SSD**: The Cohesity Cluster writes the data directly to an SSD drive for this Protection Job. Only specify this policy if you need fast ingest speed for a small number of Protection Jobs. Cohesity recommends **HDD** (the default).	All job types
Leverage Storage Snapshots for Data Protection	Toggle on to leverage storage array-based snapshots. For backups with high change rate deltas, this option can minimize the persistence time of VADP snapshots. This feature can leverage Cisco HyperFlex or Pure Flash Array Storage snapshots.	Virtual Server (VMware only)
Pre & Post Scripts	Edit this option to run scripts on the protected server before and/or after a Protection Job runs. If configured, the scripts are run every time an object is backed up by a Job Run.	Physical Server, MS SQL, Oracle Database, NAS
Skip Files on Errors	Toggled on by default. The Protection Job continues to run even if it encounters errors on files, such as permissions errors. If files are skipped, the job run details page indicates a warning status and provides additional information. If toggled off, the Protection Job stops when it encounters an error.	NAS **NOTE**: This setting is always enabled automatically for file-based Physical Server backups.

FIELD	DESCRIPTION	APPLICABLE JOB TYPE
Exclusions and Inclusions	Everything is included by default. Toggle on Exclusions and Inclusions if you want to exclude or include locations. By creating exclusion and inclusion rules, you can limit the Protection Job to a specific set of files and directories and therefore minimize the disk space used to store the data. Cohesity automatically excludes the following NetApp system files: .vtoc_internal and .bplusvtoc_internal files **.copy-offload** directory and **.tokens** file **WARNING**: Always specify forward slashes (/) even for Windows systems. For Windows, do not specify the drive letter and colon in front of directory path.	Virtual Server, NAS, Office 365
Retain on Pure Storage Array	Enter the number of the most recent snapshots to retain on the Pure Storage array. Retaining the last 5 snapshots is the default. You can change this number or choose to retain **All Snapshots**. To take incremental snapshots, this value must be at least 1.	Pure Storage Volume
Alerts (optional)	Select one or more of the following settings if you want Alerts to be created for the following triggers: **Success**: Create an Informational Alert when a Protection Job completes successfully. Emails are not sent when Informational Alerts are created. **Failure**: Create a Critical Alert if the Protection Job fails to complete. Emails are sent when Critical Alerts are created. **SLA Violation**: Create a Warning Alert if the Protection Job takes longer than the time period specified in the SLA field. Emails are sent when Warning Alerts are created.	All job types
Priority	Select a priority for the Protection Job execution. Cohesity supports concurrent backups, but if the number of Jobs exceeds the ability to process Jobs, they are executed in priority order: **High** first, then **Medium**, and then **Low**.	All job types
Email Recipients	You can add email addresses to a Protection Job to notify the email recipients when Alerts are triggered for the Job.	All job types

FIELD	DESCRIPTION	APPLICABLE JOB TYPE
Incremental After Restart	For Windows physical Servers, toggle on if you want the first scheduled backup performed after an intentional restart of the Server to be an incremental backup instead of a full backup. If an incremental backup is not possible, for example, the Server restarts after a power failure, a full backup is performed.	Physical Server (block-based only)
Indexing	Indexing is required for file recovery. The Cohesity Cluster will scan all the files in the Protection Job and create an internal index that can be used later by a Recover task to locate files by name. When creating a volume-based SQL job, indexing is not turned on automatically. Cohesity recommends turning indexing on because indexing is required to restore .mdf, .ldf and .ndf files from the cloud.	Virtual Server, Physical Server, MS SQL, Office 365, NAS
Abort in Blackouts	Available only if the selected Policy has at least one Blackout period. Toggle it on to specify that all currently executing Job Runs should abort if a blackout period specified for the Protection Job starts. By default, this toggle is off, which means after a Job Run starts, it continues to execute even when a black period specified for this Job starts. However, a new Job Run will not start during a blackout period.	All job types **NOTE**: This setting only applies to local backups and not to replication and archival. For archival, the blackout is controlled at the External Target level.
Exclude Disks	By default, all the volumes on a selected Server are protected. Toggle on to select disks to exclude for all VMs is this Job. Provide the type of controller, bus number and unit number for each disk to exclude. Excluded disks are not backed up and are not recovered during VM recovery. **NOTE**: If you exclude a disk that is part of a striped volume, the Cohesity Cluster does not index the volume even if Indexing is toggled on for this Job. You will not be able to search for or recover individual files in that volume.	Virtual Server
Exclude Physical RDM Volumes	Toggle on to exclude VMs that have Physical Disks with Raw Device Mappings (RDMs). If toggled off, those VMs will not be backed up and Job Run will fail. (Creating Snapshots of VMs that have Physical Disks with Raw Device Mappings (RDMs) is not supported by VMware vSphere or the Cohesity Cluster.)	Virtual Server

FIELD	DESCRIPTION	APPLICABLE JOB TYPE
App Consistent Backups	Toggle **App-Consistent backups** for a Protection Job if you want the guest Operating Systems of all the VMs in the Job to be quiesced before Snapshots of these VMs are created. If this option is selected, the Cohesity Cluster makes a request to the VMware vSphere software to create a quiesced VM Snapshot by invoking VMware Tools (installed on the guest Operating Systems of the VM). The VMware Tools requests that applications on the guest OS quiesce their state so application-consistent Snapshots can be created. This quiescing of VMs prior to capturing Snapshots ensures the integrity of the data saved in the Snapshots. App-Consistent backups apply to VMs only. For physical Servers, Windows is app-consistent by default and Linux is crash-consistent. For more information, see Creating Application-Consistent Snapshots in the online Help. If the **App-Consistent backups** toggle is on, the **Take a Crash Consistent backup if unable to perform an App Consistent backup** toggle is available. Toggle it on if you want the Cohesity Cluster to capture a Crash-Consistent Snapshot if the Cohesity Cluster fails to capture an App-Consistent Snapshot. For example, the Cluster may be unable to perform an App-Consistent backup when VMware Tools is not installed on the VM, the VM is powered off, or the VM cannot be quiesced. If the **Take a Crash Consistent backup if unable to perform an App Consistent backup** toggle is off and the Cohesity Cluster is unable to perform a App-Consistent backup of a VM, a Snapshot is not captured.	Virtual Server
SLA	The Service-Level Agreement (SLA) defines how long the administrator expects a Job Run to take. **Incremental**: Enter the number of minutes you expect an incremental backup job run to complete. An incremental backup captures only the differences (changed blocks) since the last job run. **Full**: Enter the number of minutes you expect a full backup job run to complete. A full backup captures the entire object (all blocks).	All job types

FIELD	DESCRIPTION	APPLICABLE JOB TYPE
Cloud Migration	Enable this option to support the Cloud Migration of VMs between hypervisors (such as vCenter) servers and Cloud Services for failover and failback. Cohesity agents must be installed on the Windows VMs prior to backing up the Snapshots on the on-premises Cohesity Cluster. Disaster Recovery using Cloud Migration is currently supported for Windows VMs backed up from a VMware vCenter Source. For more information, see Disaster Recovery of VMs using Cloud Edition in the online Help. After enabling Cloud Migration, you can download the Cohesity Agent directly from the Cohesity Dashboard.	Virtual Server
Backup Method	Two backup methods are available for MS SQL on physical Servers: **Volume-based** protects MS SQL at the Server level, meaning all databases on a Server. **File-based** protects only the specific MS SQL databases that you select. If you want to change the **Backup Method** to volume-based after Objects are selected, you must select a Server Object. A file-based Job can protect any Object.	MS SQL
Make Full Backups Copy-only	Toggle on if you want full backups to be copy-only backups so they do not affect the differential base. Copy-only full backups do not take log backups even if they are scheduled by the policy.	MS SQL
Databases to Backup	Select which databases to back up.	MS SQL

Use these settings when you are setting up your Protection Job.

Your Feedback

Was this document helpful? Send us your feedback!

About the Authors

Bart Abicht is Senior Technology Writer and Editor for Technical Marketing & Solutions Engineering at Cohesity. In his role, Bart focuses on creating white papers, solution guides, and references that are accessible, enabling, and delightful for Cohesity's partners and customers.

Adaikkappan Arumugam is Senior Manager, Tech Marketing, Solutions Engineering, & Tech Pubs at Cohesity. In his role, Adai focuses on connecting the technical expertise of Cohesity's developer and product management staff with the needs and feedback from Cohesity's customers, support staff, and sales enablement staff.

Other essential contributors include:

- Dayanand Sharma, Product Manager
- Kevin Hill, Cloud Solutions Architect
- Praveen Yarlagadda, CloudArchive Lead Engineer
- Radhani Guturi, Cloud Engineering Director
- Sai Krishna Mukundan, Director, Product Management

Document Version History

VERSION	DATE	DOCUMENT HISTORY
1.1	July 2024	Republishing
1.0	May 2019	First full release

ABOUT COHESITY

Cohesity is a leader in AI-powered data security and management. Aided by an extensive ecosystem of partners, Cohesity makes it easier to protect, manage, and get value from data – across the data center, edge, and cloud. Cohesity helps organizations defend against cybersecurity threats with comprehensive data security and management capabilities, including immutable backup snapshots, AI-based threat detection, monitoring for malicious behavior, and rapid recovery at scale. Cohesity solutions are delivered as a service, self-managed, or provided by a Cohesity-powered partner. Cohesity is headquartered in San Jose, CA, and is trusted by the world's largest enterprises, including six of the Fortune 10 and 44 of the Fortune 100.

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