COHESITY

Protect NetApp with Cohesity

Bringing Scalability and Simplicity to NetApp Data Protection

Version 1.3

May 2025

ABSTRACT

Cohesity streamlines the protection of the petabytes of NetApp NAS data you manage. Furthermore, it allows you to archive the backups to any public cloud or tape storage for long-term retention, replication, and recovery to a different location for disaster recovery.

Table of Contents

Figures

Tables

Introduction to NAS Data Protection

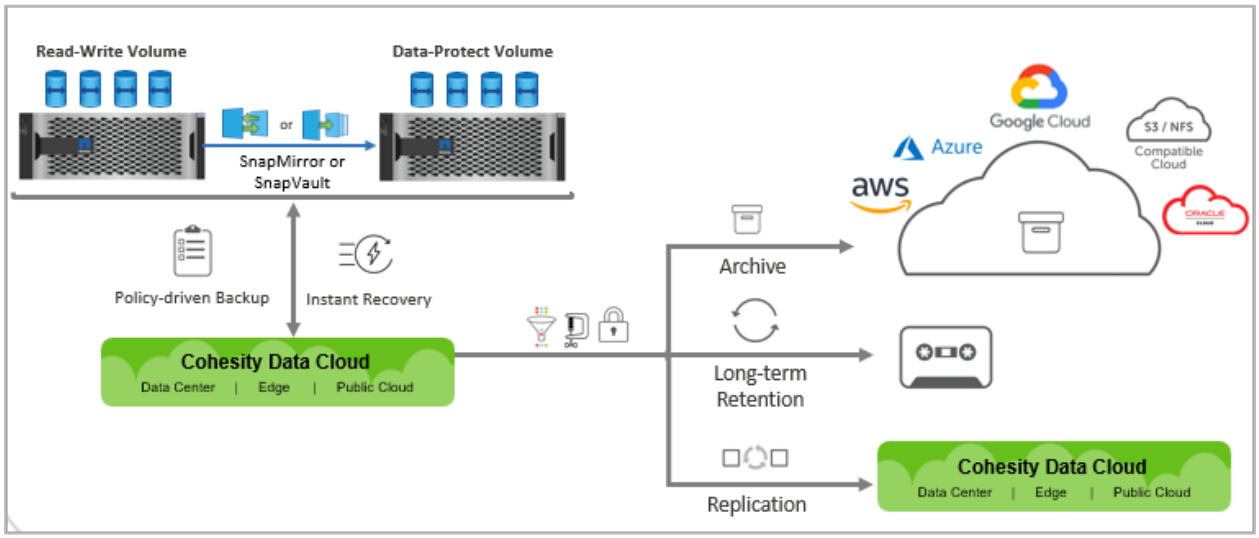
Modern enterprise data centers contain massive amounts of structured and unstructured data in many forms, including log directories, home directories, departmental shares, engineering repositories, and application datasets. This critical data requires new data protection and recovery solutions that can efficiently protect it via NAS protocols. The solution must adhere to the organizational data protection SLAs and, at the same time, provide better storage efficiency and data reusability.

Cohesity Data Cloud simplifies data protection, consolidates file and object services, provides instant access to dev/test copies, and performs in-place searches and analytics, all on a software-defined platform that spans from the edge to the cloud. Cohesity's integrated backup, archive, replication, disaster recovery, and public/private cloud support, combined with its inherent context awareness, eliminate the need for cataloging software, separate backup software, and other ancillary backup infrastructure.

Cohesity Data Protection Architecture for NetApp

With Cohesity, you can protect your NetApp Read-Write (RW) *and* Data-Protect (DP) volumes. While NetApp provides the option of a SnapMirror or SnapVault of your primary NetApp volume to a secondary NetApp volume, Cohesity gives you the flexibility to back up your data from primary *or* secondary NetApp arrays.

Figure 1: Protect NetApp Data with Cohesity



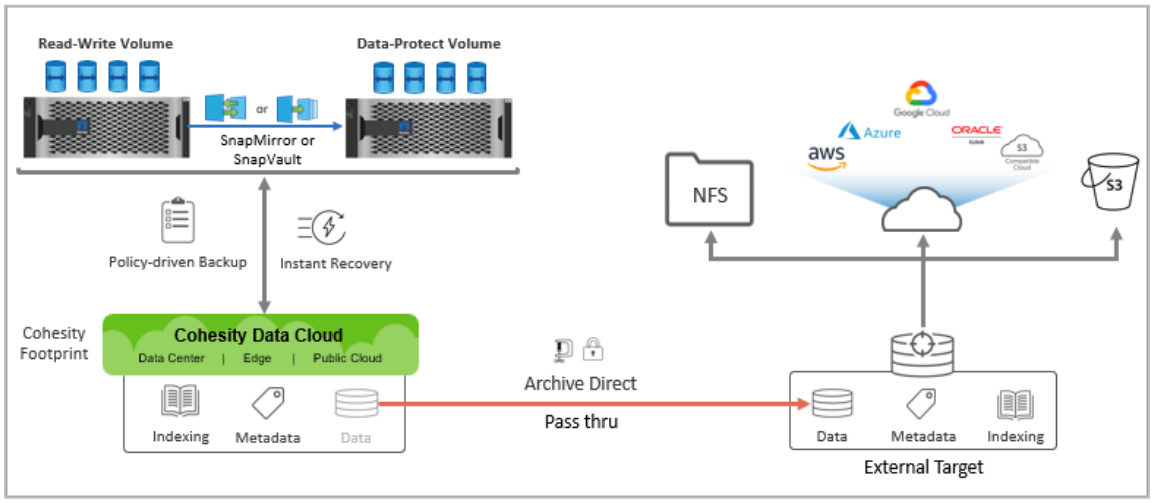
Once you back up your NetApp data via your primary or secondary NetApp array, you can also:

- Archive it to cloud/NFS/S3 storage.
- Send it to tape for long-term retention.
- Replicate it to another on-premises or cloud Cohesity cluster.

Cohesity CloudArchive Direct Architecture for NetApp

Cohesity has built CloudArchive Direct for NAS, a cost-efficient solution that processes and streams the data directly to lower-cost storage on External Targets using object storage in the public/private cloud or NFS. The footprint/capacity requirements of the Cohesity cluster are dramatically reduced by eliminating the need to store a copy locally before archiving. Only the metadata and indexes that enable quick search and recovery are stored on the Cohesity cluster. The entire NAS dataset (the data and metadata and indexes) is stored only on the External Target.

Figure 2: Making NetApp Data Archival Cost-effective with Cohesity's CloudArchive Direct



CloudArchive Direct is a policy-driven feature with seamless integration with all major cloud vendors like AWS, Azure, GCP, Oracle, or any S3-compatible object store. You can configure it with compression and encryption to achieve maximum storage efficiency and security.

NOTE: Refer to Archive Your Data Directly with Cohesity Cloud Archive Direct. for more details on Encryption and compression support.

Cohesity's Solution for NetApp NAS Data Protection

Organizations that rely on NetApp as their primary NAS need a fast, powerful, and simple backup and recovery solution that scales well to grow with their ever-growing data.

To meet these needs in a reliable and efficient ecosystem, Cohesity provides a solution that eliminates the complexities and operational inefficiencies of traditional NAS protection solutions by unifying your data protection and recovery infrastructure — including target storage, backup, replication, disaster recovery, and cloud archiving — on a single platform.

Features and Benefits

As data grows exponentially, the need for a modern approach towards data protection and backup solutions has become critical. Cohesity offers agentless, policy-based data protection, granular file-level restoration capabilities, replication and cloud archiving.

What's more, Cohesity provides:

- **Snapshot-based Backups**. Cohesity leverages its native snapshot capabilities to take snapshots of storage volumes. Point-in-time (PIT) snapshots are captured and mounted locally for faster backups.
- **Incremental Forever**. Using native NFS- and SMB-based backups, Cohesity offers true 'incremental forever' functionality. It requires performing only one full backup followed by incremental backups forever. This reduces the time required for backups and recovery and simplifies operations.
- **Multithreaded File Discovery**. Cohesity backups are much faster because it uses high-speed multithreaded file discovery.
- **Distributed and Parallel Ingest**. Cohesity's intelligent data transfer logic creates an efficient backup plan, assigns backup streams across all nodes, and performs distributed and parallel ingest in the cluster, ensuring faster backups.
- **Storage Efficiency.** Inline and post-process variable-length deduplication makes efficient use of storage and lowers the total cost of ownership.
- **Instant Recovery**. Cohesity enables instant NAS volume recovery to restore to any point-in-time (PIT) copy. Upon restore, Cohesity creates an instantaneous clone of the snapshot. The NAS volume can be accessed directly from the clone, with storage running directly from the Cohesity cluster. This eliminates the need to move data from secondary to primary systems before initiating a restore.
- **No Vendor Lock-in.** Backups are taken using native NFS and SMB protocols, making data recovery platform-independent. For example, a NetApp backup can be restored to Isilon and vice versa. You can also mobilize data directly to your cloud of choice using CloudArchive Direct.
- **Faster and Granular Restores.** The file metadata is indexed to power Google-like search in your backups, enabling very fast, granular file-level recovery to any point in time across billions of files, whether your data is backed up to a local cluster or the cloud.
- **Cloud Integration.** Natively integrated with major public cloud providers, it provides smooth archival and tiering to the cloud.

Best Practices

Cohesity protects NetApp ONTAP volumes by communicating with a NetApp ONTAP cluster or Storage Virtual Machine (SVM). As a security best practice, you must enable SSL before registering the cluster or SVM as a source on the Cohesity cluster.

To enable SSL, run the following ONTAP CLI commands on your NetApp ONTAP cluster:

1. To verify the settings, run the following command:
   ```
   netapp_cluster::> security ssl show -vserver <cluster/vservername>
   ```
2. To enable SSL, run the following command:
   ```
   netapp_cluster::> security ssl modify -vserver <cluster/vservername> -
   server-enabled true
   ```
3. After enabling SSL, run the following command to verify the settings:
   ```
   netapp_cluster::> security ssl show -vserver <cluster/vservername>.
   ```

Considerations

Review and understand the following considerations before you protect your NetApp ONTAP data:

- Instant Volume Mount for NetApp ONTAP stub file is not supported.
- You cannot restore the NetApp Data-Protect volume to the original location or to an alternate Data-Protect volume because the Data-Protect volume is a read-only volume.
- To backup NetApp DP volumes with SMB share, the volume needs to be in the junction path on the NetApp.
- Cohesity skips the backup of the following security style Qtrees:
  - Mixed-mode security style Qtrees
  - UNIX security style Qtree when root volume security style is NTFS.
- Volume recovery: The restored volume (target volume) should have the same security style as that of the volume being protected.
- When volumes with Qtrees are recovered and the Qtrees have a different export policy than the parent volume, update the export policy on the recovered Qtrees to match the original Qtrees.
- If volume export policies are not in use, but export policies are applied at the Qtree level, Cohesity IP addresses must be manually added to the Qtree-level export policies to ensure proper access.
- Recovering files/folders with Qtrees: The security style of the location (target) to which the files or folders are restored should have the same security style as the restored files and folders.
- Cohesity supports the backup of the symbolic link reparse point, IO_REPARSE_TAG_SYMLINK in an SMB share.

- Cohesity recommends a first full and incremental forever backup approach to back up your NAS sources.
- When browsing files or folders for recovery, Cohesity does not display files or folders with an absolute path length exceeding approximately 3,500 characters. To recover these files or folders, select the parent directory.

The following are the considerations for NetApp NFSv4.1 volumes:

- Encryption is not supported.
- Kerberos is not supported.
- If the NFS backup preference selected in the Cohesity UI is NFSv4.1, and if the NetApp Vserver has NFSv4.1 disabled, all the volumes belonging to this Vserver will use the NFSv3 protocol for backup.
- Recovering an NFS 4.1 volume to NFSv3 is not supported.
- Recovering an NFS 4.1 volume to the Cohesity View is not supported.

Explore Cohesity's NetApp Adapter's Capabilities

Cohesity's NetApp adapter provides a wide range of NAS data protection capabilities for your NetApp backup, including:

- **Choose Backup Scope.** You can configure it to back up your NetApp data at the cluster or Storage Virtual Machine (SVM) levels. With cluster-level backup, you can discover all the volumes across different SVMs of the cluster to back up. With SVM-level backup, you can find the volumes at the SVM level to back up.
- **Auto Protect New Volumes.** With the auto protection capability, Cohesity automatically detects and protects each new volume you add to your NetApp storage, saving you from manually configuring the Cohesity Protection Group to back up the new volumes each time you add them. (A new volume can be an addition to an existing SVM or a new SVM.)
- Cohesity allows you to enable auto protect at the cluster or SVM levels. It also gives you the flexibility to exclude the volumes of your choice from auto-protection. For example, imagine you have 100 volumes across 20 SVMs on a NetApp cluster where you frequently add new volumes to different SVMs. Out of 100 volumes, you need to back up only 70 volumes; however, their respective SVMs require auto-protection. To address this, you can enable **Auto Protect** at the cluster level in the Protection Group settings and exclude the 30 volumes from the **Auto Protect** rule.
  - **Auto Protect at the Cluster Level.** If you enable Auto Protect at the cluster level, both new SVMs and new volumes that are added to an existing auto-protected SVM are also protected. If necessary, you can exclude any specific volumes from the auto protection.
  - **Auto Protect at the SVM Level.** If you enable auto protect at the SVM level, new volumes that are added to an existing auto-protected SVM are protected. You can exclude any existing volumes from the auto protection if necessary.
- **Backup Cross-Security Style Qtrees.** With Cohesity, you can back up Qtrees even if their security style does not match the root-volume security style.

Table 1: Supported Security Style of Qtree with Root Volume's Security Style

SECURITY STYLE		
ROOT VOLUME	QTREE	
	SUPPORTED	NOT SUPPORTED
NTFS	NTFS	UNIX, MIXED
UNIX	NTFS, UNIX	MIXED
MIXED	NTFS, UNIX	MIXED

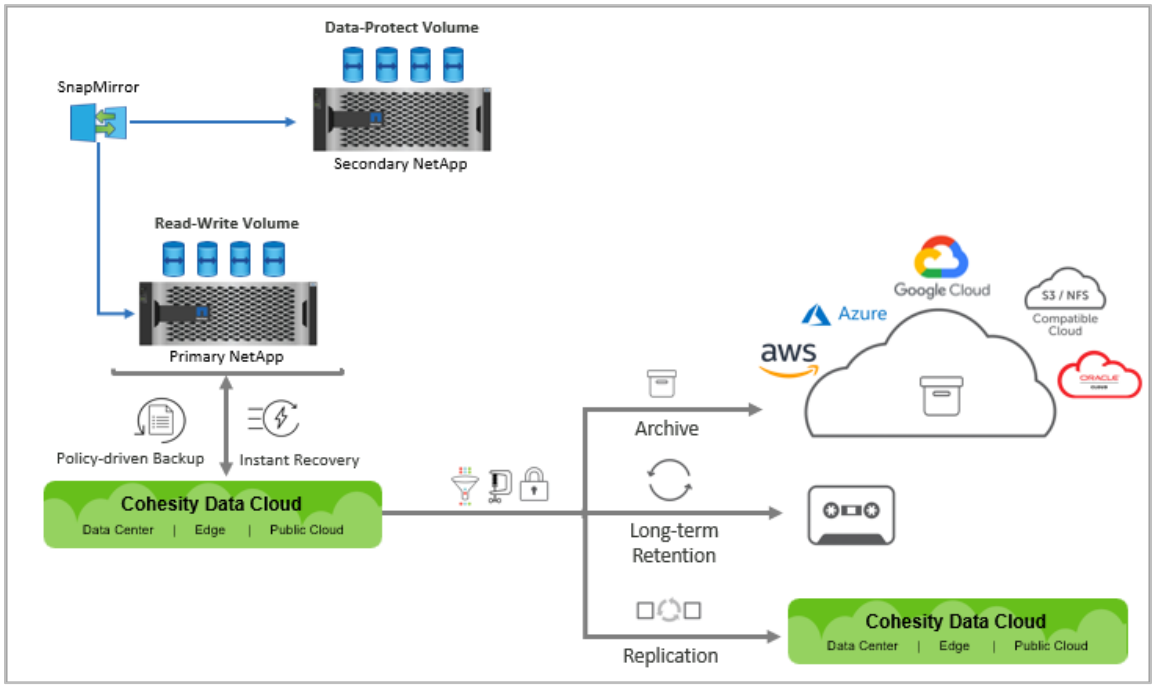
NOTE:

- Recovery to a Non-NetApp Alternate Volume treats Qtrees as directories.
- As a best practice, NetApp does not recommend using mixed security styles unless your application requires it.

- **Backup Read-Write and Data-Protect Volumes.** Cohesity allows you to backup both Read-Write (RW) as well as Data-Protect (DP) volumes.

  While protecting the data from the *primary* NetApp array, you can deploy Cohesity with the primary NetApp array and configure it to take backups of *Read-Write (RW) volumes*. In addition to backups, you can configure archiving, replication, and long-term retention for these RW volumes.

Figure 3: Back Up Primary Volume

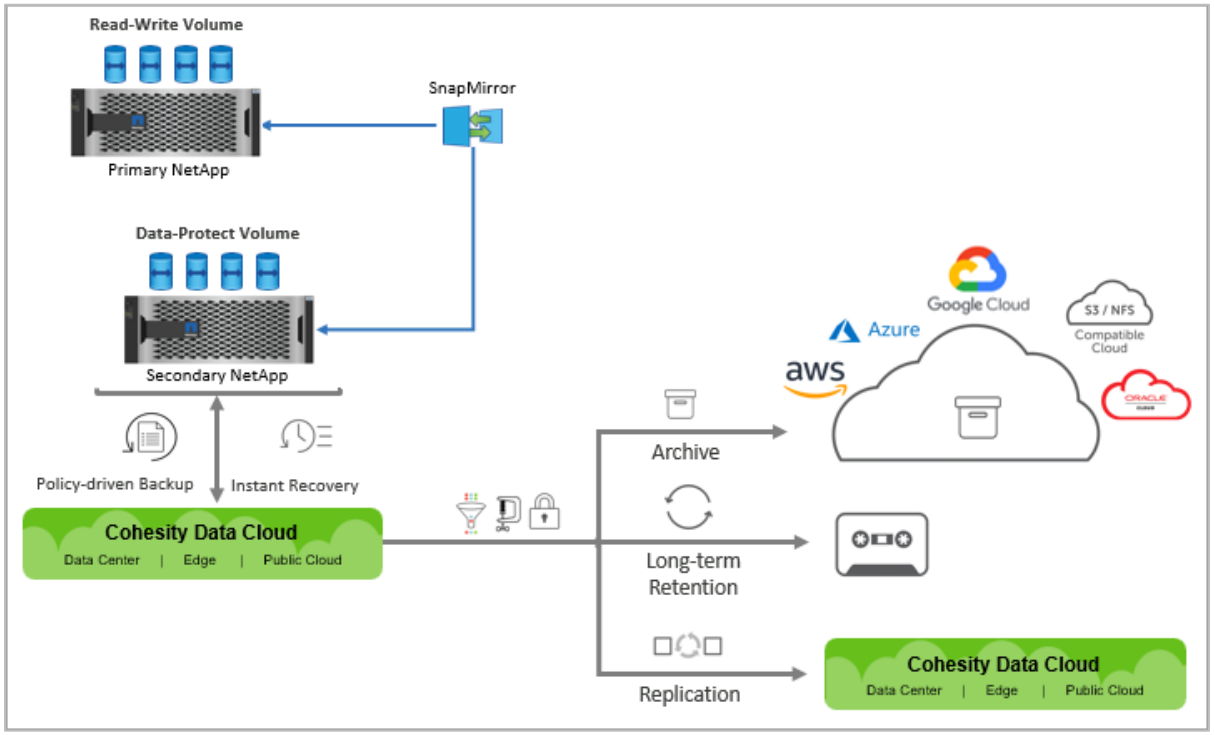


While protecting the data from the *secondary* NetApp array, you can deploy Cohesity with the secondary NetApp array and configure it to take backups of *Data-Protect (DP) volumes*. In addition to backups, you can choose to configure archiving, replication, or long-term retention for these DP volumes.

Customers might choose to protect their NetApp data from a secondary NetApp array when:

- One secondary NetApp array acts as a disaster recovery (DR) copy for multiple primary NetApp clusters (Fan-in SnapMirror), and you want to consolidate the multi-site backup from the secondary NetApp array.
- The primary NetApp array is dedicated to serving production users and cannot be used for any other workload.
- You need to make optimal use of the secondary NetApp array. Most of the time, secondary NetApp arrays are passive and only used during DR scenarios.
- The RPO and RTO for backups are less aggressive than those for SnapMirror/SnapVault, as the backup from secondary storage depends on SnapMirror/SnapVault snapshots.

Figure 4: Back Up Secondary Volume



There are some points to consider while configuring data protection from a secondary NetApp, namely:

- Network and SMB/NFS protocols must be configured on the secondary NetApp.
- The admin state of the SVM is in the starting state.
- A Junction Path should exist for each volume you plan to protect.
- To restore data to a primary NetApp array:
  - Register the primary NetApp array on Cohesity.
  - Open network communication between the primary NetApp array and Cohesity.
  - Recover to a new location.
- To restore data to any NAS volume, the target volume must be Read-Write (RW).
- As Data-Protect volumes are read-only, you cannot recover them to their original location.

Table 2: NetApp Backups — Read-Write (RW) vs Data-Protect (DP) Volumes

FEATURES	NETAPP RW VOLUME	NETAPP DP VOLUME
Snapshot Methodology for Backups	Create a snapshot.	Use an existing snapshot.
NetApp Volume Type	Read-Write Flex Volume	Data-Protect (Destination SnapMirror and SnapVault volume)
Additional Protection Group Configuration	No	Requires snapshot labels (Full and Incremental). **If multiple snapshots match, Cohesity backs up the latest snapshot.*

NOTE: You must ensure that the snapshot on the Data-Protect volume that matches the snapshot label (specified in the Protection Group) is not deleted within the configured backup duration in the associated Protection Policy.

- **Encrypt Backup Traffic Between NetApp and Cohesity.** You can encrypt your backup traffic between NetApp and Cohesity just by enabling encryption in the Protection Group's advanced settings.

For SMB backup, if encryption is enabled in the Protection Group, Cohesity starts an encrypted SMB session with NetApp to access the SMB volume data to back up. The encrypted SMB session takes care of encrypting all the session traffic between NetApp and Cohesity.

NOTE:

- Encryption should be enabled at the SVM or share level in your NetApp configuration.

Table 3: Encryption Behavior Relationship with SMB Encryption Status in NetApp

COHESITY VERSION	ENCRYPTION ENABLED IN PROTECTION GROUP?	OPERATION	SMB ENCRYPTION *ENABLED* ON NetApp SVM/SHARE	SMB ENCRYPTION *DISABLED* ON NetApp SVM/SHARE
6.8.1, 7.0, 7.0.1, 7.1, 7.1.1 and 7.1.2	No	Backup	Fail	Pass
		Restore	Fail	Pass
		Backup traffic encryption	NA, as backup failed	No
	Yes	Backup	Pass	Fail
		Restore	Pass	Fail
		Backup traffic encryption	Yes	N/A, as backup failed

For NFS backup traffic, if encryption is enabled in a Protection Group's advanced settings, then Cohesity reads the NFS volume data over Kerberos to ensure the backup traffic is encrypted between Cohesity and NetApp.

NOTE:

- Configure NFS Kerberos on NetApp.
- Configure NFS export policy to allow Kerberos krb5p.
- For Kerberos access, enable Cohesity as an NFS client.
- Join the Cohesity cluster to your domain.
- This is not supported with NFSv4.1.
- For more information, see NetApp's NFS Kerberos in ONTAP with the Microsoft Active Directory guide.

- **Download the List of Skipped Files to a Local .csv File**. You can download the list of all entities (files and folders) that were skipped during backup, along with the most applicable reason. Log in to Cohesity to download this list to a .csv file on your local machine.
- **Filter IPs (Allow\Deny) for backup communication**. Enable this option to filter the IP addresses of the NetApp ONTAP cluster. By filtering IP addresses, you can allow or deny the communication of the Cohesity cluster to specific IP addresses or subnets of the NetApp ONTAP cluster.

To filter the IP addresses of the NetApp ONTAP cluster, select one of the following options:

- **Allow IPs**: Select this option and specify the IP addresses of the NetApp ONTAP source through which the communication to the Cohesity cluster must happen. You can provide the IP addresses in a comma-separated list or in a CIDR format.
- **Deny IPs**: Select this option and specify the IP addresses of the NetApp ONTAP source through which communication to the Cohesity cluster must not happen. You can provide the IP addresses in a comma-separated list or in CIDR format.

- **File DataLock to Preserve Access Time of NetApp SnapLock Directories.** You can enable **File DataLock** in Cohesity Protection Group's advanced settings for NAS data protection. If enabled, Cohesity preserves the write once read many (WORM) attributes for the files and folders of the protected WORM directories, which provides the access time, lock period of the files/folders, and other information. These attributes are applied to the files and folders when they are recovered to Cohesity View, thus making the data immutable in the View. Cohesity also allows you to:
  - Override the locking period of the files/folders while recovering them to Cohesity View.
  - Set the locking period for the new files/folders that are created in the recovered Cohesity View.

NOTE:

- File DataLock supports Enterprise or Compliance modes and preserves the access times of the hard and symbolic links.
- You cannot enable or disable File DataLock once the Protection Group is created.

See File DataLock in the online Help for more details.

Understand Cohesity's NAS Backup Approach

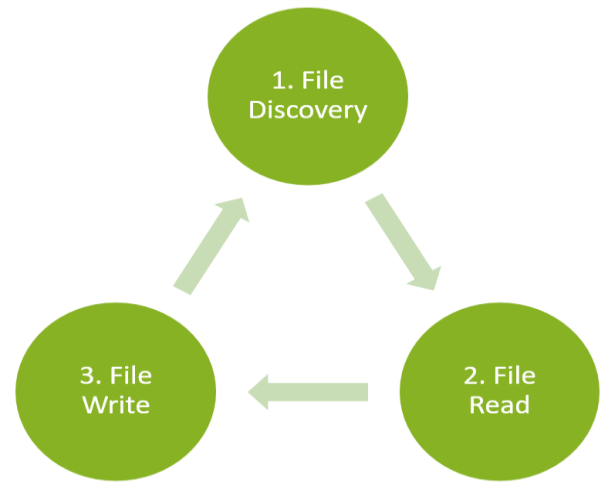
Cohesity offers two methods to back up and recover NetApp ONTAP volume data.

- NetApp Volume backup using Cohesity Built-in CFT (Streaming Diff) – A File-Runner approach.
- NetApp Volume backup using NetApp SnapDiff APIs – A SnapDiff API approach.

NetApp Volume backup using Cohesity Built-in CFT (Streaming Diff) – *A File-Runner approach*

Cohesity uses a simple three-step approach to protecting NetApp data. Each step is optimized with modern techniques like API integration, adaptive data tasking, and distributed, parallel data streaming.

To understand how Cohesity NAS Data protection works, it helps to understand precisely what is going on in each phase of the process:



1. **File Discovery**: Discover the files to back up.
2. **File Read**: Read the discovered files over NAS protocols and divide the files into multiple data chunks.
3. **File Write**: Write the data chunks to Cohesity using distributed and parallel streams.

File Discovery

At a high level, for every NAS backup run, the Cohesity file runner discovers the list of files and folders that need to be backed up from the user-selected NetApp object in the Cohesity Protection Group. You can choose to back up at the cluster level or SVM level. You can also use exclusion and inclusion rules to define the objects to be protected.

NOTE:

- To add an exclusion or inclusion, you must prefix a forward slash ('/') or suffix an asterisk ('*') to the path or a particular file within the protected object. For example, '/test' or '*.txt'.
- Cohesity does not support regular expressions for inclusions.
- Cohesity supports complex regular expressions for exclusions such as '`/Vol1_Folder2/*.txt`' or '`/Vol1_Folder*/File1.txt`.' See Step 21 in Create a Protection Group for NAS Volumes in the online Help for more on exclusions and inclusions.

Cohesity performs slightly different file-discovery processes during full and incremental backups. See Cohesity NAS Backup Workflows below for details.

File Read

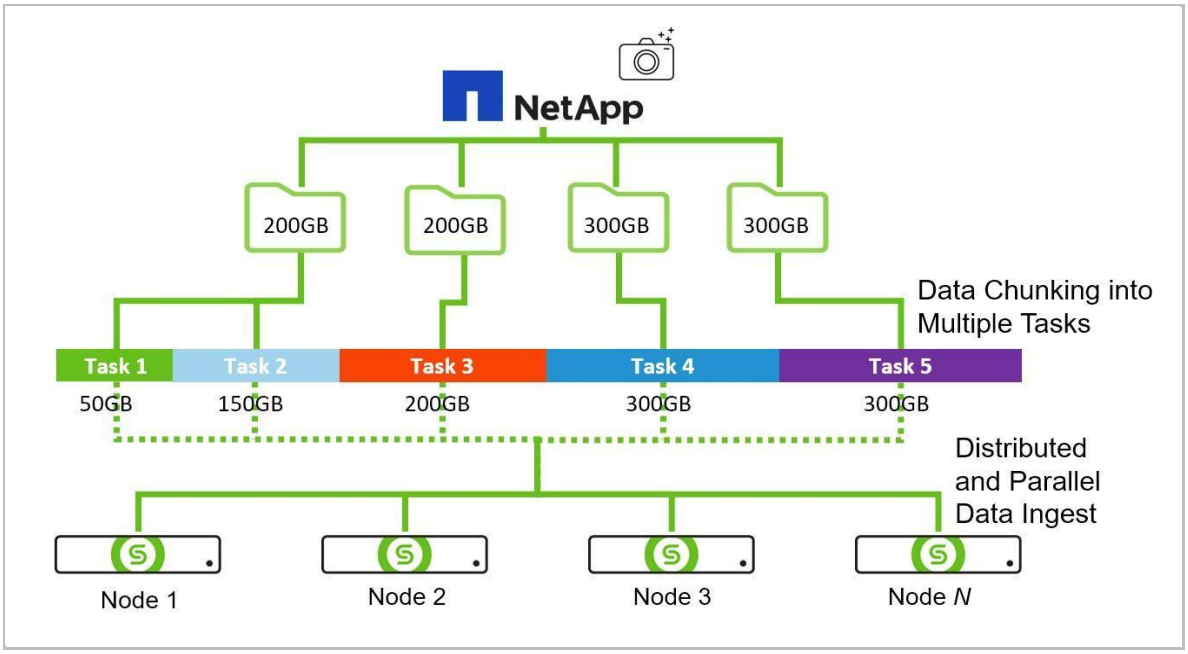
Cohesity uses a parallel and distributed architecture to read the NetApp NAS dataset using native SMB/NFS protocols identified during the *File Discovery* phase. It uses an intelligent algorithm to divide the identified dataset into multiple tasks (data chunks) based on the size and the number of files. This adaptive task chunking enables Cohesity to back up data in a more efficient way, reducing the backup window and improving backup SLAs.

Once the dataset is divided into multiple tasks, Cohesity writes them over parallel streams to different Cohesity cluster nodes. As all the nodes are involved in writing the data, backup throughput is much improved.

After a Protection Run is completed, all files are indexed by Cohesity to enable global search and rapid recovery.

In the example below, identified datasets in the File Discovery phase have been divided into five tasks of varying sizes. Each task is then ingested in parallel and distributed to all Cohesity nodes.

Figure 5: File Read



File Write

File Write is the last phase in the backup process, where files and folders are written to the cluster or, with CloudArchive Direct, are streamed directly to an archive storage target in the public cloud.

Cohesity intelligently selects the node for data placement based on multiple factors such as capacity, performance, Quality of Service (QoS) policy, and system state of the node, which helps improve RTOs (Recovery Time Objectives) and RPOs (Recovery Point Objectives). The ingest engine also ensures that data is optimally placed onto the SSD or spinning disk tier that best suits the data stream's profile.

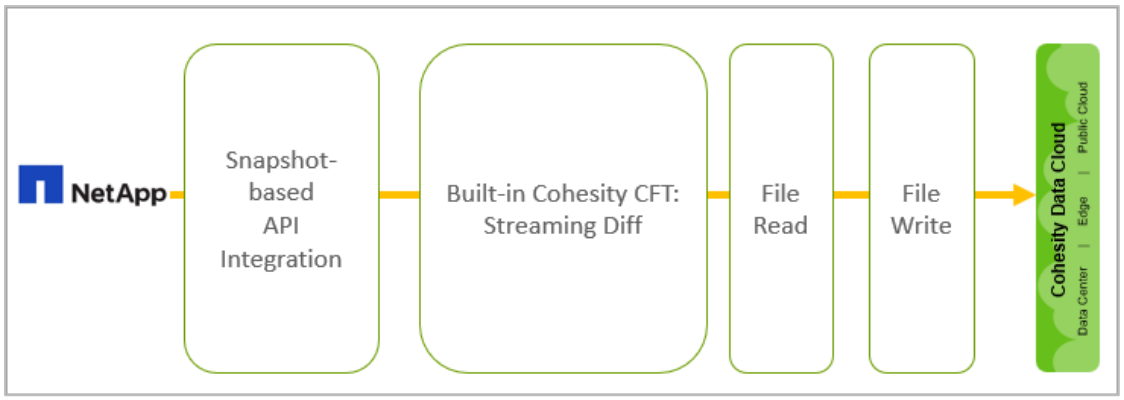
Cohesity also provides data encryption, both at rest and in flight, with AES 256-bit encryption. Once a data ingestion task is completed, Cohesity unmounts the previous snapshot from all cluster nodes, regardless of the Protection Run status (success, warning, or failure), and uses checkpoint files to identify changed data. The snapshots themselves are deleted, again regardless of the Protection Run results.

Cohesity NAS Backup Workflows

As you prepare to protect your NetApp data, it helps to understand the various workflows and choices available in Cohesity's solution. With our approach of taking a full backup once and following it with incremental forever backups, each phase involves slightly different operations:

- Full backups: Cohesity executes high-speed file discovery using a file runner.
- Incremental backups: After the full backup, Cohesity takes a forever-incremental approach with Cohesity's built-in Changed File Tracking (CFT) streaming diff technology.

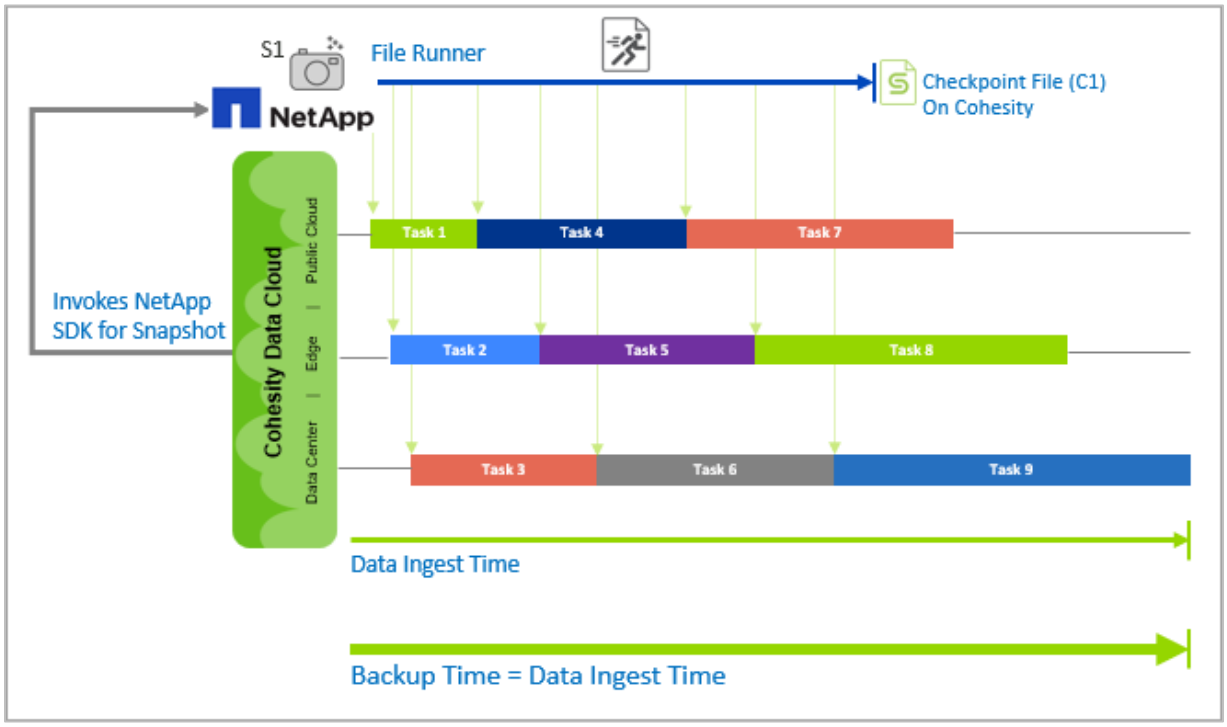
Figure 6: Cohesity's Approach to Protecting NetApp NAS Data



Full Backup with High-speed File Discovery

The initial backup is always a full backup and copies all the data. Cohesity leverages the NetApp snapshot SDK to create a snapshot for a point-in-time (PIT) backup by performing a high-speed file discovery using Cohesity's file runner.

Figure 7: Cohesity's Initial, Full NetApp Backup Process



During the initial, full backup, Cohesity creates snapshot S1 using the NetApp snapshot SDK and performs the following operations in parallel, thereby dramatically reducing backup times:

- **File Runner:** Starts the high-speed file runner on the snapshot (S1) and discovers the files & folders to back up by accessing the snapshot on all the nodes of the Cohesity cluster.
- **Checkpoint File:** The file runner stores the metadata in a checkpoint file (C1).

- **Data Ingest:** This function creates data-ingest tasks as new files and folders are discovered and distributes them in parallel across all nodes in the cluster.

After the first full backup is completed, it is followed by incremental forever backups that send only changed data.

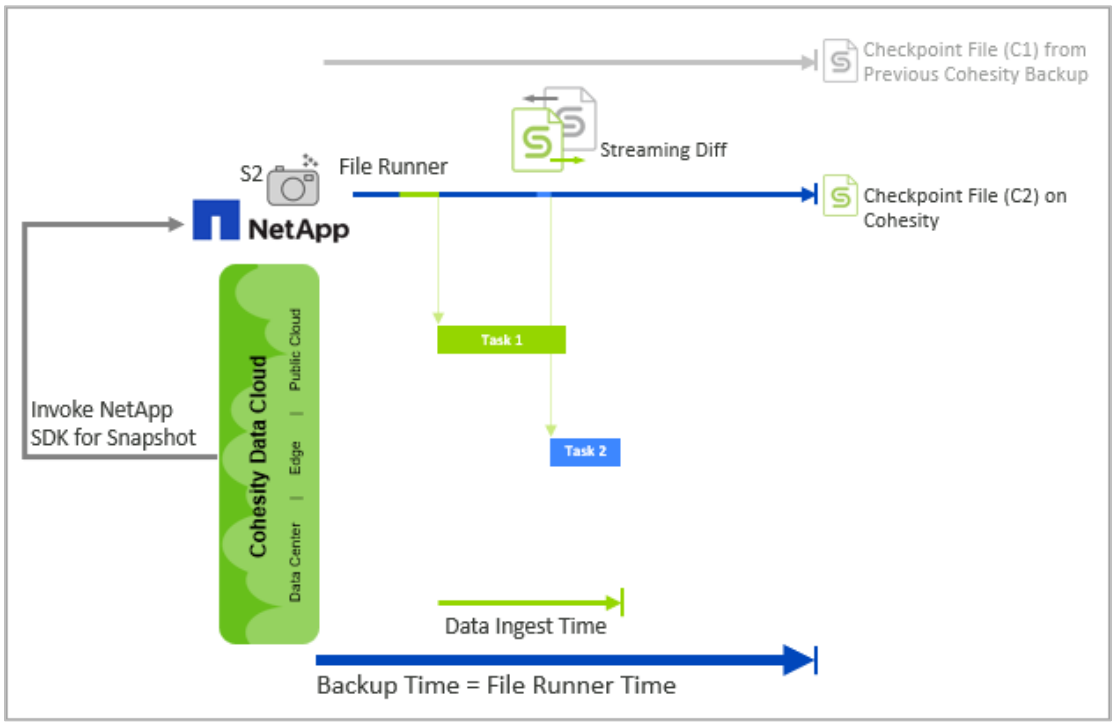
For incremental backups, Cohesity uses the built-in CFT, which uses checkpoint files to identify changed data. So, after the full backup run, the snapshot (S1) is deleted.

NOTE: For Data-Protect volumes, Cohesity uses the existing snapshot.

Incremental Forever Backups with Built-in Cohesity CFT

With a full backup, Cohesity uses an incremental forever approach for all subsequent backups. The goal of an incremental backup is to locate and transfer only the data that has changed since the last backup. Cohesity identifies the changed data for incremental backup by performing its own built-in CFT using 'streaming diff' technology to discover the changed files and folders to back up.

Figure 8: Cohesity's Incremental NetApp Backup Process



During the incremental backup, Cohesity uses the NetApp snapshot SDK to create a new snapshot (S2) and performs the following operations in parallel, thereby dramatically reducing backup times:

- **File Runner**. Starts the high-speed file runner on the new snapshot (S2) and discovers the files & folders by accessing the snapshot on all the nodes of the Cohesity cluster.
- **Checkpoint File**. The file runner stores the metadata in a new checkpoint file (C2).
- **Streaming Diff**. This operation compares the file runner output with the previous checkpoint file (C1) and identifies the changed files to protect.

- **Data Ingest**. It creates data-ingest tasks as changed files are discovered and distributes them in parallel across all nodes in the cluster. Cohesity creates the tasks as changed files are identified and does not wait for the new checkpoint file (C2) to complete.

NOTE: Because the forever-incremental backup approach relies on built-in CFT (comparing checkpoint files), the snapshots are deleted after every backup operation.

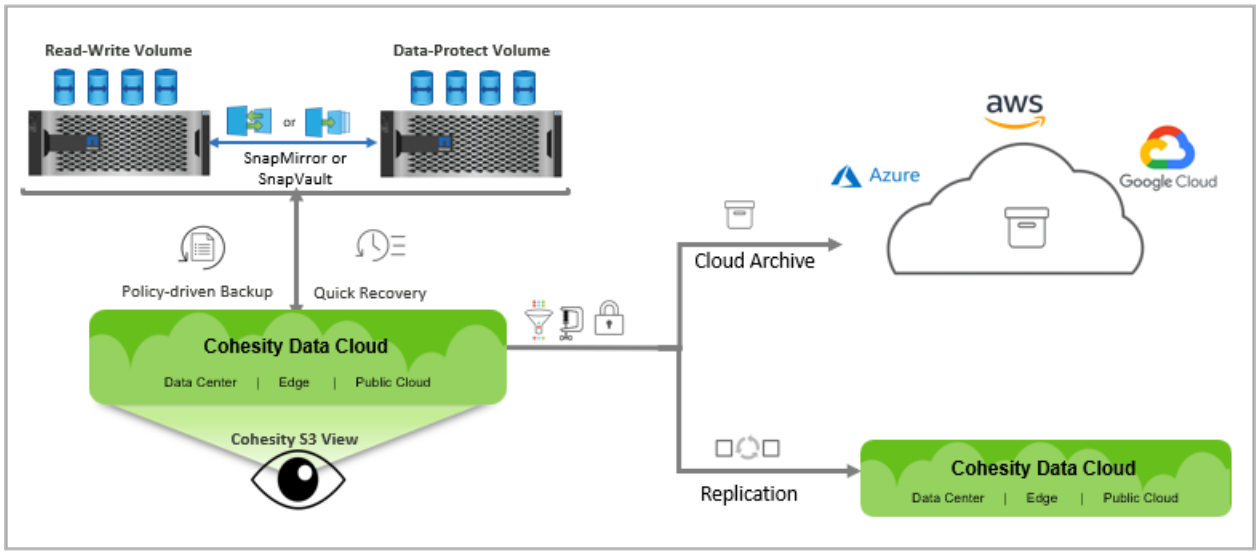
NOTE: Cohesity uses the existing snapshot for incremental backups for Data-Protect volumes.

NetApp Volume backup using NetApp SnapDiff APIs – *A SnapDiff API approach*

In this method, Cohesity leverages NetApp's SnapMirror cloud functionality to move the data to the Cohesity cluster (in NetApp's proprietary format) and NetApp's SnapDiff functionality to get the changed file list.

NetApp SnapDiff Backup Architecture

Figure 9: NetApp SnapDiff Backup Architecture



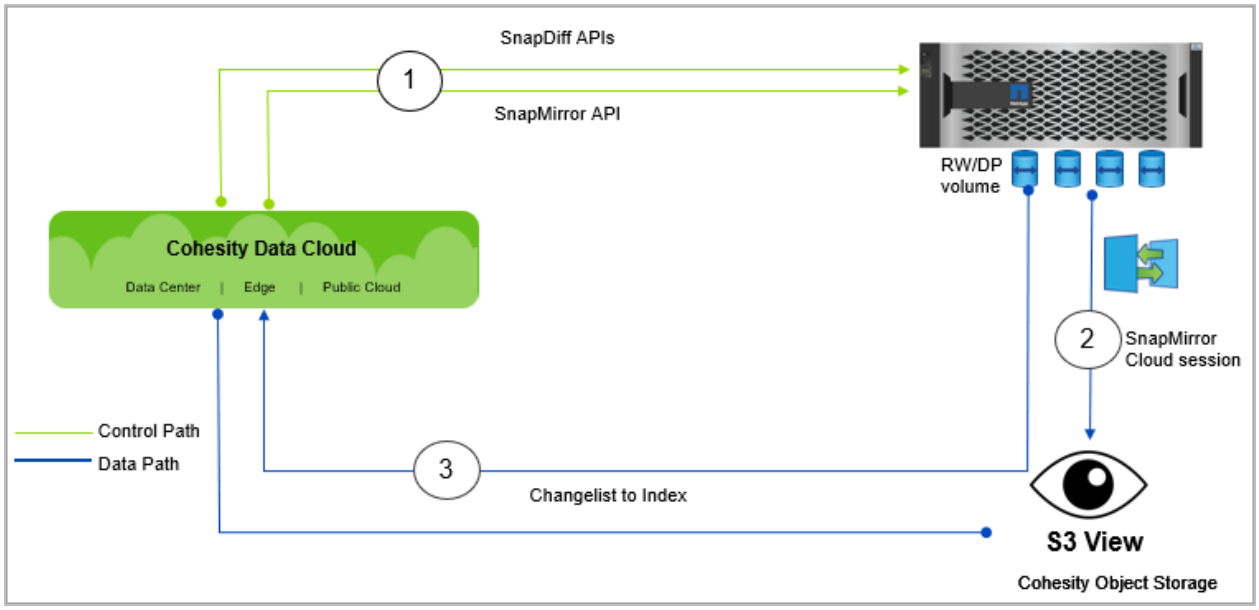
Cohesity can protect NetApp's Read-Write Flex volumes and Data-Protect Flex volumes using NetApp's native change file tracking (SnapDiff API) onto a Cohesity's S3 view (object store). You can configure the backed-up data to be replicated to a remote Cohesity Cluster or archived to an external target.

NOTE:

- As of release 7.1.2, Cohesity's SnapDiff backups only support CloudArchive with an incremental forever format to an S3-compatible target, NAS, or Tape.
- CloudArchive with Periodic full is not supported.
- CloudArchive Direct is not supported.

SnapDiff Full Backup Workflow

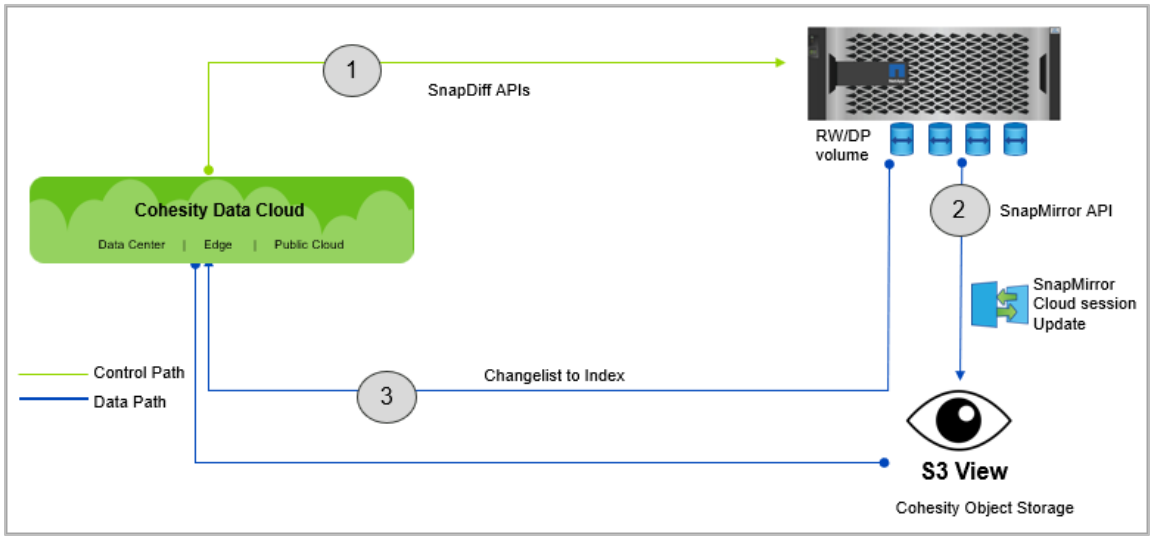
Figure 10: SnapDiff Full Backup Workflow



4. API communication with NetApp to perform the following:
   a. **SnapMirror API**–to Initialize the SnapMirror session between NetApp volume and Cohesity S3 view (Object storage). This creates the SnapMirror Cloud session between RW/DP volume and Cohesity S3 view (Object storage) for data transfer.
   b. **SnapDiff APIs**–to create the change list from the snapmirror snapshot.
5. Initialize the SnapMirror Cloud session to backup the data on Cohesity S3 view (Object storage). This transfers the SnapMirror base snapshot to Cohesity S3 View (Object storage)
6. The Changelist returned from SnapDiff APIs (list of files/folders that are backed up) are indexed and stored on Cohesity internal view.

SnapDiff Incremental Backup Workflow

Figure 11: SnapDiff Incremental Backup Workflow



1. Cohesity queries\receives the change file list between the previous SnapMirror snapshot and the current snapshot using SnapDiff APIs.
2. Updates the SnapMirror Cloud session to backup the change list to the Cohesity S3 view using the SnapMirror APIs.
3. Changelist returned from SnapDiff APIs (list of files/folders that are backed up) are indexed and stored on Cohesity internal view.

SnapDiff Backup Considerations

Cohesity Platform	NetApp Cluster
• The NetApp source should be registered at the Cluster Level. • Once SnapDiff backup is enabled for a Protection Group (PG), it cannot be disabled. • A DP Flex volume cannot be protected in two different SnapDiff-enabled PGs. • A DP Flex volume cannot be protected from 2 different Cohesity clusters with SnapDiff Backup. • In the case of SMB backup, the Mtime of the files is not preserved. • Hardlinks and symlinks are not supported. • Do not modify/delete the Cohesity S3 view involved in the SnapDiff backup. • Recovery to Cohesity View is not supported; hence, Instant access to backup data is not supported.	• Do not modify/delete the SnapMirror session created between NetApp volume and Cohesity S3 view. • Cohesity will pick the latest available NetApp snapshot of the volume being protected (internally using the snapshot prefix provided in the protection group) based on policy scheduled on NetApp and any snapshot updates scheduled on the DP volume from NetApp policy. • NFS must be enabled on the data LIFs to run SnapDiff APIs (for File Level Recovery & to allow RPC initialization).

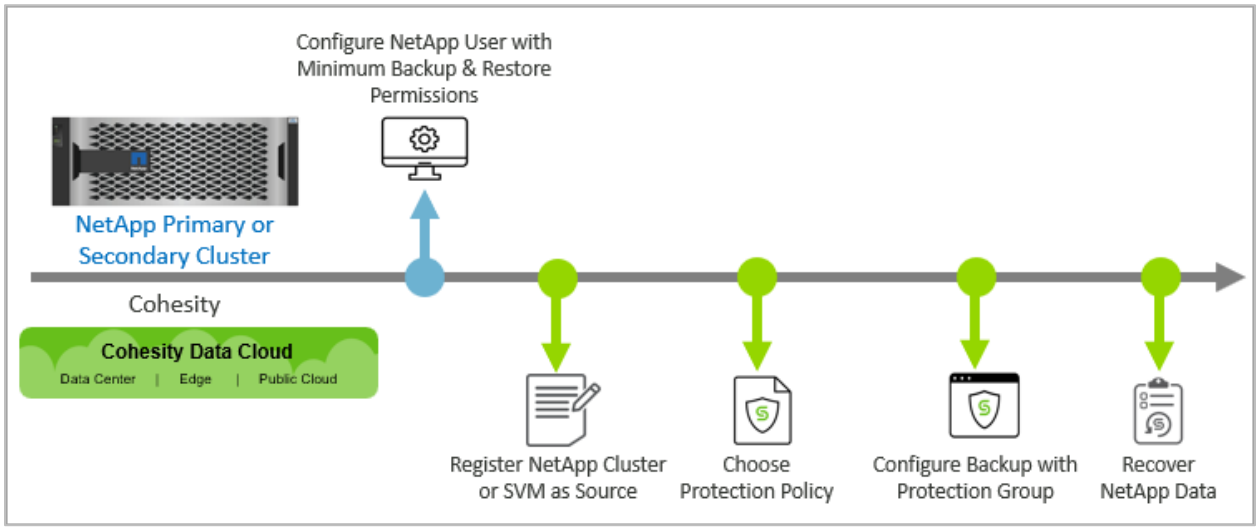
Protect and Recover NetApp with Cohesity

Using Cohesity, you can back up one or more NetApp NFS mount points or SMB shares while preserving ACLs and Extended attributes on SMB at a cluster or SVM level.

To prepare Cohesity to back up your NetApp cluster or SVM, ensure your Cohesity cluster and NetApp cluster or SVM meet the prerequisites for File-Runner backup and SnapDiff backup and then:

1. Configure the NetApp user with the minimum required backup and restore permissions.
   - Minimum permissions for File-Runner-based NetApp ONTAP backup
   - Minimum permissions for SnapDiff-based NetApp ONTAP backup
2. Register your NetApp cluster or SVM as a source in Cohesity.
   - For File-Runner-based backup, register the source as NetApp Cluster or SVM.
   - For SnapDiff-based backup, you must register the source as a NetApp Cluster only.
3. Create a Cohesity Protection Policy.
4. Create a Protection Group to protect your NAS volumes.
   - Create Protection Group for File-runner-based backup of NetApp ONTAP
   - Create Protection Group for SnapDiff-based backup of NetApp ONTAP
5. Recover your NetApp NAS data using Cohesity.

Figure 12: Protect NetApp NAS with Cohesity



Cohesity's Solution for NAS Data Recovery

NAS data recovery is essentially the rebuilding of NAS data that is lost due to unfortunate incidents such as media or storage failure, application failure, or data deletion due to human errors. Cohesity offers efficient NAS data recovery options by protecting your data in backups and allowing you to instantly recover it whenever necessary, with uncompromised data availability and reduced downtimes.

Overview

Cohesity recovers NAS data from snapshots of storage volumes created earlier by a Protection Group. Restoring data for NetApp using Cohesity is simple, fast, and intuitive. Restoring can be performed at various levels of granularity, including at the Files/Folders and Volume levels. NAS volumes, files, and folders can be recovered to their original location or to a newly specified location, which can be in the original source or a different NAS source. You can also perform an instant mount of your backup and download files from specific snapshots that were created by a Cohesity Protection Group.

Features and Benefits of Cohesity's NAS Recovery Solution

NAS data recovery with Cohesity has many benefits, including:

- **Reduced restore time**. Most current backup solutions store file-level backup metadata in a Table Of Contents (TOC). A TOC can take a longer time to load during recovery when a single file needs to be recovered from the backup image that contains millions of data files. Cohesity addresses this challenge by indexing individual files (simple files, directories, and symlinks) in the backup data after every successful Protection Run instead of relying on a TOC to store metadata. This is similar to the WAFL (Write Anywhere File Layout) that ONTAP uses, thus reducing recovery time.
- **Flexible restore targets with File-Runner-based backups**. Most current backup solutions do not support data restoration to different vendor devices, as their backups are not written in a native format. With Cohesity, backup administrators have the flexibility to restore data to original or alternate locations.

  > **NOTE**: SnapDiff based backups are written to Cohesity in NetApp's proprietary format.

- **Search needle in a haystack**. Unlike legacy recovery solutions, Cohesity negates the need to bring old backups back online to analyze data to search and locate specific files. In Cohesity, you can perform a search across different File-Runner-based backups, at different levels of granularity, by entering search terms to locate by Source Name, Protection Group, Backup Start Date, etc.

  > **NOTE**: This does not apply to SnapDiff-based backups as the data is written in NetApp proprietary format.

- **Instant Recovery**. In most legacy solutions, backup administrators can provide users access to backup data only *after* the recovery is complete. This means significant wait times, as it involves reading the data first and only *then* writing it to a specified location. With Cohesity, you can instantly recover the File-Runner-based backup to a Cohesity View and provide immediate access to the backup data via NFS or SMB protocols.

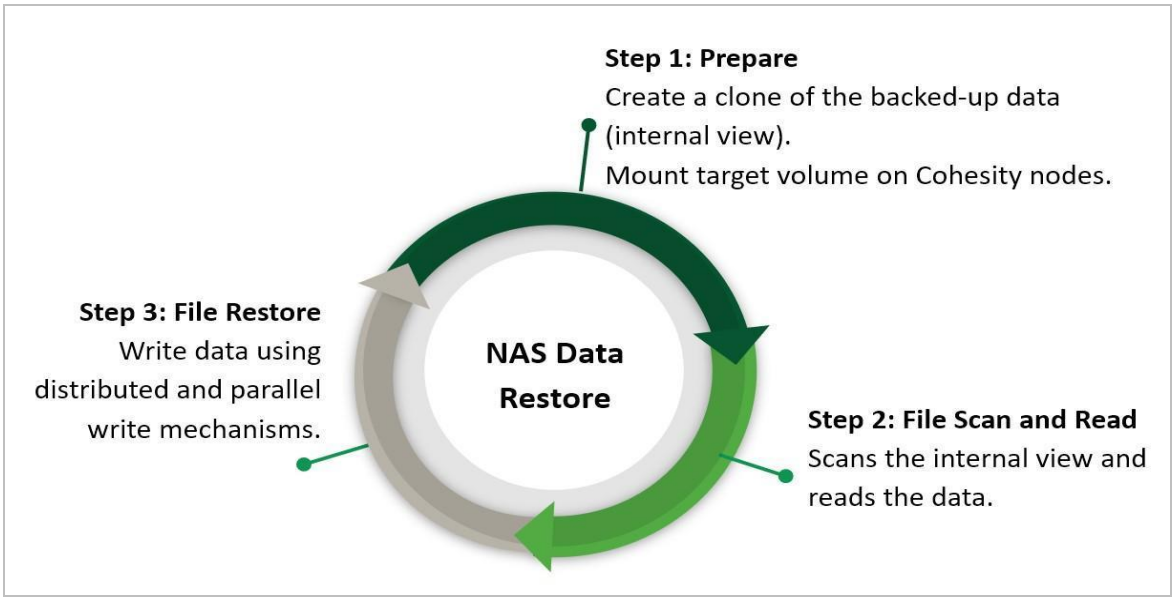
  NOTE: Recover to Cohesity View for instant access is not supported with SnapDiff base backups.

Understand NAS Restore Internal Workflow

Once NetApp data is backed up with Cohesity, you can start a recovery task that restores specific NetApp volumes or files as per your business requirements. During the recovery task run, there are multiple operations Cohesity runs in the background to complete the restoration successfully.

NAS Restore Internal Workflow of File-Runner-based backup

Figure 13: NAS Restore Internal Workflow of File-Runner-based backup



In Cohesity's NAS restore internal sequence, Cohesity does the following:

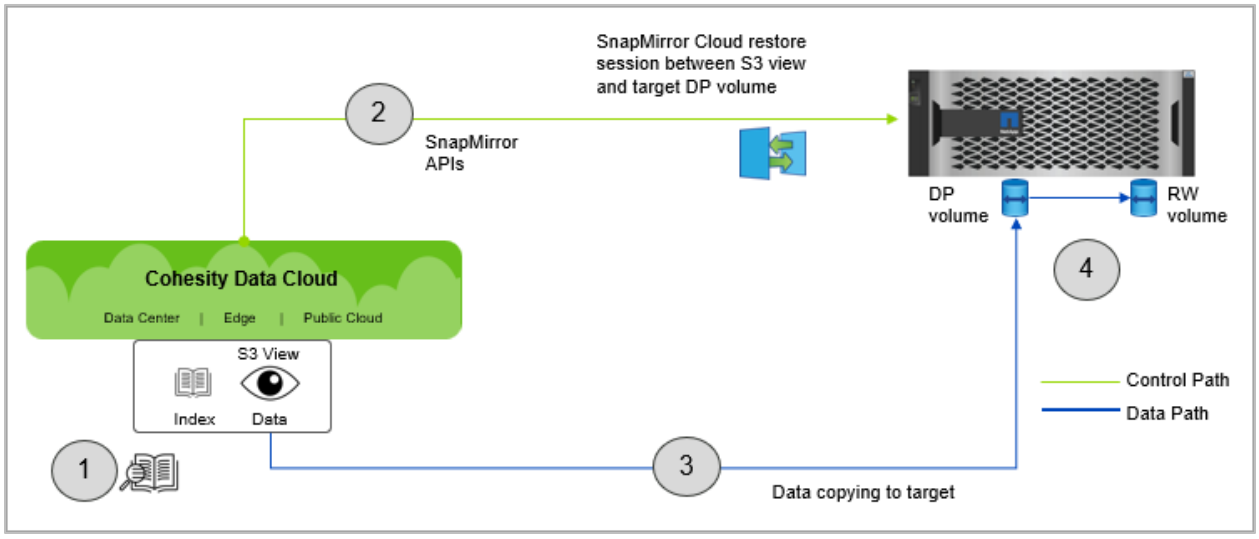
6. Prepare: Creates a clone of the data (an internal view) that needs to be recovered and mounts the target volume on the Cohesity nodes.
7. File Scan and File Read: Performs a file scan on the cloned internal view and reads the data.
8. File Restore: Performs writes on the target in batches, using distributed and parallel write mechanisms.

   NOTE: Before the write operation begins, the required minimum permissions on the target are checked for the user account that was used to register the NetApp NAS with Cohesity. If permissions are not in place, the recovery task fails.

See Appendix A: Restore Write Behavior for more details on recovery operations.

NAS Restore Internal Workflow of SnapDiff-based backup

Figure 14: NAS Restore Internal Workflow of SnapDiff-based backup



1. **Search Object**–Search for the volume\files to restore. Faster search results since indexing metadata is stored in the Cohesity internal database.
2. **Create Restore SnapMirror relationship**–RST SnapMirror relationship between the S3 view and the destination DP volume is created using SnapMirror APIs.
3. **Restore**–The entire content of the Snapshot copy selected to be restored is copied to the destination DP volume.
4. **Volume Type Conversion**–The destination volume is made read-write, and the RST SnapMirror relationship is deleted.

Recover NetApp Data Using Cohesity

Cohesity offers two levels of recovery granularity for NetApp data:

1. Recover Storage Volume
2. Recover Files or Folders

In each case, you can search for the specific data you need and choose how and where to recover it.

Recover NetApp Data from File-Runner-based backups

The following figures illustrate the phases and choices you encounter in a restore workflow of File-runner and Snapdiff based backup.

Figure 15: NAS Data Restore Decision Tree for File-Runner-based backup

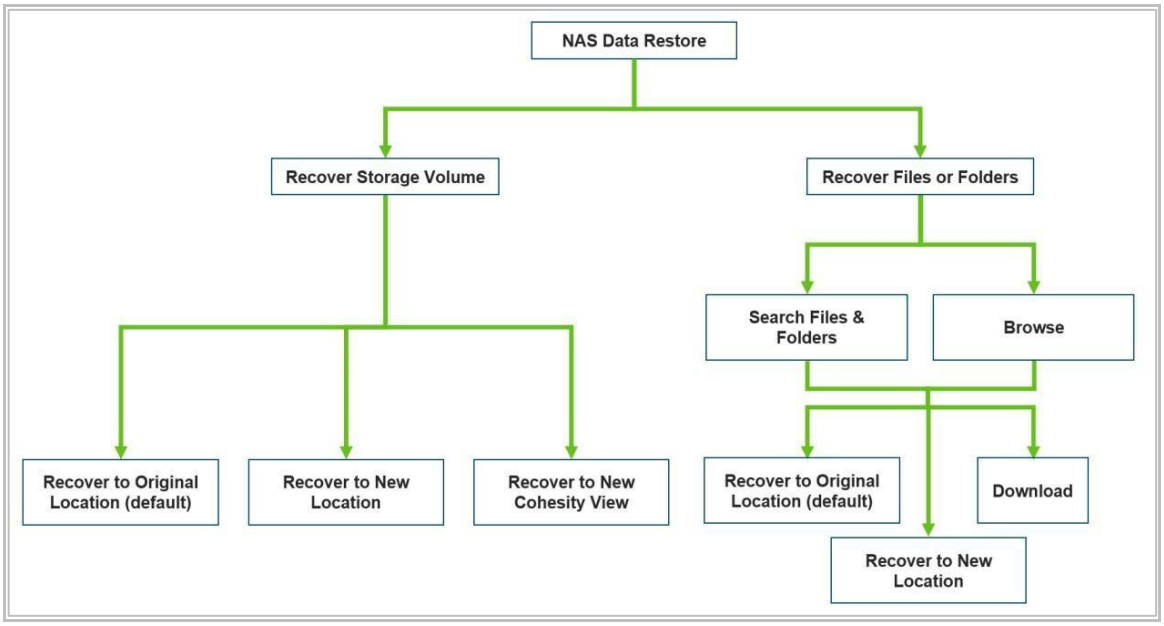
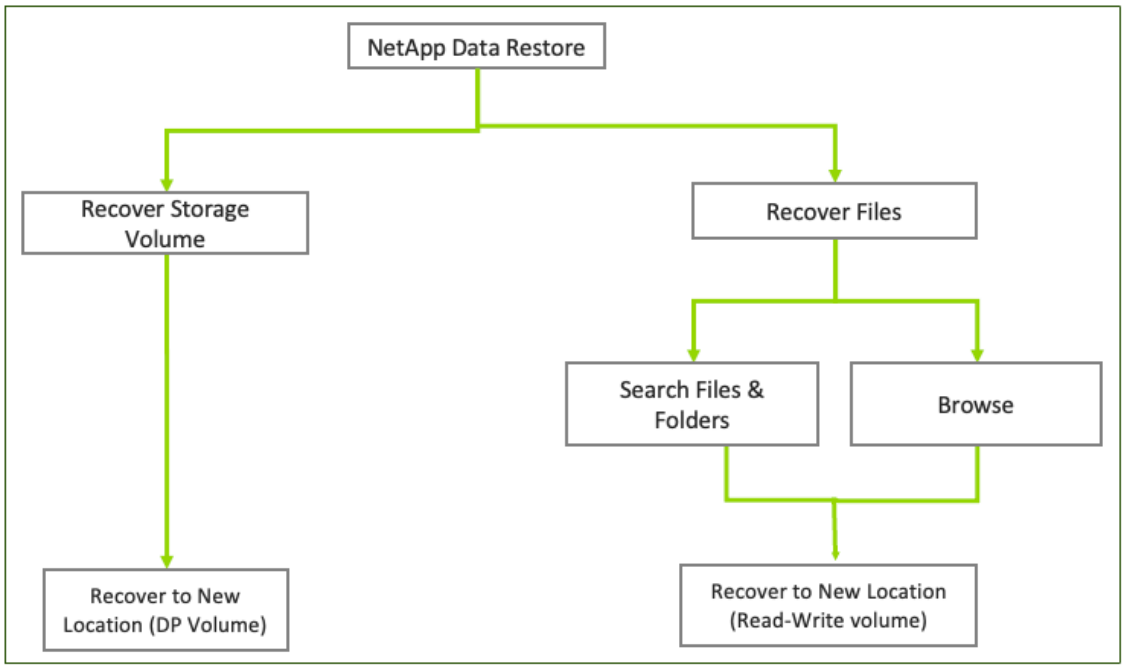


Figure 16: NAS Data Restore Decision Tree for SnapDiff-based backup.



Recover Storage Volume

Cohesit's *Recover Storage Volume* capability allows IT administrators to select and restore specific NetApp shares or exports from any previous backup. With this recovery approach, you can recover NetApp shares or exports to:

- The original location (default).
- A new location.
- A new Cohesity View.

See Recover Storage Volumes in the online Help for more details.

> **NOTE**: Before you can recover storage volumes, ensure that snapshots of those volumes exist in the Protection Groups on your Cohesity cluster.

Recover to Original Location (Default)

NOTE: Recovery to the original location is not supported when recovering from SnapDiff-based backups.

If your data becomes unavailable or is lost from the source from which it has been backed up but your original source infrastructure is still functional, you can recover it to your original NAS location (target) with this option.

NOTE:

- Data is recovered along with metadata, including permissions.
- Recovery to the Original location (Original source RW/DP Volume) is not supported when recovering from SnapDiff-based backups.
- Recovery to the Original NetApp cluster on a new DP Flex Volume is supported.

Recover to a New Location

Your data can be recovered to a new NAS location under circumstances such as source unavailability or source data migration. Using this option with File-Runner based backup, you can restore your NetApp NAS data easily and quickly to different NAS locations in the same or different sources like NetApp, Dell EMC, or any generic NAS. Since SnapDiff based backups are written to Cohesity in NetApp proprietary format, you can restore your NetApp NAS data to different NAS locations (DP volume) in the same or different NetApp source only.

NOTE:

- The target NAS must be registered with the Cohesity cluster as a source.
- The new location has to be a NetApp cluster registered as a source when recovering from a SnapDiff-based backup as the backup on Cohesity is written in NetApp proprietary format.
- When recovering a volume from a SnapDiff-based backup, the new location has to be a DP Flex volume and should not be involved in any SnapMirror relationship.
- When recovering a volume from a SnapDiff-based backup, the target DP Flex volume can only be used in one recovery task; after recovery, it will be converted to an RW Flex volume.

Recover to a New Cohesity View

This option presents you with the flexibility to repurpose the backup data without delay and disruption. You can instantly create dev/test environments from the backup as required because recovering to a Cohesity View doesn't require the data to be written to a location. When recovering to a Cohesity View, Cohesity clones the selected backup to a new View within the Cohesity cluster and provides you instant access to it.

NOTE:

- Recovery to a new Cohesity View is not supported when recovering from a SnapDiff-based backup. Hence, Instant Access to the backup data is not supported with SnapDiff backups.

Recover Files or Folders

Cohesity 's *Recover Files or Folders* capability allows IT administrators to search and recover specific files and whole folders from any previous File-Runner-based NetApp backup. When recovering from a SnapDiff-based backup, you can search and recover specific files only, as folder recovery is not supported.

For File-Runner-based backups– This feature allows you to restore the files or folders to their original location or to a newly specified location, which can be within the original source or a different one, without losing the original permissions and attributes. You can also download specific files and folders from any snapshot that was created by a Cohesity Protection Group.

For SnapDiff-based backups– This feature allows you to restore files to a new location, which can be within the original source NetApp Cluster or a different source NetApp Cluster, without losing the original permissions and attributes. The new location where the files will be recovered should be any RW Flex volume other than the original source volume where the files reside. In one recovery job, you can only recover eight files at a time. The download files feature is not supported with SnapDiff-based backups.

NOTE:

- Before you try to recover files or folders, ensure that snapshots of these files and folders appear on the Cohesity cluster in a Protection Group.
- For backed-up SMB volumes, the user account must have full control of the target, as Cohesity uses that user to perform recovery.
- You can also use the wildcard character * or narrow the search results by specifying filter criteria. For example, you can filter the search results by a specific Protection Group. Click the Add Filters icon, choose a filter, and click Add.

The next steps in the procedure depend on the type of search you select.

- To search by file or path name, see Search Files and Folders next.
- To browse, see Browse for Search below.

See Recover Files or Folders in the online Help for more details.

Search Files and Folders

The **Files or Folders** search option allows you to search and recover files using file or folder names from any backup available on Cohesity.

> **NOTE**: To use this search option, you need to enable indexing for the Protection Group. This ensures backup metadata is indexed and allows Google-like search, enabling instant granular file-level recovery to any point in time across billions of files.
>
> To enable indexing, see Appendix B: Index for Faster Granular-level Recovery.

Browse for Search

The **Browse** search option allows you to search and recover files using the name of the server or Protection Group from any backup available on Cohesity.

See Appendix A: Restore Write Behavior (File-Runner approach) for more details on recovery operations.

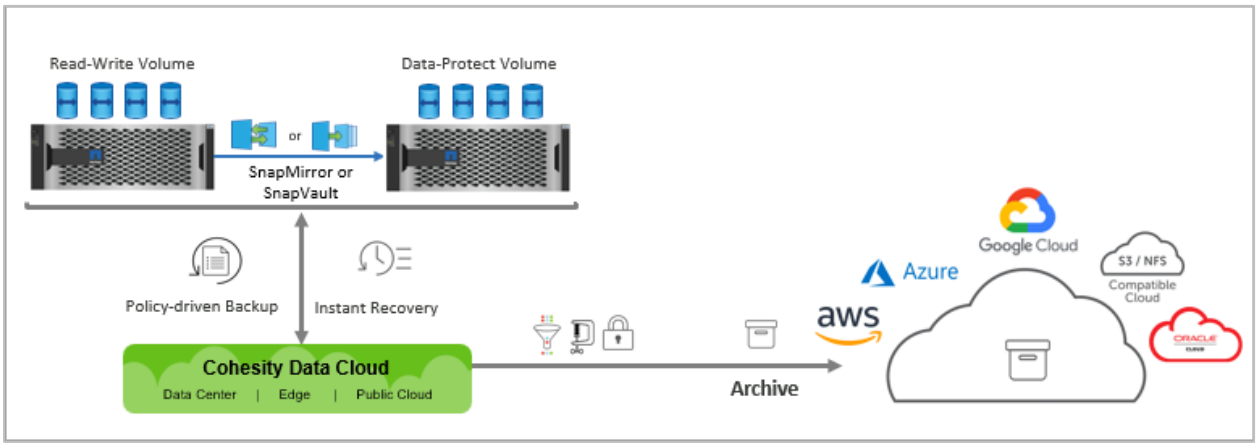
Use CloudArchive for Long-term Retention

The exponential growth of data volumes and the resulting IT management demands have prompted businesses to seek more cost-effective, reliable data storage and protection solutions. Cohesity provides a policy-based method to archive to public clouds (AWS, Azure, GCP), to any S3-compatible storage, tape, and/or to any NFS mount point. Cohesity CloudArchive offers a complete, self-contained copy of your backup, containing backup data, backup metadata, indexing data, and deduplication fingerprints.

NAS administrators can take advantage of Cohesity CloudArchive to address long-term data retention requirements. The archived data is efficiently transferred and stored by sending only deduplicated, compressed, incremental backups, thereby reducing network and storage utilization.

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

Figure 17: Leverage Public Cloud Infrastructure for Long-term Data Retention and Archival



CloudArchive - SnapDiff Considerations

1. As of release 7.1.2, Cohesity SnapDiff backups only support CloudArchive with an incremental forever format to an S3-compatible target, NAS, or Tape.
2. CloudArchive with Periodic full is not supported.
3. CloudArchive Direct is not supported.

Maintain Business Continuity with Disaster Recovery

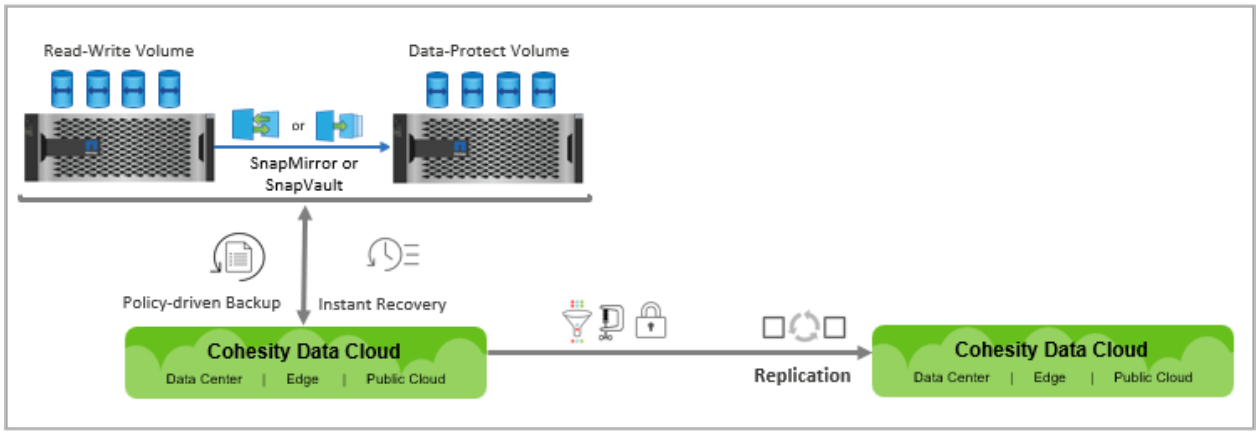
Cohesity provides two mechanisms for protecting your data from disruptions and disasters:

- **Replication** provides a simple way to store and retrieve data in the event of major business disruptions, such as natural disasters and IT failures.
- **CloudRetrieve** works with CloudArchive to restore your data to an alternate Cohesity cluster.

Replicate Backups to Other Cohesity Clusters

NAS administrators can take advantage of Cohesity 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, Cohesityalways performs source-side 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 fail over to the DR site for backup and recovery of their data.



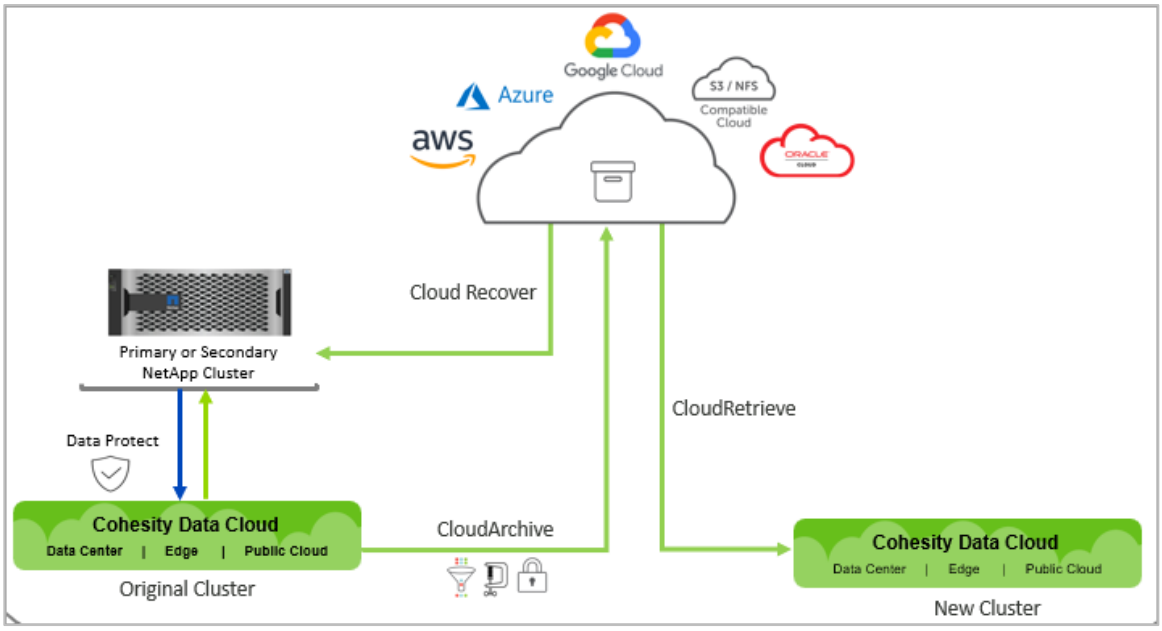
For more, see About Replication in the online Help.

Access Your Cloud-stored Data

Once the data is archived, NAS administrators can also take advantage of the Cloud Recover and CloudRetrieve features:

- **Cloud Recover** to source cluster: Recover entire objects to your original cluster.
- **CloudRetrieve** to new cluster: Retrieve your previously archived data onto an entirely new cluster as a cost-effective alternative for disaster recovery, geo-redundancy, and business continuity.

Figure 18: Cloud Recover to Original Source & CloudRetrieve to New Cluster



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

Best Practices for Protecting NetApp NAS

Some tips for getting the best performance from our solution for protecting your NetApp NAS data:

1. Prefix the usernames used for backup with ‘Cohesity’ for easy identification during troubleshooting scenarios.
2. Avoid adding the same shares to multiple Protection Groups.
3. Filter Ips to Deny the NetApp LIFs (Logical Interfaces) that are unreachable from Cohesity to avoid backup/restore failures.
4. Use NTP to avoid time skew as errors may occur for SnapMirror Cloud (SnapDiff Backup) if time skew is detected.
5. When configuring SnapDiff backup Protection Groups, ensure that the backup policy schedule is in sync or less aggressive than the NetApp snapshot (used for incremental backup) schedule to ensure the snapshot's existence for the incremental backup run.
6. When recovering 100s of files from SnapDiff backups; it is recommended to use volume-level recovery workflow.

Appendix A: Restore Write Behavior (File-Runner approach)

Restore Behavior with and without “Overwrite Existing File/Folder”

As you restore files and folders from your NetApp backups, the restore operations behave differently depending on the restore location, target volume, and the “Overwrite Existing File/Folder” option. See Table 4 below for details.

Table 4: Restore Behavior with and without “Overwrite Existing File/Folder”

RESTORE TO	SELECT VOLUME	OVERWRITE EXISITNG FILE/FOLDER	RESTORE BEHAVIOR
Original Location	N/A	Enabled	Data is restored to the original location. Any duplicate folders are merged, and any duplicate files are overwritten. Unique data on the target (data that is not present in the recovery point snapshot) is not touched during the restore task run.
		Disabled	Data is restored to the original location. Any duplicate folders are merged, and any duplicate files are skipped. Unique data on the target (data that is not present in the recovery point snapshot) is not touched during the restore task run.
New Location	Same as the source volume	Enabled	Same behavior as "Restore to Original Location."
		Disabled	Same behavior as "Restore to Original Location."
	Alternate volume	Enabled	Data is restored to the alternate location. Any duplicate folders are merged, and any duplicate files are overwritten. Unique data on the target (data that is not present in the recovery point snapshot) is not touched during the restore task run.
		Disabled	Data is restored to the alternate location. Any duplicate folders are merged, and any duplicate files are skipped. Unique data on the target (data that is not present in the recovery point snapshot) is not touched during the restore task run.

Appendix B: Index for Faster Granular-level Recovery

Cohesity allows you to index the backup content, which enables you to get faster search results even when you are searching for a single file among billions of files.

Indexing scans the backup data with simple, text-based search-and-restore functionality to find files quickly, thus enabling instantaneous data retrieval from the backed-up snapshots.

Improved Indexing

Cohesity's indexing engine performs incremental indexing to deliver better performance and faster results.

- **Indexing in File-Runner backup–**In Cohesity's File-Runner-based backup approach, the indexing engine only scans the data that has changed since the previous snapshot. This is much less resource-intensive and results in faster indexing.
- **Indexing in SnapDiff backups–**Cohesity communicates with NetApp using SnapDiff APIs to query for a list of changed files between current and previous SnapMirror snapshots for backup. After the SnapMirror data transfer to the Cohesity S3 view is complete, the SnapDiff APIs send the change list of the files that were backed up to Cohesity for indexing.

Enable Indexing

To enable "Files and Folders" search, enable **Indexing** in the Protection Group. See Create a Protection Group for NAS Volumes in the online Help.

Once indexing is enabled in the Protection Group, everything in the protected SMB/NFS share is indexed by default. However, because indexing is resource-intensive, we recommend excluding the files and directories where file-level recovery is not required, such as scratch spaces, binaries, temp files, etc. You can limit the indexing to a specific set of files and directories by defining them in the **Include** and **Exclude** settings under **Indexing** in the Protection Group settings.

NOTE: In a SnapDiff Protection Group, you can enable Indexing by toggling the Indexing option under Advanced settings of the Protection Group. However, Indexing rules to Include\Exclude files is not supported with SnapDiff.

NOTE: During a Protection Run, the indexing engine scans the objects (NAS volume or file/folders) that were successfully backed up and skips any objects that failed to back up. The indexing engine runs for all successfully backed-up objects, even if a Protection Run is completed with warnings.

Your Feedback

Was this document helpful? Send us your feedback!

About the Authors

Sadik Sayed is a Staff Technical Solutions Engineer at Cohesity. In his role, he focused on VMware data protection, NAS data protection, and SiteContinuity.

Other significant contributors included:

- Adaikkappan Arumugam, Sr. Manager, Technical Marketing
- Aditya Vasudevan, VP, Customer Success
- Rich Kuhn, Product Solutions

Document Version History

VERSION	DATE	DOCUMENT HISTORY
1.3	May 2025	Added SnapDiff best practice section Added FileRunner Best Practice and Considerations
1.2	May 2024	SnapDiff backup updates, other minor changes
1.1	July 2021	Cohesity rebranding updates
1.0	Aug 2023	First 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 42 of the Fortune 100.

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