COHESITY

Version 1.1

November 2021

Bootstrap Your Archive to GCP with Cohesity & Google Transfer Appliance

Seed Your Data to the Cloud with Cohesity & Google Transfer Appliance

ABSTRACT

Transferring large amounts of data to the cloud can be fraught with network bandwidth and connectivity hiccups. Cohesity's partnership with GCP offers a solution that meets your ever-growing and ever-evolving data needs. Cohesity's integration with Google Transfer Appliance gives you the power to move massive amounts of data safely and securely.

Table of Contents

Figures

Tables

The Need for Google Transfer Appliance

As organizations become overwhelmed with the exponential growth of data they must store and keep secure, it is more and more urgent for them to find other, more efficient solutions for storing large amounts of data. The most compelling of those alternatives is cloud storage. Cloud storage is a scalable and highly available storage solution with well-defined Recovery Point Objectives (RPOs) and can be used for long-term data retention, which is often mandatory to meet a host of compliance, regulatory, and legal requirements. However, archiving your data to the cloud comes with some challenges. Although cloud storage saves organizations in costs and operational overhead, once your data scales to hundreds of terabytes or petabytes, it can become a challenge to move the data to the cloud.

The main questions that IT administrators face with protecting their data and leveraging the cloud at the same time are:

- How do I get my existing data to the cloud in the first place?
- Is my data network robust enough to meet the initial data-transfer requirements?
- How long will it take for the initial data transfer to the cloud?

An integral part of reducing your data transfer requirements is to reduce the amount of data you need to send to the cloud in the first place. Cohesity leverages its robust global deduplication and compression capabilities to dramatically reduce the amount of storage required to protect your data.

For example:

You want to transfer 100 TB of data to the cloud with approximately 100 Mbps of bandwidth dedicated to data transfer. Assuming no connectivity hiccups, it will still take over 100 days to transmit your data over that network connection. For many organizations, that's simply too long.

Table 1 outlines some approximate data-transfer times, based on data size and network bandwidth.

Table 1: Data Transfer Times by Size & Network Speed

NETWORK BANDWIDTH	DATA SIZE			
	100 TB	200 TB	400 TB	500 TB
100 Mbps	124 days	248 days	1.35 yrs	1.69 yrs
1 Gbps	12 days	24 days	48 days	60 days
10 Gbps	1.25 days	2.5 days	5 days	6.25 days
100 Gbps	3 hrs	6 hrs	12 hrs	15 hrs

Google Transfer Appliance (TA) is a high-capacity storage device that enables you to securely transfer and ship your data to Google Cloud Storage in Google Cloud Platform (GCP). Using your Internet connection to upload large amounts of data directly to the cloud can take weeks before your data is available. Using data-transfer appliances like Google TA can shorten that timeline exponentially.

Read the chapters below to learn the steps and best practices to seed initial data from your Cohesity cluster to GCP using a Transfer Appliance. Once you do, you'll be able to resume your Cohesity

incremental archival runs to GCP. In the Appendices, you'll find useful details and learn how to consolidate your data by seeding it from multiple devices to a single Google Storage bucket.

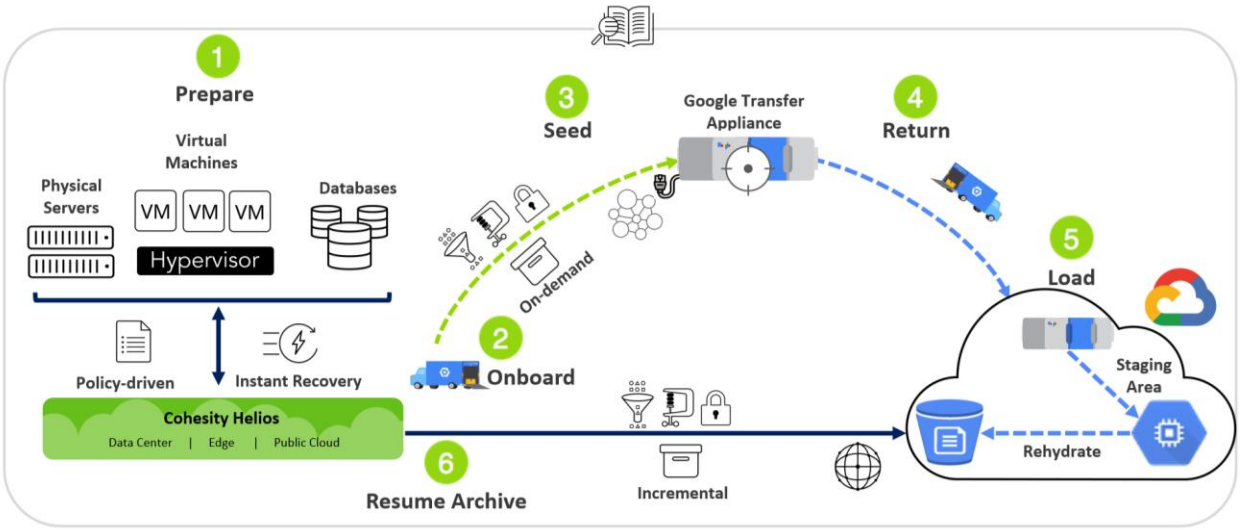
Use Google Transfer Appliance with Cohesity

Google Transfer Appliance is a physical data-transport device that you can order from your Google Cloud Console and is best suited for IT administrators who need to move terabytes of data into the GCP faster than they can over the Internet. The data stored in a Google Transfer Appliance is AES256-bit encrypted. The device comes with two configuration options—100 TB or 480 TB of raw capacity for large-scale data transfer to the cloud.

When you return the Transfer Appliance, GCP initiates an importing process to transfer the data from the Transfer Appliance to a temporary storage platform. The data is then transferred to your Cloud Storage bucket when you initiate the rehydration process.

When the rehydration process is complete, validate the data in the Google Storage bucket to ensure data integrity and then resume the incremental archival of data, but now to the Google Storage bucket instead of the Transfer Appliance.

Figure 1: Seed Your Cohesity Archive with the Google Transfer Appliance



The seeding process involves several distinct phases:

1. **Prepare**. Protect your data in Cohesity and order the Transfer Appliance from GCP.

   NOTE: During the Protection Run, Cohesity's advanced deduplication and compression dramatically reduce storage consumption. When the Protection Run completes, you can retrieve the size of the protected data and decide how many Transfer Appliances to order.

2. **Onboard**. Connect and configure the Transfer Appliance.

   NOTE: When you configure the Transfer Appliance, you will have to create a password and passphrase that are used to encrypt the data with an AES 256 algorithm on the Transfer Appliance (whether or not it is already encrypted by Cohesity).

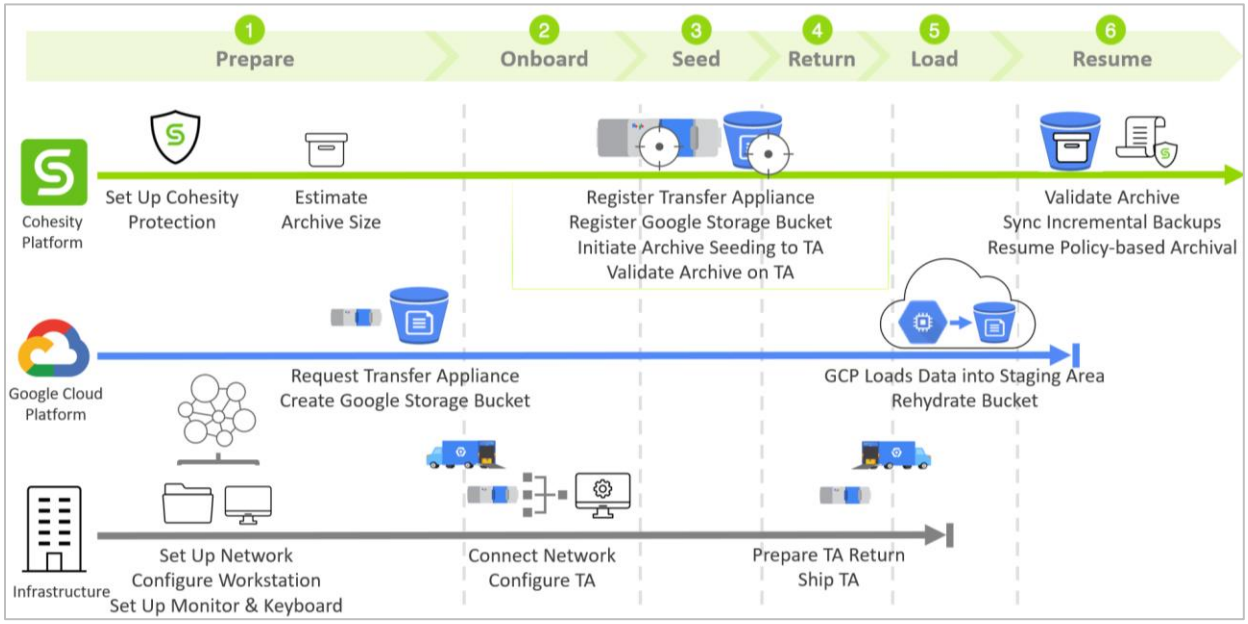
3. **Seed**. Register the Transfer Appliance as an External Target in Cohesity to seed it with your protected data. In most cases, customers find that they need to transfer the full backup followed by several incremental runs to ensure that the seeded data is as up to date as possible.

4. **Return**. Send the Transfer Appliance back to GCP's upload facility and initiate the External Target swap.
5. **Load & Rehydrate**. When GCP receives the Transfer Appliance, they load the data to their cloud staging area. Once the data is loaded, you will use the passphrase and password you created earlier to rehydrate (decrypt and move) it into your Google Storage bucket.
6. **Resume**. When the rehydrating process is complete, validate the data in the Google Storage bucket to ensure data integrity, and then resume the scheduled incremental archival of the data to that Google Storage bucket.

Review Your Archive Seeding Plan

Archival to the cloud using a seeding device involves coordination between several internal and external components. It is important to review the phases involved to optimize the archival process. When you identify the dependencies between each phase, you can plan the activities that can be executed in parallel.

Figure 2: Phases of the Archive Seeding Process



Archiving your data to GCP using a Transfer Appliance involves three components:

- Cohesity platform
- GCP
- Infrastructure

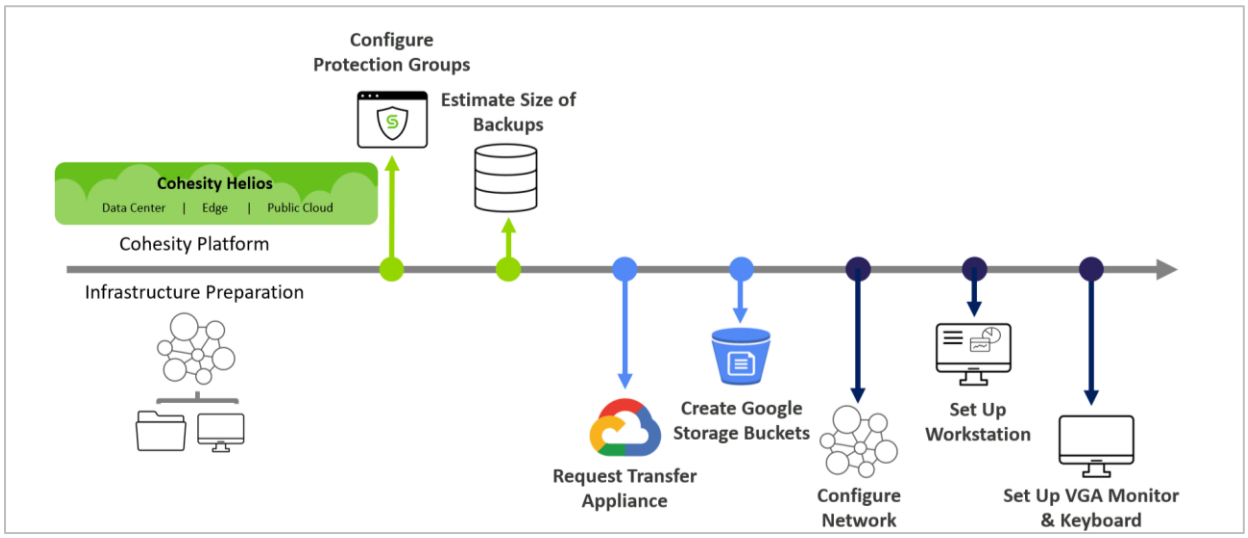
The process is chronologically organized into six phases: **Prepare, Onboard, Seed, Return, Load & Rehydrate**, and **Resume**.

Figure 2 above gives you a timeline view of the activities involved in the process. Within each phase, several activities can be executed in parallel. For Example, in the Prepare phase, you can get Cohesity protection for your data set up and started while you set up the network.

Prepare to Seed Transfer Appliance

To maximize the benefit of using the Google Transfer Appliance, we recommend that you have the facilities in place before receiving the device. Doing so helps you keep the whole process as short as possible.

Figure 3: Prepare Software and Infrastructure for Transfer Appliance



To prepare for seeding data to GCP using the Transfer Appliance:

1. Set up Cohesity protection for your data.
2. Estimate the size of the data to be archived.
3. Request a Transfer Appliance from the GCP Management Console.
4. Prepare your data center infrastructure.

Set Up Cohesity Protection for Your Data

You prepare Cohesity platform by identifying the data objects that need to be archived and backing them up by adding them to a Cohesity Protection Group. In addition to being protected with backups, your data will be compressed and globally deduplicated, resulting in a dramatically reduced storage footprint and improved data transfer times. Once the first Protection Run completes, you can determine the storage consumption of your protected data objects, which will help you estimate the size of data that will be transferred to your cloud archives via the Transfer Appliance.

NOTE: If you plan to use an existing Cohesity Protection Group, you can skip the steps in this section.

To set up Cohesity protection:

1. **Register Sources**. Identify the data objects that need to be archived and register them as Cohesity Sources. See Manage Sources in the online Help.
2. **Choose a Protection Policy**. If necessary, create a new Protection Policy or edit an existing one to set the frequency of the Protection Runs and the retention period for the snapshots. For details, see Create or Edit a Standard Policy in the online Help.

   > **NOTE**: To avoid shipping and other delays causing snapshots to expire during the process, add a few weeks of buffer to the retention period of the backups. You can reset the retention period later.

   > **IMPORTANT**: Ensure that the Protection Policy you use has no archival schedule.
3. **Create a Protection Group**. Add the Cohesity Sources to be protected to a new Protection Group and attach the Protection Policy.

   > **NOTE**: To parallelize and streamline the process, you can create separate Protection Groups for different Sources.
4. **Validate Backup**. Confirm that the first Protection Run of each Protection Group involved was a successful, full backup.

   > **NOTE**: For a list of supported workflows, see Supported Workflows and External Targets in the online Help.

Estimate Data Size for Archival

Before you can request a Transfer Appliance, you need to estimate the size of the data to be archived. This estimation is needed to determine the number of Transfer Appliances you'll have to order.

Some guidelines that can help you in estimating the data size:

- The storage metrics for each Protection Run of a Protection Group are updated in Cohesity.
- Identify the storage consumed by the full backup runs of the Protection Group configured in the previous step.

  > **NOTE**: In addition to the full backups, account for any incremental runs you plan to archive to the Transfer Appliance, as well.
- Add a 10% buffer to your estimate.
- If you have multiple Protection Groups, aggregate the storage consumed by all the Protection Groups you plan to transfer via the Transfer Appliance.

  > **IMPORTANT**: You can't seed a Protection Group across multiple Transfer Appliances. If the estimated archival size for a particular Protection Group exceeds the maximum storage capacity of the Transfer Appliance, split it up into smaller Protection Groups.

Request Google Transfer Appliance

It is a multi-step process to import data from your on-premises data center to the Google Storage bucket. The process starts by checking the requirements and your environment and then placing a request for the Transfer Appliance. From the GCP console, you will be tasked to plan the complete import process.

To prepare for the Google Transfer Appliance:

1. Ensure your environment meets the specific requirements for a Transfer Appliance. Carefully read and follow all of the instructions in Before you begin in the Google Cloud documentation.
2. Log in to Google Cloud Platform.
3. Choose the project in which you want the data to be moved.
4. From the Navigation menu, under **Storage**, select **Data Transfer** > **Transfer Appliance**.
5. Determine the amount of data to be moved to GCP. To calculate the total capacity you will need, see Estimate Data Size for Archival above.
6. Submit a Request. Google will process your request and reach out via contact email.

Create Google Storage Bucket

The goal of archiving data using Google's Transfer Appliance is to move data from an on-premises data center to Google Storage. It is imperative to plan and create the Google Storage buckets in advance to ensure a smooth rehydration process. The number of storage buckets required depends on how you plan to maintain the data in the cloud. For example, you can consolidate the data into a single storage bucket for a simplified management experience or globally distribute the data in different storage buckets for compliance and regulatory reasons.

To create a Google Storage bucket:

1. Log in to Google Cloud Platform.
2. Choose the project from which you issued the request for the Transfer Appliance(s).
3. From the Navigation menu, under **Storage**, select **Storage** > **Browser**.
4. Select **Create Bucket** and follow the instructions.
5. Repeat Step 4 to create additional storage buckets.

For more, see Creating storage buckets in the Google Cloud documentation.

Prepare Your Data Center for Transfer Appliance

It takes at least 2-3 business days for GCP to respond to the request and at least 1-2 business days for the Transfer Appliance to be delivered. We can use this time to prepare the infrastructure required to boot the Transfer Appliance when it arrives. This step will optimize the time it takes to onboard the device.

To prepare your infrastructure for the Transfer Appliance:

1. **Set up the network**.
    a) Allocate a dedicated IP address for the Transfer Appliance.
    b) Connect the Transfer Appliance to your network via Ethernet or fiber-optic network cables.
    c) Configure your network firewalls for communication between Cohesity, a client workstation, and the Transfer Appliance.
    d) Optimize network performance by configuring minimum network segments (hops) between the source and the Google Transfer Appliance.
2. **Dedicate a workstation or server** with a recent version of the Chrome browser to access the Transfer Appliance's Web User Interface.
3. **Connect a VGA monitor and USB Keyboard** directly to the Transfer Appliance.

For detailed configuration instructions, see Acquiring the necessary hardware in the Google Cloud documentation.

Set Up Transfer Appliance

When the Transfer Appliance is delivered, inspect the device for any damage or tampering. If you find evidence of damage or tampering, contact Google Cloud support for a replacement. Once the inspection is complete, unpack the device and configure it to be a NAS endpoint. The data from Cohesity will be seeded to this endpoint.

To configure the Transfer Appliance:

1. Unpack the device. Follow the steps in Unpack Transfer Appliance in the Google Cloud documentation.
2. Onboard the appliance. In this step, you need to create an encryption password and passphrase to seed data to the Transfer Appliance. See Configuring Transfer Appliance in the Google Cloud documentation. (Later, once GCP has loaded the data into the cloud, you will enter those credentials to rehydrate your data into your Google Storage bucket.)
3. Assign the IP address to allow a connection to the appliance. See Setting up an IP address in the Google Cloud documentation.
4. Use the NFS Export data capture method to seed data to the Transfer Appliance. In this method, you mount the NFS export by exporting the directories on the appliance via NFS. See Mount NFS exports in the Google Cloud documentation.
5. Fetch the exported NFS endpoint from the Transfer Appliance Web User Interface.

   Endpoint Information:

   ```
   NAS Host: IP address assigned to the appliance
   Mount Path :/nfs/<SHARE_NAME>
   ```

   For example:

   - NAS Host: **192.168.5.10**
   - Mount Path If Share Name is transfer-appliance=nfs-archive-seeding. Then the Mount Path is:**/nfs/transfer-appliance=nfs-archive-seeding**

Seed Transfer Appliance with Your Protected Data

Once the Transfer Appliance setup is complete, you can register it as a NAS External Target in Cohesity. After the Transfer Appliance is registered, you will edit the Protection Group you created above to add an on-demand archival task that seeds the Transfer Appliance with the data in that Protection Group.

For the initial seeding of the Protection Group to the Transfer Appliance, the Protection Policy includes *only a backup schedule and no archival schedule*. The Transfer Appliance is seeded using the on-demand archive feature in the Protection Runs of the Protection Group. However, to resume archival after GCP finishes loading and rehydrating your data from the Transfer Appliance to the Google Storage bucket, you will edit the Protection Policy to *add the archival schedule* and specify the corresponding Google Storage External Target.

To seed the data from Cohesity to the Transfer Appliance:

- Register the Transfer Appliance and the Google Storage bucket as Cohesity External Targets.
- Initiate an on-demand archival run to the Transfer Appliance.
- Validate the data in Transfer Appliance.

Register External Targets

There are two endpoints you need to register as External Targets with Cohesity:

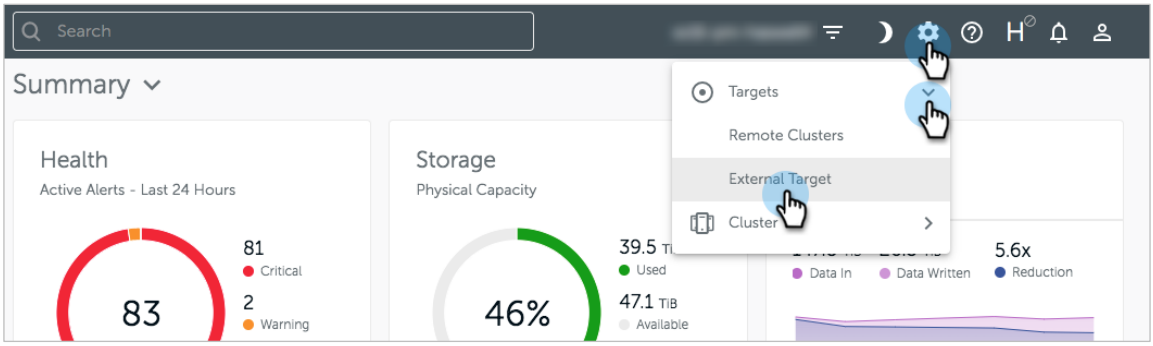
- **Transfer Appliance**, to move your data to it.
- **Google Storage bucket**, the long-term cloud home of your archived data, to resume incremental archives once GCP loads your data from the Transfer Appliance into your Google Storage bucket.

Register Transfer Appliance as NAS External Target

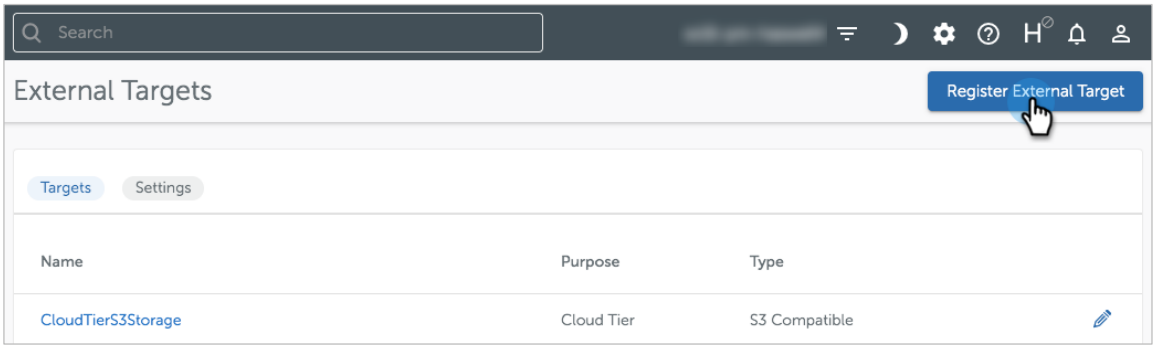
To register the Google Transfer Appliance as a NAS External Target in Cohesity, you will need the Transfer Appliance's endpoint (NAS host) details and the mount path details you fetched while setting up the Transfer Appliance.

To register the Transfer Appliance as a NAS External Target:

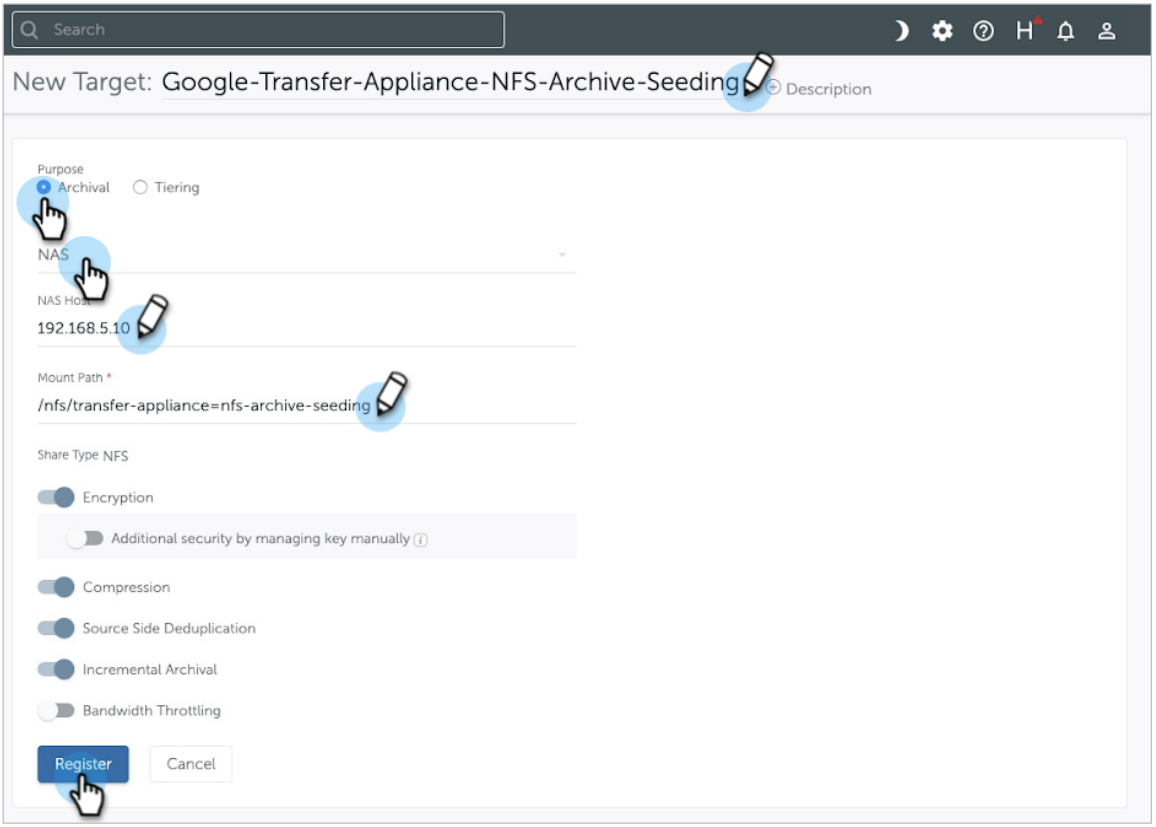
1. Log in to Cohesity platform.
2. Under the **Settings** gear icon, select **Targets** > **External Target**.



3. Click **Register External Target**.



4. In the form that opens, enter the **New Target** name and other parameters that you fetched from the Transfer Appliance earlier, and then click **Register**.
    - **Purpose**: Archival
    - **Type**: NAS
    - **NAS Host**: The IP address that is assigned to the Transfer Appliance
    - **Mount Path**: The path to the share name



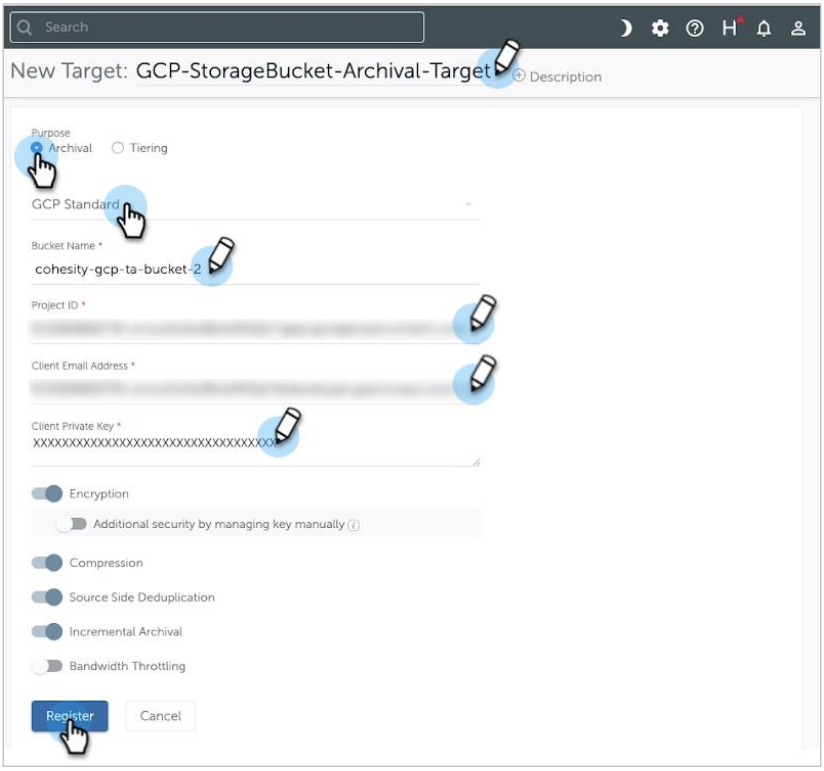
Register Google Storage Bucket as Google Storage External Target

The data seeded to the Transfer Appliance will be loaded by GCP. After that, you will rehydrate your data into the Google Storage bucket that you created during the Create Google Storage bucket step. This Google Storage bucket is registered as a Cohesity External Target to ensure the continuous archival of data. Find the access and endpoint details of the Google Storage bucket in your Google Cloud Platform console.

To register the Google Storage bucket as a Cohesity External Target:

1. Log in to Cohesity platform.
2. Under Settings, select **Targets** > **External Target**.
3. Click **Register External Target**.

4. In the form that opens, enter the **New Target** name and other parameters as specified, and then click **Register**.
    - **Purpose**: Archival
    - **Type**: Select GCP and choose the appropriate bucket type. See Appendix A: Supported Google Storage Classes.
    - **Bucket Name**: Enter the bucket name. The data from Cohesity will be archived to this bucket after the initial seeding is completed.
    - **Project Id**: Fetched from GCP console
    - **Client Email Address**: Fetched from GCP console
    - **Client Private Key**: Fetched from GCP console



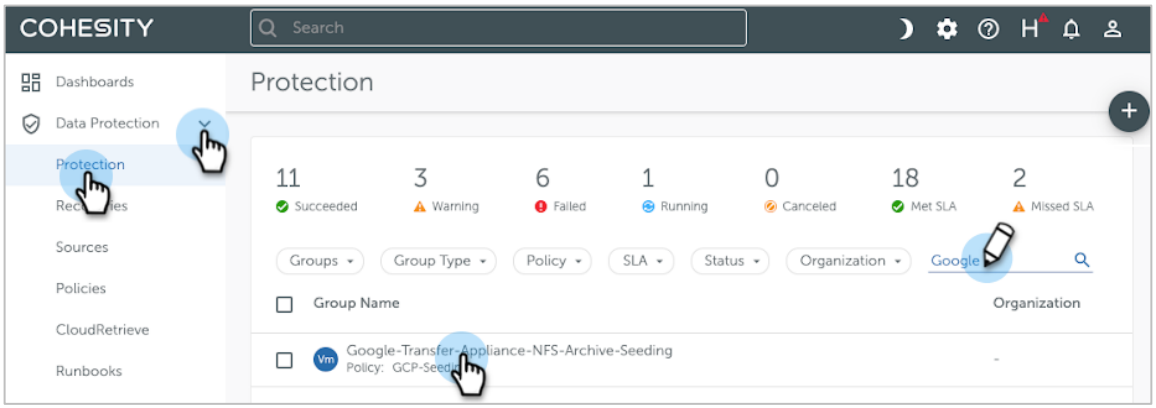
Initiate On-demand Archive Seeding to Transfer Appliance

Once the Transfer Appliance is registered as a Cohesity (NAS) External Target, you can archive data directly to it on demand.

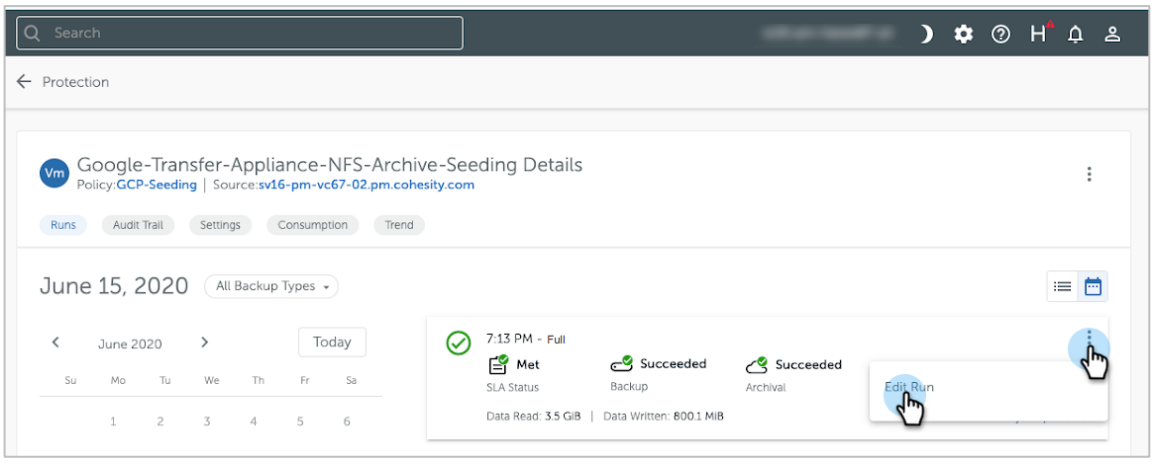
To start an on-demand archival run:

1. Log in to Cohesity platform.

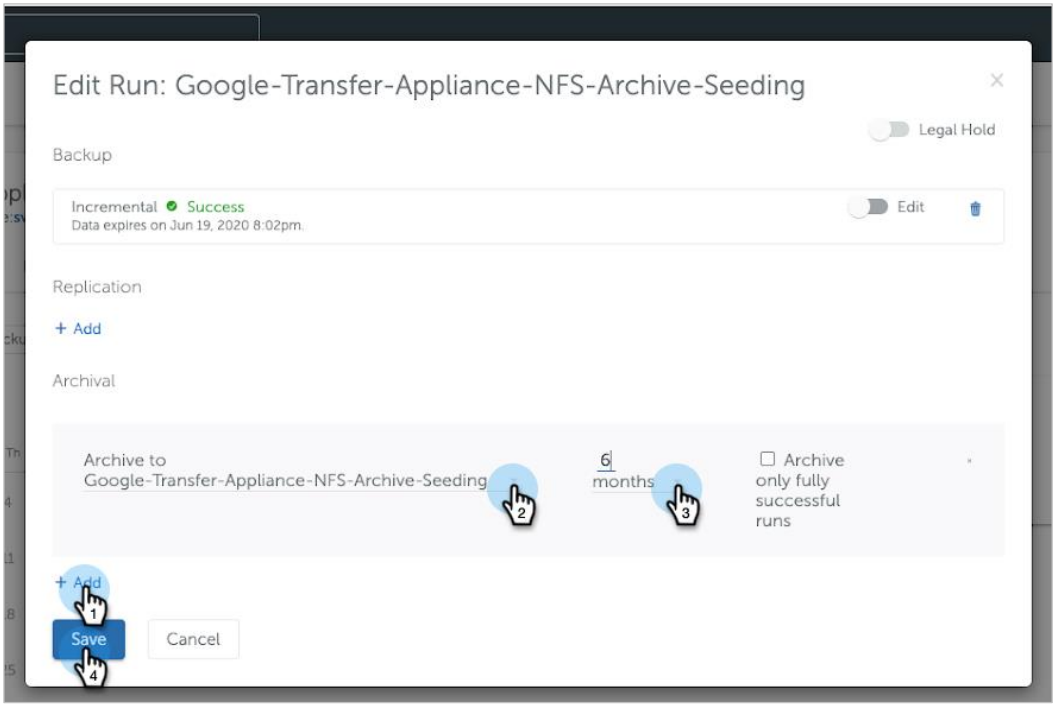
2. Navigate to **Data Protection** > **Protection** and click the Protection Group you created earlier. (Use the **Search** box to narrow the results.)



3. In the Protection Group, select the first successful **Full** backup run, click the **Action** menu and select **Edit Run**.



4. In the dialog, under **Archival**, click **Add** and select the Transfer Appliance External Target as the **Archive to** endpoint. Add the retention period to the archival task and click **Save**.



> **TIP**: Add some buffer to the retention period of the archival. For example, if your planned retention is 90 days, add 30 days, for a total retention of 120 days. You can reset the retention period later.

5. When the archival of the full backup is complete, check the Protection Group for possible errors and warnings. If the archival task fails, you will need to clean up the Transfer Appliance before you retry. See Appendix B: Clean Up Transfer Appliance for Retry for instructions.
6. If you need to, you can initiate the on-demand archival for the remaining incremental Protection Runs in the Protection Group. Repeat the previous steps to add the incremental protection runs to the Transfer Appliance.

> **IMPORTANT**: Contact Cohesity Support to complete this verification step before initiating on-demand archivals to the Transfer Appliance.

If there are other Protection Groups that need to be archived to the Transfer Appliance, repeat this procedure for each Protection Group.

Validate the Data in the Transfer Appliance

It is important to validate the data in the Transfer Appliance after the archival run to the device completes. This validation step ensures there is no corruption in the data seeded to the Transfer Appliance, and, if there are discrepancies, gives you an opportunity to start over.

> **NOTE**: Contact Cohesity Support to complete this data-validation step.

Return, Initiate Target Swap, and Rehydrate Data

When all the data that was marked for archival is seeded to the Transfer Appliance, prepare the appliance to be returned to the Google Cloud upload facility.

For a smooth return process, ensure the data transfer to the Transfer Appliance is finished and the validation step completed, and:

1. Make note of the Application ID. The Application ID is available in the email from Google that contains the login information. You will use this ID when you rehydrate the data to select the appliance that was used to capture the data.
2. Finalize the data. Run the Finalize job from your GCP console to remove the encryption credentials from the memory of the Transfer Appliance and to perform other internal processes. See Preparing an appliance for shipping in the Google Cloud documentation.
3. Shut down the Transfer Appliance.
4. Ship the Transfer Appliance back to GCP.

Initiate External Target Swapping

Once the Transfer Appliance is returned and the Google Storage bucket is loaded, Cohesity will be able to resume archiving, but going forward now, to the Google Storage bucket and not the Transfer Appliance. To be able to pick up where you left off, you need to ask Cohesity Support to help you replace the Transfer Appliance External Target with the External Target for the Google Storage bucket, known as External Target swapping.

Because it takes at least 2-3 days after you ship the device back for GCP to receive it and load the data into the Google Storage bucket, you can take advantage of this window and start the process of swapping External Targets, to ensure continuous incremental archival runs after the seeding process completes.

To initiate External Target swapping, contact Cohesity Support.

Rehydrate Your Data and Validate Archive in Storage Bucket

When seeding is complete, you ship Transfer Appliance back to Google Cloud Platform. When Google receives the Transfer Appliance, the encrypted data is first loaded to the cloud staging environment. Once the data is in the staging environment, it's ready to be rehydrated into your Google Storage bucket.

The rehydration process decrypts the data using the same password and passphrase that you created when you configured the Transfer Appliance. As the data is rehydrated, it is moved to the assigned Google Storage bucket. The process of data rehydration and validation involves:

1. Launch and configure the rehydrator instance
2. Initiate the rehydration process
3. Validate the archived data
4. Run a post-hydration cleanup

Launch and Configure a Rehydrator Instance

The first step in the rehydration process is to launch and configure a Rehydrator instance, which is a virtual Compute Engine instance that is hosted on Google Cloud Platform.

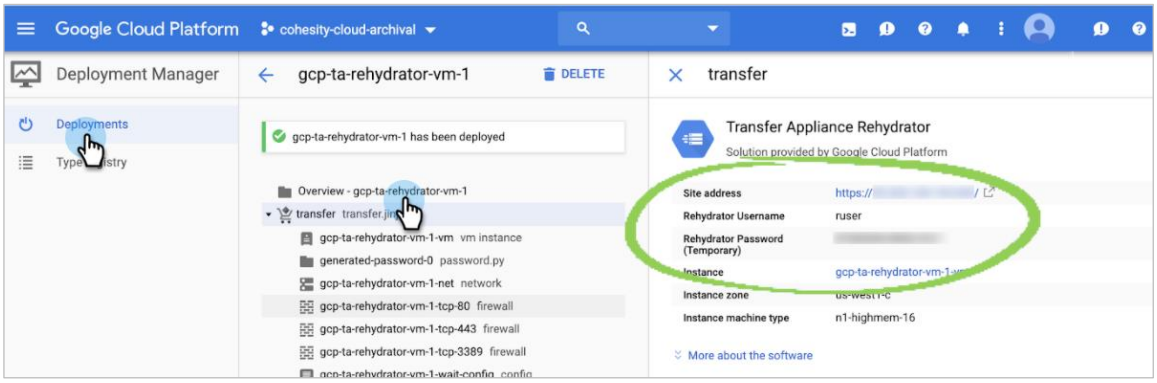
See Launching and configuring a rehydrator instance in the Google Cloud documentation.

Initiate Rehydration Process

Once the rehydrator instance is configured, you can launch the Transfer Appliance Rehydrator User Interface and start the rehydration job. In this process, the data is rehydrated into your Cloud Storage bucket.

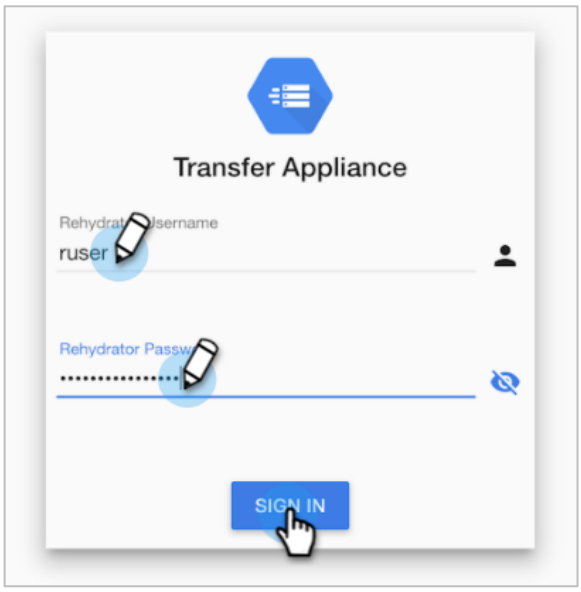
To initiate the rehydration process:

1. Fetch the **Site address**, **Rehydrator Username**, and **Rehydrator Password** from the rehydration instance.

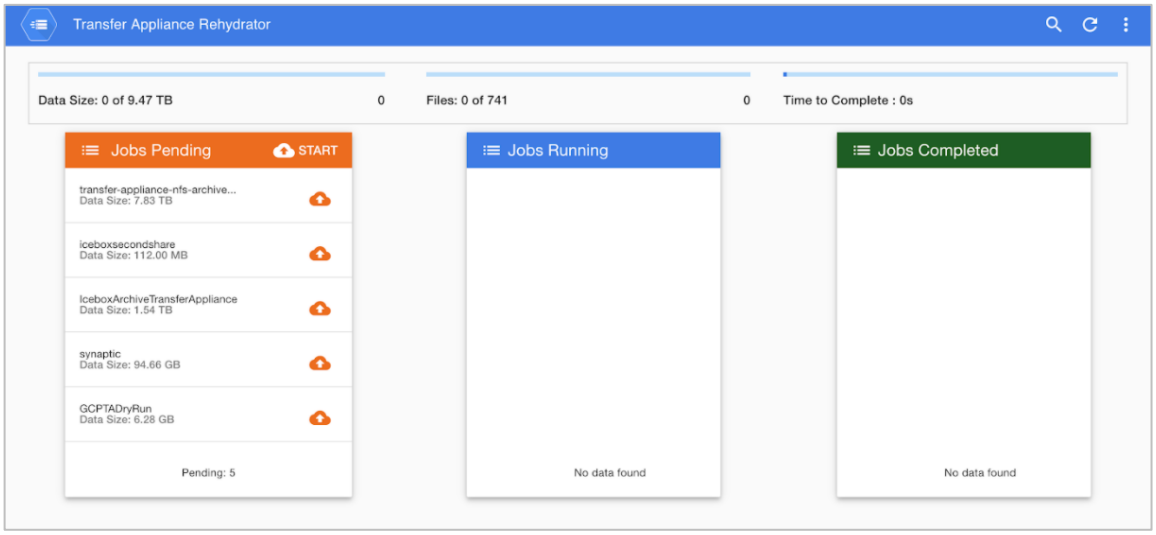


2. Launch a browser and navigate to the **Site address** to open the Transfer Appliance Rehydrator User Interface.

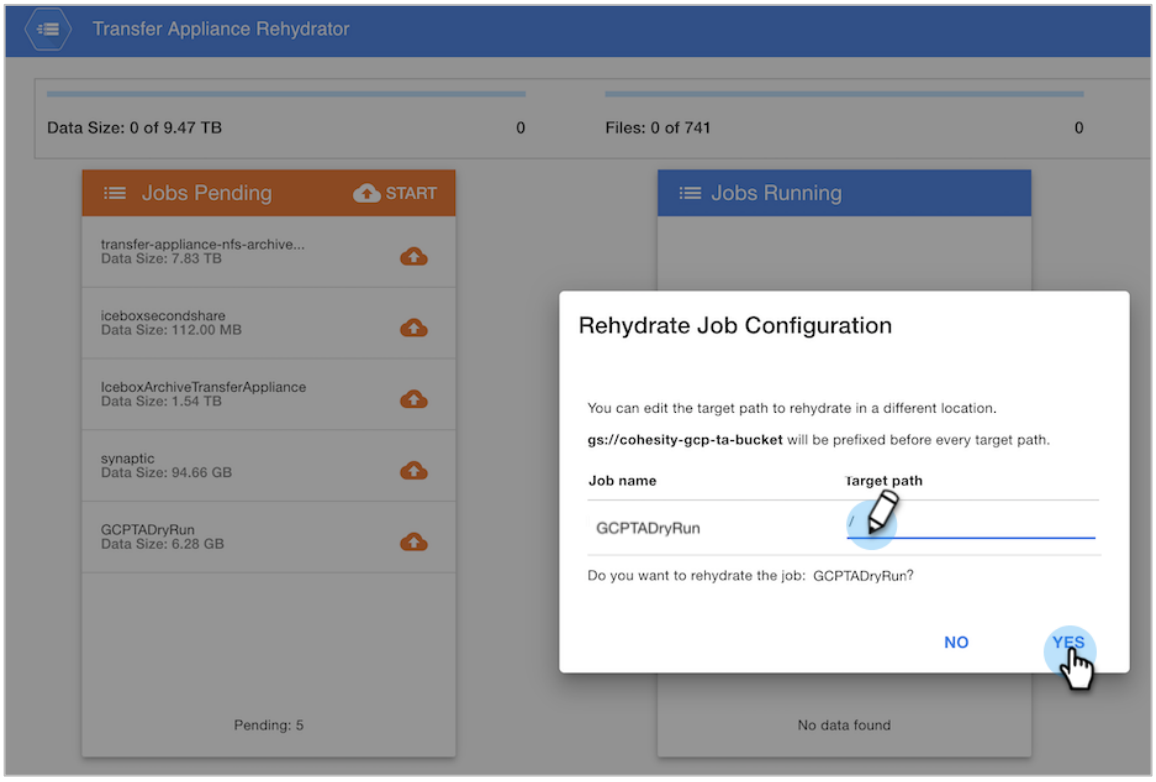
3. Sign in with the **Rehydrator Username** and **Rehydrator Password**.



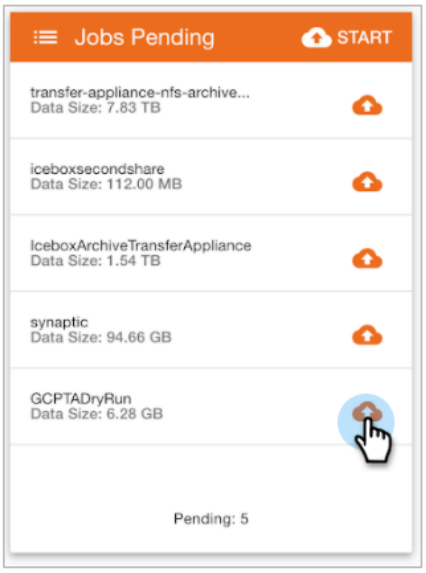
4. Once you log in, you can find the archived jobs listed in the **Jobs Pending** pane. Select the rehydration jobs you want to run from the list.



5. In the dialog that appears, provide the **Target path**. The data will be rehydrated into the Cloud Storage bucket and folder that make up the target path. By default, the target path is set to *gs://<destination-bucket-name>/<job-name>*.

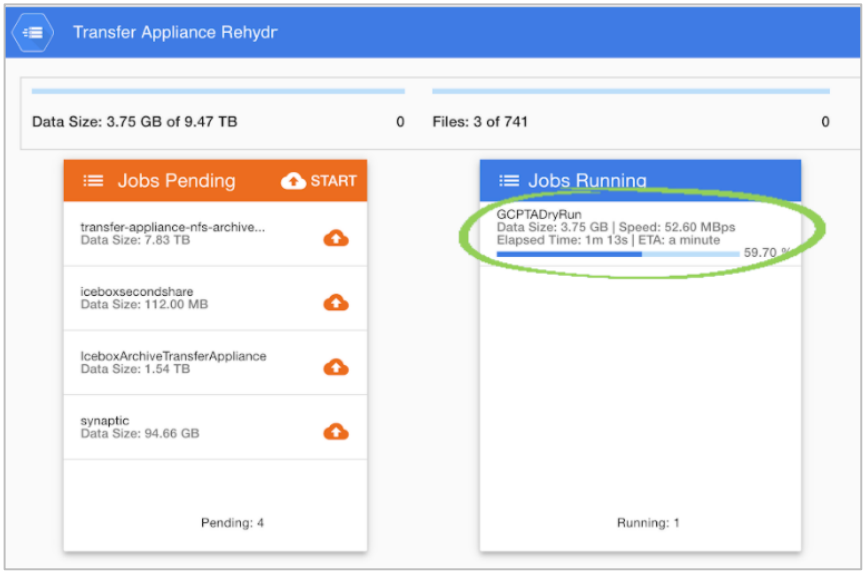


6. To start the hydration run for the selected job, click the **START** icon associated with that job in the **Jobs Pending** pane and click **OK**.

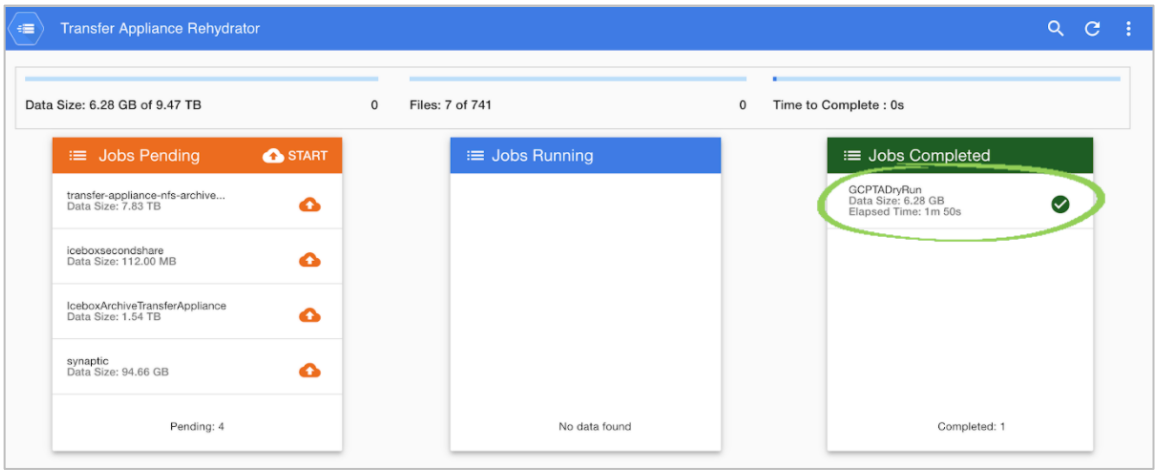


> **NOTE**: To run all rehydration jobs in parallel, click the **START** icon at the top of the **Jobs Pending** pane.

7. The hydration job you started is now displayed in the **Jobs Running** pane.



8. Once the hydration of data to the Google Storage bucket is complete, the hydration job will be moved to the **Jobs Completed** pane.



If you haven't launched all the jobs in parallel, when the first job completes, start the next one, until all pending jobs have started and successfully completed.

Post-hydration Instance Cleanup

Delete the rehyderator instance once you verify that the data hydration has completed.

See Cleaning up in the Google Cloud documentation.

> **NOTE**: Once the data transfer to your Google Storage bucket is complete, the Transfer Appliance is completely wiped following the NIST 800-88 Guidelines for Media Sanitization.

Validate the Archived Data in the Google Storage Bucket

The archived data is available once the rehydration jobs complete. You should validate the data in the Google Storage bucket using one or both of these methods:

- **Inspect the Google Storage bucket**. Check the availability of data in the Google Storage bucket by logging in to your GCP Console and navigating to the Google Storage bucket you registered as your Google Storage bucket External Target. Confirm that the data is there.
- **Initiate recovery from Cohesity**. Search for a file or a VM that was archived using the Transfer Appliance and recover the file or the VM from the Google Storage bucket. Check the recovery task and verify that the recovered file or VM matches the source. For more information, see Recovery in the online Help.

Resume Archival to Google Storage

Now that you have validated the data that you moved to the cloud, you can start the on-demand archival to Google Storage bucket for the new Protection Runs that were completed during the return period. You will trigger on-demand archival runs in chronological order and Cohesity will ensure that the archival runs are executed in that order. Once the on-demand archival is triggered for the completed Protection Runs, you can edit the Protection Policy to add an archival schedule. This way, all future runs of that Protection Group will also be archived to the Google Storage bucket automatically.

To resume archiving your data:

1. Start the on-demand archivals for the new Protection Runs.
2. Add an archival schedule to the Protection Policy.
3. Reset the retention buffer that you increased when you seeded the Transfer Appliance.

Sync Incremental Backups to Google Storage External Target

To start, run on-demand archival tasks for the Protection Runs that completed during the Transfer Appliance return window to the Google Storage bucket External Target. You need this step to keep the archival process consistent.

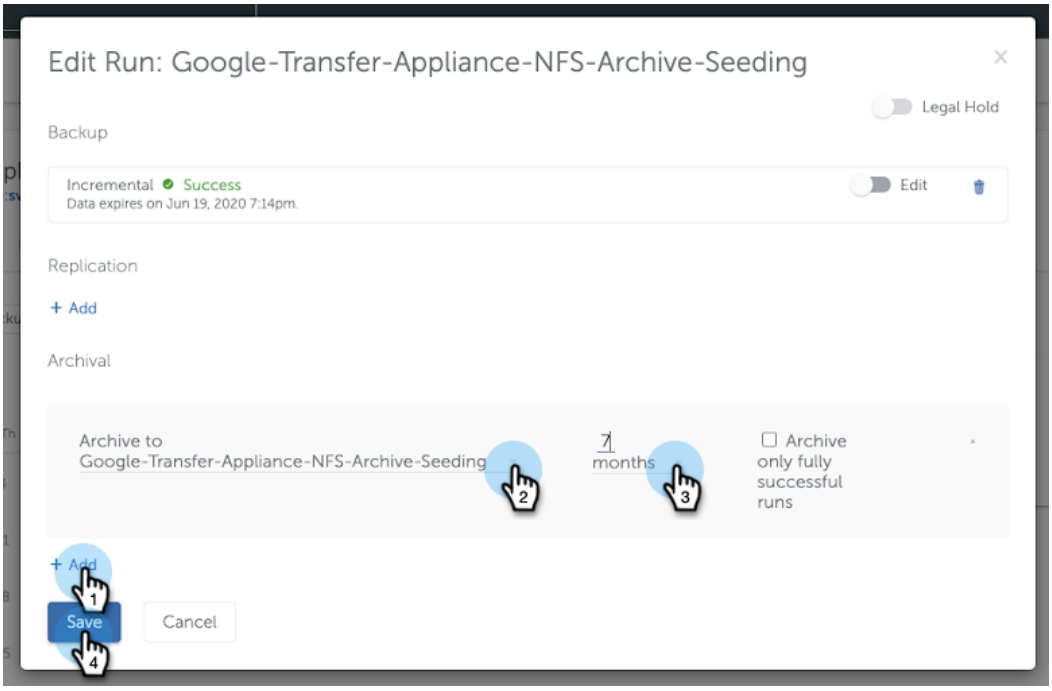
To synchronize the unarchived Protection Runs with the Google Storage bucket External Target:

1. Log in to Cohesity platform.
2. Navigate to **Data Protection > Protection** and click the Protection Group you created earlier. (Use the **Search** box to narrow the results.)
3. Select the successful Protection Runs that are yet to be archived.

> **NOTE**: The archival of Protection Runs must occur in chronological order. If you miss a Protection Run in the on-demand archival process, then the next Protection Run will archive a full instead of incremental backup.

4. For each run, click the **Action** menu and select **Edit Run**.

5. In the dialog, under **Archival**, click **Add** and select the registered Google Storage External Target as the **Archive to** endpoint. Set the retention period and click **Save**.



6. Schedule a sequential on-demand archival to the Google Storage bucket External Target for all the remaining Protection Runs. Ensure that all on-demand runs are scheduled in proper chronological order.

You can now edit the Protection Policy to add an archival schedule to the Protection Group.

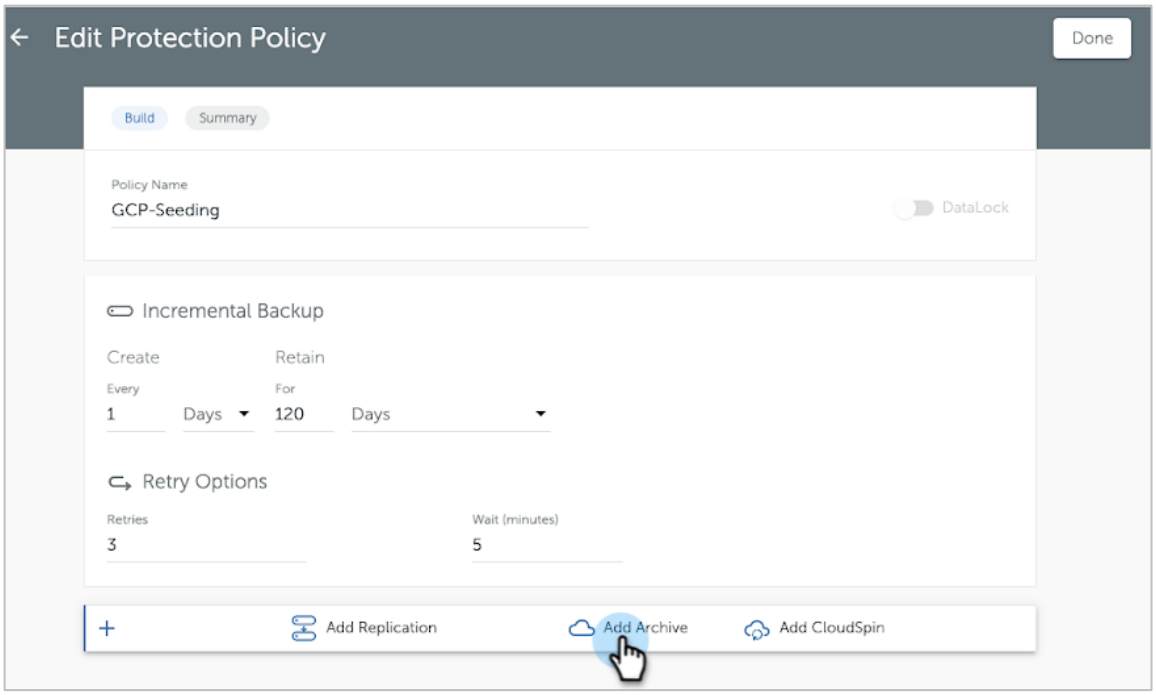
Add Archival Schedule to Protection Policy

To continue with incremental archivals, Cohesity executes the archival tasks, both on-demand and policy-driven, in order. Any new archival runs are added to the queue and executed in the order of arrival. This ordered execution by Cohesity minimizes user wait times by allowing the user to add an archival schedule to the Protection Policy up front and not wait for the existing archival schedule to complete.

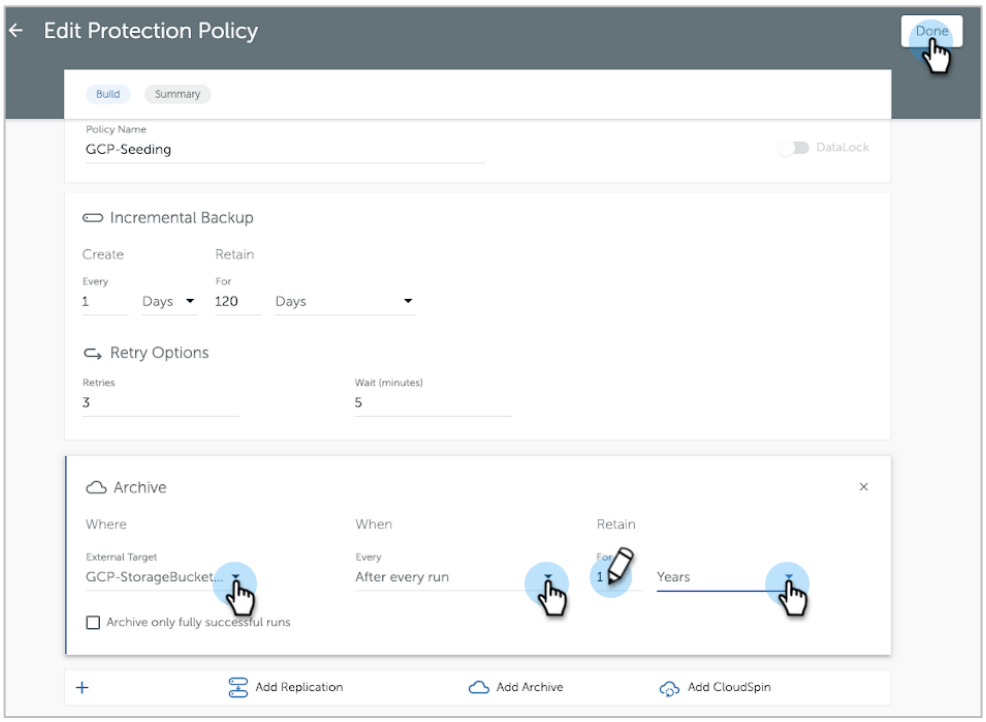
To add an archival schedule to the Protection Policy:

1. Log in to Cohesity platform.
2. Navigate to **Data Protection > Protection** and click the Protection Group you created earlier. (Use the **Search** box to narrow the results.)

3. Click **Add Archive** from the menu at the bottom to add archival to the Protection Policy you associated with the Protection Group.



4. In the **Archive** section, select the Google Storage bucket External Target (**Where**), the appropriate archival frequency (**When**), and how long to keep the archived data (**Retain**). Finally, click **Done** in the top-right corner.



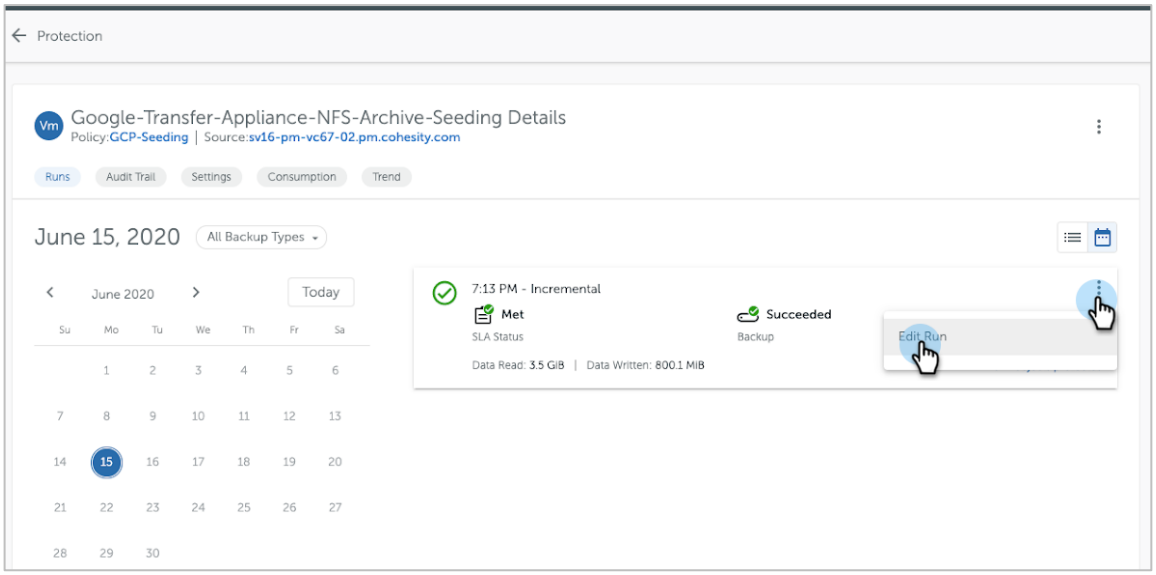
5. Once you save these changes to your Protection Policy, the associated Protection Group starts archiving the data to the Google Storage bucket External Target.

Reset Retention Period

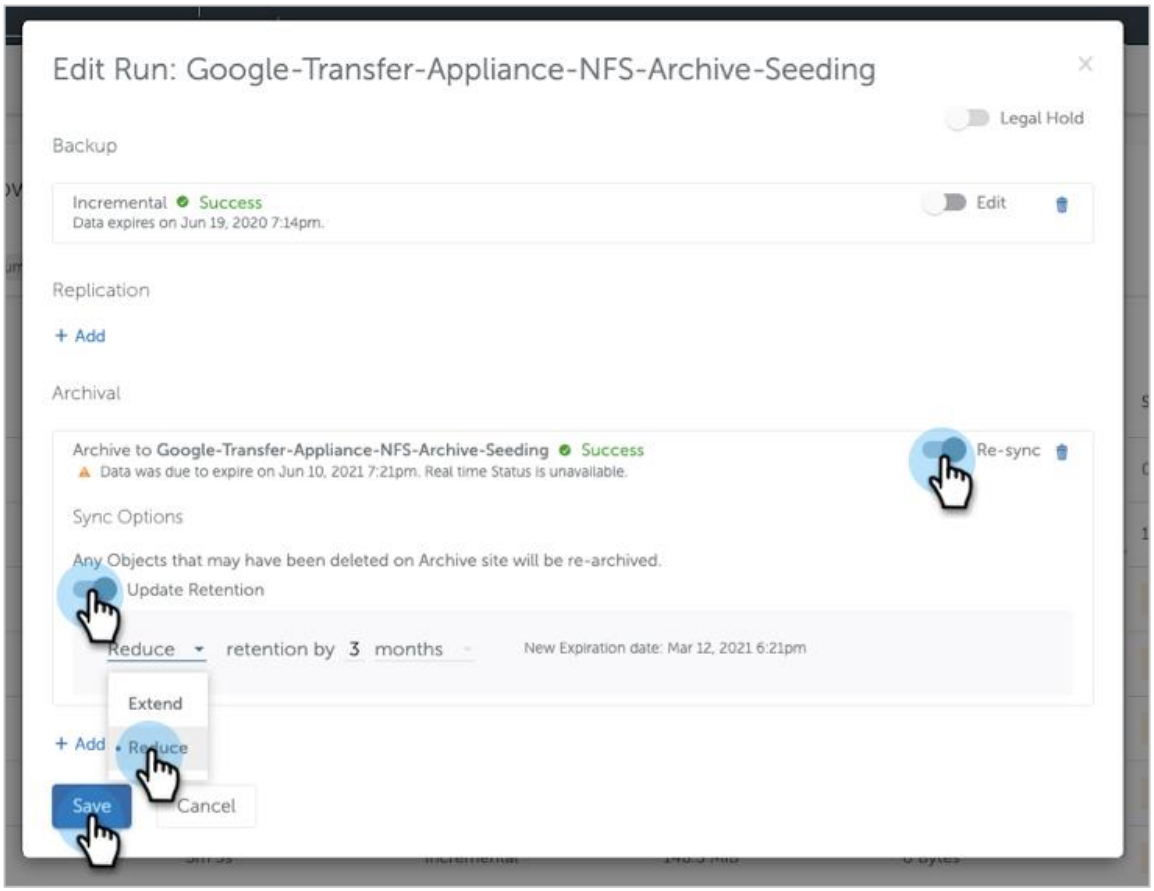
You can now edit the Protection Groups to remove the buffer you added to the retention period earlier.

To reduce the retention period:

1. Log in to Cohesity platform.
2. Navigate to **Data Protection > Protection** and click the Protection Group you created earlier. (Use the **Search** box to narrow the results.)
3. Find the Protection Runs to which you added the buffer and select **Edit Run** from the **Action** menu.



4. In the dialog, under **Archival**, enable **Re-sync** and **Update Retention**, and select the amount of time by which to reduce the retention period. Click **Save.**



Appendix A: Supported Google Storage Classes

When you register the Google Storage bucket as a Cohesity External Target, be sure to select the appropriate storage class for the archived data.

Cohesity supports all Google Storage classes currently available: Nearline, Coldline, Multi-Regional, Regional, and Standard.

Appendix B: Clean Up Transfer Appliance for Retry

Transferring large datasets to a Transfer Appliance involves many components. You might encounter issues with networking connectivity or data corruption problems that can affect or even disrupt the data transfer. If you have found yourself in a situation that requires you to restart the process, you need to prepare the Transfer Appliance by cleaning up the stale data.

To clean up a Transfer Appliance:

1. Connect to the Transfer Appliance Web User Interface and delete the data transfer Jobs.
2. Reset the Transfer Appliance by using either the Web User Interface or the Console User Interface.

See Deleting Data Capture Jobs and Resetting Transfer Appliance in the Google Cloud documentation.

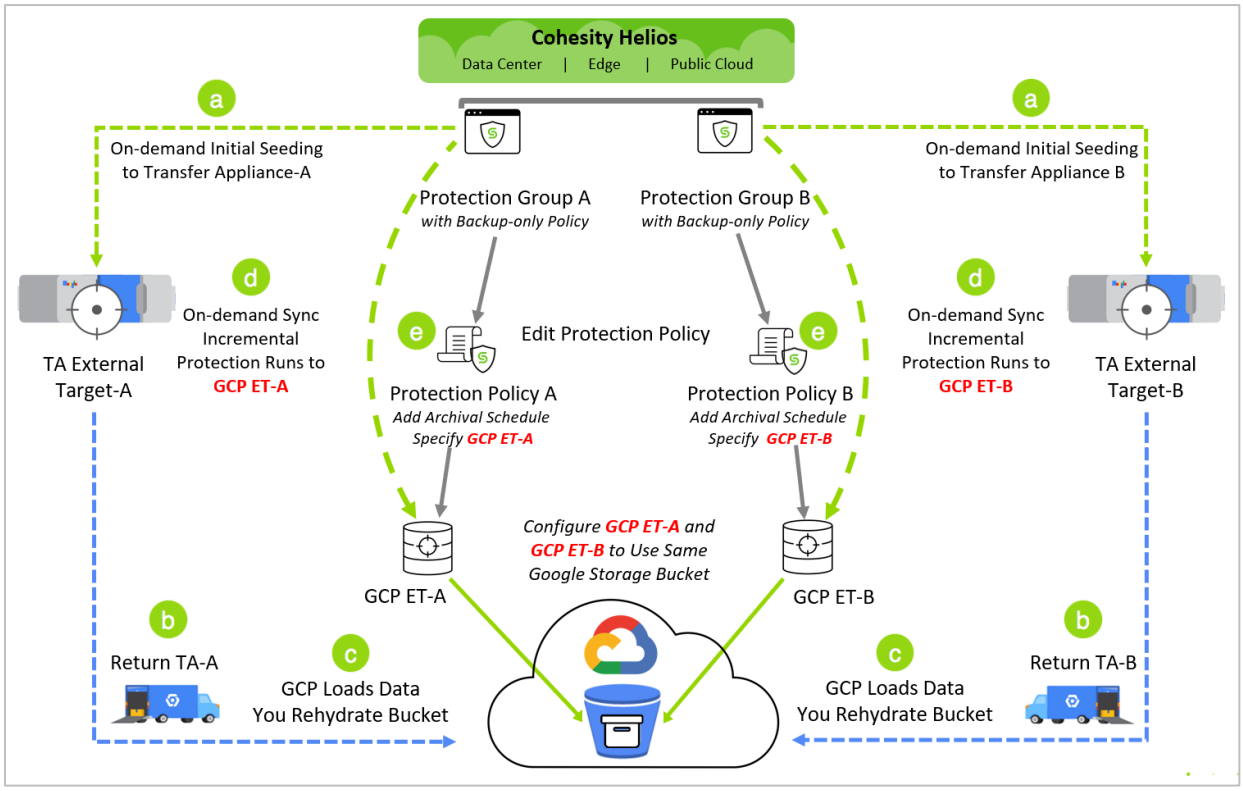
Appendix C: Seed Data with Multiple Devices to a Single Google Storage Bucket

Managing data in the cloud across a number of cloud objects is an administrative hassle as well as operational overhead. Hence, it is beneficial to consolidate the cloud objects for simplified management and reduced operational cost.

In most cases, when the data is migrated to the cloud using a seeding device, there is a limitation on the usable storage capacity for each device. For example, the 'Google Transfer Appliance' comes in two sizes and has a 100 TB or 480TB storage capacity with 80 TB or 384 TB usable storage respectively. This limitation often forces backup administrators to use multiple seeding devices.

With Cohesity, you can set up your archival to the cloud from multiple seeding devices into a single cloud container. As mentioned earlier, you execute the initial seeding of each device using the on-demand archive feature in the Protection Groups. Later, once you have migrated your data into the Google Storage bucket, you will edit the Protection Policy to add the archival schedule and specify the corresponding Google Storage bucket External Target. The trick is to register multiple Google Storage bucket External Targets *with the same Google Storage bucket name*.

Figure 4: Seed Data with Multiple Devices to a Single Storage Bucket



To seed data with multiple Transfer Appliances to a single Google Storage bucket:

1. Configure Cohesity protection. Create multiple Protection Groups.
2. Request the Google Transfer Appliance.

3. Register the Transfer Appliance as a Cohesity External Target, and register a separate External Target for every Transfer Appliance.
4. Initiate on-demand archival seeding to the Transfer Appliances.

Before you initiate seeding to each Transfer Appliance:

a) Ensure the storage consumed by a Protection Group does not exceed the maximum storage capacity of the Transfer Appliance.

b) Seed data from a Protection Group only to a single Cohesity (NAS) External Target.

c) Make note of which Protection Group is using which External Target, as this mapping is necessary to ensure an incremental archival.

5. Return the Transfer Appliances.
6. Register the Google Storage buckets as Cohesity External Targets. Register multiple External Targets, one per Transfer Appliance. If you plan to use the same Google Storage bucket for all your Transfer Appliances, make sure the destination bucket is the same for all External Targets.
7. Initiate External Target swap. Contact Cohesity Support to swap the Transfer Appliance External Target with the Google Storage bucket External Target for each Protection Group. Note that you need to do this for each Transfer Appliance.
8. Rehydrate the data to the same Google Storage bucket for each Transfer Appliance.
9. Sync incremental backups to the Google Storage bucket External Target. Ensure that the External Target selected for the on-demand archival to the cloud is the Google Storage bucket that you swapped in in the previous step.
10. Add archival schedule. Ensure that External Target selected for the archival schedule is the Google Storage bucket to which incremental backups were synchronized in the previous step.

That completes the process! Now that your archives to the cloud are running on a schedule, you're ready to brew yourself a cup of Nesquik® and explore all the other ways Cohesity helps you mine your data for insights, value, and endless possibilities in our fresh online Help and technical guides portal.

Your Feedback

Was this document helpful? Send us your feedback!

About the Authors

Surya Swaminathan is a Sr. Technical Marketing Engineer at Cohesity. In his role, Surya focuses on the cloud, manageability, and disaster recovery.

Adaikkappan Arumugam is Sr. 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:

- Arvind Chandra, Sr. MTS, QA
- Bart Abicht, Sr. Technology Writer and Editor at Cohesity
- Sarthak Agarwal, MTS, Engineering
- Siddhesh Rumde, Sr. MTS, QA

Document Version History

VERSION	DATE	DOCUMENT HISTORY
1.0	Jun 2020	First release
1.1	Nov 2021	Rebranding updates

ABOUT COHESITY

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

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