



Version 1.0

May 2020

# Use Snowball Edge with Cohesity to Bootstrap Archive to AWS

*Seed Your Data to the Cloud with AWS Snowball Edge & Cohesity*

## **ABSTRACT**

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

# Table of Contents

The Need for AWS Snowball Edge .....	4
Use Snowball Edge with Cohesity Platform .....	5
Review Your Archive Seeding Plan .....	6
Prepare to Seed Snowball Edge .....	7
Set Up Cohesity Protection for Your Data.....	7
Estimate Data Size for Archival .....	8
Create Snowball Edge Import Job .....	9
Prepare Your Data Center Infrastructure for Snowball Edge .....	9
Set Up Snowball Edge .....	11
Onboard Snowball Edge.....	11
Unlock Snowball Edge.....	11
Configure and Fetch the S3-compatible Endpoint.....	13
Seed Snowball Edge with Your Protected Data.....	16
Register External Targets .....	16
<i>Register Snowball Edge as an S3-compatible External Target .....</i>	<i>17</i>
<i>Register Amazon S3 Bucket as External Target .....</i>	<i>18</i>
Initiate On-demand Archive Seeding to Snowball Edge.....	20
Validate Data in Snowball Edge.....	22
Return Snowball Edge and Initiate Target Swap .....	23
Initiate External Target Swapping .....	23
Validate the Archived Data in the Amazon S3 Bucket .....	24
Resume Archive to Amazon S3 Bucket .....	25
Sync Incremental Backups to AWS S3 External Target.....	25
Add Archival Schedule to Protection Policy .....	26
Reset Retention Period.....	28
Appendix A: Supported AWS S3 Storage Classes .....	30
Appendix B: Clean Up Snowball Edge for Retry .....	31

Appendix C: Seed Data with Multiple Devices to a Single Amazon S3 Bucket ..	32
Your Feedback .....	34
About the Authors.....	34
Document Version History.....	34

## Figures

Figure 1: Seed Your Cohesity Archive with Snowball Edge .....	5
Figure 2: Phases of the Archive Seeding Process .....	6
Figure 3: Prepare Software and Infrastructure for Snowball Edge .....	7
Figure 4: Seed Data with Multiple Devices to a Single Bucket.....	32

## Tables

Table 1: Data Transfer Times by Size & Network Speed .....	4
Table 2: AWS Parameters.....	15
Table 3: Supported AWS Storage Classes and Cloud Categories.....	30

## The Need for AWS Snowball Edge

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 challenges with protecting your data and leveraging the cloud at the same time are:

- How to get your existing data to the cloud in the first place.
- Is your data network robust enough to facilitate your initial data transfer requirements?
- The time it will 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 Platform 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 transfer rate.

Table 1: Data Transfer Times by Size & Network Speed

TRANSFER RATE	DATA SIZE			
	10 TB	20 TB	40 TB	80 TB
100 Mbps	12 days	24 days	48 days	96 days
1 Gbps	1.25 days	2.5 days	5 days	12.5 days
10 Gbps	3 hrs	6 hrs	12 hrs	1.25 days

Snowball Edge is a terabyte-scale, durable data transport solution, with onboard storage and computing power, that is used to transfer large amounts of data to and from Amazon Web Services (AWS). With Cohesity Platform's cloud-native integration with AWS and Snowball Edge, you now have a much faster, cost-effective solution for transferring large amounts of data to the AWS cloud.

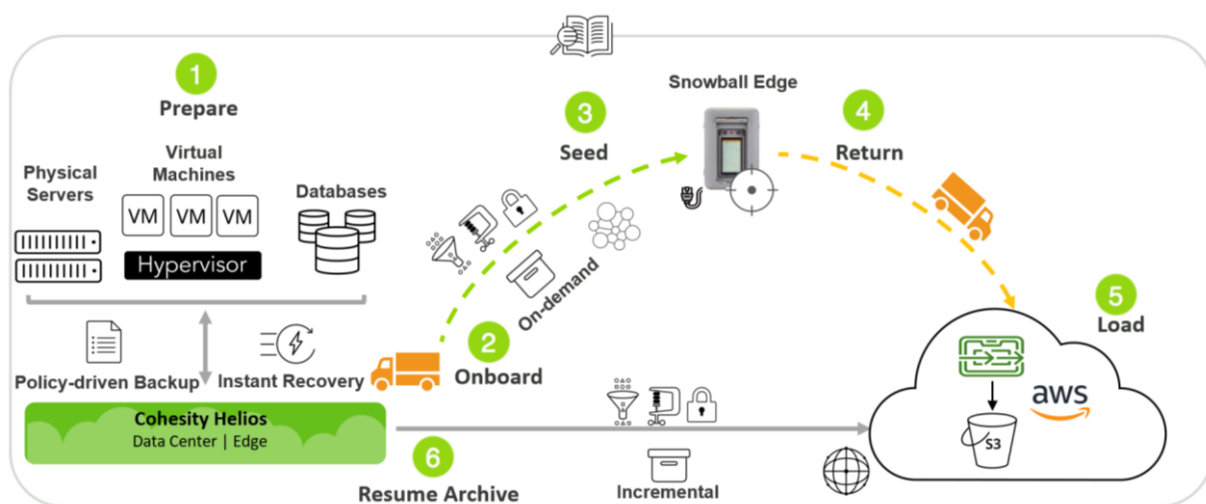
Read the chapters below to learn the steps and best practices to seed initial data from your Cohesity cluster to AWS using a Snowball Edge device. Once you do, you'll be able to resume your Cohesity archival runs to AWS. In the Appendices, you'll find useful details and learn how to consolidate your data by [seeding it from multiple devices to a single AWS bucket](#).

## Use Snowball Edge with Cohesity Platform

Snowball Edge is a physical data-transport device that you can order from your [AWS Management Console](#) and is best suited for IT administrators who need to move terabytes of data into the AWS cloud faster than they can over the Internet. The data stored in an [AWS Snowball Edge](#) device is 256-bit encrypted and is mapped to a pre-configured Amazon S3 bucket. The device comes with multiple configuration options and the *AWS Snowball Edge Storage Optimized* option is best for large-scale data transfer.

**NOTE:** The original Snowball devices were transitioned out of service and Snowball Edge Storage Optimized are now the primary devices used for data transfer. See the [AWS Snowball FAQs](#) for more information.

Figure 1: Seed Your Cohesity Archive with Snowball Edge



The seeding process involves several distinct phases:

1. **Prepare.** Protect your data in Cohesity Platform and order the Snowball Edge device from AWS.

**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 Snowball Edges to order.

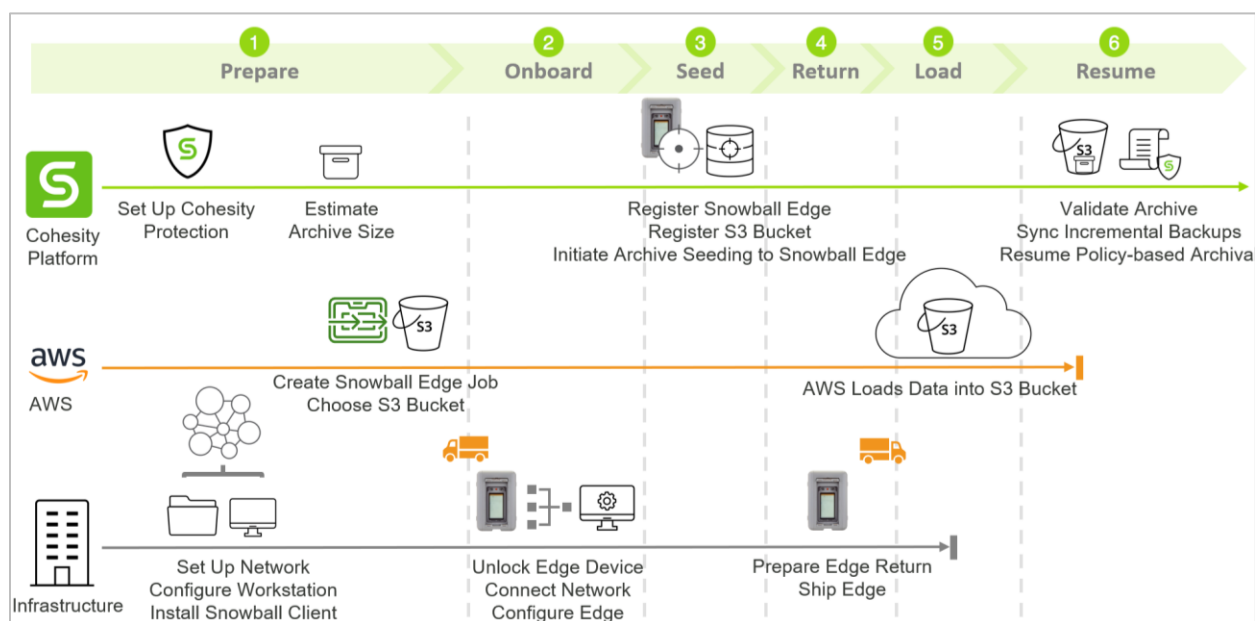
2. **Onboard.** Connect, unlock, and configure the Snowball Edge device.
3. **Seed.** Register the Snowball Edge as an External Target in Cohesity Platform to seed it with your protected data. In most cases, customers find that they need to transfer the full backup followed by several incremental runs to ensure that the seeded data is as up to date as possible.
4. **Return.** Send the device back to AWS and initiate External Target swap.
5. **Load.** AWS loads your data from the Snowball Edge into an Amazon S3 bucket and notifies you once complete.

6. **Resume.** When the AWS import process is complete, validate the data in the Amazon S3 bucket to ensure data integrity, and then resume continuous incremental archival of the data to that Amazon S3 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 AWS using Snowball Edge involves three components:

- Cohesity Platform
- AWS
- Infrastructure

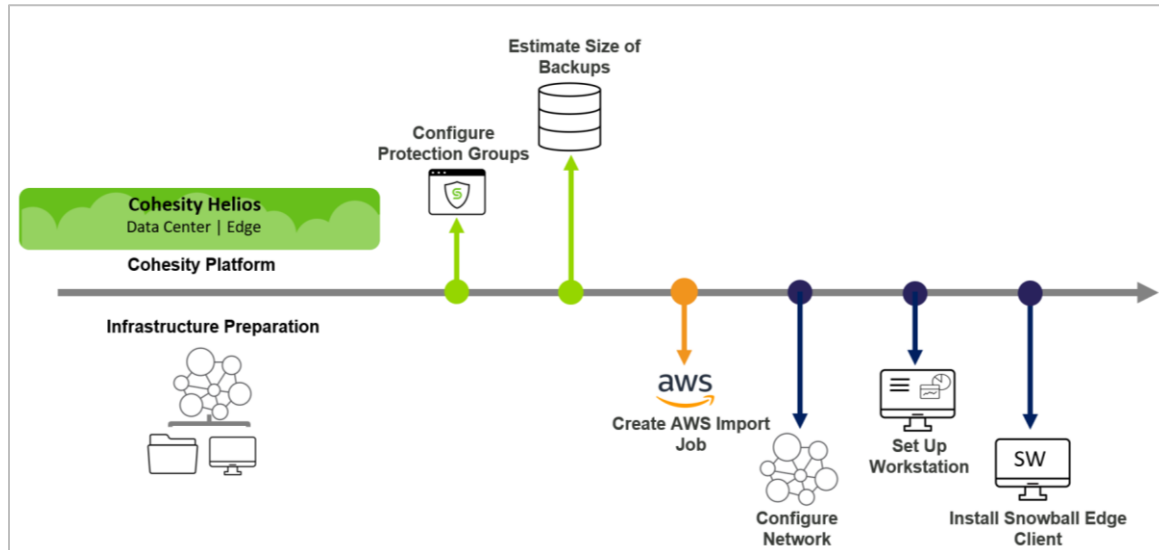
The process is chronologically organized into six phases: Prepare, Onboard, Seed, Return, Load, and Validation.

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 start while you set up the network, configure your workstation, and install the Snowball software client.

## Prepare to Seed Snowball Edge

To maximize the benefit of using the AWS Snowball Edge, Cohesity recommends 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 Snowball Edge



To prepare for seeding data to AWS using Snowball Edge:

1. [Set up Cohesity protection for your data.](#)
2. [Estimate the size of the data to be archived.](#)
3. [Create an import job from the AWS 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 needs to be transferred to your cloud archives via the Snowball Edge.

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

To set up Cohesity Platform 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 [modify 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 create a Snowball Edge import job, you need to estimate the size of the data to be archived. This estimation is needed to determine the number of Snowball Edge devices 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 Platform.
- 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 Snowball Edge device, 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 Snowball Edge.

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

## Create Snowball Edge Import Job

Once you know the size of the Protection Groups, you're ready to create an import job from your AWS Management Console. In the AWS Management Console, you will plan the complete import process, including the selection of your Amazon S3 bucket, the region to host the data, the number of import jobs you need to transfer all the data, and the shipping address where they can send the Snowball Edge device(s).

To create the Snowball Edge import Job:

1. Log in to your [AWS Management Console](#) and create a Snowball Edge import job.
2. Choose **Snowball Edge Storage Optimized (For Data Transfer)** as the device for your Snowball Edge import job.
3. Specify the Amazon S3 bucket into which you will be seeding your data. You can create a new Amazon S3 bucket or choose an existing bucket.

**NOTE:** All Amazon S3 Storage Classes are supported, with the exception of legacy Glacier (AWS S3 Glacier).

4. Enter the shipping address for the AWS Snowball Edge device. The device will be delivered to this address.

For more, see [Create Your First Job](#) in the AWS documentation.

## Prepare Your Data Center Infrastructure for Snowball Edge

Once the import job is created, it takes at least one business day for the Snowball Edge device to be delivered. We can use this time to prepare the infrastructure required to boot the AWS Snowball Edge device when it arrives. This step will optimize the time it takes to onboard the device.

To prepare your infrastructure for AWS Snowball Edge:

1. **Set up the network.**
  - a) Allocate a dedicated IP address for the Snowball Edge device.
  - b) Configure your network firewalls for communication between Cohesity Platform, a client workstation, and the Snowball Edge device. See [Ports Required to Use AWS Services on an AWS Snowball Edge Device](#) in the online Help.
  - c) Optimize network performance by configuring minimum network segments (hops) between the source and the AWS Snowball Edge device.
2. **Dedicate a workstation** or server to facilitate communication with the Snowball Edge.

3. **Download and install the Snowball Edge client software** on the workstation. This client software is required for locking and unlocking the Snowball Edge. For instructions, see [Download and Install the Snowball client](#) in the AWS documentation.
  - a) **Download and install the AWS CLI** on the workstation. See *Downloading and Installing the AWS CLI* in [Transferring Files Using the Amazon S3 Adapter](#) in the AWS documentation.
  - b) **Download your Snowball Edge credentials** from your [AWS Management Console](#).

With each import job, AWS creates a unique set of credentials and associates the credentials with the Job. The credentials include an encrypted manifest job file and an unlock code. You'll need these credentials to unlock the Snowball Edge device. Download the credentials from the AWS Snowball Management Console to your laptop or desktop. For instructions, see [Get Your Credentials and Tools](#) in the AWS documentation.

**NOTE:** As a security best practice, Cohesity strongly recommends you *not* download the credentials to the workstation you will use to unlock the Snowball Edge device.

## Set Up Snowball Edge

When the Snowball Edge arrives, inspect the device for any damage or tampering. If you find any evidence of damage or tampering, contact [AWS Support](#) immediately. Once you confirm the device is undamaged, you're ready to onboard and configure it as an S3-compatible endpoint. Once connected, you can seed your data from Cohesity Platform to this endpoint.

To configure the Snowball Edge device:

1. [Onboard it](#).
2. [Unlock it](#).
3. [Configure and fetch the S3-compatible endpoint](#).

## Onboard Snowball Edge

Start by attaching the Snowball Edge device to a network. Then:

1. Power on the device by pressing the **Power** button above the LCD display screen.

**NOTE:** At this point, if necessary, you can change the network connection using the LCD display.

2. See the LCD display on the device to confirm that the device is ready for use.

For detailed instructions, see [Connect to Your Local Network](#) in the AWS documentation.

## Unlock Snowball Edge

To ensure the integrity of the data transferred using the Snowball Edge device, AWS ships the device in a locked state. To unlock the device, you'll need the Snowball Edge client software, the encrypted job manifest file, and a uniquely generated unlock code. The credentials are cryptologically attached to your account and can be downloaded from the AWS Management Console.

To unlock a Snowball Edge device:

1. Log in to the workstation on which you installed the AWS Snowball Edge client software.
2. Copy the encrypted job manifest file and the unlock code to the workstation. Use credentials you fetched in the [Infrastructure preparation](#) step above.

**NOTE:** As a security best practice, Cohesity strongly recommends you delete the encrypted job manifest file and the unlock code after unlocking the Snowball Edge device.

3. Configure a profile for the Snowball Edge device. This helps you attach the credentials to the profile and use the profile to run subsequent Snowball Edge client commands.

To configure the profile, run the **snowballEdge configure** command in the AWS CLI and enter:

- **Snowball Edge Manifest File Path:** The path to the encrypted manifest file.
- **Unlock Code:** The 29-character unlock code.
- **Default Endpoint:** The endpoint is the IP address with the `https://` prefix. Locate this IP address on the AWS Snowball Edge device's LCD display.

**For example:**

```
snowballEdge configure --profile <Profile_Name>
Configuration will stored at: <Path_to_Encrypted_Manifest_File>
Snowball Edge Manifest Path: /tmp/JID2EXAMPLE-0c40-49a7-9f53-
916aEXAMPLE81-manifest.bin
Unlock Code: 12345-abcde-12345-ABCDE-12345
Default Endpoint: https://192.0.2.0:8443
```

4. Use the **unlock-device** command and the configured profile as input to unlock the Snowball Edge device.

```
$ snowballEdge unlock-device --profile <Profile_Name>
```

**Output:**

```
Your Snowball Edge device is unlocking. You may determine the unlock state
of your device using the describe-device command. Your Snowball Edge device
will be available for use when it is in the UNLOCKED state.
```

5. Use the **describe-device** command to verify the state of the Snowball Edge device.

```
$ snowballEdge describe-device --profile <Profile_Name>
```

**Sample Output:**

```
{
  "DeviceId" : "JID14a49670-98d6-57ba-92f9-b27a5133a44f",
  "UnlockStatus" : {
    "State" : "UNLOCKED"
  },
  "ActiveNetworkInterface" : {
    "IpAddress" : "192.2.88.987"
  },
  "PhysicalNetworkInterfaces" : [ {
    "PhysicalNetworkInterfaceId" : "s.ni-89e7e5i698e07d63f",
    "PhysicalConnectorType" : "RJ45",
    "IpAddressAssignment" : "DHCP",
    "IpAddress" : "192.168.88.110",
    "Netmask" : "255.255.255.0",
    "DefaultGateway" : "192.168.88.1",
    "MacAddress" : "ah:1v:6n:08:77:z4"
  }
}
```

For more, see [Unlock the Snowball Edge](#) in the AWS documentation.

## Configure and Fetch the S3-compatible Endpoint

Once the device is unlocked, the next steps involve fetching the Access Key and the Secret Access Key of the admin user of your AWS account on the Snowball Edge device. These credentials are then used to configure the root user credentials to authenticate your requests through the AWS CLI.

To configure the root user credentials and fetch the S3 compatible endpoints:

1. List the Access Keys stored in the Snowball Edge device.

```
$ snowballEdge list-access-keys --profile <Profile_Name>
```

**Sample Output:**

```
{
  "AccessKeyIds" : [ "AHFICJFQ4BL864204K2B" ]
}
```

2. Provide the Access Key as input to fetch the Secret Access Key.

```
$ snowballEdge get-secret-access-key --accesskey-id  
AHFICJFQ4BL864204K2B --profile <Profile_Name>
```

**Output:**

```
aws_access_key_id = AHFICJFQ4BL864204K2B  
aws_secret_access_key = qBavloXUtnWLDI/T7KDEJR/bLyRsiGNESAMPLEKEY
```

3. Use the **aws configure** command in the AWS CLI to configure the root user credentials. At the prompt, enter:

- **Profile Name:** Enter the name of the profile.
- **AWS Access Key ID:** The output of the **snowballEdge list-access-keys** command.
- **AWS Secret Access Key:** The output of the **snowballEdge get-secret-access-key** command.
- **Default region name:** Region of your Amazon S3 bucket.
- **Default output format [json]:** The supported formats are **JSON** (default), **YAML**, **text**, and **table**.

**Sample Command:**

```
$ aws configure --profile <Profile_Name>  
AWS Access Key ID [None]: AHFICJFQ4BL864204K2B  
AWS Secret Access Key [None]:qBavloXUtnWLDI/T7KDEJR/bLyRsiGNESAMPLEKEY  
Default region name [None]: us-east-1  
Default output format [json]: json
```

4. Once the root user credentials are configured, you can use the **s3 ls** command to list the Amazon S3 bucket:

**Sample Command:**

```
$ aws s3 ls --profile <Profile_Name>
```

**Sample Output:**

```
2019-10-12 10:50:30 snowballs3importbucket
```

5. Fetch and make note of all the information you will need to [register the AWS Snowball Edge device as an External Target](#) in Cohesity Platform.

Table 2: AWS Parameters

PARAMETER	TO FETCH PARAMETER VALUE, USE:
Bucket Name	<code>aws s3 ls --profile &lt;Profile_Name&gt;</code>
Access Key	<code>snowballEdge list-access-keys --profile &lt;Profile_Name&gt;</code>
Secret Access Key	<code>snowballEdge get-secret-access-key --accesskey-id &lt;id&gt; --profile &lt;Profile_Name&gt;</code>
Endpoint	Fetch the endpoint IP address from the LCD display of the Snowball Edge device. For example: <code>https://192.0.2.0:8443</code>

## Seed Snowball Edge with Your Protected Data

Once the Snowball Edge setup is complete, you can register it as an S3-compatible External Target in Cohesity Platform. And once the Snowball Edge is registered, you will edit the [Protection Group you created above](#) to add an on-demand archival task that seeds the Snowball Edge with the data in that Protection Group.

For the initial seeding of the Protection Group to the device, the Protection Policy includes *only a backup schedule and no archival schedule*. The Snowball Edge device is seeded using the on-demand archive feature in the Protection Runs of the Protection Group. However, to resume archival after AWS finishes loading your data from the Snowball Edge to the Amazon S3 bucket, you will edit the Protection Policy to *add the archival schedule* and specify the corresponding AWS S3 External Target.

To seed the data from Cohesity Platform to the Snowball Edge:

1. [Register the Snowball Edge and the Amazon S3 bucket as Cohesity External Targets](#).
2. [Initiate an on-demand archival run to the Snowball Edge](#).
3. [Validate the data in Snowball Edge](#).

## Register External Targets

There are two S3 endpoints you need to register as External Targets with Cohesity Platform.

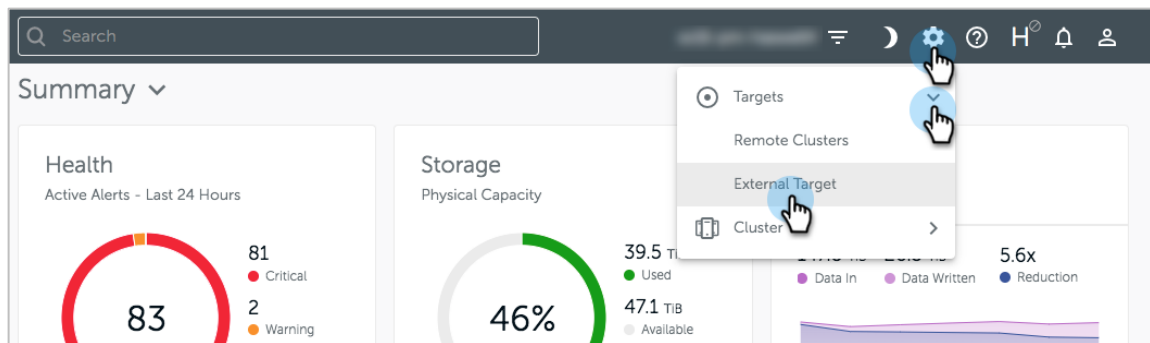
1. [Snowball Edge device](#), to move your data to it.
2. [Amazon S3 bucket](#), the long-term cloud home of your archived data, to resume incremental archives once AWS loads your data from the Snowball Edge into your Amazon S3 bucket.

## Register Snowball Edge as an S3-compatible External Target

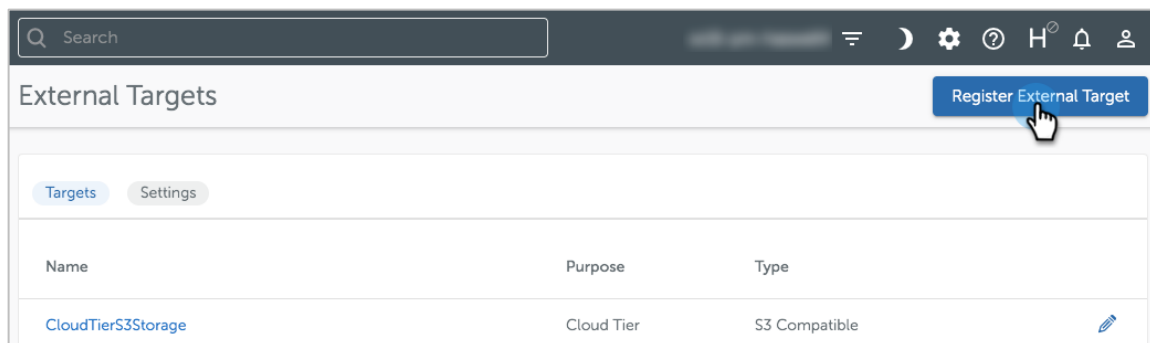
To register the AWS Snowball Edge device as an S3-compatible External Target in Cohesity Platform, you will need the Snowball Edge's bucket name, endpoint details, and the Access Keys [you fetched while setting up the Snowball Edge](#).

To register the Snowball Edge as an S3-compatible External Target:

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



3. Click **Register External Target**.



- In the form that opens, enter the **New Target** name and other parameters [that you fetched from the Snowball Edge device](#) earlier, and then click **Register**.
  - Purpose:** Archival
  - Type:** S3 Compatible
  - Bucket Name:** The bucket name [you fetched from the Snowball Edge device](#).
  - Access Key ID:** The Access Key of the AWS account on the Snowball Edge device.
  - Secret Access Key:** The Secret Access Key of the AWS account on the Snowball Edge device.
  - Endpoint:** The default configured endpoint of the Snowball Edge device.

Search

New Target: SnowballEdge-Seeding-S3-Archive-Target Description

**Purpose**  
 Archival  Tiering

**S3 Compatible**

**Bucket Name \***

**Access Key ID \***

**Secret Access Key \***

**Endpoint \***  
https://

Secure Connection (HTTPS)

AWS Signature Version  ver 2  ver 4

Encryption

Additional security by managing key manually (i)

Compression

Source Side Deduplication

Incremental Archival

Bandwidth Throttling

**Register** **Cancel**

## Register Amazon S3 Bucket as External Target

The data seeded to the AWS Snowball Edge device will be loaded by AWS into the Amazon S3 bucket that you created during the [Create Import Job](#) step. This Amazon S3 bucket is registered as a Cohesity External Target to ensure the continuous archival of data. Find the access and endpoint details of the Amazon S3 bucket in your [AWS Management Console](#).

To register the Amazon S3 bucket as a Cohesity External Target:

1. Log into 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 the appropriate bucket type. See [Appendix A: Supported AWS S3 Storage Classes](#).
  - **Bucket Name:** Bucket name from the [Snowball Edge import job](#) you created earlier.
  - **Region:** Enter the AWS region that the bucket belongs to.
  - **Access Key ID:** Fetched from the AWS Management Console.
  - **Secret Access Key:** Fetched from the AWS Management Console.

The screenshot shows the 'New Target' registration form in the Cohesity platform. The form is titled 'New Target: AWS-S3-Archive-Target' and includes a search bar at the top. The form fields are as follows:

- Purpose:** Archival (selected), Tiering
- AWS S3:** (selected)
- Category:** Standard (selected), Gov, C2S
- Bucket Name:** snowballs3importbucket
- Region:** us-east-1
- Access Key ID:** (redacted)
- Secret Access Key:** (redacted)
- Encryption:** (checked)
- Additional security by managing key manually:** (unchecked)
- Compression:** (checked)
- Source Side Deduplication:** (checked)
- Incremental Archival:** (checked)
- Bandwidth Throttling:** (unchecked)

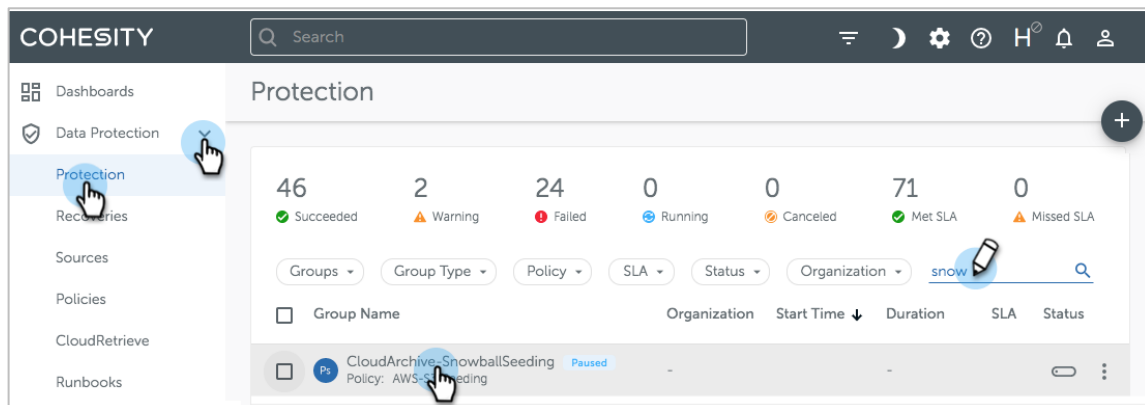
At the bottom of the form, there are two buttons: 'Register' and 'Cancel'.

## Initiate On-demand Archive Seeding to Snowball Edge

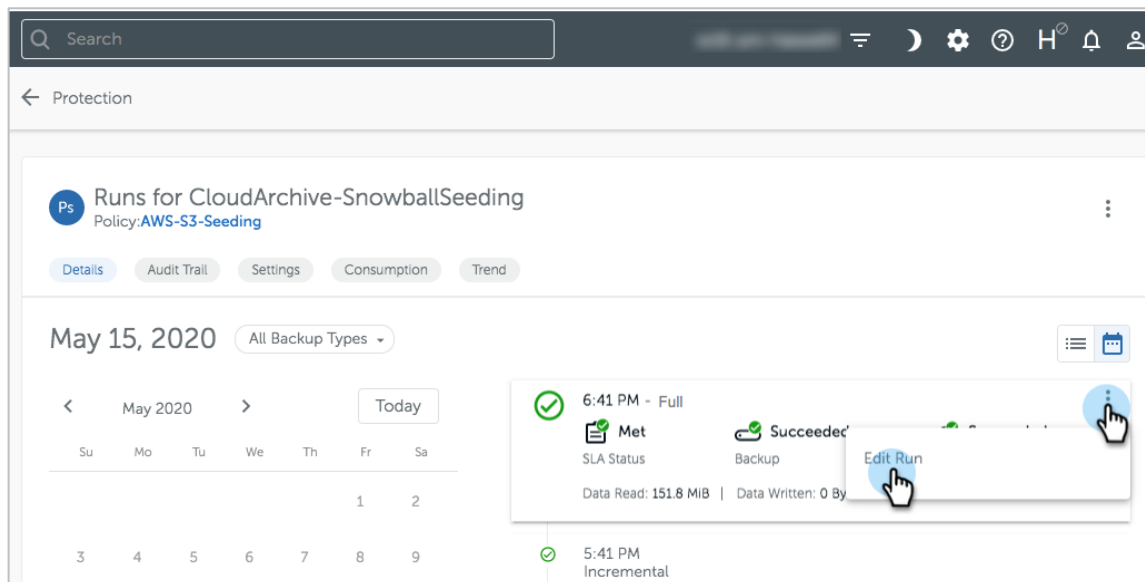
Once the Snowball Edge is registered as a Cohesity External Target, you can archive data directly to it on demand.

To start an on-demand archival run:

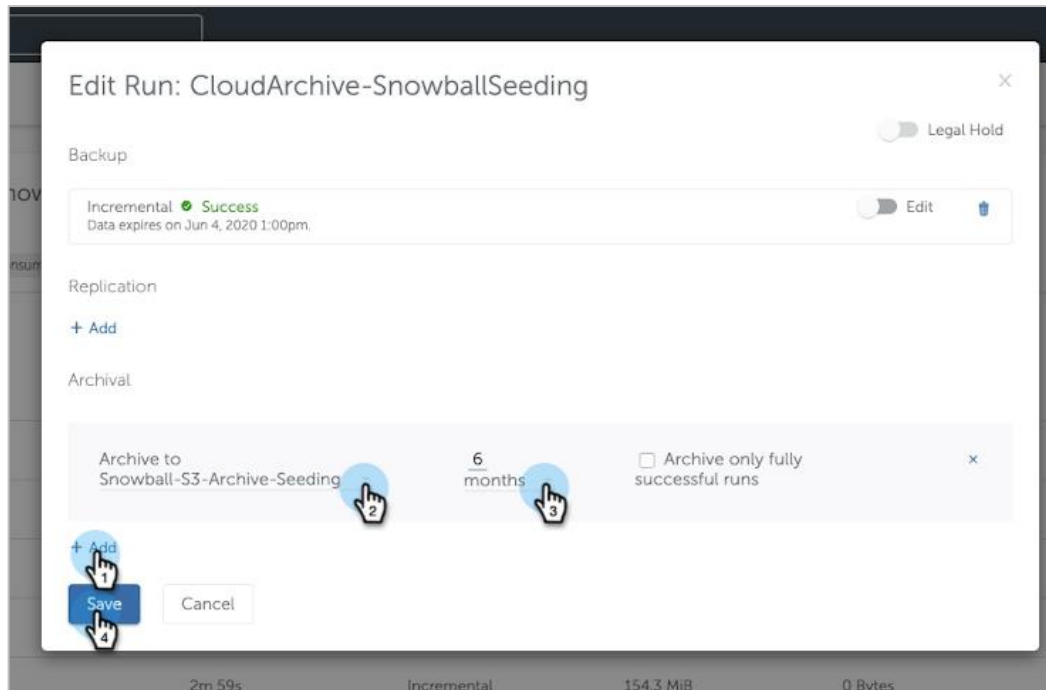
1. Log in to Cohesity Platform.
2. Navigate to **Data Protection > Protection**, and click the [Protection Group you created earlier](#). (Use the **Search** box to narrow the results.)



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



- In the dialog, under **Archival**, click **Add** and select the Snowball Edge 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 [modify the retention period](#) later.

- 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 Snowball Edge before you retry. See [Appendix B: Clean Up Snowball Edge for Retry](#) for instructions.
- 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 more recent incremental archival runs.

**IMPORTANT:** Contact [Cohesity Support](#) to complete [this verification step](#) before initiating on-demand archivals to the Snowball Edge device.

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

## Validate Data in Snowball Edge

It is important to validate the data in the Snowball Edge after the archival run to the device completes.

This validation step ensures there is no corruption in the data seeded to Snowball Edge, and, if there are discrepancies, gives you an opportunity to start over.

**IMPORTANT:** Contact [Cohesity Support](#) to complete [this data-validation step](#).

## Return Snowball Edge and Initiate Target Swap

When all the data that was marked for archival is seeded to the AWS Snowball Edge device, prepare the device to be returned to the AWS upload facility. When AWS receives the device, they will load the data into the [Amazon S3 bucket you registered earlier](#) and notify you when the seeding process completes.

For a smooth return process, ensure that data transfer to the Snowball Edge device is finished and the [validation step](#) completed, and:

1. Press the **Power** button above the LCD Display on the Snowball Edge device to turn it off. For details, see [Stop the Snowball client, and Power Off the Snowball Edge](#) in the AWS documentation.
2. Disconnect the AWS Snowball Edge device cables. See [Disconnect the Snowball Edge](#) in the AWS documentation.
3. Prepare the device for shipment. See [Return the Snowball Edge](#) in the AWS documentation.
4. Track the return and monitor the progress of the import through your [AWS Management Console](#). See [Monitor the Import Status](#) in the AWS documentation.

Whenever data is imported into Amazon S3, you get a downloadable PDF job report. Check the import job report for errors. See [Get Your Job Completion Report and Logs in the Console](#) in the AWS documentation.

## Initiate External Target Swapping

Once the device is returned and the Amazon S3 bucket is seeded, Cohesity Platform will be able to resume archiving, but going forward now, to the Amazon S3 bucket and not the Snowball Edge device. To be able to pick up where you left off, you need to ask Cohesity Support to help you replace the Snowball Edge External Target with the AWS S3 External Target, known as External Target swapping.

Because it takes at least 2-3 days after you ship the device back for AWS to receive it and load the data into the Amazon S3 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](#).

## Validate the Archived Data in the Amazon S3 Bucket

After AWS has finished loading your data into the Amazon S3 bucket, you have several options to validate the data in Amazon S3.

- **Inspect the Amazon S3 bucket.** Check the availability of data in the Amazon S3 bucket by logging in to your [AWS Management Console](#) and navigating to the Amazon S3 bucket [you registered as your AWS S3 External Target](#). Confirm that the data is there.
- **Initiate recovery from Cohesity Platform.** Search for a file or a VM that was archived using Snowball Edge and recover the VM from the Amazon S3 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.

**NOTE:** Once your data is loaded into the Amazon S3 bucket, AWS performs a software erasure of the Snowball Edge device. For more information on security, see [Protecting Data On Your Device](#) in the AWS documentation.

By default, AWS loads the data from the seeding device to an Amazon S3 bucket in the S3 Standard storage class. You can change the storage class of the data in the Amazon S3 bucket, but to maintain consistency, ensure the storage class of the data in the Amazon S3 bucket and the [storage class selected during the registration of AWS S3 External Target](#) is the same.

You can do this manually or by setting up a [life cycle policy](#) on the bucket. For more, see [Amazon S3 storage classes](#) in the AWS documentation.

## Resume Archive to Amazon S3 Bucket

Now that you have validated the data that you moved to the cloud, you can start the on-demand archival to Amazon S3 bucket for the new Protection Runs that were completed during the return period. You will trigger on-demand archival runs in a chronological order and Cohesity Platform 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 Amazon S3 bucket automatically.

To resume archiving your data:

1. Start the [on-demand archival](#) for the new Protection Runs.
2. [Add an archival schedule](#) to the Protection Policy.
3. [Reduce the retention buffer](#) that you increased when you [seeded the Snowball Edge device](#).

## Sync Incremental Backups to AWS S3 External Target

To start, run on-demand archival tasks for the Protection Runs that completed during the device return window to the [AWS S3 External Target](#). You need this step to keep the archival process consistent.

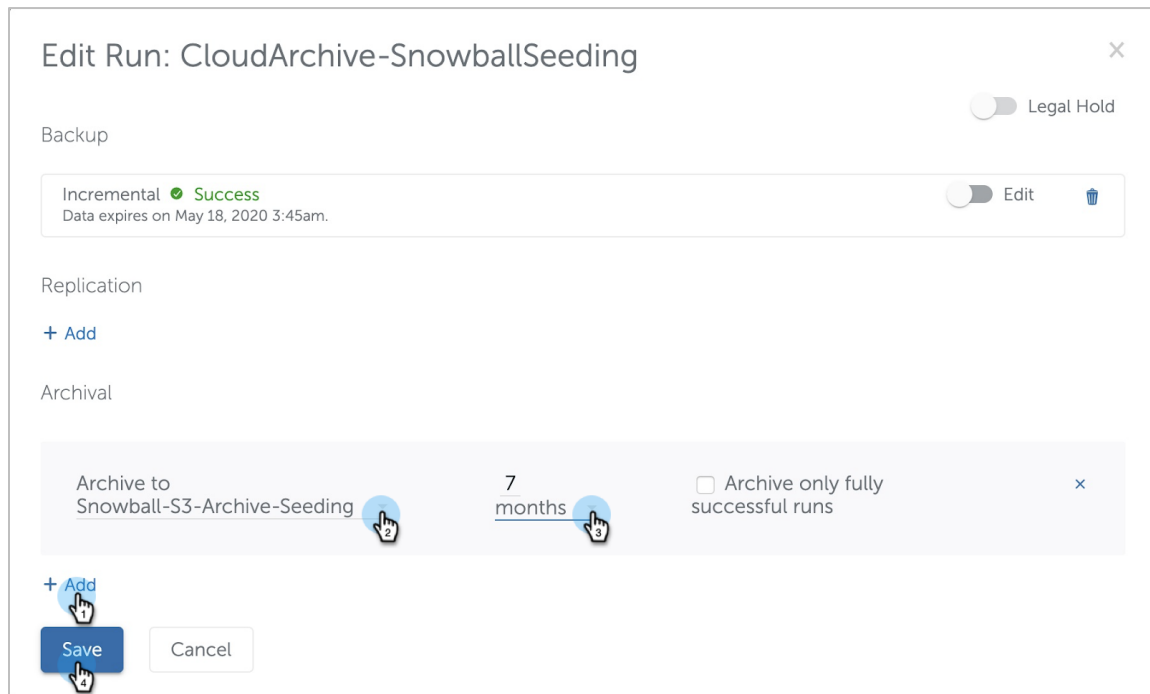
To synchronize the unarchived Protection Runs with the AWS S3 External Target:

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**.

- In the dialog, under **Archival**, click **Add** and select the registered AWS S3 External Target as the **Archive to** endpoint. Set the retention period and click **Save**.



- You can schedule a sequential on-demand archival to the AWS S3 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 maintain consistency among archivals, Cohesity Platform executes the archival tasks, both on-demand and policy-driven, on a schedule. Any new archival runs are added to the queue and executed in the order of arrival. This ordered execution by Cohesity Platform 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.

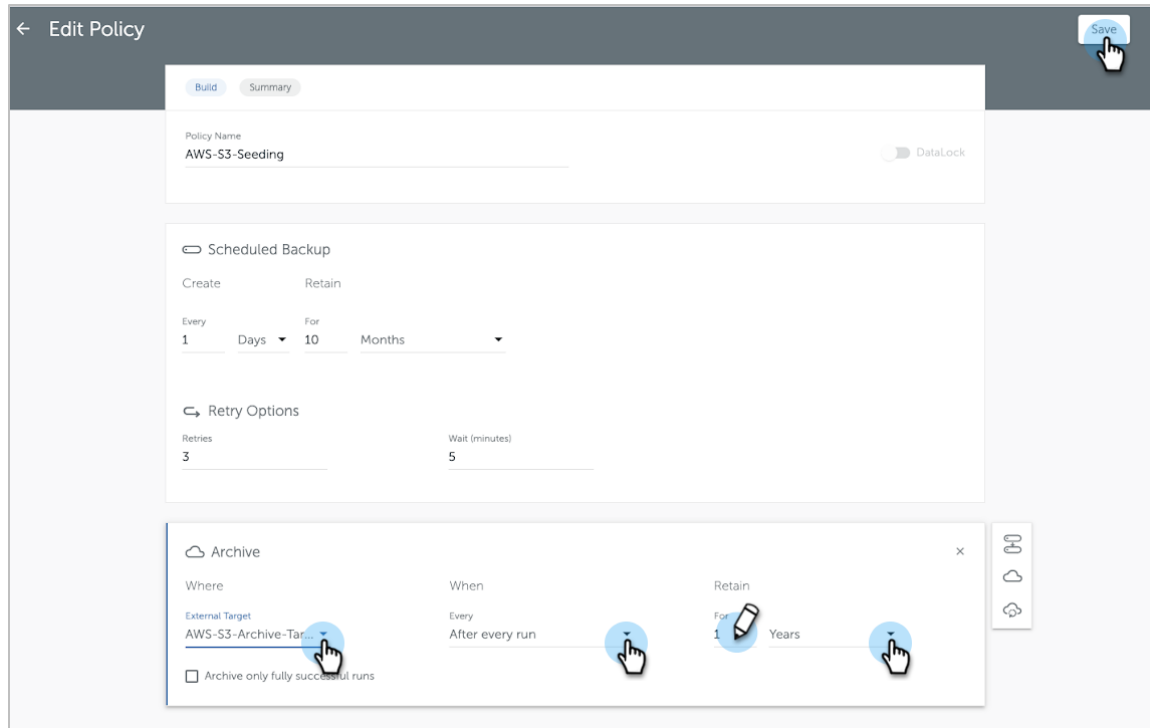
To add an archival schedule to the Protection Policy:

- Log in to Cohesity Platform.
- Navigate to **Data Protection > Protection** and click the [Protection Group you created earlier](#). (Use the **Search** box to narrow the results.)

3. Click **Archive** from the menu on the right to add archival to the Protection Policy you associated with the Protection Group.

The screenshot displays the 'Edit Policy' interface for a policy named 'AWS-S3-Seeding'. The interface includes a 'Save' button in the top right corner and two tabs: 'Build' (active) and 'Summary'. The 'Policy Name' field contains 'AWS-S3-Seeding' and a 'DataLock' toggle is visible. Below this is the 'Scheduled Backup' section, which includes 'Create' and 'Retain' options, and a frequency setting of 'Every 1 Days For 10 Months'. The 'Retry Options' section shows 'Retries' set to 3 and 'Wait (minutes)' set to 5. On the right side, a vertical menu contains several icons, with the 'Archive' icon (a cloud with a checkmark) highlighted by a mouse cursor.

4. In the **Archive** section, select the AWS S3 External Target (**Where**), the appropriate archival frequency (**When**), and how long to keep the archived data (**Retain**). Finally, click **Save** 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 AWS S3 External Target.

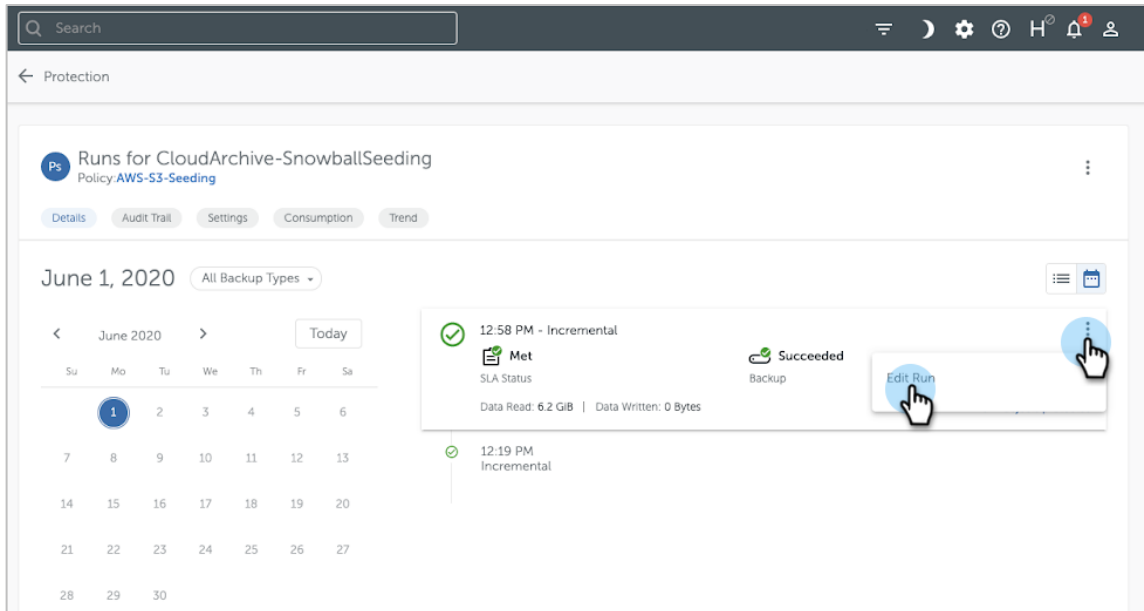
## Reset Retention Period

You can now edit the Protection Groups to remove the buffer [you added to the retention period earlier](#).

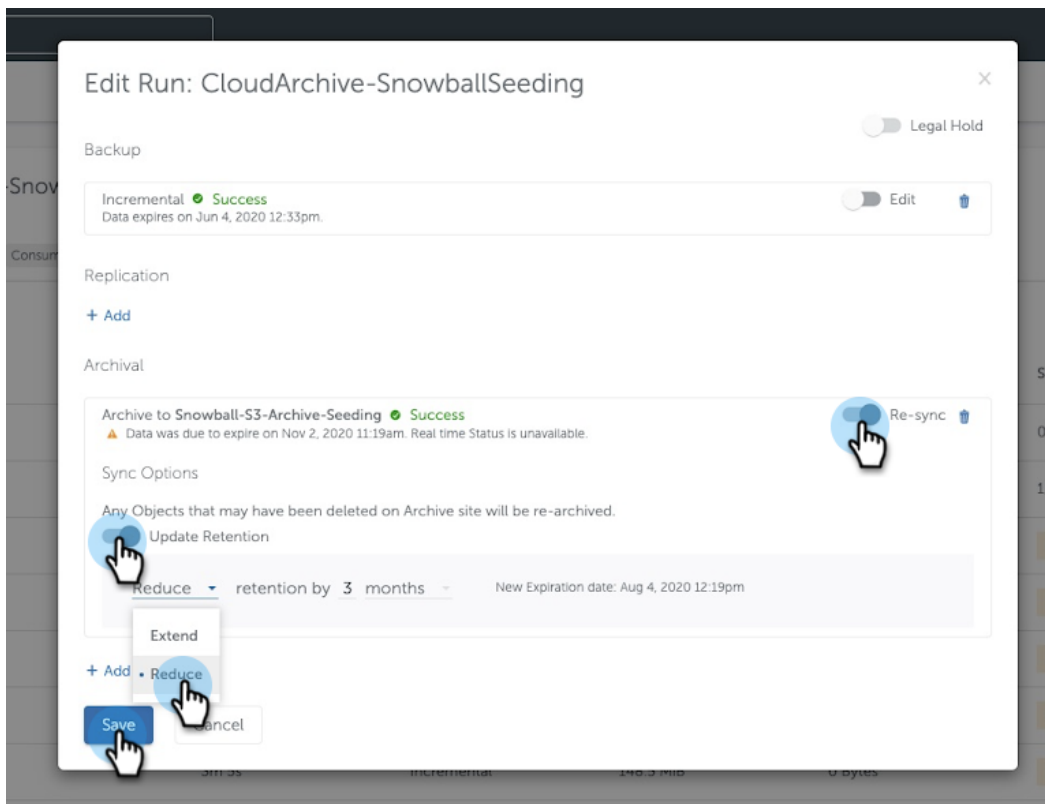
To reduce the retention period:

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.)

- Find the Protection Runs to which you added the buffer and select **Edit Run** from the **Action** menu.



- 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 AWS S3 Storage Classes

Use Table 3 below to choose the appropriate AWS S3 storage class when you [register the Amazon S3 bucket as a Cohesity External Target](#). This step ensures that the archived data is assigned to the appropriate S3 storage class.

Table 3: Supported AWS Storage Classes and Cloud Categories

AWS STORAGE CLASS	STANDARD	GOV	C2S
S3	Yes	Yes	Yes
S3-One-Zone	Yes	Yes	N/A
S3-IA	Yes	N/A	Yes
S3-Intelligent	Yes	N/A	N/A
S3-Glacier	Yes	Yes	N/A
S3-Glacier-Deep-Archive	Yes	Yes	N/A
AWS S3 Glacier (Legacy)	No	No	No

## Appendix B: Clean Up Snowball Edge for Retry

Transferring large datasets to a seeding device involves many components. You might encounter issues with networking connectivity, data corruption, or manifest file 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 Snowball Edge device by cleaning up the stale data.

To clean up a Snowball Edge device:

1. Log in to the workstation on which the AWS Snowball Edge was installed.
2. Use the `aws s3 ls` command in the AWS CLI to list the contents of the Amazon S3 bucket. For example:

```
$ aws s3 ls s3://<Bucket_Name> --profile <Profile_Name>
```

3. Delete the objects in the bucket using the `aws s3 rm` command. Add the `--recursive` option to delete all the objects in the bucket's hierarchy. For example:

```
$ aws s3 rm s3://<Bucket_Name>/ --recursive --profile <Profile_Name>
```

For more, see [Troubleshooting for an AWS Snowball Edge](#) in the AWS documentation.

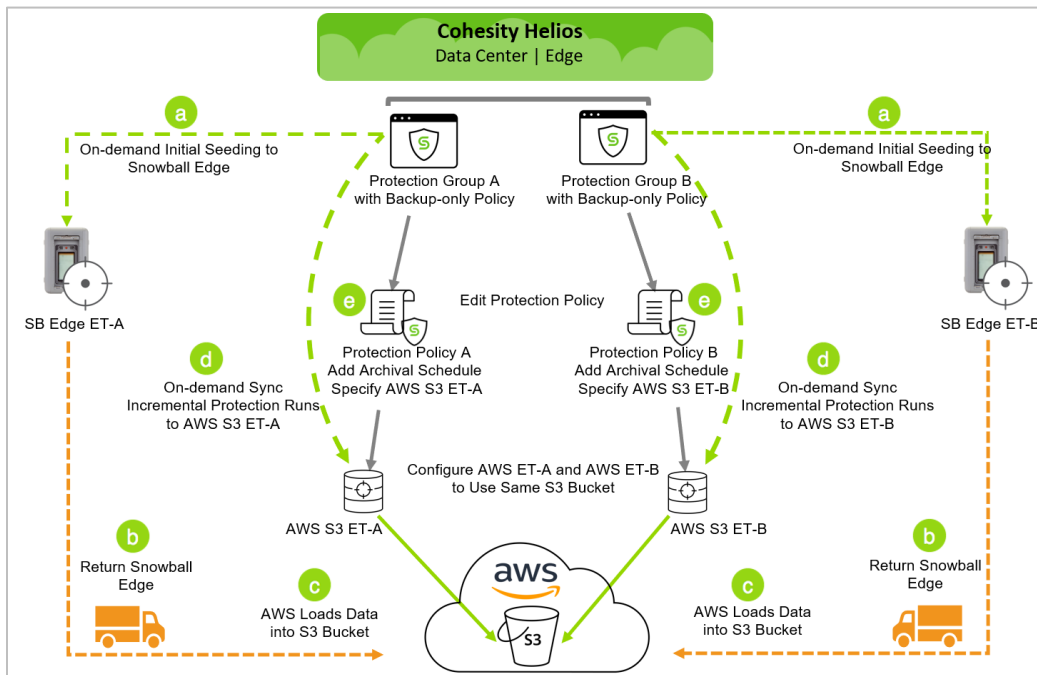
## Appendix C: Seed Data with Multiple Devices to a Single Amazon S3 Bucket

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 ‘Snowball Edge Storage Optimized (for data transfer)’ device has a 100 TB storage capacity with 80 TB usable storage. This limitation often forces backup administrators to use multiple seeding devices.

With Cohesity Platform, 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 AWS finishes loading your data into the Amazon S3 bucket, you will edit the Protection Policy to add the archival schedule and specify the corresponding AWS S3 External Target. The trick is to register multiple AWS S3 External Targets *with the same Amazon S3 bucket name*.

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



To seed data with multiple devices to a single Amazon S3 bucket:

1. [Configure Cohesity Protection](#). Create multiple Protection Groups.
2. [Create Snowball Edge Import Jobs](#). Create multiple import jobs that are each configured to use the same Amazon S3 bucket.
3. [Register Snowball Edge as S3-compatible External Target](#). Register an External Target for every Snowball Edge device.
4. [Initiate On-demand Archival Seeding to the Snowball Edge Devices](#).

Before you initiate seeding each device:

- a) Ensure the storage consumed by a Protection Group does not exceed the maximum storage capacity of the Snowball Edge device.
  - b) Seed data from a Protection Group only to a single S3-compatible External Target.
  - c) Make note of which Protection Group is using which External Target, as this mapping is necessary to ensure continuous archival.
5. [Return the Snowball Edge devices](#).
  6. [Register Amazon S3 Buckets as Cohesity External Targets](#). Register multiple External Targets, one per Snowball Edge device. If you plan to use the same Amazon S3 bucket for all your Snowball Edge devices, make sure the destination bucket is the same for all External Targets.
  7. [Initiate External Target Swap](#). Contact [Cohesity Support](#) to swap the Snowball Edge External Target with the Amazon S3 External Target for each Protection Group. Note that you need to do this for each Snowball Edge device.
  8. [Sync Incremental Backups to AWS S3 External Target](#). Ensure that External Target selected for the on-demand archival to the cloud is the Amazon S3 bucket that you swapped in the previous step.
  9. [Add Archival Schedule](#). Ensure that External Target selected for the archival schedule is the Amazon S3 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 tea and explore all the other ways Cohesity helps you mine your data for insights, value, and endless possibilities in our [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 Senior Manager, Tech Marketing, Solutions Engineering, and 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, Senior Technology Writer and Editor at Cohesity
- Benjamin Rualthanzauva, Technical Lead, SRE
- Kevin Hill, Senior Solutions Architect
- Sarthak Agarwal, MTS, Engineering
- Sharada Jayaprakash, Technical Writer
- Siddhesh Rumde, Sr. MTS, QA

## Document Version History

VERSION	DATE	DOCUMENT HISTORY
1.0	May 2020	First release

## 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](#).

© 2022. Cohesity, Inc. All Rights Reserved. The information supplied herein is the confidential and proprietary information of Cohesity and may only be used (a) by the intended recipients and (b) in conjunction with validly licensed Cohesity software and services. Find the terms of Cohesity licenses at [www.cohesity.com/agreements](http://www.cohesity.com/agreements).

Cohesity, the Cohesity logo, SnapTree, SpanFS, DataProtect, Helios, and other Cohesity marks are trademarks or registered trademarks of Cohesity, Inc. in the US and/or internationally. Other company and product names may be trademarks of the respective companies with which they are associated. This material (a) is intended to provide you information about Cohesity and our business and products; (b) was believed to be true and accurate at the time it was written, but is subject to change without notice; and (c) is provided on an "AS IS" basis. Cohesity disclaims all express or implied conditions, representations, warranties of any kind.