COHESITY

Version 1.6

April 2024

Protect Your EC2 Data with Cohesity

Cohesity for AWS Native Backup Best Practice Guide

ABSTRACT

Data loss is inevitable in any data center, even in data centers hosted by cloud service providers. With Cohesity AWS Native Backup, you can now protect your AWS Elastic Compute Cloud (EC2) and Elastic Block Store (EBS) Volumes against data loss. Our solution also provides the ability to restore data at granular levels to the same or different AWS accounts.

Table of Contents

Figures

Tables

AWS Native Backup Using Cohesity

Amazon Web Service (AWS) Elastic Compute Cloud (EC2) is a service that provides secure and on-demand compute resources in the cloud. With more and more organizations using EC2 instances to deploy enterprise workloads, it's imperative to think about backing up data on EC2 instances. Now you can protect your EC2 data while enjoying the granularity and flexibility of Cohesity platform.

AWS Backup Service Vs Cohesity AWS Native Backups

AWS provides a backup service for EC2 instances, but there are some caveats. Table 1 below compares the out-of-the-box backup features provided by the AWS Backup Service with AWS native backups using Cohesity platform.

Table 1: AWS Backup Service vs Cohesity AWS Native Backups

FEATURES	AWS BACKUP SERVICE	COHESITY AWS NATIVE BACKUPS
Scheduled Backup	Takes backups on a schedule.	Schedules backups. Administrators just need to specify the schedule and backups happen automatically.
Granular Recovery	No granular recovery. Recovery at the snapshot level.	Recovery can be done at different levels. Data can be recovered from an entire Protection Group, per EC2 instance, or at the files-and-folders level.
Choice of Target	Only AWS Backup Vault	Multiple target options, including on-premises, private clouds, and public clouds

Using Cohesity platform to perform your AWS backups gives you the ability to take control of your AWS infrastructure and associated data by providing the ability to do automated point-in-time backups of the EC2 instances and the associated EBS volumes. It allows you to restore at different levels of granularity, from entire EC2 instances to individual files and folders. What's more, you can use Cohesity platform to protect and recover your AWS data across multiple AWS accounts.

AWS EC2 Data Protection Methods

Cohesity platform offers two different ways to protect EC2 instances, Cohesity Native Snapshot and Cloud Snapshot Manager.

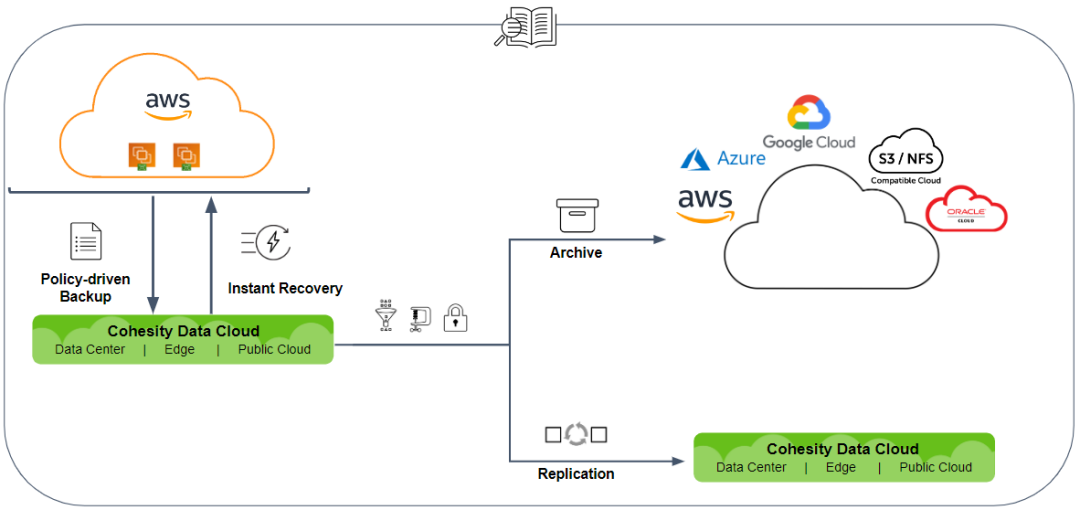
- **Native Snapshot**: The primary use case for Native Snapshot is protection of native cloud VMs from within the cloud. Data and metadata are stored on the Cohesity cluster. Cohesity recommends using Native Snapshot as the default method for protecting native cloud VMs because data is stored on the Cohesity cluster. With Cohesity snapshots, we can extend the full benefits of using the Cohesity platform to the protected data.
- **Cloud Snapshot Manager (CSM)**: The primary use for Cloud Snapshot Manager (CSM) is the protection of native cloud VMs from on-prem (outside of the cloud) or across Cloud Service Provider (CSP) regions that would incur data egress charges. The CSM method manages native cloud snapshots, with only metadata being stored on the Cohesity cluster (no data is stored on the Cohesity cluster).

Cohesity's AWS Native Snapshot Solution

The AWS native snapshot feature of Cohesity platform gives you the ability to back up AWS EC2 instances along with the attached EBS volumes. This solution helps you protect your entire AWS infrastructure with backup and restore features that keep your data safe when you need to move it around. With your EC2 backups in Cohesity, you can protect them further with flexible replication to any Cohesity storage as well as archival to almost any storage platform.

NOTE: You can use Cohesity On-Premise or Cloud Edition (CE) clusters to back up your AWS EC2 instances.

Figure 1: Use Cohesity platform to Protect EC2s with Backup and Archive to Any Storage



How AWS Native Snapshot Works with Cohesity Platform

There are several steps involved in setting up Cohesity platform to take backups of AWS EC2 instances.

1. **Create AWS IAM User with access to EC2s**. To add your AWS data to Cohesity platform as a source, create an AWS IAM (Identity and Access Management) user (or role) with access to your EC2s. See Appendix C: Create an IAM User and Appendix D: Register cloud source using IAM role. Review AWS Permissions.

   > **NOTE**: Roles are only supported for CE clusters.

2. **Register AWS account as a source in Cohesity platform**. Use the IAM user (or role) and its Access Key ID, Secret Access Key, and Amazon Resource Name (ARN) to add your AWS account as a source in Cohesity platform. See Register or Edit an AWS Cloud Source in the online Help.

3. **Add EC2 instances to Cohesity Protection Group**. Once your AWS IAM user, with access to your EC2s, is registered with Cohesity platform, you can discover and add EC2 instances to a Protection Group. From there, you can filter discovered EC2 instances using EC2 instance names or tags, and then add them to a Protection Group.

   > **NOTE**: You can also use EC2 names and tags to create Auto-Protect rules that will automatically include future EC2 instances that match the Auto-Protect rules.

   a) With 6.6.0D onwards, Cohesity allows excluding of EBS volumes from EC2 instance backup. Specific volumes can be deselected while creating or editing a Protection Group.

   > **NOTE**: Excluding root disk is not supported. Recoveries will fail if the excluded volume is necessary for booting up the VM.

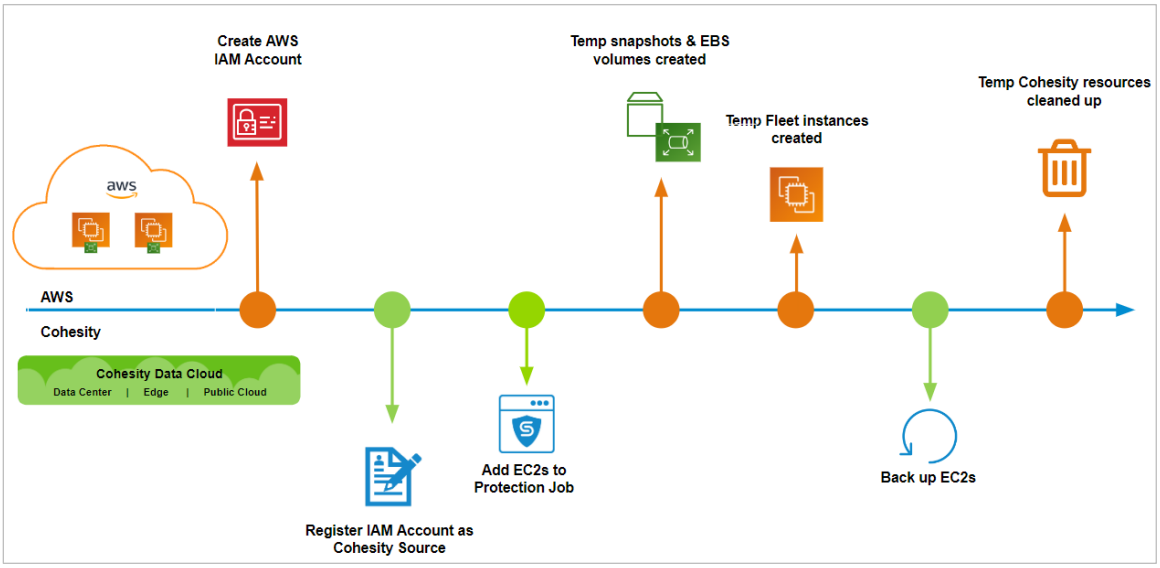
4. **Cohesity platform creates snapshots and temporary EBS volumes**. Snapshots are created for the EBS volumes associated with an EC2 instance, and temporary EBS volumes are created from these snapshots.

5. **Cohesity platform creates fleet instances**. Cohesity platform launches fleet instances and attaches the temporary EBS volumes to the fleet instances.

6. **Backup data to Cohesity platform**. The Protection Group copies data from the fleet instance to Cohesity platform using the remote agent (which is preinstalled on the fleet instance).

7. **Resource Cleanup**. Resources created for the backup/restore process — such as fleet instances and temporary snapshots and EBS volumes — are cleaned up after the backup or restore process completes.

Figure 2: AWS Native Backups with Cohesity Workflow



Understand Cohesity Fleet Instances

Cohesity fleet instances are a crucial component of Cohesity's backup and recovery solution for AWS EC2 instances. Fleet instances are ephemeral EC2 instances created by Cohesity platform on AWS. They act as proxies to which the temporary EBS volumes are attached.

There are three types of Cohesity fleet instances. See Table 2 below.

Table 2: Cohesity Platform Fleet Instance Types

FLEET INSTANCE TYPE	DESCRIPTION
Generic Fleet Instance	Fleet instances are C5.large (if not available, then m5.large) type EC2 instance that is created by Cohesity platform. The Generic Fleet Instance acts as a proxy host, used to mount the temporary EBS volumes. Multiple fleet instances can be instantiated, depending upon the number of EC2 instances being backed up. A Generic Fleet Instance is launched when no marketplace identifiers are found for the EC2 instance. They are terminated after backup completion if there is no other backup activity for that AWS source for 30 mins.
Marketplace Fleet Instance	These will be used if the source EC2 instance being protected has an EBS volume with marketplace code. A new Marketplace Fleet Instance is created for each source EC2 instance. Each instance is started and shut down before and after backup.

FLEET INSTANCE TYPE	DESCRIPTION
Restore Fleet Instance	Restore fleet instances are launched for performing restores. When the EBS volume being restored is greater than 1 TB in size, a new Restore Fleet Instance is launched for each EBS volume needed to perform the restore. This is done to comply with the object and part size restrictions in the API for S3 multipart uploads.

Each fleet instance has the Cohesity Remote Agent pre-installed. The remote agent is responsible for doing the backup of data from the EBS volumes to Cohesity clusters.

A Cohesity fleet instance is instantiated based on the rules outlined in the next section, Factors Governing Fleet Instance Creation.

NOTE: Like any running EC2 instance, fleet instances incur some additional costs to customers.

Factors Governing Fleet Instance Creation

Table 3 below describes the factors that govern the creation and termination of Cohesity fleet instances.

Table 3: Criteria Governing Fleet Instance Creation

CRITERIA	CONSIDERATION
Region	Fleet instance is instantiated per account per region for the EC2s being backed up. For example, if you have two EC2 instances being backed up that are in different regions, two different fleet instances are instantiated.
Size of EBS volumes	One fleet instance per 1 TB of data. A new fleet instance is spun up if the sum total size of the EBS volumes being backed up exceeds 1 TB. If the size of the volume is greater than 1 TB, a dedicated fleet instance is created for it. To change the size limit with the gflag, contact Cohesity Support.
Number of EBS volumes	The Default fleet instance type C5.large, allows 4 EBS volumes per instance. No more than 4 EBS volumes can be attached on a single fleet instance. If the number of volumes exceeds 4, a new fleet instance is spun up. You can change the instance type to increase the EBS volume count. E.g.: t3.micro, small, and medium supports up to 8 EBS volumes. With t3.large instance, you can connect a maximum of 12 EBS volumes.
Activity	A fleet instance is terminated if there is no activity for 30 mins. Otherwise, it is reused for incoming backup tasks. Restore fleet instances are terminated immediately after the restore task is finished.

NOTE: By default, Cohesity platform creates a maximum of 4 concurrent fleet instances. The limit is set to optimize the workloads running on the Cohesity cluster, and to accommodate AWS cloud limits. Cohesity doesn't allow users to manually spin up fleet instances to speed up the backup or restore process.

Native Snapshot Using AWS EBS Direct APIs

Cohesity also provides flexibility to use AWS EBS direct APIs to back up and restore the EC2 instances. By using the EBS direct APIs, Cohesity can identify the delta from the previous snapshots and copies only the incremental changes to Cohesity storage. Contact Cohesity Support to enable this feature.

NOTE: Enabling the backup gflag is a cluster-wide setup and cannot be set at the Protection Group level. Once enabled, all native snapshot Protection Groups will use EBS direct APIs method.

Figure 3: Native Snapshot Using AWS EBS Direct APIs

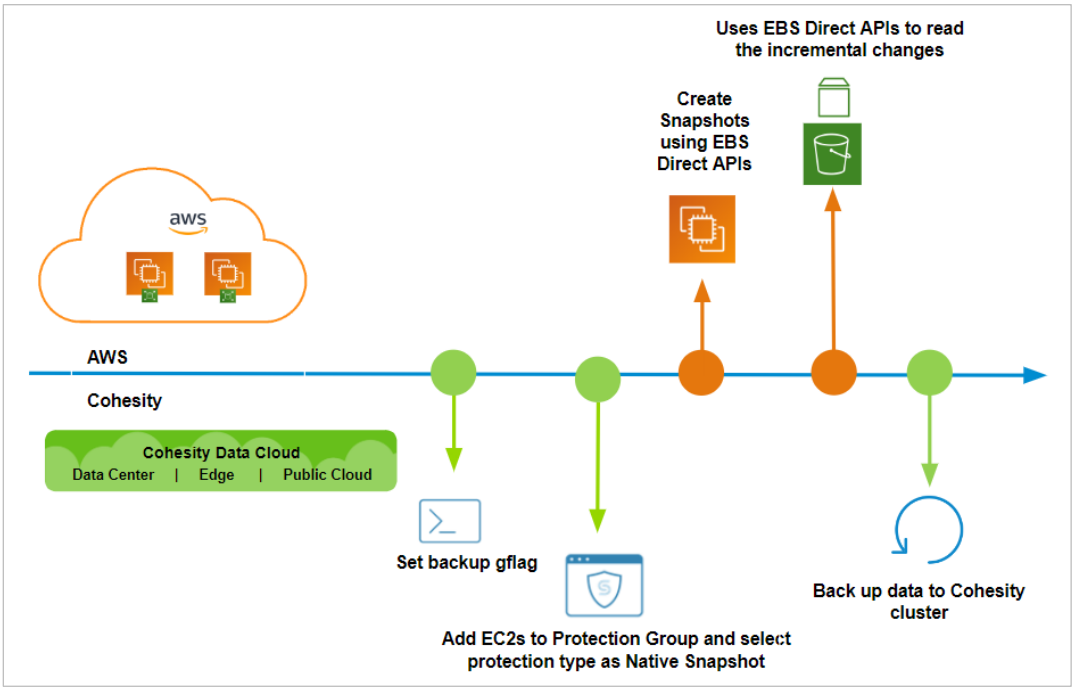


Table 4: Native Backup Using EBS Direct APIs Workflow

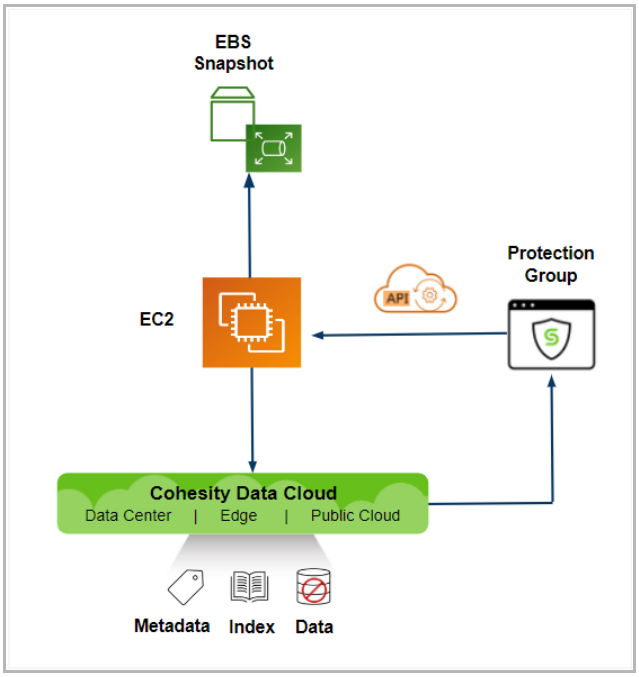
PHASE	NATIVE BACKUP USING EBS DIRECT APIS WORKFLOW
1	To enable the feature, contact Cohesity Support.
	Add EC2s to protection group and select the protection type as Native Snapshot.
2	Snapshot is created for every attached disk using the Create Snapshot API.
3	For the initial backup, `ListSnapshotBlocks` API is called to get the blocks in the EBS snapshot.
4	For incremental backup, `ListChangedBlocks` API is called to identify the delta.
5	Data is copied to Cohesity storage.

NOTE: Using the EBS direct API backup method will incur additional cost as Cohesity keeps the last outstanding snapshot for each disk(s) to identify the incremental changes.

Cohesity's Cloud Snapshot Manager (CSM) Backup Solution

From 6.4.1 onwards, Cohesity also provides an alternative EC2 backup solution called Cohesity's Cloud Snapshot Manager (CSM) backup solution. Cohesity uses CSM APIs to back up the EC2 instances. When CSM is used, only metadata and index data is stored on the Cohesity cluster. All snapshot data gets stored in AWS itself. This means there is very little data egress associated with CSM (metadata and index only). As a result, this is an efficient approach for using an on-prem Cohesity cluster to back up native cloud VMs. CSM backup does not support storage efficiency features like deduplication compression. Additionally, you have to account for the relatively small egress charges with respect to the metadata size.

Figure 4: CSM Backup Solution

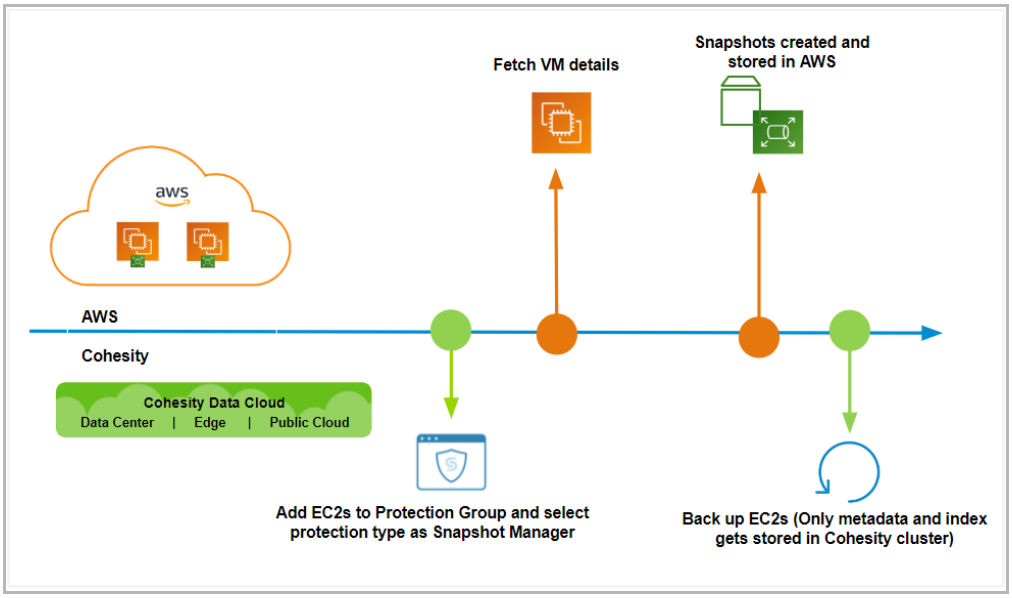


How Cloud Snapshot Manager Backup Works

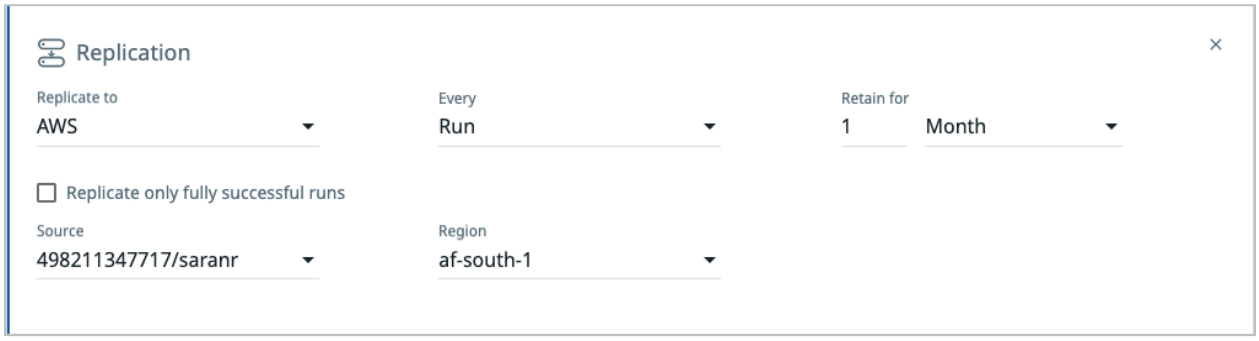
Setting up Cohesity platform to take backups of AWS EC2 instances using CSM backup involves the following four steps:

1. **Select Protection Group type as Snapshot Manager**. While creating the Protection Group, select the protection type as Snapshot Manager.
2. **Fetch VM details.** Cohesity fetches the VMs details using the GET APIs.
3. **Create EBS Snapshot.** Creates EBS Snapshots using APIs and stores it in AWS.
4. **Back up metadata and index.** Once EBS Snapshots are created, Cohesity backs up the metadata and index to the source storage.

Figure 5: CSM Backup Workflow



If the policy is attached with replication to AWS, Cloud Snapshot Manager can replicate the data to different AWS regions or accounts. To replicate EBS snapshots to another AWS region, select the **AWS** option for the **Replicate to** field. A registered AWS hypervisor source (IAM user/IAM role) and the region that snapshots will be replicated into must be selected. This type of replication does not require a Cohesity cluster to serve as the destination for the replication data. The AWS region serves as the destination. Cohesity replicates the snapshots to the EBS bucket in the selected region.



Design Decisions and Best Practices

You must make important decisions about choosing a particular backup or recovery workflow that solves a given problem. This section provides a decision tree on when you should use a workflow and the best practices around that.

Backup Using Native Snapshot

Let's understand when you can use this workflow for EC2 protection:

- Use if you are looking for AWS instance protection from within the cloud.
- Use if you want to recover file/folder-level granularity.
- Use if you want to reduce cloud storage costs with data deduplication and compression.
- Use if you want to leverage Cohesity archival & cluster level replication.

Backup Using Native Snapshot with EBS Direct APIs

Apart from the details listed under section Backup Using Native Snapshot, the following details help you in considering EBS direct API backup:

- Use if you want faster backup than the native snapshot method.
- Use if you understand the additional storage costs for the outstanding snapshots for every disk in VMs.

Backup Using Cloud Snapshot Manager

Let's understand when you can use this workflow for EC2 protection:

- Use if you want to protect the AWS instances from outside of the cloud (e.g.: protect the AWS instances using an on-prem Cohesity cluster).
- Use if you want to recover only at the instance level (no granular recoveries).
- Use if you want to replicate the EBS snapshots to another AWS region without the need of an additional Cohesity cluster.
- Use if you want quicker recovery of instances.

Backup Using a Combination of Native Snapshot and CSM Methods

If you have larger VMs and want to achieve quicker backup and require granular recoveries, then you can use both the methods. The CSM method will provide a quicker backup compared to long-running backup operations in the native snapshot method. Along with the CSM method, many customers also enable the native snapshot method (in a different protection group) to leverage the granular recovery options as well.

Design Considerations

- Cohesity always deploys the fleet instances in the same VPC where the Cloud Edition is running. If the customer has multiple accounts and account A is running Cloud Edition and wants to protect EC2 instances in Account B, deploying fleet instances in account A will invite additional egress cost. If you are protecting instances from a different account, ensure that you select a VPC from the same account to launch the fleet instances.
- Cohesity fleet instances must have port 50051 and 22 open, so that the Cohesity cluster can communicate with fleet instances.
- Cohesity creates one network security group in each VPC and attaches to the fleet instance. By default, the inbound rules at the security group will be configured with the CIDR range of the network where the Cohesity cluster is configured.
- Cohesity requires ports 443, 50051 enabled on the target for File and Folder recovery. The EC2 instance should have an SSM agent installed and the *AmazonEC2RoleforSSM* policy attached to it.
- Delete is also a required permission you need to grant to Cohesity EC2 backup solution. Although Cohesity doesn't require delete permission for backup, whenever the customer wants to delete any of the backups, Cohesity requires delete permissions to clean the resources created by the Cohesity backup solution.
- From 6.8 onwards, Cohesity locks down the fleet instance's security group to the VPC. If you want to back up or recover to a different subnet or VPC, the operation will fail unless you unlock the security group hardening. Contact Cohesity Support to unlock the security group hardening.

AWS EC2 Recovery Methods

Based on the protection type you have selected, Cohesity allows you to recover the data at different levels of granularity.

Cohesity AWS Native snapshot Recovery

Recovery from Cohesity AWS native snapshot can be performed at different levels of granularity, and to both the original and alternate locations. Data can be restored across AWS accounts, regions, and Amazon Virtual Private Clouds (VPCs), as well.

The restore granularity levels are:

1. **Protection Group**. All EC2 instances that are part of a Protection Group can be recovered to a previous point in time using backup snapshots in a single click.
2. **EC2 Instance**. EC2 instances can be searched using EC2 instance names, or the tags associated with them, and recovered from a backup snapshot.
3. **Files and Folders**. Starting with version 6.4, files and folders that reside in an EC2 instance can be restored to the EC2 instance. The EC2 instance should have an SSM agent installed and the *AmazonEC2RoleforSSM* policy attached to it.

> **NOTE**: Cohesity platform currently doesn't have the option to recover specific EBS volume(s) and attach them to a different instance. You can achieve this by recovering the entire instance and then reattach the volume(s) to the required instance.

How Recovery Works with AWS Native Backup

Recovery is the most important aspect of any backup solution, and as discussed above, with Cohesity platform, you can recover data at every level of granularity.

Although there are slightly different workflows, they all involve these steps:

1. Go to **Protection > Recovery**, click **Recover** and select **VMs**.
2. Search for an EC2 instance using its name or assigned tags.
3. Select the EC2 instance to be restored and initiate a restore.

Different Restore Workflows

There are different workflows for data recovery, based on the following factors:

- **EBS volume size**. Having different workflows for EBS volumes of different sizes enables optimized restore performance, and minimizes the cost incurred for starting fleet instances for small EC2 instances.
- **Operating System (OS) on the EC2 being restored**. The operating system running on EC2 being restored determines the restore workflow. This is because, for Linux distributions, AWS allows Linux AMI (Amazon Machine Image) creation from a snapshot, meaning both the disks and the AMI can be created from the snapshot copied during backup. However, for the Windows OS, AWS doesn't allow AMI creation from snapshots, which is where the restore process differs between Linux and Windows.
- **EBS volume encryption**. If the EBS volumes being backed up are encrypted, the restore approach is different from the one followed for unencrypted volumes.

> **NOTE**: The Cohesity VM import role is needed by Cohesity when doing EC2 instance recovery in AWS. If this role does not already exist in your account for the AWS region, Cohesity creates it during the recovery operation.

EC2 with Linux and Disk Size Less Than 1 TB

For recovering a Linux instance with an EBS volume size less than 1 TB, Cohesity platform copies the backup files to an S3 bucket and uses the files to restore the EC2 instance. Figure 6 shows the general restore workflow in these parameters, followed by detailed steps for both unencrypted and encrypted volumes.

Figure 6: Restore Workflow for Linux Instances with EBS Volumes < 1 TB

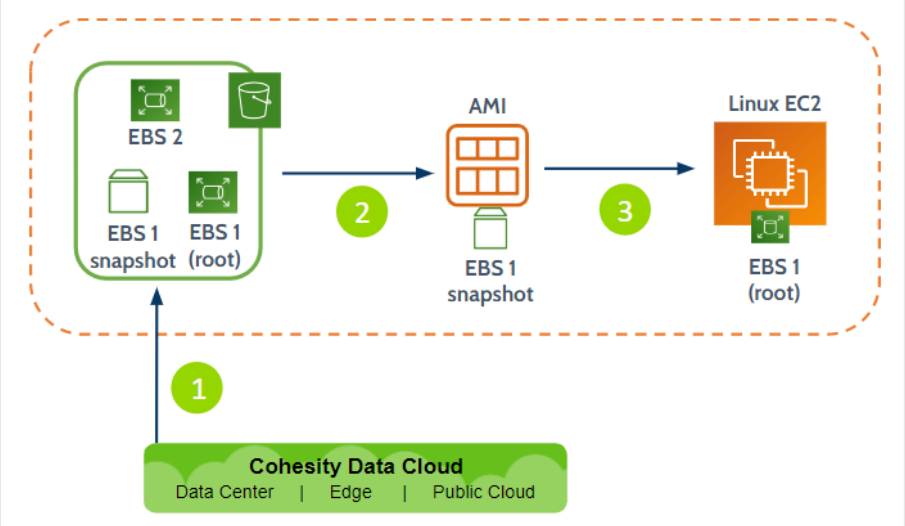


Table 5: EC2 Restore Process on Linux Instances with EBS Volumes < 1 TB

<table>
<tr><th rowspan="2">PHASE</th><th colspan="2">RESTORE STEPS FOR EBS VOLUMES < 1 TB (LINUX)</th></tr>
<tr><th>UNENCRYPTED</th><th>ENCRYPTED</th></tr>
<tr><td rowspan="4">1</td><td colspan="2">The raw disk files for the instance being restored are copied from Cohesity platform to an S3 bucket created using the IAM user credentials of the account to which the restore is being done.

NOTE: If the VM import role is not present, Cohesity will create the role to perform the recovery.</td></tr>
<tr><td colspan="2">For root and data volumes, we import the raw disk files from S3 into the EC2 service, to create EBS snapshots.</td></tr>
<tr><td>n/a</td><td>For an EC2 instance that has encrypted volumes, a copy of this snapshot is created by encrypting the data using the KMS key alias.</td></tr>
<tr><td colspan="2"></td></tr>
<tr><td>2</td><td colspan="2">A Linux AMI is created using the snapshots of the root and data disks imported from S3.</td></tr>
<tr><td>3</td><td colspan="2">An EC2 instance is launched using the AMI and the instance metadata stored on the Cohesity cluster. This creates the EBS volumes for the root and data disks and attaches them to the instance.</td></tr>
</table>

EC2 with Linux and Disk Size Greater Than 1 TB

For recovering a Linux instance with an EBS volume size greater than 1 TB, Cohesity copies the backup files to an S3 bucket and creates fleet instances to use them to restore data. Figure 7 shows the general restore workflow in these parameters, followed by the detailed steps for both unencrypted and encrypted EBS volumes.

Figure 7: Restore Workflow for Linux Instances with EBS Volumes > 1 TB

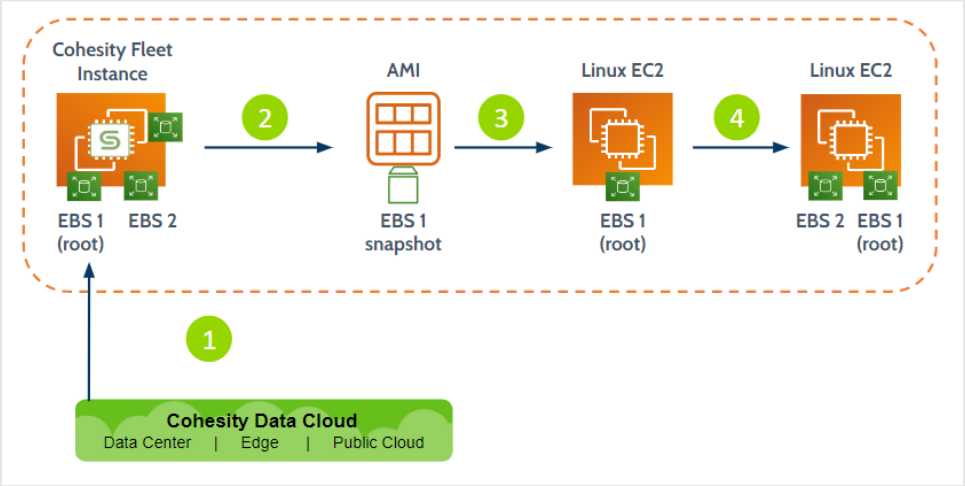


Table 6: EC2 Restore Process on Linux Instances with EBS Volumes > 1 TB

PHASE	RESTORE STEPS FOR EBS VOLUMES > 1 TB (LINUX)	
	UNENCRYPTED	ENCRYPTED
1	A Cohesity fleet instance is created.	
	Empty EBS volumes are created corresponding to the root and data volumes for the EC2 instance being restored.	n/a
	n/a	Empty encrypted EBS volumes are created corresponding to the root and data volumes for the EC2 instance being restored.
	These volumes are attached to the fleet instance. If the volume has marketplace codes, the fleet instance is stopped before the volumes can be attached.	
	Data from Cohesity platform is copied to these EBS volumes using the Cohesity Remote Agent.	
2	A snapshot is created for the root volume.	

PHASE	RESTORE STEPS FOR EBS VOLUMES > 1 TB (LINUX)	
	UNENCRYPTED	ENCRYPTED
	An AMI is created from the root volume snapshot.	
3	An AMI is launched with the snapshot, and we have an EC2 instance with the root volume.	
4	The data volumes are then detached from the fleet instance and attached to the EC2 created from the AMI.	

EC2 with Windows and Disk Size Less Than 1 TB

For recovering a Windows instance with an EBS volume size less than 1 TB, Cohesity copies the backup files to an S3 bucket and uses the files to restore the EC2 instance. Figure 8 shows the general restore workflow in these parameters, followed by the detailed steps for both unencrypted and encrypted EBS volumes.

Figure 8: Restore Workflow for Windows Instances with EBS Volumes < 1 TB

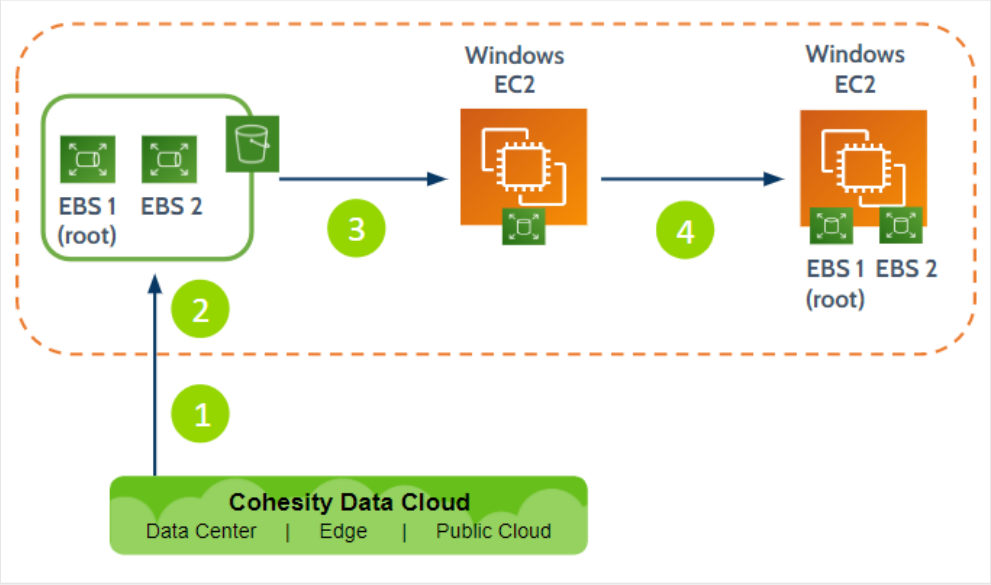


Table 7: EC2 Restore Process on Windows Instances with EBS Volumes < 1 TB

PHASE	RESTORE STEPS FOR EBS VOLUMES > 1 TB (LINUX)	
	UNENCRYPTED	ENCRYPTED
1	The raw disk files for the instance being restored are copied from Cohesity platform to an S3 bucket created by using the IAM user credentials of the account to which the restore is being done.	
2	For root and data volumes, we import the raw disk files from S3 into the EC2 service, to create EBS snapshots.	
	EBS volumes for root and data disks are created from the imported snapshots.	n/a
	n/a	For an EC2 instance that has encrypted volumes, a copy of this snapshot is created by encrypting the data using the KMS key alias, and the EBS volumes for root and data disks are created from the encrypted snapshots.
3	A Windows EC2 instance is launched using a Windows AMI (hosted publicly by Cohesity).	
4	The Windows EC2 instance is stopped, the root volume is replaced with the one created from backed up data, and the data volumes are attached.	

EC2 with Windows and Disk Size Greater Than 1 TB

For recovering a Windows instance with an EBS volume size greater than 1 TB, Cohesity copies the backup files to an S3 bucket and creates fleet instances to restore the data. Figure 9 shows the general restore workflow in these parameters, followed by the detailed steps for both unencrypted and encrypted EBS volumes.

Figure 9: Restore Workflow for Windows Instances with EBS Volumes > 1 TB

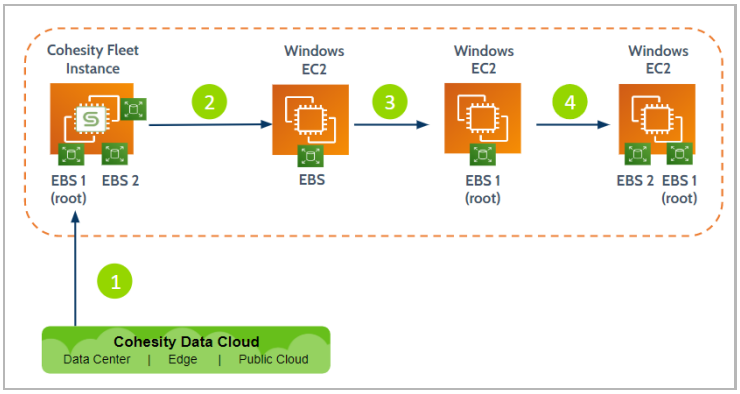


Table 8: EC2 Restore Process on Windows Instances with EBS Volumes > 1 TB

PHASE	RESTORE STEPS FOR EBS VOLUMES > 1 TB (WINDOWS)	
	UNENCRYPTED	ENCRYPTED
1	A Cohesity Fleet Instance is created.	
	Empty EBS volumes are created corresponding to the root and data volumes for EC2 being restored and are attached to the fleet instance.	n/a
	n/a	Encrypted empty EBS volumes are created corresponding to the root and data volumes on EC2 being restored and are attached to the fleet instance.
	Data from Cohesity platform is copied onto these EBS volumes.	
2	A Windows instance is created from an AMI (publicly hosted by Cohesity).	
3	The Windows instance is stopped, and the root volume of the instance is replaced with the root volume restored on the fleet instance.	
4	The data volumes are detached from the fleet instance and attached to the EC2 instance.	
	The EC2 instance is started.	

IMPORTANT

For encrypted volumes:

- During backup, the data is read using a fleet instance. It is unencrypted by AWS as fleet instances read the data. The decryption is performed in advance by AWS, ‘under the hood'.
- For restore, if the volume was encrypted with a key having the alias ‘ProdKey’ at the time of backup, we look for a key that has the same alias in the region and account to do the restore.
  - If the key has been deleted, the customer can manually create another key with the same alias as the original and the restore will use it.
  - In case of key rotation by AWS, a new key with the same alias is automatically generated by AWS, so the backup and restore operations proceed seamlessly.
- Encryption with cross-account Customer Managed Keys (CMK):
  - In the AWS CMK account, you must specify the other different accounts that can access this key in the AWS Key Management Service (KMS).
  - In the KMS Key policy, update Allow use of the key section and Allow attachment of persistent resources section with AWS IAM account (or role) that is used for source registration on the Cohesity cluster.
  - In the accounts that are going to use this cross-account CMK, you must create an IAM policy and assign it to the respective IAM users (or roles) that are being used for AWS source registration on the Cohesity cluster.

Cohesity's Cloud Snapshot Manager (CSM) Recovery

During the restore process, no data is transferred from Cohesity to AWS as the data is stored in AWS itself. When the user selects a snapshot to recover, Cohesity gets the Snapshot from AWS and restores the instance to the original or a new location.

Figure 10: CSM Recovery

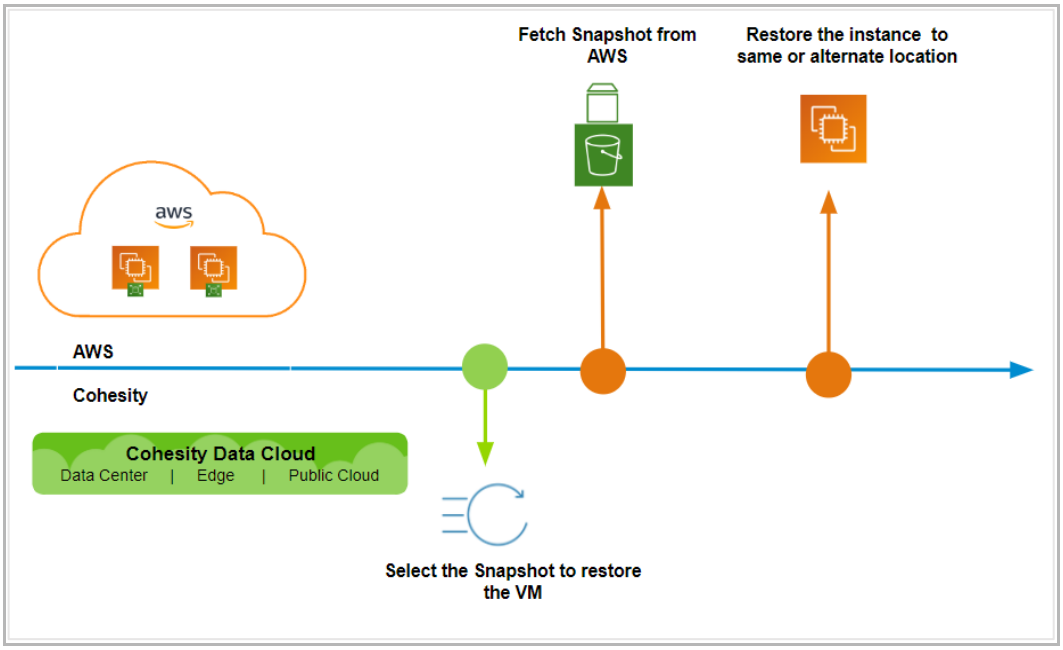


Table 9: CSM Recovery Workflow

PHASE	CSM RECOVERY WORKFLOW
1	Select the desired snapshot from which you want to restore the VM.
	The metadata is stored in Cohesity, and this offers quicker snapshot search and selection.
2	Fetches the snapshot from AWS.
3	Restores the instance to the same location.
4	If an alternate location is selected, Cohesity first copies the snapshot to the new location and then restores the instance.

NOTE: File and folder restore is not supported for CSM recovery.

AWS Resources Created by Cohesity Platform

While running AWS native snapshots, Cohesity platform creates temporary AWS resources under the IAM user account.

Table 10: Resources Created in AWS by Cohesity Platform

RESOURCE	WORKFLOW	COUNT	REASON	LIFE CYCLE	NAMING CONVENTION
Security group	Backup	One per account per region per VPC protected	Fleet Instances	Permanent	Security “Group Name” will be cohesity_fleet_sg
Generic Fleet Instances	Backup	>=1 per account per region. Could be more if multiple backups are running in parallel	Fleet Instances	Temporary, terminated after 30 min idle time	Name: gen_cohesity_fleet_ <task_id>
				Tag: cohesity_fleet_instance	
Marketplace Fleet Instance	Backup	>=1 per account per region of the source VM being backed up	Fleet Instances	Permanent	Name: mkt_cohesity_fleet_ <task_id>
				Tag: cohesity_fleet_instance	
Restore Fleet Instance	Backup	>=1 per EC2 instance being restored	Fleet Instances	Temporary, until the data is uploaded to AWS	Name: rec_cohesity_fleet_ <task_id>
				Tag: cohesity_fleet_instance	
VMImport role	Restore	One per S3 bucket created	For import snapshot	Permanent	vmimport-chsty-<cluster_id>-<cluster_incarnation_id>_<region>
S3 bucket	Restore	One per account per region	For import snapshot	Permanent	chsty_<cluster_id>_<cluster_incarnation_id>_<region>

RESOURCE	WORKFLOW	COUNT	REASON	LIFE CYCLE	NAMING CONVENTION
S3 objects	Backup	One per backup run	For import snapshot	Temporary, until backup run finishes	source_volume_id>_<job_uid>_<job_instance_id>_<task_id>
EBS Snapshots	Backup/Restore	One per backup run per backup disk	For backup/restore	Temporary, until backup run finishes	cohesity_<cluster_id>_<cluster_incarnation_id>_snap_<source_volume_id>
EBS Volumes	Backup	One per backup run per backup disk	For backup/restore	Temporary until backup run finishes	cohesity_<cluster_id>_<cluster_incarnation_id>_vol_<source_volume_id>

Appendix A: AWS Native Snapshot Terminology

There are several terms that are especially important to understand as you learn about the architecture of AWS native snapshots on Cohesity platform.

Table 11: AWS Native snapshot Terminology

TERM	DEFINITION
Elastic Cloud Compute (EC2)	Compute allocation in the cloud. EC2 being backed up is addressed as “Source VM”.
Elastic Block Storage (EBS)	Block storage allocated in the form of disks to a EC2 instance.
Virtual Private Cloud (VPC)	Virtual network dedicated to your AWS account.
VPC Gateway	A virtual private gateway is the VPN concentrator on the Amazon side of a Site-to-Site VPN connection. You create a virtual private gateway and attach it to the VPC from which you want to create the Site-to-Site VPN connection.
Customer Gateway	A physical device or a software application on the customer's side of a Site-to-Site VPN connection.
Tags	A tag is a label that you or AWS assigns to an AWS resource. Each tag consists of a key and a value.
CloudFormation Template	A template is a declaration of the AWS resources that make up a stack. The template is stored as a text file whose format complies with the JavaScript Object Notation (JSON) or YAML standard.

Appendix B: Egress Cost Considerations

Table 12 below outlines the deployment models and associated egress cost for the backup and restore of data using Cohesity's AWS native backup solution.

Table 12: Egress Cost Considerations

COHESITY	REGION	BACKUP	RESTORE
On-Premise	NA	Yes	No
Cloud Edition	Same Region	No	No
	Different Region	Yes	Yes

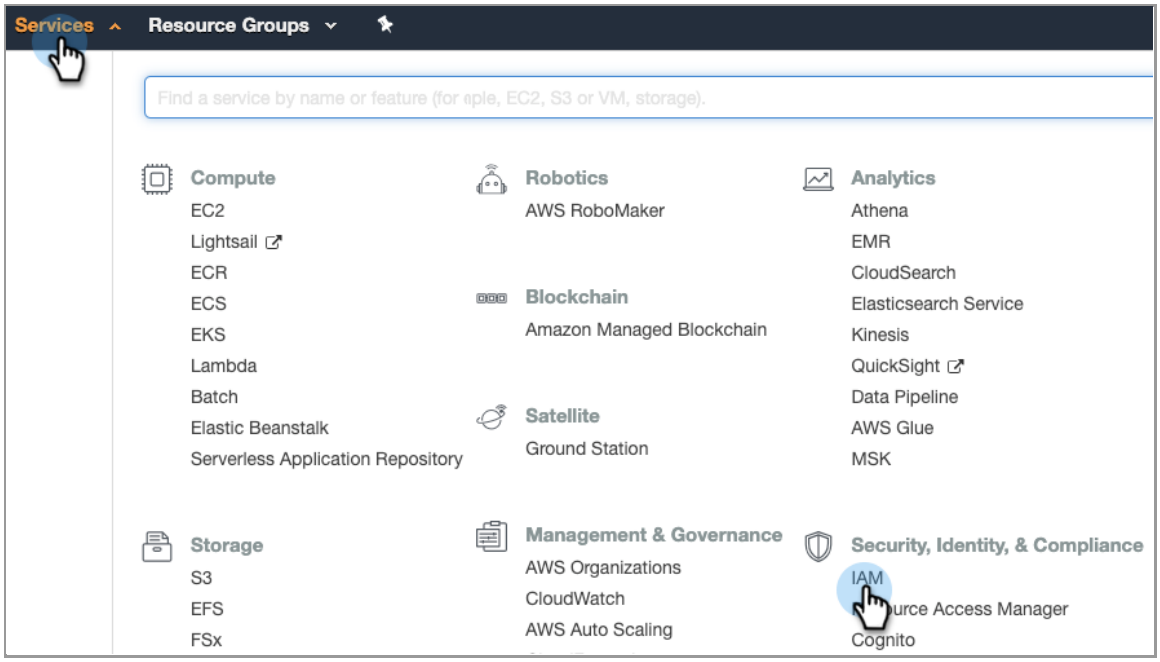
Appendix C: Create an IAM user to Register AWS Source with Cohesity Platform

Cohesity platform uses the AWS Cloud Formation Service to back up an EC2 in AWS. To restrict the Cohesity Control Instance's access to your AWS resources, Cohesity recommends that you create a new IAM user and grant that user the permissions required to back up an EC2 in AWS, as described below.

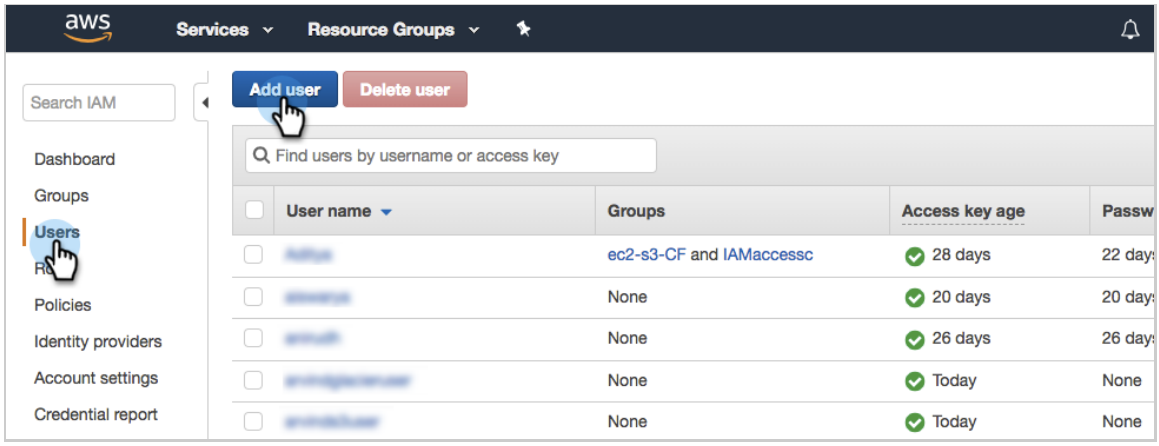
Create IAM User

To create a new IAM user:

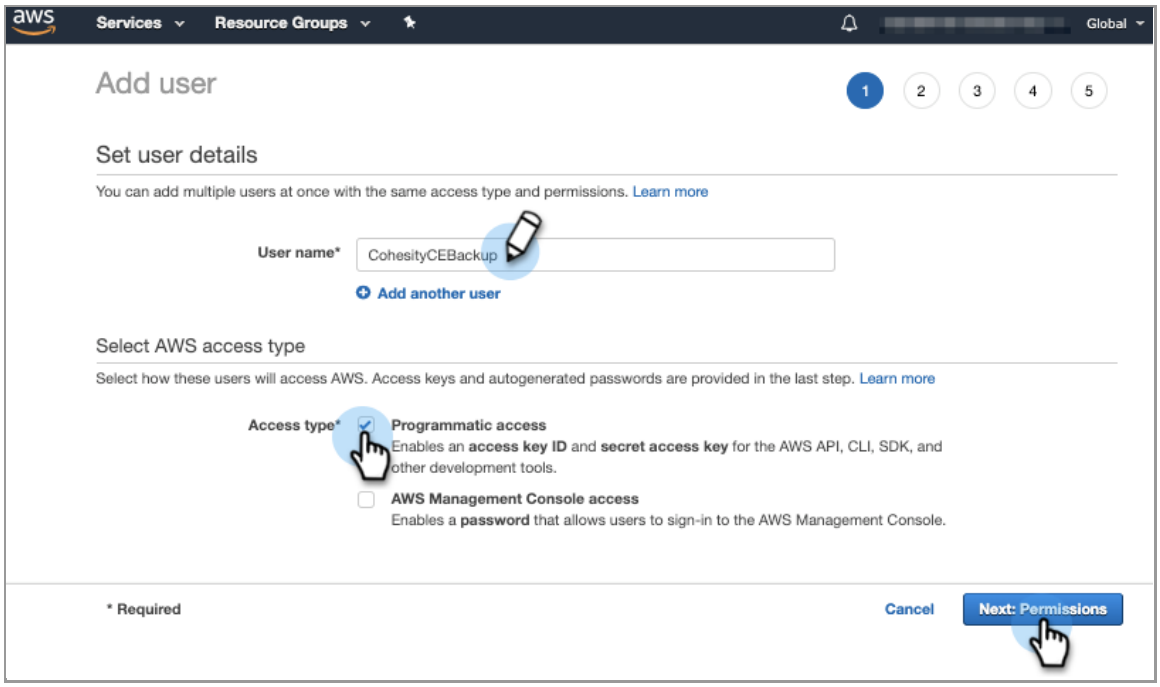
1. Log in to the Amazon AWS console with a user account that has the permissions to create IAM Policies.
2. In the top left, click **Services**, then, under **Security, Identity, & Compliance**, click **IAM**.



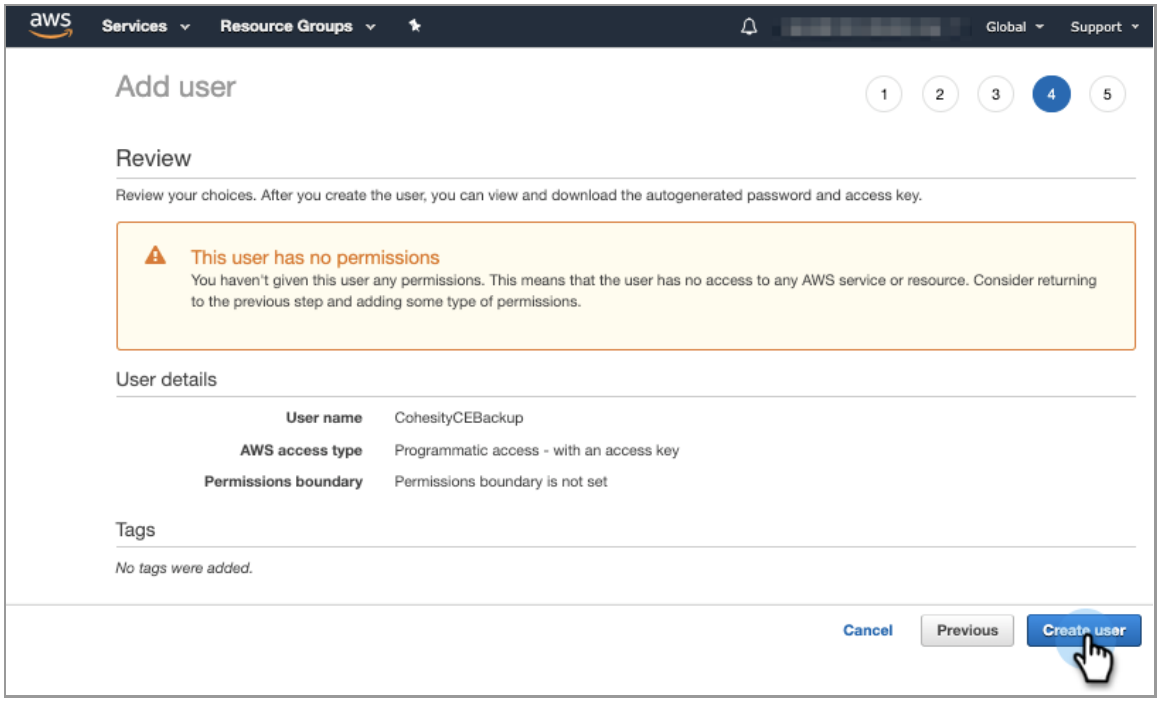
3. Click **Users** > **Add user**.



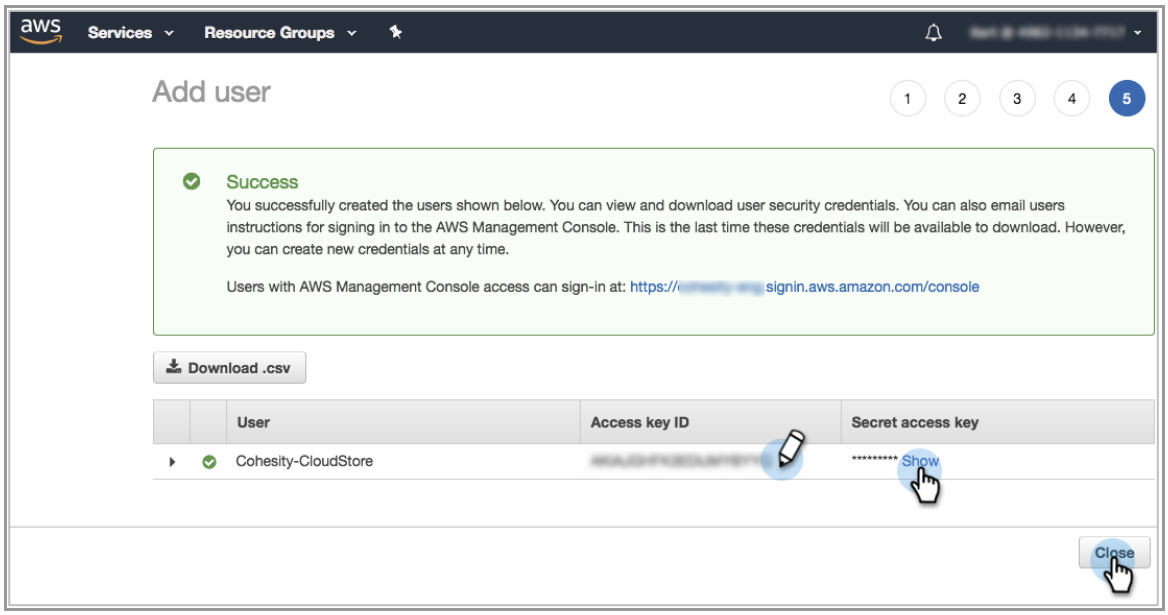
4. In the **User name** field, enter **CohesityCEBackup**, and for **Access type**, choose **Programmatic access**. Then click **Next: Permissions**.



5. In the **Set permissions** and **Add tags** pages, do not change any settings, and click **Next.**
6. In the **Review** page, ignore the “no permissions” warning and click **Create user**.



7. In the next screen, click **Show** to view and copy the **Access key ID** and **Secret access key**. Paste the **Access key ID** and the **Secret access key** into a settings text file. (You can also click **Download .csv** to save your key information for later use.) You will need these keys to register an AWS cloud source with Cohesity platform. Click **Close**.

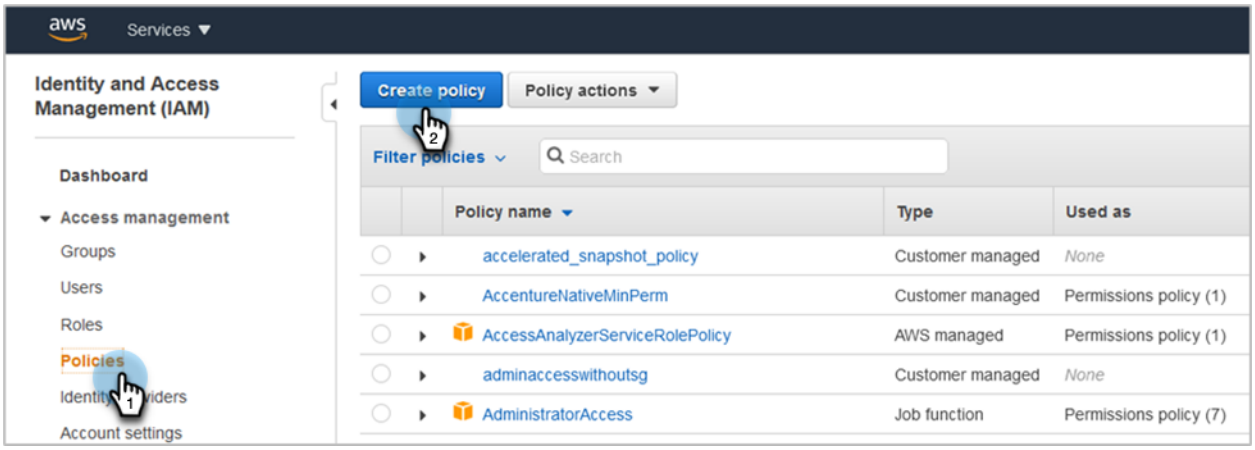


Your new user appears in the Users list.

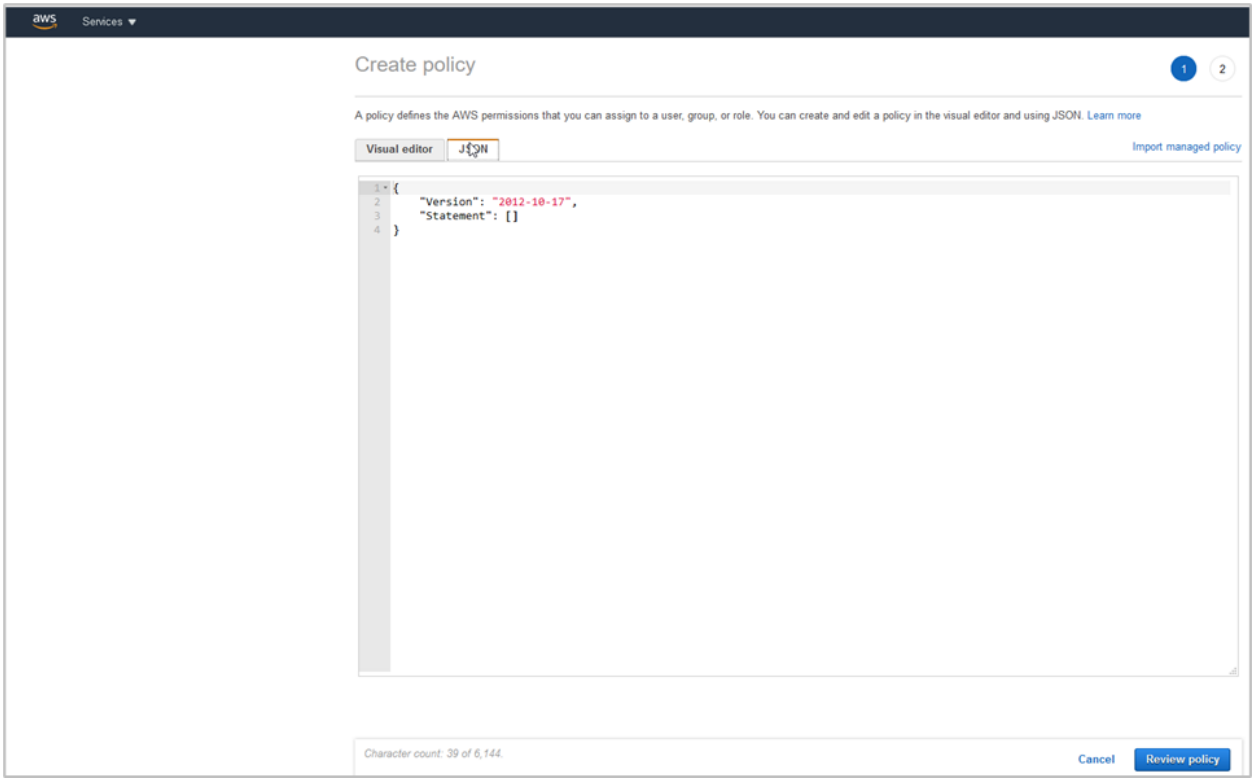
Create an IAM Policy

Create a new IAM policy and attach it to an IAM user to have the required permissions.

1. Go to **Services** > **Security, Identity & Compliance** > **IAM**.
2. On the left, click **Policies** and click **Create policy** to create a policy.



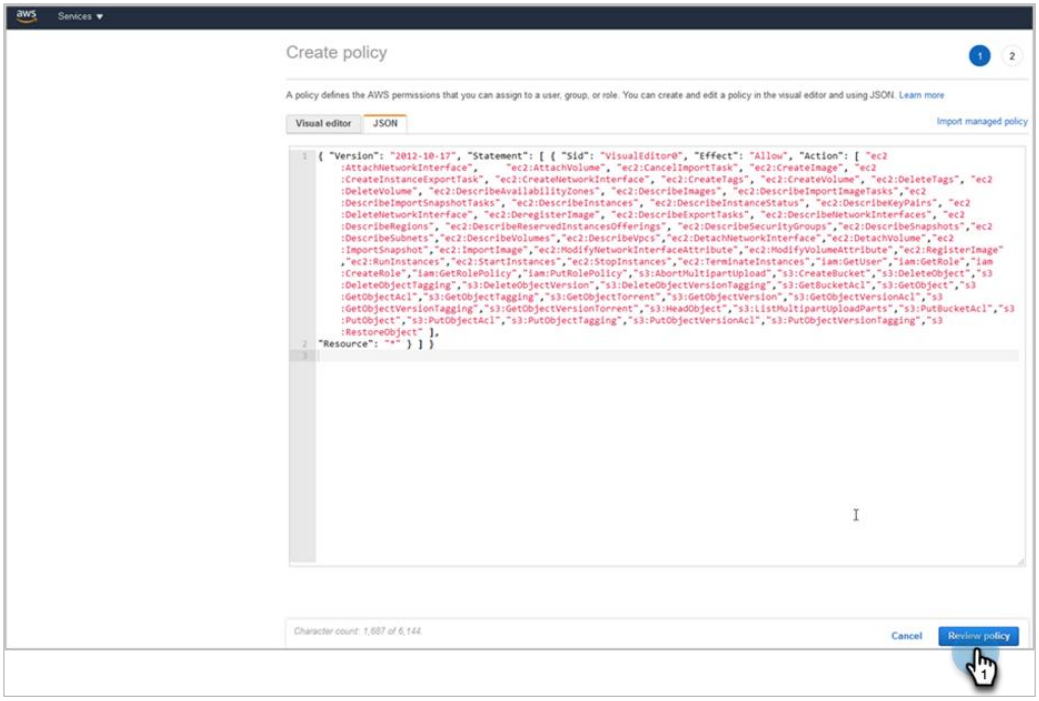
3. In the **Create policy** page, select the **JSON** tab.



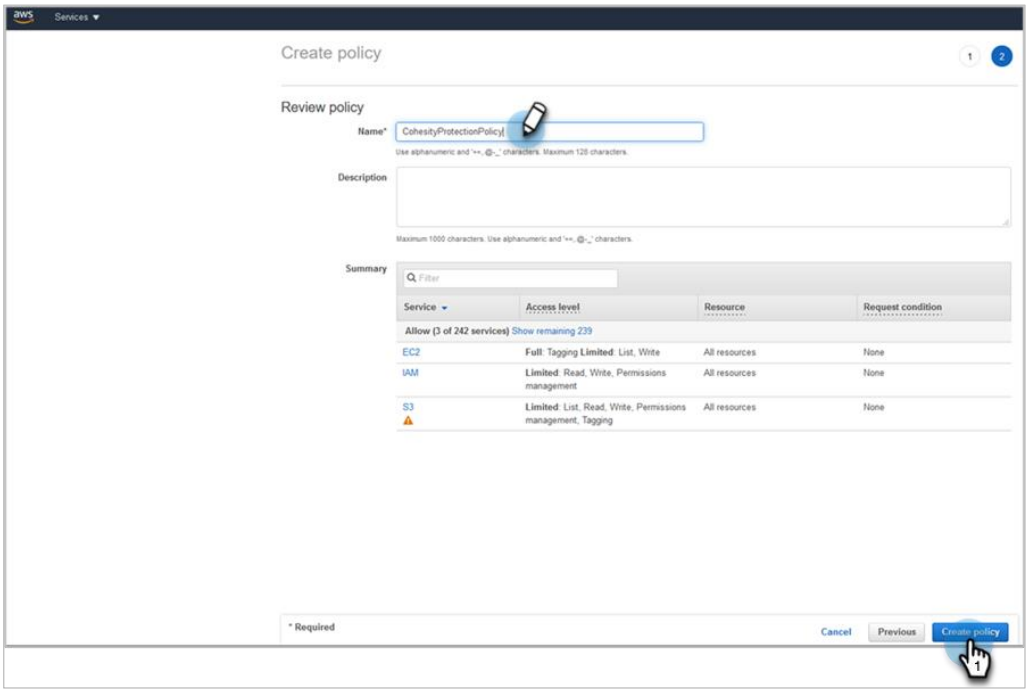
4. Delete the default JSON code. Copy the permissions from the Cohesity site.

> **NOTE**: Always review the latest permissions on the Cohesity site.

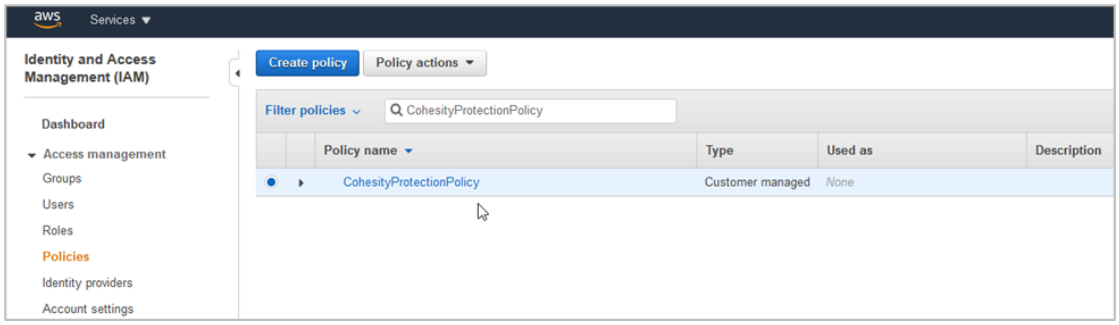
5. Click **Review policy** and continue.



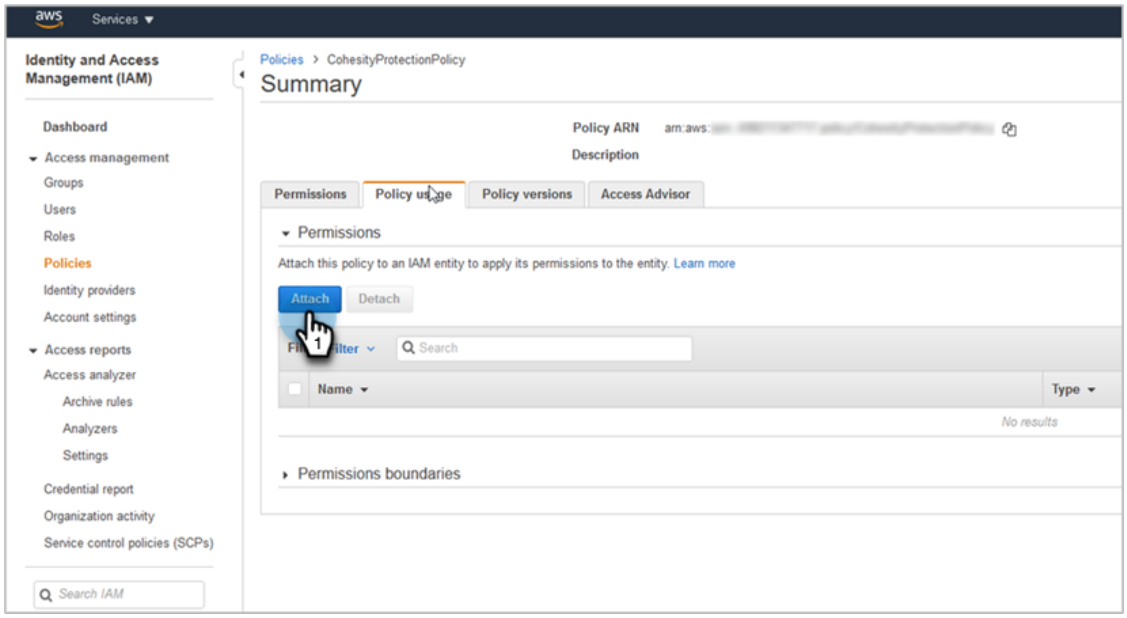
6. Enter **CohesityProtectionPolicy** as the policy **Name** and click **Create policy**.



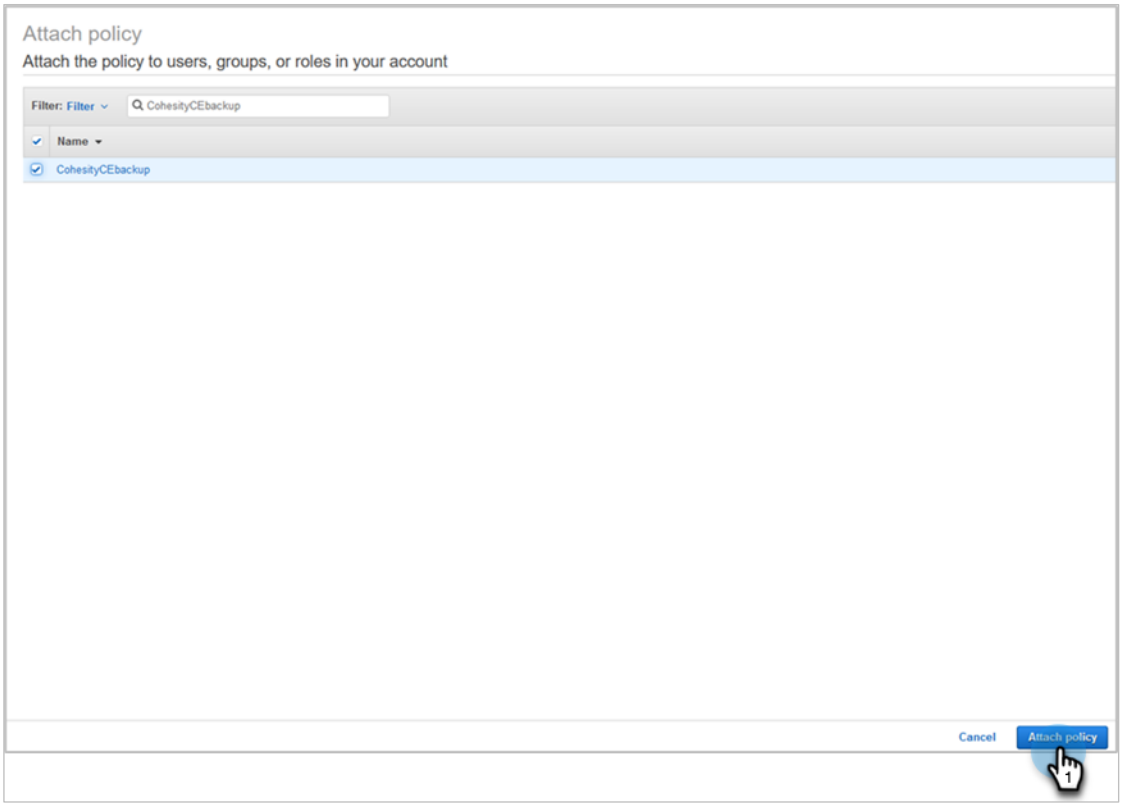
7. Once the policy is created, attach it to the user who is designated as a data protector user. Select the policy and click the link to the next step.



8. Select **Policy usage** tab and click **Attach** to go to the next step.



9. Search for the appropriate user and click **Attach policy** to activate the policy for the user**.**



Once the AWS IAM user is created with the required policies, you can use this user's **Access key ID** and **Secret access key** to register the AWS environment as a source with Cohesity.

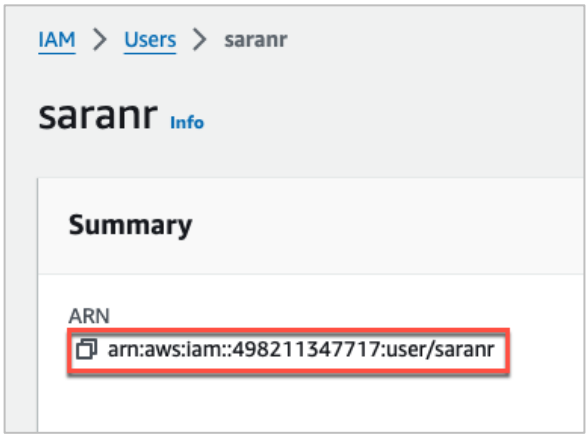
For instructions, see Register or Edit an AWS Cloud Source.

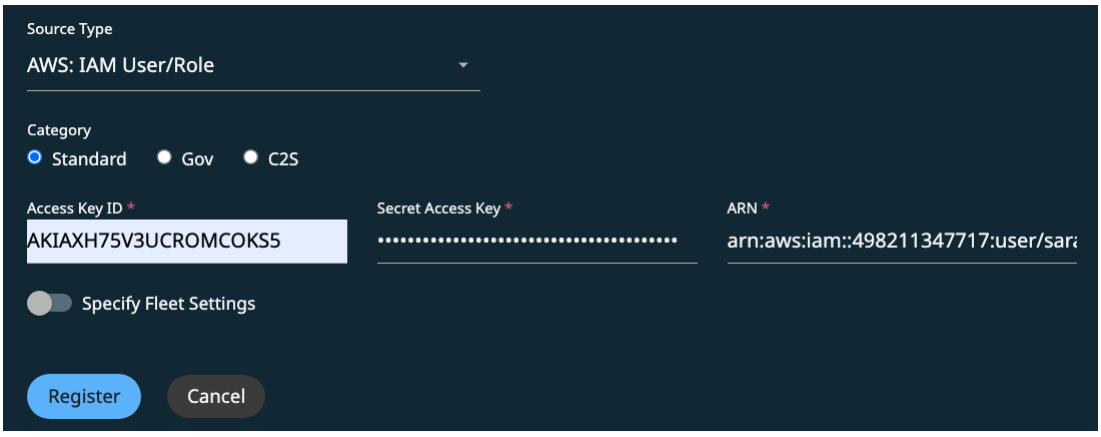
Appendix D: Register AWS Source with Cohesity Using IAM Role

If you have a Cloud Edition, you can register the EC2 instances using the IAM role.

Register Cloud Source using IAM Role

1. Select **Data Protection > Sources.**
2. Select **Register** > **Virtual Machines.**
3. From the **Select Hypervisor Source Type drop-down,** select **AWS: IAM User/Role.**
4. Choose **Category** as **Standard, Gov** or **C2S.**
5. Enter the Access Key ID, Secret Access key, and ARN (go to **IAM** > **Users** > and select the user and copy the **ARN**)





6. Click **Register** to complete source registration.

Your Feedback

Was this document helpful? Send us your feedback!

About the Author

Saran Ravi is a Staff Technical Solutions Engineer at Cohesity. In his role, he focuses on Cloud and Cohesity Cloud Services.

Other essential contributors included:

- Himanshu Srivastava, Engineering
- Saurabh Singh, Product Management
- Dinesh Pathak, Engineering
- Anirudh Kumar, Engineering
- James White, Principal Solutions Engineer
- Kevin Hill, Senior Solution Architect

Document Version History

VERSION	DATE	DOCUMENT HISTORY
1.6	Apr 2024	Minor updates
1.5	Jan 2024	Minor updates
1.4	June 2023	Minor updates
1.3	Mar 2023	Rebranding updates
1.2	Nov 2022	Updated to 6.8
1.1	Oct 2021	Rebranding updates
1.0	Sep 2019	Original document

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.