

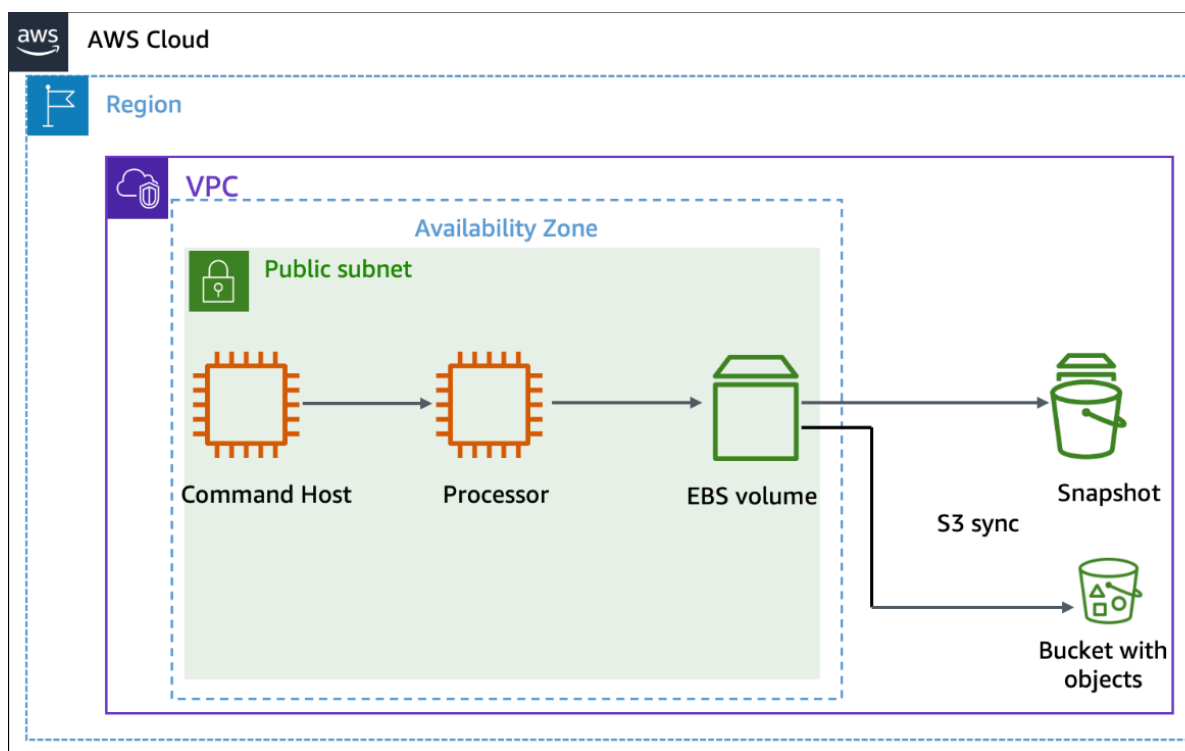
183-[JAWS]-Lab - Managing Storage

Managing Storage

Lab overview

AWS provides multiple ways to manage data on Amazon Elastic Block Store (Amazon EBS) volumes. In this lab, you use AWS Command Line Interface (AWS CLI) to create snapshots of an EBS volume and configure a scheduler to run Python scripts to delete older snapshots.

In the challenge section of this lab, you are challenged to sync the contents of a directory on an EBS volume to an Amazon Simple Storage Service (Amazon S3) bucket using an Amazon S3 sync command.



Your lab environment consists of a virtual private cloud (VPC) with a public subnet. Amazon Elastic Compute Cloud (Amazon EC2) instances named "Command Host" and "Processor" have already been created in this VPC for you as part of this lab.

The "Command Host" instance will be used to administer AWS resources including the "Processor" instance.

Objectives

By the end of this lab, you will be able to do the following:

- Create and maintain snapshots for Amazon EC2 instances.
- Use Amazon S3 sync to copy files from an EBS volume to an S3 bucket.
- Use Amazon S3 versioning to retrieve deleted files.

Duration

This lab will require approximately **45 minutes** to complete.

Accessing the AWS Management Console

1. At the top of these instructions, choose **Start Lab** to launch your lab.

A Start Lab panel opens displaying the lab status.

2. Wait until you see the message "**Lab status: ready**", then choose **X** to close the Start Lab panel.

3. At the top of these instructions, choose **AWS**.

This opens the AWS Management Console in a new browser tab. The system automatically logs you in.

Tip: If a new browser tab doesn't open, there's typically a banner or icon at the top of your browser indicating that your browser is preventing the site from opening pop-up windows. Choose the banner or icon and choose "Allow pop ups."

4. Arrange the AWS Management Console tab so that it displays along side these instructions. Ideally, you will be able to see both browser tabs at the same time to make it easier to follow the lab steps.

Task 1: Creating and configuring resources

In this task, you create an Amazon S3 bucket and configure the "Command Host" EC2 instance to have secure access to other AWS resources.

Task 1.1: Create an S3 bucket

In this task, you create an S3 bucket to sync files from an EBS volume.

5. On the **AWS Management Console**, in the **Search** bar, enter and choose **S3** to open the **S3 Management Console**.
6. On the console, choose **Create bucket**.
7. In the **Create bucket** section, configure the following:
 - **Bucket name:** Enter a bucket name. Use a combination of characters and numbers to keep it unique.
This will be referred to as "s3-bucket-name" throughout the lab.
 - **Region:** Leave as default.
8. Scroll and choose **Create bucket**.

Task 1.2: Attach instance profile to Processor

In this task, you attach a pre-created IAM role as an instance profile to the EC2 instance "Processor," giving it the permissions to interact with other AWS services such as EBS volumes and S3 buckets.

9. On the **AWS Management Console**, in the **Search** bar, enter and choose **EC2** to open the **EC2 Management Console**.
10. In the navigation pane, choose **Instances**.
11. Choose **Processor** from the list of EC2 instances.
12. Choose **Actions > Security > Modify IAM role**.
13. Choose the **S3BucketAccess** role in the **IAM role** dropdown list.
14. Choose **Update IAM role**.

Task 2: Taking snapshots of your instance

In this section, you use the AWS Command Line Interface (AWS CLI) to manage the processing of snapshots of an instance.

Task 2.1: Connecting to the Command Host EC2 instance

In this task, you use EC2 Instance Connect to connect to the "Command Host" EC2 instance.

15. On the **AWS Management Console**, in the **Search** bar, enter and choose **EC2** to open the **EC2 Management Console**.
16. In the navigation pane, choose **Instances**.
17. From the list of instances, choose **Command Host**.
18. Choose **Connect**.
19. On the **EC2 Instance Connect** tab, choose **Connect**.

This option opens a new browser tab with the **EC2 Instance Connect** terminal window.

Note: If you prefer to use an SSH client to connect to the EC2 instance, see the guidance provided in the additional references section.

You use this terminal window to complete the tasks throughout the lab. If the terminal becomes unresponsive, refresh the browser or use the steps in this task to connect again.

Task 2.2: Taking an initial snapshot

In this task, you identify the EBS volume that's attached to the "Processor" instance and take an initial snapshot. To do so, you run commands in the **EC2 Instance Connect terminal window**. You can copy the command output to a text editor for subsequent use.

20. To display the EBS volume-id, run the following command:

```
aws ec2 describe-instances --filter 'Name=tag:Name,Values=Processor' --query  
'Reservations[0].Instances[0].BlockDeviceMappings[0].Ebs.{VolumeId:VolumeId}'
```

Note: The command returns a response similar to this: "Volumeld": "vol-1234abcd".

You use this value for Volumeld throughout the lab steps when prompted.

21. Next, you take snapshot of this volume. Prior to this, you shut down the "Processor" instance, which requires its instance ID. Run the following command to obtain the instance ID:

```
aws ec2 describe-instances --filters 'Name=tag:Name,Values=Processor' --query  
'Reservations[0].Instances[0].InstanceId'
```

The command returns a value for INSTANCE-ID similar to this: "i-0b06965263c7ac08f"

The command returns a value for INSTANCE-ID similar to this: "i-0b06965263c7ac08f"

22. To shut down the "Processor" instance, run the following command and replace "INSTANCE-ID" with the instance-id that you retrieved earlier:

```
aws ec2 stop-instances --instance-ids INSTANCE-ID
```

23. To verify that the "Processor" instance stopped, run the following command and replace "INSTANCE-ID" with your instance id.

```
aws ec2 wait instance-stopped --instance-id INSTANCE-ID
```

When the instance stops, the command returns to a prompt.

24. To create your first snapshot of the volume of your "Processor" instance, run the following command and replace "VOLUME-ID" with the Volumeld that you retrieved earlier:

```
aws ec2 create-snapshot --volume-id VOLUME-ID
```

The command returns a set of information that includes a SnapshotId similar to "snap-0643809e73e6cce13"

25. To check the status of your snapshot, run the following command and replace "SNAPSHOT-ID" with the SnapshotId that you retrieved earlier:

```
aws ec2 wait snapshot-completed --snapshot-id SNAPSHOT-ID
```

Continue to next steps after the command returns to the prompt.

26. To restart the "Processor" instance, run the following command and replace "INSTANCE-ID" with the instance-id that you retrieved earlier:

```
aws ec2 start-instances --instance-ids INSTANCE-ID
```

After a couple minutes, the instance will be in the running state.

Task 2.3 Scheduling the creation of subsequent snapshots

Using the Linux scheduling system (cron), you can set up a recurring snapshot process so that new snapshots of your data are taken automatically.

For the purposes of this lab, you schedule a snapshot creation every minute so that you can verify the results of your work.

In this task, you create a cron job to manage the number of snapshots that are maintained for a volume.

Note: This section of the lab doesn't stop the instance in order to create a large number of snapshots for the next step.

27. To create and schedule a cron entry that runs a job every minute, run the following command and replace "VOLUME-ID" with the Volumeld that you retrieved earlier:

```
echo "* * * * * aws ec2 create-snapshot --volume-id VOLUME-ID 2>&1 >> /tmp/cronlog" > cronjob
crontab cronjob
```

Note: This will take 1-2 minutes.

28. To verify that subsequent snapshots are being created, run the following command and replace "VOLUME-ID" with the Volumeld that you retrieved earlier:

```
aws ec2 describe-snapshots --filters "Name=volume-id,Values=VOLUME-ID"
```

Re-run the command after few minutes to see more snapshots.

29. Wait a few minutes so that a few more snapshots are generated before beginning the next task.

Task 2.4 Retaining the last two snapshots

In this task, you run a Python script that maintains only the last two snapshots for any given EBS volume.

30. To stop the cron job, run the following command:

```
crontab -r
```

31. To examine the contents of the Python script "snapshotter_v2.py", run the following command:

```
more /home/ec2-user/snapshotter_v2.py
```

The script finds all EBS volumes that are associated with the current user's account and takes snapshots. It then examines the number of snapshots that are associated with the volume, sorts the snapshots by date, and removes all but the two most recent snapshots.

32. Before running snapshotter_v2.py, run the following command and replace "VOLUME-ID" with the Volumeld that you retrieved earlier:

```
aws ec2 describe-snapshots --filters "Name=volume-id,Values=VOLUME-ID" --query 'Snapshots[*].SnapshotId'
```

The command returns the multiple snapshot IDs that were returned for the volume. These are the snapshots that were created by your cron job before you stopped it.

33. Run the the "snapshotter_v2.py" script using following command:

```
python3.8 snapshotter_v2.py
```

The script runs for a few seconds, and then it returns a list of all of the snapshots that it deleted:

```
[ec2-user@ip-]*]$ python3.8 snapshotter_v2.py
```

```
Deleting snapshot snap-e8128a20
```

```
Deleting snapshot snap-d0d34818
```

```
Deleting snapshot snap-ded14a16
```

```
Deleting snapshot snap-e8d74c20
```

```
Deleting snapshot snap-25d54eed
```

```
Deleting snapshot snap-4acb5082
```

34. To examine the new number of snapshots for the current volume, re-run the following command from an earlier step:

```
aws ec2 describe-snapshots --filters "Name=volume-id, Values=VOLUME-ID" --query 'Snapshots[*].SnapshotId'
```

The command returns only **two** snapshot IDs.

Task 3: Challenge: Synchronize files with Amazon S3

In this task, you are challenged to sync the contents of a directory with the Amazon S3 bucket that you created earlier.

Note: It's recommended that you try this challenge by yourself using the information that's provided in this section **before** reading the detailed solution. When you complete the challenge, check your work by reviewing the detailed solution.

Challenge Description

Run the following command in the terminal to download a sample set of files:

```
wget https://aws-tc-largeobjects.s3.us-west-2.amazonaws.com/CUR-TF-100-RSJAWS-3-124627/183-1ab-JAWS-managing-storage/s3/files.zip
```

Unzip these files, and then, using the AWS CLI, figure out how to accomplish the following:

- Activate versioning for your Amazon S3 bucket.
- Use a single AWS CLI command to sync the contents of your unzipped folder with your Amazon S3 bucket.
- Modify the command so that it deletes a file from Amazon S3 when the corresponding file is deleted locally on your instance.
- Recover the deleted file from Amazon S3 using versioning.

You can use the solution summary as a guide to complete the challenge yourself. Use the links in additional references section for details on using required AWS CLI commands.

Solution Summary

The solution involves the following steps:

- To activate versioning for the bucket, use the `aws s3api put-bucket-versioning` command.
- To synchronize the local files with Amazon S3, use the `aws s3 sync` command on the local folder.
- Delete a local file.
- To force Amazon S3 to delete any files that aren't present on the local drive but present in Amazon S3, use the `--delete` option with the `aws s3 sync` command.
- Because there's no direct command in Amazon S3 to restore a previous version of a file, to download a previous version of the deleted file from Amazon S3, use the `aws s3api list-object-versions` and `aws s3api get-object` commands. You can then restore the file to Amazon S3 by using `aws s3 sync`.

Task 3.1: Downloading and unzipping sample files

The sample file package contains a folder with three text files: file1.txt, file2.txt, and file3.txt. These are the files that you will sync with your Amazon S3 bucket.

35. Connect to the "Processor" instance using EC2 Instance Connect.

Note: Refer to the earlier steps that you used to connect to the "Command Host" instance.

You run following AWS CLI commands in the EC2 Instance Connect terminal window.

36. To download the sample files on the "Processor" instance, run the following command from within your instance:

```
wget https://aws-tc-largeobjects.s3.us-west-2.amazonaws.com/CUR-TF-100-RSJAWS-3-124627/183-lab-JAWS-managing-storage/s3/files.zip
```

37. To unzip the directory, use the following command:

```
unzip files.zip
```

Task 3.1: Downloading and unzipping sample files

The sample file package contains a folder with three text files: file1.txt, file2.txt, and file3.txt. These are the files that you will sync with your Amazon S3 bucket.

35. Connect to the "Processor" instance using EC2 Instance Connect.

Note: Refer to the earlier steps that you used to connect to the "Command Host" instance.

You run following AWS CLI commands in the EC2 Instance Connect terminal window.

36. To download the sample files on the "Processor" instance, run the following command from within your instance:

```
wget https://aws-tc-largeobjects.s3.us-west-2.amazonaws.com/CUR-TF-100-RSJAWS-3-124627/183-lab-JAWS-managing-storage/s3/files.zip
```

37. To unzip the directory, use the following command:

```
unzip files.zip
```

Task 3.2: Syncing files

38. Before syncing content with your Amazon S3 bucket, you need to activate versioning on your bucket.

Run the following command and replace "S3-BUCKET-NAME" with your bucket name:

```
aws s3api put-bucket-versioning --bucket S3-BUCKET-NAME --versioning-configuration Status=Enabled
```

39. To sync the contents of the files folder with your Amazon S3 bucket, run the following command and replace "S3-BUCKET-NAME" with your bucket name:

```
aws s3 sync files s3://S3-BUCKET-NAME/files/
```

The command confirms that three files were uploaded to your S3 bucket.

40. To confirm the state of your files, run the following command and replace "S3-BUCKET-NAME" with your bucket name:

```
aws s3 ls s3://S3-BUCKET-NAME/files/
```

41. To delete one of the files on the local drive, run the following command:

```
rm files/file1.txt
```

42. To delete the same file from the S3 bucket, use the `--delete` option with the `aws s3 sync` command.

Run the following command and replace "S3-BUCKET-NAME" with your bucket name:

```
aws s3 sync files s3://S3-BUCKET-NAME/files/ --delete
```

The command returns a message similar to following:

```
delete: s3://examplebucketname/files/file1.txt
```

43. To verify that the file was deleted from the bucket, run the following command and replace "S3-BUCKET-NAME" with your bucket name:

```
aws s3 ls s3://S3-BUCKET-NAME/files/
```

44. Now, try to recover the old version of file1.txt. To view a list of previous versions of this file, run the following command and replace "S3-BUCKET-NAME" with your bucket name:

```
aws s3api list-object-versions --bucket S3-BUCKET-NAME --prefix files/file1.txt
```

The output contains a `DeleteMarkers` and a `Versions` block. `DeleteMarkers` indicates where the delete marker is. For example, if you perform an `aws s3 rm` operation (or an `aws s3 sync` operation with the `--delete` option), this is the next version that the file will revert to.

The `Versions` block contains a list of all available versions. You should have only a single versions entry. Note the value for `VersionId` for use later.

45. Because there's no direct command to restore an older version of an Amazon S3 object to its own bucket, you need to re-download the old version and sync again to Amazon S3.

To download the previous version of file1.txt, run the following command and replace "S3-BUCKET-NAME" with your bucket name:

```
aws s3api get-object --bucket S3-BUCKET-NAME --key files/file1.txt --version-id VERSION-ID
files/file1.txt
```

46. To verify that the file was restored locally, run the following command:

```
ls files
```

The command shows all three files listed.

47. To re-sync the contents of the files/ folder to Amazon S3, run the following command from within your instance and replace "S3-BUCKET-NAME" with your bucket name:

```
aws s3 sync files s3://S3-BUCKET-NAME/files/
```

48. Finally, to verify that a new version of file1.txt was pushed to the S3 bucket, run the following command and replace "S3-BUCKET-NAME" with your bucket name:

```
aws s3 ls s3://S3-BUCKET-NAME/files/
```

Conclusion

Congratulations! You successfully completed the following:

- Created and maintain snapshots for Amazon EC2 instances.
- Used Amazon S3 sync to copy files from an EBS volume to an S3 bucket.
- Used Amazon S3 versioning to retrieve deleted files.

Lab Complete

Congratulations! You completed the lab.

49. Choose **End Lab** at the top of this page, and then choose **Yes** to confirm that you want to end the lab.

A panel appears, indicating that "DELETE has been initiated... You may close this message box now."

50. Choose **X** in the top right corner to close the panel.

- - -