Caltech | Center for Technology & Management Education

Post Graduate Program in Cloud

simplilearn

Cloud

Caltech | Center for Technology & Management Education

AWS Certified SysOps Administrator – Associate Level

simplilearn

Data Management

Learning Objectives

By the end of this lesson, you will be able to:

- Check S3 versioning on AWS Console
- Work with default encryption and bucket policies
- Upgrade and change an EC2 volume
- Create a query in Athena to perform operations on a specific bucket in S3

A Day in the Life of an AWS Administrator

You work as an architect for an e-commerce company that primarily deals with data. Your company is seeking cloud services to assist them in better managing their data while maintaining its security and availability. You've been asked to assist them with a few AWS services which will help them in meeting the following requirements:

- The company wants a database that can provide industry-leading scalability, data availability, security, and performance. They'd also be able to store and retrieve any amount of data from any location, at any time.
- They want a feature that will allow them to preserve, retrieve, and restore every version of every object stored in the database system.
- They also want to make certain that data is encrypted.

A Day in the Life of an AWS Administrator

- Additionally, they should also have a solution in place to evaluate data using standard SQL.
- The company would also like to find a database solution that allows them to upgrade the volume type to improve or decrease the database's performance depending on their needs.

To achieve all the above along with some additional features, you will be learning a few concepts in this lesson that will help you find solutions for the above-given scenario.

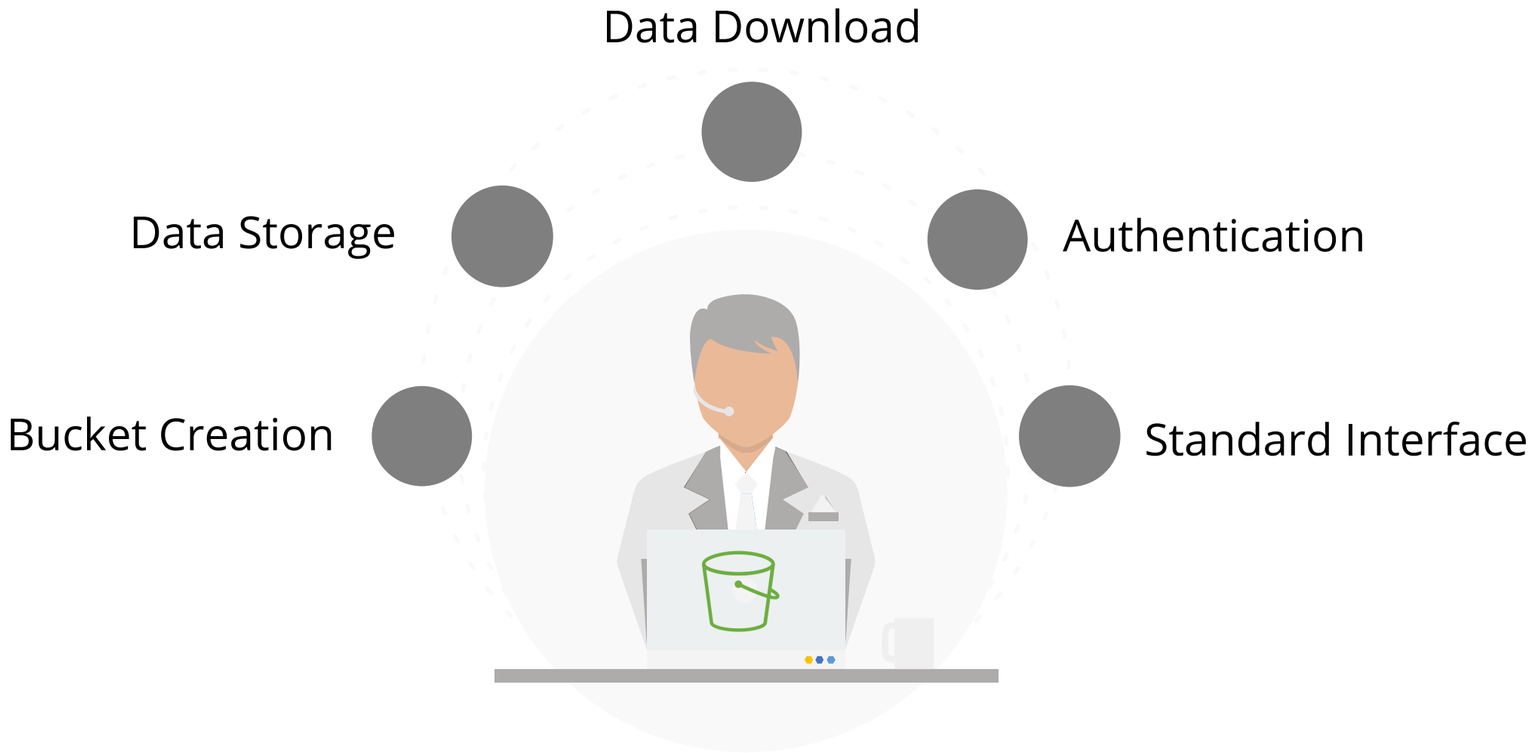
Introduction to S3

What Is S3?

S3 is defined as the storage to make web-scale computing easier for developers.

It has a web service interface that can be used to store, get data from anywhere and anytime on the web.

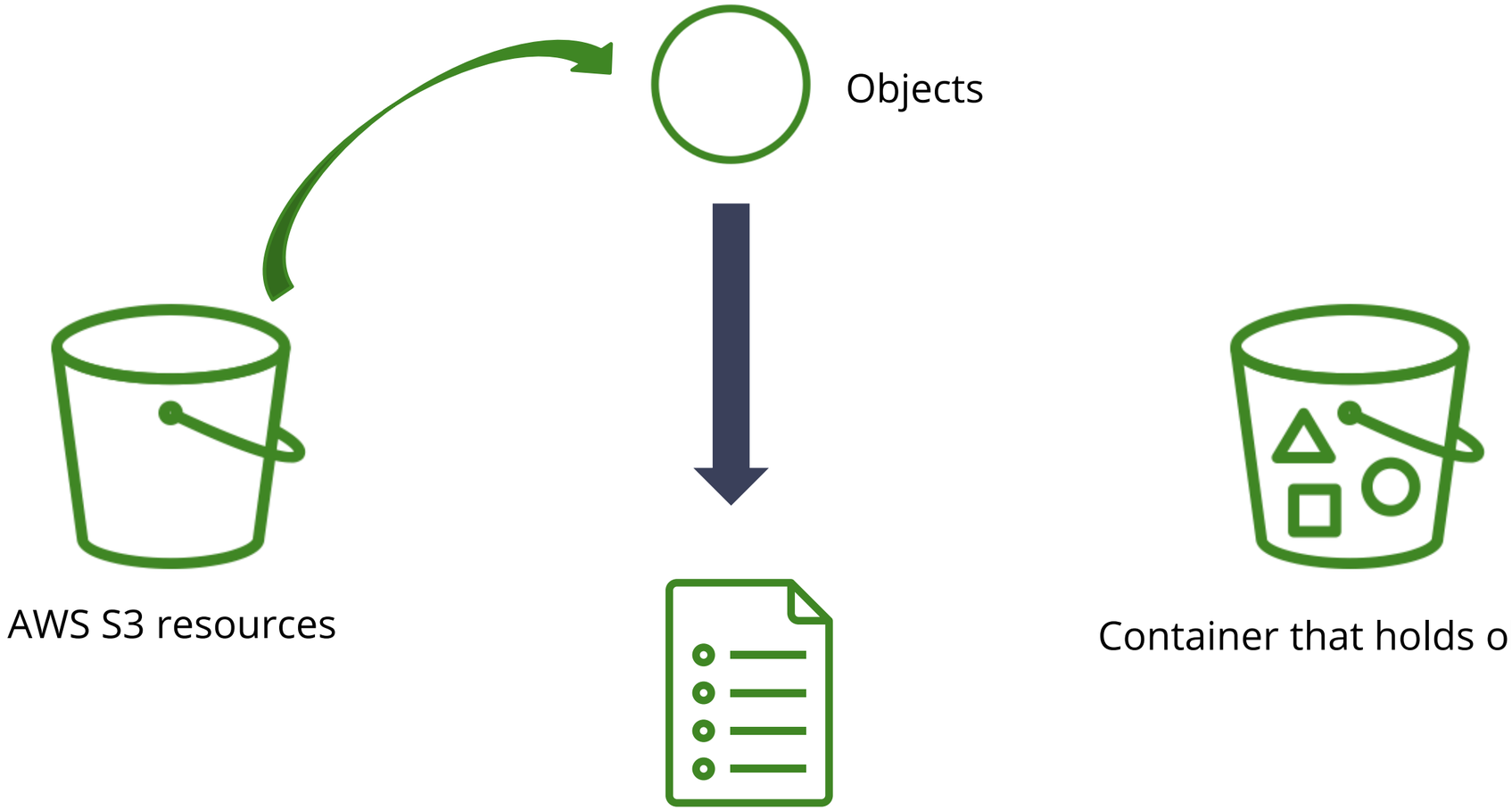
S3: Advantages



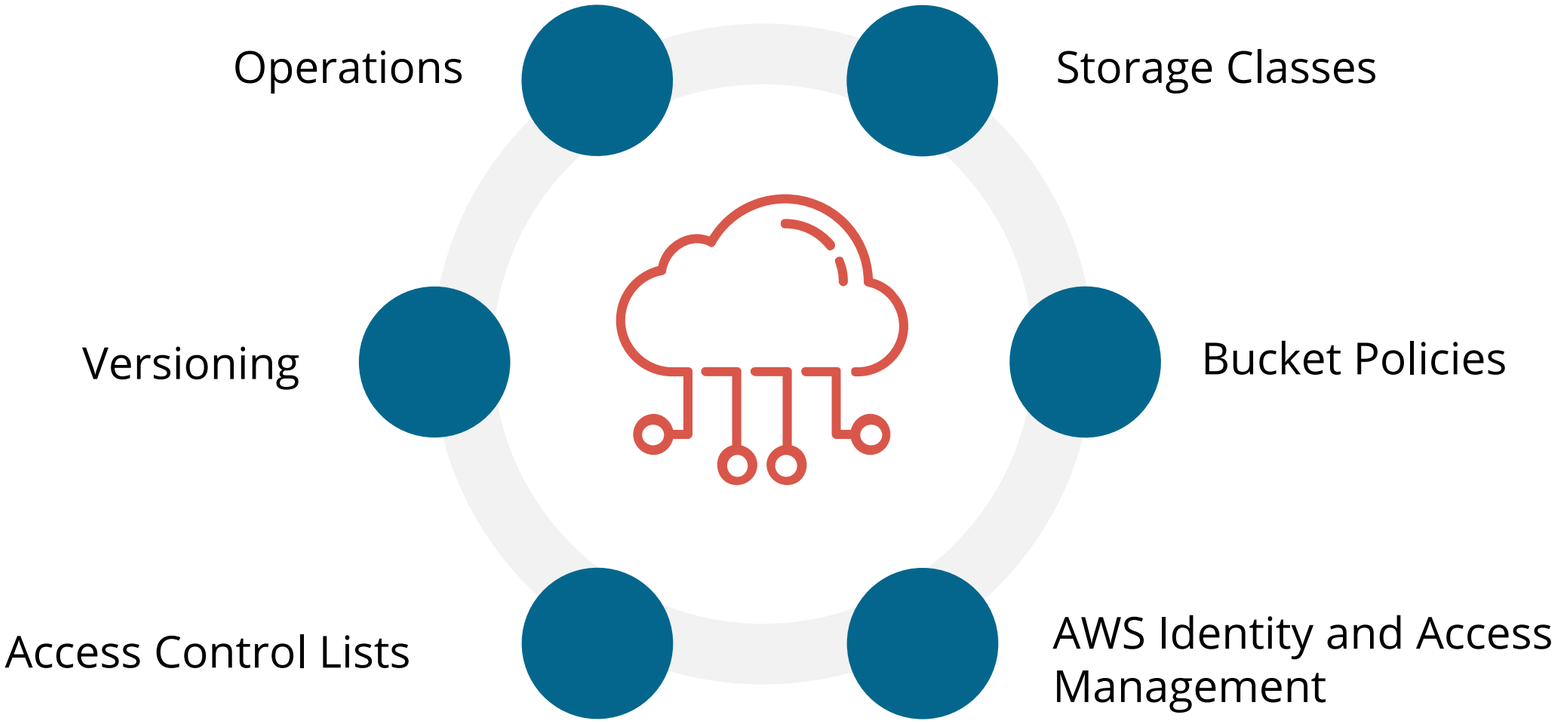
Concepts of S3

Buckets	Objects	Keys	Regions	Amazon S3 Data Consistency Model
They are defined as containers for objects stored in Amazon S3.	They are entities stored in S3 and consist of object data and metadata.	They are unique identifiers for objects that are stored in a bucket. Every object has a unique key.	They are geographical AWS regions where the buckets created in S3 are stored.	It provides the read-after-write consistency for the PUTS of new objects in the S3 bucket in all regions with a warning.

Concepts of S3



Features of S3



Assisted Practice

Creating a S3 Bucket

Duration: 10 Min.

Problem Statement:

You are given a project to create an S3 bucket which is an object storage service that offers industry-leading scalability, data availability, security, and performance. Amazon S3 allows a user to store and retrieve any amount of data at any time, from anywhere.

Assisted Practice: Guidelines

Steps to create a S3 Bucket:

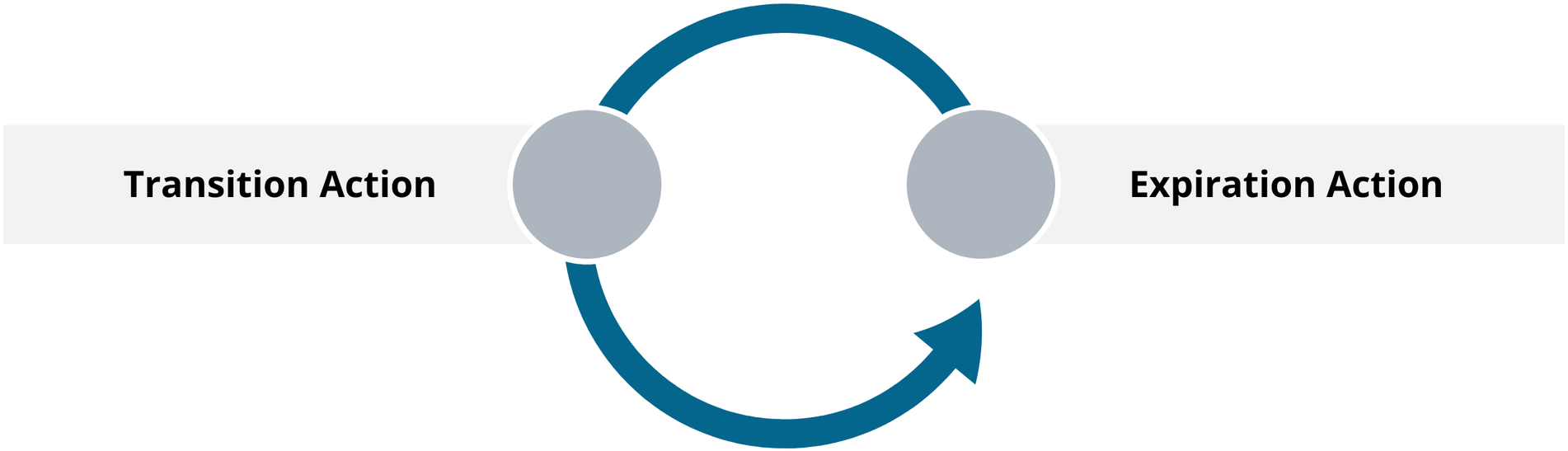
1. Log in to your AWS lab
2. Click on **Amazon S3** and create a bucket with all the relevant detail.

S3 Lifecycle

S3 Lifecycle Policies

Amazon S3 Lifecycle policies are configured to manage the objects to be stored effectively throughout the lifecycle.

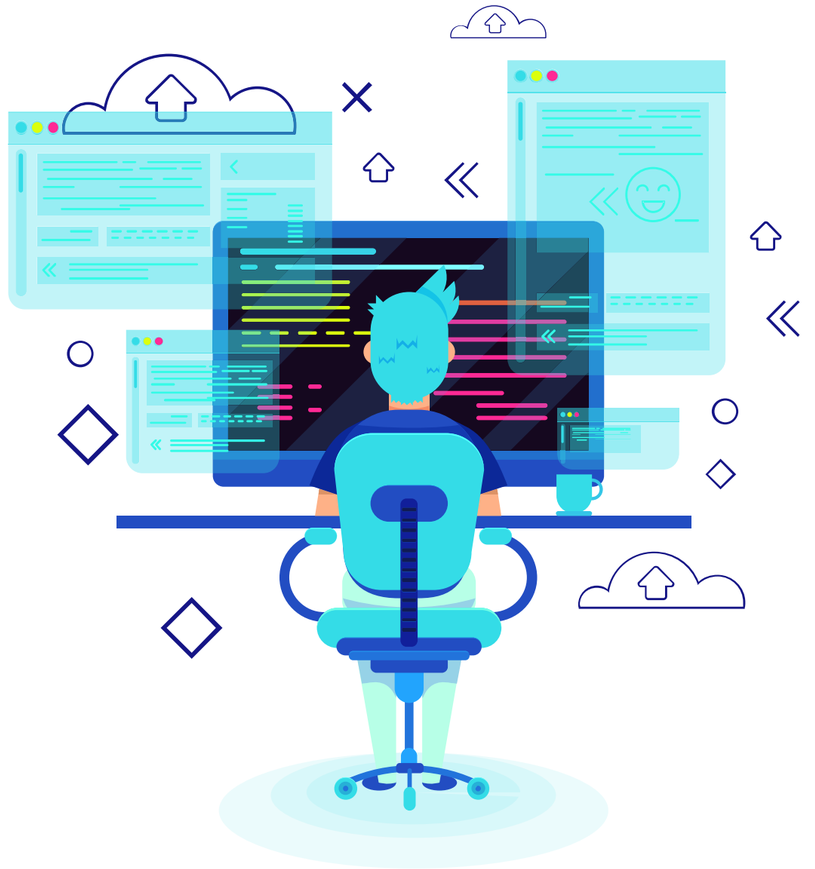
Below are the two types of actions:



S3 Lifecycle Policies: Example

Below is an example of lifecycle configuration to abort multipart uploads API that is used to upload large objects in parts.

```
<LifecycleConfiguration>
   <Rule>
      <ID>sample-rule</ID>
      <Filter>
        <Prefix>SomeKeyPrefix/</Prefix>
      </Filter>
      <Status>rule-status</Status>
      <AbortIncompleteMultipartUpload>
        <DaysAfterInitiation>7</DaysAfterInitiation>
      </AbortIncompleteMultipartUpload>
   </Rule>
</LifecycleConfiguration>
```

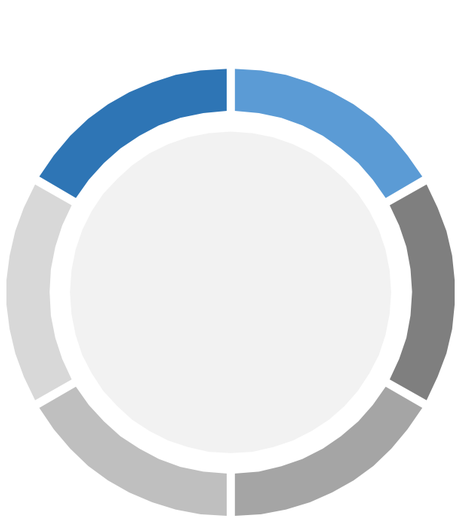


S3 Versioning and Cross-Region Replication

S3 Versioning

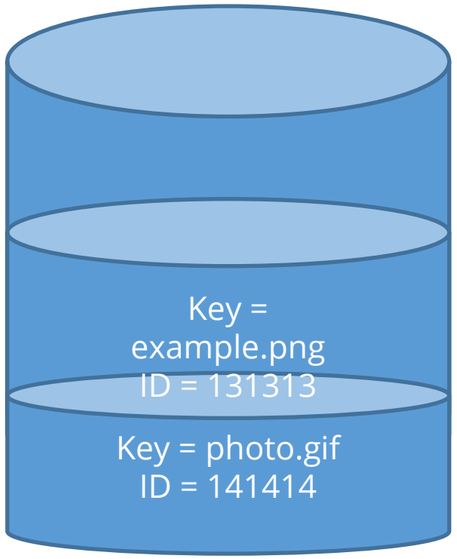
Versioning refers to storing multiple variants of an object in the same bucket. Below are the benefits of versioning:

- Enables us to recover objects from accidental deletion or overwrite
- Used to preserve and restore every version of each object stored



- Helps Amazon S3 to automatically generate the unique version ID for the object to be stored
- Used to recover from application failures and unintended user actions

S3 Versioning: Example



We can have two objects in a bucket with the same key but different version IDs, such as example.png (version 131313) and photo.gif (version 141414).

Assisted Practice

S3 Versioning

Duration: 10 Min.

Problem Statement:

You are given a project to demonstrate S3 versioning in Amazon S3 which is a means of keeping multiple variants of an object in the same bucket. You can use the S3 Versioning feature to preserve, retrieve, and restore every version of every object stored in your buckets.

Assisted Practice: Guidelines

Steps to check S3 versioning:

1. Log in to your AWS lab
2. Go to the Amazon dashboard and choose Amazon S3
3. Create a bucket
4. Configure the S3 bucket to enable versioning
5. Upload the files, verify the versioning, and reupload the bucket to check consistency

MFA Delete

MFA Delete

Multi-Factor Authentication (MFA) is used to add another layer of security by configuring the bucket.

MFA requires the authentication for the below operations:

Changing the versioning state of the bucket

Permanently deleting the version of an object

MFA Delete

MFA Delete requires the below two forms of authentications together:

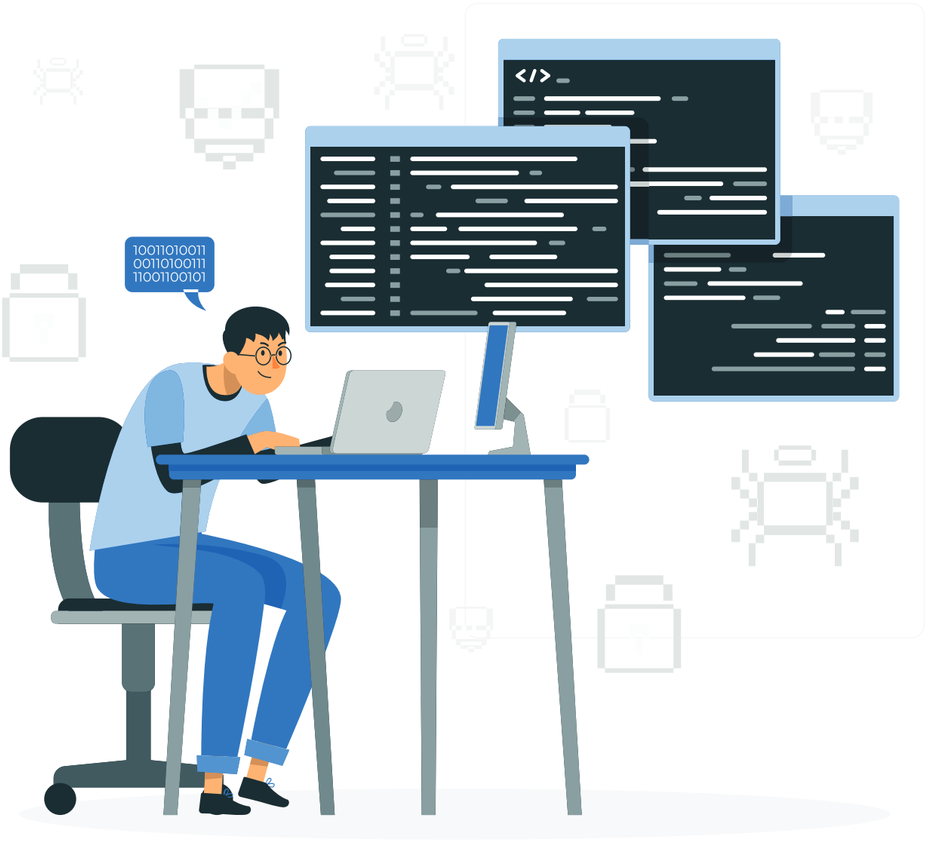
- User security credentials
- Concatenation of a serial number, space, and the six-digit code displayed on an approved authentication device

MFA Delete

MFA Delete can help prevent accidental bucket deletions by doing the following:

- Adding a user to initiate the delete action to prove the physical possession of an MFA device
- Adding an extra layer of security and friction to the delete action

MFA Delete




```
<VersioningConfiguration
xmlns="http://s3.amazonaws.com/doc/2006-03-01/">
  <Status>VersioningState</Status>
  <MfaDelete>MfaDeleteState</MfaDelete>
</VersioningConfiguration>
```

Amazon S3 Replication

Replication allows objects to be copied asynchronously between Amazon S3 buckets.

- Object replication buckets can be owned by the same AWS account or by distinct AWS accounts.
- Objects can be copied to either a single or multiple destination buckets.
- The destination buckets might be in different AWS Regions or the same as the source bucket.
- After replication is enabled, it only allows a user to copy new Amazon S3 items by default.
- A user can use replication to copy existing objects by contacting AWS support.

Why Use S3 Replication?

Replication can help a user do the following:

- Replicate objects while retaining metadata
- Replicate objects into different storage classes
- Maintain object copies under different ownership
- Keep objects stored over multiple AWS Regions
- Replicate objects within 15 minutes

Cross-Region Replication

S3 Cross-Region Replication (CRR) is a feature of Amazon S3 that allows a user to copy objects between buckets in different AWS Regions.

CRR can assist a user in completing the following tasks:

- Meet compliance requirements
- Minimize latency
- Increase operational efficiency

Same-Region Replication

S3 Same-Region Replication (SRR) is a feature of Amazon S3 that allows a user to copy objects between buckets in the same AWS Regions.

SRR can assist a user in completing the following tasks:

- Aggregate logs into a single bucket
- Configure live replication between production and test accounts
- Abide by data sovereignty laws

S3 Encryption

S3 Encryption

Amazon S3 default encryption is used to set the default encryption behavior for an S3 bucket. This is done so that all the new objects are encrypted when they are stored in the bucket.

The objects are encrypted using server-side encryption with either Amazon S3-managed keys (SSE-S3) or customer master keys (CMKs) stored in AWS Key Management Service (AWS KMS).

S3 Encryption

When you use server-side encryption, Amazon S3 encrypts an object before saving it to the disk and decrypts it when you download the objects.

Assisted Practice

Enable S3 Encryption

Duration: 10 Min.

Problem Statement:

You are given a project to demonstrate encryption on an S3 bucket so that all objects are encrypted when they are stored in the bucket.

Assisted Practice: Guidelines

Steps to work with default encryption:

1. Log in to your AWS lab
2. Click on **Amazon S3**, and create a bucket with all the relevant details
3. Click on the **Automatically encrypt objects when stored in S3** checkbox
4. Verify the working of default encryption
5. Upload the file and check the encryption

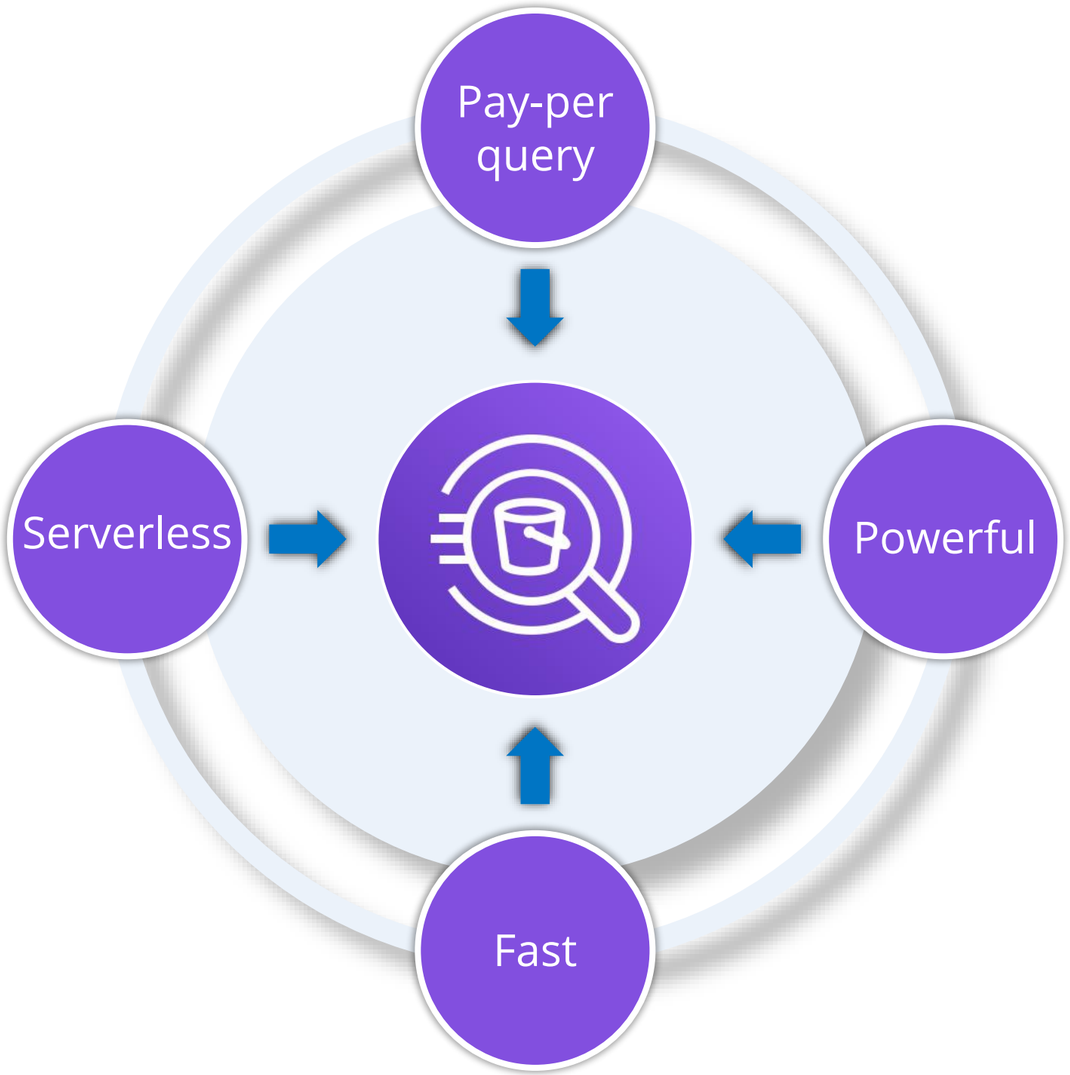
Introduction to Athena

What is Athena?

It is an interactive query service that is used to analyze data directly in Amazon S3 using the standard SQL.

Athena can be pointed to the data stored in S3 using a few actions from the AWS Management Console, and standard SQL can be used to run ad-hoc queries to get results in seconds.

Benefits of Athena



Features of Athena

01. Serverless – Zero Administration

02. Easy to Get Started

03. Uses Standard SQL

04. Pay Per Query

05. Fast Performance

06. Highly Available and Durable

07. Integrated and Secured

08. Machine Learning

Assisted Practice

Athena

Duration: 10 Min.

Problem Statement:

You have been given a project to demonstrate Athena which is an interactive query service that makes it easy to analyze data in Amazon S3 using standard SQL.

Assisted Practice: Guidelines

Steps to create a query in Athena:

1. Log in to your AWS lab
2. Create two S3 buckets
3. Add logging to the bucket
4. Create and execute Athena queries

EC2 Volume Types

Volume Types

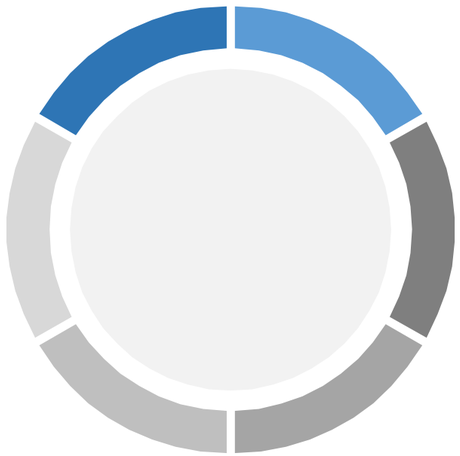
Characteristics				
Volume Types	General Purpose SSD (gp2)	Provisioned IOPS SSD (io2)	Throughput Optimized HDD (st1)	Cold HDD (sc1)
Description	It balances price and performance for workloads	It is meant for mission-critical, low-latency, or high-throughput workloads	It is a low-cost HDD volume that is designed for frequently accessed workloads	It is the lowest cost HDD volume that is designed for less frequently accessed workloads
Durability	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)	99.999% durability (0.001% annual failure rate)	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)
Volume size	1 GiB - 16 TiB	4 GiB - 16 TiB	500 GiB - 16 TiB	500 GiB - 16 TiB
Max IOPS per volume	16,000 (16 KiB I/O) *	64,000 (16 KiB I/O)	500 (1 MiB I/O)	250 (1 MiB I/O)

Terminating an Instance

When an instance is terminated:

The data on instance store volumes linked with that instance is deleted.

It automatically gets deleted from the console after a short while.

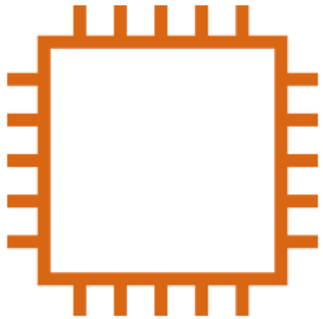


Amazon EBS root device volumes are automatically deleted.

All the resources get disassociated from the instance.

Amazon EC2 Instance Store

An instance store is used to provide temporary block-level storage for the instance that is located on disks attached to the computer.

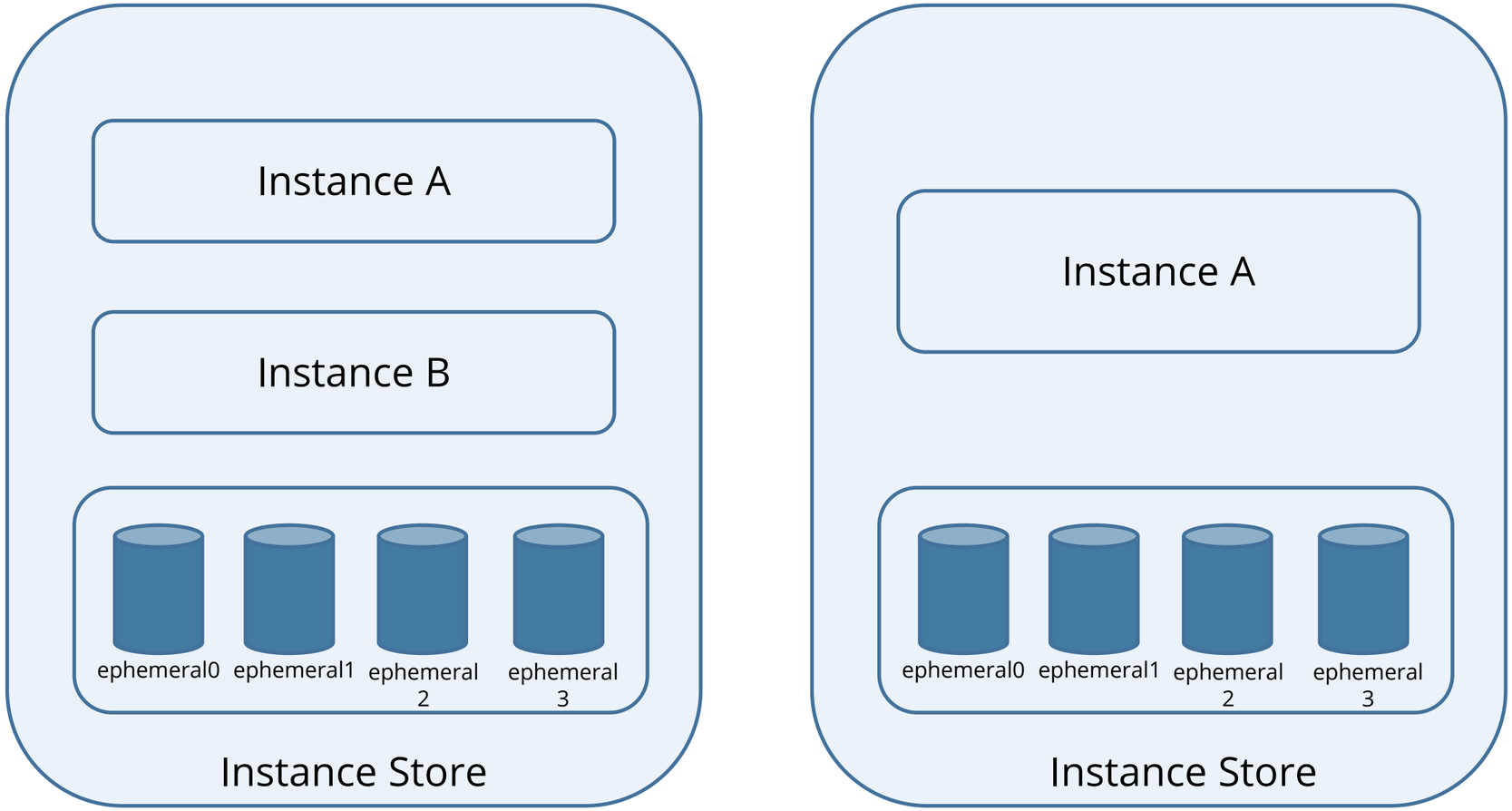


01 Ideal for the temporary storage of information that changes frequently

02 Consists of one or more instance store volumes exposed as block devices

03 Size increases with the increase in the number of devices available

Amazon EC2 Instance Store



Assisted Practice

Upgrading EC2 Volume and Changing Volume Types

Duration: 10 Min.

Problem Statement:

You have been given a project to demonstrate how to upgrade the EC2 Volume type to increase or decrease the performance of the database based upon the requirement.

Assisted Practice: Guidelines

Steps to upgrade volume of an EC2 instance:

1. Log in to your AWS lab
2. Go to the Amazon dashboard, and click on EC2 instance
3. Choose the AMI machine type, and select **t2.micro** as the instance type
4. Add storage, and configure **Security Group**
5. Launch instance, and then configure EBS storage and storage space
6. Create a volume and attach the instance

Key Takeaways

- Amazon S3 lifecycle policies are configured to store objects effectively throughout the lifecycle.
- Versioning refers to storing multiple variants of an object in the same bucket.
- Replication allows objects to be copied asynchronously between Amazon S3 buckets.
- Athena can be pointed to the data using a few actions from the AWS Management Console, and standard SQL can be used to run ad-hoc queries to get results in seconds.

Lesson-End Project

Duration: 60 min.

Run Batch Operations to Encrypt Data in an S3 Bucket

Problem Statement:

Use S3 Batch Operations to encrypt the existing data in the existing S3 bucket, and use Athena to check the bytes of data uploaded and downloaded from the monitored bucket.

Background of the problem statement:

Your company has got a contract for storage and data management from a client. As you are a senior SysOps engineer, this task has been assigned to you.