

## Chapter 4: Deep Dive into Storage Services

### Objective:

Equip students with the skills to create and manage AWS storage solutions effectively, including S3 buckets, storage classes, lifecycle policies, EBS volumes, snapshots, and EFS for shared storage.

### 1. Creating and Managing S3 Buckets

What is Amazon S3?

Amazon Simple Storage Service (S3) is a highly scalable, durable, and secure object storage service.

Stores data as objects inside buckets (containers).

Objects include files like images, videos, documents, backups, and more.

Creating an S3 Bucket

Open AWS Management Console → S3 service.

Click Create bucket.

Enter a unique bucket name (globally unique across AWS).

Select the AWS Region closest to your users for low latency.

Configure options like versioning, logging, and tags (optional).

Set permissions carefully to control who can access the bucket.

Click Create bucket.

Managing Buckets and Objects

Upload objects via console, CLI, SDKs, or APIs.

Set permissions with Access Control Lists (ACLs) or bucket policies.

Enable versioning to keep multiple versions of objects (useful for recovery).

Enable server-side encryption to protect data at rest.

Configure event notifications to trigger Lambda functions or workflows when objects are added or deleted.

### 2. Understanding S3 Storage Classes and Lifecycle Policies

S3 Storage Classes

AWS offers multiple storage classes designed to optimize cost and access speed based on data usage patterns:

| Storage Class | Description | Use Case | Cost |
|---------------|-------------|----------|------|
|---------------|-------------|----------|------|

|                                    |                                                                        |                                              |
|------------------------------------|------------------------------------------------------------------------|----------------------------------------------|
| S3 Standard                        | High durability, availability, and performance                         | Frequently accessed data                     |
| Highest                            |                                                                        |                                              |
| S3 Intelligent-Tiering             | Automatic cost optimization between frequent and infrequent access     |                                              |
|                                    | Unknown or changing access patterns                                    | Medium                                       |
| S3 Standard-IA (Infrequent Access) | Lower cost for data accessed less frequently but requires rapid access | Backup, disaster recovery                    |
|                                    |                                                                        | Lower                                        |
| S3 One Zone-IA                     | Same as Standard-IA but data stored in a single AZ                     | Non-critical data that can be recreated      |
|                                    |                                                                        | Lowest                                       |
| S3 Glacier Instant Retrieval       | Archive storage with milliseconds retrieval                            | Long-term archives needing occasional access |
|                                    |                                                                        | Low                                          |
| S3 Glacier Flexible Retrieval      | Archive with minutes to hours retrieval                                | Deep archive with less frequent access       |
|                                    |                                                                        | Lowest                                       |
| S3 Glacier Deep Archive            | Cheapest storage, retrieval within 12 hours                            | Long-term data archival                      |
|                                    |                                                                        | Lowest                                       |

### Lifecycle Policies

Lifecycle policies automate moving objects between storage classes or deleting them after a certain period.

Example policy: Move objects to Glacier after 30 days, delete after 365 days.

Helps optimize cost by keeping rarely accessed data in cheaper storage classes.

How to create a lifecycle policy?

Go to your bucket → Management → Lifecycle rules → Create rule.

Define the rule scope (all objects or specific prefixes/tags).

Define transitions (e.g., move to Glacier after X days).

Define expiration (e.g., delete after Y days).

### 3. Working with EBS Volumes and Snapshots

What is Amazon EBS?

Elastic Block Store (EBS) provides persistent block storage volumes for EC2 instances.

Similar to virtual hard drives that you attach to your EC2 instances.

Data persists independently of the instance lifecycle.

### Types of EBS Volumes

| Volume Type                    | Description                     | Use Case                              |
|--------------------------------|---------------------------------|---------------------------------------|
| General Purpose SSD (gp3/gp2)  | Balanced price and performance  | Boot volumes, transactional workloads |
| Provisioned IOPS SSD (io1/io2) | High performance and throughput | Databases, critical business apps     |
| Throughput Optimized HDD (st1) | Low cost, high throughput       | Big data, log processing              |

Cold HDD (sc1)      Lowest cost, infrequent access      Archives, backups

#### Creating and Attaching EBS Volumes

Create EBS volumes via AWS Console → EC2 → Elastic Block Store → Volumes → Create Volume.

Select type, size, and availability zone (must match your EC2 instance's AZ).

Attach the volume to a running EC2 instance.

On the instance, format and mount the volume to use it.

#### Snapshots

Snapshots are point-in-time backups of EBS volumes stored in S3.

Useful for backup, recovery, or creating new volumes.

Incremental by nature—only changes since last snapshot are saved, saving space.

Creating a snapshot:

Go to EC2 → Volumes → Select volume → Create snapshot.

Snapshots can be copied across regions for disaster recovery.

#### 4. Introduction to EFS for Shared File Storage

What is Amazon EFS?

Elastic File System (EFS) is a fully managed, scalable network file system.

Provides shared file storage that multiple EC2 instances can access simultaneously.

Supports standard file system interfaces and semantics (NFS protocol).

#### Use Cases for EFS

Shared content repositories.

Big data analytics.

Media workflows.

Lift-and-shift migration of applications that require shared file access.

#### Key Features

Automatically scales storage as files are added or removed.

High availability and durability across multiple Availability Zones.

Supports concurrent access from multiple instances.

Supports POSIX permissions, making it compatible with many Linux applications.

How to create and mount an EFS file system

Create an EFS file system from the AWS Console.

Configure network access via VPC security groups.

Mount the EFS file system on your EC2 instances using standard Linux mount commands.

Use the shared file system just like any local disk but accessible by multiple instances.

Summary

Amazon S3 provides scalable object storage with flexible classes for cost and access optimization.

Lifecycle policies automate data movement to cheaper storage tiers, reducing costs.

Amazon EBS offers persistent block storage with snapshots for backup and disaster recovery.

Amazon EFS provides shared file storage for multiple EC2 instances, ideal for scalable applications requiring shared access.