Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

CSU44000 Internet Applications

Week 6 Lecture 2

Pre-lab 9

AWS Storage

Conor Sheedy

Prelab for Week 9 lab, Why?

Last Lecture for a While

- Next Week is reading week (no lectures or labs, i.e. no ”Lab 7”)
- The week after has the October bank holiday on Monday (no Monday lecture or labs, i.e. no “Lab 8”)

Deadline for Assessment 1 at the end of this week

- I want to leave you with plenty of material to study during reading week
- To prepare for Assessment 2 (Week 11)
- Can complete Lab 9 and Lab 10 if you like (they will run in Week 9 and 10 and so can be assessed in Assessment 2)
- Can read Lecture notes, Week 8, Week 9 (Lecture 1 and 2) and Week 10 Lecture 1 (All can be assessed in Assessment 2, together with everything else up to that point.)

And if I finish all of that?

Decentralised Internet

- Week 10 lecture 2 and Week 11
- Will be Assessed in week 12
- MCQ worth **20%**

Register for an Account

Go to this link and register with your tcd email account

https://aws.amazon.com/education/awseducate/

- This may take a few days to be approved so do it in advance of the lab.
- Find and complete:
  - "Getting Started with Storage (Lab) " – Lab 9
  - "Getting Started with Compute (Lab) " – Lab 10
    - We will do a "Prelab" for this when we get back

Objective

To Familiarize yourself with how to deploy your Internet Application to the cloud.

- **Issues**.
  - Cloud resources are not free
    - The ‘Free Tier’ will only be free up to a certain usage level
    - It can be difficult to control your usage
    - Mistakes will occur
      - Have you ever accidentally created an infinite loop?
    - Alerts will typically let you know after you have exceeded a threshold
      - You could find yourself asking AWS to cancel a charge
        - » They may well cancel a charge for an honest mistake
        - » So do ask if you ever find yourself in this situation

Objective

How to deploy your Internet Application to the cloud.

- **Issues**.
  - Vendor lock in
    - Cloud providers often offer equivalent services
      - Infrastructure as a service
      - Platform as a service
      - Storage
      - Serverless
      - etc
    - Sometimes they will use their product name
      - Try to familiarise yourself with what terms are generic and which are specific to the provider

Approach for the Lab

Choose one example cloud provider.

- **AWS**.
  - The most popular
  - The first

Use their ‘sandbox’ learning environment.

- **No Credit Card Required**
- Limits what you can do

Approach to assessment

Assessment for the course will not require that you provide your Credit Card Details.

- **It is recommended that once you familiarise yourself with the AWS platform that you:**
  - Sign up for your own real account
    - With a Credit or Debit card
  - Practice deploying your Internet Application to the cloud
  - Sign up for their main competitors
    - Azure
    - Google Cloud Platform
  - Deploy your Internet Application there too

This will not be assessed.

You will be able to complete the module without handing over your Credit Card details

After the Lab

- **Experiment with the Sandbox to find its limits**
- **Then sign up for an account to try the other services**
  - Can you use Database services?
    - Dynamo DB?
  - Can you use serverless?
    - AWS Lambda?
  - Can you use Platform as a service?
    - AWS Elastic Beanstalk?

Then sign up for their competitors

- Google Cloud Platform
- Azure

The Lab Structure

- **The lab is divided into three sections.**
  - first section, recorded slides, 45 minutes
  - second section, hands-on lab, 45 minutes
  - third section, assessment, 20 minutes

The Pre Lab

- **Based on the first Section.**
  - Should help when you do it yourself
    - You can ask questions
    - Discuss with your classmates

Choosing a storage class – case 1

A user uploads a video to your application and your application generates a thumbnail preview of the video. Your user agreement permits a user to access the video thumbnail for one year. However, your usage data indicates that most thumbnail previews are not often accessed after 60 days but must be accessed quickly.

Which storage class might you recommend? *Choose the storage class icon.*

Choosing a storage class – case 2

A business wants to implement stronger business continuity practices and keep backups of their data for a long period of time. They do not anticipate needing to access the data often. But they want to be able to have the data within a few hours in the event of disaster recovery.

Which storage class might you recommend? *Choose the storage class icon.*

Choosing a storage class – case 2

A business wants to implement stronger business continuity practices and keep backups of their data for a long period of time. They do not anticipate needing to access the data often. But they want to be able to have the data within a few hours in the event of disaster recovery.

Which storage class might you recommend? *Choose the storage class icon.*

That's correct! Amazon S3 Glacier is a low-cost option for long-term storage. Data can be retrieved in a number of hours with standard retrieval or 1-5 minutes with expedited retrieval.

Amazon S3 Standard

Amazon S3 Standard-IA

Amazon S3 One Zone-IA

Amazon S3 Glacier

Amazon S3 Glacier Deep Archive

Amazon S3 Intelligent-Tiering

Choosing a storage class – case 3

A company builds a cloud-native application with Amazon S3 storage for school exams. The application gets some use during the school term but then is used very heavily during the exam season near the end of the term. The application has unpredictable use patterns, and the company is trying to control costs.

Which storage class might you recommend? *Choose the storage class icon.*

That's correct! Amazon S3 Intelligent-Tiering should be used when workloads are unpredictable. Amazon S3 Intelligent-Tiering will optimize storage costs by automatically moving data to the most cost-effective access tier when access patterns change.

Create a bucket

1. **Create a bucket.**
2. Configure the bucket.
3. Upload objects.
4. Work with objects.
5. Delete objects and buckets.

Choose a globally unique bucket name

- 3-63 characters
- Lowercase letters
- Numbers
- Dots (.)
- Hyphens(-)

Create a bucket

1. **Create a bucket.**
2. Configure the bucket.
3. Upload objects.
4. Work with objects.
5. Delete objects and buckets.

photobucket

myphotos.zip

https://**photobucket**.s3.amazonaws.com/**2021-04-13/myphotos.zip**

Bucket name

Key

Configure the bucket – choose a Region

1. Create a bucket.
2. **Configure the bucket.**
3. Upload objects.
4. Work with objects.
5. Delete objects and buckets.

Configure the bucket – object ownership and access

1. Create a bucket.
2. **Configure the bucket.**
3. Upload objects.
4. Work with objects.
5. Delete objects and buckets.

Object ownership

Ownership is controlled with an access control list (ACL). ACLs are the lists that you use to say who can communicate with what.

Default = Turn off ACL

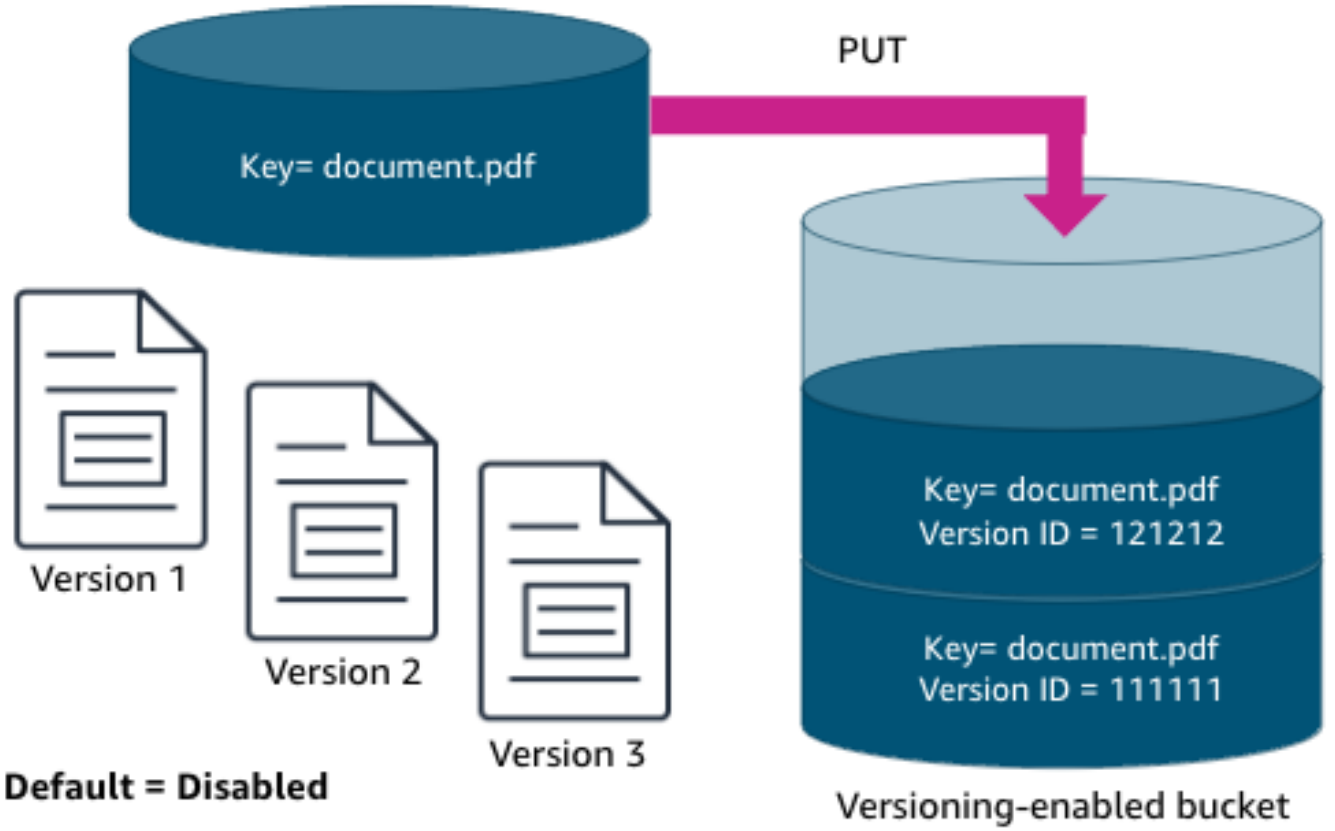
Block public access

By default, new buckets, access points, and objects don't allow public access. However, users can modify bucket policies, access point policies, or object permissions to allow public access.

Default = No public access

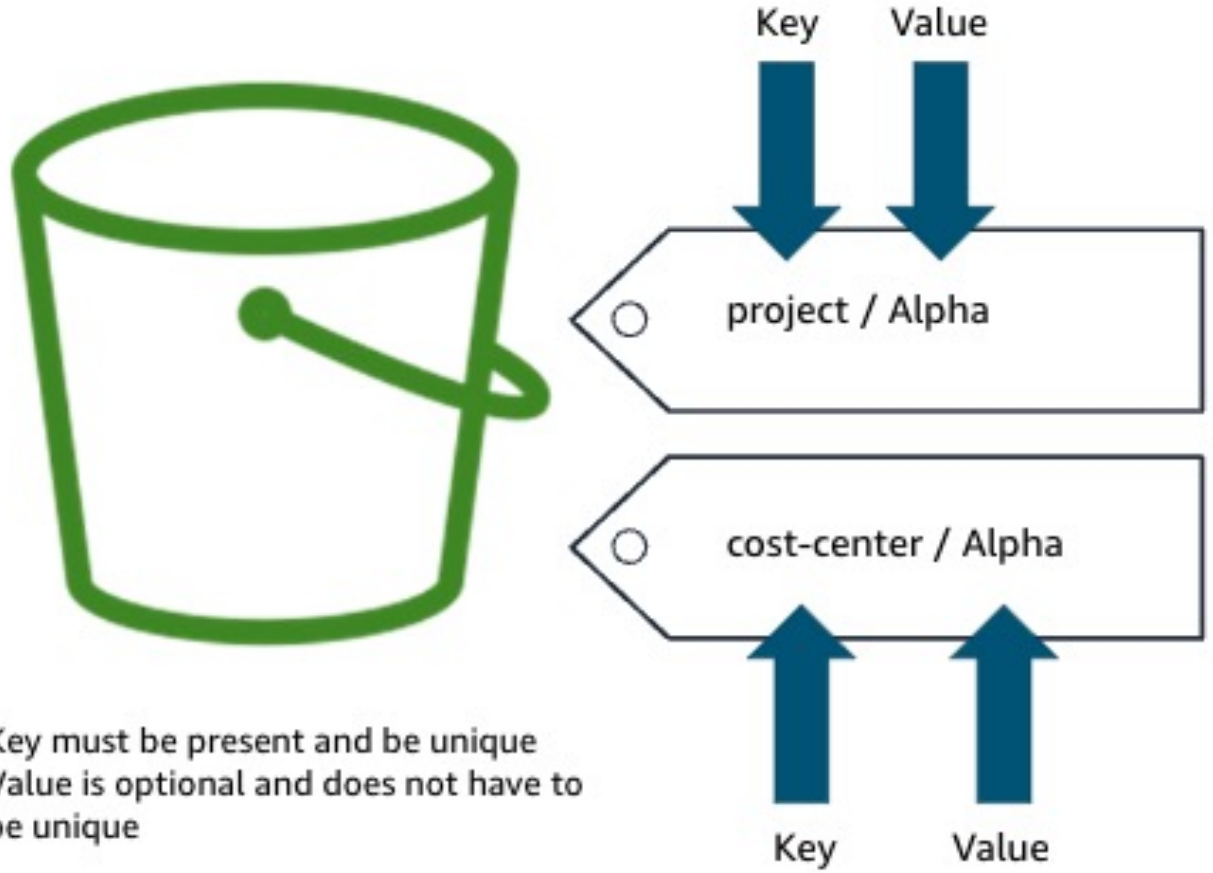
Configure the bucket – bucket versioning

1. Create a bucket.
2. **Configure the bucket.**
3. Upload objects.
4. Work with objects.
5. Delete objects and buckets.



Configure the bucket – tags

1. Create a bucket.
2. **Configure the bucket.**
3. Upload objects.
4. Work with objects.
5. Delete objects and buckets.



- Key must be present and be unique
- Value is optional and does not have to be unique

Upload objects

1. Create a bucket.
2. Configure the bucket.
3. **Upload objects.**
4. Work with objects.
5. Delete objects and buckets.

- Unlimited data
- 5 TB object limit
- 160 GB upload limit from the console

Upload objects – choose storage class

1. Create a bucket.
2. Configure the bucket.
3. **Upload objects.**
4. Work with objects.
5. Delete objects and buckets.

Default = Amazon S3 Standard

Upload objects – multipart upload

1. Create a bucket.
2. Configure the bucket.
3. **Upload objects.**
4. Work with objects.
5. Delete objects and buckets.

Use multipart for objects over 100 MB.

Work with objects

1. Create a bucket.
2. Configure the bucket.
3. Upload objects.
4. **Work with objects.**
5. Delete objects and buckets.

Copy objects

- Create a copy
- Move an object
- Rename an object
- Change object metadata

Copy URL

Bucket name

https://**photobucket**.s3.amazonaws.com/
2021-04-13/myphotos.zip

Key

Download objects

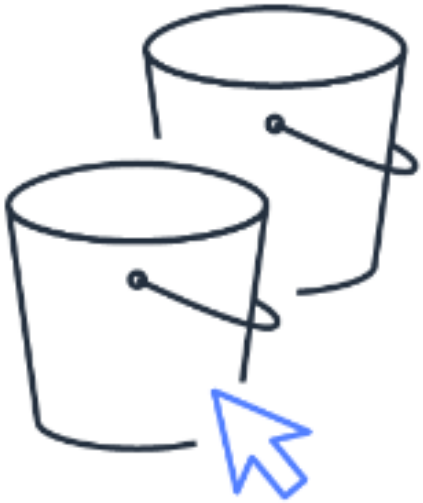
Delete objects

1. Create a bucket.
2. Configure the bucket.
3. Upload objects.
4. Work with objects.
5. **Delete objects and buckets.**

Multi-factor authentication (MFA) delete

Additional features

Lifecycle rules

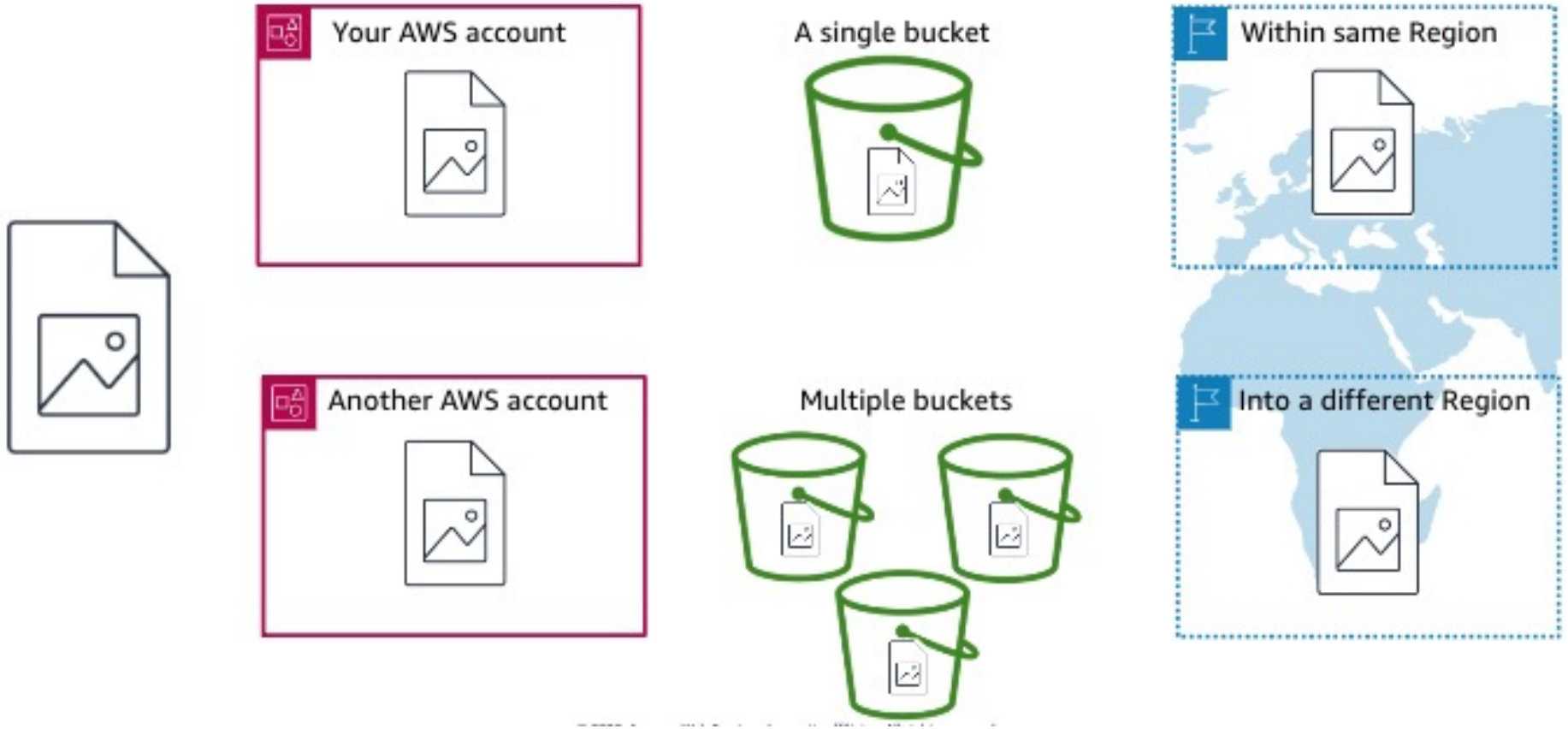


Replication rules

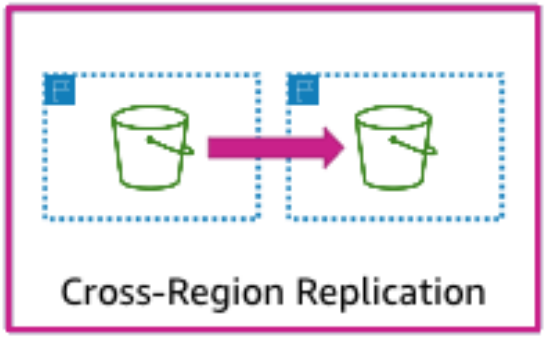
Security

Additional features – lifecycle rules

Additional features – replication rules



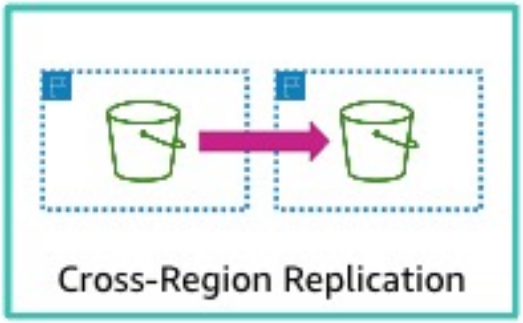
Additional features - Cross-Region and Same-Region Replication

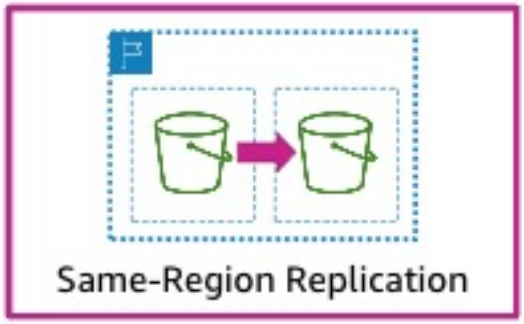


You can use Cross-Region Replication to copy objects across Amazon S3 buckets in different AWS Regions. Consider this type of replication when you need to:

- Meet compliance requirements that dictate that you store data at even greater distances for disaster recovery.
- Minimize latency by maintaining object copies in AWS Regions that are geographically closer to your users.
- Increase operational efficiency for compute clusters in two different AWS Regions by maintaining object copies in those Regions.

Additional features - Cross-Region and Same-Region Replication





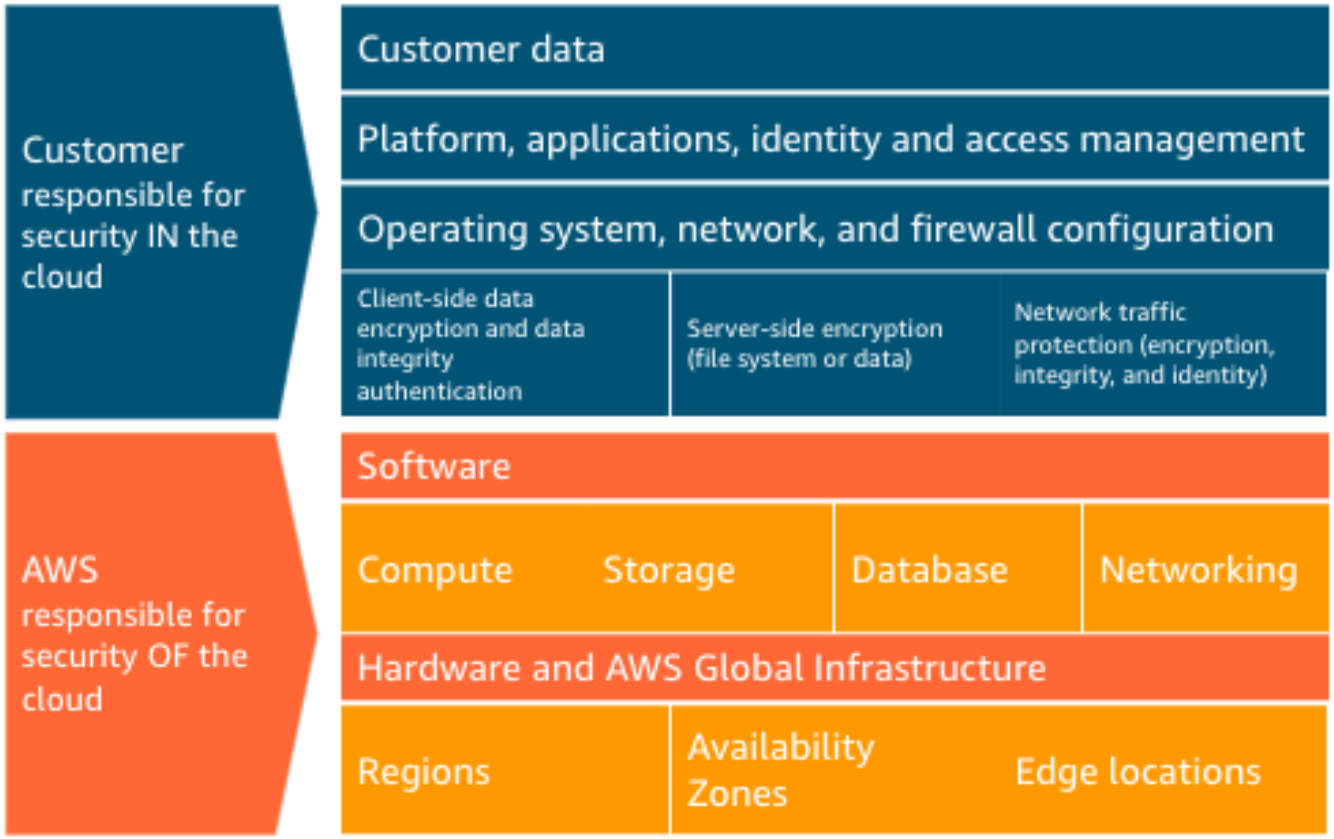
Same-Region Replication (SRR) is used to copy objects across Amazon S3 buckets in the same AWS Region. SRR can help you do the following:

- Aggregate logs into a single bucket for processing of logs in a single location.
- Configure live replication between production and test accounts that use the same data. You can replicate objects between those accounts while maintaining object metadata.
- Abide by data sovereignty laws by storing multiple copies of your data in separate AWS accounts within a certain Region.

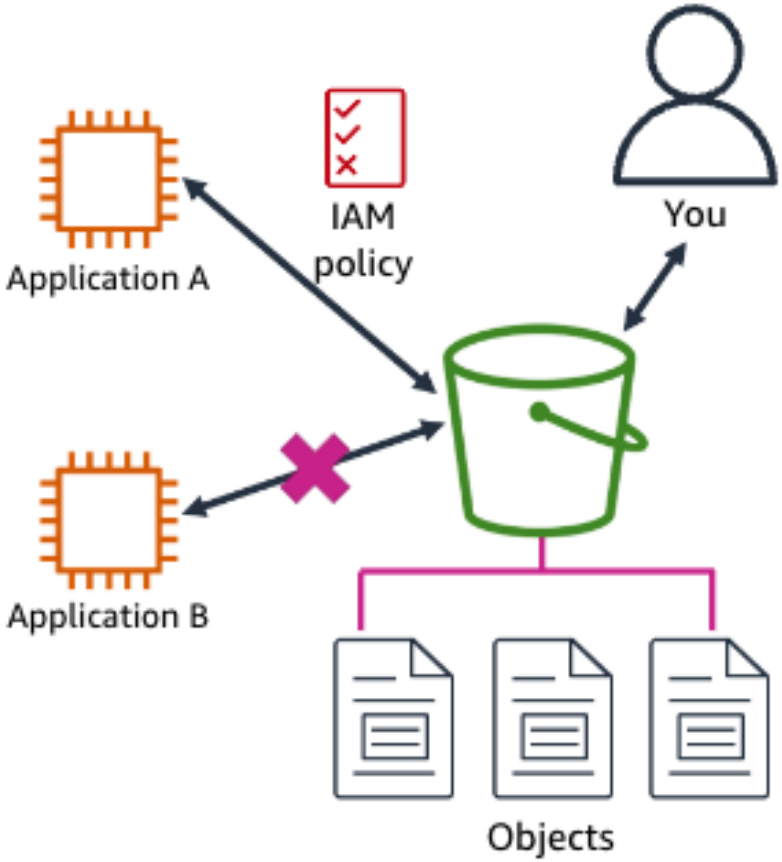
Bucket security

- AWS Identity and Access Management (IAM) policies
- Bucket policies
- Bucket encryption

For more information, see the AWS Shared Responsibility Model.



Bucket security – IAM policy



Bucket security – bucket policy

```
{
  "Version": "2012-10-17",
  "Id": "S3PolicyId1",

  "Statement": [

    {
      "Sid": "IPAllow",
      "Effect": "Deny",
      "Principal": "*",
      "Action": "s3:*",
      "Resource": "arn:aws:s3:::examplebucket/*",
      "Condition": {
         "NotIpAddress": {"aws:SourceIp":
"54.240.143.0/24"}

      }
    }
  ]
}
```

Bucket security – encryption

What is encryption? Encryption is the process of transforming readable data into another form by using a code so that only people with the secret key can read the data.

Client-side encryption - For data in transit and data at rest

Server-side encryption - For data at rest

Moving large amounts of data into Amazon S3

- **S3 Transfer Acceleration**
- AWS Snowcone
- AWS Snowball
- AWS Snowmobile

Amazon S3 Transfer Acceleration offers fast and straightforward data transfer into an S3 bucket by taking advantage of many Amazon CloudFront globally distributed edge locations. S3 Transfer Acceleration shortens the distance between client applications and Amazon S3 by using the global network of hundreds of CloudFront edge locations. Uploads and downloads are automatically routed through the closest edge locations to your application.

Consider Transfer Acceleration when you:

- Have customers all over the world who upload to a centralized bucket.
- Transfer gigabytes or terabytes of data across continents on a regular basis.

Choose the tabs to read about AWS data migration services.

Moving large amounts of data into Amazon S3

S3 Transfer Acceleration

AWS Snowcone

AWS Snowball

AWS Snowmobile

AWS Snowcone is a portable, rugged, and secure device for edge computing and data transfer. You can use a Snowcone device to collect, process, and move data to the AWS Cloud. You can move the data either offline by shipping the device to AWS, or online by using AWS DataSync.

Snowcone HDD		
Two vCPUs	4 GB of memory	8 TB of hard disk drive (HDD) based storage
Snowcone SSD		
Two vCPUs	4 GB of memory	14 TB of solid state drive (SSD) based storage

You can use AWS Snowcone devices for one-time data migration scenarios where your data is ready to be transferred. Snowcone offers a quick and low-cost way to transfer up to 8 TB or 14 TB of data to the AWS Cloud by shipping the device back to AWS.

For more information, see AWS Snowcone.

Choose the tabs to read about AWS data migration services.

Moving large amounts of data into Amazon S3

S3 Transfer Acceleration

AWS Snowcone

AWS Snowball

AWS Snowmobile

The AWS Snowball service uses physical storage devices to transfer large amounts of data between Amazon S3 and your onsite data storage location at faster-than-internet speeds.

The two different job types are import jobs and export jobs.

- Import jobs - An import job is the transfer of 72 TB or less of your data (located in an on-premises data source). It is copied onto a single Snowball, and then moved into Amazon S3.
- Export jobs - An export job is the transfer of any amount of data (located in Amazon S3). It is copied onto any number of Snowballs, and then moved one Snowball at a time into your on-premises data destination.

You can use AWS Snowball devices to migrate petabytes of data into Amazon S3 for processing and analysis.

For more information, see AWS Snowball.

Choose the tabs to read about AWS data migration services.

Moving large amounts of data into Amazon S3

- S3 Transfer Acceleration
- AWS Snowcone
- AWS Snowball
- AWS Snowmobile

AWS Snowmobile is an exabyte-scale data transfer service that is used to move extremely large amounts of data to AWS. You can transfer up to 100 PB per Snowmobile, a 45-foot long ruggedized shipping container, pulled by a semi-trailer truck. Snowmobile makes it easy to move massive volumes of data to the cloud, including video libraries, image repositories, or even a complete data center migration. Transferring data with Snowmobile is more secure, fast, and cost effective.

Using Amazon S3 summary

Next up

Additional AWS Storage Services

AWS has storage services for block and file storage that you might want to use for particular workloads. In the next section, you will learn more about Amazon Elastic Block Store and Amazon Elastic File System.

In this section, you have learned about how to use Amazon S3.

The key takeaways from this section are:

- In a few steps, you can create a bucket and begin uploading objects.
- Objects can be copied, downloaded, or deleted.
- Additional features include lifecycle policies, replication rules, and security.

Additional resources:

Getting started with Amazon S3

Bucket naming rules

Amazon S3 video tutorials

Managing your storage lifecycle

Replicating objects

AWS Snow Family

In this module

Choose the icon to review each section.

Additional AWS Storage Services

Amazon S3
object storage

Amazon EBS
block storage

Amazon EFS
file Storage

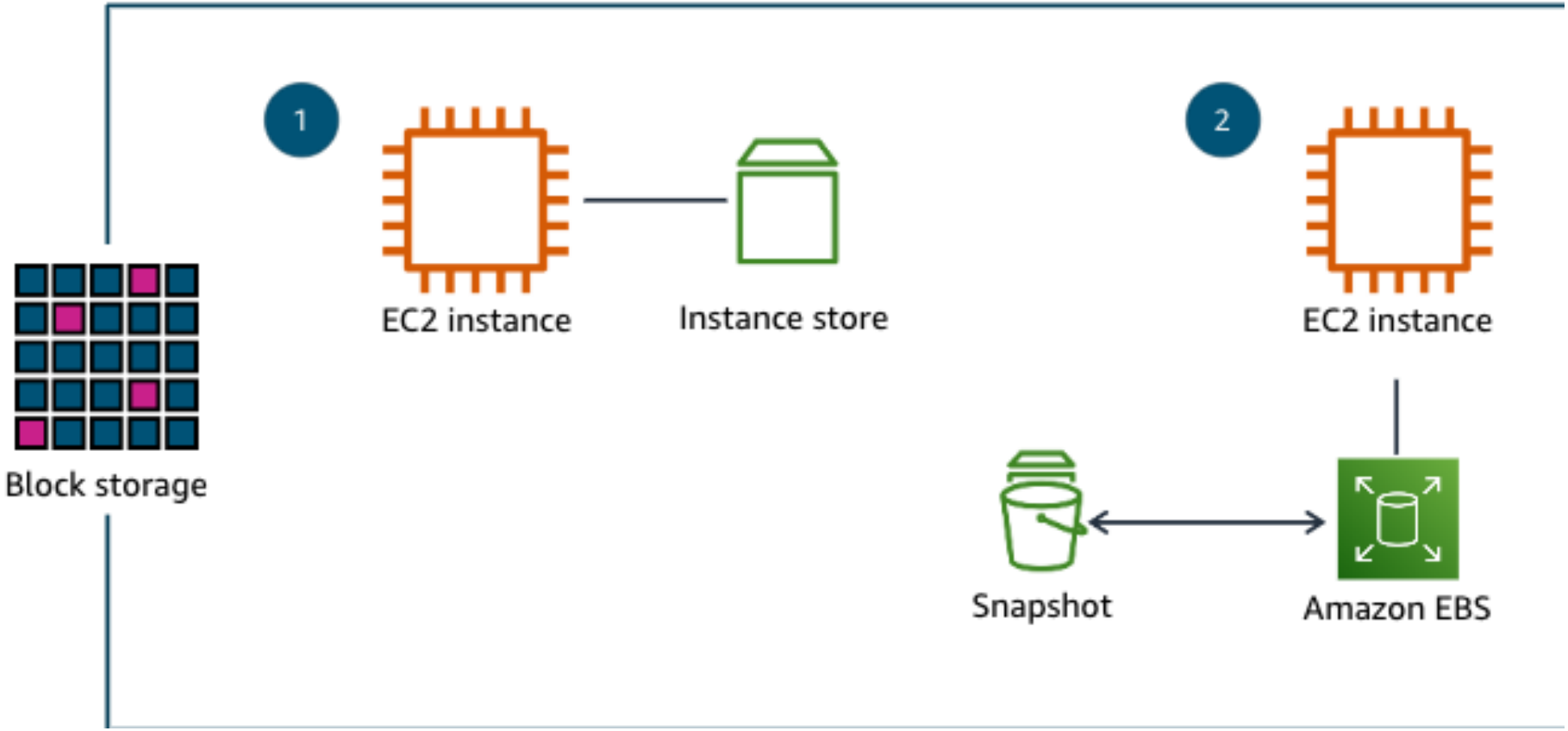
Introduction
to Storage

Introduction to
Amazon S3

Using
Amazon S3

Additional AWS
Storage Services

Block storage at AWS



Amazon EBS key features and benefits

- Data availability
- Data persistence
- Data encryption
- Data security
- Snapshots
- Flexibility

When you create an EBS volume, it is automatically replicated within its Availability Zone to prevent data loss due to failure of any single hardware component. You can attach an EBS volume to any EC2 instance in the same Availability Zone. After you attach a volume, it appears as a native block device similar to a hard drive or other physical device. At that point, the instance can interact with the volume just as it would with a local drive.

Amazon EBS key features and benefits

- Data availability
- Data persistence
- Data encryption
- Data security
- Snapshots
- Flexibility

An EBS volume is off-instance storage that can persist independently from the life of an instance. You continue to pay for the volume usage as long as the data persists.

EBS volumes can be set to automatically detach from the instance with the data intact and can be reattached to a new instance, enabling quick recovery.

If you are using an EBS-backed instance, you can stop and restart that instance without affecting the data stored in the attached volume. The volume remains attached throughout the stop-start cycle.

Amazon EBS key features and benefits

- Data availability
- Data persistence
- Data encryption
- Data security
- Snapshots
- Flexibility

For data encryption, you can create encrypted EBS volumes with the Amazon EBS encryption feature. All EBS volume types support encryption. Amazon EBS encryption uses 256-bit Advanced Encryption Standard algorithms (AES-256) and an Amazon-managed key infrastructure. The encryption occurs on the server that hosts the EC2 instance, providing encryption of data-in-transit from the EC2 instance to Amazon EBS storage.

Amazon EBS key features and benefits

- Data availability
- Data persistence
- Data encryption
- **Data security**
- Snapshots
- Flexiblity

Amazon EBS volumes are presented to you as raw, unformatted block devices. These devices are logical devices that are created on the EBS infrastructure. The Amazon EBS service verifies that the devices are logically empty prior to any use or reuse by a customer.

Amazon EBS key features and benefits

- Data availability
- Data persistence
- Data encryption
- Data security
- Snapshots
- Flexibility

You can back up the data on your Amazon EBS volumes to Amazon S3 by taking point-in-time snapshots.

Snapshots are incremental backups, which means that only the blocks on the device that have changed after your most recent snapshot are saved. This method minimizes the time required to create the snapshot and saves on storage costs by not duplicating data.

Each snapshot contains all of the information that is needed to restore your data (from the moment when the snapshot was taken) to a new EBS volume.

Amazon EBS key features and benefits

- Data availability
- Data persistence
- Data encryption
- Data security
- Snapshots
- Flexibility

Amazon EBS is flexible, which means that you can make changes without service interruptions. You can modify the volume type, volume size, and IOPS capacity while in production.

When you make a change, you are only charged for the new volume configuration after the change is in effect.

Amazon EBS volume types

Solid state drives (SSD)

General purpose
Balance of price and performance suitable for most workloads

Provisioned IOPS
High performance for mission-critical, low-latency, or high-throughput workloads

Hard disk drives (HDD)

Throughput optimized HDD
Low-cost HDD for frequently accessed, throughput-intensive workloads

Cold HDD
The lowest-cost HDD design for less frequently accessed workloads

Use cases for EBS volume types

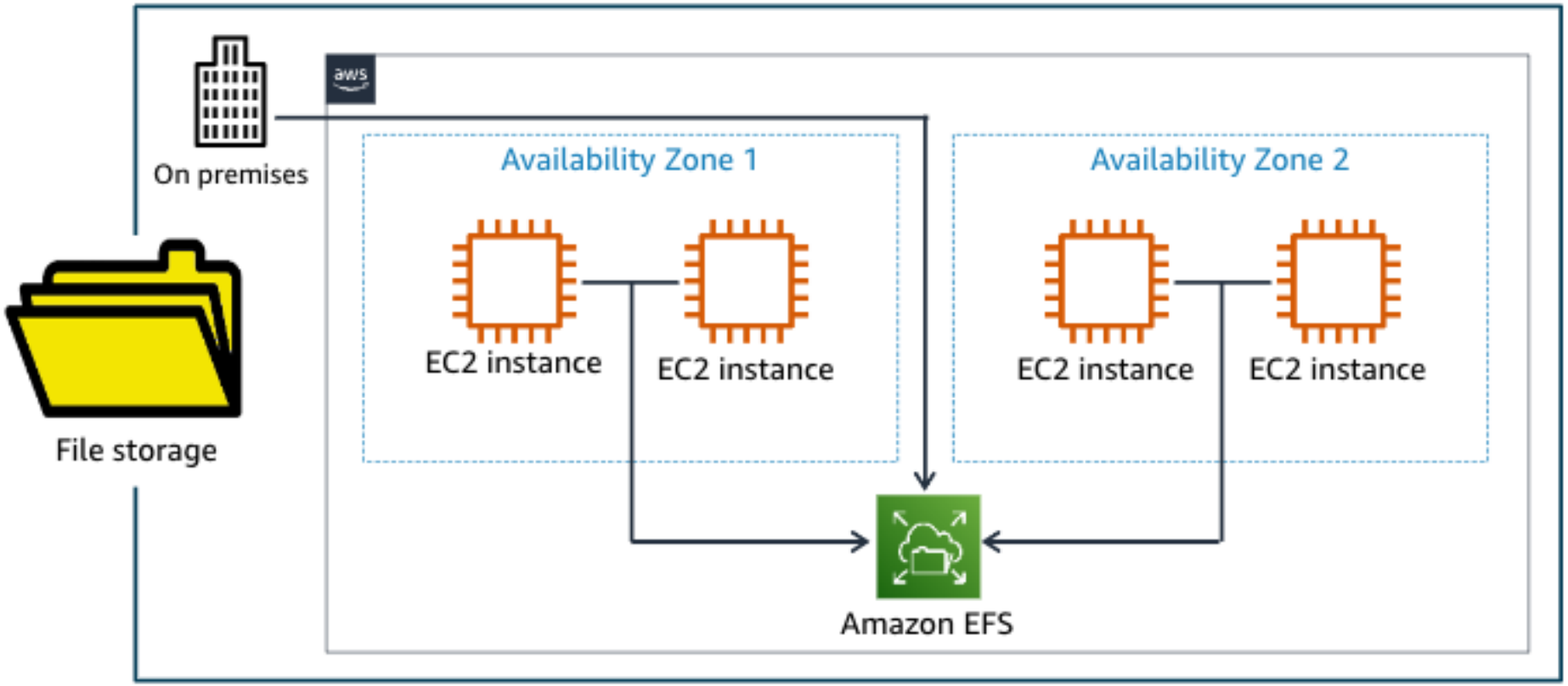
Solid state drives (SSD)

General purpose SSD	Provisioned IOPS SSD
• Transactional workloads • Virtual desktops • Medium-sized databases • Low-latency interactive applications • Boot volumes • Development and test environments	• Workloads that require sustained IOPS performance • I/O-intensive database workload

Hard disk drives (HDD)

Throughput Optimized HDD	Cold HDD
• Big data • Data warehouses • Log processing	• Throughput-oriented storage for data that is infrequently accessed • Scenarios where the lowest storage cost is important

File storage at AWS



Amazon EFS key features and benefits

- Fully managed
- Highly available and durable
- Elastic and scalable
- Data encryption

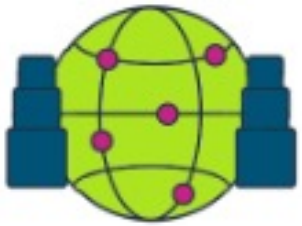
Amazon EFS is a fully managed service. As a result, you don't need to worry about managing file servers or storage, updating hardware, configuring software, or performing backups. AWS does all of those things for you.

You can set up Amazon EFS in seconds to create and configure file systems. As you use Amazon EFS, all that you are responsible for is your data.

Amazon EFS key features and benefits

- Fully managed
- Highly available and durable
- Elastic and scalable
- Data encryption

Amazon EFS is highly available and durable. It was designed for 99.999999999 percent (11 9s) of durability and up to 99.99 percent (4 9s) of availability. By default, Amazon EFS redundantly stores every file system object (in other words, a directory, file, and link) across multiple Availability Zones for file systems. It uses Standard storage classes.



Amazon EFS key features and benefits

- Fully managed
- Highly available and durable
- Elastic and scalable
- Data encryption

With Amazon EFS, storage capacity is elastic. Therefore, the storage capacity grows and shrinks automatically as you add and remove files. This elasticity provides storage capacity to your applications as needed.

With elastic capacity, provisioning is unnecessary, and you're billed only for what you use. Amazon EFS is designed to be highly scalable both in storage capacity and throughput performance. With Amazon EFS, throughput and IOPS scale as your file system grows, while file operations are delivered with consistent low latencies.

Amazon EFS key features and benefits

- Fully managed
- Highly available and durable
- Elastic and scalable
- Data encryption

Amazon EFS provides a comprehensive encryption solution to secure both your stored data and data in flight. Data at rest is transparently encrypted by using encryption keys that the AWS Key Management Service (AWS KMS) manages. Encryption of data in transit uses industry-standard Transport Layer Security (TLS) to secure network traffic.

Amazon EFS use cases

Content management

Media processing workflows

Shared and home directories

Database backups

Developer and application tools

Big data analytics

Other AWS Storage Services summary

Next up

Wrap up and Amazon S3 Lab

You have reached the end of the content and next is a quick wrap up of the course. Once you have completed the course material you will move to the Amazon S3 Lab where you will practice your new knowledge and skills.

In this section, you have learned about other AWS storage services and their uses.

The key takeaways from this section are:

- Amazon EBS is a block storage service.
- Block, file, and object are the categories of storage.
- Choose the appropriate storage based on workload criteria.
- AWS has several storage solutions to choose from.

Additional resources:

Amazon Elastic Block Store (EBS)

Amazon EBS volumes

Amazon EBS Backup & Restore using AWS Backup

Amazon Elastic File System

What is Amazon Elastic File System?

Create a Network File System with Amazon EFS

Summary

	**Introduction to storage**	Storage basics Types of storage	How to choose storage Storage use cases
	**Introduction to Amazon S3**	Key concepts of Amazon S3 Storage classes	 Storage class use cases
	**Using Amazon S3**	Create a bucket Manage storage	Configure a bucket Manage additional settings
	**Additional AWS storage services**	Amazon EBS Amazon EFS	