



JAMES COOK
UNIVERSITY
AUSTRALIA

EE4500

Module 7: Storage

AWS Academy Cloud Foundations

Module overview

Topics

- Amazon Elastic Block Store (Amazon EBS)
- Amazon Simple Storage Service (Amazon S3)
- Amazon Elastic File System (Amazon EFS)
- Amazon Simple Storage Service Glacier

Demos

- Amazon EBS console
- Amazon S3 console
- Amazon EFS console
- Amazon S3 Glacier console

Lab

- Working with Amazon EBS

Activities

- Storage solution case study



Knowledge check

Module objectives

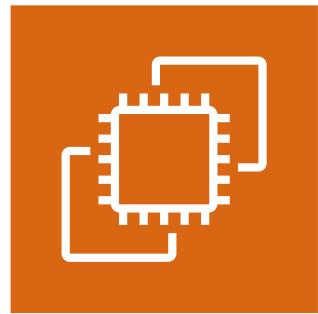
After completing this module, you should be able to:

- Identify the different types of storage
- Explain Amazon S3
- Identify the functionality in Amazon S3
- Explain Amazon EBS
- Identify the functionality in Amazon EBS
- Perform functions in Amazon EBS to build an Amazon EC2 storage solution
- Explain Amazon EFS
- Identify the functionality in Amazon EFS
- Explain Amazon S3 Glacier
- Identify the functionality in Amazon S3 Glacier
- Differentiate between Amazon EBS, Amazon S3, Amazon EFS, and Amazon S3 Glacier

Core AWS services



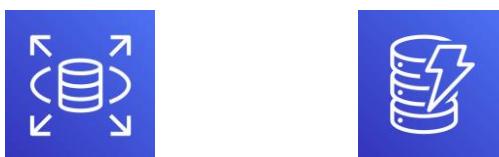
**Amazon Virtual
Private Cloud
(Amazon VPC)**



**Amazon Elastic
Compute Cloud
(Amazon EC2)**



Database



Amazon Relational
Database Service Amazon
DynamoDB



**AWS Identity and Access
Management (IAM)**

Section 1: Amazon Elastic Block Store (Amazon EBS)

Module 7: Storage

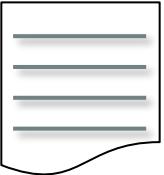


Storage

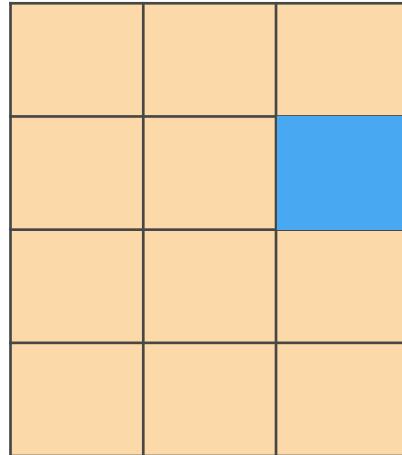


Amazon Elastic Block Store
(Amazon EBS)

AWS storage options: Block storage versus object storage



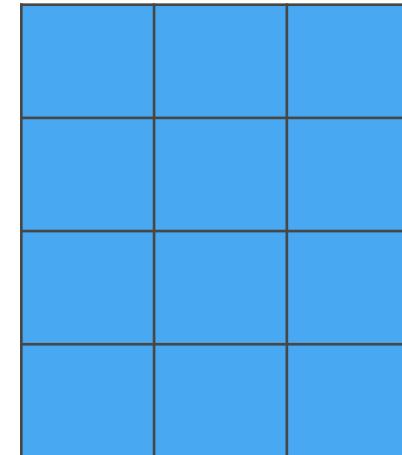
What if you want to change **one character** in a 1-GB file?



Block storage

Change one block (piece of the file)

that contains the character



Object storage

Entire file must be updated

Amazon EBS

Amazon EBS enables you to **create individual storage volumes** and **attach them** to an Amazon EC2 instance:

- Amazon EBS offers block-level storage.
- Volumes are automatically replicated within its Availability Zone.
- It can be backed up automatically to Amazon S3 through snapshots.
- Uses include –
 - Boot volumes and storage for Amazon Elastic Compute Cloud (Amazon EC2) instances
 - Data storage with a file system
 - Database hosts
 - Enterprise applications



Amazon EBS volume types

Maximum Volume Size
Maximum IOPS/Volume
Maximum Throughput/Volume

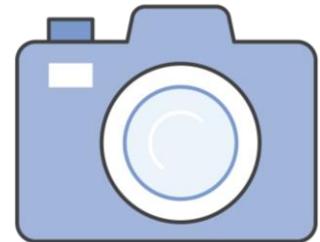
| | Solid State Drives (SSD) | | Hard Disk Drives (HDD) | |
|---------------------------|--------------------------|------------------|------------------------|-----------|
| | General Purpose | Provisioned IOPS | Throughput-Optimized | Cold |
| Maximum Volume Size | 16 TiB | 16 TiB | 16 TiB | 16 TiB |
| Maximum IOPS/Volume | 16,000 | 64,000 | 500 | 250 |
| Maximum Throughput/Volume | 250 MiB/s | 1,000 MiB/s | 500 MiB/s | 250 MiB/s |

Amazon EBS volume type use cases

| Solid State Drives (SSD) | | Hard Disk Drives (HDD) | |
|---|--|---|---|
| General Purpose | Provisioned IOPS | Throughput-Optimized | Cold |
| <ul style="list-style-type: none">This type is recommended for most workloads | <ul style="list-style-type: none">Critical business applications that require sustained IOPS performance, or more than 16,000 IOPS or 250 MiB/second of throughput per volumeLarge database workloads | <ul style="list-style-type: none">Streaming workloads that require consistent, fast throughput at a low price | <ul style="list-style-type: none">Throughput-oriented storage for large volumes of data that is infrequently accessed |
| <ul style="list-style-type: none">System boot volumesVirtual desktopsLow-latency interactive applicationsDevelopment and test environments | | <ul style="list-style-type: none">Big dataData warehousesLog processing | <ul style="list-style-type: none">Scenarios where the lowest storage cost is importantIt cannot be a boot volume |
| | | <ul style="list-style-type: none">It cannot be a boot volume | |

Amazon EBS features

- Snapshots –
 - Point-in-time snapshots
 - Recreate a new volume at any time
- Encryption –
 - Encrypted Amazon EBS volumes
 - No additional cost
- Elasticity –
 - Increase capacity
 - Change to different types



Amazon EBS: Volumes, IOPS, and pricing

1. Volumes –

- Amazon EBS volumes persist independently from the instance.
- All volume types are charged by the amount that is provisioned per month.

2. IOPS –

- General Purpose SSD:
 - Charged by the amount that you provision in GB per month until storage is released.
- Magnetic:
 - Charged by the number of requests to the volume.
- Provisioned IOPS SSD:
 - Charged by the amount that you provision in IOPS (multiplied by the percentage of days that you provision for the month).

Amazon EBS: Snapshots and data transfer

3. Snapshots –

- Added cost of Amazon EBS snapshots to Amazon S3 is per GB-month of data stored.

4. Data transfer –

- Inbound data transfer is free.
- Outbound data transfer across Regions incurs charges.



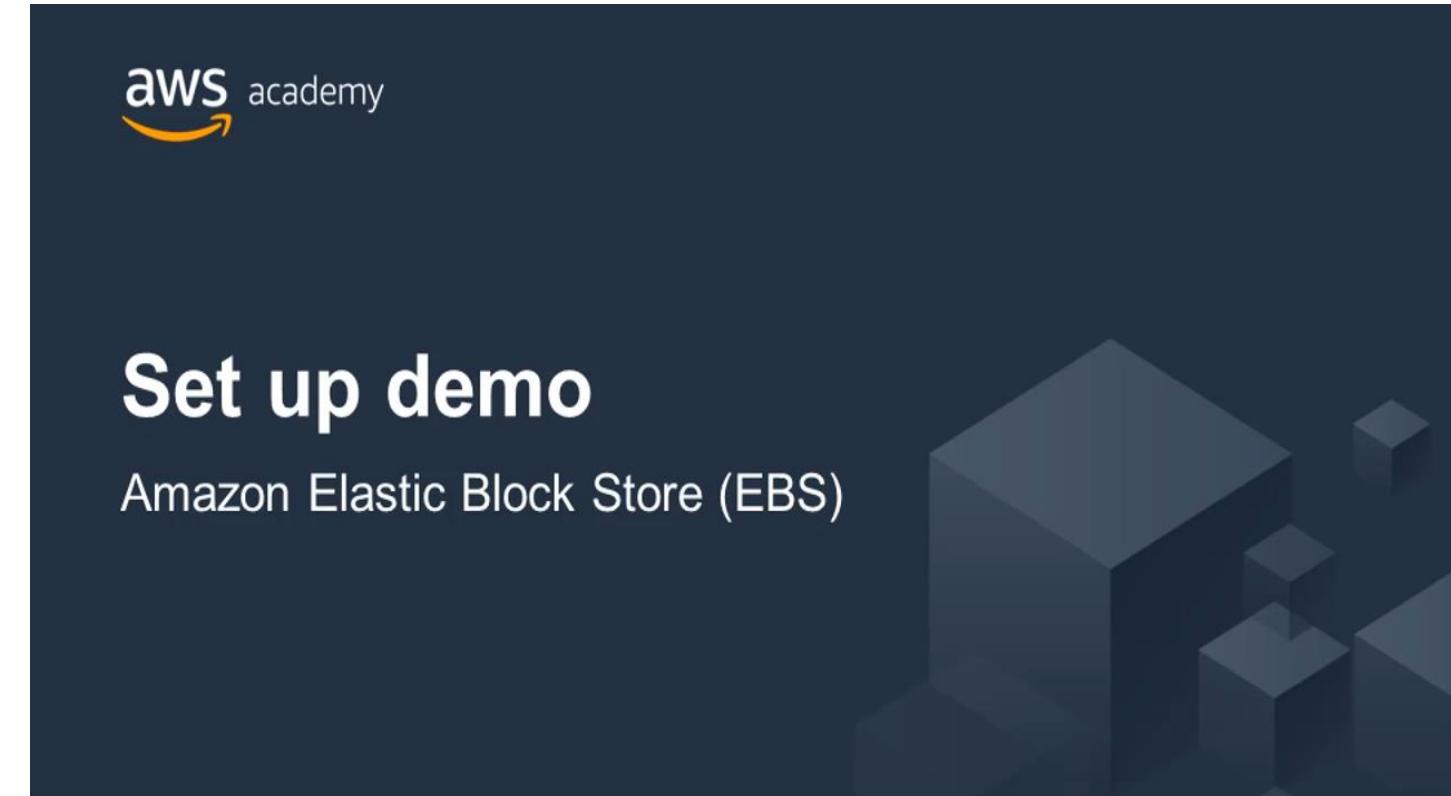
Section 1 key takeaways



Amazon EBS features:

- Persistent and customizable block storage for Amazon EC2
- HDD and SSD types
- Replicated in the same Availability Zone
- Easy and transparent encryption
- Elastic volumes
- Back up by using snapshots

Recorded demo: Amazon Elastic Block Store



Lab 4: Working with Amazon EBS

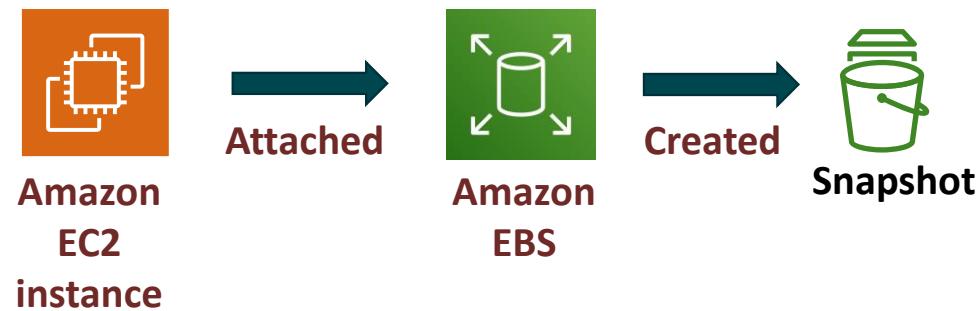


Lab 4: Scenario

This lab is designed to show you how to create an Amazon EBS volume. After you create the volume, you will attach the volume to an Amazon EC2 instance, configure the instance to use a virtual disk, create a snapshot and then restore from the snapshot.



Lab 4: Final product





~ 30 minutes



Begin Lab 4: Working with Amazon EBS

Lab debrief: Key takeaways



Section 2: Amazon Simple Storage Service (Amazon S3)

Module 7: Storage



Storage



Amazon Simple Storage Service
(Amazon S3)

Amazon S3 overview

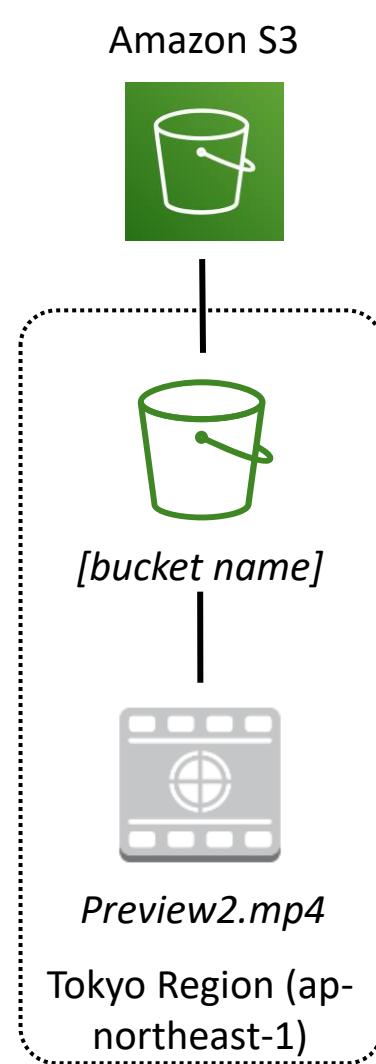
- Data is stored as objects in buckets
- Virtually unlimited storage
 - Single object is limited to 5 TB
- Designed for 11 9s of durability
- Granular access to bucket and objects

Amazon S3 storage classes

Amazon S3 offers a range of object-level storage classes that are designed for different use cases:

- Amazon S3 Standard
- Amazon S3 Intelligent-Tiering
- Amazon S3 Standard-Infrequent Access (Amazon S3 Standard-IA)
- Amazon S3 One Zone-Infrequent Access (Amazon S3 One Zone-IA)
- Amazon S3 Glacier
- Amazon S3 Glacier Deep Archive

Amazon S3 bucket URLs (two styles)



To upload your data:

1. Create a **bucket** in an AWS Region.
2. Upload almost any number of **objects** to the bucket.

Bucket path-style URL endpoint:

<https://s3.ap-northeast-1.amazonaws.com/bucket-name>

Region code

Bucket name

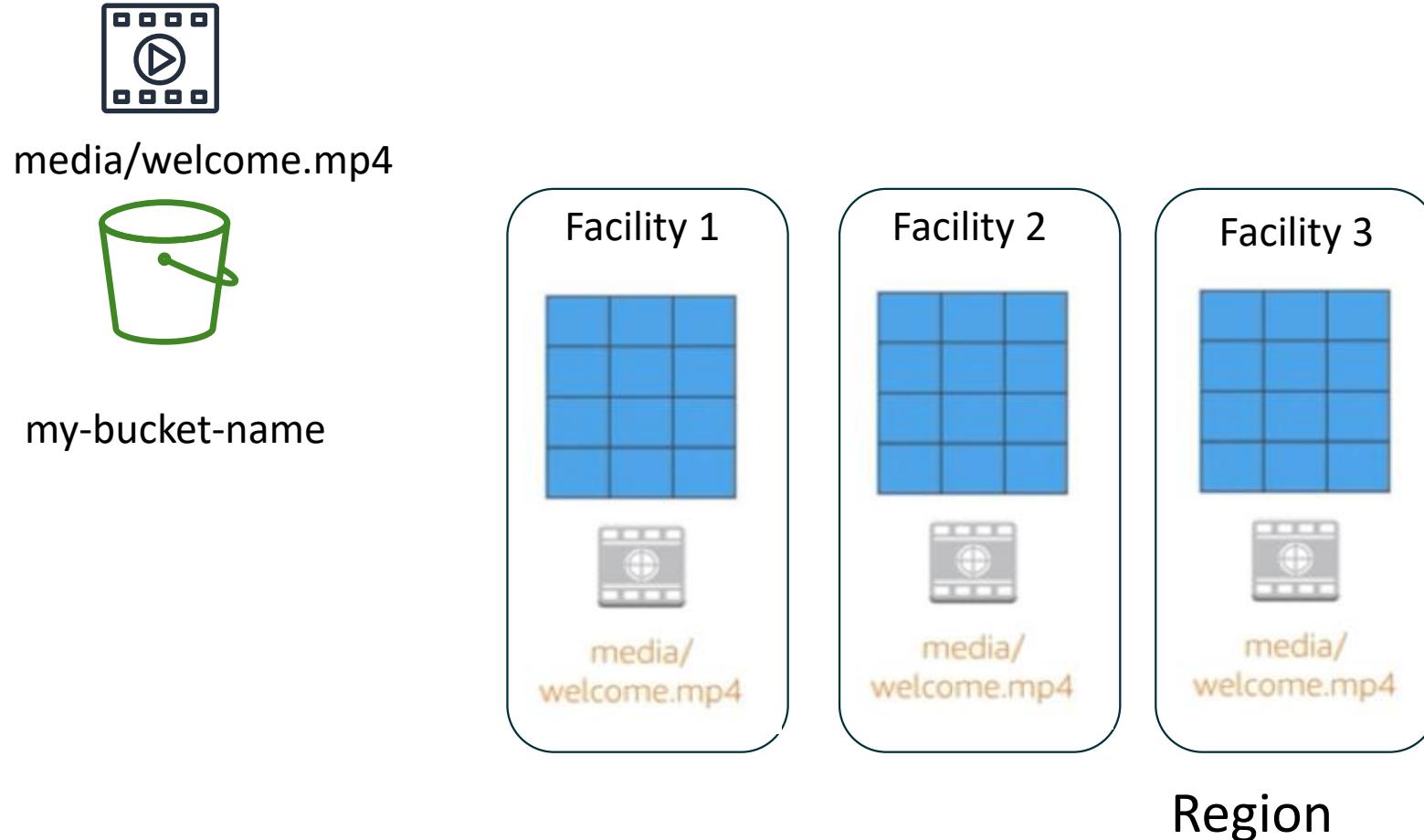
Bucket virtual hosted-style URL endpoint:

<https://bucket-name.s3-ap-northeast-1.amazonaws.com>

Bucket name

Region code

Data is redundantly stored in the Region



Designed for seamless scaling



my-bucket-name



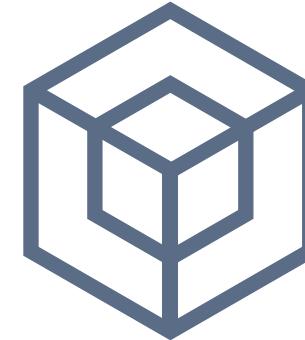
Access the data anywhere



AWS Management
Console



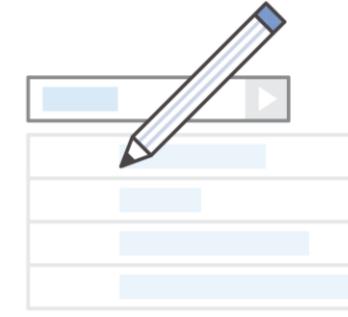
AWS Command Line
Interface



SDK

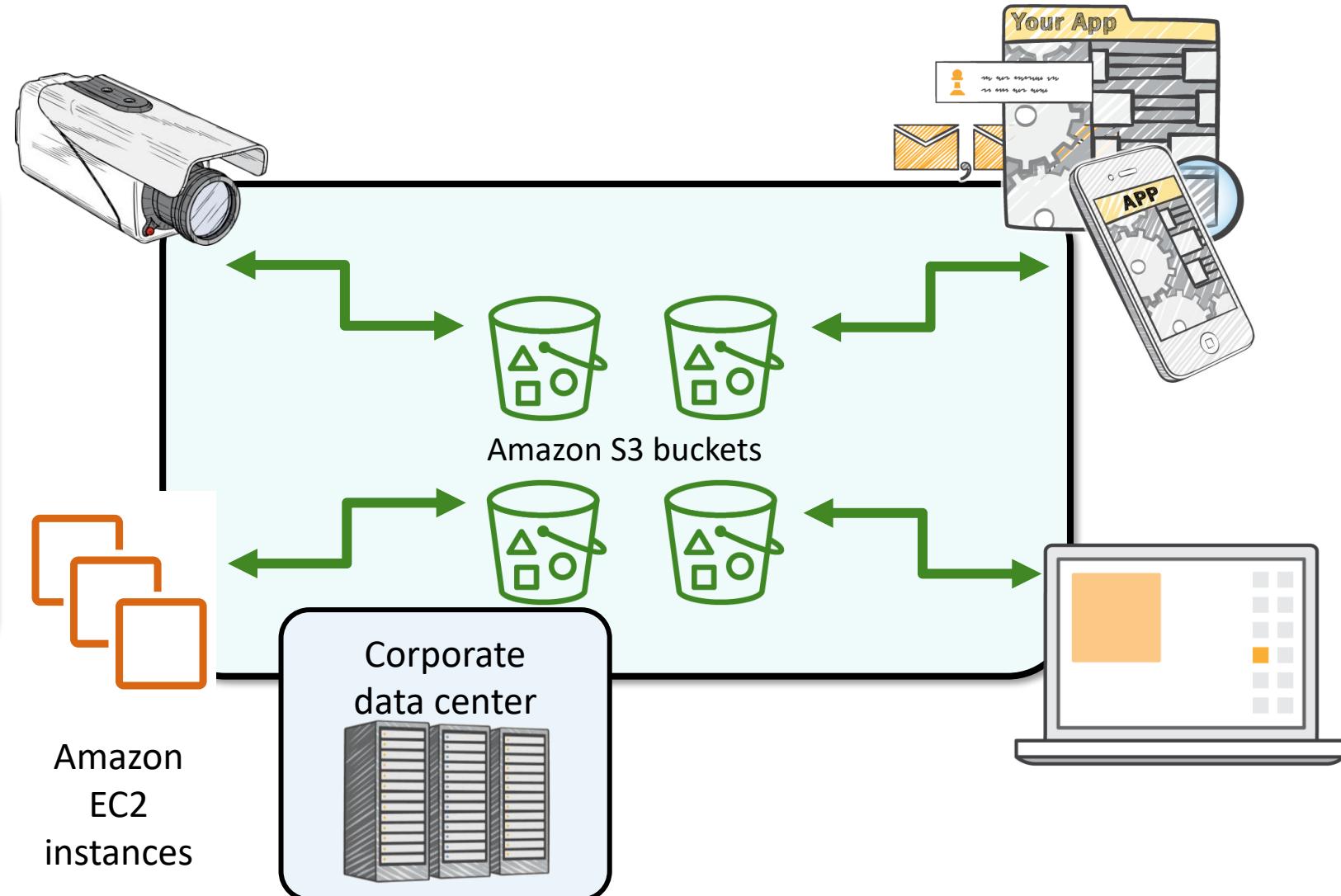
Common use cases

- Storing application assets
- Static web hosting
- Backup and disaster recovery (DR)
- Staging area for big data
- *Many more....*



Amazon S3 common scenarios

- Backup and storage
- Application hosting
- Media hosting
- Software delivery



Amazon S3 pricing

- Pay only for what you use, including –
 - GBs per month
 - Transfer OUT to other Regions
 - PUT, COPY, POST, LIST, and GET requests
- You do not pay for –
 - Transfers IN to Amazon S3
 - Transfers OUT from Amazon S3 to Amazon CloudFront or Amazon EC2 in the same Region

Amazon S3: Storage pricing (1 of 2)

To estimate Amazon S3 costs, consider the following:

1. Storage class type –

- Standard storage is designed for:
 - 11 9s of durability
 - Four 9s of availability
- S3 Standard-Infrequent Access (S-IA) is designed for:
 - 11 9s of durability
 - Three 9s of availability

2. Amount of storage –

- The number and size of objects

Amazon S3: Storage pricing (2 of 2)

3. Requests –

- The number and type of requests (**GET, PUT, COPY**)
- Type of requests:
 - Different rates for GET requests than other requests.

4. Data transfer –

- Pricing is based on the amount of data that is transferred out of the Amazon S3 Region
- Data transfer in is free, but you incur charges for data that is transferred out.

Section 2 key takeaways



- Amazon S3 is a fully managed cloud storage service.
- You can store a virtually unlimited number of objects.
- You pay for only what you use.
- You can access Amazon S3 at any time from anywhere through a URL.
- Amazon S3 offers rich security controls.

Recorded demo:
Amazon Simple
Storage System



Section 3: Amazon Elastic File System (Amazon EFS)

Module 7: Storage



Storage

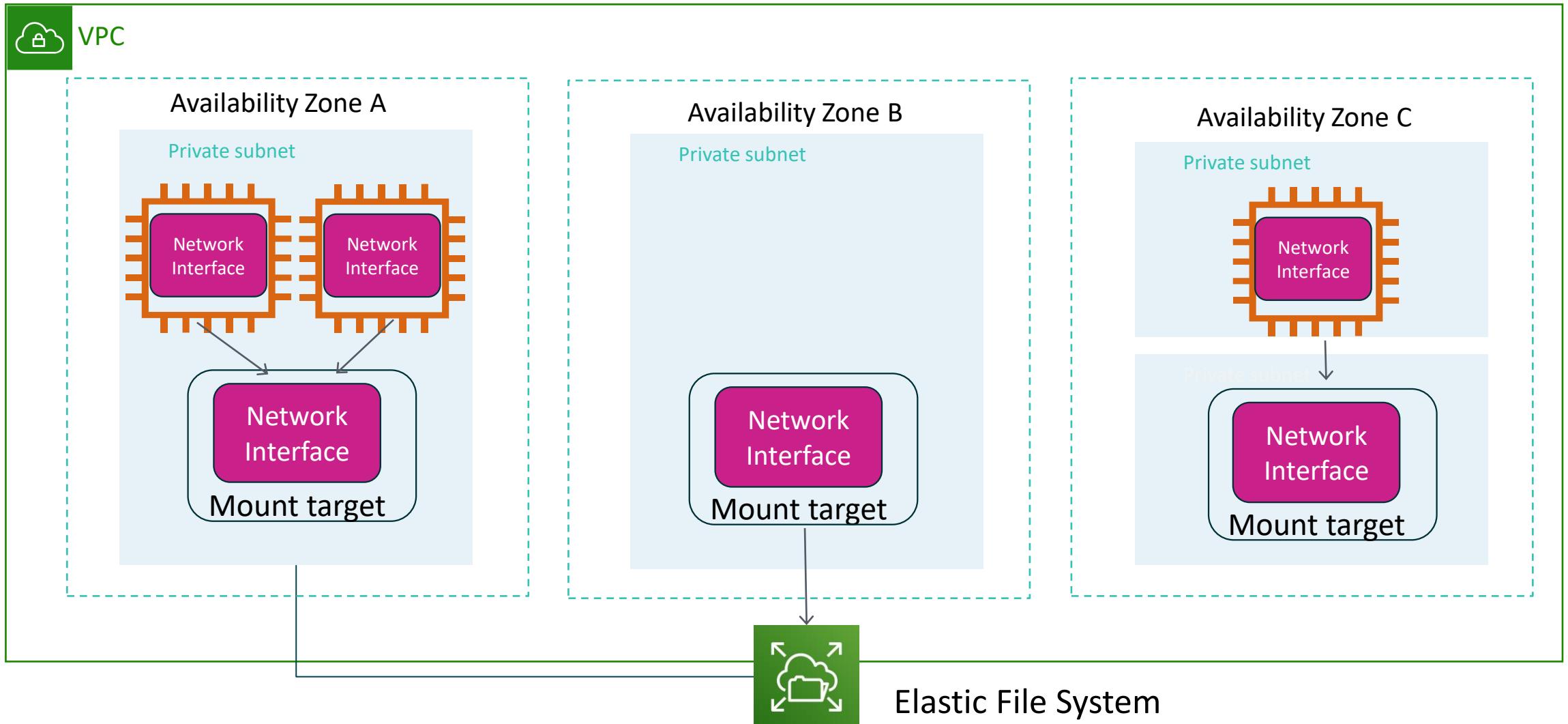


Amazon Elastic File System
(Amazon EFS)

Amazon EFS features

- File storage in the AWS Cloud
- Works well for big data and analytics, media processing workflows, content management, web serving, and home directories
- Petabyte-scale, low-latency file system
- Shared storage
- Elastic capacity
- Supports Network File System (NFS) versions 4.0 and 4.1 (NFSv4)
- Compatible with all Linux-based AMIs for Amazon EC2

Amazon EFS architecture



Amazon EFS implementation

- 1** Create your Amazon EC2 resources and launch your Amazon EC2 instance.
- 2** Create your Amazon EFS file system.
- 3** Create your mount targets in the appropriate subnets.
- 4** Connect your Amazon EC2 instances to the mount targets.
- 5** Verify the resources and protection of your AWS account.

Amazon EFS resources

File system

- Mount target
 - Subnet ID
 - Security groups
 - One or more per file system
 - Create in a VPC subnet
 - One per Availability Zone
 - Must be in the same VPC
- Tags
 - Key-value pairs



Section 3 key takeaways



- Amazon EFS provides file storage over a network.
- Perfect for big data and analytics, media processing workflows, content management, web serving, and home directories.
- Fully managed service that eliminates storage administration tasks.
- Accessible from the console, an API, or the CLI.
- Scales up or down as files are added or removed and you pay for what you use.

Recorded demo: Amazon Elastic File System



Section 4: Amazon S3 Glacier

Module 7: Storage



Storage



Amazon S3 Glacier

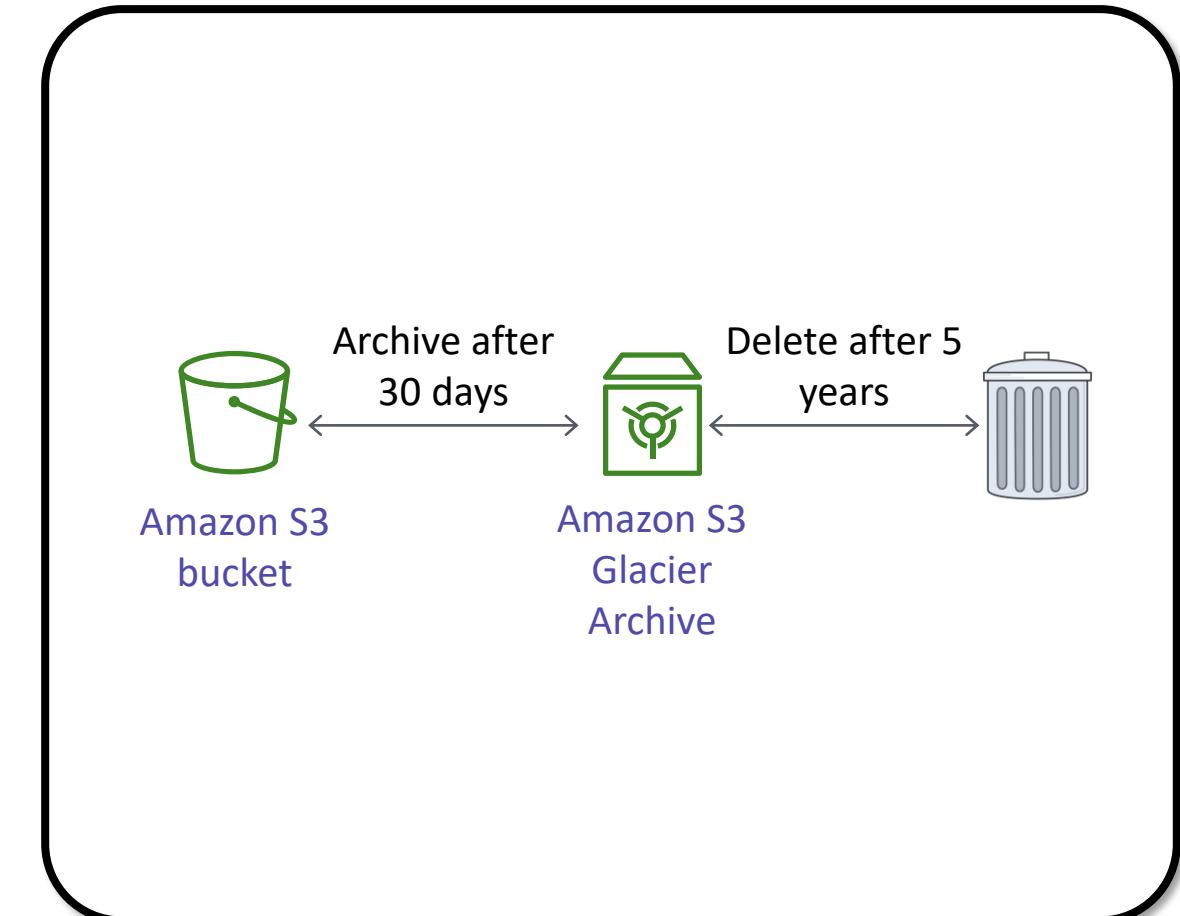
Amazon S3 Glacier review

Amazon S3 Glacier is a data archiving service that is designed for **security**, **durability**, and an **extremely low cost**.

- Amazon S3 Glacier is designed to provide 11 9s of durability for objects.
- It supports the encryption of data in transit and at rest through Secure Sockets Layer (SSL) or Transport Layer Security (TLS).
- The Vault Lock feature enforces compliance through a policy.
- Extremely low-cost design works well for long-term archiving.
 - Provides three options for access to archives—expedited, standard, and bulk—retrieval times range from a few minutes to several hours.

Amazon S3 Glacier

- Storage service for low-cost data archiving and long-term backup
- You can configure lifecycle archiving of Amazon S3 content to Amazon S3 Glacier
- Retrieval options –
 - Standard: 3–5 hours
 - Bulk: 5–12 hours
 - Expedited: 1–5 minutes



Amazon S3 Glacier use cases



Media asset archiving



Healthcare information archiving



Regulatory and compliance archiving



Scientific data archiving

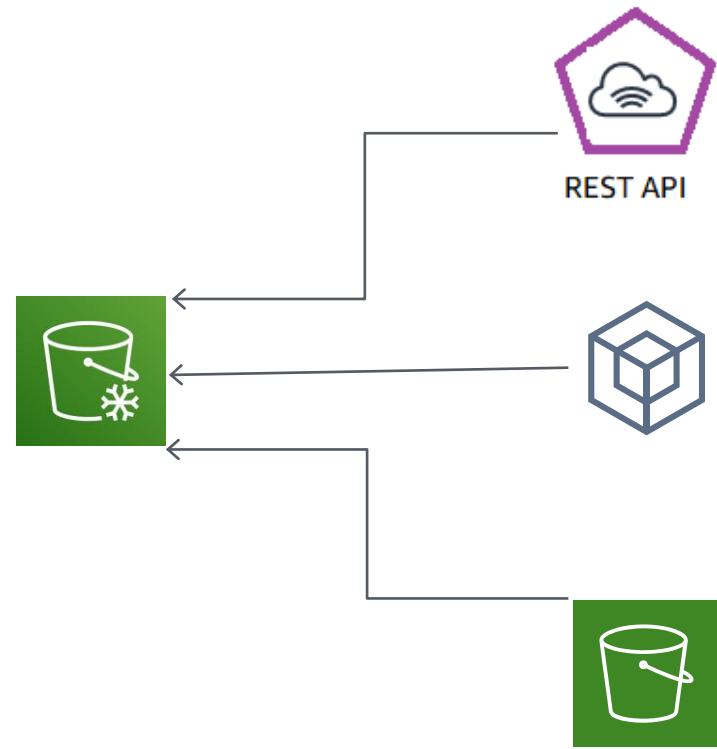


Digital preservation



Magnetic tape replacement

Using Amazon S3 Glacier



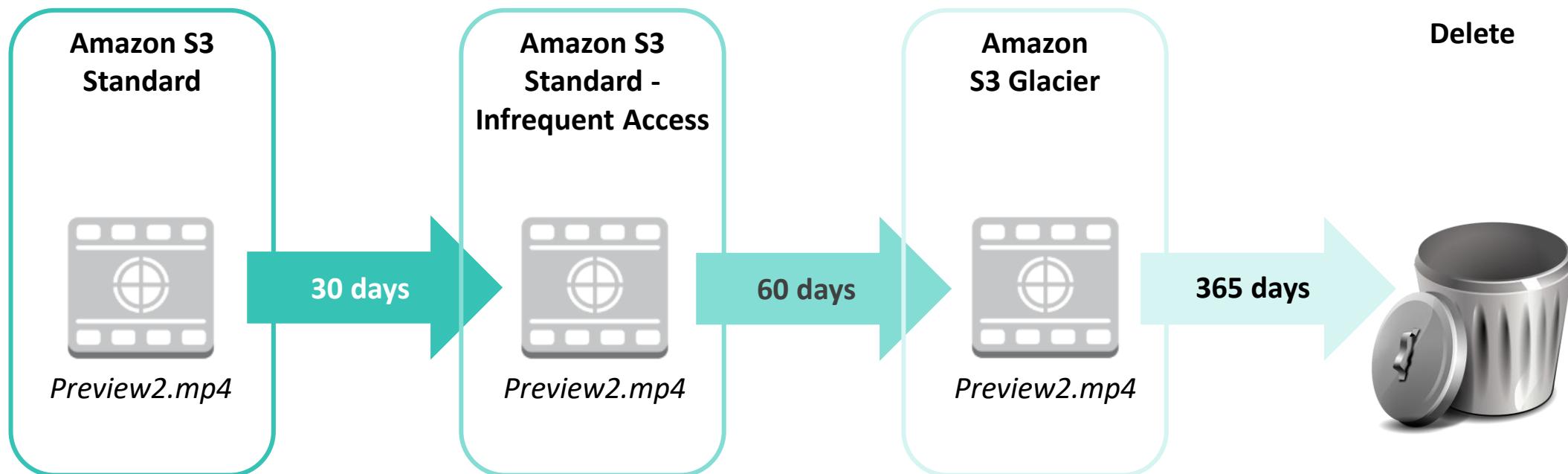
RESTful
web services

Java or .NET
SDKs

Amazon S3 with
lifecycle policies

Lifecycle policies

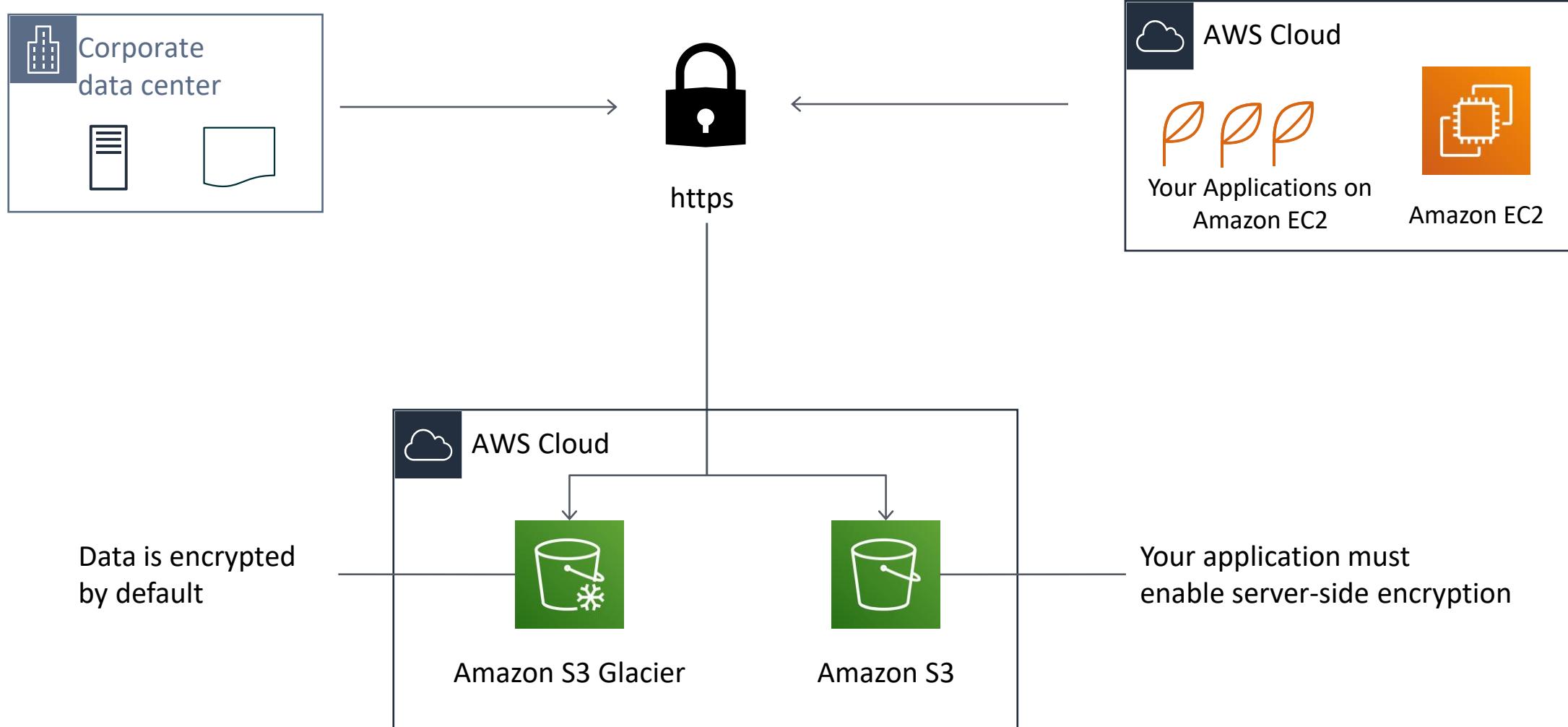
Amazon S3 lifecycle policies enable you to delete or move objects based on age.



Storage comparison

| | Amazon S3 | Amazon S3 Glacier |
|-------------------|-----------------------------------|------------------------------|
| Data Volume | No limit | No limit |
| Average Latency | ms | minutes/hours |
| Item Size | 5 TB maximum | 40 TB maximum |
| Cost/GB per Month | Higher cost | Lower cost |
| Billed Requests | PUT, COPY, POST, LIST, and GET | UPLOAD and retrieval |
| Retrieval Pricing | ¢ Per request | ¢¢ Per request and per GB |

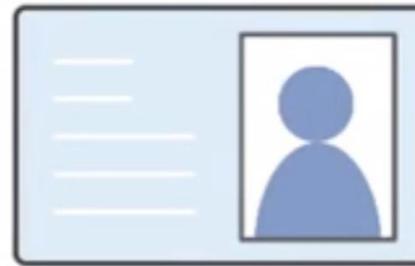
Server-side encryption



Security with Amazon S3 Glacier



**Amazon S3
Glacier**



**Control access with
IAM**



**Amazon S3 Glacier encrypts
your data with AES-256**



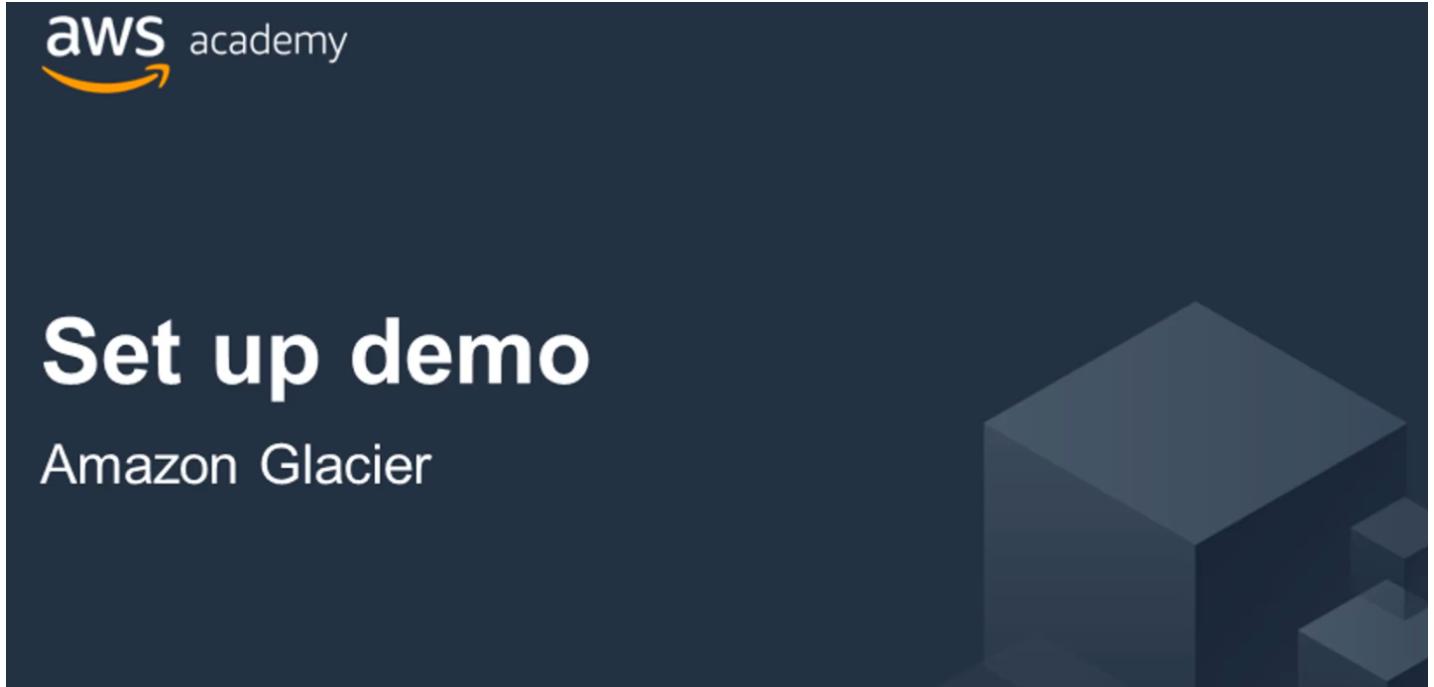
**Amazon S3 Glacier manages
your keys for you**

Section 4 key takeaways



- Amazon S3 Glacier is a data archiving service that is designed for security, durability, and an extremely low cost.
- Amazon S3 Glacier pricing is based on Region.
- Its extremely low-cost design works well for long-term archiving.
- The service is designed to provide 11 9s of durability for objects.

Recorded demo:
Amazon S3 Glacier



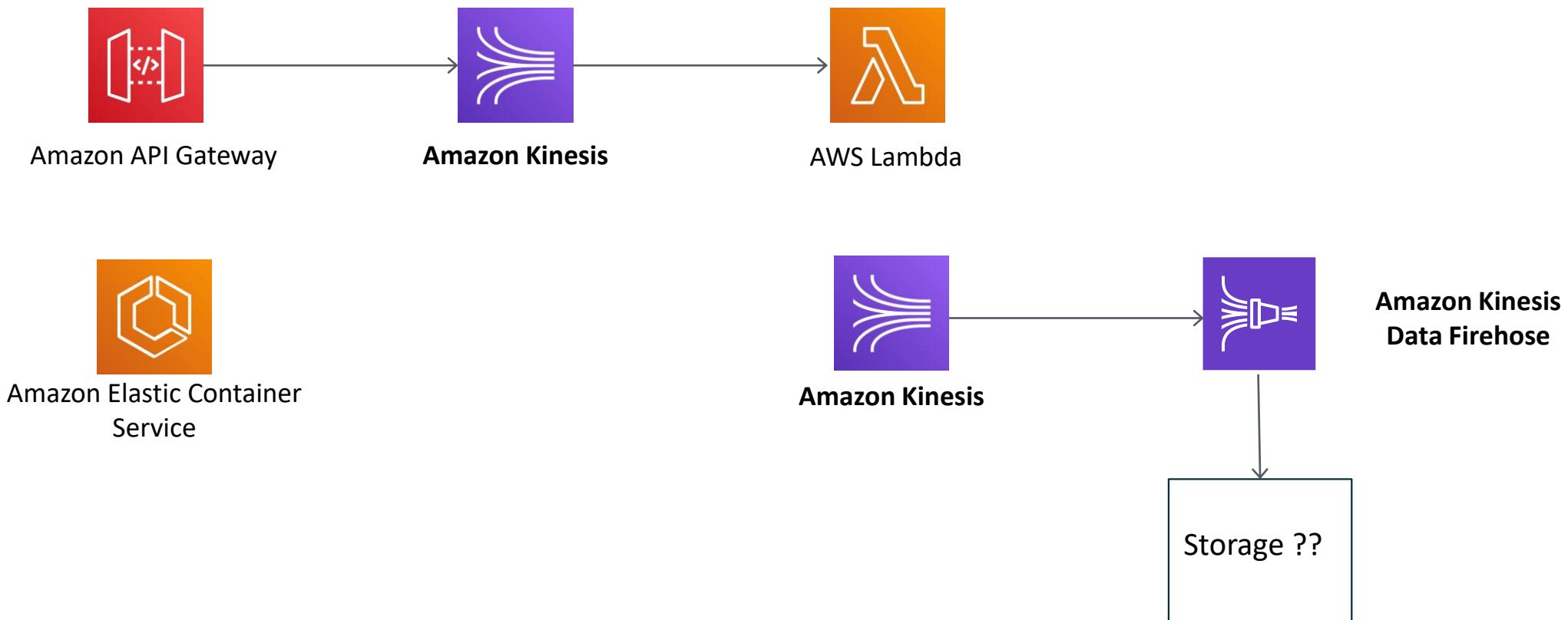
Activity: Storage Case Studies



Photo by Pixabay from Pexels.

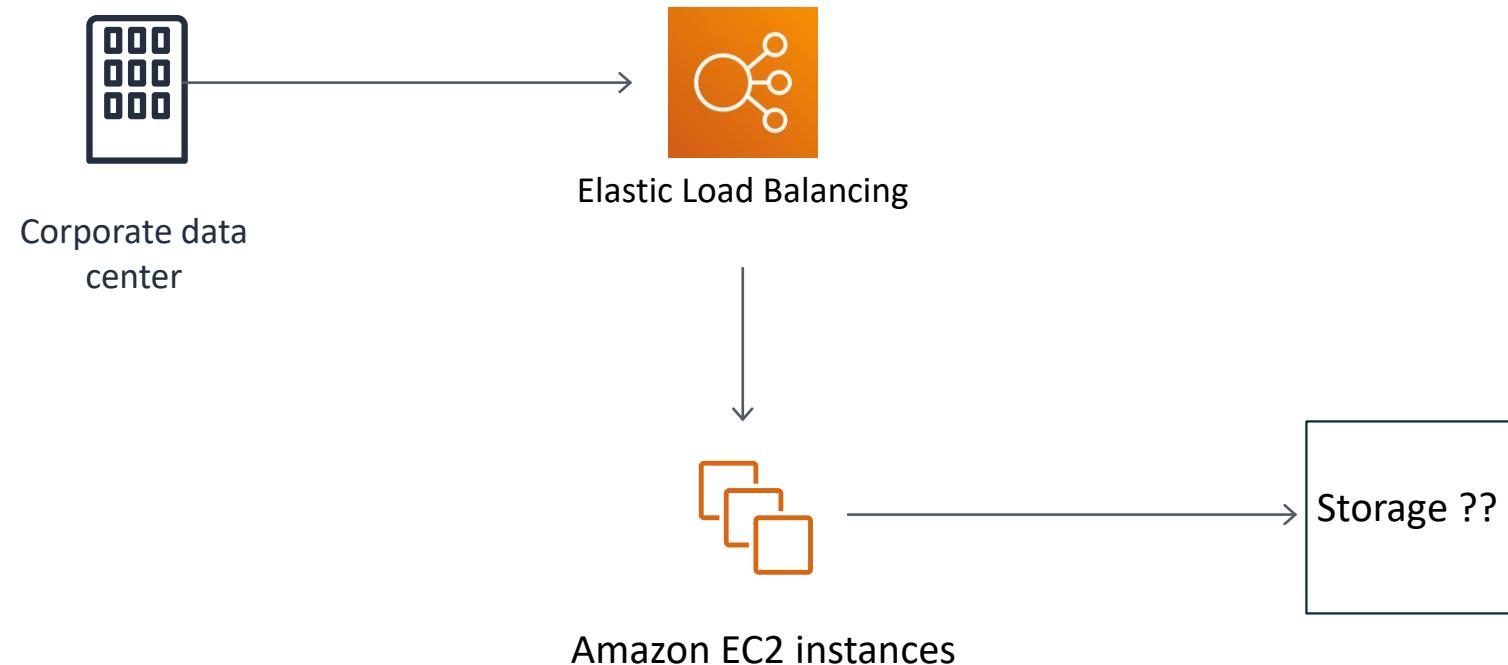
Storage case study activity (1 of 3)

Case 1: A data analytics company for travel sites must store billions of customer events per day. They use the data analytics services that are in the diagram. The following diagram illustrates their architecture.



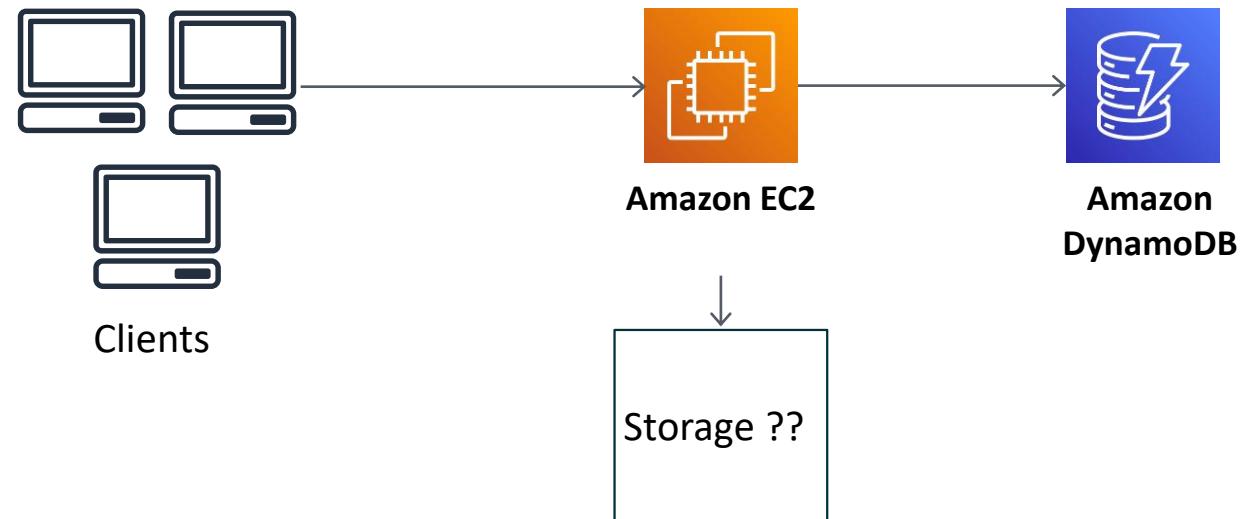
Storage case study activity (2 of 3)

Case 2: A collaboration software company processes email for enterprise customers. They have more than 250 enterprise customers and more than half a million users. They must store petabytes of data for their customers. The following diagram illustrates their architecture.



Storage case study activity (3 of 3)

Case 3: A data protection company must be able to ingest and store large amounts of customer data and help their customers meet compliance requirements. They use Amazon EC2 for scalable compute and Amazon DynamoDB for duplicate data and metadata lookups. The following diagram illustrates their architecture.



Module wrap-up

Module 7: Storage



Module summary

In summary, in this module, you learned how to:

- Identify the different types of storage
- Explain Amazon S3
- Identify the functionality in Amazon S3
- Explain Amazon EBS
- Identify the functionality in Amazon EBS
- Perform functions in Amazon EBS to build an Amazon EC2 storage solution
- Explain Amazon EFS
- Identify the functionality in Amazon EFS
- Explain Amazon S3 Glacier
- Identify the functionality in Amazon S3 Glacier
- Differentiate between Amazon EBS, Amazon S3, Amazon EFS, and Amazon S3 Glacier

Complete the knowledge check



Sample exam question

A company wants to store data that is not frequently accessed. What is the best and cost-effective solution that should be considered?

| Choice | Response |
|--------|---|
| A | AWS Storage Gateway |
| B | Amazon Simple Storage Service Glacier |
| C | Amazon Elastic Block Store (Amazon EBS) |
| D | Amazon Simple Storage Service (Amazon S3) |

Sample exam question answer

A company wants to store data that is not frequently accessed. What is the best and cost-effective solution that should be considered?

The correct answer is **B. Amazon Simple Storage Service Glacier**

The keywords in the question are “not frequently accessed” and “cost-effective solution.”

Additional resources

- AWS Storage page: <https://aws.amazon.com/products/storage/>
- Storage Overview: <https://docs.aws.amazon.com/whitepapers/latest/aws-overview/storage-services.html>
- Recovering files from an Amazon EBS volume backup: <https://aws.amazon.com/blogs/compute/recovering-files-from-an-amazon-ebs-volume-backup/>
- Confused by AWS Storage Options? S3, EFS, EBS Explained: <https://dzone.com/articles/confused-by-aws-storage-options-s3-ebs-and-efs-explained>

Thank you

All trademarks are the property of their owners.

