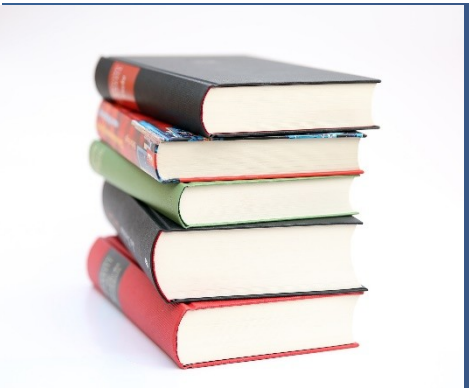
UNIVERSITY OF ŽILINA
Faculty of Management Science and Informatics

Lecture 3 - Storage

- AWS M7 - Storage

Vytvorené v rámci projektu **KEGA 026TUKE-4/2021**

aws academy



Outline

- **AWS M7 - Storage**
  - Amazon Elastic Block Store (Amazon EBS)
  - Amazon Simple Storage Service (Amazon S3)
  - Amazon Elastic File System (Amazon EFS)
  - Amazon Simple Storage Service Glacier

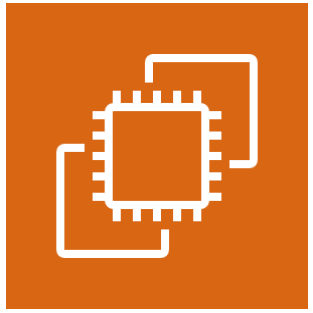
Lecture objectives

After completing this lecture, 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

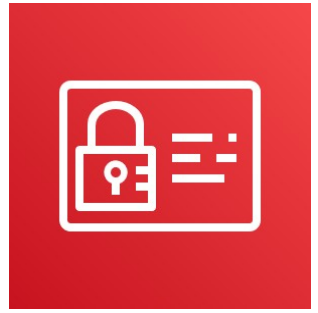
Core AWS services

Amazon Virtual Private Cloud (Amazon VPC)



Amazon Elastic Compute Cloud (Amazon EC2)

Storage



AWS Identity and Access Management (IAM)

Amazon Relational Database Service

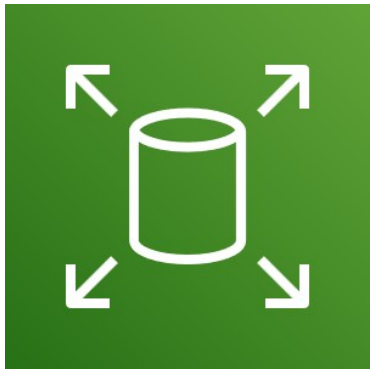
Amazon DynamoDB

Database

Amazon Elastic Block Store (Amazon EBS)

Amazon EBS Storage

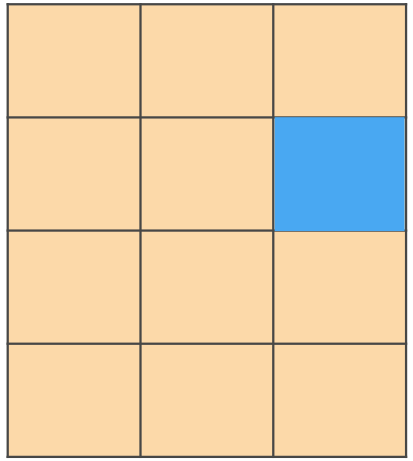
- persistent block storage volumes for use with Amazon EC2 instances
- Persistent storage
  - data storage device that retains data after power is shut off
- Each Amazon EBS volume is automatically replicated within AZ
- high availability and durability
- consistent and low-latency performance



Amazon Elastic Block Store (Amazon EBS)

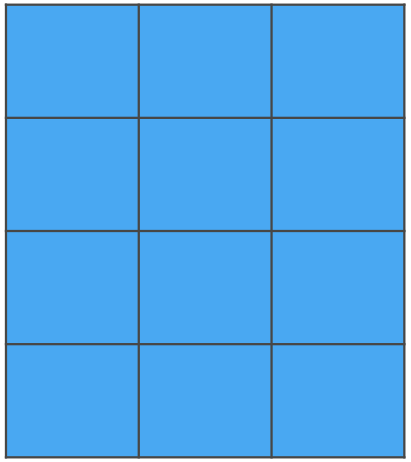
AWS storage options: Block storage versus object storage

What if you want to change <u>one character</u> 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

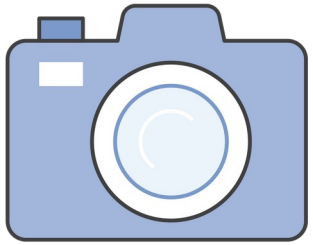
	Solid State Drives (SSD)		Hard Disk Drives (HDD)	
	General Purpose	Provisioned IOPS	Throughput-Optimized	Cold
Maximum Volume Size	16 TiB	16 TiB	16 TiB	16 TiB
Maximum IOPS/Volume	16,000	64,000	500	250
Maximum Throughput/Volume	250 MiB/s	1,000 MiB/s	500 MiB/s	250 MiB/s

Amazon EBS volume type use cases

Solid State Drives (SSD)		Hard Disk Drives (HDD)	
General Purpose	**Provisioned IOPS**	**Throughput-Optimized**	**Cold**
• This type is recommended for most workloads • System boot volumes • Virtual desktops • Low-latency interactive applications • Development and test environments	• Critical business applications that require sustained IOPS performance, or more than 16,000 IOPS or 250 MiB/second of throughput per volume • Large database workloads	• Streaming workloads that require consistent, fast throughput at a low price • Big data • Data warehouses • Log processing • It cannot be a boot volume	• Throughput-oriented storage for large volumes of data that is infrequently accessed • Scenarios where the lowest storage cost is important • 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.

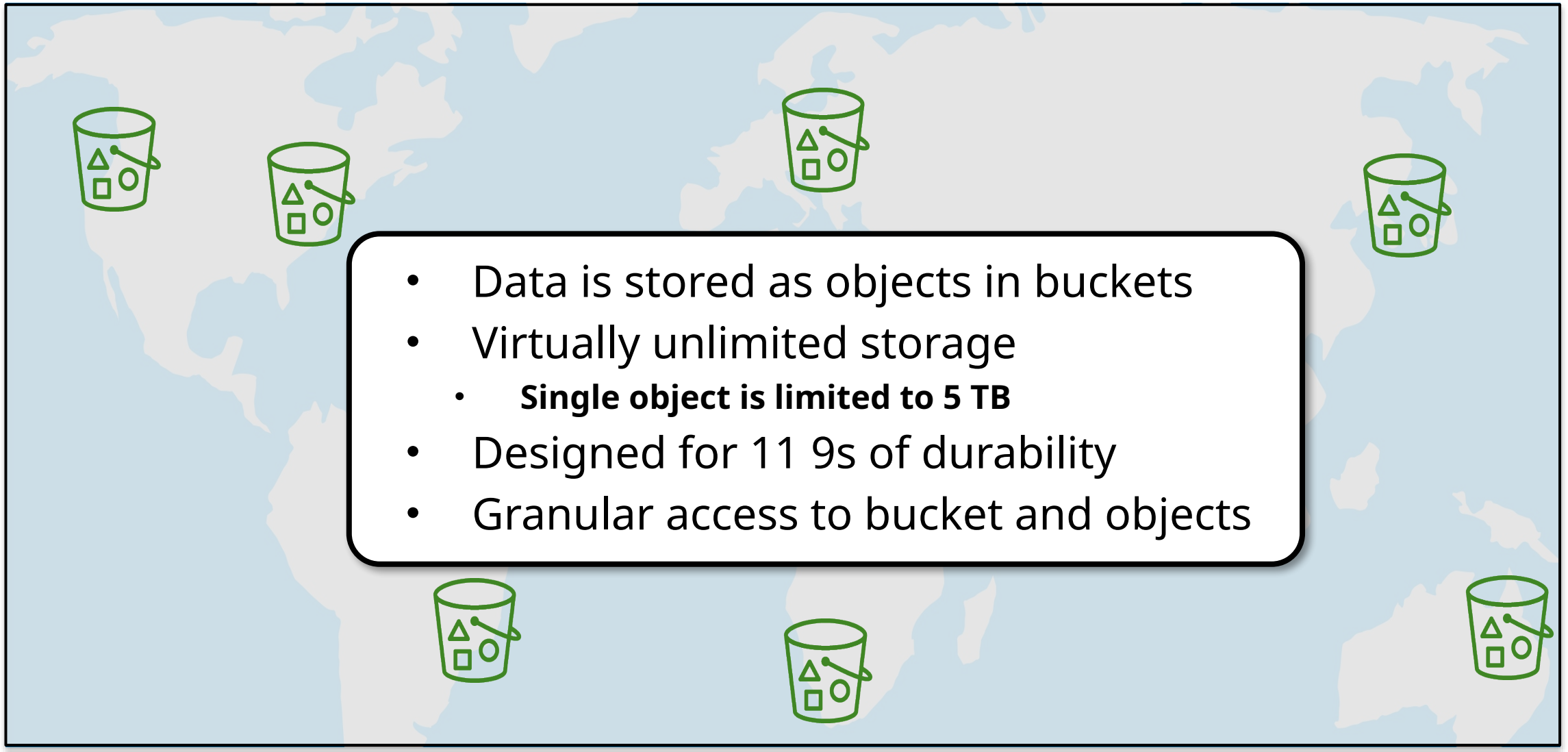
Amazon Simple Storage Service (Amazon S3)

Amazon S3

- object-level storage
  - if you want to change a part of a file, you must make the change and then re-upload the entire modified file
  - stores data as objects within resources that are called **buckets**

Amazon Simple Storage Service (Amazon S3)

Amazon S3 overview



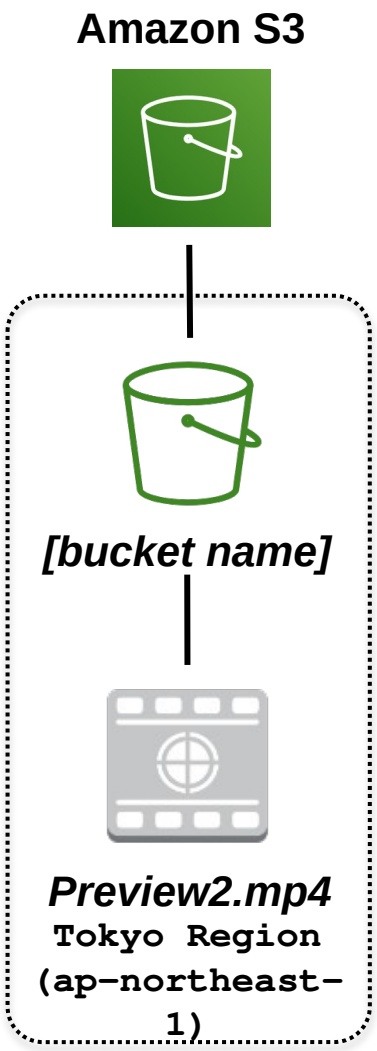
- Data is stored as objects in buckets
- Virtually unlimited storage
  - **Single object is limited to 5 TB**
- Designed for 11 9s of durability
- Granular access to bucket and objects

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:

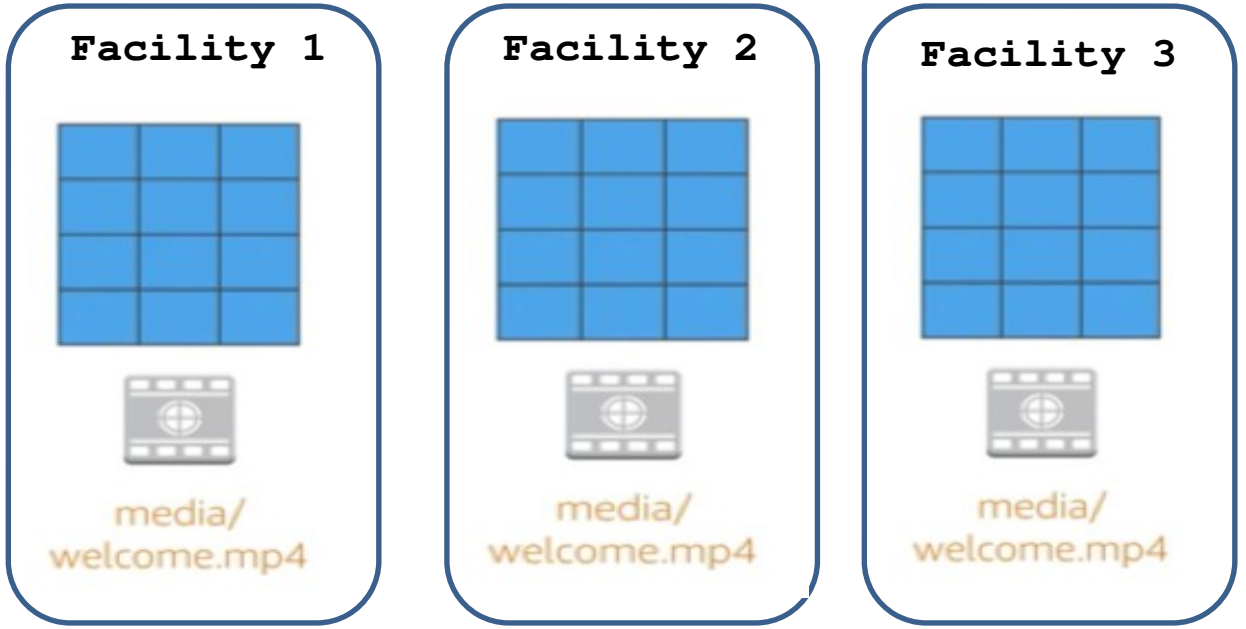
https:// bucket-name.s3-ap-northeast-1.amazonaws.com

Bucket name — Region code

Data is redundantly stored in the Region

media/welcome.mp4

my-bucket-name



Designed for seamless scaling

my-bucket-name

media/welcome.mp4

prod2.mp4

prod3.mp4

prod4.mp4

prod5.mp4

prod6.mp4

prod7.mp4

prod8.mp4

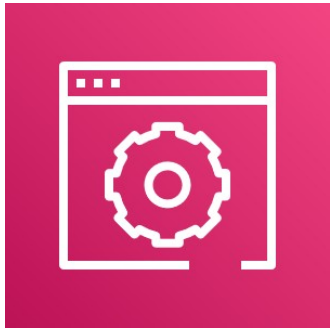
prod9.mp4

prod10.mp4

prod11.mp4

prod12.mp4

Access the data anywhere



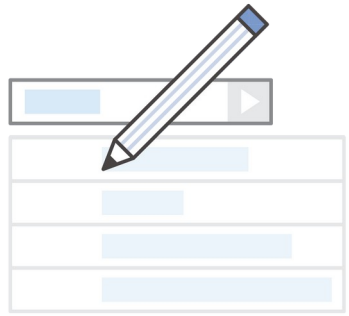
AWS Management Console

AWS Command Line Interface

SDK

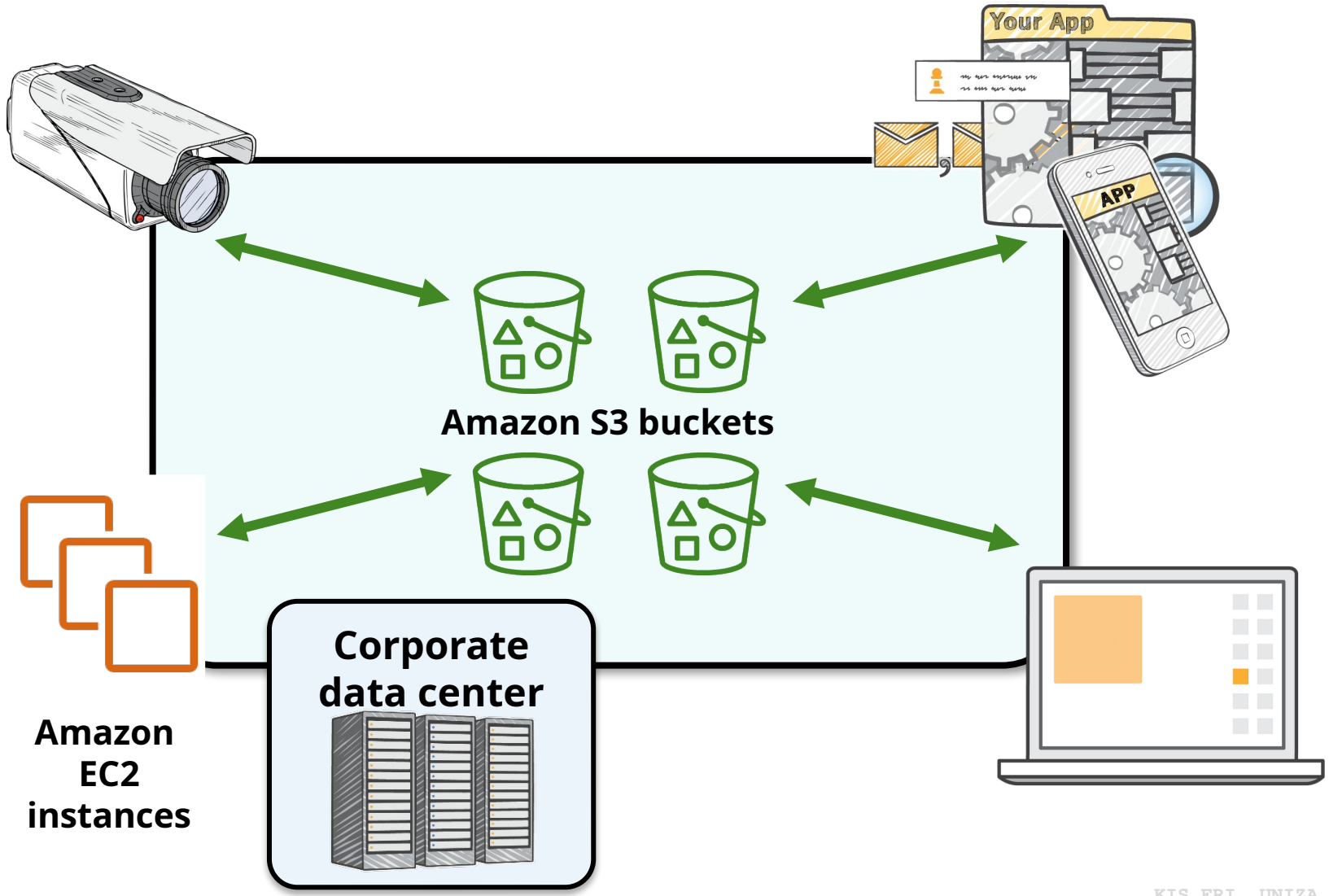
Common use cases

- 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

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

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.

Amazon Elastic File System (Amazon EFS)

Amazon EFS

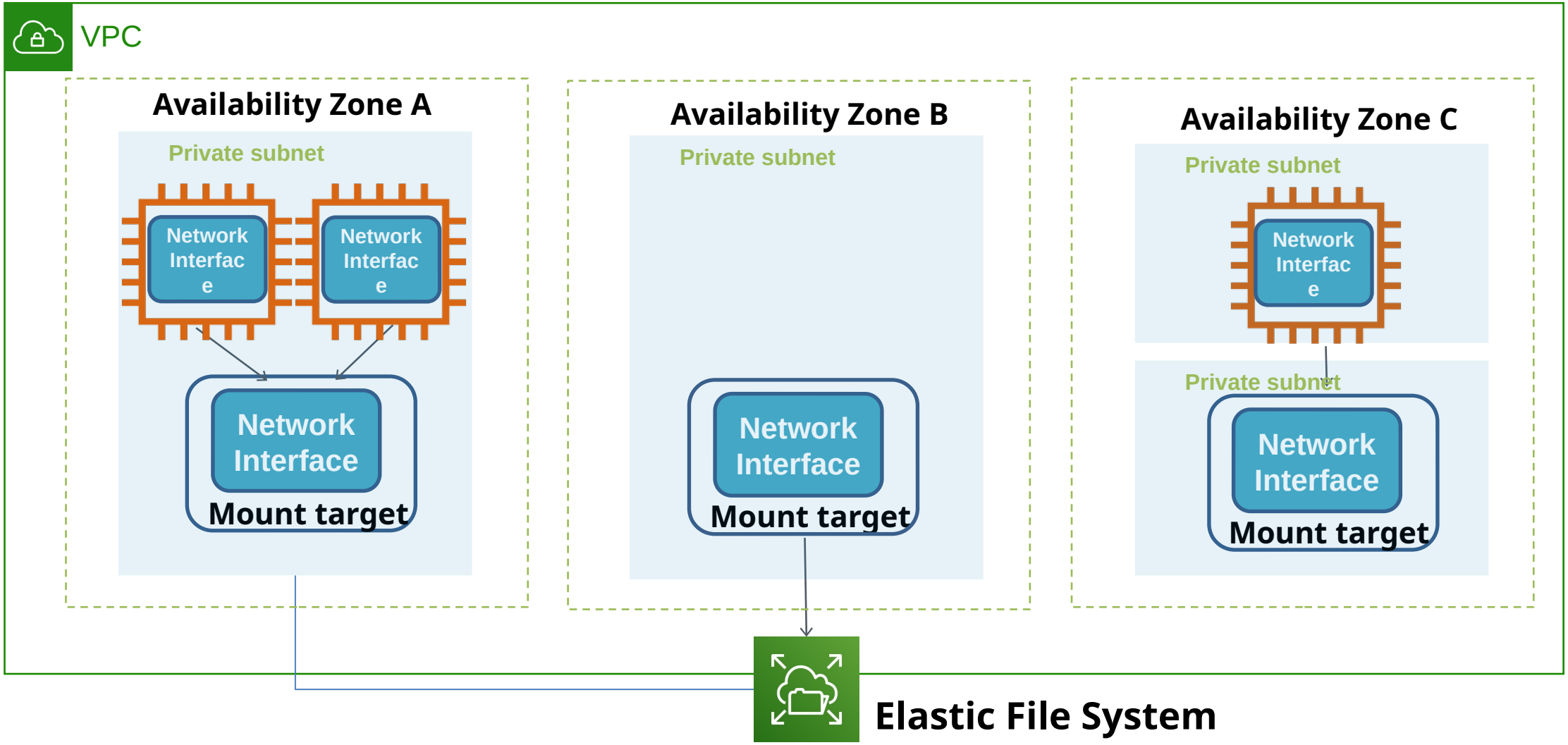
- simple, scalable, elastic file storage
- simple interface that enables you to create and configure file systems quickly and easily
- built to dynamically scale on demand without disrupting applications
- it will grow and shrink automatically as you add and remove files

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

Amazon S3 Glacier

Amazon S3 Glacier

- data archiving service that is designed for security, durability, and an extremely low cost
- pricing is based on Region
- works well for long-term archiving

Amazon S3 Glacier

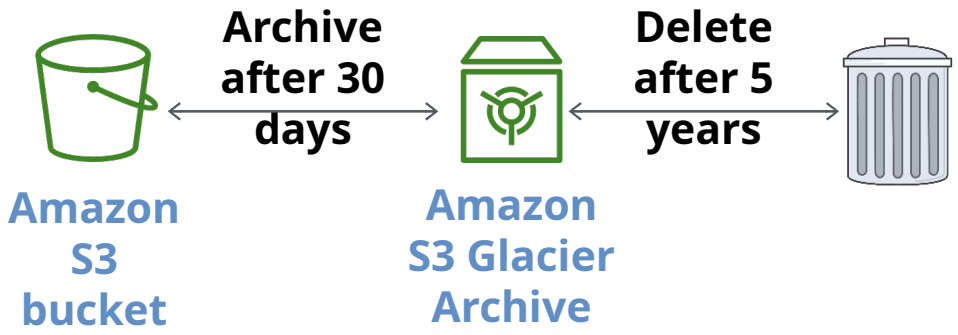
Amazon S3 Glacier review

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

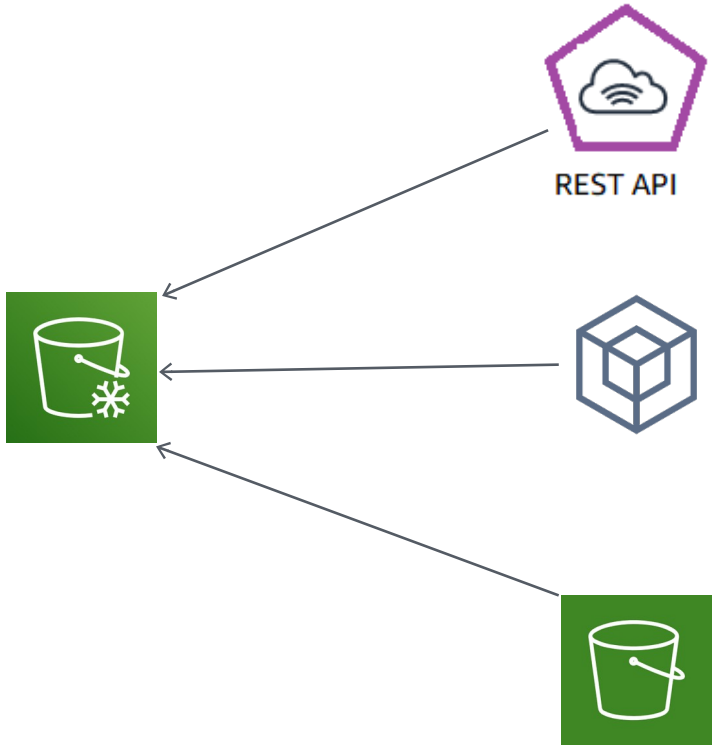

Regulatory and compliance archiving


Scientific data archiving


Digital preservation


Magnetic tape replacement

Using Amazon S3 Glacier



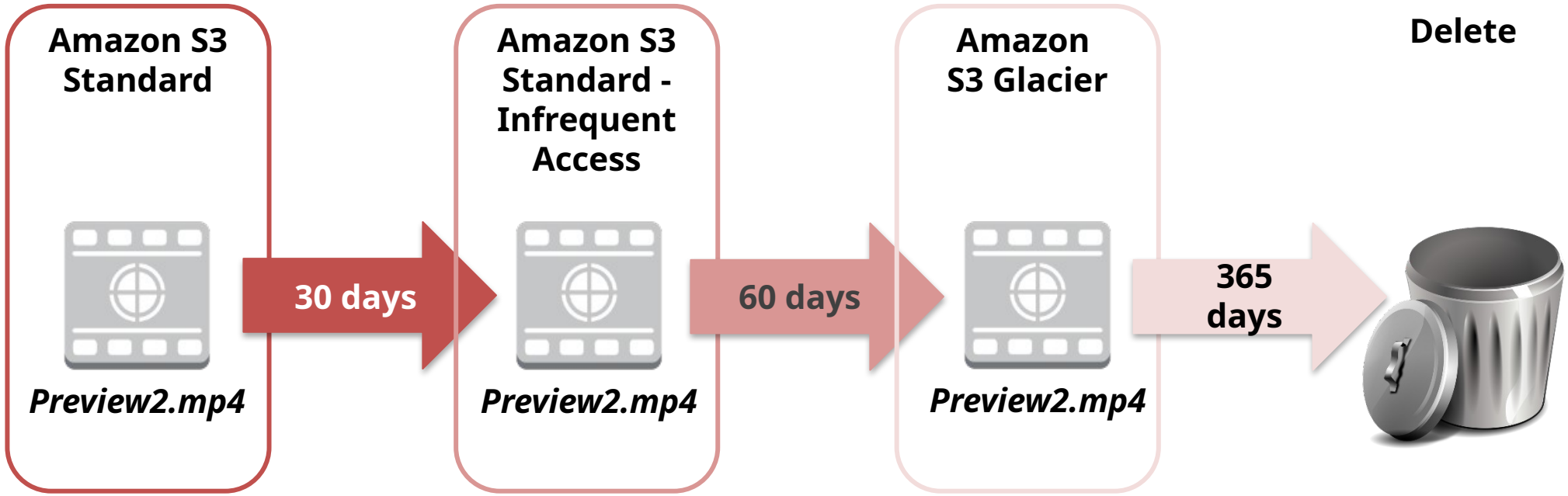
RESTful web services

Java or .NET SDKs

Amazon S3 with lifecycle policies

Lifecycle policies

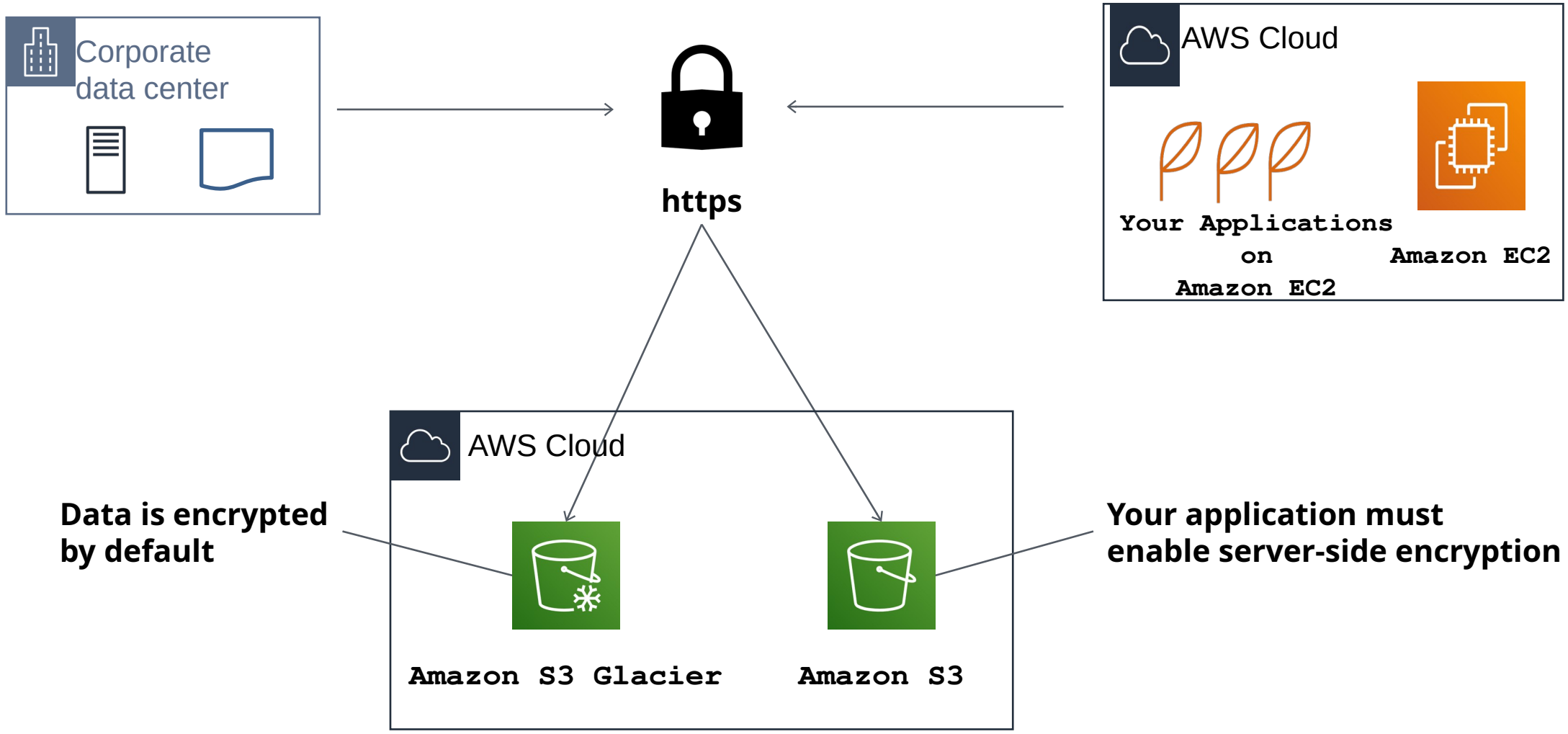
Amazon S3 lifecycle policies enable you to delete or move objects based on age.



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



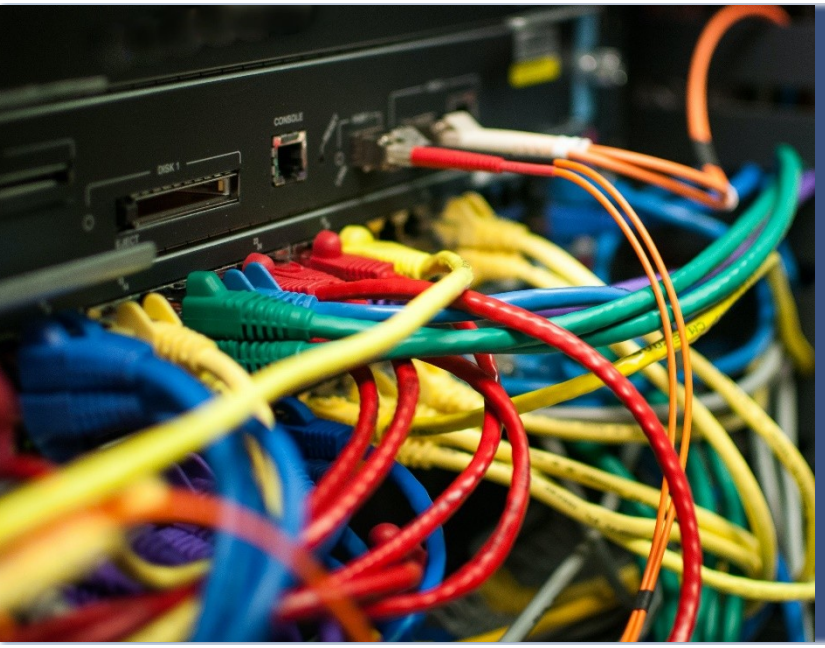
Security with Amazon S3 Glacier

Amazon S3 Glacier

Control access with IAM

Amazon S3 Glacier encrypts your data with AES-256

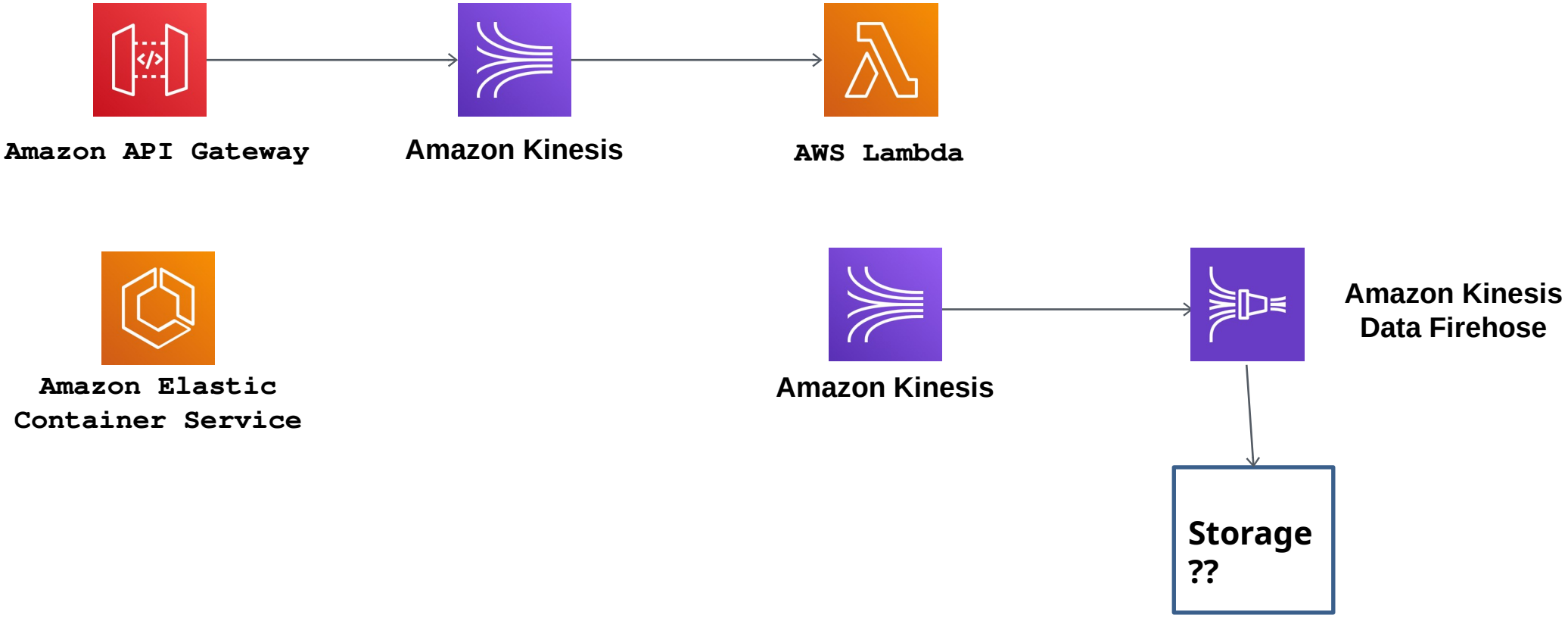
Amazon S3 Glacier manages your keys for you



Activity: Storage Case Studies

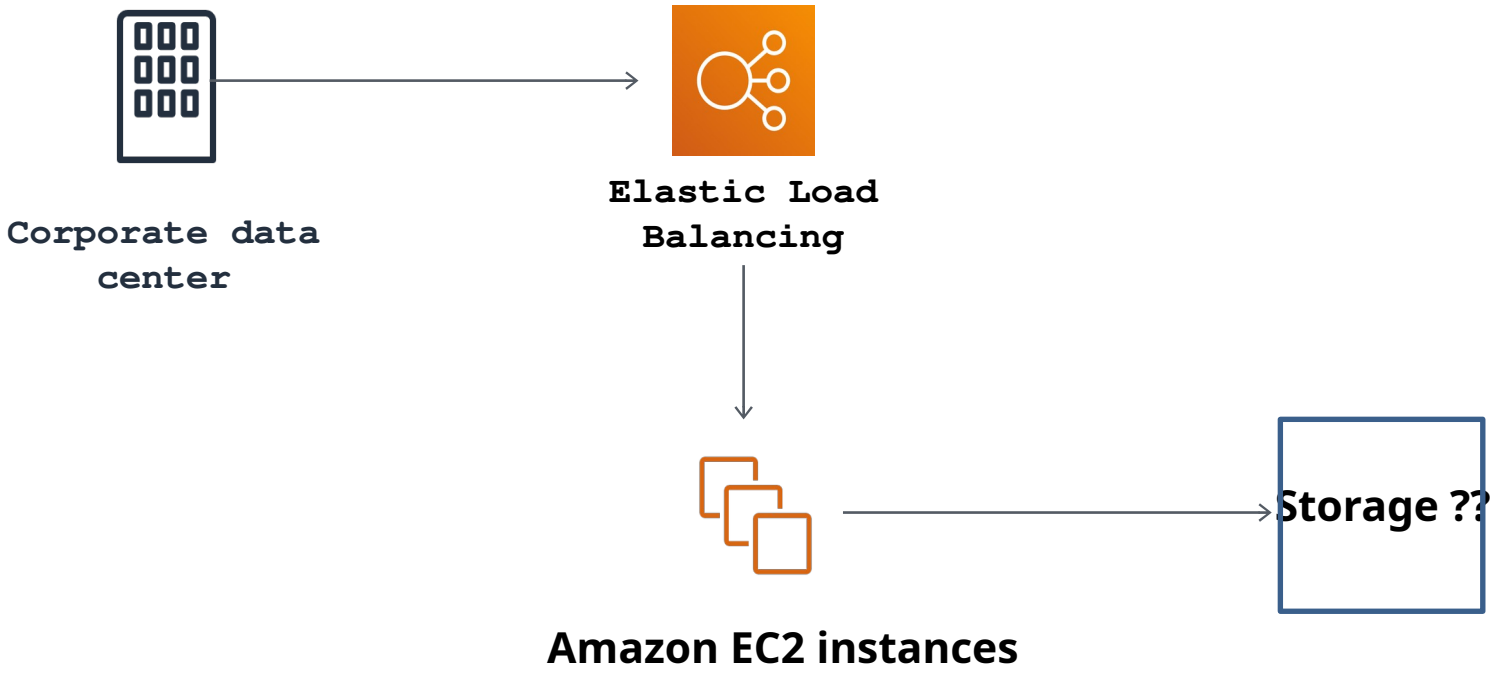
Storage case study activity

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.



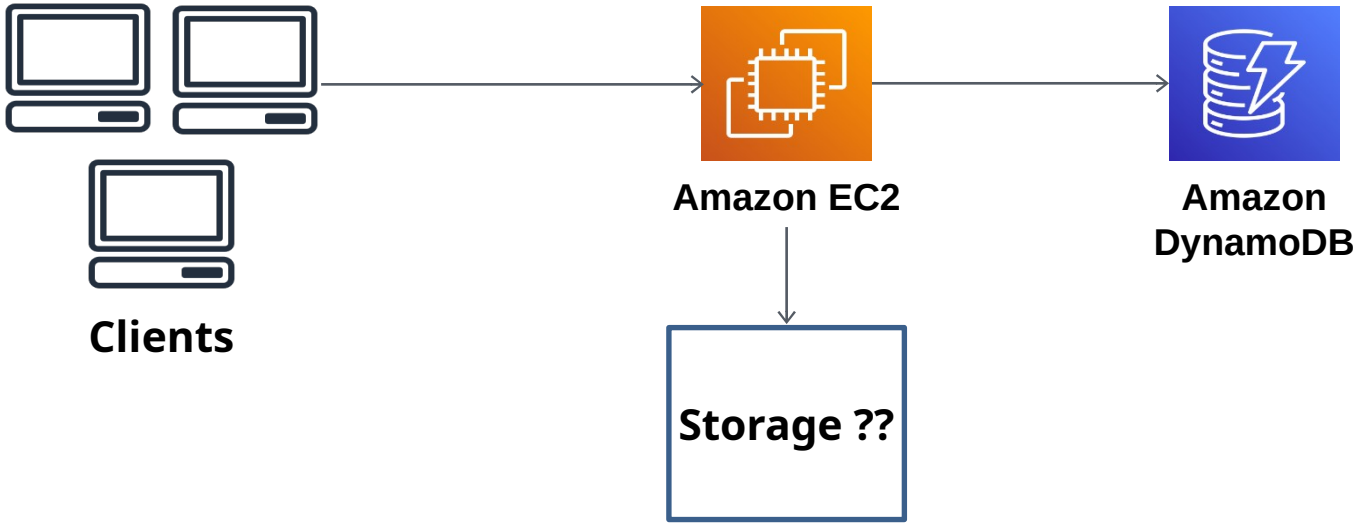
Storage case study activity

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

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 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

Additional resources

- AWS Storage page
- Storage Overview
- Recovering files from an Amazon EBS volume backup
- Confused by AWS Storage Options? S3, EFS, EBS Explained

UNIVERSITY OF ŽILINA
Faculty of Management Science and Informatics

MINISTERSTVO ŠKOLSTVA, VEDY, VÝSKUMU A ŠPORTU SLOVENSKEJ REPUBLIKY

Thank you for your attention.

The content was chapter from AWS Foundations Modules

- AWS M7 - Storage

Vytvorené v rámci projektu **KEGA 026TUKE-4/2021**