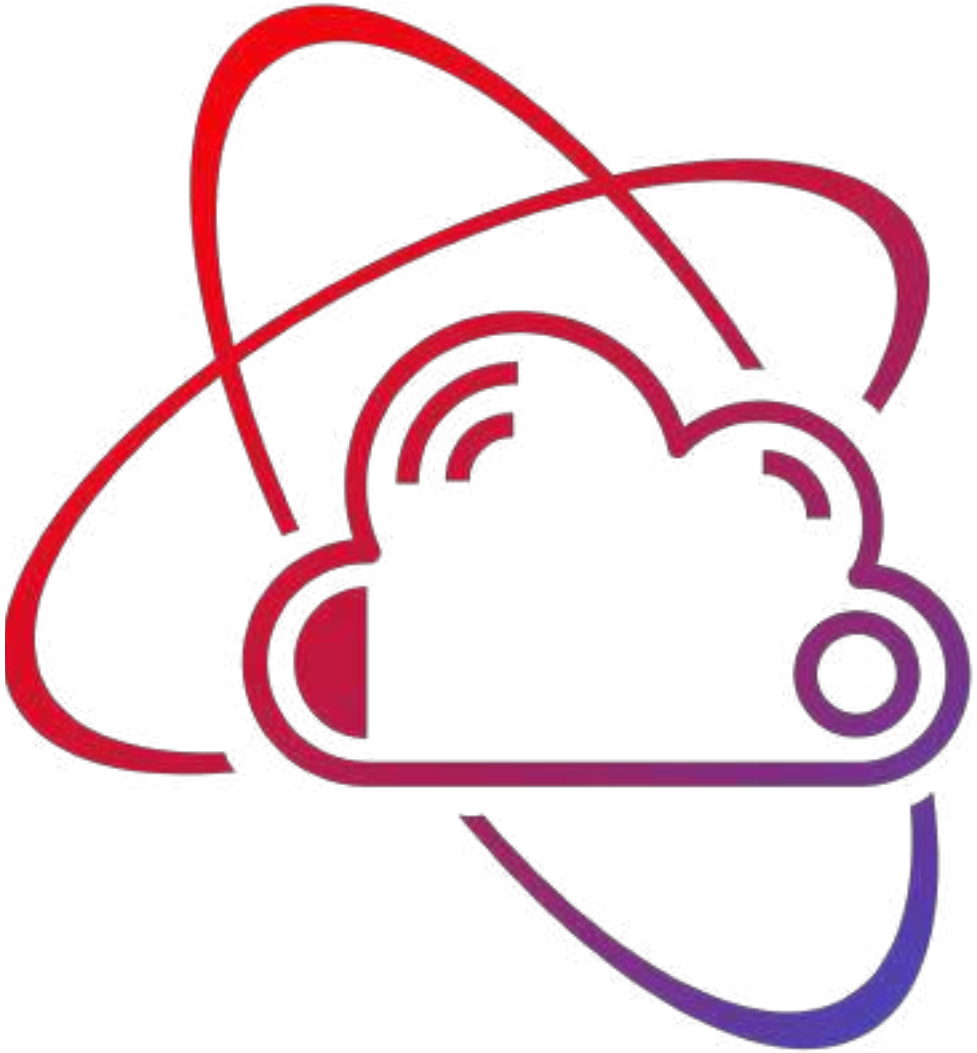
Cloud Computing Architecture

Week 4 – Introduction



SWINBURNE UNIVERSITY OF TECHNOLOGY

Acknowledgement of Country

We respectfully acknowledge the Wurundjeri People of the Kulin Nation, who are the Traditional Owners of the land on which Swinburne's Australian campuses are located in Melbourne's east and outer-east, and pay our respect to their Elders past, present and emerging.

We are honoured to recognise our connection to Wurundjeri Country, history, culture, and spirituality through these locations, and strive to ensure that we operate in a manner that respects and honours the Elders and Ancestors of these lands.

We also respectfully acknowledge Swinburne's Aboriginal and Torres Strait Islander staff, students, alumni, partners and visitors.

We also acknowledge and respect the Traditional Owners of lands across Australia, their Elders, Ancestors, cultures, and heritage, and recognise the continuing sovereignties of all Aboriginal and Torres Strait Islander Nations.

In this Presentation:

- AWS Storage Services Overview

- Introduction to Amazon S3

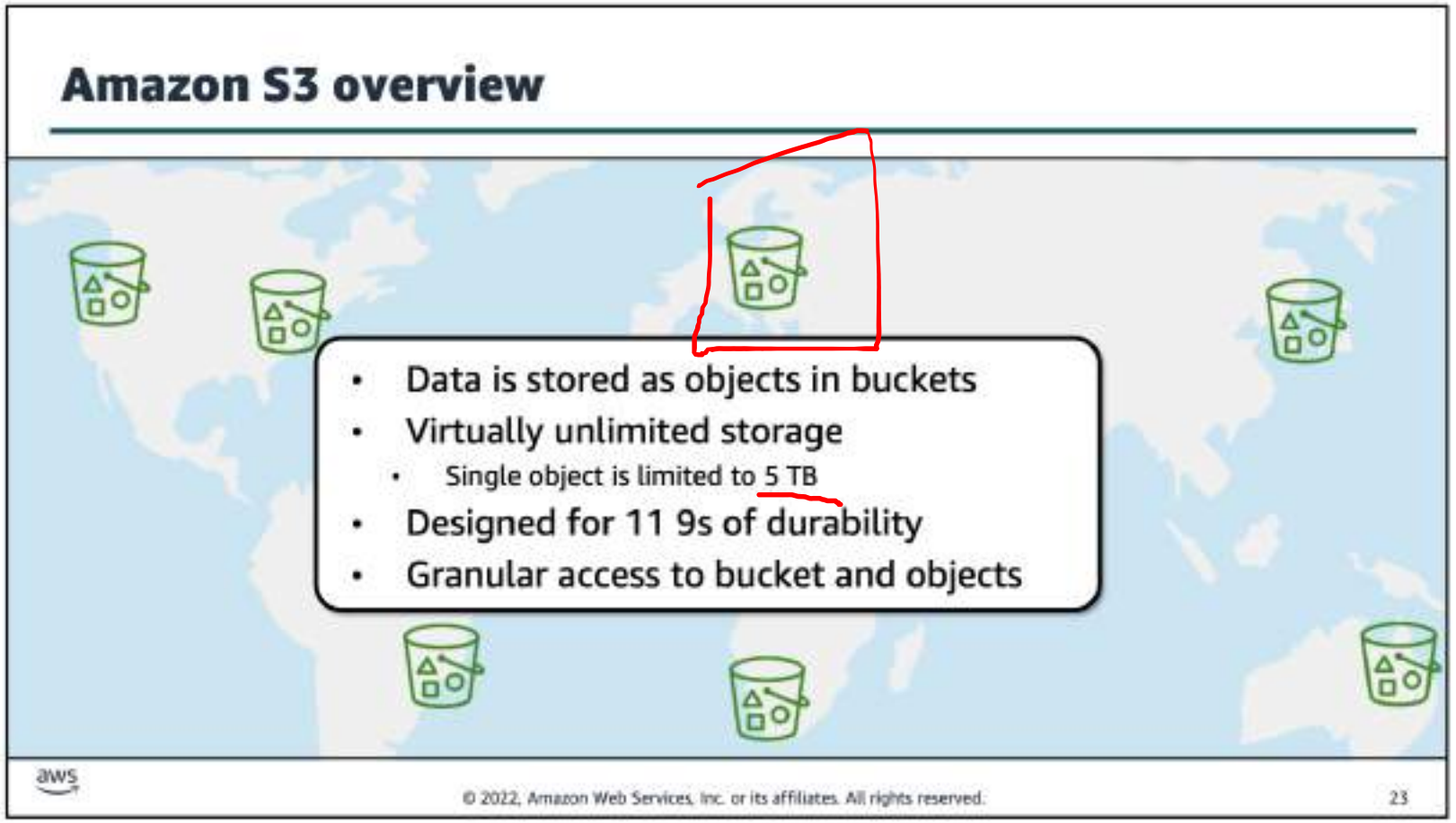
Image from: https://aws.plainenglish.io/optimize-your-aws-s3-performance-27b057f231a3

AWS Storage Services Overview

AWS Storage Services

Image from: https://blog.jineshkumar.com/aws-cloud-storage-services

Introduction to 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
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.
23

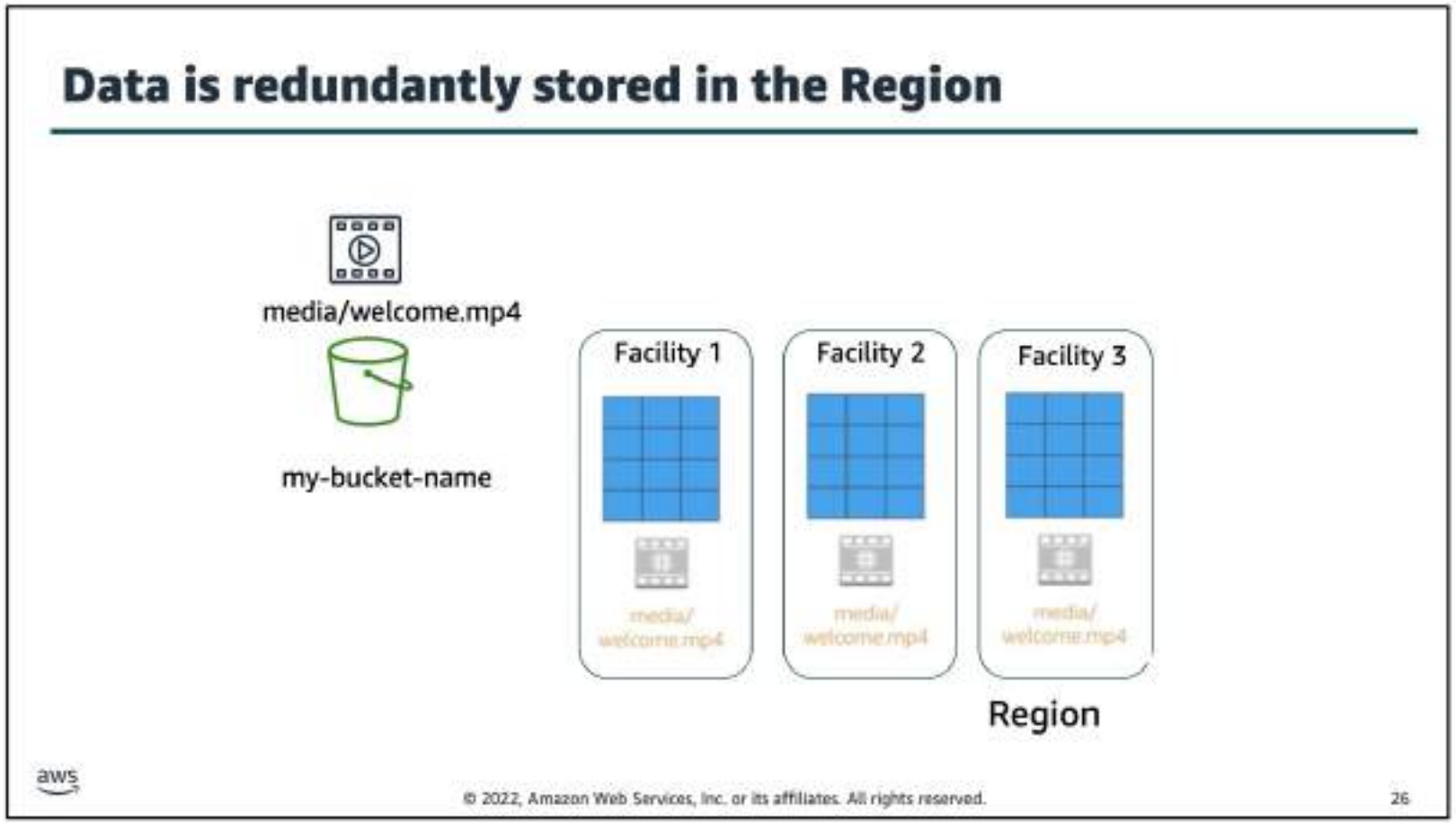
Amazon S3 storage classes

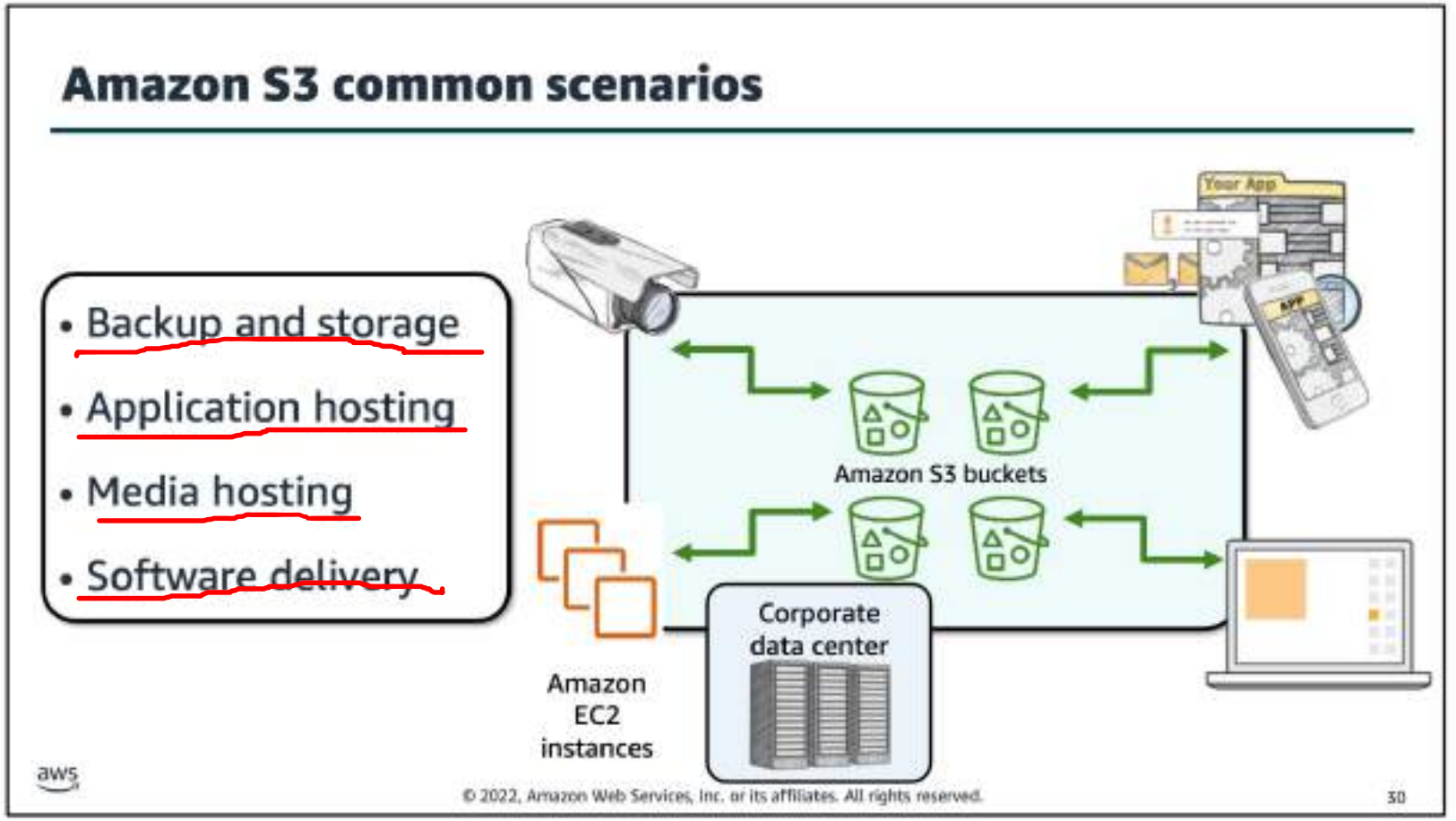
Amazon S3 offers a range of object-level storage classes that are designed for different use cases:

- 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

aws

24


Data is redundantly stored in the Region
media/welcome.mp4
my-bucket-name
Facility 1
Facility 2
Facility 3
media/
welcome.mp4
Region
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.
26

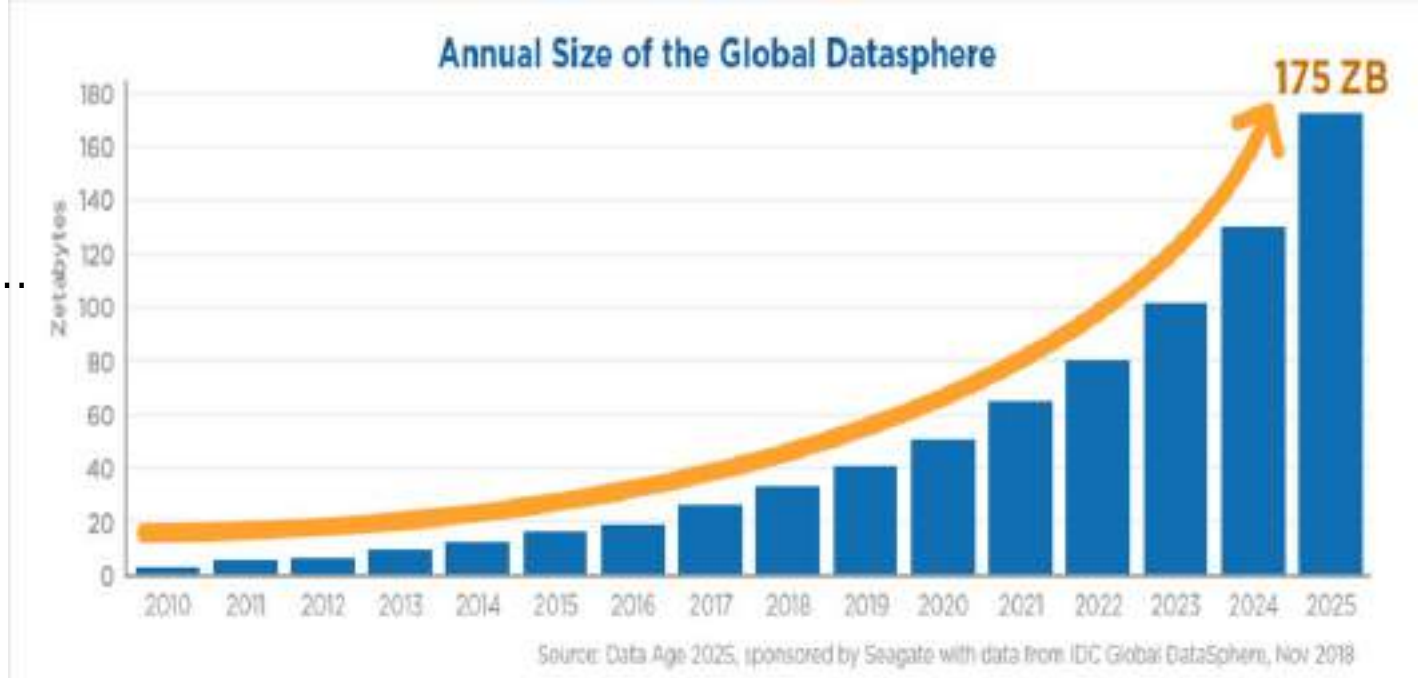

Amazon S3 common scenarios
• Backup and storage
• Application hosting
• Media hosting
• Software delivery
Your App
Amazon S3 buckets
Amazon EC2 instances
Corporate data center
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.
30

This week

- **Storage in the Cloud**
  - **Big Data**
  - **NFS**
  - **Distributed File Systems (databases next week)**
- **AWS Storage services (ACF Module 7)**
  - Amazon Elastic Block Store (Amazon EBS)
    - Plus some extra notes on Instance Storage
  - Amazon Elastic File System (Amazon EFS)
  - Amazon Simple Storage Service (Amazon S3)
  - Amazon Glacier

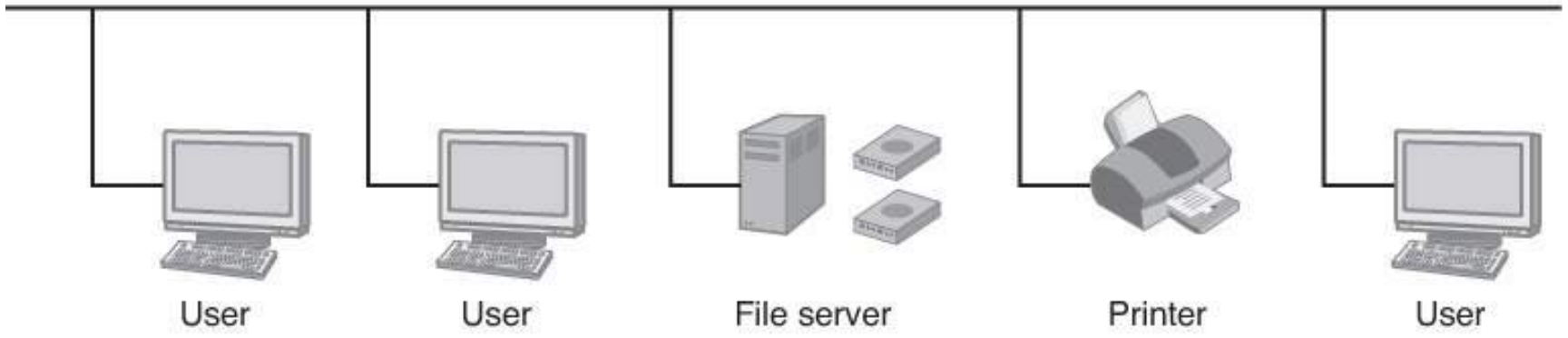
Big Data – Centre of the Universe

- Drivers
  - Internet commerce, Mobile, Social media, IoT and sensors, Science (e.g. biology, astronomy, meteorology), Health, Spooks, …
- Data sizes
  - KB (10^3), MB (10^6), GB (10^9), TB (10^{12}), PB(10^{15}), exabyte (EB, 10^{18}), zettabyte (ZB, 10^{21}), yottabyte (YB,10^{24})
  - Everyday: Facebook 10T, Twitter 7T, Youtube 4.5T+
- 4Vs: volume, variety, velocity, and veracity



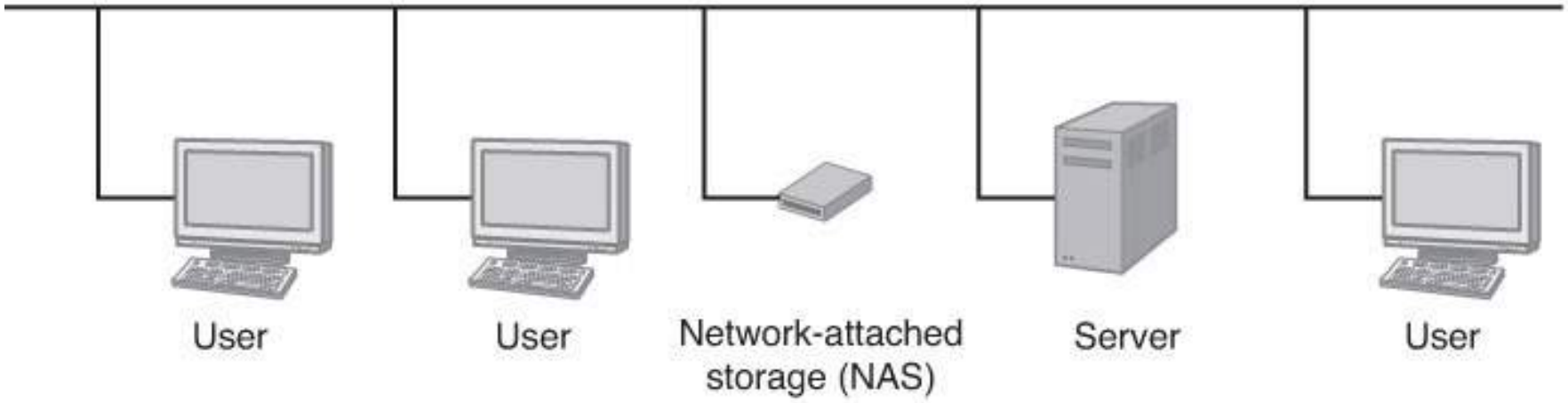
Network Storage Began with File Servers

Years ago, local-area networks used special servers, called file servers, to support file sharing, file replication, and storage for large files.



Network-Attached Storage (NAS)

Plug directly into the network.



Advantages of NAS

- **Reliability**: A NAS device typically provides advanced data striping across multiple volumes within the device. If one (or more) volumes fail, the data striping would maintain the data and allow reconstruction of the file contents.
- **Performance**: Because the NAS device did not run a complete operating system, the hardware had less system overhead, which allowed it to outperform a file server.
- **Compatibility**: NAS devices normally support common file systems, which, in turn, make them fully compatible with common operating systems.
- **Ease of performing backups**: NAS devices are commonly used for backup devices. Within a home, for example, all devices can easily access and back up files to a NAS device.

Cloud-Based Storage

- Cloud-based data storage is the next step in the evolution of NAS devices.
- Across the web (the cloud), many providers offer data storage that resides in the cloud.
- Data may be accessible as follows:
  - Through a web browser interface
  - Through a mounted disk drive
  - Through a set of API (application program interface) calls

Advantages of Cloud-Based Storage

- **Scalability**: Most cloud-based data storage providers let users scale their storage capacity (up or down) to align with their storage needs.
- **Pay for use**: With most cloud-based data storage facilities, users pay only for the storage (within a range) that they need.
- **Reliability**: Many cloud-based data storage facilities provide transparent data replication.
- **Ease of access**: Most cloud-based data storage facilities support web-based access to files from any place, at any time, using a variety of devices.
- **Ease of use**: Many cloud-based data storage solutions let users map a drive letter to the remote file storage area and then access the files through the use of a logical drive.

Disadvantages of Cloud-Based Storage

- **Performance**: Because the cloud-based disk storage devices are accessed over the Internet, they will never be as fast as local drives.
- **Security**: Some users will never feel comfortable with their data in the cloud.
- **Data orphans**: Users may abandon data in cloud storage facilities, leaving confidential private or company data at risk.

Cloud-Based Block Storage

- In the simplest sense, a block of data storage is a fixed-sized sequence of bits. The size of the block normally corresponds to an underlying unit of storage on the **cloud-based block storage device**.
- Some applications work with very large blocks of data, the format of which has meaning only to the application itself—meaning that the data may not map well to storage within a file system or database.

File Systems

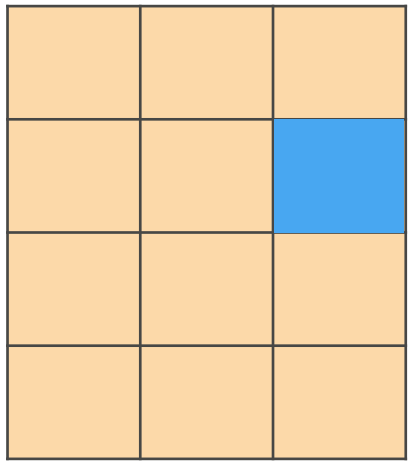
- Operating systems exist to allow users to run programs and to store and retrieve data (files) from one user session to the next.
- Within the operating system, special software, called the file system, oversees the storage and retrieval of files to and from a disk.
- When you copy a file, delete a file, or create and move files between folders, the file system is performing the work.
- Initially, file systems allowed users to manipulate only local files that reside on one of the PC's disk drives.
- As networks became more prevalent, so too did network operating systems, which allow users and programs to manipulate files residing on a device across the network.

Real World: Hadoop Distributed File System

- Apache Hadoop is an open source project, the goal of which is to support reliable, scalable distributed computing.
- Part of the project includes the Hadoop Distributed File System (HDFS), a Java-based file system that is well suited for cloud-based storage.
- HDFS is designed to be highly fault tolerant and robust to maintain operation in the event of a device failure.

AWS Storage Options: Block vs. Object Storage

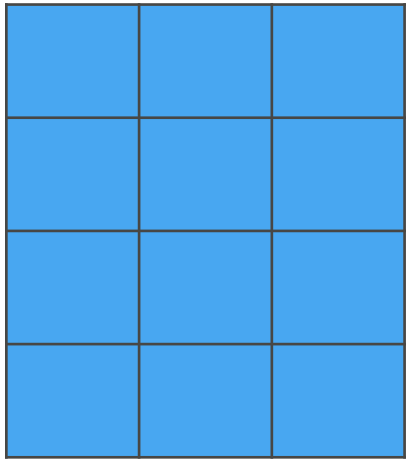
What if you want to **change** one character in a 1-GB file?



Block Storage

Change one block (piece of the file)

that contains the character



Object Storage

Entire file must be updated

This week

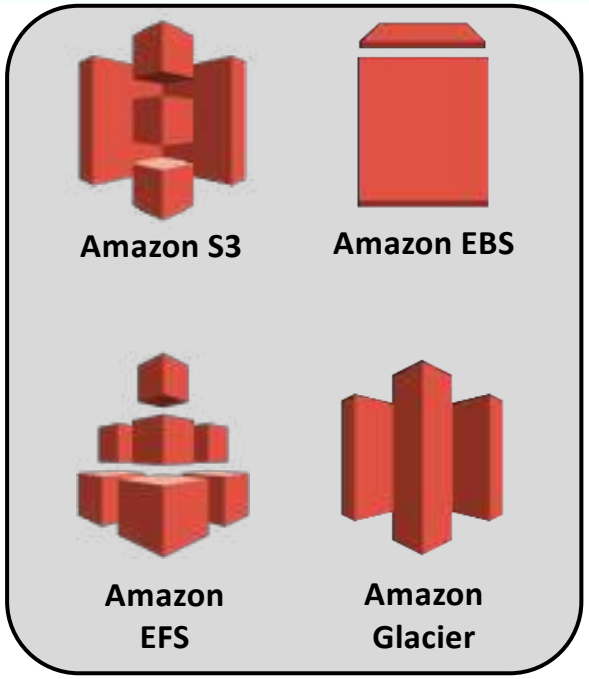
- Storage in the Cloud
  - Big Data
  - NFS
  - Distributed File Systems (databases next week)
- **AWS Storage services (ACF Module 7)**
  - Amazon Elastic Block Store (Amazon EBS)
    - Plus some extra notes on Instance Storage
  - Amazon Elastic File System (Amazon EFS)
  - Amazon Simple Storage Service (Amazon S3)
  - Amazon Glacier

Introduction to Storage Services

Core AWS Services

Amazon VPC

Amazon EC2



Storage

Database

AWS IAM

Amazon Elastic Block Store (Amazon EBS)

Storage

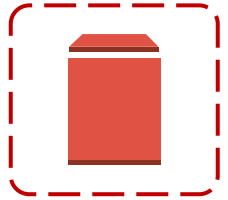
Amazon Elastic Block Store (Amazon EBS)

Amazon EBS Review

*Amazon EBS allows 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
- Can be backed up automatically to Amazon S3
- Uses:
  - Boot volumes and storage for Amazon EC2 instances
  - Data storage with a file system
  - Database hosts
  - Enterprise applications

Amazon EBS Lifecycle



Amazon EBS Volume Types

	Solid-State Drives (SSD)		Hard Disk Drives (HDD)	
	General Purpose	Provisioned IOPS	Throughput-Optimized	Cold
Max volume size	16 TiB	16 TiB	16 TiB	16 TiB
Max IOPS/volume	10,000	32,000	500	250
Max throughput/volume	160 MiB/s	500 MiB/s	500 MiB/s	250 MiB/s

Amazon EBS Volume Types

	Solid-State Drives (SSD)		Hard Disk Drives (HDD)	
	General Purpose	**Provisioned IOPS**	**Throughput-Optimized**	**Cold**
Use Cases	• Recommended for most workloads • System boot volumes • Virtual desktops • Low-latency interactive apps • Development and test environments	• I/O-intensive workloads • Relational DBs • NoSQL DBs	• Streaming workloads requiring consistent, fast throughput at a low price • Big data • Data warehouses • Log processing • 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 • Cannot be a boot volume

Amazon EBS

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 and IOPS

1. **Volumes**
   - Amazon EBS volumes persist independently from the instance
   - All volume types are charged by the amount provisioned per month
2. **Input Output Operations per Second (IOPS)**
   - General Purpose (SSD)
     - Charged by the amount your provision in GB per month until storage is released
   - Magnetic
     - Charged by the number of requests to volume
   - Provisioned IOPS (SSD)
     - Charged by the amount you provision in IOPS (by % of day / month used)

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 charges are tiered

In Review

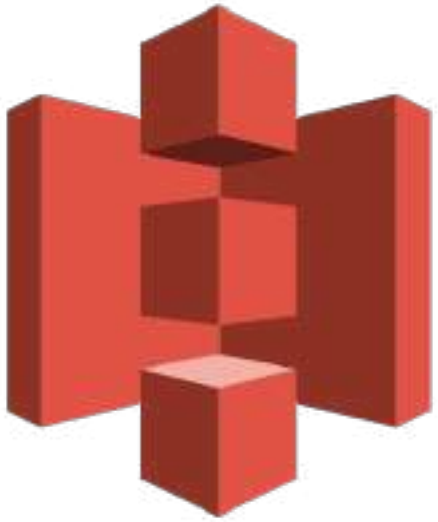
Amazon EBS Features:

- 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 using snapshots

This week

- Storage in the Cloud
  - Big Data
  - NFS
  - Distributed File Systems (databases next week)
- AWS Storage services (ACF Module 7)
  - Amazon Elastic Block Store (Amazon EBS)
    - Plus some extra notes on Instance Storage
  - Amazon Elastic File System (Amazon EFS)
  - Amazon Simple Storage Service (Amazon S3)
  - Amazon Glacier

Part 3:
Amazon Simple Storage Service (S3)



Amazon Simple Storage Service (Amazon S3)

Amazon S3 Review

Managed cloud storage solution** designed to **scale seamlessly and** provide **99.999999999% durability.

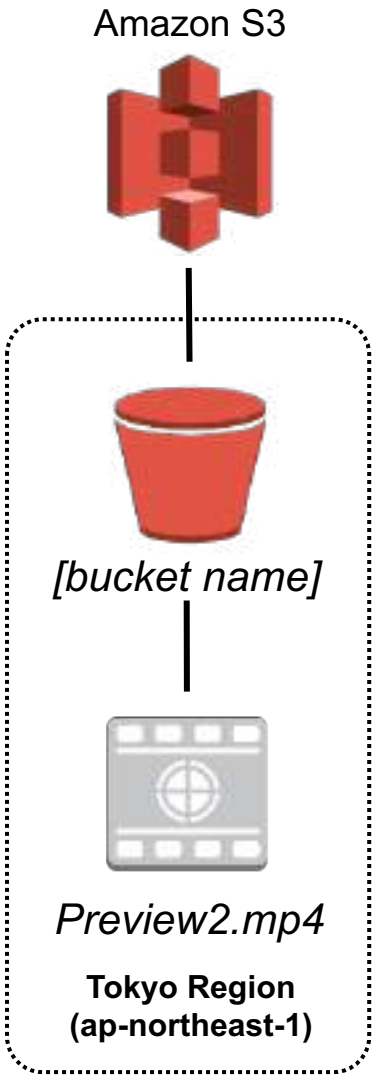
- Store as many objects as you want.
- Bucket names must be unique across all existing bucket names in Amazon S3.
- Amazon S3 cannot be used as a bootable drive.
- Data is stored redundantly.
- Access Amazon S3 with the AWS Management Console, one of the AWS SDKs, or a third-party solution.
- Object uploads or deletes can trigger notifications, workflows, or even scripts.
- Data in transit and at rest can be encrypted automatically.
- Storage class analysis (Amazon S3 Analytics) to analyze storage access patterns and transition the right data to the right storage class.

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 Review



To upload your data (photos, videos, documents, etc.):

1. Create a bucket in one of the AWS Regions.
2. Upload any number of objects to the bucket.

Bucket

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

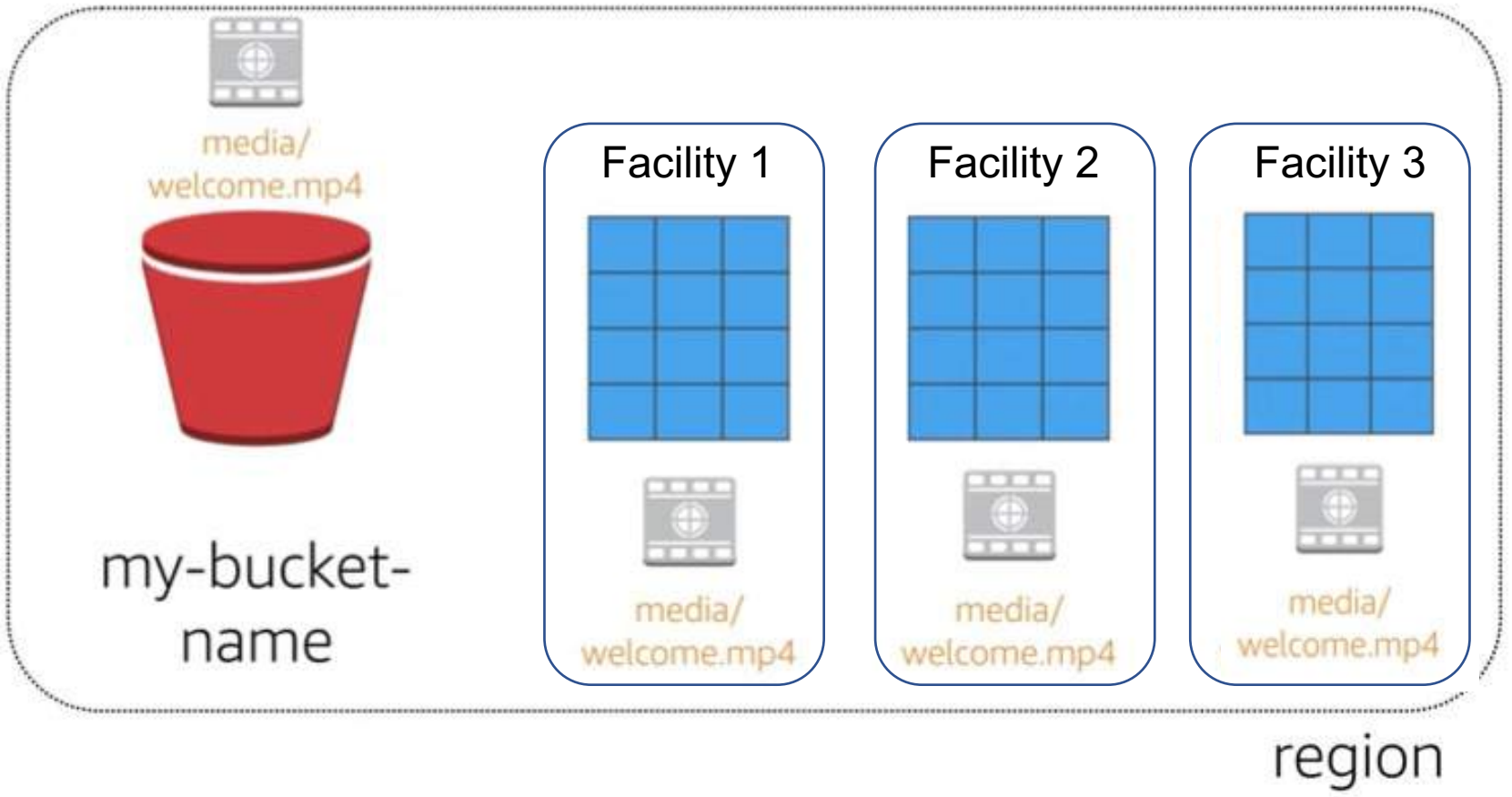
Region code — Bucket name

Object

https://s3-ap-northeast-1.amazonaws.com/[bucket name]/Preview2.mp4

Key

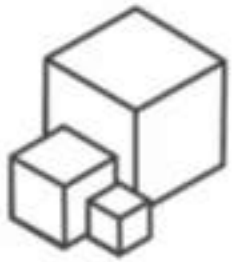
Data Redundantly Stored in Region



Designed for Seamless Scaling

Access the Data Anywhere

AWS Management Console



AWS CLI

AWS SDKs

Common Use Cases

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

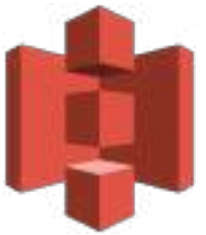
Amazon S3 Pricing

- 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 have to pay for:
  - Transfer IN to Amazon S3
  - Transfer OUT from Amazon S3 to Amazon CloudFront or Amazon EC2 in the same region.

Amazon S3: Storage Pricing

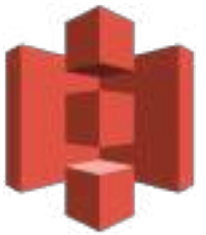
To estimate Amazon S3 costs, consider the following:

1. **Types of storage classes**
   - Standard Storage
     - 99.999999999% durability
     - 99.99% availability
   - Standard-Infrequent Access (SIA)
     - 99.999999999% durability
     - 99.9% availability
2. **Amount of storage**
   - The number and size of objects



Amazon S3: Storage Pricing

3. **Requests:**
   - The number of requests **(GET, PUT, COPY):**
   - Type of requests
     - Different rates for GET requests than other requests
4. **Data Transfer:**
   - Pricing based on the amount of data transferred out of the Amazon S3 region
     - Data transfer in is free, charge for data transfer out



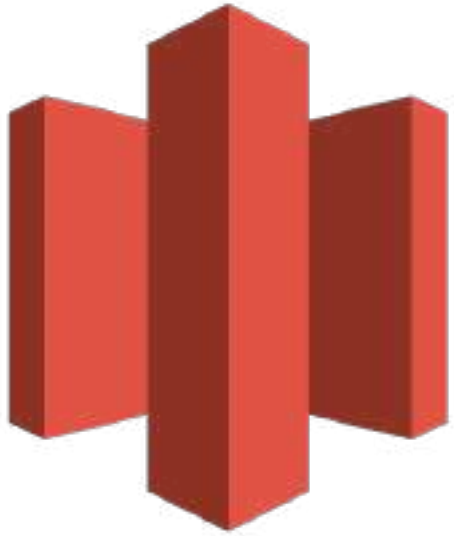
In Review

- Amazon S3 is a fully managed cloud storage service
- Store a virtually unlimited number of objects
- Pay for only what you use
- Access at any time, from anywhere
- Amazon S3 offers rich security controls

This week

- Storage in the Cloud
  - Big Data
  - NFS
  - Distributed File Systems (databases next week)
- AWS Storage services (ACF Module 7)
  - Amazon Elastic Block Store (Amazon EBS)
    - Plus some extra notes on Instance Storage
  - Amazon Elastic File System (Amazon EFS)
  - Amazon Simple Storage Service (Amazon S3)
  - Amazon Glacier

Amazon Glacier



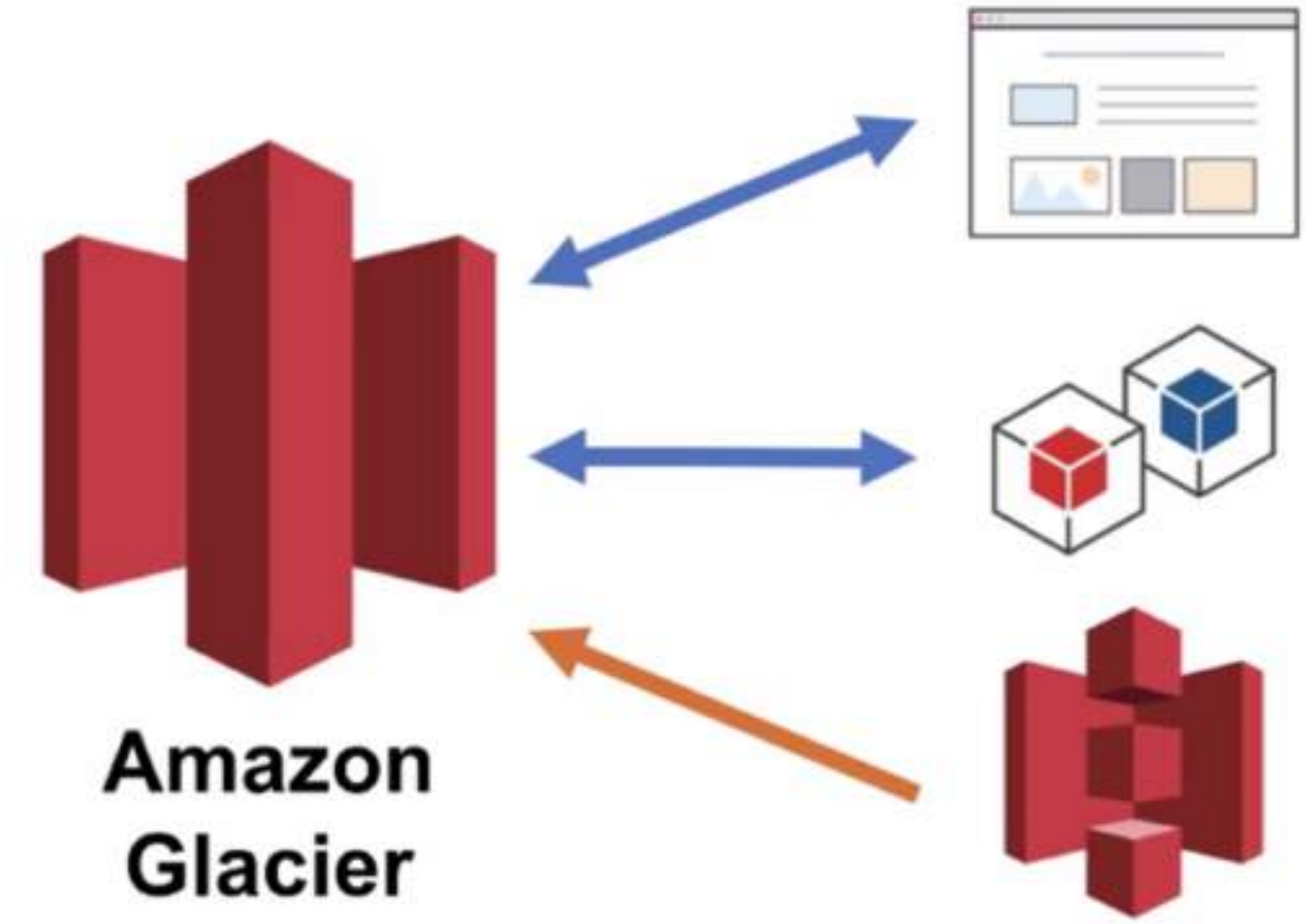
Amazon Glacier

Amazon Glacier Review

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

- Designed for durability of 99.999999999% of objects
- Supports SSL/TLS encryption of data in transit and at rest
- The Vault Lock feature enforces compliance via a lockable policy
- Extremely low-cost design is ideal for long-term archiving
  - Provides three options for access to archives (Expedited, Standard, and Bulk) from a few minutes to several hours

Using Amazon Glacier



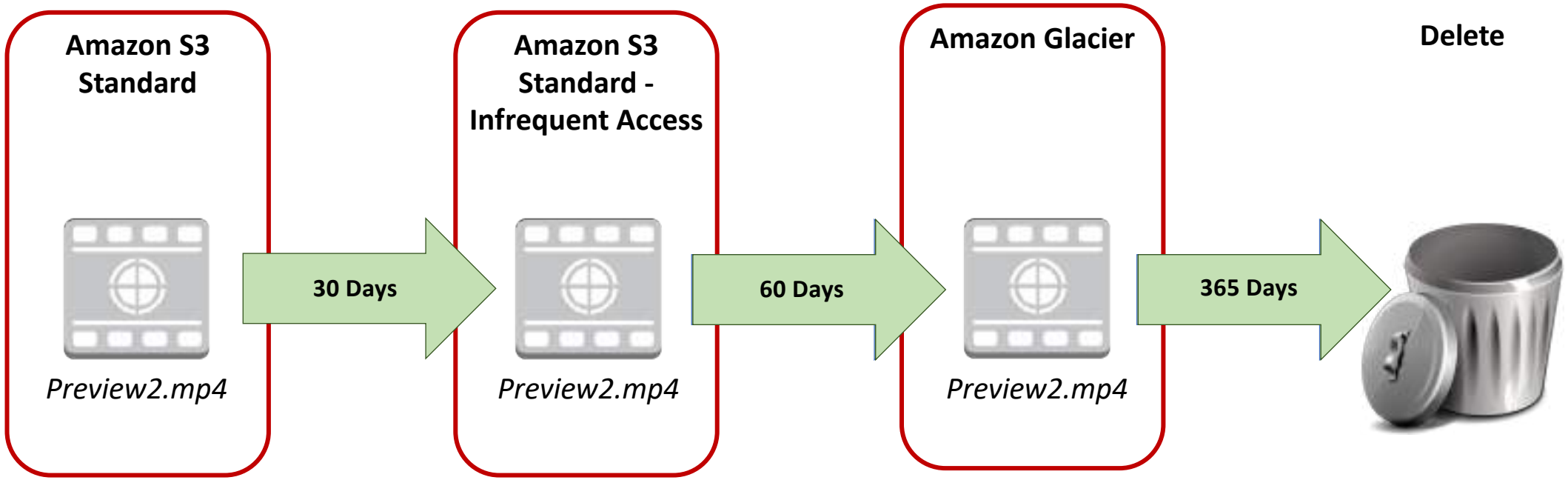
RESTful
Web services

Java or .NET
SDKs

Amazon S3 with
lifecycle policies

Lifecycle Policies

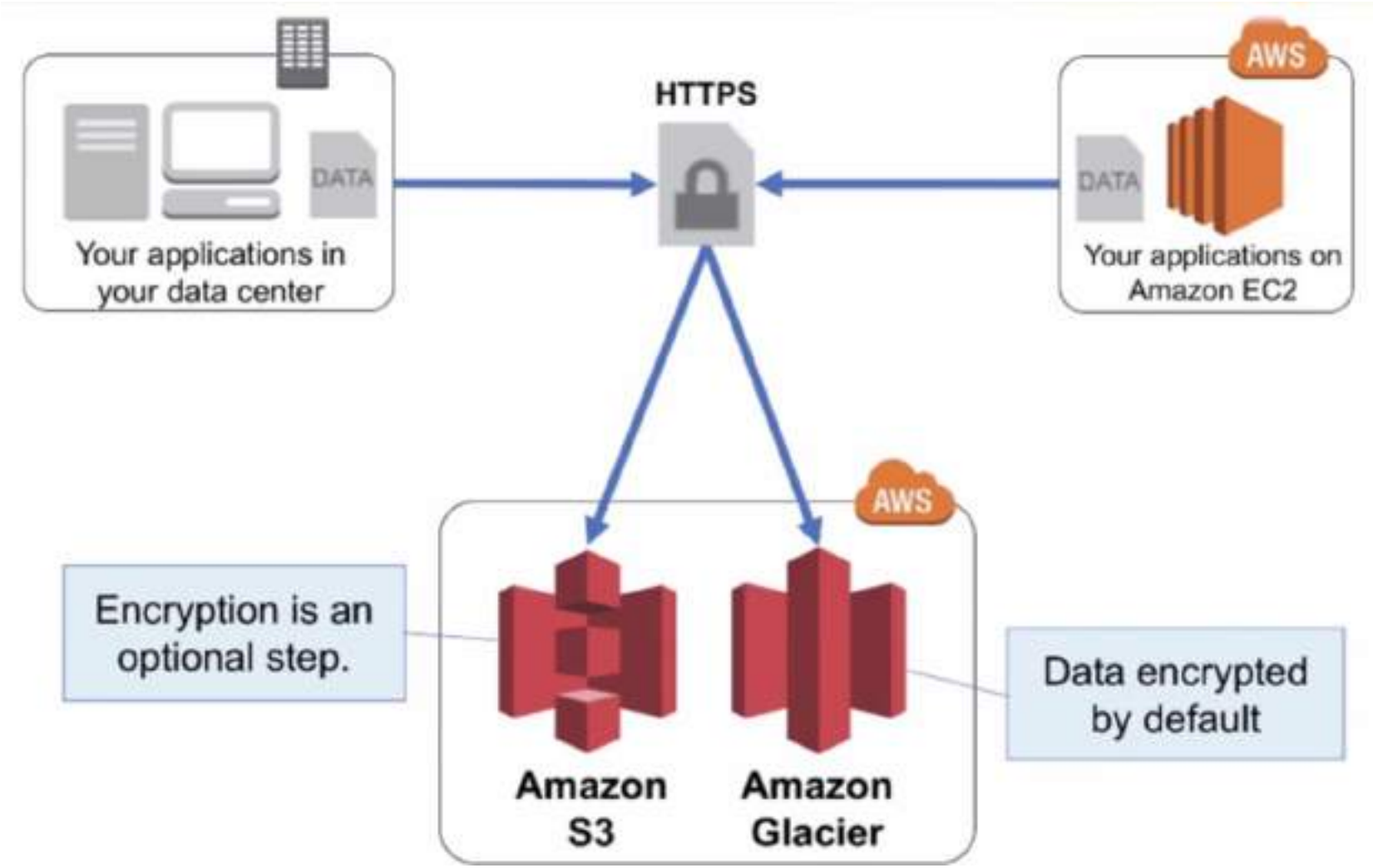
Amazon S3 lifecycle policies allow you to delete or move objects based on age.



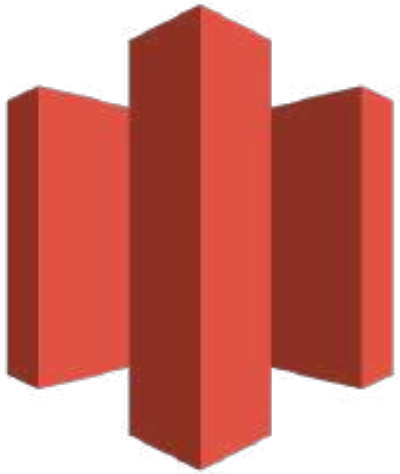
Storage Comparison

	Amazon S3	Amazon Glacier
Data volume	No limit	No limit
Average latency	ms	min/hrs
Item size	5 TB max	40 TB max
Cost/GB per month	¢¢	¢
Billed requests	PUT, COPY, POST, LIST, and GET	UPLOAD and retrieval
Retrieval pricing	¢ Per request	¢¢ Per request and per GB

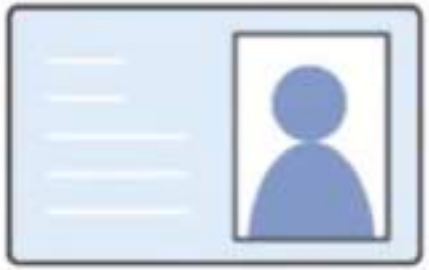
Server-Side Encryption



Security with Amazon Glacier



Amazon Glacier



Control access with AWS IAM

Amazon Glacier encrypts your data with AES-256

Amazon Glacier manages your keys for you

In Review

- Amazon Glacier is a data archiving service designed for security, durability, and an extremely low cost.
- Amazon Glacier pricing is region-based.
- Extremely low-cost design is ideal for long-term archiving.
- The service is designed for durability of 99.999999999% of objects.