

CS 4037

Introduction to Cloud Computing

Lecture 27

Danyal Farhat
FAST School of Computing
NUCES Lahore

AWS Storage

Lecture's Agenda

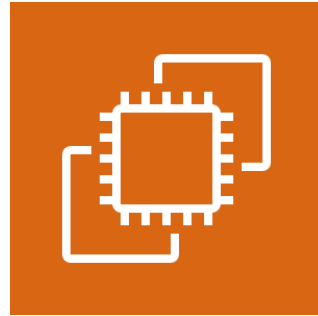
- **Amazon Elastic Block Store (Amazon EBS)**
- Amazon Simple Storage Service (S3)
- Amazon Elastic File System (Amazon EFS)
- Amazon Simple Storage Service Glacier



Core AWS services



Amazon Virtual Private Cloud (Amazon VPC)



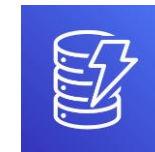
Amazon Elastic Compute Cloud (Amazon EC2)



Storage



Amazon Relational Database Service



Amazon DynamoDB

Database



AWS Identity and Access Management (IAM)

Amazon Elastic Block Store (Amazon EBS)

- Provides **persistent** (non-volatile) block storage volume for use with EC2 instances
- Each EBS volume is **replicated** within its Availability Zone
- Can be backed up to S3 storage through **snapshots**

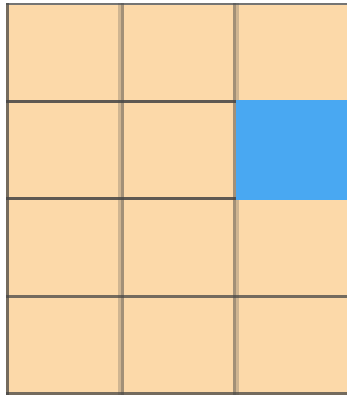


Amazon Elastic Block Store
(Amazon EBS)

AWS Storage Options: Block Storage versus Object Storage

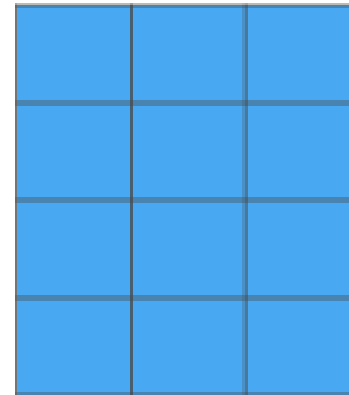


What if you want to change **one character** in a 1-GB file?



Block storage

Change one block (piece of the file)
that contains the character



Object storage

Entire file must be updated

Amazon EBS Common Uses

- **EC2 Boot volumes & EC2 additional volumes**
- **Data storage with a file system**
- **Database hosts**
- **Enterprise applications**

Amazon EBS Volume Types

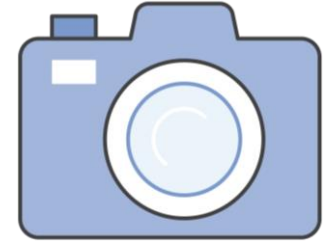
	Solid State Drives (SSD)		Hard Disk Drives (HDD)	
	General Purpose	Provisioned IOPS	Throughput-Optimized	Cold
	Maximum Volume Size	16 TiB	16 TiB	16 TiB
	Maximum IOPS/Volume	16,000	500	250
Maximum Throughput/Volume	250 MiB/s	1,000 MiB/s	500 MiB/s	250 MiB/s

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 volume 	<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 volumes 	<ul style="list-style-type: none"> Large database workloads 	<ul style="list-style-type: none"> Big data 	<ul style="list-style-type: none"> Scenarios where the lowest storage cost is important
<ul style="list-style-type: none"> Virtual desktops 		<ul style="list-style-type: none"> Data warehouses 	<ul style="list-style-type: none"> It cannot be a boot volume
<ul style="list-style-type: none"> Low-latency interactive applications 		<ul style="list-style-type: none"> Log processing 	
<ul style="list-style-type: none"> Development and test environments 		<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 Demo



Set up demo

Amazon Elastic Block Store (EBS)



Lab 4: Working with Amazon EBS

Lab 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: Working with Amazon EBS (Cont.)

Lab Tasks:

- **Task 1 – Create a New EBS Volume**
- **Task 2 – Attach the Volume to an Instance**
- **Task 3 – Connect to the Amazon EC2 Instance**
- **Task 4 – Create and Configure the File System**
- **Task 5 – Create an Amazon EBS Snapshot**
- **Task 6 – Restore the Amazon EBS Snapshot**
 - **Create a Volume Using the Snapshot**
 - **Mount the Restored Volume**

Amazon EBS – Key Points

- Persistent and customizable **block storage** for Amazon EC2
- HDD and SSD types
- Replicated in the same Availability Zone
- Easy and transparent **encryption**
- Elastic volumes
 - Instance don't required to power-off during resizing of volume
- Back up by using **snapshots**

Lecture's Agenda

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



Amazon Simple Storage Service (Amazon S3)

- S3 is **managed object-level** storage
 - If user want to change a part of a file, user must make the change and then **re-upload the entire modified** file
- 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
 - Bucket names are **universal** and must be **unique** worldwide across all existing buckets



Amazon S3

Amazon S3 (Cont.)

- **Amazon S3 includes event notifications**
 - Enables to set up **automatic notifications** when certain events occur
 - ✓ Example: An object is uploaded to a bucket or deleted from a bucket
 - Notifications can be sent to customer
 - Notifications can be used to trigger other processes like Lambda functions
- **You get fine-grained control over who can access S3 data**
 - IAM policies
 - S3 bucket policies
 - Per-object access control lists
- **Amazon S3 is accessible through AWS Management Console, AWS CLI, AWS SDK, or third party tools**

S3 Storage Classes

Amazon S3 Standard:

- Designed for **frequently accessed data**
- Delivers **low latency** and **high throughput**
- Appropriate for **cloud applications**, dynamic websites, content distribution, mobile and gaming applications, and big data analytics

S3 Storage Classes (Cont.)

Amazon S3 Standard-Infrequent Access:

- Used for data that is **accessed less frequently**, but requires rapid access when needed
- Designed to provide S3 Standard, with a **low per-GB storage price** and **per-GB retrieval fee**
- Good for **long-term storage and backups**, and as a data store for disaster recovery files

S3 Storage Classes (Cont.)

Amazon S3 Intelligent-Tiering

- Designed to optimize costs by automatically moving data to the most **cost-effective access tier**, without performance impact or operational overhead
- Monitors access patterns of the objects and moves the objects that have not been accessed for 30 consecutive days to the **infrequent access tier**
 - If an object in the infrequent access tier is accessed, it is automatically moved back to the frequent access tier

S3 Storage Classes (Cont.)

Amazon S3 One Zone-Infrequent Access:

- Stores data in a **single Availability Zone**
 - Unlike other S3 storage classes, which store data in a minimum of three AZs
- Costs **less than** Amazon S3 Standard-IA
- Works well for customers who want a **lower-cost option** for infrequently accessed data
- Good choice for **storing secondary backup copies** of on-premises data or easily re-creatable data

S3 Storage Classes (Cont.)

Amazon S3 Glacier:

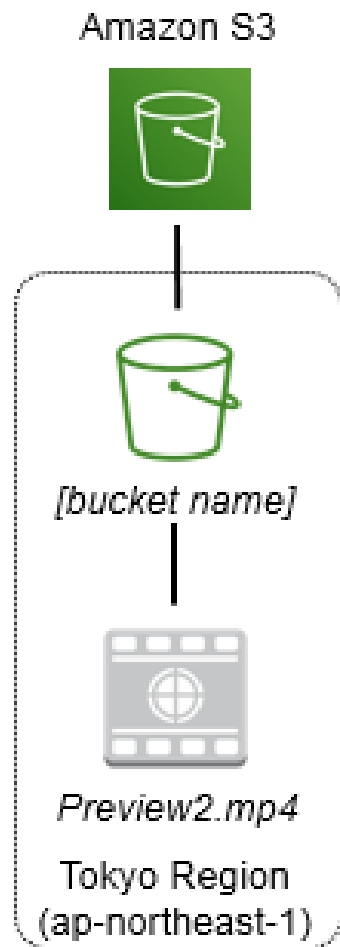
- Low-cost storage class for **data archiving**
- Provides **three retrieval options** that range from a few minutes to hours
- Customer can **upload objects** directly to Amazon S3 Glacier
- Customer can use **S3 lifecycle policies** to transfer data between any of the S3 storage classes

S3 Storage Classes (Cont.)

Amazon S3 Glacier Deep Archive:

- Lowest-cost storage class
- Supports **long-term retention and digital preservation** for data that might be accessed once or twice in a year
- Designed for customers in **highly regulated industries**
 - Example: Financial services, healthcare, and public sectors
 - ✓ Retain datasets for 7–10 years (or more) to meet regulatory compliance requirements
- All objects are **replicated** across at least three Availability Zones, and these objects can be restored within 12 hours

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>

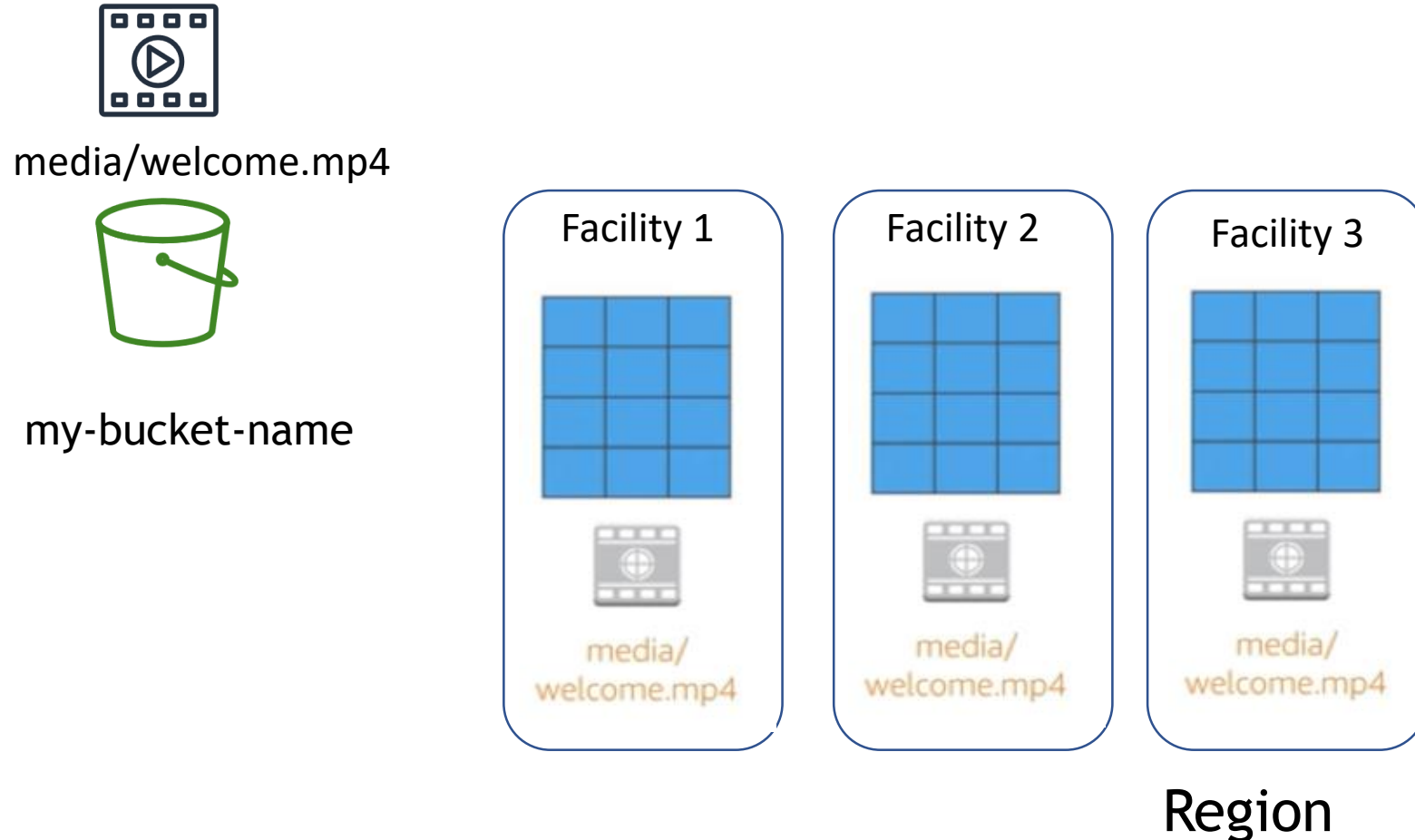
Region code Bucket name

Bucket virtual hosted-style URL endpoint:

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

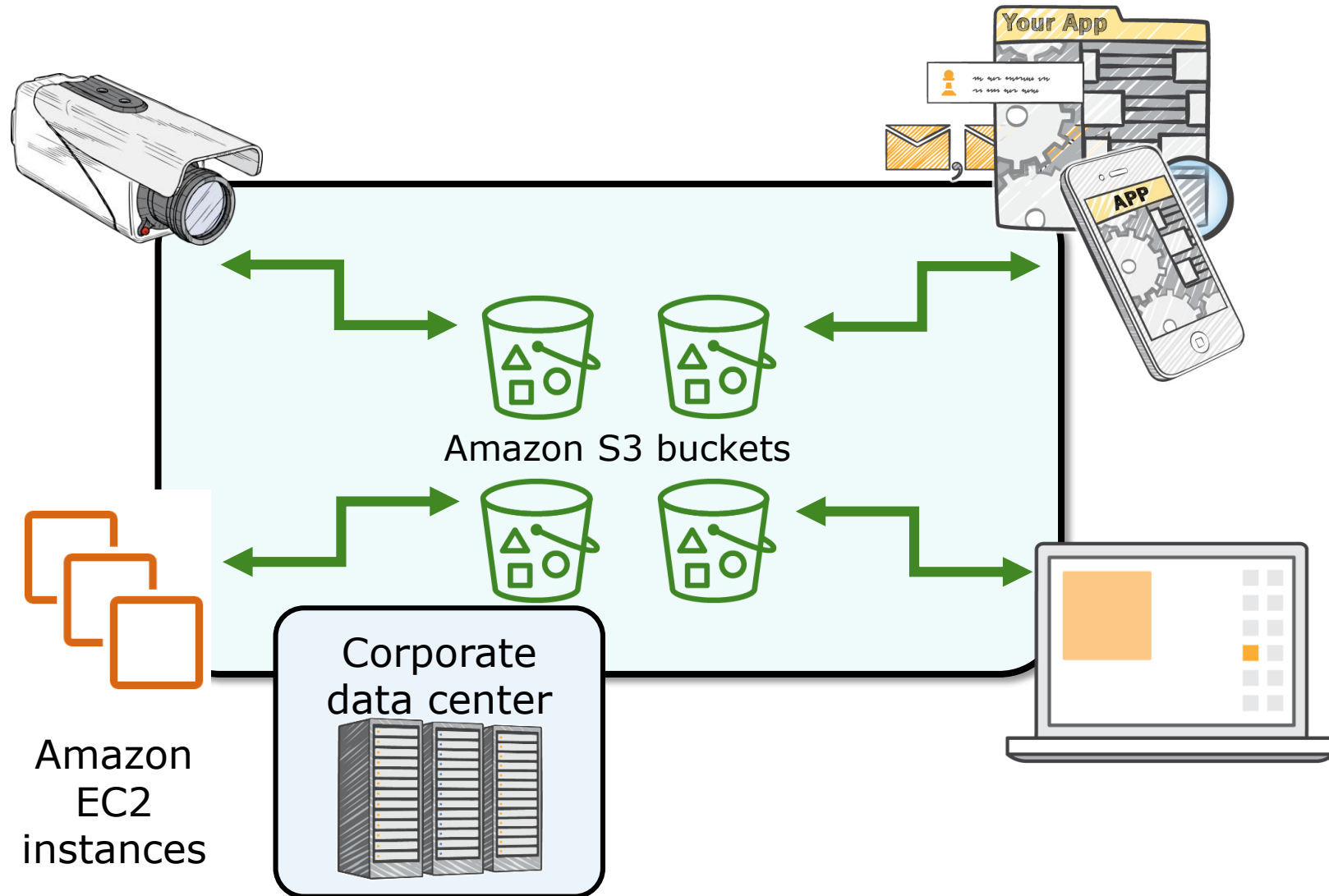
Bucket name Region code

Data is Redundantly Stored in the Region



Amazon S3 Common Scenarios

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



Amazon S3 Pricing

- **You pay for**
 - 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 Factors

- To estimate Amazon S3 costs, consider the following:

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

Amount of Storage:

- The number and size of objects

Amazon S3: Storage Pricing Factors (Cont.)

Requests:

- The number and type of requests (GET, PUT, COPY)
 - Different rates for different request types

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 S3 Demo



Set up demo

Amazon S3



Amazon S3 – Key Points

- Amazon S3 is a **fully managed** cloud storage service
- You can **store** a virtually unlimited number of objects
- You can **access** Amazon S3 at any time from anywhere through a URL
- Amazon S3 **offers** rich security controls

Lecture's Agenda

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



Amazon Elastic File System

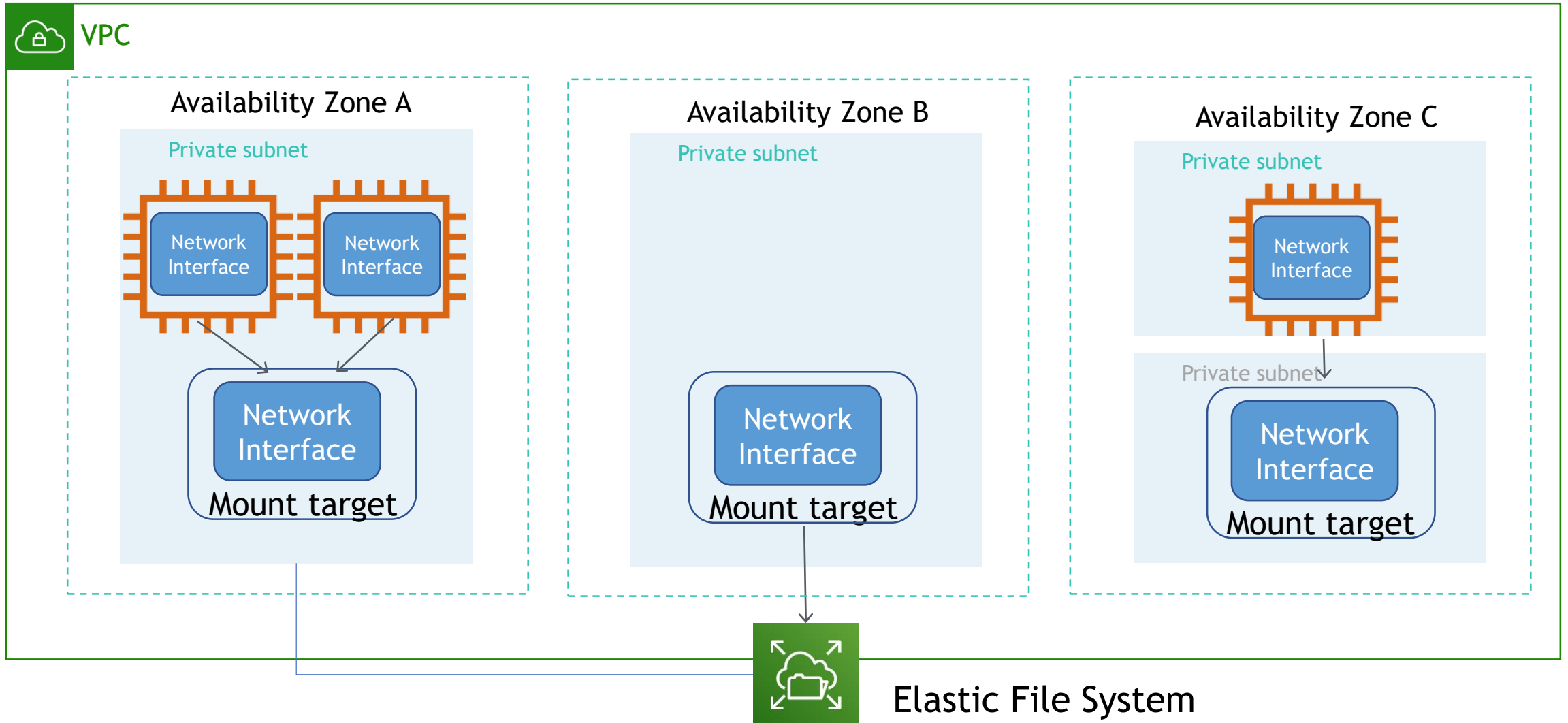
- Amazon EFS implements storage for EC2 instances that **multiple virtual machines** can access at the same time
- It is implemented as a **shared file system** that uses the Network File System (NFS) protocol



Amazon EFS

- Compatible with all **Linux-based AMIs** for Amazon EC2
- Works well for **big data analytics**, media processing workflows, content management, web serving, and home directories

Amazon EFS Architecture



Amazon EFS Implementation

- **Create EC2 instance**
- **Launch EC2 instance**
- **Create EFS file system**
- **Create mount targets in the appropriate subnets**
- **Connect Amazon EC2 instances to the mount targets**
- **Verify the resources and protection of the 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 EFS – Key Points

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

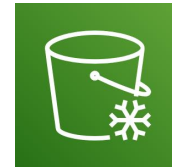
Lecture's Agenda

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



Amazon S3 Glacier

- A **data archiving service** designed for security, durability, and an extremely low cost
- Supports **encryption** of data in transit and at rest through Transport Layer Security (TLS)
- Three retrieval options
 - Expedited: 1–5 minutes
 - Standard: 3–5 hours
 - Bulk: 5–12 hours



Amazon S3 Glacier



Amazon S3 Glacier – Key Terms

Archive:

- Any **object** that you store in Amazon S3 Glacier
- It is the **base unit of storage** in Amazon S3 Glacier
- Each archive has its own **unique ID**
- Each archive can also have a **description**

Amazon S3 Glacier – Key Terms (Cont.)

Vault:

- A **container** for storing archives
- When you create a vault, you specify the **vault name** and the **Region** where you want to locate the vault

Amazon S3 Glacier – Key Terms (Cont.)

Vault Access Policy:

- Determine who can **access the data** stored in the vault
- Determine what **operations** users can perform
- One vault **access permissions policy** can be created for each vault to manage access permissions for that vault
- You can also use a **vault lock policy** to make sure that a vault cannot be altered
- Each vault can have **one vault access policy** and **one vault lock policy** that are attached to it

Amazon S3 Glacier Use Cases



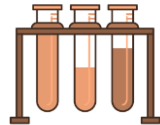
Media asset archiving



Healthcare information archiving



Regulatory and compliance archiving



Scientific data archiving

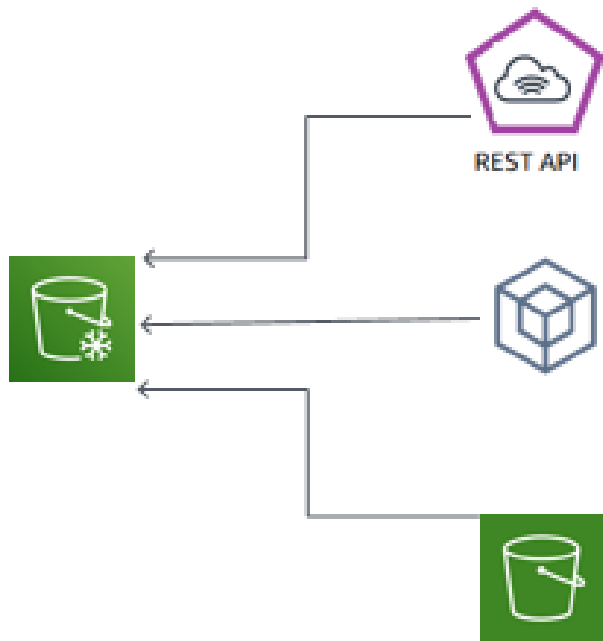


Digital preservation



Magnetic tape replacement

Using Amazon S3 Glacier



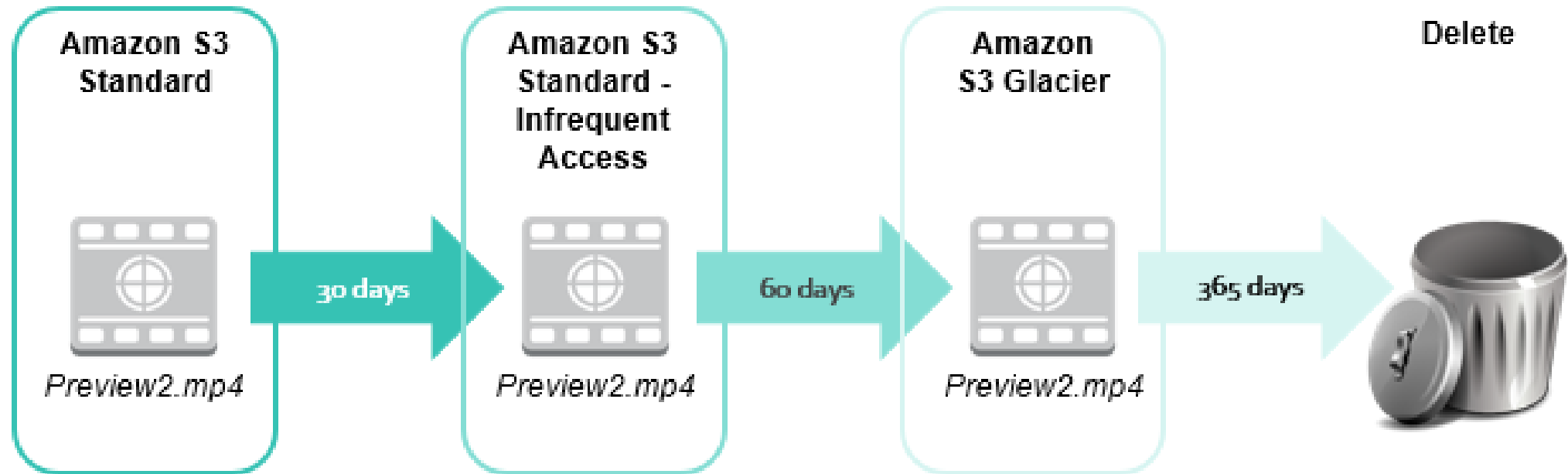
RESTful
web services

Java or .NET
SDKs

Amazon S3 with
lifecycle policies

Lifecycle Policies

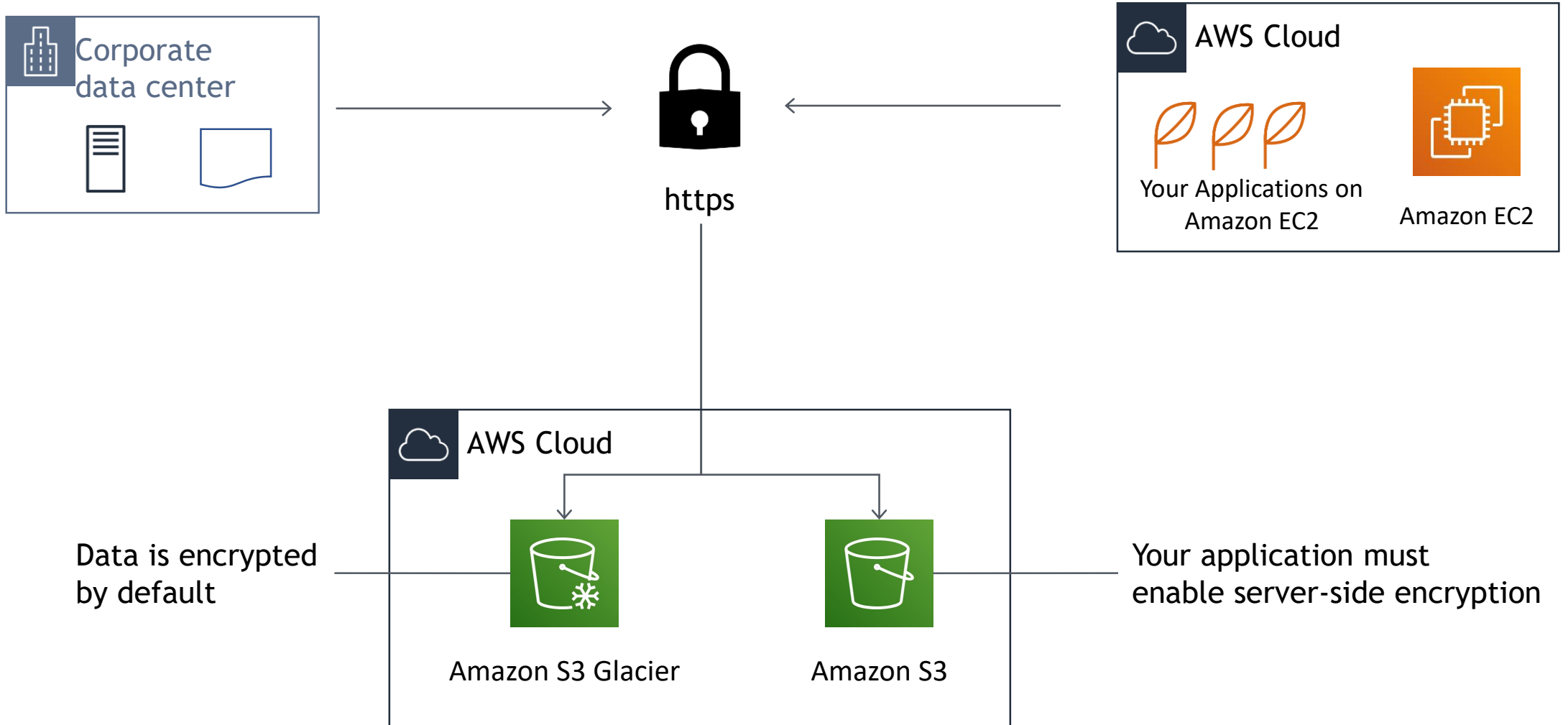
- Amazon S3 lifecycle policies **enable customer** to delete or move objects based on age



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

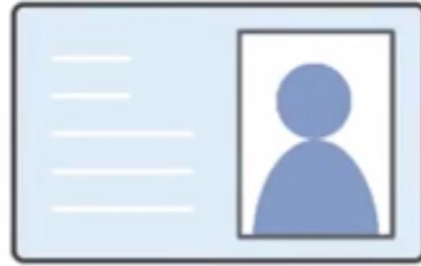
Server-Side Encryption



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**

Amazon S3 Glacier – Key Points

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

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-amp-efs-explained>

Questions?