Simple Storage Services (S3) , Cloudfront and SnowBall

Agenda

- ➤ What is AWS S3
- ➤ S3 Advantages
- ➤ S3 Buckets , Objects , Keys and Endpoints
- ➤ S3 Data Consistency Model
- ➤ S3 Transfer Acceleration
- ➤ S3 Storage Types
- ➤ S3 Versioning
- ➤ S3 Life Cycle Management
- ➤ S3 Data Protection
- ➤ S3 Cross-Region Replication
- ➤ S3 Hosting a Static Website
- ➤ Overview of Cloudfront
- ➤ Overview of Snowball
- ➤ Hand-on Lab

What is AWS S3

❑ Amazon Simple Storage Service is storage for the Internet.

❑ Amazon S3 has a simple web services interface that you can use to store and retrieve any amount of data, at any time, from anywhere on the web.

❑ Accessible via simple web service interface.

❑ Use alone or with other service of AWS services.

❑ Amazon S3 provide 99.999999999% [eleven 9's] durability. If you store 10,000 objects, on average you may lose one of them every 10 million years or so.

AWS S3: Advantages

Create Buckets

Store data in Buckets

Download Data

Permissions

Standard Interfaces

Transfer Acceleration

AWS S3 Concepts: Buckets

❑ A bucket is a container for objects stored in Amazon S3.

❑ Every object is contained in a bucket.

❑ Organise the AWS S3 namespace at the highest level.

❑ Each object in Amazon S3 has a unique HTTP URL.

❑ For example, if the object named awsdoc.txt is stored in the "awss3" bucket, then it is addressable using the URL https://s3.amazonaws.com/awss3/awsdoc.txt

❑ You can choose any AWS region that is geographically close to you to optimize latency, minimize costs, or address regulatory requirements.

AWS S3 Concepts: Objects

❑ Objects are fundamental entities stored in S3.

❑ Object is uniquely identified within a bucket by key(name) and a version ID.

❑ Each object in Amazon S3 has a unique HTTP URL.

❑ Objects can be 1 byte to 5 terabytes (TiB).

❑ Objects of differing storage classes can exist in the same bucket.

AWS S3 Concepts: Keys

❑ A key is the unique identifier for an object within a bucket.

❑ Unique Version IDs are randomly generated, 1024 bytes long.

❑ Every object in a bucket has exactly one key.

❑ Every object in Amazon S3 can be uniquely addressed through the, bucket name + key + version (optional).

  o http://s3.amazonaws.com/doc/folder1/amazons3.jpg

  o "doc" is the name of the bucket

  o "folder1/amazons3.jpg" is the key

AWS S3 Concepts: Endpoints

❑ If you create a client by specifying the US East (N. Virginia) Region, it uses the following endpoint to communicate with Amazon S3

  o <s3.amazonaws.com>

❑ If you create a client by specifying any other AWS Region, each of these regions maps to the region- specific endpoint:

  o <s3-*region*.amazonaws.com>

  o Example, s3-eu-west-1.amazonaws.com (for eu-west-1 region)

AWS S3: Data Consistency Model

❑ Amazon S3 provides read-after-write consistency for PUTS of new objects in your S3 bucket.

❑ Amazon S3 offers eventual consistency for overwrite PUTS and DELETES.

❑ If a PUT request is successful, your data is safely stored. However, information about the changes must replicate across Amazon S3, which can take some time, and so you might observe the following behaviors:

  o A process writes a new object to Amazon S3 and immediately lists keys within its bucket. Until the change is fully propagated, the object might not appear in the list.

  o A process replaces an existing object and immediately attempts to read it. Until the change is fully propagated, Amazon S3 might return the prior data.

  o A process deletes an existing object and immediately attempts to read it. Until the deletion is fully propagated, Amazon S3 might return the deleted data.

  o A process deletes an existing object and immediately lists keys within its bucket. Until the deletion is fully propagated, Amazon S3 might list the deleted object.

AWS S3: Transfer Acceleration

❑ Amazon S3 Transfer Acceleration enables fast, easy, and secure transfers of files.

❑ Transfer Acceleration takes advantage of Amazon CloudFront's globally distributed edge locations.

❑ When using Transfer Acceleration, additional data transfer charges may apply.

❑ After enabling Amazon S3 Transfer Acceleration feature a new name will be generated.

❑ Using the Transfer Acceleration, allows to speed up upload by multiple times.

❑ Example,
  o If the bucket's initial name was mybucket.s3.amazonaws.com
  o The new one will be mybucket.s3-accelerate.amazonaws.com
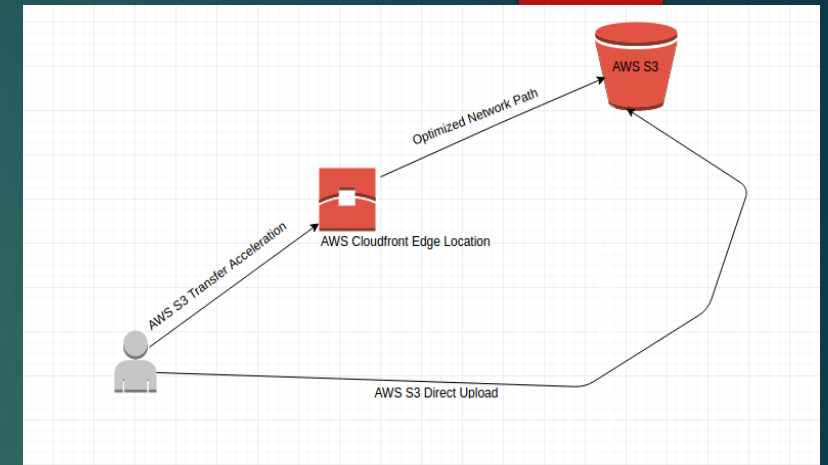
Comparison of S3 Storage Types

	Standard	Standard - Infrequent Access	Reduced Redundancy Storage
Durability	99.999999999%	99.999999999%	99.99%
Availability	99.99%	99.9%	99.99%
Concurrent facility fault tolerance	2	2	1
SSL support	Yes	Yes	Yes
First byte latency	Milliseconds	Milliseconds	Milliseconds
Lifecycle Management Policies	Yes	Yes	Yes

AWS S3: Versioning

❑ When you PUT an object in a versioning-enabled bucket, the noncurrent version is not overwritten.

❑ When a new version of photo.gif is PUT into a bucket that already contains an object with the same name, the original object   (ID = 111111) remains in the bucket, Amazon S3 generates a new version ID (121212), and adds the newer version to the bucket.

❑ When you DELETE an object, all versions remain in the bucket and Amazon S3 inserts a delete marker.

❑ The delete marker becomes the current version of the object.

❑ When the current version is a delete marker returns a "404 not found" error.
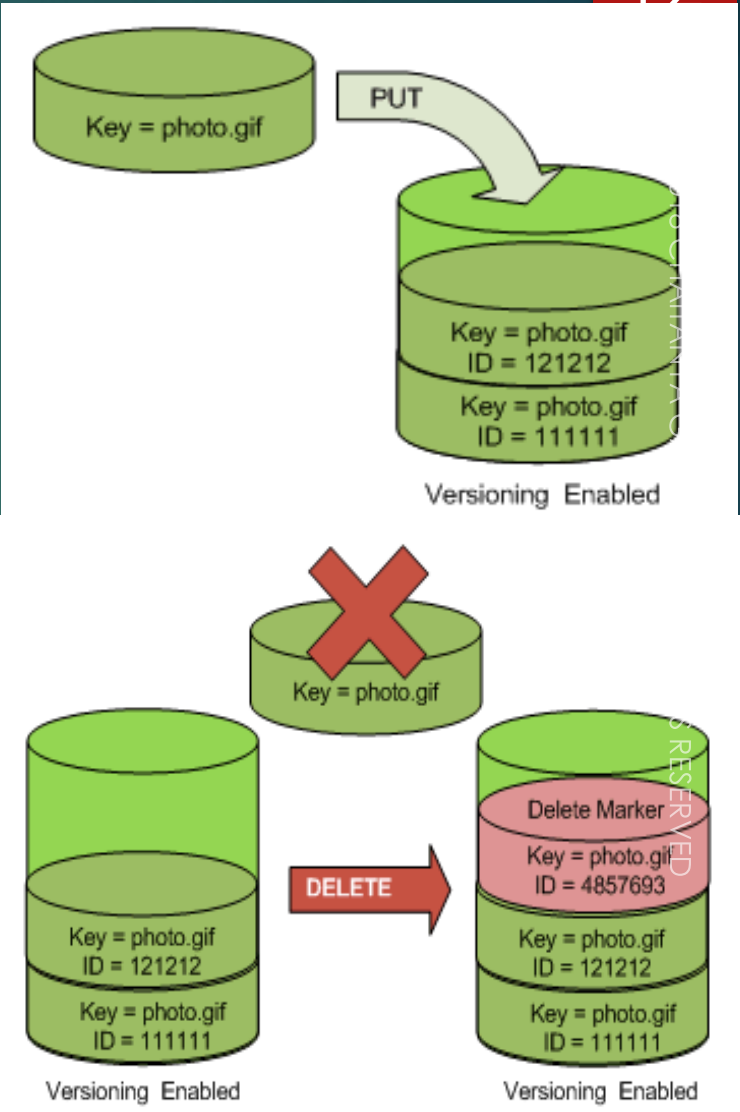
AWS S3: Life Cycle Management

❑ Lifecycle configuration enables you to specify the lifecycle management of objects in a bucket.

❑ These actions can be classified as follows:
   o **Transition actions**: You define when objects transition to another storage class.
      ➢ Transition objects to the STANDARD_IA storage class 30 days after creation
      ➢ Archive objects to the GLACIER storage class one year after creation.
   o **Expiration actions**: You specify when the objects expire.
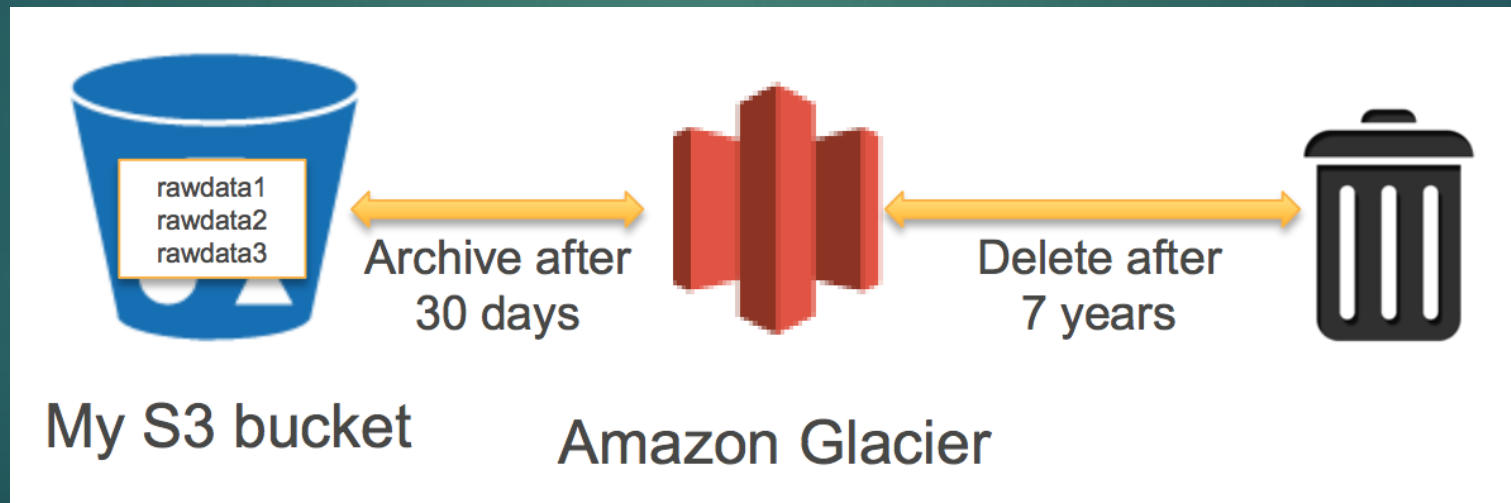
AWS S3: Data Protection

❑ You can protect data in transit by using encryption.

❑ You have the following options of protecting data at rest in Amazon S3.

- o **Use Server-Side Encryption** – You request Amazon S3 to encrypt your object before saving it on disks in its data centers and decrypt it when you download the objects.

- o **Use Client-Side Encryption** – You can encrypt data client-side and upload the encrypted data to Amazon S3.

AWS S3: Cross-Region Replication

❑ Cross-region replication is a bucket-level feature that enables automatic, asynchronous copying of objects across buckets in different AWS regions.

❑ Amazon S3 replicates the new objects created after enabling the replication configuration.

❑ Amazon S3 does not replicate objects in the source bucket for which the bucket owner does not have permissions.

❑ Objects in the source bucket that are replicas, created by another cross-region replication, are not replicated.
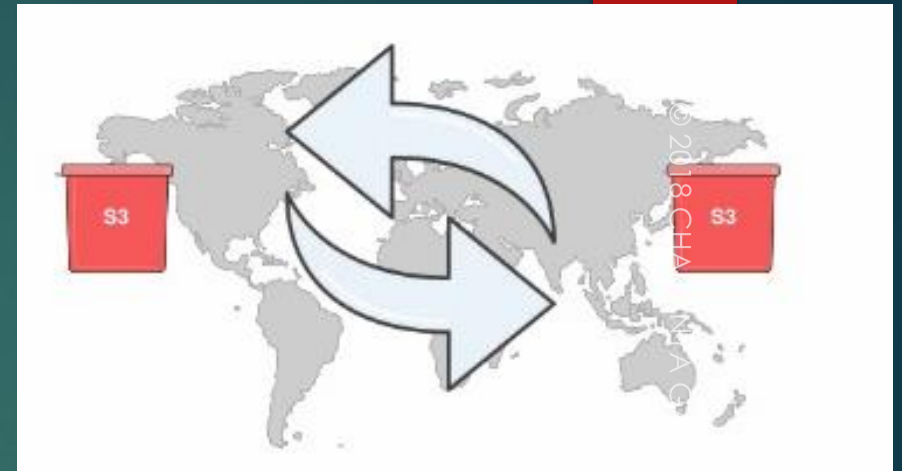
AWS S3: Hosting a Static Website

❏ You can host a static website on Amazon Simple Storage Service (Amazon S3).

❏ To configure a bucket for static website hosting, you add a website configuration to your bucket.

❏ To host a static website on Amazon S3, you need only provide the name of the index document.

AWS S3: Limitation

Description	Limitation
Buckets in each accounts	100
Bucket naming	3 to 63 characters long

Hands-On-Lab

Hands-on Lab

- ❑ Create Bucket
- ❑ Upload Object
- ❑ Access Object
- ❑ Configure versioning
- ❑ Host static website
- ❑ Configure cross-zone replication
- ❑ Enable Transfer Acceleration
- ❑ Configure Life Cycle Management

Amazon Cloudfront

Amazon Cloudfront

▶ Amazon CloudFront is a web service that speeds up distribution of your static and dynamic web content, for example, .html, .css, .php, image, and media files, to end users.

▶ CloudFront delivers your content through a worldwide network of edge locations.

▶ When an end user requests content that you're serving with CloudFront, the user is routed to the edge location that provides the lowest latency, so content is delivered with the best possible performance.

▶ If the content is already in that edge location, CloudFront delivers it immediately.

▶ If the content is not currently in that edge location, CloudFront retrieves it from an Amazon S3 bucket or an HTTP server (for example, a web server) that you have identified as the source for the definitive version of your content.

AWS Cloudfront terms

- Edge Location: An edge location is an amazon facility where your content will be cached. CloudFront edge locations are present all major the world. Contains over 50 edge locations spread across the world.

- Origin: An origin is a place where your original files that have to be served are present. An origin could be a web server (outside AWS), ec2 instance, s3, ELB, Route   another other storage which is accessible through the internet.

- Distribution: This is the name of the CDN which consists of several edge locations for your content distribution. Basically, this would be a unique URL for your content distribution.

- Web Distribution: Used typically for websites.

- RTMP: Typically for streaming media which supports Adobe flash media server RTMP protocol.

How Cloudfront Works
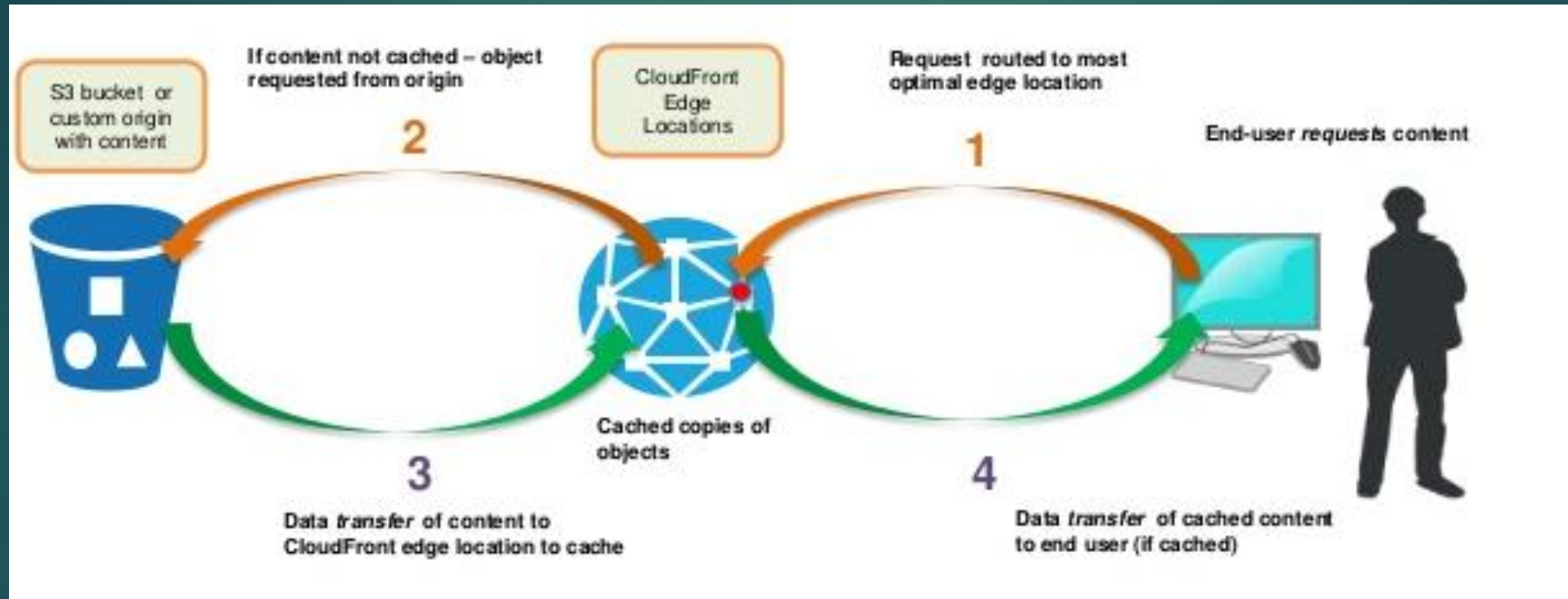
Snowball

What is AWS Snowball

❑ AWS Snowball is a service that accelerates transferring large amounts of data into and out of AWS using physical storage appliances, bypassing the Internet.

❑ Using Snowball helps to eliminate challenges that can be encountered with large-scale data transfers including high network costs, long transfer times, and security concerns.

AWS Snowball: How it Works

Plan your job

AWS Import/Export Snowball imports large amounts of data into Amazon S3 using an AWS-provided appliance to safely transfer your data. The following graph shows how Snowball works. Learn more.

CREATE A JOB

Create a job in the AWS Import/Export Snowball Management Console. AWS will ship you a Snowball for your job through UPS.

CONNECT THE SNOWBALL

Plug the Snowball into your local network. Download and run the Snowball client with your credentials to connect to the Snowball.

COPY DATA TO THE SNOWBALL

Copy your data onto the Snowball. Once complete, disconnect the Snowball and ship it back as-is, no packaging required.

AWS WILL MOVE YOUR DATA TO S3

After AWS gets the Snowball, your data will be moved into S3.

| Job created | In transit to you | Delivered to you | | In transit to AWS | At AWS | Completed |

AWS Snowball: Features

❑ Snowball with the Snowball appliance has the following features:

    o You can import and export data between your on-premises data storage locations and Amazon S3.

    o Snowball has an 80 TB model available in all regions, and a 50 TB model only available in the US regions.

    o Encryption is enforced, protecting your data at rest and in physical transit.

    o You don't have to buy or maintain your own hardware devices.

What is AWS Snowball: Pricing

❑ Snowball Edge has two pricing elements.
   o Service fee per data transfer job, which includes 10 days of onsite Snowball Edge device usage.
      ▪ Shipping days, including the day the device is received and the day it is shipped back to AWS, are not counted toward the 10 days.
   o Second, if the device is kept for more than 10 days, you can incur an additional fee for each additional day.

❑ Data transferred into AWS does not incur any data transfer fees.

❑ Snowball 50 TB: $200
❑ Snowball 80 TB: $250
❑ Extra day charge: $15

Thank You