AWS Certified Developer Associate

Lesson 4: Simple Storage Service (S3)

simplilearn

What You'll Learn

- Describe Simple Storage Service
- Identify different storage classes in S3
- Define Version Control and Lifecycle Management
- Define Buckets, ACLs, and User Policies
- Host a static website on S3
- Define Cross Origin Resource Sharing
- Define Tagging, Analytics, and Cloud-watch Metrics with S3

Basic Concepts of S3

S3 Overview

S3 provides a secure, durable, highly scalable, and object based storage.

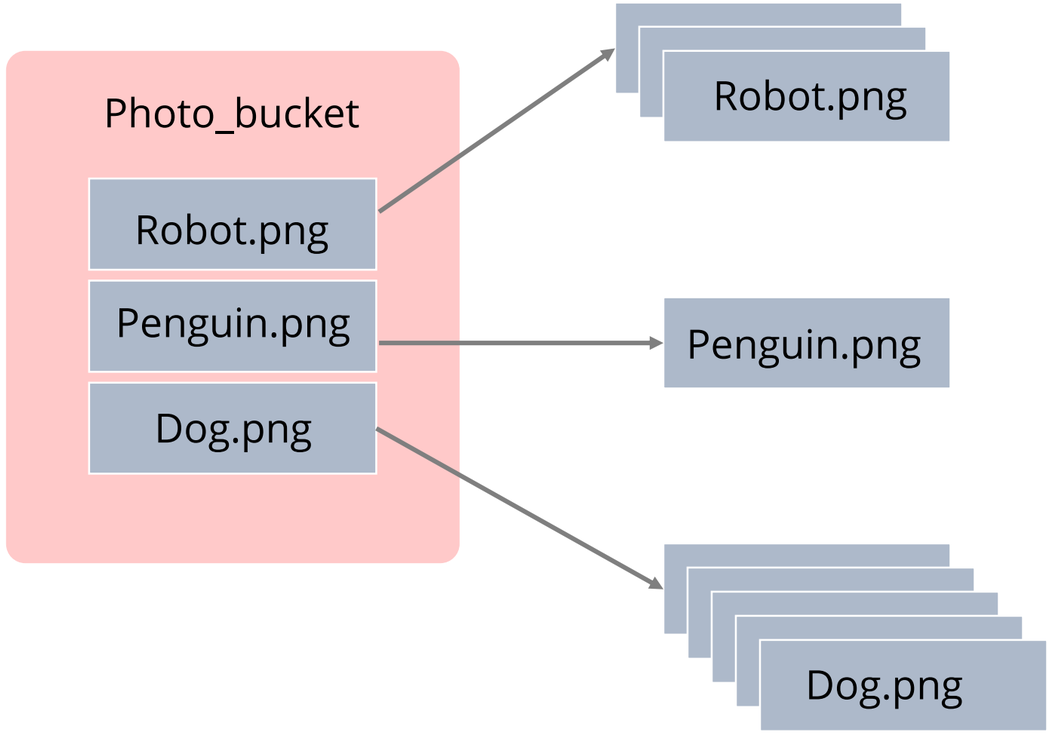
It can store unlimited data, with no restrictions on data format.

An object is a combination of a file and any meta data that describes that file.

S3 objects are stored in S3 bucket. Object sizes can vary from 1 MB to 5 TB.

Amazon S3 bucket must have a unique name as it shares a global name space. S3 supports access control policies and access logs.

S3 Bucket – Deep Dive



1. Organize and identify your content at the highest level
2. Identify who or which AWS account is responsible for data stored
3. Control who can read, write, modify, or delete the data
4. Unit of aggregation for usage reporting

S3 Bucket – Deep Dive, Cont.

Virtual Hosted Style URL	Path-Style URL
The bucket name is part of the domain name in the URL.	The bucket name is not part of the domain.
Example: http://simplilearn.s3.amazonaws.com	Example: http://s3.amazonaws.com/simplilearn

CompleteMultiPartUpload API

You can upload the objects in any sequence while using the Multipart Upload API to upload the objects.

S3 Storage Tiers: Standard & IA

Amazon S3 is highly available as it synchronizes the data to two other availability zones. Data stored in S3 is spread across multiple devices and facilities to prevent data loss. S3 offers four different storage tiers.

- **Standard Storage**
- **Infrequently Access**
- **Reduced Redundancy**
- **Glacier**

Standard Storage

S3 Standard Storage provides 99.99% availability and 99.999999999% or “11 nines” durability of data, which is stored redundantly across multiple devices in multiple facilities.
It is good for frequently accessed data and supports a wide variety of use cases, including dynamic websites, content distribution, and mobile and gaming applications.

S3 Storage Tiers: Standard & IA

Amazon S3 is highly available as it synchronizes the data to two other availability zones. Data stored in S3 is spread across multiple devices and facilities to prevent data loss. S3 offers four different storage tiers.

- **Standard Storage**
- **Infrequently Access**
- **Reduced Redundancy**
- **Glacier**

Infrequently Access

Infrequently Access or IA is used for data that is accessed less frequently, but it provides rapid access when needed. Because of the low cost and high performance, IA is good for offsite storage, backups, and disaster recovery data store.

S3 Storage Tiers: Standard & IA

Amazon S3 is highly available as it synchronizes the data to two other availability zones. Data stored in S3 is spread across multiple devices and facilities to prevent data loss. S3 offers four different storage tiers.

- **Standard Storage**
- **Infrequently Access**
- **Reduced Redundancy**
- **Glacier**

Glacier

- Secure, durable, and extremely low-cost storage service
- Retrieval of data can take up to 3-5 hours
- Ideal for long-term archive and offsite backup solutions

S3 Storage Tiers: Standard & IA

Amazon S3 is highly available as it synchronizes the data to two other availability zones. Data stored in S3 is spread across multiple devices and facilities to prevent data loss. S3 offers four different storage tiers.

Standard Storage

Infrequently Access

Reduced Redundancy

Glacier

Reduced Redundancy

- Guarantees 99.99% availability and 99.99% durability
- May lose one object out of 10,000 per year
- Ideal for non-critical data that can be reproduced

Data Consistency Models

Rules for order and visibility of newly created objects or any updates on existing objects.

1st Rule	S3 supports Read after write consistency for PUTS of new objects in all regions.
2nd Rule	S3 supports Eventual consistency for overwrite PUTS and DELETES.
NOTE	If two PUT requests are simultaneously made, the request with the latest time stamp wins.

Knowledge Check

KNOWLEDGE CHECK

Which of the following storage tiers provide 99.999999999% (11 nines) durability?

a. Glacier

b. Reduced Redundancy Storage

c. S3 Standard Storage

d. S3 Standard Storage - IA

KNOWLEDGE CHECK

Which of the following storage tiers provide 99.999999999% (11 nines) durability?

a. Glacier

b. Reduced Redundancy Storage

c. S3 Standard Storage

d. S3 Standard Storage - IA

The correct answers are **S3 Standard Storage & S3 Standard Storage - IA**

Explanation : S3 Standard Storage and IA provide 99.99% availability and 99.999999999% or “11 nines” durability of data, which is stored redundantly across multiple devices in multiple facilities.

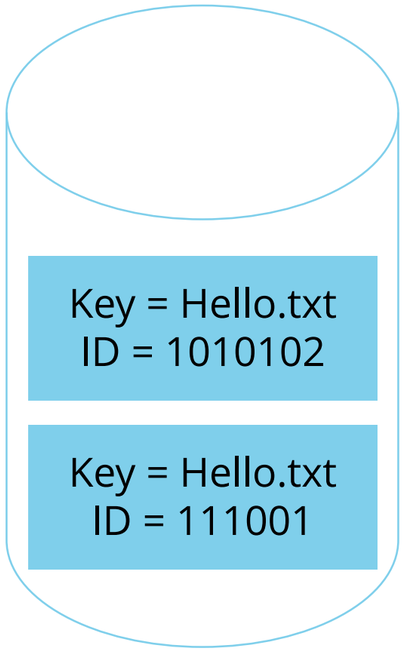
S3 Advanced Features

Versioning

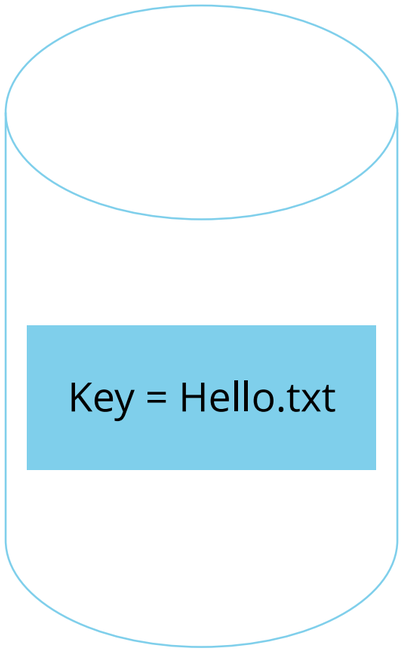
Describe Versioning

- Helps to keep multiple states of all objects within a bucket
- Helps to store, retrieve, and restore every version of every object
- Disabled by default
- Used with MFA, which provides an additional layer of security
- When a version-enabled object gets deleted, delete marker is placed
- Needs to be enabled on source bucket for Cross Region Replication
- Works in conjunction with lifecycle rules

Example



Versioning Enabled



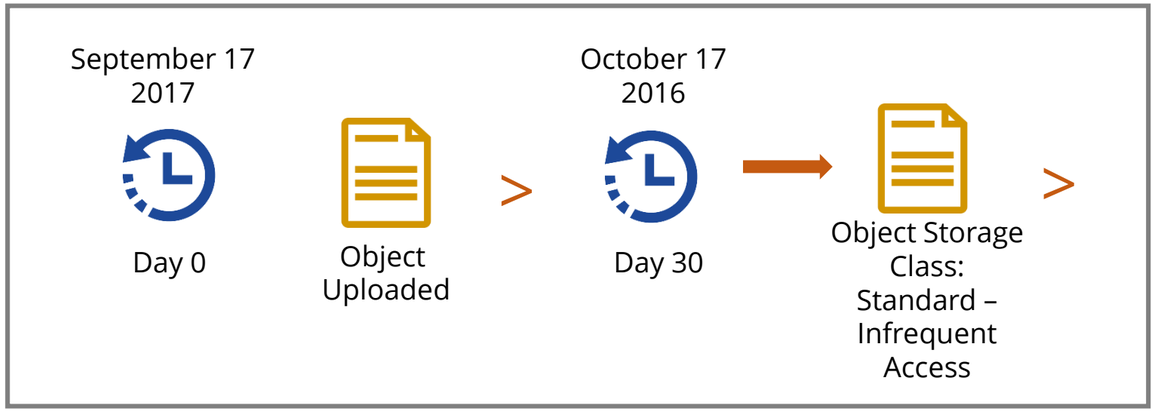
Versioning NOT Enabled

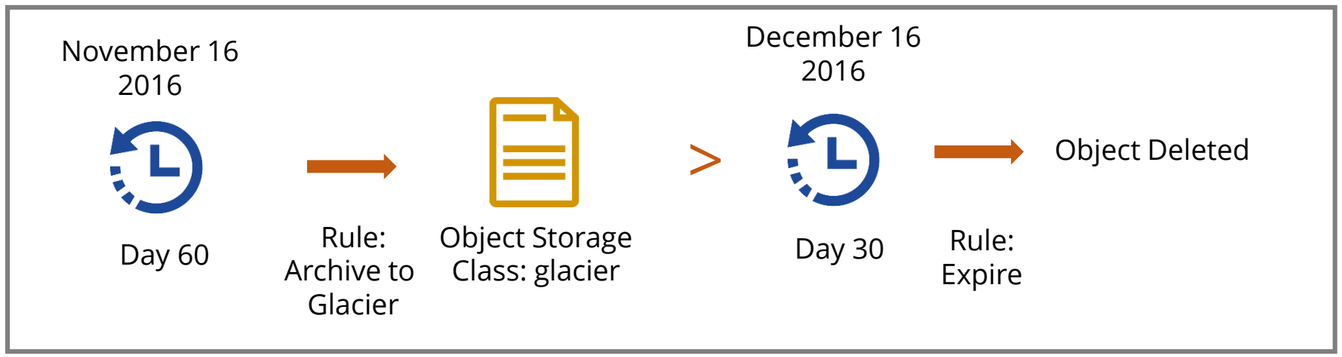
Life Cycle Management

Life Cycle Management

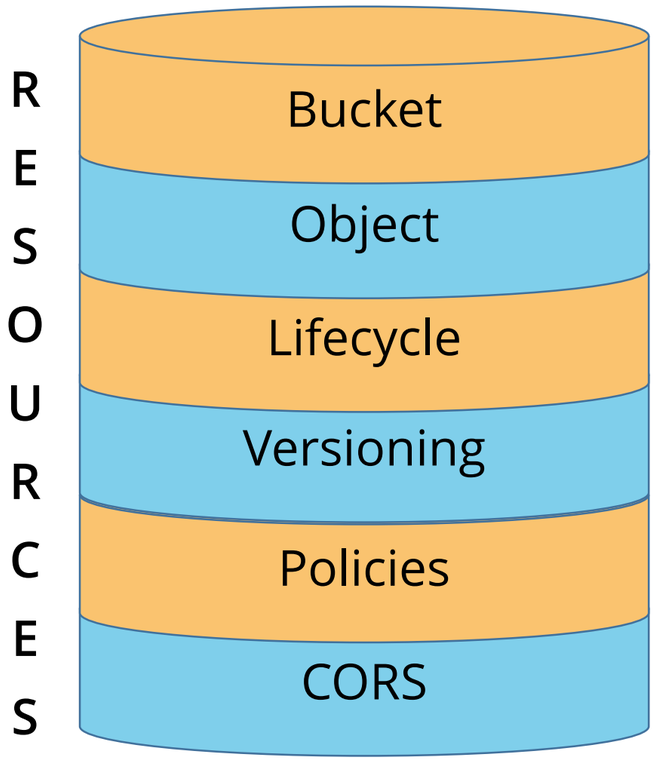
- Manages the object during its lifetime or based on the age of object
- Moves objects between different storage tires to reduce the cost of storage
- Gives you the ability to create up to 1000 lifecycle rules on a single bucket
- Uses a lifecycle configuration as XML
- Used with versioning and is applied to current or previous versions of the objects
- Implies some restriction on transitioning to Standard IA storage class

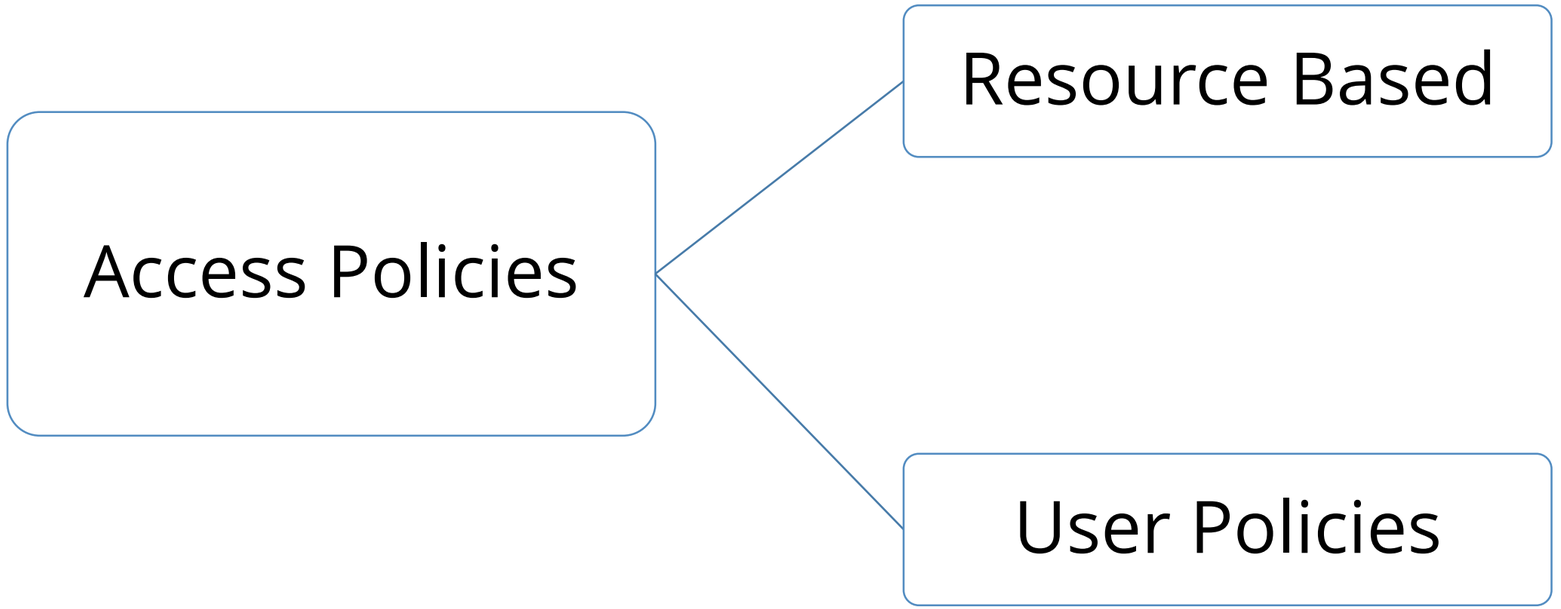
Example



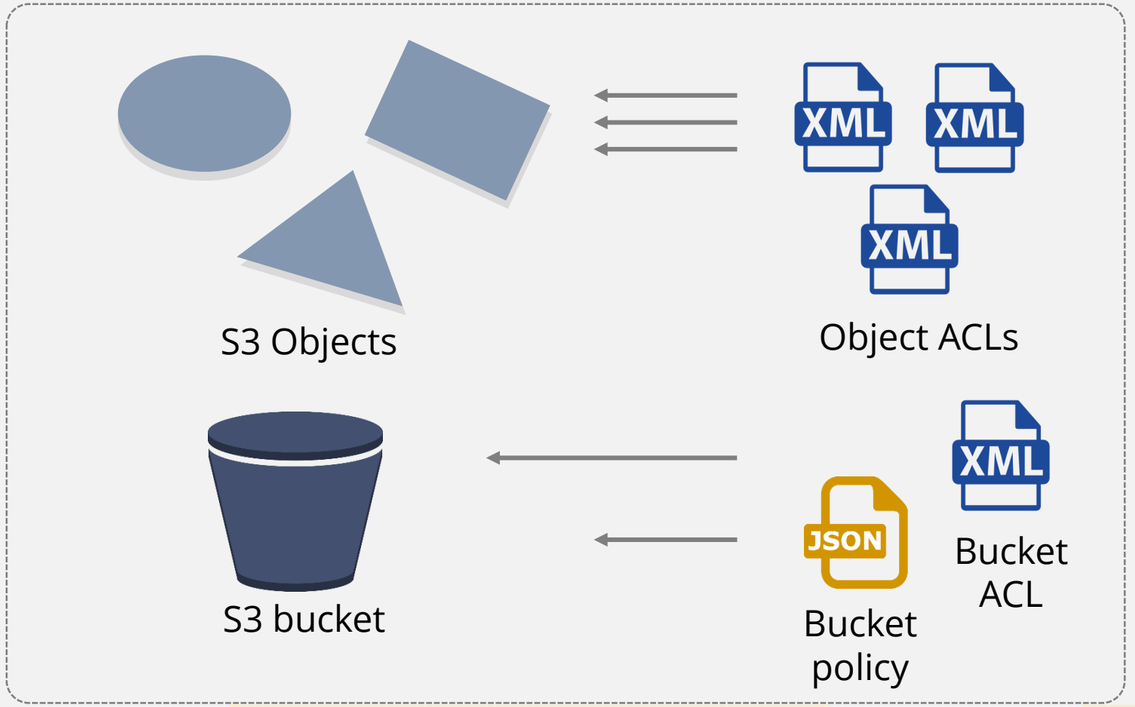


Resources and Access Policies





Bucket, User and ACL Policies, Cont.



ACLs

Use XML schema, and can be attached to both bucket and object

Bucket Policies

Use JSON style syntax that defines what permissions are allowed for a principal to take on an attached resource

User Policies

Use IAM policy document to decide what permissions or effects are allowed

Knowledge Check

KNOWLEDGE CHECK

Select two types of S3 access policies.

a. Resource based policies

b. Lifecycle policies

c. User policies

d. Versioning policies

KNOWLEDGE CHECK

Select two types of S3 access policies.

a. Resource based policies

b. Lifecycle policies

c. User policies

d. Versioning policies

The correct answer is **Resource based policies & Versioning policies**

Explanation : Amazon S3 supports two types of access policies: resource based policies and user policies. Resource based policies can be attached to the resources. Bucket policies and access control lists or ACLs are resource-based because you can attach them to your Amazon S3 resources.

S3 Deep Dive

Hosting a Static Website on S3

The website is available at the region-specific endpoint

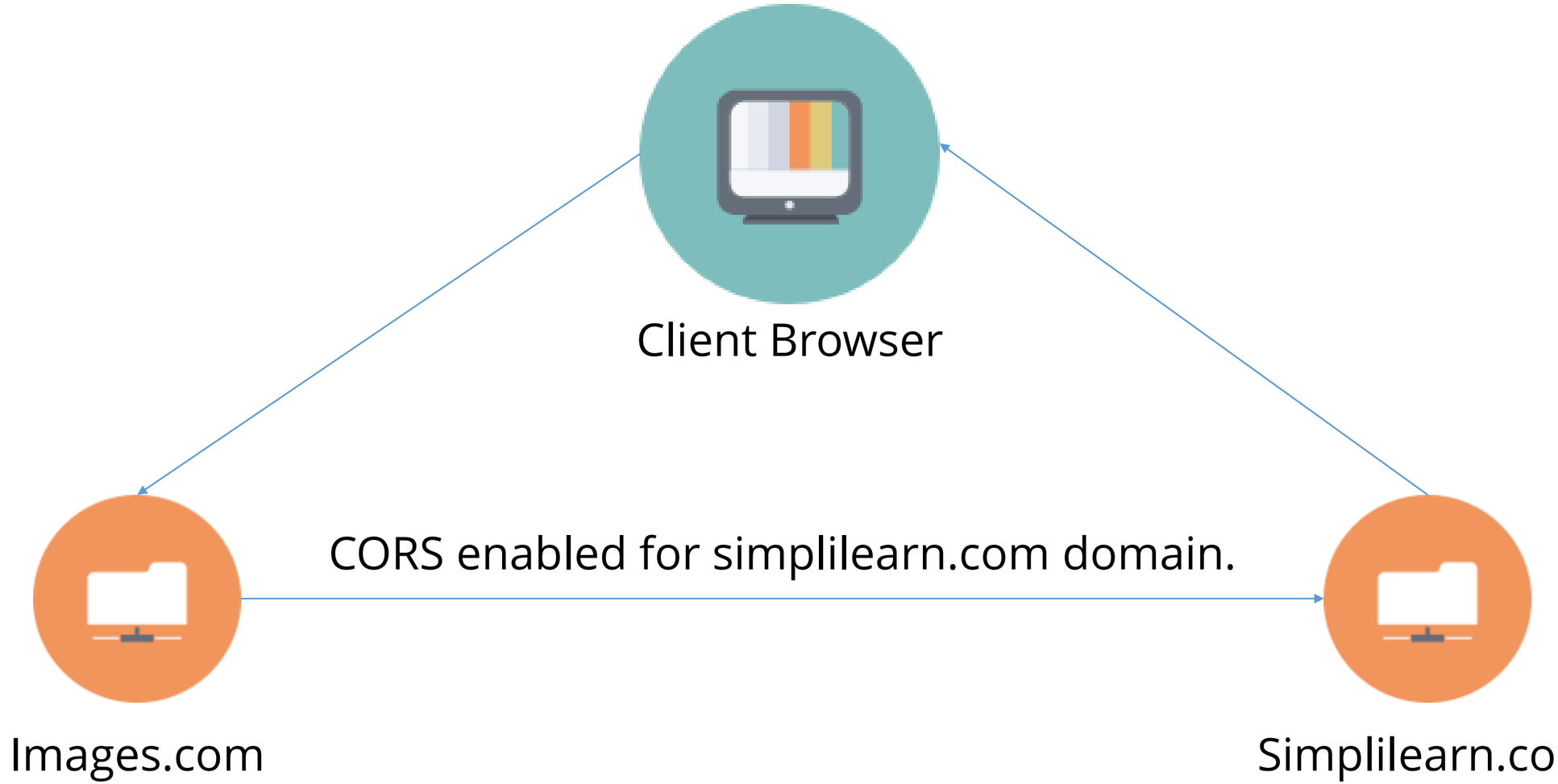
<bucket name>.s3-website-<AWS-region>.amazonaws.com

To access the image exam.jpg, the URL is

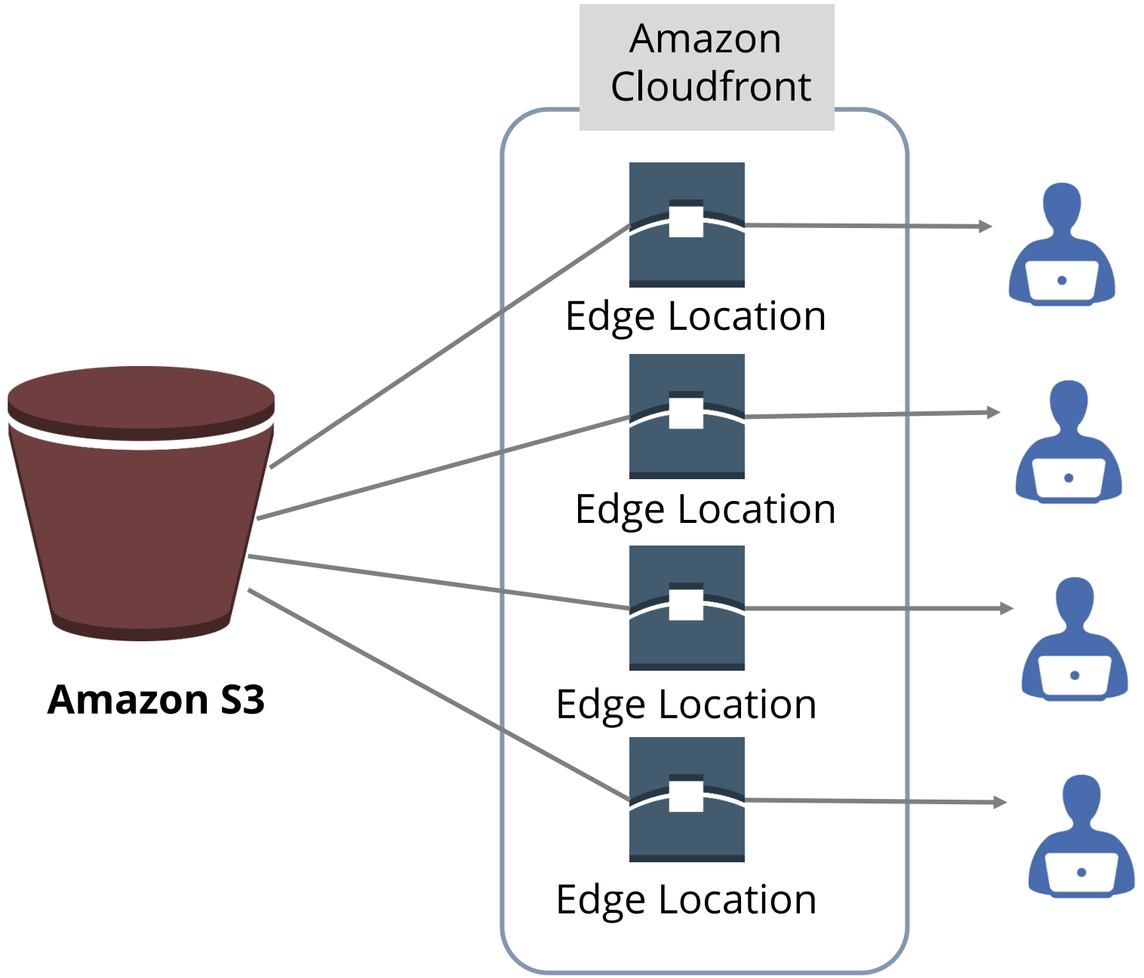
http://simplilearn.s3-website-us-east-1.amazonaws.com/examm.jpg

Supports own domain, such as simplilearn.com

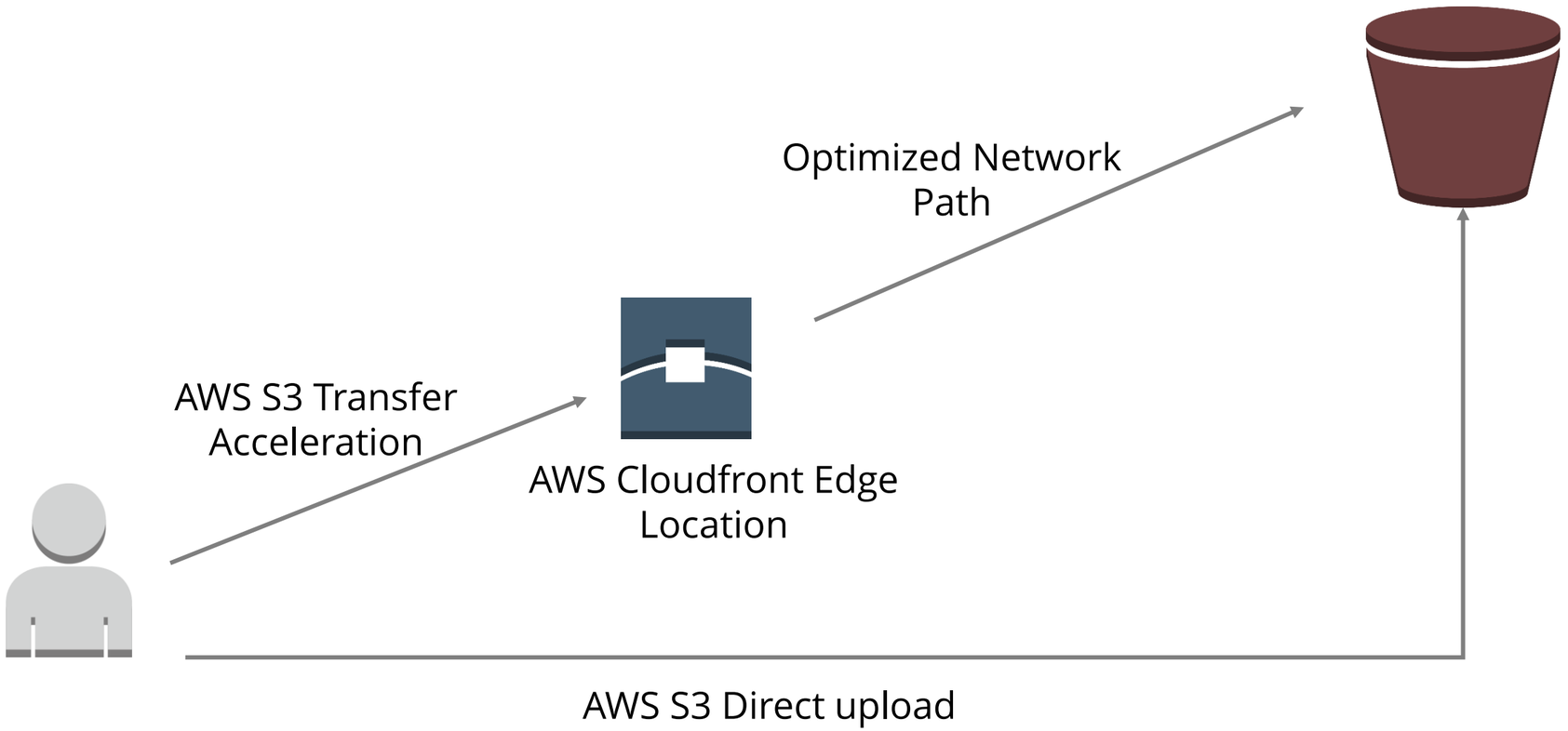
Cross Origin Resource Sharing (CORS)



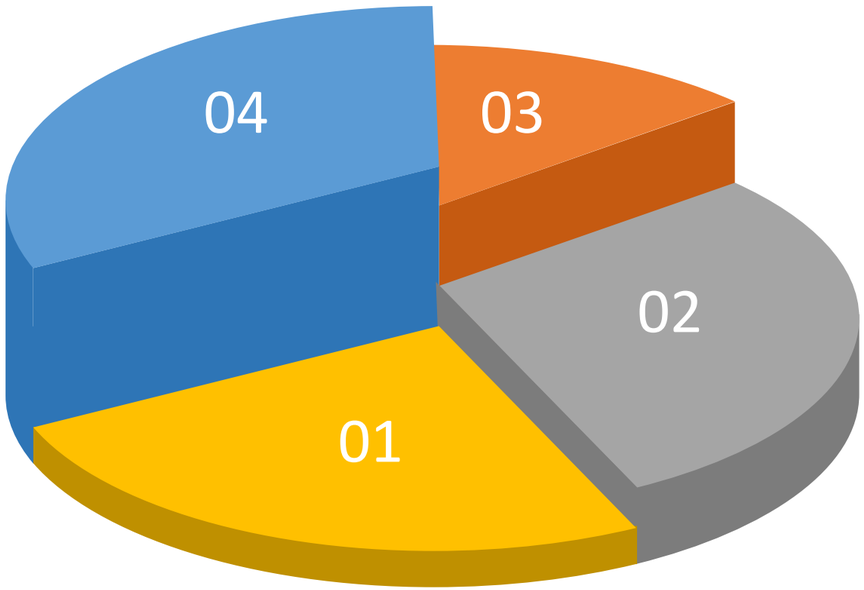
S3 with Cloud Front



Transfer Acceleration

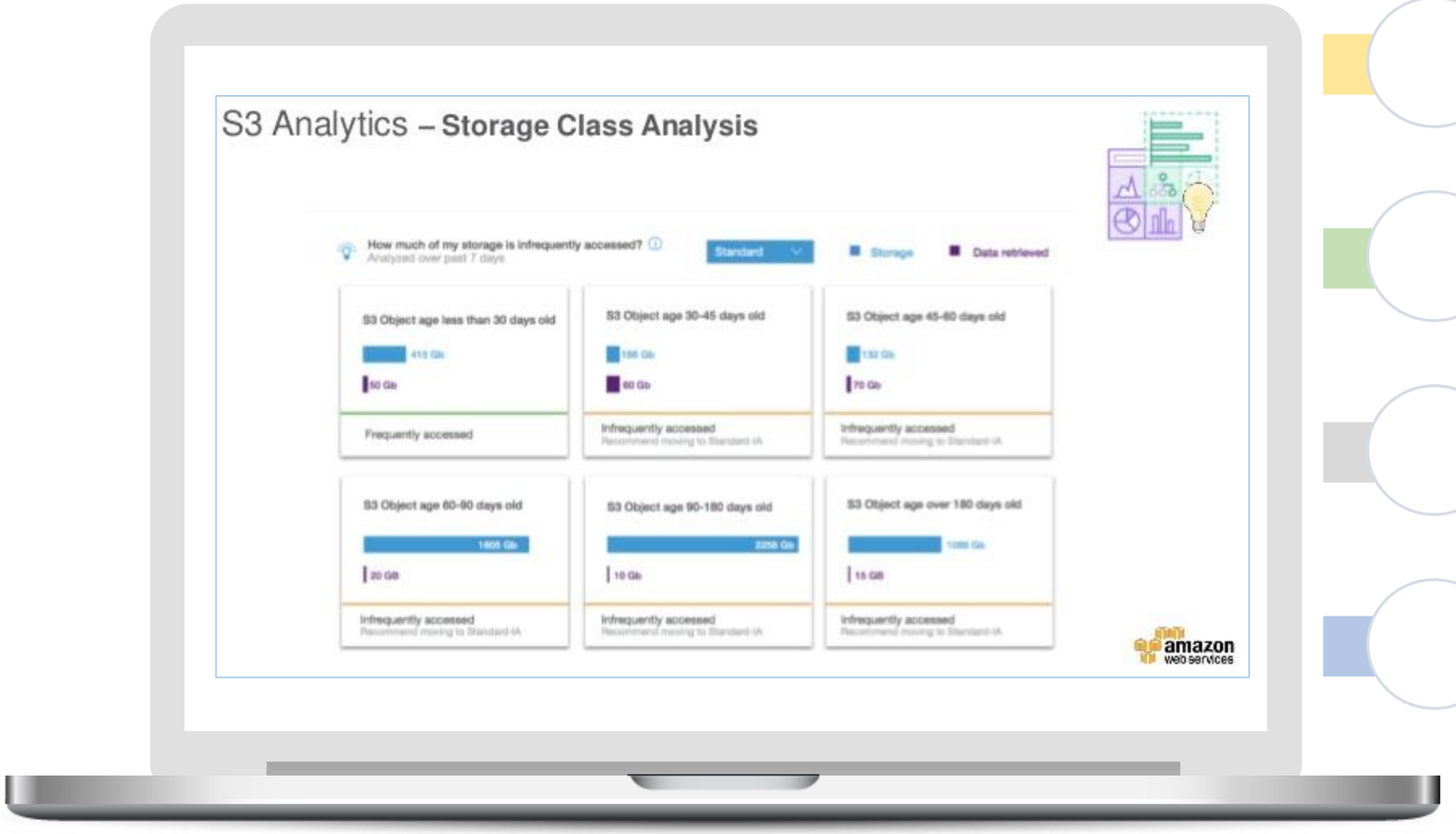


Tagging Objects



- **Tagging with S3 objects helps you manage access at object level**
- **Like normal tagging, these are key value pairs used to identify the objects and can be altered at any time**
- **With Tags, one can perform the following tasks:**
  - Create Identity and Access Management (IAM) policies
  - Setup S3 Lifecycle policies,
  - Customize storage metrics
- **These object-level tags can then manage transitions between storage classes and expire objects in the background**

S3 Analytics



- Better known as storage class analysis
- Patterns of stored objects are identified and analysis is done for transition of those objects to the correct storage class
- If an object is less frequently accessed, it will identify this and advise you to move the object to SIA
- Storage class analysis also provides daily visualizations of your storage usage

S3 Inventory

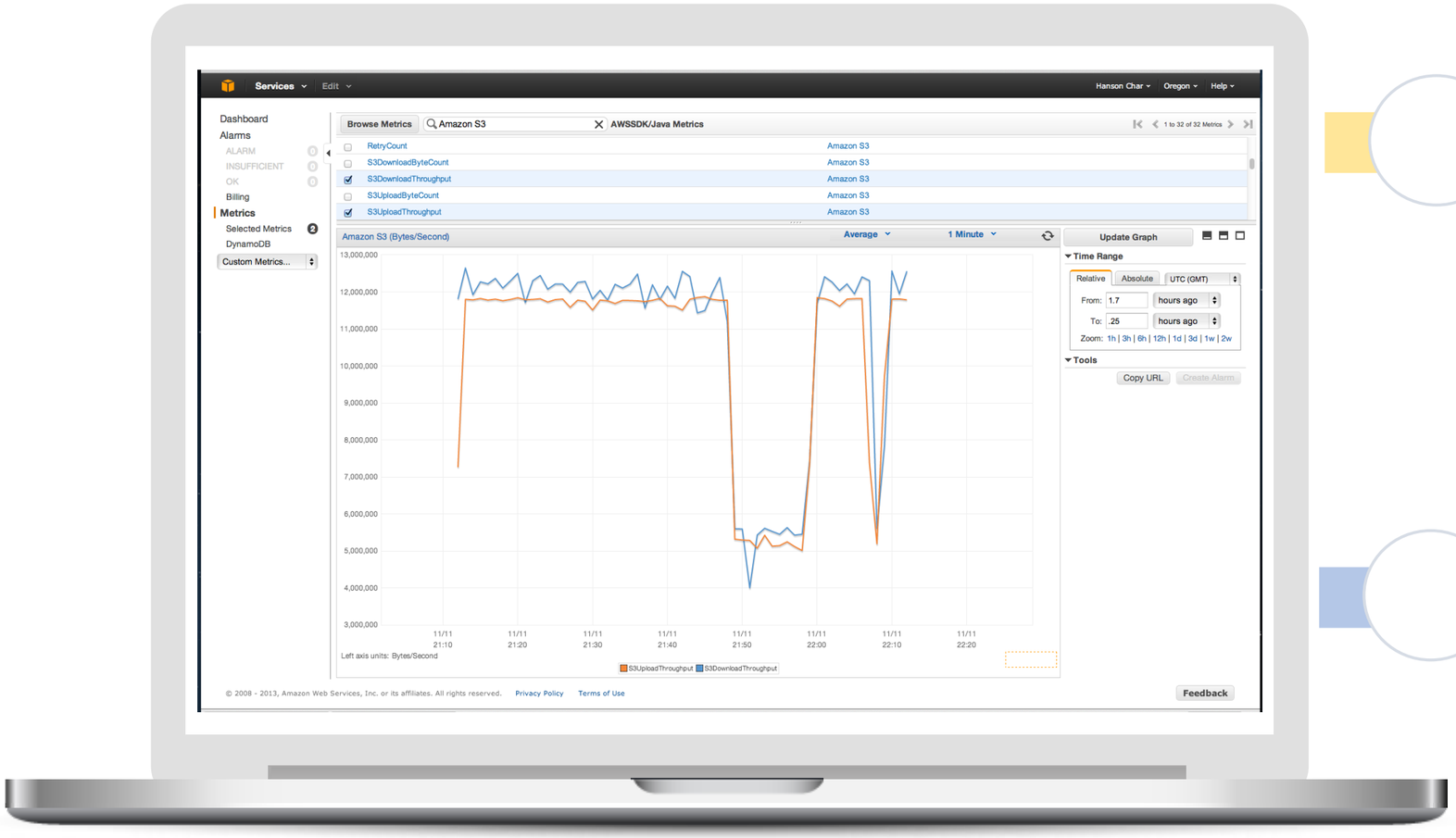
Bucket	Key	VersionId	IsLatest	IsDeleteMaker	Size	LastModifiedDate	Etag	StorageClass	MultipartUploaded	ReplicationStatus
example-bucket	object1			FALSE	2.4E+08	2016-08-11T01:19	e80d8eda4	STANDARD	TRUE	
example-bucket	object2			FALSE	0	2016-08-10T22:23	d41d8cd98	STANDARD	FALSE	
example-bucket	object3			FALSE	9	2016-08-10T20:18	9090441e4	STANDARD_IA	FALSE	
example-bucket	object4			FALSE	9	2016-08-10T20:36	9090441e4	STANDARD_IA	FALSE	
example-bucket	object1			FALSE	22	2016-08-10T20:35	9090441e4	STANDARD	FALSE	
example-bucket	object1			FALSE		2016-08-10T20:34	9090441e4	REDUCED_RED	FALSE	
example-bucket	object1			FALSE		2016-08-10T21:13	9090441e4	GLACIER	FALSE	

Export the object and the metadata of the object using the Inventory feature

Manage all S3 storage as csv files

You can store these files in another bucket

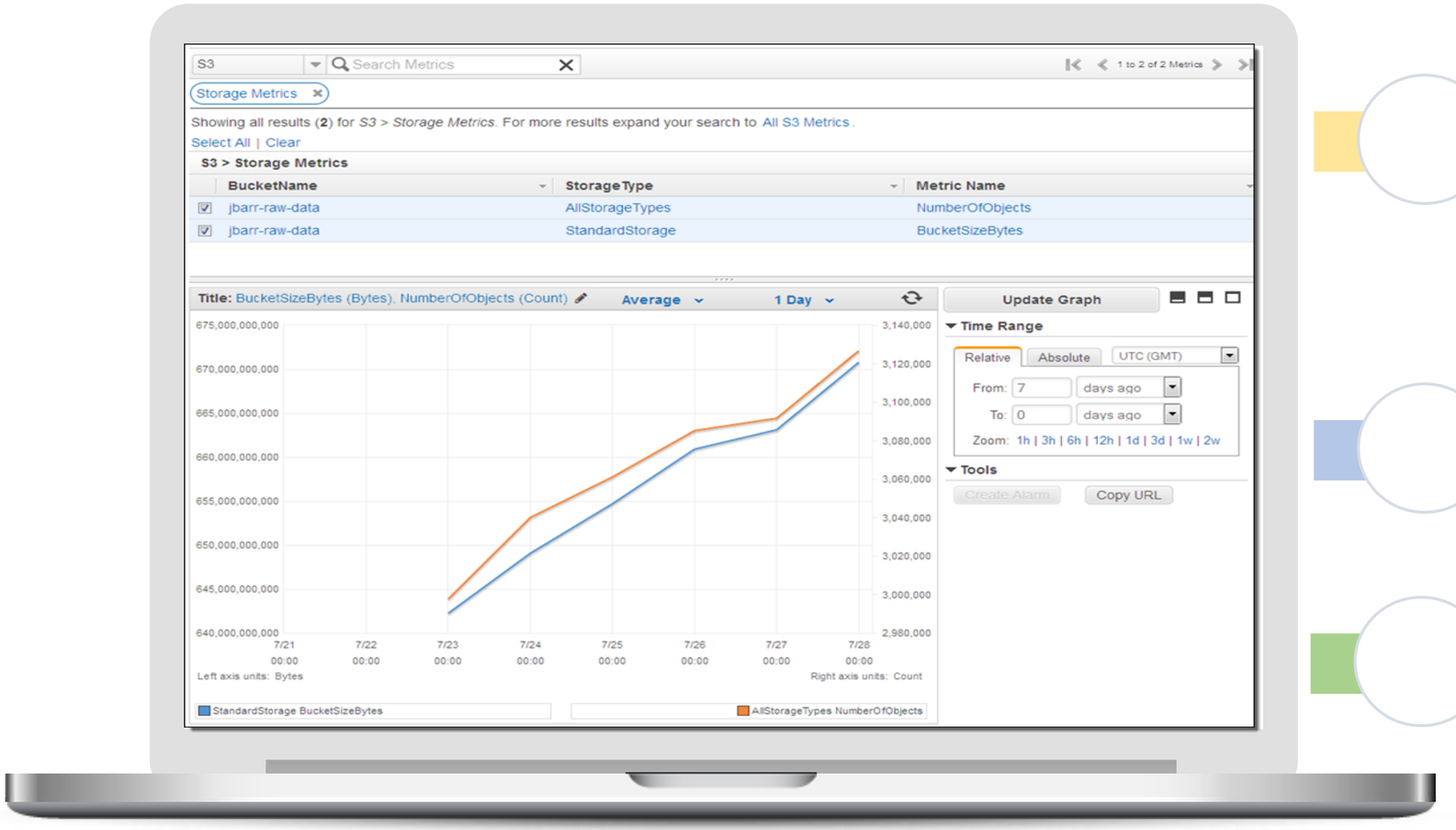
CloudWatch Metrics with S3



13 new CloudWatch metrics for S3 to monitor your bucket and improve the performance of the applications that use S3

Metric types include
Daily Storage Metrics for buckets
and Request Metrics

CloudWatch Metric Types



Bucket Level Metrics include::

- Total bytes for standard storage
- Total bytes for standard-IA storage
- Total bytes for reduced redundancy storage
- Number of objects in a bucket

Alarm on S3 bucket::
Set threshold for the alarms

Daily Alarms:
Metrics released daily

Knowledge Check

KNOWLEDGE CHECK

Which of the following configurations is necessary to enable static website hosting?

a. Index document

b. Policy Document

c. Error Document

d. CORS

KNOWLEDGE CHECK

Which of the following configurations is necessary to enable static website hosting?

a. Index document

b. Policy Document

c. Error Document

d. CORS

The correct answer is **Index Document & Error Document**

Explanation : Index document is the default home page when someone is trying to reach your domain. When you configure a bucket for website hosting, you must map an index document with one of your files from the bucket. By default, Amazon returns HTML error document. If you want any custom error message, you must map an error document with one of your files from the bucket. Again, this is optional.

KNOWLEDGE CHECK	Storage class analysis observes the access patterns of a filtered object data set for 30 days or longer to gather enough information for the analysis.

a. True

b. False

c.

d.

The correct answer is **a**

True : Storage class analysis observes the access patterns of a filtered object data set for 30 days or longer to gather enough information for the analysis. After storage class analysis has gathered sufficient information, you'll see a message in the Amazon S3 console stating and classifying the objects

Practice Assignment: Perform Amazon S3 bucket operations.

Create a bucket, add object into a bucket, view the bucket contents, move the object, and then delete the object.

Create a S3 Bucket and perform Operations on the bucket

You are conducting research to come up with the most efficient way to handle the execution of a project for your company. You have a huge amount of documentation that needs to be stored in an S3 bucket and you need to perform bucket operations such as adding, viewing, moving, and deleting on it.

Prerequisites:
AWS Account

Task:
To create a bucket in S3 and perform bucket operations such as adding, viewing, moving, and deleting an object from the bucket.

Quiz

QUIZ 1	Does Amazon S3 Support block based storage?

a. Yes

b. No

QUIZ 1	Does Amazon S3 Support block based storage?

a. Yes

b. No

The correct answer is **NO**

Explanation: S3 supports Object based storage. It does not support block based storage.

QUIZ 2

Which consistency model does S3 follow for PUTS of new objects?

a. Read after write consistency

b. Write after read consistency

c. Eventual Consistency

d. No Consistency

QUIZ 2	Which consistency model does S3 follow for PUTS of new objects?

a. Read after write consistency

b. Write after read consistency

c. Eventual Consistency

d. No Consistency

The correct answer is **Read after write consistency**

Explanation: S3 allows you to retrieve objects immediately after creating them for Read after write Consistency for PUTS of new objects.

QUIZ 3	What is the data durability on RRS?

a. 99.9%

b. 999999999.99%

c. 99.99%

d. 99.90%

QUIZ 3	What is the data durability on RRS?

a. 99.9%

b. 999999999.99%

c. 99.99%

d. 99.90%

The correct answer is **99.99%**

Explanation: RSS provides a data Durability and availability of 99.99%

QUIZ 4	How can you delete a large number of objects?

a. Multi-Object Delete

b. Multi-Select and delete

c. Select all and delete

d. Multiple delete options

QUIZ

4

How can you delete a large number of objects?

a. Multi-Object Delete

b. Multi-Select and delete

c. Select all and delete

d. Multiple delete options

The correct answer is **Multi-Object Delete**

Explanation: Multi-Object Delete is used to delete a large number of objects.

QUIZ 5

How can data be replicated in different AWS Regions?

a. Cross-Region Replication

b. Cross-zonal Replication

c. Cross-Region Replace

d. Cross-Region Release

QUIZ 5	How can data be replicated in different AWS Regions?

a. Cross-Region Replication

b. Cross-zonal Replication

c. Cross-Region Replace

d. Cross-Region Release

The correct answer is **Cross-Region Replication**

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

QUIZ 6	Choose the correct statements.

a. One can associate multiple tags with single object

b. Once can associate single tag with multiple objects

c. Key value pairs are key sensitive

d. Tagging is optional with objects

QUIZ 6

Choose the correct statements.

a. One can associate multiple tags with single object

b. Once can associate single tag with multiple objects

c. Key value pairs are key sensitive

d. Tagging is optional with objects

The correct answer is **A, C, and D**

You can associate up to 10 tags with an object. Tags associated with an object must have unique tag keys.

A tag key can be up to 128 Unicode characters in length and tag values can be up to 256 Unicode characters in length.

Key and values are case sensitive.

QUIZ 7

What all is included in S3 inventory?

a. Bucket Name, Key name, and Version ID

b. Bucket Name, Key name, Version ID, size, and storage class

c. Only Bucket name

d. Only Bucket Name and Size

QUIZ 8

What all is included in S3 inventory?

a. Bucket Name, Key name, and Version ID

b. Bucket Name, Key name, Version ID, size, and storage class

c. Only Bucket name

d. Only Bucket Name and Size

The correct answer is

Key Takeaways

- S3 supports object sizes between 1 MB to 5 TB. Objects are the combination of a file and any Meta data that describes that file
- Amazon S3 bucket must have a unique name as it shares a global name space.
- S3 Standard Storage provides 99.99% availability and 99.999999999% durability of data.
- In Glacier, retrieval of data can take up to 5 hours and cannot be used when you need to retrieve data in less than 4 hours
- S3 supports Read after write consistency for PUTS of new objects in all regions. S3 supports Eventual consistency for overwrite PUTS and DELETES
- When the static website is enabled, the website endpoint of the bucket is bucket name.s3-website-AWS-region.amazonaws.com
- Cross-Origin Resource Sharing is a mechanism that allows a browser to show the resources of a domain that are referenced or used in a different domain

This concludes ‘Simple Storage Service S3.’

The next lesson is ‘S3 LAB.’