

What Is AWS S3?



What is S3?

What Is S3?

Simple Storage Service







Data Availability

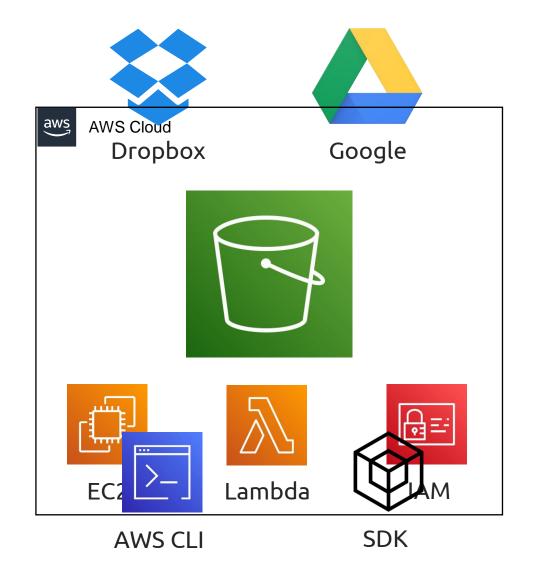


Security



Performance

What Is S3?







What Is S3?

Object-based Storage

File-based Storage





Block-based Storage



S3 Use Cases

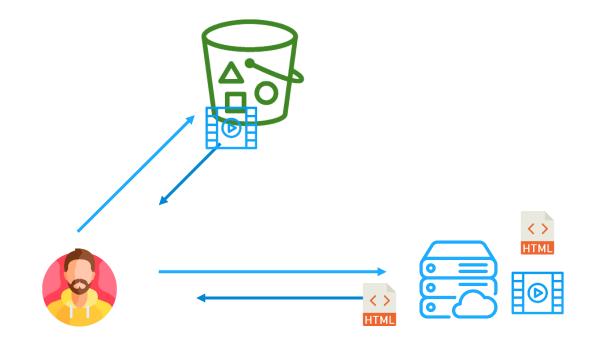






S3 Use Case





Bucket



Bucket



Bucket







Objects

Objects are files that are uploaded to S3

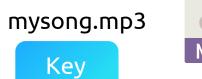


An object has:

Key – The file name

Value – File data

VersionID/Metadata/Other information





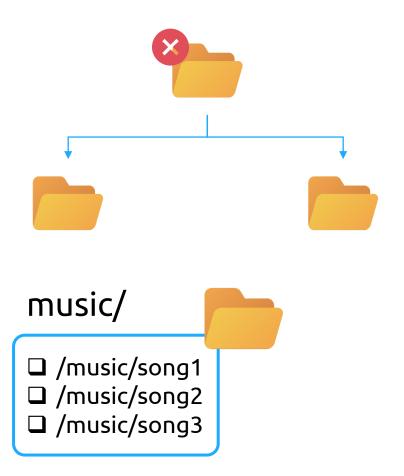


S3 File Structure

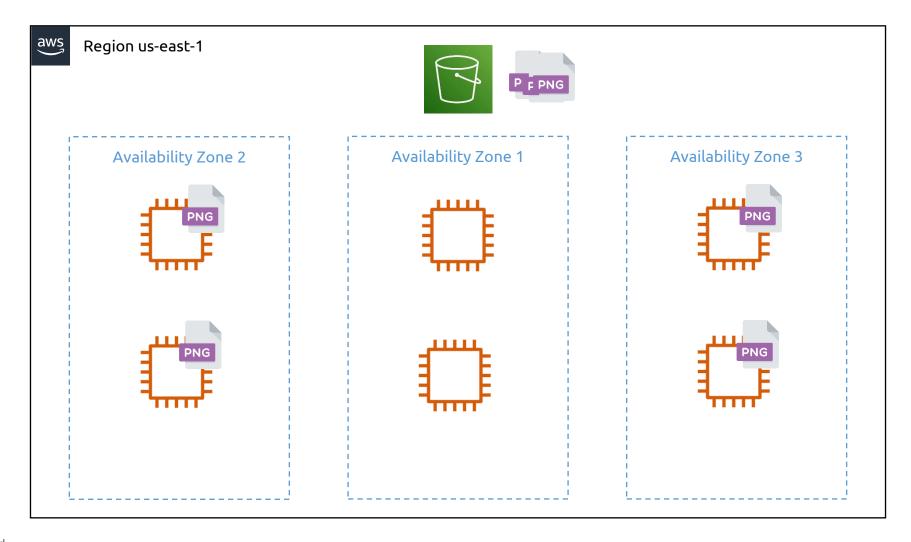
S3 Buckets have a flat file structure



- ☐ File1.txt
- ☐ File2.txt
- ☐ File3.txt
- ☐ File4.txt



Availability



S3 Bucket Names

Important

S3 bucket names must be unique globally across all AWS accounts



https://kodekloud.s3.amazonaws.com
Bucket Name



S3 Restrictions



S3 can handle unlimited number of objects



Maximum size of a single file is 5TB



An AWS account supports 100 buckets by default, but this number can be increased to 1,000 by requesting a service limit increase



Storage Classes

Storage Classes



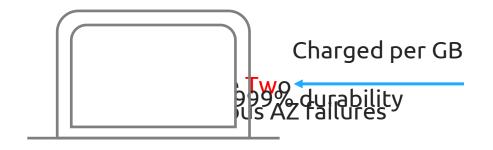


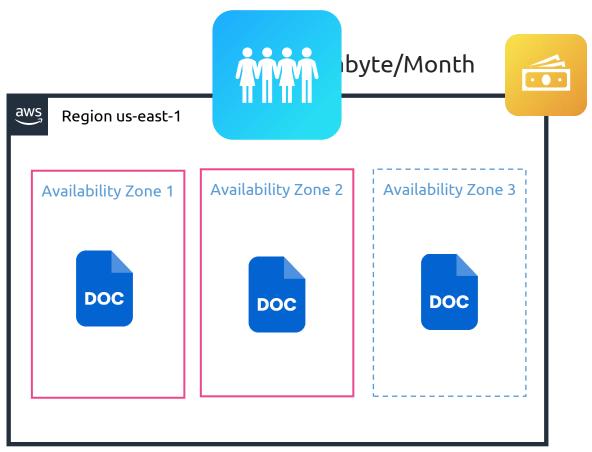




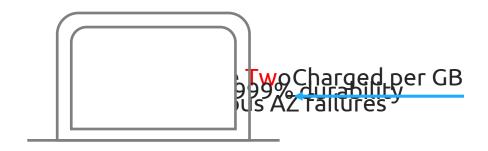
© Copyright KodeKloud

S3 Standard (default)



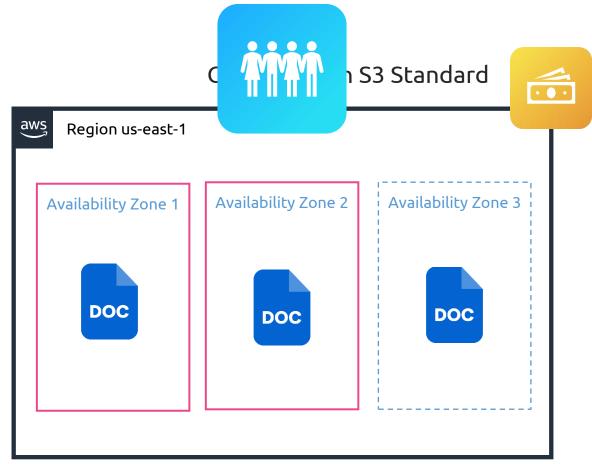


S3 Standard-IA



Has a retrieval fee

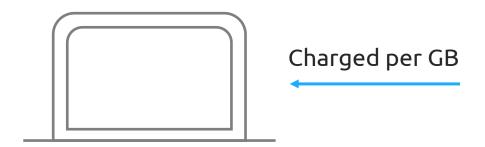
Minimum duration charge of 30 days Minimum size charge of 128 KB per object



S3 One Zone-IA

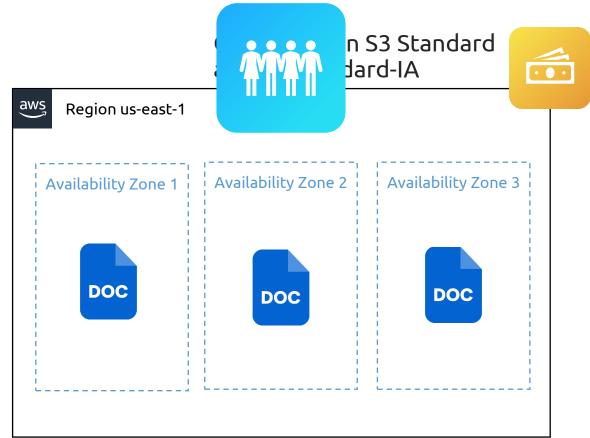
Note

Designed for IA data Not required to handle AZ failure Replication still occurs within AZ



Has a retrieval fee

Minimum duration charge of 30 days Minimum size charge of 128 KB per object

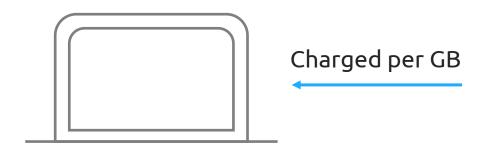


S3 Glacier-Instant

Summary

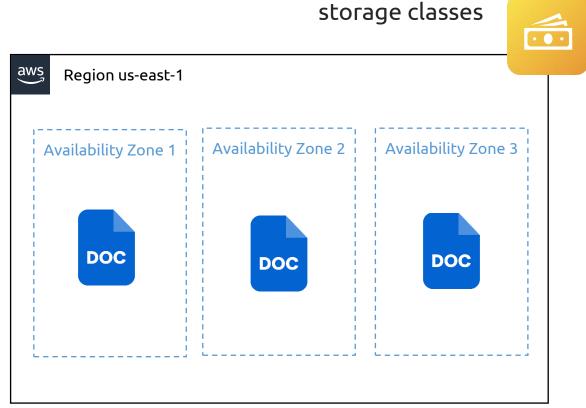
Wewy constappation for rarely accessed data Stightepeeloien a hoests that of S3

Longer minimum duration



Has a retrieval fee

Minimum duration charge of 90 days Minimum size charge of 128 KB per object

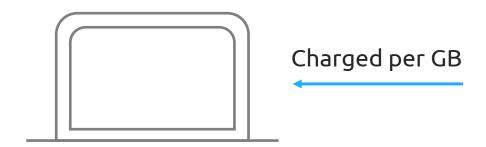


Cheaper than the mentioned

S3 Glacier-Flexible

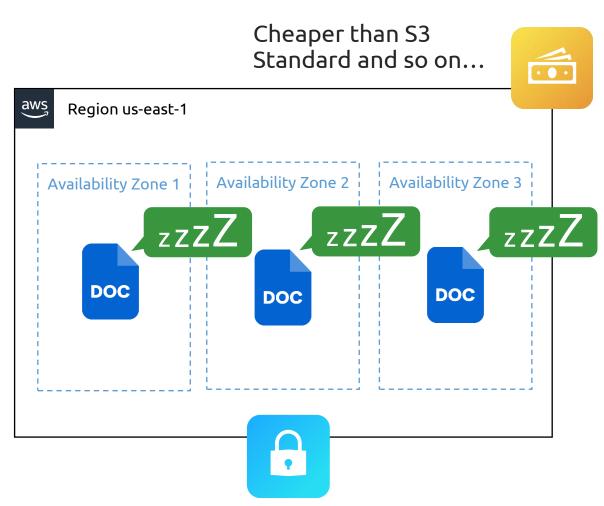
Options

Expedited: 1–5 Minutes
During retrieval, objects are stored
Standard-IA class temporarily
Bulk: 5–12 Hours



Has a retrieval fee

Minimum duration charge of 90 days Minimum size charge of 40 KB per object

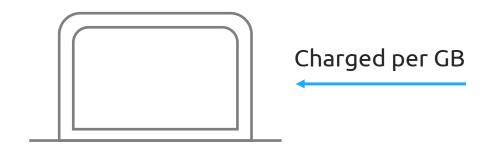


S3 Glacier Deep Archive

Options

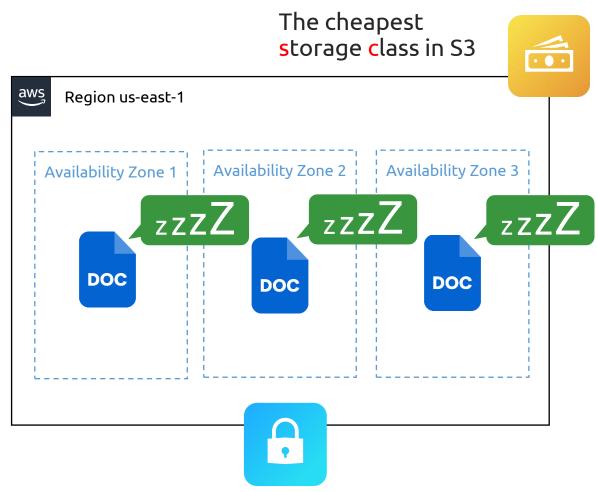
Standard: 12 Hours

Bulk : 48 Hours



Has a retrieval fee

Minimum duration charge of 180 days Minimum size charge of 40 KB per object



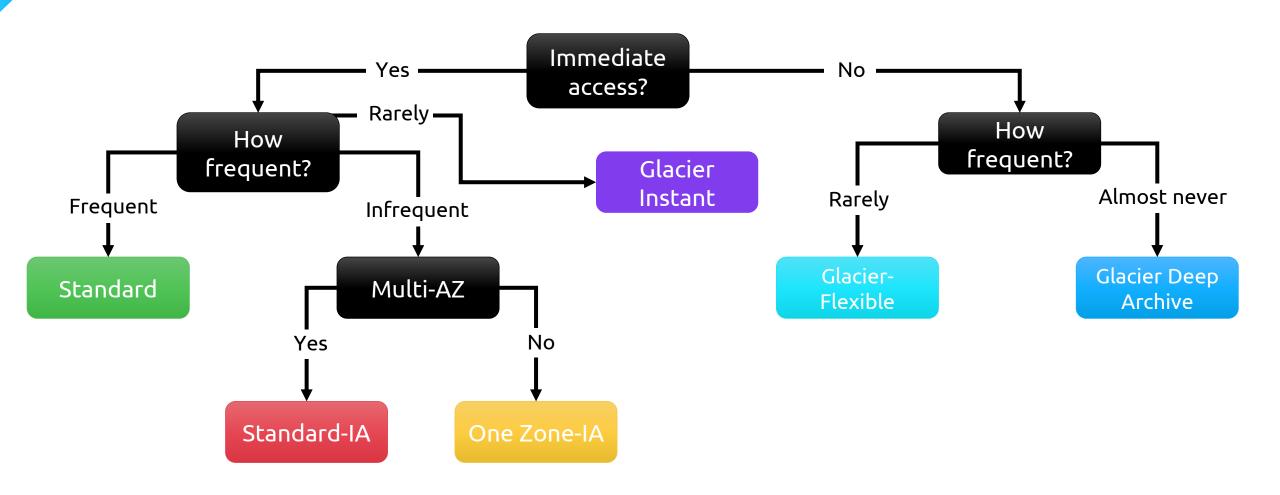
S3 Intelligent-Tiering

01

Automatically reduces storage costs by intelligently moving data to the most cost-effective access tier

02

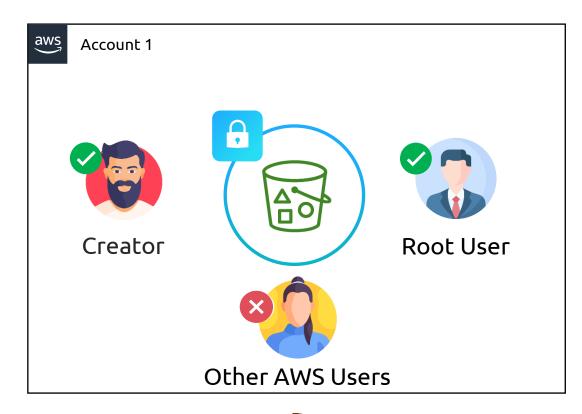
Apart from the cost of a storage class an object gets assigned to, all objects will also incur a monitoring/automation cost per 1,000 objects

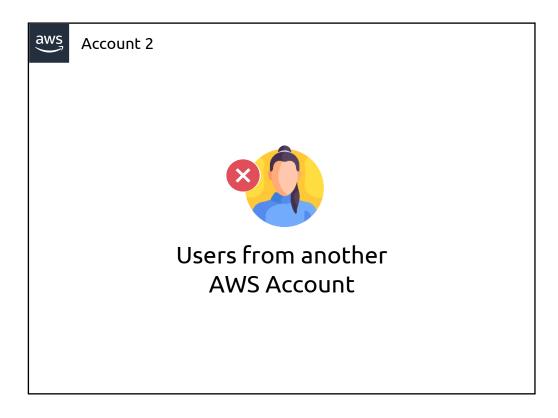




ACLs and Resource Policies

S3 Access







S3 Bucket Policies

Resource Policy

Determines who has access to an S3 resource

S3 Bucket Policy

Determines who can have access to the bucket and what operations they can perform

S3 Bucket Policies

```
S3 Bucket policies
                                                                          are written in JSON
Terminal
"Version": "2012-10-17",
                                                                  Version
 "Statement": [
         "Sid": "AllowRule"
                                                                  Sid
                                                                   Principal
         "Principal": {
             "AWS":
                 "arn:aws:iam::111122223333:user/JohnDoe"
         "Effect" · "Allow"
                                                                   Effect
         "Action": ["s3:GetObject"],
                                                                   Action
         "Resource": |"arn:aws:s3:::DOC-EXAMPLE-BUCKET/*"
                                                                   Resource
```

```
"Version": "2012-10-17",
                             "Statement": [
                                      "Sid" · "AllowAll"
Applies to all users
                                      "Principal": "*",
                                      "Effect": "Allow",
                                      "Action": ["s3:GetObject"],
                                      "Resource": ["arn:aws:s3:::DOC-EXAMPLE-BUCKET/*"]
                                  },
                                      "Sid": "DenyDaisy",
                                      "Principal": {
                                          "AWS": [
                                              "arn:aws:iam::666438:user/DaisyM"
                                      "Effect": "Deny",
                                      "Action": ["s3:GetObject"],
                                      "Resource": ["arn:aws:s3:::DOC-EXAMPLE-BUCKET/*"]
```

```
Terminal
   "Version": "2012-10-17",
    "Statement": [
            "Sid": "AllowDaisy",
            "Principal": {
                "AWS": [
                    "arn:aws:iam::666438:user/DaisyM"
            "Effect": "Allow",
            "Action": ["s3:GetObject"],
            "Resource": ["arn:aws:s3:::DOC-EXAMPLE BUCKET/media/*"]
```

```
Terminal
"Id": "PolicyId2",
"Version": "2012-10-17",
"Statement": [
        "Sid": "AllowIP",
        "Effect": "Allow",
        "Principal": "*",
        "Action": "s3:*",
        "Resource":
            "arn:aws:s3:::DOC-EXAMPLE-BUCKET",
            "arn:aws:s3:::DOC-EXAMPLE-BUCKET1/*"
        ]
        "Condition": {
            "IpAddress": {
                "aws:SourceIp": [
                    "192.0.2.0/24",
```

```
Terminal
"Id": "PolicyId2",
"Version": "2012-10-17",
"Statement": [
        "Sid": "AllowIP",
         "Effect": "Allow",
         "Principal": "*",
         "Action": "s3:*",
         "Resource": [
             "arn:aws:s3:::DOC-EXAMPLE-BUCKET",
             "arn:aws:s3:::DOC-EXAMPLE-BUCKET1/*"
         ]
         "Condition": {
             "StringEquals": {
                 "s3:prefix": ["audio/", "video/"],
                 "s3:delimiter": ["/"]
```

Block Public Access

✓ Block all public access

Turning this setting on is the same as turning on all four settings below. Each of the following settings are independent of one another.

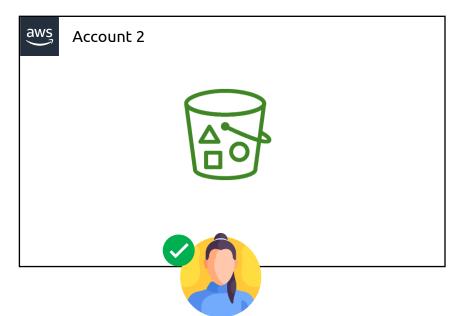
- Block public access to buckets and objects granted through *new* access control lists (ACLs)

 S3 will block public access permissions applied to newly added buckets or objects, and prevent the creation of new public access ACLs for existing buckets and objects. This setting doesn't change any existing permissions that allow public access to S3 resources using ACLs.
- Block public access to buckets and objects granted through *any* access control lists (ACLs)

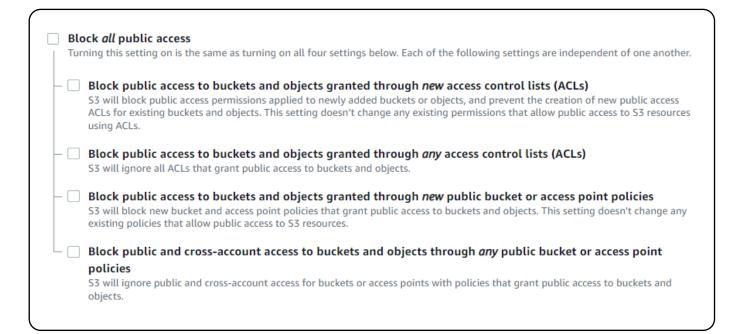
 S3 will ignore all ACLs that grant public access to buckets and objects.
- Block public access to buckets and objects granted through *new* public bucket or access point policies

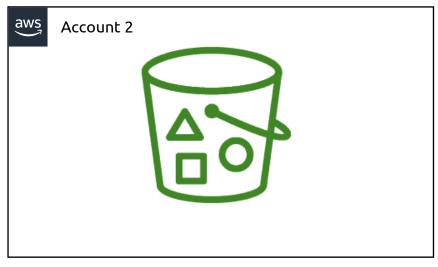
 S3 will block new bucket and access point policies that grant public access to buckets and objects. This setting doesn't change any existing policies that allow public access to S3 resources.
- ☑ Block public and cross-account access to buckets and objects through any public bucket or access point policies
 - S3 will ignore public and cross-account access for buckets or access points with policies that grant public access to buckets and objects.

```
"Id": "PolicyId2",
"Version": "2012-10-17",
"Statement": [
        "Sid": "AllowAll",
        "Effect": "Allow",
        "Principal": "*",
        "Action": "s3:*",
         Resource": [
            "arn:aws:s3:::DOC-EXAMPLE-BUCKET1/*"
        ],
```

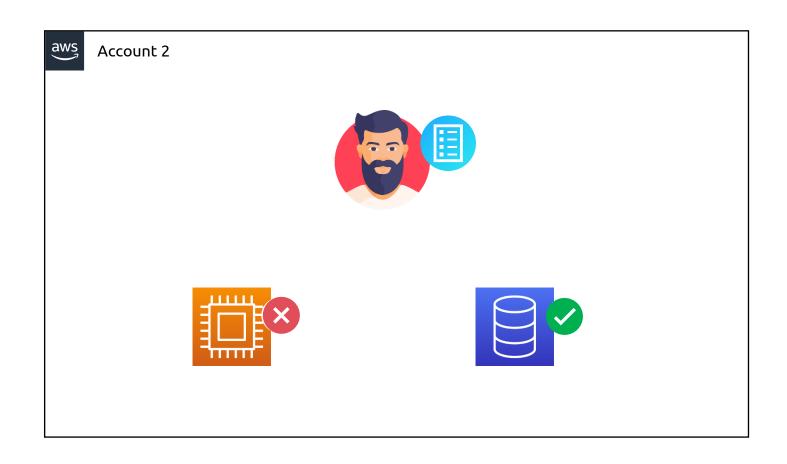


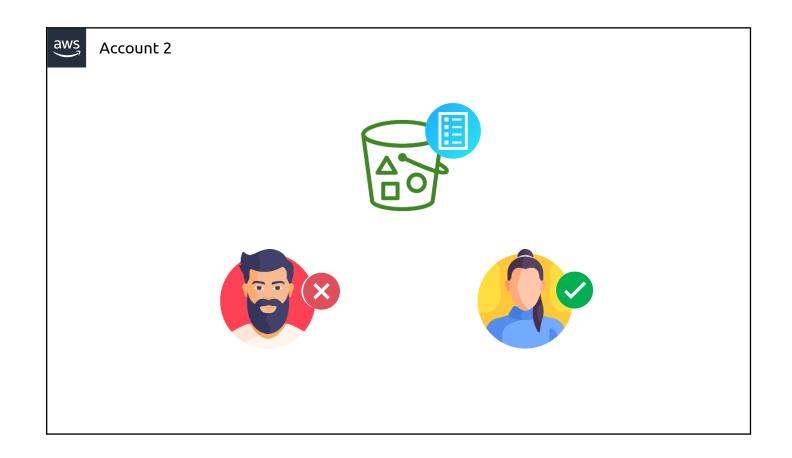
Anonymous/Public











IAM Policy



Can only be applied to authenticated AWS users

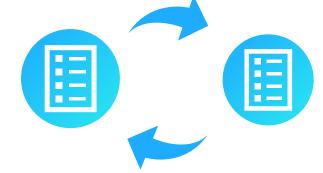
Cannot be applied to anonymous users

Resource Policy



Since the policy is applied to the resource, rules can be added for anonymous/public users

























S3 ACLs



Have a legacy access control mechanism that predates IAM



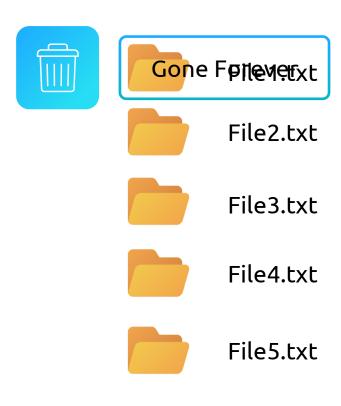
Are inflexible and provide only a limited set of rules

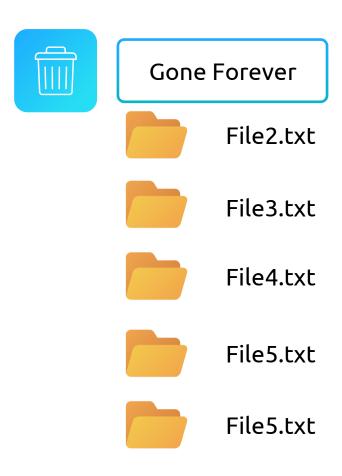
Cannot be applied to a group of objects

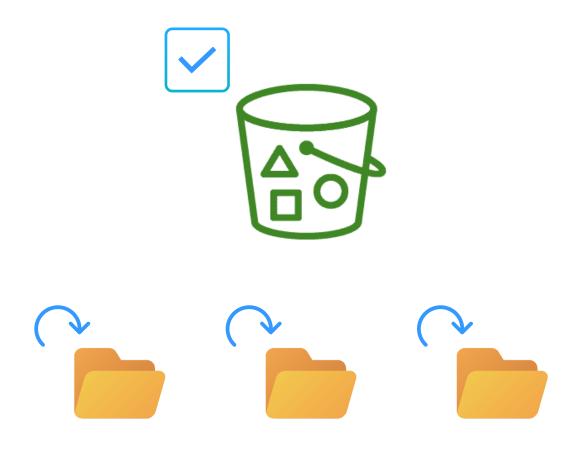
Can be used but it is not recommended

ACL permissions		
Permission	When granted on a bucket	When granted on an object
READ	Allows grantee to list the objects in the bucket.	Allows grantee to read the object data and its metadata
WRITE	Allows grantee to create new objects in the bucket. For the bucket and object owners of existing objects, also allows deletions and overwrites of those objects.	Not applicable
READ_ACP	Allows grantee to read the bucket ACL	Allows grantee to read the object ACL
WRITE_ACP	Allows grantee to write the ACL for the applicable bucket	Allows grantee to write the ACL for the applicable object
FULL_CONTROL	Allows grantee the READ, WRITE, READ_ACP, and WRITE_ACP permissions on the bucket	Allows grantee the READ, READ_ACP, and WRITE_ACP permissions on the object



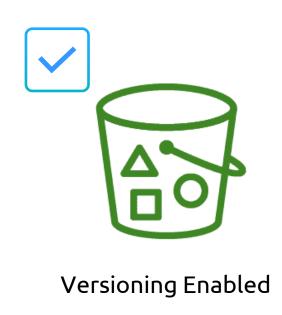


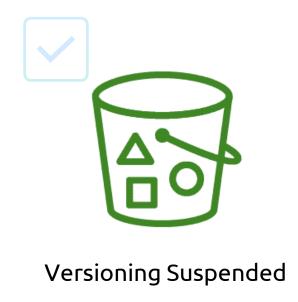




Three States







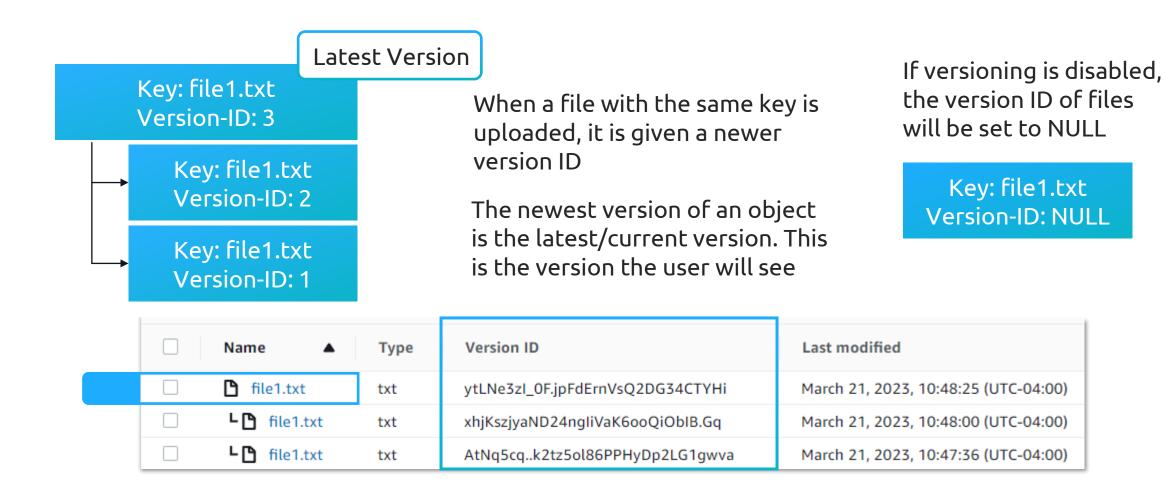
How Versioning Works

Key: file1.txt Version-ID: 3

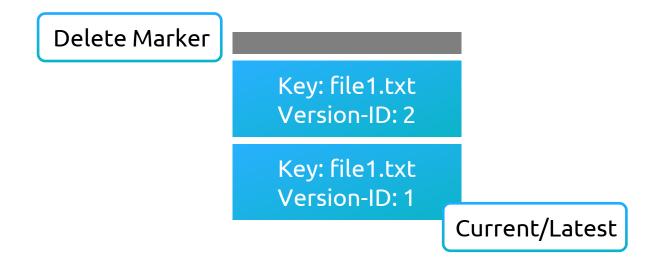
Key: file1.txt Version-ID: 2

Key: file1.txt Version-ID: 1 When a file with the same key is uploaded, it is given a newer version ID

How Versioning Works



Deleting File Versions



Versioning Price

Key: file1.txt Version-ID: 2 Size: 15 GB

Key: file1.txt Version-ID: 1 Size: 10 GB



Version Suspending

Key: file1.txt
Version-ID: 3

Key: file1.txt
Version-ID: 2

Key: file1.txt
Version-ID: 1

Versioning Enabled

Key: file1.txt
Version-ID: NULL

Key: file1 txt

Key: file1.txt Version-ID: 3

Key: file1.txt Version-ID: 2

Key: file1.txt Version-ID: 1 Key: file1.txt Version-ID: NULL

> Key: house.png Version-ID: NULL

Suspended Versioning

Multi-Factor Authentication (MFA) Delete



Multi-Factor Authentication (MFA) Delete





Note

When the feature is enabled, MFA is required to change the versioning state of the bucket

MFA is required to delete versions and can only be enabled using CLI





S3 Standard

S3 Standard-IA

S3 One Zone-IA

S3 Glacier Instant Retrieval

S3 Glacier Flexible Retrieval

S3 Glacier Deep Archive

After 30 Days



S3 Standard

S3 Standard-IA

S3 One Zone-IA

S3 Glacier Instant Retrieval

S3 Glacier Flexible Retrieval

S3 Glacier Deep Archive

S3 Standard

S3 Standard-IA

S3 One Zone-IA

S3 Glacier Instant Retrieval

S3 Glacier Flexible Retrieval

S3 Glacier Deep Archive

After 90 Days



After 1 Year



S3 Standard

S3 Standard-IA

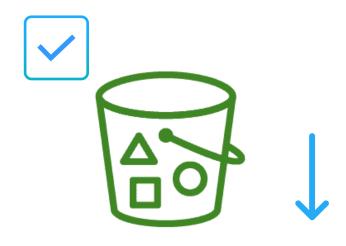
S3 One Zone-IA

S3 Glacier Instant Retrieval

S3 Glacier Flexible Retrieval

S3 Glacier Deep Archive





S3 Standard

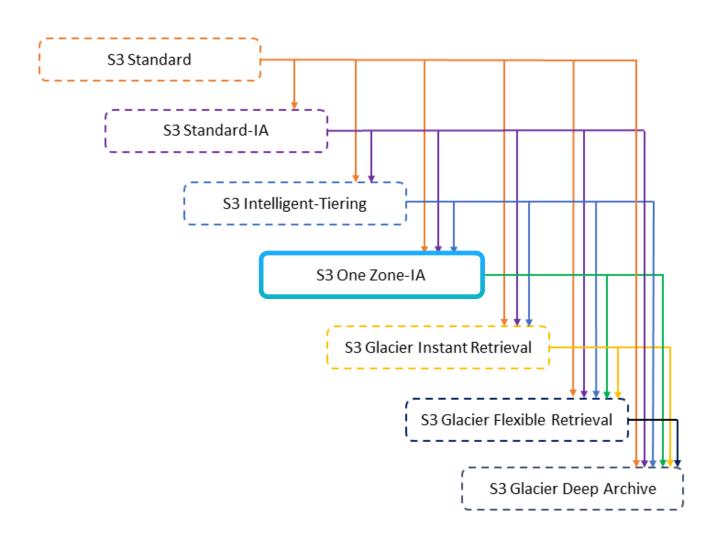
S3 Standard-IA

S3 One Zone-IA

S3 Glacier Instant Retrieval

S3 Glacier Flexible Retrieval

S3 Glacier Deep Archive



Note

For the following transitions, S3 does not transition objects smaller than 128 KB

S3 Standard

S3 Standard-IA



S3 Intelligent-Tiering

S3 Glacier Instant Retrieval

Note

For the following transitions, S3 does not transition objects smaller than 128 KB

S3 Standard



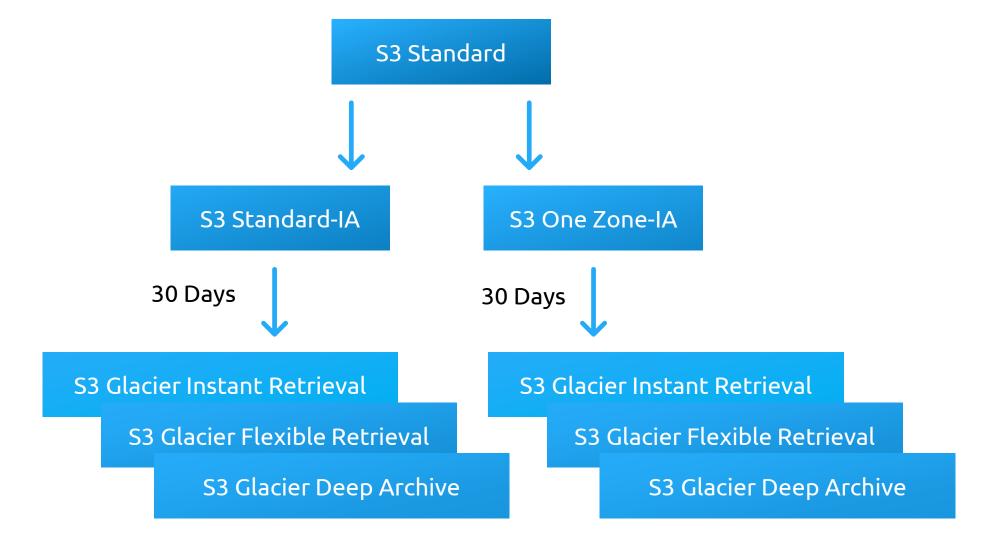
S3 Standard-IA

S3 One Zone-IA

S3 Standard 30 Days S3 Standard-IA

S3 Standard-IA

S3 One Zone-IA





Static Hosting







Note

A website is just an HTML file



Structure/Content of the website



Adds colors and provides visual elements



Adds dynamic functionality



Images/Videos/Audio





Allows access to website files through HTTP



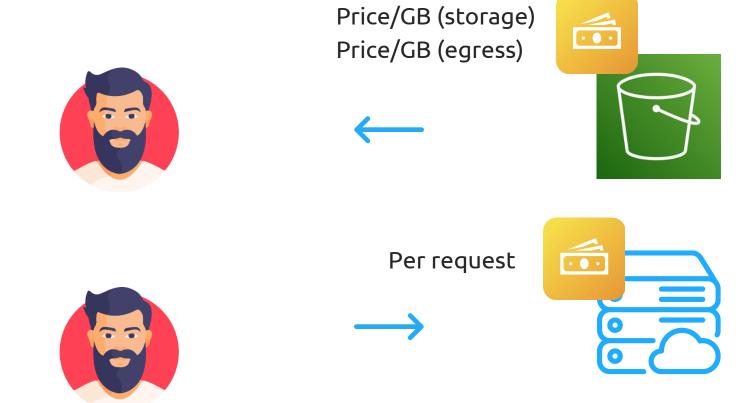
S3 gives the URL through which you can access the website

Note

It is used only for static websites

To customize a domain for your website, the bucket must follow a specific format

Pricing



Pricing

	PUT, COPY, POST, LIST requests (per 1,000 requests)	GET, SELECT, and all other requests (per 1,000 requests)	Lifecycle Transition requests into (per 1,000 requests)	Data Retrieval requests (per 1,000 requests)	Data retrievals (per GB)
S3 Standard	\$0.005	\$0.0004	n/a	n/a	n/a
S3 Intelligent - Tiering *	\$0.005	\$0.0004	\$0.01	n/a	n/a
Frequent Access	n/a	n/a	n/a	n/a	n/a
Infrequent Access	n/a	n/a	n/a	n/a	n/a
Archive Instant	n/a	n/a	n/a	n/a	n/a
Archive Access, Standard	n/a	n/a	n/a	n/a	n/a
Archive Access, Bulk	n/a	n/a	n/a	n/a	n/a
Archive Access, Expedited	n/a	n/a	n/a	\$10.00	\$0.03
Deep Archive Access, Standard	n/a	n/a	n/a	n/a	n/a
Deep Archive Access, Bulk	n/a	n/a	n/a	n/a	n/a

[©] Copyright KodeKloud

Custom Domain Name

http://bucketname.s3-website-<region-name>.amazonaws.com



http://bucketname.s3-website-<region-name>.amazonaws.com





Custom Domain Name













Prevents permanently deleting or overwriting data

Meets regulatory requirements

Enforces write-once-readmany (WORM) model







Object Lock Modes

Governance Mode



Compliance Mode



Legal Hold



Use Case

Documents used during active litigation

To Use

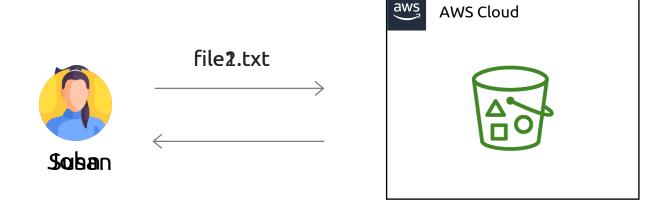
Versioning must be enabled on the bucket

Object locking on the bucket must be enabled



Access Logs

Access Logging



John [06/Feb/2019:00:00:38 +0000] GET /file1.txt Susan [08/Feb/2019:00:00:38 +0000] GET /file2.txt

Access Logs

Uses

For security and auditing access

For better understanding of how the buckets are used and thus finding the right storage class

Fact

Stored in a separate S3 Bucket

Log Format

Access logs contain the following details:

Bucket Owner

Bucket Name

Timestamp

Remote IP

Requester (User ID of the requester)

Request ID (Unique ID for each request)

Operation (GET/PUT/DELETE)

Key

Version ID

Status Code

Моге...



Inventory Reports

Inventory Reports









Apache Parquet

Inventory Reports

Information included in Inventory Reports:

Bucket Name

Key

Version ID

Size

Last Modified Date

Storage Class

Replication Status

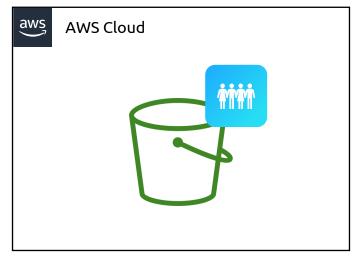
Encryption Status

Object Lock Status

More...

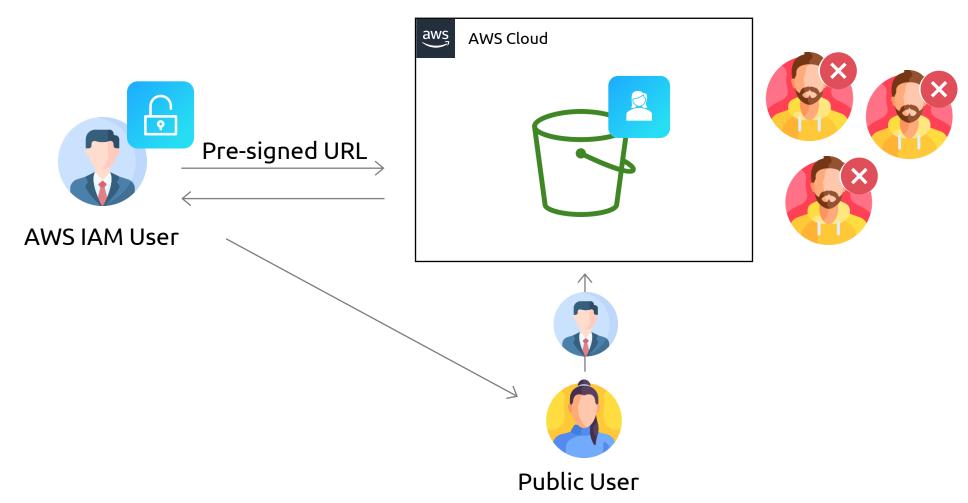




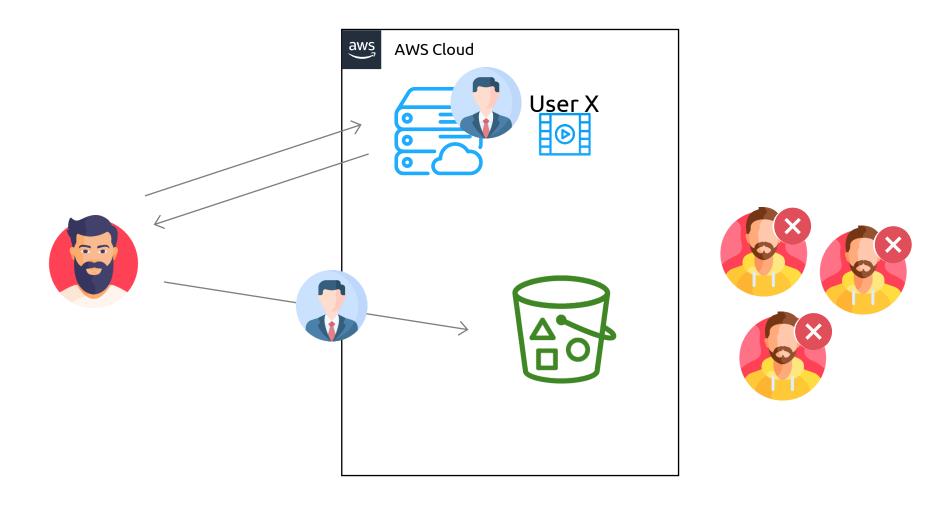


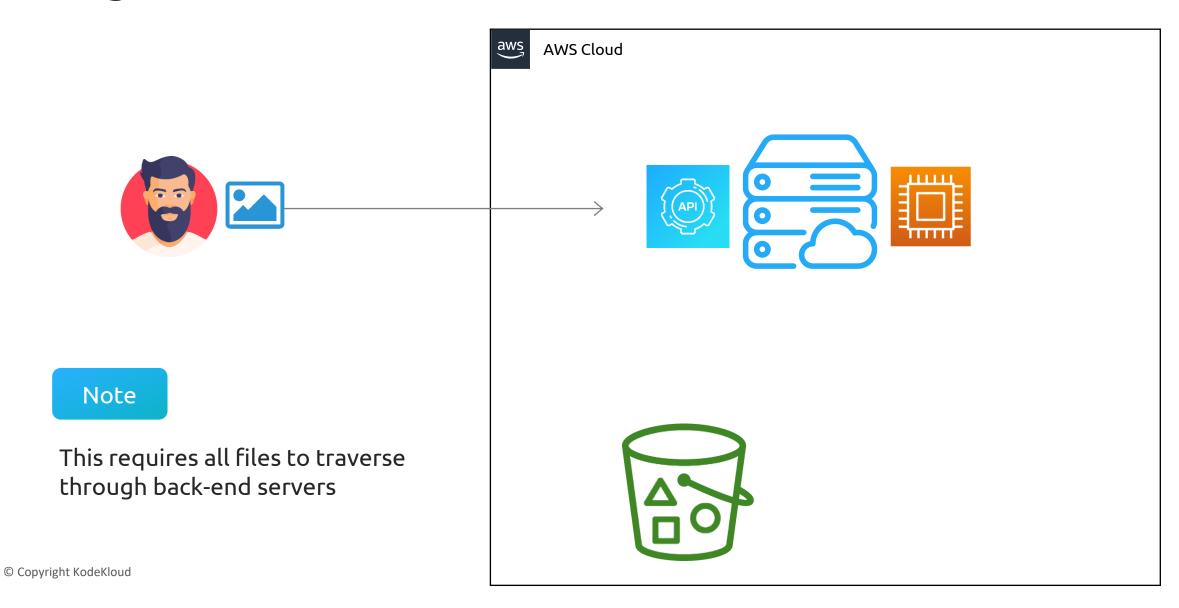


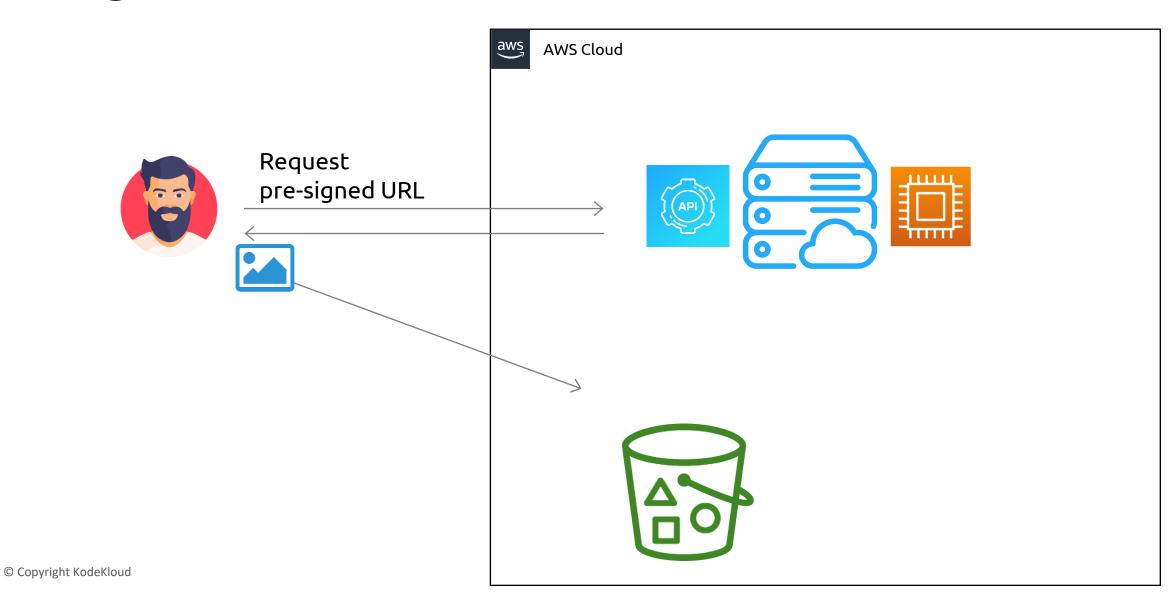




Pre-Signed URL Use Case



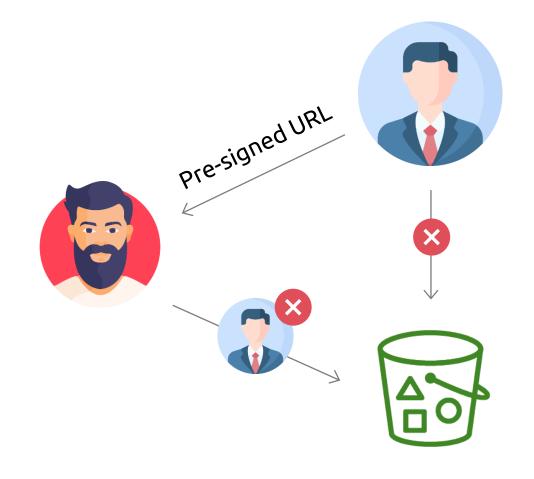




Note

When creating pre-signed URLs, an expiration date must be provided Expiration duration of maximum 7 days using an IAM user is provided If an IAM user does not have access to an S3 bucket, a pre-signed URL can still be generated using that account

The pre-signed URL does not give you access to a bucket; however, it allows you to send a request to S3 as the user that generated the URL



IAM User with no access to S3 Bucket



Replication

S3 Replication



Uses of Replication



Ensures data protection



Is a compliance requirement



Helps store data closer to users



Keeps data near their servers

Types of Replication

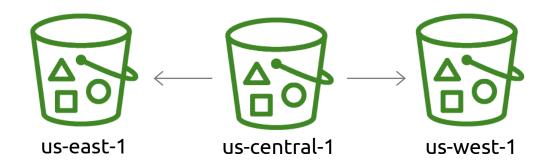
Same-Region Replication (SRR)



Cross-Region Replication (CRR)

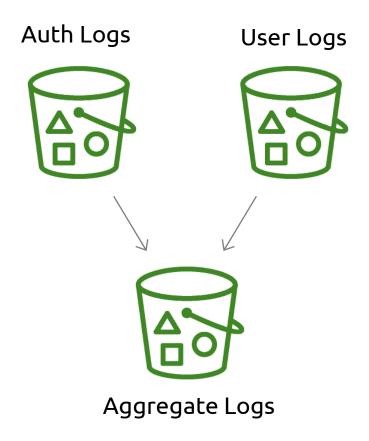


Multi-Destination Replication

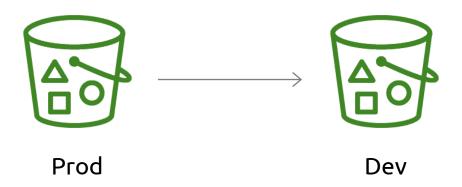


Same-Region Replication Use Cases

Aggregate logs in to a single bucket



Live replication between production and test environments



Cross-Region Replication Use Cases



Compliance requirements



Minimizing latency



Operational efficiency

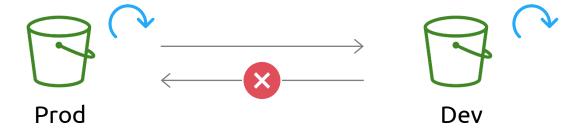
Bidirectional Replication

By default, replication is a one-way process



Bidirectional Replication

Bidirectional replication can be configured manually



Replication Requirements



Versioning on both buckets



AWS S3 permission to replicate



S3 object lock on both buckets

Other Details

Only objects created after enabling replication will be replicated (For any previous objects, a batch replication job can be run to get them replicated)

Objects with encryption enabled (SSE-C, SSE-S3, or SSE-KMS) will be replicated

Objects in the S3 Glacier Flexible Retrieval and S3 Glacier Deep Archive storage classes do not get replicated

Object metadata and tags will be replicated

After replication, objects can be moved from one storage class to another

Other Details

S3 Standard



Source

S3 Standard-IA



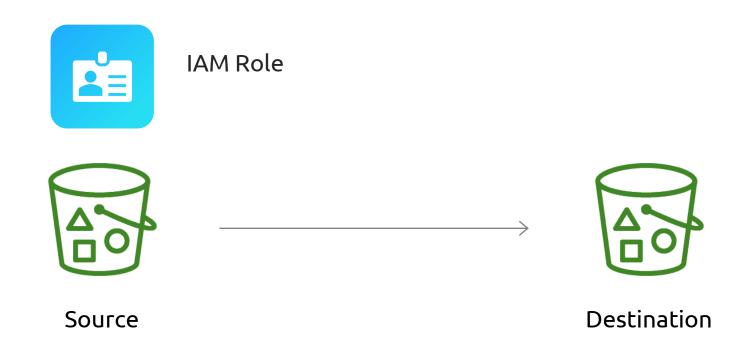
Destination

Deletion with Replication

By default, delete markers do not get replicated (Enable delete marker replication)

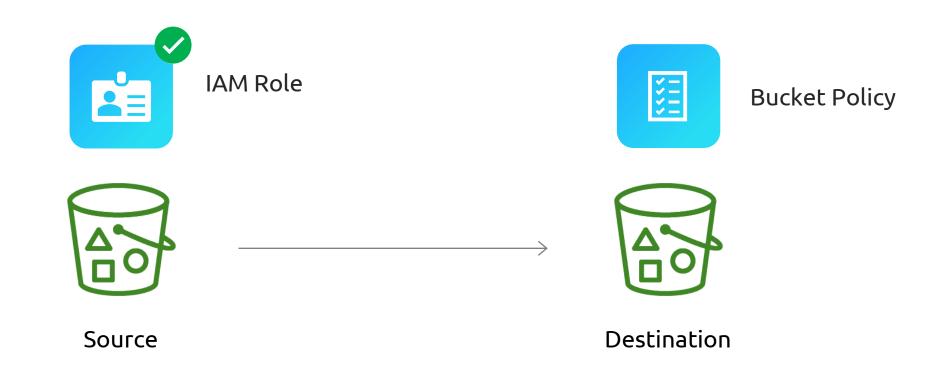
If a specific version of an object is deleted on the source bucket, that version will not be deleted on the destination bucket (This protects data from malicious deletions)

Permissions



Replication within the same AWS account

Permissions



Replication to a different AWS account

Replication Time Control (RTC)



Replication does not occur instantly and can take some time before a file shows up in the destination bucket



S3 Replication Time Control (S3 RTC) replicates objects within 15 minutes

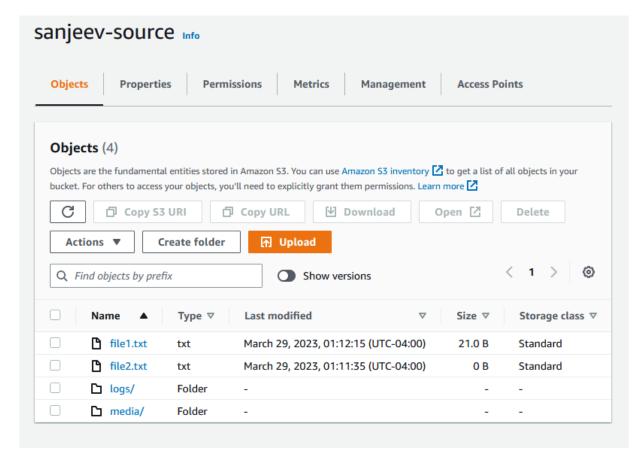
(Great for meeting compliance or business requirements)



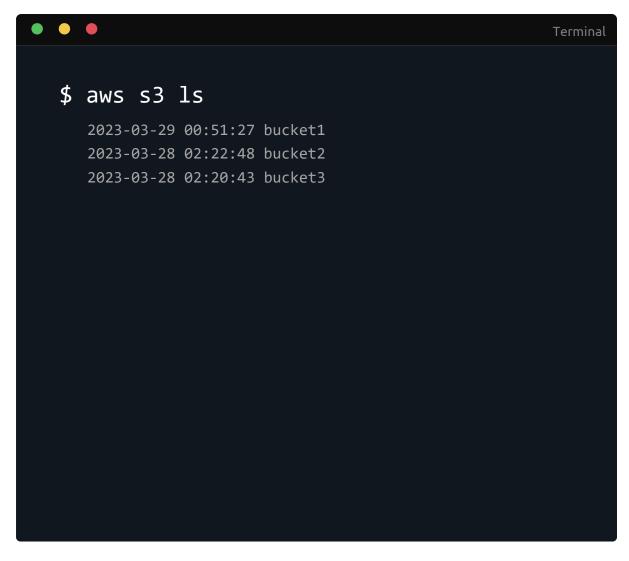
Creating a Bucket Using AWS CLI

AWS CLI

Console



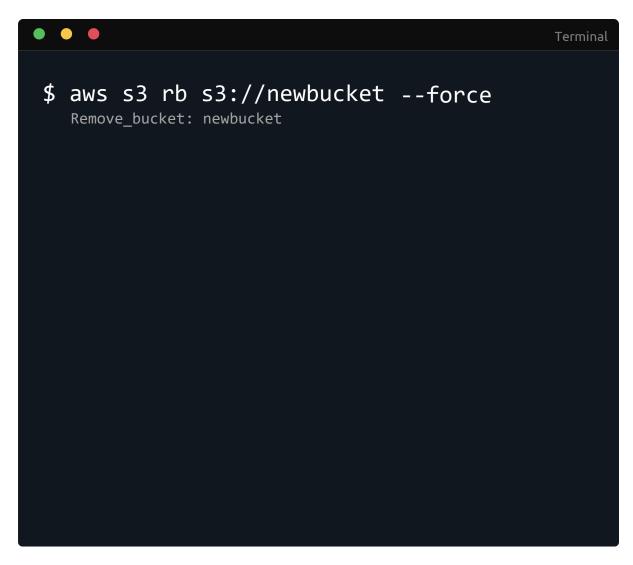
List All S3 Buckets



Create New Bucket

```
Terminal
$ aws s3 mb s3://newbucket --region us-east-1
  make_bucket: newbucket
```

Delete Bucket



Note

AWS CLI will not delete buckets that contain files

Add the --force flag to delete buckets that contain files

List Objects in Bucket

```
Terminal
 $ aws s3 ls s3://newbucket
                             PRE logs/
                             PRE media/
                              0 file1.txt
   2023-03-29 01:38:08
   2023-03-29 01:38:09
                              0 file2.txt
$ aws s3 ls s3://newbucket --recursive
   2023-03-29 01:38:08
                              0 file1.txt
                              0 file2.txt
   2023-03-29 01:38:09
   2023-03-29 01:38:09
                              0 logs/log1
   2023-03-29 01:38:09
                              0 logs/log2
   2023-03-29 01:38:10
                           24599 media/images/image1.png
                           21420 media/images/image2.png
   2023-03-29 01:38:10
```

Note

Use the --recursive flag to recursively print all files within all directories

Copy Files

Terminal \$ aws s3 cp file1 s3://newbucket upload: ./file1 to s3://newtbucket/file1 \$ aws s3 cp s3://newbucket/file1 /tmp download: s3://sanjeeevtestbucket/file1.txt to /tmp/file1 \$ aws s3 cp s3://bucket1/file1 s3://bucket2 copy: s3://bucket1/file1.txt to s3://bucket2/file1.txt

From local machine to Bucket

From S3 to local machine

From S3 Bucket to another Bucket

Delete Objects

```
Terminal
$ aws s3 ls s3://bucket
  2023-04-03 07:42:25
                         458 404.html
  2023-04-03 07:42:25 414 index.css
  2023-04-03 07:42:25
                    1126 index.html
$ aws s3 rm s3://bucket/404.html
  delete: s3://bucket/404.html
$ aws s3 ls s3://bucket
  2023-04-03 07:42:25
                         414 index.css
  2023-04-03 07:42:25 1126 index.html
```

Copy Folder

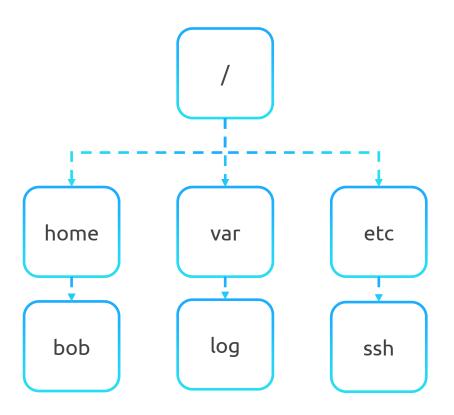
```
Termina
$ aws s3 cp media/ s3://newbucket --recursive
$ aws s3 cp s3://newbucket/media/ /tmp/ --recursive
$ aws s3 cp s3://newbucket/media/ s3://bucket2/media/ --recursive
```

Move Files

```
Terminal
$ aws s3 mv file1 s3://newbucket
$ aws s3 mv s3://newbucket/file1 /tmp
$ aws s3 mv s3://bucket1/file1 s3://bucket2
```

S3 CLI

\$ aws sync / s3://bucket upload: home\bob to s3://bucket/home/bob upload: var\log to s3://bucket/var/log \$ aws sync / s3://bucket upload: etc\ssh to s3://sanjeeevtestbucket/etc/ssh

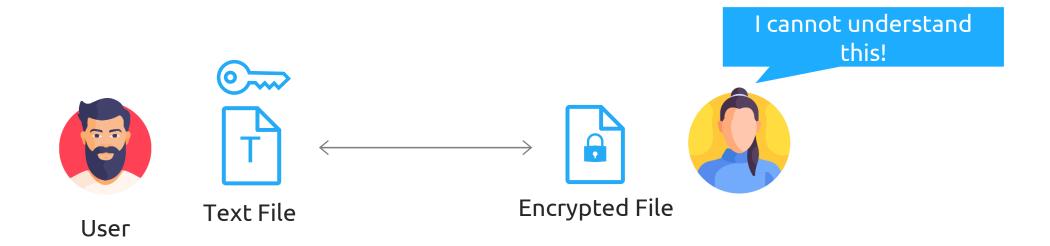




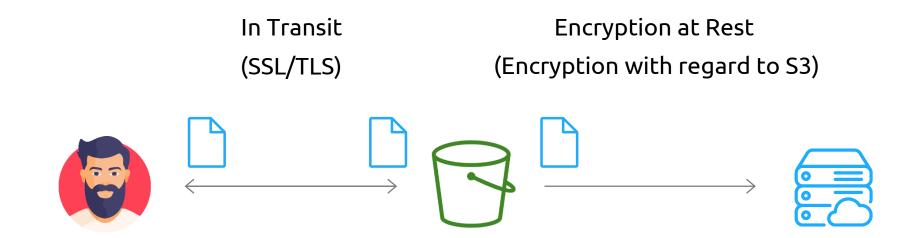
Encryption

Encryption

To scramble data so that only authorized parties can understand the information



Types of Encryption



Client-Side vs Server-Side Encryption

Client-Side Encryption



Server-Side Encryption



Encryption Methods

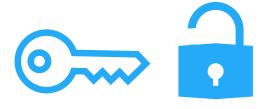
Server-Side Encryption with Amazon S3-Managed Keys (SSE-S3) Server-Side Encryption with Customer-Provided Keys (SSE-C) Server-Side Encryption with Key Management Service Keys (SSE-KMS)

> Encryption Methods



Generating and managing keys

> Encryption Methods



Using the keys to perform

Encryption

Note

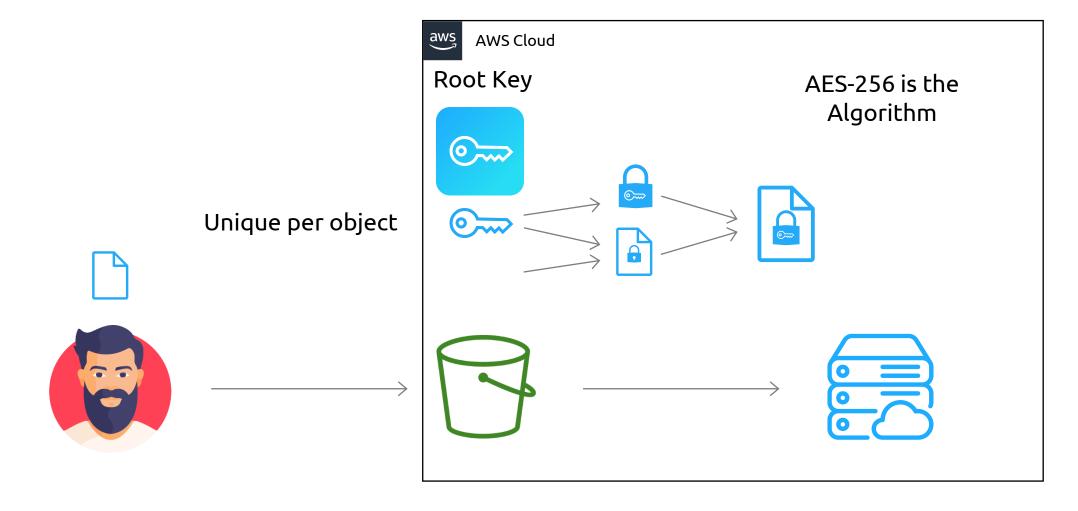
Encryption occurs on a per object basis

(One object can use one type of encryption, while another object can use a different type)

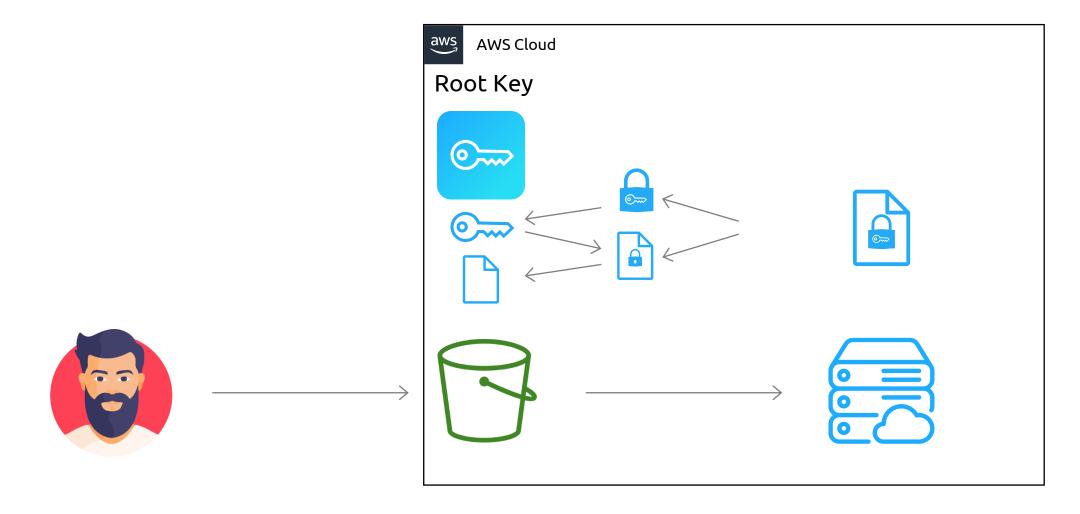
A default encryption method can be configured on a bucket

(All objects where user does not specify the encryption method will use the bucket default encryption)

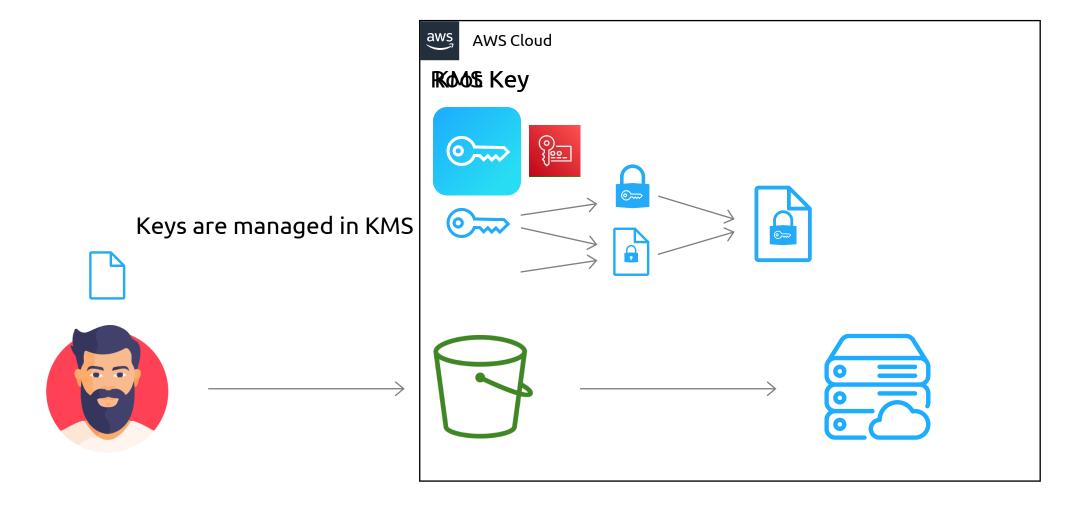
SSE-S3 Encryption



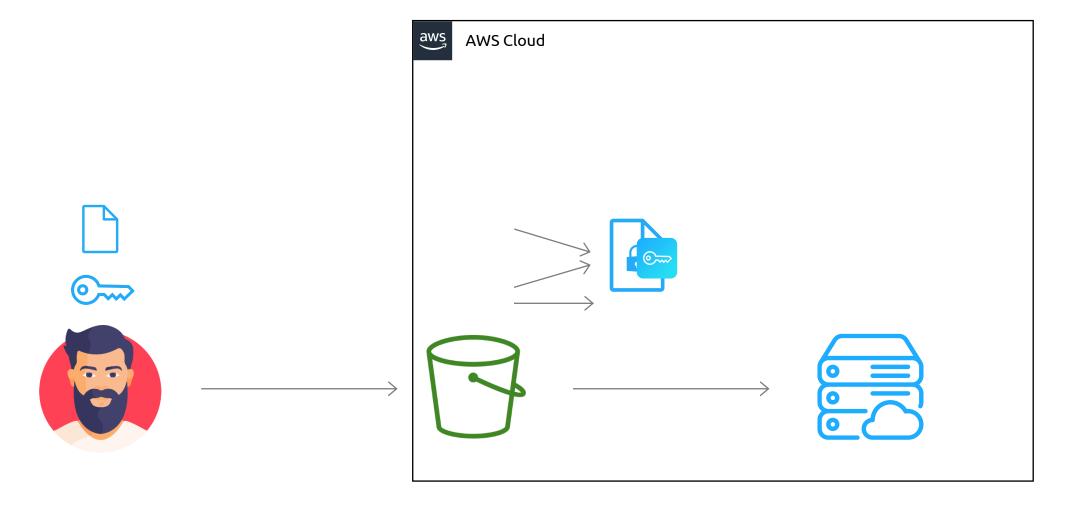
SSE-S3 Decryption



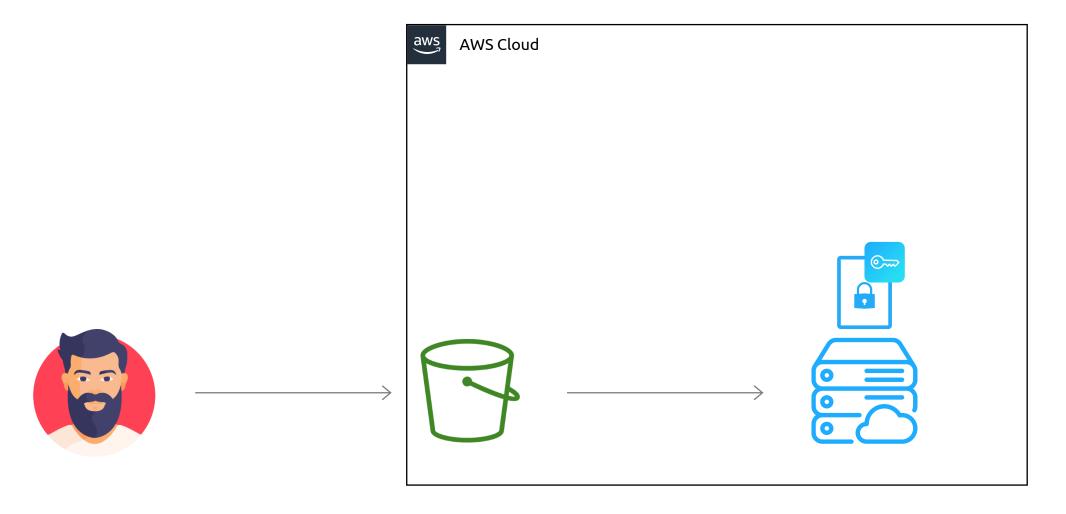
SSE-KMS



SSE-C



SSE-C



Headers

Name	Description
x-amz-server-side- encryption- customer-algorithm	Use this header to specify the encryption algorithm. The header value must be AES256.
x-amz-server-side- encryption- customer-key	Use this header to provide the 256-bit, base64-encoded encryption key for Amazon S3 to use to encrypt or decrypt your data.
x-amz-server-side- encryption- customer-key-MD5	Use this header to provide the base64-encoded 128-bit MD5 digest of the encryption key according to RFC 1321 . Amazon S3 uses this header for a message integrity check to ensure that the encryption key was transmitted without error.

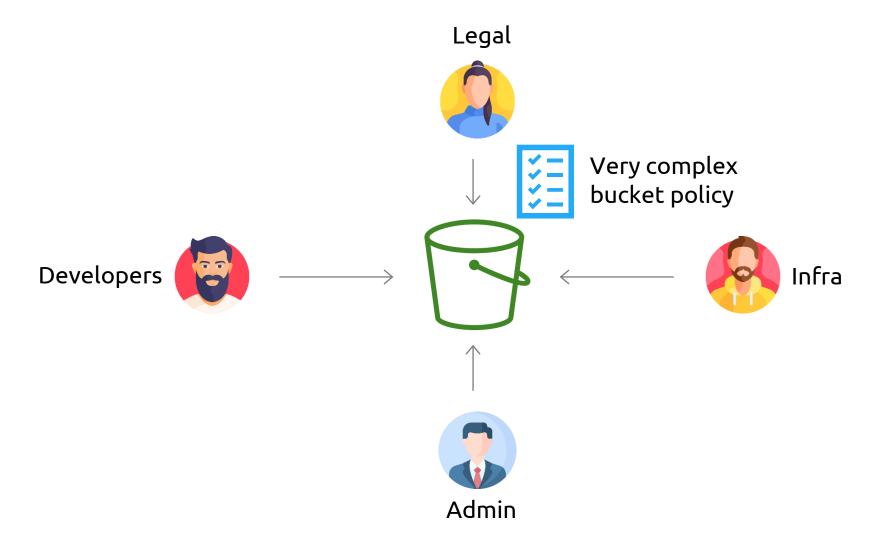
Summary

	Client Side	SSE-C	SSE-S3	SSE-KMS
Generate keys	Customer	Customer	S3	KMS + S3
Encrypt/decrypt	Customer	S3	S3	S3

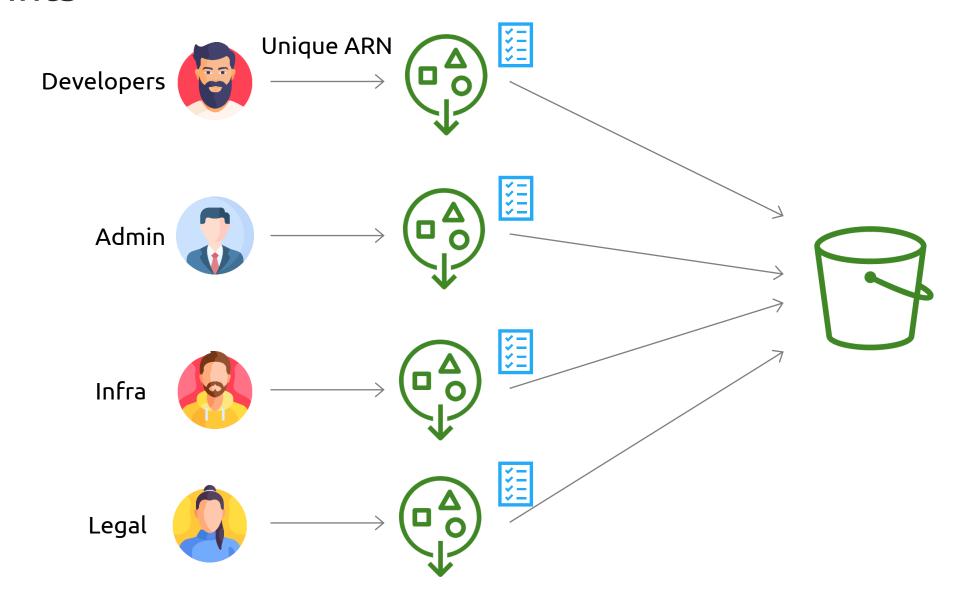


Access Points

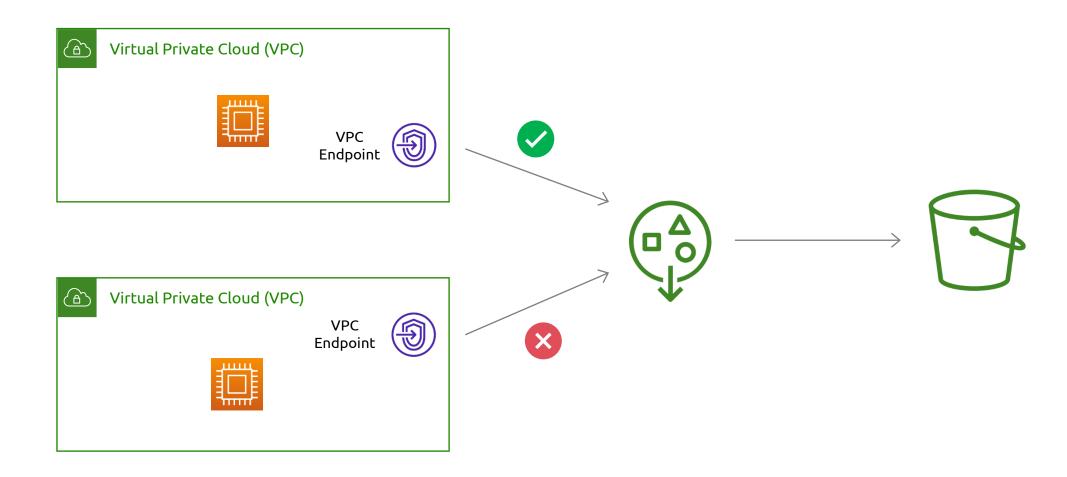
Access Points



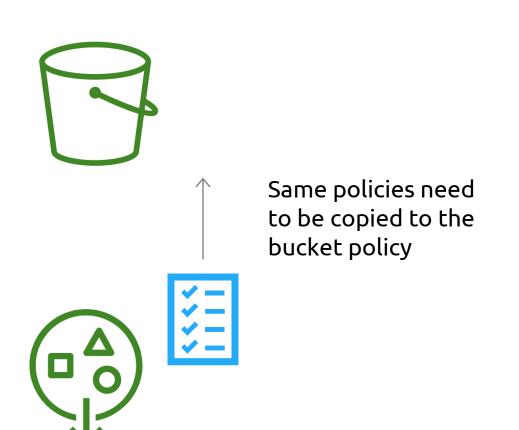
Access Points

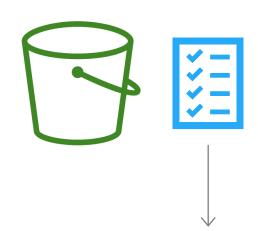


Access Point Restricting VPCs



Access Point Policy





Delegate policies to the access point



