AWS Solutions Architect : Associate Level

Source: https://docs.aws.amazon.com/

TECHNOLOGY

Amazon Storage Services

simplilearn

Learning Objectives

By the end of the lesson, you will be able to:

- Create Amazon S3 buckets
- Enable versioning in Amazon S3 buckets
- Demonstrate static web hosting using Amazon S3
- Share Amazon S3 buckets between multiple accounts
- Backup and sync data in Amazon S3
- Transfer files from Amazon S3 to on-premise storage
- Use CloudFront, Snowball, Athena, and Macie services

TECHNOLOGY

Introduction to Amazon S3

simplilearn

What Is Amazon S3?

Amazon Simple Storage Service (Amazon S3) is a web-based storage service offered by AWS for online backup and archiving of data. Users can store and retrieve any amount of data, at any time, from anywhere on the internet.

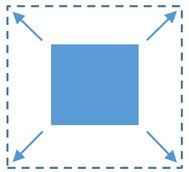
Amazon S3

Features of Amazon S3

The following are some of the features of Amazon S3:

01 **Availability**

02 **Scalability**



03 **Durability**

04 **Performance**

05 **Cost-efficiency**

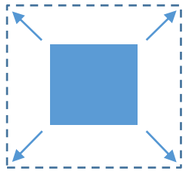
06 **Security**

Features of Amazon S3

Amazon S3 is designed for 99.99% availability. Users can choose the AWS region to store their data for optimizing latency and minimizing costs.

01 **Availability**

02 **Scalability**



03 **Durability**

04 **Performance**

05 **Cost-efficiency**

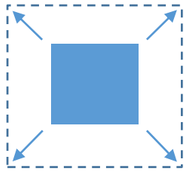
06 **Security**

Features of Amazon S3

Amazon S3 allows users to store data as per their needs and scales up or down the storage as required.

01 Availability

02 Scalability



03 Durability

04 Performance

05 Cost-efficiency

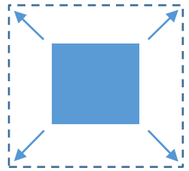
06 Security

Features of Amazon S3

Amazon S3 stores data redundantly across multiple facilities and multiple devices in each facility making it extremely durable.

01 **Availability**

02 **Scalability**



03 **Durability**

04 **Performance**

05 **Cost-efficiency**

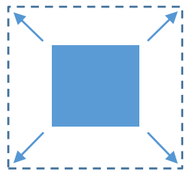
06 **Security**

Features of Amazon S3

Amazon S3 supports multipart uploads, which maximizes network throughput and resilience.

01 **Availability**

02 **Scalability**



03 **Durability**

04 **Performance**

05 **Cost-efficiency**

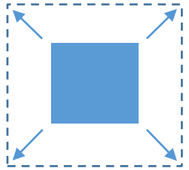
06 **Security**

Features of Amazon S3

Amazon S3 uses a pay-per-use model that allows users to store large data sets at a very low cost.

01 **Availability**

02 **Scalability**



03 **Durability**

04 **Performance**

05 **Cost-efficiency**

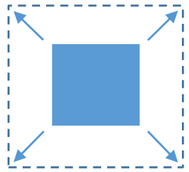
06 **Security**

Features of Amazon S3

Amazon S3 supports SSL (Secure Sockets Layer) data transfer and data encryption once the data is uploaded.

01 **Availability**

02 **Scalability**



03 **Durability**

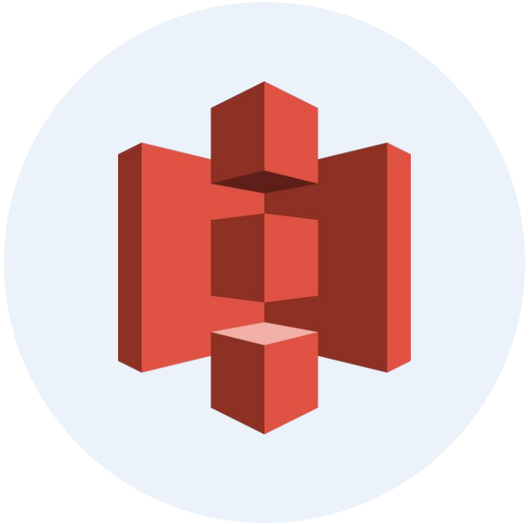
04 **Performance**

05 **Cost-efficiency**

06 **Security**

Components of Amazon S3

Components of Amazon S3



01

Buckets:
Buckets are containers for storing and organizing data of any format in Amazon S3.

02

Keys:
Keys are unique identifiers for each piece of data stored in Amazon S3 buckets.

03

Region:
Regions are geographical locations where Amazon stores the S3 buckets created by its users.

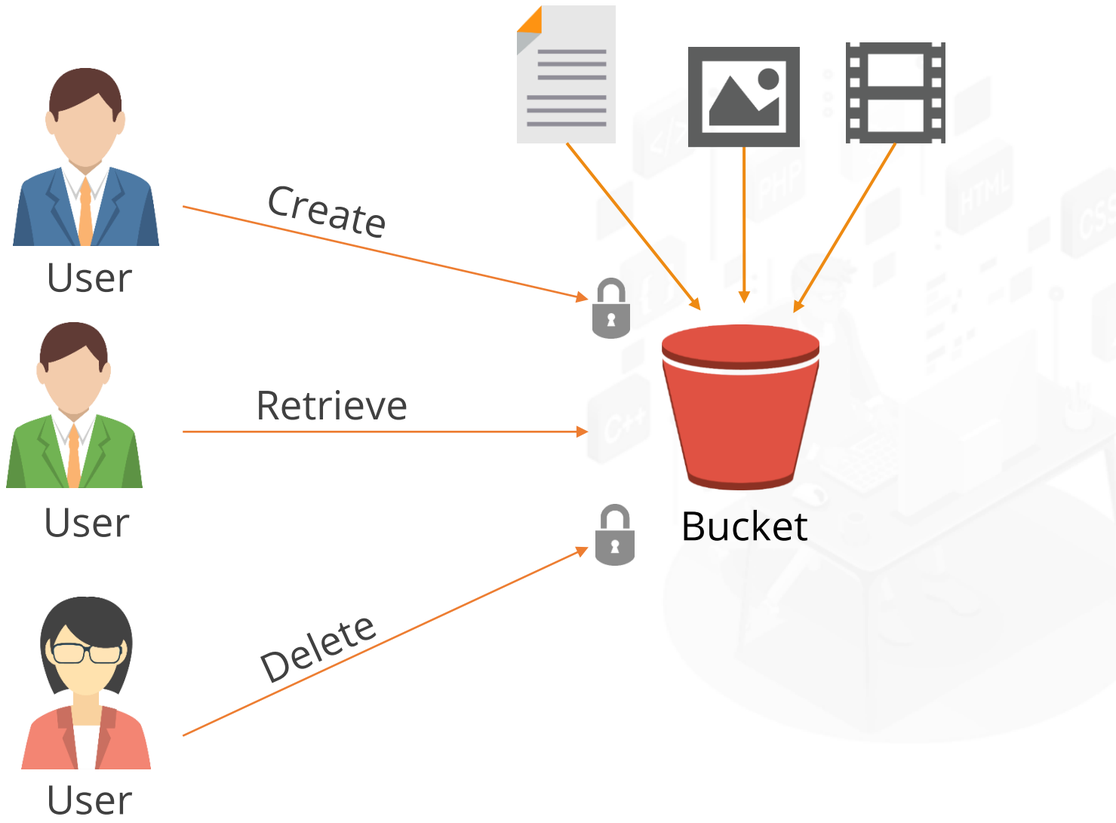
Components of Amazon S3

More about Amazon S3 buckets:

01 Buckets can be used to store text files, images, videos, and more.

02 Any number of objects can be stored in a bucket, given that the total bucket size is 5TB.

03 Access to each bucket action can be controlled by the user.



Creating an Amazon S3 Bucket

Duration: 15 min.

Problem Statement:

Create an Amazon S3 bucket using the Amazon console

ASSISTED PRACTICE

Assisted Practice: Guidelines to Create an Amazon S3 Bucket

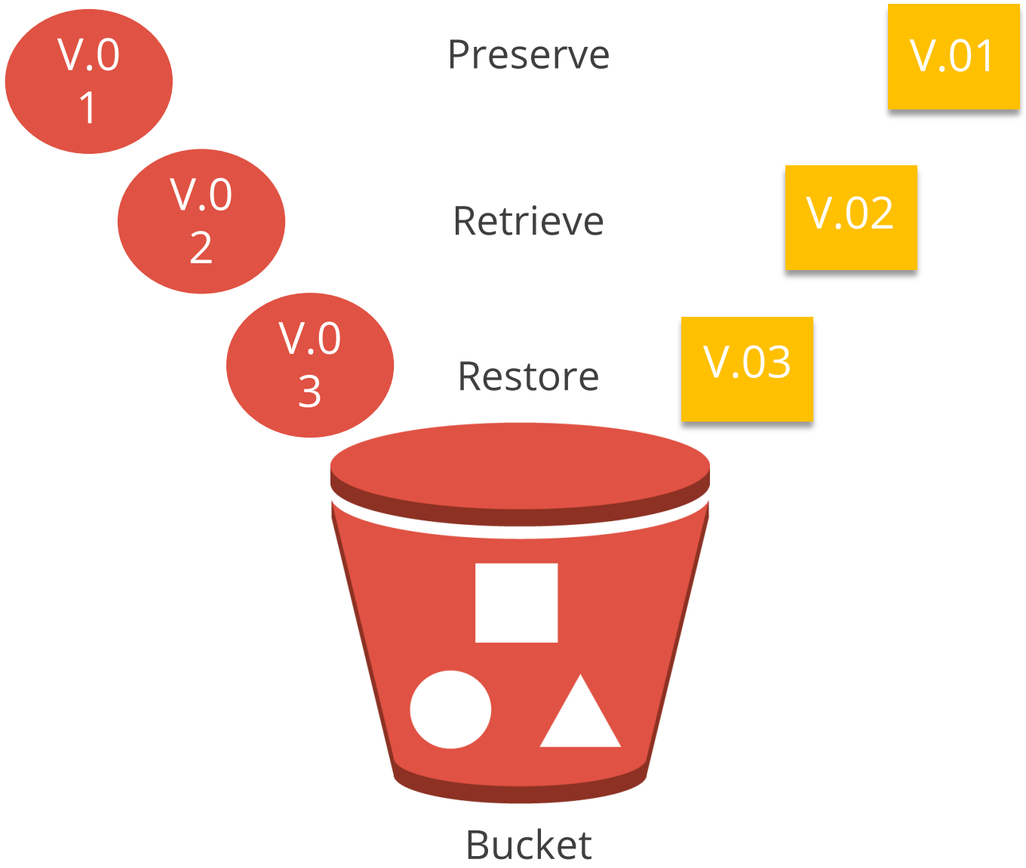
Steps to perform:

1. Go to your Amazon Console
2. Open the Amazon S3 dashboard
3. Click on the Create bucket button
4. Fill in the details about the bucket
5. Skip to the review page and click on the Create bucket button

Version Control in Amazon S3

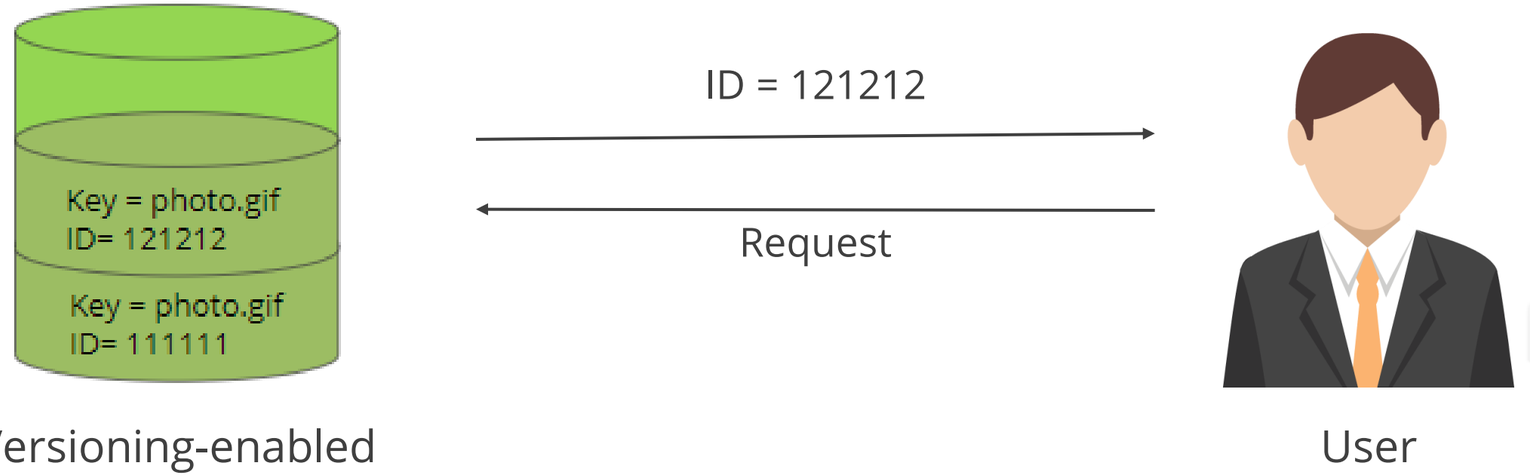
Version Control

Versioning is used to keep and maintain multiple variants of an object in the same bucket. It is used to preserve, retrieve, and restore earlier versions of every object stored in Amazon S3 buckets.



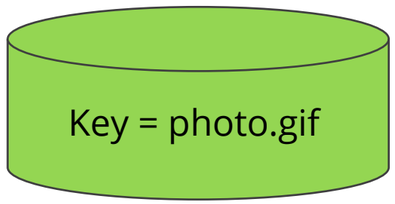
Version Control

Versioning allows you to recover your files from accidental deletion or overwrite.

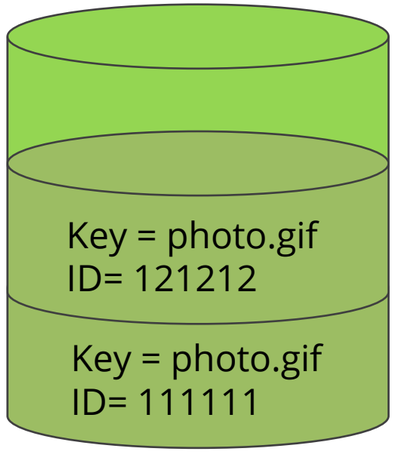


Version Control States

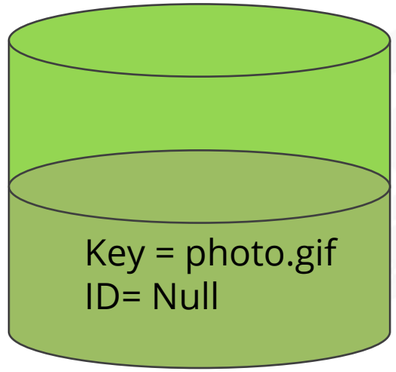
S3 buckets can be in one of the three states of versioning that are unversioned, versioning-enabled, and versioning-suspended.



Unversioned



Versioning-enabled

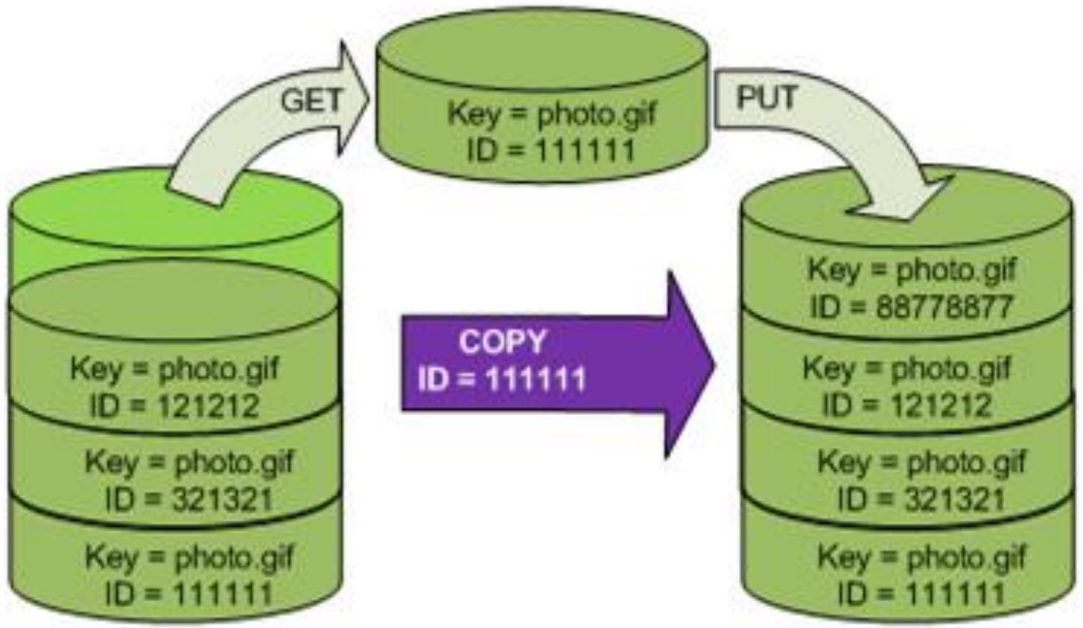


Versioning-suspended

Restoring a Previous Version

There are two approaches to restore a previous version:

1. Copy the previous version of an object into the bucket



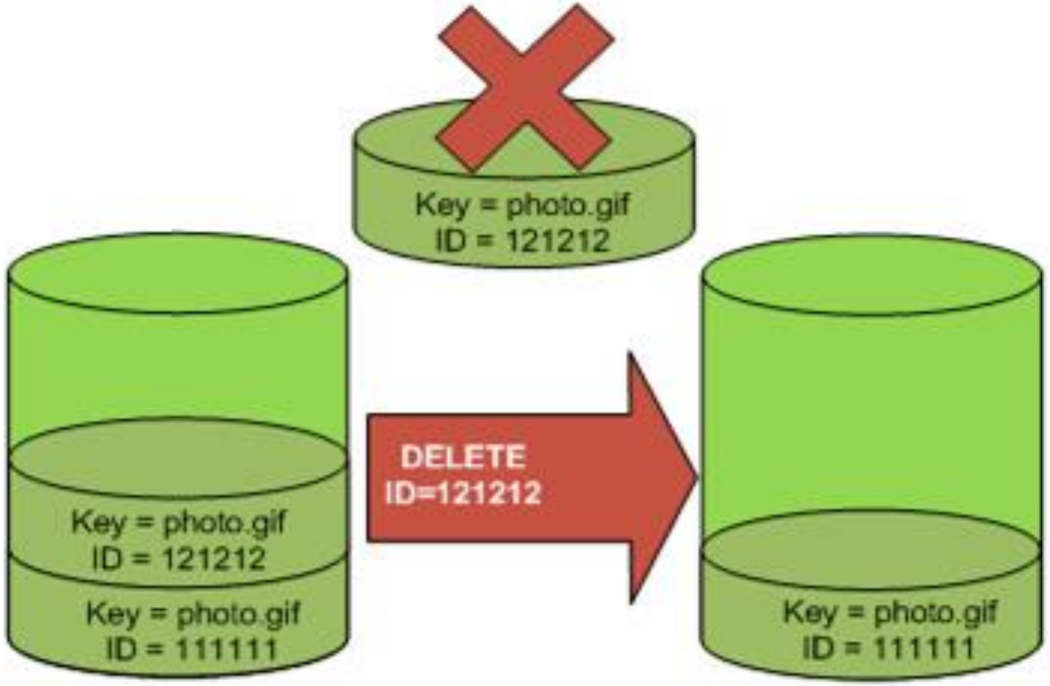
Versioning-enabled

Versioning-enabled

Source: https://docs.aws.amazon.com/AmazonS3/latest/dev/RestoringPreviousVersions.html

Restoring a Previous Version

2. Delete the current version of an object



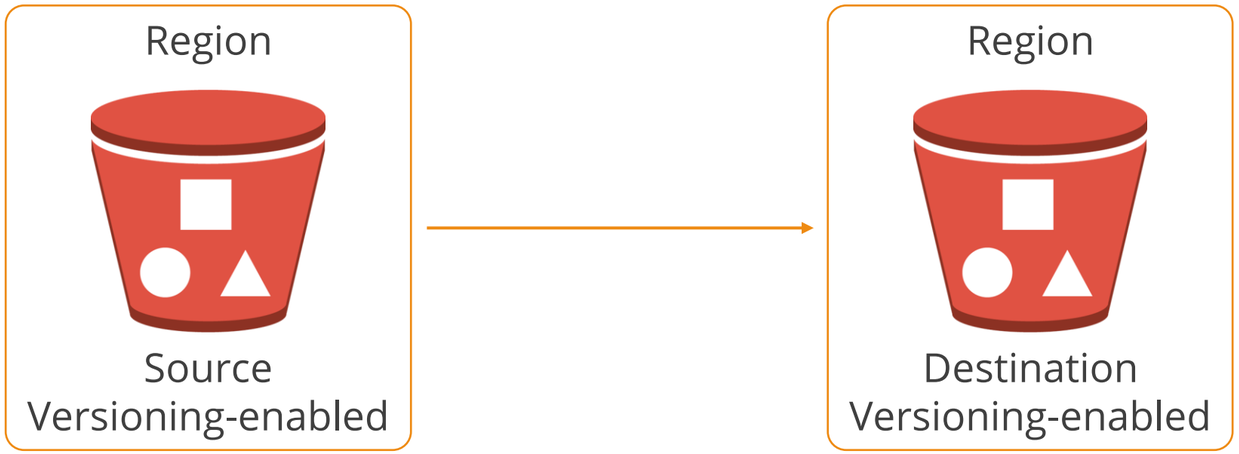
Versioning-enabled

Versioning-enabled

Source: https://docs.aws.amazon.com/AmazonS3/latest/dev/RestoringPreviousVersions.html

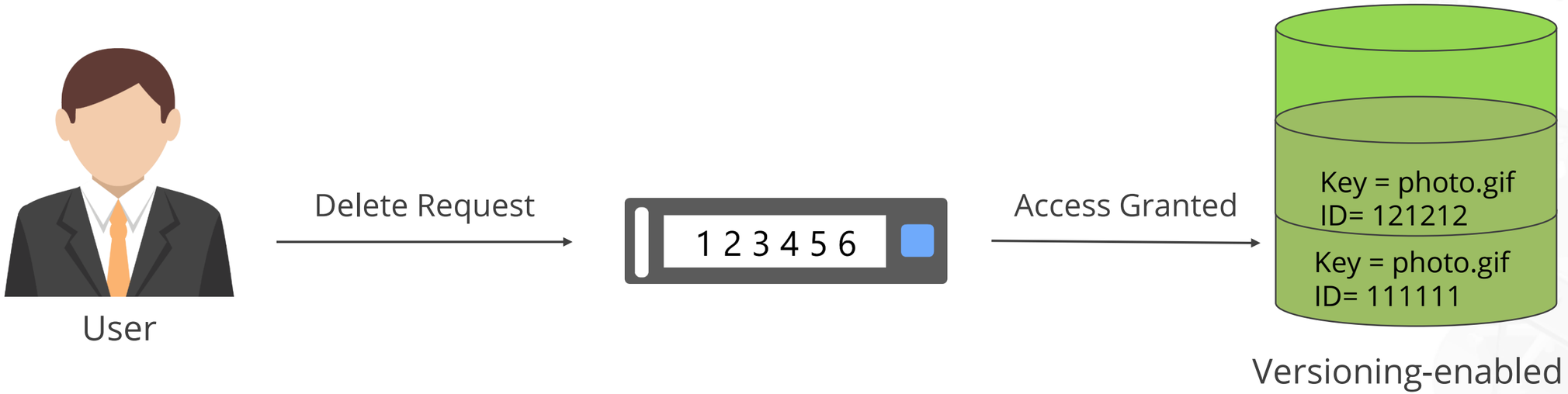
Cross-Region Replication

Cross-region replication is a bucket-level feature that enables automatic, asynchronous copying of objects across buckets in different AWS regions. For cross-region replication to work, versioning needs to be enabled on both the source and destination buckets.



MFA Delete

Amazon S3 allows users to protect their data by enabling Multi-Factor Authentication (MFA) delete. This provides additional authentication for operations such as changing the versioning state of a bucket permanently and deleting an object version.



Configure an S3 Bucket for Versioning

Duration: 5 min.

Problem Statement:

Configure an S3 bucket for versioning using Amazon S3 console

ASSISTED PRACTICE

Assisted Practice: Guidelines to Configure Bucket Versioning

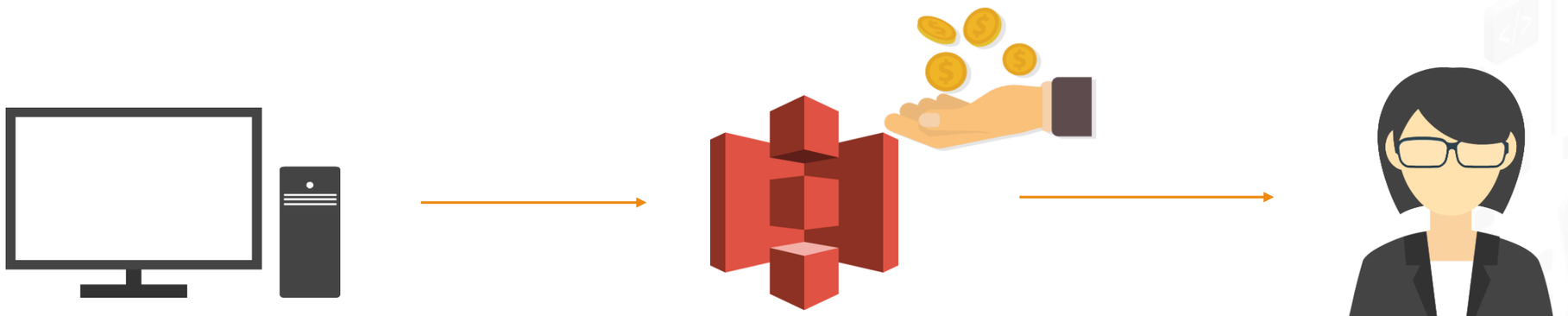
Steps to perform:

1. Go to your Amazon Console
2. Open the Amazon S3 dashboard
3. Create a bucket
4. Select the bucket and go to Properties tab
5. Click on the Enable versioning button

Static Web Hosting

Static Web Hosting

Amazon S3 allows users to host static websites at a low cost by providing a highly available hosting solution. A static website contains web pages with static content and client-side scripts.



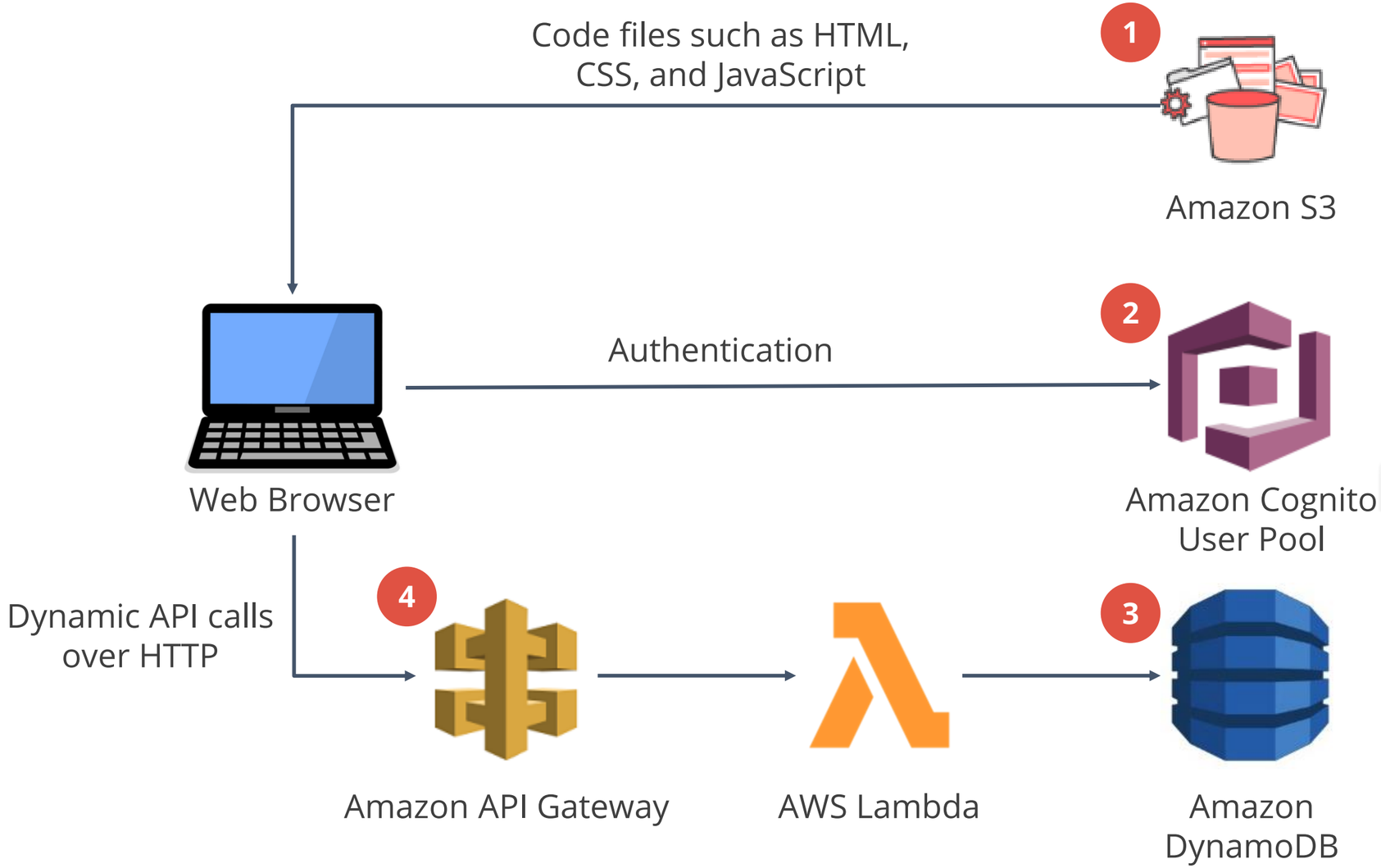
Website

Amazon S3

User

<bucket-name>.s3-website-<AWS-region>.amazonaws.com is the naming convention used for hosting static websites.

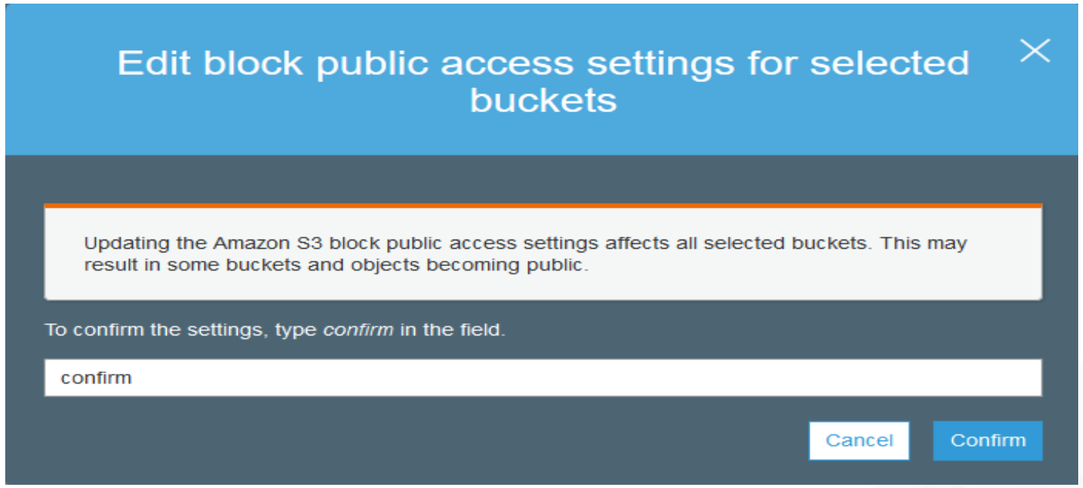
Static Web Hosting



Static Web Hosting

To configure an Amazon S3 bucket for static website hosting in the Amazon S3 console, follow the steps mentioned below:

1. Enable static website hosting



2. Edit block public access settings

Static Web Hosting

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "PublicReadGetObject",
            "Effect": "Allow",
            "Principal": "*",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::example.com/*"
            ]
        }
    ]
}
```

3. Add a bucket policy in the bucket policy editor

```
<html xmlns="http://www.w3.org/1999/xhtml" >
<head>
    <title>My Website Home Page</title>
</head>
<body>
  <h1>Welcome to my website</h1>
  <p>Now hosted on Amazon S3!</p>
</body>
</html>
```

4. Configure an index document

TECHNOLOGY

Amazon S3 Policies

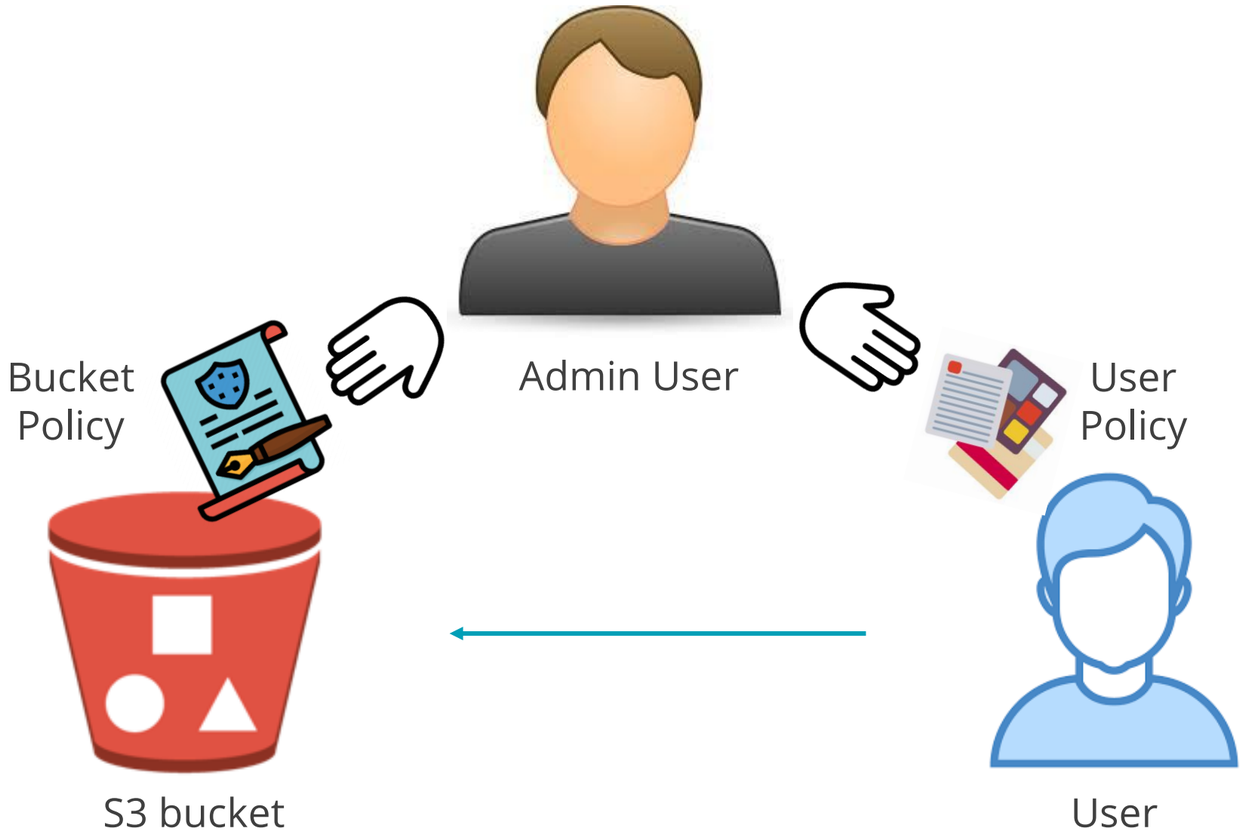
Amazon S3 Policies

An S3 bucket policy is a resource-based AWS Identity and Access Management (IAM) policy that grants other AWS accounts or IAM users access to the objects in an S3 bucket.

Amazon S3 bucket policy

Object permissions apply only to the objects created by the bucket owner. S3 policies supplement, and in many cases, replace ACL(Access Control List)-based access policies.

Amazon S3 Policies



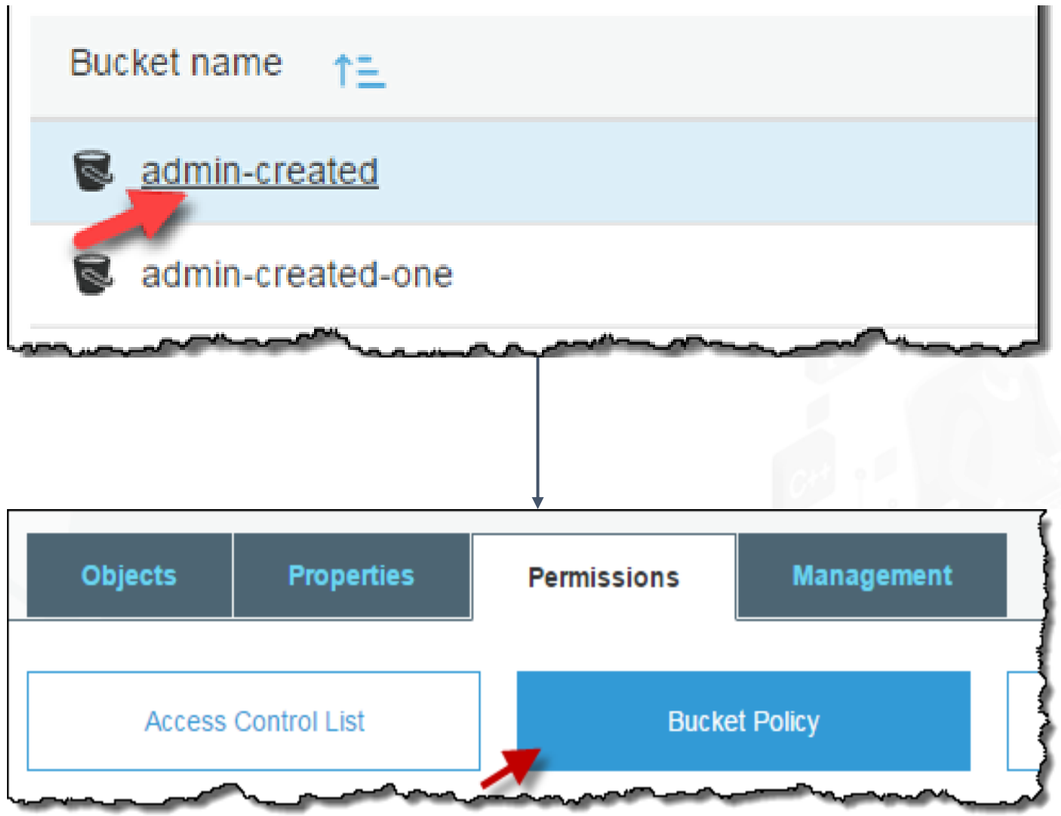
Bucket owner grants permissions to users through bucket policies.

Amazon S3 Policies

1. Sign in to the AWS Management Console and open the Amazon S3 console at https://console.aws.amazon.com/s3/

2. Choose a bucket name from the **Bucket name** list for which the bucket policy is to be created

3. Choose **Permissions**, and then select **Bucket Policy**

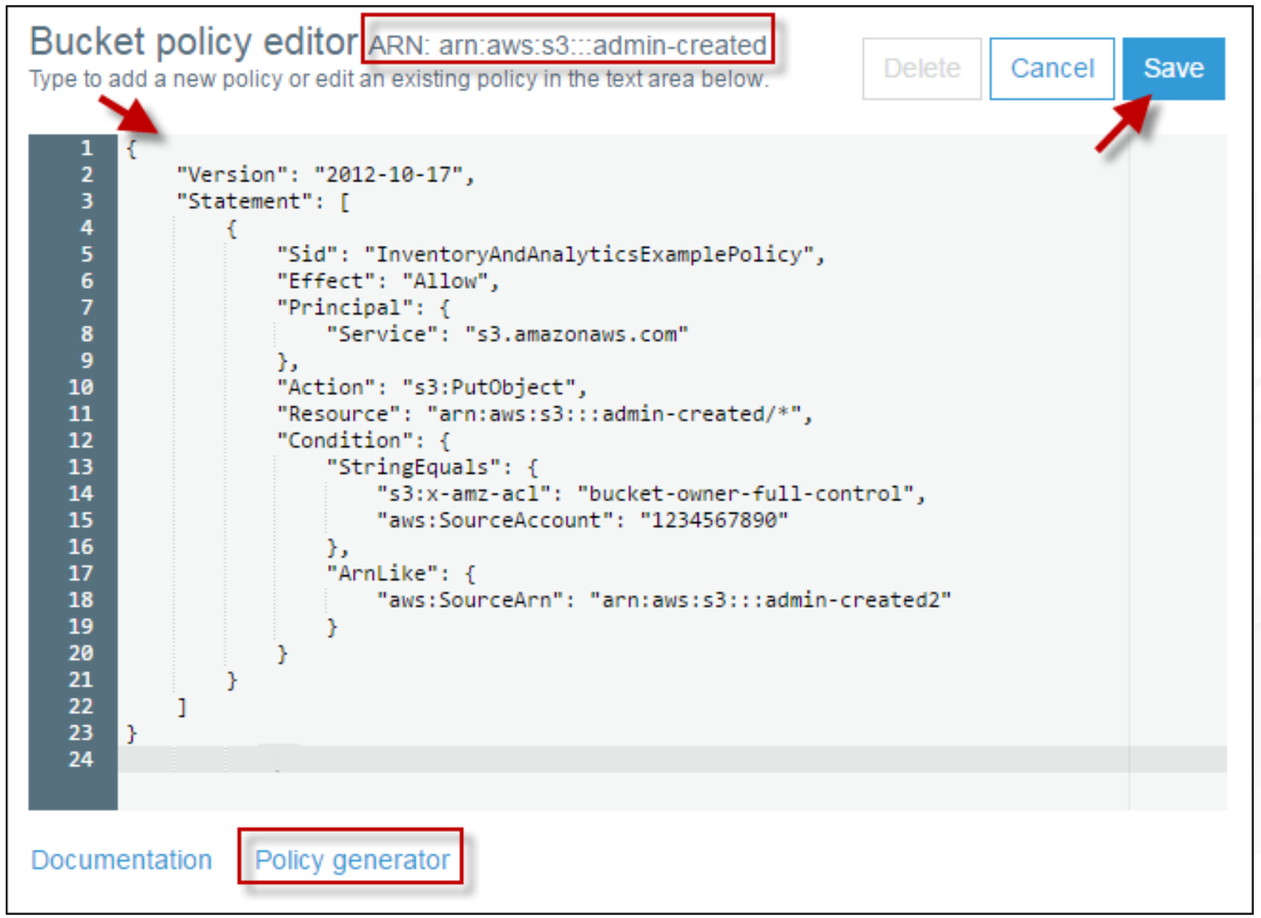


Source: https://docs.aws.amazon.com/AmazonS3/latest/user-guide/add-bucket-policy.html

Amazon S3 Policies

4

In the **Bucket policy editor**, add a new bucket policy or edit an existing policy. Add a valid JSON in the policy, and then click **Save** to save the bucket policy



Source: https://docs.aws.amazon.com/AmazonS3/latest/user-guide/add-bucket-policy.html

Configure an S3 Bucket for Static Website Hosting

Duration: 20 min.

Problem Statement:

Configure an S3 bucket for static website hosting using the Amazon S3 console

ASSISTED PRACTICE

Assisted Practice: Guidelines to Configure an S3 Bucket for Static Website Hosting

Steps to perform:

1. Go to your Amazon Console
2. Open the Amazon S3 dashboard
3. Create a bucket
4. Select the bucket and go to the Static website hosting tab
5. Upload the HTML files in the bucket
6. Edit Block access and Bucket policies

Amazon S3 Storage Classes

Amazon S3 Storage Classes

Amazon S3 comes in the following range of storage classes:

- Amazon S3 Standard
- Amazon S3 Intelligent-Tiering
- Amazon S3 Standard-Infrequent Access
- Amazon S3 One Zone-Infrequent Access
- Amazon S3 Glacier
- Amazon S3 Glacier Deep Archive

Amazon S3 Standard

Features of Amazon S3 Standard

- Stores objects that are frequently accessed with high availability and durability
- Low latency and high throughput performance
- Backed with Amazon S3 Service Level Agreement for availability
- Ideal for dynamic websites, cloud and mobile applications, and file storage
- S3 Lifecycle management for an automatic transition of objects between S3 storage classes

Amazon S3 Intelligent-Tiering

Features of Amazon S3 Intelligent-Tiering

- Stores objects in two access tiers optimized for frequent and infrequent access
- Designed for high availability of 99.99% and 11 9’s of durability
- Same low latency and high throughput performance of S3 Standard
- Small monthly monitoring and auto-tiering fee
- Ideal for long-lived data with unpredictable access patterns
- Automatically moves objects between two access tiers based on changing access patterns

Amazon S3 Standard-Infrequent Access

Features of Amazon S3 Standard-Infrequent Access

- Stores objects that are accessed less frequently
- Designed for high availability of 99.99% and 11 9’s of durability
- Same low latency and high throughput performance of S3 Standard
- Lower per GB storage price and per GB retrieval fee
- Ideal for long-term storage, backups, and as a data store for disaster recovery files

Amazon S3 One Zone-Infrequent Access

Features of Amazon S3 One Zone-Infrequent Access

- Stores objects that are accessed less frequently, but require rapid access when needed
- Stores data in a single AZ and costs 20% less than S3 Standard-IA
- Lower per GB storage price and per GB retrieval fee
- Ideal for customers who want a lower-cost option for infrequently accessed data
- Good for storing secondary backup copies of on-premises data

Amazon S3 Glacier

Features of Amazon S3 Glacier

- Secure, durable, and low-cost storage class for archiving rarely accessed data
- Provides three configurable retrieval options, varying from minutes to hours
- Lower per GB storage price and per GB retrieval fee
- Offers a secure vault lock feature that enforces compliance
- Data is resilient in case of an entire Availability Zone destruction
- Ideal for database backups, compliance data, or audit log files that are accessed rarely

Amazon S3 Glacier Deep Archive

Features of Amazon S3 Glacier Deep Archive

- Lowest-cost storage class supporting long-term retention and digital preservation
- Stores data that may be accessed once or twice in a year
- Used for backup and disaster recovery
- Data replicated and stored in at least three geographically-dispersed AZs
- Data retrieval time within 12 hours
- Good for customers from highly-regulated industries who retain data for 7-10 years

Performance Chart

	S3 Standard	S3 Intelligent-Tiering*	S3 Standard-IA	S3 One Zone-IA†	S3 Glacier	S3 Glacier Deep Archive
Designed for durability	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)
Designed for availability	99.99%	99.9%	99.9%	99.5%	99.99%	99.99%
Availability SLA	99.9%	99%	99%	99%	99.9%	99.9%
Availability Zones	≥3	≥3	≥3	1	≥3	≥3
Minimum capacity charge per object	N/A	N/A	128KB	128KB	40KB	40KB

Performance across the S3 storage classes

Source: https://aws.amazon.com/s3/storage-classes/

Performance Chart

	S3 Standard	S3 Intelligent-Tiering*	S3 Standard-IA	S3 One Zone-IA†	S3 Glacier	S3 Glacier Deep Archive
Minimum storage duration charge	N/A	30 days	30 days	30 days	90 days	180 days
Retrieval fee	N/A	N/A	per GB retrieved	per GB retrieved	per GB retrieved	per GB retrieved
First byte latency	milliseconds	milliseconds	milliseconds	milliseconds	select minutes or hours	select hours
Storage type	Object	Object	Object	Object	Object	Object
Lifecycle transitions	Yes	Yes	Yes	Yes	Yes	Yes

Performance across the S3 storage classes

Source: https://aws.amazon.com/s3/storage-classes/

Amazon S3 Pricing

Amazon S3 Pricing

Amazon S3 offers a pay-as-you-go model for pricing, which means users pay only for the individual services they use, for as long as they use them, and without requiring any long-term contracts or complex licensing.

Amazon S3 pricing

Source: https://aws.amazon.com/pricing/

Amazon S3 Pricing

The following are the four cost components to consider when deciding on which S3 storage class best fits the data profile:

Storage pricing

Request and data retrieval pricing

Data transfer and acceleration pricing

Data management features and replication pricing

Amazon S3 Pricing

- Storage pricing is the fee paid by users for storing objects in the S3 buckets.
- A monthly monitoring and automation fee per object is charged for the objects stored in the S3 Intelligent-Tiering storage class.
- Per-request ingest fee is also charged while using PUT, COPY, or lifecycle rules to move data into any S3 storage class.

Storage pricing

Amazon S3 Pricing

	Storage pricing
S3 Standard - General purpose storage for any type of data, typically used for frequently accessed data	
First 50 TB / Month	$0.025 per GB
Next 450 TB / Month	$0.024 per GB
Over 500 TB / Month	$0.023 per GB
S3 Intelligent - Tiering * - Automatic cost savings for data with unknown or changing access patterns	
Frequent Access Tier, First 50 TB / Month	$0.025 per GB
Frequent Access Tier, Next 450 TB / Month	$0.024 per GB
Frequent Access Tier, Over 500 TB / Month	$0.023 per GB
Infrequent Access Tier, All Storage / Month	$0.019 per GB
Monitoring and Automation, All Storage / Month	$0.0025 per 1,000 objects

Storage pricing for Asia Pacific region

Source: https://aws.amazon.com/s3/pricing/

Amazon S3 Pricing

S3 Standard - Infrequent Access * - For long lived but infrequently accessed data that needs millisecond access	
All Storage / Month	$0.019 per GB
S3 One Zone - Infrequent Access * - For re-createable infrequently accessed data that needs millisecond access	
All Storage / Month	$0.0152 per GB
S3 Glacier ** - For long-term backups and archives with retrieval option from 1 minute to 12 hours	
All Storage / Month	$0.005 per GB
S3 Glacier Deep Archive ** - For long-term data archiving that is accessed once or twice in a year and can be restored within 12 hours	
All Storage / Month	$0.002 per GB

Storage pricing for Asia Pacific region

Source: https://aws.amazon.com/s3/pricing/

Amazon S3 Pricing

- Requests and data retrieval pricing is the fee paid by users for requests made against the S3 buckets and objects.
- Browsing charges are also incurred by the user for using the Amazon S3 console to browse the storage.
- Charges are accrued at the same rate as requests that are made using the API/SDK.
- DELETE and CANCEL requests are free, but LIST requests for any storage class are charged at the same rate as S3 Standard PUT, COPY, and POST requests.

Requests and data retrieval pricing

Amazon S3 Pricing

	PUT, COPY, POST, LIST requests (per 1,000 requests)	GET, SELECT, and all other requests (per 1,000 requests)	Lifecycle Transition requests (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
S3 Standard - Infrequent Access*	$0.01	$0.001	$0.01	n/a	$0.01
S3 One Zone - Infrequent Access*	$0.01	$0.001	$0.01	n/a	$0.01
S3 Glacier **	$0.06	$0.0004	$0.06	See below	See below
Expedited	n/a	n/a	n/a	$12.00	$0.036
Standard	n/a	n/a	n/a	$0.06	$0.012
Bulk	n/a	n/a	n/a	$0.03	$0.003
Provisioned Capacity Unit ***	n/a	n/a	n/a	n/a	$120.00 per unit
S3 Glacier Deep Archive **	$0.07	$0.0004	$0.07	See below	See below
Standard	n/a	n/a	n/a	$0.12	$0.024
Bulk	n/a	n/a	n/a	$0.03	$0.005

Requests and data retrieval pricing for Asia Pacific region

Source: https://aws.amazon.com/s3/pricing/

Amazon S3 Pricing

- Data management features pricing is the fee paid by the users for the storage management features that are enabled on their account's buckets.

- Replication pricing is the fee paid by the users for S3 Replication, which includes:
  - S3 charges for storage in the selected destination S3 storage class
  - Storage charges for the primary copy
  - Charges for replication PUT requests and applicable infrequent access storage retrieval fee

Data management features and replication pricing

Amazon S3 Pricing

- For Cross-Region Replication, users pay a fee for inter-region Data Transfer OUT From S3 to their destination region.
- For S3 Replication Time Control, users pay a Replication Time Control Data Transfer fee and S3 Replication Metrics charges.
- For the replicated copy, Storage and PUT request pricing is based on the destination AWS Region.

Data management features and replication pricing

Amazon S3 Pricing

S3 Inventory††	$0.0028 per million objects listed
S3 Analytics Storage Class Analysis††	$0.10 per million objects monitored per month
S3 Object Tagging	$0.01 per 10,000 tags per month
†† The files produced by S3 Inventory and S3 Storage Class Analysis exports are stored in your specified S3 bucket, and are subject to S3 Standard storage charges.	
S3 Replication Time Control data transfer†	$0.015 per GB
† Amazon S3 Replication Time Control Data Transfer pricing is the same in all AWS Regions. Replication Time Control is available in all commercial AWS Regions excluding AWS GovCloud (US) Regions, the AWS China (Beijing) Region, and the AWS China (Ningxia) Region.	

Data management features and replication pricing for Asia Pacific region

Source: https://aws.amazon.com/s3/pricing/

Amazon S3 Pricing

- Data transfer and acceleration pricing is the fee paid by the users for all the bandwidth into and out of Amazon S3, except for the following:
  - Data transferred in from the internet
  - Data transferred out to an Amazon EC2 instance, when the instance is in the same AWS Region
  - Data transferred out to Amazon CloudFront
- Users have to pay for any data transferred using Amazon S3 Transfer Acceleration.

Data transfer and acceleration pricing

Amazon S3 Pricing

	Price
Data Transfer IN To Amazon S3 From Internet	
All data transfer in	$0.00 per GB
Data Transfer OUT From Amazon S3 To Internet	
Up to 1 GB / Month	$0.00 per GB
Next 9.999 TB / Month	$0.1093 per GB
Next 40 TB / Month	$0.085 per GB
Next 100 TB / Month	$0.082 per GB
Greater than 150 TB / Month	$0.08 per GB
Data Transfer OUT From Amazon S3 To	
Amazon CloudFront	$0.00 per GB

Data transfer and acceleration pricing for Asia Pacific region

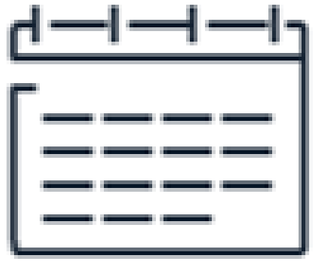
Source: https://aws.amazon.com/s3/pricing/

AWS Free Tier

AWS offers a Free Tier to new customers upon sign up to get started with Amazon S3 for free. It provides free, hands-on experience with the AWS platform, products, and services.

These are the different types of free offers, based on the product usage:

Always free



12-months free

Trials

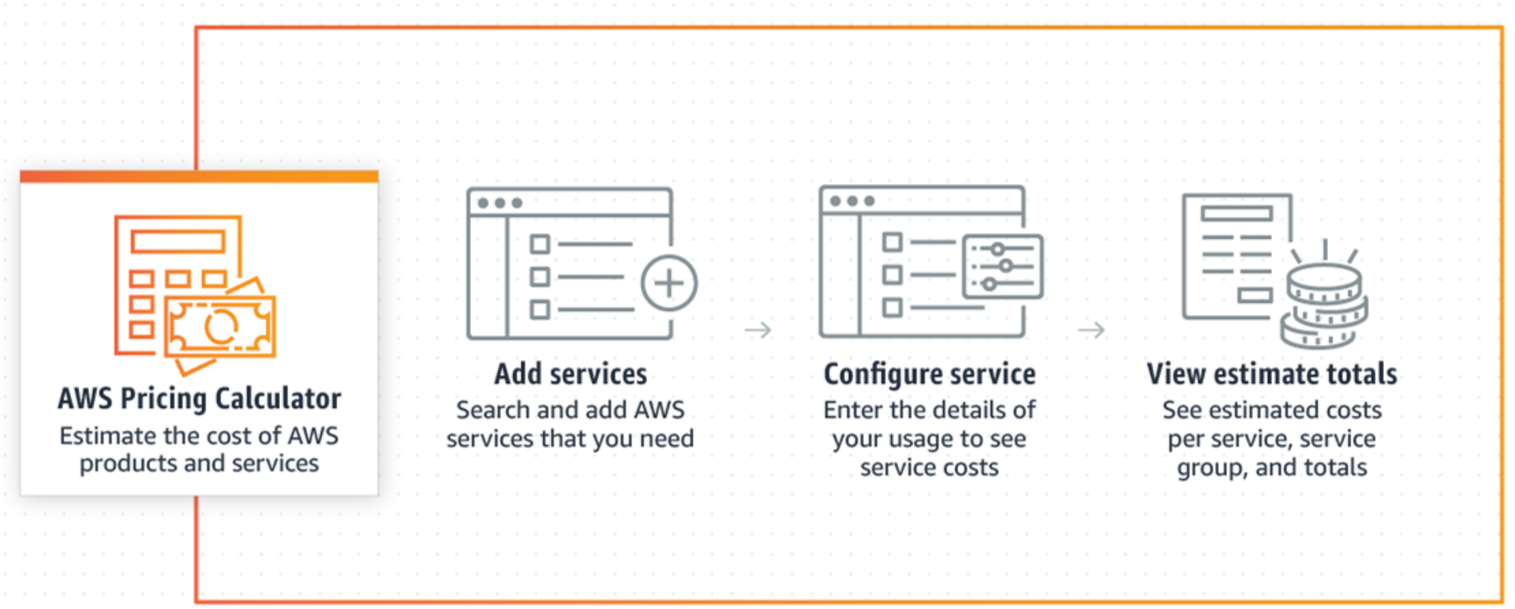
AWS Free Tier

In AWS Free Tier, new customers receive the following monthly benefits for a year:

- 5GB of Amazon S3 storage in the S3 Standard storage class
- 20,000 GET Requests
- 2,000 PUT, COPY, POST, or LIST Requests
- 15GB of Data Transfer Out

AWS Pricing Calculator

AWS pricing calculator estimates the cost for a storage solution. Users can configure a cost estimate that fits their individual business or personal requirements with Amazon S3.



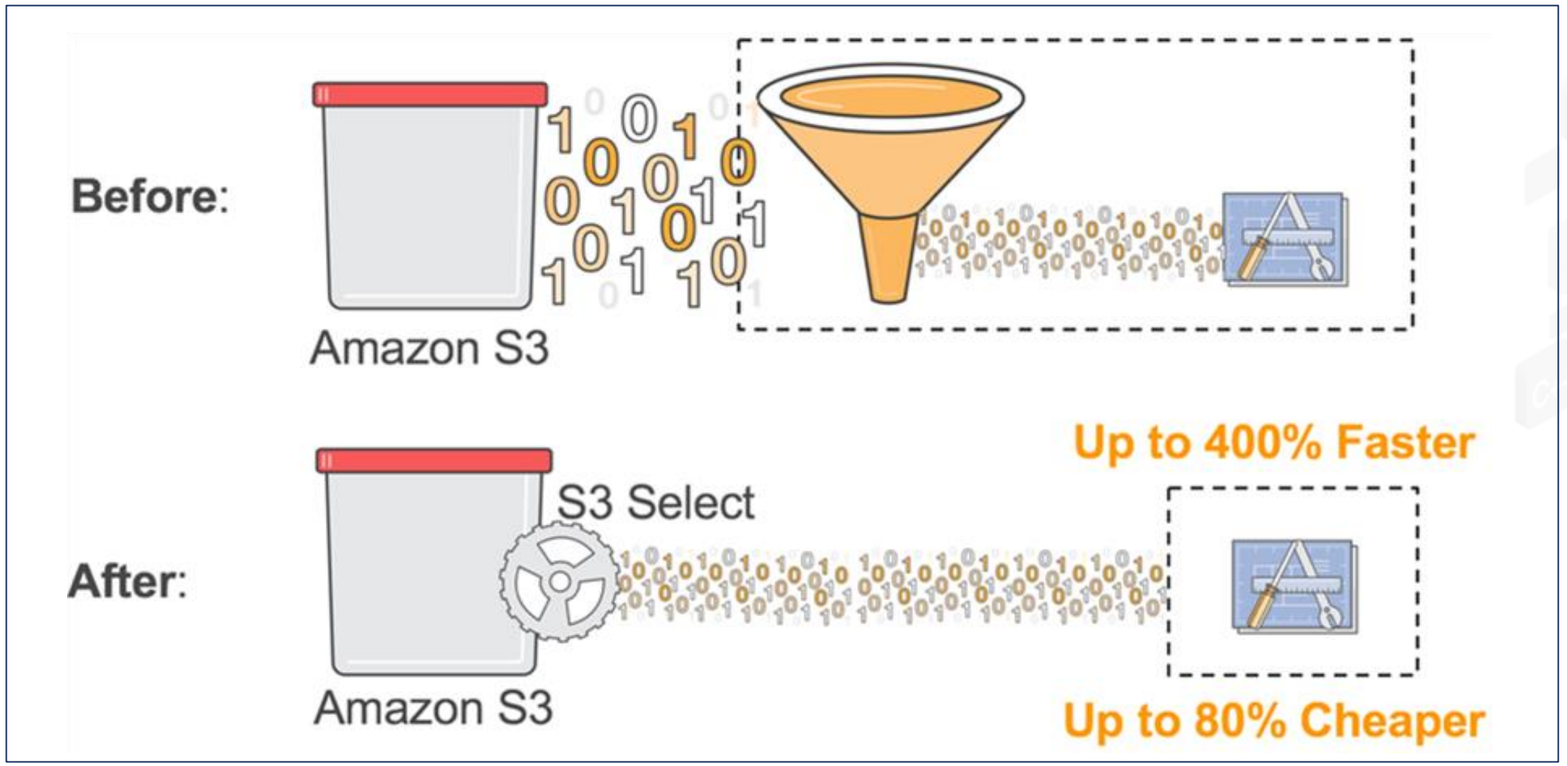
Working of AWS Pricing Calculator

Source: https://calculator.aws/#/

Amazon S3 Select and S3 Glacier Select

Amazon S3 Select

Amazon S3 Select allows applications to fetch only a subset of data from an object by using simple SQL expressions, helping users to achieve drastic performance increases.



Data fetching using Amazon S3 Select

Source: https://aws.amazon.com/blogs/aws/s3-glacier-select/

Amazon S3 Glacier Select

- Amazon S3 Glacier Select enables users to perform data filtering directly in a Glacier object using standard SQL statements.
- It is mostly available in all the commercial regions that have S3 Glacier.
- Glacier is priced in 3 dimensions:
  - GB of Data Scanned
  - GB of Data Returned
  - Select Requests

Amazon S3 Glacier Select

```
import boto3
glacier = boto3.client("glacier")

jobParameters = {
    "Type": "select", "ArchiveId": "ID",
    "Tier": "Expedited",
    "SelectParameters": {
        "InputSerialization": {"csv": {}},
        "ExpressionType": "SQL",
        "Expression": "SELECT * FROM archive WHERE _5='498960'",
        "OutputSerialization": {
            "csv": {}
        }
    },
    "OutputLocation": {
        "S3": {"BucketName": "glacier-select-output", "Prefix": "1"}
    }
}

glacier.initiate_job(vaultName="reInventSecrets", jobParameters=jobParameters)
```

Example of Amazon S3 Glacier Select

Source: https://aws.amazon.com/blogs/aws/s3-glacier-select/

TECHNOLOGY

Sharing an S3 Bucket Between Multiple Accounts

Sharing an S3 Bucket Between Multiple Accounts

Depending on the type of access that a user wants to provide, the following methods can be used to grant cross-account access for sharing objects between multiple accounts:

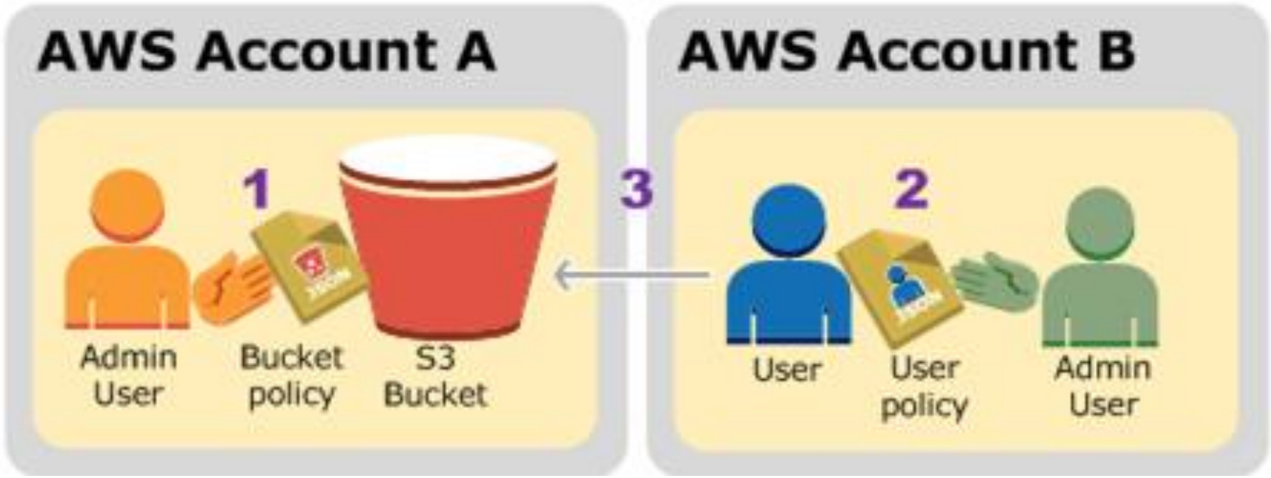
- Resource-based policies and AWS IAM policies
- Resource-based Access Control List (ACL) and IAM policies
- Cross-account IAM roles

Steps to Share an S3 Bucket Between Multiple Accounts

1. Account A admin user attaches a bucket policy granting cross-account permissions to Account B.

2. Account B admin user attaches a user policy authorizing the permissions it received from Account A.

3. User in Account B verifies the permissions by accessing the bucket owned by Account A.

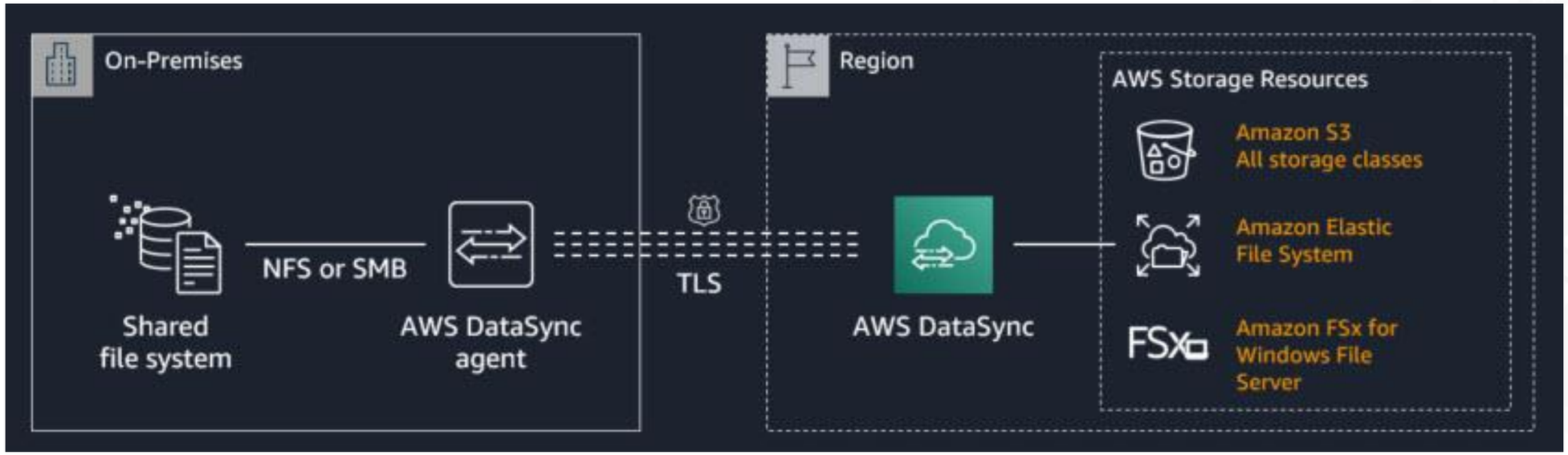


Source: https://docs.aws.amazon.com/AmazonS3/latest/dev/example-walkthroughs-managing-access-example2.html

Backup and DataSync

AWS DataSync

AWS DataSync allows fast and simple transfer of large amounts of data online between on-premises storage and Amazon S3, Amazon EFS, or Amazon FSx for Windows File Server.

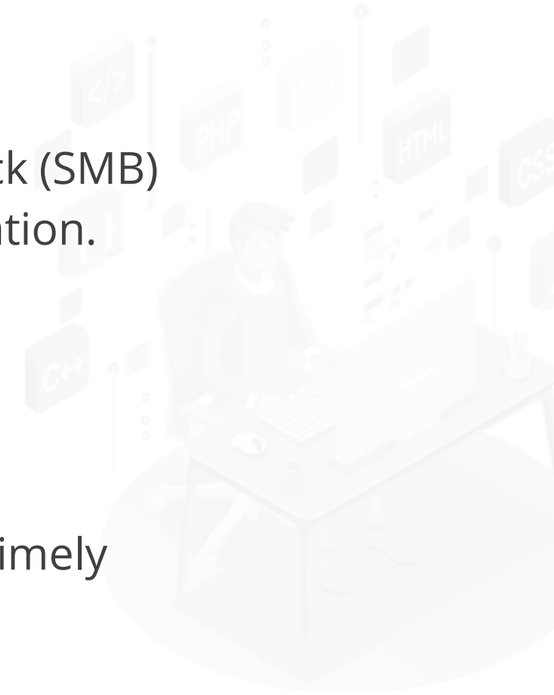


Working of AWS DataSync

Source: https://aws.amazon.com/blogs/storage/protect-your-file-and-backup-archives-using-aws-datasync-and-amazon-s3-glacier/

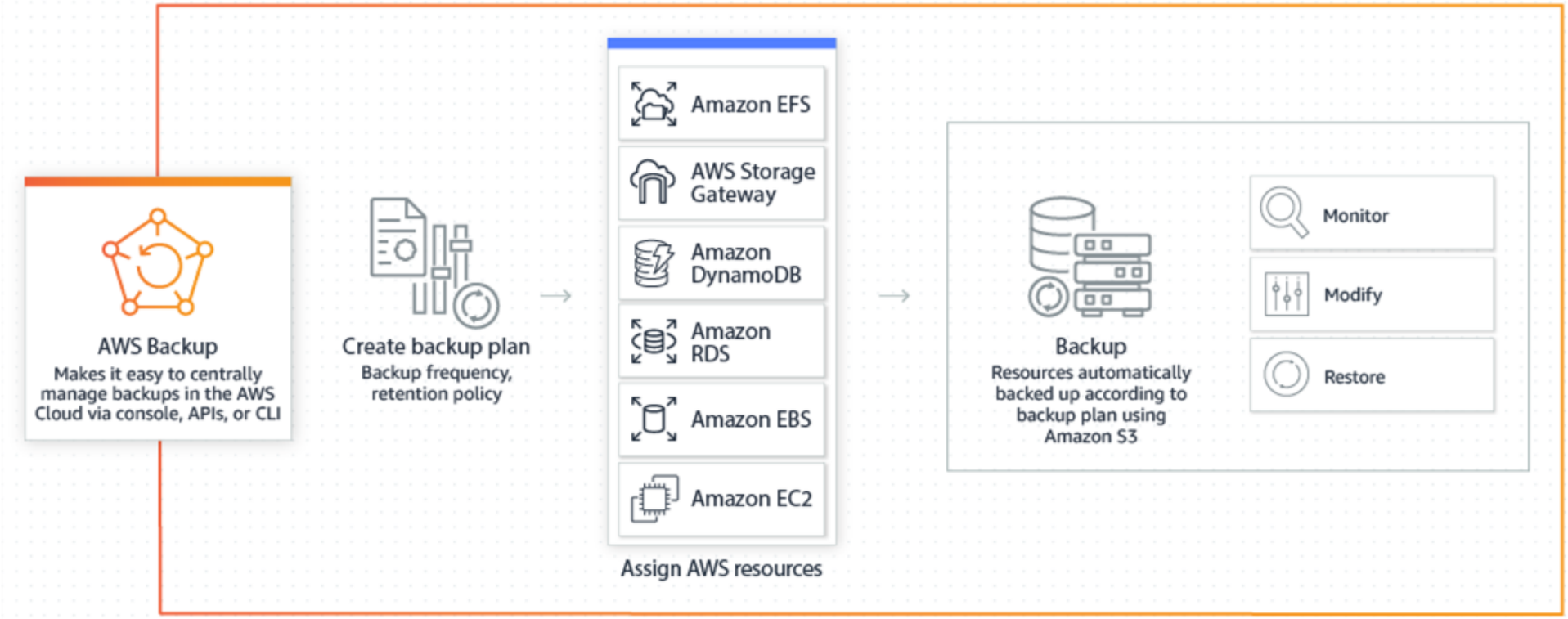
AWS DataSync

- Automatically handles tasks such as scripting copy jobs, scheduling and monitoring transfers, validating data, and optimizing network utilization.
- Agent connects to the user's Network File System (NFS), Server Message Block (SMB) storage, or self-managed object storage to avoid modifications in the application.
- Migrates hundreds of terabytes of data at a very high speed.
- Transfers active data sets or archives to AWS, transfer data to the cloud for timely analysis and processing, or replicate data to AWS for business continuity.



AWS Backup

AWS Backup provides centralized and automated data backup across AWS resources. It can be used to centrally configure backup policies and monitor backup activities.

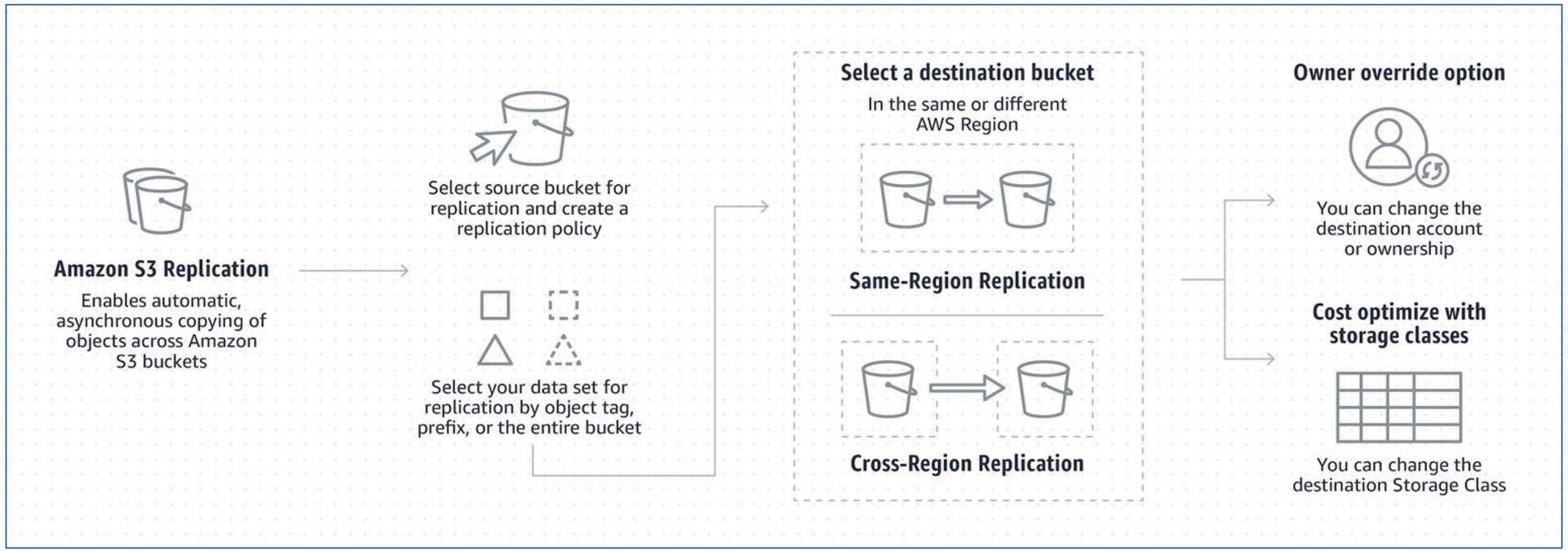


Working of AWS DataSync

Source: https://aws.amazon.com/backup/

Amazon S3 Replication

Amazon S3 Replication is an elastic, fully managed, low cost feature that replicates objects between buckets by providing great flexibility and functionality in the cloud storage.

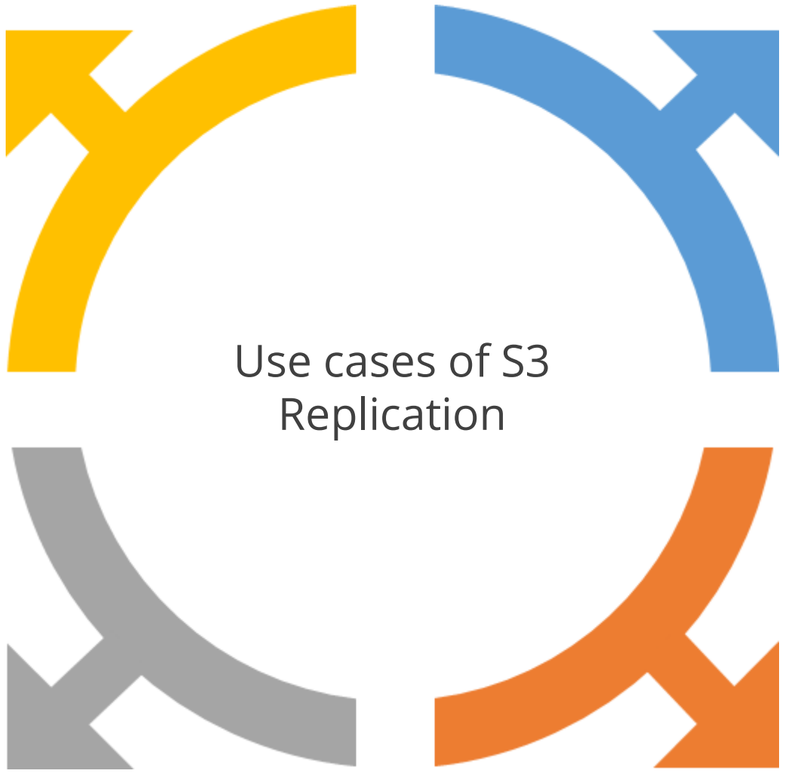


Working of S3 Replication

Source: https://aws.amazon.com/s3/features/replication/

S3 Replication Use Cases

Replicate objects while retaining metadata

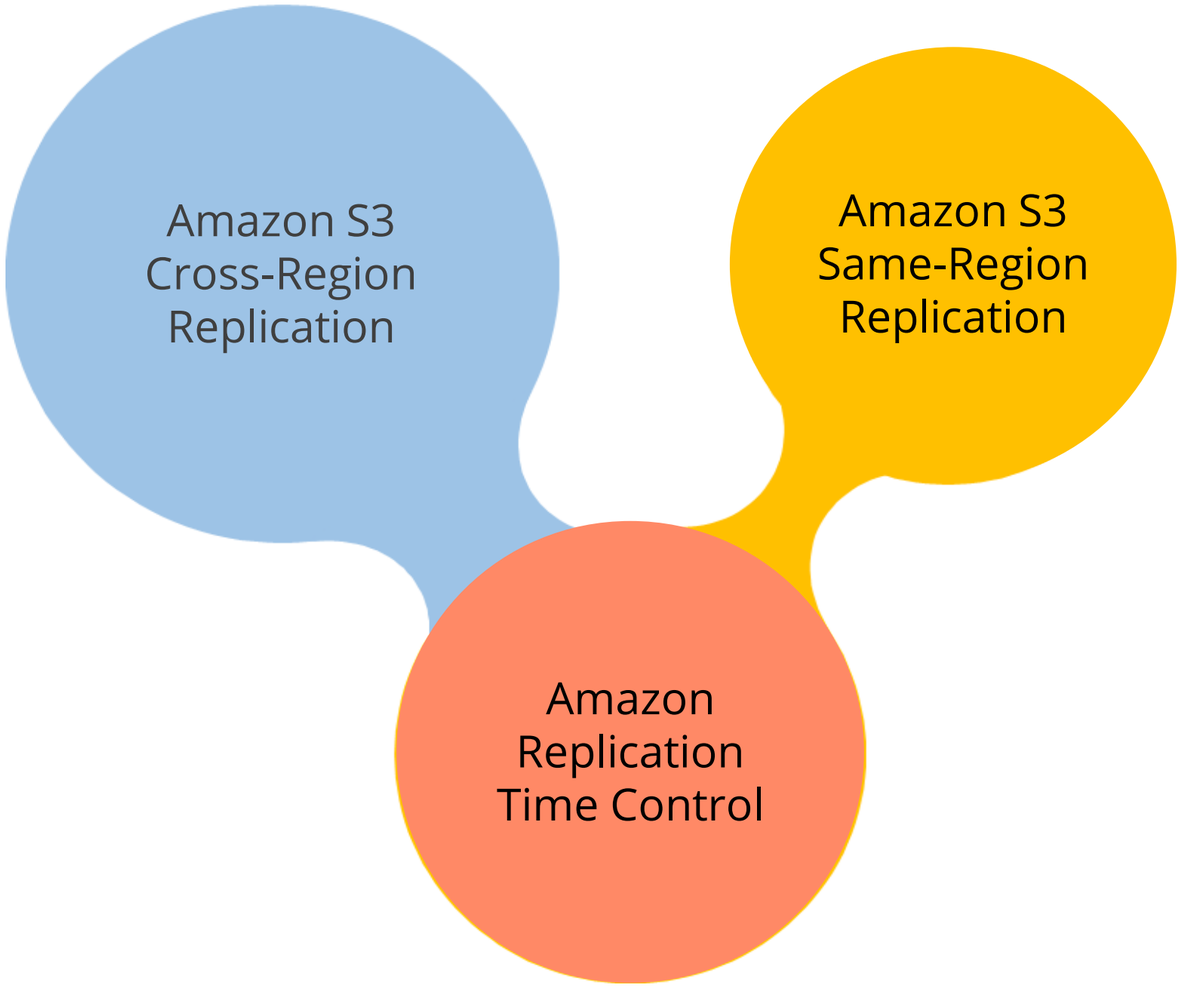


Replicate objects to more cost-effective storage classes

Replicate objects within 15 minutes

Maintain object copies under different ownership

Types of S3 Replication



Types of S3 Replication

Amazon S3 Cross-Region Replication (CRR)

- Automatically replicates data between buckets across different AWS Regions
- Configures to a source S3 bucket and replicates objects into a destination bucket in another AWS Region
- Replicates data at a bucket level, a shared prefix level, or an object level using S3 object tags
- Provides lower-latency data access in different geographic regions

Types of S3 Replication

Amazon S3 Same-Region Replication (SRR)

- Automatically replicates data between buckets within the same AWS Region
- Replicates data at a bucket level, a shared prefix level, or an object level using S3 object tags
- Addresses data sovereignty and compliance requirements
- Changes account ownership for the replicated objects to protect accidental data deletion
- Collects logs from various S3 buckets for in-region processing, or configures live replication

Types of S3 Replication

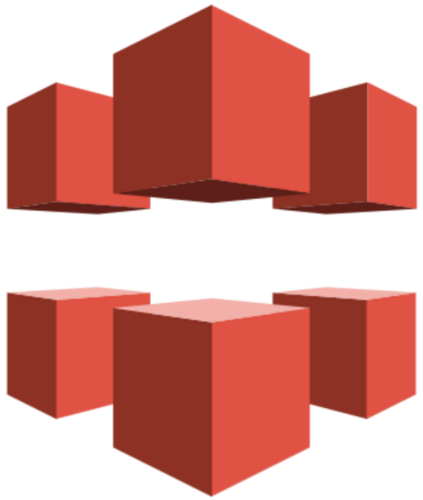
Amazon Replication Time Control

- Replicates 99.99% of new objects stored in Amazon S3 within 15 minutes
- Meets compliance or business requirements for data replication and provides visibility into S3 Replication activity
- Provides S3 Replication metrics that monitor:
  - The total number of S3 API operations that are pending replication
  - The total size of objects pending replication
  - The maximum replication time to the destination AWS Region

Amazon CloudFront

Amazon CloudFront

Amazon CloudFront is a global content delivery network (CDN) service that securely delivers content to the end users with low latency and high transfer speeds.



Amazon CloudFront

Amazon CloudFront

Amazon CloudFront Points of Presence (POPs)

Source: https://aws.amazon.com/blogs/aws/200-amazon-cloudfront-points-of-presence-price-reduction/

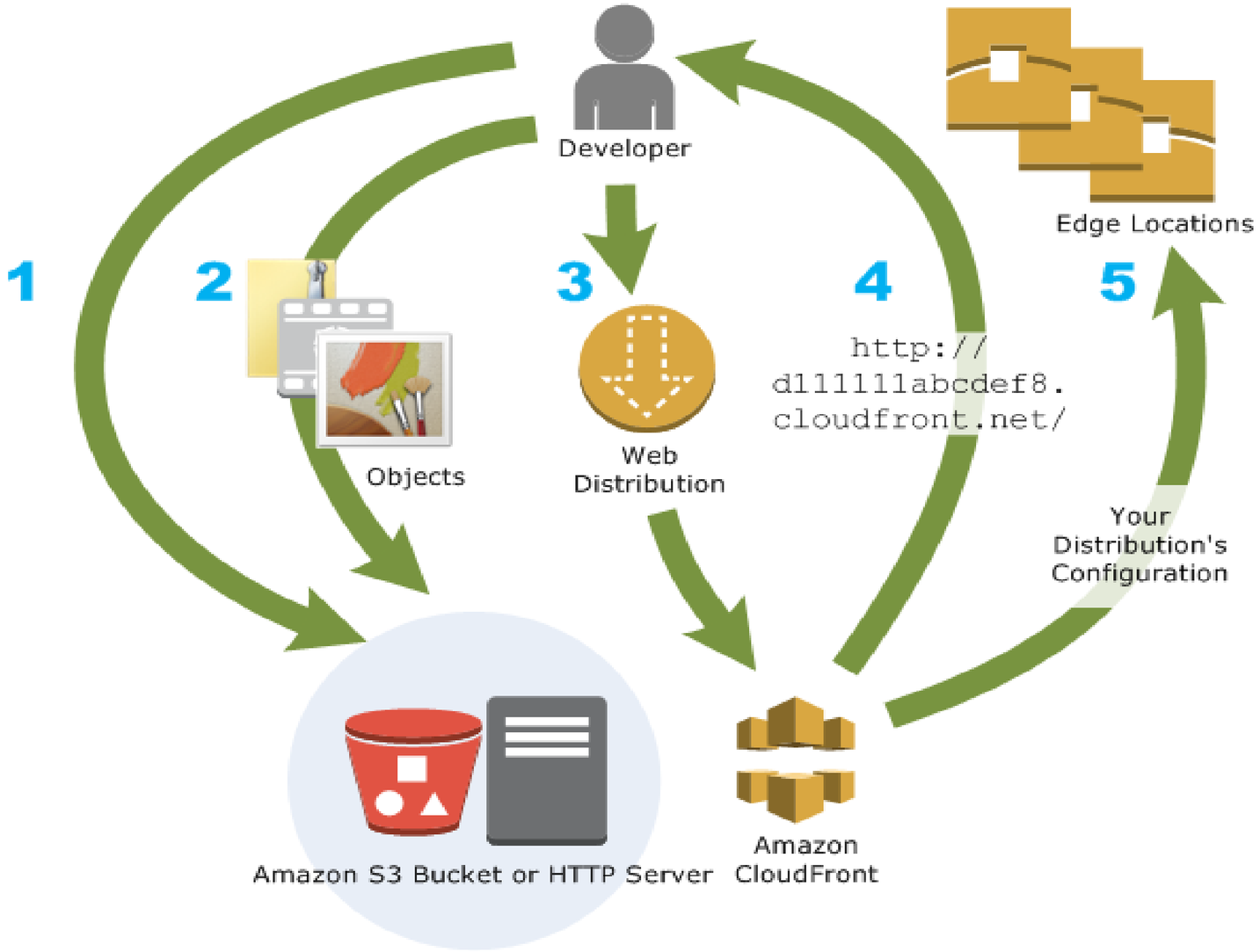
Content Delivery Using CloudFront

Steps to setup a CloudFront distribution to deliver content:

1. Specify origin servers
2. Upload data files to the origin servers
3. Create a CloudFront distribution

CloudFront assigns a domain name to the new distribution and sends the distribution's configuration to all of its edge locations

Content Delivery Using CloudFront



Steps to set up a CloudFront distribution to deliver content

Source: https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/Introduction.html

Configure an Amazon CloudFront Distribution

Duration: 15 min.

Problem Statement:

Configure an Amazon CloudFront distribution to serve a static website hosted on Amazon S3

ASSISTED PRACTICE

Assisted Practice: Guidelines to Configure an Amazon CloudFront distribution

Steps to perform:

1. Go to your Amazon Console
2. Open the CloudFront Management Console
3. Click on the Create distribution button
4. Fill in the details about the distribution
5. On review page, click on the Create distribution button

AWS Snowball

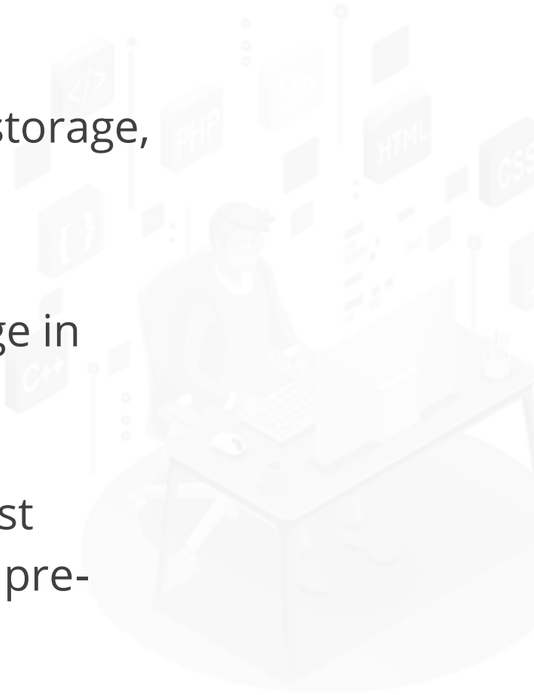
AWS Snowball

AWS Snowball, a part of the AWS Snow Family, is a petabyte-scale edge computing, data migration, and edge storage device that transfers large amounts of data between Amazon S3 and on-site data storage location at a faster-than-internet speed.

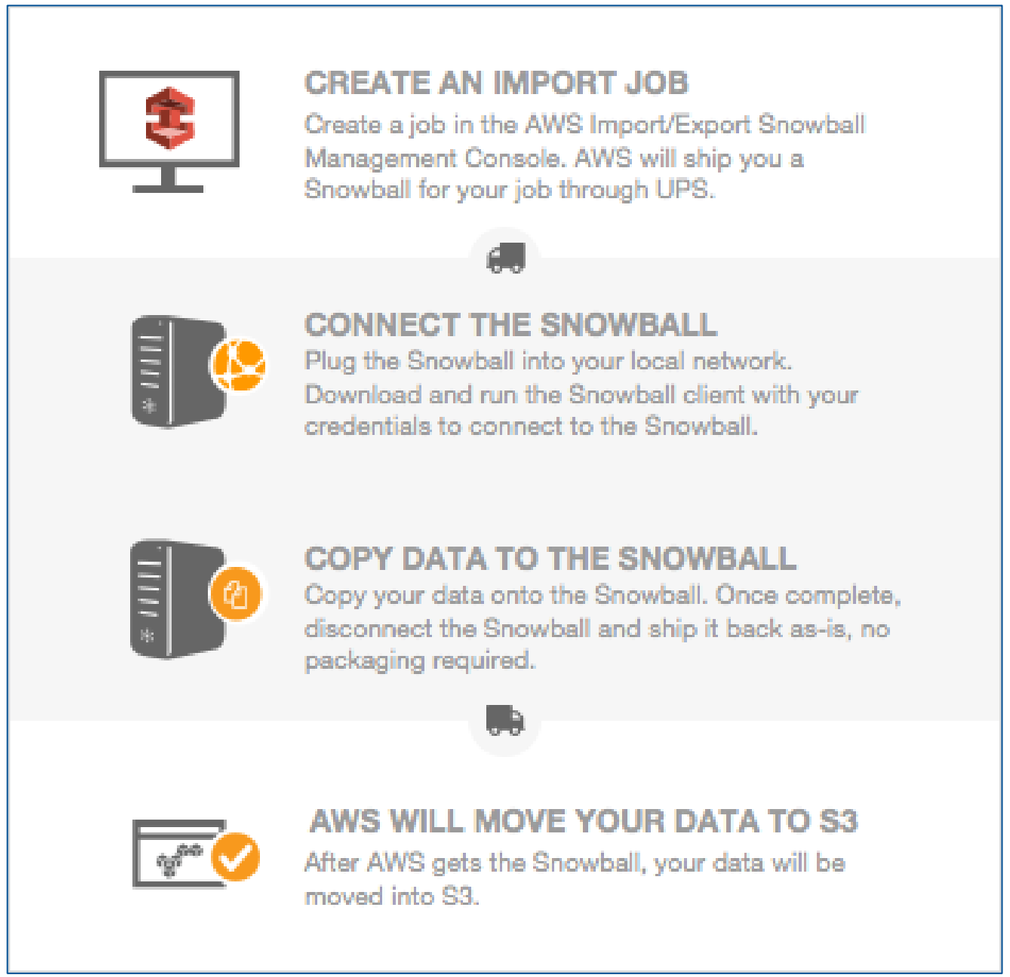
AWS Snowball device

AWS Snowball

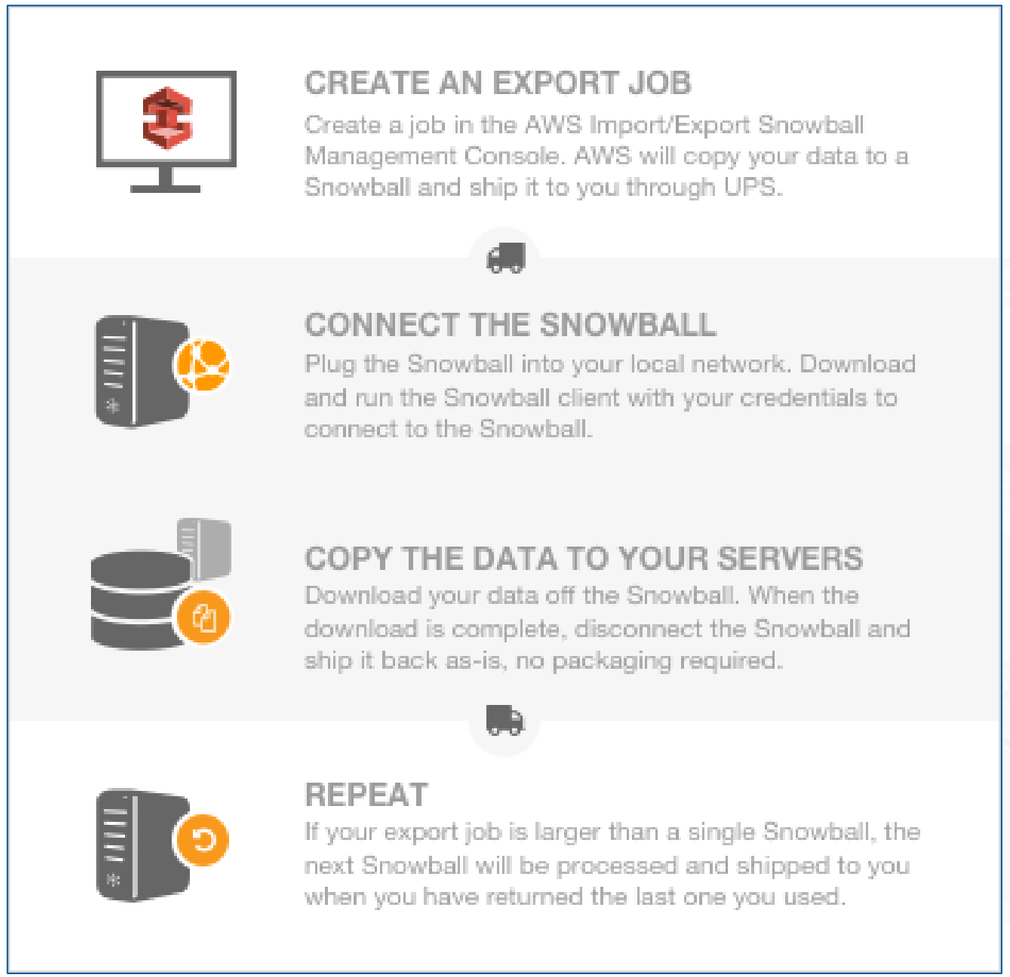
- Snowball Edge Storage Optimized devices provide both block storage and Amazon S3-compatible object storage, and 40 vCPUs.
- Snowball Edge Compute Optimized devices provide 52 vCPUs, block and object storage, and an optional GPU.
- These devices offer data collection, machine learning and processing, and storage in environments with intermittent connectivity or in extremely remote locations.
- It supports Amazon EC2 instances and AWS Lambda functions to develop and test applications in AWS cloud, and then deploy them on Snowball devices to collect, pre-process, and ship data to AWS.



Working of AWS Snowball



Importing data into Amazon S3 with Snowball



Exporting data from Amazon S3 with Snowball

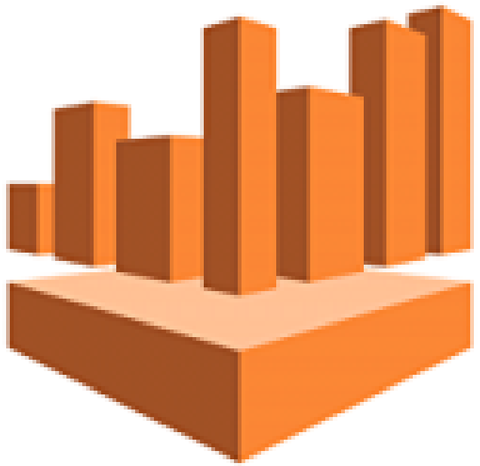
Source: https://docs.aws.amazon.com/snowball/latest/ug/getting-started.html

TECHNOLOGY

Amazon Athena and Macie

Amazon Athena

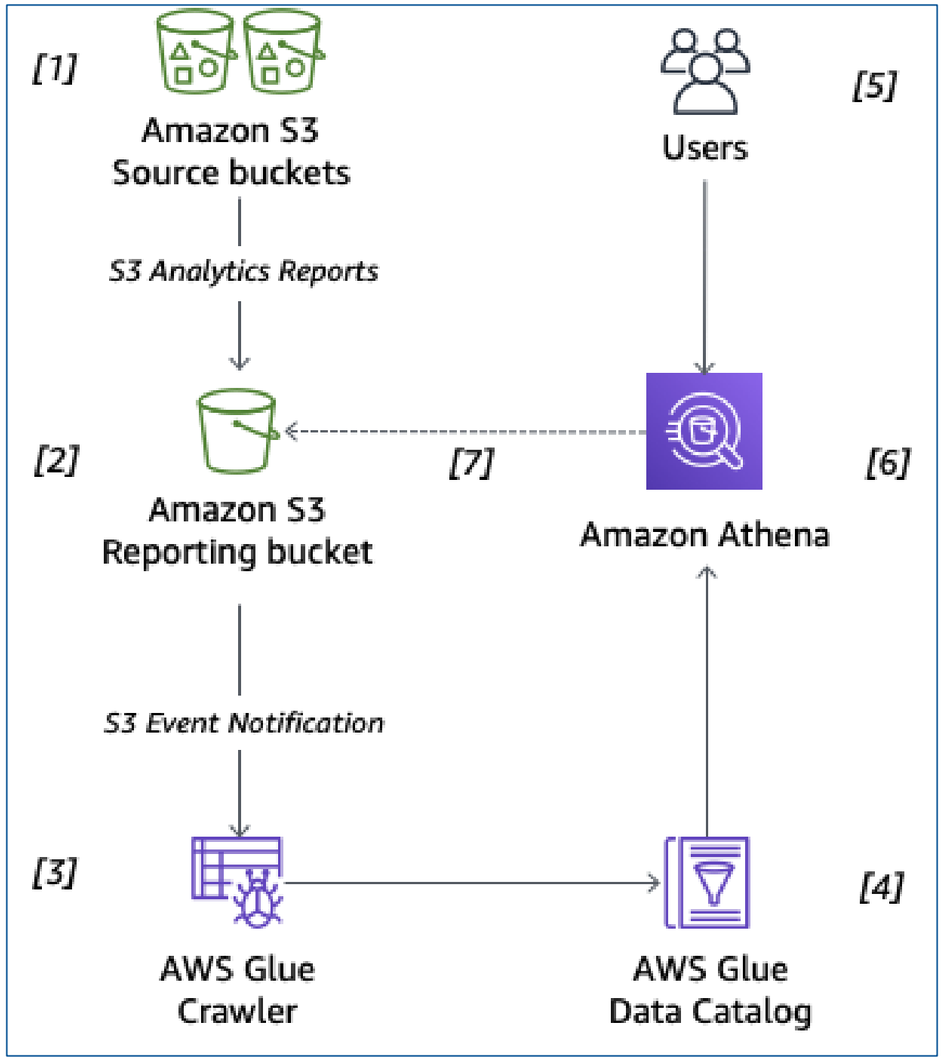
Amazon Athena is an interactive query service that allows easy analysis of data in Amazon S3 using standard SQL. It is serverless, so there is no infrastructure to manage, and users pay only for the queries that they run.



Amazon Athena

Amazon Athena

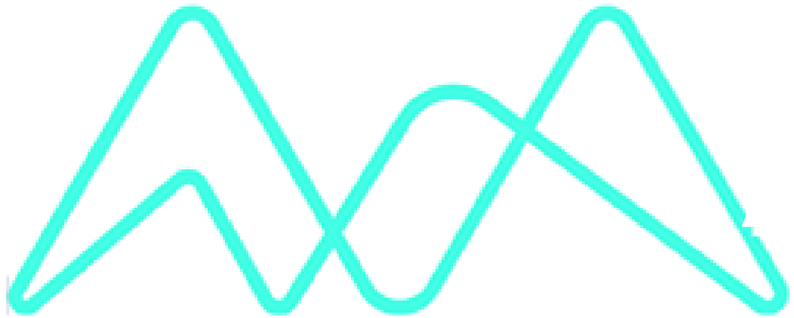
Steps to query S3 analytics data with Amazon Athena:



Source: https://aws.amazon.com/blogs/storage/query-amazon-s3-analytics-data-with-amazon-athena/

Amazon Macie

Amazon Macie is a fully managed data security and data privacy service that uses machine learning and pattern matching to discover and protect the sensitive data at scale.



Amazon Macie

Amazon Macie Features

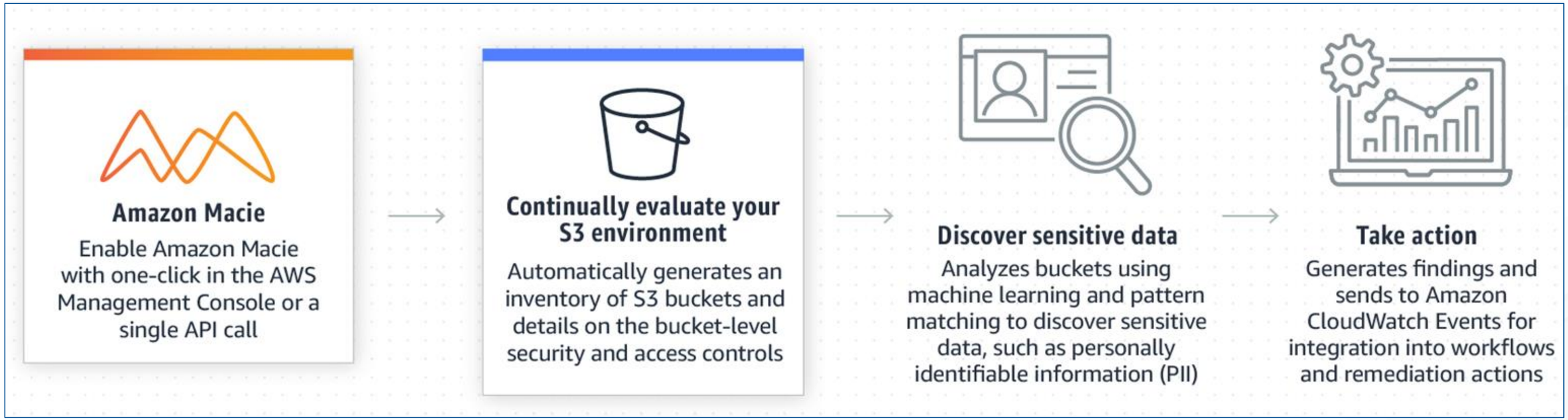
Data Security Automation

Data Security and Monitoring

Data Visibility for Proactive Loss Prevention

Data Research and Reporting

Working of Amazon Macie



Steps to demonstrate the working of Amazon Macie

Source: https://aws.amazon.com/blogs/storage/query-amazon-s3-analytics-data-with-amazon-athena/

Key Takeaways

- Amazon S3 is a web-based storage service for online backup and archiving of data.
- Unversioned, versioning-enabled, and versioning-suspended are the three states of versioning.
- An S3 bucket policy is a resource-based AWS IAM policy that grants other AWS accounts or IAM users access to the objects in an S3 bucket.
- Amazon S3 Standard, Intelligent-Tiering, Standard-IA, One Zone-IA, Glacier, and Glacier Deep Archive are the storage classes of Amazon S3.
- Amazon S3 Select allows applications to fetch only a subset of data from an object using simple SQL expressions.

Key Takeaways

- AWS DataSync allows fast transfer of data online between on-premises storage and Amazon storage services.
- Amazon S3 Cross-Region Replication, Same-Region Replication, and Amazon Replication Time Control are the three types of S3 Replication.
- Amazon CloudFront is a global CDN service that offers fast and secure content delivery to end users.
- Amazon Athena is an interactive query service that allows easy data analysis in Amazon S3 using standard SQL.
- Amazon Macie is a data security and privacy service that discovers and protects the sensitive data at scale.

Setting Up a CloudFront Distribution with an S3 Bucket

Duration: 60 min.

Problem Statement:

As an AWS solutions architect, you have been asked to set up a CloudFront distribution with an S3 bucket to host a web app.

Perform the following:

- Create an S3 bucket with versioning enabled
- Configure the S3 bucket for hosting a static website
- Create a CloudFront distribution
- Verify the distribution with the domain name

LESSON-END PROJECT