Cloud Computing

By

Dr. P. Sreedhar

S3 (Simple Storage service)

AWS S3

- AWS Simple Storage Service (S3) is a web service exposed over the Internet to store

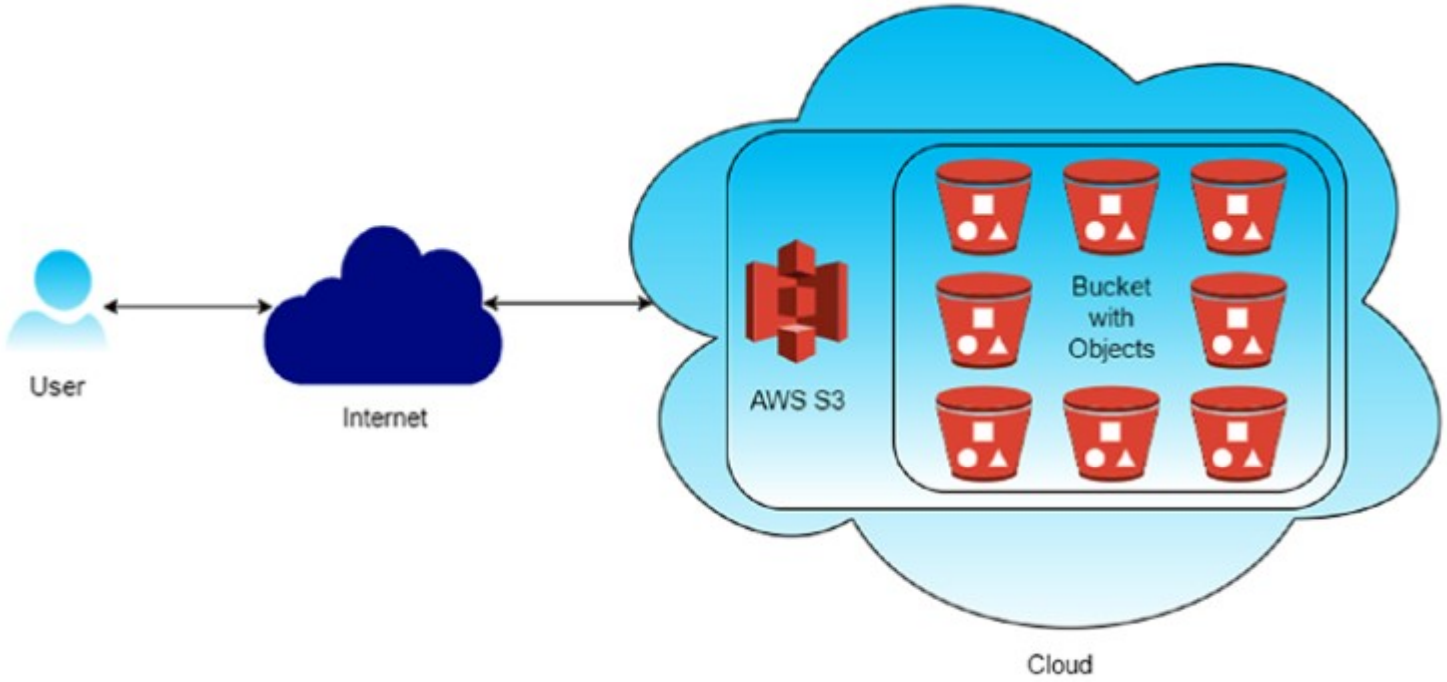
  and retrieve data.
- It provides infrastructure that is reliable, scalable, and cost-efficient.
- Amazon uses AWS S3 to serve its global network of web sites. You can even host static web

  sites over AWS S3.
- AWS S3 stores data objects in buckets. A bucket is simply a logical container that is used to identify the namespace of data objects.
- Each AWS user account can create up to 100 buckets.
- You can also create folders under the bucket, and inside folders you can store

  data objects. Consequently, it becomes easier to separate the logical storage of data objects.
- Each data object has a unique key within a bucket. You can access the data objects using the following URL: http://**BUCKET_NAME**.s3.amazonaws.com/**DATA_OBJECT_KEY**.

Here **BUCKET_NAME** will be the actual bucket name in which **DATA_OBJECT_KEY** resides.

AWS S3



AWS S3

- The infinite object based storage facility in the cloud.
- A RESTful (or SOAP) data storage API.
- Full access control per file or user.
- Preauthorize direct uploads by users.
- Billed by capacity stored and transfer rates
- Individual Amazon S3 objects can range in 1 byte to 5 terabytes.

AWS S3

- **Creating/Listing Buckets**

  Globally unique “buckets” hold files on S3.

- **Uploading/Listing Files**

  Files are “objects” stored under “keys” on S3.

- **Downloading a File**

  You can stream files to and from S3.

AWS S3 features

- **Scalability**

AWS S3 provides infrastructure to store and retrieve data objects based on configuration. It scales automatically when we store a large number of data objects. There won't be any configuration changes from the user's end.

- **Availability**

AWS S3 ensures that data objects are available at all times. You don't need to manage anything if you have heavy traffic overnight.

- **Cost Efficiency**

AWS S3 follows the pay-as-you-go strategy, meaning you only have to pay for the data objects you have stored and retrieved on AWS S3. For startups and small-scale companies, it becomes inexpensive to use data storage infrastructure.

- **Static Web Hosting**

AWS S3 provides static web site hosting that can be mapped. When you have static HTML web sites, you can opt for hosting on S3 rather than purchasing the hosting space.

AWS S3 features

- The Good Scalable: effectively "unlimited" storage.
- Reliable: 99.999999999% guaranteed uptime and very redundant
- Inexpensive: rates for GB in cents

  First 1 TB / month $0.0300 per GB
- Universal: everything supports it
- Objects in S3 Trillions of Objects (000,000,000,000s) Servicing over 2 million requests per Second
- Secure

  Supports data transfer over SSL and automatic encryption of your data once it is uploaded. You can also configure bucket policies to manage object permissions and control access to your data using AWS IAM.

AWS S3 features

- **Send Event Notifications**

  Amazon S3 can send event notifications when objects are uploaded to Amazon S3. Amazon S3 event notifications can be delivered using Amazon SQS or Amazon SNS, or sent directly to AWS Lambda, enabling you to trigger workflows, alerts, or other processing.

- **High Performance**

  Amazon S3 supports multi-part uploads to help maximize network throughput and resiliency, and lets you choose the AWS region to store your data close to the end user and minimize network latency.

- **Integrated**

  Amazon S3 is integrated with other AWS services to simplify uploading and downloading data from Amazon S3 and make it easier to build solutions that use a range of AWS services. Amazon S3 integrations include Amazon CloudFront, Amazon Kinesis, Amazon RDS, Amazon Glacier, Amazon EBS, Amazon DynamoDB, Amazon Redshift, Amazon Route 53, Amazon EMR, and AWS Lambda.

AWS S3 Configuring

S3 can be used using AWS management console as well as awscli.

- Setting up aws cli:

  Linux

  Download the AWS CLI Bundled Installer using wget or curl.

  Unzip the package.

  Run the install executable.

  On Linux and OS X, here are the three commands that correspond to each step:

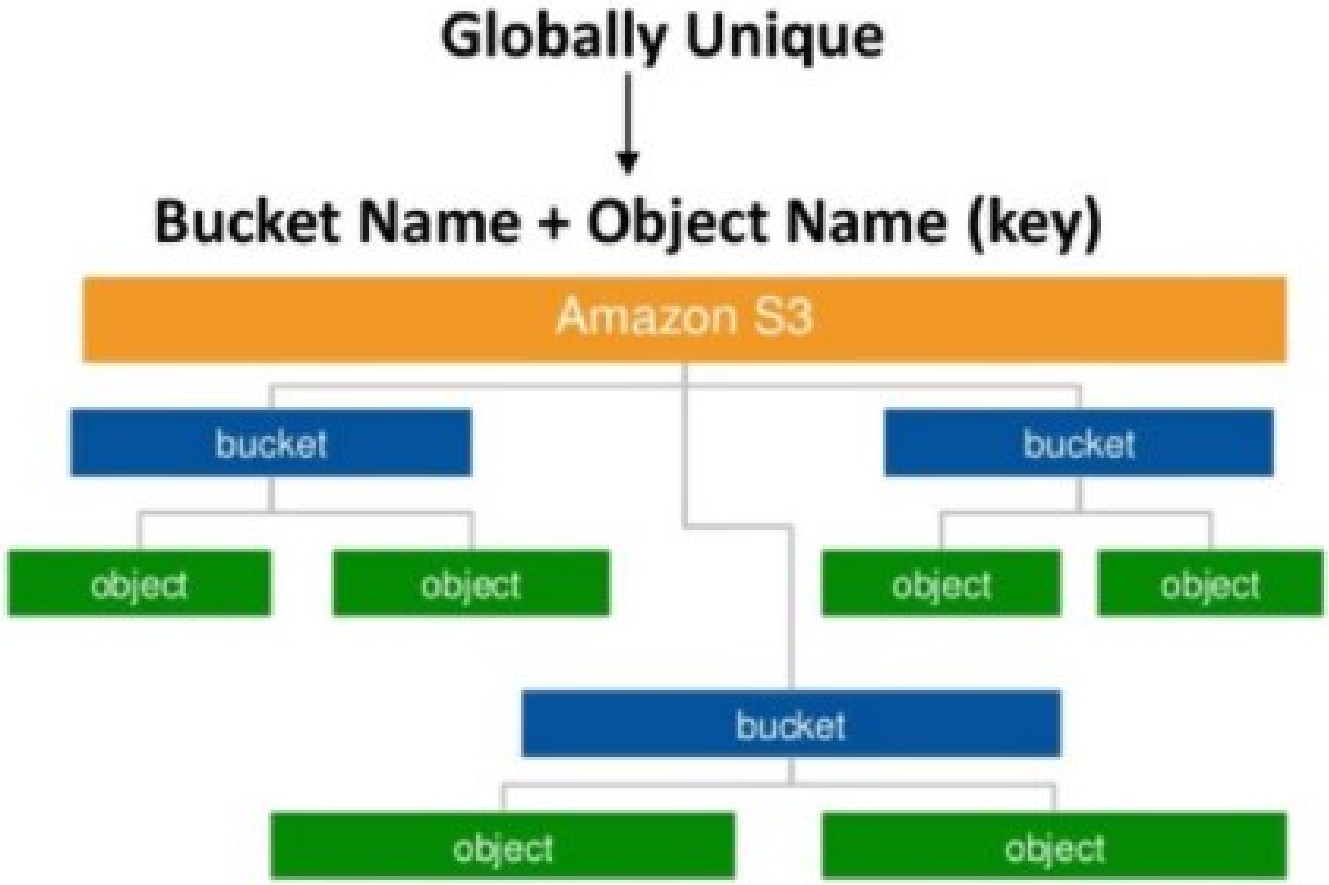
```
$ curl "https://s3.amazonaws.com/aws-cli/awscli-bundle.zip" -o "awscli-bundle.zip"
$ unzip awscli-bundle.zip
$ sudo ./awscli-bundle/install -i /usr/local/aws -b /usr/local/bin/aws
```

  Windows:

  https://s3.amazonaws.com/aws-cli/AWSCLI64.msi

S3 Buckets

S3 Namespace



AWS S3

- **Bucket**

  - Container for objects stored in S3
  - Unlimited size
  - Organize the namespace at the highest level
  - Internet accessible storage via HTTP/HTTPS
  - Global unique name

AWS S3 Bucket

- Each AWS account can own up to 100 buckets at a time.
- Bucket name must be globally unique.
- Bucket ownership is not transferable; however, if a bucket is empty, you can delete it.
- Default Bucket permissions for Account owner is “full”.
- Bucket names must be at least 3 and no more than 63 characters long.
- Bucket names must be a series of one or more labels. Adjacent labels are separated by a single period (.). Bucket names can contain lowercase letters, numbers, and hyphens. Each label must start and end with a lowercase letter or a number.
- Bucket names must not be formatted as an IP address (e.g., 192.168.5.4).

AWS S3 Bucket

- **Creating a Bucket**
- **Emptying a bucket**
  - You can use the Amazon S3 console to empty a bucket, which deletes all of the objects in the bucket without deleting the bucket.
  - When you empty a bucket that has S3 Bucket Versioning enabled, all versions of all the objects in the bucket are deleted.
- **Deleting a Bucket**
  - You can delete an empty Amazon S3 bucket

Viewing Bucket properties:

AWS S3 Bucket Properties

- **Bucket Versioning** – Keep multiple versions of an object in one bucket by using versioning. By default, versioning is disabled for a new bucket.
- **Tags** – With AWS cost allocation, you can use bucket tags to give billing for your use of a bucket. A tag is a key-value pair that represents a label that you assign to a bucket. To add tags, choose **Tags**, and then choose **Add tag**.
- **Default encryption** – Enabling default encryption provides you with automatic server-side encryption. Amazon S3 encrypts an object before saving it to a disk and decrypts the object when you download it.
- **Server access logging** – Get detailed records for the requests that are made to your bucket with server access logging. By default, Amazon S3 doesn't collect server access logs.
- **AWS CloudTrail data events** – Use CloudTrail to log data events. By default, trails don't log data events. Additional charges apply for data events.
- **Event notifications** – Enable certain Amazon S3 bucket events to send notification messages to a destination whenever the events occur. To enable events, choose **Create event notification**, and then specify the settings you want to use.
- **Transfer acceleration** – Enable fast, easy, and secure transfers of files over long distances between your client and an S3 bucket.
- **Object Lock** – Use S3 Object Lock to prevent an object from being deleted or overwritten for a fixed amount of time or indefinitely.
- **Requester Pays** – Enable Requester Pays if you want the requester (instead of the bucket owner) to pay for requests and data transfers.
- **Static website hosting** – You can host a static website on Amazon S3. To enable static website hosting, choose **Static website hosting**, and then specify the settings you want to use.

S3 Objects

AWS S3

- **Object**

  - Similar to files.
  - No hierarchy.
  - Objects are immutable.
  - Size up to 5 TB
  - Uniquely identified within a bucket by a key(name) and a version ID

AWS S3 Objects

- Amazon S3 is an object store that uses unique key-values to store as many objects as you want.
- You store these objects in one or more buckets, and each object can be up to 5 TB in size.

An object consists of the following:

1. Key
2. Version ID
3. Value
4. Metadata
5. Subresources
6. ACL (Access Control Information)

Key → The name that you assign to an object. You use the object key to retrieve the object.

Version ID → Within a bucket, a key and version ID uniquely identify an object. The version ID is a string that Amazon S3 generates when you add an object to a bucket.

Value → The content that you are storing.

- ✓ An object value can be any sequence of bytes. Objects can range in size from zero to 5 TB.

Metadata → A set of name-value pairs with which you can store information regarding the object. You can assign metadata, referred to as user-defined metadata, to your objects in Amazon S3. (Ex: object creation date, Is server-side-encryption enabled?, content MD5).

AWS S3 Objects

Sub resources → used to store object specific additional information.

Access Control Information → You can control access to the objects you store in Amazon S3. Amazon S3 supports both the resource-based access control, such as an access control list (ACL) and bucket policies, and user-based access control.

S3 Object keys

- The *object key* (or key name) uniquely identifies the object in an Amazon S3 bucket.
- When you create an object, you specify the key name, which uniquely identifies the object in the bucket.
- ***The Amazon S3 data model is a flat structure:*** *You create a bucket, and the bucket store objects. There is no hierarchy of subbuckets or subfolders. However, you can infer logical hierarchy using key name prefixes and delimiters as the Amazon S3 console does.*
- Amazon S3 supports buckets and objects, and there is no hierarchy. However, by using prefixes and delimiters in an object key name, the Amazon S3 console and the AWS SDKs can infer hierarchy and introduce the concept of folders.
- object key names can have alpha numeric, - _ . ‘ etc

AWS S3 Logging

Logging:

You can record the actions that are taken by users, roles, or AWS services on Amazon S3 resources and maintain log records for auditing and compliance purposes. To do this, you can use server access logging, AWS CloudTrail logging, or a combination of both. We recommend that you use AWS CloudTrail for logging bucket and object-level actions for your Amazon S3 resources.

AWS S3 Events

Events:

- You can configure the events on the bucket activities. When a new object is added to the

bucket, notification should be sent to SNS or SQS or trigger the lambda function to carry

out some processing.

For example, logs are being added to the bucket. You want to filter out logs so you can configure an event that will trigger whenever a log file is added to the bucket.

AWS S3 Versioning

Versioning:

- Versioning helps to store every version of objects. Versioned objects can be used to

recover the accidental override or expiration of an object. If bucket versioning is enabled,

then it can't be disabled. It can only be suspended.

AWS S3 Versioning

Life Cycle:

- ✓ The life cycle of objects can be managed by creating life-cycle rules. You can create rules

to move objects to other storage locations where objects are not accessed frequently, or

you can move objects to Glacier storage or even set an expiry date for the object**.**

AWS S3 – Cross Region Replication

Cross-Region Replication

- Cross-region replication will replicate the bucket with objects to another region. To enable cross-region replication, you need to enable versioning.

S3 using AWS CLI

S3 using AWS CLI

- ✓ **Create Bucket**
- ✓ **Bucket Permissions**
- ✓ ***Tags***
- ✓ ***Create, Empty and Delete Buckets***
- ✓ ***Upload, Downloading & Deleting objects***
- ✓ ***Creating Folders***

- ➢ **Logging**
- ➢ **Events**
- ➢ **Versioning**
- ➢ **Life Cycle**
- ➢ **Cross Region Replication**
- ➢ **Transferring Acceleration**

S3 using AWS CLI

Create Bucket

C:\> aws s3api create-bucket --bucket "bucket1withcli"

Ex:

1. aws s3api create-bucket --bucket "bucket2withcli"

2. aws s3api create-bucket --bucket "bucket2withcli" **--region** ap-south-1 **--create-bucket-configuration LocationConstraint=**ap-south-1

3. aws s3api create-bucket --bucket "bucket3withcli" **--acl** "public-read" **--region** ap-south-1 **--create-bucket-configuration LocationConstraint=**ap-south-1

4.

Upload, Downloading & Deleting objects

Ex: Uploading Object

1. ***aws s3api put-object --bucket*** *"bucket3withcli"* ***--key*** *"arrow.gif"* ***--body*** *D:/arrow.gif*
2. ***aws s3api delete-object --bucket*** *"bucket3withcli"* ***--key*** *"elecbike.jpg"*
3. ***aws s3api put-object --bucket*** *"bucket3withcli"* ***--key*** *"elecbike.jpg"* ***--body*** *D:/elecbike.jpg* ***--server-side-encryption*** *AES256*

Deletion of Object:

Ex:

1. **aws s3api delete-object --bucket** "bucket3withcli" **--key** "elecbike.jpg"

4. ***Create, Empty and Delete Buckets***

S3 using AWS CLI

Deletion of Bucket:

aws s3api delete-bucket --bucket "bucket3withcli“

Emptying Bucket:

aws s3 rm s3://*bucket3withcli* ***–recursive***

aws s3 rm s3://*bucket-name/doc* ***–recursive***

→ *removes objects that have the key name prefix doc, for example, doc/doc1 and doc/doc2.*

Hosting Static WS in S3

S3 using AWS CLI – Hosting ws

Hosting a Static Website on Amazon S3:

Steps:

1. *Create a Bucket*
2. *select 'Enable Website Hosting' option in properties window of Bucket.*
3. *Create Index.html file (or all files of website) and upload it to bucket*
4. *Make index.html (or all files) status as 'Make public' from the actions menu of the index.html (or all files of the website).*
5. *Identify the URL of the website and test it from the browser.*

Working with S3 using Java API

S3 using AWS SDK - Objects

Working with Objects:

- ✓ *Upload an Object*
- ✓ *List Objects*
- ✓ *Download an Object*
- ✓ *Copy, Move or Rename an Object*
- ✓ *Delete an Object*
- ✓ *Delete Multiple Objects at Once*

Working with ACLs and Access Policies

Bucket ACLs and Access Policies

Object ACLs and Access Policies