# AWS Simple Storage Service (S3)

Cloud Computing

Brenden west

# Contents

*Learning Outcomes*

- Features & components of AWS S3
- How to project data in S3 and manage access
- S3 object operations
- How to interact with S3 via AWS CLI or SDKs

Reading

- AWS Cloud Foundations - module 7 (section 2)
- AWS Cloud Developing - module 3

# Terminology

- **Bucket -**
- Object
-
-

# What is Amazon S3

- Object storage service offering scalability, availability, security, & performance
- Large-scale key-value data store, where values are file objects
- Objects can be any static content - e.g. documents, images, videos, audio, etc.
- S3 can generate event notifications when an object is created, deleted, restored, or replicated
- Can use storage policies to archive rarely used objects

# AWS S3 Components

- **Bucket** - container for objects that organize content & identify the responsible account.
  - Each bucket lives in a specific region.
  - Can maintain object versions (off by default)
- **Object key** - unique identifier for an object
- **Object** - any type of file.
  - Can use name prefix to organize in pseudo folder structure  (e.g. 2022/file1.txt)
  - Can be referred to by URL,
  - Consists of data & metadata (a set of name/value pairs that describe the object).
    - Metadata includes system-defined - e.g. creation date, size, version
    - Metadata can include user-defined values
  - Can be locked to prevent data changes

# Creating S3 Buckets

- Buckets are created via Management Console, CLI, or AWS SDK
- Bucket is created in a specific region & region can’t be changed later
- Name must be globally unique
- Name must be 3-63 characters, lowercase, alphanumeric and hyphens only

# Working with S3 Objects

- **PUT** - method to upload or copy an object to a bucket
  - Requires write permission
  - Must use multi-part upload for objects > 5 GB & should use for objects > 100 MB
- **GET** - method to retrieve objects from S3
  - Requires read access to the bucket
  - Can retrieve a complete object or a range of bytes (partial object)
  - Returns 404 for invalid or deleted objects
- **SELECT** - query S3 using SQL & return data for matching objects
  - Requires s3,GetObject permissions
  - Command includes a SQL expression and response format (e.g. JSON, CSV)
  - Can return all or specific data columns
- **DELETE** - method to permanently delete one or more objects
  - Need to specify version ID in case of versioned objects

# Securing S3 Data

- **Encryption** - data can by encrypted in transit & at rest
    - At rest encryption can be managed by client or by AWS S3 or KMS
- **Identity-based policies** - grant permissions to users, groups, or roles in same AWS account
- **Access control lists (ACLs)** - resource-based policies that manage access to buckets and objects
- **Bucket policies** - resource-based policy to supplement or replace ACLs.
    - Can grant permissions to other AWS accounts or IAM users
    - Object permissions apply only to objects the bucket owner creates
- **Pre-signed URLs** - provide PUT or GET access using permissions of user who created the URL
- **Cross-origin resource sharing (CORS)** - specific rules to identify origins that can access a bucket & allowed operations