



CLOUD COMPUTING APPLICATIONS

Amazon S3 BLOB Storage

Roy Campbell & Reza Farivar

Definition

Online file storage web service offered by Amazon Web Services. Amazon S3 provides storage through web service interfaces REST, SOAP, BitTorrent. (wikipedia)

<https://aws.amazon.com/s3/>

Use case

- Scalability, high availability, low latency – 99.99% availability
 - Files up to 5 terabytes
 - Objects stored in buckets owned by users
 - User assigned keys refer to objects
-
- Amazon Machine Images (exported as a bundle of objects)
 - SmugMug, Hadoop file store, Netflix, reddit, Dropbox, Tumbler

Simple Storage Service (S3)

- A **bucket** is a container for objects and describes location, logging, accounting, and access control.
 - A bucket has a name that must be **globally unique**.
 - `http://bucket.s3.amazonaws.com`
 - `http://bucket.s3-aws-region.amazonaws.com`.
- A bucket can hold any number of **objects**, which are files of up to 5TB.
 - `http://bucket.s3.amazonaws.com/object`
 - `http://johnsmith.s3.amazonaws.com/photos/puppy.jpg`

Fundamental operations corresponding to HTTP actions:

`http://bucket.s3.amazonaws.com/object`

- POST a new object or update an existing object.
- GET an existing object from a bucket.
- DELETE an object from the bucket
- LIST keys present in a bucket, with a filter.

A bucket has a **flat directory structure**

S3 Weak Consistency Model

“Updates to a single key are **atomic**....”

“Amazon S3 achieves high availability by replicating data across multiple servers within Amazon's data centers. If a PUT request is successful, your data is safely stored. However:

- A process writes a new object to Amazon S3 and immediately attempts to read it. Until the change is fully propagated, Amazon S3 might report "key does not exist."
- A process writes a new object to Amazon S3 and immediately lists keys within its bucket. Until the change is fully propagated, the object might not appear in the list.
- A process replaces an existing object and immediately attempts to read it. Until the change is fully propagated, Amazon S3 might return the prior data.
- A process deletes an existing object and immediately attempts to read it. Until the deletion is fully propagated, Amazon S3 might return the deleted data.”

S3 Command Line Interface

```
aws s3 mb s3://bucket
...    cp localfile s3://bucket/key
        mv s3://bucket/key s3://bucket/newname
        ls s3://bucket
        rm s3://bucket/key
        rb s3://bucket

aws s3 help
aws s3 ls help
```