

Cloud Storage

Big Data Systems
School of Data Science
University of Virginia

Agenda

- > Objects
- > Amazon S3
- > HDFS

What are Objects?

Objects are fundamental units of data storage in cloud computing

Stored in flat namespace (not in a hierarchical structure)

Each object has three parts:

- > data
- > metadata
- > unique identifier (key)

Objects: Additional Properties

- > Objects are often immutable.
Typically overwrite an object rather than making edit.
- > Since objects are identified by unique keys, they can be spread across multiple servers and data centers

Amazon S3

Introducing Amazon S3

- > Amazon's cloud storage solution
- > Simple Storage Service; not so simple any more
- > Handles durability, replication, scalability behind the scenes
- > Object storage built to store and retrieve any amount of data from anywhere

Amazon S3 Durability

> Amazon S3 Standard and most S3 storage classes designed to provide 99.999999999% durability (“eleven nines”) over a year.

What this means:

For every 10,000 objects stored, you'd expect to lose one object on average every 10 million years

Amazon S3 Replication

- > S3 is durable thanks to **replication** across multiple data centers and advanced data protection
- > AWS automatically copies data across at least three physically separated Availability Zones within a single region.

Amazon S3 Summary

Amazon S3 on the Free Tier

PRODUCT	DESCRIPTION	FREE TIER OFFER DETAILS	PRODUCT PRICING
Amazon S3	Amazon Simple Storage Service (Amazon S3) is an object storage service that offers industry-leading scalability, data availability, security, and performance.	Use credits to access features in both Free and Paid plans , including: Storage of data across multiple zones with 99.99999999% durability Protection of data with encryption, access control, and Block Public Access settings	Amazon S3 Pricing

Amazon S3 Pricing

Amazon S3

Overview

Features ▾

Storage classes ▾

Pricing

Security

Resources ▾

FAQs

Please note that we list Storage Requests and Data Retrievals Pricing below the Storage Pricing table.

Region:

US East (N. Virginia) ▾

Storage pricing

S3 Standard - General purpose storage for any type of data, typically used for frequently accessed data

First 50 TB / Month

\$0.023 per GB

Next 450 TB / Month

\$0.022 per GB

Over 500 TB / Month

\$0.021 per GB

S3 Intelligent - Tiering * - Automatic cost savings for data with unknown or changing access patterns

Monitoring and Automation, All Storage / Month (Objects > 128 KB)

\$0.0025 per 1,000 objects

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

HDFS

About HDFS

> Lorem Ipsum