

AWS STORAGE SERVICES

Agenda

- ▶ Storage Services Overview
- ▶ Elastic Block Store (EBS)
- ▶ Simple Storage Service (S3)
- ▶ Elastic File System (EFS)
- ▶ Glacier
- ▶ Concepts and Usecases
- ▶ Pricing
- ▶ Hands-On

Storage Services



Amazon EBS (Elastic Block Storage)



Amazon S3 (Simple Storage Service)



Amazon Elastic File System (EFS)



Amazon FSx



Amazon Glacier



AWS Storage Gateway

Block vs File vs Object Storage



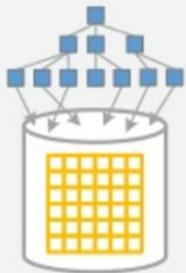
Block Storage

Raw Storage

Data organized as an array of unrelated blocks

Host File System places data on disk

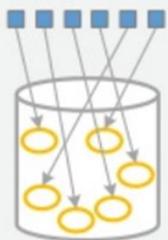
e.g.: Microsoft NTFS, Unix ZFS



File Storage

Unrelated data blocks managed by a file (serving) system

Native file system places data on disk



Object Storage

Stores Virtual containers that encapsulate the data, data attributes, metadata and Object IDs

API Access to data

Metadata Driven, Policy-based, etc

Storage – Characteristics

Durability	Availability	Security	Cost	Scalability	Performance	Integration
Measure of expected data loss	Measure of expected downtime	Security measures in place	Amount per storage unit, e.g. \$ / GB	Upward flexibility	Performance metrics	Ability to interact with other technologies

Amazon EBS Overview

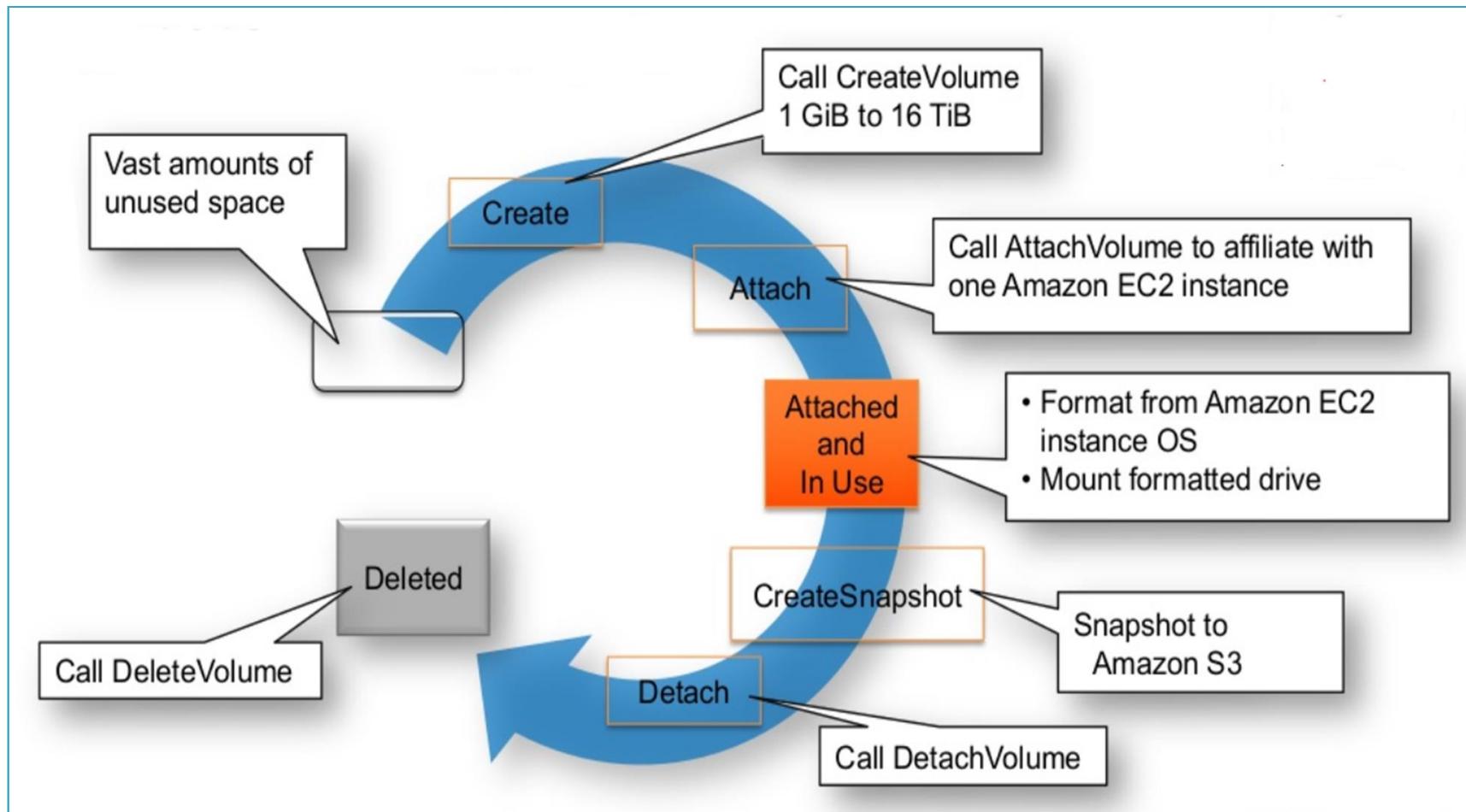
Amazon Elastic Block Storage (Amazon EBS) provides scalable persistent storage capacity in the Amazon Web Services (AWS) cloud



Amazon
EBS

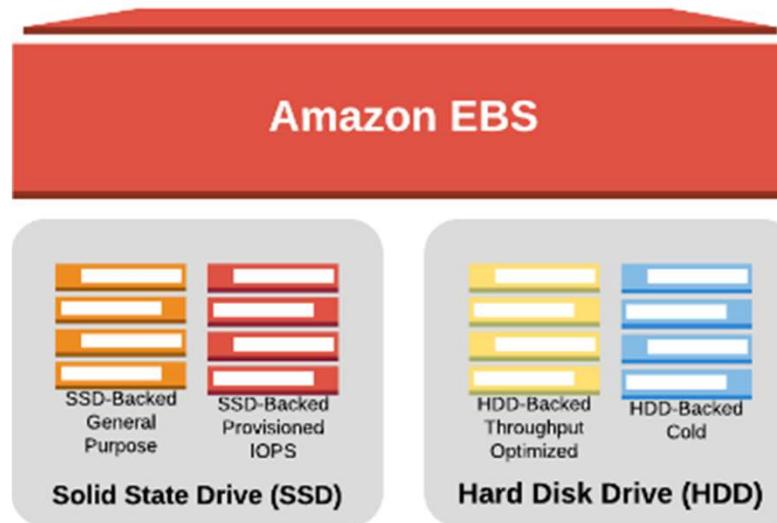
- **Persistent block level storage** volumes offer consistent and low-latency performance.
- Stored data is automatically replicated within its Availability Zone.
- Snapshots are stored durably in Amazon S3.

Amazon EBS Lifecycle



Amazon EBS Volume Types

- SSD-backed volumes are
 - Optimized for **transactional** workloads that involve **frequent read/write** operations with **small I/O size**.
 - Dominant in **IOPS** performance.
- HDD-backed volumes are
 - Optimized for **large streaming** workloads.
 - Dominant in **throughput** (measured in MiB/s).



Amazon EBS Facts

- EBS is recommended when data must be **quickly accessible** and requires **long-term persistence**.
- You can launch your EBS volumes as **encrypted** volumes – data stored at rest on the volume, disk I/O, and snapshots created from the volume are all encrypted.
- You can create **point-in-time snapshots** of EBS volumes, which are persisted to Amazon S3.

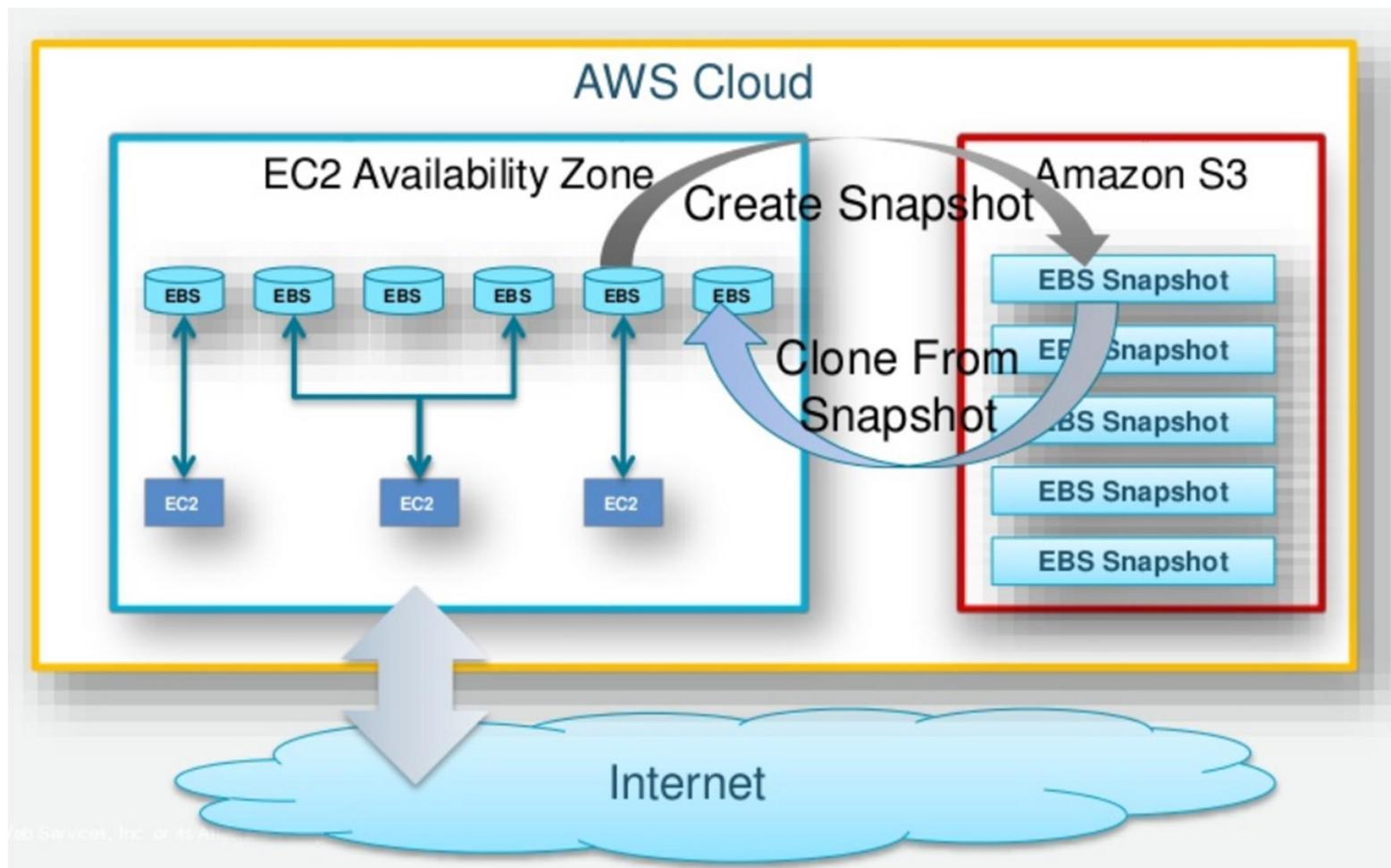
Amazon EBS Usecases

- **OS:** Use for boot/root volume, secondary volumes
- **Databases:** Scales with your performance needs
- **Enterprise applications:** Provides reliable block storage to run mission-critical applications
- **Business continuity:** Minimize data loss and recovery time by regularly backing up using EBS Snapshots
- **Applications:** Install and persist any application

Amazon EBS Volume Type Comparison

	Magnetic	General Purpose (SSD)	Provisioned IOPS (SSD)
Performance	Lowest Cost	Burstable	Predictable
Use Cases	Infrequent Data Access	Boot volumes Small to Medium DBs Dev & Test	I/O Intensive Relational & NoSQL
Media	Magnetic (HDD)	SSD	SSD
Max IOPS	100 on average with the ability to burst to hundreds of IOPS	Baseline 3 IOPS/GB Burstable to 3,000 IOPS	Consistently performed at provisioned level, up to 64,000 IOPS
Price	\$.05/GB/Month \$.05/million I/O	\$.10/GB/Month I/O Operations - Free	\$.125/GB/Month \$.065/provisioned IOPS

Amazon EBS Snapshots



Amazon S3 Overview

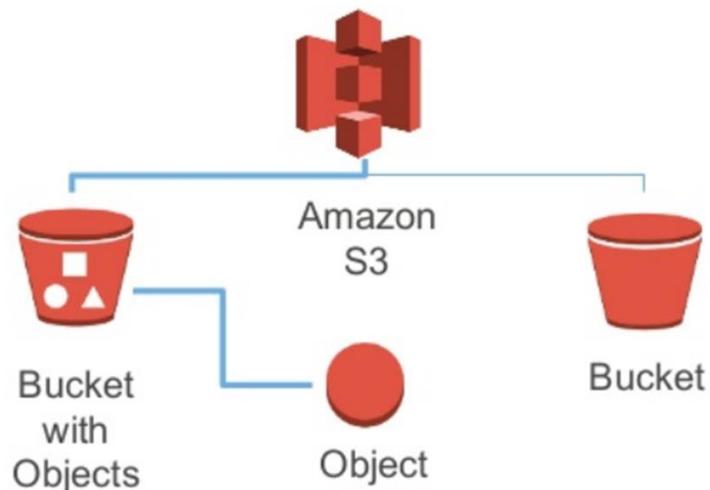
Amazon Simple Storage Service (Amazon S3) provides scalable object storage capacity in the Amazon Web Services (AWS) cloud



Amazon S3

- Storage for the Internet
- Natively online, HTTP access
- Storage that allows you to store and retrieve **any amount of data**, any time, from anywhere on the web
- **Highly scalable**, reliable, fast and durable

Amazon S3 Concepts



- Amazon S3 stores data as objects within **buckets**
- An object is composed of a file and optionally any **metadata** that describes that file
- You can have **up to 100 buckets** in each account
- You can **control access** to the bucket and its objects

Amazon S3 Facts

- Can store an **unlimited number of objects** in a bucket
- Objects can be **up to 5 TB**; no bucket size limit
- Designed for **99.999999999%** durability and **99.99%** availability of objects over a given year
- Can use **HTTP/S** endpoints to store and retrieve any amount of data, at any time, from anywhere on the web
- Is highly scalable, reliable, fast, and inexpensive
- Can use optional server-side **encryption** using AWS or customer-managed provided client-side encryption
- Auditing is provided by access logs
- Provides standards-based **REST** and SOAP interfaces

Amazon S3 Object Keys

An object key is the unique identifier for an object in a bucket.

<http://doc.s3.amazonaws.com/2006-03-01/AmazonS3.html>

Bucket

object/Key

Amazon S3 Usecases

- Storage and backup
- Application file hosting
- Media hosting
- Software delivery
- Store AMIs and snapshots

Object Storage Tiering

Data tiering using S3 Life Cycle Policies

S3 Standard

Primary data
Big Data Analytics
Small objects
Temporary scratch space

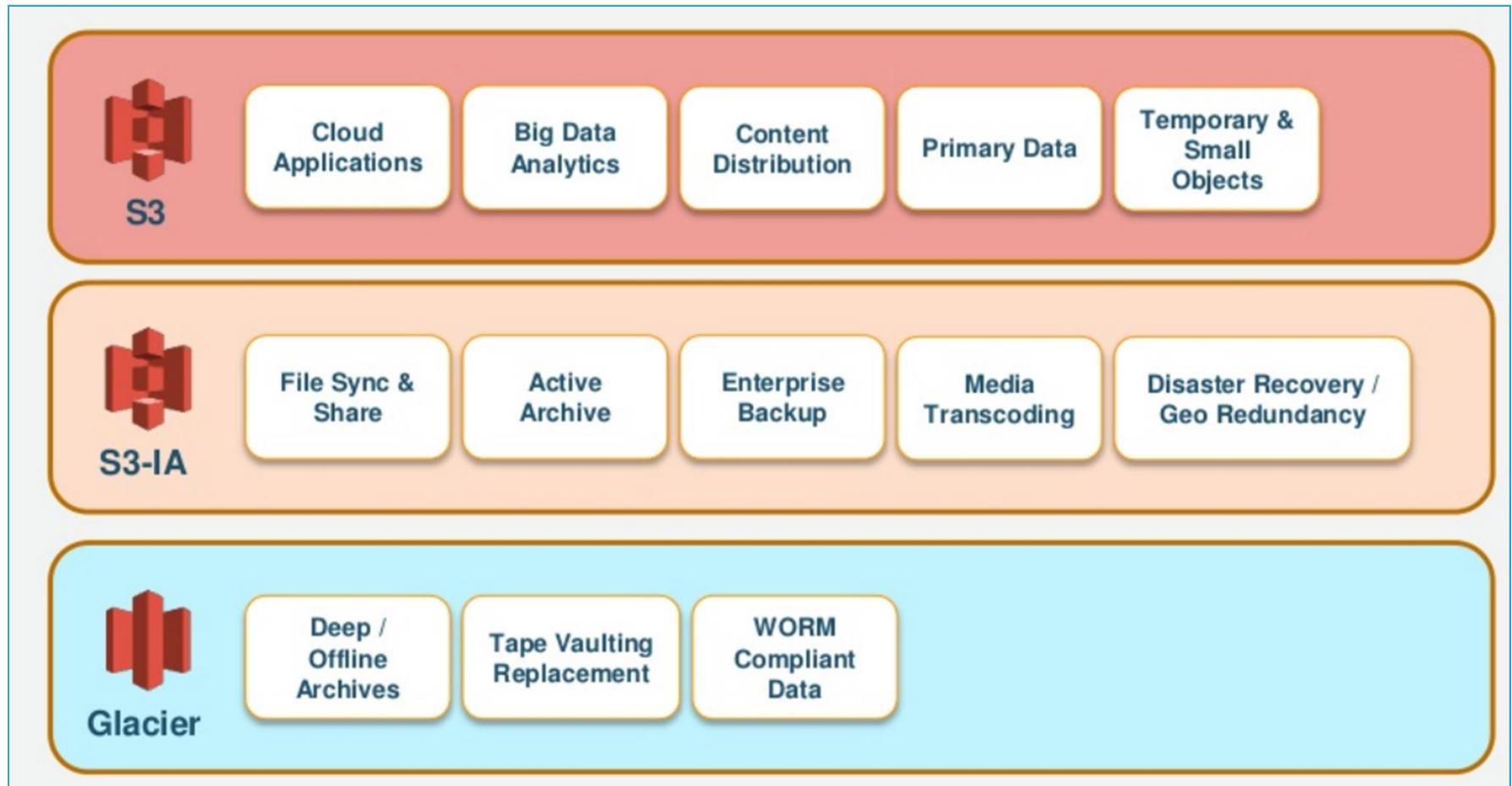
S3 - IA

File sync and share
Active Archive
Enterprise backup
Media transcoding
Geo-redundancy/DR

Glacier

Deep/offline archives
Tape vaulting replacement
WORM-compliant data

Object Storage Usecases



S3 Security

- You can **control access** to buckets and objects with:
 - Access Control Lists (ACLs)
 - Bucket policies
 - Identity and Access Management (IAM) policies
- You can upload or download data to Amazon S3 via **SSL** encrypted endpoints.
- You can **encrypt data** using AWS SDKs.

S3 Versioning

- Protects from **accidental overwrites and deletes** with no performance penalty.
- Generates a **new version with every upload**.
- Allows easily retrieval of deleted objects or **roll back** to previous versions.
- Three states of an Amazon S3 bucket
 - Un-versioned (default)
 - Versioning-enabled
 - Versioning-suspended



Versioning Enabled

S3 Object Lifecycle

Lifecycle management defines how Amazon S3 manages objects during their lifetime. Some objects that you store in an Amazon S3 bucket might have a well-defined lifecycle:

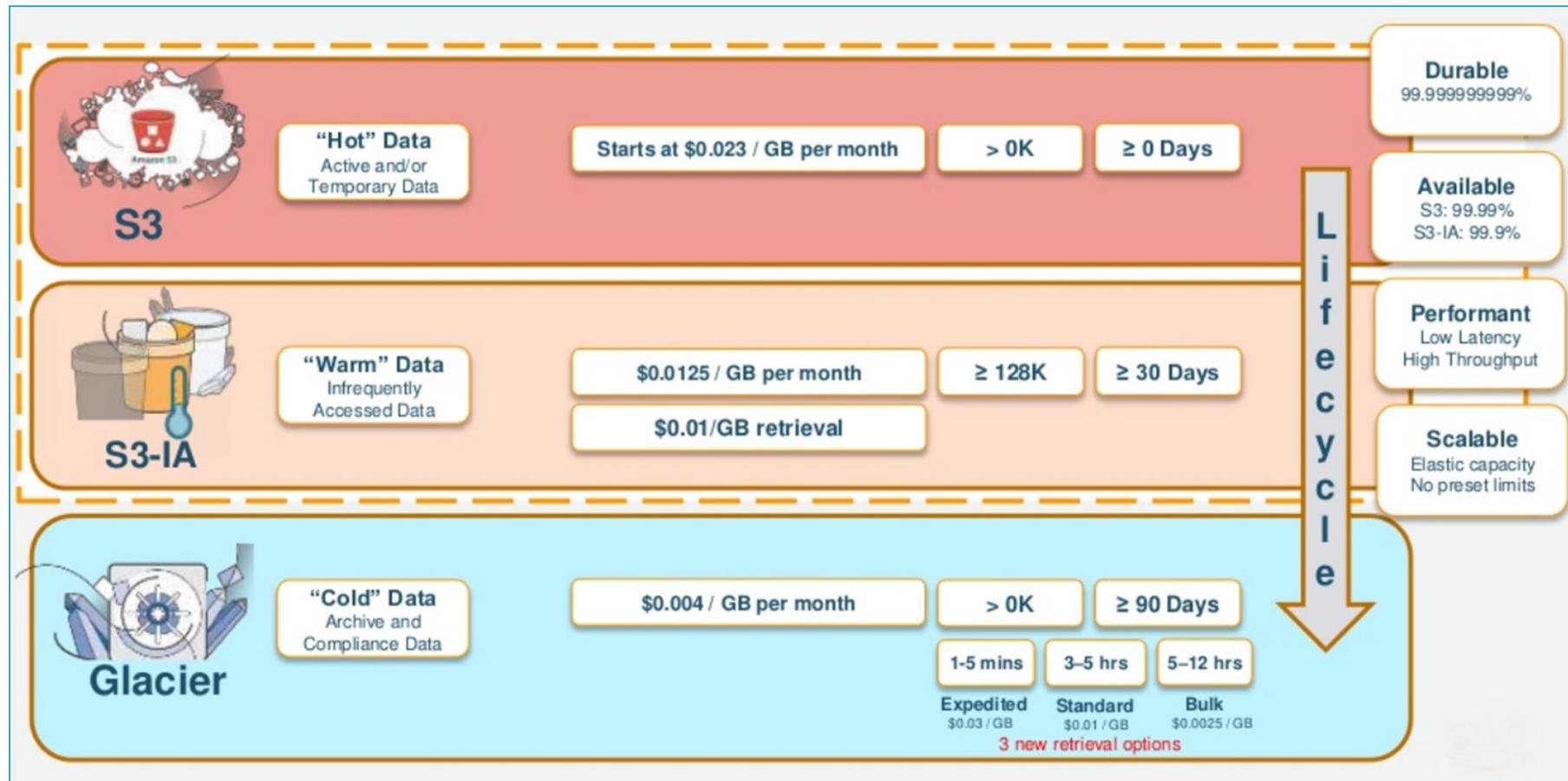
- Log files
- Archive documents
- Digital media archives
- Financial and healthcare records
- Raw genomics sequence data
- Long-term database backups
- Data that must be retained for regulatory compliance

S3 Pricing

- Pay only for what you use
- No minimum fee
- Prices based on location of your Amazon S3 bucket
- Estimate monthly bill using the **AWS Simple Monthly Calculator**
- Pricing is available as:
 - Storage Pricing
 - Request Pricing
 - Data Transfer Pricing: data transferred out of Amazon S3



Storage Class Tiering and Pricing



Amazon EBS vs Amazon S3

	Amazon EBS 	Amazon S3 
Paradigm	Block storage with file system	Object store
Performance	Very fast	Fast
Redundancy	Across multiple servers in an Availability Zone	Across multiple facilities in a Region
Security	EBS Encryption – Data volumes and Snapshots	Encryption
Access from the Internet?	No (1)	Yes (2)
Typical use case	It is a disk drive	Online storage

(1) Accessible from the Internet if mounted to server and set up as FTP, etc.
(2) Only with proper credentials, unless ACLs are world-readable

Amazon Glacier

- Low-Cost Archival Storage
- Secure
 - SSL & AES-256
- Durable
 - Designed for 99.99999999% durability
- Optimized for data archiving and backup
 - Suitable for RTO measured in hours
 - Includes storage costs and retrieval costs
- Three retrieval options: Expedited, Standard, Bulk
- As little as \$0.004 per GB/month
- Integrated with S3

Amazon EFS Overview

Amazon Elastic File System (Amazon EFS) provides scalable shared file storage capacity in the Amazon Web Services (AWS) cloud

Fully managed file system for EC2 instances

Provides standard file system semantics

Works with standard operating system APIs

Sharable across thousands of instances

Elastically grows to petabyte scale

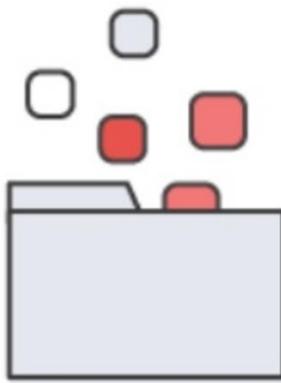
Delivers performance for a wide variety of workloads

Highly available and durable

NFS v4-based

Accessible from on-prem servers **New!**

Amazon EFS is Simple



Fully managed

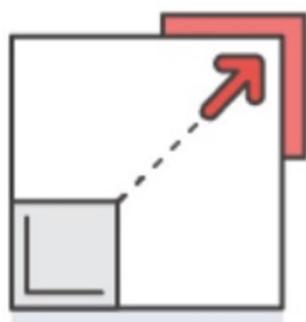
- No hardware, network, file layer
- Create a scalable file system in seconds!

Seamless integration with existing tools and apps

- NFS v4.1—widespread, open
- Standard file system access semantics
- Works with standard OS file system APIs

Simple pricing = simple forecasting

Amazon EFS is Elastic

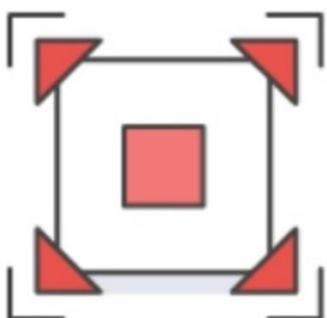


File systems grow and shrink automatically as you add and remove files

No need to provision storage capacity or performance

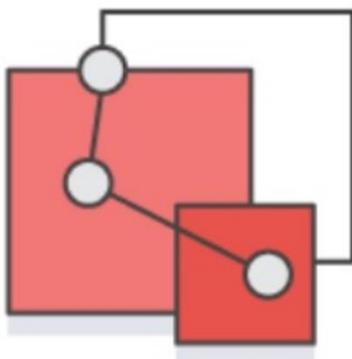
You pay only for the storage space you use, with no minimum fee

Amazon EFS is Scalable



- File systems can grow to petabyte scale
- Throughput and IOPS scale automatically as file systems grow
- Consistent low latencies regardless of file system size
- Support for thousands of concurrent NFS connections

Highly Durable and Highly Available



Designed to sustain AZ offline conditions
Resources aggregated across multiple AZs
Superior to traditional NAS availability models
Appropriate for Production / Tier 0 applications

Usecases

Big Data Analytics

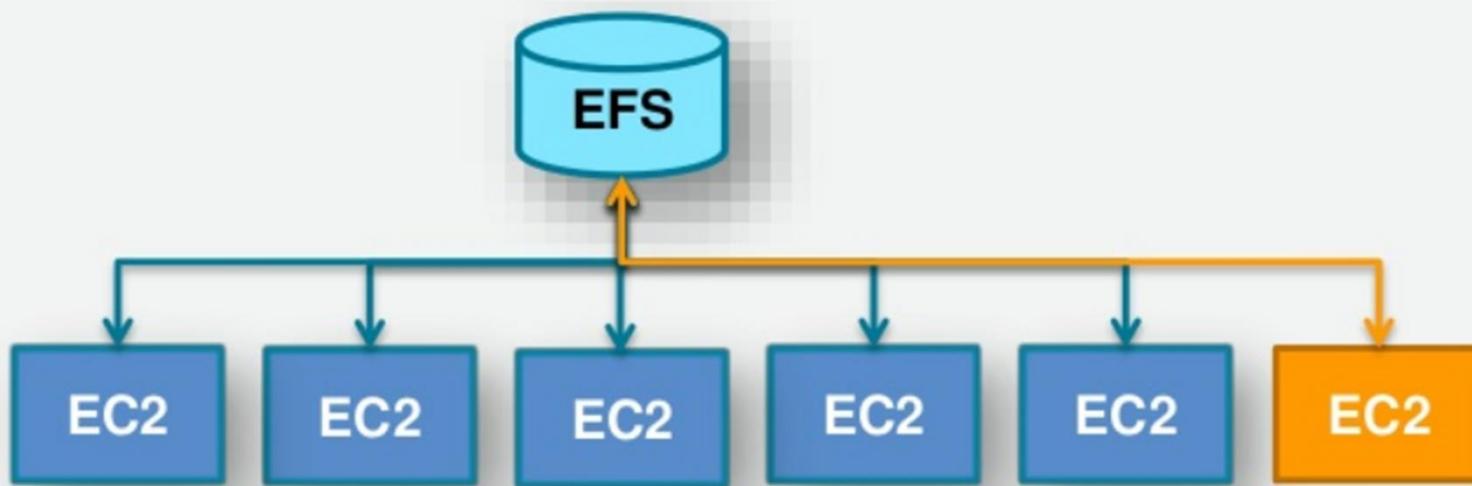
Media Workflow Processing

Web Serving

Content Management

Home Directories

EFS – Mounting



EFS DNS Name

availability-zone.file-system-id.efs.aws-region.amazonaws.com

Mount on machine

```
sudo mount -t nfs4 mount-target-DNS:/ ~/efs-mount-point
```

Thank You!