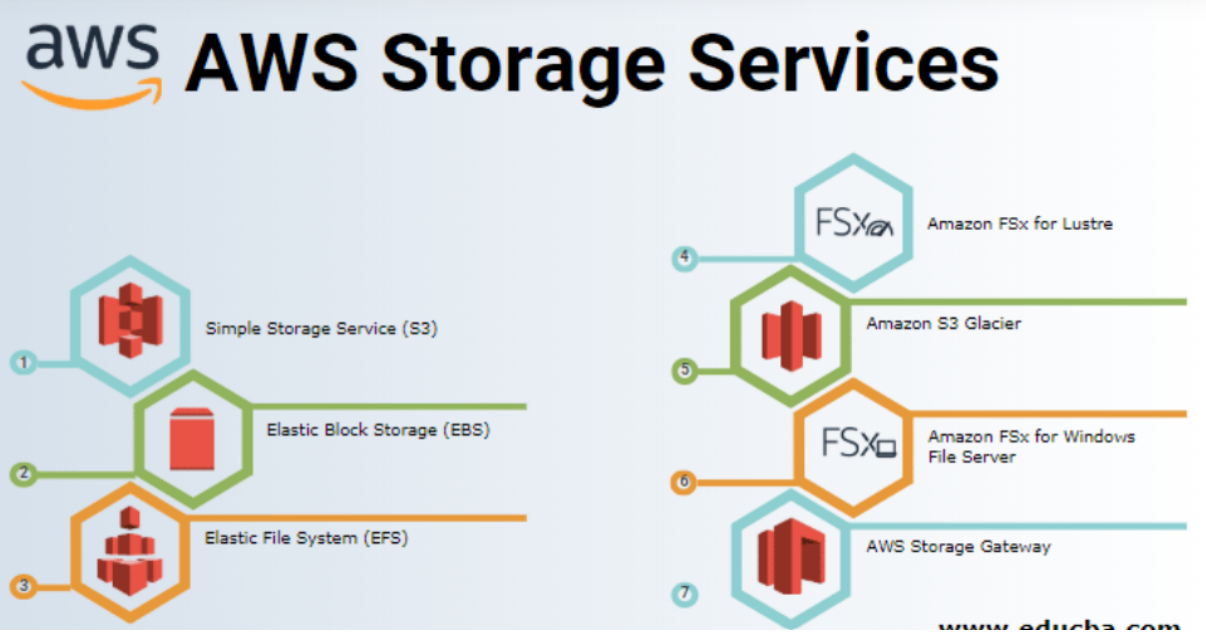
AWS Storage Services



Storage is divided into 3 family

1. Block storage (instance storage,EBS) Hdd is attached
2. Object storage (S3) upload, download file using rest API
3. File storage (EFS for Linux, FSx for windows) share storage with multiple devices

Amazon Simple Storage Service (S3)

Amazon S3 is object storage **built to store and retrieve any amount of data** from anywhere – web sites and mobile apps, corporate applications, and data from IoT sensors or devices.

You can store any type of file in S3.

S3 is designed to deliver 99.999999999% 11 ,9s durability, and stores data for millions of applications used by market leaders in every industry.

S3 provides comprehensive security and compliance capabilities that meet even the most stringent regulatory requirements.

S3 gives customers flexibility in the way they manage data for cost optimization, access control, and compliance.

Typical use cases include:

- **Backup and Storage** – Provide data backup and storage services for others.
- **Application Hosting** – Provide services that deploy, install, and manage web applications.
- **Media Hosting** – Build a redundant, scalable, and highly available infrastructure that hosts video, photo, or music uploads and downloads.
- **Software Delivery** – Host your software applications that customers can download.
- **Static Website** – you can configure a static website to run from an S3 bucket.

S3 provides query-in-place functionality, allowing you to run powerful analytics directly on your data at rest in S3. And Amazon S3 is the most supported cloud storage service available, with integration from the largest community of third-party solutions, systems integrator partners, and other AWS services.

Files can be anywhere from 0 bytes to 5 TB.

There is unlimited storage available.

Files are stored in buckets.

Buckets are root level folders.

Any subfolder within a bucket is known as a “folder”.

S3 is a universal namespace so bucket names must be unique globally.

There are seven S3 storage classes.

- S3 Standard (durable, immediately available, frequently accessed).
- S3 Intelligent-Tiering (automatically moves data to the most cost-effective tier).
- S3 Standard-IA (durable, immediately available, infrequently accessed).

- S3 One Zone-IA (lower cost for infrequently accessed data with less resilience).
- S3 Glacier Instant Retrieval (data that is rarely accessed and requires retrieval in milliseconds).
- S3 Glacier Flexible Retrieval (archived data, retrieval times in minutes or hours).
- S3 Glacier Deep Archive (lowest cost storage class for long term retention).

When you successfully upload a file to S3 you receive a HTTP 200 code.

S3 is a persistent, highly durable data store.

Persistent data stores are non-volatile storage systems that retain data when powered off.

This contrasts with transient data stores and ephemeral data stores which lose the data when powered off.

The following table provides a description of persistent, transient, and ephemeral data stores and which AWS service to use:

Storage Type	Description	Examples
Persistent Data Store	Data is durable and sticks around after reboots, restarts, or power cycles	S3, Glacier, EBS, EFS
Transient Data Store	Data is just temporarily stored and passed along to another process or persistent store	SQS, SNS
Ephemeral Data Store	Data is lost when the system is stopped	EC2 Instance Store, Memcached

Bucket names must follow a set of rules:

- Names must be unique across all of AWS.
- Names must be 3 to 63 characters in length.
- Names can only contain lowercase letters, numbers, and hyphens.
- Names cannot be formatted as an IP address.

Objects consist of:

- Key (name of the object).
- Value (data made up of a sequence of bytes).
- Version ID (used for versioning).
- Metadata (data about the data that is stored).

Subresources:

- Access control lists.
- Torrent.

Object sharing – the ability to make any object publicly available via a URL.

Lifecycle management – set rules to transfer objects between storage classes at defined time intervals.

Versioning – automatically keep multiple versions of an object (when enabled).

Encryption can be enabled for bucket.

Data is secured using ACLs and bucket policies.

Charges:

- Storage.
- Requests.
- Storage management pricing.
- Data transfer pricing.
- Transfer acceleration.

When you create a bucket you need to select the region where it will be created.

It is a best practice to create buckets in regions that are physically closest to your users to reduce latency.

Additional S3 Capability	**How it Works**
Transfer Acceleration	Speed up data uploads using CloudFront in reverse
Requester Pays	The requester rather than the bucket owner pays for requests and data transfer
Tags	Assign tags to objects to use in costing, billing, security etc.
Events	Trigger notifications to SNS, SQS, or Lambda when certain events happen in your bucket
Static Web Hosting	Simple and massively scalable static website hosting
BitTorrent	Use the BitTorrent protocol to retrieve any publicly available object by automatically generating a .torrent file

Amazon Elastic Block Store (EBS)

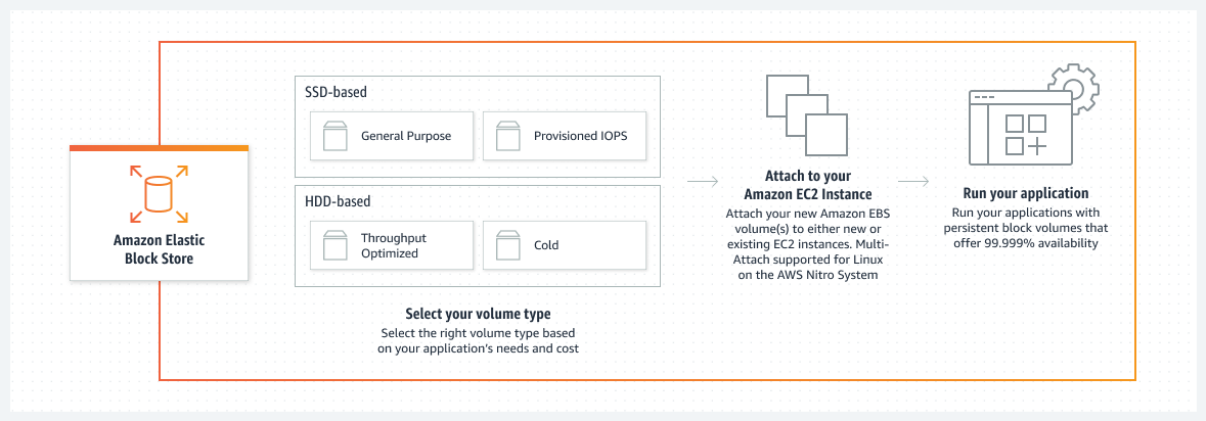
Amazon Elastic Block Store (Amazon EBS) provides persistent block storage volumes for use with Amazon EC2 instances in the AWS Cloud.

Each Amazon EBS volume is automatically replicated within its Availability Zone to protect you from component failure, offering high availability and durability.

Amazon EBS volumes offer the consistent and low-latency performance needed to run your workloads. With Amazon EBS, you can scale your usage up or down within minutes – all while paying a low price for only what you provision.

Benefits of Elastic Block Store

1. Scale fast
2. High performance
3. Optimize storage and cost
4. Secure
5. Simple data protection



The following EBS volumes appear most often on the AWS exams:

Volume Type	EBS Provisioned IOPS SSD (io1/io2)	EBS General Purpose SSD (gp2/gp3)	Throughput Optimized HDD (st1)	Cold HDD (sc1)
Short Description	Highest performance SSD volume designed for latency-sensitive transactional workloads	General Purpose SSD volume that balances price performance for a wide variety of transactional workloads	Low-cost HDD volume, designed for frequently accessed. Throughput intensive workloads	Lowest cost HDD volume designed for less frequently accessed workloads

Use Cases	I/O-intensive NoSQL and relational databases	Boot volumes, low-latency interactive apps, dev & test	Big-data, data warehouses, log processing	Colder data requiring fewer scans per day
Volume Size	4 GiB – 16 TiB	1 GiB – 16 TiB	125 GB – 16 TiB	125 GB – 16 TiB
Max IOPS / Volume**	64,000	16,000	500	250
Max Throughput*Volume**	1,000 MiB/s	250 MiB/s (gp2) 1000 MiB/s (gp3)	500 MiB/s	250 MiB/s
Can be boot volume?	Yes	Yes	No	No
EBS Multi-attach	Supported	Not Supported	Not Supported	Not Supported

EBS volume data persists independently of the life of the instance.

EBS volumes do not need to be attached to an instance.

You can attach multiple EBS volumes to an instance.

You cannot attach an EBS volume to multiple instances (use Elastic File Store instead).

EBS volumes must be in the same AZ as the instances they are attached to.

Termination protection is turned off by default and must be manually enabled (keeps the volume/data when the instance is terminated).

Root EBS volumes are deleted on termination by default.

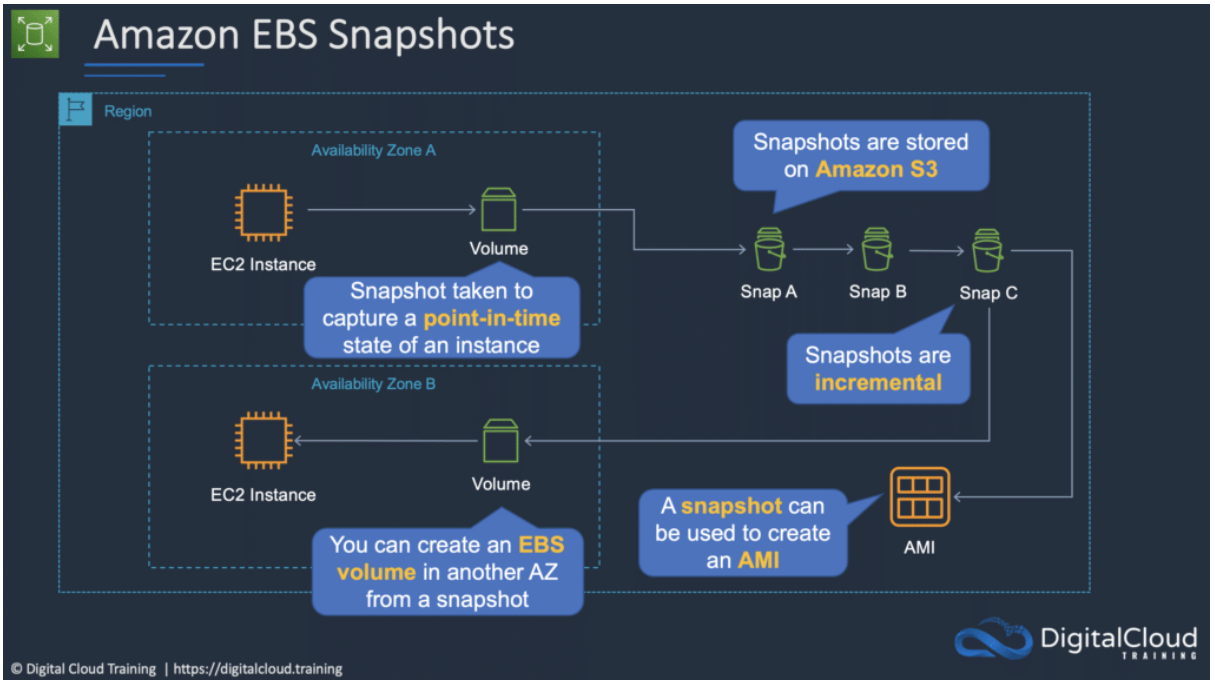
Extra non-boot volumes are not deleted on termination by default.

The behavior can be changed by altering the “DeleteOnTermination” attribute.

EBS Snapshots:

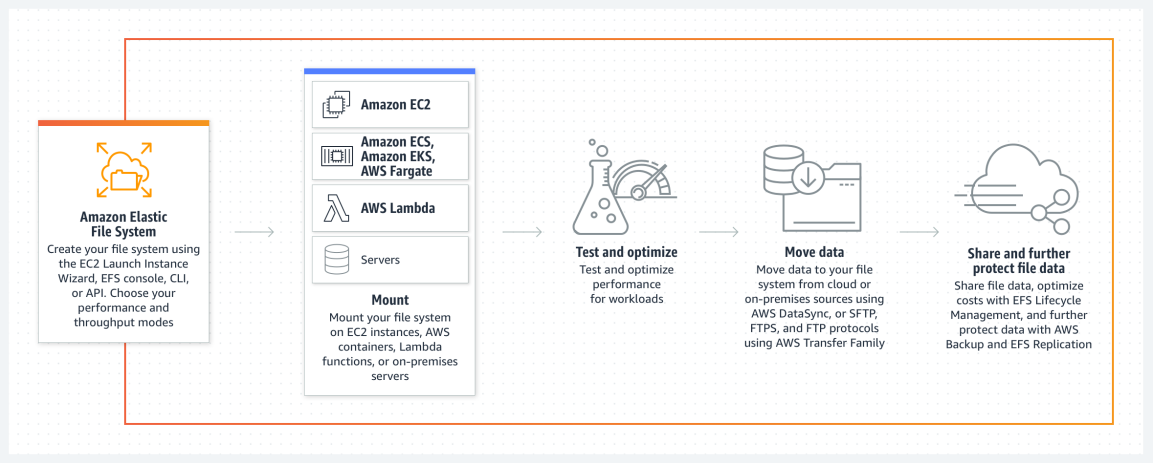
- Snapshots capture a point-in-time state of an instance.
- Snapshots are stored on S3.
- Does not provide granular backup (not a replacement for backup software).
- If you make periodic snapshots of a volume, the snapshots are incremental, which means that only the blocks on the device that have changed after your last snapshot are saved in the new snapshot.
- Even though snapshots are saved incrementally, the snapshot deletion process is designed so that you need to retain only the most recent snapshot to restore the volume.
- Snapshots can only be accessed through the EC2 APIs.

- EBS volumes are AZ specific, but snapshots are region specific.



Amazon Elastic File System

- Amazon Elastic File System (EFS) automatically grows and shrinks as you add and remove files with no need for management or provisioning.
- EFS is a managed network file system that is easy to set up right from the amazon console or CLI.
- When you have multiple EC2 instances needed to access the same file system EFS helps in this situation.
- Unlike EBS, EFS is built using the NFS4.x protocol on SSDs and have a much faster throughput. This also means that EFS is much more expensive than EBS as it can be used on very large analytical workloads.
- EFS scales up or down based on the size of the files you store and is also accessible from multiple availability zones.



Use cases

- Simplify DevOps
- Modernize application development
- Accelerate data science
- Enhance content management systems

AWS Storage Gateway

AWS Storage Gateway is a hybrid cloud storage service that gives you on-premises access to virtually unlimited cloud storage.

Customers use Storage Gateway to simplify storage management and reduce costs for key hybrid cloud storage use cases.

AWS Backup

The AWS Backup service offers a centralized backup console, APIs, and command line interface for managing backups across AWS services, including:

- Amazon Simple Storage Service (S3)
- Amazon Elastic Block Store (EBS)
- Amazon FSx
- Amazon Elastic Compute Cloud (EC2)
- Amazon Relational Database Service (RDS)
- Amazon DynamoDB
- Amazon Elastic File System (EFS)
- AWS Storage Gateway
- ... and more

AWS Snowball

With AWS Snowball (Snowball), you can transfer hundreds of terabytes or petabytes of data between your on-premises data centers and Amazon Simple Storage Service (Amazon S3).

Uses a secure storage device for physical transportation.

AWS Snowball Client is software that is installed on a local computer and is used to identify, compress, encrypt, and transfer data.

Uses 256-bit encryption (managed with the AWS KMS) and tamper-resistant enclosures with TPM.

The table below describes the AWS Snow offerings at a high-level:

Service	What it Is
AWS Snowball	Bulk data transfer, edge storage, and edge compute
AWS Snowmobile	A literal shipping container full of storage (up to 100PB) and a truck to transport it
AWS Snowcone	The smallest device in the range that is best suited for outside the data center

Snowball can import to S3 or export from S3.

Import/export is when you send your own disks into AWS – this is being deprecated in favor of Snowball.

Snowball must be ordered from and returned to the same region.

To speed up data transfer it is recommended to run simultaneous instances of the AWS Snowball Client in multiple terminals and transfer small files as batches.