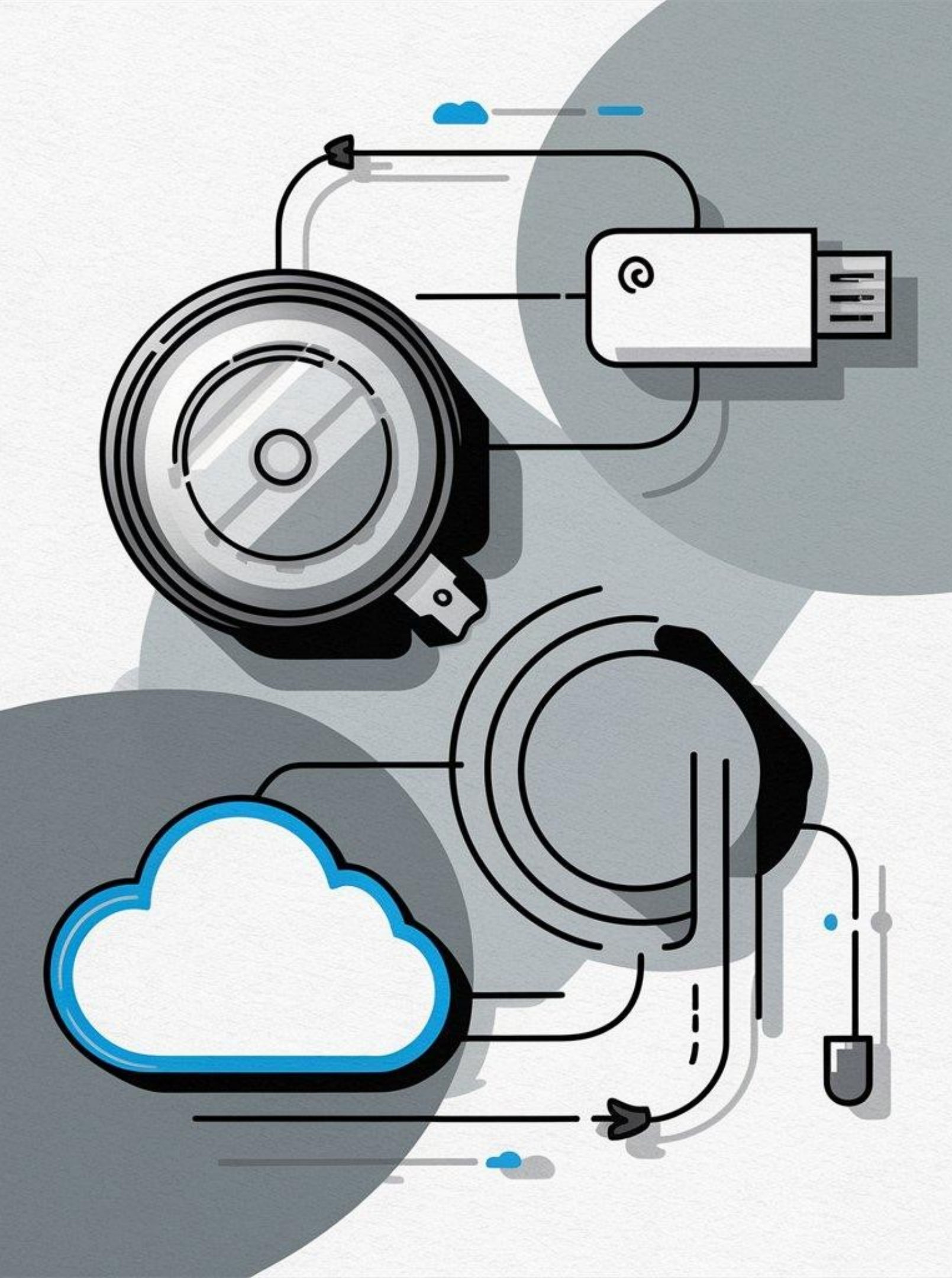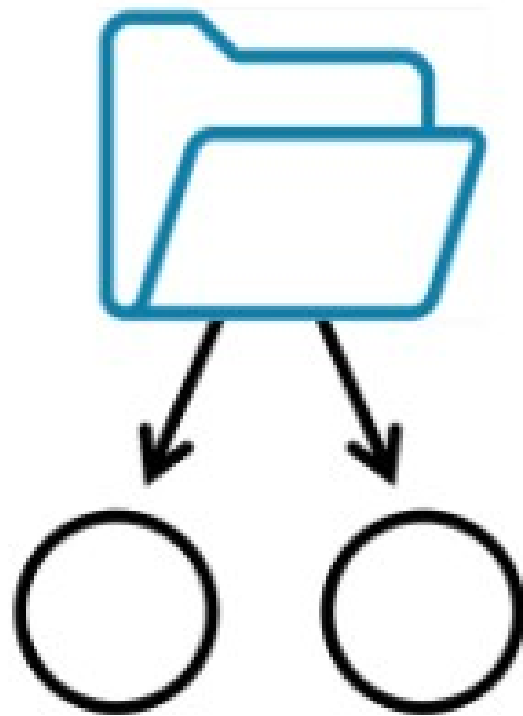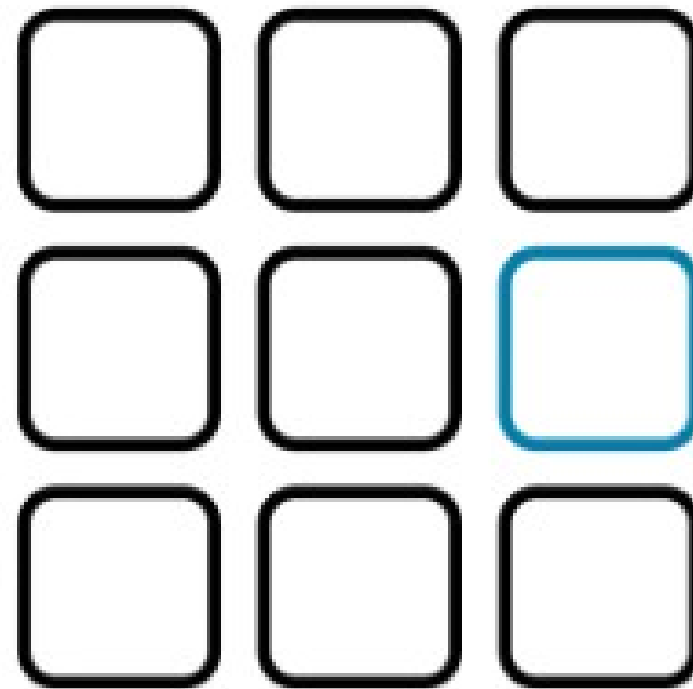# Storage

- **What is Storage?:** A place where data is saved, such as hard drives, USBs, or the cloud.
- **Types of Storage:** Includes physical storage (e.g., disks) and digital storage (e.g., cloud services).
- **Importance of Storage:** Helps keep information safe, organized, and easy to access when needed.

AWS storage services are grouped into three categories: **File storage, Block storage** and **Object storage.**
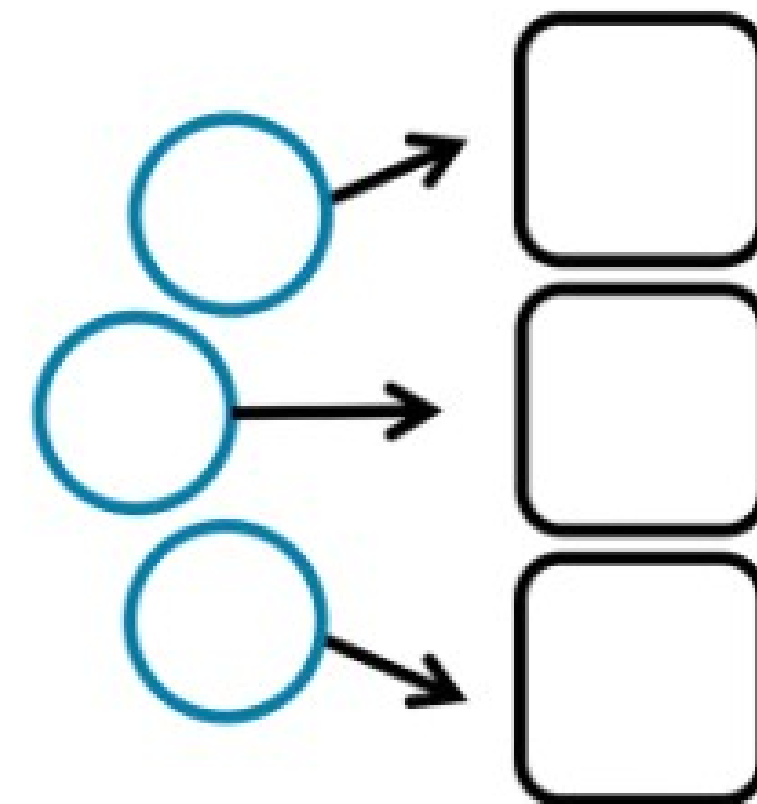
File storage

Block storage

Object storage

File Storage

- Organizes data in a tree-like hierarchy of folders and subfolders.

- Commonly used in systems like Windows File Explorer or macOS Finder.

- Ideal for structured data management, such as organizing application files.

- Example: A folder named Cat photos inside Application files for better organization.

File Storage with Amazon EFS

- Scales automatically as files are added or removed.

- Allows multiple compute instances to access with consistent performance.

- Easy setup, pay only for storage used.

- **Storage Classes:**

- **Standard/Standard-IA**: Multi-AZ, high durability.

- **One Zone/One Zone-IA**: Single-AZ, cost-effective.

File Storage with Amazon FSx

- Fully managed, reliable, secure, and scalable file system service.

- Supports four file systems: Lustre, NetApp ONTAP, OpenZFS, and Windows File Server.

- Choose file systems based on workload needs, features, performance, or familiarity.

- Provides high-performance cloud file systems tailored for specific use cases.

Block Storage
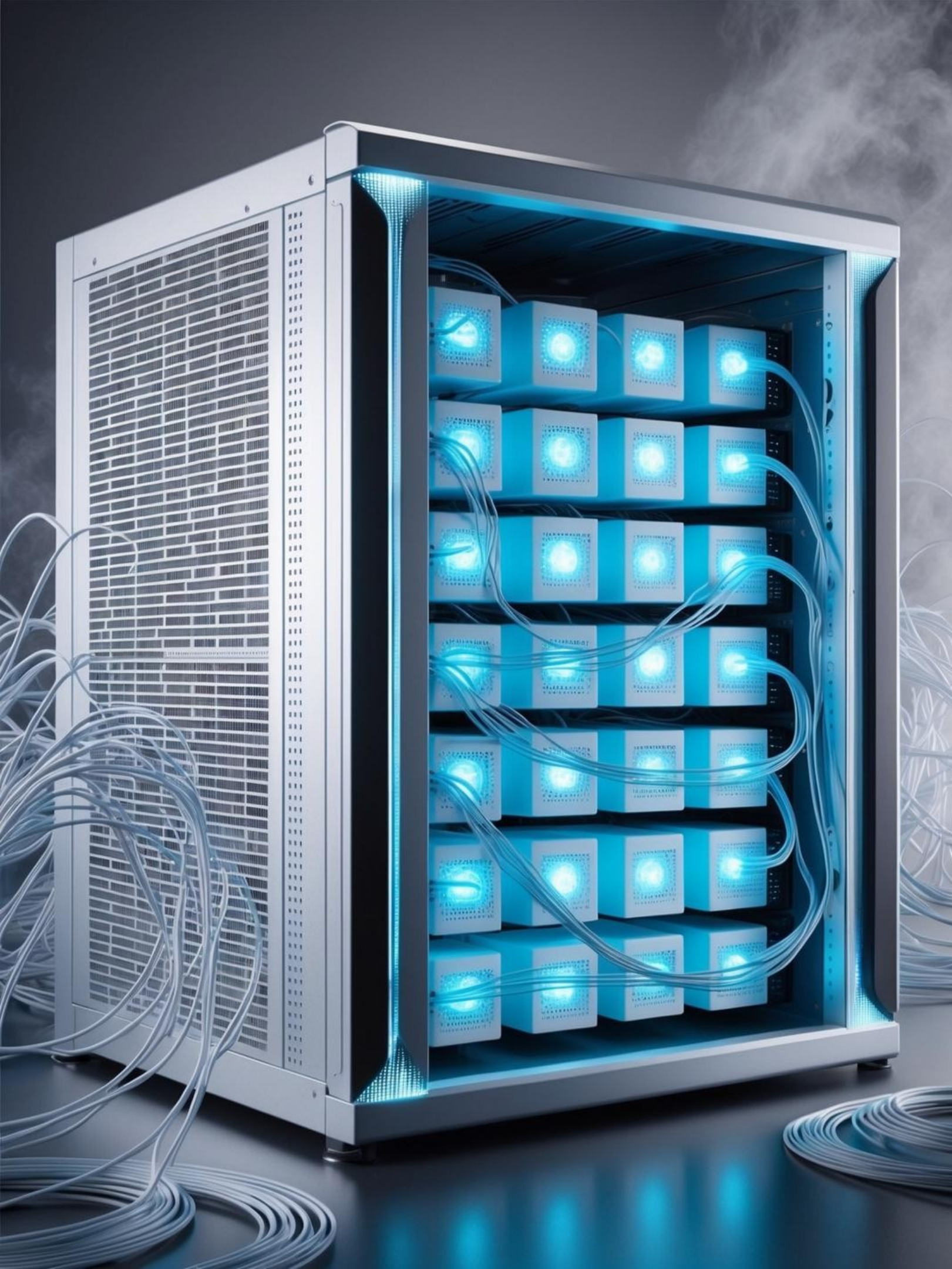
- Splits files into fixed-size chunks called blocks, each with its own address.

- Enables efficient retrieval of data by addressing individual blocks directly.

- Reassembles blocks in the correct order to form complete files for requests.

- Example: A database storing and retrieving data quickly for high-performance applications.

Block Storage with Amazon EC2 Instance Store

Amazon EC2 instance

Instance store

- Temporary block storage tied to EC2 instance lifecycle; deleted when the instance is terminated.

- Ideal for temporary data like caches, buffers, or scratch data.

- Suitable for distributed applications like Hadoop clusters needing fast, local storage.

Block Storage with Amazon EBS


Amazon EBS (Elastic Block Store)

- Block-level storage attached to EC2 instances, similar to external drives for laptops.

- Detachable: Move volumes between EC2 instances in the same Availability Zone.

- Distinct: Data remains intact even if the EC2 instance fails.

- Size-limited: Volumes have a fixed size limit, similar to external drives.

- 1-to-1 Connection: Most volumes can only be attached to one EC2 instance at a time.

Scaling Amazon EBS Volumes

Increase Volume Size: Expand volume size up to the maximum limit (64 TiB, depending on volume type). Example: A 5-TiB volume can be scaled to 64 TiB.
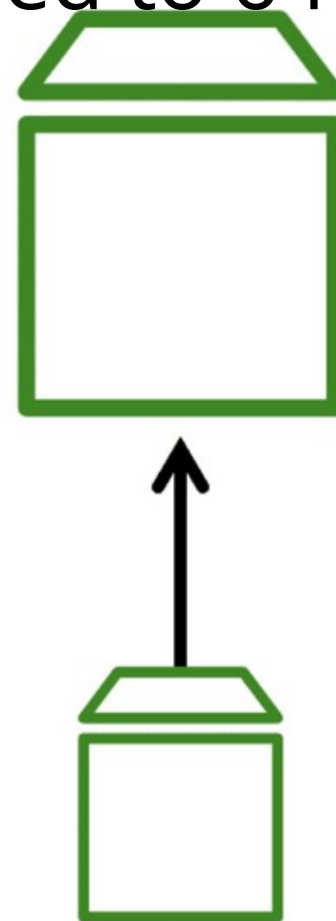
Attach Multiple Volumes: Connect multiple EBS volumes to a single EC2 instance, either during or after instance creation, to increase storage capacity.

Amazon EBS snapshots

- **Incremental Backups:** Only changes since the last snapshot are stored.
- **Redundant Storage:** Snapshots are stored across multiple Availability Zones in Amazon S3.
- **Volume Creation:** Snapshots can be used to create new, identical volumes.

Block storage

Amazon EC2 instance

Amazon EC2 instance

Instance store

Snapshot

Amazon EBS

Object Storage
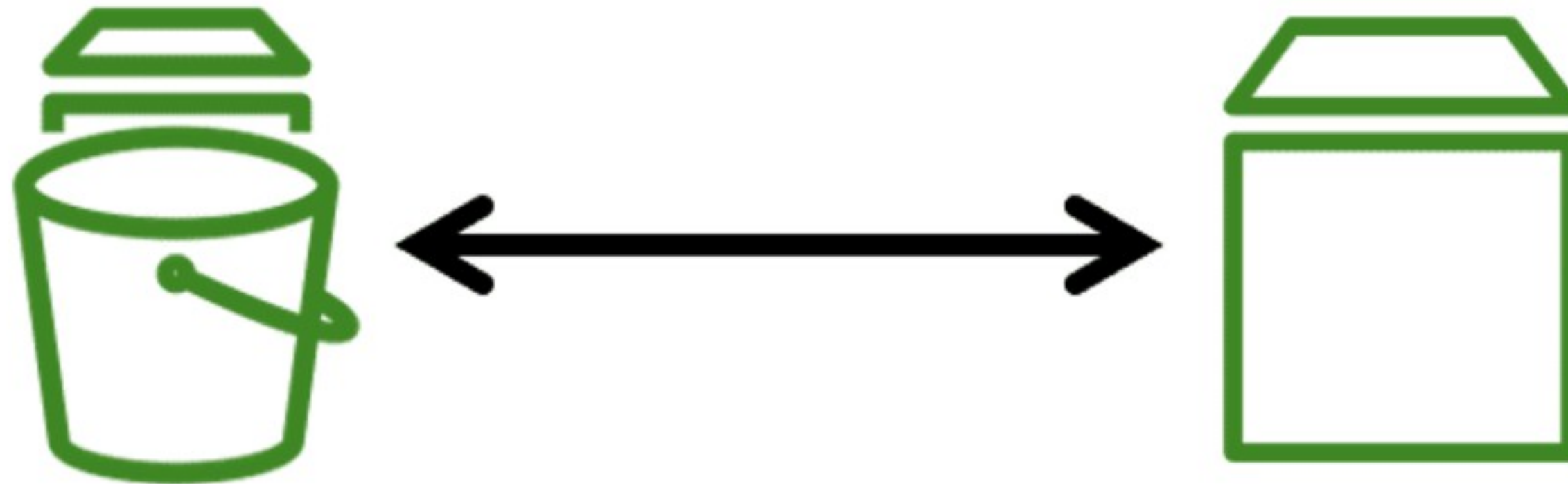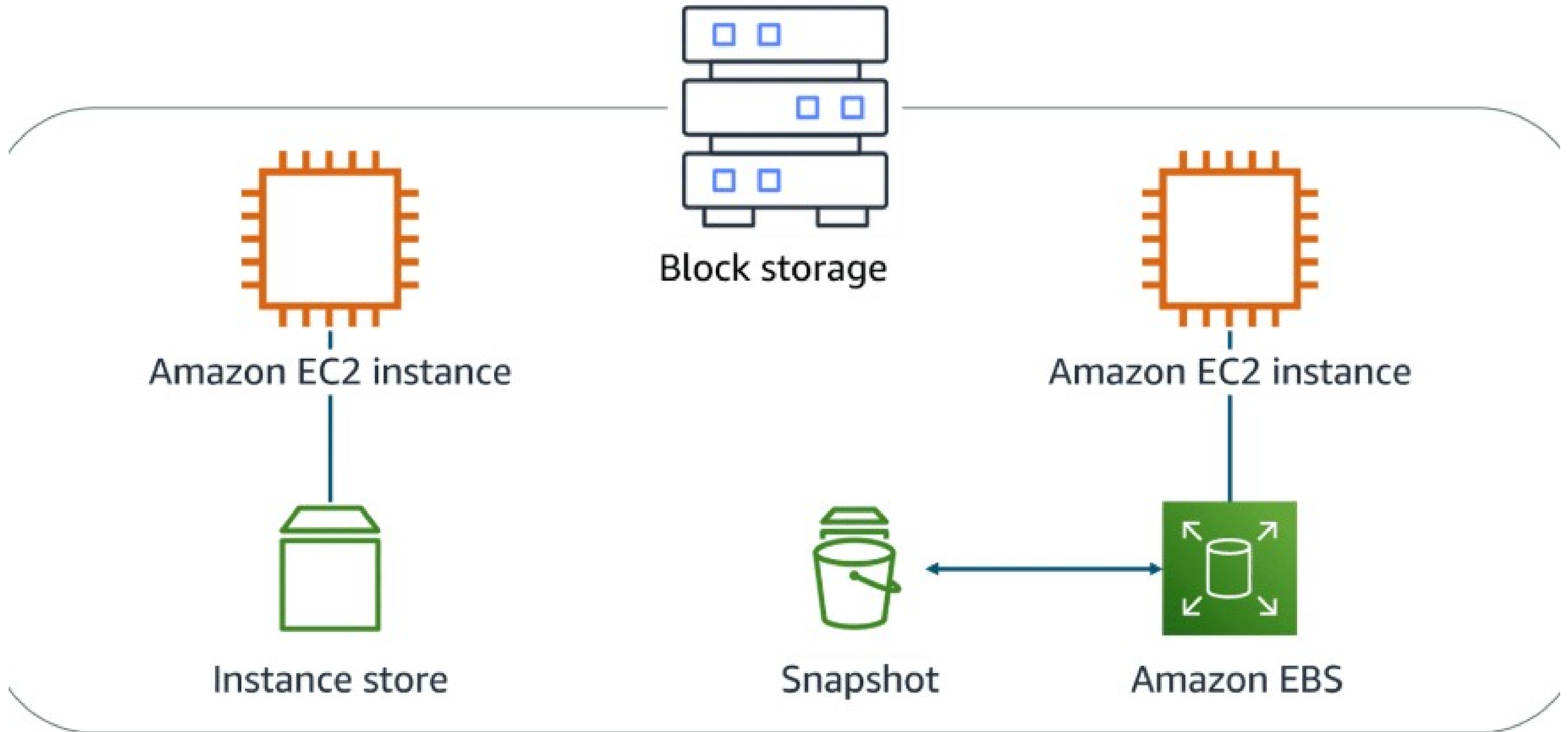
- Files are stored as distinct objects, each with a unique identifier.

- Uses a flat structure in buckets without folders or complex hierarchies.

- Bundles metadata with data for easy retrieval and organization.

- Example: Storing images and videos in Amazon S3 for scalable access.
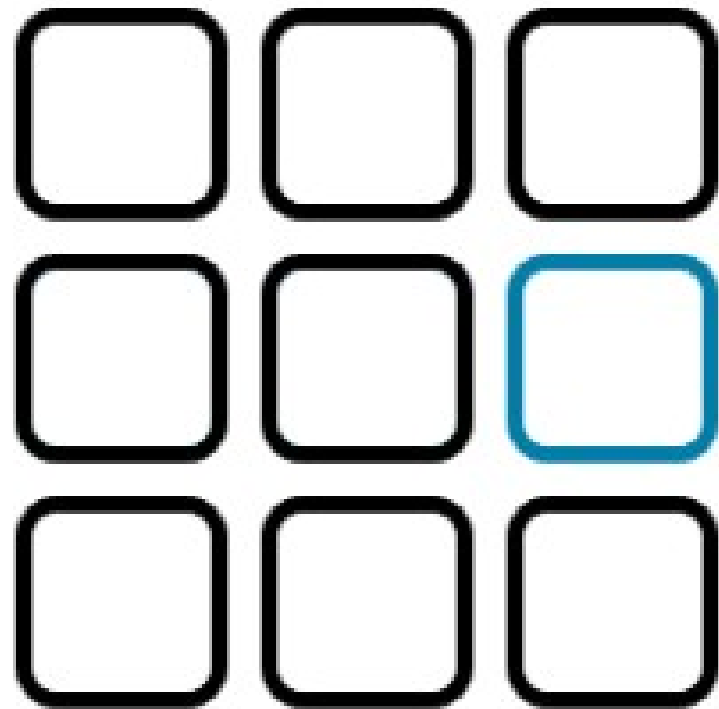
Block storage

Change one block (piece of
the file) that contains the
character
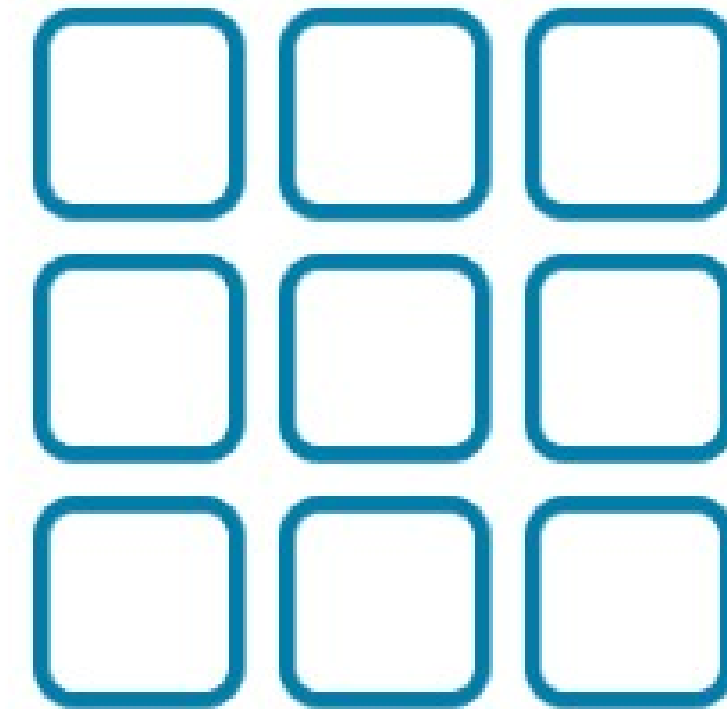
Object storage

The entire file must be updated

Changing just one character in an object is more difficult than with block storage.
When you want to change one character in an object, the entire object must be
updated.

AWS Storage

Object Storage with Amazon S3

- **Standalone Storage:** Amazon S3 is independent of compute resources and accessible from anywhere on the web.

- **Object Storage:** Data is stored as objects (files + metadata) in a flat structure.

- **Scalable:** Unlimited storage for objects, offering flexibility for diverse data needs.

Amazon S3 concepts

Amazon S3

Objects   Keys   Buckets   Regions
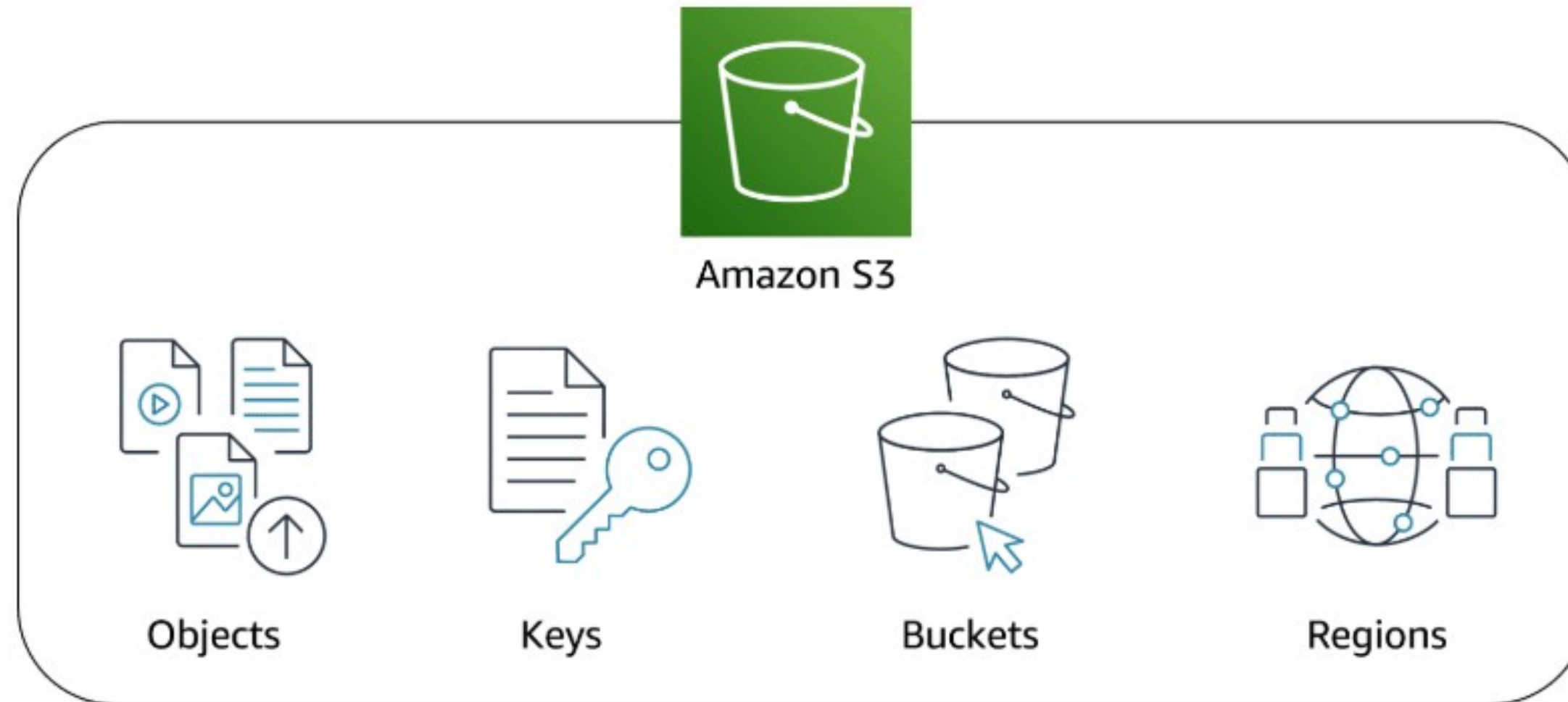
- **Objects:** Data files with metadata stored in S3.

- **Keys:** Unique identifiers for objects.

- **Buckets:** Containers for storing objects.

- **Regions:** Locations where S3 stores data.

Amazon S3 Bucket Naming Rules:

- **Uniqueness:** Bucket names must be globally unique across all AWS accounts and regions within a partition.

- **Length:** Bucket names must be between 3 and 63 characters long.

- **Valid Characters:** Only lowercase letters, numbers, dots (.), and hyphens (-) are allowed.

- **Start/End:** Must start and end with a letter or number.

- **No IP format:** Buckets cannot be named as IP addresses.

What is S3 Versioning?

A feature in Amazon S3 that keeps multiple versions of an object in the same bucket.

- **Purpose of Versioning:** Protects against accidental deletions and overwrites, allowing recovery of previous object versions.
- **Key Benefits:** Enables data recovery, provides audit trails, and supports compliance with data retention policies.

Security in Amazon S3
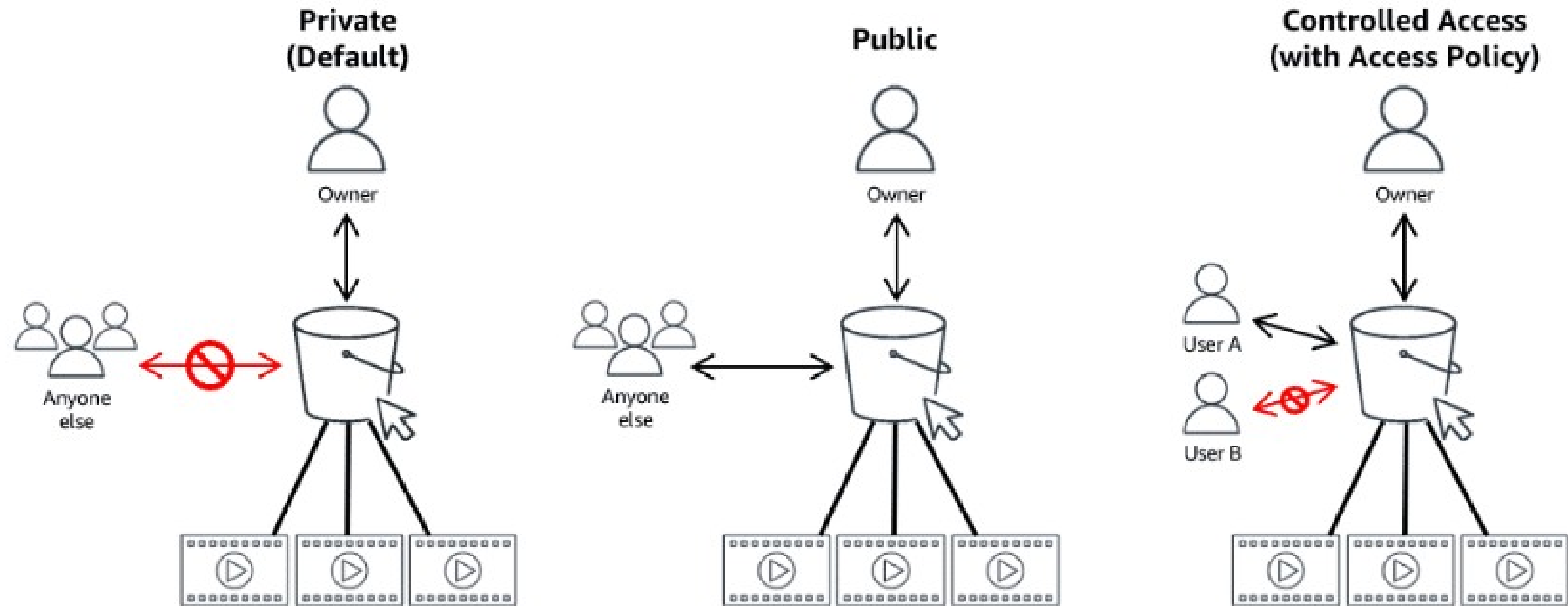
# Amazon S3 Bucket Policies:

- Written in JSON format.

- Attached to S3 buckets, not users.

- Applies to all objects in the bucket.

Example of bucket policy

```
{

   "Version":"2012-10-17",

   "Statement":[

    {

       "Sid":"PolicyForAllowUploadWithACL",

       "Effect":"Allow",

       "Principal":{"AWS":"111122223333"},

       "Action":"s3:PutObject",

       "Resource":"arn:aws:s3:::amzn-s3-demo-bucket/*",

       "Condition": {

          "StringEquals": {"s3:x-amz-acl":"bucket-owner-full-control"}

       }

    }
```

Amazon S3 Storage Classes
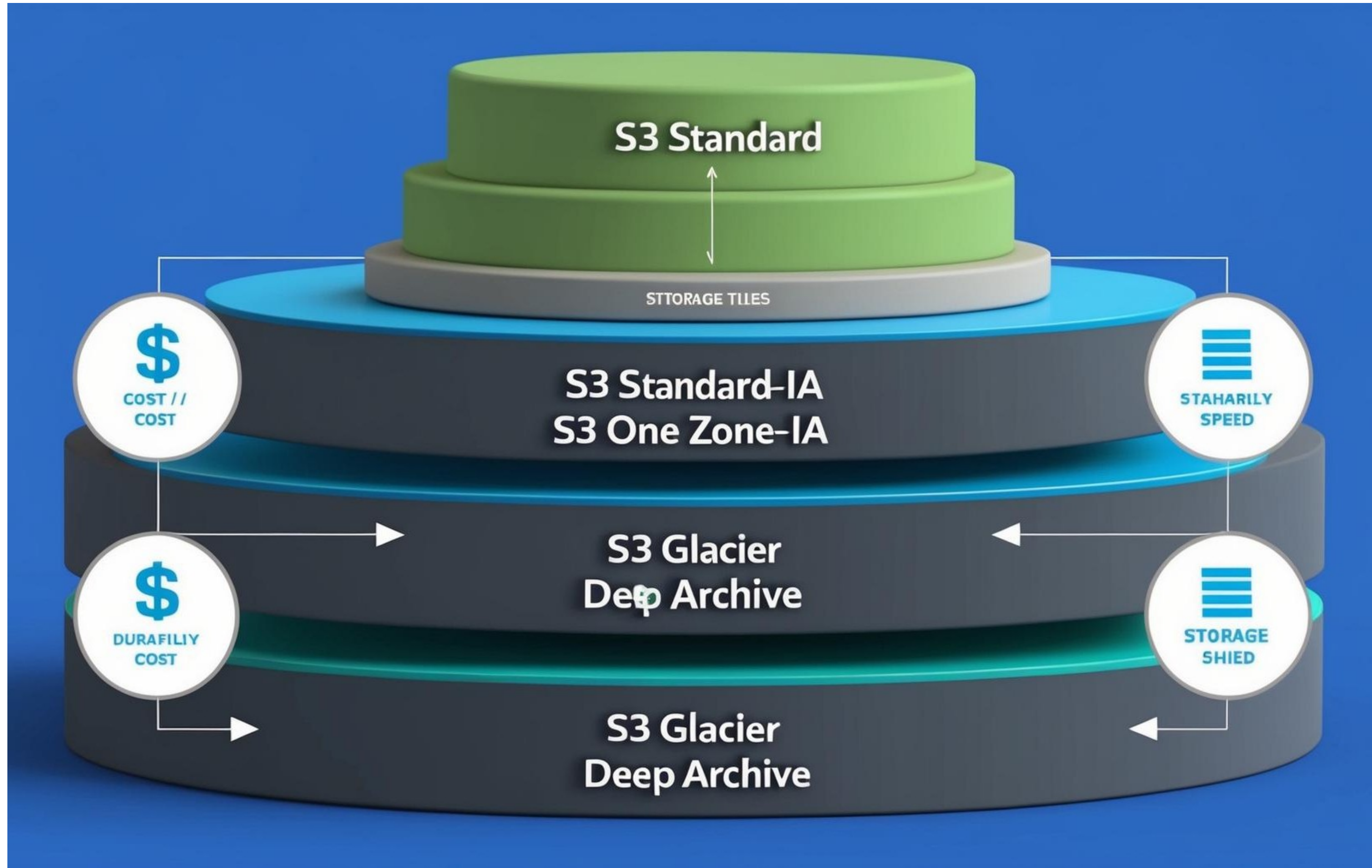
Amazon S3 Storage Classes

S3 Standard

General-purpose storage for cloud apps, websites, and analytics.

S3 Intelligent-Tiering

Automatically moves data to the most cost-effective tier based on access patterns.

S3 Standard-IA

Low-cost storage for infrequent access data, with fast retrieval.

S3 One Zone-IA

Inexpensive storage for infrequent access in a single Availability Zone.

S3 Glacier Instant Retrieval

Low-cost archival storage with millisecond retrieval.

S3 Glacier Flexible Retrieval

Archive storage with retrieval in 1-5 minutes or up to 12 hours.

S3 Glacier Deep Archive

Cheapest storage for long-term retention with 12-hour retrieval.

S3 on Outposts

On-premises storage for local data residency and performance.

Choosing the Right Storage Service

- **Amazon EC2 Instance Store:** Temporary block storage tied to EC2 instances.

- **Amazon EBS:** Persistent block storage for EC2 instances.

- **Amazon S3:** Scalable object storage for data and backups.

- **Amazon EFS:** Managed file storage for multiple EC2 instances.

- **Amazon FSx:** Managed file systems for specialized workloads.

Amazon S3 use cases

Amazon S3 is a widely used **storage service**, with far more

use cases than could fit on one screen.
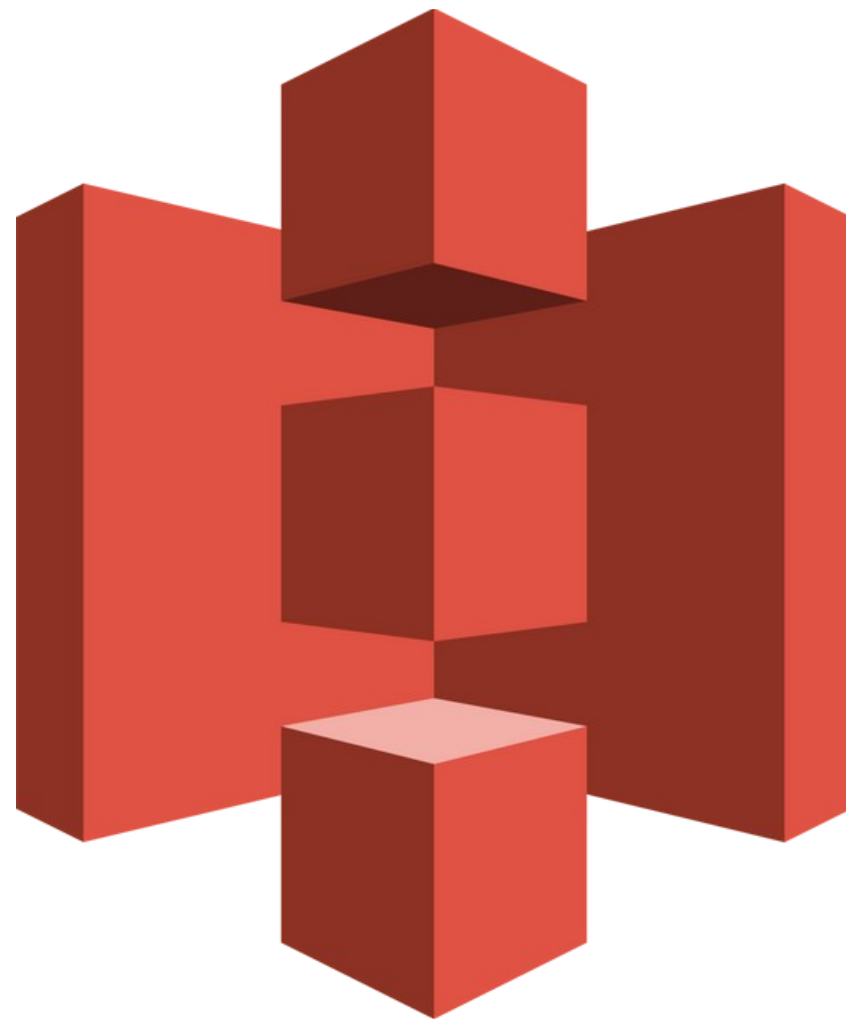
- Backup and storage
- Media hosting
- Software delivery
- Data lakes
- Static websites
- Static content