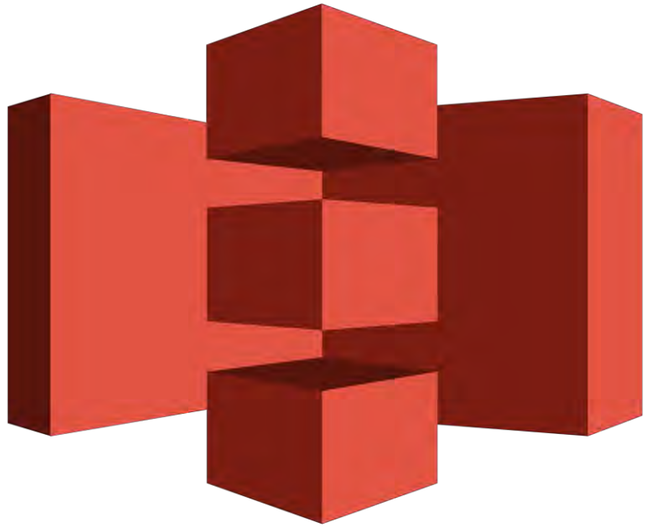
iamneo

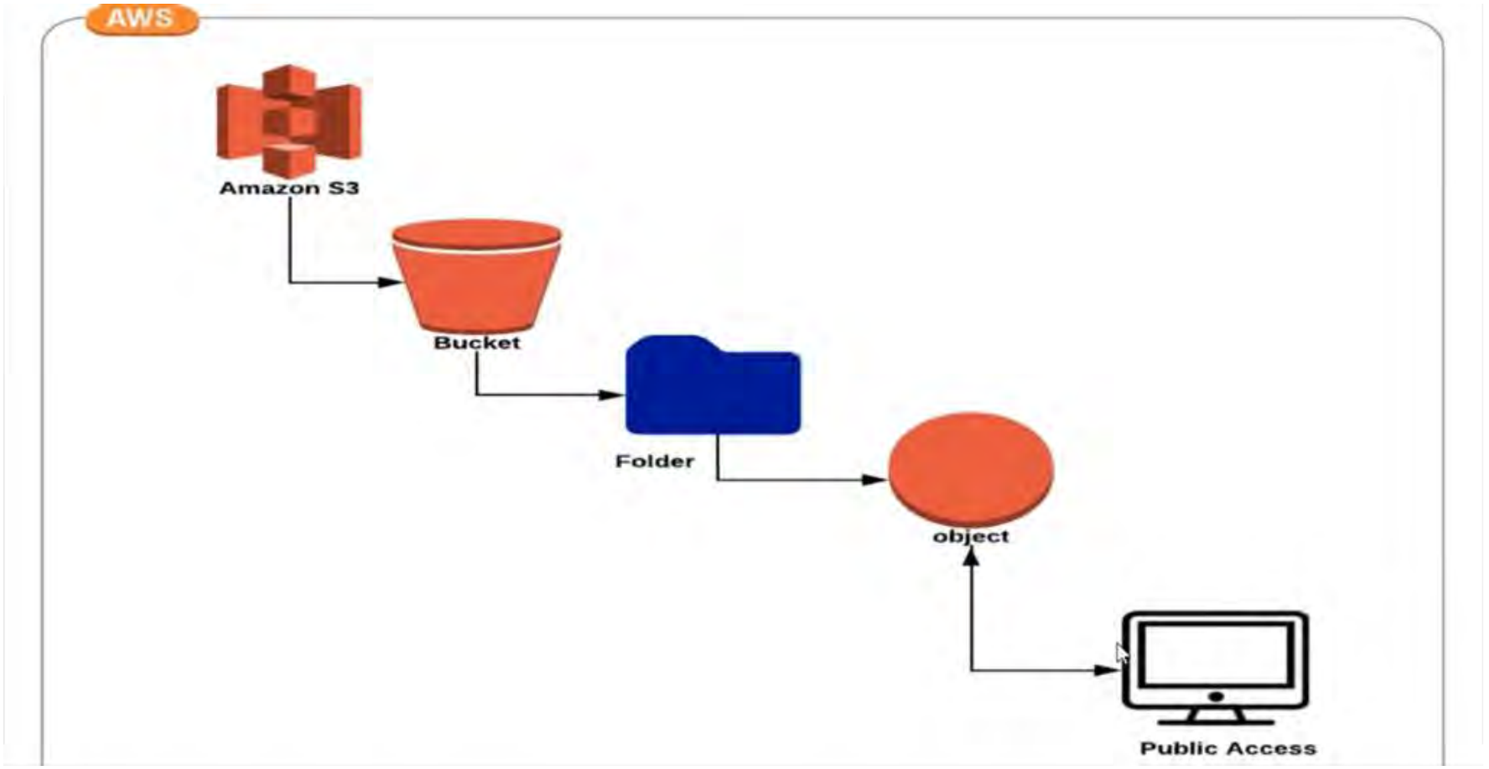
Amazon S3

Introduction to Amazon S3



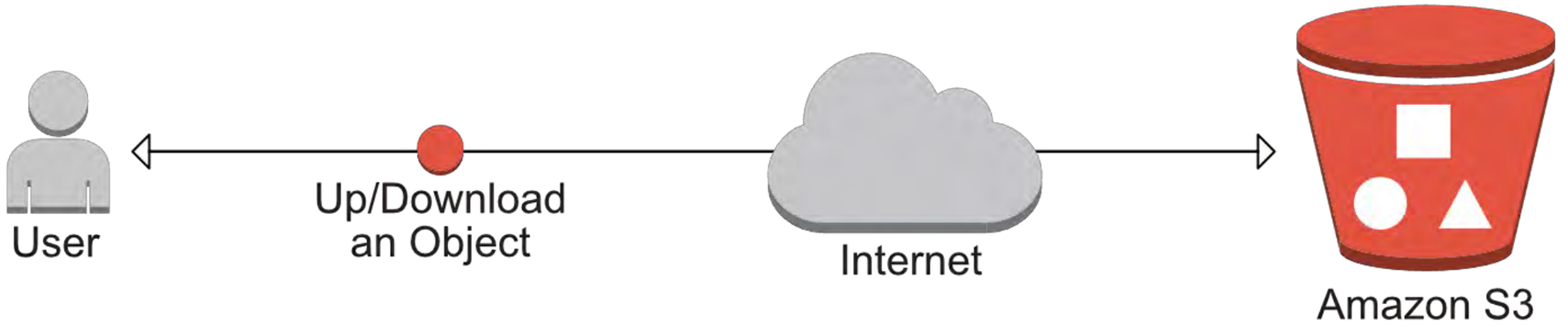
- Amazon Simple Storage Service (Amazon S3) is an object storage service offering industry-leading scalability, data availability, security, and performance.
- Customers of all sizes and industries can store and protect any amount of data for virtually any use case, such as data lakes, cloud-native applications, and mobile apps.
- With cost-effective storage classes and easy-to-use management features, you can optimize costs, organize data, and configure fine-tuned access controls to meet specific business, organizational, and compliance requirements.

S3 Bucket Concepts



S3 buckets are the containers for storing objects. Buckets are accessed using a unique Amazon Resource Name (ARN).

Object Storage



- Object storage is a technology that stores and manages data in an unstructured format called objects.
- Modern organizations create and analyze large volumes of unstructured data such as photos, videos, email, web pages, sensor data, and audio files.
- Cloud object storage systems distribute this data across multiple physical devices but allow users to access the content efficiently from a single, virtual storage repository.

Key Features

- Amazon S3 has various features you can use to organize and manage your data in ways that support specific use cases, enable cost efficiencies, enforce security, and meet compliance requirements.
- Data is stored as objects within resources called “buckets”, and a single object can be up to 5 terabytes in size.
- S3 features include capabilities to append metadata tags to objects, move and store data across the S3 Storage Classes, configure and enforce data access controls, secure data against unauthorized users, run big data analytics, monitor data at the object and bucket levels.
- Objects can be accessed through S3 Access Points or directly through the bucket hostname.

Basic Configuration of Amazon S3

1

Choose a Region

Choose the AWS Region where you want to create the S3 bucket. Regions define the geographical location where your data is stored.

2

Create an S3 Bucket

Configure the S3 bucket properties, such as bucket name and storage class, and set up Bucket Policies to define who can access your data and how to access it.

3

Set Up Lifecycle Management

Set up rules to move objects to different storage classes based on their age, transition to different storage systems, or delete them when no longer needed.

Creating S3 Buckets and Setting Bucket Policies

Bucket Naming Rules

A bucket name must be unique, DNS-compliant, and follow specific naming rules, such as no uppercase letters or underscores, and between 3-63 characters long.

Bucket Properties

You can set properties such as Location, Permissions, Versioning, Logging, and Tagging.
Versioning enables you to store multiple versions of an object in the same bucket.

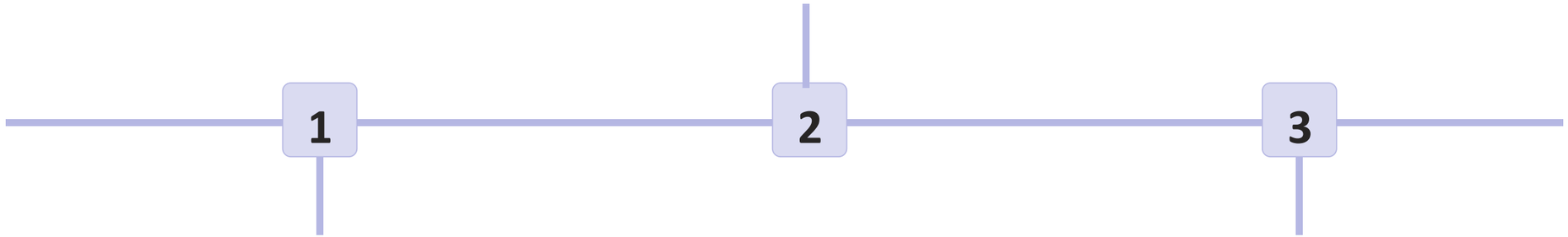
Bucket Policies

You can create policies to define who can perform specific actions on the bucket, such as listing objects or uploading new objects. You can define users, groups, or roles and use specific conditions to limit access.

Configuring S3 Object Permissions and Access Control

S3 Object Tags

You can tag objects with metadata that enables you to categorize and search for objects based on their properties. Tags can be used to manage object lifecycle, to control access permissions, or for billing and cost management purposes.



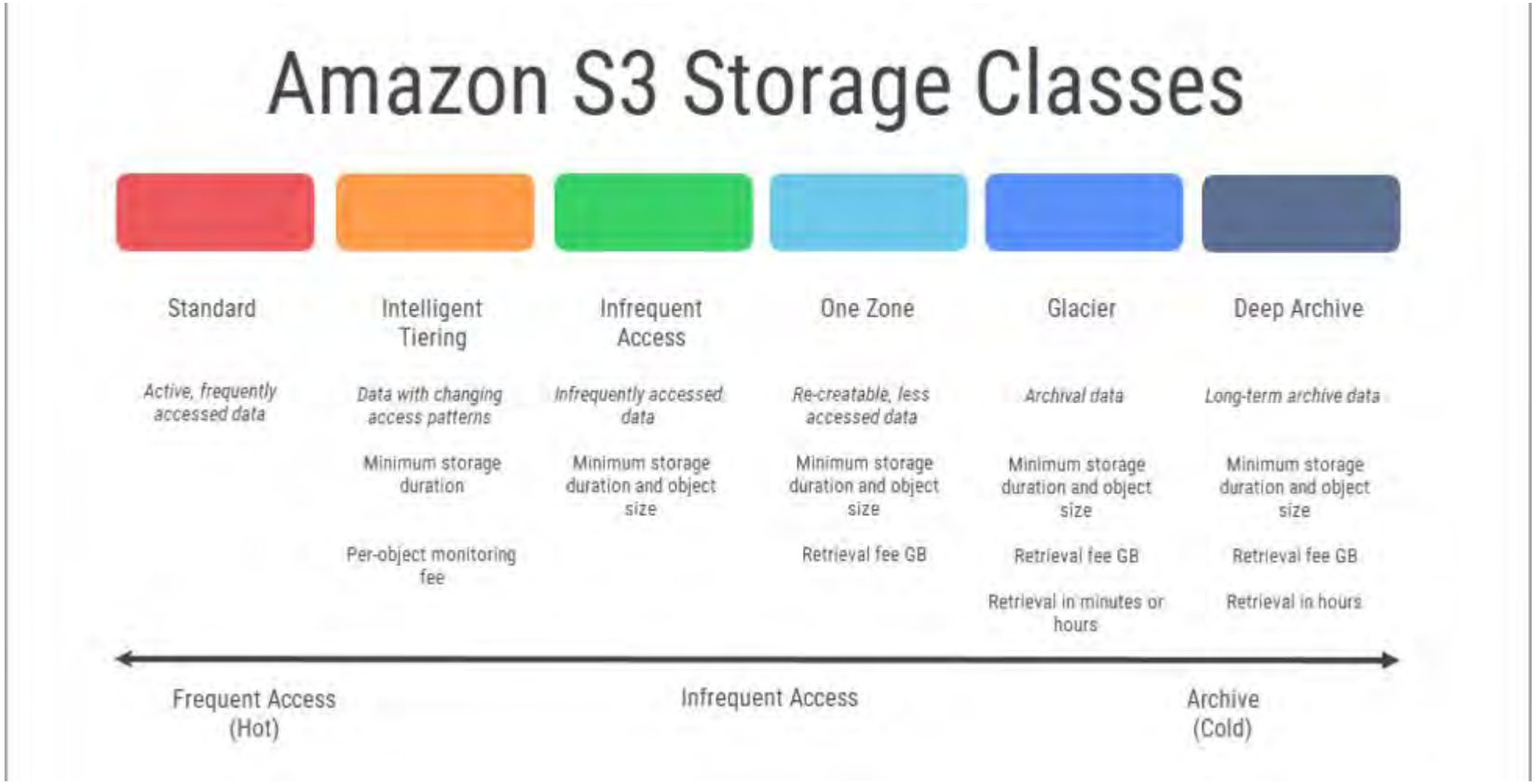
Access Control Lists (ACLs)

ACLs provide finer-grained permissions to individual objects within the bucket. You can define permissions for specific users, groups, or public access. ACLs can be modified at any time on a per-object basis.

Pre-Signed URLs

Pre-Signed URLs enable you to grant specific users temporary access to private objects in the bucket. You can set an expiration time and permissions for each URL.

Understanding S3 Storage Classes



Understanding S3 Storage Classes

Standard

The default storage class. It is designed for frequently accessed data that requires low latency and high throughput

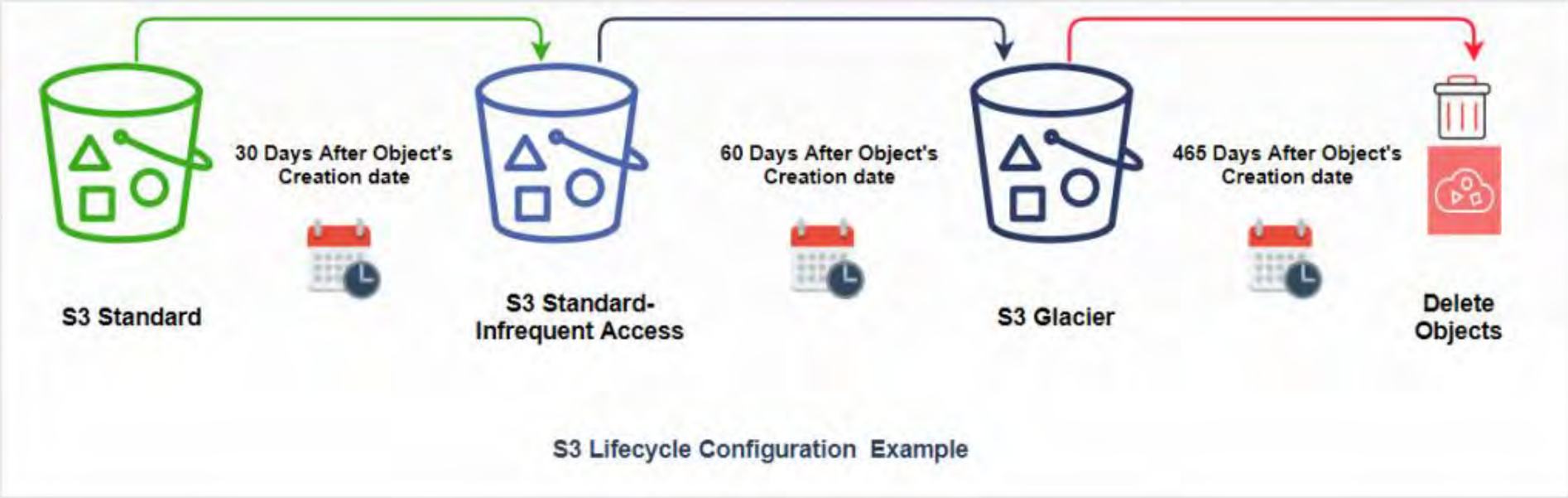
Intelligent-Tiering

Designed for data with unknown or changing access patterns. It automatically moves objects between two access tiers based on changing access patterns.

Glacier

Designed for data archiving and long-term backup. Data is stored for months or years and can take several hours to retrieve.

Managing S3 Storage Lifecycle



Lifecycle Management

Lifecycle Configuration

You can create rules that automatically transition objects to different storage tiers based on their age. Storage tiers include S3 Intelligent-Tiering, S3 Standard-Infrequent Access, and Glacier

Delete Markers

When an object is deleted, it's not immediately removed but instead marked with a delete marker. You can restore an object before the expiration period if needed.

S3 Inventory

S3 Inventory provides CSV or ORC files that list all your objects, their location, size, and metadata. You can use S3 Inventory to audit your objects, prepare for compliance audits, or analyze cost and usage trends.

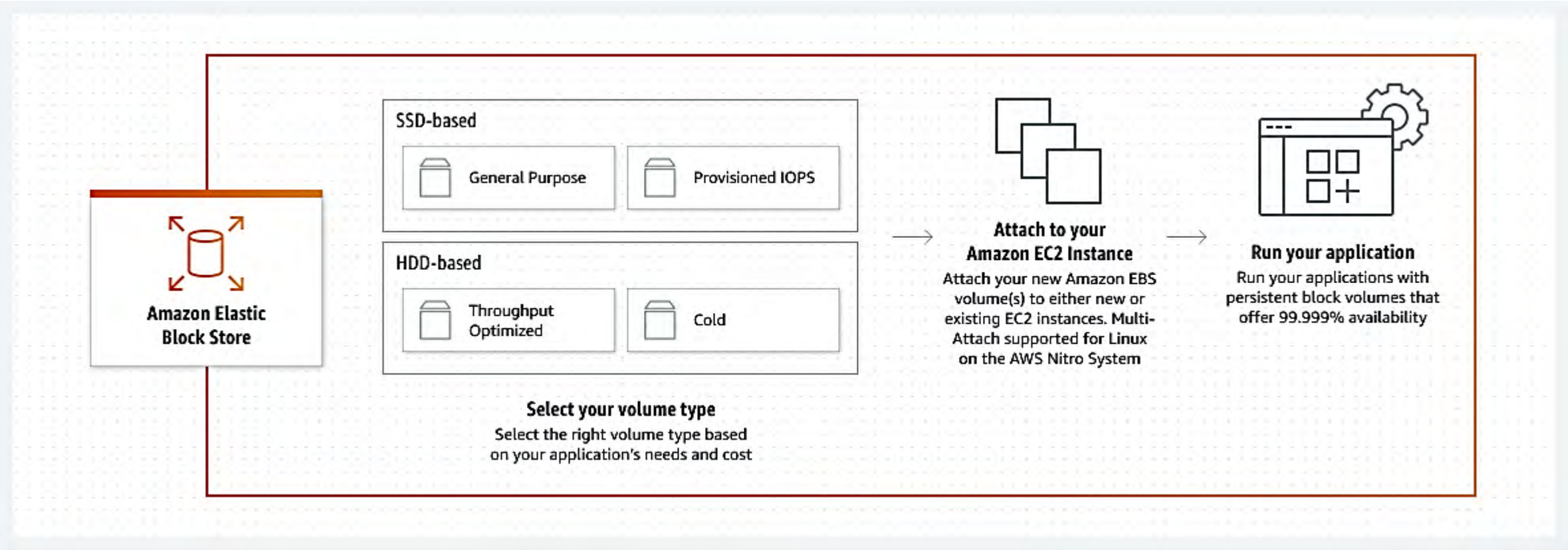
iamneo

Amazon EBS

Understanding S3 Storage Classes

- Amazon Elastic Block Store (Amazon EBS) provides block level storage volumes for use with EC2 instances. EBS volumes behave like raw, unformatted block devices.
- You can mount these volumes as devices on your instances. EBS volumes that are attached to an instance are exposed as storage volumes that persist independently from the life of the instance. You can create a file system on top of these volumes, or use them in any way you would use a block device.
- You can dynamically change the configuration of a volume attached to an instance.

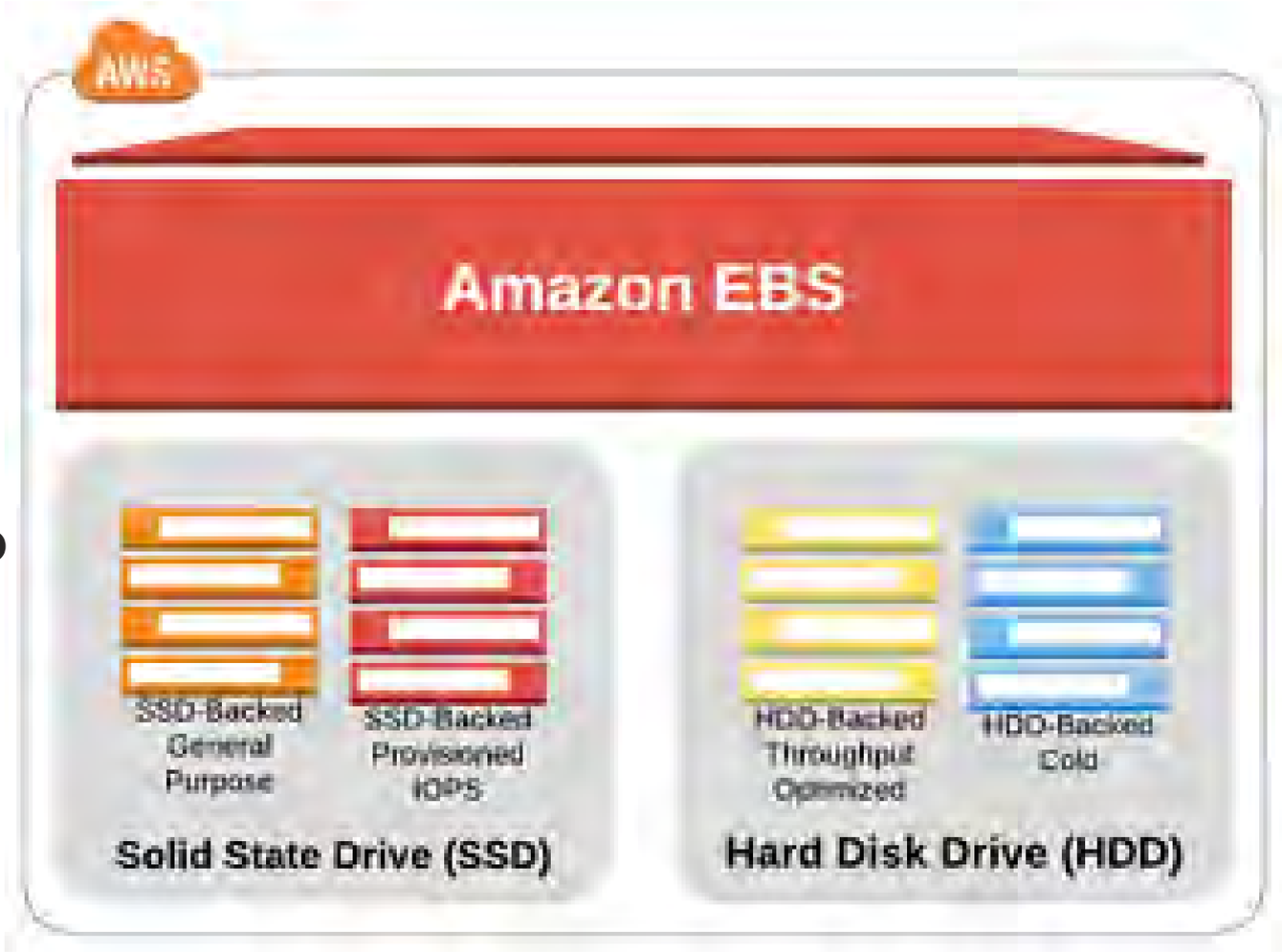
EBS Working


Amazon Elastic Block Store
SSD-based
General Purpose
Provisioned IOPS
HDD-based
Throughput Optimized
Cold
Select your volume type
Select the right volume type based on your application's needs and cost
Attach to your Amazon EC2 Instance
Attach your new Amazon EBS volume(s) to either new or existing EC2 instances. Multi-Attach supported for Linux on the AWS Nitro System
Run your application
Run your applications with persistent block volumes that offer 99.999% availability

EBS Volume Types

General Purpose SSD (gp2)

Provisioned IOPS SSD (io1)



Throughput Optimized HDD (st1)

Cold HDD (sc1)

EBS Volume Types Comparison

	General Purpose SSD		Provisioned IOPS SSD		
Volume type	gp3	gp2	io2 Block Express ‡	io2	io1
Durability	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)	99.999% durability (0.001% annual failure rate)	99.999% durability (0.001% annual failure rate)	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)
Use cases	• Low-latency interactive apps • Development and test environments		Workloads that require: • Sub-millisecond latency • Sustained IOPS performance • More than 64,000 IOPS or 1,000 MiB/s of throughput	• Workloads that require sustained IOPS performance or more than 16,000 IOPS • I/O-intensive database workloads	
Volume size	1 GiB - 16 TiB		4 GiB - 64 TiB	4 GiB - 16 TiB	
Max IOPS per volume (16 KiB I/O)	16,000		256,000	64,000 †	
Max throughput per volume	1,000 MiB/s	250 MiB/s *	4,000 MiB/s	1,000 MiB/s †	
Amazon EBS Multi-attach	Not supported		Supported		
Boot volume	Supported				

EBS Use cases

- **Block-Level Storage for EC2 Instances:**
  - •The primary use case of AWS EBS is to provide block-level storage volumes for EC2 instances.
  - •EBS volumes act as durable, persistent storage that can be attached and detached from EC2 instances as needed.
  - •It enables data storage and retrieval for applications running on EC2 instances, offering flexibility and scalability.

- **Database Storage**:
  - •EBS volumes are commonly used for database storage, including both relational databases and NoSQL databases.
  - •EBS provides consistent, low-latency storage performance, making it well-suited for database workloads that require high I/O throughput and low latency access to data.

EBS Use cases

- **Application and Web Server Storage:**
  - •EBS volumes are used as the primary storage for application and web servers hosted on EC2 instances.
  - •Application code, web content, log files, and other static or dynamic assets can be stored on EBS volumes, allowing for data persistence and easy management.

- **Disaster Recovery and Backup:**
  - •EBS snapshots enable efficient backup and disaster recovery solutions for EC2 instances and EBS volumes.
  - •Snapshots can be used to create point-in-time copies of EBS volumes, which can be stored in Amazon S3 for long-term durability and used to restore data in case of data loss or system failures.

EBS Use cases

- **Big Data Analytics:**
  - •EBS volumes are commonly used for storing and processing large datasets in big data analytics scenarios.
  - •EBS provides the required performance and capacity to handle the massive volumes of data generated by analytics workloads, enabling efficient data processing with services like Amazon EMR (Elastic MapReduce) or self-managed analytics frameworks.

- **Content Management and Media Workloads:**
  - •EBS volumes can be used to store and manage content for content management systems (CMS) or media workloads.
  - •It provides fast and reliable storage for content repositories, media files, and other assets, allowing for efficient retrieval and distribution of content to end-users.

EBS Use cases

- **Development and Testing Environments:**
  - •EBS volumes are commonly used in development and testing environments to store code repositories, development environments, and test data.
  - •EBS snapshots can be used to create consistent and reproducible copies of environments, simplifying the process of creating development and testing instances.