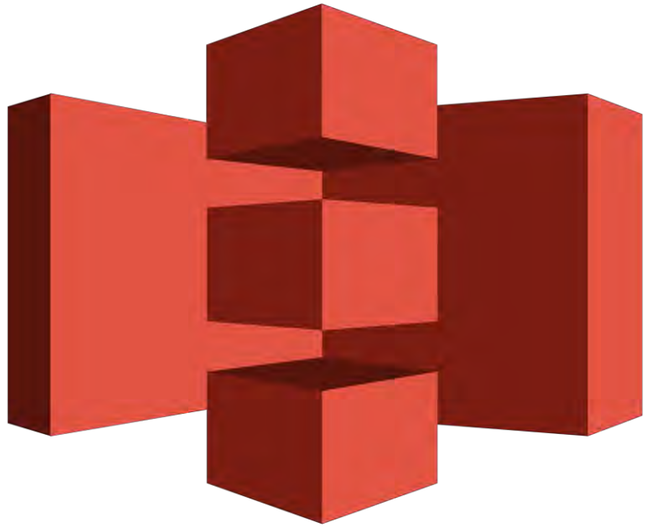
iamneo

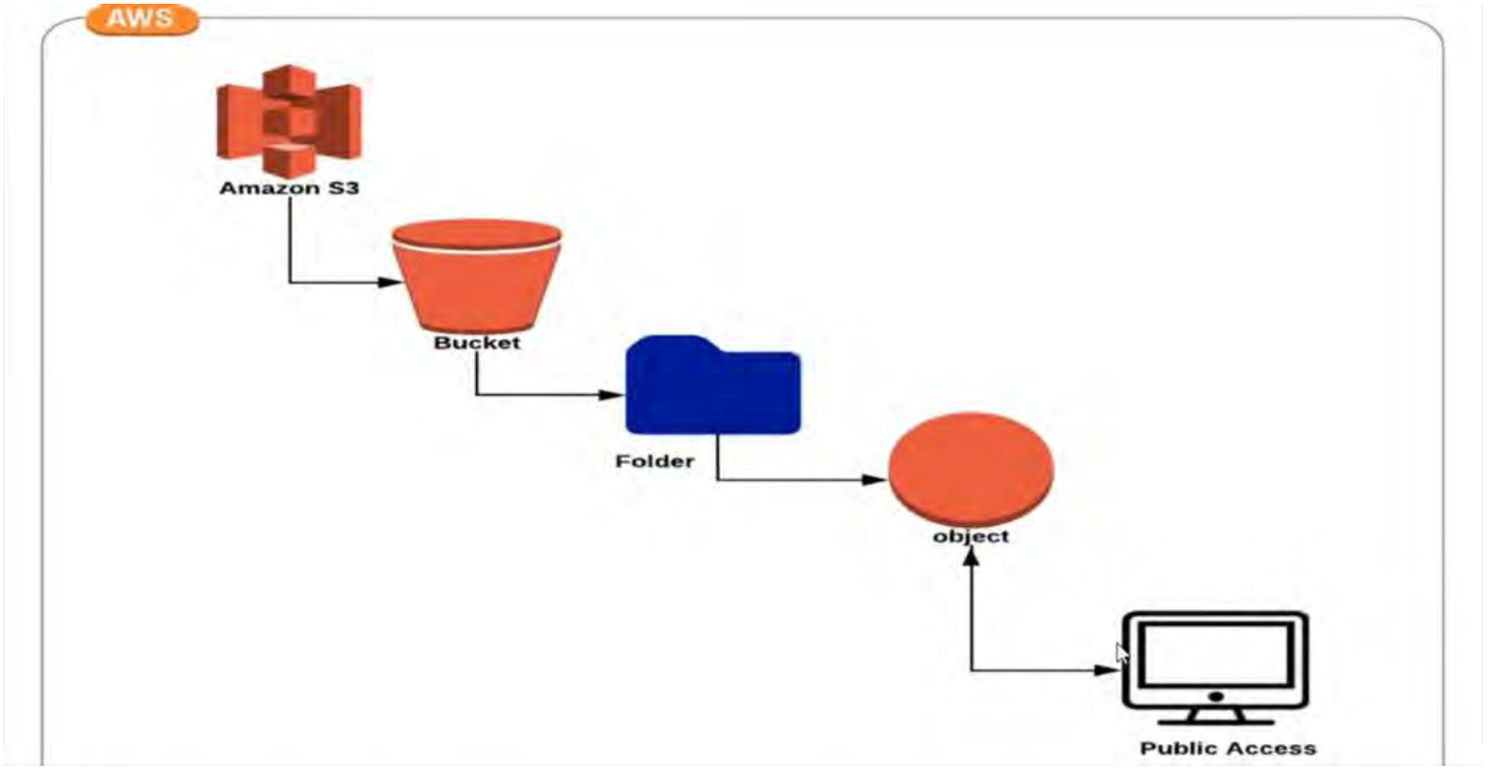
Amazon S3

Introduction to Amazon S3



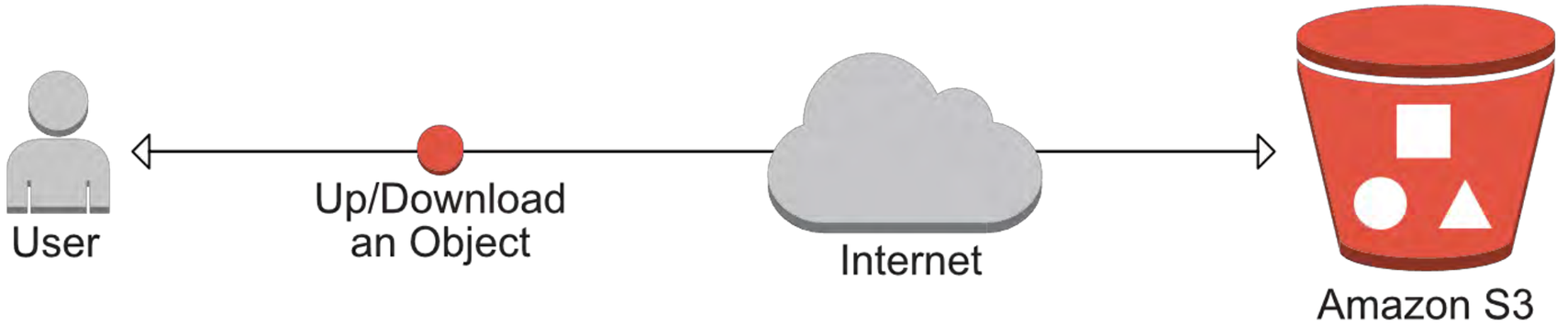
- Amazon Simple Storage Service (Amazon S3) is an object storage service offering industry-leading scalability, data availability, security, and performance.
- Customers of all sizes and industries can store and protect any amount of data for virtually any use case, such as data lakes, cloud-native applications, 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.

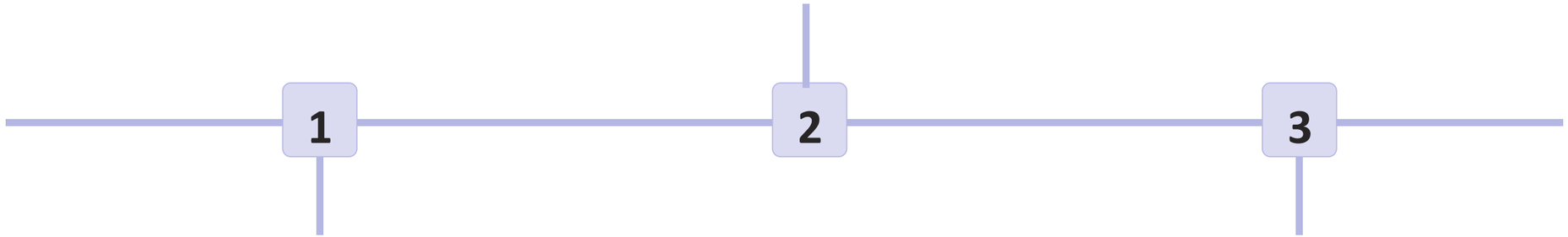
Bucket Policies

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.



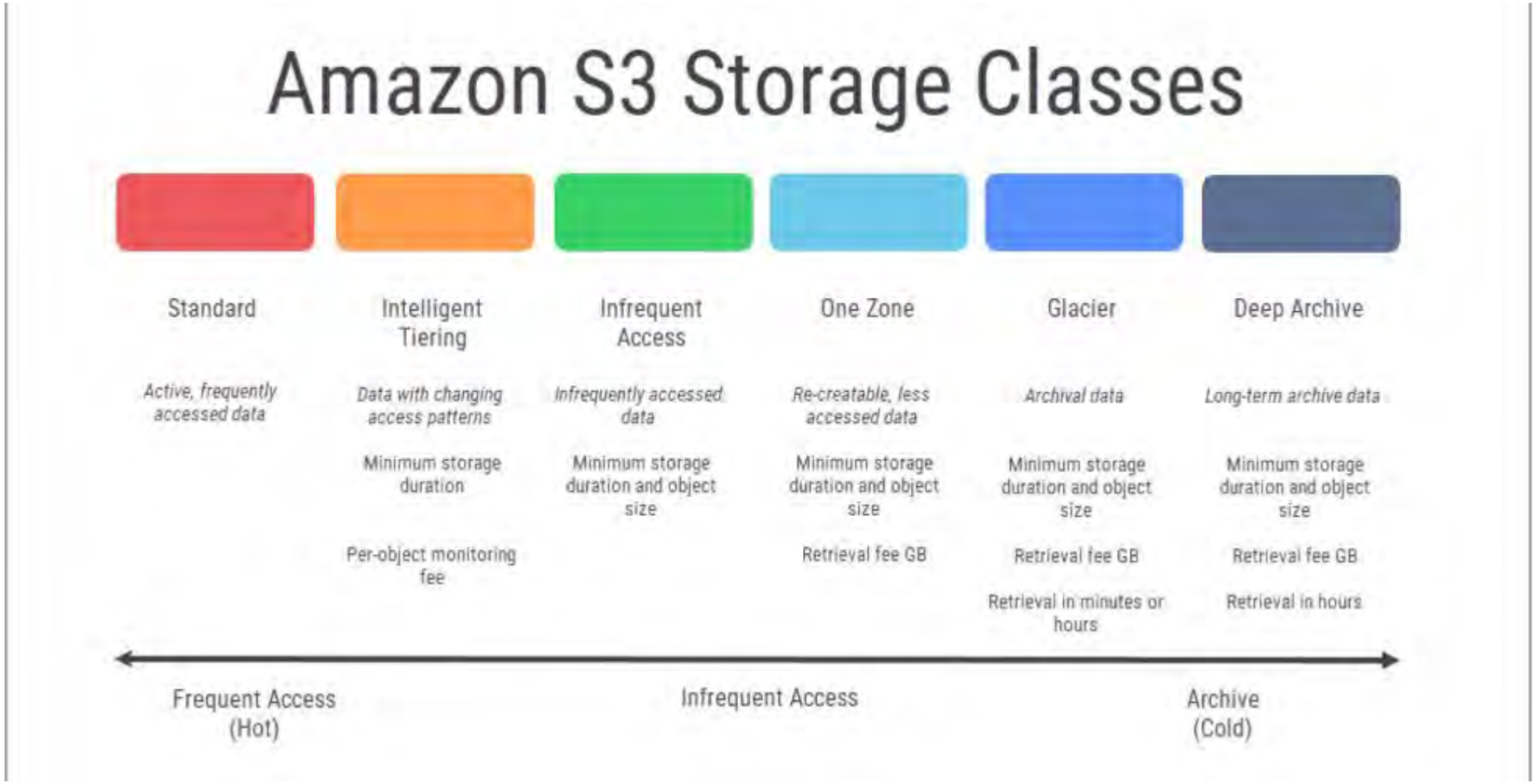
Access Control Lists (ACLs)

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

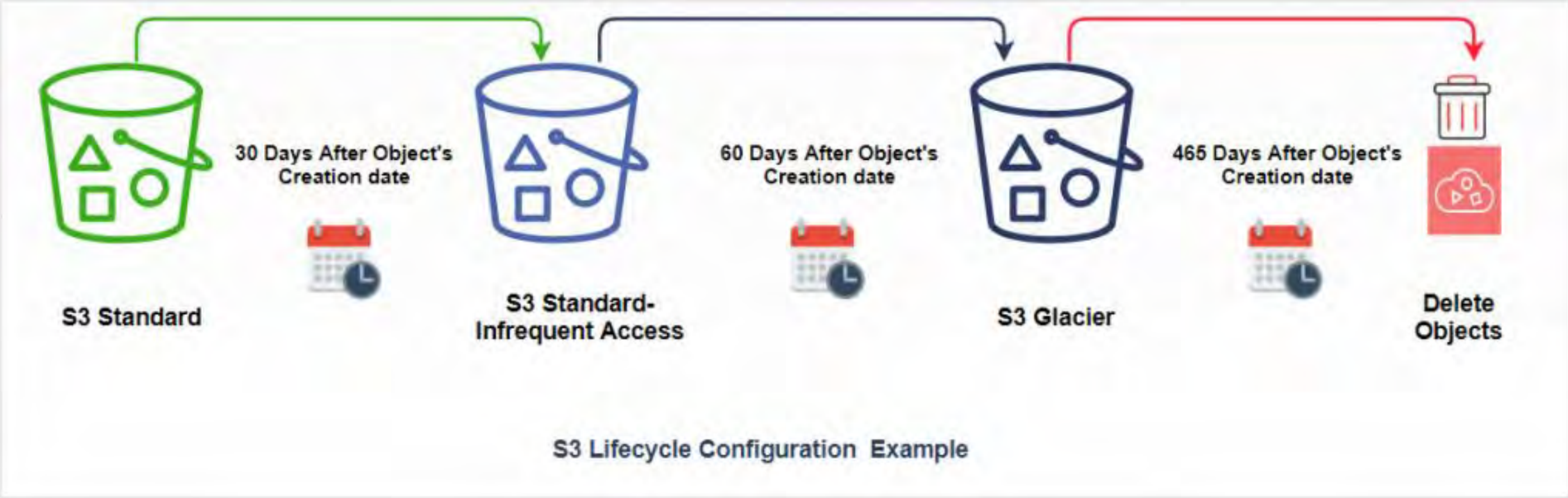
Intelligent-Tiering

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.

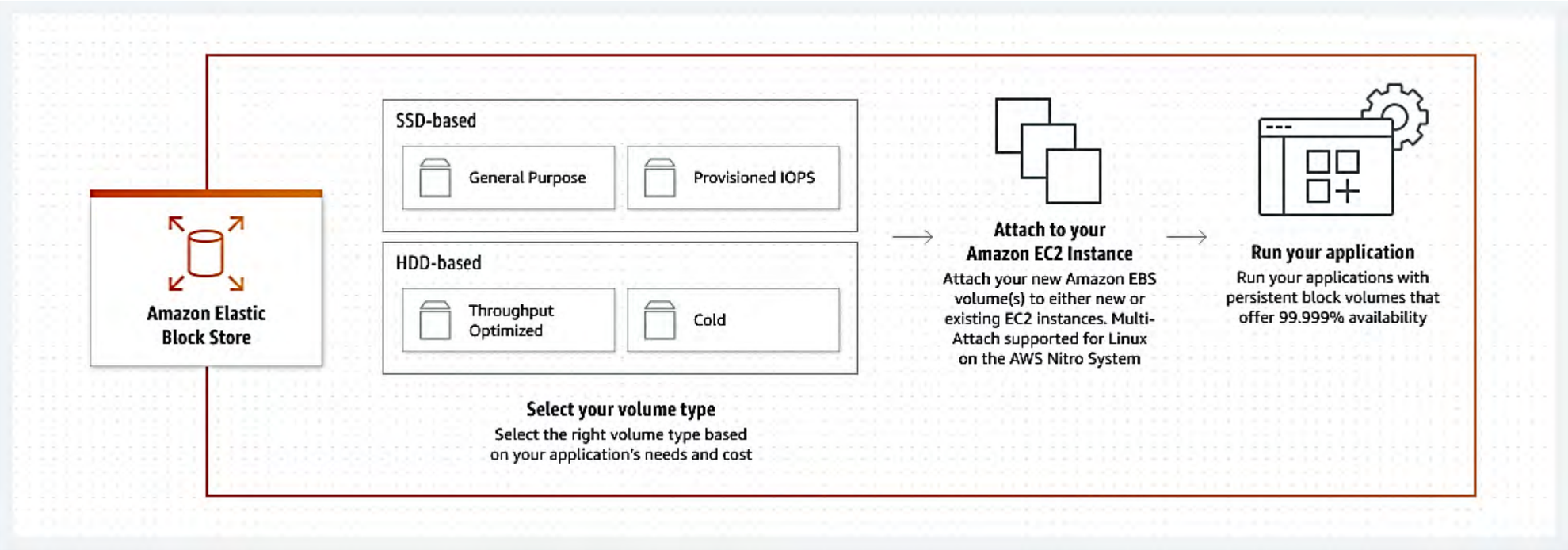
iamneo

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.

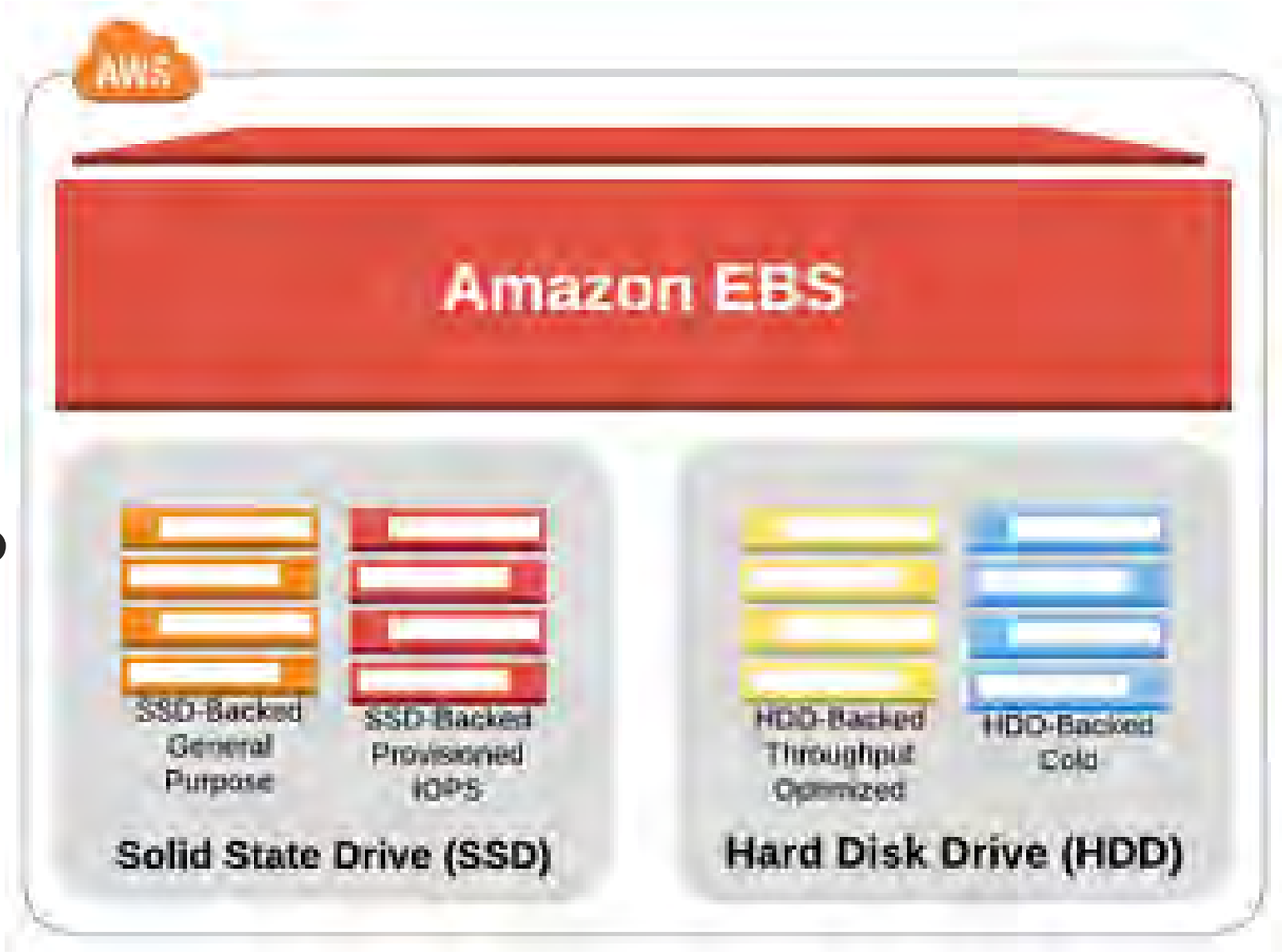
EBS Working


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

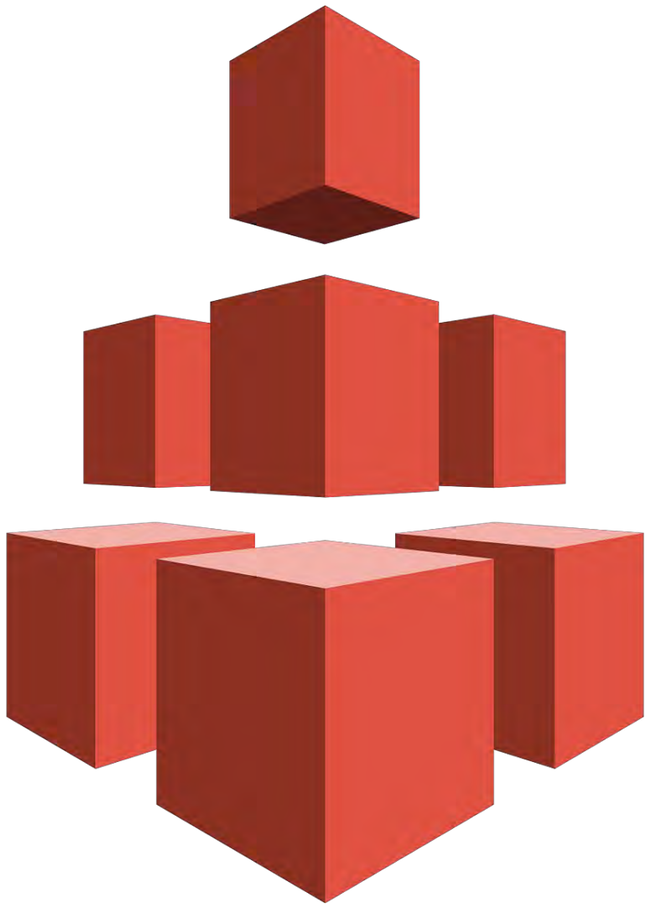
- **Development and Testing Environments:**
  - •EBS volumes are commonly used in development and testing environments to store code repositories, development environments, and test data.
  - •EBS snapshots can be used to create consistent and reproducible copies of environments, simplifying the process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

iamneo

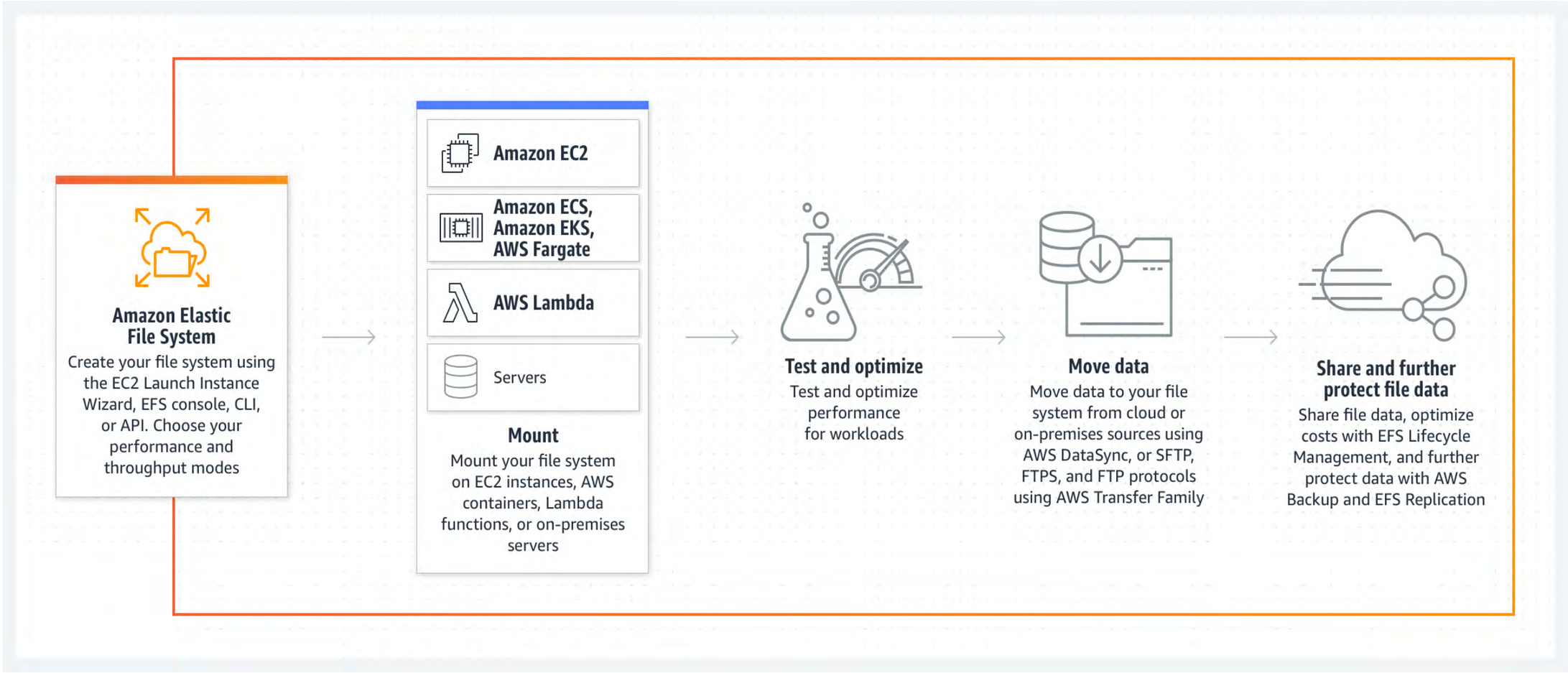
Amazon EFS

Exploring Amazon EFS

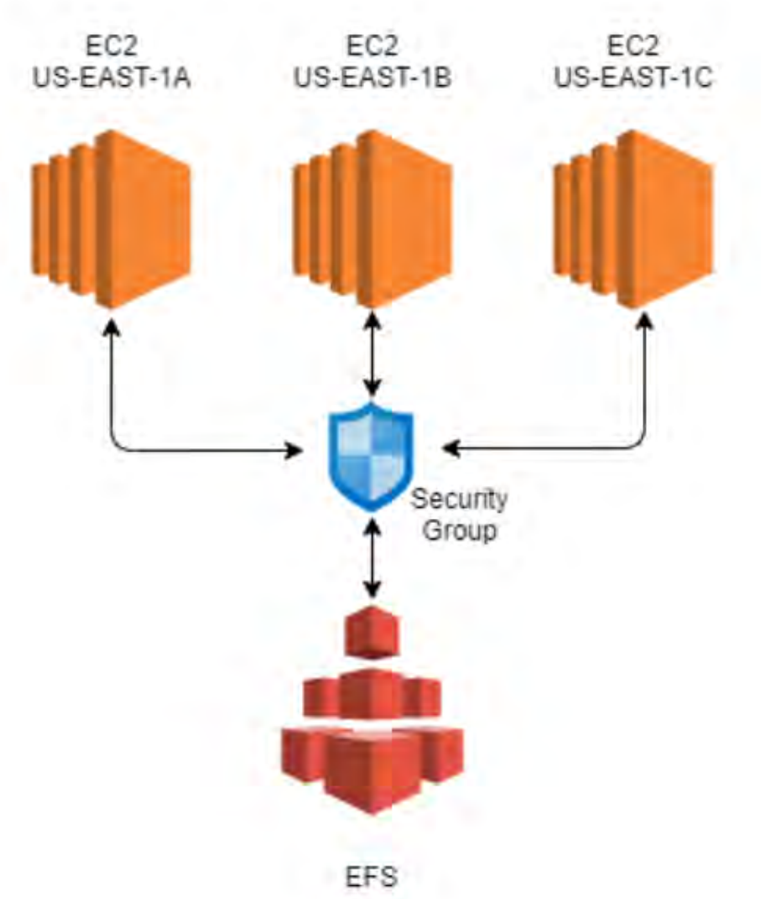
Amazon EFS (Elastic File System) is a scalable, secure, and fully-managed cloud-based file system with high availability and easy access to data from multiple instances. Let's dive deeper into its features, benefits, and potential use cases.



How it Works


Amazon Elastic File System
Create your file system using the EC2 Launch Instance Wizard, EFS console, CLI, or API. Choose your performance and throughput modes
Amazon EC2
Amazon ECS, Amazon EKS, AWS Fargate
AWS Lambda
Servers
Mount
Mount your file system on EC2 instances, AWS containers, Lambda functions, or on-premises servers
Test and optimize
Test and optimize performance for workloads
Move data
Move data to your file system from cloud or on-premises sources using AWS DataSync, or SFTP, FTPS, and FTP protocols using AWS Transfer Family
Share and further protect file data
Share file data, optimize costs with EFS Lifecycle Management, and further protect data with AWS Backup and EFS Replication

Elastic File System (EFS)



- It is a Managed NFS that can be mounted on many EC2 instances within the same region but as many Availability zones you want.
- EFS works with EC2 instances in Multi-AZ.
- Highly available, Scalable, expensive (3 times more then Gp2), but you can pay for what you use.
- Use Cases: Content Management, web serving, Word press and much more.
- Uses NFSv4.1 protocol.

Elastic File System (EFS)

- Uses security group to control access to EFS.
- Compatible with Linux Based AMI (Not Windows).
- Performance:
  - General Purpose (by default)
  - Max I/O – used when thousands of EC2 are using the EFS.
- File Sync to Sync from on-premise File system to EFS.
- Backup EFS-to-EFS (it is incremental, and you can choose frequency)
- Encryption at rest using KMS.

The Many Types of Amazon EFS

Standard

A general-purpose file system for Big Data, content repositories, web server logs, and more.

Infrequent Access

A cost-optimized file system for infrequently accessed data, with lower storage prices. Ideal for backups, logs, and workflows.

One Zone

A fully-managed file system with data stored in a single availability zone, useful for workloads that do not need multi-AZ resiliency.

Lifecycle Management

An automated feature that moves your files between different storage classes based on the frequency of access, reducing costs and optimizing performance.

Features That Make EFS Stand Out

1. Easy to Use
2. High Performance
3. Secure and Compliant
4. Scalable and Flexible

Use Cases for Amazon EFS

Big Data Analytics

EFS can store large volumes of data generated by data processing workflows, Big Data analytics, and machine learning models.

Content Repositories

EFS provides a shared file system for content creation, document management, and collaborative work, accessible by all team members and devices.

Web Applications

EFS can support highly-available and scalable web applications, allowing file sharing across elastic deployments, caching layers, and fault-tolerant architectures.

EFS Performance and Scalability

Throughput Modes

EFS supports two throughput modes which you can switch dynamically: bursting and provisioned. Bursting mode offers up to 10Gbps throughput when a file system is idle, and it applies a bursting credit approach, while provisioned mode offers a predictable throughput level, ranging from 0.001 - 1024 MB/s, based on your setup.

1

I/O Operations

EFS uses an optimized Linux file system driver to minimize write and metadata operations, and support sustained read/write performance at scale.

2

3

Multi-AZ Resiliency

EFS offers data replication across multiple Availability Zones (AZs) within a region, ensuring high durability levels and preventing data loss in case of an outage.

Managing Your Data with Amazon EFS

Backing Up Your Data

- Use Lifecycle Management to move files to S3 buckets on a regular basis for long-term storage.
- Use EFS-to-EFS backup to copy file systems across regions or accounts for higher durability and disaster recovery capabilities.

Securing Your Data

- Encrypt data both in transit and at rest using AWS KMS or your own customer-managed CMK.
- Use IAM access controls to secure your file systems and manage user permissions.

Monitoring Your Data

- Use Amazon CloudWatch metrics to monitor file system performance, I/O activities, and throughput utilization.
- Use Amazon CloudTrail logs to track file system events and changes for auditing and compliance purposes.

Pricing and Billing for Amazon EFS

1

Pay-per-Use Model

EFS charges a simple monthly fee based on usage of storage and data transfer, without any upfront fees or minimum commitments.

2

Price per GB

Prices start at $0.30 per GB and decrease based on usage tier and volume; data transfer prices start at $0.01 per GB for transfers that exceed 5% of the total GB-months of storage used per month.

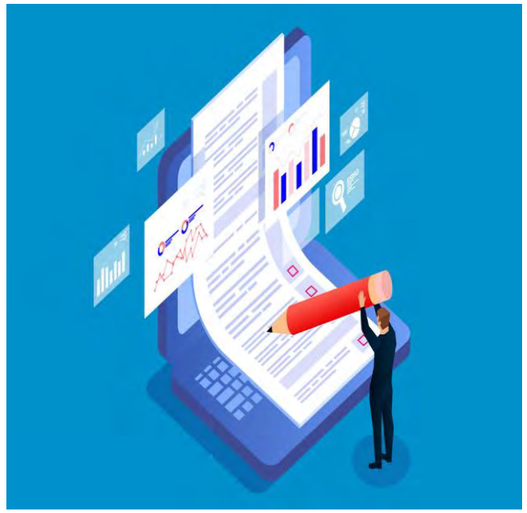
3

Additional Features

EFS offers additional features like Lifecycle Management, Backup, and Multi-AZ support at no extra cost.

Best Practices for Using Amazon EFS

Data Management



Documentation and Expertise

Security and Compliance

iamneo

S3 Glacier

Overview of S3 Glacier

Amazon S3 Glacier is a secure, durable, and cost-effective cloud storage service for data archiving and long-term backup. It is designed to deliver 99.999999999% durability and provide comprehensive data management.

Storage Classes and Pricing

S3 Glacier

- S3 Glacier is a low-cost storage service designed for data archiving. It's ideal for data that is seldom accessed but needs to be retained for a long period of time.
- However, retrieval times can range from minutes to hours, so it may not be the best choice for data that needs to be accessed frequently or quickly.
- One key feature of S3 Glacier is the ability to set up lifecycle policies, which automatically transition objects to Glacier from other S3 storage classes based on the age of the object.
- This can help reduce costs by moving less frequently accessed data to a lower-cost storage class, and by removing data that is no longer needed.

Storage Classes and Pricing

S3 Glacier Deep Archive

- S3 Glacier Deep Archive is the lowest-cost storage service available for long-term retention. It's designed for data that is rarely accessed and needs to be retained for up to several decades.
- The price per GB starts from $0.00099/month, making it one of the most affordable storage options in the market

- One important thing to note is that S3 Glacier Deep Archive has a minimum storage duration of 180 days.
- This means that if you delete an object or remove it from the service before the 180 days are up, you will still be charged for the full 180-day duration.

Retrieval Options and Costs

1

Expedited

Retrieval time of 1-5 minutes. Price starts from $0.03 per GB retrieved.

2

Standard

Retrieval time of 3-5 hours. Price starts from $0.01 per GB retrieved.

3

Bulk

Retrieval time of 5-12 hours. Price starts from $0.0025 per GB retrieved.

Configuring S3 Glacier Vaults and Lifecycle Policies

1 Creating a Glacier Vault

Use S3 console, command line tools, or AWS SDKs to create a Glacier vault.

2 Lifecycle Policies

Automate the migration of objects between Amazon S3 and S3 Glacier, and between S3 Glacier and S3 Glacier Deep Archive using lifecycle policies.

3 Notifications

Set up SNS notifications to get notified on vault and inventory events.

Overview of S3 Glacier Deep Archive

Cost-effective Archive Storage

S3 Glacier Deep Archive is the most cost-effective storage option for long-term data retention, backup, and archival.

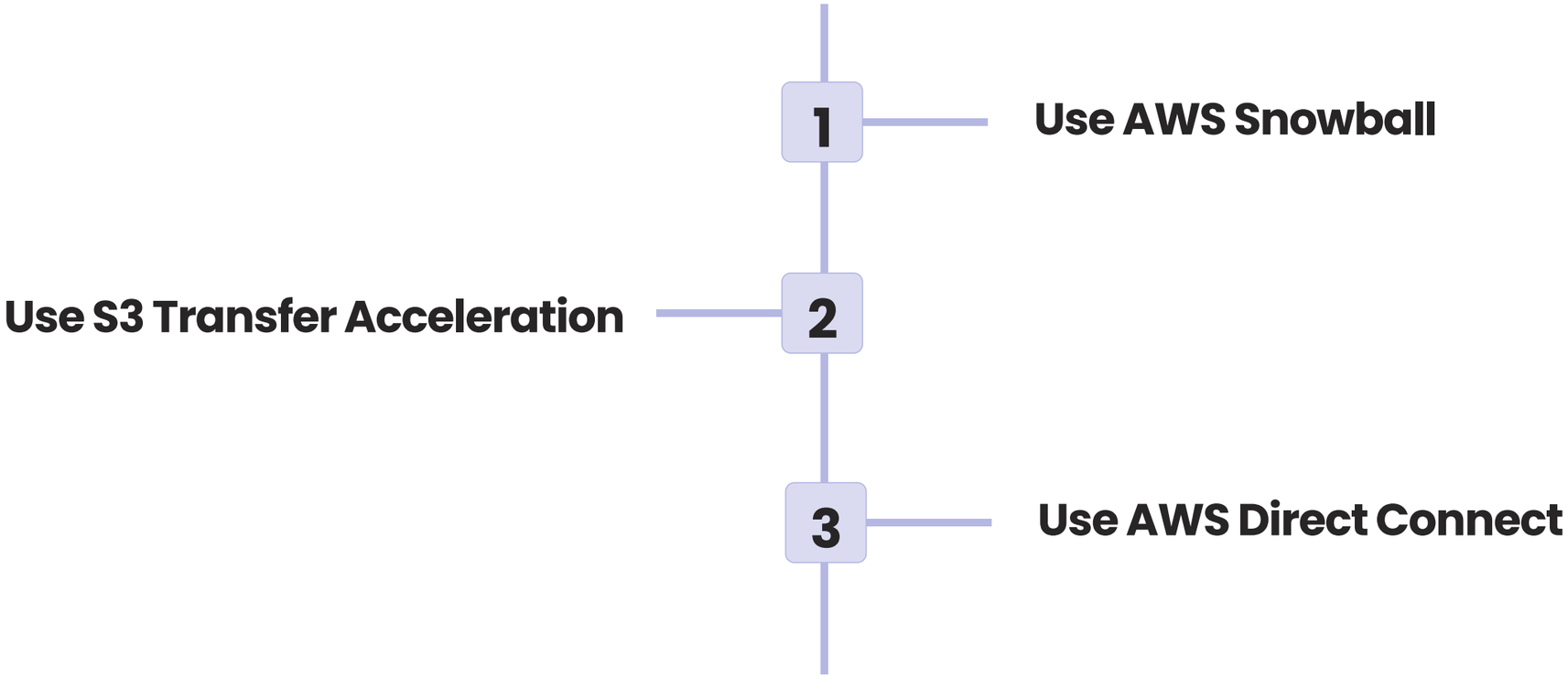
Retention Periods

S3 Glacier Deep Archive provides retention periods ranging from seven years to 30 years.

Secure Storage

S3 Glacier Deep Archive provides industry-standard AES-256 encryption for data at rest and in transit.

Data Transfer



Use Cases

1 Data Archiving

Archive and manage data that may be rarely accessed, but still needs to be available for compliance.

2 Backup and Disaster Recovery

Ensure reliable backup and recovery of critical business data in the cloud.

3 Media Archives

Preserve and manage media assets, such as movies, audio files, and original recordings, for long-term storage.

Benefits and Drawbacks

Benefits

- Cost-effective
- Secure
- Reliable
- Scalable for all business sizes.
- Provides low-cost storage for rarely accessed but important data.

Drawbacks

- Retrieving data can take a while depending on the retrieval option selected.
- Large object transfers can be challenging and can take time.

iamneo

S3 Snowball

Overview of AWS Snowball

Hardware Device

- AWS Snowball is a physical device that organizations can use to migrate large amounts of their data to AWS.
- The device comes with a built-in E Ink shipping label, making it easy to track and manage the device during shipping.

Data Migration

- Snowball enables organizations to migrate huge volumes of data to AWS with ease and confidence.
- It also provides tools to help you automate the data transfer process, saving you time and reducing the risk of error.

Accelerated Transfer

- Snowball uses multiple network connections, helping to accelerate the data transfer process.
- This means you can transfer large amounts of data to AWS in less time than other data transfer methods.

Overview of AWS Snowball

Easier Integration

- Snowball integrates seamlessly with Amazon S3 and other AWS storage services to boost data transfer ease and consistency.
- This means you can easily integrate Snowball into your existing workflows and data management systems.

Increased Security

- Snowball provides built-in security features to keep your data safe during the transfer process.
- The device is tamper-resistant and uses 256-bit encryption to secure your data.
- You also have the option to use your own encryption keys for added security.

Benefits of using AWS Snowball

1. **Fast and Reliable Data Transfer**
2. **Cost-Effective Solution**
3. **Highly Secure**

Use cases for AWS Snowball

1

Content Distribution

Snowball allows content providers to distribute video content and large files to end-users quickly and seamlessly.

2

Backup and Disaster Recovery

Snowball enables companies to back up huge amounts of data to the cloud, ensuring safety from the risk of data loss and ensuring continuity of operations in the event of a disaster.

3

Big Data Applications

Snowball provides the perfect infrastructure for big data applications by enabling organizations to quickly move large amounts of data from their data center to their AWS environment.

How to request and set up an AWS Snowball

Open AWS Management Console

Log into the AWS Management Console and choose Amazon Snowball to create a new job.

Specify Job Details

Enter details about the job and the data transfer, including pick-up and drop-off addresses, destination S3 bucket, and other details.

Connect and Use the Snowball

Snowball gets shipped to the client, and then the data gets loaded onto the Snowball with the help of a user-friendly interface designed for novice users.

How to transfer data with AWS Snowball

1

Step 1. Create an Export Job

The first step is to use the AWS Management Console to create a new export job. This job contains information about the data you want to transfer.

2

Step 2. Load Data on the Device

Once you have created your export job, it's time to load your data onto the Snowball device. This is done using a user-friendly interface that is designed for novice users.

3

Step 3. Connect the Device and Create a Job

After your data is loaded onto the Snowball, it's time to connect the device to your AWS account and create an import job. Then, plug the Snowball into a power source and wait for the data to transfer.

AWS Snowball pricing and limitations

Data Transfer Cost: $0.03/GB	**Snowball Size Limit:** 80TB starting at $300 + shipping costs
Data Export: Nothing	**Imported Data:** $15 per Snowball transfer job
Data Eras: Nothing	**Snowball Capacity:** 80-100TB depending on the device version

Conclusion and next steps

1 Cost-Effective Data Migration

With AWS Snowball, organizations can transfer significant amounts of data to AWS in a simple yet cost-effective way.

2 Flexible and Highly Secure

Snowball provides the perfect solution for transferring data, back up, and disaster recovery, as well as handling big data applications.

3 Start Your Data Transfer Strategy Now

Sign up for AWS Snowball and make your data transfer easier and faster.