AWS

What is AWS?

AWS (Amazon Web Services) is a comprehensive, evolving cloud computing platform provided by Amazon. It includes a mixture of infrastructure-as-a-service (IaaS), platform-as-a-service (PaaS) and packaged software-as-a-service (SaaS) offerings. AWS offers tools such as compute power, database storage and content delivery services.

AWS Cloud Use Cases

1. AWS helps you create powerful and flexible apps that can grow as needed.

2. Suitable for many different types of businesses.

3. Here are some simplified use cases:-

a. Enterprise IT & Analytics: Secure data management and insights.

b. Hosting & Apps: Scalable support for websites and mobile apps.

c. Gaming: Cloud-based game development and hosting.

Global Services

1. Identity and Access Management (IAM)
2. Route 53 (DNS service)
3. CloudFront (Content Delivery Network)
4. WAF (Web Application Firewall)

Region-scoped Services

1. Amazon EC2 (Infrastructure as a Service)
2. Elastic Beanstalk (Platform as a Service)
3. Lambda (Function as a Service)
4. Rekognition (Software as a Service)

IAM (AWS Identity and Access Management)

What is IAM:-

AWS Identity and Access Management (IAM) is like a bouncer for your AWS resources. It's a tool that lets you decide who gets into the party (your AWS resources) and what they're allowed to do once they're inside. So, you can manage who can access your stuff and what actions they can take, all from one central place.

IAM features

Shared access, granular permissions, secure EC2 access, MFA, identity federation, compliance, AWS integration, consistency, free to use.

Accessing IAM

1.AWS Management Console (protected by password + MFA)

2.AWS Command Line Interface (CLI): protected by access keys

3.AWS Software Developer Kit (SDK) - for code: protected by access keys

4.IAM Query API

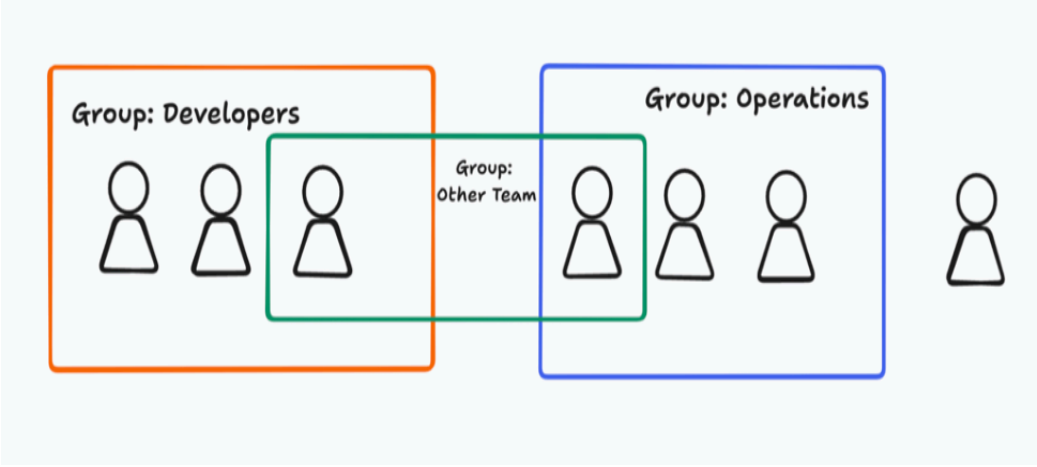
Users & Groups

1.Root account is automatically made, but don't use or share it.

2.Users are people within your organization, and can be grouped

3.Groups only contain users, not other groups

4.Users don't have to belong to a group, and user can belong to multiple groups



Permissions

1.Users or Groups can be assigned JSON documents called policies

2.These policies define the permissions of the users

3.In AWS you apply the least privilege principle: don't give more permissions than a user needs

Policies Structure

Consists of:

Version: Always "2012-10-17"

Id: Optional identifier

Statement: One or more rules

Statement Components:-

Sid: Optional statement ID

Effect: Allow/Deny access

Principal: User/role the policy applies to

Action: Allowed/denied actions

Resource: Affected AWS resources

Condition: Optional conditions

Example Policy: Grants password change, S3 bucket listing, and conditional S3 data access with MFA.

Roles for Services

Some AWS service will need to perform actions on your behalf

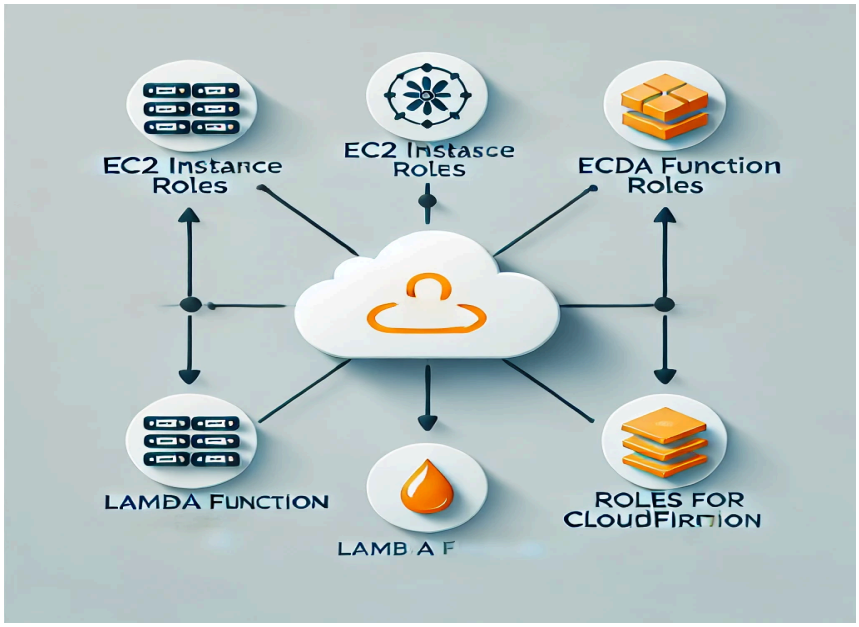
To do so, we will assign permissions to AWS services with IAM Roles

Common roles:

a. EC2 Instance Roles

b. Lambda Function Roles

c. Roles for CloudFormation



Security Tools

IAM Credentials Report (account-level)

a report that lists all your account's users and the status of their various credentials

IAM Access Advisor (user-level):-

Access advisor shows the service permissions granted to a user and when those services were last accessed.

You can use this information to revise your policies.

Guidelines & Best Practices

Avoid using the root account

One AWS user per person

Use groups for permission management

Enforce strong passwords & MFA

Assign roles for AWS services

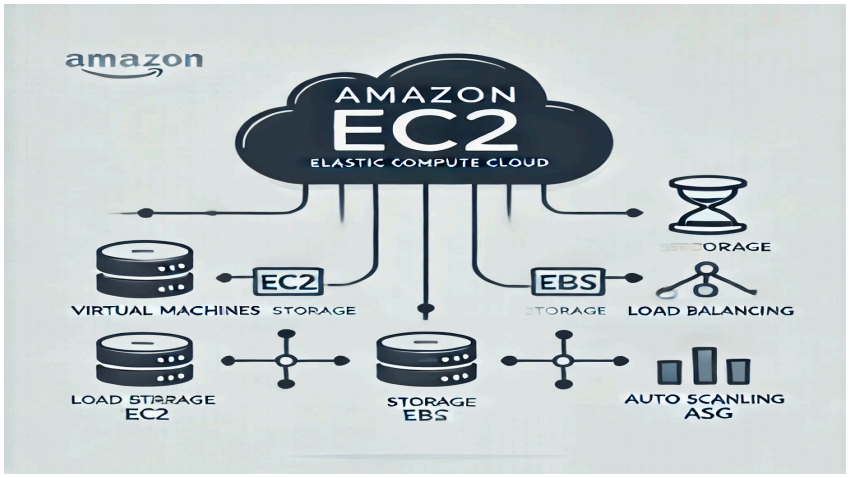
Use access keys for CLI/SDK

Audit permissions regularly

Never share IAM users or keys

Amazon EC2

Amazon EC2 (Elastic Compute Cloud): A cloud-based IaaS that allows users to rent virtual machines, store data using EBS (Elastic Block Store), distribute traffic with ELB (Elastic Load Balancer), and scale resources automatically using Auto Scaling Groups (ASG) for flexibility and performance.



EC2 sizing & configuration options

OS: Linux, Windows, or Mac OS

Compute: CPU & RAM selection

Storage: EBS, EFS, or Instance Store

Network: Speed, Public IP

Security: Firewall (Security Groups)

Automation: EC2 User Data for boot tasks (updates, software installs, file downloads)

Runs as Root User

EC2 InstanceTypes

General Purpose (M, T, A Series): Balanced compute, memory, and networking for diverse workloads.

Compute Optimized (C Series): High-performance CPUs for web servers, batch processing, and HPC.

Memory Optimized (R, X, Z Series): Fast processing for large in-memory datasets, databases, and BI.

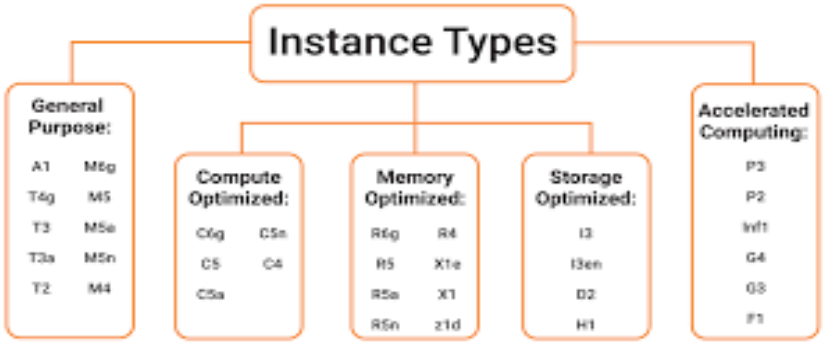
Accelerated Computing (P, G, F Series): Uses GPUs/accelerators for AI, ML, and graphics processing.

Storage Optimized (I, D, H Series): High-speed storage for databases, analytics, and data warehousing.

Key Features: Scalability, security, monitoring, multiple storage options, cost-effectiveness, and high reliability.

Performance Testing: Measure CPU, memory, disk, and network performance to choose the right instance.

Limits: Default 20 instances per region, more require AWS approval.



Introduction to Security Groups

Controls inbound/outbound traffic for EC2 instances.

Acts as a firewall with only allow rules based on IP or security groups.

Manages ports, IP ranges (IPv4/IPv6), and access control.

Inbound traffic blocked by default, outbound allowed by default.

Best practice: Separate security group for SSH access.

Can be attached to multiple instances.

Classic Ports

22 = SSH (Secure Shell) - log into a Linux instance

21 = FTP (File Transfer Protocol) – upload files into a file share

22 = SFTP (Secure File Transfer Protocol) – upload files using SSH

80 = HTTP – access unsecured websites

443 = HTTPS – access secured websites

3389 = RDP (Remote Desktop Protocol) – log into a Windows instance

SSH Summary Table

	SSH	Putty	EC2 Instance Connect
Mac	✓		✓
Linux	✓		✓
Window < 10		✓	✓
Window >= 10	✓	✓	✓

Amazon EC2 – Solutions Architect Associate Level

Private vs Public IP (IPv4):

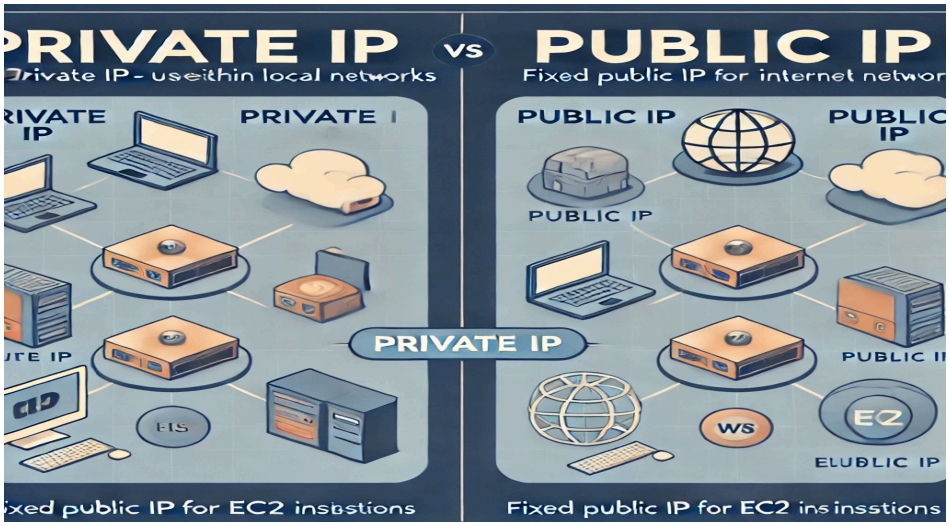
IPv4 (common) & IPv6 (newer, IoT support)

Public IP: Unique on the internet, geo-locatable.

Private IP: Used within private networks, can be duplicated across networks.

Elastic IP: Fixed public IPv4 for EC2, limited to 5 per account (can request more).

NAT + Internet Gateway allows private IPs to access the web.



Elastic Network Interfaces (ENI)

A virtual network card in a VPC with:-

Primary & secondary private IPv4

Primary & multiple IPv6 addresses

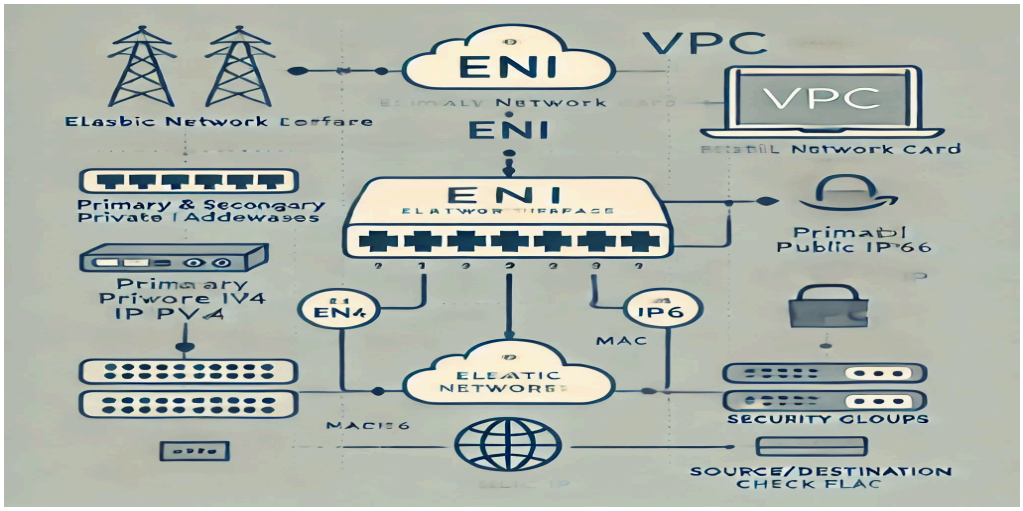
One Elastic IP (IPv4) per private IP

One public IPv4

Security groups, MAC address

Source/destination check flag

Custom description



EC2 Hibernate

Preserves in-memory (RAM) state for faster boot.

RAM state is saved to an encrypted root EBS volume.

Use Cases:

Long-running processes, saving RAM state, slow-initializing services.

Key Points:-

Supports C3, C4, C5, I3, M3, M4, R3, R4, T2, T3, etc.

RAM must be <150GB, not for bare metal instances.

Works with Amazon Linux 2, Ubuntu, RHEL, CentOS, Windows.

Root volume must be encrypted EBS, not instance store.

Available for On-Demand, Reserved, and Spot instances.

Cannot hibernate for more than 60 days.

Amazon EC2 – Instance Storage

EBS Overview:-

Elastic Block Store (EBS): Network drive for EC2 instances.

Persists data even after instance termination.

Mounted to one instance at a time (except multi-attach).

Bound to an Availability Zone (AZ).

Snapshot required to move across AZs

EBS Volume Types

General Purpose SSD (gp3, gp2) – Balanced price & performance.

Provisioned IOPS SSD (io1, io2) – High performance, low latency.

Throughput Optimized HDD (st1) – High throughput, frequently accessed.

Cold HDD (sc1) – Lowest cost, infrequent access.

Only gp2/gp3 and io1/io2 Block Express can be boot volumes.

EBS Features

Multi-Attach: Attach one volume to multiple instances (same AZ).

Encryption: Uses KMS (AES-256), automatic for snapshots & volumes.

Snapshots: Backups, can be copied across regions.

Fast Snapshot Restore (FSR): Reduces latency.

Recycle Bin: Recover deleted snapshots.

EBS Snapshot Archive: Long-term storage (24-72h restore).

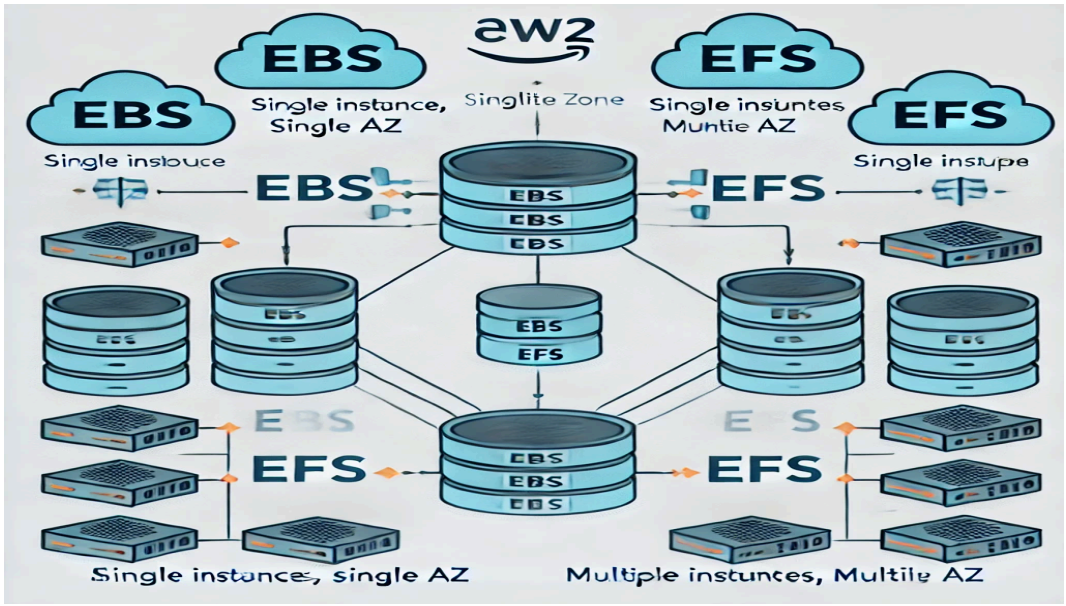
Amazon EFS (Elastic File System)

Managed NFS, multi-AZ support, auto-scaling.

Use Cases: Web hosting, content management, data sharing.

Compatible with Linux AMIs, uses NFSv4.1 protocol.

Encryption at rest via KMS.



EFS Performance & Storage

Performance Modes:

General Purpose (low latency)and Max I/O (high throughput, big data).

Throughput Modes:

Bursting, Provisioned, Elastic.

Storage Tiers:

Standard: Frequent access and Infrequent Access (EFS-IA): Cost-effective.

EBS vs. EFS

Feature	EBS	EFS
Scope	Single instance	Multiple instances
Availability	Single AZ	Multi-AZ
Performance	High IOPS	Scalable
Use Cases	Databases, persistent storage	Web hosting, file sharing
Pricing	Lower	Higher
Migration	Snapshot-based	Lifecycle policy (IA)

AWS High Availability & Scalability

Scalability

Ability of a system to handle increased load.

Vertical Scaling: Increase instance size (e.g., t2.micro → t2.large).

Common for RDS, ElastiCache.

Horizontal Scaling: Increase number of instances. Used in distributed systems.

High Availability (HA)

Ensures uptime by running in multiple Availability Zones (AZs).

Passive HA: e.g., RDS Multi-AZ.

Active HA: Used with horizontal scaling.

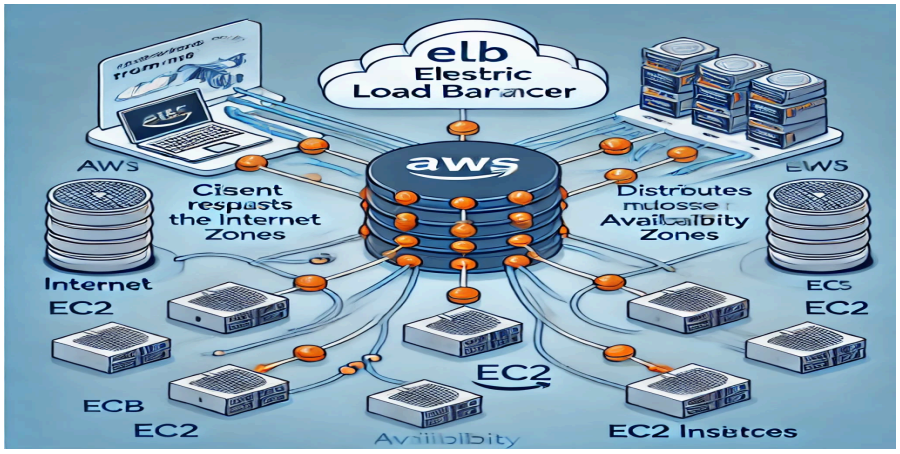
Elastic Load Balancing (ELB)

Distributes traffic across multiple instances.

Benefits: Load distribution, failure handling, health checks, SSL termination, stickiness, HA across AZs.

Types of Load Balancers

Load Balancer	Layer	Protocols	Use Case
Classic LB (CLB)	4	HTTP, HTTPS, TCP, SSL	Legacy apps
Application LB (ALB)	7	HTTP, HTTPS, WebSocket	Microservices, Routing
Network LB (NLB)	4	TCP, TLS, UDP	High performance
Gateway LB (GWLB)	3	IP Protocol	Firewalls, IDS/IPS



Application Load Balancer (ALB)

Layer 7: HTTP-based routing.

Features: Path-based & hostname-based routing, WebSockets, auto-scaling.

Target Groups: EC2, ECS, Lambda, Private IPs.

Network Load Balancer (NLB)

Layer 4: Handles millions of requests/sec.

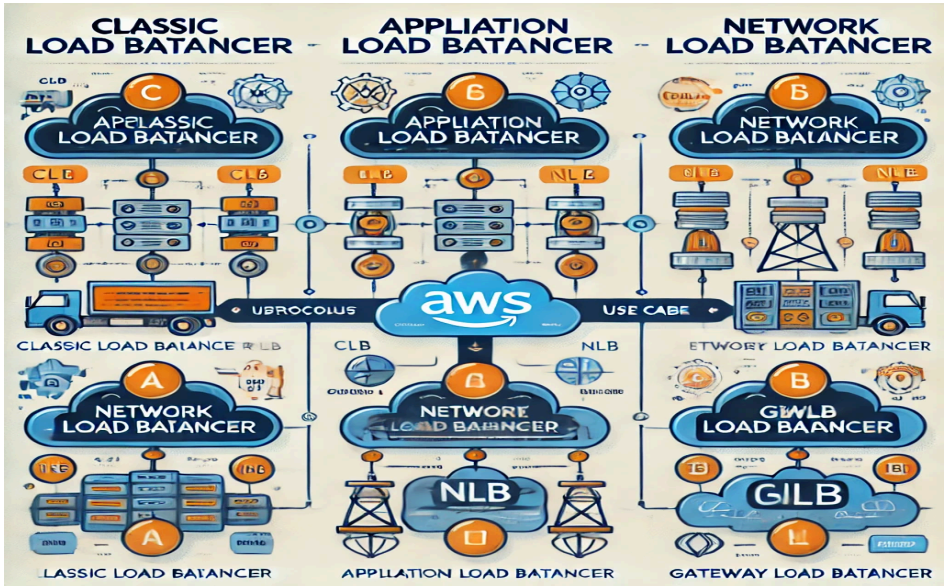
Lower latency (~100ms).

Fixed/static IP per AZ.

Gateway Load Balancer (GWLB)

Layer 3: Used for virtual appliances like firewalls, IDS.

Uses GENEVE protocol (port 6081).



Auto Scaling Groups (ASG)

Automatically adjusts EC2 instances based on demand.

Components:-

a. Launch Template: AMI, instance type, security groups, etc.

Scaling Policies:-

a. Dynamic Scaling: Based on CloudWatch metrics.

b. Scheduled Scaling: Predefined time-based scaling.

c. Predictive Scaling: AI-based forecasted scaling.

Scaling Type	Mechanism	Use Case
Target Tracking	Maintain a target metric (e.g., CPU 40%)	Simple setup
Step Scaling	Scale based on thresholds (e.g., +10 instances @ 60% CPU)	Granular control
Scheduled Scaling	Scale based on time (e.g., +10 instances @ 5 PM)	Predictable workloads
Predictive Scaling	Machine learning-based scaling	Seasonal trends

Standby State

Instances in Standby Mode remain in ASG but do not handle requests.

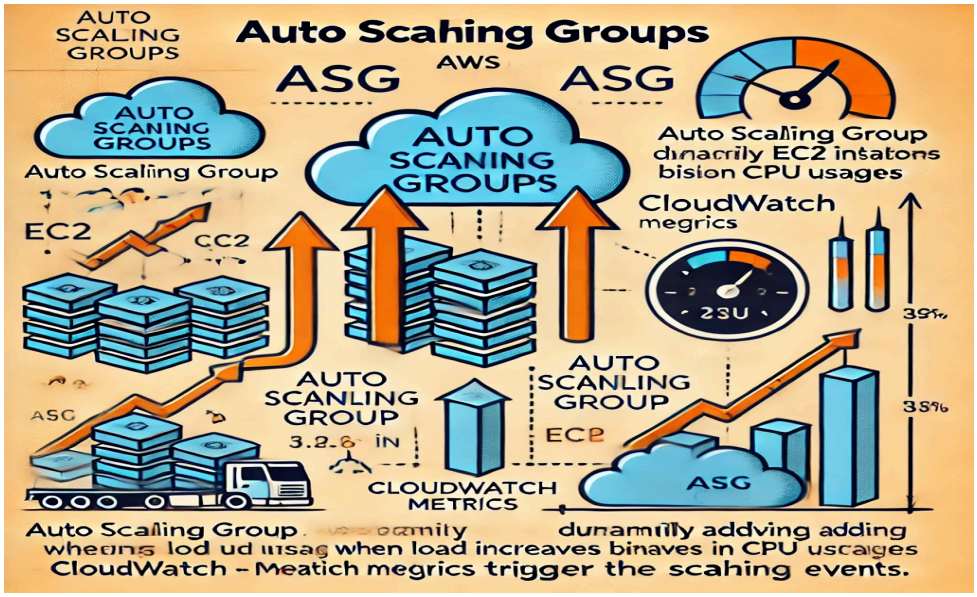
Useful for maintenance without terminating instances.

Connection Draining (Deregistration Delay):-

CLB: Connection Draining

ALB & NLB: Deregistration Delay

Allows instances to complete in-flight requests before termination.



Elastic Load Balancer - SSL & Stickiness

a. SSL/TLS Certificates: Encrypt traffic.

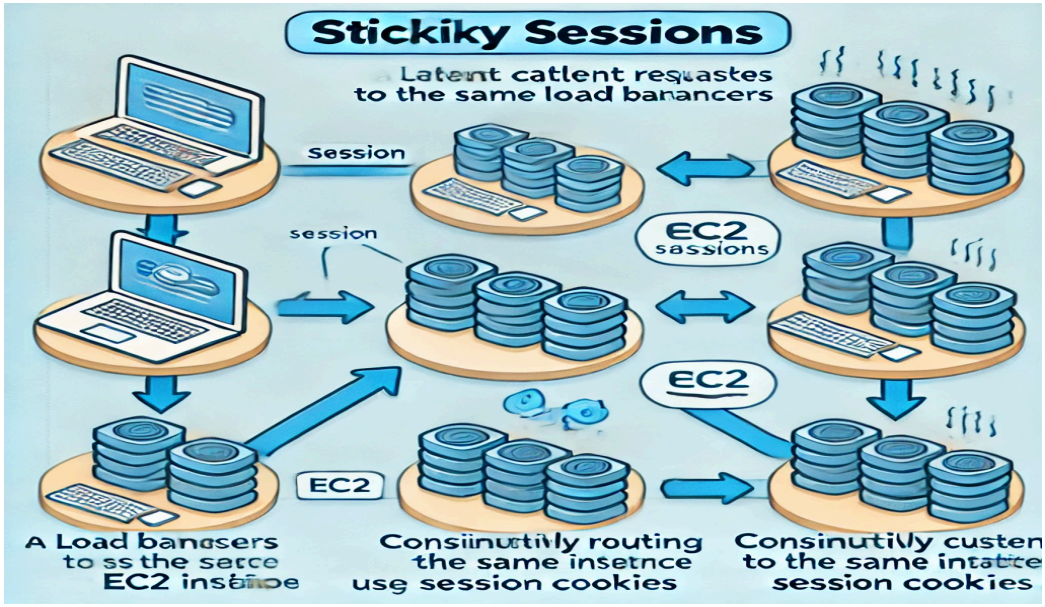
b. Managed via ACM (AWS Certificate Manager).

c. Server Name Indication (SNI): Supports multiple SSL certs per LB (ALB & NLB only).

d. Stickiness: Ensures same client is routed to same instance (via cookies).

Cross-Zone Load Balancing

Load Balancer	Default	Inter-AZ Cost
ALB	Enabled	Free
NLB	Disabled	Charged
CLB	Disabled	Free



Amazon RDS Overview

RDS (Relational Database Service): Managed service for SQL-based databases:

Supported Engines: PostgreSQL, MySQL, MariaDB, Oracle, Microsoft SQL Server, Aurora (AWS Proprietary).

Advantages over EC2-hosted DB:-

Automated provisioning, OS patching.

Continuous backups & Point-in-Time Restore.

Read replicas for performance, Multi-AZ for DR.

Vertical & horizontal scaling.

Storage: EBS-backed (gp2 or io1).

Limitations: No SSH access

RDS Storage Auto Scaling

Automatically scales storage when free space < 10%.

Supports all RDS engines.

Requires Maximum Storage Threshold setting.

RDS Read Replicas vs Multi-AZ

Read Replicas (for scaling):

Up to 15 replicas, within AZ, cross-AZ, or cross-region.

Asynchronous replication (eventual consistency).

Used for read-heavy workloads (SELECT queries).

Multi-AZ (for HA & DR):

Synchronous replication, automatic failover.

Single DNS endpoint, no manual intervention.

Not for scaling but ensures high availability.

Read Replicas can be Multi-AZ for disaster recovery.

RDS Custom:-

Managed Oracle & SQL Server with OS & DB customization.

Allows SSH access, patching, and advanced settings.

Amazon Aurora

AWS Proprietary DB (supports MySQL & PostgreSQL).

5x performance vs MySQL RDS, 3x vs PostgreSQL RDS.

Storage: Auto-scales in 10GB increments, up to 128TB.

Up to 15 replicas with faster replication (sub-10ms lag).

Failover < 30s, native HA.

Costs ~20% more than RDS but more efficient.

Aurora High Availability & Scaling:-

6 copies across 3 AZs (4 needed for writes, 3 for reads).

Self-healing with peer-to-peer replication.

Automated failover, cross-region replication supported.

Aurora Advanced Features:-

Aurora Serverless: Auto-scales based on usage, pay-per-second.

Global Aurora:-

1 Primary (Read/Write), Up to 5 Secondary Regions (Read-Only).

<1s replication lag, DR promotion in <1 min.

Aurora Machine Learning (integrates with AWS ML services like SageMaker, Comprehend).

RDS & Aurora: Backup & Monitoring

Automated Backups: Daily full backups + 5-min transaction log backups.

Retention: 1-35 days (0 to disable).

Manual Snapshots: Retained indefinitely.

Restore: Creates a new DB instance.

Aurora Cloning: Faster than snapshots, copy-on-write protocol.

Security & RDS Proxy

Encryption:-

At-rest via AWS KMS, must be set at launch.

In-flight via TLS (AWS TLS root certificates).

IAM Authentication: Use IAM roles instead of username/password.

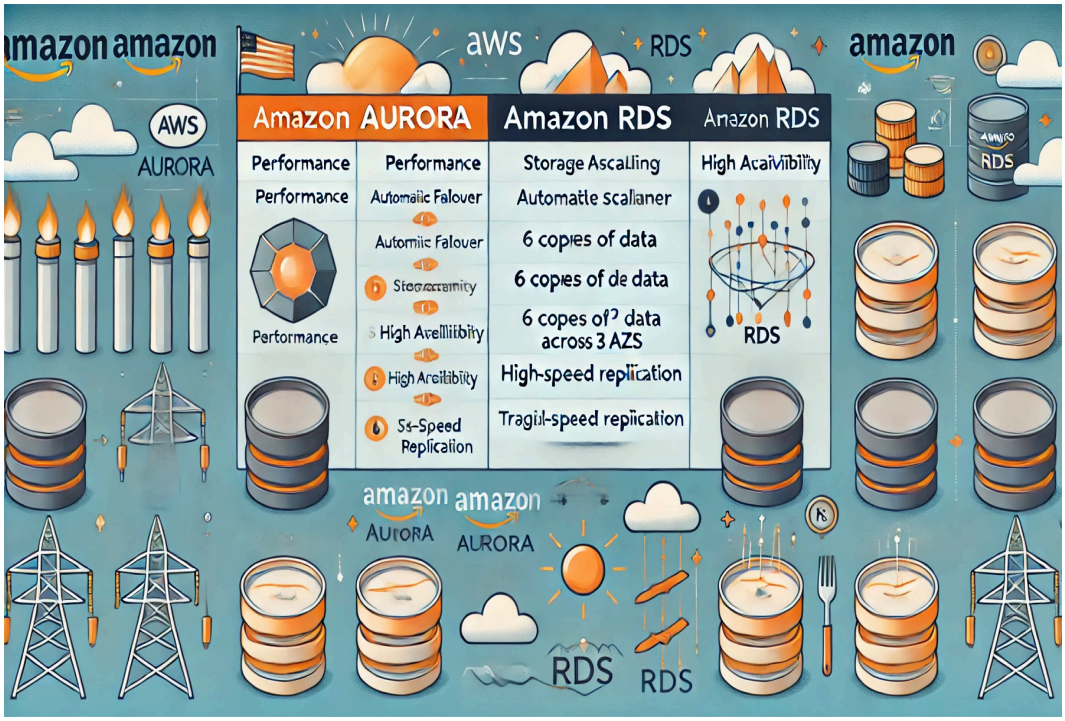
Security Groups: Control network access.

RDS Proxy:-

Pooled DB connections to reduce resource usage.

Serverless, multi-AZ, reduces failover time by 66%.

Supports RDS & Aurora (MySQL, PostgreSQL, SQL Server, MariaDB).



Amazon ElastiCache

Use Case:-

Caching layer for high-performance, low-latency data retrieval.

Reduces database load for read-heavy workloads.

Supports session storage (e.g., user sessions).

Architecture:-

App queries ElastiCache → If data not in cache, fetch from DB and store in cache.

Requires cache invalidation strategy to prevent stale data.

Redis vs Memcached

Feature	Redis	Memcached
High Availability	Multi-AZ with failover	No HA
Read Scaling	Read replicas	Sharding only
Persistence	AOF & backups	No persistence
Backup & Restore	Yes	No
Data Structures	Sets, Sorted Sets	Simple key-value
Performance	Single-threaded	Multi-threaded

ElastiCache Security:-

Redis AUTH: Password/token for cluster security.

Supports IAM authentication & SSL encryption.

Memcached: SASL-based authentication.

Caching Patterns:-

Lazy Loading: Cache read data, updates only when missing.

Write Through: Cache updated when DB is updated.

Session Store: Stores temporary session data (TTL-based).

Redis Use Case:

Gaming Leaderboards

Sorted Sets ensure real-time ranking & uniqueness.

Important Ports

Service	Port
SSH / SFTP	22
HTTP / HTTPS	80 / 443
PostgreSQL	5432
MySQL / MariaDB	3306
Oracle RDS	1521
MSSQL Server	1433
Aurora	5432 (Postgres) / 3306 (MySQL)

Amazon Route 53

What is DNS?

DNS (Domain Name System) translates human-friendly hostnames into machine IP addresses (e.g., www.google.com → 172.217.18.36).

Key Terminologies:

Domain Registrar: Amazon Route 53, GoDaddy, etc.

DNS Records: A, AAAA, CNAME, NS, etc.

Zone File: Contains DNS records.

Name Server: Resolves DNS queries (Authoritative/Non-Authoritative).

TLD (Top-Level Domain): .com, .org, .gov, etc.

SLD (Second-Level Domain): amazon.com, google.com, etc.

Route 53 Overview:-

Fully managed, highly available, and scalable DNS service with a 100% availability SLA.

Authoritative DNS (you can update DNS records).

Functions as a Domain Registrar and checks resource health.

"53" refers to the DNS port number.

Route 53 – Records:-

Define how traffic is routed for a domain.

Each record includes:

Name: e.g., example.com

Record Type: A, AAAA, CNAME, etc.

Value: IP address or another domain name

Routing Policy: Determines DNS response behavior

TTL (Time to Live): Cache duration in resolvers

Supported DNS record types:

Basic: A, AAAA, CNAME, NS

Advanced: MX, TXT, PTR, SRV, etc.

Route 53 – Hosted Zones:-

Public Hosted Zone: Routes traffic on the internet (e.g., app.example.com).

Private Hosted Zone: Routes traffic within VPCs (app.company.internal).

Route 53 - TTL (Time To Live):-

High TTL (e.g., 24 hrs): Less traffic but possibly outdated records.

Low TTL (e.g., 60 sec): More traffic but records update faster.

TTL is mandatory for all records except Alias records.

Route 53 - CNAME vs. Alias

Feature	CNAME	Alias
Maps to	Any hostname	AWS Resource (e.g., ELB, CloudFront)
Root Domain?	❌ No	✅ Yes
Cost	Paid	Free
Health Check	❌ No	✅ Yes

Route 53 – Alias Record Targets:-

Elastic Load Balancer

CloudFront

API Gateway

S3 Websites

VPC Endpoints

Route 53 – Routing Policies:-

Define how Route 53 responds to DNS queries.

Types of Routing Policies:-

Simple Routing: Single resource, random selection if multiple values.

Weighted Routing: Control % of traffic to each resource.

Latency-based Routing: Redirect to the lowest latency resource.

Failover Routing: Active-passive failover setup.

Geolocation Routing: Route traffic based on user's physical location.

Geoproximity Routing: Traffic shifting based on location bias.

IP-based Routing: Directs traffic based on client IP address ranges.

Multivalue Answer Routing: Returns multiple values/resources.

Route 53 – Health Checks:-

Types of Health Checks:

Endpoint Health Check: Monitor application/server health.

Calculated Health Check: Combine multiple checks using AND/OR logic.

CloudWatch Alarm Health Check: Monitor AWS CloudWatch alarms.

Key Features:-

Uses HTTP, HTTPS, TCP protocols.

15+ global health checkers verify endpoint health.

Healthy threshold: Default 3, customizable.

Passes only for 2xx & 3xx HTTP status codes.

Private resources require a workaround via CloudWatch.

Domain Registrar vs. DNS Service:-

Domain Registrar: Where you purchase/register a domain (e.g., GoDaddy, Amazon Route 53).

DNS Service: Manages DNS records (e.g., Route 53, Cloudflare).

You can register a domain with one provider and use another DNS service.

Steps to use Route 53 with a 3rd party domain:

Create a Hosted Zone in Route 53.

Update NS Records on the 3rd party website to Route 53's Name Servers.

Elastic Beanstalk

Elastic Beanstalk is a PaaS (Platform-as-a-Service) for deploying and scaling web applications.

Developer-centric service that abstracts infrastructure management.

Uses EC2, Auto Scaling Groups (ASG), Load Balancer (ELB), RDS, etc.

Fully managed service:

Handles capacity provisioning, load balancing, scaling, and health monitoring.

Only the application code is managed by the developer.

Full control over configuration if needed.

Beanstalk itself is free, but you pay for the underlying AWS resources.

Elastic Beanstalk – Components:-

Application: Collection of all Beanstalk components (environments, versions, configs).

Application Version: A specific version/iteration of application code.

Environment:

A collection of AWS resources running one application version.

Environment Tiers:

Web Server Environment Tier (handles HTTP requests).

Worker Environment Tier (processes background tasks).

Supports multiple environments like Dev, Test, Prod.

Elastic Beanstalk – Supported Platforms

Programming Languages & Frameworks:

Go

Java SE & Java with Tomcat

.NET Core (Linux) & .NET (Windows Server)

Node.js

PHP

Python

Ruby

Docker Support:

Single-container Docker

Multi-container Docker

Preconfigured Docker

Custom Platforms:

If not supported, you can create a custom platform using Packer Builder (advanced).

Amazon S3

S3 Basics:-

Amazon S3: Scalable storage used by websites & AWS services.

Use Cases: Backup, DR, archives, hybrid cloud, media hosting, data lakes, software delivery, static websites.

Buckets & Objects:-

Buckets: Store objects (files), globally unique names, defined at the region level.

Objects: Identified by a key (full path), max size 5TB (multipart upload >5GB).

No directories: Just keys structured with "/".

Metadata: Key-value pairs, tags (up to 10), version ID (if versioning is enabled).

Security & Policies:-

IAM Policies: User-based permissions.

Bucket Policies: JSON-based, allow/deny access, cross-account sharing.

ACLs: Object & bucket level (can be disabled).

Encryption: Encrypt objects using keys.

S3 Website Hosting:-

Supports static websites: URL format →

```
http://bucket-name.s3-website.aws-region.amazonaws.com
```

S3 Versioning & Replication:-

Versioning: Protects against unintended deletes, enables rollback.

Replication:

CRR (Cross-Region Replication): Compliance, cross-account sharing.

SRR (Same-Region Replication): Log aggregation, prod-test sync.

Only new objects replicated, old ones require Batch Replication.

DELETE rules: Can replicate delete markers (optional).

S3 Storage Classes

Storage Class	Availability	Use Case
S3 Standard	99.99%	Frequent access, big data, gaming
S3 Standard-IA	99.9%	Disaster recovery, backups
S3 One Zone-IA	99.5%	Secondary backups, recreatable data
S3 Glacier Instant Retrieval	99.9%	Archiving, access once per quarter
S3 Glacier Flexible Retrieval	99.9%	Deep archive, 1 min to 12-hour retrieval
S3 Glacier Deep Archive	99.9%	Long-term storage (12–48 hr retrieval)
S3 Intelligent Tiering	99.9%	Auto-moves between tiers based on usage

Durability: 11 9's (99.999999999%) across all storage classes.

Availability varies per class (e.g., S3 Standard = 99.99% = 53 min downtime/year).

Advanced Amazon S3

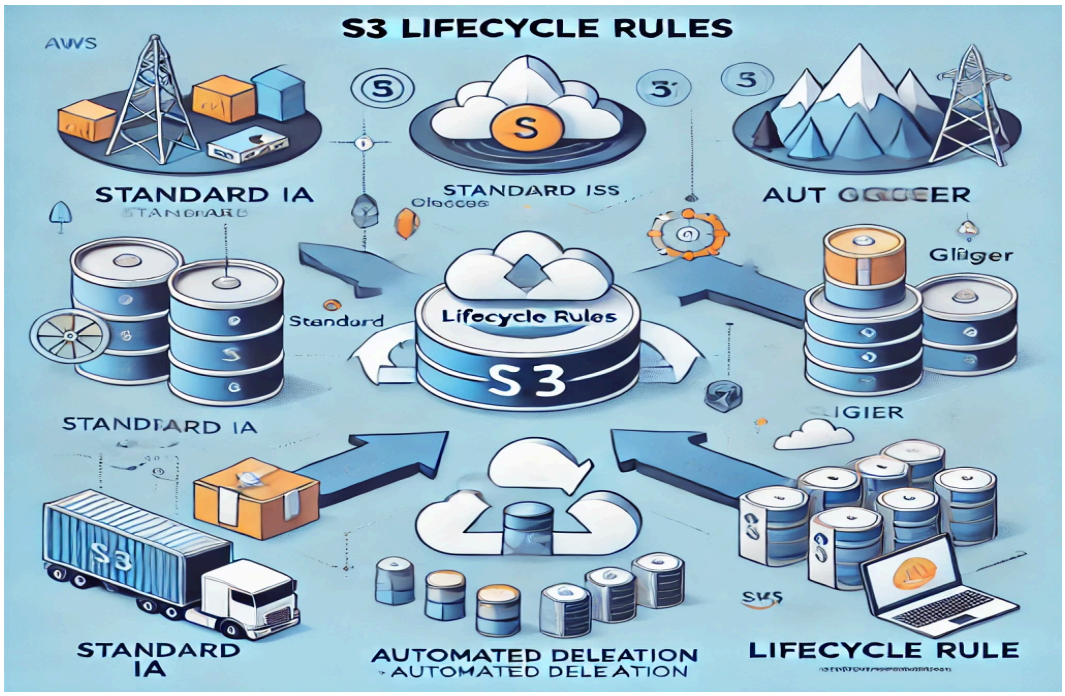
S3 Lifecycle Rules:-

Automate storage transitions & deletions using lifecycle rules.

Move infrequent data to Standard IA, archives to Glacier.

Expire/delete old files, versions, and incomplete uploads.

Use analytics for storage class recommendations.



S3 Requester Pays:-

Requesters pay for data transfer instead of bucket owners.

Useful for sharing large datasets.

Requires authentication (no anonymous access).

S3 Event Notifications:-

Triggers events (e.g., create, delete, restore) for automation.

Supports filtering by object name (*.jpg).

Works with EventBridge for advanced filtering & multiple destinations.

S3 Performance:-

Baseline: 3,500 write or 5,500 read requests/sec per prefix.

Multi-Part Upload: Speeds up large file transfers (>100MB).

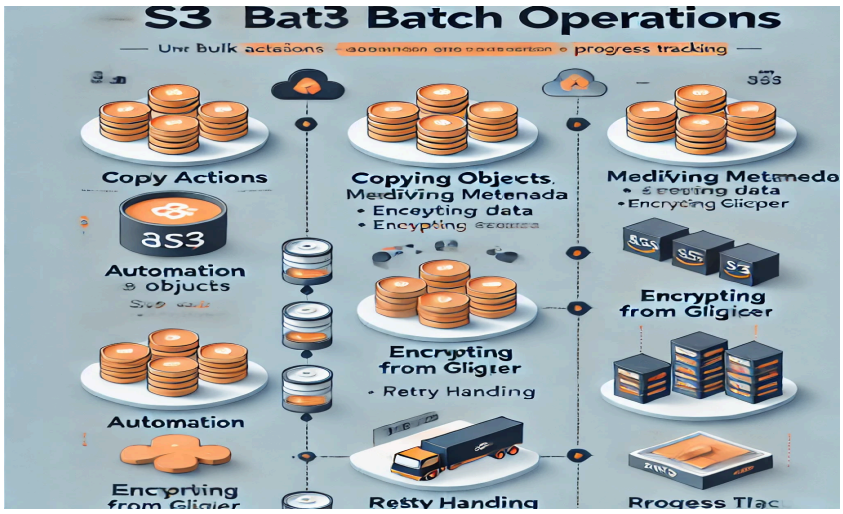
Transfer Acceleration: Uses AWS Edge Locations for faster uploads.

Byte-Range Fetches: Parallelizes downloads for better performance.

S3 Batch Operations:-

Bulk actions on multiple S3 objects (copy, tag, encrypt, restore).

Automates retries, progress tracking, and notifications.



S3 Select & Glacier Select:-

SQL-based filtering to retrieve only needed data.

Reduces network transfer & CPU costs.

Amazon S3 Security

S3 Encryption:-

Amazon S3 provides multiple encryption methods to secure data:

Server-Side Encryption (SSE):

SSE-S3: Uses Amazon S3-managed keys (AES-256) (Enabled by default).

SSE-KMS: Uses AWS KMS for key management, providing user control and audit logging.

SSE-C: Customers manage their own encryption keys; AWS does not store them.

Client-Side Encryption: Data is encrypted before being uploaded and decrypted upon retrieval by the client.

Encryption in Transit: Uses SSL/TLS for secure data transmission.

Default Encryption vs. Bucket Policies: SSE-S3 encryption is automatically applied to new objects, but policies can enforce stricter encryption requirements.

S3 CORS (Cross-Origin Resource Sharing):-

CORS allows web applications in different domains to access S3 resources.

Requires proper CORS headers (Access-Control-Allow-Origin).

Supports specific origins or wildcard *.

S3 MFA Delete:-

Adds an extra security layer requiring MFA for:

Permanently deleting object versions.

Suspending versioning.

Requires versioning to be enabled.

Only the bucket owner can enable/disable MFA Delete.

S3 Access Logs:-

Logs all access attempts (authorized or denied) into a designated S3 bucket for auditing.

Warning: Logging into the same monitored bucket causes an infinite loop.

S3 Pre-Signed URLs:-

Temporary URLs allow users to securely upload/download objects.

Expiration:

S3 Console: 1 to 720 minutes.

AWS CLI: Up to 168 hours (--expires-in parameter).

Use cases:

Allow temporary access to premium content.

Dynamic user-based downloads.

Temporary uploads to a specific bucket location.

. S3 Glacier Vault Lock:-

WORM (Write Once, Read Many) compliance model.

Vault Lock Policy: Locks policies to prevent future modifications.

Used for long-term compliance and data retention.

S3 Object Lock:-

Ensures data is not deleted or modified for a set period.

Modes:

Compliance Mode: Prevents deletion or modification, even by the root user.

Governance Mode: Restricts deletion but allows specific users to modify settings.

Retention Period: Protects objects for a defined time.

Legal Hold: Indefinite protection that can be freely placed or removed.

S3 Access Points:-

Simplifies security management for large-scale S3 usage.

Each access point has its own DNS name and policy.

VPC Access Points require a VPC Endpoint for controlled internal access.

S3 Object Lambda:-

Modifies data before retrieval using AWS Lambda.

Use cases:

Redacting sensitive data (e.g., PII in logs).

Data format conversions (e.g., XML to JSON).

On-the-fly image transformations (resizing, watermarking).

CloudFront & AWS Global Accelerator

CloudFront Overview:-

Content Delivery Network (CDN) improves read performance by caching content at 216+ edge locations globally.

Security: DDoS protection, AWS Shield & Web Application Firewall (WAF).

CloudFront Origins:-

S3 Bucket (OAC replaces OAI for enhanced security).

Custom Origin (HTTP): ALB, EC2, S3 website, or any HTTP backend.

CloudFront vs. S3 Cross-Region Replication:-

CloudFront: Global caching (TTL-based), best for static content.

S3 CRR: Real-time updates, read-only, best for dynamic content.

CloudFront – ALB or EC2 as an Origin:-

EC2 as Origin: Must be public, allow Edge Location IPs.

ALB as Origin: Private EC2 allowed, ALB must be public, allow security group access.

CloudFront - Geo Restriction:-

Allowlist (access granted to specific countries) & Blocklist (deny access to specific countries) based on Geo-IP.

Use Case: Compliance with copyright laws.

CloudFront - Price Classes:-

Reduce edge locations to save costs.

Price Class All (best performance, highest cost).

Price Class 200 (most locations, excludes costly ones).

Price Class 100 (only least expensive regions).

CloudFront – Cache Invalidation:-

Refresh cache manually before TTL expiry using CloudFront Invalidation.

Can invalidate all files (*) or specific paths (/images/).

AWS Global Accelerator

Overview:-

Uses AWS internal network for faster global access.

Provides 2 Anycast IPs for your application.

Routes traffic via Edge Locations → Your application.

Performance & Health Checks:-

Low latency & fast failover (<1 min) using intelligent routing.

Health checks monitor applications and reroute traffic when needed.

Security:-

Only 2 external IPs need whitelisting (simplifies security).

DDoS protection via AWS Shield.

AWS Global Accelerator vs. CloudFront

Feature	CloudFront	Global Accelerator
Optimized for	Cacheable (static) & dynamic content	Low-latency TCP/UDP traffic
Routing	Serves content at Edge	Routes packets via AWS network
Use Cases	Website, APIs, video streaming	Gaming (UDP), IoT (MQTT), VoIP
Failover	DNS-based	Fast regional failover

AWS Storage Extras

AWS Snow Family Overview:-

Secure, portable devices for data migration & edge computing.

Snowcone (8TB HDD / 14TB SSD) – Small, lightweight, used in constrained spaces.

Snowball Edge (80TB HDD) – Large-scale migration, disaster recovery.

Snowmobile (100PB per unit) – Transfer exabytes of data.

Edge Computing: Use Snowball Edge / Snowcone for ML, preprocessing, and media transcoding.

Process: Request device → Copy data → Ship to AWS → Auto-import to S3.

Amazon FSx – Managed File Systems:-

Fully managed file systems for high-performance computing.

FSx for Windows: SMB, NTFS support, multi-AZ, Active Directory integration.

FSx for Lustre: Parallel, high-speed (HPC, ML, financial modeling).

FSx for NetApp ONTAP: Compatible with NFS, SMB, iSCSI.

FSx for OpenZFS: Optimized for Linux workloads, fast cloning, snapshots.

FSx Integration: Works with EC2, EKS, VMware, AppStream, AWS Direct Connect.

AWS Storage Gateway – Hybrid Cloud Storage:-

Connects on-prem storage to AWS cloud.

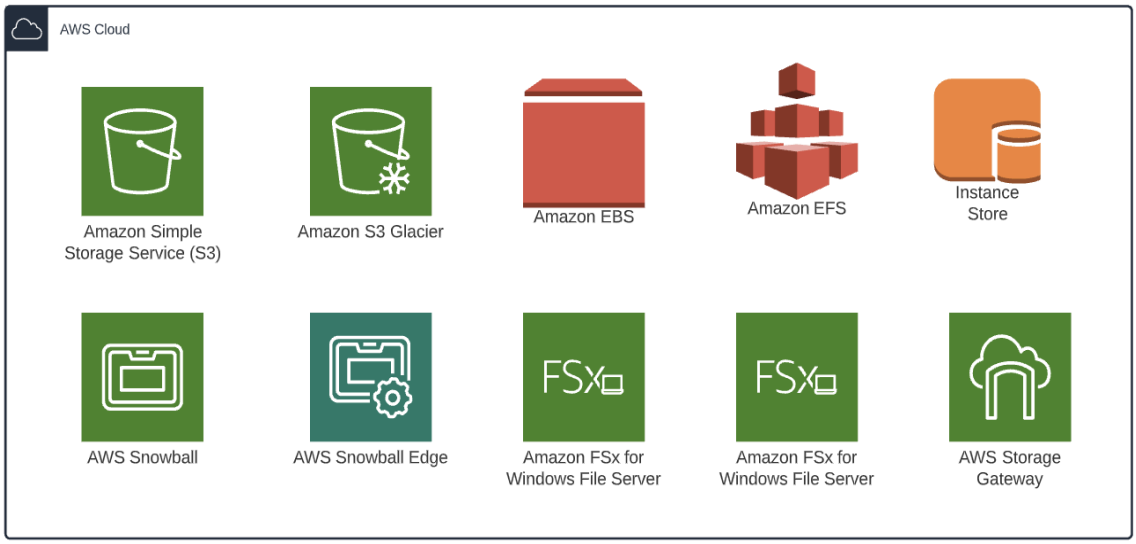
S3 File Gateway: Access S3 using NFS/SMB.

FSx File Gateway: Native Windows compatibility.

Volume Gateway: Block storage (backed by EBS snapshots).

Tape Gateway: Replaces physical tape backups with S3/Glacier storage.

Hardware Appliance: Physical device to use Storage Gateway without virtualization.



AWS Transfer Family – FTP Services:-

Managed service for file transfers to S3/EFS over:

FTP, FTPS, SFTP

Integrated with Active Directory, Cognito, Okta

Used for file sharing, CRM, ERP, public datasets

AWS DataSync – Large-Scale Data Transfers:-

Moves data on-prem & between AWS services.

Supports S3, EFS, FSx, HDFS, SMB, NFS.

Scheduled tasks (hourly, daily, weekly).

Preserves file permissions & metadata.

AWS Storage Options Comparison

Storage Type	Use Case
S3	Object storage
S3 Glacier	Archival storage
EBS	Block storage for EC2
Instance Store	Temporary storage for EC2
EFS	Scalable Linux file system
FSx for Windows	Windows file system
FSx for Lustre	High-performance file system
FSx for NetApp	Multi-protocol storage
FSx for OpenZFS	ZFS-based file system
Storage Gateway	Hybrid cloud storage
Transfer Family	Managed FTP, SFTP, FTPS
DataSync	Automated data transfers
Snow Family	Physical data migration