CHEAT SHEET

Quick review highlights for your upcoming exam!

AWS Certified Cloud Practitioner (CLF-C02)

https://shapingpixel.com


aws
Analytics

Amazon Athena

What is Amazon Athena?

Amazon Athena is an interactive serverless service used to analyze data directly in Amazon Simple Storage Service using standard SQL ad-hoc queries.

Pricing Details:

- Charges are applied based on the amount of data scanned by each query at standard S3 rates for storage, requests, and data transfer.
- Canceled queries are charged based on the amount of data scanned.
- No charges are applied for Data Definition Language (DDL) statements.
- Charges are applied for canceled queries also based on the amount of data scanned.
- Additional costs can be reduced if data gets compressed, partitioned, or converted into a columnar format.

Functions of Athena:

- It helps to analyze different kinds of data (unstructured, semi-structured, and structured) stored in Amazon S3.
- Using Athena, ad-hoc queries can be executed using ANSI SQL without actually loading the data into Athena.
- It can be integrated with Amazon Quick Sight for data visualization and helps to generate reports with business intelligence tools.
- It helps to connect SQL clients with a JDBC or an ODBC driver.
- It executes multiple queries in parallel, so no need to worry about compute resources.
- It supports various standard data formats, such as CSV, JSON, ORC, Avro, and Parquet.

Amazon OpenSearch Service

OpenSearch Service is a free and open-source search engine for all types of data like textual, numerical, geospatial, structured, and unstructured.

What is Amazon ES?

Amazon Elasticsearch Service (Amazon ES) is a managed service that allows users to deploy, manage, and scale Elasticsearch clusters in the AWS Cloud. Amazon ES provides direct access to the Elasticsearch APIs.

Amazon OpenSearch Service can be integrated with following services:

- Amazon CloudWatch
- Amazon CloudTrail
- Amazon Kinesis
- Amazon S3
- AWS IAM
- AWS Lambda
- Amazon DynamoDB

- ✔ Amazon OpenSearch Service with Kibana (visualization) & Logstash (log ingestion) provides an enhanced search experience for the applications and websites to find relevant data quickly.
- ✔ Amazon OpenSearch Service launches the Elasticsearch cluster's resources and detects the failed Elasticsearch nodes and replaces them.
- ✔ The OpenSearch Service cluster can be scaled with a few clicks in the console.

Pricing Details:

- Charges are applied for each hour of use of EC2 instances and storage volumes attached to the instances.
- Amazon OpenSearch Service does not charge for data transfer between availability zones.

Download the Latest Q&A Materials From www.shapingpixel.com

Amazon EMR

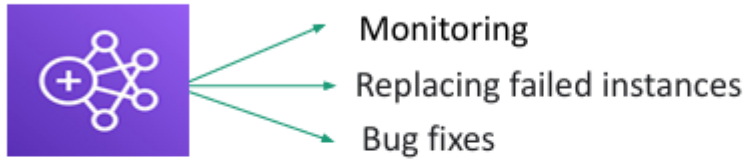
What is Amazon EMR?

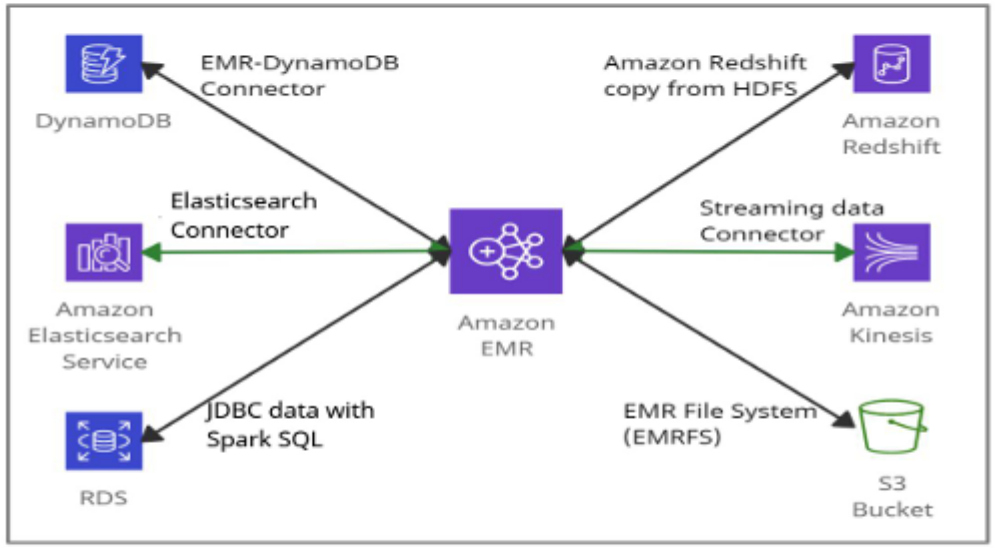
Amazon EMR (Elastic Map Reduce) is a service used to process and analyze large amounts of data in the cloud using Apache Hive, Hadoop, Apache Flink, Spark, etc.

- ✔ The main component of EMR is a cluster that collects Amazon EC2 instances (also known as nodes in EMR).
- ✔ It decouples the compute and storage layer by scaling independently and storing cluster data on Amazon S3.
- ✔ It also controls network access for the instances by configuring instance firewall settings.
- ✔ It offers basic functionalities for maintaining clusters such as monitoring, replacing failed instances, bug fixes, etc.
- ✔ It analyzes machine learning workloads using Apache Spark MLlib and TensorFlow, clickstream workloads using Apache Spark and Apache Hive, and real-time streaming workloads from Amazon Kinesis using Apache Flink.
- ✔ It provides more than one compute instances or containers to process the workloads and can be executed on the following AWS services:

Amazon EC2 **Amazon EKS** **AWS Outposts**

It offers basic functionalities for maintaining clusters such as





Amazon EMR can be accessed in the following ways:

- EMR Console
- AWS Command Line Interface (AWS CLI)
- Software Development Kit (SDK)
- Web Service API

Amazon Kinesis Data Streams

What are Amazon Kinesis Data Streams?

Amazon Kinesis Data Streams (KDS) is a scalable real-time data streaming service. It captures gigabytes of data from sources like website clickstreams, events streams (database and location-tracking), and social media feeds.

Amazon Kinesis Data Streams

Amazon Kinesis is a service used to collect, process, and analyze real-time streaming data. It can be an alternative to Apache Kafka.

- ❑ Kinesis family consists of Kinesis Data Streams, Kinesis Data Analytics, Kinesis Data Firehose, and Kinesis Video Streams.
- ❑ The Real-time data can be fetched from Producers that are Kinesis Streams API, Kinesis Producer Library (KPL), and Kinesis Agent.
- ❑ It allows building custom applications known as Kinesis Data Streams applications (Consumers), which reads data from a data stream as data records.

Amazon Kinesis Data Streams

Data Streams are divided into Shards / Partitions whose data retention is 1 day (by default) and can be extended to 7 days

Each shard provides a capacity of 1MB per second input data and 2MB per second output data.

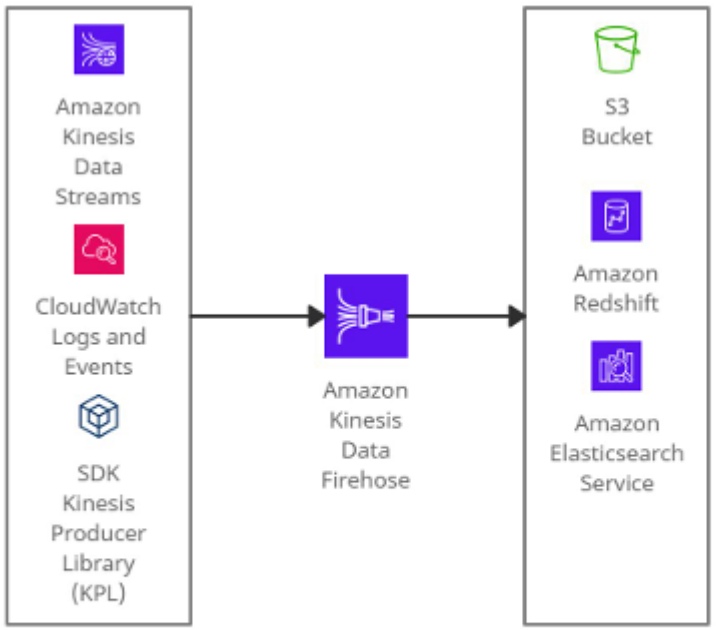
Amazon Kinesis Data Firehose

What is Amazon Kinesis Data Firehose?

Amazon Kinesis Data Firehose is a serverless service used to capture, transform, and load streaming data into data stores and analytics services.

- It synchronously replicates data across three AZs while delivering them to the destinations.
- It allows real-time analysis with existing business intelligence tools and helps to transform, batch, compress and encrypt the data before delivering it.
- It creates a Kinesis Data Firehose delivery stream to send data. Each delivery stream keeps data records for one day.
- It has 60 seconds minimum latency or a minimum of 32 MB of data transfer at a time.
- Kinesis Data Streams, CloudWatch events can be considered as the source(s) to Kinesis Data Firehose.

It delivers streaming data to the following services:



AWS Kinesis Data Firehose

Amazon Managed Streaming for Apache Kafka

What is Amazon MSK?

Amazon MSK is a managed cluster service used to build and execute Apache Kafka applications for processing streaming data.

It helps to populate machine learning applications, analytical applications, data lakes, and stream changes to and from databases using Apache Kafka APIs.

It provides multiple kinds of security for Apache Kafka clusters, including:

- AWS IAM for API Authorization
- Encryption at Rest
- Apache Kafka Access Control Lists (ACLs)

✔ It easily configures applications by removing all the manual tasks used to configure.

The steps which Amazon MSK manages are:

- ❖ Replacing servers during failures
- ❖ Handling server patches and upgrades with no downtime
- ❖ Maintenance of Apache Kafka clusters
- ❖ Maintenance of Apache ZooKeeper
- ❖ Multi-AZ replication for Apache Kafka clusters
- ❖ Planning scaling events

Amazon MSK Integrates with:

AWS Glue: To execute Apache Spark job on Amazon MSK cluster

Amazon Kinesis Data Analytics: To execute Apache Flink job on Amazon MSK cluster

Lambda Functions

Amazon Redshift

What is Amazon Redshift?

Amazon Redshift is a fast and petabyte-scale, SQL based data warehouse service used to analyze data easily.

Pricing Details:

- It offers on-demand pricing that will charge by the hour with no commitments and no upfront costs.
- Charges are applied based on the type and number of nodes used in the Redshift Cluster.
- Charged based on the number of bytes scanned by Redshift Spectrum, rounded up to 10MB minimum per query.

Functions of Redshift:

- It supports Online Analytical Processing (OLAP) type of DB workloads and analyzes them using standard SQL and existing Business Intelligence (BI) tools (AWS QuickSight or Tableau).
- It is used for executing complex analytic queries on semi-structured and structured data using query optimization, columnar-based storage, and Massively Parallel Query Execution (MPP).
- Redshift Spectrum helps to directly query from the objects (files) on S3 without actually loading them.
- It has the capability to automatically copy snapshots (automated or manual) of a cluster to another AWS Region

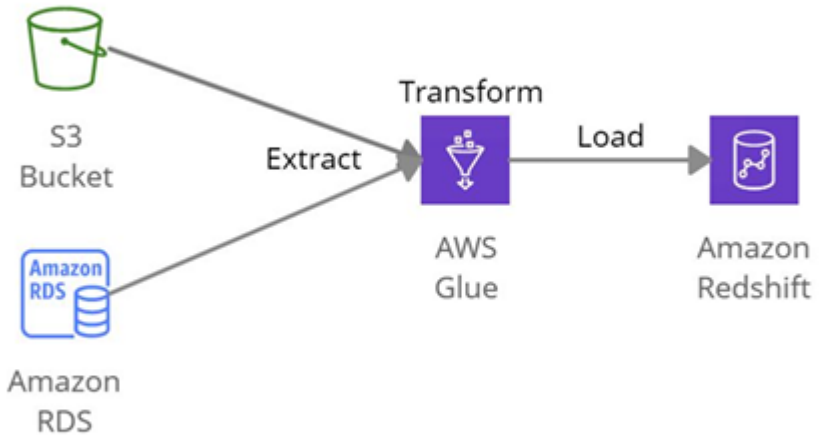
AWS Glue

What is AWS Glue?

AWS Glue is a serverless ETL (extract, transform, and load) service used to categorize data and move them between various data stores and streams.

AWS Glue works with the following services:

- **Redshift** - for data warehouses
- **S3** - for data lakes
- **RDS or EC2 instances** - for data stores



Properties of AWS Glue:

- It supports data integration, preparing and combining data for analytics, machine learning, and other applications' development.
- It has a central repository known as the AWS Glue Data Catalog that automatically generates Python or Scala code.
- It processes semi-structured data using a simple 'dynamic' frame in the ETL scripts similar to an Apache Spark data frame that organizes data into rows and columns.
- It helps execute the Apache Spark environment's ETL jobs by discovering data and storing the associated metadata in the AWS Glue Data Catalog.
- AWS Glue and Spark can be used together by converting dynamic frames and Spark data frames to perform all kinds of analysis.
- It allows organizations to work together and perform data integration tasks, like extraction, normalization, combining, loading, and running ETL workloads.

AWS Lake Formation

A data lake is a secure repository that stores all the data in its original form and is used for analysis.

What is AWS Lake Formation?

AWS Lake Formation is a cloud service that is used to create, manage and secure data lakes. It automates the complex manual steps required to create data lakes.

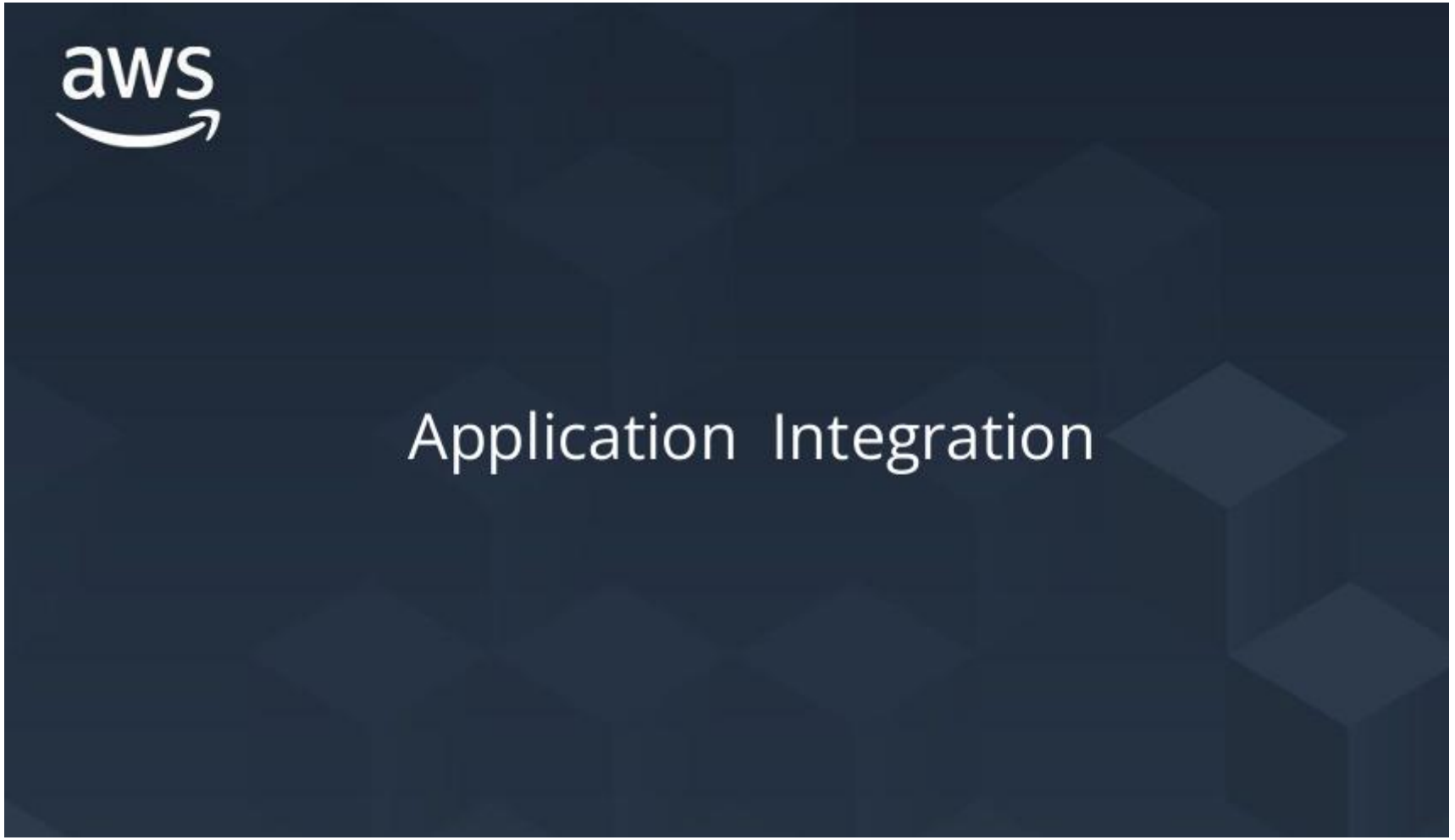
AWS Lake Formation integrates with:

- *Amazon CloudWatch*
- *Amazon CloudTrail*
- *Amazon Glue:* Both use same Data Catalog
- *Amazon Redshift Spectrum*
- *Amazon EMR*
- *AWS Key Management Service*
- *Amazon Athena:* Athena's users can query those AWS Glue catalog which has Lake Formation permissions on them.

- Lake Formation is pointed at the data sources, then crawls the sources and moves the data into the new Amazon S3 data lake.
- It integrates with AWS Identity and Access Management (IAM) to provide fine-grained access to the data stored in data lakes using a simple grant/revoke process

Pricing Details:

- Charges are applied based on the service integrations (AWS Glue, Amazon S3, Amazon EMR, Amazon Redshift) at a standard rate


aws
Application Integration

AWS Step Functions

What is AWS Step Functions?

AWS Step Functions is a serverless orchestration service that converts an application's workflow into a series of steps by combining AWS Lambda functions and other AWS services.

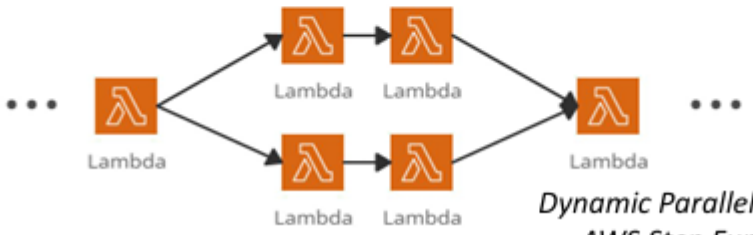
Standard Workflows

- It executes once in a workflow execution for up to one year.
- They are ideal for long-running and auditable workflows.

Express Workflows

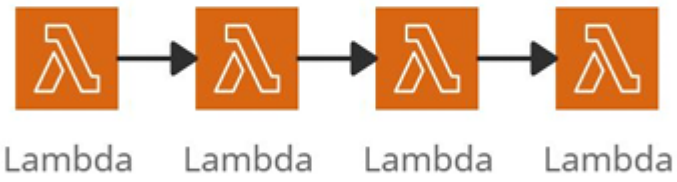
- It executes at-least-once in a workflow execution for up to five minutes.
- They are ideal for high-processing workloads, such as streaming data processing and IoT data ingestion.

Executions are the instances where workflow runs to perform tasks.



Dynamic Parallelism using AWS Step Functions

- AWS Step Functions resembles state machines and tasks. Each step in a workflow is a state. The output of one step signifies an input to the next results in functions orchestration.
- It helps to execute each step in an order defined by the business logic of the application.
- It provides some built-in functionalities like sequencing, error handling, timeout handling, and removing a significant operational overhead from the team.
- It can control other AWS services, like AWS Lambda (to perform tasks), processing machine learning models, AWS Glue (to create an extract, transform, and load (ETL) workflows), and automated workflows that require human approval.
- It provides multiple automation features like routine deployments, upgrades, installations, migrations, patch management, infrastructure selection, and data synchronization



Functions Orchestration using AWS Step Function

Amazon EventBridge

What is Amazon EventBridge?

Amazon EventBridge is a serverless event bus service that connects applications with data from multiple sources.

Amazon EventBridge integrates with the following services:

- AWS CloudTrail
- AWS CloudFormation
- AWS Config
- AWS Identity and Access Management (IAM)
- AWS Kinesis Data Streams
- AWS Lambda

Functions of Amazon EventBridge:

- An event bus is an entity that receives events, and rules get attached to that event bus that matches the events received.
- It helps to build loosely coupled and distributed event-driven architectures.
- It connects applications and delivers the events without the need to write custom code.
- It delivers a stream of real-time data from SaaS applications or other AWS services and routes that data to different targets such as Amazon EC2 instances, Amazon ECS tasks, AWS CodeBuild projects, etc.
- It sets up routing rules that determine the targets to build application architectures that react according to the data sources.
- The EventBridge schema registry stores a collection of event structures (schemas) and allows users to download code for those schemas in the IDE representing events as objects in the code.

Amazon SNS

What is Amazon SNS?

Amazon Simple Notification Service (Amazon SNS) is a serverless notification service that offers message delivery from publishers to subscribers.

SNS helps to publish messages to many subscriber endpoints:

- Amazon SQS Queues
- AWS Lambda Functions
- Email
- Amazon Kinesis Data Firehose
- Mobile push
- SMS

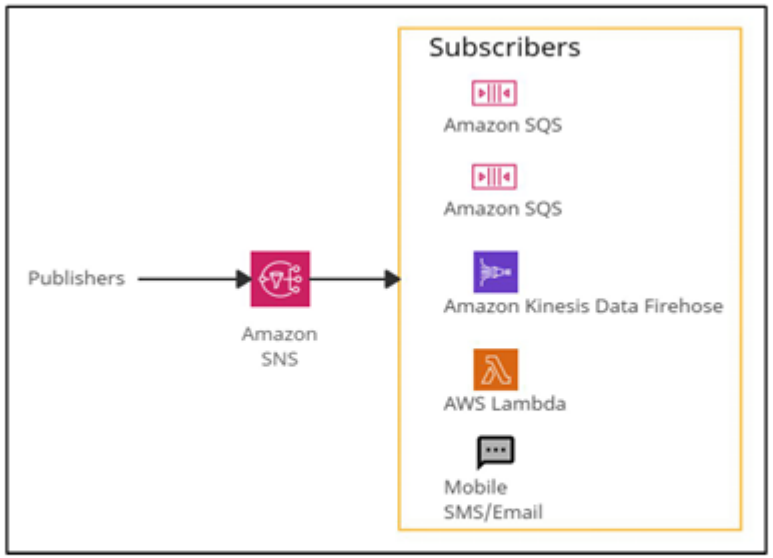
✔ It creates asynchronous communication between publishers and subscribers by sending messages to a 'topic.'

✔ It supports application-to-application subscribers that include Amazon SQS and other AWS services and Application-to-person subscribers that include Mobile SMS, Email, etc.

- The producer sends one message to one SNS topic.
- Multiple receivers (subscribers) listen for the notification of messages.
- All the subscribers will receive all the messages.

Example:

1 message, 1 topic, 10 subscribers so that a single message will be notified to 10 different subscribers.



Amazon SNS

Amazon Simple Queue Service (SQS)

What are Amazon Simple Queue Service (SQS)?

Amazon Simple Queue Service (SQS) is a serverless service used to decouple (loose couple) serverless applications and components.

- The queue represents a temporary repository between the producer and consumer of messages.
- It can scale up to 1-10000 messages per second.
- The default retention period of messages is four days and can be extended to fourteen days.
- SQS messages get automatically deleted after being consumed by the consumers.
- SQS messages have a fixed size of 256KB.

There are two SQS Queue types:

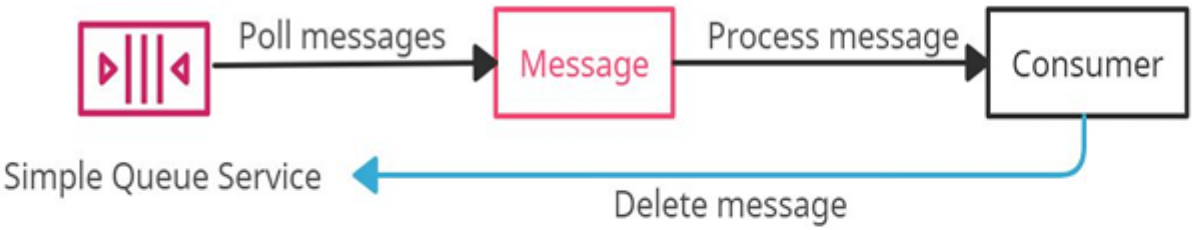
Standard Queue -

- The unlimited number of transactions per second.
- Messages get delivered in any order.
- Messages can be sent twice or multiple times.

FIFO Queue -

- 300 messages per second.
- Support batches of 10 messages per operation, results in 3000 messages per second.
- Messages get consumed only once.

Dead-Letter Queue is a queue for those messages that are not consumed successfully. It is used to handle message failure.



Delay Queue is a queue that allows users to postpone/delay the delivery of messages to a **queue** for a specific number of seconds. Messages can be delayed for 0 seconds (default) -15 (maximum) minutes.

Visibility Timeout is the amount of time during which SQS prevents other consumers from receiving (poll) and processing the messages.
Default visibility timeout - 30 seconds
Minimum visibility timeout - 0 seconds
Maximum visibility timeout - 12 hours

AWS AppSync

What is AWS AppSync?

AWS AppSync is a serverless service used to build GraphQL API with real-time data synchronization and offline programming features.

GraphQL is a data language built to allow apps to fetch data from servers.

The different data sources supported by AppSync are:

- Amazon DynamoDB tables
- RDS Databases
- Amazon Elasticsearch
- AWS Lambda Functions
- Third Party HTTP Endpoints

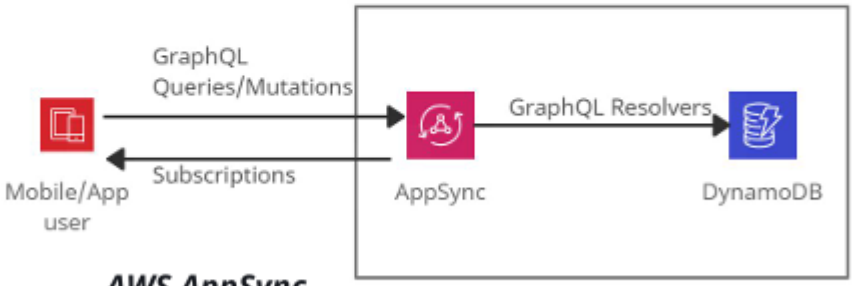
- It replaces the functionality of Cognito Sync by providing offline data synchronization.
- It improves performance by providing data caches, provides subscriptions to support real-time updates, and provides client-side data stores to keep off-line clients in sync.
- It offers certain advantages over GraphQL, such as enhanced coding style and seamless integration with modern tools and frameworks like iOS and Android
- AppSync interface provides a live GraphQL API feature that allows users to test and iterate on GraphQL schemas and data sources quickly.
- Along with AppSync, AWS provides an Amplify Framework that helps build mobile and web applications using GraphQL APIs.

AWS AppSync

Queries: For fetching data from the API

Mutations: For changing data via API

Subscriptions: The connections for streaming data from API



AWS AppSync

Amazon Simple Workflow Service

What is Amazon Simple Workflow Service?

Amazon Simple Workflow Service (Amazon SWF) is used to coordinate work amongst distributed application components.

A task is a logical representation of work performed by a component of the application.

Tasks are performed by implementing workers and execute either on Amazon EC2 or on on-premise servers (which means it is not a serverless service).

Amazon SWF stores tasks and assigns them to workers during execution.

It controls task implementation and coordination, such as tracking and maintaining the state using API.

It helps to create distributed asynchronous applications and supports sequential and parallel processing.

It is best suited for human-intervened workflows.

Amazon SWF is a less-used service, so AWS Step Functions is the better option than SWF.


aws
AWS Cost Management

AWS Cost Explorer

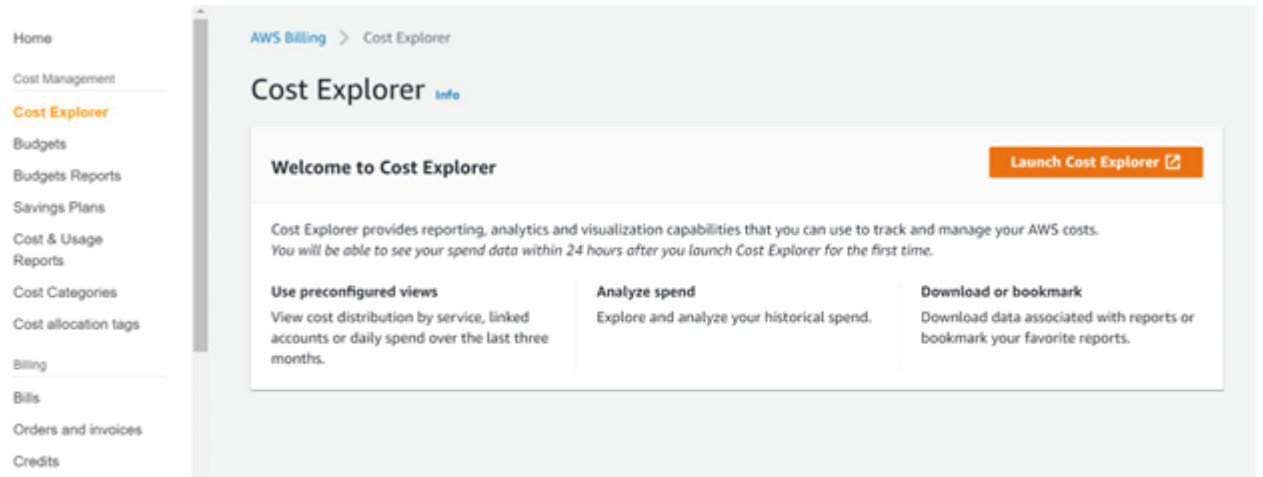
What is AWS Cost Explorer?

AWS Cost Explorer is a UI-tool that enables users to analyze the costs and usage with a graph, the Cost Explorer cost and usage reports, and the Cost Explorer RI report. It can be accessed from the Billing and Cost Management console.

- ✔ The first time the user signs up for Cost Explorer, it directs through the console's main parts.
- ✔ It prepares the data regarding costs & usage and displays up to 12 months of historical data (might be less if less used), current month data, and then calculates the forecast data for the next 12 months.

The default reports provided by Cost Explorer are:

- Cost and Usage Reports
- Reserved Instance Reports



AWS Cost Explorer

AWS Budgets

What is AWS Budgets?

AWS Budgets enables the customer to set custom budgets to track cost and usage from the simplest to the complex use cases.

AWS Budgets can be used to set reservation utilization or coverage targets allowing you to get alerts by email or SNS notification when the metrics reach the threshold.

Reservation Alerts feature is provided to:

- Amazon EC2
- Amazon RDS
- Amazon Redshift
- Amazon Elasticache
- Amaxon ElasticSearch

- ❑ AWS Budgets can be accessed from the AWS Management Console's service links and within the AWS Billing Console.
- ❑ Budgets API or CLI (command-line interface) can also be used to create, edit, delete and view up to 20,000 budgets per payer account.

Users can set up five alerts for each budget. But the most important are:

i. Alerts when current monthly costs exceed the budgeted amount.
ii. Alerts when current monthly costs exceed 80% of the budgeted amount.
iii. Alerts when forecasted monthly costs exceed the budgeted amount.

AWS Budgets can now be created monthly, quarterly, or annual budgets for the AWS resource usage or the AWS costs.

Types of Budgets:

- Cost budgets
- Usage budgets
- RI utilization budgets
- RI coverage budgets
- Savings Plans utilization budgets
- Savings Plans coverage budgets

AWS Cost & Usage Report

What is AWS Cost and Usage Report?

AWS Cost & Usage Report is a service that allows users to access the detailed set of AWS cost and usage data available, including metadata about AWS resources, pricing, Reserved Instances, and Savings Plans.

- ❑ AWS Cost & Usage Report is a part of AWS Cost Explorer.

AWS Cost and Usage Reports functions:

- ❑ It sends report files to your Amazon S3 bucket.
- ❑ It updates reports up to three times a day.

- ✔ For viewing, reports can be downloaded from the Amazon S3 console; for analyzing the report, Amazon Athena can be used, or upload the report into Amazon Redshift or Amazon QuickSight.
- ✔ Users with IAM permissions or IAM roles can access and view the reports.
- ✔ If a member account in an organization owns or creates a Cost and Usage Report, it can have access only to billing data when it has been a member of the Organization.
- ✔ If the master account of an AWS Organization wants to block access to the member accounts to set-up a Cost and Usage Report, Service Control Policy (SCP) can be used.

Reserved Instance Reporting

What is Reserved Instance Reporting?

Reserved Instance Reporting is a service used to summarize Reserved Instance (RIs) usage over a while.

Reserved Instance Reporting

RI Utilization reports can be visualized by exporting to both PDF and CSV formats.

RI coverage reports:

RI coverage report is used to visualize RI coverage and monitor against a RI coverage threshold.

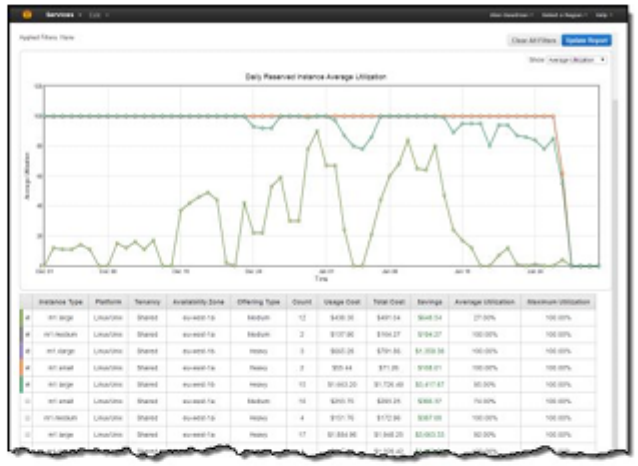
RI utilization reports:

RI utilization report is used to visualize daily RI utilization.

RI Coverage Report

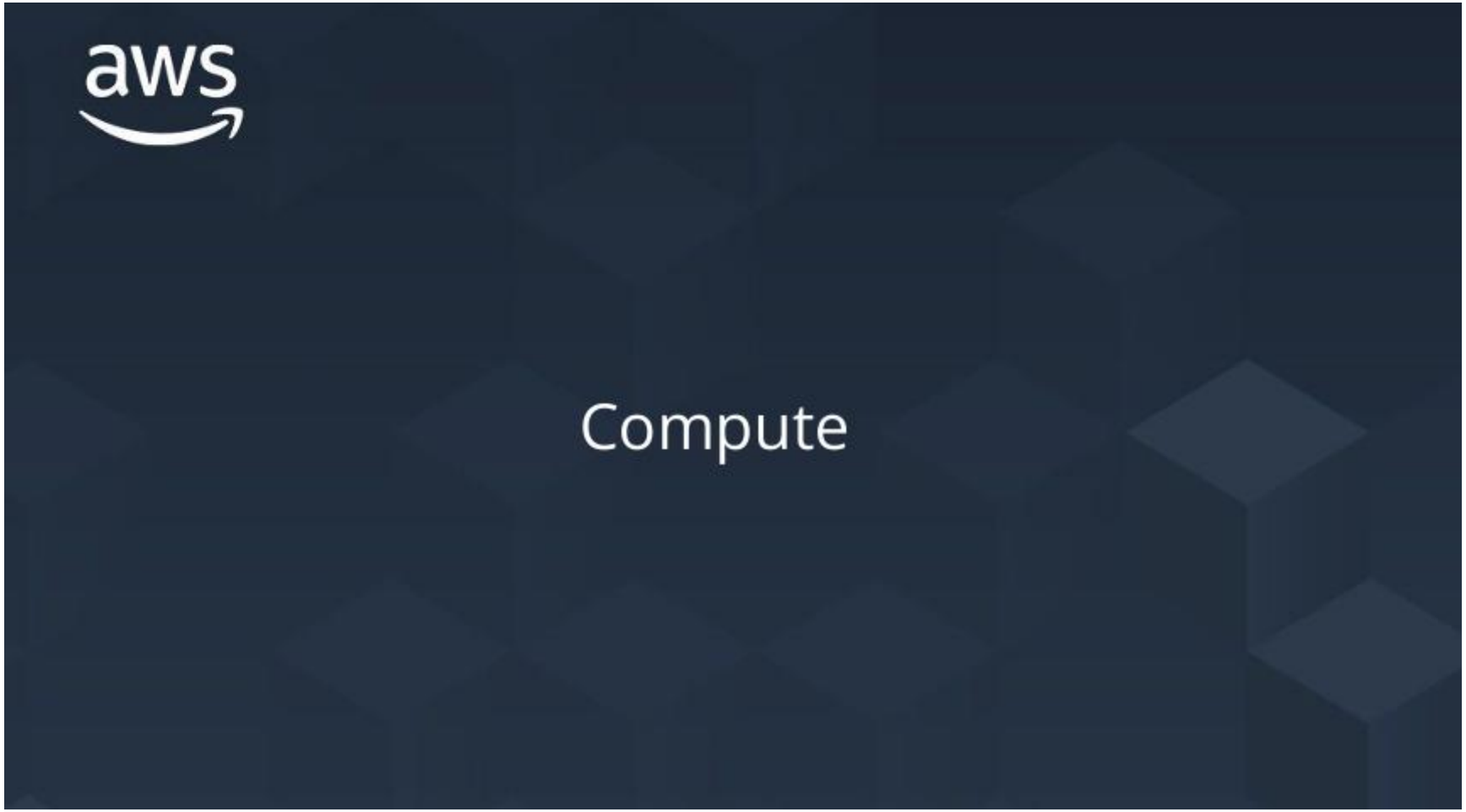
Target utilization (threshold utilization) of **RI utilization reports** are represented with the dotted line in the chart with different colored status:

- Red bar - RIs with no hours used.
- Yellow bar - Under the utilization target.
- Green bar - Reached utilization target.
- Gray bar - instances not using reservations.



RI Utilization Report

Along with AWS Cost Explorer, it increases cost savings as compared to On-Demand instance prices.


aws
Compute

Amazon EC2

What is Amazon EC2?

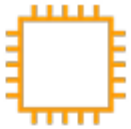
Amazon Elastic Compute Cloud (Amazon EC2) is a service that provides secure and scalable compute capacity in the AWS cloud. It falls under the category of Infrastructure as a Service (IAAS).

It provides the different type of instances based on the pricing models:

On-Demand Instances

- ✔ Useful for short-term needs, unpredictable workloads.
- ✔ No advance payment, no prior commitment.

Spot Instances



- ✔ No advance payment, no prior commitment.
- ✔ Useful for cost-sensitive compute workloads.

Reserved Instances

- ✔ Useful for long-running workloads and predictable usage.
- ✔ Offer to choose from No upfront, Partial upfront, or All upfront.

Dedicated Instances

- ✔ Instances run on hardware dedicated to a single user.
- ✔ Other customers can not share the hardware.

Dedicated Hosts

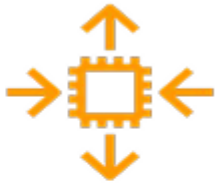
- ✔ A whole physical server with an EC2 instance allocates to an organization.

- It provides different compute platforms and instance types based on price, CPU, operating system, storage, and networking, and each instance type consists of one or more instance sizes. Eg., t2.micro, t4g.nano, m4.large, r5a.large, etc.
- It provides pre-configured templates that package the operating system and other software for the instances. This template is called Amazon Machine Images (AMIs).
- It helps to login into the instances using key-pairs, in which AWS manages the public key, and the user operates the private key.
- It also provides firewall-like security by specifying IP ranges, type, protocols (TCP), port range (22, 25, 443) using security groups.
- It provides temporary storage volumes known as instance store volumes, which are deleted if the instance gets stopped, hibernated, or terminated. It also offers non-temporary or persistent volumes known as Amazon EBS volumes.
- It enables users to write scripts under the option 'User data,' used at the instances' launch.
- It offers to choose from three IP addresses, which are Public IP address (Changes when the instance is stopped or refreshed), Private IP address (retained even if the model is stopped), Elastic IP address (static public IP address).

Amazon EC2 Auto Scaling

What is Amazon EC2 Auto Scaling?

Amazon EC2 Auto Scaling is a region-specific service used to maintain application availability and enables users to automatically add or remove EC2 instances according to the compute workloads.



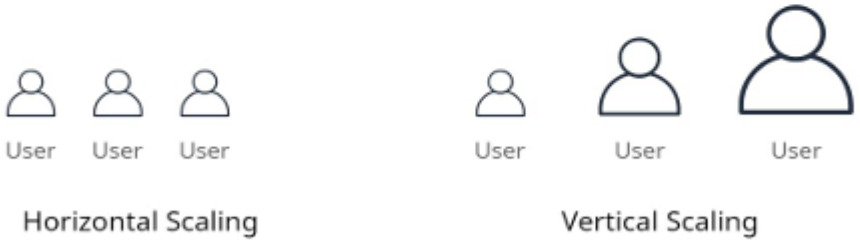
- The Auto Scaling group is a collection of the minimum number of EC2 used for high availability.
- It enables users to use Amazon EC2 Auto Scaling features such as fault tolerance, health check, scaling policies, and cost management.
- The scaling of the Auto Scaling group depends on the size of the desired capacity. It is not necessary to keep DesiredCapacity and MaxSize equal.

  E.g.,
  DesiredCapacity: '2' - There will be total 2 EC2 instances
  MinSize: '1'
  MaxSize: '2

- EC2 Auto Scaling supports automatic Horizontal Scaling (increases or decreases the number of EC2 instances) rather than Vertical Scaling (increases or decreases EC2 instances like large, small, medium).

Launch Configuration	Launch Template
A **launch configuration** is a configuration file used by an Auto Scaling group to launch EC2 instances	A **launch template** is similar to launch configuration with extra features as below
It launches any one of the Spot or On-Demand instances	It launches both Spot and On-Demand instances.
It specifies single instance types.	It specifies multiple instance types
It specifies one launch configuration at a time	It specifies multiple launch templates.

It scales across multiple Availability Zones within the same AWS region.



Amazon EC2 Auto Scaling

The ways to scale Auto Scaling Groups are as follows:

Manual Scaling
Update the desired capacity of the Auto Scaling Group manually.

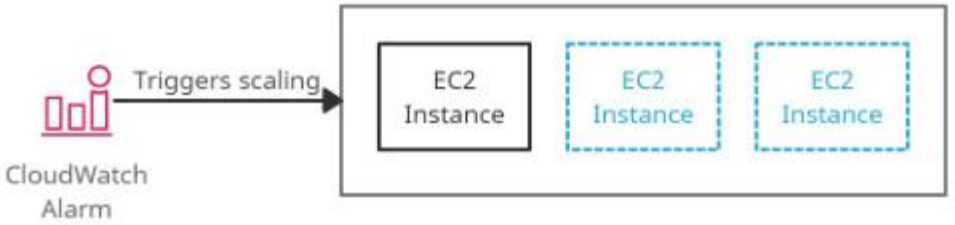
Scheduled Scaling
This scaling policy adds or removes instances based on predictable traffic patterns of the application.

Example:
Scale-out on every Tuesday or Scale in on every Saturday

Dynamic Scaling

- *Target tracking scaling policy*: This scaling policy adds or removes instances to keep the scaling metric close to the specified target value.
- *Simple Scaling Policy:* This scaling policy adds or removes instances when the scaling metric value exceeds the threshold value.
- *Step Scaling Policy:* This scaling policy adds or removes instances based on step adjustments (lower bound and upper bound of the metric value).

- The Cooldown period is the time during which an Auto Scaling group doesn't launch or terminate any instances before the previous scaling activity completes.



Amazon EC2 Auto Scaling using CloudWatch Alarm

AWS Batch

What is AWS Batch?

AWS Batch is a fully managed and regional batch processing service that allows developers, scientists, and engineers to execute large amounts of batch computing workloads on AWS.

It provides a correct amount of memory and can efficiently execute 100,000s of batch computing workloads across AWS compute services such as:

1. AWS Fargate
2. Amazon EC2
3. Spot Instances

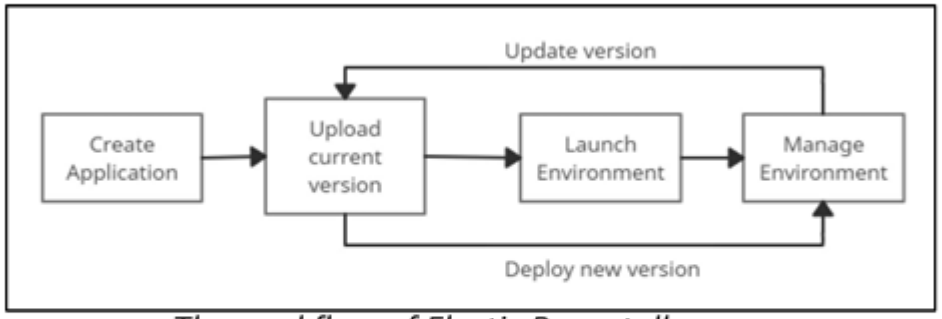
- It submits a job to a particular job queue and schedules them in a computing environment.
- A job is a work unit such as a shell script, a Linux executable, or a Docker container image.
- AWS Batch can be integrated with AWS data stores like Amazon S3 or Amazon DynamoDB to retrieve and write data securely.

AWS Elastic Beanstalk

What is AWS Elastic Beanstalk?

AWS Elastic Beanstalk is a service used to quickly deploy, scale, and manage applications in the AWS Cloud with automatic infrastructure management.

- It falls under the category of Platform as a Service (PaaS)
- It is also defined as a developer-centric view of deploying an application on AWS. The only responsibility of the developer is to write, and Elastic Beanstalk handles code and the infrastructure
- An Elastic Beanstalk application comprises components, including environments, versions, platforms, and environment configurations.



The workflow of Elastic Beanstalk

- Elastic Beanstalk console offers users to perform deployment and management tasks such as changing the size of Amazon EC2 instances, monitoring (metrics, events), and environment status.
- It supports web applications coded in popular languages and frameworks such as Java, .NET, Node.js, PHP, Ruby, Python, Go, and Docker.
- It uses Elastic Load Balancing and Auto Scaling to scale the application based on its specific needs automatically.

It provides multiple deployment policies such as:

- All at once, Rolling
- Rolling with an additional batch
- Immutable
- Traffic splitting

AWS CloudFormation vs. AWS Elastic Beanstalk

AWS CloudFormation	AWS Elastic Beanstalk
It deploys infrastructure using YAML/JSON template files. It can deploy Elastic Beanstalk environments.	It deploys applications on EC2. It cannot deploy Cloud Formation templates.

AWS Lambda

What is AWS Lambda?

AWS Lambda is a serverless computing service that allows users to run code as functions without provisioning or managing servers.

It helps to run the code on highly-available computing infrastructure and performs administrative tasks like server maintenance, logging, capacity provisioning, and automatic scaling and code monitoring.

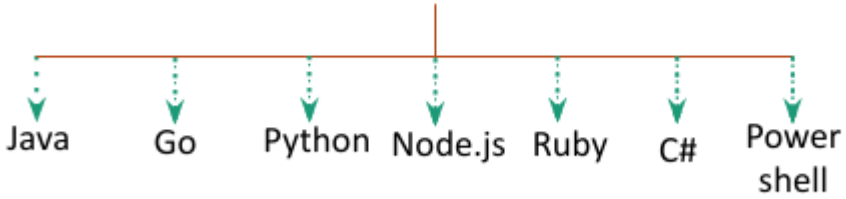
Using AWS Lambda, one can build serverless applications composed of Lambda functions triggered by events and can be automatically deployed using AWS CodePipeline and AWS CodeBuild.

Amazon EC2	Amazon Lambda
They are termed virtual servers in the AWS cloud.	They are termed virtual functions.
It is limited to instance types (RAM and CPU).	Limited by time (less execution time of 300 seconds).
It runs continuously.	It runs on demand.
Scaling computing resources is manual.	It has automated scaling.

✔ The memory allocated to AWS Lambda for computing is 128MB (minimum) to 3008MB (maximum). Additional memory can be requested in an increment of 64MB between 128MB - 3008MB.

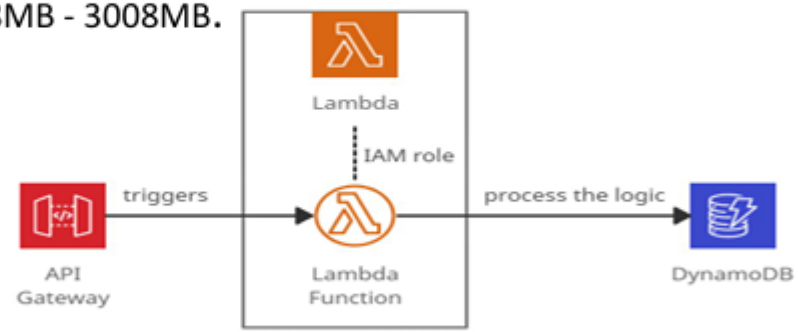
✔ The default execution time for AWS Lambda is 3 seconds, and the maximum is 15 minutes (900 seconds).

Lambda Functions supports the following languages:



Pricing details:

Charges are applied based on the number of requests for the functions and the time taken to execute the code



AWS Lambda

AWS Serverless Application Repository

What is AWS Serverless Application Repository?

AWS Serverless Application Repository is a managed repository used by developers and organizations to search, assemble, publish, deploy and store serverless architectures.

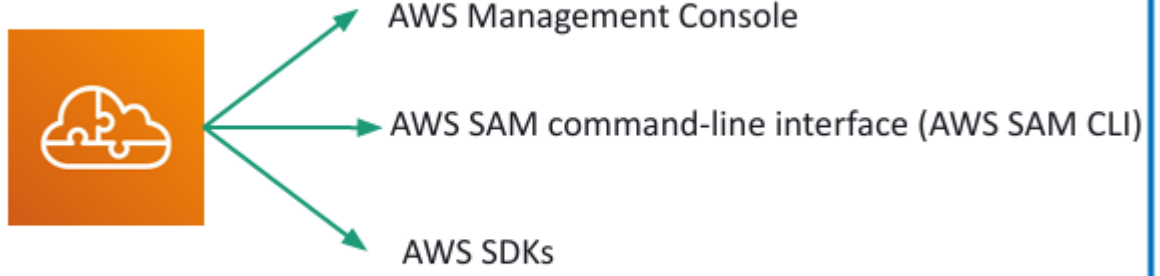
- It helps share reusable serverless application architectures and compose new serverless architectures using AWS Serverless Application Model (SAM) template.
- It uses pre-built applications in serverless deployments, eliminating the need to re-build and publish code to AWS.
- It discovers and offers best practices for serverless architectures to provide consistency within the organizations or provide permissions to share applications with specific AWS accounts.
- It integrates with AWS Lambda that allows developers of all levels to work with serverless computing by using re-usable architectures.

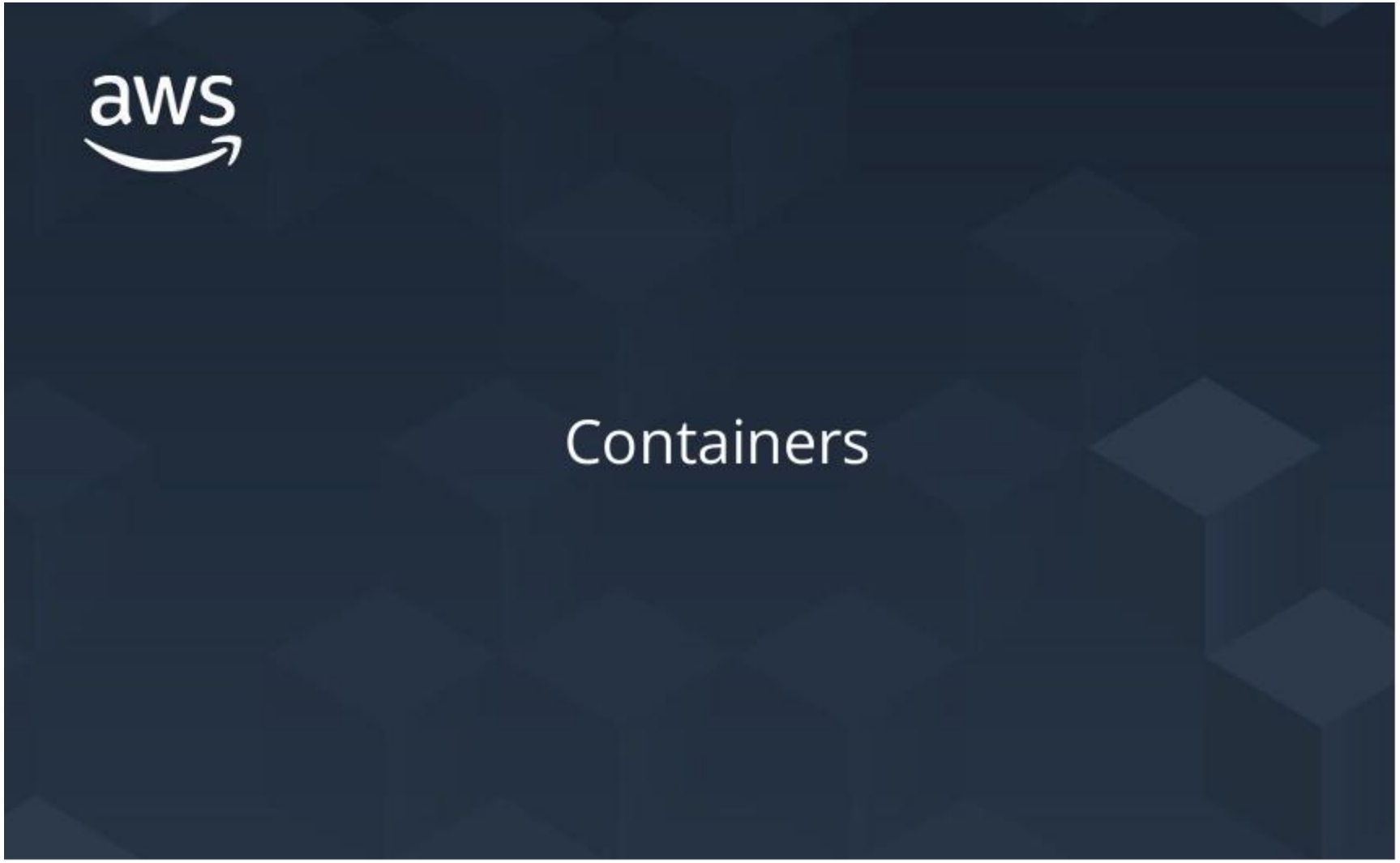
There are two ways to work with the AWS Serverless Application

Publishing Applications:
Upload and publish applications to be used by other developers.

Deploying Applications:
Search for applications with their required files and deploy them.

AWS Serverless Application Repository can be accessed in the following ways:




aws
Containers

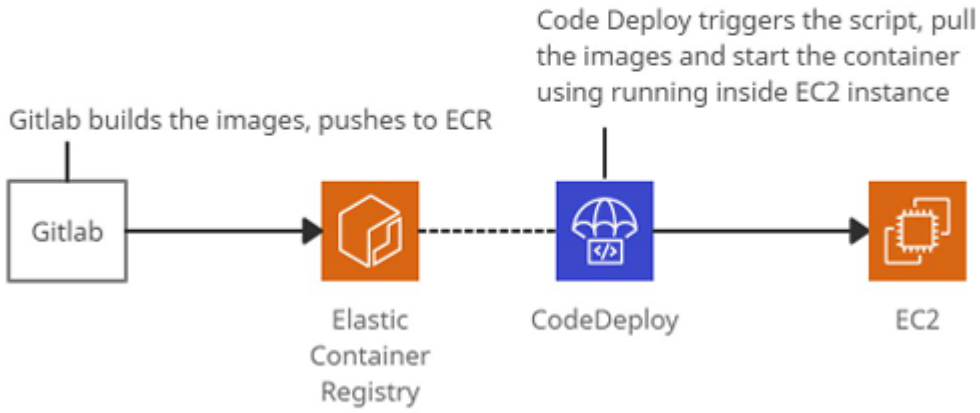
Amazon Elastic Container Registry

What is Amazon Elastic Container Registry?

Amazon Elastic Container Registry (ECR) is a managed service that allows users to store, manage, share, and deploy container images and other artifacts.

- It stores both the containers which are created, and any container software bought through AWS Marketplace.
- It is integrated with the following services:
  - Amazon Elastic Container Service(ECS)
  - Amazon Elastic Kubernetes Service(EKS)
  - AWS Lambda
  - Docker CLI
  - AWS Fargate for easy deployments

- AWS Identity and Access Management (IAM) enables resource-level control of each repository within ECR.
- Amazon Elastic Container Registry (ECR) supports public and private container image repositories. It allows sharing container applications privately within the organization or publicly for anyone to download.
- Images are encrypted at rest using Amazon S3 server-side encryption or using customer keys managed by AWS Key Management System (KMS).
- Amazon Elastic Container Registry (ECR) is integrated with continuous integration, continuous delivery, and third-party developer tools.
- Image scanning allows identifying vulnerabilities in the container images. It ensures that only scanned images are pushed to the repository



Amazon ECR example

Amazon Elastic Container Service

What is Amazon Elastic Container Service?

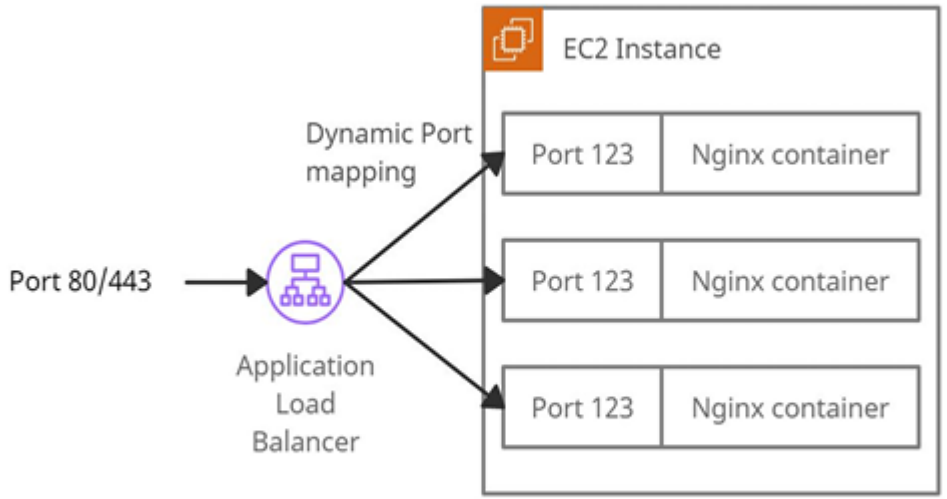
Amazon Elastic Container Service is a regional and docker-supported service that allows users to manage and scale containers on a cluster.

- ECS cluster is a combination of tasks or services executed on EC2 Instances or AWS Fargate.
- It offers to scale ECS clusters using Autoscaling based on CPU usage and other Autoscaling rules.
- Using Application Load Balancer, ECS enables dynamic port mapping and path-based routing.
- It provides Multi-AZ features for the ECS clusters.

Two main use cases of Amazon ECS are:

Microservices are built by the architectural method that decouples complex applications into smaller and independent services.

Batch Jobs - Batch jobs are short-lived packages that can be executed using containers.



Amazon ECS with Application Load Balancer

Amazon Elastic Kubernetes Service(EKS)

What is Amazon Elastic Kubernetes Service?

Amazon Elastic Kubernetes Service (Amazon EKS) is a service that enables users to manage Kubernetes applications in the AWS cloud or on-premises.

Amazon EKS

Using Amazon EKS, Kubernetes clusters and applications can be managed across hybrid environments without altering the code.

The EKS cluster consists of two components:

- ❑ Amazon EKS control plane
- ❑ Amazon EKS nodes

- The **Amazon EKS control plane** consists of nodes that run the Kubernetes software, such as etcd and the Kubernetes API server.
- To ensure high availability, Amazon EKS runs Kubernetes control plane instances across multiple Availability Zones.
- It automatically replaces unhealthy control plane instances and provides automated upgrades and patches for the new control planes.

- Users can execute batch workloads on the EKS cluster using the Kubernetes Jobs API across AWS compute services such as Amazon EC2, Fargate, and Spot Instances.
- The two methods for creating a new Kubernetes cluster with nodes in Amazon EKS:
  - **eksctl** - A command-line utility that consists of kubectl for creating/managing Kubernetes clusters on Amazon EKS.
  - AWS Management Console and AWS CLI

Amazon Elastic Kubernetes Service is integrated with many AWS services for unique capabilities:

- ❖ Images - Amazon ECR for container images
- ❖ Load distribution - AWS ELB (Elastic Load Balancing)
- ❖ Authentication - AWS IAM
- ❖ Isolation - Amazon VPC

AWS Fargate

What is AWS Fargate?

AWS Fargate is a serverless compute service used for containers by Amazon Elastic Container Service (ECS) and Amazon Elastic Kubernetes Service (EKS).

In the AWS Management Console, ECS clusters containing Fargate and EC2 tasks are displayed separately.

- It executes each task of Amazon ECS or pods of Amazon EKS in its kernel as an isolated computing environment and improves security.
- It packages the application in containers by just specifying the CPU and memory requirements with IAM policies. Fargate task does not share its underlying kernel, memory resources, CPU resources, or elastic network interface (ENI) with another task.
- It automatically scales the compute environment that matches the resource requirements for the container.

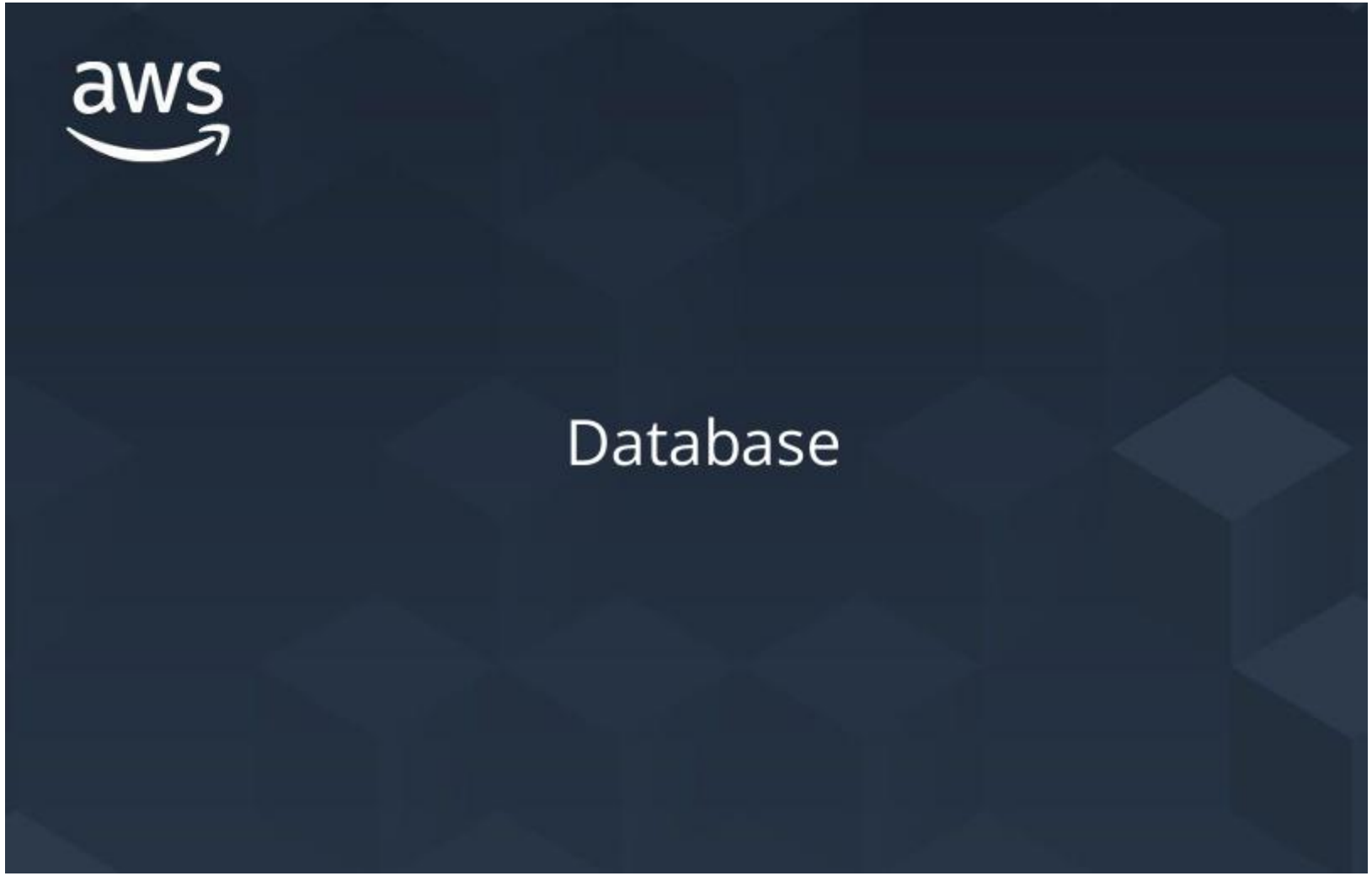
Security groups for pods in EKS cannot be used when pods are running on Fargate.

AWS Fargate

Storage Types supported for Fargate Tasks

- Amazon EFS volume for persistent storage
- Ephemeral Storage for non-persistent storage

Difference between EC2 instance and AWS Fargate


aws
Database

Amazon Aurora

What is Aurora?

Amazon Aurora is a MySQL and PostgreSQL-compatible, fully managed relational database engine built to enhance traditional enterprise databases' performance and availability.

- Is a part of the fully managed Amazon Relational Database Service (Amazon RDS).

Features include:

- RDS Management Console
- CLI commands and API operations for patching
- Backup
- Recovery
- Database Setup
- Failure Detection and repair

- Amazon Aurora replicates 2 copies of data in each availability zone (minimum of 3 AZ). So a total of 6 copies per region.

Data Replication : 2 Types

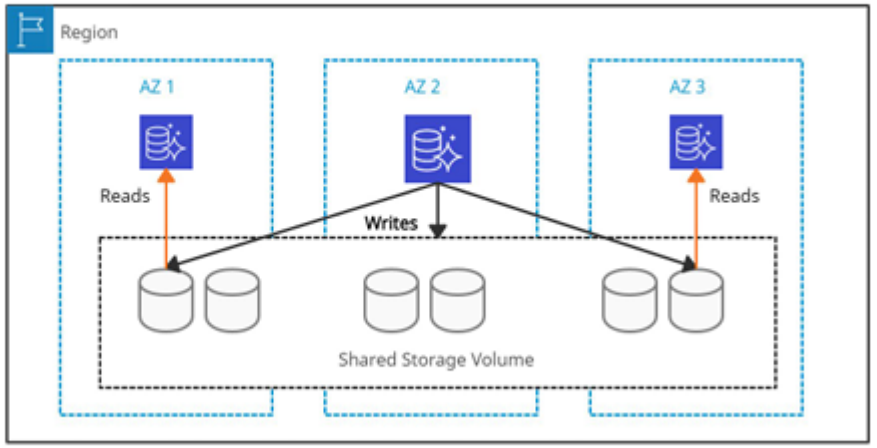
Aurora replica (in-region)

It can provide 15 read replicas.

MySQL Read Replica (cross-region)

It can provide 5 read replicas.

Amazon Aurora Cross-Region read replicas help to improve disaster recovery and provide fast reads in regions closer to the application users.



Performance

5x greater than MySQL

MySQL on RDS

3x greater than PostgreSQL

PostgreSQL on RDS

Amazon DocumentDB

What is Amazon DocumentDB?

Amazon DocumentDB is a fully managed NoSQL database service that manages MongoDB databases in AWS.

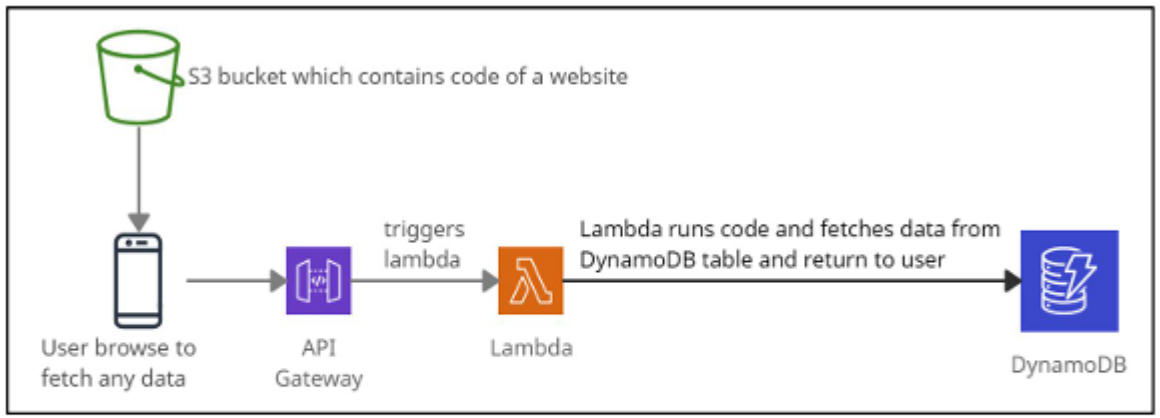
- It is a non-relational database service and supports document data structures.
- Using DocumentDB with Amazon CloudWatch helps to monitor the health and performance of the instances in a cluster.
- It works by building clusters that consist of 0 - 16 database instances (1 primary and 15 read replicas) and a cluster storage volume.

- It provides 99.99% availability by copying the cluster's data in three different Availability Zones.
- It helps to scale storage and compute services independently.
- It provides automatic failover either to one of up to 15 replicas created in other Availability Zones or to a new instance if no replicas have been provisioned.
- It provides backup capability and point-in-time recovery for the cluster. It has a backup retention period of up to 35 days.
- It is best suited for TTL and Timeseries Workloads and supports ACID properties based on transactions across one or more documents.

Amazon DynamoDB

What is Amazon DynamoDB?

Amazon DynamoDB is a serverless NoSQL database service that provides fast and predictable performance with single-digit millisecond latency.



Amazon DynamoDB example

- It provides a push button scaling feature, signifying that DB can scale without any downtime.
- It is a multi-region cloud service that supports key-value and document data structure.
- It provides high availability and data durability by replicating data synchronously on solid-state disks (SSDs) across 3 AZs in a region.
- It helps to store session states and supports ACID transactions for business-critical application
- It provides the on-demand backup capability of the tables for long-term retention and enables point-in-time recovery from accidental write or delete operations.
- Amazon DynamoDB Accelerator (DAX) is a highly available in-memory cache service that provides data from DynamoDB tables. DAX is not used for strongly consistent reads and write-intensive workloads.
- It supports Cross-Region Replication using DynamoDB Global Tables. Global Tables helps to deploy a multi-region database and provide automatic multi-master replication to AWS regions.

Amazon ElastiCache

What is Amazon ElastiCache?

ElastiCache is a web service used to manage and run in-memory data stores Redis and Memcached in the cloud.

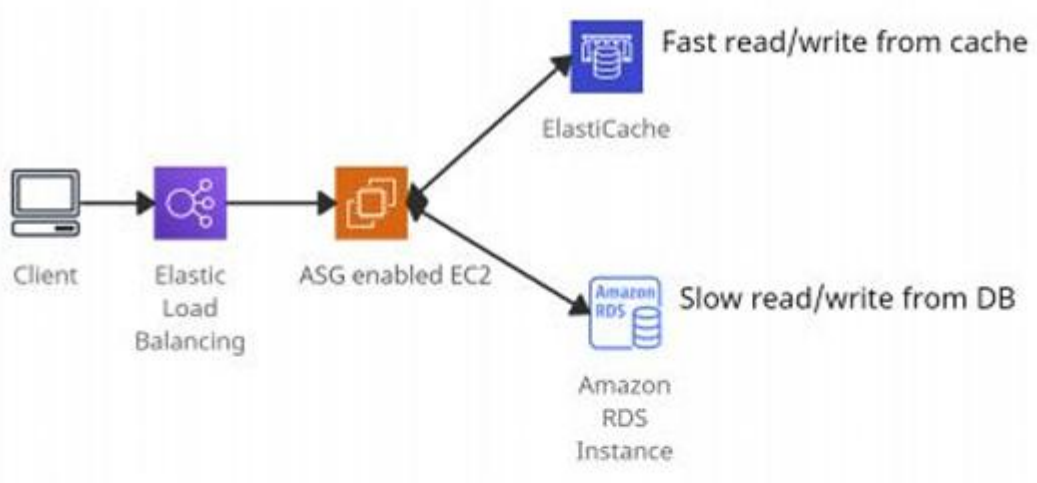
- It is best suited for Online Analytical Processing (OLAP) transaction workloads and for storing session states.
- It has in-memory caching features to provide sub-millisecond latency for read-heavy application workloads.

Amazon ElastiCache for Redis:

- It is useful for gaming applications, geospatial services, caching, session stores, and replication.
- Data is persistent.
- It is not multi-threaded.
- It supports Multi-AZ using read replicas.

Amazon ElastiCache for Memcached:

- It is useful for building applications that require caching layers.
- Data is not persistent.
- It supports multi-threading.
- It does not support Multi-AZ failover.
- It does not support snapshots.



Amazon ElastiCache

Amazon Keyspaces

What is Amazon Keyspaces?

Amazon Keyspaces (for Apache Cassandra) is a serverless service used to manage Apache Cassandra databases in AWS.

It provides the following throughput capacity modes for reads and writes:

On-demand
Charges are applied for the reads and write performed.

Provisioned
Charges are minimized by specifying the number of reads and writes per second in advance.

Using Amazon Keyspaces, tables can be scaled automatically, and read-write costs can be optimized by choosing either on-demand or provisioned capacity mode.

Functions of Keyspaces:

- It helps to run existing Cassandra workloads on AWS without making any changes to the Cassandra application code.
- It eliminates the developers' operational burden such as scaling, patching, updates, server maintenance, and provisioning.
- It offers high availability and durability by maintaining three copies of data in multiple Availability Zones.
- It implements the Apache Cassandra Query Language (CQL) API for using CQL and Cassandra drivers similar to Apache Cassandra.
- It helps to build applications that can serve thousands of requests with single-digit-millisecond response latency.
- It continuously backups hundreds of terabytes of table data and provides point-in-time recovery in the next 35 days.

Amazon Neptune

What is Amazon Neptune?

Amazon Neptune is a graph database service used as a web service to build and run applications that require connected datasets

- The graph database engine helps to store billions of connections and provides milliseconds latency for querying them.
- It offers to choose from graph models and languages for querying data.
- Property Graph (PG) model with Apache TinkerPop Gremlin graph traversal language.
- W3C standard Resource Description Framework (RDF) model with SPARQL Query Language.

Functions of Amazon Neptune:

- It is highly available across three AZs and automatically fails over any of the 15 low latency read replicas.
- It provides fault-tolerant storage by replicating two copies of data across three availability zones.
- It provides continuous backup to Amazon S3 and point-in-time recovery from storage failures.
- It automatically scales storage capacity and provides encryption at rest and in transit.

Amazon RDS

What is Amazon RDS?

Amazon Relational Database Service (Amazon RDS) is a service used to build and operate relational databases in the AWS Cloud

RDS provides read replicas of reading replicas and can also read replicas as a standby DB like Multi-AZ.

Read replicas feature is not available for SQL Server.

- It is best suited for structured data and Online Transaction Processing (OLTP) types of database workloads such as InnoDB.

It supports the following database engines:

- SQL Server
- PostgreSQL
- Amazon Aurora
- MYSQL
- MariaDB
- Oracle

- AWS KMS provides encryption at rest for RDS instances, DB snapshots, DB instance storage, and Read Replicas. The existing database cannot be encrypted.
- Amazon RDS only scales up for compute and storage, with no option for decreasing allocated storage
- It provides Multi-AZ and Read Replicas features for high availability, disaster recovery, and scaling.
  - **Multi-AZ Deployments** - Synchronous replication
  - **Read Replicas** - Asynchronous replication.

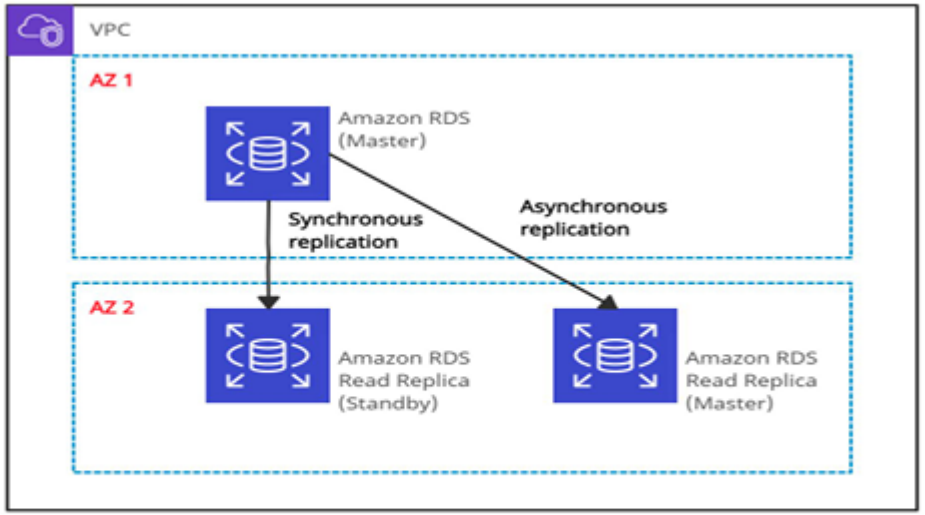
✔ If there is a need for unsupported RDS database engines, DB can be deployed on EC2 instances.

The following tasks need to be taken care of manually.

Encryption and Security

Updates and Backups

Disaster Recovery



Amazon RDS


aws
Developer Tools

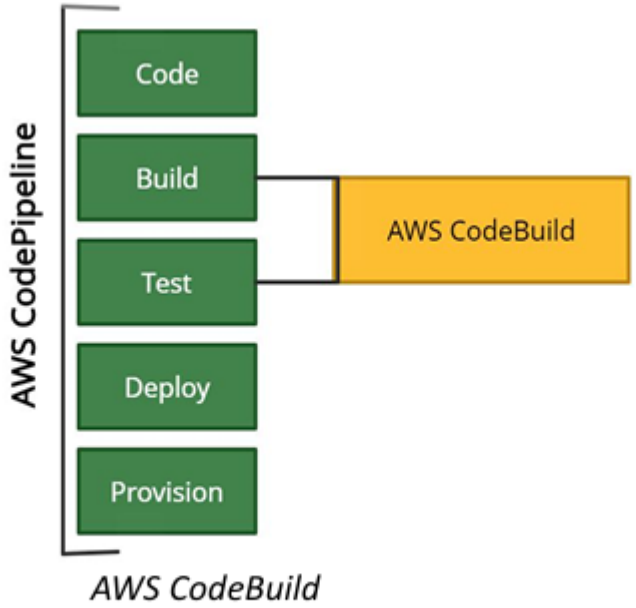
AWS CodeBuild

What is AWS CodeBuild?

AWS CodeBuild is a continuous integration service in the cloud used to compile source code, run tests, and build packages for deployment.

- AWS Code Services family consists of AWS CodeBuild, AWS CodeCommit, AWS CodeDeploy, and AWS CodePipeline that provide complete and automated continuous integration and delivery (CI/CD).
- It provides prepackaged and customized build environments for many programming languages and tools.
- It scales automatically to process multiple separate builds concurrently.
- It can be used as a build or test stage of a pipeline in AWS CodePipeline.
- It requires VPC ID, VPC subnet IDs, and VPC security group IDs to access resources in a VPC to perform build or test.

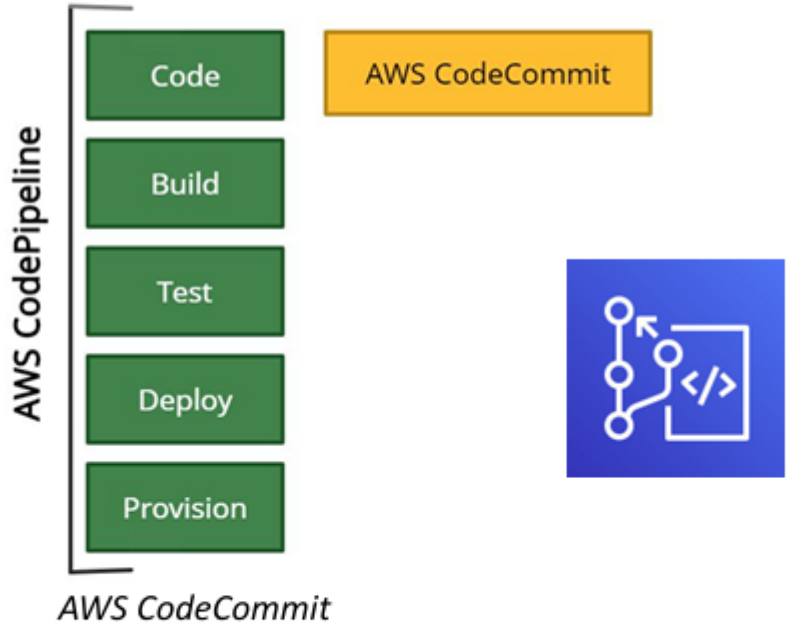
- Charges are applied based on the amount of time taken by AWS CodeBuild to complete the build.
- The following ways are used to run CodeBuild:
  - AWS CodeBuild
  - AWS CodePipeline console
  - AWS Command Line Interface (AWS CLI)
  - AWS SDKs



AWS CodeCommit

What is AWS CodeCommit?

AWS CodeCommit is a managed source control service used to store and manage private repositories in the AWS cloud, such as Git.



Functions of AWS CodeCommit:

- It works with existing Git-based repositories, tools, and commands in addition to AWS CLI commands and APIs.
- It provides high availability, durability, and redundancy.
- It eliminates the need to back up and scale the source control servers.
- CodeCommit repositories support pull requests, version differencing, merge requests between branches, and notifications through emails about any code changes.
- As compared to Amazon S3 versioning of individual files, AWS CodeCommit support tracking batched changes across multiple files.
- It provides encryption at rest and in transit for the files in the repositories.

AWS CodeDeploy

What is AWS CodeDeploy?

AWS CodeDeploy is a service that helps to automate application deployments to a variety of compute services such as Amazon EC2, AWS Fargate, AWS ECS, and on-premises instances.

AWS CodeDeploy

- ❑ Using Amazon EKS, Kubernetes clusters and applications can be managed across hybrid environments without altering the code.
- ❑ It can fetch the content for deployment from Amazon S3 buckets, Bitbucket, or GitHub repositories.
- ❑ It can deploy different types of application content such as Code, Lambda functions, configuration files, scripts and even Multimedia files.
- ❑ It can scale with the infrastructure to deploy on multiple instances across development, test, and production environments.
- ❑ It can integrate with existing continuous delivery workflows such as AWS CodePipeline, GitHub, Jenkins.

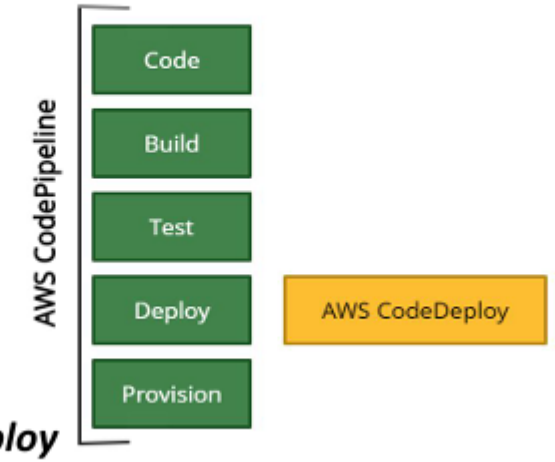
It provides the following deployment type to choose from:

In-place deployment:

- All the instances in the deployment group are stopped, updated with new revision and started again after the deployment is complete.
- Useful for EC2/On-premises compute platform.

Blue/green deployment:

- The instances in the deployment group of the original environment are replaced by a new set of instances of the replacement environment.
- Using Elastic Load Balancer, traffic gets rerouted from the original environment to the replacement environment and instances of the original environment get terminated after the deployment is complete.
- Useful for EC2/On-Premises, AWS Lambda and Amazon ECS compute platform



AWS CodeDeploy

AWS X-Ray

What is AWS X-Ray?

AWS X-Ray is a service that allows visual analysis or allows to trace microservices based applications.

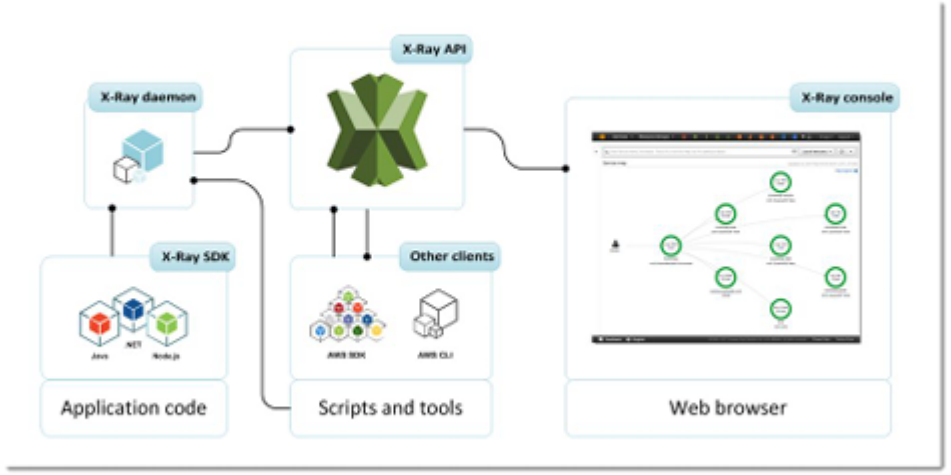
- ✔ It provides end-to-end information about the request, response and calls made to other AWS resources by travelling through the application's underlying components consisting of multiple microservices.
- ✔ It creates a service graph by using trace data from the AWS resources.
  - The graph shows the information about front-end and backend services calls to process requests and continue the flow of data.
  - The graph helps to troubleshoot issues and improve the performance of the applications.

It works with the following AWS services:

- AWS EC2 (Applications deployed on Instances)
- AWS Elastic Load Balancer
- AWS Elastic BeanStalk
- AWS Lambda
- Amazon ECS (Elastic Container Service)
- Amazon API Gateway

The X-Ray SDKs are available for the following languages:

- Go
- Java
- Node.js
- Python
- Ruby
- .Net



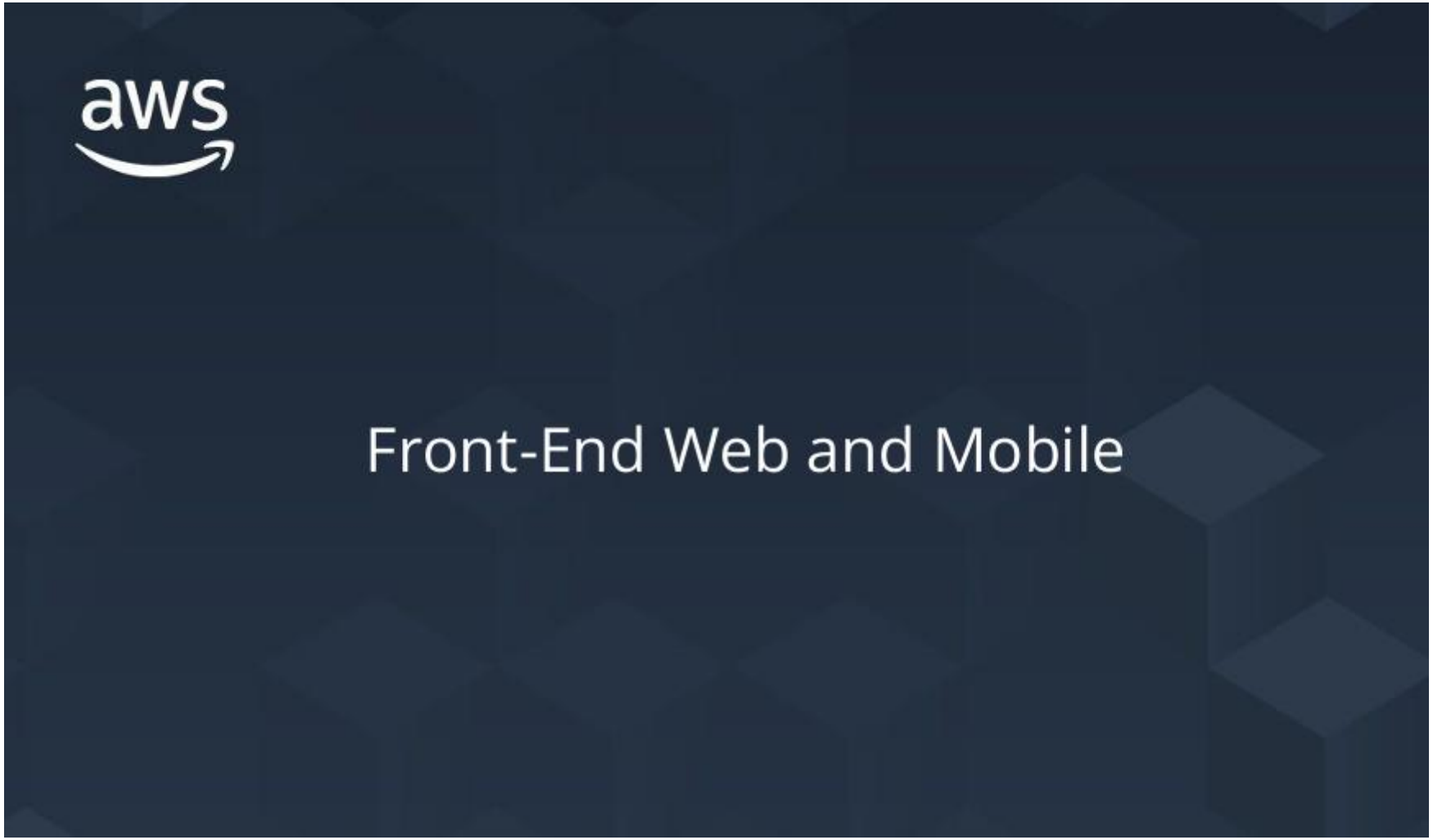
Amazon WorkSpaces

What is Amazon WorkSpaces?

Amazon WorkSpaces is a managed service used to provision virtual Windows or Linux desktops for users across the globe.

- It helps to eliminate the management of on-premise VDIs (Virtual Desktop Infrastructure).
- It offers to choose PCoIP protocols (port 4172) or WorkSpaces Streaming Protocol (WSP, port 4195) based on user's requirements such as the type of devices used for workspaces, operating system, and network conditions.
- Amazon WorkSpaces Application Manager (Amazon WAM) helps to manage the applications on Windows WorkSpaces.
- Multi-factor authentication (MFA) and AWS Key Management Service (AWS KMS) is used for account and data security.
- Each WorkSpace is connected to a virtual private cloud (VPC) with two elastic network interfaces (ENI) and AWS Directory Service.

- Amazon WorkSpaces can be accessed with the following client application for a specific device:
  - ✔ Android devices, iPads
  - ✔ Windows, macOS, and Ubuntu Linux computers
  - ✔ Chromebooks
  - ✔ Teradici zero client devices -supported only with PCoIP
- For Amazon WorkSpaces, billing takes place either monthly or hourly.


aws
Front-End Web and Mobile

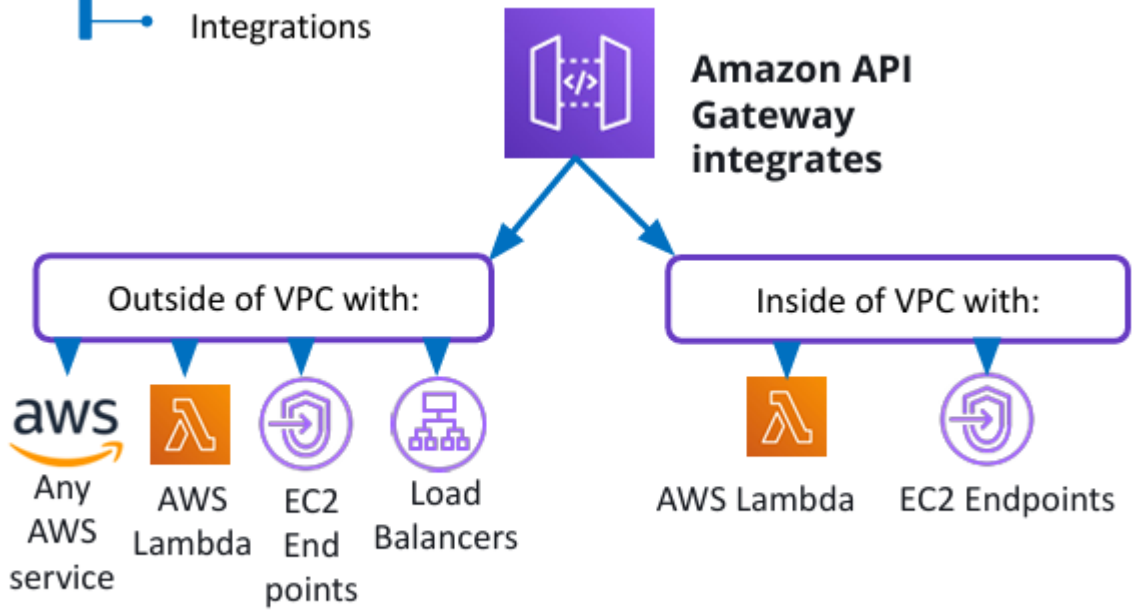
Amazon API Gateway

What is Amazon API Gateway?

Amazon API Gateway is a service that maintains and secures APIs at any scale. It is categorized as a serverless service of AWS.

API Gateway consists of:

- Stages
- Resources
- Methods
- Integrations

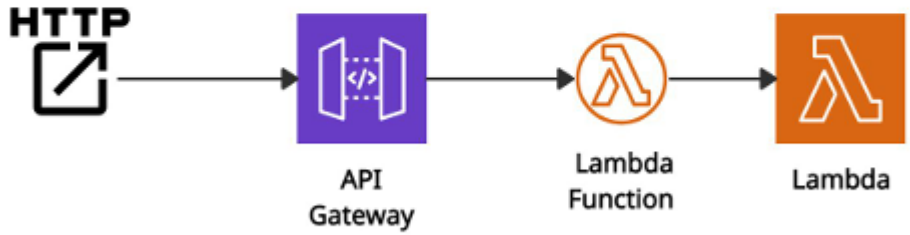


Amazon API Gateway:

- ✔ Acts as a front door for real-world applications to access data, business logic from the back-end services, such as code running on AWS Lambda, or any web application.
- ✔ Handles the processing of hundreds of thousands of existing API calls, including authorization, access control, different environments (dev, test, production), and API version management.
- ✔ Helps to create web APIs that route HTTP requests to Lambda functions

Example:

When a request is sent through a browser or HTTP client to the public endpoint, API Gateway API broadcasts the request and sends it to the Lambda function. The Function calls the Lambda API to get the required data and returns it to the API.



AWS Lambda + API Gateway = No need to manage infrastructure


aws
Internet of Things

AWS IoT Core

What is AWS IoT Core?

AWS IoT Core is a cloud service that enables users to connect IoT devices (wireless devices, sensors, and smart appliances) to the AWS cloud without managing servers.

- It supports devices and clients that use the following protocol:

AWS IoT Core

MQTT (Message Queuing and Telemetry Transport) - publish and subscribe messages

MQTT over WSS protocols - publish and subscribe messages

HTTPS protocol - publish messages

- It provides secure and bi-directional communication with all the devices, even when they aren't connected.
- It consists of a device gateway and a message broker that helps connect and process messages and routes those messages to other devices or AWS endpoints.
- It helps developers to operate wireless LoRaWAN (low-power long-range Wide Area Network) devices.
- It helps to create a persistent Device Shadow (a virtual version of devices) so that other applications or devices can interact.

It integrates with Amazon services like Amazon CloudWatch, AWS CloudTrail, Amazon S3, Amazon DynamoDB, AWS Lambda, Amazon Kinesis, Amazon SageMaker, and Amazon QuickSight to build IoT applications.

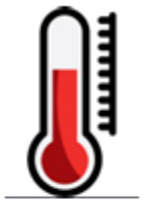
AWS IoT Events

What is AWS IoT Events?

AWS IoT Events is a monitoring service that allows users to monitor and respond to devise fleets' events in IoT applications.

It builds event monitoring applications in the AWS Cloud that can be accessed through the AWS IoT Events console.

It detects events from IoT sensors such as temperature, motor voltage, motion detectors, humidity.



It helps to create event logic using conditional statements and trigger alerts when an event occurs.

AWS IoT Events accepts data from many IoT sources like sensor devices, AWS IoT Core, and AWS IoT Analytics.

AWS IoT Greengrass

What is AWS IoT Greengrass?

AWS IoT Greengrass is a cloud service that groups, deploys, and manages software for all devices at once and enables edge devices to communicate securely.

- It is used on multiple IoT devices in homes, vehicles, factories, and businesses.
- It provides a pub/sub message manager that stores messages as a buffer to preserve them in the cloud

It synchronizes data on the device using the following AWS services:

- Amazon Simple Storage Service (Amazon S3)
- Amazon Kinesis
- AWS IoT Core
- AWS IoT Analytics

- The Greengrass Core is a device that enables the communication between AWS IoT Core and the AWS IoT Greengrass.
- Devices with IoT Greengrass can process data streams without being online.
- It provides different programming languages, open-source software, and development environments to develop and test IoT applications on specific hardware.
- It provides encryption and authentication for device data for cloud communications.
- It provides AWS Lambda functions and Docker containers as an environment for code execution.

FreeRTOS

What is FreeRTOS?

FreeRTOS is an open-source operating system for microcontrollers that enables devices to connect, manage, program, deploy and scale.

It helps securely connect small devices to AWS IoT Core or the devices running AWS IoT Greengrass.

The microcontroller is a kind of processor available in many devices like industrial automation, automobiles, sensors, appliances.

It acts as a multitasking scheduler and provides multiple memory allocation options, semaphore, task notifications, message queues, and message buffers.


aws
Machine Learning

Amazon SageMaker

What is Amazon SageMaker?

Amazon SageMaker is a cloud service that allows developers to prepare, build, train, deploy and manage machine learning models.

- It scales up to petabytes level to train models and manages all the underlying infrastructure.
- Amazon SageMaker notebook instances are created using Jupyter notebooks to write code to train and validate the models.
- Amazon SageMaker gets billed in seconds based on the amount of time required to build, train, and deploy machine learning models.

- It provides a secure and scalable environment to deploy a model using SageMaker Studio or the SageMaker console.
- It has pre-installed machine learning algorithms to optimize and deliver 10X performance.

Amazon Polly

What is Amazon Polly?

Amazon Polly is a cloud service used to convert text into speech.

It supports many different languages, and Neural Text-to-Speech (NTTS) voices to create speech-enabled applications.

It requires no setup costs, only pay for the text converted.

It offers caching and replays of Amazon Polly's generated speech in a format like MP3.

Amazon Transcribe

What is Amazon Transcribe?

Amazon Transcribe is a service used to convert audio (speech) to text using a Deep Learning process known as automatic speech recognition (ASR).

Amazon Transcribe Medical is used to convert medical speech to text for clinical documentation.

It is best suited for customer service calls, live broadcasts, and media subtitling.

It automatically matches the text quality similar to the manual transcription. For transcribe, charges are applied based on the seconds of speech converted per month.


aws
Management and Governance

Amazon CloudWatch

What is Amazon CloudWatch?

Amazon CloudWatch is a service that monitors based on multiple metrics of AWS and on-premises resources.

Amazon CloudWatch

AWS CloudWatch monitors AWS resources such as Amazon RDS DB instances, Amazon EC2 instances, Amazon DynamoDB tables, and any log files generated by the applications.

- Collects and correlates monitoring data in logs, metrics, and events from AWS resources, applications, and services that run on AWS and on-premises servers.
- Offers dashboards and creates graphs to visualize cloud resources.
- Visualizes logs to address issues and improve performance by performing queries.

- Alarms can be created using CloudWatch Alarms that monitors metrics and send notifications.
- CloudWatch Agent or API can be used to monitor hybrid cloud architectures.
- CloudWatch Container Insights and Lambda Insights both provide dashboards to summarize the performance and errors for a selected time window.

Amazon CloudWatch is used alongside the following applications:

- Amazon Simple Notification Service (Amazon SNS)
- Amazon EC2 Auto Scaling
- AWS CloudTrail
- AWS Identity and Access Management (IAM)

AWS CloudFormation

What is AWS CloudFormation?

AWS CloudFormation is a service that collects AWS and third-party resources and manages them throughout their lifecycles by launching them together as a stack.

AWS CloudFormation

Template:

- A **template** is used to create, update, and delete an entire stack as a single unit without managing resources individually.
- CloudFormation provides the capability to reuse the template to set the resources easily and repeatedly.

Stacks:

- **Stacks** can be created using the AWS CloudFormation console and AWS Command Line Interface (CLI).
- **Nested Stacks** are stacks created within another stack by using the 'AWS::CloudFormation::Stack' resource attribute.
- The main stack is termed as parent stack, and other belonging stacks are termed as child stack, which can be implemented by using ref variable '! Ref'.

AWS does not charge for using AWS CloudFormation, and charges are applied for the CloudFormation template services.

Example: CloudFormation template for creating EC2 instance

```
EC2Instance:
    Type: AWS::EC2::Instance
    Properties:
        ImageId: 1234xyz
        KeyName: aws-keypair
        InstanceType: t2.micro
        SecurityGroups:
            - !Ref EC2SecurityGroup
        BlockDeviceMappings:
            - DeviceName: /dev/sda1
              Ebs:
                  VolumeSize: 50
```

AWS CloudTrail

What is AWS CloudTrail?

AWS CloudTrail is a service that gets enabled when the AWS account is created and is used to enable compliance and auditing of the AWS account.

- ✔ It offers to view, analyze, and respond to activity across the AWS infrastructure.
- ✔ It records actions as an event by an IAM user, role, or an AWS service.
- ✔ CloudTrail records can download Cloud Trial events in JSON or CSV file.
- ✔ **CloudWatch** monitors and manages the activity of AWS services and resources, reporting on their health and performance. Whereas **CloudTrail** resembles logs of all actions performed inside the AWS environment.
- ✔ **IAM log file -**

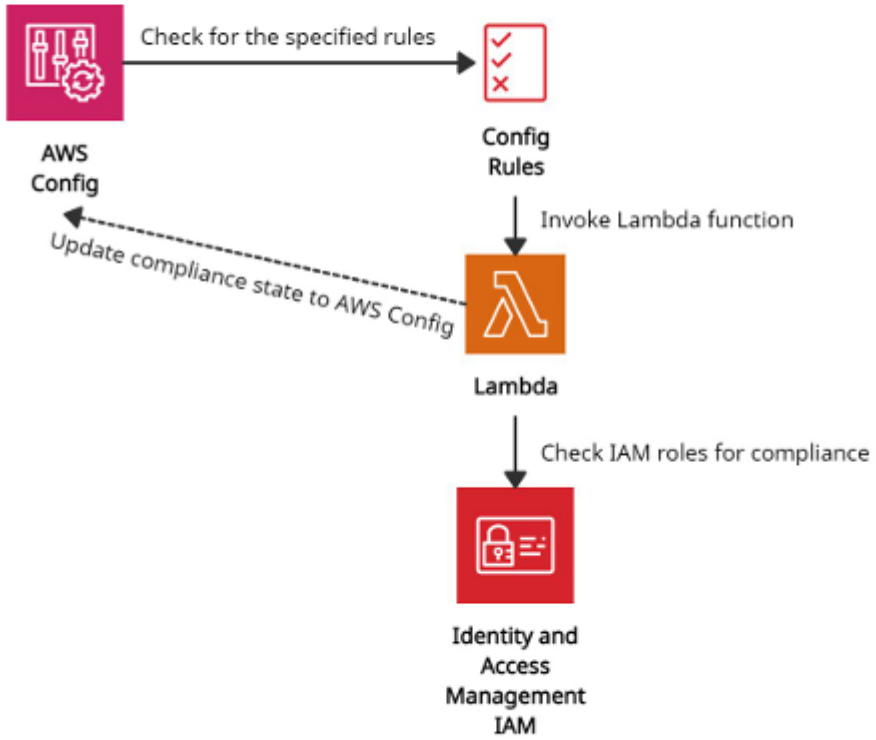
  The below example shows that the IAM user Rohit used the AWS Management Console to call the AddUserToGroup action to add Nayan to the administrator group.


```
Records": [{
    "eventVersion": "1.0",
    "userIdentity": {
        "type": "IAMUser",
        "principalId": "PR_ID",
        "arn":
"arn:aws:iam::210123456789:user/Rohit",
        "accountId": "210123456789",
        "accessKeyId": "KEY_ID",
        "userName": "Rohit"
    },
    "eventTime": "2021-01-24T21:18:50Z",
    "eventSource": "iam.amazonaws.com",
    "eventName": "CreateUser",
    "awsRegion": "ap-south-2",
    "sourceIPAddress": "176.1.0.1",
    "userAgent": "aws-cli/1.3.2 Python/2.7.5
Windows/7",
    "requestParameters": {"userName": "Nayan"},
    "responseElements": {"user": {
        "createDate": "Jan 24, 2021 9:18:50 PM",
        "userName": "Nayan",
        "arn": "arn:aws:iam::128x:user/Nayan",
        "path": "/",
        "userId": "12xyz"
    }}
}]}
```

AWS Config

What is AWS Config?

AWS Config is a service that allows users to determine the quality of a resource's configuration in the AWS account.



AWS Config in action

Functions of AWS Config:

- It helps to monitor configuration changes performed over a specific period using AWS Config console and AWS CLI and generates notifications about changes.
- It offers a dashboard to view compliance status for an account across regions.
- It uses Config rules to evaluate configuration settings of the AWS resources.
- It captures the history of configurations and tracks relationships of resources before making changes.
- Using AWS CloudTrail, AWS Config helps to identify and troubleshoot issues by capturing API calls as events.

AWS License Manager

What is AWS License Manager?

AWS License Manager is a service used to centralize the usage of software licenses across the environment.

- It supports Bring-Your-Own-License (BYOL) feature, which means that users can manage their existing licenses for third-party workloads (Microsoft Windows Server, SQL Server) to AWS.
- It enables administrators to create customized licensing rules that help prevent licensing violations (using more licenses than the agreement).
- It provides a dashboard to control the visibility of all the licenses to the administrators.

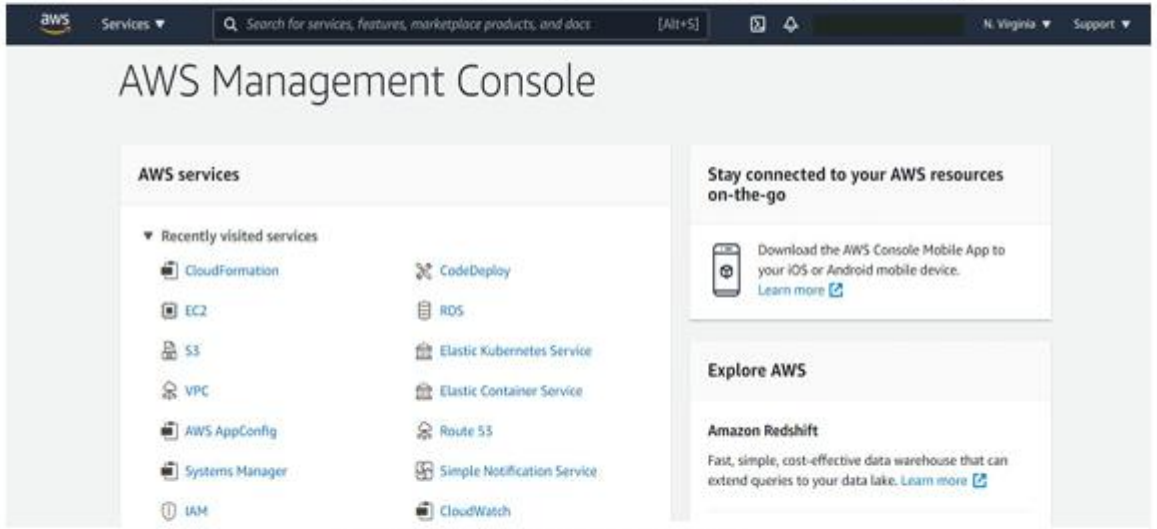
- It allows administrators to specify Dedicated Host management preferences for allocation and capacity utilization.
- AWS License Manager's managed entitlements provide built-in controls to software vendors (ISVs) and administrators so that they can assign licenses to approved users and workloads.
- AWS Systems Manager can manage licenses on physical or virtual servers hosted outside of AWS using AWS License Manager.
- AWS Organizations and AWS License Manager help to allow cross-account disclosure of computing resources in the organization.

AWS Management Console

What is AWS Management Console?

AWS Management Console is a web console with multiple wizards and services used to manage Amazon Web Services.

- It can be visible when a user first-time signs in. It provides access to other service consoles and a user interface for exploring AWS.

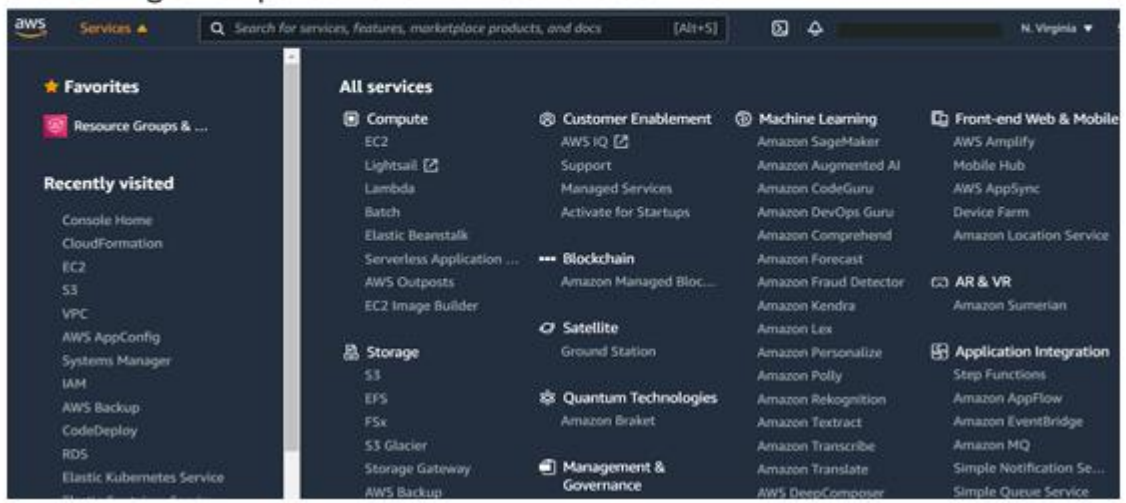


AWS Management Console

- AWS Management Console provides a Services option on the navigation bar that allows choosing services from the **Recently visited** list or the **All services** list.

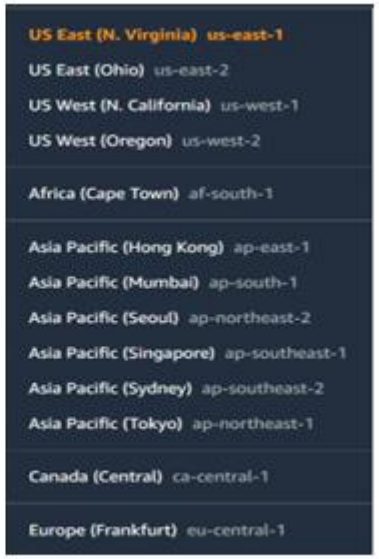
- A GUI Console is available as an app for Android and iOS for a better experience.

- There is a **Search** box on the navigation bar to search for AWS services by entering all or part of the name of the service



AWS Services Console

- On the navigation bar, there is an option to select **Regions** from.



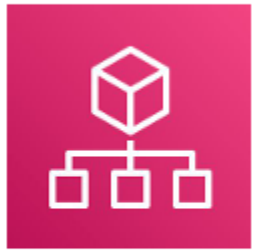
AWS Regions

AWS Organizations

What are AWS Organizations?

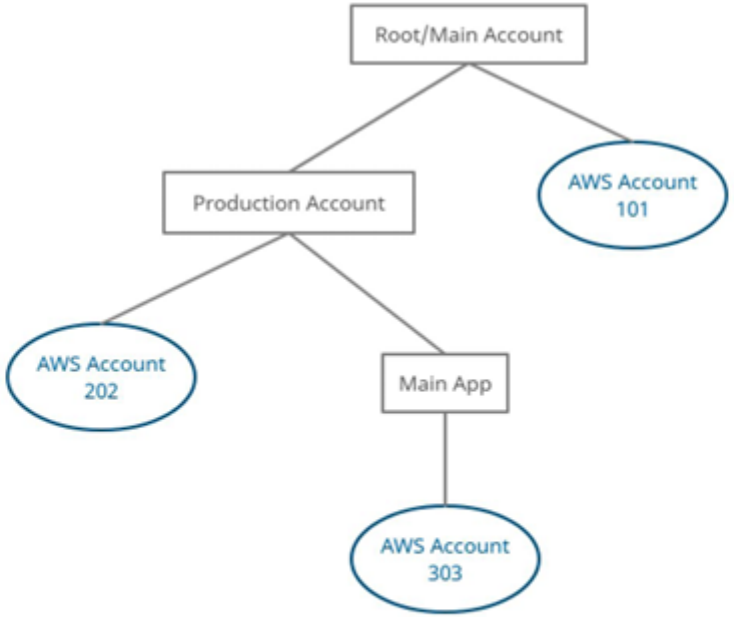
AWS Organizations is a service that allows users to manage multiple AWS accounts grouped into a single organization.

It includes account management and combined billing capabilities to meet the business's budgetary and security needs.



AWS Organizations

- It easily shares critical common resources across the accounts.
- It organizes accounts into organizational units (OUs), which are groups of accounts that serve specified applications.



Service Control Policies

Service Control Policies (SCPs) can be created to provide governance boundaries for the OUs. SCPs ensure that users in the accounts only perform actions that meet security requirements.

The master account is responsible for paying charges of all resources used by the accounts in the organization.

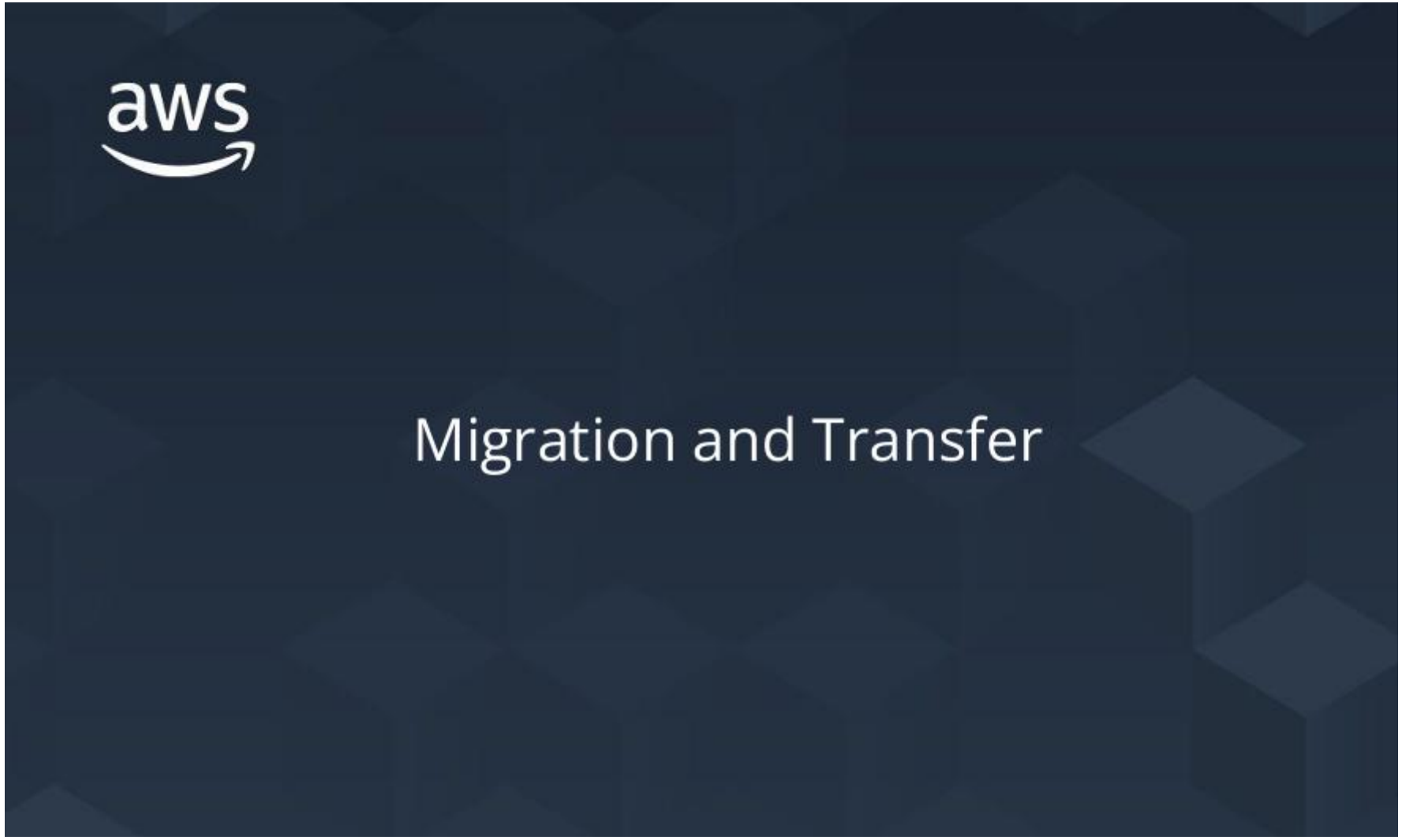
AWS Systems Manger

What is AWS Systems Manager?

AWS Systems Manager (SSM) is a service that allows users to centralize or group operational data using multiple services and automate operations across AWS infrastructure.

- ✔ It simplifies maintenance and identifies issues in the resources that may impact the applications.
- ✔ It displays the operational data, system and application configurations, software installations, and other details on a single dashboard known as AWS Systems Manager Explorer.
- ✔ It manages secrets and configuration data and separates them from code using a centralized store known as Parameter Store.
- ✔ It helps to communicate with the Systems Manager agent installed on AWS servers and in an on-premises environment. Agents are installed to manage resources on servers using different operating systems.

- It helps to manage servers without actually logging into the server using a web console known as Session Manager.
- It helps to automate repetitive operations and management tasks using predefined playbooks.
- It connects with Jira Service Desk and ServiceNow to allow ITSM platform users to manage AWS resources.
- Systems Manager Distributor helps to distribute software packages on hosts along with versioning.


aws
Migration and Transfer

AWS Database Migration Service

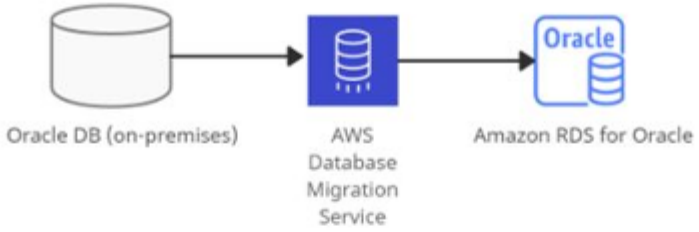
What is AWS Database Migration Service?

AWS Database Migration Service is a cloud service used to migrate relational databases from on-premises, Amazon EC2, or Amazon RDS to AWS securely.

Amazon Database Management Service

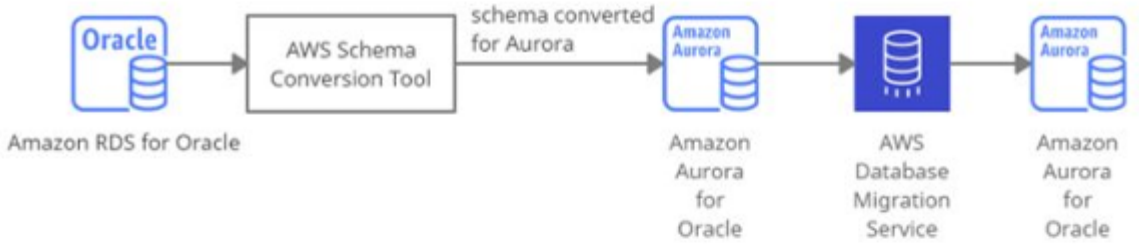
AWS DMS does not stop the running application while performing the migration of databases, resulting in downtime minimization.

Homogeneous migration



Homogeneous migration

Heterogeneous migration

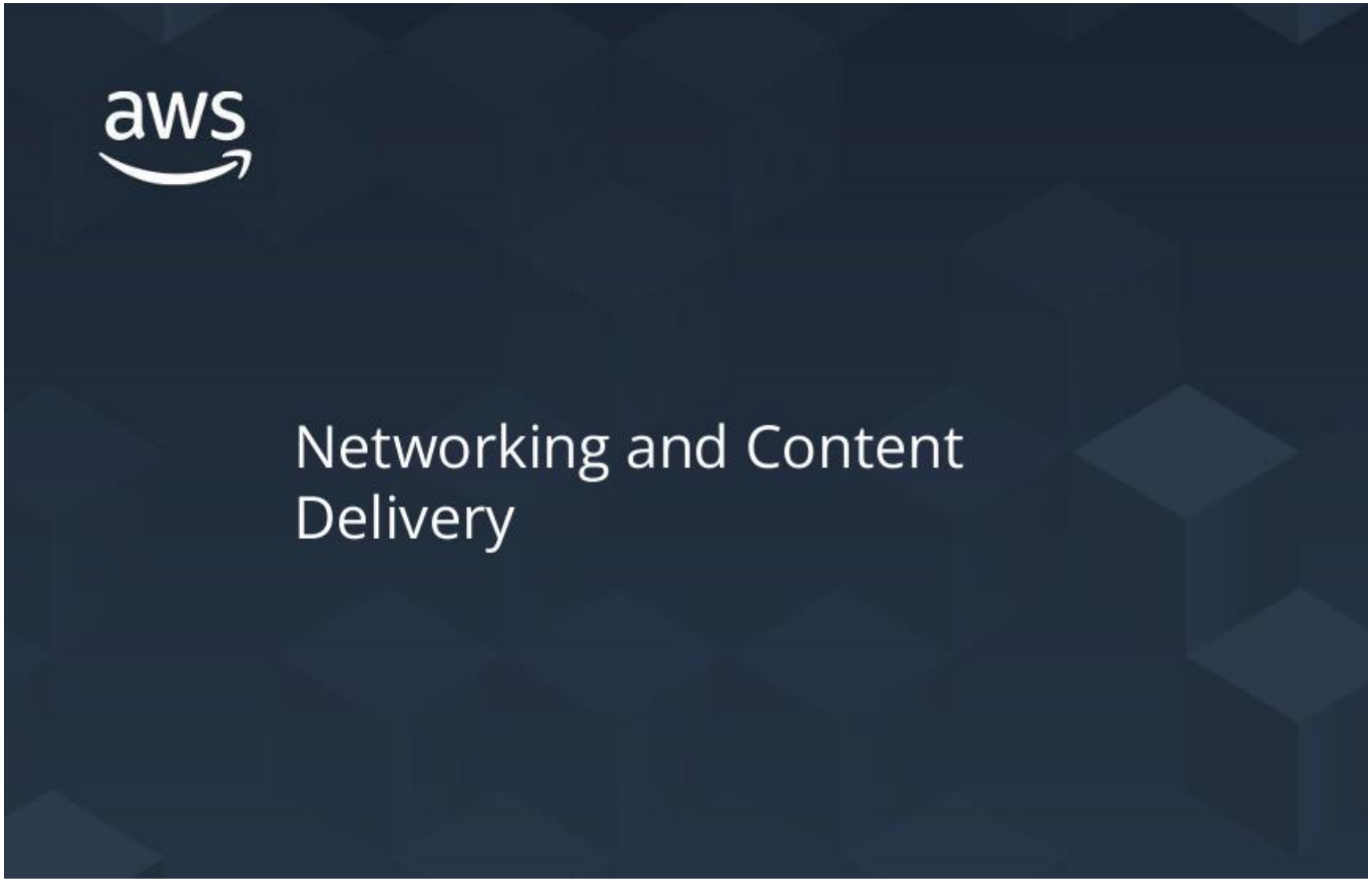


Heterogeneous migration

AWS DMS supports the following data sources and targets engines for migration:

- ❑ Sources: Oracle, Microsoft SQL Server, PostgreSQL, Db2 LUW, SAP, MySQL, MariaDB, MongoDB, and Amazon Aurora.
- ❑ Targets: Oracle, Microsoft SQL Server, PostgreSQL, SAP ASE, MySQL, Amazon Redshift, Amazon S3, and Amazon DynamoDB.

- ❑ It performs all the management steps required during the migration, such as monitoring, scaling, error handling, network connectivity, replicating during failure, and software patching.
- ❑ AWS DMS with AWS Schema Conversion Tool (AWS SCT) helps to perform heterogeneous migration.


aws
Networking and Content Delivery

Amazon VPC

What is Amazon VPC?

Amazon Virtual Private Cloud is a service that allows users to create a virtual dedicated network for resources.

Private subnet - A subnet that does not have internet access is termed a private subnet.

Public subnet - A subnet that has internet access is termed a public subnet.

VPN only subnet - A subnet that does not have internet access but has access to the virtual private gateway for a VPN connection is termed a VPN-only subnet.

- It includes many components such as Internet gateways, VPN tools, CIDR, Subnets, Route tables, VPC endpoint, NAT instances, Bastion servers, Peering Connection, and others.
- It spans across multiple Availability Zones (AZs) within a region.
- The first four IP and last one IP addresses are reserved per subnet.
- It creates a public subnet for web servers that uses internet access and a private subnet for backend systems, such as databases or application servers.
- It can monitor resources using Amazon CloudWatch and Auto Scaling Groups.

- Every EC2 instance is launched within a default VPC with equal security and control like normal Amazon VPC. Default VPC has no private subnet.
- It uses Security Groups and NACL (Network Access Control Lists) for multi-layer security.
- Security Groups (stateful) provide instance-level security, whereas NACLs (stateless) provide subnet-level security.
- VPC sharing is a component that allows subnets to share with other AWS accounts within the same AWS Organization.

Amazon CloudFront

What is Amazon CloudFront?

Amazon CloudFront is a content delivery network (CDN) service that securely delivers any kind of data to customers worldwide with low latency, low network, and high transfer speeds.

- It makes use of Edge locations (worldwide network of data centers) to deliver the content faster.
- Without edge locations, it retrieves data from an origin such as an Amazon S3 bucket, a Media Package channel, or an HTTP server.

CloudFront provides some security features such as:

- **Field-level encryption with HTTPS** - Data remains encrypted throughout starting from the upload of sensitive data.
- **AWS Shield Standard** - Against DDoS attacks.
- **AWS Shield Standard + AWS WAF + Amazon Route 53** - Against more complex attacks than DDoS.

CloudFront is integrated with AWS Services such as:

- Amazon S3
- Amazon EC2
- Elastic Load Balancing
- Amazon Route 53
- AWS Essential Media Services

Amazon CloudFront Access Controls:

Signed URLs:

- Use this to restrict access to individual files.

Signed Cookies:

- Use this to provide access to multiple restricted files.
- Use this if the user does not want to change current URLs.

Geo Restriction:

- Use this to ***restrict*** access to the data based on the geographic location of the website viewers.

Origin Access Identity (OAI):

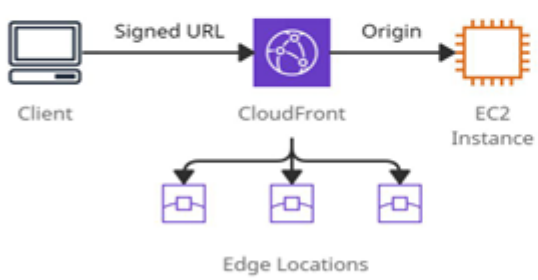
- Outside access is restricted using signed URLs and signed cookies, but what if someone tries to access objects using Amazon S3 URL, bypassing CloudFront signed URL and signed cookies. To restrict that, OAI is used.
- Use OAI as a special CloudFront user and associate it with your CloudFront distribution to secure Amazon S3 content.

CloudFront Signed URL:

- It allows access to a path, no matter what is the origin
- It can be filtered by IP, path, date, expiration
- It leverages caching features

S3 Pre-Signed URL:

- It issues a request as the person who pre-signed the URL.



CloudFront Signed URL

S3 Pre-Signed URL

Amazon Route 53

What is Route 53?

Route 53 is a managed DNS (Domain Name System) service where DNS is a collection of rules and records intended to help clients/users understand how to reach any server by its domain name.

The most common records supported in Route 53 are:

- A: hostname to IPv4
- AAAA: hostname to IPv6
- CNAME: hostname to hostname
- Alias: hostname to AWS resource

- **Route 53 hosted zone** is a collection of records for a specified domain that can be managed together.
- There are two types of zones:
  - **Public Hosted Zone** - Determines how traffic is routed on the Internet.
  - **Private Hosted Zone** - Determines how traffic is routed within VPC.

Route 53 CNAME	Route 53 Alias
It points a hostname to any other hostname.(app.mything.com -> abc.anything.com)	It points a hostname to an AWS Resource.(app.mything.com -> abc.amazonaws.com)
It works only for the non-root domains.(abcxyz.maindomain.com)	It works for the root domain and non-root domain. (maindomain.com)
It charges for CNAME queries.	It doesn't charge for Alias queries.
It points to any DNS record that is hosted anywhere.	It points to an ELB, CloudFront distribution, Elastic Beanstalk environment, S3 bucket as a static website, or another record in the same hosted zone.

Amazon Route 53

Route 53 Routing Policies:

Simple:
- It is used when there is a need to redirect traffic to a single resource.
- It does not support health checks.

Weighted:
- It is similar to simple, but you can specify a weight associated with resources.
- It supports health checks.

Failover:
- If the primary resource is down (based on health checks), it will route to a secondary destination.
- It supports health checks.

Geo-location:
- It routes traffic to the closest geographic location you are in.

Geo-proximity:
- It routes traffic based on the location of resources to the closest region within a geographic area.

Latency based:
- It routes traffic to the destination that has the least latency.

Multi-value answer:
- It distributes DNS responses across multiple IP addresses.

AWS Direct Connect

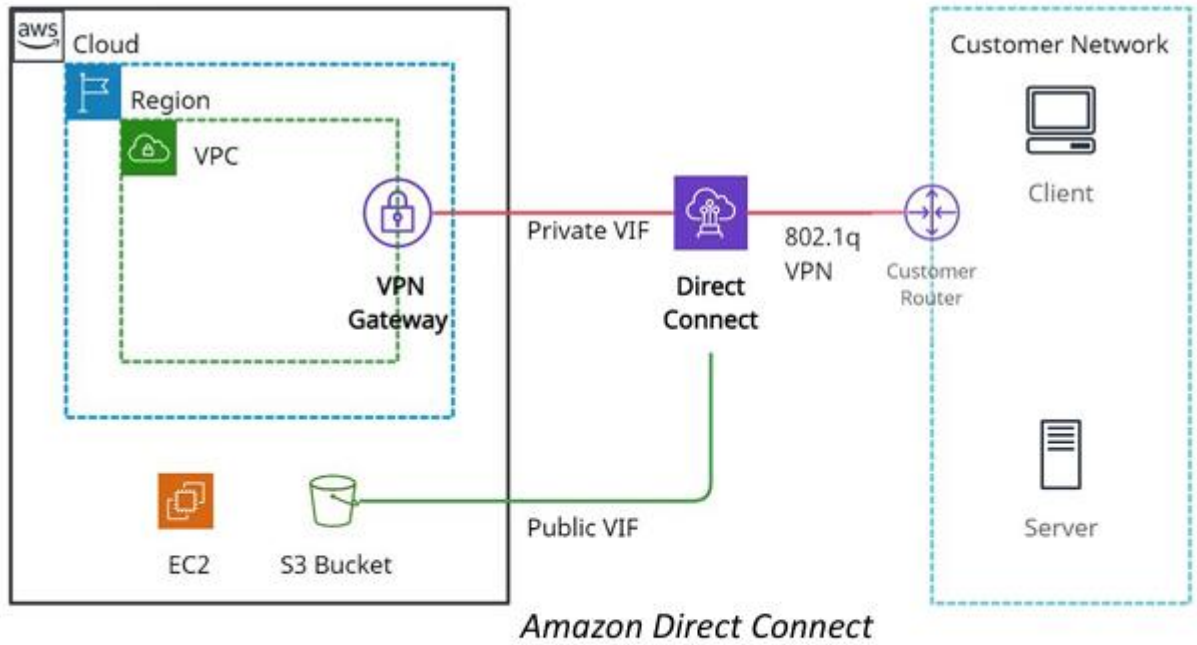
What is AWS Direct Connect?

AWS Direct Connect is a cloud service that helps to establish a dedicated connection from an on-premises network to one or more VPCs and other services in the same region.

- ✔ With the help of industry-standard 802.1Q virtual LANs (VLANs), the dedicated connection can be partitioned into multiple virtual interfaces.
- ✔ Virtual interfaces can be reconfigured at any time to meet the changing needs.

Private virtual interface:
It helps to connect an Amazon VPC using private IP addresses.

Public virtual interface:
It helps to connect AWS services located in any AWS region (except China) from your on-premises data center using public IP addresses.



Amazon Direct Connect

Pricing details:

- Port hours - charges are determined by capacity and connection type
- Outbound data transfer

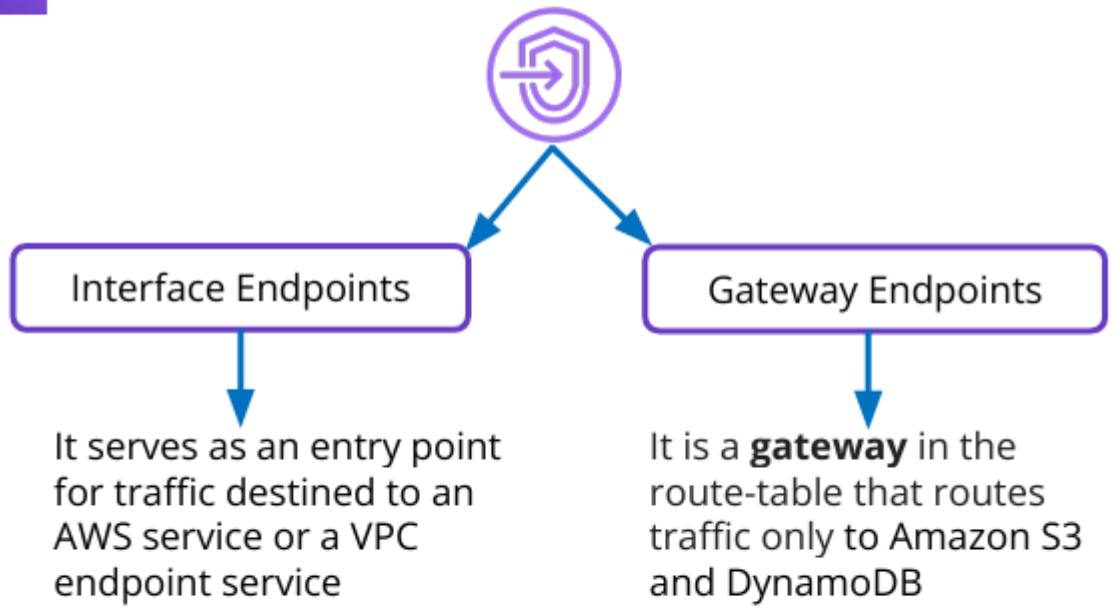
AWS PrivateLink

What is PrivateLink?

AWS PrivateLink is a network service used to connect to AWS services hosted by other AWS accounts (referred to as *endpoint services*) or AWS Marketplace.

- It is used for scenarios where the **source VPC** acts as a service provider, and the **destination VPC** acts as a service consumer.
- So, **service consumers** use an interface endpoint to access the services running in the **service provider**.
- It provides security by not allowing the public internet and reducing the exposure to threats, such as brute force and DDoS attacks.

Types of VPC End Points



AWS Transit Gateway

What is AWS Transit Gateway?

AWS Transit Gateway is a network hub used to interconnect multiple VPCs. It can be used to attach all hybrid connectivity by controlling your organization's entire AWS routing configuration in one place

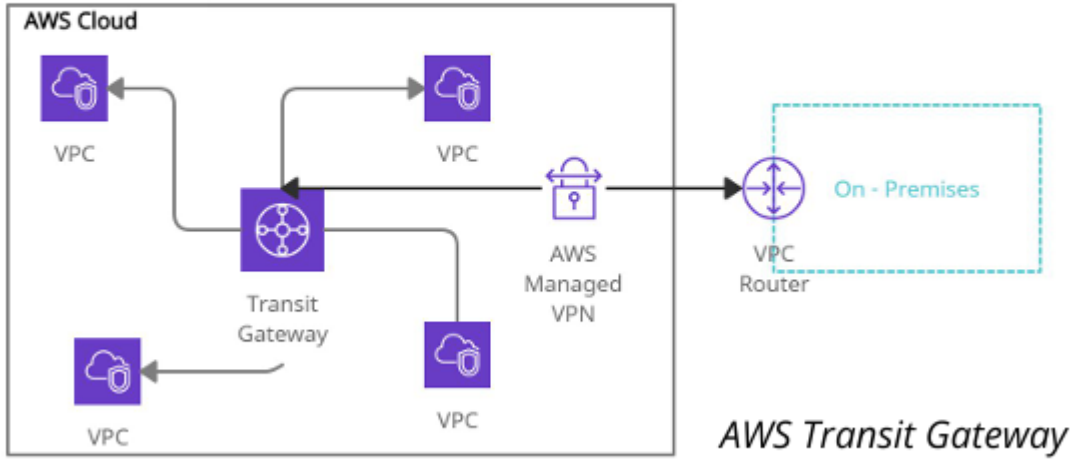
It can be more than one per region but can not be peered within a single region. It supports attaching Amazon VPCs with **IPv6** CIDRs.

It helps to solve the problem of complex VPC peering connections.

Transit Gateway reduces the complexity of maintaining VPN connections with hundreds of VPCs, which become very useful for large enterprises.

Transit Gateway vs. VPC Peering

Transit Gateway	VPC Peering
It has an hourly charge per attachment in addition to the data transfer fees.	It does not charge for data transfer.
Multicast traffic can be routed between VPC attachments to a Transit Gateway.	Multicast traffic cannot be routed to peering connections.
It provides Maximum bandwidth (burst) of 50 Gbps per Availability Zone per VPC connection.	It provides no aggregate bandwidth.
Security groups feature does not currently work with Transit Gateway	Security groups feature works with intra-Region VPC peering.



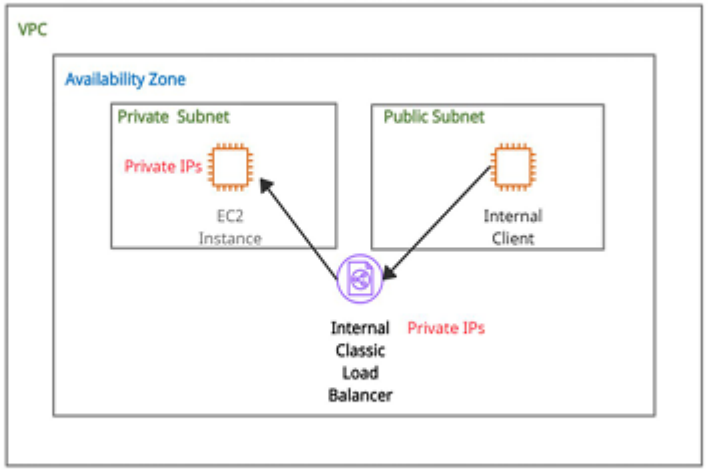
AWS Transit Gateway

Elastic Load Balancing (ELB)

What is Elastic Load Balancing?

Elastic Load Balancing is a managed service that allows traffic to get distributed across EC2 instances, containers, and virtual appliances as target groups.

Elastic Load Balancing



Classic Load Balancer (Internal)

Elastic Load Balancer types are as follows:

Classic Load Balancer:

- Oldest and less recommended load balancer.
- Routes TCP, HTTP, or HTTPS traffic at layer 4 and layer 7.
- They are used for existing EC2-Classic instances.

Application Load Balancer:

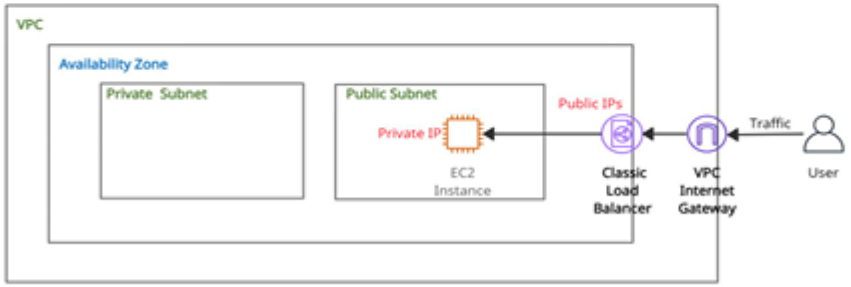
- Routes HTTP and HTTPS traffic at layer 7.
- Offers path-based routing, host-based routing, query-string, parameter-based routing, and source IP address-based routing.

Network Load Balancer:

- Routes TCP, UDP, and TLS traffic at layer 4.
- Suitable for high-performance and low latency applications.

Gateway Load Balancer:

- Suitable for third-party networking appliances.
- It simplifies tasks to manage, scale, and deploy virtual appliances.



Classic Load Balancer (Internet-Facing)

- ELB integrates with every AWS service throughout the applications.
- It is tightly integrated with Amazon EC2, Amazon ECS/EKS.
- ELB integrates with Amazon VPC and AWS WAF to offer extra security features to the applications.
- It helps monitor the servers' health and performance in real-time using Amazon CloudWatch metrics and request tracing.
- ELB can be placed based on the following aspects:
- Internet-facing ELB:
  - Load Balancers have public IPs.
- Internal only ELB:
  - Load Balancers have private IPs.
- ELB offers the functionality of Sticky sessions. It is a process to route requests to the same target from the same client.


aws
Security, Identity, and
Compliance

Amazon Identity and Access Management (IAM)

What is Amazon IAM ?

AWS Identity and Access Management is a free service used to define permissions and manage users to access multi-account AWS services.

Amazon Identity and Access Management

Amazon Identity and Access Management allows:

- users to analyze access and provide MFA (Multi-factor authentication) to protect the AWS environment.
- managing IAM users, IAM roles, and federated users.

IAM Policies

Policies are documents written in JSON (key-value pairs) used to define permissions.

IAM Users

User can be a person or service.

IAM Groups

Groups are collections of users, and policies are attached to them. It is used to assign permissions to users.

IAM Roles

IAM users or AWS services can assume a role to obtain temporary security credentials to make AWS API calls.

Amazon Cognito

What is Amazon Cognito?

Amazon Cognito is a service used for authentication, authorization, and user management for web or mobile applications.

- Amazon Cognito allows customers to sign in through social identity providers such as Google, Facebook, and Amazon, and through enterprise identity providers such as Microsoft Active Directory via SAML

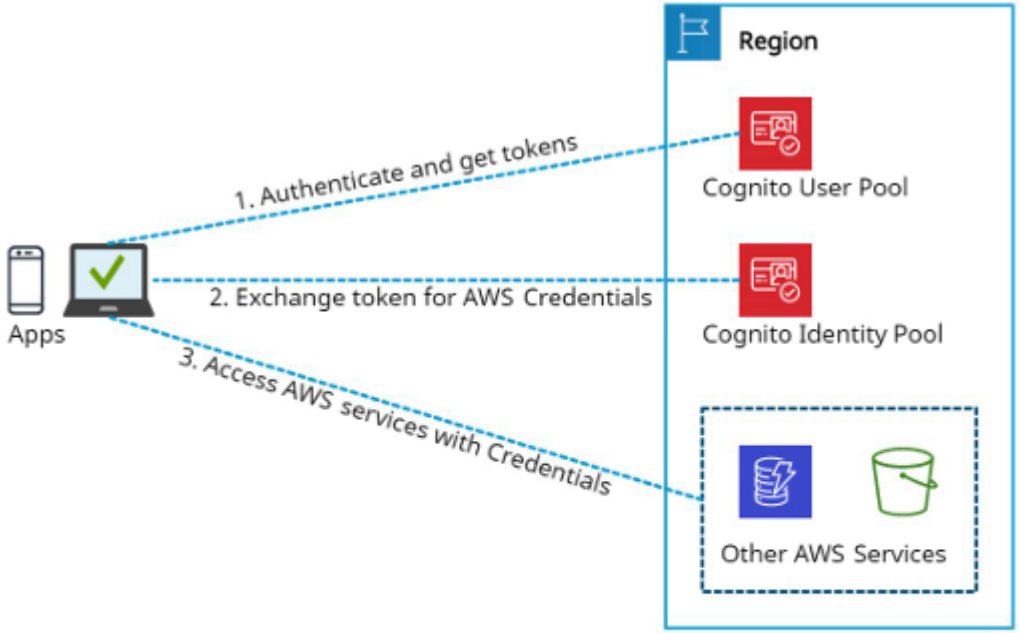
- Amazon Cognito User Pools is a standards-based Identity Provider and supports Oauth 2.0, SAML 2.0, and OpenID Connect. Amazon Cognito identity pools are useful for both authenticated and unauthenticated identities.

- Amazon Cognito is capable enough to allow usage of user pools and identity pools separately or together

The two main components of Amazon Cognito are as follows:

User pools are user repositories (*where user profile details are kept*) that provide sign-up and sign-in options for your app users.

Identity pools are user repositories of an account, which provide temporary and limited-permission AWS credentials to the users so that they can access other AWS resources without re-entering their credentials.



AWS Certificate Manager

What is AWS Certificate Manager?

AWS Certificate Manager is a service that allows a user to protect AWS applications by storing, renewing, and deploying public and private SSL/TLS X.509 certificates.

- HTTPS transactions require server certificates X.509 that bind the public key in the certificate to provide authenticity.
- The certificates are signed by a certificate authority (CA) and contain the server's name, the validity period, the public key, the signature algorithm, and more.
- It centrally manages the certificate lifecycle and helps to automate certificate renewals.
- SSL/TLS certificates provide data-in-transit security and authorize the identity of sites and connections between browsers and applications.
- The certificates created by AWS Certificate Manager for using ACM-integrated services are free.
- With AWS Certificate Manager Private Certificate Authority, monthly charges are applied for the private CA operation and the private certificates issued.

The types of SSL certificates are:

Extended Validation Certificates (EV SSL)
Most expensive SSL certificate type

Organization Validated Certificates (OV SSL)
Validates a business' creditably.

Domain Validated Certificates (DV SSL)
Provides minimal encryption

Wildcard SSL Certificate
Secures base domain and subdomains.

Multi-Domain SSL Certificate (MDC)
Secure up to hundreds of domain and subdomains.

Unified Communications Certificate (UCC)
Single certificate secures multiple domain names.

Ways to deploy managed X.509 certificates:

AWS Certificate Manager (ACM)
Useful for customers who need a secure and public web presence.

ACM Private CA
Useful for customers that are intended for private use within an organization.

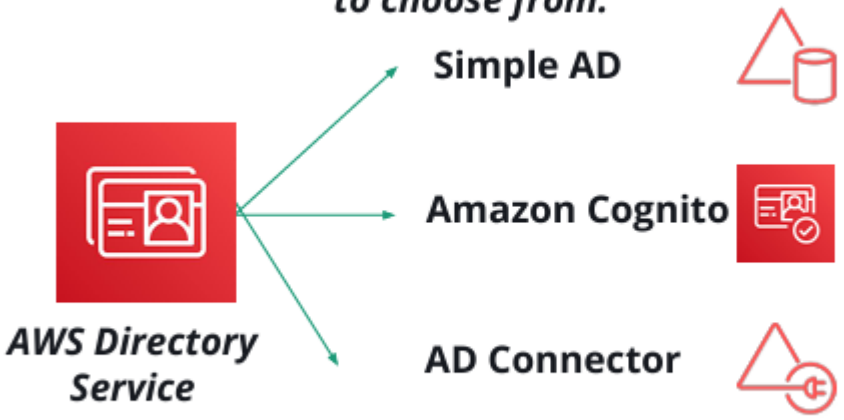
AWS Directory Service

What is AWS Directory Service?

AWS Directory Service, also known as AWS Managed Microsoft Active Directory (AD), enables multiple ways to use Microsoft Active Directory (AD) with other AWS services.

- Using AWS Managed Microsoft AD, it becomes easy to migrate AD-dependent applications and Windows workloads to AWS.
- A trust relationship can be created between AWS Managed Microsoft AD and existing on-premises Microsoft Active using single sign-on (SSO).

AWS Directory Service provides the following directory types to choose from:



Simple AD

- It is an inexpensive Active Directory-compatible service driven by SAMBA 4.
- It can be used when there is a need for less than 5000 users.
- It does not support Multi-factor authentication (MFA).

Amazon Cognito

- It is a user directory type that provides sign-up and sign-in for the application using Amazon Cognito User Pools.

AD Connector

- It is like a gateway used for redirecting directory requests to the on-premise Active Directory.
- For this, there must be an existing AD, and VPC must be connected to the on-premise network via VPN or Direct Connect.
- It supports multi-factor authentication (MFA) via existing RADIUS-based MFA infrastructure.

AWS Key Management Service

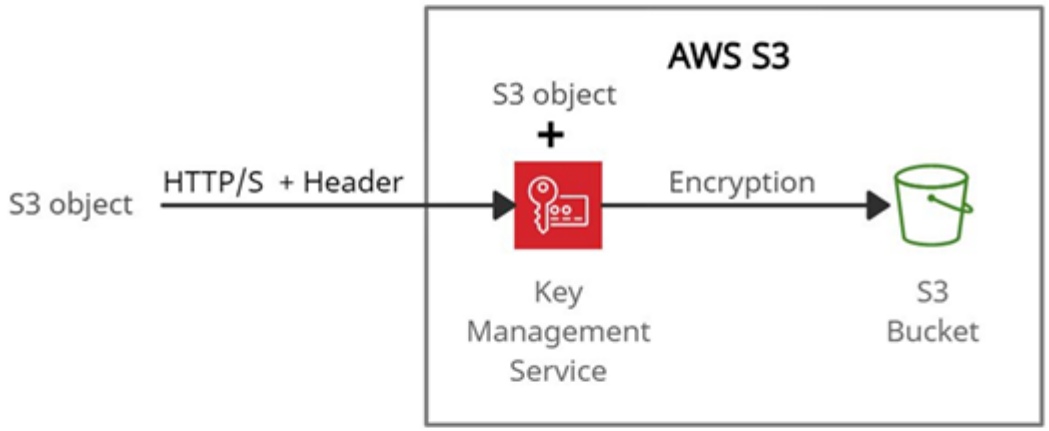
What is AWS Key Management Service?

AWS Key Management Service is a global service that creates, stores, and manages encryption keys.

AWS Key Management Service

- Provides data security at rest using encryption keys and provides access control for encryption, decryption, and re-encryption.
- Offers SDKs for different languages to add digital signature capability in the application code.
- Allows rotation of master keys once a year using previous versions of keys.
- AWS KMS produces new cryptographic data for the KMS key once a year, when automatic key rotation is turned on for a KMS key.
- AWS KMS preserves all previous iterations of the cryptographic information so that you can decrypt any data that has been encrypted using that KMS key. Until the KMS key is deleted, AWS KMS does not remove any rotated key material.

Encryption using AWS KMS



Customer Managed CMKs:

The CMKs created, managed, and used by users are termed as Customer managed CMKs and support cryptographic operations.

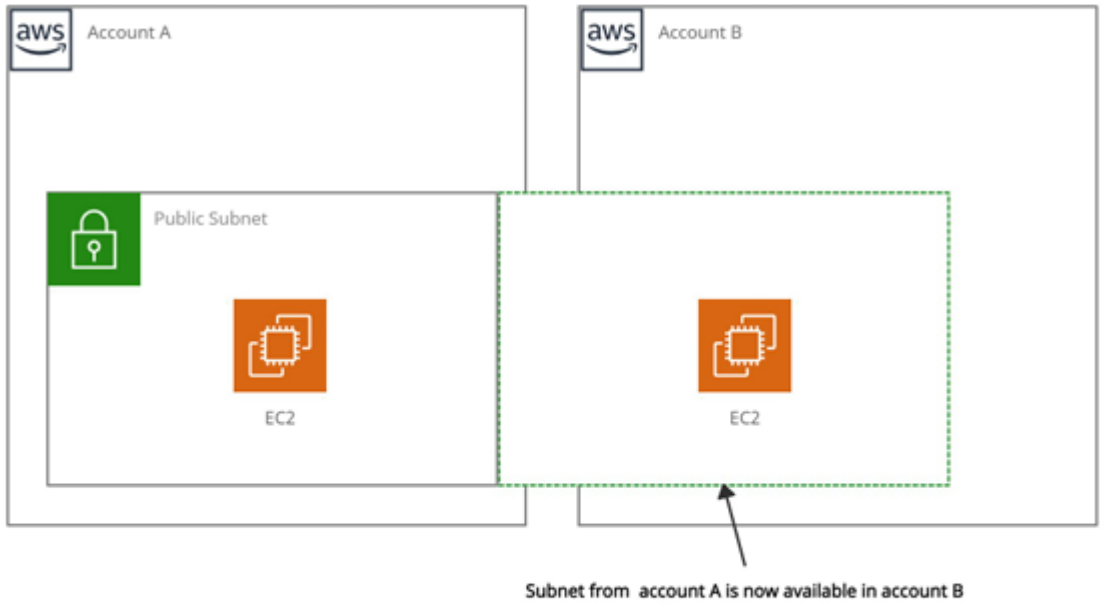
AWS Managed CMKs:

The CMKs created, managed, and used by AWS services on the user's behalf are termed AWS-managed CMKs.

AWS Resource Access Manager

What is AWS Resource Access Manager?

AWS Resource Access Manager (RAM) is a service that allows resources to be shared through AWS Organizations or across AWS accounts.



AWS Resource Access Manager

- The resource sharing feature of AWS RAM reduces customers' need to create duplicate resources in each of their accounts.
- It controls the consumption of shared resources using existing policies and permissions.
- It can be integrated with Amazon CloudWatch and AWS CloudTrail to provide detailed visibility into shared resources and accounts.
- Access control policies in AWS IAM and Service Control Policies in AWS Organizations provide security and governance controls to AWS Resource Access Manager (RAM).

AWS Secrets Manager

What is AWS Secrets Manager?

AWS Secrets Manager is a service that prevents secret credentials from being hardcoded in the source code.

AWS Secrets Manager

Secrets Manager can be accessed using the following ways:

- AWS Management Console
- AWS Command Line Tools
- AWS SDKs
- HTTPS Query API

- ❑ It provides security and compliance facilities by rotating secrets safely without the need for code deployment.
- ❑ It integrates with AWS CloudTrail and AWS CloudWatch to log and monitor services for centralized auditing.
- ❑ It integrates with AWS Config and facilitates tracking of changes in Secrets Manager.

AWS Secrets Manager:

- ❑ Ensures in-transit encryption of the secret between AWS and the system to retrieve the secret.
- ❑ Rotates credentials for AWS services using the Lambda function that instructs Secrets Manager to interact with the service or database.
- ❑ Stores the encrypted secret value in SecretString or SecretBinary field.
- ❑ Uses open-source client components to cache secrets and updates them when there is a need for rotation.

Secret rotation is supported with the below Databases:

- MySQL, PostgreSQL, Oracle, MariaDB, Microsoft SQL Server, on Amazon RDS
- Amazon Aurora on Amazon RDS
- Amazon DocumentDB
- Amazon Redshift

AWS Security Hub

What is AWS Security Hub?

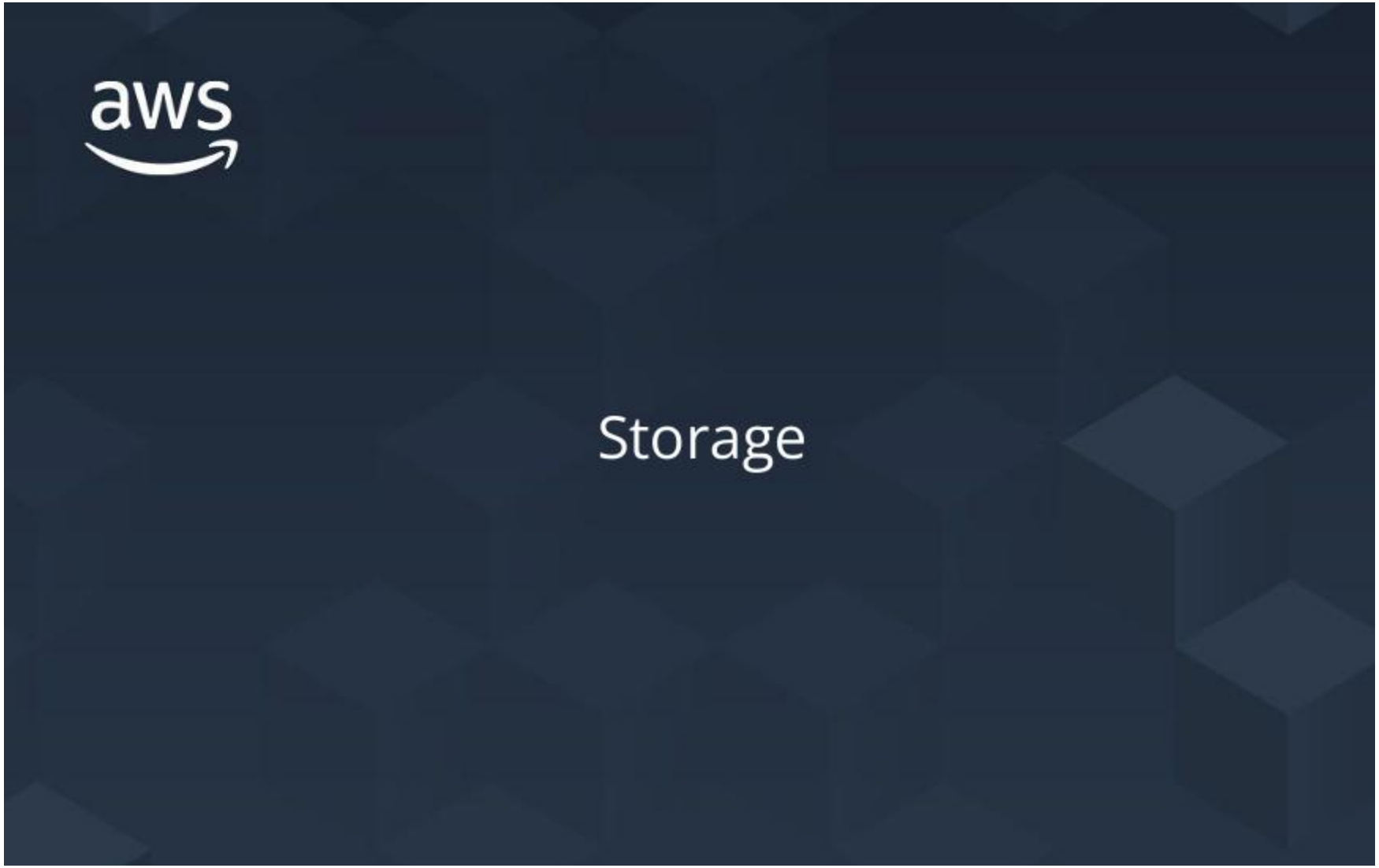
AWS Security Hub is a service that offers security aspects to protect the environment using industry-standard best practices.

- AWS Security Hub helps the Payment Card Industry Data Security Standard (PCI DSS) and the Center for Internet Security (CIS) AWS Foundations Benchmark with a set of security configuration best practices for AWS.

Enabling (or disabling) Can quickly do AWS Security Hub through:

- AWS Management Console
- AWS CLI
- By using Infrastructure-as-Code tools -- Terraform

- It collects findings or alerts from multiple AWS accounts. Then it analyzes security trends and identifies the highest priority security issues.
- AWS Security Hub provides an option to aggregate, organize, and prioritize the security alerts or findings from multiple AWS services.
- It automatically checks the compliance status using CIS AWS Foundations Benchmark.
- The security alerts or findings can be investigated using Amazon Detective or Amazon CloudWatch Event rules.
- It collects data from AWS services across accounts and reduces the need for time-consuming data conversion efforts
- It uses integrated dashboards to show the current security and compliance status.
- Charges are applied only for the current Region, not for all Regions in which Security Hub is enabled.


aws
Storage

Amazon Simple Storage Service (S3)

What is Amazon Simple Storage Service?

Amazon S3 is a simple service used to provide key-based object storage across multiple availability zones (AZs) in a specific region.

- ❑ S3 is a global service with region-specific buckets.
- ❑ It is also termed a static website hosting service.
- ❑ It provides 99.999999999% (11 9's) of content durability.
- ❑ S3 offers strong read-after-write consistency for any object.
- ❑ Objects (files) are stored in a region-specific container known as Bucket.
- ❑ Objects that are stored can range from 0 bytes - 5TB.

- It provides 'Multipart upload' features that upload objects in parts, suitable for 100 MB or larger objects.
- It offers to choose 'Versioning' features to retain multiple versions of objects, must enable versioning at both source and destination.
- Amazon S3 Transfer Acceleration allows fast and secure transfer of objects over long distances with minimum latency using Amazon CloudFront's Edge Locations.
- Amazon S3 uses access control lists (ACL) to control access to the objects and buckets.
- Amazon S3 provides Cross-Account access to the objects and buckets by assuming a role with specified privileges.

Amazon S3

Amazon S3 uses the following ways for security:

User-based security

- IAM policies

Resource-Based

- Bucket Policies
- Bucket Access Control List (ACL)
- Object Access Control List (ACL)

Amazon S3 provides the following storage classes used to maintain the integrity of the objects:

- ❑ **S3 Standard** - offers frequent data access.
- ❑ **S3 Intelligent-Tiering** - automatically transfer data to other cost-effective access tiers.
- ❑ **S3 Standard-IA** - offers immediate and infrequent data access.
- ❑ **S3 One Zone-IA** - infrequent data access.
- ❑ **S3 Glacier** - long-term archive data, cheap data retrieval.
- ❑ **S3 Glacier Deep Archive** - used for long-term retention.

Amazon S3 offers to choose from the following ways to replicate objects:

- Cross-Region Replication - used to replicate objects in different AWS Regions.
- Same Region Replication - used to replicate objects in the same AWS Region.

Amazon Elastic Block Store

What is Amazon Elastic Block Store?

Amazon Elastic Block Store is a service that provides the block-level storage drive to store persistent data.

- Multiple EBS volumes can be attached to a single EC2 instance in the same availability zone.
- A single EBS volume can not be attached to multiple EC2 instances.
- Amazon EBS Multi-Attach is a feature used to attach a single Provisioned IOPS SSD (io1 or io2) volume to multiple instances in the same Availability Zone.
- EBS volumes persist independently after getting attached to an instance, which means the data will not be erased even if it terminates.
- By default, the root EBS volume gets terminated when the instance is terminated.

- By default, the non-root EBS volume does not get affected when the instance is terminated.
- Amazon EBS can be attached and detached to an instance and can be reattached to other EC2 instances.
- Amazon EBS easily scales up to petabytes of data storage.
- Amazon EBS volumes are best suited for database servers with high reads and write and throughput-intensive workloads with continuous reads and write.
- Amazon EBS uses AWS KMS service with AES-256 algorithm to support encryption.
- Amazon EBS offers point-in-time snapshots for volumes to migrate to other AZs or regions.
- EBS snapshots are region-specific and are incrementally stored in Amazon S3.

Amazon Elastic Block Store

EBS volumes types are as follows:

SSD (Solid-state drives)

General Purpose SSD:

- Useful for low-latency applications, development, and test environments.
- Supports volume size from 1 GiB to 16 TiB.
- Allows 16,000 as maximum IOPS per volume.
- Allows 1000 MiB/s as maximum throughput per volume.

Provisioned IOPS SSD:

- Useful for I/O-intensive database workloads and provide sub-millisecond latency.
- Supports volume size from 4 GiB to 64 TiB.
- Allows 256,000 as maximum IOPS per volume.
- Allows 4,000 MiB/s as maximum throughput per volume.
- The multi-Attach feature is supported for io1 and io2

HDD (Hard disk drives)

Throughput Optimized HDD:

- Useful for Big data and Log processing workloads.
- Supports volume size from 125 GiB to 16 TiB.
- Allows 500 as maximum IOPS per volume.
- Allows 500 MiB/s as maximum throughput per volume.

Cold HDD:

- Useful for infrequently accessed data and lowest cost workloads.
- Supports volume size from 125 GiB to 16 TiB.
- Allows 250 as maximum IOPS per volume.
- Allows 250 MiB/s as maximum throughput per volume.

Amazon Elastic File System (EFS)

What is Amazon Elastic File System?

Amazon Elastic File System is a managed service used to create and scale file storage systems for AWS and on-premises resources.

Amazon EFS

It offers the following storage classes for file storage:

- EFS Standard storage class
- EFS Infrequent Access storage class - can store less frequently accessed files.

- ❑ It spans multiple availability zones and regions.
- ❑ It uses EFS Mount Target to share a file system with multiple availability zones and VPCs.
- ❑ It is best suited for Linux-based workloads and applications.
- ❑ Multiple instances can access it at the same time leads to high throughput and low latency IOPS.
- ❑ It automatically scales storage capacity up to petabyte.
- ❑ It supports file locking and strong data consistency.
- ❑ It offers data encryption at rest and in-transit using AWS KMS and TLS, respectively.
- ❑ It uses POSIX permissions to control access to files and directories.

It offers the following modes to ease the file storage system:

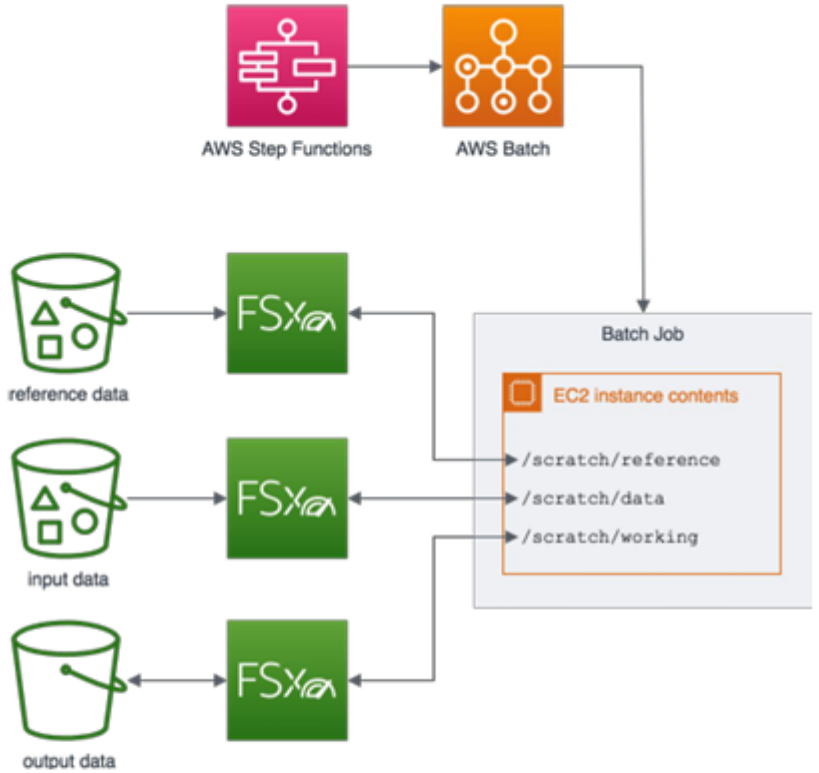
- ❑ Performance modes -
  - General Purpose performance mode: Useful for low-latency workloads.
  - Max I/O mode: High throughput workloads.
- ❑ Throughput modes -
  - Bursting Throughput mode: Throughput increases based on the file system storage.
  - Provisioned Throughput mode: Throughput changes are independent of the file system storage.
- ❑ It provides EFS lifecycle management policies based on the number of days ranges from 7-90 days to automatically move files from Standard storage class to EFS IA storage class.

Amazon FSx for Lustre

What is Amazon FSx?

Amazon FSx for Lustre is an FSx solution that offers scalable storage for the Lustre system (parallel and high-performance file storage system).

Amazon FSx



Using Amazon FSx

- Supports fast processing workloads like custom electronic design automation (EDA) and high-performance computing (HPC).
- Offers to choose between SSD and HDD for storage.
- Stores datasets in S3 as files instead of objects and automatically updates with the latest data to run the workload.
- Offers to select unreplicated file systems for shorter-term data processing.
- FSx can be used with existing Linux-based applications without any changes.
- Offers network access control using POSIX permissions or Amazon VPC Security Groups.
- FSx easily provides data-at-rest and in-transit encryption.

AWS Backup can also be used to backup Lustre file systems.

Amazon FSx for Windows File Server

What is Amazon FSx for Windows File Server??

Amazon FSx for Windows File Server is an FSx solution that offers a scalable and shared file storage system on the Microsoft Windows server.

Amazon FSx

- Using the Server Message Block (SMB) protocol with Amazon FSx Can access file storage systems from multiple windows servers.
- Using SMB protocol, Amazon FSx can connect file systems to Amazon EC2, Amazon ECS, Amazon WorkSpaces, Amazon AppStream 2.0 instances, and on-premises servers using AWS Direct Connect or AWS VPN.

- FSx provides high availability (Multi-AZ deployments) with an active and standby file server in separate AZs.
- It automatically and synchronously replicates data in the standby Availability Zone (AZ) to manage failover.
- Using AWS DataSync with Amazon FSx helps to migrate self-managed file systems to Windows storage systems.

Amazon FSx uses SSD storage, offers high throughput and IOPS with sub-millisecond latencies for Windows workloads.

Amazon FSx offers identity-based authentication using Microsoft Active Directory (AD).

Amazon S3 Glacier

What is Amazon S3 Glacier?

Amazon S3 Glacier is a web service with vaults that offer long-term data archiving and data backup.

It is the cheapest S3 storage class and offers 99.999999999% of data durability.

S3 Glacier provides the following data retrieval options:

Expedited retrievals -

- It retrieves data in 1-5 minutes.

Standard retrievals -

- It retrieves data between 3-5 hours.

Bulk retrievals -

- It retrieves data between 5-12 hours.

- S3-Standard, S3 Standard-IA, and S3 Glacier storage classes, objects, or data are automatically stored across availability zones in a specific region.
- A vault is a place for storing archives with a unique address.
- Amazon S3 Glacier jobs are the select queries that execute to retrieve archived data. It uses Amazon SNS to notify when the jobs complete.
- Amazon S3 Glacier does not provide real-time data retrieval of the archives.
- Amazon S3 Glacier uses 'S3 Glacier Select' to query archive objects in uncompressed CSV format and store the output to the S3 bucket.
- Amazon S3 Glacier Select uses common SQL statements like SELECT, FROM, and WHERE.
- It offers only SSE-KMS and SSE-S3 encryption.

Amazon Backup

What is Amazon Backup?

AWS Backup is a secure service that automates and governs data backup (protection) in the AWS cloud and on-premises.

Amazon Backup

AWS Backup provides the following features:

- Scheduled backup plans (policies) to automate backup of AWS resources across AWS accounts and regions.
- Incremental backup to minimize storage costs.
- Backup retention plans to retain and expire backups automatically.
- Dashboard in the AWS Backup console to monitor backup and restore activities.
- Different encryption keys for encrypting multiple AWS resources.
- Lifecycle policies configured to transition backups from Amazon EFS to cold storage automatically.

AWS Snowball

What is AWS Snowball?

AWS Snowball is a data transfer service that uses storage gadgets to transfer a huge amount of data ranging from 50TB - 80TB between Amazon Simple Storage Service and onsite data storage location at high speed.

- It makes use of AWS Key Management Service to protect data in transit securely.
- If data transfer is less than 10 TB, no need to use Snowball.
- The Snowball client and the Amazon S3 Adapter for Snowball are used to perform data transfers on the Snowball device locally.
- **Snowball's size** - *50 TB (42 usable) and 80 TB (72 usable). To move a petabyte (1024 TB) of data, 14 Snowballs can be used.*

- If data transfers involve large files and multiple jobs, you might separate the data into several smaller data segments. Parallelization helps to transfer data with Snowball at a faster rate.
  E.g., ten segments of 7 TB each in a size of 80 TB Snowball.
- AWS Snowball is integrated with other AWS services such as AWS CloudTrail to capture all API calls as events and with Amazon Simple Notification Service (Amazon SNS) to notify about data transfer.
- **AWS Snowball Edge** is a type of Snowball device that can transport data faster and process edge-computing workloads between the local environment and the AWS Cloud.
- Using Snowball Edge devices, one can execute EC2 AMIs and deploy AWS Lambda code on the devices to perform processing and analysis with the applications.

AWS Storage Gateway

What is AWS Storage Gateway?

AWS Storage Gateway is a virtual device installed as a hypervisor or VM at the on-premises data center to integrate with AWS storage services.

It offers secure, scalable, and cost-effective storage management.

It offers the following types of storage gateways:

File Gateway

- It uses S3 Standard, S3 Standard-IA, and S3 One Zone-IA as storage interfaces.
- It supports NFS and SMB protocol to store and retrieve files.
- It supports the WORM (Write Once Read Many) based file system.

Tape Gateway -

- It uses Amazon Glacier to archive backup data.
- It stores data on virtual tape cartridges using VTL (virtual tape library) interface with an iSCSI connection.
- It supports WORM based file system.

Volume Gateway -

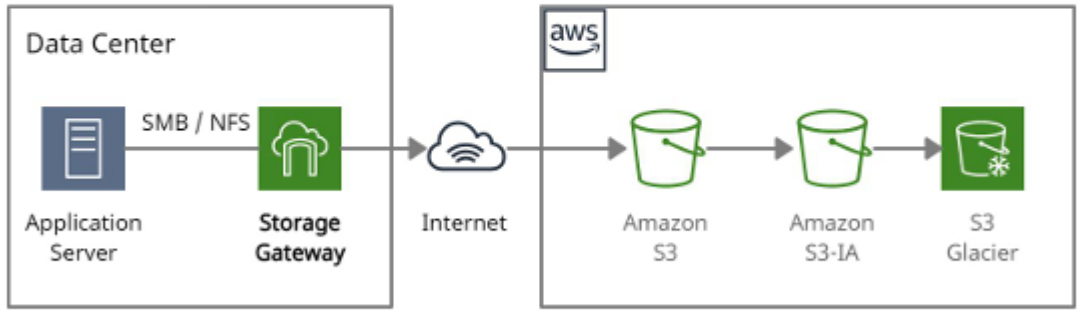
- It uses block storage (EBS volumes) as a storage interface.
- It supports the iSCSI block protocol.

Stored Volumes:

- It stores the entire data locally.
- It offers low-latency access to the entire data and backup.
- It can use 32 volumes with size ranges from 1 GiB - 16 TiB

Cached Gateway:

- It stores the most recent data locally in the storage gateway and the rest of the data in Amazon S3.
- It can use 32 volumes with size ranges from 1 GiB - 32 TiB.



AWS Storage Gateway