Qiniso Sakhiwo Mtshali
CONTACT: 0798111015
EMAIL: Qiniso.sakhiwo.mtshali@gmail.com

2024

GOOD LUCK!!!!!!!!!!

About Me

I am Qiniso Sakhiwo Mtshali, an Advanced Diploma Information Technology Multimedia graduate from Tshwane University of Technology as of 2024. I am passionate about cloud computing and am currently pursuing a career as an AWS Cloud Engineer and Developer. As I embark on this journey, I am planning to take the AWS Certified Cloud Practitioner exam to solidify my foundational knowledge and kickstart my professional development in the AWS ecosystem.

About AWS Cloud Practitioner Guide

This guide is designed to provide an in-depth understanding of the key concepts, services, and best practices of Amazon Web Services (AWS) for aspiring AWS Cloud Practitioners. Whether you are new to cloud computing or looking to enhance your foundational knowledge of AWS, this document covers essential topics that will help you develop a comprehensive understanding of AWS cloud infrastructure.

Key Areas Covered:

AWS Global Infrastructure: Learn about AWS regions, availability zones, and the backbone network that ensures reliable and high-performing cloud services globally.

Core AWS Services: Gain insights into fundamental AWS services including Amazon EC2, Amazon S3, Amazon RDS, and AWS Lambda, understanding their roles in building scalable and resilient applications.

Security and Compliance: Explore the AWS Shared Responsibility Model, security best practices, and compliance frameworks that help protect your data and applications.

Networking and Content Delivery: Understand the basics of Amazon VPC, AWS Direct Connect, and services like Amazon CloudFront that enhance connectivity and content delivery.

Storage Solutions: Compare different storage options including Amazon EBS, Amazon S3, and Amazon EFS, and learn when to use each service for optimal performance and cost efficiency.

IAM and Access Management: Delve into AWS Identity and Access Management (IAM) to learn how to manage users, roles, and policies securely.

Support Plans: Get an overview of AWS support plans, including Enterprise On-Ramp and Enterprise Support, to understand the different levels of technical assistance available.

This guide aims to equip you with the knowledge required to effectively leverage AWS services, ensuring you are prepared to navigate the AWS ecosystem and apply best practices in your cloud journey. Whether preparing for the AWS Certified Cloud Practitioner exam or enhancing your professional cloud skills, this document serves as a valuable resource for mastering the essentials of AWS.

AWS Certified Cloud Practitioner (CLF-C02) Exam Notes

AWS Basics

Why use AWS?

With AWS, you have the ability to only pay for what you use. For example, you could use Amazon EC2 to launch as many or as few virtual servers as you need.

What is Cloud Computing?

Cloud computing is the on-demand delivery of IT resources over the internet with pay-as-you-go pricing.

What are the 6 main benefits of using AWS?

1. Trade fixed expense for variable expense (pay-as-you-go)
2. Benefit from massive economies of scale
3. Stop guessing capacity
4. Increase speed and agility
5. Stop spending money running data centers
6. Go global in minutes

How is AWS able to offer pay-as-you-go?

AWS separates different virtual machines from each other on the same hardware. So for example, my company and John's company could be sharing the exact same physical server owned by AWS, but our virtual machines would still be completely separated. This ensures that servers are put to their highest and best use, allowing cheaper server costs to us as users than what we'd typically have to pay for to manage our own data server (unless my company becomes as big as Amazon). **YOU PAY ONLY FOR WHAT YOU USE**

AWS Service Offerings

1. Compute
2. Storage
3. Network Security
4. Blockchain
5. Machine learning
6. Artificial intelligence
7. Robot development
8. Video production management
9. Orbital satellites

Deployment models for cloud computing

When selecting a cloud strategy, a company must consider factors such as required cloud application components, preferred resource management tools, and any legacy IT infrastructure requirements.

1. Cloud Based Deployment Model
2. Private Cloud/ On Premises Deployment Model
3. Hybrid Cloud Deployment Model

Cloud Deployment Models

- Cloud deployment models describe the environment in which cloud services are deployed. These models dictate how cloud resources are utilized, managed, and who has access to them. The primary cloud deployment models are:

Public Cloud Deployment Model:

- **Description**: Services are provided over the public internet and shared among multiple organizations (tenants).
- **Providers**: Examples include Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP).
- **Use Cases**: Ideal for applications with unpredictable traffic, development and testing environments, and public-facing services.

On-Premises Deployment is also known as a private cloud deployment:

- **Description**: Dedicated to a single organization, offering more control and customization. Can be hosted on-premises or by a third-party provider.
- **Use Cases**: Suitable for organizations with strict regulatory requirements, sensitive data, or high performance and control needs.

Hybrid Cloud Deployment Model:

- **Description**: Combines public and private clouds, allowing data and applications to be shared between them.
- **Use Cases**: Best for organizations that need to balance between scalability and security, or those transitioning to the cloud.

Community Cloud Deployment Model:

- **Description**: Shared by several organizations with common goals and requirements. Managed internally or by a third-party.
- **Use Cases**: Ideal for government agencies, healthcare organizations, and educational institutions with shared objectives and standards.

Cloud Service Models

- Cloud service models describe the type of services offered over the cloud. These models dictate the level of management and control provided to the user. The primary cloud service models are:

Infrastructure as a Service (IaaS) Cloud Service Model:

- **Description**: Provides virtualized computing resources over the internet, including servers, storage, and networking.

- **Examples**: AWS EC2, Microsoft Azure Virtual Machines, Google Cloud Compute Engine.
- **Use Cases**: Hosting websites and applications, data storage and backup, development and testing environments.

Platform as a Service (PaaS) Cloud Service Model:

- **Description**: Provides a platform allowing customers to develop, run, and manage applications without dealing with the underlying infrastructure.
- **Examples**: Google App Engine, Microsoft Azure App Service, AWS Elastic Beanstalk.
- **Use Cases:** Application development and deployment, API development and management.

Software as a Service (SaaS) Cloud Service Model:

- **Description**: Delivers software applications over the internet, on a subscription basis, accessible via a web browser.
- **Examples**: Google Workspace, Microsoft 365, Salesforce.
- **Use Cases**: Email and collaboration tools, customer relationship management (CRM), project management.

Function as a Service (FaaS) Cloud Service Model:

- **Description**: Allows users to execute code in response to events without provisioning or managing servers. Also known as serverless computing.
- **Examples**: AWS Lambda, Google Cloud Functions, Microsoft Azure Functions.
- **Use Cases**: Real-time data processing, event-driven applications, microservices.

Cloud Deployment Models and Cloud Service Models Summary

- **Cloud Deployment Models**: Define where the cloud services are hosted and who has access to them (Public, Private, Hybrid, Community).
- **Cloud Service Models**: Define the type of cloud services provided and the level of control and management (IaaS, PaaS, SaaS, FaaS).

Understanding both sets of models helps organizations to choose the right combination of deployment and service models that best fit their business needs and technical requirements.

AWS Cloud Concepts

1. AWS Global Infrastructure:

- **Regions**: Physical locations with multiple Availability Zones. Each region is a separate geographic area. Examples include US East (N. Virginia), EU (Ireland).

- **Availability Zones (AZs)**: Data centers within a region.

Each region has multiple, isolated locations known as Availability Zones.

- **Edge Locations**:Deliver content closer to users.

  These are endpoints for AWS used for caching content. Typically used with CloudFront to deliver content faster to end-users.

2. Core Services:

1. **Compute**: Amazon EC2, Lambda, Elastic Beanstalk.
2. **Storage**: S3, EBS, Glacier.
3. **Databases**: RDS, DynamoDB, Redshift.
4. **Networking**: VPC, Route 53, CloudFront.

Compute:

1. **Amazon EC2**: Virtual servers in the cloud.

- **Amazon EC2 Auto Scaling**: Automatically adjusts the number of EC2 instances in response to demand.
- **Amazon Elastic Container Service (ECS) and EKS**: Manage containers using Docker and Kubernetes.

2. **AWS Lambda**: Run code without provisioning servers. Charged only for the compute time you consume.
3. **AWS Elastic Beanstalk**: Deploy and manage applications quickly without worrying about the underlying infrastructure.

Storage:

- **Amazon S3**: Scalable object storage.
- **Amazon EBS**: Block storage volumes for use with EC2 instances.
- **Amazon EFS**: Scalable file storage for use with EC2.
- **Amazon Glacier**: Long-term, low-cost storage for backups and archival.

Databases:

- **Amazon RDS:** Managed relational database service (supports MySQL, PostgreSQL, MariaDB, Oracle, and SQL Server).
- **Amazon DynamoDB:** Managed NoSQL database.
- **Amazon Redshift**: Data warehousing service.
- **Amazon Aurora:** High-performance managed relational database.

Networking:

- **Amazon VPC**: Virtual Private Cloud to launch AWS resources in a virtual network.
- **Amazon Route 53**: Scalable Domain Name System (DNS) web service.
- **Amazon CloudFront**: Content Delivery Network (CDN).
- **AWS Direct Connect**: Dedicated network connection to AWS.

Amazon Elastic Compute Cloud (Amazon EC2)

- Provides secure, resizable compute capacity in the cloud as Amazon EC2 instances.

With an Amazon EC2 instance you can use a virtual server to run applications in the AWS Cloud.

- You can provision and launch an Amazon EC2 instance within minutes.
- You can stop using it when you have finished running a workload.
- You pay only for the compute time you use when an instance is running, not when it is stopped or terminated.
- You can save costs by paying only for server capacity that you need or want.

What OS can you use with EC2?

- You can use Windows or Linux OS.

What software can you run on the EC2 instance?

Almost anything. Software, databases, web apps, scripts, and more.

What are the EC2 instance types?

- There are a variety of tasks that can be done, so different EC2 instances should be used for specific jobs.
- Each instance type is grouped under an instance family, similar to how employees of a big company are run under teams (marketing, sales, engineering, HR, etc).

What are the EC2 instance families?

- General Purpose
- Compute Optimized
- Memory Optimized
- Accelerated Computing
- Storage Optimized

What is the General Purpose EC2 family used for?

provide a balance of compute, memory, and networking resources. You can use them for a variety of workloads, such as:

- application servers
- gaming servers
- backend servers for enterprise applications
- small and medium databases

What is the Compute Optimized EC2 family used for?

- Are ideal for high-performance web servers, compute-intensive applications servers, and dedicated gaming servers.

What is the Memory Optimized EC2 family used for?

- Are designed to deliver fast performance for workloads that process large datasets in memory

What is the Accelerated Computing EC2 family used for?

- Floating point number calculation, graphics processing, streaming data, or data pattern matching.

What is the Storage Optimized EC2 family used for?

- High performance for locally stored data.
- are designed for workloads that require high, sequential read and write access to large datasets on local storage. E.g distributed file systems, data warehousing applications, and high-frequency online transaction processing (OLTP) systems.

Amazon EC2 pricing

With Amazon EC2, you pay only for the compute time that you use. Amazon EC2 offers a variety of pricing options for different use cases.

- On-Demand Instances
- Reserved Instances
- EC2 Instance Savings Plans
- Spot Instances
- Dedicated Hosts

What are the EC2 purchase options?

- You can pay for it **ON DEMAND** (per hour or per second, depending on instance type and OS).Are ideal for short-term, irregular workloads that cannot be interrupted. No upfront costs or minimum contracts apply

- You can buy a **SAVINGS PLAN** (lower prices in exchange for a commitment of dollars per hour for 1 or 3 year term). This can save up to 72% of AWS cost. Reduce your EC2 instance costs when you make an hourly spend commitment to an instance family and Region for a 1-year or 3-year term. good option if you need flexibility in your Amazon EC2 usage over the duration of the commitment term.

- You can also have **RESERVED INSTANCES**. These are good when you know you will need a steady amount of resources that is unlikely to change. **Standard Reserved Instances**: good fit if you know the EC2 instance type and size you need for your steady-state applications and in which AWS Region you plan to run them. If you need to run your EC2 instances in different Availability Zones or different instance types, then **Convertible Reserved Instances** is a good choice.

- **Spot Instances** are ideal for workloads with flexible start and end times, or that can withstand interruptions. Spot Instances use unused Amazon EC2 computing capacity and offer you cost savings at up to 90% off of On-Demand prices.

- **The Dedicated Hosts** are physical servers with Amazon EC2 instance capacity that is fully dedicated to your use.
  You can use your existing per-socket, per-core, or per-VM software licenses to help maintain license compliance.

Scaling Amazon EC2

Scalability

involves beginning with only the resources you need and designing your architecture to automatically respond to changing demand by scaling out or in. As a result, you pay for only the resources you use.

Amazon EC2 Auto Scaling.

Is the AWS service that provides the scaling process to happen automatically for Amazon EC2 instances.

Amazon EC2 Auto Scaling enables you to automatically add or remove Amazon EC2 instances in response to changing application demand. By automatically scaling your instances in and out as needed, you can maintain a greater sense of application availability.

Within Amazon EC2 Auto Scaling, you can use two approaches:

- dynamic scaling
- predictive scaling.

Dynamic scaling responds to changing demand.

Predictive scaling automatically schedules the right number of Amazon EC2 instances based on predicted demand.

To scale faster, you can use dynamic scaling and predictive scaling together.

Example: Amazon EC2 Auto Scaling

In the cloud, computing power is a programmatic resource, so you can take a more flexible approach to the issue of scaling. By adding **Amazon EC2 Auto Scaling** to an application, you can add new instances to the application when necessary and terminate them when no longer needed.

Suppose that you are preparing to launch an application on Amazon EC2 instances. When configuring the size of your Auto Scaling group, you might set the minimum number of Amazon EC2 instances at one. This means that at all times, there must be at least one Amazon EC2 instance running.

1. When you create an Auto Scaling group, you can set **the minimum number of Amazon EC2 instances.** The minimum capacity is the number of Amazon EC2 instances that launch immediately after you have created the Auto Scaling group.
2. Next, you can set **the desired capacity at two Amazon EC2 instances** even though your application needs a minimum of a single Amazon EC2 instance to run.
3. The third configuration that you can set in an Auto Scaling group is t**he maximum capacity**. For example, you might configure the Auto Scaling group to scale out in response to increased demand, but only to a maximum of four Amazon EC2 instances.

Because Amazon EC2 Auto Scaling uses Amazon EC2 instances, you pay for only the instances you use, when you use them. You now have a cost-effective architecture that provides the best customer experience while reducing expenses.

Directing Traffic with Elastic Load Balancing

Elastic Load Balancing

- is the AWS service that automatically distributes incoming application traffic across multiple resources, such as Amazon EC2 instances.
- A load balancer serves as the single entry point for incoming web traffic to an Auto Scaling group, routing requests to multiple Amazon EC2 instances.
- As traffic fluctuates and instances are added or removed, the load balancer distributes the workload evenly among the available instances, preventing any single instance from being overwhelmed.
- While **Elastic Load Balancing** and **Amazon EC2 Auto Scaling** are separate services, they work together to ensure high performance and availability for applications running on Amazon EC2.

Messaging and Queuing

Monolithic applications and microservices

Applications are made of multiple components. The components communicate with each other to transmit data, fulfill requests, and keep the application running.

Monolithic application.

- An application with tightly coupled components. These components might include databases, servers, the user interface, business logic, and so on.
- In this approach to application architecture, if a single component fails, other components fail, and possibly the entire application fails.

microservices approach

- In a microservices approach, application components are loosely coupled.
- In this case, if a single component fails, the other components continue to work because they are communicating with each other.
- The loose coupling prevents the entire application from failing.

Amazon Simple Notification Service (Amazon SNS) and Amazon Simple Queue Service (Amazon SQS).

When designing applications on AWS, you can take a microservices approach with services and components that fulfill different functions.

Amazon Simple Notification Service (Amazon SNS)

- Is a publish/subscribe service. Using Amazon SNS topics, a publisher publishes messages to subscribers.
- In Amazon SNS, subscribers can be web servers, email addresses, AWS Lambda functions, or several other options.

Amazon Simple Queue Service (Amazon SQS)

- is a message queuing service.
- Using Amazon SQS, you can send, store, and receive messages between software components, without losing messages or requiring other services to be available.
- In Amazon SQS, an application sends messages into a queue. A user or service retrieves a message from the queue, processes it, and then deletes it from the queue.

Additional Compute Services

Serverless computing

- Serverless means that your code runs on servers, but you do not need to provision or manage these servers.
- With serverless computing, you can focus more on innovating new products and features instead of maintaining servers.
- Another benefit of serverless computing is the flexibility to scale serverless applications automatically.
- Serverless computing can adjust the applications' capacity by modifying the units of consumptions, such as throughput and memory.

An AWS service for serverless computing is AWS Lambda.

AWS Lambda

- is a service that lets you run code without needing to provision or manage servers.
- While using AWS Lambda, you pay only for the compute time that you consume. Charges apply only when your code is running. You can also run code for virtually any type of application or backend service, all with zero administration.

How AWS Lambda works

1. You upload your code to Lambda.
2. You set your code to trigger from an event source, such as AWS services, mobile applications, or HTTP endpoints.
3. Lambda runs your code only when triggered.
4. You pay only for the compute time that you use.

In AWS, you can also build and run containerized applications.

Containers

- Containers provide you with a standard way to package your application's code and dependencies into a single object.
- You can also use containers for processes and workflows in which there are essential requirements for security, reliability, and scalability.

Amazon Elastic Container Service (Amazon ECS)

is a highly scalable, high-performance container management system that enables you to run and scale containerized applications on AWS.

Amazon ECS supports **Docker containers**. **Docker** is a software platform that enables you to build, test, and deploy applications quickly. AWS supports the use of open-source Docker Community Edition and subscription-based Docker Enterprise Edition. With Amazon ECS, you can use API calls to launch and stop Docker-enabled applications.

Amazon Elastic Kubernetes Service (Amazon EKS)

- Amazon EKS is a fully managed service that you can use to run Kubernetes on AWS.
- Kubernetes is open-source software that enables you to deploy and manage containerized applications at scale. A large community of volunteers maintains Kubernetes, and AWS actively works together with the Kubernetes community. As new features and functionalities release for Kubernetes applications, you can easily apply these updates to your applications managed by Amazon EKS.

AWS Fargate

- AWS Fargate is a serverless compute engine for containers. It works with both Amazon ECS and Amazon EKS.
- When using AWS Fargate, you do not need to provision or manage servers. AWS Fargate manages your server infrastructure for you. You can focus more on innovating and developing your applications, and you pay only for the resources that are required to run your containers.

AWS Global Infrastructure

Selecting a Region

Four business factors to consider When determining the right Region for your services, data, and applications

- Compliance with Data governance and legal requirements
- Proximity to Users (Customers)
- Available services within a region
- Pricing

Compliance and Data Residency

- **Regulatory Requirements**: Ensure the region complies with local data residency regulations and standards.
- **Industry Standards**: Certain industries like finance and healthcare have specific compliance needs.

Proximity to Users

- **Latency**: Choose a region that is geographically closest to your end-users to reduce latency and improve performance.
- **User Experience**: Better user experience with faster response times due to lower latency.

Service Availability

- Sometimes, the closest Region might not have all the features that you want to offer to customers.

Pricing

- AWS pricing varies by region. Evaluate the cost implications for compute, storage, data transfer, etc., in different regions.

Availability Zones

- is a single data center or a group of data centers within a Region.
- Availability Zones are located tens of miles apart from each other. This is close enough to have low latency *(the time between when content requested and received)* between Availability Zones.
- However, if a disaster occurs in one part of the Region, they are distant enough to reduce the chance that multiple Availability Zones are affected.

You'll want to run your application/instance on at least 2 availability zones. In case something happens to one, you'll have another zone up and running.

Edge locations

- Is a site that Amazon CloudFront uses to store cached copies of your content closer to your customers for faster delivery.
- So if your data is in the Ohio Region but you have customers in China, you can cache a copy of your data somewhere close to China so your customers in China will receive your data faster.

Amazon CloudFront

- is a content delivery service. It uses a network of edge locations to cache content and deliver content to customers all over the world. When content is cached, it is stored locally as a copy. This content might be video files, photos, webpages, and so on.

How to Provision AWS Resources

Ways to interact with AWS services

- The Management Console
- The Command Line Interface (CLI)
- Software Development Kits (SDK).

AWS Management Console

- Is a web-based interface for accessing and managing AWS services.
- You can quickly access recently used services and search for other services by name, keyword, or acronym.
- The console includes wizards and automated workflows that can simplify the process of completing tasks.

You can also use the **AWS Console mobile application** to perform tasks such as monitoring resources, viewing alarms, and accessing billing information. Multiple identities can stay logged into the AWS Console mobile app at the same time.

AWS Command Line Interface (AWS CLI).

- Save time when making API requests
- Enables you to control multiple AWS services directly from the command line within one tool.
- AWS CLI is available for users on Windows, macos, and Linux.

By using AWS CLI, you can automate the actions that your services and applications perform through scripts. For example, you can use commands to launch an Amazon EC2 instance, connect an Amazon EC2 instance to a specific Auto Scaling group, and more.

The software development kits (sdks)

- SDKs make it easier for you to use AWS services through an API designed for your programming language or platform.
- SDKs enable you to use AWS services with your existing applications or create entirely new applications that will run on AWS.

AWS Elastic Beanstalk

With AWS Elastic Beanstalk, you provide code and configuration settings, and Elastic Beanstalk deploys the resources necessary to perform the following tasks:

- Adjust capacity
- Load balancing
- Automatic scaling
- Application health monitoring

AWS CloudFormation

- With AWS CloudFormation, you can treat your infrastructure as code. This means that you can build an environment by writing lines of code instead of using the AWS Management Console to individually provision resources.
- AWS CloudFormation provisions your resources in a safe, repeatable manner, enabling you to frequently build your infrastructure and applications without having to perform manual actions. It determines the right operations to perform when managing your stack and rolls back changes automatically if it detects errors.

Connectivity to AWS

AWS Networking

Amazon Virtual Private Cloud (Amazon VPC)

Is a networking service that you can use to establish boundaries around your AWS resources**.**

- Amazon VPC enables you to provision an isolated section of the AWS Cloud.
- In this isolated section, you can launch resources in a virtual network that you define.
- Within a virtual private cloud (VPC), you can organize your resources into subnets.
- A **subnet** is a section of a VPC that can contain resources such as Amazon EC2 instances.

Internet gateway

- To allow public traffic from the internet to access your VPC, you attach an internet gateway to the VPC.
- An **internet gateway** is a connection between a VPC and the internet. You can think of an internet gateway as being similar to a doorway that customers use to enter the coffee shop. Without an internet gateway, no one can access the resources within your VPC.

Virtual private gateway

- To access private resources in a VPC, you can use a virtual private gateway.
- The virtual private gateway is the component that allows protected internet traffic to enter into the VPC.

A virtual private gateway enables you to establish a virtual private network (VPN) connection between your VPC and a private network, such as an on-premises data center or internal corporate network. A virtual private gateway allows traffic into the VPC only if it is coming from an approved network.

AWS Direct Connect

is a service that lets you to establish a dedicated private connection between your data center and a VPC.

Subnets and Network Access Control Lists

- A subnet is a section of a VPC in which you can group resources based on security or operational needs. Subnets can be public or private.
- Public subnets contain resources that need to be accessible by the public, such as an online store’s website.

- Private subnets contain resources that should be accessible only through your private network, such as a database that contains customers' personal information and order histories.

- In a VPC, subnets can communicate with each other. For example, you might have an application that involves Amazon EC2 instances in a public subnet communicating with databases that are located in a private subnet.

Network traffic in a VPC

- When a customer requests data from an application hosted in the AWS Cloud, this request is sent as a packet. A ***packet*** is a unit of data sent over the internet or a network.

- It enters into a VPC through an internet gateway. Before a packet can enter into a subnet or exit from a subnet, it checks for permissions. These permissions indicate who sent the packet and how the packet is trying to communicate with the resources in a subnet.

- The VPC component that checks packet permissions for subnets is a ***network access control list (ACL)***.

Network ACLs

- A network ACL is a virtual firewall that controls inbound and outbound traffic at the subnet level.

- Each AWS account includes a default **network ACL**. When configuring your VPC, you can use your account's default network ACL or create custom network ACLs.

- By default, your account's default network ACL allows all inbound and outbound traffic, but you can modify it by adding your own rules. For custom network ACLs, all inbound and outbound traffic is denied until you add rules to specify which traffic to allow. Additionally, all network ACLs have an explicit deny rule. This rule ensures that if a packet doesn't match any of the other rules on the list, the packet is denied.

Stateless packet filtering

- Network ACLs perform **stateless packet filtering**. They remember nothing and check packets that cross the subnet border each way: inbound and outbound.

- When a packet response for that request comes back to the subnet, the network ACL does not remember your previous request. The network ACL checks the packet response against its list of rules to determine whether to allow or deny.

- After a packet has entered a subnet, it must have its permissions evaluated for resources within the subnet, such as Amazon EC2 instances.

- The VPC component that checks packet permissions for an Amazon EC2 instance is a **security group**.

Security group

- Is a virtual firewall that controls inbound and outbound traffic for an Amazon ec2 instance.
- By default, a security group denies all inbound traffic and allows all outbound traffic. You can add custom rules to configure which traffic should be allowed; any other traffic would then be denied
- If you have multiple Amazon EC2 instances within the same VPC, you can associate them with the same security group or use different security groups for each instance.

Stateful packet filtering

- Security groups perform stateful packet filtering. They remember previous decisions made for incoming packets.
- When a packet response for that request returns to the instance, the security group remembers your previous request. The security group allows the response to proceed, regardless of inbound security group rules.

***The Network ACL** is like a Passport Control Traffic officer checking traffic in and out of the country, while a Security Group checks traffic going into a building (but doesn't care about traffic going out).*

*The Security Group is also **statefull** (while the Network ACL is **stateless**), which means the Security Group can remember which packets are already cleared for entry and therefore does not need to check a packet's validity if it's already been accepted through the instance before. Meanwhile, the Network ACL checks a packet's validity every single time it passes through a subnet, regardless if it's already been accepted before.*

VPC component recall

A. **Private subnet:**
B. **Virtual private getaway:**
C. **Public subset:**
D. **AWS Direct Connect:**

Private Subnet

Purpose:

- is used to host resources that should not be directly accessible from the internet.

Key Characteristics:

- Resources in a private subnet can communicate with the internet via a Network Address Translation (NAT) gateway or NAT instance, but they are not directly accessible from the internet.
- Typically used for backend services, databases, and application servers that do not require direct internet access.
- Enhances security by isolating internal resources from external access.

Virtual Private Gateway

Purpose:

- is used to enable communication between your AWS VPC and your on-premises network over an IPsec VPN connection or AWS Direct Connect.

Key Characteristics:

- Acts as a VPN concentrator on the AWS side of a VPN connection.
- Allows secure, encrypted communication between your VPC and your on-premises data center or office network.
- Facilitates hybrid cloud setups where part of your infrastructure is on-premises and part is in the AWS cloud.

Public Subnet

Purpose:

- is used to host resources that need to be directly accessible from the internet.

Key Characteristics:

- Resources in a public subnet have direct access to the internet via an Internet Gateway (IGW).
- Typically used for front-end services such as web servers, bastion hosts, and load balancers.
- Ensures that resources requiring public access can receive traffic from and send traffic to the internet.

AWS Direct Connect

Purpose:

- is a network service that provides a dedicated, private connection between your on-premises network and your AWS VPC.

Key Characteristics:

- Provides high bandwidth and low latency connectivity, bypassing the public internet.
- Offers more consistent network performance compared to internet-based connections.
- Useful for workloads that require high data transfer rates or stable network performance, such as large-scale data migration, real-time applications, and hybrid cloud setups.
- Can reduce network costs over time by minimizing data transfer charges associated with internet usage.

Summary

- **Private Subnet**: Secure, isolated subnet for resources not needing direct internet access.
- **Virtual Private Gateway (VGW)**: Enables secure VPN or Direct Connect connections between VPC and on-premises networks.
- **Public Subnet**: Subnet for resources requiring direct internet access.
- **AWS Direct Connect**: Dedicated private network connection from on-premises to AWS for enhanced performance and reliability.

Global Networking

Amazon Route 53

- Route 53 is Amazon's DNS.
- It gives developers and businesses a reliable way to route end users to internet applications hosted in AWS.
- Connects user requests to infrastructure running in AWS (such as Amazon EC2 instances and load balancers). It can route users to infrastructure outside of AWS.
- Manage the DNS records for domain names. You can register new domain names directly in Route 53. You can also transfer DNS records for existing domain names managed by other domain registrars. This enables you to manage all of your domain names within a single location.

Domain Name System DNS

The Domain Name System translates website names into IP addresses that computers can read.

Instance Stores and Amazon Elastic Block Store (Amazon EBS)

Instance stores

- Provides temporary **block-level storage** for an Amazon EC2 instance. An instance store is disk storage that is physically attached to the host computer for an EC2 instance, and therefore has the same lifespan as the instance. When the instance is terminated, you lose any data in the instance store.
- ***Block-level storage volumes behave like physical hard drives.***

Amazon Elastic Block Store (Amazon EBS)

- is a service that provides block-level storage volumes that you can use with Amazon EC2 instances. If you stop or terminate an Amazon EC2 instance, all the data on the attached EBS volume remains available.
- To create an EBS volume, you define the configuration (such as volume size and type) and provision it. After you create an EBS volume, it can attach to an Amazon EC2 instance.
- Because EBS volumes are for data that needs to persist, it's important to back up the data. You can take incremental backups of EBS volumes by creating Amazon EBS snapshots.

An EBS snapshot

- is an incremental backup. This means that the first backup taken of a volume copies all the data. For subsequent backups, only the blocks of data that have changed since the most recent snapshot are saved.

- Incremental backups are different from full backups, in which all the data in a storage volume copies each time a backup occurs. The full backup includes data that has not changed since the most recent backup.

Object storage

In object storage, each object consists of **data**, **metadata**, and a **key**.

- The **data** might be an image, video, text document, or any other type of file.
- **Metadata** contains information about what the data is, how it is used, the object size, and so on.
- An object's **key** is its unique identifier.

when you modify a file in block storage, only the pieces that are changed are updated. When a file in object storage is modified, the entire object is updated.

Amazon Simple Storage Service (Amazon S3)

- is a service that provides object-level storage. Amazon S3 stores data as objects in buckets.
- You can upload any type of file to Amazon S3, such as images, videos, text files, and so on. For example, you might use Amazon S3 to store backup files, media files for a website, or archived documents. Amazon S3 offers unlimited storage space. The maximum file size for an object in Amazon S3 is 5 TB.
- When you upload a file to Amazon S3, you can set permissions to control visibility and access to it. You can also use the Amazon S3 versioning feature to track changes to your objects over time.

Amazon S3 storage classes

With Amazon S3, you pay only for what you use. You can choose from a range of storage classes to select a fit for your business and cost needs.

1. **S3 Standard**: Frequent access, high durability and availability.
2. **S3 Intelligent-Tiering**: Automatic tiering based on access patterns.
3. **S3 Standard-IA**: Lower-cost for infrequent access, but still needs rapid access.
4. **S3 One Zone-IA**: Lower-cost for infrequent access in a single AZ.
5. **S3 Glacier Instant Retrieval**: Low-cost, millisecond access for archival data.
6. **S3 Glacier Flexible Retrieval**: Very low-cost, flexible retrieval times.
7. **S3 Glacier Deep Archive**: Lowest cost, long-term archival storage.
8. **S3 Outposts**: On-premises object storage for low-latency and local data residency.

When selecting an Amazon S3 storage class, consider these two factors:

- How often you plan to retrieve your data
- How available you need your data to be

1. S3 Standard

Purpose: Designed for frequently accessed data.

Key Characteristics:

- High durability (99.999999999%, or 11 nines) and availability.
- Low latency and high throughput.
- Stores data in a minimum of three Availability Zones
- Suitable for a wide variety of use cases such as cloud applications, dynamic websites, content distribution, and data analytics.
- has a higher cost than other storage classes intended for infrequently accessed data and archival storage.

2. S3 Intelligent-Tiering

Purpose: Ideal for data with unknown or changing access patterns.

Key Characteristics:

- Automatically moves data between two access tiers (frequent and infrequent) based on changing access patterns.
- Designed to optimize storage costs by moving objects between access tiers when access patterns change.
- No retrieval charges in the frequent access tier, small charge for monitoring and automation.

3. S3 Standard-IA (Infrequent Access)

Purpose: Suitable for data that is accessed less frequently but requires rapid access when needed.

Key Characteristics:

- Lower storage cost compared to S3 Standard.
- Retrieval fee for accessing data.
- Ideal for long-term storage, backups, and disaster recovery.

4. S3 One Zone-IA (Infrequent Access)

Purpose: For infrequently accessed data that does not require multiple availability zone resilience.

Key Characteristics:

- Lower storage cost than S3 Standard-IA.
- Data stored in a single availability zone, so it's less resilient to AZ failures.
- Suitable for secondary backups or data that can be easily recreated.

5. S3 Glacier Instant Retrieval

Purpose: Designed for long-term storage of rarely accessed data that requires milliseconds retrieval.

Key Characteristics:

- Lower cost compared to S3 Standard and S3 Standard-IA.
- Suitable for medical images, news media archives, and other data that must be retained for regulatory compliance but accessed rarely.

6. S3 Glacier Flexible Retrieval (formerly S3 Glacier)

Purpose: Ideal for long-term storage of rarely accessed data.

Key Characteristics:

- Extremely low cost.
- Retrieval times ranging from minutes to hours.
- Suitable for archival storage and digital preservation.

7. S3 Glacier Deep Archive

Purpose: Lowest-cost storage class for data that is rarely accessed and has a retrieval time of hours.

Key Characteristics:

- Extremely low storage cost.
- Retrieval times of 12 hours or more.
- Ideal for long-term data archiving, such as compliance archives and digital preservation.

8. S3 Outposts

Purpose: For storing data on-premises using AWS Outposts hardware.

Key Characteristics:

- Provides object storage to your on-premises environments using the same APIs as S3.
- Ideal for workloads with local data residency requirements or where you need to process and store data locally for latency-sensitive applications.

Comparing Amazon EBS and Amazon S3

Amazon Elastic Block Store (EBS) and Amazon Simple Storage Service (S3) are both storage solutions offered by AWS, but they are designed for different use cases and have distinct characteristics.

Purpose and Use Cases

Amazon EBS:

Purpose: Provides persistent block storage for use with Amazon EC2 instances.

Use Cases:

- Primary storage for data that requires frequent and rapid access, such as databases, applications, and boot volumes.
- Suitable for transactional workloads that require consistent performance and low latency.

Amazon S3:

Purpose: Object storage service for storing and retrieving any amount of data at any time from anywhere on the web.

Use Cases:

- Storing large amounts of unstructured data such as backups, logs, media files, and big data analytics.
- Data archiving and long-term storage with various access frequency and durability requirements.

Storage Type

Amazon EBS:

Block Storage: Data is stored in fixed-sized blocks and managed by the operating system as a disk volume.

Volumes: Can be attached to EC2 instances and used like a physical hard drive.

Amazon S3:

Object Storage: Data is stored as objects in buckets, and each object consists of data, metadata, and a unique identifier.

Buckets: Used to organize and manage data stored as objects.

Performance and Latency

Amazon EBS:

Performance: Offers high IOPS and low latency suitable for running high-performance workloads.

Latency: Typically lower latency compared to S3, making it ideal for performance-sensitive applications.

Amazon S3:

Performance: Optimized for high throughput rather than low latency.

Latency: Slightly higher latency than EBS, suitable for less time-sensitive operations.

Durability and Availability

Amazon EBS:

Durability: Replicates data within the same Availability Zone to protect against component failure.

Availability: Designed for high availability within a single Availability Zone. For cross-AZ redundancy, snapshots can be taken and stored in S3.

Amazon S3:

Durability: 99.999999999% (11 nines) durability, replicating data across multiple facilities and Availability Zones within a region.

Availability: 99.99% availability SLA, designed for high availability and can be accessed from anywhere.

Scalability

Amazon EBS:

Scalability: Limited to the size of the volumes that can be attached to EC2 instances (up to 16 TiB per volume).

Elasticity: Volumes can be resized, but there is a maximum size limit per volume and per EC2 instance.

Amazon S3:

Scalability: Virtually unlimited storage capacity, can store any amount of data.

Elasticity: Automatically scales as you add more data without the need to provision additional capacity.

Pricing

Amazon EBS:

Pricing: Charged based on the provisioned size of the volume and the type of volume (e.g., SSD, HDD).

Snapshot Costs: Additional costs for snapshots stored in S3.

Amazon S3:

Pricing: Charged based on the amount of data stored, requests, and data transfer out of S3.

Storage Classes: Different pricing for different storage classes (e.g., Standard, Intelligent-Tiering, Glacier).

Data Access and Management

Amazon EBS:

Access: Data can be accessed only from within the same Availability Zone and must be attached to an EC2 instance.

Management: Managed through AWS Management Console, CLI, and APIs. Requires management of volume attachment and resizing.

Amazon S3:

Access: Data can be accessed from anywhere via the internet using REST APIs. Supports fine-grained access control with IAM policies and bucket policies.

Management: Easy management with features like versioning, lifecycle policies, and cross-region replication.

Backup and Recovery

Amazon EBS:

Backup: Supports point-in-time snapshots which are stored in S3. Snapshots can be used to create new EBS volumes.

Recovery: Snapshots enable recovery from failures, but recovery time depends on the size of the volume and the speed of restoring data.

Amazon S3:

Backup: Built-in durability and replication features reduce the need for additional backup.

Recovery: Objects are easily recoverable due to the high durability and availability of S3.

Amazon EBS and Amazon S3 Summary

Amazon EBS:

- Best for block storage needs with low latency and high IOPS.
- Suitable for primary storage of data that requires frequent and rapid access.
- Used for attaching to EC2 instances as persistent storage.

Amazon S3:

- Best for object storage with high durability and scalability.
- Suitable for storing large amounts of unstructured data, backups, and archives.
- Can be accessed from anywhere and offers various storage classes to optimize cost.

Amazon Elastic File System (Amazon EFS)

- is a scalable file system used with AWS Cloud services and on-premises resources. As you add and remove files, Amazon EFS grows and shrinks automatically. It can scale on demand to petabytes without disrupting applications.

File storage

- In **file storage**, multiple clients (such as users, applications, servers, and so on) can access data that is stored in shared file folders. In this approach, a storage server uses block storage with a local file system to organize files. Clients access data through file paths.
- Compared to block storage and object storage, **file storage** is ideal for use cases in which a large number of services and resources need to access the same data at the same time.

Comparing Amazon EBS and Amazon EFS

Amazon EBS (Elastic Block Store)

- is designed for use as block storage that can be attached to a single Amazon EC2 instance.
- This service provides high-performance storage that is well-suited for applications that require low latency and high IOPS, such as databases, enterprise applications, and boot volumes. EBS volumes can be scaled up to 16 TiB and offer various types to cater to different performance needs, including SSD and HDD options.
- EBS volumes are limited to the instance they are attached to and must be manually resized if additional capacity is required.

- Data stored in EBS is replicated within a single Availability Zone, providing high durability and availability within that zone. Pricing for EBS is based on the provisioned size and type of the volume.

Amazon EFS (Elastic File System)

- offers a fully managed network file system that can be accessed by multiple Amazon EC2 instances, ECS containers, EKS pods, and even on-premises servers.
- EFS is ideal for scenarios where multiple instances need to share access to a file system, such as content management, web serving, big data analytics, and shared development environments.
- Unlike EBS, EFS automatically scales up and down to accommodate the amount of data stored, providing virtually unlimited capacity without the need for manual intervention.
- It also offers scalable performance with two modes: General Purpose for latency-sensitive workloads and Max I/O for high-throughput workloads.
- Data in EFS is replicated across multiple Availability Zones within a region, ensuring high availability and durability. EFS uses a pay-as-you-go pricing model based on the amount of data stored, with options for **lifecycle management** to optimize costs.

What is S3 Lifecycle Management?

- Amazon S3 Lifecycle Management is where you can automatically move data to different tiers.
- For example, if you wanted data to be accessible in S3 Standard for 90 days, then moved to Standard-IA for another 30 days, and then moved to S3 Glacier after that, lifecycle policies can do this automatically.

Amazon Relational Database Service (Amazon RDS)

- Is a service that enables you to run relational databases in the AWS Cloud.
- Is a managed service that automates tasks such as hardware provisioning, database setup, patching, and backups.
- Supports all the major relational database management systems (RDBMS) like MySQL, PostgreSQL, and more.
- Amazon RDS also automatically handles things for you like backups, redundancy, and disaster recovery.
- You can integrate Amazon RDS with other services

Relational databases

- Data is stored in a way that relates it to other pieces of data.
- Use structured query language (SQL) to store and query data. This approach allows data to be stored in an easily understandable, consistent, and scalable way.

Amazon RDS engines provides different security options for:

- Encryption at rest (protecting data while it is stored)
- Encryption in transit (protecting data while it is being sent and received).

Amazon RDS database engines

- ❖ Amazon RDS is available on six database engines, which optimize for memory, performance, or input/output (I/O).

Supported database engines include:

- ❖ Amazon Aurora
- ❖ PostgreSQL
- ❖ MySQL
- ❖ MariaDB
- ❖ Oracle Database
- ❖ Microsoft SQL Server

Amazon Aurora

- ❖ is an enterprise-class relational database. It is compatible with MySQL and PostgreSQL relational databases. It is up to five times faster than standard MySQL databases and up to three times faster than standard PostgreSQL databases.

- ❖ Amazon Aurora helps to reduce your database costs by reducing unnecessary input/output (I/O) operations, while ensuring that your database resources remain reliable and available.

- ❖ Amazon Aurora is best for workloads that require high availability. It replicates six copies of your data across three Availability Zones and continuously backs up your data to Amazon S3.

Amazon DynamoDB

- ❖ Is a key-value database service. It delivers single-digit millisecond performance at any scale.

- ❖ Is a serverless, non-relational, NoSQL database, fully manage and automatic scaling. This is great for datasets that have variations from item to item. You can query this database through key-value pair relationships (think JSON or Python Dictionaries).

Amazon Redshift

- ❖ Is a **data warehousing** service that you can use for big data analytics. It offers the ability to collect data from many sources and helps you to understand relationships and trends across your data.

What is a Data Warehouse good for?

- ❖ A Data Warehouse can be good for historical data which we will want to run Business Intelligence Analytics on. Historical data is set and will never change, making a Data Warehouse a good place to store it.

AWS Database Migration Service

- Amazon DMS can migrate data from one source database to another target database (which can be the same type or different type).
- If the database types are the same, this is known as a homogenous migration, and is a one-step (one-click) process.
- If the source and target database types are different, this is known as a heterogeneous migration. You'll need to use the AWS Schema Conversion Tool first to convert the source schema and code to match that of the target database. Then you can use DMS to migrate into the target database.

AWS Database Migration Service (AWS DMS)

- Enables you to migrate relational databases, nonrelational databases, and other types of data stores.
- With AWS DMS, you move data between a **source** database and a **target** database. The source and target databases can be of the same type or different types. During the migration, your source database remains operational, reducing downtime for any applications that rely on the database.
- For example, suppose that you have a MySQL database that is stored on premises in an Amazon EC2 instance or in Amazon RDS. Consider the MySQL database to be your **source** database. Using AWS DMS, you could migrate your data to a target database, such as an Amazon Aurora database.

What are other use cases for Amazon DMS?

- You can create a test database migration, so you can test applications against production data without affecting production users.
- You can consolidate several databases into one database.
- You can do continuous replication to send ongoing copies of your data to target sources (instead of a one-time migration).

Additional database services

- Amazon DocumentDB
- Amazon Neptune
- Amazon Quantum Ledger Database (Amazon QLDB)
- Amazon Managed Blockchain
- Amazon ElastiCache
- Amazon DynamoDB Accelerator (DAX)

Amazon DocumentDB

- Is a document database service that supports MongoDB workloads.
- (MongoDB is a document database program.)

Amazon Neptune

- Is a graph database service.
- You can use Amazon Neptune to build and run applications that work with highly connected datasets, such as recommendation engines, fraud detection, and knowledge graphs

Amazon Quantum Ledger Database (Amazon QLDB)

- Is a ledger database service.
- You can use Amazon QLDB to review a complete history of all the changes that have been made to your application data.

Amazon Managed Blockchain

- Is a service that you can use to create and manage blockchain networks with open-source frameworks.
- Blockchain is a distributed ledger system that lets multiple parties run transactions and share data without a central authority.

Amazon ElastiCache

- Is a service that adds caching layers on top of your databases to help improve the read times of common requests.
- It supports two types of data stores: Redis and Memcached.

Amazon DynamoDB Accelerator (DAX)

- Is an in-memory cache for DynamoDB.
- It helps improve response times from single-digit milliseconds to microseconds.

AWS Security and Compliance

The AWS Shared Responsibility Model

- is a framework that delineates the security and compliance responsibilities between AWS and its customers, ensuring clarity on which security tasks are managed by AWS and which are managed by the customer.

AWS is responsible

- For the security **of** the cloud infrastructure, which includes the physical security of data centers, the security of hardware and software, and the overall networking and facilities management. This encompasses managing the security configurations and patching of managed services such as RDS, DynamoDB, and S3, and ensuring compliance with global standards like ISO, SOC, and PCI DSS.

Customers, on the other hand, are responsible

- For security **in** the cloud. This includes securing their data through encryption, managing user access via IAM policies, ensuring application security by patching and updating systems, and configuring network settings like VPCs and security groups. Customers must also ensure their use of AWS services complies with relevant legal and regulatory requirements, conducting audits and maintaining data privacy.

Shared Responsibility Model:

- **AWS Responsibility**: Security "of" the cloud (hardware, software, networking)

  Infrastructure, including hardware, software, networking, and facilities.

- **Customer Responsibility**: Security "in" the cloud (data, access management)

  Data, identity, and access management, OS, network, and firewall configuration.

User Permissions and Access

AWS Identity and Access Management (IAM):

Enables you to manage access to AWS services and resources securely. IAM gives you the flexibility to configure access based on your company's specific operational and security needs. You do this by using a combination of IAM features,

- **IAM Users**: Entities that can interact with AWS services.
- **IAM Groups**: Collection of users to apply policies.
- **IAM Roles**: Define permissions and can be assumed by users or services.
- **IAM Policies**: JSON documents that define permissions for users, groups, and roles.
- **MFA (Multi-Factor Authentication)**: Adds an additional layer of security. E.g password and then a second form of authentication, such as a random code sent to your phone.

IAM users

- An IAM user is an identity that you create in AWS. It represents the person or application that interacts with AWS services and resources. It consists of a name and credentials.
- You can create IAM users, and by default, will have no permissions.
- You should follow the principle of least privilege: A user should be granted access only to what they need.

What are IAM groups?

- IAM groups are groups of users with the same permissions (which are implemented via policies).

What are roles in IAM?

- Roles have associated permissions that allow or deny specific actions (and can be assumed for temporary amounts of time). They are basically temporary permissions for users.

IAM policies

- An IAM policy is a document that allows or denies permissions to AWS services and resources.
- IAM policies enable you to customize users' levels of access to resources. For example, you can allow users to access all of the Amazon S3 buckets within your AWS account, or only a specific bucket.

AWS account root user

- When you first create an AWS account, you begin with an identity known as the root user.
- Is accessed by signing in with the email address and password that you used to create your AWS account.
- The root user has access and control of any resource in the account. It is important to also implement MFA (multi-factor authentication).

Root User Best practice:

- Do not use the root user for everyday tasks.
- Instead, use the root user to create your first IAM user and assign it permissions to create other users.
- Then, continue to create other IAM users, and access those identities for performing regular tasks throughout AWS. Only use the root user when you need to perform a limited number of tasks that are only available to the root user. Examples of these tasks include changing your root user email address and changing your AWS support plan.

AWS Organizations

- AWS Organizations is a central location to manage multiple AWS accounts. It can also consolidate billing (using the primary account to pay for all accounts). You can also create hierarchical groupings of accounts. You will also have control over the AWS services and APIs that each account can access.
- You can use Service Control Policies (SCPs) to specify the maximum permissions for accounts in the organization. SCPs can applied to Organizations Units (OUs) or individual member accounts. to make it easier to manage accounts with similar business or security requirements.
- IAM policies can be applied to IAM users, IAM groups, and IAM roles.
- When you create an organization, AWS Organizations automatically creates a root, which is the parent container for all the accounts in your organization.

Compliance

Depending on your company's industry, you may need to uphold specific standards. An audit or inspection will ensure that the company has met those standards.

AWS Artifact

- Is a service that provides on-demand access to AWS security and compliance reports and select online agreements.
- Provides access to AWS security and compliance documents, such as AWS ISO certifications, Payment Card Industry (PCI) reports, and Service Organization Control (SOC) reports.

AWS Artifact consists of two main sections:

AWS Artifact Agreements:

- A place where you can sign an agreement with AWS regarding your use of certain types of information throughout AWS services
- A place where you can review, accept, and manage agreements for an individual account and for all your accounts in AWS Organizations

AWS Artifact Reports

- A place where a member of your company's development team can access more information about their responsibility for complying with certain regulatory standards.
- AWS Artifact Reports provide compliance reports from third-party auditors.
- You can provide the AWS audit artifacts to your auditors or regulators as evidence of AWS security controls.
- AWS Artifact Reports remains up to date with the latest reports released.

What is the Customer Compliance Center?

- Is a place that contains resources to help you learn more about compliance.

AWS Security Services:

- **AWS Shield:** DDoS (distributed denial-of-service) protection .
- **AWS WAF:** Web Application Firewall.
- **AWS KMS:** Key Management Service for creating and managing cryptographic keys.
- **AWS CloudTrail:** Logs API calls.
- **AWS CloudWatch:** Monitoring and management service.
- **AWS Inspector:** Automated security assessment.

Denial-of-Service Attacks

- A denial-of-service (DoS) attack is a deliberate attempt to make a website or application unavailable to users.
- For example, an attacker might flood a website or application with excessive network traffic until the targeted website or application becomes overloaded and is no longer able to respond. If the website or application becomes unavailable, this denies service to users who are trying to make legitimate requests.

Distributed denial-of-service (DDoS) attack

- Multiple sources are used to start an attack that aims to make a website or application unavailable. This can come from a group of attackers, or even a single attacker. The single attacker can use multiple infected computers (also known as “bots”) to send excessive traffic to a website or application.

AWS Shield

- Is a service that protects applications against DDoS attacks.

AWS Shield provides two levels of protection:

- **AWS Shield Standard** automatically protects all AWS customers at no cost. It protects your AWS resources from the most common, frequently occurring types of DDoS attacks. As network traffic comes into your applications, AWS Shield Standard uses a variety of analysis techniques to detect malicious traffic in real time and automatically mitigates it.
- **AWS Shield Advanced** is a paid service that provides detailed attack diagnostics and the ability to detect and mitigate sophisticated DDoS attacks. It also integrates with other services such as Amazon CloudFront, Amazon Route 53, and Elastic Load Balancing and even WAF (Web Application Firewall).

AWS WAF

- Uses a **web application firewall** to filter incoming traffic with the signatures of bad actors (it uses a web access control list, or web ACL). It has ML abilities and can detect threats as they evolve.
- AWS WAF works together with Amazon CloudFront and an Application Load Balancer.

AWS Key Management Service (AWS KMS)

- Enables you to perform encryption operations through the use of cryptographic keys.
- A cryptographic key is a random string of digits used for locking (encrypting) and unlocking (decrypting) data.
- Secure data while in storage (encryption at rest) and while it is transmitted, (encryption in transit).
- You can use AWS KMS to create, manage, and use cryptographic keys.

Amazon Inspector

- Helps to improve the security and compliance of applications by running automated security assessments.
- It checks applications for security vulnerabilities and deviations from security best practices, such as open access to Amazon EC2 instances and installations of vulnerable software versions.
- it provides you with a list of security findings. The list prioritizes by severity level, including a detailed description of each security issue and a recommendation for how to fix it.

Amazon GuardDuty

- Is a service that provides intelligent threat detection for your AWS infrastructure and resources.
- It identifies threats by continuously monitoring the network activity and account behavior within your AWS environment.
- Analyzes your metadata to identify threats. It runs independently and does not affect performance of your workloads.
- If GuardDuty detects any threats, you can review detailed findings about them from the AWS Management Console. Findings include recommended steps for remediation.

AWS Monitoring and Analytics

What is monitoring?

- Monitoring means observing systems, collecting **metrics**, and then using that data to make decisions.

Amazon CloudWatch

- Is a web service that enables you to monitor and manage various metrics and configure alarm actions based on data from those metrics. You can connect it to AWS SNS to send a message to a user to alert them as well.
- Uses metrics to represent the data points for your resources. AWS services send metrics to CloudWatch. CloudWatch then uses these metrics to create graphs automatically that show how performance has changed over time.
- You can create a CloudWatch dashboard to show us everything we want in one place. For example, you can use a CloudWatch dashboard to monitor the CPU utilization of an Amazon EC2 instance, the total number of requests made to an Amazon S3 bucket, and more.
- **CloudWatch** allows you to monitor your AWS infrastructure and your applications in real-time.

CloudWatch alarms

- Allows you to create alarms that automatically perform actions if the value of your metric has gone above or below a predefined threshold.
- **Example**: you could create a **CloudWatch** alarm that automatically stops an Amazon EC2 instance when the CPU utilization percentage has remained below a certain threshold for a specified period.
- When configuring the alarm, you can specify to receive a notification whenever this alarm is triggered.

AWS CloudTrail

- Is an API auditing tool
- Records API calls for your account. The recorded information includes the identity of the API caller, the time of the API call, the source IP address of the API caller, etc
- Every request made to AWS gets logged in the CloudTrail engine who made it, when, what was the IP address, and what was the response, was the request denied.
- You can think of CloudTrail as a “trail” of breadcrumbs (or a log of actions) that someone has left behind them.
- You can also enable **CloudTrail Insights**, which can detect unusual API activities in your account.

AWS Trusted Advisor

- Is a web service that inspects your AWS environment and provides real-time recommendations in accordance with AWS best practices.The inspection includes security checks, such as Amazon S3 buckets with open access permissions.

- Trusted Advisor compares its findings to AWS best practices in five categories: **cost optimization**, **performance**, **security**, **fault tolerance**, and **service limits**. For the checks in each category, Trusted Advisor offers a list of recommended actions and additional resources to learn more about AWS best practices.

For each category **AWS Trusted Advisor** indicates:

- The green check = **no problems**.
- The orange triangle = **recommended investigations**.
- The red circle = **recommended actions**.

AWS Pricing and Support

The AWS Free Tier

- Enables you to begin using certain services without having to worry about incurring costs for the specified period.

Three types of offers are available:

- Always Free
- 12 Months Free
- Trials

For 12 months after you first sign up for an AWS account, you can take advantage of offers in the 12 Months Free category. Examples of offers in this category include specific amounts of Amazon S3 Standard Storage, thresholds for monthly hours of Amazon EC2 compute time, and amounts of Amazon CloudFront data transfer out.

Here are some examples:

- AWS Lambda allows 1 million free requests per month (and up to 3.2 million seconds of compute time… that’s 888 hours).

- AWS S3 is free for 12 months for up to 5 GBs of storage.

- AWS Lightsail offers a 1 month trial of up to 750 hours of usage.

- A few more examples are AWS SageMaker, DynamoDB (25 GB of free storage per month), SNS, and more.

AWS Pricing Concepts

How AWS pricing works

AWS offers a range of cloud computing services with pay-as-you-go pricing.

- Pay for what you use.
- For each service, you pay for exactly the amount of resources that you actually use, without requiring long-term contracts or complex licensing.

- Pay less when you reserve.
- Some services offer reservation options that provide a significant discount compared to On-Demand Instance pricing.

- Pay less with volume-based discounts when you use more.
- Some services offer tiered pricing, so the per-unit cost is incrementally lower with increased usage.

AWS Pricing Calculator

- lets you explore AWS services and create an estimate for the cost of your use cases on AWS. You can organize your AWS estimates by groups that you define.
- When you have created an estimate, you can save it and generate a link to share it with others.

AWS pricing examples

For AWS Lambda

- You are charged based on the number of requests for your functions and the time that it takes for them to run.
- AWS Lambda allows 1 million free requests and up to 3.2 million seconds of compute time per month.
- You can save on AWS Lambda costs by signing up for a Compute Savings Plan. A Compute Savings Plan offers lower compute costs in exchange for committing to a consistent amount of usage over a 1-year or 3-year term. This is an example of **paying less when you reserve**.

With Amazon EC2

- You pay for only the compute time that you use while your instances are running.
- For some workloads, you can significantly reduce Amazon EC2 costs by using Spot Instances.
- For example, suppose that you are running a batch processing job that is able to withstand interruptions. Using a Spot Instance would provide you with up to 90% cost savings while still meeting the availability requirements of your workload.

For Amazon S3 pricing, consider the following cost components:

- **Storage** - You pay for only the storage that you use. You are charged the rate to store objects in your Amazon S3 buckets based on your objects' sizes, storage classes, and how long you have stored each object during the month.
- **Requests and data retrievals** - You pay for requests made to your Amazon S3 objects and buckets. For example, suppose that you are storing photo files in Amazon S3 buckets and hosting them on a website. Every time a visitor requests the website that includes these photo files, this counts towards requests you must pay for.
- **Data transfer** - There is no cost to transfer data between different Amazon S3 buckets or from Amazon S3 to other services within the same AWS Region. However, you pay for data that you transfer into and out of Amazon S3, with a few exceptions. There is no cost for data transferred into Amazon S3 from the internet or out to Amazon CloudFront. There is also no cost for data transferred out to an Amazon EC2 instance in the same AWS Region as the Amazon S3 bucket.
- **Management and replication** - You pay for the storage management features that you have enabled on your account's Amazon S3 buckets. These features include Amazon S3 inventory, analytics, and object tagging.

AWS Billing Cost Management Dashboard

- Use to pay your AWS bill, monitor your usage, and analyze and control your costs.
- Compare your current month-to-date balance with the previous month, and get a forecast of the next month based on current usage.
- View month-to-date spend by service.
- View Free Tier usage by service.
- Access Cost Explorer and create budgets.
- Purchase and manage Savings Plans.
- Publish AWS Cost and Usage Reports.

AWS Consolidated billing

- Enables you to receive a single bill for all AWS accounts in your organization. By consolidating, you can easily track the combined costs of all the linked accounts in your organization.
- The default maximum number of accounts allowed for an organization is 4, but you can contact AWS Support to increase your quota, if needed.

AWS Budgets

- You can create budgets to plan your service usage, service costs, and instance reservations.
- The information in AWS Budgets updates three times a day. This helps you to accurately determine how close your usage is to your budgeted amounts or to the AWS Free Tier limits.
- In AWS Budgets, you can also set custom alerts when your usage exceeds (or is forecasted to exceed) the budgeted amount.

AWS Cost Explorer

- Is a tool that lets you visualize, understand, and manage your AWS costs and usage over time.
- Its stores data up to 12 months
- You can filter by tag, group or service
- AWS Cost Explorer includes a default report of the costs and usage for your top five cost-accruing AWS services.
- You can apply custom filters and groups to analyze your data. For example, you can view resource usage at the hourly level.

AWS Support Plans

- **Basic**: Free with all accounts. Access to whitepapers, documentation, and support communities, You can also contact AWS for billing questions and service limit increases, You can use the AWS Personal Health Dashboard.
- **Developer**: Basic support plus business hours access. Email access to Cloud Support Associates during business hours, Best practice guidance, Client-side diagnostic tools, Building-block architecture support, which consists of guidance for how to use AWS offerings, features, and services together.
- **Business**: 24/7 support with faster response times. Email, chat, and phone support. Access to AWS Trusted Advisor, Use-case guidance to identify AWS offerings, features, and services that can best support your specific needs, Limited support for third-party software, such as common operating systems and application stack components.
- **Enterprise**: 24/7 support. All features of Business, plus a dedicated Technical Account Manager (TAM) and 5-minute response time for mission-critical workloads.
- **Enterprise On-Ramp:** Offers proactive engagement and advanced support for businesses needing guidance and optimization for their AWS environments. It includes a designated Technical Account Manager (TAM), 24/7 technical support with fast response times, regular well-architected reviews, cost and operational optimization, and comprehensive architectural guidance.

Technical Account Manager (TAM)

- The Enterprise On-Ramp and Enterprise Support plans include access to a Technical Account Manager (TAM).
- The **TAM** is your primary point of contact at AWS. If your company subscribes to Enterprise Support or Enterprise On-Ramp, your TAM educates, empowers, and evolves your cloud journey across the full range of AWS services.
- TAMs provide expert engineering guidance, help you design solutions that efficiently integrate AWS services, assist with cost-effective and resilient architectures, and provide direct access to AWS programs and a broad community of experts.

AWS Marketplace

- Is a digital catalog that includes thousands of software listings from independent software vendors. You can use AWS Marketplace to find, test, and buy software that runs on AWS.
- For each listing in AWS Marketplace, you can access detailed information on pricing options, available support, and reviews from other AWS customers.
- You can also explore software solutions by industry and use case.

AWS Marketplace offers products in several categories, such as

- Infrastructure Software,
- DevOps,
- Data Products,
- Professional Services,
- Business Applications,
- Machine Learning,
- Industries, and Internet of Things (IoT).

AWS Migration and Innovation

AWS Cloud Adoption Framework (AWS CAF)

- The AWS Cloud Adoption Framework is a framework to help users migrate on-premise data (or cloud data) into AWS.
- At the highest level, **the AWS Cloud Adoption Framework (AWS CAF)** organizes guidance into ***SIX AREAS OF FOCUS, CALLED PERSPECTIVES.***
- In general, the **Business**, **People**, and **Governance** Perspectives focus on business capabilities, whereas the **Platform**, **Security**, and **Operations** Perspectives focus on technical capabilities.

Six Core Perspectives of the AWS Cloud Adoption Framework

Business Perspective:

- Ensures that IT aligns with business needs and that IT investments link to key business results. It emphasizes creating a strong business case for cloud adoption, ensuring that cloud initiatives drive business value.
- **Common roles in the Business Perspective include**: Business managers, Finance managers, Budget owners, Strategy stakeholders

People Perspective:

- Focuses on organizational change management, preparing the workforce for cloud adoption through training and upskilling and identify gaps, defining new roles and responsibilities, and fostering a culture of change and innovation.
- **Common roles in the People Perspective include**: Human resources, Staffing, People managers.

Governance Perspective:

- Focuses on aligning IT strategy with business strategy to maximize value and minimize risks. It involves updating staff skills and processes for effective governance in the cloud, and managing and measuring cloud investments to evaluate business outcomes.
- **Common roles in the Governance Perspective include**: Chief Information Officer (CIO), Program managers, Enterprise architects, Business analysts, Portfolio managers

Platform Perspective:

- Includes principles and patterns for implementing new cloud solutions and migrating on-premises workloads. Helps you design, implement, and optimize your AWS infrastructure based on your business goals and perspectives .
- **Common roles in the Platform Perspective include:** Chief Technology Officer (CTO), IT managers, Solutions architects

Security Perspective:

- Ensures that the organization meets security objectives for visibility, auditability, control, and agility. Used to structure the selection and implementation of security controls that meet the organization’s needs.
- **Common roles in the Security Perspective include**: Chief Information Security Officer (CISO), IT security managers, IT security analysts

Operations Perspective:

- Helps enable, run, use, operate, and recover IT workloads to meet business stakeholder agreements. It defines daily, quarterly, and annual operations, aligning with business needs. It also helps stakeholders define current procedures and identify necessary process changes and training for successful cloud adoption.
- **Common roles in the Operations Perspective include**: IT operations managers, IT support managers

Migration Strategies

6 most common **migration strategies** you can use when migrating applications to the cloud.

Also known as the 6 R's of migration

Once you know exactly what you have in your existing environment, you decide which of the 6 R's are best for your migration:

- **Rehosting** (lift and shift): Simply lifting your app/infrastructure and shifting it to AWS without changes.
- **Replatforming** (lift, tinker, and shift): You may make a few cloud optimizations first before lifting and shifting to the cloud. Optimization is achieved without changing the core architecture of the application.
- **Retire:** Some parts of your infrastructure may no longer be needed
- **Retain:** Some parts of your infrastructure may need to be retained on-premise but not necessarily migrated to AWS (and can be retired eventually)
- **Repurchase:** Abandon legacy software and move to something new on the cloud
- **Refactoring/ Re-architecting:** You'll need to write new code, you'll need to add new features, and likely change the architecture of the code to make sure everything works once moved to AWS

AWS Snow Family

The **AWS Snow Family** is a collection of physical devices that help physically transport data into and out of AWS. For example, Amazon can ship you **AWS Snowcone**, you plug it into your data server or computer, copy all the data you'd like onto it, then ship it back to Amazon. Then, you will find all of your data copied into an S3 bucket, as an example.

AWS Snowcone

- Is a small, rugged, and secure edge computing and data transfer device.
- It features 2 CPUs, 4 GB of memory, and up to 14 TB of usable storage.

AWS Snowball offers two types of devices:

Snowball Edge Storage Optimized: This device is ideal for large-scale data migrations and recurring transfer workflows, as well as local computing with high capacity requirements.

- It provides 80 TB of HDD capacity for block volumes and Amazon S3-compatible object storage, along with 1 TB of SATA SSD for block volumes.
- The compute resources include 40 vCPUs and 80 GiB of memory, supporting Amazon EC2 sbe1 instances similar to C5 instances.

Snowball Edge Compute Optimized: Designed for intensive computing tasks such as machine learning, video analysis, analytics, and local computing stacks, this device offers robust computing capabilities.

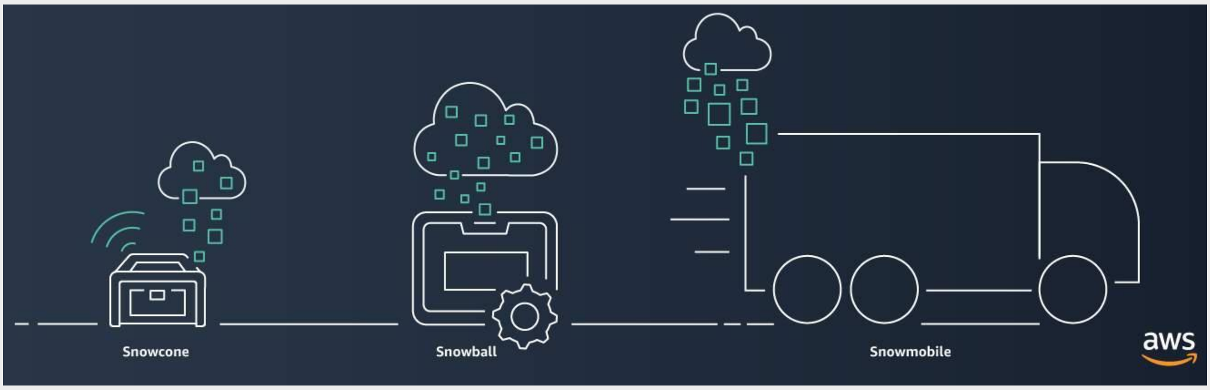
- It features 80 TB of usable HDD capacity for Amazon S3-compatible object storage or Amazon EBS-compatible block volumes, and 28 TB of usable NVMe SSD capacity specifically for Amazon EBS block volumes.

- The compute power includes 104 vCPUs, 416 GiB of memory, and optionally, an NVIDIA Tesla V100 GPU. It runs Amazon EC2 sbe-c and sbe-g instances equivalent to C5, M5a, G3, and P3 instances, providing flexibility for demanding workloads.

AWS Snowmobile

- Is an exabyte-scale data transfer service used to move large amounts of data to AWS.

- You can transfer up to 100 petabytes of data per Snowmobile, a 45-foot long ruggedized shipping container, pulled by a semi trailer truck.



Data Security in Transit and at Rest with AWS Snowball

Data Security in Transit:

- **Encryption**: All data transferred to and from AWS Snowball devices is encrypted using 256-bit encryption keys, managed by you through the AWS Key Management Service (KMS). This ensures that data is secure during transport.

- **Tamper-Resistant Enclosure**: The devices are physically rugged and tamper-resistant to protect against physical tampering during transit.

- **Chain of Custody**: AWS employs a strict chain of custody protocol, including tracking and logging the status of the Snowball device at every step of its journey, to ensure the device's integrity during transport.

Data Security at Rest:

- **Encryption and Tamper-Evident Design plus the following**

- **Access Controls**: Access to the data on the device is restricted through IAM roles and policies, ensuring that only authorized users can decrypt and access the data.

- **Secure Erase**: Once data transfer is complete and verified, AWS erases the data from the device using a secure erase process to ensure that no data remnants remain.

Innovate with AWS Services

AWS VMware Cloud

- When it comes to migrating onto AWS, The same VMware based infrastructure that you use on premises can in many cases, just be lifted up and dropped onto AWS via VMware Cloud on AWS.

Machine Learning And Artificial Intelligence

AWS has the broadest and deepest set of ML and AI services for your business. Allowing businesses to derive insights and automate processes more effectively.

Machine Learning (ML)

Traditional **machine learning (ML**) development is complex, expensive, time consuming, and error prone.

AWS offers ML Tools that any business can build upon with ease

- **Amazon SageMaker** to empower you to build, train, and deploy ML models quickly.
- **AWS DeepRacer** One of the newest branches of machine learning algorithms for developers to experiment with reinforcement learning
- **Amazon Augmented AI**, or **Amazon A2I**

Artificial intelligence

AWS offers a variety of services powered by artificial intelligence (AI).
For example, you can perform the following tasks:

- Convert speech to text with **Amazon Transcribe**.
- Discover patterns in text with **Amazon Comprehend**.
- Identify potentially fraudulent online activities with **Amazon Fraud Detector**.
- Build voice and text chatbots with **Amazon Lex**.
- Extract text and data from documents with **Amazon Textract.**

AWS IoT services

- AWS offers brand new technologies in things like Internet of Things. Enabling connected devices to communicate all around the world.

AWS Ground Station

- launch your own satellite and only pay for the satellite time you actually need

AWS Serverless applications

- With AWS, serverless refers to applications that don't require you to provision, maintain, or administer servers. You don't need to worry about fault tolerance or availability. AWS handles these capabilities for you.

- AWS Lambda is an example of a service that you can use to run serverless applications. If you design your architecture to trigger Lambda functions to run your code, you can bypass the need to manage a fleet of servers.

- Building your architecture with serverless applications enables your developers to focus on their core product instead of managing and operating servers.

The AWS Well-Architected Framework

- Is designed to enable architects, developers, and users of AWS to build secure, high-performance, resilient, and efficient infrastructure for their applications.
- It provides a way for you to consistently measure your architecture against best practices and design principles and identify areas for improvement.

Six pillars of The Well-Architected Framework

Operational Excellence:

- Focuses on running and monitoring systems to deliver business value, and, continually improving processes and procedures.
- Includes the ability to run workloads effectively and gain insights into their operations
- For example, automating changes with deployment pipelines or responding to events that are triggered.

Security:

- Security is priority number one at AWS. And this pillar exemplifies it by checking integrity of data.
- For example, protecting systems by using encryption.

Reliability:

- Focuses on recovery planning from failure or disruption of infrastructure or AWS, being able to scale horizontally to meet demand and automatically recovering from failure.
- For example Amazon DynamoDB disruption or an Amazon EC2 node failure

Performance efficiency:

- Choosing the right services (and the right types of a chosen service) for efficiency and to maintain that efficiency as demand changes and technologies evolve.
- For example, using the right Amazon EC2 type based on workload and memory requirements.

Cost optimization:

- Optimizing for full cost. This is controlling where money is spent.
- And, for example, checking if you have overestimated your Amazon EC2 server size. You can then lower cost by choosing a more cost-effective size.

Sustainability:

- Reducing energy consumption by minimizing the total resources needed an maximizing the benefits from the resources you are using to continually improve sustainability

What is the AWS Well-Architected Tool?

The Well-Architected Tool is a tool you can run against your framework which can help check if your architecture meets the Well-Architected Framework standards (and can note areas of improvement).

Advantages of cloud computing

Trade upfront expense for variable expense.

- Upfront expenses include data centers, physical servers, and other resources that you would need to invest in before using computing resources.
- Instead of investing heavily in data centers and servers before you know how you're going to use them, you can pay only when you consume computing resources.

Benefit from massive economies of scale.

- Economies of scale translate into lower pay-as-you-go prices.

Stop guessing capacity.

- With cloud computing, you don't have to predict how much infrastructure capacity you will need before deploying an application.
- With AWS, you don't need to guess your capacity. Instead, provision the resources you need for the now and then use the scaling mechanisms for each resource to scale up or down based on the real life capacity it needs from day to day as scaling can take minutes with AWS instead of weeks or months with on-premises resources.

Increase speed and agility

- The flexibility of cloud computing makes it easier for you to develop and deploy applications.
- This flexibility also provides your development teams with more time to experiment and innovate.

Stop spending money running and maintaining data centers.

- Cloud computing in data centers often requires you to spend more money and time managing infrastructure and servers.
- A benefit of cloud computing is the ability to focus less on these tasks and more on your applications and customers.

Global Reach

- With AWS's global infrastructure, businesses can deploy applications globally with low latency, reaching customers in different regions efficiently. This global footprint supports international expansion and improves user experience.

Additional Notes

What are shared controls?

Shared controls are controls where both AWS and the user share responsibility, such as:

- Patch management
- Configuration management
- Awareness and training

Note about S3

- S3 cannot be scaled manually, it is instead scaled automatically. It can also provide unlimited storage for any type of data, but each individual object does have a maximum size of 5 TB. It is “infinitely scalable.”

Why use Amazon ElastiCache?

- ElastiCache is a web service that makes it easy to deploy and manage an in-memory data store or cache in the cloud. It’s great for applications that need fast retrieval of data (querying a database is always slower than getting a copy of that data in the cache).

What are the 3 AWS Cloud Computing Models?

The 3 AWS Cloud Computing Models are:

- Infrastructure as a Service (IaaS)
- Platform as a Service (PaaS)
- Software as a Service (SaaS)

Note: Networking is a part of Infrastructure as a Service.

What is an AWS-Managed Service?

- An AWS-managed service is a service where AWS is responsible for the operational and maintenance burdens of running them.
- Examples are services like DynamoDB, Lambda, Elastic MapReduce, S3, and more.
- IAM , VPC, and EC2 are NOT examples, since you are the one responsible for managing them.

What is Infrastructure Event Management?

- Infrastructure Event Management is a short-term service you can get from AWS Support (which is included in the Enterprise Support plan and available for purchase in the Business Support plan) that helps provide architectural guidance for big traffic events (like product launches and advertising launches).

Summary

Cloud Technology and Services

1. Compute:

- **EC2**: Virtual servers.
- **Lambda**: Serverless compute.
- **Elastic Beanstalk**: Managed service for deploying applications.

EC2 Instance Types: Varieties include General Purpose, Compute Optimized, Memory Optimized, Storage Optimized, and Accelerated Computing.

Elastic Load Balancing (ELB): Distributes incoming application traffic across multiple targets.

2. Storage:

- **S3**: Object storage.
- **EBS**: Block storage for EC2.
- **Glacier**: Low-cost archival storage.
- **S3 Storage Classes**: Standard, Intelligent-Tiering, Standard-IA, One Zone-IA, Glacier, Glacier Deep Archive.
- **S3 Lifecycle Policies**: Define actions to transition or expire objects based on age.

3. Databases:

- **RDS**: Offers automated backups, snapshots, and read replicas. ***Managed relational database***.
- **DynamoDB**: Provides fast and flexible NoSQL database service.
- **Redshift Spectrum**: Query data in your S3 data lake without moving it, ***Data warehousing.***

4. Networking:

- **VPC Subnets**: Public and private subnets.
- **NAT Gateway**: Enables instances in a private subnet to connect to the internet.
- **VPC Peering**: Connects VPCs across regions.
- **Elastic IPs**: Static IPv4 addresses designed for dynamic cloud computing.
- **VPC**: Isolated network.
- **Route 53**: DNS and domain registration.
- **CloudFront**: Content delivery network.

Billing and Pricing

1. Cost Management:

AWS Pricing Models:

- **On-Demand:** Pay for compute or database capacity with no long-term commitments.
- **Reserved Instances**: Up to 75% savings by reserving instances for 1 or 3 years.
- **Spot Instances**: Bid for unused EC2 capacity, offering savings of up to 90%.
- **Savings Plans**: Flexible pricing model that provides savings on AWS usage.

2. Cost Management Tools:

- **AWS Cost Explorer:** Visualize and manage AWS costs and usage over time.
- **AWS Budgets**: Set custom cost and usage budgets and receive alerts.
- **AWS Pricing Calculator**: Estimate the cost of AWS services.

3. Support Plans:

Basic: Free with all accounts. Access to whitepapers, documentation, and support communities.

Developer: Basic support plus business hours access. Email access to Cloud Support Associates during business hours.

Business: 24/7 support with faster response times. Email, chat, and phone support. Access to AWS Trusted Advisor.

Enterprise: 24/7 support. All features of Business, plus a dedicated Technical Account Manager (TAM).

AWS Well-Architected Framework

1. Pillars:

- **Operational Excellence**: Focus on operations and monitoring to continually improve processes.
- **Security**: Protect data, systems, and assets with detective controls, infrastructure protection, data protection, and incident response.
- **Reliability**: Ensure workloads recover from failures and meet customer demands.
- **Performance Efficiency**: Use IT and computing resources efficiently.
- **Cost Optimization**: Avoid unnecessary costs with the right resource selection and pricing model.

Additional Topics

- 1. **AWS Marketplace**: A curated digital catalog for software.

- 2. **AWS Free Tier:** Offers free usage for new customers to experiment with AWS services.
- 3. **AWS Management Console**: A web-based interface for accessing and managing AWS services.

Exam Preparation Tips

Official AWS Training:

- **AWS Cloud Practitioner Essentials**: Comprehensive training module provided by AWS.

Whitepapers and Documentation:

- **Overview of Amazon Web Services**: Understand the broad scope of AWS services.
- **AWS Well-Architected Framework**: Learn best practices for architecting in the cloud.
- **AWS Security Best Practices**: Detailed practices for securing AWS environments.

Hands-On Practice:

- Use the AWS Free Tier to gain practical experience.
- Complete tutorials and labs on the AWS website.

Mock Exams and Practice Questions:

- Use AWS's official practice exams and sample questions.
- Platforms like Whizlabs, A Cloud Guru, and Udemy offer practice tests.

Join Study Groups and Online Communities:

- Participate in forums like Reddit's r/AWSCertifications or LinkedIn groups.
- Join study groups to share knowledge and resources.