



1.cloud computing:.....

- Cloud computing is the on-demand delivery of IT resources over the internet with pay-as-you-go pricing.
- cloud computing is a range of services delivered over the internet, or “the cloud.” It means using remote servers to store and access data instead of relying on local hard drives and private datacenters.
- The three cloud computing deployment models are cloud-based, on-premises, and hybrid.
 1. cloud-based deployment
 2. on-premises deployment
 3. hybrid deployment

Benefits of cloud computing:

- Trade upfront expense for variable expense
- Stop spending money to run and maintain data centers
- Stop guessing capacity
- Benefit from massive economies of scale
- Increase speed and agility
- Go global in minute.

Questions:

1.What is cloud computing?

- Backing up files that are stored on desktop and mobile devices to prevent data loss
- Deploying applications connected to on-premises infrastructure
- Running code without needing to manage or provision servers
- ✓ On-demand delivery of IT resources and applications through the internet with pay-as-you-go pricing

2.What is another name for on-premises deployment?

- ✓ Private cloud deployment
- Cloud-based application
- Hybrid deployment
- AWS Cloud

3.How does the scale of cloud computing help you to save costs?

- You do not have to invest in technology resources before using them.
- ✓ The aggregated cloud usage from a large number of customers results in lower pay-as-you-go prices.
- Accessing services on-demand helps to prevent excess or limited capacity.
- You can quickly deploy applications to customers and provide them with low latency.

2.COMPUTE IN THE CLOUD.....

Amazon Elastic Compute Cloud (Amazon EC2):

- When you're working with AWS, those servers are virtual. And the service you use to gain access to virtual servers is called EC2.
- Using EC2 for compute is highly flexible, cost effective, and quick when you compare it to running your own servers on premises in a data center that you own. The time and money it takes to get up and running with on-premises resources is fairly high.
- [Amazon Elastic Compute Cloud \(Amazon EC2\)](#) provides secure, resizable compute capacity in the cloud as Amazon EC2 instances.
- Imagine you are responsible for the architecture of your company's resources and need to support new websites. With traditional on-premises resources, you have to do the following:
 - Spend money upfront to purchase hardware.
 - Wait for the servers to be delivered to you.
 - Install the servers in your physical data center.
 - Make all the necessary configurations.
- By comparison, 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.

Amazon EC2 instance types:

- [Amazon EC2 instance types](#) are optimized for different tasks. When selecting an instance type, consider the specific needs of your workloads and applications. This might include requirements for compute, memory, or storage capabilities.
- **General purpose instances**
General purpose instances 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
- **Compute optimized instances**
Compute optimized instances are ideal for compute-bound applications that benefit from high-performance processors. Like general purpose instances, you can use compute optimized instances for workloads such as web, application, and gaming servers.
- **Memory optimized instances**
Memory optimized instances are designed to deliver fast performance for workloads that process large datasets in memory. In computing, memory is a temporary storage area. It holds all the data and instructions that a central processing unit (CPU) needs to be able to complete actions. Before a computer program or application is able to run, it is loaded from storage into memory. This preloading process gives the CPU direct access to the computer program.
- **Accelerated computing instances**
Accelerated computing instances use hardware accelerators, or coprocessors, to perform some functions more efficiently than is possible in software running on CPUs. Examples of these functions include floating-point number calculations, graphics processing, and data pattern matching.

- **Storage optimized instances**

Storage optimized instances are designed for workloads that require high, sequential read and write access to large datasets on local storage.

Examples of workloads suitable for storage optimized instances include distributed file systems, data warehousing applications, and high-frequency online transaction processing (OLTP) systems.

- Ideal for high-performance databases-----→ Memory optimized
- Suitable for data warehousing applications-----→ Storage optimized
- Balances compute, memory, and networking resources-----→ General purpose
- Offers high-performance processors-----→ Compute optimized

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. For example, if your use case can withstand interruptions, you can save with Spot Instances. You can also save by committing early and locking in a minimum level of use with Reserved Instances.
- **On-Demand**
On-Demand Instances are ideal for short-term, irregular workloads that cannot be interrupted. No upfront costs or minimum contracts apply. The instances run continuously until you stop them, and you pay for only the compute time you use
- **Amazon EC2 Savings Plans**
AWS offers Savings Plans for several compute services, including Amazon EC2. **Amazon EC2 Savings Plans** enable you to reduce your compute costs by committing to a consistent amount of compute usage for a 1-year or 3-year term. This term commitment results in savings of up to 66% over On-Demand costs.
- **Reserved Instances**
Reserved Instances are a billing discount applied to the use of On-Demand Instances in your account. You can purchase Standard Reserved and Convertible Reserved Instances for a 1-year or 3-year term, and Scheduled Reserved Instances for a 1-year term. You realize greater cost savings with the 3-year option.
- **Spot Instances**
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.
- **Dedicated Hosts**
Dedicated Hosts are physical servers with Amazon EC2 instance capacity that is fully dedicated to your use.

Scalability

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. You don't have to worry about a lack of computing capacity to meet your customers' needs.

If you wanted the scaling process to happen automatically, which AWS service would you use? The AWS service that provides this functionality for Amazon EC2 instances is **Amazon EC2 Auto Scaling**.

Amazon EC2 Auto Scaling

If you've tried to access a website that wouldn't load and frequently timed out, the website might have received more requests than it was able to handle. This situation is similar to waiting in a long line at a coffee shop, when there is only one barista present to take orders from customers.

Example: Amazon EC2 Auto Scaling

In the cloud, computing power is a programmable 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.

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 acts as a single point of contact for all incoming web traffic to your Auto Scaling group. This means that as you add or remove Amazon EC2 instances in response to the amount of incoming traffic, these requests route to the load balancer first. Then, the requests spread across multiple resources that will handle them. For example, if you have multiple Amazon EC2 instances, Elastic Load Balancing distributes the workload across the multiple instances so that no single instance has to carry the bulk of it.

Example: Elastic Load Balancing

Low-demand period

Here's an example of how Elastic Load Balancing works. Suppose that a few customers have come to the coffee shop and are ready to place their orders.

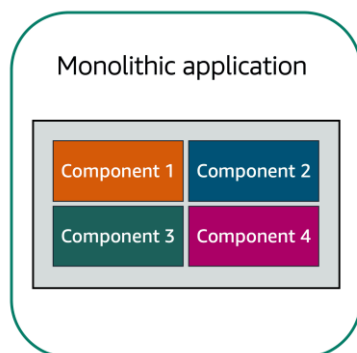
If only a few registers are open, this matches the demand of customers who need service. The coffee shop is less likely to have open registers with no customers. In this example, you can think of the registers as Amazon EC2 instances.

High-demand period

Throughout the day, as the number of customers increases, the coffee shop opens more registers to accommodate them. In the diagram, the Auto Scaling group represents this.

Additionally, a coffee shop employee directs customers to the most appropriate register so that the number of requests can evenly distribute across the open registers. You can think of this coffee shop employee as a load balancer.

Monolithic applications and microservices

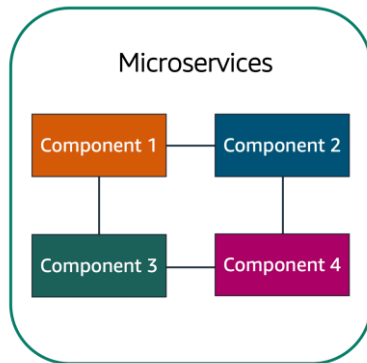


- Applications are made of multiple components. The components communicate with each other to transmit data, fulfill requests, and keep the application running.

Suppose that you have an application with tightly coupled components. These components might include databases, servers, the user interface, business logic, and so on. This type of architecture can be considered a **monolithic application**.

In this approach to application architecture, if a single component fails, other components fail, and possibly the entire application fails.

To help maintain application availability when a single component fails, you can design your application through a **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.

When designing applications on AWS, you can take a microservices approach with services and components that fulfill different functions. Two services facilitate application integration: Amazon Simple Notification Service (Amazon SNS) and Amazon Simple Queue Service (Amazon SQS).

Amazon Simple Notification Service (Amazon SNS)

Amazon Simple Notification Service (Amazon SNS) is a publish/subscribe service. Using Amazon SNS topics, a publisher publishes messages to subscribers. This is similar to the coffee shop; the cashier provides coffee orders to the barista who makes the drinks.

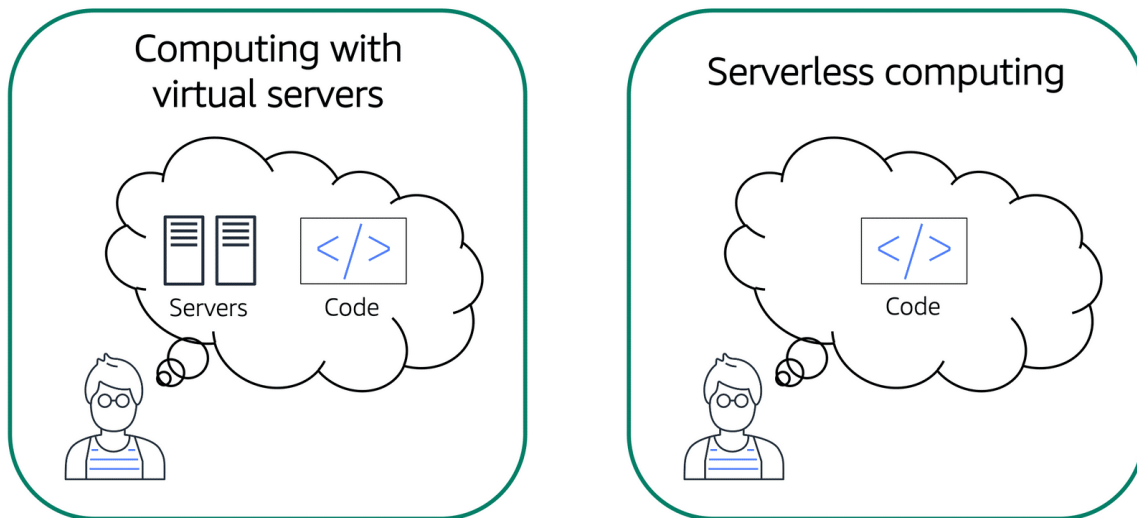
In Amazon SNS, subscribers can be web servers, email addresses, AWS Lambda functions, or several other options.

Serverless computing

Earlier in this module, you learned about Amazon EC2, a service that lets you run virtual servers in the cloud. If you have applications that you want to run in Amazon EC2, you must do the following:

- Provision instances (virtual servers).

- Upload your code.
→ Continue to manage the instances while your application is running.



The term “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

[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.

For example, a simple Lambda function might involve automatically resizing uploaded images to the AWS Cloud. In this case, the function triggers when uploading a new image.

Amazon Elastic Container Service (Amazon ECS)

[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 Elastic Kubernetes Service \(Amazon EKS\)](#) 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.

Questions:

1. You want to use an Amazon EC2 instance for a batch processing workload. What would be the best Amazon EC2 instance type to use?
 - General purpose
 - Memory optimized
 - ✓ Compute optimized
 - Storage optimized
2. What are the contract length options for Amazon EC2 Reserved Instances? (Select TWO.)
 - ✓ 1 year
 - 2 years
 - ✓ 3 years
 - 4 years
 - 5 years

3. You have a workload that will run for a total of 6 months and can withstand interruptions. What would be the most cost-efficient Amazon EC2 purchasing option?

- Reserved Instance
- ✓ Spot Instance
- Dedicated Instance
- On-Demand Instance

4. Which process is an example of Elastic Load Balancing?

- ✓ Ensuring that no single Amazon EC2 instance has to carry the full workload on its own
- Removing unneeded Amazon EC2 instances when demand is low
- Adding a second Amazon EC2 instance during an online store's popular sale
- Automatically adjusting the number of Amazon EC2 instances to meet demand

5. You want to deploy and manage containerized applications. Which service should you use?

- AWS Lambda
- Amazon Simple Notification Service (Amazon SNS)
- Amazon Simple Queue Service (Amazon SQS)
- ✓ Amazon Elastic Kubernetes Service (Amazon EKS)

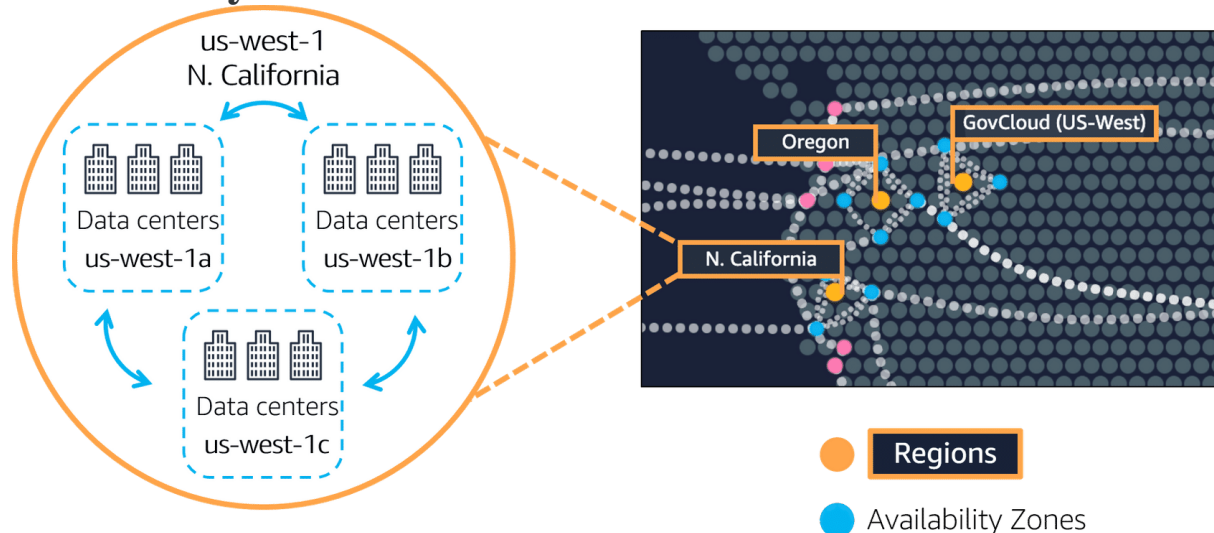
3.GLOBAL INFRA & RELIABILITY..

AWS global infrastructure

four key factors to choose a Region:

- Compliance,
- proximity,
- feature availability,
- and pricing.

Availability Zones



An **Availability Zone** 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.

So here's the key points. Number one,

Regions are **1.geographically isolated areas,**

where you can access services needed to run your enterprise.

2.Regions contain Availability Zones,

that allow you to run across physically separated buildings, tens of miles of separation, while keeping your application logically unified. Availability Zones help you solve high availability and disaster recovery scenarios, without any additional effort on your part,

3.AWS Edge locations run Amazon

CloudFront to help get content closer to your customers, no matter where they are in the world.

Edge locations

An **edge location** is a site that Amazon CloudFront uses to store cached copies of your content closer to your customers for faster delivery.

How to provision AWS resources:

For example, you can launch an EC2 instance or you can create an AWS Lambda function. Each of those would be different requests and different API calls to AWS. You can use the

Ways to interact with AWS services

1. **AWS Management Console,**

2. **the AWS Command Line Interface,**

3. **the AWS Software Development Kits,**

or various other tools like AWS CloudFormation, to create requests to send to AWS APIs to create and manage AWS resources.

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.

Amazon Virtual Private Cloud (Amazon VPC)

Imagine the millions of customers who use AWS services. Also, imagine the millions of resources that these customers have created, such as Amazon EC2 instances. Without boundaries around all of these resources, network traffic would be able to flow between them unrestricted.

A networking service that you can use to establish boundaries around your AWS resources is [Amazon Virtual Private Cloud \(Amazon VPC\)](#).

4.NETWORKING.....

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.

Subnets and network access control lists

WS has a wide range of tools that cover every layer of security: network hardening, application security, user identity, authentication and authorization, distributed denial-of-service or DDoS prevention, data integrity, encryption, much more.

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 access control lists (ACLs)

A network access control list (ACL) is a virtual firewall that controls inbound and outbound traffic at the subnet level.

For example, step outside of the coffee shop and imagine that you are in an airport. In the airport, travelers are trying to enter into a different country. You can think of the travelers as packets and the passport control officer as a network ACL. The passport control officer checks travelers' credentials when they are both entering and exiting out of the country. If a traveler is on an approved list, they are able to get through. However, if they are not on the approved list or are explicitly on a list of banned travelers, they cannot come in.

Domain Name System (DNS)

Suppose that AnyCompany has a website hosted in the AWS Cloud. Customers enter the web address into their browser, and they are able to access the website. This

happens because of **Domain Name System (DNS)** resolution. DNS resolution involves a customer DNS resolver communicating with a company DNS server.

Amazon Route 53

[Amazon Route 53](#) is a DNS web service. It gives developers and businesses a reliable way to route end users to internet applications hosted in AWS.

Amazon Route 53 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.

5.STORAGES.....

Storages: 1.EBS(Elastic Block Storage)

2.S3(Simple Storage Service)

3.EFS(Elastic File System)

Instance stores

Block-level storage volumes behave like physical hard drives.

An [instance store](#) 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.

[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.

You can take incremental backups of EBS volumes by creating Amazon EBS snapshots.

[Amazon Simple Storage Service \(Amazon S3\)](#) is a service that provides object-level storage. Amazon S3 stores data as objects in buckets.

- 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.

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.

[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.

DATABASEs:

Relational databases

In a **relational database**, data is stored in a way that relates it to other pieces of data.

An example of a relational database might be the coffee shop's inventory management system. Each record in the database would include data for a single item, such as product name, size, price, and so on. Relational databases use **structured query language (SQL)** to store and query data

Amazon Relational Database Service (Amazon RDS) is a service that enables you to run relational databases in the AWS Cloud.

Amazon RDS is a managed service that automates tasks such as hardware provisioning, database setup, patching, and backups.

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**
 - [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.

DynamoDB: is a non-relational, NoSQL database. It is purpose built. Meaning it has specific use cases, and it isn't the best fit for every workload out there. It has millisecond response time. It's fully managed, and it's highly scalable.

Amazon DynamoDB

[Amazon DynamoDB](#) is a key-value database service. It delivers single-digit millisecond performance at any scale.

Amazon Redshift

[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.

AWS Database Migration Service (AWS DMS)

[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.

Additional database services

1. Amazon DocumentDB

2. Amazon Neptune

3. Amazon Quantum Ledger Database (Amazon QLDB)

4.Amazon Managed Blockchain

5.Amazon ElastiCache

6.Amazon DynamoDB Accelerator

6.Security.....

- Explain the benefits of the shared responsibility model.
- Describe multi-factor authentication (MFA).
- Differentiate between the AWS Identity and Access Management (IAM) security levels.
- Explain the main benefits of AWS Organizations.
- Describe security policies at a basic level.
- Summarize the benefits of compliance with AWS.
- Explain additional AWS security services at a basic level.

The shared responsibility model, AWS controls security of the cloud and customers control security in the cloud. We, as AWS, control the data centers, security of our services, and all the layers in this section.

The AWS shared responsibility model

Throughout this course, you have learned about a variety of resources that you can create in the AWS Cloud. These resources include Amazon EC2 instances, Amazon S3 buckets, and Amazon RDS databases. Who is responsible for keeping these resources secure: you (the customer) or AWS?

The answer is both. The reason is that you do not treat your AWS environment as a single object. Rather, you treat the environment as a collection of parts that build upon each other. AWS is responsible for some parts of your environment and you (the customer) are responsible for other parts. This concept is known as the [shared responsibility model](#).

The shared responsibility model divides into **customer responsibilities** (commonly referred to as “security in the cloud”) and **AWS responsibilities** (commonly referred to as “security of the cloud”).

AWS Identity and Access Management (IAM)

[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, which are explored in detail in this lesson:

- IAM users, groups, and roles
- IAM policies
- Multi-factor authentication

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.

IAM policies

An **IAM policy** is a document that allows or denies permissions to AWS services and resources.

IAM groups

An IAM group is a collection of IAM users. When you assign an IAM policy to a group, all users in the group are granted permissions specified by the policy.

IAM roles

In the coffee shop, an employee rotates to different workstations throughout the day. Depending on the staffing of the coffee shop, this employee might perform several duties: work at the cash register, update the inventory system, process online orders, and so on.

When the employee needs to switch to a different task, they give up their access to one workstation and gain access to the next workstation. The employee can easily switch between workstations, but at any given point in time, they can have access to only a single workstation. This same concept exists in AWS with IAM roles.

An IAM role is an identity that you can assume to gain temporary access to permissions.

Multi-factor authentication

Have you ever signed in to a website that required you to provide multiple pieces of information to verify your identity? You might have needed to provide your password and then a second form of authentication, such as a random code sent to your phone. This is an example of [multi-factor authentication](#).

In IAM, multi-factor authentication (MFA) provides an extra layer of security for your AWS account.

AWS Organizations

Suppose that your company has multiple AWS accounts. You can use [AWS Organizations](#) to consolidate and manage multiple AWS accounts within a central location.

When you create an organization, AWS Organizations automatically creates a **root**, which is the parent container for all the accounts in your organization.

In AWS Organizations, you can centrally control permissions for the accounts in your organization by using [service control policies \(SCPs\)](#). SCPs enable you to place restrictions on the AWS services, resources, and individual API actions that users and roles in each account can access.

[AWS Artifact](#) is a service that provides on-demand access to AWS security and compliance reports and select online agreements. AWS Artifact consists of two main sections: AWS Artifact Agreements and AWS Artifact Reports.

Customer Compliance Center

The [Customer Compliance Center](#) contains resources to help you learn more about AWS compliance.

In the Customer Compliance Center, you can read customer compliance stories to discover how companies in regulated industries have solved various compliance, governance, and audit challenges.

You can also access compliance whitepapers and documentation on topics such as:

- AWS answers to key compliance questions
- An overview of AWS risk and compliance
- An auditing security checklist

Denial-of-service attacks

- A **denial-of-service (DoS) attack** is a deliberate attempt to make a website or application unavailable to users.
- **Distributed denial-of-service attacks**
- Now, suppose that the prankster has enlisted the help of friends.
- The prankster and their friends repeatedly call the coffee shop with requests to place orders, even though they do not intend to pick them up. These requests are coming in from different phone numbers, and it's impossible for the coffee shop to block them all. Additionally, the influx of calls has made it increasingly difficult for customers to be able to get their calls through. This is similar to a **distributed denial-of-service attack**.
- **AWS Shield**
- AWS Shield is a service that protects applications against DDoS attacks. AWS Shield provides two levels of protection:

1. **Standard and**

2. **Advanced.**

AWS Key Management Service (AWS KMS)

The coffee shop has many items, such as coffee machines, pastries, money in the cash registers, and so on. You can think of these items as data. The coffee shop owners want to ensure that all of these items are secure, whether they're sitting in the storage room or being transported between shop locations.

In the same way, you must ensure that your applications' data is secure while in storage (**encryption at rest**) and while it is transmitted, known as **encryption in transit**.

[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. You can use AWS KMS to create, manage, and use cryptographic keys. You can also control the use of keys across a wide range of services and in your applications.

AWS WAF

[AWS WAF](#) is a web application firewall that lets you monitor network requests that come into your web applications.

Amazon Inspector

Suppose that the developers at the coffee shop are developing and testing a new ordering application. They want to make sure that they are designing the application in accordance with security best practices. However, they have several other applications to develop, so they cannot spend much time conducting manual assessments. To perform automated security assessments, they decide to use [Amazon Inspector](#). Amazon Inspector helps to improve the security and compliance of applications by running automated security assessments.

Amazon GuardDuty

[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.

7.MONITORING.....

This idea of observing systems, collecting metrics, evaluating those metrics over time, and then using them to make decisions or take actions, is what we call monitoring.

Amazon CloudWatch

[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.

CloudWatch 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.

CloudWatch alarms

With CloudWatch, you can create [alarms](#) that automatically perform actions if the value of your metric has gone above or below a predefined threshold.

For example, suppose that your company’s developers use Amazon EC2 instances for application development or testing purposes. If the developers occasionally forget to stop the instances, the instances will continue to run and incur charges.

In this scenario, 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

[AWS CloudTrail](#) 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, and more. You can think of CloudTrail as a “trail” of breadcrumbs (or a log of actions) that someone has left behind them.

Recall that you can use API calls to provision, manage, and configure your AWS resources. With CloudTrail, you can view a complete history of user activity and API calls for your applications and resources.

AWS Trusted Advisor

[AWS Trusted Advisor](#) is a web service that inspects your AWS environment and provides real-time recommendations in accordance with AWS best practices.

Trusted Advisor compares its findings to AWS best practices in five categories:

- 1.cost optimization,
- 2.performance,
- 3.security,
- 4.fault tolerance, and
- 5.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.

8.PRICING AND SUPPORT.....

AWS Free Tier

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

Consolidated billing(all bills containing in a single bill)

- In an earlier module, you learned about AWS Organizations, a service that enables you to manage multiple AWS accounts from a central location. AWS Organizations also provides the option for [consolidated billing](#).
- The consolidated billing feature of AWS Organizations 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

- In [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

- [AWS Cost Explorer](#) is a tool that enables you to visualize, understand, and manage your AWS costs and usage over time.
- 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

AWS offers four different [Support plans](#) to help you troubleshoot issues, lower costs, and efficiently use AWS services.

You can choose from the following Support plans to meet your company's needs:

- Basic
- Developer
- Business
- Enterprise

Basic Support

Basic Support is free for all AWS customers. It includes access to whitepapers, documentation, and support communities. With Basic Support, you can also contact AWS for billing questions and service limit increases.

With Basic Support, you have access to a limited selection of AWS Trusted Advisor checks. Additionally, you can use the **AWS Personal Health Dashboard**, a tool that provides alerts and remediation guidance when AWS is experiencing events that may affect you.

Developer, Business, and Enterprise Support

The Developer, Business, and Enterprise Support plans include all the benefits of Basic Support, in addition to the ability to open an unrestricted number of technical support cases. These three Support plans have pay-by-the-month pricing and require no long-term contracts.

Technical Account Manager (TAM)

The Enterprise Support plan includes access to a **Technical Account Manager (TAM)**.

If your company has an Enterprise Support plan, the TAM is your primary point of contact at AWS. They provide guidance, architectural reviews, and ongoing communication with your company as you plan, deploy, and optimize your applications.

Your TAM provides expertise across the full range of AWS services. They help you design solutions that efficiently use multiple services together through an integrated approach.

AWS Marketplace

[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.

9.MIGRATION AND INNOVATION.....

- Understand migration and innovation in the AWS Cloud.
- Summarize the AWS Cloud Adoption Framework (AWS CAF).
- Summarize the six key factors of a cloud migration strategy.
- Describe the benefits of AWS data migration solutions, such as AWS Snowcone, AWS Snowball, and AWS Snowmobile.
- Summarize the broad scope of innovative solutions that AWS offers.

Six core perspectives of the Cloud Adoption Framework

- At the highest level, the [AWS Cloud Adoption Framework \(AWS CAF\)](#) organizes guidance into six areas of focus, called **Perspectives**. Each Perspective addresses distinct responsibilities. The planning process helps the right people across the organization prepare for the changes ahead.
- In general, the **Business, People, and Governance** Perspectives focus on business capabilities,
- whereas the **Platform, Security, and Operations** Perspectives focus on technical capabilities.

1. Business Perspective:

The **Business Perspective** ensures that IT aligns with business needs and that IT investments link to key business results.

Use the Business Perspective to create a strong business case for cloud adoption and prioritize cloud adoption initiatives. Ensure that your business strategies and goals align with your IT strategies and goals.

Common roles in the Business Perspective include:

- Business managers
- Finance managers
- Budget owners
- Strategy stakeholders

2. People Perspective:

The **People Perspective** supports development of an organization-wide change management strategy for successful cloud adoption.

Use the People Perspective to evaluate organizational structures and roles, new skill and process requirements, and identify gaps. This helps prioritize training, staffing, and organizational changes.

Common roles in the People Perspective include:

- Human resources
- Staffing
- People managers

3. Governance Perspective:

The **Governance Perspective** focuses on the skills and processes to align IT strategy with business strategy. This ensures that you maximize the business value and minimize risks.

Use the Governance Perspective to understand how to update the staff skills and processes necessary to ensure business governance in the cloud. Manage and measure 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

4.Platform Perspective :

The **Platform Perspective** includes principles and patterns for implementing new solutions on the cloud, and migrating on-premises workloads to the cloud.

Use a variety of architectural models to understand and communicate the structure of IT systems and their relationships. Describe the architecture of the target state environment in detail.

Common roles in the Platform Perspective include:

- Chief Technology Officer (CTO)
- IT managers
- Solutions architects

5.Security Perspective:

The **Security Perspective** ensures that the organization meets security objectives for visibility, auditability, control, and agility.

Use the AWS CAF 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

6.Operations Perspective:

The **Operations Perspective** helps you to enable, run, use, operate, and recover IT workloads to the level agreed upon with your business stakeholders.

Define how day-to-day, quarter-to-quarter, and year-to-year business is conducted. Align with and support the operations of the business. The AWS CAF helps these stakeholders define current operating procedures and identify the process changes and training needed to implement successful cloud adoption.

Common roles in the Operations Perspective include:

- IT operations managers
- IT support managers

6 strategies for migration

When migrating applications to the cloud, six of the most common [migration strategies](#) that you can implement are:

- Rehosting
- Replatforming

- Refactoring/re-architecting
- Repurchasing
- Retaining
- Retiring

Rehosting:

- **Rehosting** also known as “lift-and-shift” involves moving applications without changes.
- In the scenario of a large legacy migration, in which the company is looking to implement its migration and scale quickly to meet a business case, the majority of applications are rehosted.

Replatforming:

- **Replatforming**, also known as “lift, tinker, and shift,” involves making a few cloud optimizations to realize a tangible benefit. Optimization is achieved without changing the core architecture of the application.

Refactoring/re-architecting:

- **Refactoring** (also known as **re-architecting**) involves reimagining how an application is architected and developed by using cloud-native features. Refactoring is driven by a strong business need to add features, scale, or performance that would otherwise be difficult to achieve in the application’s existing environment.

Repurchasing:

- **Repurchasing** involves moving from a traditional license to a software-as-a-service model.
- For example, a business might choose to implement the repurchasing strategy by migrating from a customer relationship management (CRM) system to Salesforce.com.

Retaining:

- **Retaining** consists of keeping applications that are critical for the business in the source environment. This might include applications that require major refactoring before they can be migrated, or, work that can be postponed until a later time.

Retiring:

- **Retiring** is the process of removing applications that are no longer needed.

AWS Snow Family members

- The [AWS Snow Family](#) is a collection of physical devices that help to physically transport up to exabytes of data into and out of AWS.
- AWS Snow Family is composed of **AWS Snowcone**, **AWS Snowball**, and **AWS Snowmobile**.



AWS Snowcone

is a small, rugged, and secure edge computing and data transfer device.

It features 2 CPUs, 4 GB of memory, and 8 TB of usable storage.

AWS Snowball offers two types of devices:

- **Snowball Edge Storage Optimized** devices are well suited for large-scale data migrations and recurring transfer workflows, in addition to local computing with higher capacity needs.
 - Storage: 80 TB of hard disk drive (HDD) capacity for block volumes and Amazon S3 compatible object storage, and 1 TB of SATA solid state drive (SSD) for block volumes.
 - Compute: 40 vCPUs, and 80 GiB of memory to support Amazon EC2 sbe1 instances (equivalent to C5).
- **Snowball Edge Compute Optimized** provides powerful computing resources for use cases such as machine learning, full motion video analysis, analytics, and local computing stacks.
 - Storage: 42-TB usable HDD capacity for Amazon S3 compatible object storage or Amazon EBS compatible block volumes and 7.68 TB of usable NVMe SSD capacity for Amazon EBS compatible block volumes.
 - Compute: 52 vCPUs, 208 GiB of memory, and an optional NVIDIA Tesla V100 GPU. Devices run Amazon EC2 sbe-c and sbe-g instances, which are equivalent to C5, M5a, G3, and P3 instances.

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.

Innovate with AWS Services

When examining how to use AWS services, it is important to focus on the desired outcomes. You are properly equipped to drive innovation in the cloud if you can clearly articulate the following conditions:

- The current state
- The desired state
- The problems you are trying to solve

Consider some of the paths you might explore in the future as you continue on your cloud journey.

10.THE CLOUD JOURNEY.....

The AWS Well-Architected Framework

The [AWS Well-Architected Framework](#) helps you understand how to design and operate reliable, secure, efficient, and cost-effective systems in the AWS Cloud. It provides a way for you to consistently measure your architecture against best practices and design principles and identify areas for improvement.

The Well-Architected Framework is based on five pillars:

- Operational excellence
- Security
- Reliability
- Performance efficiency
- Cost optimization

Advantages of cloud computing

Operating in the AWS Cloud offers many benefits over computing in on-premises or hybrid environments.

In this section, you will learn about six advantages of cloud computing:

- Trade upfront expense for variable expense.
- Benefit from massive economies of scale.
- Stop guessing capacity.
- Increase speed and agility.
- Stop spending money running and maintaining data centers.
- Go global in minutes.