

Module 1

Introduction to the Cloud

Outline

- **MODULE 1 INTRODUCTION**
 - Welcome to AWS Cloud Practitioner Essentials
- **THE AWS CLOUD**
 - What is Cloud Computing?
 - Benefits of the AWS Cloud
 - Introduction to AWS Global Infrastructure
 - The AWS Shared Responsibility Model
- **CLOUD IN REAL LIFE**
 - Applying Cloud Concepts to Real Life Use Cases

What is Cloud Computing?

In this lesson, you will learn:

- Define cloud computing.
- Describe and differentiate between cloud deployment types.

Defining cloud computing

- It is on-demand delivery of IT resources over the internet with pay-as-you-go pricing:
- **On-demand delivery**
 - *On-demand delivery* means that customers can access computing resources, such as storage or compute power, within seconds and as needed.
 - Users can scale their resource usage up or down based on current requirements without lengthy provisioning processes.
- **of IT resources**
 - The *of IT resources* aspect highlights the wide array of information technology assets in the cloud-computing space.
 - These resources include servers, storage solutions, databases, networking components, artificial intelligence and machine learning (AI/ML) tools, and more.
 - Customers can use these resources to build, deploy, and manage applications and services through the cloud infrastructure.

Defining cloud computing

- **over the internet**
 - *Over the internet* signifies that cloud computing delivers IT resources through internet connectivity.
 - This means that users access and use these resources through web-based services rather than maintaining local hardware or software.
 - The internet acts as the conduit, which provides remote access to compute power, storage, and applications from anywhere in the world.
- **with pay-as-you-go-pricing**
 - Flexible pricing is a fundamental economic aspect of cloud computing.
 - Users pay only for the resources they actually consume, rather than committing to fixed, long-term contracts.
 - This usage-based pricing model offers cost efficiency and financial flexibility.

Cloud deployment types

- **Cloud**
 - In a cloud-based deployment model, you have the flexibility to migrate your existing resources to the cloud, design and build new applications within the cloud environment, or use a combination of both.
 - For instance, a company might migrate data resources to the cloud, then develop an application comprised of virtual servers, databases, and networking components entirely hosted in the cloud.
- **On-premises**
 - Deploying resources on premises using virtualization and resource management tools does not provide many of the benefits of cloud computing.
 - However, it is sometimes sought for its ability to provide dedicated resources and low latency.
 - In most cases this deployment model is the same as legacy IT infrastructure while using application management and virtualization technologies to try increasing resource utilization.

Cloud deployment types

- **Hybrid**
 - In a hybrid deployment, cloud-based resources and on-premises infrastructure work together. This approach is ideal for situations where legacy applications must remain on premises due to maintenance preferences or regulatory requirements.
 - For instance, a company might choose to retain certain regulated legacy applications on-premises while using cloud services for advanced data processing and analytics.
 - Multi-cloud deployments can also be considered *hybrid deployments*.

Benefits of the AWS Cloud

In this lesson, you will learn

- Describe the six key benefits of cloud computing.

The six key benefits of the AWS Cloud

1. Trade fixed expense for variable expense

By using the AWS Cloud, businesses can transition from fixed investments to variable costs.

With variable costs, customer expenses are better aligned with actual usage, thus creating more financial flexibility.

2. Benefit from massive economies of scale

Like buying a product in bulk can result in lower prices per unit, the vast global infrastructure of AWS can result in lower costs for customers.

This means that AWS can be used by many organizations, from small startups to major corporations.

Businesses big and small can access advanced technologies that were previously only accessible to large enterprises.

The six key benefits of the AWS Cloud

3. Stop guessing capacity

Customers can dynamically scale AWS Cloud resources up or down based on real-time demand.

This means businesses can achieve optimal performance without provisioning more or less infrastructure than they need.

4. Increase speed and agility

With the cloud, businesses can rapidly deploy applications and services, accelerating time to market and facilitating quicker responses to changing business needs and market conditions.

The six key benefits of the AWS Cloud

5. Stop spending money to run and maintain data centers

The AWS Cloud eliminates the need for businesses to invest in physical data centers.

This means customers aren't required to spend time and money on utilities and ongoing maintenance.

With AWS taking care of the physical infrastructure of the cloud, customer resources can be reallocated to more strategic initiatives.

6. Go global in minutes

Businesses don't need to set up their own infrastructure to expand internationally.

AWS provides a robust global infrastructure that customers can use to deploy applications and services across multiple areas in minutes.

Introduction to AWS Global Infrastructure

In this lesson, you will learn:

- Define AWS Regions and Availability Zones.
- Explain the benefits of high availability and fault tolerance.

AWS Regions and Availability Zones

- AWS Global Infrastructure consists of physical locations around the world that contain groups of data centers.



AWS Regions and Availability Zones

- **AWS Regions**

- AWS Regions are physical locations around the world that contain groups of data centers.
- These groups of data centers are called Availability Zones.
- Each AWS Region consists of a minimum of three physically separate Availability Zones within a geographic area.

- **Availability Zones**

- An Availability Zone consists of one or more data centers with redundant power, networking, and connectivity.
- Regions and Availability Zones are designed to provide low-latency, fault-tolerant access to services for users within a given area.

Achieving high availability with AWS Global Infrastructure

- AWS infrastructure is designed with high availability and fault tolerance in mind.
- Availability Zones (AZs) are configured as isolated resources, and they are each equipped with independent power, networking, and connectivity.
- It's recommended to distribute your resources across multiple AZs.
- That way, if one AZ encounters an outage, your business applications will continue to operate without interruption.
- With this approach of redundancy and resource isolation, AWS customers can achieve the benefits of high availability and fault tolerance.

The AWS Shared Responsibility Model

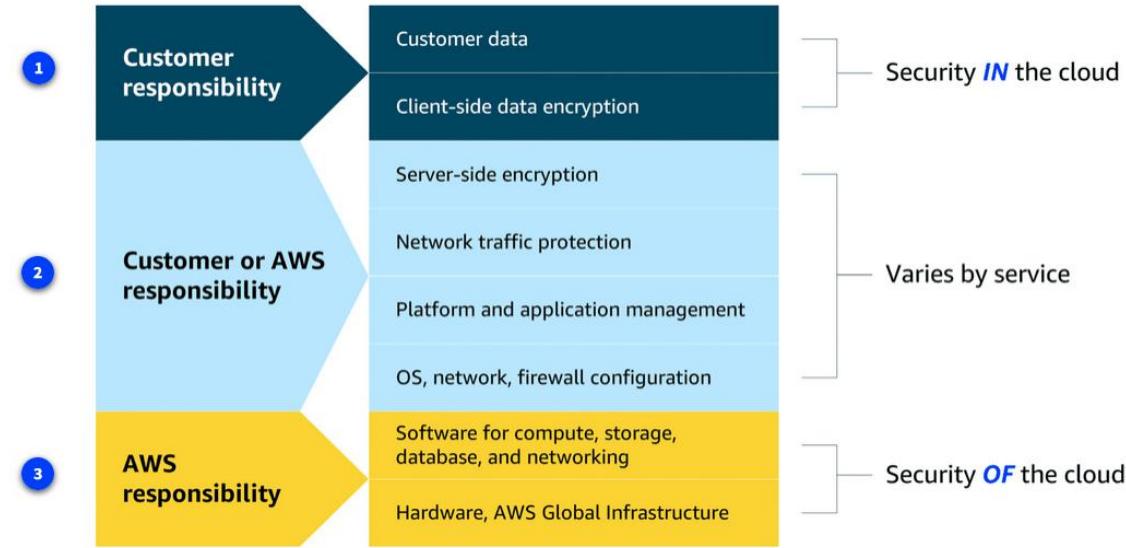
In this lesson, you will learn

- Describe and differentiate between customer responsibilities, AWS responsibilities, and shared responsibilities in the AWS Cloud.
- Describe the components of the AWS Shared Responsibility Model.

Components of the AWS Shared Responsibility Model

- **Customer responsibilities**

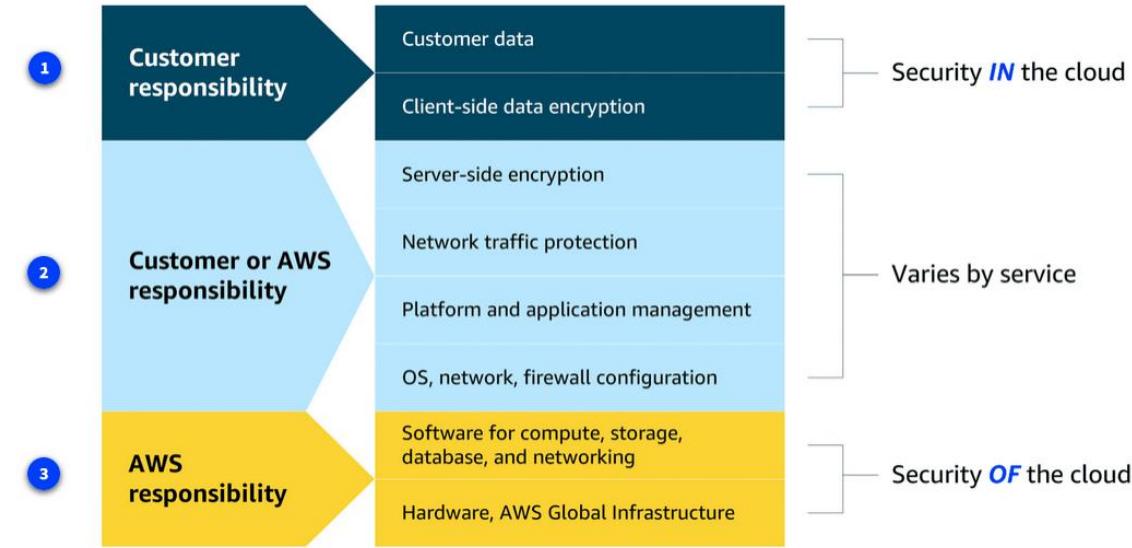
- Customers are responsible for managing security requirements for their data, including which data they store on AWS and who has access to that data.
- Customers also control how access to the data is granted, managed, and revoked.
- Additionally, customers are responsible for client-side encryption.



Components of the AWS Shared Responsibility Model

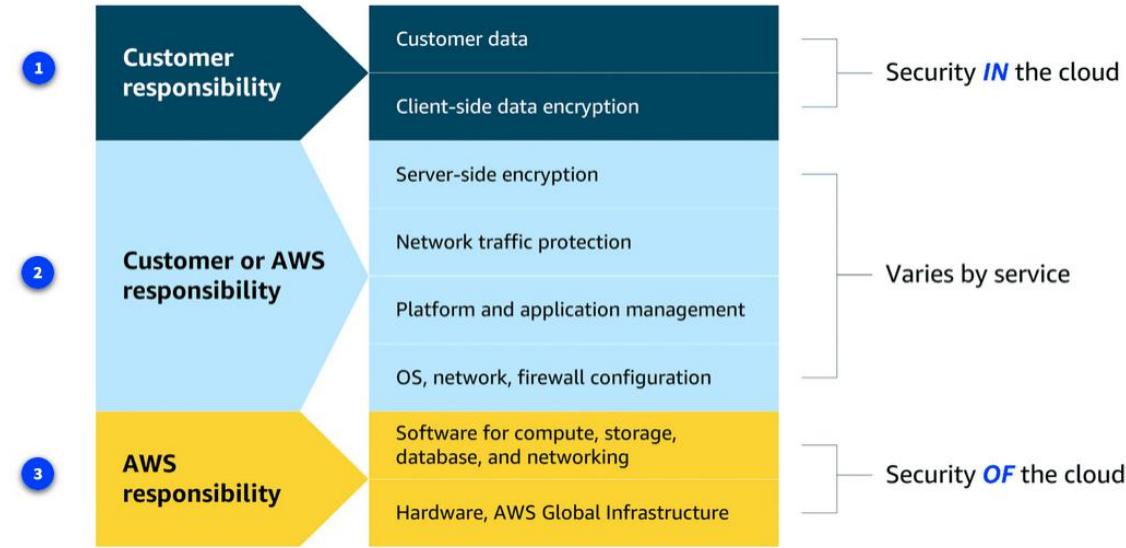
- **Shared responsibilities**

- Depending on the service used, responsibilities might shift between the customer and AWS.
- Components such as server-side encryption, network traffic protection, platform and application management, and OS, network, and firewall configuration vary by service in terms of who is responsible for these items.
- As you learn more about different types of services, you will see specific examples of how these elements are divided between the customer and AWS.



Components of the AWS Shared Responsibility Model

- **AWS responsibilities**
 - AWS is responsible for protecting the infrastructure that runs all of the services offered in the AWS Cloud.
 - This infrastructure is composed of the hardware, software, networking, and facilities that run AWS Cloud services.



Applying Cloud Concepts to Real Life Use Cases

In this lesson, you will learn

- Explain how fundamental cloud concepts, such as the AWS Global Infrastructure and AWS Shared Responsibility Model, work together to form real-world business solutions.

Cloud in real life: Infrastructure and shared responsibility

- **AWS Global Infrastructure – Real-World Impact**
 - **Problem:** An e-commerce company wants to expand globally.
 - Customers in Europe and Asia need fast, reliable access.
 - **Solution with AWS:**
 - Deploy workloads in **Regions close to customers** (e.g., eu-west-1 in Ireland, ap-southeast-1 in Singapore).
 - Reduce **latency** by leveraging AWS's pre-built global footprint.
 - Startups benefit since they don't need to build or buy data centers in those geographies
 - cloud lowers the barrier to global reach.

Cloud in real life: Infrastructure and shared responsibility

- **High Availability & Fault Tolerance**
 - **Problem:** E-commerce apps can't afford downtime—one AZ outage should not kill the website.
 - **Solution with AWS:**
 - Deploy in **multiple Availability Zones (AZs)** within a Region.
 - If one AZ goes down, traffic can fail over to another.
 - This design provides **fault tolerance and high availability** without building redundant physical data centers.

Cloud in real life: Infrastructure and shared responsibility

- **Shared Responsibility Model**
 - **Problem:** Who handles security for global expansion—AWS or the customer?
 - **Solution with AWS:**
 - **AWS responsibility (“of the cloud”)** → securing physical infrastructure (data centers, networking, hardware).
 - **Customer responsibility (“in the cloud”)** → securing applications, configurations, user access, and compliance with standards (like PCI DSS for credit card data).
 - This split allows businesses to **focus on business-critical security** instead of racking servers or guarding facilities.

Recap

- In this section of the training, you learned fundamental concepts of cloud computing.
- You explored the definition and benefits of the cloud, and you were introduced to AWS Global Infrastructure.
- You also explored the AWS Shared Responsibility Model to clarify the division of responsibilities between AWS and customers.

Resources

Resource link

Description

[What is Cloud Computing?](#)

Learn more about the basics of cloud computing.

[AWS Shared Responsibility Model](#)

Learn more about the AWS Shared Responsibility Model.

[Regions and Availability Zones](#)

View a complete list of AWS Regions.