Series Agenda

Part 1

- Introduction: Course Overview
- Module 1: Introduction to Amazon Web Services
- Module 2: Global Infrastructure and Reliability
- Module 3: Networking
- Module 4: Object Storage

Part 2

- Module 5: Security
- Module 6: Block and File Storage
- Module 7: Compute in the Cloud

Part 3

- Module 8: AWS Frameworks, Pricing, and Support
- Module 9: Applications in the Cloud
- Module 10: Databases
- Module 11: Monitoring and Analytics

Introduce yourself

- Name
- Where are you from?
- What do you hope to get out of this class?
- What is your experience level with AWS?

Your introduction helps the trainer understand the experience level of the class and what you each want to get out of your time here.

Setup and books: One-time steps

Resources you will need during this course:

Demonstrations

- Be sure to create your Skill Builder account so you can follow along.

Digital Books

- Are available in your SkillBuilder account following class.

Course slides and notes can be accessed following the class.
Interactive Demos are performed using Skill Builder's AWS Labs.

This course offers many interactive activities

Instructor demonstrations

- You observe as the instructor demonstrates parts of the AWS console and features related to your discussion.

Interactive demo labs

- You access a lab environment where you are able to follow along with instructor-led demonstrations.

Knowledge Checks

- You answer questions to reinforce what you have learned over the course of the module.

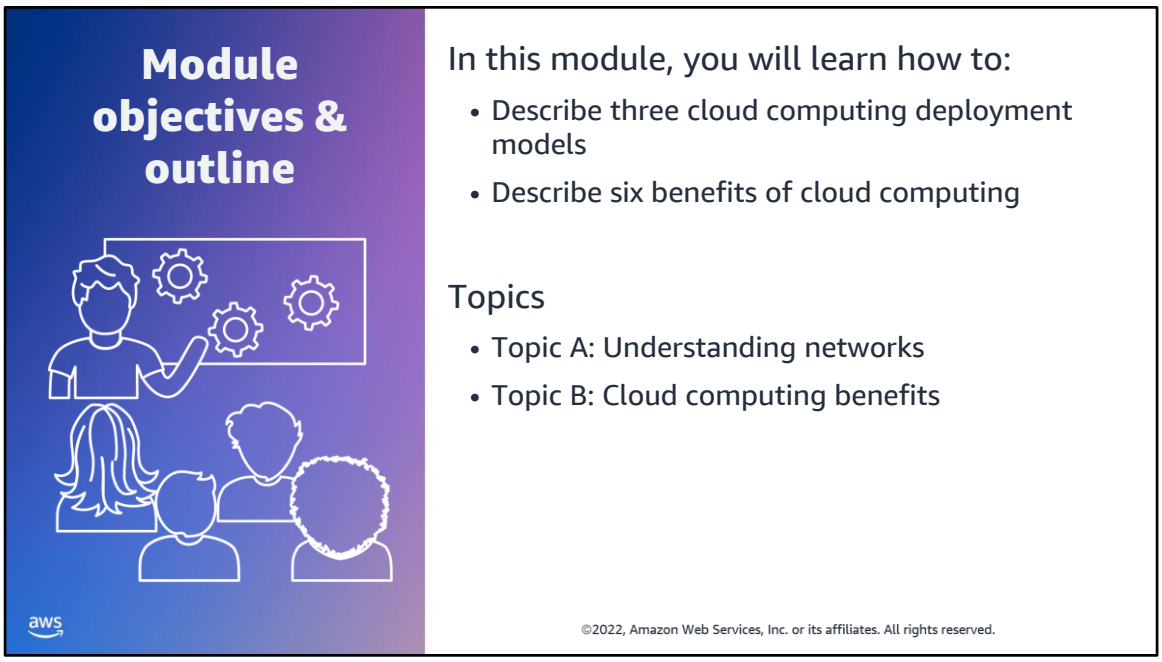
In this module, you will learn how to describe cloud computing deployment models and define the six benefits of cloud computing.

Module Overview:

- Describe three cloud computing deployment models
- Describe six benefits of cloud computing

Topics:

- Topic A: Understanding networks
- Topic B: Cloud computing benefits



In this module, you will learn how to describe cloud computing deployment models and define the six benefits of cloud computing.

Module Overview:

- Describe three cloud computing deployment models
- Describe six benefits of cloud computing

Topics:

- Topic A: Understanding networks
- Topic B: Cloud computing benefits

Acronyms covered in this module

AWS — **Amazon Web Services**
Is a cloud services platform. Simply, it provides cloud services to both individuals and businesses.

PC — **Personal Computer**
A small scale computer made to be used by an individual.
- Laptop or desktops

ISP — **Internet Service Provider**
A company that provides access to the Internet.

CSP — **Cloud Service Provider**
A company that provides cloud-based services such as platform, infrastructure, application, or storage.
- AWS, Google, Azure

IT/IS — **Information Technology / Information Systems**
Often used to refer to the department in a company that is responsible for installing and maintaining computer hardware and software.

aws

Terminology covered in this module (1 of 3)

Device

A piece of hardware used to connect you to the applications, files, or services you need.

- Cell phone, laptop, tablet, PC, or server.

Client

A device that gathers information and directions from another device.

- Most commonly a laptop or PC. Can be a cell phone or tablet.

Server

A very powerful computer that provides access to applications, files, and services.

- Print server, file server, network server, or application server.

aws

Terminology covered in this module (2 of 3)

On-premises / On-prem

IT hardware and software applications hosted where the business operates or in a physical data center.

Data center

A facility dedicated to supporting a very large number of powerful servers used by organizations for remote storage and to prevent failures (fault tolerance).

Deploy

It a term used to describe the process of installing and configuring and new virtual server or application. This can be used in the context of on-prem or cloud environments.

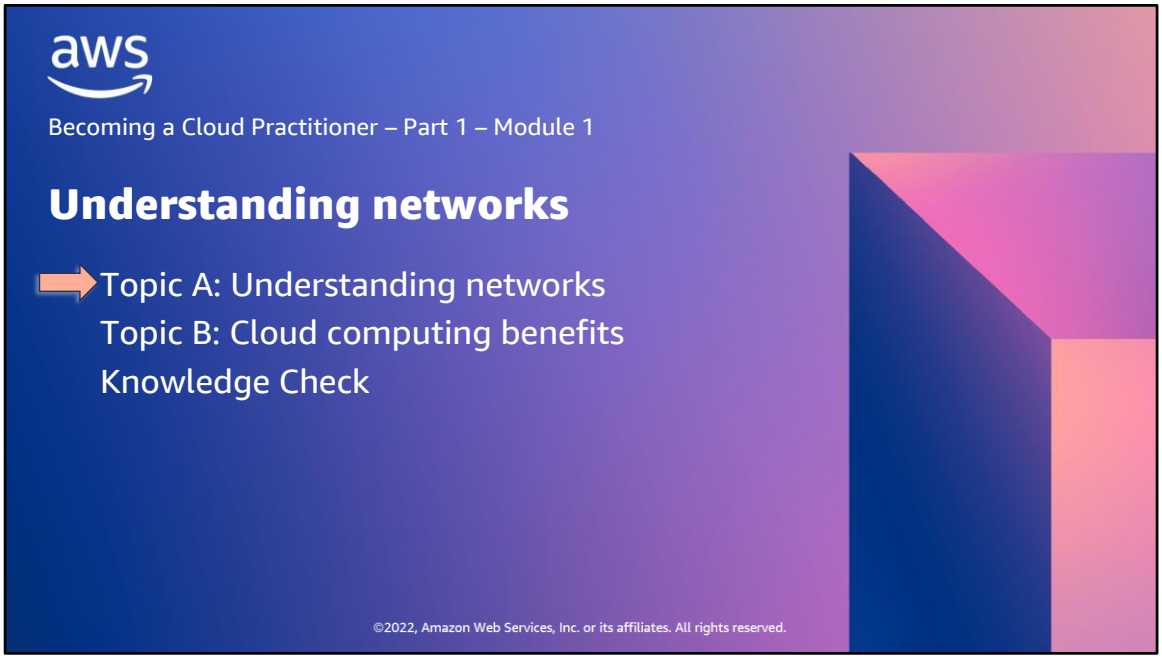
- Synonymous with launch

aws

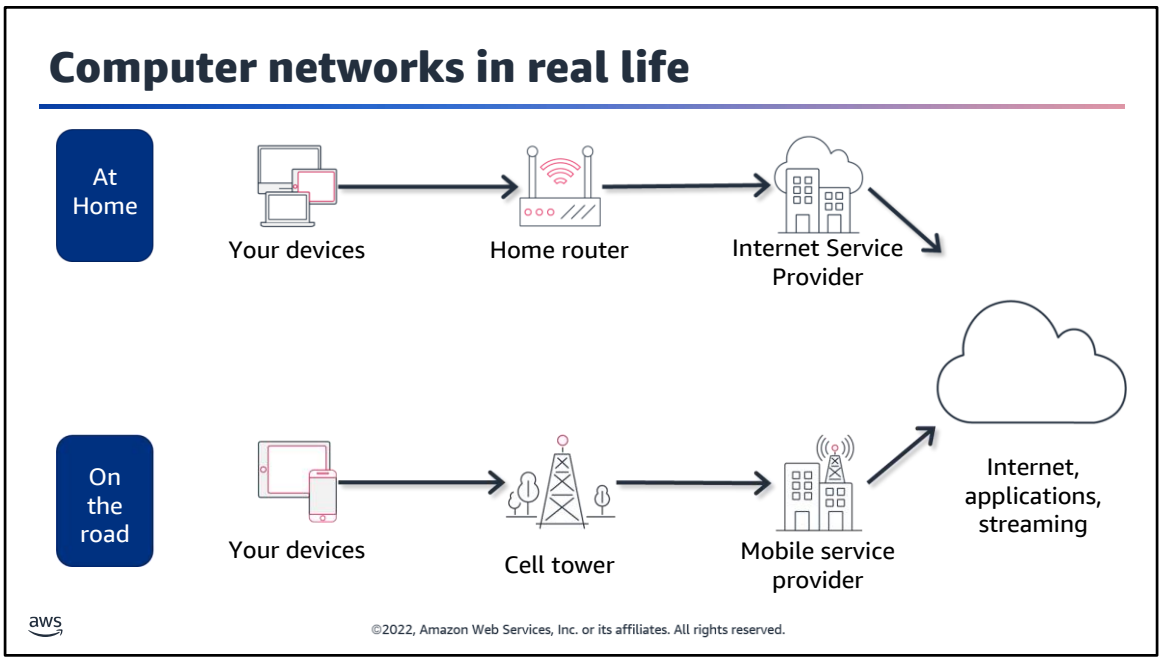
Terminology covered in this module (3 of 3)

Cloud-native

This describes an application or feature that was designed specifically to run in the cloud.



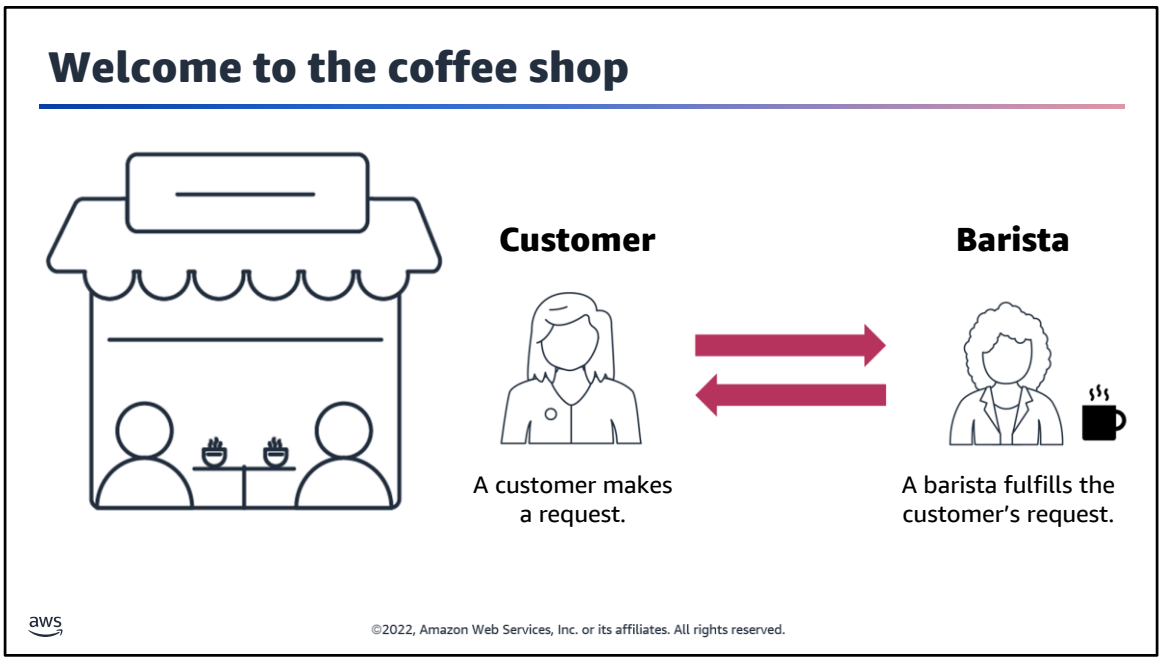
In this topic, you will learn about the Amazon EC2 service.



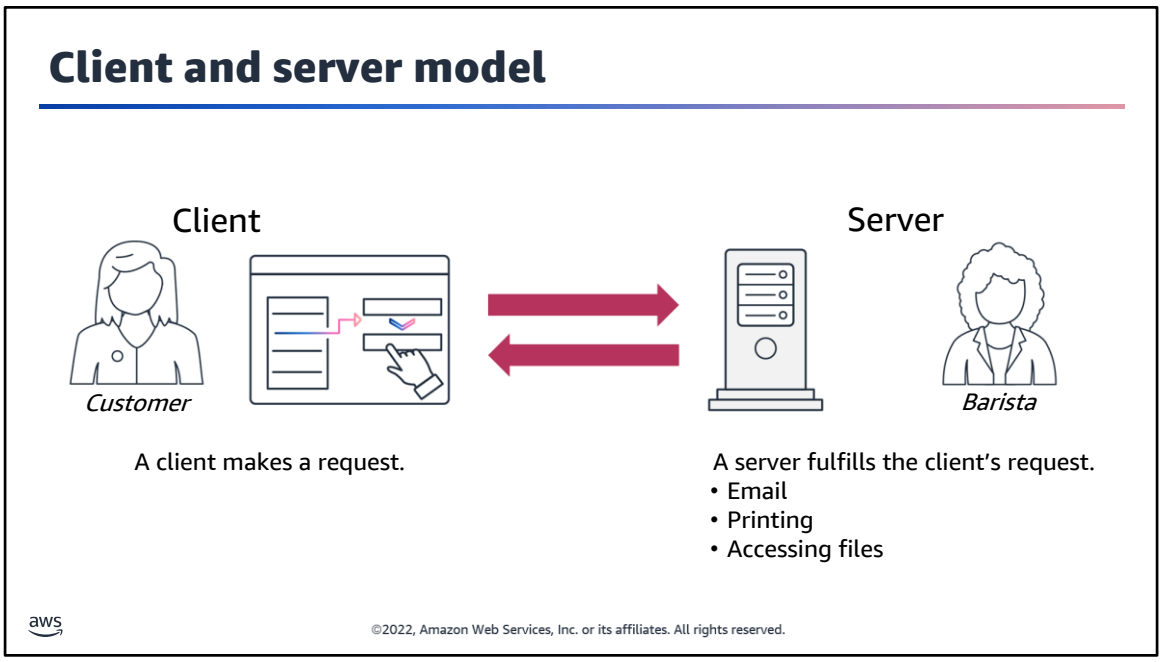
Throughout this course, you will learn essential information about the breadth and depth of Amazon Web Services (**AWS**) offerings, and how AWS products and services can benefit companies.

Understanding computer networks starts by understanding what many people have right at home. You connect your tablet or cell phone to your home router. This is often a box with several blinking lights. It may or may not have an external antenna. Your home router connects you to the internet. It is your Internet Service Provider (**ISP**) that allows you to connect to the internet.

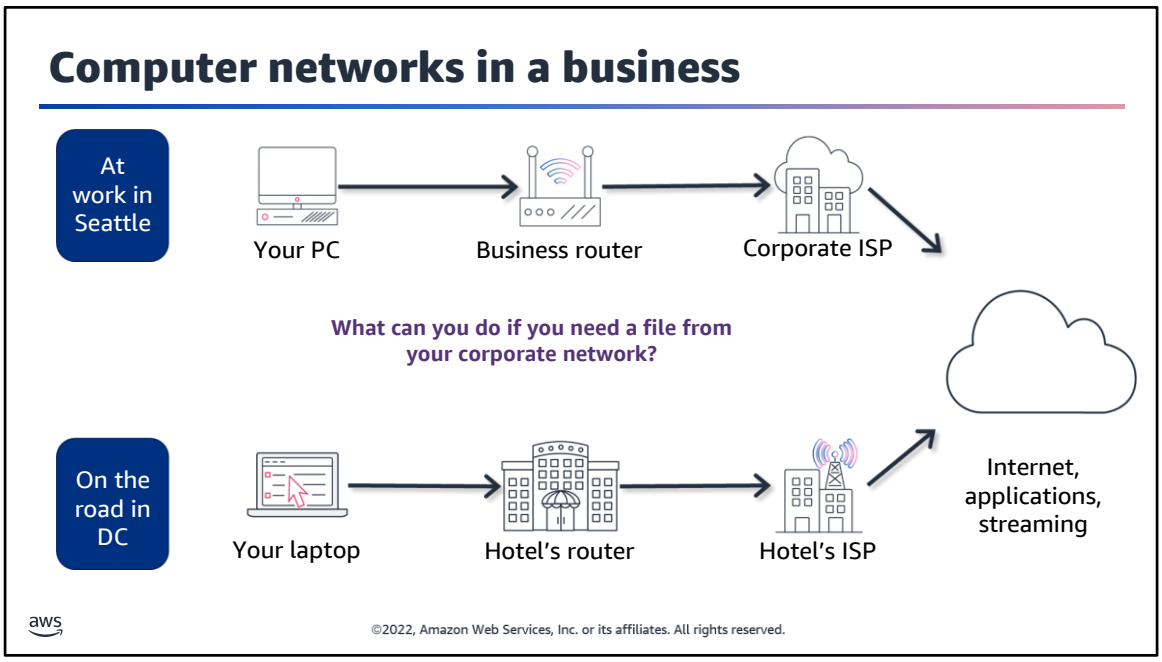
ISPs offer many different services for both residential and commercial customers. They provide your connection to the internet and even cable TV. Today, many cell phone services are also ISPs.



To help you understand the concept of networking, we will use the metaphor of a coffee shop.
Almost all modern computing centers around a basic client-server model, which uses transactions similar to coffee shop transactions. In a coffee shop, a customer makes a request for a cup of coffee. This request is fulfilled when the barista prepares the order and provides it to the customer.



In computing, the "customer" is the client. A client can be a web browser or desktop application that a person interacts with to make a request to a computer server. The server is like the "barista" in the coffee shop. For example, suppose that a client requests a coffee order, coupons, promotions, and so on. The server evaluates the request and fulfills it by returning the information to the client.

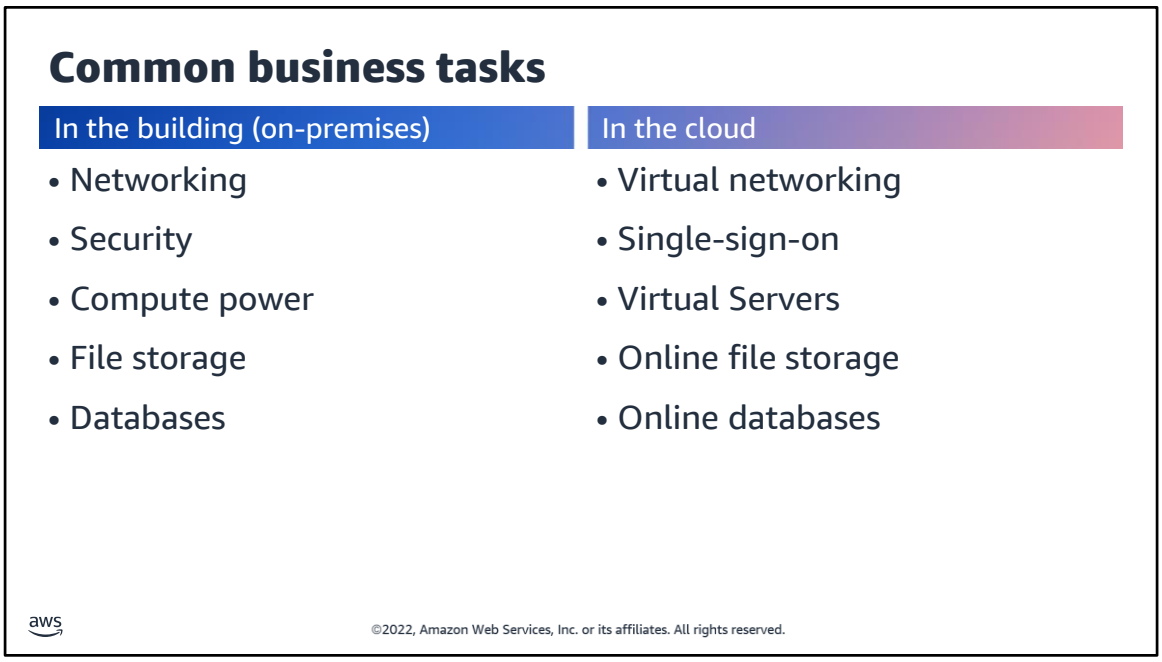


You have learned what is involved with connecting to the internet at home or on your cell phone. How does it work when you are at work?

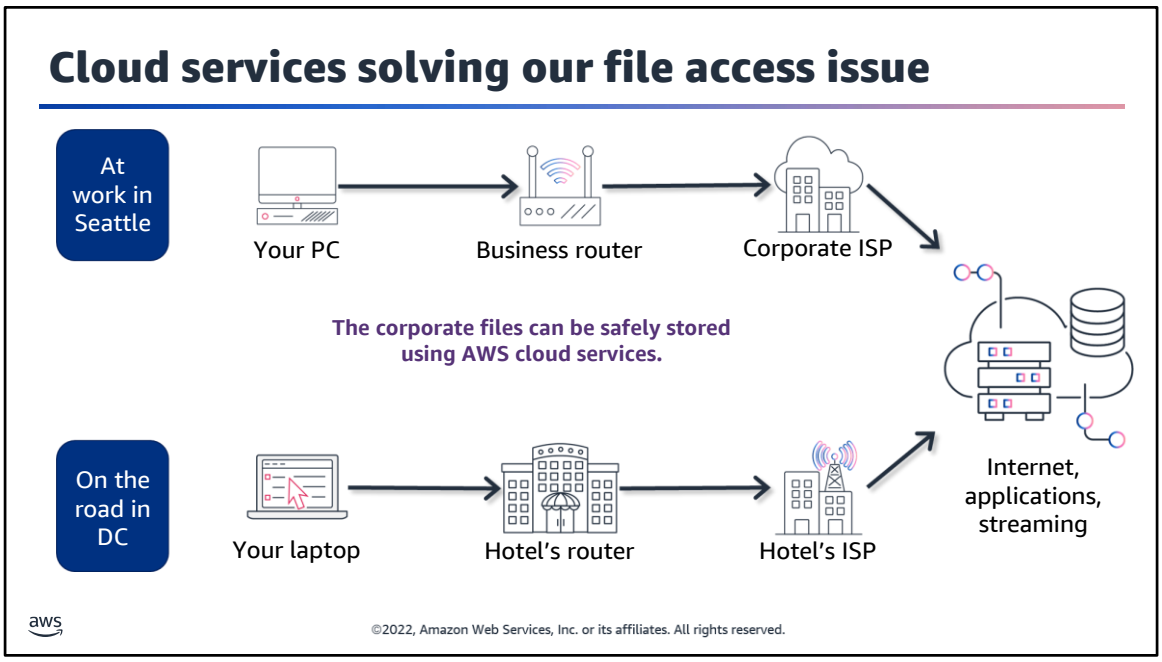
Your PC is connected to a computer network. The network is a series of wires or wireless signals that allow you to communicate with a business router. Just like at home, this router connects the corporate network to the internet.

When you are on the road for work, your laptop will connect to the hotel's router. Just like at home and at work, the router connects you to the hotel's ISP which then connects you to the internet.

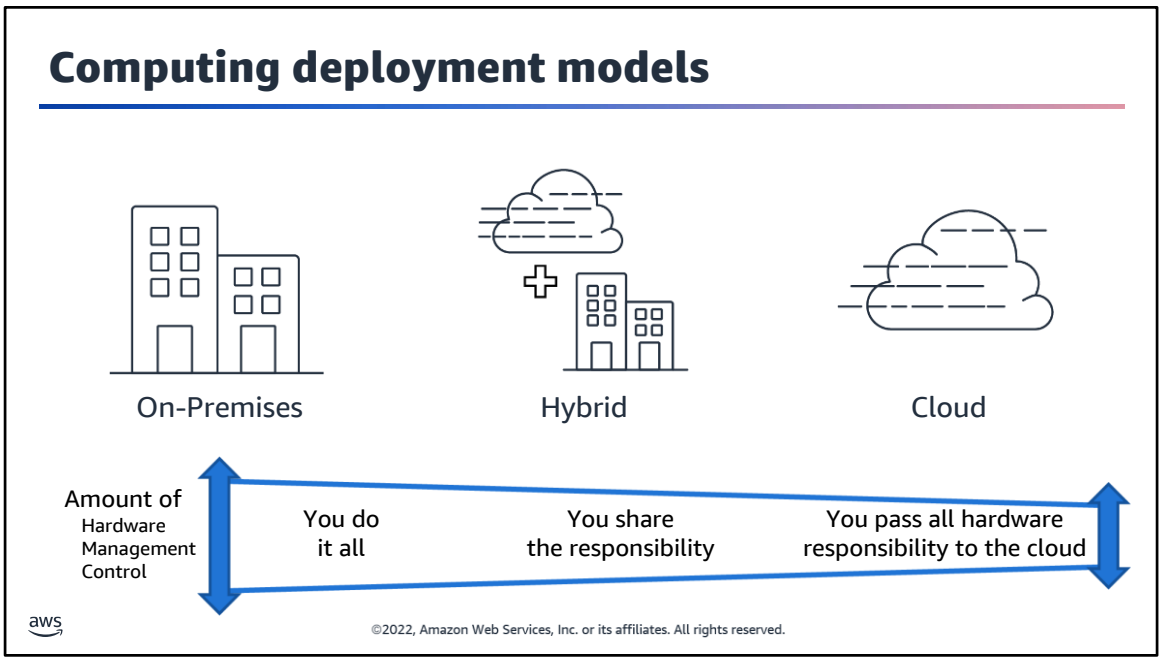
What can you do if you need a file from your corporate network when you are on the road?



Networking: Businesses can choose to use the AWS Cloud networking services to replace some or all of the networking components in their physical locations.
Security: Businesses often need to have security measures in place that may be far more robust than what they may require if they did not have an internet presence. Adding AWS Enhanced security can enable advanced functionality and security features in the cloud that would be far to expensive to implement on-premises.
Compute power: Let face it, buying a super computer is far out of reach for many small to mid-sized business. But, with AWS you can purchase small segments of time on a super computer instead! The cost of hardware is no longer a consideration when using AWS compute services.
File Storage: Online photo storage and music storage has become almost mainstream in many countries around the world. These services are offered by Cloud Service Providers like AWS.
Databases: All of the most powerful websites on the internet are run with the help of equally powerful databases. Often times these databases run in the cloud. Businesses can take advantage of cloude database and realize the same benefits of cloud compute.



One solution for this problem is to move those corporate files to a secure online cloud service. AWS offers these services along with many others that we will cover in this course.
Now that the files are online, you do not have to figure out how to get back onto your corporate network when travelling. You simply need an internet connection and you can connect to your files in the cloud.



The way that a company chooses to organize their IT resources is known as a Computing Deployment Model. The model determines how users and applications access files, data, and other applications. In a world that is becoming every more connected, having access to these resources anywhere in the world is a more common requirement.

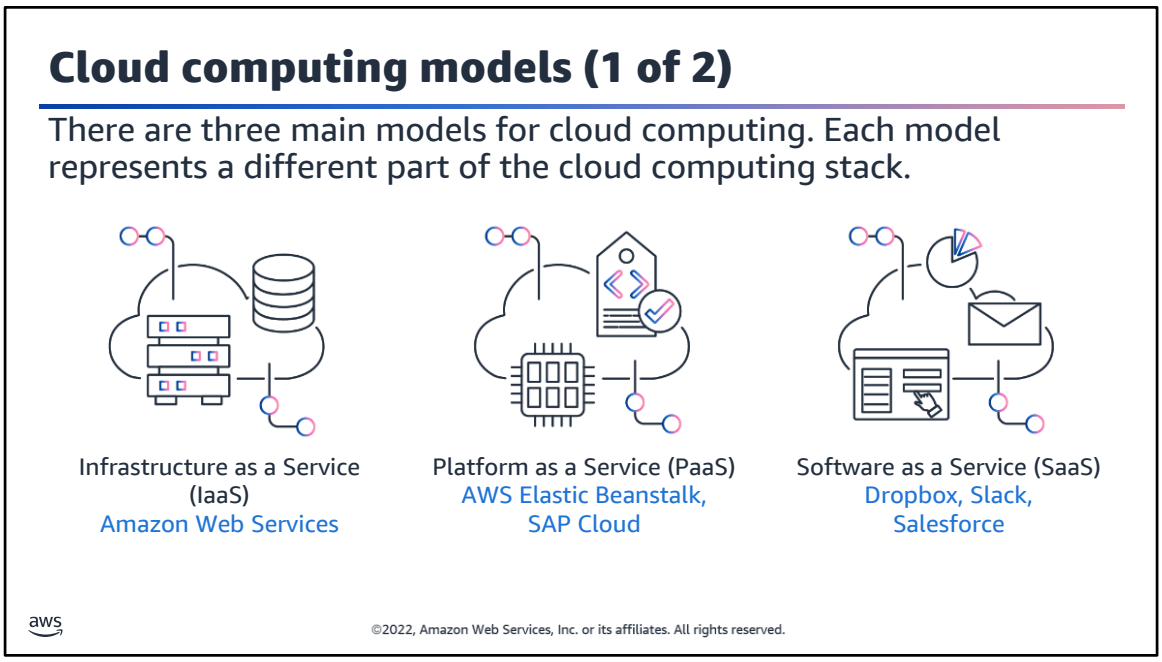
There are three commonly accepted computing deployment models. They are cloud-based, on-premises, and hybrid deployments. When selecting a deployment strategy, a company must consider factors such as required application components, preferred resource management tools, and legacy IT infrastructure requirements.

On-premises deployment is the most common. In this model, resources are deployed in facilities owned by the company using virtualization and resource-management tools. For example, you might have applications that run in your on-premises data center. As a company grows, they may use a central facility that serves only as a data center for the company. The services offered in the data center would be accessed over virtual private connections. When this method is implemented, it is known as a private cloud.

In a **cloud-based deployment** model, you can migrate existing applications to the cloud, or you can design and build new applications in the cloud. You can build the applications on low-level infrastructure that requires your IT staff to manage them. Or, you can build them using higher-level services that reduce the management, architecting, and scaling requirements of the core infrastructure. For example, a company might create an application consisting of virtual servers, databases, and networking components that are fully based in the cloud. This would reduce the company's management, architecting, and scaling requirements.

In a **hybrid deployment model**, cloud-based resources are connected to an on-premises infrastructure. You might want to use this approach in a number of situations. For example, you might have legacy applications that are better maintained on premises, or government regulations require your business to keep certain records on premises.

Suppose that a company wants to use cloud services that can automate batch data processing and analytics. However, the company has several legacy applications that are more suitable on premises and will not be migrated to the cloud. With a hybrid deployment, the company could keep the legacy applications on premises, and use the data and analytics services that run in the cloud.

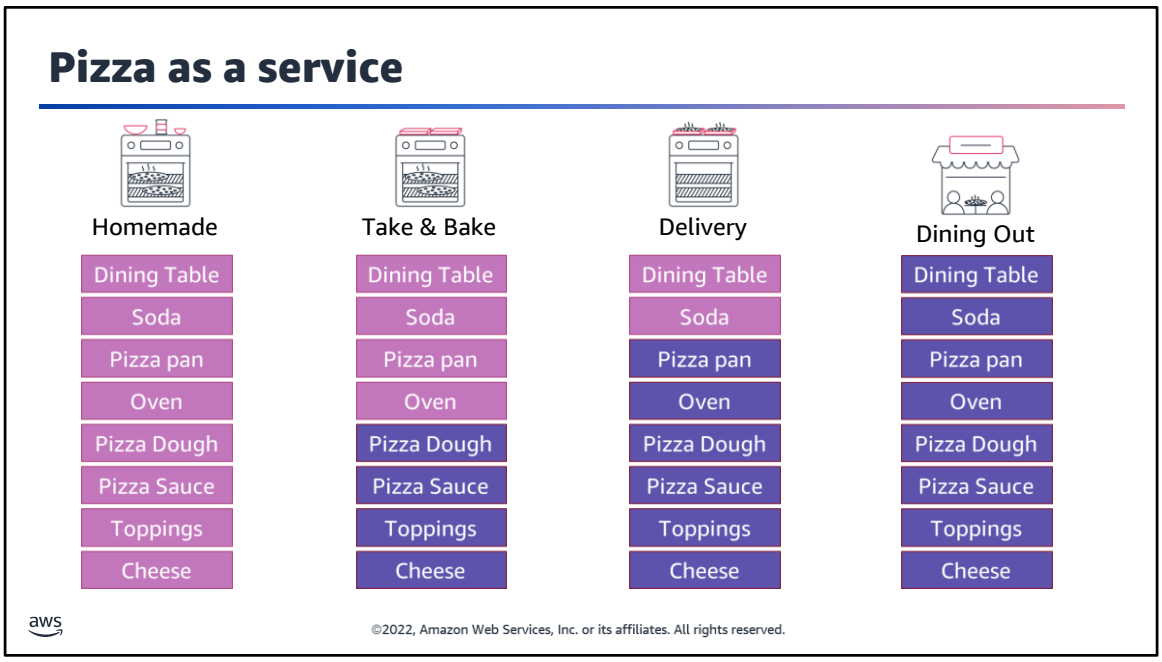


When selecting cloud services, there are three common phrases used; Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS).

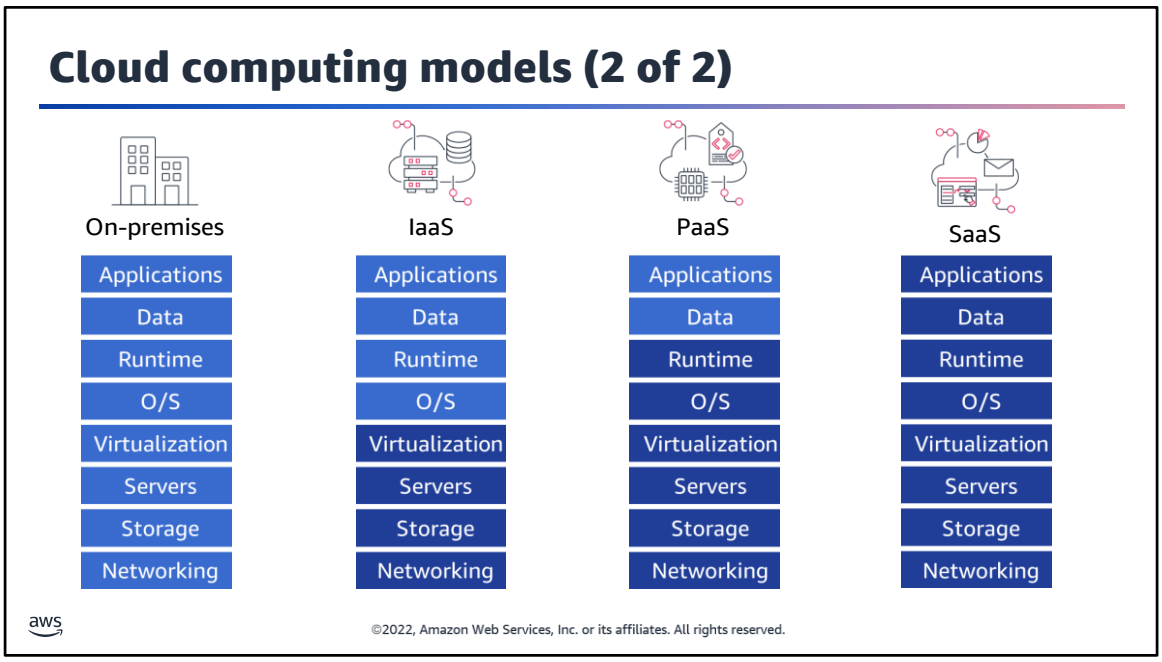
Infrastructure as a Service (IaaS) contains the basic building blocks for cloud IT and typically provide access to networking features, computers (virtual or on dedicated hardware), and data storage space. Infrastructure as a Service provides you with the highest level of flexibility and management control over your IT resources and is most similar to existing IT resources that many IT departments and developers are familiar with today. IaaS includes networking, computing, and data storage components that you manage.

Platform as a Service (PaaS) removes the need for organizations to manage the underlying infrastructure (usually hardware and operating systems) and allow you to focus on the deployment and management of your applications. This helps you be more efficient as you don't need to worry about resource procurement, capacity planning, software maintenance, patching, or any of the other undifferentiated heavy lifting involved in running your application. PaaS solutions are fully managed. You are able to focus on the applications and data without concern over complex networking and compute.

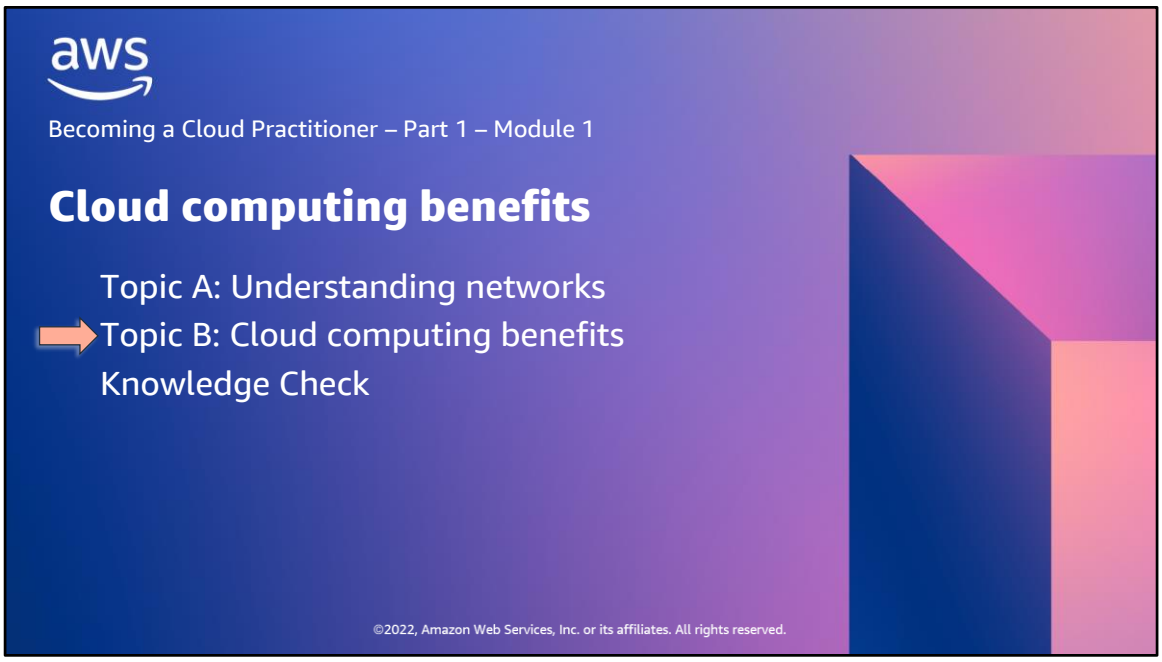
Software as a Service (SaaS) provides you with a completed product that is run and managed by the service provider. In most cases, people referring to Software as a Service are referring to end-user applications. With a SaaS offering you do not have to think about how the service is maintained or how the underlying infrastructure is managed; you only need to think about how you will use that particular piece of software. A common example of a SaaS application is web-based email where you can send and receive email without having to manage feature additions to the email product or maintaining the servers and operating systems that the email program is running on.



We all love to eat pizza but there are different ways that we can get our pizza.
Homemade pizza is great when you want to pick the type of crust and every topping that goes on your pizza. You even get to pick how it is cooked and for how long.
Take and bake pizza is best when you don't want to worry about making the dough and having all of the ingredients on hand. But, you do want to choose how it is baked and for how long.
Delivery pizza is best when you just want to eat great pizza at home.
Dining out is best when you want everything done for you.
So. How does this relate to cloud computing models?

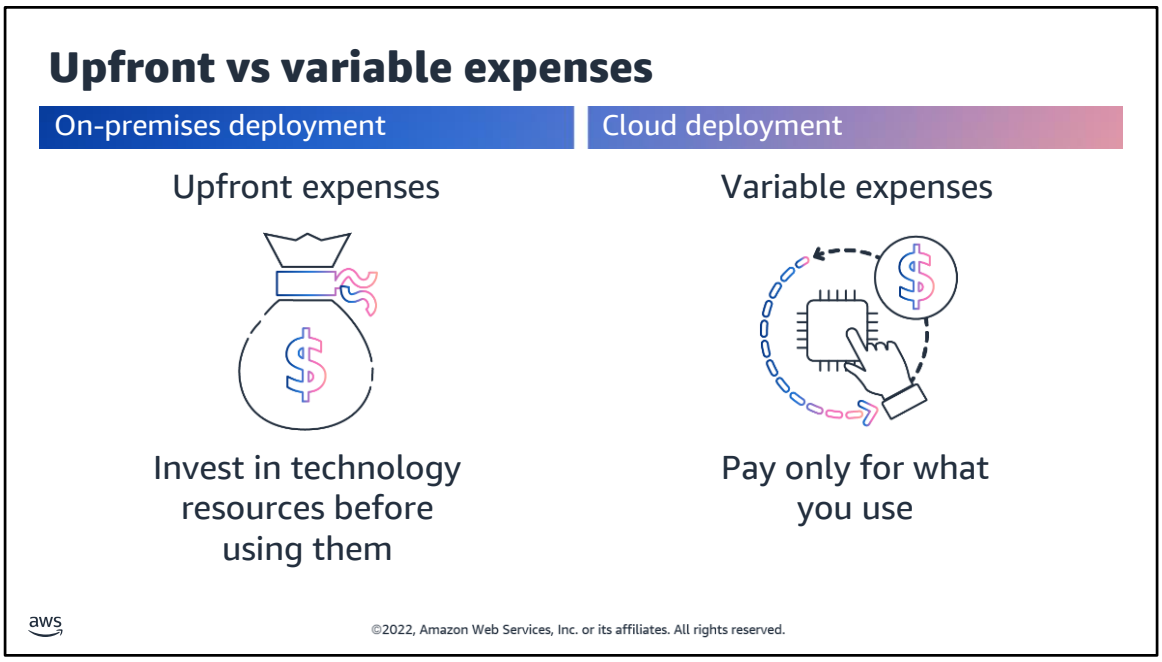


Computing models help us to define what responsibilities the company administrators are responsible and what a service provider will be responsible for.
In an **on-premises deployment** the administrator is responsible for everything.
In an **IaaS deployment** the administrator is responsible for the application, data, runtime, and operating system. The service provider is responsible for the virtualization, servers, storage, and networking.
In a **PaaS deployment** the administrator is only responsible for the application and data. The service provider is responsible for everything else.
In a **SaaS deployment** the administrator is no responsible for anything other that the internal application management. The service provider is responsible for everything else.

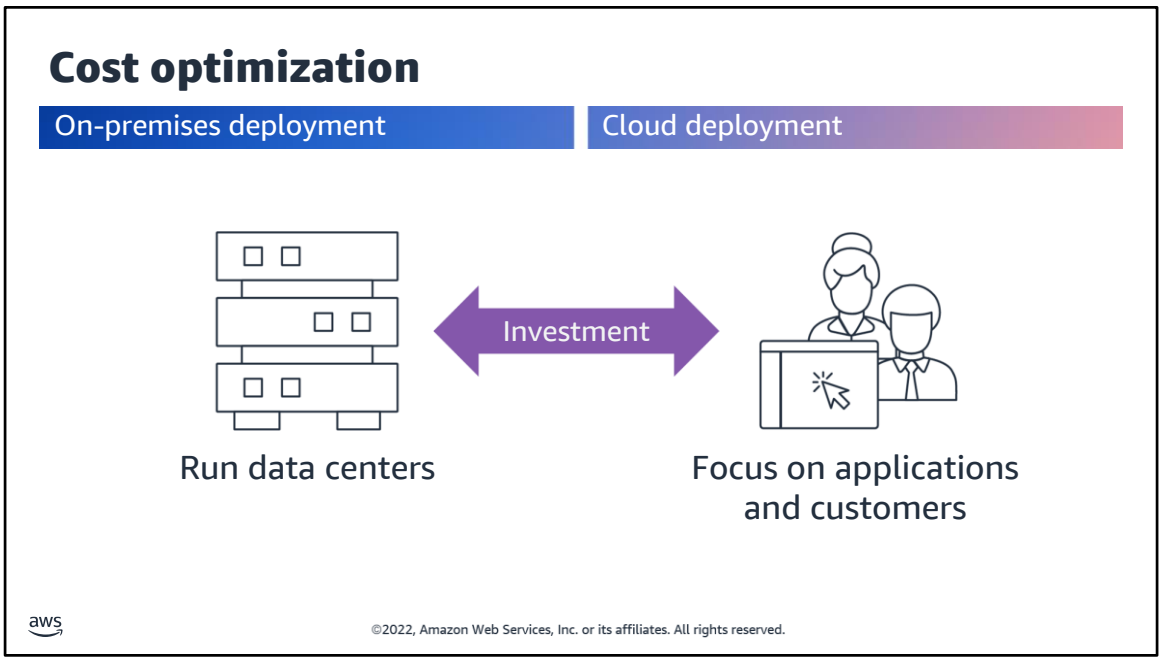


As you begin to learn about cloud computing, also consider *why* a company might choose a particular cloud computing approach to address their business needs.

This topic of the course explores six benefits of cloud computing. As you learn about each benefit, think about how you might have experienced the benefit at work or during other computing activities.

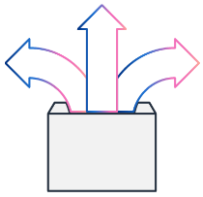


With cloud computing, you can trade upfront expenses for variable expenses.
Upfront expenses refer to data centers, physical servers, and other resources that you invest in before using them. Variable expenses mean that you only pay for computing resources that you consume instead of investing heavily in data centers and servers before you know how you will use them.
By taking a cloud computing approach that offers the benefit of variable expenses, companies can implement innovative solutions while saving on costs.



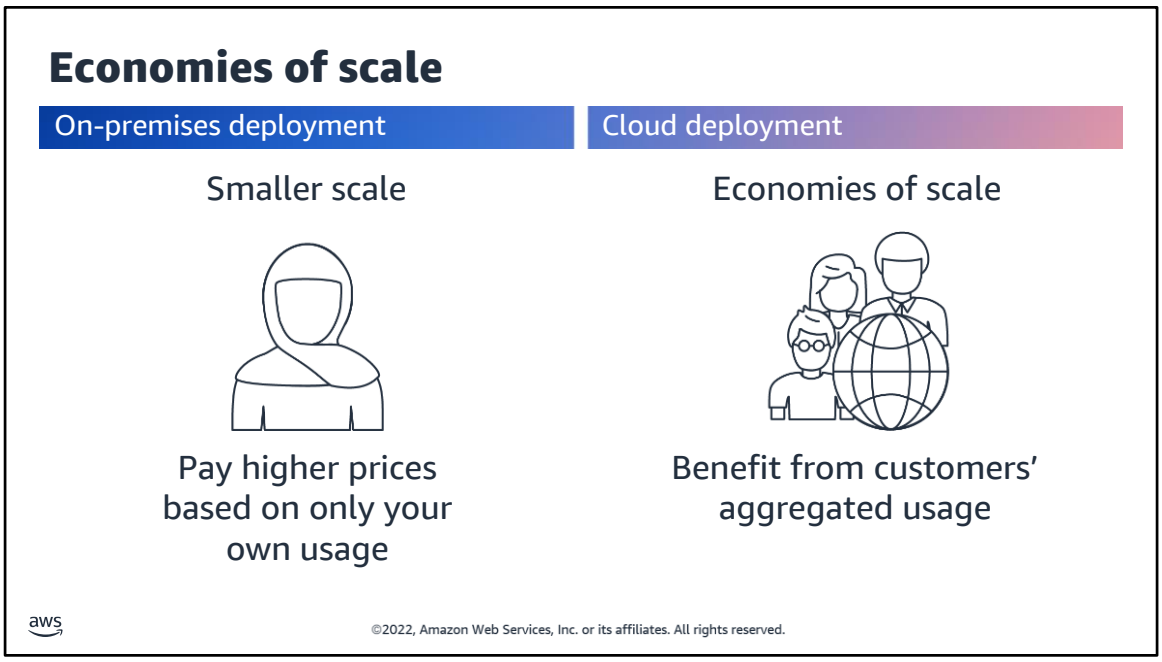
With cloud computing, you can optimize costs by no longer needing to spend money on running and maintaining data centers.
Computing in data centers often requires you to spend more money and time managing infrastructure and servers.
A benefit of cloud computing is the ability to focus less on those tasks and more on applications and customers.

Capacity

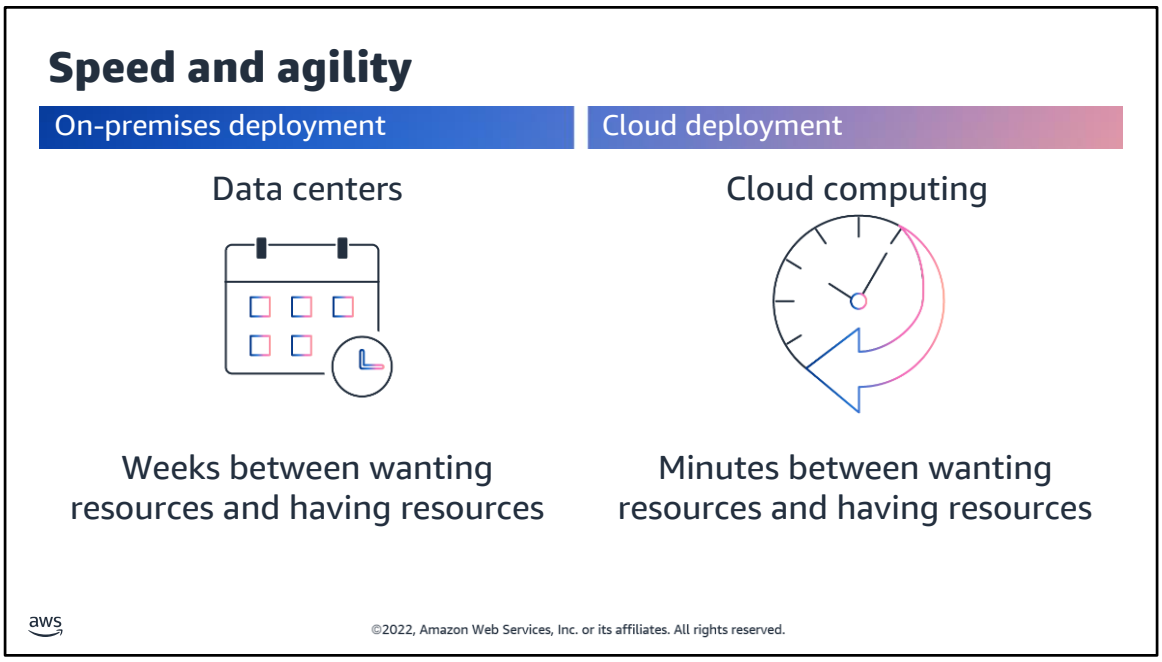
On-premises deployment	Cloud deployment
	
Guessing on your infrastructure capacity needs	Scale in and scale out as needed

aws

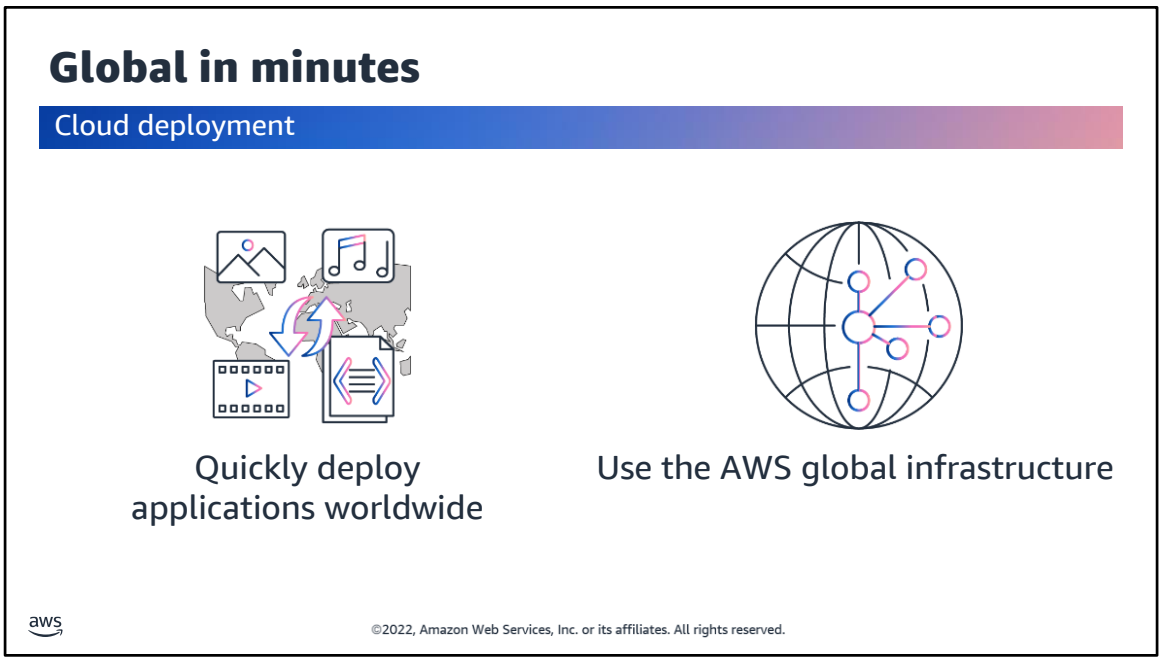
You can stop guessing capacity.
With cloud computing, you don’t have to predict how much infrastructure capacity you will need before deploying an application.
For example, you can launch Amazon EC2 instances when needed, and pay only for the compute time you use. Instead of paying for unused resources or dealing with limited capacity, you can access only the capacity that you need. You can scale in or scale out in response to demand.



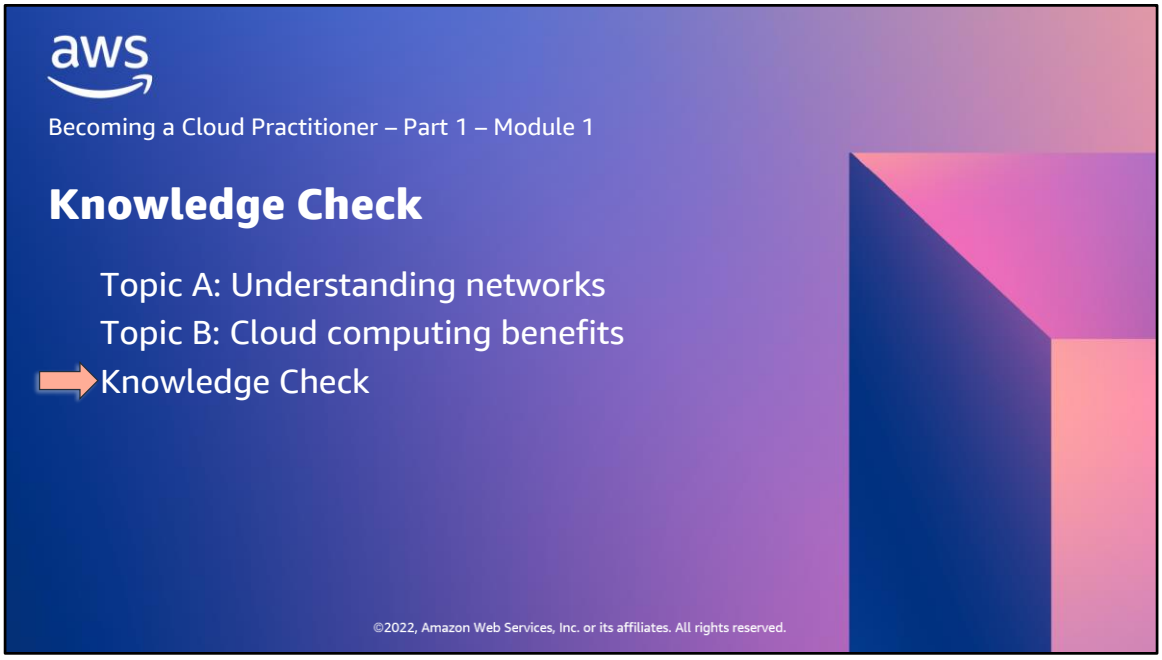
You can benefit from massive economies of scale.
By using cloud computing, you can achieve a lower variable cost than you can get on your own.
Because usage from hundreds of thousands of customers can aggregate in the cloud, AWS can achieve higher economies of scale. The economy of scale translates into lower pay-as-you-go prices.



You can increase speed and agility.
The flexibility of cloud computing makes it easier for you to develop and deploy applications. This flexibility provides you with more time to experiment and innovate.
When computing in data centers, it might take weeks to obtain new resources that you need. In comparison, cloud computing enables you to access new resources in minutes.



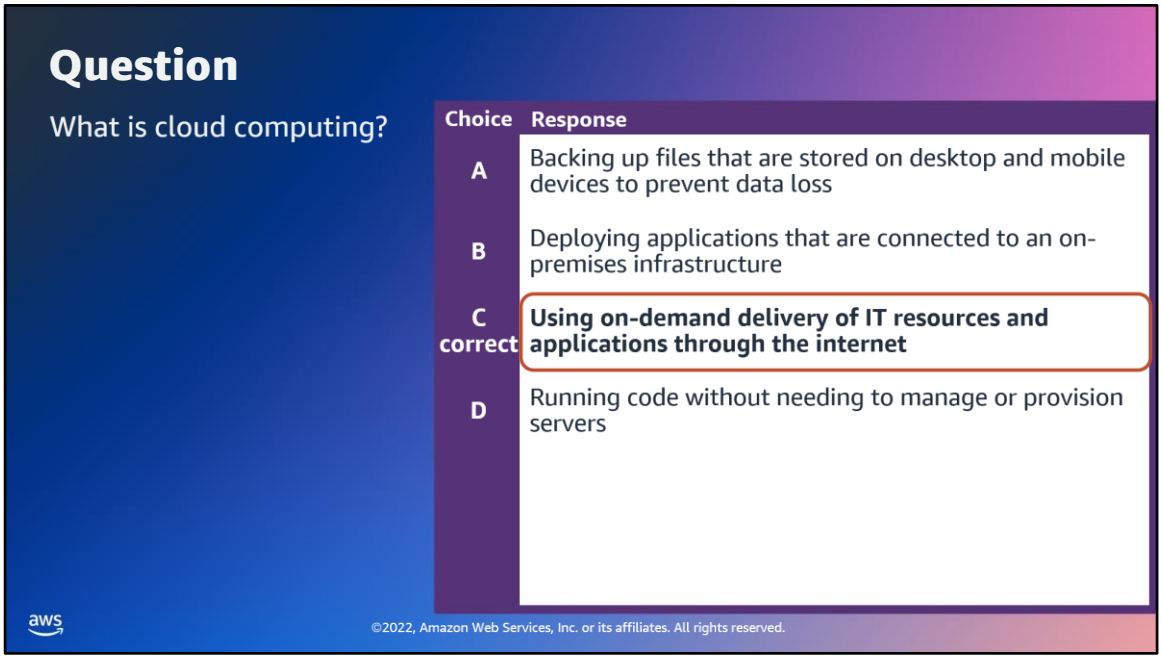
You can go global in minutes.
The global footprint of the AWS Cloud enables you to deploy applications to customers around the world quickly, while providing them with low latency. This means that even if you are located in a different part of the world than your customers, customers can access your applications with minimal delays.
Later in this course, you will explore the AWS global infrastructure in greater detail. You will examine some of the services that you can use to deliver content to customers around the world.



Question

What is cloud computing?

Choice	Response
A	Backing up files that are stored on desktop and mobile devices to prevent data loss
B	Deploying applications that are connected to an on-premises infrastructure
C	Using on-demand delivery of IT resources and applications through the internet
D	Running code without needing to manage or provision servers

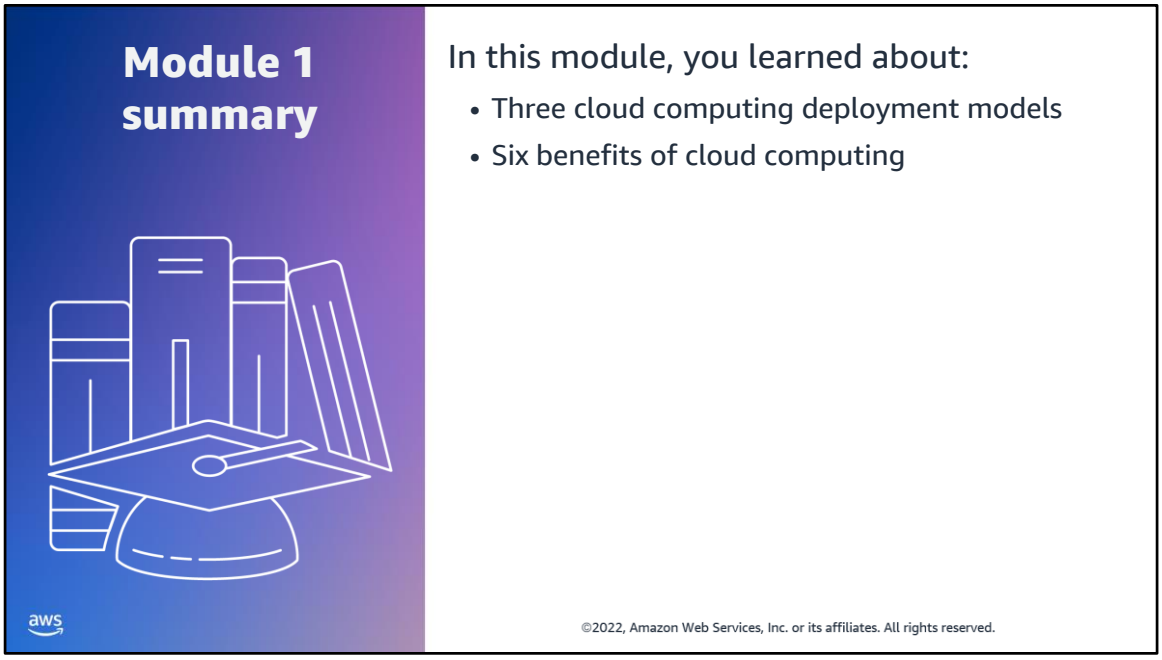


The correct response option is **Using on-demand delivery of IT resources and applications through the internet**.

- As we learned, cloud computing allows us to take advantage of computing services delivered over the internet.

The other response options are incorrect because:

- While you can back up files to the cloud, this response option does not describe cloud computing as a whole.
- Response B describes a sample use case for a hybrid cloud deployment. Remember that cloud computing has cloud and on-premises (or private cloud) deployment models.
- AWS Lambda is an AWS service that lets you run code without needing to manage or provision servers. This description does not describe cloud computing as a whole. AWS Lambda is explained in greater detail in the next module.



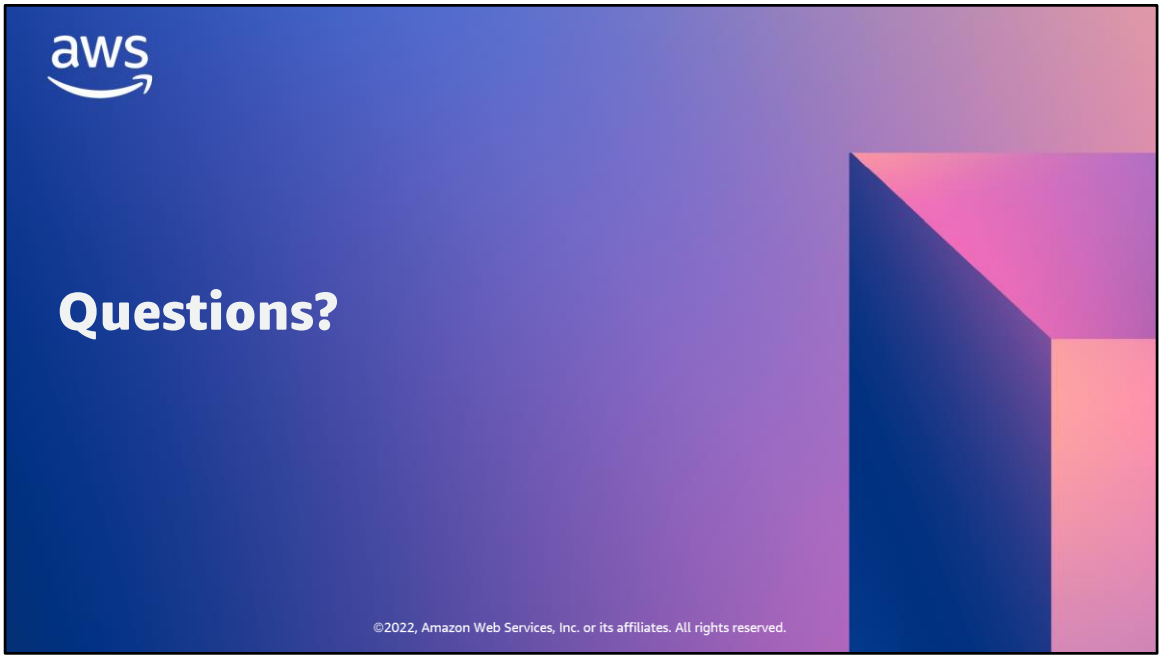
In this module, you learned about three cloud computing deployment models:

- Cloud
- On premises
- Hybrid

You also explored six benefits that cloud computing offers:

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

With cloud computing, you can access services on demand, provision computing resources as needed, and save on costs by paying only for what you use. The next module explores some of the compute services that AWS offers.



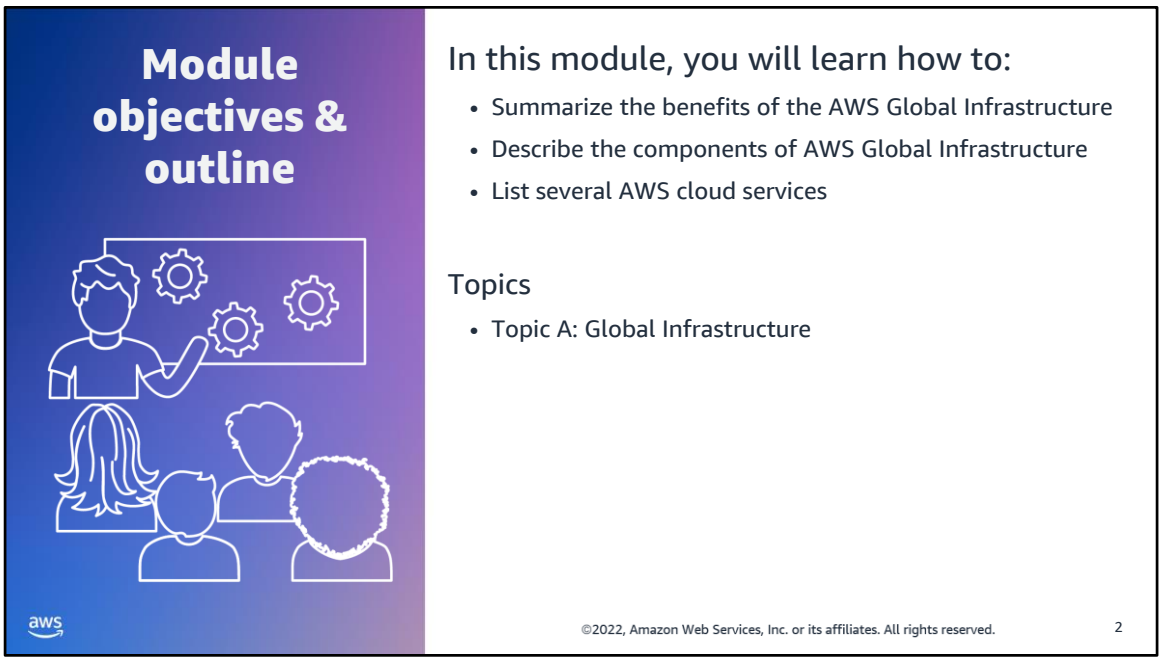
In this module, you will learn about the AWS Global Infrastructure including what Regions and Availability Zones are and how common AWS services work with them.

Module Objectives:

- Summarize the AWS Global Infrastructure benefits
- Describe Availability Zones
- Compare methods for provisioning AWS services.

Topic:

- Topic A: Global Infrastructure



In this module, you will learn about the AWS Global Infrastructure including what Regions and Availability Zones are and how common AWS services work with them.

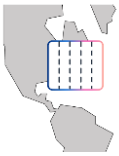
Module Objectives:

- Summarize the AWS Global Infrastructure benefits
- Describe Availability Zones
- Compare methods for provisioning AWS services.

Topic:

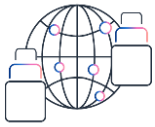
- Topic A: Global Infrastructure

Terminology covered in this module (1 of 3)



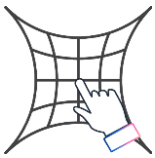
Availability Zone (AZ)
A distinct location within a Region that's insulated from failures in other Availability Zones.

Region
A named set of AWS resources that's in the same geographical area.



Zonal service / Regional service
Zonal services require the assignment of a Region and an AZ. Regional services only require the assignment of a Region.

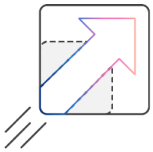
Terminology covered in this module (2 of 3)



Elasticity

The ability of a Cloud service to dynamically grow and shrink based on demands of a workload.

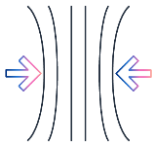
- *Amazon Auto Scaling*



Scalability

The ability of a Cloud service to grow manually as the demands of a workload change over time.

- *Horizontal or vertical scaling mechanisms*



Resiliency

Indicates the ability of a system to both recover and continue operating in the event of a disruption.

- *Clustered servers, redundant workloads, failover mechanisms.*

Terminology covered in this module (3 of 3)

Durability

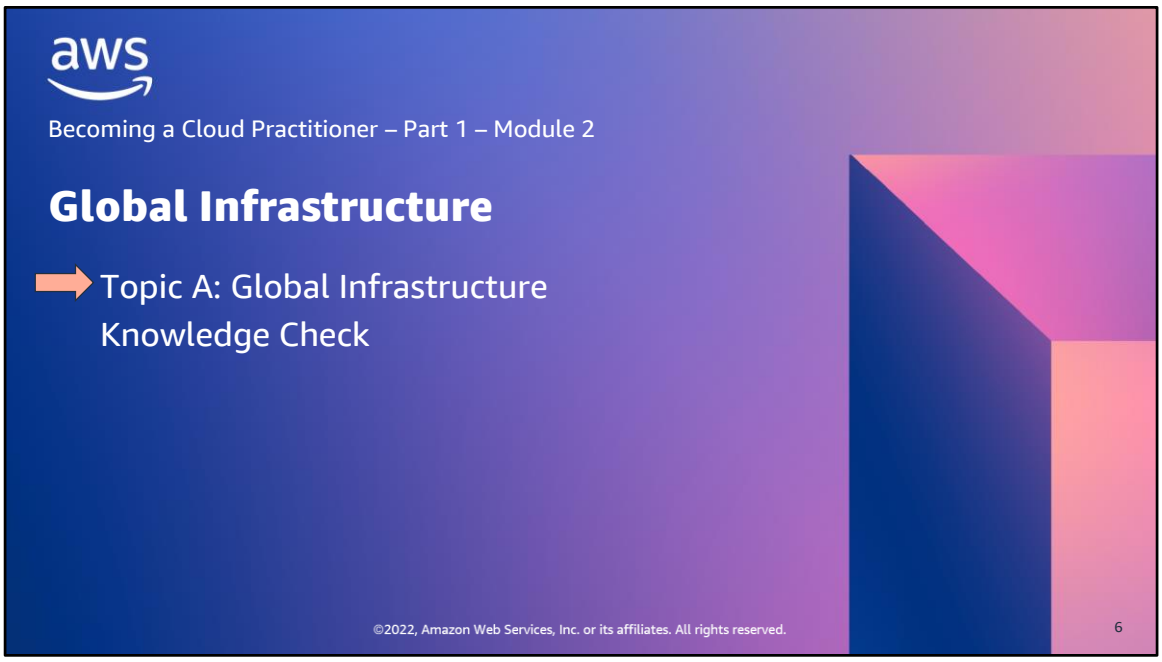
The ability of a Cloud service to ensure long term data stability.

- Data durability of Amazon S3 99.999999999%

Availability

The ability of a Cloud service to be available when it is needed.

- Deploying into multiple AZs and/or Region.

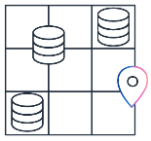


In this topic, you will learn about AWS's global infrastructure and you will be introduced to several of the cloud services that are available.

What is a Region?

A **physical location** (Region) where AWS builds a cluster of data centers.

There are **27 regions** as of September 2022.



A **data center** is a building that houses the physical computing equipment that the AWS region runs on.

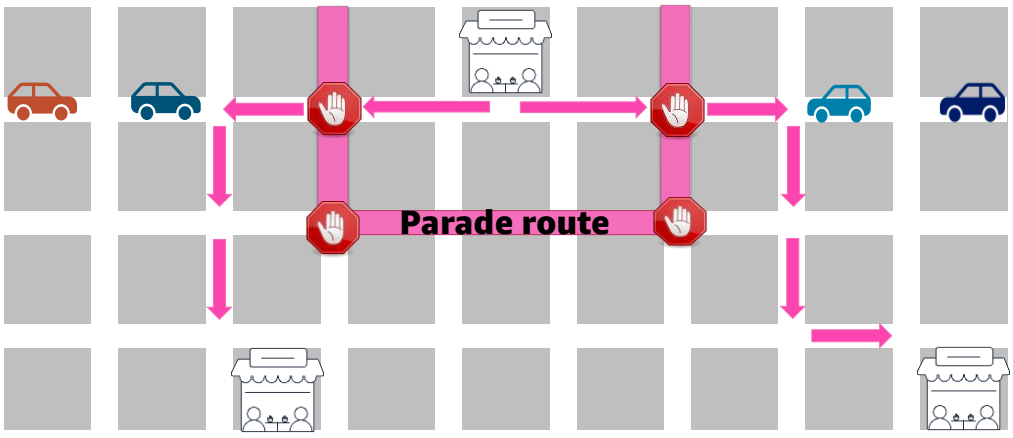
A Region is what connects the physical and logical.
Multiple, physically distant data centers provide **resiliency**.

aws

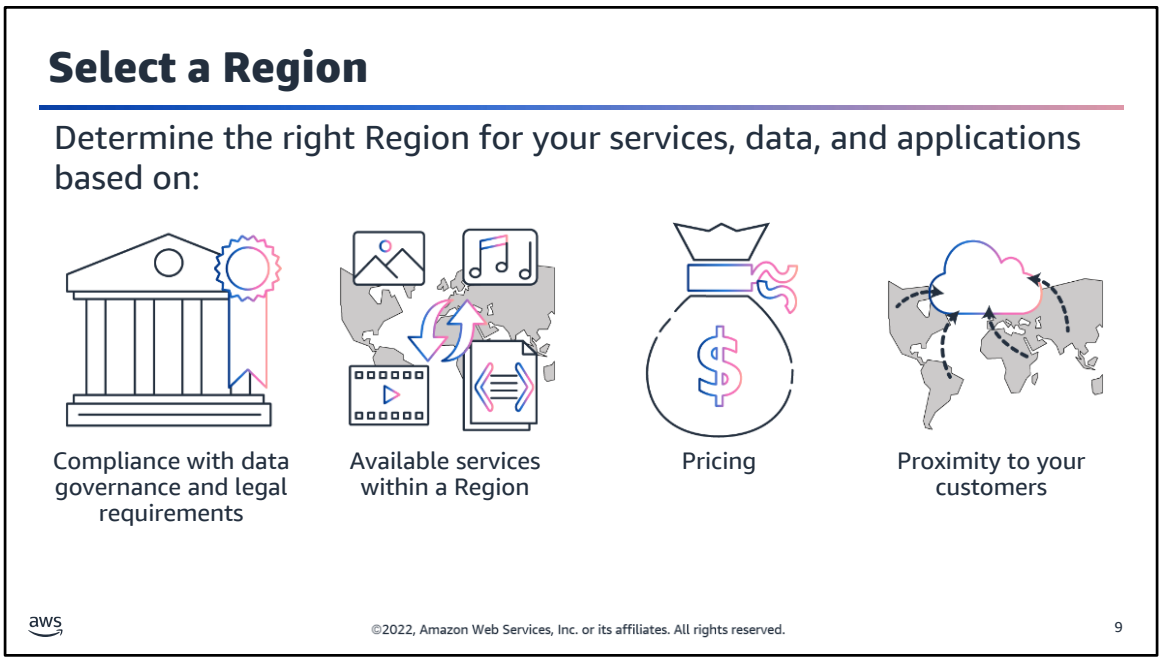
AWS has the concept of a Region, which is **a physical location around the world where we cluster data centers**. Each AWS Region consists of multiple, isolated, and physically separate datacenters within a geographic area.

The AWS Cloud spans **26 geographic regions around the world**, with announced plans for 8 more AWS Regions in Australia, Canada, India, Israel, New Zealand, Spain, Switzerland, and United Arab Emirates (UAE).

Region availability



So what happens if a data center goes offline? Or a region is unavailable?
To understand the AWS Global Infrastructure, begin with an example from the coffee shop. Customers are unable to get to the coffee shop because a parade is blocking the road.
If an event such as a parade, flood, or power outage impacts one location, customers can still get their coffee by visiting a different location only a few blocks away.
This is similar to how the AWS Global Infrastructure works.



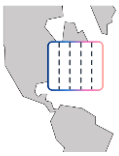
When determining the right Region for your services, data, and applications, consider these four business factors:

- **Compliance with data governance and legal requirements** – Depending on your company and location, you might need to run your data out of specific areas. For example, if your company requires all of its data to reside within the boundaries of the UK, you would choose the London Region. Not all companies have location-specific data regulations, so you might need to focus more on the other three factors.
- **Proximity to your customers** – Selecting a Region that is close to your customers will help you to get content to them faster. For example, your company is based in Washington, DC, and many of your customers live in Singapore. You might consider running your infrastructure in the Northern Virginia Region to be close to company headquarters, and run your applications from the Singapore Region.
- **Available services within a Region** – Sometimes, the closest Region might not have all the features that you want to offer to customers. AWS is frequently innovating by creating new services and expanding on features within existing services. However, making new services available around the world sometimes requires AWS to build out physical hardware one Region at a time. Suppose that your developers want to build an application that uses Amazon Braket (AWS quantum computing platform). As of this course, Amazon Braket is not yet available in every AWS Region around the world, so your developers would have to run it in one of the Regions that already offers it.
- **Pricing** – Suppose that you are considering running applications in both the United States and Brazil. The way Brazil's tax structure is set up, it might cost 50% more to run the same workload out of the São Paulo Region compared to the Oregon Region. You will learn in more detail that several factors determine pricing, but for now know that the cost of services can vary from Region to Region.

Spanning multiple Regions helps to keep your applications and data safe from disasters. However, this isn't the only way to get high availability and fault tolerance in the AWS Global Infrastructure. Regions are made up of multiple **Availability Zones**.

What is an Availability Zone (AZ)?

Availability Zones are distinct locations within a Region that are engineered to be isolated from failures in other AZs.



An AZ can only exist in a single Region.

AZs are logical representations of how cloud resources are grouped.

The hardware that supports AZs is located in **data centers**.

The AWS Cloud spans **87 AZs** as of September 2022.

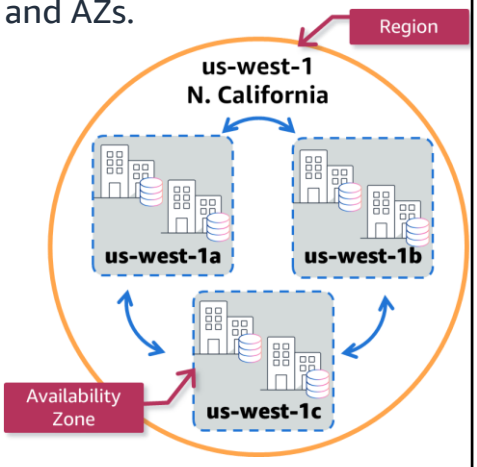
aws

Availability Zones are **distinct locations within an AWS Region that are engineered to be isolated from failures in other Availability Zones**. They provide inexpensive, low-latency network connectivity to other Availability Zones in the same AWS Region. Each region is completely independent.
An Availability Zone is a single data center or a group of data centers within a Region. If a disaster occurs in one part of the Region, the geographical distance helps to ensure that not all Availability Zones are affected.
Availability Zones are located tens of miles apart from each other. This helps them to provide interconnectivity to support the services and applications that run within a Region.
Even though Availability Zones are close enough to have low latency (the time between when content is requested and received), they are not built directly next to one another.
The AWS Cloud spans **84 Availability Zones**, with announced plans for 24 more Availability Zones in Australia, Canada, India, Israel, New Zealand, Spain, Switzerland, and United Arab Emirates (UAE).

Regions and Availability Zones

Naming conventions are used for Regions and AZs.

- Region name has two parts
  - Region identifier (us-west-1)
  - Familiar name (N. California)
- AZs are named based on their Region
  - AZ identifier (us-west-1a)
  - AZ identifier (us-west-1b)
  - AZ identifier (us-west-1c)



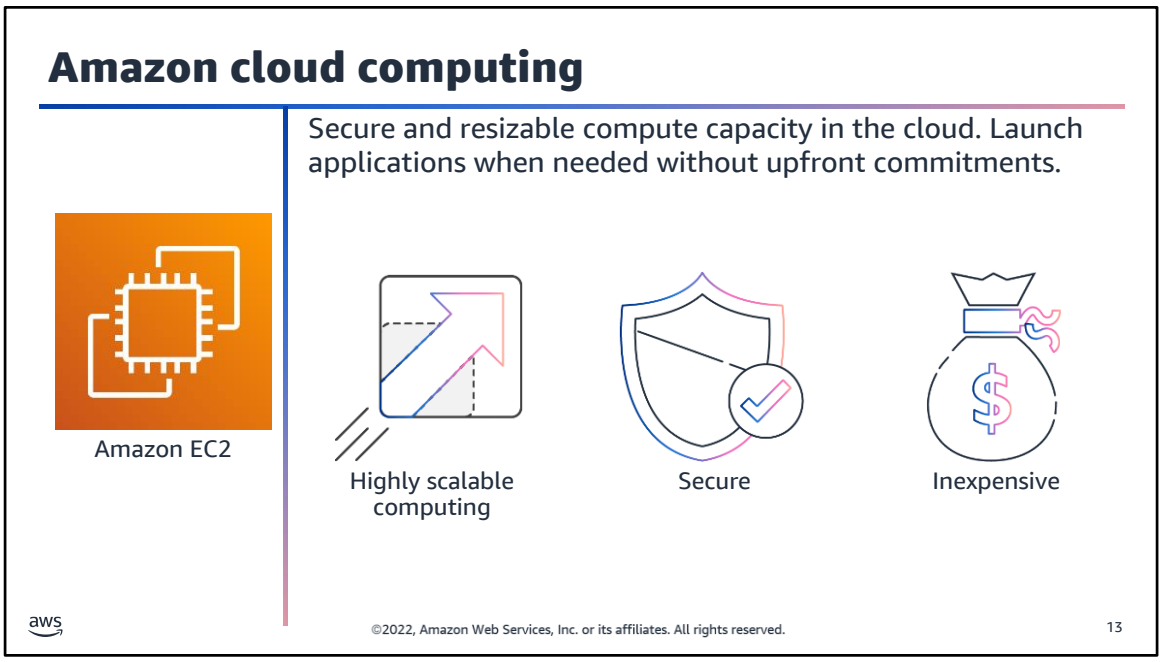
This slide contains…
****For Accessibility:** info **End Description**

What is a Cloud service?

Cloud services are products which deliver compute power, data storage, applications, and more over the internet.

- Office 365
- Gmail
- DropBox
- Amazon EC2
- Amazon S3
- Amazon RDS

Cloud services are products which deliver compute power over the internet including servers, storage, databases, networking, software, analytics, and intelligence.

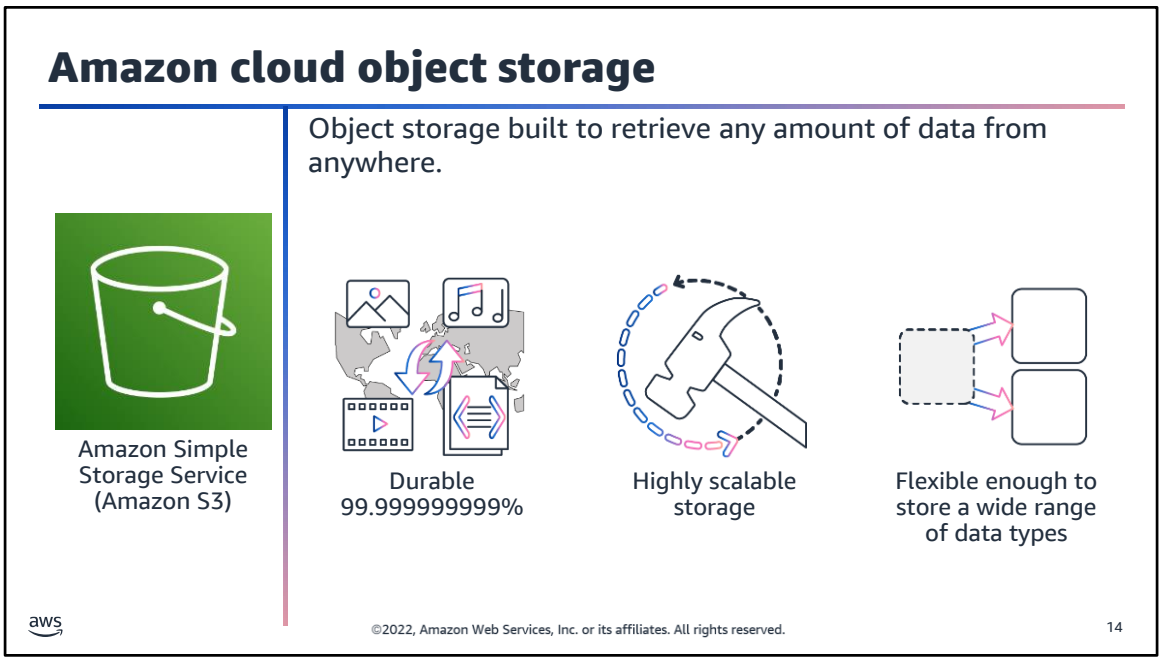


Amazon EC2 provides secure, resizable compute capacity in the cloud as Amazon EC2 instances.
Imagine that you are responsible for the architecture of your company's resources and must support new websites. With traditional on-premises resources, you would:

1. Spend money upfront to purchase hardware.
2. Wait for the servers to be delivered to you.
3. Install the servers in your physical data center.
4. Make all the necessary configurations.

By comparison, with an Amazon EC2 instance, you would use a virtual server to run applications in the AWS Cloud. You could:

- Provision and launch an Amazon EC2 instance within minutes
- Stop using it when you finish running a workload
- Pay only for the compute time you use when an instance is running, not when it is stopped or shut down
- Save costs by paying only for server capacity that you need or want



Amazon S3 offers a range of storage classes designed for different use cases. For example, you can store mission-critical production data in S3 Standard for frequent access, save costs by storing infrequently accessed data in S3 Standard-IA or S3 One Zone-IA, and archive data at the lowest costs in S3 Glacier Instant Retrieval, S3 Glacier Flexible Retrieval, and S3 Glacier Deep Archive.

You can store data with changing or unknown access patterns in S3 Intelligent-Tiering, which optimizes storage costs by automatically moving your data between four access tiers when your access patterns change. These four access tiers include two low-latency access tiers optimized for frequent and infrequent access, and two opt-in archive access tiers designed for asynchronous access for rarely accessed data.

Cloud services, Regions, and AZs

All AWS services logically reside within a Region.

Some services reside directly in a Region, **Regional services**.

Some services reside in AZs within a Region, **Zonal services**.

Amazon S3

Each bucket is created in a specific **Region**.

Amazon EC2

Each instance is assigned to a specific **AZ**.

aws

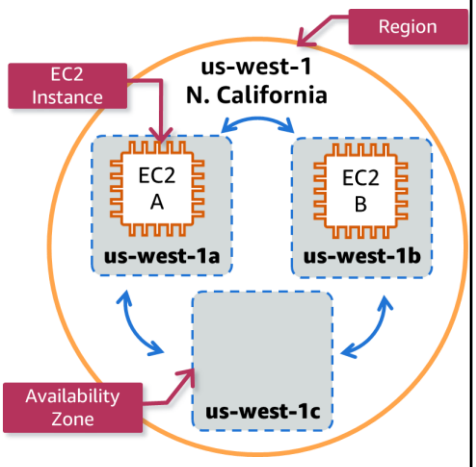
Cloud services are products which deliver compute power over the internet including servers, storage, databases, networking, software, analytics, and intelligence.

All AWS Services are available within a Region. Different regions may offer different services. Within the region, some services may need to be assigned to an Availability Zone as well. Services such as Amazon Simple Storage Service, or Amazon S3, do not reside within an AZ. However, the majority of AWS services, like Amazon EC2, require you to select a specific AZ when you launch an instance of the service.

Building service resilience

Not all services are resilient by default.

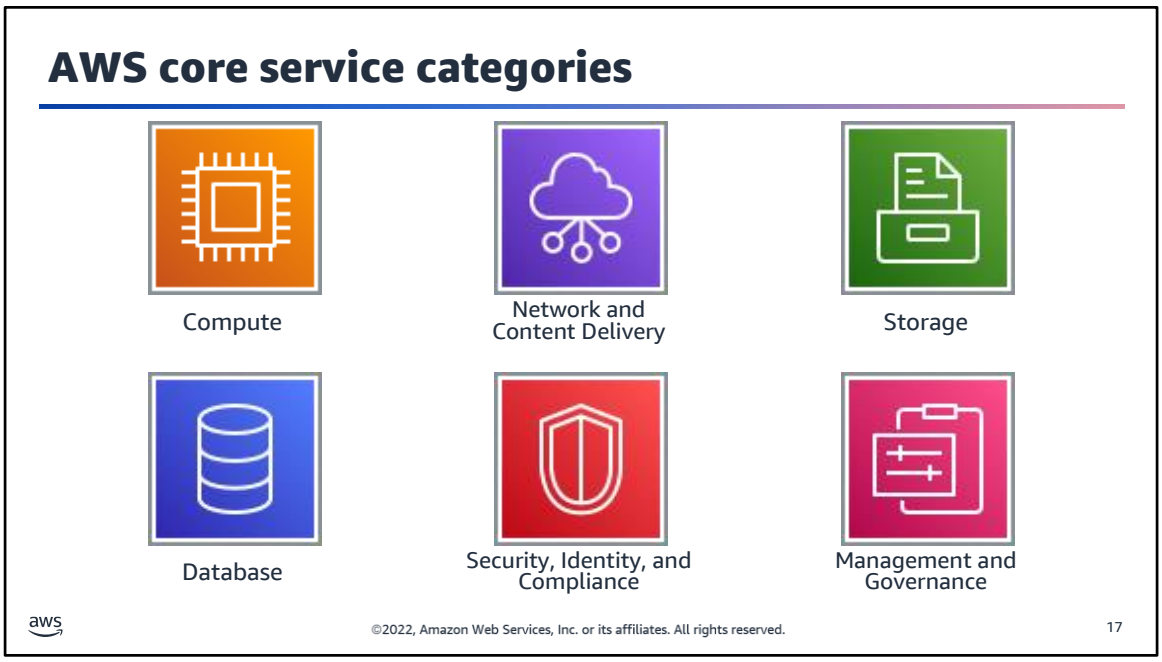
- **Servers**:
  - Build applications to use Amazon EC2 clusters that contain instances in multiple AZs.
- **Storage**:
  - Amazon S3 is resilient across AZs by default.
  - An Amazon S3 bucket exists in only one region.
    - Multi-region replication can be enabled.



Here's an example. Suppose that you're running an application on a single Amazon EC2 instance in the Northern California Region. The instance is running in the us-west-1a Availability Zone. If us-west-1a were to fail, you would lose your instance.

A best practice is to build applications using groups of Amazon EC2 instances, called clusters, across at least two AZs. In this example, you might choose to run a second Amazon EC2 instance in us-west-1b. If us-west-1a were to fail, your application would still be running in us-west-1b.

Object storage with Amazon S3 is resilient across AZs within the region. A single S3 bucket exists in a single region and cannot extend beyond that region. To provide additional resiliency and availability, you can enable multi-region replication. This allows content from the bucket to be duplicated in a bucket that exists in another region.



Throughout this course, you will explore AWS services in categories such as:

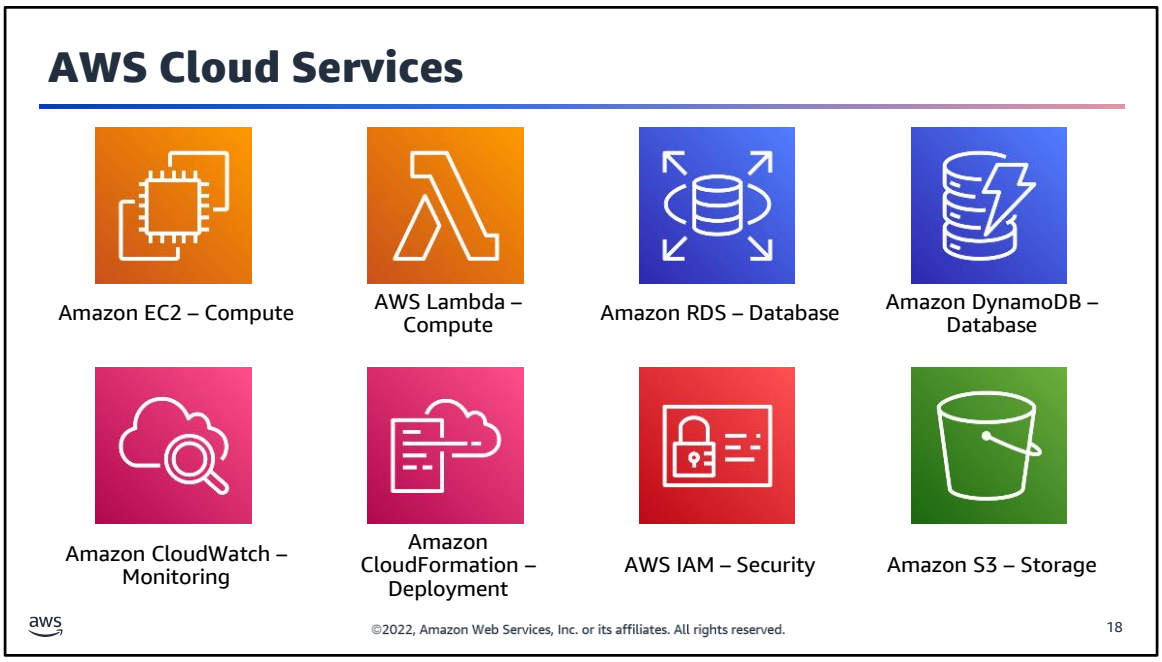
- Compute
- Networking and Content Delivery
- Storage
- Database
- Security, Identity, and Compliance
- Management and Governance

As you begin to learn about AWS services, consider the features and benefits of a single service in addition to how you can potentially combine the services to offer a solution for your business needs.

For example, suppose that the owners of the coffee shop want to create a new application for customers. They might use the following types of services:

- Database service to store customer information
- Management and governance services to ensure that the application is designed in accordance with AWS best practices
- Networking and content delivery services to deliver websites and videos to customers

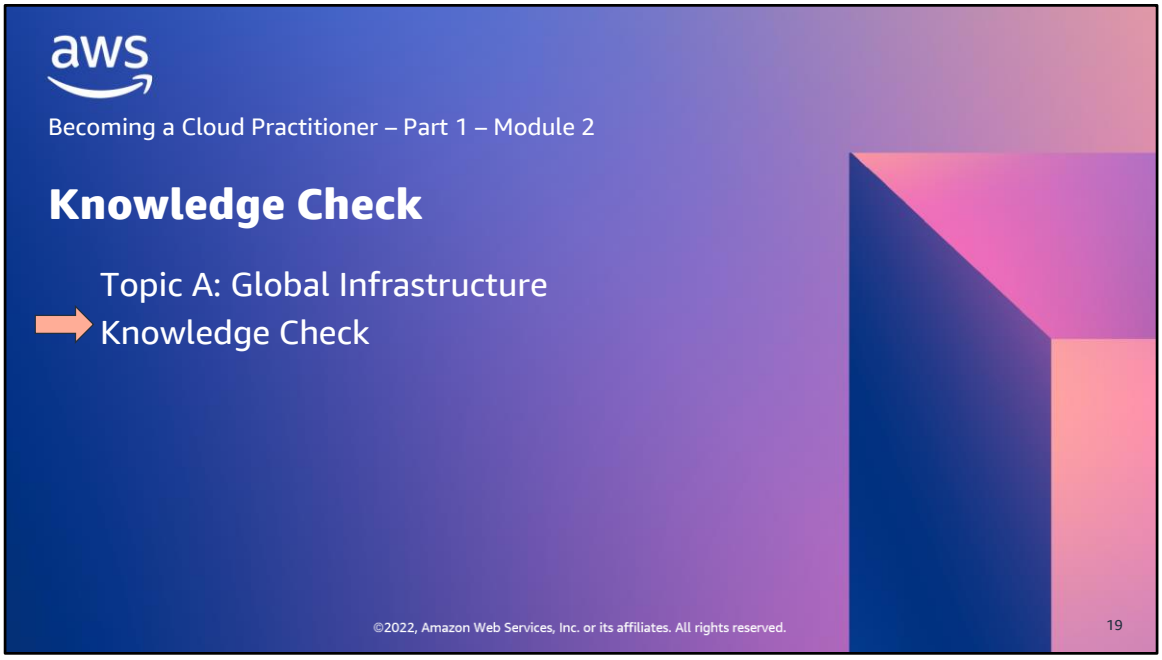
As a reminder, you do not need to know all the services that AWS offers. This course focuses on foundational concepts of the AWS Cloud.



In 2006, AWS began offering IT infrastructure services to businesses as web services. This is now commonly known as cloud computing.

AWS services can be implemented for use cases across industries and businesses of varying sizes. These businesses include enterprises, startups, small- and medium-sized businesses, and AWS customers in the public sector. Even if you have not yet created your own AWS account, you have probably used applications that run in the AWS Cloud. Some examples of applications that run in the AWS Cloud include video streaming services, photo hosting applications, hotel reservation websites, and more.

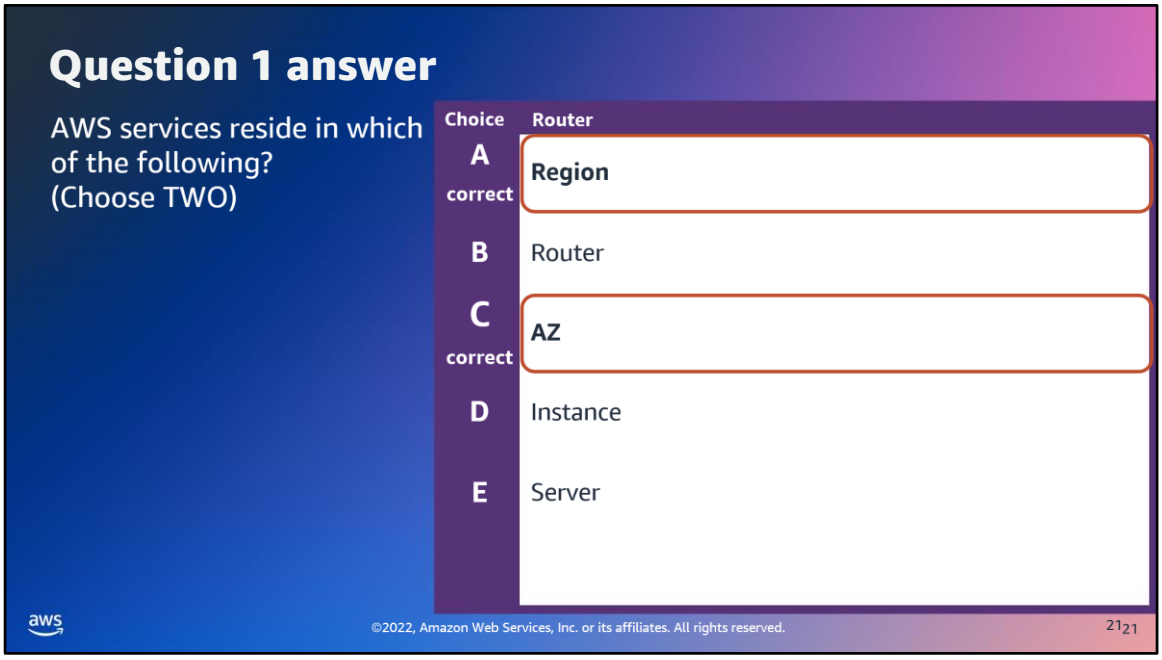
The AWS Cloud is a comprehensive and broadly adopted cloud service. AWS offers more than 200 services from data centers globally. Some of the service categories included in the console are Compute, Containers, Storage, and Database. You do not need to know all the services. This course focuses on foundational concepts of the AWS Cloud. In the next section, you will explore some of the benefits of cloud computing.


aws
Becoming a Cloud Practitioner – Part 1 – Module 2
Knowledge Check
Topic A: Global Infrastructure
Knowledge Check
©2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.
19

Question 1

AWS services reside in which of the following?
(Choose TWO)

Choice	Response
A	Region
B	Router
C	AZ
D	Instance
E	Server



The correct answers are **Region** and **AZ**.

- AWS Services are classified as *Regional*, like Amazon S3, or *zonal*, like Amazon RDS.

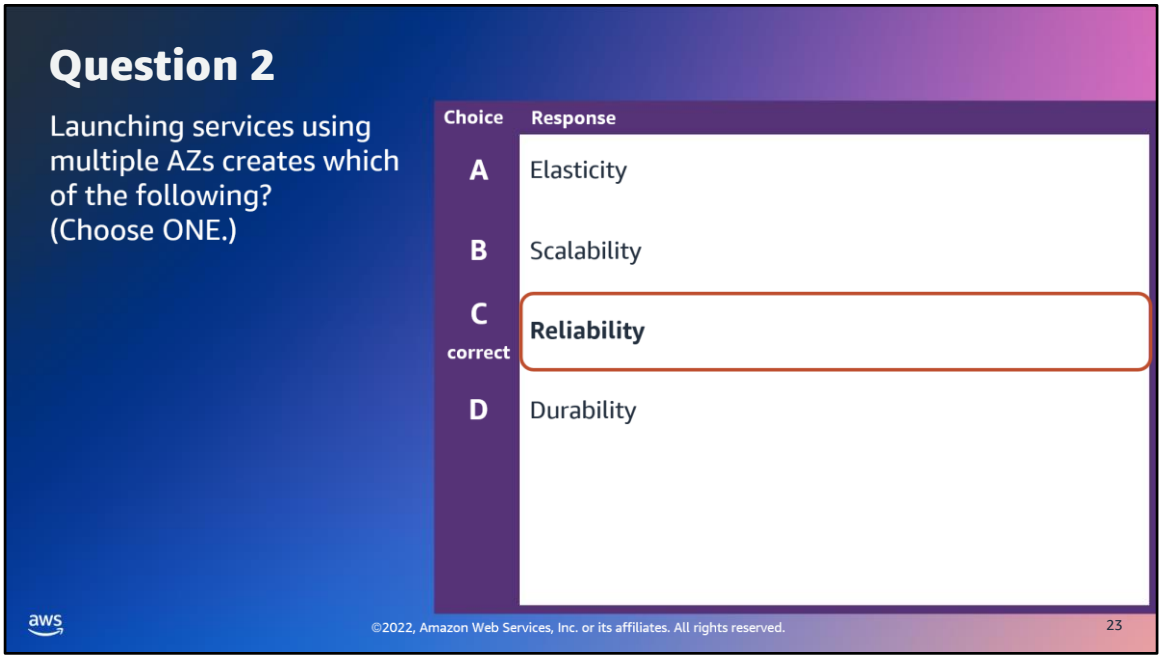
The other response options are incorrect because:

- Services are not attached to physical devices such as routers or servers. Although an AWS service may use instances, the service itself does not reside in the service.

Question 2

Launching services using multiple AZs creates which of the following? (Choose ONE.)

Choice	Response
A	Elasticity
B	Scalability
C	Reliability
D	Durability



The correct answer is **Reliability**.

- Implementing multiple AZs within a service creates reliability in the event of a failure of an instance in a single AZ.

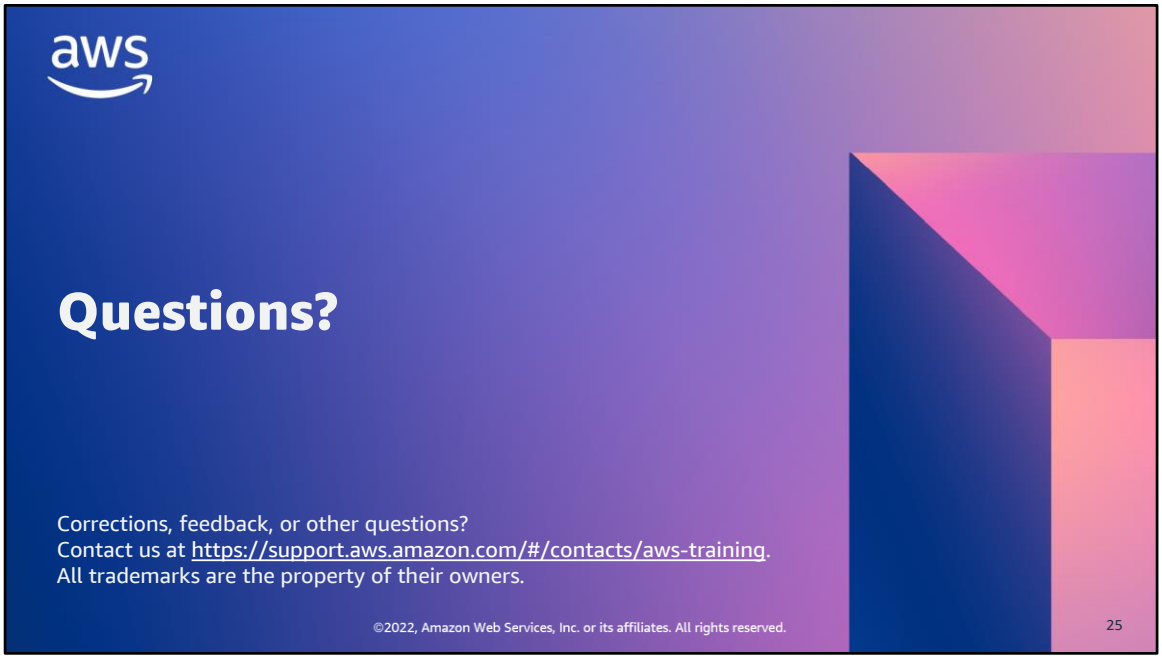
The other response options are incorrect because:

- Elasticity is the ability of a Cloud service to dynamically grow and shrink based on demands of a workload. In the example of Amazon EC2, Elasticity is called auto-scaling.
- Scalability is the ability of a Cloud service to grow with the growth of a workloads needs over time.
- Durability is the ability of a Cloud service to ensure long term data stability.

Module summary

Covered in this module:

- Summary of the benefits of the AWS Global Infrastructure
- Described the components of AWS Global Infrastructure
- Listed several AWS cloud services


aws
Questions?
Corrections, feedback, or other questions?
Contact us at https://support.aws.amazon.com/#/contacts/aws-training.
All trademarks are the property of their owners.
©2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.
25

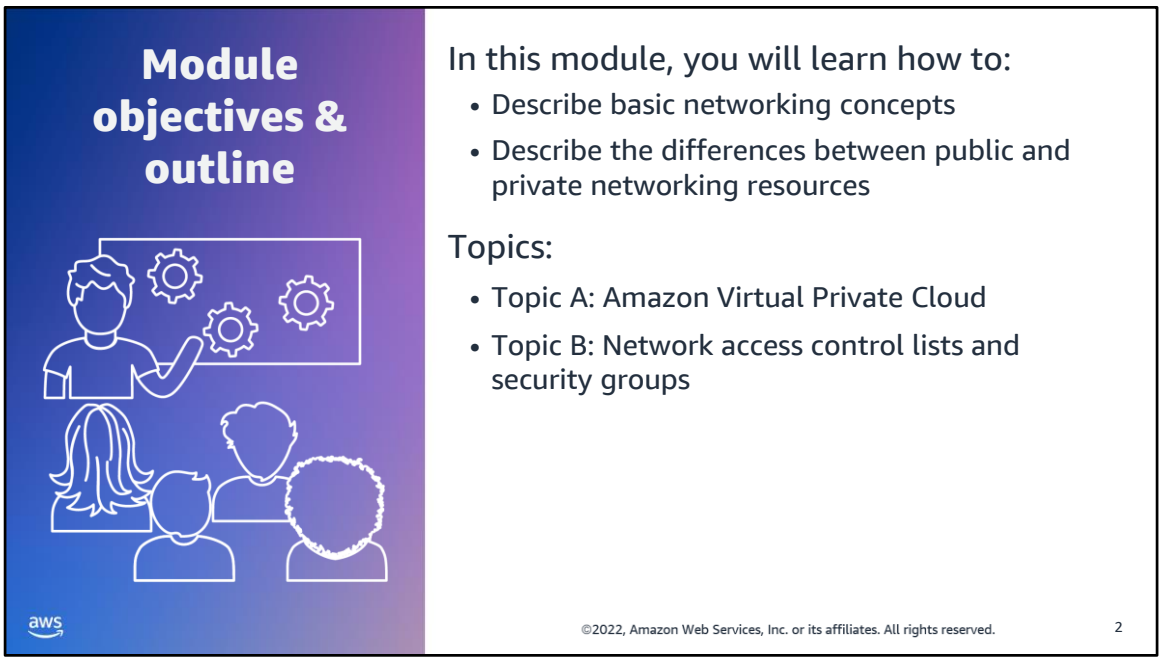
In this module, you will learn about Amazon Virtual Private Cloud (Amazon VPC), network access control lists (ACLs), and security groups.

Module Objectives:

- Describe basic networking concepts
- Describe the differences between public and private networking resources

Topics:

- Topic A: Amazon Virtual Private Cloud (Amazon VPC)
- Topic B: Network access control lists and security groups



In this module, you will learn about Amazon Virtual Private Cloud (Amazon VPC), network access control lists (ACLs), and security groups.

Module Objectives:

- Describe basic networking concepts
- Describe the differences between public and private networking resources

Topics:

- Topic A: Amazon Virtual Private Cloud (Amazon VPC)
- Topic B: Network access control lists and security groups

Acronyms covered in this module

OS — Operating System
A program that communicates between the hardware and software supported by that system.

DNS — Domain Name System
A naming system used to identify servers reachable over the internet or private IP networks.

CIDR — Classless Inter-domain Routing
A method for allocating IP addresses and performing routing.

HTTP — Hypertext Transfer Protocol
This is a protocol used to transmit data over the internet.

IP — Internet Protocol
This is a protocol that defines addressing on a computer network.
- IPv4, IPv6

ACL — Access Control List
A virtual firewall that controls inbound and outbound traffic at the subnet level.

aws

Terminology covered in this module (1 of 3)

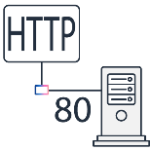
Data Packets

A unit of data made into a small package for travel along a network path.

Protocol

A set of rules that determine how data is transmitted between devices in the same network.

- HTTP, SMTP, SSL.



Port/Port number

A **Port** is a logical construct that identifies a specific process or type of network service or protocol.

A well-known **Port number** is the aligned with a specific transport.

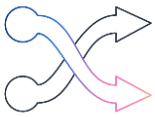
- Port 80 HTTP; Port 443 HTTPS; port 53 DNS.

aws

Terminology covered in this module (2 of 3)

IP Address

A unique number that identifies a device on the internet or local network.
Addresses are made up of 4 numbers with values of 0 – 255.
- 127.0.0.1; 169.254.0.1; 176.16.0.1; 192.168.0.1;



Routing

The process of selecting a path across one or more networks and transmitting data from the source network to the destination network.

Subnet/Subnetting/Subnet Mask

A logical subdivision on an IP network. Subnetting is the process of creating the subdivision. This is enforced using a subnet mask.

aws

Terminology covered in this module (3 of 3)

Switch

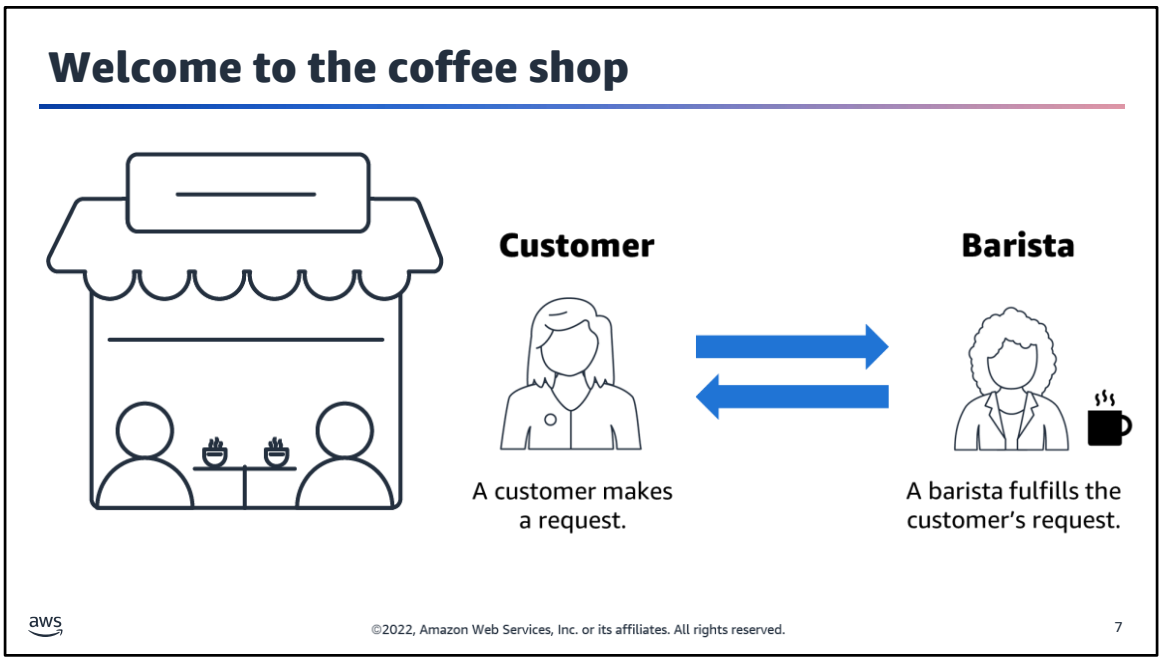
A device that forwards data packets between devices in a single network. It is not aware of network addressing.

Router

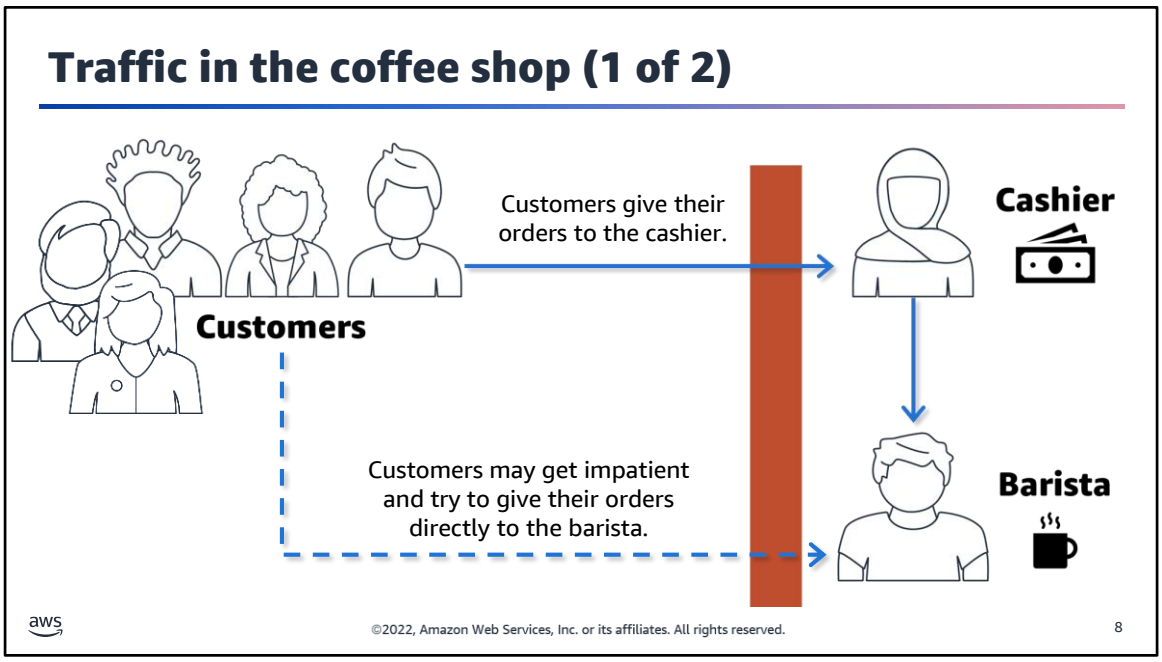
A device that forwards data packets between networks.

Gateway

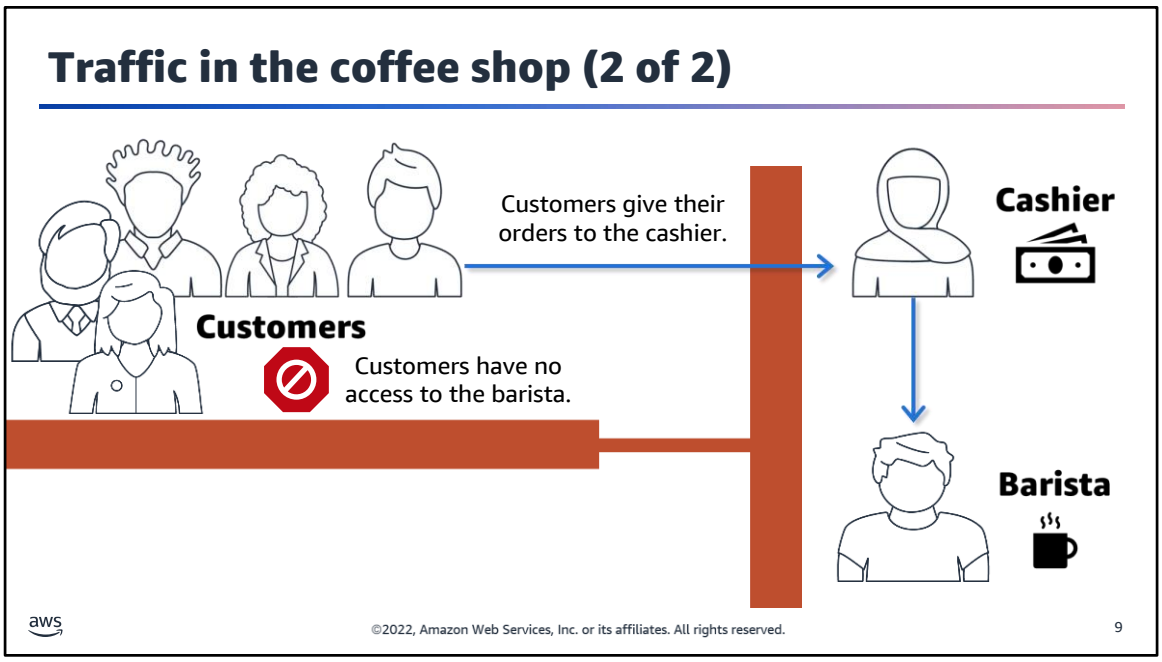
A device or node that connects two different networks by **translating communications** from one protocol to another.



We are going to continue using the example of our coffee chop. Remember, in a coffee shop, a customer makes a request, and then, a barista fulfills the customer's request. Think of a barista as a virtual server that fulfills requests. A barista can fulfill requests by providing customers with items such as coffee, tea, or pastries. A virtual server can fulfill requests by providing a client with items such as videos, photos, or static webpages.



Our coffee shop is starting to become popular. Suppose that some customers try to skip the cashier line and give their orders directly to the barista. This disrupts the flow of traffic and results in customers accessing a part of the coffee shop that is off limits to them. This can cause every customers order to be delayed as the barista now has to focus on redirecting the customers to the cashier instead of making coffee.
****For Accessibility:** This image depicts customers waiting in line to pay a cashier for their order and the order being passed to a Barista. **End Description****



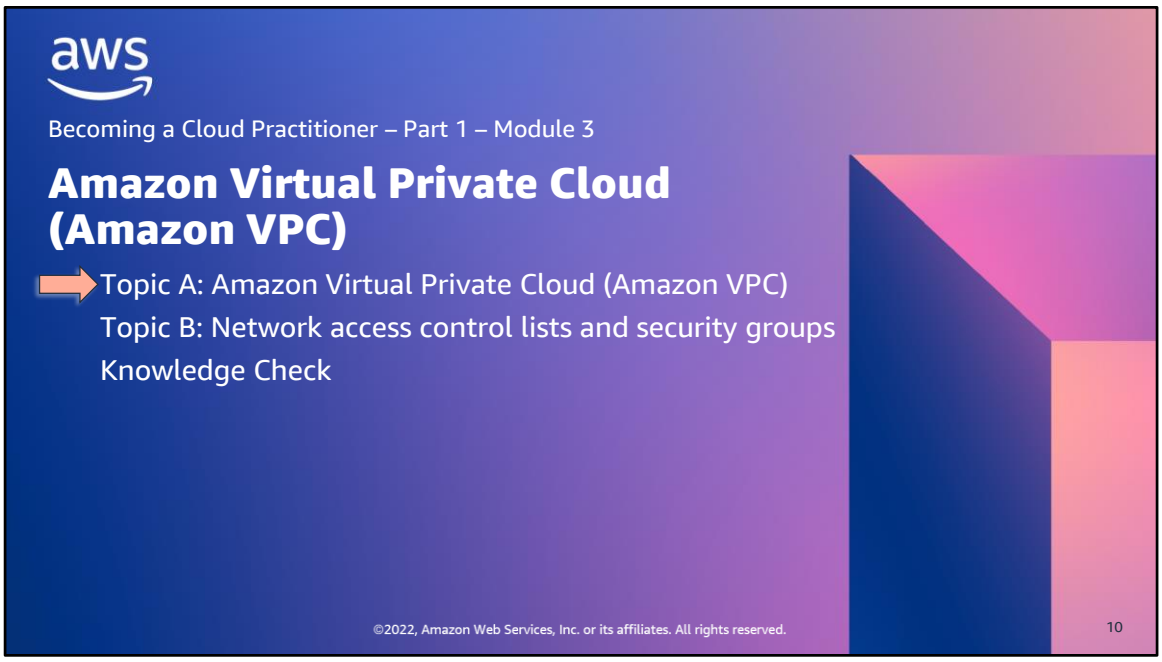
To fix this, the owners of the coffee shop completely separate the customer counter area by placing the cashier and the barista in separate workstations. The cashier's workstation is public facing and designed to receive customers. The barista's area is private. The barista can still receive orders from the cashier but not directly from customers.

This is similar to how you can use AWS networking services to isolate resources and determine exactly how network traffic flows.

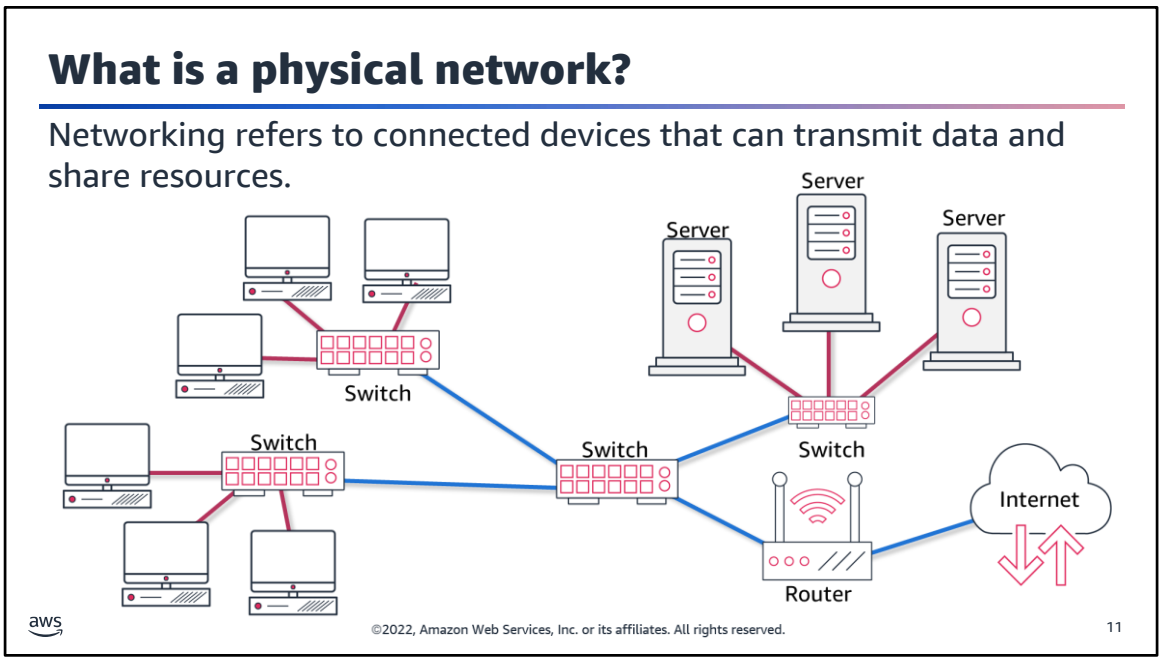
Imagine the millions of customers who use AWS services. Also, imagine the millions of resources that these customers have created, such as Amazon EC2 instances. Without boundaries around all of these resources, network traffic would be able to flow between them, unrestricted.

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

****For Accessibility:** This image depicts the new coffee shop layout that separates the customers from the barista to ensure that all orders go to the cashier. **End Description****



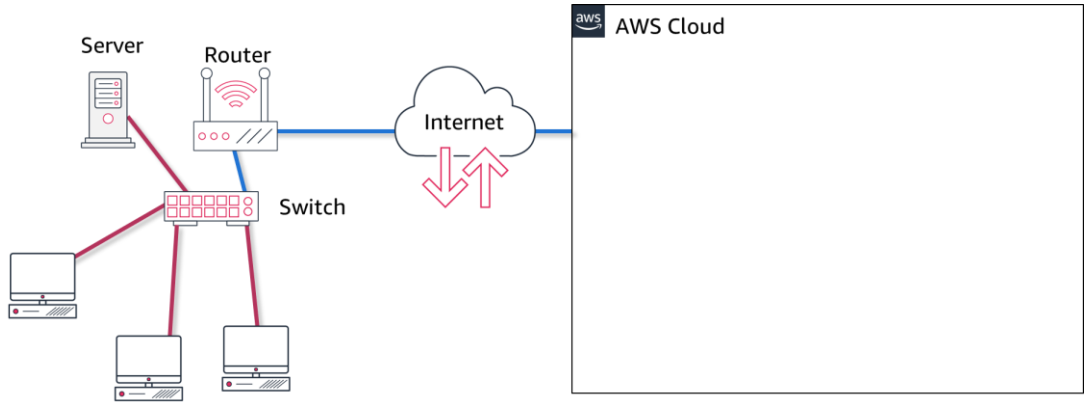
In this topic, you will learn about the Amazon VPC service.



Networking refers to connected devices that can transmit data and share resources. Computers cannot communicate with other systems unless there is a **network** connecting the computers to each other.
Your PC connects to a switch. There can be multiple switches on a network. Servers are computers that runs software or provides storage for other computers on a network. Servers must also be connected to one another using switches.
Access to the internet or a network in a separate building is provided by **routers**. A **Router** is a device that connects a network to the internet or another network.

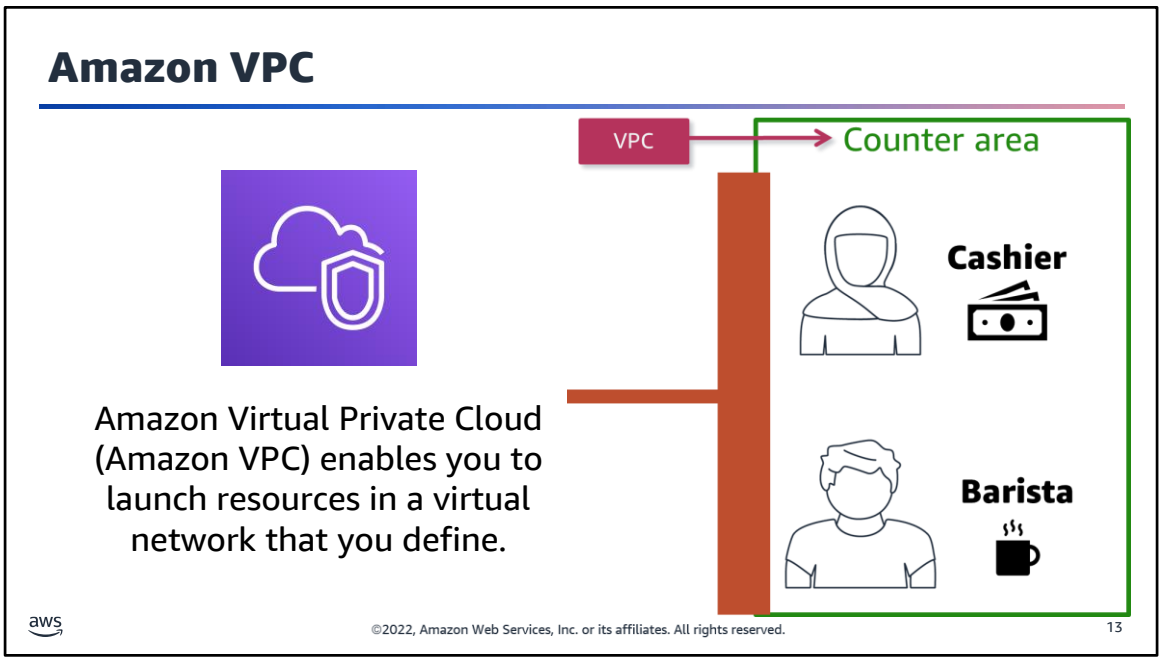
How does a physical network connect to the cloud?

The internet is used to connect a physical network to the AWS cloud.

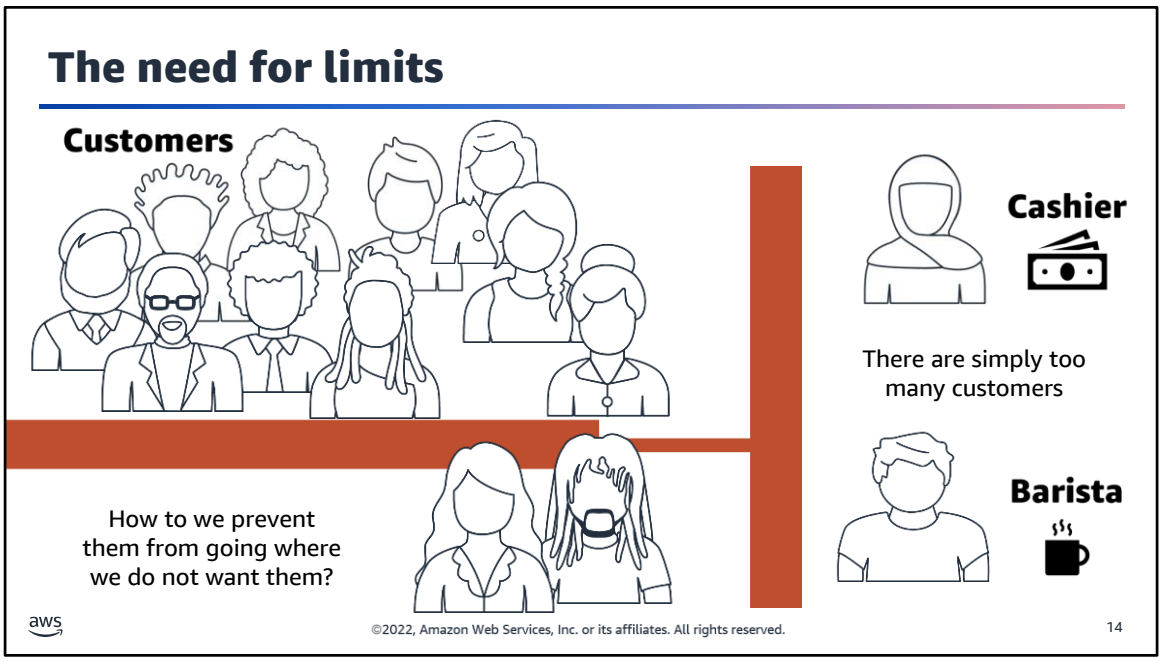


12

Once you have connected to the internet, you also have the ability to connect to the AWS cloud and use the services that AWS has to offer.



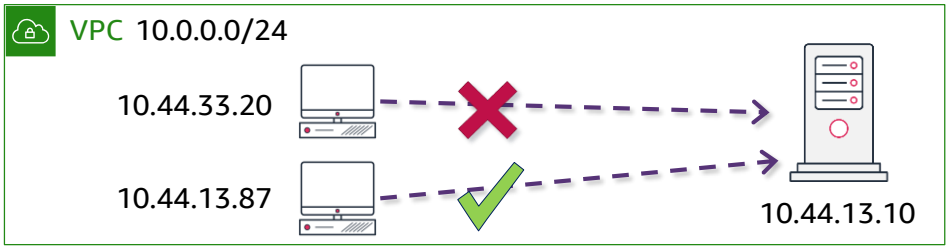
Amazon VPC enables you to provision an isolated section of the AWS Cloud. In this isolated section, you can launch resources in a virtual network that you define. Within a virtual private cloud (VPC), you can organize your resources into subnets. A **subnet** is a section of a VPC that can contain resources such as Amazon EC2 instances. In the coffee shop, you can think of the counter area as a VPC. The counter area divides into two separate areas for the cashier's workstation and the barista's workstation.
****For Accessibility:** This image depicts how the coffee shop maps to a virtual network. The counter area is related to the idea of a Virtual Private Network. **End Description****



The coffee shop has become overwhelmed with customers trying to place orders. Not only that, there are even customers trying to get into places they are not allowed.
When you consider a Cloud network, the coffee shop is the VPC as you just learned. A VPC requires a way to identify each of the connections that are wanting access. In our example, it would be similar to requiring a membership card to come in.
****For Accessibility:** This image depicts the new coffee shop layout that separates the customers from the barista to ensure that all orders go to the cashier. **End Description****

IP Addresses

- Is assigned to a Network Interface Card (NIC)
- Uniquely identify devices/resources
- IP addresses must be unique
- Communication requires devices/resources to be in the same CIDR range



aws

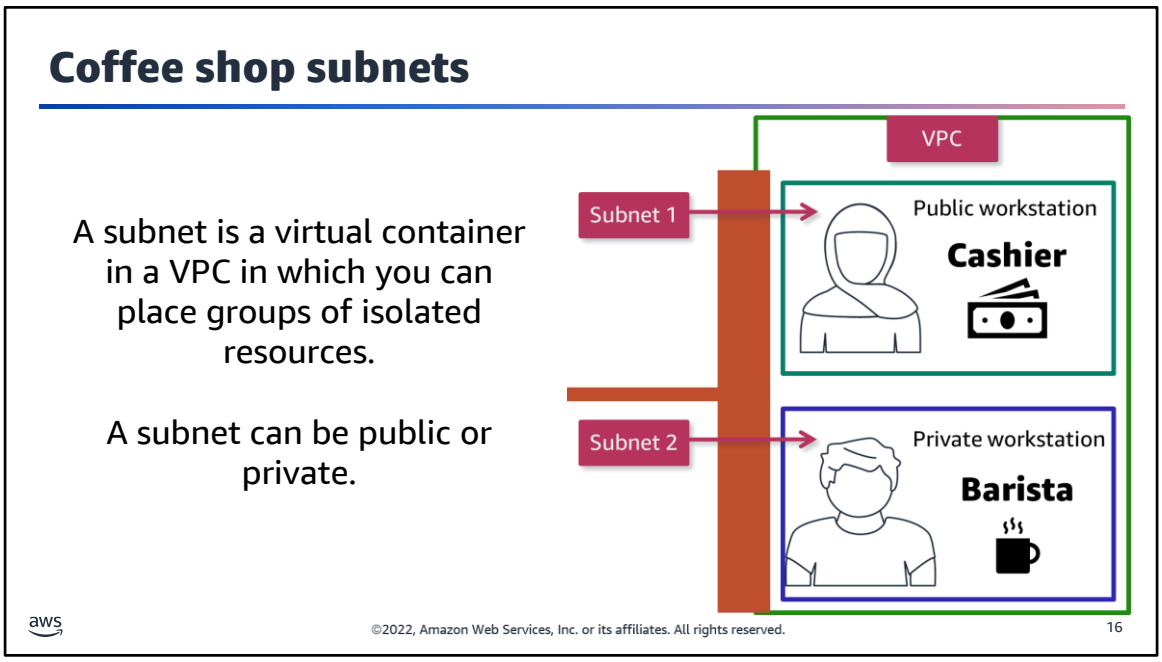
15

IP Addresses (IPs) are assigned to the *Network Interface Card (NIC)* on a device. This is true for services and instances in the Cloud as well. This interface is what allows the service or instance to communicate with other services or instances. IPs are assigned to devices like laptops, printers, database services, file servers, and Amazon EC2 instances. In order fore these entities to communicate on a network, they must have an IP that is unique on that network. Think of it as your house number. There cannot be two houses on the same street with the same house number.

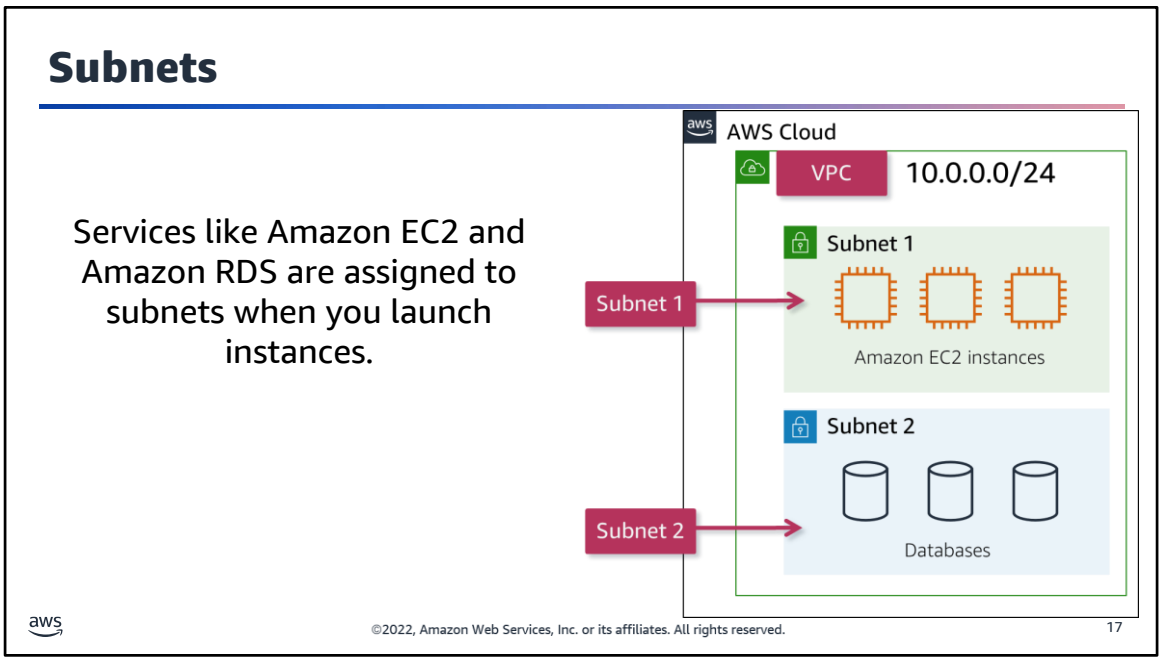
VPCs use IP Addresses to identify the entity trying to communicate with it. In order for a device to communicate, its IP must be in the same range as the device they are trying to communicate with. The range is called a *CIDR range*. When you create a VPC you must define the CIDR range.

In our example, there is a VPC that has a CIDR range of 10.0.0.0/24. The /24 indicates that the first three numbers in the IP must be the same to communicate with each other. The Cloud instance in this example has an IP of 10.44.13.10. A device would like to access the server. The device has an IP of 10.44.33.20. Since the first three numbers do not match, the connection is denied. A second device attempts to connect to the server. This device has an IP of 10.44.13.87. The first three numbers match this time so the connection is allowed.

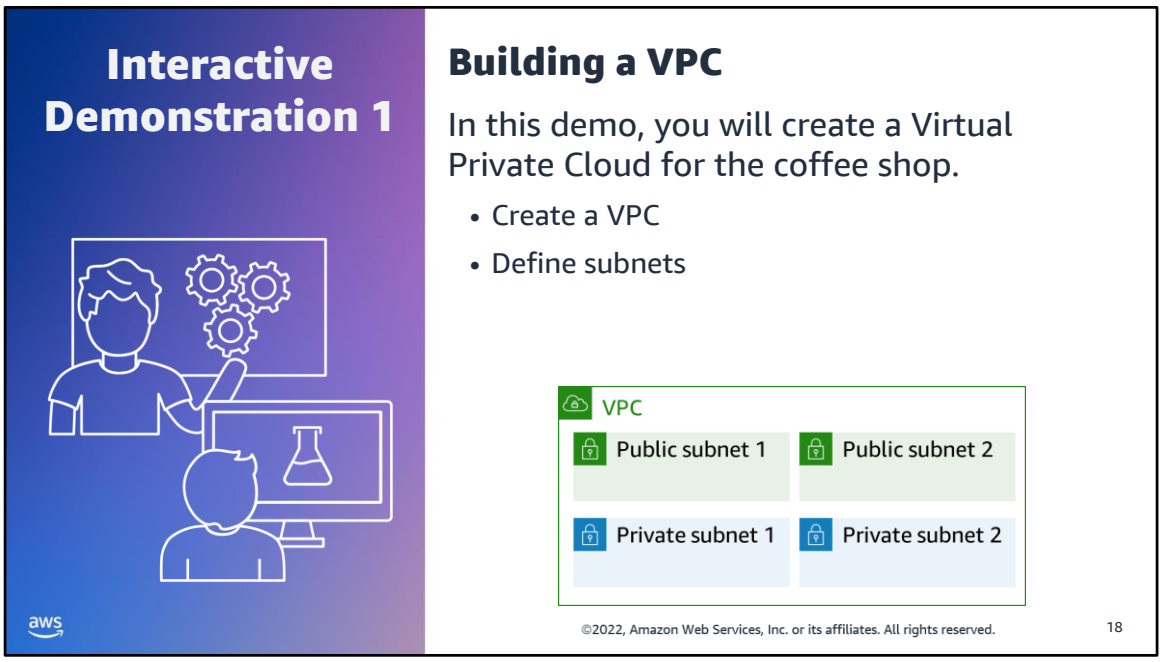
This is a very simplified example but is a good place to start your understanding of how IP addresses and VPC work to control traffic on the network.



In our coffee shop example, the cashier and barista can be separated from each other using subnets. You can designate the workstation with the cashier as public and the workstation with the barista as private. This helps you to limit the customers to only give orders to the cashier.

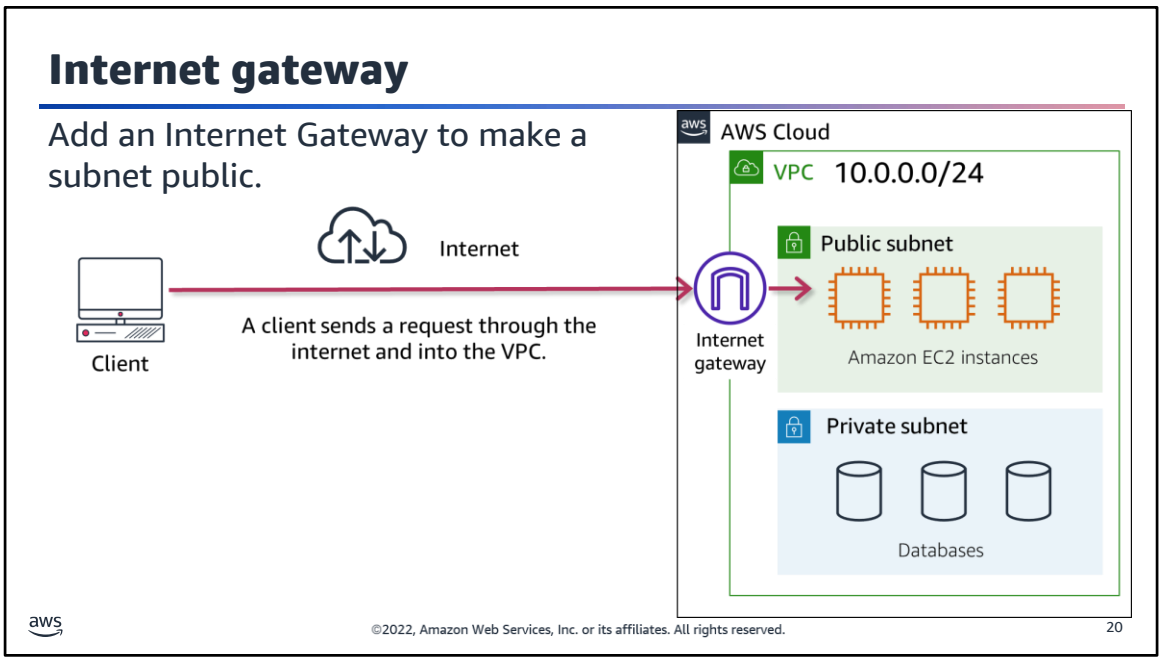


Within a VPC, the workspace in the coffee shop, subnets can communicate with each other. Just like the barista and the cashier can talk freely. The same is true with within the VPC. For example, you might have an application that uses Amazon EC2 instances in a public subnet. When the applications running on the instances need data, they can communicate with databases that are located in a private subnet.
****For Accessibility:** The VPC contains a public subnet with Amazon EC2 instances and a private subnet with databases. **End Description****



Please pay attention to the instructor as you complete this interactive demonstration.
The Lab directions are found in the Lab Guide.

In this topic, you will learn about the additional components within a VPC and how to connect a subnet to the internet.

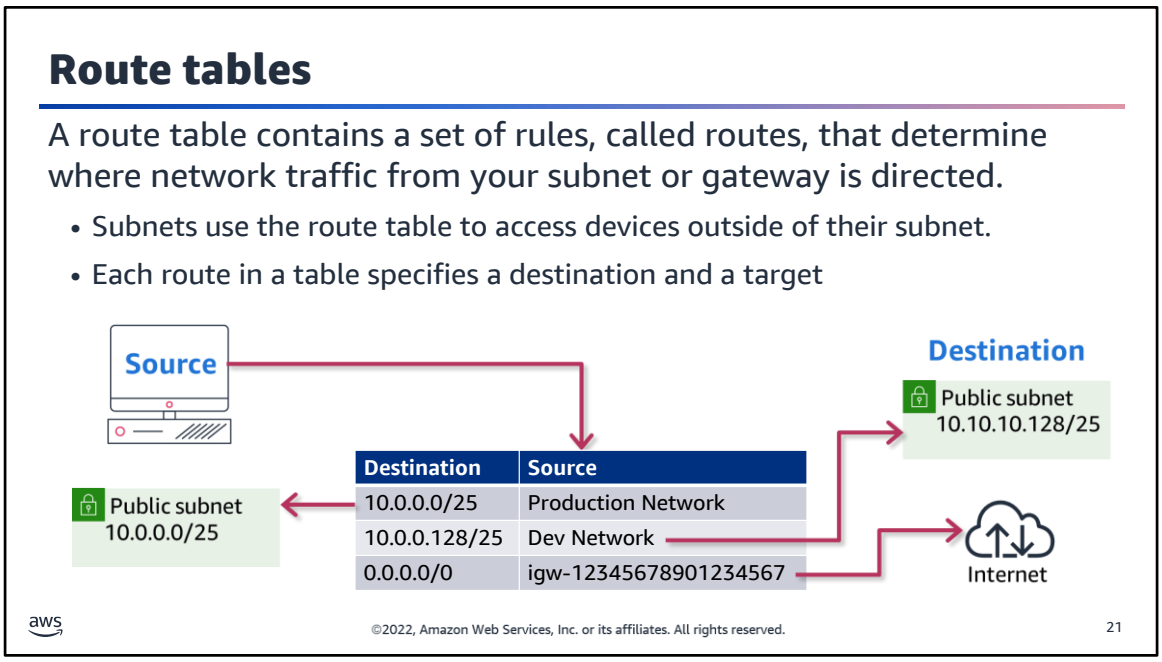


So let's talk about the technology behind subnets.

- *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.

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.

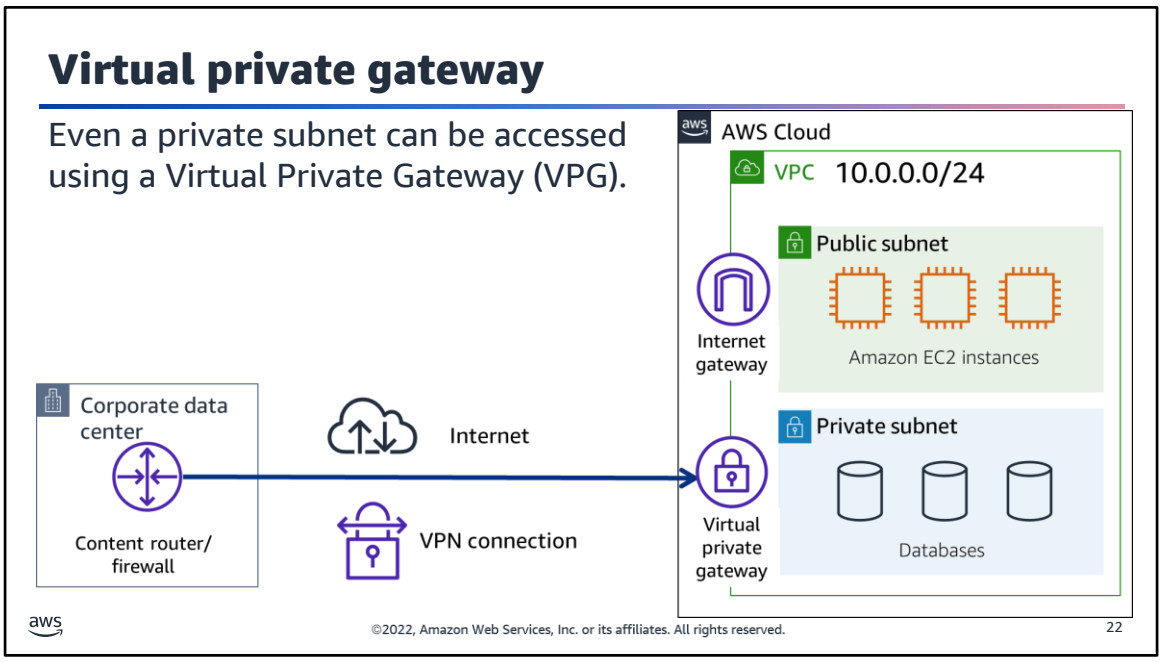


A *route table* contains a set of rules, called *routes*, that determine where network traffic from your subnet or gateway is directed.

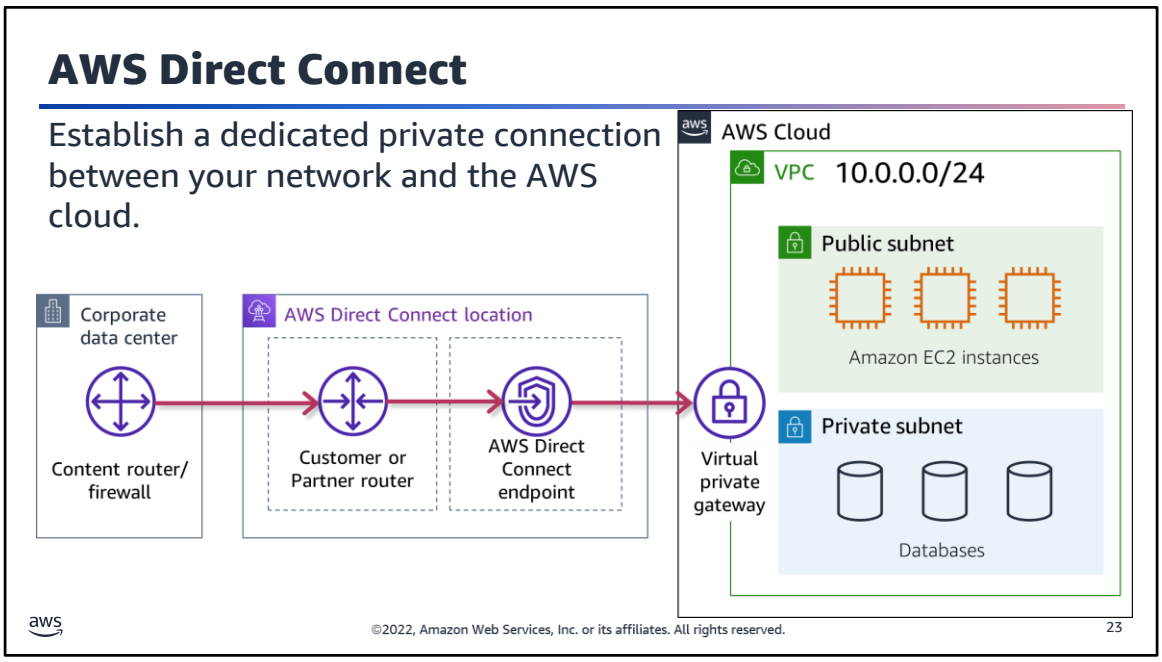
Your VPC has an implicit router, and you use route tables to control where network traffic is directed. Each subnet in your VPC must be associated with a route table, which controls the routing for the subnet (subnet route table). A subnet can only be associated with one route table at a time, but you can associate multiple subnets with the same subnet route table.

Each *route* in a table specifies a destination and a target. For example, to enable your subnet to access the internet through an internet gateway, add the following route to your subnet route table.

When you create a VPC, it automatically has a main route table. When a subnet does not have an explicit routing table associated with it, the main routing table is used by default.



To access private resources in a VPC, you can use a **virtual private gateway**.
Here's an example of how a virtual private gateway works. You can think of the internet as the road between your home and the coffee shop. Suppose that you are traveling on this road with a bodyguard to protect you. You are still using the same road as other customers, but with an extra layer of protection.
The bodyguard is like a VPN connection that encrypts (or protects) your internet traffic from all the other requests around it.
The virtual private gateway is the component that allows protected internet traffic to enter into the VPC.
Even though your connection to the coffee shop has extra protection, traffic jams are possible because you're using the same road as other customers.
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.
Another option that you can use to get from your private network to the VPC is AWS Direct Connect.

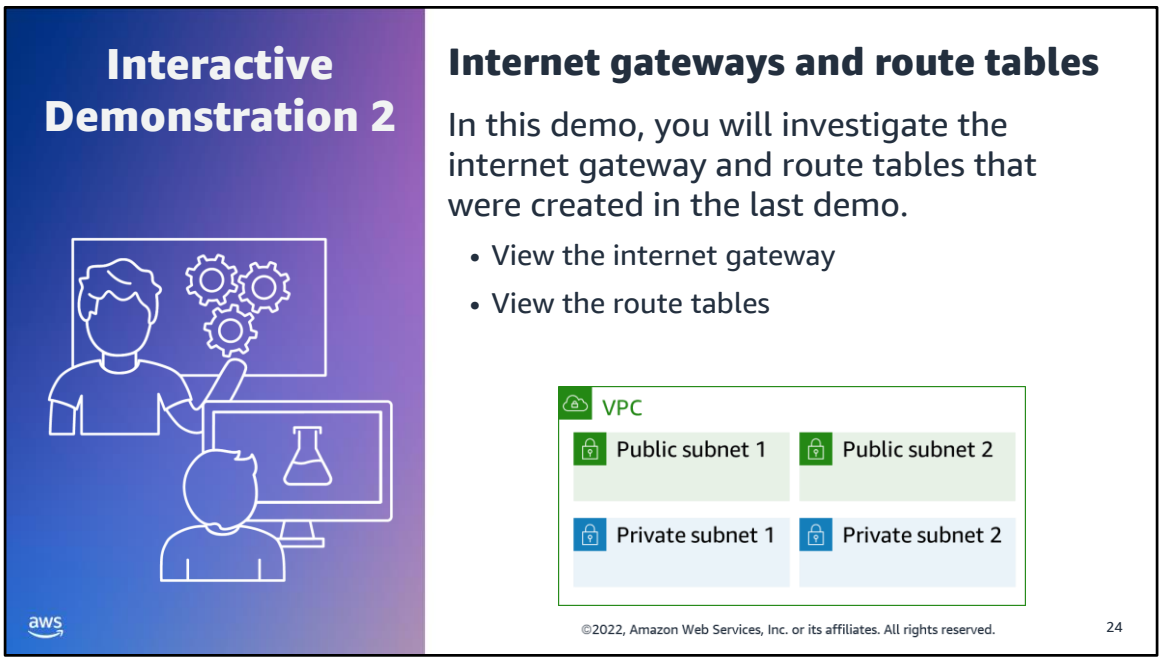


AWS Direct Connect is a service that enables you to establish a dedicated private connection between your data center and VPC.

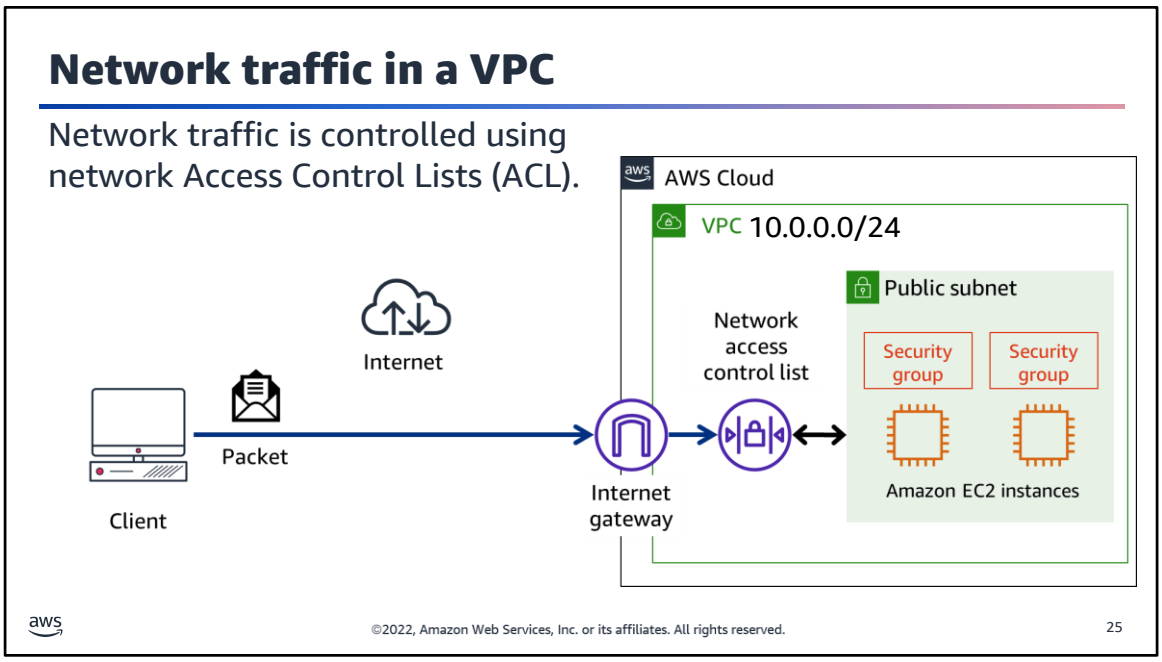
Suppose that there is an apartment building with a hallway directly linking the building to the coffee shop. Only the residents of the apartment building are allowed to travel through this hallway.

This private hallway provides the same type of dedicated connection as AWS Direct Connect. Residents are able to get into the coffee shop without needing to use the public road that is shared with other customers.

The private connection that AWS Direct Connect provides helps you to reduce network costs and increase the amount of bandwidth that can travel through your network.



Please pay attention to the instructor as you complete this interactive demonstration.
The Lab directions are found in the Lab Guide.



This section explores how resources in a VPC can communicate through the internet.
When a customer requests data from an application that is 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 within a subnet.
The VPC component that checks packet permissions for subnets is a **network access control list (NACL)**.

Network access control lists

A network access control list (network ACL) is a virtual firewall for a subnet.

- A network ACL can be used by multiple subnets.
- Network ACLs use rules

Rule #	Type	Protocol	Port range	Source	Allow/Deny
100	All IPv4 traffic	SSH	All	0.0.0.0/0	Allow
*	All IPv4 traffic	All	All	0.0.0.0/0	Deny

aws

26

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

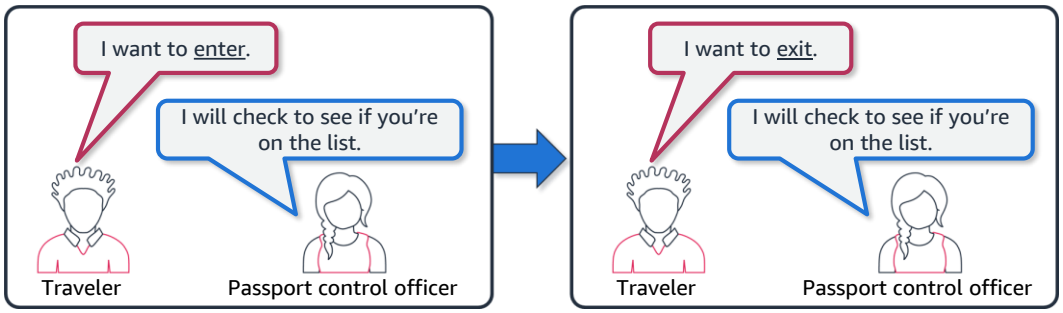
For this example, step outside of the coffee shop and imagine that you are in an airport. In the airport, travelers are trying to enter into a different country. You can think of the travelers as packets and the passport control officer as a network ACL. The passport control officer checks travelers' 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.

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 should be allowed. 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.
- Before a packet can exit a subnet, it must be checked against the outbound rules.



aws

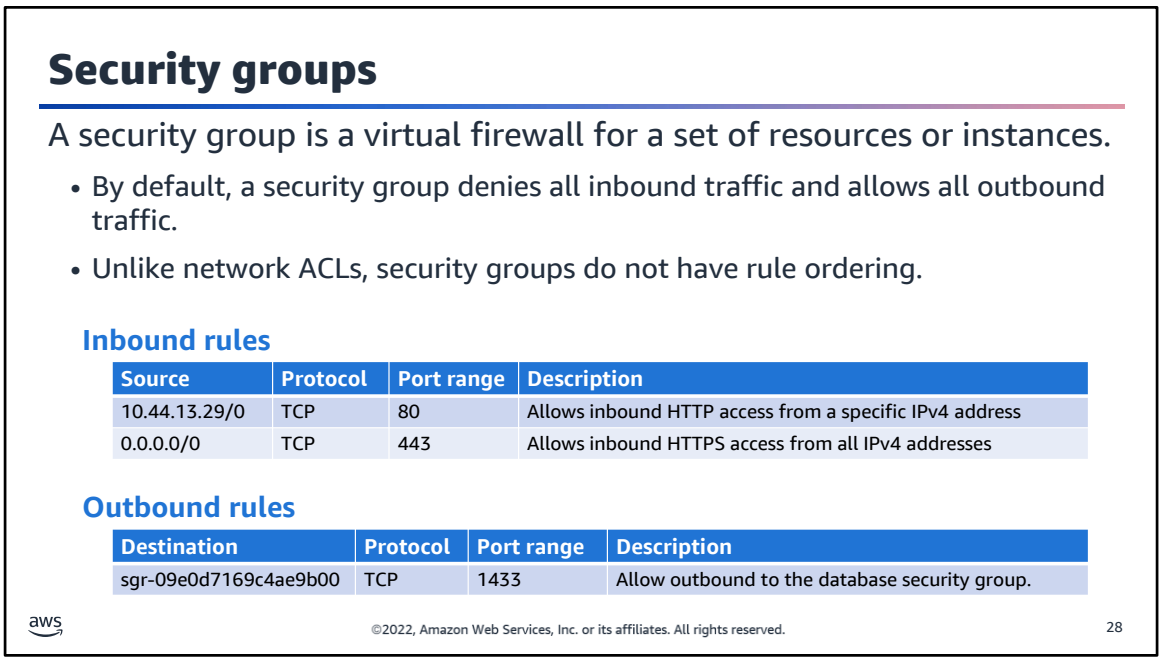
27

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

Recall the previous example of a traveler who wants to enter into a different country. This is similar to sending a request out from an Amazon EC2 instance and to the internet.

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 it is allowed or denied.

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**.



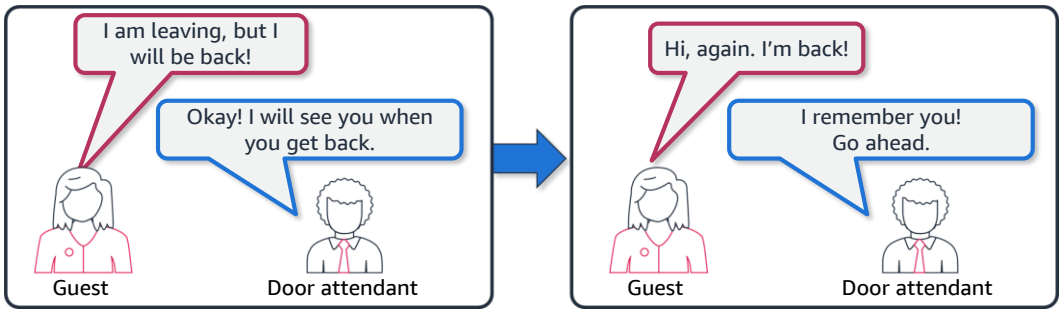
A **security group** is a virtual firewall that controls inbound and outbound traffic for an Amazon EC2 instance. By default, a security group denies all inbound traffic and allows all outbound traffic. You can add custom rules to configure which traffic should be allowed or denied.

For this example, suppose that you are in an apartment building with a door attendant who greets guests in the lobby. You can think of the guests as packets and the door attendant as a security group. As guests arrive, the door attendant checks a list to ensure they can enter the building. However, the door attendant does not check the list again when guests are exiting the building.

If you have multiple Amazon EC2 instances within a subnet, 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 that were made for incoming packets.

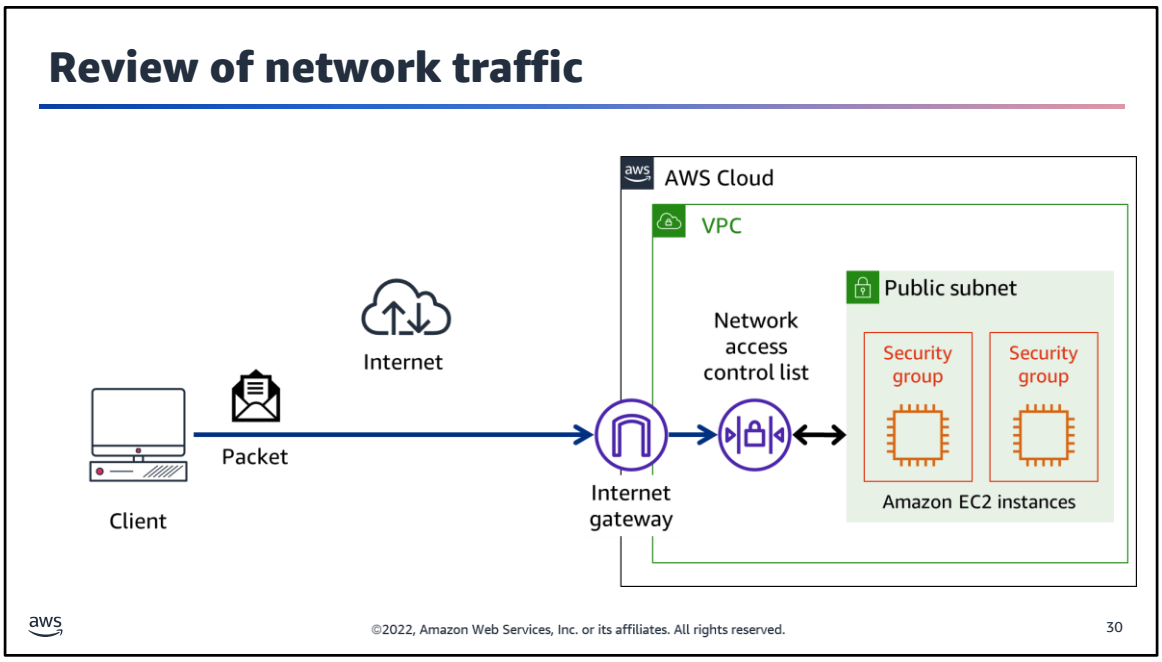


aws

29

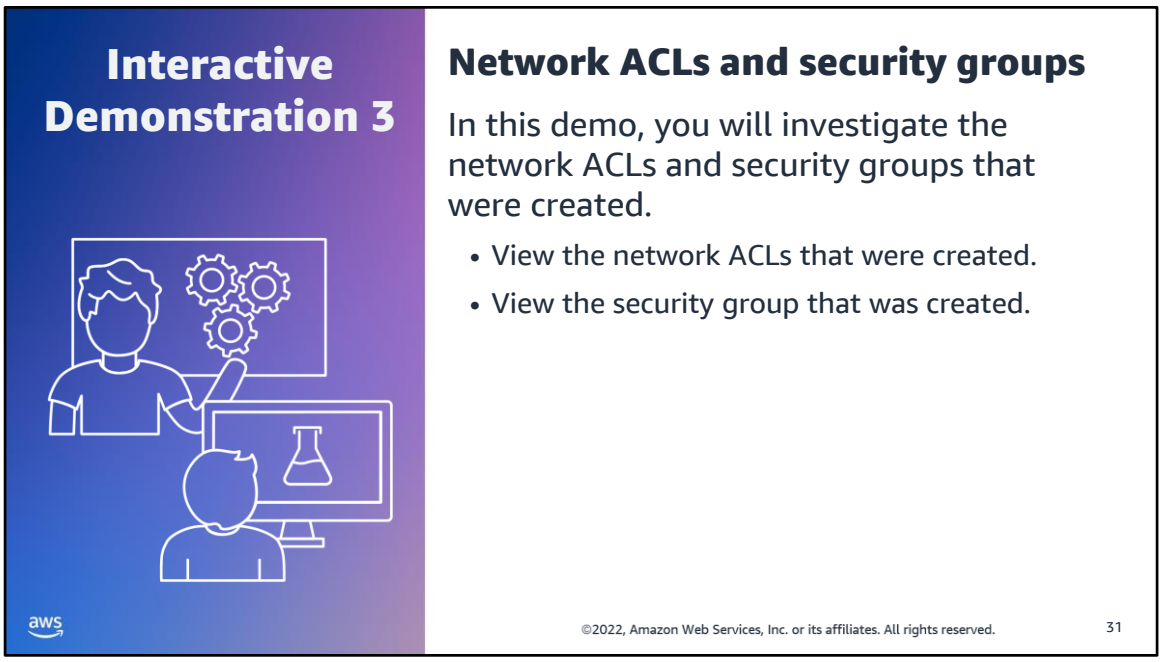
Security groups perform **stateful** packet filtering. They remember previous decisions that were made for incoming packets.

Consider the same example of sending a request out from an Amazon EC2 instance to the internet. When a packet response for that request comes back to the instance, the security group remembers your previous request and allows the response to proceed, regardless of inbound security group rules.

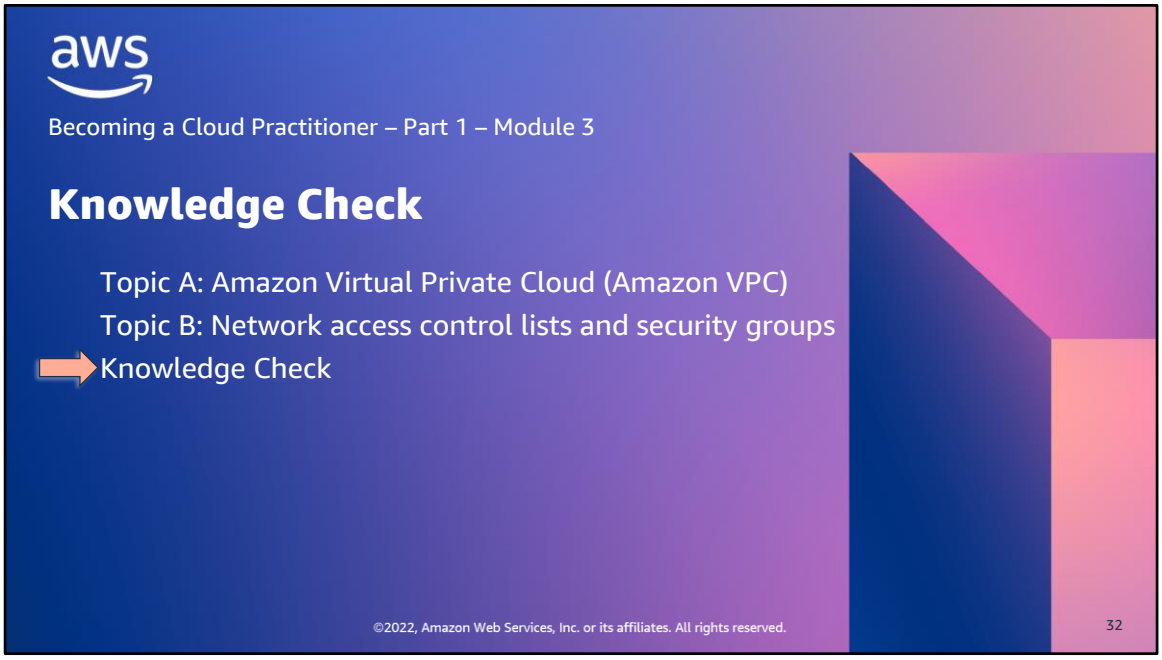


You learned how packets travel into a virtual private cloud (VPC). A packet is a request from the internet. Packets come into a VPC through an internet gateway. Before a packet can enter into a subnet, it is evaluated against rules in a network ACL. Then if the packet wants to access an Amazon EC2 instance, its permissions are evaluated by a security group.

Both network ACLs and security groups enable you to configure custom rules for the traffic in your VPC. As you continue to learn more about AWS security and networking, make sure to understand the differences between network ACLs and security groups.



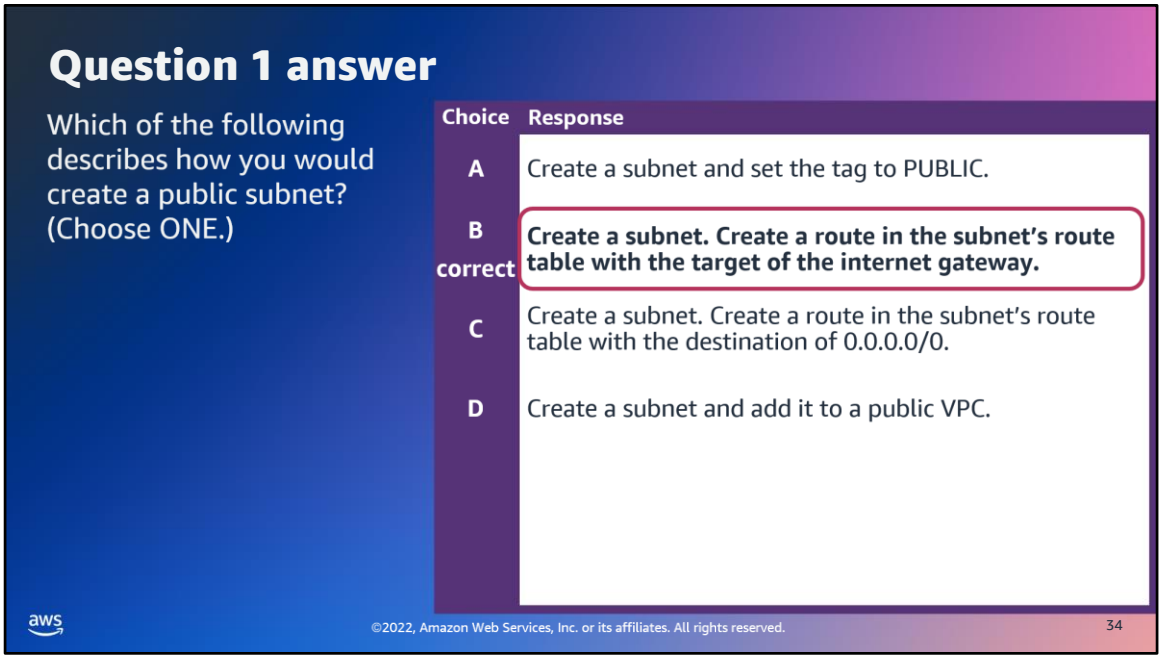
Please pay attention to the instructor as you complete this interactive demonstration.
The Lab directions are found in the Lab Guide.



Question 1

Which of the following describes how you would create a public subnet? (Choose ONE.)

Choice	Response
A	Create a subnet and set the tag to PUBLIC.
B	Create a subnet. Create a route in the subnet's route table with the target of the internet gateway.
C	Create a subnet. Create a route in the subnet's route table with the destination of 0.0.0.0/0.
D	Create a subnet and add it to a public VPC.



The correct answer is **Create a subnet. Create a route in the subnet's route table with the target of the internet gateway.**

- Creating a public subnet is done by adding a route to an Internet Gateway.

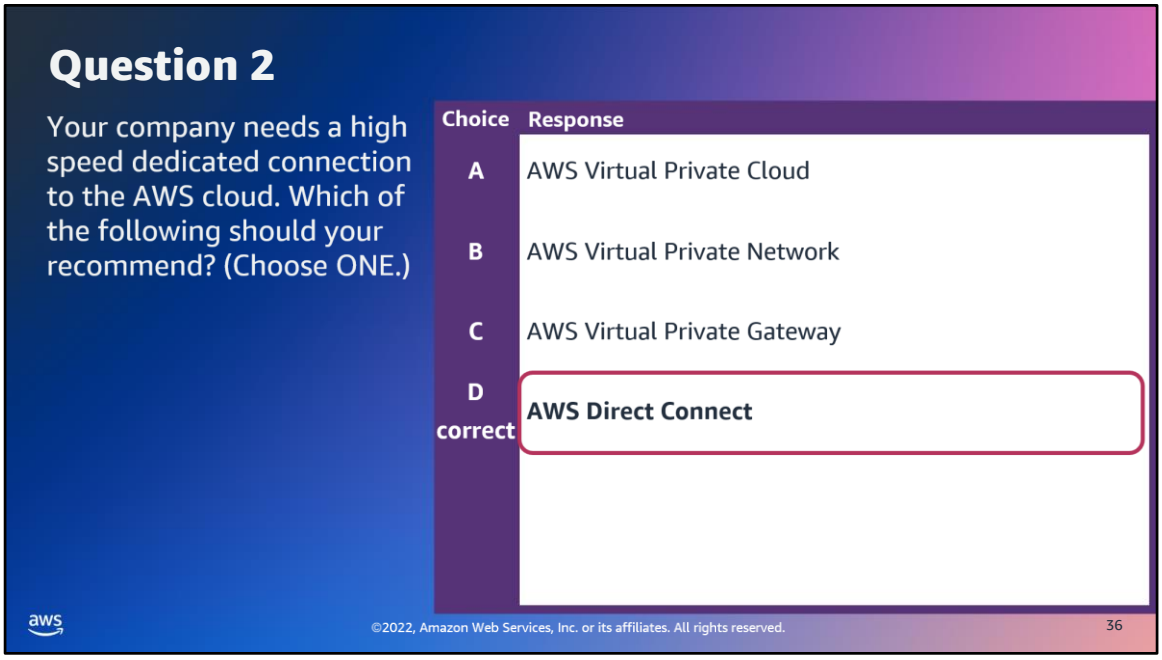
The other response options are incorrect because:

- All subnets are private unless a route to an internet gateway is added to its route table.
- There is no setting or tag that makes a subnet public.
- Adding a route with a destination of 0.00.0/0 is not a complete rule. A rule must contain both a destination and a target.

Question 2

Your company needs a high speed dedicated connection to the AWS cloud. Which of the following should your recommend? (Choose ONE.)

Choice	Response
A	AWS Virtual Private Cloud
B	AWS Virtual Private Network
C	AWS Virtual Private Gateway
D	AWS Direct Connect



The correct answer is **AWS Direct Connect**.

- AWS Direct Connect is a service that enables you to establish a dedicated private connection between your data center and VPC.

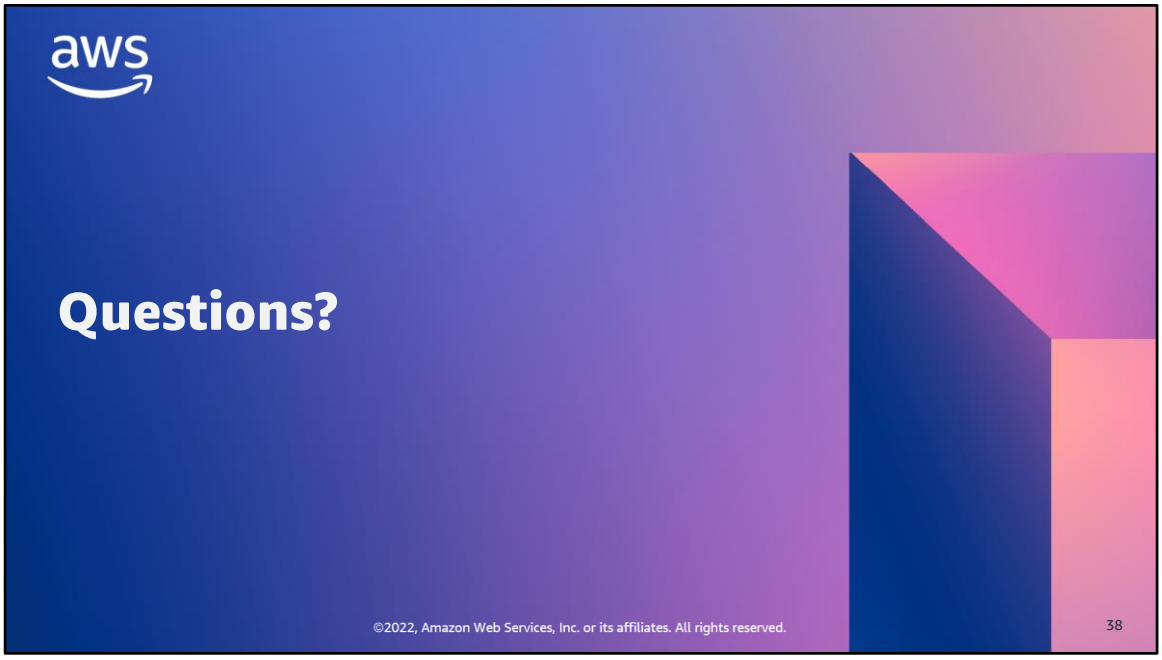
The other response options are incorrect because:

- AWS Virtual Private Cloud is the service that creates private clouds within the AWS Cloud. It is not a dedicated connection.
- AWS Virtual Private Networks create a private connection but the connection is made over the internet (public network).
- AWS Virtual Private Gateway is a service that connects VPCs and an existing Direct Connect connection to each other.

Module 3 summary

In this module, you learned how to:

- Describe basic networking concepts
- Describe the differences between public and private networking resources

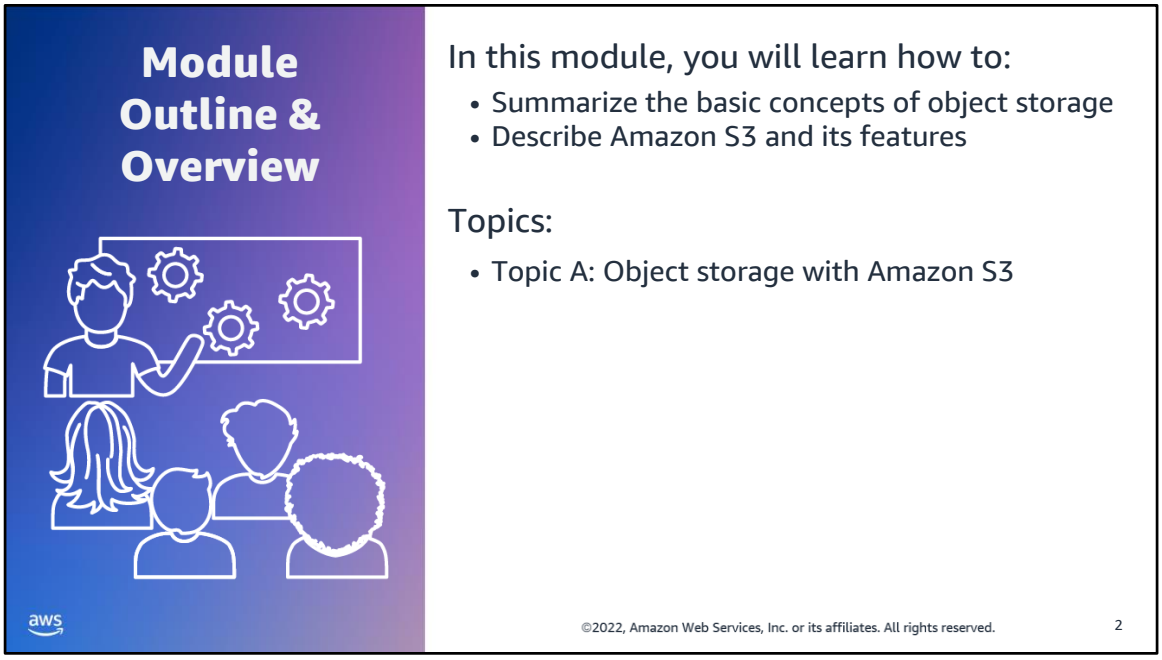


In this module, you will learn about Amazon S3 and the features this service provides.
Module Objectives:

- Summarize the basic concepts of object storage
- Describer Amazon S3 and its features

Topics:

- Topic A: Object storage with Amazon S3

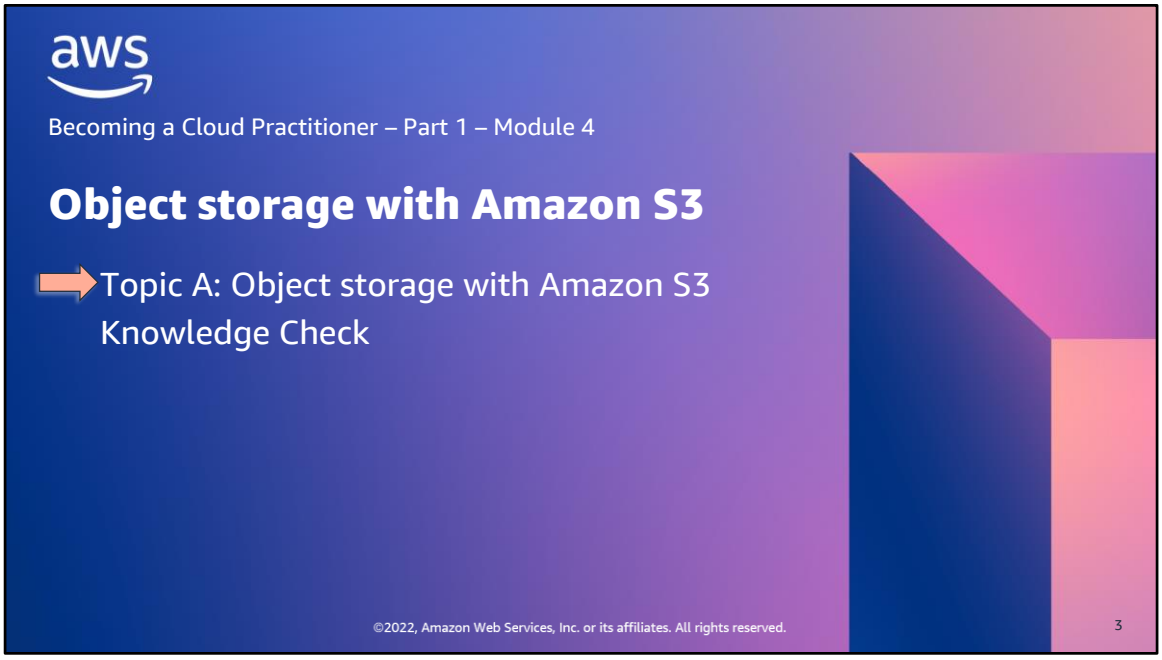


In this module, you will learn about Amazon S3 and the features this service provides.
Module Objectives:

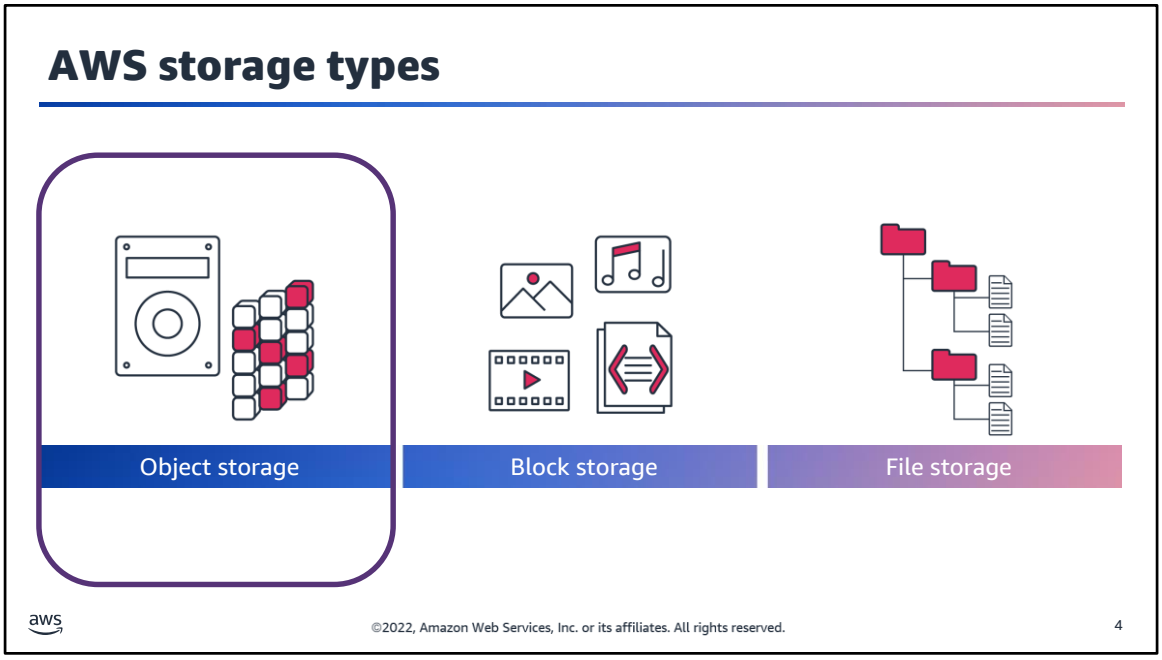
- Summarize the basic concepts of object storage
- Describer Amazon S3 and its features

Topics:

- Topic A: Object storage with Amazon S3



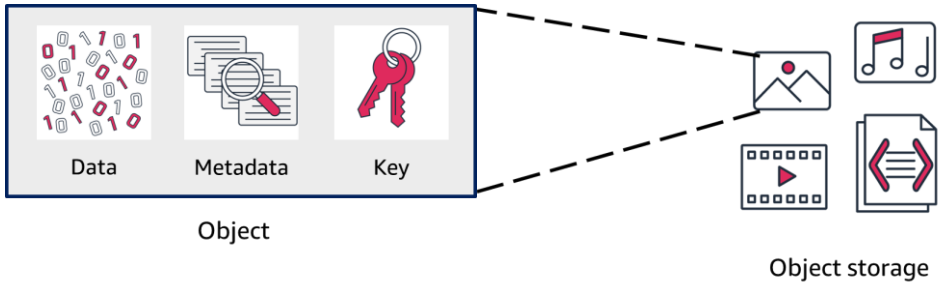
In this topic, you will learn about object storage with Amazon S3.



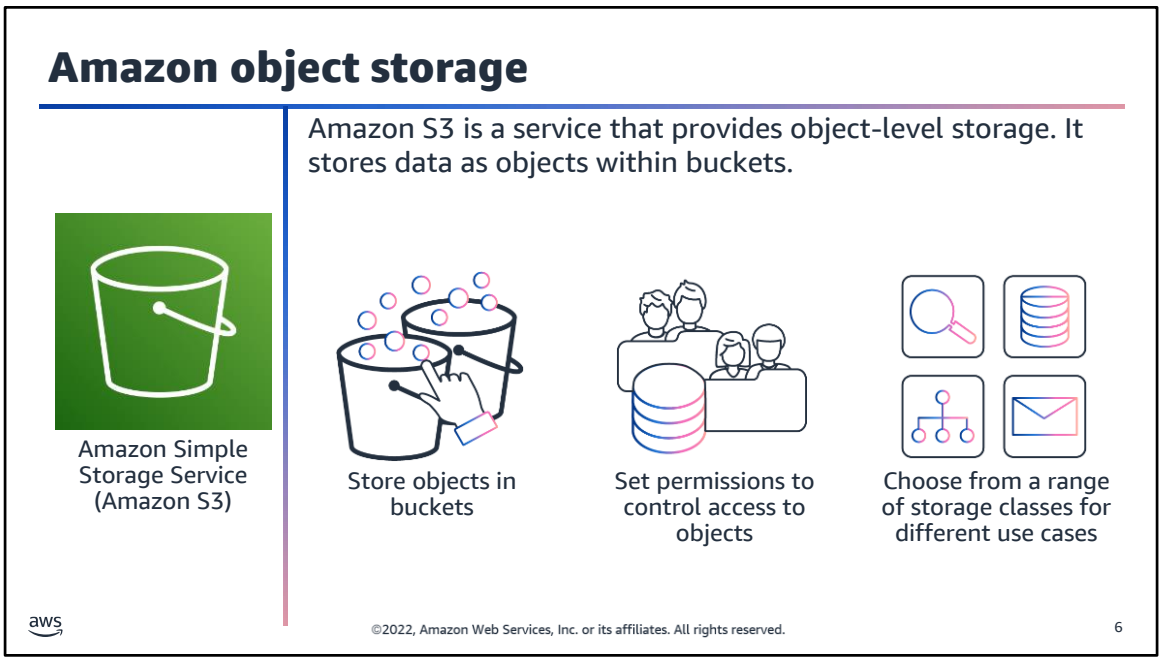
This module explores three types of storage: object storage, block storage, and file storage.
You will learn about AWS services that offer each type of storage and discuss how these services are used.
To begin, you will learn about block storage.

Object storage

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



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 contextual information about what the data is, what it should be used for, the object size, and so on. An object's key is its unique identifier. For example, an object key name might be "4my-organization" or "my.great_photos-2014/jan/myvacation.jpg."
Recall that 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.
To learn more about object storage, you need to examine **Amazon Simple Storage Service (Amazon S3)**.



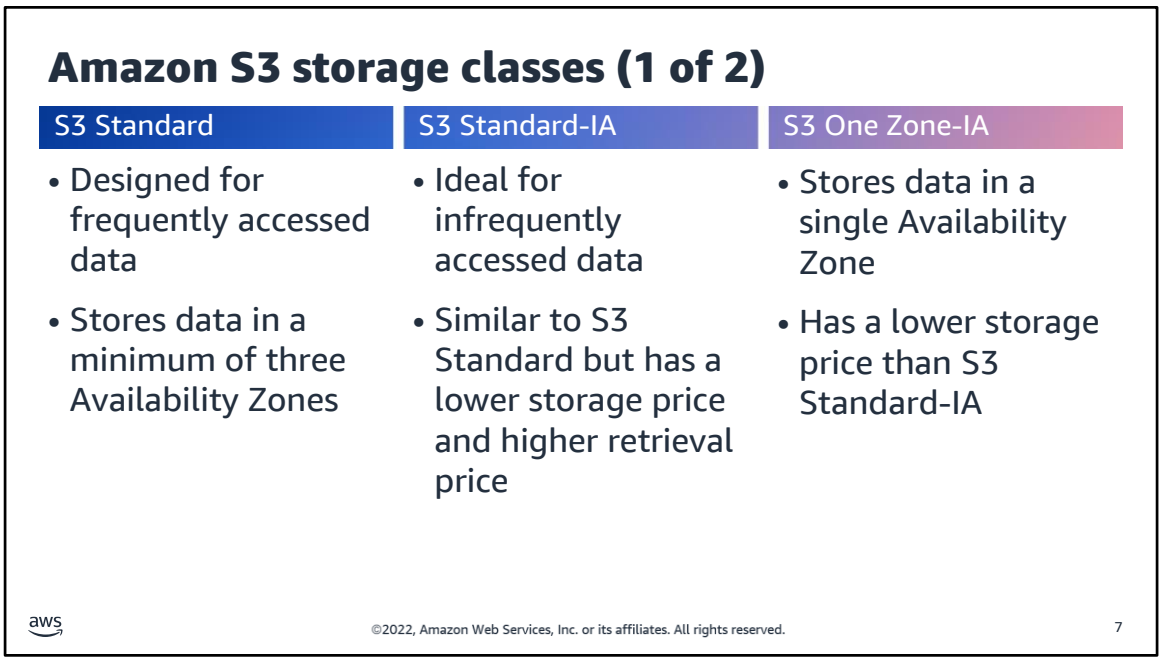
Amazon Simple Storage Service (Amazon S3) is a service that provides object-level storage. Amazon S3 stores data as objects within buckets.

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 Amazon S3 versioning feature to track changes to your objects over time.

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. 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



The **S3 Standard** storage class provides high availability for objects. This makes it a good choice for a wide range of use cases, such as websites, content distribution, and data analytics. S3 Standard has a higher cost than other storage classes intended for infrequently accessed data and archival storage.

Suppose that some of the objects you are storing in Amazon S3 are backup files and accessed rarely. However, you still want to ensure that when you need the backup files, you can access them quickly.

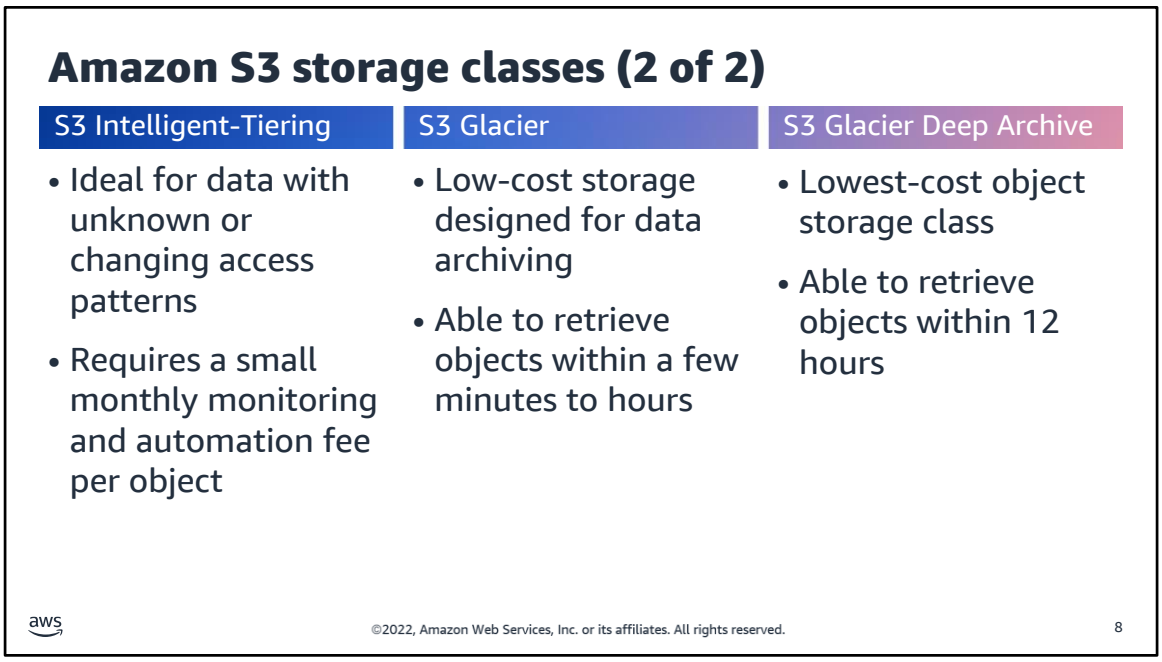
For this type of scenario, **S3 Standard-Infrequent Access (S3 Standard-IA)** might be a good fit. S3 Standard-IA is ideal for data infrequently accessed but requires high availability when needed. Both S3 Standard and S3 Standard-IA store data in a minimum of three Availability Zones. S3 Standard-IA provides the same level of availability as S3 Standard but with a lower storage price and a higher retrieval price.

Compared to S3 Standard, S3 Standard-IA has a lower per-TB storage price. Here's an example of what this means: Suppose that you have a few objects that you access only once a month. If you stored these objects in the S3 Standard storage class, you would pay a higher storage price for objects that are suited for S3 Standard-IA.

Compared to S3 Standard, S3 Standard-IA has a higher per-GB retrieval price. Suppose that you have a few objects that you need to frequently access throughout the day. S3 Standard-IA would most likely not be the optimal choice for these objects. You would pay a higher per-GB retrieval price as data is frequently accessed throughout the day. In this example, S3 Standard would be more cost-efficient.

Another storage class to consider for infrequently accessed data is **S3 One Zone-Infrequent Access (S3 One Zone-IA)**. S3 One Zone-IA is ideal for infrequently accessed data that does not require high availability. Compared to S3 Standard and S3 Standard-IA, which store data in a minimum of three Availability Zones, S3 One Zone-IA stores data in a single Availability Zone. This makes it a good storage class to consider if the following conditions apply:

- You want to save costs on storage
- You do not require the higher availability that S3 Standard and S3 Standard-IA offer



The scenarios discussed involved objects with access patterns that are known and consistent. If you have data with access patterns that are unknown or changing, you might consider storing this data in the **S3 Intelligent-Tiering** storage class.

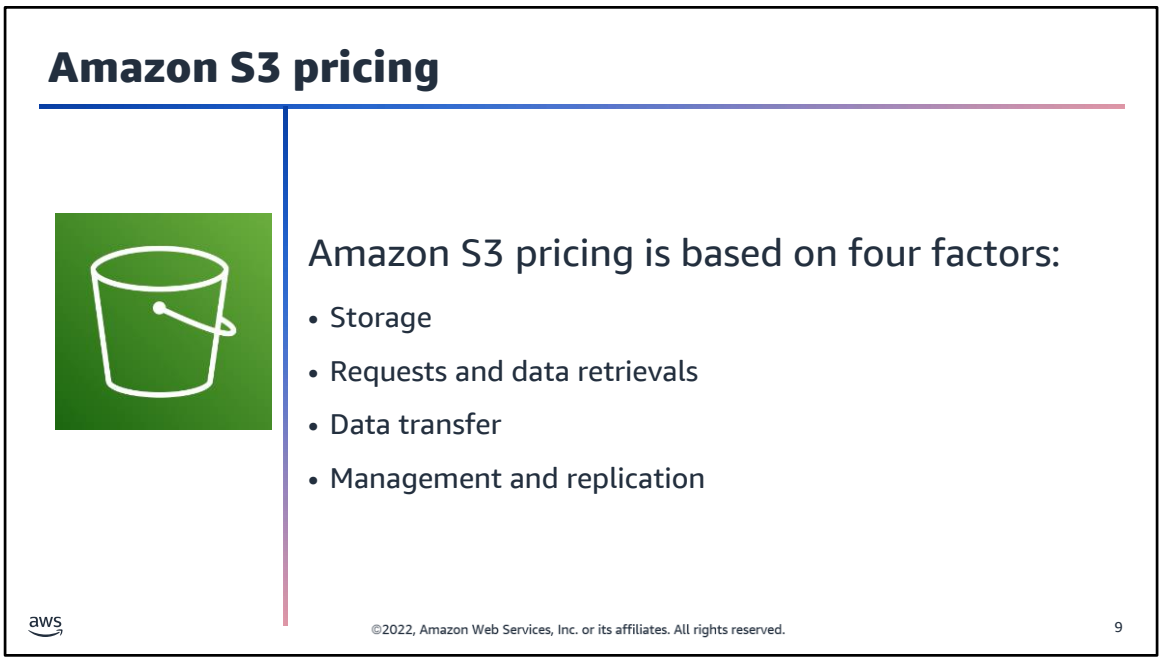
In the S3 Intelligent-Tiering storage class, Amazon S3 monitors objects' access patterns. If you haven't accessed an object for 30 consecutive days, Amazon S3 automatically moves it to the infrequent access tier, S3 Standard-IA. If you access an object in the infrequent access tier, Amazon S3 automatically moves it to the frequent access tier, S3 Standard.

With S3 Intelligent-Tiering, there is a small monthly monitoring and automation fee per object. However, a key benefit of the S3 Intelligent-Tiering storage class is the ability to save time by not having to manually monitor your objects' access patterns. You also do not have to manually move them to different storage classes.

The final Amazon S3 storage classes are **Amazon S3 Glacier** and **S3 Glacier Deep Archive**. These are low-cost storage classes that are ideal for data archiving. For example, you might use these storage classes to store archived customer records or earlier photos and video files.

When deciding between Amazon S3 Glacier and Amazon S3 Glacier Deep Archive, consider how quickly you must retrieve archived objects. You can retrieve objects stored in the Amazon S3 Glacier storage class within a few minutes to a few hours. By comparison, you can retrieve objects stored in the S3 Glacier Deep Archive storage class within 12 hours.

Next, you will review the Amazon S3 storage classes.



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 length of storage for 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 the 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 in the same AWS Region. However, you pay for data that you transfer into and out of Amazon S3, with a few exceptions. Data transferred into Amazon S3 from the internet or out to Amazon CloudFront incurs no cost. In addition, data transferred out to an Amazon EC2 instance in the same AWS Region as the Amazon S3 bucket incurs no cost.
- **Management and replication:** You pay for the storage management features that you enabled on your account's Amazon S3 buckets. These features include Amazon S3 inventory, analytics, and object tagging.

Example: Amazon S3 service charges

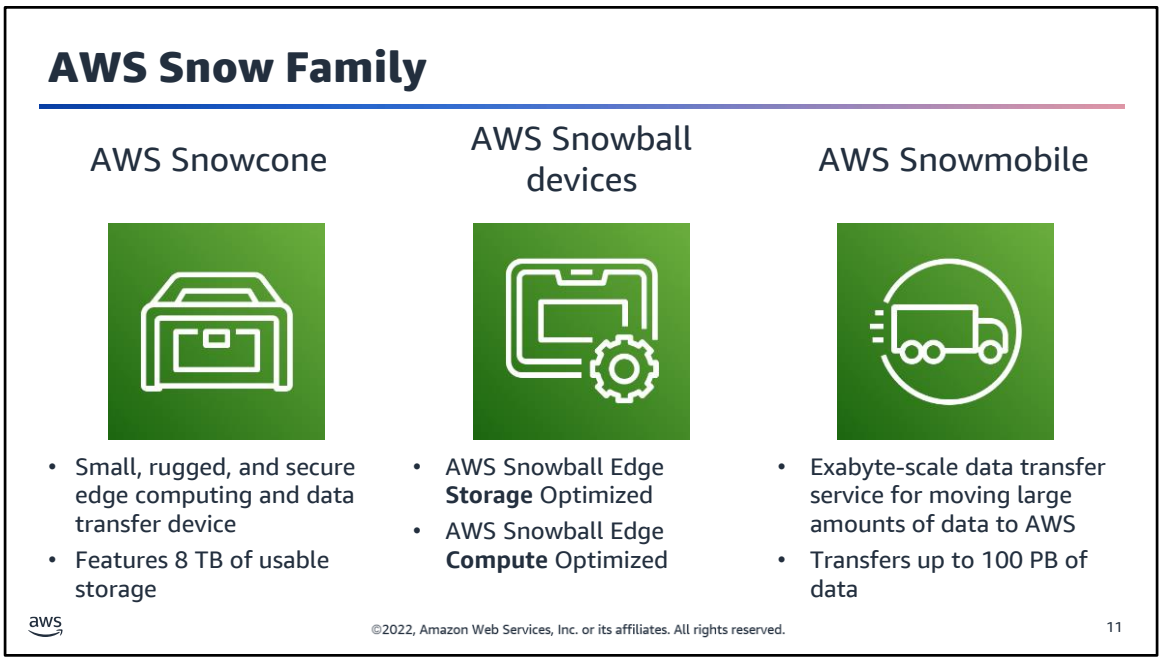
Simple Storage Service		**$0.00**
US East (N. Virginia)		**$0.00**
Amazon Simple Storage Service Requests-Tier1		$0.00
$0.00 per request - PUT, COPY, POST, or LIST requests under the monthly global free tier	185.000 Requests	$0.00
Amazon Simple Storage Service Requests-Tier2		$0.00
$0.00 per request - GET and all other requests under the monthly global free tier	923.000 Requests	$0.00
Amazon Simple Storage Service TimedStorage-ByteHrs		$0.00
$0.000 per GB - storage under the monthly global free tier	0.159 GB-Mo	$0.00
US East (Ohio)		**$0.00**
Amazon Simple Storage Service USE2-Requests-Tier2		$0.00
$0.00 per request - GET and all other requests under the monthly global free tier	4.000 Requests	$0.00
Amazon Simple Storage Service USE2-TimedStorage-ByteHrs		$0.00
$0.000 per GB - storage under the monthly global free tier	0.000001 GB-Mo	$0.00

aws

The AWS account in this example uses Amazon S3 in two Regions: Northern Virginia and Ohio. For each Region, itemized charges are based on the following factors:

- Number of requests to add or copy objects into a bucket
- Number of requests to retrieve objects from a bucket
- Amount of storage space used

All the usage for Amazon S3 in this example is under the AWS Free Tier limits, so the account owner does not have to pay for any Amazon S3 usage in this month.



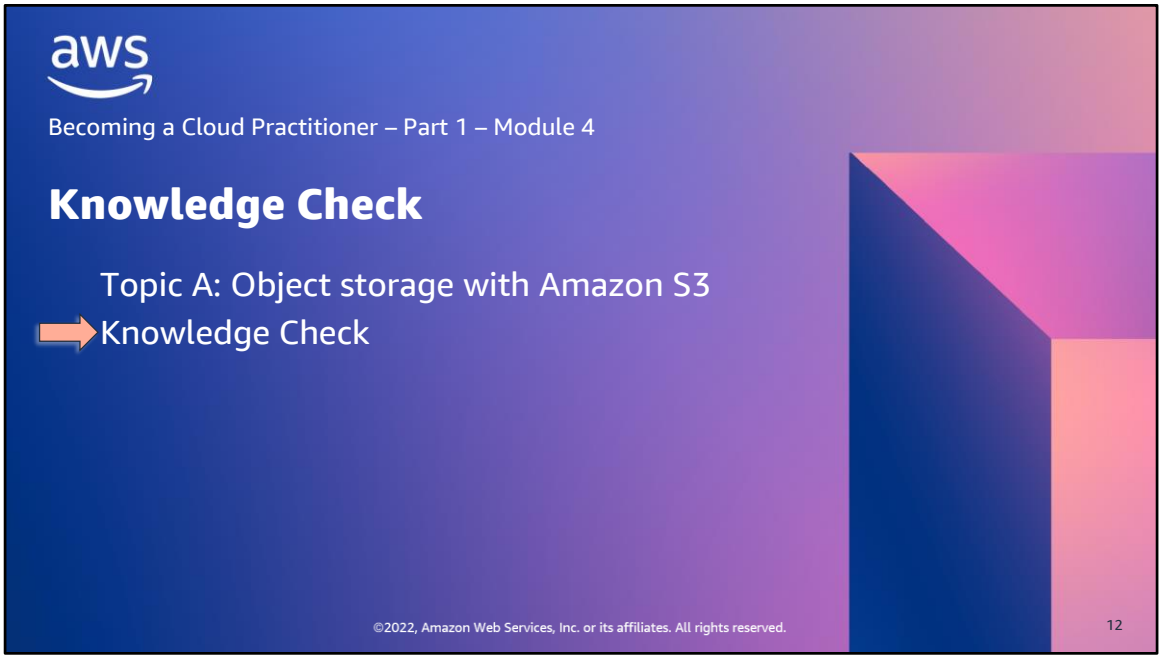
The AWS Snow Family is composed of **AWS Snowcone**, **AWS Snowball**, and **AWS Snowmobile**. These devices offer different capacity points, and most include built-in computing capabilities. AWS owns and manages the Snow Family devices and integrates with AWS security, monitoring, storage management, and computing capabilities.

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 offers two types of **AWS Snowball** 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; 1 TB of SATA solid state drive (SSD) for block volumes
    - Compute: 40 vCPUs; 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; 7.68 TB of usable NVMe SSD capacity for Amazon EBS compatible block volumes
    - Compute: 52 vCPUs; 208 GiB of memory; optional NVIDIA Tesla V100 GPU; 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, which is a 45-foot long ruggedized shipping container, pulled by a semitrailer truck.

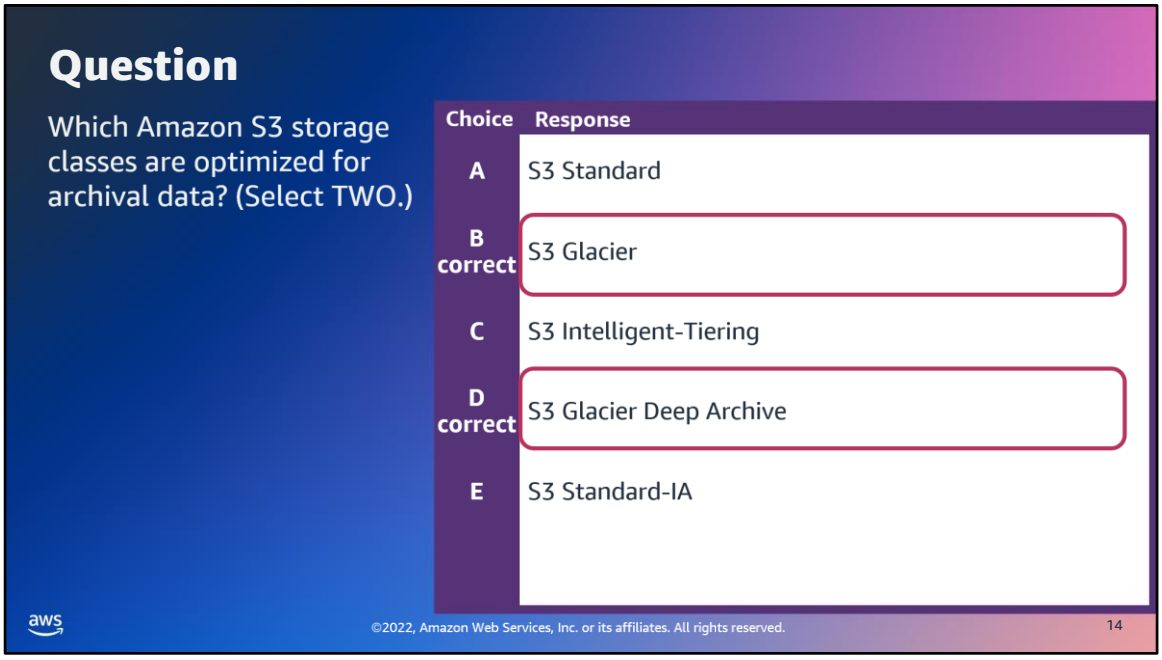


In this section, you will learn about the additional components within a VPC and how to connect a subnet to the internet.

Question

Which Amazon S3 storage classes are optimized for archival data? (Select TWO.)

Choice	Response
A	S3 Standard
B	S3 Glacier
C	S3 Intelligent-Tiering
D	S3 Glacier Deep Archive
E	S3 Standard-IA



The correct answer is **S3 Glacier and S3 Glacier Deep Archive**.

- Objects stored in the S3 Glacier storage class can be retrieved within a few minutes to a few hours. By comparison, objects that are stored in the S3 Glacier Deep Archive storage class can be retrieved within 12 hours.

The other response options are incorrect because:

- S3 Standard is a storage class that is ideal for frequently accessed data, not archival data.
- S3 Intelligent-Tiering monitors access patterns of objects and automatically moves them between the S3 Standard and S3 Standard-IA storage classes. It is not designed for archival data.
- S3 Standard-IA is ideal for data that is infrequently accessed but requires high availability when needed.

Module 4 summary

In this module, you learned how to:

- Summarize the basic concepts of object storage
- Describe Amazon S3 and its features