Academy Cloud Foundations (ACF)
Module 01 Student Guide
Version 1.0.5

100-ACFNDS-10-EN-SG

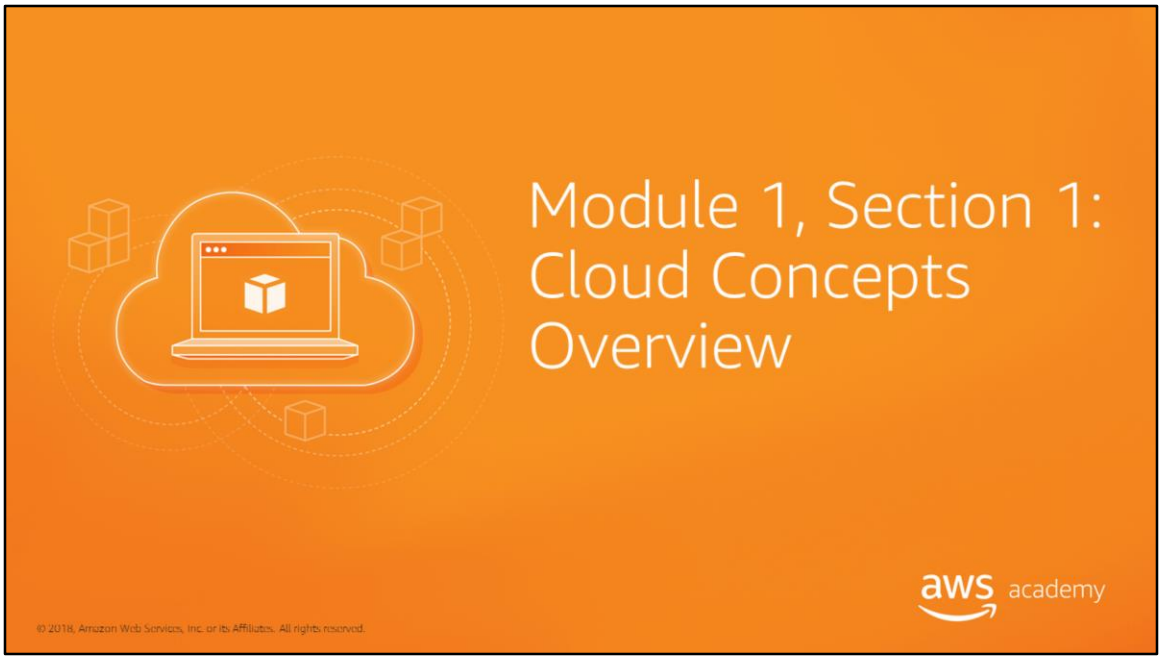
Corrections or feedback on the course, please email us at:
aws-course-feedback@amazon.com.
For all other questions, contact us at:
https://aws.amazon.com/contact-us/aws-training/.

Contents

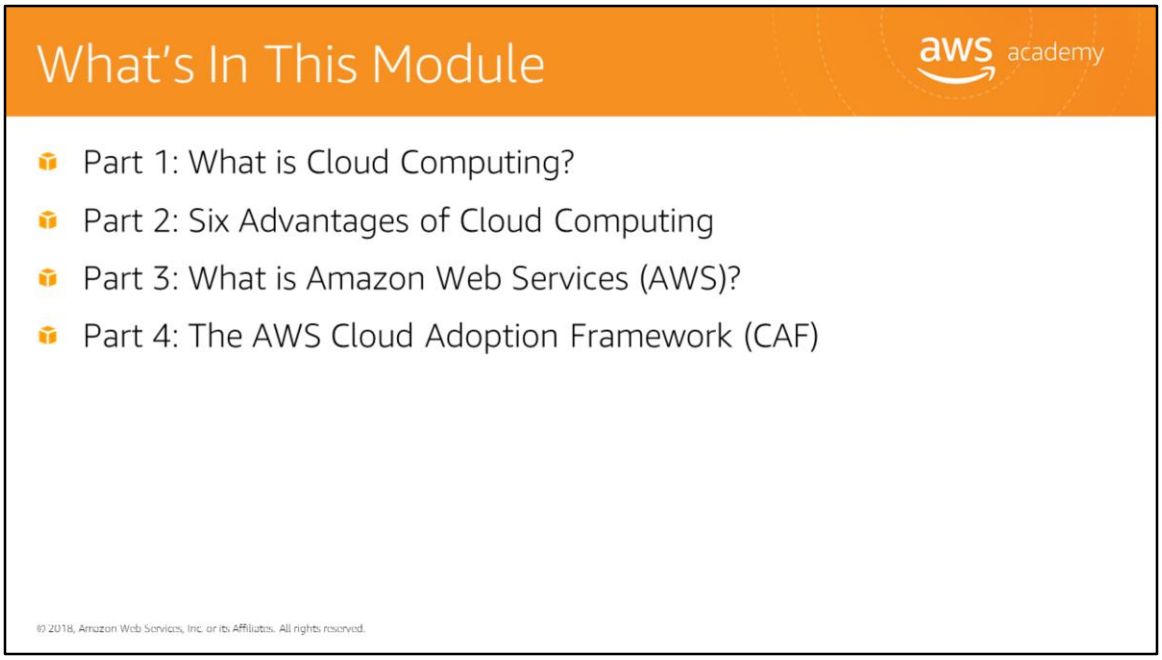
Module 1.1: Cloud Concepts Overview 4

Module 1.2: Cloud Economics 57

Module 1.3: AWS Infrastructure Overview 97

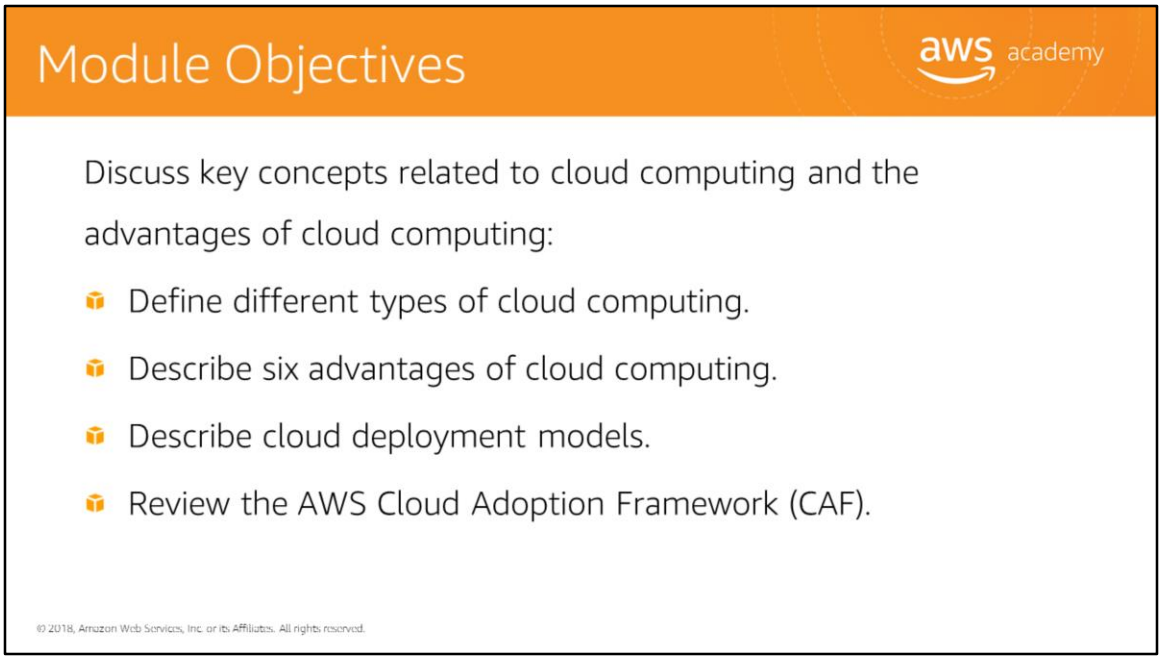


Welcome to Module 1, Section 1 – Cloud Concepts Overview.



In this module, we are going to discuss the basics of cloud computing. In part one, you'll learn what cloud computing is. In part two, you'll discover the six advantages of cloud computing. In part three, we'll reveal what Amazon Web Services is, and in part four, you'll discover the AWS Cloud Adoption Framework.

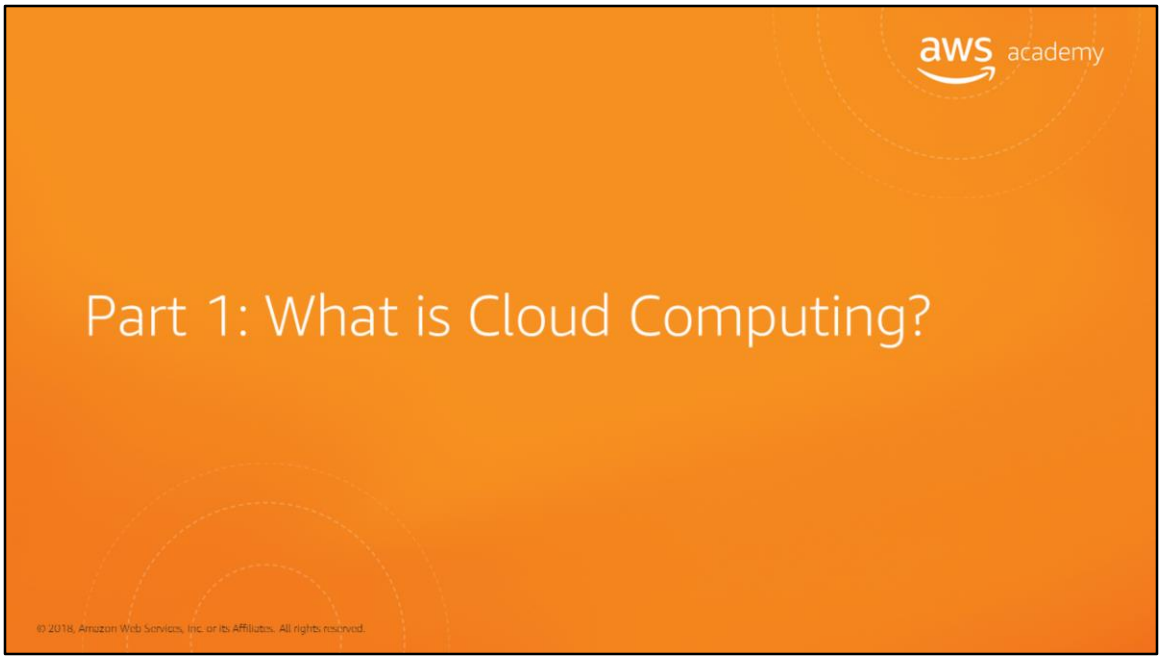
This course assumes that you have a non-IT background, as it will not teach you how to build applications in the cloud. This course will give you a general conceptual understanding about the cloud and AWS.



The goal of this module is to discuss key concepts related to cloud computing and the advantages of cloud computing with Amazon Web Services.

We will:
• Define different types of cloud computing to understand internet-based computing and three broad categories of cloud computing.
• Describe the six advantages of cloud computing that are helping organizations make the decision to get out of the low-value parts of IT and focus on things that drive business success.
• Describe three cloud deployment models to understand alternative models of cloud usage.
• Review the AWS Cloud Adoption Framework which helps organizations understand how cloud adoption transforms the way they work.

After this overview, you will have the opportunity to complete a Knowledge Assessment.



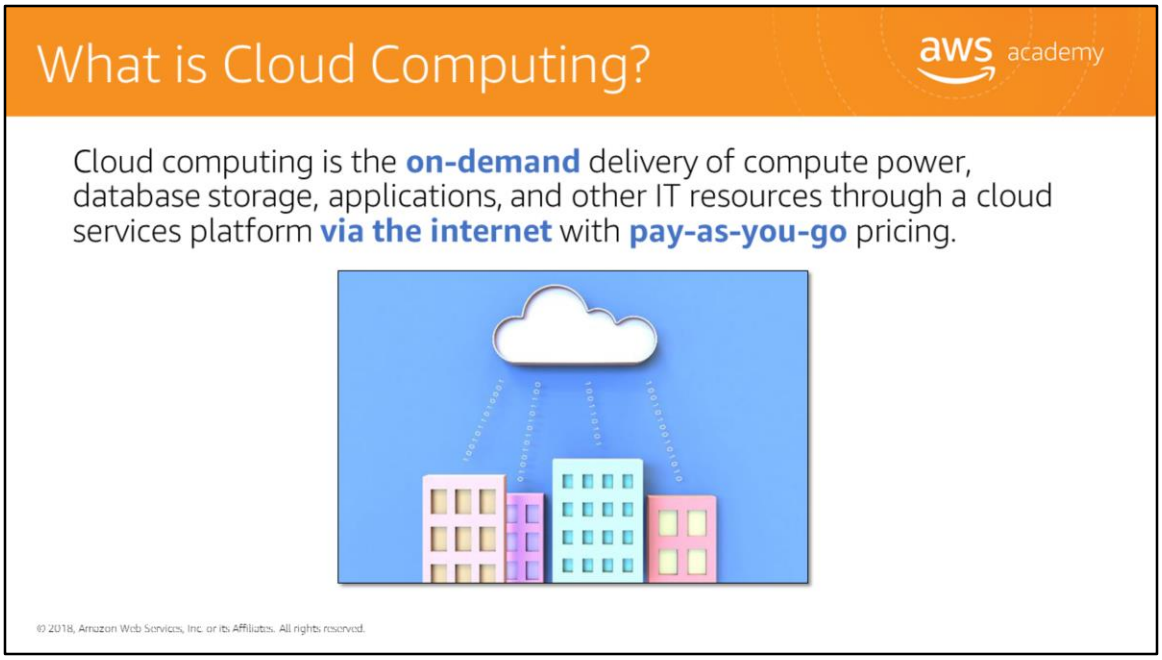
In part one, we'll answer the question, "What is cloud computing?"

What is Cloud Computing?

What does cloud computing mean to you?

Take a moment to provide a quick sentence on what you understand cloud computing to be.

Note: There’s no wrong answer.



Cloud computing is the on-demand delivery of compute power, database storage, applications, and other IT resources through a cloud services platform via the internet, with pay-as-you-go pricing.

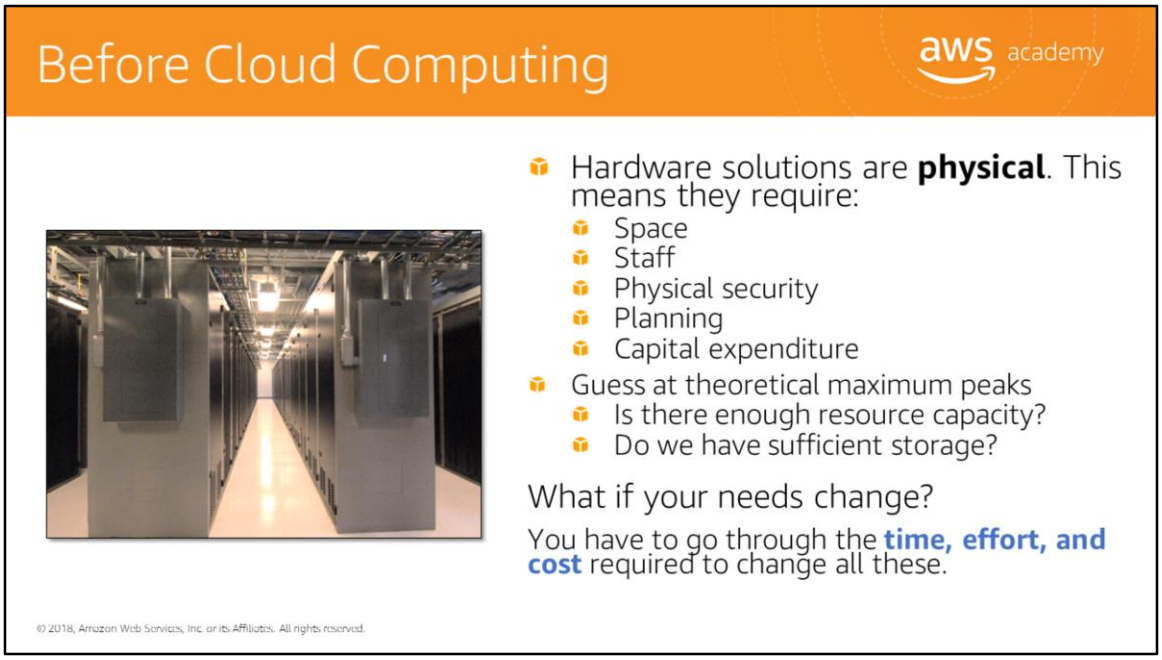
The most basic way to define what the “cloud” is that it is a computer located somewhere else that is accessed via the Internet and utilized in some way. Web services is also another name for what people call the cloud.

The cloud is comprised of server computers located in large data centers in different locations around the world. When you use a cloud service like Amazon Web Services (AWS), you are utilizing the computers owned by AWS. AWS is a cloud services provider.

The computers contain various technology features and services, like building blocks, that can be used to assemble solutions that help a user meet their business goals and technology requirements. With cloud computing, organizations can consume on-demand computing and storage resources rather than building, operating, and improving infrastructure on their own.

Visit the link to learn more
https://aws.amazon.com/what-is-cloud-computing/.

Cloud computing enables you to stop thinking of your infrastructure as hardware, and instead think of it (and use it) as software. Before cloud computing, you would have to provision capacity based on guessing theoretical maximum peaks. If you didn't meet your projected maximum peaks, or you exceeded them, you would be paying for expensive resources that would stay idle or have insufficient capacity to meet your needs.



Managing hardware takes away time and resources you could be using to improve your architecture and your application. Hardware solutions are physical. This means they require space, staff, physical security, planning and capital expenditure.

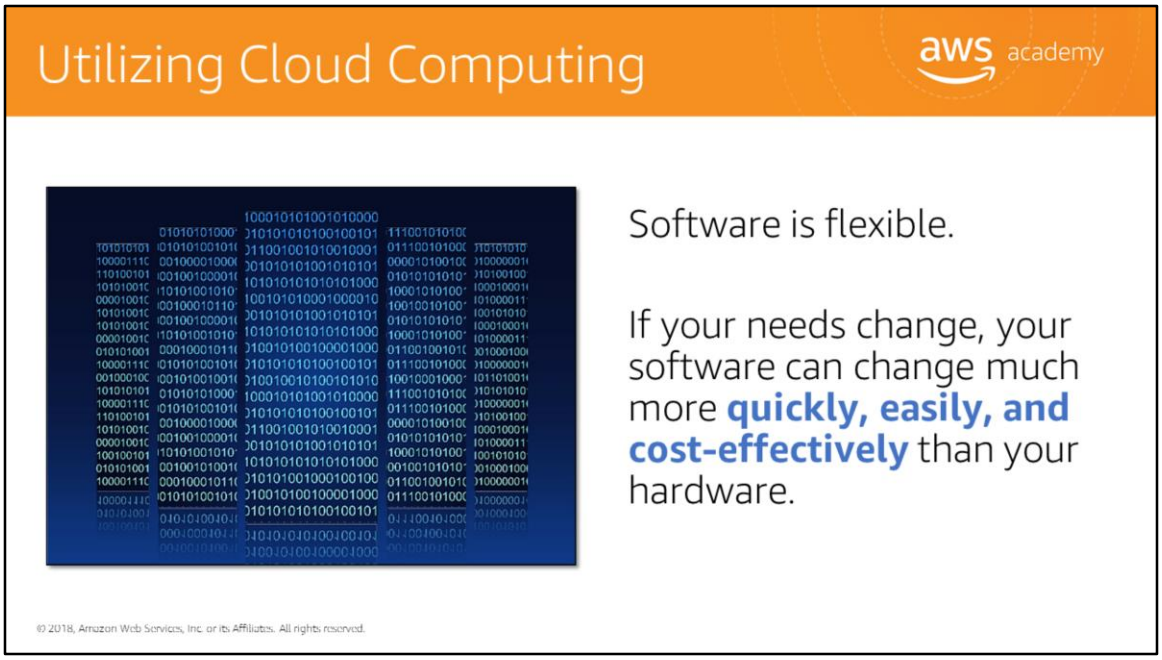
You have to guess at theoretical maximum peaks, asking if there is enough resource capacity or if you have sufficient storage. What if your needs change? You have to go through the time, effort, and cost required to change all of these.

For example, if you want to provision a new web site, you would have to go out and buy the hardware, rack and stack it, put it in a data center, and then manage it or have someone else manage it. This approach is very expensive.

Cloud computing addresses some of the issues in the traditional computing model. One of the most prohibitive aspects of traditional computing is the significant up-front investment of acquiring, provisioning, and maintaining on-premises infrastructure. Cloud computing can get businesses up and running with a new solution in place quickly and with very **low up-front costs**. Then, you can **elastically** scale up and down in an automated fashion so that you **pay only for what you use**.

Cloud computing also allows you to select the services that best match your needs, giving you

flexibility with a wide range of choices and the ability to change your configuration at will. All of these services are provided on a **secured** infrastructure.

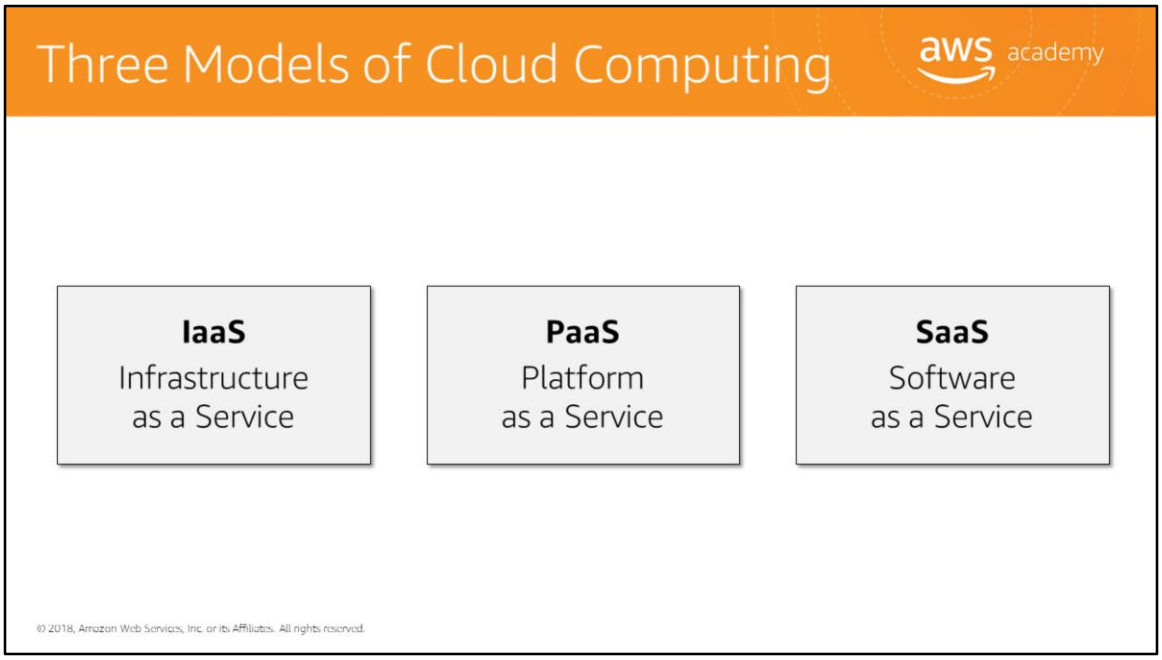


Accessing and using your infrastructure as software offers a number of benefits-flexibility in particular. If your needs change, your software can change much more quickly, easily, and cost-effectively than your hardware.

With a cloud services provider like AWS, you don't have to anticipate your hardware needs ahead of time and then order, install, and set it up at your data center. You also don’t need to undergo a long procurement cycle. With a few clicks, you can provision exactly what you need-and it will be available to you in a few minutes.

That means you can provision and terminate resources as necessary on AWS, instead of paying for hardware when you’re not using it. You can treat resources as temporary and disposable resources, free from the inflexibility and constraints of a fixed and finite IT infrastructure.

By harnessing the power of AWS, you can be more agile and efficient with change management, testing, reliability, and capacity planning.



Cloud services can fall into one of three primary categories, based mainly around how much control and responsibility you have over how the service is configured.

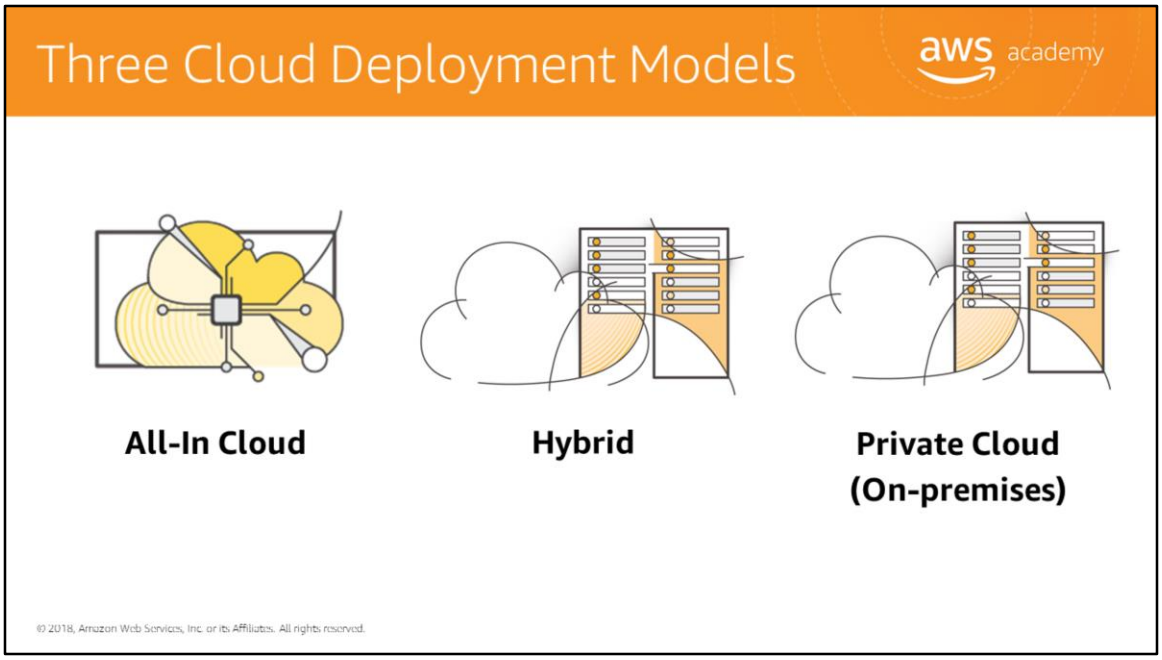
With IaaS (or Infrastructure as a Service), you manage the server, which can be physical or virtual, as well as the operating system (Windows or Linux). In general, the data center provider has no access to your server.
Basic building blocks for cloud IT include:

- Networking features
- Compute, and
- Data storage space

With **PaaS (or Platform as a Service),** someone else manages the underlying hardware and operating systems. This enables you to run applications without managing underlying infrastructure (for example -- patching, updates, maintenance, hardware and operating systems). PaaS also provides a framework for developers that they can build upon to create customized applications.

With **SaaS (or Software as a Service),** you manage your files, while the service provider takes care of all of the data centers, servers, networks, storage,

maintenance, patching, etc. All you worry about is the software and how you want to use it. You are provided with a complete product that is run and managed by the service provider. Facebook and Dropbox are examples of SaaS. You manage your Facebook contacts and Dropbox files, and the service providers manage the systems.



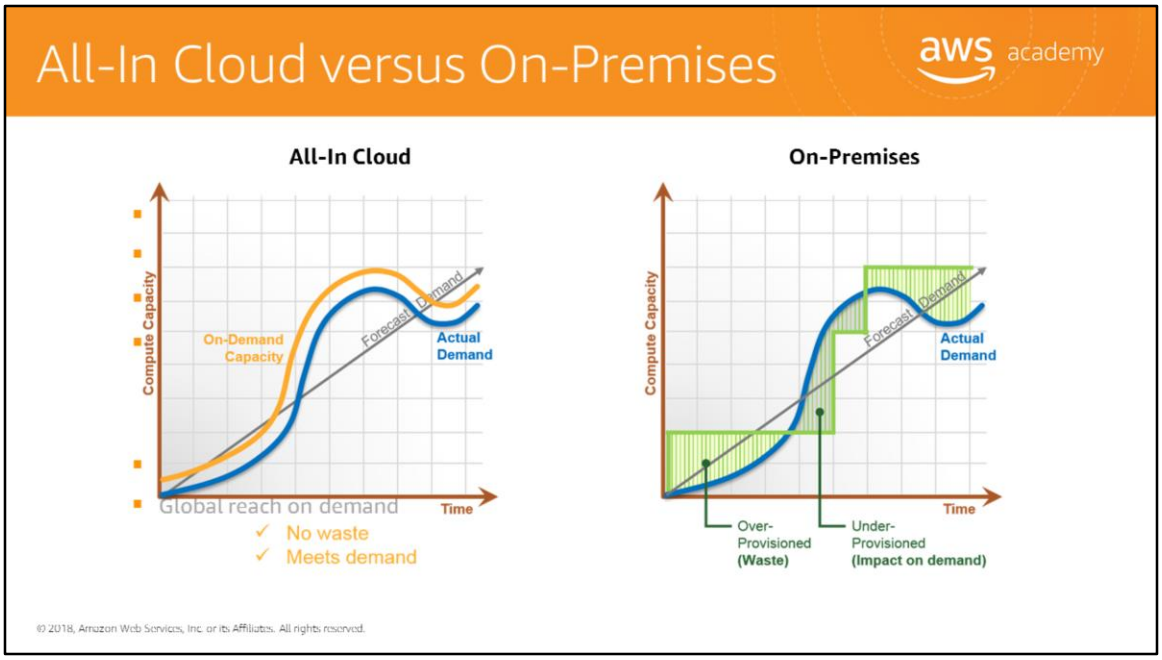
Now, let’s reveal the three cloud deployment models.

"All-In" Cloud is a cloud-based application that is fully deployed in the cloud, and all parts of the application run in the cloud. Applications in the cloud have either been created in the cloud or have been migrated from an existing infrastructure. Cloud-based applications can be built on low-level infrastructure pieces (for example, networking, compute or storage) or can use higher-level services that provide abstraction from the management, architecting, and scaling requirements of core infrastructure.

A hybrid deployment is a way to connect infrastructure and applications between cloud-based resources and existing resources that are not located in the cloud. The most common method of hybrid deployment is between the cloud and existing on-premises infrastructure (sometimes called on-prem). On-premises infrastructure is located within the physical confines of an enterprise, often in the company's data center. A hybrid deployment model is used to extend an organization's infrastructure into the cloud while connecting cloud resources to an internal system. For more information on how AWS can help you with your hybrid deployment, visit the link https://aws.amazon.com/enterprise/hybrid/.

When you run a cloud infrastructure from your own data center, that’s called on-premises or private cloud. While this kind of deployment lacks many of the benefits of cloud computing, it does provide dedicated resources and is a popular choice for organizations who need to meet certain compliance standards. In most cases, this deployment model is the same as legacy IT

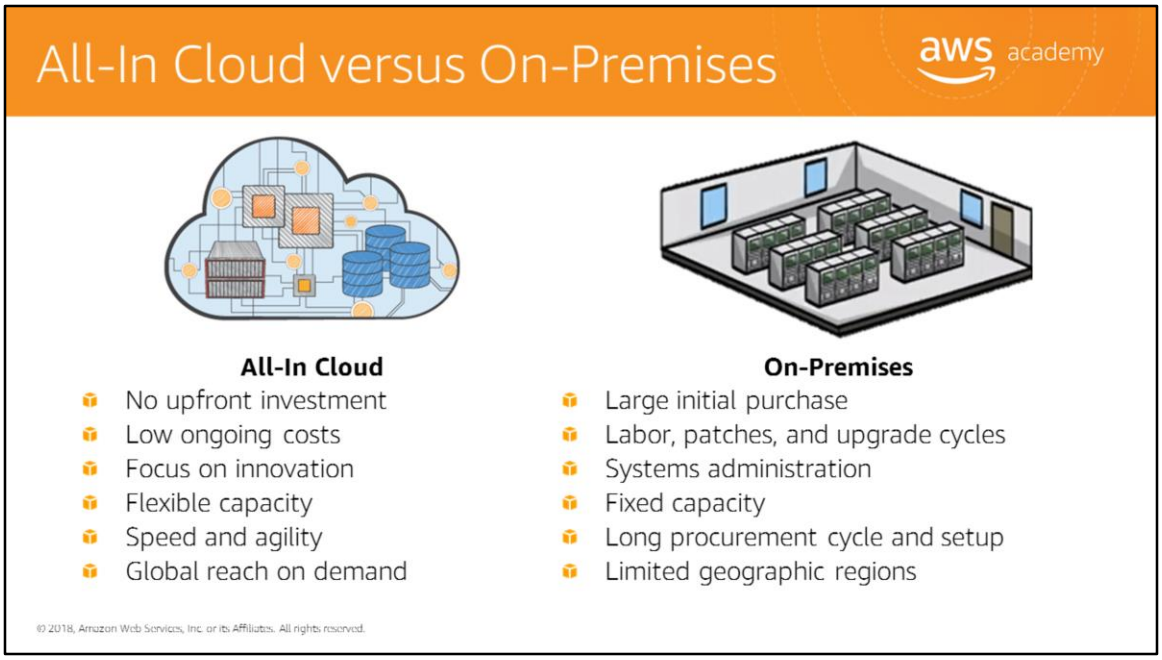
infrastructure while using application management and virtualization to increase resource utilization.



Let's take a closer look at **capacity** in the All-In Cloud and On-Premises solutions.

In the "All In" solution, capacity is in sync with demand. Resources are provisioned and decommissioned in response to demand with only a couple clicks.

In contrast, in "On-Premises" deployments, because you rely on physical hardware, you have to forecast your capacity needs well in advance of the actual demand. Instead of resources that expand and contract with demand, the on-premises solution results in idle, wasted resources waiting for demand to catch up. If demand suddenly outpaces capacity, the shortfall may result in unhappy customers. Your ability to respond quickly to this situation can be limited by long procurement cycles or by constraints on where you house your IT resources. Furthermore, building an on-premises infrastructure can be slow and expensive.

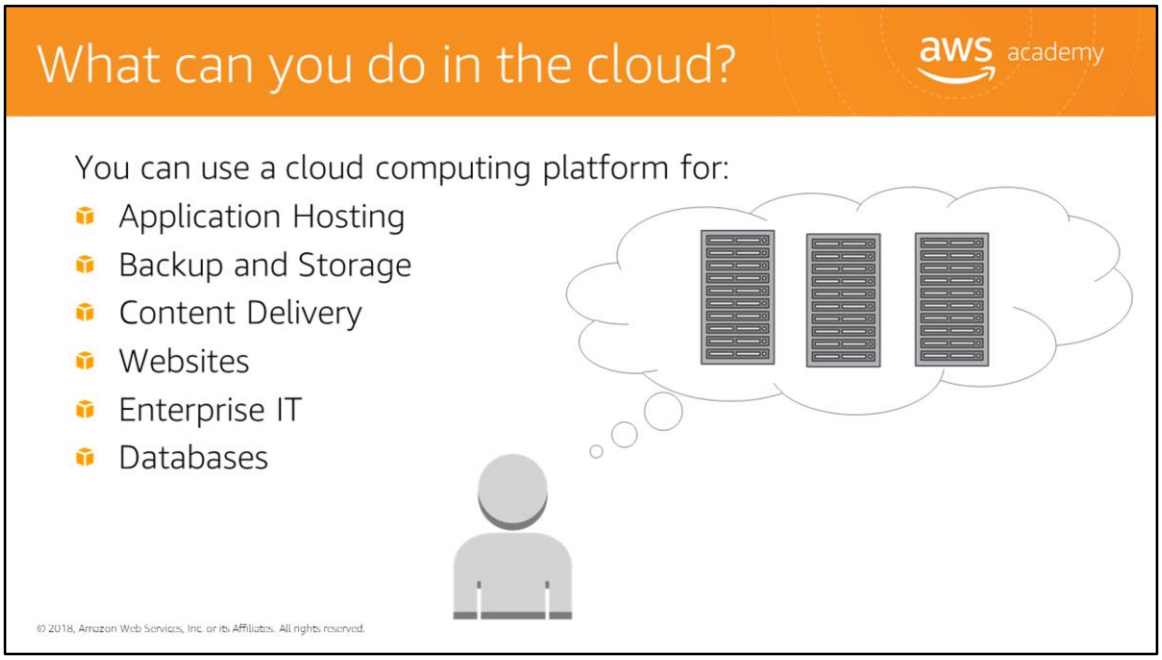


To summarize:

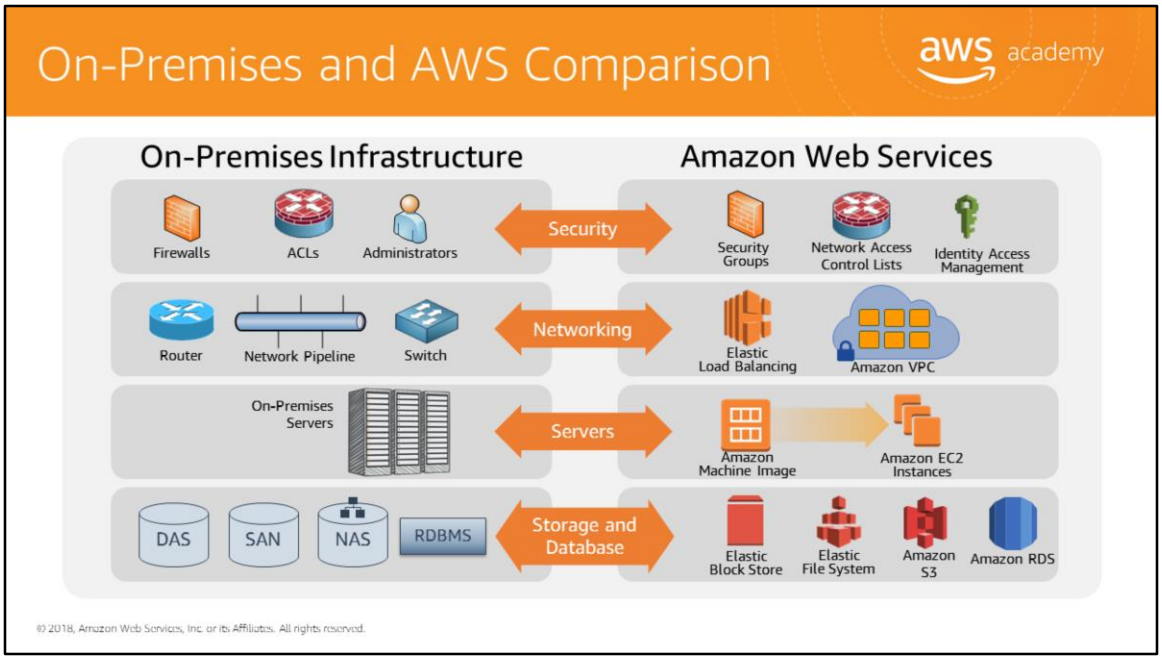
- With the All-In solution, there is no upfront investment, so you avoid the large capital purchases required for an On-Premises solution. You have immediate access to resources without having to procure, install, and configure cabling, racks, servers, and storage in a physical location with appropriate facilities like cooling and power. Instead, you just click to order and pay for the resources you need, which are available almost immediately.
- Cloud computing helps you reduce ongoing IT costs in multiple ways. AWS continually lowers prices due to massive economies of scale and continual improvements. Multiple pricing options also help you optimize costs based on your unique workloads. You pay only for what you use on a variable, monthly basis. On-premises solutions typically require upgrades on 1-year, 3-year, or 5-year cycles.
- Cloud gives you managed IT resources on demand, at a fraction of the cost of traditional infrastructure. This cost savings empowers organizations to shift resources toward innovative new projects that grow their business by focusing on “apps not ops.”
- Predicting how customers are going to adopt your new application is complex, making it difficult to estimate your infrastructure capacity needs. Flexible capacity means that your resources are dynamic. You can quickly provision resources as demand goes up and turn off what you don’t need as demand declines.
- Cloud computing’s speed and agility makes it possible for you to respond to changing market conditions. With AWS, resources can be provisioned as needed. This self-service environment changes how you develop and deploy applications, allowing your team to experiment more quickly

and more frequently. The amount of time it takes to get a server procured, delivered, and running limits this in a traditional infrastructure.

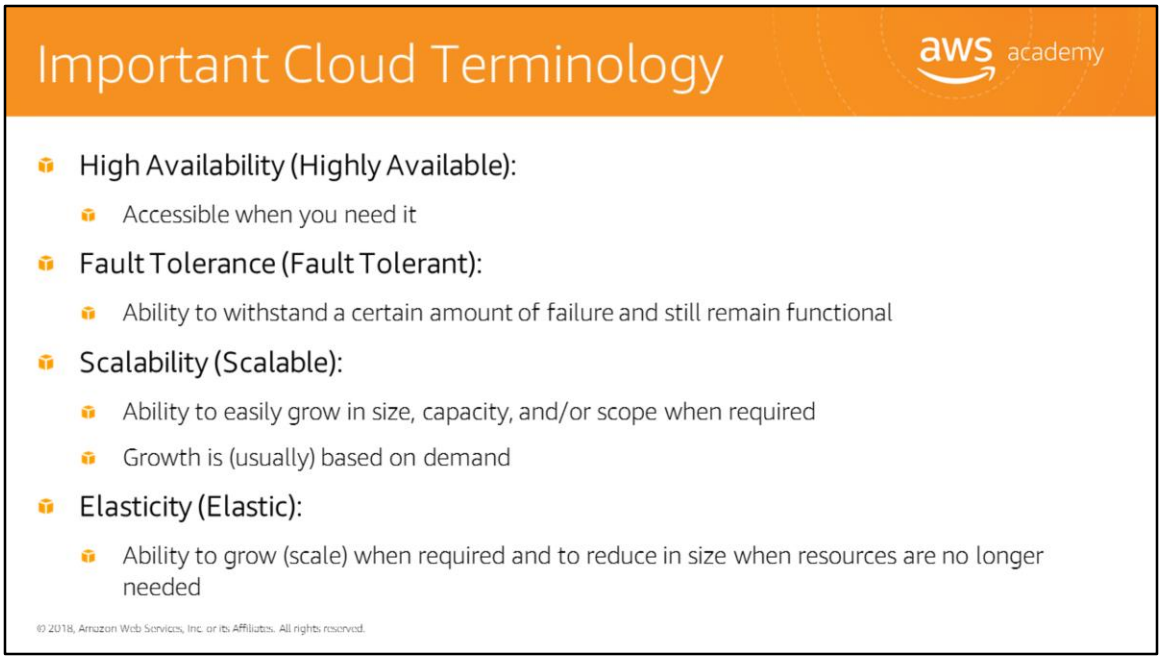
- With on-premises, it is hard to deliver great performance to a distributed user base. The initial purchase is large, it's labor intensive with patches, upgrade cycles and systems administration with fixed capacity. There are long procurement cycles and setup, so companies focus on one geographic region at a time to save costs and time. Without geographical limitations, you can deploy your application in any of the AWS regions around the world with lower latency and at minimal cost.



You can use a cloud computing platform for the following:
- **Application Hosting** for an on-demand infrastructure to host internal or SaaS applications.
- **Backup and Storage** to store data and build dependable backup solutions.
- **Content Delivery** to distribute content worldwide, with high data transfer speeds.
- Host static and dynamic **websites**.
- **Enterprise IT** to host internal- or external-facing IT applications in AWS's secure environment.
- Use a variety of scalable **database** solutions, from hosted enterprise database software to non-relational database solutions.



Many AWS services have analogs in the traditional IT space and terminology. This side-by-side comparison shows how Amazon Web Services (AWS) products and services relate to a traditional infrastructure.



Now let's review some important cloud terminology. High availability, fault tolerance, scalability, and elasticity are four terms often used when discussing the cloud. These concepts are the fundamental building blocks of AWS and will be referred to throughout the course.

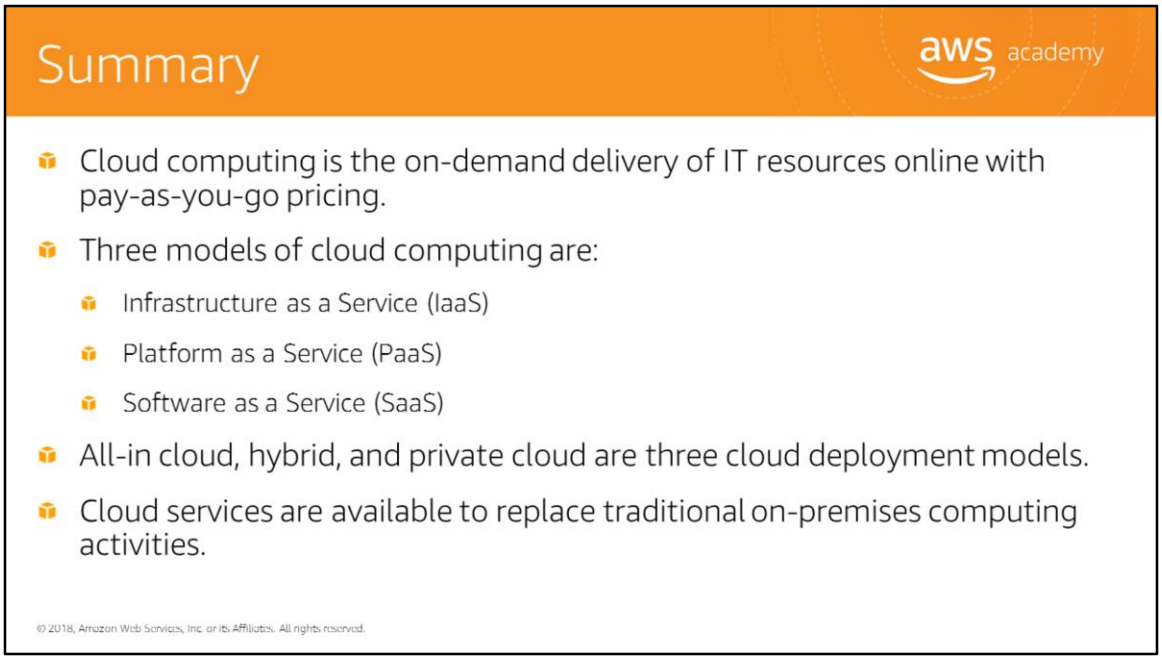
High availability refers to a resource that is accessible when you attempt to access it. For example, if every time you go to the ATM to make a withdrawal it works as expected the ATM is highly available; however, if you go to use it and there is a sign on the front that says "Out of Order", it is not highly available.

Fault tolerance is the ability to withstand a certain amount of failure and still remain functional. It also refers to the ability of a system to be self-healing and return to full capacity despite a failure. It is the ability of a system to fail in some way but still remain functional.

Scalability is the ability to easily grow in size, capacity, and/or scope when required particularly in response to demand. If something cannot quickly grow in an easy manner it is not scalable.

Elasticity is the ability to not only grow (or scale) when required, but also to reduce or contract in size as needed. A system that is elastic can scale to grow as needed usually based

on demand and contract as demand decreases.

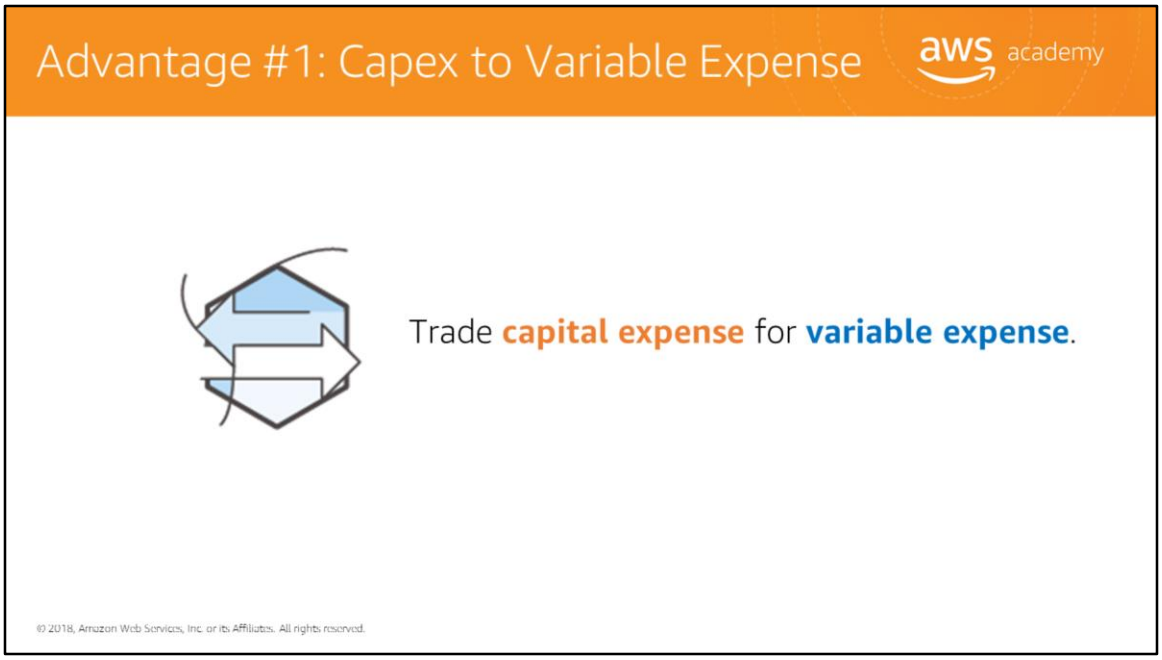


In summary, cloud computing is the on-demand delivery of IT resources online with pay-as-you-go pricing.
The three models of cloud computing are:
Infrastructure as a Service (IaaS)
Platform as a Service (PaaS), and
Software as a Service (SaaS)

All-in cloud, hybrid, and private cloud, are three cloud deployment models.

Cloud services are available to replace traditional on-premises computing activities.

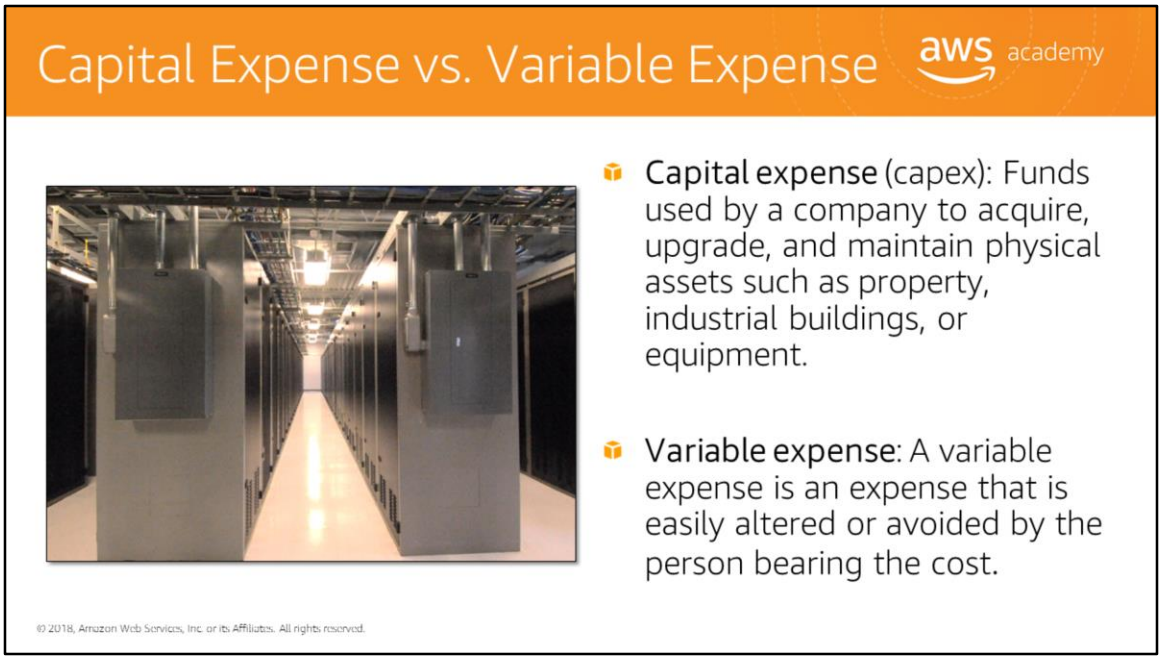
Why are so many companies interested in moving to the cloud? Let's take a look at the six benefits companies can realize by moving to the cloud.



Advantage 1: Trade Capital Expense for Variable Expense

The first advantage is to trade capital expense for variable expense.

Instead of having to invest heavily in data centers and servers before you know how you're going to use them, you can pay only when you use computing resources, and pay for how much you use.
Select the link to learn more https://d1.awsstatic.com/whitepapers/aws-overview.pdf.

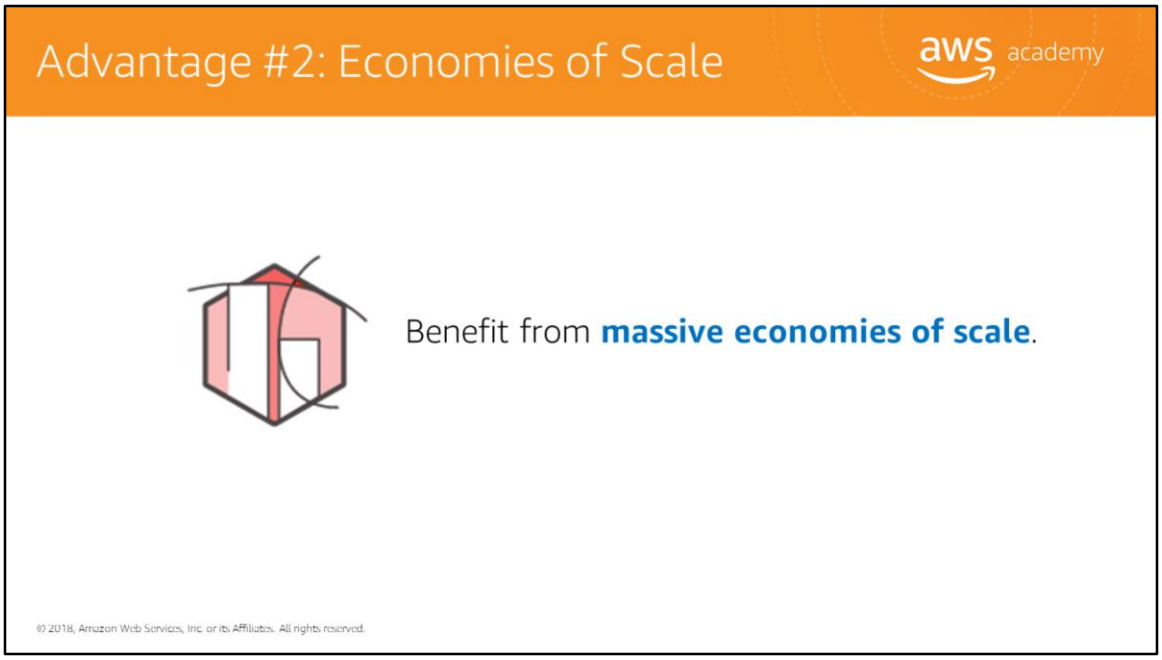


Capital expenses, or capex, are funds used by a company to acquire, upgrade, and maintain physical assets such as property, industrial buildings, or equipment.

Remember the data center example where we racked and stacked the hardware and then had to manage it all? You have to pay for everything in the data center whether you use it or not. Furthermore, when you purchase a hardware solution, you cannot take advantage of the same massive economies of scale that Amazon can.

A variable expense is an expense that is easily altered or avoided by the person bearing the cost. By using the cloud, businesses won't have to invest money into data centers and servers, and can pay for only what they use in a pay-as-you-go fashion. This lets businesses save money on technology and enables them to adapt to new applications with as much space as needed in minutes, rather than weeks or days. Maintenance is reduced so the business can spend more time focusing on the core goals of the business.
See the links to learn more.

Source:
https://www.investopedia.com/terms/c/capitalexpenditure.asp
Overview of Amazon Web Services: https://aws.amazon.com/whitepapers/overview-of-amazon-web-services.

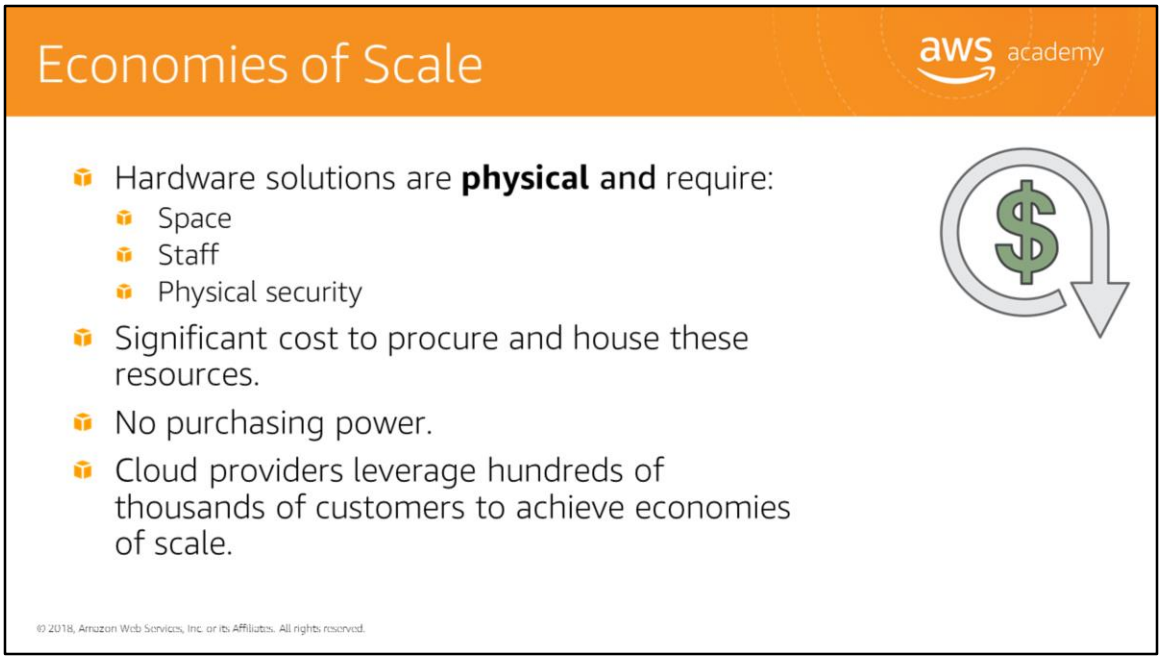


Advantage 2: Benefit from massive economies of scale.

Advantage number two is that 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.

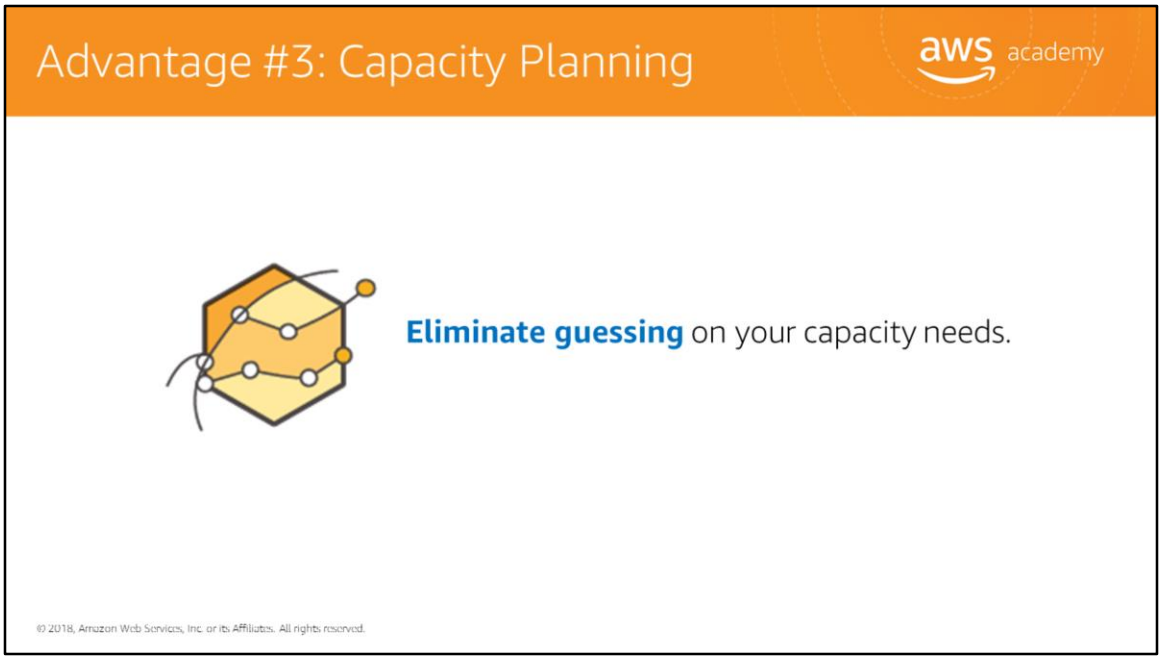
Select the link to learn more https://d1.awsstatic.com/whitepapers/aws-overview.pdf.



Data centers require hardware solutions, which are physical, and require space, staff, and physical security. Significant cost and time is associated with the procurement of these resources. Additionally, purchasing power is limited to the size of their individual purchases.

In contrast, with usage from hundreds of thousands of customers are aggregated in the cloud, providers such as Amazon Web Services can achieve higher economies of scale, which translates into lower pay as-you-go prices.

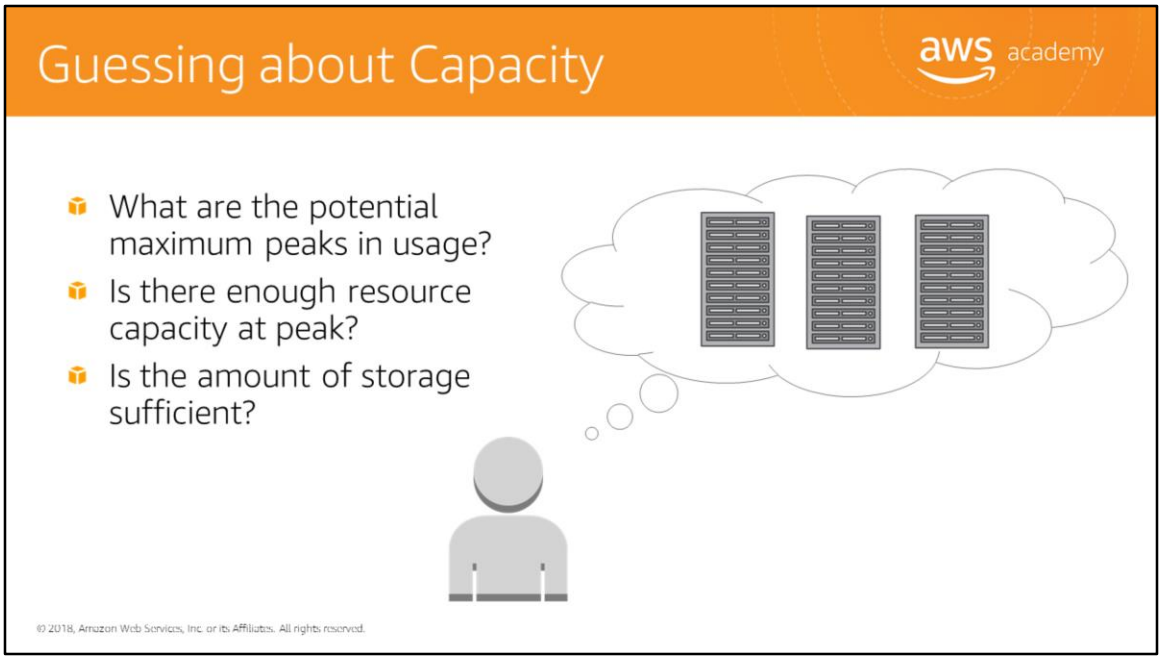
Sources:
AWS | What Is Cloud Computing - Benefits of the Cloud. Amazon Web Services, Inc.N.p., n.d.Web. 31 Aug. 2014.



Advantage 3: **Stop guessing about capacity.**

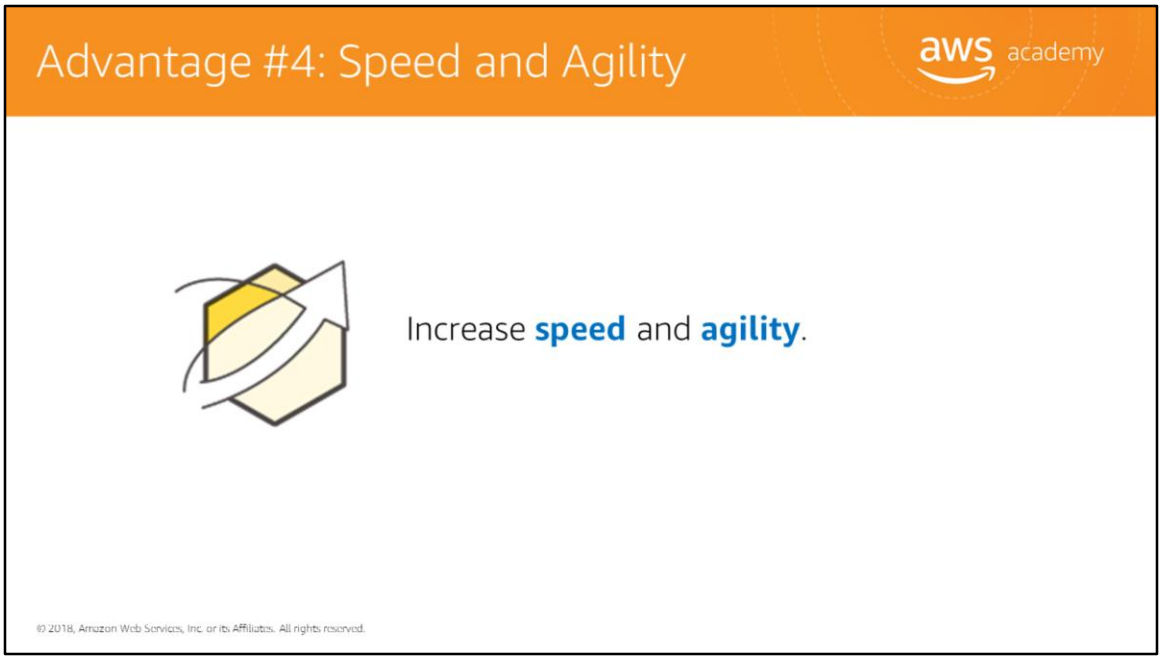
Advantage number three is to eliminate guessing on your infrastructure capacity needs.

Source:
https://d1.awsstatic.com/whitepapers/aws-overview.pdf



Before cloud computing, we had to guess about how much resources would be required to accommodate maximum usage peaks. That method also assumed that you could accurately predict the usage peaks, if there's enough resource capacity and if the amount of storage is sufficient. When you guess, it is very likely that you will probably either buy too much or too little. If you buy too much, you've wasted money. If you buy too little, you will have downtime.

With cloud computing, these problems go away. You can access as much or as little as you need, and scale up and down, scale in and out as required with only a few minutes notice.



Advantage 4: Increase speed and agility.

Advantage four is increased speed and agility.

In a cloud computing environment, new IT resources are only a click away, which means you reduce the time it takes to make those resources available to your developers from weeks to just minutes. This results in a dramatic increase in agility for the organization, since the cost and time it takes to experiment and develop is significantly lower.

Source:
https://d1.awsstatic.com/whitepapers/aws-overview.pdf

Increase Speed and Agility

- Rapid availability of new resources
  - Provision resources in minutes, not weeks.
- Increase Innovation
  - Quick, low cost experimentation.
  - Leverage pre-fabricated functionality without requiring in-house expertise. (i.e., data warehousing, analytics)
- Increase experimentation
  - Explore new avenues of business with minimal risk and expense.
  - Test with different configurations.

Cloud services provide instant global reach and rapid availability of new resources that enables you to quickly change or scale your technology in minutes, not weeks.

Organizational agility is a core differentiator in today's rapidly changing business environment. For organizations to succeed in a constantly changing world, they need to improve their ability to change and adapt.

You can safely experiment with new ideas and encourage innovation at a very low cost and leverage pre-fabricated functionality without requiring in-house expertise, such as data warehousing and analytics.

You may even find success with ideas that simply weren't feasible in the past due to hardware or budget constraints. You're able to explore new avenues of business with minimal risk and expense, and can test with different configurations.

Advantage 5: Stop spending money running and maintaining data centers.

Advantage five is to stop spending money running and maintaining data centers.

Focus on projects that differentiate your business, not the infrastructure. Cloud computing lets you focus on your own customers, rather than on the heavy lifting of racking, stacking, and powering servers.

Stop Spending Money on Data Centers

- Focus on customers
- Focus on projects that differentiate the business
- Delegate the racking, stacking and powering of servers to the cloud provider

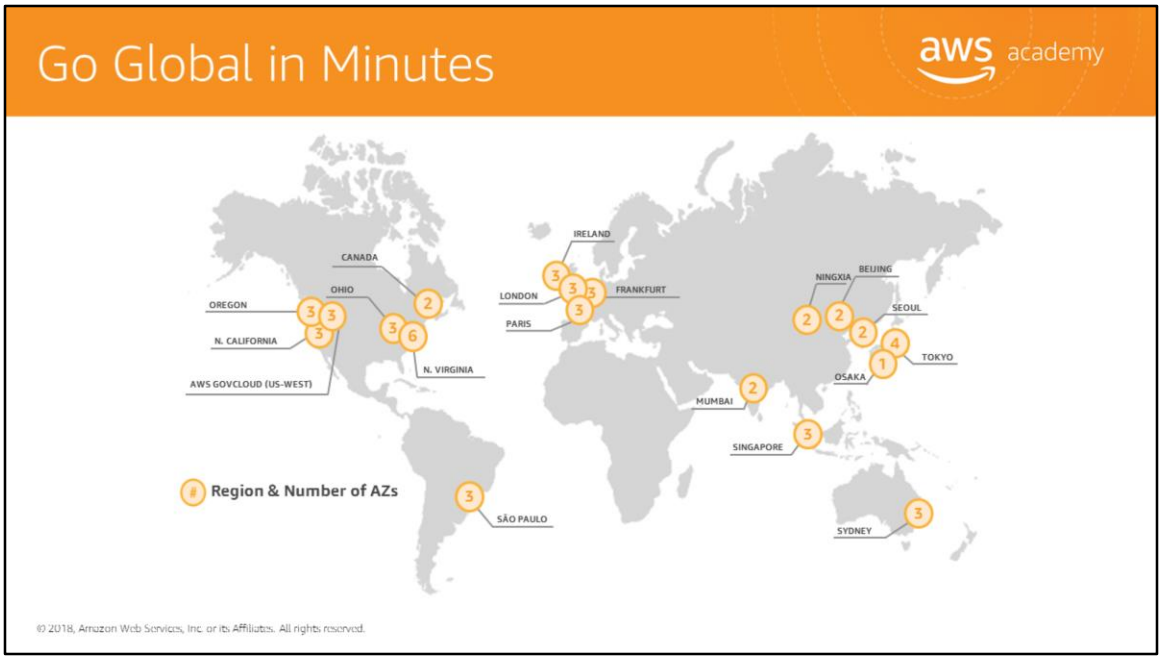
Cloud computing allows you to focus on your customers and your core business - what you are good at - rather than on managing infrastructure. Let someone else manage that for you. Focus on projects that differentiate the business, not the infrastructure. Delegate the racking, stacking and powering of servers to the cloud provider!

Advantage 6: Go global in minutes.

Advantage six is the ability to go global in minutes.

Easily deploy your application in multiple regions around the world with just a few clicks. This means you can provide a lower latency and better experience for your customers simply, and at minimal cost.

Select the link to learn more.
https://www.youtube.com/watch?v=JIQETrFC_SQ.



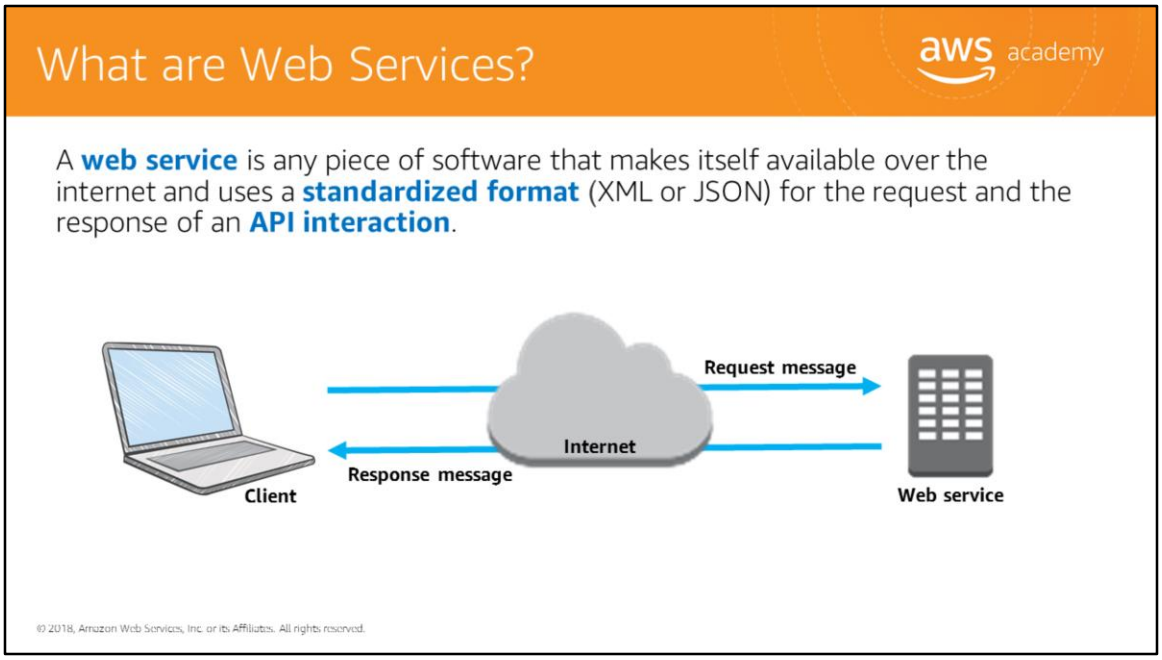
You can be operating locally today and become available globally in minutes by leveraging **Amazon CloudFront** and our global network of **Edge Locations**. Deploying in the region where your customers are helps you provide lower latency and a better experience at minimal cost.

To learn more about specific locations, visit the link. https://aws.amazon.com/about-aws/global-infrastructure/.

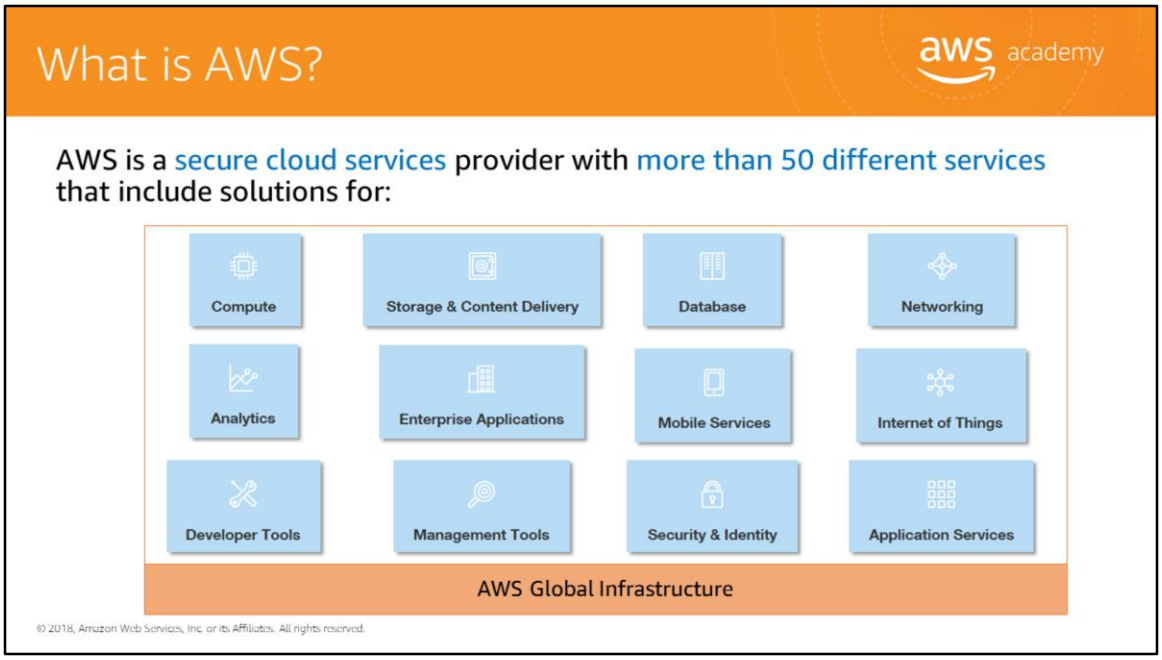
The six benefits discussed in this section provide a strong value proposition for moving to the cloud. Each of these benefits should be considered when deciding between an on-premises or cloud solution.

For more information, visit the link. https://d1.awsstatic.com/whitepapers/aws-overview.pdf.

In part three, we'll learn what Amazon Web Services is.

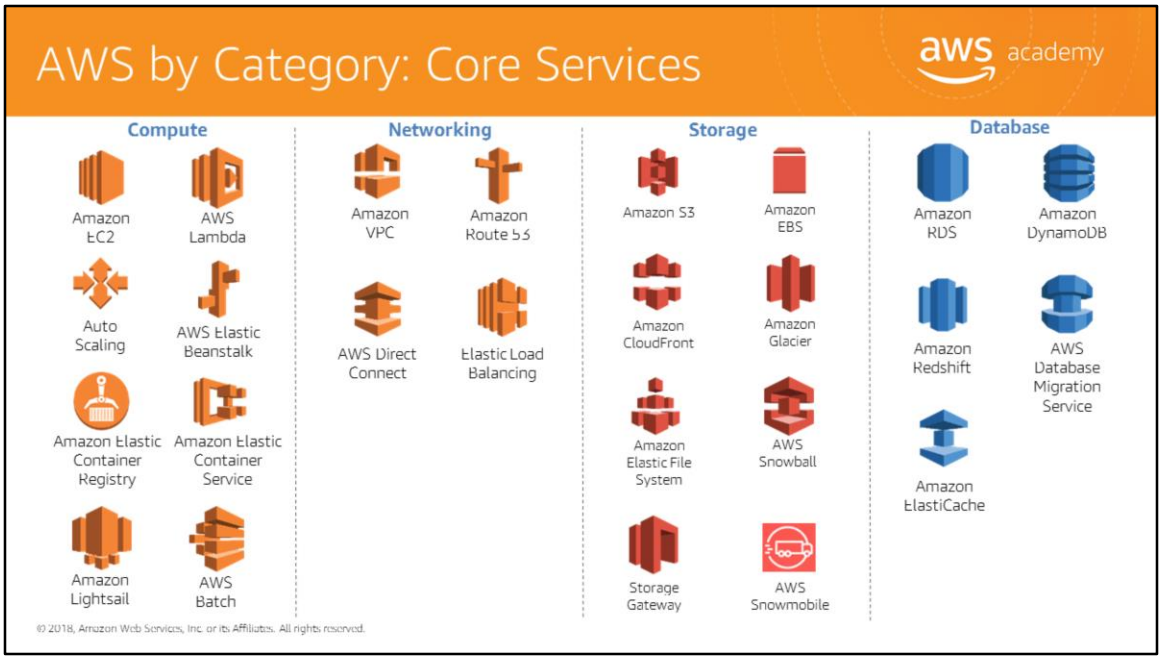


A web service is any piece of software that makes itself available over the Internet or on private (intranet) networks. A web service uses a standardized format (such as XML or JSON) for the request and the response of an API interaction. It is not tied to any one operating system or programming language, it's self-describing via an interface definition file and is discoverable.



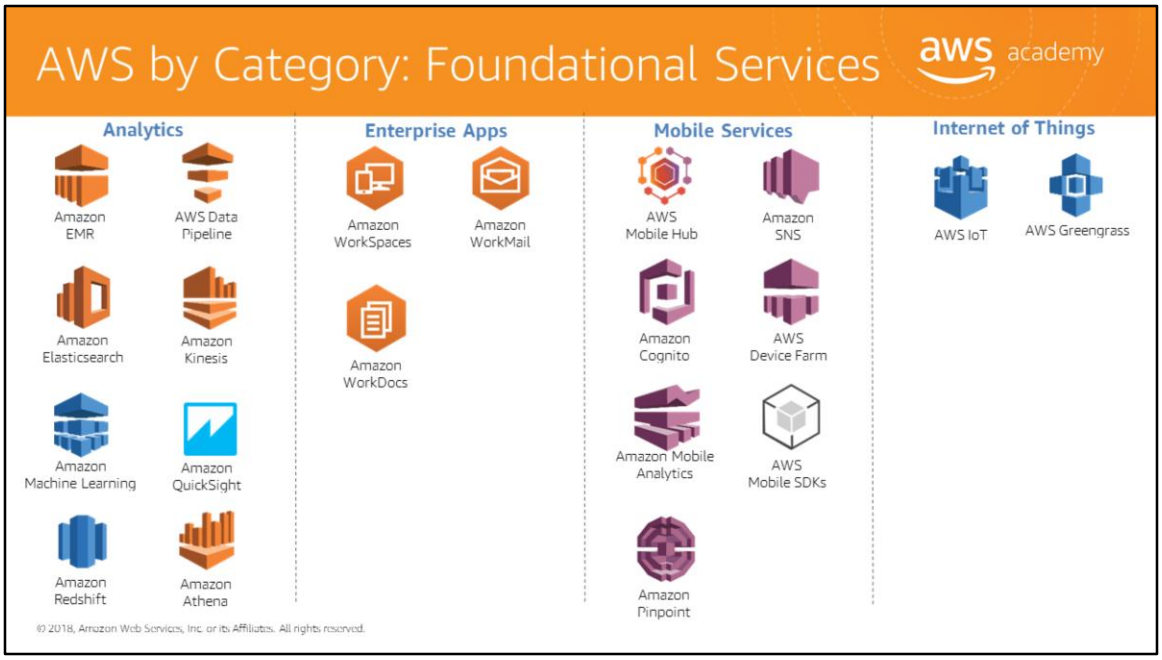
AWS is a secure cloud services provider with more than fifty services to help businesses scale and grow.

The AWS cloud provides a broad set of infrastructure services, such as compute power, storage options, networking, and databases delivered as an on-demand utility that is available in seconds, with pay-as-you-go pricing. All of these services reside on AWS global infrastructure.



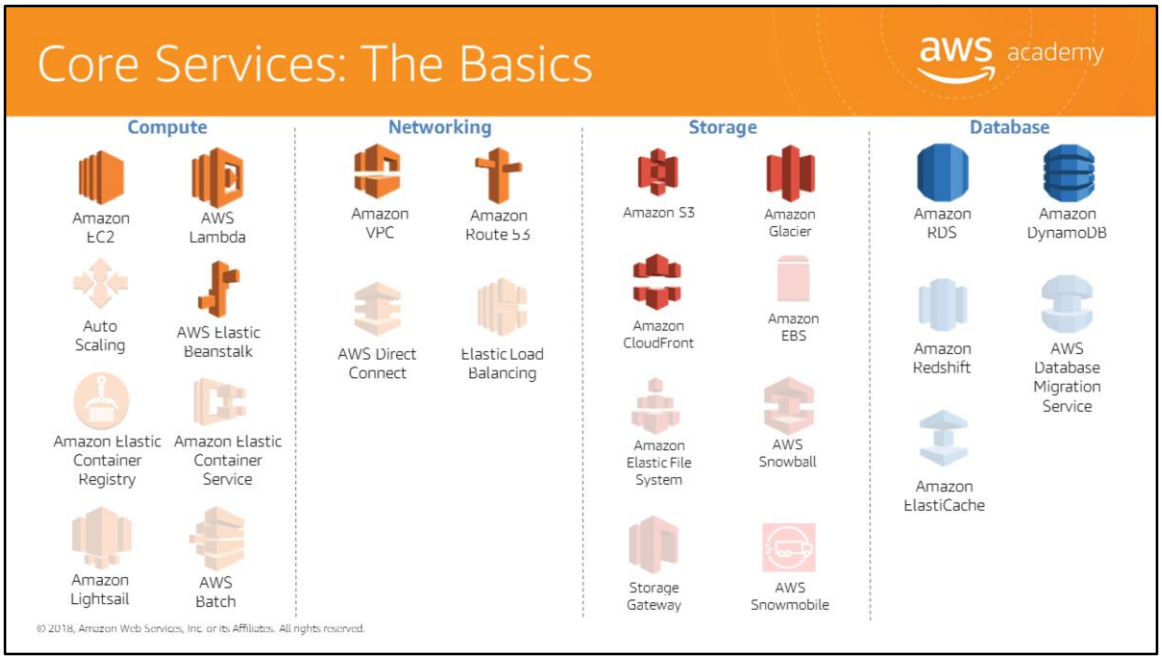
AWS offers many services. **Core Services** refers to a broad and deep group of core cloud infrastructure services. The blue label denotes the service group category. Each of the individual icons represents a service within that group.

AWS services are typically grouped into different categories like compute, networking, storage, applications, databases, and analytics.



Foundational Services refers to groups of services that offer cloud-based solutions for the analytics, enterprise, mobile, and Internet of Things (IoT) platforms.

The **AWS Developer Tools** is a set of services designed to enable developers and IT operations professionals practicing DevOps to rapidly and safely deliver software via Management Tools, security and identify, and app services.



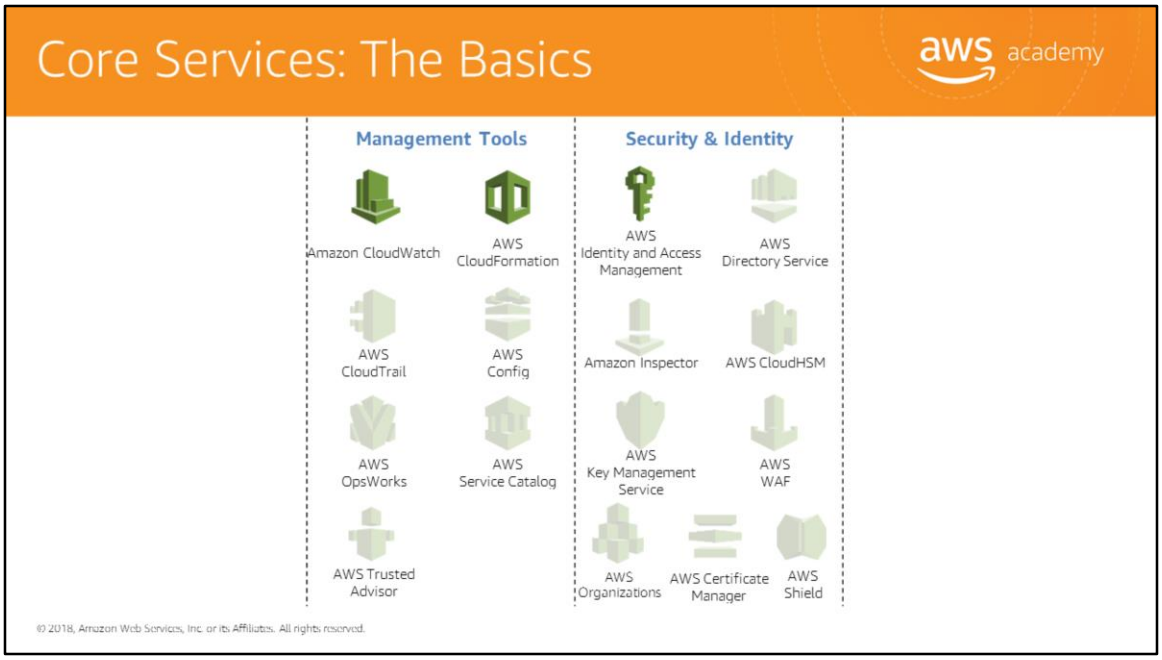
The array of AWS services can be intimidating as you start your journey into the cloud. Initially, you only need to focus on a few ”core” services. You will need to understand the AWS Global Infrastructure, several Compute services, Networking & Content Delivery, Storage, Databases, Security & Identity Access Management, and, finally, Management Tools. Specifically, you should understand the following services from the core service group:

• Compute -- including Amazon Elastic Compute Cloud (or Amazon EC2), AWS Lambda, and AWS Elastic Beanstalk.

• Networking -- including Amazon Virtual Private Cloud (or VPC), Amazon Route 53, and Domain Name Services.

• Storage -- including Amazon S3 (or Simple Storage Service), and Amazon Glacier.

• Databases -- including Amazon RDS (or Relational Database Service), and Amazon DynamoDB (or Non-Relational Database).



From the Developer and Operations group, you should also understand the following:

- Management Tools including CloudWatch and AWS CloudFormation.
- Security and Identity -- including AWS Identity and Access Management (IAM)

Access to AWS Services

- AWS Management Console
  - Access on the go with AWS Console Mobile App
- AWS Command Line Interface (AWS CLI)
- Software Development Kits (SDK)

You may wonder how to access this broad array of services. You can access them in any of three ways: using the **AWS Management Console,** the **AWS Command Line Interface (AWS CLI),** and via **Software Development Kits (SDKs)**. For access on the go, you can use the AWS Console Mobile App to quickly view resources on the go.

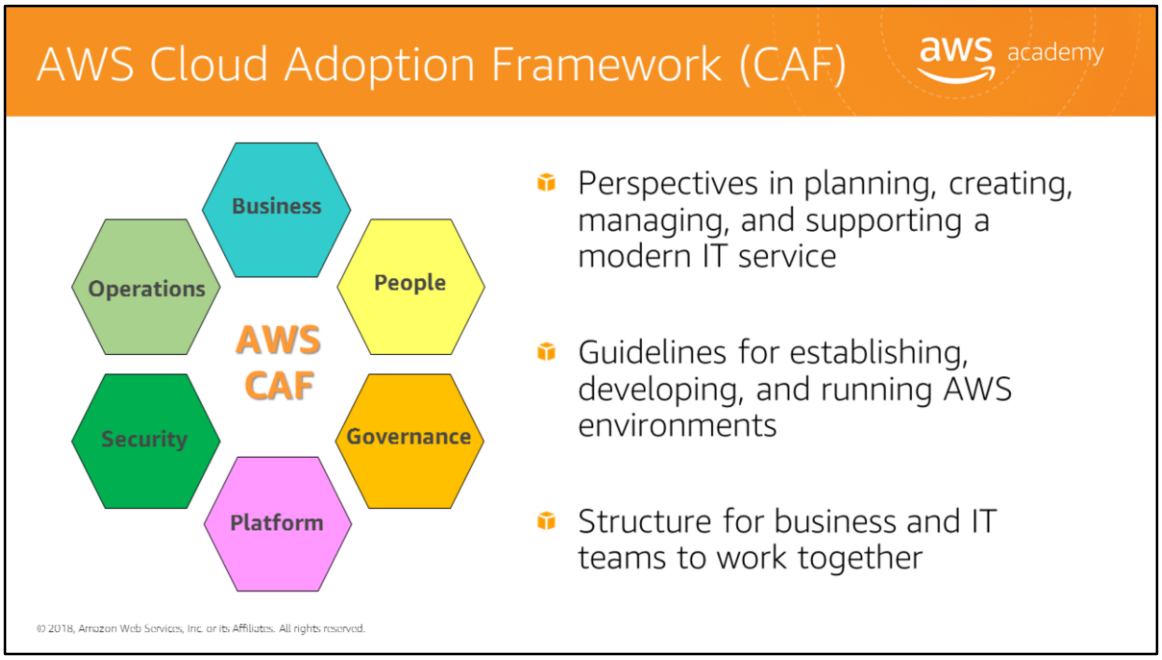
The AWS Management console breaks down AWS services into separate categories, like Compute, Management Tools, Mobile Services, etc. Each category has a number of different services in it, and all of the services can be accessed from the console. For example, the Compute category contains Amazon EC2, Amazon EC2 Container Service, Amazon Lightsail, AWS Elastic Beanstalk, AWS Lambda, and AWS Batch.

The AWS CLI is a unified tool to manage your AWS services. With just one tool to download and configure, you can control multiple AWS services from the CLI and automate them through scripts. The CLI User Guide instructs you on how to install and configure the tool. After that, you can begin making calls to your AWS services from the CLI.

Simplify using AWS services in your applications with an Application Programming Interface (or API) tailored to your programming language or platform.

To learn more, visit the links.
https://aws.amazon.com/console/mobile/
https://docs.aws.amazon.com/cli/latest/userguide/cli-chap-welcome.html
https://aws.amazon.com/tools/
https://docs.aws.amazon.com/apigateway/latest/developerguide/welcome.html

In part four, we'll discuss the AWS Cloud Adoption Framework, which helps organizations understand how cloud adoption transforms the way they work, and it provides structure to identify and address gaps in skills and processes.

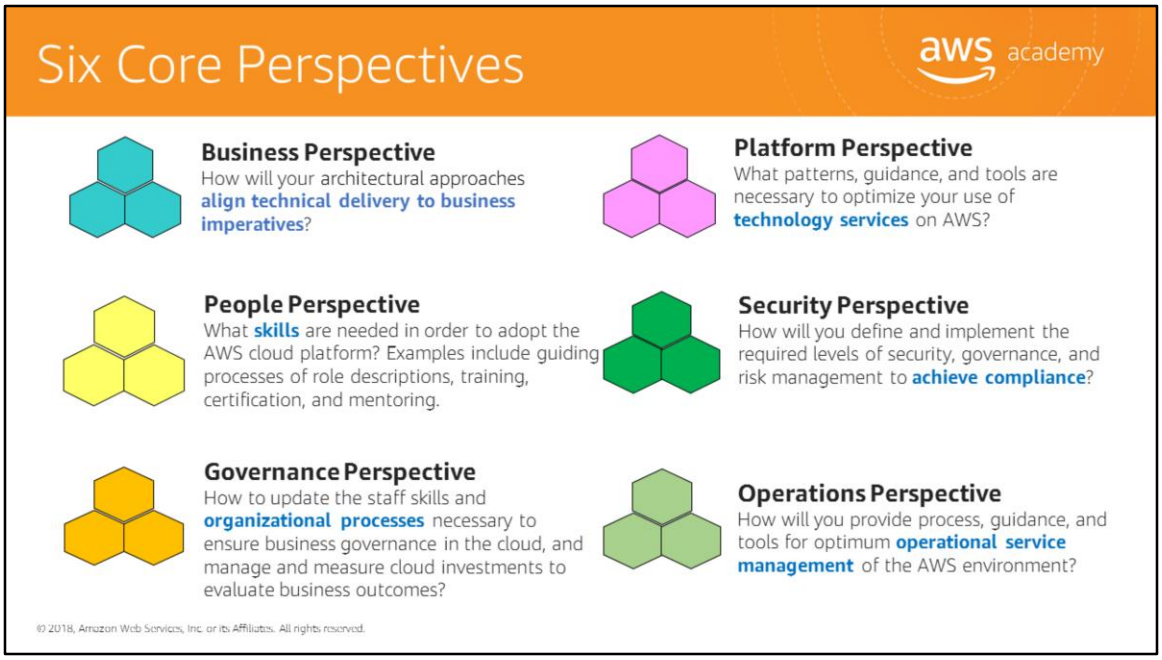


AWS Professional Services created the AWS Cloud Adoption Framework (AWS CAF) to help organizations develop efficient and effective plans for their cloud adoption journey. The guidance and best practices provided by the framework help you build a comprehensive approach to cloud computing across your organization, and throughout your IT lifecycle.

The AWS CAF breaks down the complex process of planning a move to the cloud into manageable pieces called perspectives. Perspectives represent essential areas of focus that span people, processes, and technology. Capabilities within each perspective identify which areas of your organization require attention. From that, actions are organized into prescriptive work streams that support a successful cloud journey.

The AWS Cloud Adoption Framework provides guidelines for establishing, developing, and running AWS environments. You'll receive guidance that supports each unit in your organization so that each area understands how to update skills, adapt existing processes, and introduce new processes to take maximum advantage of the services provided by cloud computing.

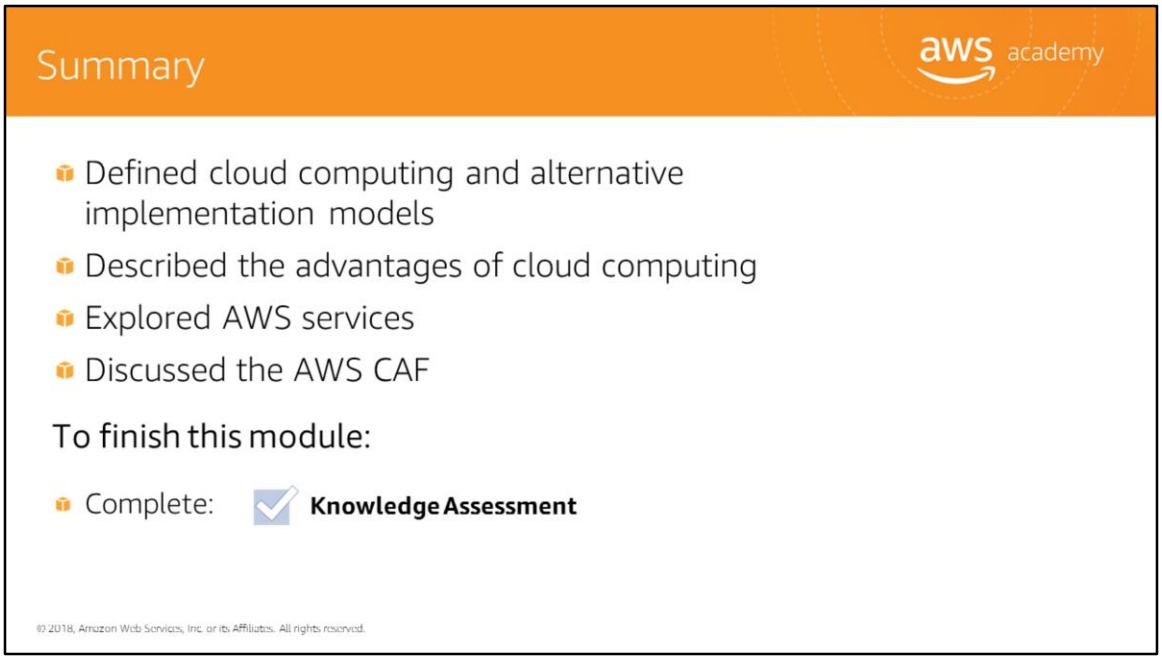
This provides a structure for business and IT teams to work together. Thousands of organizations around the world have successfully migrated their businesses to the cloud, relying on the AWS Cloud Adoption Framework to guide their efforts. AWS and our partners provide tools and services that can help you every step of the way to ensure complete understanding and transition.



At the highest level, the AWS Cloud Adoption Framework (AWS CAF) organizes guidance into six areas of focus, called perspectives.

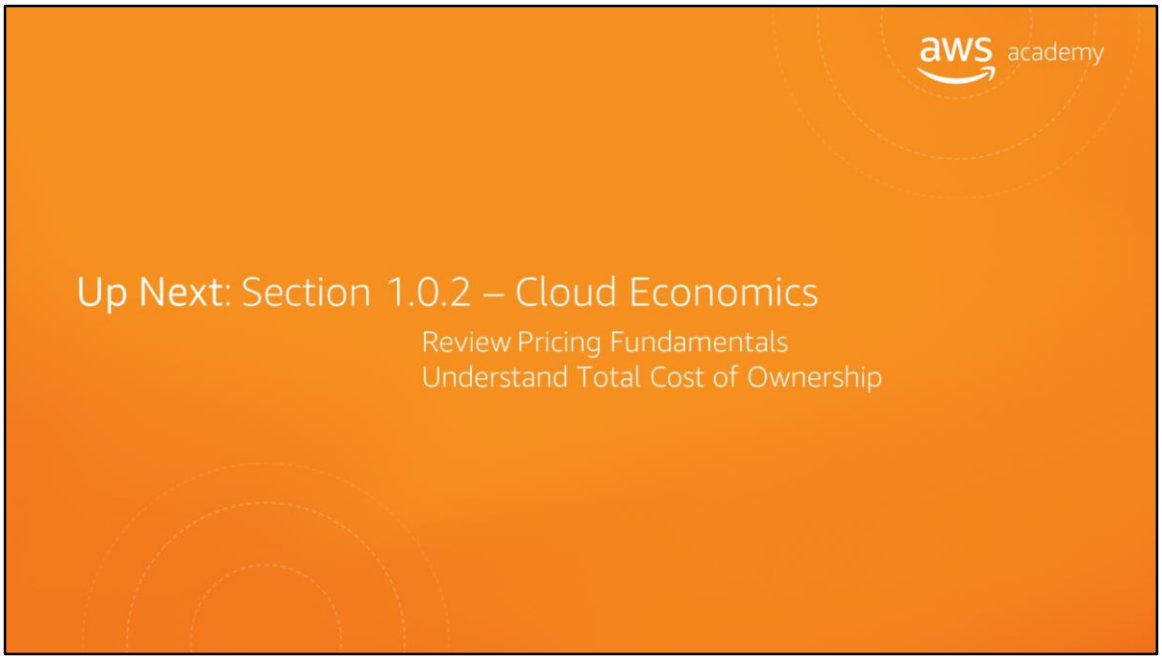
Each perspective covers distinct responsibilities owned or managed by functionally related stakeholders. In general, the Business, People, and Governance Perspectives focus on business capabilities; while the Platform, Security, and Operations Perspectives focus on technical capabilities.

For more information about the AWS CAF visit the link http://bit.ly/AWSCAF.



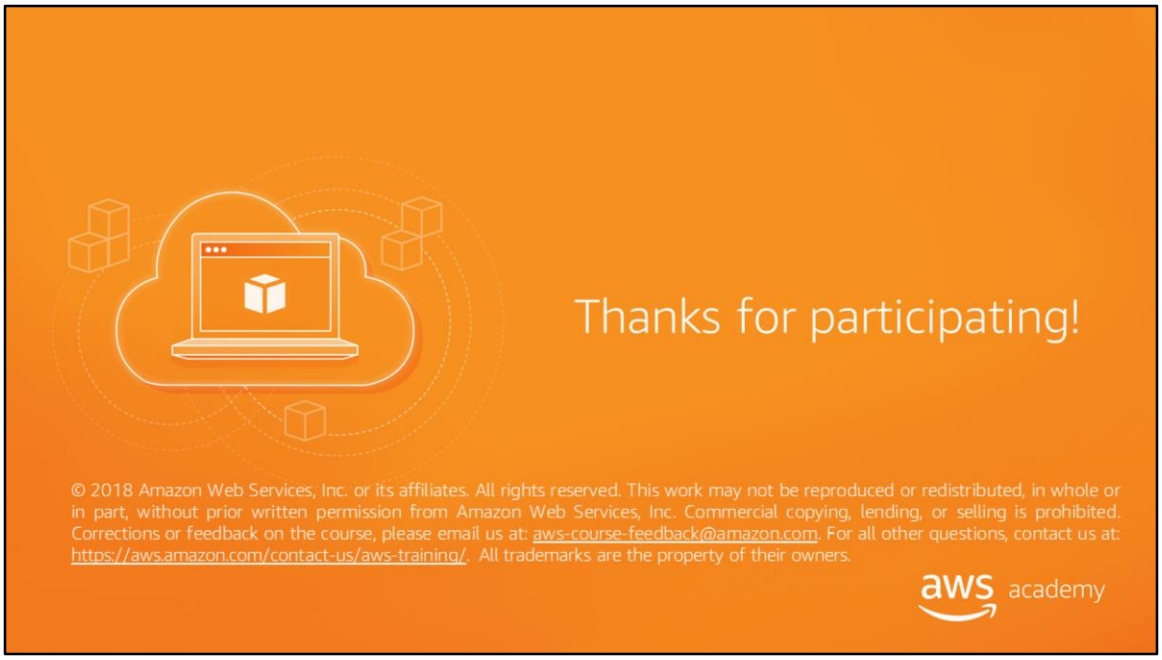
In summary, we defined cloud computing and alternative implementation models. We described the advantages of cloud computing, explored AWS services, and discussed the AWS Cloud Adoption Framework.

To finish this module, complete the knowledge assessment.



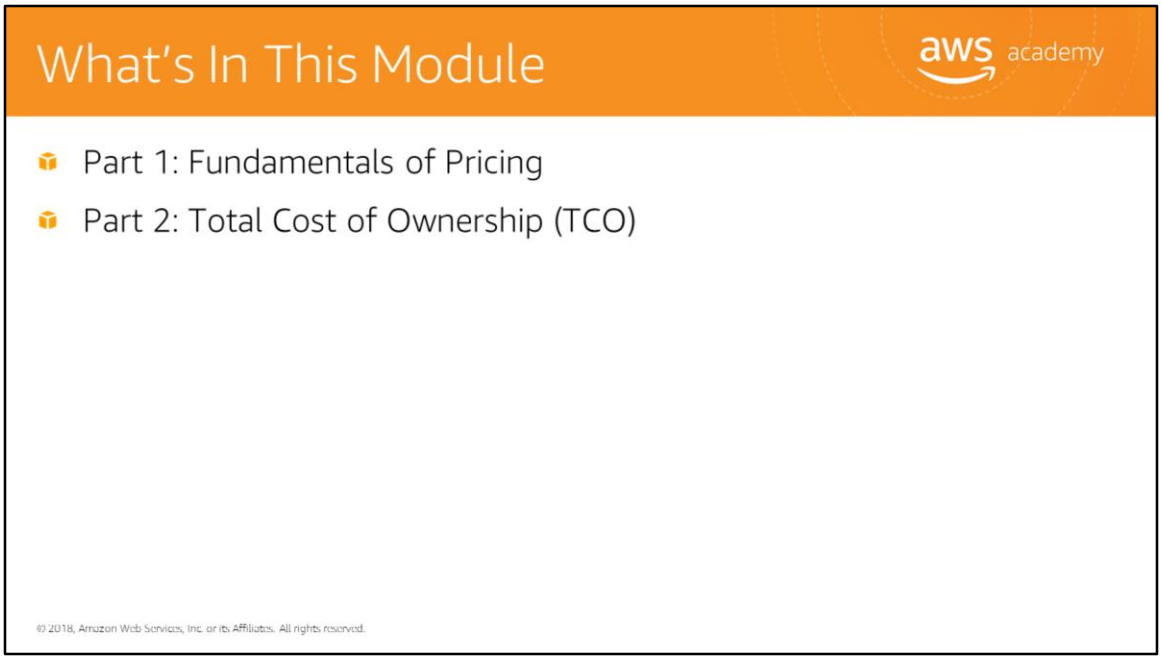
Up next, we will move on to Section 1.0.2 where we review the basics of cloud economics.

We'll review pricing fundamentals and understand the total cost of ownership.



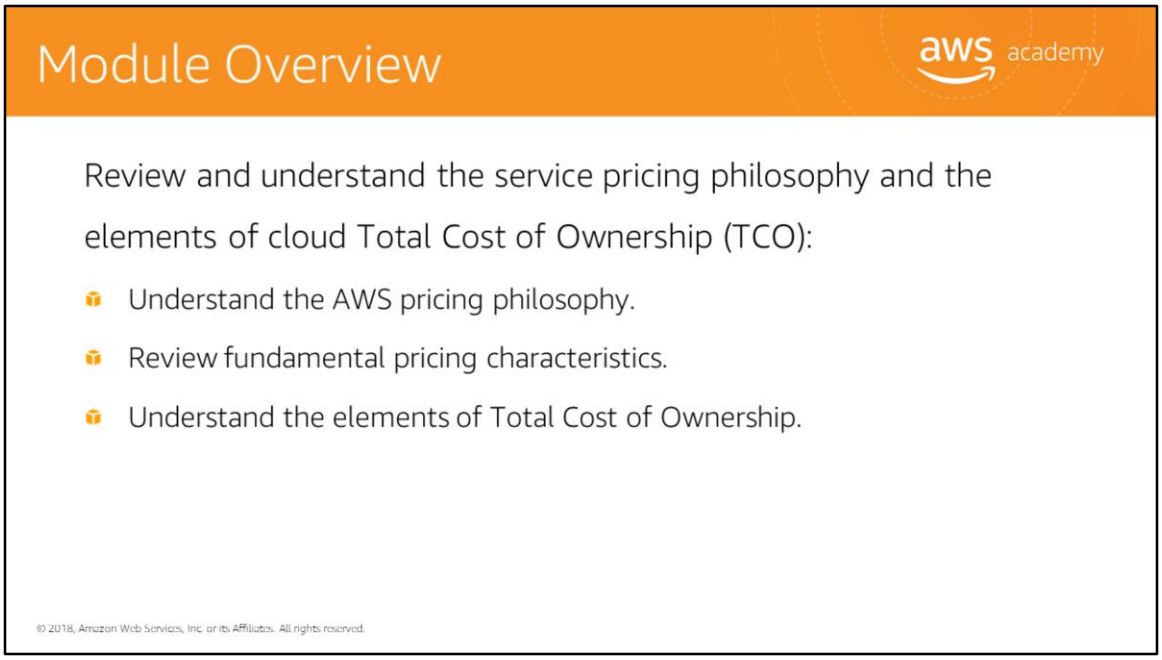
Thanks for participating!

Welcome to Module 1, Section 2 – Cloud Economics.

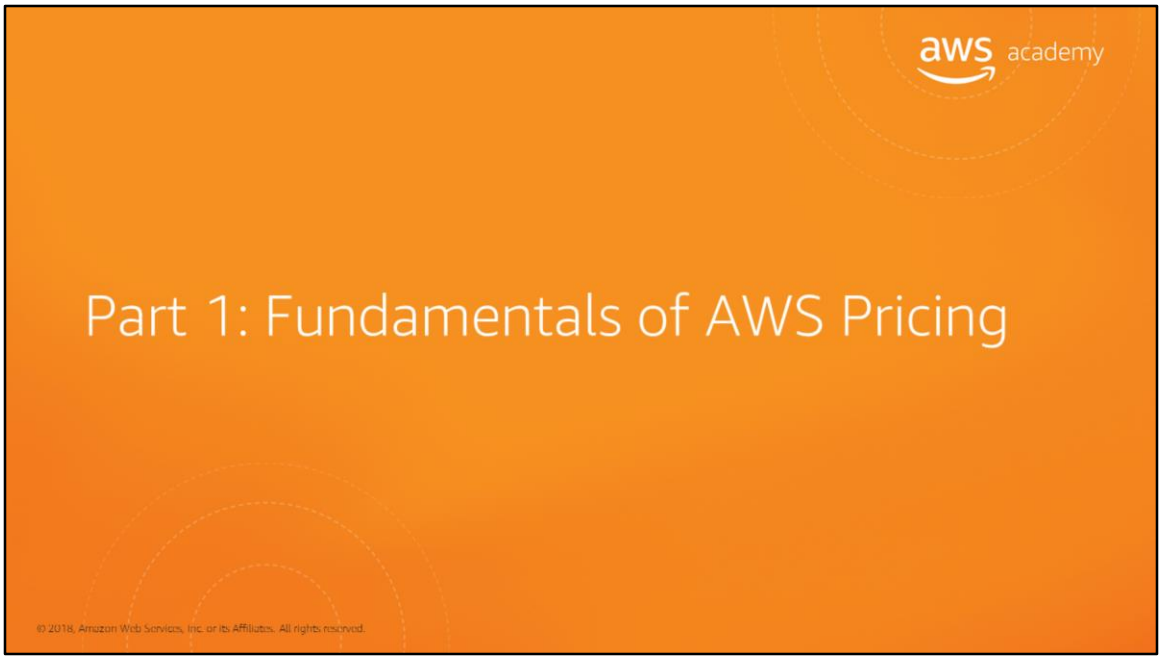


In this module, we are going to discuss the economics of cloud computing.

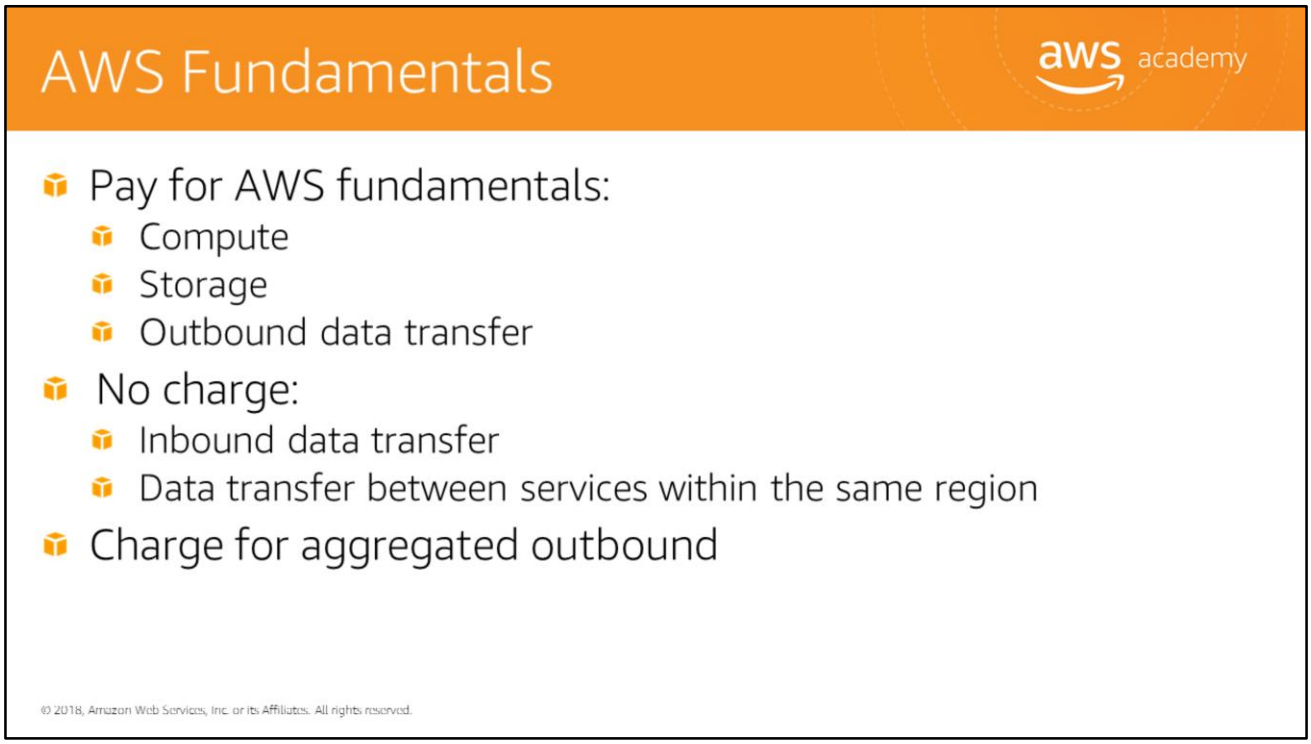
Part one will introduce you to the fundamentals of pricing, and in part two we'll review total cost of ownership.



Upon completing this module, you'll understand the Amazon Web Services pricing philosophy, review fundamental pricing characteristics and understand the elements of total cost of ownership, which is used to calculate the total cost of purchasing and operating a technology product during its useful life. Total cost of ownership gives businesses a framework to evaluate competing solutions to a problem and will be part of cloud conversations with your business partners.



Introducing Part 1: Fundamentals of AWS Pricing.



There are three fundamental characteristics you pay for with AWS: **compute**, **storage**, and **outbound data transfer**. These characteristics vary slightly depending on the AWS product you are using. However, these are the core characteristics that have the greatest impact on cost.

There is no charge for inbound data transfer or for data transfer between other AWS services within the same AWS Region, although there is a charge for aggregated outbound data transfer.

The outbound data transfer is aggregated across:
Amazon Elastic Compute Cloud (or Amazon EC2),
Amazon Simple Storage Service (or Amazon S3),
Amazon Relational Database Service (or Amazon RDS),
Amazon SimpleDB,
Amazon Simple Queue Service (or Amazon SQS),
Amazon Simple Notification Service (or Amazon SNS), and
Amazon Virtual Private Cloud (or Amazon VPC), and then charged at the outbound data transfer rate. This charge appears on the monthly statement as AWS Data Transfer Out.

AWS Pricing Model

AWS Pricing Policy:

While the number and types of services offered by AWS have increased dramatically, our philosophy on pricing has not changed: at the end of each month, you pay for what you use, You can start or stop using a product at any time. No long-term contracts required.

- Pay for what you use
- Pay less when you reserve
- Pay less when you use more
- Pay even less as AWS grows

This philosophy is what underlies AWS pricing. While the number and types of services offered by AWS have increased dramatically, our philosophy on pricing has not changed. At the end of each month, you pay for what you use. You can start or stop using a product at any time. No long-term contracts are required.

AWS offers a range of cloud computing services. For each service, you pay for exactly the amount of resources you actually need. This utility-style pricing model includes:

- Pay for what you use,
- Pay less when you reserve,
- Pay less when you use more, and
- Pay even less as AWS grows

Let's take a closer look at these core concepts of pricing.

Select the link to review the AWS pricing white paper.
https://d0.awsstatic.com/whitepapers/aws_pricing_overview.pdf

Pay for What You Use

Pay only for the services you consume, with no large upfront expenses.

- Lower variable costs
- Adapt to changing business needs

Benefits:

- Adapt to changing business needs
- Redirect focus on innovation and invention

Unless you build data centers for a living, you have likely spent too much time and money building them. With AWS, you pay only for the services you consume with no large upfront expenses. You're able to lower variable costs, so you no longer need to dedicate valuable resources to building costly infrastructure, including purchasing servers, software licenses, or leasing facilities.

Quickly adapt to changing business needs and redirect your focus on innovation and invention by paying only for what you use and for as long as you need it. All AWS services are available on demand, require no long-term contracts, and have no complex licensing dependencies.

Pay Less When You Reserve

Invest in Reserved Instances (RIs):

- Save up to 75%
- Options:
  - All Up-front Reserved Instance (AURI) → largest discount
  - Partial Up-front Reserved Instance (PURI) → lower discounts
  - No Upfront Payments Reserved Instance (NURI) → smaller discount

Benefits:

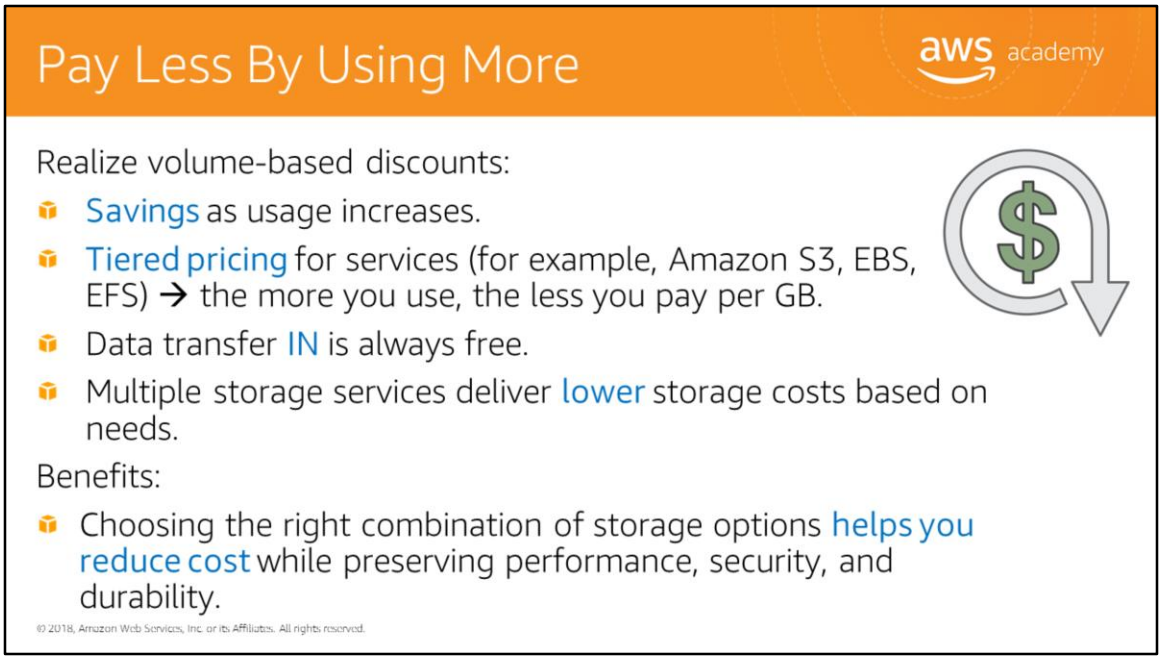
- Minimize risk
- Predictably manage budgets
- Comply with policies that require longer-term commitments

For certain services like Amazon EC2 and Amazon RDS, you can invest in reserved capacity. With Reserved Instances, you can save up to 75 percent over equivalent on-demand capacity. Reserved Instances are available in 3 options:

- All up-front reserved instance (or AURI)
- Partial up-front reserved instance (or PURI), and
- No upfront payments reserved instance (or NURI)

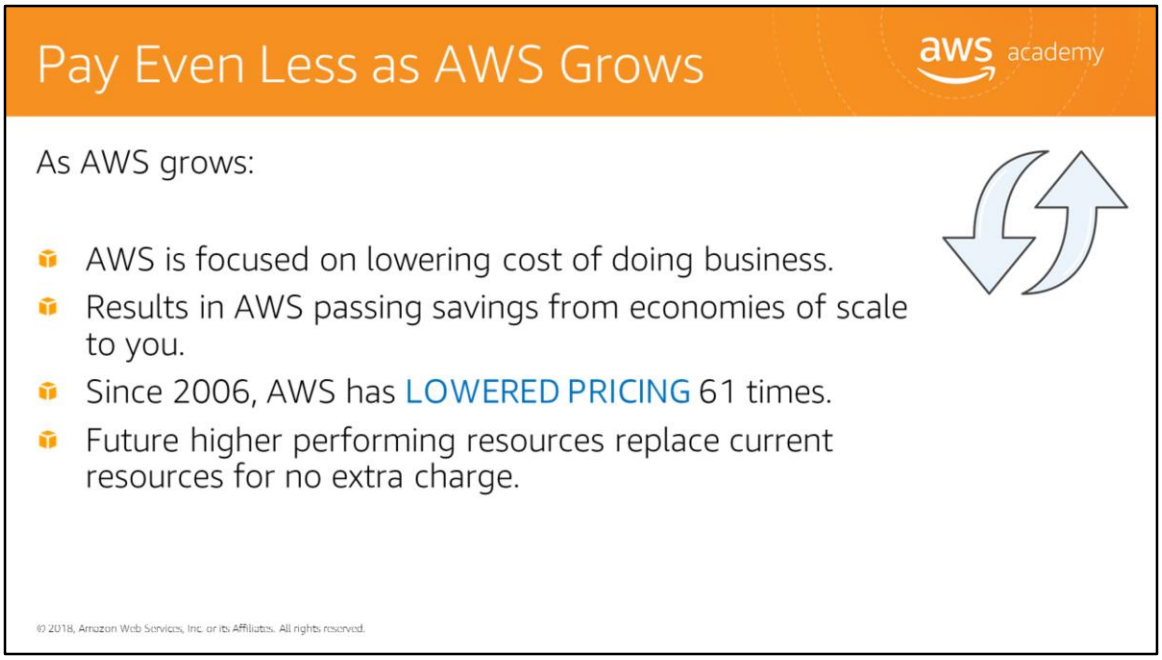
When you buy Reserved Instances, the larger the upfront payment, the greater the discount. To maximize your savings, you can pay all up-front and receive the largest discount. Partial up-front RIs offer lower discounts but give you the option to spend less up front. Lastly, you can choose to spend nothing up front and receive a smaller discount, allowing you to free up capital to spend in other projects.

By using reserved capacity, your organization can minimize risks, more predictably manage budgets, and comply with policies that require longer-term commitments.



With AWS, you can get volume based discounts and realize important savings as your usage increases. For services like Amazon S3, pricing is tiered, meaning the more you use, the less you pay per GB. In addition, data transfer IN is always free of charge. Multiple storage services deliver lower storage costs based on your needs. As a result, as your AWS usage needs increase, you benefit from the economies of scale that allow you to increase adoption and keep costs under control.

As your organization evolves, AWS also gives you options to acquire services that help you address your business needs. For example, the AWS storage services portfolio offers options to help you lower pricing based on how frequently you access data and the performance you need to retrieve it. To optimize your savings, choose the right combination of storage solutions that help you reduce costs while preserving performance, security and durability.



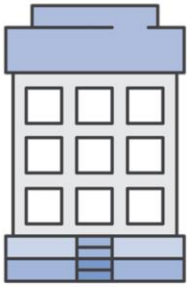
AWS is constantly focused on reducing data center hardware costs, improving operational efficiencies, lowering power consumption, and generally lowering the cost of doing business.

These optimizations and AWS's substantial and growing economies of scale result in passing savings back to you in the form of lower pricing. Since 2006, AWS has lowered pricing **61** times!

Another benefit of AWS growth is that future, higher performing resources replace current ones for no extra charge.

Custom Pricing

- Meet varying needs through custom pricing.
- Available for high-volume projects with unique requirements.



AWS realizes that every customer has different needs. If none of the AWS pricing models work for your project, custom pricing is available for high-volume projects with unique requirements.

AWS Free Tier

aws academy

AWS Free Tier helps customers get started in the cloud.

- Limitations:
  - Only new customers
  - Up to one year
  - Applicable to only certain services and options

For more details: www.aws.amazon.com/free

To help new AWS customers get started in the cloud, AWS offers a free usage tier for new customers up to one year, applicable to certain services and options. If you're a new AWS customer, you can run a free Amazon EC2 T1 micro instance for a year, while also leveraging a free usage tier for Amazon S3, Amazon Elastic Block Store, Elastic Load Balancing, AWS data transfer, and other AWS services.

Select the link for more details.
https://aws.amazon.com/free/

No Charge

AWS services for no additional charge:

- Amazon VPC
- AWS Identity and Access Management (IAM)
- Consolidated Billing
- AWS Elastic Beanstalk**
- AWS CloudFormation**
- Automatic Scaling**
- AWS OpsWorks**

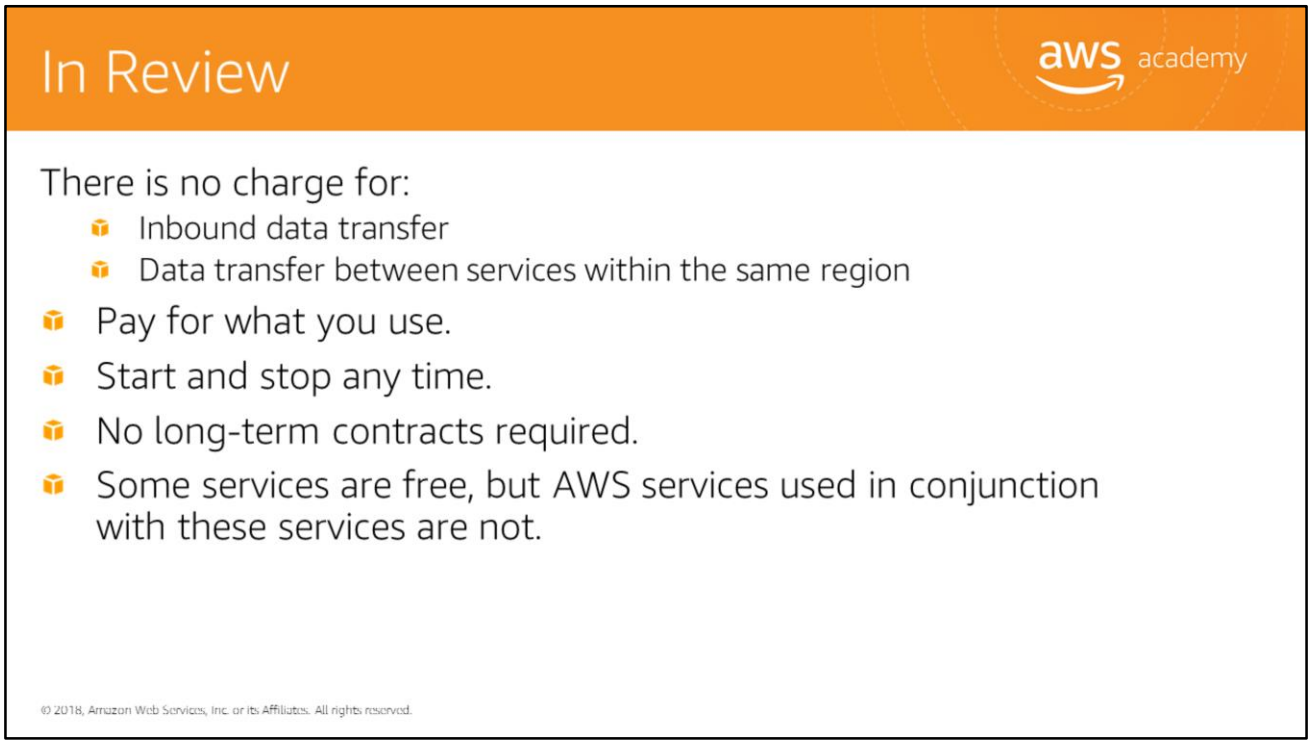
**Note: There may be charges associated with other AWS services used in conjunction with these services.

AWS also offers a variety of services for no additional charge.

- **Amazon VPC** lets you provision a logically isolated section of the AWS Cloud where you can launch AWS resources in a virtual network that you define.
- **AWS Identity and Access Management (or IAM)** controls your users' access to AWS services and resources.
- **Consolidated Billing** is a billing feature in AWS Organizations to consolidate payment for multiple AWS accounts or multiple AISPL accounts. Consolidated billing provides:
    - **One bill** for multiple accounts.
    - The ability to **easily track** each account's charges.
    - The opportunity to decrease charges as a result of volume pricing discounts from **combined usage**.
    - And you can consolidate all of your accounts using Consolidated Billing and get tiered benefits.
- **AWS Elastic Beanstalk** is an even easier way for you to quickly deploy and manage applications in the AWS cloud.
- **AWS CloudFormation** gives developers and systems administrators an easy way to create a collection of related AWS resources and provision them in an orderly and predictable fashion.
- **Automatic Scaling** automatically adds or removes resources according to conditions you define. The resources you are using increase seamlessly during demand spikes to maintain performance and decrease automatically during demand lulls to minimize costs.

- **AWS OpsWorks** is an application management service that makes it easy to deploy and operate applications of all shapes and sizes.

Note that although there is no charge for these services, there may be charges associated with other AWS services used in conjunction with these services. For example, when auto scaling additional EC2 instances, there will be charges for those instances.



In summary, while the number and types of services offered by AWS have increased dramatically, our philosophy on pricing has not changed. At the end of each month, you pay only for what you use, and you can start or stop using a product at any time. No long-term contracts are required.

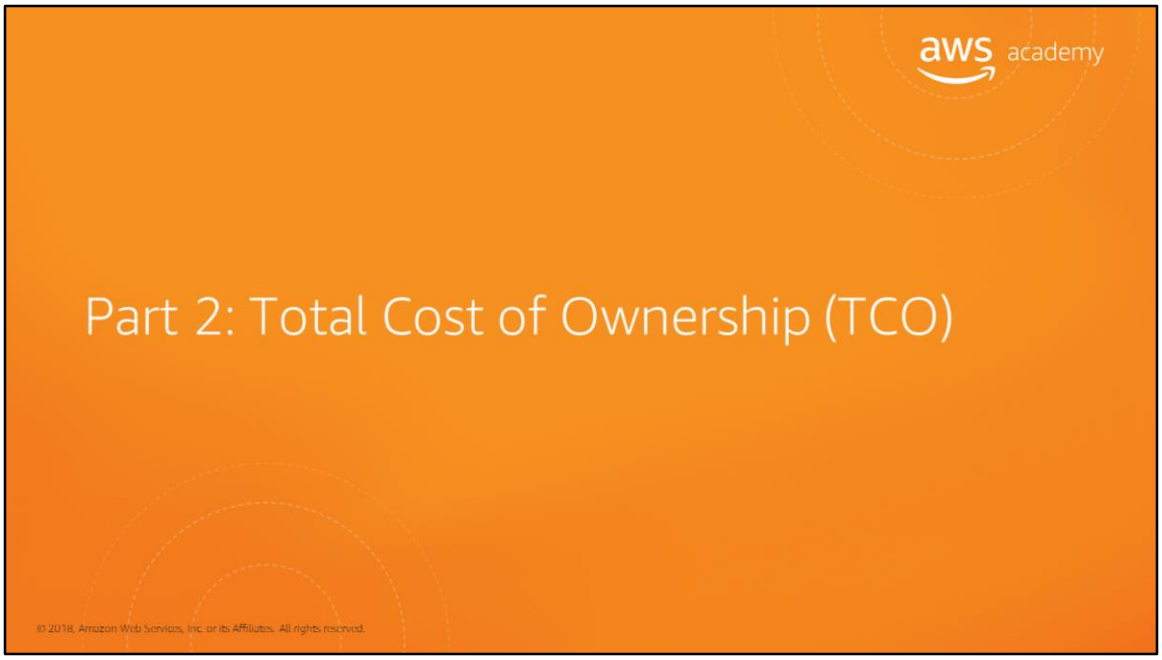
The best way to estimate costs is to examine the fundamental characteristics for each AWS service, estimate your usage for each characteristic, and then map that usage to the prices posted on the website. The service pricing strategy gives you tremendous flexibility to choose the services you need for each project and to pay only for what you use.

There are a number of free AWS services, including:
- Amazon VPC,
- Elastic Beanstalk,
- AWS CloudFormation,
- IAM,
- Automatic Scaling,
- AWS OpsWorks, and
- Consolidated Billing

While the services themselves are free, the resources that they provision are not. Additionally, there is no charge for inbound data or data transfer between services within the

same region; however, outbound data transfer costs are tiered.

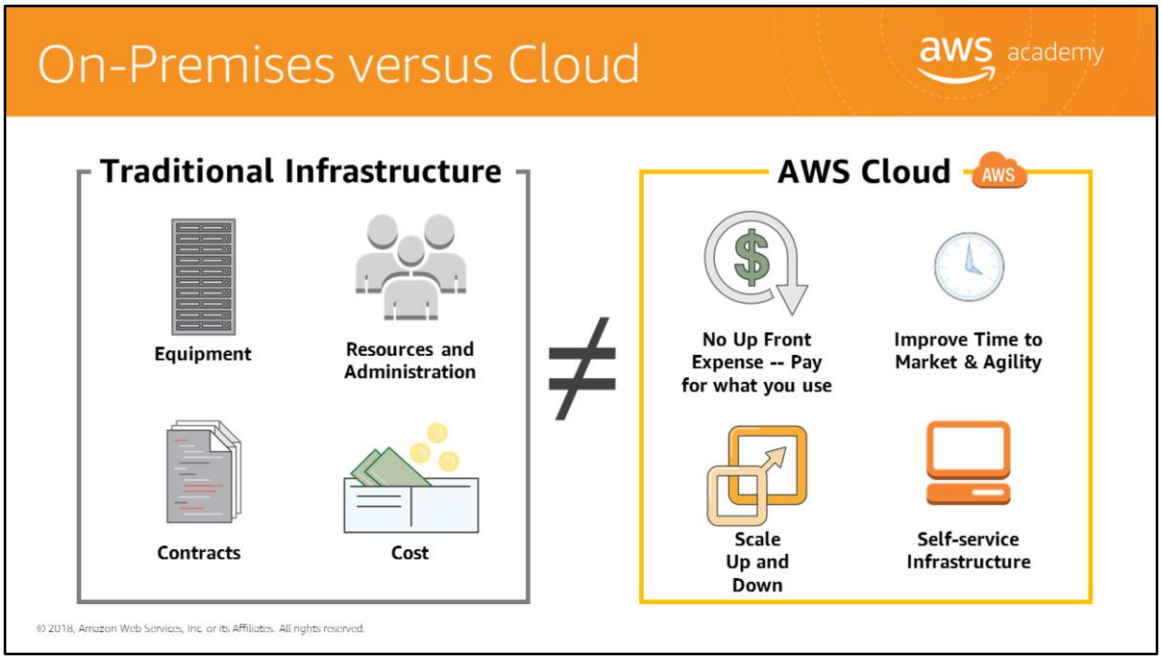
Select the links to learn more.
http://aws.amazon.com/pricing/.
https://d0.awsstatic.com/whitepapers/aws_pricing_overview.pdf.



Now that you understand the AWS pricing philosophy and how different AWS services are priced, let's take a look at Part 2: Total Cost of Ownership.

In addition to prices, businesses often want to understand the total cost of ownership, which is a financial estimate that helps buyers and owners determine the direct and indirect costs of a product or system. It reflects the purchase price of an asset *plus* the costs of operation. Total cost of ownership information is especially helpful when you are making the decision of whether or not to deploy on Amazon Web Services.

Let's move forward to the final part of our economic discussion and review both TCO and the TCO Calculator.



On-premises versus cloud is a question being asked by many businesses. The difference between these two options is how they are deployed.

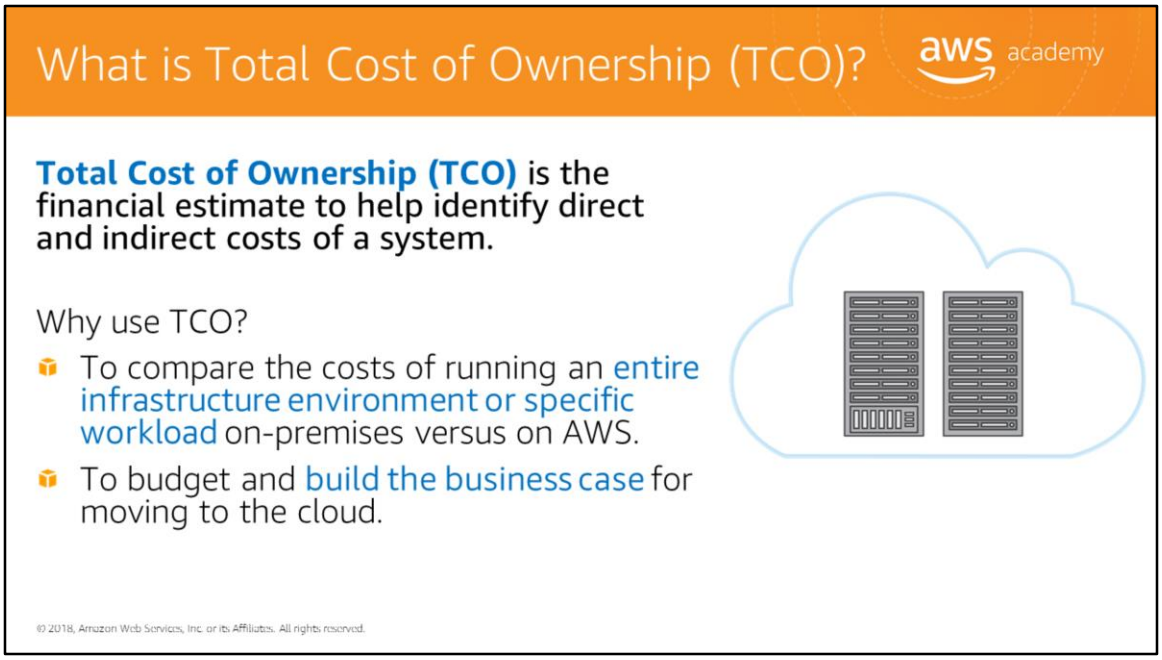
An on-premises infrastructure is installed locally on a company's own computers and servers. There are several fixed costs, also known as capital expenses, associated with the traditional infrastructure including facilities, hardware, licenses, and maintenance staff. Scaling up can be expensive and time consuming. Scaling down does not reduce fixed costs.

A cloud infrastructure is purchased from a service provider who builds and maintains the facilities, hardware, and maintenance staff. A customer pays for what is used. Scaling up or down is simple. Costs are easy to estimate because they depend on service usage.

It is difficult to compare an on-premises IT delivery model with the AWS cloud. The two are so very different that they use different languages.

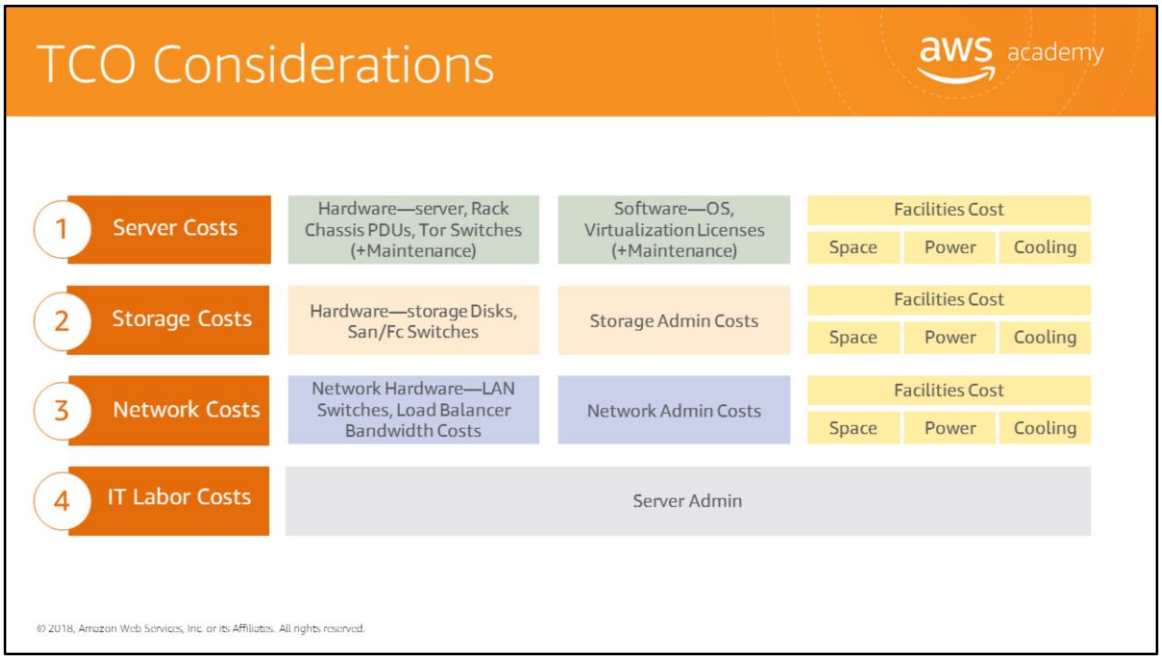
On-premises IT is a discussion based on capital expenditure, long planning cycles, and multiple components to buy, build, manage, and refresh over time.
AWS is a discussion about flexibility, agility, and consumption based costs.

So, how can we identify the best option?



We can identify the best option by comparing the on-premises solution to a cloud solution. Total Cost of Ownership (or TCO) is a tool that can be used for this comparison. TCO is a financial estimate intended to help buyers and owners determine the direct and indirect costs of a product or system. TCO includes the cost of a service plus all the costs associated with owning the service.

In the cloud environment, TCO is used for comparing the costs of running an entire infrastructure environment for a specific workload in an on-premises or co-location facility, to the same workload running on a cloud-based infrastructure. This comparison is done for budgeting purposes or to build a business case for business decisions regarding the optimal deployment solution.



So, what are some of the costs associated with data center management? These costs include:

- **Server** costs for both hardware and software, along and facilities costs to house the equipment.
- **Storage** costs are associated with the hardware, administration and facilities.
- **Network** costs are similar to the storage costs and include hardware, administration, and facilities costs.
- And **IT labor** costs that are required to administer the entire solution.

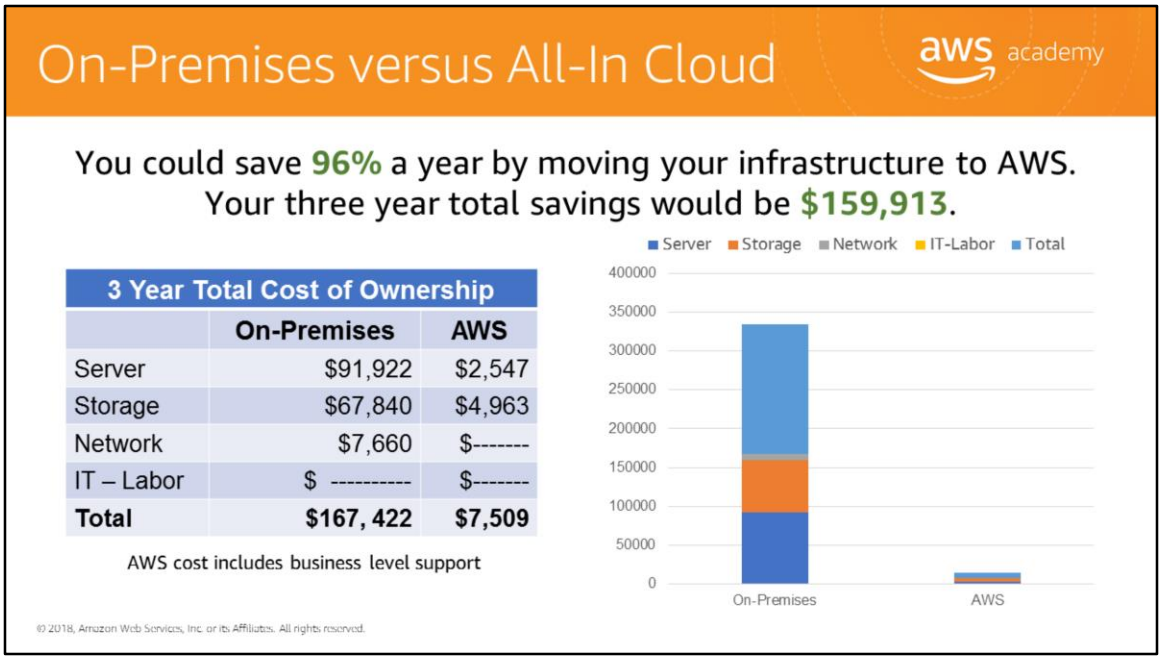
When comparing an on-premises to cloud solution, it is important to accurately assess the true costs of both options. With the Cloud, most costs are upfront and readily calculated. For example, cloud providers give transparent pricing based on different usage metrics, such as RAM, storage, and bandwidth, among others. Pricing is frequently fixed per unit of time.

Customers gain certainty over pricing and are then able to readily calculate costs based on several different usage estimates.

Compare this to on-premise technology. Although they are sometimes difficult to determine, calculations of in-house costs must take into account all:

- **Direct costs** that accompany running a server like power, floor space, storage, and IT operations to manage those resources
- **Indirect costs** of running a server like network and storage infrastructure

Please note that this diagram is conceptual and does not include every cost item. For example, depending on the solution you are implementing, software costs can include database, management, and middle-tier costs. Facilities costs can include upgrades, maintenance, building security, taxes, and so on. IT labor costs can include security admin and application admin costs. This is an abbreviated list to demonstrate the type of costs that are involved in data center maintenance.



Let's take a look at a cost comparison. This example shows a cost comparison for an on-premises and cloud solution over 3 years. For this comparison, two similar environments were constructed to represent the on-premises and AWS environments. Note that additional direct and indirect costs associated with the on-premises solution have not been included.
The components of the on-premises solution include:

- 1 virtual machine with 4 CPUs, 16 GB of RAM, and a Linux operating system
- Average utilization is 100%
- Optimized by RAM

The components of a comparable AWS environment include:

- 1 m4.xlarge instance with 4 CPUs, 16 GB of RAM,
- The Instance type is a 3 year partial upfront reserved instance

The cost difference between these solutions is significant - a 96% annual savings on cloud infrastructure versus an on-premises solution. The three year total savings would be $159,913!

The on-premise three year total solution cost is $502,266. The AWS Cloud three year total solution cost is $22,537 for a 96% savings. This comparison helps a business clearly understand the differences between the alternatives.

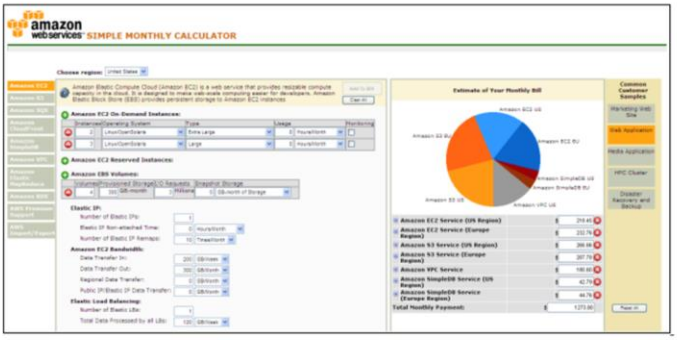
So, what is the difference in the costs?
Remember, the on-premises solution is “predicted” and then continues to incur costs whether or not the capacity is utilized.

In contrast, the AWS solution is commissioned when needed and decommissioned when the resources are no longer in use, resulting in the lower overall costs.

AWS Simple Monthly Calculator

Use the Simple Monthly Calculator to:

- Estimate monthly costs
- Identify opportunities to reduce monthly costs
- Use templates to compare services and deployment models



Access the Simple Monthly Calculator. http://calculator.s3.amazonaws.com/index.html

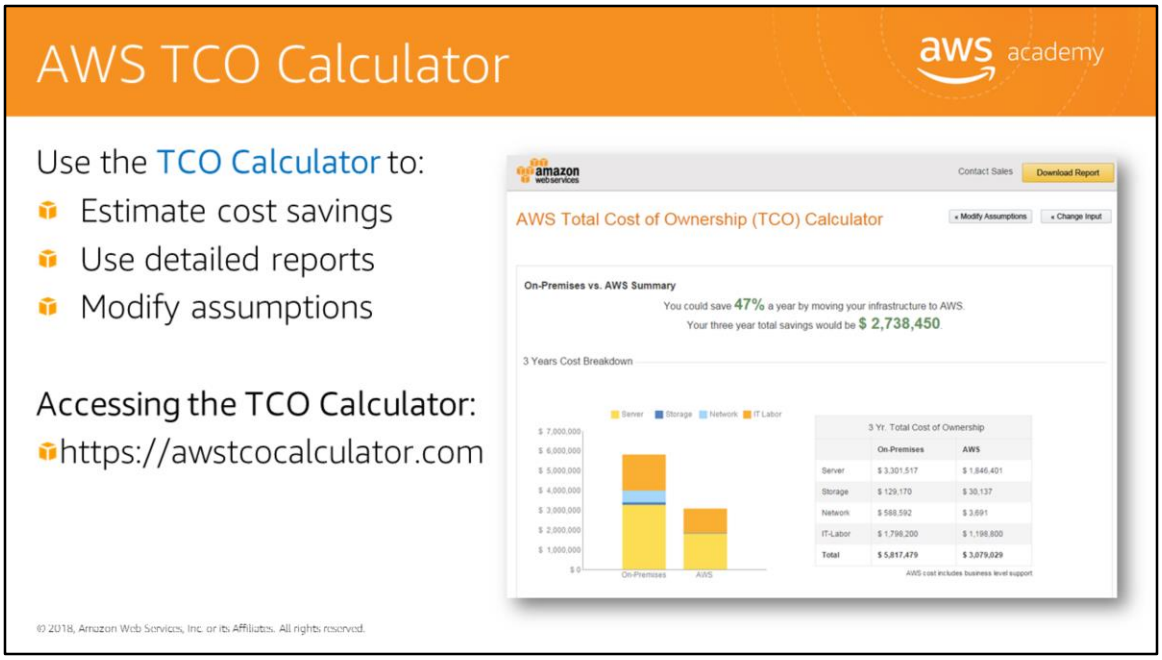
There are tools available to assist you with these comparisons. The **AWS Simple Monthly Calculator** helps estimate a monthly AWS bill. Using this tool, you can add, modify and remove services from your 'bill' and it will recalculate the estimated monthly charges automatically.

The calculator incorporates a wide array of pricing calculations across all services in all regions. It also shows a breakdown of features for each service in each region.

The **Simple Monthly Calculator** is a tool that helps you:
Estimate monthly services costs when using AWS
Identify opportunities for cost reduction
And use templates to model solutions to compare services and deployment models

The calculator also shows common customer samples and their usage. You can click on the “Disaster Recovery and Backup” sample or “Web Application” sample and see the uses of each service.

Select the link to launch the Simple Monthly calculator.



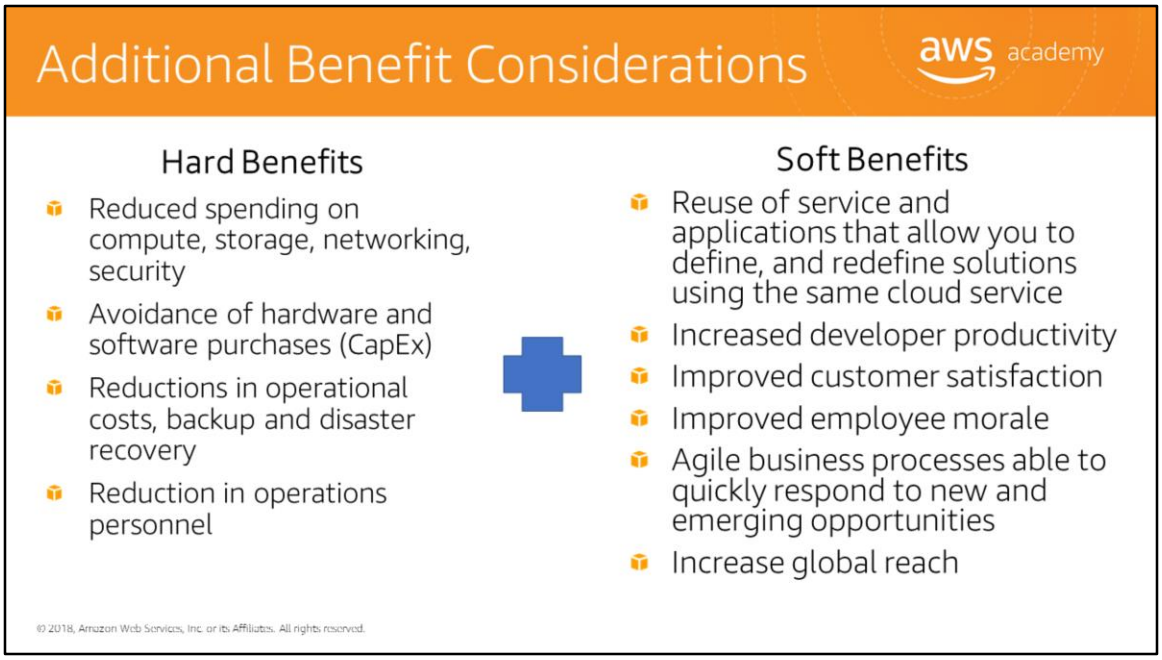
The **AWS TCO Calculator** helps you evaluate the total cost of ownership of a solution. Eliminating the need to invest in large capital expenditures, or **capex**, and providing a pay-as-you-go model that empowers you to invest in the capacity you need, and use it only when you require it, helps reduce total cost of ownership.

The TCO calculator is a tool that helps you:

- Estimate cost savings when using AWS
- Uses a detailed set of reports, which can be used in executive presentations.
- And modify assumptions that best meets your needs.

An additional benefit of the calculator includes the ability to weigh the financial considerations of owning and operating a data center versus using a cloud infrastructure. Also, the TCO calculator explains the assumptions and the methodology behind the calculations.

Select the link to launch the TCO calculator.
https://awstcocalculator.com



Hard benefits include reduced spending on compute, storage, networking and security. Avoid hardware and software purchases and reductions in operational costs, backup and disaster recovery and a reduction in operations personnel.

Cloud Total Cost of Ownership defines what will be spent on the technology after adoption - or what it costs to "run the engine". Typically, a TCO analysis looks at the "as is" on-premises infrastructure and compares this with the cost of the "to be" infrastructure state in the cloud. While this is easy to calculate, it may only provide a narrow view of the total financial impact of moving to the cloud.

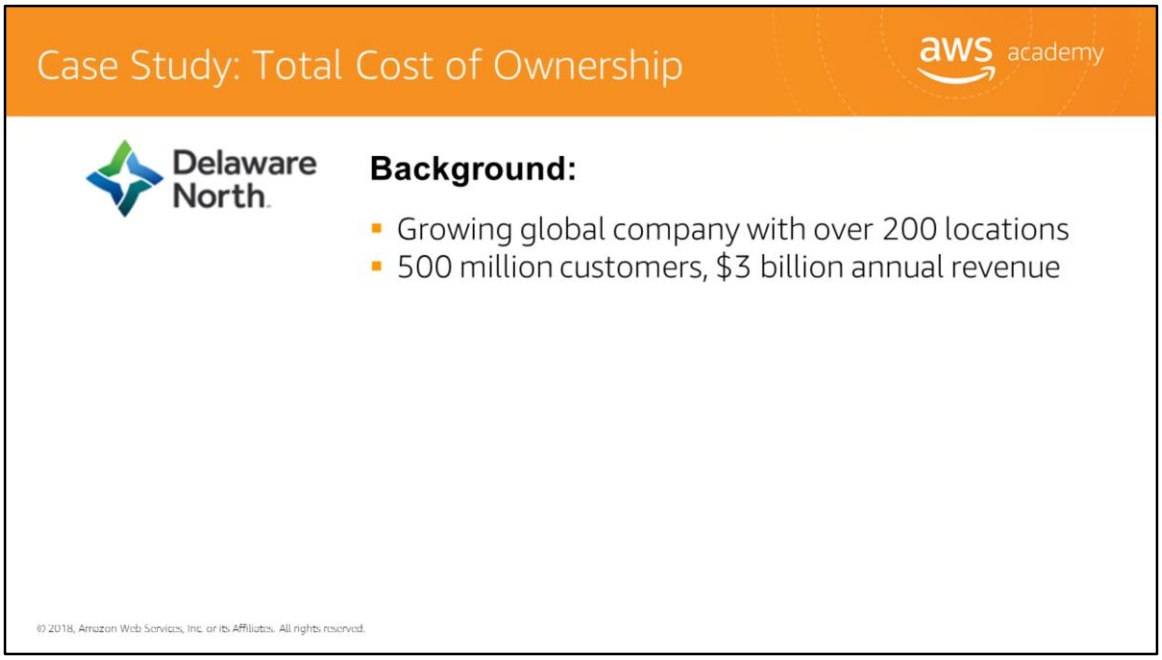
A **Return on Investment** analysis can be used to determine the value generated while taking spending and saving into consideration. This analysis starts by identifying the hard benefits in terms of direct and visible cost reductions and efficiency improvements.

Next, **soft savings** are identified. Soft savings are value points that are challenging to accurately quantify but can be more valuable than the hard savings. It is important for you to understand both hard and soft benefits to understand the full value of cloud. Soft benefits include:

- Reusing service and applications that allow you to define, and redefine solutions using the same cloud service

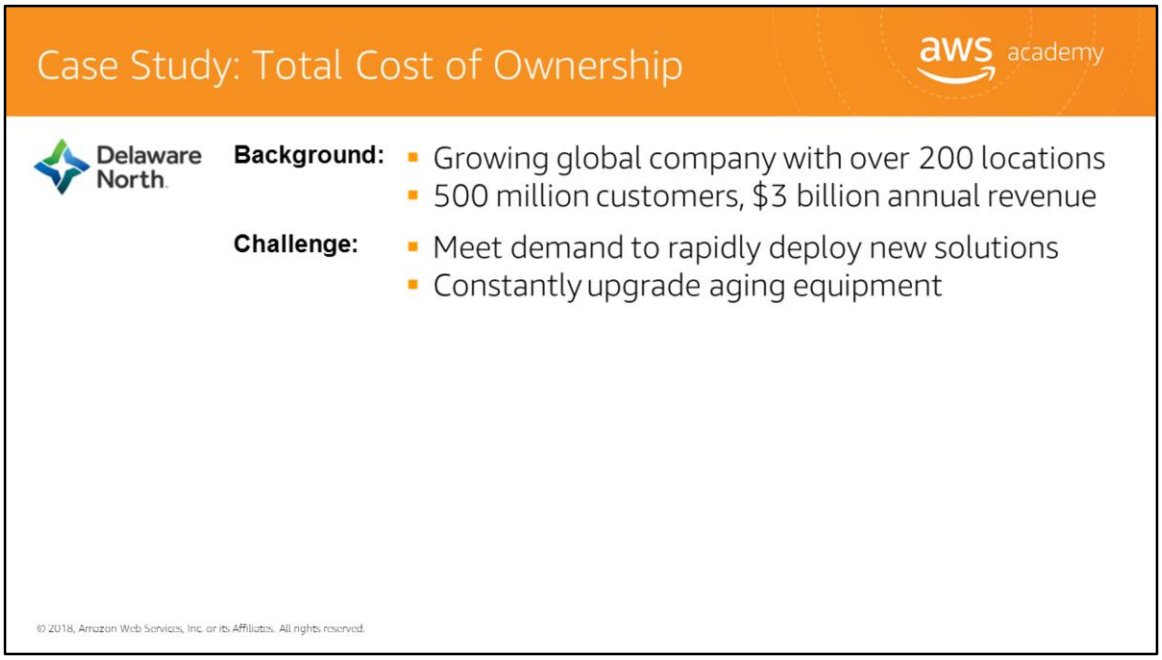
- Increased developer productivity
- Improved customer satisfaction
- Improved employee morale
- Agile business processes able to quickly respond to new and emerging opportunities, and
- Increased global reach

Now, let's review a case study from Delaware North to see an actual TCO example.



Background:
Delaware North originated in 1915 as a peanut and popcorn concessions vendor; today, it's a major food and hospitality company. Although the company deliberately keeps a low profile, it is a leader in the food-service and hospitality industry.

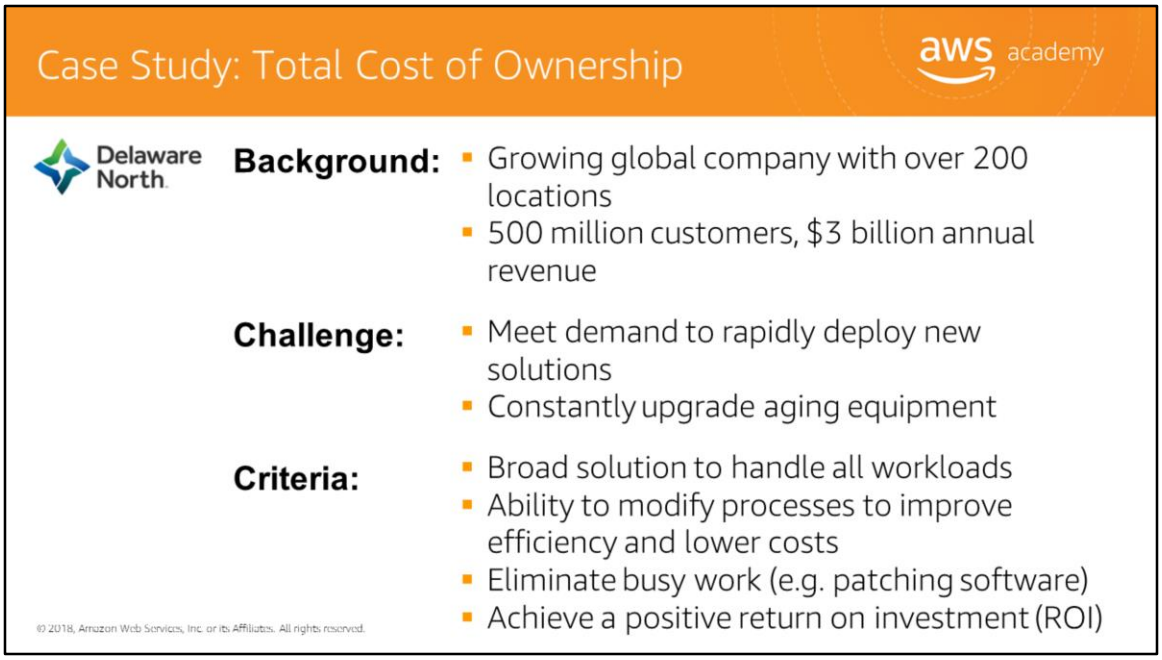
Delaware North serves more than **500 million customers** annually at more than **200 locations** around the world, including venues as diverse as the Kennedy Space Center in Florida, London Heathrow Airport, Kings Canyon Resort in Australia, and the Green Bay Packers' Lambeau Field in Wisconsin. This global presence has turned Delaware North into a **$3 billion enterprise**.



The company's on-premises data center was becoming too expensive and inefficient to support its global business operations.

Kevin Quinlivan, Delaware North's Chief Information Officer, says, "As the company continued to grow, the **demand to rapidly deploy new solutions** to meet customer requirements increased as well. This fact, combined with the **need to constantly upgrade aging equipment**, required an even greater commitment of resources on our part. We had to find a better strategy."

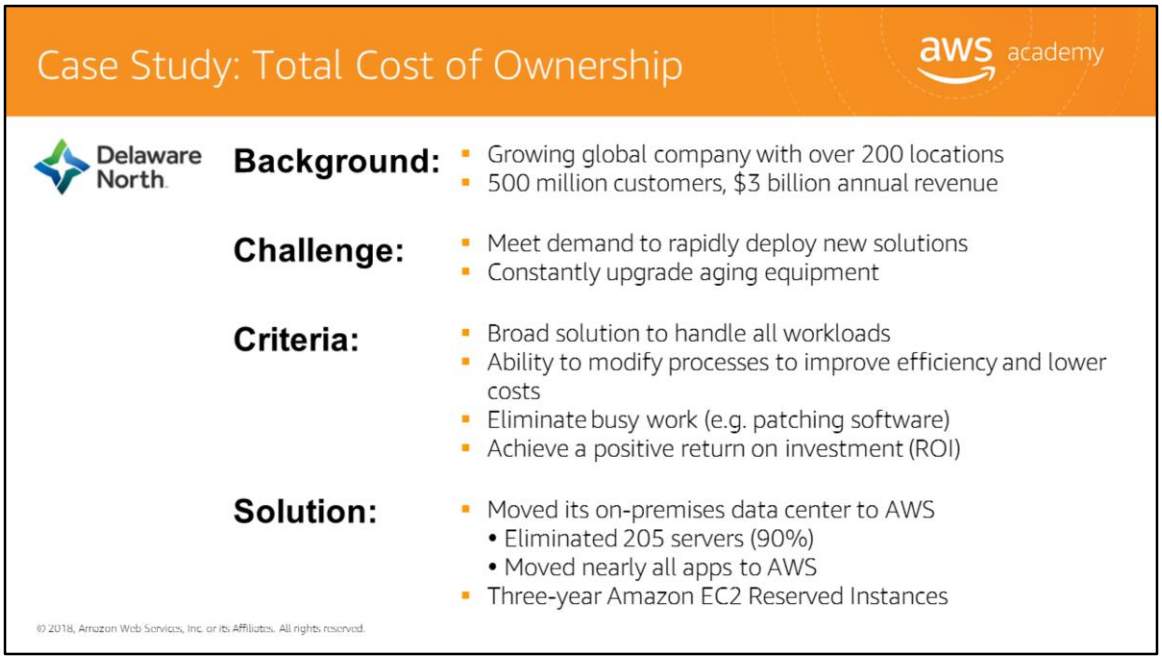
Delaware North turned to AWS for a solution.



After a successful migration of about 50 websites to AWS in 2013, Delaware North evaluated the cost benefit and total cost of ownership to move their IT infrastructure to AWS. Their focus was to answer C-Suite level business demands for measurable benefits that could convince an executive committee that the AWS cloud was the right approach.

The evaluation process centered on three criteria:

- First, a cloud solution needed a broad set of technologies that could **handle all of Delaware North's enterprise workloads** while delivering support for critical functions.
- From an operational perspective, Delaware North wanted the features and flexibility to **modify core IT processes to improve efficiencies and lower costs**. This included **eliminating redundant or time-consuming tasks** like patching software and pushing test and development tasks through outdated systems that, in the past, added months to the deployment of new services.
- Finally, financial requirements needed to **demonstrate a return on investment** with a solid cost-benefit justification for moving away from their existing data center environment.



A cost comparison completed by Delaware North demonstrated that it could save 3.5 million dollars based on a five-year run rate by **moving its on-premises data center to AWS** and using three-year Amazon EC2 Reserved Instances and Reserved Instance renewals.

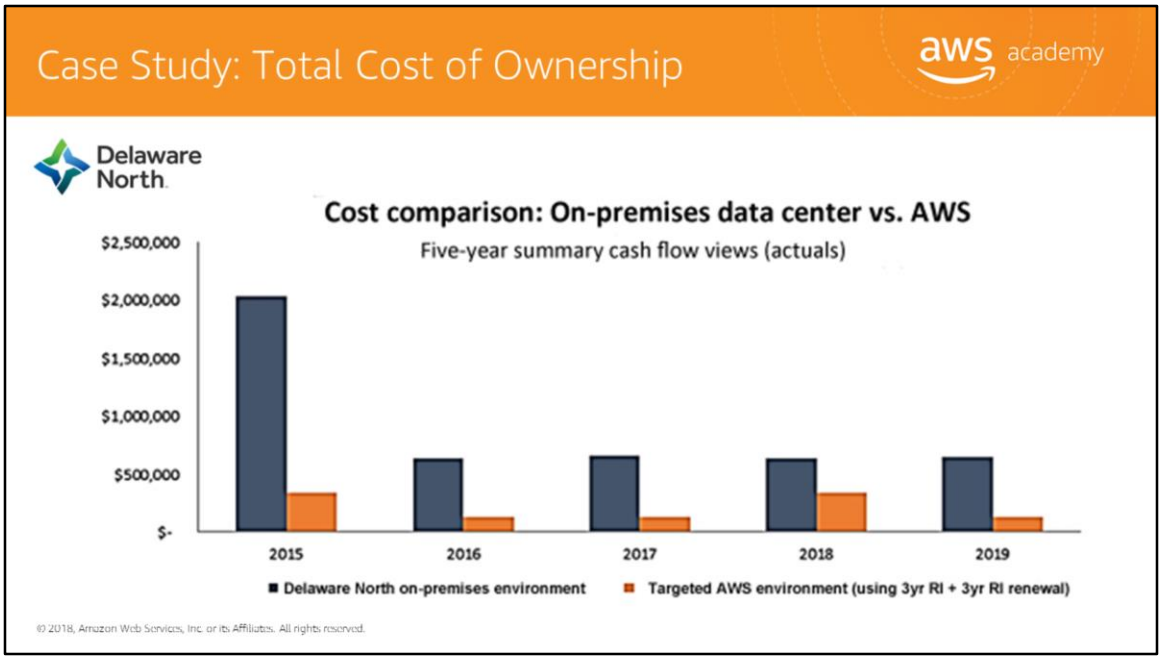
Quinlivan noted that the deep technology stack available on AWS was more than sufficient to meet the company's technical and operational requirements. The pricing structure of the AWS offerings, which includes paying only for what is used, provided total cost of ownership benefits which was presented to senior leaders.

Quinlivan stated, "We compared the costs of keeping our on-premises data center versus moving to the AWS cloud, measuring basic infrastructure items such as hardware cost and maintenance." He also says "We estimate that moving to AWS will save us at least $3.5 million over five years by **reducing our server hardware by more than 90 percent**. However, the cost savings will likely be greater due to additional benefits, like the increased compute capacity we can get using AWS. That lets us continually add more and larger workloads than we could using a traditional data center infrastructure, and achieve savings by only paying for what we use."

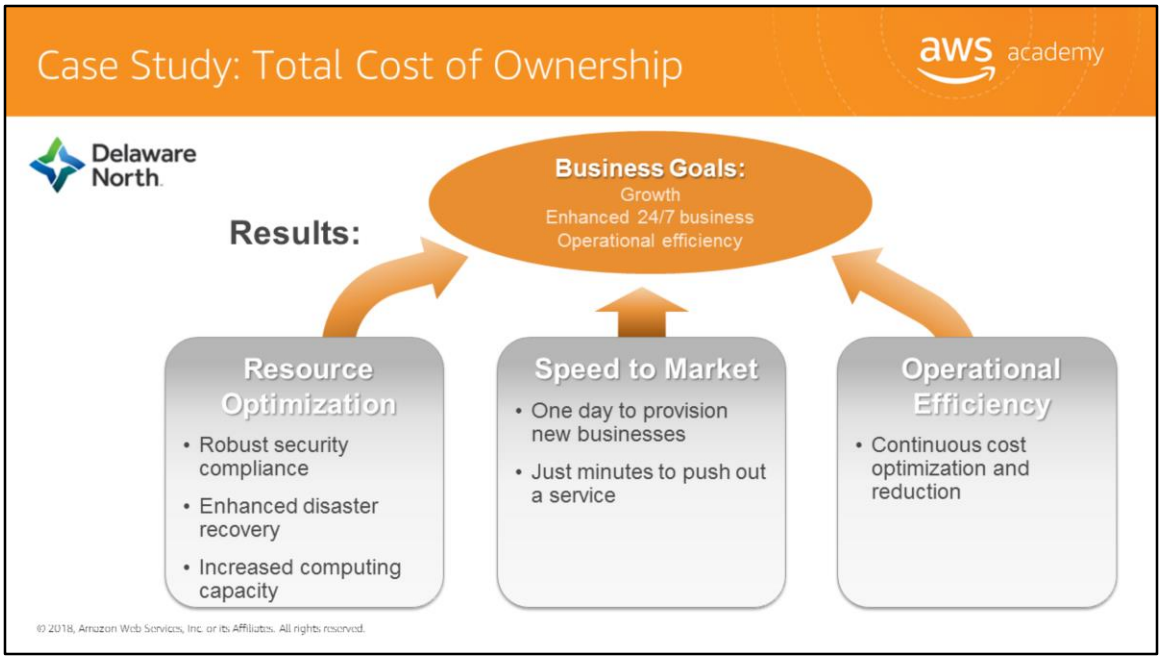
Delaware North moved almost all of its applications to AWS, including enterprise software such as its Fiorano middleware, Crystal Reports and QLIK business intelligence solutions, its Citrix virtual desktop system, and Microsoft System Center Configuration Manager, which is used to manage workstations.

The most dramatic physical change was the **elimination of 205 servers.** Everything running on that hardware was migrated to AWS. The IT department decided to keep about 20 servers on-premises at the new headquarters building to run communications and file-and-print tasks.

“We erred on the side of caution to ensure there is no latency with these tasks, but once we reach a certain comfort level, we may move these to the cloud as well,” says Brian Mercer, Senior Software Architect for the project.



This chart displays the cost comparison done by Delaware North showing the costs of their on-premises environment and the proposed AWS environment. The estimates showed a $3.5 million based on a five-year run rate by moving from an on-premises data center to AWS.



Six months into its cloud migration, Delaware North was realizing benefits in addition to its data center consolidation, including cost-effective security compliance, enhanced disaster recovery, and faster deployment times for new services.

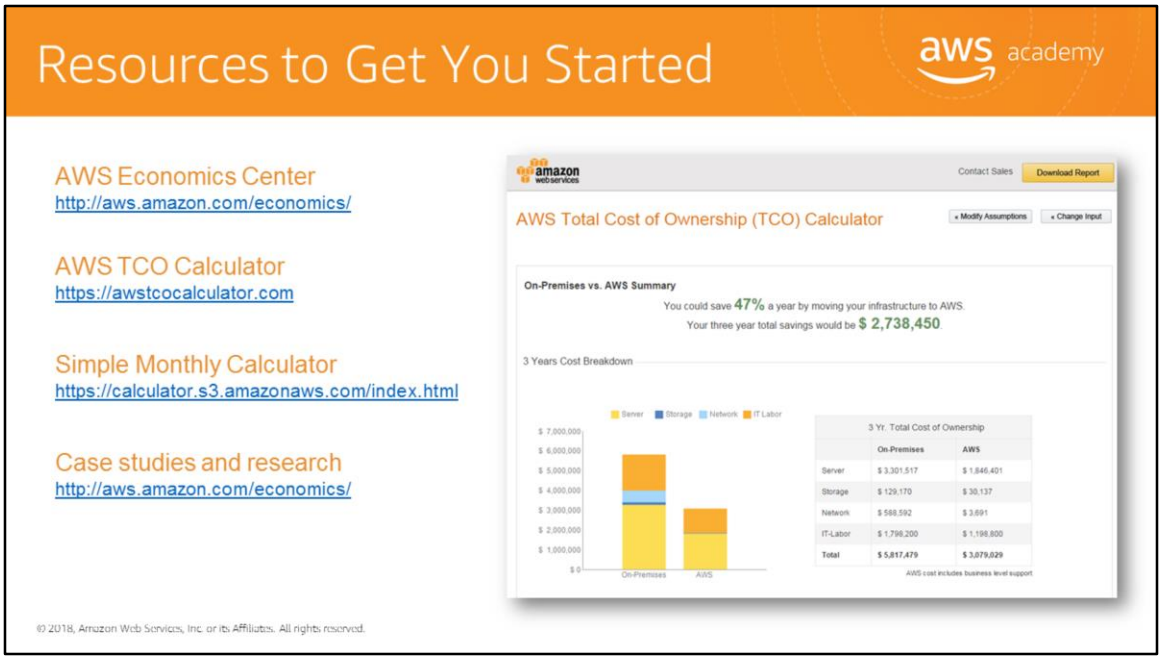
“Robust security in a retail environment is critical for us because of our many retail operations, and AWS is enormously helpful for that,” says Brian Mercer, the senior software architect for the project. “By leveraging the security best practices of AWS, we’ve been able to eliminate a lot of compliance tasks that in the past took up valuable time and money.”

He adds that the company also has increased its disaster recovery capabilities at a lower cost than what was available in its previous data center deployment. “It significantly improved our business continuity capabilities, including seamless failovers,” he says.

The solution is also helping Delaware North operate with greater speed and agility. For example, it can bring in new businesses - either through contracts or acquisitions - and get them online much faster than in the past by eliminating the need for traditional IT procurement and provisioning. It used to take between two and three weeks to provision new business units; now it takes one day. The Delaware North IT team is also using AWS to overhaul its operations by eliminating outdated and cumbersome processes, cleaning up documentation, and leveraging the benefits of running test and development tasks in

combination with rapid deployment of services through the cloud.

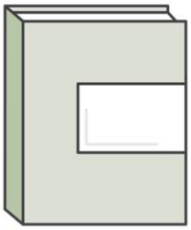
“Our DevOps team can now spin up the resources to push out a service in just minutes, compared to the weeks it used to take,” says Brian Mercer. “With AWS, we can respond much faster to business needs. And we can start repurposing time and resources to deliver more value and services to our internal teams and to our customers.”



If you are interested in learning more, there are a number of resources available. Here are some links for you to look at later. Try out the simple online calculator, and access additional resources.

In Review

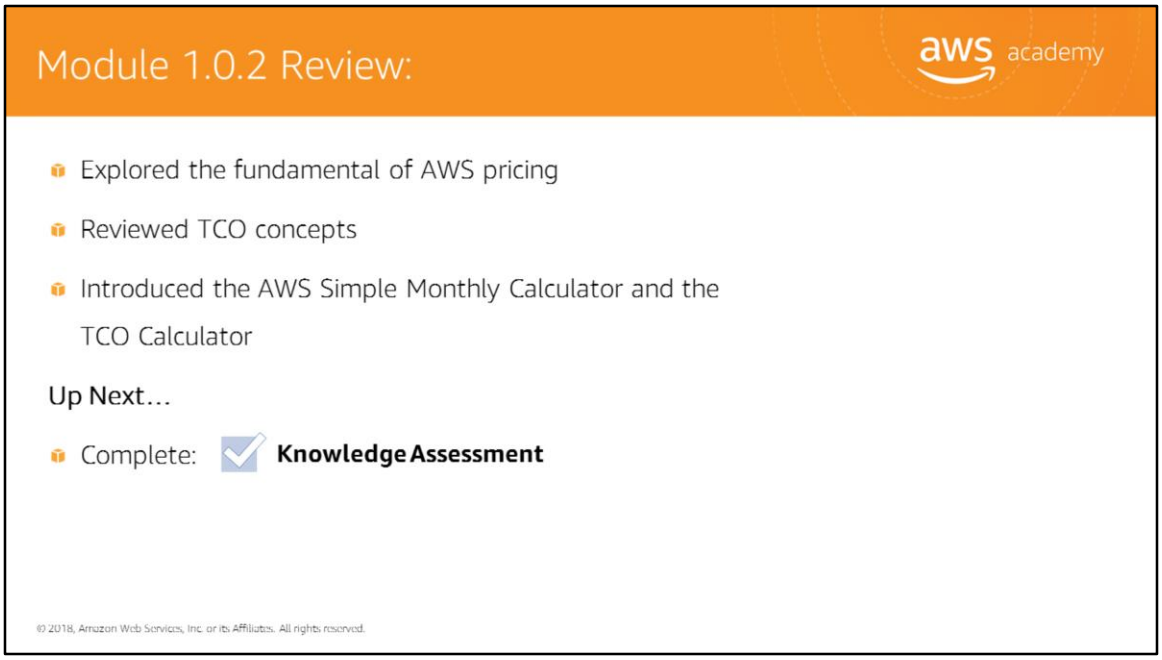
- TCO is a valuable tool.
- The AWS Simple Monthly Calculator can be used to provide accurate cost estimates.
- The AWS TCO Calculator can be used to estimate cost savings.
- The TCO Calculator:
  - Detailed reports with 3-year TCO comparison by cost categories
  - Executive reports
  - Assumptions that can be tailored to business needs



In summary, the total cost of ownership is a valuable tool that can be used to understand and compare the costs associated to different deployments. AWS provides the AWS Simple Monthly Calculator and the TCO Calculator to assist you with the calculations needed to estimate cost savings.

Use the AWS **Simple Monthly Calculator** to:
- Estimate monthly costs
- Identify opportunities to reduce monthly costs, and
- Use templates to compare services and deployment models

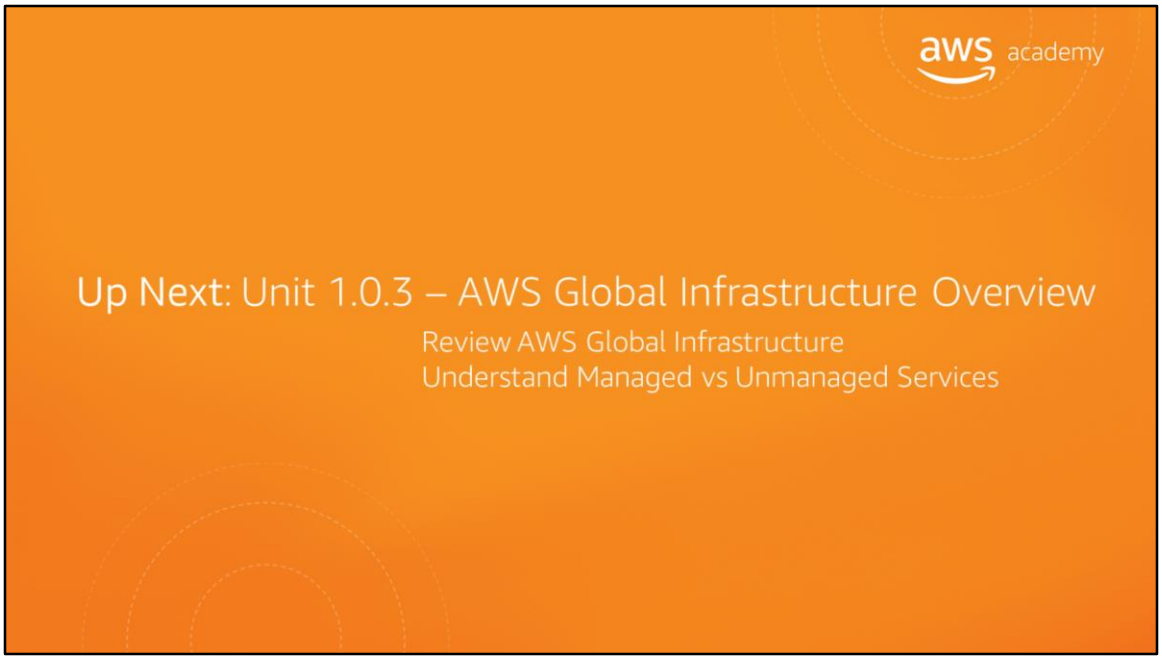
Use the **TCO Calculator to**:
- Analyze detailed reports that show a 3-year TCO comparison by cost categories
- Reports that are appropriate for inclusion in executive presentations, and
- The ability to modify assumptions for business needs



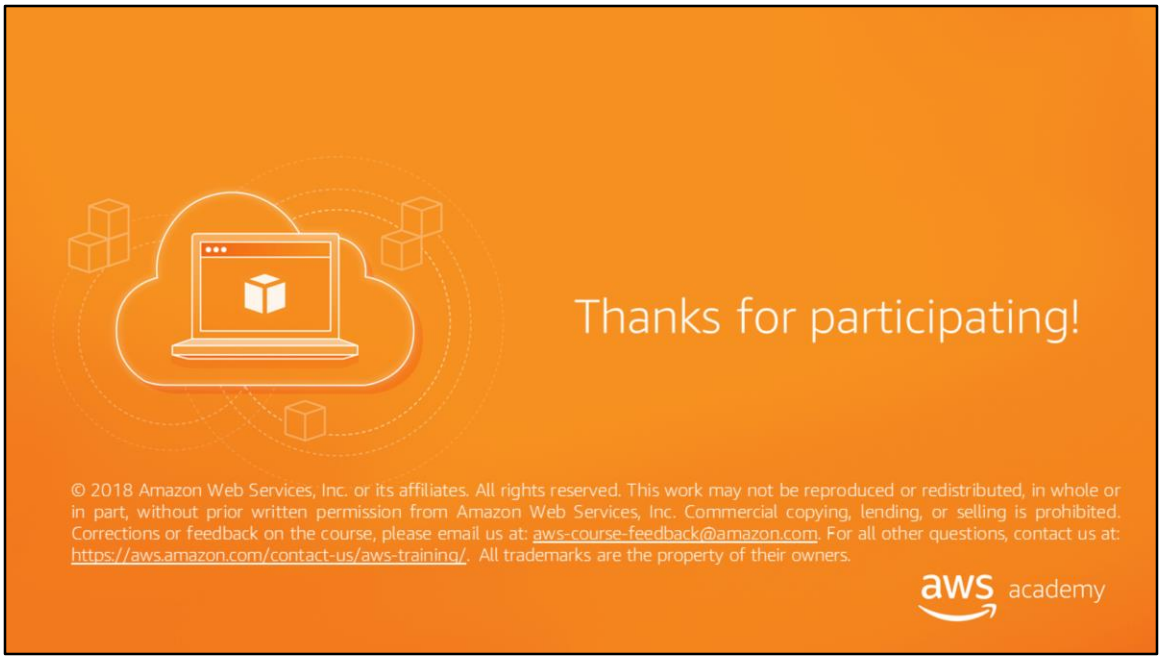
In review, we:

- Explored the fundamental of AWS pricing
- Reviewed the Total Cost of Ownership concepts, and
- Introduced the AWS Simple Monthly Calculator and the AWS TCO calculator.

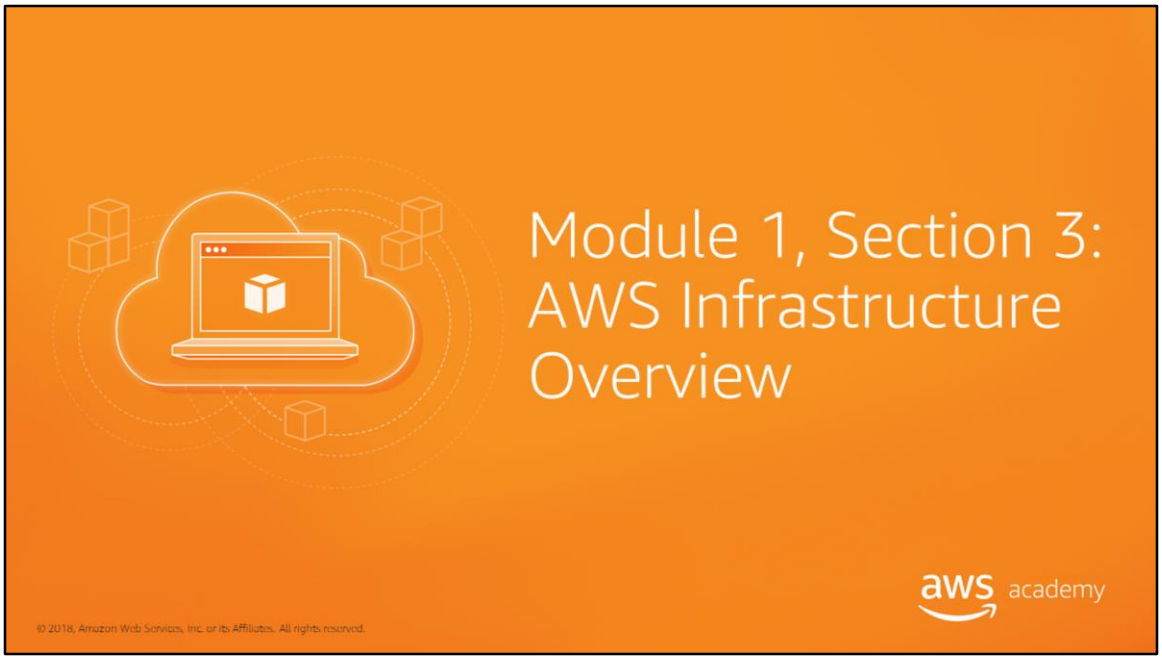
To complete this module, complete the knowledge assessment.



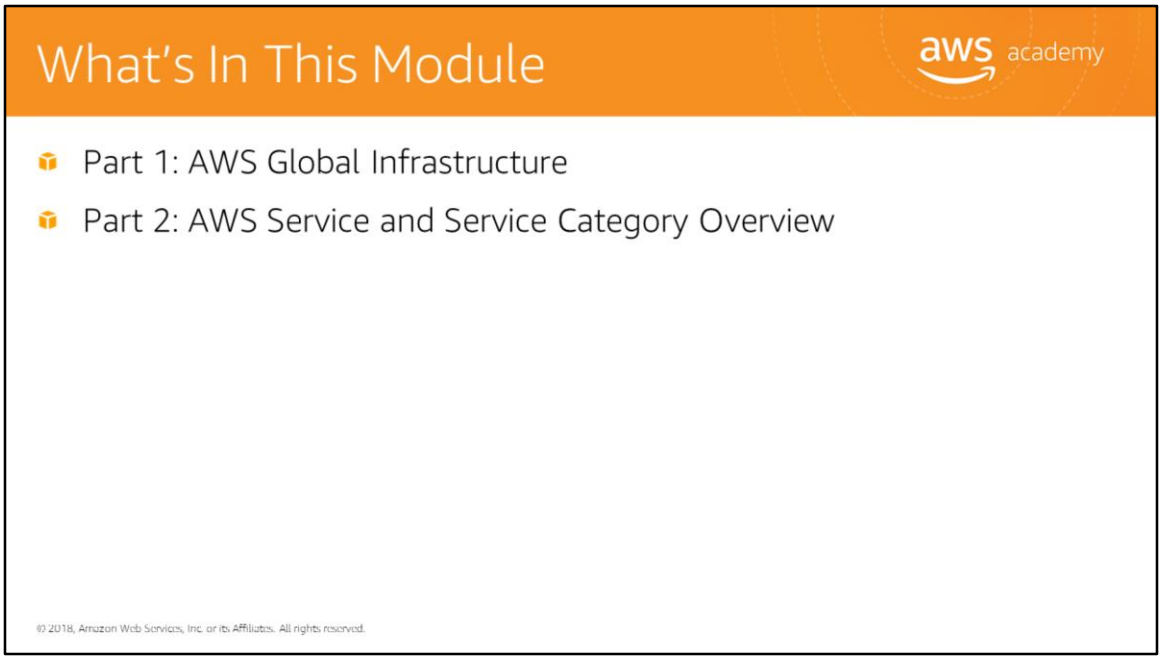
Next, we'll review the AWS Global Infrastructure and understand managed versus unmanaged services.



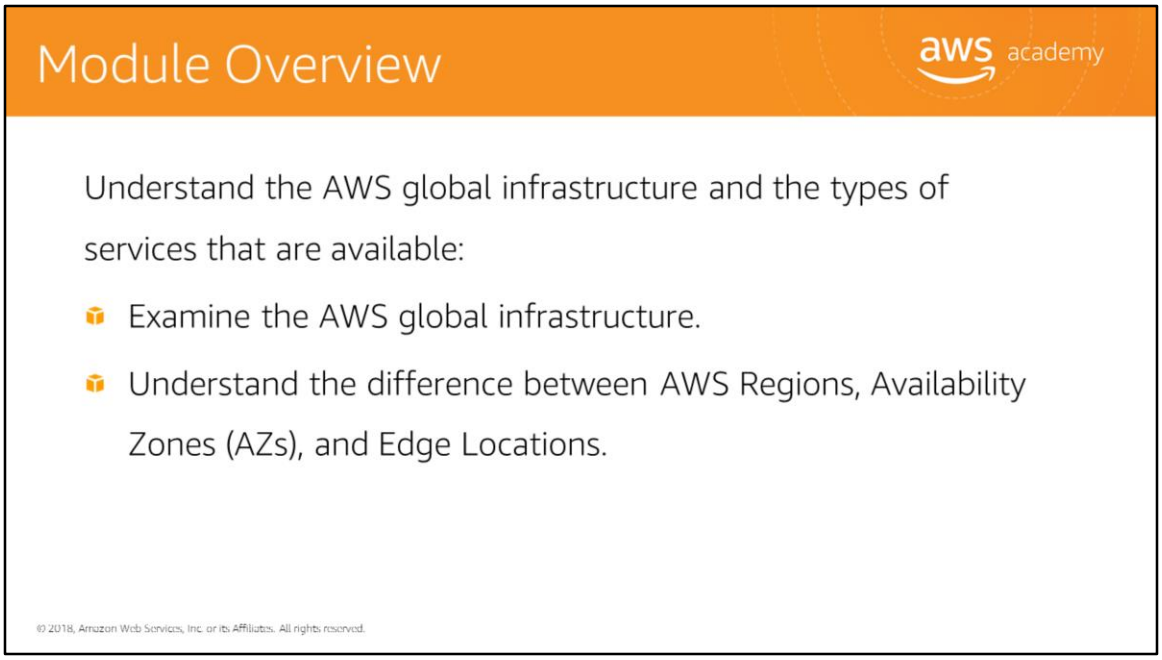
Thanks for participating! You may now exit this module.



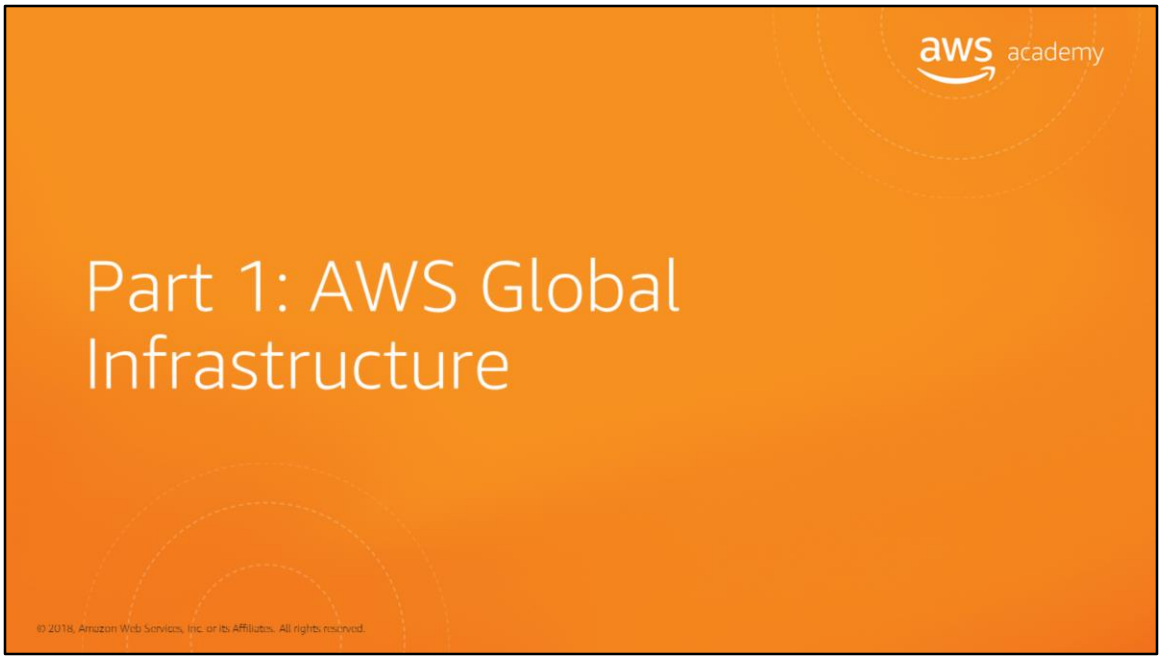
Welcome to Module 1, Section 3: AWS Infrastructure Overview.



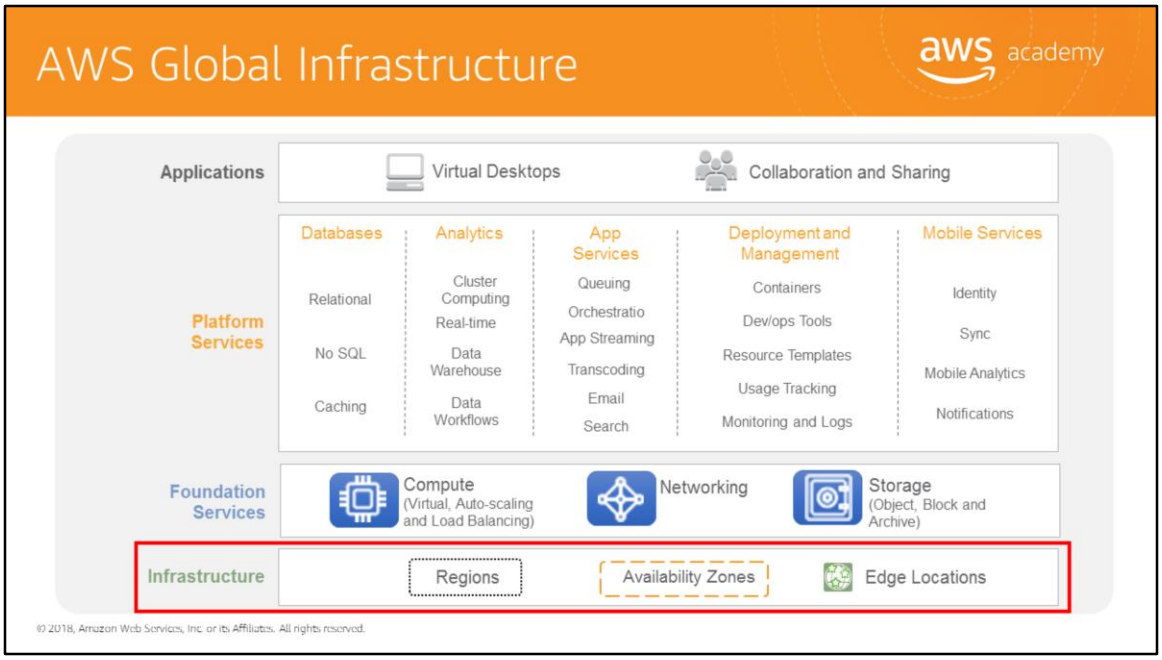
In part one, we will review the AWS Global Infrastructure, and in part two we'll provide an AWS service and service category overview.



The goal of this module is to understand the AWS global infrastructure and the types of services that are available. We'll examine the AWS Global Infrastructure to gain a clear understanding of what the infrastructure includes and understand the differences between AWS Regions, Availability Zones, and Edge Locations.



Introducing Part 1: AWS Global Infrastructure.



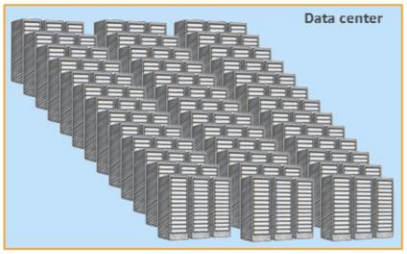
As we discussed earlier, AWS provides a broad set of services, such as compute power, storage options, networking, and databases, delivered as an on-demand utility that is available in seconds, with pay-as-you-go pricing. All of these services reside on the AWS global infrastructure.

AWS's global infrastructure can be broken down into three elements: Regions, Availability Zones, and Edge Locations.

Let's take a more in depth look at the AWS infrastructure and see what these are.

AWS Data Centers

- Data centers are securely designed
- A datacenter is a location where actual physical data resides
- A data center typically has 50,000 to 80,000 physical servers
- All data centers are online. No data center is "cold"
- AWS custom network equipment:
  - Multi-ODM sourced
  - Amazon custom network protocol stack



The foundation for the AWS infrastructure are the data centers. A data center is a location where the actual physical data resides. AWS data centers are built in clusters in various global regions.

Data centers are securely designed with several factors in mind.

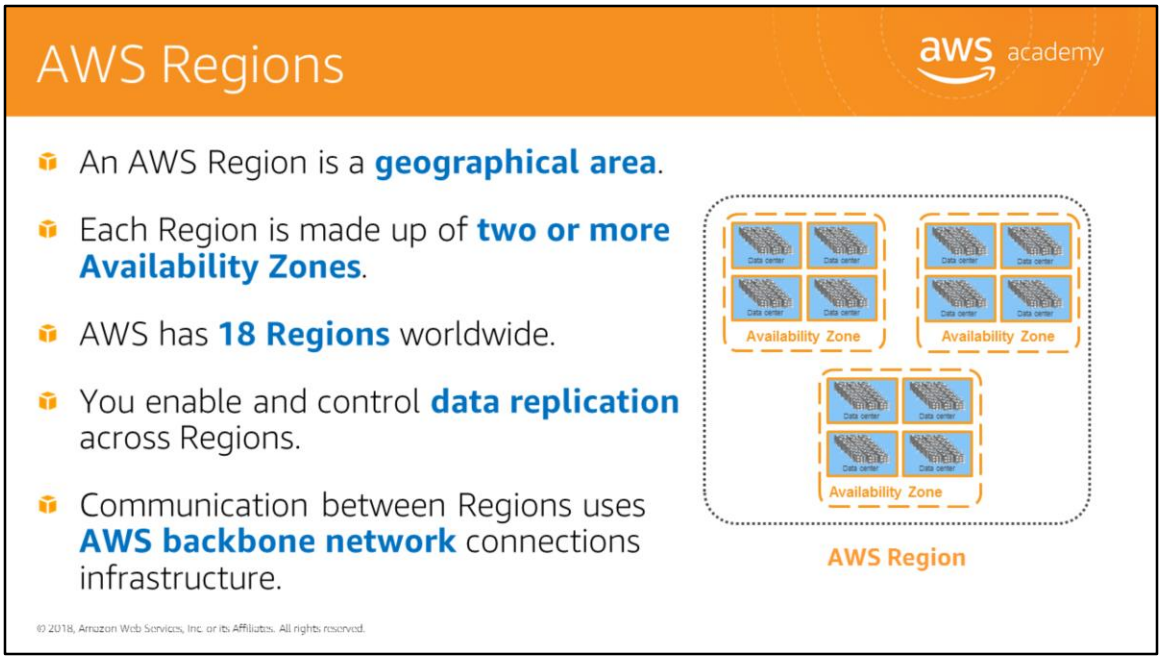
Each location is carefully evaluated to mitigate environmental risk.
Data centers have a redundant design that anticipates and tolerates failure while maintaining service levels.
To ensure availability, critical system components are backed up across multiple isolated locations known as Availability Zones.
To ensure capacity, AWS continuously monitors service usage to deploy infrastructure to support availability commitments and requirements.
Data center locations are not disclosed and all access to them is restricted.
In case of failure, automated processes move customer data traffic away from the affected area.

A single data center typically houses 50,000 to 80,000 physical servers, as larger data centers are undesirable.

All data centers are online and serving customers, so no data center is “cold.”

AWS utilizes custom, multi-ODM sourced network equipment. Original Design Manufacturer (or ODM) designs and manufactures products based on specifications from a second company. The second company then rebrands the products for sale.

Select the link to learn more.
https://aws.amazon.com/compliance/data-center/.



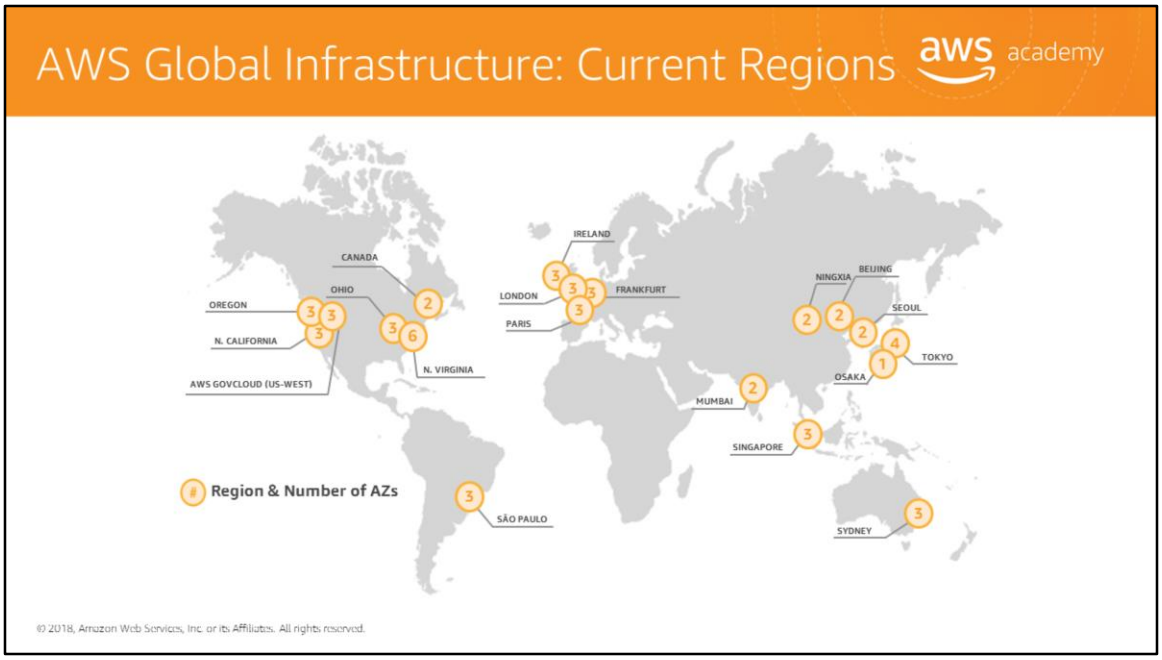
The AWS Cloud infrastructure is built around Regions and Availability Zones.

An AWS Region is a physical geographical location in the world where we have multiple Availability Zones. To achieve fault tolerance and stability, Regions are isolated from one another. Resources in one region are not automatically replicated to other regions. Each AWS Region contains two or more Availability Zones. AWS has 18 regions worldwide.

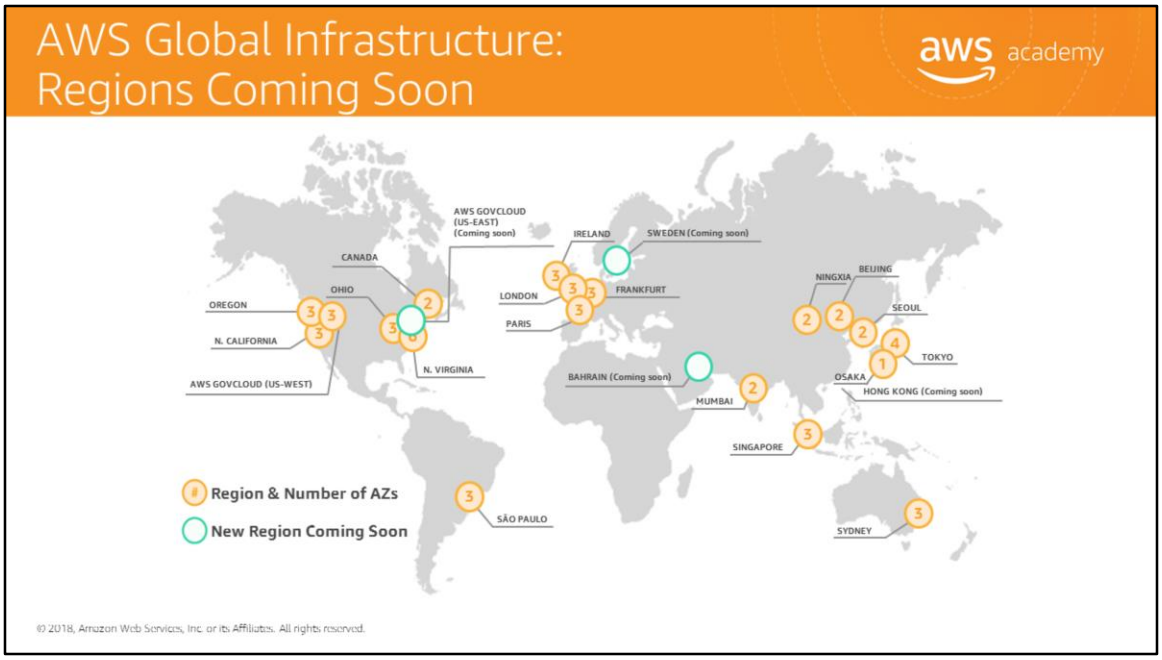
When you store data in a specific region, it is not replicated outside that region. AWS *never* moves your data out of the region you put it in. It is your responsibility to replicate data across regions, if your business needs require that. AWS provides information about the country, and, where applicable, the state where each region resides. You are responsible for selecting the region to store data in, based on your compliance and network latency requirements. When you distribute applications across multiple Availability Zones, be aware of location-dependent privacy and compliance requirements, such as the EU Data Privacy Directive. When selecting a Region, it is also important to consider which region will help you optimize latency while minimizing costs and adhering to whatever regulatory requirements you may have.

Let’s dive deeper on this point. If you are leveraging cloud computing services, you can easily deploy your application in multiple regions. For instance, you can have an application in a

region nearest your headquarters, such as San Diego, and then also have a deployable application in a region in the East Coast. Let's say your largest customer base is located in Virginia. With a few clicks, you can easily deploy in the US East region to provide a better experience for your customers located there. You will be minimizing latency and increasing agility for your organization within minutes and with minimal cost.



To see the regions currently available, navigate to the AWS homepage (https://aws.amazon.com) and scroll down to the global network of regions and edge locations. Edge Locations can be found by selecting the link https://aws.amazon.com/cloudfront/details/).

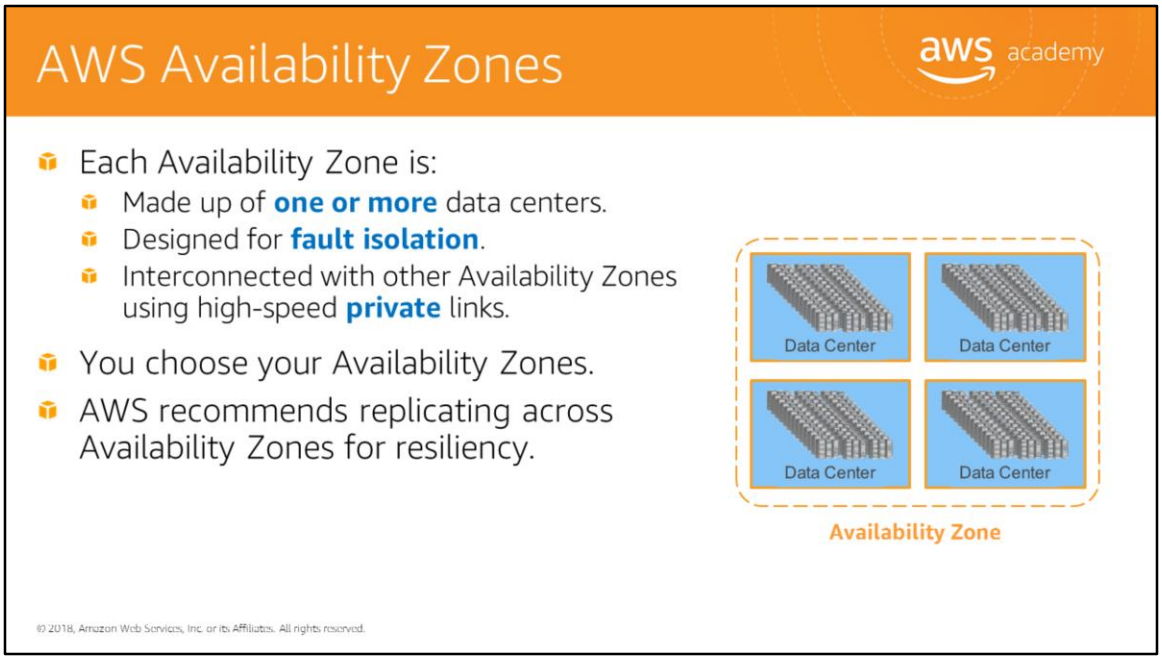


AWS products and services are available by region, so you may not see all regions available for a given service. AWS is steadily expanding its global infrastructure to help customers achieve lower latency and higher throughput, and to ensure that your data resides only in the Region you specify.

The AWS Cloud has announced plans to expand with 17 new Availability Zones in four new geographic Regions: Bahrain, Hong Kong, Sweden, and a second AWS GovCloud Region in the Eastern US.

The isolated GovCloud (US) Region is designed to allow US government agencies and customers to move sensitive workloads into the cloud by addressing their specific regulatory and compliance requirements.

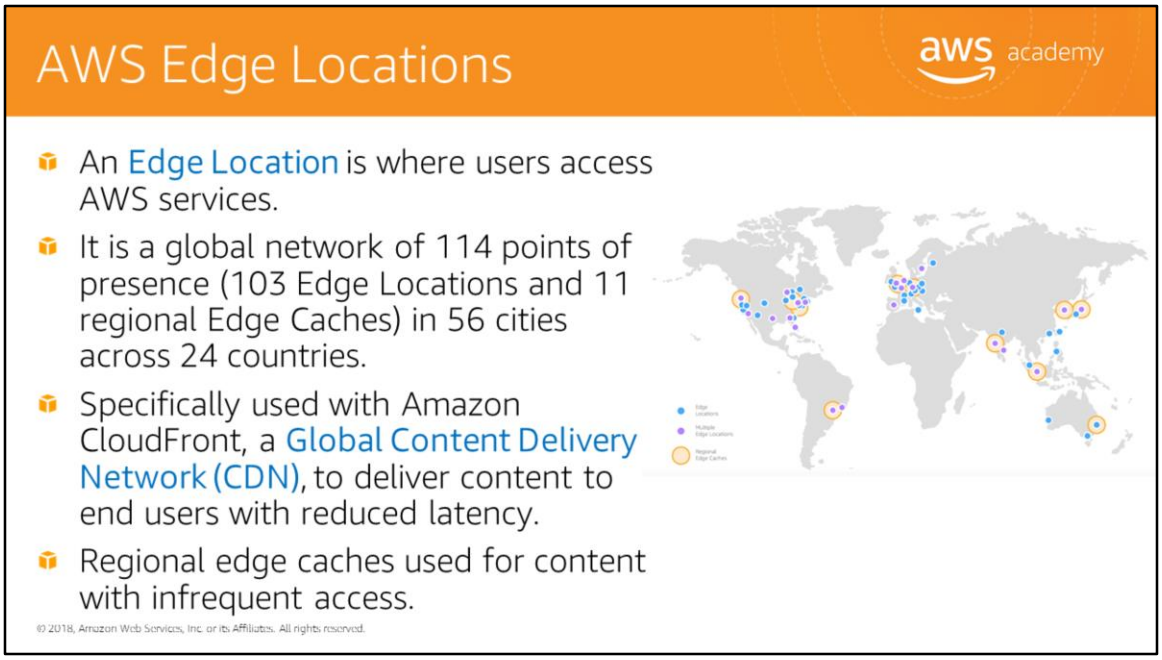
For more information about global infrastructure, select the link.
http://aws.amazon.com/about-aws/globalinfrastructure/



Availability Zones consist of one or more discrete data centers designed for fault isolation, each with redundant power, networking, and connectivity housed in separate facilities. They are interconnected with other Availability Zones using high-speed private links. Some Availability Zones have as many as six data centers; however, no data center can be part of two Availability Zones.

Each Availability Zone is designed as an independent failure zone. This means that Availability Zones are physically separated within a typical metropolitan region and are located in lower-risk flood plains with specific flood-zone categorization that varies by region. In addition to having a discrete uninterruptable power supply and onsite backup generation facilities, they are each fed via different grids from independent utilities to further reduce single points of failure. Availability Zones are all redundantly connected to multiple tier-1 transit providers. Availability Zones in a region are connected through low-latency links.

You are responsible for selecting the Availability Zones where your systems will reside. Systems can span across multiple Availability Zones. AWS recommends replicating across Availability Zones for resiliency. You should design your systems to survive temporary or prolonged failure of an Availability Zone if a disaster occurs. Distributing applications across multiple Availability Zones allows them to remain resilient in most failure situations, including natural disasters or system failures.

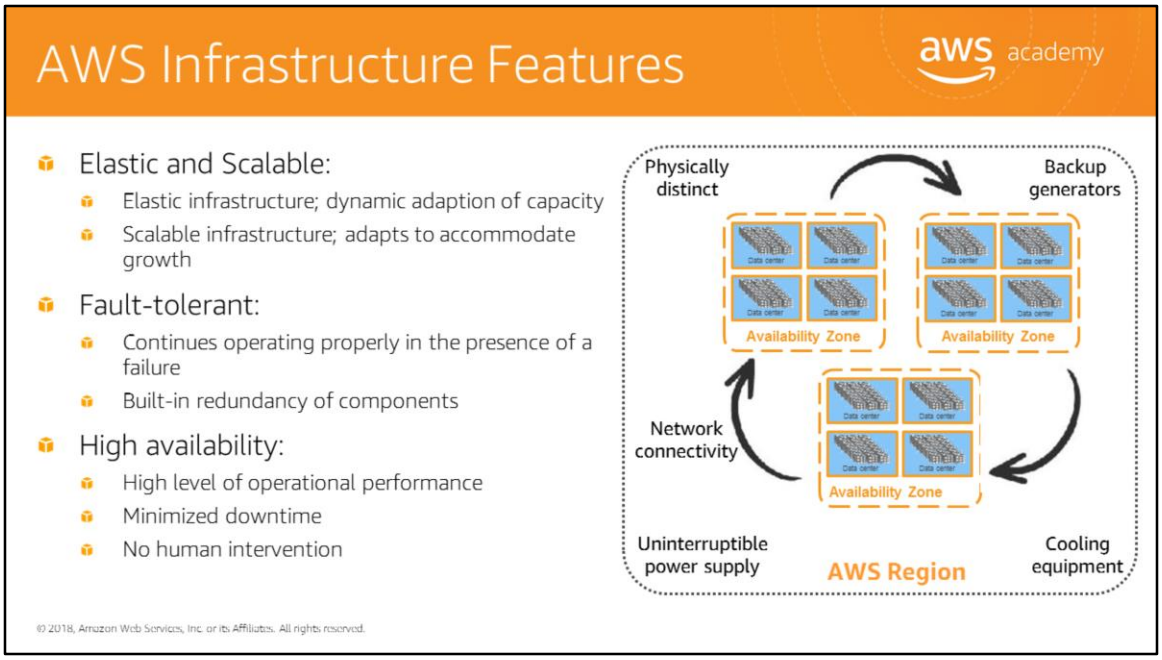


An Edge Location is where end users access AWS services.

It is a global network of 114 points of presence, with 103 Edge Locations and 11 regional Edge Caches in 56 cities, across 24 countries. They are located in most of the major cities around the world and serve requests for Amazon CloudFront and Amazon Route 53. Edge Locations are currently located in North America, Europe, Asia, Australia, and South America. AWS Edge locations offer CloudFront, Amazon Route 53, AWS Shield, and AWS Web Application Firewall (or WAF) services.

CloudFront is a Content Delivery Network (or CDN) used to distribute content to end users to reduce latency. Amazon Route 53 is a DNS service. Requests going to either one of these services will be routed to the nearest Edge Location automatically.

Regional Edge Caches, used by default with Amazon CloudFront, are utilized when you have content that is not accessed frequently enough to remain in an Edge Location. Regional Edge Caches absorb this content and provide an alternative to that content having to be fetched from the origin server.

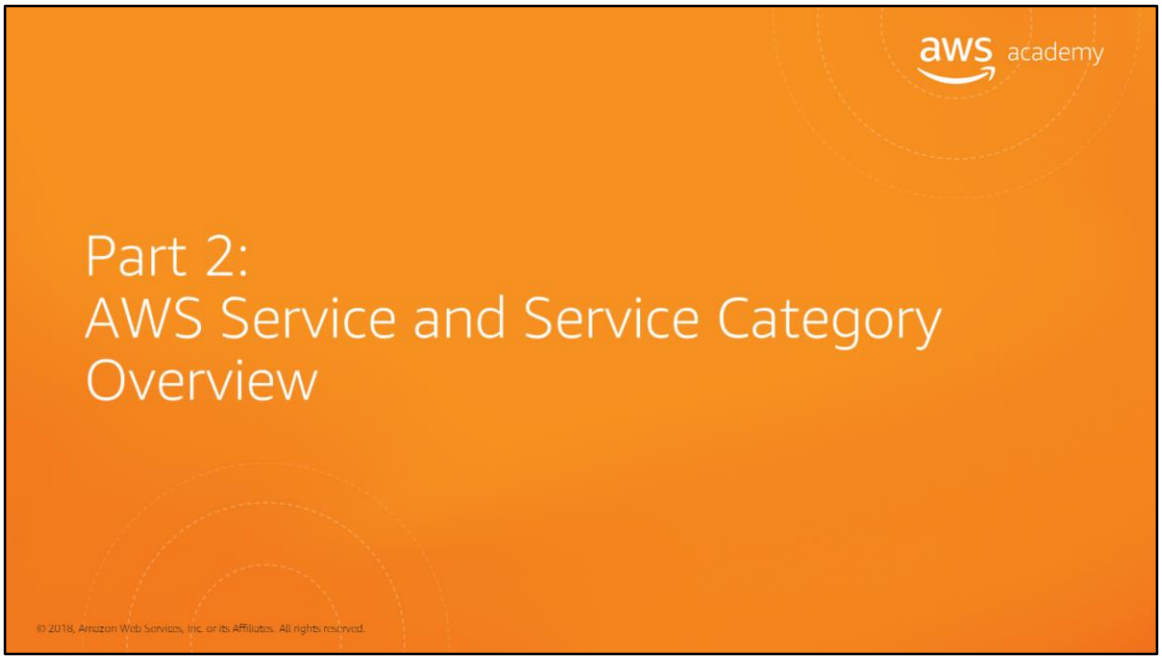


The AWS Cloud infrastructure is built around Regions and Availability Zones. AWS Regions provide multiple, physically separated, and isolated Availability Zones. An AWS Region contains two or more Availability Zones.

An Availability Zone is a data center or collection of data centers that are connected with low latency, high throughput, and highly redundant networking. Availability Zones are physically distinct and each has equipment like Uninterruptible Power Supplies, cooling equipment, backup generators, and security, to ensure uninterrupted operations.

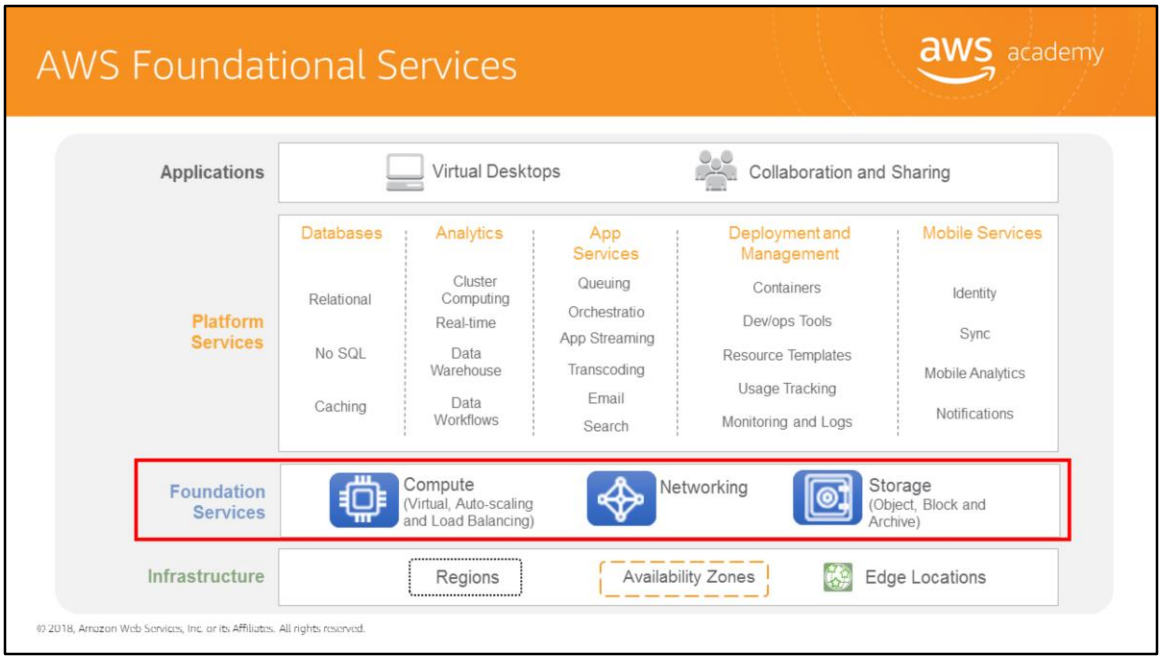
This infrastructure has several valuable features:

- First, it is elastic and scalable. This means resources can dynamically adjust to increases or decreases in capacity requirements. It can also rapidly adjust to accommodate growth.
- Second, this infrastructure is fault tolerant, which means it has built-in component redundancy which enables it to continue operations despite a failed component.
- Finally, it requires minimal to no human intervention, while providing high availability with minimal down time.



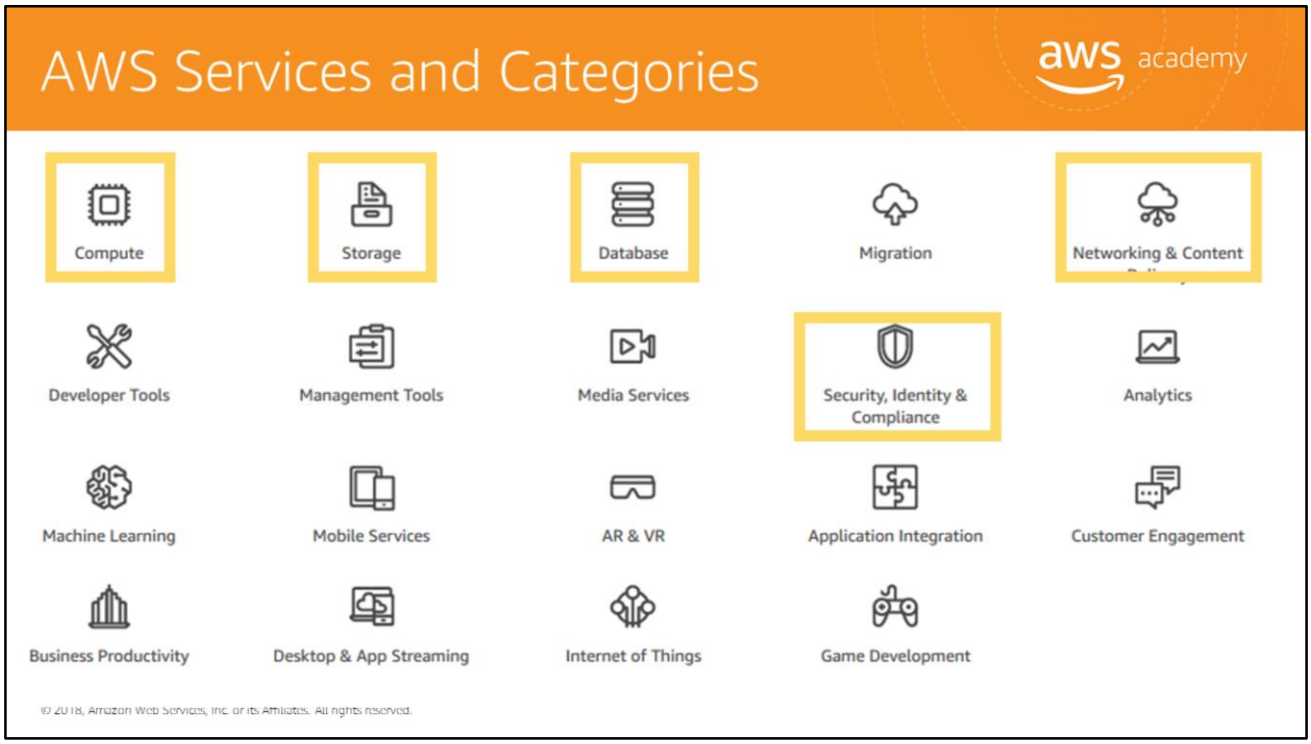
Introducing Part 2: AWS Service and Service Category Overview.

AWS offers a broad set of global cloud based products that can be used as building blocks for common cloud architectures. Let's look at how these cloud based products are organized.



As discussed previously, AWS's global infrastructure can be broken down into three elements: Regions, Availability Zones, and Edge Locations. This infrastructure provides the platform for a broad set of services, such as networking, storage, compute power, and databases delivered as an on-demand utility that is available in seconds, with pay-as-you-go pricing.

Now, let's shift our focus to the core services and take a more in-depth look at what these are and what each offers you for building your cloud solution.

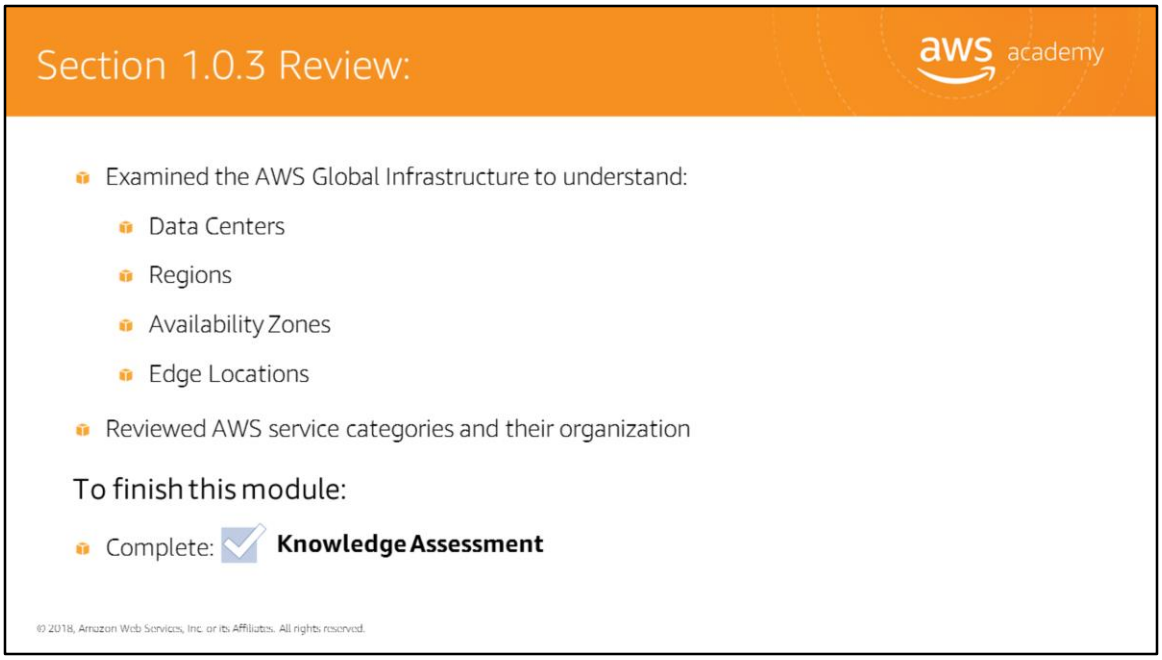


AWS offers a broad set of global cloud-based services that can be used as building blocks for common cloud architectures. Some of the categories we will discuss in this module include Compute, Storage, Database, Networking & Content Delivery and Security, Identity & Compliance.

If you go to the AWS front page, aws.amazon.com, and scroll down to find the section that allows you to explore the products. It places all of the products and services into different categories. For example, click on Compute and you will see Amazon EC2 is first on the list. There are also a lot of other products and services that appear in the compute category.

If you click Amazon EC2, it brings you to the Amazon EC2 main page. Select the link to review. It gives you a detailed description of the product and lists some of the benefits. Additionally, there are links for Product Details, Instance Types, Pricing, Getting Started, FAQs, and Resources. When you click on Product Details there is more detailed information about Amazon EC2.

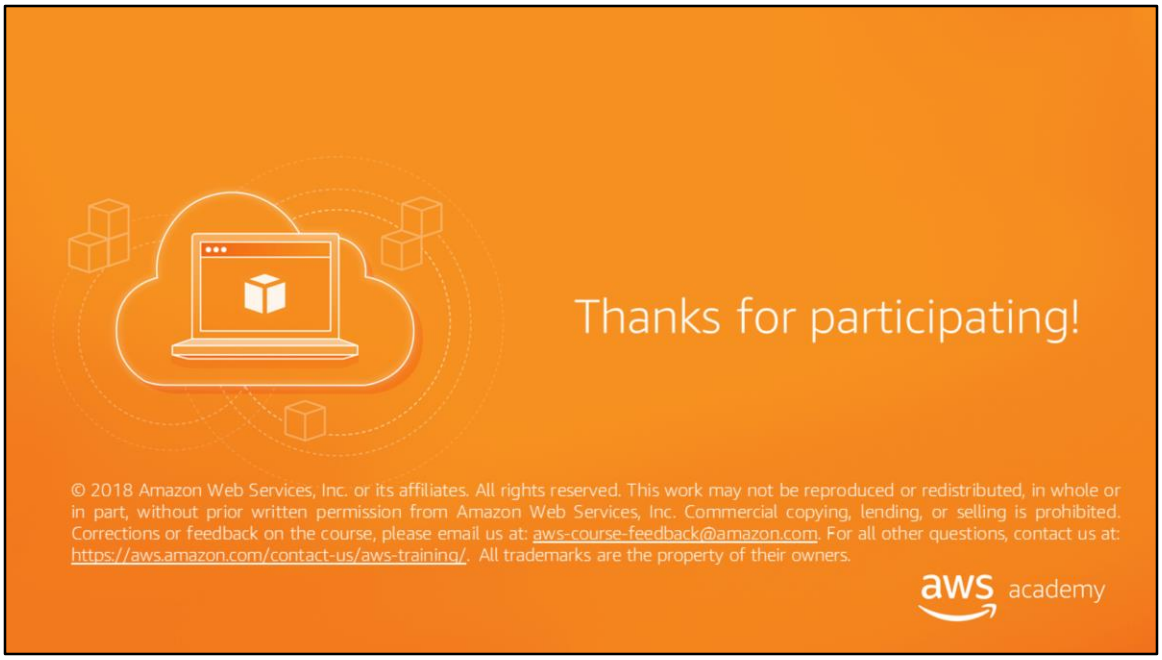
Explore the different service groups to understand the categories and services within them. Now that you know how to locate information about different services, let’s narrow our discussion to the AWS Core Services.



In summary, we examined the AWS Global Infrastructure to understand Data Centers, Regions, Availability Zones, and Edge Locations, and we reviewed different categories of AWS service categories and their organization.

To finish this module, please complete the corresponding knowledge assessment.

In Unit 2, we will look at the core of Amazon Web Services to better understand the specifics of each service. We will start in Module 2, Unit 1 with an introduction of compute services.



Thanks for participating!

