What is AWS?

AWS stands for Amazon Web Services that offers various IT services on demand using distributed IT infrastructure and offers flexible, reliable, scalable, and cost-effective cloud computing solutions.

Amazon Web Services (AWS) is the world's most comprehensive and broadly adopted cloud platform, offering nearly 200 fully-featured services from data centers globally. Millions of customers—including the fastest-growing startups, largest enterprises, and leading government agencies—are using AWS to lower costs, become more agile, and innovate faster.

AWS provides different services such as infrastructure as a service (IaaS), platform as a service (PaaS), software as a service (SaaS), and a wide range of workloads, including game development, data processing, warehousing, performance growth, and more.

AWS was invented by Amazon.com and it's the origin as a developer platform can be traced back to 2002 when an initial beta called Amazon.com Web Service was launched.

The Amazon logo was created to represent the message that it sells everything from A to Z (the arrow connects the two letters) and also represents the smile that customers would experience by shopping on the Amazon.com Web site (the arrow becomes a smile).

In 2003, when Chris Pinkham and Benjamin Black presented a paper describing a vision of Amazon's retail computing infrastructure, the AWS concept was publicly reformulated.

In 2004, the first AWS service launched for public usage: Simple Queue Service (SQS).

In 2006, Amazon Web Services LLC (AWS), an Amazon.com company, officially began offering developer customers access to in-the-cloud infrastructure services based on Amazon's own back-end technology platform. The three initial service offerings of Amazon S3 cloud storage, SQS, and EC2 were combined.

In 2007, Amazon launched Amazon SimpleDB, which enables organizations, academics, data analysts and developers to process large amounts of data easily and cheaply.

In 2008, Amazon announced Elastic IPs, Amazon Elastic Block Store (EBS), Amazon CloudFront,

In 2009, Amazon launched Amazon Elastic MapReduce (EMR), Elastic Load Balancing (ELB), Virtual Private Cloud (VPC), Amazon Relational Database Service (RDS), EC2 Spot Instances, Amazon Route 53, AWS Elastic Beanstalk,

In 2010, it was announced that all the retail sites of Amazon.com had moved to AWS. AWS launched Simple Notification Service and AWS CloudFormation,

In 2011, AWS announced the launch of Amazon Simple Email Service (SES),

In 2012, AWS launched Amazon DynamoDB, AWS Identity and Access Management (IAM) for EC2, Amazon Glacier, Amazon Redshift,

In 2013, AWS announced AWS CloudTrail, released Amazon Kinesis, AWS Lambda,

In 2014, AWS announced Amazon Aurora, EC2 Container Service (ECS),

In 2015, AWS launched AWS API Gateway Service, AWS Elasticsearch Service, Snowball, Internet of Things platform, Amazon Elastic Container Registry (ECR).

In 2016, AWS announced Auto Scaling for Amazon EC2 Container Service (ECS), Elastic File System (EFS), AWS Snowmobile, Snowball Edge, Amazon Lightsail and AWS acquired Cloud9.

In 2017, AWS announced Amazon Glue, Amazon SageMaker, AWS CloudWatch agent.

In 2018, AWS announced AWS Elastic Kubernetes Service (EKS).

Today, AWS is a comprehensive cloud services platform, offering to compute power, storage, content delivery, and other functionality that enables businesses to cost-effectively deploy applications and services with greater flexibility, scalability, and reliability. The AWS mission is to enable developers and businesses to use web services to easily build and be paid for sophisticated, scalable applications.

AWS Features

Amazon Web Services has a variety of features that make it consistent across different companies. AWS's characteristics are:

Easy to use:

AWS is designed to allow application providers, ISVs, and vendors to quickly and securely host your applications – whether an existing application or a new SaaS-based application. You can use the AWS Management Console or well-documented web services APIs to access AWS's application hosting platform.

Reliable:

With AWS, you take advantage of a scalable, reliable, and secure global computing infrastructure, the virtual backbone of Amazon.com's multi-billion dollar online business that has been honed for over a decade.

Flexible:

AWS enables you to select the operating system, programming language, web application platform, database, and other services you need. With AWS, you receive a virtual environment that lets you load the software and services your application requires. This eases the migration process for existing applications while preserving options for building new solutions.

Scalable and high-performance:

Using AWS tools, Auto Scaling, and Elastic Load Balancing, your application can scale up or down based on demand. Backed by Amazon's massive infrastructure, you have access to compute and storage resources when you need them.

Cost-Effective:

You pay only for the compute power, storage, and other resources you use, with no long-term contracts or up-front commitments. For more information on comparing the costs of other hosting alternatives with AWS, see the AWS Economics Center.

Secure:

AWS utilizes an end-to-end approach to secure and harden the infrastructure, including physical, operational, and software measures. For more information, see the AWS Security Center.

AWS Pros

AWS is a leading brand in the field and has important features that enable it.

- AWS offers more than 5 times the computation capacity in use compared to the aggregate of the other 14 leading market providers.
- With hundreds of thousands of customers operating every possible use case on AWS in over 190 countries.
- Thousands of independent software vendors such as SAP, Oracle, Adobe, Microsoft, etc. also made their applications available to customers on AWS.
- Amazon Web Services currently supports over 2,000 government agencies and 5,000 educational institutions. AWS also attained most industry-standard compliance certifications like HIPAA, SOC 1/SSAE 16/ISAE 3402 (formerly SAS70), SOC 2, SOC 3, PCI DSS Level 1, ISO 27001, FedRAMP, DIACAP and FISMA, ITAR, FIPS 140-2, CSA, MPAA.

Following are some disadvantages of AWS services:

- Sometimes difficult learning curve for large companies
- Billing can be confusing
- Amazon's EC2 Limits
- Common Cloud Computing problems

AWS Product Categories

Here are AWS Main product categories;

- Analytics
- Application Integration
- AR & VR
- AWS Cost Management
- Blockchain
- Business Applications
- Compute
- Customer Engagement
- Database
- Developer Tools
- End-User Computing
- Game Tech
- Internet of Things
- Machine Learning
- Management & Governance
- Media Services
- Migration & Transfer
- Mobile
- Networking & Content Delivery
- Quantum Technologies
- Robotics
- Satellite
- Security, Identity, & Compliance
- Storage

AWS Infrastructure

Introduction
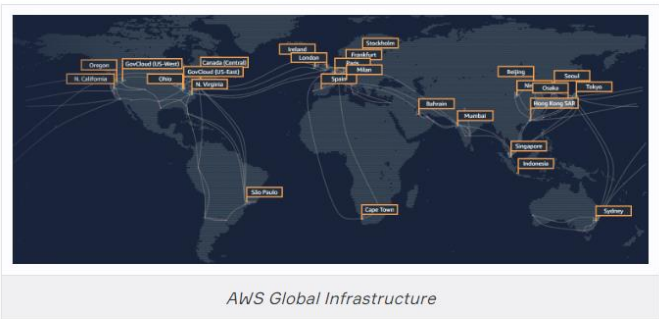
AWS Global Infrastructure

AWS serves over a million active customers in more than 190 countries. AWS expanding global infrastructure to help their customers achieve lower latency and higher throughput, and to ensure that their data resides only in the AWS Region they specify. As customers grow their businesses, AWS will continue to provide the infrastructure that meets their global requirements.

AWS serves over a million active customers in more than 190 countries. AWS expanding global infrastructure to help their customers achieve lower latency and higher throughput, and to ensure that their data resides only in the AWS Region they specify. As customers grow their businesses, AWS will continue to provide the infrastructure that meets their global requirements.
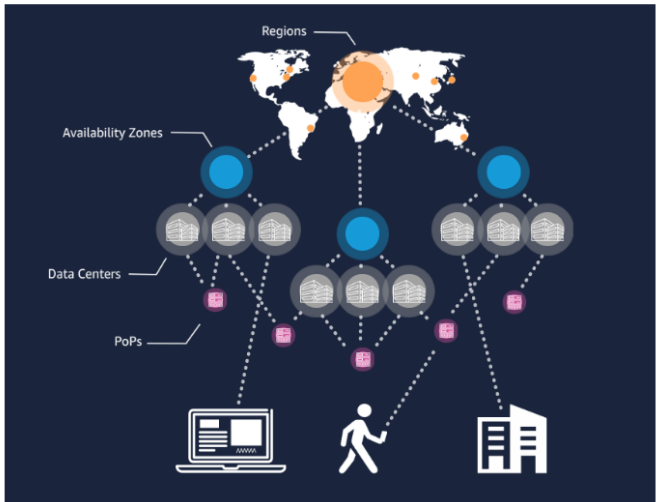
The AWS Cloud infrastructure is built around AWS Regions and Availability Zones.

- An AWS Region is a physical location in the world where we have multiple Availability Zones.
- Availability Zones consist of one or more discrete data centers, each with redundant power, networking, and connectivity, housed in separate facilities.
- These Availability Zones offer you the ability to operate production applications and databases that are more highly available, fault-tolerant, and scalable than would be possible from a single data center.
- The AWS Cloud operates in over 77 Availability Zones within over 24 geographic Regions around the world, with announced plans for more Availability Zones and Regions.
- For more information on the AWS Cloud Availability Zones and AWS Regions, see AWS Global Infrastructure.

AWS has the concept of a Region, which is a physical location around the world where we cluster data centers.

- Each group of logical data centers is called Availability Zone (AZ).
- Each AWS Region consists of multiple, isolated, and physically separate AZ's within a geographic area.
- Unlike other cloud providers, who often define a region as a single data center, the multiple AZ design of every AWS Region offers advantages for customers.
- Each AZ has independent power, cooling, and physical security and is connected via redundant, ultra-low-latency networks. AWS customers focused on high availability can design their applications to run in multiple AZ's to achieve even greater fault-tolerance.
- AWS infrastructure Regions meet the highest levels of security, compliance, and data protection.

AWS Regions

AWS now spans 76 Availability Zones within 24 geographic regions around the world and has announced plans for thirteen more Availability Zones and four more AWS Regions in Indonesia, Italy, South Africa, and Spain.

An Availability Zone (AZ) is one or more discrete data centers with redundant power, networking, and connectivity in an AWS Region.

AWS AZs
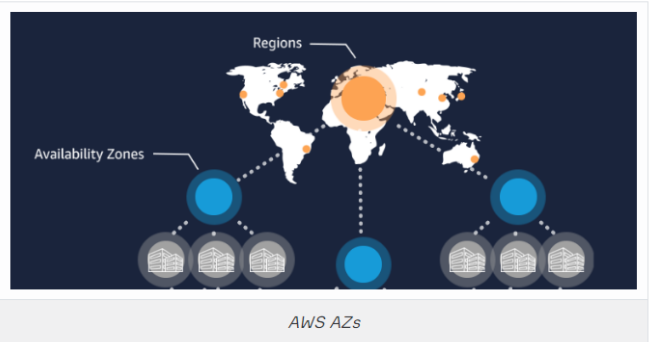
- AZ's give customers the ability to operate production applications and databases that are more highly available, fault-tolerant, and scalable than would be possible from a single data center.
- All AZ's in an AWS Region are interconnected with high-bandwidth, low-latency networking, over fully redundant, dedicated metro fiber providing high-throughput, low-latency networking between AZ's.
- All traffic between AZ's is encrypted.
- The network performance is sufficient to accomplish synchronous replication between AZ's.
- AZ's make partitioning applications for high availability easy. If an application is partitioned across AZ's, companies are better isolated and protected from issues such as power outages, lightning strikes, tornadoes, earthquakes, and more.
- AZ's are physically separated by a meaningful distance, many kilometers, from any other AZ, although all are within 100 km (60 miles of each other).

AWS Data Centers

A data center is a facility that centralizes the IT processes and resources of an enterprise, as well as where the data are processed, handled and disseminated.

AWS pioneered cloud computing in 2006, creating cloud infrastructure that allows you to securely build and innovate faster. AWS is continuously innovating the design and systems of data centers to protect them from man-made and natural risks.

AWS Data Centers

AWS data centers are secure by design. Before building a data center, AWS spend countless hours considering potential threats and designing, implementing, and testing controls to ensure the systems, technology, and people they deploy counteract risk. To help customers fulfill their own audit and regulatory requirements, AWS provides some of the physical and environmental controls below.

- Secure Design (Site Selection, Redundancy, Availability, Capacity Planning)
- Business Continuity & Disaster Recovery (Business Continuity Plan, Pandemic Response)
- Physical Access (Employee Data Center Access, Third-Party Data Center Access, AWS GOVCLOUD Data Center Access)
- Monitoring & Logging (Data Center Access Review, Data Center Access Logs, Data Center Access Monitoring)
- Surveillance & Detection (Closed Circuit Television Camera (CCTV), Data Center Entry Points, Intrusion Detection)
- Device Management (Asset Management, Media Destruction)
- Operational Support Systems (Power, Climate & Temperature, Fire Detection & Suppression, Leakage Detection)
- Infrastructure Maintenance (Equipment Maintenance, Environment Management)
- Governance & Risk (Ongoing Data Center Risk Management, Third-Party Security Attestation)

AWS Data Center Layers

Perimeter Layer

AWS data center physical security begins at the Perimeter Layer. This layer includes a number of security features depending on the location, such as security guards, fencing, security feeds, intrusion detection technology, and other security measures.

Infrastructure Layer

The Infrastructure Layer is the data center building and the equipment and systems that keep it running. Components like back-up power equipment, the HVAC system, and fire suppression equipment are all part of the Infrastructure Layer. These devices and systems help protect servers and ultimately your data.

Data Layer

The Data Layer is the most critical point of protection because it is the only area that holds customer data. Protection begins by restricting access and maintaining a separation of privilege for each layer. In addition, AWS deploys threat detection devices, video surveillance, and system protocols, further safeguarding this layer.

Environmental Layer

The Environmental Layer is dedicated to environmental considerations from site selection and construction to operations and sustainability. AWS carefully chooses our data center locations to mitigate environmental risks, such as flooding, extreme weather, and seismic activity.

Edge Locations

An edge location is where end-users access services located at AWS and used for caching content. Edge locations serve requests for CloudFront and Route 53.

- CloudFront is a content delivery network, while Route 53 is a DNS service.
- Requests going to either one of these services will be routed to the nearest edge location automatically. This allows for low latency no matter where the end-user is located.
- They are located in most of the major cities around the world and are specifically used by CloudFront (CDN) to distribute content to end-user to reduce latency.
- It is like a frontend for the service we access which are located in the AWS cloud.

AWS Local Zones

AWS Local Zones place compute, storage, database, and other select AWS services closer to end-users.

- With AWS Local Zones, you can easily run highly-demanding applications that require single-digit millisecond latencies to your end-users such as media & entertainment content creation, real-time gaming, reservoir simulations, electronic design automation, and machine learning.
- Each AWS Local Zone location is an extension of an AWS Region where you can run your latency-sensitive applications using AWS services such as Amazon Elastic Compute Cloud, Amazon Virtual Private Cloud, Amazon Elastic Block Store, Amazon File Storage, and Amazon Elastic Load Balancing in geographic proximity to end-users.
- AWS Local Zones provide a high-bandwidth, secure connection between local workloads and those running in the AWS Region, allowing you to seamlessly connect to the full range of in-region services through the same APIs and toolsets.
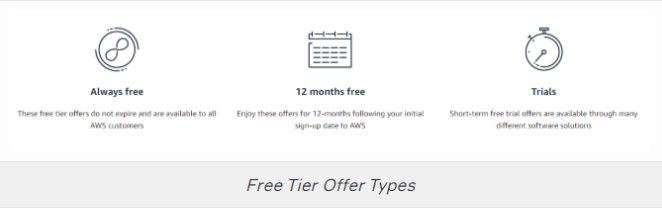
AWS Free Tier

What is AWS Free Tier?

The AWS Free Tier provides customers the ability to explore and try out AWS services free of charge up to specified limits for each service. You can explore more than 60 products and start building on AWS using the free tier.
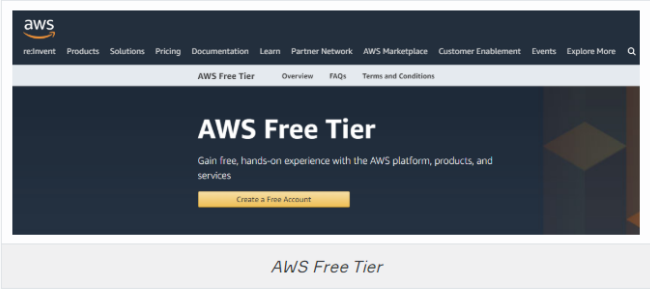
AWS Free Tier

- The Free Tier is designed to give you hands-on experience with a range of AWS services at no charge.
- For example, you can explore AWS as a platform for your business by setting up a test website with a server, alarms, and database.
- You can also try out services for developers, such as AWS CodePipeline, AWS Data Pipeline, and AWS Device Farm.
- If you don't use the full benefits provided by the Free Tier in a given month, the benefits don't roll over to the next month.
- To maximize your Free Tier benefits, be sure to spend time with AWS each month, trying out the services that you're curious about.
- For more detailed information, please follow this link.

Types of Offers

The Free Tier is comprised of three different types of offerings, a 12-month Free Tier, an Always Free offer, and short term trials.

Free Tier Offer Types

- Services with a **12-month Free Tier** allow customers to use the product for free up to specified limits for one year from the date the account was created.
- Services with an **Always Free** offer allow customers to use the product for free up to specified limits as long as they are an AWS customer.
- Services with a **short term trial** are free to use for a specified period of time or up to a one-time limit depending on the service selected.

The 12 Month Free Tier is only available to new AWS customers and is available for 12 months following your AWS sign-up date. The Other Offers are available to both existing and new AWS customers and may be limited in duration (such as for trials) or in available free usage (such as the amount of free storage for a database Offer).

You will not be eligible for any **Offers** if you or your entity create(s) more than one account to receive additional benefits under the Offers. An Organization (under AWS Organizations) can only benefit from Offers from one account in the Organization, and to calculate the Organization's use of AWS Services under any Offers, AWS will aggregate the usage across all accounts in the Organization. You will be charged standard rates for use of AWS Services if AWS determines that you are not eligible for an Offer.

Billing Policy

To avoid charges while on the Free Tier, you must keep your usage below the Free Tier limits.

- You are charged for any usage that exceeds the limits.
- To help you stay within the limits, you can track your Free Tier usage and set a billing alarm to notify you if you start incurring charges.
- If your application use exceeds the free tier limits, you simply pay standard, pay-as-you-go service rates.
- AWS Free Tier is applied to your monthly usage. It will expire on the 1st day of each month and does not accumulate.

You can see current and past usage activity by service and region by logging into your account and going to the Billing & Cost Management Dashboard. From there you can manage your costs and usage using AWS Budgets, visualize your cost drivers and usage trends via Cost Explorer, and dive deeper into your costs using the Cost and Usage Reports.

Limits

All services that offer a Free Tier have limits on what you can use without being charged. Many services have multiple types of limits.

- For example, Amazon EC2 has limits on both the type of instance you can use and how many hours you can use in one month.
- Amazon S3 has a limit on how much storage you can use and on how often you can call certain operations each month.
- For example, the Free Tier covers the first 20,000 times you retrieve a file from Amazon S3, but you're charged for additional file retrievals.
- Each service has limits that are unique to that service.

Some of the most common limits are by time, such as hourly or by the minute, or by requests, which are the requests you send to the service, also known as API operations.

AWS Free Tier Limits