

AWS GLOBAL CLOUD INFRASTRUCTURE: REGIONS, ZONES & MORE



Amazon Web Services (AWS) pioneered the modern-day cloud computing [IT service delivery model](#). Amazon's [cloud service delivery](#) success relies on the technology's accessibility and availability for end-users.

The AWS Cloud has grown rapidly since its inception, expanding its services across worldwide geographic locations. Its vast global cloud infrastructure footprint sets it apart from leading competitors like Microsoft Azure and Google Cloud.

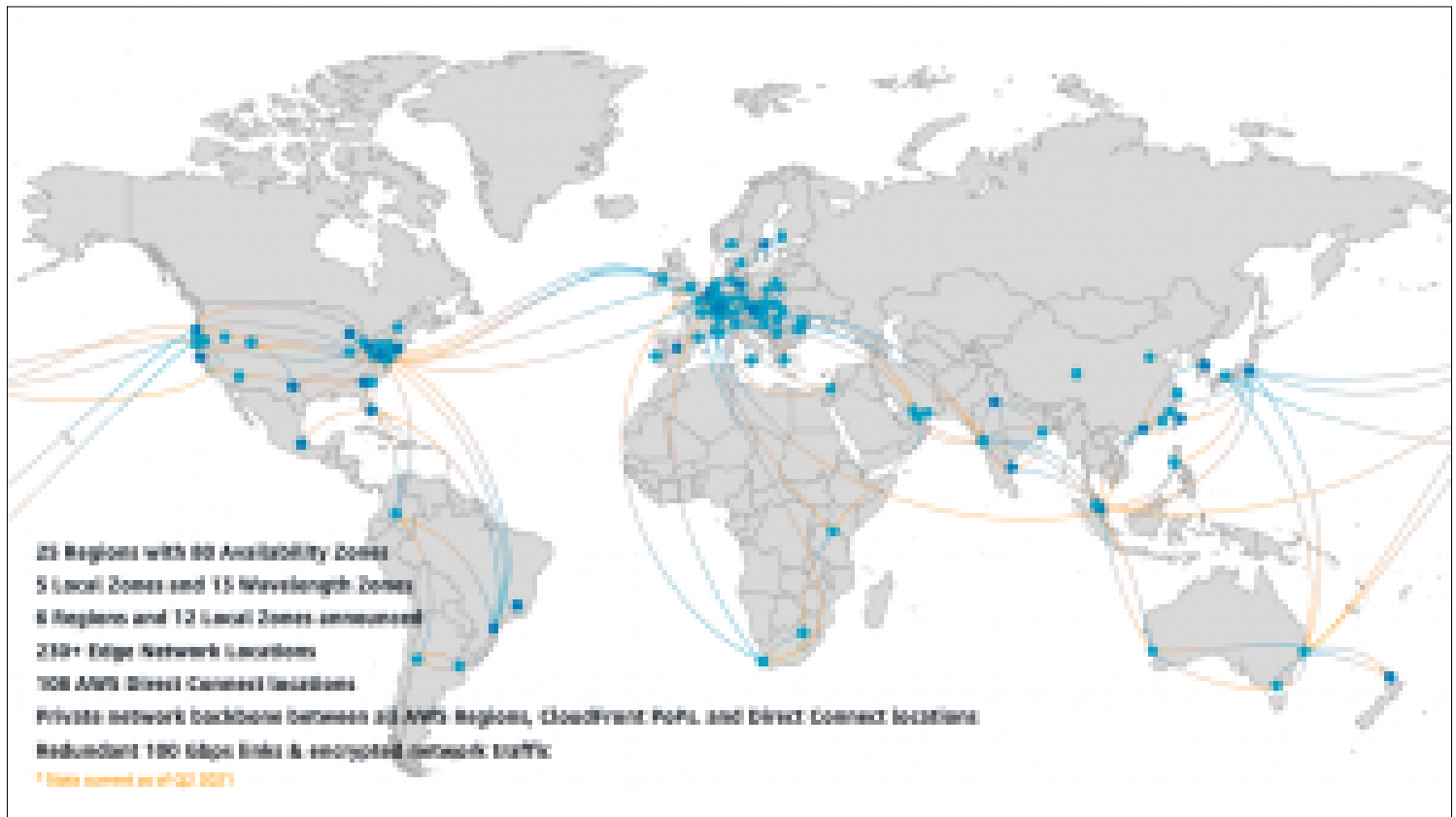
This primer describes how the AWS global cloud infrastructure works and what that means for AWS customers. We'll look at the following items:

- [What is the AWS Global Cloud Infrastructure?](#)
- [AWS Cloud Infrastructure Delivery Environments](#)
- [How Big is the AWS Cloud?](#)
- [Considerations When Choosing AWS Global Infrastructure Components](#)

(This tutorial is part of our [AWS Guide](#). Use the right-hand menu to navigate.)

What is the AWS Global Cloud Infrastructure?

The AWS Global Cloud infrastructure is the backbone network of global data centers and other platforms that Amazon uses to deliver application workloads and AWS services.



The AWS Global Cloud Infrastructure is the backbone network for delivering AWS workloads and services (Source)

For cloud application and service delivery, customers provision and connect their end users and organizational environments to the following AWS global infrastructure components:

- AWS Regions
- AWS Availability Zones
- AWS Local Zones
- AWS Wavelength Zones

AWS Regions

The heart of the AWS Global Cloud. AWS Regions are physical locations around the world where Amazon clusters data centers for application and service delivery in AWS Availability Zones. Regions also provide extensions for other delivery options, such as AWS Local Zones.

Each AWS Region may offer different service quality in terms of latency, solutions portfolio, and cost, based on its geographic location and distance from customer sites.

(Explore regions from other cloud providers in [Availability Regions & Zones for AWS, Azure & GCP.](#))

AWS Availability Zones

An Availability Zone (AZ) is a grouping of one or more discrete data centers that provide applications and services in an AWS region.

Each AZ contains redundant connectivity, power, and networking capabilities, and individual AZs are physically separated (isolated) from each other by a meaningful distance. All AZs in an AWS Region are connected through low latency and high throughput networking channels.

Because of their connectivity and redundancy, AZs provide customer application and database operating environments that are more scalable and fault tolerant. Because regional AZs are physically isolated from each other, applications can be partitioned across multiple AZs for [high availability](#).

AWS Local Zones

AWS Region extensions that place compute, [storage](#), [databases](#), and other AWS services in closer proximity to large populations, IT centers, and industries.

AWS Local Zones are provisioned to run high-speed applications—such as media, entertainment, real-time gaming, [live video streaming](#), and [machine learning](#)—that require single-digit millisecond latency to service users in specific geographic locations.

AWS Wavelength Zones

Wavelength zones provide [5G telecommunications connections](#) inside AWS Regions. Wavelength zones embed AWS compute and storage capabilities within communication service providers' (CSPs) data centers at the edge of their 5G networks.

5G devices can reach apps running in Wavelength Zones without ever leaving the 5G network, allowing them to take advantage of 5G bandwidth and latency.

AWS Outposts

AWS Outposts is a [fully managed service](#) that allows customers to build and deploy AWS storage capacity at customer sites. Outposts extend AWS infrastructure, services, tools, and APIs to customer locations, enabling a hybrid delivery experience.

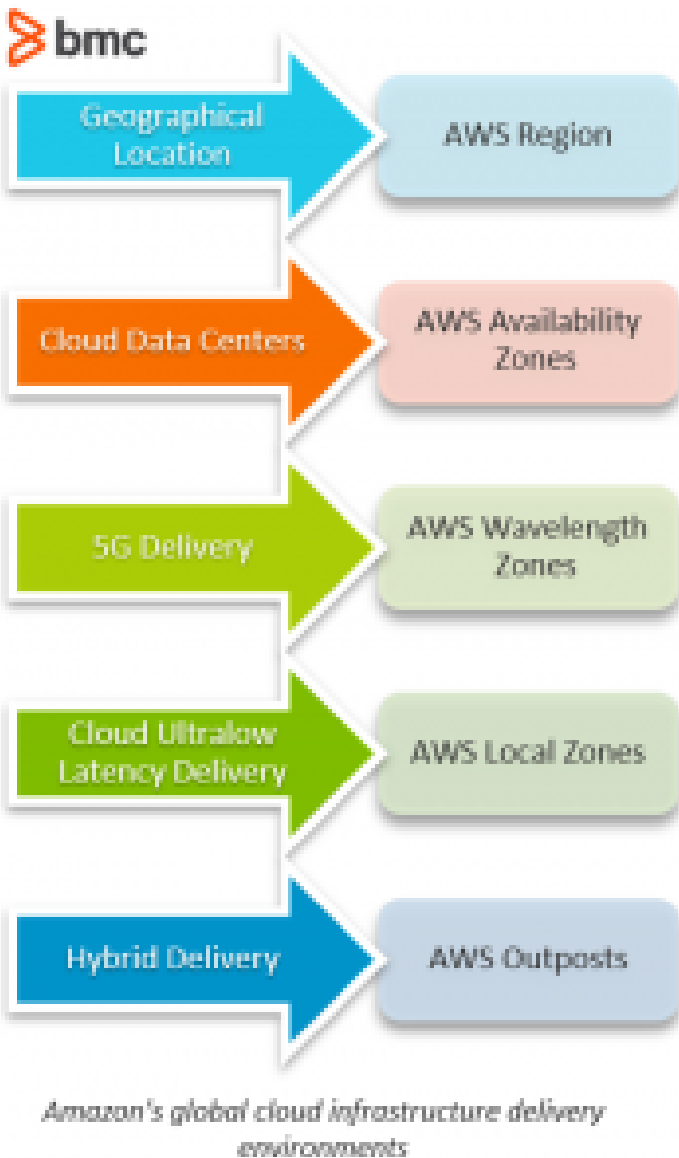
AWS Outposts can be created in almost any customer provided space, including:

- Data centers
- Co-location facilities
- Customer on-premises facilities

AWS cloud infrastructure delivery environments

Customers can use AWS global cloud infrastructure zones and outposts to deliver workloads and services to users inside specific geographical locations (AWS Regions), through the following infrastructure delivery environments:

- **Cloud data centers:** Scalable, clustered, and redundant Amazon provided data centers (AWS Availability Zones).
- **Cloud single-digit millisecond latency environments:** Amazon ultralow latency data centers that are significantly closer to users in large population centers, IT centers, and industries (AWS Local Zones).
- **5G environments:** 5G telecommunication network delivery for single-digit millisecond latency access for mobile devices and end users (AWS Wavelength Zones).
- **Hybrid delivery:** Customer hosted sites where workloads and services can be delivered with the same infrastructure, services, tools, and APIs used in AWS facilities (AWS Outposts).



How big is the AWS cloud?

As of this writing, here's how big the AWS cloud is:

- The AWS cloud spans 81 Availability Zones within 25 worldwide geographic regions. Announced plans include 21 more Availability Zones and seven more regions in Europe, Asia, and Australia.
- There are currently eight AWS Local Zones for ultralow latency applications, with nine more Local Zones on the way.
- There are 17 Wavelength Zones available for ultralow latency and 5G processing.

Up-to-date information on the location and count of AWS Regions, Availability Zones, Local Regions, and Wavelength Zones can be found on the [official AWS Global Infrastructure page](#).

How to choose AWS global infrastructure components

The choice for choosing your AWS global cloud infrastructure may come as a tradeoff for organizations based on a range of factors, including:

Latency & proximity

The distance between the cloud deployments (AWS Regions and Zones) and end-users is a key factor that determines the latency and network performance of the cloud service.

- For ultralow latency workloads, you may need to provision AWS Local Zones to meet user needs.
- For 5G access, there may be no other choice than to provision an AWS Wavelength Zone.

Performance is further affected when the cloud solution is integrated with on-premise legacy technologies and apps as part of a hybrid cloud strategy (AWS Outposts). In a hybrid strategy, performance is highly influenced by local factors, such as on-premises network and telecommunications speed, that AWS capabilities may not be able to compensate for.

Selecting the AWS Region closest to customer or end-user proximity helps ensure the best user experience. The closest region is also usually the least expensive option when compared to choosing an AWS Region in a distant geographic location.

Amazon also offers services such as [AWS Route 53](#) to automatically direct global network traffic through the most optimal channels to maximize availability and performance.

Cost

Pricing across AWS Regions varies, because of different [CapEx, OpEx](#), and regulations in different geographic locations. Organizations may need to identify the optimal tradeoff between the cost and other factors—including [service catalog items](#), latency, network performance, regulatory compliance—when configuring their AWS infrastructure.

AWS offers a [cost calculator](#) to estimate the expected costs of AWS services in different regions. The more complex your AWS environment is, the harder it will be to accurately estimate your AWS global cloud infrastructure costs. The cost calculator is a good place to start for estimating costs on expected AWS regions used.

Service catalog availability

Amazon offers a vast portfolio of cloud-based solutions spanning AWS Regions. While the most popular AWS services are available across all AWS Regions, not every region offers all services. Consult the [AWS Regional Services page](#) to determine AWS service availability in every Region.

AWS Region choices should be based on current and future workload needs. Workload requirement changes may require additional AWS service investments in different regions, particularly when servicing many worldwide locations.

Regulatory compliance & security

[Regulatory compliance and security](#) can also affect which AWS regions and zones will host your workloads. Consider the following factors when selecting which AWS global cloud infrastructure components to deploy for applications and services.

- **Regulatory compliance.** Specific industry and regulatory specifications—such as the EU's General Data Protection Regulation (GDPR) and state, provincial, or local locality data handling

regulations—may require sensitive end-user data to be processed in specific geographic locations.

- **Data mobility.** Cloud computing makes it easy for organizations to transfer, store, and process information in data centers at distant locations, which may be considered as a violation of compliance regulations that may lead to costly lawsuits and damages to the brand reputation.
- **Security requirements.** Organizations may also be obliged to distribute workloads across multiple geographically disparate cloud data centers to ensure high availability and security standards of sensitive business information and IT-enabled services.

Service level agreements

To meet the desired standards of IT service availability and performance, AWS provides many [Service Level Agreements \(SLAs\)](#) covering service uptime and credits for service failure.

Consult the [AWS Service Level Agreement](#) page for more information on the SLAs Amazon offers for delivered services.

The green factor

Many organizations are pushing to achieve carbon neutrality and adopting environmentally friendly business practices. Amazon is no different. Using renewable energy, AWS [aims to operate](#) a net carbon neutral global infrastructure by contributing toward various renewable energy and data center projects.

The location of AWS data centers is therefore one criteria to consider in your annual sustainability report that could serve as a competitive differentiation toward [your sustainability efforts](#).

Considering your cloud strategy

AWS offers a range of global cloud infrastructure offerings that can fit well in your IT strategy, whether you're focusing on service uptime, product portfolio, availability, compliance, or sustainability factors.

[CIOs](#) need to make well-informed decisions pertaining to AWS global cloud infrastructure with considerations for both the near-term and long-term implications of their IT investments.

(See how to build your [multi-cloud strategy](#).)

Related reading

- [BMC Multi-Cloud Blog](#)
- [The AWS Well-Architected Framework: 5 Pillars & Best Practices](#)
- [Cloud Service Providers \(CSPs\) Explained](#)
- [Cloud Governance Best Practices for Efficient Cloud Operations](#)
- [Common Roles in Cloud Computing](#)
- [AWS Certifications: An Introduction](#)