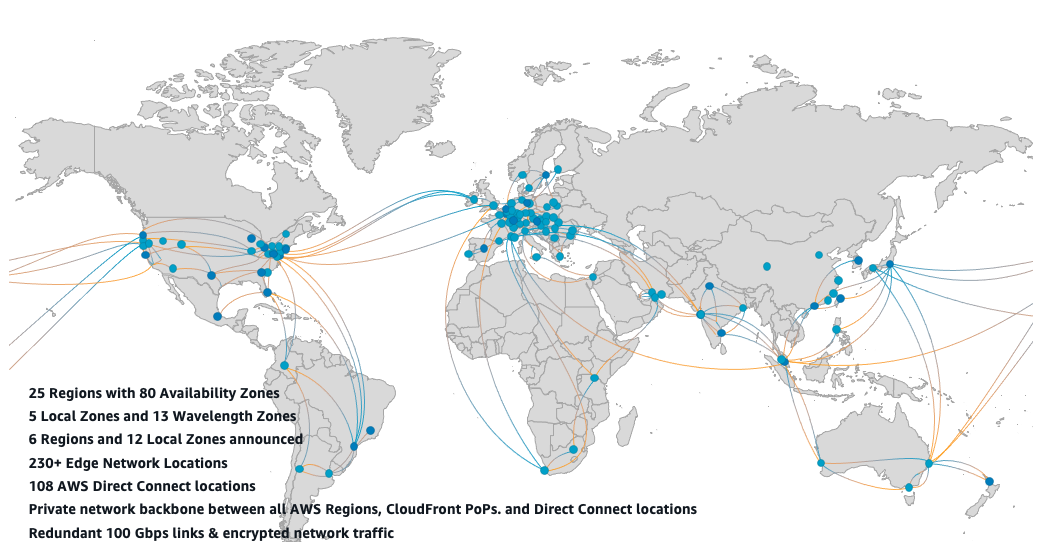
AWS NETWORKING

AWS Networking :

AWS networking allows you to build, manage, and secure virtual networks within the AWS Cloud, offering a wide array of services that are foundational for deploying any application.

AWS Global Infrastructure :

* The AWS global infrastructure is a vast network of physical data centers and points of presence that AWS uses to host and deliver its cloud services worldwide.
* It is built for security, high availability, and performance, with a layered architecture that includes regions, Availability Zones (AZs), and edge locations.



REGIONAL EDGE CACHES

EDGE LOCATIONS

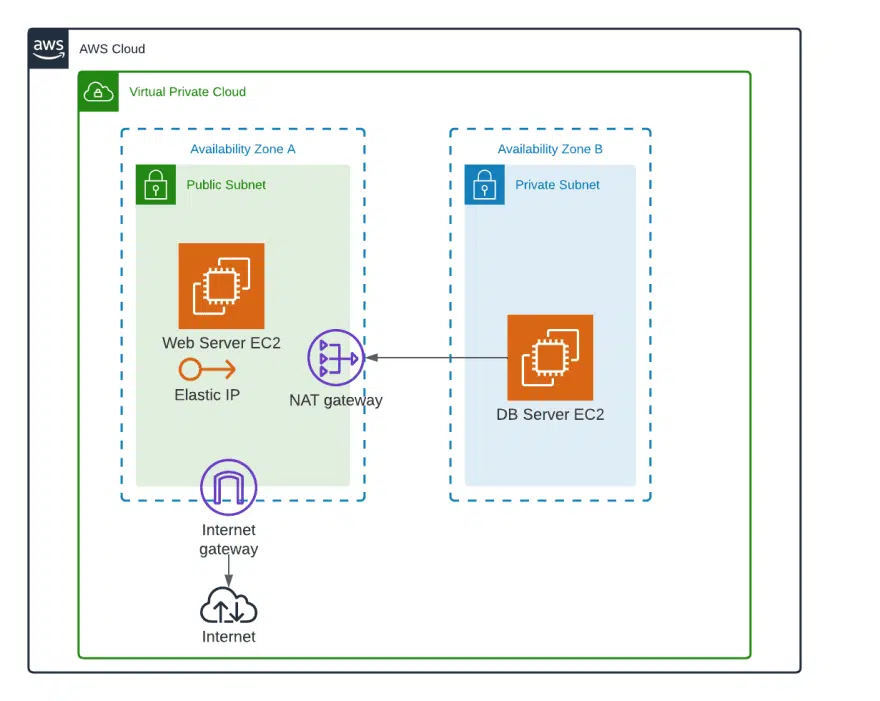
REGIONS

AVAILABILITY

ZONES

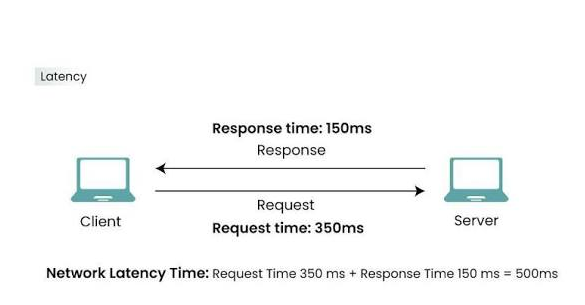
Amazon Virtual Private Cloud (Amazon VPC) :

* Amazon Virtual Private Cloud (Amazon VPC) is a service that allows you to create a logically isolated virtual network within the AWS cloud.
* Within your VPC, you can define your own virtual network environment and launch AWS resources, such as Amazon Elastic Compute Cloud (EC2) instances. A VPC offers the security, control, and scalability necessary for running your applications on AWS.



Latency :

Latency is the time delay between a user's action and the application's response.



Factors influencing latency :

* Distance: The greater the physical distance between the client and the server, the higher the latency. Services like AWS Local Zones and Content Delivery Networks (CDNs) place resources closer to users to reduce this.
* Network Congestion: High network traffic can slow data transmission, increasing latency, especially during peak hours.
* Network Hops: Each time a data packet passes through a router, it adds a small processing delay. More hops mean higher latency.
* Transmission Medium: The physical medium for transmitting data affects latency. For instance, fiber-optic cables generally have lower latency than wireless connections.
* Server Processing Time: A slow-performing server can be a source of latency, as it takes longer to process requests and generate a response.
* Inefficient Protocols: Some network protocols, such as TCP, require more setup and acknowledgment steps, adding a small amount of delay compared to others like UDP.

How to Measure Latency ?

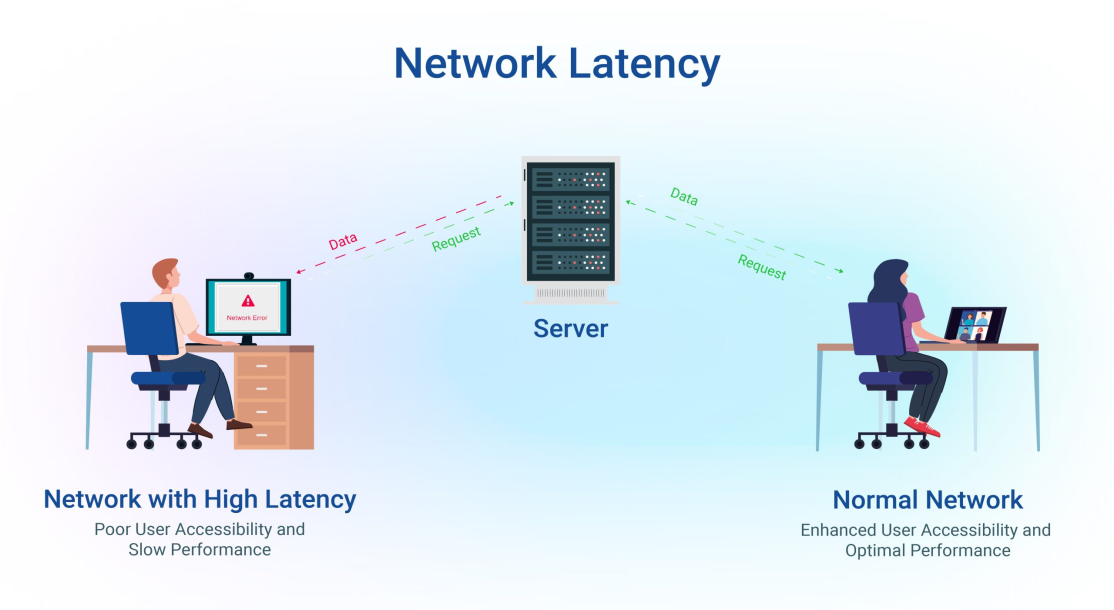
Measure latency: The most common way to measure network latency is with the ping command, which sends a packet to a destination and measures the Round-Trip Time (RTT).

How to Reduce Latency?

* Deploy closer to users
* Use Content Delivery Networks
* Optimize Routing
* Upgrade Hardware
* Choose the right protocol

Network Latency :

Network latency is the time delay for data to travel from its source to its destination across a network.



How to Reduce Network Latency?

* Reduce physical distance
* Optimize routing
* Upgrade hardware
* Use wired connections
* Optimize applications
* Limit network congestion

How to reduce latency in AWS services ?

* Amazon CloudFront: This is a Content Delivery Network (CDN) that caches your static and dynamic content at a global network of over 700 edge locations. When a user requests content, it is served from the closest edge location, which significantly reduces latency.
* AWS Global Accelerator: This service improves performance for a wide range of applications by routing user traffic to the nearest AWS edge location using the highly optimized AWS global network.
* AWS Local Zones: These are infrastructure deployments that place compute, storage, and other select AWS services close to large population and industry centers. They are an extension of a parent AWS region, offering single-digit millisecond latency to local users.
* AWS Wavelength Zones: These embed AWS compute and storage services within 5G network providers' data centers. They are designed for applications requiring ultra-low latency for 5G devices, like real-time gaming or autonomous vehicles.
* Multi-Region or Multi-AZ architecture: You can deploy your applications across multiple AWS regions or Availability Zones (AZs). This enables users to connect to the nearest data center, minimizing latency and providing greater resilience.

Elastic IP Addresses :

An Elastic IP (EIP) is a static, public IPv4 address that you can allocate to your AWS account and associate with a specific resource, like an Amazon EC2 instance.

USERS

ELASTIC IP

APP SERVER

EC2

Default VPC Privacy :

By default when you create a new vpc in aws It’s private. This means that resources within the vpc cannot directly access the internet or be accessed from the internet. This offers a fundamental layer of security by isolating your resources from outside traffic.

There are two primary options for connecting your vpc to the internet :

1. Internet Gateway
2. E Egress-Only Internet Gateway(EIGW)

1.Internet Gateway :

An Internet Gateway is a network component that enables communication between a private network and the public internet.

* Inbound traffic refers to data packets originating from an external source, such as the public internet, that are received by a web server or other resource within your private network.
* Outbound traffic is any data originating from inside a private network that is destined for an external source, such as the public internet.

REGION

VPC

SUBNET

INSTANCES

SUBNET

INSTANCES

SUBNET

INSTANCES



AVAILABILITY ZONES AVAILABILITY ZONES AVAILABILITY ZONES

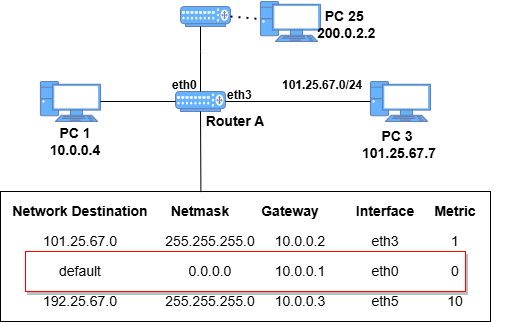
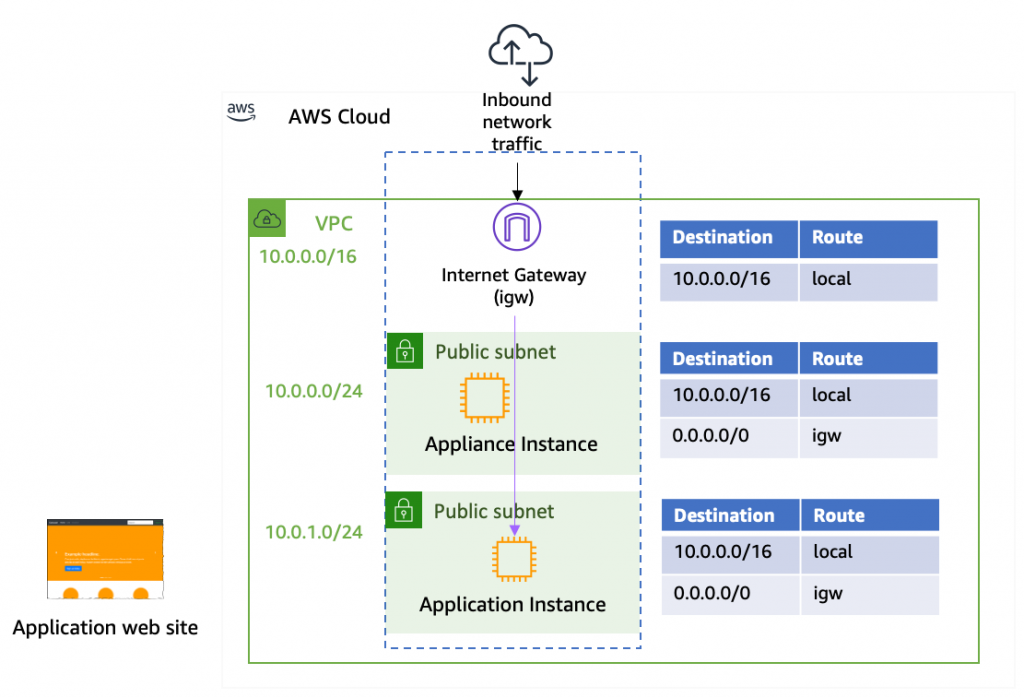
2.E Egress-Only Internet Gateway(EIGW):

An Egress-Only Internet Gateway (EIGW) is a highly available and scalable Amazon VPC component that enables instances in your Virtual Private Cloud (VPC) to initiate secure outbound communication over IPv6 to the internet.

* You create and attach an EIGW to your VPC.
* You update the route table for the private subnet where your IPv6-only instances are located.
* You add a route with the destination ::/0 (representing all IPv6 addresses) and the target set to your new EIGW.

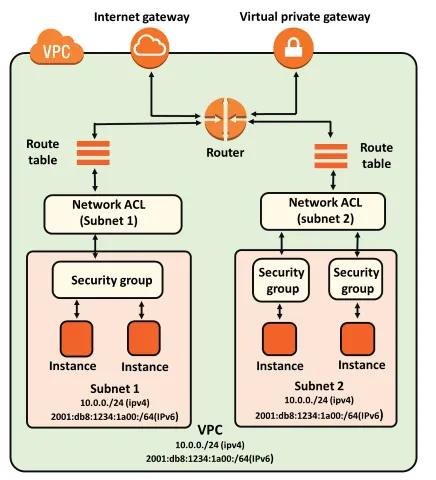
Route Table :

A route table is a set of rules used by routers and network devices to determine the most efficient path for data packets traveling across an Internet Protocol (IP) network. It is essentially a map that directs network traffic, specifying where to send traffic based on its destination IP address.

Network ACLs (NACLs) :

Network Access Control Lists (NACLs) are an optional layer of security in Amazon Web Services (AWS) that function as a virtual firewall for controlling traffic in and out of one or more subnets.



Security Groups :

In AWS, a security group acts as a virtual firewall for your Amazon EC2 instances and other related resources, controlling the traffic that is allowed to reach and leave them.

