**Types of cloud platforms**

**Public clouds** are the most common type of cloud computing deployment. The cloud resources (like servers and storage) are owned and operated by a third-party cloud service provider and delivered over the internet. With a public cloud, all hardware, software and other supporting infrastructure are owned and managed by the cloud provider. Microsoft Azure , AWS ,GCP , is are example of a public cloud.

Advantages of public clouds:

* **Lower costs**—no need to purchase hardware or software and you pay only for the service you use.
* **No maintenance**—your service provider provides the maintenance.
* **Near-unlimited scalability**—on-demand resources are available to meet your business needs.
* **High reliability**—a vast network of servers ensures against failure.

A **private cloud** consists of cloud computing resources used exclusively by one business or organization. The private cloud can be physically located at your organization’s on-site datacenter or it can be hosted by a third-party service provider. But in a private cloud, the services and infrastructure are always maintained on a private network and the hardware and software are dedicated solely to your organization

Advantages of a private cloud:

* **More flexibility**—your organisation can customise its cloud environment to meet specific business needs.
* **More control**—resources are not shared with others, so higher levels of control and privacy are possible.
* **More scalability**—private clouds often offer more scalability compared to on-premises infrastructure.

A **hybrid cloud** platform gives organizations many advantages—such as greater flexibility, more deployment options, security, compliance and getting more value from their existing infrastructure. When computing and processing demand fluctuates, hybrid cloud computing gives businesses the ability to seamlessly scale up their on-premises infrastructure to the public cloud to handle any overflow—without giving third-party datacenters access to the entirety of their data.

.

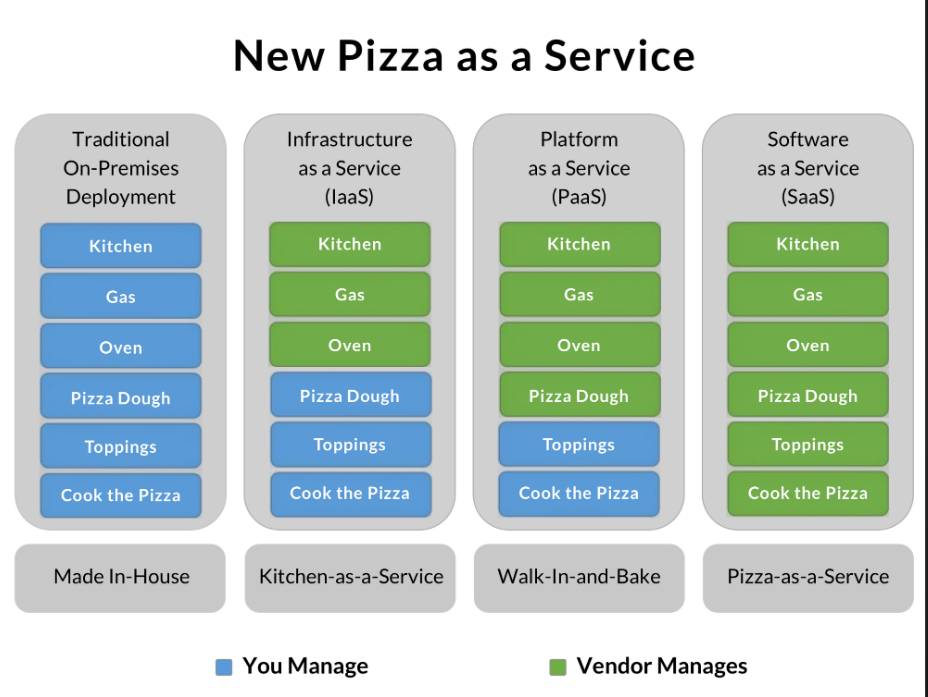
Advantages of the hybrid cloud:

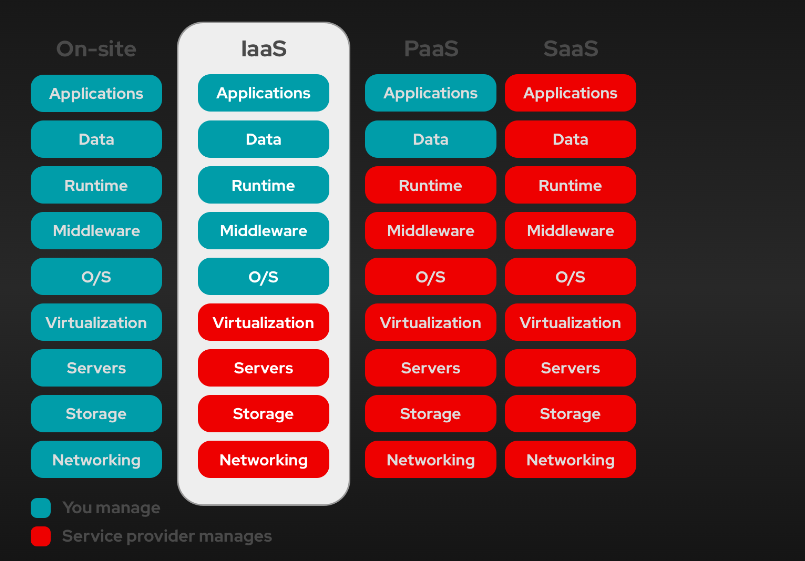
* **Control**—your organisation can maintain a private infrastructure for sensitive assets or workloads that require low latency.
* **Flexibility**—you can take advantage of additional resources in the public cloud when you need them.
* **Cost-effectiveness**—with the ability to scale to the public cloud, you pay for extra computing power only when needed.
* **Ease**—transitioning to the cloud does not have to be overwhelming because you can migrate gradually—phasing in workloads over time.

**Virtualization will transfer hardware into software**

Managed AWS till recent times was using the Xen hypervisor

IAAS PAAS SAAS





**How is an IP Address Formed?**

An **IP address is just a number** that follows a specific format.

**IPv4 – The Most Common Type of IP Address**

* **An IPv4 address is made up of 32 bits (binary numbers of 0s and 1s).**
* These 32 bits are split into **4 parts (called octets).**
* Each **octet** contains **8 bits**.
* It is written in **decimal format** (so humans can read it easily).

192.168.1.10 :

Octet4 Octet3 octet2 octet1

11000000.10101000.00000001.00001010

Decimal Binary Representation

192 11000000

168 10101000

1 00000001

10 00001010

**Types of IP Addresses**

1. **IPv4 (Internet Protocol version 4)**
   * Formed using **32 bits** (4 octets)
   * Written in **dotted decimal notation** (e.g., 192.168.1.1)
   * Each octet ranges from **0 to 255**
   * Example: 172.16.254.1
2. **IPv6 (Internet Protocol version 6)**
   * Formed using **128 bits** (8 groups of 4 hexadecimal digits)
   * Written in **colon-separated hexadecimal notation** (e.g., 2001:0db8:85a3:0000:0000:8a2e:0370:7334)
   * Allows for a vastly larger number of unique addresses

**Formation of an IP Address**

An IP address consists of two main parts:

1. **Network Portion**: Identifies the network to which the device belongs.
2. **Host Portion**: Identifies the specific device (host) within that network.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Class** | **Starting IP** | **Ending IP** | **First Octet Range** | **Usage** |
| **Class A** | 1.0.0.0 | 126.255.255.255 | 1 - 127 | Large networks (ISPs, big companies) |
| **Class B** | 128.0.0.0 | 191.255.255.255 | 128 - 191 | Medium-sized businesses |
| **Class C** | 192.0.0.0 | 223.255.255.255 | 192 - 223 | Small businesses & home networks |
| **Class D** | 224.0.0.0 | 239.255.255.255 | 224 - 239 | Multicast communication |
| **Class E** | 240.0.0.0 | 255.255.255.255 | 240 - 255 | Experimental & research |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **IP Class** | **IP Range** | **Subnet Mask** | **Network Part** | **Host Part** | **Example** |
| Class A | 1.0.0.0 - 126.255.255.255 | 255.0.0.0 (/8) | First 8 bits (1st octet) | Remaining 24 bits (3 octets) | 10.5.25.3 (Network: 10, Host: 5.25.3) |
| Class B | 128.0.0.0 - 191.255.255.255 | 255.255.0.0 (/16) | First 16 bits (2 octets) | Remaining 16 bits (2 octets) | 172.16.5.25 (Network: 172.16, Host: 5.25) |
| Class C | 192.0.0.0 - 223.255.255.255 | 255.255.255.0 (/24) | First 24 bits (3 octets) | Remaining 8 bits (1 octet) | 192.168.1.50 (Network: 192.168.1, Host: 50) |
| Class D (Multicast) | 224.0.0.0 - 239.255.255.255 | N/A | N/A | N/A | N/A (Used for multicast communication) |
| Class E (Experimental) | 240.0.0.0 - 255.255.255.255 | N/A | N/A | N/A | N/A (Reserved for experimental use) |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Class** | **IP Range** | **Default Subnet Mask** | **Network Bits** | **Host Bits** | **Number of Networks** | **Number of Hosts per Network** |
| **Class A** | 1.0.0.0 - 126.255.255.255 | 255.0.0.0 (/8) | 8 bits | 24 bits | 128 | 1,67,77,214 |
| **Class B** | 128.0.0.0 - 191.255.255.255 | 255.255.0.0 (/16) | 16 bits | 16 bits | 16,384 | 65,534 |
| **Class C** | 192.0.0.0 - 223.255.255.255 | 255.255.255.0 (/24) | 24 bits | 8 bits | 20,97,152 | 254 |

**Subnet Masking** is a technique used in networking to divide an IP address into two parts: the **network part** and the **host part**. This process helps in organizing and managing IP addresses more efficiently, particularly in larger networks.

* **Network Part**: Identifies the specific network to which the device belongs.
* **Host Part**: Identifies the specific device (or host) within that network

**Subnet Mask** is a 32-bit number, similar to an IP address, written in the same format (four octets). The **'1' bits** in the subnet mask represent the **network portion**, and the **'0' bits** represent the **host portion**.

**Loopback Address**

* **Definition:** A loopback address is a special IP address used for testing network connections on the local host (your own computer).
* **Value:** The most common loopback address is **127.0.0.1**.

**VPC**

CIDR (Classless Inter-Domain Routing) is a method used to allocate IP addresses and route IP packets. A **CIDR range** is a notation for specifying IP addresses and their associated network mask.

In CIDR notation, an IP address is followed by a slash (/) and a number that represents the subnet mask. The subnet mask defines the size of the network and how many IP addresses are available in the network.

For example:

**CIDR Block**: 172.16.0.0/26

* 172.16.0.0 is the IP address.
* /26 indicates the subnet mask. This means that the first 24 bits are part of the network address, and the remaining 8 bits are available for host addresses within that network.

**Total IP Range**:

* The subnet mask is /26, which means the network has 32 - 26 = 6 bits for hosts.
* The number of IPs in a /26 subnet is 2^6 = 64 IPs.

However, two of these IP addresses are reserved:

* + **Network Address** (the first IP in the range)
  + **Broadcast Address** (the last IP in the range)

So, the **usable IP addresses** are 64 - 2 = 62 IP addresses.

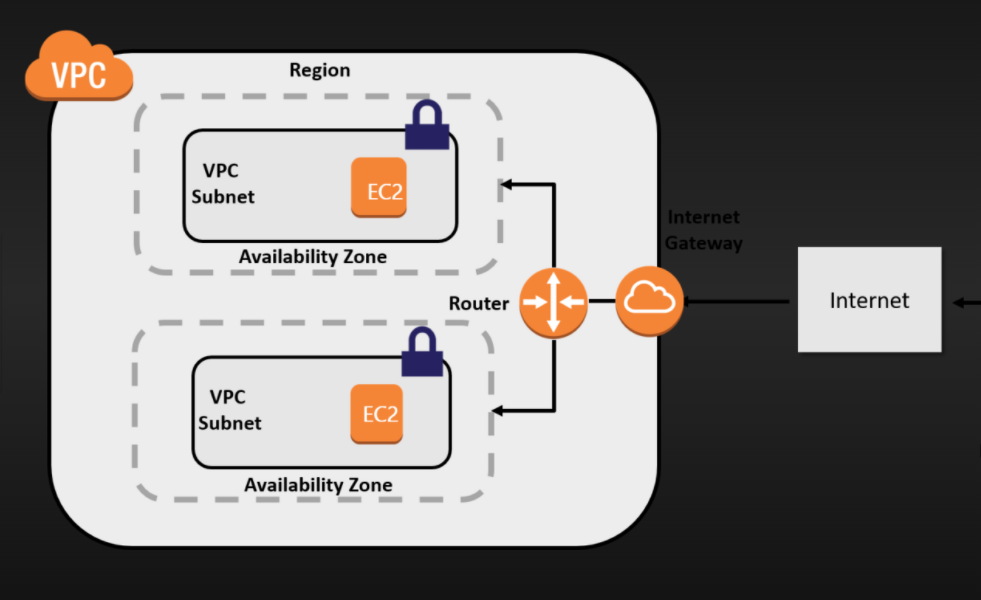
Lets create **8 subnets**, which means we need to borrow 3 more bits from the host portion (2^3 = 8 subnets).

This will give us a **/29** subnet mask (26 + 3 = 29).

**CIDR Block**: 172.16.0.0/26

Subnets:

|  |  |  |  |
| --- | --- | --- | --- |
| **Subnet** | **Network Address** | **Usable IP Range** | **Broadcast Address** |
| **Subnet 1** | 172.16.0.0/29 | 172.16.0.1 to 172.16.0.6 | 172.16.0.7 |
| **Subnet 2** | 172.16.0.8/29 | 172.16.0.9 to 172.16.0.14 | 172.16.0.15 |
| **Subnet 3** | 172.16.0.16/29 | 172.16.0.17 to 172.16.0.22 | 172.16.0.23 |
| **Subnet 4** | 172.16.0.24/29 | 172.16.0.25 to 172.16.0.30 | 172.16.0.31 |
| **Subnet 5** | 172.16.0.32/29 | 172.16.0.33 to 172.16.0.38 | 172.16.0.39 |
| **Subnet 6** | 172.16.0.40/29 | 172.16.0.41 to 172.16.0.46 | 172.16.0.47 |
| **Subnet 7** | 172.16.0.48/29 | 172.16.0.49 to 172.16.0.54 | 172.16.0.55 |
| **Subnet 8** | 172.16.0.56/29 | 172.16.0.57 to 172.16.0.62 | 172.16.0.63 |

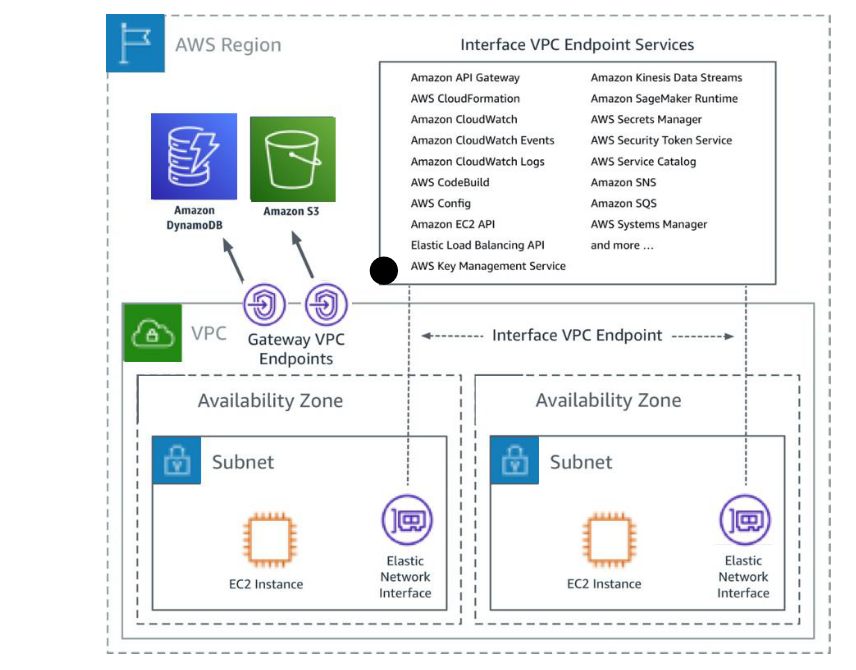


1. **What is VPC (virtual private cloud)?**  
   It is a virtual network dedicated to your AWS account, it logically isolates from other virtual networks in the AWS cloud, and where you can launch your AWS instance.  
   VPC consist of subnets, internet gateway ,nat gateway & Routing table.  
     
   **What is a subnet?**  
   Subnet is a logical subdivision of IP network. The practice of dividing a network into 2 or more networks is called sub netting.
2. **What is Route table?**  
   A set of rules called routes that are used to determine where network traffic is directed.
3. **What is Internet gateway?**  
   A gateway that you attach to your VPC to enable communication between resources in your VPC and the internet.
4. **What is NAT gateway?**

NAT Gateway is used to connect to the Internet from instances within a private subnet in the VPC.

Nat gateway will be created from public subnet and attach to the private subnet.

1. **What is VPC endpoint?**  
   It enables you to privately connect your VPC to supported AWS services and VPC endpoint services powered by private link without requiring an internet gateway, NAT device, VPN connection or AWS direct connection.

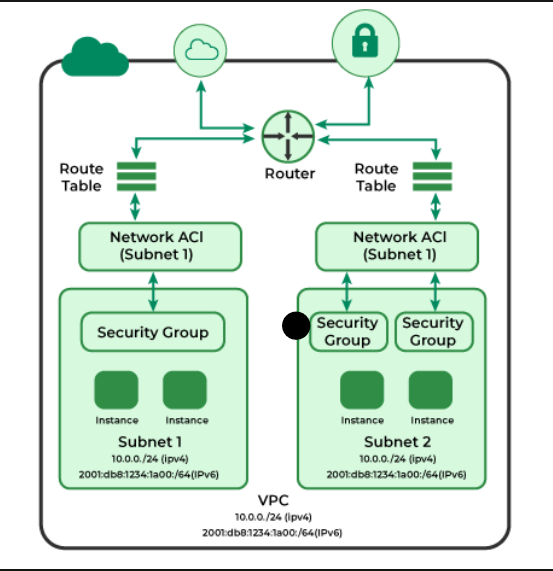


1. **What is NACL (Network ACL)? Network Access Control Lists**  
   NACL are firewall at the subnet level.

* It is a stateless means that any changes made in the inbound rule will not reflect the outbound rule, i.e., you need to add the outbound rule separately. For example, if you add an inbound rule port number 80, then you also have to explicitly add the outbound rule.

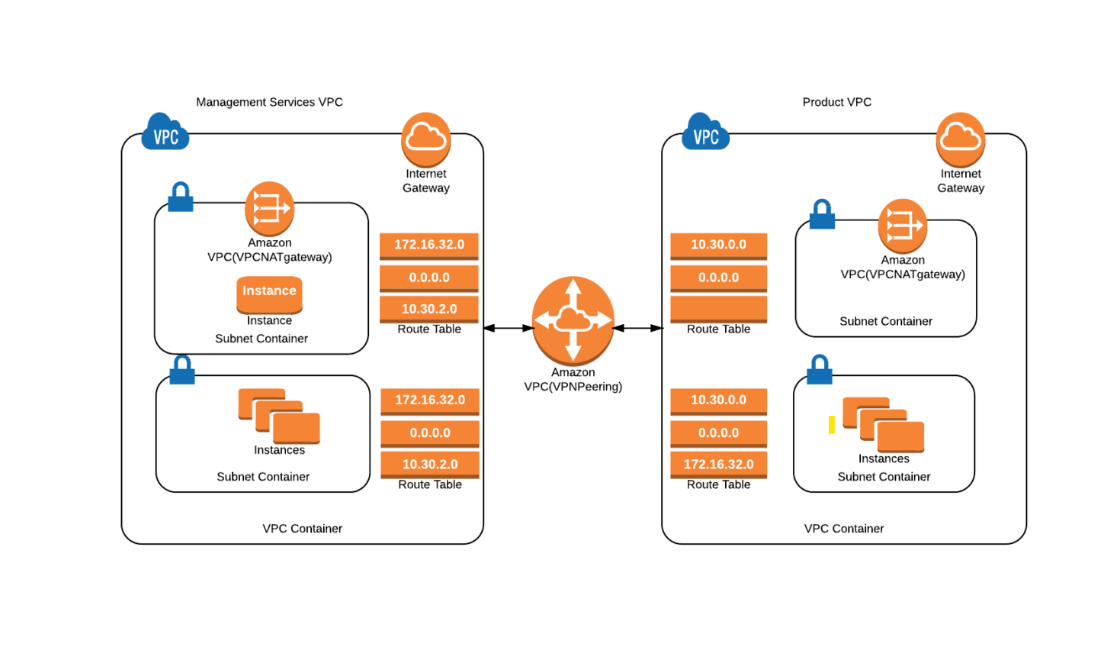
1. **What is a security group?**  
   A security group acts as a virtual firewall for your instance to control inbound and outbound traffic.  
   When you launch an instance in a VPC, you can assign up to 5 security groups to the instance.  
   Security group act at instance level and not at subnet level.

* It is a stateful means that any changes made in the inbound rule will be automatically reflected in the outbound rule. For example, If you are allowing an incoming port 80, then you also have to add the outbound rule explicitly.



1. **Difference between public subnet and private subnet**  
   Public Subnet – Users can access resources from the internet. Internet traffic is routed via internet gateway. Applications are stored in public subnet.  
     
   Private Subnet - Users cannot access resources from the internet. Internet traffic is routed via NAT gateway. Data is stored in private subnet (database, API calls)

VPC PEERING:



**A VPC peering** connection is a networking connection between two VPCs that enables you to route traffic between them using private IPv4 addresses or IPv6 addresses. Instances in either VPC can communicate with each other as if they are within the same network. You can create a VPC peering connection between your own VPCs, or with a VPC in another AWS account.

Types of VPC peering

* [VPCs Peering same account and Region](https://docs.aws.amazon.com/vpc/latest/peering/create-vpc-peering-connection.html#same-account-same-region)
* [VPCs peering same account and different Regions](https://docs.aws.amazon.com/vpc/latest/peering/create-vpc-peering-connection.html#same-account-different-region)
* [VPCs peering different accounts and the same Region](https://docs.aws.amazon.com/vpc/latest/peering/create-vpc-peering-connection.html#different-account-same-region)
* [VPCs peering different accounts and Regions](https://docs.aws.amazon.com/vpc/latest/peering/create-vpc-peering-connection.html#different-account-different-region)

**1. General Conditions**

* Both VPCs must have **non-overlapping CIDR** blocks.
* VPC peering is **not transitive** (i.e., if VPC A is peered with VPC B and VPC B is peered with VPC C, A cannot communicate with C unless an explicit peering connection is established).

✅ **Allowed Peering (Non-Overlapping CIDRs)**

* **VPC A**: 10.0.0.0/16
* **VPC B**: 192.168.1.0/24

❌ **Not Allowed Peering (Overlapping CIDRs)**

* **VPC A**: 10.0.0.0/16
* **VPC B**: 10.0.1.0/24 (overlaps with 10.0.0.0/16)

VPC PEERING LIMITATIONS:

There is a quota on the number of active and pending VPC peering connections per VPC. For more information

* You cannot have more than one VPC peering connection between two VPCs at the same time.
* Any tags that you create for your VPC peering connection are only applied in the account or Region in which you create them.
* You cannot create a VPC peering connection between VPCs that have matching or overlapping IPv4 CIDR blocks.
* You cannot create a VPC peering connection between VPCs that have matching or overlapping IPv6 CIDR blocks.
* VPC peering does not support transitive peering relationships. For example, if there are VPC peering connections between VPC A and VPC B, and between VPC A and VPC C, you can't route traffic from VPC B to VPC C through VPC A. To route traffic between VPC B and VPC C, you must create a VPC peering connection between them.
* You cannot create a security group rule that references a peer VPC security group

**VPC Endpoints (AWS) - Definition**

A **VPC Endpoint** allows secure and private connectivity between **AWS services** and your **VPC** without using the **internet** or a **NAT gateway**.

**Types of VPC Endpoints**

| **Type** | **Description** | **Example Services** |
| --- | --- | --- |
| **Interface Endpoint** | Uses **AWS PrivateLink** to connect to AWS services via an **ENI (Elastic Network Interface)** in your VPC. | S3, DynamoDB, SNS, SQS, API Gateway |
| **Gateway Endpoint** | Uses a **gateway** to route traffic to supported services **without internet access**. | Amazon S3, DynamoDB |
| **Gateway Load Balancer Endpoint** | Provides **private access** to third-party virtual appliances in your VPC via a **GWLB**. | Firewalls, Intrusion Detection Systems |

|  |  |  |
| --- | --- | --- |
| **Name** | **Default** | **Adjustable** |
| Active VPC peering connections per VPC | 50 | [Yes](https://console.aws.amazon.com/servicequotas/home/services/vpc/quotas/L-7E9ECCDB)  (up to 125) |
| Outstanding VPC peering connection requests | 25 | [Yes](https://console.aws.amazon.com/servicequotas/home/services/vpc/quotas/L-DC9F7029) |
| Expiry time for an unaccepted VPC peering connection request | 1 week (168 hours) | No |

In AWS, "user data" refers to a set of instructions or scripts provided when launching an EC2 instance, allowing you to automatically configure and customize the instance during its initial boot process, typically used to install software, set up configurations, or run specific commands without manual intervention after launch

1**. Availability Zones (AZs)**

Definition:

An Availability Zone is a physically separate data center within an AWS Region. Each AZ consists of multiple redundant power, networking, and connectivity resources

Key Features:

• Multiple AZs exist within a single AWS region.

• Each AZ is physically isolated from other AZs in the same region.

• Connected via high-speed, low-latency private networking.

• Helps in fault tolerance and high availability of applications.

2. **Edge Locations**

Definition:

An Edge Location is a global content delivery endpoint that AWS uses to cache data and reduce latency for users.

Key Features:

• Edge locations are not part of AWS Regions.

• They are closer to end users than Availability Zones.

• Used primarily by Amazon Cloud Front (CDN) and AWS Global Accelerator.

• Improves content delivery speed and reduces latency.

As of February 2025, Amazon Web Services (AWS) operates 114 Availability Zones across 36 geographic regions worldwide. Additionally, AWS has announced plans to launch 12 more Availability Zones and four more AWS Regions in New Zealand, the Kingdom of Saudi Arabia, Taiwan, and the AWS European Sovereign Cloud

**EC2 classification by payment models:**

1. On demand or capacity reservation:

• These instances work as pay as you go model.

• long time commitment is not required for this instance.

• These are bit costlier

2. Spot instances:

• These are called as bidding instances.

• We can bid these instances as per the requirement.

3. Dedicated hosts:

• Basically, dedicated instances are provided with the hardware configurations.

4. Reserved instances:

• These are utilized for the longer time frame.

• These would be having discounts in terms of pricing because of longer utilization time frame.

• Will get to know platform, tenancy, instance type and payment options for reserved instances.

**Classification of EC2 by configuration:**

GENERAL PURPOSE:

These instances are ideal for applications that use these resources in equal proportions such as web servers and code repositories.

Series: t2, T3, M4 A1.

Compute Optimized:

well suited for High performance computing (HPC), batch processing, ad serving, video encoding, gaming, scientific modelling, distributed analytics, and CPU-based machine learning inference.

Series: c7g , c6g , c6i , c6a

Memory Optimized:

Memory optimized instances are designed to deliver fast performance for workloads that process large data sets in memory.

Series: R6g, R6i, R5

Accelerated Computing:

Machine learning, high performance computing, computational fluid dynamics, computational finance, seismic analysis, speech recognition, autonomous vehicles, and drug discovery.

Series: p4, p3, p2, Dl1.

Storage Optimized:

These instances maximize the number of transactions processed per second (TPS) for I/O intensive and business-critical workloads which have medium size data sets and can benefit from high compute performance and high network throughput such as relational databases (MySQL, MariaDB, and PostgreSQL), and NoSQL databases (KeyDB, Scylla DB, and Cassandra).

Series: l4i, l3, I3N, D2, D3.

**AMI:**

Amazon Machine Image (or AMI) offers an easy and visual mode of launching instances of your virtual machine on the cloud platform.

Additionally, you can use AMI to create multiple instances of different virtual machines when you require instances of different configurations, we can also create our own custom AMI

**Can you illustrate the relationship between an instance and AMI?**

With the help of just a single AMI, you can launch multiple instances and that to even different types.At the same time, an instance type is characterized by the host.

**Can you define EIP?**

EIP stands for Elastic IP address.

It is a static Ipv4 address that is provided by AWS to administer dynamic cloud computing services.

Can you name some AWS services that are not region-specific?

• IAM

• Route 53

• Web application firewall

• CloudFront

• S3 bucket

1. How to take snapshot backup( all volume backup)

2. How to take specific volume backup?

3. How to take ami back up?.

3 In which Sincero have u go through change instance type or upgrade and what r the pre requisites and steps to follow?