

# Amazon Virtual Private Cloud (VPC)

A virtual private cloud (VPC) is a virtual network dedicated to your AWS account.

Analogous to having your own DC inside AWS.

It is logically isolated from other virtual networks in the AWS Cloud.

Provides complete control over the virtual networking environment including selection of IP ranges, creation of subnets, and configuration of route tables and gateways.

You can launch your AWS resources, such as Amazon EC2 instances, into your VPC.

When you create a VPC, you must specify a range of IPv4 addresses for the VPC in the form of a Classless Inter-Domain Routing (CIDR) block; for example, 10.0.0.0/16.

This is the primary CIDR block for your VPC.

A VPC spans all the Availability Zones in the region.

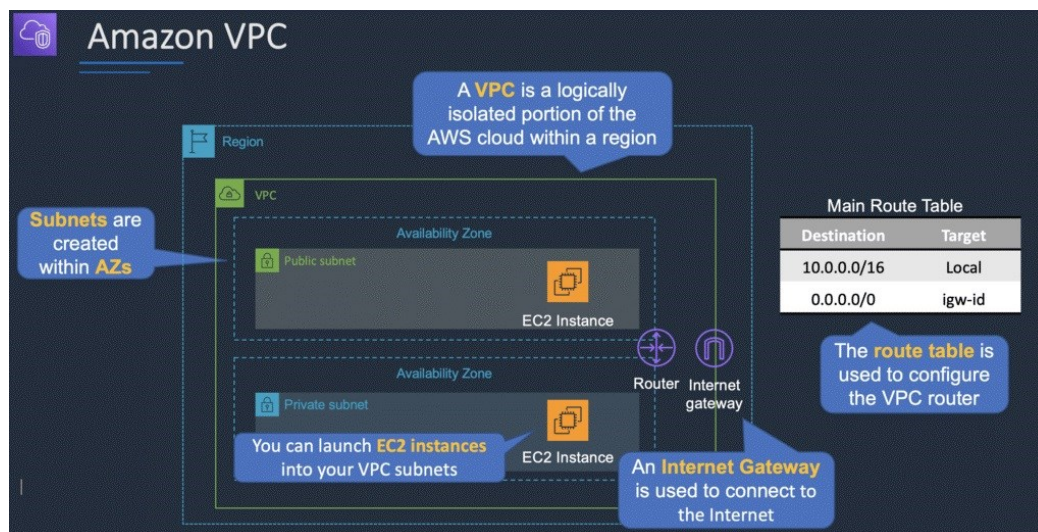
You have full control over who has access to the AWS resources inside your VPC.

You can create your own IP address ranges, and create subnets, route tables and network gateways.

When you first create your AWS account a default VPC is created for you in each AWS region.

A default VPC is created in each region with a subnet in each AZ.

By default, you can create up to 5 VPCs per region.



You can define dedicated tenancy for a VPC to ensure instances are launched on dedicated hardware (overrides the configuration specified at launch).

A default VPC is automatically created for each AWS account the first time Amazon EC2 resources are provisioned.

The default VPC has all-public subnets.

Public subnets are subnets that have:

- “Auto-assign public IPv4 address” set to “Yes”.
- The subnet route table has an attached Internet Gateway.

Instances in the default VPC always have both a public and private IP address.

AZs names are mapped to different zones for different users (i.e. the AZ “ap-southeast-2a” may map to a different physical zone for a different user).

Components of a VPC:

- **A Virtual Private Cloud:** A logically isolated virtual network in the AWS cloud. You define a VPC’s IP address space from ranges you select.
- **Subnet:** A segment of a VPC’s IP address range where you can place groups of isolated resources (maps to an AZ, 1:1).
- **Internet Gateway:** The Amazon VPC side of a connection to the public Internet.
- **NAT Gateway:** A highly available, managed Network Address Translation (NAT) service for your resources in a private subnet to access the Internet.
- **Hardware VPN Connection:** A hardware-based VPN connection between your Amazon VPC and your datacenter, home network, or co-location facility.
- **Virtual Private Gateway:** The Amazon VPC side of a VPN connection.
- **Customer Gateway:** Your side of a VPN connection.
- **Router:** Routers interconnect subnets and direct traffic between Internet gateways, virtual private gateways, NAT gateways, and subnets.
- **Peering Connection:** A peering connection enables you to route traffic via private IP addresses between two peered VPCs.
- **VPC Endpoints:** Enables private connectivity to services hosted in AWS, from within your VPC without using an Internet Gateway, VPN, Network Address Translation (NAT) devices, or firewall proxies.
- **Egress-only Internet Gateway:** A stateful gateway to provide egress only access for IPv6 traffic from the VPC to the Internet.

Options for securely connecting to a VPC are:

- AWS managed VPN – fast to setup.
- Direct Connect – high bandwidth, low-latency but takes weeks to months to setup.
- VPN CloudHub – used for connecting multiple sites to AWS.
- Software VPN – use 3rd party software.

An Elastic Network Interface (ENI) is a logical networking component that represents a NIC.

ENIs can be attached and detached from EC2 instances and the configuration of the ENI will be maintained.

Flow Logs capture information about the IP traffic going to and from network interfaces in a VPC.

Flow log data is stored using Amazon CloudWatch Logs.

Flow logs can be created at the following levels:

- VPC.
- Subnet.
- Network interface.

Peering connections can be created with VPCs in different regions (available in most regions now).

## Subnets

After creating a VPC, you can add one or more subnets in each Availability Zone.

When you create a subnet, you specify the CIDR block for the subnet, which is a subset of the VPC CIDR block.

Each subnet must reside entirely within one Availability Zone and cannot span zones.

Types of subnet:

- If a subnet's traffic is routed to an internet gateway, the subnet is known as a public subnet.
- If a subnet doesn't have a route to the internet gateway, the subnet is known as a private subnet.
- If a subnet doesn't have a route to the internet gateway, but has its traffic routed to a virtual private gateway for a VPN connection, the subnet is known as a VPN-only subnet.

An Internet Gateway is a horizontally scaled, redundant, and highly available VPC component that allows communication between instances in your VPC and the internet.

## Firewalls

Network Access Control Lists (ACLs) provide a firewall/security layer at the subnet level.

Security Groups provide a firewall/security layer at the instance level.

The table below describes some differences between Security Groups and Network ACLs:

Security Group	Network ACL
Operates at the instance (interface) level	Operates at the subnet level
Supports allow rules only	Supports allow and deny rules
Stateful	Stateless
Evaluates all rules	Processes rules in order
Applies to an instance only if associated with a group	Automatically applies to all instances in the subnets its associated with

# VPC Wizard

The VPC Wizard can be used to create the following four configurations:

VPC with a Single Public Subnet:

- Your instances run in a private, isolated section of the AWS cloud with direct access to the Internet.
- Network access control lists and security groups can be used to provide strict control over inbound and outbound network traffic to your instances.
- Creates a /16 network with a /24 subnet. Public subnet instances use Elastic IPs or Public IPs to access the Internet.

VPC with Public and Private Subnets:

- In addition to containing a public subnet, this configuration adds a private subnet whose instances are not addressable from the Internet.
- Instances in the private subnet can establish outbound connections to the Internet via the public subnet using Network Address Translation (NAT).
- Creates a /16 network with two /24 subnets.
- Public subnet instances use Elastic IPs to access the Internet.
- Private subnet instances access the Internet via Network Address Translation (NAT).

VPC with Public and Private Subnets and Hardware VPN Access:

- This configuration adds an IPsec Virtual Private Network (VPN) connection between your Amazon VPC and your data center – effectively extending your data center to the cloud while also providing direct access to the Internet for public subnet instances in your Amazon VPC.
- Creates a /16 network with two /24 subnets.
- One subnet is directly connected to the Internet while the other subnet is connected to your corporate network via an IPsec VPN tunnel.

VPC with a Private Subnet Only and Hardware VPN Access:

- Your instances run in a private, isolated section of the AWS cloud with a private subnet whose instances are not addressable from the Internet.
- You can connect this private subnet to your corporate data center via an IPsec Virtual Private Network (VPN) tunnel.
- Creates a /16 network with a /24 subnet and provisions an IPsec VPN tunnel between your Amazon VPC and your corporate network.

## NAT Instances

NAT instances are managed **by** you.

Used to enable private subnet instances to access the Internet.

When creating NAT instances always disable the source/destination check on the instance.

NAT instances must be in a single public subnet.

NAT instances need to be assigned to security groups.

## NAT Gateways

NAT gateways are managed **for** you by AWS.

NAT gateways are highly available in each AZ into which they are deployed.

They are preferred by enterprises.

Can scale automatically up to 45Gbps.

No need to patch.

Not associated with any security groups.

The table below describes some differences between NAT instances and NAT gateways:

NAT Instance	NAT Gateway
Managed by you (e.g. software updates)	Managed by AWS
Scale up (instance type) manually and use enhanced networking	Elastic scalability up to 45 Gbps
No high availability – scripted/auto-scaled HA possible using multiple NATs in multiple subnets	Provides automatic high availability within an AZ and can be placed in multiple AZs
Need to assign Security Group	No Security Groups
Can use as a bastion host	Cannot access through SSH
Use an Elastic IP address or a public IP address with a NAT instance	Choose the Elastic IP address to associate with a NAT gateway at creation
Can implement port forwarding through manual customisation	Does not support port forwarding

# Direct Connect

AWS Direct Connect is a network service that provides an alternative to using the Internet to connect a customer's on-premise sites to AWS.

Data is transmitted through a private network connection between AWS and a customer's datacenter or corporate network.

Benefits:

- Reduce cost when using large volumes of traffic.
- Increase reliability (predictable performance).
- Increase bandwidth (predictable bandwidth).
- Decrease latency.

Each AWS Direct Connect connection can be configured with one or more virtual interfaces (VIFs).

Public VIFs allow access to public services such as S3, EC2, and DynamoDB.

Private VIFs allow access to your VPC.

From Direct Connect you can connect to all AZs **within the Region**.

You can establish IPSec connections over public VIFs to remote regions.

Direct Connect is charged by port hours and data transfer.

Available in 1Gbps and 10Gbps.

Speeds of 50Mbps, 100Mbps, 200Mbps, 300Mbps, 400Mbps, and 500Mbps can be purchased through AWS Direct Connect Partners.

Each connection consists of a single dedicated connection between ports on the customer router and an Amazon router.

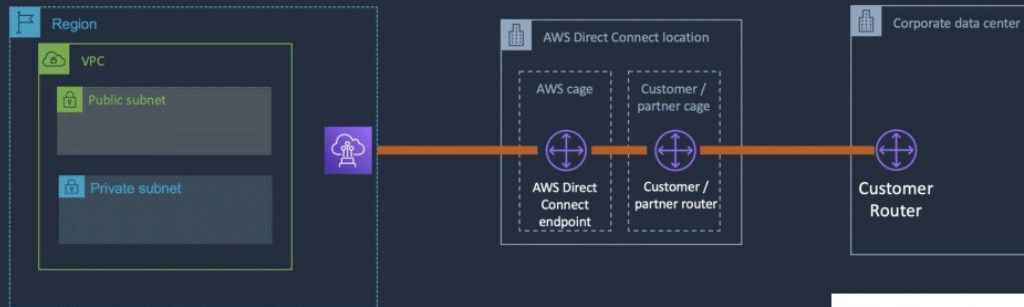
for HA you must have 2 DX connections – can be active/active or active/standby.

Route tables need to be updated to point to a Direct Connect connection.



## AWS Direct Connect

- **Private** connectivity between AWS and your data center / office
- Consistent network experience – increased **speed/latency** & **bandwidth/throughput**
- Lower costs for organizations that transfer **large** volumes of data





# AWS Global Accelerator

AWS Global Accelerator is a service that improves the availability and performance of applications with local or global users.

It provides static IP addresses that act as a fixed entry point to application endpoints in a single or multiple AWS Regions, such as Application Load Balancers, Network Load Balancers or EC2 instances.

Uses the AWS global network to optimize the path from users to applications, improving the performance of TCP and UDP traffic.

AWS Global Accelerator continually monitors the health of application endpoints and will detect an unhealthy endpoint and redirect traffic to healthy endpoints in less than 1 minute.

## Details and Benefits

Uses redundant (two) static anycast IP addresses in different network zones (A and B).

The redundant pair are globally advertized.

Uses AWS Edge Locations – addresses are announced from multiple edge locations at the same time.

Addresses are associated to regional AWS resources or endpoints.

AWS Global Accelerator's IP addresses serve as the frontend interface of applications.

Intelligent traffic distribution: Routes connections to the closest point of presence for applications.

Targets can be Amazon EC2 instances or Elastic Load Balancers (ALB and NLB).

By using the static IP addresses, you don't need to make any client-facing changes or update DNS records as you modify or replace endpoints.

The addresses are assigned to your accelerator for as long as it exists, even if you disable the accelerator and it no longer accepts or routes traffic.

# AWS Outposts

AWS Outposts is a fully managed service that offers the same AWS infrastructure, AWS services, APIs, and tools to virtually any datacenter, co-location space, or on-premises facility for a truly consistent hybrid experience.

AWS Outposts is ideal for workloads that require low latency access to on-premises systems, local data processing, data residency, and migration of applications with local system interdependencies.

AWS compute, storage, database, and other services run locally on Outposts, and you can access the full range of AWS services available in the Region to build, manage, and scale your on-premises applications using familiar AWS services and tools.

Outposts is available as a 42U rack that can scale from 1 rack to 96 racks to create pools of compute and storage capacity.

Services you can run on AWS Outposts include:

- Amazon EC2.
- Amazon EBS.
- Amazon S3.
- Amazon VPC.
- Amazon ECS/EKS.
- Amazon RDS.
- Amazon EMR.