AWS Solution Architect Associate Certification Training – Module 11

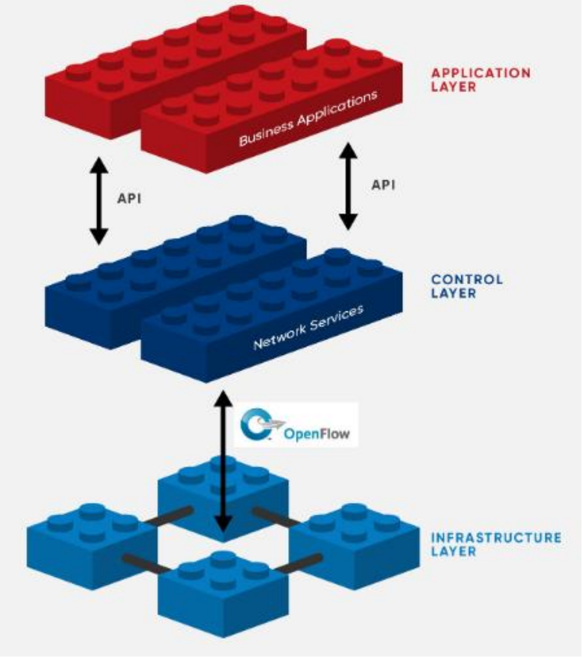
11. Virtual Private Cloud (VPC)

Software-defined networking

Software-defined networking technology is an approach to cloud computing that facilitates network management and enables programmatically efficient network configuration in order to improve network performance and monitoring.

What is SDN? The physical separation of the network control plane from the forwarding plane, and where a control plane controls several devices.

Software-Defined Networking (SDN) is an emerging architecture that is dynamic, manageable, cost-effective, and adaptable, making it ideal for the high-bandwidth, dynamic nature of today's applications. This architecture decouples the network control and forwarding functions enabling the network control to become directly programmable and the underlying infrastructure to be abstracted for applications and network services.



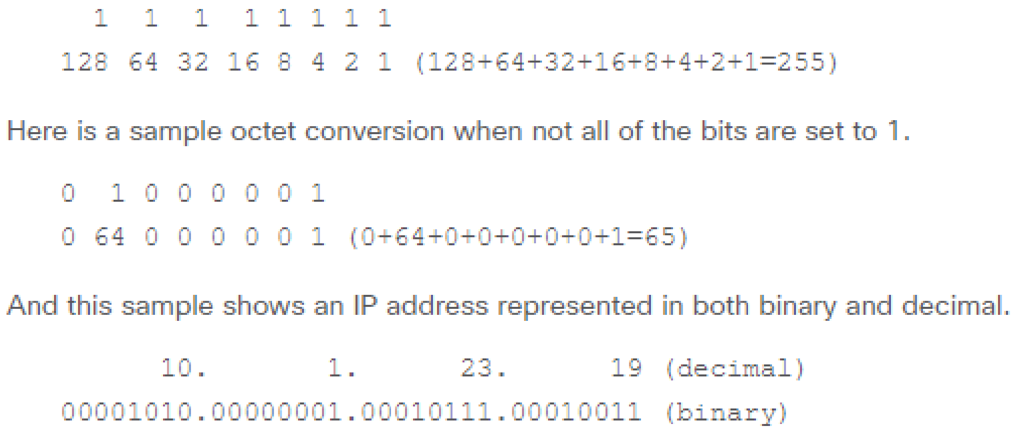
Internet Protocols

The Internet Protocol (IP) is the method or protocol by which data is sent from one computer to another on the Internet. Each computer (known as a host) on the Internet has at least one IP address that uniquely identifies it from all other computers on the Internet.

IP Address: An IP address, or simply an "IP," is a unique address that identifies a device on the Internet or a local network. It allows a system to be recognized by other systems connected via the Internet protocol.

The address is made up of 32 binary bits, which can be divisible into a network portion and host portion with the help of a subnet mask. The 32 binary bits are broken into four octets (1 octet = 8 bits). Each octet is converted to decimal and separated by a period (dot). For this reason, an IP address is said to be expressed in dotted decimal format (for example, 172.16.81.100). The value in each octet ranges from 0 to 255 decimal, or 00000000 - 11111111 binary.

Here is how binary octets convert to decimal:



```
  1   1   1   1 1 1 1 1
128  64  32  16 8 4 2 1 (128+64+32+16+8+4+2+1=255)
```

Here is a sample octet conversion when not all of the bits are set to 1.

```
0  1 0 0 0 0 0 1
0 64 0 0 0 0 0 1 (0+64+0+0+0+0+0+1=65)
```

And this sample shows an IP address represented in both binary and decimal.

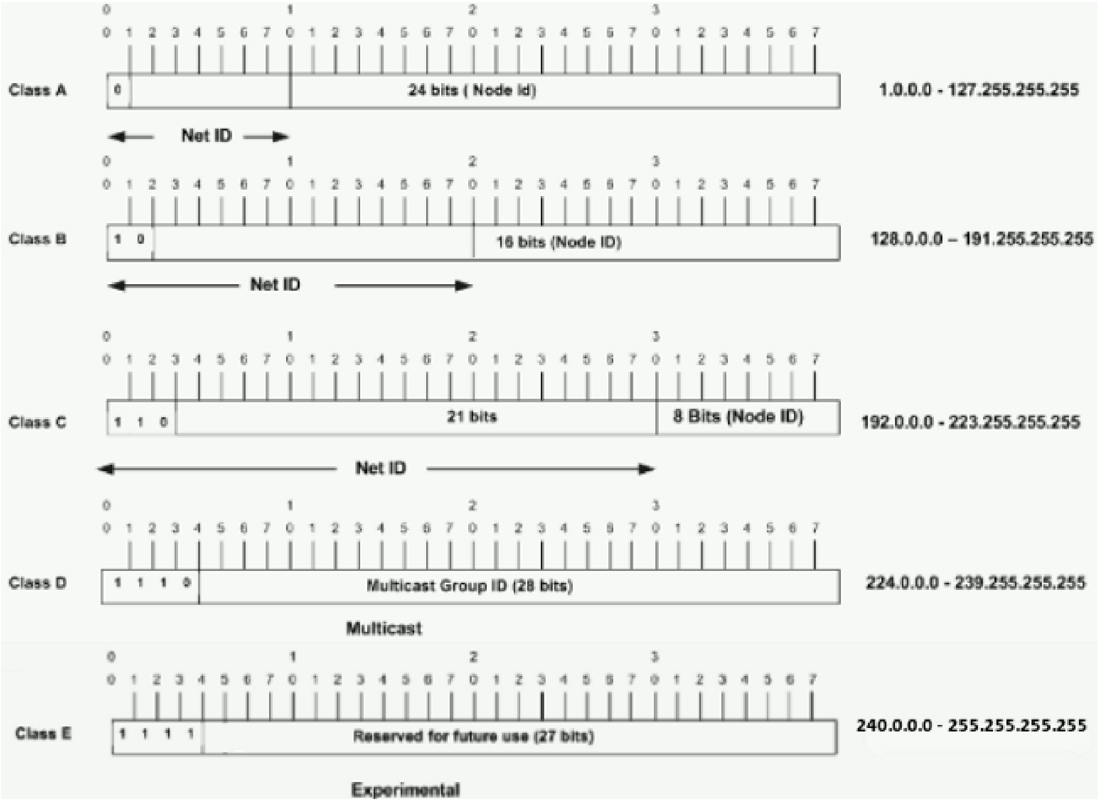
```
      10.       1.      23.      19 (decimal)
00001010.00000001.00010111.00010011 (binary)
```

IPv4

An IPv4 address consist of four sets of numbers from 0 to 255, separated by three dots. For example, the IP address of TechTerms.com is 67.43.14.98. This number is used to identify the TechTerms website on the Internet. When you visit http://techterms.com in your web browser, the DNS system automatically translates the domain name "techterms.com" to the IP address "67.43.14.98."

There are three classes of IPv4 address sets that can be registered through the InterNIC. The smallest is Class C, which consists of 256 IP addresses (e.g. 123.123.123.xxx — where xxx is 0 to 255). The next largest is Class B, which contains 65,536 IP addresses (e.g. 123.123.xxx.xxx). The largest block is Class A, which contains 16,777,216 IP addresses (e.g. 123.xxx.xxx.xxx).

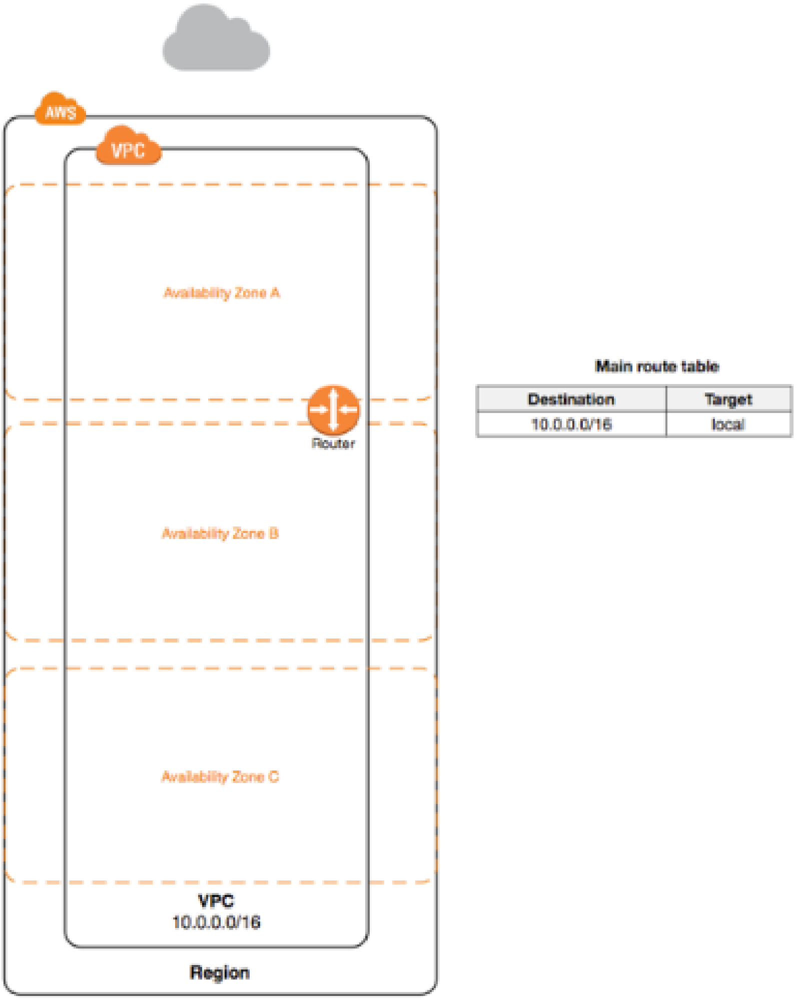
The total number of IPv4 addresses ranges from 000.000.000.000 to 255.255.255.255. Because 256 = 28, there are 28 x 4 or 4,294,967,296 possible IP addresses. While this may seem like a large number, it is no longer enough to cover all the devices connected to the Internet around the world. Therefore, many devices now use IPv6 addresses.



Overview of AWS VPC

Amazon Virtual Private Cloud (Amazon VPC) enables you to launch AWS resources into a virtual network that you've defined. This virtual network closely resembles a traditional network that you'd operate in your own data center, with the benefits of using the scalable infrastructure of AWS. A *virtual private cloud* (VPC) is a virtual network dedicated to your AWS account. It is logically isolated from other virtual networks in the AWS Cloud. You can launch your AWS resources, such as Amazon EC2 instances, into your VPC. You can specify an IP address range for the VPC, add subnets, associate security groups, and configure route tables.

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. 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. Availability Zones are distinct locations that are engineered to be isolated from failures in other Availability Zones. By launching instances in separate Availability Zones, you can protect your applications from the failure of a single location. We assign a unique ID to each subnet.

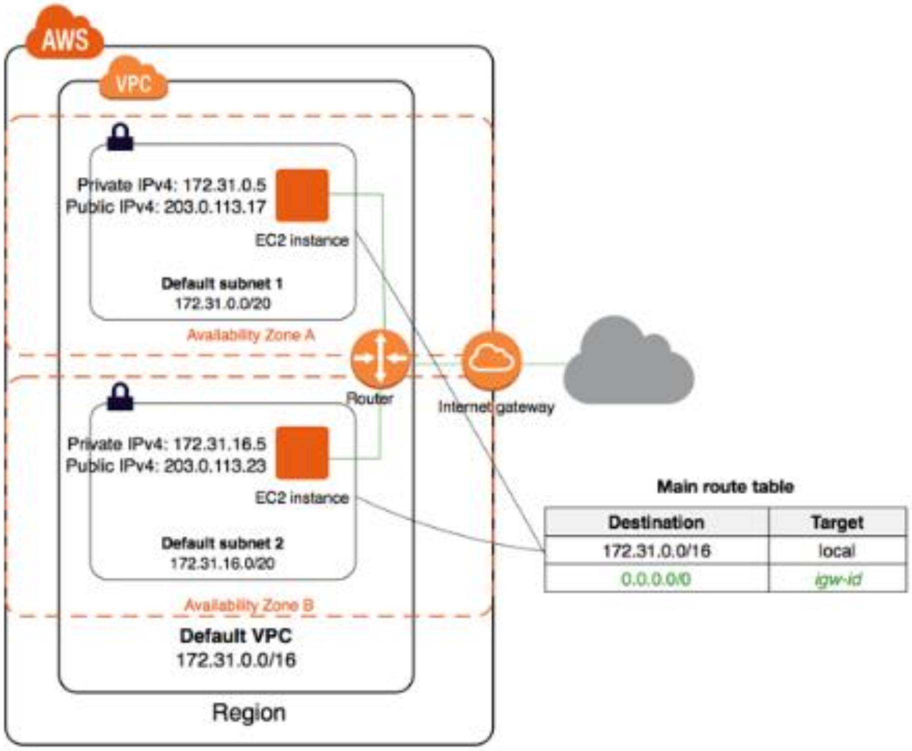


Main route table

Destination	Target
10.0.0.0/16	local

Accessing the Internet

You control how the instances that you launch into a VPC access resources outside the VPC.Your default VPC includes an internet gateway, and each default subnet is a public subnet. Each instance that you launch into a default subnet has a private IPv4 address and a public IPv4 address. These instances can communicate with the internet through the internet gateway. An internet gateway enables your instances to connect to the internet through the Amazon EC2 network edge.



Main route table

Destination	Target
172.31.0.0/16	local
0.0.0.0/0	igw-id

Default Vs. Custom VPC

If your account supports the EC2-VPC platform only, it comes with a *default VPC* that has a *default subnet* in each Availability Zone. A default VPC has the benefits of the advanced features provided by EC2-VPC, and is ready for you to use. If you have a default VPC and don't specify a subnet when you launch an instance, the instance is launched into your default VPC. You can launch instances into your default VPC without needing to know anything about Amazon VPC.

Regardless of which platforms your account supports, you can create your own VPC, and configure it as you need. This is known as a non-default *VPC*. Subnets that you create in your non-default VPC and additional subnets that you create in your default VPC are called non-default *subnets*.

Components of VPC – Subnets, Route Tables, Gateways, Routes

A subnet is a range of IP addresses in your VPC. You can launch AWS resources into a specified subnet. Use a public subnet for resources that must be connected to the internet, and a private subnet for resources that won't be connected to the internet.

To protect the AWS resources in each subnet, you can use multiple layers of security, including security groups and network access control lists (ACL).

Route Tables

A route table contains a set of rules, called routes, that are used to determine where network traffic is directed.

Each subnet in your VPC must be associated with a route table; the table controls the routing for the subnet. A subnet can only be associated with one route table at a time, but you can associate multiple subnets with the same route table.

Internet Gateway

An Internet Gateway is a horizontally scaled, redundant, and highly available VPC component that allows communication between instances in your VPC and the internet. It therefore imposes no availability risks or bandwidth constraints and network traffic.

An Internet gateway servers two purposes: to provide a target in your VPC route tables for internet-routable traffic, and to perform network address translation (NAT) for instances that have been assigned public IPV4 addresses.

An internet gateway supports IPV4 and IPV6 traffic.

Security Groups, Network ACL's

A network access control list (ACL) is an optional layer of security for your VPC that acts as a firewall for controlling traffic in and out of one or more subnets. You might set up network ACLs with rules similar to your security groups in order to add an additional layer of security to your VPC.

Network ACL basics

The following are the basic things that you need to know about network ACLs:

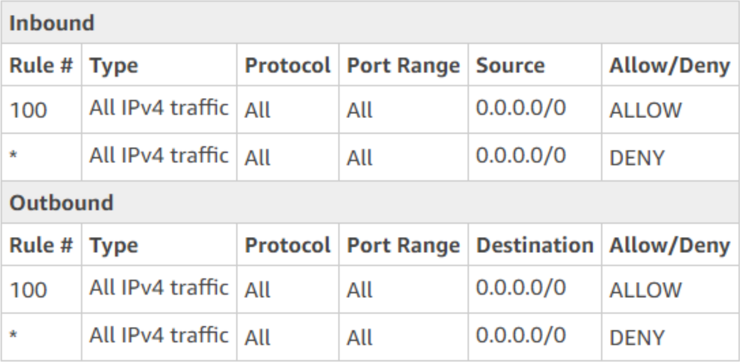
- Your VPC automatically comes with a modifiable default network ACL. By default, it allows all inbound and outbound IPv4 traffic and, if applicable, IPv6 traffic.
- You can create a custom network ACL and associate it with a subnet. By default, each custom network ACL denies all inbound and outbound traffic until you add rules.
- Each subnet in your VPC must be associated with a network ACL. If you don't explicitly associate a subnet with a network ACL, the subnet is automatically associated with the default network ACL.
- You can associate a network ACL with multiple subnets; however, a subnet can be associated with only one network ACL at a time. When you associate a network ACL with a subnet, the previous association is removed.
- A network ACL contains a numbered list of rules that we evaluate in order, starting with the lowest numbered rule, to determine whether traffic is allowed in or out of any subnet associated with the network ACL. The highest number that you can use for a rule is 32766. We recommend that you start by creating rules in increments (for example, increments of 10 or 100) so that you can insert new rules where you need to later on.
- A network ACL has separate inbound and outbound rules, and each rule can either allow or deny traffic.

- Network ACLs are stateless; responses to allowed inbound traffic are subject to the rules for outbound traffic (and vice versa).

Default Network ACL

The default network ACL is configured to allow all traffic to flow in and out of the subnets with which it is associated. Each network ACL also includes a rule whose rule number is an asterisk. This rule ensures that if a packet doesn't match any of the other numbered rules, it's denied. You can't modify or remove this rule.

The following is an example default network ACL for a VPC that supports IPv4 only.



Inbound					
Rule #	Type	Protocol	Port Range	Source	Allow/Deny
100	All IPv4 traffic	All	All	0.0.0.0/0	ALLOW
*	All IPv4 traffic	All	All	0.0.0.0/0	DENY
Outbound					
Rule #	Type	Protocol	Port Range	Destination	Allow/Deny
100	All IPv4 traffic	All	All	0.0.0.0/0	ALLOW
*	All IPv4 traffic	All	All	0.0.0.0/0	DENY

Security Groups

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 five security groups to the instance. Security groups act at the instance level, not the subnet level. Therefore, each instance in a subnet in your VPC could be assigned to a different set of security groups. If you don't specify a particular group at launch time, the instance is automatically assigned to the default security group for the VPC.

For each security group, you add rules that control the inbound traffic to instances, and a separate set of rules that control the outbound traffic.

Security Groups Basics

The following are the basic characteristics of security groups for your VPC:

- You have limits on the number of security groups that you can create per VPC, the number of rules that you can add to each security group, and the number of security groups you can associate with a network interface.
- You can specify allow rules, but not deny rules.

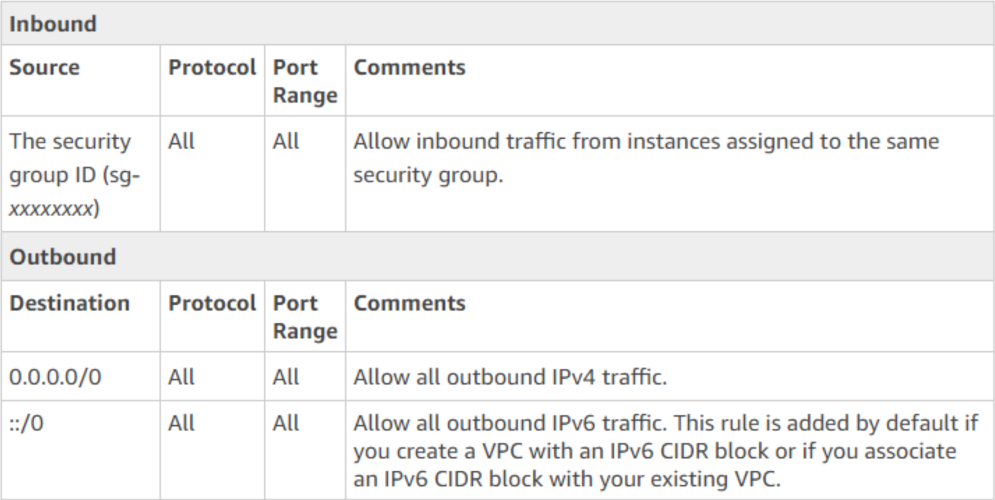
- You can specify separate rules for inbound and outbound traffic.
- When you create a security group, it has no inbound rules. Therefore, no inbound traffic originating from another host to your instance is allowed until you add inbound rules to the security group.
- By default, a security group includes an outbound rule that allows all outbound traffic. You can remove the rule and add outbound rules that allow specific outbound traffic only. If your security group has no outbound rules, no outbound traffic originating from your instance is allowed.
- Security groups are stateful — if you send a request from your instance, the response traffic for that request is allowed to flow in regardless of inbound security group rules. Responses to allowed inbound traffic are allowed to flow out, regardless of outbound rules.
- When you create a security group, you must provide it with a name and a description. The following rules apply:

  i) Names and descriptions can be up to 255 characters in length.
  ii) Names and descriptions are limited to the following characters: a-z, A-Z, 0-9, spaces, and ._-:/()#,@[]+=&;{}!$*.
  iii) A security group name cannot start with sg-.
  iv) A security group name must be unique within the VPC.

Default Security Group for Your VPC

Your VPC automatically comes with a default security group. Each EC2 instance that you launch in your VPC is automatically associated with the default security group if you don't specify a different security group when you launch the instance.

The following table describes the default rules for a default security group.



Inbound			
Source	**Protocol**	**Port Range**	**Comments**
The security group ID (sg-*xxxxxxxx*)	All	All	Allow inbound traffic from instances assigned to the same security group.
Outbound			
Destination	**Protocol**	**Port Range**	**Comments**
0.0.0.0/0	All	All	Allow all outbound IPv4 traffic.
::/0	All	All	Allow all outbound IPv6 traffic. This rule is added by default if you create a VPC with an IPv6 CIDR block or if you associate an IPv6 CIDR block with your existing VPC.

Public and Private Subnets

The instances in the public subnet can send outbound traffic directly to the Internet, whereas the instances in the private subnet can't. Instead, the instances in the private subnet can access the Internet by using a network address translation (NAT) gateway that resides in the public subnet.

Network Address Translation

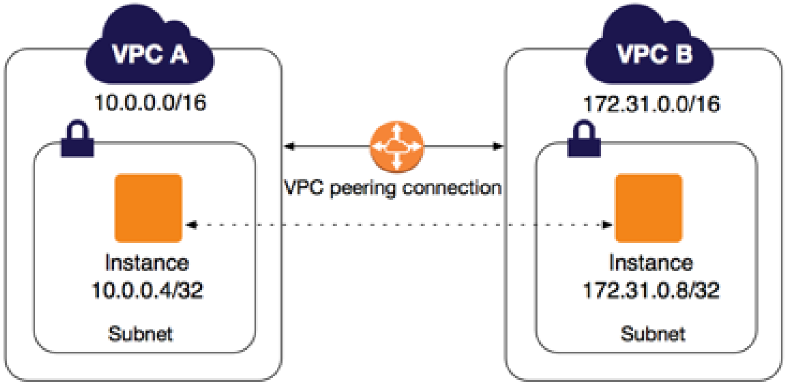
You can use a NAT device to enable instances in a private subnet to connect to the internet (for example, for software updates) or other AWS services, but prevent the internet from initiating connections with the instances. A NAT device forwards traffic from the instances in the private subnet to the internet or other AWS services, and then sends the response back to the instances. When traffic goes to the internet, the source IPv4 address is replaced with the NAT device's address and similarly, when the response traffic goes to those instances, the NAT device translates the address back to those instances' private IPv4 addresses.

NAT devices are not supported for IPv6 traffic

AWS offers two kinds of NAT devices—a *NAT gateway* or a *NAT instance*. We recommend NAT gateways, as they provide better availability and bandwidth over NAT instances. The NAT Gateway service is also a managed service that does not require your administration efforts. A NAT instance is launched from a NAT AMI. You can choose to use a NAT instance for special purposes.

Introduction to VPC Peering

Amazon Virtual Private Cloud (Amazon VPC) enables you to launch AWS resources into a virtual network that you've defined. 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. The VPCs can be in different regions (also known as an inter-region VPC peering connection).


VPC A
10.0.0.0/16
Instance
10.0.0.4/32
Subnet
VPC peering connection
VPC B
172.31.0.0/16
Instance
172.31.0.8/32
Subnet