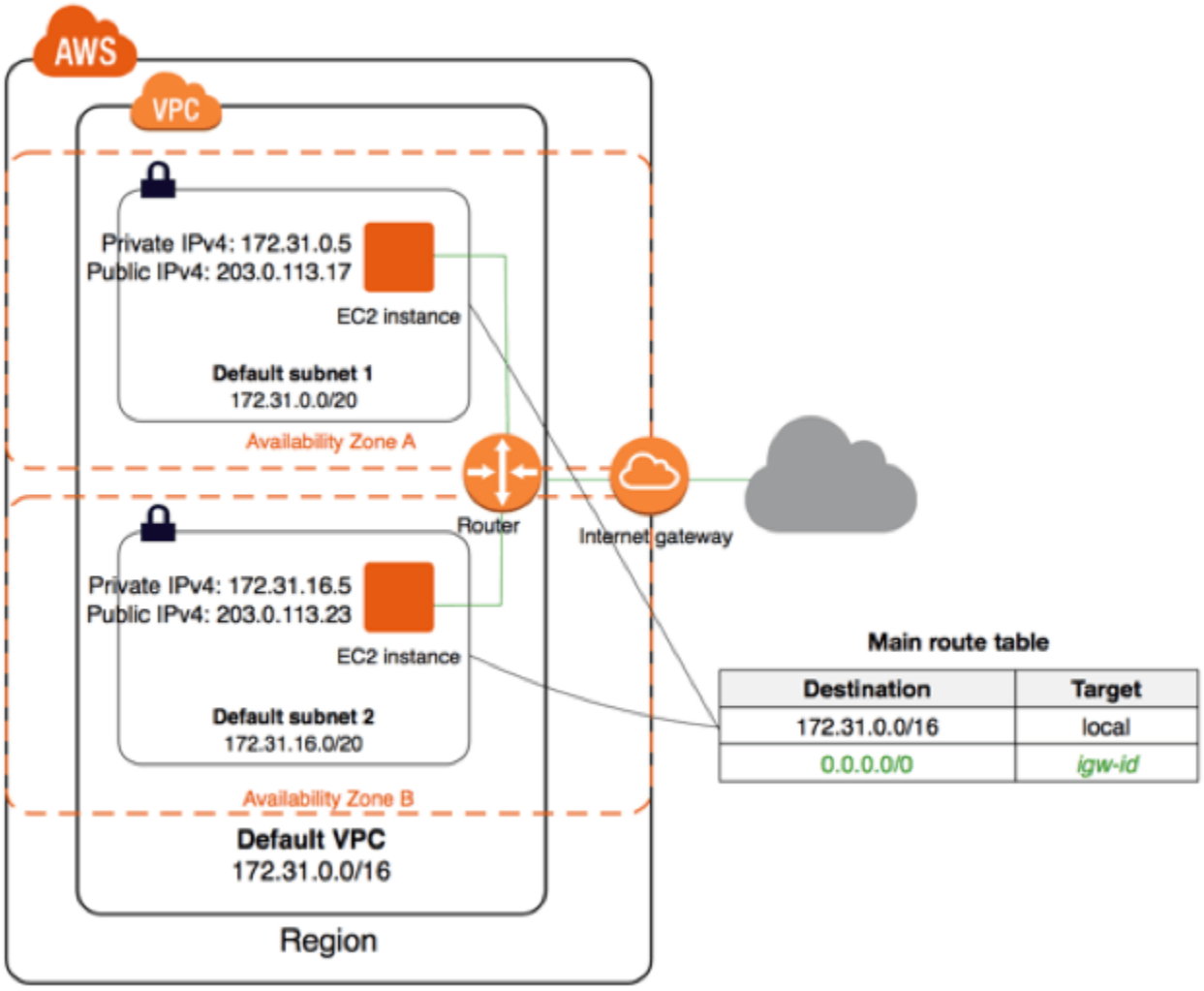
A list of AWS networking services cloud users should know

Understanding the Default Virtual Private Cloud



I walked through the basics of the AWS Global Infrastructure in class.

Understanding these concepts is foundational for understanding the AWS Virtual Private Cloud (VPC) and how it enables advanced networking capabilities for your AWS resources.

A VPC is a logically isolated virtual network, spanning an entire AWS Region, where your EC2 instances are launched.

A VPC is primarily concerned with enabling the following capabilities:

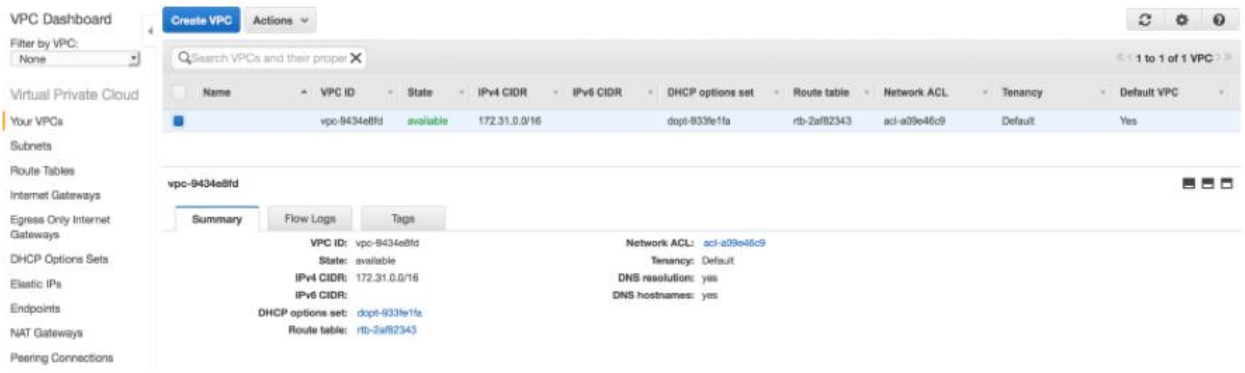
- Isolating your AWS resources from other accounts
- Routing network traffic to and from your instances
- Protecting your instances from network intrusion

There are six core components which are fundamental to a VPC and will be created by a user or by AWS as part of a default VPC. These components are:

- VPC CIDR Block
- Subnet
- Gateways
- Route Table
- Network Access Control Lists (ACLs)
- Security Group
- VPC Pairing

VPC CIDR block

Select "Your VPCs" in the left sidebar and the dashboard will display all your VPCs in a particular AWS Region, including the default VPC. A region can only have one default VPC. Although you can have up to five VPCs in a region, only the initial VPC that AWS creates for you can be the default VPC. As of now, we can create DEFAULT VPC as we wish



Every VPC is associated with an IP address range that is part of a **Classless Inter-Domain Routing (CIDR)** block which will be used to allocated **private IP addresses to EC2 instances**.

AWS recommends that VPCs use private ranges that are defined in RFC 1918.

3. Private Address Space

```
The Internet Assigned Numbers Authority (IANA) has reserved the
following three blocks of the IP address space for private internets:

  10.0.0.0        -   10.255.255.255  (10/8 prefix)
  172.16.0.0      -   172.31.255.255  (172.16/12 prefix)
  192.168.0.0     -   192.168.255.255 (192.168/16 prefix)
```

These private ranges are reserved by the Internet Assigned Numbers Authority (IANA) and cannot be routed to the Internet.

Different sized ranges with different allocations of IP addresses can be assigned to a VPC depending on need.

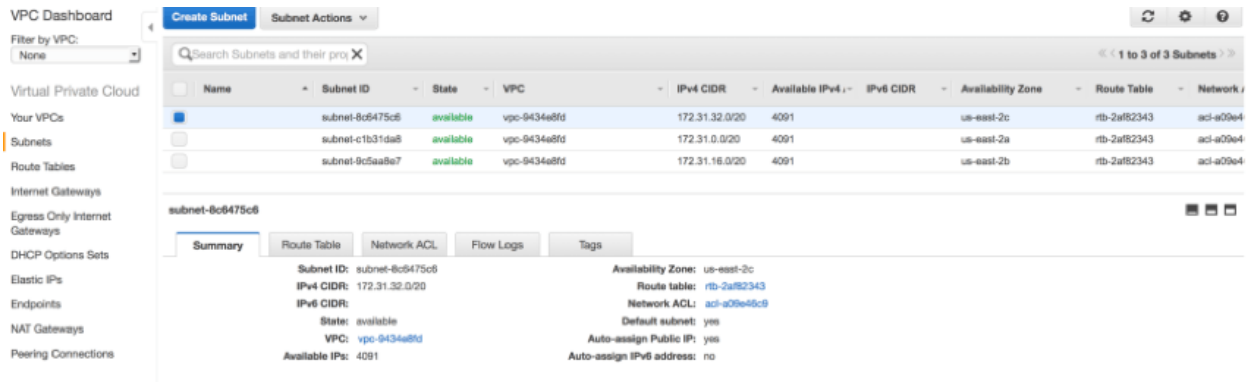
All default VPCs will be associated with an IPv4 CIDR block with a 172.31.0.0/16 address range.

This will give you 65,536 possible IP addresses, minus some AWS reserved addresses.

VPCs can be created with smaller CIDR blocks, such as a /20, which would yield 4091 possible addresses.

Subnet

Next, if you go to the "Subnets" screen, you will see that multiple default subnets have already been assigned to your default VPC, one subnet for each availability zone.



A subnet is always associated with a single availability zone and cannot span multiple zones. However, an availability zone can host multiple subnets.

Each subnet in a VPC is associated with an IPv4 CIDR block that is a subset of the /16 CIDR block of its VPC.

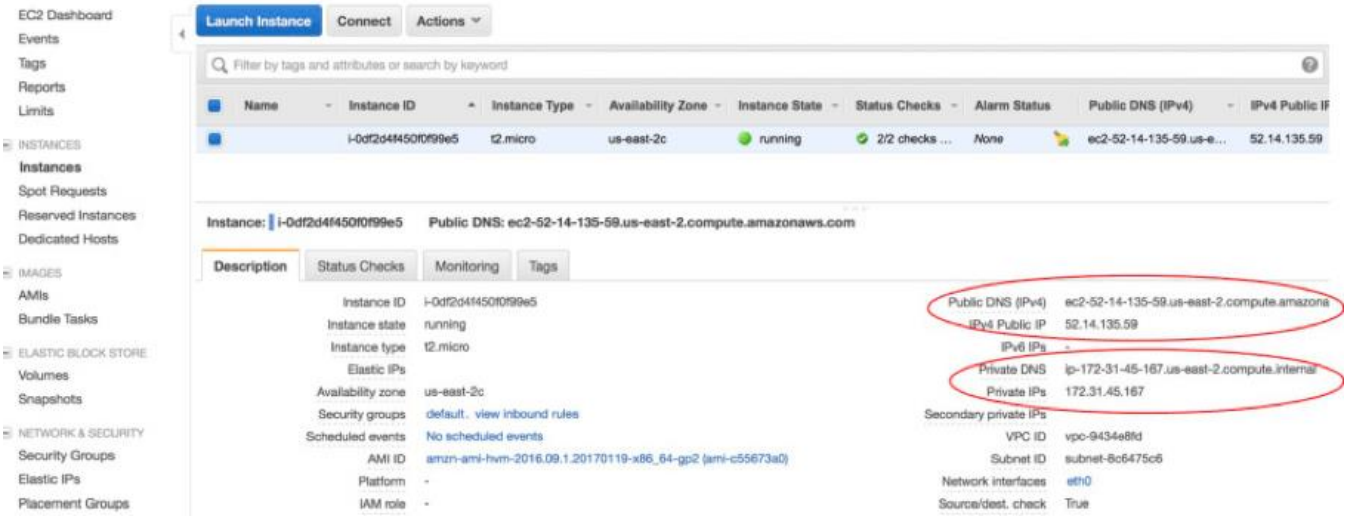
In a default VPC, each default subnet is associated with /20 CIDR block address range which will **have 4091 possible IP addresses minus the five addresses AWS always reserves.**

Note that two subnets cannot have overlapping address ranges.

When you launch an EC2 instance into a default VPC without specifying a specific subnet, it's automatically launched in one of the default subnets.

Every instance in a default subnet receives a private IP address from the pool of addresses associated with that subnet and a private DNS hostname.

In a default subnet, an instance will also receive a public IP address from the pool of addresses owned by AWS along with a public DNS hostname, which will facilitate Internet access for your instances.



Gateways

Frequently, your EC2 instances will require connectivity outside of AWS to the Internet or to a user's corporate network via the use of gateways.

For communication with the Internet, a VPC must be attached to an Internet gateway.

An Internet gateway is a fully managed AWS service that **performs bi-direction source and destination network address translation for your EC2 instances**.

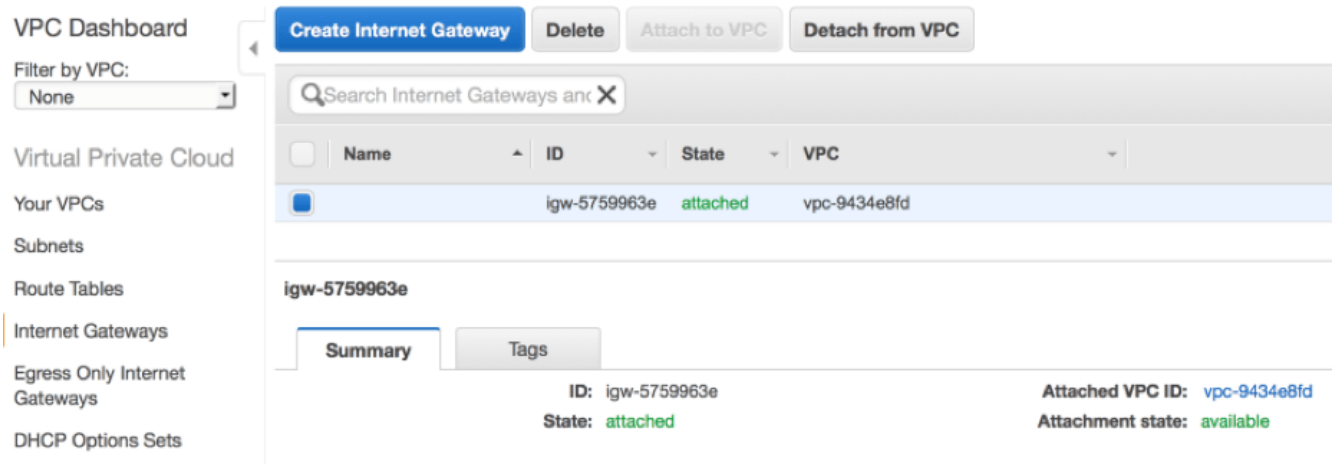
Optionally, a VPC may use **a virtual private gateway** to grant instances secure access to a user's **corporate network via VPN or direct connect links**.

Instances in a subnet can also be granted outbound only Internet access through a NAT gateway.

A subnet that provides its instances a route to an Internet gateway is considered a public subnet.

A private subnet may be in a VPC with an attached Internet gateway but will not have a route to that gateway.

In a default VPC, all default subnets are public subnets and will have a route to a default gateway.



Route table

I have mentioned routing several times while talking about the Internet gateway.

Every VPC is attached to an implicit router.

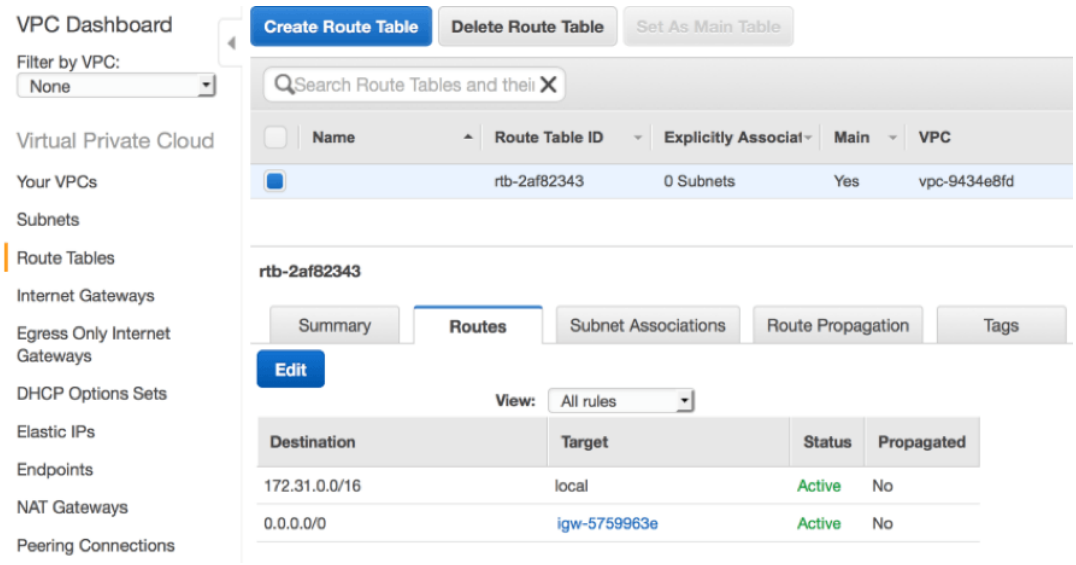
This router is not visible to the user and is fully managed and scaled by AWS.

What is visible is the **route table associated with each subnet**, which is used by the VPC router to **determine the allowed routes for outbound network traffic leaving a subnet**.

Note from the screenshot below that every route table contains **a default local route** to facilitate communication between instances in the same VPC, even across subnets.

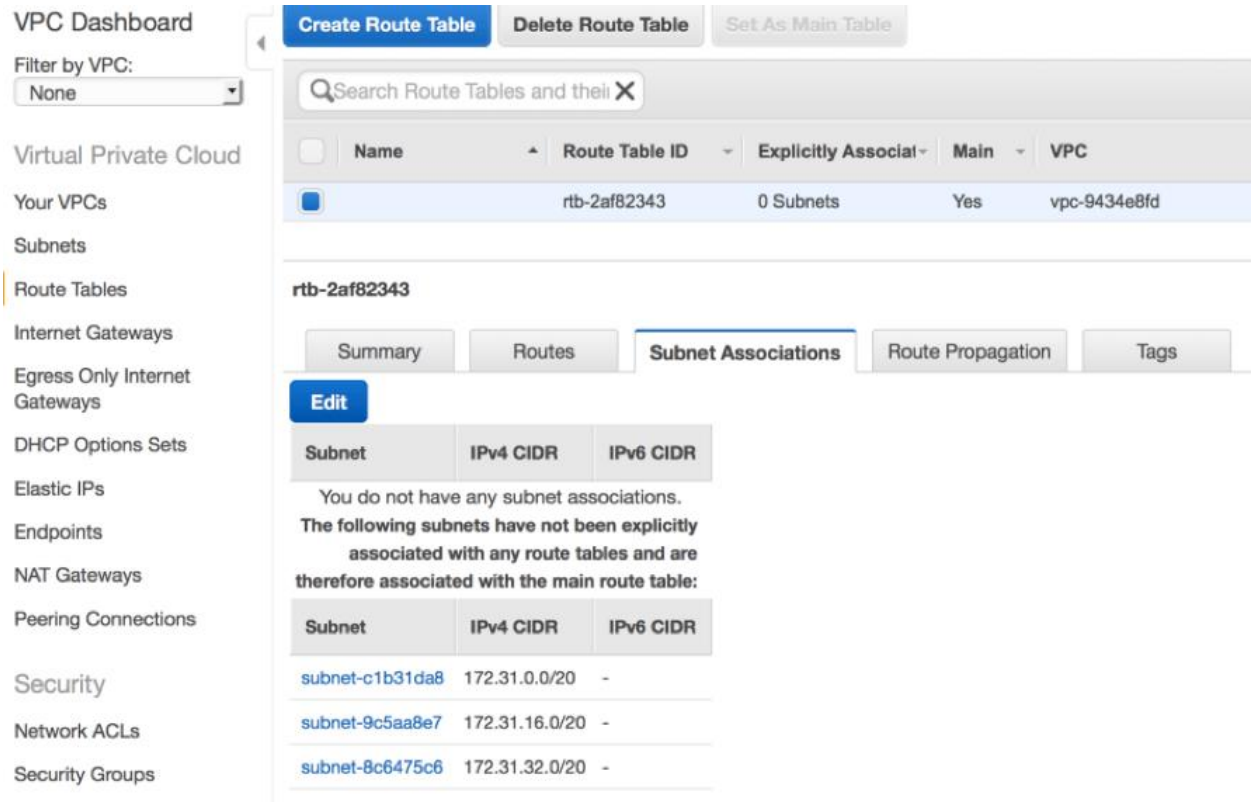
This intra-VPC local route is implied and cannot be changed.

In the case of the main route table that is associated with a default subnet, there will also be a route out to the Internet via the default gateway for the VPC.



Also note that every subnet must be associated with a route table.

If the association is not explicitly defined, then a subnet will be implicitly associated with the main route table.



Network security

One concern you may have is network security, particularly if all default subnets in a default VPC are public and open to Internet traffic.

AWS provides security mechanisms for your instances in the form of network ACLs and security groups.

These two mechanisms can work together to provide layered protection for your EC2 instances.

An ACL acts as a firewall that controls network traffic in and out of a subnet.

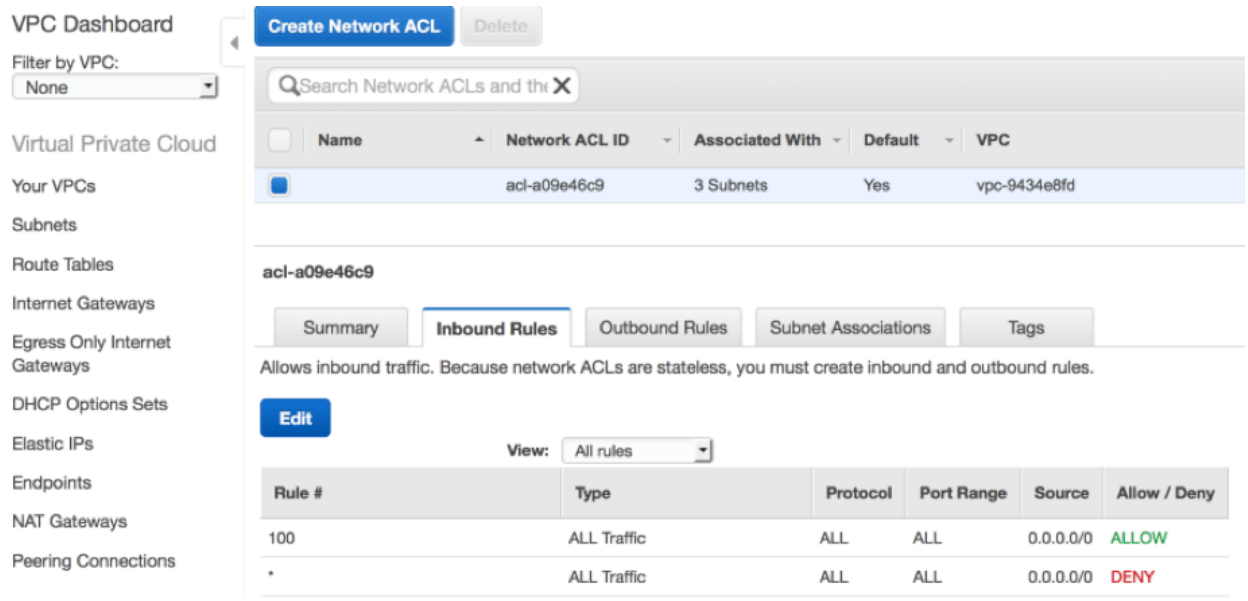
You create rules for allowing or denying network traffic for specific protocols, through specific ports and for specific IP address ranges.

Network ACLs are stateless and have separate inbound and outbound rules.

A network ACL has separate inbound and outbound rules, and each rule can either allow or deny traffic. Network ACLs are stateless, which means that responses to allowed inbound traffic are subject to the rules for outbound traffic (and vice versa).

Stateful

Security groups are stateful. For example, if you send a request from an instance, the response traffic for that request is allowed to reach the instance **regardless of the inbound security group rules. Responses to allowed inbound traffic are allowed to leave the instance, regardless of the outbound rules.**



This means both inbound and outbound rules have to be created to allow certain network traffic to enter the subnet and for responses to go back through.

A number is assigned to each rule and all rules are evaluated starting with the lowest numbered rule.

When traffic hits the firewall, it is evaluated against the rules in ascending order. As soon as a rule is evaluated that matches the traffic being considered, it is applied regardless of what is indicated in a subsequent rule.

As indicated above, the default ACL in a default VPC is configured with lower-numbered rules for both inbound and outbound traffic which combine to explicitly allow bi-directional communication for any protocol, through any port and to and from any source or destination.

You can associate an ACL with multiple subnets, but any single subnet can only be associated with one ACL.

If you do not specifically associate an ACL with a subnet, the subnet is automatically associated with the default ACL.

This is the case with your default VPCs which have all subnets associated with the default ACL.

Security groups

Security groups are considered the first line of defense and consist of a firewall that's applied at the instance level.

This means **only instances explicitly associated with a security group will be subject** to its rules while **all instances in a subnet are impacted by the network ACL applied** to that subnet.

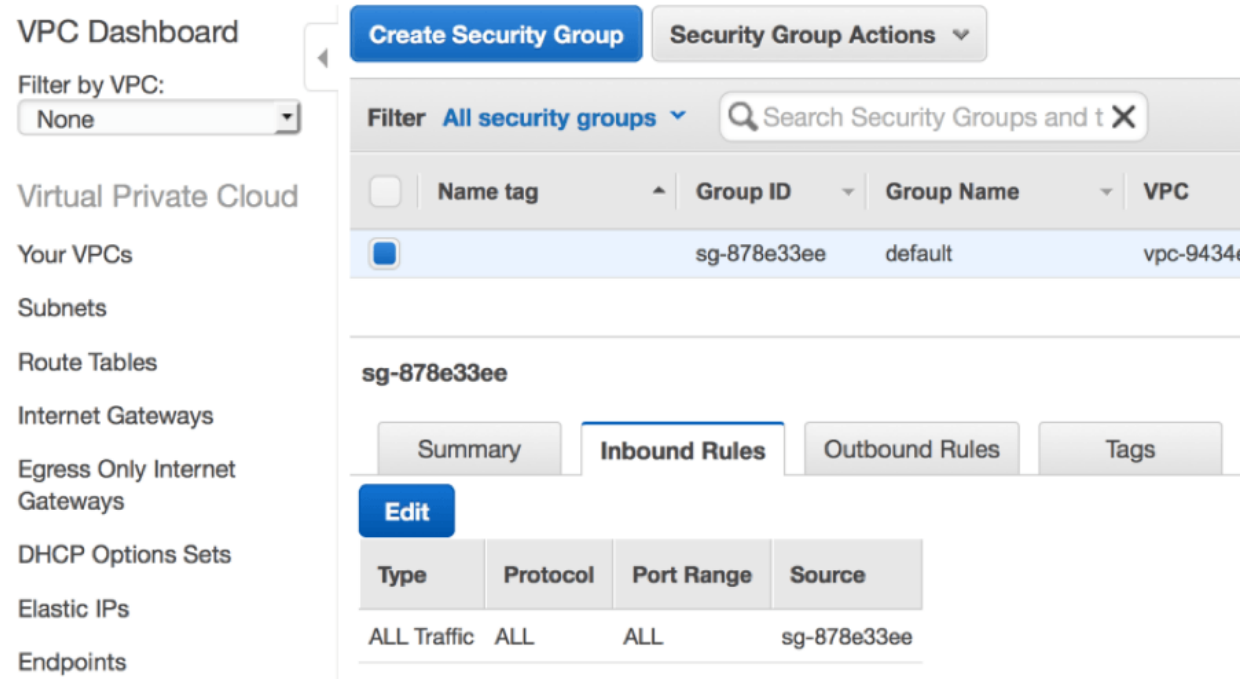
Like ACLs, you create inbound and outbound traffic rules based on protocol, port and source or destination IP. However, there are some differences:

You can specify rules to allow network traffic but cannot create rules to deny specific types of traffic.

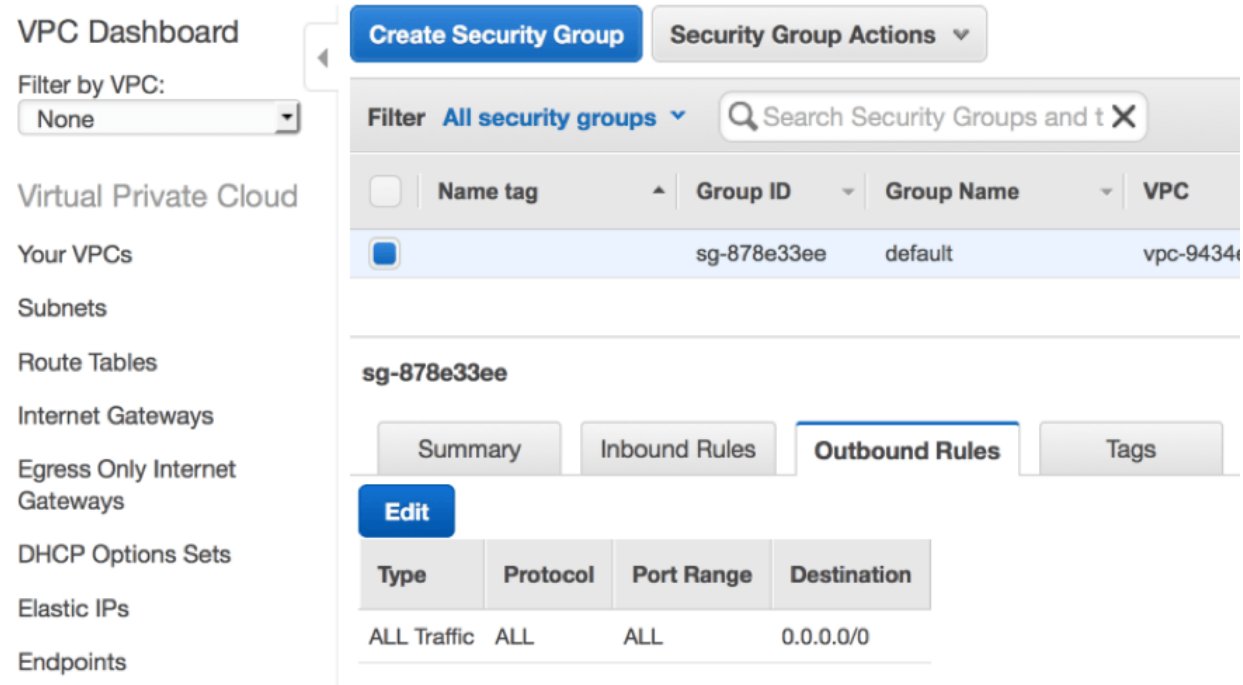
In essence, all traffic is denied except for traffic you explicitly allow.

Security groups **are stateful**, so if you create a rule to allow a certain type of traffic in, then outbound traffic in response is also allowed even if there is no explicit outbound rule to allow such traffic.

Every instance must be associated with a security group **and if a security group is** not specified at launch time, then that instance will be associated **with a default security group**.



You can see from the screenshot above that a default security group will have a rule that only allows inbound traffic from other instances that are associated with the same default security group. No other inbound traffic is allowed.



Looking at the outbound rules above, all network traffic out is allowed by the default security group.

This includes traffic out to the Internet since a default VPC will have a route to a default internet gateway.

While the default security group can be modified, we don’t recommend making any changes to it. Instead, Rackspace recommends creating new security groups that can be attached to resources.

As you can imagine, a default VPC may be suitable for a small non-critical single tier application but is not ideal for a robust production environment.

That's often why users will modify their default VPC configurations once they become more familiar with the technology, and even create custom VPCs for production use.

VPC Pairing

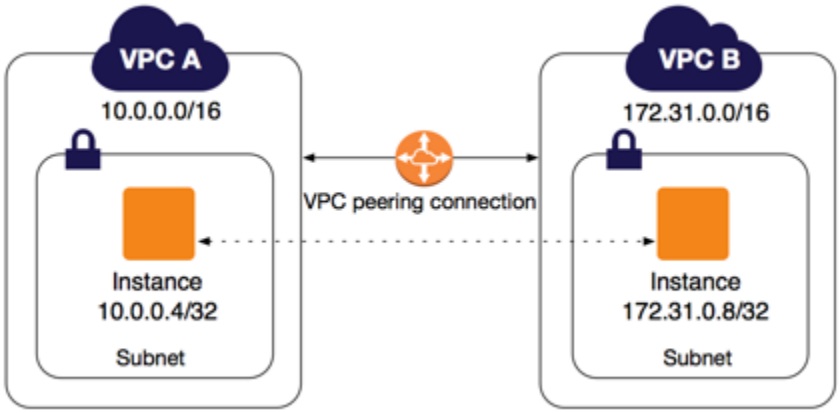
Amazon Virtual Private Cloud (Amazon VPC) enables you to launch AWS resources into a virtual network that you've defined.

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).



AWS uses the existing infrastructure of a VPC to create a VPC peering connection; it is neither a gateway nor a VPN connection and does not rely on a separate piece of physical hardware.

There is no single point of failure for communication or a bandwidth bottleneck.

A VPC peering connection helps you to facilitate the transfer of data.

For example, if you have more than one AWS account, you can peer the VPCs across those accounts to create a file sharing network.

You can also use a VPC peering connection to allow other VPCs to access resources you have in one of your VPCs.