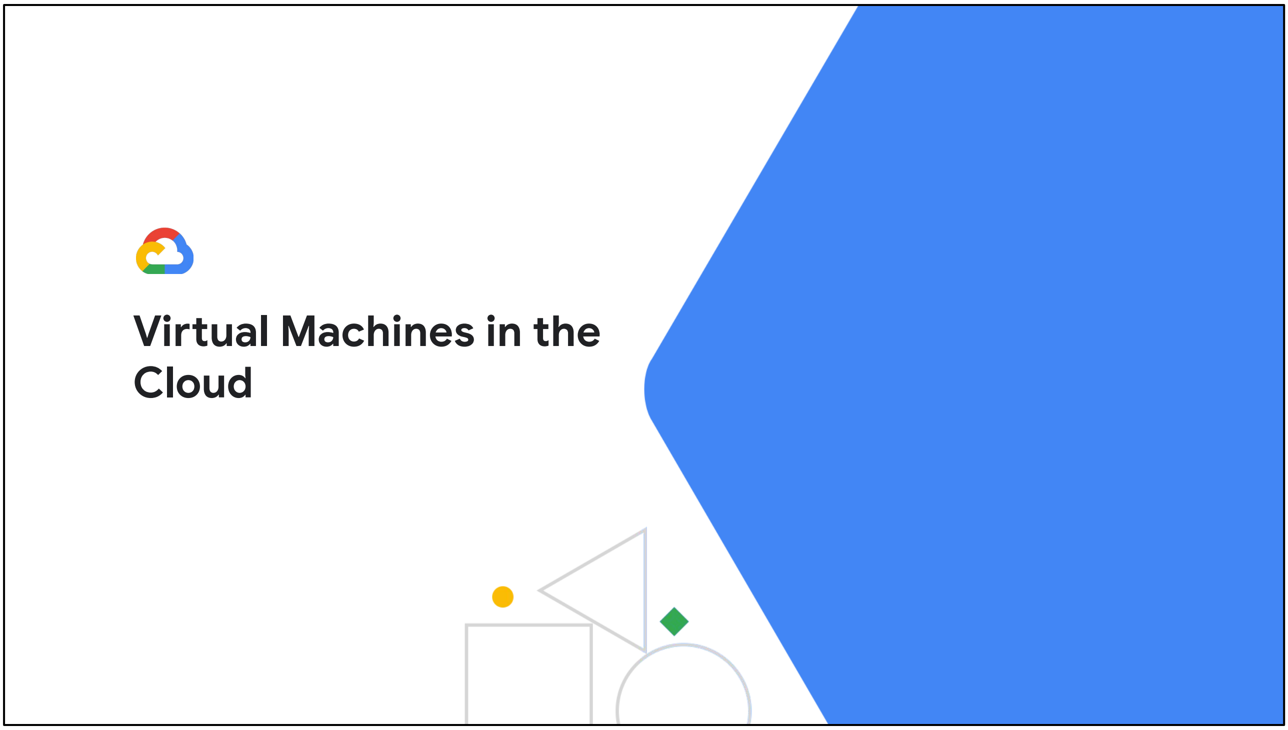
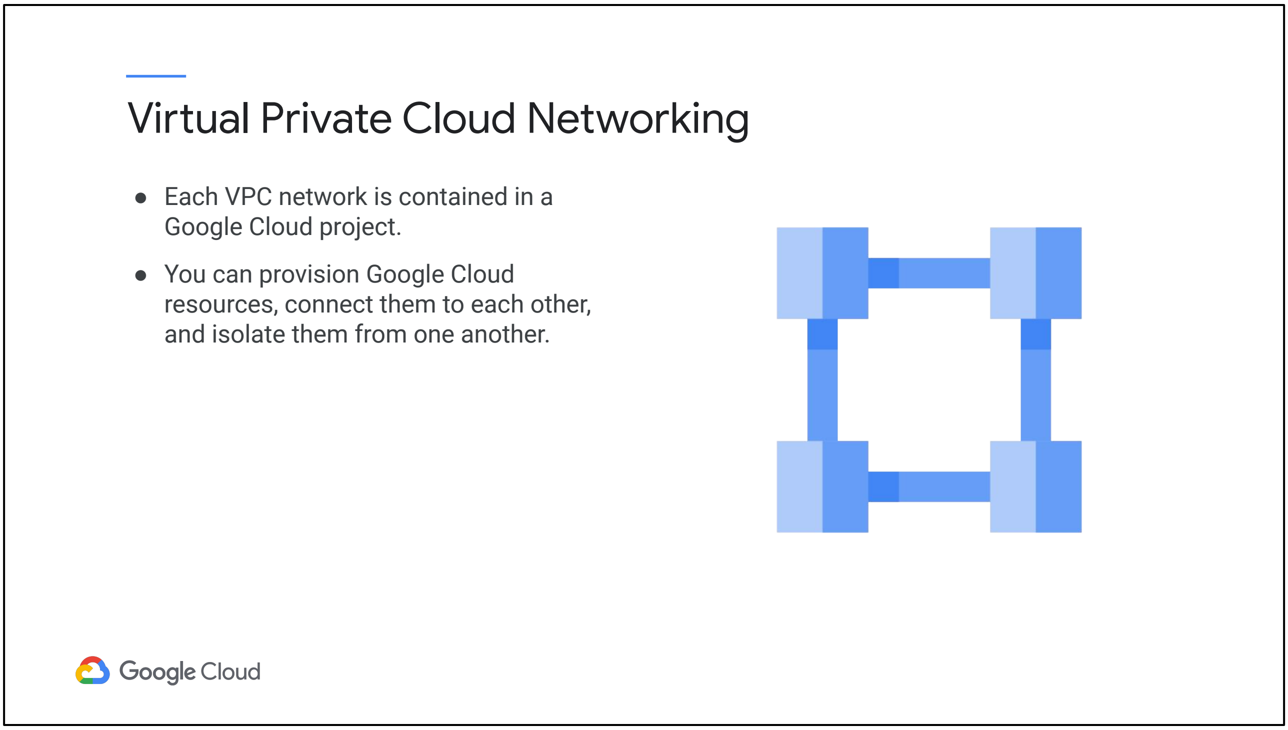


Virtual Machines in the Cloud

Of all the ways you can run workloads in the cloud, virtual machines may be the most familiar. Compute Engine lets you run virtual machines on Google's global infrastructure. In this module, we'll learn how Compute Engine works, with a focus on Google virtual networking.

One of the nice things about virtual machines is that they have the power and generality of a full-fledged operating system in each. You configure a virtual machine much like you build out a physical server: by specifying its amounts of CPU power and memory, its amounts and types of storage, and its operating system. You can flexibly reconfigure them. And a VM running on Google's cloud has unmatched worldwide network connectivity.

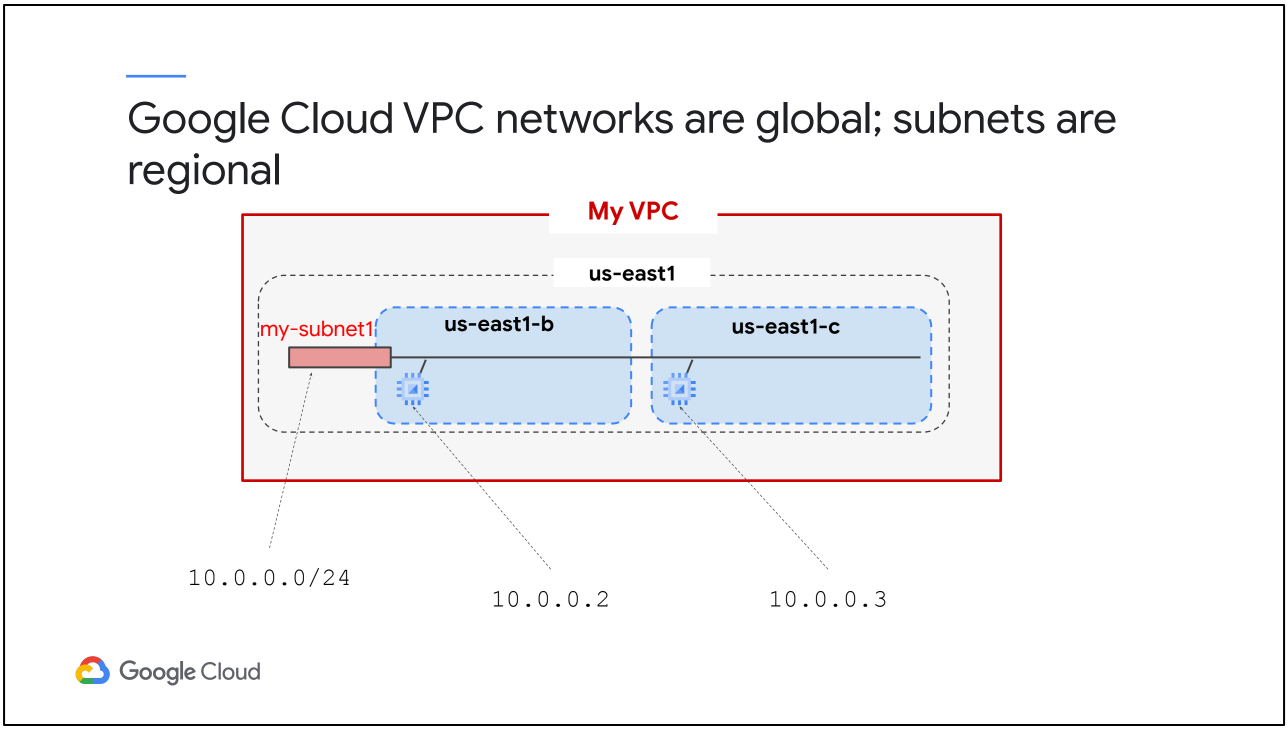


Virtual Private Cloud Networking

- Each VPC network is contained in a Google Cloud project.
- You can provision Google Cloud resources, connect them to each other, and isolate them from one another.

Google Cloud

The way a lot of people get started with Google Cloud is to define their own Virtual Private Cloud inside their first Google Cloud project. Or they can simply choose the default VPC and get started with that. Regardless, your VPC networks connect your Google Cloud resources to each other and to the internet. You can segment your networks, use firewall rules to restrict access to instances, and create static routes to forward traffic to specific destinations.

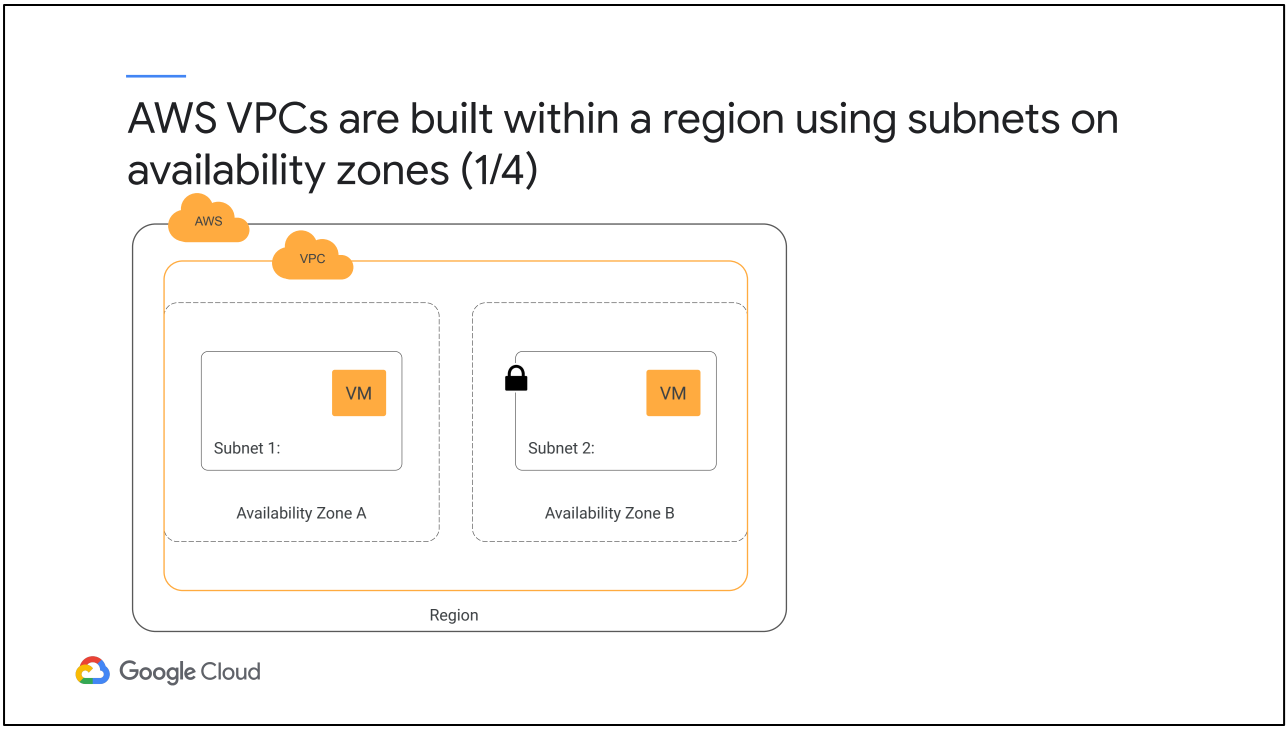


Google Cloud VPC networks are global; subnets are regional

Here’s something that surprises a lot of people who are new to Google Cloud. The Virtual Private Cloud networks that you define have global scope. They can have subnets in any Google Cloud region worldwide. Subnets can span the zones that make up a region. This architecture makes it easy for you to define your own network layout with global scope. You can also have resources in different zones on the same subnet.

You can dynamically increase the size of a subnet in a custom network by expanding the range of IP addresses allocated to it. Doing that doesn’t affect already configured VMs.

In this example, your VPC has one network. So far, it has one subnet defined, in Google Cloud’s us-east1 region. Notice that it has two Compute Engine VMs attached to it. They’re neighbors on the same subnet even though they are in different zones! You can use this capability to build solutions that are resilient but still have simple network layouts.



AWS VPCs are built within a region using subnets on availability zones (1/4)

AWS

VPC

VM

Subnet 1:

Availability Zone A

VM

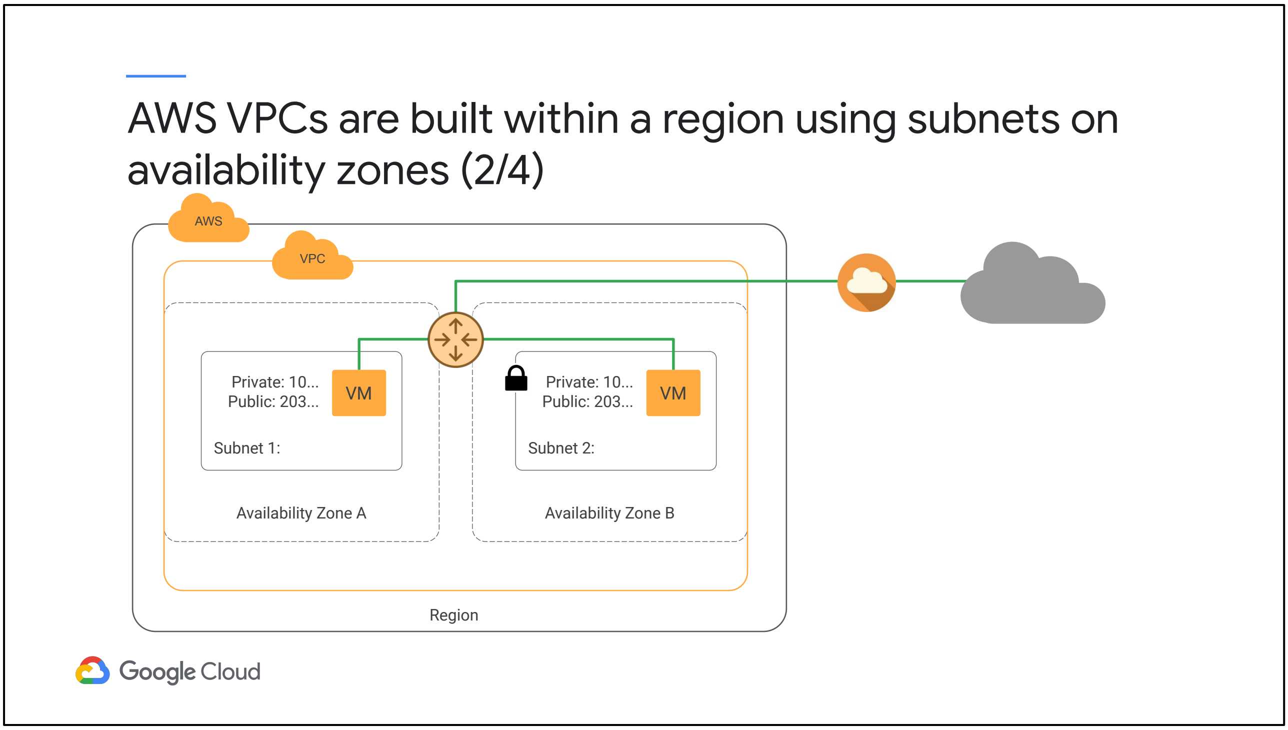
Subnet 2:

Availability Zone B

Region

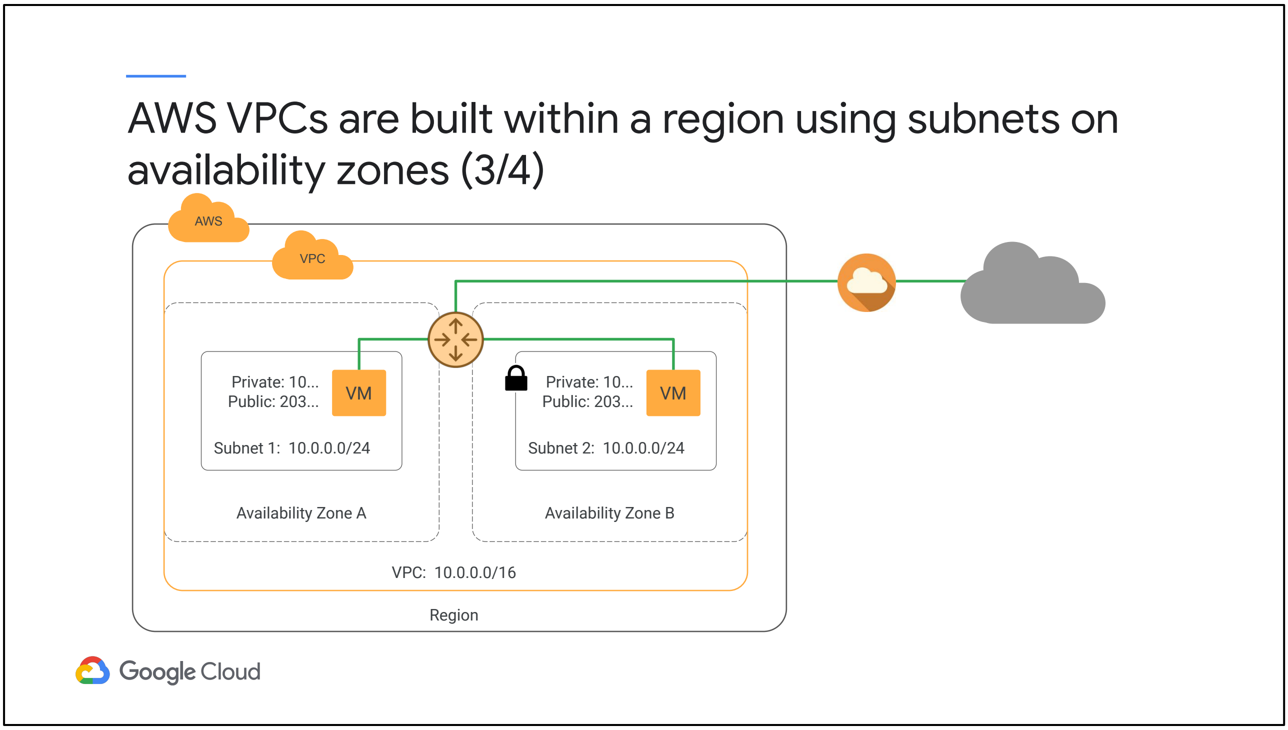
Google Cloud

The VPC is made up of subnets, which must be built on availability zones in the region.



AWS VPCs are built within a region using subnets on availability zones (2/4)

A subnet can be either public or private. A public subnet can route traffic to the internet. A private subnet's traffic never leaves the VPC.



AWS VPCs are built within a region using subnets on availability zones (3/4)

AWS

VPC

Private: 10...
Public: 203...

VM

Subnet 1: 10.0.0.0/24

Availability Zone A

Private: 10...
Public: 203...

VM

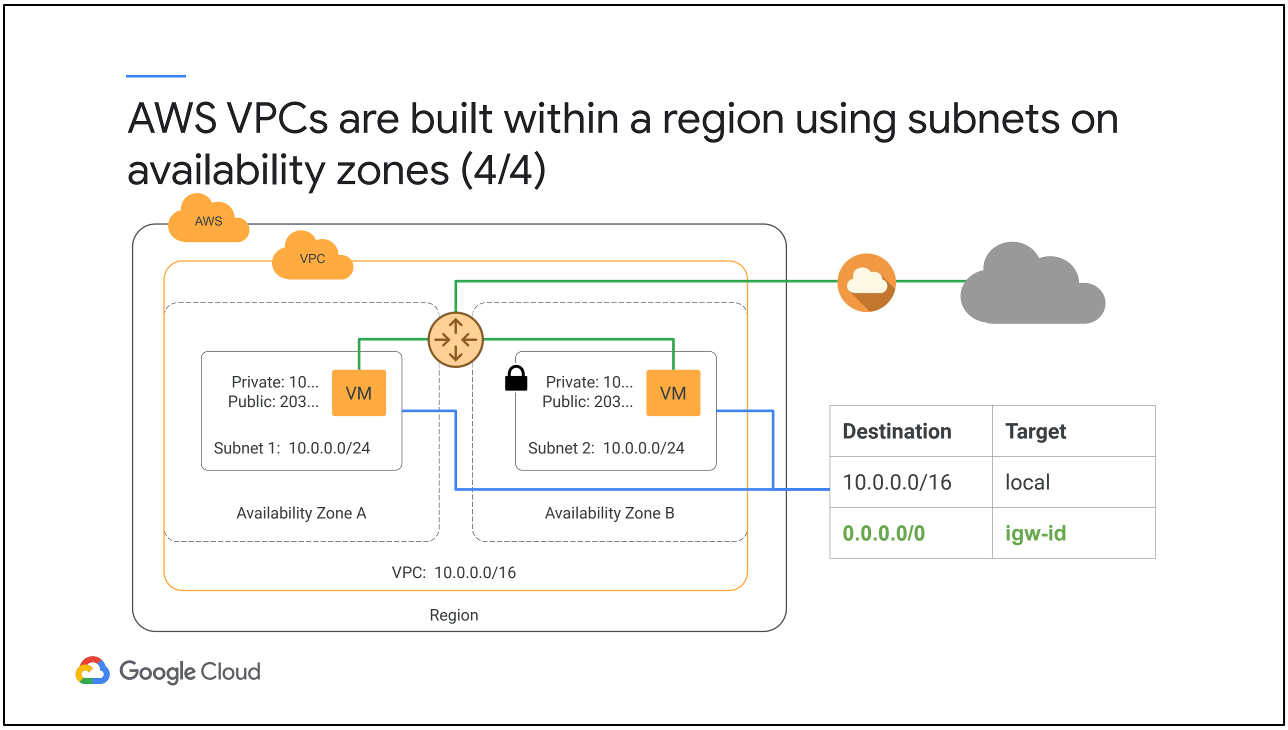
Subnet 2: 10.0.0.0/24

Availability Zone B

VPC: 10.0.0.0/16

Region

A VPC must be built with a CIDR, or Classless Inter-Domain Routing, range of private IP addresses that conform to RFC 1918. All subnets in a VPC must have private IP ranges that are part of the VPC CIDR range.



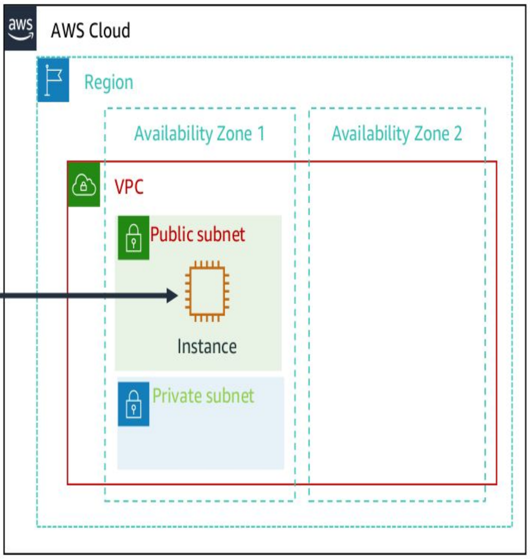
AWS VPCs are built within a region using subnets on availability zones (4/4)

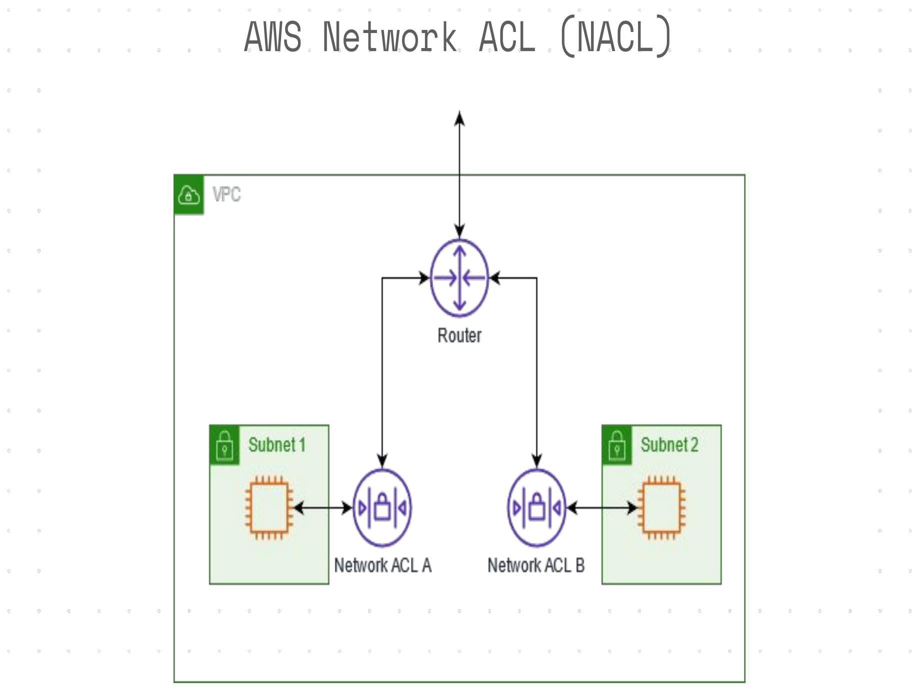
Destination	Target
10.0.0.0/16	local
0.0.0.0/0	**igw-id**

Route tables are built for each VPC that configure paths for traffic. Traffic cannot flow outside the network without Security groups, which are firewalls that can be applied to the virtual machine or network. Network Access Control Lists can be configured to allow and deny traffic to a subnet, but they are not stateful.

AWS Network Settings

Example: specify to deploy the instance here





https://docs.aws.amazon.com/vpc/latest/userguide/vpc-network-acls.html

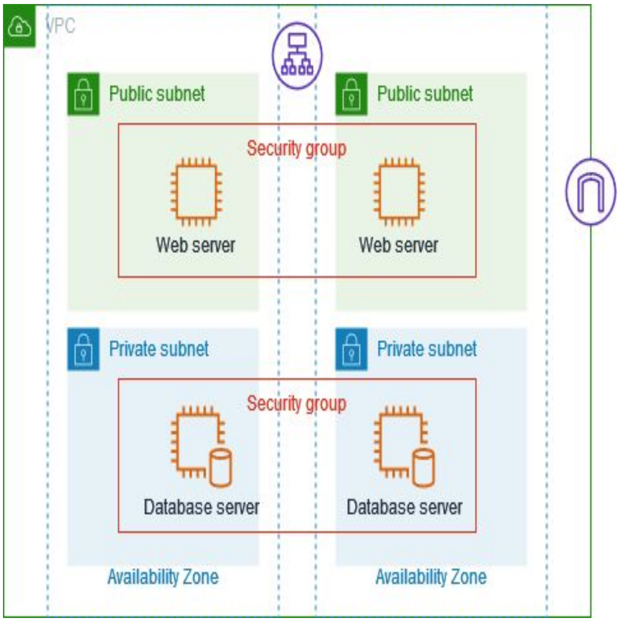
Network Access Control List (NACL)

- It allows or denies specific inbound or outbound traffic at the subnet level.
- You can use the default Network ACL for your VPC, or you can create a custom network ACL for your VPC with rules that are similar to the rules for your **Security Groups** in order to add an additional layer of security to your VPC.
- It provides an additional layer of defense if **Security Group** rules are too permissive.
- There is no additional charge for using network ACLs.

The above diagram shows a VPC with two subnets.

- Each subnet has a network ACL.
- When traffic enters the VPC (as from a peered VPC, VPN connection, or the internet), the router sends the traffic to its destination.
- **Network ACL A** determines which traffic destined for subnet 1 is allowed to enter subnet 1, and which traffic destined for a location outside subnet 1 is allowed to leave subnet 1.
- **Network ACL B** determines which traffic is allowed to enter and leave subnet 2.

AWS Security Group (SG)



https://docs.aws.amazon.com/vpc/latest/userguide/vpc-security-groups.html

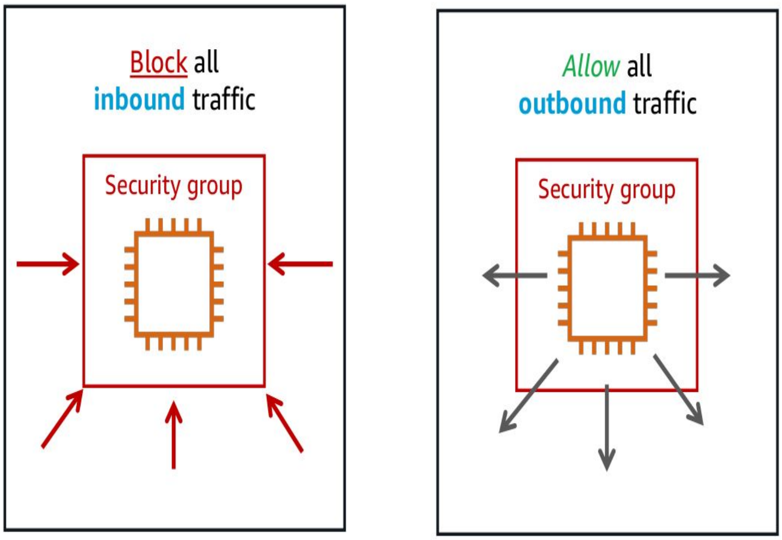
Security Group

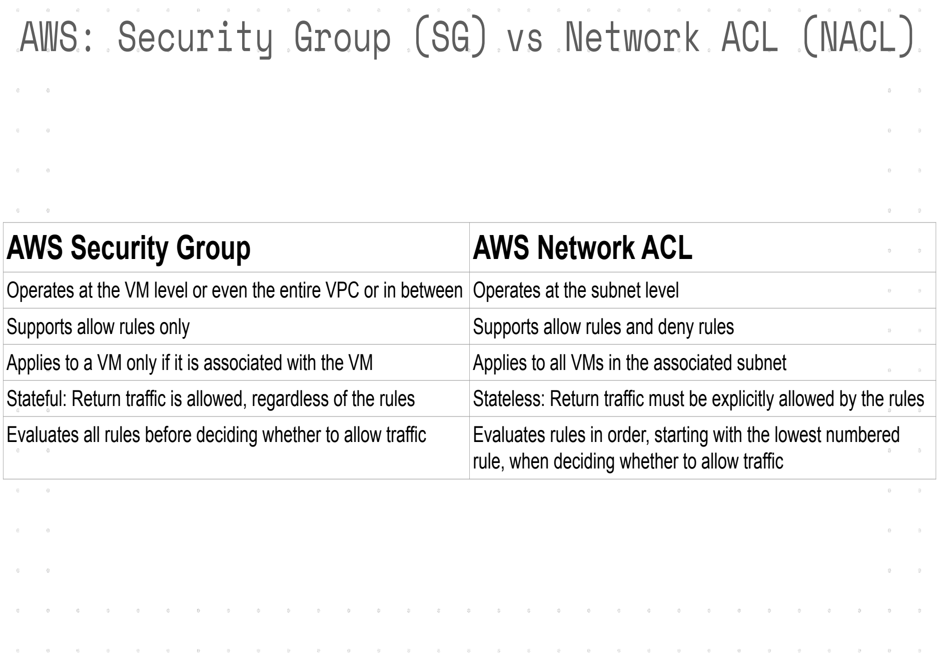
- It controls the traffic that is allowed to reach /leave the resources that it is associated with
- For example, after you associate a **Security Group** with an EC2 instance, it controls the inbound and outbound traffic for the instance.
- You can associate a **Security Group** only with resources in the VPC for which it is created.
- When you create a VPC, it comes with a default security group.
- You can create additional **Security Groups** for each VPC.
- There is no additional charge for using **Security Groups**.

The above diagram shows a VPC with two subnets in two AZs

- The diagram also shows an Application Load Balancer.
- Each AZ has a public subnet for web servers and a private subnet for database servers.
- There are separate **Security Groups** for the load balancer, the web servers, and the database servers.
  a. You can add rules to the **Security Group** for the load balancer to allow HTTP and HTTPS traffic from the internet.
  b. You can add rules to the **Security Group** for the web servers to allow traffic only from the load balancer.
  c. You can add rules to the **Security Group** for the database servers to allow only database requests from the web servers.

AWS Security Group





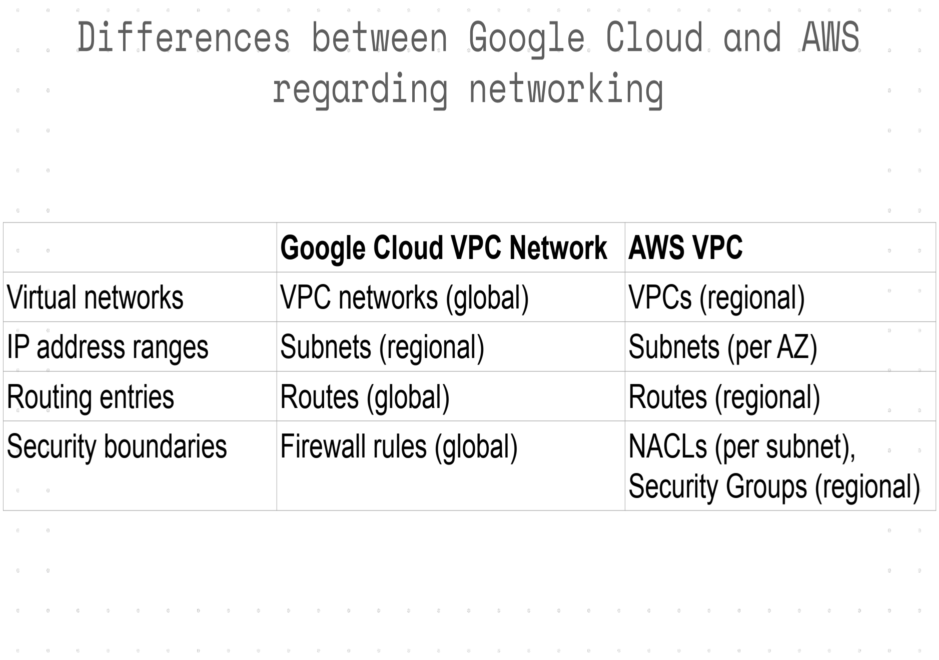
AWS: Security Group (SG) vs Network ACL (NACL)

AWS Security Group	AWS Network ACL
Operates at the VM level or even the entire VPC or in between	Operates at the subnet level
Supports allow rules only	Supports allow rules and deny rules
Applies to a VM only if it is associated with the VM	Applies to all VMs in the associated subnet
Stateful: Return traffic is allowed, regardless of the rules	Stateless: Return traffic must be explicitly allowed by the rules
Evaluates all rules before deciding whether to allow traffic	Evaluates rules in order, starting with the lowest numbered rule, when deciding whether to allow traffic

https://docs.aws.amazon.com/vpc/latest/userguide/infrastructure-security.html

- Use **Security Groups** as the primary mechanism for controlling network access to your VPCs.
- When necessary, use **network ACLs** to provide stateless, coarse-grain network control.
- **Security groups** are more versatile than **network ACLs**, due to their ability to
  - perform stateful packet filtering and
  - create rules that reference other security groups.
- **Network ACLs** can be effective as a
  - secondary control (for example: to deny a specific subset)
  - high-level subnet guard rails.

- Because **network ACLs** apply to an entire subnet, they can be used as defense-in-depth in case an instance is ever launched without the correct **Security Group**.



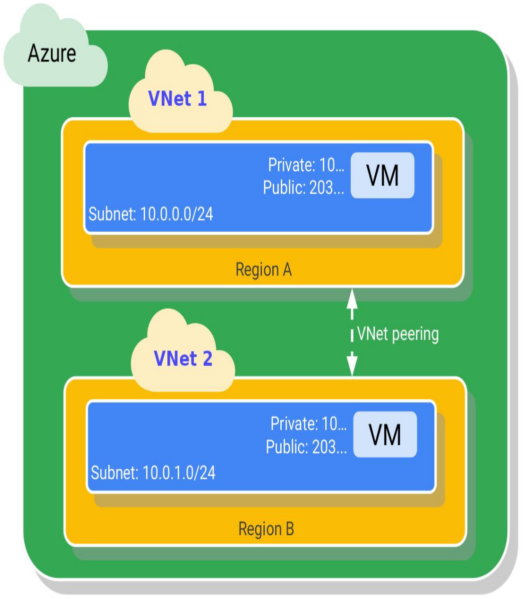
Differences between Google Cloud and AWS regarding networking

	Google Cloud VPC Network	AWS VPC
Virtual networks	VPC networks (global)	VPCs (regional)
IP address ranges	Subnets (regional)	Subnets (per AZ)
Routing entries	Routes (global)	Routes (regional)
Security boundaries	Firewall rules (global)	NACLs (per subnet), Security Groups (regional)

The networking services that Google and Amazon provide can vary in terms of their scope, as indicated here in parentheses.

- The key takeaways are that Google VPC networks are global and subnets span regions, not AZ as in the case of AWS.
- Global networks offer farther, out-of-the-box reach, which means you can create a single private network that is global, without having to connect multiple private networks and manage those spaces separately.
- In Google Cloud, you can also define multiple VPC networks per project for added flexibility.

Azure VNets are regional

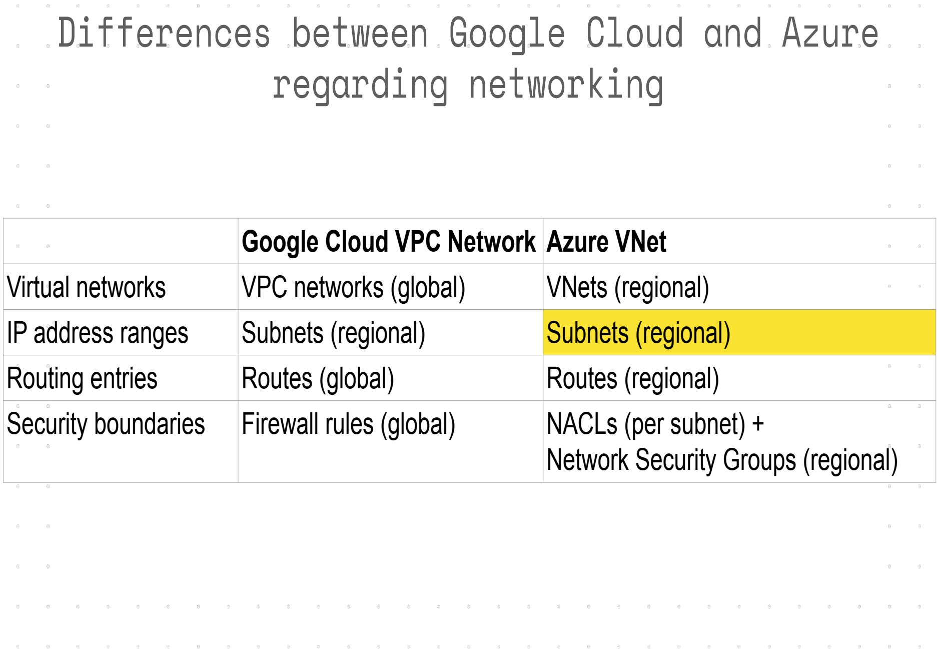


VNets have a slightly different internal structure to VPCs.

- The VNets are not global like VPC networks; instead, each VNet is built in a region and contains 1 or more subnets.
- Multiple VNets from different regions can be connected together using **VNet Peering**.
- When creating a VNet, you must specify a custom private IP address space.
- A VNet must be built with a **CIDR range** of private IP addresses that conform to **RFC 1918**.
- CIDR stands for Classless Inter-Domain Routing.
- All subnets in a VNet must have private IP ranges that are part of the VNet CIDR range.

- You must make sure your VNet address space in this CIDR range does not overlap with your organization's other network ranges.
- Route tables are built for each VNet that configure paths for traffic.

- Traffic cannot flow outside the network without **Network Security Groups**, which are firewalls that can be applied to the VM or network.
- Network Access Control Lists can be configured to allow and deny traffic to a subnet, but they are not stateful.



Differences between Google Cloud and Azure regarding networking

	Google Cloud VPC Network	Azure VNet
Virtual networks	VPC networks (global)	VNets (regional)
IP address ranges	Subnets (regional)	Subnets (regional)
Routing entries	Routes (global)	Routes (regional)
Security boundaries	Firewall rules (global)	NACLs (per subnet) + Network Security Groups (regional)

The networking services that Google Cloud and Azure provide can vary in terms of their scope, as indicated here in parentheses.

- The key takeaways are that Google Cloud VPC networks are global, and subnets span regions; Azure maintains the scope of its networks to a regional level.