# VPCs & Load Balancing

Group 7

# What is a VPC?

A Virtual Private Cloud lets you ‘launch resources into a virtual network you define’.

This network exists in a ‘logically isolated section of the Amazon Web Services cloud’. (Amazon Web Services, 2017)

You have added, and more granular, control over how this network and the instances interact with the internet (Amazon Web Services, 2015, 2017; Cheng 2014).

Image Source:
https://www.newegg.ca/Servers-Workstations/Category/ID-271

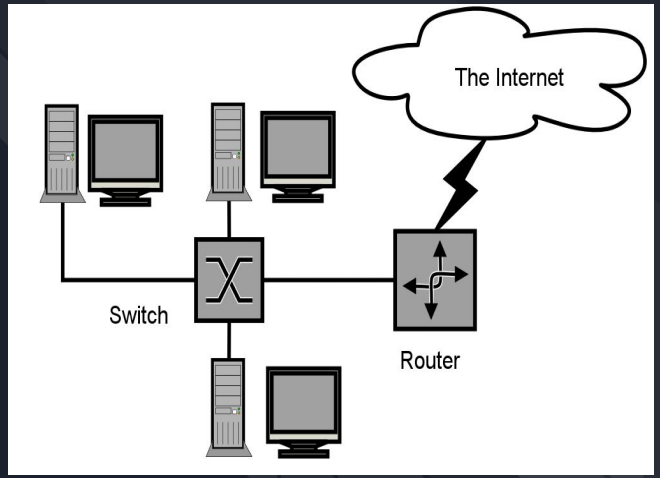
Image Source:
https://pixabay.com/en/workstation-computer-office-desktop-147953/

# VPCs Continued

This virtual network resembles a traditional one (Simplilearn, 2016).

You can have extra control over this working environment, such as:

- Selecting your IP address range
- Creating your own subnets
- Configuring route tables
- Configuring network gateways
- Adding network access control lists (‘Firewalls’)




Adapted From: https://en.wikipedia.org/wiki/Computer_network_diagram

Customize the network configuration based on different needs

(Different use cases benefit from different setups).

(Amazon Web Services, 2015, 2017; CBT Nuggets, 2012)

# AWS VPC Configurations

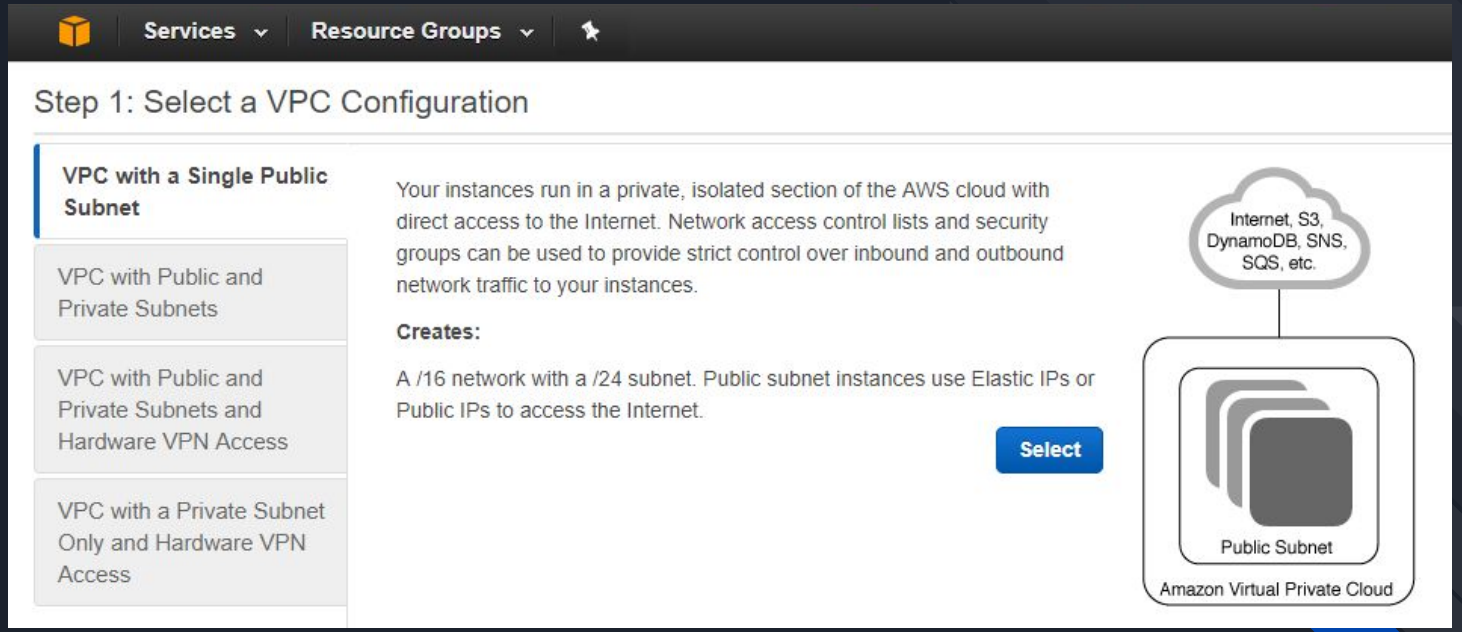



Image Source:
https://ap-southeast-2.console.aws.amazon.com/vpc/home?region=ap-southeast-2#wizardSelector:

# AWS VPC Config: Single Public Subnet

This configuration is recommended for running a single-tier, public-facing web application. Examples include:

- Blogs
- Simple websites

(Amazon Web Services, 2017)

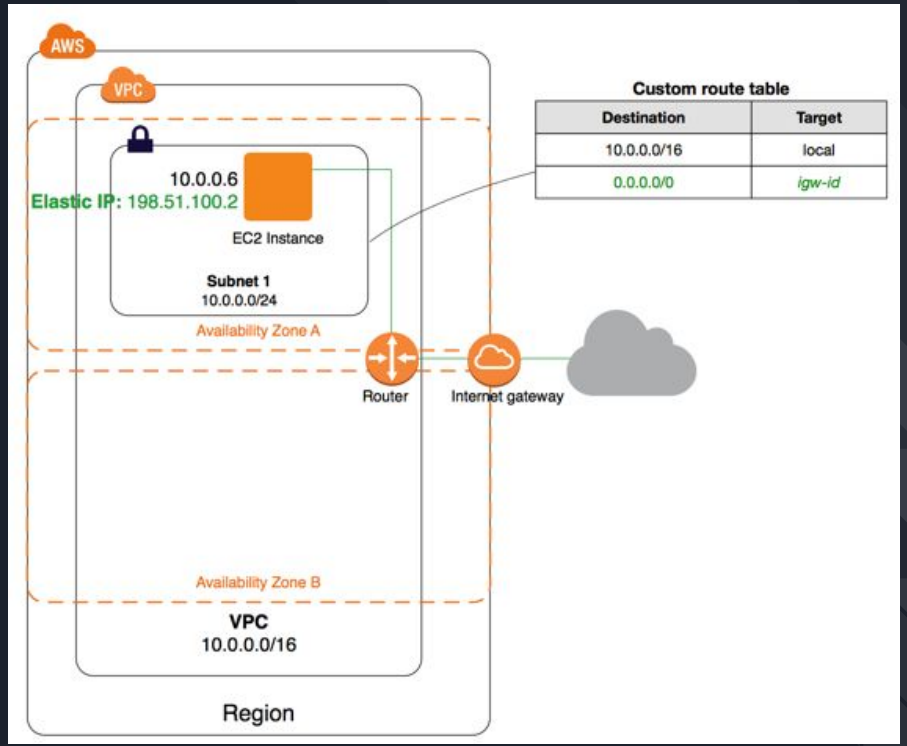



Image Source: http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/VPC_Scenario1.html

# AWS VPC Config: Public & Private Subnets

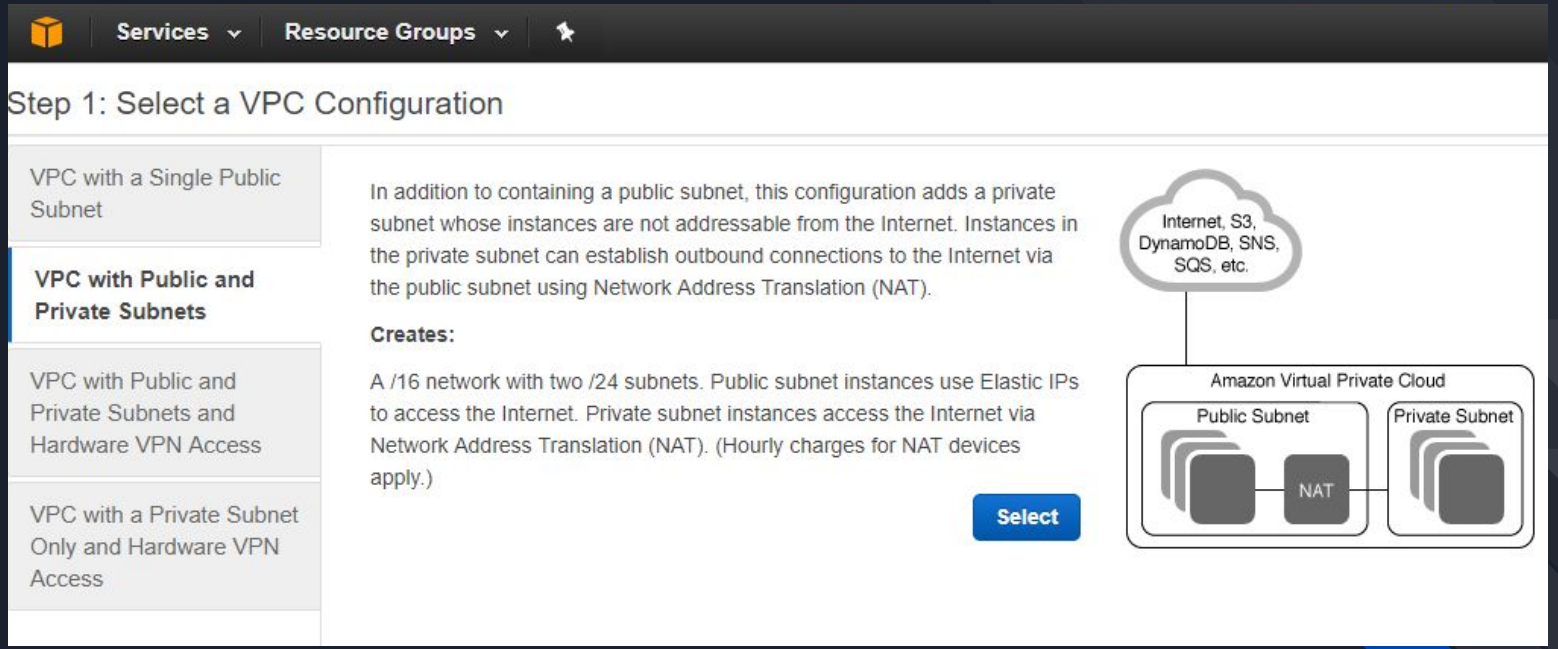



Image Source:
https://ap-southeast-2.console.aws.amazon.com/vpc/home?region=ap-southeast-2#wizardSelector:

# AWS VPC Config: Public & Private Subnets

This configuration is recommended for running a public-facing web application, while maintaining back-end servers that aren't publicly accessible. Examples include:

- A multi-tier website, with the web servers in a public subnet and the database servers in a private subnet.

(Amazon Web Services, 2017)

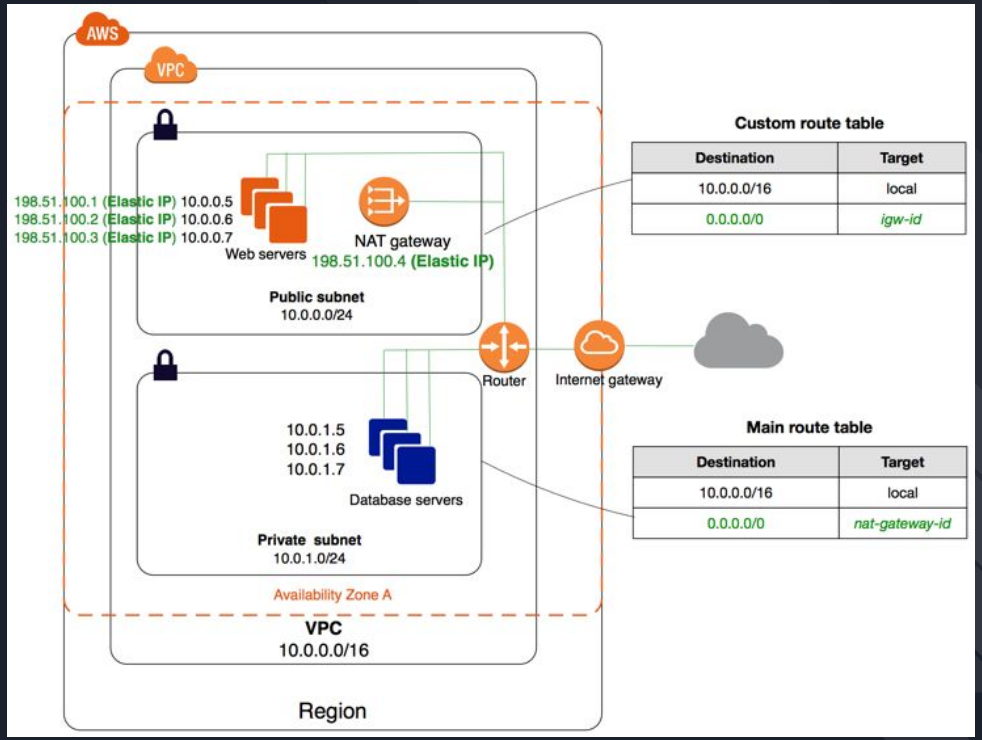



Image Source: http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/VPC_Scenario2.html

# NAT: Network Address Translation

*'Network Address Translation* (NAT) is the process where a network device ... assigns a public address to a computer (or group of computers) inside a private network.' (WhatIsMyIPAddress, 2017)

Public

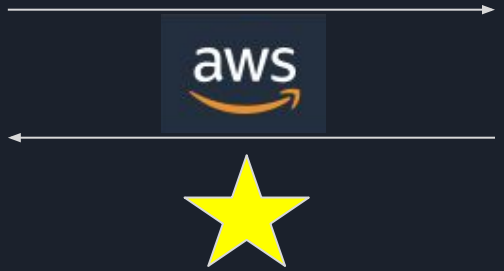



Image Source:
https://www.medicalnewstoday.com/articles/306488.php

# NAT: Network Address Translation

In the case of a VPC with public and private subnets, you can set up an (NAT) instance in the public subnet to act as the necessary device in your VPC to enable instances in the private subnet to initiate outbound IPv4 traffic to the Internet or other AWS services, but prevent the instances from receiving inbound traffic initiated by someone on the Internet (Amazon Web Services 2017).


Image Source:
https://www.medicalnewstoday.com/articles/306488.php

# Configuring ★

To allow instances in your (Public) VPC subnet to access the internet you must attach a gateway, which provides a target for internet traffic as well as potentially acting as an NAT device for a private subnet (Amazon Web Services, 2017). The NAT device can be given further security setting via the use of Network Access Control Lists (Network ACLs).

Each subnet must be associated with a route table, which controls the routing for the subnet and determines where traffic is directed. These tables are customisable, or you can create your own (Amazon Web Services, 2017).

This is on top of your Security Group, providing inbound and outbound filtering at the subnet level as well as the instance level (Amazon Web Services, 2015).

Extra security! If you know what you are doing… (Rando, 2015)

# AWS VPC Config: Private Subnet & Hardware VPN Access

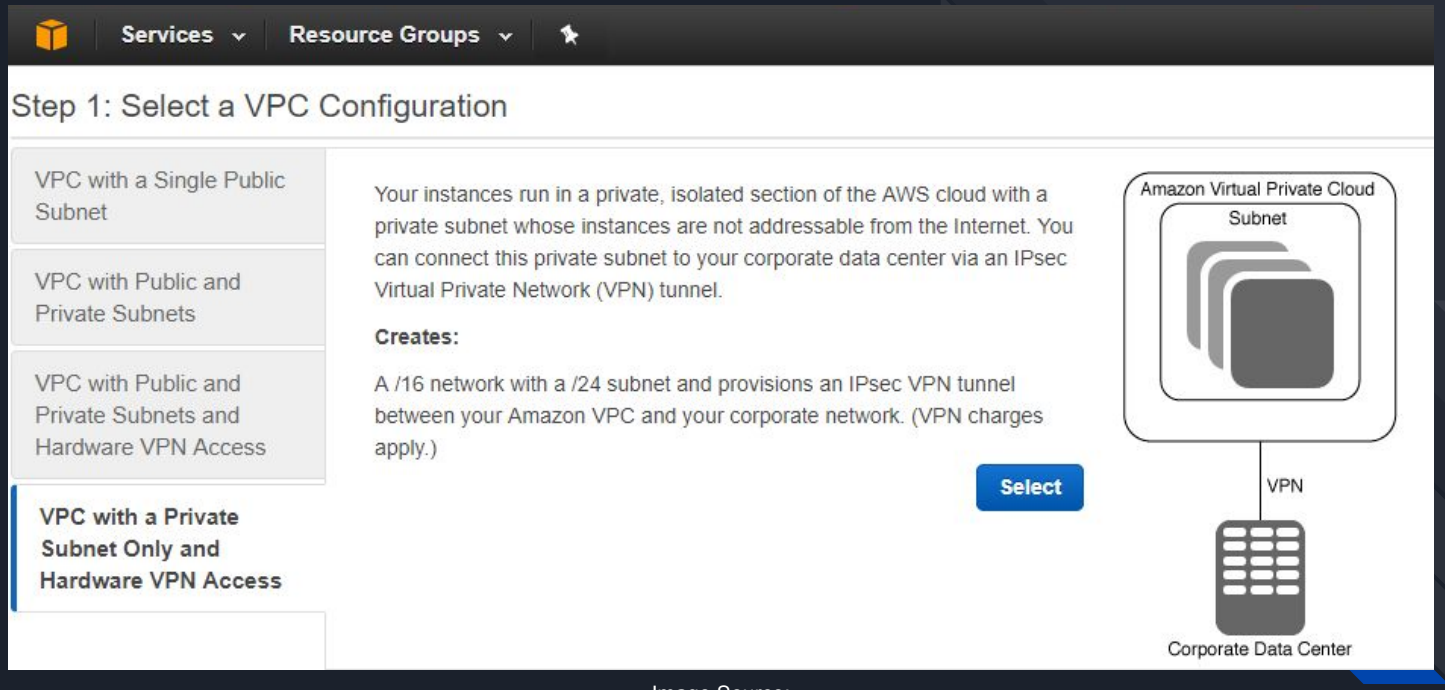



Image Source:
https://ap-southeast-2.console.aws.amazon.com/vpc/home?region=ap-southeast-2#wizardSelector:

# AWS VPC Config: Private Subnet & Hardware VPN Access

This configuration is recommended for extend your network into the cloud using Amazon's infrastructure without exposing your network to the Internet. Examples Include:

- Seamlessly extending corporate data center resources beyond physical capability.

(Amazon Web Services, 2017; Earls, 2017)

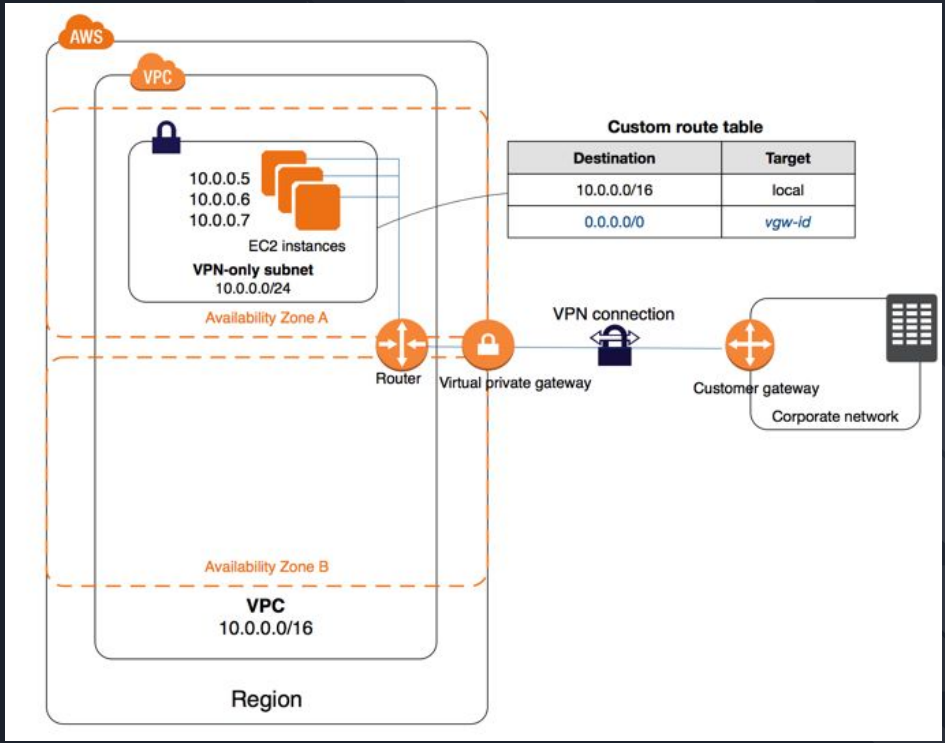



Image Source: http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/VPC_Scenario4.html

# AWS VPC Config: Private & Public Subnets & Hardware VPN Access

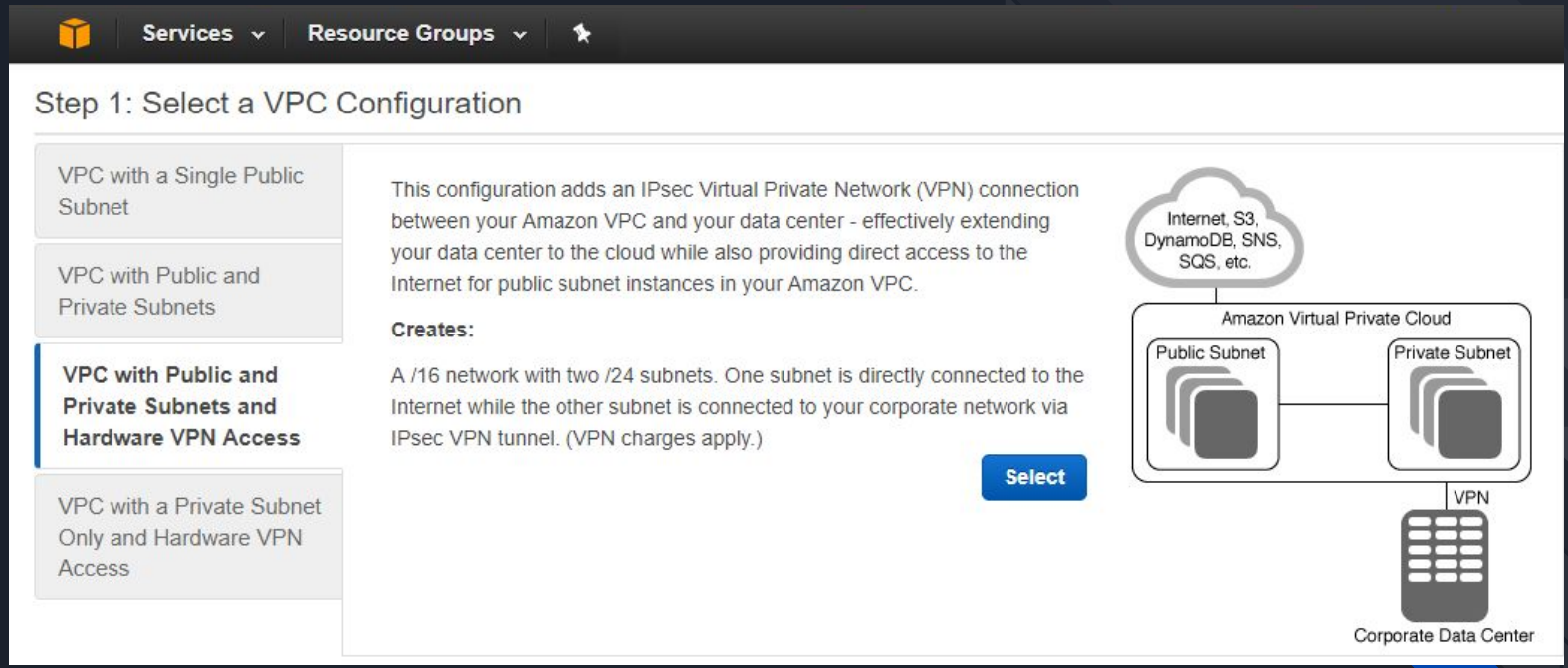



Image Source:
https://ap-southeast-2.console.aws.amazon.com/vpc/home?region=ap-southeast-2#wizardSelector:

# AWS VPC Config: Private & Public Subnets & Hardware VPN Access

This configuration is recommended for extending networks into the cloud while also directly accessing the Internet from your VPC. Examples include:

- Multi-tiered applications with web servers on the front end in the public subnet, and application servers in a private subnet connected to databases in a corporate data center.

(Amazon Web Services, 2017)

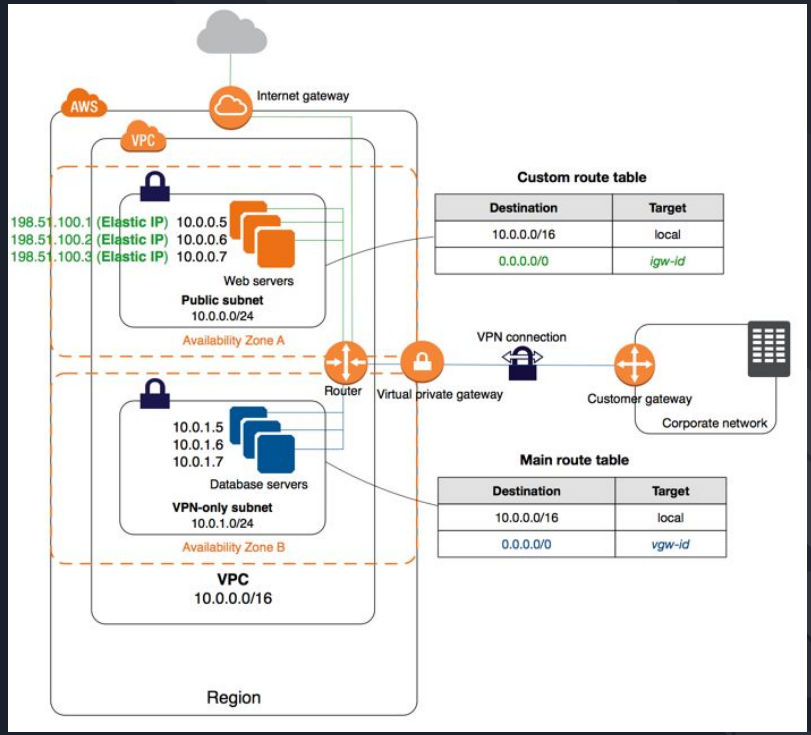



Image Source: http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/VPC_Scenario3.html

# Benefits & Drawbacks of VPC

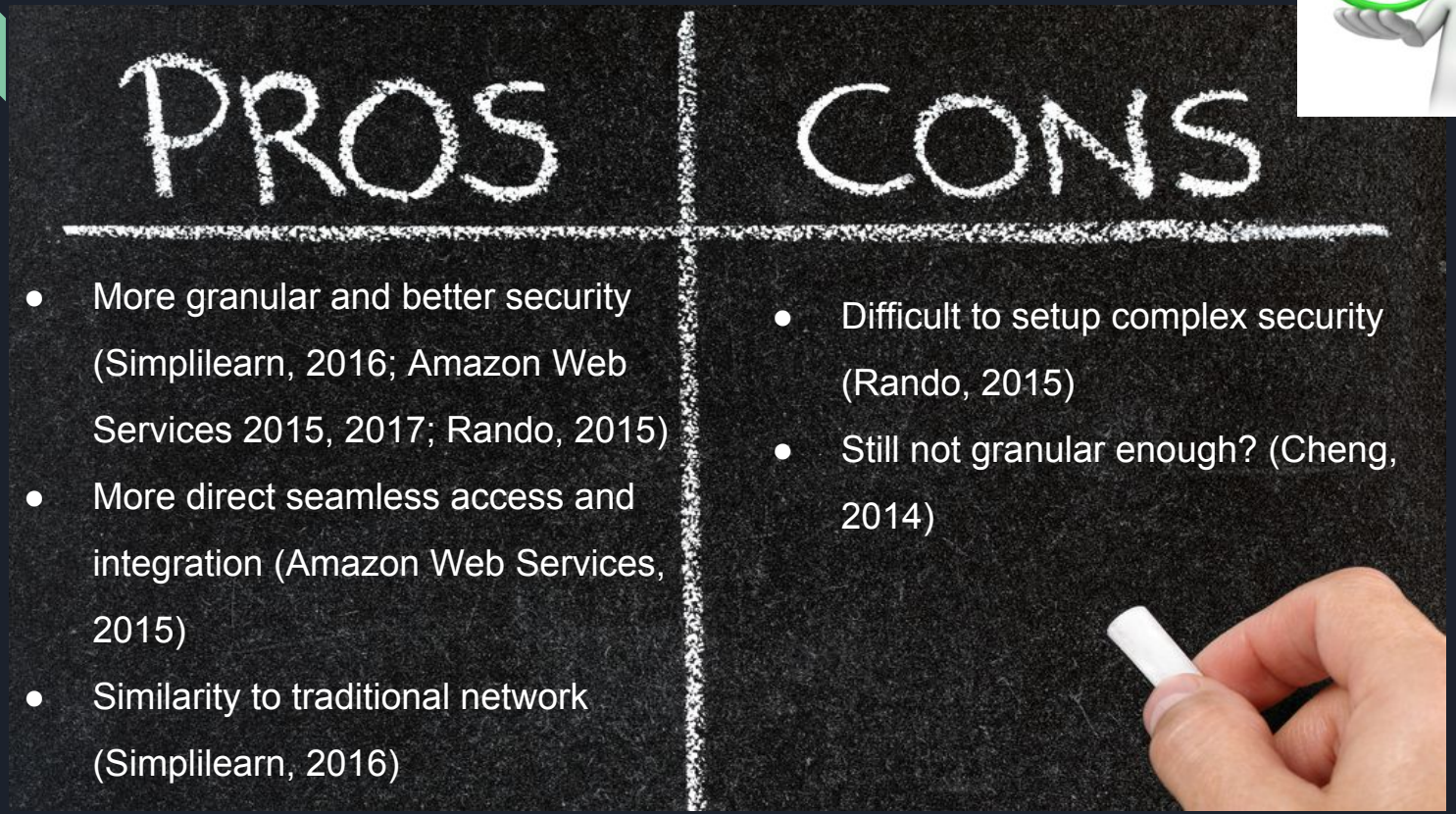


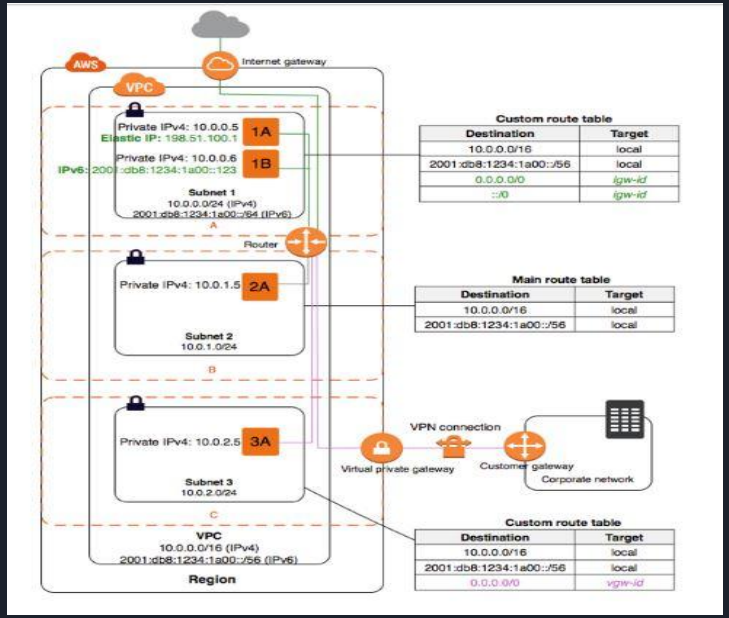

Image Sources: https://www.planetminecraft.com/blog/pros-and-cons-of-starting-your-own-server/
And https://www.solsyst.com/single-post/2014/09/11/Microsoft-Dynamics-NAV-Vs-NetSuite

# VPCs and Subnets

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

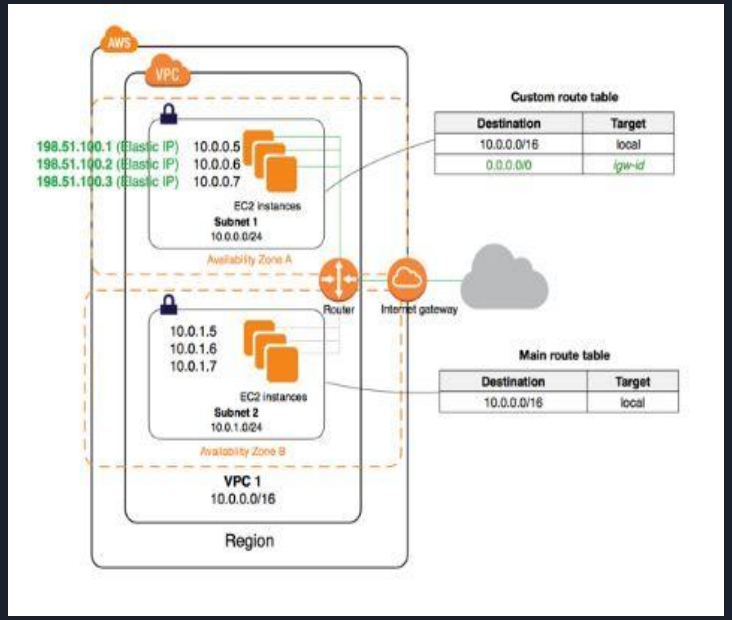
# VPCs and Subnets

- 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.



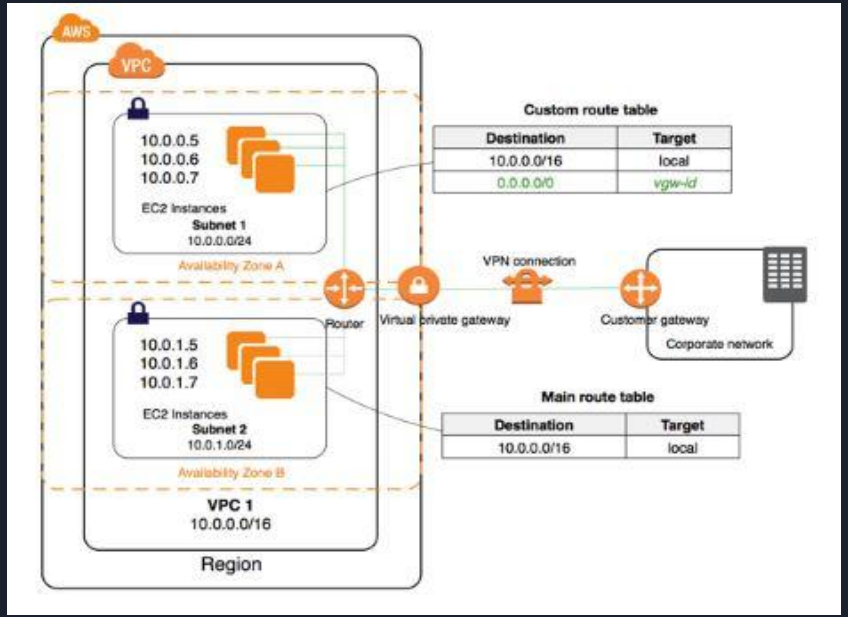
# Accessing the Internet

- By default, each instance that you launch into a nondefault subnet has a private IPv4 address, but no public IPv4 address. These instances can communicate with each other, but can't access the Internet.
- You can enable Internet access for an instance launched into a nondefault subnet by attaching an Internet gateway to its VPC and associating an Elastic IP address with the instance.
- Alternatively, you can allow an instance to connect to the Internet by using a NAT device.



# Accessing a Corporate or Home Network

- You can optionally connect your VPC to your own corporate data center using an IPsec hardware VPN connection, making the AWS cloud an extension of your data center.
- A VPN connection consists of a virtual private gateway attached to your VPC and a customer gateway located in your data center.
- A virtual private gateway is the VPN concentrator on the Amazon side of the VPN connection.
- A customer gateway is a physical device or software appliance on your side of the VPN connection.



# IP Addressing in VPCs

- IP addresses enable resources in your VPC to communicate with each other, and with resources over the Internet. Amazon EC2 and Amazon VPC support the IPv4 and IPv6 addressing protocols.
- When you create a VPC, you must assign it an IPv4 CIDR block. Private IPv4 addresses are not reachable over the Internet. To connect to your instance over the Internet, you need to assign a globally-unique public IPv4 address to your instance.
- You can optionally associate an IPv6 CIDR block with your VPC and subnets, and assign IPv6 addresses from that block to the resources in your VPC. IPv6 addresses are public and reachable over the Internet.
- Your VPC can operate in dual-stack mode: your resources can communicate over IPv4, or IPv6, or both. IPv4 and IPv6 addresses are independent of each other; you must configure routing and security in your VPC separately for IPv4 and IPv6.

# IPv4 vs IPv6

| ***IPv4*** | ***IPv6*** |
| --- | --- |
| The format is 32-bit, 4 groups of 4 numerical digits. | The format is 128-bit, 8 groups of 4 hexadecimal digits. |
| Default and required for all VPCs; cannot be removed. | Opt-in only. |
| The VPC CIDR block size can be from /16 to /28. | The VPC CIDR block size is fixed at /56. |
| The subnet CIDR block size can be from /16 to /28. | The subnet CIDR block size is fixed at /64. |
| Supported for VPC VPN connections and customer gateways, NAT devices, and VPC endpoints. | Not supported for VPC VPN connections and customer gateways, NAT devices, and VPC endpoints. |

# CIDR Blocks

| ***CIDR Block Prefix*** | ***# of Host Addresses*** |
|---|---|
| /27 | 32 hosts |
| /26 | 64 hosts |
| /25 | 128 hosts |
| /24 | 256 hosts |
| /23 | 512 hosts |
| /22 | 1,024 hosts |
| /21 | 2,048 hosts |
| /20 | 4,096 hosts |

# Security in VPC

Amazon VPC provides features that you can use to increase and monitor the security for your VPC:

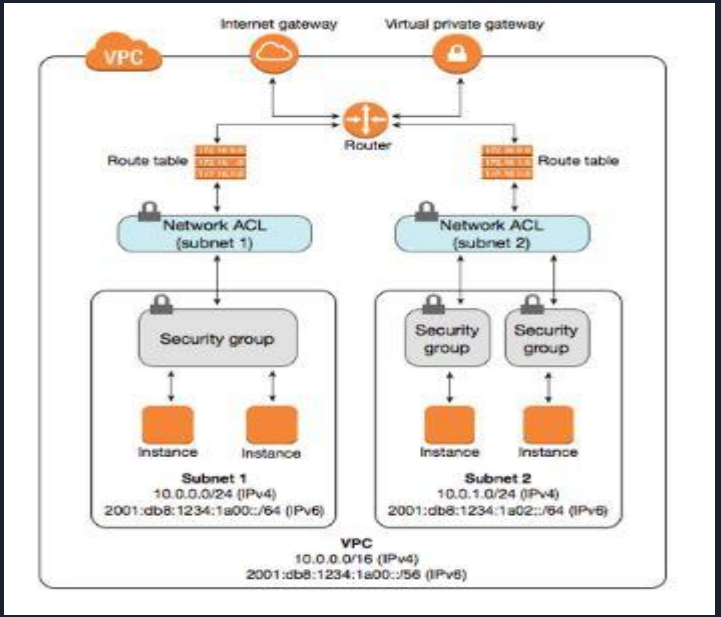
- Security groups — Act as a firewall for associated Amazon EC2 instances, controlling both inbound and outbound traffic at the instance level
- Network access control lists (ACLs) — Act as a firewall for associated subnets, controlling both inbound and outbound traffic at the subnet level
- Flow logs — Capture information about the IP traffic going to and from network interfaces in your VPC

# Security Groups vs Network ACLs

| ***Security Groups*** | ***Network ACLs*** |
| --- | --- |
| Operates at the instance level (first layer of defense) | Operates at the subnet level (second layer of defense) |
| Supports allow rules only | Supports allow rules and deny rules |
| Is stateful: Return traffic is automatically allowed, regardless of any rules | Is stateless: Return traffic must be explicitly allowed by rules |
| We evaluate all rules before deciding whether to allow traffic | We process rules in number order when deciding whether to allow traffic |
| Applies to an instance only if someone specifies the security group when launching the instance, or associates the security group with the instance later on | Automatically applies to all instances in the subnets it's associated with (backup layer of defense, so you don't have to rely on someone specifying the security group) |

# Layers of Security in VPC

- Traffic from an Internet gateway is routed to the appropriate subnet using the routes in the routing table.
- The rules of the network ACL associated with the subnet control which traffic is allowed to the subnet.
- The rules of the security group associated with an instance control which traffic is allowed to the instance.



# Elastic Load Balancing

# What is Load Balancing?

Load balancing means that distributing the total amount of work that a single computer has to do, into two or more computers. As results , computers can to do  more work in the same amount of time and users get served faster . Typically it reduces the time and increases the speed. Load balancing can be implemented with hardware, software, or a combination of both.

Arshdeep Kaur
22326267

# Cloud Load Balancing

Cloud load balancing is the process of distributing workloads and computing resources in a cloud computing environment.It allows enterprises to manage application or workload demands by allocating resources among multiple computers, networks or servers.

Arshdeep Kaur
22326267

Arshdeep Kaur
22326267

# Cloud load balancing technologies

| | |
|---|---|
| Google | Google Cloud Platform offers load balancing for its infrastructure as a service, Google Compute Engine, which distributes network traffic between VM instances. |
| Amazon Web Services (AWS) | AWS offers Elastic Load Balancing, which distributes workloads and traffic among EC2 instances |
| Rackspace | Microsoft Azure's Traffic Manager distributes traffic for its cloud services across multiple data centres. |
| Microsoft Azure | Rackspace's Cloud Load Balancers use multiple servers for workload distribution. |

# AWS Elastic Load Balancing

Elastic Load Balancing automatically distributes incoming application traffic across multiple targets, such as Amazon EC2 instances, containers, and IP addresses. It monitors the health of registered targets and routes traffic only to the healthy targets. It can handle the varying load of your application traffic in a single Availability Zone or across multiple Availability Zones.

Arshdeep Kaur
22326267

# Features of Elastic Load Balancing

1.Security
2.Health Checks
3.High availability
4. Layer 4 or Layer 7 Load Balancing
5. TLS Termination
6. Operational Monitoring

Arshdeep Kaur
22326267

# Elastic Load Balancing Use Cases

1. Achieve Better Fault Tolerance for Your Applications.
2. Automatically Load Balance your Containerized Applications
3. Automatically Scale Your Applications
4. Using Elastic Load Balancing in your Amazon Virtual Private Cloud (Amazon VPC)
5. Hybrid Load Balancing with Elastic Load Balancing

Arshdeep Kaur
22326267

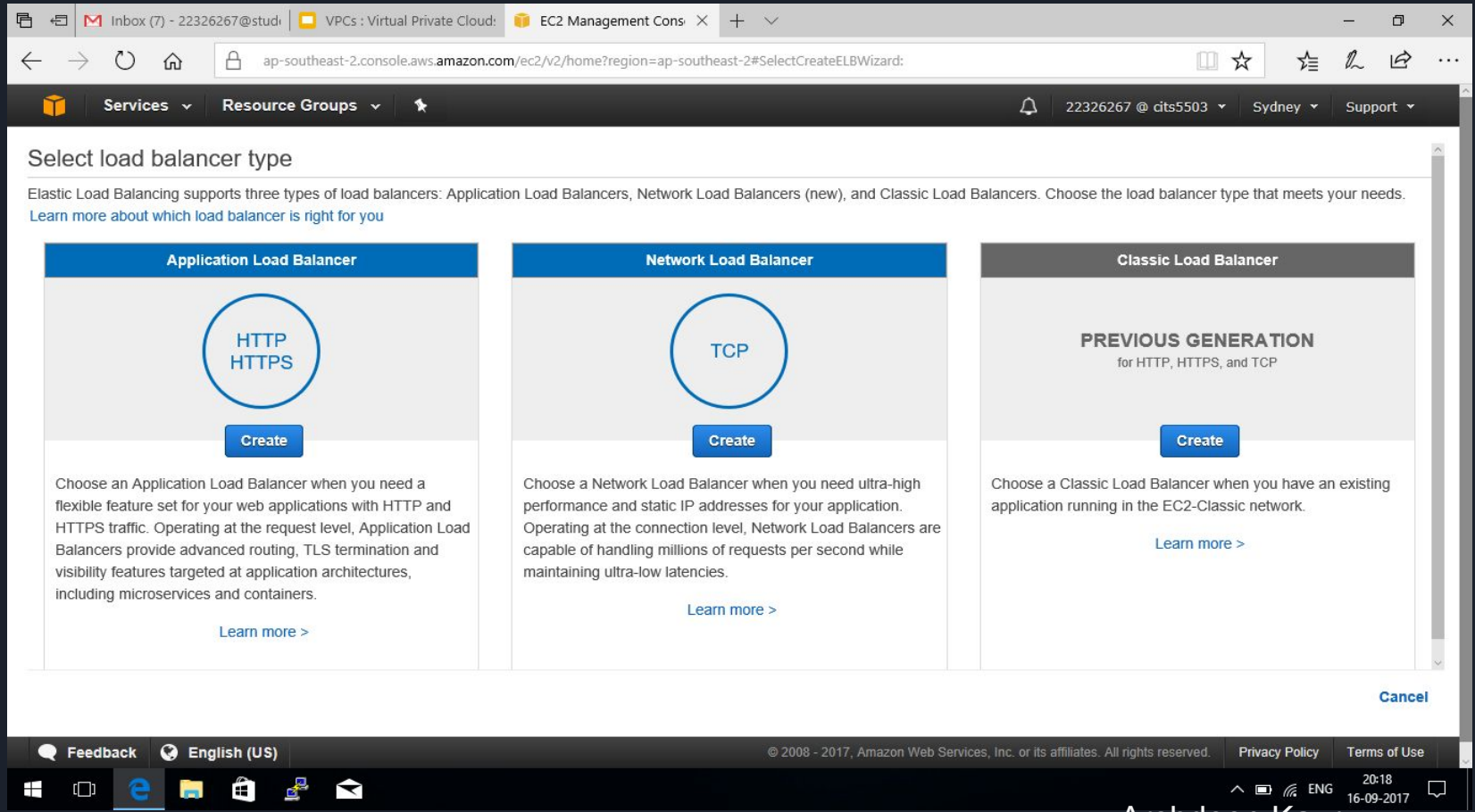
# Types of Elastic Load Balancing

1. Application Load Balancer

2. Network Load Balancer

3. Classic Load Balancer

Arshdeep Kaur
22326267




https://ap-southeast-2.console.aws.amazon.com/ec2/v2/home?region=ap-southeast-2#SelectCreateELBWizard:

Arshdeep Kaur
22326267

# Application Load Balancer

- Application Load Balancer operates at the request level (layer 7), routing traffic to targets - EC2 instances, containers and IP addresses based on the content of the request.
- Ideal for advanced load balancing of HTTP and HTTPS traffic, Application Load Balancer provides advanced request routing targeted at delivery of modern application architectures, including microservices and container-based applications.
- Application Load Balancer simplifies and improves the security of your application, by ensuring that the latest SSL/TLS ciphers and protocols are used at all times.

Arshdeep Kaur
22326267

# Network Load Balancer

- Network Load Balancer operates at the connection level (Layer 4), routing connections to targets - Amazon EC2 instances, micro services and containers – within Amazon Virtual Private Cloud (Amazon VPC) based on IP protocol data.
- Ideal for load balancing of TCP traffic, Network Load Balancer is capable of handling millions of requests per second while maintaining ultra-low latencies.
- Network Load Balancer is optimized to handle sudden and volatile traffic patterns while using a single static IP address per Availability Zone.
- It is integrated with other popular AWS services such as Auto Scaling, Amazon EC2 Container Service (ECS), and Amazon Cloud Formation.

Arshdeep Kaur
22326267

# Classic Load Balancer

- Classic Load Balancer provides basic load balancing across multiple Amazon EC2 instances and operates at both the request level and connection level.
- Classic Load Balancer is intended for applications that were built within the EC2-Classic network. We recommend Application Load Balancer for Layer 7 and Network Load Balancer for Layer 4

  when using Virtual Private Cloud (VPC).

Arshdeep Kaur
22326267

# Reference

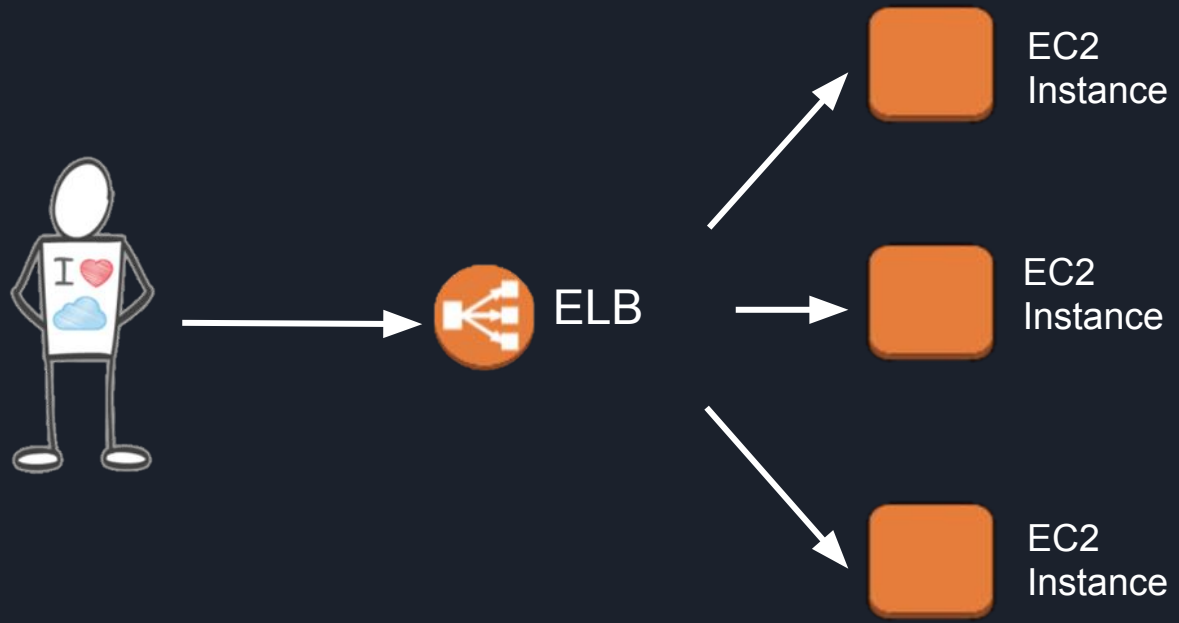
http://searchcloudcomputing.techtarget.com/definition/cloud-load-balancing

https://aws.amazon.com/elasticloadbalancing

Arshdeep Kaur
22326267


EC2
Instance

Yuntian He
22017827




Load Balancer used to route incoming requests to multiple EC2 instances.

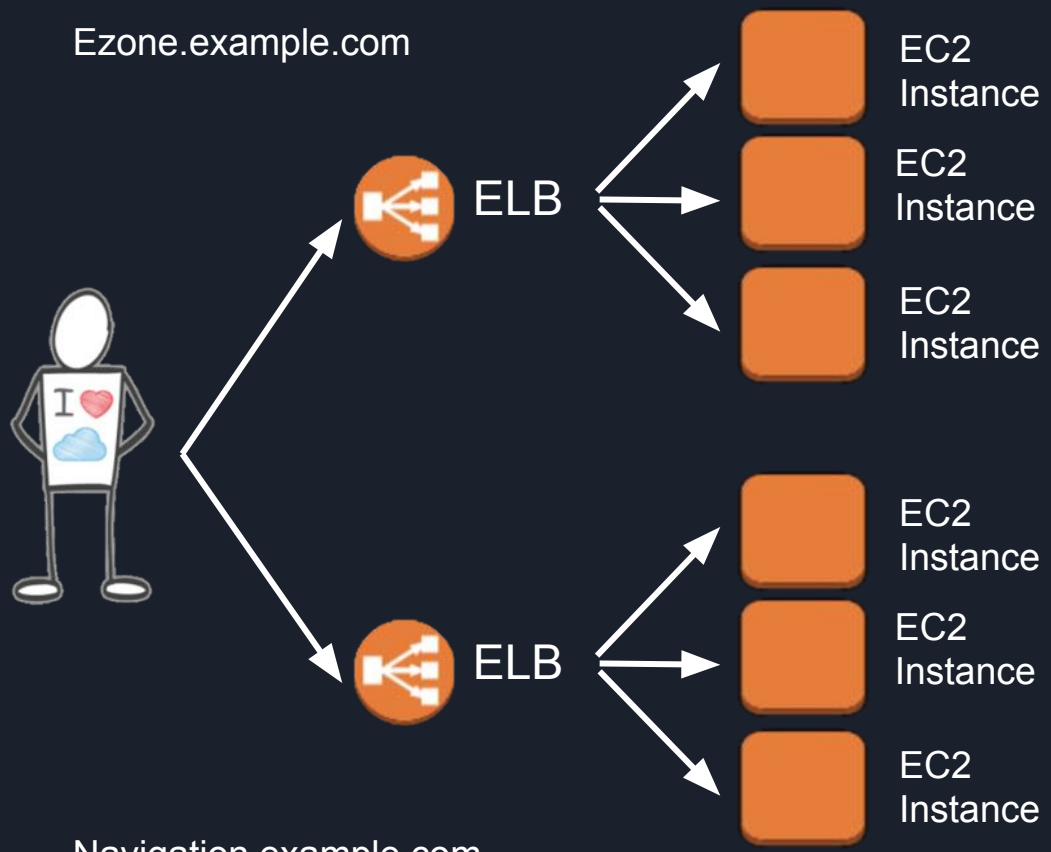
Yuntian He
22017827

| | Classic | Application |
| --- | --- | --- |
| Protocol | TCP, SSL, HTTP, HTTPS | HTTP, HTTPS |
| Platforms | EC2-Classic, EC2-VPC | EC2-VPC |
| Health checks | ✔ | Improved |
| CloudWatch metrics | ✔ | Improved |
| Path-based routing | | ✔ |
| Container support | | ✔ |
| WebSocket & HTTP/2 | | ✔ |

Yuntian He
22017827

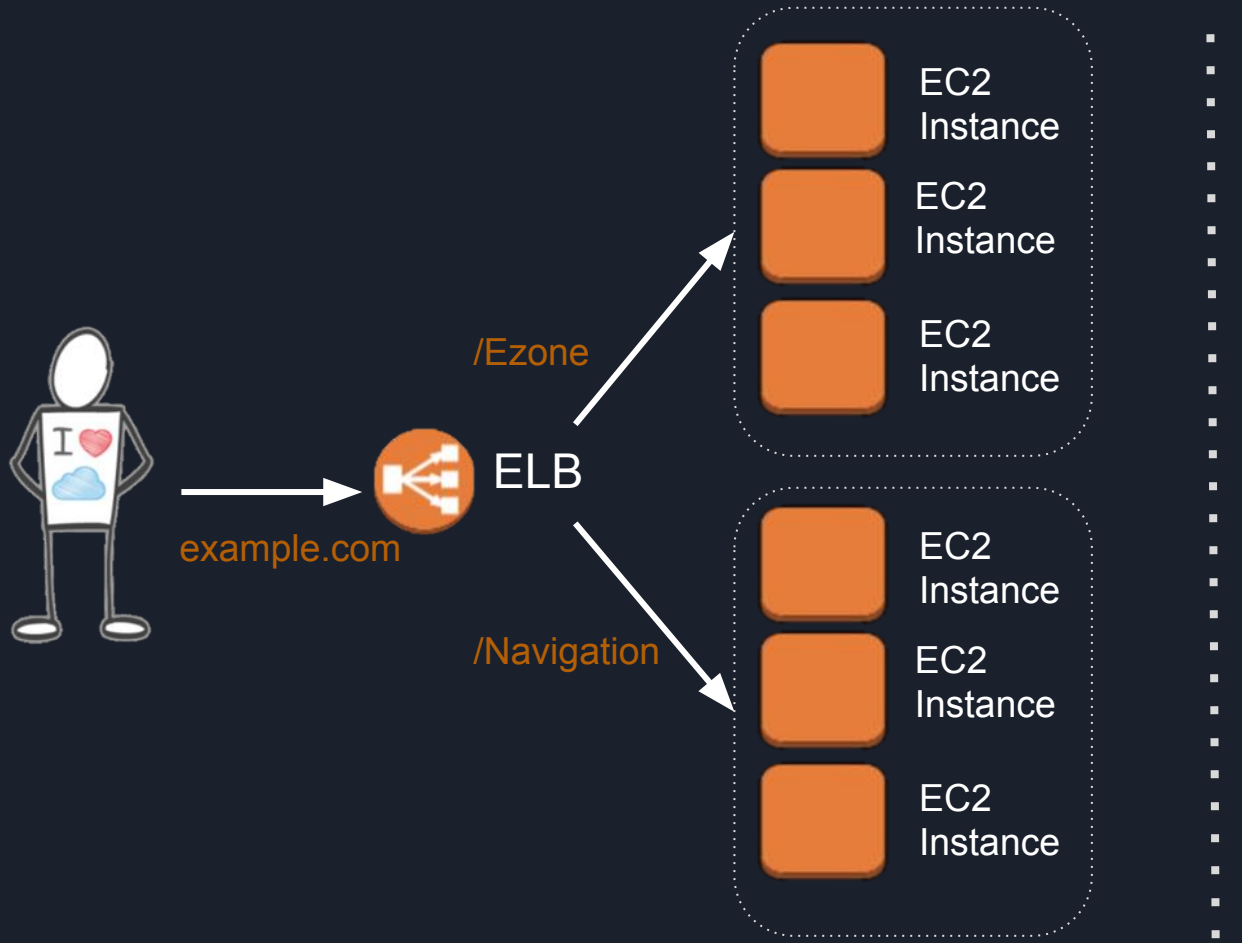
Application Load Balancer allows for multiple applications to be hosted behind a single load balancer

Yuntian He
22017827




Running two separate applications with Classic Load Balancer requires multiple load balancers

Yuntian He
22017827




Application Load Balancer allows for multiple applications to be hosted behind a single load balancer

Yuntian He
22017827

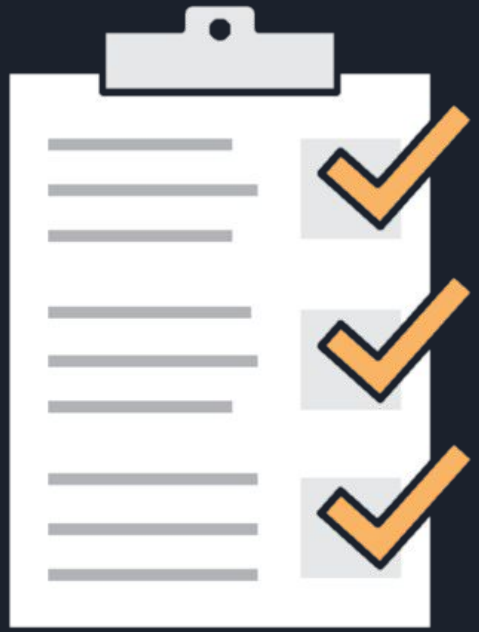

Multiple applications behind a single load balancer provides a significant cost saving

Yuntian He
22017827

# Application Load Balancer provides native support for microservice and container-based architectures

Yuntian He
22017827

# Application Load Balancer

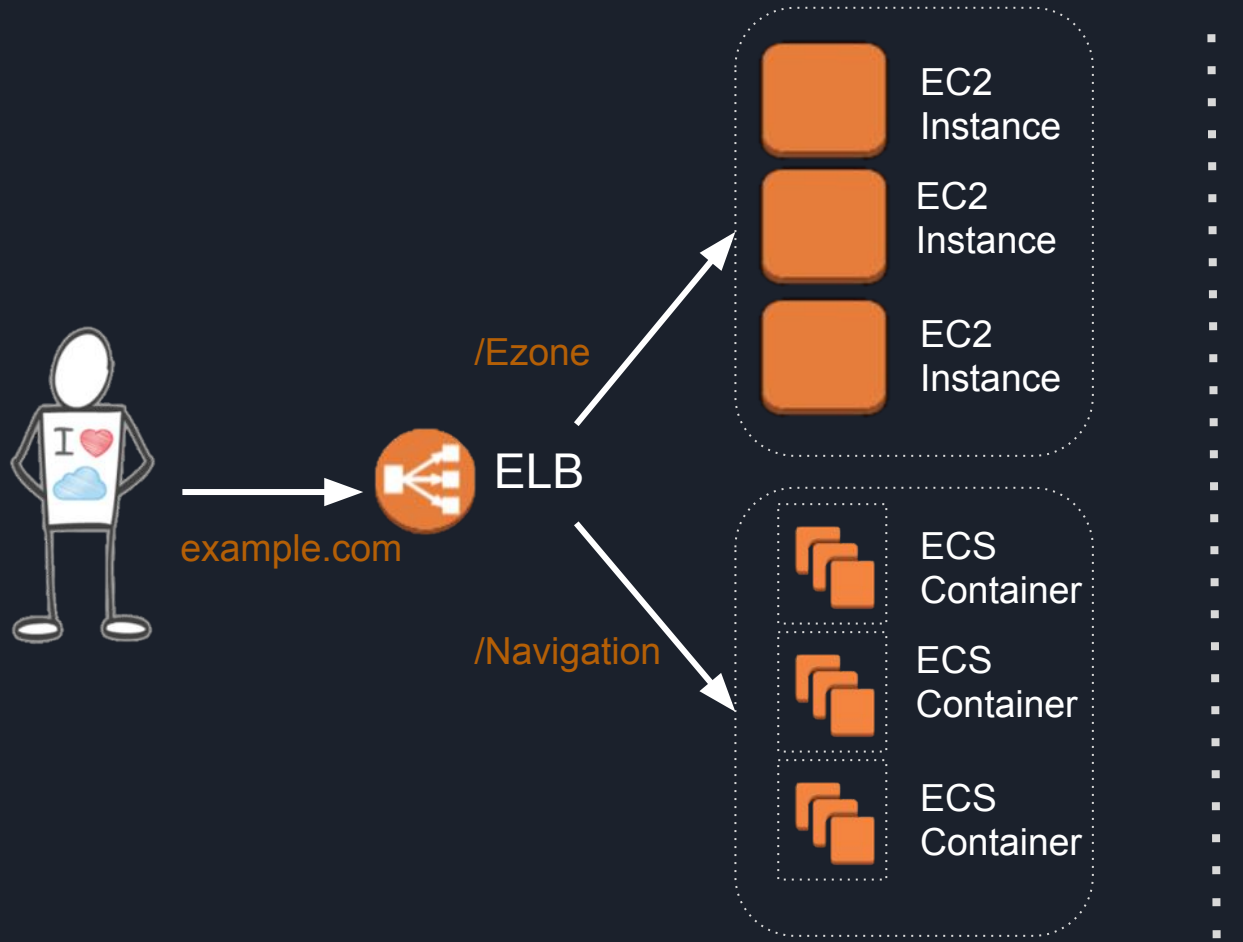


Instances can be registered with multiple ports, allowing for requests to be routed to multiple containers on a single instance

Amazon ECS will automatically register tasks with the load balancer using a dynamic port mapping

Can also be used with other container technologies

Yuntian He
22017827




Application Load Balancer allows containers to be registered with the load balancer

Yuntian He
22017827


Microservice and container-based architectures provide further cost savings by improving resource utilization

Yuntian He
22017827

# Improvements to application scalability, availability and security

Yuntian He
22017827

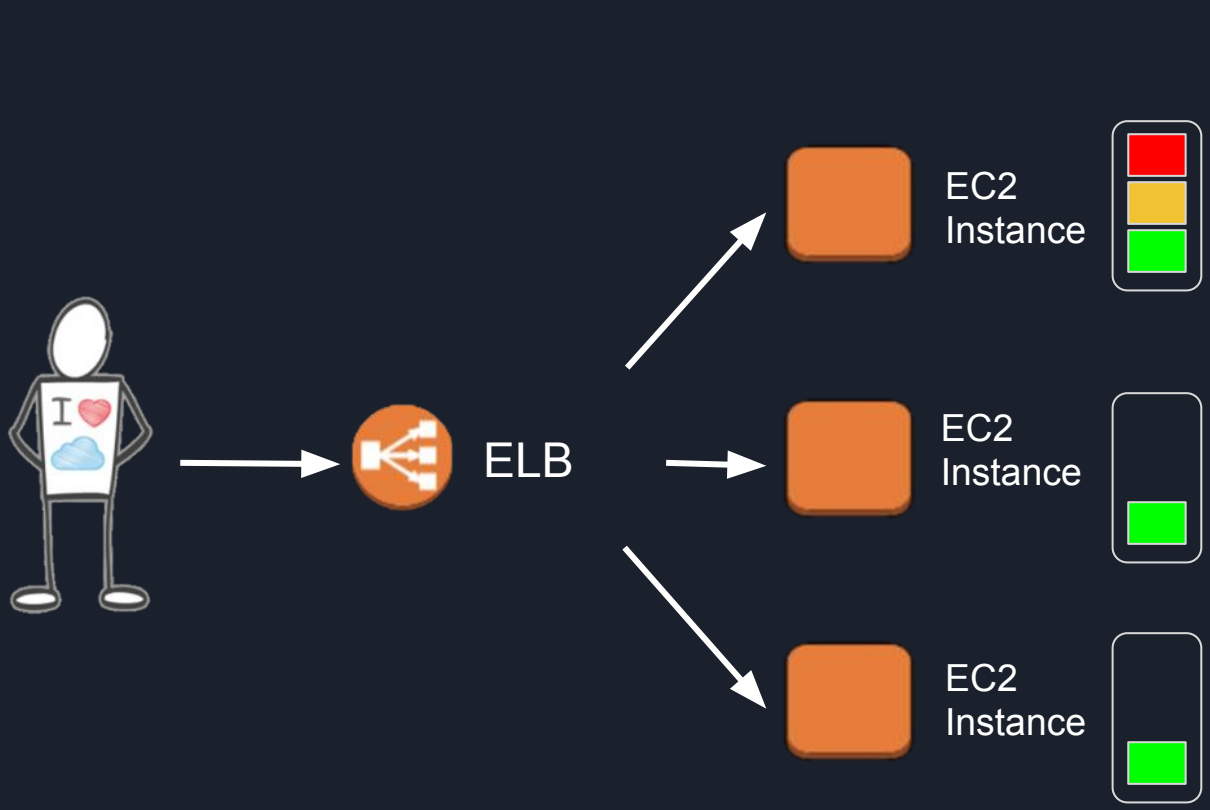

?

EC2
Instance

Yuntian He
22017827

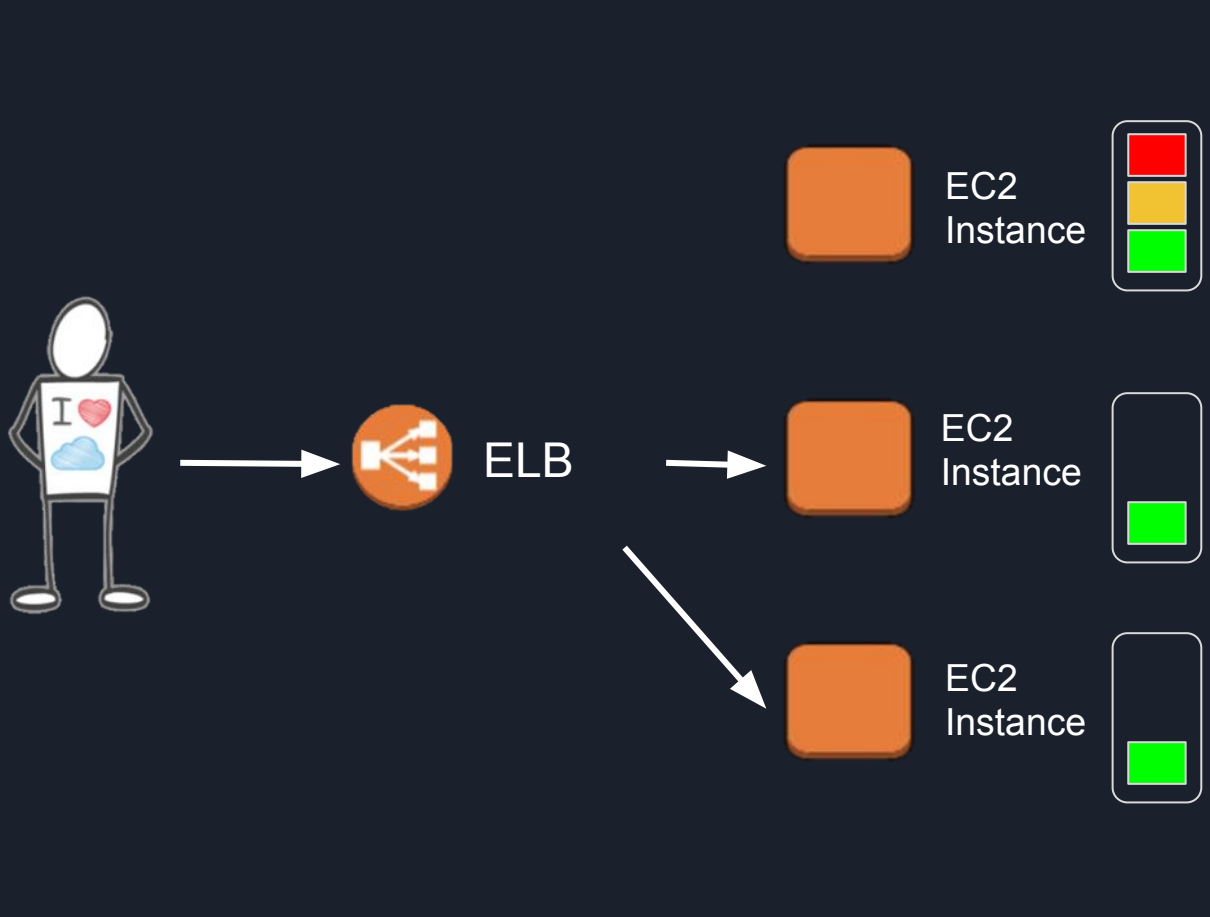
Health checks allow for traffic to be shifted away from impaired or failed instances

Yuntian He
22017827



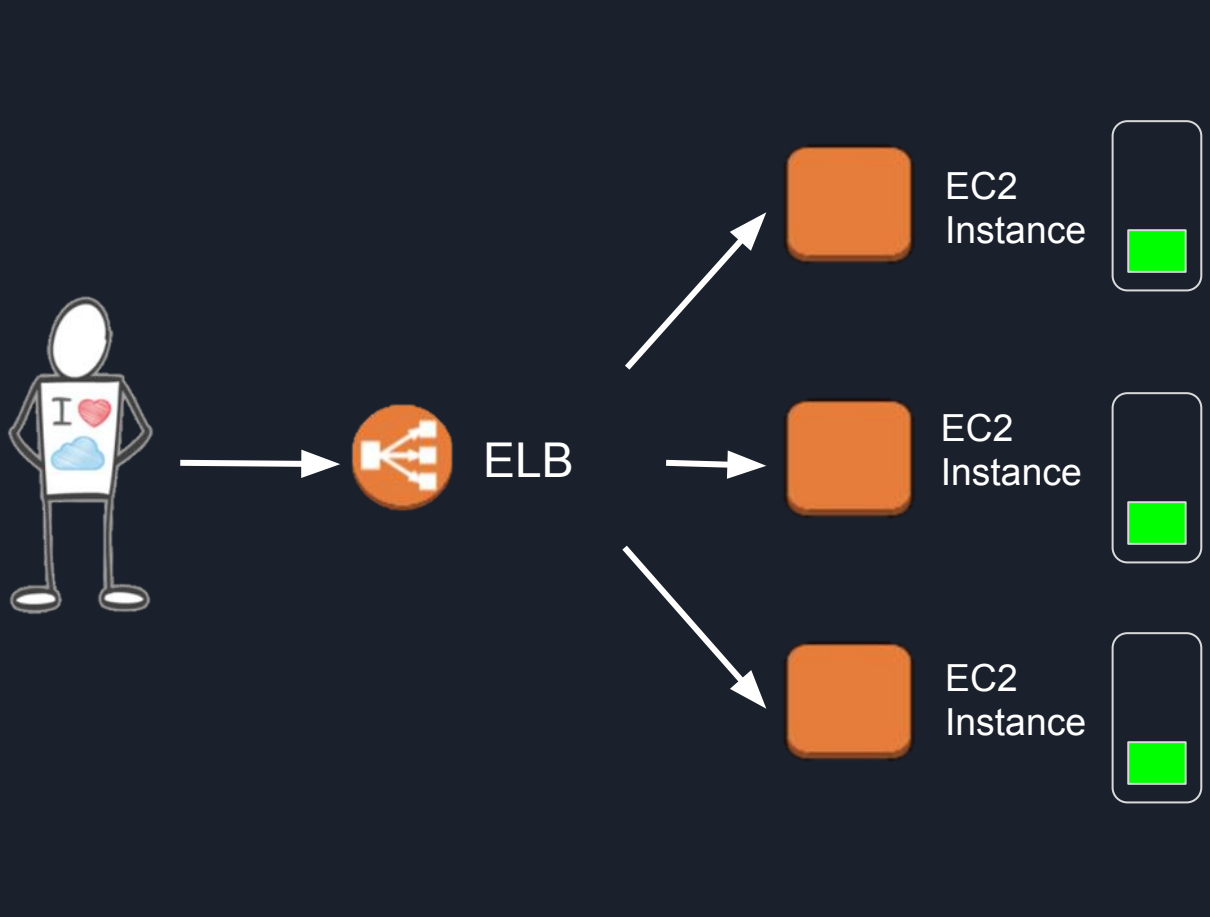

Health checks ensure that request traffic is shifted away from a failed instance

Yuntian He
22017827




Health checks ensure that request traffic is shifted away from a failed instance

Yuntian He
22017827

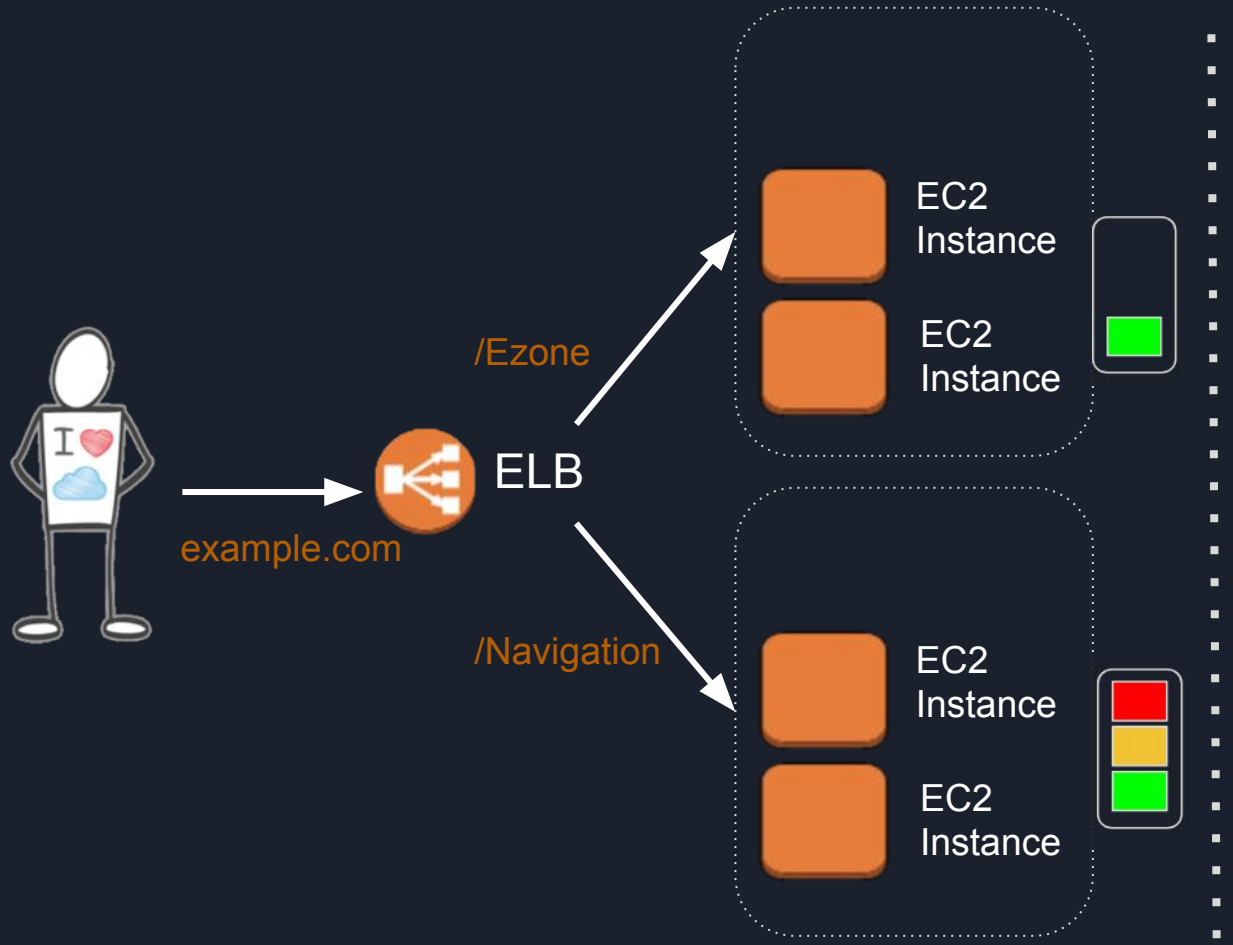



Health checks ensure that request traffic is shifted away from a failed instance

Yuntian He
22017827

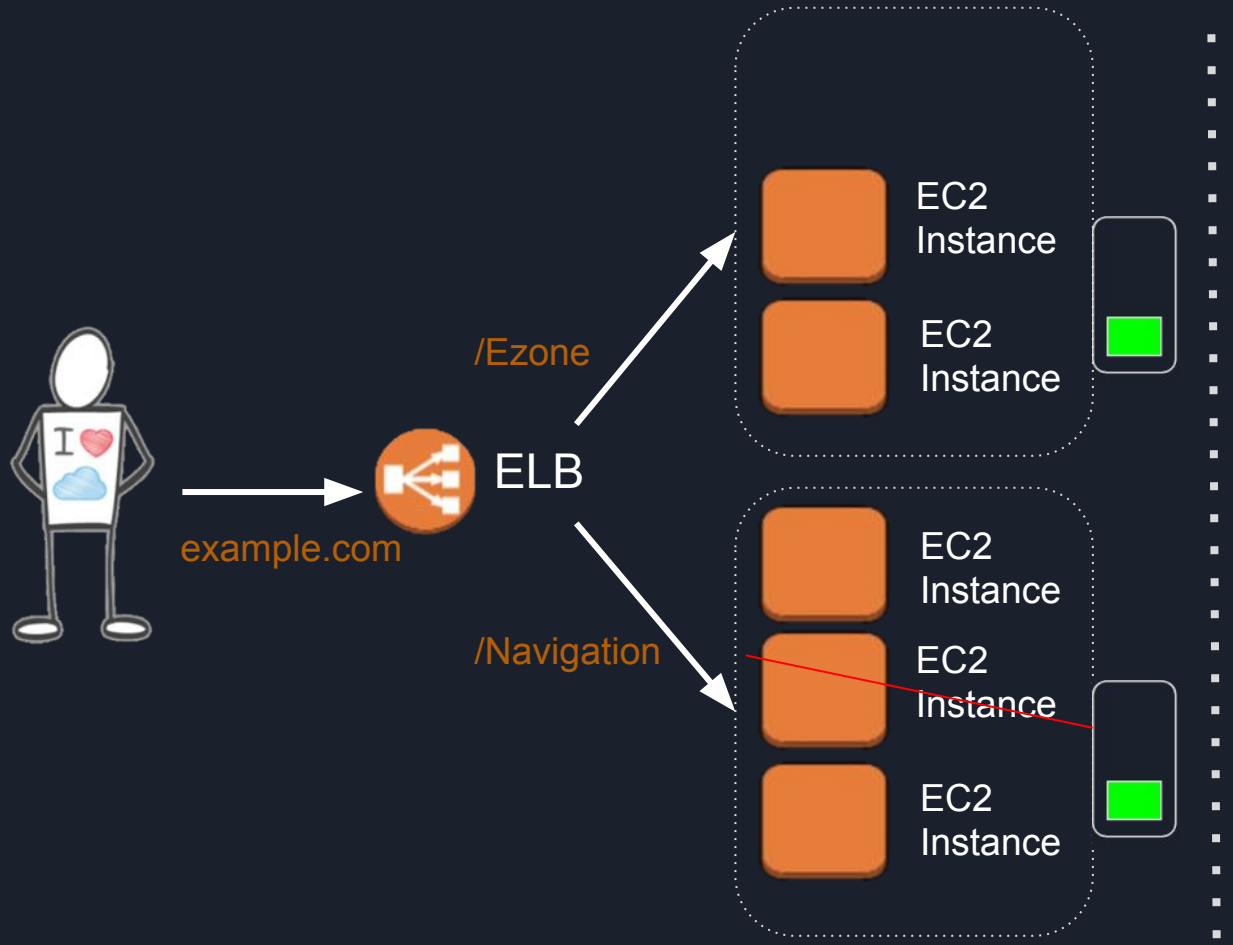
# Auto Scaling now supports the scaling of applications at the target group level

Yuntian He
22017827




Application Load Balancer integrated with Auto Scaling to manage the scaling of each target group independently

Yuntian He
22017827




Application Load Balancer integrated with Auto Scaling to manage the scaling of each target group independently

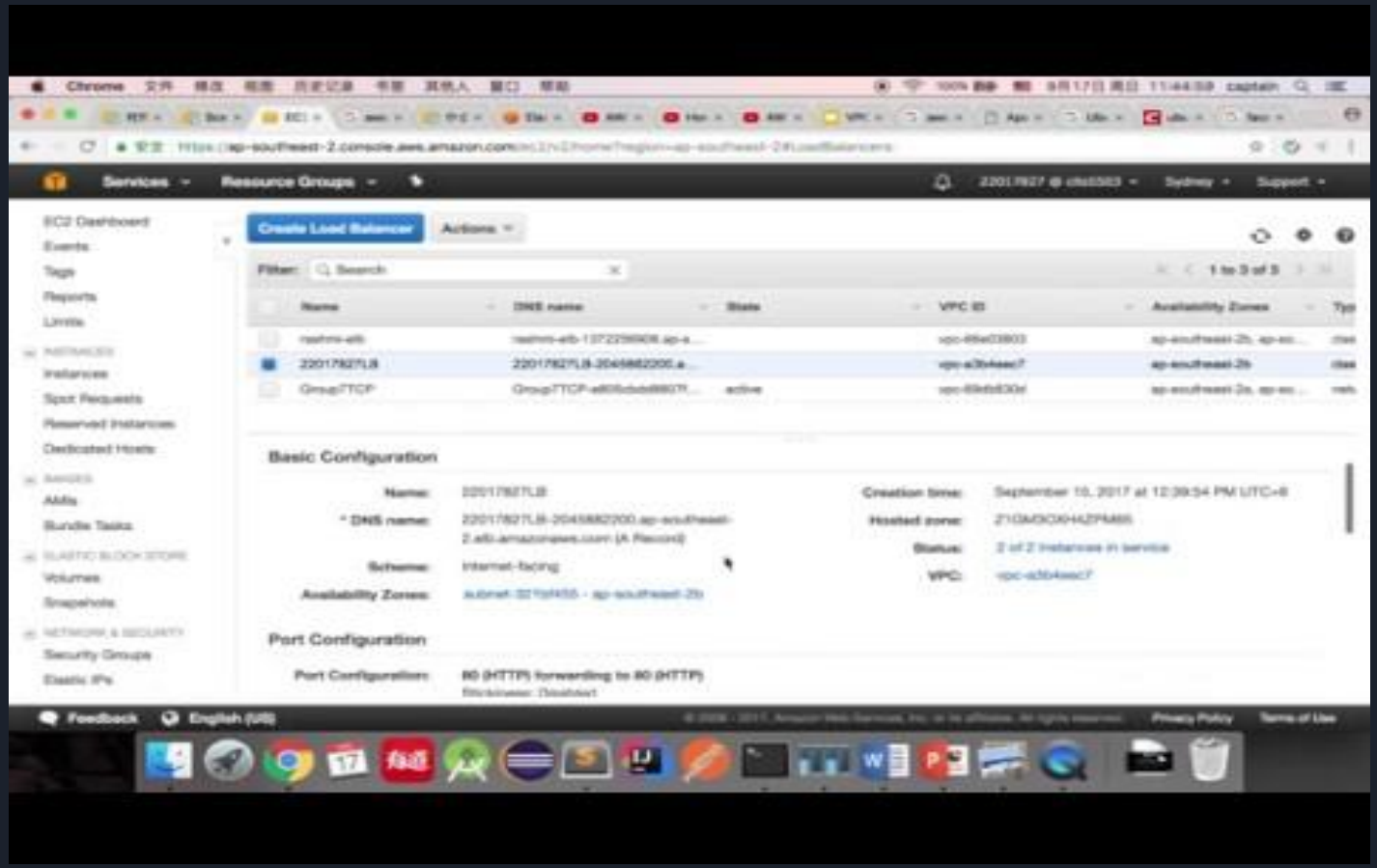
Yuntian He
22017827

# Website Application Firewall

Monitors requests and protects web applications from malicious activities at the load balancer level

Block, allow, or count web requests based on WAF rules and conditions

Yuntian He
22017827




Yuntian He
22017827

# References


Amazon Web Services, 2015. Technical Introduction to Amazon VPC - Virtual Private Cloud on AWS. [Online Video]
Available at: https://www.youtube.com/watch?v=jcyZmj6Ywh4&t=1s [Accessed 10 September 2017].

Amazon Web Services, 2017. AWS Documentation. [Online]
Available at: https://aws.amazon.com/?icmpid=docs_menu_internal [Accessed 10 September 2017].

CBT Nuggets, 2012. MicroNugget: What is Amazon VPC. [Online Video]
Available at: https://www.youtube.com/watch?v=rv7-ec9Wt_U&t=127s [Accessed 10 September 2017].

Cheng, L., 2014. Do Amazon VPC benefits outweigh downsides?. [Online]
Available at: http://searchaws.techtarget.com/tip/Do-Amazon-VPC-benefits-outweigh-downsides [Accessed 12 September 2017].

Earls, A. R., 2017. Virtual private clouds offer an alternative to on-premises computing. [Online]
Available at: http://searchcloudcomputing.techtarget.com/tip/Virtual-private-clouds-offer-an-alternative-to-on-premises-computing
[Accessed 13 September 2017].

Rando, N., 2015. Ensure project success with private cloud planning. [Online]
Available at: http://searchcloudcomputing.techtarget.com/feature/Ensure-project-success-with-private-cloud-planning
[Accessed 12 September 2017].

Simplilearn, 2016. Amazon Virtual Private Cloud (VPC) | AWS Tutorial For Beginners | Simplilearn. [Online Video]
Available at: https://www.youtube.com/watch?v=fpxDGU2KdkA&t=205s
[Accessed 10 September 2017].

What Is My IP Adress.com, 2017. What is Network Address Translation?. [Online]
Available at: http://whatismyipaddress.com/nat [Accessed 15 September 2017]

# References


- AWS re:Invent 2016: Elastic Load Balancing Deep Dive and Best Practices https://www.youtube.com/watch?v=qy7zNaDTYGQ
- What is Amazon VPC?: http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/VPC_Introduction.html
- VPCs and Subnets: http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/VPC_Subnets.html
- IP Addressing in Your VPC: http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/vpc-ip-addressing.html
- Security: http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/VPC_Security.html
- Classless Inter-Domain Routing (CIDR) Overview: http://www.wirelesstek.com/cidr.htm