Virtual Private Cloud

Consider a village, where people are lazy and they do not want to build & construct their own houses. Seeing this, a businesswoman sees profit, acquires a big plot of land, and starts building houses to be rented. She builds houses, rents them make a lot of money. But, in the village, as in any village, people are of different status, profession and cultures. Some people go to the businesswoman showing concern of security in the currently built houses that the houses are so nearby that a person from one house can easily be a threat to another. In sum, the concern was regarding isolation & security. So, the businessman decides to build houses in colonies where a set of people who identify themselves distinctly in the village can buy/rent the whole colony.

The colony consisted of well-mapped houses, with an entrance-gateway. The gateway had security to allow people with authorization only. Secondly, it had a guide near the entrance to direct the incoming person to the address he wants to reach. Thirdly, at the address there is one more security person to authorize the entry of the person in the house. Hence, she solved the problem of security.

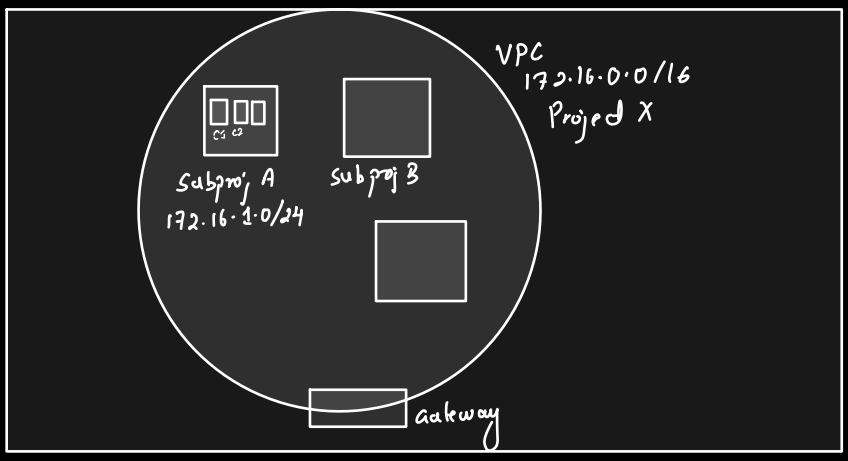
Now, how does this scenario relate to VPC?

In the wake of internet revolution, large amount of data was downloaded and uploaded daily by large number of companies. Therefore, it became inefficient for companies to construct, maintain and upgrade their data centers timely. Therefore AWS saw this opportunity and it build a large data center for rent. Companies immidiately switched to this rental data centers.

Until 2013-14, everything was smooth. But, soon after that companies became reluctant to rely upon aws due to security issues. It is because, although AWS did build multiple servers inside same data center, multiple companies operated their EC2 instances under same server. Now, ofcourse if one application in one instance was hacked all the other companies had to suffer. This was a big problem.

Now, AWS decided to build something called VPC, which is very similar to the colonies mentioned above. VPC is a logical grouping of servers in a specified network. The servers that you are going to deploy in VPC will be completely isolated from the other servers that are deployed in VPC. Now, practically, a devops engineer of a company requests for a VPC from AWS. AWS then gives the devops engineer an

IP address range. Say, 172.16.0.0/16. Now, say the devops engineer (DE), needs this for a project, he can create a subnet within this range and split for the project. Say the project has 20 subproject, 300 components and needs 60,000 instances / servers or whatever. In that range 172.16.0.0/16, he can deploy 255x255 of such servers. For each component he will deploy a range, for each subproject also a range.

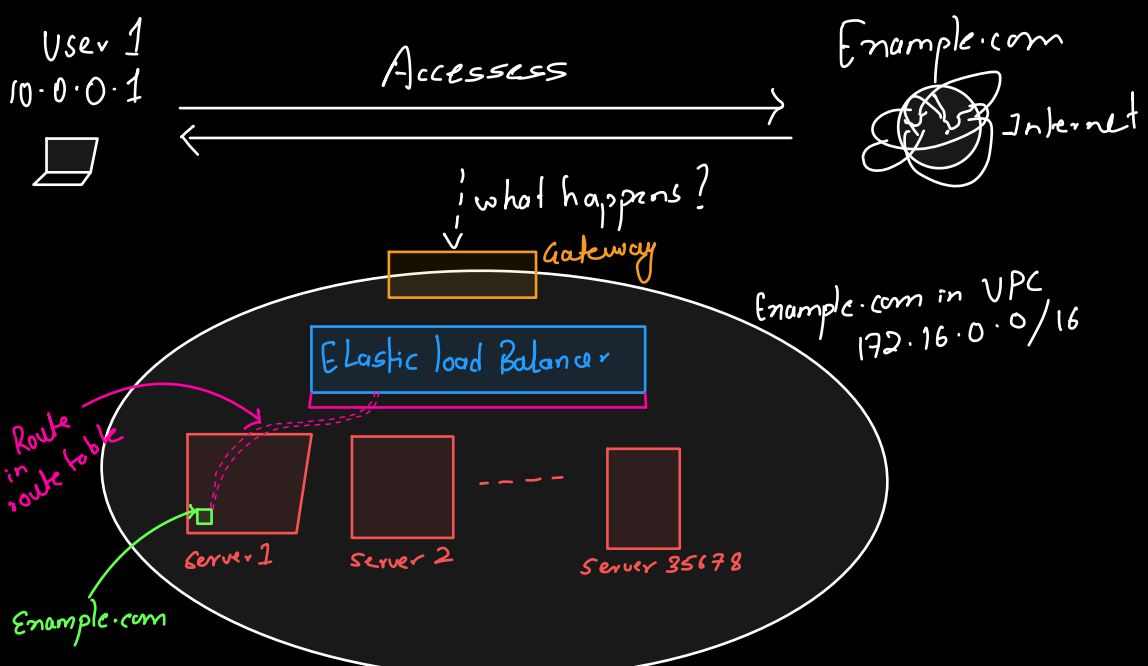


Now, how do individuals in this private IP range connect to the internet. ∴ The devops engineer will also deploy a public subnet. So, through the public subnet and internet gateway in VPC, any server can connect

to the internet.

Now, since there are thousands of IP addresses in the VPC, how will a request make it to each address. Thus, after public subnet there is a load balancer which is connected to all the servers via routers / route tables. Route table defines the path within the network. But there is a firewall / security group that sets the ingress / egress rules to each subnetwork.

A visual representation:



En: VPC

• What is ELB?

Elastic load balancer is a service provided by Amazon in which the incoming traffic is efficiently and automatically distributed across a group of backend servers in a manner that increases speed and performance. It helps to improve the scalibility of your application and secures your application

Advantages of Elastic Load Balancer

- ELB automatically distributes incoming application traffic across multiple targets, such as EC2 instances, containers, and IP addresses, to achieve high availability.
- It can automatically scale to handle changes in traffic demand, allowing you to maintain consistent application performance.
- It can monitor the health of its registered targets and route traffic only to the healthy targets.
- It evenly distributes traffic across all availability zones in a region, improving fault tolerance.

doing a