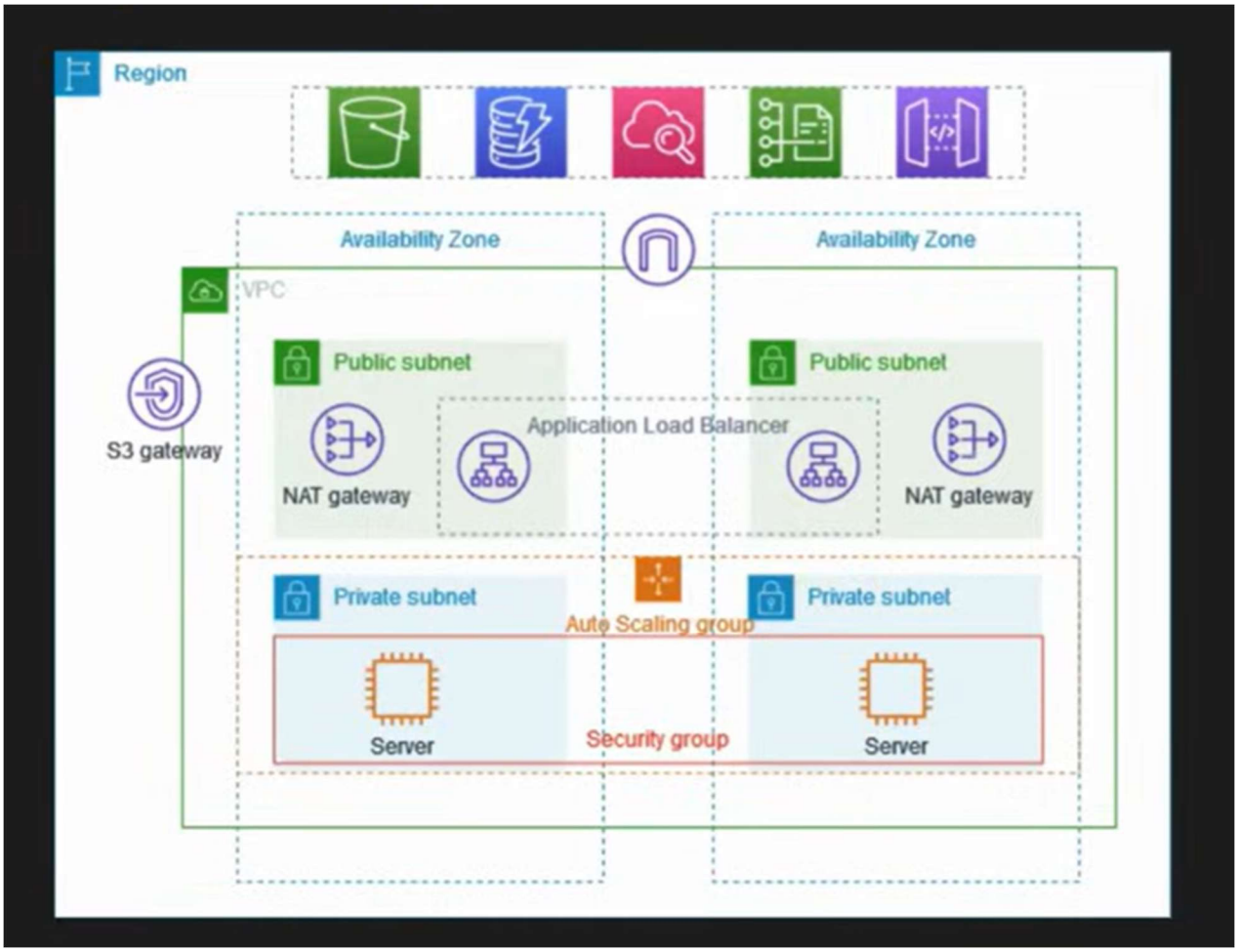
Name:Roshan Saral Kumar

AWS-PROJECT-1
VPC WITH PUBLIC-PRIVATE SUBNET IN PRODUCTION

- **ARCHITECTURE:-**



- **DESCRIPTION:-**

➢ Using Amazon Web Services (AWS), this project showcases a scalable and reliable cloud infrastructure design that is suited for production-grade deployments. To guarantee high availability and fault tolerance, the architecture makes use of a Virtual Private Cloud (VPC) that is configured with both public and private subnets deployed over two availability zones.

- Key Components and Features:

➢ VPC Configuration: A logically separated network environment that safely hosts all Amazon Web Services resources.

➢ Public Subnets: These subnets are deployed in each Availability Zone and host NAT Gateways, allowing private instances to have secure outbound internet access.

- Private Subnets: Dedicated to application servers and shielded from direct internet access for added security.

- The Application Load Balancer (ALB): distributes incoming traffic across many instances across private subnets to ensure load balancing and fault tolerance.

- Auto Scaling Group: Optimises cost and performance by automatically adjusting the number of EC2 instances based on traffic demand.

- Security Groups: Serve as instance-level virtual firewalls to regulate incoming and outgoing traffic.

- Accessing Amazon S3 securely and effectively without using the public internet is made possible by the S3 Gateway Endpoint

- **ADVANTAGES:-**

  - High Availability: Resilience against zone-level failures is ensured by multi-AZ deployment.

  - Scalability: Adapting to traffic patterns, Auto Scaling and ALB offer dynamic resource management.

  - Security: A secure working environment is ensured by NAT, security groups, and the separation of public and private resources.

  - Cost optimisation: Effective usage of scaling tools and NAT gateways aids in controlling operating expenses.

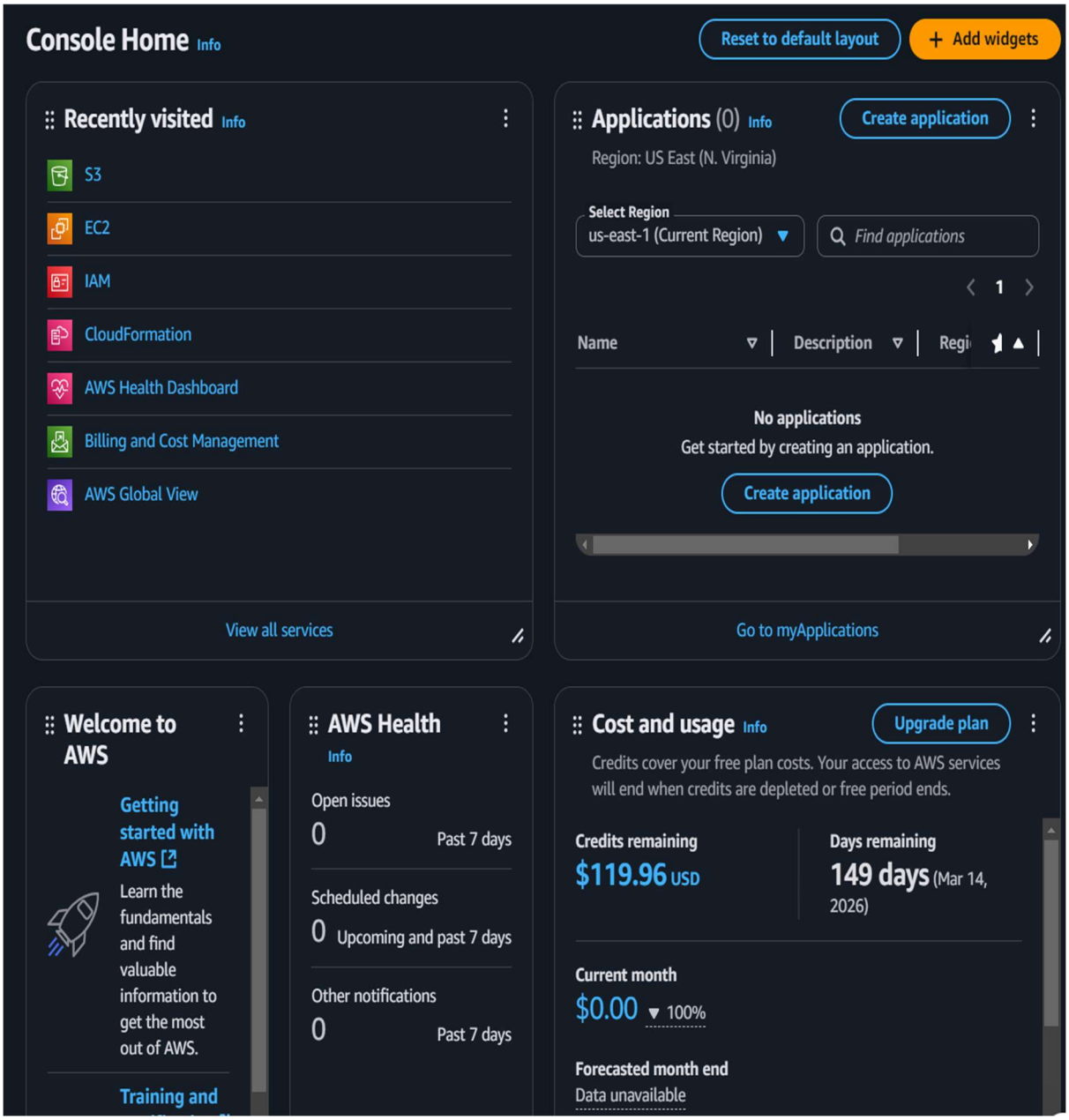
- **ABOUT THE PROJECT:-**

  - This example demonstrates how to create a VPC that you can use for servers in a production environment.

  - To improve resiliency, you deploy the servers in two Availability Zones, by using an Auto Scaling group and an Application Load Balancer. For additional security, you deploy the servers in private subnets. The servers receive requests through the load balancer. The servers can connect to the internet by using a NAT gateway. To improve resiliency, you deploy the NAT gateway in both Availability Zones.

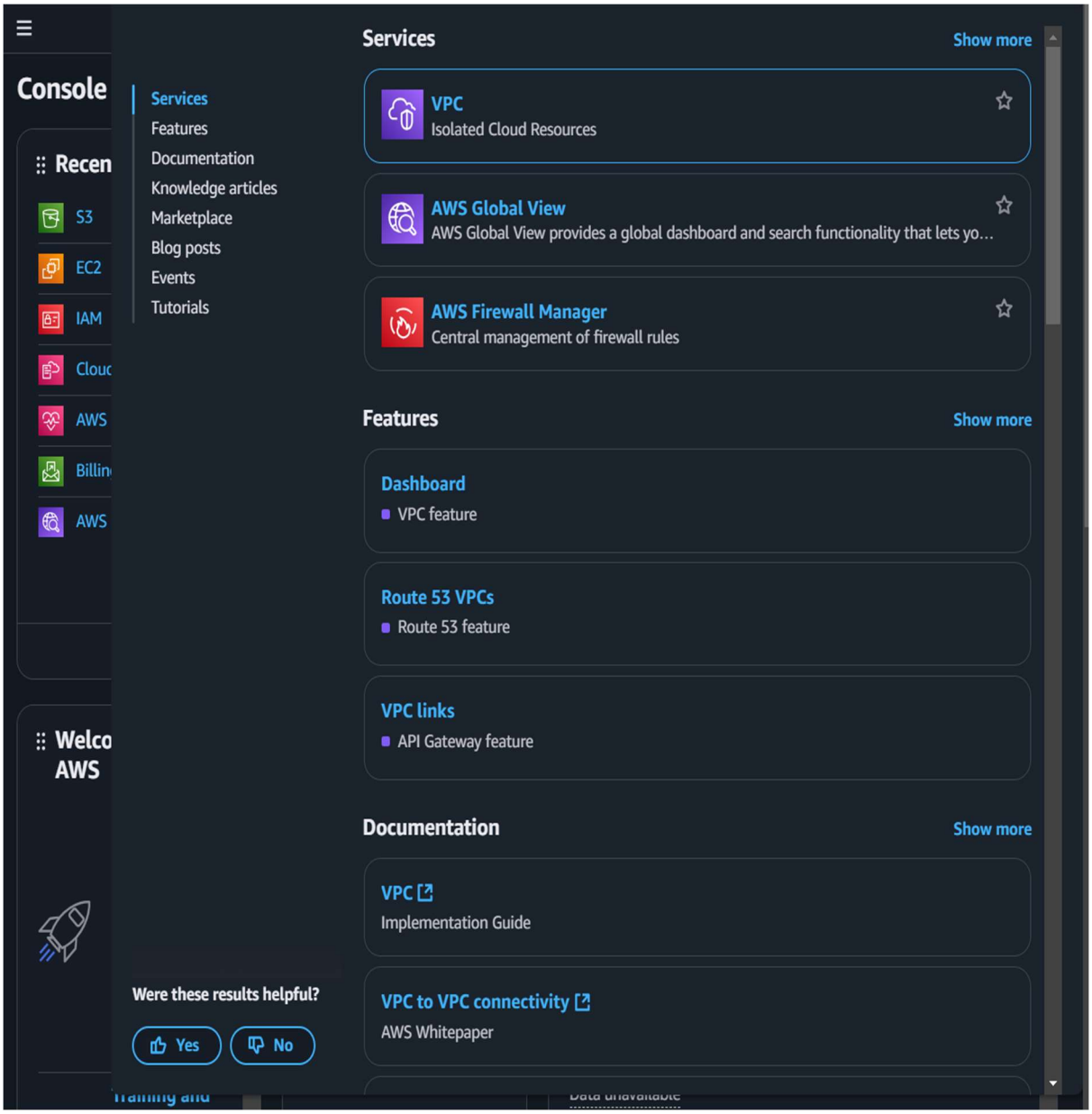
- PROJECT IMPLEMENTATION:-

STEP 1:- Logging in as root user and entering into my aws management console.

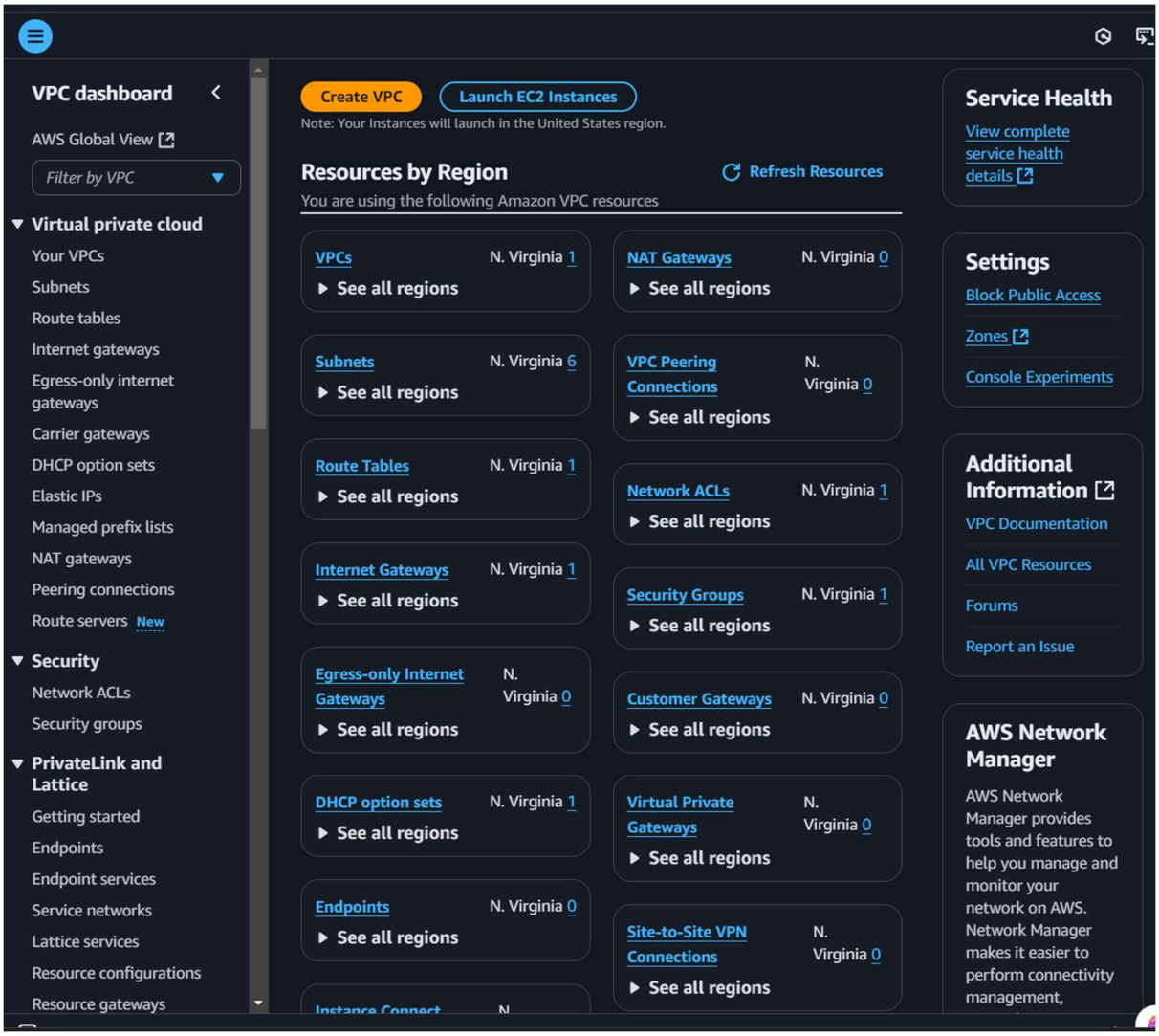


\

Step 2:- SEARCHING FOR VPC ISOLATED CLOUD SERVICES IN AWS MANAGEMENT CONSOLE.

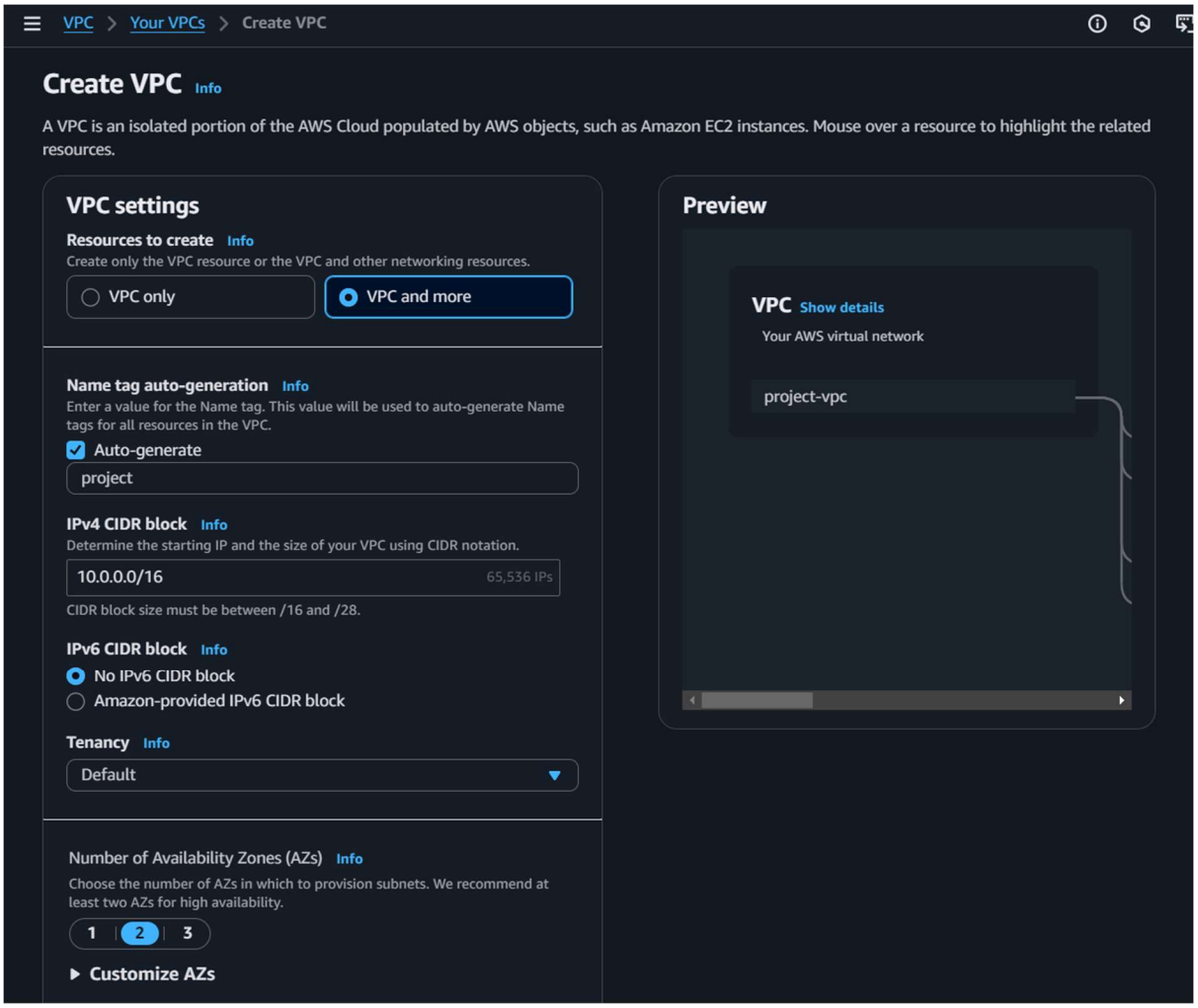


STEP 3:- Once we click on VPC we have to create the VPC AT THE VPC DASHBOARD BY CLICKING ON " Create VPC ".

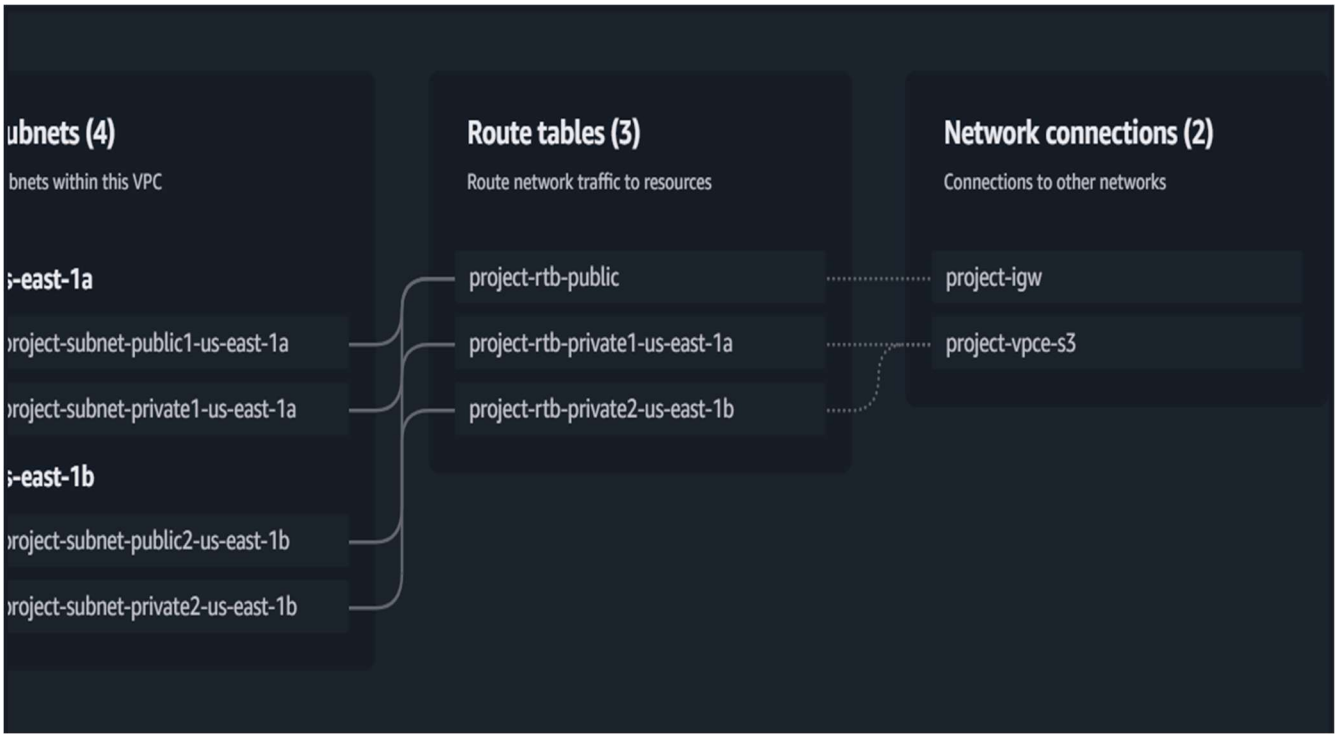


.

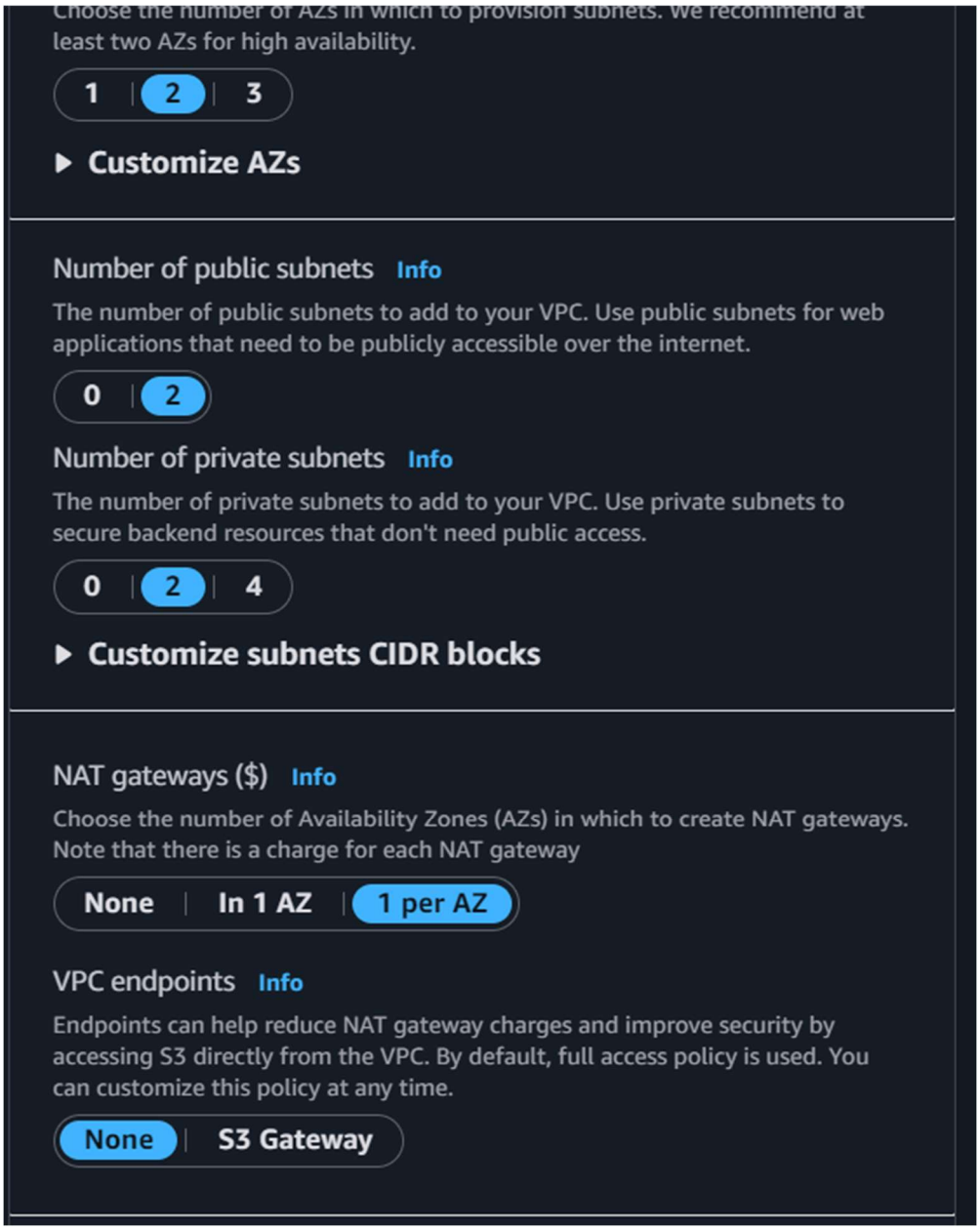
STEP 4: AFTER CLICKING ON CREATE VPC WE HAVE TWO OPTIONS ONE IS VPC ONLY AND THE OTHER IS VPC AND MORE. CHOOSE VPC AND MORE SO THAT WE DON'T HAVE TO MAKE MANUAL SUBNETS IF WE WANT MANUAL SUBNETS OR VPCS WE CAN GO FOR VPC ONLY.



➔ VPC AND MORE GIVES US A PREVIEW DIAGRAMS:-

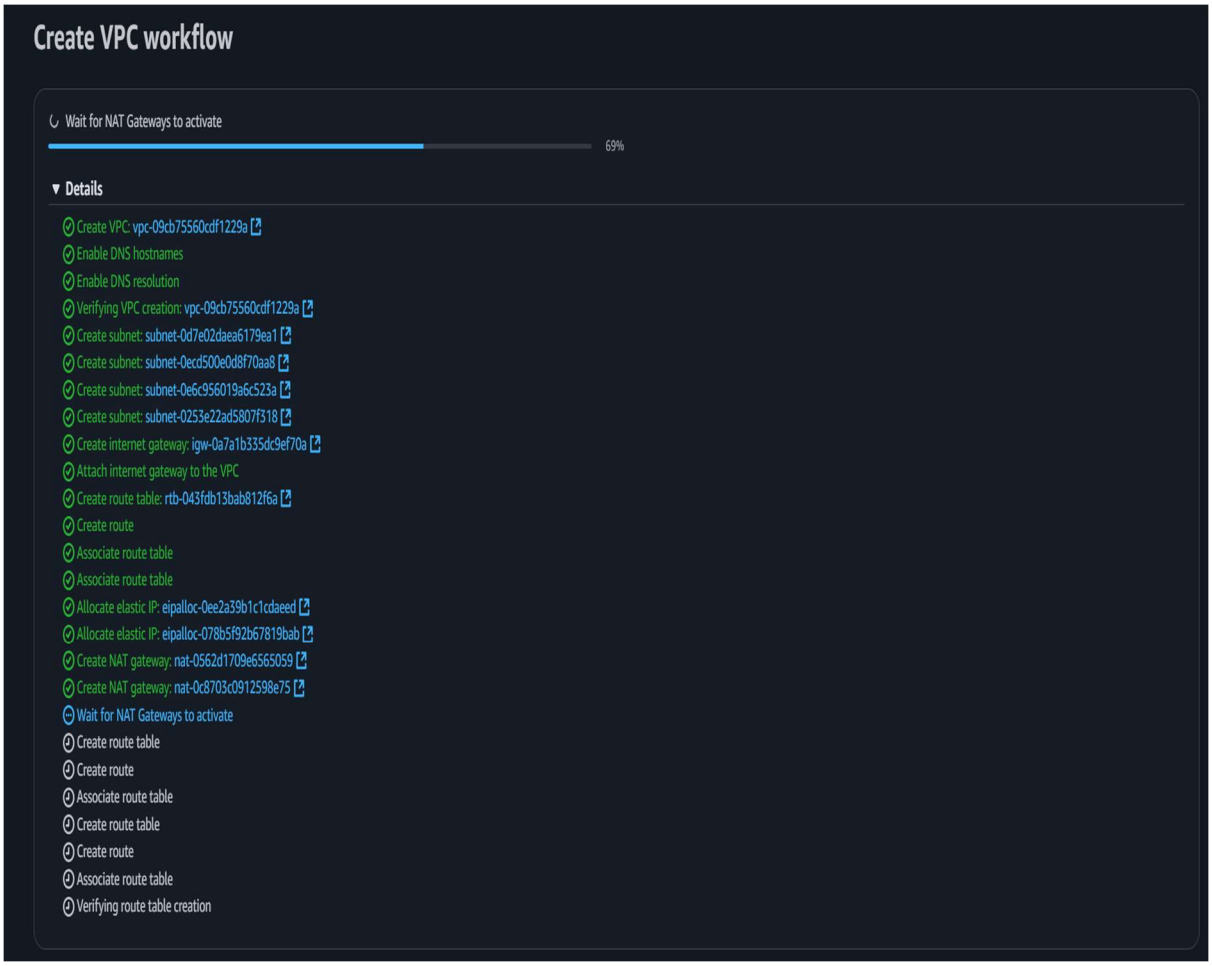


Step 5:AFTER THE PREVIEW IS SEEN WE ARE TRYING TO ELIMINATE THE S3 BUCKET AS IT IS OFF NOT MUCH USE IN THE FOLLOWING OPTIONS AND WE ARE TRYING TO RENAME THE PROJECT AS "aws-prod-example" and everything that is private and public subnets are made by default . so in the ARCHITECTURE GIVEN WE NEED 2 PUBLIC SUBNETS AND 2 PRIVATE SUBNETS ALONG WITH A NAT GATEWAY PER AVAILABILITY ZONE.SO WE ARE MODIFYING THE OPTIONS AND REMOVING S3 BY CLICKING ON THE NONE OPTION FOR VPC ENDPOINTS.

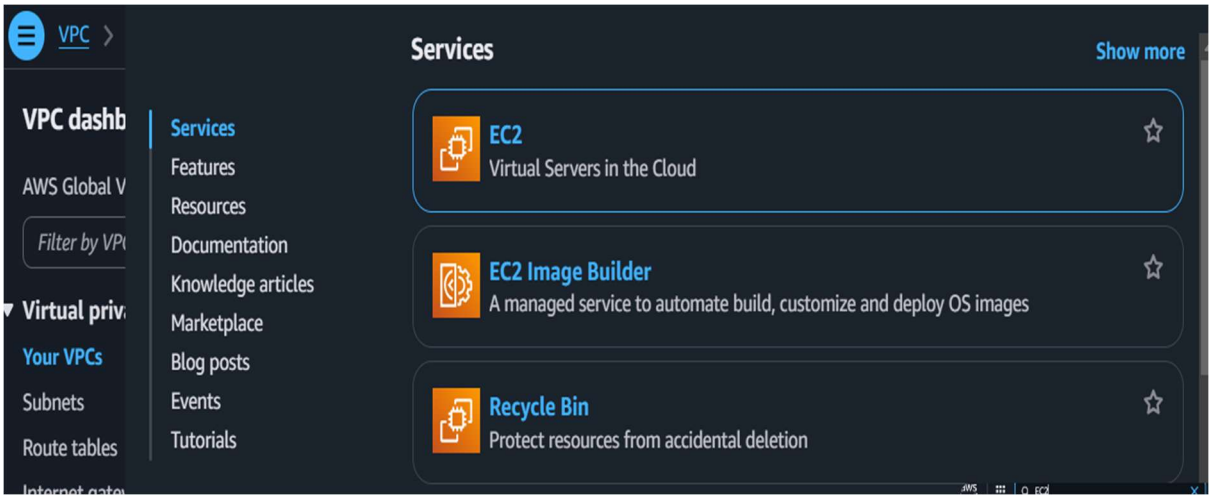


AFTER SELECTING THE RIGHT OPTIONS CLICK ON CREATE VPC.

VPC WORKFLOW AND WE HAVE TO WAIT FOR THE NAT GATEWAYS TO ACTIVATE

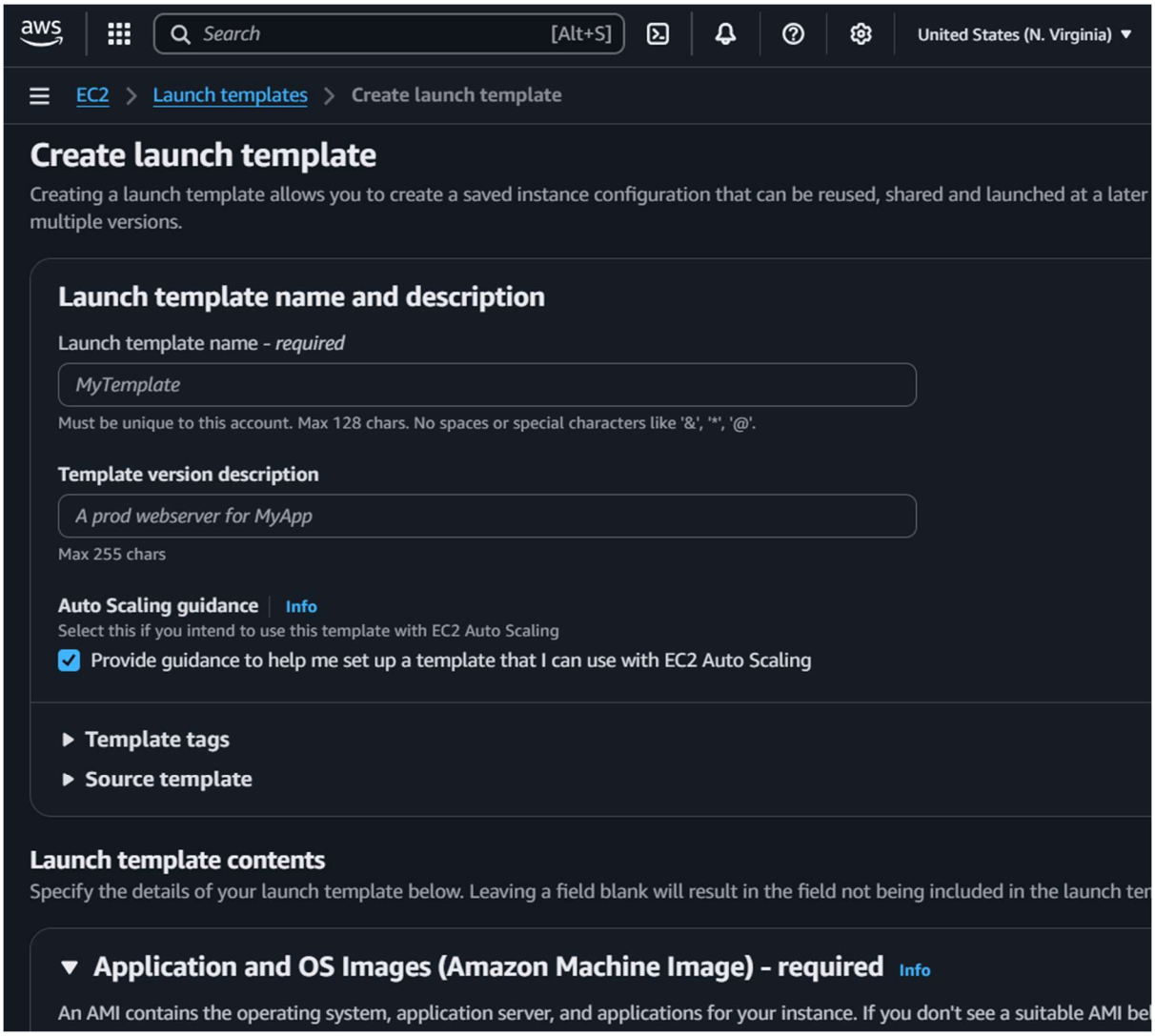


STEP 6: Search for AUTOSCALING GROUP IN EC2 IN AWS MANAGEMENT CONSOLE.



STEP 7:- CLICK ON AUTOSCALING GROUPS IN THE EC2 DASHBOARD AND CLICK ON CREATE AUTOSCALING GROUPS. AND THEN CLICK ON LAUNCH TEMPLATES TO

PROVIDE AUTO SCALING GROUPS EASILY FOR THE TRAFFIC FLOWING IN.
AFTER THIS WE ARE CREATING A LAUNCH TEMPLATE



Select the required options like os,instance type etc don’t chose default vpc choose the one which we made that is “aws-prod-example-vpc” allow ssh and http traffic that is inbound traffic

Create launch template

Creating a launch template allows you to create a saved instance configuration that can be reused, shared and launched at a later
multiple versions.

Launch template name and description

Launch template name - *required*

aws-prod-example

Must be unique to this account. Max 128 chars. No spaces or special characters like '&', '*', '@'.

Template version description

proof of concept for app deployment in aws private subnet

Max 255 chars

Auto Scaling guidance | Info

Select this if you intend to use this template with EC2 Auto Scaling

☑ Provide guidance to help me set up a template that I can use with EC2 Auto Scaling

Quick Start

Amazon Linux | macOS | Ubuntu | Windows | Red Hat | SUSE Linux | Debian

Browse more AMIs

Including AMIs from AWS, Marketplace and the Community

Amazon Machine Image (AMI)

Ubuntu Server 24.04 LTS (HVM), SSD Volume Type — Free tier eligible

ami-0360c520857e3138f (64-bit (x86)) / ami-026fccd88446aa0bf (64-bit (Arm))

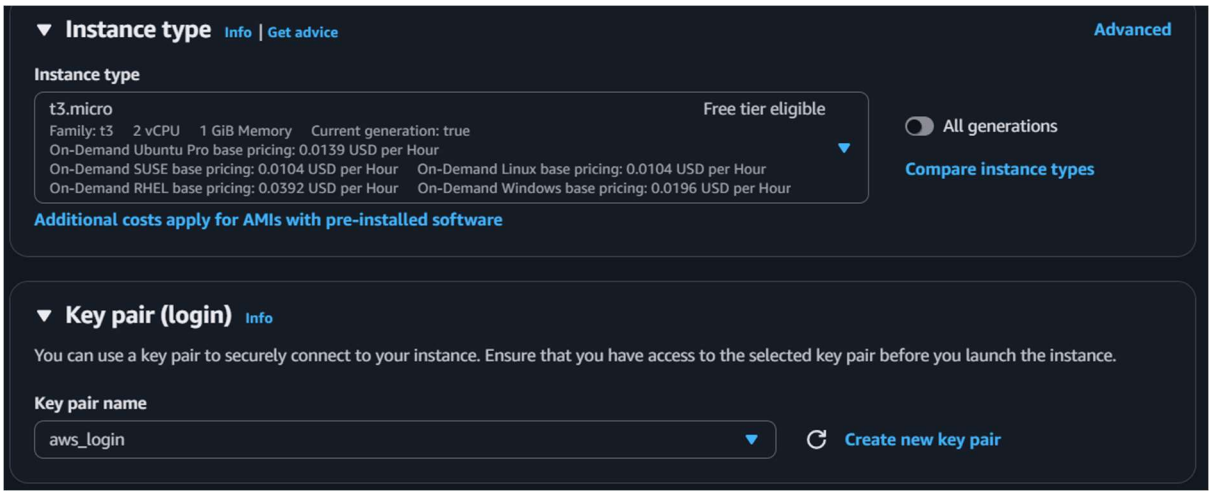
Virtualization: hvm ENA enabled: true Root device type: ebs

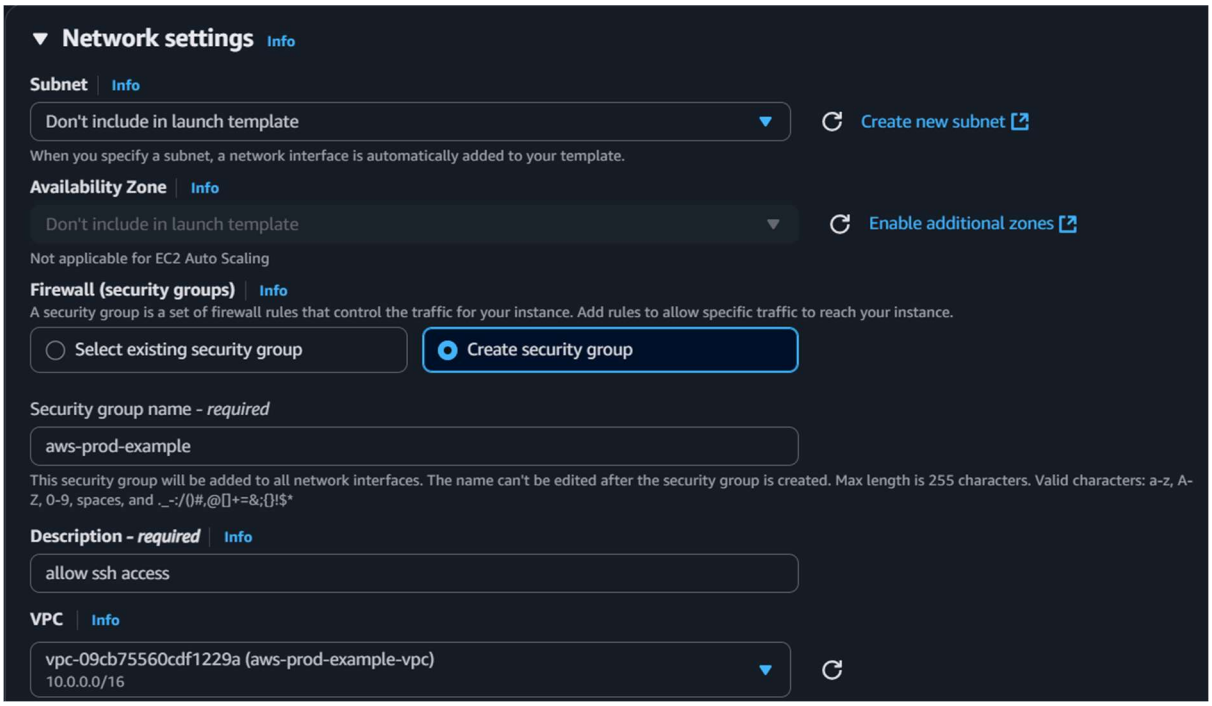
Description

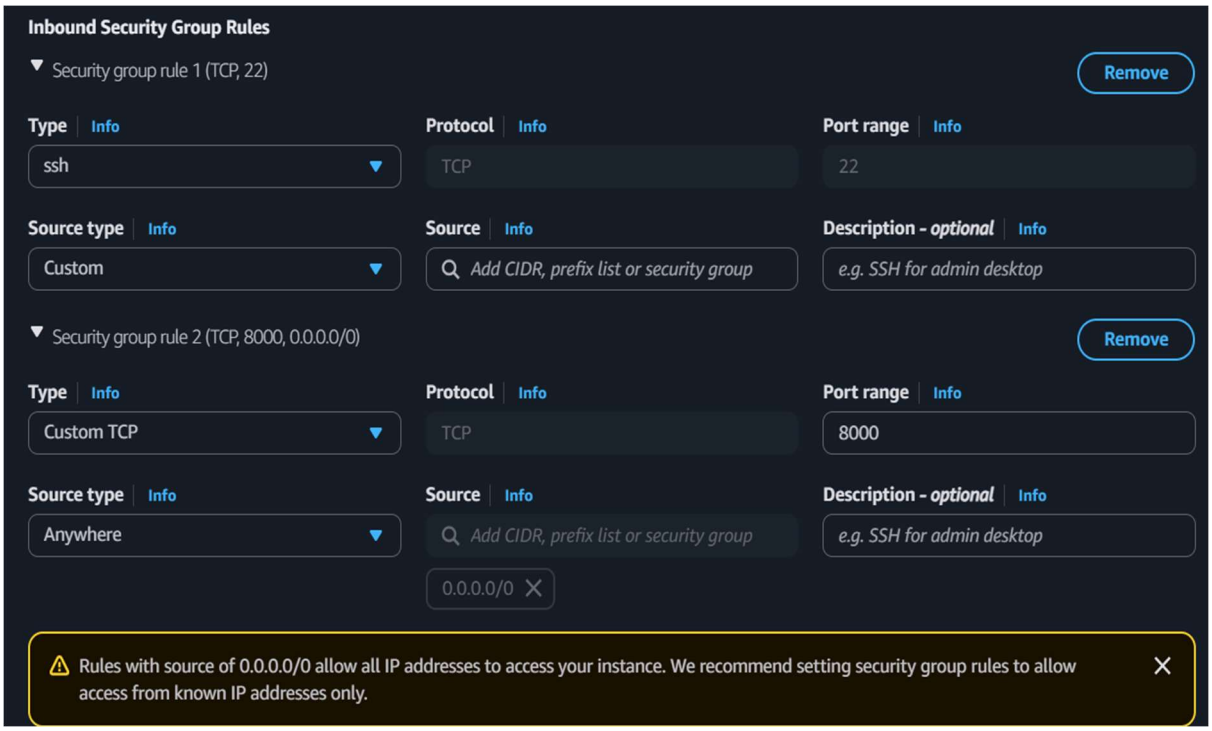
Ubuntu Server 24.04 LTS (HVM),EBS General Purpose (SSD) Volume Type. Support available from Canonical (http://www.ubuntu.com/cloud/services).

Canonical, Ubuntu, 24.04, amd64 noble image

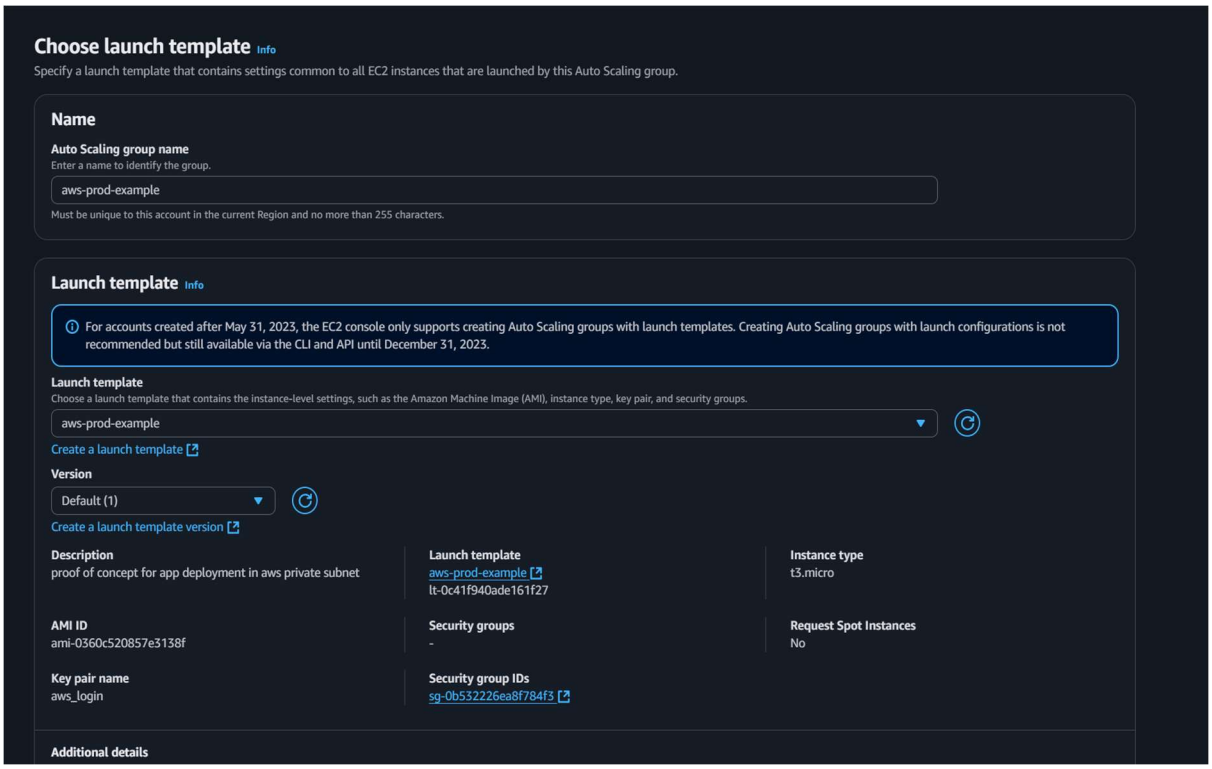
Architecture	AMI ID	Publish Date	Username	
64-bit (x86)	ami-0360c520857e3138f	2025-08-21	ubuntu	Verified provider


Instance type Info | Get advice
Advanced
Instance type
t3.micro
Free tier eligible
Family: t3 2 vCPU 1 GiB Memory Current generation: true
On-Demand Ubuntu Pro base pricing: 0.0139 USD per Hour
On-Demand SUSE base pricing: 0.0104 USD per Hour
On-Demand Linux base pricing: 0.0104 USD per Hour
On-Demand RHEL base pricing: 0.0392 USD per Hour
On-Demand Windows base pricing: 0.0196 USD per Hour
All generations
Compare instance types
Additional costs apply for AMIs with pre-installed software
Key pair (login) Info
You can use a key pair to securely connect to your instance. Ensure that you have access to the selected key pair before you launch the instance.
Key pair name
aws_login
Create new key pair

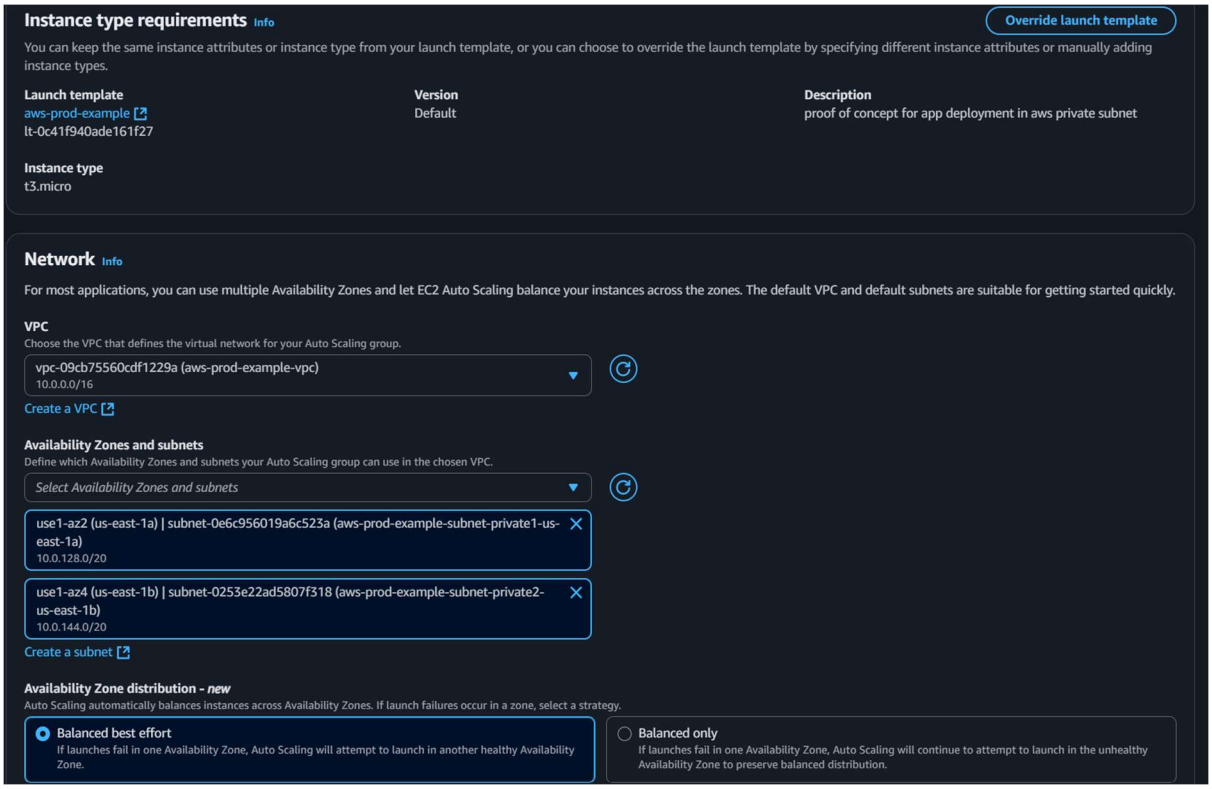

Network settings Info
Subnet Info
Don't include in launch template
Create new subnet
When you specify a subnet, a network interface is automatically added to your template.
Availability Zone Info
Don't include in launch template
Enable additional zones
Not applicable for EC2 Auto Scaling
Firewall (security groups) Info
A security group is a set of firewall rules that control the traffic for your instance. Add rules to allow specific traffic to reach your instance.
Select existing security group
Create security group
Security group name - required
aws-prod-example
This security group will be added to all network interfaces. The name can't be edited after the security group is created. Max length is 255 characters. Valid characters: a-z, A-Z, 0-9, spaces, and ._-:/()#,@[]+=&;{}!$*
Description - required Info
allow ssh access
VPC Info
vpc-09cb75560cdf1229a (aws-prod-example-vpc)
10.0.0.0/16



AFTER WE HAVE LAUNCHED THE TEMPLATE WIHOUT ADDING ANY EBS VOLUMES IT WILL BE CREATED THEN GO TO THE AUTOSCALING GROUP IN THE EC2 INSTANCE DASHBOARD AND REFRESH THE PAGE SO THAT WE CAN SEE THE TEMPLATE THEAT WAS LAUNCHED RECENTLY WHCICH IS “aws-prod-example”.

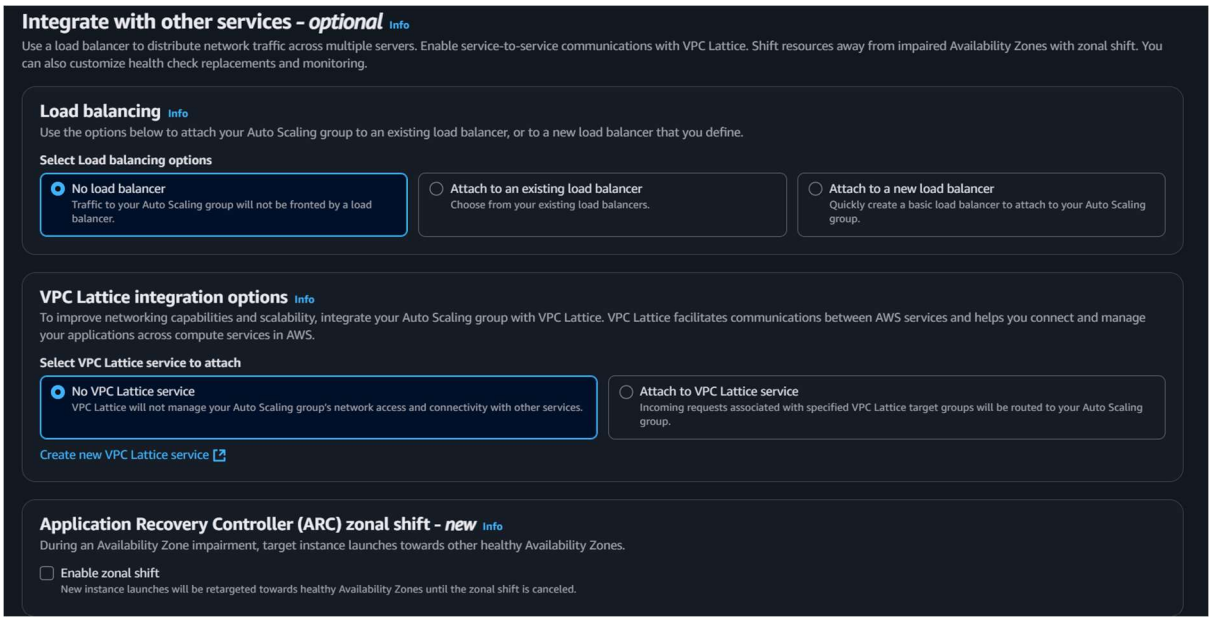


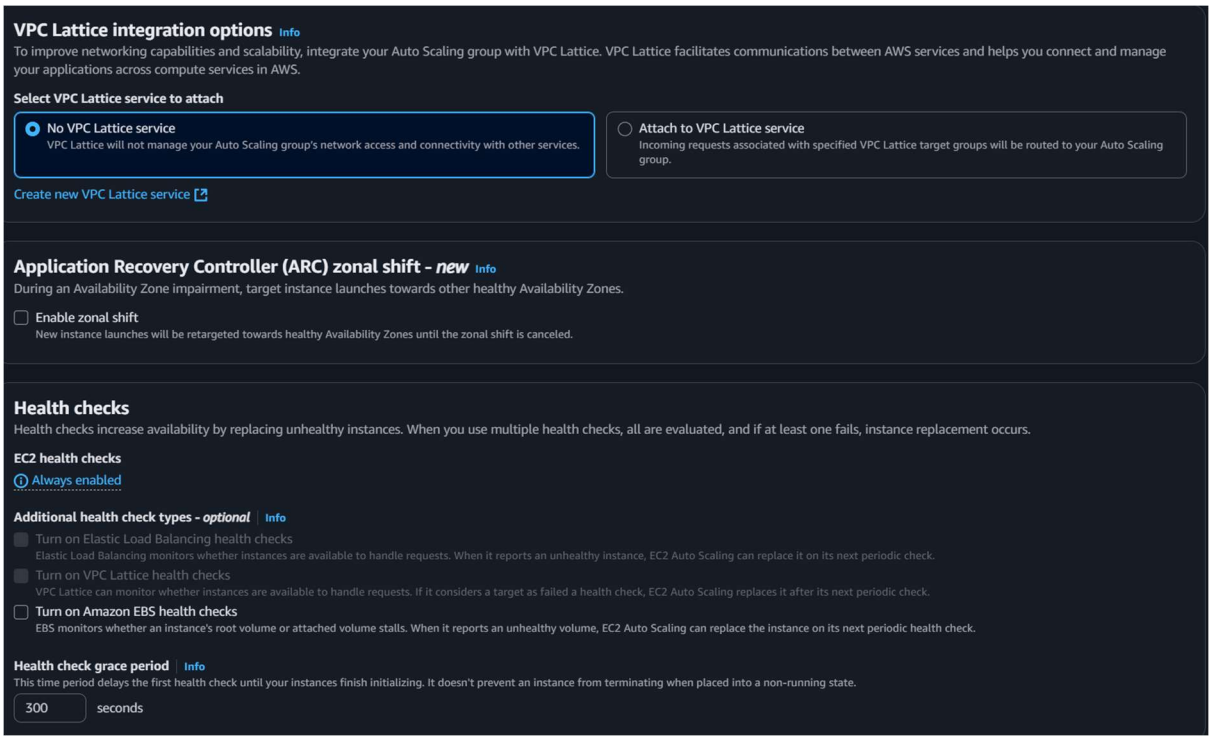
INSTANCE TYPE REQUIREMENTS:-



BOTH INSTANCES SHOULD BE RUNNING IN MY PRIVATE SUBNETS AS IN THE ARCHTECTURE.

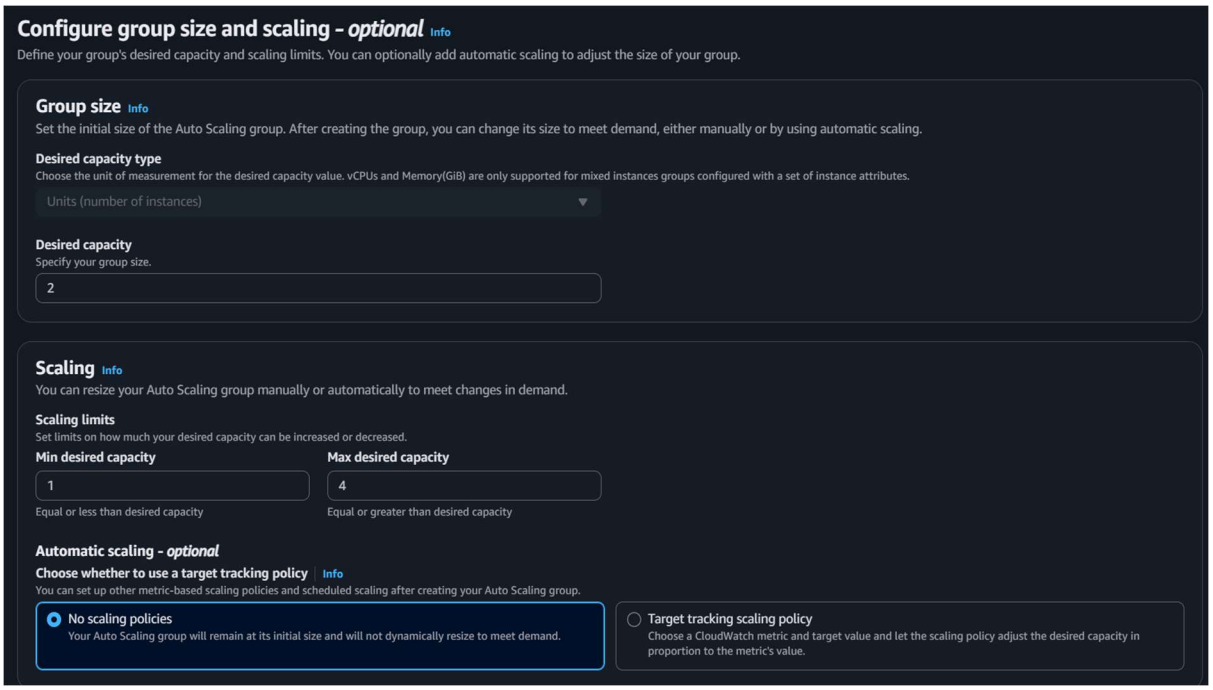
IN THE BELOW PICTURE U MIGHT SEE ME NOT CLICKING ANY LOAD BALNCER BECAUSE I DON’T WANT ANY LOAD BALANCER IN MY ARCHITECTURE FOR PRIVATE SUBNET INSTEAD I WILL CREATE AN APPLICATION LOAD BALNCER FOR MY PUBLIC SUBNET.



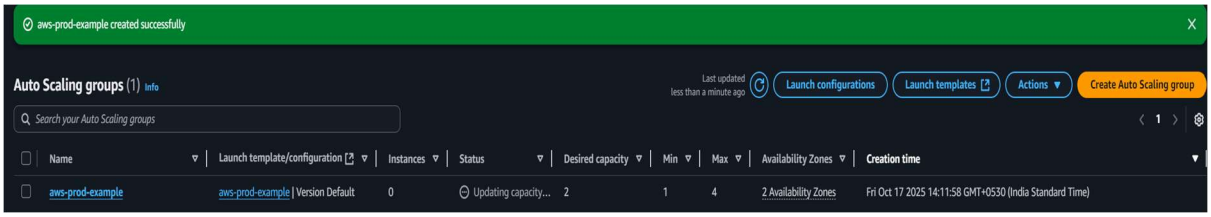


Taking default options and moving on health checks is 300 seconds.

MAXIMUM DESIRED CAPACITY TAKING AS 4 INSTANCES AND MINMUM DESIRED CAPACITY IS 1 INSTANCE AND THE DESIRED CAPACITY IS 2 INSTANCES

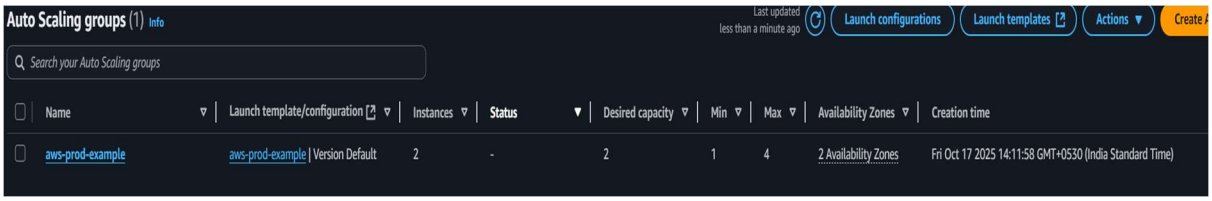


IAM NOT ADDING ANY NOTIFICATIONS LIKE SNS FOR AUTO SCALING GROUPS AND CLICKING ON NEXT FOR LAUNCHING THE AUTOSCALING GROUPS.

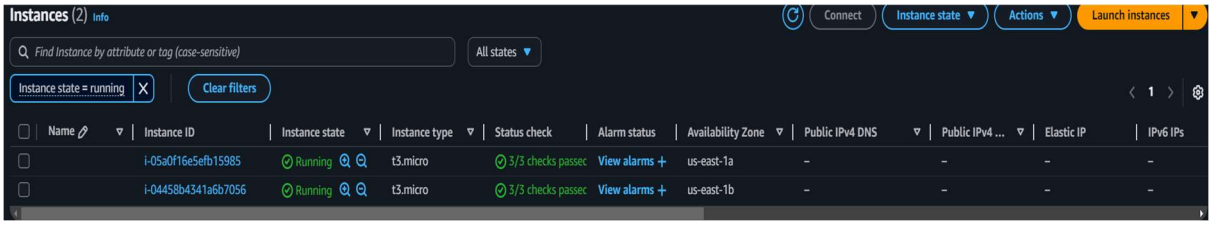


MY AUTOSCALING GROUP IS LAUNCHED SUCCESSFULLY.

MY UPDATING CAPACITY STATE in autoscaling group has disappeared and 2 instances are launched in EC2.GO TO EC2 AND VERIFY.

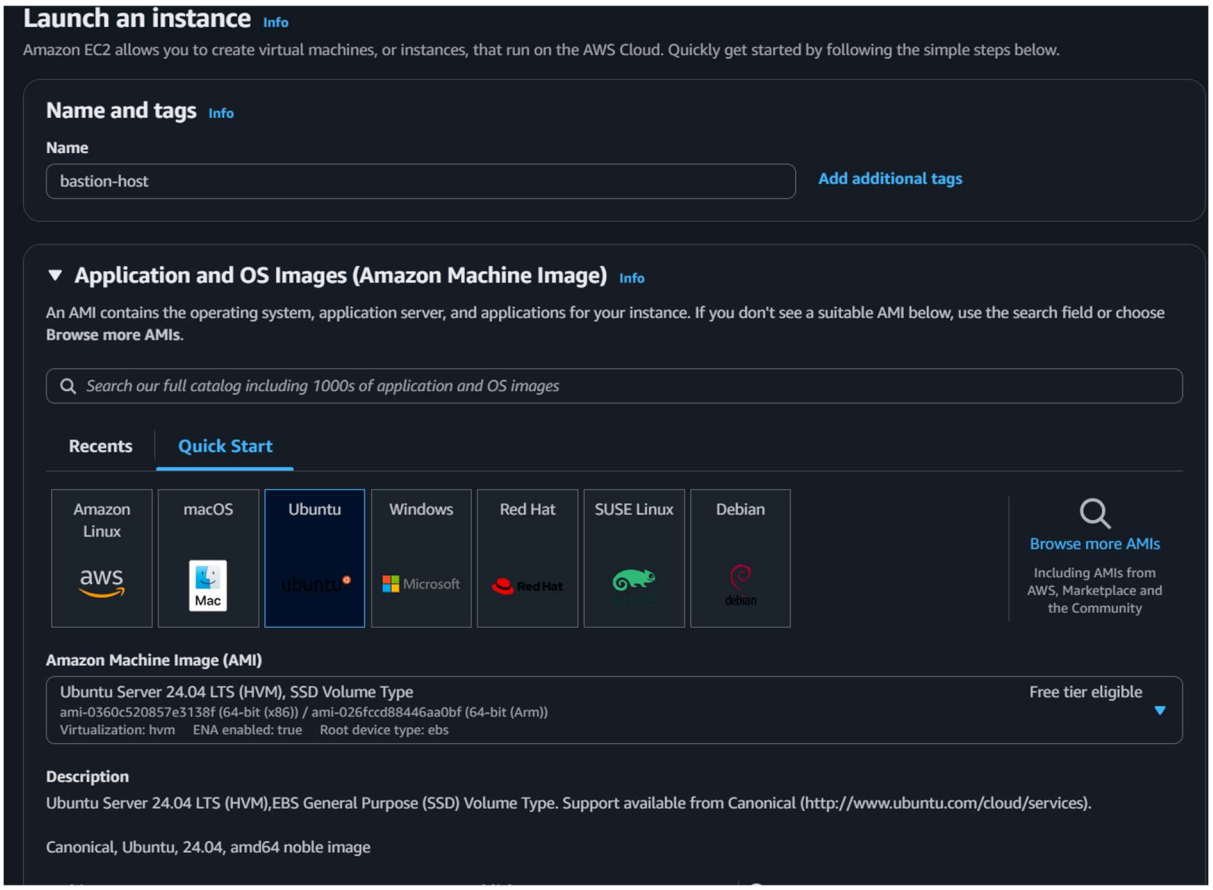


IN EC2:- 2 INSTANCES ARE LAUNCHED.

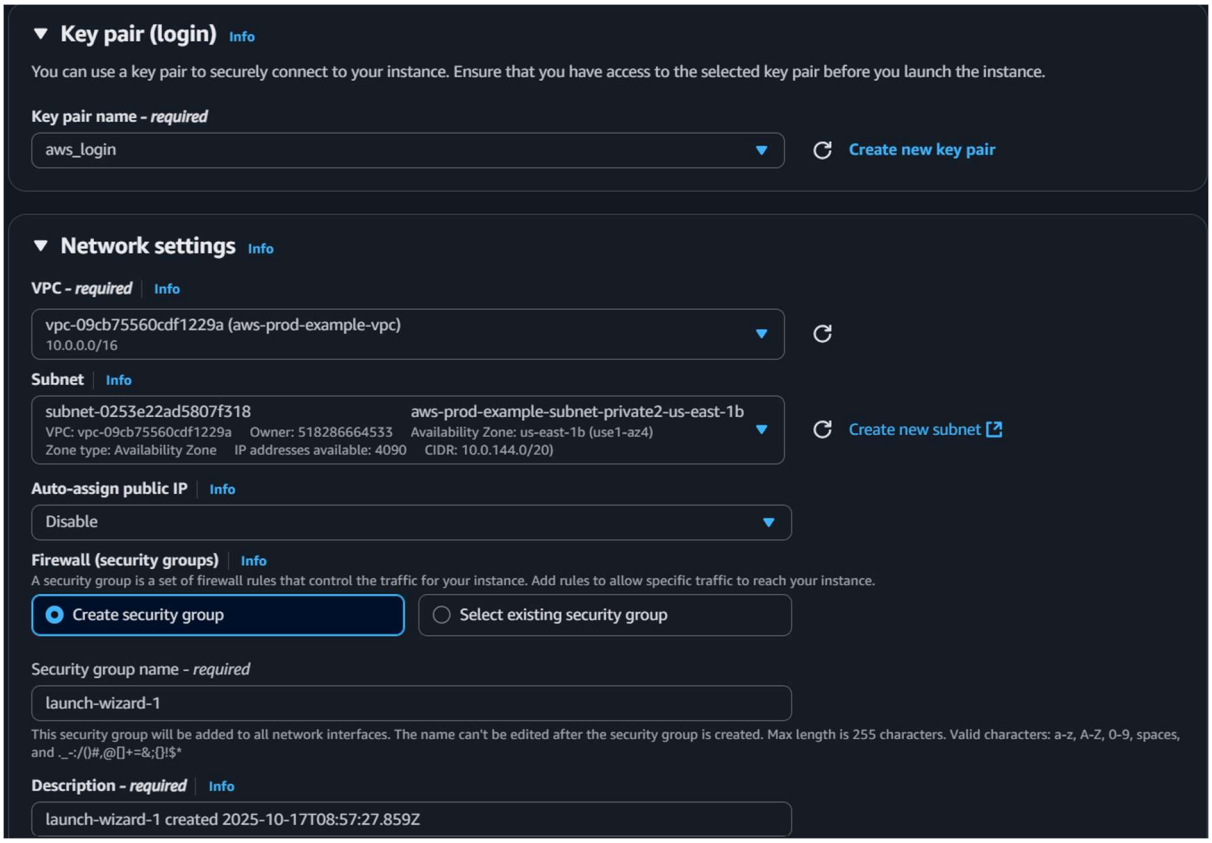


AFTER WE HAVE LAUNCHED THE INSTANCES IN THE PRIVATEE SUBNET WE HAVE TO INSTALL THE APPLICATIONS FIRST IN THE PRIVATE SUBNET BY ENTERING INTO THE EC2 INSTANCES WHICH ARE PRIVATE WHICH MEANS THEY DON'T HAVE AN PUBLIC IPV4 ADDRESS INSTEADTHEY HAVE PRIVATE IPV4 ADDRESS SO WE HAVE TO USE "BASTON HOST" TO ENTER INTO THES INSTANCES WHICH ACTS AS A MEDIATOR BETWEEN THE PUBLIC AND PRIVATE SUBNETS.

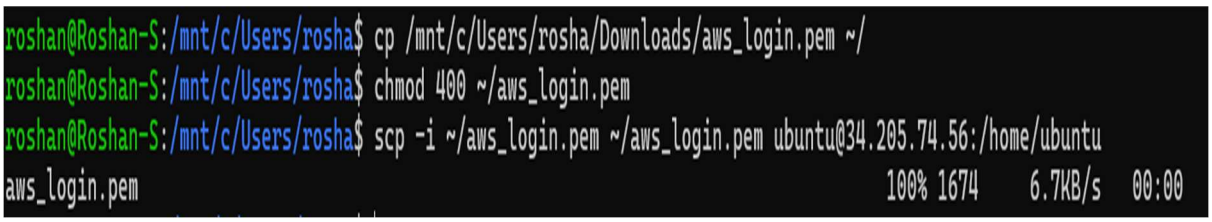
CREATING BASTION HOST BY LAUNCHING AN INSTANCE



EDIT THE KEY PAIR AND NETWORK SETTINGS THE NETWORK CONFIGURATION SHOULD CHANGE TO aws-prod-vpc that is the vpc which we created otherwise bastion-host will not be able to access the ec2 instances which are present in the private subnet.



Entering into the bastion-host in wsl.



Iam doing ssh to my bastion-host to enter the bastion instance pin my public subnet to access the ec2 instance running in my private subnet.

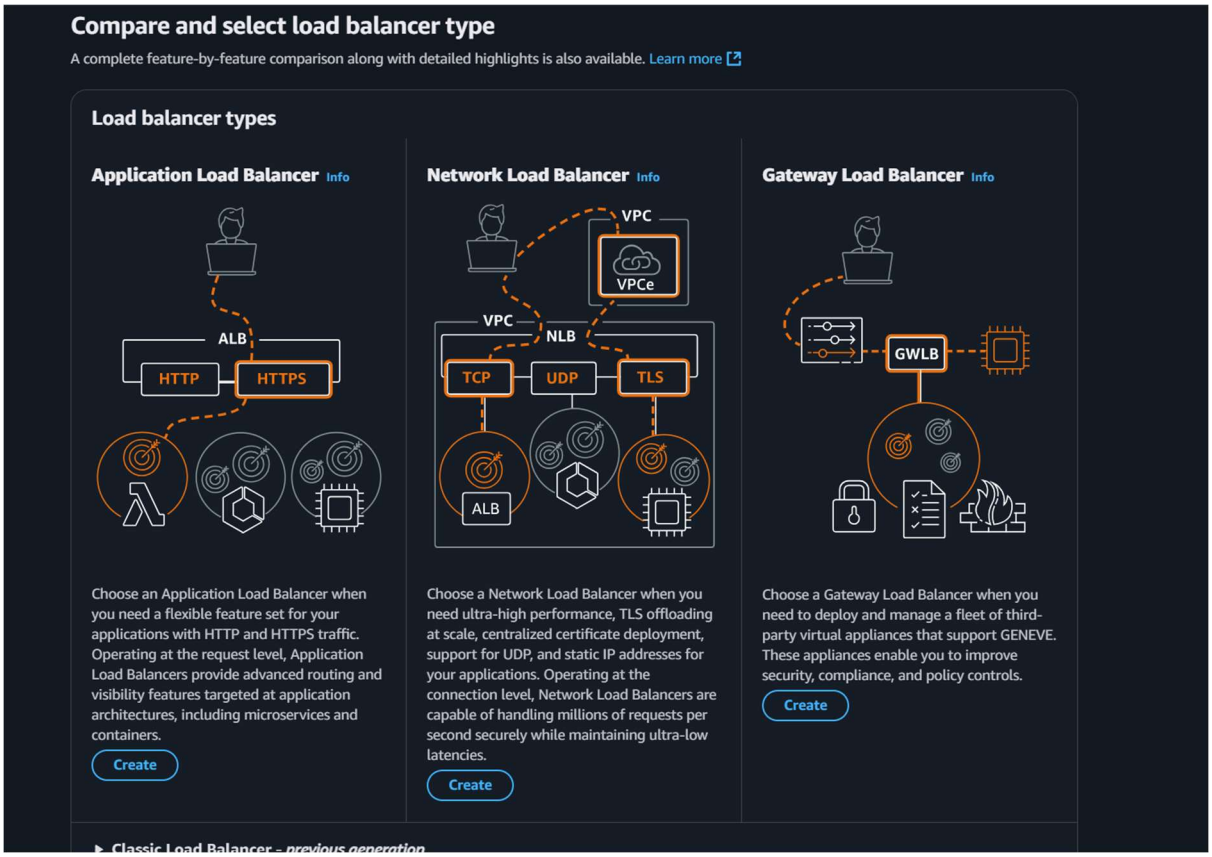
```
roshan@Roshan-S:/mnt/c/Users/rosha$ ssh -i ~/aws_login.pem ubuntu@34.205.74.56
Welcome to Ubuntu 24.04.3 LTS (GNU/Linux 6.14.0-1011-aws x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/pro

 System information as of Fri Oct 17 09:32:10 UTC 2025

  System load:  0.0               Temperature:           -273.1 C
  Usage of /:   26.0% of 6.71GB   Processes:             111
  Memory usage: 22%               Users logged in:       0
  Swap usage:   0%                IPv4 address for ens5: 10.0.0.115


Expanded Security Maintenance for Applications is not enabled.

0 updates can be applied immediately.

Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status
```

AS WE CAN SEE BY DOING THIS I HAVE SUCCESSFULLY GOT INTO MY PRIVATE SUBNET IN WSL . BY DOING SSH TO MY PRIVATE SUBNET IN USE-EAST-1A.

```
ubuntu@ip-10-0-0-115:~$ ls
aws_login.pem
ubuntu@ip-10-0-0-115:~$ ssh -i aws_login.pem ubuntu@10.0.133.242
The authenticity of host '10.0.133.242 (10.0.133.242)' can't be established.
ED25519 key fingerprint is SHA256:CPYKSvzPAX6oX9CNtkrOYjB26imp0/sipYOl+zoNDAA.
This key is not known by any other names.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added '10.0.133.242' (ED25519) to the list of known hosts.
Welcome to Ubuntu 24.04.3 LTS (GNU/Linux 6.14.0-1011-aws x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/pro

 System information as of Fri Oct 17 09:36:10 UTC 2025

  System load:  0.0               Temperature:           -273.1 C
  Usage of /:   25.8% of 6.71GB   Processes:             110
  Memory usage: 25%               Users logged in:       0
  Swap usage:   0%                IPv4 address for ens5: 10.0.133.242

Expanded Security Maintenance for Applications is not enabled.

0 updates can be applied immediately.

Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status


The list of available updates is more than a week old.
```

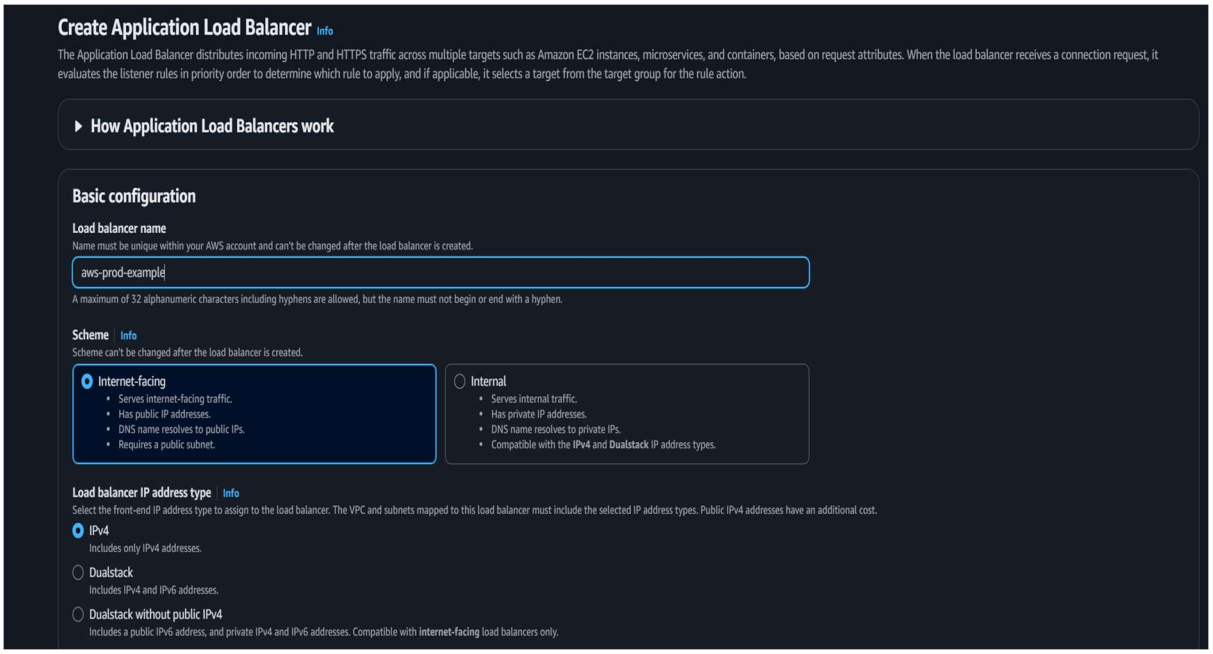
In the private subnet we are creaing a simple index.html file. And running the server in 8000 port which was developed in the aws management console.

CREATE LOAD BALANCERS FOR SENDING THE TRAFFIC TO THESE EC2 INSTANCES CREATED IN THE PRIVATE SUBNETS THE VERIFIFVATIO IS THAT IN USE-EAST-1A THE PYTHON3 IS INSTALLED SO IT HITS AND GIVES A MESSAGE WITHOUT ANY ERROR BUT WHEN WE ARE TRYING TO ACCESS THE SECOND PRIVATE SUBNET EC2 INSTANCE IT WILL GIVE AN ERROR MESSAGE AS PYTHON IS NOT INSTALLED

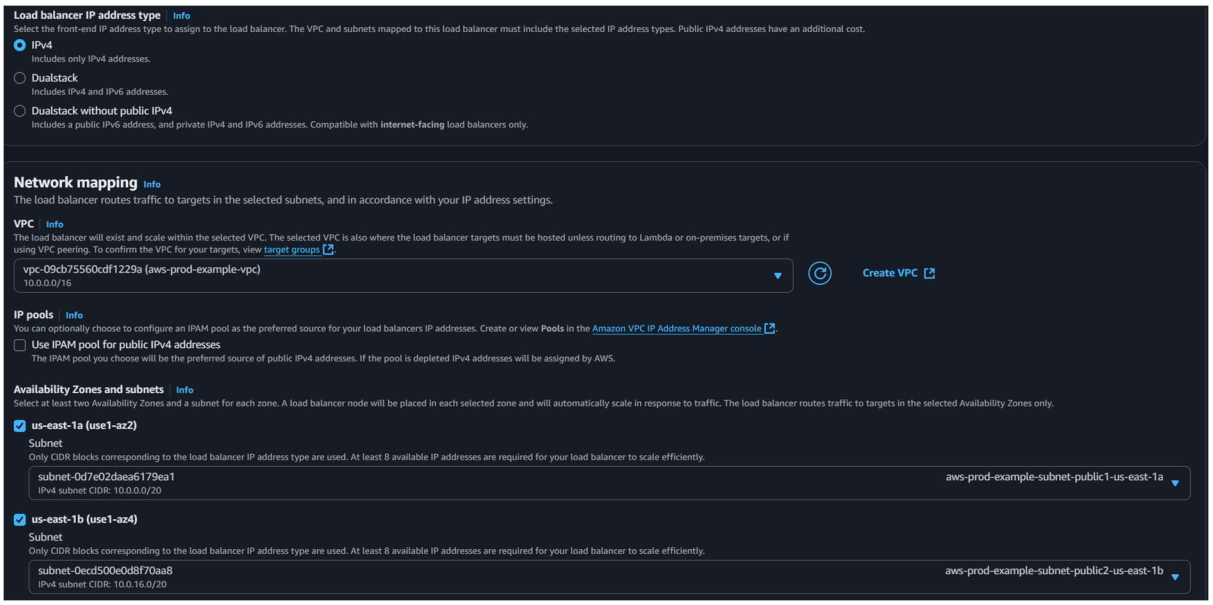


DIFFERENT TYPES OF LOAD BALANCERS WE CAN CHOOSE.

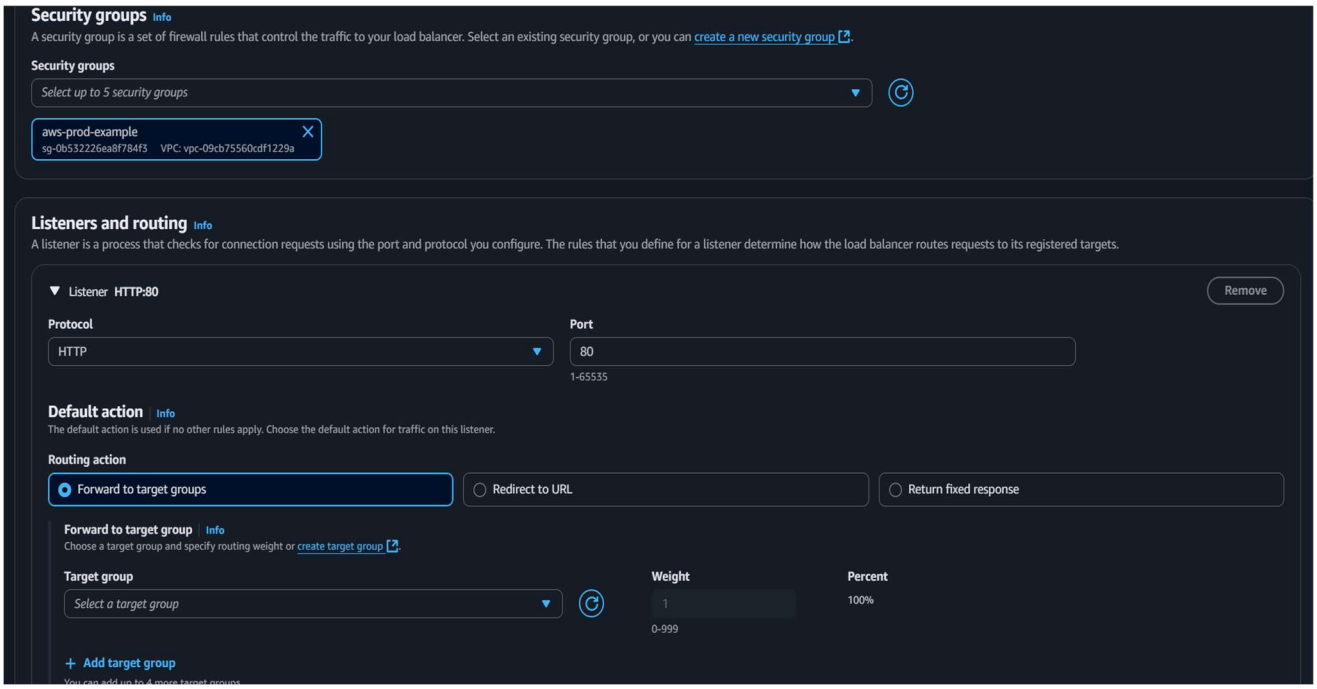
LOAD BALNCER SHOULD BE INTERNET FACING BECAUSE WE ARE TRYING TO ACCESS VIA THE PUBLIC SUBNET.



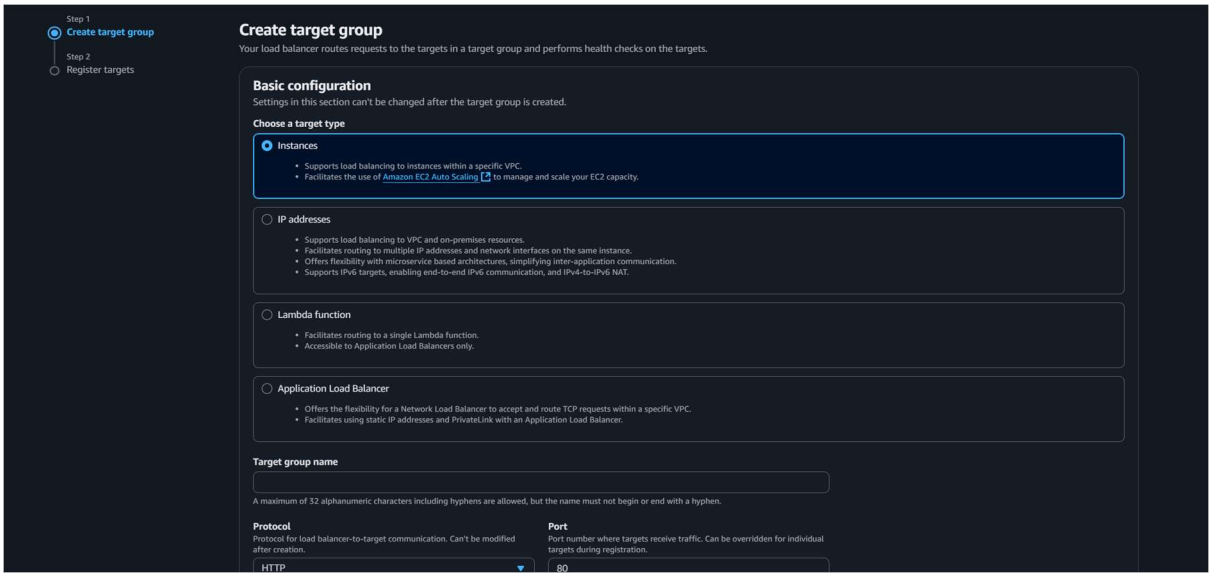
NETWORK MAPPING IS FOR THE PUBLIC SUBNETS NOT PRIVATE SUBNETS FOR APPLICATION LOAD BALNCER COOHSE ONLY PUBLIC SUBNETS.



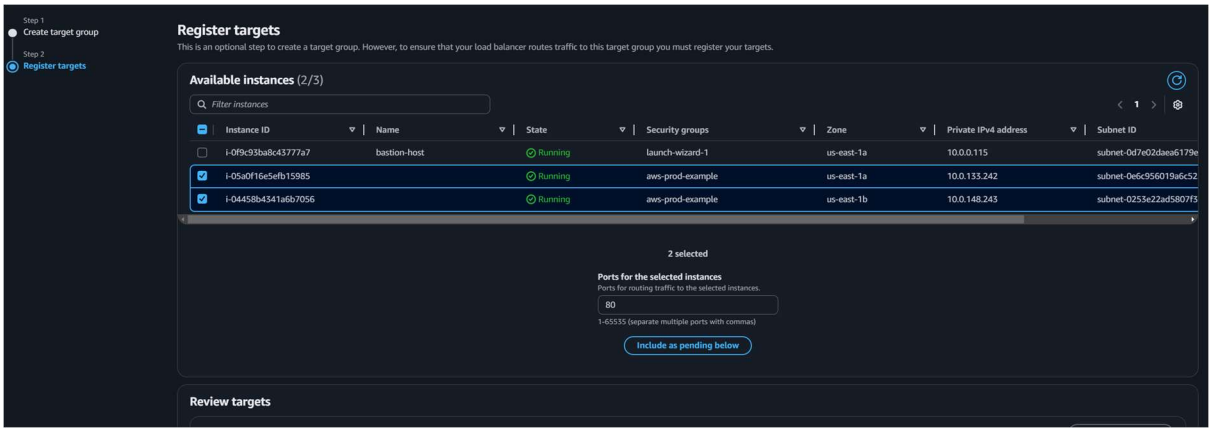
ADD SECURITY GROUP aws-prod-example and make a target group by adding target groups mention target group as ec2 instances.



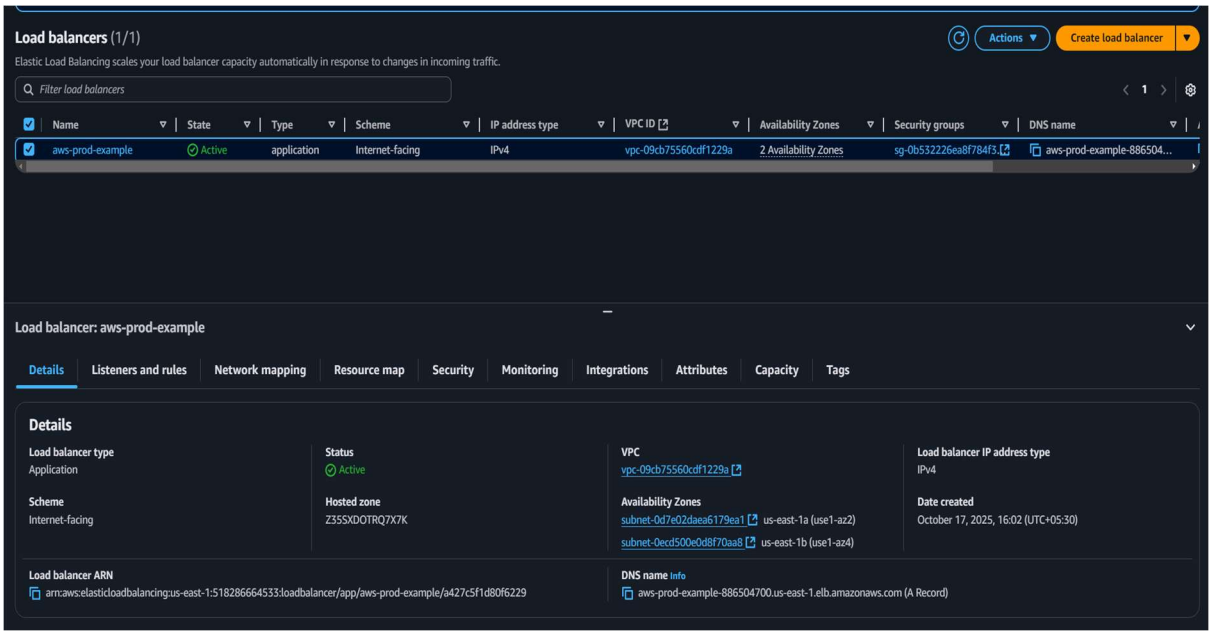
CREATING TARGET GROUPS:-



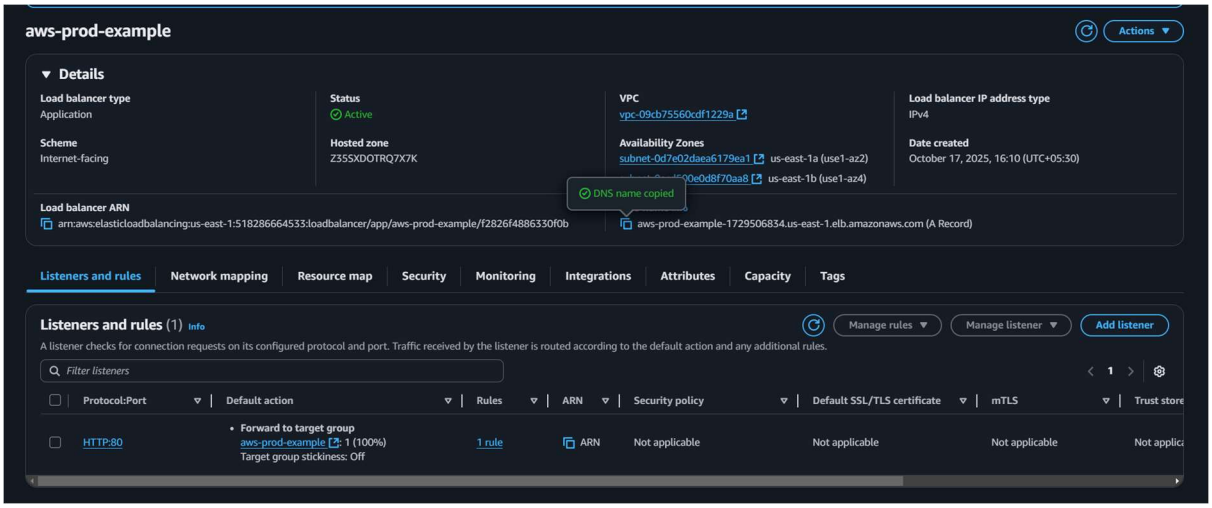
WHILE REGISTERING THE TARGETS DON’T SELECT BASTION-HOST SELECT THE OTHER TWO INSTANCES.



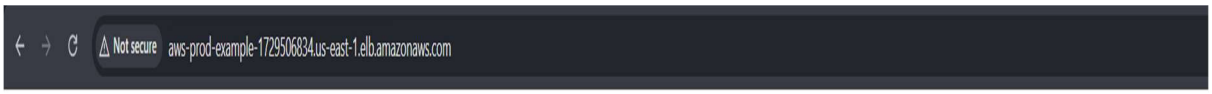
THE LOAD BALANCER is in the active state.



BY using the DNS NAME AT THE APPLICATION LOAD BALANCER AND COPY PASTING IT IN A NEW TAB I WILL BE ABLE TO CREATE MY APPLICATION IN A PRIVATE SUBNET



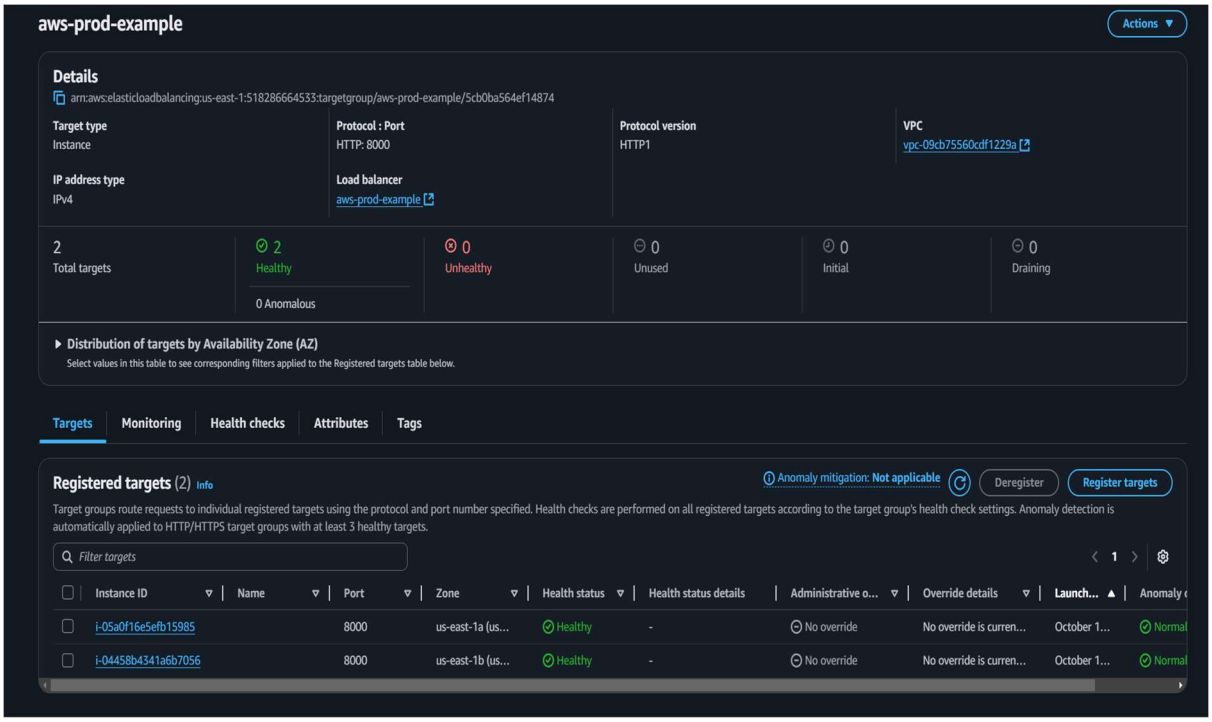
OUTPUT FOR First Private subnet (us-east-1a)



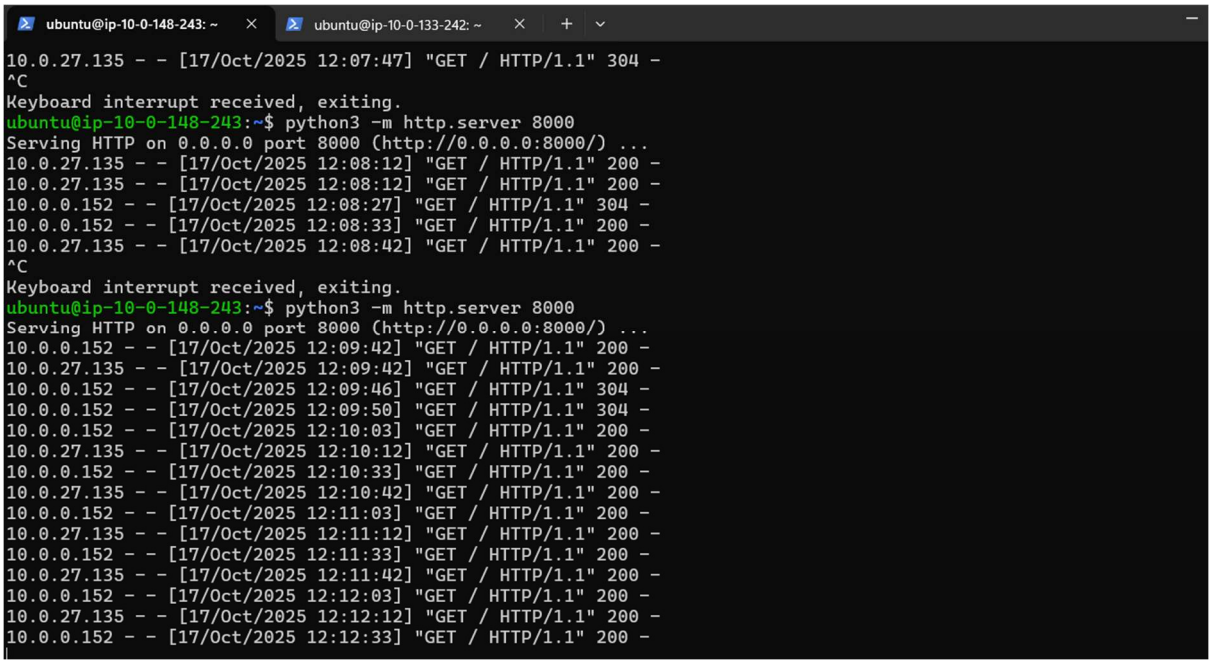
MY first AWS PROJECT to demonstrate apps in private subnet

This is a paragraph.

I have also made both the applications healthy and saw how one was healthy and unhealthy using the following DNS NAME of load balancer and I was able to understand how the load balancer functions



The below picture represents the way I was accessing the ec2 instance in my private subnets and checking if there were healthy or not in the load balncer.so what we understand is even if one server is down the other server is able to run in a different availability zone for backups.



OUTPUT FOR second private subnet (US-EAST-1B)

While refreshing the pages we will be able to see different outputs and health checks can be maintained by configuring the health checks in the application load balancer.

NEW TERMINOLOGY:- “BASTION-HOST”

A **Bastion Host** is a special-purpose server used to securely access servers in a private network from an external network, usually the internet. It acts as a “jump box” and is heavily secured since it’s exposed to potential attacks. Here are the common **applications and use cases** of a bastion host:

---

1. Secure Remote Access

- Allows system administrators to access internal servers (like database servers or application servers) in private subnets.
- Example: SSH or RDP access to EC2 instances in AWS private VPCs.

---

2. Jump Server / Jump Box

- Serves as the only entry point to a network segment.
- All traffic to sensitive servers goes through the bastion host for monitoring and control.

---

3. Network Segmentation Enforcement

- Helps enforce security policies by isolating public-facing and internal servers.
- Reduces the attack surface since only the bastion host is exposed to the internet.

---

4. Auditing and Logging

- Tracks all connections and actions for compliance purposes.
- Example: Logging all SSH sessions from admins for audit purposes.

---

5. VPN Alternative

- Sometimes used as a lightweight alternative to full VPN for secure remote access.

---

6. Temporary Administrative Access

- Can provide temporary access to developers or support staff without opening direct access to internal servers.

---

7. Multi-Cloud or Hybrid Cloud Access

- Acts as a secure entry point for managing servers across multiple cloud providers or on-premises environments.

---

8. Security Layer for Automation

- Scripts or automation tools can route through the bastion host for tasks like updates, backups, or monitoring internal servers securely.

Example in AWS:

- An **EC2 instance** in a public subnet acts as a bastion host.
- Admins SSH into this instance, then use it to SSH into EC2 instances in private subnets.
- Combined with security groups, multi-factor authentication, and CloudWatch logging for secure operations.