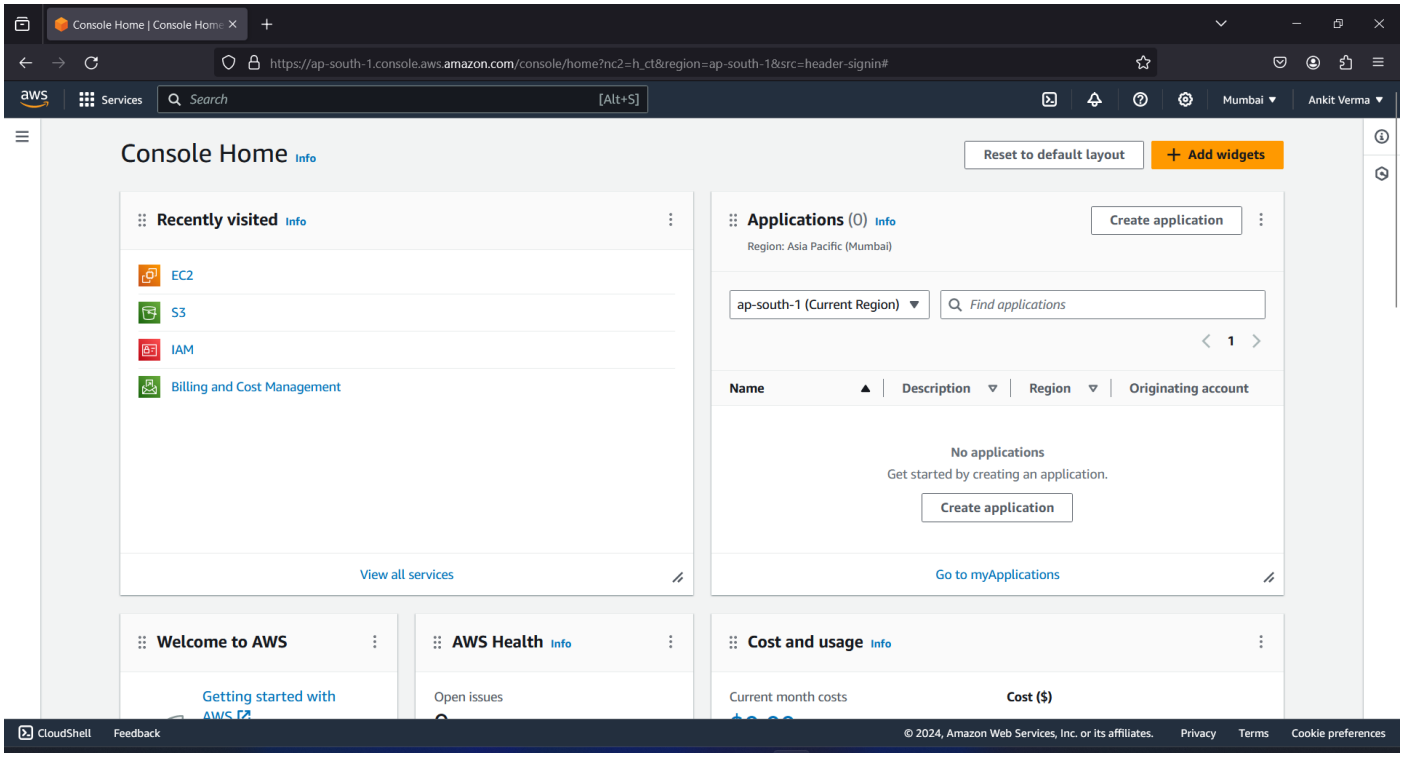
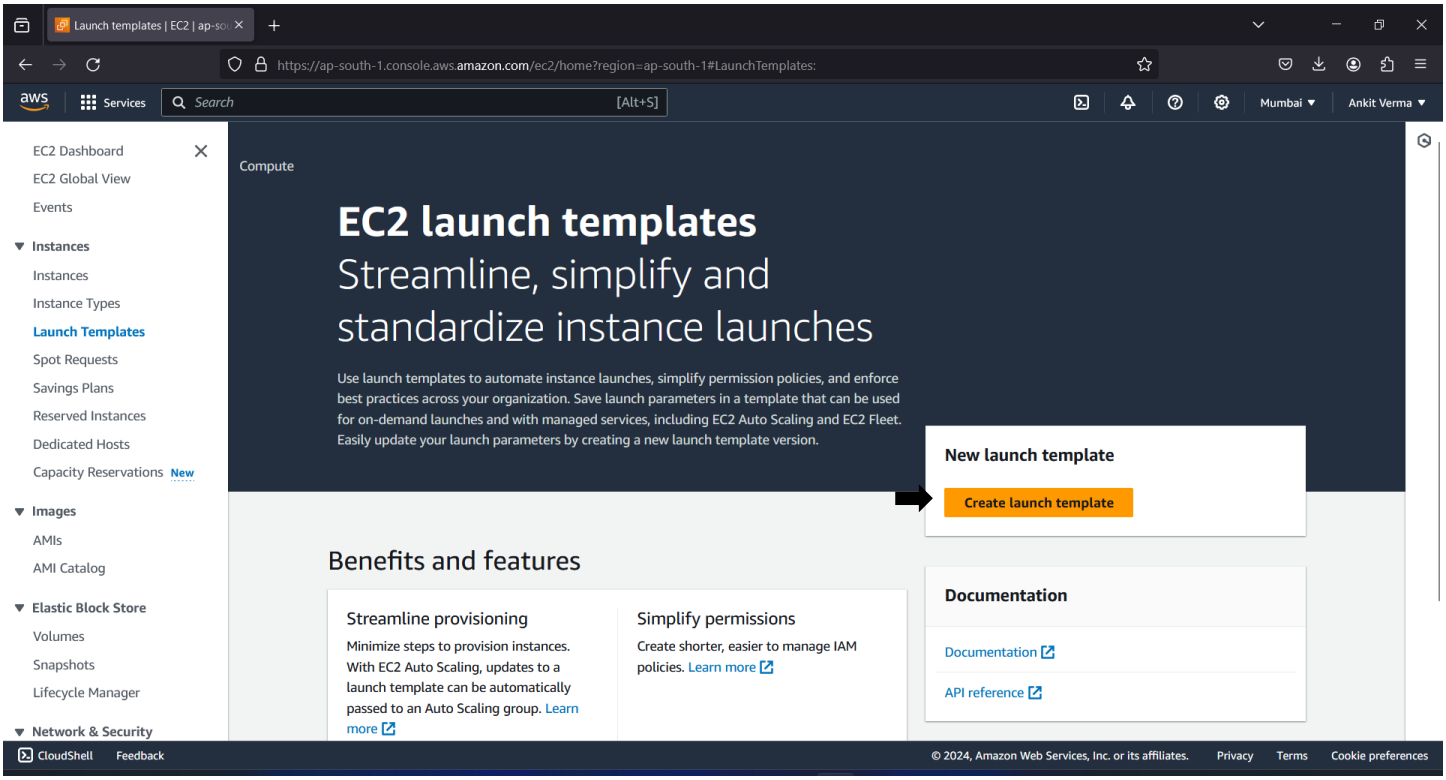
Assignment: 11

Problem Statement: Build scaling plans in AWS that balances the load on different EC2 instances.

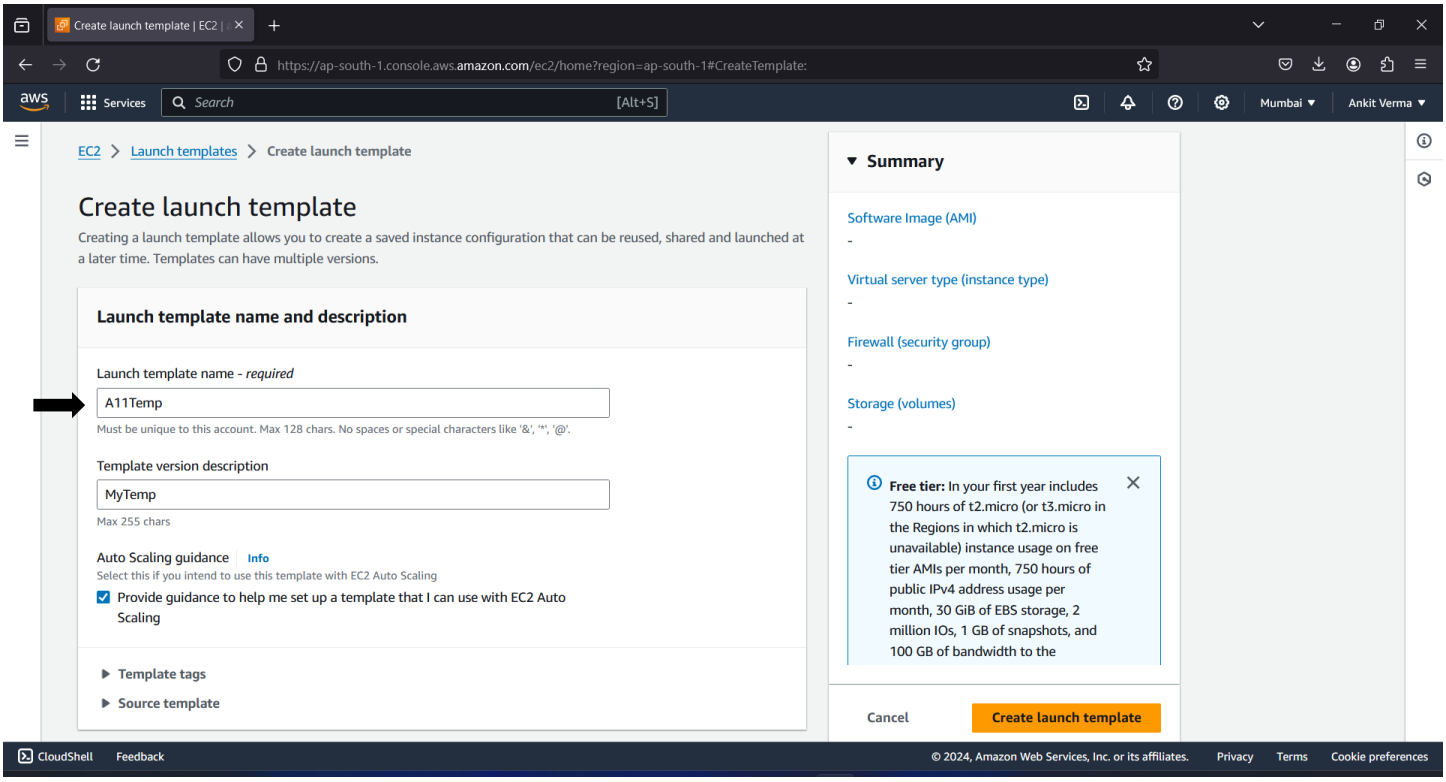
Step 1: Select EC2.



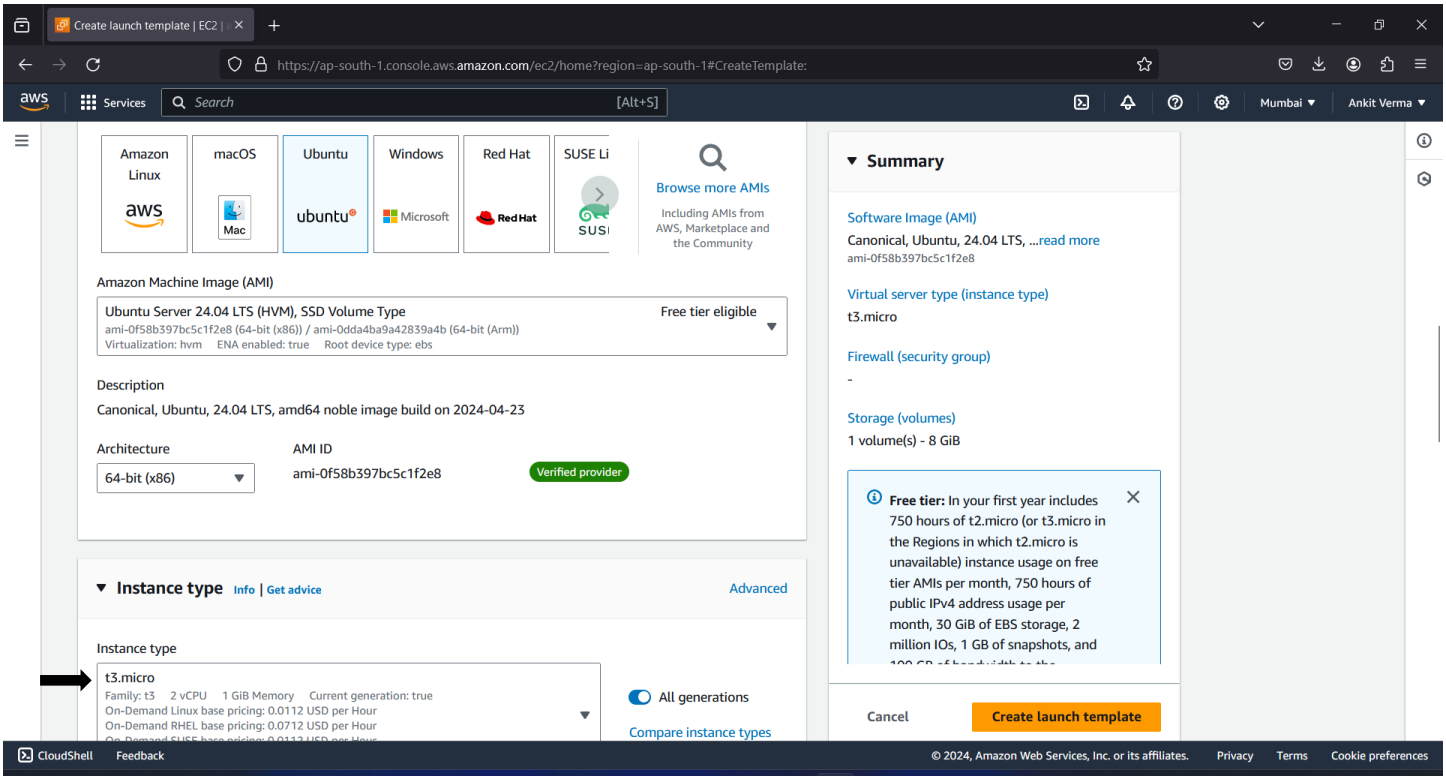
Step 2: Under the Launch Templates, click on Create Launch Template



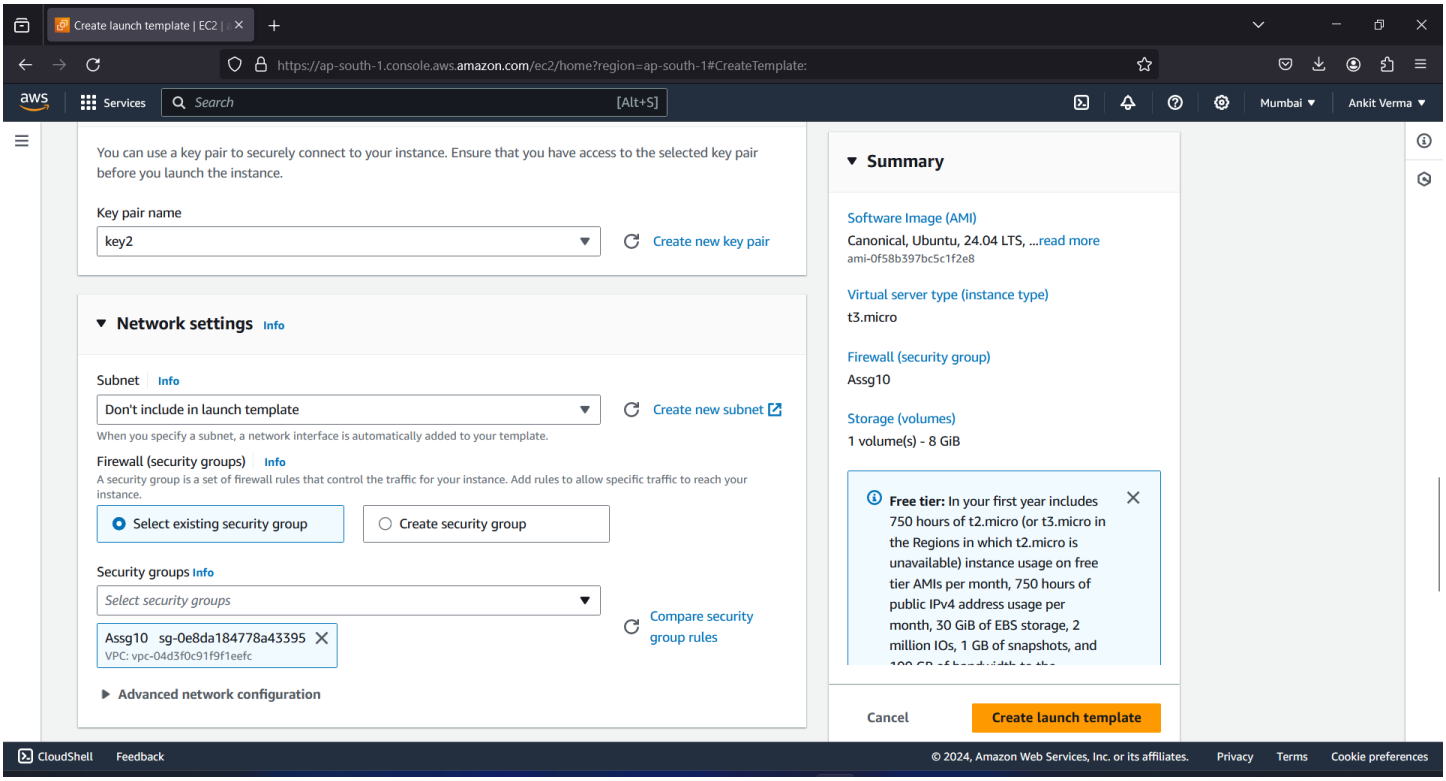
Step 3: Give a name and description to the template. Check the Auto Scaling Guidance checkbox.



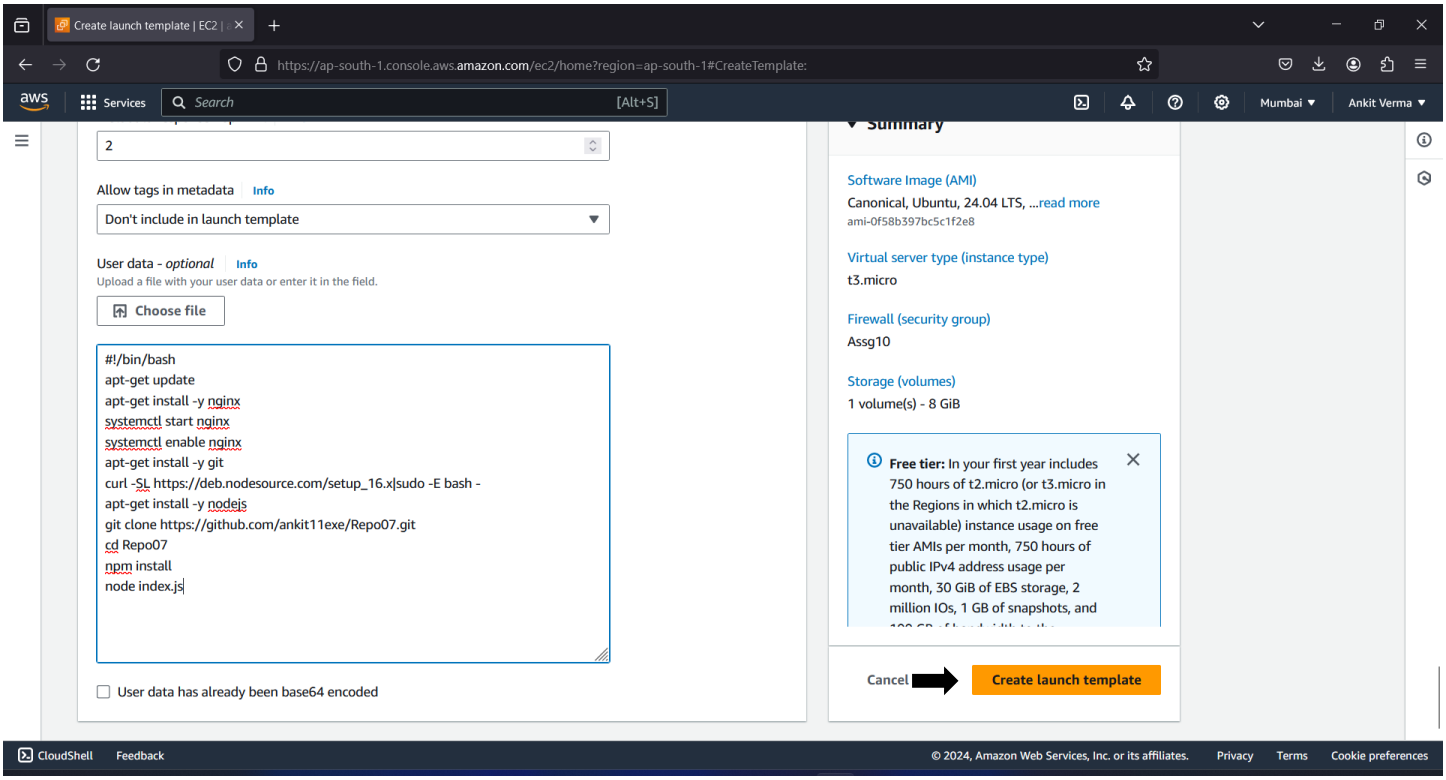
Step 4: Select Ubuntu & under instance type, select t3.micro .



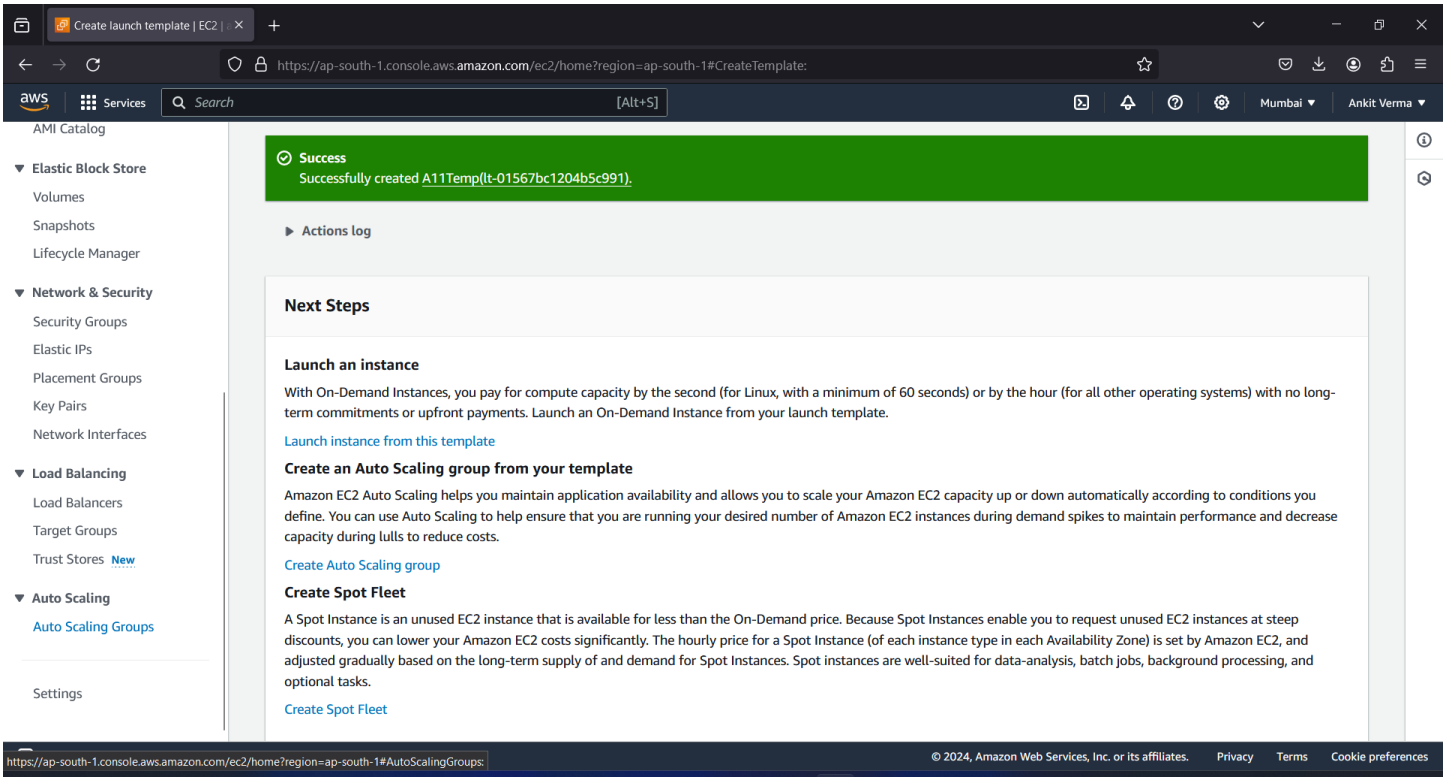
Step 5: Under key pair, select an existing key and select the user created Security group.



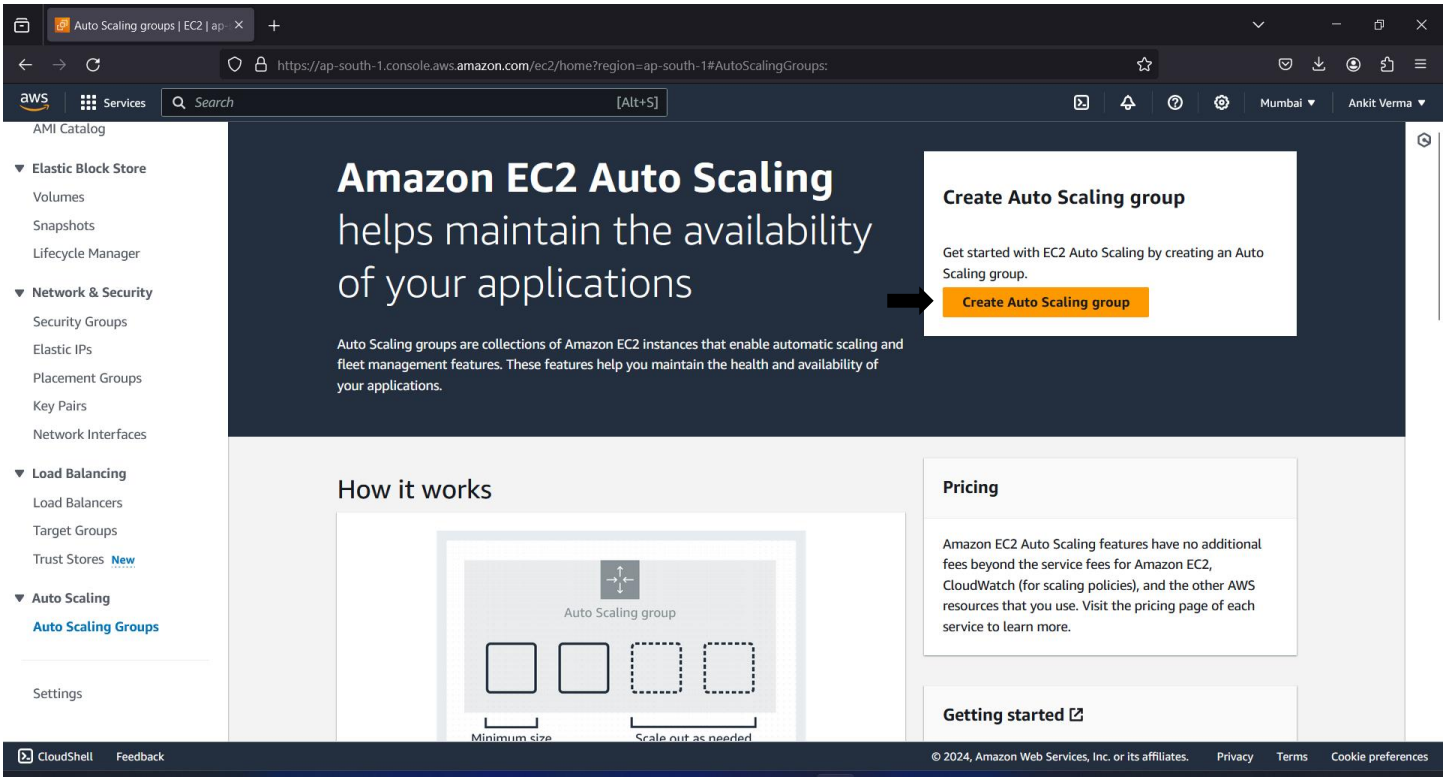
Step 6: Expand the Advanced Details tab & Scroll down to the bottom, in the bash console type the following commands, Then click on Create Launch Template.



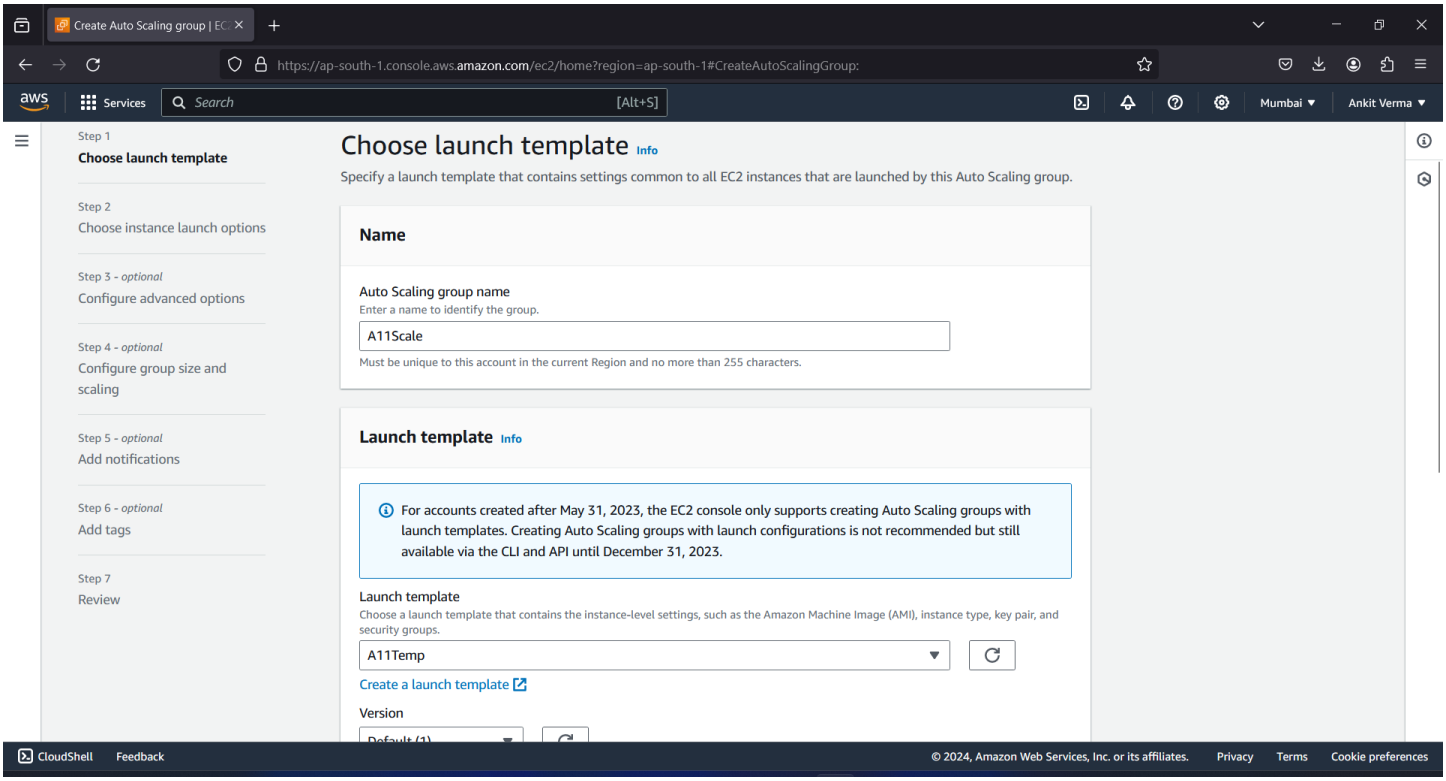
Step 7: Click on Auto Scaling Group.



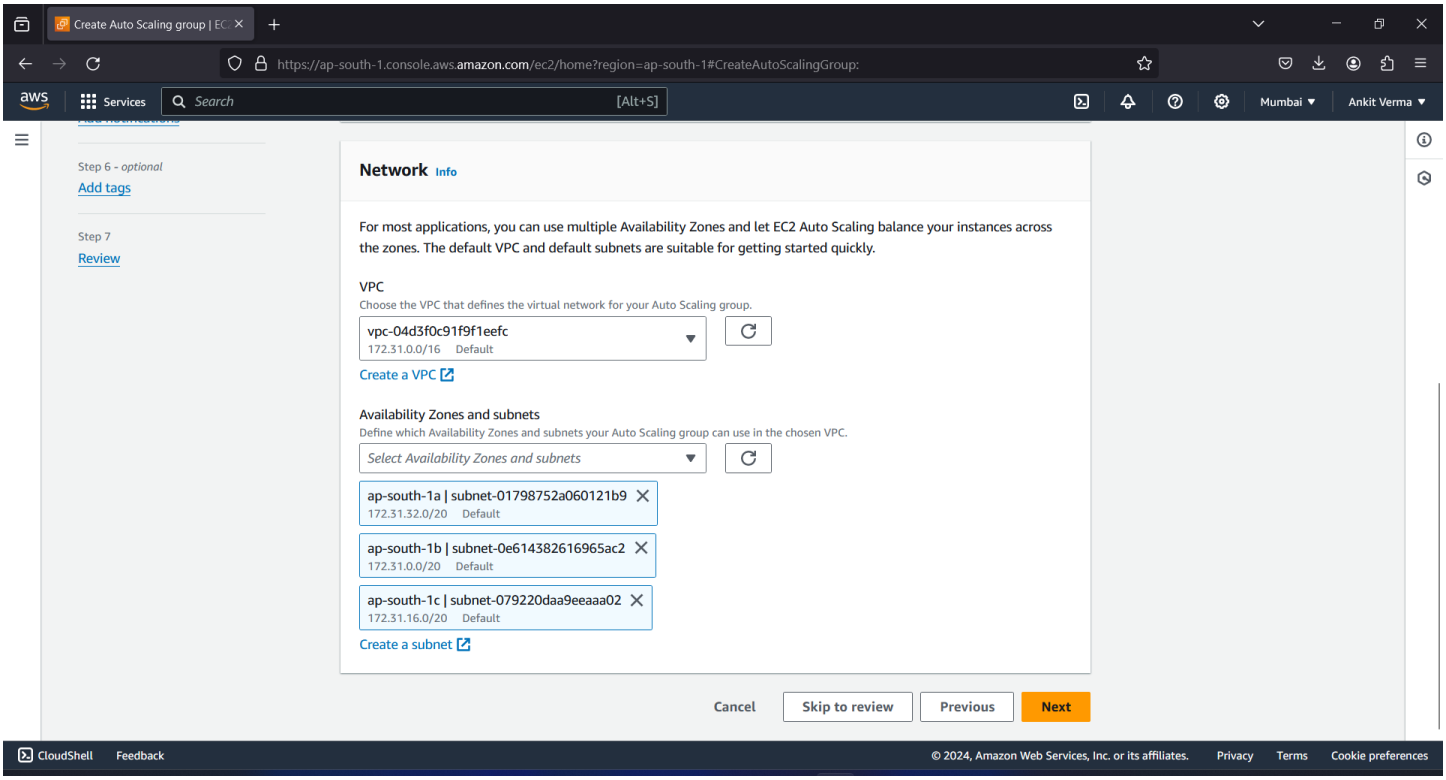
Step 8: Click on Create Auto Scaling Group.



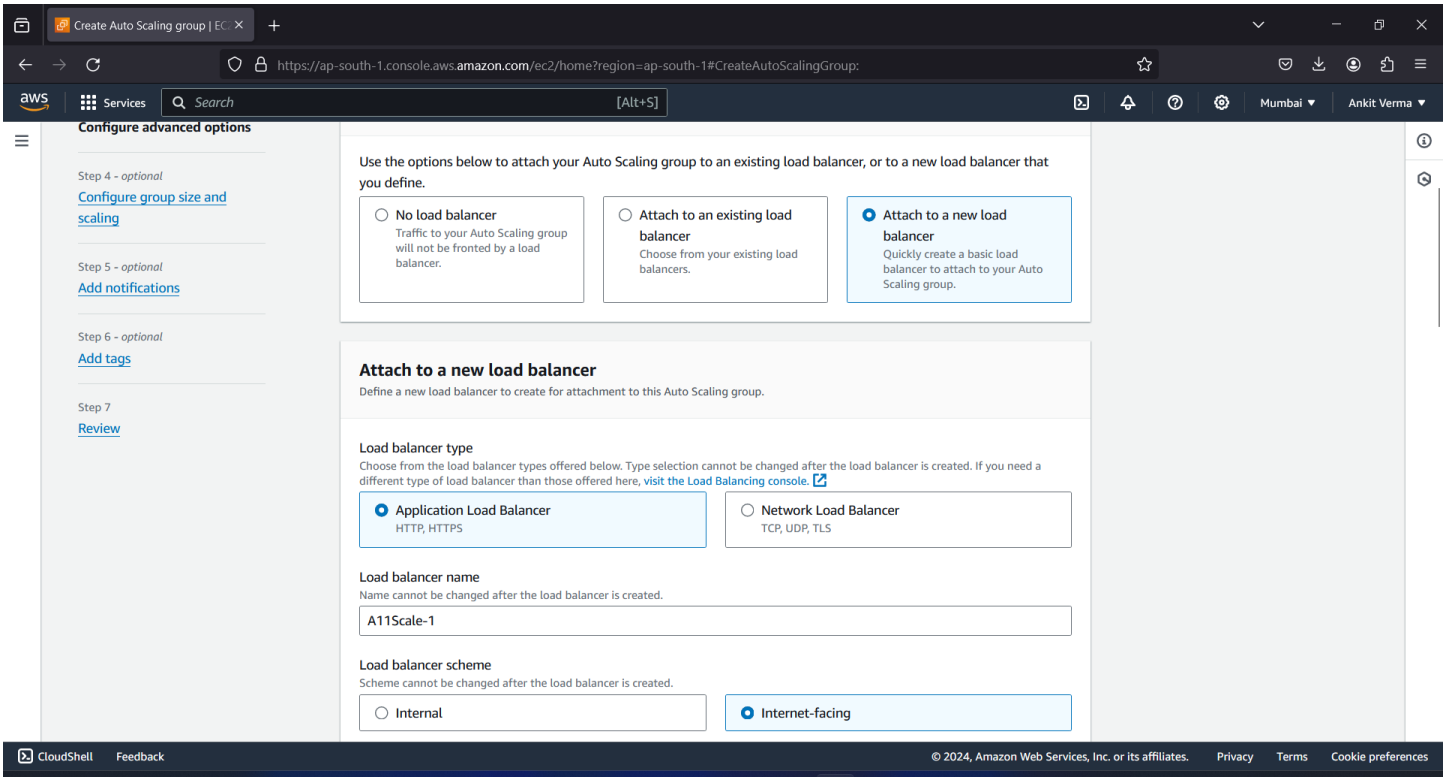
Step 9: Give a name and select the newly created Template. Then, click on Next.



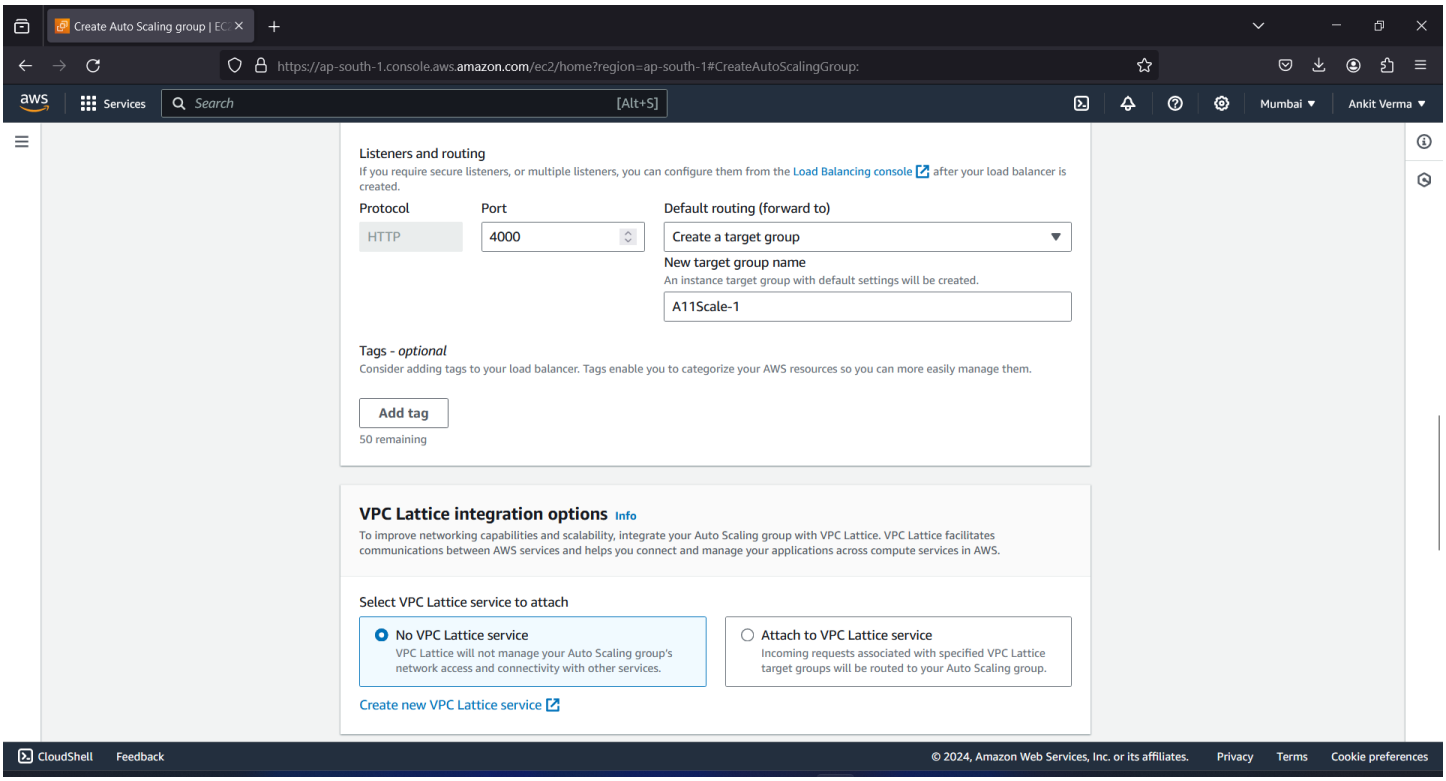
Step 10: In Network tab select all the available zones. Then click on next.



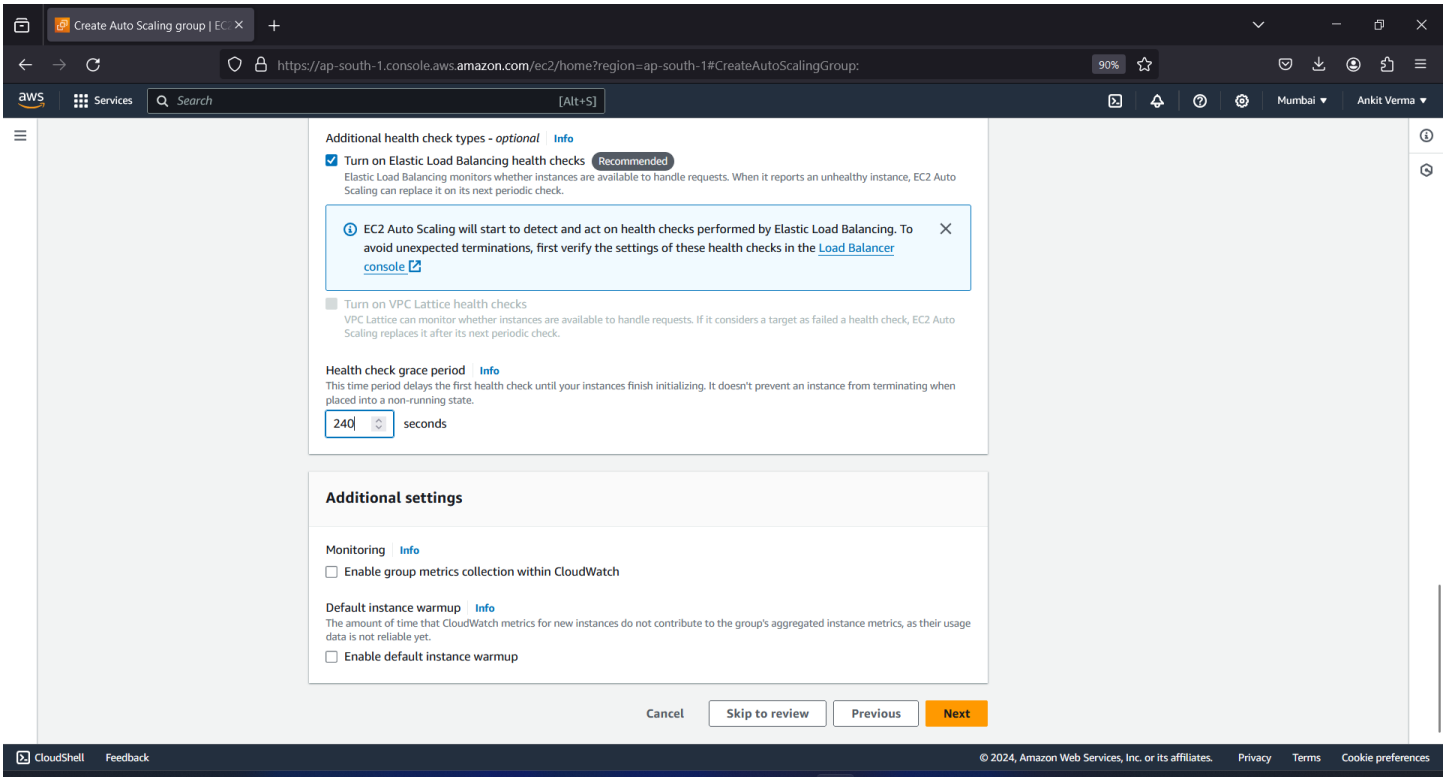
Step 11: Select Attach a new load balancer, select Application Load Balancer & select Internet Facing.



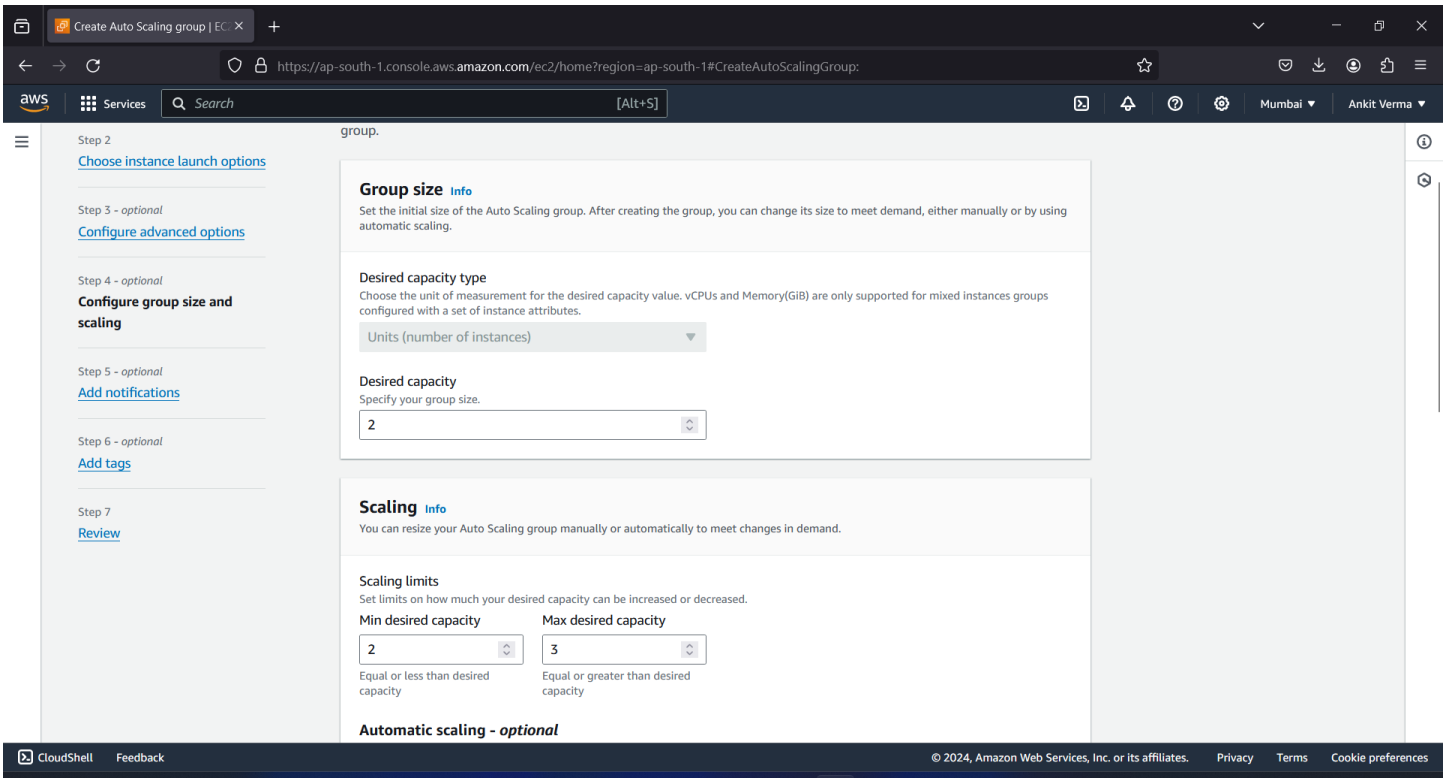
Step 12: Give the port no. 4000 & select Create a target group. Then select No VPC Lattice Service.



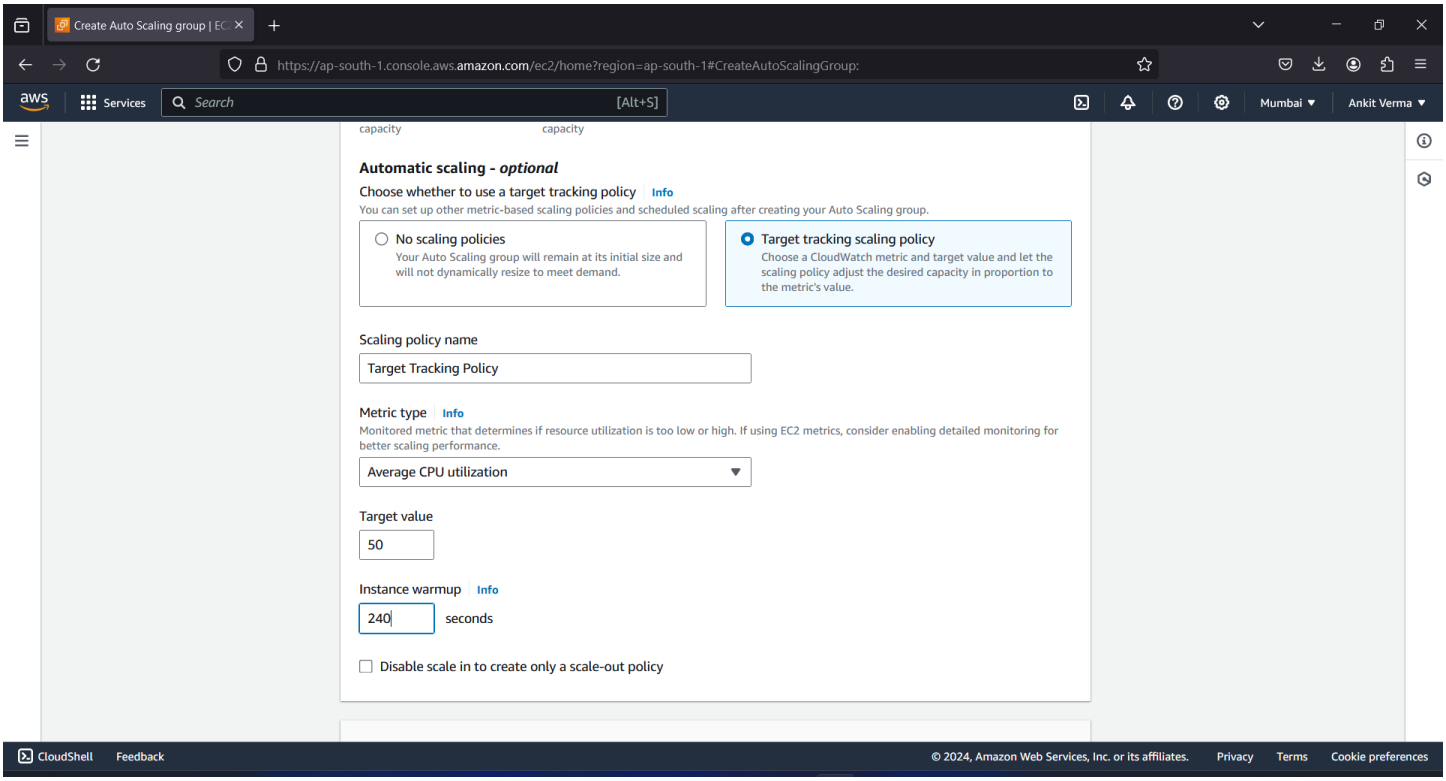
Step 13: Check the Turn on Elastic Load Balancing Health checks checkbox. Give the Health Check Grace Period of 240 seconds. Click on NEXT.



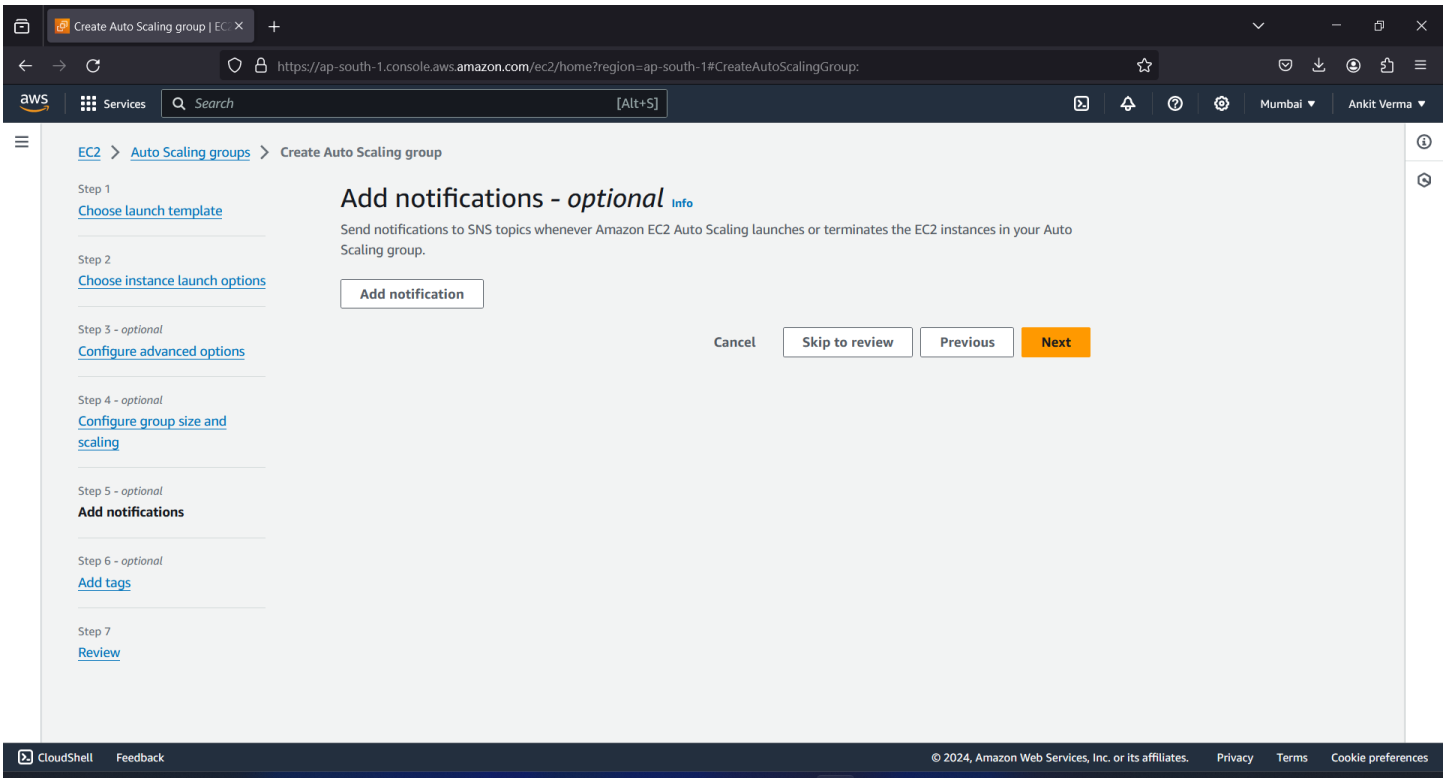
Step 14: Under Desired capacity, give a size of 2. Under Scaling, give min capacity 2 & max capacity 3.



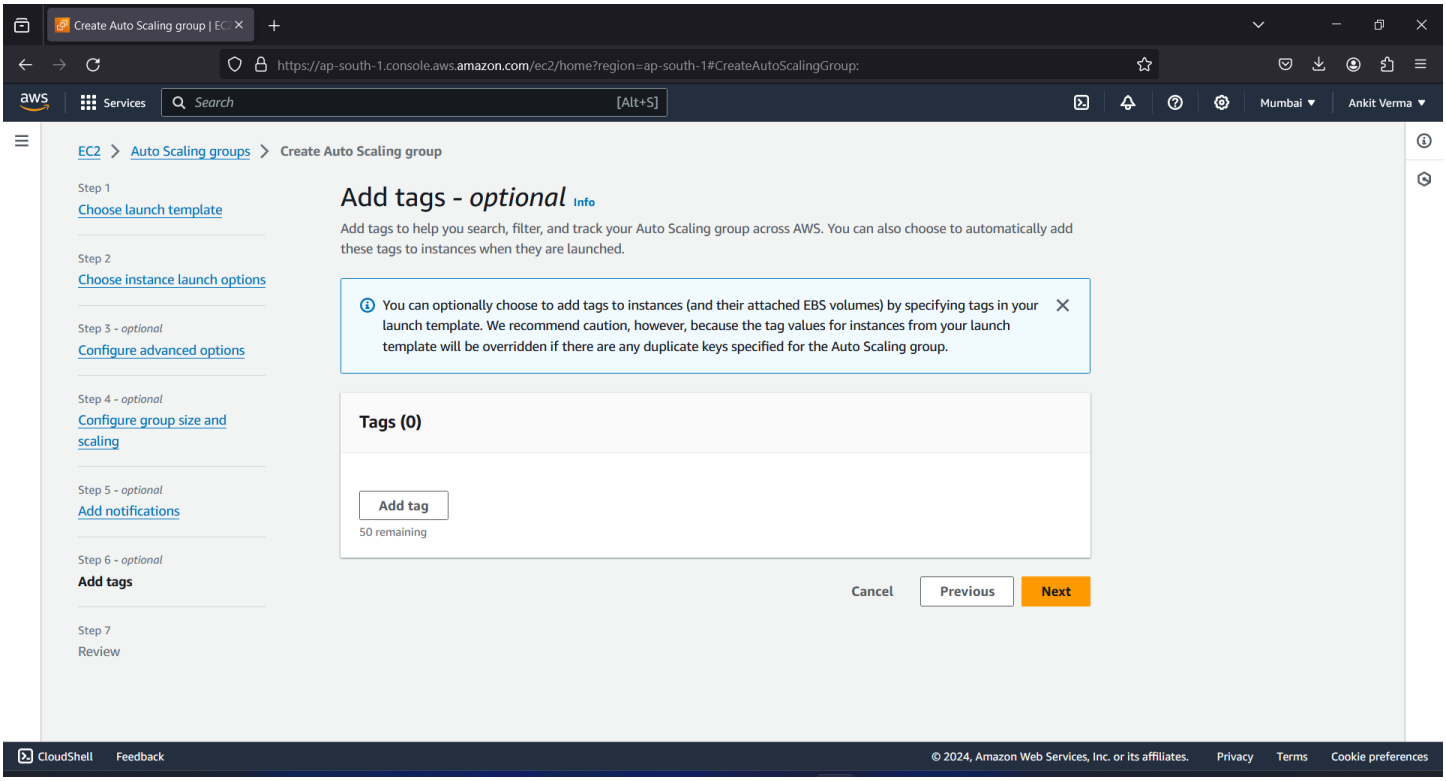
Step 15: Select Target tracking scaling policy. And give the instance warmup time of 240 seconds. Then click on ‘Next’.



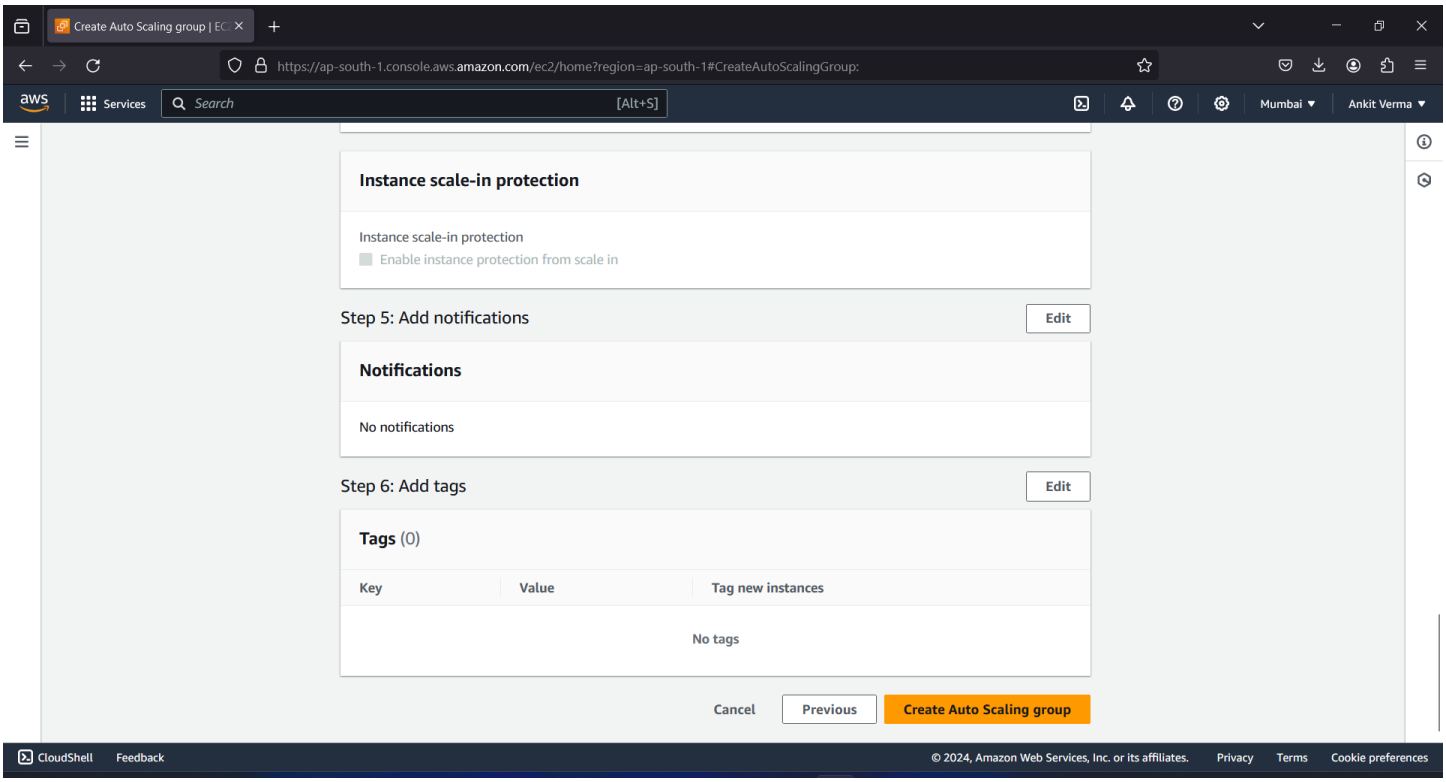
Step 16: Click next.



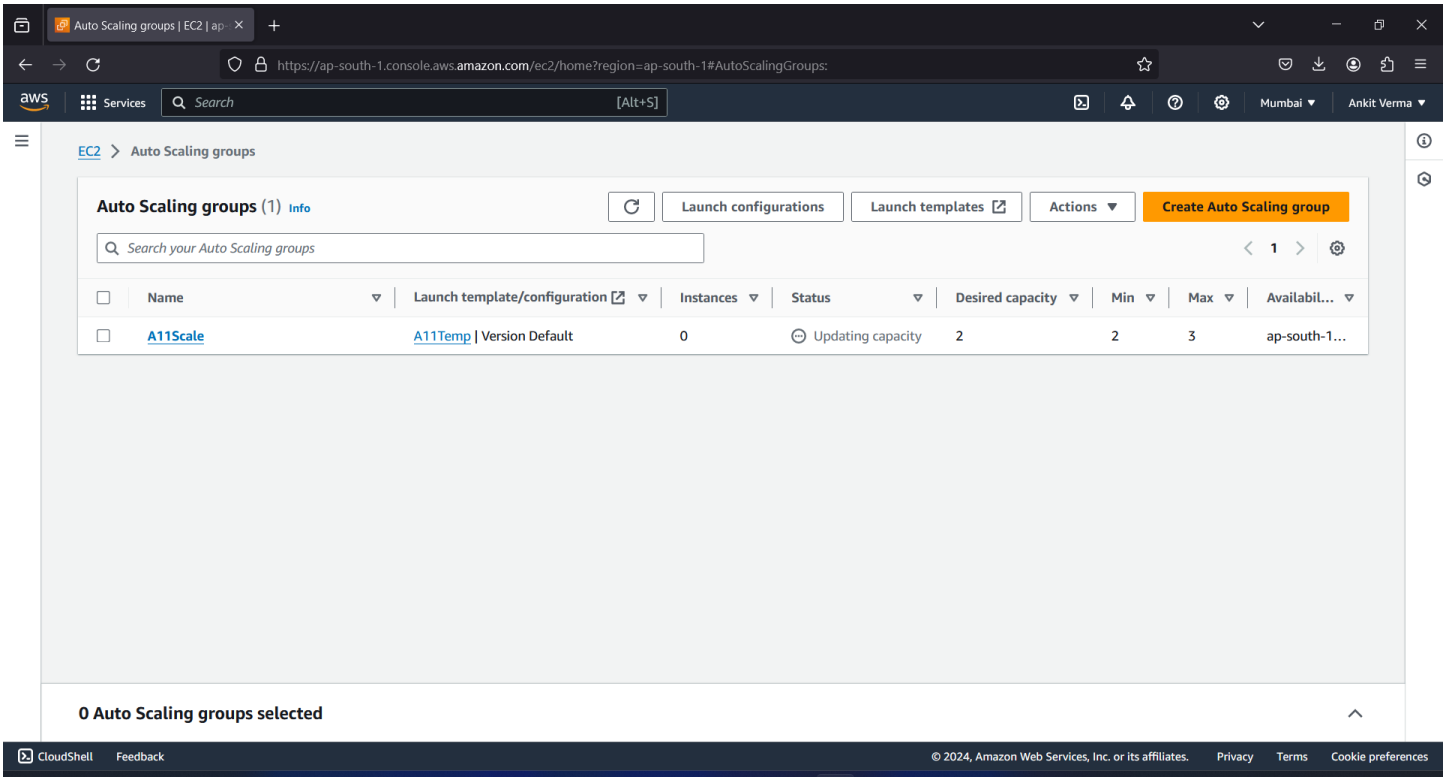
Step 17: Click on next.



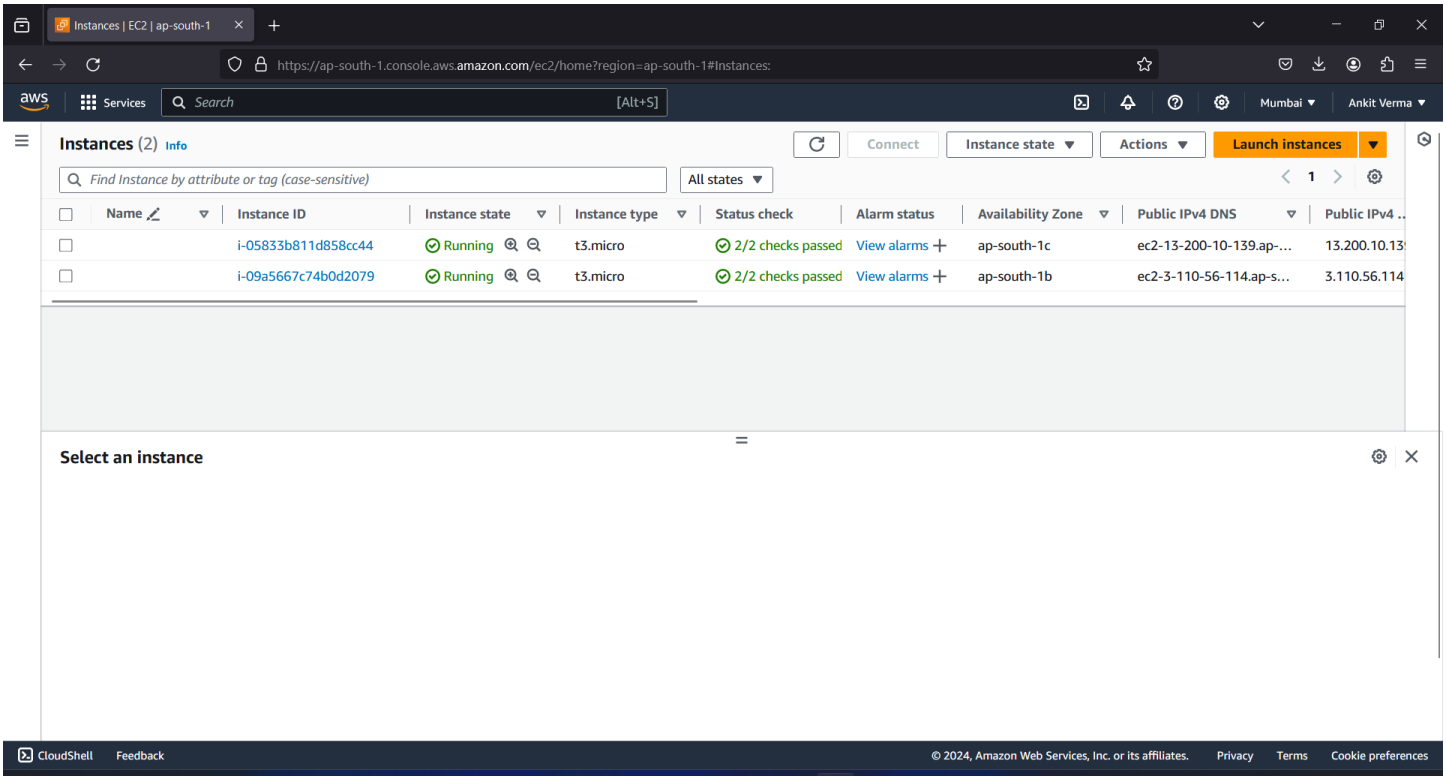
Step 18: Review all the data of the group to be created and click on Create Auto Scaling Group.



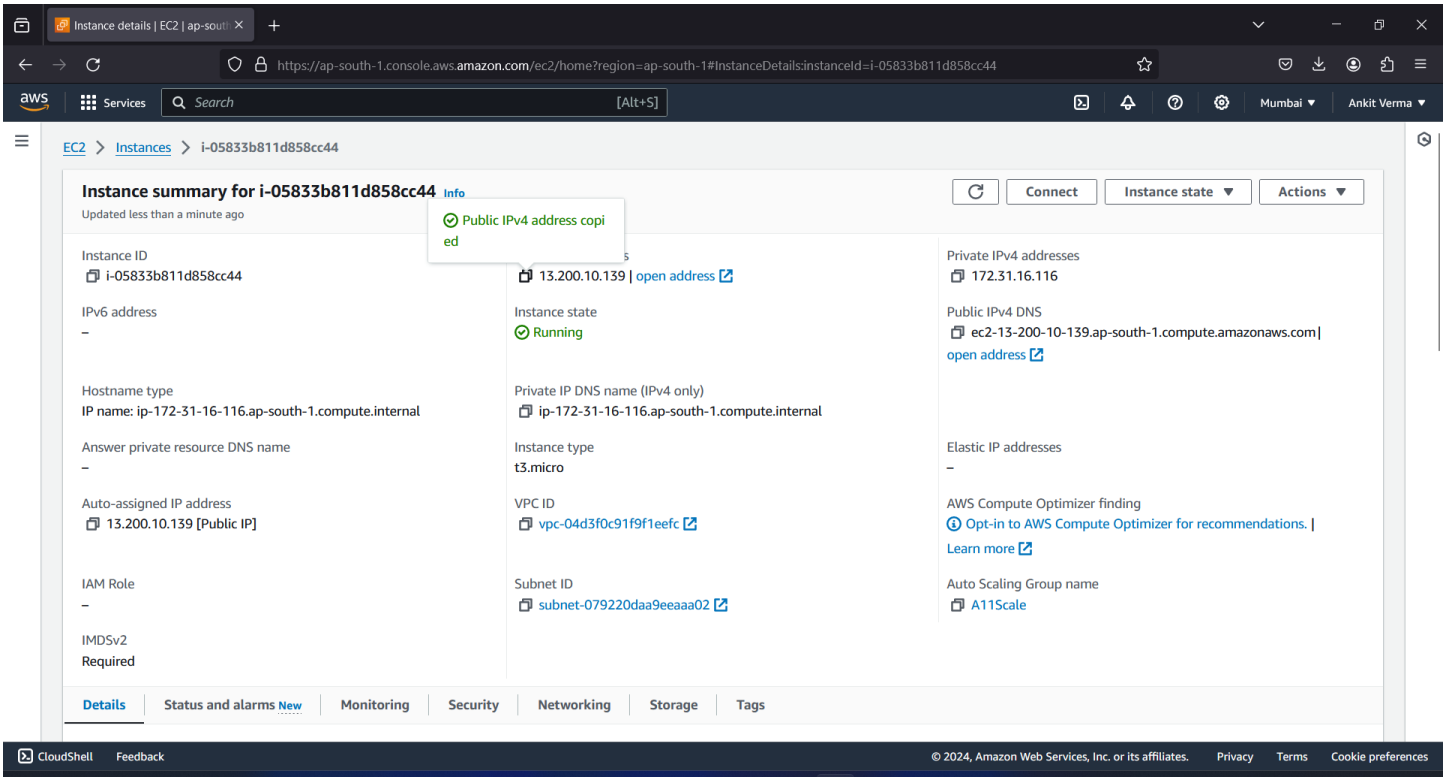
Step 19: After creating the scaling group, go back to Instances from the left side menu.



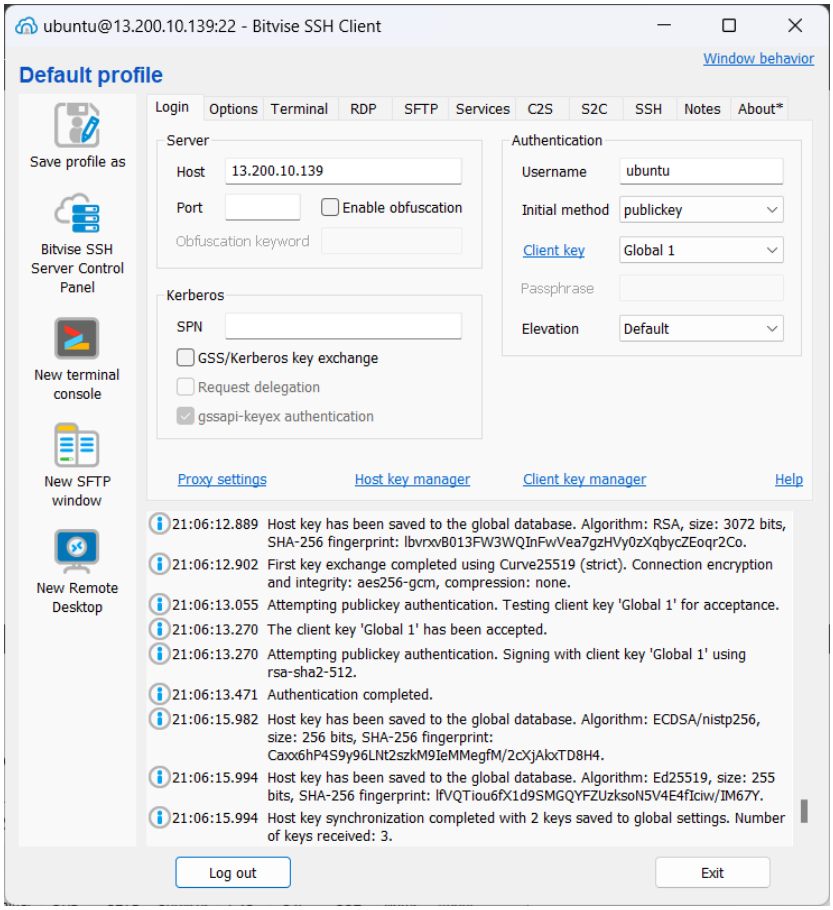
Step 20: Since the capacity was given as 2, two instances are created. Now open any one of the instances by clicking on its id.



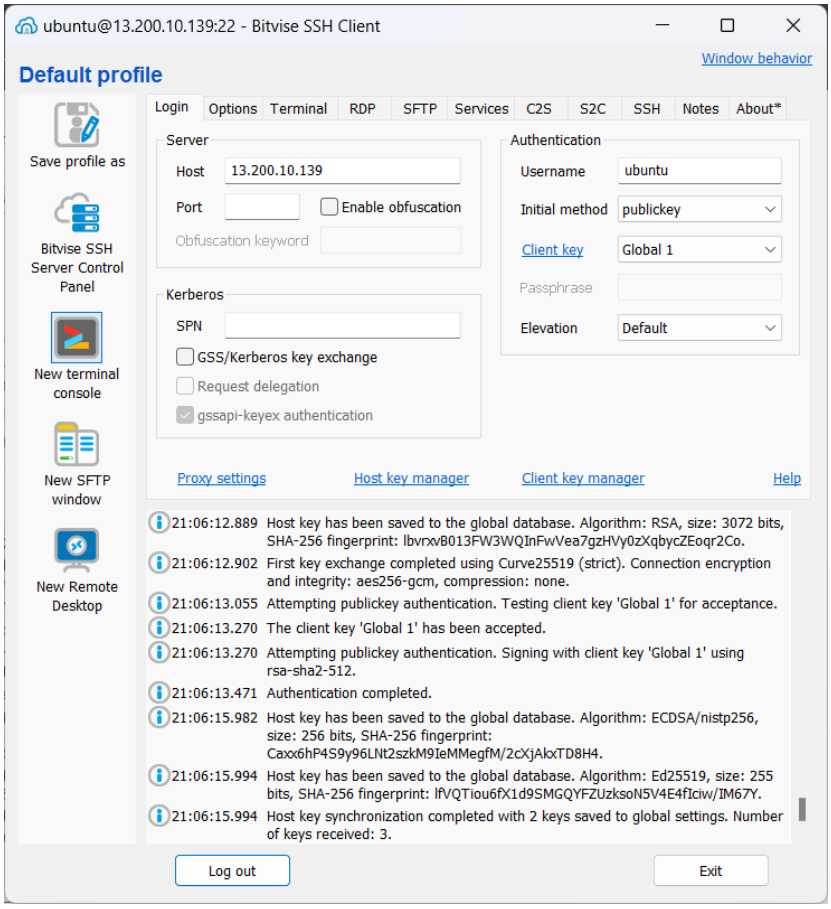
Step 21: Copy its Public IPv4 Address.



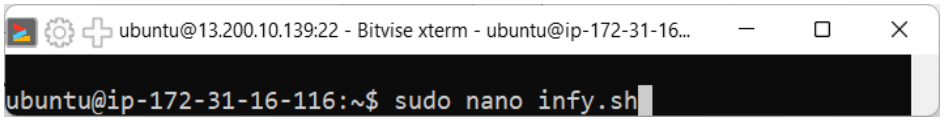
Step 22: Paste the copied address and click on Log in.



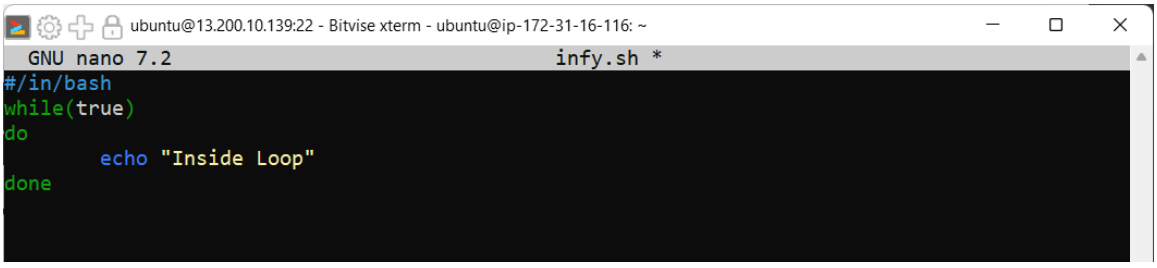
Step 23: Click on New Terminal Console.



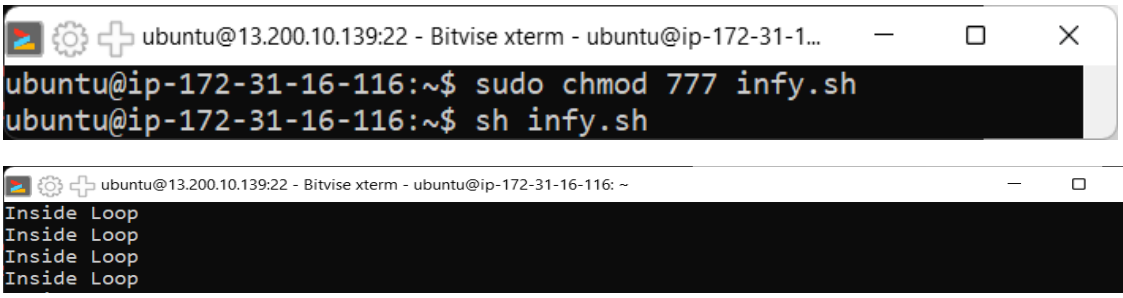
Step 24: Type the command:



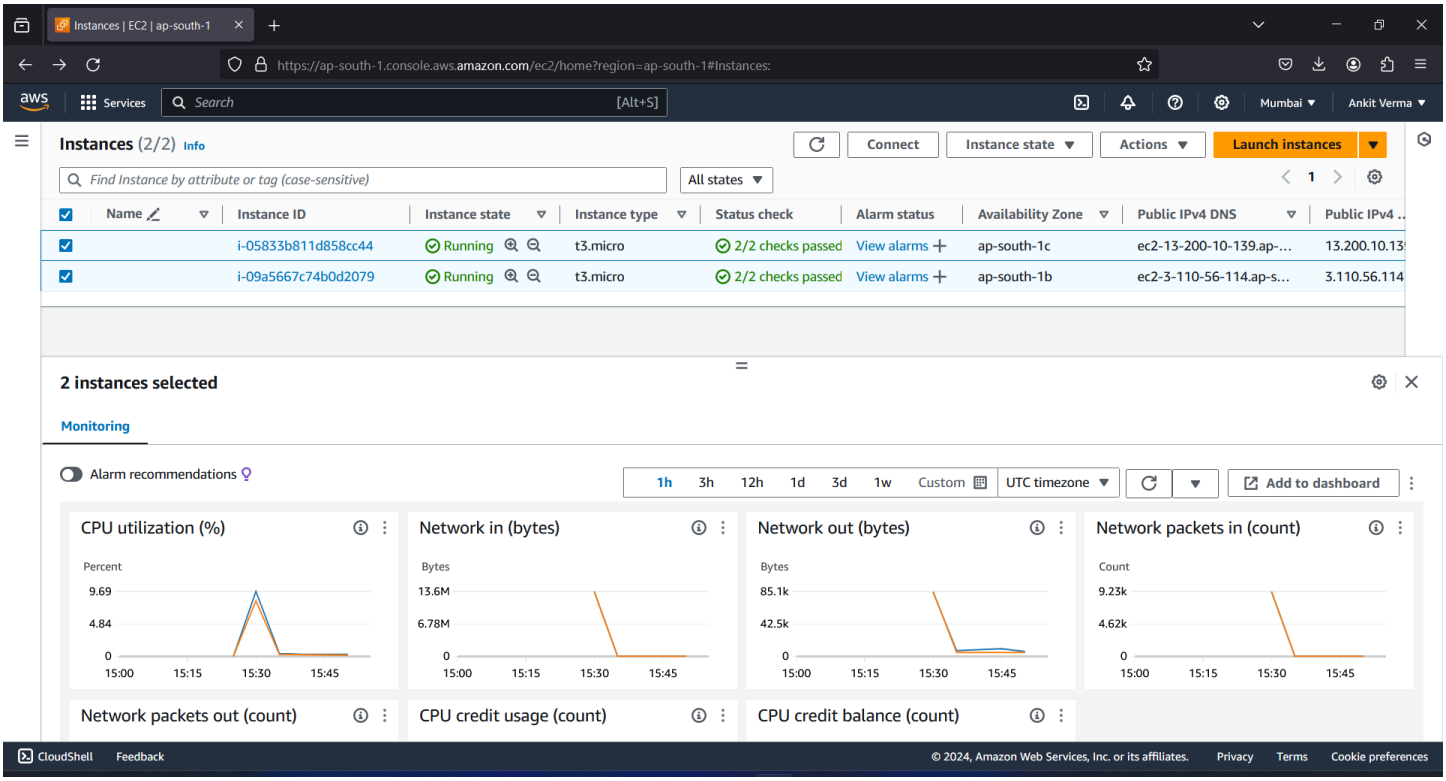
Step 25: Write the following code for an infinite loop in the infy.sh file.



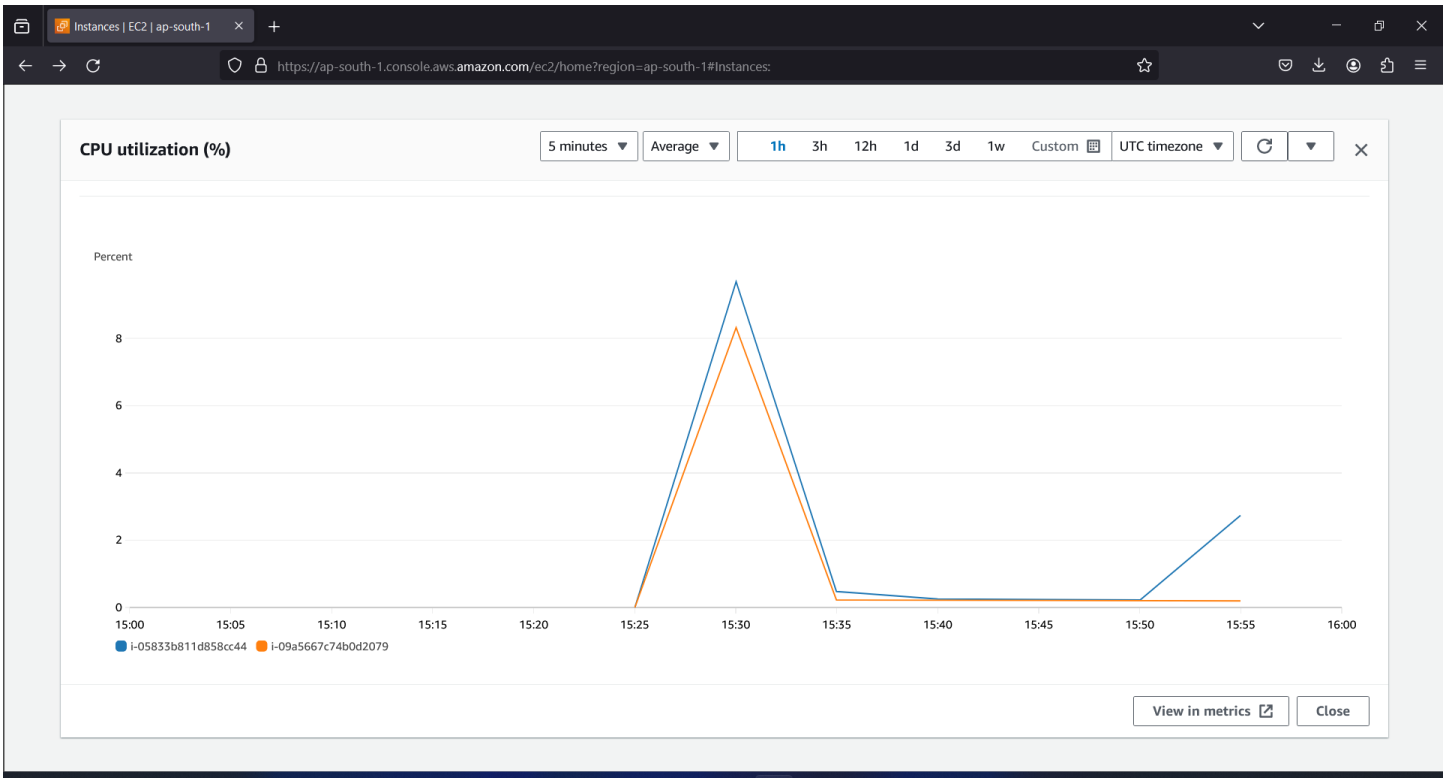
Step 26: Write the following commands in the terminal:



Step 27: Select both the instances, then under monitoring go to CPU utilization and enlarge it.



Step 28: The graph shows the CPU Utilization for both the instances.



When the CPU utilization exceed the limit for both the instances, a new instance will be created.