Autoscaling - Launch Configurations & Auto Scaling Groups

Instructions

Steps:

1.Creating launch configuration
2.Creating Auto Scaling with load balancer
3.Apply load using jmeter and verify new instances are launched based on Autoscaling (jmeter script attached)

Pre-requests

1.Launch a new ec2 instance and install and run nginx server alone
2.Create an image based on the above instance where you have nginx running

Steps to create Launch configuration

1.Navigate to Auto scaling services -> Launch configurations -> create launch configurations
2.Provide the Name, choose the AMI that you have created in the pre-requests
3.Choose Instance type ad t2.micro
4.Additional configuration , EBS volume as default settings, also enable public assign a public ip option
5.You can create or select existing security group
6.Choose from existing key pair
7.Create launch configuration

Create a Load Balancer for Auto scaling groups

1.Navigate to Load Balancers -> Target groups click create Target group
2.Target Type as Instance, provide Target group name, Protocol http : 80 (default)
3.Select your VPC
4.In the Advanced health check settings you can give your custom values in the traffic port.(you can leave as default if you don't wish to change it)
5.Add tag and click Next
6.In the Registered target, do not select and Instance
7.Click ‘Create Target group’ and it will be created successfully
8.Create a load balancer by selecting Application Load Balancer
9.Provide the load balancer name, select Internet-facing, and IPv4 address type

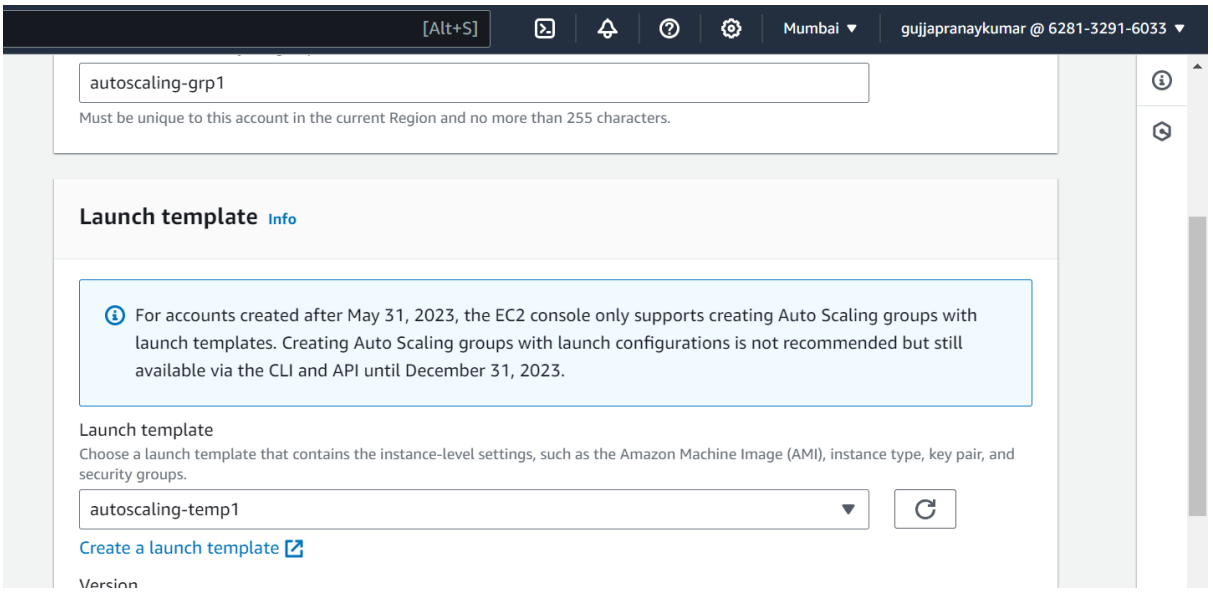
10. In the network mapping select your VPC
11. In the subnet mapping select the availability zones and select public subnets that you have created from the dropdown
12. Select the security group from existing that you have created for previous Load Balancer assignment
13. In Listeners and Routing select the target group that you have created above
14. Provide the Tags and create the load balancer

Steps to Create Auto Scaling group

1.Navigate to Auto Scaling group -> Create an Auto scaling group

2.Provide Auto scaling group name

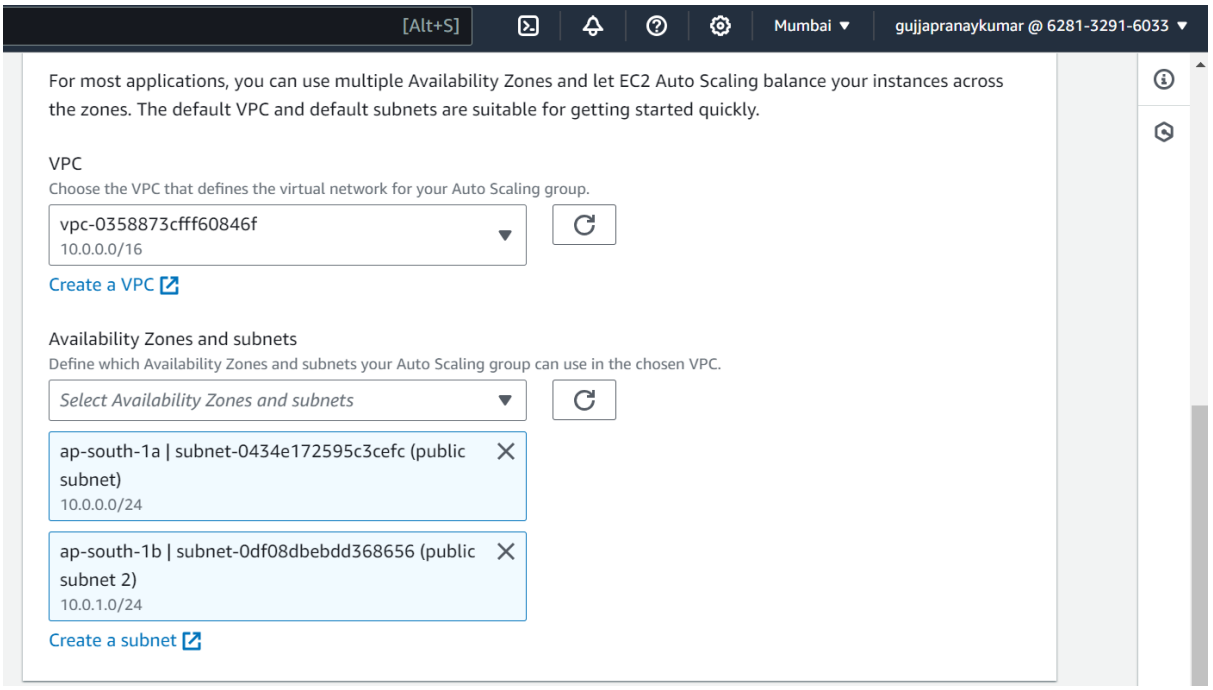
3.In the Launch template option click switch to launch configuration
A:



4.Select the Launch configuration that you have created and click Next

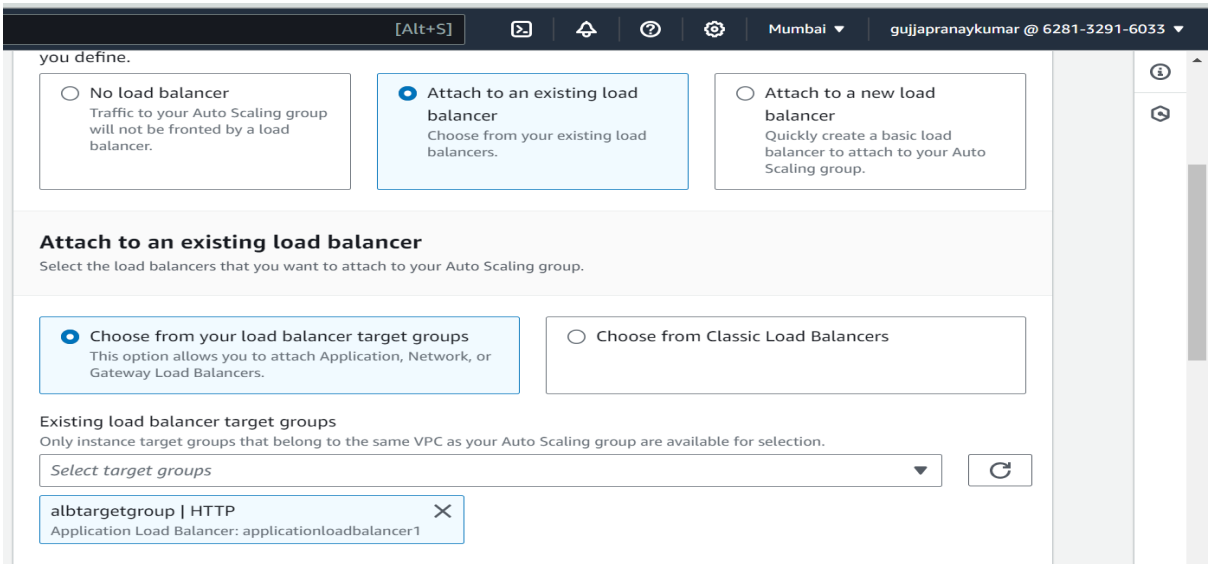
5.In the Network , choose your vpc, choose the different availability zones of public subnet

6.Click Next
A:



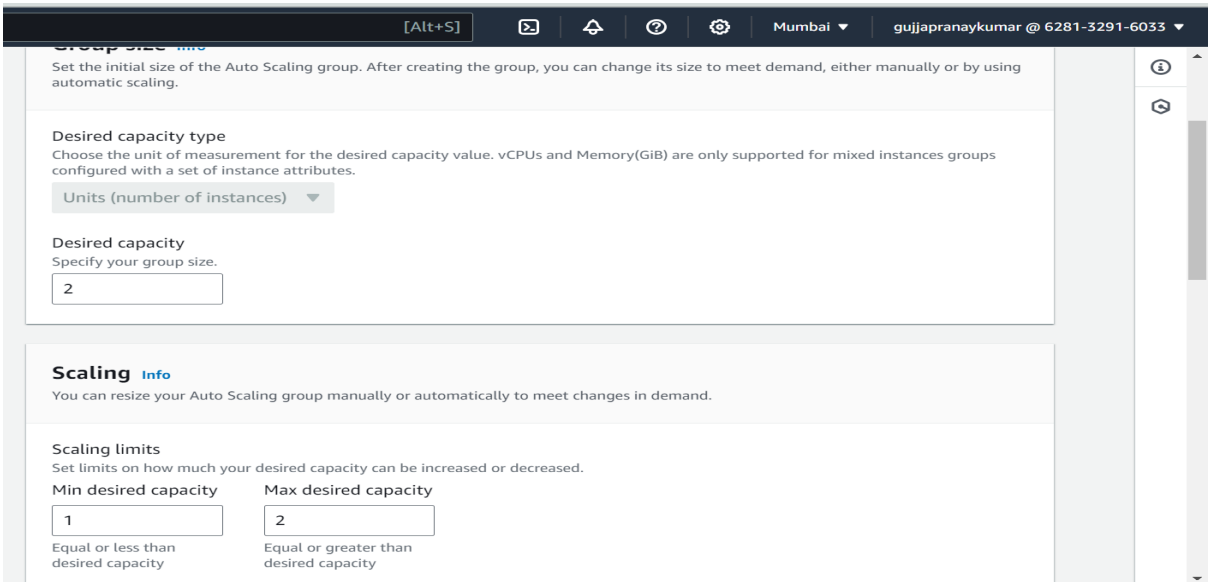
7.Now select attach to existing Load Balancer
8.Select the target group that you have created for Autoscaling
9.Leave Health checks and Additional settings as default
10. Click Next

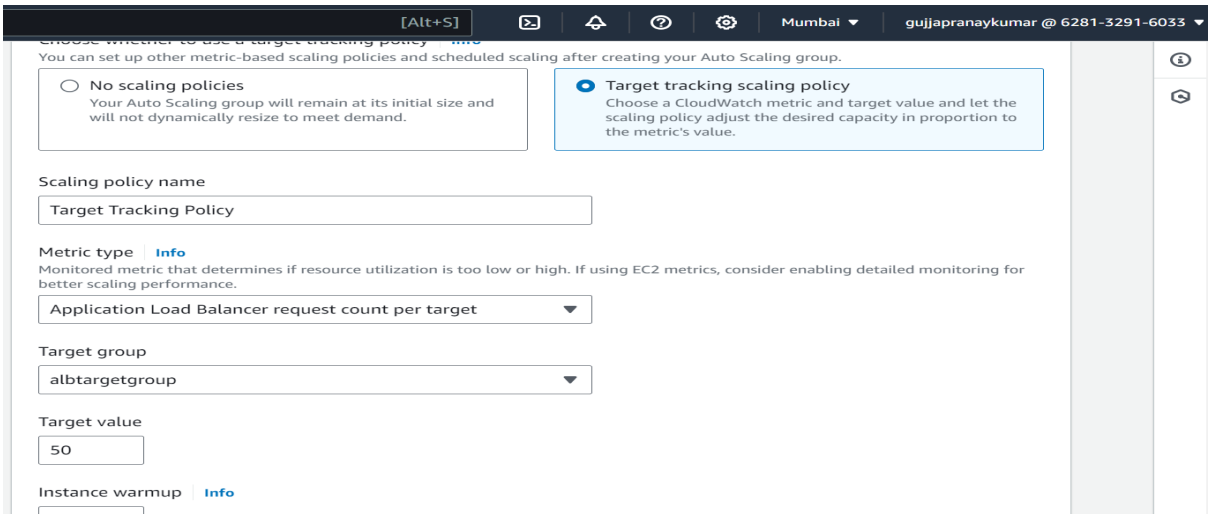
A:



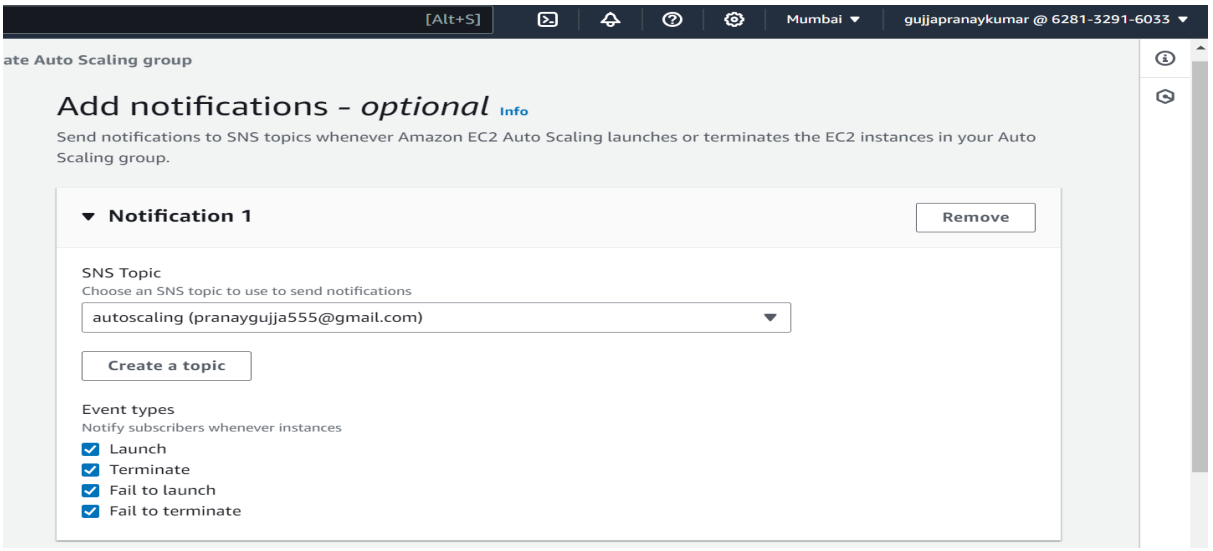
11. In the Configure group size set as below 2,1,2
12. In the Scaling policies , select Target scaling policy and provide the options as below
13. Click Next

A:





14. In Add notification click on create a topic,provide your email address
A:

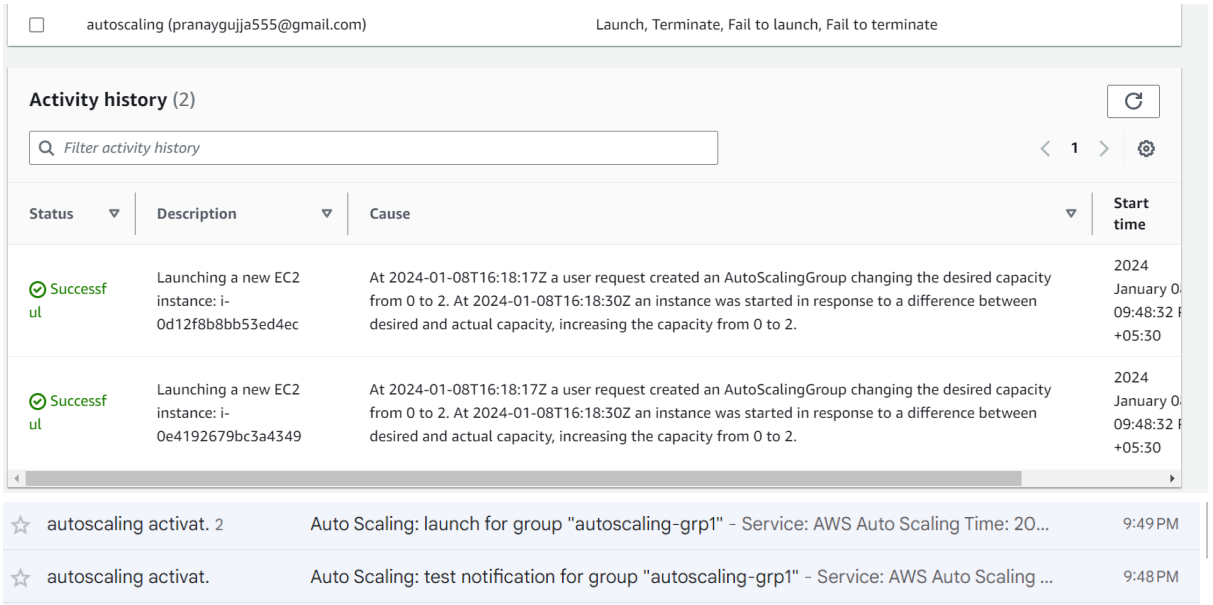


15. Provide the Tags for new instances

16. Verify all the configurations and click on ‘create auto scaling groups’

17. Auto Scaling group will be created as below

18. Once Auto scaling group created you can see the instance will be launched automatically as per the minimum capacity value. You can check the target details in your target group
A:



19. Also you can see the instances list, where you can see Auto scaled instance will be up and running

Test the Auto scaling by applying Loads using jmeter(install in windows)

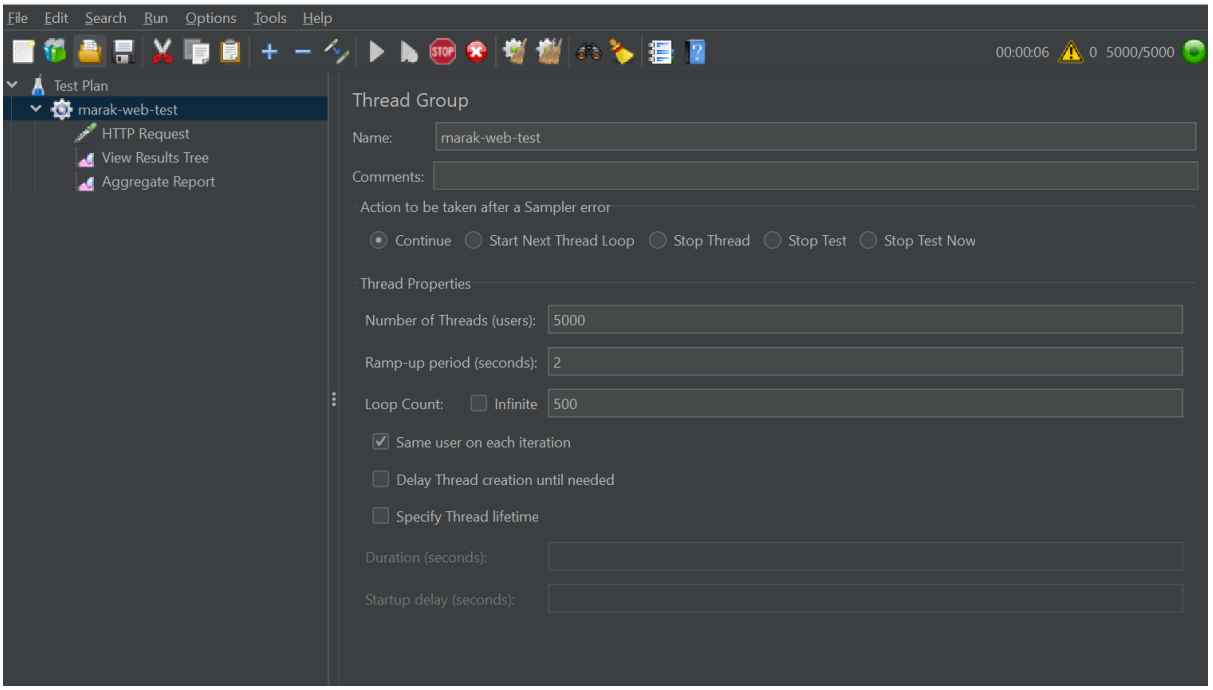
1.Download and install Apache jmeter 5.5 from this url
https://jmeter.apache.org/download_jmeter.cgi

2.Run jmeter batch file inside bin folder to open the console

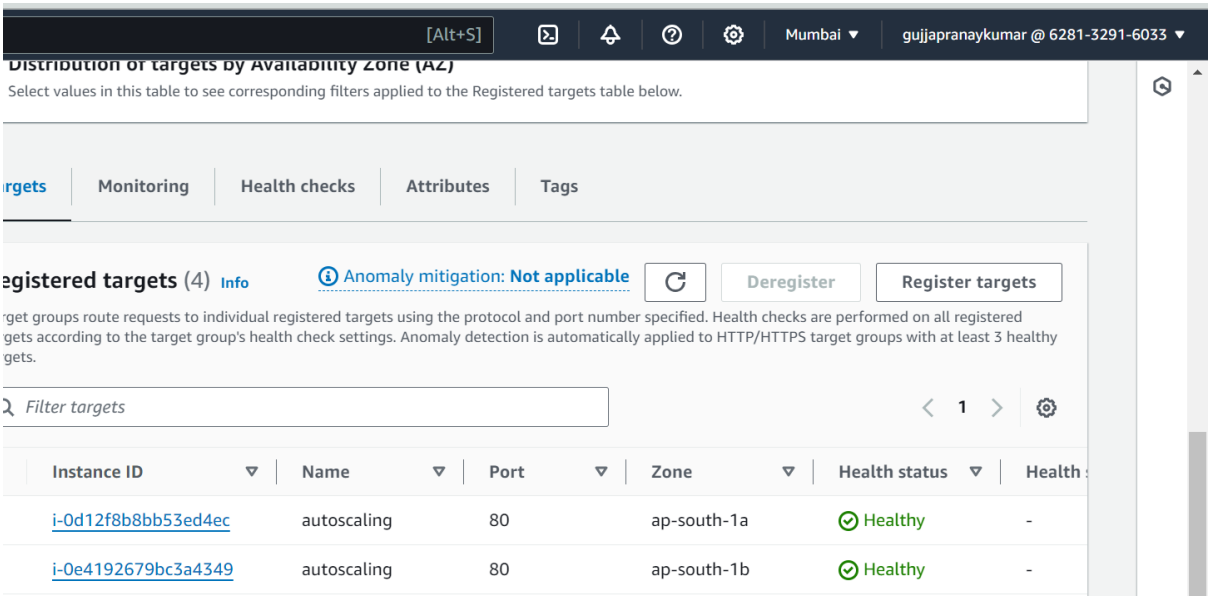
3.Open web-test.jmx in the below link your_Load_Balancer_url update your url and save changes

https://drive.google.com/drive/folders/1HlpPpbsMu75aCLpnQogfPhb4XBB84aUW

4.In the jmeter console open the web-test.jmx file and click play button and it starts to push the load to your Load Balancer
A:

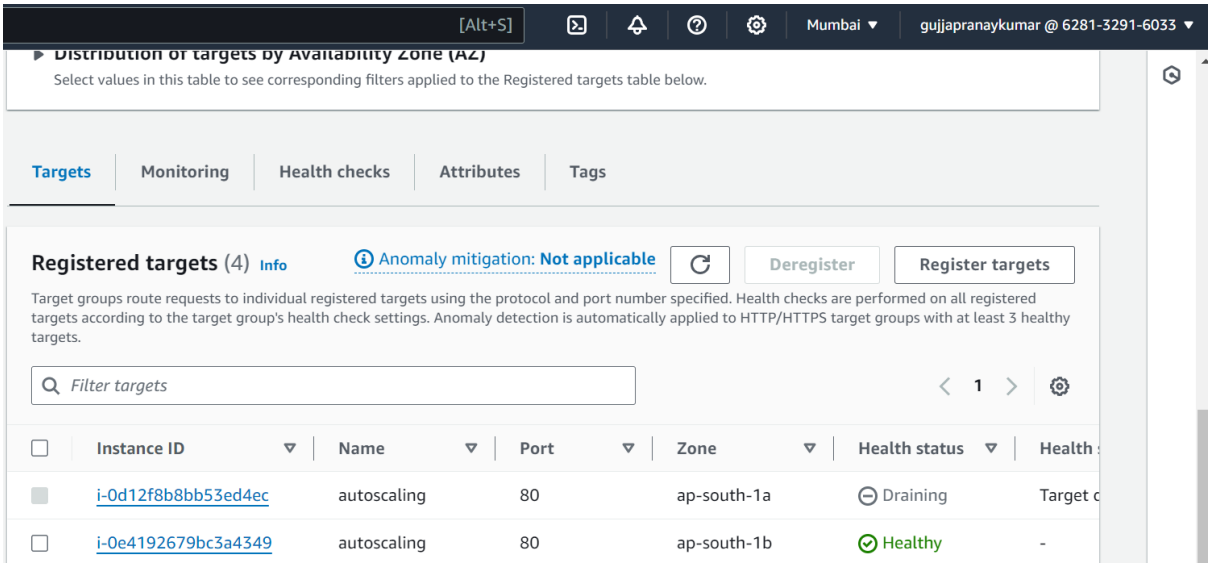


5.Check the Target group Registered targets after 5 minutes you will see a new instance launched as per Auto scaling configuration
A:



6.Once after the Load run gets completed in the jmeter the 2nd instance will drain automatically after some 15 minutes of time and the instance will get terminated eventually as seen in the picture below

A:



7.Once Auto scaled instance gets drained it will get terminated automatically you can verify it in the Instance list

_____END_____