

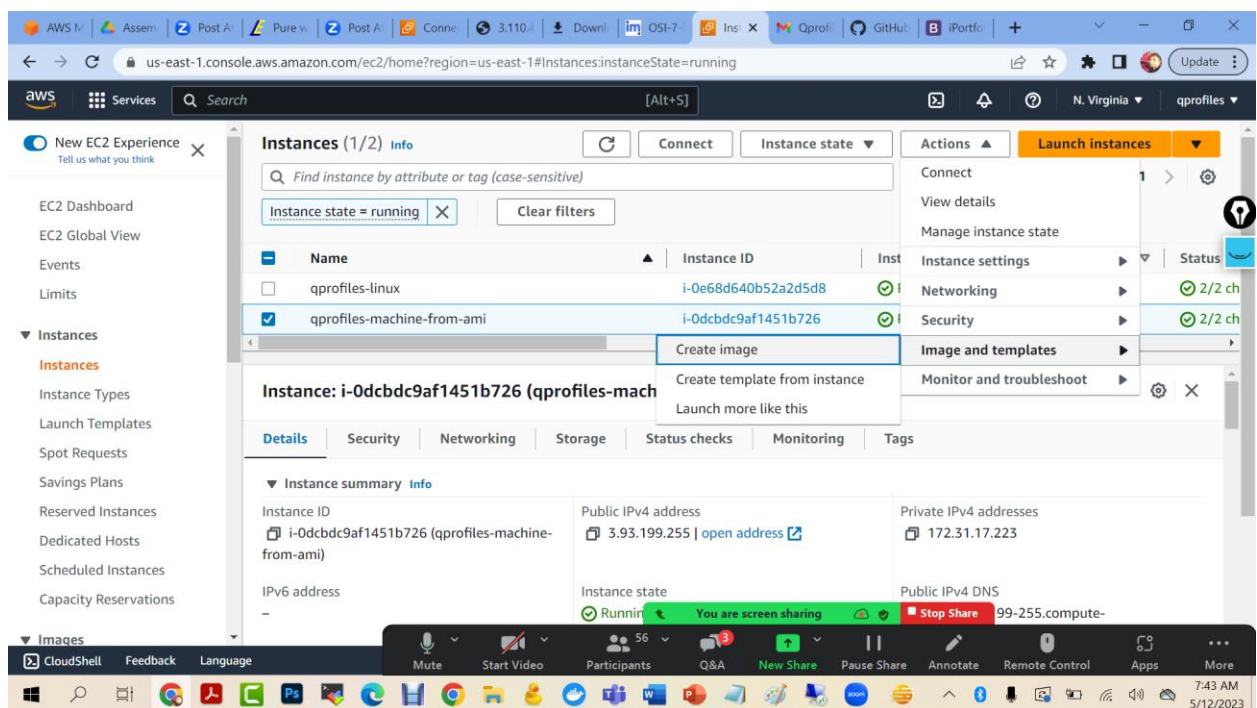
Autoscaling :

Increase/ decrease of the number of server as per the need and requirement .

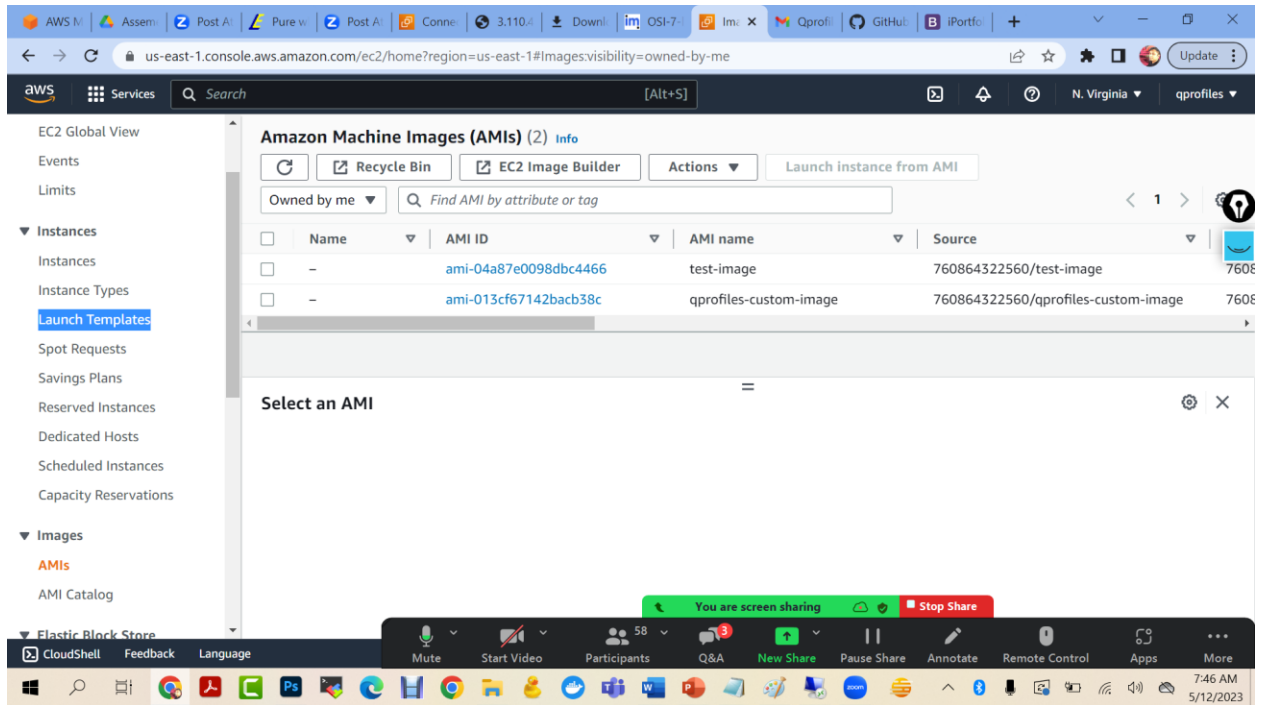
For example : if my server is capable of handling 1000 users/min but currently I have 1200 users/min ??  
what about these 200 extra users? For this we will enable autoscaling in which automatically new machine is created when the traffic increases and vice versa

Lab:

- 1) Create a windows ec2
- 2) Launch the website in ec2 machine
- 3) Create a image of the ec2 machine .  
Select the machine -> actions -> images and templates -> select the create image



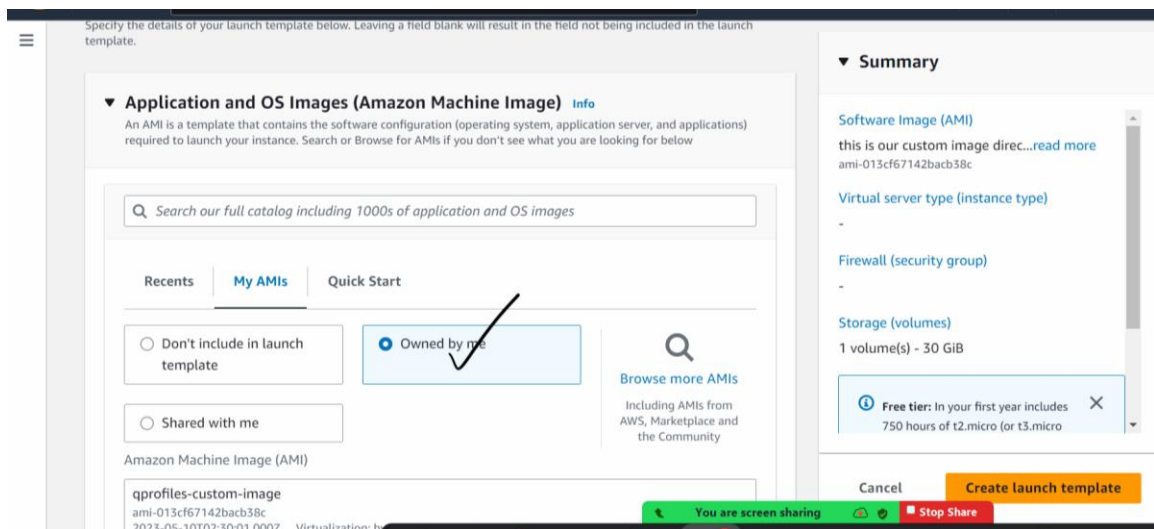
- 4) Now we will create a launch template



5) Launch template -> create launch template

Give some name

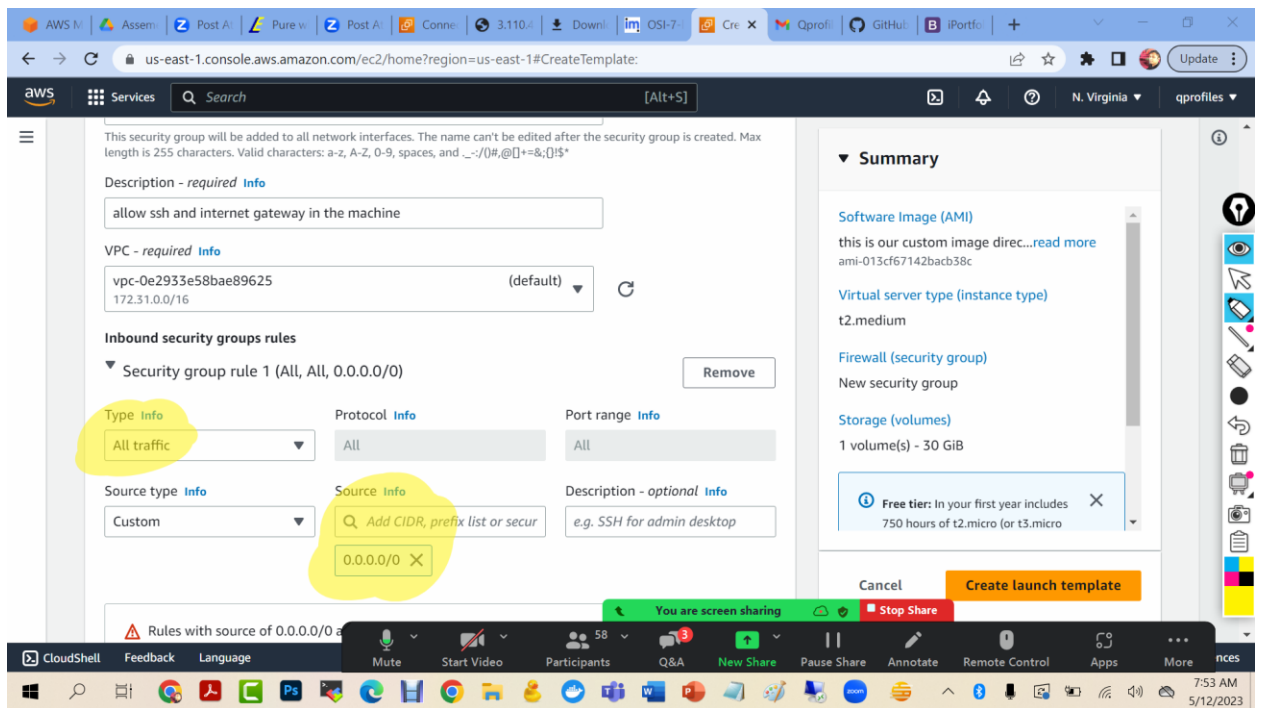
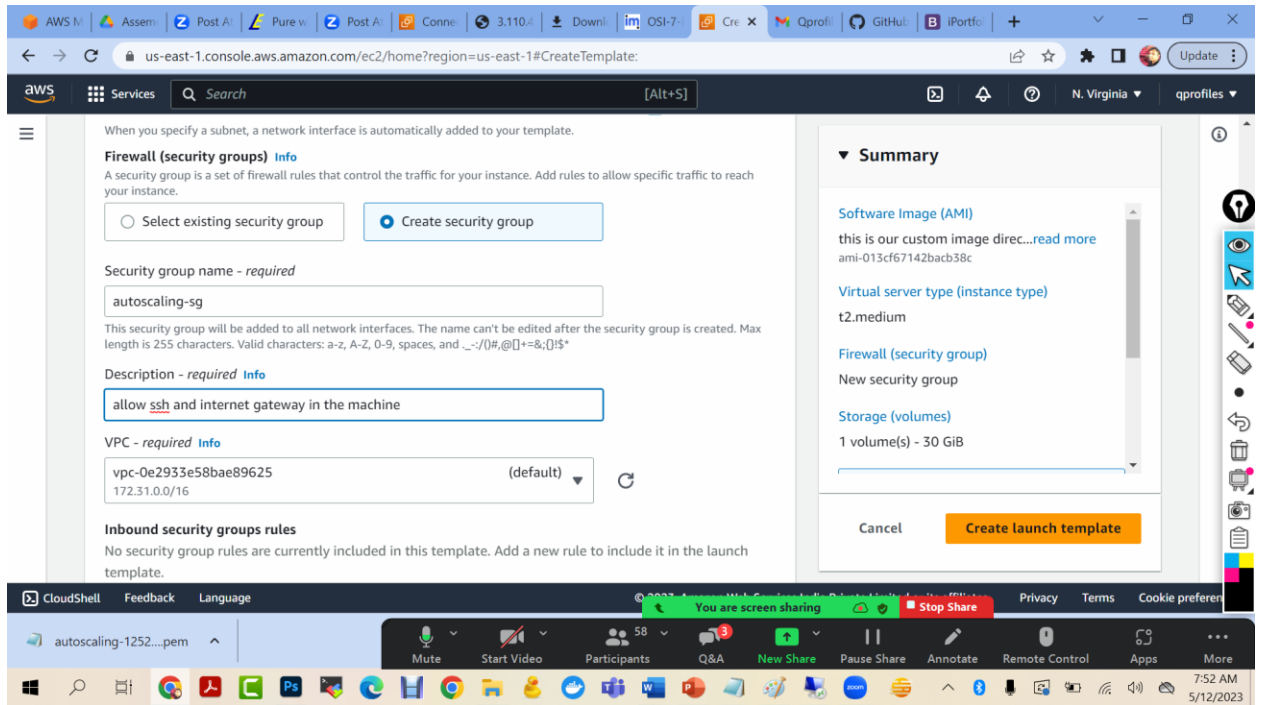
In ami -> select owned by me -> select the custom image which u have created



In instance type -> t2.medium (it means during autoscaling the machine which would be created would be t2.medium)

Use existing or create a new keypair

In security group -> create a new security group



Create a launch template

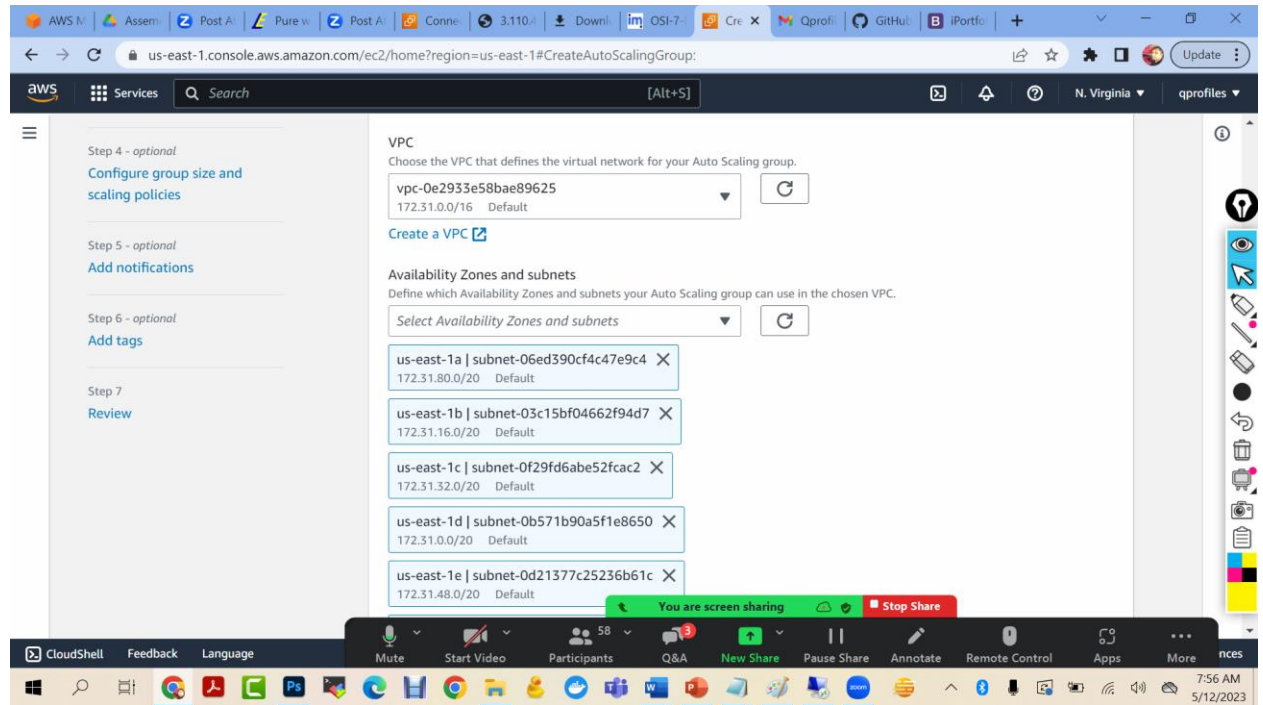
- 6) Now we will create autoscaling group  
Go to autoscaling group -> create autoscaling group

Give the name of autoscaling group

Select the launch template which u created in step 5

Next

Select all the subnets

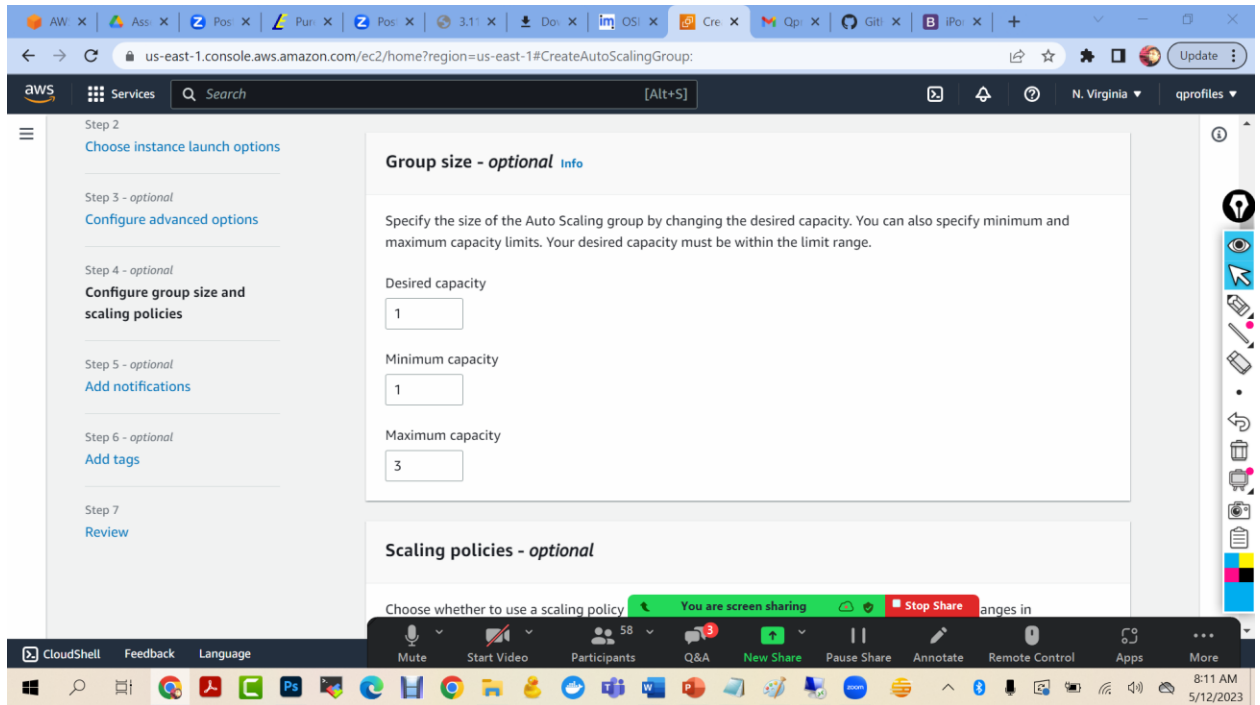


Next

Select no load balancer

No vpc lattice service

Next



## Scaling policies : Condition when to create new machine

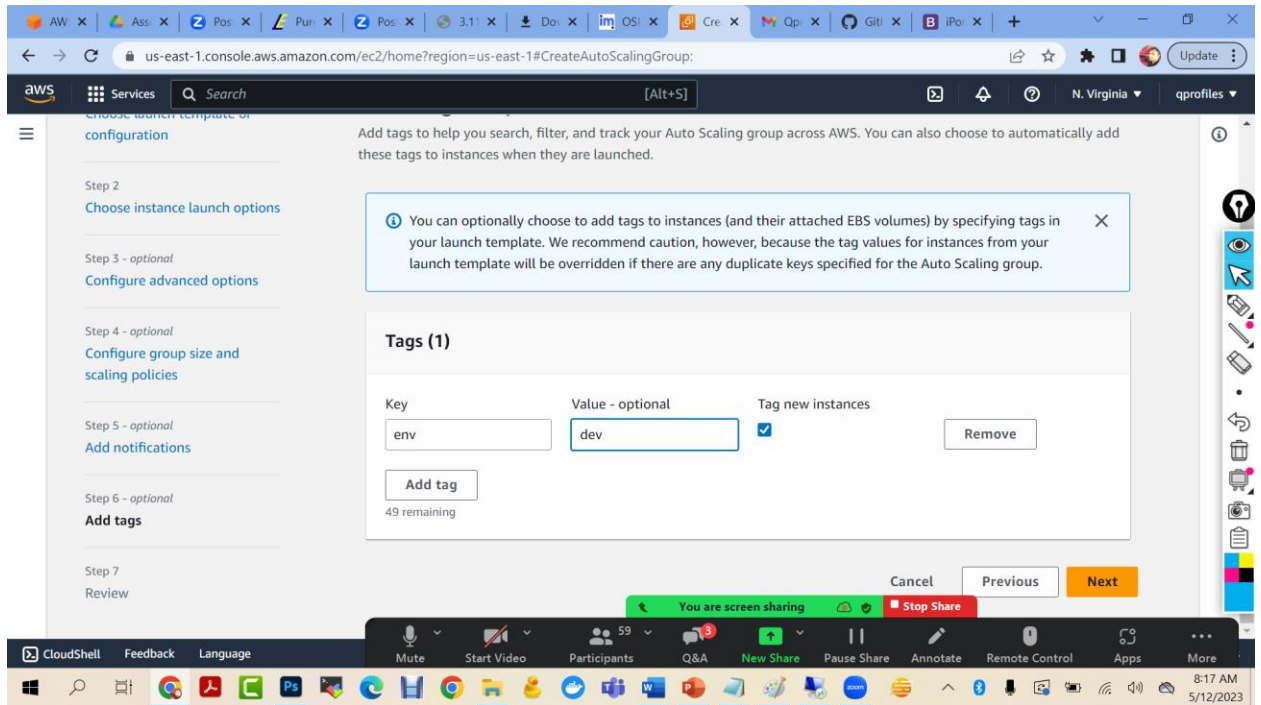
If cpu utilization goes above 20% then a new machine should be created up

Target tracking -> 20%

Next

Next

Next



Create a autoscaling group

Wait for some time

Let it update .... It would be creating one instance

Go to ec2 -> you will see extra instance