1. Run 2 web servers behind ALB.  
   * + Create SG for Application Load balancer for port 80 to access anywhere

Graphical user interface, application

Description automatically generated

b. Create an SG for an EC2 instance. Open up port 80 for the ALB SG.

A screenshot of a computer

Description automatically generated

c. Create 2 servers in 2 different AZs (us-east-1a, us-east-1b) with different HTML content.

Refer to the code snippet below and put it into “User Data”. Hit Edit in the “Network Settings” and select us-east-1a in instance1 and us-east1b in instance2. Select the SG for the webserver you created in the previous step.

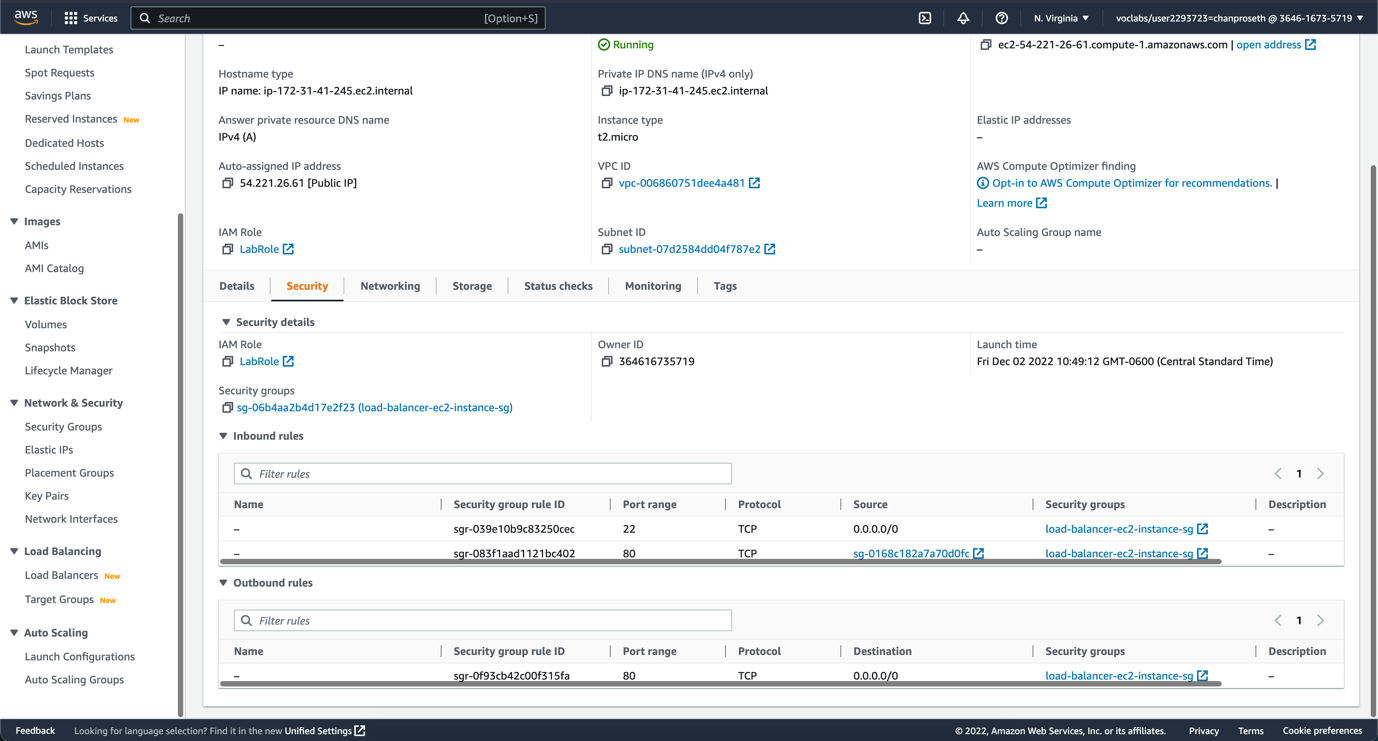
Graphical user interface

Description automatically generated

* + - Add ec2-lb-0102122022 and ec2-lb-0202122022 instances security group that allow only load balancer to access

Graphical user interface, application

Description automatically generated



* + - Add code snipet user data to each instance

Graphical user interface, application

Description automatically generated

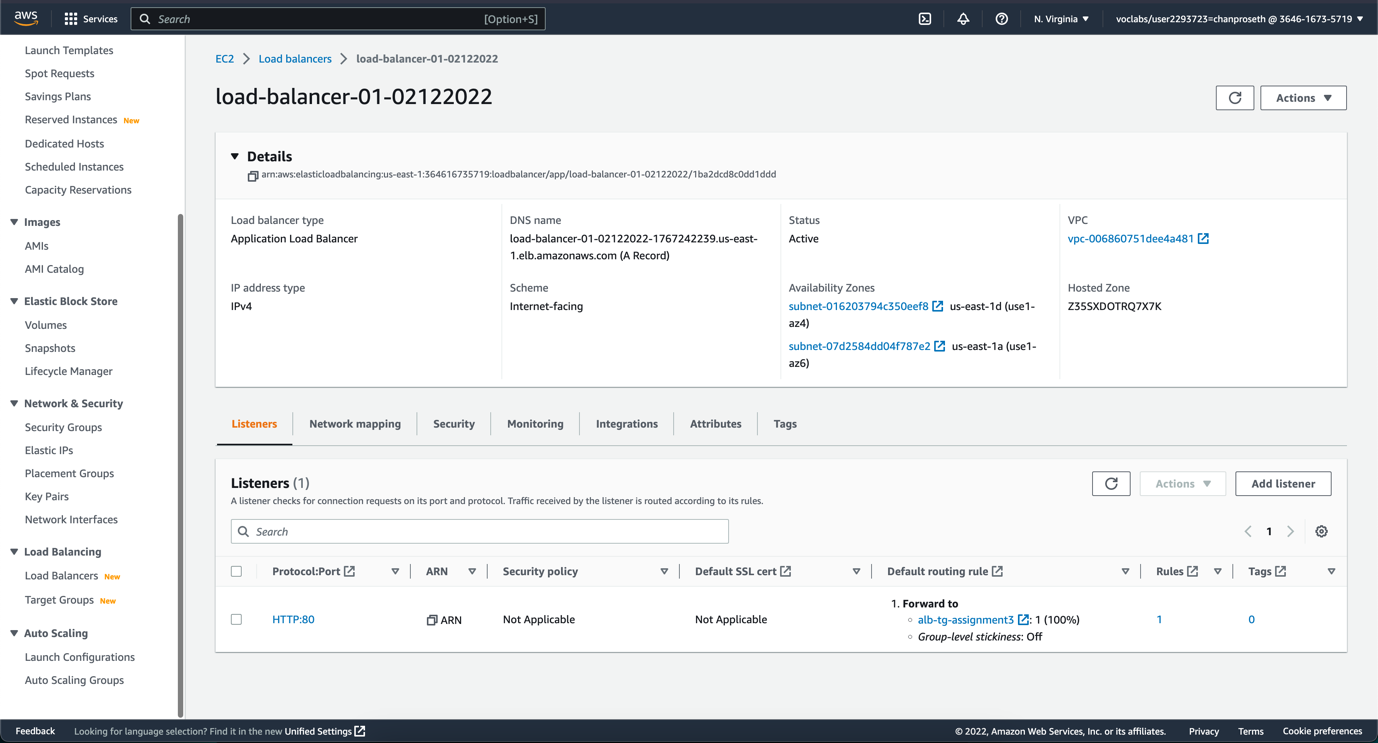
d. Create ALB. Select 2 AZs (us-east-1a, us-east-1b) for HA (High Availability). Create the TG and register the servers. And select the TG you created.

* + - Create Target Group

Graphical user interface, text, application

Description automatically generated

* + - Create Application Load Balancer



Graphical user interface, application

Description automatically generated

* + - Result of load balancer sequentially change page each time we refresh

Graphical user interface

Description automatically generated

Graphical user interface

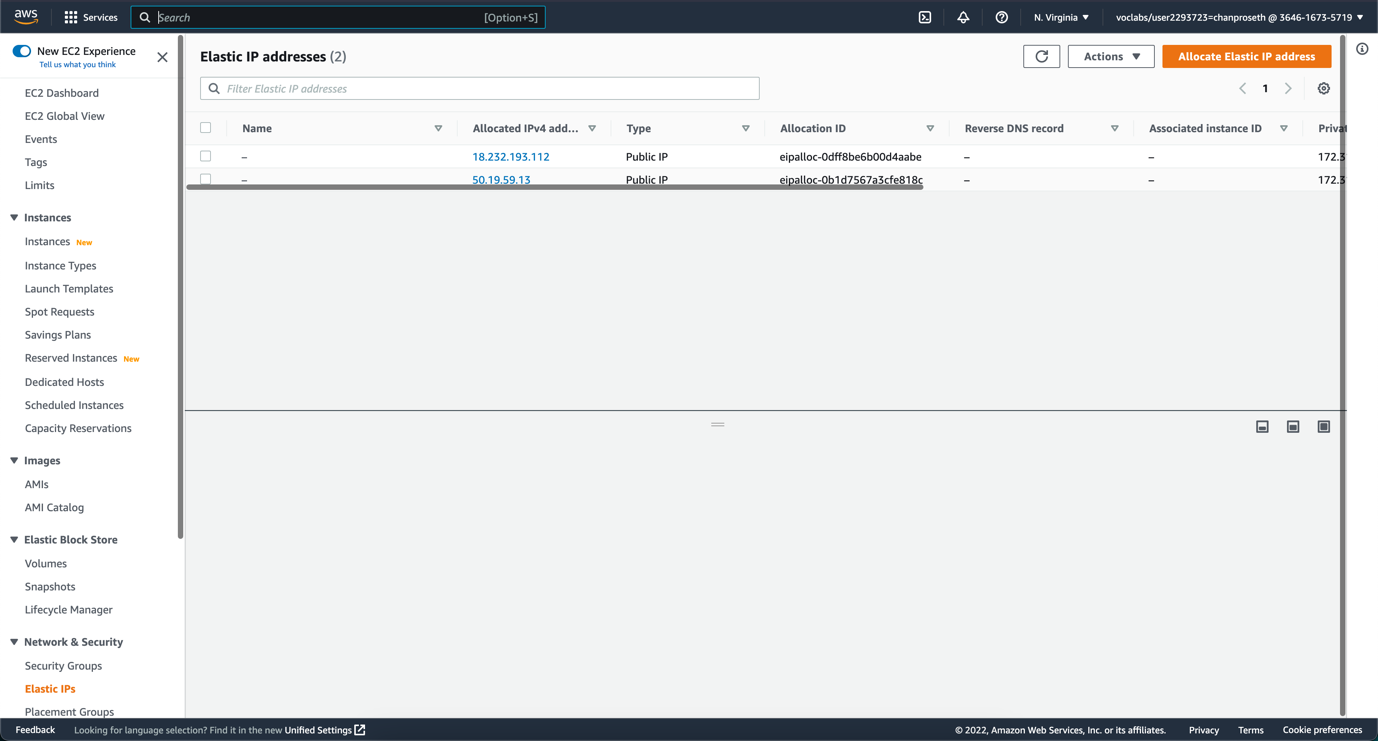
Description automatically generated

1. Run web servers behind NLB.

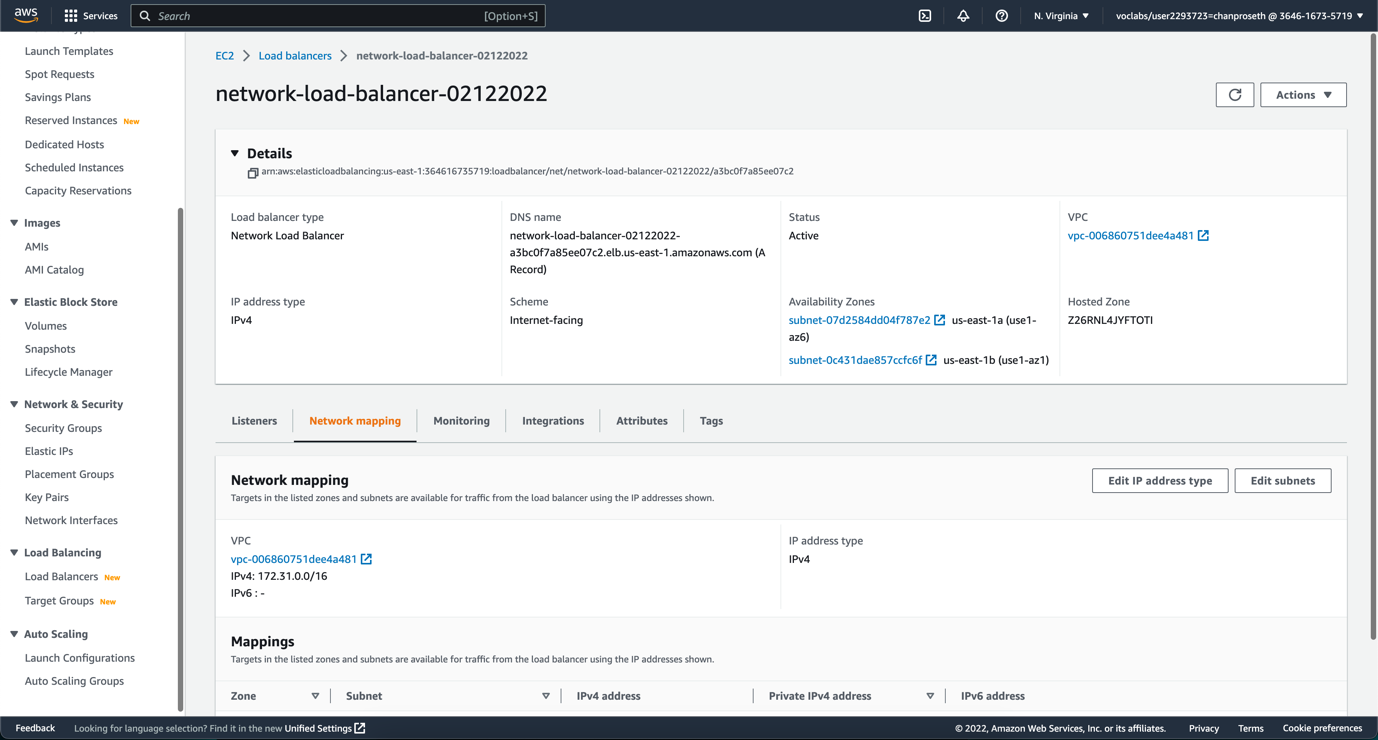
NLB operates at layer 4. There is no SG. How you are going to whitelist access only from NLB in your EC2 SG?

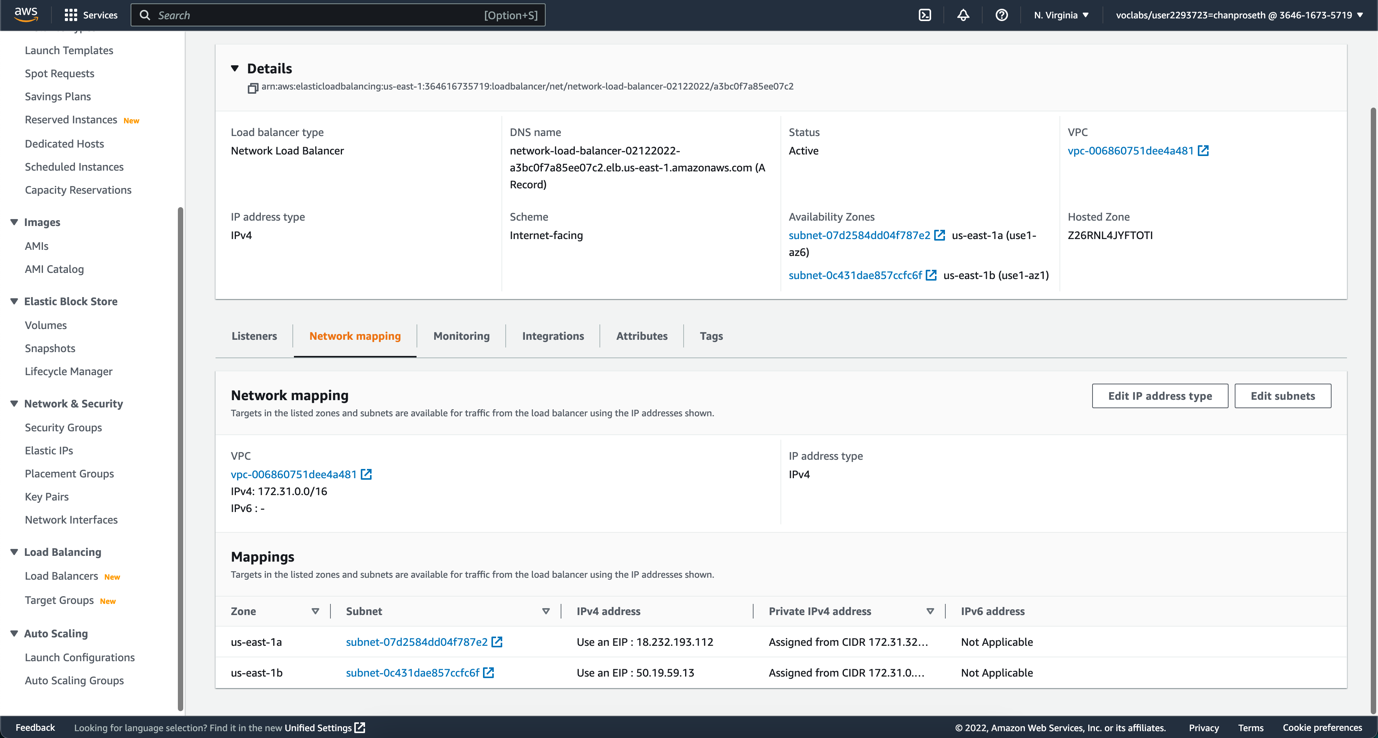
The solution is to use elastic IPs in the NLB nodes. So we can whitelist those static IPs in the EC2 SG. We whitelist the private IPs of the elastic IP. Because all traffic inside the same network (VPC) uses private IPs.

1. Create 2 elastic IPs.



1. Create NLB and associate the nodes in AZ with the elastic IPs.





1. Add the instances in us-east-1a, and us-east-1b that you created in task 1 to the target group of

the NLB. Protocol must be TCP, not HTTP.

Graphical user interface, application

Description automatically generated

1. Once NLB is being provisioned, you will find private IPs in the Elastic IP. Grab those IPs, and

allow it in in the EC2 SG.

Graphical user interface, application, Teams

Description automatically generated

Graphical user interface, application, Teams

Description automatically generated

1. Update the target group and deselect Preserve client IP addresses.

Graphical user interface, application, Teams

Description automatically generated

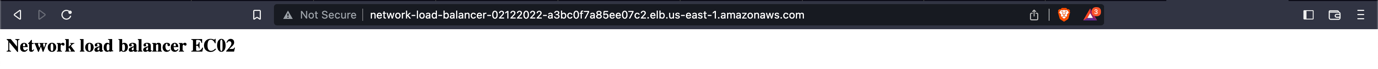
1. Grab the private IP of the elastic IP in the VPC elastic IP section. Update the private instance’s

security group to allow private IPs of the elastic IPs that you created in the first step.

Graphical user interface, application

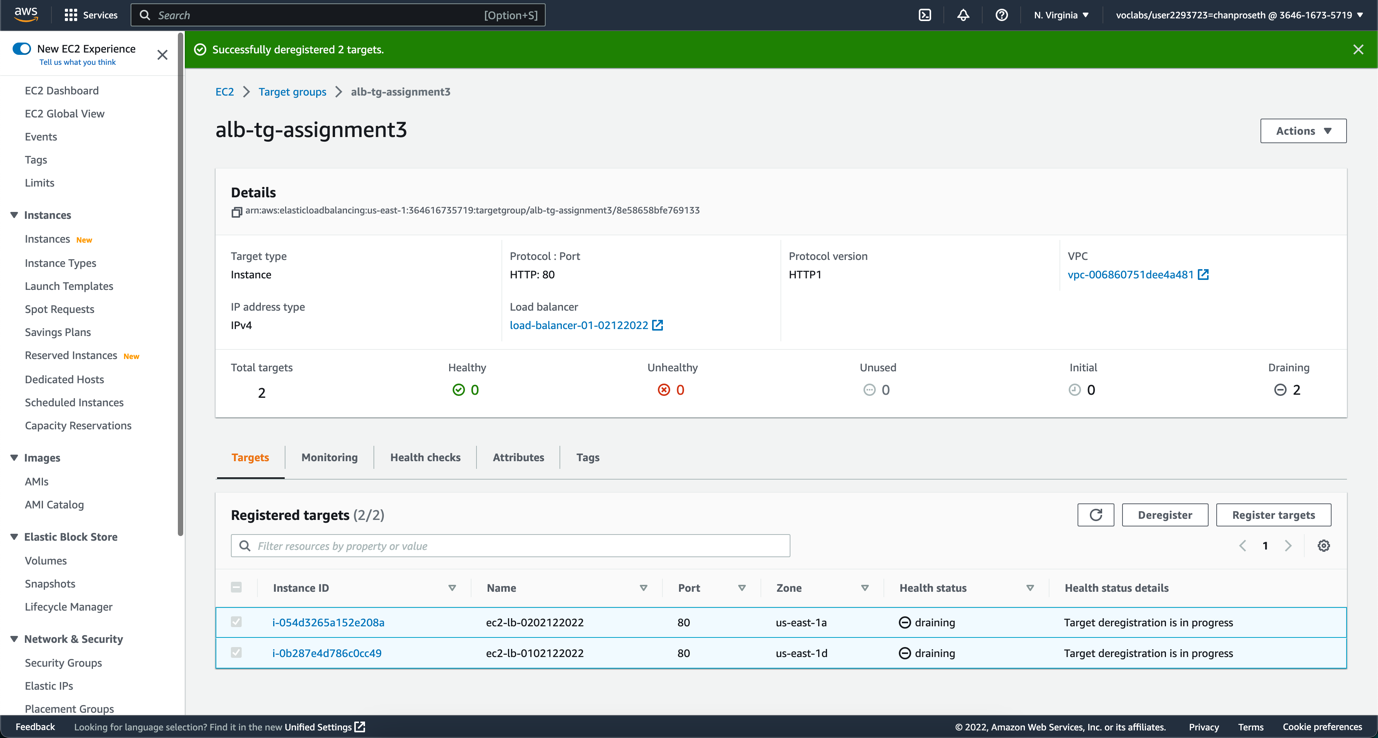
Description automatically generated

* + - Result of network load balancer

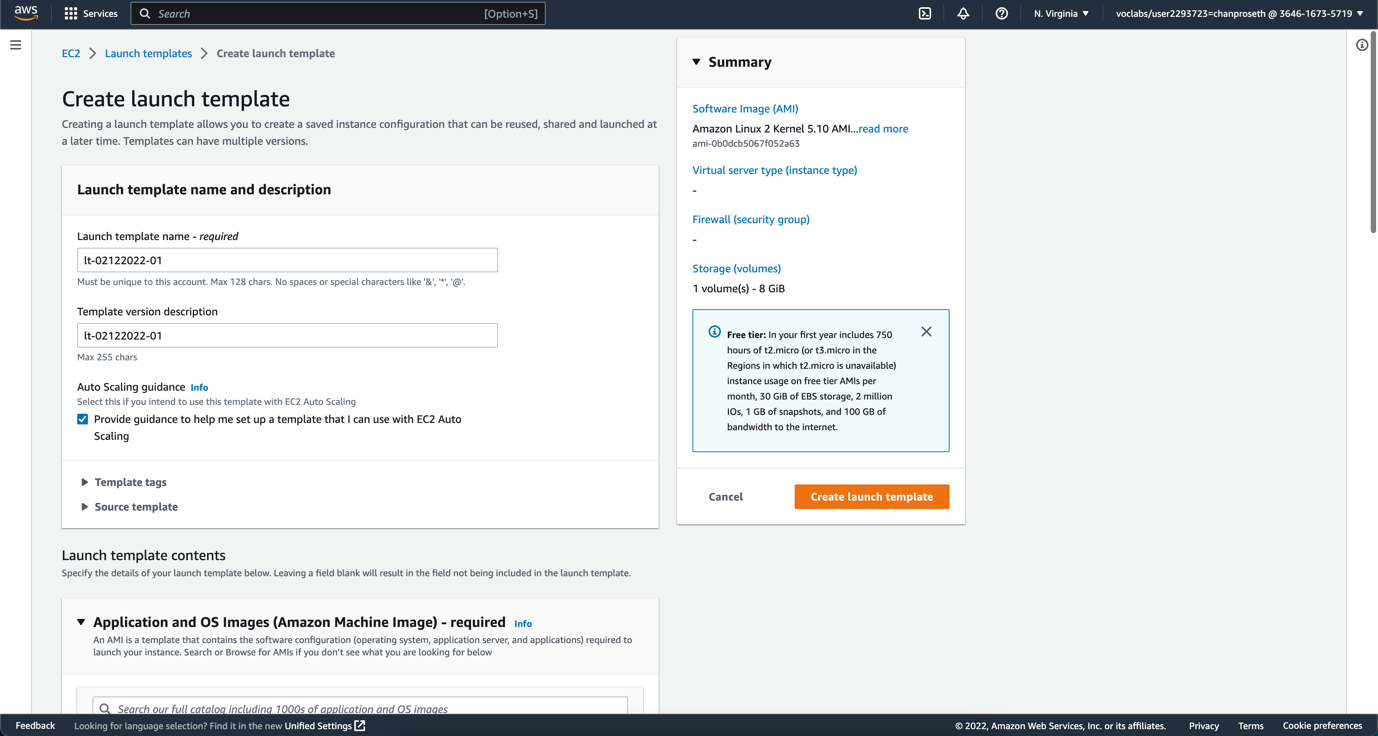


3. Run the web server behind the ALB in ASG.

1. Deregister instances behind the ALB. We will register them inside ASG so it can scale automatically.

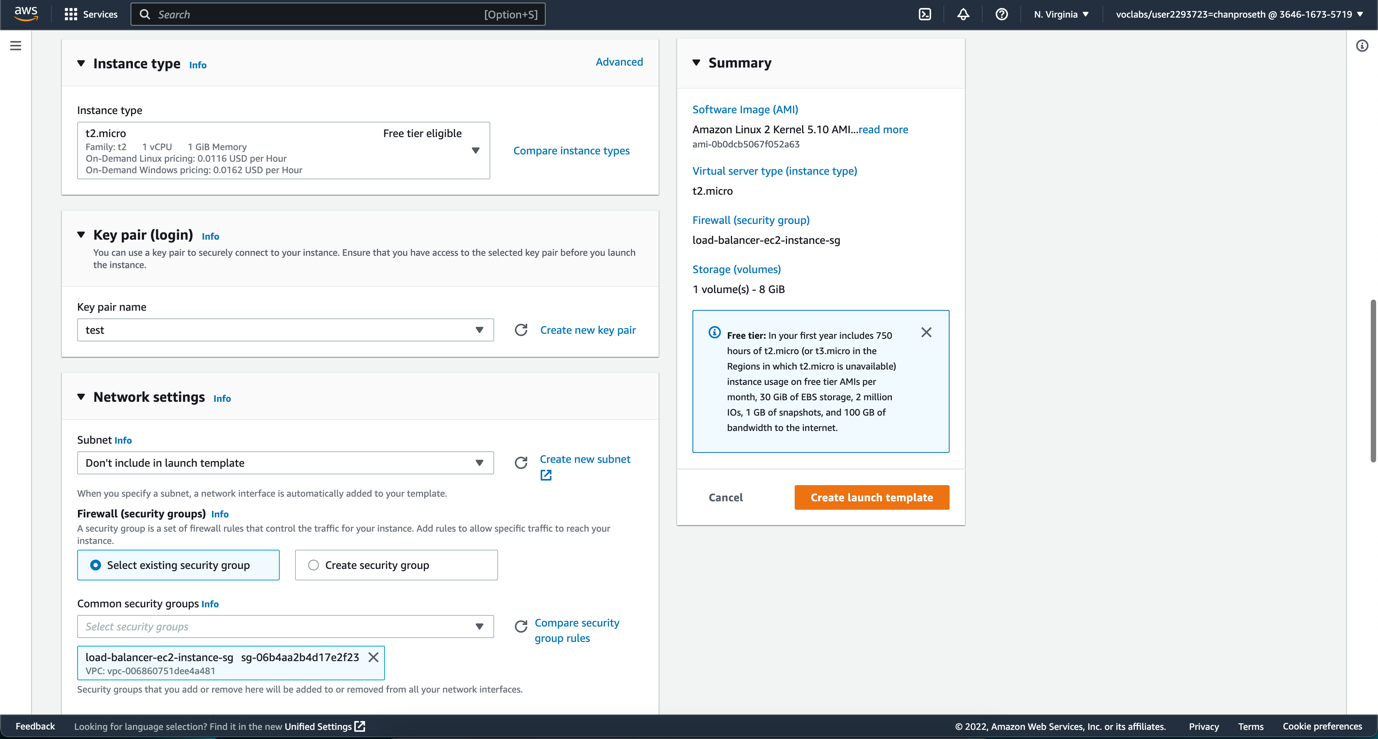


1. Create a launch configuration. Launch template is recommended. You can use that instead.
2. Give it a name
3. Select the Amazon Linux AMI. You can find the AMI ID from EC2 creation wizard.
4. Select instance type, t2.micro.
5. Expand advanced. Select IAM profile.
6. Enter the User Data below.
7. Select the web server’s SG. Created in the task 1.
8. Select any key pair. It doesn’t matter. Because we use Session Manager to SSH into the instance.



Graphical user interface, application

Description automatically generated

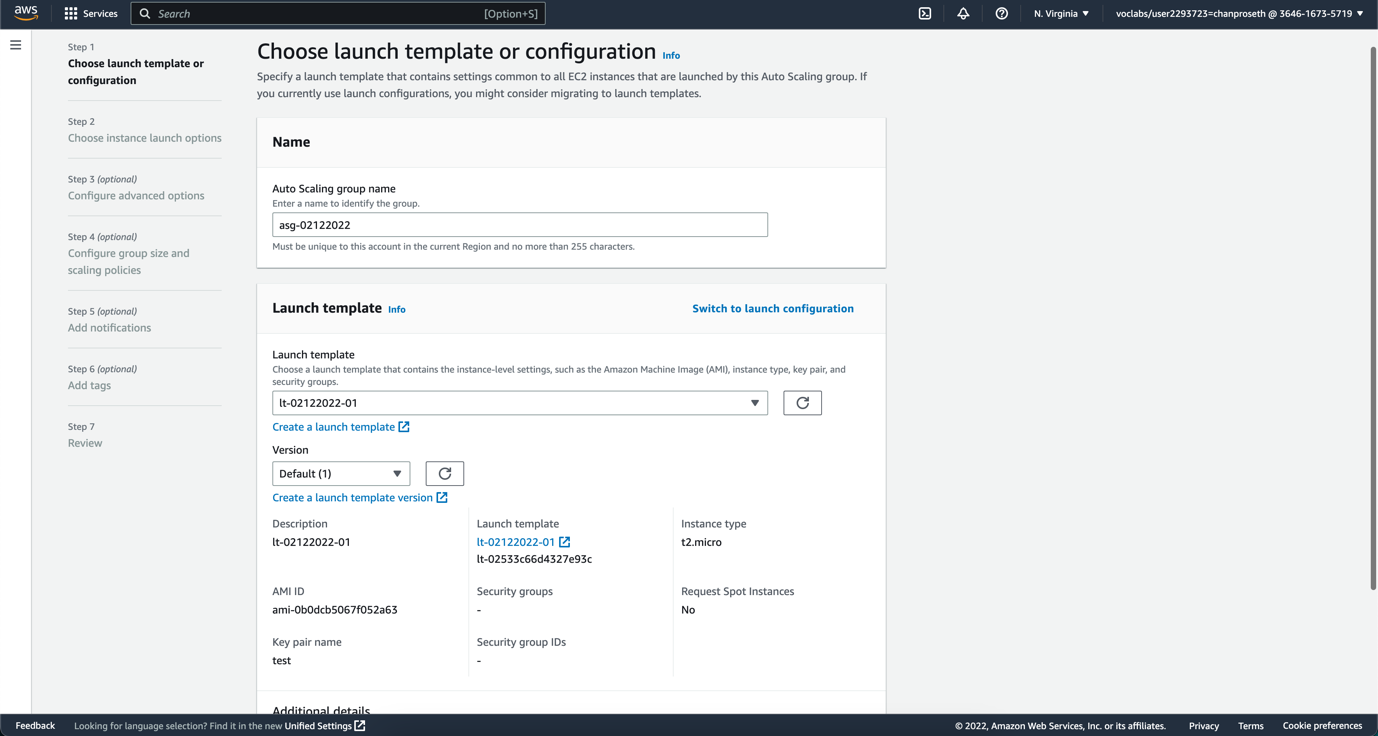


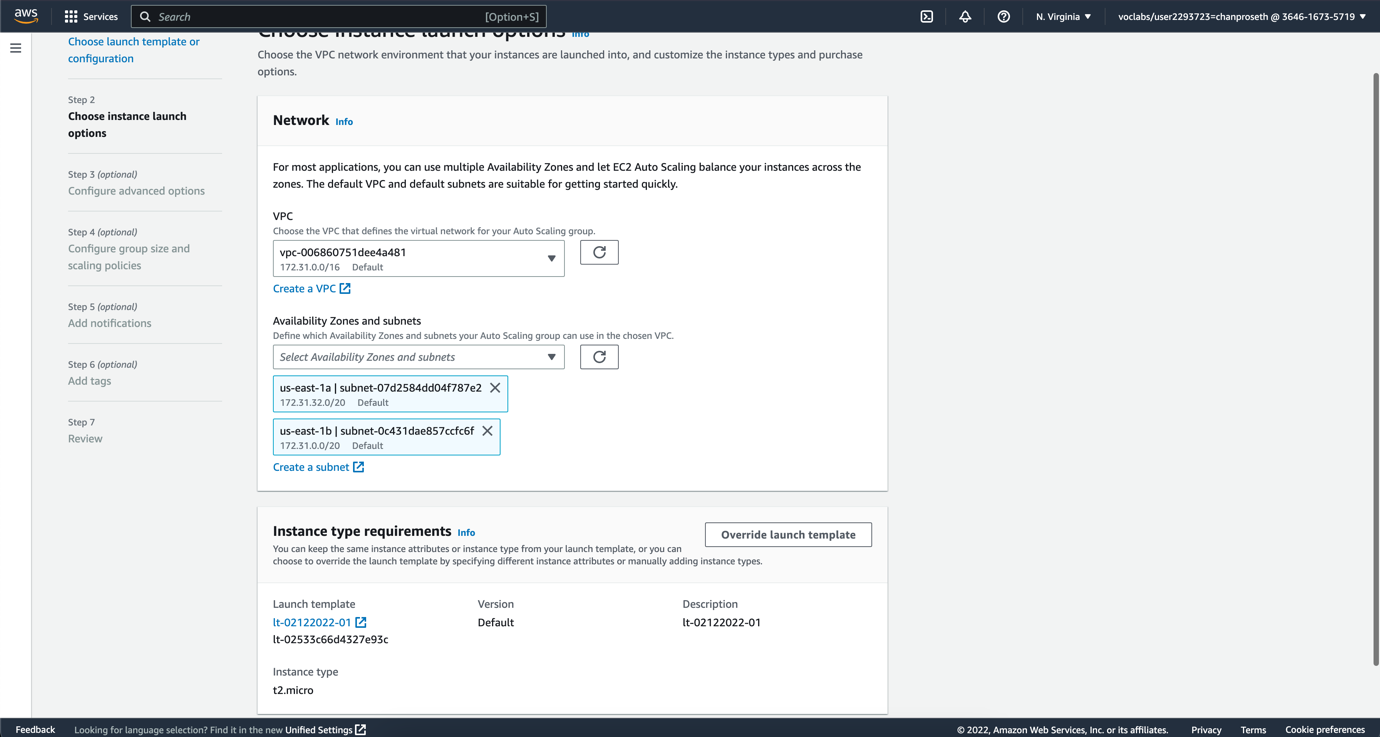
Graphical user interface, application

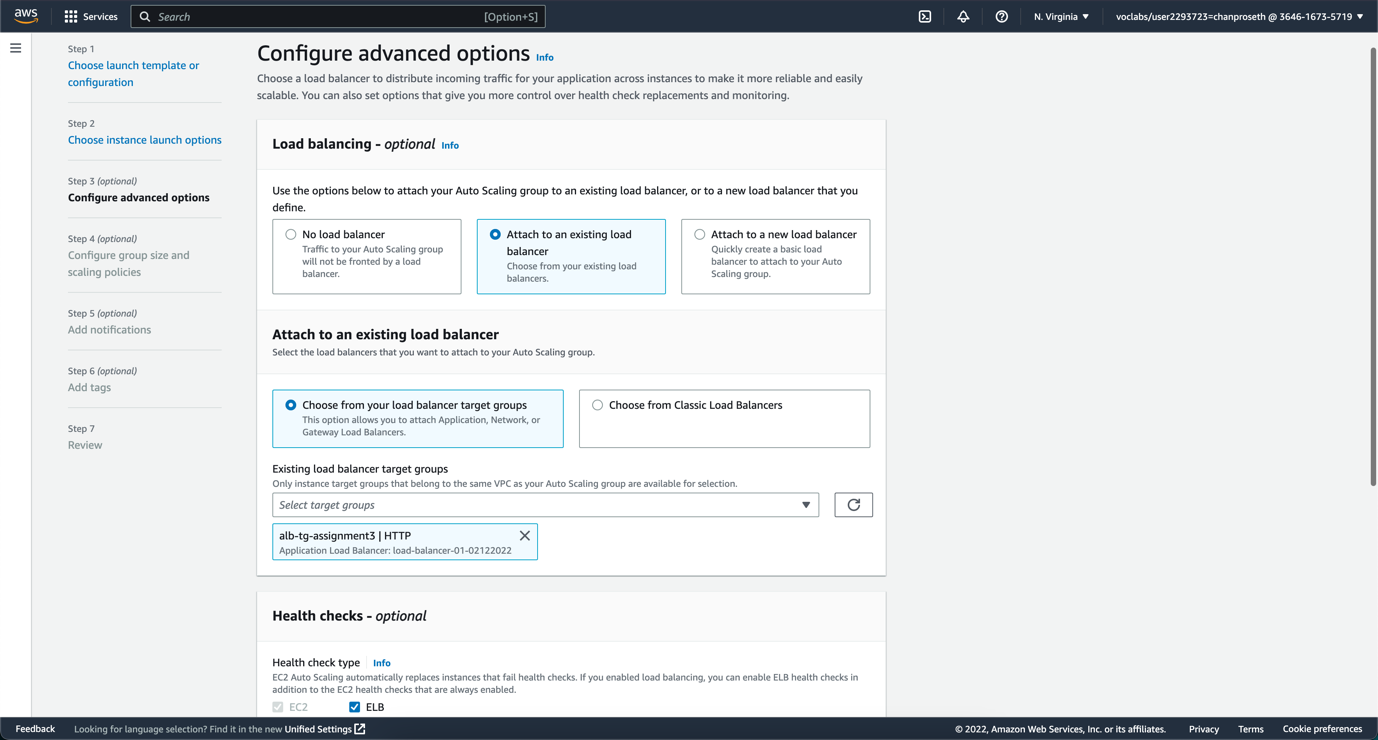
Description automatically generated

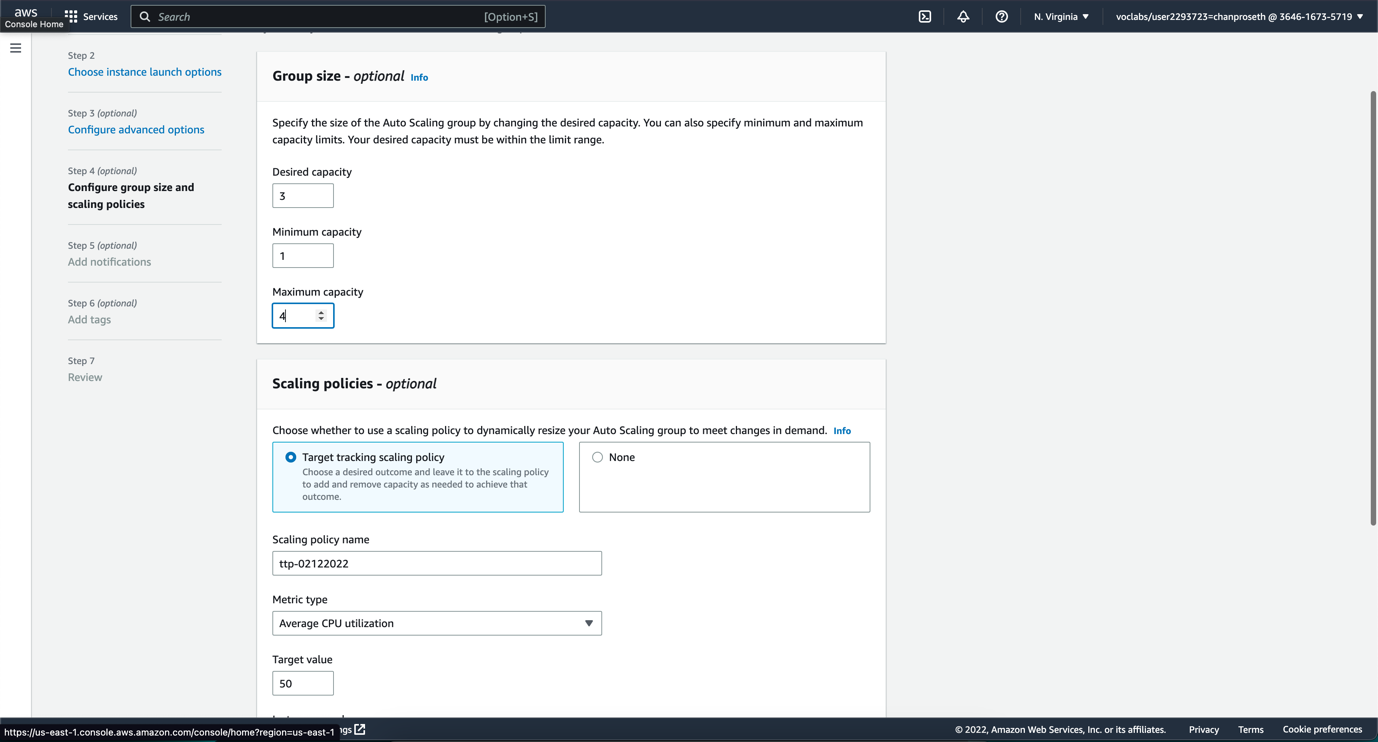
c. Create the Auto Scaling Group.

1. Select launch template/configration.
2. Select AZs (Subnets). That is where your instances launched.
3. Click on attach to an existing load balancer and select the default TG of the ALB.
4. Select ELB in health checks panel.
5. Set desired, min, and max capacity. Set a target tracking scale with a target value of 50.









Graphical user interface, application

Description automatically generated

d. Mimic

the high CPU utilization with a library to test scaling out.

Extra:  
• Establish a secure connection between EC2 and ELB.