

EXPERIMENT: 7

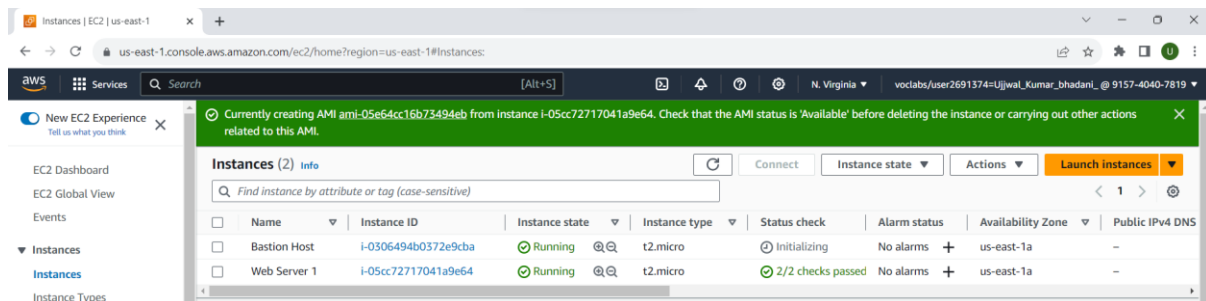
AIM: Explore Scaling and Load Balancing in Cloud Architecture

Elastic Load Balancing automatically distributes incoming application traffic across multiple Amazon EC2 instances. It enables you to achieve fault tolerance in your applications by seamlessly providing the required amount of load balancing capacity needed to route application traffic.

Auto Scaling helps you maintain application availability and allows you to scale your Amazon EC2 capacity out or in automatically according to conditions you define. You can use Auto Scaling to help ensure that you are running your desired number of Amazon EC2 instances. Auto Scaling can also automatically increase the number of Amazon EC2 instances during demand spikes to maintain performance and decrease capacity during lulls to reduce costs.

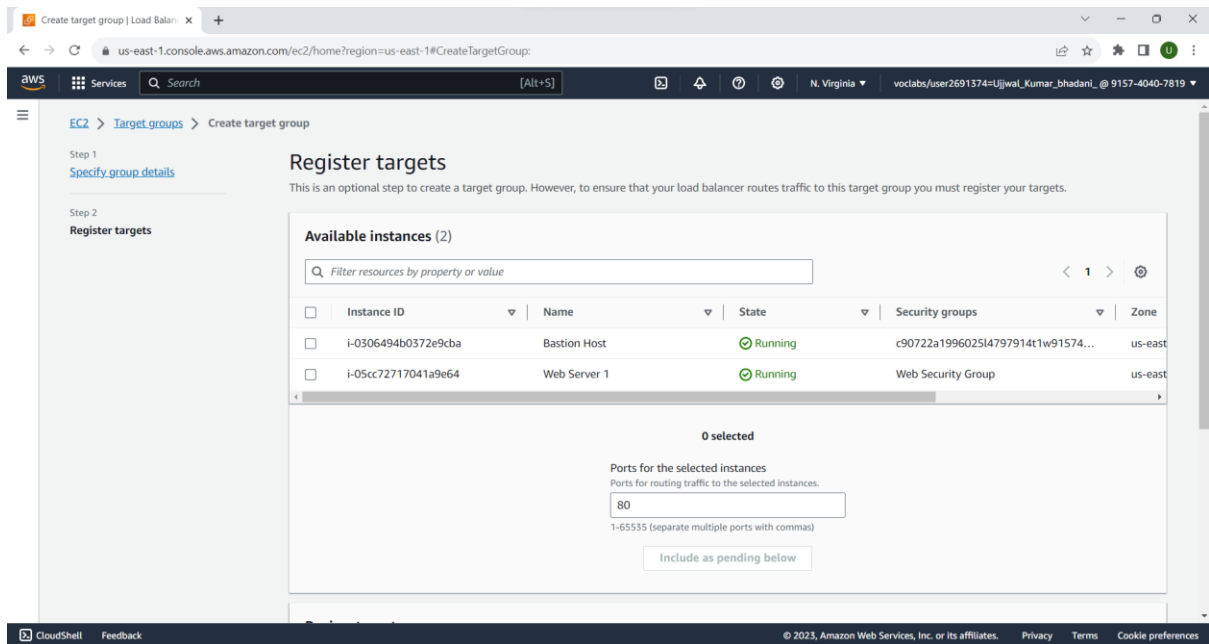
Task 1: Create an AMI for Auto Scaling

1. In the AWS Management Console, in the search box next to **Services**, search for and select EC2.
2. In the left navigation pane, choose Instances.
3. Wait until the Status Checks for Web Server 1 displays *2/2 checks passed*.
4. Select Web Server 1.
5. In the **Actions** menu, choose Image and templates > Create image, then configure:
 - Image name: **WebServerAMI**
 - Image description: **Lab AMI for Web Server**
6. Choose **Create image**
 - A confirmation banner displays the AMI ID for your new AMI.

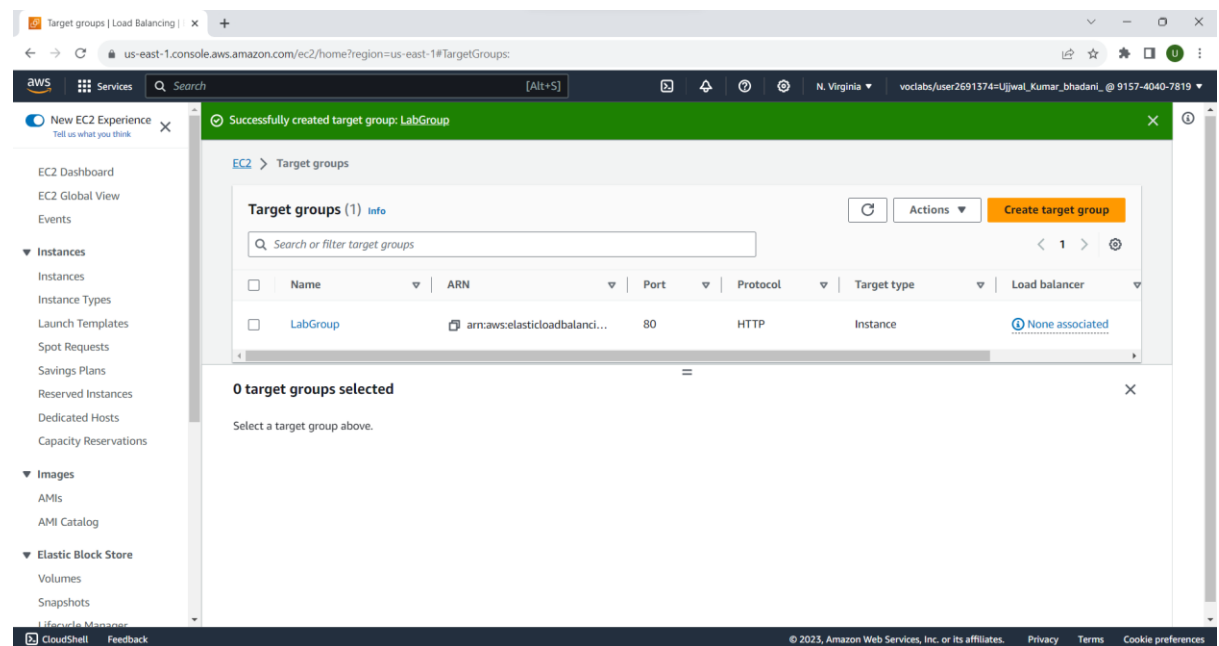


Task 2: Create a Load Balancer

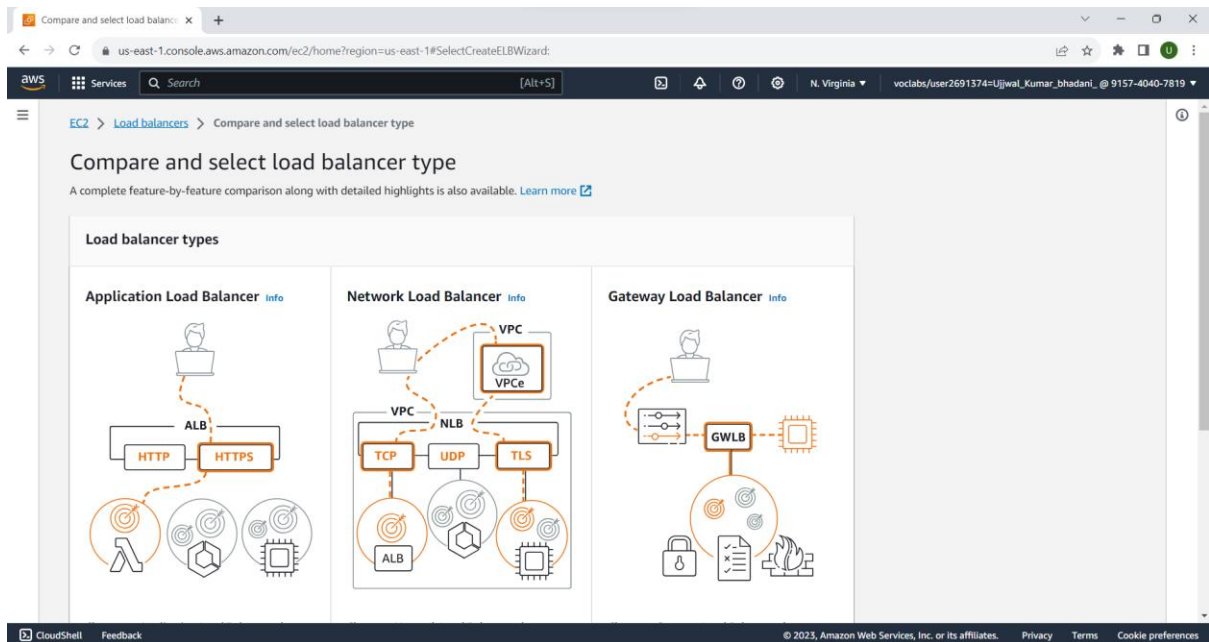
7. In the left navigation pane, choose Target Groups.
 - a. Choose **Create target group**
 - b. Choose a target type: Instances
 - c. Target group name, enter: **LabGroup**
 - d. Select Lab VPC from the VPC drop-down menu.



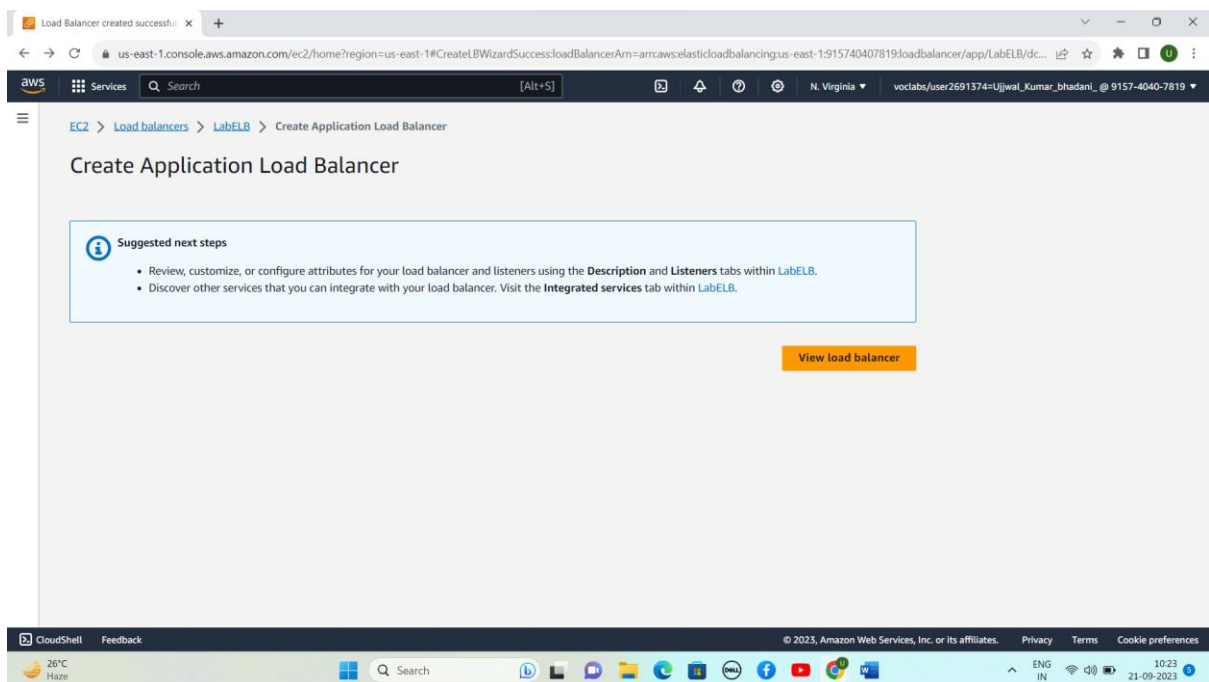
8. Choose **Next**. The Register targets screen appears.
9. Review the settings and choose **Create target group**



10. In the left navigation pane, choose Load Balancers.
11. At the top of the screen, choose **Create load balancer**.
12. Under Application Load Balancer, choose **Create**

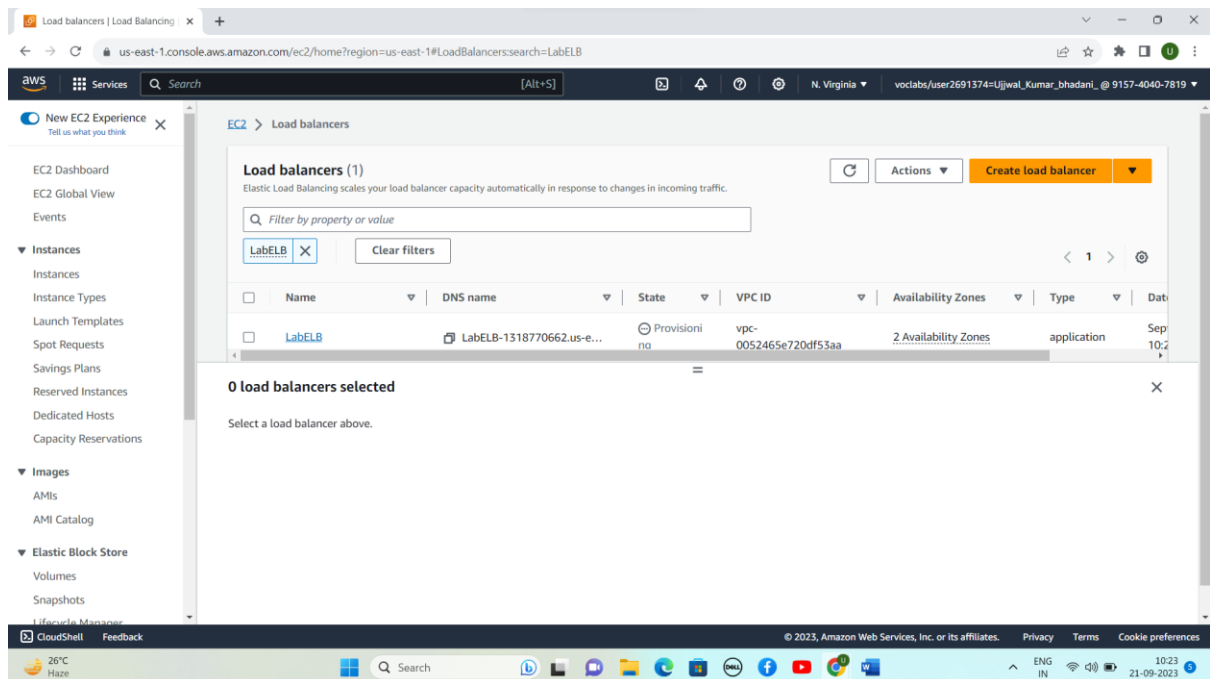


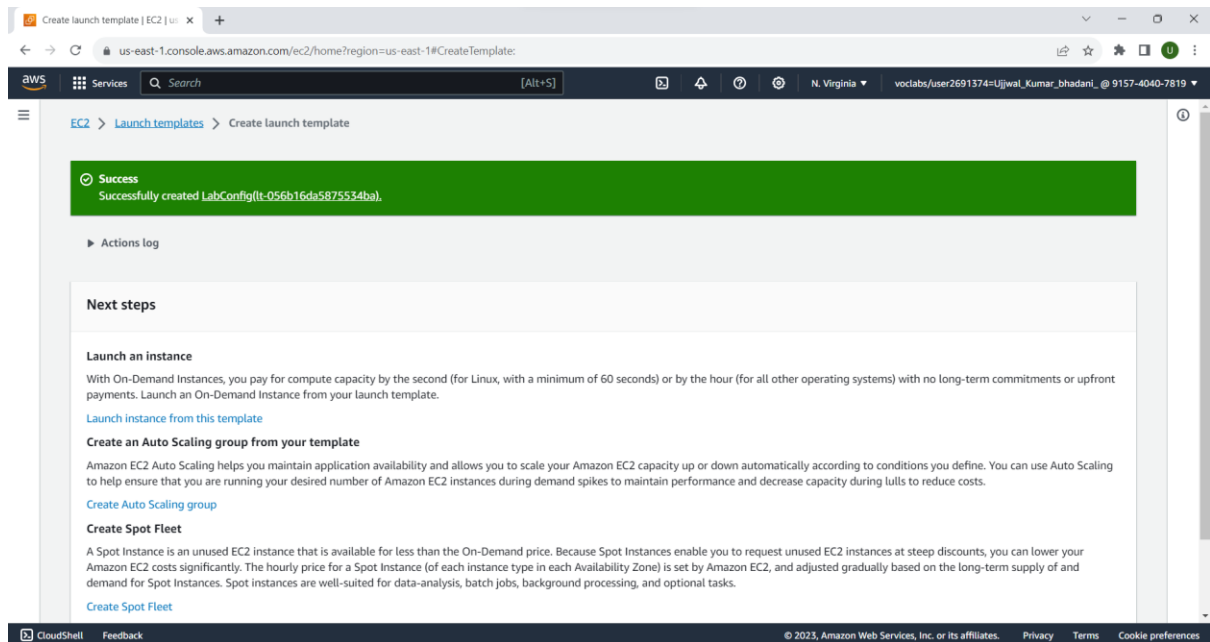
13. Under Load balancer name, enter: **LabELB**
14. Scroll down to the Network mapping section, then:
 - a. For VPC, choose Lab VPC
 - b. Choose the first displayed Availability Zone, then select Public Subnet 1 from the Subnet drop down menu that displays beneath it.
 - c. Choose the second displayed Availability Zone, then select Public Subnet 2 from the Subnet drop down menu that displays beneath it.
15. In the Security groups section:
 - a. Choose the Security groups drop down menu and select Web Security Group
 - b. Below the drop down menu, choose the X next to the default security group to remove it.
16. For the Listener HTTP:80 row, set the Default action to forward to LabGroup.
17. Scroll to the bottom and choose **Create load balancer**. The load balancer is successfully created.
 - a. Choose **View load balancer**



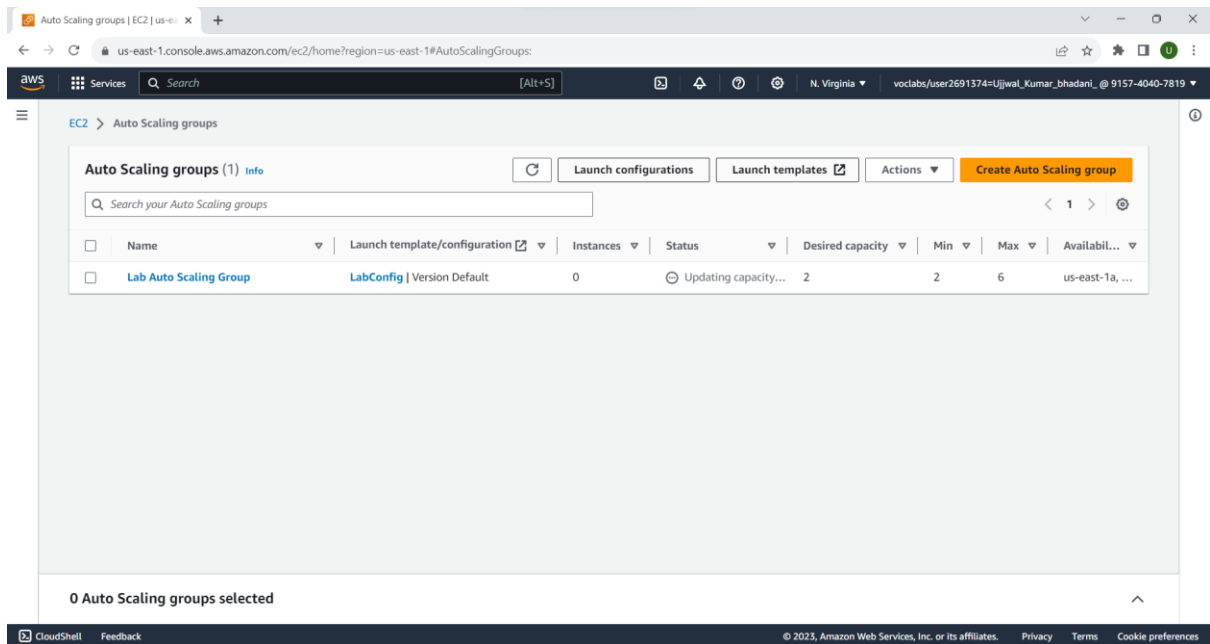
Task 3: Create a Launch Template and an Auto Scaling Group

18. In the left navigation pane, choose Launch Templates.
19. Choose **Create launch template**
20. Configure the launch template settings and create it:
 - a. Launch template name: **LabConfig**
 - b. Under Auto Scaling guidance, select *Provide guidance to help me set up a template that I can use with EC2 Auto Scaling*
 - c. In the Application and OS Images (Amazon Machine Image) area, choose *My AMIs*.
 - d. Amazon Machine Image (AMI): choose *Web Server AMI*
 - e. Instance type: choose *t2.micro*
 - f. Key pair name: choose *vockey*
 - g. Firewall (security groups): choose *Select existing security group*
 - h. Security groups: choose *Web Security Group*
 - i. Scroll down to the Advanced details area and expand it.
 - j. Scroll down to the Detailed CloudWatch monitoring setting. Select *Enable*
 - k. Choose **Create launch template**



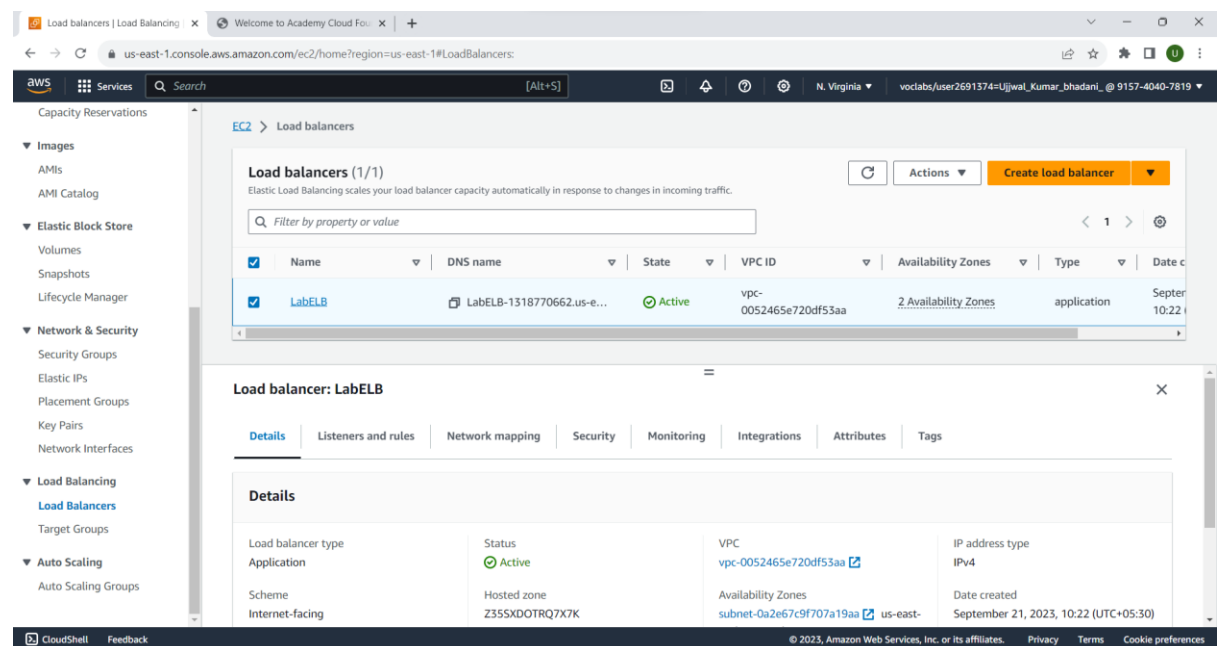


21. In the Success dialog, choose the LabConfig launch template.
22. From the **Actions** menu, choose *Create Auto Scaling group*
23. Configure the details in Step 1 (Choose launch template or configuration):
 - a. Auto Scaling group name: **Lab Auto Scaling Group**
 - b. Launch template: confirm that the *LabConfig* template you just created is selected.
 - c. Choose **Next**
24. Configure the details in Step 2 (Choose instance launch options):
 - a. VPC: choose *Lab VPC*
 - b. Availability Zones and subnets: Choose *Private Subnet 1* and then choose *Private Subnet 2*. Choose **Next**
25. Configure the details in Step 3 (Configure advanced options):
 - a. Choose Attach to an existing load balancer
 - i. Existing load balancer target groups: select *LabGroup*.
 - b. In the Additional settings pane:
 - i. Select Enable group metrics collection within CloudWatch Choose **Next**
26. Configure the details in Step 4
 - a. Under Group size, configure:
 - i. Desired capacity: 2
 - ii. Minimum capacity: 2
 - iii. Maximum capacity: 6
 - b. Under Scaling policies, choose *Target tracking scaling policy* and configure:
 - i. Scaling policy name: **LabScalingPolicy**
 - ii. Metric type: *Average CPU Utilization*
 - iii. Target value: **60**
27. Configure the details in Step 5 (Add notifications - optional):
 - a. Auto Scaling can send a notification when a scaling event takes place. You will use the default settings. Choose **Next**
 - b.
28. Configure the details in Step 6 (Add tags - optional):
 - a. Tags applied to the Auto Scaling group will be automatically propagated to the instances that are launched. Choose **Add tag** and Configure the following: Key: **Name** , Value: **Lab Instance** , Choose **Next**
29. Configure the details in Step 6 (Review):
 - a. Review the details of your Auto Scaling group Choose **Create Auto Scaling group**.

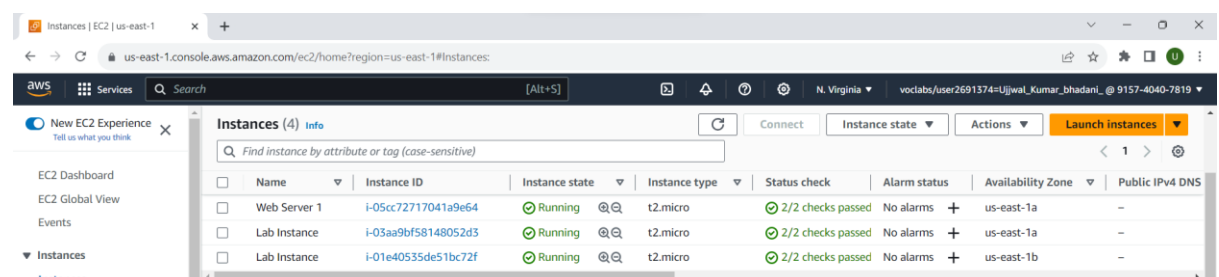


Task 4: Verify that Load Balancing is Working

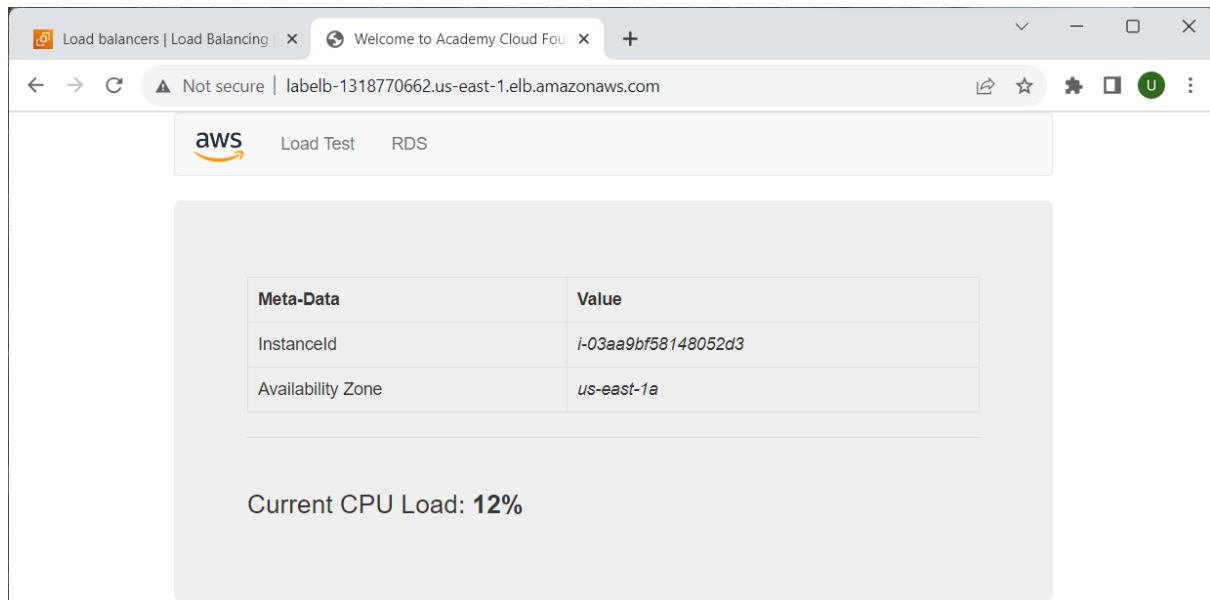
30. In this task, you will verify that Load Balancing is working correctly.



31. In the left navigation pane, choose Instances.

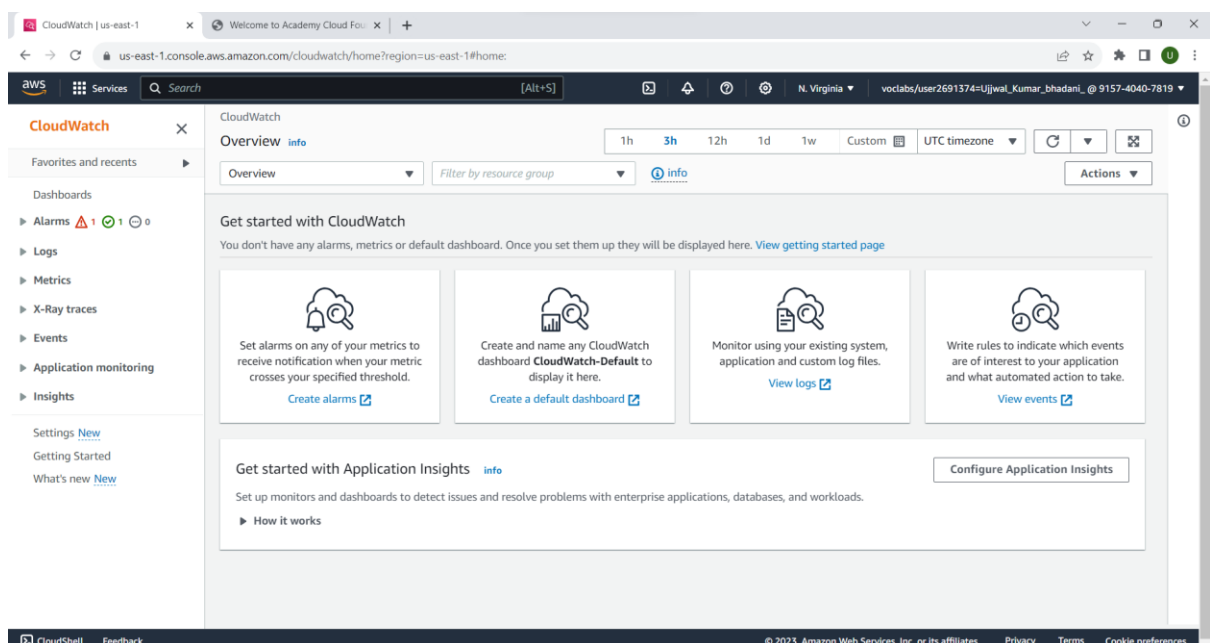


32. In the left navigation pane, choose Target Groups. Select *LabGroup*
33. Choose the Targets tab.
34. Wait until the Status of both instances transitions to *healthy*.
35. In the left navigation pane, choose Load Balancers.
36. Select the *LabELB* load balancer.
37. In the Details pane, copy the DNS name of the load balancer, making sure to omit "(A Record)".
38. Open a new web browser tab, paste the DNS Name you just copied, and press Enter.



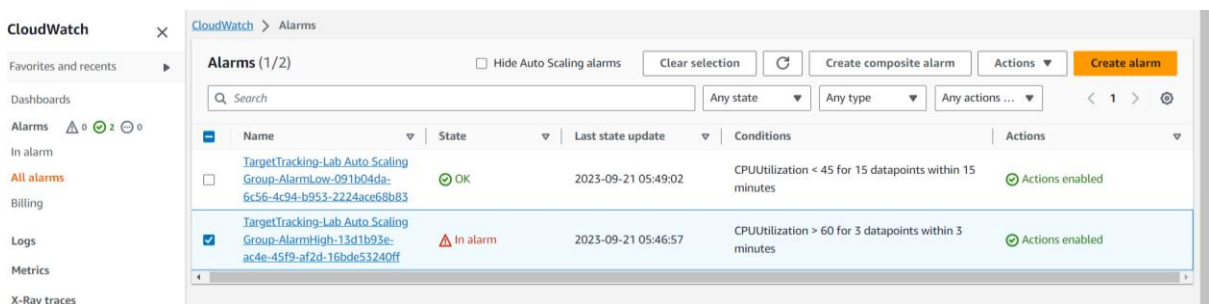
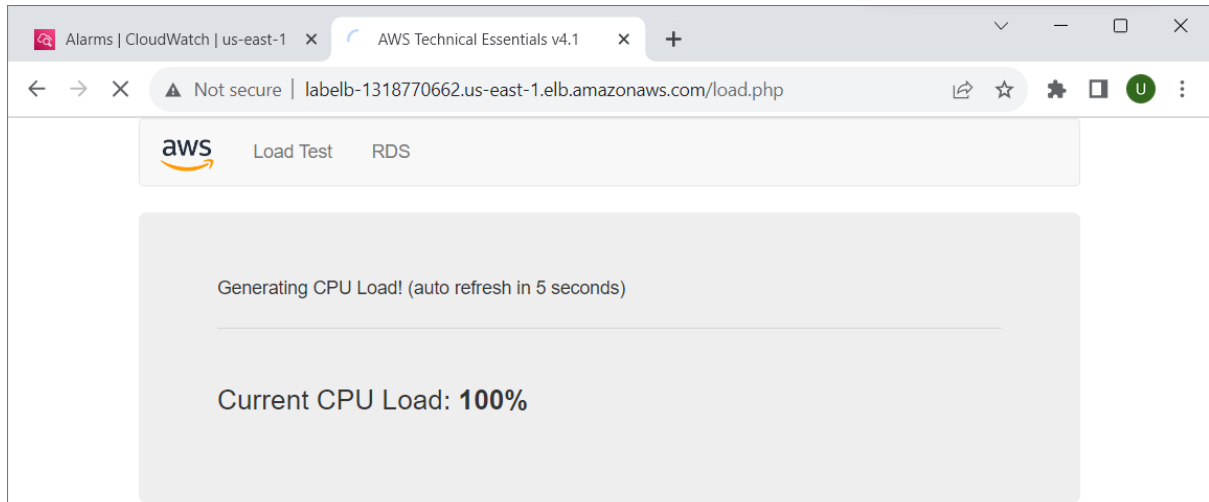
Task 5: Test Auto Scaling

39. Return to the AWS Management Console, but do not close the application tab — you will return to it soon.
40. In the search box next to **Services**, search for and select CloudWatch.

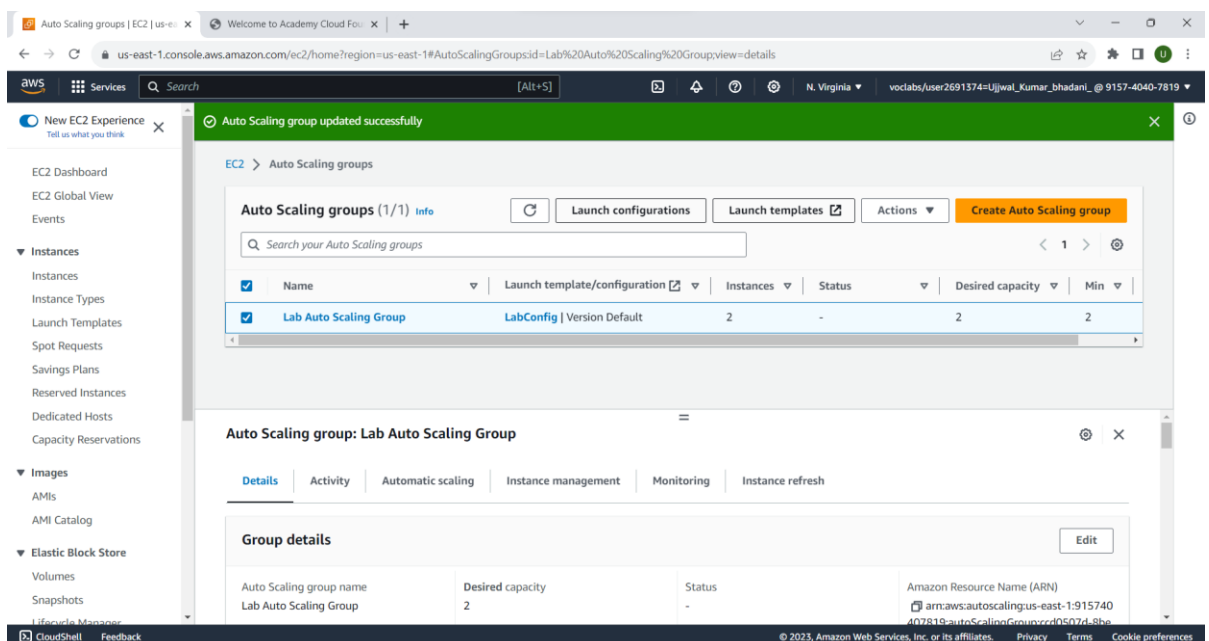


41. In the left navigation pane, choose All alarms.
42. Choose the OK alarm, which has *AlarmHigh* in its name.

43. Return to the browser tab with the web application.
44. Choose Load Test beside the AWS logo.
45. Return to browser tab with the CloudWatch console.
 - a. In less than 5 minutes, the AlarmLow alarm should change to OK and the AlarmHigh alarm status should change to *In alarm*.
46. Wait until the AlarmHigh alarm enters the *In alarm* state.



47. In the search box next to **Services**, search for and select EC2.
48. In the left navigation pane, choose Instances.



| | Name | Instance ID | Instance state | Instance type | Status check | Alarm status | Availability Zone | Public IPv4 DNS |
|--------------------------|--------------|---------------------|----------------|---------------|-------------------|--------------|-------------------|-----------------|
| <input type="checkbox"/> | Bastion Host | i-085822a9a4e9f33a5 | Running | t2.micro | 2/2 checks passed | No alarms | us-east-1a | - |
| <input type="checkbox"/> | Lab Instance | i-02933b7b93fc31e35 | Running | t2.micro | 2/2 checks passed | No alarms | us-east-1a | - |
| <input type="checkbox"/> | Lab Instance | i-0da66ec2e49d690de | Running | t2.micro | Initializing | No alarms | us-east-1a | - |
| <input type="checkbox"/> | Web Server 1 | i-083eefad37208766a | Running | t2.micro | 2/2 checks passed | No alarms | us-east-1a | - |
| <input type="checkbox"/> | Lab Instance | i-0783fbd1f6e48962e | Running | t2.micro | 2/2 checks passed | No alarms | us-east-1b | - |
| <input type="checkbox"/> | Lab Instance | i-0e5daa35e523efc49 | Running | t2.micro | 2/2 checks passed | No alarms | us-east-1b | - |

Task 6: Terminate Web Server 1

49. Select Web Server 1 (and ensure it is the only instance selected).
50. In the Instance state menu, choose Instance State > Terminate Instance.
51. Choose **Terminate**

| | Name | Instance ID | Instance state | Instance type | Status check | Alarm status | Availability Zone | Public IPv4 DNS |
|--------------------------|--------------|---------------------|----------------|---------------|--------------|--------------|-------------------|-----------------|
| <input type="checkbox"/> | Lab Instance | i-09092116b713bbace | Terminated | t2.micro | - | No alarms | us-east-1b | - |
| <input type="checkbox"/> | Lab Instance | i-0c61704f49901a7f2 | Terminated | t2.micro | - | No alarms | us-east-1b | - |
| <input type="checkbox"/> | Lab Instance | i-05bd436d732af6154 | Terminated | t2.micro | - | No alarms | us-east-1a | - |
| <input type="checkbox"/> | Lab Instance | i-0c59b1ec671d85cbf | Terminated | t2.micro | - | No alarms | us-east-1a | - |
| <input type="checkbox"/> | Web Server 1 | i-05cc72717041a9e64 | Shutting-down | t2.micro | - | No alarms | us-east-1a | - |

Submitted by

| | |
|-----------|----------------------|
| NAME : | Ujjwal Kumar Bhadani |
| ROLL NO: | CIT21046 |
| REG NO: | 2101020286 |
| GROUP: | 4 |
| SEMESTER: | 5th |