

# Lab 06

## CLOUD COMPUTING

CS-4075

Course Instructor:

Sir Zaheer Sani

**Name:** Zeeshan Ali

**Roll No:** 20i-2465

**Section:** SE-A

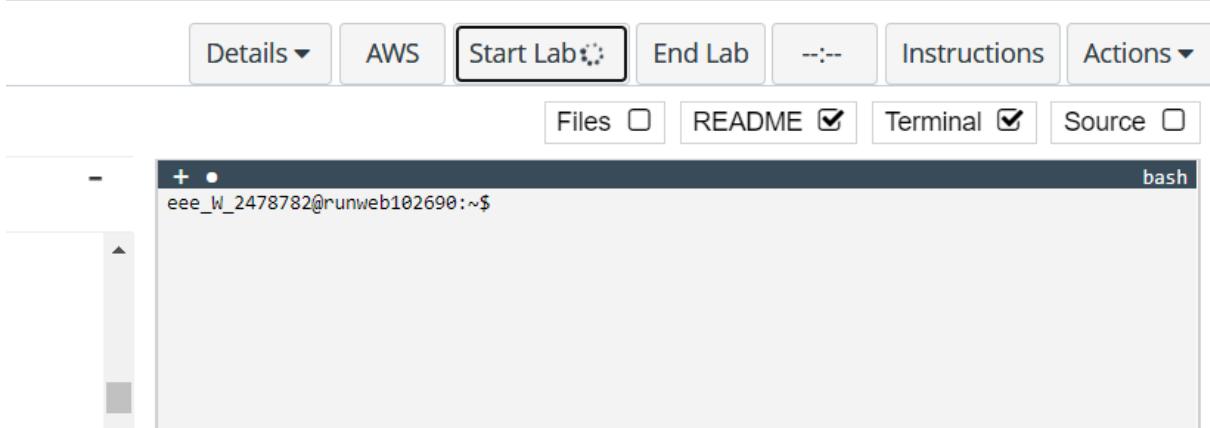
**Due Date:** 00, 2023



# Lab 6: Scale and Load Balance Your Architecture

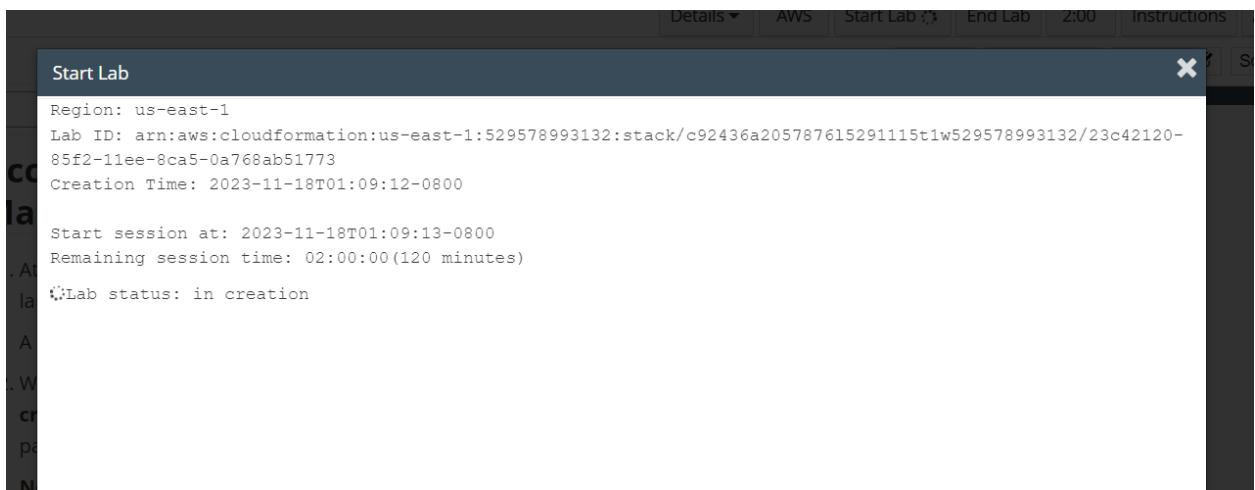
## Accessing the AWS Management Console

- At the top of these instructions, choose **Start Lab** to launch your lab.

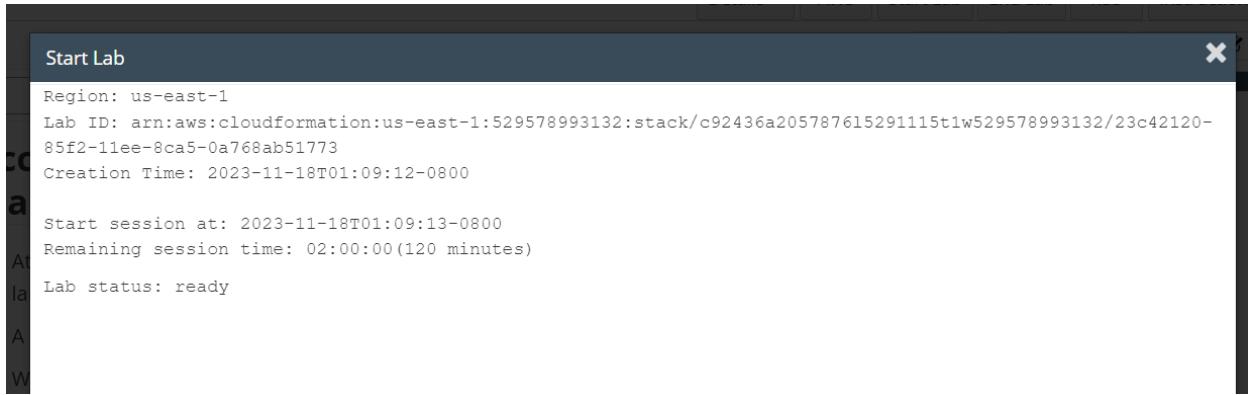


A Start Lab panel opens displaying the lab status.

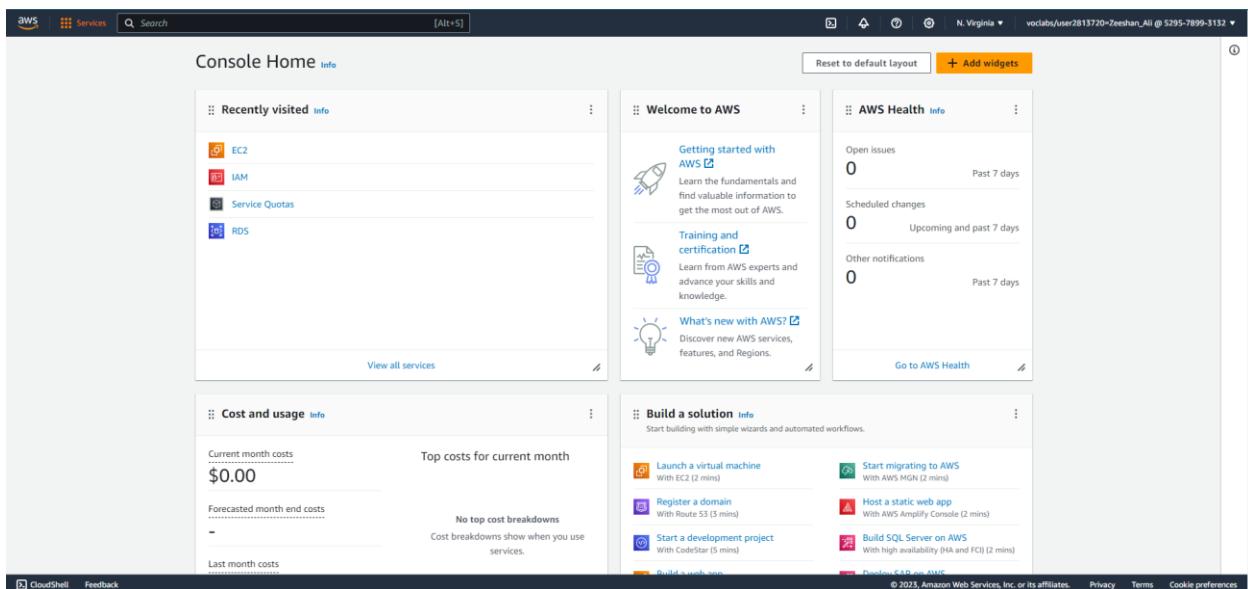
- Wait until you see the message "**Lab status: in creation**", then choose the **X** to close the Start Lab panel.



**Note:** It may take approximately 10 minutes or longer for the lab status to change to ready.



- At the top of these instructions, choose **AWS**



## Task 1: Create an AMI for Auto Scaling

In this task, you will create an AMI from the existing *Web Server 1*. This will save the contents of the boot disk so that new instances can be launched with identical content.

- In the **AWS Management Console**, in the search box next to **Services**, search for and select **EC2**.

The screenshot shows the AWS EC2 Dashboard. The left navigation pane includes options like EC2 Dashboard, EC2 Global View, Events, Instances (with sub-options: Instances, Instance Types, Launch Templates, Spot Requests, Savings Plans, Reserved Instances, Dedicated Hosts, Capacity Reservations), Images (AMIs, AMI Catalog), Elastic Block Store (Volumes, Snapshots, Lifecycle Manager), Network & Security (Security Groups, Elastic IPs, Placement Groups), and more. The main content area displays 'Resources' (Instances running: 2, Auto Scaling Groups: 0, Dedicated Hosts: 0; Elastic IPs: 1, Instances: 2, Key pairs: 1; Load balancers: 0, Placement groups: 0, Security groups: 5; Snapshots: 0, Volumes: 2) and 'Service health' (Region: US East (N. Virginia)). The right sidebar shows 'Account attributes' (Default VPC: vpc-0621d4e255bd5e34, Settings, Data protection and security, Zones, EC2 Serial Console, Default credit specification, Console experiments) and 'Additional information' (Getting started guide, Documentation, All EC2 resources, Forums, Pricing, Contact us).

6. In the left navigation pane, choose **Instances**. First, you will confirm that the instance is running.
7. Wait until the **Status Checks** for **Web Server 1** displays *2/2 checks passed*. If necessary, choose to refresh to update the status. You will now create an AMI based upon this instance.

The screenshot shows the 'Instances (2) Info' page. The left navigation pane is identical to the previous dashboard. The main table lists two instances:

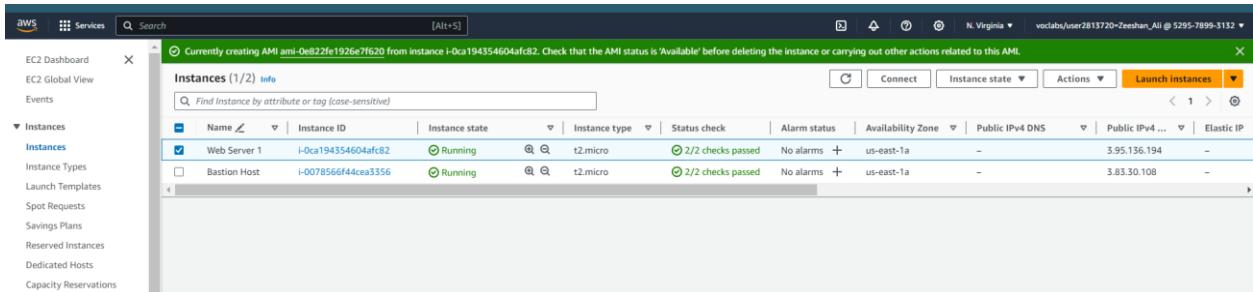
Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public
Web Server 1	i-0ca194354604afc82	Running	t2.micro	2/2 checks passed	No alarms	us-east-1a	-
Bastion Host	i-0078566f44cea3356	Running	t2.micro	2/2 checks passed	No alarms	us-east-1a	-

8. Select **Web Server 1**.
9. In the **Actions** menu, choose **Image and templates > Create image**, then configure:

- **Image name:** WebServerAMI
- **Image description:** Lab AMI for Web Server

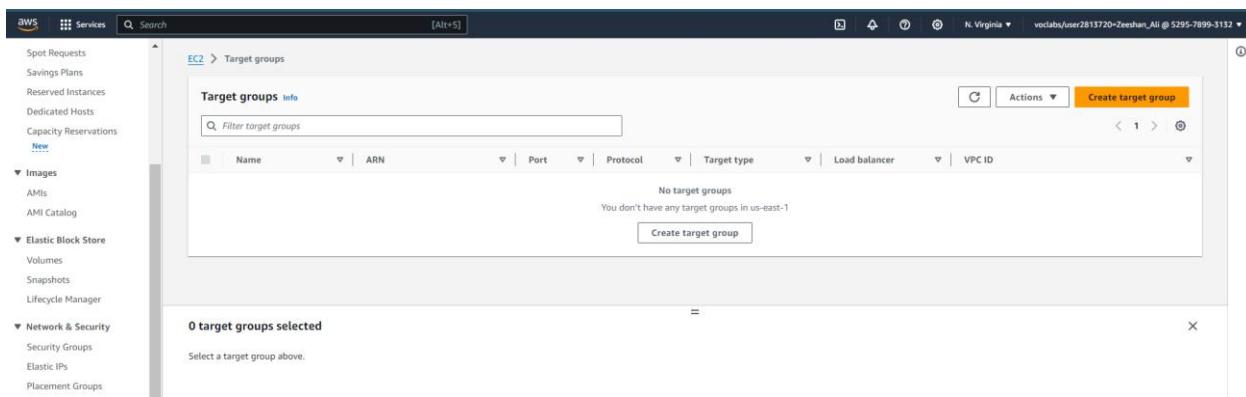
10. Choose **Create image**. A confirmation banner displays the **AMI ID** for your new AMI.

<https://us-east-1.console.aws.amazon.com/ec2/v2/home?region=us-east-1#Images:visibility=owned-by-me;imageId=ami-0e822fe1926e7f620>



## Task 2: Create a Load Balancer

11. In the left navigation pane, choose **Target Groups**.



- Choose **Create target group**
- Choose a target type: **Instances**
- **Target group name**, enter: **LabGroup**
- Select **Lab VPC** from the **VPC** drop-down menu.

aws Services Search [Alt+S]

EC2 > Target groups > Create target group

Step 1 Specify group details

Step 2 Register targets

## Specify group details

Your load balancer routes requests to the targets in a target group and performs health checks on the targets.

### Basic configuration

Settings in this section can't be changed after the target group is created.

Choose a target type

Instances

- Supports load balancing to instances within a specific VPC.
- Facilitates the use of [Amazon EC2 Auto Scaling](#) to manage and scale your EC2 capacity.

IP addresses

- Supports load balancing to VPC and on-premises resources.
- Facilitates routing to multiple IP addresses and network interfaces on the same instance.
- Offers flexibility with microservice based architectures, simplifying inter-application communication.
- Supports IPv6 targets, enabling end-to-end IPv6 communication, and IPv4-to-IPv6 NAT.

Lambda function

- Facilitates routing to a single Lambda function.
- Accessible to Application Load Balancers only.

Application Load Balancer

- Offers the flexibility for a Network Load Balancer to accept and route TCP requests within a specific VPC.
- Facilitates using static IP addresses and PrivateLink with an Application Load Balancer.

aws Services Search [Alt+S]

Target group name  
LabGroup

A maximum of 32 alphanumeric characters including hyphens are allowed, but the name must not begin or end with a hyphen.

Protocol : Port  
HTTP 80 1-65535

IP address type  
Only targets with the indicated IP address type can be registered to this target group.

IPv4  
Each instance has a default network interface (eth0) that is assigned the primary private IPv4 address. The instance's primary private IPv4 address is the one that will be applied to the target.

IPv6  
Each instance you register must have an assigned primary IPv6 address. This is configured on the instance's default network interface (eth0). [Learn more](#)

VPC  
Select the VPC with the instances that you want to include in the target group. Only VPCs that support the IP address type selected above are available in this list.  
Lab VPC  
vpc-0c6453b6cdef0ac02  
IPv4: 10.0.0.0/16

Protocol version  
 HTTP1  
Send requests to targets using HTTP/1.1. Supported when the request protocol is HTTP/1.1 or HTTP/2.

HTTP2  
Send requests to targets using HTTP/2. Supported when the request protocol is HTTP/2 or gRPC, but gRPC-specific features are not available.

gRPC  
Send requests to targets using gRPC. Supported when the request protocol is gRPC.

12. Choose **Next**. The **Register targets** screen appears.

EC2 > Target groups > Create target group [Alt+S]

Step 1  
Specify group details

Step 2  
Register targets

Register targets

This is an optional step to create a target group. However, to ensure that your load balancer routes traffic to this target group you must register your targets.

Available instances (2)

Instance ID	Name	State	Security groups	Zone	Private IPv4 address
i-0078566f44cea3356	Bastion Host	Running	c92436a20578761529115t1w52957...	us-east-1a	10.0.0.130
i-0ca194354604afc82	Web Server 1	Running	Web Security Group	us-east-1a	10.0.0.43

0 selected

Ports for the selected instances  
Ports for routing traffic to the selected instances.  
80

13. Review the settings and choose **Create target group**

The screenshot shows the AWS Elastic Load Balancing Target Groups page. A green banner at the top indicates "Successfully created the target group: LabGroup." The main section displays the "LabGroup" target group details. The "Details" table includes columns for Target type (Instance), Protocol (HTTP: 80), Protocol version (HTTP1), VPC (vpc-0c6453b6cd0ac02), and various metrics like Total targets (0), Healthy (0), Unhealthy (0), Unused (0), Initial (0), and Draining (0). Below the table is a navigation bar with tabs: Targets (selected), Monitoring, Health checks, Attributes, and Tags. Under the "Registered targets (0)" heading, there is a table with columns: Instance ID, Name, Port, Zone, Health status, Health status details, and Launch time. A message states "No registered targets" and "You have not registered targets to this group yet." At the bottom right of the main content area, there are links for "Deregister" and "Register targets". The left sidebar lists various AWS services and features.

14. In the left navigation pane, choose **Load Balancers**.

The screenshot shows the AWS Elastic Load Balancing Load balancers page. A message at the top states "Elastic Load Balancing scales your load balancer capacity automatically in response to changes in incoming traffic." Below this is a table with columns: Name, DNS name, State, VPC ID, Availability Zones, Type, and Date created. A message in the center says "No load balancers" and "You don't have any load balancers in us-east-1". At the bottom right of the main content area, there is a "Create load balancer" button. The left sidebar lists various AWS services and features.

15. At the top of the screen, choose **Create load balancer**.

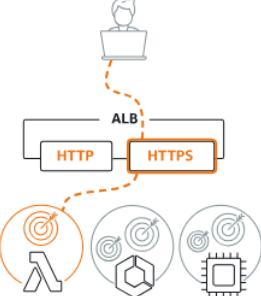
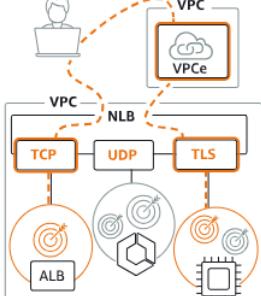
16. Under **Application Load Balancer**, choose **Create**

AWS Services Search [Alt+S]

## Compare and select load balancer type

A complete feature-by-feature comparison along with detailed highlights is also available. [Learn more](#)

### Load balancer types

Application Load Balancer	Network Load Balancer	Gateway Load Balancer
 <p>Info</p> <p>Choose an Application Load Balancer when you need a flexible feature set for your applications with HTTP and HTTPS traffic. Operating at the request level, Application Load Balancers provide advanced routing and visibility features targeted at application architectures, including microservices and containers.</p> <p><a href="#">Create</a></p>	 <p>Info</p> <p>Choose a Network Load Balancer when you need ultra-high performance, TLS offloading at scale, centralized certificate deployment, support for UDP, and static IP addresses for your applications. Operating at the connection level, Network Load Balancers are capable of handling millions of requests per second securely while maintaining ultra-low latencies.</p>	 <p>Info</p> <p>Choose a Gateway Load Balancer when you need to deploy and manage a fleet of third-party virtual appliances that support GENEVE. These appliances enable you to improve security, compliance, and policy controls.</p> <p><a href="#">Create</a></p>

CloudShell Feedback

17. Under **Load balancer name**, enter: LabELB

The screenshot shows the AWS CloudFormation console interface. The top navigation bar includes the AWS logo, a services menu, a search bar, and a keyboard shortcut [Alt+S]. The main content area is titled "Create Stack" and shows the "Template" tab selected. The template URL is set to a standard CloudFormation template. The stack name is "LabVPC". The region is set to "us-east-1". There are no tags or outputs defined yet.

18. Scroll down to the **Network mapping** section, then:

- o For **VPC**, choose **Lab VPC**

You will now specify which *subnets* the Load Balancer should use. The load balancer will be internet facing, so you will select both Public Subnets.

- o Choose the **first** displayed Availability Zone, then select **Public Subnet 1** from the Subnet drop down menu that displays beneath it.
- o Choose the **second** displayed Availability Zone, then select **Public Subnet 2** from the Subnet drop down menu that displays beneath it.

You should now have two subnets selected: **Public Subnet 1** and **Public Subnet 2**.

**Network mapping** Info

The load balancer routes traffic to targets in the selected subnets, and in accordance with your IP address settings.

**VPC** Info

Select the virtual private cloud (VPC) for your targets or you can [create a new VPC](#). Only VPCs with an internet gateway are enabled for selection. The selected VPC can't be changed after the load balancer is created. To confirm the VPC for your targets, view your [target groups](#).

**Lab VPC**  
vpc-0c6453b6cdef0ac02  
IPv4: 10.0.0.0/16

**Mappings** Info

Select at least two Availability Zones and one subnet per zone. The load balancer routes traffic to targets in these Availability Zones only. Availability Zones that are not supported by the load balancer or the VPC are not available for selection.

**us-east-1a (use1-az2)**

Subnet

subnet-0f95eb7df3a4a6b9a Public Subnet 1 ▾

IPv4 address  
Assigned by AWS

**us-east-1b (use1-az4)**

Subnet

subnet-05a5be69d5e77436d Public Subnet 2 ▾

IPv4 address  
Assigned by AWS

19. In the **Security groups** section:

- Choose the Security groups drop down menu and select **Web Security Group**
- Below the drop down menu, choose the **X** next to the default security group to remove it.

The **Web Security Group** security group should now be the only one that appears.

20. For the Listener HTTP:80 row, set the Default action to forward to **LabGroup**.

21. Scroll to the bottom and choose **Create load balancer**

**Security groups** [Info](#)

A security group is a set of firewall rules that control the traffic to your load balancer. Select an existing security group, or you can create a new security group [Create](#).

Security groups

Select up to 5 security groups

Web Security Group sg-045d26a11321d71a0 VPC: vpc-0c6453b6cdef0ac02 [X](#) [C](#)

**Listeners and routing** [Info](#)

A listener is a process that checks for connection requests using the port and protocol you configure. The rules that you define for a listener determine how the load balancer routes requests to its registered targets.

▼ Listener HTTP:80 [Remove](#)

Protocol	Port	Default action	Info
HTTP	: 80 1-65535	Forward to	LabGroup Target type: Instance, IPv4 <a href="#">HTTP</a> <a href="#">C</a>

[Create target group](#) [C](#)

Listener tags - optional

► Load balancer tags - optional

Consider adding tags to your load balancer. Tags enable you to categorize your AWS resources so you can more easily manage them. The 'Key' is required, but 'Value' is optional. For example, you can have Key = production-webserver, or Key = webserver, and Value = production.

**Summary**

Review and confirm your configurations. [Estimate cost](#)

Basic configuration <a href="#">Edit</a>	Security groups <a href="#">Edit</a>	Network mapping <a href="#">Edit</a>	Listeners and routing <a href="#">Edit</a>
LabELB <ul style="list-style-type: none"> <li>Internet-facing</li> <li>IPv4</li> </ul>	• Web Security Group <a href="#">sg-045d26a11321d71a0</a> <a href="#">C</a>	VPC <a href="#">vpc-0c6453b6cdef0ac02</a> <a href="#">C</a> Lab VPC <ul style="list-style-type: none"> <li>us-east-1a <a href="#">subnet-0f95eb7df3a4a6b9a</a> <a href="#">C</a> Public Subnet 1</li> <li>us-east-1b <a href="#">subnet-05a5be69d5e77436d</a> <a href="#">C</a> Public Subnet 2</li> </ul>	• HTTP:80 defaults to <a href="#">LabGroup</a> <a href="#">C</a>
Add-on services <a href="#">Edit</a>		Tags <a href="#">Edit</a>	
None		None	
Attributes			
<p><a href="#">?</a> Certain default attributes will be applied to your load balancer. You can view and edit them after creating the load balancer.</p>			

[Cancel](#) [Create load balancer](#)

The load balancer is successfully created.

- Choose [View load balancer](#)

**Successfully created load balancer: LabELB**

Note: It might take a few minutes for your load balancer to be fully set up and ready to route traffic. Targets will also take a few minutes to complete the registration process and pass initial health checks.

EC2 > Load balancers > LabELB > Create Application Load Balancer

Create Application Load Balancer

**Suggested next steps**

- Review, customize, or configure attributes for your load balancer and listeners using the Description and Listeners tabs within LabELB.
- Discover other services that you can integrate with your load balancer. Visit the Integrated services tab within LabELB.

**View load balancer**

EC2 Dashboard

EC2 Global View

Events

Instances

Load balancers (1)

Load balancers (1)

Elastic Load Balancing scales your load balancer capacity automatically in response to changes in incoming traffic.

Filter load balancers

1 match

Name	DNS name	State	VPC ID	Availability Zones	Type	Date created
LabELB	LabELB-1639598004.us-east-1.elasticloadbalancing.amazonaws.com	Provisioning...	vpc-0c6453b6cdf0ac02	2 Availability Zones	application	November 18, 2023, 14:34 (UTC+05:00)

Actions

Create load balancer

0 load balancers selected

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

22. In the left navigation pane, choose **Launch Templates**.

EC2 launch templates

Streamline, simplify and standardize instance launches

Use launch templates to automate instance launches, simplify permission policies, and enforce best practices across your organization. Save launch parameters in a template that can be used for on-demand launches and with managed services, including EC2 Auto Scaling and EC2 Fleet. Easily update your launch parameters by creating a new launch template version.

New launch template

Create launch template

Benefits and features

Streamline provisioning

Minimize steps to provision instances. With EC2 Auto Scaling, updates to a launch template can be automatically passed to an Auto Scaling group. Learn more

Simplify permissions

Create shorter, easier to manage IAM policies. Learn more

Documentation

Documentation

API reference

23. Choose **Create launch template**

24. Configure the launch template settings and create it:

- **Launch template name:** LabConfig
- Under **Auto Scaling guidance**, select *Provide guidance to help me set up a template that I can use with EC2 Auto Scaling*

The screenshot shows the AWS EC2 'Create launch template' interface. At the top, there's a navigation bar with the AWS logo, 'Services' dropdown, a search bar containing 'Search', and a keyboard shortcut '[Alt+S]'. Below the navigation, the breadcrumb trail shows 'EC2 > Launch templates > Create launch template'. The main title is 'Create launch template'. A sub-instruction says 'Creating a launch template allows you to create a saved instance configuration that can be reused, shared and launched at a later time. Templates can have multiple versions.' The first section is 'Launch template name and description'. It has a 'Launch template name - required' field containing 'LabConfig', with a note below it: 'Must be unique to this account. Max 128 chars. No spaces or special characters like '&', '\*', '@'.' Below that is a 'Template version description' field containing 'A prod webserver for MyApp', with a note: 'Max 255 chars'. The next section is 'Auto Scaling guidance' with an 'Info' link. It includes a checkbox labeled 'Provide guidance to help me set up a template that I can use with EC2 Auto Scaling', which is checked. The final section is 'Launch template contents', which is currently collapsed. It contains two expandable sections: 'Application and OS Images (Amazon Machine Image)' and 'Amazon Machine Image (AMI)'. The 'Application and OS Images (Amazon Machine Image)' section is expanded, showing the 'My AMIs' list.

- In the Application and OS Images (Amazon Machine Image) area, choose *My AMIs*.
- **Amazon Machine Image (AMI)**: choose *Web Server AMI*

The screenshot shows the AWS Management Console with the AWS logo and Services navigation bar. The main title is "Launch template contents". A sub-section titled "Application and OS Images (Amazon Machine Image) - required" is expanded. It contains a search bar and tabs for "Recents", "My AMIs" (which is selected), and "Quick Start". Below these are two buttons: "Owned by me" (selected) and "Shared with me". To the right is a search icon and a link to "Browse more AMIs" which includes "Including AMIs from AWS, Marketplace and the Community". Under "Amazon Machine Image (AMI)", there is a list with one item: "WebServerAMI" (ami-0e822fe1926e7f620, created 2023-11-18T09:18:47.000Z, Virtualization: hvm, ENA enabled: true, Root device type: ebs). The "Description" field contains "Lab AMI for Web Server". The "Architecture" field shows "x86\_64" and the "AMI ID" field shows "ami-0e822fe1926e7f620".

- **Instance type:** choose *t2.micro*
- **Key pair name:** choose *vockey*
- **Firewall (security groups):** choose *Select existing security group*
- **Security groups:** choose *Web Security Group*

The screenshot shows the AWS Launch Wizard interface for creating a new Amazon EC2 instance. The configuration steps are as follows:

- Instance type**: Set to t2.micro, which is Free tier eligible. It includes details about the family (t2), vCPU count (1), memory (1 GiB), current generation (true), On-Demand Windows base pricing (0.0162 USD per Hour), On-Demand SUSE base pricing (0.0116 USD per Hour), On-Demand RHEL base pricing (0.0716 USD per Hour), and On-Demand Linux base pricing (0.0116 USD per Hour). A note states: "Additional costs apply for AMIs with pre-installed software".
- Key pair (login)**: Set to "vokey". A "Create new key pair" button is available.
- Network settings**:
  - Subnet**: Set to "Don't include in launch template". A "Create new subnet" button is available.
  - Firewall (security groups)**: A note says: "A security group is a set of firewall rules that control the traffic for your instance. Add rules to allow specific traffic to reach your instance." Two options are shown: "Select existing security group" (selected) and "Create security group".

- Scroll down to the **Advanced details** area and expand it.
- Scroll down to the **Detailed CloudWatch monitoring** setting. Select *Enable*

Not applicable for Amazon EC2 Auto Scaling.

Termination protection | [Info](#)

Don't include in launch template ▾

Stop protection | [Info](#)

Don't include in launch template ▾

Detailed CloudWatch monitoring | [Info](#)

Enable ▾

[Additional charges apply](#)

Elastic GPU | [Info](#)

Don't include in launch template ▾

Elastic inference | [Info](#)

- o Choose **Create launch template**

25. In the Success dialog, choose the **LabConfig** launch template.

The screenshot shows the AWS EC2 Launch Templates interface. At the top, there's a navigation bar with 'Services' and a search bar. Below it, the breadcrumb trail shows 'EC2 > Launch templates > Create launch template'. A green success message box displays the text: 'Success Successfully created LabConfig(lt-0459c1fe796f963b4).'. To the right of the message box is a small 'Actions log' button. Below the message box is a section titled 'Next Steps' with three main options: 'Launch an instance', 'Launch instance from this template', and 'Create an Auto Scaling group from your template'. Each option has a brief description and a link to further details. At the bottom right of the page is a yellow 'View Launch templates' button.

## ▼ Summary

### Software Image (AMI)

Lab AMI for Web Server  
ami-0e822fe1926e7f620

### Virtual server type (instance type)

t2.micro

### Firewall (security group)

Web Security Group

### Storage (volumes)

1 volume(s) - 8 GiB

**i** **Free tier:** In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance usage on free tier AMIs per month, 30 GiB of EBS storage, 2 million IOs, 1 GB of snapshots, and 100 GB of bandwidth to the internet. X

Cancel

**Create launch template**

The screenshot shows the AWS EC2 Launch Templates page. The top navigation bar includes 'EC2 > Launch templates > LabConfig'. The main title is 'LabConfig (lt-0459c1fe796f9b3b4)'. On the right, there are 'Actions' and 'Delete template' buttons. Below the title, there's a 'Launch template details' section with fields: 'Launch template ID' (lt-0459c1fe796f9b3b4), 'Launch template name' (LabConfig), 'Default version' (1), and 'Owner' (arn:aws:sts::529578993132:assumed-role/voclabs/user2813720=Zeeshan\_Ali). There are tabs for 'Details' (selected), 'Versions', and 'Template tags'. Below this is a 'Launch template version details' section for version 1 (Default). It shows 'Description' (-), 'Date created' (2023-11-18T09:43:08.000Z), and 'Created by' (arn:aws:sts::529578993132:assumed-role/voclabs/user2813720=Zeeshan\_Ali). Under 'Instance details', it lists 'AMI ID' (ami-0e822fe1926e7f620), 'Instance type' (t2.micro), 'Availability Zone' (-), and 'Key pair name' (vockey). Other tabs include 'Storage', 'Resource tags', 'Network interfaces', and 'Advanced details'.

26. From the **Actions** menu, choose *Create Auto Scaling group*

The screenshot shows a dropdown menu from the 'Actions' button. The menu items are: 'Launch instance from template', 'Modify template (Create new version)', 'Set default version', 'Create Auto Scaling group' (which is highlighted with a gray background), and 'Create Spot Fleet'. To the left of the menu, there's a partial view of the 'Create role' wizard. Below the menu, there's a 'Key pair name' field containing 'vockey'.

27. Configure the details in Step 1 (Choose launch template or configuration):

- o **Auto Scaling group name:** Lab Auto Scaling Group
- o **Launch template:** confirm that the *LabConfig* template you just created is selected.
- o Choose **Next**

The screenshot shows the AWS EC2 Auto Scaling group creation wizard. The left sidebar lists steps: Step 1 (Choose launch template), Step 2 (Choose instance launch options), Step 3 - optional (Configure advanced options), Step 4 - optional (Configure group size and scaling), Step 5 - optional (Add notifications), Step 6 - optional (Add tags), and Step 7 (Review). The main content area is titled "Choose launch template". It asks to specify a launch template for all instances. A "Name" field contains "Lab Auto Scaling Group". A note states: "For accounts created after May 31, 2023, the EC2 console only supports creating Auto Scaling groups with launch templates. Creating Auto Scaling groups with launch configurations is not recommended but still available via the CLI and API until December 31, 2023." Below this is a "Launch template" section with a dropdown set to "LabConfig", a "Create a launch template" button, and a "Version" dropdown set to "Default (1)". A "Create a launch template version" button is also present. To the right, there's a table showing "Description" (empty), "Launch template" (LabConfig), and "Instance type" (t2.micro).

28. Configure the details in Step 2 (Choose instance launch options):

- o **VPC:** choose *Lab VPC*
- o **Availability Zones and subnets:** Choose *Private Subnet 1* and then choose *Private Subnet 2*.
- o Choose **Next**

The screenshot shows the AWS Auto Scaling configuration interface. On the left, a sidebar lists steps: Step 3 - optional (Configure advanced options), Step 4 - optional (Configure group size and scaling), Step 5 - optional (Add notifications), Step 6 - optional (Add tags), and Step 7 (Review). The main area is titled "Instance type requirements" and includes a "Launch template" section with "LabConfig" selected, a "Version" of "Default", and a "Description" of "-". It also shows an "Instance type" of "t2.micro". A "Network" section follows, with a note about using multiple Availability Zones. It shows a VPC dropdown set to "vpc-0c6453b6cdef0ac02 (Lab VPC)" and a subnet selection dropdown containing two entries: "us-east-1a | subnet-063469bba9c357f23 (Private Subnet 1)" and "10.0.1.0/24", and "us-east-1b | subnet-01c8452553da8e29d (Private Subnet 2)" and "10.0.3.0/24". A "Create a subnet" link is also present.

29. Configure the details in Step 3 (Configure advanced options):

- o Choose **Attach to an existing load balancer**
  - **Existing load balancer target groups:** select *LabGroup*.

Search [Alt+S]

[Create template](#)

[Advanced options](#)

[Launch options](#)

## Configure advanced options - optional Info

Integrate your Auto Scaling group with other services to distribute network traffic across multiple servers using a load balancer or to establish service-to-service communications using VPC Lattice. You can also set options that give you more control over health check replacements and monitoring.

### Load balancing Info

Use the options below to attach your Auto Scaling group to an existing load balancer, or to a new load balancer that you define.

No load balancer  
Traffic to your Auto Scaling group will not be fronted by a load balancer.

Attach to an existing load balancer  
Choose from your existing load balancers.

Attach to a new load balancer  
Quickly create a basic load balancer to attach to your Auto Scaling group.

#### Attach to an existing load balancer

Select the load balancers that you want to attach to your Auto Scaling group.

Choose from your load balancer target groups  
This option allows you to attach Application, Network, or Gateway Load Balancers.

Choose from Classic Load Balancers

Existing load balancer target groups  
Only instance target groups that belong to the same VPC as your Auto Scaling group are available for selection.

Select target groups ▾ LabGroup | HTTP X C

LabGroup | HTTP  
Application Load Balancer: LabELB

- In the **Additional settings** pane:
  - Select **Enable group metrics collection within CloudWatch Metrics**
- Choose **Next**

The screenshot shows the AWS CloudWatch Metrics console. At the top, there's a navigation bar with 'Services', a search bar, and a keyboard shortcut '[Alt+S]'. Below the navigation bar, the main content area has two sections:

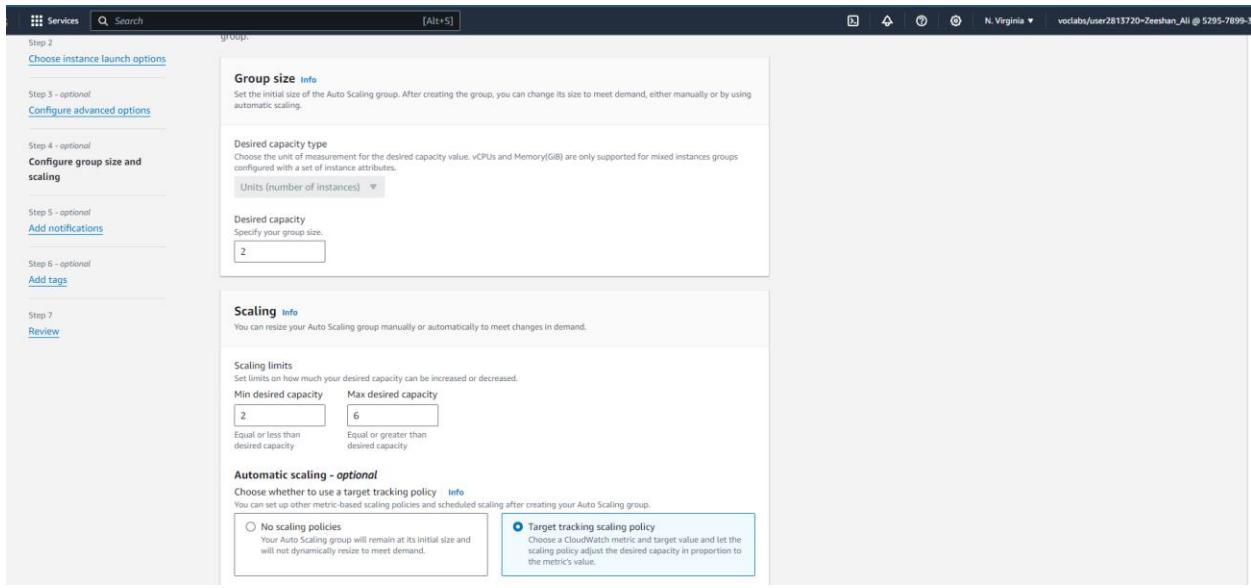
- Health checks**:
  - EC2 health checks**:
    - Always enabled**
  - Additional health check types - optional**:
    - Turn on Elastic Load Balancing health checks** (**Recommended**)  
Elastic Load Balancing monitors whether instances are available to handle requests. When it reports an unhealthy instance, EC2 Auto Scaling can replace it on its next periodic check.
    - Turn on VPC Lattice health checks**  
VPC Lattice can monitor whether instances are available to handle requests. If it considers a target as failed a health check, EC2 Auto Scaling replaces it after its next periodic check.
- Health check grace period**:
  - Info**
  - This time period delays the first health check until your instances finish initializing. It doesn't prevent an instance from terminating when placed into a non-running state.
  - 300 seconds**

- Additional settings**:
- Monitoring**:
  - Enable group metrics collection within CloudWatch**
- Default instance warmup**:
  - Info**
  - The amount of time that CloudWatch metrics for new instances do not contribute to the group's aggregated instance metrics, as their usage data is not reliable yet.
  - Enable default instance warmup**

At the bottom right, there are buttons for **Cancel**, **Skip to review**, **Previous**, and **Next**.

30. Configure the details in Step 4 (Configure group size and scaling policies - optional):

- o Under **Group size**, configure:
  - **Desired capacity:** 2
  - **Minimum capacity:** 2
  - **Maximum capacity:** 6



- Under **Scaling policies**, choose *Target tracking scaling policy* and configure:
  - **Scaling policy name:** LabScalingPolicy
  - **Metric type:** Average CPU Utilization
  - **Target value:** 60
- Choose **Next**

Scaling policy name: LabScalingPolicy

Metric type: Average CPU utilization

Target value: 60

Instance warmup: 300 seconds

Disable scale in to create only a scale-out policy

**Instance maintenance policy - new**

An instance maintenance policy determines how much availability your application has when EC2 Auto Scaling replaces instances. It also establishes guardrails that limit the amount of capacity that can be added or removed when replacing instances.

Control availability and cost during replacement events

Choose a replacement behavior depending on your availability requirements

<b>Mixed behavior</b>	<b>Prioritize availability</b>	<b>Control costs</b>	<b>Flexible</b>
<input checked="" type="radio"/> No policy For rebalancing events, new instances will launch before	<input type="radio"/> Launch before terminating Launch new instances	<input type="radio"/> Terminate and launch Terminate and launch	<input type="radio"/> Custom behavior Set custom values for the minimum and maximum amount of

31. Configure the details in Step 5 (Add notifications - optional): Auto Scaling can send a notification when a scaling event takes place. You will use the default settings.

- o Choose **Next**

Step 1: Choose launch template

Step 2: Choose instance launch options

Step 3 - optional: Configure advanced options

Step 4 - optional: Configure group size and scaling

Add notifications - optional

Add notification

Cancel Skip to review Previous Next

32. Configure the details in Step 6 (Add tags - optional): Tags applied to the Auto Scaling group will be automatically propagated to the instances that are launched.

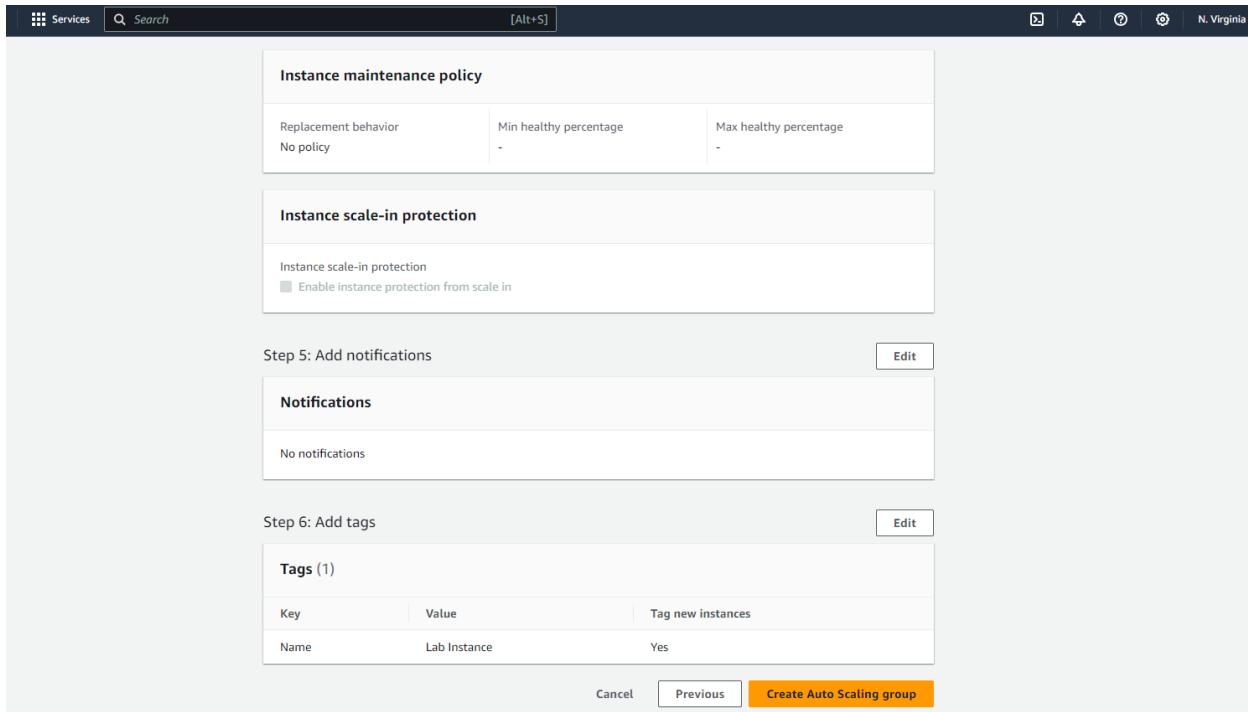
- o Choose **Add tag** and Configure the following:

- **Key:** Name
- **Value:** Lab Instance
- Choose **Next**

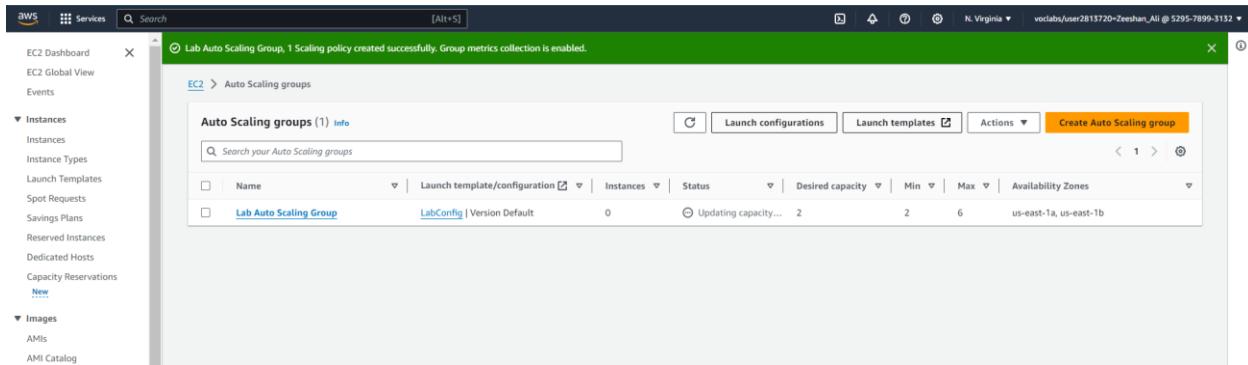
The screenshot shows the AWS EC2 Auto Scaling group creation wizard. The current step is Step 6: **Add tags - optional**. On the left, a sidebar lists steps from 1 to 7. Step 6 is highlighted. The main area shows a table for adding tags. One tag is already added: Key "Name" and Value "Lab Instance". A checkbox "Tag new instances" is checked. Below the table, there is an "Add tag" button and a note saying "49 remaining". At the bottom, there are "Cancel", "Previous", and "Next" buttons. The "Next" button is highlighted.

### 33. Configure the details in Step 6 (Review):

- Review the details of your Auto Scaling group
- Choose **Create Auto Scaling group**



Your Auto Scaling group will initially show an instance count of zero, but new instances will be launched to reach the **Desired** count of 2 instances.



## Task 4: Verify that Load Balancing is Working

44. In the left navigation pane, choose **Instances**. You should see two new instances named **Lab Instance**. These were launched by Auto Scaling. If the instances or names are not displayed, wait 30 seconds and choose refresh in the top-right. Next, you will confirm that the new instances have passed their Health Check.

The screenshot shows the AWS EC2 Instances page. The left navigation pane includes options like EC2 Dashboard, EC2 Global View, Events, Instances, Instance Types, Launch Templates, Spot Requests, Savings Plans, Reserved Instances, Dedicated Hosts, Capacity Reservations, Images, AMIs, AMI Catalog, and Elastic Block Store. The main content area displays a table of four instances:

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS	Public IPv4 IP	Elastic IP
Web Server 1	i-0ca194354604afc82	Running	t2.micro	2/2 checks passed	No alarms	us-east-1a	-	3.95.156.194	-
Lab Instance	i-06bc98ef805defe2	Running	t2.micro	2/2 checks passed	No alarms	us-east-1a	-	-	-
Bastion Host	i-0078566f44cea3356	Running	t2.micro	2/2 checks passed	No alarms	us-east-1a	-	3.83.30.108	-
Lab Instance	i-05d4f2ace4c8765f2	Running	t2.micro	2/2 checks passed	No alarms	us-east-1b	-	-	-

45. In the left navigation pane, choose **Target Groups**.

46. Select *LabGroup*

47. Choose the **Targets** tab. Two target instances named Lab Instance should be listed in the target group.

48. Wait until the **Status** of both instances transitions to *healthy*. Choose Refresh in the upper-right to check for updates if necessary. *Healthy* indicates that an instance has passed the Load Balancer's health check. This means that the Load Balancer will send traffic to the instance. You can now access the Auto Scaling group via the Load Balancer.

The screenshot shows the AWS Target Groups page. The left navigation pane includes options like EC2, Target groups, Images, AMIs, AMI Catalog, Elastic Block Store, Network & Security, Load Balancing, Auto Scaling, and Load Balancers. The main content area shows a table of target groups:

Name	ARN	Port	Protocol	Target type	Load balancer	VPC ID
LabGroup	arn:aws:elasticloadbalancing:us-east-1:123456789012:targetgroup/LabGroup/09876543210987654321	80	HTTP	Instance	LabELB	vpc-0c6453b6ccdef0ac02

Below this, the Details pane shows the Target group: LabGroup. The Targets tab is selected, displaying the Registered targets (2) table:

Instance ID	Name	Port	Zone	Health status	Health status details	Launch time
i-06bc98ef805defe2	Lab Instance	80	us-east-1a	Healthy	-	November 18, 2023, 15:...
i-05d4f2ace4c8765f2	Lab Instance	80	us-east-1b	Healthy	-	November 18, 2023, 15:...

49. In the left navigation pane, choose **Load Balancers**.

50. Select the *LabELB* load balancer.

51. In the Details pane, copy the **DNS name** of the load balancer, making sure to omit "(A Record)". It should look similar to: *LabELB-1998580470.us-west-2.elb.amazonaws.com*

52. Open a new web browser tab, paste the DNS Name you just copied, and press Enter.

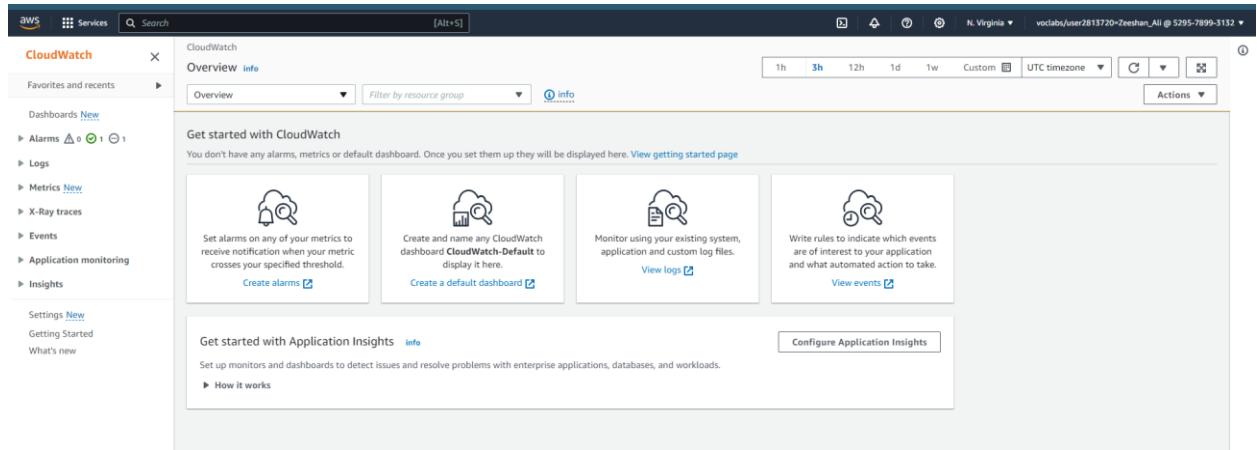
Meta-Data	Value
InstanceId	i-05d4f2ace4c8765f2
Availability Zone	us-east-1b

Current CPU Load: 0%

The application should appear in your browser. This indicates that the Load Balancer received the request, sent it to one of the EC2 instances, then passed back the result.

## Task 5: Test Auto Scaling

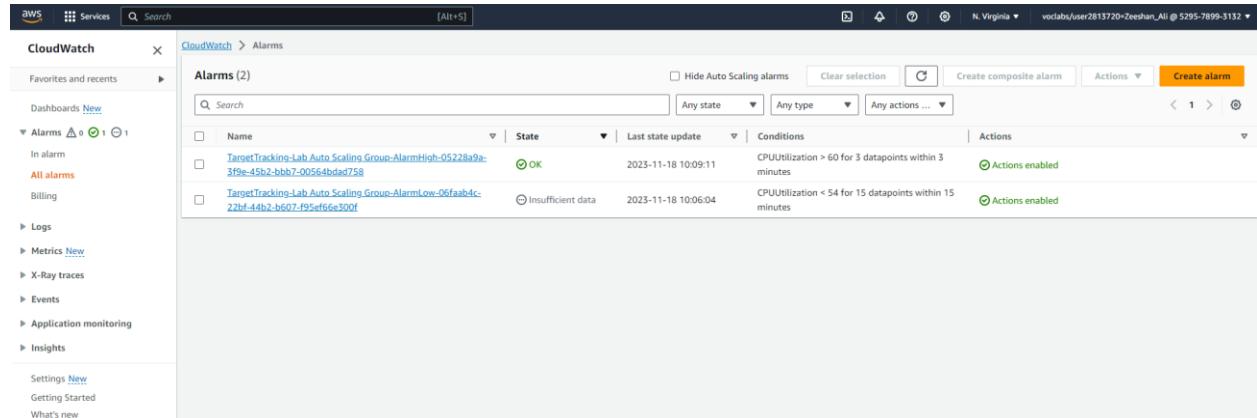
52. Return to the AWS Management Console, but do not close the application tab — you will return to it soon.
53. in the search box next to **Services**, search for and select **CloudWatch**.



54. In the left navigation pane, choose **All alarms**

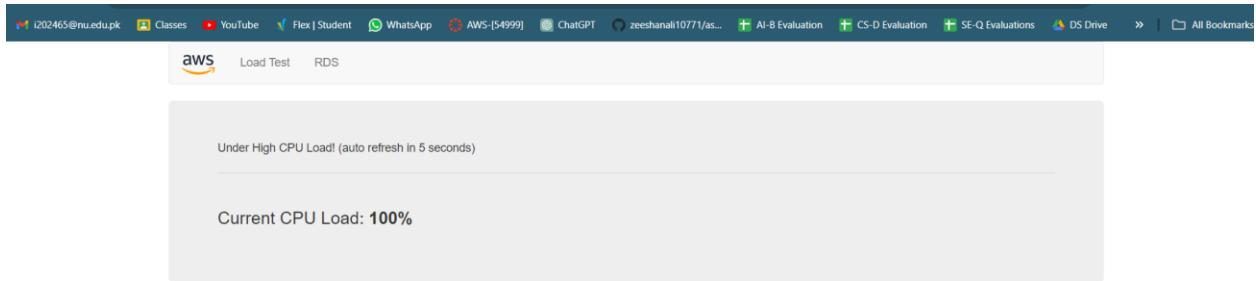
Two alarms will be displayed. These were created automatically by the Auto Scaling group. They will automatically keep the average CPU load close to 60% while also staying within the limitation of having two to six instances.

55. Choose the **OK** alarm, which has *AlarmHigh* in its name. The **OK** indicates that the alarm has *not* been triggered. It is the alarm for **CPU Utilization > 60**, which will add instances when average CPU is high. The chart should show very low levels of CPU at the moment. You will now tell the application to perform calculations that should raise the CPU level.



56. Return to the browser tab with the web application.

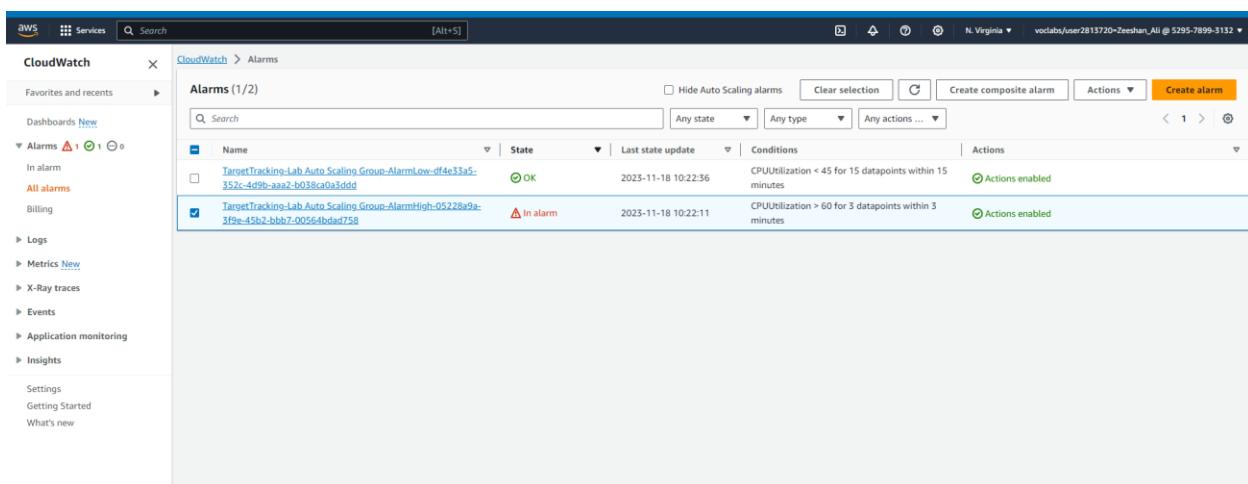
57. Choose **Load Test** beside the AWS logo. This will cause the application to generate high loads. The browser page will automatically refresh so that all instances in the Auto Scaling group will generate load. Do not close this tab.



58. Return to browser tab with the **CloudWatch** console. In less than 5 minutes, the **AlarmLow** alarm should change to **OK** and the **AlarmHigh** alarm status should change to *In alarm*. You can choose Refresh in the top-right every 60 seconds to update the display.

You should see the **AlarmHigh** chart indicating an increasing CPU percentage. Once it crosses the 60% line for more than 3 minutes, it will trigger Auto Scaling to add additional instances.

59. Wait until the **AlarmHigh** alarm enters the *In alarm* state. You can now view the additional instance(s) that were launched.



Name	State	Last state update	Conditions	Actions
TargetTracking-Lab Auto Scaling Group-AlarmLow-0f4e33a5-352c-4d9b-aaa2-b038ca0a3dd0	OK	2023-11-18 10:22:36	CPUUtilization < 45 for 15 datapoints within 15 minutes	Actions enabled
TargetTracking-Lab Auto Scaling Group-AlarmHigh-05228a9a-3f9e-45b2-bbb7-00564bded758	In alarm	2023-11-18 10:22:11	CPUUtilization > 60 for 3 datapoints within 3 minutes	Actions enabled

60. In the search box next to **Services**, search for and select **EC2**.

61. In the left navigation pane, choose **Instances**.

More than two instances labeled **Lab Instance** should now be running. The new instance(s) were created by Auto Scaling in response to the CloudWatch alarm.

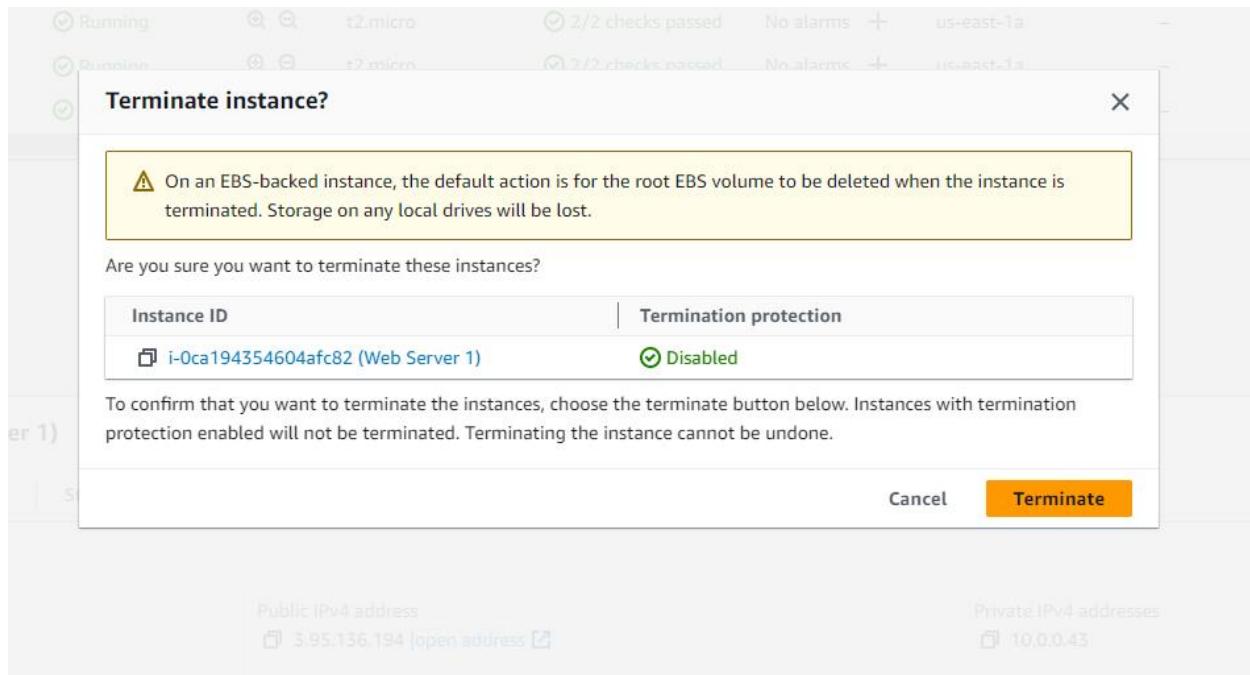
Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS	Public IPv4 ...	Elastic IP
Lab Instance	i-0919de0c8b0370349	Running	t2.micro	2/2 checks passed	No alarms	us-east-1b	-	-	-
Web Server 1	i-0ca194354604afc82	Running	t2.micro	2/2 checks passed	No alarms	us-east-1a	-	3.95.136.194	-
Lab Instance	i-0c6bc98ef803defe2	Running	t2.micro	2/2 checks passed	No alarms	us-east-1a	-	-	-
Bastion Host	i-007856f44ce3556	Running	t2.micro	2/2 checks passed	No alarms	us-east-1a	-	3.83.30.108	-
Lab Instance	i-05d4f2ace4c8765f2	Running	t2.micro	2/2 checks passed	No alarms	us-east-1b	-	-	-

## Task 6: Terminate Web Server 1

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

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS	Public IPv4 ...	Elastic IP
Lab Instance	i-0919de0c8b0370349	Running	t2.micro	2/2 checks passed	No alarms	us-east-1b	-	-	-
<b>Web Server 1</b>	i-0ca194354604afc82	Running	t2.micro	2/2 checks passed	No alarms	us-east-1a	-	3.95.136.194	-
Lab Instance	i-0c6bc98ef803defe2	Running	t2.micro	2/2 checks passed	No alarms	us-east-1a	-	-	-
Bastion Host	i-007856f44ce3556	Running	t2.micro	2/2 checks passed	No alarms	us-east-1a	-	3.83.30.108	-
Lab Instance	i-05d4f2ace4c8765f2	Running	t2.micro	2/2 checks passed	No alarms	us-east-1b	-	-	-

64. Choose **Terminate**



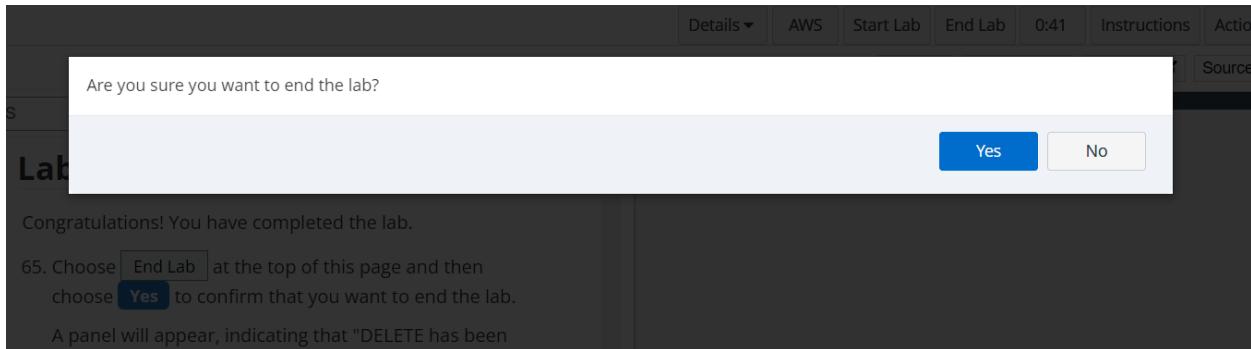
The screenshot shows the EC2 Instances page in the AWS Management Console. The left sidebar shows 'Instances' selected. The main area displays a table of instances. A success message at the top says 'Successfully terminated i-0ca194354604afc82'. The table shows five instances: 'Lab Instance' (Running, t2.micro), 'Web Server 1' (Shutting-down, t2.micro), 'Lab Instance' (Running, t2.micro), 'Bastion Host' (Running, t2.micro), and 'Lab Instance' (Running, t2.micro). The 'Web Server 1' row is highlighted. The 'Public IPv4 address' column shows '3.95.136.194' with a link to 'open address'. The 'Private IPv4 addresses' column shows '10.0.0.43'. The status bar at the bottom indicates the instance ID is 'i-0ca194354604afc82'.

This screenshot is identical to the one above, showing the EC2 Instances page after the 'Web Server 1' instance has been terminated. The 'Web Server 1' row now shows 'Terminated' under 'Instance state' instead of 'Shutting-down'. The rest of the instance details and the status bar remain the same.

# Lab Complete

Congratulations! You have completed the lab.

65. Choose **End Lab** at the top of this page and then choose **Yes** to confirm that you want to end the lab.



A panel will appear, indicating that "DELETE has been initiated... You may close this message box now."

66. Choose the **X** in the top right corner to close the panel.

