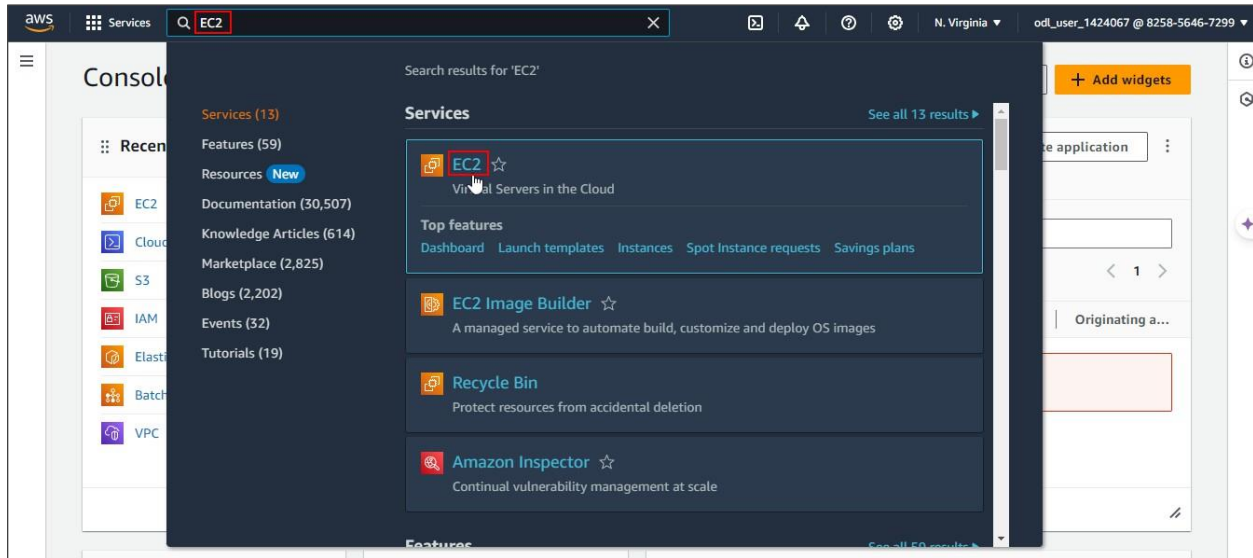


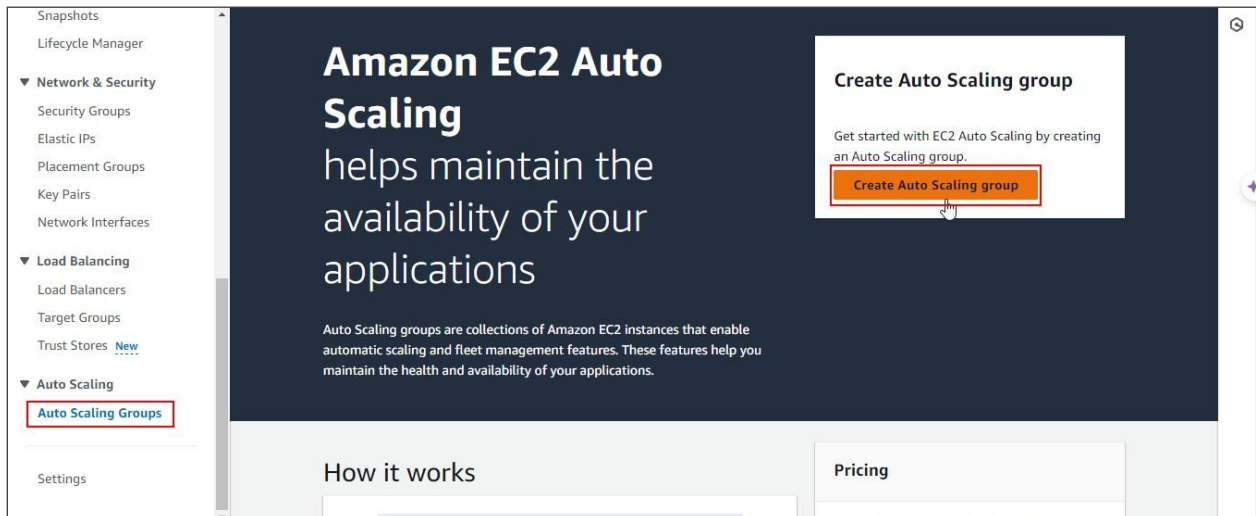
Auto Scaling

Step 1: Set up a predefined auto-scaling group

1.1 Navigate to the AWS console home dashboard, search for and click on EC2



1.2 Navigate to Auto Scaling Groups in the Auto Scaling section, and click on Create Auto Scaling group



1.3 Add the name as **Testing**, select **MyTestTemplate** in the **Launch template**, and click **Next**.

If you don't have any Launch Template then you will need to create one to be used with this practice session.

The screenshot shows the AWS Auto Scaling console during the 'Launch template' step. The left sidebar contains a navigation menu with the following steps:

- Step 3 - optional: Configure advanced options
- Step 4 - optional: Configure group size and scaling policies
- Step 5 - optional: Add notifications
- Step 6 - optional: Add tags
- Step 7: Review

The main content area is titled 'Launch template' with an 'Info' link and a 'Switch to launch configuration' link. It contains the following fields and options:

- Auto Scaling group name:** A text input field containing 'Testing'.
- Launch template:** A dropdown menu showing 'MyTestTemplate'.
- Version:** A dropdown menu showing 'Default (1)'.
- Description:** A text input field containing 'MyTestTemplate'.
- Launch template:** A text input field containing 'MyTestTemplate'.
- Instance type:** A text input field containing 't2.micro'.

There are also links to 'Create a launch template' and 'Create a launch template version'.

1.4 Select the availability zones and subnets as **us-east-1a** and **us-east-1b**, then click **Next**

The screenshot shows the 'Network' configuration step in the AWS Management Console. On the left, a sidebar indicates 'Step 6 - optional' and 'Add tags'. The main content area is titled 'Network' with an 'Info' icon. It lists several subnets for selection. Two subnets are selected and highlighted with a red box: 'us-east-1a | subnet-08219e2328d8d7c3b' and 'us-east-1b | subnet-0081290ff8abfd41f'. Below the list, a dropdown menu shows 'Select Availability Zones and subnets'. At the bottom of the main content area, there are four buttons: 'Cancel', 'Skip to review', 'Previous', and 'Next'. The 'Next' button is highlighted with a red box.

1.5 Click on **Next**

The screenshot shows the 'Additional settings' configuration step in the AWS Management Console. The main content area contains information about 'Health check grace period' and 'Additional settings'. Under 'Additional settings', there are two sections: 'Monitoring' and 'Default instance warmup'. The 'Next' button at the bottom right is highlighted with a red box.

1.6 Add the **Desired capacity**, **Minimum capacity**, and **Maximum capacity** as **2**, and click on **Next**

This screenshot shows the 'Configure group size and scaling' step in the AWS IAM console. The left sidebar contains a navigation menu with links for 'Choose instance launch options', 'Configure advanced options', 'Add notifications', 'Add tags', and 'Review'. The main content area is divided into three sections: 'Group size', 'Desired capacity type', and 'Scaling'. The 'Group size' section has a description and a 'Units (number of instances)' dropdown menu. The 'Desired capacity type' section has a description and a 'Desired capacity' input field with the value '2'. The 'Scaling' section has a description and 'Scaling limits' with 'Min desired capacity' and 'Max desired capacity' input fields, both with the value '2'. The 'Automatic scaling - optional' section is at the bottom.

Group size Info
Set the initial size of the Auto Scaling group. After creating the group, you can change its size to meet demand, either manually or by using automatic scaling.

Desired capacity type
Choose the unit of measurement for the desired capacity value. vCPUs and Memory(GiB) are only supported for mixed instances groups configured with a set of instance attributes.

Units (number of instances) ▼

Desired capacity
Specify your group size:

Scaling Info
You can resize your Auto Scaling group manually or automatically to meet changes in demand.

Scaling limits
Set limits on how much your desired capacity can be increased or decreased.

Min desired capacity **Max desired capacity**

Equal or less than desired capacity Equal or greater than desired capacity

Automatic scaling - optional

This screenshot shows the 'Instance scale-in protection' step in the AWS IAM console. The main content area has a description and a 'Choose a replacement behavior depending on your availability requirements' section with four options: 'Mixed behavior', 'Prioritize availability', 'Control costs', and 'Flexible'. The 'Mixed behavior' option is selected. Below this is the 'Instance scale-in protection' section with a description and a checkbox to 'Enable instance scale-in protection'. At the bottom, there are four buttons: 'Cancel', 'Skip to review', 'Previous', and 'Next'.

Control your Auto Scaling group's availability during instance replacement events. This includes health checks, instance refreshes, maximum instance lifetime features and events that happen automatically to keep your group balanced, called rebalancing events.

Choose a replacement behavior depending on your availability requirements

Mixed behavior
☒ **No policy**
For rebalancing events, new instances will launch before terminating others. For all other events, instances terminate and launch at the same time.

Prioritize availability
☐ **Launch before terminating**
Launch new instances and wait for them to be ready before terminating others. This allows you to go above your desired capacity by a given percentage and may temporarily increase costs.

Control costs
☐ **Terminate and launch**
Terminate and launch instances at the same time. This allows you to go below your desired capacity by a given percentage and may temporarily reduce availability.

Flexible
☐ **Custom behavior**
Set custom values for the minimum and maximum amount of available capacity. This gives you greater flexibility in setting how far below and over your desired capacity EC2 Auto Scaling goes when replacing instances.

Instance scale-in protection
Scale-in protection prevents newly launched instances from being terminated by scaling activities. Make sure to remove scale-in protection for the group or individual instances when instances are ready to be terminated.

☐ Enable instance scale-in protection

Cancel Skip to review Previous **Next**

1.7 Review the steps, and click **Create Auto Scaling group**

Instance scale-in protection

Instance scale-in protection

☐ Enable instance protection from scale in

Step 5: Add notifications Edit

Notifications

No notifications

Step 6: Add tags Edit

Tags (0)

Key	Value	Tag new instances
No tags		

Cancel Previous **Create Auto Scaling group**

EC2 > Auto Scaling groups

Auto Scaling groups (1) [Info](#) Refresh Launch configurations Launch templates Actions **Create Auto Scaling group**

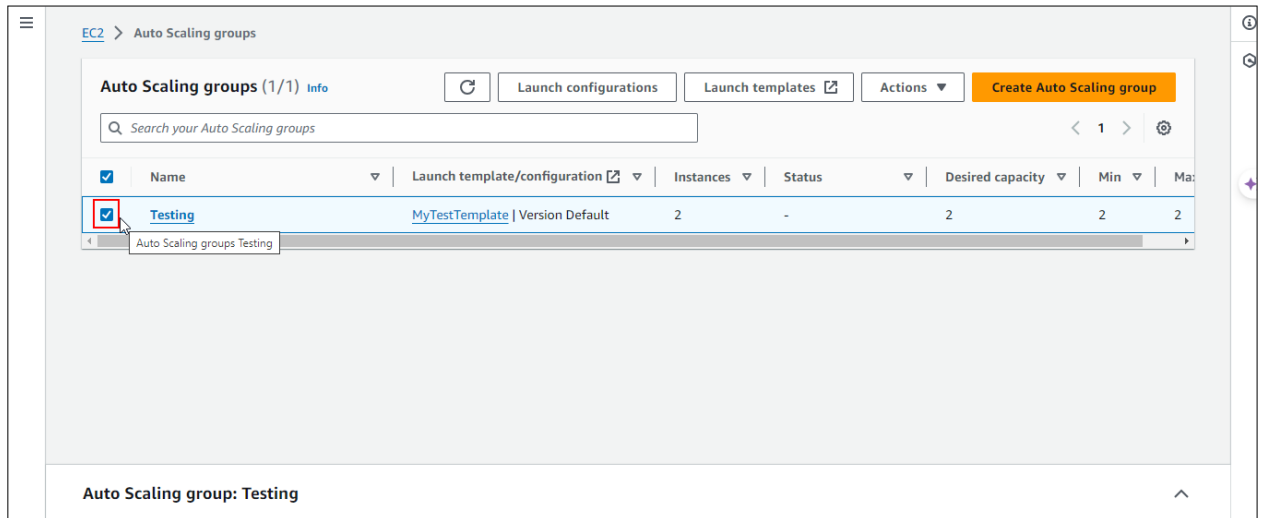
<input type="checkbox"/>	Name	Launch template/configuration	Instances	Status	Desired capacity	Min	Max
<input type="checkbox"/>	Testing	MyTestTemplate Version Default	2	-	2	2	2

0 Auto Scaling groups selected

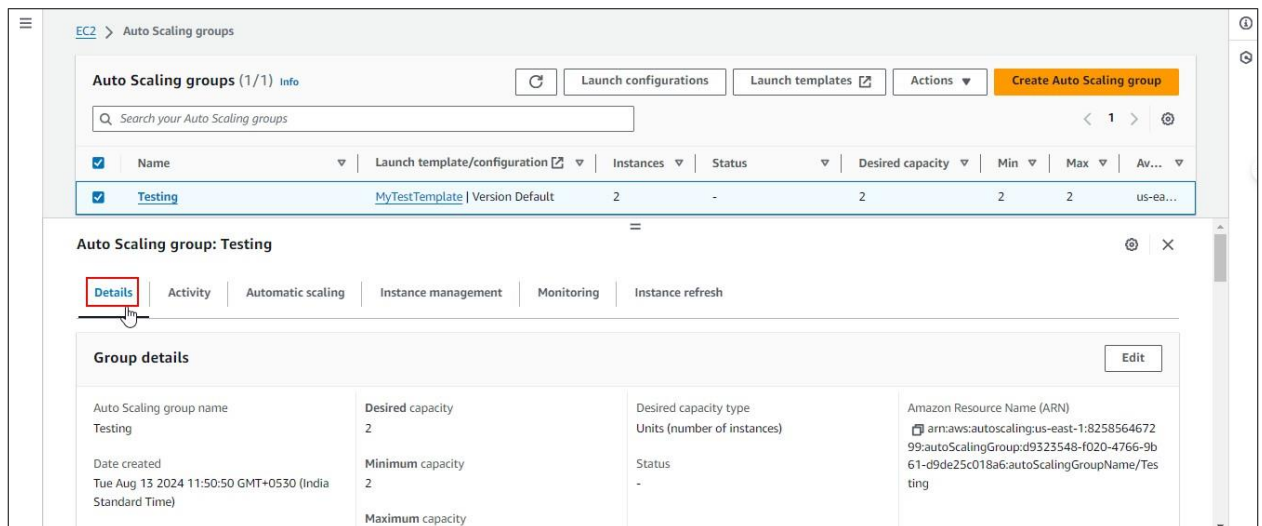
Auto-scaling groups have been created successfully.

Step 2: Set up EC2 Auto Scaling with a Load Balancer

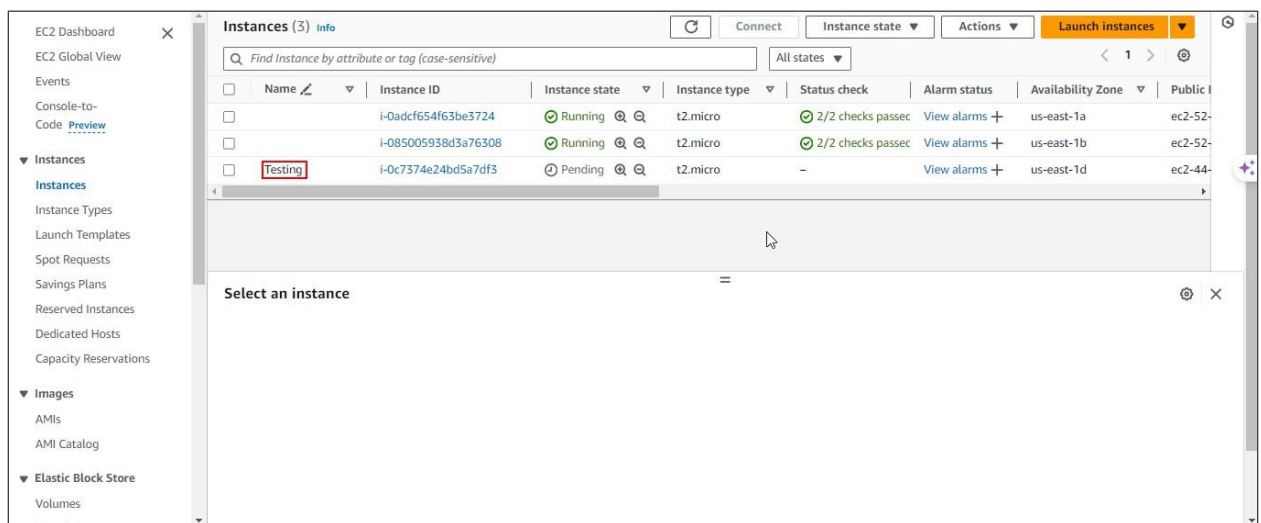
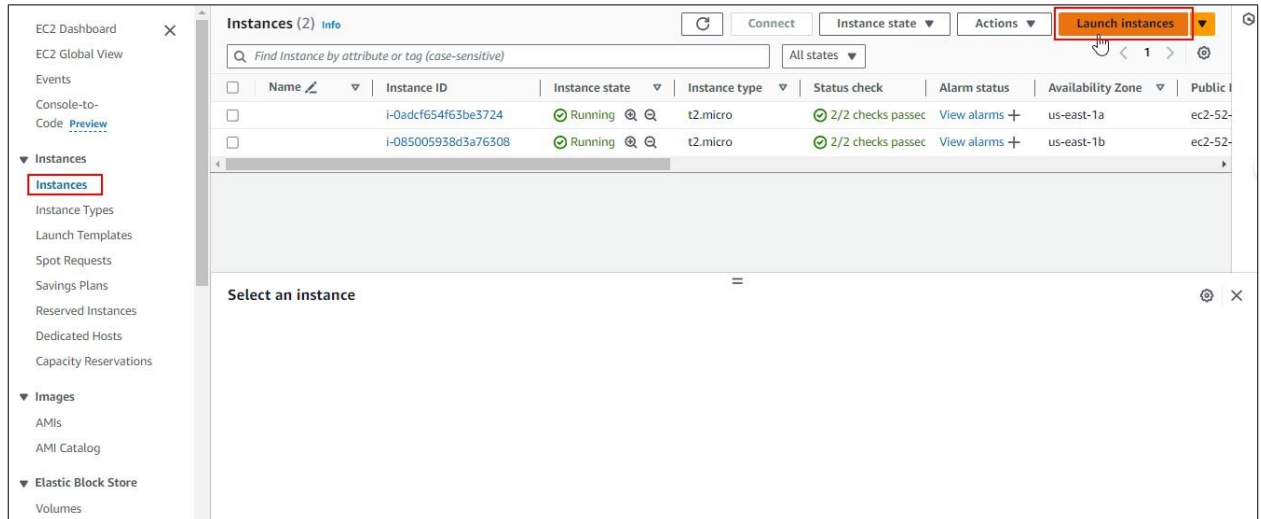
2.1 Select the previously created auto-scaling group as shown:



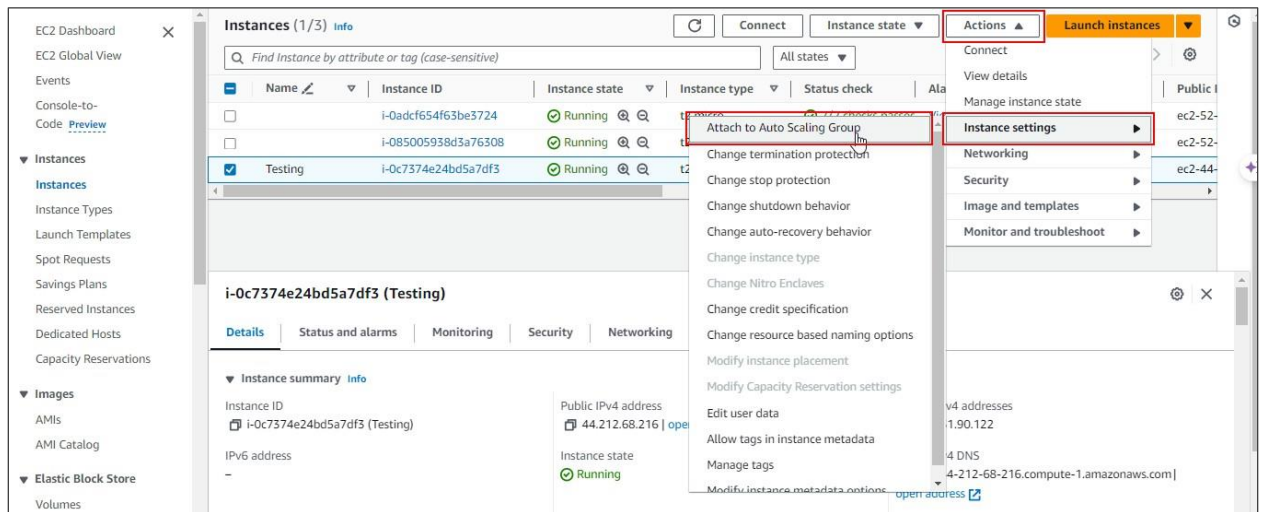
2.2 Click on **Details** to verify the group details



2.3 Navigate to the **Instances**, and click **Launch instances** to create a new instance named **Testing**



2.4 Click on **Actions**, then select **Instance settings**, and choose **Attach to Auto Scaling Group**



2.5 Select the **Auto Scaling Group** name **Testing**, and click **Attach**

