Lesson 03 Demo 10

Configuring the Manual and Dynamic Scaling

Objective: To understand how to configure manual and dynamic scaling for your application using Amazon Web Services (AWS) tools and services.

Tools required: AWS workspace

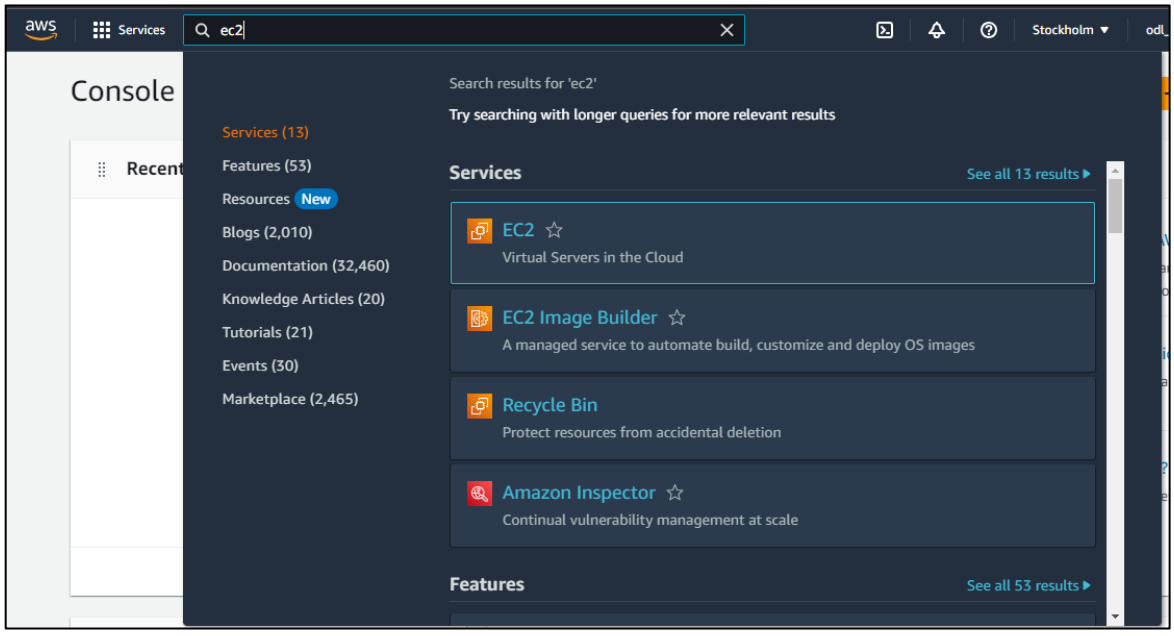
Prerequisites: Create an EC2 instance named S3

Steps to be followed:

1. Set up predefined auto scaling group
2. Set up EC2 Auto Scaling with Load Balancer

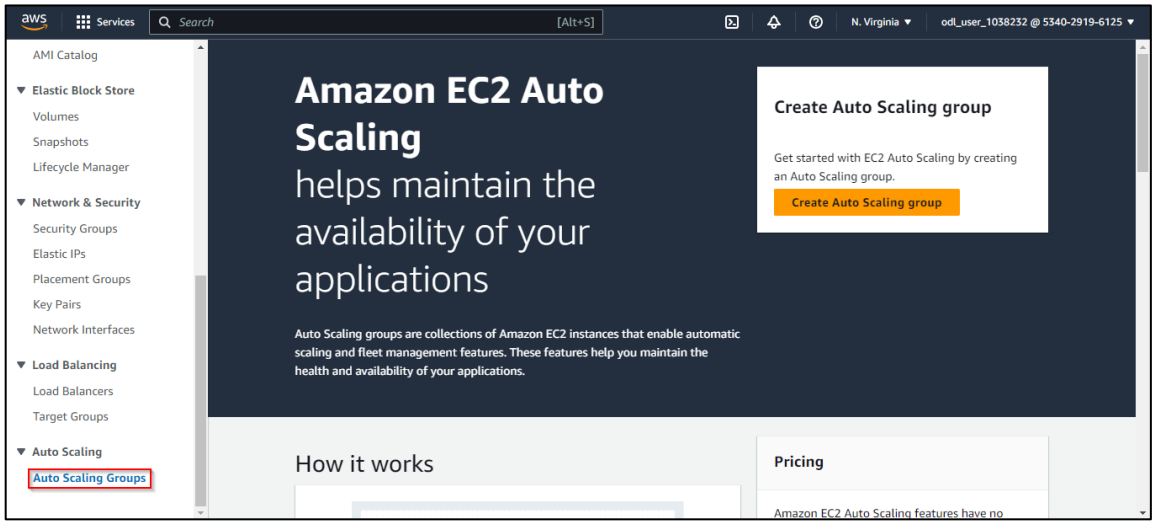
Step 1: Set up predefined auto scaling group

1.1 Open the AWS console, and search and click on **EC2**

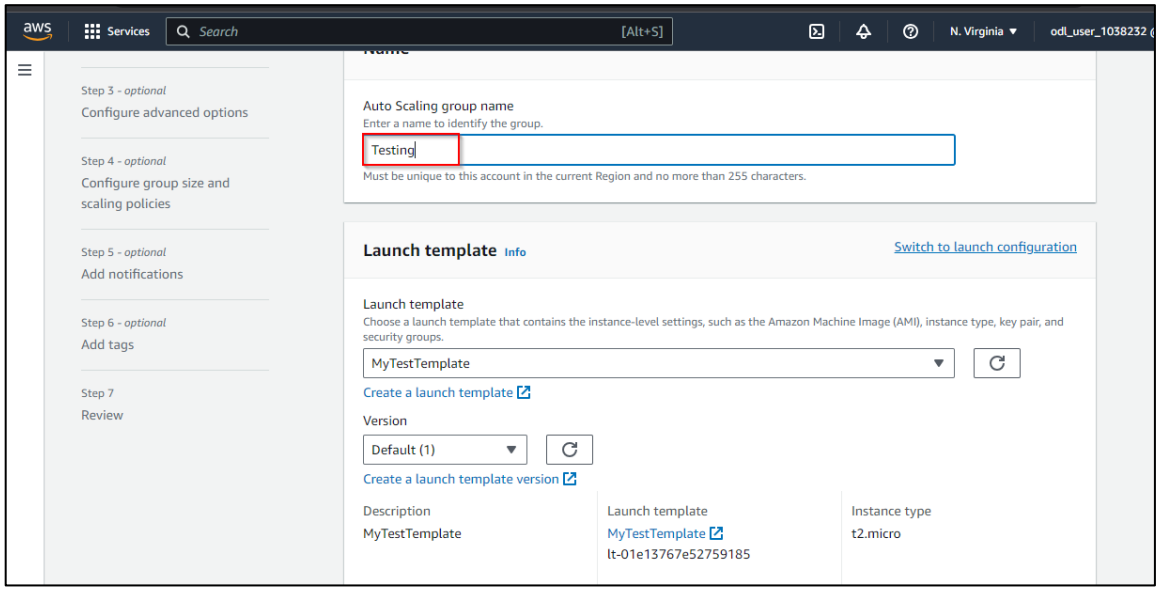


Set the region **US East (N. Virginia) us-east-1** in all the demos.

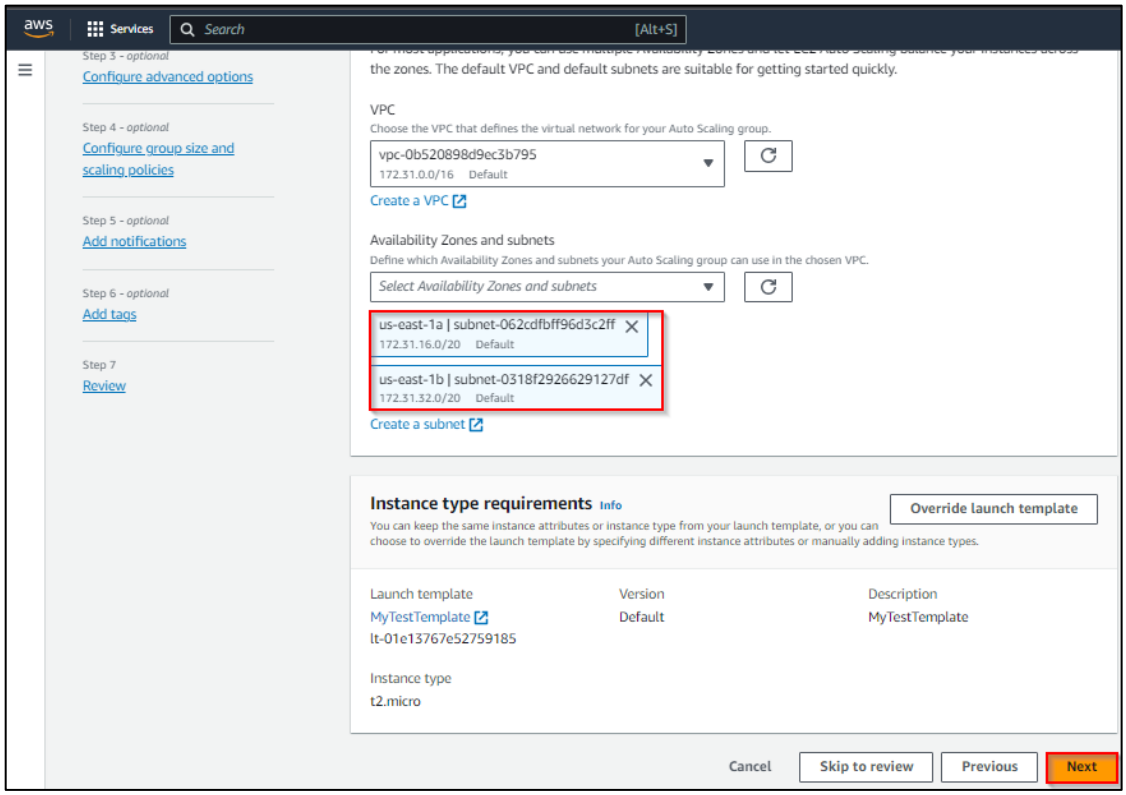
1.2 Click on **Create Auto Scaling group**



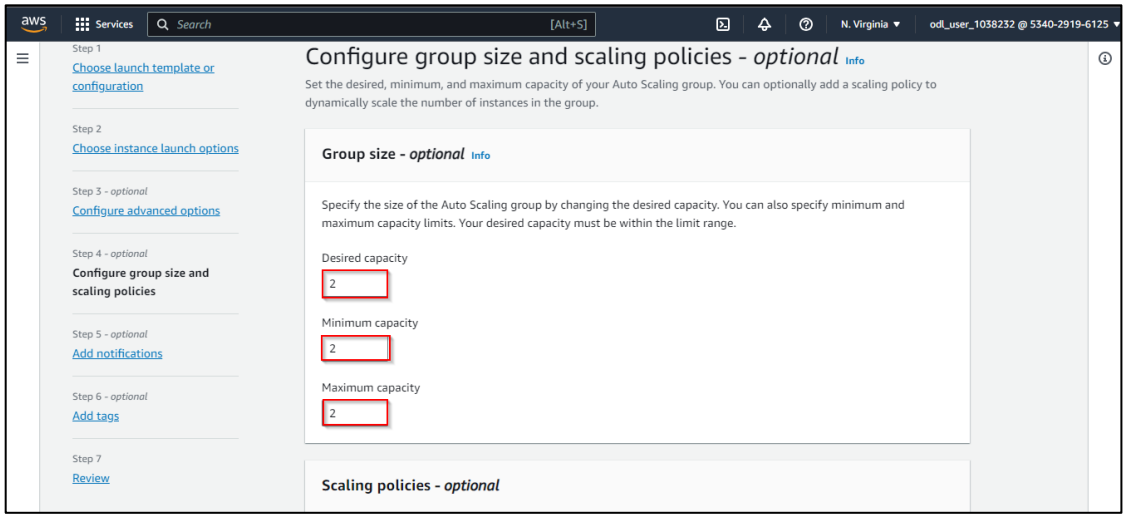
1.3 Enter the name as **Testing**, select **MyTestTemplate** in the Launch template, and click on **Next**



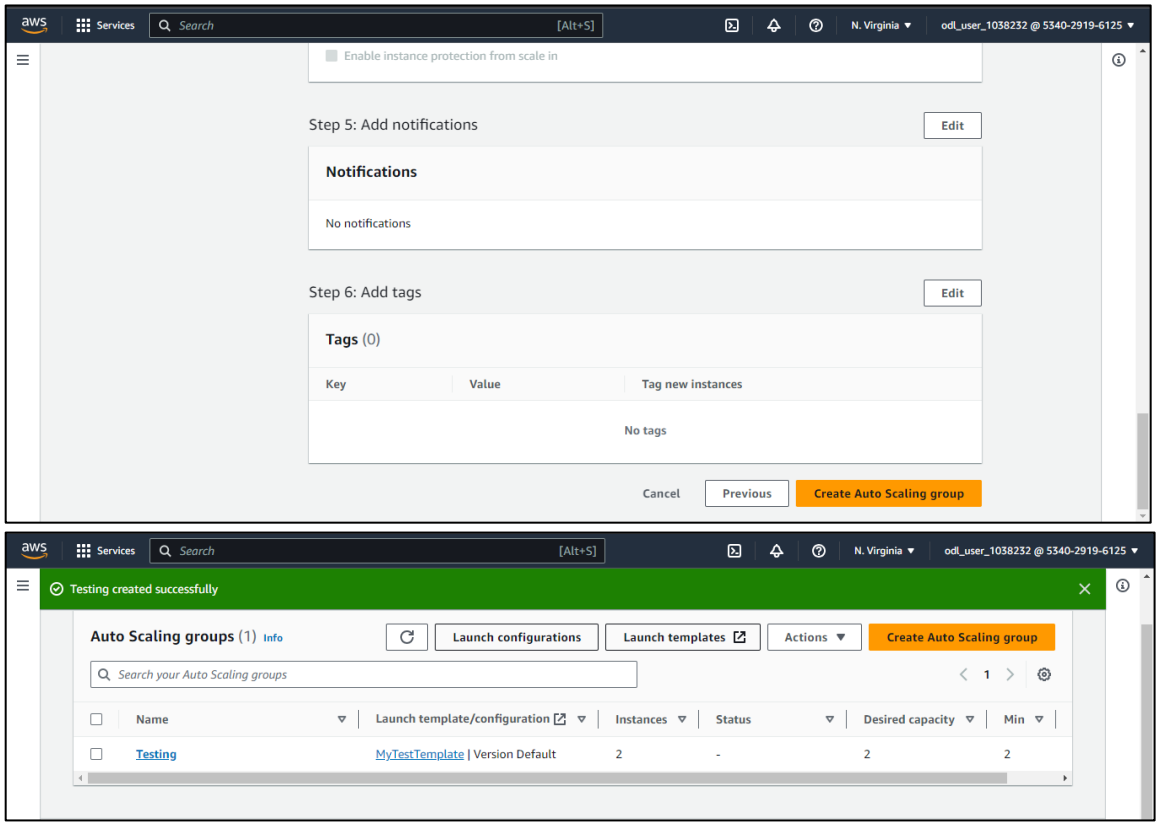
1.4 Click on the Availability Zones and subnets as **us-east-1a** and **us-east-1b,** click on **Next**



1.5 Enter the **Desired capacity**, **Minimum capacity**, and **Maximum capacity** as **2**, and click on **Next**



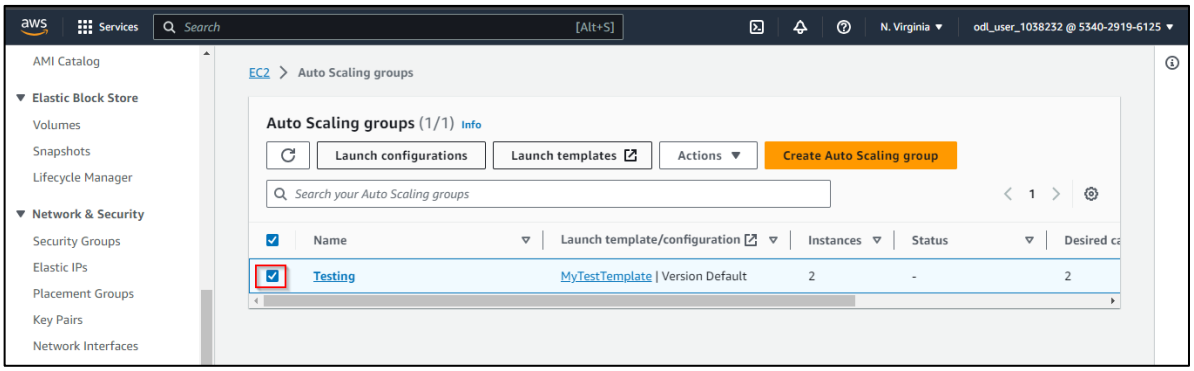
1.6 Review the steps, and click on the **Create Auto Scaling group**



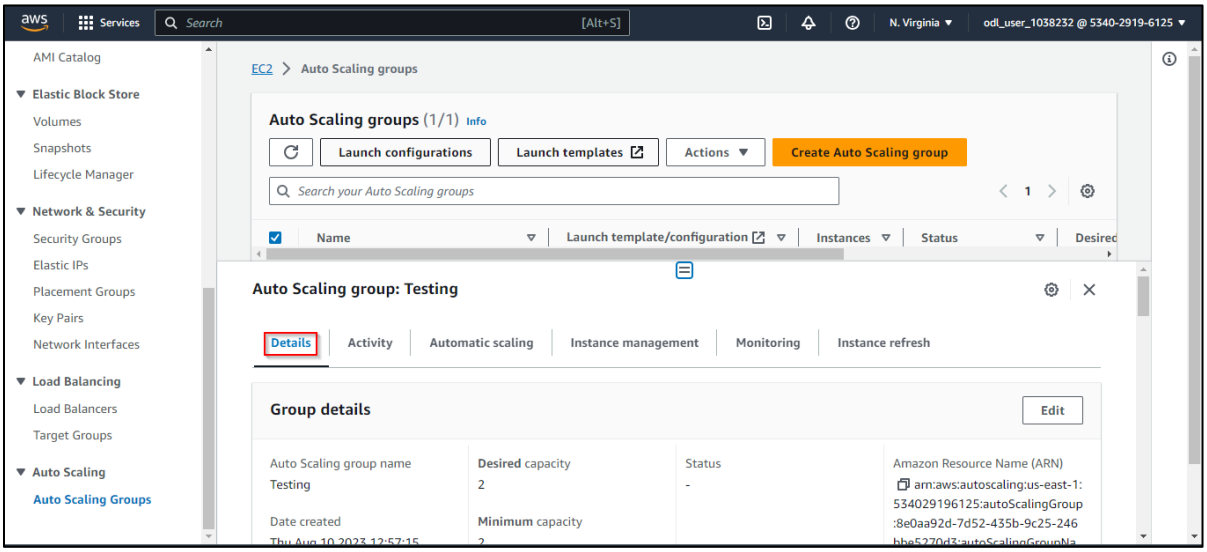
Auto scaling groups are created successfully.

Step 2: Set up EC2 Auto Scaling with Load Balancer

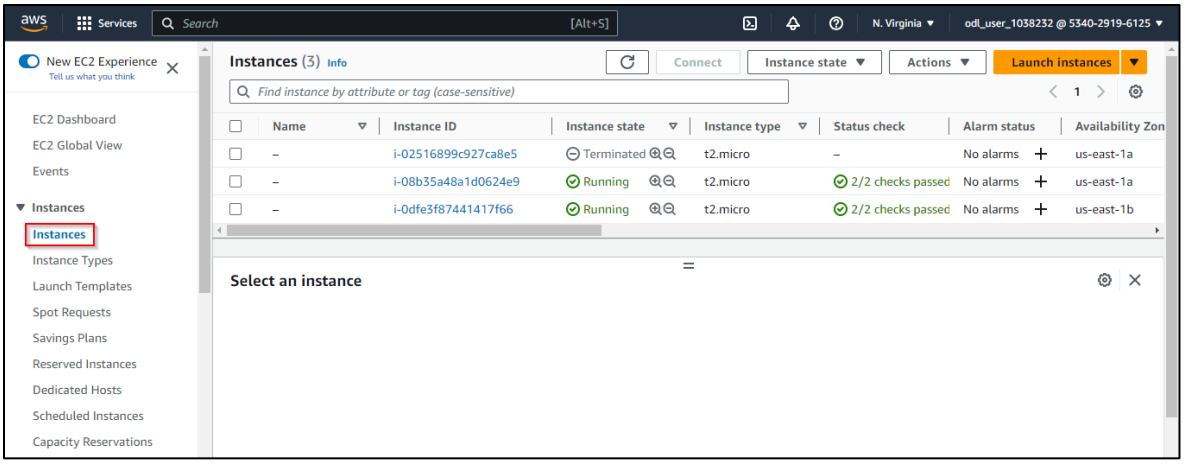
2.1 Open **the Auto scaling groups** section created previously



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

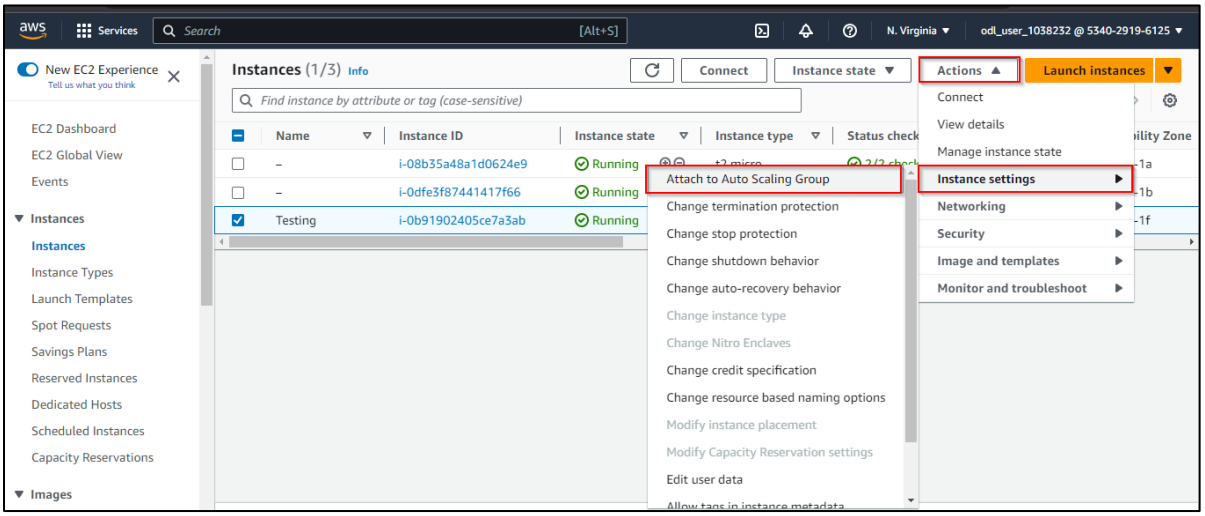


2.3 Click on **Instances**, and then create a new instance by clicking on Launch instances

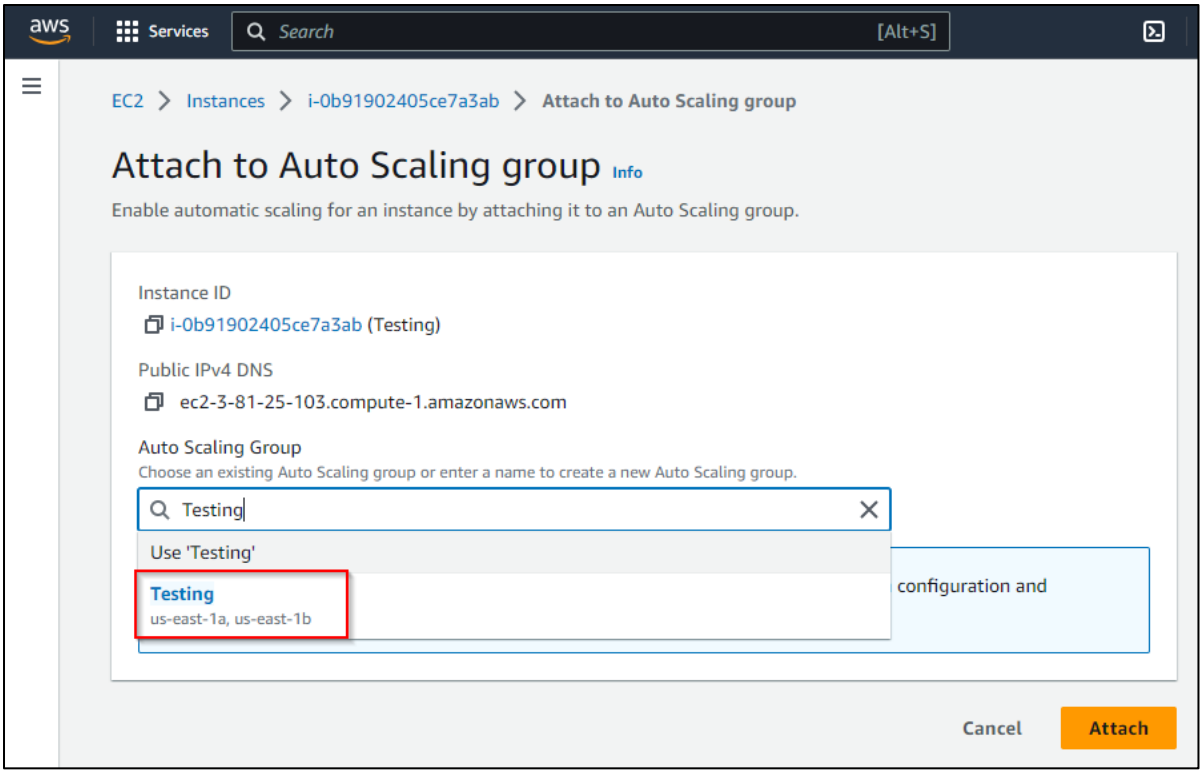


Instances are created successfully now; check out the previous demos to learn how to create instances.

2.4 Click on **Actions** > **Instance settings** > **Attach to Auto Scaling group**



2.5 Select the Auto Scaling Group name, **Testing**; click on the **Attach** button



Use different Auto Scaling group names.

We navigated through the process of configuring both manual and dynamic scaling techniques within the AWS workspace.

By following these steps, you will be able to successfully configure manual and dynamic scaling techniques. They provide the flexibility to manage resources efficiently, ensuring that your application remains responsive and available to users while adapting to changing traffic patterns and demands.