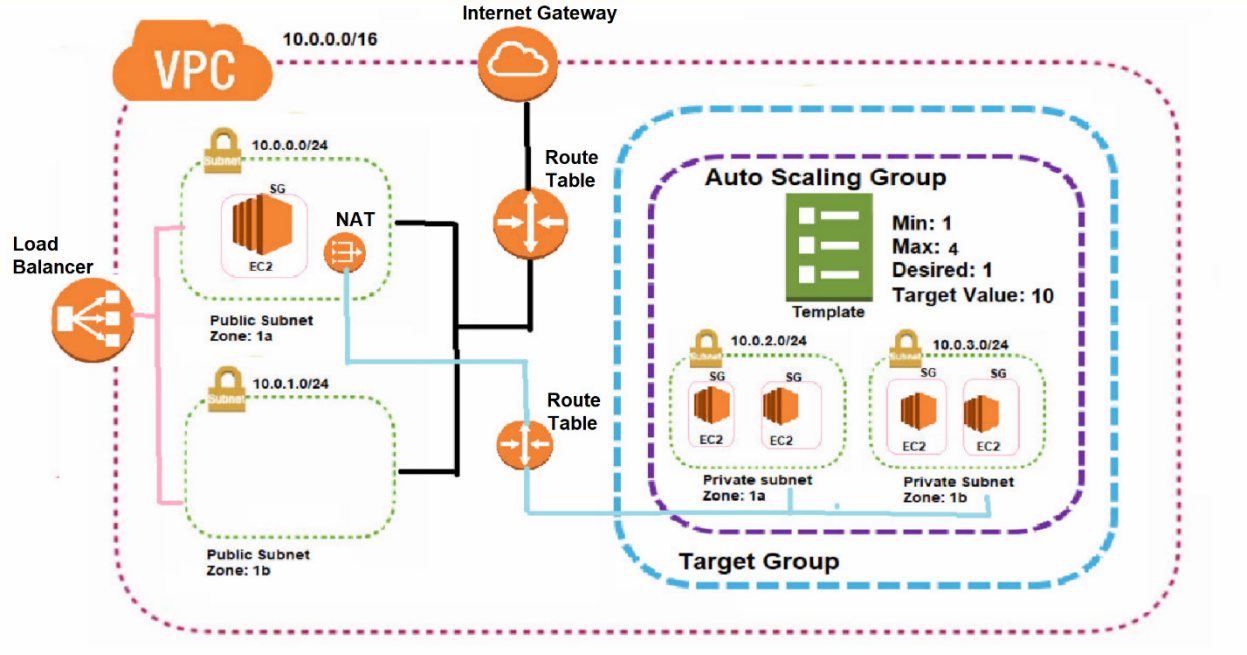
Notes-AutoScaling

STEPS TO CONFIGURE AUTO SCALING

This document deals with how to work with autoscaling group

Steps to Configure Auto Scaling

Diagram



High-level steps

1. Create a VPC.
2. Create an Internet Gateway and attach it to the VPC.
3. Create two public subnets: one in Availability Zone 2a and another in 2b. Also, create a private subnet in 2a.
4. Create a route table
5. Associate the Internet Gateway, and link it to the public subnets in Availability Zones 2a and 2b.
6. Create a NAT Gateway using the public subnet in AZ 2a.
7. Create a route table for the private subnet in AZ 2a
8. Associate it with the NAT Gateway and private subnet 2a
9. Create a security group with inbound rules for SSH and HTTP access.
10. Launch an EC2 instance in the public subnet in AZ 2a and another EC2 instance in the private subnet in AZ 2a.
    a. Access instance created on private subnet through instance created on public subnet.
    b. Install Apache webserver on private server using below commands
        i. *sudo su*
        ii. *sudo yum install httpd -y*
        iii. *sudo systemctl start httpd*
        iv. *sudo systemctl enable httpd*
        v. *cd /var/www/html*

vi. *vi index.html*
vii. *add blow html code*
<h1>Welcome </h1>
viii. *ESC then :wq*
ix. *Hit enter*

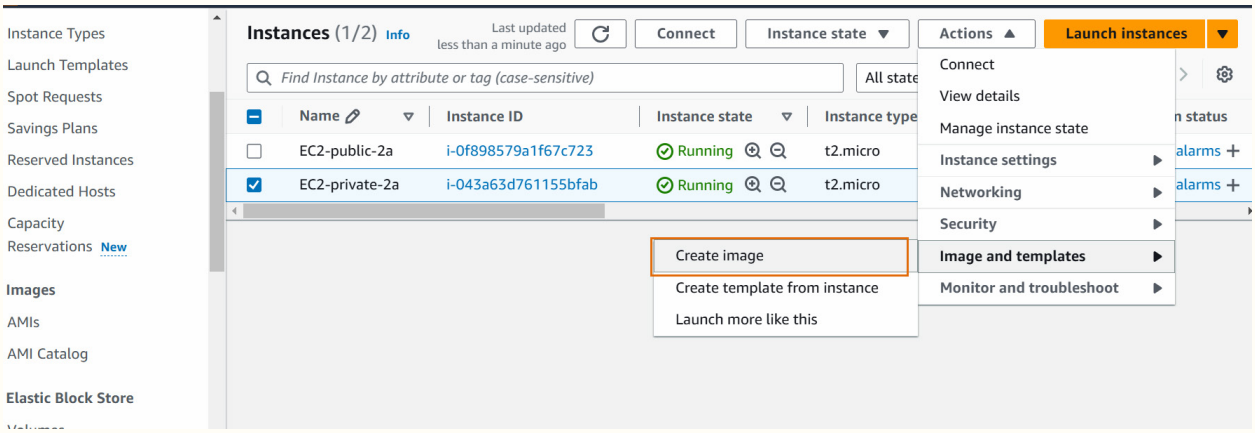
11. Create a target group and add the private instance in AZ 2a.
12. Set up a Load Balancer with Availability Zones 2a and 2b.
13. Copy the DNS name of the Load Balancer and open it in an internet browser.
14. If all configurations are correct, the index home page should be displayed.

Once above steps completed and successfully able to access the index home page using DNS name of the load balancer, please proceed with below steps.

15. Create an AMI image using an instance in the private subnet within Availability Zone 2a.
16. Create a template that includes the AMI image, instance type, and key pair information.
17. Set up an Auto Scaling Group using the created template.
    a. **Specify the private subnets,**
    b. **Attach to existing load balancer – (Your Target Group)**
    c. **Group size (desired capacity: 1)**
    d. **Scaling information (minimum: 1, maximum: 3),**
    e. **Target tracking scaling policy (target value: 10 for testing),**
    f. **Tag informatio**n (key: Name, value: ASGInstance).
18. Verify that one instance is created and initiated through Auto Scaling.
19. For testing, adjust the target tracking policy on the Auto Scaling Group by changing the target value to 0.1 to observe the creation of multiple instances.

Below screenshots are related to from step 15

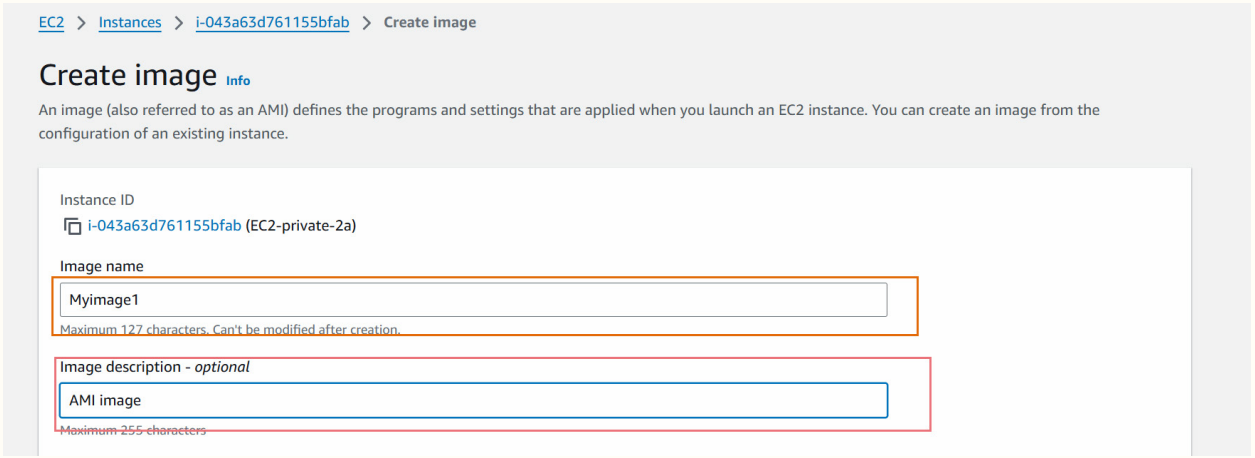
Screenshots for AMI Creation Process



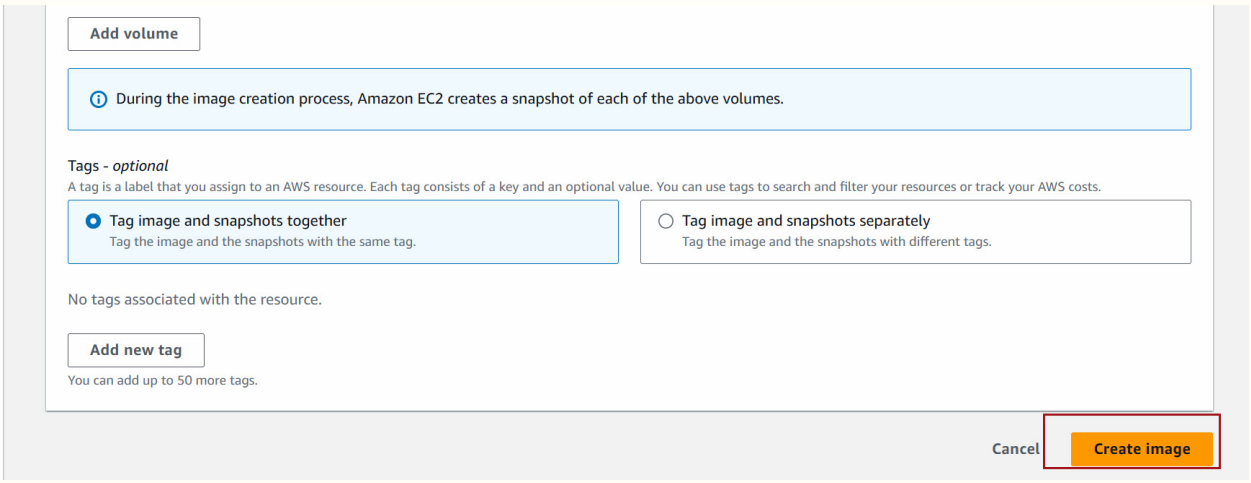
Configure Image Settings:

Image Name: Enter a descriptive name for your AMI.

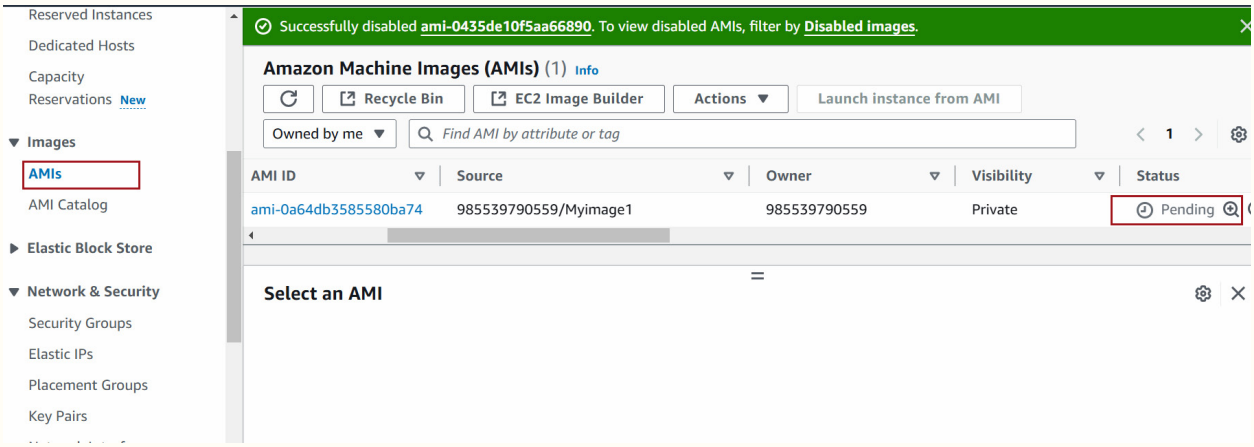
Image Description (optional): Provide a brief description for future reference



Create the Image: Click **Create Image** at the bottom of the page. AWS will initiate the AMI creation process.



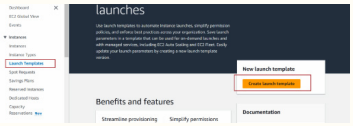
Monitor Image Creation: In the left navigation pane, go to **AMIs** to monitor the status of your new image. It will appear with the status **Pending** and switch to **Available** once ready.



Once available, you can proceed to next steps

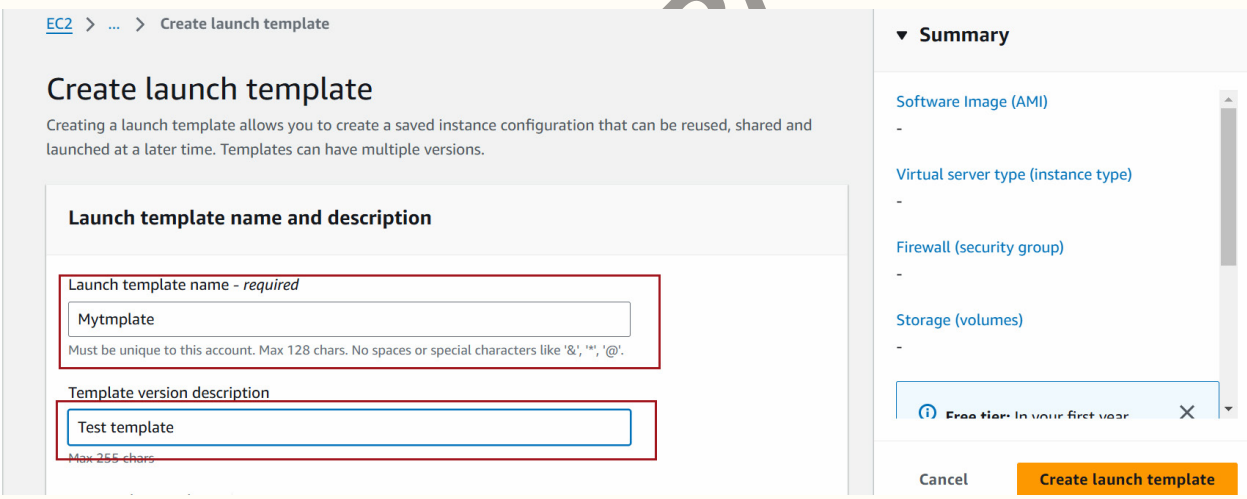
Screenshots for Template creation process

On the left sidebar, scroll down to **Instances** and click on **Launch Templates**. Click **Create launch template**.

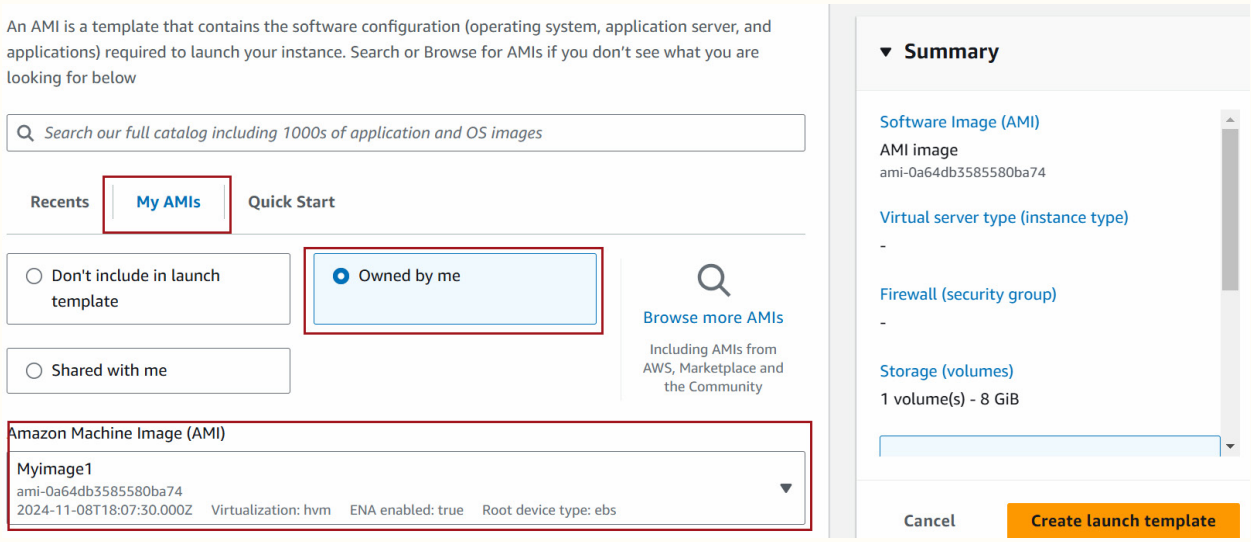


Fill out the fields for the **Launch Template**:

- **Template Name**: Provide a name for your template (e.g., MyTemplate).
- **Version Description** : Optional description for the version.

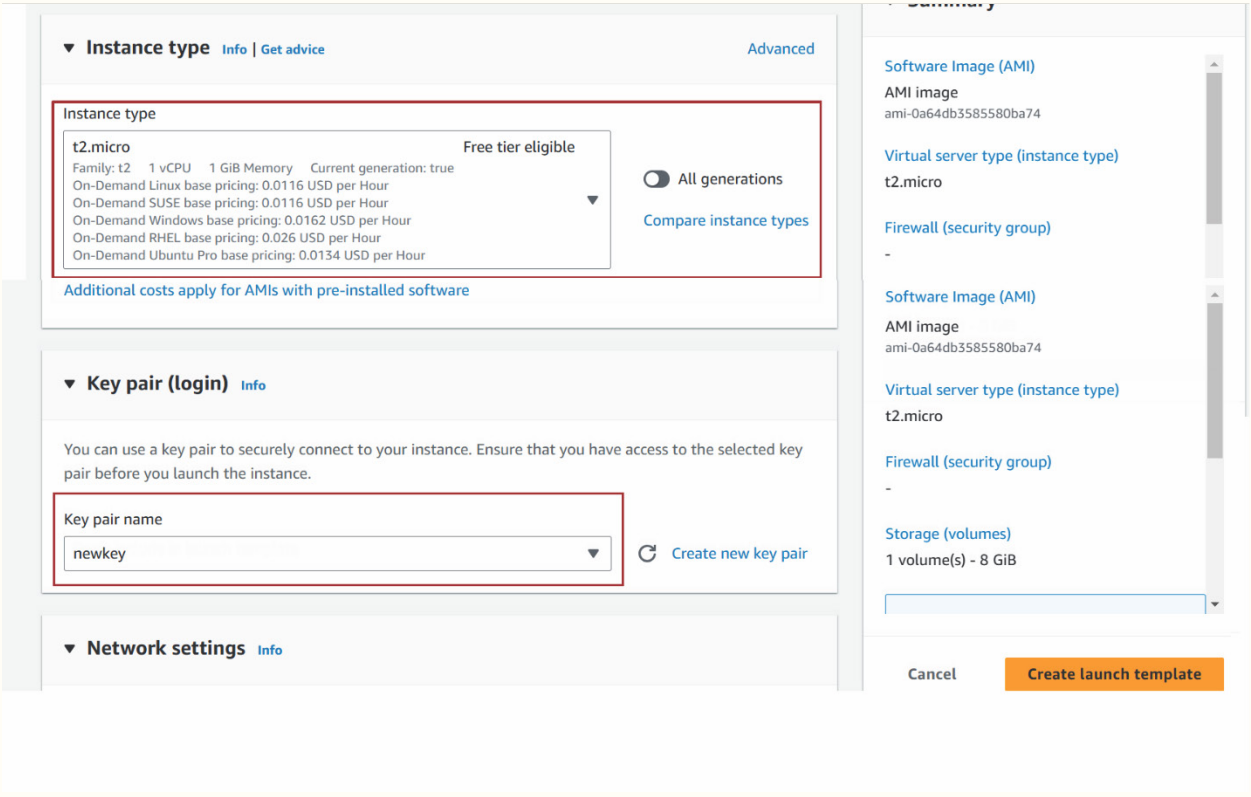


AMI: Select the AMI ID you want to use

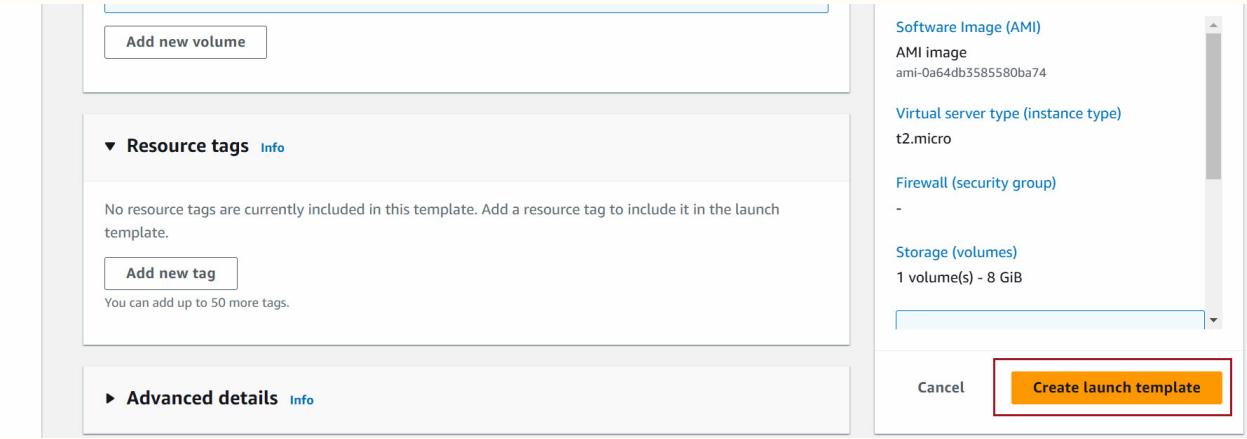


Instance Type: Choose the instance type (e.g., t2.micro).

Key Pair: Select an existing key pair or create one

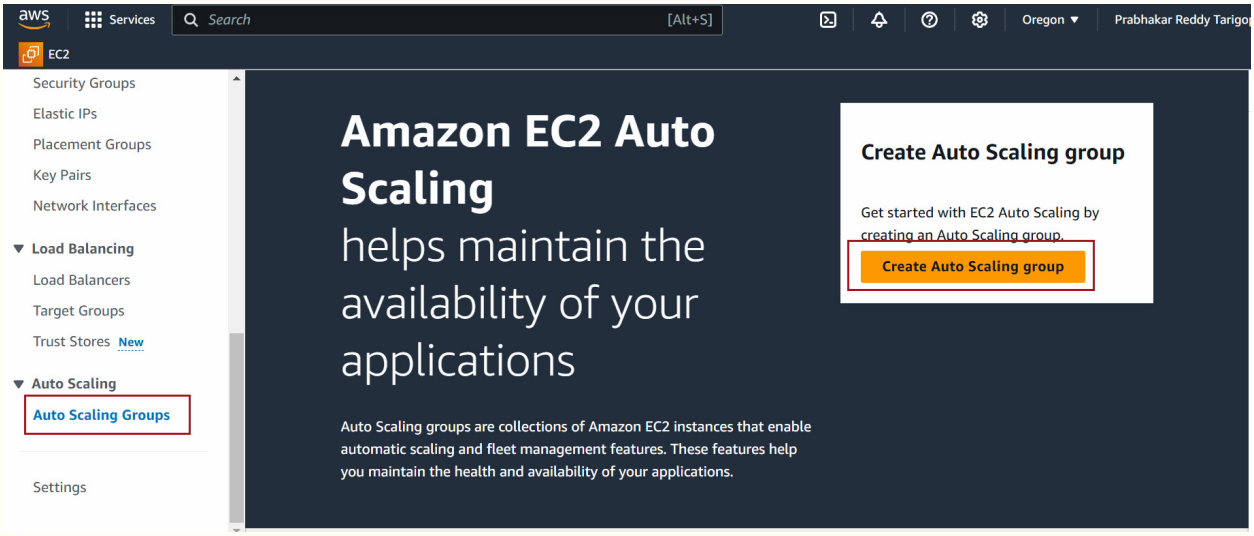


click **Create launch template**. This template is now available for launching instances or referencing in other templates.

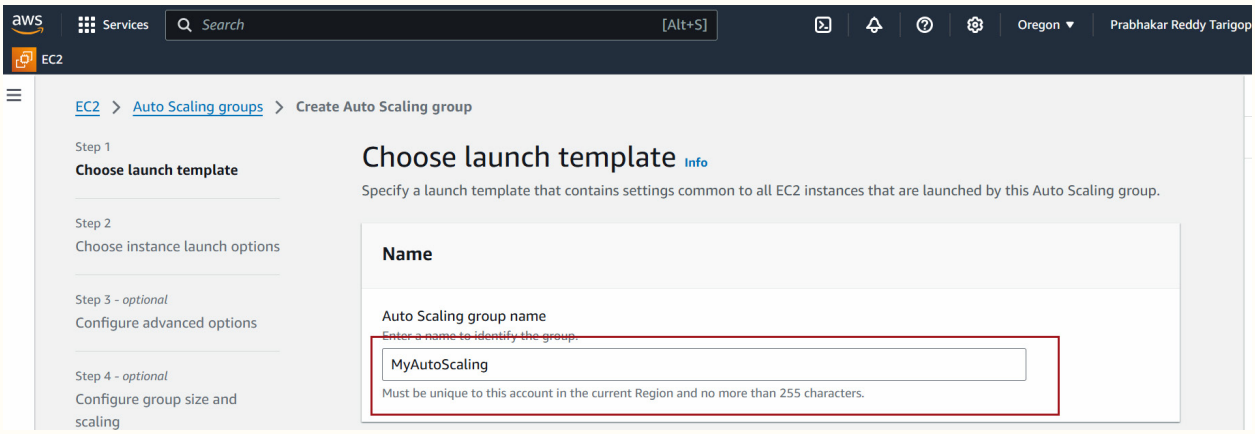


Screenshots for Auto Scaling Groups creation process

In the left-hand menu, scroll down to **Auto Scaling** and select **Auto Scaling Groups**. Click on **Create Auto Scaling group**.

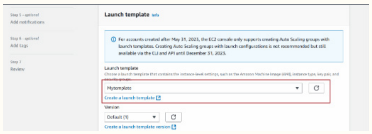


Specify the ASG Name - Enter a name for your Auto Scaling Group

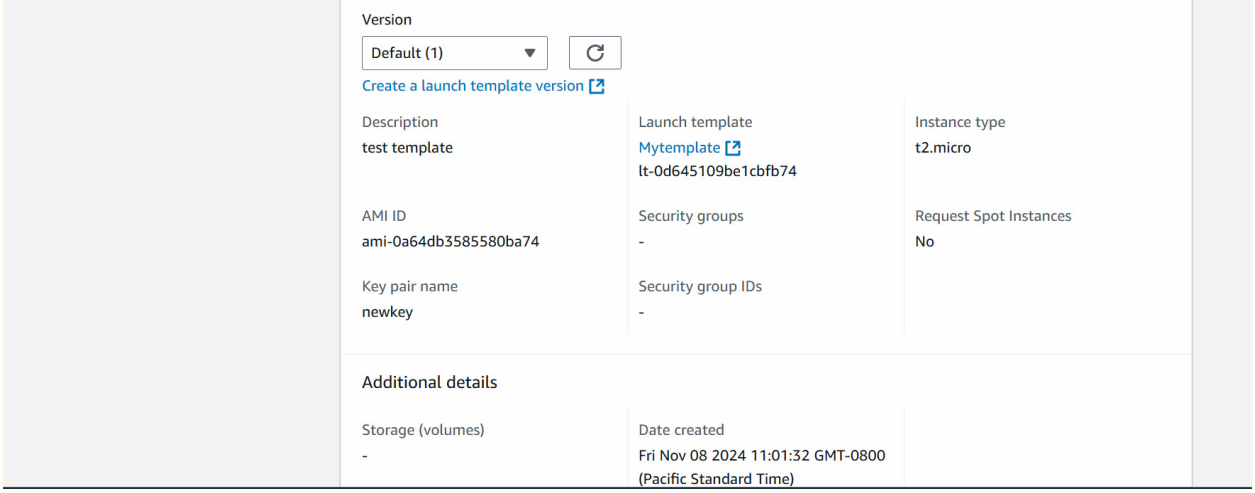


In the **Launch Template** section:

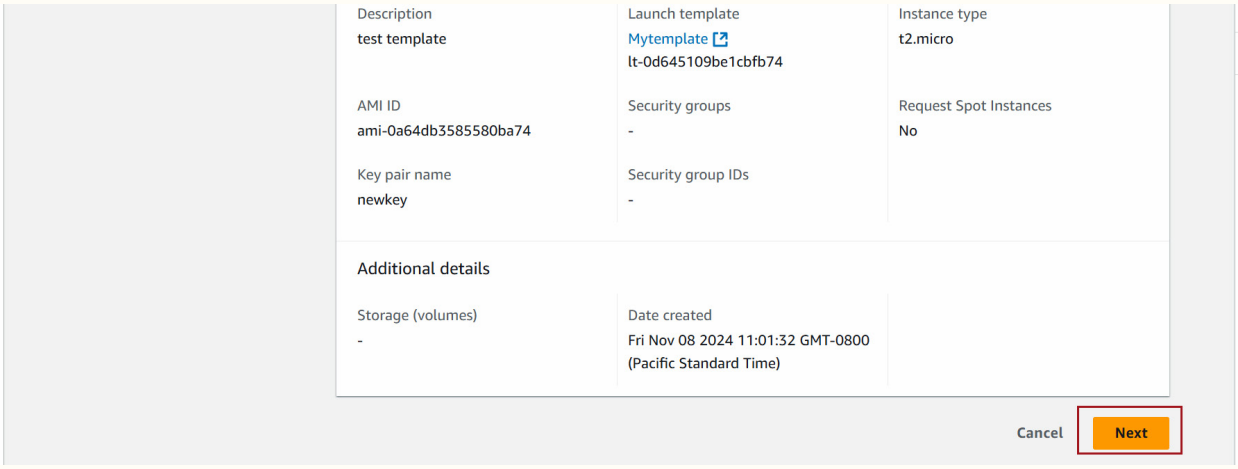
- Select **Launch Template**.
- Choose your existing **Launch Template** from the dropdown menu.



Verify the details from the template



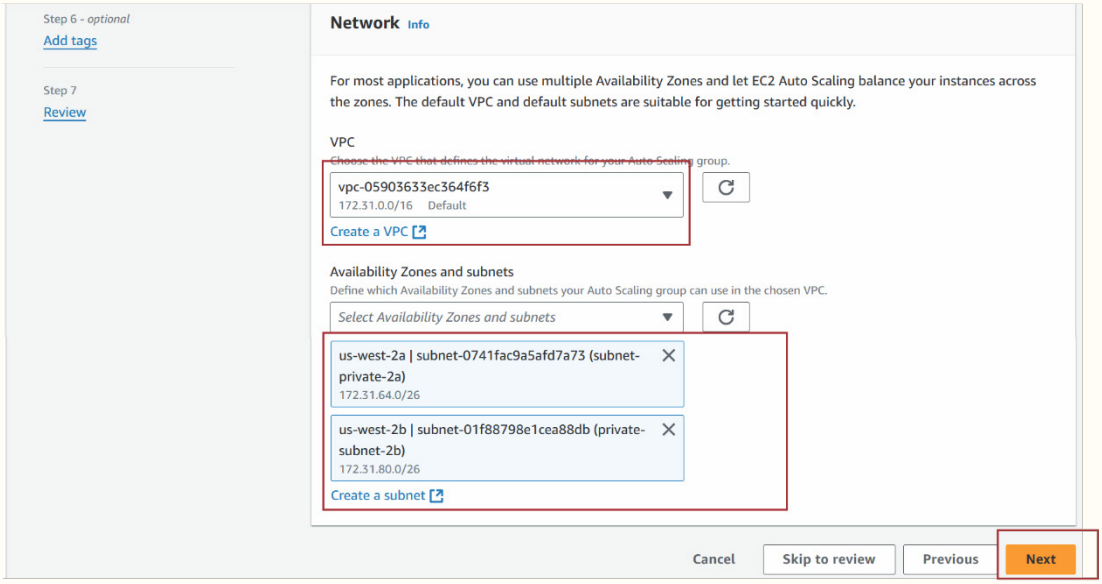
Click Next



VPC and Subnet Selection:

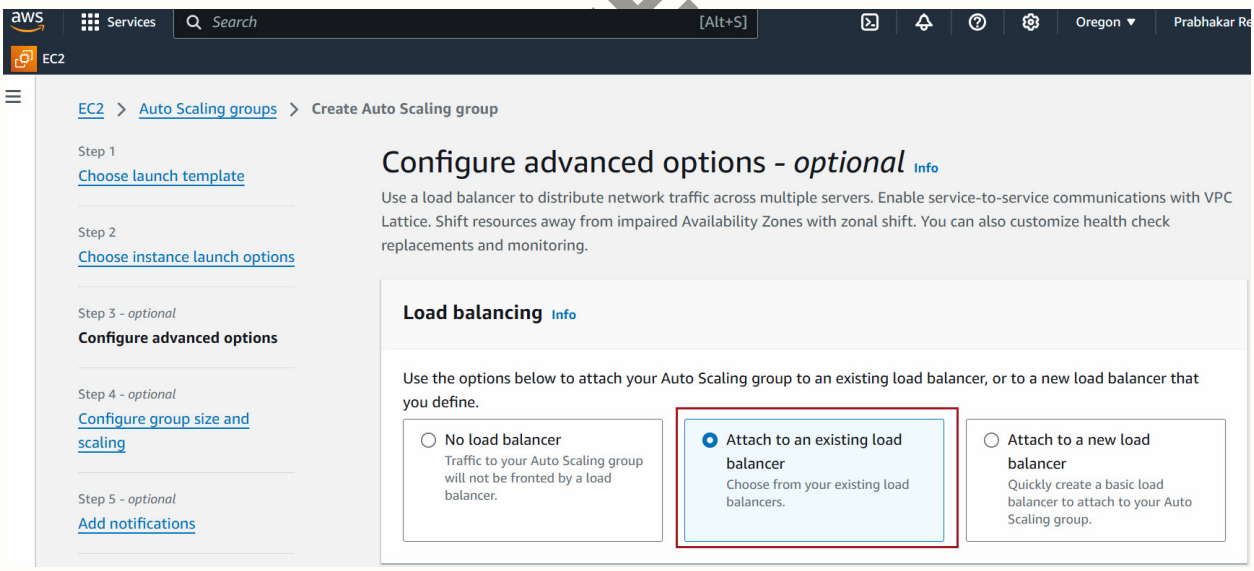
Choose the **VPC** and **subnets** where your ASG should launch instances. Select multiple subnets (private subnets) to enable your ASG to span multiple Availability Zones.

Click Next

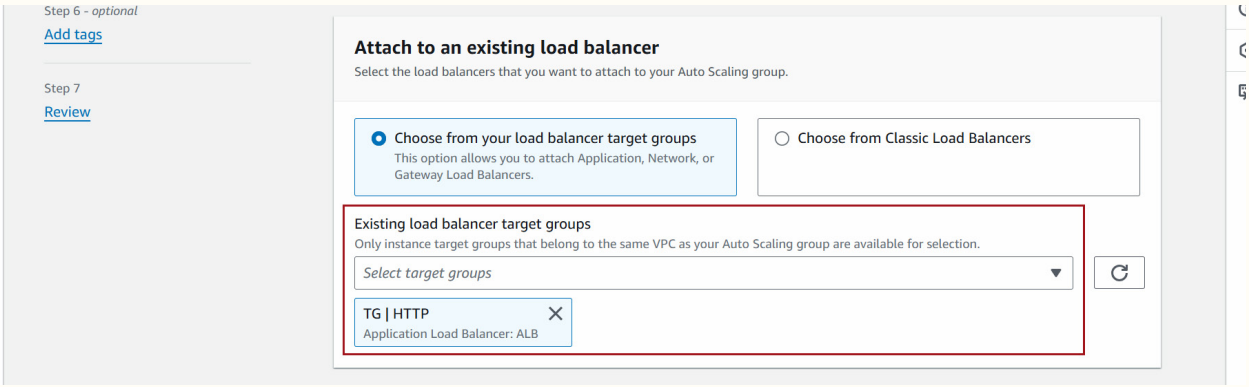


Load balancing Selection:

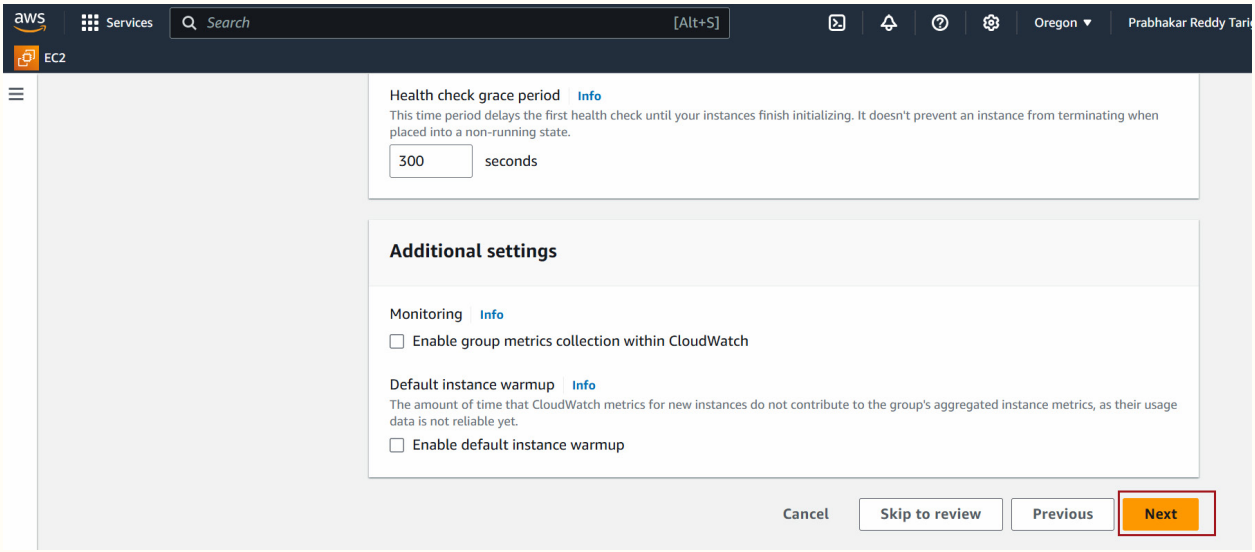
Choose Attached to an existing load balancer



Choose existing load balancer target group

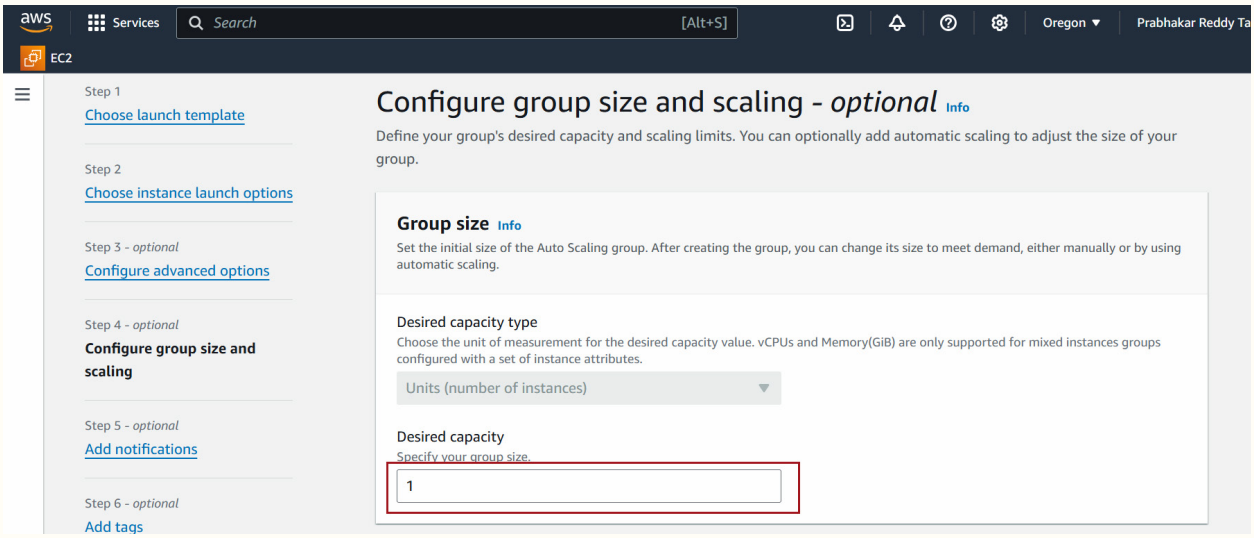


Click Next

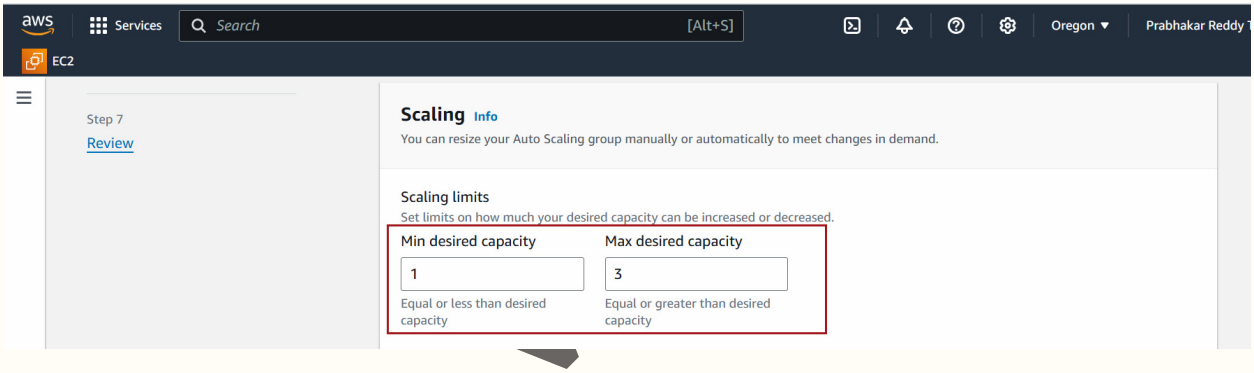


Instance Scale Options:

Choose the number of instances you want to start with in the **Desired Capacity** field.



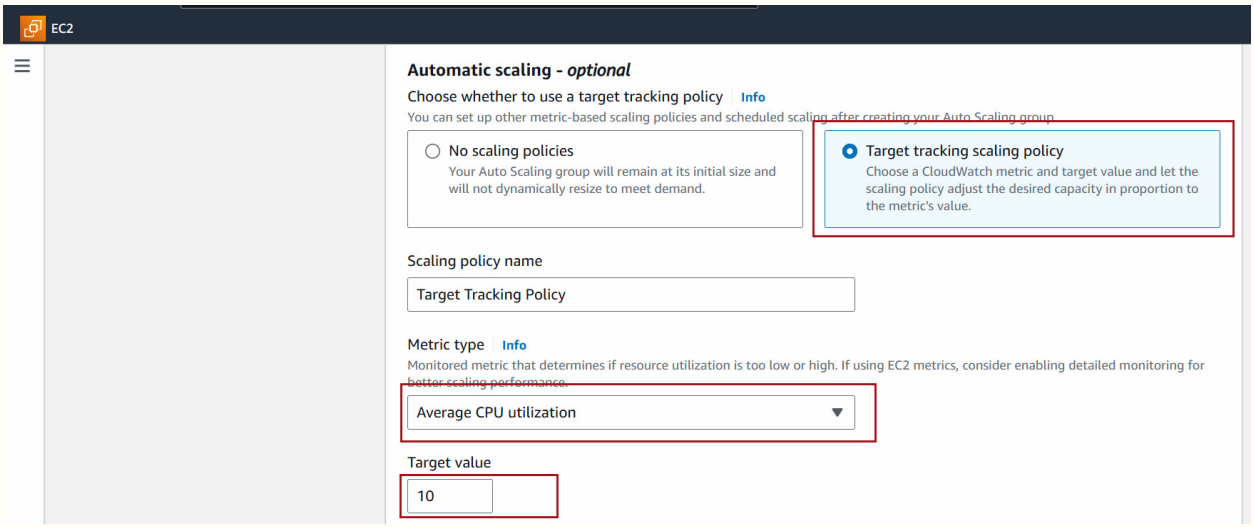
Define **Minimum** and **Maximum Capacity** based on your scaling requirements.



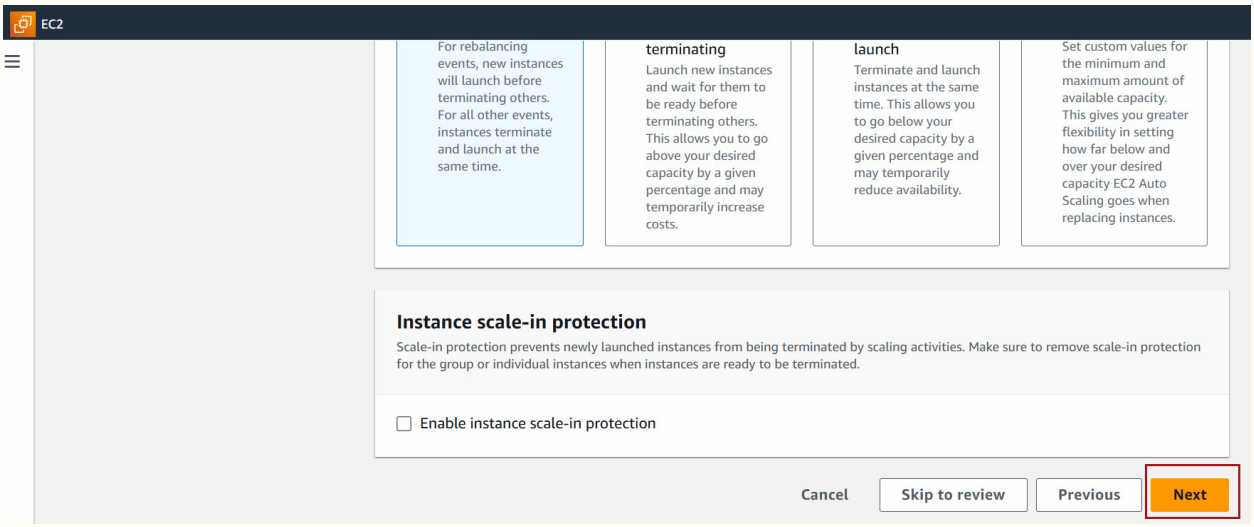
Configure Scaling Policies (Optional)

In the **Scaling Policies** section, you can configure policies to scale in response to demand.

- **Target Tracking**: Set a target metric

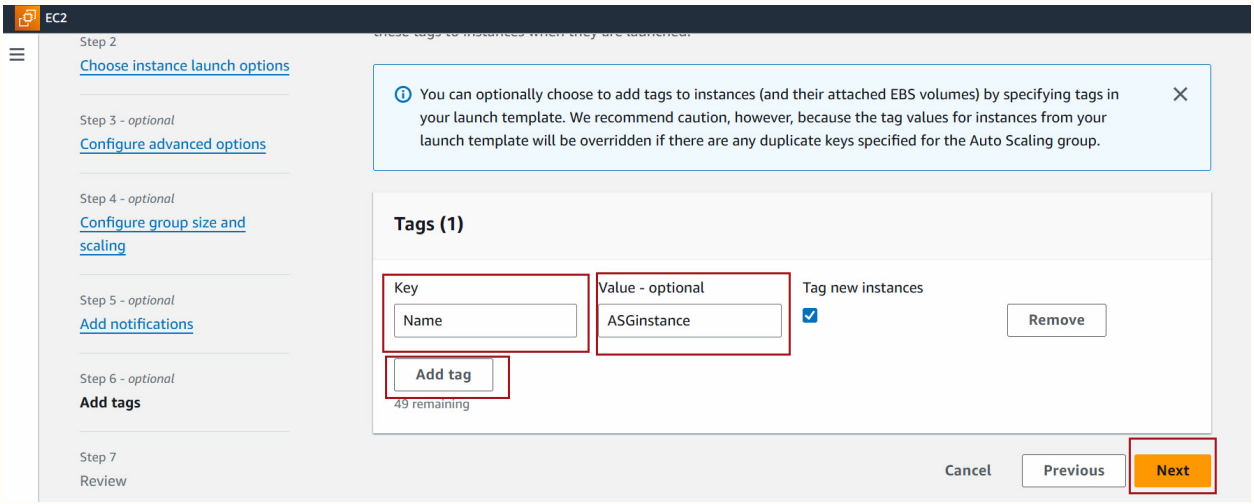


Click Next



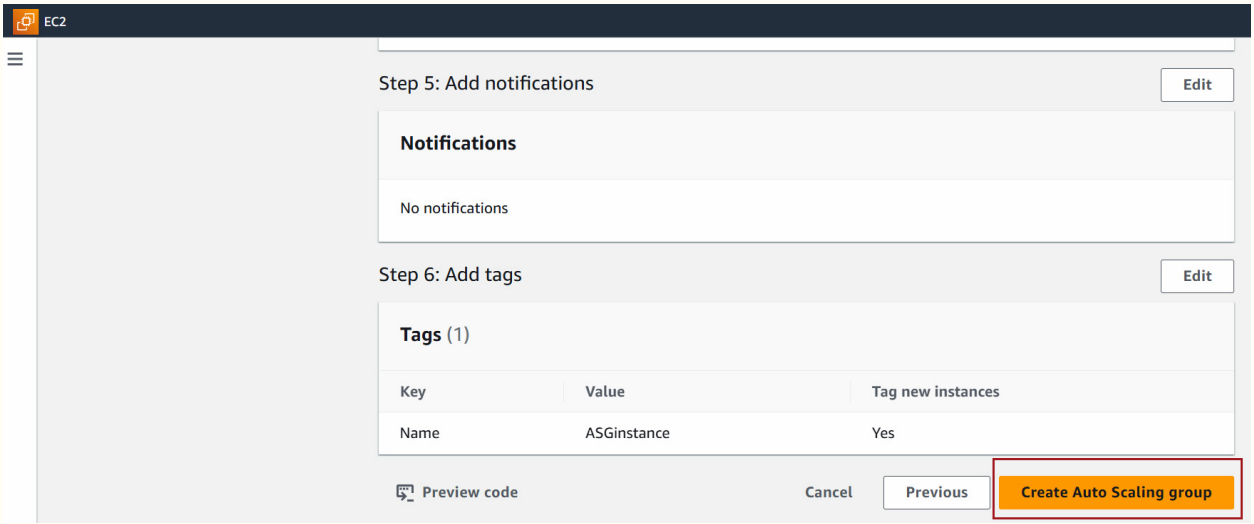
Configure Tags

Add tags for easier identification of resources. For instance, you might tag with **Name** etc.



Review and Create

- Review all the settings to ensure everything is correctly configured.
- Click **Create Auto Scaling group** to launch your ASG.



Verify ASG Creation

- After creating the ASG, it will appear in your **Auto Scaling Groups** list.

- The ASG will start launching instances based on your specified settings and scaling policies.