

**Placement Empowerment Program**

***Cloud Computing and DevOps Centre***

Implement Auto-scaling in the CloudSet up an

autoscaling group for your cloud VMs to handle

variable workloads

**Name: Kirthika S Department: CSE**

A black and white logo

Description automatically generated

**Introduction**

As modern applications face varying workloads, ensuring optimal

performance and availability is critical. Auto Scaling, a feature

provided by cloud platforms like AWS, dynamically adjusts

computing resources in response to demand changes. This Proof of

Concept (PoC) demonstrates how to set up an Auto Scaling Group

(ASG) for virtual machines (VMs) to handle fluctuating workloads

effectively. It explores defining launch configurations, setting scaling

policies, and testing automatic scaling based on CPU usage.

**Objective**

The primary objective of this PoC is to:

1. Implement an **Auto Scaling Group (ASG)** to manage

workloads effectively.

2. Define and configure a **Launch Template** for virtual machines.

3. Set up and test **scaling policies** based on predefined metrics,

such as CPU utilization.

4. Validate the scaling process by simulating real-world scenarios

(e.g., high CPU usage).

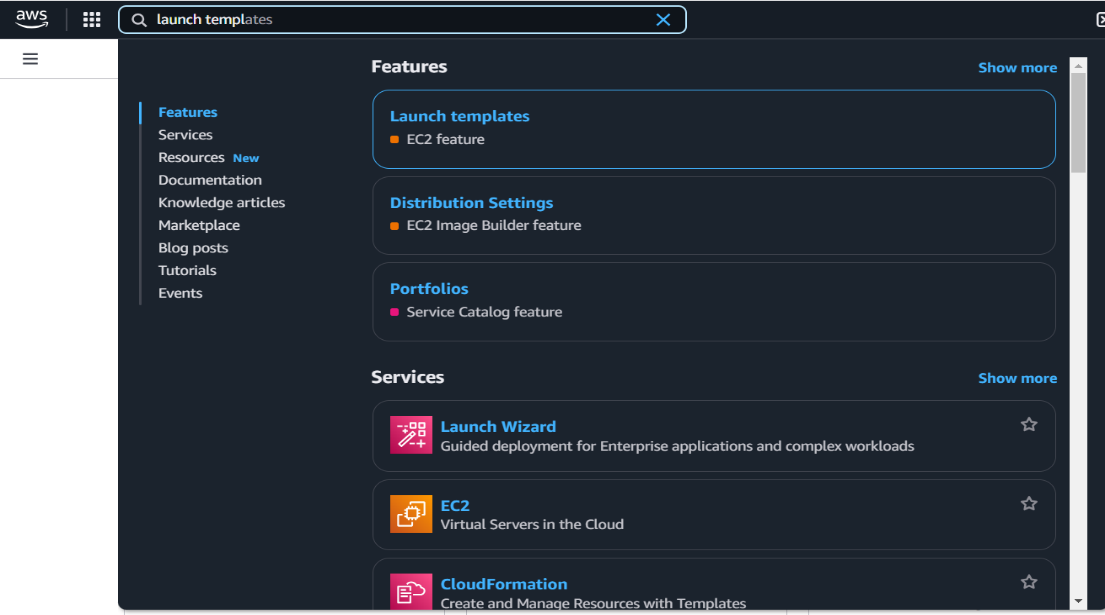
**Step 1:**

Go to AWS Management Console.

Enter your username and password to log in

Search for Launch Templates.

Click on the Create launch template



**Step 2:**

Create a **Launch Template** named **Demotemp** using an

**Amazon Machine Image (AMI)** like Amazon Linux 2 or any default

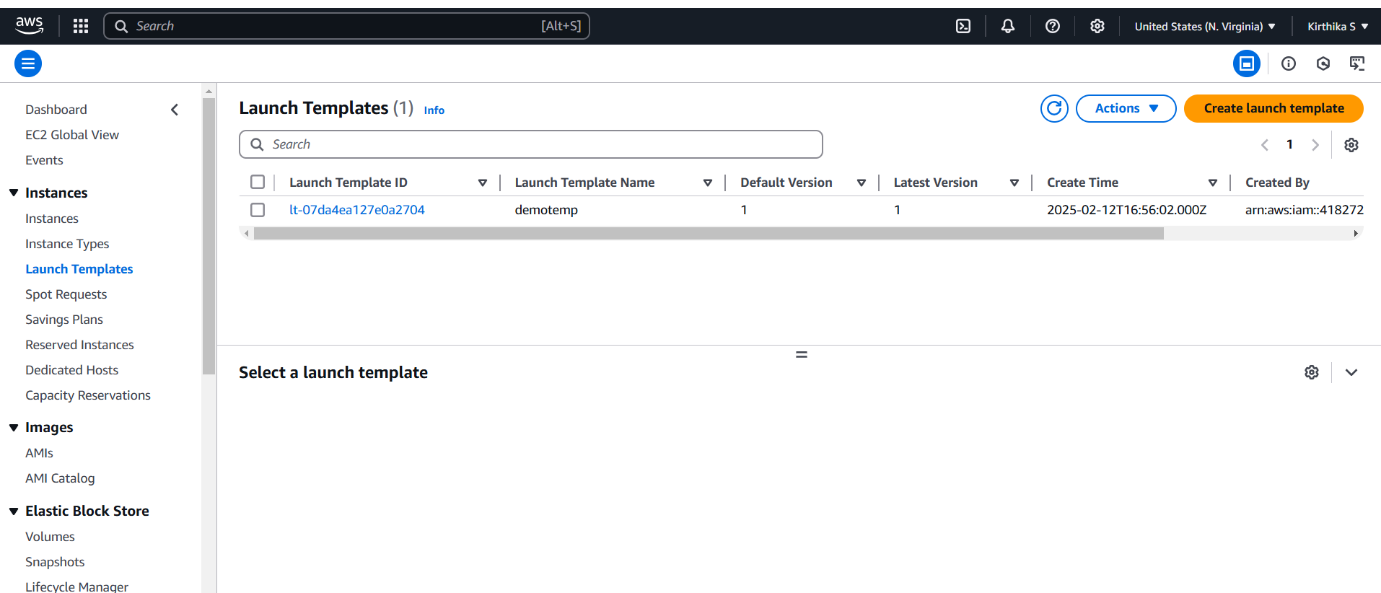
image, and choose an **instance type** such as **t2.micro** for free-tier

eligibility. Select an **existing key pair** (or create a new one) to enable

SSH access, and configure a **security group** that allows HTTP (port

80) and SSH (port 22). Once all details are filled out, click **Create**

**launch template** to complete the setup.



**Step 3:**

Go to the **EC2 Dashboard** . On the left sidebar, click on **Auto**

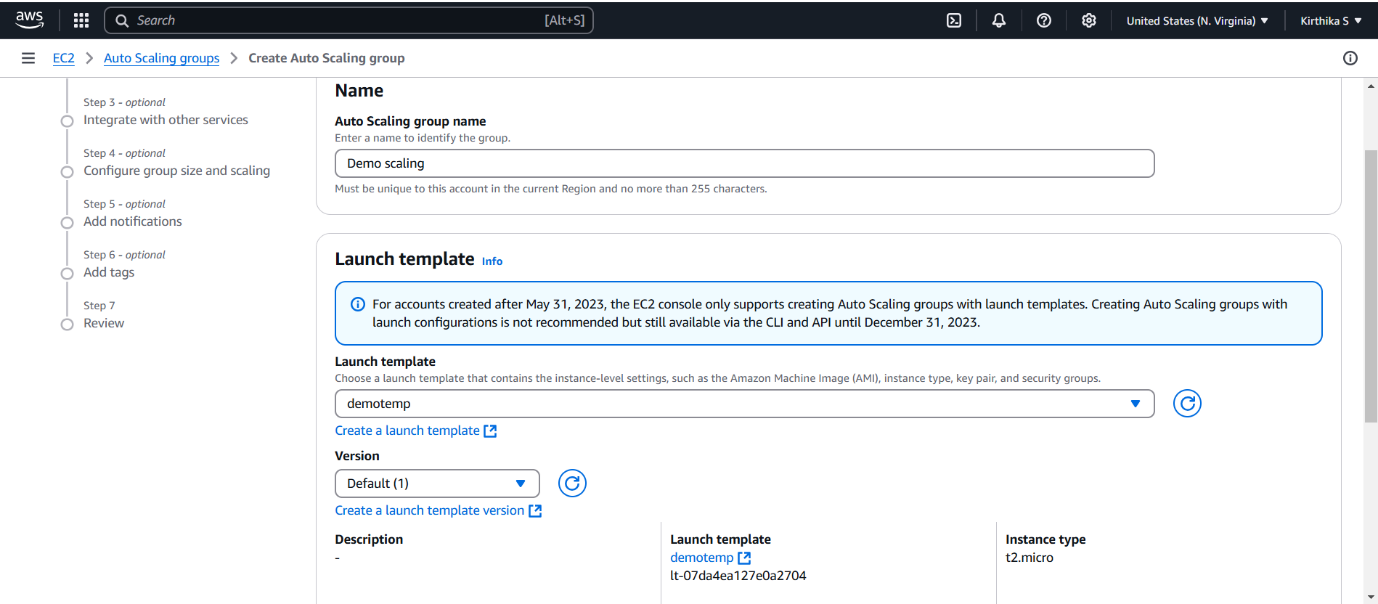
**Scaling Groups**. Click on **Create an Auto Scaling group**

**Auto Scaling group name**: Give it a name (e.g.,

Demo scaling).

**Launch Template**: Select the launch template you created

earlier (AutoScalingTemplate).

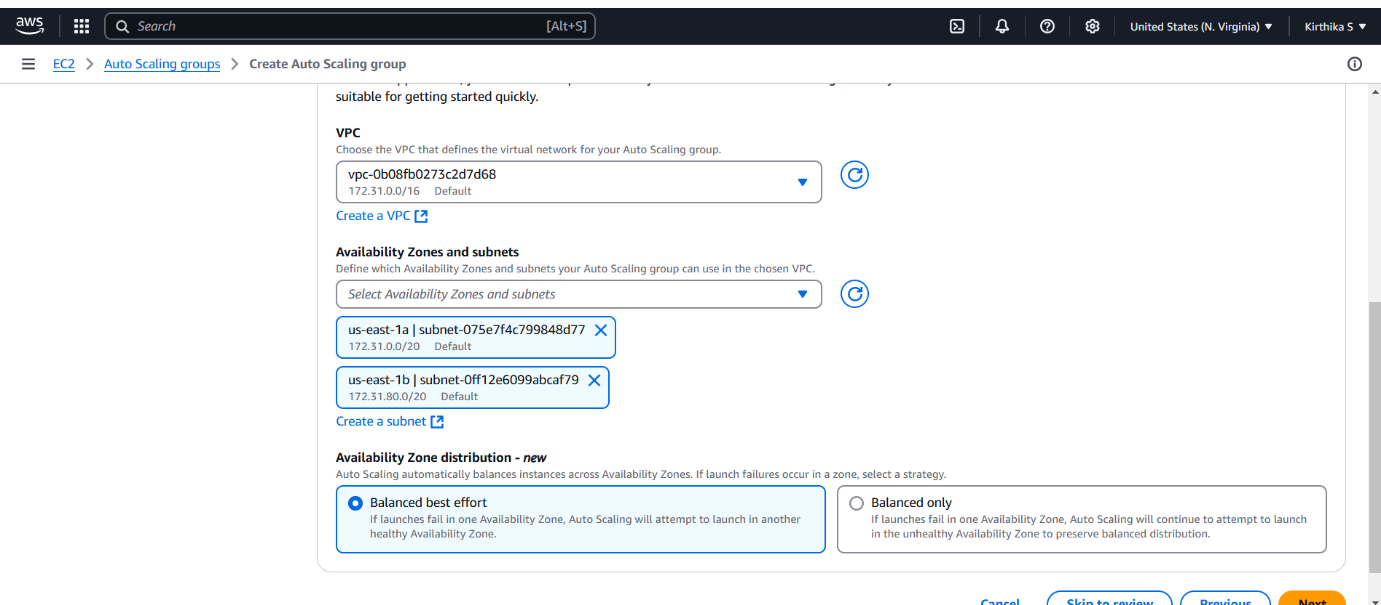


**Step 4:**

**VPC and Subnets**: Choose your **VPC** (it’s fine to use the default

one). Select at least two subnets in different Availability Zones (this

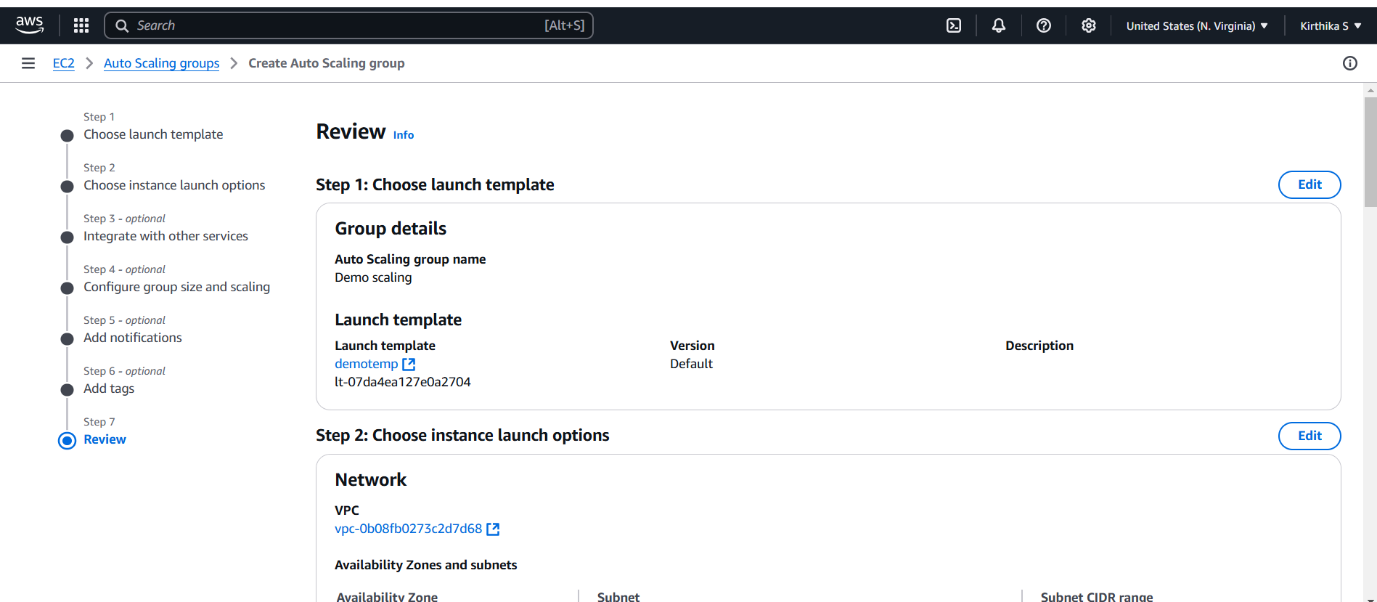
ensures high availability).



**Step 5:**

For this PoC leave the next settings as default and click next .

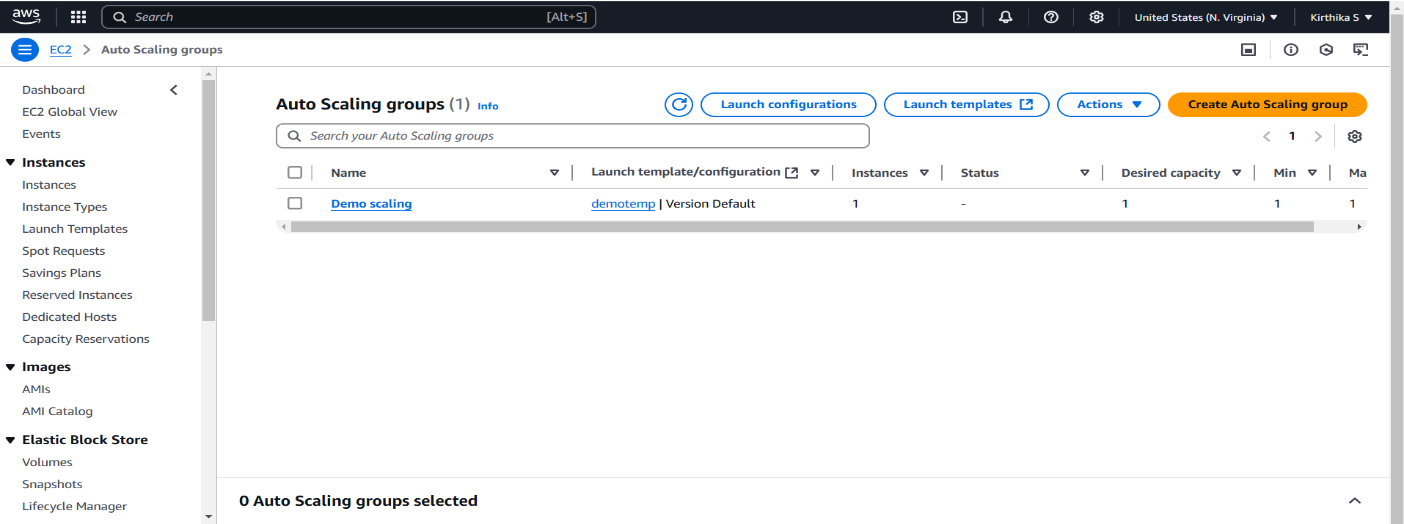
Review all the settings you've configured

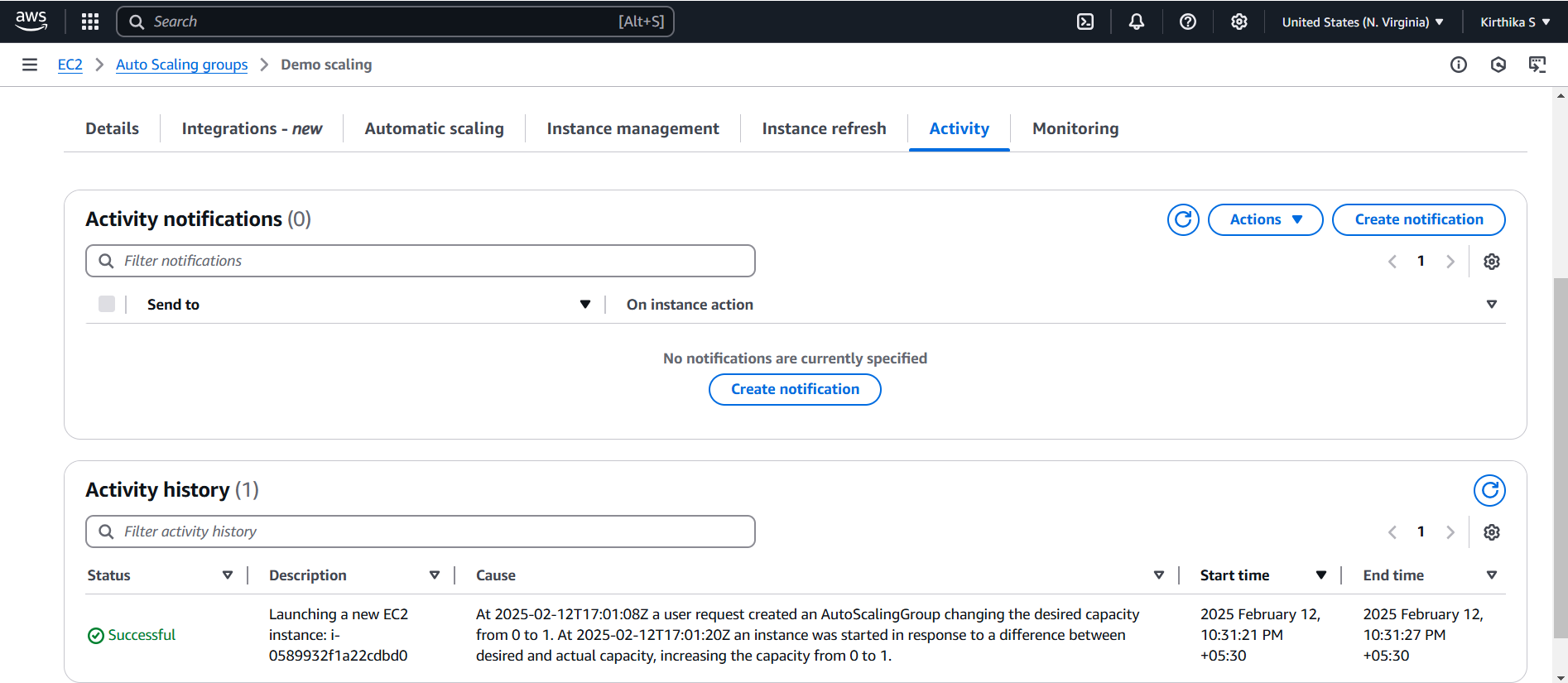


**Step 6:**

Once satisfied, click

**Create Auto Scaling Group**.





**Outcome**

This Proof of Concept (PoC) aimed to implement Auto Scaling in

AWS to dynamically manage EC2 instances based on workload

demand, ensuring efficient resource utilization and cost-effectiveness.