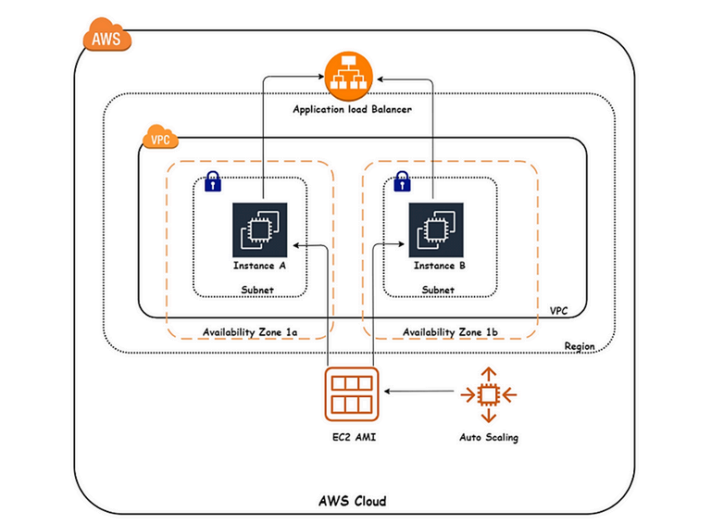
**Lab 4: Configure Load Balancer and Auto Scaling on EC2**

**Objective:**

Learn to configure Load Balancers and Auto Scaling on EC2 to enhance application performance and resilience.

**Architectural diagram:**



**Steps:**

* Launch two web servers in separate Availability Zones.
* Create an Application Load Balancer to distribute traffic.
* Set up and test Auto Scaling with simulated CPU load.

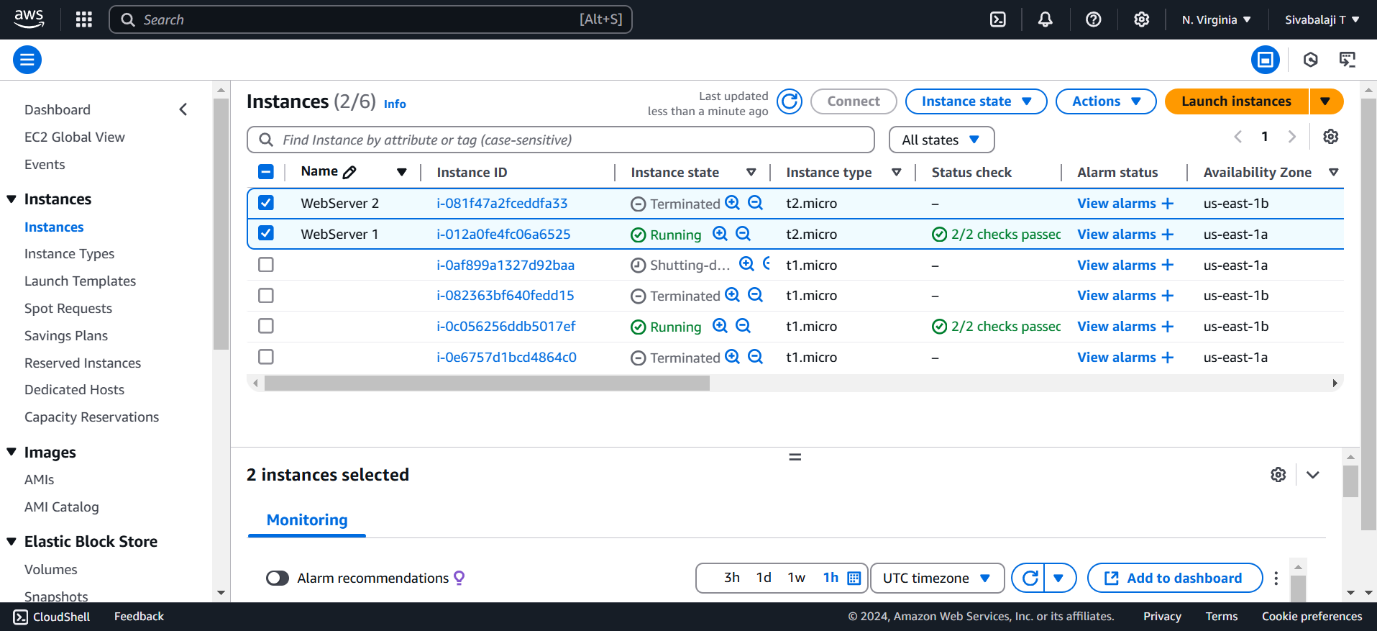
**Step-by-Step Instructions:**

**1. Prerequisites**

* An AWS account.
* Basic knowledge of EC2, Load Balancer, and Auto Scaling concepts.
* A configured VPC, Subnets, and Security Groups.

**2. Launch EC2 Instances**

* Launch Ec2 Instance in Two separates availability Zones.



* **Configure Security Group**:

Allow inbound HTTP (port 80) and SSH (port 22).

* **Launch the Instance**:

Select or create a key pair for SSH access.

**3.Install Apache Servers on the instaces which has in availability zone**

1.Update and install Apache or Nginx (e.g., for Apache):

* + sudo yum update -y
  + sudo yum install -y httpd

2.Start the service

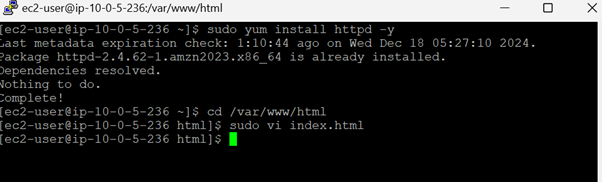
* + sudo service httpd start

3.Change the Directory to the html

* + cd /var/www/html

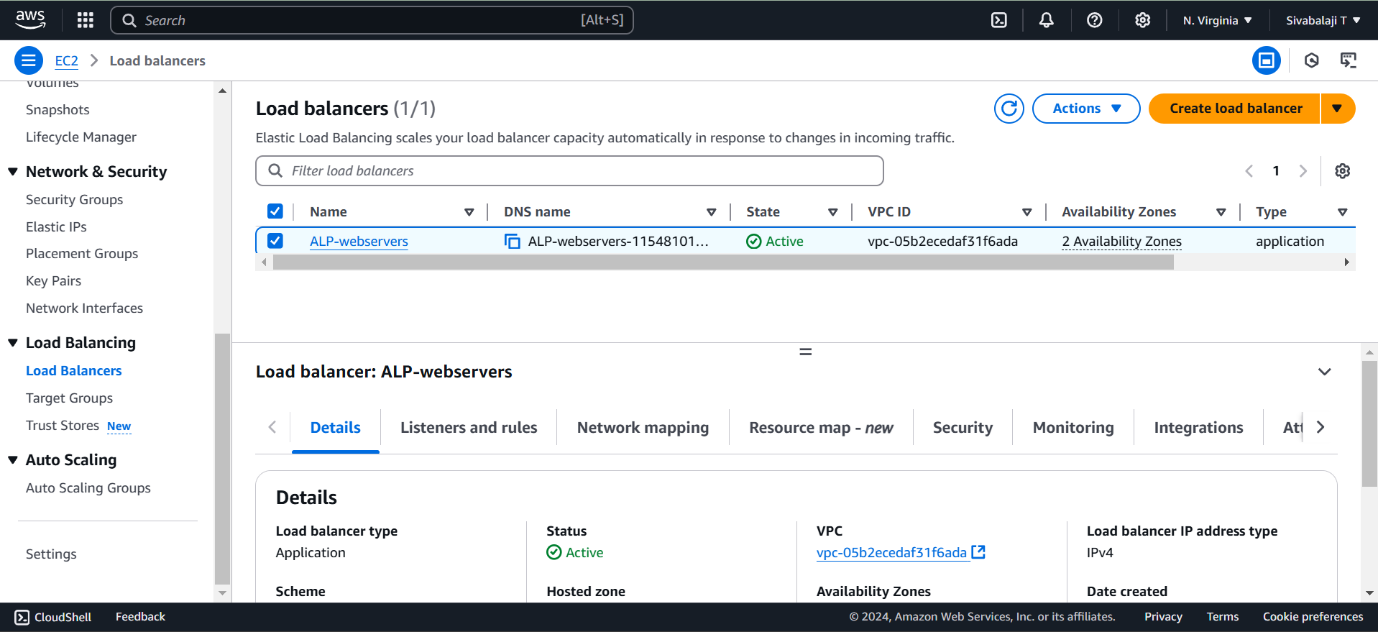
4.Open vi edior

* + sudo vi editor



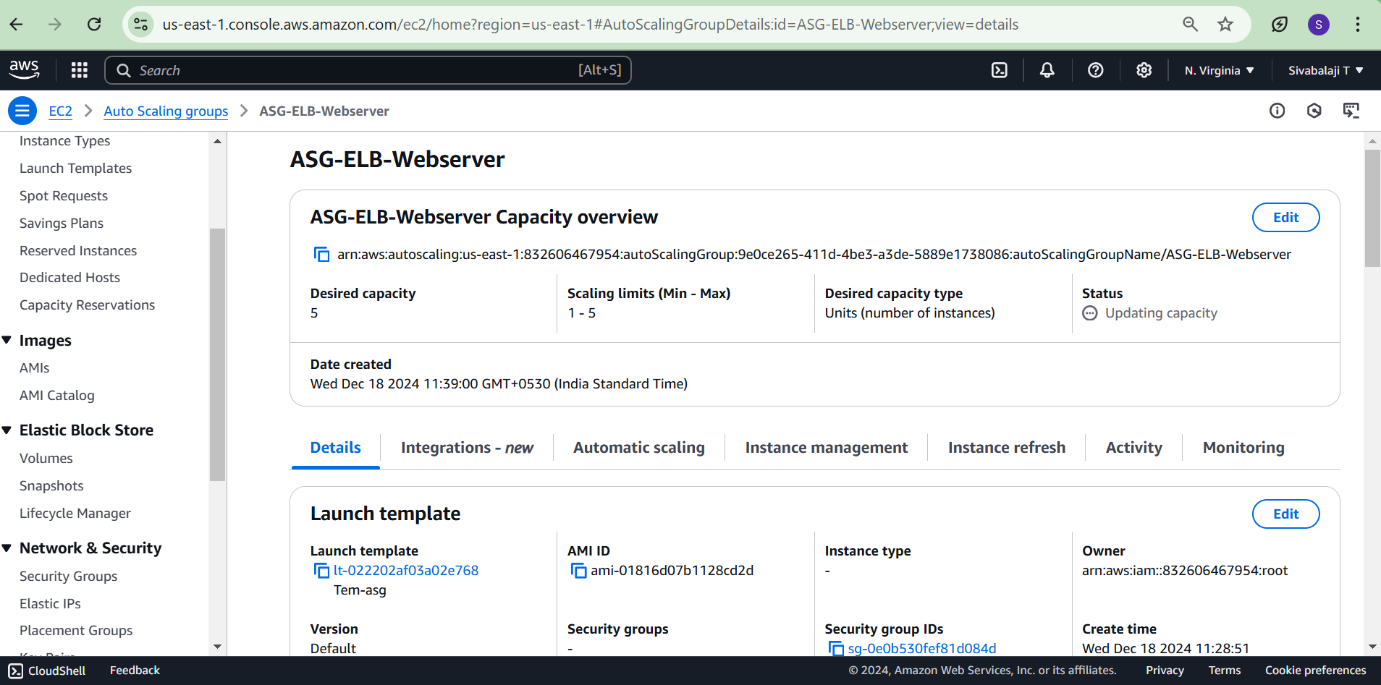


**4. Configure the Load Balancer**

1. **Go to the EC2 Console** > **Load Balancers** > **Create Load Balancer**.
2. **Select Application Load Balancer (ALB)**:
   * Choose the Internet-facing option.
   * Assign the Load Balancer to the same VPC and public subnets.
3. **Configure Security Group**:
   * Allow inbound HTTP traffic (port 80).
4. **Configure Target Group**:
   * Create a new target group (Instance type).
   * Register your EC2 instance.
5. **Configure Listener**:
   * Add a listener for HTTP (port 80) that forwards traffic to the target group.
6. **Review and Create**:
   * Verify settings and create the Load Balancer.

**5. Configure Auto Scaling**

1. **Go to Auto Scaling Groups** > **Create Auto Scaling Group**.
2. **Define Launch Template**:
   * Create a Launch Template using the same configuration as your running EC2 instance.
3. **Configure Auto Scaling Group**:
   * Assign it to the Load Balancer target group.
   * Set desired, minimum, and maximum instance counts (e.g., 1, 1, and 3).
4. **Define Scaling Policies**:
   * Enable dynamic scaling (e.g., based on CPU utilization).
     + Target Value: 50% CPU utilization.
   * Optionally, enable scheduled scaling for predictable traffic.
5. **Review and Create**:
   * Confirm the settings and create the Auto Scaling group.



**6. Test the Setup**

1. **Load Balancer**:
   * Access the Load Balancer DNS name (e.g., http://<Load-Balancer-DNS>).
   * Verify the web application is served.
2. **Auto Scaling**:
   * Simulate load (e.g., using a load testing tool like Apache Bench).
   * Monitor the creation or termination of instances in the Auto Scaling Group.

