AWS Auto Scaling with Application Load Balancer (ALB)

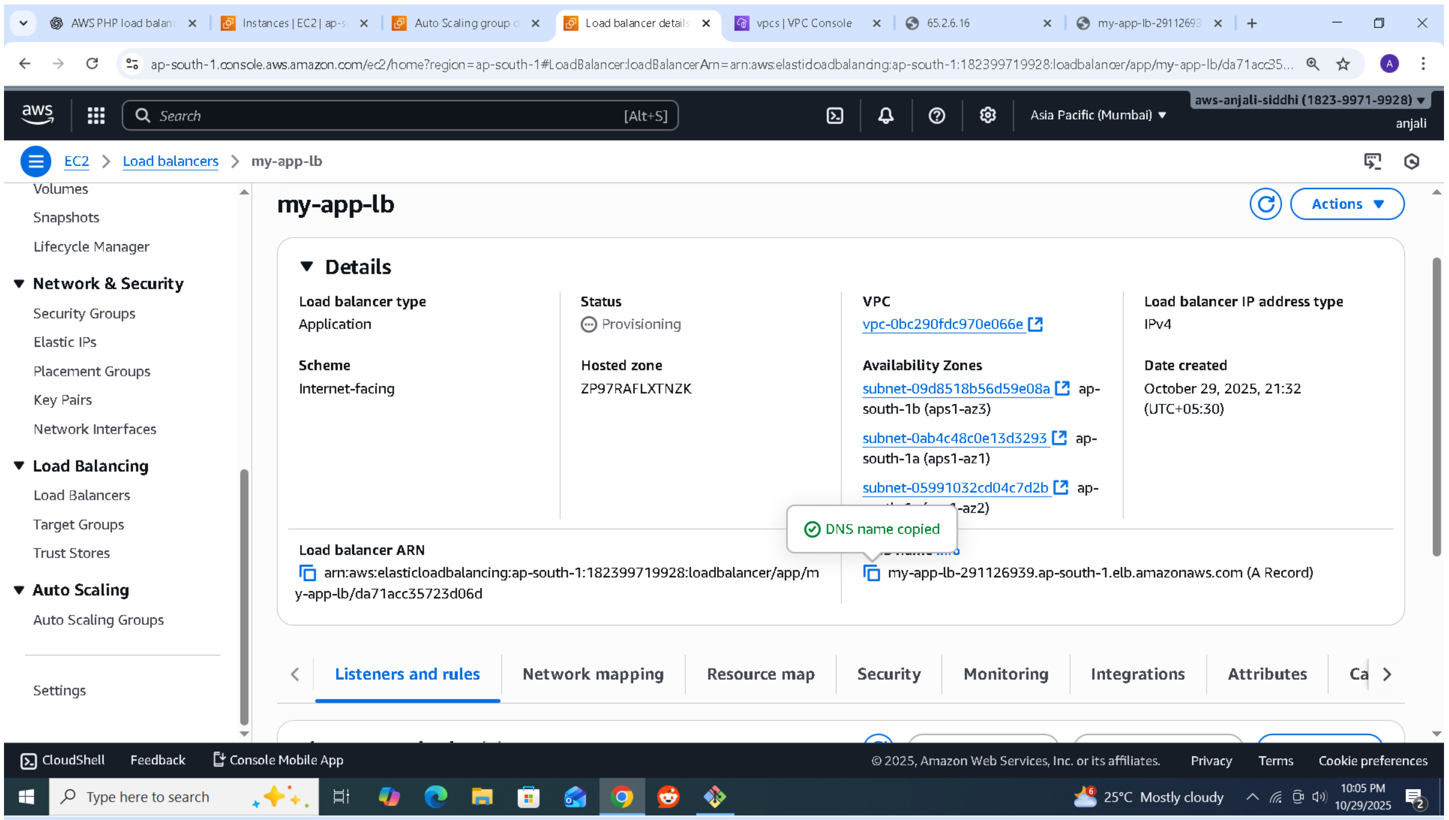
Step 1: Host a Simple PHP Application in EC2

1. Launch an EC2 instance (Amazon Linux 2 preferred).
2. Connect via SSH:
```
  ssh -i test.pem ec2-user@65.2.6.16
```
3. Install Apache and PHP:
```
  sudo yum install -y httpd php
```
4. Start and enable Apache:
```
  sudo systemctl start httpd
  sudo systemctl enable httpd
```
5. Create a PHP test page:
```
  echo "<?php phpinfo(); ?>" | sudo tee /var/www/html/index.php
```
6. Open HTTP (port 80) in the instance's Security Group.
7. Test by visiting http://65.2.6.16/ in a browser.

Step 2: Configure an Application Load Balancer (ALB)

1. Go to EC2 → Load Balancers → Create Load Balancer → Application Load Balancer.
2. Select scheme: Internet-facing, and IP type: IPv4.
3. Choose at least two subnets from different Availability Zones.
4. Select a Security Group allowing inbound HTTP (port 80).
5. Create a Target Group:
   - Target type: Instance
   - Protocol: HTTP
   - Port: 80
   - Health check path: /
6. Register your running EC2 instance with the target group.
7. Finish ALB setup and note the DNS name (e.g., my-alb-12345.ap-south-

1.elb.amazonaws.com).



AWS PHP load balan | Instances | EC2 | ap-s | Auto Scaling group | Load balancer details | vpcs | VPC Console | 65.2.6.16 | my-app-lb-29112693

ap-south-1.console.aws.amazon.com/ec2/home?region=ap-south-1#LoadBalancer:loadBalancerArn=arn:aws:elasticloadbalancing:ap-south-1:182399719928:loadbalancer/app/my-app-lb/da71acc35...

aws-anjali-siddhi (1823-9971-9928) — anjali — Asia Pacific (Mumbai)

EC2 > Load balancers > my-app-lb

Volumes
Snapshots
Lifecycle Manager
Network & Security
Security Groups
Elastic IPs
Placement Groups
Key Pairs
Network Interfaces
Load Balancing
Load Balancers
Target Groups
Trust Stores
Auto Scaling
Auto Scaling Groups
Settings

my-app-lb

Details

Load balancer type Application	**Status** Provisioning	**VPC** vpc-0bc290fdc970e066e	**Load balancer IP address type** IPv4
Scheme Internet-facing	**Hosted zone** ZP97RAFLXTNZK	**Availability Zones** subnet-09d8518b56d59e08a ap-south-1b (aps1-az3); subnet-0ab4c48c0e13d3293 ap-south-1a (aps1-az1); subnet-05991032cd04c7d2b ap-…-az2)	**Date created** October 29, 2025, 21:32 (UTC+05:30)

DNS name copied

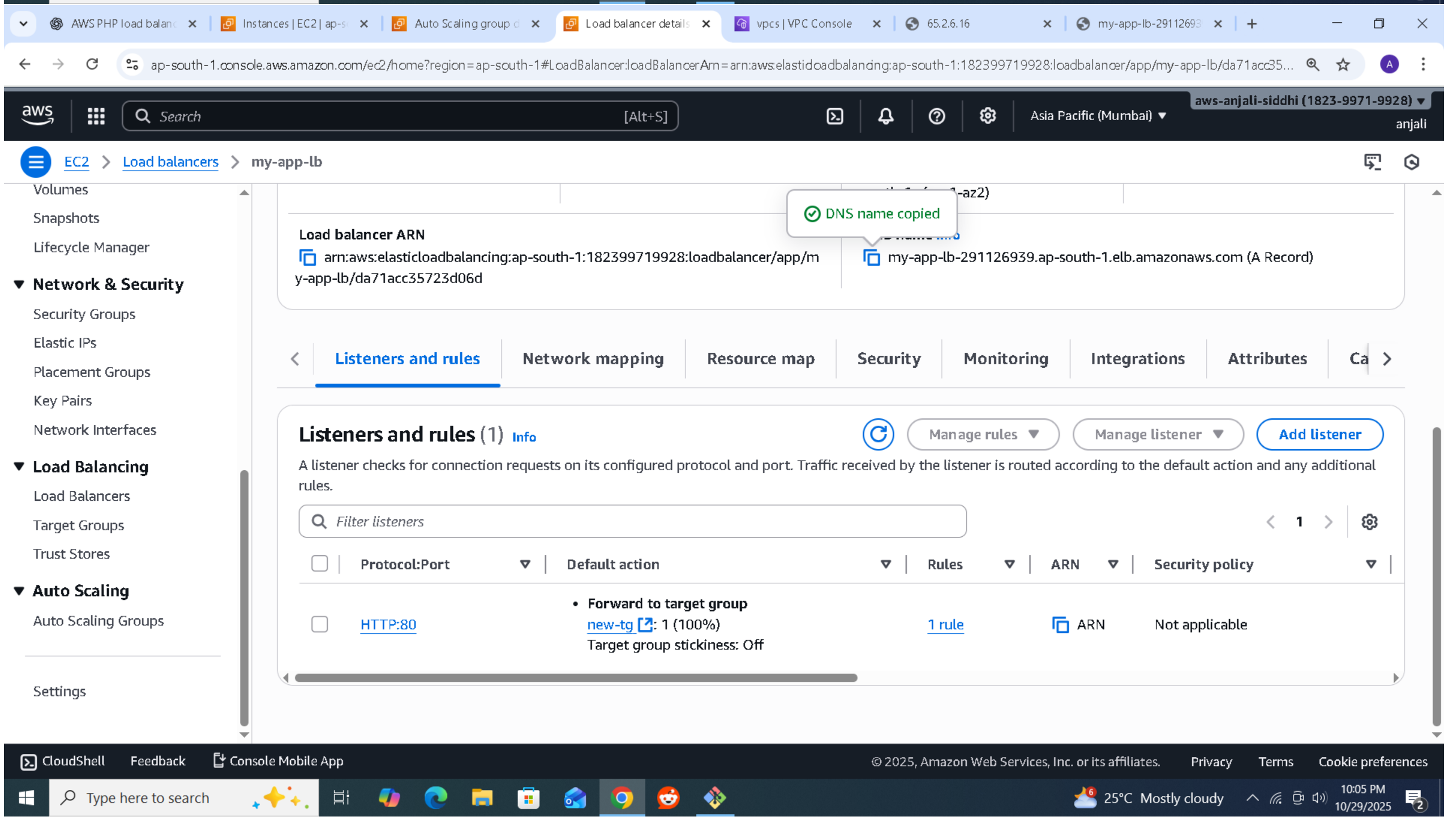
Load balancer ARN
arn:aws:elasticloadbalancing:ap-south-1:182399719928:loadbalancer/app/my-app-lb/da71acc35723d06d

my-app-lb-291126939.ap-south-1.elb.amazonaws.com (A Record)

Listeners and rules | Network mapping | Resource map | Security | Monitoring | Integrations | Attributes | Ca

© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

25°C Mostly cloudy 10:05 PM 10/29/2025



EC2 > Load balancers > my-app-lb

DNS name copied

Load balancer ARN
arn:aws:elasticloadbalancing:ap-south-1:182399719928:loadbalancer/app/my-app-lb/da71acc35723d06d

my-app-lb-291126939.ap-south-1.elb.amazonaws.com (A Record)

Listeners and rules | Network mapping | Resource map | Security | Monitoring | Integrations | Attributes | Ca

Listeners and rules (1) Info

Manage rules | Manage listener | Add listener

A listener checks for connection requests on its configured protocol and port. Traffic received by the listener is routed according to the default action and any additional rules.

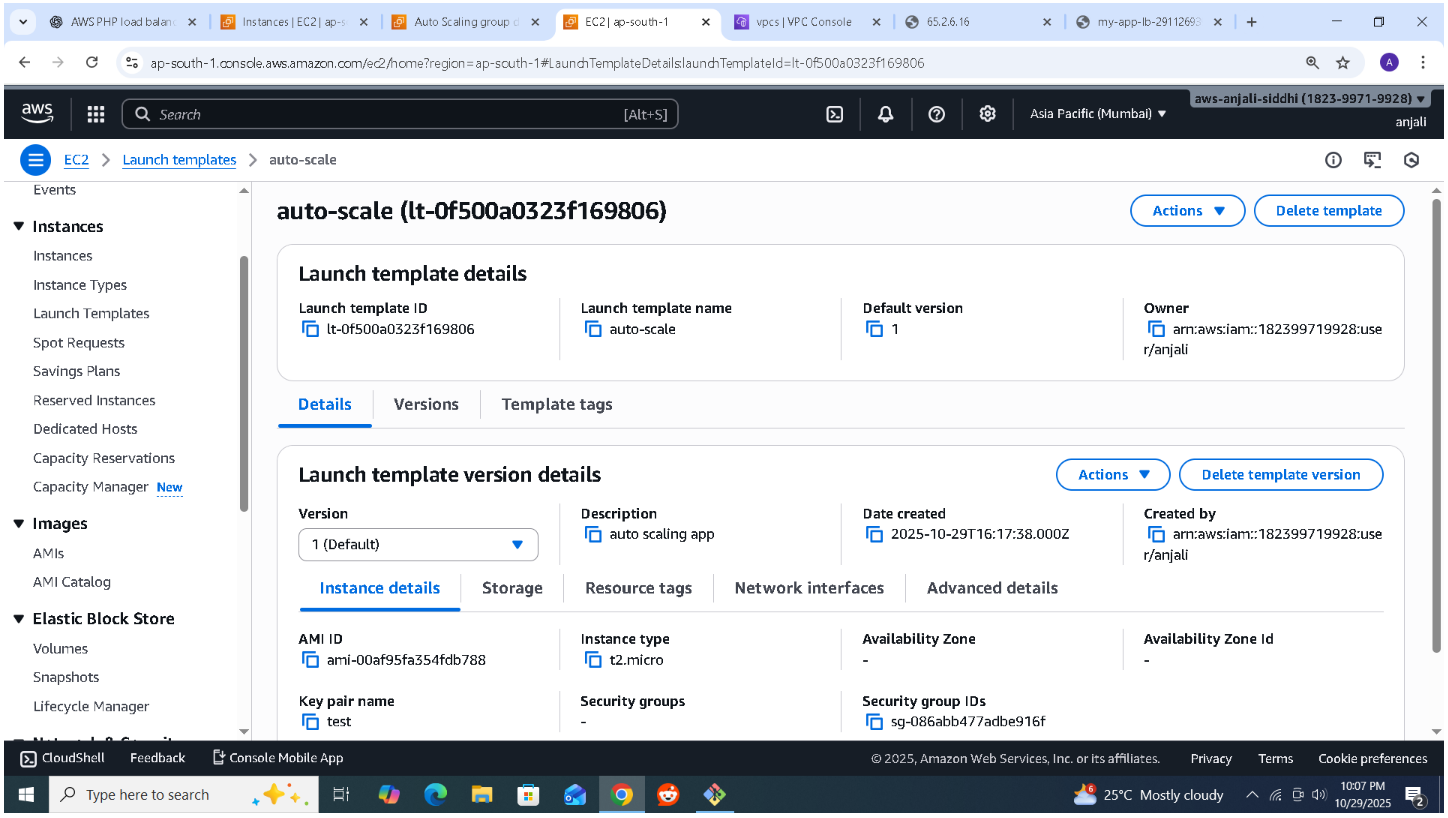
Protocol:Port	Default action	Rules	ARN	Security policy
HTTP:80	• **Forward to target group** new-tg: 1 (100%) Target group stickiness: Off	1 rule	ARN	Not applicable

© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

25°C Mostly cloudy 10:05 PM 10/29/2025

Step 3: Create Launch Template for Auto Scaling

1. Go to EC2 → Launch Templates → Create Launch Template.
2. Configure:
   - Name: auto-scale
   - AMI: same as your instance
   - Instance type: t2.micro
   - Key pair: test
   - Security Group: same as your instance (allow port 80)
3. Leave subnet blank (ASG will handle multi-AZ placement).
4. Save the template.



auto-scale (lt-0f500a0323f169806)

Launch template details

Launch template ID	Launch template name	Default version	Owner
lt-0f500a0323f169806	auto-scale	1	arn:aws:iam::182399719928:user/anjali

Launch template version details

Version	Description	Date created	Created by
1 (Default)	auto scaling app	2025-10-29T16:17:38.000Z	arn:aws:iam::182399719928:user/anjali

Instance details

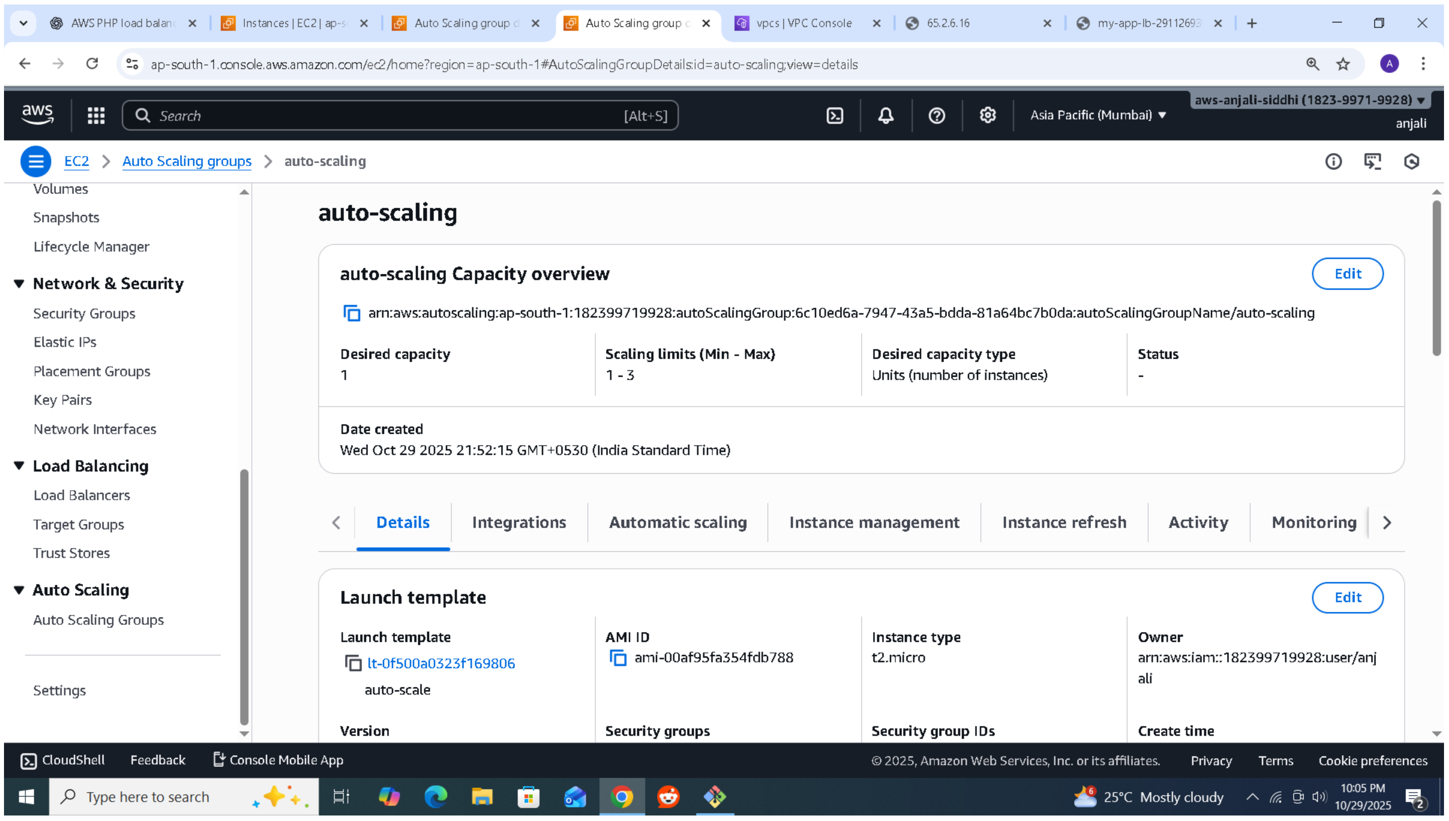
AMI ID	Instance type	Availability Zone	Availability Zone Id
ami-00af95fa354fdb788	t2.micro	-	-

Key pair name	Security groups	Security group IDs
test	-	sg-086abb477adbe916f

Step 4: Create an Auto Scaling Group (ASG)

1. Go to EC2 → Auto Scaling Groups → Create Auto Scaling Group.
2. Select the Launch Template created earlier.
3. Set:
   - Desired capacity: 1
   - Minimum capacity: 1

  - Maximum capacity: 3
4. Choose the same VPC and subnets as the ALB.
5. Attach the previously created Target Group.
6. Enable EC2 and ELB health checks (recommended).
7. Set Health check grace period: 300 seconds.
8. Skip or disable VPC Lattice integration unless required.
9. Finish creation.



auto-scaling

auto-scaling Capacity overview

arn:aws:autoscaling:ap-south-1:182399719928:autoScalingGroup:6c10ed6a-7947-43a5-bdda-81a64bc7b0da:autoScalingGroupName/auto-scaling

Desired capacity	Scaling limits (Min - Max)	Desired capacity type	Status
1	1 - 3	Units (number of instances)	-

Date created
Wed Oct 29 2025 21:52:15 GMT+0530 (India Standard Time)

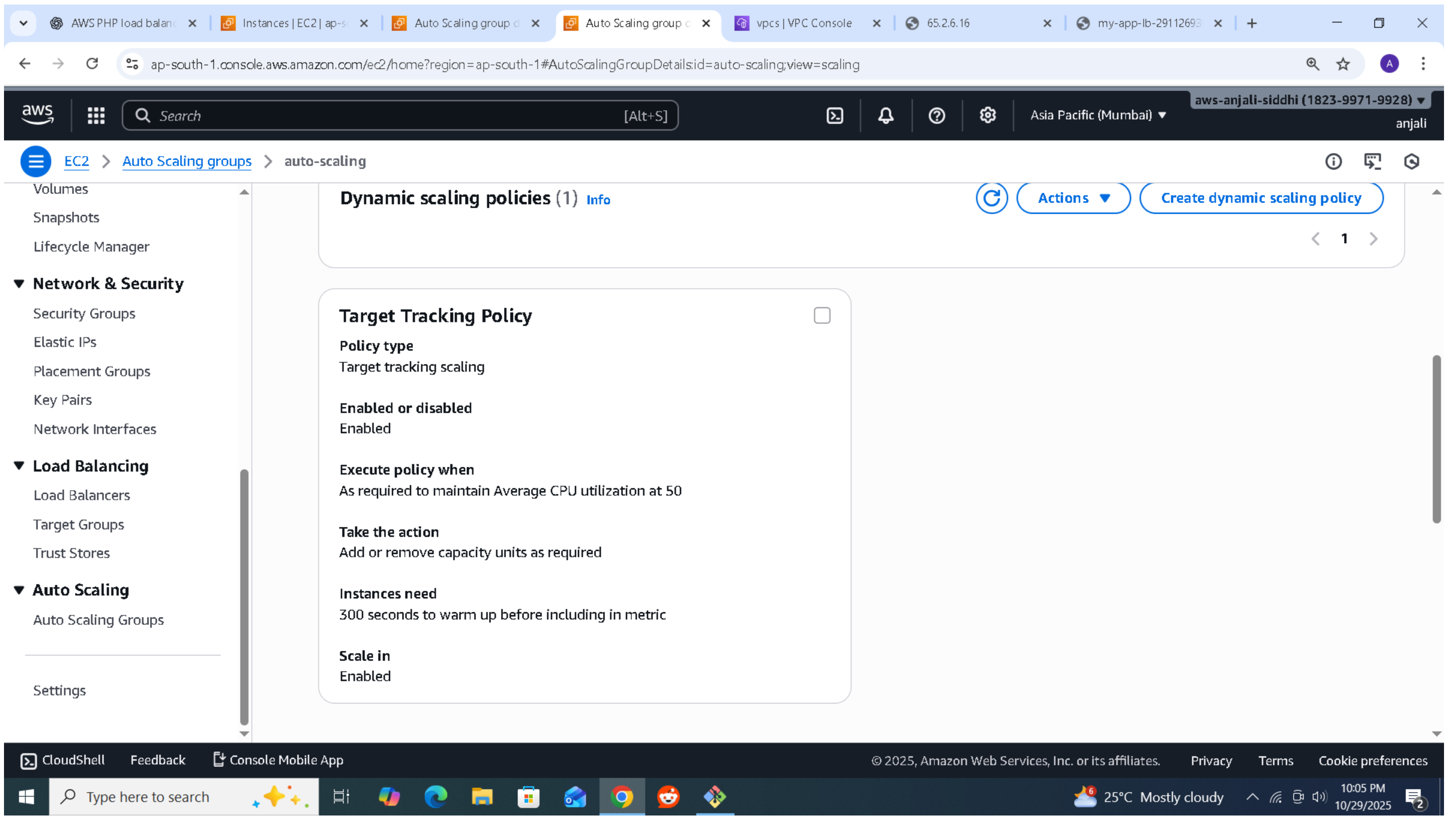
Launch template

Launch template	AMI ID	Instance type	Owner
lt-0f500a0323f169806 auto-scale	ami-00af95fa354fdb788	t2.micro	arn:aws:iam::182399719928:user/anjali

Version | Security groups | Security group IDs | Create time

Step 5: Configure Scaling Policy (CPU-based)

1. Go to your Auto Scaling Group → Automatic Scaling tab.
2. Click “Add policy” → Choose “Target tracking scaling policy.”
3. Configure as:
  - Policy name: CPU50-TargetTracking
  - Metric type: Average CPU utilization
  - Target value: 50%
  - Cooldown period: 300 seconds (default)
4. Save the policy.


Dynamic scaling policies (1) Info

Target Tracking Policy

Policy type
Target tracking scaling

Enabled or disabled
Enabled

Execute policy when
As required to maintain Average CPU utilization at 50

Take the action
Add or remove capacity units as required

Instances need
300 seconds to warm up before including in metric

Scale in
Enabled

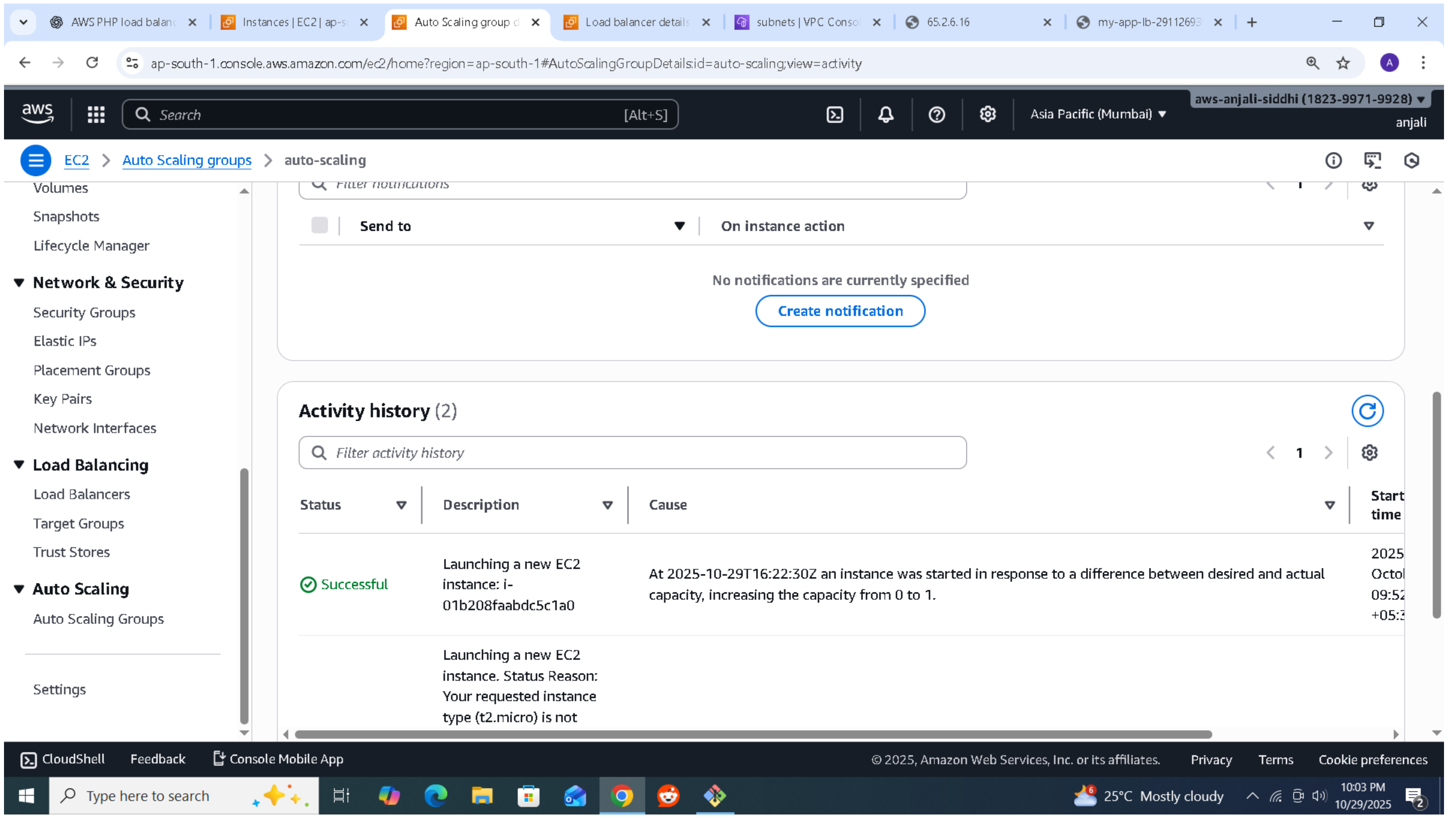
Step 6: Test Scaling Behavior

1. SSH into your instance:
   ssh -i test.pem ec2-user@<instance-public-ip>
2. Install the stress tool:
   sudo yum install -y stress
3. Run the load test:
   stress --cpu 2 --timeout 300
4. Monitor scaling:
   - In EC2 → Auto Scaling Group → Activity tab, watch for new instance launches.
   - In EC2 → Instances, new instances will appear.
   - In Load Balancer → Target Groups, all instances will show as healthy.
5. Once CPU usage drops below 50%, the ASG will terminate extra instances.

Step 7: Verification and Monitoring

1. Verify in CloudWatch:
   - Go to CloudWatch → Metrics → EC2 → CPUUtilization.
   - Observe the spike during the stress test and the automatic scaling response.
2. Visit your ALB DNS URL and ensure the load is distributed across instances.

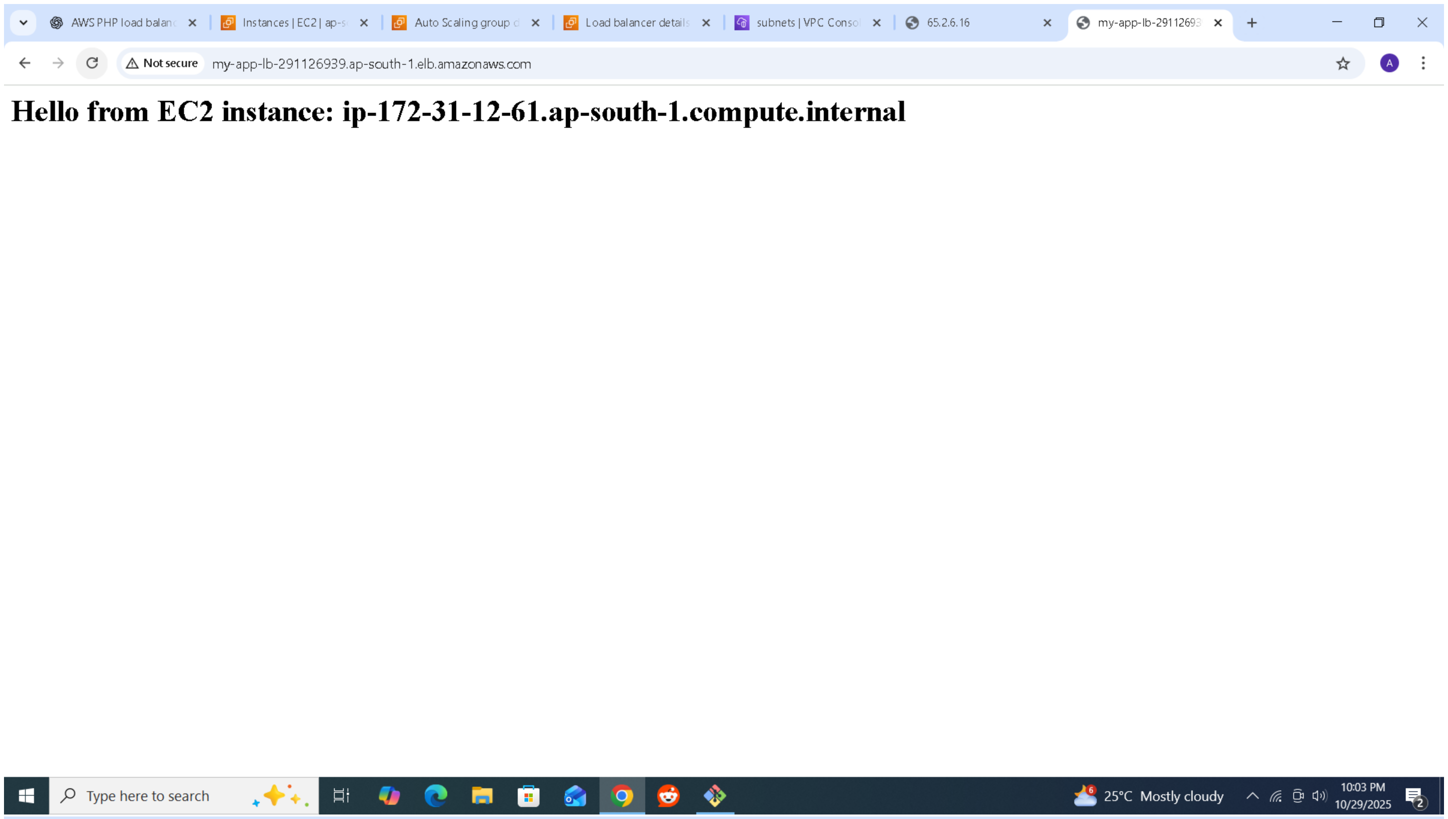
✓Successfully configured AWS Auto Scaling with an Application Load Balancer for a PHP web application.



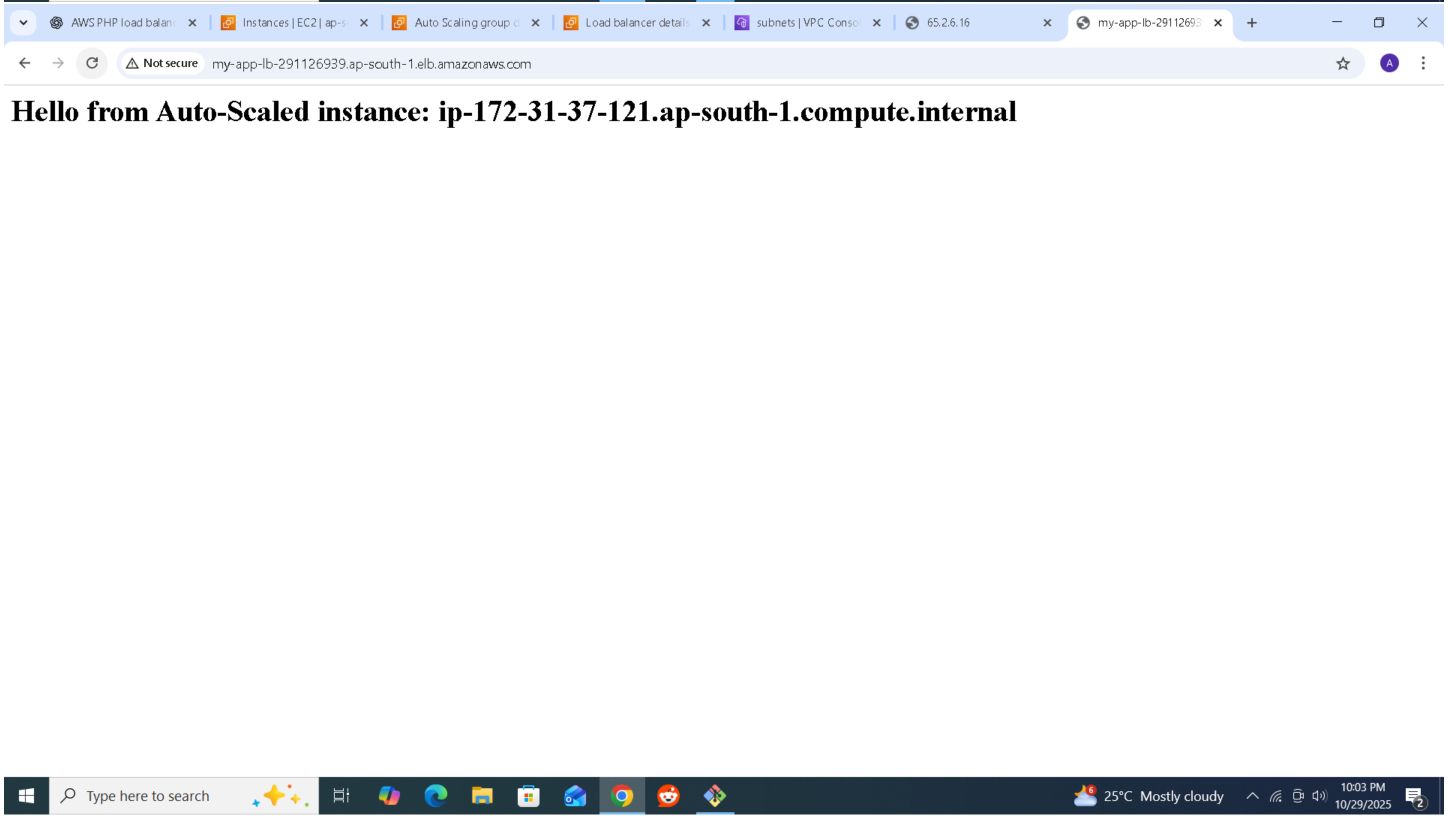
No notifications are currently specified

Activity history (2)

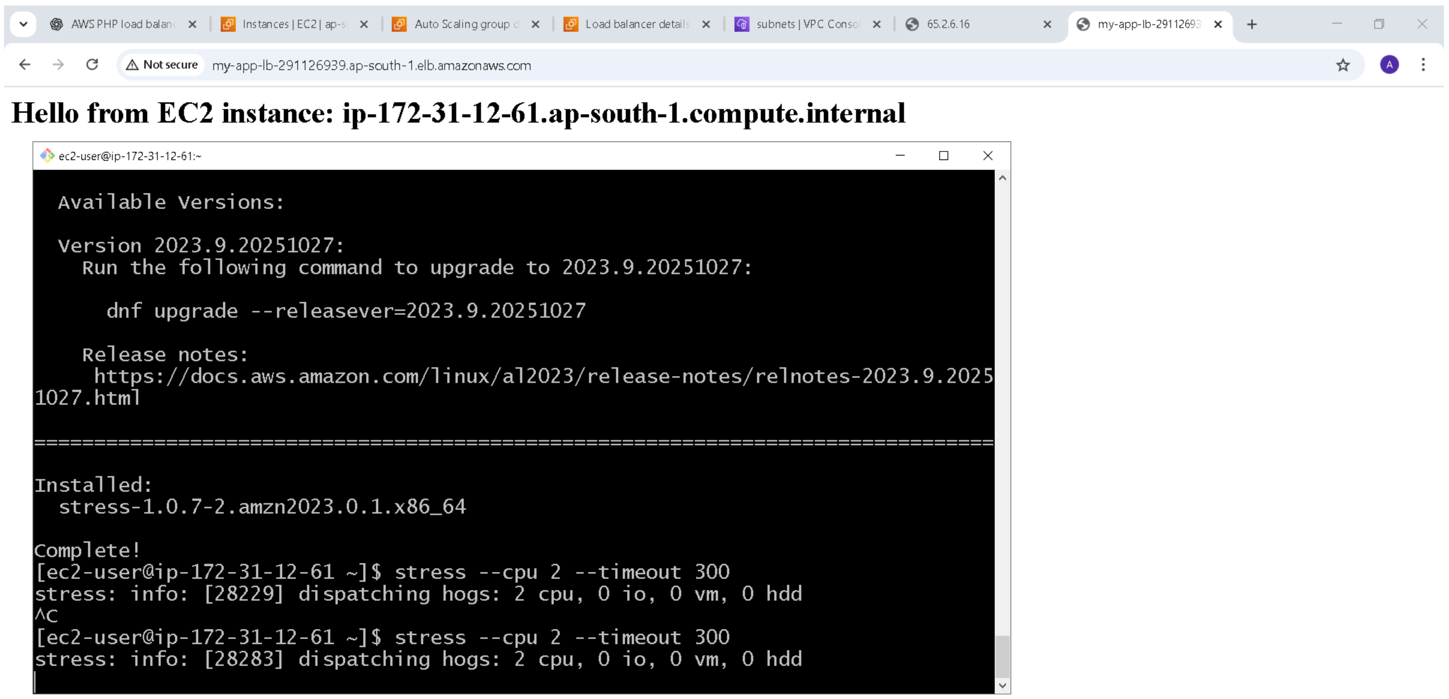
Status	Description	Cause	Start time
Successful	Launching a new EC2 instance: i-01b208faabdc5c1a0	At 2025-10-29T16:22:30Z an instance was started in response to a difference between desired and actual capacity, increasing the capacity from 0 to 1.	2025 Octol 09:52 +05:3
	Launching a new EC2 instance. Status Reason: Your requested instance type (t2.micro) is not		



Hello from EC2 instance: ip-172-31-12-61.ap-south-1.compute.internal



Hello from Auto-Scaled instance: ip-172-31-37-121.ap-south-1.compute.internal



AWS PHP load balan | Instances | EC2 | ap-s | Auto Scaling group d | Load balancer details | subnets | VPC Conso | 65.2.6.16 | my-app-lb-29112693

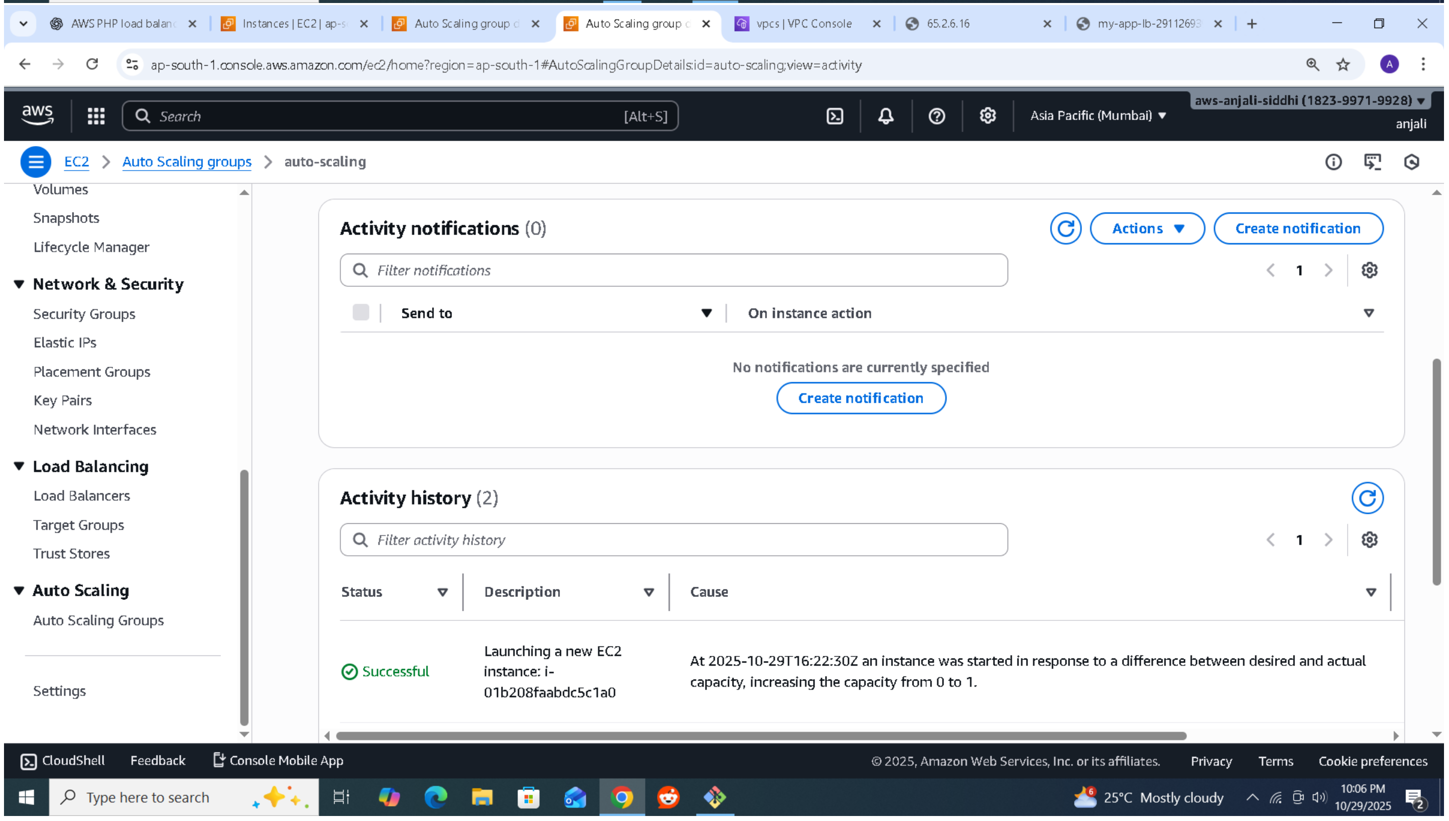
Not secure my-app-lb-291126939.ap-south-1.elb.amazonaws.com

Hello from EC2 instance: ip-172-31-12-61.ap-south-1.compute.internal

ec2-user@ip-172-31-12-61:~

```
  Available Versions:

  Version 2023.9.20251027:
    Run the following command to upgrade to 2023.9.20251027:

      dnf upgrade --releasever=2023.9.20251027

    Release notes:
     https://docs.aws.amazon.com/linux/al2023/release-notes/relnotes-2023.9.2025
1027.html

================================================================================

Installed:
  stress-1.0.7-2.amzn2023.0.1.x86_64

Complete!
[ec2-user@ip-172-31-12-61 ~]$ stress --cpu 2 --timeout 300
stress: info: [28229] dispatching hogs: 2 cpu, 0 io, 0 vm, 0 hdd
^C
[ec2-user@ip-172-31-12-61 ~]$ stress --cpu 2 --timeout 300
stress: info: [28283] dispatching hogs: 2 cpu, 0 io, 0 vm, 0 hdd
```

Type here to search | 25°C Mostly cloudy | 10:04 PM 10/29/2025



AWS PHP load balan | Instances | EC2 | ap-s | Auto Scaling group d | Auto Scaling group d | vpcs | VPC Console | 65.2.6.16 | my-app-lb-29112693

ap-south-1.console.aws.amazon.com/ec2/home?region=ap-south-1#AutoScalingGroupDetails:id=auto-scaling;view=activity

Search [Alt+S] | Asia Pacific (Mumbai) | aws-anjali-siddhi (1823-9971-9928) | anjali

EC2 > Auto Scaling groups > auto-scaling

Volumes
Snapshots
Lifecycle Manager
Network & Security
Security Groups
Elastic IPs
Placement Groups
Key Pairs
Network Interfaces
Load Balancing
Load Balancers
Target Groups
Trust Stores
Auto Scaling
Auto Scaling Groups
Settings

Activity notifications (0) | Actions | Create notification

Filter notifications

Send to	On instance action

No notifications are currently specified

Create notification

Activity history (2)

Filter activity history

Status	Description	Cause
Successful	Launching a new EC2 instance: i-01b208faabdc5c1a0	At 2025-10-29T16:22:30Z an instance was started in response to a difference between desired and actual capacity, increasing the capacity from 0 to 1.

CloudShell | Feedback | Console Mobile App | © 2025, Amazon Web Services, Inc. or its affiliates. | Privacy | Terms | Cookie preferences

Type here to search | 25°C Mostly cloudy | 10:06 PM 10/29/2025