

Practical Lab on EC2 Auto Scaling using AWS

Step 1: Create an Amazon Machine Image (AMI)

1. Log in to your AWS Management Console.
2. Navigate to the EC2 Dashboard.
3. Launch an EC2 instance to serve as your base image. Install all necessary software on this instance, such as a web server (e.g., Apache or Nginx) and your web application code.
4. Once the instance is configured, select it from the EC2 Instances list.
5. Go to Actions -> Image and templates -> Create image.
6. Give the image a name (e.g., ALB-GA-AMI) and a description.
7. Click Create image. This will create a snapshot of your instance's root volume and create a new AMI.

Amazon Machine Images (AMIs) (1) [Info](#)

Owned by me ▾

[Recycle Bin](#)

[EC2 Image Builder](#)

[Actions ▾](#)

[Launch instance from AMI](#)

<input type="checkbox"/>	Name ↗	AMI name	AMI ID	Source	Owner	Visibility	Status	Creation date	Platform	Root device
<input type="checkbox"/>		ALB-GAMI	ami-0dea3c1816fb3788b	398211280527/ALB-GAMI	398211280527	Private	Available 🔍 🔍	2025/09/12 16:31 GMT+5:30	Linux/UNIX	ebs

Step 2: Create Security Groups

1. From the EC2 Dashboard, navigate to Security Groups.
2. Click Create security group to create the following three security groups:
 - **ALB Security Group:**
 - Give it a name (e.g., elb-SG).
 - Add an **Inbound rule** for HTTP traffic:
 - **Type:** HTTP
 - **Protocol:** TCP
 - **Port range:** 80
 - **Source:** Anywhere (0.0.0.0/0)
 - This security group allows all incoming HTTP traffic to the load balancer.
 - **Web Server Security Group:**
 - Give it a name (e.g., web-server-SG).
 - Add an **Inbound rule** for HTTP traffic:
 - **Type:** HTTP
 - **Protocol:** TCP
 - **Port range:** 80
 - **Source:** Custom, and enter the security group ID for your elb-SG to allow traffic only from the ALB.
 - Add an inbound rule for SSH access for management.
 - This security group allows incoming HTTP traffic only from the load balancer.

- **EFS Security Group:**
 - Give it a name (e.g., `EFS-SG`).
 - Add an **Inbound rule** for EFS traffic:
 - **Type:** NFS
 - **Protocol:** TCP
 - **Port range:** 2049
 - **Source:** Custom, and enter the security group ID for your `web-server-SG` to allow traffic only from the web servers.
 - This security group allows the EC2 instances to communicate with the EFS file system.

Security Groups (4) [Info](#)

Find security groups by attribute or tag

Actions Export security groups to CSV Create security group

<input type="checkbox"/>	Name	Security group ID	Security group name	VPC ID	Description	Owner	Inbound rules count
<input type="checkbox"/>	-	sg-0fedf9364a8f51342	web-server-SG	vpc-0f14e8837e8e713bc	web-server-SG for hosting web pages	398211280527	3 Permission entries
<input type="checkbox"/>	-	sg-0f682968c5328f1f7	EFS-SG	vpc-0f14e8837e8e713bc	EFS-SG for shared storage to all connec...	398211280527	1 Permission entry
<input type="checkbox"/>	-	sg-01b487e42432f0bba	elb-SG	vpc-0f14e8837e8e713bc	alb-SG for web servers	398211280527	2 Permission entries
<input type="checkbox"/>	-	sg-04b35d82c1f5f6bd0	default	vpc-0f14e8837e8e713bc	default VPC security group	398211280527	1 Permission entry

Step 3: Create an Application Load Balancer (ALB)

1. From the EC2 Dashboard, navigate to Load Balancers.
2. Click Create Load Balancer.
3. Choose Application Load Balancer.
4. Configure the load balancer:
 - **Name:** `asg-alb`
 - **Scheme:** Internet-facing
 - **VPC:** Select your default VPC.
 - **Availability Zones:** Select at least two Availability Zones (e.g., `ap-south-1b` and `ap-south-1c`).
 - **Security Groups:** Select the `elb-SG` security group

Step 4: Create a Target Group

1. From the EC2 Dashboard, navigate to Target Groups.
2. Click Create target group.
3. Choose Instances.
4. **Target group name:** asg-tg
5. **Protocol:** HTTP, **Port:** 80.
6. **VPC:** Select your default VPC.

asg-TG Actions ▾

Details

arn:aws:elasticloadbalancing:ap-south-1:398211280527:targetgroup/asg-TG/5a8f7323a0ed0d42

Target type Instance	Protocol : Port HTTP: 80	Protocol version HTTP1	VPC vpc-0f14e8837e8e713bc
IP address type IPv4	Load balancer asg-alb		

2 Total targets

2 Healthy

0 Anomalous

0 Unhealthy

0 Unused

0 Initial

0 Draining

► **Distribution of targets by Availability Zone (AZ)**
Select values in this table to see corresponding filters applied to the Registered targets table below.

Targets | Monitoring | Health checks | Attributes | Tags

Registered targets (2) [info](#)

Anomaly mitigation: **Not applicable** ⌂ Deregister Register targets

Target groups route requests to individual registered targets using the protocol and port number specified. Health checks are performed on all registered targets according to the target group's health check settings. Anomaly detection is automatically applied to HTTP/HTTPS target groups with at least 3 healthy targets.

Filter targets

<input type="checkbox"/>	Instance ID	Name	Port	Zone	Health status	Health status details	Administrative override	Override details
<input type="checkbox"/>	i-04809380fcc90afac	ASG-instance	80	ap-south-1b (a...	Healthy	-	No override	No override is currently active on target
<input type="checkbox"/>	i-031904d6442f7001c	ASG-instance	80	ap-south-1c (a...	Healthy	-	No override	No override is currently active on target

Step 5: Configure the ALB Listener

1. Go back to your asg-alb in the Load Balancers section.
2. Click on the Listeners tab.
3. Click Add listener.
4. **Protocol:** HTTP, **Port:** 80.
5. **Default action:** Forward to, and select the asg-tg target group you just created.
6. Click Add.

asg-alb

▼ Details

Load balancer type

Application

Scheme

Internet-facing

Status

Active

Hosted zone

ZP97RAFLXTNZK

VPC

vpc-0f14e8837e8e713bc

Availability Zones

subnet-04396ce67770e72ed ap-south-1c (aps1-az2)

subnet-0ec8a554198204dd2 ap-south-1b (aps1-az3)

subnet-0c6c345f38458c250 ap-south-1a (aps1-az1)

Load balancer IP address type

IPv4

Date created

September 12, 2025, 16:49 (UTC+05:30)

Load balancer ARN

arn:aws:elasticloadbalancing:ap-south-1:398211280527:loadbalancer/app/asg-alb/f7173ba91e1ca594

DNS name info

asg-alb-522802902.ap-south-1.elb.amazonaws.com (A Record)

Listeners and rules

Network mapping

Resource map

Security

Monitoring

Integrations

Attributes

Capacity

Tags

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.

Filter listeners

< 1 >

Protocol:Port	Default action	Rules	ARN	Security policy	Default SSL/TLS certificate	mTLS	Trust store
<input type="checkbox"/> HTTP:80	<div> <div>Forward to target group</div> <div> <div>asg-TG 1 (100%)</div> <div>Target group stickiness: Off</div> </div> </div>	1 rule	<input type="checkbox"/> ARN	Not applicable	Not applicable	Not applicable	Not applica

Step 6: Create an Auto Scaling Group (ASG)

- From the EC2 Dashboard, navigate to Auto Scaling Groups.
- Click Create Auto Scaling group.
- Name:** Give it a name, such as my-asg.
- Launch Template:** Create a new launch template or select an existing one. In the template, specify:
 - The AMI you created in Step 1 (ALB-GA-AMI).
 - The instance type (c5a.large).
 - The security group you created for the web server (web-server-SG).
- Network:**
 - VPC:** Select your default VPC.
 - Subnets:** Select the same subnets that you used for your ALB.
- Load balancing:**
 - Attach to an existing load balancer:** Choose your asg-tg target group.
- Group size:** Set your desired, minimum, and maximum capacity (e.g., **Desired:** 4, **Minimum:** 1, **Maximum:** 6).
- Scaling policies:** Configure scaling policies based on metrics like CPU utilization or network traffic.
- Review and click Create Auto Scaling group.

Linkedin : <https://www.linkedin.com/in/uppala-venkata-sai/>

Github : <https://github.com/UppalavenkataSai>

Edit scheduled action

×

Name

ontime

ⓘ Provide at least one value for Desired, Min, or Max Capacity

Desired capacity

Min

Max

4

1

6

Recurrence

Once

Time zone

Asia/Kolkata

Current time in selected time zone is 2025-09-12/17:50 IST

Specific start time

Schedule a specific date and time for the first scheduled action to run. Interpreted in recurrence time zone: Asia/Kolkata

2025/09/12

17:53

Asia/Kolkata

[Learn more about scheduled scaling](#)

Cancel

Save changes

Step 7: Configure Scheduled Scaling

1. Select your newly created Auto Scaling Group.
2. Go to the Automatic scaling tab.
3. Click Create scheduled action.
4. **Name:** ontime
5. **Desired capacity:** 4, **Min capacity:** 1, **Max capacity:** 6.
6. **Recurrence:** Once.
7. **Start Time:** Set the specific date and time for the scheduled action to occur (e.g., 2025/09/12 at 17:53).
8. Click Create.

Instances (9) Info											
Find instance by attribute or tag (case-sensitive)				All states		Last updated less than a minute ago					
	Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS	Public IPv4 ...	Elastic IP	IPv6 IPs
<input type="checkbox"/>		i-0789eddc058fd15c7	Running	c6a.large	3/3 checks passed	View alarms +	ap-south-1b	-	-	-	-
<input type="checkbox"/>		i-040cccf17d8e7c69d	Running	c6a.large	3/3 checks passed	View alarms +	ap-south-1b	-	-	-	-
<input type="checkbox"/>		i-00a9cbb9dce0e8bde	Running	c5a.large	Initializing	View alarms +	ap-south-1a	ec2-15-127-48-191.ap-...	13.127.48.191	-	-
<input type="checkbox"/>		i-0a870ec2b3956ee33	Running	c5a.large	Initializing	View alarms +	ap-south-1c	ec2-13-200-21-130.ap-...	13.200.21.130	-	-
<input type="checkbox"/>		i-04809380fcc90afac	Terminated	c6a.large	-	View alarms +	ap-south-1b	-	-	-	-
<input type="checkbox"/>		i-05ad6213a764cfb13	Terminated	c6a.large	-	View alarms +	ap-south-1b	-	-	-	-
<input type="checkbox"/>	base-server	i-07e389a38bcd15dbc	Terminated	t2.micro	-	View alarms +	ap-south-1a	-	-	-	-
<input type="checkbox"/>		i-032e059d02f72fe73	Terminated	c5a.large	-	View alarms +	ap-south-1a	-	-	-	-
<input type="checkbox"/>		i-031904d6442f7001c	Terminated	c5a.large	-	View alarms +	ap-south-1c	-	-	-	-

In this lab, Auto Scaling was successfully demonstrated using Amazon EC2 instances. The **EC2 console** shows a total of **9 instances**, with multiple instances being launched and terminated automatically according to the scaling policies.

- **Running Instances:**
 - 4 instances (c6a.large and c5a.large types) are currently in the *Running* state.
 - Two of them (i-0789ed... and i-040cce...) have passed the **3/3 status checks**, indicating they are fully operational.
 - The other two (i-00a9cb... and i-0a87ce...) are in the **Initializing** state, showing that Auto Scaling has just launched them as part of the scaling activity.
- **Terminated Instances:**
 - 5 instances have been terminated (c6a.large, c5a.large, and t2.micro), which confirms that Auto Scaling removed excess capacity when it was no longer needed.
- **Availability Zones:**
 - Instances are distributed across multiple AZs (ap-south-1a, ap-south-1b, and ap-south-1c) for high availability and fault tolerance.
- **Public IPs:**
 - Some running instances have public IPv4 addresses assigned (13.127.48.191 and 13.200.21.130), making them accessible over the internet.

This file contains the full HTML code for the webpage. You can use the guide provided earlier to place this code on your EC2 instance. This should complete the test and confirm that your infrastructure is working correctly.

