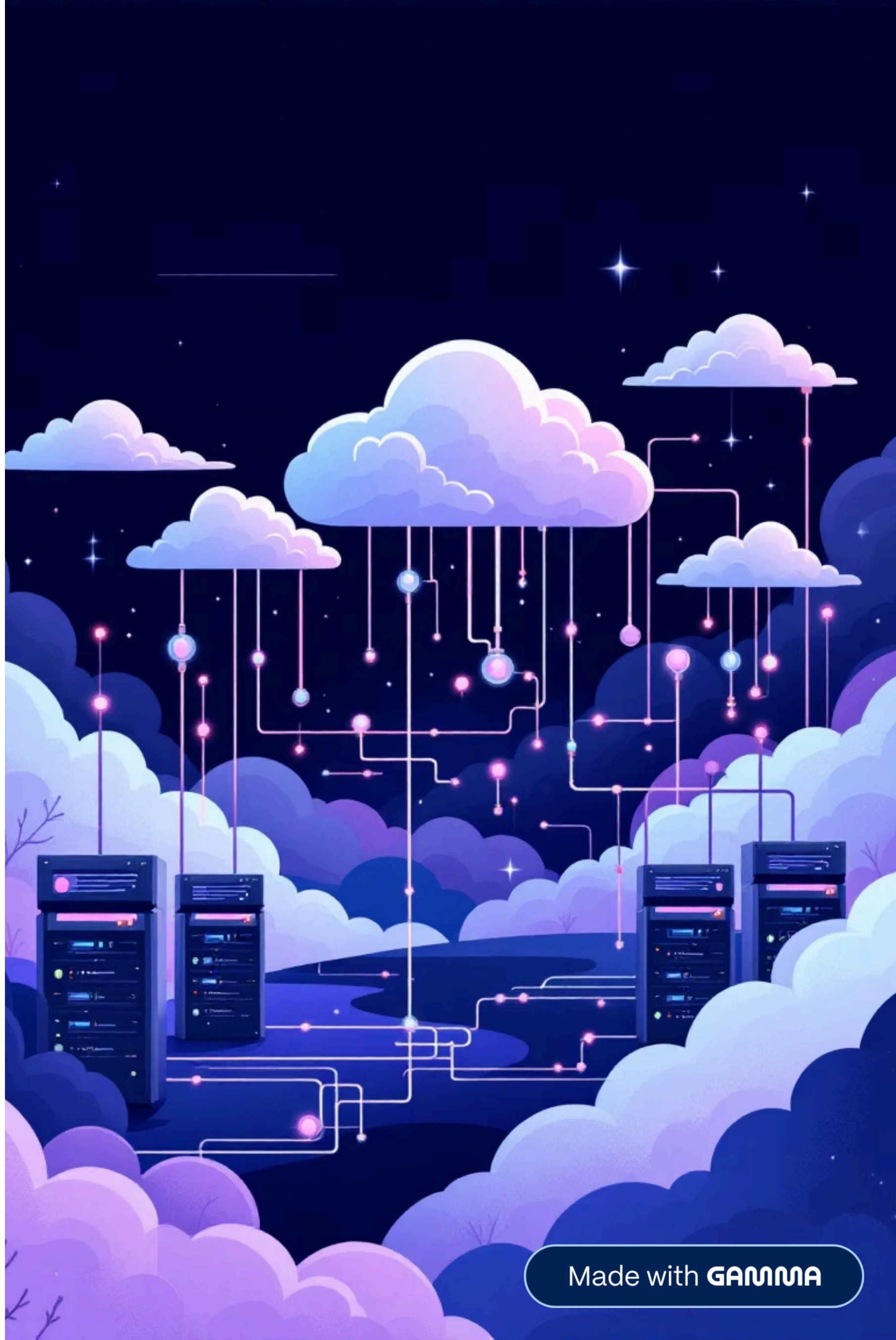
AWS Application Load Balancer with EC2 - High Availability Architecture

A production-ready architecture demonstrating traffic distribution and fault tolerance using AWS services



Project Overview

What This Project Demonstrates

This project showcases how an AWS Application Load Balancer (ALB) intelligently distributes incoming web traffic across multiple EC2 instances to achieve high availability and fault tolerance.

Rather than depending on a single server, the application operates across multiple EC2 instances, with the load balancer routing user requests exclusively to healthy instances.

This architecture represents a real-world production setup commonly deployed in enterprise cloud-based applications.

Why Application Load Balancer?

The Problem

Single EC2 deployments create critical vulnerabilities:

- Single point of failure
- No fault tolerance
- Limited scalability

The Solution

Application Load Balancer addresses these challenges by:

- Distributing traffic intelligently
- Handling instance failures automatically
- Ensuring continuous availability

The ALB functions as a unified entry point whilst efficiently managing backend server resources.

Architecture Advantages

High Availability

Application remains accessible even when individual EC2 instances experience failures or maintenance downtime.

Fault Tolerance

Traffic routes exclusively to healthy instances based on continuous health check monitoring.

Production-Ready Design

Architecture matches real industry standards and easily integrates with Auto Scaling for future growth.

AWS Services Used

1

Amazon EC2

Virtual compute service hosting the web application instances across multiple availability zones.

2

Application Load Balancer

Distributes incoming HTTP traffic across EC2 instances based on health checks and routing rules.

3

Target Groups

Logical grouping of EC2 instances that receive and handle traffic forwarded from the load balancer.

4

Virtual Private Cloud

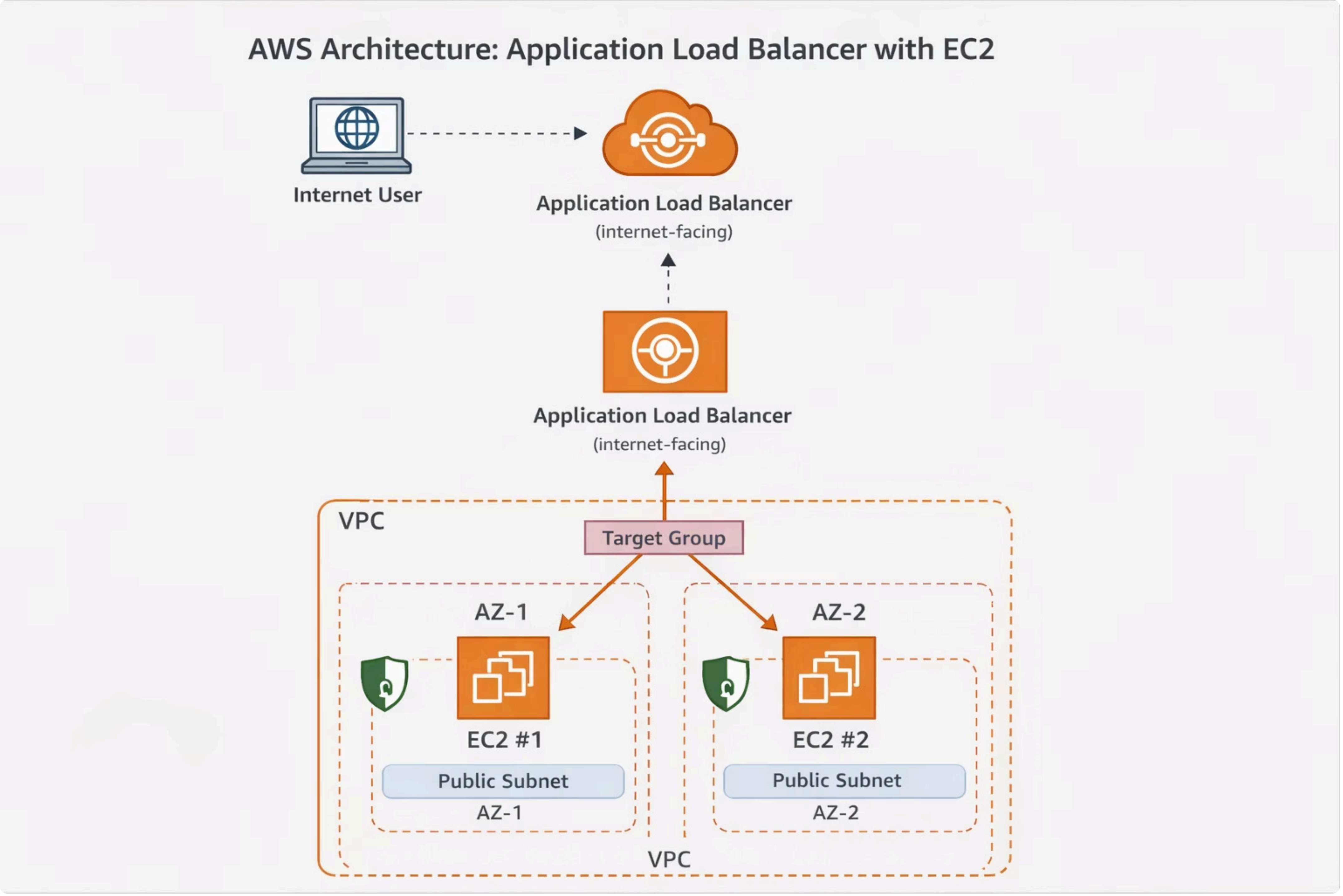
Provides isolated networking environment with subnets, route tables, and internet gateway configuration.

5

Security Groups

Acts as virtual firewall controlling inbound and outbound traffic to EC2 instances and ALB.

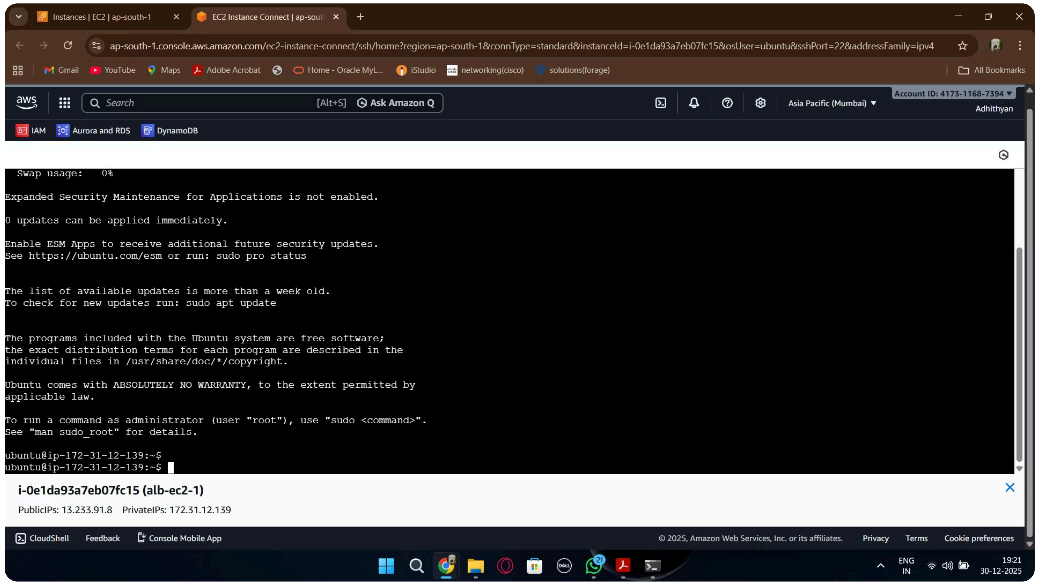
Architecture Diagram



This architecture shows how an Application Load Balancer distributes incoming traffic across multiple EC2 instances deployed in different Availability Zones. The design ensures high availability, fault tolerance, and scalable web application hosting on AWS.

EC2 Instance Connect & Apache Installation

EC2 Instance Connect (Terminal Access)



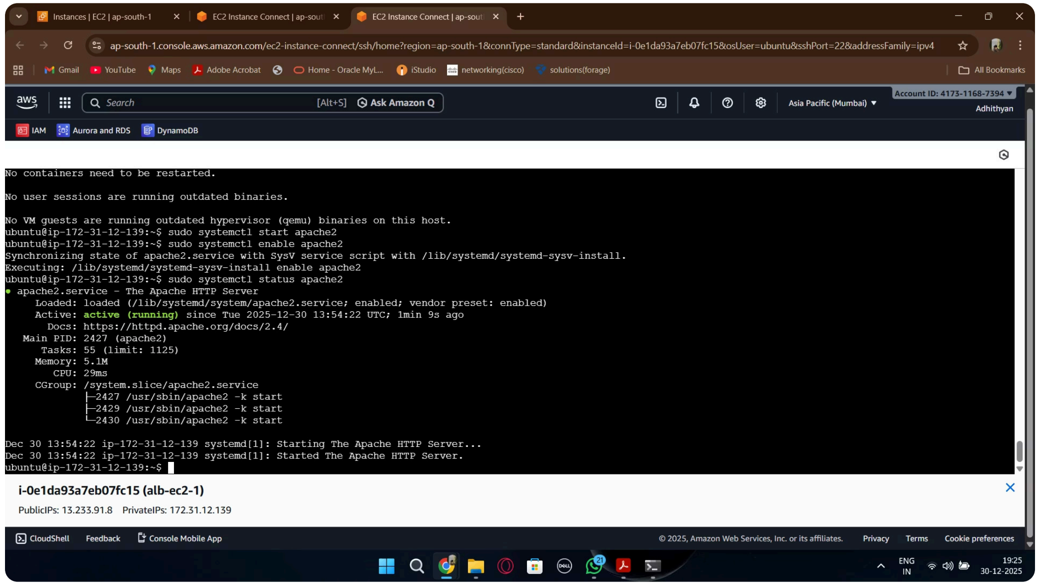
```
 Swap usage:   0%

Expanded Security Maintenance for Applications is not enabled.

0 updates can be applied immediately.

Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status


The list of available updates is more than a week old.
To check for new updates run: sudo apt update


The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

ubuntu@ip-172-31-12-139:~$
ubuntu@ip-172-31-12-139:~$
```

i-0e1da93a7eb07fc15 (alb-ec2-1)

PublicIPs: 13.233.91.8 PrivateIPs: 172.31.12.139

This screenshot shows successful connection to the EC2 instance using EC2 Instance Connect. It confirms that the instance is running and accessible for server configuration.

Apache Service Installation & Status



```
No containers need to be restarted.

No user sessions are running outdated binaries.

No VM guests are running outdated hypervisor (qemu) binaries on this host.
ubuntu@ip-172-31-12-139:~$ sudo systemctl start apache2
ubuntu@ip-172-31-12-139:~$ sudo systemctl enable apache2
Synchronizing state of apache2.service with SysV service script with /lib/systemd/systemd-sysv-install.
Executing: /lib/systemd/systemd-sysv-install enable apache2
ubuntu@ip-172-31-12-139:~$ sudo systemctl status apache2
● apache2.service - The Apache HTTP Server
     Loaded: loaded (/lib/systemd/system/apache2.service; enabled; vendor preset: enabled)
     Active: active (running) since Tue 2025-12-30 13:54:22 UTC; 1min 9s ago
       Docs: https://httpd.apache.org/docs/2.4/
   Main PID: 2427 (apache2)
      Tasks: 55 (limit: 1125)
     Memory: 5.1M
        CPU: 29ms
     CGroup: /system.slice/apache2.service
             ├─2427 /usr/sbin/apache2 -k start
             ├─2429 /usr/sbin/apache2 -k start
             └─2430 /usr/sbin/apache2 -k start

Dec 30 13:54:22 ip-172-31-12-139 systemd[1]: Starting The Apache HTTP Server...
Dec 30 13:54:22 ip-172-31-12-139 systemd[1]: Started The Apache HTTP Server.
ubuntu@ip-172-31-12-139:~$
```

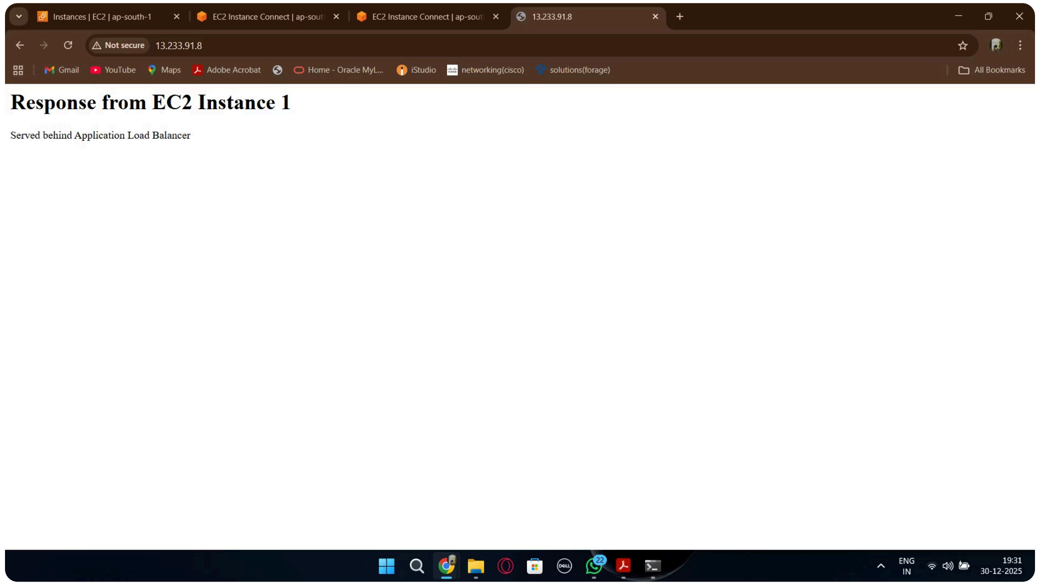
i-0e1da93a7eb07fc15 (alb-ec2-1)

PublicIPs: 13.233.91.8 PrivateIPs: 172.31.12.139

Apache web server is installed and started on the EC2 instance using systemctl commands. The service status shows Apache is active and running successfully.

Web Page Response from EC2 Instances

Web Page Response from EC2 Instance 1

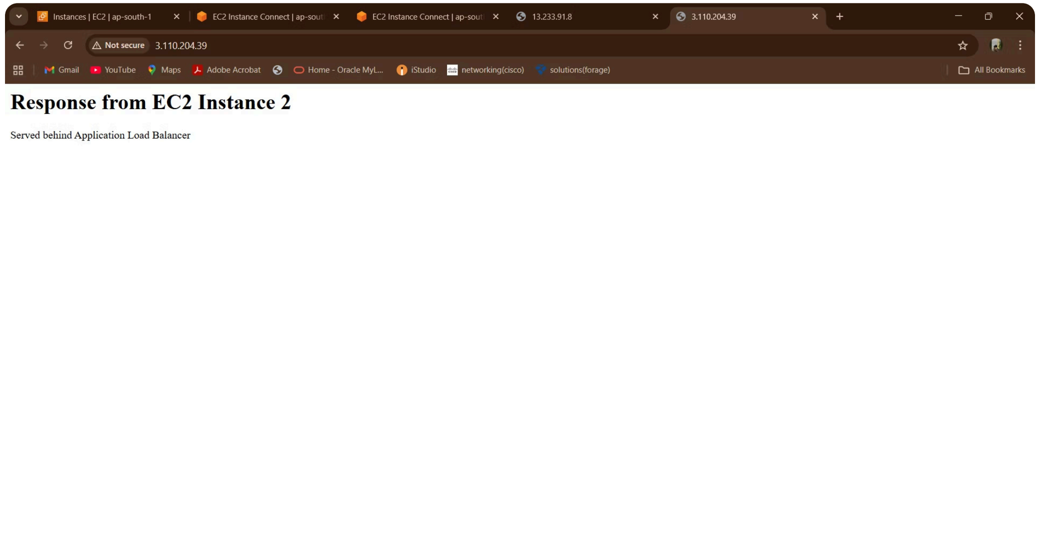


Response from EC2 Instance 1

Served behind Application Load Balancer

This output confirms that EC2 Instance 1 is serving HTTP requests correctly. The custom message helps identify responses coming from this specific instance.

Web Page Response from EC2 Instance 2



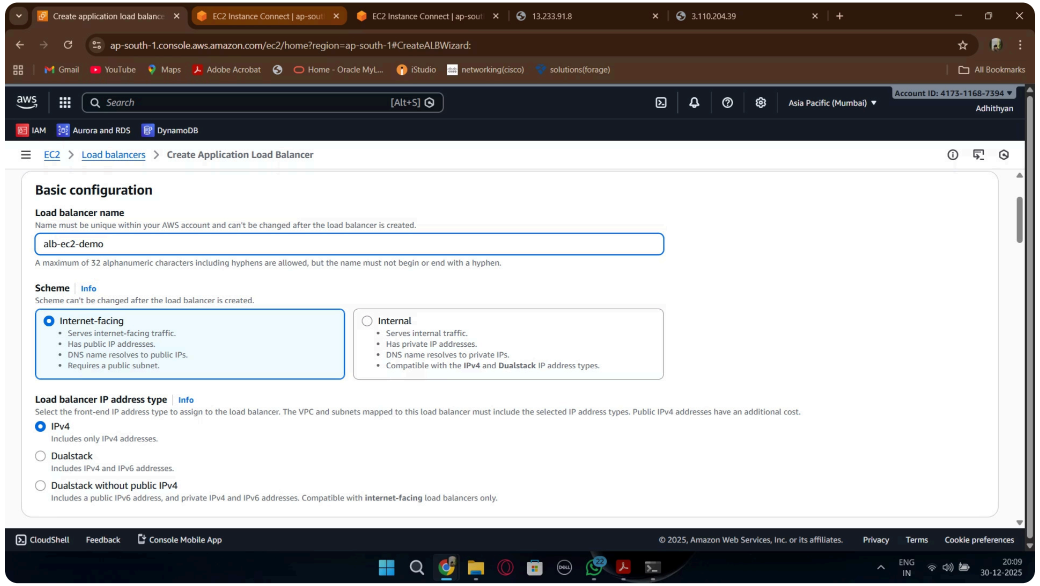
Response from EC2 Instance 2

Served behind Application Load Balancer

This output verifies that EC2 Instance 2 is also serving HTTP traffic properly. Different response text is used to distinguish traffic routing between instances.

Application Load Balancer Creation (Basic Configuration)

This screen shows the creation of an Internet-facing Application Load Balancer. The ALB is configured to distribute incoming HTTP traffic across multiple AZs.



EC2 > Load balancers > Create Application Load Balancer

Basic configuration

Load balancer name

Name must be unique within your AWS account and can't be changed after the load balancer is created.

alb-ec2-demo

A maximum of 32 alphanumeric characters including hyphens are allowed, but the name must not begin or end with a hyphen.

Scheme Info

Scheme can't be changed after the load balancer is created.

● Internet-facing
- Serves internet-facing traffic.
- Has public IP addresses.
- DNS name resolves to public IPs.
- Requires a public subnet.

○ Internal
- Serves internal traffic.
- Has private IP addresses.
- DNS name resolves to private IPs.
- Compatible with the **IPv4** and **Dualstack** IP address types.

Load balancer IP address type Info

Select the front-end IP address type to assign to the load balancer. The VPC and subnets mapped to this load balancer must include the selected IP address types. Public IPv4 addresses have an additional cost.

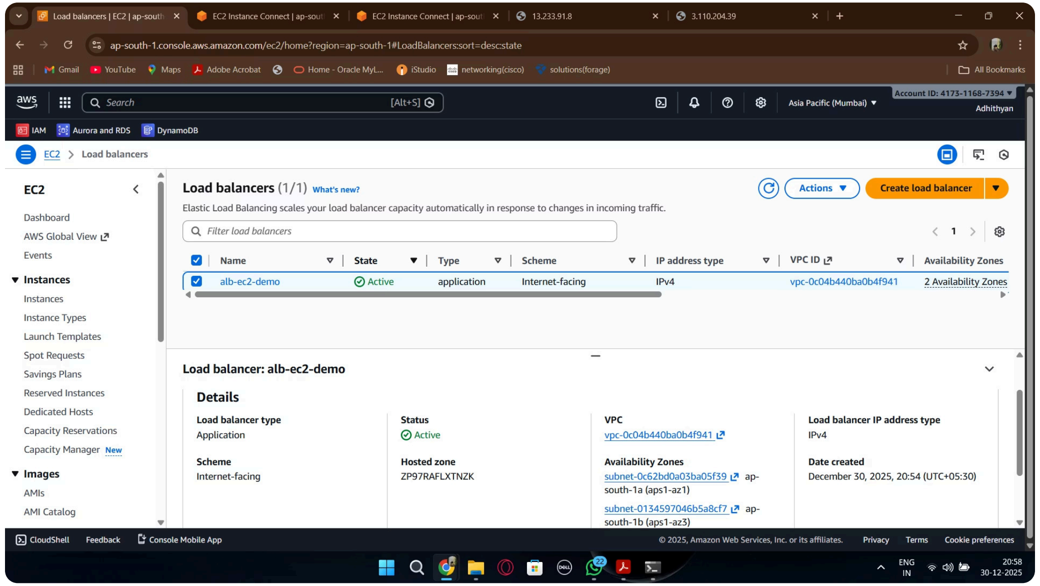
● IPv4
Includes only IPv4 addresses.

○ Dualstack
Includes IPv4 and IPv6 addresses.

○ Dualstack without public IPv4
Includes a public IPv6 address, and private IPv4 and IPv6 addresses. Compatible with **internet-facing** load balancers only.

Load Balancer Active State

The Application Load Balancer is successfully created and in an active state. It confirms that the ALB is ready to accept and route client requests.



EC2 > Load balancers

Load balancers (1/1) What's new?

Elastic Load Balancing scales your load balancer capacity automatically in response to changes in incoming traffic.

☑	Name	State	Type	Scheme	IP address type	VPC ID	Availability Zones
☑	alb-ec2-demo	✓ Active	application	Internet-facing	IPv4	vpc-0c04b440ba0b4f941	2 Availability Zones

Load balancer: alb-ec2-demo

Details

Load balancer type
Application

Scheme
Internet-facing

Status
✓ Active

Hosted zone
ZP97RAFLXTNZK

VPC
vpc-0c04b440ba0b4f941

Availability Zones
subnet-0c62bd0a03ba05f39 ap-south-1a (aps1-az1)
subnet-0134597046b5a8cf7 ap-south-1b (aps1-az3)

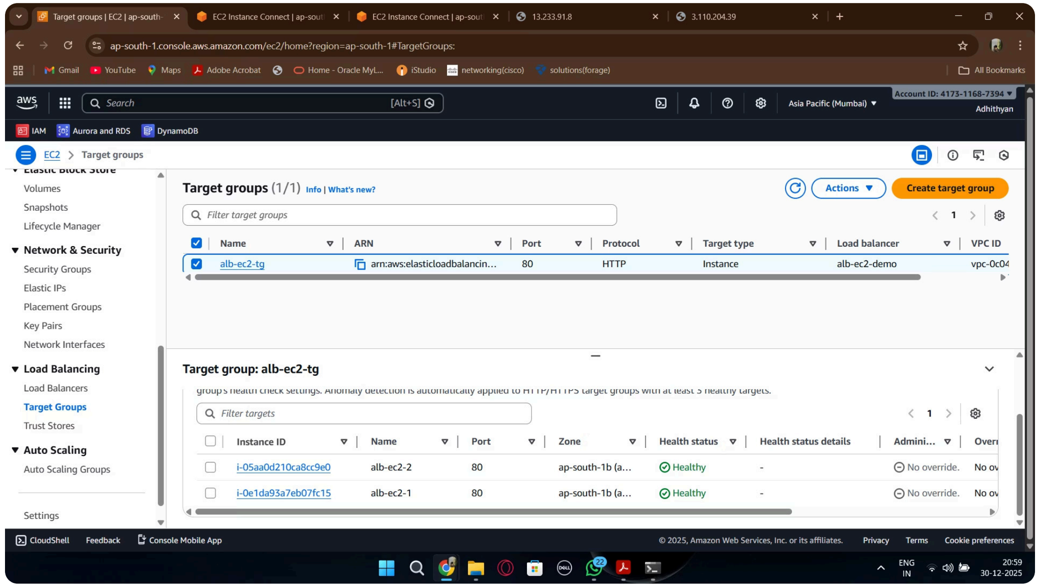
Load balancer IP address type
IPv4

Date created
December 30, 2025, 20:54 (UTC+05:30)

Target Group & Load Balancer Configuration

Target Group Attached to Load Balancer

This screenshot shows the target group linked to the Application Load Balancer. Traffic received by the ALB will now be forwarded to healthy EC2 instances.



Target groups (1/1) Info | What's new?

☑	Name	ARN	Port	Protocol	Target type	Load balancer	VPC ID
☑	alb-ec2-tg	arn:aws:elasticloadbalancin…	80	HTTP	Instance	alb-ec2-demo	vpc-0c04

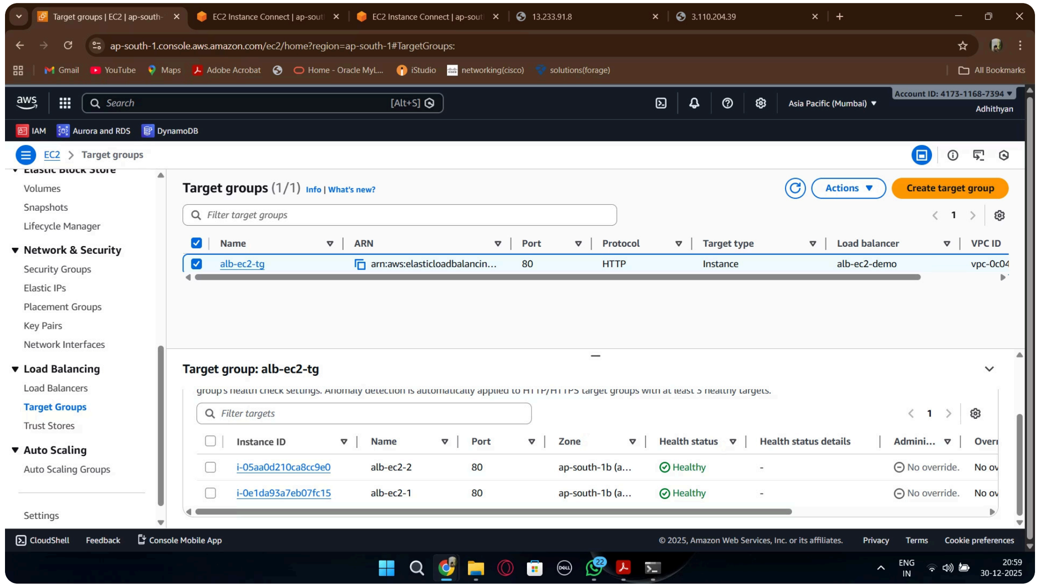
Target group: alb-ec2-tg

group's health check settings. Anomaly detection is automatically applied to HTTP/HTTPS target groups with at least 3 healthy targets.

□	Instance ID	Name	Port	Zone	Health status	Health status details	Admini…	Overr
□	i-05aa0d210ca8cc9e0	alb-ec2-2	80	ap-south-1b (a…	✓ Healthy	-	No override.	No ov
□	i-0e1da93a7eb07fc15	alb-ec2-1	80	ap-south-1b (a…	✓ Healthy	-	No override.	No ov

Target Group Details

Complete target group configuration showing health check settings and instance registration. This ensures only healthy instances receive traffic from the load balancer.



Target groups (1/1) Info | What's new?

☑	Name	ARN	Port	Protocol	Target type	Load balancer	VPC ID
☑	alb-ec2-tg	arn:aws:elasticloadbalancin…	80	HTTP	Instance	alb-ec2-demo	vpc-0c04

Target group: alb-ec2-tg

group's health check settings. Anomaly detection is automatically applied to HTTP/HTTPS target groups with at least 3 healthy targets.

□	Instance ID	Name	Port	Zone	Health status	Health status details	Admini…	Overr
□	i-05aa0d210ca8cc9e0	alb-ec2-2	80	ap-south-1b (a…	✓ Healthy	-	No override.	No ov
□	i-0e1da93a7eb07fc15	alb-ec2-1	80	ap-south-1b (a…	✓ Healthy	-	No override.	No ov

Key Takeaways

This project demonstrates the deployment of a highly available web application using AWS Application Load Balancer and EC2 instances. By distributing traffic across multiple EC2 instances in different Availability Zones, the solution ensures improved reliability, scalability, and fault tolerance.

Hands-on AWS Experience

Practical implementation of EC2, ALB, Target Groups, Security Groups, and VPC networking concepts.

Production-Ready Architecture

Real-world cloud architecture practices reflecting industry standards for resilient applications.

Foundation for Scaling

Solid groundwork for integrating Auto Scaling groups and advanced load balancing strategies.