

Task 1 –

Create a Highly Available Web App Using ALB + EC2 Auto Scaling with architect diagram and explanation.

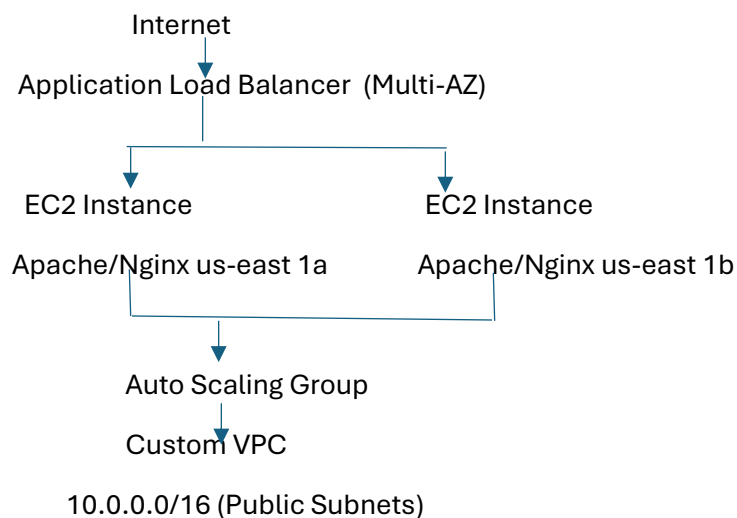
Goal: Deploy a simple web application that can scale and stay available.

Steps:

1. Launch 2 EC2 instances in different AZs with a simple Apache/Nginx page.
2. Create a Target Group and register instances.
3. Create an Application Load Balancer (ALB).
4. Set up Auto Scaling Group with:
 - a. Min: 1
 - b. Max: 3
 - c. CPU > 60% → scale out
 - d. CPU < 20% → scale in
5. Test failover by stopping one instance—web app should still load.

Ans ->

Diagram ->



Flow of request ->

1. User opens ALB DNS in browser

2. Request hits Application Load Balancer
3. ALB checks target health
4. Traffic forwarded to a healthy EC2
5. Apache serves response
6. ASG maintains desired capacity