AWS Project: Scalable EC2 Deployment with Load Balancer

This documentation showcases the process, architecture, and validation of deploying a scalable web application using EC2 instances behind a Load Balancer in a private VPC, with Nginx serving a custom page on each node.

Architecture Overview

The solution leverages the following AWS resources:

- Virtual Private Cloud (VPC) with public and private subnets in two Availability Zones
- Linux web servers (EC2) running Nginx in private subnets
- Application Load Balancer distributing HTTP traffic
- Jump server in a public subnet for SSH
- Security group and route table setup for segmentation

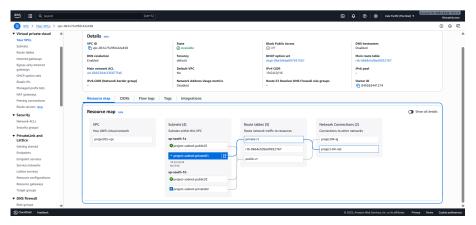
Step-by-Step Process

1. VPC and Subnet Design

Custom VPC with /16 range.

Two public and two private subnets in two Availability Zones.

Route tables and Internet Gateway for connectivity.

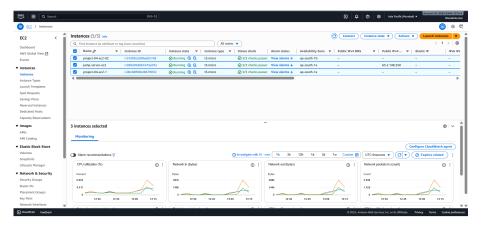


2. EC2 Instances Provisioning

Launched two Linux servers in private subnets for Nginx web hosting.

Jump-host in public subnet for SSH access.

Verified instance health and connectivity.



3. Security Group & Key Management

Configured SSH key for secure EC2 login.

Security groups allow SSH to jump host and HTTP from Load Balancer to web servers.

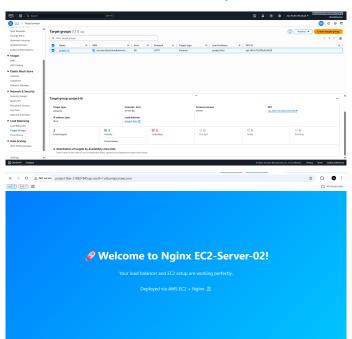


4. Nginx Setup and Validation

Nginx installed and enabled.

Customized default welcome page per instance to confirm load balancing.

Validated service health with systemctl and curl.

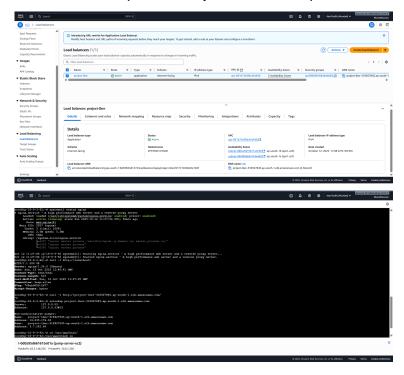


5. Load Balancer Configuration

Deployed Application Load Balancer across Availability Zones.

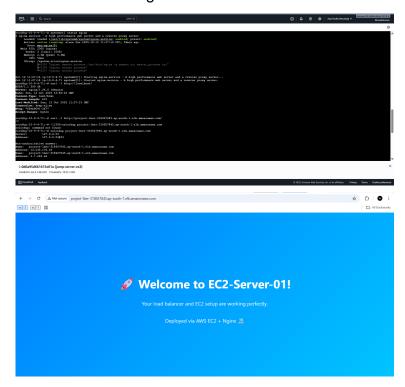
Registered web servers in target group for HTTP:80 traffic.

Both servers report healthy and serve requests.



6. End-to-End Validation

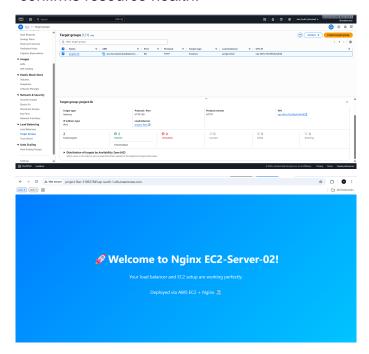
Accessed Load Balancer endpoint in browser to confirm unique greeting pages. Successful round-robin routing demonstrated.



Additional Validation

DNS resolution and HTTP headers confirmed via CLI from jump server. Network monitoring

confirms resource health.



Conclusion

This setup demonstrates the principles of secure, resilient architecture on AWS:

- Isolated subnets for security
- Centralized jump host access
- Automated high-availability load balancing
- Web application scaling foundation

All steps are validated by screenshots for reproducibility.