

The above diagram is the creation of VPC that you can use for servers

To improve resiliency, you deploy servers in two availability zones, by using an Auto scaling group and an Application load balancer. For additional security, you deploy servers in private subnets. The servers receive requests through load balancer. The servers can connect to internet using NAT gateway. To improve resiliency, you deploy NAT gateway in both availability zones.

**PROCEDURE:**

1. Create a VPC with a public subnet and private subnet in two availability zones each in a defined region
2. Create the internet gateway for public subnets and private subnets each and attach it to the VPC.
3. Create a Route Tables and associate routes between internet gateways and subnets.
4. Create NAT gateways in two availability zones and update the route association through NAT gateway for private subnets.
5. Create launch template and specify configuration with VPC we have created before.
6. Create Auto-Scaling group with previously created launch template with defined VPC and in private subnets in both availability zones.
7. Create a load balancer, which distributes traffic evenly across the instances in your Auto Scaling group, and attach the load balancer to your Auto Scaling group.

1. Create a VPC

Create a VPC with a CIDR block (e.g., 10.0.0.0/16).

This VPC will host all your AWS resources.

2. Create Subnets

Public Subnets: Create at least two public subnets, one in each Availability Zone (AZ).

Public subnet(used for all public openly)

Private Subnets: Create two private subnets, again one in each AZ.

Private subnets (used by organizations, personal use only)

3. Internet Gateway and Route Tables

Its a gateway for accessing the internet through public and private networks.

Route table will be able to route the traffic and route the request b/w our igw into public and private network.

Attach an Internet Gateway (IGW) to the VPC for internet access.

Create a route table for the public subnets and add a route to the IGW (0.0.0.0/0 → IGW).

Associate the public subnets with this route table.

4. NAT Gateway (One per AZ)

Launch NAT Gateways in each public subnet (high availability).

Allocate Elastic IPs for each NAT Gateway.

Create route tables for private subnets with a route to the NAT Gateways (0.0.0.0/0 → NAT Gateway), allowing outbound internet access from private instances.

5. Create Security Groups

Security Group for ALB (Application Load Balancer): Allow HTTP/HTTPS traffic from the internet.

Security Group for EC2 instances: Allow traffic only from the ALB security group on the required port (e.g., port 80).

6. Set Up the Application Load Balancer (ALB)

Deploy the ALB in the public subnets across AZs.

Configure target groups pointing to EC2 instances in private subnets.

Set up listeners (HTTP/HTTPS) and routing rules.

7. Launch EC2 Instances in Private Subnets

Create an Auto Scaling Group (ASG) that launches EC2 instances in private subnets.

Associate the EC2 instances with the target group of the ALB.

Ensure EC2s are part of the correct security group.

8. Configure Auto Scaling Group

Define launch templates/configurations for EC2s.