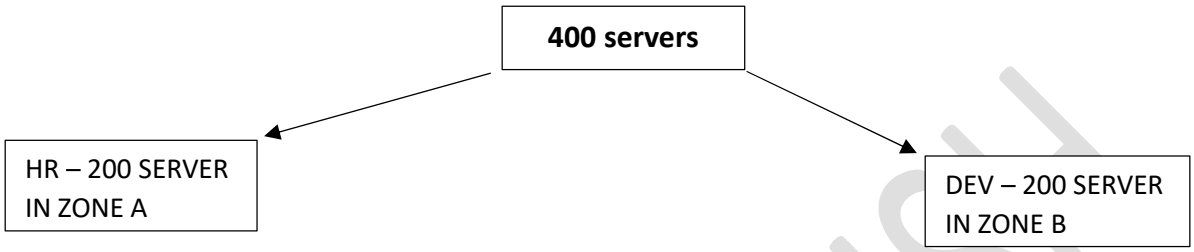
Create 400 Servers in Cloud, but here 200 Server Work in Zone A & 200 Server Work in Zone B



WHAT IS ZONE?

"ZONE" in cloud computing like a specific area or location where your data and applications are stored and processed.

Imagine a big library with many buildings (regions) and each building has multiple floors (zones). Each floor has its own set of books and resources, and they're all connected, but separate from other floors.

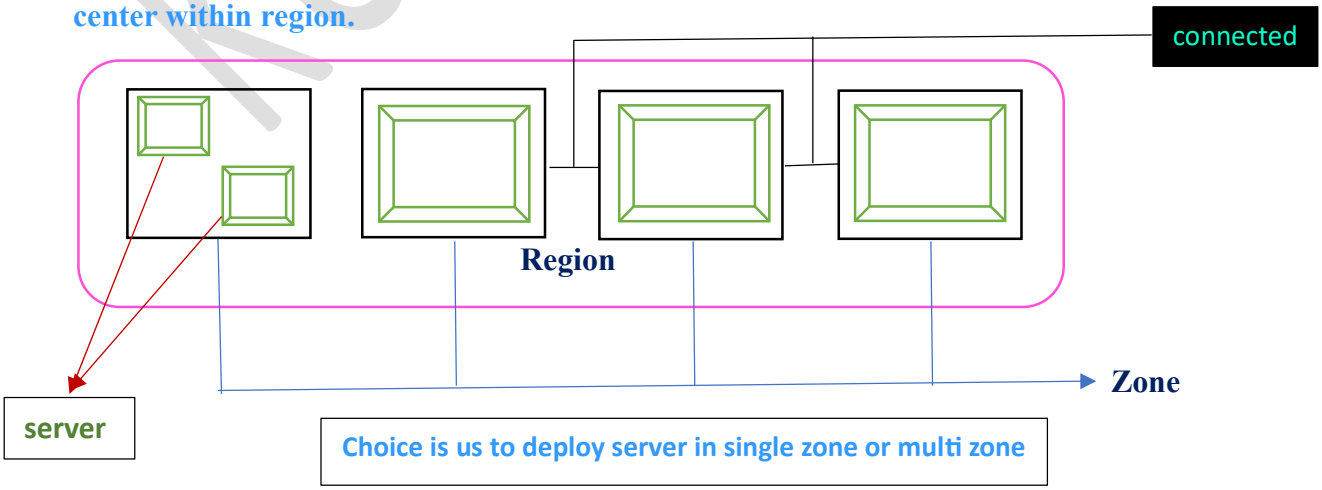
In cloud computing, zones are like those floors. They're separate areas within a larger region, and they can have their own:

- Resources (like computers or storage)

- Security settings

- Network connections

Zone: - zone work within region, cloud service provider creates multiple physical data center within region.

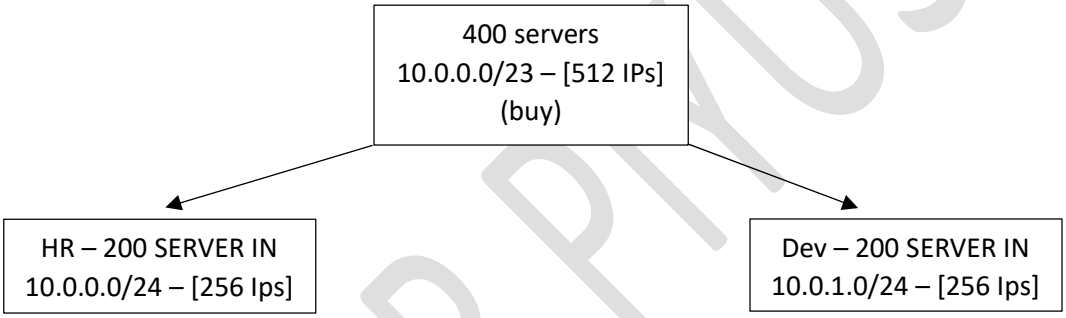


If we create our server in same zone and if one server is crash due to some reason then another server gives support.

Same as if we create server in different zones and these servers are connected to each other and if one server is crashed then other server will gives the support to another server.

But if zone is crashed then whole server is lost in multi-server zone.

But if one zone is crashed in single-server zone then other server gives support to another zone.



When we allocate Ips to any distributed server then we have to

SELECT ZONE

What are IPs (IP addresses)?

An **IP address** is a unique identifier assigned to devices on a network, like your phone, laptop, or server. IPv4 addresses are typically leased in **blocks** (like 256 or 512 addresses at a time), not individually.

"Jaise saman 185 rupaye ka hai par 200 ka note dena padta hai kyunki 185 ka note nhi hai."

Just like that. The government doesn't print a ₹185 note, so you give ₹200 and get ₹15 back.

SIMILARLY:

"400 IPs chahiye but vo exact milte nhi, to 512 IPs lene padte hain."

Exactly. In networking, IPs are allocated in blocks of powers of 2:

/24 → 256 IPs

/23 → 512 IPs

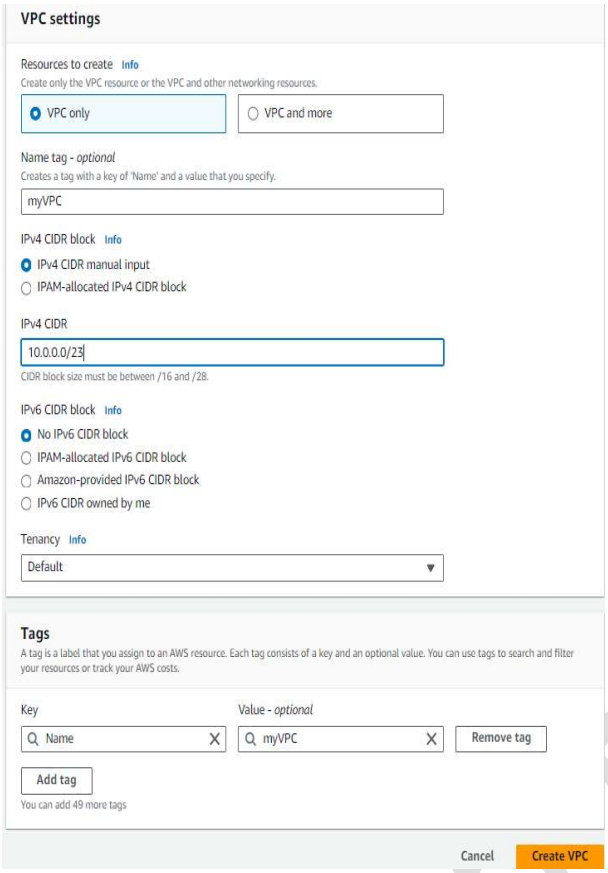
/22 → 1024 IPs ...and so on.

So, if you need 400 IPs, you can't buy exactly 400.

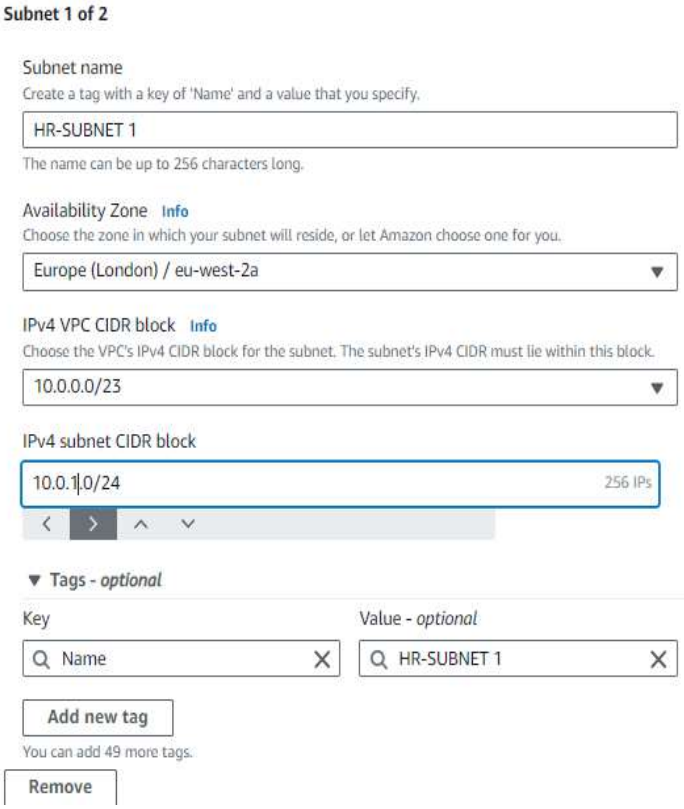
The nearest block that includes at least 400 IPs is /23 = 512 IPs. That's the smallest possible block that satisfies your requirement.

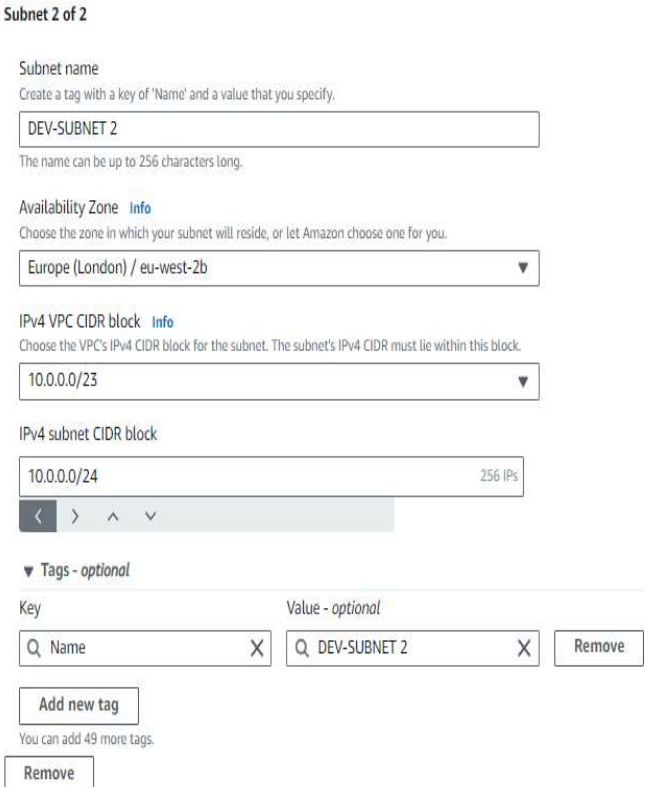
Plan to Create 400 Servers in AWS, split across ZONE in each subnet (HR – A) & (DEV – B), with internet and public IP access

1. Login
2. Select (Region)
3. VPC (Create or Use existing)
4. Subnet (Create 2 – one in Zone A, one in Zone B)
5. IGW (Attach Internet Gateway to VPC)
6. Route (Update Route Table to allow internet access)
7. SG (Create Security Group for public access)
8. Keypair (Create for SSH)
9. AMI (Choose base image – Amazon)
10. Launch (Use EC2 Launch Template)
11. Configure (Select subnet A or B, enable Auto-assign Public IP)
12. Auto Scale (Create two Auto Scaling Groups:
    – 200 in Subnet A
    – 200 in Subnet B
13. Verify (Check public IPs and internet access)


VPC settings
Resources to create Info
Create only the VPC resource or the VPC and other networking resources.
VPC only
VPC and more
Name tag - optional
Creates a tag with a key of 'Name' and a value that you specify.
myVPC
IPv4 CIDR block Info
IPv4 CIDR manual input
IPAM-allocated IPv4 CIDR block
IPv4 CIDR
10.0.0.0/23
CIDR block size must be between /16 and /28.
IPv6 CIDR block Info
No IPv6 CIDR block
IPAM-allocated IPv6 CIDR block
Amazon-provided IPv6 CIDR block
IPv6 CIDR owned by me
Tenancy Info
Default
Tags
A tag is a label that you assign to an AWS resource. Each tag consists of a key and an optional value. You can use tags to search and filter your resources or track your AWS costs.
Key
Value - optional
Name
myVPC
Remove tag
Add tag
You can add 49 more tags
Cancel
Create VPC


VPC
VPC ID
Create subnets in this VPC.
vpc-0e3cd64e603ce064d (myVPC)
Associated VPC CIDRs
IPv4 CIDRs
10.0.0.0/23


Subnet 1 of 2
Subnet name
Create a tag with a key of 'Name' and a value that you specify.
HR-SUBNET 1
The name can be up to 256 characters long.
Availability Zone Info
Choose the zone in which your subnet will reside, or let Amazon choose one for you.
Europe (London) / eu-west-2a
IPv4 VPC CIDR block Info
Choose the VPC's IPv4 CIDR block for the subnet. The subnet's IPv4 CIDR must lie within this block.
10.0.0.0/23
IPv4 subnet CIDR block
10.0.1.0/24
256 IPs
Tags - optional
Key
Value - optional
Name
HR-SUBNET 1
Add new tag
You can add 49 more tags.
Remove

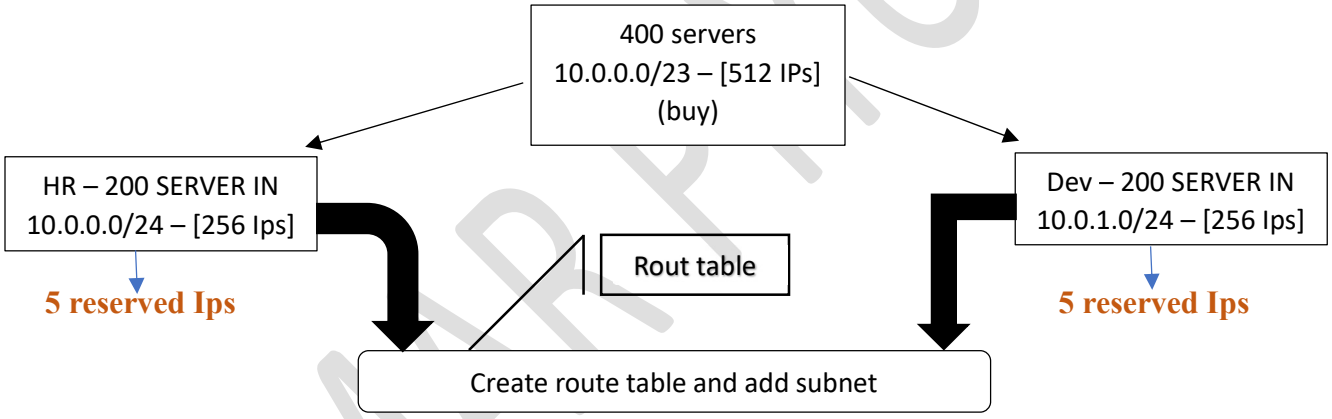

Subnet 2 of 2
Subnet name
Create a tag with a key of 'Name' and a value that you specify.
DEV-SUBNET 2
The name can be up to 256 characters long.
Availability Zone Info
Choose the zone in which your subnet will reside, or let Amazon choose one for you.
Europe (London) / eu-west-2b
IPv4 VPC CIDR block Info
Choose the VPC's IPv4 CIDR block for the subnet. The subnet's IPv4 CIDR must lie within this block.
10.0.0.0/23
IPv4 subnet CIDR block
10.0.0.0/24
256 IPs
Tags - optional
Key
Value - optional
Name
DEV-SUBNET 2
Remove
Add new tag
You can add 49 more tags.
Remove

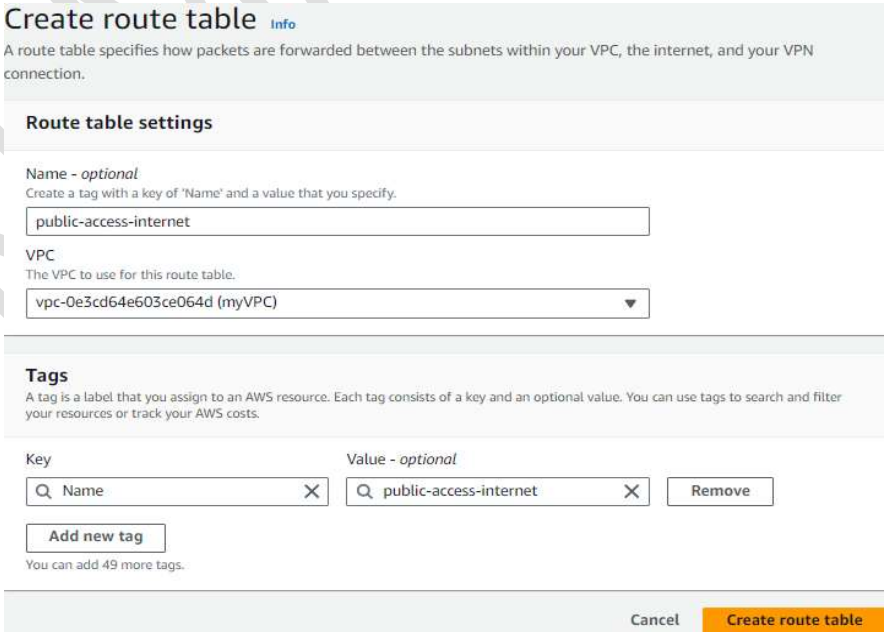
Name	Subnet ID	State	VPC	IPv4 CIDR	IPv6 CIDR	Available IPv4 addresses	Availability Zone
HR-SUBNET 1	subnet-02af4f1798c6f9f60	Available	vpc-0e3cd64e603ce064d | myVPC	10.0.1.0/24	-	251	eu-west-2a
DEV-SUBNET 2	subnet-05b44e982d0e07251	Available	vpc-0e3cd64e603ce064d | myVPC	10.0.0.0/24	-	251	eu-west-2b

Here, we allocate 256 IPs to each subnet, but we see that only 251 IPs are usable because AWS reserves 5 IPs in each subnet for internal communication.

These reserved IPs are:

1. NETWORK IP
2. GATEWAY IP
3. DNS 1
4. DNS 2
5. BROADCAST





Then, go to the Route Tables, select the created route table, click on Actions, and choose Edit Subnet Associations.

Select both subnets and save the changes.

NOTE:-

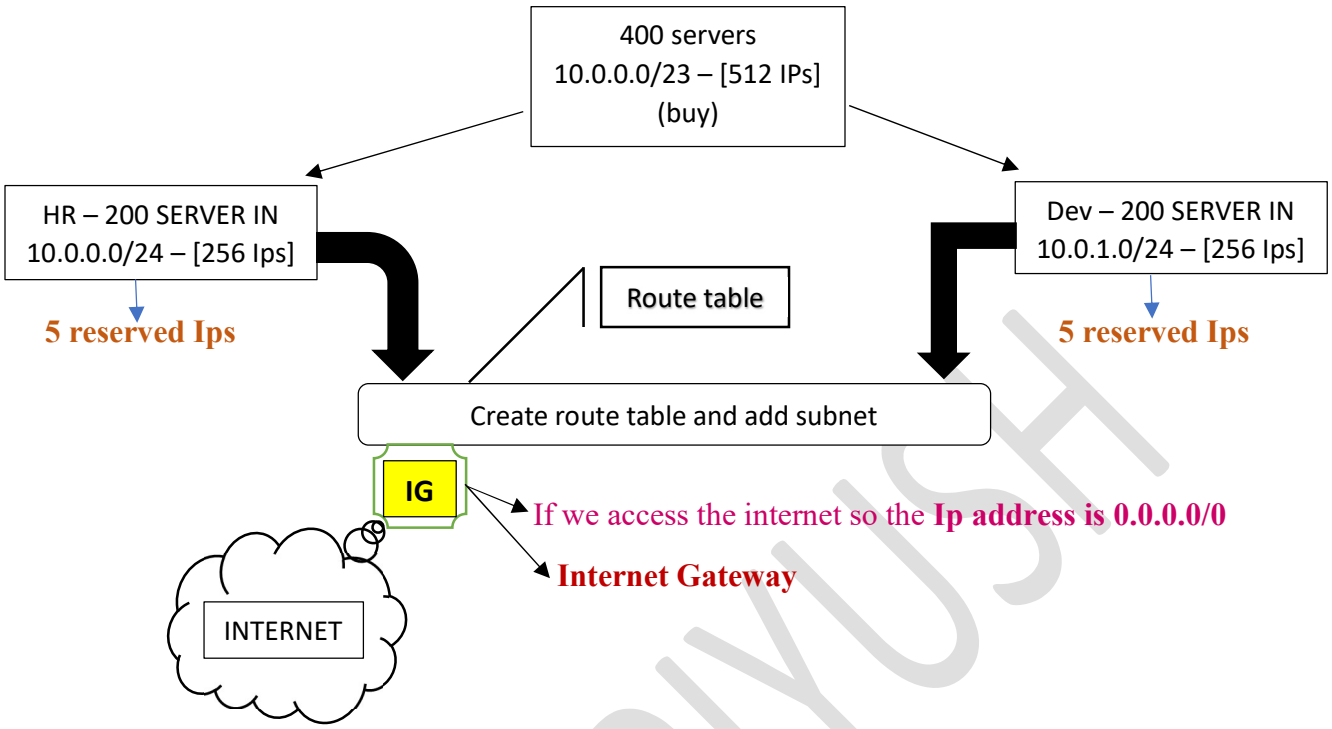
- Abhi tak hamara network traffic route table tak pahuch gaya hai, lekin ab usse internet par jane ke liye gateway ki zarurat hogi.
- Jaise hum jab ghar se bahar nikalte hain to humein gate (gateway) ki need hoti hai, waise hi cloud ke traffic ko bhi bahar nikalne ke liye Internet Gateway (IGW) chahiye hota hai.
- In short, agar aap chahte hain ki EC2 instance ka traffic internet se communicate kare, to aapko Internet Gateway VPC ke sath attach karna hoga aur route table me uska route define karna padega.

IG – INTERNET GATEWAY

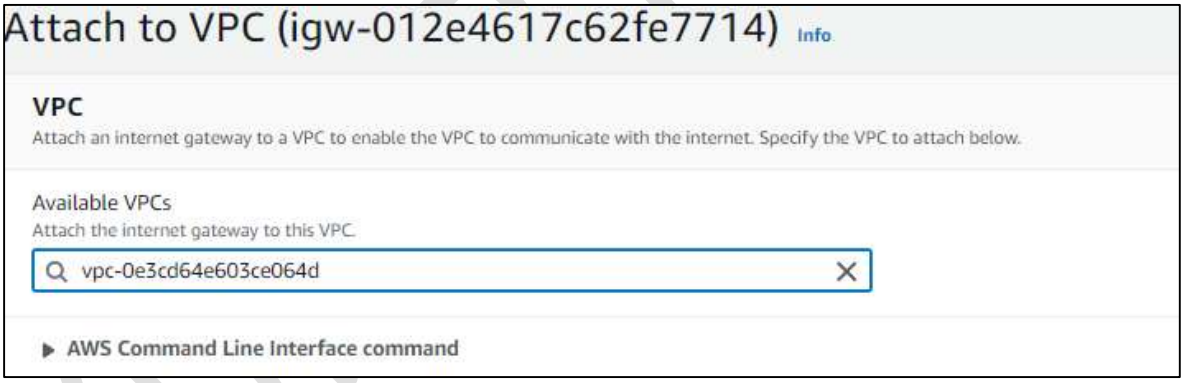
CONNECTED VIRTUAL MACHINE AND EXCESS IN WORLDWIDE VIA BY PUBLIC IP & ALSO ACCESS INTERNET.

- PRIVATE IP – **CLOUD TO CLOUD COMMUNICATION**
- PUBLIC IP – **CLOUD TO OTHER COMMUNICATION**

AND HERE WE CHOOSE PUBLIC IP .



Only one VPC can connect only one IG



We create 200 servers in each zone to ensure:

1. High Availability (HA)

If one zone (e.g., us-east-1a) goes down due to failure or maintenance, the other zone (e.g., us-east-1b) still has 200 servers running. This ensures continuous service without downtime.

2. Fault Tolerance

Distributing servers across zones helps us tolerate failures in one zone. AWS Availability Zones are physically separate, so issues like power failure, natural disasters, or hardware failure in one zone do not impact the other.

3. Load Balancing

By splitting servers evenly, traffic can be balanced between both zones using a load balancer, improving performance and reducing latency.

4. Scalability & Resource Distribution

Some AWS regions have quotas or limits per zone. Splitting helps stay within limits, and improves resource availability (CPU, IPs, etc.).

In Short:

Distributing 200-200 EC2 instances across 2 AZs = better uptime, reliability, and performance.