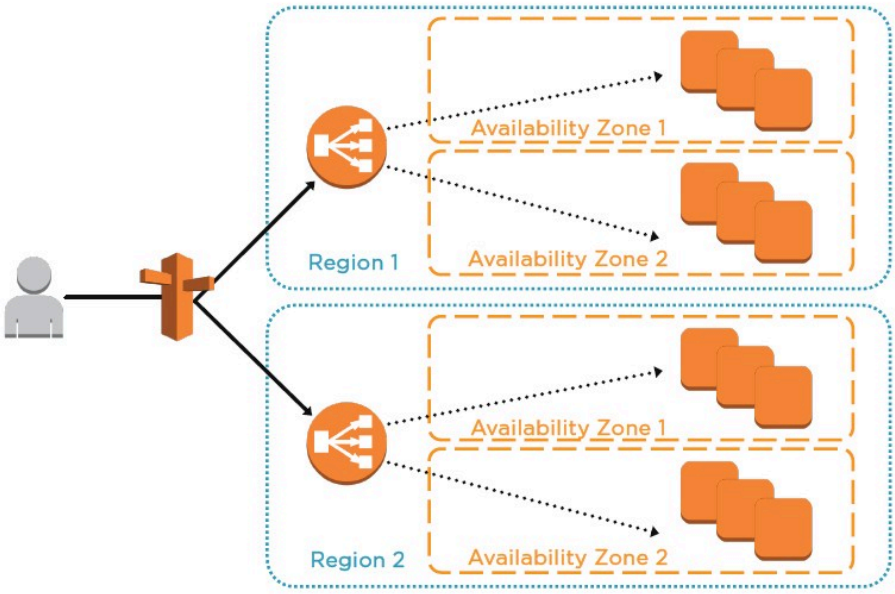
Elastic Load Balancer

Elastic Load Balancing (ELB) is an AWS service that automatically distributes incoming application traffic across multiple targets—such as EC2 instances, containers, and IP addresses—in one or more Availability Zones (AZs). It improves fault tolerance, availability, and scalability by balancing traffic and performing health checks.



The clients send the requests t, the request goes to the load balancer and it is send to the healthy region/AZ.

Types :

- Application load balancer
- Network load balancer
- Classic load balancer

- Gateway load balancer

ELB Type	When It's Used	Protocol	Example Use Case
ALB (Application Load Balancer)	For web apps needing HTTP/HTTPS with routing logic	Layer 7 (HTTP/HTTPS)	Microservices, API Gateway-style routing
NLB (Network Load Balancer)	For high-performance, low-latency TCP/UDP traffic	Layer 4 (TCP/UDP)	Gaming, IoT, video streaming, financial apps
CLB (Classic Load Balancer)	Legacy apps needing simple TCP/HTTP load balancing	Layer 4 & 7	Old web apps with minimal routing logic
GWLB (Gateway Load Balancer)	For traffic inspection and third-party appliance insertion	Layer 3/4	Firewall, Intrusion Detection (IDS/IPS)