



Amazon **ELB**

❖ ELASTIC LOAD BALANCER:

- Amazon Elastic Load Balancing automatically distributes incoming application traffic across multiple Amazon EC2 instances. This increases the fault tolerance of your applications.
- An ELB can load balance traffic to multiple EC2 instances located across multiple availability zones.
- ELB should be paired with Auto Scaling to enhance high availability and fault tolerance, and allow for automated scalability and elasticity.
- An ELB has its own DNS record set that allows for direct access from the open internet access.
- ELB's will automatically stop serving traffic to an instance that becomes unhealthy. (Via health checks)
- When you create a load balancer, you must choose whether to make it an internal load balancer or an internet-facing load balancer.

INTERNET-FACING:

- The nodes of an internet-facing load balancer have public IP addresses. The DNS name of an internet-facing load balancer is publicly resolvable to the public IP addresses of the nodes. Therefore, internet-facing load balancers can route requests from clients over the internet.

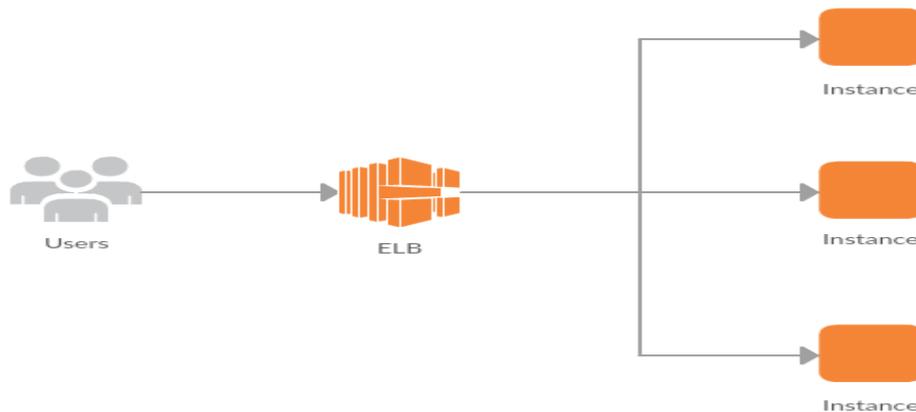
INTERNAL-FACING:

- The nodes of an internal load balancer have only private IP addresses. The DNS name of an internal load balancer is publicly resolvable to the private IP addresses of the nodes. Therefore, internal load balancers can only route requests from clients with access to the VPC for the load balancer.
- Both internet-facing and internal load balancers route requests to your targets using private IP addresses. Therefore, your targets do not need public IP addresses to receive requests from an internal or an internet-facing load balancer.
- If your application has multiple tiers, for example web servers that must be connected to the internet and database servers that are only connected to the web servers, you can design an architecture that uses both internal and internet-facing load balancers.

➤ **WHY USE ELB:**

- Spread load across multiple downstream instances
- Expose a single point of access (DNS) to your application
- Seamlessly handle failures of downstream instances
- Do regular health checks to your instances
- Provide SSL termination (HTTPS) for your websites
- High availability across zones
- Separate public traffic from private traffic

APPLICATION AWS INFRASTRUCTURE WITH ELB & MULTIPLE EC2 INSTANCES:



➤ **BENEFITS:**

- Highly available
- Secure
- Elastic
- Flexible
- Robust monitoring & auditing
- Hybrid load balancing
- Health checks
- Layer 4 or Layer 7 load balancing
- TLS termination

➤ **USE CASES:**

- Achieve better fault tolerance for your applications
- Automatically load balance your containerized applications
- Automatically scale your applications
- Using Elastic Load Balancing in your Amazon Virtual Private Cloud
- Hybrid load balancing with Elastic Load Balancing
- Invoking Lambda functions over HTTP(S)

➤ **ELB TYPES (PRODUCTS):**

- Elastic Load Balancing (ELB) supports four types of load balancers. You can select the appropriate load balancer based on your application needs.
 - Application Load Balances (ALB)
 - Network Load Balancer (NLB)
 - Gateway Load balancer (GLB)
 - Classic Load Balancer (CLB)

APPLICATION LOAD BALANCER (ALB):

- An “Application” elastic load balancer is designed for complex balancing of traffic to multiple EC2 instances using “Content-based” rules.
- If you need to load balance HTTP requests, we recommend you use the Application Load Balancer (ALB).
- The seventh layer of the Open Systems Interconnection (OSI) model.
- An Application Load Balancer supports targets with any operating system currently supported by the Amazon EC2 service.
- An Application Load Balancer supports load balancing of applications using HTTP and HTTPS (Secure HTTP) protocols.
- You can perform load balancing for the following TCP ports: 1-65535

CONTENT-BASED RULES:

HOST-BASED: Route traffic based on the host field of the HTTP header.

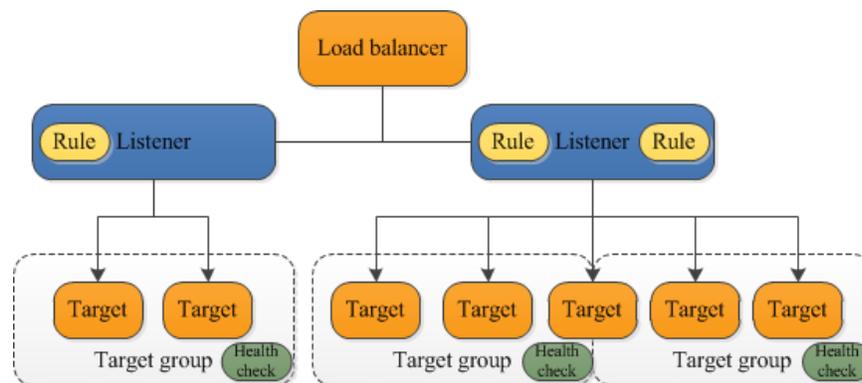
PATH-BASED: Route traffic based on the URL path of the HTTP header.

HTTP HEADER-BASED ROUTING: You can route a client request based on the value of any standard or custom HTTP header.

HTTP METHOD-BASED ROUTING: You can route a client request based on any standard or custom HTTP method.

QUERY STRING PARAMETER-BASED ROUTING: You can route a client request based on query string or query parameters.

SOURCE IP ADDRESS CIDR-BASED ROUTING: You can route a client request based on source IP address CIDR from where the request originates.

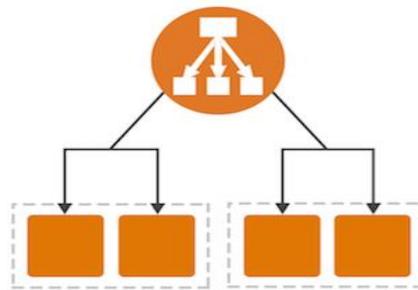


KEY FEATURES:

- Layer-7 Load Balancing
- HTTPS Support
- Server Name Indication (SNI)
- IP addresses / Lambda functions as Targets
- High Availability
- Security Features
- Content-based Routing
- Containerized Application Support
- HTTP/2 Support
- WebSocket's Support
- Sticky Sessions
- Operational Monitoring
- Request Tracing
- Web Application Firewall
- Slow Start Mode with Load-Balancing Algorithm
- User Authentication
- Redirects
- Fixed Response

NETWORK LOAD BALANCER (NLB):

- A Network Load Balancer functions at the fourth layer of the Open Systems Interconnection (OSI) model.
- It can handle millions of requests per second.
- Support for static IP and Elastic IP.
- Network Load Balancers are mostly used for extreme performance.
- It is integrated with other popular AWS services such as Auto Scaling, Amazon EC2 Container Service (ECS), Amazon CloudFormation, and AWS Certificate Manager (ACM).
- Overall, the creation process is the same as Application Load Balancers.

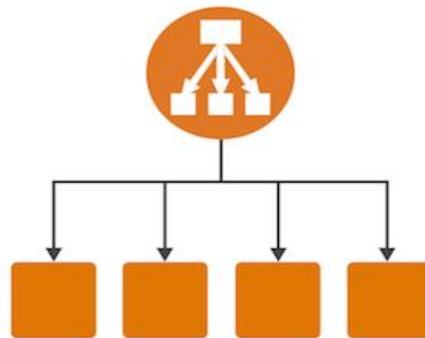


GATEWAY LOAD BALANCER (GLB):

- Gateway Load Balancer helps you easily deploy, scale, and manage your third-party virtual appliances.
- It enables you to deploy, scale, and manage virtual appliances, such as firewalls, intrusion detection and prevention systems, and deep packet inspection systems.
- It combines a transparent network gateway (that is, a single entry and exit point for all traffic) and distributes traffic while scaling your virtual appliances with the demand.
- A Gateway Load Balancer operates at the third layer of the Open Systems Interconnection (OSI) model, the network layer. It listens for all IP packets across all ports and forwards traffic to the target group that's specified in the listener rule.
- Gateway Load Balancers use Gateway Load Balancer endpoints to securely exchange traffic across VPC boundaries. A Gateway Load Balancer endpoint is a VPC endpoint that provides private connectivity between virtual appliances in the service provider VPC and application servers in the service consumer VPC.

CLASSIC LOAD BALANCER (CLB):

- CLB provides basic load balancing across multiple Amazon EC2 instances and operates at both the request level and connection level.
- Classic Load Balancer is intended for applications that are built within the EC2-Classic network.
- We recommend Application Load Balancer for Layer 7 traffic and Network Load Balancer for Layer 4 traffic when using Virtual Private Cloud (VPC).



➤ HEALTH CHECKS:

- Health Checks are crucial for Load Balancers
- They enable the load balancer to know if instances it forwards traffic to are available to reply to requests
- The health check is done on a port and a route (/health is common)
- If the response is not 200 (OK), then the instance is unhealthy

