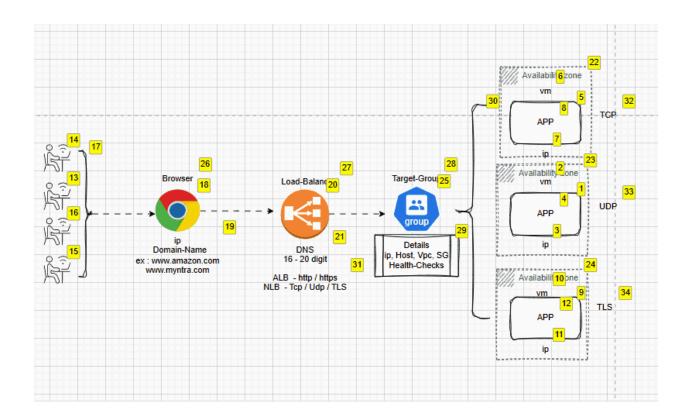
A **Network Load Balancer (NLB)** is a type of load balancer that distributes incoming network (usually TCP/UDP) traffic across multiple backend servers (also called targets) at the transport layer (Layer 4) of the OSI model. It is designed to handle high volumes of traffic with ultra-low latency.



# Ney Components of a Network Load Balancer

Here are the main components of a Network Load Balancer:

#### 1. Listeners

**Definition**: A listener checks for connection requests from clients.

- **Details**: It defines the protocol (e.g., TCP, UDP) and the port (e.g., 80, 443) that the NLB uses to listen for incoming traffic.
- **Example**: A listener on port 80 using TCP.

## 2. Targets

- **Definition**: These are the destination systems where the NLB sends the traffic.
- **Types**: Targets can be:
  - EC2 instances (in AWS)
  - Containers (e.g., in Kubernetes)
  - o IP addresses
- **Target Groups**: Targets are grouped together into target groups for better management.

# 3. Target Groups

- **Definition**: Logical groupings of backend targets.
- Functions:
  - Health checks are defined per group.
  - Routing decisions are made based on group rules.
- **Example**: All web servers might be in one target group.

#### 4. Health Checks

• **Definition**: Automated tests performed by the NLB to check if a target is healthy (able to handle traffic).

#### • Parameters:

- Protocol (e.g., TCP)
- o Port
- Thresholds (e.g., 3 failed checks = unhealthy)
- **Purpose**: Ensures traffic is only sent to healthy targets.

## 5. DNS Name

- **Definition**: The entry point for client traffic, provided by the NLB.
- **Example**: When a client wants to connect, it resolves the NLB's DNS name to the appropriate IP addresses.

#### 6. Elastic IPs / Static IPs

- **Definition**: NLB can be associated with static IPs for predictable access.
- Use: Useful for whitelisting or DNS configurations.

### 7. Cross-Zone Load Balancing (optional)

- **Definition**: Distributes traffic evenly across all targets in all enabled Availability Zones.
- **Benefit**: Prevents overloading targets in a single zone.

# How It Works (Flow)

- 1. A client makes a request to the NLB's DNS name.
- 2. The DNS resolves to one or more static IPs assigned to the NLB.
- 3. The NLB listener receives the traffic on a specified port and protocol.
- 4. The load balancer picks a healthy target from the target group.
- 5. The request is forwarded directly (Layer 4 routing) to the target.
- 6. The target processes the request and returns the response directly to the client.

# Advantages of NLB

- High throughput & low latency
- Supports millions of requests per second
- Works at Layer 4 (TCP, UDP)
- Static IP and Elastic IP support
- Can preserve the client IP for backend logging or security

# Comparison: NLB vs. Other Load Balancers

Feature	Network LB (NLB)	Application LB (ALB)	Classic LB (CLB)
Layer	4 (Transport)	7 (Application)	4 & 7
Protocols	TCP, UDP, TLS	HTTP, HTTPS, gRPC	HTTP, HTTPS, TCP
Latency	Very low	Higher than NLB	Medium
Client IP Preservation	Yes	Optional	Yes