AWS Load Balancer – Detailed Notes

🔹 What is a Load Balancer?

An **AWS Load Balancer** is a **fully managed service** that automatically distributes incoming application traffic across multiple targets (like EC2, Lambda, containers) to increase fault tolerance and availability.

AWS ELB is part of the **Elastic Load Balancing** service.

🔹 Types of AWS Load Balancers

Load Balancer Type	Protocols	Best For
Application (ALB)	HTTP, HTTPS	Web applications (Layer 7)
Network (NLB)	TCP, UDP, TLS	High performance, low latency apps (Layer 4)
Gateway (GWLB)	All IP traffic	Third-party virtual appliances (firewalls, IDS)
Classic (CLB) *(Legacy)*	HTTP, HTTPS, TCP	Basic Layer 4 and Layer 7 – use ALB/NLB instead

---

1. Application Load Balancer (ALB)

Layer 7 (HTTP/HTTPS) – routes traffic based on content (host, path, headers, etc.)

Key Features:

- **Content-based routing** (e.g., /api, /images)
- **Host-based routing** (e.g., app1.example.com)
- **Supports containers** via **dynamic port mapping**
- Native **WebSocket** and **HTTP/2** support
- Integrates with **WAF**, **Cognito**, and **OIDC**
- **Fixed IP via ALB endpoints**

Target Types:

- EC2 instances
- IP addresses (internal or external)
- Lambda functions

Use Cases:

- Microservices architecture

- API Gateway replacement
- Web apps with SSL offloading

---

2. Network Load Balancer (NLB)

Layer 4 (TCP/UDP/TLS) – ultra-low latency with millions of requests per second.

Key Features:

- **High-performance, low-latency**
- Supports **TLS termination**
- Static **IP address** per AZ
- Can expose services on **non-HTTP ports**
- Integrates with **PrivateLink** and **Global Accelerator**
- **Zonal isolation** – survives AZ failures

Use Cases:

- Real-time applications (VoIP, gaming)
- Legacy systems requiring TCP
- IoT backends

---

3. Gateway Load Balancer (GWLB)

For **third-party virtual appliances** (firewalls, intrusion detection, etc.).

Key Features:

- Operates at **Layer 3 (IP)**
- Routes all traffic to the virtual appliance before reaching backend
- Uses **GWLBe (Gateway Load Balancer Endpoint)** in VPC
- Transparent traffic redirection
- Helps with **centralized inspection**

Use Cases:

- In-line security appliances
- Network traffic filtering
- Service chaining

---

4. Classic Load Balancer (CLB) – *Legacy*

Supports Layer 4 (TCP) and Layer 7 (HTTP/HTTPS).

⚠️ Recommended only for legacy applications. Use ALB/NLB for new designs.

🔹 **Key Components of ELB**

Component	Description
Target Group	Defines the targets (EC2, Lambda, etc.) for a load balancer.
Listener	A process that checks for connection requests.
Rules	Used by ALB to route traffic based on conditions.
Health Check	Periodic checks to determine the availability of targets.

🔹 **Load Balancer Listener Configuration**

- Each **Listener** has:
  - Protocol (HTTP/HTTPS for ALB, TCP for NLB)
  - Port (e.g., 80, 443)
  - Default target group
  - Optional **routing rules** (for ALB)

🔹 **Load Balancer Pricing (2024)**

Resource	Pricing Model
Load Balancer Hours	Charged per hour per Load Balancer
LCU (Load Balancer Capacity Unit)	Based on:

- New connections/sec
- Active connections
- Processed bytes
- Rule evaluations (ALB only) |
  | **Data Processed** | Charged per GB |

Use the **AWS Pricing Calculator** for exact estimates.

🔹 **Security and Monitoring**

- Integrated with **IAM** for access control.

- Supports **SSL termination** and **TLS listeners**.
- Can attach **AWS WAF** for protection (ALB only).
- Metrics via **CloudWatch**:
    - RequestCount
    - HealthyHostCount
    - TargetResponseTime
- Logging via **Access Logs** (stored in S3)

🔹 **High Availability & Resilience**

- Load balancers are **regional services** spanning multiple **Availability Zones**.
- Automatically distribute traffic across **healthy targets**.
- **Health checks** ensure only healthy targets receive traffic.

🔹 **Comparison Summary**

Feature	ALB	NLB	GWLB
OSI Layer	7	4	3
Protocols	HTTP, HTTPS	TCP, UDP, TLS	All IP
SSL Termination	Yes	Yes	No
IP Static	Via IP mode	Native	Native
Content-based Routing	Yes	No	No
Use Lambda	Yes	No	No
WebSocket	Yes	No	No

🔹 **Use Cases**

Use Case	Recommended ELB
Host-based and path-based routing	ALB
TCP or UDP load balancing	NLB
Firewall/inspection appliance	GWLB

Use Case	Recommended ELB
IoT or game traffic	NLB
API Gateway alternative	ALB
WebSocket apps	ALB

- **Best Practices**
  - **Use ALB for web and microservices**, NLB for high-performance needs.
  - Always enable **cross-zone load balancing** for even distribution.
  - Use **HTTPS with SSL certificates** (via ACM) to secure traffic.
  - Set **short health check intervals** for quicker failover.
  - Enable **access logs** and **CloudWatch alarms**.
  - Place load balancer in **multiple AZs** for high availability.