# Types of Elastic Load Balancers

As mentioned in the last section, there are different types of Elastic Load Balancers, each of which have different characteristics and serve different use cases. These are as follows:

## Application Load Balancer (ALB):

Application Load Balancers operate at the application layer, the seventh layer of the OSI model. This makes it particularly suited for dealing with HTTP and HTTPS traffic, two prominent application layer protocols. Its primary strength lies in advanced routing capabilities, which include host-based and path-based routing. This functionality allows ALBs to intelligently direct requests to different backend services based on the content of the request.

Additionally, they provide features like sticky sessions and request tracing, enhancing the user experience by maintaining session continuity which combined with their intelligent routing capabilities have led to the widespread adoption of ALBs in dynamic web applications, RESTful APIs, microservice architectures and containerized applications; all systems and services that require and/or benefit from having fine-grained routing logic.

## Network Load Balancer (NLB):

In contrast, Network Load Balancers operate at the transport layer, the fourth layer of the OSI model and is optimized for handling traffic utilizing the TCP and UDP protocols. NLBs are designed to handle millions of requests per second while maintaining ultra-low latency, making them an excellent choice for applications that demand high performance, such as gaming servers, VoIP applications, and real-time data processing.

NLBs can also operate in scenarios where static IP addresses are required, providing a more stable endpoint for applications. Furthermore, NLBs also increase a systems capacity to handling sudden spikes in traffic, and are thus often utilized in high-throughput environments.

## Gateway Load Balancer (GLB):

The Gateway Load Balancer is a newer addition to AWS's suite of load balancers, operating primarily at the third layer of the OSI model. Its primary purpose is to facilitate the deployment and scaling of third-party virtual appliances such as firewalls, intrusion detection systems, and other network security devices.

A more structured, tabular comparison of the three different types of load balancers is given below:

| **Feature** | **Application Load Balancer (ALB)** | **Network Load Balancer (NLB)** | **Gateway Load Balancer (GLB)** |
|---|---|---|---|
| **OSI Layer** | Layer 7 (Application Layer) | Layer 4 (Transport Layer) | Layer 3 (Network Layer) |
| **Protocol Support** | HTTP, HTTPS, WebSocket, HTTP/2 | TCP, UDP | Primarily IP traffic (integrates with other services) |
| **Routing Capability** | Advanced routing (host-based, path-based) | Simple routing based on IP and port | Distributes traffic to virtual appliances |
| **Performance** | Optimized for complex application traffic | High performance, low latency, millions of requests | Scalable for virtual appliances, but latency depends on appliances |
| **Use Case** | Microservices, dynamic web applications | Gaming, VoIP, real-time data processing | Security appliance integration, network monitoring |
| **Health Checks** | Application-level health checks | Transport-level health checks | Monitors appliance health, not application health |
| **Sticky Sessions** | Supports sticky sessions | Does not support sticky sessions | Not applicable |