perplexity

## Key Note: Differences Between Application Load Balancer (ALB), Network Load Balancer (NLB), and Gateway Load Balancer (GLB)

Understanding load balancer types is essential for architecting scalable, resilient, and secure cloud solutions. Here's a concise summary and technical analysis of the three main cloud load balancer types:

### 1. Application Load Balancer (ALB)

**Purpose:**

- Handles HTTP/HTTPS traffic at the application layer (Layer 7).
- Provides sophisticated routing decisions based on content, such as path, host, HTTP headers, query strings, and source IP.

**Key Features:**

- Path-based and host-based routing.
- SSL termination for secure traffic management.
- Supports WebSocket, HTTP/2, and gRPC.
- Integrates with AWS Lambda for serverless architectures.
- Active health checks for backend services.
- Session persistence (stickiness).

**Best Use Cases:**

- Modern web applications, microservices architectures, and APIs requiring complex, content-based routing.

### 2. Network Load Balancer (NLB)

**Purpose:**

- Fast, high-performance traffic distribution at the transport layer (Layer 4, TCP/UDP/TLS).
- Designed for extreme performance and low latency, suitable for millions of requests per second.

**Key Features:**

- Preserves client source IP, allowing backend services to see client IP.

- Supports static and Elastic IPs.
- Flow-based forwarding: routes traffic based on connection/flow, not HTTP headers.
- No application-level awareness (cannot inspect HTTP parameters).
- Load balancing for any TCP/UDP service.

**Best Use Cases:**

- Gaming, media streaming, or high-frequency trading applications where speed and stability are priorities.
- Applications using protocols other than HTTP/HTTPS.

### 3. Gateway Load Balancer (GLB)

**Purpose:**

- Serves as a transparent network gateway for deploying, scaling, and managing *virtual appliances* (e.g., firewalls, intrusion detection/prevention systems) at OSI Layer 3 (network layer).
- Enables *service chaining*—traffic always passes through these appliances before reaching apps.

**Key Features:**

- Works with any IP-based protocol.
- Maintains flow stickiness and symmetry for appliances requiring stateful inspection.
- Uses route table rules for traffic routing.
- Deploys and manages appliances without requiring application or architecture changes.
- Uses GENEVE/VXLAN protocol for packet encapsulation (AWS/Azure).
- Enables secure connectivity across Virtual Private Clouds (VPCs)/virtual networks.

**Best Use Cases:**

- Security, analytics, packet inspection, or compliance enforcement across hybrid/multi-cloud deployments.
- Use cases requiring traffic mirroring, DDoS protection, or mandatory network appliance traversal.

### Technical Note: Comparative Table

| Feature | Application Load Balancer (ALB) | Network Load Balancer (NLB) | Gateway Load Balancer (GLB) |
|---|---|---|---|
| OSI Layer | Layer 7 (Application) | Layer 4 (Transport) | Layer 3 (Network), Layer 7 (Hybrid) |

| Feature | Application Load Balancer (ALB) | Network Load Balancer (NLB) | Gateway Load Balancer (GLB) |
|---|---|---|---|
| Protocols Supported | HTTP, HTTPS, gRPC | TCP, UDP, TLS | Any IP-based; GENEVE/VXLAN tunneling |
| Routing Logic | Content-based: path, host, headers | Flow hash, connection/flow-based | Routing table, flow stickiness/symmetry |
| Target Types | IP, Instance, Lambda | IP, Instance, ALB | IP, Instance (virtual appliance) |
| Source IP Preservation | No | Yes | Yes |
| SSL Termination | Yes | Yes | No |
| Session Persistence | Yes | Yes (less flexible) | No |
| Performance/Scalability | High, Web-scale | Extreme, millions of requests/sec | High, transparent scaling |
| Use Cases | Web apps, APIs, microservices | Gaming, streaming, non-HTTP/TCP | Security appliances, VPC traffic |
| Appliance Management | No | No | Yes (scale, manage, insert appliances) |
| Health Checking | Application-level | Network-level | Appliance-level |

## Additional Insights

- **Proxy Behavior:** ALB and NLB terminate client connections and proxy traffic, while GLB typically forwards traffic to backend appliances without fully terminating the flow.[1] [2] [3]
- **Static IP Assignment:** NLB supports static/Elastic IPs, valuable when predictable endpoint addresses are required.[4]
- **Advanced Security/Compliance:** GLB enables mandatory inspection and chaining of traffic through security functions (firewalls, IDS/IPS).[2] [3] [1]

## Summary

- **Choose ALB** for advanced HTTP routing and web/application scalability.
- **Choose NLB** for speed, simplicity, and TCP/UDP saturation.
- **Choose GLB** for integrating third-party network appliances or implementing mandatory security, audit, or compliance controls with full IP transparency.

This technical distinction ensures you select the ideal load balancer for any cloud architecture—balancing performance, flexibility, and operational requirements.[3] [1] [2]

⁂

1. https://aws.amazon.com/compare/the-difference-between-the-difference-between-application-network-and-gateway-load-balancing/

2. https://docs.aws.amazon.com/elasticloadbalancing/latest/gateway/introduction.html
3. https://learn.microsoft.com/en-us/azure/load-balancer/gateway-overview
4. https://www.site24×7.com/learn/clb-vs-alb-vs-nlb.html