

Elastic Load Balancing

Inhalt

- Wieso braucht man einen Load Balancer?
- Was ist ein Elastic Load Balancer?
- Welche Typen gibt es?
- Anwendungsfälle

Wieso braucht man einen Load Balancer?

- Für horizontale Skalierung (horizontal scaling) / Hochverfügbarkeit (High Availability)
- Verteilt den Datenverkehr auf mehrere Instanzen
- Erhöht die Verfügbarkeit
- Erhöht die Skalierbarkeit
- Alle EC2 Instanzen sind mit einer URL / IP erreichbar (Traffic wird gleichmäßig verteilt)

Was ist ein Elastic Load Balancer?

- Automatische Verteilung des eingehenden Anwendungsverkehrs auf mehrere Ziele wie EC2 Instanzen, Container, IP-Adressen und Lambda Funktionen
- Elastic bedeutet, dass der Service im Hintergrund auch skaliert wird und hochverfügbar ist (High Availability)

Funktionen

- Hohe Verfügbarkeit (High Availability, HA)
- Zustandsprüfungen (Health Checks)
- Sicherheitsfunktionen
- Beendigung des Transport Layer Security (TLS)
- Layer 4 oder Layer 7 Lastenverteilung
- Betriebsüberwachung

Welche Typen gibt es?

Es gibt in AWS 3 verschiedene Typen von Load Balancern:

- Classic Load Balancer (CLB)
- Application Load Balancer (ALB)
- Network Load Balancer (NLB)

Classic Load Balancer (CLB)

- Vorherige Generation
- HTTP, HTTPS und TCP
- Nur URL / keine statische IP
- Sollte nicht für neue Projekte im Produktivbetrieb verwendet werden

Application Load Balancer (ALB)

- Optimiert für HTTP und HTTPS
- Flexible Anwendungsverwaltung
- Bietet erweiterte Lastenverteilung für HTTPS Datenverkehr
- Wird auf Anfrage-Ebene ausgeführt (Layer 7)
- Pfadbasiertes / Hostbasiertes Routing
- Unterstützung IPv6
- Dynamische Ports
- Verbesserte Metriken und Zugriffsprotokolle
- Gezielte Zustandsprüfung

Network Load Balancer (NLB)

- Optimiert für TCP-Anwendungen mit hoher Geschwindigkeit
- Extreme Leistung
- Statische IP Adresse pro AZ
- Datenverteilung für TCP Datenverkehr
- Wird auf Verbindungs-Ebene ausgeführt (Layer 4)
- Plötzlich auftretende und unbeständige Lasten
- Geeignet für Anwendung die extreme Leistung fordern

Anwendungsfälle für ELB

- **Sicherheit:** Zugriff über einen einzigen Punkt
- **Entkoppeln:** Entkoppeln der Anwendungsumgebung
- **Fehlertoleranz:** High Availability & Durability

Fragen

- Welcher ELB sollte für eine Website verwendet werden?
- Statische IP wird gebraucht, welcher ELB sollte verwendet werden?

Zusammenfassung

- ELB ist ein Service zur Lastenverteilung
- Load Balancer verteilen automatisch die eingehende Datenverkehrslast
- Es gibt 3 Arten:
 - **Application Load Balancer (ALB):**
 - HTTP / HTTPS
 - Fortschrittliches Routing
 - Applikationen, Microservices, Container, Lambda
 - **Network Load Balancer (NLB):**
 - TCP
 - Höchstleistung
 - **Classic Load Balancer (CLB):**
 - Grundlegende Lastenverteilung
 - Mehrere EC2 Instanzen