Wie arbeitet ein Load Balancer? – Verteilung von Netzwerkverkehr für hohe Verfügbarkeit

Inhaltsverzeichnis

- 1 Funktionen und Vorteile eines Load Balancers
- 2 Wie funktioniert ein Load Balancer?
- 3 Typen von Load Balancern
- 4 Load-Balancing-Methoden
- 5 Wann sollte ein Load Balancer eingesetzt werden?
- 6 Alternativen zu Load Balancern
- 7 Fazit

Ein Load Balancer verteilt Netzwerkverkehr auf mehrere Server, um Performance, Skalierbarkeit und Verfügbarkeit zu gewährleisten. Er verbessert die Lastverteilung, schützt vor Überlastung und ermöglicht einen effizienten Betrieb von Webanwendungen und IT-Systemen.

Ein **Load Balancer** (Lastverteiler) ist eine **Netzwerkkomponente oder Software**, die eingehende Anfragen auf mehrere Server verteilt. Dies verbessert die **Leistung, Skalierbarkeit und Verfügbarkeit** von Webanwendungen und IT-Infrastrukturen.

1 Funktionen und Vorteile eines Load Balancers

- **Verteilung des Traffics auf mehrere Server** – Reduziert Serverlast und verbessert die Performance.
- **Erhöhung der Ausfallsicherheit** – Wenn ein Server ausfällt, übernimmt ein anderer.
- **Optimierte Ressourcennutzung** – Stellt sicher, dass keine Server überlastet sind.
- **SSL/TLS-Terminierung** – Entlastet Backend-Server von Verschlüsselungsaufgaben.
- **Schutz vor DDoS-Angriffen** – Verhindert einzelne Serverüberlastungen durch gezieltes Traffic-Management.

2 Wie funktioniert ein Load Balancer?

1. **Ein Client sendet eine Anfrage** (z. B. Zugriff auf eine Website oder API).
2. **Der Load Balancer empfängt die Anfrage** und entscheidet, an welchen Server sie weitergeleitet wird.
3. **Die Lastverteilung erfolgt nach einem Algorithmus**, um eine gleichmäßige Verteilung zu gewährleisten.
4. **Der ausgewählte Server verarbeitet die Anfrage** und sendet die Antwort über den Load Balancer an den Client zurück.

3 Typen von Load Balancern

Typ	Beschreibung	Beispiel
Hardware Load Balancer	Physische Geräte mit spezialisierter Hardware für hohe Performance	F5 Big-IP, Citrix ADC
Software Load Balancer	Läuft als Software auf einem Server oder in der Cloud	HAProxy, Nginx, Traefik
Cloud Load Balancer	Load Balancing als Dienst eines Cloud-Anbieters	AWS Elastic Load Balancer, Google Cloud Load Balancer

4 Load-Balancing-Methoden

Methode	Funktionsweise
Round Robin	Anfragen werden reihum an die Server verteilt
Least Connections	Der Server mit den wenigsten aktiven Verbindungen erhält die nächste Anfrage
IP Hashing	Clients werden basierend auf ihrer IP immer demselben Server zugewiesen
Weighted Load Balancing	Server mit höherer Kapazität erhalten mehr Anfragen
Dynamic Load Balancing	Echtzeit-Monitoring passt die Verteilung basierend auf Serverauslastung an

5 Wann sollte ein Load Balancer eingesetzt werden?

- **Bei hohem Traffic** – Verhindert Überlastung einzelner Server.
- **Für hochverfügbare Webanwendungen** – Sorgt für Ausfallsicherheit.
- **In Microservices-Architekturen** – Verteilt Anfragen auf verschiedene Dienste.
- **Für verteilte Cloud-Umgebungen** – Verbessert Skalierbarkeit und Performance.

6 Alternativen zu Load Balancern

- **DNS Load Balancing** – Verteilt Traffic bereits auf DNS-Ebene.
- **Reverse Proxy (z. B. Nginx, Apache)** – Leitet Anfragen weiter und übernimmt Caching.
- **Content Delivery Networks (CDN)** – Reduziert Last durch verteilte Bereitstellung.

7 Fazit

Ein Load Balancer ist ein **zentrales Element moderner IT-Infrastrukturen**, das **Skalierbarkeit, Ausfallsicherheit und optimale Ressourcennutzung** sicherstellt. Durch verschiedene Methoden kann der Netzwerkverkehr effizient gesteuert werden, um eine hohe Verfügbarkeit und Performance zu gewährleisten.
