
TECH STARTER
Load Balancer

Wichtige Instanztyp Kategorien

- General Purpose (inklusive t2 und t3!)
- Compute Optimized
- Memory Optimized
- Storage Optimized

Instanz-Kaufoptionen

- On-Demand: Pro Sekunde bezahlen, kurze Workloads (teuer!)
- Saving Plans: Auf konstante Nutzung festlegen (1 oder 3 Jahre)
- Reserved Instances: Auf Instanz festlegen (1 oder 3 Jahre)
- Spot Instances: Ungenutzte EC2-Instanzen nutzen (günstigste!)
- Dedicated Hosts: Einen tatsächlichen Server mieten
- Dedicated Instances: Instanz auf reservierter Hardware
- Kapazitätsreservierung: Kapazität in AZ reservieren

Heutige Inhalte

- **Scalability + High Availability**
- **Elastic Load Balancer**
  (Gruppenarbeit)

Je nach Tempo:

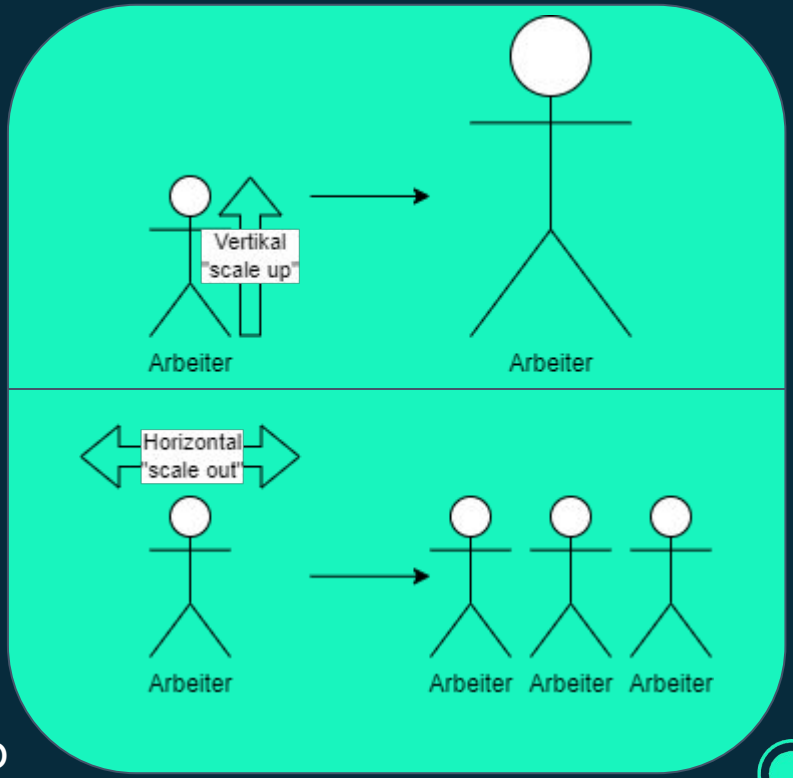
- *(Auto-Scaling-Groups)*

TECH STARTER

Skalierbarkeit

Wir unterscheiden zwischen zwei Arten des Skalierens:

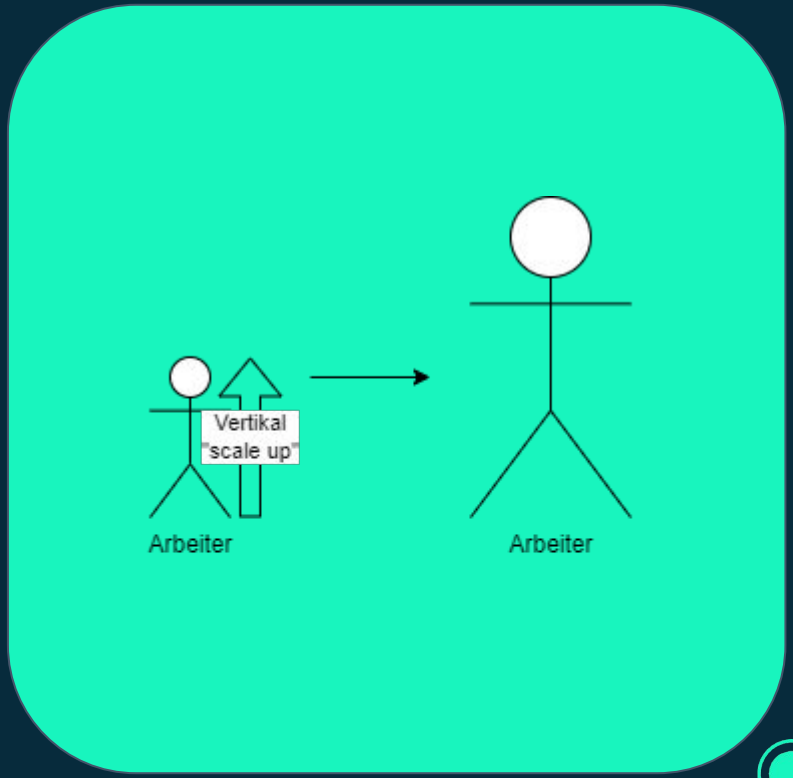
- Vertikal (“scale up”)
  - Meinen Arbeiter stärker machen
  - Aus t3.nano, mach t3.large
- Horizontal (“scale out”)
  - Mehr Arbeiter hinzufügen
  - Aus 1x t3.nano, mach 4x t3.nano



TECH STARTER

Vertikal Skalieren

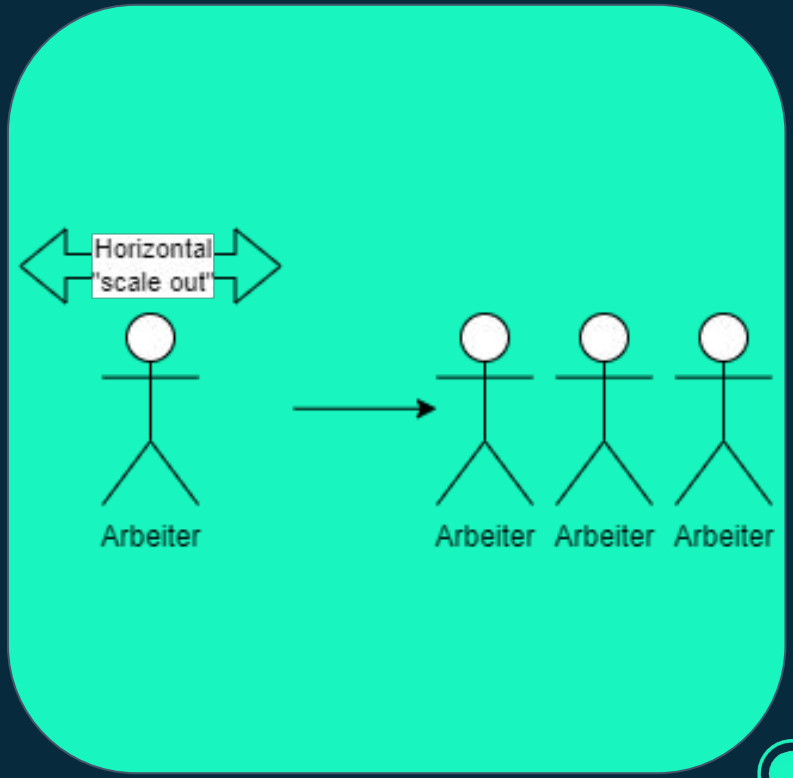
- Kommt häufig bei Relationalen Datenbanken vor, die sich oft schlecht auf mehrere Systeme verteilen lassen
- Es gibt ein Limit, bis zu dem man Vertikal Skalieren kann (Stärkste aktuelle Hardware)



TECH STARTER

Horizontal Skalieren

- Horizontale Skalierung benötigt ein verteiltes System
- Bildet die übliche Variante für moderne Web-Anwendungen
- Horizontal Skalieren ist flexibler und einfacher als Vertikal



TECH STARTER

Single Point of Failure (SPOF)

- Fällt ein einziger Bestandteil unserer Infrastruktur aus, ist unser Service nicht mehr nutzbar

- Beispiele: Eine Instanz; Ein Server; Ein Router; Ein Standort; …

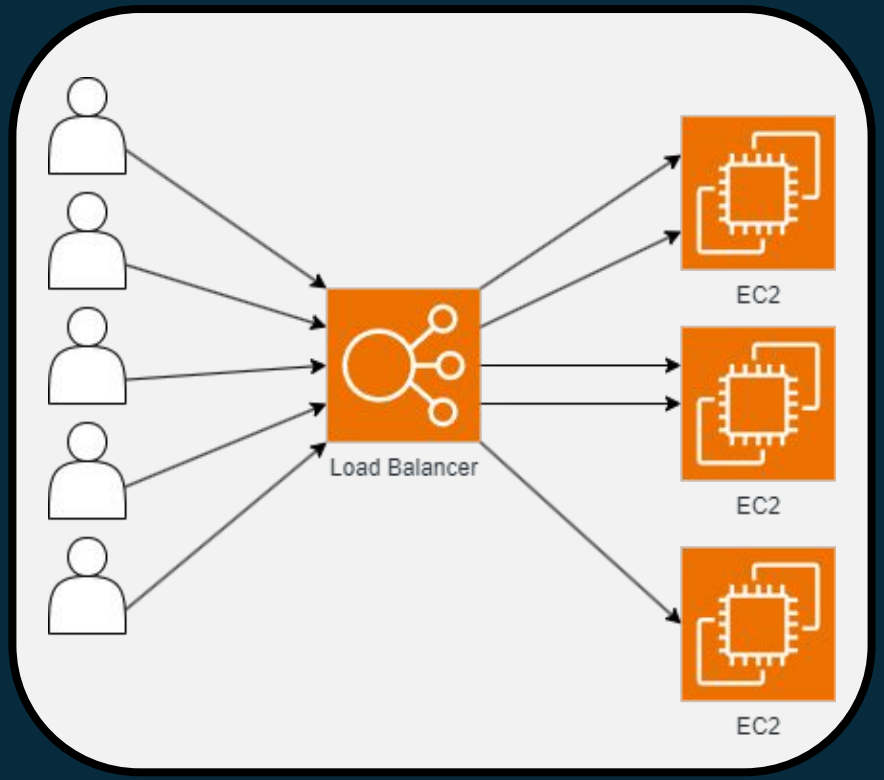
TECH STARTER

High Availability / Hohe Verfügbarkeit

- Ziel: Single Point of Failure entgegenwirken
- Wie?: Unser System / unseren Service in wenigstens 2 Availability Zones gleichzeitig laufen lassen (“multi AZ”)
- Dazu bräuchten wir mindestens 2 Instanzen

TECH STARTER

Load Balancer



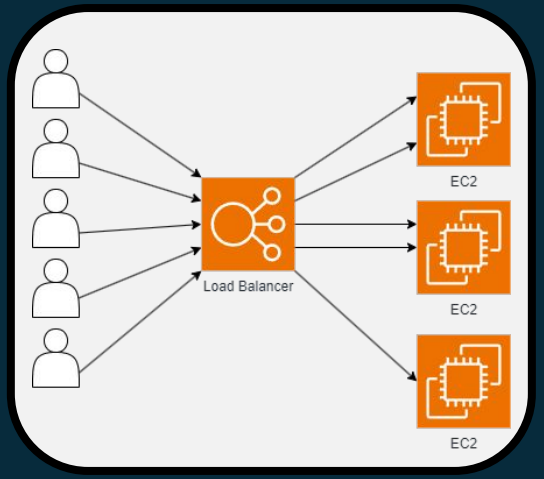
TECH STARTER

Load Balancer

- Load Balancer sind separate Server, welche jeglichen Datenverkehr an andere Server (unsere Instanzen) weiterleitet

- Vorteile:
  - mehrere Instanzen nebeneinander schaltbar (Lastverteilung)
  - Instanzen über mehrere AZs verteilbar
  - User behalten eine einzige Adresse um unseren Service zu erreichen

Load Balancer



Was passiert, wenn eine Instanz ausfällt?

- Load Balancer bemerkt dies (*Health Checks*)
- Verkehr wird auf alle “*healthy*” Instanzen geleitet

Was passiert, wenn der Load Balancer ausfällt?

- Service wäre nicht (normal) zu erreichen. **SPOF**!

TECH STARTER

Load Balancer - Selber machen?

Wir *könnten* selbst einen Server / eine Instanz erstellen, die eingehenden Verkehr an andere Instanzen weiterleitet.

- Spart Geld
- Sorgt für mehr Arbeit und Verantwortung

- AWS bietet ***E****lastic* ***L****oad* ***B****alancer* als *verwalteten Load Balancer*, der uns diese Arbeit und Verantwortung abnimmt

Elastic Load Balancer

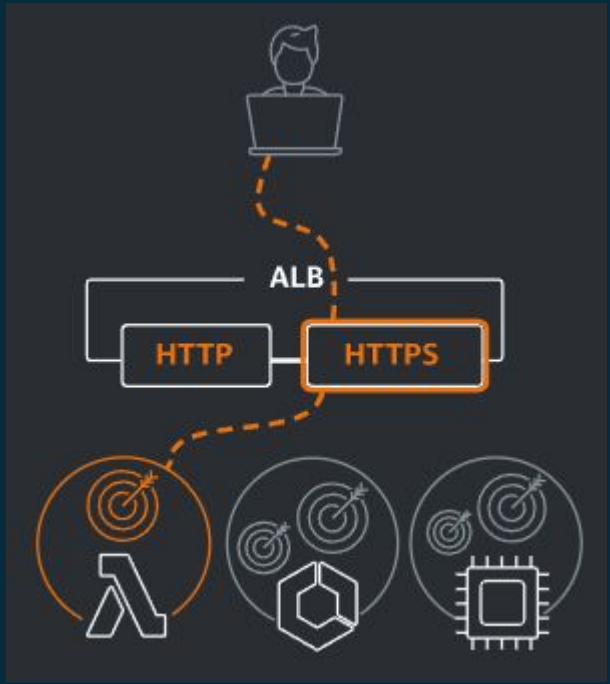
- AWS bietet **ELB** als einen *Verwalteten Load Balancer*
- AWS kümmert sich um:
  - hohe Ausfallsicherheit, Upgrades / Updates, Wartung,...
- Uns bleiben nur wenige Einstellung (und wenig Verantwortungen)

TECH STARTER

AWS Load Balancer Typen

Application Load Balancer

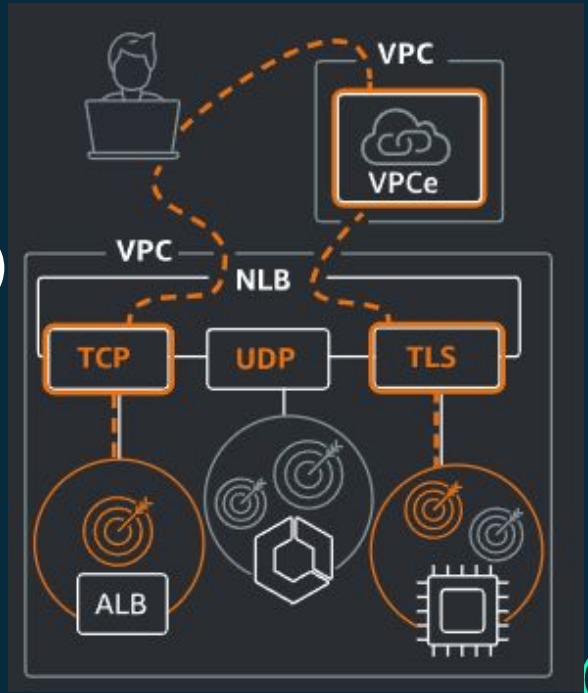
- Für typische Webanwendungen
- Nutzt HTTP / HTTPS
- Liefert Usern einen *Single Point of Entry* über statische DNS
- (diesen nutzen wir am ehesten)



TECH STARTER

Network Load Balancer

- Für extrem hohe Leistungen
  (Millionen von Anfragen pro Sekunde)
- Nutzt TCP/UDP (und TLS)
- Extrem niedrige Latenzen



TECH STARTER

Gateway Load Balancer

- Nutzt IP-Pakete (GENEVE Protocol)

- Typischerweise für Firewalls, Intrusion Detection, Prävention
- Sicherheitsbedrohungen in Datenpaketen identifizieren

TECH STARTER

Classic Load Balancer

- Die ursprüngliche Variante
- Wurde durch die Spezialisierten Load Balancer Varianten abgelöst
- Sollte ab 2023 nicht mehr genutzt werden

Arten von AWS Load Balancer

- **Application Load Balancer** ← Diesen wollen wir heute nutzen!
  (für HTTP/HTTPS)
- Network Load Balancer
  (TCP/UDP, Millionen Anfragen pro Sekunden)
- Gateway Load Balancer
  (IP, Security)
- Classic Load Balancer
  (Nicht länger unterstützt)

TECH STARTER