

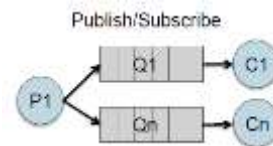
# Mögliche Klausurfragen

## 0 Einfuehrung

- Was versteht man im Cloud Computing unter Ressourcen Pools?
  - Verfügbarkeit von **scheinbar unbegrenzten Ressourcen**, die Anfragen verteilt verarbeiten.
- Was sind zwei mögliche Gründe für Unternehmen, einen Multi-Cloud-Ansatz zu nutzen? Nenne mindestens zwei Gründe.
  - *Ausfallsicherheit durch Redundanz (unterschiedliche Standorte)*
  - *Lokalität (Anfragen aus Amerika an Cloud in NewYork, Anfragen aus Deutschland an Cloud in Berlin)*
  - *Geringere Kosten*

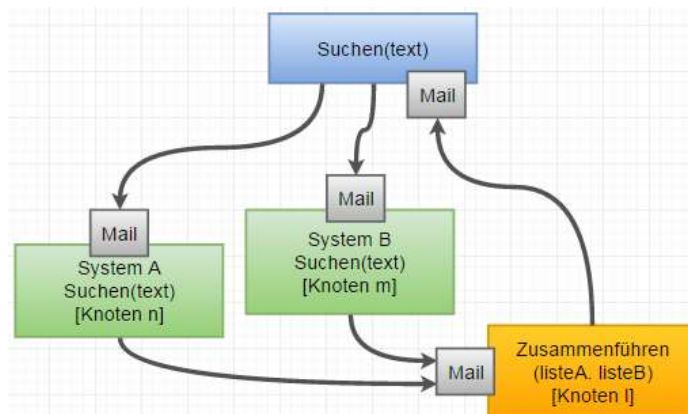
## 1 Kommunikation

- Was ist Push-Kommunikation?
  - *Ist eine Benachrichtigung in eine Richtung bei der der Server mit der Kommunikation beginnt*
- Was versteht man bei der Internet-Kommunikation unter einem Service?
  - *Ein Service (definiert durch eine Serviceschnittstelle) ist eine Funktionalität, die über eine definierte Schnittstelle zur Verfügung stellt*
  - *Serviceschnittstelle ist ein Vertrag zwischen Nutzer und Anbieter über Syntax und Semantik der Service-Nutzung*
- Stellen sie das Messaging-Kommunikationsmuster Publish/Subscribe grafisch dar und beschreiben sie einen typischen Anwendungsfall für dieses Muster.
  - *Anwendungsfall: Wenn Daten Unterschiedlich verarbeitet werden  
Beispielsweise Erfolgt die verarbeitung von Bildern in Service A  
Und die verarbeitung von Text in Service B*



## 2 Programmiermodelle

- Was sind die vier wesentlichen Forderungen des ReactiveManifesto?
  - *Systeme müssen stets **antwortbereit, widerstandsfähig, elastisch und nachrichtenorientiert** sein*
- Welche Vorteile bietet FRP gegenüber synchronen Systemen?
  - *Höhere Prozessorauslastung: Der Prozessor befindet sich deutlich weniger Zeit in Wartezuständen (IO-Wait, Lock-Wait, ...)*
  - *Höherer Parallelisierungsgrad möglich und damit höherer möglicher Speedup*
- Stellen sie ein Aktorensystem grafisch dar, das folgende Funktion erfüllt
  - Ein Aufrufer übergibt eine Zeichenkette nach der er suchen will
  - Das System stößt parallel zueinander zwei Service-Aufrufe an. Einen an System A den anderen an System B. Die Systeme können zu einer Zeichenkette eine Liste an Ergebnissen liefern.
  - Die Liste der Ergebnisse wird zusammengefasst und an den Aufrufer übergeben

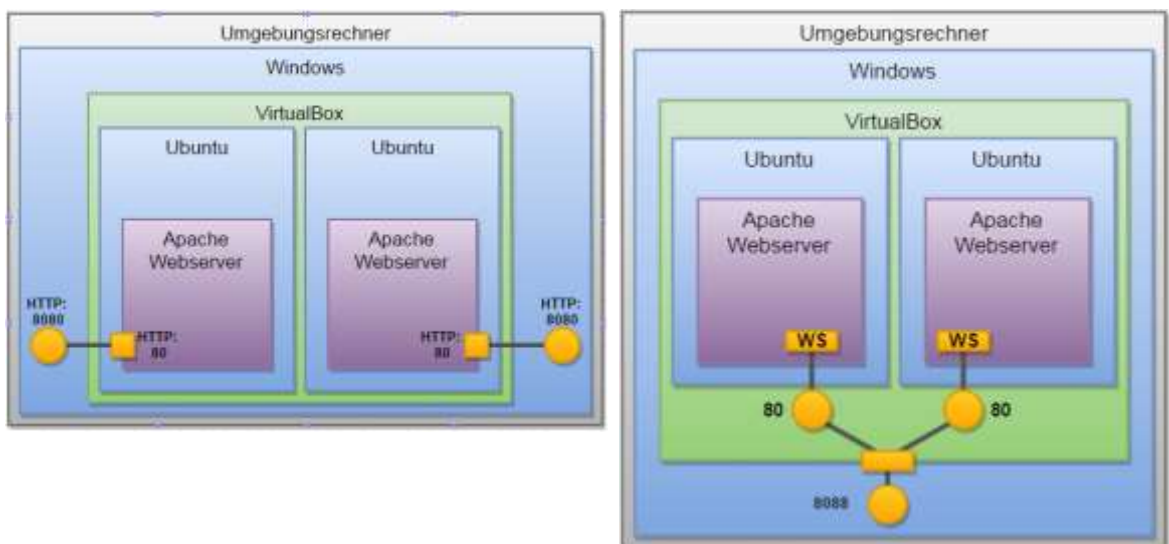


### 3 Virtualisierung

- Was ist der allgemeine Zweck von Virtualisierung? Wie kann dieser Zweck in Bezug auf die Hardware Virtualisierung interpretiert werden?
  - Zweck:
    - Multiplizität *Erzeugung mehrerer virtueller Realitäten innerhalb einer physikalischen Realität*
    - Entkopplung *Bindung und Abhängigkeit zur Realität auflösen*
    - Isolation *Physikalische Seiteneffekte zwischen den virtuellen Realitäten vermeiden*
  - Bezug auf die Hardware Virtualisierung
    - Multiplizität *Viele Virtuelle Maschinen auf einem Rechner*
    - Entkopplung *Die Software ist unabhängig von der darunterliegenden HW*
    - Isolation *VMs laufen isoliert vom Host und sind Gekapselt und gegenseitig Unabhängig*
- Welche vier Ressourcen-Arten werden mindestens virtualisiert?
  - Prozessor *Virtuelle Rechenkerne*
  - Hauptspeicher *Virtuelle Hauptspeicher-Partition*
  - Netzwerk *Virtuelle Netzwerkschnittstellen und virtuelle Netzwerk-Infrastrukturen (VLAN))*
  - Storage *Virtuelle Festplatten-Laufwerke*
- Was ist bei Docker der Unterschied zwischen Images und Containern?
  - Image: *Ruhender und transportierbarer Zustand*
  - Laufender Zustand (Laufendes Image): *Ein Container läuft so lange wie sein Entrypoint-Prozess im Vordergrund läuft. Docker merkt sich den Container-Zustand.*
- Was bewirkt der Docker Befehl „run“?
  - *Das Container auf Basis eines Images Starten*

### 4 Provisionierung

- Was versteht man unter dem Begriff Provisionierung?
  - *Überführen eines Systems vom ist Zustand zum soll Zustand*
  - *(Bezeichnung für die automatisierte Bereitstellung von IT-Ressourcen)*
- Welche Zusicherungen muss ein Provisionierungsmechanismus machen? Erläutern sie diese Zusicherungen. Welchen Zweck verfolgen diese Zusicherungen aus ihrer Sicht?
  - Idempotenz: *Das selbe Ergebnis wird erzeugt egal wie oft die Aktion durchgeführt wird*
  - Konsistenz: *Nach Ausführung der Aktionen herrscht ein konsistenter Systemzustand. Egal ob Aktionen gescheitert sind. Zustand ist danach entweder der Ausgangszustand oder der Zielzustand aber nichts dazwischen*
  - Zweck:
    - *Ausfallsicherheit, Fehler im Betrieb vermeiden*
- Auf ihrem Windows-PC laufen zwei virtuelle Maschinen auf Basis VirtualBox. In diesen Maschinen befindet sich ein Ubuntu Linux Betriebssystem mit installiertem Apache Webserver. Dieser stellt ein Website über den Port 80 zur Verfügung. Dieser Port ist im Host-Betriebssystem unter dem Port 8088 erreichbar. Stellen sie diese Umgebung grafisch dar.



## 5 Iaas

- Was versteht man unter Opportunitätskosten in Bezug auf die Betriebskosten von Hardware?
  - *Betriebskosten von ungenutzten HW Ressourcen*
- Gegeben sei eine IaaS-Architektur entsprechend dem Beispiel von Eucalyptus: Was passiert, wenn der Cluster Controller vollständig ausfällt? Was ist weiterhin in der IaaS Cloud noch möglich?
  - *Ein Cluster ist beim Ausfall des Cluster Controllers von außen nicht mehr nutzbar.*
  - *Andere Cluster sind davon unbetroffen, solange der Cloud Controller fehlerfrei läuft*
  - *Das betroffene Cluster wird vom Cluster Controller aus dem Betrieb genommen*
  - *Rechner im betroffenen Cluster können weiterlaufen, sind von außen jedoch nicht erreichbar => bestehende Jobs können weiterlaufen*

## 6 Cluster Scheduling

- Was ist der Unterschied zwischen statischer und dynamischer Partitionierung im Scheduling?
  - *Bei statischen Scheduling werden Ressourcen fest an einen Prozess zugewiesen (keine dynamische Änderung)*
  - *Beim dynamischen Scheduling weist der Scheduler die Ressourcen den einzelnen Prozessen dynamisch zur Laufzeit je nach Anforderung zu*
- Erläutern Sie die beiden mindestens zu unterscheidenden Job-Arten im Cluster-Scheduling.
  - *Batch-Jobs: Ausführungszeit im Minuten- bis Stundenbereich. Eher niedrige Priorität und gut unterbrechbar. Müssen i.d.R. bis zu einem bestimmten Zeitpunkt abgeschlossen sein. Zustandsbehaftet.*
  - *Service-Jobs: Sollen auf unbestimmte Zeit unterbrechungsfrei laufen. Haben hohe Priorität und sollten nicht unterbrochen werden. Sind teilweise auch zustandslos.*
- Was versteht man unter einer dominanten Ressource beim DRF-Scheduling-Algorithmus?
  - *Die Ressource, die pro Task/Team am meisten benötigt wird (die Ressource mit dem höchsten Bedarf)*

## 7 Orchestrierung

- In welchen typischen Fällen führt ein Orchestrator ein Re-Scheduling von Containern durch?
  - *Im Fehlerfall oder zur Performance Optimierung*
- Erläutern Sie das Konzept eines ReplicaSets in Kubernetes.
  - *ReplicaSets / Replication Controller: stellen sicher, dass eine spezifizierte Anzahl an Instanzen pro Pod ständig läuft*
- Was versteht man unter Auto-Skalierung?
  - *Das **vollkommen Automatisch** auf die Last Anforderungen dynamisch reagiert wird*

## 8 Cloud Architektur

- Welche typischen Aufgaben übernimmt ein Edge Server in einer Cloud Architektur? Nennen Sie mindestens 3 Aufgaben.
  - *Typische Aufgaben:*
    - *Authentifizierung*                      *Load Balancing*                      *Request Validierung*
    - *Caching*                                      *Logging*
- Was besagt das CAP-Theorem?
  - *Es gibt drei wesentliche Eigenschaften (Konsistenz, Verfügbarkeit, PartitionsToleranz), von denen ein verteiltes System nur zwei gleichzeitig haben kann*
- Nennen Sie die Protokollart für einen verteilten Konfigurationsspeicher, die sowohl konsistent als auch partitions-tolerant ist.
  - *Konsens                      Konsens                      und nochmal Konsens, aber momentan da war was, ja stimmt Konsens*

## 9 Big Data

- Ab welcher Datenmenge spricht man in der Regel von Big Data?
  - *Ab 100Gb bzw. 1Tb*
- Welche Vorteile bietet die Parallelität im Großen im Vergleich zur Parallelität im Kleinen?
  - *Ausfallsicherheit, Mehr Leistung, Hoher Durchsatz, Horizontale Skalierbarkeit, Keine hardwarebedingte Limitation*
- Was ist der Unterschied zwischen einer Action- und einer Transformation-Methode bei einem RDD?
  - *Transformations:*                      *Aus RDD Daten werden RDD Daten*
    - *Verändert nur die Sicht auf die Daten*
  - *Actions:*                                      *Aus RDD Daten werden andere Daten (skalärer Typ, Collection, Storage)*
    - *Eine Action macht was Aktives mit den Daten*

## 10 Paas

- Wie wird eine Anwendung auf einer PaaS deployed?
  - *Indem Sourcecode zur Verfügung gestellt wird, dann als eine ausführbare Einheit in die PaaS transferiert wird.*
  - *Oder in form einer App-Paker*
- Wie sind die Begriffe Deployment und Provisionierung in Bezug auf die Building Blocks einer PaaS Cloud zu unterscheiden?
  - *Deployment: Aus Quellcode wird eine Applikation erstellt*
  - *Provisionierung: Kümmert sich um Bereitstellung der Umgebung*