

CC: DATENSCHUTZ

Dipl.-Medieninf. Hai Dang Le
Software Engineer
hhdang.88@gmail.com

SS 2017



AGENDA

4. DATENSCHUTZ & CLOUD VERTRAG

- a. Studentenvorträge
- b. Autoscaling & Microservices (Abschluss VL2)
- c. Datenschutz

4. AUTO-SCALING (AUS VL2)

A. AUTO-SCALING & MICROSERVICES

AUTO-SCALING

- Essentieller Mechanismus im Cloud Computing um auf den Bedarf von Ressourcen zu reagieren
- dynamische Bereitstellung von VMs, Container, Speicher, Storage
- bidirektional: Scale-Up, Scale-Down
- Nutzen:
 - Elastizität: bessere Auslastung, Kostenoptimierung (Kosteneinsparung)
 - Verfügbarkeit und Zuverlässigkeit: durch Sicherstellung von min. Instanzen

AUTOSCALING STRATEGIEN

- Reaktiv: Cloud Umgebung beobachtet Kernmetriken (CPU, Memory, Traffic) und reagiert auf Last
- Proaktiv:
 - Scheduling (scheduled scaling): Zeitlich geplante Skalierung auf Basis von Erfahrungswerten
 - Prädiktiv (predictive scaling): durch Vorhersage auf Basis von prädiktiver Analyse

BEISPIEL: AUTO-SCALING BEI AWS (IAAS)

- CPU, Speicherauslastung, Netzwerk Traffic wird durch AWS EC2 überwacht
- alle Logs/Stdout werden an AWS Cloudwatch übermittelt (Logging/Monitoring-Service)

<http://docs.aws.amazon.com/autoscaling/latest/userguide/WhatIsAutoScaling.html>

- AMIs/VMs können zu Auto-Scaling Groups hinzugefügt werden, Konfiguration von Scaling-Plans:
 - Scaling Policies: reaktiv (standard), scheduled-scaling, dynamisch (auf Basis von selbstdefinierten Events/Metriken aus Cloudwatch)
 - min./max. und gewünschte Anzahl an Instanzen

BEISPIEL: AUTO-SCALING BEI AWS (IAAS)

- Ausführung: auf Basis von Scaling Plans werden AMI Instanzen hinzugefügt / terminiert
- hinzufügen von Instanzen (auf Basis von Launch-Configuration):
 - Instanziierung aus AMI
 - boot-up Phase, Initialisierung, Konfigurierung per Skript (siehe AMI Design Strategien)
 - Anbindung an Cloud-Storage
 - Registrierung ins Virtual Private Cloud (Netzwerk)
 - Registrierung beim Elastic Load-Balancer

BEISPIEL: AUTO-SCALING BEI AWS (IAAS)

- Status Überwachung: Health-Checking
 - über HTTP-Endpunkt: z.B. /heathcheck
 - healthy flag

BEISPIEL: AUTO-SCALING BEI CLOUD-FOUNDRY (PAAS)

- Monitoring / Überwachung von Basis-Metriken:
analog zu AWS EC2
- Skalierung auf Container / App-Ebene
- Skalierung über Einbinden eines "Auto-Scaler"-
Services an einer App
- Definition von Basis-Metriken:
 - min./max. Instanzen
 - CPU-, Traffic-, Speicher-Grenzen
- nur reaktiv

BEISPIEL: AUTO-SCALING BEI CLOUD-FOUNDRY (PAAS)

- Ausführung: auf Basis der Auto-Scaler Konfiguration werden Container-Instanzen hinzugefügt / terminiert
- hinzufügen von Instanzen:
 - Container Instanz wird aus dem gebautem Image instanziert
 - Einbinden der Container Instanz ins private virtual network
 - Container Prozess wird gestartet (run-Befehl)
 - Registrierung beim Load-Balancer

BEISPIEL: AUTO-SCALING BEI CLOUD-FOUNDRY (PAAS)

- Status Überwachung: Health-Checking
 - über HTTP-Endpunkt: z.B. /heathcheck, (http-response-Code: 200 bedeutet "OK", ansonsten "NOK")
 - Überwachung des Container Prozesses: terminiert der Container Prozess, stoppt der Container und crasht, es wird ein neuer Container gestartet

BEISPIEL: AUTO-SCALING BEI KUBERNETES (CAAS)

- Monitoring / Überwachung von Basis-Metriken: analog zu Cloud Foundry
- Skalierung auf Pod-Ebene (Kubernetes Konzept: Zusammenfassung von Containern)
- Skalierung über die Definition einer Auto-Scaling Konfiguration
- Konfiguration wird über Selektoren auf Kubernetes Bausteine (Pod, ReplicaSets, Deployments) angewendet
- Auto-Scaling ist reaktiv und basiert nur auf CPU-Last-Grenzen

BEISPIEL: AUTO-SCALING BEI KUBERNETES (CAAS)

- Ausführung: auf Basis der Auto-Scaler Konfiguration werden Container-Instanzen hinzugefügt / terminiert
- hinzufügen von Instanzen:
 - Pod Instanz wird erstellt, alle Container Definitionen werden aus dem konfiguriertem Image instanziert
 - Einbinden der Pod/Container Instanz ins private virtual network
 - Container Prozess wird gestartet (run-Befehl)
 - Registrierung beim Service und Load-Balancer

BEISPIEL: AUTO-SCALING BEI KUBERNETES (CAAS)

- Status Überwachung: Health-Checking
 - über HTTP-Endpunkt: z.B. /heathcheck, (http-response-Code: 200 bedeutet "OK", ansonsten "NOK")
 - Überwachung der Pod/Container Prozesse: terminiert der Pod/Container Prozess, stoppt der Pod/Container, es wird ein neuer Pod/Container gestartet

MICROSERVICES

- modernes / populäres Architekturmuster in der Cloud
- eignen sich für "App-Containerisierung"
- eignen sich für die horizontale Skalierung in der Cloud
- eignen sich für "Continuous Delivery"

PHILOSOPHIE

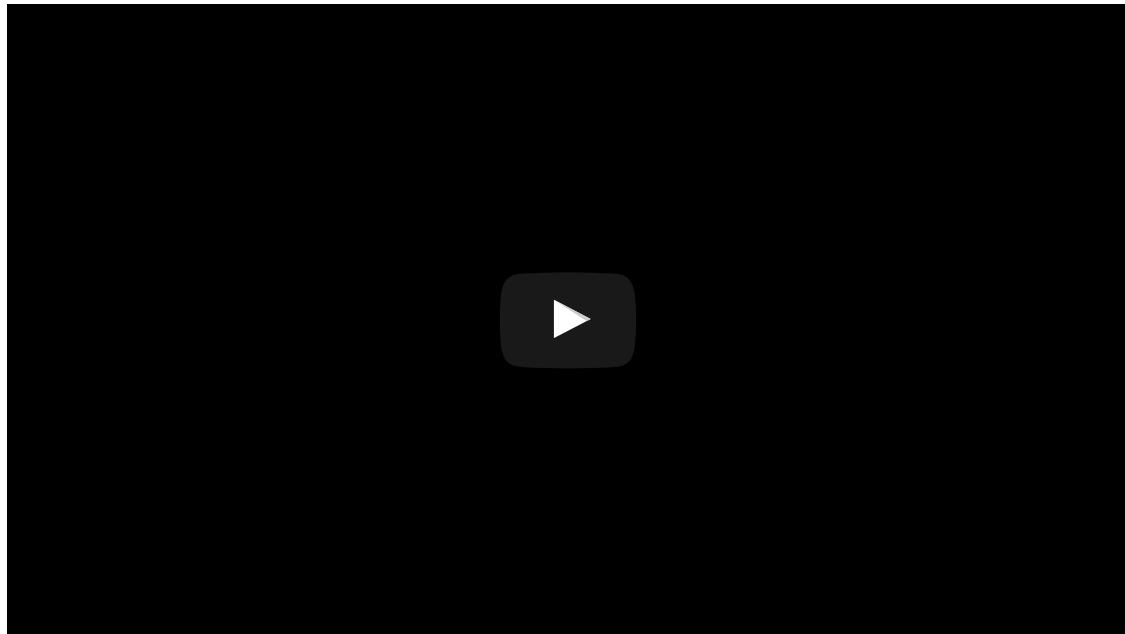
- Entkopplung von Funktionalitäten in kleine eigenständig lebende Funktionen
- Dekomposition von Applikationen in abgegrenzte, isolierte Module
- Überschaubarkeit, Einfachheit von Komponenten
- Offene Schnittstellen, keine proprietäre Protokolle, meistens Http + JSON/XML
- Programmiersprachenunabhängigkeit von Komponenten
- Entkopplung von Entwicklungsteams

VORTEILE

- unabhängige Skalierbarkeit von Microservices
- bessere Wartbarkeit, einfache Neuimplementierung
- versteckte Abhängigkeiten werden über die API erkennbar
- Sprachenunabhängig, Teams können nach ihren Wünschen den Technologiestack auswählen
- unabhängige Entwicklungszyklen, Parallelisierung der Entwicklung
- bei Überlastung (DDos) können sekundäre Services zugunsten von primären Services abgeschaltet werden

NACHTEILE

- Performance Overhead durch Netzwerk-Kommunikation
- Erhöhte Komplexität von Tests: Unit/Komponenten-Tests werden Integrationstests
- Erhöhte Komplexität von Monitoring: zentralisiertes Monitoring, Splunk, AppDynamics etc.
- Deployment Prozess muss ggfs. zwischen Teams abgestimmt werden (API Breaking Changes)
- Einführung von generellen Problemen von verteilten Anwendungen: z.B. Lastverteilung, Zeitsynchronisation, distributed Locking



<https://www.youtube.com/embed/CKL3fV5UR8w>

4. DATENSCHUTZ

B. PERSONENBEZOGENE DATEN & AUFTRAGSDATENVERARBEITUNG

DATENSCHUTZ

Darf ich Daten erheben und bei einem Cloud-Provider speichern der nicht in D (oder EU) ansässig ist?

DATENSCHUTZ

- Recap: Besonderheiten bei Cloud Computing
 - keine technische Beschränkung durch natürliche Grenzen
 - Undurchsichtigkeit der Datenverarbeitung
 - Geographische Verteilung, Grenzüberschreitender Datenverkehr
 - Kosteneinsparung durch Übertragung von Verantwortung

WELCHE DATEN SIND SCHUTZBEDÜRFIG ?

- Beispiele:
 - Steuerlich relevante Daten (§146 Abs. 2 S1 AO)
 - müssen im Inland oder in einem Mitgliedstaat der EU/EWR (durch Einwilligung der zust. Finanzbehörde) gespeichert werden (§146 Abs. 2a AO)

- Handelsdaten z.B. Buchungsbelege, Handelsbriefe (§257 Abs.4 HGB)
 - müssen im Inland für 6 bzw. 10 Jahre aufbewahrt werden
- Personenbezogene Daten (§3 Abs.1 BDSG)
 - Vertraulichkeit und Integrität muss bewahrt sein
- Quelle:

<https://www.datenschutzzentrum.de/cloud-computing/20100617-cloud-computing-und-datenschutz.html>

WAS SIND PERSONENBEZOGENE DATEN ?

- §3 Abs. 1 BDSG:
 - "(1) Personenbezogene Daten sind Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbaren natürlichen Person (Betroffener)"
 - Daten die eindeutig einer Person zugeordnet werden können

WAS SIND PERSONENBEZOGENE DATEN ?

- die Identität einer Person aus dem Inhalt des Datums oder mit Zusatzwissen sich herstellen lässt
 - Quellen:
- https://www.gesetze-im-internet.de/bdsg_1990/__3.html
- https://www.ldi.nrw.de/mainmenu_Datenschutz/Inhalt/FAQ/PersonenbezogeneDaten.php

WAS SIND PERSONENBEZOGENE DATEN ?

- Beispiele:
 - Alter, Augenfarbe, Geburtsort
 - Telefonnummer, E-Mail Adresse
 - KFZ-Kennzeichen

BESONDERS SCHUTZWÜRDIGE PERSONENBEZOGENE DATEN

- "(9) Besondere Arten personenbezogener Daten sind Angaben über die rassische und ethnische Herkunft, politische Meinungen, religiöse oder philosophische Überzeugungen, Gewerkschaftszugehörigkeit, Gesundheit oder Sexualleben"
- Verlust kann erhebliche Konsequenzen für die betroffene Person haben
- es gelten erhöhte Schutzmaßnahmen

ANWENDBARKEIT

- anonymisierte Daten
 - BDSG findet keine Anwendung wenn die Daten so anonymisiert sind dass nur mit unverhältnismäßigem Aufwand die Daten einer Person zugeordnet werden können
- pseudonymisierte Daten
 - schließen die Anwendbarkeit des BDSG nicht aus

ANWENDBARKEIT

- zentrale Bedeutung hat die "verantwortliche Stelle" (Person die Daten für sich selbst erhebt und verarbeitet bzw. für sich verarbeiten lässt)
- befindet sich diese in der EU/EWR oder Deutschland findet das jeweilige Nationale Recht bzw. deutsches Datenschutzrecht Anwendung
- befindet sich die verantwortliche Stelle außerhalb der EU/EWR, aber werden Daten innerhalb der EU (Server) verarbeitet gilt das "Territorialprinzip"

ANWENDBARKEIT

- Sitzlandprinzip (europäisches Recht)
 - Europ. Datenschutzrichtlinie (EU-DSRL) Art. 4 Abs. 1a), b): kommt es für die Anwendbarkeit einzelstaatlichen Rechts darauf an, in welchem Mitgliedstaat die Daten verarbeitende Niederlassung ihren Sitz hat.

- findet Anwendung bei grenzüberschreitenden Datenverkehr, regelt welches Nationale Recht Anwendung findet
- befindet sich eine Niederlassung eines Unternehmens/verantwortliche Stelle in Deutschland (und der Hauptsitz in der EU), in dem die Datenerhebung/-Verarbeitung stattfindet, dann findet das BDSG Anwendung
- Relevant ist das Land in dem sich die Niederlassung der verantwortlichen Stelle befindet, das jeweilige Nationale Recht gilt

ANWENDBARKEIT

- Territorialprinzip
 - das jeweilige Nationale Recht findet Anwendung in der die Datenerhebung, -verarbeitung, -nutzung stattfindet, d.h. in dem Land an dem sich die Server befinden
- Quelle: <https://www.datenschutzbeauftragter-info.de/ist-das-bundesdatenschutzgesetz-bdsg-im-internationalen-datenschutz-anwendbar/>

AUFTAGSDATENVERARBEITUNG

Welche Regelungen gibt es damit Daten zur Verarbeitung in eine Cloud transferiert werden können
?

AUFTAGSDATENVERARBEITUNG (§11 BDSG)

- betrifft das Outsourcing von Datenverarbeitung
- regelt die Verantwortlichkeit im Sinne des Datenschutzes, zwischen dem Auftraggeber und Auftragsnehmer
- Verantwortlich ist der Auftragsgeber der die Datenverarbeitung für sich durchführen lässt
- der Auftragsgeber ist die verantwortliche Stelle, hat Kontrollpflichten ggü. dem Auftragsnehmer
- bei Datenverlust/Leck hat die verantwortliche Stelle eine Meldepflicht

BEISPIEL CLOUD COMPUTING: SAAS

- Rollen: Cloud-Nutzer (Person o. Firma), Cloud Anbieter, Rechenzentrumsbetreiber
- Verantwortliche Stelle ist der Cloud Nutzer, dieser lässt über Auftragsdatenverarbeitung personenbezogene Daten durch den Cloud Anbieter und Rechenzentrumsbetreiber (Subbeauftragung) verarbeiten
- der Cloud Nutzer ist verantwortlich für den Schutz der Daten
- der Cloud Nutzer hat die Pflicht den Cloud Provider sorgfältig auszuwählen

MINDESTANFORDERUNGEN FÜR DIE AUFTAGSDATENVERARBEITUNG

1. Gegenstand und Dauer des Auftrags
2. Umfang, Art und Zweck der Verarbeitung, Art der Daten und Kreis der Betroffenen
3. die Datensicherungsmaßnahmen nach § 9 BDSG
4. Berichtigung, Löschung und Sperrung der Daten
5. die (Kontroll-)Pflichten der Auftragnehmer (AN)
6. Unterauftragsverhältnisse
7. Kontrollrechte der Auftraggeber (AG)
8. Mitteilungspflichten der AN bei Verstößen
9. Weisungsbefugnisse
10. Datenlöschung beim AN. Nach § 11 Abs. 2 S. 4, 5 BDSG muss sich der AG regelmäßig über die Beachtung der Datensicherungsmaßnahmen überzeugen

VERARBEITUNG VON PERSONENBEZOGENEN DATEN AUSSERHALB DER EU

- nach BDSG nur erlaubt wenn die Einwilligung aller Betroffenen eingeholt wurde oder ein angemessenes Datenschutzniveau im Drittland sichergestellt ist (§ 4b Abs. 2, 3 BDSG)
- angemessenes Datenschutzniveau besteht nach Attestierung der Europ. Kommission in Andorra, Argentinien, Australien, Faroer Inseln, Guernsey, Israel, Isle of Man, Jersey, Kanada, Schweiz, Uruquay und Neuseeland
- für die Nutzung in Staaten außerhalb der EU muss ein angemessenes Datenschutzniveau gewährleistet werden, z.B. durch gesonderte Abkommen:
 - Bsp.: Safe-Harbor (nicht ausreichend), Binding Corporate Rules (muss durch die Datenschutzaufsichtsbehörden genehmigt werden)

FAZIT

- Es ist am sichersten dass Daten die innerhalb der EU erhoben/verarbeitet werden, die EU Grenzen nicht verlassen
- als Maßnahmen zur Erfüllung der Kontrollpflicht als verantwortliche Stelle können Zertifizierungen des Cloud Providers herangezogen werden (ISO 27001, Trusted Cloud)
- neben des Standorts, sollte der Cloud Provider sorgfältig nach folgenden Kriterien ausgesucht werden:
 - Verschlüsselungsmöglichkeiten, Anonymisierung von Daten
 - werden Daten in ein Drittland repliziert ?
 - Verbindliche Zusagen in Sachen Transparenz, SLAs, Vertragliche Zusagen ausreichend? (z.B. Mindestanforderungen für ADV)

4. ZUSAMMENFASSUNG

END. ZUSAMMENFASSUNG VL 'CLOUD COMPUTING'

ZUSAMMENFASSUNG

Was haben wir (kennen-) gelernt ?

VORLESUNG 1:

- Begriffsdefinition 'Cloud Computing'
 - Eigenschaften, Liefermodelle, Servicemodelle
 - Verantwortlichkeiten
 - Vorteile / Nachteile Cloud Computing
- Marktübersicht
 - Hemnisse im dt. Markt

VORLESUNG 2 (SOFTWAREENTWICKLUNG):

- Virtualisierung / Virtualisierungsformen
 - Systemvirtualisierung, Betriebssystem-basierte Virtualisierung, Emulation
 - Bedeutung für IaaS, PaaS, CaaS
 - Deployment: "BYOCode", "BYOContainer"
- Cloud Trends
 - Kein Vendor-Lockin durch Open-Source Plattformen
 - Containerisierung, Orchestrierung, Serverless
- Auto-Scaling, Microservices

VORLESUNG 3 (BETRIEB & SICHERHEIT):

- Rechenzentrumsbetrieb
 - Organisation, Technik
 - Ausfallsicherheit, Sicherheit
- Sicherheit
 - Angriffspotentiale, Risiken bei Cloud Computing,
Top 12 Cloud Risiken
 - Maßnahmen, Verantwortlichkeiten
 - Identity Management für Cloud Computing, On-Premise Integration

VORLESUNG 4 (DATENSCHUTZ):

- Datenschutz
 - Rechtliches, Personenbezogene Daten
 - Auftragsdatenverarbeitung
 - Anforderungen an Cloud Dienste

STUDENTENVORTRÄGE:

- XaaS:
 - MBaaS, SECaaS, Storage-as-a-Service, Cognitive-Services
- Cloud Computing Markt:
 - SAP Hana, MS Azure, OwnCloud, Steuerberater-Cloud, Fog-Computing
- Technik:
 - Loadbalancer, AMQP
- Rechtliches:
 - Datenschutz

ABSCHLUSS

Vielen Dank für Eure Aufmerksamkeit !



