

BACHELORARBEIT

Serverless Architekturen für Konventionelle Webanwendungen

Vorgelegt von: Dragoljub Milasinovic
Matrikelnummer: 20140076
am: 23. September 2017

zum
Erlangen des akademischen Grades

BACHELOR OF SCIENCE
(B.Sc.)

Erstbetreuer: Prof. Dr.-Ing. Schafföner
Zweitbetreuer: Jonas Brüstel, M.Sc.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Vorgehen	2
1.2	Ziel	2
1.3	Aufbau der Arbeit	3
2	Klassische Service-Modelle	5
2.1	Cloud Eigenschaften	5
2.2	IaaS	6
2.3	PaaS	7
2.4	SaaS	8
3	FaaS	9
4	Serverless-Angebote und Architekturen	11
4.1	Serverless	13
4.1.1	Pipes and Filters, Compute as a Glue	14
4.1.2	Legacy Api Proxy	15
4.1.3	Compute as a Backend	15
4.1.4	Graph Query	15
4.1.5	Real time processing	16
4.1.6	Priority Queue	16
4.1.7	Fan Out	17
4.1.8	Federated Identity	17
5	AWS-Serverless-Angebote	19
6	KOMA, eine Beispielanwendung	23
6.0.1	Anforderungsanalyse	24
6.0.2	ER-Modell	24
6.0.3	Komponentenübersicht	25
6.1	Umsetzung	26
6.1.1	Datenspeicherung- Analyse und Auswahl	27
6.2	RESTful API	30
6.2.1	Datenhaltung Synchronisierung und Replizieren	37
6.2.2	Datenverarbeitung	38

6.2.3	Abtrennung des Monoliths	38
6.2.4	Single Page Application	38
7	Bewertung	41
8	Ausblick	45

Eidesstattliche Erklärung

Ich versichere hiermit, dass ich die von mir eingereichte Masterarbeit selbstständig verfasst, ausschließlich die angegebenen Hilfsmittel benutzt und sowohl wörtliche, als auch sinngemäße entlehnte Stellen als solche kenntlich gemacht habe. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen.

Brandenburg an der Havel, 21. September 2017

Dragoljub Milasinovic

Abstrakt

Pomodoro @Deutsch @English

1 Einleitung

Idee-Ausführung-Markt

Ideen entstehen, verändern sich, werden im Laufe der Zeit vergessen, manchmal begeistern sie. Ihnen Form und Inhalt zu geben, also sie umzusetzen, ist die Voraussetzung, um nachzuvollziehen ob die ursprüngliche Idee wirklich ausgebaut und verstanden worden ist.

Die Technik kann als Medium für den Ausdruck solcher Ideen eingesetzt werden. Diese kann so komplex werden, dass sie eine Barriere in Form eines Wissensmonopols darstellt, die hinderlich für die Umsetzung neuer Ideen ist.

Die Faktoren am Anfang einer technologischen Umsetzung einer Idee sind:

- Konzeptioneller Beweis
- Vorlaufzeit,Produkteinführungszeit (Time-To-Market)
- Personalkosten und Mangel von Fertigkeiten
- Technische und technologische Details
- Rentabilität

Ein Zeichen für die Existenz dieser Komplexität im Rahmen des Cloudcomputings ist die Entstehung neuer Technologien für die Vereinfachung der Entwicklungsprozesse eines Projekts.

Je mehr Anforderungen auf einem System z.B. Webanwendung, desto komplexer wird es. Je mehr Softwarekomponente, desto mehr Verwaltungsaufwand mittels Load Balancing, Messaging usw. Je mehr Veränderlichkeit, desto schwieriger ihre Integration und Skalierung. [HW04]

Die Cloud Anbieter versuchen diesen Verwaltungsaufwand, Skalierung und Integrationschwierigkeiten mittels einem neuen Architekturstil Serverless.

Um die Umsetzungsvorgänge einer Webanwendung möglichst simpel zu halten, werden in der vorliegende Arbeit die Serverless Architekturen für konventionelle Webanwendungen untersucht.

1.1 Vorgehen

Auf dem Weg zur technologischen Umsetzung einer neuen Idee liegen unbekannte Schwierigkeiten bei den Entscheidungen über ihrer Umsetzung. Problematisch können sich der Architektorentwurf, die IT Infrastruktur, die Drittanbieter von Software, die Auswahl der Infrastruktur usw. gestalten. Hinzu kommen Schwierigkeiten, die spezialisierte Kompetenzen, Fertigkeiten und „Know-How“ erfordern. Diese gehören jedoch nicht immer zum Problem der Domain der Anwendung.

Der Begriff Serverless weist darauf hin, dass die Verwaltung der zugrunde liegenden Serverinfrastruktur der Anwendung von Cloudanbietern übernommen wird.

Für dieses Problem wurde Function as a Service (FaaS) ³, als Lösung unter der Rubrik „Serverless“ ^{4.1} von den Hauptanbietern von „Cloud“ Technologien vorgestellt.

FaaS definiert das Programmiermodel, eine Funktion oder auch „Nano-Microservice“ genannt, um den serverless Architekturstil zu adoptieren.

Im Rahmen des Cloud-Computing handelt es sich in dieser Arbeit um eine Untersuchung der Serverless Architekturen am Beispiel einer konventionellen Webanwendung. Dabei wird besonders beachtet, ob und wie solche Technologien die Umsetzung erleichtern. Die Entwurfsmuster und die Kernfunktionalität werden ausschließlich mit Serverless Technologien am Beispiel einer Webanwendung (Kompetenz Matrix (KOMMA), siehe Kapitel 6), mit AWS umgesetzt.

1.2 Ziel

Das Ziel ist ein Minimal Viable Product (Mivip), in Form von einer Single Page Application (SPA) ^{6.2.4}, ausschließlich mit Serverless Technologien zur Verfügung zu stellen.

Nach der Umsetzung werden die Erfahrungen und Ergebnisse ausgewertet, um bei zukünftigen Entscheidungsprozessen bei der Umsetzung einer Webanwendung zu unterstützen.

Die Webanwendung soll möglichst flexibel für zukünftige Änderungen sein.

1.3 Aufbau der Arbeit

Die vorliegende Arbeit beschäftigt sich mit dem Serverless Ausschnitt der Cloud Dienste.

Zuerst wird der Leser in die klassischen Servicemodelle 2 eingeführt. Zunächst werden die technischen Anforderungen und die dazugehörigen Beispiele des Serverless Ansatzes erläutert. Das nächste Kapitel überblickt die aktuellen Serverless Angebote der größten Cloud Anbieter. Den Kern der Arbeit bildet die Analyse und Darstellung von Serverless Architekturen 5 und sie fokussiert sich auf Amazon Web Services (AWS). In diesem Abschnitt 5 werden die Serverless4.1 Architekturen und das Programmiermodell vorgestellt, sowie die Entscheidungsprinzipien4.1 erläutert. Der praktische Teil beschäftigt sich mit der Umsetzung und Bewertung von der oben genannten Serverless Webanwendung KOMA 6.

Am Ende erfolgt eine Diskussion darüber, welche Trade-offs entstehen und welche Zukunftsperspektiven Serverless Technologien bieten.

2 Klassische Service-Modelle

Cloud Computing beschreibt die Bereitstellung von IT-Infrastruktur und IT-Leistungen wie beispielsweise Speicherplatz, Rechenleistung oder Anwendungssoftware als Service über das Internet. [Cha14]

2.1 Cloud Eigenschaften

Als Softwarearchitekt, Entwickler oder Projektmanager es ist wichtig, die spezifischen Eigenschaften von Cloud Angeboten zu verstehen. Aus einem Meer von Cloud Diensten ist die richtige Auswahl je nach Anforderungen und Art der technologischen Umsetzung schwer zu treffen.

Im Allgemeinen teilen Cloud Angebote laut Chandrasekaran (2014) folgende Eigenschaften:

- On-Demand Self-Service - Nutzer können die IT-Kapazitäten, die sie benötigen, selbständig ordern und einrichten. Der Anbieter muss in den Prozess nicht eingebunden werden.
- Broad Network Access bezeichnet den standardbasierten Netzzugriff von verschiedenen Endgeräten (z.B. Smartphones, Tablets, Laptops, PCs) aus.
- Measured Service bietet eine automatische Kontrolle und Optimierung der genutzten Ressourcen durch Metering, wodurch Transparenz für Anbieter und Nutzer sichergestellt wird. Somit bezahlen Kunden nur die Dienstleistungen, die sie auch tatsächlich in Anspruch nehmen.
- Resource Pooling - 3 – Ressourcen des Anbieters (z.B. Speicher oder Bandbreite) werden gebündelt, multimandantenfähig bereitgestellt und nach Bedarf zugewiesen.

- Rapid Elasticity - 2.1 – Kapazitäten sind schnell und dynamisch verfügbar und können je nach Bedarf skaliert werden. [Cha14]

Daher ergeben sich für die, in der Cloud betriebenen Anwendungen, folgende Eigenschaften:

- Isolated state - Der Zustand wird in kleinen Einheiten der Anwendung isoliert, so dass sie besser skaliert. Eine zustandslose IT Ressource kann ohne Synchronisierungen aggregiert oder gelöscht werden. Dieser Zustand bezieht sich nicht nur auf die Verwaltung der Interaktionen eines Clients, sondern auch auf dessen Datenverarbeitung.
- Distribution - Anwendungen müssen so in Komponenten zerlegt werden, dass sich ihre Ressourcen weltweit auf-/verteilen.
- Elasticity - Die Anpassung sowohl auf die Anzahl als auch auf die Leistungsfähigkeit der zu benutzenden IT Ressourcen kann im Sinne einer Addition oder Subtraktion erfolgen. Im ersten Fall nimmt auf der Ebene der horizontalen Skalierung (scale out) die Anzahl der Server zu. Während bei der vertikalen Skalierung (scale up) die Leistungsfähigkeit der Ressourcen der Server steigt.
- Automated management - Die konstanten Aufgaben zur Verwaltung von Elasticity sollen automatisiert werden, um eine Cloudanwendung fehlerresistent auf Ressourcenebene zu implementieren.
- Loose coupling - Die Minimierung von Abhängigkeiten einer Anwendung von IT Ressourcen vereinfacht die Bereitstellung, die Fehlerkontrolle und Wiederverwendung von Komponenten. 4 [Cha14]

2.2 IaaS

Infrastructure as a Service (IaaS) kann als ein Service beschrieben werden, der Abstraktionen für Hardware, Server und Netzwerkkomponenten bereitstellt. Der Serviceanbieter besitzt die Ausrüstung und ist für die Unterbringung, die Inbetriebnahme und die Wartung der Server verantwortlich [You15]. Der Benutzer bezahlt nicht für die Hardware, deren Lagerung und den Zugriff auf sie, sondern für die Nutzung des gesamten Servicemodells z.B.: Zahlung nach benutzten Stunden, Ressourcen usw.

Die Aufgaben für Systeme mit Softwareelementen wie Load Balancing, Transaktionen, Clustering, Caching, Messaging, und Datenredundanz werden komplexer. Diese Elemente fordern an, dass Server verwaltet, gewartet, patched und gesichert werden brauchen. In einer nicht-trivialen Systemumgebung sind solche Aufgaben zeitintensiv und ebenso aufwändig fertigzustellen, wie effizient zu betreiben. Infrastruktur und Hardware sind zwar nötige Komponenten für jegliche IT-Systeme, aber gleichzeitig stellen sie nur das Medium für deren Anwendung dar - sei es Geschäftslogik, oder ein darauf bauender Dienst.

2.3 PaaS

Platform as a Service (PaaS), kann als ein Service beschrieben werden, der eine Rechenplattform liefert, z.B. ein Betriebssystem, eine Ausführungsumgebung für Programmiersprachen (siehe ElasticBeanstalk), eine Datenbank oder einen Webserver. Dieser Dienst übernimmt je nach benutzerdefinierter Konfiguration sowohl die Wartung der Datenbank, des Webserver und der Versionen des Laufzeitquellcodes, als auch deren Skalierbarkeit. [You15].

Inkonsistenzen in der Infrastruktur oder den Umgebungen können durch den hohen Aufwand der Serververwaltung entstehen. Sie werden durch standardisierte und automatisierte Angebote von PaaS umgangen. Deren effiziente Benutzung ist abhängig davon, wie gezielt der Quellcode auf die Features der Plattform abgestimmt ist. Dies ergibt auf einer Seite weniger Wartung, aber andererseits erfordern die importierten Anwendungen (z.B. für ein Standalone Server) eine Anpassung an die Plattform.

Die Containerisierung ist eine Isolierung der Anwendung von ihrer Umgebung. Die Konfiguration des Containers, sowie dessen Deployment ist nicht trivial und erfordert daher spezialisiertes Wissen über Containerisierung. Für das Monitoring werden bestimmte Tools wie Boot2Docker [Goo17a] oder cAdvisor [Goo17b] benötigt. Jedoch bietet die Containerisierung eine ausgezeichnete Lösung für Anwendungen mit starker Kopplung zu anderen Softwarekomponenten. [You15]

2.4 SaaS

Software as a Service (SaaS) kann als ein Service beschrieben werden, der OS-Images mit konfigurierbaren Diensten wie Datenbanken, Webanwendungen usw. bereitstellt. SaaS gestaltet sich benutzerfreundlich, da die Konfiguration und das Deployment dieser Softwaredienste nicht erlernt werden müssen, um sie in eine größere Anwendung einzubinden. Anfallende Gebühren berechnen sich nach der Nutzungsdauer. Viele traditionelle Software bietet seine SaaS nicht an. Dies Impliziert laut Chandrasekaran, K. dass dieses Serviceliefermodell sich für Anwendungen nicht gut eignet wegen folgenden Punkten:

- Geringe Latenz kann durch die Entfernung der gespeicherten Daten für Echtzeitanwendungen nicht gewährleistet werden.
- Die Datensicherheit kann nicht sichergestellt werden, da die mitbeteiligten Drittanbieter bei SaaS die Service Level Agreement (SLA) von Kunden nicht immer erfüllen.
- Anforderungen bestimmter Software verlangen eine Zentralisierung und eine Lokalisierung vor Ort, anders als bei SaaS.

3 FaaS

Function as a Service (FaaS) kann als ein Rechenservice beschrieben werden, der nach Anfrage isoliert, unabhängig und granular ausgeführt wird. Komplexe Probleme wie horizontale und vertikale Skalierbarkeit, Fehlertoleranz und Elascitity werden von Kunden nur noch nach Bedarf konfiguriert und von dem Anbieter verwaltet. Die Besonderheit von FaaS ist die „Unit of Deployment“ und die Skalierung in Form einer Funktion. [You15]

Events innerhalb eines verteilten Systems müssen verwaltet werden. Technologien wie Virtualisierte oder Containerized Server erzeugen neue Serverinstanzen zur Verarbeitung von einer Kette von variablen Events, die danach gelöscht wird. [Kin16] Die entstehende Problematik ergab sich durch eine starke Zunahme an Elementen, die einen hohen Verwaltungsaufwand forderten. Was zurück auf die oben genannten Aufgaben 2.2 führte.

Polling ist der Ausdruck für eine zyklische Abfrage über einen Status z.B. von Ports oder Locks über Ressourcen. Die Verwendung von Systemressourcen ist ineffizient im Vergleich zu Alternativansätzen wie z.B in dem Push- oder Pull- Kommunikationsmodell. [Kin16]

Funktionale Programmierung ist ein Programmierparadigma, in dem Funktionen nicht nur definiert und angewendet werden können, sondern auch wie Daten miteinander verknüpft, als Parameter verwendet und als Funktionsergebnisse auftreten können. Zustand und mutable Daten werden vermieden damit Seiteneffekten nicht entstehen und die Komposition flexibler wird. Das stellt einen Vorteil für die Skalierung eines Softwarekomponenten dar. [Ray13]

Die Implementierung einer solchen Funktion geschieht durch die Auswahl der Programmiersprache und der von dem Cloud-Anbieter vorgegebenen Funktionsfassade.

Diese wird vom Cloud-Anbieter aufgerufen, stellt aber keine zusätzlichen Bibliotheken bereit, daher ist es nötig, dass die auszuführende Datei alle Abhängigkeiten enthält.

Der Fokus bei FaaS liegt auf der Quellcodeentwicklung und nicht auf dem Provisioning von Servern, der Installation von Software, dem Deployment von Containern oder auf konkreten Details der Infrastruktur.

Für die Betrachtung, ob FaaS eine Lösung für eine konkrete Problemstellung ist, folgt eine Auflistung von Kriterien:

Es ist nicht empfehlenswert FaaS zu benutzen, wenn:

- der Entwickler Rootzugriffsrechte auf alle Ressourcen eines Servers benötigt.
- die Priorisierung von Betriebssystemattributen wie CPU, GPU, Networking oder Speichergeschwindigkeit angefordert ist.
- Sicherheit relevant ist. Unautorisierte Zugriffe können mit FaaS nur auf Systemebene erkannt werden.
- dauerhafte Prozesse angefordert sind.

Es ist empfehlenswert FaaS zu benutzen, wenn:

- Aufgaben als Reaktion auf Events erledigt werden.
- ein Scriptbehälter für z.B Cron Aufgaben benötigt wird. Die Zugriffsrechte sind beschränkter, Fehler sind einfacher zu erkennen und an einer Stelle aggregiert (siehe CloudWatch), des weiteren können Deployments einfacher angestoßen werden.
- die Skalierung des Servers bei einer ressourcenintensiven Verarbeitung vermieden werden soll.4.1.1
- Services vorgegeben sind, die selten benutzt werden.
- die Verwaltung von API-Server umgangen werden soll.

[Kin16]

4 Serverless-Angebote und Architekturen

Die Softwarearchitekturen helfen uns zu kommunizieren, welchen Zweck unsere Software erreichen möchte. Ihre **Entwurfsmuster** bieten generische Lösungen für wiederkehrende Probleme bei der Softwareentwicklung.

Auf konventionelle Webanwendungen wird hier Bezug genommen, als ein System, das über Presentation-, Data-, and Logik-Tiers verfügt. Jedes Tier kann mehrere Logik-Layers enthalten, die für unterschiedliche Funktionalitäten der Domains verantwortlich sind. Logging ist ein Beispiel für Cross-Cutting Concern, das Layers überspannt. Die Komplexität der Anwendung wächst zusammen mit der Beschichtung. Eine Überprüfung der Architektur ist sinnvoll, wenn eine erfolgreiche Codeänderung von einer Anderen abhängt.

Service Oriented Architecture (SOA) unterlegt der Annahme, dass ein System aus mehreren kleinen, austauschbaren, wiederverwendbaren und entkoppelten Diensten besteht. Für die Entwicklung und Integration solcher Systeme bietet SOA eine Menge von Entwurfsprinzipien und Standards. Entwickler stellen autonome Services her, die durch Nachrichtenübergabe kommunizieren und oft einen Schema haben oder eine Schnittstelle, die definiert wie die Nachrichten erzeugt werden. [Cha14]

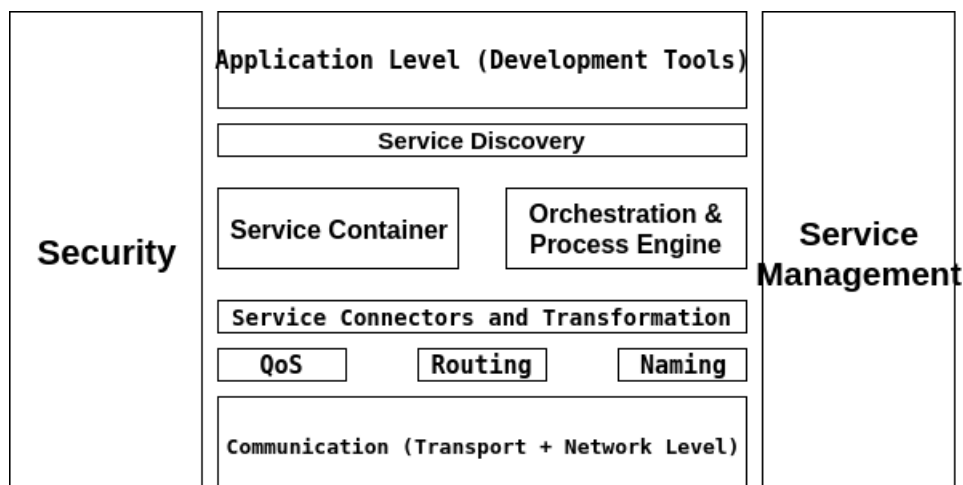


Abbildung 4.1: SOA Architektureferenz [Sta07]

SOA ermöglicht gegenseitigen Datenaustausch zwischen Programmen von unterschiedlichen Anbietern, ohne dass zusätzliche Änderungen an den Services vorzunehmen sind. Die Bestandteile eines solchen Architekturstils sind sowohl Standardschnittstellen als auch voneinander unabhängige Services [Sta07].

Der Fokus in der Cloud liegt daher auf Service und Servicekomposition [Cha14].

Microservices und Serverless versuchen die Komplexität der SOA (Abbildung 4.1) anzusprechen. Beide Ansätze führen Separation of Concerns, häufige Deployments und Heterogene Domain Specific Language (DSL) mit sich. [H⁺17]

Auf einer Seite können Microservices ihren Zustand und Daten speichern und mit Hilfe von Frameworks implementiert werden. Auf der anderen Seite sind Serverless zustandlos, ihre Datenspeicherung ist zeitlich begrenzt und sie unterstützen Frameworks nicht direkt. ConnectWise [Con17a], Netflix [Net17] und UNLESS [UNL17] sind Beispiele für Unternehmen, die auch von Serverless Architekturen profitieren.

Event Driven Architecture (EDA) ist eine Softwarearchitektur, in der das Zusammenspiel der Komponenten durch Ereignisse gesteuert wird. Die Ereignisorientierung besitzt das Potenzial, dass die Architekturen von Anwendungen agiler, reaktions-schneller und echtzeitfähig werden. Laut Ralf Bruns und Jürgen Dunkel ist EDA für komplexe Fachlogik, große Datenvolumina, geringe Latenzzeit, Skalierbarkeit und Agilität geeignet. [RB]

Die Serverless Technologien können durch Benachrichtigungen gestartet werden. Dieser EDA-Stil verstärkt die Entkopplung auf einer temporären Ebene zwischen Producer

und Consumer. Weiterhin ermöglicht ein Kommunikationskanal zur Benachrichtigung eine asynchrone Verarbeitung, ohne dass das System auf Grund von Fehlern abstürzt. [HW04]

Zusammenfassend sind folgende Vorteile ersichtlich:

- Die Wiederverwendung von Services in unterschiedlichen Anwendungen senkt die Entwicklungskosten und den Time-To-Market.
- Durch die Standarisierung der Services kann ein System mit einer Rekonfiguration und ohne Weiterentwicklung schnell auf die geschäftliche oder externe Bedürfnisse angepasst werden. Somit wird ein agiles Arbeiten möglich.
- Das Monitoring hilft Fehler zu erkennen und die Leistung zu messen.
- Aggregate von bereitgestellten Services können komplexere und domainübergreifende Aufgaben ausführen.

[Cha14]

Im späteren Kapitel wird REST als Teil des EDA-Architekturstils vorgestellt.

4.1 Serverless

Serverless kann als ein Ansatz beschrieben werden, der die Verwendung von einem Rechenservice, Dienste von Drittanbietern, von Application Interface (API)s und die Anwendung von Architekturmustern fördert. Ein solcher Anwendungsfall ist die Kommunikation mithilfe eines „Delegation-Tokens“ zwischen den Front- und Back-End Diensten. FaaS ist nur ein Aspekt dessen.

Serverless übernimmt die Entwurfsprinzipien von SOA und EDA 4 und daher ergeben sich laut Sbarski folgende Richtlinien:

- Ein Rechenservice wird genutzt, um Quellcode auf Anfrage auszuführen, kein Server.
- Zustandslose Funktionen unterliegen dem SRP Prinzip.
- Für den Architekturentwurf werden Push basierte ereignisorientierte Pipelines genutzt.
- Front-Ends werden durch die Einbettung von mehr Zuständigkeiten verstärkt.

- Dienste von Drittanbieter dem Schreiben von eigenem Quellcode bevorzugt.

[Sba17]

Die Vernetzung von zustandslosen Funktionen erlaubt, komplexe Systeme zu entwerfen, die einfach zu skalieren sind. Die Komplexität und längerfristige Wartbarkeit des Systems lässt sich dadurch reduzieren, dass der Controller und/oder Router aus dem Model View Controller (MVC) [Fow17a] vom Back- zum Front-End verschoben wird und Dienste von Drittanbieter hinzugefügt werden. [You15]

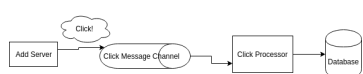


Abbildung 4.2: Classic Add Server

In einem beispielhaften konventionellen AdServer wird nach einem Click auf eine Werbung eine Nachricht über ein Kanal an einen Clickprozessor geschickt, der innerhalb einer Anwendung ausgeführt wird.

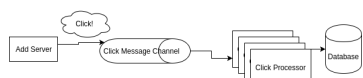
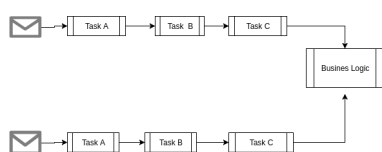


Abbildung 4.3: Serverless Add Server

Mit dem Serverless Ansatz wird dieser Clickprozessor pro Nachricht als eine neue Instanz der Funktion ausgeführt. Ihre Laufzeitumgebung und ihr Messagebroker wird von dem Cloudanbieter verwaltet. [Fow17b]

Dazu wird für den Entwurf von Serverless Architekturen eine Reihe von Mustern von unterschiedlichen Autoren vorgeschlagen.

4.1.1 Pipes and Filters, Compute as a Glue

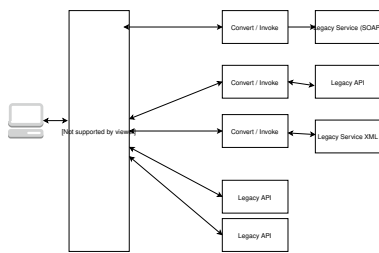


Eine Anwendung kann Aufgaben von unterschiedlicher Komplexität ausführen. In einem monolithischen Modul sind das Refactoring, die Optimierung und die Wiederverwendung erschwert.

Zunächst werden die Aufgaben in diskrete Elemente (oder Filter) nach dem SRP zerteilt und in einer Pipeline kombiniert. Dies hilft

redundanten Quellcode zu vermeiden, ihn zu löschen, zu ersetzen oder zu integrieren in zusätzliche Komponenten, sich sobald die Aufgabenanforderungen ändern [HSB⁺14]. Ein weiterer Vorteil besteht darin, dass ein Flaschenhalseffekt vermieden wird, in dem mehreren Instanzen erzeugt werden, falls ein Element nicht genug Ressourcen für die Verarbeitung hat.

4.1.2 Legacy Api Proxy



Wenn eine API veraltet oder schwer zu benutzen ist, kann eine extra (RESTful) API in den Vordergrund gestellt werden, die in gesonderten Prozessen Daten transponieren und für die angeforderten Formate aufstellen (marshall). Dies ist besonders nützlich, wenn die Legacy-Services selten benutzt werden. Zusätzlich erleichtert die Api Proxy die Integration mit anderen

Architekturansätzen.

4.1.3 Compute as a Backend

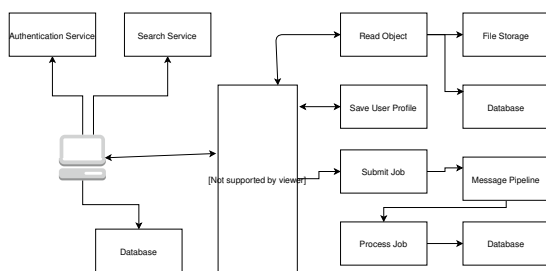


Abbildung 4.4: Compute as a Backend

Obwohl der Client direkt mit Services kommunizieren kann, müssen vertraute Informationen geschützt werden, indem sie das Back-End verarbeitet [Sba17]. Diese Aufgaben können hinter einer REST Schnittstelle koordiniert werden. Wie in Abschnitt 4.1 erwähnt, minimieren die Einbettung der Dienste von Drittanbietern und verstärkte Front-Ends

den Fußabdruck des eigenen Back-Ends.

4.1.4 Graph Query

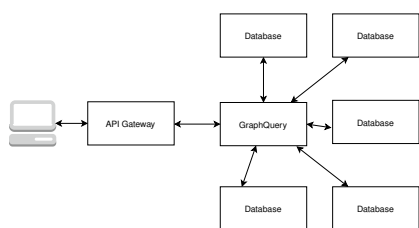


Abbildung 4.5: Graph Query

Wenn mit einer Anfrage mehrere Datenbanken abgefragt werden, entstehen multiple Paketumlaufzeiten (Round-Trip). Stellt eine REST Schnittstelle in einer Anfrage zu wenig Queryparameter zur Verfügung, dann entsteht Overfetching, weil die Anfrage nicht präzise genug ist. Dagegen kann der Client die Parameter für die Abfrage spezifizieren und das Back-End

baut sie zusammen und führt sie aus.

4.1.5 Real time processing

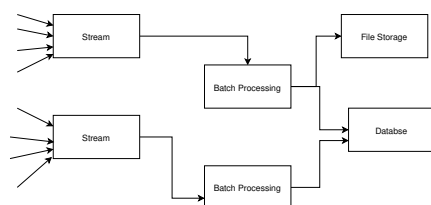


Abbildung 4.6: Real Time Processing

Die Verarbeitung von Streams in Echtzeit kann durch einen Buffer, der je nach Konfiguration die Daten weiter an den Worker leitet. Einen wesentlichen Vorteil stellt die unabhängige Skalierung des Streams und des Workers je nach Anfrage dar. Bei fehlerhafter Verarbeitung werden die Prozesse neu angestoßen.

4.1.6 Priority Queue

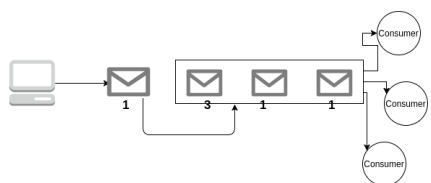


Abbildung 4.7: Priority Queue

Anwendungen können spezifische Aufgaben delegieren, wie z.B die Integration mit anderen Anwendungen und Services. Am Beispiel einer FIFO (First In First Out) Queue können die Nachrichten nach Priorität automatisch sortiert und asynchron verarbeitet werden. Für Systeme ohne integrierte Priorisierung können mehrere Queues für unterschiedliche Prioritäten benutzt werden und die Anzahl von Consumerprozessen wird entsprechend angepasst. Im letzteren Fall wird dabei die Starvation von Nachrichten, mit geringer Priorität, vermieden.

Das Priority Queue wird im Rahmen des Entwurfsmusters Competing Consumers als ein Consumerservice aufgefasst. Steigt stark die Anzahl an Anfragen die Kapazitäten des Systems, kommt zu seiner Überbelastung. Ein Consumerservice als Moderator von Anfragen ist im Stande das zu verhindern.

4.1.7 Fan Out

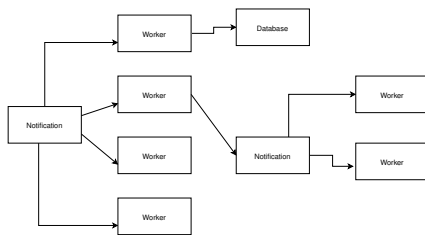


Abbildung 4.8: Fan Out

Bei dem Eintritt eines Events können ein oder mehrere Subscriber durch die gepush-te Benachrichtigung angestoßen werden. Damit wird ein Kommunikationskanal 4.1.6 zur Verwaltung von Nachrichten wiederverwendet und extra Geschäftslogik umgangen, z.B. kein Command Pattern für die gleiche Funktionalität [Sba17].

4.1.8 Federated Identity

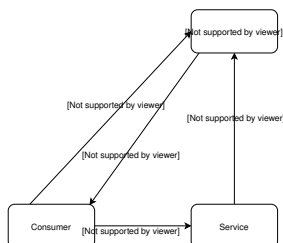


Abbildung 4.9: Federated Identity

Authentisierung vermieden werden .

Benutzer arbeiten mit multiplen Anwendungen von unterschiedlichen Organisationen. Um gleiche Zugangsdaten für Benutzer wiederverwenden zu können, wird dem Nutzer eine Identität bei einem Drittanbieter zugeteilt, diese wird mit einem Token an die Anwendung weitergeleitet. Der Authentisierungscode kann daher von dem Anwendungscode abgetrennt werden, damit kann zusätzliche Codierung für die

5 AWS-Serverless-Angebote

AWS ist ein Cloudanbieter



Lambda ist ein Rechenservice, der aus Quellcode und dessen Abhängigkeiten besteht. Die horizontale Skalierung erfolgt automatisch, elastisch und vom Cloudanbieter verwaltet; währenddessen die vertikale nach Konfiguration erfolgt. Lambda wird als Einheit für die Skalierung und das Deployment benutzt. AWS unterstützt Javascript, Python, C# und Java. Die letzte Programmiersprache betrifft das Konzept „Cold- und Warm run“ besonders, da es nötig ist, die Java Virtual Machine (JVM) in ihrer Laufzeitumgebung hochzufahren (cold run). Der Cloudanbieter stützt sich auf das Pull- und Push- Kommunikationsmodell. Bei Pull überprüft die Lambda Laufzeitumgebung in regelmäßigen zeitlichen Abständen, ob Events z.B. bei Kinesis eingetreten sind und ruft die entsprechende Lambda Funktion mit einer Event-Datennutzlast auf. Bei Push ruft die Eventquelle z.B S3 je nach Konfiguration (event source mapping) die entsprechende Lambda Funktion auf.



API Gateway ist eine Fassade, um Operationen sicher auszuführen, wie die Benachrichtigung von Kunden per Email und deren Identitätsüberprüfung. Weil AWS die Skalabilität von API Gateway und von Lambda übernimmt, ist die Bereitstellung und Wartung von EC2 Instanzen und die Konfiguration deren Load Balancer nicht mehr nötig.



Simple Notification Service (SNS) erweitert das schon gut etablierte Beobachtermuster, in dem es einen Kanal für Events hinzufügt. Die konzeptionelle Technologie hinter SNS wird „Publish-Subscribe

Channel [HW04]“ Muster genannt und entspricht den oben genannten „Fan Out“ 4.1.7 Architekturmustern. AWS kann durch Redundanz mindestens eine Lieferung der Nachricht gewährleisten.



Simple Storage Service (S3) ist ein Speicherdienst, der durch SNS 5 die Events (z.B das Löschen oder Erzeugen eines Objektes) an SNS5, Simple Queue Service (SQS) 5 oder Lambda schicken kann. „Buckets“ sind Verzeichnisse auf der höchsten Ebene des Verzeichnissystems. Ein Objekt ist eine Kombination von Daten, Metadaten und eines Keys, der innerhalb des Buckets eindeutig ist.



DynamoDB ist eine NoSQL Datenbank. Ihre Tabellen bestehen aus Items (Zeilen) und deren Attributen (Spalten). Die Datenbank hat laut AWS unendliche Datenkapazitäten und der Datenverkehr ist unbegrenzt. Durch automatische Skalierung mindert sich die Leistung nicht. Bei der Änderung eines Zeilenwertes in DynamoDB ist die Konfiguration eines Anstoßes der Lambdafunktion möglich.



Simple Queue Service (SQS) ist eine message queue. Sie erlaubt die Interaktion von mehreren Publishers und Consumers in einer SQS und automatisch den Lebenszyklus der Nachrichten verwaltet und kontrolliert Auszeiten (Time out) oder individuelle Verzögerungen.



Kinesis Streams ist ein Service für Echtzeitverarbeitung von Datenstreams. Es wird für Logging, Datenimport, Metriken, Analytics und Reporting benutzt. Ein Kinesis Stream ist eine sortierte Folge von Datensätzen, die auf „Shards“ verteilt sind. Diese definieren die Kapazität des 'Durchsatzes' (Throughput) von einem Stream und können nach Bedarf vergrößert werden.



Relational Database Service (RDS) hilft bei dem Setup und Wartung von MySQL, MariaDB, Oracle, MS-SQL, PostgreSQL und Amazon Aurora mit automatischer Bereitstellung (provisioning), Sicherung, patching,



recovery, repair und Fehlererkennung. RDS kann durch SNS 5 über eigene Events berichten.



Simple Email Service (SES) behandelt die Absendung und die Empfangsoperationen wie Spamfilterung, Virus Scann und Ablehnung nicht vertrauter Quellen. Emails können weiterhin in S3 gespeichert, an Lambda versendet oder eine SNS Benachrichtigung erstellen.

6 KOMA, eine Beispielanwendung

Im Folgenden wird eine Beispielanwendung, unter Verwendung der in Kapitel 5 genannten Entwurfsmuster, implementiert.

Die Beschleunigung der Veränderungen in der heutigen Gesellschaft und der technologischen Landschaft prägt sich in Bildung und Beruf in so fern aus, dass die heutigen Rahmenlehrpläne nicht mehr fachlich, sondern an der Kompetenzentwicklung orientiert sind, um u.a Kompetenzprofile für Lerner zu erstellen. Es existiert bereits ein anerkannter Europäischer Rahmen [EQF17] für Kompetenzbildung: European Qualifications Framework (EQF). Am Beispiel von Sachsen-Anhalt [SA17] werden diese Kompetenzorientierung auf die spezifischen Bedürfnisse der Schule beschrieben und die Unterrichtsstunden entsprechend gestaltet.

Die Anwendung soll für die pädagogische Diagnostik und Intervention genutzt werden. Ziel der Umsetzung ist es daher, auf einer Seite die von einem Schüler erworbenen und zu erwerbenden Kompetenzen und deren Niveau nachzuvollziehen. Andererseits bietet sie das Potenzial die Bildungs-, Unterrichts- und Stundenplanung zu unterstützen.

Kommt diese Anwendung in Bildungsinstitutionen zum Einsatz, dienen die Rahmenlehrpläne als Leitpfad für die Bezeichnungen und Anforderungen der Kompetenzen, wohingegen KOMA für die Organisation der einzelnen Fachrichtungen oder Lehrveranstaltungen zuständig ist. Da der EQF als Basis mit internationaler Anerkennung genutzt wird, den Kompetenzstand eines Individuums abzubilden, wird die Bildungsqualität international vergleichbar.

Den Kern von KOMA bildet die Zuweisung von Aktivitäten auf vordefinierte Kompetenzen. Erstere lassen sich einzeln oder in einer Sequenz anordnen. Sequenzen werden in Lehrveranstaltungen zusammengestellt. So können Aktivitäten, Sequenzen und Kompetenzen als Gestaltungsmittel für Lehrveranstaltungen genutzt werden.

Das folgende Beispiel beschreibt einen fachorientierten Ansatz zur Gestaltung von Lehrveranstaltungen:

Das Fach „Web Computing“ lässt sich mit einer Sequenz von Lernaktivitäten (Unterrichtseinheiten) gestalten. Die zu behandelnden Themen erfordern grundlegendes Wissen und Fertigkeiten wie das Beschreiben von Kommunikationsprotokollen und Bash Scripting. Dabei ist zu beachten, dass Wissen und Fertigkeiten kumulativ wachsen und damit aufeinander aufbauen.

Im Gegensatz zu dem fachorientierten Ansatz wird nun der kompetenzorientierten Ansatz erläutert: Das Kompetenzmodell differenziert unterschiedliche Kompetenzen in vielfältigen Zusammensetzungen. Ihre Fertigkeiten und/oder Wissen werden während der Lernaktivitäten, bei der die Auswahl von Themen freigestellt ist, erworben. Durch die Sequenzierung von Lernaktivitäten baut sich das Kompetenzmodell aus.

6.0.1 Anforderungsanalyse

Die Auflistung 6.0.1 stellt einen für diese Arbeit angepassten Ausschnitt der Anforderungen für KOMA dar:

- Mit dem EQF vergleichbare Kompetenzdefinitionen
- Berücksichtigen zukünftiger Erweiterungen
- Abrufbarkeit durch den Browser
- Private Datenspeicherung u.d Login
- Ertragen von großen Nutzlastschwankungen

6.0.2 ER-Modell

Aus der Beschreibung von KOMA Kapitel 6 ergibt sich folgendes Modell. Unterabschnitt 6.0.1

Dieses ER-Diagramm Unterabschnitt 6.0.1 definiert die Beziehungen zwischen den in Kapitel 6 beschriebenen Konzepten. Eine Kompetenz (Competence) besteht aus Fertigkeiten (Skill), die in einer Lernaktivität (LearningActivity) erworben werden. Die Letztere gehört zu einer Sequenz von Lernaktivitäten (LearningSequence). Erworbene Kompetenzen können nach Klassenstufen und Fächern aufgerufen werden, damit

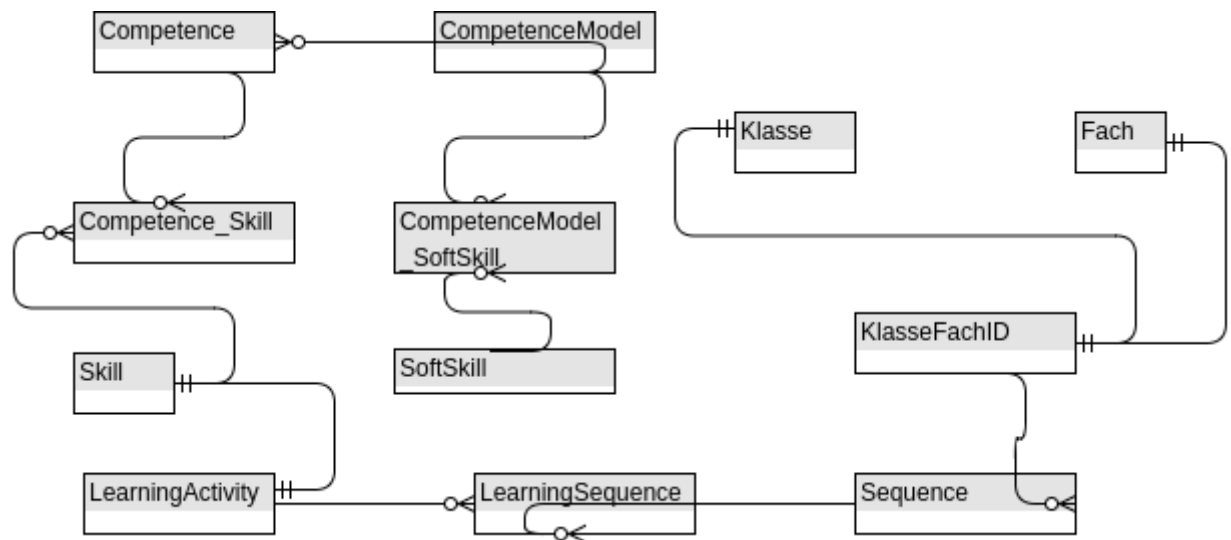


Abbildung 6.1: Entity-Relationship Modell für KOMA

ein Kompetenzstand abgeleitet werden kann. Das Modell ermöglicht die Gestaltung von EQF-konformen Kompetenzprofilen (`CompetenceModel`).

6.0.3 Komponentenübersicht

Um einen Anhaltspunkt zu geben, werden hier die Softwarekomponenten beschrieben. Bei dem Aufruf der Website im Browser präsentiert sich ein „Sign in“ Knopf zum Einloggen und verschiedene Möglichkeiten zur Formulierung einer Abfrage. Diese ist in einem S3 Bucket als statische Webseite gelagert. Wird der Knopf zum Einloggen gedrückt, fordert eine Weiterleitung (mit dem dargestellten Schlüssel) die Authentisierung des Benutzers an. Mit der erfolgreichen Operation kehrt die Nutzeransicht mit einem „Token“ auf die Webseite zurück. Der erhaltene Token wird als Autorisierung-Parameter in dem „HTTP Request Header“ mit den nächsten Anfragen an das Back-End geschickt.

Nach dem Login kann der Browser durch Absenden der nächsten Anfragen seinen Token zur Überprüfung übergeben. Die API Gateway empfängt die Anfragen und transponiert deren Parameter, um zunächst die entsprechende Lambdafunktion synchron aufzurufen. Da die API nach REST entworfen ist, werden die Lambdafunktionen nach der HTTP-Methode und der URL abgebildet.

Dieser Loginprozess entspricht dem in Unterabschnitt 4.1.8 beschrieben Federated

Identity Muster.

Eine Anfrage an die GET `https://<host>/page/{individual}` URL führt die Lambdafunktion „OWL Parser“ aus. Die Funktion liest die Datenbasis von S3 „OWL Storage“ und extrahiert den Wert des `{individual}`-Parameters, um eine Antwort zu generieren. Diese abstrahierte Zuweisung von URLs auf Back-End Dienste entspricht einem REST Ansatz. Dagegen erwartet die POST `https://<host>/sparql` URL eine Abfrage im „Request Body“ der Anfrage mit dem Key „Query“, um sie auf der oben genannten Datenbasis auszuführen. Dabei unterstützt die Java Bibliothek Jena ARQ [Fou17]

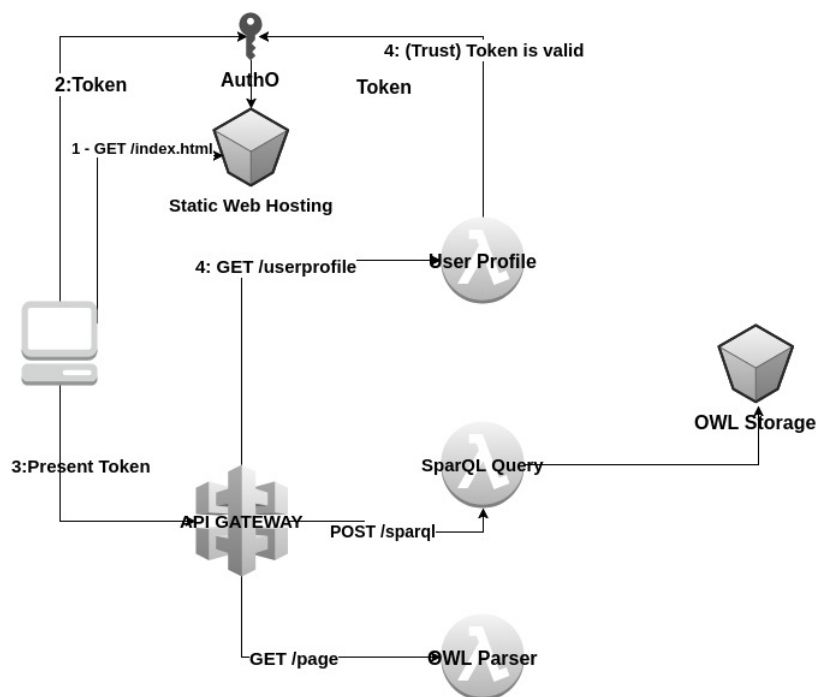


Abbildung 6.2: KOMA Components

6.1 Umsetzung

Die grundlegende Vorgehensweise bei der Umsetzung dieser Webanwendung gliedert sich zunächst in die Datenspeicherung und deren Analyse bzw. Auswahl. Als Zweites wird zwischen dem Entwurf mit den oben vorgestellten Serverless Architekturmustern und Technologien und deren Implementierung iteriert, um ein schnelles Feedback zu erhalten.

6.1.1 Datenspeicherung- Analyse und Auswahl

Die Gestaltung von Kompetenzmodellen und deren zukünftige Weiterentwicklung hängt stark von den spezifischen Bedürfnissen der jeweiligen Schulen ab. Die möglichen Erweiterungen oder Anpassungen des Modells stellt die Benutzung des einer relationalen Datenbank für KOMA in Frage. In der folgenden Tabelle werden die Eigenschaften von relationalen mit ontologischen Schemas verglichen.

Tabelle 6.1: Vergleich relationalem mit ontologischem 6.1.1 Schema

Eigenschaft	Relational	Ontologisch
Weltannahme	Existiert nur	Existiert mindestens
Individual	muss Unique	kann ≥ 1
Info	Ableitung = x	ja
Orientation	Data	Bedeutung

Das ausgewählte Datendarstellungsformat ist das ontologische Schema, da es den Fokus auf Erweiterbarkeit und semantische Konzepte legt. Im „Semantic Web“ profitieren „Linked Data Driven Web Applications“ von den vernetzten Datenbanken. [Con17b]

Das „Semantic Web“ ist eine Erweiterung des herkömmlichen Web, in der Informationen mit eindeutigen Bedeutungen versehen werden [GOS09]. Das World Wide Web Consortium (W3C) spezifiziert eine Zusammenstellung von Standards und best practices für die Mitteilung von Daten und deren semantischer Darstellung. [Bob13]. Diese Bedeutungen werden für Maschinen durch Ontologien dargestellt, welche in „owl“ Dateien gespeichert werden. [Con17b]

Eine Ontologie ist eine formale Spezifikation über eine Konzeptualisierung [SBF98]. Die Denotation der dargestellten Signifikanten lässt sich durch ihre weltweit eindeutigen Präfix identifizieren z.B: PREFIX owl: <http://www.w3.org/2002/07/owl#> [Con17b]. Deren Beziehungen können zu externen Ontologie-signifikanten verweisen und dadurch ein Consensus über Begrifflichkeiten erreichen.

Der Entwurf der Ontologie wurde nach Ontology-Engineering-101 durchgeführt:

Während der Umsetzung wurde Protégé [?] als unterstützende Anwendung benutzt.

Um bereits vorhandene Technologien zu nutzen, wurde ein aktueller öffentlicher graphischer ontologischer Entwurf [RMG14] (siehe Abbildung 6.3), der in Moodle mit einer relationalen Datenbasis und PHP umgesetzt wurde, untersucht.

Eine Implementation dieser Ontologie ist jedoch nicht von Watson [Wat17] und LOD [LOD17] auffindbar.

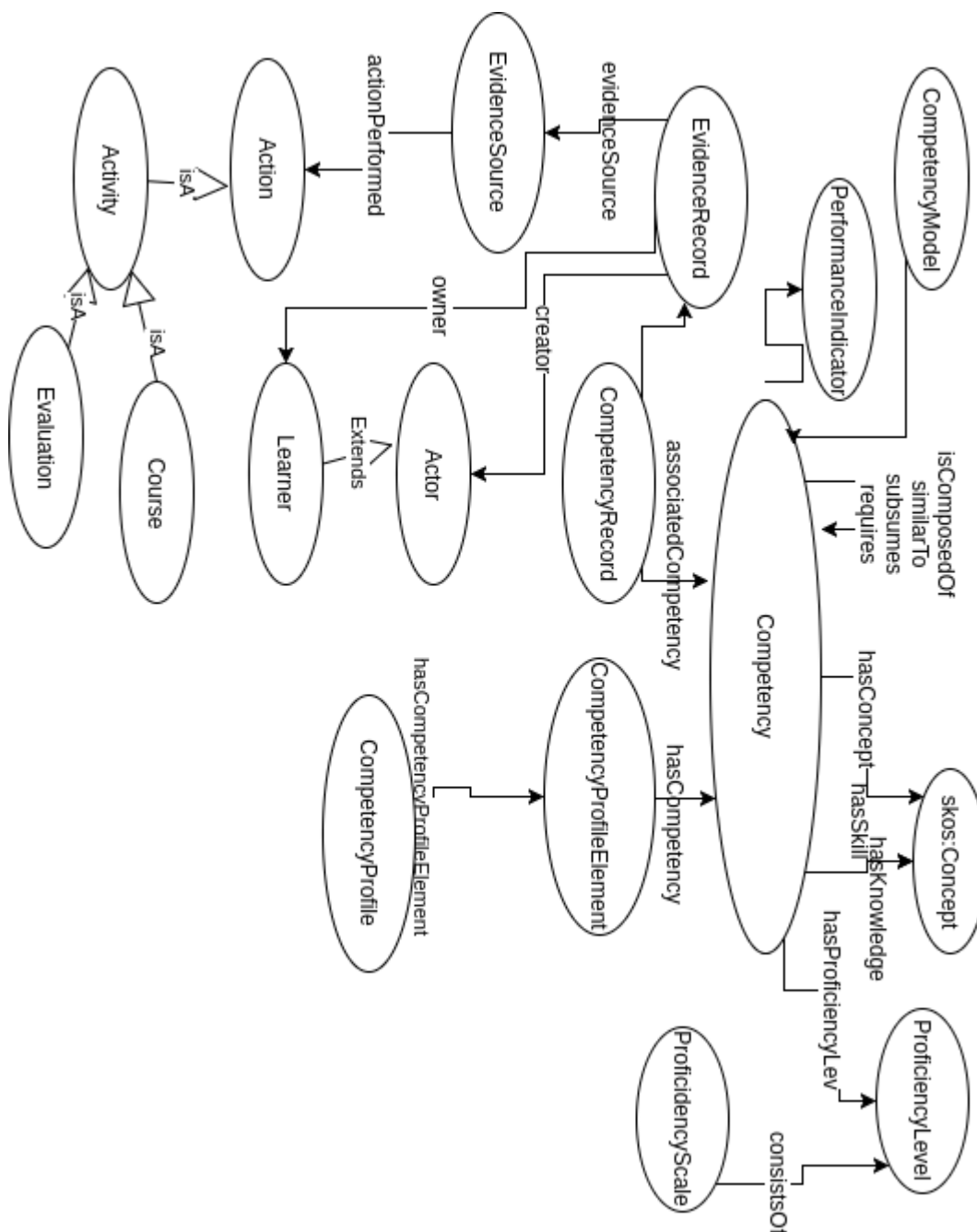


Abbildung 6.3: Kompetenzontologie

Der Entwurf und seine Dokumentation lassen freie Interpretation über Begriffe und deren Zweck, z.B. „isComposedOf“, „subsumes“. Ein Standard zur graphischen Darstellung ist zur Zeit noch nicht anerkannt. Graphische Benutzeroberflächen zur Darstellung und zum Entwurf von Ontologien sind derzeit entwickelt, z.B. Graphol [CLSS14].

Daher folgt eine beispielhafte Erklärung, die auf den Anwendungsfall KOMA eine angepasste und ergänzende Interpretation der dargestellten Terminologie des Entwurfs und der RCD (s.u.) liefert.

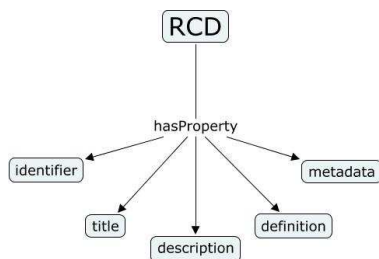


Abbildung 6.4: Reusable Competency Definition

Der EQF wurde für die ontologische Darstellung beschrieben und bietet eine europäisch anerkannte Definition von Kompetenz, die Reusable Competency Definition (RCD) [DCAB17]. Diese ist jedoch nicht in einer veröffentlichten Datenbasis umgesetzt worden.

Die zwei Leitmotive sind auf der einen Seite Kompetenzanforderungen, „die festlegen, über welche Kompetenzen ein Schüler, eine Schülerin verfügen muss, wenn wichtige Ziele der Schule als erreicht gelten sollen.

Systematisch geordnet werden diese Anforderungen in Kompetenzmodellen, die Aspekte, Abstufungen und Entwicklungsverläufe von Kompetenzen darstellen“ [Kli03].

Auf der anderen Seite bildet die Definition von Kompetenz eine handlungsleitende Grundlage. Nach Weinert werden Kompetenzen als „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ [Wei02].

Die bisherige Analyse des Domainproblems wird nun anhand von Protégé in einen Resource Description Framework (RDF) Format bzw. Terse RDF Triple Language (TURTLE) beschrieben. TURTLE besteht aus einer für Menschen lesbaren Syntax und kann sowohl Ontologien in OWL als auch Sparql Abfragen darstellen.

Das folgende Listing 6.1 zeigt eine beispielhafte Darstellung von ontologischen Fakten, den sogenannten „Triples“ . Diese bestehen aus Subjekt, Merkmal und Objekt. Die erste Zeile lässt sich wie folgt interpretieren: dem Namen einer fiktiven Schülerin Alice (Subjekt) wird ein Merkmal in Form einer Kompetenzausprägung (Property) innerhalb eines Unterrichtsfachs, das in Unterrichtsreihen oder -einheiten sequenziert werden kann, wie z.B Mathematik II (Objekt).

Listing 6.1: Darstellung von Triples in TURTLE

```
1 :Alice :hasCompetency :Math_II .
2 :EvidenceRecord rdf:type owl:Class .
3
4 :actionPerformed rdf:type owl:ObjectProperty ;
5   rdfs:domain :EvidenceSource ;
6   rdfs:range :Action .
```

Um aus Ontologien Informationen zu entnehmen, wird die Abfragesprache Protocol And RDF Query Language (Sparql) verwendet. Diese ähnelt der traditioneller SQL.

Die einfachste Abfrage in Sparql wählt alle Triples von dem abgefragten Datenmodell (oder Graph) wie in Listing 6.2 gezeigt wird.

Listing 6.2: Sparql SELECT ALL

```
1 SELECT * WHERE { ?s ?p ?o . }
```

Hilfsvariablen können deklariert werden, um Ergebnisse aus einem Triple als Parameter für das nächste zu benutzen. Die folgende Abfrage ließe sich wie folgt formulieren: „Wähle alle Properties des Graphes und wähle alle dessen Subjekten mit Alice als Objekt“ .

```
1 PREFIX owl: <http://www.w3.org/2002/07/owl#>
2 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
3 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
4 PREFIX koma: <https://s3-us-west-2.amazonaws.com/ontology.thb.de/koma-complex.owl#>
5 SELECT ?x WHERE {
6   ?y rdf:type owl:ObjectProperty .
7   ?x ?y koma:Alice .
8 }
```

Wie in Unterabschnitt 6.0.3 erwähnt wurde, werden Sparqlabfragen durch die API Gateway für Lambda transponiert.

6.2 RESTful API

Der Entwurf von einem benutzerfreundlichen Hypertext Transfer Protocol (HTTP) API beinhaltet die Abstraktion von komplexer Geschäftslogik und Datenverarbeitung in den vier Operationen Create, Read, Update, Delete (CRUD).

Die Komplexität des darunterliegenden Datenmodells erlaubt einer REST Schnittstelle nur einfache Abfragen zu formulieren. [H⁺17] Daher stellt KOMA zusätzlich einen Sparql-Endpunkt für komplexe Sparql Abfragen zur Verfügung. Die letzte Zeile der Tabelle 6.2 ist eine solche komplexe Abfrage.

Tabelle 6.2: RESTful API

Methode	URL	Rückgabe
GET	/ontology	Information über KOMA
GET	/ontology/{individual}	RDF von Individual
GET	/page	Auflistung von Entitäten
GET	/page/{individual}	Information über diesen Fakt
POST	/sparql	Abfragenergebnis

AWS API Gateway ermöglicht die Definition, Konfiguration und das Importieren von Schnittstellen. Beispielsweise kann der Anfrageparameter in `GET https://<host>/page/{individual}` mithilfe des Musters in Listing 6.3 an den Key „individual“ zugewiesen werden. Hier handelt sich um einen „body mapping template“ für den Inhaltstyp (content type) „application/json“ .

Listing 6.3: API Gateway Request Mapping Template

```

1 https://<host>/page/{individual}
2 ...
3 {
4   "individual" : "$input.params('individual')"
5 }
```

Nachdem eine Anfrage eingetreten ist und deren Parameter transponiert sind, erfolgt ein synchroner Aufruf der zuständigen Lambdafunktion. Die Verwaltung dieser Instanziierung entspricht dem Priority Queue Muster in Unterabschnitt 4.1.6, wobei die API Gateway der Queue und die Lambdafunktion dem Consumer entspricht.

Das folgende Listing 6.4 zeigt die Implementierung einer Lambdafunktionsfassade, die für die oben genannte URL zuständig ist.

Listing 6.4: Lambda Javascript Funktionsfassade

```

1 ...
2 exports.handler = function (event, context, callback) {
3   reqIndividual = event.individual; // check possible exception
4   async.waterfall([createBucketParams
5     , getS3ObjectBody
6     , parseOntology
7   ],
8     function (err, result) {
9       if (err) {
10        callback(res.createErrorResponse(500, err));
11      } else {
12        if (Object.keys(result).length === 0 && result.constructor === Object) {
13          callback(null, res.createErrorResponse(404, "there was no result on the
14            search"));
15        } else {
16          callback(null, res.createSuccessResponse(result));
17        }
18      }
19    }
20  );
21 }

```

Die zweite Zeile entspricht der Fassadensignatur von Lambda für Javascript. Der erste Parameter, „event“, enthält in diesem Fall Informationen über die ursprüngliche Anfrage. Der Zweite, „context“ erlaubt den Zugang auf die Laufzeitumgebung von Lambda und kann definieren, ob die Ausführung erfolgreich (`context.success(Object result)`), fehlerhaft (`context.fail(Error error)`) oder beides (`context.done(Error error, Object result)`) war.

In der vierten Zeile befindet sich der Aufruf `async.waterfall(...)`. Dieser entspricht dem Waterfall Muster dessen Verarbeitungsschritte in der Abbildung 6.5 dargestellt werden. Dieses Muster erlaubt es, eine Reihe von Funktionen so zu verketteten, dass das Ergebnis einer Funktion der Eingabeparameter der Nächsten ist. Wenn ein Fehler in einer Funktion entsteht, hält der Wasserfall an.

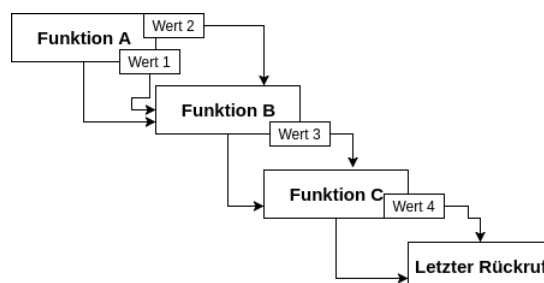


Abbildung 6.5: Waterfall Muster

Um die Wiederverwendbarkeit zu gewährleisten, ist die Abtrennung der Abhängigkeiten

in Bibliotheken der Lambdafunktion sinnvoll. Dieser Ansatz wird von dem SOA-Architekturstil Abschnitt 4 gefordert.

Bei einer Abfrage an die `POST https://<host>/sparql` URL werden die Abfrageparameter ohne die Konfiguration von „body mapping“ weiter an die zuständige Lambdafunktion geleitet.

Die Funktion hinter dieser URL, wie in Listing 6.5 gezeigt, ist in Java 8 implementiert. Sie fordert die Implementierung der Schnittstelle `RequestHandler`. Die Generics-Parameter definieren den Klassentyp der Anfrage und des Ergebnisses, wie es im Fall der Abbildung „RequestClass“ als Anfrage und „String“ als Ergebnis ist.

Der Nutzer kann Umgebungsvariablen bei der Konfiguration von Lambdafunktionen anlegen und während Laufzeit auf sie zugreifen, z.B in der Zeile 7 mittels `System.getenv (StringConstante)`.

Listing 6.5: Lambda Sparql-Endpoint

```
1 public class Handler implements RequestHandler<RequestClass, String> {
2 ...
3 @Override
4 public String handleRequest(RequestClass input, Context context) {
5     context.getLogger();
6     request = input;
7     return new Controller(System.getenv(ENV_BUCKET), Regions.US_WEST_2.getName())
8         .executeQuery(request.getQuery(), request.getBucketKey());
9 }
```

Die Klasse „Controller“ (Zeile 7) benutzt die Jena ARQ Bibliothek, um Sparqlabfragen auszuführen. Ein Beispiel ihrer Benutzung wird in Listing 6.6 dargestellt.

Die Abfrage an die Datenbasis lässt sich mit `request.getQuery()` (Zeile 8) aus dem Bodyparameter „query“ extrahieren.

Listing 6.6: Lambda Sparql-Endpoint-Controller

```
1 ObjectMapper mapper = new ObjectMapper();
2 QueryResultWithMap resultWithMap =
3     new QueryResultWithMap();
4 Map<String, String> tmp;
5
6 Query query = QueryFactory.create(aQuery);
7 try (QueryExecution qexec = QueryExecutionFactory.create(query, model)) {
8
9     ResultSet results = qexec.execSelect();
```

```

10  while (results.hasNext()) {
11      tmp = new LinkedHashMap<>();
12
13      QuerySolution soln = results.nextSolution();
14
15      for(String v : results.getResultVars()){
16          RDFNode node = soln.get(v);
17          tmp.put(v, node.isResource() ?
18              soln.getResource(v).getLocalName() :
19              soln.getLiteral(v).getString());
20      }
21      resultWithMap.getBody().add(tmp);
22
23  }
24  toClient = mapper.writeValueAsString(resultWithMap);
25  ...

```

Nach der Ausführung der Abfrage (Zeile 7) werden die Ergebnisse in Zeile 15 iteriert und anschließend wird eine Referenz zu einem JSON-Objekt zurück geliefert.

Mit der bisher beschriebenen Implementierung der Webanwendung können Daten abgefragt und verarbeitet werden. Diese Ergebnisse stellen ein Feedback dar, womit Konzepte bewiesen (siehe Kapitel 1), ausgewertet und weiterentwickelt werden können.

Ein wesentlicher Merkmal von den oben beschriebenen Lambdafunktionen ist, dass sie für jeden Aufruf die komplette Datenbasis neu herunterladen. Für die Gebrauchstauglichkeit der Anwendung werden im Folgenden Verbesserungsansätze beschrieben.

Caching zielt auf die Verbesserung der Performance und Skalabilität eines Systems, in dem die oft gelesene Daten temporär zwischenspeichert. Daher können wiederholte Zugriffe auf die Ontologie in einer Cache geladen werden. Hierzu bietet AWS „ElastiCache“ als Serverless Dienst.

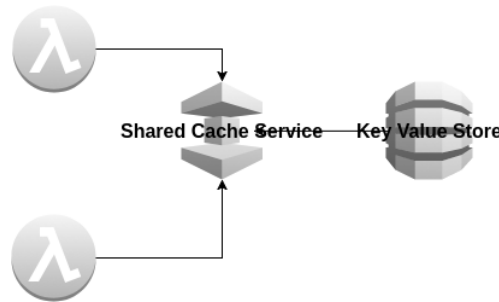


Abbildung 6.6: Shared Cache Muster

Es gibt zwei Typen von Cache: „Shared-Cache“ und „In-Memory“ [HSB⁺14]. Die letztere teilt die Umgebung des Consumers (Lambda in dem Fall) mit. Da die Laufzeitumgebung von Lambda ohne Vorwarnung wechseln kann und nur von nicht persistenten Speicher verfügt [AWS17], lässt sich die In-Memory Cachevariante nicht sicher implementieren.

In dem eine Cache als Puffer-Speicher zwischen der Datenbasis und den Lambdafunktionen (siehe Abbildung 6.6) kann ein schnellerer Zugriff gewährleistet werden. Da die Consumers den gleichen „Snapshot“ der Daten abrufen sinken die inkonsistenten Ergebnisse. [HSB⁺14]

Eine zusätzliche Maßnahme zur Optimierung des Datenzugriffes ist die Datenpartitionierung. Diese kann je nach Benutzungsmuster horizontal, vertikal oder funktional erfolgen.

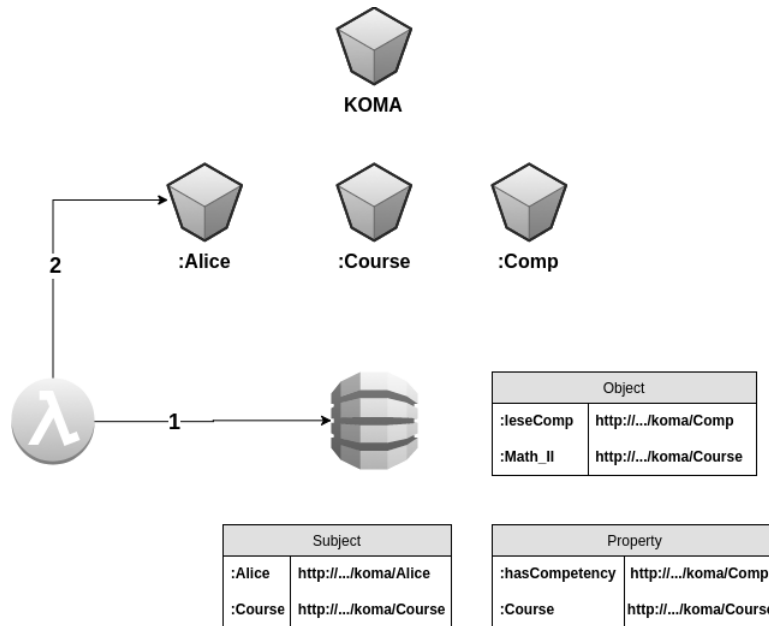


Abbildung 6.7: Partitionierung der Datenbasis

Im folgendem wird die Anwendung als Beispiel zur Erklärung benutzt. Wird die OWL Datei nach „Individuals“ in unterschiedlichen Buckets getrennt, handelt es sich um horizontale Datenpartitionierung. Bei der Vertikalen wird jeweils eine Tabelle für Instanzen (Individuals) von Subjekten, Merkmalen (Properties) und von Objekten aufgestellt. Dafür eignet sich DynamoDB als „Key-Value Store“ um für Keys die Individuals und für Values dessen Universal Ressource Identifier (URI)s darzustellen. [AMMH07]

Mit der Benutzung der Anwendung werden voraussichtlich die OWL Dateien vergrößert und somit die Latenz bei Abfragen verlängert.

Cloud Dienste werden oft in unterschiedlichen Datazentren oder Regionen deployed. Um die Verfügbarkeit, Performance und Konsistenz zu maximiereun und die Datenübertragungskosten zu minimieren, kann eine Master-Datenbasis mit erlaubte Read und Write Operationen und eine Subordinate-Datenbasis mit nur Read Operationen definiert werden.

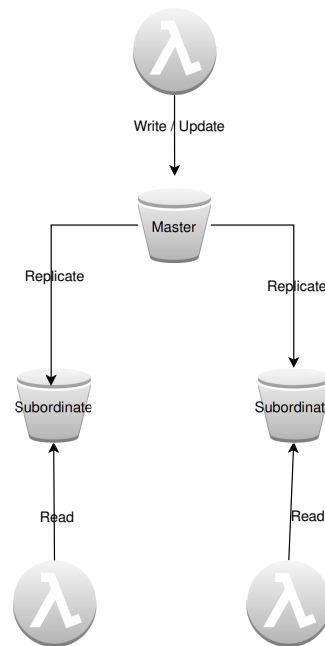
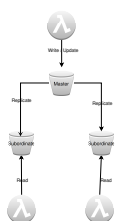


Abbildung 6.8: Replizieren der Datenbasis

Der Master pusht oder (one-way) synchronisiert die Subordinate-Replikas. Diese Aufteilung favorisiert Read Operationen.

Mit der Benutzung der Anwendung werden voraussichtlich die OWL Dateien vergrößert und somit die Latenz bei Abfragen verlängert.

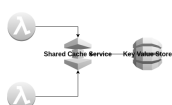
6.2.1 Datenhaltung Synchronisierung und Replizieren



Master - Subordinate Cloud Dienste werden oft in unterschiedlichen Datazentren oder Regionen deployed. Um die Verfügbarkeit, Performance und Konsistenz zu maximieren und die Datenübertragungskosten zu minimieren, kann eine Master-Datenbasis mit erlaubte Read und Write Operationen und eine Subordinate-Datenbasis mit nur Read Operationen definiert werden.

Der Master pusht oder (one-way) synchronisiert die Subordinate-Replikas.

Diese Aufteilung favorisiert Read Operationen.

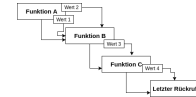


cache shared Caching zielt auf die Verbesserung der Performance und Skalabilität eines Systems, in dem die oft gelesene Daten temporär zwischenspeichert. Es gibt zwei Typen von Cache: In-Memory und

Shared-Cache. Die zustandslose Natur der Lambda kompliziert die Benutzung der ersten Cachevariante. Die Shared-Cache hilft das Problem zu beheben, bei Inkonsistenzen unter verteilten Caches.

6.2.2 Datenverarbeitung

Waterfall Dieser Muster erlaubt eine Menge von Funktionen so zu verketteten, dass das Ergebnis der Einer als Eingabeparameter der Nächsten ist. Wenn ein Fehler in eine Funktion entsteht, hält das Waterfall an.



6.2.3 Abtrennung des Monoliths

Im Folgendem es wird beispielhaft eine abtrennung einer JEE Anwendung.

6.2.4 Single Page Application

Da KOMA ohne Vorkenntnisse gebrauchsfertig sein soll, mit dem Fakt dass Milliarden von Desktop Geräte die Web mit einem Browser erkundigen können, lässt sich die Entscheidung über die Art der Benutzeroberfläche leicht Treffen.

Die Web Anwendung ist für alle Rechenaufgaben verantwortlich die im Browser aus dem Sicherheitssichtpunkt kein Gefahr darstellen, um den Backend oder Servers möglichst wenig auszulasten. Deswegen bietet sich eine Single Page Application an. Die SPA besteht aus ein einziges HTML Dokument. Dies vereinfacht man die Konfiguration der Authentifizierung und unterbricht den Fluss der UI-Darstellung zwischen Seiten.

Ein konfiguriertes Anfangsprojekt/Quickstart kann mithilfe von Initializr [Ini17] oder JHipster [Jhi17]. Für die lokale Entwicklung der Webseite werden anhand von NodeJS und NPM folgende Bibliotheken als Abhängigkeiten verwaltet: Bootstrap als Stylesheet und jQuery als Javascript-Bibliothek. Die Webseite wird Statisch mittels S3 geliefert. Dies geschieht mit einem Befehl:

Listing 6.7: Webseite veröffentlichen

```
1 $ aws --region us-west-2 s3 website --index-document index.html --error-document error.
   html 's3://koma.thb.de'
```



Da der Zugriff auf die Datenspeicherung gesichert werden soll, wird die Login-Funktionalität hinzugefügt.

7 Bewertung

Regeln für Anfragen Der Zugang auf die Schnittstelle wird durch CORS konfiguriert um deren Ausnutzung zu vermeiden.

Zur Skalabilität Skalabilität in Datenhaltung -> Entwerfe für Distribution + Vorteile von Lokalität. Read replikas -> evtl. Konsistent. Viele Perspektiven von Daten -> Lebenszyklus von Daten. — Bounded Context — Service ist nicht nur Funktion oder nur DB. Entkopplung fordert enkapsulation und Cohesion.

Event als Businessmanager -> Coordinator / Orchestrator -> Lambda Als Finite State Machine oder Workflow

Anwendung Latenz Die entstandene Webanwendung befindet sich in US-WEST-2, Oregon, in den USA. Da keine Cache oder CDN Funktionalität weder Implementiert noch konfiguriert ist, ist die Latenz direkt proportional zur Ausführungsdauer der Lambda Funktion. @Benchmark testing curl @Lambda Monitoring

In Zeiten des Cloud computings

Frameworks und FaaS Frameworks helfen aber sind platform abhängig. Entweder JEE und JVM oder PHP. Es kann auf die Layer of Abstraction in FW verzichtet werden. Die Ersetzbarkeit des FaaS entkoppelt die Anwendung und den Entwickler von der darunterliegende Technologie.

DevOps Frameworks Die benötigte Fertigkeiten für die Umsetzung einer Serverless Anwendung werden mithilfe von Deploymentframeworks gemindert. Die Aufnahme von 3.Anbieter ist deswegen notwendig. Es existieren bereits solche Hilfe wie z.B Serverlessframework@Ref

Risiko: Entwickler brauchen einen guten Testplan und eine gute DevOps Strategie. <- skills shortage

Transaktionen Transaktionen können nicht parallel ausgeführt werden. Sequenziell aka Messaging Pattern. Zusammenspiel Arch. interfaces prog.modell und FW Arch
1st -> def interfaces and interactions. to program to a interface

Eventual consistency -> event driven + ontology quality Consistency -> kommandostandalone <- transaction mgn

Vorteile Automatische Skalierung <!-- große und kleine Apps -> und Fehlertoleranz
Automatisches Kapazitätsmanagement Flexible Ressourcenverwaltung Schnelle Bereitstellung der Ressourcen Exakte nutzungsabhängige Abrechnung der Ressourcen
Konzentration auf den Kern des Source-Codes

Nachteile SLA Service Level Agreement: Latency, Bank:High volume Transactions, Decentralisation of Services = Challenge = Overhead, time, energy <- orchestration of events. Decentralisation vs monolithik != -komplexity Kontrollverlust Erhöhtes Lock-in Risiko

kurzlebige konfigurationen herausfinden ?? tracking? viel Konfiguration, kaum Konvention -> .json 4 everything local testing braucht event-simulation.json

Zur Entwicklung Die Starke Komponentisierung und Dezentralisierung von Software, die Variabilität von Programmiermodellen, Frameworks, Tools, -Sprachen und dessen Entwicklungsumgebung erhöht die Komplexität des Entwicklungszyklus und hervorhebt die Bedürfnis von Tools zur Automatisierung von Tests, Deployment und Konfiguration. Also ein wohldefiniertes Handlungsplan bei der Softwareentwicklung dass von der nicht zu bearbeitende Details abstrahiert. Die DevOps Kultur spricht solche Probleme an. Neben dem Entwurf der Softwarearchitektur muss, um derer Umsetzung Zeitgemäß zu gewährleisten, eine zum Projekt passende DevOps Strategie. Um Vorteil von der neuen Technologien zu nehmen, ist die Recherche nach schon existierenden DevOps Frameworks besonders wichtig. Dessen Integration in der DevOps Strategie diene für eine Agile Entwicklung.



Zum Datenmodel Aus der Anforderungs analyse einer Informations Technologie Web Anwendung sind die Builder, Texte und dessen Darstellung das ergebniss, dass ohne Daten inhaltlos wäre. Auf eine Seite Das Relationale Datenschema stellt keine Semantik für sich dar, sondern durch von der Software entstandene Verknüpfung zwischen dem Endergebnis und dem Datenschema. Auf der Anderen Seite die RDF Daten einer Ontologie *is* das Modell.

Zum API Gateway Bei Frameworks wie JEE werden Schnittstellen zwischen Layers und Tiers bereitgestellt und diese am Laufzeit entdeckt aka Service Discovery. Im Fall der API Gateway wird die Kopplung bei derer Konfiguration festgelegt wo derer Rekonfiguration ein neues Deployment ohne Downtime bedeutet. Die Der Quellcode der Dienste bleiben unberührt und kein Load Balancer muss rekonfiguriert werden.

Zum Serverless In dieser Arbeit wurde eine "nach buch"weise die Architektur gestaltet. Die unterschiedliche Interpretationen des Begriffs Serverless kann auch zu Kreativen anzätzen führen@AdamBien JEE Es kann daher auch als Serverless betrachtet wenn neue Quelldataien eine Docker Instance neu Erzeugen oder nur Updaten, dessen LoadBalancing auch als Serverless Quellcode verpakt werden kann.

Zur Annahme dass Quellcode schneller entwickelt werden kann, wenn der Entwickler sich nur damit beschäftigt.

8 Ausblick

RESTful UI RESTful Anfragen für bestimmte UI formate.

Listings

6.1	Darstellung von Triples in TURTLE	29
6.2	Sparql SELECT ALL	30
6.3	API Gateway Request Mapping Template	31
6.4	Lambda Javascript Funktionsfassade	31
6.5	Lambda Sparql-Endpoint	33
6.6	Lambda Sparql-Endpoint-Controller	33
6.7	Webseite veröffentlichen	38

Tabellenverzeichnis

6.1	Vergleich relationalem mit ontologischem 6.1.1 Schema	27
6.2	RESTful API	31

Abkürzungen

GC Garbage Collection

„Garbage Collection“ bezeichnet die automatische Speicherwaltung zur Minimierung des Speicherbedarfes eines Programmes. Garbage Collection (GC) wird zur Laufzeit durch Identifikation von nicht mehr benötigten Speicherbereichen ausgeführt. Im Vergleich zur manuellen Speicherverwaltung benötigt GC mehr Ressourcen.

Literaturverzeichnis

- [AMMH07] Daniel J. Abadi, Adam Marcus, Samuel R. Madden, and Kate Hollenbach. Scalable semantic web data management using vertical partitioning. In *Proceedings of the 33rd International Conference on Very Large Data Bases*, VLDB '07, pages 411–422. VLDB Endowment, 2007.
- [AWS17] AWS. Aws lambda – häufig gestellte fragen, 2017.
- [Bob13] DuCharme Bob. Learning sparql. sl, 2013.
- [Cha14] K Chandrasekaran. *Essentials of cloud computing*. CRC Press, 2014.
- [CLSS14] Marco Console, Domenico Lembo, Valerio Santarelli, and Domenico Fabio Savo. *Graphical Representation of OWL 2 Ontologies through Graphol*, volume 1272. 10 2014.
- [Con17a] ConnectWise. Connectwise, 2017.
- [Con17b] World Wide Web Consortium. World wide web consortium, 2017.
- [DCAB17] Diego Duran, Gabriel Chanchí, Jose Luis Arciniegas, and Sandra Baldassarri. A semantic recommender system for idtv based on educational competencies. In *Applications and Usability of Interactive TV*. Springer, January 2017.
- [EQF17] EQF. Eqf, 2017.
- [Fou17] The Apache Software Foundation. Arq - a sparql processor for jena, 2017.
- [Fow17a] M Fowler. Serverless architectures, 2017.
- [Fow17b] M Fowler. Serverless architectures, 2017.
- [Goo17a] Google. Docker, 2017.
- [Goo17b] Google. Docker monitoring, 2017.

- [GOS09] Nicola Guarino, Daniel Oberle, and Steffen Staab. What is an ontology? In *Handbook on ontologies*, pages 1–17. Springer, 2009.
- [H⁺17] II Hunter et al. Advanced microservices: A hands-on approach to micro-service infrastructure and tooling. 2017.
- [HSB⁺14] A. Homer, J. Sharp, L. Brader, M. Narumoto, and T. Swanson. *Cloud Design Patterns: Prescriptive Architecture Guidance for Cloud Applications*. Patterns & practices. Microsoft Developer Guidance, 2014.
- [HW04] Gregor Hohpe and Bobby Woolf. *Enterprise integration patterns: Designing, building, and deploying messaging solutions*. Addison-Wesley Professional, 2004.
- [Ini17] Initializr. Initializr, 2017.
- [Jhi17] Jhipster. Jhipster, 2017.
- [Kin16] W. King. *AWS Lambda: The Complete Guide to Serverless Microservices - Learn Everything You Need to Know about AWS Lambda!* AWS Lambda for Beginners, Serverless Microservices Series. CreateSpace Independent Publishing Platform, 2016.
- [Kli03] Eckhard Klieme. ua: Zur entwicklung nationaler bildungsstandards–eine expertise. *Berlin 2003*, 2003.
- [LOD17] LOD. Lod, 2017.
- [Net17] Netflix, 2017.
- [Ray13] Nilanjan Raychaudhuri. *Scala in action*. Manning Publications Co., 2013.
- [RB] Jürgen Dunkel Ralf Bruns. *Event-Driven Architecture: Softwarearchitektur für ereignisgesteuerte Geschäftsprozesse*.
- [RMG14] Kalthoum Rezgui, Hédia Mhiri, and Khaled Ghédira. Extending moodle functionalities with ontology-based competency management. *Procedia Computer Science*, 35:570–579, 2014.
- [SA17] Sachsen-Anhalt. Sachsen-anhalt, 2017.
- [Sba17] P. Sbarski. *Serverless Architectures on AWS: With Examples Using AWS Lambda*. Manning Publications Company, 2017.

- [SBF98] Rudi Studer, V Richard Benjamins, and Dieter Fensel. Knowledge engineering: principles and methods. *Data & knowledge engineering*, 25(1-2):161–197, 1998.
- [Sta07] Gernot Starke. *SOA-Expertenwissen: Methoden, Konzepte und Praxis serviceorientierter Architekturen*. dpunkt, 2007.
- [UNL17] UNLESS, 2017.
- [Wat17] Watson. Watson, 2017.
- [Wei02] F.E. Weinert. *Leistungsmessungen in Schulen*. Beltz Pädagogik. Beltz, 2002.
- [You15] Marcus Young. *Implementing Cloud Design Patterns for AWS*. Packt Publishing Ltd, 2015.