

UNIVERSITÄT LEIPZIG

Institut für Informatik – Abteilung Datenbanken

Cloud-Infrastrukturen

Seminararbeit

Leipzig, 04.01.2010

vorgelegt von:

Richard Beyer

geb. am: 30.01.1981

Betreuer:

Dr. Thor / Dipl. Ing. Aumüller

Inhaltsverzeichnis

1	Einleitung	1
1.1	Gegenstand	1
1.2	Problematik	1
1.3	Motivation	1
1.4	Zielsetzung	2
2	Grundlagen	2
3	Amazon Web Services	4
3.1	Allgemeines.....	4
3.2	Amazon Elastic Compute Cloud (Amazon EC2).....	4
3.2.1	EC2 Merkmale.....	5
3.2.2	EC2 Dienste.....	5
3.2.3	Instanztypen und Preise	6
3.3	Amazon Simple Storage Service (S3).....	9
3.4	Weitere AWS-Dienste (Auswahl).....	10
4	Yahoo! Cloud Computing	12
4.1	Allgemeines.....	12
4.2	Anforderung an die Cloud seitens Yahoo!	12
4.3	Architektur der Yahoo!-Cloud	13
4.4	Bestandteile der Yahoo!-Cloud (Auswahl)	14
5	Zusammenfassung und Vergleich	16
6	Diskussion	18

Begriffs- und Abkürzungsverzeichnis

AWS	Amazon W eb S ervices
AMI	Amazon M achine I mage
NAS	N etwork A ttached S torage
SAN	S torage A rea N etwork
SQS	Amazon S imple Q ueue S ervice
VPN	V irtual P rivate N etwork

1 Einleitung

1.1 Gegenstand

In den letzten Jahren wird über Cloud Computing in den Medien und in der Fachliteratur immer öfter als Zukunft der Computertechnik berichtet. Die Möglichkeit, mit Hilfe mehrerer Rechner gleichzeitig viel komplexere Berechnungen in kürzerer Zeit durchzuführen als es bislang möglich war, bietet zum Beispiel Forschern die Chance, sehr große Mengen an Daten verarbeiten zu können und somit schneller und effizienter Forschungsergebnisse zu erhalten. Doch auch für private Anwender oder kleine Firmen bietet Cloud Computing interessante Ansätze. Um sein Geld nicht an der falschen Stelle zu investieren, muss der Anwender allerdings die verschiedenen Architekturstile der verschiedenen Cloudplattformen kennen, um die optimale Plattform für sein Problem nutzen zu können.

1.2 Problematik

Mittlerweile existieren verschiedene Cloud-Computing-Anbieter, welche sich allerdings in ihrer Architektur, Funktionsweise und Bezahlmodell stark unterscheiden. Eine der grundlegenden Unterschiede der Anbieter zeigt sich im allgemeinen Aufbau, der Infrastruktur der Plattformen. Wenn der Einsatz von Cloud-Computing für ein Projekt geplant wird, so ist die Kenntnis der Plattforminfrastruktur von entscheidender Bedeutung, um keine finanziellen und zeitlichen Ressourcen zu verschwenden.

1.3 Motivation

Diese Seminararbeit soll den Teilnehmern des Seminars „Cloud Data Management“ und weiteren Lesern einen Überblick über die Architektur wichtiger Anbieter von Cloud-Diensten geben und damit die Grundlage bilden, mit der weiterführende komplexere Themen dieses Fachgebiets behandelt werden können.

1.4 Zielsetzung

Es sollen zwei wichtige Vertreter des Cloud-Computings vorgestellt werden, deren Infrastruktur sich offenbar stark voneinander unterscheidet. Dabei werden

1. Amazons Web Service EC2 und
2. die Yahoo! Cloud

beschrieben und miteinander verglichen.

Weitere Cloud-Anbieter wie Google oder Salesforce werden in [Hütter 2010] näher betrachtet. Da diese andere Ansätze verfolgen, sollen sie in dieser Arbeit nicht näher betrachtet werden.

2 Grundlagen

Aufgrund der Neuartigkeit des Cloud Computing existieren mehrere Definitionen des Begriffs, wobei keine davon bislang als allgemeingültig anerkannt wurde. Die einfachste Beschreibung des Begriffs findet sich vielleicht in dieser Definition von Lewis Cunningham:

“Cloud computing is using the internet to access someone else's software running on someone else's hardware in someone else's data center.”

Cloud Computing wird darin als ein Konzept betrachtet, welches den Betrieb, Bereitstellung, Support und Management von IT als eine Dienstleistung darstellt. Dabei werden Hardware und Software nicht mehr vom Anwender selbst betrieben, sondern durch einen Anbieter im Internet oder in einem Netzwerk bereitgestellt.

Eine weitere mittlerweile etablierte Form der Darstellung ist die Verwendung eines 3-Schichten-Modells, welche die unterschiedlichen Typen von Clouds aufzeigt:

Anwendung

Plattform

Infrastruktur

- 1) **Infrastruktur:** Die unterste Schicht repräsentiert Anbieter von Cloud-Diensten, bei denen man eine Umgebung aus virtuellen Servern zur Verfügung hat und diese sich sehr frei skalieren lassen. Je nach Anforderung können Ressourcen bereitgestellt oder

wieder freigegeben werden. Der Anwender hat dabei die Möglichkeit, die Server selbst zu administrieren und eigene Anwendungen aufzuspielen. Ein Beispiel dafür ist Amazon EC2.

- 2) Plattform: Dienste auf der Plattformschicht werden beispielsweise von Google unter dem Namen App-Engine angeboten und ermöglichen dem Anwender den Betrieb eigener erstellter Software auf der bereitgestellten Hardware. In diesem Fall hat der Kunde keinen direkten Einfluss auf die Administration der Server, beziehungsweise die Aufteilung oder Zuweisung von Ressourcen an bestimmte Lastpunkte. Die Administration wird komplett vom Anbieter übernommen.
- 3) Anwendung: Dieser Dienst wird auch als Software-as-a-Service bezeichnet, da dem Anwender „nur“ die reine Anwendung zur Verfügung gestellt wird, mit deren Hilfe er die Vorzüge des Cloud-Computings nutzen kann. Veränderungen an Hardware oder Software seitens des Kunden sind hierbei meist ausgeschlossen. Als Beispiel dafür ist Microsofts Sharepoint Online zu nennen.

Ein weiteres allgemeines Unterscheidungskriterium bilden die Begriffe

- On-Demand-Computing-Instances und
 - On-Demand-Computing-Capacity.
- 1) On-Demand-Computing-Instances bezeichnet dabei die Möglichkeit, Instanzen von einer Gesamtheit an Rechnerkapazität zu kaufen oder zu mieten und bildet damit eine Schnittmenge der Infrastruktur- und Plattformschicht der oben genannten Unterteilung.
 - 2) On-Demand-Computing-Capacity hingegen beschreibt das Mittel des Kaufs oder Mietens von Rechen- oder Datenkapazität eines Anbieters zur Unterstützung besonders leistungsintensiver Algorithmen. Ein Beispiel dafür ist die Yahoo! Cloud.

In den folgenden Abschnitten der Arbeit wird Amazon EC2 als Anbieter von On-Demand-Computing-Instances und die Yahoo! Cloud als Anbieter von On-Demand-Computing-Capacity vorgestellt und miteinander verglichen.

3 Amazon Web Services

3.1 Allgemeines

Die angebotenen Cloud-Dienste von Amazon werden unter dem allgemeinen Begriff *Amazon Web Services* zusammengefasst. Diese beinhalten nicht nur die Möglichkeit Daten- und Rechenkapazität zu mieten oder zu kaufen, sondern bieten auch spezielle Extradienste an, mit deren Hilfe die gemietete Cloud-Struktur noch erweitert, beziehungsweise verbessert werden kann.

3.2 Amazon Elastic Compute Cloud (Amazon EC2)

Amazon selbst ([AWS EC2 2009]) beschreibt sein Angebot wie folgt:

„Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud.“

Wenn man die Definition aus Abschnitt 2 allerdings zu Hilfe nimmt, müsste diese Beschreibung leicht geändert werden, denn man kann nicht nur reine Kapazität mieten, sondern kann mit EC2 ganze Instanzen von virtuellen Rechenmaschinen mieten. Dabei sind den Kunden viele Optionen zur Gestaltung dieser Instanzen freigestellt, er kann zum Beispiel von Rechenleistung über Speicher, Speicherplatz und Zugriffsgeschwindigkeit frei wählen. Die Grenze dabei bildet dabei der Preis, jeder Dienst und jedes Extra wird einzeln berechnet (siehe dazu Abschnitt 3.2.3).

Um mit EC2 starten zu können, wählt man als User über den Browser aus voreingestellten Images eine sogenannte AMI, ein Amazon machine Image, aus. Dieses beinhaltet Betriebssystem und Anwendungen. Eine andere Möglichkeit besteht darin, zum Beispiel aus bestehenden Serverkonfigurationen von Windows 2003 ein AMI zu erstellen und dieses bei Amazon S3 (dem Speicherdienst von Amazon) hochzuladen. Nach Konfiguration von Netzwerkdiensten und bestimmten Sicherheitsoptionen wählt man bei AWS die Instanztypen aus, auf denen das Image laufen soll und hinterlegt seine Abrechnungsinformationen. Dann ist das Image startbereit und kann in Amazons Cloud gestartet und genutzt werden. Kosten entstehen nur bei entsprechendem Betrieb der AMI und bestimmten gewählten Zusatzoptionen.

3.2.1 EC2 Merkmale

Amazon selbst definiert bestimmte Anforderungen und Merkmale die AWS an seine Kunden richten soll, darunter folgende:

- **Elastisch:** EC2 Instanzen sollen innerhalb von Minuten umkonfigurierbar sein können, sowohl vom User selbst per Hand, als auch automatisch nach Leistungsbedarf.
- **Komplett kontrollierbar:** Instanzen können vom User selbstständig gestartet, beendet, modifiziert und ausgelesen werden, als würde man mit einen lokalen Rechner arbeiten.
- **AWS Design:** Alle AWS Dienste sind darauf ausgelegt, dass sämtliche Komponenten aufeinander zugreifen und mit den gemeinsamen Daten arbeiten können.
- **Zuverlässigkeit:** Amazon garantiert 99.95%ige Verfügbarkeit der Server in der jeweiligen Zugriffsregion.
- **Sicherheit:** Amazon stellt verschiedene Techniken zur Erhöhung der Sicherheit der EC2-Maschinen zur Verfügung (siehe [AWS SEC 2009]), darunter Firewalls, begrenzte IP-Bereiche, Backups, begrenzten physischen Zugang und VPNs.
- **Preiswert:** Amazon verfolgt das Modell, dass man nur für das bezahlt, was man wirklich verbraucht, das heißt, man zahlt für Speicherplatz den man mietet und sekundengenau für die verbrauchte Zeit an Rechenleistung die man genutzt hat.

3.2.2 EC2 Dienste

Folgende Dienste werden von Amazon zusätzlich gewährleistet, um einen performanten und reibungslosen Betrieb der Cloud zu gewährleisten:

- **Multiple Hosting-Zonen:** Amazon speichert Instanzen und Daten in mehreren geografischen Regionen um Latenzzeiten zu minimieren. Des Weiteren kann man Instanzen mehrfach in verschiedenen Regionen starten, um im Falle eines Ausfalls ein bereits laufendes Backupsystem zur Verfügung zu haben.

- **Elastische IP Adressen:** Statische IP-Adressen werden nicht mit einem Server, sondern mit einem AWS-Account verknüpft und können daher für alle Instanzen benutzt werden, die man startet.
- **Amazon Virtual Private Clouds:** Über eine VPN-Verbindung können damit bereits existierende EC2-Instanzen in ein Firmennetz mit eingebunden werden und erscheinen dadurch im internen Firmennetz als lokale Netzteilnehmer.
- **Cloudwatch, Loadbalancing, AutoScaling:** Amazon stellt des Weiteren Tools zur Verfügung, um die Auslastung der gemieteten Instanzen überwachen zu können. Die Automatische Lastverteilung vergibt bei eingehenden Daten die Aufträge automatisch an die Instanzen mit der geringsten Last. Übersteigt die Last eine angegebene Grenze, so kann per AutoScaling die gemietete Leistung nach definierten Regeln automatisch angehoben werden, bis die Lastspitze verarbeitet wurde.

3.2.3 Instanztypen und Preise

Grundsätzlich werden drei verschiedene Arten von Instanzen angeboten, welche sich in ein Standard-, ein High-Memory- und ein High-CPU-Angebot unterteilen. Folgende Tabelle zeigt die aktuell unterschiedenen Varianten:

<u>Standard</u>	<u>High-Memory</u>	<u>High-CPU</u>
Small Instance 1.7 GB memory 1 EC2 Compute Unit 160 GB instance storage 32-bit platform I/O Performance: Moderate	High-Memory Double Extra Large Instance 34.2 GB of memory 13 EC2 Compute Units 850 GB of instance storage 64-bit platform I/O Performance: High	High-CPU Medium Instance 1.7 GB of memory 5 EC2 Compute Units 350 GB of instance storage 32-bit platform I/O Performance: Moderate
Large Instance 7.5 GB memory 4 EC2 Compute Units 850 GB instance storage 64-bit platform I/O Performance: High	High-Memory Quadruple Extra Large Instance 68.4 GB of memory 26 EC2 Compute Units 1690 GB of instance storage 64-bit platform I/O Performance: High	High-CPU Extra Large Instance 7 GB of memory 20 EC2 Compute Units 1690 GB of instance storage 64-bit platform I/O Performance: High
Extra Large Instance 15 GB memory 8 EC2 Compute 1,690 GB instance storage 64-bit platform I/O Performance: High		

Zu beachten ist die Angabe der Rechenkapazität in EC2-Compute-Units, was einer relativen Leistungsangabe folgender Spezifikation gleicht:

1 EC2 Compute Unit = 1.0-1.2 GHz 2007 Opteron oder 2007 Xeon Prozessor

Folgende Software kann mit jeder AMI genutzt werden:

➤ **Betriebssysteme**

Red Hat Enterprise Linux	Windows Server 2003/2008	Oracle Enterprise Linux
OpenSolaris	openSUSE Linux	Ubuntu Linux
Fedora	Gentoo Linux	Debian

➤ **Software**

Datenbanken	Batch Verarbeitung	Webhosting
IBM DB2	Hadoop	Apache HTTP
IBM Informix Dynamic Server	Condor	IIS/Asp.Net
Microsoft SQL Server Standard 2005	Open MPI	IBM Lotus Web Content Management
MySQL Enterprise		IBM WebSphere Portal Server
Oracle Database 11g		

Anwendungsentwicklung	Anwendungsserver	Video Encoding
IBM sMash	IBM WebSphere Application Server	Wowza Media Server Pro
JBoss Enterprise Application Platform	Java Application Server	Windows Media Server
Ruby on Rails	Oracle WebLogic Server	

Amazon berechnet zum Einen gemietete Instanzen die man bei Bedarf nutzen kann und ansonsten abgeschaltet lassen kann. Dies kostet etwas mehr im Nutzungsfall, bedeutet aber, dass man keinerlei längerfristige Verpflichtungen mit Amazon eingehen muss. Zum Anderen bietet Amazon die Möglichkeit, gegen eine Jahresgebühr (1 oder 3 Jahre im Voraus) sich Instanzen „reservieren“ zu lassen. Dies schlägt sich im Resultat positiv auf die stündlichen Nutzungspreise nieder, wie die beiden folgenden Tabellen belegen:

US – N. Virginia	US – N. California	EU – Ireland
Standard On-Demand Instances		
	Linux/UNIX Usage	Windows Usage
Small (Default)	\$0.095 per hour	\$0.12 per hour
Large	\$0.38 per hour	\$0.48 per hour
Extra Large	\$0.76 per hour	\$0.96 per hour
High-Memory On-Demand Instances		
	Linux/UNIX Usage	Windows Usage
Double Extra Large	\$1.34 per hour	\$1.44 per hour
Quadruple Extra Large	\$2.68 per hour	\$2.88 per hour
High-CPU On-Demand Instances		
	Linux/UNIX Usage	Windows Usage
Medium	\$0.19 per hour	\$0.29 per hour
Extra Large	\$0.76 per hour	\$1.16 per hour

On-Demand Instanzen (aus [AWS EC2 2009])

US – N. Virginia	US – N. California	EU – Ireland	
Linux/UNIX			
One-time Fee			
Standard Reserved Instances	1 yr Term	3 yr Term	Usage
Small (Default)	\$227.50	\$350	\$0.04 per hour
Large	\$910	\$1400	\$0.16 per hour
Extra Large	\$1820	\$2800	\$0.32 per hour
High-Memory Reserved Instances	1 yr Term	3 yr Term	Usage
Double Extra Large	\$3185	\$4900	\$0.56 per hour
Quadruple Extra Large	\$6370	\$9800	\$1.12 per hour
High-CPU Reserved Instances	1 yr Term	3 yr Term	Usage
Medium	\$455	\$700	\$0.08 per hour
Extra Large	\$1820	\$2800	\$0.32 per hour

Reservierte Instanzen (aus [AWS EC2 2009])

Weitere Kosten fallen hauptsächlich für die Nutzung des Speichersystems Amazon S3, Datentransferkosten und für optionale Zusatzfeatures an, welche man unter [AWS EC2 2009] nachschlagen kann.

3.3 Amazon Simple Storage Service (S3)

Ein wichtiger Bestandteil des Amazon Web Services ist der *Amazon Simple Storage Service*, auch kurz Amazon S3 genannt. Dieser erfüllt die Funktion eines Web-Dateihosting-Services, welcher über eine Web-Schnittstelle sehr einfach angesprochen werden kann. Die Dateien können in sogenannten *Buckets* abgelegt werden, was eine Art Verzeichnis durch ein URL-Präfix der S3-Webseite darstellt. Zum Beispiel könnte man Bilder und Textdaten des Seminars „Cloud Data Management“ im Bucket *clouddatamanagement* (sofern denn bei Amazon gekauft) unter der folgenden Adresse abrufen:

http://clouddatamanagement.s3.amazonaws.com/DATEI-NAME

Pro Account bei AWS sind 100 Buckets erlaubt, die Dateien selbst können zwischen 1B und 5GB groß sein, der Name des Buckets muss „amazonweit“ einen eindeutigen Namen besitzen. Des Weiteren hat man die Auswahl, die Daten in den USA oder in Europa hosten zu lassen, was sich in Zugriffsgeschwindigkeit und leicht unterschiedlichen Preisen auswirkt.

Folgende Tabelle zeigt eine aktuelle Kostenübersicht:

US – Standard		US – N. California		EU – Ireland	
Storage		Data Transfer		Requests	
Tier	Pricing	Tier	Pricing	Tier	Pricing
First 50 TB / Month of Storage Used	\$0.150 per GB	All Data Transfer In	Free until June 30th, 2010*	PUT, COPY, POST, or LIST	\$0.01 per 1,000 Requests
Next 50 TB / Month of Storage Used	\$0.140 per GB	First 10 TB / Month Data Transfer Out	\$0.170 per GB	GET and All Other Requests*	\$0.01 per 10,000 Requests
Next 400 TB / Month of Storage Used	\$0.130 per GB	Next 40 TB / Month Data Transfer Out	\$0.130 per GB	* No charge for delete requests	
Next 500 TB Storage Used	\$0.105 per GB	Next 100 TB / Month Data Transfer Out	\$0.110 per GB		
Next 4000 TB Storage Used	\$.080 per GB	Data Transfer Out / Month Over 150 TB	\$0.100 per GB		
Storage used over 5000 TB	\$.055 per GB	* Data Transfer In will be \$0.100 per GB after June 30th, 2010			

Preise Amazon S3 in Europa (aus [AWS EC2 2009])

3.4 Weitere AWS-Dienste (Auswahl)

Jeder dieser Dienste kann gegen einen geringen Aufpreis zu einer existierenden EC2-Instanz hinzu gekauft werden, meist ist ein S3-Bucket notwendig, um anfallende Daten speichern oder verarbeiten zu können. Dies stellt nur eine Auswahl der wichtigsten Dienste dar, Amazon stellt noch weitere zur Verfügung, deren Vorstellung allerdings den Umfang der Arbeit übersteigen würde. (Weitere Informationen findet man unter [AWS EC2 2009].)

➤ Amazon Simple DB / Amazon RDS

Mit Simple DB stellt Amazon dem Anwender eine sehr einfache Datenbank zur Verfügung, mit der Aufwand des Entwurfs einer komplexen relationalen Datenbank übersprungen und deren Betrieb und Wartung minimiert werden kann. Dafür sind komplexere Abfragen oder semantische Integritätsbedingungen nicht möglich. Amazon gibt als Anwendungsbeispiele eine Datenbank für Loggingdaten, Onlinespiel-Highscores oder Metadatenindizes an, also Datenbanken mit keinem komplexen Entwurfsmodell.

Ist eine relationale Datenbank unumgänglich, so kann Amazon RDS genutzt werden, eine Implementierung von MySQL 5.1, welche man gesondert wie eine EC2-Instanz instanzieren und nutzen kann. Dies bietet vor allem bei Fragen der dynamischen Skalierung interessante Ansätze, es gilt auch hier Amazons Prinzip, nur für das zu bezahlen, was man wirklich verbraucht. Entsteht zum Beispiel im Lebenszyklus einer Social-Networking-Webseite der Bedarf, die Datenbank zu vergrößern, beziehungsweise den dahinter stehenden Rechner leistungsmäßig aufzuwerten, so kann dies bei AWS mit ein paar Mausklicks oder sogar automatisch geschehen, während man im „Normalfall“ physische Hardware nachrüsten müsste (was des Weiteren zum Beispiel auch noch Personalkosten nach sich ziehen würde und deutlich länger dauert).

➤ Amazon Cloud Front

Cloud Front ist ein sogenannter „Content-Delivery“-Dienst, welcher dafür sorgt, dass Nutzer einer Webseite bessere Latenzzeiten bei Datenübertragungen erhalten. Dies wird dadurch ermöglicht, dass Amazon Dateien aus einem S3-Bucket (siehe Abschnitt 3.3) in gewünschten geografischen Regionen gespiegelt vorgehalten werden. Ein Beispiel dafür wäre Amazons Webseite selbst, deren Bilder für Zugriffe aus Japan direkt in Japan nochmals als Kopie gehalten werden. Bei diesem Service zahlt man einen geringfügigen

Aufschlag für die Mehrfachhaltung der Daten in S3 und die Transferkosten bei der Vervielfältigung der Daten.

➤ Amazon Simple Queue Service

Um zwischen EC2-Instanzen einfache Daten austauschen zu können, bietet Amazon SQS an, ein simpler Messaging-Service, der pro Message bis zu 8KB große Daten enthalten kann. Queue-Management und Verwaltung simultaner Mehrfachzugriffe werden bereits von SQS verwaltet und müssen vom Benutzer nicht mehr beachtet werden.

➤ Amazon Elastic Map Reduce

Mit diesem Service bietet Amazon die Möglichkeit, sehr große daten- oder rechenintensive Aufgaben mit Hilfe von Hadoop (einer Opensource-Implementierung des Google MapReduce Algorithmus) parallel verarbeiten zu lassen. Dazu muss der Anwender keine tieferen Kenntnisse von Hadoop oder MapReduce besitzen, mit Hilfe einer einfachen SQL-ähnlichen Sprache können Verarbeitungs-Jobs formuliert und gestartet werden. Die zu verarbeitenden Daten werden vorher in S3 hochgeladen und können dann von dort aus direkt für die Berechnungen genutzt werden. Während der Verarbeitung kann man mit Hilfe von Monitoring-Tools den Fortschritt überwachen, die Ergebnisse werden ebenfalls zum Schluss in S3 gespeichert, um sie von dort abrufen zu können.

4 Yahoo! Cloud Computing

4.1 Allgemeines

Yahoo! verfolgt mit seinem Cloud-Computing-Ansatz ein etwas anderes Ziel als Amazon. So bestand bei Amazon die Cloud-Infrastruktur bereits vor der Veräußerung von Kapazitäten der Cloud an Unternehmen und Privatpersonen, Yahoo! versuchte dagegen, ihre bereits bestehende Nicht-Cloud-Infrastruktur durch Cloud-Technologie zu ersetzen und damit die immer weiter steigenden Anforderungen an ihre Social-Networking- und Web 2.0-Dienste zu decken ([Cooper 2009]). Als Dienstleister für Emailverkehr, Bilderhosting, Chat und als Suchmaschine (u.v.m.) stieg in den letzten Jahren Yahoo!'s Anforderungen an dynamische Skalierung, Datensicherung und –partitionierung und der allgemeine Hardwareverbrauch gewaltig. Somit war Yahoo! gezwungen, eine Cloud-Infrastruktur einzuführen, mit der sowohl alle Yahoo!-Seiten und –Dienste unterstützt, als auch deren interne Entwickler entlastet wurden.

4.2 Anforderung an die Cloud seitens Yahoo!

Yahoo! bietet seit Jahren zentralisiert verwaltete Datenmanagementdienste an. Die Anforderungen an die neu zu entwickelnde Cloud-Infrastruktur bildete sich also nicht nur aus allgemeinen Anforderungen an Clouds, sondern auch aus den bestehenden Anforderungen der existierenden Technik bei Yahoo!. Folgende wichtige Anforderungen stellt Yahoo! an sein Cloud-Rechenzentrum:

- **Mandantenfähigkeit:** Tausende User nutzen das Angebot von Yahoo!, es muss daher möglich sein, deren Accounts so auf einem System laufen zu lassen, dass niemand eines Anderes Daten einsehen oder gar beeinflussen kann und des Weiteren einzelne Dienste nicht voneinander beeinflusst werden. Das heißt, dass eine größere Lastspitze bei Yahoo! Mail nicht zu Verzögerungen beim Yahoo! Messenger führen darf.
- **Elastizität:** Es muss möglich sein, schnellstmöglich auf erhöhte Kapazitäts- und Datendurchsatzanforderungen reagieren zu können.
- **Skalierbarkeit:** Unterstützung für sehr große Datenbanken ist notwendig, ebenso soll die Erweiterung mit mehr Hardware und deren Einbinden sehr einfach und ohne großen Aufwand möglich sein.

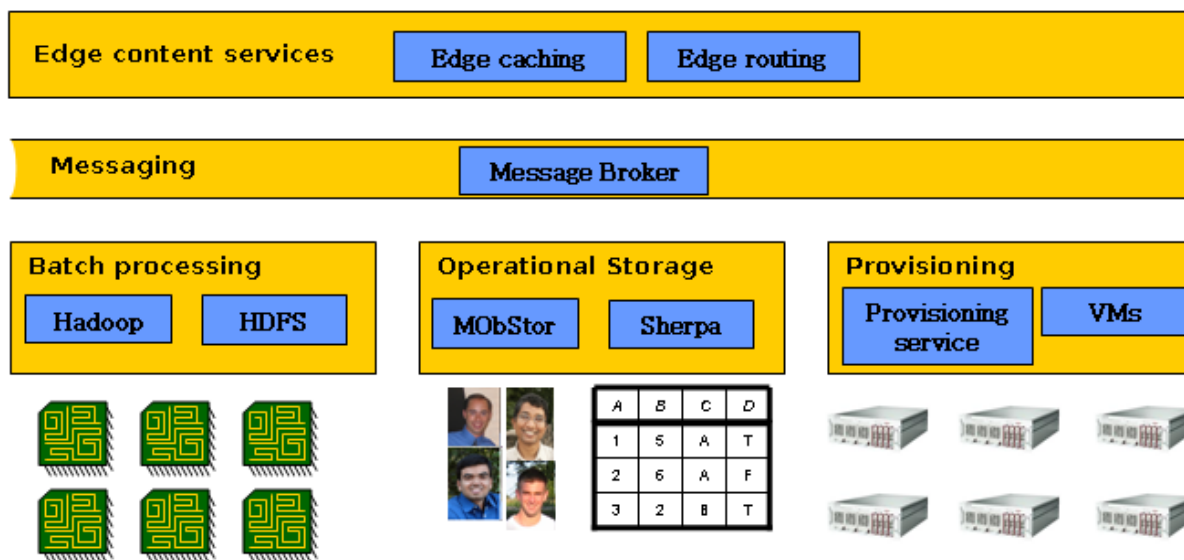
- **Lastverteilung:** Zwischen unterschiedlichen Servern soll es möglich sein, im Falle von einzelner Überlastung Aufgaben an andere Systeme zu verteilen.
- **Verfügbarkeit:** Die Anforderungen an die Cloud wird von Yahoo! sehr hoch gesteckt, selbst ein Ausfall eines gesamten Datacenters soll die Verfügbarkeit der Clouddienste nicht einschränken.
- **Globalisierung:** Auf Yahoo!-Dienste wird weltweit zugegriffen, es ist daher notwendig, Netzwerkverzögerungen und Flaschenhälse zu identifizieren und zu vermeiden, wenn ein weltweit reibungsloser Ablauf der Cloud-Dienste gewährleistet sein soll.

Wie also ist die Yahoo!-Cloud aufgebaut, um all diese Anforderungen, die an sie gestellt werden, zu erfüllen?

4.3 Architektur der Yahoo!-Cloud

Wie das Bild zeigt, unterteilt sich die Cloud von Yahoo! in drei Hauptbereiche:

1. *Edge Services,*
2. *Messaging und*
3. *Core Services.*



Aufbau der Yahoo!-Cloud (aus [Cooper 2009])

Auf das Wesentliche konzentriert sind die Core Services für die meisten serverseitigen Datenmanagementaufgaben zuständig, die Edge Services für Latenzreduzierung und Content-Delivery-Services. Dazu gehört auch Mehrfachcaching von Inhalten als auch Anfragerouting zwischen Servern und möglichen aufgetretenen Fehlerquellen. Die Messaging-Services sind für die Kommunikation zwischen sämtlichen anderen Komponenten erforderlich. Zum Beispiel für ein Update im Hauptspeicher zu einer Inkonsistenz sämtlicher Caches, Aufgabe der Message Services ist es nun, diese zu benachrichtigen, damit ein Update dieser erfolgen kann.

Die untere Zeile der Grafik teilt die Core Services nochmals in drei Teile. Die „*Batch processing*“-Systeme optimieren CPU-Zyklen bei größeren parallel ablaufenden Jobs. Im speziellen Falle von Yahoo! wird Hadoop eingesetzt (siehe auch 3.4 Weitere AWS-Dienste (Auswahl) - Amazon Elastic Map Reduce) und das dazugehörige Dateisystem HDFS. Der Bereich „*Optional storage*“ managt die Datenverwaltung fast aller Anwendungen. Er wird nochmals unterteilt in strukturierte Datensätze mittels Sherpa (auch PNUTS genannt, siehe dazu) und in sogenannte unstrukturierte „Blobs“ mittels *MObStor*, einem Cloud-Datenspeichersystem ähnlich Amazon S3 (siehe dazu auch [Navneet 2009]). Der dritte Teil der „*Provisioning Systems*“ dient der Zuweisung von virtuellen Servern für sämtliche Komponenten der Yahoo! Cloud.

4.4 Bestandteile der Yahoo!-Cloud (Auswahl)

➤ Hadoop

Die genaue Funktionsweise von Hadoop soll hier nicht vorgestellt werden, näheres dazu findet man unter [HADOOP]. Interessanter ist dessen Implementierung und Einsatzbereich bei Yahoo!. Hadoop läuft bei Yahoo! unter einem großen Cluster von Servern in der Cloud. Daten verbleiben meist ihre gesamte Lebenszeit in der Cloud im HPFS-Filesystem oder tauschen Daten mit Sherpa oder MObStor. Hadoop wird bei Yahoo! meist für Werbezwecke, machine learning bei der Such- / Inhalts-Optimierung und Spam-Reduzierung eingesetzt. Die Yahoo!-Search-WebMap ist eine Hadoop-Anwendung welche unter einem Linux mit 10000 Kernen läuft und mit mehr als 300TB komprimierten Daten arbeitet. Die Resultate sind ca. 33% schneller als ohne Hadoop-Optimierung.

➤ MObStor

Das generelle Ziel von MObStor ist es, den Users Daten schnell und einfach zugänglich zu machen und gleichzeitig Yahoo! eine gut skalierbare Massenspeicherlösung zu ermöglichen. Skalierbar heißt in diesem Falle sowohl in Hinsicht auf die Gesamtspeicherkapazität, als auch auf die Zugriffszeiten der Daten. Rein technisch gesehen ist MObStor eine Middleware, welche physische Speichersysteme wie mehrere SAN oder NAS virtualisieren und somit deren technische Betriebsbedingungen in den Hintergrund rücken. Die gespeicherten Dateien werden direkt mittels einer URL angesprochen und können damit direkt von anderen Webseiten referenziert werden. Des Weiteren unterstützt MObStor auch das Streamen und Cachen von Daten, so dass es Yahoo! ermöglicht wird, bestimmte Daten an geografische Hotspots zwischenspeichern, um die dortigen Zugriffszeiten zu senken. Anwendungsbeispiele für MObStor sind unter anderem die Werbung bei Yahoo!, Bilder für Yahoo! Maps und Mailanhänge bei Yahoo! Mail.

➤ Sherpa

Früher auch PNUTS genannt (siehe auch [Cooper 2008]), ist Sherpa ein Datenbanksystem, welches dem Anwender ein vereinfachtes relationales Datenbankmodell zur Verfügung stellt. Daten werden in regulären Tabellen, samt Datensätzen und Attributen gespeichert. Zusätzlich dazu existiert der Datentyp „blob“, welcher beliebige Strukturen in seinem Inneren zulässt. Einfache Abfrageoperationen sind möglich, komplexere Joins werden allerdings aus Performancegründen nicht unterstützt. Sherpa wurde so entworfen, dass Hadoop ebenfalls darauf zugreifen kann und statt nativem HDFS Daten auch in einer Sherpa Datenbank ablegen kann.

5 Zusammenfassung und Vergleich

Sowohl Amazons AWS Cloud-Dienste, als auch die Yahoo! Cloud sind, wie im Abschnitt 2 unterteilt, offensichtlich Cloud-Anbieter, welche eine Cloud auf der Infrastrukturebene betreiben. Amazons Cloud stellte sich des Weiteren als ein „On-Demand-Computing-Instances“-Dienst heraus (man kann also bei Amazon virtuelle Instanzen von Rechnern kaufen oder mieten und diese dann wie einen eigenen physischen Rechner betreiben), beziehungsweise Yahoo! als ein Anbieter von „On-Demand-Computing-Capacity“ (ein Angebot reiner Rechenkapazität). Beide Betreiber definieren verschiedene grundlegende Anforderungen an ihre Cloud, welche sich in den meisten Punkten sehr ähnlich sind, es wird nach Elastizität, Zuverlässigkeit, Sicherheit und Geschwindigkeit verlangt. Da Amazons Angebot sehr kommerziell ausgelegt ist, so ist auch deren angebotene Menge an Cloud-Services stark an einem wirtschaftlichen Modell angelehnt, so werden für alle optionalen Dienste extra Preise berechnet, in der Gesamtmenge übersteigen diese aber nicht den durchschnittlichen Preis eines oder mehrerer realer Rechner. Verzichtet man auf Cloud-Technologie, so entgehen einem nicht nur verschiedenste Chancen, die eigene Performance und Variabilität hinsichtlich Skalierung und Kapazität zu steigern, man muss außerdem an die vielen Zusatzkosten denken, die eine eigener Serverbetrieb mit sich bringt: Hardware- und Personalkosten, Zeit für Problembhebung, eigene Haftbarkeit. In dieser Hinsicht bietet Amazon eine hervorragende Möglichkeit, vor allem für kleine Startup-Unternehmen, welche sich zu Beginn nicht mit spezieller Technik und der Anschaffung teurer Hardware belasten möchten, als auch als Erweiterung für bestehende Unternehmen oder Privatpersonen, welche Wert auf effiziente Verarbeitung großer Datenmengen oder komplexer Algorithmen legen, die Rechenleistung in die Cloud zu verlegen. Yahoo! bietet seine Cloud momentan nur sehr zurückhaltend öffentlich an. Sie wird vor allem von Yahoo! selbst intern eingesetzt um deren eigenen angebotenen Dienste von Mail, Datenspeicher und Messagingdiensten zu unterstützen. Des Weiteren werden ausgewählte Vertretern von Universitäten oder aus dem Bereich der Forschung im Bereich Cloudtechnologien die Mitnutzung auf Anfrage gestattet, eine kommerzielle Vermarktung der Yahoo!-Cloud ist offenbar in naher Zukunft nicht in Sicht. So sehr sich die beiden Clouds in den oben genannten Dingen unterscheiden, so zeigen sich (trotz verschiedener Nomenklatur der Dienste) auch viele Gemeinsamkeiten innerhalb der Infrastrukturen. Folgende Grafik veranschaulicht, dass beide Dienste grundlegende

Basiskomponenten verwendet, die in ihrer allgemeinen Funktionsweise und Bestimmung gleich sind.

Amazon AWS	Yahoo! Cloud	Funktion
Amazon EC2	VMs (Provisioning Service)	Bereitstellen einzelner virtueller Rechner
Amazon S3	MOBStor	Simpler Filehosting-Service, mit Vorteil des schnellen Zugriffs und großer Dateien
Amazon RDS / Simple DB	Sherpa	Datenbanksystem
Amazon Elastic Map Reduce	Hadoop + HDFS	Parallelverarbeitung komplexer Prozesse
Amazon SQS	Message Broker	Simpler Nachrichtenaustausch zwischen Instanzen
Amazon Cloud Front	Edge Caching / Edge Routing	Geografisch orientiertes Cachen von Daten zur Optimierung von Zugriffszeiten

Es zeigt sich also, dass beide Systeme, trotz ihrer unterschiedlichen Intention beim momentanen Verwendungszweck, grundlegend eine sehr ähnliche Infrastruktur besitzen. Sollte Yahoo! seine Cloudtechnologie weiter optimieren können und zusätzliche Ressourcen implementieren, so stünde wahrscheinlich einer kommerziellen Vermarktung, zumindest von der technischen Seite aus gesehen, nichts im Wege.

6 Diskussion

Beide vorgestellten Anbieter benutzen ihre Clouds für ähnliche Zwecke, Amazon für das Betreiben der eigenen Seiten und als kommerzielles Angebot, Yahoo! als Entlastung der eigenen Dienste. Die Vorteile der Technologie, die beide Anbieter daraus ziehen, liegen auf der Hand. Komplexe Algorithmen in überschaubarer Zeit lösen, einfache Skalierung der Hardwareebene, automatische Lastverteilung, Reduzierung der Latenzzeiten und bei Amazon noch zusätzlich die wirtschaftliche Möglichkeit, das bereits bestehende System noch kommerziell vermarkten zu können, sind nur einige ausgewählte Punkte. Ein weiterer großer Vorteil gegenüber herkömmlicher IT-Infrastruktur ist das Bezahlen der Leistung nach Nutzungsdauer. Lokale Ressourcen und Dienstleistungen lassen sich damit sehr effektiv einsparen. Ein grundlegender Vorteil von *Cloud-Infrastruktur*diensten ist weiterhin, dass dem User relativ große Freiheiten in seiner Gestaltung der virtuellen Maschinen gelassen werden. Eigene Images hochladen, eigene Software einsetzen, spezielle Konfigurationswünsche – alles offenbar kein Problem. Doch genau das ist wahrscheinlich auch ein großer Nachteil dieser Dienste. Falsche Konfigurationen, Fehler in der Sicherheit sind mit der Verlagerung der IT in eine Cloudumgebung nicht nur gleichgeblieben, sondern nehmen wahrscheinlich auch noch zu, da vermehrt Datenverkehr zwischen Endkunde und Cloudanbieter stattfindet und spezielles Wissen der Administratoren erforderlich ist. In Internetforen werden, gerade bei Konfigurationen der AMIs bei Amazon via Terminalsitzung, Probleme bei der Bedienbarkeit und Nutzerfreundlichkeit genannt, sehr lange kryptische Konfigurationsbefehle auf Konsolenebene für die Images lassen Bedenken aufkommen. Auch besteht immer die Ungewissheit des Kunden, wo denn seine Daten nun eigentlich gespeichert sind und wie sicher der Zugriff oder der physische Zugang zu diesen nun wirklich ist. Eine weitere momentan ungeklärte Frage ist, in welchem rechtlichen Raum man sich bewegt, wenn die Daten eines deutschen Kunden zum Beispiel in den USA oder in China gehostet werden. Brisante Firmendaten im Ausland zu speichern, wo sie aufgrund der aktuellen Terrorismuspolitik dem Zugriff bestimmter Staaten relativ frei zugänglich sind, ist für viele Firmen ein Dorn im Auge, wenn es darum geht, Cloudtechnik in bereits bestehende IT zu integrieren. Gerade in dieser Hinsicht muss wohl in Zukunft sowohl technisch, als auch rechtlich noch viel verbessert werden, um die Bedenken zu vermindern - gänzlich ausgeräumt werden können sie wahrscheinlich nie.

Literaturverzeichnis

- [AWS EC2 2009]: <http://aws.amazon.com/ec2/>, aufgerufen 27.12.2009.
- [AWS SEC 2009]: <http://aws.amazon.com/security>, aufgerufen 27.12.2009.
- [Brantner 2008]: M. Brantner, D. Florescu et al., *Building a database on S3.*, SIGMOD 2008.
- [Cooper 2008]: B. F. Cooper, R. Ramakrishnan, U. Srivastava, et al., *PNUTS: Yahoo!'s hosted data serving platform.* VLDB, 1(2):1277–1288, 2008.
- [Cooper 2009]: B. Cooper, E. Baldeschwieler et al., *Building a Cloud for Yahoo!*, IEEE DE Bulletin, 2009.
- [Grossman, Gu 2009]: Robert L. Grossman, Yunhong Gu, *On the Varieties of Clouds for Data Intensive Computing*, <http://sites.computer.org/debull/A09mar/grossman.pdf>, aufgerufen 27.12.2009.
- [HADOOP 2009]: <http://hadoop.apache.org/>, aufgerufen 27.12.2009.
- [Navneet 2009]: Navneet Joneja, *MObStor: Yahoo!'s Unstructured Data Cloud*, <http://developer.yahoo.net/blog/archives/2009/07/mobstor.html>, aufgerufen 27.12.2009.
- [Hütter 2010]: Markus Hütter, *Cloud-Software-Plattformen*, Seminararbeit an der Universität Leipzig, 2010.