

Aufbau eigener Cloud-Infrastrukturen mit Eucalyptus

Andreas Ries

Fakultät für Informatik,
Hochschule Mannheim,
Paul-Wittsack-Straße 10,
68163 Mannheim
`ries.andreas@web.de`

Zusammenfassung Der Begriff Cloud Computing ist zur Zeit in aller Munde. Jeder möchte mit Hilfe der Technologie Geld und Zeit sparen. Daher gibt es am Markt eine Fülle von kommerziellen Anbietern, die ihren Service zur Verfügung stellen. Nur wenige User machen sich dabei Gedanken über die Sicherheit oder über die Techniken, die dabei zum Einsatz kommen. Dies ist oftmals auch nicht möglich, da es sich um sogenannte Private Clouds handelt. Um Einblick in die Funktionsweise und Architektur eines solchen Systems zu bekommen, besteht mit dem Open Source Projekt Eucalyptus die Möglichkeit, sich recht einfach eine eigene Cloud Struktur aufzubauen und zu verwalten. In diesem Beitrag wird der Aufbau und die Installation der Software beschrieben.

Diese Arbeit gibt einen Einblick in den Aufbau, die Installation und die Verwaltung einer eigenen Cloud-Struktur mit Eucalyptus. Nachdem einige Grundlagen erläutert sind, wird die Architektur, die einer Eucalyptus Cloud zugrunde liegt, genauer untersucht. Im folgenden Teil werden die wichtigsten Voraussetzungen und Schritte für die Installation einer eigenen Cloud aufgeführt und im letzten Teil ein kurzer Einblick in die Verwaltung der Cloud gegeben. Abschließend folgt eine kurze Zusammenfassung sowie ein Fazit des Autors.

1 Cloud Computing Begriffe

1.1 Public Cloud

Gartner defines public cloud computing as a style of computing in which scalable and elastic IT-enabled capabilities are delivered as a service to external customers using internet technologies. Private cloud computing is defined as a style of computing in which scalable and elastic IT-enabled capabilities are delivered as a service to internal customers using internet technologies [2]



Abbildung 1. Auswahl einiger Cloud Anbieter[3]

Wie Abbildung 1 zeigt, gibt es Public Clouds kommerzieller Anbieter zu Hauf im Internet. Sie bieten die Möglichkeit Ressourcen jeglicher Art zu mieten. Das Angebot umfasst dabei alle Arten von Cloud Systemen wie *Software as a Service* (SaaS), *Platform as a Service* (PaaS) und *Infrastructure as a Service* (IaaS). Durch die Spezialisierung der Anbieter auf die Versorgung ihrer Kunden mit Hardware, Plattformen, etc ist es ihnen möglich ihre Services zu sehr günstigen Preisen zur Verfügung zu stellen. Dadurch stellt sich für Unternehmen oftmals gar nicht erst die Frage, ob der Aufbau einer eigenen Cloud lohnenswert ist.

Vorteile:

- Sehr kosteneffizient
- Minimale Investitionskosten
- Schnelle Reaktionszeit
- Hohe Flexibilität

Nachteile

- Abhängigkeit vom Anbieter
- Gefahr eines Lock-In
- Vertrauliche Daten werden außerhalb der eigenen Einrichtung aufbewahrt
- Probleme beim Datenschutz

1.2 Private Cloud

Private Clouds sind weit weniger verbreitet als ihre öffentlichen Ableger. Sie sind häufig nur im wissenschaftlichen Bereich oder in Unternehmen mit sehr

vertraulichen Daten anzutreffen. Grund sind meist die immensen Kosten für die Infrastruktur, die Installation und die Wartung der Cloud. Daher kommen Entwicklungen aus dem Bereich der Private Clouds oftmals aus dem forschersichen Umfeld von Universitäten.

Vorteile:

- Voller Zugriff auf die Struktur der Cloud
- Kontrolle über alle Vorgänge
- Weiterentwicklung der Cloud
- Einbringung eigener Ideen

Nachteile

- Hohe Investitionskosten
- Verwaltungskosten
- Entsprechendes Know-How erforderlich

1.3 Infrastructure as a Service

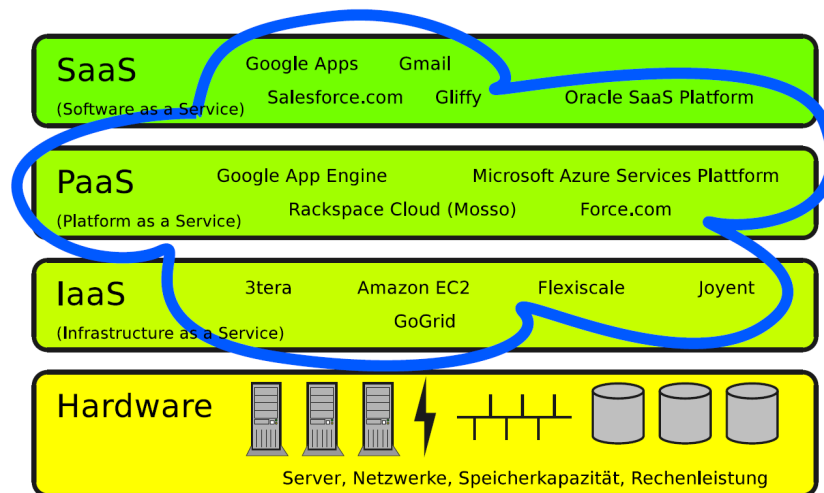


Abbildung 2. Verschiedene Arten an Cloud-Systemen[1]

Ein Merkmal des Cloud Computing ist das zur Verfügung stellen von Ressourcen. Diese Ressourcen sind im Falle des *Infrastructure as a Service* Hardwarekomponenten wie CPUs, RAM oder Festplattenspeicher. Der Kunde hat dabei

nur eine abstrakte Sicht auf die Hardware und kann diese nach seinen individuellen Bedürfnissen zusammenstellen. Auf den so erstellten virtuellen Rechnern kann er beliebige Betriebssysteme mit Administrator-Rechten betreiben und so alle relevanten Einstellungen wie Firewall, User und Storage-Management selbst vornehmen.[1]

2 Eucalyptus

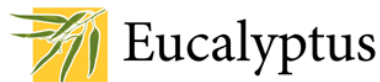


Abbildung 3. Logo der Open Source Software Eucalyptus [3]

Elastic Utility Computing Architecture for Linking Your Programs To Useful Systems, kurz Eucalyptus, ist eine Open Source Software zur Erzeugung eigener Cloud Strukturen. Entwickelt wurde Eucalyptus an der University of California, Santa Barbara. Eucalyptus entstand aus einem Forschungsprojekt und ist auf der Homepage der Universität seit Mai 2008 zum Download verfügbar. Seit dem offiziellen Release sind in relativ kurzen Abständen mehrere Versionen erschienen. Die Aktuellste ist zur Zeit Version 1.6.1 [3]

2.1 Eucalyptus Komponenten

Eine Eucalyptus Cloud besteht im wesentlichen aus vier Komponenten; dem Cloud Controller (CLC), dem Cluster Controller (CC), dem Node Controller (NC) und des Speicherverwaltungssystems Walrus, die alle als Web Services implementiert wurden. Jeder Controller repräsentiert eine Ebene in der Cloud und hat spezifische Aufgaben.[4]

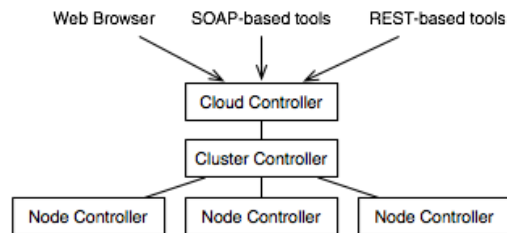


Abbildung 4. Die 3 Ebenen einer Eucalyptus Cloud [3]

Cloud Controller Der CLC repräsentiert die oberste Ebene in der Cloud. Er ist einmalig in der Cloud und die Schnittstelle zur Außenwelt. Er nimmt Aufträge entgegen und verteilt diese auf die nächste Ebene zu den Clustern. Dazu sammelt der CLC die Ressourceninformationen, die er von seinen CC's übermittelt bekommt.

Cluster Controller Ein CC verwaltet einen Pool von NC,s, welche dem CC ihre Ressourcen mitteilen. Der CC entscheidet anhand dieser Daten, welcher Node eine angefragte VM ¹ starten soll. Des weiteren ist er für die Verwaltung des Netzwerkes zuständig, das ihn und seine NC's verbindet.

Node Controller Ein NC repräsentiert die unterste Ebene in der Architektur. Er muss auf jeder Maschine installiert sein, die später VM starten soll. Der NC teilt seinem CC den aktuellen Zustand seiner Ressourcen (virtuelle Prozessoren, freier Haupt- und Festplattenspeicher) mit und ist für die von ihm gestarteten VM's zuständig.

Walrus Walrus ist der Speicherservice innerhalb der Cloud. Die Hauptaufgabe von Walrus ist das Lagern von Eucalyptus Machine Images (EMI). EMI's beinhalten die Abbilder der Betriebssysteme, die später von den NC's gestartet werden können. Zusätzlich kann Walrus als Raw Data Storage verwendet werden und so auch andere relevante Daten speichern.

[1] [4]

3 Installation

Die Entwickler von Eucalyptus werben mit der einfachen Installation ihres Produktes. In der Tat geht die Installation der Software, dank der umfangreichen Dokumentation, recht einfach von der Hand. Die Installation erfolgt mit Hilfe von Binärpaketen, die zur Zeit für CentOS, Ubuntu, openSUSE und Debian auf der Homepage zum Download bereitstehen.

3.1 Voraussetzungen

Um eine Cloud mit Hilfe von Eucalyptus aufzubauen ist mindestens ein Rechner mit einem installierten Linux und root-Zugriff erforderlich. Es ist möglich den CLC, den CC, den NC und Walrus auf ein und demselben Rechner laufen zu lassen. Als *Best Practice* hat es sich jedoch erwiesen, den CLC, den CC und Walrus zusammenzufassen und für jeden Node einen NC zu verwenden. Neben

¹ Virtual Machine

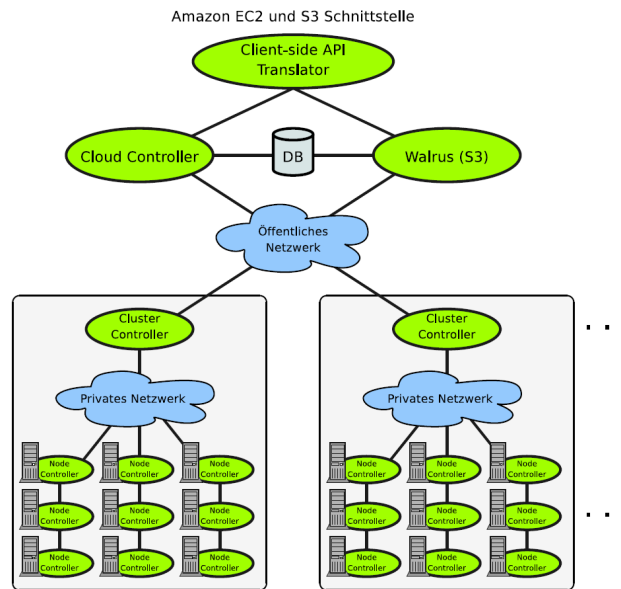


Abbildung 5. Komplette Übersicht über eine Eucalyptus Cloud [1]

Linux muss außerdem auf jedem Rechner, der später VM's starten soll, ein Xen-Hypervisor oder eine KVM² vorhanden sein.

3.2 Schritte zur Installation

- Xen auf allen Nodes installieren
- Eucalyptus Komponenten installieren
- Cluster einrichten
- Dem Cluster Nodes hinzufügen
- Ein System in ein Bucket legen (EMS)
- Instanz starten

[5]

4 Administration

Eine Eucalyptus Cloud ist kompatibel zu den Amazon EC2 Tools und kann mit diesen komplett verwaltet werden. Zur Zeit arbeitet man zusätzlich an den

² Kernel Virtual Machine

Euca2ools, um eigene Verwaltungstools zur Verfügung stellen zu können. Da alle Komponenten als Web Service realisiert sind, kann auf die meist benutzten Funktionen auch komfortabel per Web Browser zugegriffen werden.

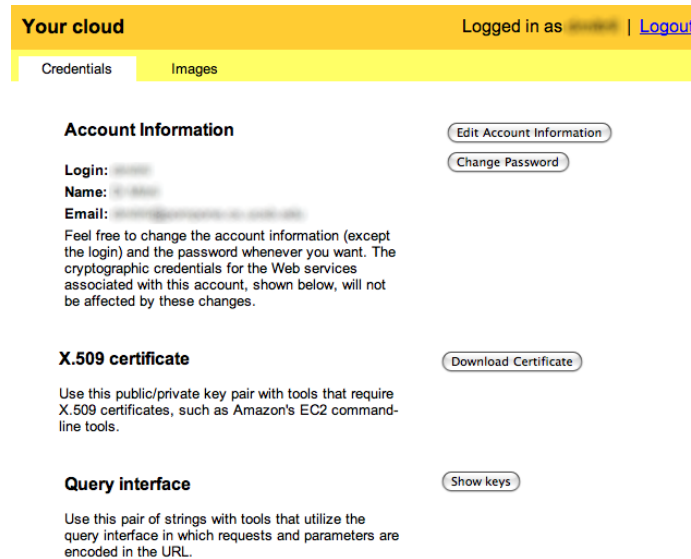


Abbildung 6. Erster Log-In in die Cloud[3]

4.1 VM Typen

VM Typen legen die Ressourcen fest, die eine Instanz dieser VM zur Verfügung stehen. Eucalyptus erlaubt die freie Konfiguration von sechs VM Typen. Standardmäßig sind diese wie folgt belegt:

| Name | CPUs | RAM | Festplatte |
|-----------|------|------|------------|
| m1.small | 1 | 128 | 10 |
| c1.medium | 1 | 256 | 10 |
| m1.large | 2 | 512 | 10 |
| m1.xlarge | 2 | 1024 | 20 |
| c1.xlarge | 4 | 2048 | 20 |

Tabelle 1. Liste der VM Typen (Standardbelegung)[3]

Mit Hilfe des Kommandos

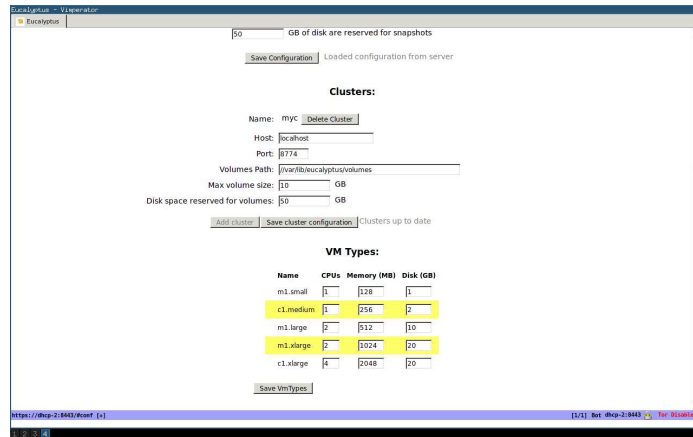


Abbildung 7. Eingabe der VM Typen in Eucalyptus[3]

EC2-DESCRIBE-AVAILABILITY-ZONES VERBOSE

wird die Liste der definierten VM Typen ausgegeben, sowie die freien und maximalen Instanzen des Clusters angeben.[1]

4.2 Netzwerk Modi

Zur Verwaltung des privaten Netzwerkes innerhalb der Cluster bietet Eucalyptus mehrere Modi an. Die drei wichtigsten werden im Folgenden kurz erklärt.

SYSTEM Mode Dies ist der einfachste Netzwerkmodus, bietet aber auch die wenigsten Möglichkeiten. In diesem Modus verteilt Eucalyptus einfach zufällig MAC Adressen an die VM Instanzen bevor diese booten und verbindet sie über die Xen Bride mit dem Netzwerk. Die Instanzen holen sich dann über einen DHCP Server ihre IP-Adressen. Daher muss dafür gesorgt sein, dass ein DHCP Server mit freien IP-Adressen zur Verfügung steht. Dieser Modus hat den Vorteil, dass Instanzen schnell und einfach gestartet werden können, was besonders zum Ausprobieren am Laptop geeignet ist.

STATIC Mode Dieser Modus gibt dem Cloud Administrator mehr Kontrolle über die Verteilung der IP-Adressen an die Instanzen. Der Administrator legt hierzu erst eine Liste mit Paaren von MAC-/IP-Adressen an. Eucalyptus arbeitet dann bei der Erstellung neuer Instanzen diese Liste ab und vergibt dabei die Paare der MAC-/IP-Adressen. Eucalyptus fungiert hier also als DHCP Server. Der Rest funktioniert ähnlich dem System Mode. Geeignet ist diese Variante für Administratoren, denen nur eine festgelegte Anzahl von Adressen zur Verfügung stehen.

MANAGED Mode Dies ist der Modus mit den meistens Möglichkeiten, aber auch den meisten Auflagen hinsichtlich der Konfiguration. Im Managed Mode übernimmt Eucalyptus, genau wie im Static Mode, die Rolle des DHCP Servers. Allerdings handelt es sich hier um ein sehr viel größeres Netzwerk. Benutzer können eine Reihe von *named networks* oder *security groups* innerhalb dieser Netzwerke definieren und ihren VM Instanzen diesen zuweisen. Auf diese Weise können die Benutzer spezielle Regeln für ihre Netzwerke und somit ihre Instanzen festlegen, wie zum Beispiel SSH-Support. Außerdem ist es möglich den Instanzen *Public IP's* zuzuweisen.³

[3]

5 Fazit

Eucalyptus ist auf einem gute Weg sich in der Cloud Computing Welt zu etablieren. Als Private Cloud steht es in Kontrast zu dem viel genutzten Angebot der Amazon EC2. Die Entscheidung, die gut bekannten Amazon Services zu kopieren und Eucalyptus kompatibel zu fast allen Features der Amazon Cloud zu machen, macht den Umstieg und die Einarbeitung in dieses neue System einfacher. Auch die Kompatibilität zwischen Walrus und Amazon S3 war gut gewählt. Dies zeigt auch die Tatsache, dass Eucalyptus in der nächsten Version von Ubuntu (Karmic Koala) fest enthalten sein wird. [3] Die Entscheidung mit den Euca2ools nun auch eigene Verwaltungsprogramme bereitzustellen zeigt, dass man nicht nur im Schatten von Amazon bleiben will, sondern sich einen eigenen Namen in der Welt der Cloud Computing machen möchte.

Die Frage, die bleibt, ist nun welche Kundengruppe nach Private Clouds wie Eucalyptus verlangt. Zur Zeit ist vor allem im wissenschaftlichen Umfeld das Interesse hoch, aber auch die Industrie wird in absehbarer Zeit ihr Interessen erkennen. Die Gefahren von Public Clouds werden auf die Dauer für einige Unternehmen nicht mehr tragbar sein und einige Prozesse lassen sich auch im Moment noch nicht aus dem Unternehmen heraus verlagern. Daher wird wohl in naher Zukunft der Markt für Private Clouds vorhanden sein; im kleinen Umfeld, wie zum Beispiel bei Schulungsfirmen, als auch bei großen Konzernen, die aufgrund des Datenschutzes die volle Kontrolle über ihre Daten behalten möchten.

Literatur

1. Christian Baun und Marcel Kunze:
Cloud-Infrastrukturen mit Linux und Eucalyptus aufbauen
Linuxtag August 2009
2. <http://it-republik.de/business-technology/news/Gartner-Unternehmen-werden-Private-Clouds-Vorzug-geben-052759.html>
Letzter Besuch: 05.01.2010

³ Vergleichbar mit den *Elastic IP's* von Amazon

3. <http://open.eucalyptus.com>
Wiki, Admin Guide, User Guide *Version 1.5*
4. Tingxi Tan and Cameron Kidle:
An Assessment of Eucalyptus Version 1.4
University of Calgary, Canada *04.05.2009*
5. Michael Stapelberg:
Aufbau von Cloud-Infrastrukturen mit Eucalyptus
Seminararbeit, University of Heidelberg *25.05.2009*