

Aufbau und Betrieb einer Private Cloud mit Open Source Software

iX CeBIT Forum – Hannover – 5. März 2010

Christian Baun, Marcel Kunze

baun | kunze@kit.edu



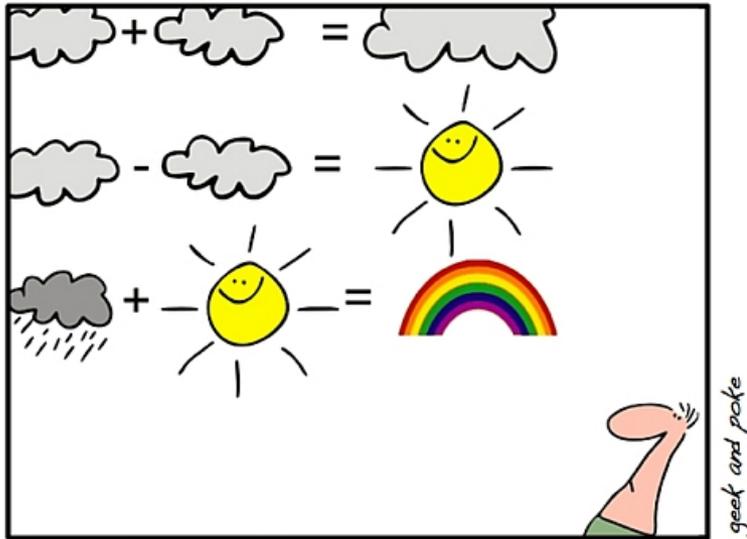
Forschungszentrum Karlsruhe
in der Helmholtz-Gemeinschaft



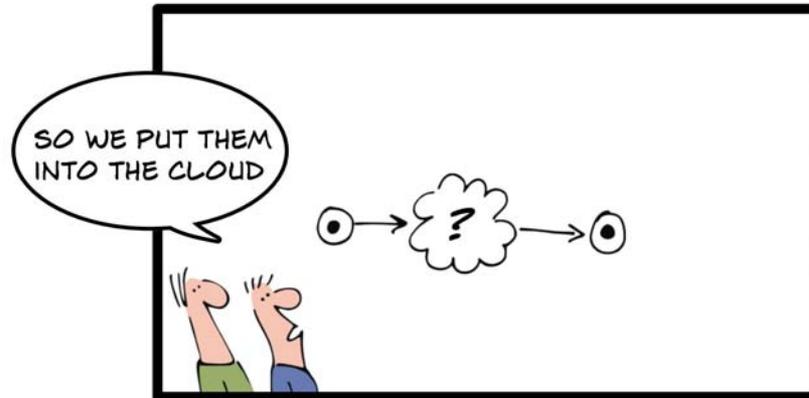
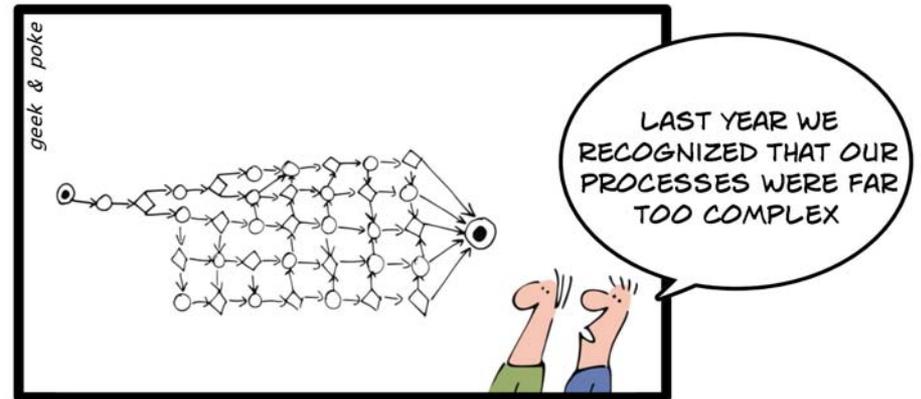
Universität Karlsruhe (TH)
Forschungsuniversität · gegründet 1825

Cloud-Computing 2009 + 2010

- Allgemein großes Interesse am Cloud-Computing
- Hype-Thema seit Ende 2008
- Gute Definitionen sind schwer zu finden
- Was steckt dahinter?



SIMPLY EXPLAINED - PART 17:
CLOUD COMPUTING



LET THE CLOUDS MAKE YOUR LIFE EASIER

Definition: Cloud-Computing

■ Definition

- Unter Ausnutzung virtualisierter Rechen- und Speicherressourcen und moderner Web-Technologien stellt Cloud-Computing skalierbare, netzwerk-zentrierte, abstrahierte IT-Infrastrukturen, Plattformen und Anwendungen als on-demand Dienste zur Verfügung. Die Abrechnung dieser Dienste erfolgt nutzungsabhängig.

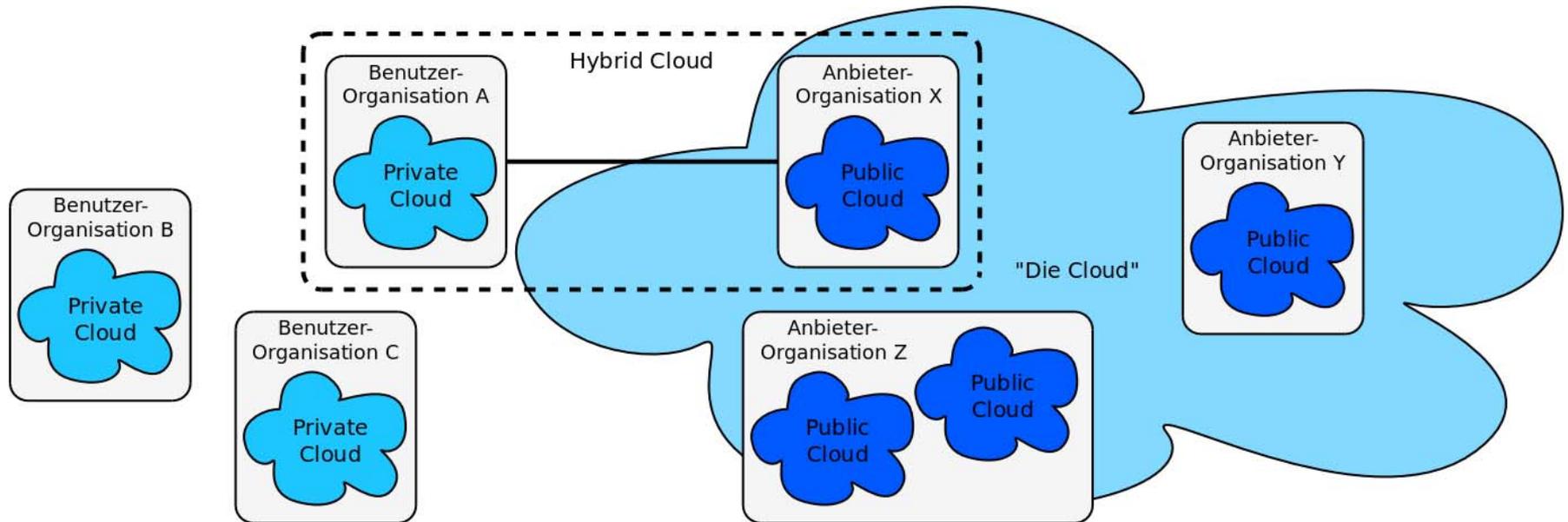
■ Übergang der IT in das Zeitalter der Industrialisierung

- Ein oder wenige Rechenzentren mit heterogenen oder homogenen Ressourcen unter zentraler Kontrolle
- Virtualisierte Ressourcen
- Verbrauchsabhängige Abrechnung
- Benutzerfreundlich (geringe Einarbeitungszeit notwendig)

mitp



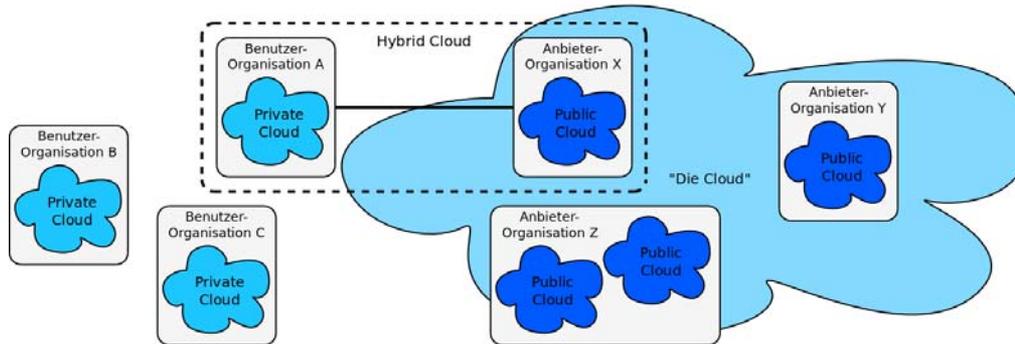
Drei organisatorische Typen von Clouds



■ Public Cloud (bzw. External Cloud)

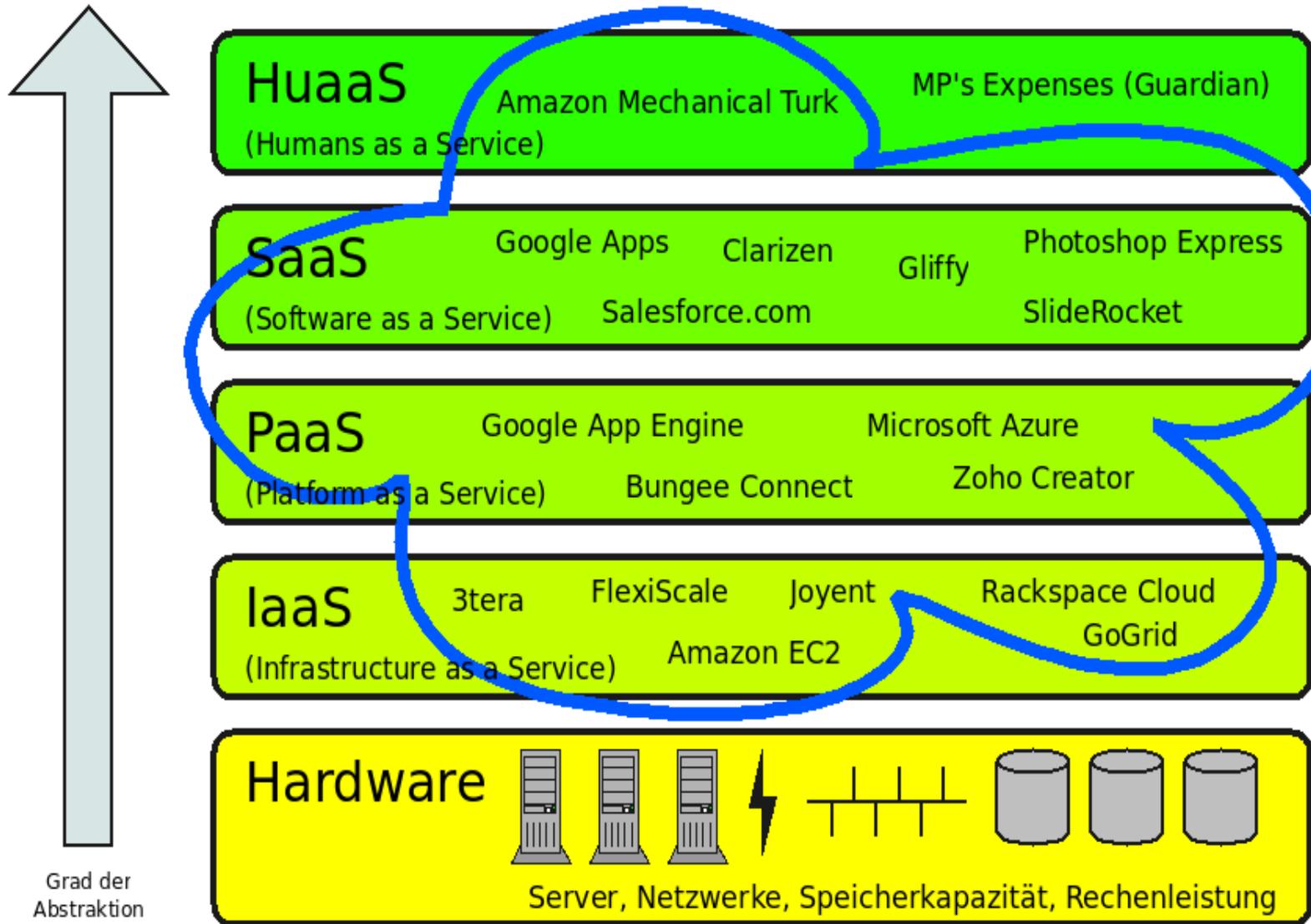
- Anbieter und Nutzer gehören unterschiedlichen Organisationen an
- Anbieter verfolgen immer kommerzielle Geschäftsinteressen
- Nutzer haben keine Kosten für Anschaffung, Betrieb und Wartung eigener Hardware
- Public Cloud Dienste liegen immer vor der eigenen Firewall
- Theoretisches Risiko des Lock-in
- Datenschutz und Datensicherheit beachten!

Drei organisatorische Typen von Clouds



- **Private Cloud (bzw. Internal Cloud oder Intra Cloud)**
 - Anbieter und Benutzer der Cloud Dienste gehören der gleichen Organisation an
 - Dienste der Private Cloud sind im Idealfall kompatibel zu den Public Cloud Diensten etablierter Anbieter
 - Keine Gefahr des Lock-in
 - Datenschutz ist kein Problem
 - Ähnlich hohe Kosten für eigene Hardware, Stellplatz und Administration wie bei einer nicht-Cloud-basierten Architektur
- **Hybrid Cloud**
 - Dienste aus Public und Private Clouds werden gemeinsam innerhalb einer Infrastruktur verwendet
 - Anwendungsgebiete:
 - Lastspitzen mit den Ressourcen von Public Clouds abfedern
 - Sicherheitskopien von Daten in Public Clouds auslagern

Die vier (wichtigsten) Kategorien von Cloud-Systemen



Humans as a Service (HuaaS)

HuaaS

Amazon Mechanical Turk

(Humans as a Service)

MP's Expenses (Guardian)

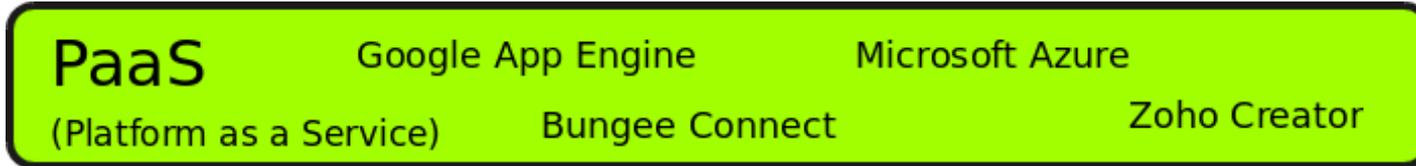
- **Prinzip des Crowdsourcing**
- **Ressource Mensch wird zu geringen Kosten oder als Spende Freiwilliger angeboten**
- **Interessant für...**
 - **niedrig qualifizierte Tätigkeiten**
 - **Tätigkeiten, die Computer gar nicht, nur schlecht oder erst nach einer unverhältnismäßig hohen Entwicklungszeit erledigen könnte**
- **Einsatzmöglichkeiten sind u.a.**
 - **Bilderkennung**
 - **subjektive (Produkt)-Bewertungen**
 - **Übersetzungen**
 - **(Produkt)-Zuordnungen zu (Produkt)-Kategorien**
- **Beispiele für Public Cloud HuaaS**
 - **Untersuchung der Zeitung The Guardian zum britischen Spesenskandal 2009**
 - **Marktplatz für HuaaS: Amazon Mechanical Turk**

Software as a Service (SaaS)



- Applikationen werden durch Dienstleister betrieben und als Service zur Verfügung gestellt
- Nutzer kümmern sich nicht um Installationen, Administration, Updates, ...
- Nutzung der Software ist kostenlos oder wird nach Verbrauch abgerechnet
- Lokale Installation der Software ist nicht vorgesehen
- Auf Seite der Nutzer ist lediglich ein Browser notwendig
- Benutzer muss dem Anbieter vertrauen im Bezug auf:
 - eigene Daten
 - Verfügbarkeit der Dienste
- Beispiele für Public Cloud SaaS
 - Google Docs, Salesforce.com, Gliffy, Clarizen, Zoho Writer, SlideRocket, Adobe Photoshop Express, Microsoft Office Live und Apple iWork.com
- Beispiele für Private Cloud SaaS
 - Sugar und Zimbra

Platform as a Service (PaaS)



- **Skalierbare, integrierte Laufzeitumgebungen**
- **Häufig auch Entwicklungsumgebungen**
- **Anbieter unterstützen meistens 1 oder 2 Programmiersprachen**
- **Erscheint als ein großer Rechner und macht es einfach zu skalieren**
- **Zielgruppe sind primär Entwickler**
- **Geringere Flexibilität als bei IaaS**
- **Kein Administrationsaufwand für die Nutzer**
 - **Anwender kümmern sich nicht um das Betriebssystem, Systemsoftware, ...**
- **Beispiele für Public Cloud PaaS**
 - **Google App Engine, Amazon Simple DB, Bungee Connect, Zoho Creator und Windows Azure Services Plattform**
- **Beispiele für Private Cloud PaaS**
 - **AppScale und 10gen**

Infrastructure as a Service (IaaS)



- Selten auch Hardware as a Service (HaaS) genannt
- Nutzer betreiben virtuelle Server-Instanzen mit (fast) beliebigen Betriebssystemen und unveränderten Anwendungen auf den Serverfarmen des Anbieters
- Nutzer haben innerhalb ihrer Instanzen Administratorenrechte
- Nutzer haben keinen Kontakt zur physischen Hardware
- Firewall-Regeln können selbst definiert werden
- Welche Betriebssysteme unterstützt werden, hängt vom Dienst-Anbieter ab
- Komplette Rechenzentren können virtuell realisiert werden
- Beispiele für Public Cloud IaaS
 - 3tera, Amazon Elastic Compute Cloud (EC2), GoGrid, Joyent, FlexiScale, Rackspace Cloud (vormals Mosso) und Zimory Public Cloud
- Beispiele für Private Cloud IaaS
 - Eucalyptus, Enomaly, OpenNebula, Nimbus und abiCloud

Amazon Web Services (AWS)

- Die Amazon Web Services (AWS) sind eine Sammlung verschiedener Cloud-Dienste
- <http://aws.amazon.com>
- Abrechnung nach Verbrauch
- Bekannte Dienste innerhalb der AWS sind u.a. EC2, S3, EBS...
- Amazon Elastic Compute Cloud (EC2)
 - Dienst für virtuelle Server (Instanzen)
- Amazon Simple Storage Service (S3)
 - Dienst für Webobjekte
- Amazon Elastic Block Store (EBS)
 - Dienst für persistente Datenspeichervolumen
- Amazon SimpleDB
 - Verteiltes Datenbankmanagementsystem
- Amazon Simple Queue Service (SQS)
 - Dienst für Nachrichtenwarteschlangen (Message Queues)
- Amazon Mechanical Turk
 - Marktplatz für HaaS/Crowdsourcing
- ...



Kommerzielle Cloud Anbieter (kleine Auswahl)



APPLIC | UTILITY COMPUTING | TECHNOLOGY | PARTNERS | GRID UNIVERSITY | COMPANY

Cloud Computing

Cloudware - Cloud Computing Without Compromise



- Kommerzielle Cloud Angebote sind meist proprietär
 - Nicht alle Aspekte der Architektur einsehbar und beeinflussbar
- Aufbau eigener Cloud-Infrastrukturen (Private Cloud) nicht immer möglich und nicht einfach
 - Aufbau einer Hybrid Cloud ist noch schwieriger

Übersicht über Private Cloud IaaS (1)

■ **abiCloud**

- <http://www.abicloud.org>
- Unterstützt VMware, KVM, Xen und Virtual Box
- Open Source (steht unter der GNU Lesser General Public License v3)
- Version 1.0.0 seit 23.2.2010
- Noch sehr neu und unbekannt

■ **OpenNebula**

- <http://www.opennebula.org>
- Unterstützt VMware, KVM und Xen (Virtual Box geplant für Version 1.4.2)
- Bietet Migration von Instanzen
- Möglichkeit zum Aufbau einer Hybrid Cloud mit EC2
- Geringe Verbreitung
- Nur ein kleiner Teil (Stand Version 1.4) der EC2-API wird unterstützt
 - upload, register und describe images
 - run und terminate instances
- Kein S3 und EBS
- Open Source (steht unter der Apache License Version 2.0)

Übersicht über Private Cloud IaaS (2)

■ Nimbus

- <http://www.nimbusproject.org>
- Alter Name: Virtual Workspace Service (VWS)
- Unterstützt KVM und Xen
- Baut auf der Grid-Middleware Globus 4 auf
- Kleiner Teil (Stand Version 2.3) der EC2-API wird unterstützt
 - describe images
 - describe, run, reboot und terminate instances
 - add und delete keypair
- EC2-kompatible Ressourcen können gesteuert werden
- Open Source (steht unter der Apache License Version 2.0)

Übersicht über Private Cloud IaaS (3)

■ Enomaly Elastic Compute Cloud (ECP)

- <http://src.enomaly.com>
- Unterstützt VMware, KVM, Xen, Virtual Box und OpenVZ
- Open Source (steht unter der Affero General Public License v3)
- Kaum Dokumentation
- Keine API für die Benutzer
- Eingeschränkte API zur Administration
- Skaliert nur bis < 10 Knoten
- Seit Herbst 2009 ist Enomaly ECP nicht mehr offiziell verfügbar
- Abkehr des Unternehmens von der Open Source Strategie

■ OpenECP

- <http://www.openecp.org>
- Fork von Enomaly ECP
- Open Source (steht unter der Affero General Public License v3)
- Projektstart: Februar 2010
- Noch sehr neu und unbekannt

- **Open-Source um eine Private Cloud IaaS auf Clustern aufzubauen**
 - Ermöglicht Ausführung und Kontrolle virtueller Instanzen (Xen oder KVM) auf verschiedenen physischen Ressourcen
- **EUCALYPTUS - Elastic Utility Computing Architecture for Linking Your Programs To Useful Systems**
- **Entwickelt an der UC Santa Barbara**
 - Weiterentwicklung durch Eucalyptus Systems, Inc.
- **Versuch, die AWS zu emulieren**
 - Schnittstelle kompatibel zu Amazon EC2
 - Enthält „Walrus“, einen S3-kompatiblen Speicher-Dienst
 - Enthält „Storage Controller“, einen EBS-kompatiblen Speicher-Dienst für persistenten Speicher
- **Nutzung bekannter Tools für AWS möglich**
 - z.B: S3 Curl, Elasticfox, s3cmd, ...
- **Schritt hin zur Etablierung von AWS als Standard für Cloud-Infrastrukturen**

Eucalyptus (v1.6) – Komponenten



Eucalyptus

Cloud Controller (CLC)

- Meta-Scheduler in der Cloud
- Sammelt Ressourcen-Informationen der CCs
- Enthält standardmäßig die S3 und EBS Dienste

Cluster Controller (CC)

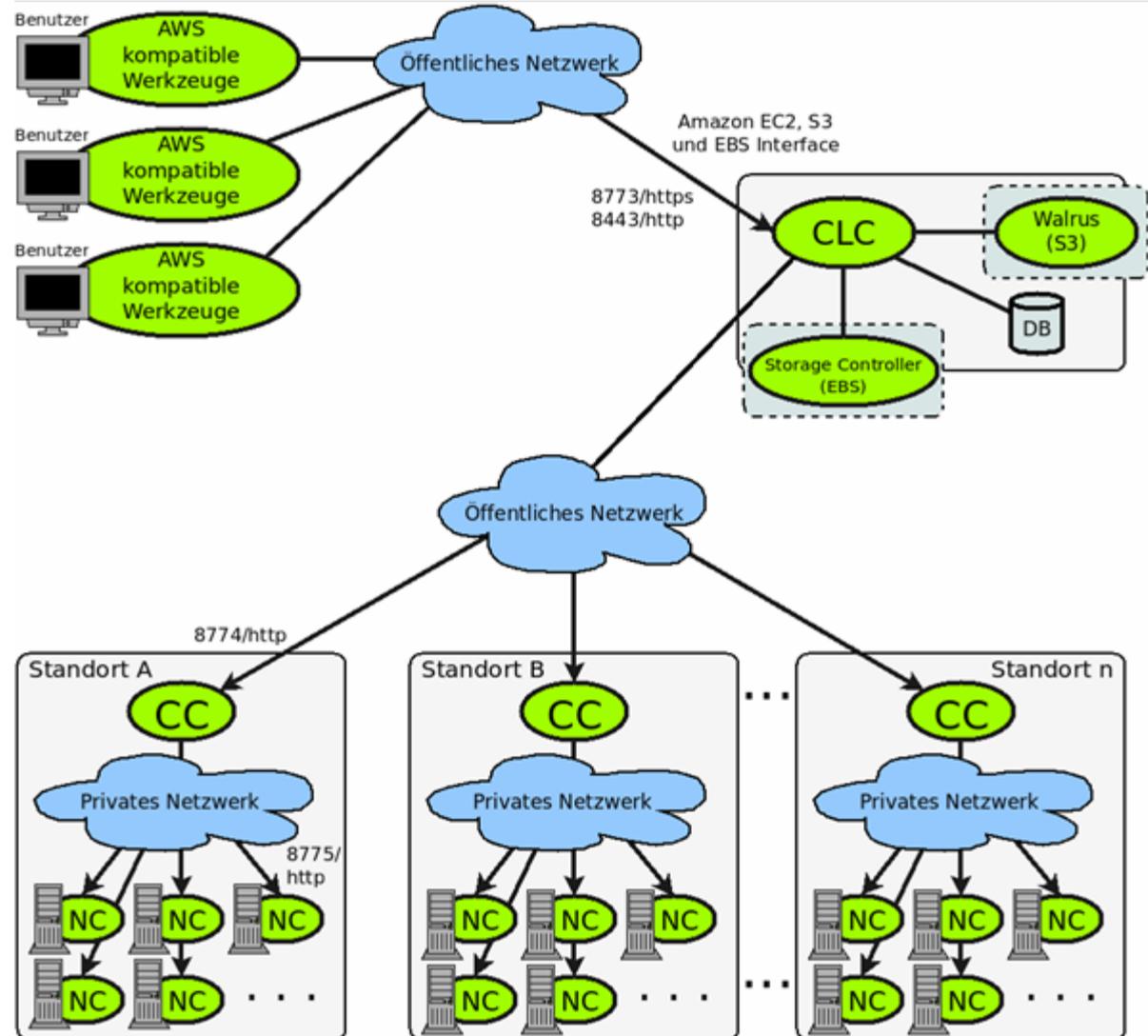
- Regelt die Verteilung der virtuellen Maschinen auf die NCs
- Sammelt Informationen über freie Ressourcen

Node Controller (NC)

- Läuft auf jedem Rechenknoten
- Steuert den Xen-Hypervisor oder KVM
- Schickt Ressourcen-Informationen an den CC

Walrus

Storage Controller



In 60 Minuten eine eigene Private Cloud (1)

- 2 Rechner
 - Rechner 1: CLC, CC, Walrus (S3), Storage Controller (EBS)
 - Rechner 2: NC
- Auf beiden Rechnern Ubuntu 9.10 Server installieren
 - *Ubuntu Enterprise-Wolke installieren*
 - Rechner 1: Cloud installation mode: Cluster
 - IP-Adressraum für die Instanzen festlegen
 - Rechner 2: Cloud installation mode: Node
 - Der CLC/CC (Rechner 1) sollte automatisch gefunden werden



Ubuntu Server installieren
Ubuntu Enterprise-Wolke installieren
CD/DVD auf Fehler prüfen
Arbeitsspeicher testen
Von der ersten Festplatte starten
Ein beschädigtes System reparieren

```
[!!] Select cloud installation mode
```

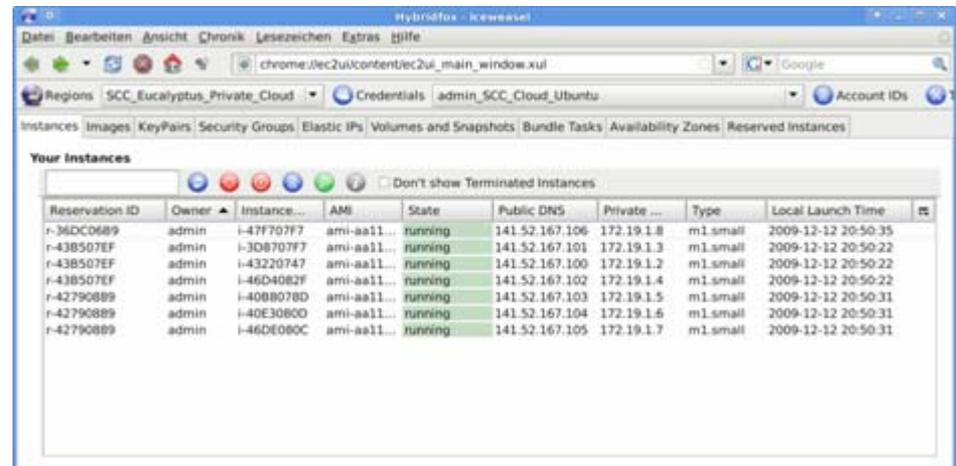
```
There is already a Eucalyptus cluster controller on your network, so  
this installation will default to installing a new computing node.  
Select "Cluster" instead if this is a mistake and you already have a  
node controller in place.
```

```
Cloud installation mode:
```

```
Cluster  
Node
```

In 60 Minuten eine eigene Private Cloud (2)

- Elasticfox oder Hybridfox unter Firefox installieren
 - <http://s3.amazonaws.com/ec2-downloads/elasticfox.xpi>
 - <http://code.google.com/p/hybridfox/>
- Zugang zur Private Cloud besorgen
 - <https://<Rechner1>:8443>
 - Apply for account
- Zugangsdaten besorgen
 - Credentials → Query interface credentials
- Zugangsdaten in Elasticfox/Hybridfox eintragen
 - Regions → Endpoint: URL: <http://<Rechner1>:8773/services/Eucalyptus>
 - Credentials: AWS Access Key und AWS Secret Access Key eintragen



Fakten zu Eucalyptus

- **Installation von Eucalyptus ist schwieriger als kommuniziert**
 - Installationsanleitungen lesen sich einfach
 - Deutliche Verbesserung der Situation seit Ubuntu 9.10 Server
 - Aber: Noch viel Raum für Verbesserungen (noch viele kleine Fehler/Bugs)
- **Stabiler Betrieb einer Eucalyptus IaaS ist möglich**
 - Einrichtung ist aber schwierig
 - Einzelne Dienste brauchen gelegentlich einen Neustart
 - Hilfe der Entwickler nicht immer hilfreich
 - Kommerzieller Support verfügbar
 - Teilweise scheinen kommerzielle Interessen dem Community-Gedanken entgegen zu stehen (u.a. Unterstützung von VMware)
- **Erweiterungen und Änderungen am Quellcode sind schwierig**
 - Quellcode und Komponenten erscheinen undurchsichtig
 - Von Seiten der Entwickler ist hier keine Hilfe zu erwarten

AppScale

<http://appscale.cs.ucsb.edu>



- **Open-Source Reimplementierung der Google App Engine (GAE)**
- **Entwickelt an der UC Santa Barbara**
- **GAE ermöglicht es, Web Anwendungen in Python und JAVA in der Google Infrastruktur zu betreiben**
- **AppScale ermöglicht den Aufbau einer eigenen PaaS, kompatibel zur GAE**
- **GAE-kompatible Anwendungen für GAE können innerhalb einer Private Cloud (Eucalyptus) und innerhalb einer Public Cloud (EC2) betrieben und getestet werden**
- **AppScale emuliert auch Google-Infrastruktur-Dienste wie Datastore, Memcache und die Authentifizierung**

AppScale - Cloud Status
The open source framework for running Google App Engine applications.

CPU / Memory Usage:

IP Address	CPU % Usage	Mem % Usage
ec2-75-101-227-238.compute-1.amazonaws.com	4.3	7.17
ec2-75-101-245-66.compute-1.amazonaws.com	4.1	6.62

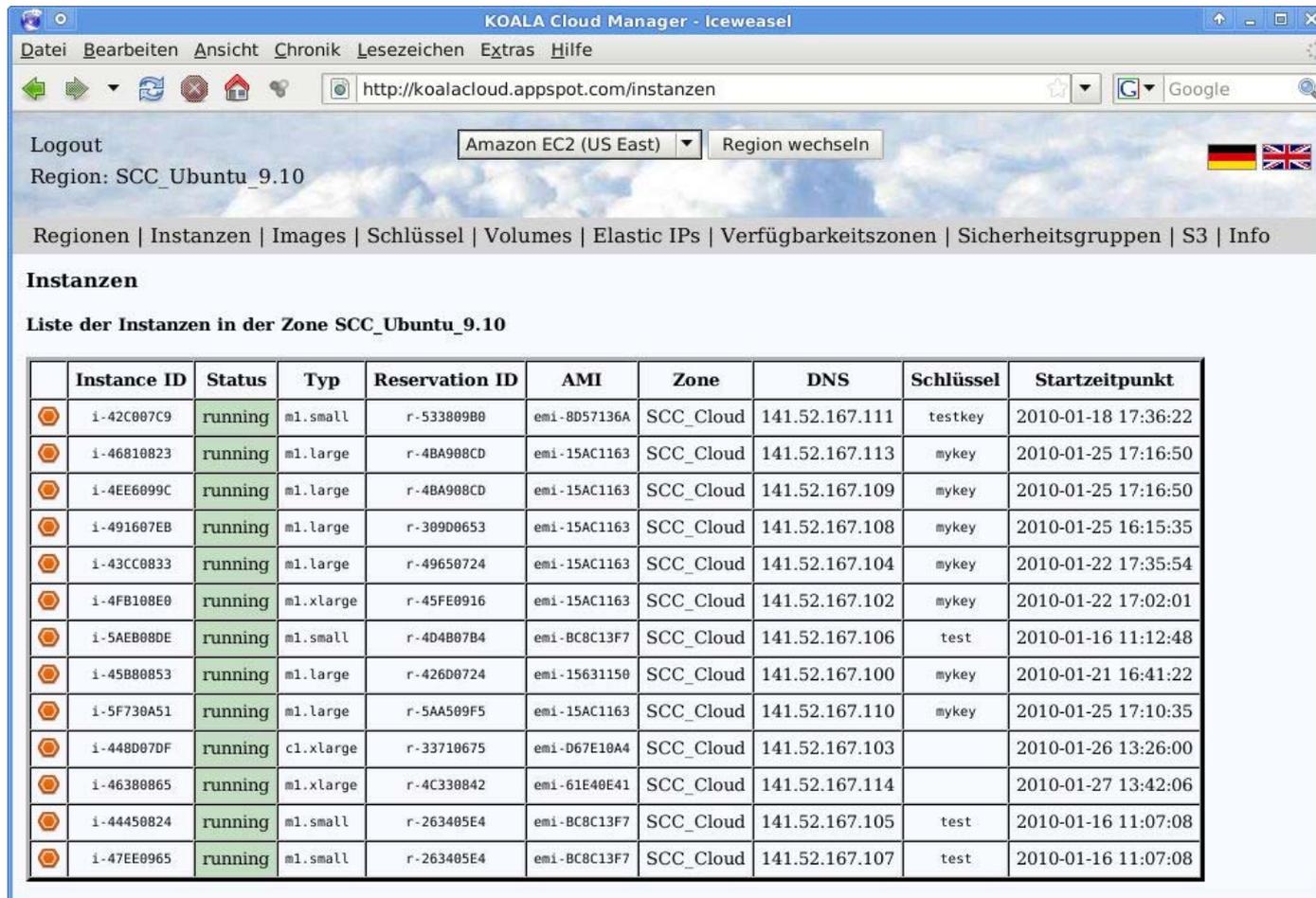
Database Information:

- Powered by voldemort
- Data is replicated 1 times

Applications Hosted in This Cloud:

- [guestbook](#)
- [petlog](#)

Web-Anwendung zur Steuerung von Public und Private Cloud Infrastrukturen via PaaS (aus der Cloud heraus)



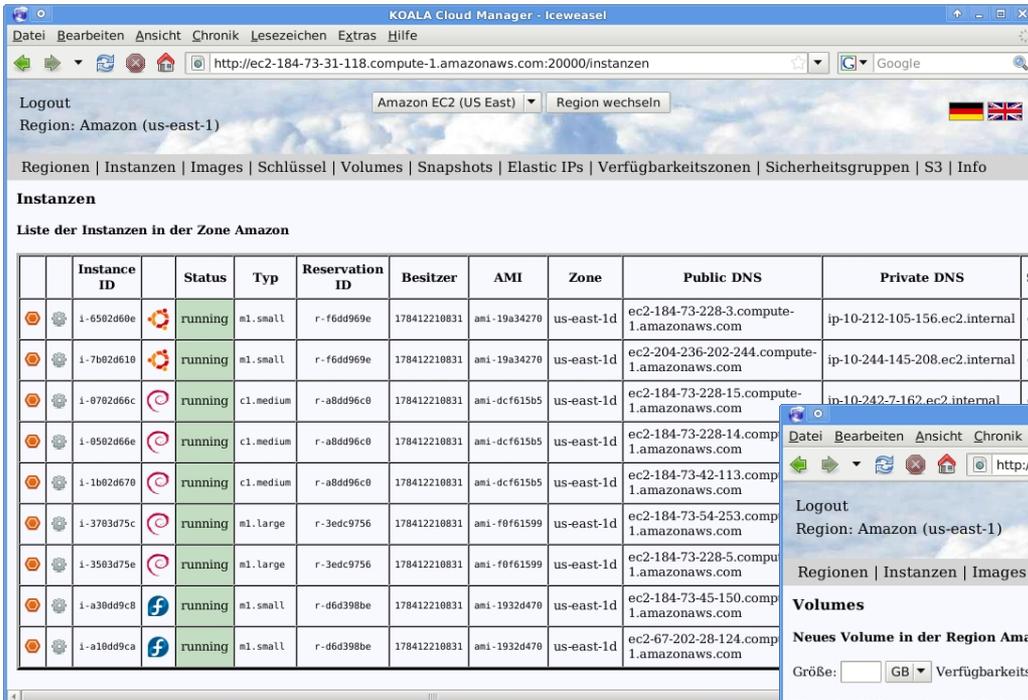
The screenshot shows the KOALA Cloud Manager interface in a browser window. The page title is "KOALA Cloud Manager - Iceweasel". The browser address bar shows "http://koalacloud.appspot.com/instanzen". The interface includes a navigation menu with options like "Logout", "Region: SCC_Ubuntu_9.10", and "Region wechseln". Below the navigation, there is a list of instance types and a table of instances.

Instance ID	Status	Typ	Reservation ID	AMI	Zone	DNS	Schlüssel	Startzeitpunkt
i-42C007C9	running	m1.small	r-533809B0	emi-8D57136A	SCC_Cloud	141.52.167.111	testkey	2010-01-18 17:36:22
i-46810823	running	m1.large	r-4BA908CD	emi-15AC1163	SCC_Cloud	141.52.167.113	mykey	2010-01-25 17:16:50
i-4EE6099C	running	m1.large	r-4BA908CD	emi-15AC1163	SCC_Cloud	141.52.167.109	mykey	2010-01-25 17:16:50
i-491607EB	running	m1.large	r-309D0653	emi-15AC1163	SCC_Cloud	141.52.167.108	mykey	2010-01-25 16:15:35
i-43CC0833	running	m1.large	r-49650724	emi-15AC1163	SCC_Cloud	141.52.167.104	mykey	2010-01-22 17:35:54
i-4FB108E0	running	m1.xlarge	r-45FE0916	emi-15AC1163	SCC_Cloud	141.52.167.102	mykey	2010-01-22 17:02:01
i-5AEB08DE	running	m1.small	r-4D4807B4	emi-BC8C13F7	SCC_Cloud	141.52.167.106	test	2010-01-16 11:12:48
i-45B00853	running	m1.large	r-426D0724	emi-15631150	SCC_Cloud	141.52.167.100	mykey	2010-01-21 16:41:22
i-5F730A51	running	m1.large	r-5AA509F5	emi-15AC1163	SCC_Cloud	141.52.167.110	mykey	2010-01-25 17:10:35
i-448D07DF	running	c1.xlarge	r-33710675	emi-D67E10A4	SCC_Cloud	141.52.167.103		2010-01-26 13:26:00
i-46300865	running	m1.xlarge	r-4C330842	emi-61E40E41	SCC_Cloud	141.52.167.114		2010-01-27 13:42:06
i-44450824	running	m1.small	r-263405E4	emi-BC8C13F7	SCC_Cloud	141.52.167.105	test	2010-01-16 11:07:08
i-47EE0965	running	m1.small	r-263405E4	emi-BC8C13F7	SCC_Cloud	141.52.167.107	test	2010-01-16 11:07:08

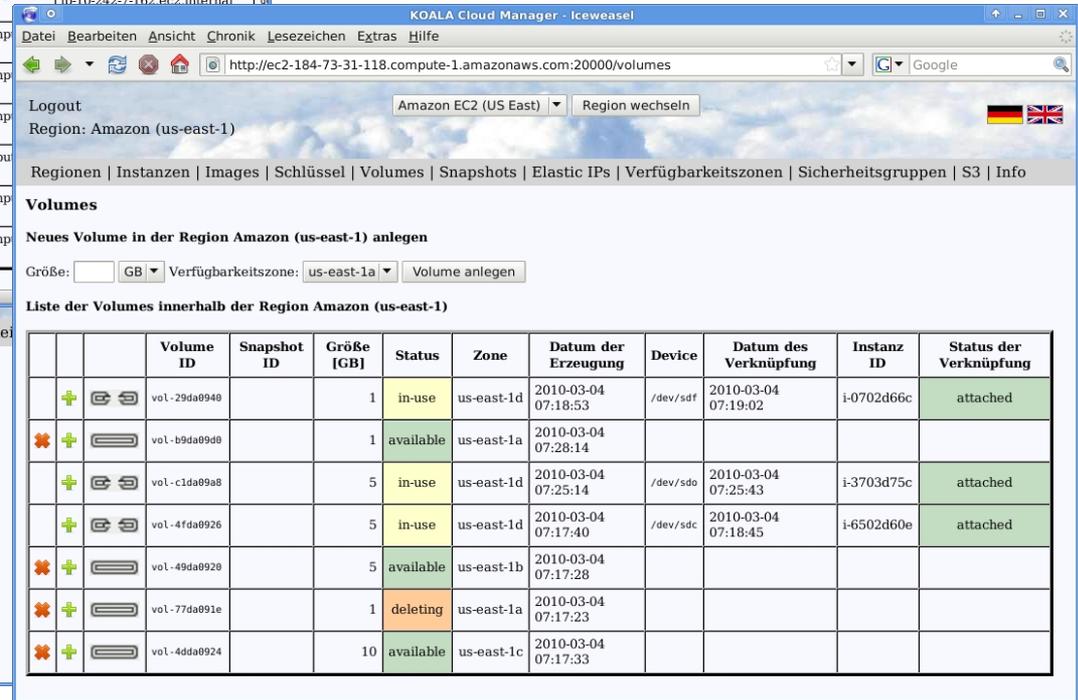
- **Vor- und Nachteile einer Cloud-basierten Steuerung für IaaS gegenüber einer lokalen Lösung (z.B. Elasticfox oder Hybridfox)**
 - **Vorteile:**
 - Flexibilität bzgl. des Browsers
 - Unterstützung von EC2/S3/EBS und Eucalyptus
 - Keine lokale Installation notwendig (Ausnahme: Private Schlüssel)
 - Cloud Installation entspricht dem Cloud-Gedanken
 - **Nachteile:**
 - Vertrauen gegenüber dem Anbieter (Google) bezüglich Datenschutz und Verfügbarkeit ist notwendig

- **Vor- und Nachteile von KOALA gegenüber den Amazon-Werkzeugen (insb. AWS Management Console)**
 - **Vorteile:**
 - Unterstützung von EC2/S3/EBS und Eucalyptus
 - KOALA selbst kann in einer Private Cloud (AppScale) betrieben werden
 - **Nachteile:**
 - Nicht alle EC2/S3/EBS-Features implementiert
 - Kein Support von Amazon

- AppScale läuft in EC2
- KOALA läuft in AppScale
- Die Cloud aus sich selbst heraus zu steuern
 - Zugangsdaten müssen nicht bei einem anderen Anbieter (Google) gespeichert werden



Instance ID	Status	Typ	Reservation ID	Besitzer	AMI	Zone	Public DNS	Private DNS
i-6502d66e	running	m1.small	r-f6dd969e	178412210831	ami-19a34270	us-east-1d	ec2-184-73-228-3.compute-1.amazonaws.com	ip-10-212-105-156.ec2.internal
i-7b02d610	running	m1.small	r-f6dd969e	178412210831	ami-19a34270	us-east-1d	ec2-204-236-202-244.compute-1.amazonaws.com	ip-10-244-145-208.ec2.internal
i-0702d66c	running	c1.medium	r-abdd96c0	178412210831	ami-dcf615b5	us-east-1d	ec2-184-73-228-15.compute-1.amazonaws.com	ip-10-242-7-162.ec2.internal
i-0502d66e	running	c1.medium	r-abdd96c0	178412210831	ami-dcf615b5	us-east-1d	ec2-184-73-228-14.compute-1.amazonaws.com	
i-1b02d670	running	c1.medium	r-abdd96c0	178412210831	ami-dcf615b5	us-east-1d	ec2-184-73-42-113.compute-1.amazonaws.com	
i-3703d75c	running	m1.large	r-3edc9756	178412210831	ami-f0f61599	us-east-1d	ec2-184-73-54-253.compute-1.amazonaws.com	
i-3503d75e	running	m1.large	r-3edc9756	178412210831	ami-f0f61599	us-east-1d	ec2-184-73-228-5.compute-1.amazonaws.com	
i-a30d9c8	running	m1.small	r-d6d398be	178412210831	ami-1932d470	us-east-1d	ec2-184-73-45-150.compute-1.amazonaws.com	
i-a10d9ca	running	m1.small	r-d6d398be	178412210831	ami-1932d470	us-east-1d	ec2-67-202-28-124.compute-1.amazonaws.com	



Neues Volume in der Region Amazon (us-east-1) anlegen

Größe: GB Verfügbarkeitszone: us-east-1a

Volume ID	Snapshot ID	Größe [GB]	Status	Zone	Datum der Erzeugung	Device	Datum des Verknüpfung	Instanz ID	Status der Verknüpfung
vol-29da0940		1	in-use	us-east-1d	2010-03-04 07:18:53	/dev/sdf	2010-03-04 07:19:02	i-0702d66c	attached
vol-b9da09d0		1	available	us-east-1a	2010-03-04 07:28:14				
vol-c1da09a8		5	in-use	us-east-1d	2010-03-04 07:25:14	/dev/sdo	2010-03-04 07:25:43	i-3703d75c	attached
vol-4fda0926		5	in-use	us-east-1d	2010-03-04 07:17:40	/dev/sdc	2010-03-04 07:18:45	i-6502d66e	attached
vol-49da0920		5	available	us-east-1b	2010-03-04 07:17:28				
vol-77da091e		1	deleting	us-east-1a	2010-03-04 07:17:23				
vol-4dda0924		10	available	us-east-1c	2010-03-04 07:17:33				

Regionen | Instanzen | Images | Schlüssel | Volumes | Snapshots | Elastic IPs | Verfügbarkeitszonen

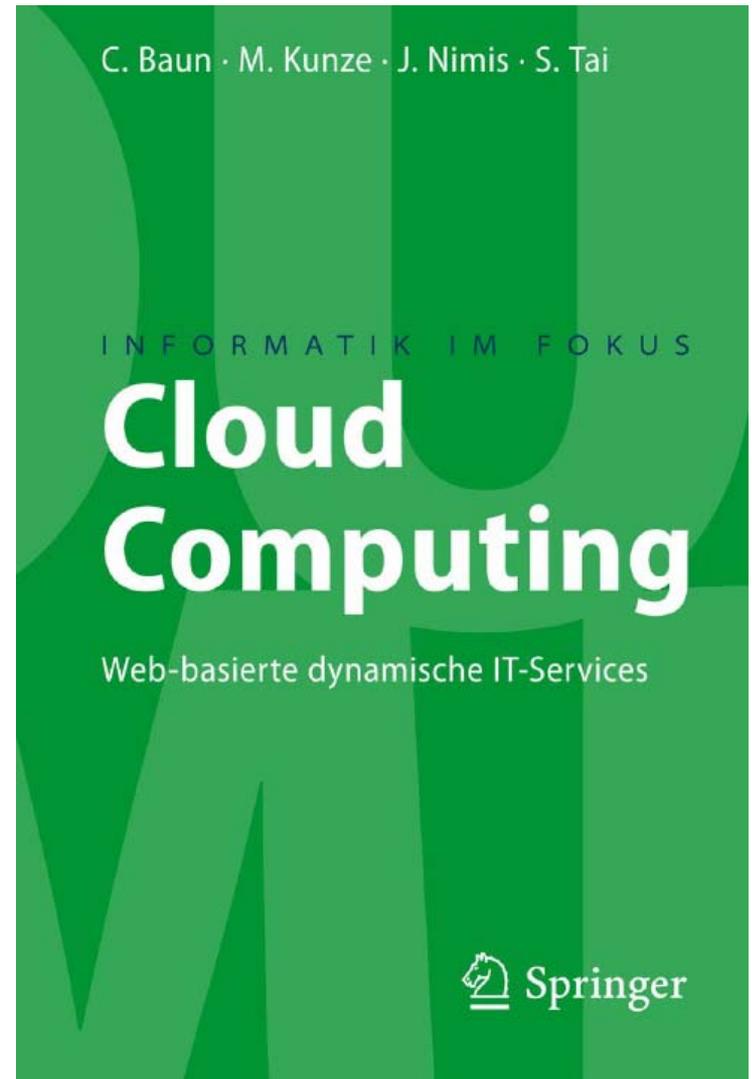
Elastische IPs

Liste der elastischen IPs in der Region Amazon (us-east-1)

Adresse	Instanz ID
184.73.228.3	i-6502d66e
184.73.228.5	i-3503d75e
184.73.228.12	
184.73.228.14	i-0502d66e
184.73.228.15	i-0702d66c

Jetzt verfügbar!

- Auf dem Markt seit November 2009
- Erstes Buch zum Cloud Computing in deutscher Sprache
- Nur €14,95



Vielen Dank für Ihre Aufmerksamkeit

