

Cluster-, Grid- und Cloud-Computing (CGC)

15.7.2010

M.Sc. Christian Baun

Aufgabe 1 (4 Punkte)

In der Vorlesung wurde behauptet, dass Cloud-Computing und Grid-Computing nicht das gleiche sind. Nennen Sie vier offensichtliche Unterschiede zwischen einer Cloud und einem Grid.

Aufgabe 2 (5+5 Punkte)

Im LHC Computing Grid fallen jährlich 15 Petabyte Daten an, die gespeichert werden müssen. Wie hoch wäre ein Stapel, wenn zur Speicherung

- CDs verwendet würden (Kapazität: 625 Megabyte, Dicke: 1,1 mm)?
- Festplatten verwendet werden (Kapazität: 2 Terabyte, Dicke: 2 cm)?

Aufgabe 3 (4+4 Punkte)

- Ordnen Sie die Cloud-Dienste-Kategorien SaaS, IaaS, PaaS und HuaaS den Ebenen zu.
- Ordnen Sie die folgenden kommerziellen Cloud-Angebote den Ebenen zu:
 - Google App Engine
 - Amazon Elastic Compute Cloud
 - Google Docs
 - Amazon Mechanical Turk

Aufgabe 4 (4+2 Punkte)

- Amazon Web Services (AWS)
 - Erklären Sie die beiden Konzepte Availability Zone und Region.
 - Erklären Sie die beiden Konzepte AMI und Instanz.
- Google App Engine (GAE)
 - Erklären Sie die Unterschiede zwischen Datastore und Memcache.

Aufgabe 5 (8 Punkte)

- Sie sollen in der Google App Engine ein Gästebuch in der Programmiersprache Python erzeugen. Glücklicherweise haben Sie eine solche Applikation bereits. Diese ist nicht nur in der Lage, neue Gästebucheinträge aufzunehmen, sondern auch die bereits vorhandenen Einträge zu löschen. Leider sind wegen eines Dateisystemfehlers ein paar Zeilen verloren gegangen. Glücklicherweise konnte die Zeilen wieder hergestellt werden. Ihre Position aber nicht. Fügen Sie in den Listings die Nummern der fehlenden Zeilen ein.

Aufgabe 6 (5+5 Punkte)

Ein wissenschaftliches Experiment erzeugt 20 Petabyte Daten pro Jahr.

- Wie lange dauert die Übertragung in einem Ethernet (LAN) mit 1000 Mbit pro Sekunde?
- Wie lange dauert die Übertragung über ADSL mit 16.000 kbit pro Sekunde?

Aufgabe 7 (2+2 Punkte)

Berechnen Sie Stromkosten. Was kostete der jährliche Dauerbetrieb (24/7) bei 0,23 €/kWh für ein Unternehmen XY mit 500 Computerarbeitsplätzen?:

a) Szenario 1: Fat Clients (PC)

- Elektrische Anschlussleistung pro Desktopsystem: 300 Watt
- Elektrische Anschlussleistung pro Bildschirm: 100 Watt

b) Szenario 2: Thin Clients mit Blade-Servern

- Elektrische Anschlussleistung pro Thin Client: 25 Watt
- Elektrische Anschlussleistung pro Bildschirm: 100 Watt
- Elektrische Anschlussleistung pro Server-Blade: 150 Watt
- Auf ein Server-Blade passen 25 virtuelle Desktopsysteme

Bedenken Sie, dass es auch Schaltjahre gibt!

Name:

Vorname:

Matr.Nr.:

Aufgabe 1)

Punkte:

Grid-Computing	Cloud-Computing
Geographisch verteilte, heterogene Ressourcen ohne zentrale Kontrolle	Ein oder wenige Rechenzentren mit heterogenen oder homogenen Ressourcen unter zentraler Kontrolle
Physische Ressourcen	Virtualisierte Ressourcen
Schwach automatisierte Dienste (<i>klassische IT</i>)	Vollautomatisierte Dienste (Industrialisierung der IT)
Lange Einarbeitungszeit notwendig (nicht besonders benutzerfreundlich)	Benutzerfreundlich
Hohe Einstiegshürden	Geringe Einstiegshürden
Einsatz primär in Wissenschaft und Forschung (z.B. LHC)	Einsatz primär in Industrie und <i>Startups</i>
Prinzip der Virtuellen Organisationen	Keine Virtuellen Organisationen
Grids basieren auf OpenSource Software	Public Clouds bestehen aus proprietärer Software. Private Clouds bestehen aus OpenSource Software
Förderung durch die öffentliche Hand. Kostenfreie Nutzung der Ressourcen nach Zustimmung durch die Ressourcenbetreiber	Nutzung kostenfrei oder verbrauchsabhängige Abrechnung (Pay-as-you-go)

Name:

Vorname:

Matr.Nr.:

Aufgabe 3)

Punkte:

HuaaS	Amazon Mechanical Turk
SaaS	Google Docs
PaaS	Google App Engine
IaaS	Amazon Elastic Compute Cloud
	Hardware

Name:

Vorname:

Matr.Nr.:

Aufgabe 4)

Punkte:

- **Amazon Web Services (AWS)**

- EC2 hat aktuell 4 Standorte, in denen sich Ressourcen befinden. Die Standorte werden als Regionen bezeichnet
- Jeder Standort enthält Availability Zonen (Verfügbarkeitszonen). Jede Verfügbarkeitszone ist wie ein in sich abgeschlossener Cluster
- Ein Amazon Machine Images (AMI) ist eine Art Blaupause für das Anlegen eines neuen virtuellen Servers.
- Eine Instanz ist ein virtueller Server, der auf den Serverfarmen von Amazon läuft. Diese virtuellen Server werden aus AMI erzeugt.

- **Google App Engine (GAE)**

- Der Datastore ist ein persistenter Speicher, der als Key/Value-Datenbank realisiert ist. Transaktionen sind atomar. Definition, Abfrage und Manipulation von Daten erfolgt über eine eigene Sprache, die GQL (Google Query Language).
- Der Memcache ist ein hochperformanter temporärer Datenspeicher aus Hauptspeicher. Der Memcache hat sehr gute Zugriffszeiten. Jeder Eintrag wird mit einem eindeutigen Schlüssel abgelegt und ist auf 1 MB beschränkt. Es wird eine Verfallszeit in Sekunden angegeben, wann der Eintrag aus dem Memcache entfernt werden soll. Daten werden je nach Auslastung des Mamcache früher wieder verdrängt

Name:

Vorname:

Matr.Nr.:

Aufgabe 5)

Punkte:

Datei app.yaml

```
application: einfaches_gaestebuch

version: 1
[1]
[15]

handlers:
- url: .*
  script: main.py
```

Datei main.py

```
#!/usr/bin/env python
# -*- coding: iso-8859-15 -*-
import os, sys
[11]
from google.appengine.ext import db
from google.appengine.ext import webapp
from google.appengine.ext.webapp import template

[4]
message = db.StringProperty(required=True)
when = db.DateTimeProperty(auto_now_add=True)
[12]

class MyHandler(webapp.RequestHandler):
    def get(self):
        shouts = db.GqlQuery('SELECT * FROM gaestebuch ORDER BY when DESC')
        values = {'shouts': shouts}
        self.response.out.write(template.render('main.html', values))
[5]
        shout = gaestebuch(message=self.request.get('message'), who=self.request.get('who'))
        shout.put()
[3]

class Loeschen(webapp.RequestHandler):
[2]
    alles_loeschen_query = gaestebuch.all(keys_only=True)
    alles_loeschen = alles_loeschen_query.fetch(300)
[14]
    self.redirect('/')

app = webapp.WSGIApplication([('/', MyHandler),
[6]
                             debug=True)

def main():
    wsgiref.handlers.CGIHandler().run(app)

[13]
main()
```


Name:

Vorname:

Matr.Nr.:

Aufgabe 5 – Fortsetzung) Punkte:

Datei main.html

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
    "http://www.w3.org/TR/html4/loose.dtd">

<html>
<head>
<title>Einfaches G&auml;stebuch mit Google App Engine</title>
</head>
<body>

[16]
<div>
    {% if shout.who %}
        <b>{{shout.who}}</b>
    {% else %}
        <b>Anonymous</b>
    [7]
    sagt:
    [8]
</div>
{% endfor %}

<p>&nbsp;</p>

[10]
<table border="0" cellspacing="5" cellpadding="5">
    <tr>
        <td align="right">Name:</td>
        <td colspan="2" align="left"><input type="text" size="40"
            name="who" value="" if="who"></td>
    </tr>
    <tr>
        <td align="right">Nachricht:</td>
        <td colspan="2" align="left"><input type="text" size="40"
            name="message" value="" if="message"></td>
    </tr>
    <tr>
        <td>&nbsp;</td>
        <td align="center"><input type="submit" value="Absenden"></td>
        <td align="center"><input type="reset" value="L&ouml;schen"></td>
    </tr>
</table>
</form>

<p>&nbsp;</p>

<form action="loeschen" method="post" accept-charset="utf-8">
<table border="0" cellspacing="5" cellpadding="5">
    <tr>
        [9]
    </tr>
</table>
</form>
</body>
</html>
```

Name:

Vorname:

Matr.Nr.:

Aufgabe 6)

Punkte:

a)

1 Gigabit (Gbit) = 10^9 Bit	1.000.000.000 Bit
Daten im Experiment	20.000.000.000.000.000 Byte
Bandbreite des Ethernet (1 Gbit/s)	1.000.000.000 Bit/s
Bandbreite des Ethernet in Byte/s	125.000.000 Byte/s
Dauer der Datenübertragung [s]	= 160.000.000 : 60
Dauer der Datenübertragung [m]	$\approx 2.666.666,67 : 60$
Dauer der Datenübertragung [h]	$\approx 44.444,45 : 24$
Dauer der Datenübertragung [d]	$\approx 1.851,85 : 365,25$
Dauer der Datenübertragung [y]	$\approx 5,07$

b)

1 Megabit (Mbit) = 10^6 Bit	1.000.000 Bit
Daten im Experiment	20.000.000.000.000.000 Byte
Bandbreite des ADSL (16 Mbit/s)	16.000.000 Bit/s
Bandbreite des Ethernet in Byte/s	2.000.000 Byte/s
Dauer der Datenübertragung [s]	= 10.000.000.000 : 60
Dauer der Datenübertragung [m]	$\approx 166.666.666,67 : 60$
Dauer der Datenübertragung [h]	$\approx 2.777.777,78 : 24$
Dauer der Datenübertragung [d]	$\approx 115.740,7408 : 365,25$
Dauer der Datenübertragung [y]	$\approx 316,88$

Name:

Vorname:

Matr.Nr.:

Aufgabe 7)

Punkte:

a)

- Stromkosten (mit Schaltjahr) pro Computerarbeitsplatz pro Jahr

$$0,4 \text{ kW} * 24 \frac{\text{h}}{\text{Tag}} * 365,25 \frac{\text{Tag}}{\text{Jahr}} * 0,23 \frac{\text{€}}{\text{kWh}} = 806,472 \frac{\text{€}}{\text{Jahr}}$$

- Stromkosten für 500 Computerarbeitsplätze pro Jahr

$$500 * 806,472 \frac{\text{€}}{\text{Jahr}} = 403236 \frac{\text{€}}{\text{Jahr}}$$

b)

- Stromkosten (mit Schaltjahr) pro Computerarbeitsplatz (ohne Server) pro Jahr

$$0,125 \text{ kW} * 24 \frac{\text{h}}{\text{Tag}} * 365,25 \frac{\text{Tag}}{\text{Jahr}} * 0,23 \frac{\text{€}}{\text{kWh}} = 252,0225 \frac{\text{€}}{\text{Jahr}}$$

- Stromkosten (mit Schaltjahr) pro Server-Blade pro Jahr

$$0,15 \text{ kW} * 24 \frac{\text{h}}{\text{Tag}} * 365,25 \frac{\text{Tag}}{\text{Jahr}} * 0,23 \frac{\text{€}}{\text{kWh}} = 302,427 \frac{\text{€}}{\text{Jahr}}$$

- Stromkosten für 500 Computerarbeitsplätze (ohne Server) pro Jahr

$$500 * 252,0225 \frac{\text{€}}{\text{Jahr}} = 126011,25 \frac{\text{€}}{\text{Jahr}}$$

- Stromkosten für 20 Server-Blades pro Jahr

$$20 * 302,427 \frac{\text{€}}{\text{Jahr}} = 6048,54 \frac{\text{€}}{\text{Jahr}}$$

- Stromkosten für Computerarbeitsplätze und Server-Blades pro Jahr

$$126011,25 \frac{\text{€}}{\text{Jahr}} = +6048,54 \frac{\text{€}}{\text{Jahr}} = 132059,79 \frac{\text{€}}{\text{Jahr}}$$