

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

25.1.2012

Dr. Christian Baum

Aufgabe 1 (4+4 Punkte)

a) Ordnen Sie die folgenden Cloud-Dienste-Kategorien den Ebenen in der Abbildung zu

- PaaS
- Cloud-Gaming
- Cloud-Printing
- IaaS
- HPCaaS
- HuaaS
- Cloud-Betriebssystem
- SaaS

b) Ordnen Sie die folgenden freien und kommerziellen Cloud-Angebote den Ebenen zu:

- Google App Engine
- Google Cloud Print
- Amazon Elastic Compute Cloud
- Amazon Mechanical Turk
- eyeOS
- EC2 Cluster Compute Instances
- Google Apps
- OnLive

Aufgabe 2 (4+2 Punkte)

- a) Nennen Sie die vier HTTP-Methoden bei REST Web Services, die an die aus dem Datenbank-Umfeld bekannten CRUD-Aktionen erinnern und beschreiben Sie kurz deren Funktion.
- b) Zusätzlich zu den vier HTTP-Methoden werden zwei weitere HTTP-Methoden häufig bei Cloud-Diensten angeboten. Nennen Sie diese und beschreiben Sie kurz deren Funktion.

Aufgabe 3 (4 Punkte)

Ordnen Sie die Eigenschaften in der Tabelle jeweils dem Cloud-Computing oder dem Grid-Computing zu. (Es genügt, wenn Sie jeweils „C“ für Cloud Computing und „G“ für Grid Computing eintragen.)

Eigenschaft	Cloud/Grid Computing
Verteilte, heterogene Ressourcen ohne zentrale Kontrolle	
Benutzerfreundliche Bedienung	
Vollautomatisierte Dienste	
Basiert auf freier, standardisierter Software und Schnittstellen	
Finanzierung primär durch Förderung durch die öffentliche Hand	
Verbrauchsabhängige Abrechnung	
Hauptsächlich physische Ressourcen	
Hauptsächlich virtualisierte Ressourcen	

Für jede korrekte Antwort gibt es 0.5 Punkte. Für jede falsche Antwort werden 0.5 Punkte abgezogen. Es können maximal 4 Punkte und nicht weniger als 0 Punkte insgesamt erreicht werden.

Aufgabe 4 (4+2 Punkte)

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

Aufgabe 5 (1+3+3 Punkte)

- a) Worin unterscheiden sich Peer-to-Peer und das Client-Server-Modell?
- b) Nennen Sie die Namen der drei Arten von Peer-to-Peer-Systemen.
- c) Beschreiben Sie in wenigen Worten, was jede der drei Arten von Peer-to-Peer-Systemen auszeichnet.

Aufgabe 6 (4 Punkte)

Kreuzen Sie bei jeder Aussage in der Tabelle an, ob sie wahr oder falsch ist.

Aussage	wahr	falsch
IBM Smart Cloud Enterprise ist eine „Infrastructure as a Service“		
Buckets in S3 haben einen hierarchischen Namensraum		
Die Google App Engine ist eine „Platform as a Service“		
Google Cloud Storage hat die gleiche Schnittstelle wie EBS		
Man kann die Firewall-Einstellungen von EC2-Instanzen mit Hilfe von Sicherheitsgruppen konfigurieren		
EBS-Volumen können zu jedem Zeitpunkt nur an eine Instanz angehängt sein		
EBS-Volumen dürfen nur das Dateisystem ext3 enthalten		
Ein verteiltes System auf Basis von BOINC ist eine „Infrastructure as a Service“		

Für jede korrekte Antwort gibt es 0.5 Punkte. Für jede falsche Antwort werden 0.5 Punkte abgezogen. Es können maximal 4 Punkte und nicht weniger als 0 Punkte insgesamt erreicht werden.

Aufgabe 7 (5+2 Punkte)

- a) Berechnen Sie die Werte der Fingertable von Knoten $n = 8$ und tragen Sie diese in die Tabelle ein.
- b) Welche beiden Formen der Suche gibt es bei verteilten Hashtabellen?

Aufgabe 8 (2+2 Punkte)

10 TB Daten sollen aus einer Cloud exportiert werden.

- a) Wie lange dauert die Übertragung via Ethernet (LAN) mit 10 Gbit/s?
- b) Wie lange dauert die Übertragung via DSL mit 16.000 Kbit/s?

Aufgabe 9 (2+1+1 Punkte)

- a) Beschreiben Sie den Unterschied zwischen **Virtualisierung** und **Emulation**.
- b) Nennen Sie ein Beispiel für Anwendungsvirtualisierung.
- c) Beschreiben Sie die Funktion des **VMM** bei vollständiger Virtualisierung.

Name:

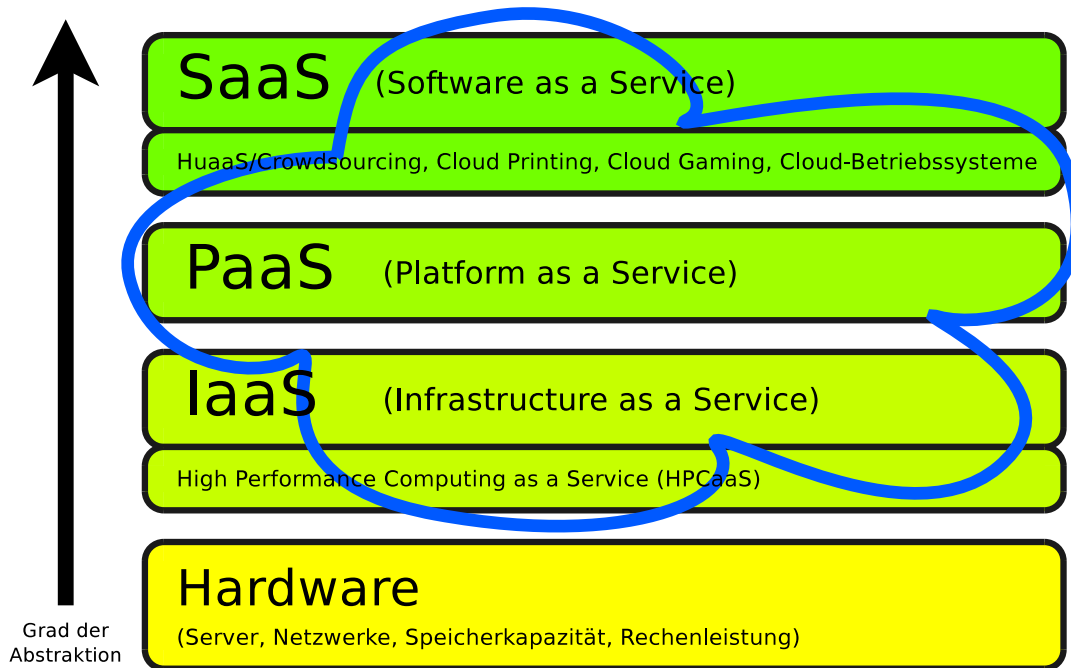
Vorname:

Matr.Nr.:

Aufgabe 1)

Punkte:

a)



b)

- | | | |
|-------------------------------|---|----------------------|
| Google App Engine | ⇒ | PaaS |
| Google Cloud Print | ⇒ | Cloud-Printing |
| Amazon Elastic Compute Cloud | ⇒ | IaaS |
| Amazon Mechanical Turk | ⇒ | HuaaS |
| eyeOS | ⇒ | Cloud-Betriebssystem |
| EC2 Cluster Compute Instances | ⇒ | HPCaaS |
| Google Apps | ⇒ | SaaS |
| OnLive | ⇒ | Cloud-Gaming |

Name:

Vorname:

Matr.Nr.:

Aufgabe 2)

Punkte:

a)

HTTP	CRUD-Aktionen	SQL	Beschreibung
PUT	Create	INSERT	Ressource erzeugen oder deren Inhalt ersetzen
GET	Read/Retrieve	SELECT	Ressource bzw. deren Repräsentation anfordern
POST	Update	UPDATE	Einer Ressource etwas hinzufügen
DELETE	Delete/Destroy	DELETE	Ressource löschen

b)

- HEAD fordert vom Server nur den Header einer Ressource (Datei) an
 - So kann sich der Benutzer des Web-Service über die Metadaten informieren, ohne die eigentlichen Ressource zu übertragen
 - Es wird der gleiche Header zurückgeliefert wie bei GET
- OPTIONS prüft welche Methoden auf einer Ressource verfügbar sind

Name:

Vorname:

Matr.Nr.:

Aufgabe 3)

Punkte:

Eigenschaft	Cloud/Grid Computing
Verteilte, heterogene Ressourcen ohne zentrale Kontrolle	Grid Computing
Benutzerfreundliche Bedienung	Cloud Computing
Vollautomatisierte Dienste	Cloud Computing
Basiert auf freier, standardisierter Software und Schnittstellen	Grid Computing
Finanzierung primär durch Förderung durch die öffentliche Hand	Grid Computing
Verbrauchsabhängige Abrechnung	Cloud Computing
Hauptsächlich physische Ressourcen	Grid Computing
Hauptsächlich virtualisierte Ressourcen	Cloud Computing

Name:

Vorname:

Matr.Nr.:

Aufgabe 4)

Punkte:

a)

- EC2 besteht aus Standorten (**Regionen**), mit Ressourcen. Jeder Standort enthält Verfügbarkeitszonen (**Availability Zones**). Jede Verfügbarkeitszone ist ein in sich abgeschlossener Cluster.
- Virtuelle Server (Instanzen) werden aus Amazon Machine Images (AMI) erzeugt. Ein AMI ist eine Blaupause für das Anlegen eines neuen virtuellen Servers.

b)

- **Datastore**

- Persistenter Speicher, realisiert als Key/Value-Datenbank
- Transaktionen sind atomar
- Definition, Abfrage und Manipulation von Daten erfolgt über eine eigene Sprache, die GQL (Google Query Language)
 - * GQL hat große Ähnlichkeiten mit SQL (Structured Query Language)

- **Memcache**

- Hochperformanter temporärer Datenspeicher aus Hauptspeicher
- Sehr gute Zugriffszeiten
- Jeder Eintrag wird mit einem eindeutigen Schlüssel abgelegt
- Jeder Eintrag 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 Memcache früher wieder verdrängt

Name:

Vorname:

Matr.Nr.:

Aufgabe 5)

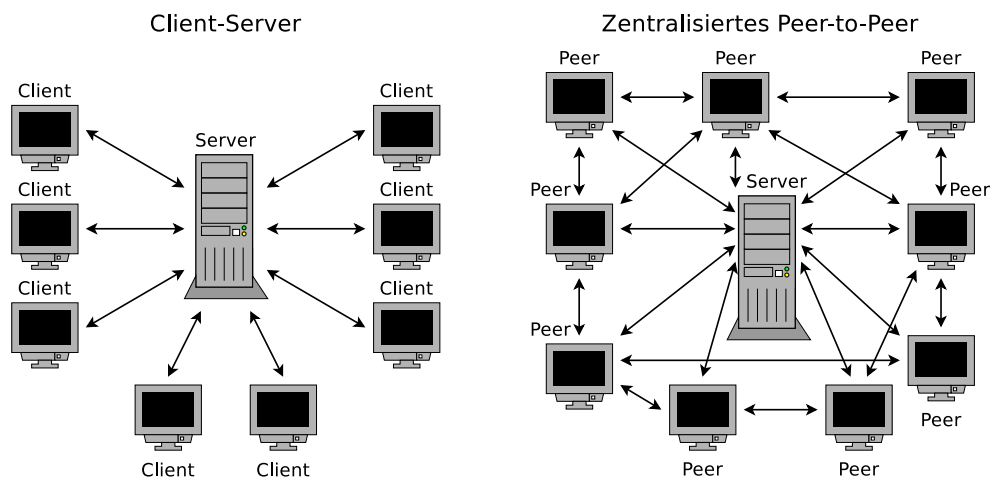
Punkte:

a)

- Ein Peer-to-Peer-System ist ein Verbund gleichberechtigter Knoten
 - Knoten werden als Peers bezeichnet
 - Knoten machen sich gegenseitig Ressourcen zugänglich
 - Jeder Knoten ist gleichzeitig Client und Server

„A Peer-to-Peer system is a self-organizing system of equal, autonomous entities (peers) which aims for the shared usage of distributed resources in a networked environment avoiding central services.“

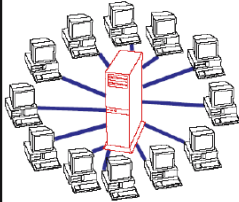
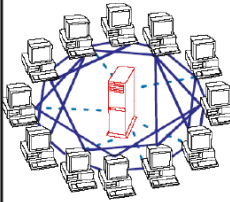
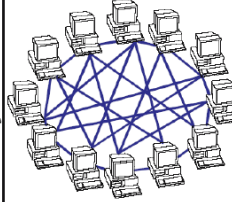
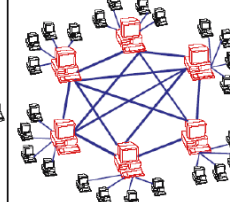
(Andy Oram)



b)

- Zentralisiertes P2P
- Reines/Pures P2P
- Hybrides P2P

c)

Client-Server	Peer-to-Peer		
	1. Resources are shared between the peers 2. Resources can be accessed directly from other peers 3. Peer is provider and requestor (Servent concept)		
	1st Generation		2nd Generation
1. Server is the central entity and only provider of service and content. → Network managed by the Server 2. Server as the higher performance system. 3. Clients as the lower performance system Example: WWW	Centralized P2P	Pure P2P	Hybrid P2P
	1. All features of Peer-to-Peer included 2. Central entity is necessary to provide the service 3. Central entity is some kind of index/group database Example: Napster	1. All features of Peer-to-Peer included 2. Any terminal entity can be removed without loss of functionality 3. → No central entities Examples: Gnutella 0.4, Freenet	1. All features of Peer-to-Peer included 2. Any terminal entity can be removed without loss of functionality 3. → dynamic central entities Example: Gnutella 0.6, JXTA
			

Quelle: Jörg Eberspächer und Rüdiger Schollmeier. *First and Second Generation of Peer-to-Peer Systems* (2005). LNCS 3485

Name:

Vorname:

Matr.Nr.:

Aufgabe 6)

Punkte:

Aussage	wahr	falsch
IBM Smart Cloud Enterprise ist eine „Infrastructure as a Service“	X	
Buckets in S3 haben einen hierarchischen Namensraum		X
Die Google App Engine ist eine „Plattform as a Service“	X	
Google Cloud Storage hat die gleiche Schnittstelle wie EBS		X
Man kann die Firewall-Einstellungen von EC2-Instanzen mit Hilfe von Sicherheitsgruppen konfigurieren	X	
EBS-Volumen können zu jedem Zeitpunkt nur an eine Instanz angehängt sein	X	
EBS-Volumen dürfen nur das Dateisystem ext3 enthalten		X
Ein verteiltes System auf Basis von BOINC ist eine „Infrastructure as a Service“		X

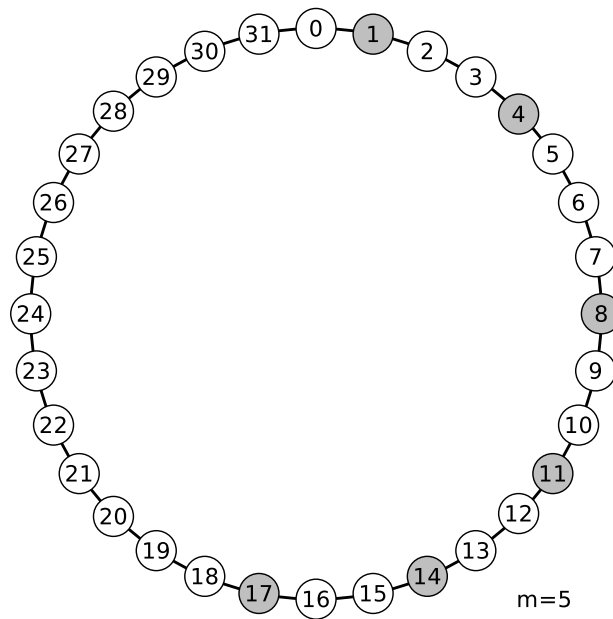
Name:

Vorname:

Matr.Nr.:

Aufgabe 7)

Punkte:



a)

Fingertable von Knoten $n = 8$

Eintrag	Start	Knoten
1	9	11
2	10	11
3	12	14
4	16	17
5	24	1

Die Tabelle hat 5 Einträge, weil m die Länge der ID in Bit ist und $m = 5$

b) Es gibt **lineare Suche** und **binäre Suche**

Name:

Vorname:

Matr.Nr.:

Aufgabe 8)

Punkte:

a)

Daten in der Cloud (10 TB) 10.000.000.000.000 Byte

Bandbreite des Ethernet (10 Gbit/s) 10.000.000.000 Bit/s

Bandbreite des Ethernet in Byte/s 1.250.000.000 Byte/s

$$10.000.000.000.000 \text{ Byte} / 1.250.000.000 \text{ Byte/s} = 8.000 \text{ s}$$

Dauer der Datenübertragung [s] = 8.000 : 60

Dauer der Datenübertragung [min] = 133, $\bar{3}$

⇒ ca. 2 Stunden, 13 Minuten

b)

Daten in der Cloud (10 TB) 10.000.000.000.000 Byte

Bandbreite des DSL (16.000 Kbit/s) 16.000.000 Bit/s

Bandbreite des DSL in Byte/s 2.000.000 Byte/s

$$10.000.000.000.000 \text{ Byte} / 2.000.000 \text{ Byte/s} = 5.000.000 \text{ s}$$

Dauer der Datenübertragung [s] = 5.000.000 : 60

Dauer der Datenübertragung [min] = 83.333, $\bar{3}$: 60

Dauer der Datenübertragung [h] = 1.388, $\bar{8}$: 24

Dauer der Datenübertragung [d] ≈ 57,87

⇒ ca. 57 Tage, 20 Stunden, 53 Minuten

Name:

Vorname:

Matr.Nr.:

Aufgabe 9)

Punkte:

a)

- Emulation bildet die **komplette Hardware** eines Rechnersystems nach, um ein **unverändertes Betriebssystem**, das für eine **andere Hardwarearchitektur** (CPU) ausgelegt ist, zu betreiben
- Durch **Virtualisierung** werden die Ressourcen eines Rechnersystems aufgeteilt und von mehreren unabhängigen Betriebssystem-Instanzen genutzt

b) Java Virtual Machine (JVM) oder VMware ThinApp

c) Der VMM verteilt die Hardwareressourcen des Rechners an die VMs. Teilweise emuliert der VMM Hardware, die nicht für den gleichzeitigen Zugriff mehrerer Betriebssysteme ausgelegt ist. Den VMM bezeichnet man auch als **Typ-2-Hypervisor**. Der VMM läuft *hosted* als Anwendung unter dem Host-Betriebssystem.