



# Sommersemester 2015

## Proseminar Speicher- und Dateisysteme



# Cloudspeicher

Sven Schmidt am 09. Juli 2015

# Viele halten sich bedeckt, wenn es um die Cloud geht

51% der Amerikaner sind sich sicher:  
Stürmisches Wetter beeinflusst Cloud-Computing!



Abbildung : Symbolbild verändert nach [13], [7]

# Gliederung

- 1 Einleitung
- 2 Über den Begriff
- 3 Warum Cloud?
- 4 Cloud-Architektur
- 5 Schlussteil

# Gliederung

1 Einleitung

**2 Über den Begriff**

3 Warum Cloud?

4 Cloud-Architektur

5 Schlussteil

# Erste Verwendung des Wortes

What's interesting [now] is that there is an emergent new model [...]



Abbildung : Symbolbild [5]

---

Zitat: [23]

# Erste Verwendung des Wortes

What's interesting [now] is that there is an emergent new model [...] It starts with the premise that the data services and architecture should be on servers.



Abbildung : Symbolbild [5]

# Erste Verwendung des Wortes

What's interesting [now] is that there is an emergent new model [...] It starts with the premise that the data services and architecture should be on servers. We call it cloud computing — they should be in a “cloud” somewhere.

— Eric Schmidt, 2006



Abbildung : Symbolbild [5]

# Erste Verwendung des Wortes

What's interesting [now] is that there is an emergent new model [...] It starts with the premise that the **data services and architecture should be on servers**. We call it cloud computing — they should be in a “cloud” somewhere.

— Eric Schmidt, 2006



Abbildung : Symbolbild [5]



# Der Begriff in der Informatik

Rechnernetze, deren Inneres unbedeutend oder unbekannt ist. [15]

— Wikipedia

# Der Begriff

## Wolken

- Diffus
- Nicht greifbar
- Metaphorisch
- Symbolisch
- Verteilt, befinden sich "irgendwo"

## Die Cloud

- Abstrakt
- Nicht greifbar
- Schwer vorstellbar
- Bedeutungsgeladen
- Verteilt, Daten befinden sich "irgendwo"

# Gliederung

1 Einleitung

2 Über den Begriff

**3 Warum Cloud?**

4 Cloud-Architektur

5 Schluss

# Server-Architekturen im Wandel

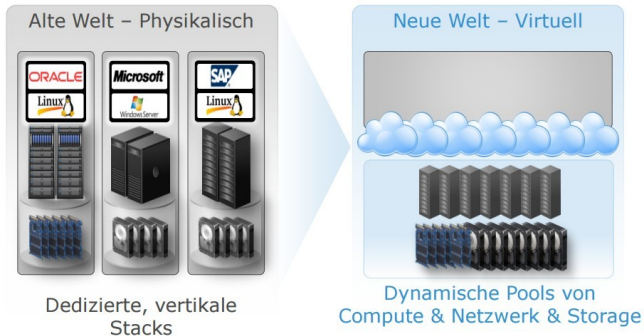


Abbildung : Vergleichsbild von [6]

# Server-Architekturen im Wandel

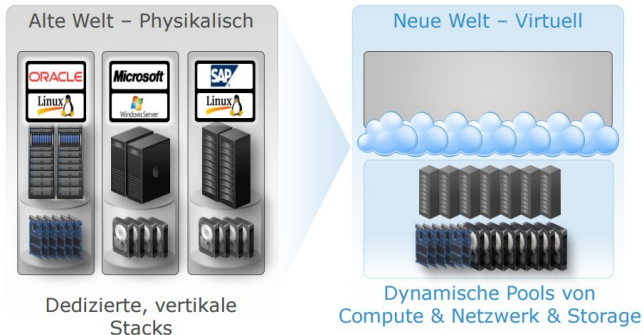
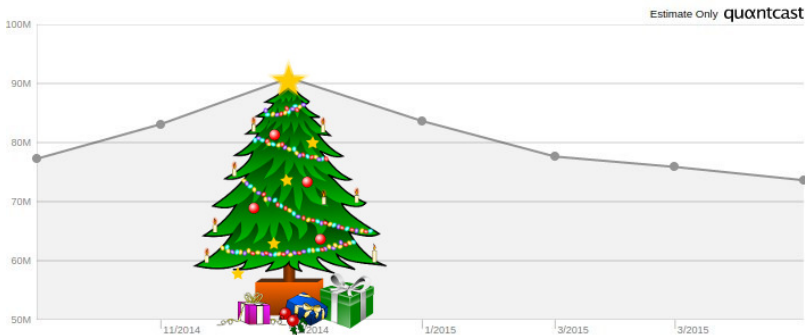


Abbildung : Vergleichsbild von [6]

Das Problem?

# Notwendigkeit für die Cloud

## Amazons “Weihnachts-Problem“



**Abbildung** : Amazon.com Zugriffszahlen 2014/15, Screenshot von [12], Weihnachtsbaum von [3]

# Lösung: Die Cloud



- Aufbau eines großen Ressourcen-Pools
- Verlagern der Infrastruktur in die Cloud
- Vermietung an End-User
- Abrechnung nach Bedarf
- Skalierbar je nach Auslastung

Abbildung : Amazon Web Services [1]

# Die 5 Charakteristika der Cloud

- On-demand self-service  
Provisionierung erfolgt automatisch



# Die 5 Charakteristika der Cloud

- On-demand self-service  
Provisionierung erfolgt automatisch
- Broad network access  
Abruf übers Netz ohne Client-Bindung

# Die 5 Charakteristika der Cloud

- On-demand self-service  
Provisionierung erfolgt automatisch
- Broad network access  
Abruf übers Netz ohne Client-Bindung
- Resource pooling  
Ressourcen liegen in einem Pool zur "Selbstbedienung"

# Die 5 Charakteristika der Cloud

- On-demand self-service  
Provisionierung erfolgt automatisch
- Broad network access  
Abruf übers Netz ohne Client-Bindung
- Resource pooling  
Ressourcen liegen in einem Pool zur "Selbstbedienung"
- Rapid elasticity  
Ressourcen sind skalierbar, erwecken die Illusion von unbegrenzter Kapazität

# Die 5 Charakteristika der Cloud

- **On-demand self-service**  
Provisionierung erfolgt automatisch
- **Broad network access**  
Abruf übers Netz ohne Client-Bindung
- **Resource pooling**  
Ressourcen liegen in einem Pool zur "Selbstbedienung"
- **Rapid elasticity**  
Ressourcen sind skalierbar, erwecken die Illusion von unbegrenzter Kapazität
- **Measured service**  
"Pay as you go "

# Gliederung

1 Einleitung

2 Über den Begriff

3 Warum Cloud?

**4 Cloud-Architektur**

5 Schlussteil

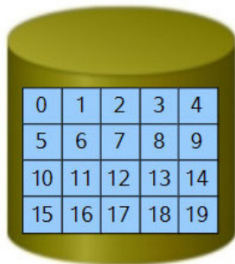
# Das Cloud-Speichersystem: Object-based storage

## Kerneigenschaften

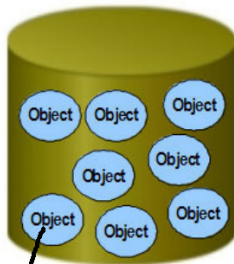
- Ablageform: binäre Objekte (statt Blöcke)
- Flache Organisation, alle Daten liegen auf derselben Ebene
- Objekte beinhalten:
  - Die Daten selbst
  - Meta-Daten (Erstelldatum, Besitzer, Dateityp...)
  - Einen globally unique identifier (GUID)
  - Kurz: Objekt = Daten + Meta-Daten
- Objekte sind selbst-beschreibend
- Objekte sind selbst-beinhaltend

# Objekte statt Blöcke

## Block-Based Disk



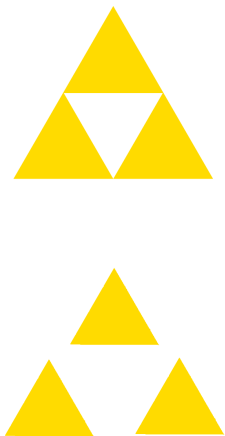
## Object-Based Disk



ID	Binary Data	Metadata
1234	01010101010101001110101010010 01011000010101001110101010010 01011000010101001110101010010	name1 value1 name2 value2 nameN valueN

Abbildung : Vergleich der Speicher-Typen nach [19], [11]

# Object-based storage: Verteilbarkeit der Daten



- Nutzung binärer Dateien
- Problemlose Aufsplittung in Einzelteile
- Partiiell speicherbar
- Mehrere Festplatten, verschiedene Orte
- Datenträger-Grenzen werden uninteressant
- Zusammensetzen in richtiger Reihenfolge ergibt Ursprungsobjekt

Abbildung : Symbolbild nach [14]



# Object-based storage: Zusammensetzen der Daten-Parts



Abbildung : Symbolbild [9]

# Zugriff Über HTTP-Schnittstellen

```
GET /ObjectName HTTP/1.1  
Host: BucketName.s3.amazonaws.com  
Date: date  
Authorization: authorization string  
Range: bytes=byte_range
```



**Abbildung** : Darstellung verändert nach [18], [4], [16]

# Cloud-Dateisysteme

## Was muss geleistet werden?

- Skalierbarkeit
- Daten-Verluste müssen kompensierbar sein
- "Selbst-Heilend"
- Standort der Daten darf keine Rolle spielen
- Latenzzeiten müssen stemmbar sein
- Verschiedene Dateisysteme sinnvoll virtualisieren
- Daten ohne Flaschenhalse verteilen
- Beispiele: Lustre und Ceph File System

# Beispiel: Lustre



## Lustre (Simplified)

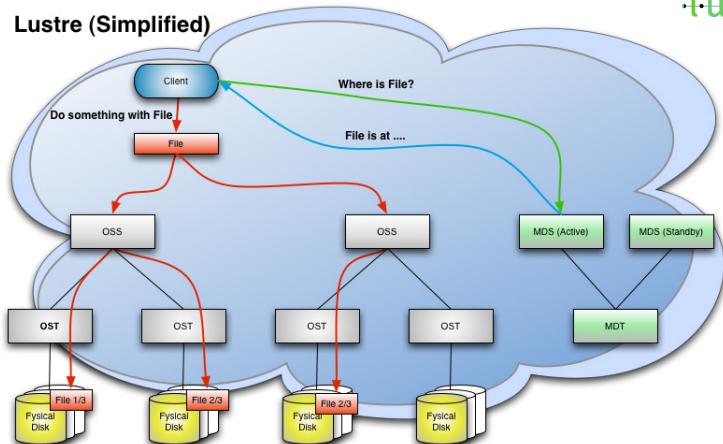


Abbildung : Lustre Übersicht [17], Logo [10]

# Beispiel: Ceph File System

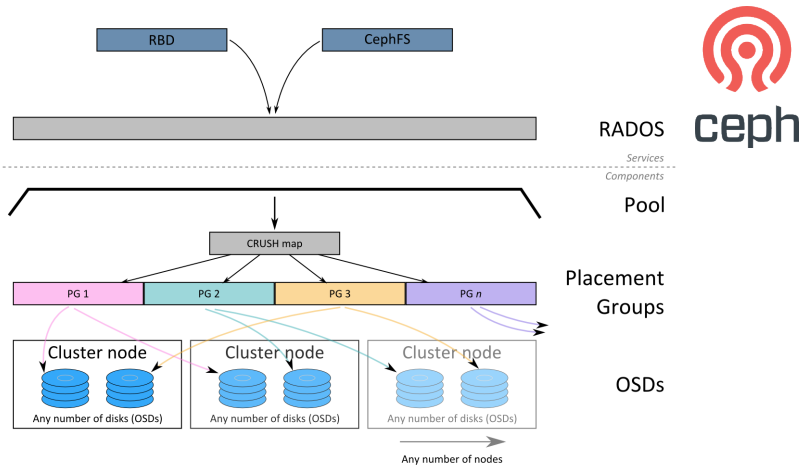


Abbildung : Ceph File System Übersicht [20], Logo [2]

# Perfekt für die Cloud

- On-demand self-service
  - Multi-Mandanten-fähig, quasi-unbegrenzter Speicher
- Broad network access
  - Breiten-Skalierung, Einfacher API-Zugriff (REST/SOAP)
- Resource pooling
  - Flacher Namespace, unbegrenzter Speicher, Verteilung, Breiten-Skalierung
- Rapid elasticity
  - Flacher Namespace, globale Adressierung
- Measured service
  - Multi-Mandanten-fähig, Metadaten-Speicherung

## Nachteile: Keine In-Place-Changes

- Nur Objekt als ganzes veränderbar, keine Teile davon
- Verändern eines Objekts erzeugt neue Objekt-Version
- Objekt-Speicher nicht geeignet für oft ändernde Daten
- Allerdings: “write once, read often“

# Nachteile: HTTP-Overhead

## ▼ General

**Remote Address:** 127.0.0.1:80  
**Request URL:** http://foo.bar/hello  
**Request Method:** GET  
**Status Code:** ● 200 OK

## ▼ Response Headers [view source](#)

**Accept-Ranges:** bytes  
**Connection:** Keep-Alive  
**Content-Length:** 27  
**Date:** Fri, 26 Jun 2015 19:56:13 GMT  
**ETag:** "1b-5197121a961f3"  
**Keep-Alive:** timeout=5, max=100  
**Last-Modified:** Fri, 26 Jun 2015 19:56:10 GMT  
**Server:** Apache/2.4.10 (Unix) OpenSSL/1.0.1j PHP/5.6.3 mod\_perl/2.0.8-dev Perl/v5.16.3

```
1 {  
2   data: 'hallo welt!'  
3 }
```

## ▼ Request Headers [view source](#)

**Accept:** \*/\*  
**Accept-Encoding:** gzip, deflate, sdch  
**Accept-Language:** de-DE,de;q=0.8,en-US;q=0.6,en;q=0.4  
**Connection:** keep-alive  
**Host:** foo.bar  
**Origin:** null  
**User-Agent:** Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/41.0.2272.76 Safari/537.36

**Abbildung :** HTTP-Request in den Google Chrome Dev Tools



# Gliederung

1 Einleitung

2 Über den Begriff

3 Warum Cloud?

4 Cloud-Architektur

**5 Schluss**

# Schlussfolgerungen: Was haben wir gelernt?

- Die Cloud ist bereits da  
Server-Architekturen verlagern sich sukzessive in die Cloud
- Der Begriff suggeriert fluffige Einfachheit  
Im Hintergrund stehen jedoch komplizierte Architekturen
- Die Cloud muss 5 Charakteristika abdecken  
on-demand self-service, broad network access, resource pooling, rapid elasticity, measured service
- Die Cloud braucht mächtige Dateisysteme  
Skalierbar, elastisch, Latenzzeiten ausgleichen, Ressourcen sinnvoll virtualisieren, ausfallsicher
- Objekt-basierter Speicher eignet sich perfekt für Cloud-Speicher  
Daten partiell speicherbar, flach organisiert, eindeutig identifizierbar

# Quellenangaben I

- [1] URL <http://image.slidesharecdn.com/overviewamazonwebservices-120503101649-phpapp02/95/overview-of-amazon-web-services-20-728.jpg%3Fcb%3D1345985130>.
- [2] URL [http://ceph.com/wp-content/uploads/2012/11/Ceph\\_Logo\\_Stacked\\_RGB\\_120411\\_fa.png](http://ceph.com/wp-content/uploads/2012/11/Ceph_Logo_Stacked_RGB_120411_fa.png). Ceph Logo.
- [3] URL <http://images.clipartpanda.com/christmas-tree-clipart-christmas-tree10.png>.
- [4] . URL <http://www.udldigital.de/wp-content/uploads/2013/03/cloud.gif>.
- [5] . URL <http://www.addsomehotsauce.com/wp-content/uploads/2013/10/clouds.jpg>.
- [6] URL <http://de.slideshare.net/StalwartAcademy/emc-it-transformation-stalwart-executive-briefing-2012>.
- [7] URL <http://www.softwarecandy.com/shop/images/brochures/fixfileassociation.com/ss-02.png>.
- [8] URL <http://cliparts.co/cliparts/rcL/ndB/rcLndB5Xi.png>.
- [9] URL [http://orig12.deviantart.net/fc2e/f/2012/220/8/1/alttpp\\_link\\_holding\\_triforce\\_sprite\\_by\\_eri\\_tchi-d4lcrgy.png](http://orig12.deviantart.net/fc2e/f/2012/220/8/1/alttpp_link_holding_triforce_sprite_by_eri_tchi-d4lcrgy.png).

## Quellenangaben II

- [10] URL [https://en.wikipedia.org/wiki/File:Lustre\\_file\\_system\\_logo.gif](https://en.wikipedia.org/wiki/File:Lustre_file_system_logo.gif).  
Lustre Logo.
- [11] URL [http://ceph.com/docs/master/\\_images/ditaa-ae8b394e1d31afd181408bab946ca4a216ca44b7.png](http://ceph.com/docs/master/_images/ditaa-ae8b394e1d31afd181408bab946ca4a216ca44b7.png).
- [12] Amazon.com traffic and demographic statistics by quantcast. URL <https://www.quantcast.com/amazon.com>.
- [13] URL [http://www.klimakasko.de/klimakasko/html/images/background\\_sturm.jpg](http://www.klimakasko.de/klimakasko/html/images/background_sturm.jpg).
- [14] URL <http://stickerish.com/wp-content/uploads/2011/09/TriForceYellowSS.png>.
- [15] URL [https://de.wikipedia.org/wiki/Cloud\\_Computing](https://de.wikipedia.org/wiki/Cloud_Computing).
- [16] URL <https://wrathofnino.files.wordpress.com/2009/07/zelda.jpg>.
- [17] July 2010. URL <http://louwrentius.com/static/images/lustre-schema.jpg>.
- [18] Amazon.com. URL <http://docs.aws.amazon.com/AmazonS3/latest/API/RESTObjectGET.html>.  
Screenshot.

## Quellenangaben III

- [19] C. Bandulet. Object-based storage devices, July 2007. URL <http://www.oracle.com/technetwork/server-storage/solaris/osd-142183.html>.
- [20] B. Desmond, September 2012. URL <http://www.anchor.com.au/blog/2012/09/a-crash-course-in-ceph/>.
- [21] P. Mell and T. Grance. The nist definition of cloud computing. URL <http://www.seu.ac.lk/careerguidanceunit/freedownload/0000%20The%20NIST%20Definition%20of%20Cloud%20Computing.pdf>.
- [22] W. Research. Many believe "the cloud" requires a rain coat, August 2012. URL <http://s3.amazonaws.com/legacy.icmp/additional/citrix-cloud-survey-guide.pdf>.
- [23] E. Schmidt. Search engine strategies conference, August 2006. URL <http://www.google.com/press/podium/ses2006.html>.

Danke für Eure Aufmerksamkeit!  
Fragen, Anregungen, Sorgen?