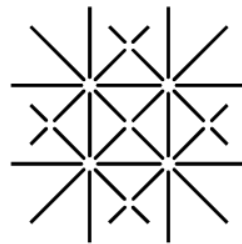


# Build your own Cloud dCache

Dietrich Zerr & Alexander Stiemer



UNI  
BASEL

# Inhalt

- Vorstellung dCache
- Umsetzung
- Probleme & Schwierigkeiten
- Zeitaufteilung
- Demo

# Vorstellung dCache

# Vorstellung

- Was ist dCache?

Das Ziel von dCache ist es, ein System zum Speichern und Abrufen großer Datenmengen, welche auf eine Vielzahl von heterogenen Server-Knoten verteilt sind, unter einem virtuellen Verzeichnisbaums mit einer Vielzahl von standardisierten Zugriffsverfahren bereitzustellen.

# Vorstellung

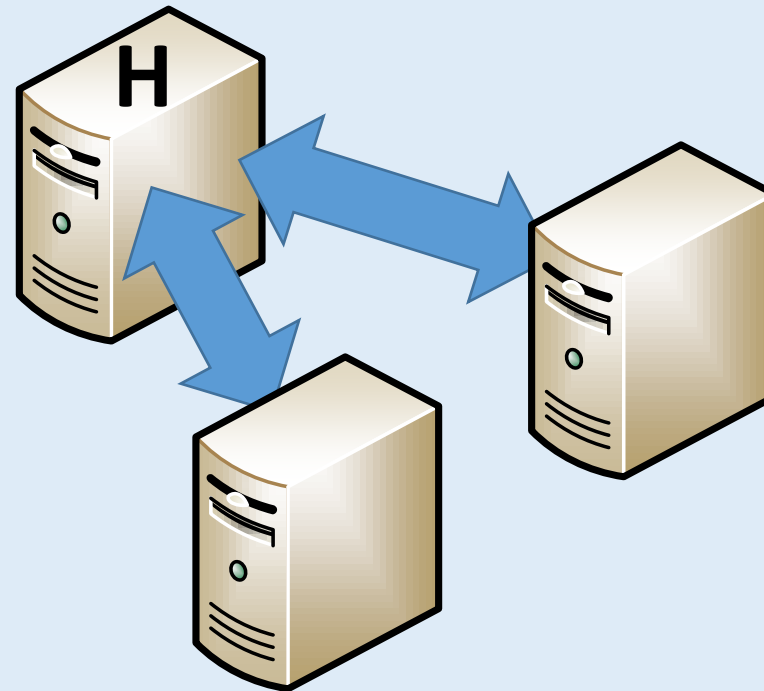
- Welche Vorteile bietet dCache?
  - OpenSource
  - Modular / Plugins
  - Skaliert, von einem Raspberry Pi bis zum Serverrack
- Was benötige ich um dCache zum laufen zu bringen?
  - GNU/Linux oder Solaris (--> Ubuntu 12.04 LTS)
  - Java 7 (--> Oracle JDK)
  - PostgreSQL > 9.2
  - Speicherplatz ;-)

# Vorstellung

- Und was ist nun das Ziel des Projektes?

- Installation eines Head Nodes
  - Anmeldung, Organisation, etc.

- Verbinden weiterer Nodes
  - weitere Speicher-Pools

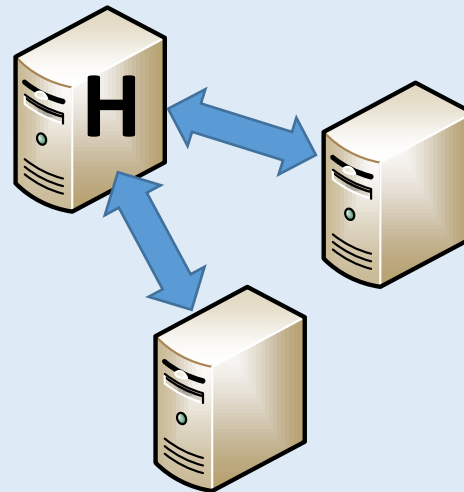


# Umsetzung

# Umsetzung

Wir bekamen 3 virtuelle Maschinen von der Uni, die im Uni-Netz liegen und nur über VPN erreichbar sind.

Eine davon sollte der Head Node sein, die anderen die Pool Nodes.





# Umsetzung

## Installation

- Java 7 installieren

```
apt-get install python-software-properties && add-apt-repository ppa:webupd8team/java  
apt-get update && apt-get install oracle-java7-installer
```

- Für Head Node: PostgreSQL 9.2 installieren

```
nano /etc/apt/sources.list.d/pgdg.list  
Einfügen>> "deb http://apt.postgresql.org/pub/repos/apt/ precise-pgdg main"  
  
wget --quiet -O - https://www.postgresql.org/media/keys/ACCC4CF8.asc | sudo apt-key add -  
apt-get update && apt-get install postgresql-9.2
```

# Umsetzung

## Installation & Konfiguration

- dCache installieren und Konfigurieren

```
wget http://www.dcache.org/downloads/1.9/repo/2.6/dcache_2.6.11-1_all.deb  
dpkg -i dcache_2.6.11-1_all.deb  
  
mv /etc/dcache/gplazma.conf /etc/dcache/gplazma.conf.bak  
touch /etc/dcache/gplazma.conf
```

→ Mehr zur Konfiguration von dCache in der Demo

# Umsetzung

## Konfiguration

- Für Head Node: PostgreSQL konfigurieren

```
nano /etc/postgresql/9.2/main/pg_hba.conf
```

```
...  
local    all             postgres               trust  
local    all             all                             trust  
host     all             all             127.0.0.1/32         trust  
host     all             all             ::1/128              trust  
...
```

# Umsetzung

## Konfiguration

- Für Head Node: PostgreSQL konfigurieren

```
createdb -U postgres chimera  
createuser -U postgres --no-superuser --no-creatorole --createdb --pwprompt chimera  
(kein Passwort nötig)  
createuser -U postgres --no-superuser --no-creatorole --createdb --pwprompt srmdcache  
(kein Passwort nötig)  
createdb -U srmdcache dcache  
createdb -O srmdcache -U postgres billing  
dcache database update
```

# Probleme & Schwierigkeiten

# Probleme & Schwierigkeiten

Mit den virtuellen Maschinen der Uni:

- Anfangs: Wie arbeitet man eigentlich damit?  
-> Tool „PuTTY“, alles ohne GUI
- Nichts vorinstalliert; die von uns benötigten Repositories nicht drin
- Liefen aufgrund von Ausfällen im November öfters nicht

# Probleme & Schwierigkeiten

Mit dCache:

- Sehr große Doku (237 Seiten PDF), funktioniert trotzdem nicht immer so, wie es drin steht (z.B. Ports)
- Uns unbekannte Technologien (webDAV, gPlazma2, ...) -> Konfiguration von Dingen, die wir nicht kennen / nicht verstehen
- Kein eigener Client, nur Schnittstellen
- Mal läuft's, mal nicht, mal was dazwischen
- Keine große Community -> kaum was Hilfreiches online

# Zeitaufteilung

- 20% dCache aufsetzen (erster Kontakt)
  - 20% Doku lesen
  - 40% dCache mit mehreren Nodes zum Laufen bekommen\*
  - 15% Client suchen (bzw. Filemanagement allgemein)
  - 5% Präsentation vorbereiten
- \* installieren, konfigurieren, rekonfigurieren, sich ärgern, neuinstallieren





# Demo