Deploying and extending on-premise cloud storage based on ownCloud

Dr. T. Hildmann and Prof. O. Kao | tubIT – IT Service Center | DCPerf Madrid 30 June 2014
Expectations – What is this talk about?

- Common cluster-architecture
- Standard LAMP-services
- Usual amounts of users (max. 40,000) and data (max. 300TB)

But as we started our service 2012...

- No reference architecture for ownCloud in larger environments
- No experience data (typical CPU / RAM / storage usage...)
- No idea if such a cloud storage service is comparable e.g. with CMS, ...

Let’s fill this gap!
Agenda

- **Motivation**: Why should we use / provide on-promise cloud storage?
- **Starting point**: Where did we come from?
- **The way**: The design that just worked for a year
- **Our experiences**: Some data (more see paper!)
- **Outlook**: Where do we go to?
- **Conclusion**
- **Invitation**
## Motivation

<table>
<thead>
<tr>
<th>2012 we already had...</th>
<th>but...</th>
</tr>
</thead>
</table>
| Andrew File System (AFS) | - Quota not enough  
- Great for Servers  
- good for Desktops  
- tricky for Laptops  
- bad for Tablets and Smartphones |
| - 1 GB / Student  
- 10GB / Staff  
- 250GB / Unit | |
| Subversion (SVN) | Too complicated for non-techies! |
| Sharepoint | Too expensive (licenses + admin), no good mobile integration. |
| Dropbox-Accounts (4.654 space race participants) | - data protection issues  
- what about NDA material etc.?  
- vendor lock-in  
- missing integration |
Starting point

So we evaluated Dropbox-like on-promised sync and share solutions.

The service should meet the following requirements

– Enough Quota: 10GB/student, 50GB/staff, 250GB/unit
– Scalable for about 50,000 users
– Sync clients for Linux, Windows, MacOS X
– Mobile clients for Android and iOS
– Integration into our directory service

After testing several products we decided to use ownCloud.
The ownCloud application

ownCloud web gui (branded)

sync client (OS X)
First extrapolation from beta-test with 150 users

DB: „We need 500 Servers for 50,000 users!“

SAN: „May be not enough IOPs!“

Servers: „We don‘t know if this crashes the virtualization infrastructure.“
Architecture

- Cisco ACE Load Balancer
- ownCloud Application Server
- MySQL Cluster with 'Galera'
- GPFS (Disks from SAN)

Deploying on-premise cloud storage using ownCloud | T. Hildmann | DCPerf'14
Component diagramm
Number of Users

Today about 6,000 users

8.5.2014 ~1,200

29.5.2014 ~2,200
How to increase users numbers?

- Integration in **study classes** (e.g. share documents, software or folder for homework, …)
- **Auto-mount** ownCloud directories (WebDAV) in **PC-pools** (no problem with Linux but…)
- Linux image with **pre-installed environment** (Knoppix-like)
- **Flyers** with installation instructions during matriculation
- Just **wait!** For many users migration (from other s&s services) takes some time.
- Supporting the **federated service** for other universities (usage grows with cooperation).
Filecache

ownCloud bug fixed here

Today about 8 mio. files
Some new data (not in paper)

- < 1GB: 69%
- < 10GB: 25%
- ≥ 10GB: 6%

- Daily: 60%
- Monthly: 7%
- Yearly: 33%
## Human Resources

<table>
<thead>
<tr>
<th>Project Phase</th>
<th>Admin / Techn.</th>
<th>Helpdesk</th>
<th>Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation</td>
<td>🧑‍💼 1 ½ 🧑‍💼</td>
<td>-</td>
<td>🧑 ½</td>
</tr>
<tr>
<td>Implementation</td>
<td>🧑 ½ 🧑</td>
<td>-</td>
<td>🧑 ¼</td>
</tr>
<tr>
<td>Beta-Testing</td>
<td>🧑 🧑</td>
<td>🧑 ½</td>
<td>🧑 ¼</td>
</tr>
<tr>
<td>Roll-Out</td>
<td>🧑 🧑 🧑</td>
<td>🧑 🧑</td>
<td>🧑</td>
</tr>
<tr>
<td>1st Year</td>
<td>¼ 🧑</td>
<td>½ 🧑</td>
<td>¼ 🧑</td>
</tr>
<tr>
<td>Upgrade*</td>
<td>🧑 🧑 🧑</td>
<td>🧑</td>
<td>🧑 ½</td>
</tr>
</tbody>
</table>

*Upgrade = OC5 → OC6 EE, federated service, hardware upgrade, 250GB for units incl. self-service portal.*
ownCloud @ TU Berlin today / next steps

Done:

- IBM GSS cluster 300 TB
- MySQL Galera cluster with 3 DB Servers
- TUB offers ownCloud as a federated DFN service

Work in progress:

- ownCloud 6 Enterprise Edition
  - branded and preconfigured clients
  - professional support
  - shibboleth integration
- Extended Groups
Conclusion

1. The **need** of an on-premise cloud storage is **recognized**.
   - Access to data from many different mobile smart devices
   - Address compliance and data protection laws
2. User **expectations**:
   - High integration
   - High security (integration into IDM, backup, incident-handling)
3. **ownCloud** is a **standard LAMP** application
   - We are concentrating on sync and share service.
   - Can be driven up to 5,000 active users with 2 web-frontends, 3 db-servers and 2 filer.
   - No “Big Bang” – usage is growing slightly.
4. **Awesome positive feedback** from users!
   - For some users ownCloud became the most important service beside e-mail.
Invitation

- Run your own on-premise Sync & Share service or use federated services from other universities
- Start with a minimal setup and let it grow
- Cloud storage is no rocket science
- Share our paper with the key persons at your institution

If there are any questions (now or at home), do not hesitate to ask me:

thomas.hildmann@tu-berlin.de