Designing a Large Scale Cooperative Cloud Storage Platform for the Academic Community in Northrhine-Westfalia (NRW)

EUNIS Riga, June 13, 2013
Agenda

1. Why Sync&Share NRW?
   - Motivation from user survey
2. Market and products for Sync&Share
3. Designing a large scale Sync&Share service
   - Service characteristics
   - Sizing
   - System architecture and Service operation
4. Timeline for Sync&Share NRW
Why Sync&Share NRW?

• 2006: Amazon EC2/S3
• 2007: Dropbox
• 2009: NIST definitions for Cloud Computing
• Academic community uses free and/or paid public cloud services (e.g., Dropbox for scientific project communication)
• Data privacy/confidentiality concerns – esp. for non-EU services
• Lack of on-premise alternatives compliant with EU/german data protection laws (Datenschutzgesetz, Auftragsdatenverarbeitung)
• Demand imminent by late 2011: discussion at universities
Sync&Share NRW Timeline

- 25.04.2012 Uni Münster students petition for S&S service
- 07.05.2012: NRW Uni IT director workshop on Sync&Share
- May 2012: Uni Münster user survey on cloud storage demand
- 27.6.2012: System Architecture design workshop
- Mid 2012: Sync&Share Software Evaluation: favorit -> OwnCloud
- Oct 2012: first proposal by system vendor with OwnCloud
- Nov 2012: discuss funding possibilities with Ministry and DFG
- 15.01.2013: final workshop for S&S NRW consortium formation
- 22.05.2013: workshop on legal aspects (DFN Berlin)
- 28.05.2013: funding proposal filed at Science Ministry by Univ. of Münster as lead in Sync&Share NRW consortium – DFG peer review currently underway
- Dec 2013: still hoping to get Sync&Share NRW started!
Previous usage of cloud services by WWU users (students and employees)

- **47.8%** used
- **35.6%** known, but unused
- **16.6%** unknown

$n=2.493$
Storage Cloud Service: Usage amongst WWU students and employees

- Yes, I am already using cloud storage services: 52.0% students, 36.6% WWU employees
- I know cloud storage service, but I am not using them: 32.6% students, 44.0% WWU employees
- Cloud storage services were unknown to me: 15.5% students, 19.4% WWU employees
Reasons for not using cloud storage: Students vs. employees

- **no additional value**
  - Students: 49.5%
  - Employees: 50.6%

- **too complicated**
  - Students: 13.2%
  - Employees: 20.8%

- **not enough storage space**
  - Students: 3.7%
  - Employees: 6.2%

- **deficient data security**
  - Students: 63.0%
  - Employees: 56.8%

- **too expensive**
  - Students: 12.0%
  - Employees: 7.8%

- **other**
  - Students: 11.1%
  - Employees: 17.3%
Willingness to use on-premise storage cloud: Students vs. employees
Intended usage for on premise cloud storage:
Students vs. employees

- Data backup:
  - Students: 49.6%
  - WWU Employees: 52.8%

- Data exchange within the WWU / the UKM:
  - Students: 55.2%
  - WWU Employees: 68.2%

- Data exchange universities NRW:
  - Students: 15.7%
  - WWU Employees: 16.2%

- Data exchange universities Germany:
  - Students: 17.2%
  - WWU Employees: 17.4%

- Data exchange external users:
  - Students: 20.8%
  - WWU Employees: 21.4%
Sync&Share NRW Timeline

• 25.04.2012 Uni Münster students petition for S&S service
• 07.05.2012: NRW Uni IT director workshop on Sync&Share
• May 2012: Uni Münster user survey on cloud storage demand
• 27.6.2012: System Architecture design workshop
• Mid 2012: Sync&Share Software Evaluation: favorit -> OwnCloud
• Oct 2012: first proposal by system vendor with OwnCloud
• Nov 2012: discuss funding possibilities with Ministry and DFG
• 15.01.2013: final workshop for S&S NRW consortium formation
• 22.05.2013: workshop on legal aspects (DFN Berlin)
• 28.05.2013: funding proposal filed at Science Ministry by Univ. of Münster as lead in Sync&Share NRW consortium – DFG peer review currently underway
• Dec 2013: still hoping to get Sync&Share NRW started!
Market and Products for Sync&Share

• Evaluated software solutions for Sync&Share in Summer of 2012:
  • iFolder (students recommendation; Open Source)
  • Powerfolder (tests at GWDG Göttingen; commercial)
  • TeamDrive (commercial)
  • OwnCloud (then new; Open Source)
• Sync&Share Solutions: Software + Plattform
  • IBM GPFS GSS appliance with OwnCloud
  • EMC ISILON with Syncplicity (or OwnCloud)
  • Object stores: **DELL (Grau Data)**, DDN (Mezeo), HDS (HCP Anywhere)
Sync&Share NRW - service characteristics

- User self enrolment (keep management effort down)
- Aim for 30 GB per user; 400GB for scientific work groups (better than Dropbox)
  - -> 6 PB highly available storage space envisioned to be needed
- Sync clients for common platforms (Win, Mac, Linux, Android, iOS)
- Web interface: eg. to share data with other users
- Limited external (non academic community) users possible
- No additional backup of stored data – Sync&Share!
- Approx. 60% of NRW Uni FTE/students committed; possibly more
- Service provided by 3 university computer centers (MS, BN, DU-E) for the consortium - best effort service
Schematic view of system architecture

Service distributed to three university data centers to keep resource demand in acceptable limits.
Model prediction:
Data volume in storage cloud

Model prediction:
Network bandwidth demand (all 3 sites)
Sync&Share NRW Timeline

- 25.04.2012 Uni Münster students petition for S&S service
- 07.05.2012: NRW Uni IT director workshop on Sync&Share
- May 2012: Uni Münster user survey on cloud storage demand
- 27.6.2012: System Architecture design workshop
- Mid 2012: Sync&Share Software Evaluation: favorit -> OwnCloud
- Oct 2012: first proposal by system vendor with OwnCloud
- Nov 2012: discuss funding possibilities with Ministry and DFG
- 15.01.2013: final workshop for S&S NRW consortium formation
- 22.05.2013: workshop on legal aspects (DFN Berlin)
- 28.05.2013: funding proposal filed at Science Ministry by Univ. of Münster as lead in Sync&Share NRW consortium – DFG peer review currently underway
- **Dec 2013: still hoping to get Sync&Share NRW started!**
Thank you very much for your attention!