

# HU Box

## Cloud Storage Deployment and Operation at Humboldt-Universität zu Berlin

Karsten Asshauer, Enno Gröper



Computer- und Medienservice, Humboldt-Universität zu Berlin

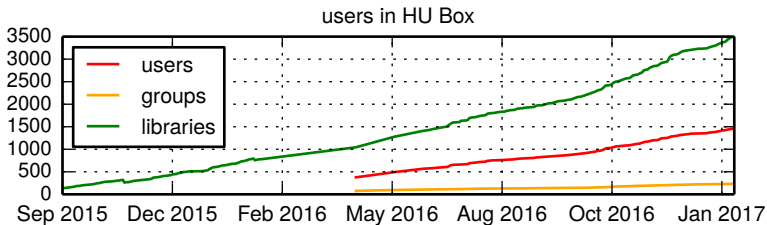
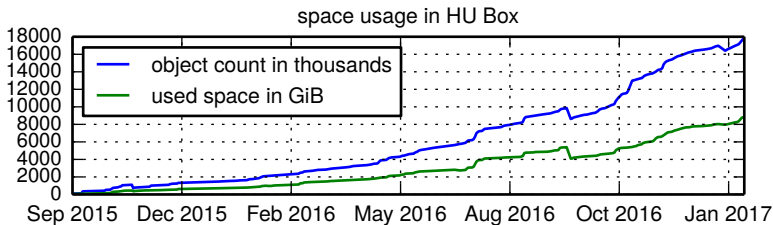
31 January 2017

# Overview



- Statistics
- Our Deployment
- Join efforts!

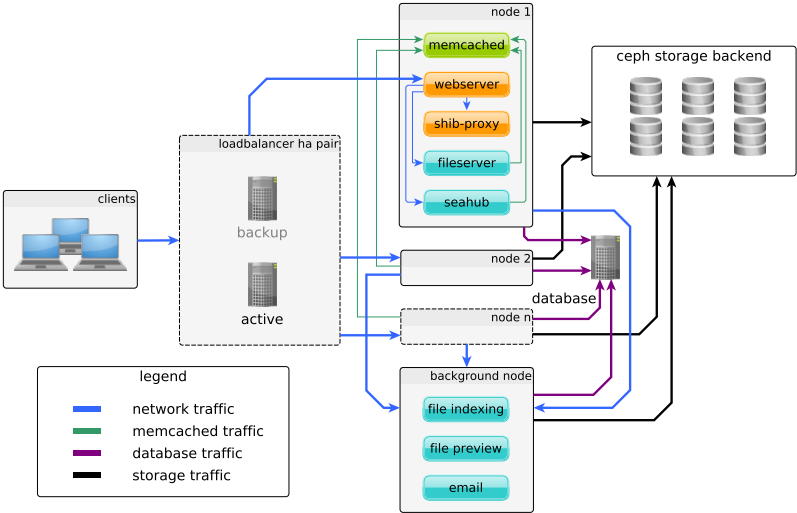
# Statistics!



- ▶ Seafile Professional (currently 6.x)
- ▶ Running on OpenStack with CEPH as storage backend
- ▶ Using single central database server
- ▶ Shibboleth for Authentication
  - ▶ Local users
  - ▶ Guests (DFN-AAI, eduGAIN) (in progress)
- ▶ Generic local guest accounts using invitation feature of Seafile (planned)

- ▶ All users have to accept current terms and conditions (currently self-made solution)
- ▶ Shibboleth affiliation can be mapped to Seafire roles
- ▶ Apply guest role to all external users by default
- ▶ Including Shibboleth Embedded Discovery Service to select identity provider
- ▶ Clients use out-of-band authentication via tokens
  - ▶ By using authentication tokens you bypass central authentication servers
  - ▶ Token lifetime has to be limited

# Cluster setup



# Concerns?



- ▶ Backup and restore
  - ▶ Replication of CEPH objects into filesystem backend representation of Seafile (objects as files) in our SAN
  - ▶ Backup file representation to TSM (tape)
  - ▶ If ever needed: Restore from tape will take a *long time*
- ▶ Memcached high availability (HA)
  - ▶ Memcached cluster mainly to improve memory usage / share load, not for HA
  - ▶ Memcached cluster logic located in the different client implementations
  - ▶ Consistency issues could arise → active/passive using pacemaker?
- ▶ Deletion of user data: a users personal data vs. collaborative data (shared libraries)

# Join efforts!



The screenshot shows the GitHub interface for the repository 'hu-berlin-cms / ansible-seafile'. The URL in the browser is 'https://github.com/hu-berlin-cms/ansible-seafile'. The repository has 6 Unwatch, 2 Stars, and 0 Forks. The main content area shows the repository name and navigation tabs for Code, Issues (12), Pull requests (0), Projects (0), Wiki, Pulse, Graphs, and Settings. Below this, there is a section titled 'ConfigStack for deploying Seafile Professional in a cluster' with an 'Edit' button. A summary bar displays '258 commits', '3 branches', '0 releases', and '2 contributors'. Below the summary bar, there are buttons for 'Branch: master', 'New pull request', 'Create new file', 'Upload files', 'Find file', and 'Clone or download'. A commit message by 'egroeper' is visible: 'Merge branch 'feature/helper\_scripts''. The latest commit hash is '199ff1c' and it was made 7 days ago.

## Goals:

- ▶ Join efforts
- ▶ Be flexible and extensible
- ▶ Get simple way to deploy production cluster (on Ubuntu)



# Further collaboration



- ▶ seafile-python-api-scripts
  - ▶ Trying to establish collection of helpful scripts using python API of Seafile
- ▶ seafile-ceph-tools
  - ▶ Tools for ceph backend, like replication  
CEPH → filesystem or CEPH usage accounting
- ▶ Seafile Educational User Community (since 2015)
  - ▶ Meetings of academic users and Seafile developers (Seafile Ltd.)
  - ▶ Seafile Educational Community Forum (HU Moodle)
  - ▶ Next Seafile Workshop: September 2017, Université de Strasbourg, France



Thank you for your attention!

Karsten Asshauer - [karsten.asshauer@cms.hu-berlin.de](mailto:karsten.asshauer@cms.hu-berlin.de)  
Enno Gröper - [enno.groeper@cms.hu-berlin.de](mailto:enno.groeper@cms.hu-berlin.de)

Feel free to contact us for further questions!