CS3 community summary – animated CS3map for continuous site report surveys

ETH Zürich

polybox.ethz.ch

Request for site survey
26th February 2024

27 answers received from:

2023: 12 answers received

- CNRS
- CERN
- University of Oslo
- University of Oslo
- AGH University of Kracow
- Nextcloud
- CESNET a.l.e.
- INPE
- King's College London - e-Research
- Humboldt University Berlin
- MAX IV Laboratory
- Austrian Academy of Sciences, Vienna Biocenter
- Helmholtz-Zentrum Berlin GmbH
- Collabora

- SWITCH
- Foschungszentrum Jülich GmbH
- INPE
- Sunet
- thegood.cloud (formerly www.360ict.nl)
- CESNET
- Max-Planck Digital Library
- ETH Zurich
- University of Münster
- Max Planck Computing and Data Facility (MPCDF)
- Ludwig Maximilians Universität, Fakultät für Physik
- Technical University Berlin

Thank you for your contribution!

Q: Role of EFSS (Enterprise File Sync&Share) services in your organization?

- **Administration workflows and processes** (document handling, process support,...); General project collaboration (sharing of project data and documents in technical and engineering context)

- **Sharing and collaborating outside of organization** and across other organizations

- **General project collaboration** (sharing of project data and documents in technical and engineering context); Data workflows (collection, access and sharing of datasets or files considered as "data" independently of their type)

- **Research collaboration** (sharing code, data and other files; integrated access to software, computing, analytics and similar resources)

- **Education and outreach use-cases**

Q: Describe some important workflows that have been enabled by EFSS services by your organisation?

- Dropbox like service on prem'

- concurrent editing/collaboration, secure sharing of internal documents, data analysis

- standard sync & share, collaboration with office files

- Publishing research data to EUDAT B2SHARE

- Collaboration on research data

- sync&share, collaboration, data certification, archiving

- Collaboration, Online Exams, Course Datastore, Research Data distribution

- Google drive replacement

- very diverse

Q: Must your use cases be hosted on premise?

- **9 * yes**

- **1* Depends on company data policy**

Q: Describe the developments of your service over the past 5 y?

- Building of Reva, used as the building block of ownCloud OCIS; usability improvements; integration with SAMBA, integration with CEP

- provide stable and reliable Sync & Share service for the community

- The service moved from a singe server, hosting all components, to a redundant multiple server setup

- We started in 2021 and have since onboarded about 20 paying customers and we also have 34 free limited instances

- Diversing from Nextcloud hosting to M365/Gworkspace alternative.

- growth

- Well established and liked service over the whole campus. Strongly used over COVID period.

- project lead

Q: Describe the developments of your service over the past 5 y?

- At the begining, some custom apps for SSO and guests users, among others. It is available now on community edition.

- Integration with CEPH, multiple access protocols including FUSE and SAMBA

- project-folders (technical accounts, data belong to the institution, not to a person)

- Setpup authentication, mfazones, jupyter notebooks, Sciebo RDS

- yes, per organisation the requirements are very diverse

- Integration with our Perun AAI system (for group management), customized openidconnect app.

- data certification (bloxberg), archiving

- We wrote additional apps

Q: Are you planning on future developments and if so, what are they and why do you need them?

- Integration of backup self-restore to reduce user support; better support for CEPH, allowing the usage of apps; better integration with SWAN jupyter notebooks; better collaborative tools via the clients

- sharing - depending on development of ownCloud

- Yes, we are continually developing the service in contact with users

- Enterprise email, rich text chat client.

- Infrastructure modernisation - migration to new storage backend (Ceph), deployment model (k8s).

- Migration to other EFSS software; create research data workflow

- Looking for ways to handle larger (TB+) datasets, ideally with POSIX access for efficient data transfer. Potentially ownCloud Infinite Scale with CephFS or EOS.

Q: What is your vision on the future of EFSS services - where do you see these kind of services going to coming 5 years or so?

- Replacement of ownCloud based service by a NextCloud based one

- More integration with other web services, centralizing users' data; blured lines between local and web apps

- provide a secure and proven alternative to commercial cloud services

- We want to be an academic toolbox where the complete lifecycle of research data has a home

- EFSS is already a commodity, look at the focus of CS3. I expect the added value (in Europe the privacy aspect) will be more important.

- Seeing how much things are outsourced today, probably many will be migrated/moved to OneDrive/Google Suite etc.

- I see them as an infrastructure for research data. You have to have an EFSS for a data privacy friendly research environment.

Q: What do you think what the added value is compared to "hyper scaler" offerings?

- **Data sovereignty, more freedom to do integrations and adapt to use particular cases, full ownership allows better support to users**

- **service on premise, more trust in SWITCH than in hyper scaler**

- **GDPR compliance and security**

- **Not so much in EFSS in my opinion. Eventually services need to be paid and I assume the cost benefit in most cases isn't that big.**

- **privacy, data ownership & control, customisability, not depending on closed proprietary SW**

- **You keep your data in-house. Especially on a university, research data is the direct output of the core business.**

- **Data privacy, more research oriented, more flexible**

Q: Would your organisation in principle be interested in joining a federative infrastructure of EFSS services to enable easy sharing across instances?

- 8 * YES

- 1 * MAY BE (but it raise a lot of security questions)

- 1 * in principle, yes, depends on actual conditions & technical implementation of such infrastructure

Q: Is federated sharing enabled/supported?

- 11 * YES

- 5 * NO

Q: Would you be interested in joinly collaborating with other organisations?

- 9 * YES   ;   3 * NO

Q: Additional usage of public cloud services?

- **6 / 18 : yes**
- **12 /18 : no**

Q: Comments on public cloud usage ?

- **5 answers:**
  Google Workspaces
- Microsoft Office 365
- We also use box and google services
- OneDrive and Google Suite are used here as well. OneDrive is used as personal profile storage.
- Users most probably use these services, but are probably not allowed to ;)

Q: Number of users?

- **530000 reported within 15 answers**

Q: Comments on public cloud usage – amount of shares?

- **2.3 Mio shares reported within 12 answers**

Q: Technologies used (Owncloud/Nextcloud/Seafile/…..)?

- **8 / 15 : owncloud**
- **6 / 15 : nextcloud**
- **1 / 15 : seafile**

Q: used storage / objects ?

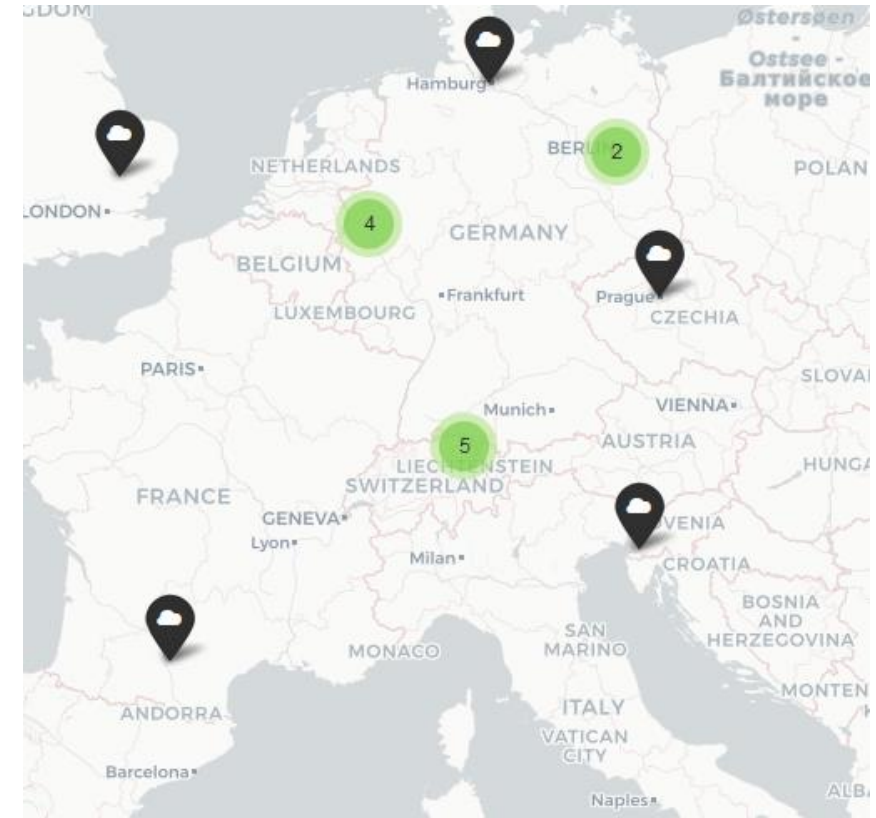- **24 PB / 3 billions**

Q: Integration with applications and other software: Integrated software?

- **7 * Collabora Online**
- **6 * OnlyOffice**
- **3 * Jupyter Notebooks**
- **2 * Microsoft Office Online / Sharepoint**

out of 18 answers

https://cs3map.ethz.ch

feedbacks are welcome

Summary

- This year 27 feedbacks (former times 12-16) – thanks a lot for your contribution!!!

- Installations since 2013 are reported

- Installations from more than 2500 users up to 230k users are running

- 530k user accounts have been reported within the CS3 community

- 24 PB used sync and share storage and 3 billions objects within the CS3 community

- 1 million+ public shares (anonymous links) are reproted form 11 feedbacks

- Special Use Cases have to be hosted on-premise

- # PLEASE PARTICIPATE: https://cs3map.ethz.ch

Site Reports until 17:30

- 16:10 - 16:25:
Combining NextCloud with Direct Access to dCache at DESY, Christian Voss

- 16:25 - 16:40:
The S3 Object Storage Service on INFN Cloud, Dr. Alessandro Costantini

- 16:40 - 16:55:
Publication of open research data with sync&share storage, ETH Zurich, Andreas la Roi, Gianluca Caratsch - Speaker

- 16:55 - 17:10:
Sunet Drive - Sweden - Community Site Report, Magnus Andersson et al.

- 17:00 - 17:15:
Federated Sync&Share, Enrico Signoretti

- 17:20 - 17:30:
Community Site Report – Summary, Dr. Tilo Steiger

ETH *zürich*

Now the site reports will follow:
Thank you for your attention.

Manual 1/3

CS3 Community Map

Please Visit: https://cs3map.ethz.ch/

For making an entry please generate a user:

https://cs3map.ethz.ch/register

Register:

Username:

Please enter a username

Password:

Please enter a password

Confirm password:

Please confirm your password

Register

Return

Manual 2/3

Please add your institute:

https://cs3map.ethz.ch/forms

Add Institute

Institute:

Name of the institute e.g. ETH Zürich

Number of terrabytes:

Number of terrabytes e.g. 5.5

Location X:

Enter the X coordinate e.g. 47.376332

Location Y:

Enter the Y coordinate e.g 8.547531

Number of users:

Number of users e.g. 5000

Link of the institute:

Enter the link of the institute e.g. https://ethz.ch

Submit   Return

Manual 3/3

Modify your entry:

https://cs3map.ethz.ch/profile

Profile of admin

Name	X	Y	Edit
ETH Zürich	47.376234	8.547117	

To change a created entry, please click on the green button. After that you will find a list of all entries that you have created.

To delete something, you have to press the red button. You will be asked again if you are sure.

Delete Entry forever

Institute:

ETH Zürich

Number of terrabytes:

20000

Location X:

47.376234

Location Y:

8.547117

Number of users:

10000

Link of the institute:

https://www.ethz.ch

Delete forever   Cancel

Questions? Please feel free to contact:

Dr. Tilo Steiger
steiger@ethz.ch
+41 44 676 2275