



**CS<sup>3</sup>  
MESH<sup>4</sup>  
EOSC**

**Connecting European Data**



## ScienceMesh-Seafile

Update on the integration work within CS3MESH4EOSC



CS3MESH4EOSC has received funding from the European Union's Horizon 2020 Research and Innovation programme under **Grant Agreement No. 863353**.

## \* Seafile sites

- \* PSNC - runs Seafile-based sync & share since 2015
- \* 3k+ users, some power users :
  - \* Biology-related research: **100TB**
  - \* Video codecs-related research: **50TB**
- \* Other sites - several dozens 1000's of users each:
  - \* HU Box - Humboldt University Berlin
  - \* Seafile @ Johannes Gutenberg University Mainz
  - \* Keeper - Max Planck Digital Library data archival service



## \* Integration is a must!

- \* We have chosen Seafile for its performance and reliability
- \* BUT isolation is not an option!



Credits to Jonathan @Seafile (from presentation during day 1 of CS3 2022)

### Work positioning within the project:

- \* Dedicated task for EFSS-CS3API integration within the project including Seafire - CS3API effort
- \* PSNC represents Seafire sites in CS3MESH4EOSC
- \* More sites to collaborate in 2022:
  - \* Currently in collaboration with Humboldt University Berlin
  - \* Discussions with MPDL
  - \* Collaboration with Uni Mainz planned

### Status:

- \* Main body of work to be taken up in 2022
- \* Assumption to have a working prototype in 4Q 2022
- \* Planning:
  - \* Discussing and defining the scope and “depth” of the integration
  - \* Organising the team:
    - \* PSNC / Humboldt University - coordination
    - \* Seafire Ltd <sup>1)</sup> / DataMate <sup>2)</sup> - industrial partners
    - \* Universities - we’re open for collaboration
  - \* Setting up the plan
    - \* Starting from testbeds for OCM
    - \* Examining the OCM support (using OCM testing suite - see Michiel’s @Pondersource presentation on Monday)
    - \* Defining the integration approach
    - \* Actual integration

1) <https://www.seafire.com/en/about/>

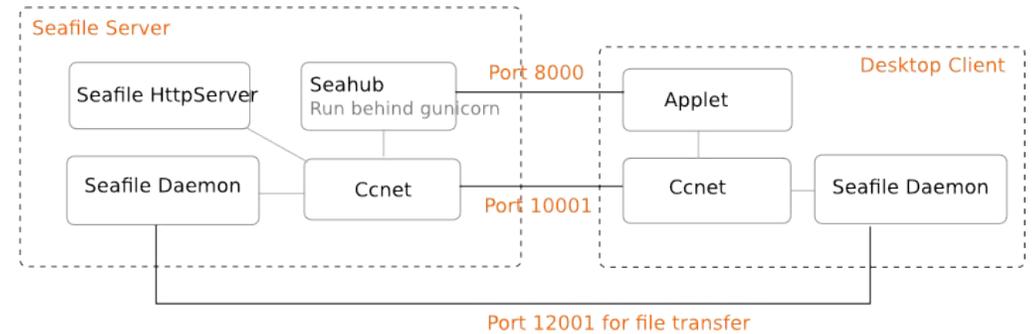
2) <https://www.datamate.org/seafire/>

## Seafile architecture (vs ownCloud / NextCloud):

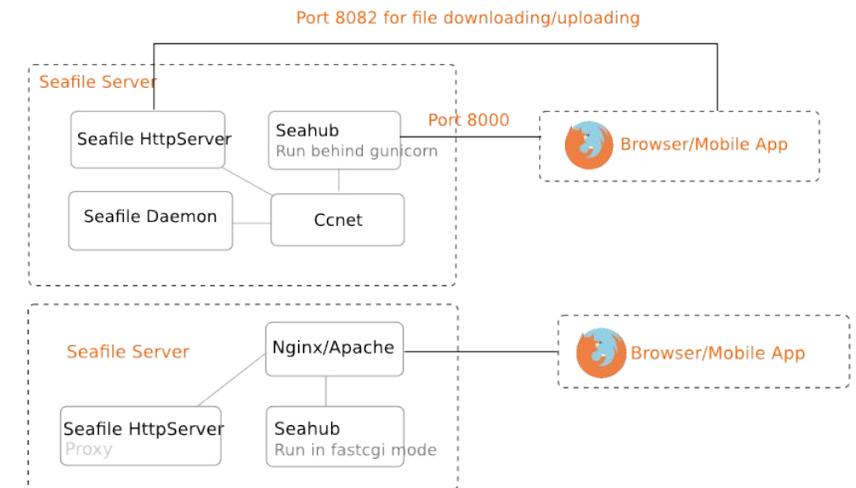
- \* Data organised in so-called libraries
- \* Internal data structures similar to git repositories
- \* **No 1-1 mapping of user-level files/objects to storage level files/objects!**

## Seafile implementation features:

- **Modularity** - all Seafile clients are built on top of **Seafile WebAPI** including:
  - Web interface (Seahub)
  - Sync desktop clients for: Windows, Linux, MacOS
  - Mobile applications: Android, iOS
  - Drive clients for: Windows, Linux, MacOS
- Well-documented **Seafile Web API**:
  - <https://download.seafile.com/published/web-api/v2.1>
- Web UI hooks possible - see examples on following slides and: <https://mpdl.zendesk.com/hc/en-us/articles/360011432700-Archiving>



Seafile desktop clients - data synchronization with Seafile Server



The picture below shows how Seafile mobile client interacts with Seafile server: two options - with nginx configure or simple setup

Source: <https://manual.seafile.com/develop/server-components/>

### \* **Deep integration:**

#### OCM/CS3APIs implementation mainly in Seafile server:

- \* Seafile already supports OCM - good starting point!
- \* OCM support allows connection to NextCloud
- \* CS3APIs integration would extend OCM implementation
- \* *Deep integration would involve Seafile Ltd*

### \* **Lightweight integration:**

#### Proxy / overlay on top of the Seafile API

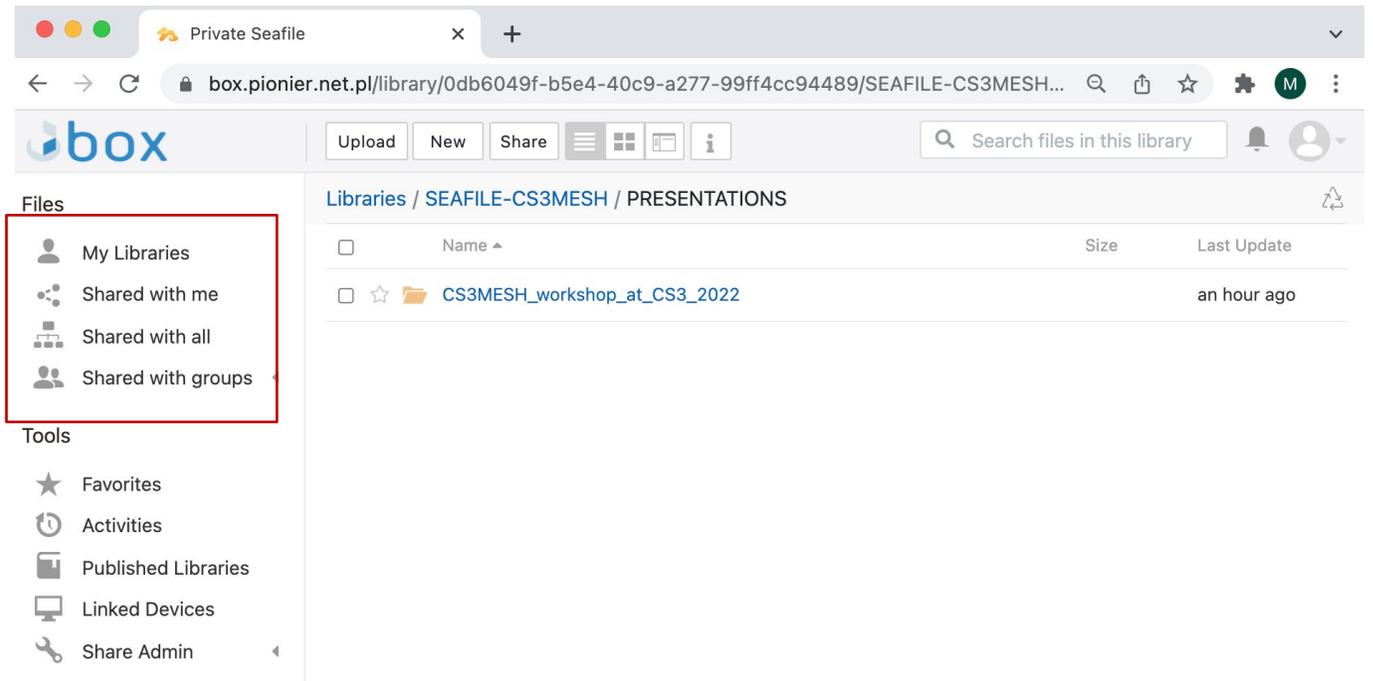
- \* CS3APIs-related extensions - to be implemented as an overlay / proxy on top of Seafile server
- \* *Limited intervention into the Seafile Server code!*

### \* **Possibly two-step approach:**

#### Depending on the analysis results:

- \* lightweight integration for fast prototyping
- \* then deep integration for sustainability

- \* Federation features to be 'presented'
  - \* Share with ... Federated users...
    - \* Currently:
      - \* *My Libraries.*
      - \* *Shared with me*
      - \* *Shared with all*
      - \* *Share with groups*
    - \* *To be added:*
      - \* *Shared with federated user*
      - \* *Shared with federated group*
  - \* Shares presentation:
    - \* *to be discussed and designed*
    - \* *to be consistent with other EFSS platforms GUI integrations*
- \* Support for invitation workflow:
  - \* *To be analysed*



\* **Basic stuff / existing stuff:**

\* **Office integration - “open with... (federated app)”**

\* **Open Data - data archival + repository deposit:**

\* *For instance based on Keeper<sup>1)</sup> @MPDL reliable data archival functionality<sup>2)</sup>*

\* extensions available in GitHub<sup>3)</sup>  
we’re in contact with MPDL team

\* **Other options: analysis & discussion:**

\* **Data science environments:**  
there were trials to bind JupyterLab and Seafile - see Human Brain Project<sup>4)</sup>

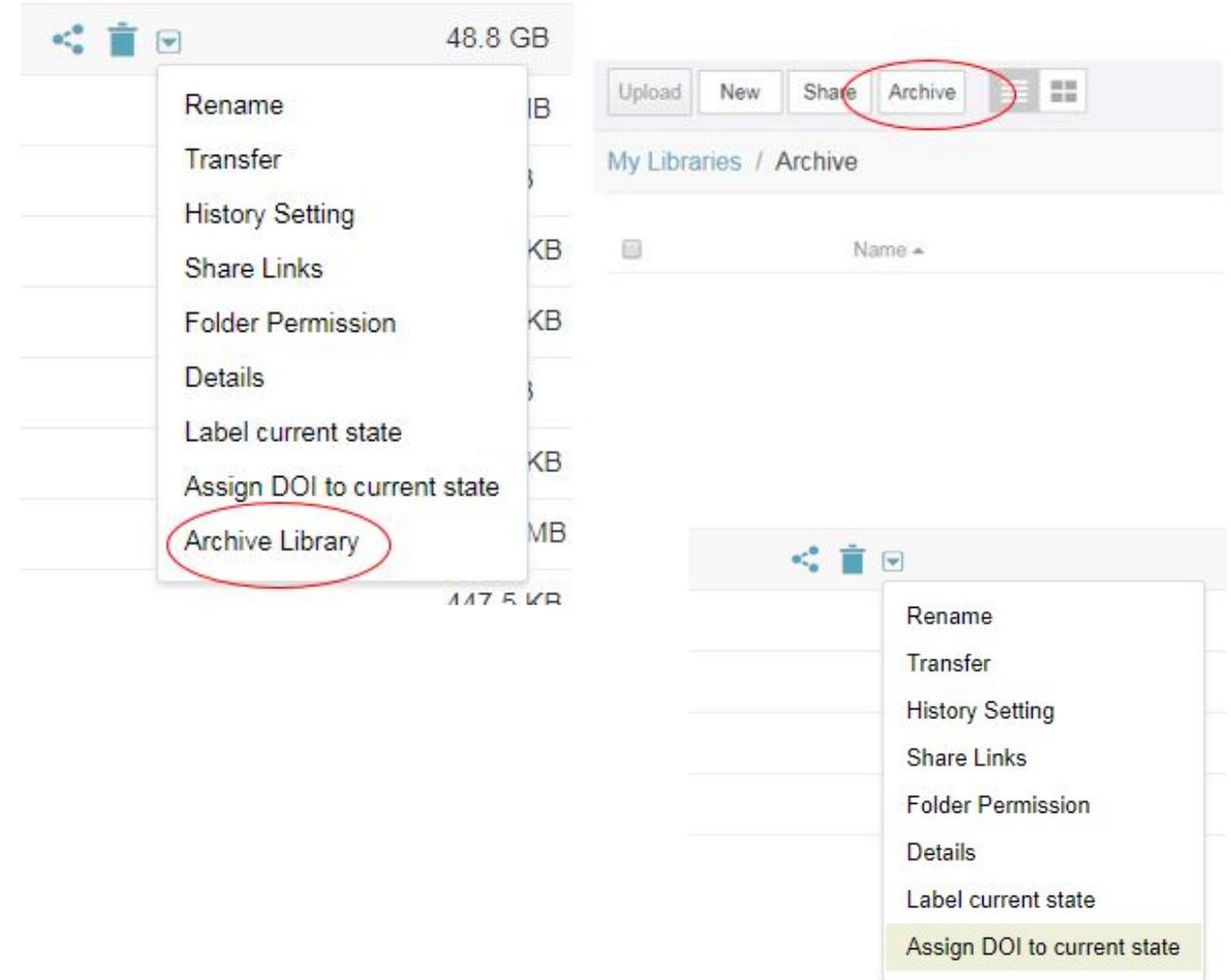
\* **Massive on-demand data transfer** - EFSS to be an user-facing interface to large data storage infrastructure and repositories - e.g. LOFAR

1) <https://mpdl.zendesk.com/hc/en-us/categories/360001234340-Keeper>

2) <https://mpdl.zendesk.com/hc/en-us/articles/360011432700-Archiving>

3) <https://github.com/MPDL/KEEPER/>

4) <https://github.com/HumanBrainProject/clb-jupyterhub-openshift>



Source: Keeper documentation:

<https://mpdl.zendesk.com/hc/en-us/categories/360001234340-Keeper>

### The work is ongoing...

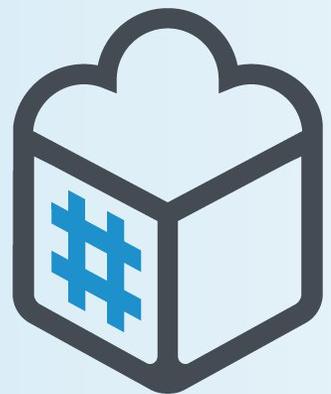
- \* **Defining the scope and depth** of integration is happening now
- \* **Looking for partners:**
  - \* PSNC / Humboldt University - coordination
  - \* DataMate / Seafire - industrial partners
  - \* Universities / other partners - welcome!
- \* **You're invited to join!**
  - \* Analysis & brainstorming
  - \* Requirement definition
  - \* Design and implementation
  - \* Testing

### Contacts:

- \* Maciej Brzeźniak <maciekb@man.poznan.pl>
- \* Karsten Asshauer <karsten.asshauer@hu-berlin.de>

**We will be organising periodic technical telcos starting in March 2022 - you're welcome!**

**Stay tuned!**



**CS<sup>3</sup>  
MESH<sup>4</sup>  
EOSC**

---

**Connecting European Data**

**Thank you!**  
Discover more on...

 [cs3mesh4eosc.eu](https://www.cs3mesh4eosc.eu)

 [company/cs3mesh4eosc](https://www.linkedin.com/company/cs3mesh4eosc)

 [CS3org](https://twitter.com/CS3org)

 [CS3MESH4EOSC Project](https://www.youtube.com/channel/UCHKcZEKMaXjCvc3MLFjFxbw)  
<https://www.youtube.com/channel/UCHKcZEKMaXjCvc3MLFjFxbw>

