



Workshop on Cloud Services for Synchronisation and Sharing

A Summary

by Pedro Ferreira

8th edition!

24 - 27 Jan 2022

- 2014 - Genève
- 2016 - Zürich
- 2017 - Amsterdam
- 2018 - Kraków
- 2019 - Roma
- 2020 - København
- 2021 - **Virtual**
- 2022 - **Virtual**



What is it about?

- **Sync & Share Cloud Storage, on-premises, FOSS**
 - Sites and Services: *DESY Sync & Share, CERNBox, BNLBox...*
 - Storage Technology: *EOS, GPFS, CEPH, dCache, ...*
 - EFSS Companies: *ownCloud, Nextcloud, Seafile...*



CERNBox



Nextcloud



What is it about?

- **Integrated Services**
 - Collaborative Tools: *OnlyOffice, Collabora, ...*
 - Research Services: *Jupyter, Open Data Repositories, FAIR metadata ...*
- **Projects and Federations**
 - *European Open Science Cloud, ScienceMesh, HIFIS ...*



What is it about?

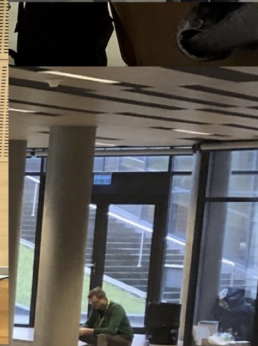
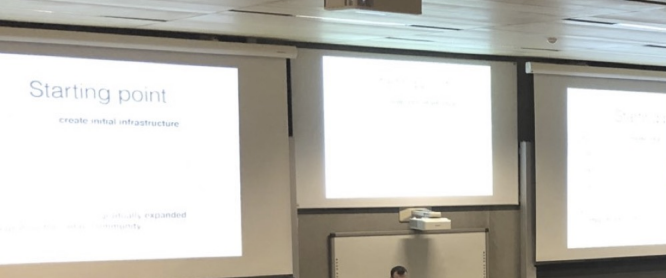
- New: **Decentralized Web and Storage**
 - IPFS, Solid, ...
 - web3, Fediverse...
- **Thematic workshops**
 - Interoperability protocols and APIs
 - ScienceMesh



**Science
Mesh**

Attendance

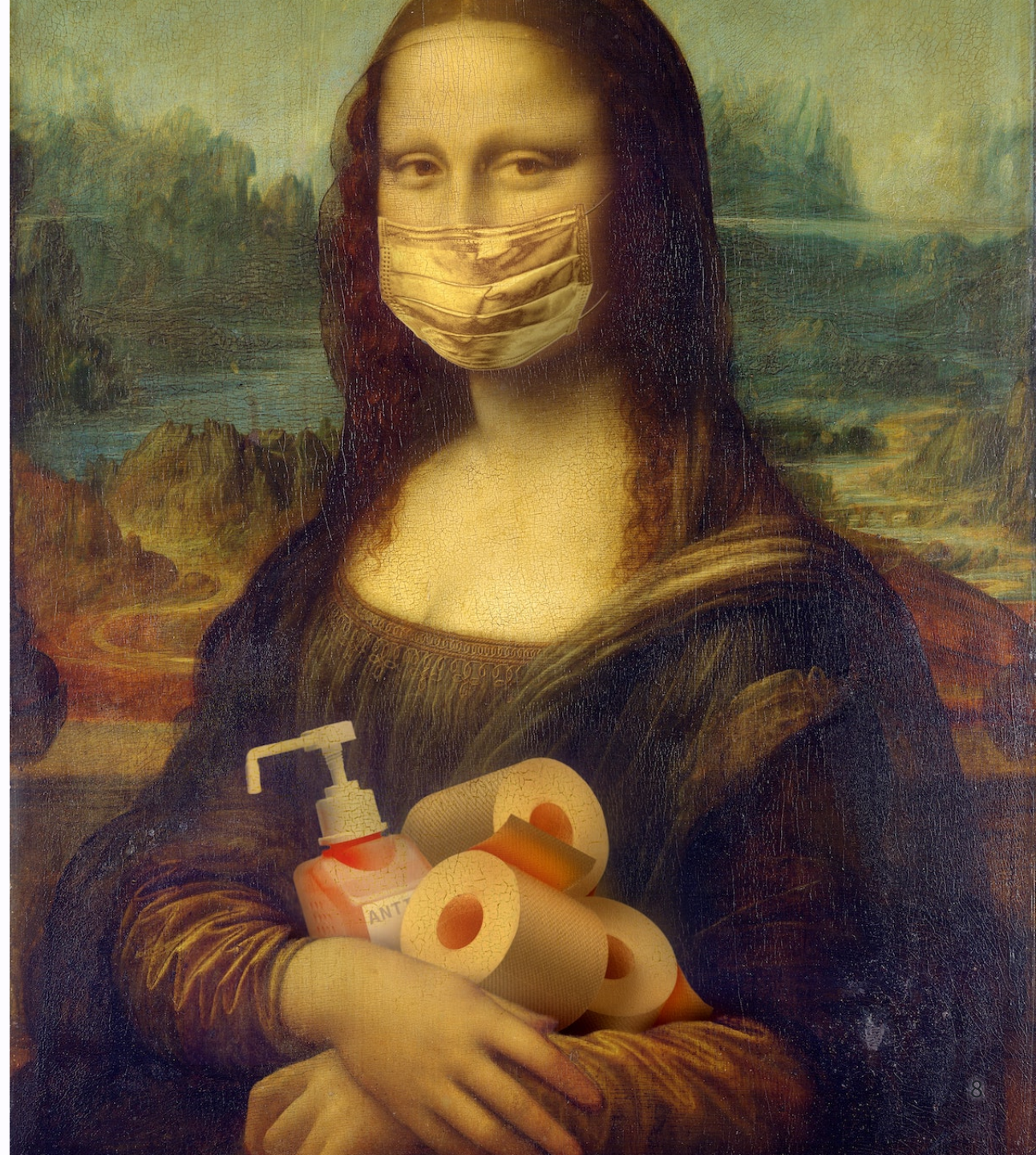
- **280+** registered participants (~310 previous year)
- **~120** different organizations
 - Universities, International Organizations, NRENs, Computing Centers, SMEs, Enterprises, European Commission...



2021 - 2022

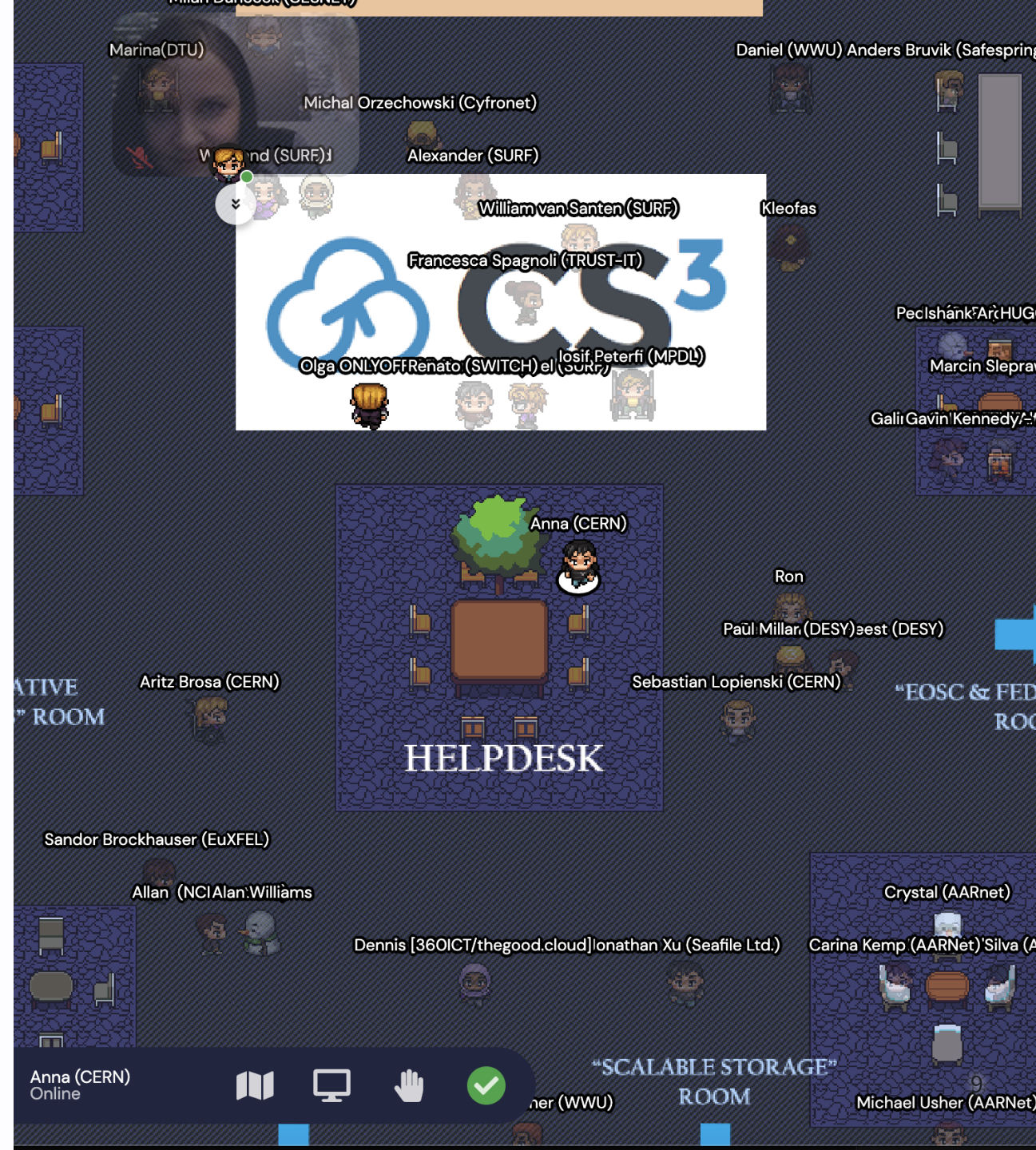
- Needed to go **full virtual** (Zoom)
- Collaborative **tools** (Zoom, gather.town, Indico)

Photo by Yaroslav Danylchenko from Pexels



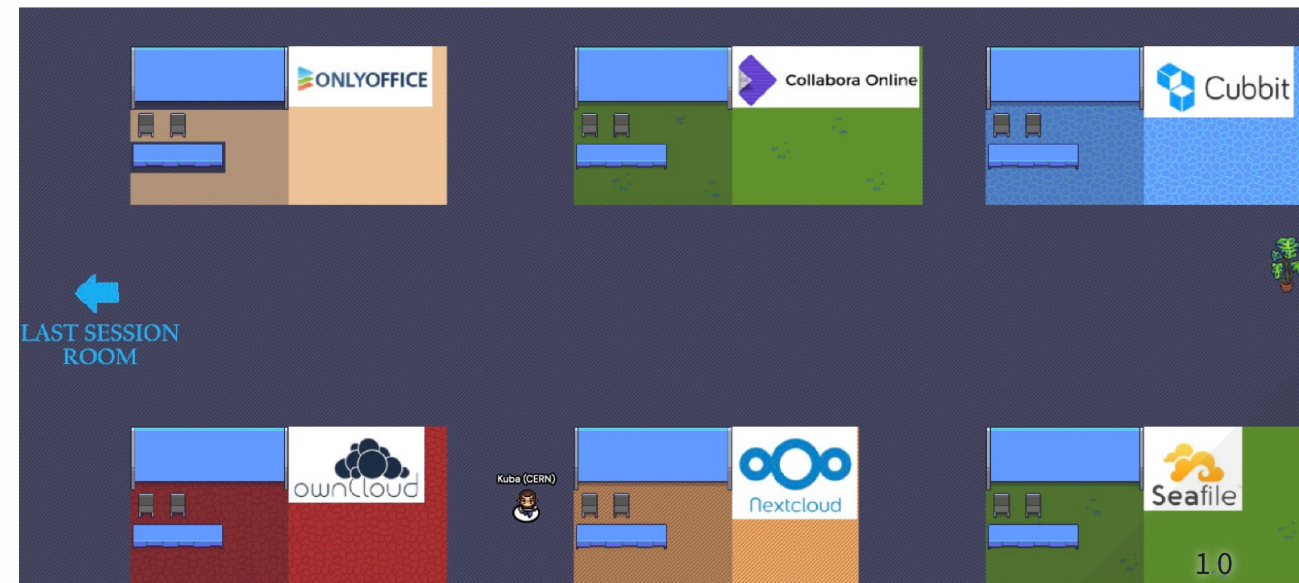
gather.town

- **8-bit-RPG-like** interface 🤖
- Game-like **controls** (arrow keys)
- "Line of sight" **video chats**



This time

- More or less same layout
- Rooms for discussions + exhibitor **booths**
- Usage visibly **lower** than previous year...
- Maybe "Zoom fatigue"?



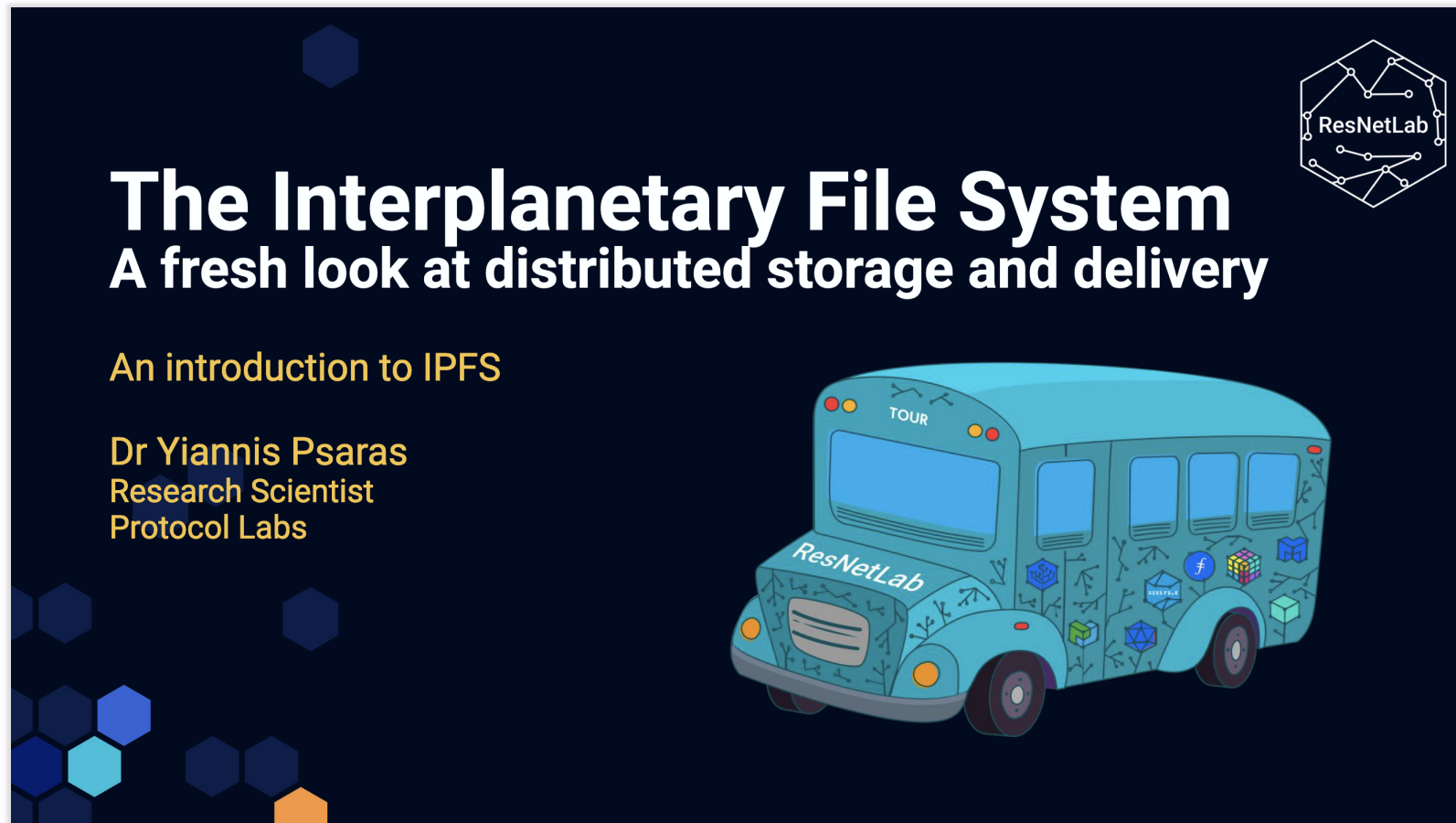
Schedule

| Monday, 24 January 2022 | | Tuesday, 25 January 2022 | | Wednesday, 26 January 2022 | | Thursday, 27 January 2022 | |
|-------------------------|-------------------------------|--------------------------|--|----------------------------|---------------------------|---------------------------|---|
| 08:45 | Good Morning Coffee | 08:45 | Good Morning Coffee | 08:45 | Good Morning Coffee | 08:45 | Good Morning Coffee |
| 09:00 | Introduction and Welcome | 09:00 | Keynote | 09:00 | Keynote | 09:00 | Decentralized Web and Storage Architectures |
| 09:15 | Keynote | 09:45 | Coffee break | 09:45 | Coffee Break | 09:36 | User Stories |
| 10:00 | Coffee break | 10:00 | Federated Infrastructures & Clouds | 10:00 | Scalable Storage Backends | 10:00 | Coffee Break |
| 10:20 | Site Reports | 11:30 | Lunch break | 11:00 | Coffee Break | 10:15 | User Stories |
| 11:30 | Lunch break | 13:00 | Collaboration Products | 11:20 | User Stories | 11:15 | Coffee Break |
| 13:00 | EFSS Products | 14:00 | Coffee Break | 12:20 | Lunch break | 11:30 | Technology for Application Integration |
| 14:30 | Coffee Break | 14:15 | CS3 Org: Governance Campfire Discussion (CS3APIs, REVA, OCM, WOPI,...) | 14:00 | ScienceMesh Workshop | 12:30 | Summary and Conclusions |
| 15:00 | OCM Interoperability Workshop | 15:45 | Chat-away Coffee | 17:45 | Chat-away Coffee | | |
| 16:40 | Chat-away Coffee | | | | | | |

Highlights

Keynote: IPFS: Interplanetary filesystem

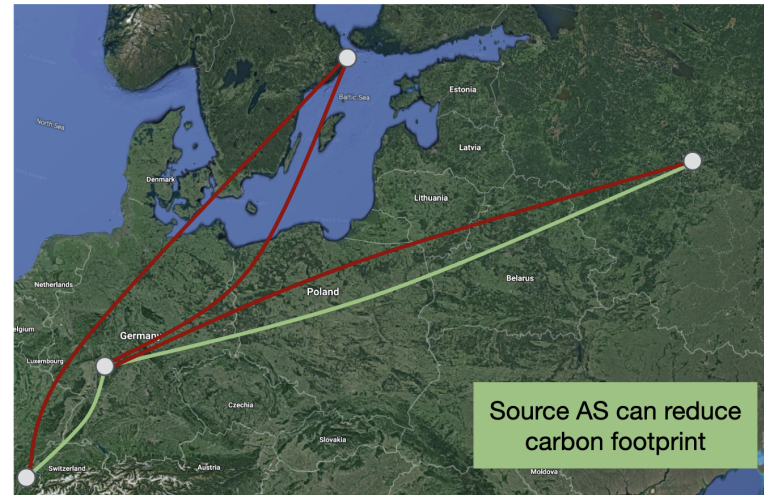
Yiannis Psaras (ProtocolLabs)



Keynote: SCION: Experiencing a new Internet Architecture

Adrian Perrig (ETHZ)

Carbon-intensity difference between paths (1/2)



Path between Swisscom (Geneva) and Extreme Ltd (Moscow)

BGP Path:

Geneva - Stockholm - Frankfurt - Moscow

Greenest SCION Path:

Geneva - Frankfurt - Moscow

Emission savings

Absolute: 0.13 g/GB

Relative: 93%

ETH zürich

SCION

11

Keynote: Digital Market: a level playing field for EU Tech sector

Frank Karlitschek (Nextcloud)

User Data Manifesto 2.0

Defining basic rights for people to control their own data in the internet age



Control
over user data access

Data explicitly and willingly uploaded by a user should be under the ultimate control of the user. Users should be able to decide whom to grant direct access to their data and with which permissions and licenses such access should be granted.



Knowledge
of how the data is stored

When the data is uploaded to a specific service provider, users should be informed about where that specific service provider stores the data, how long, in which jurisdiction the specific service provider operates, and which laws apply.



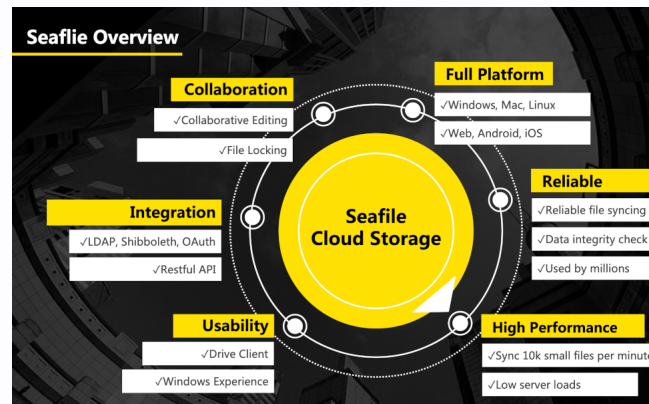
Freedom
to choose a platform

Users should always be able to extract their data from the service at any time without experiencing any vendor lock-in. Open standards for formats and protocols are necessary to guarantee this.

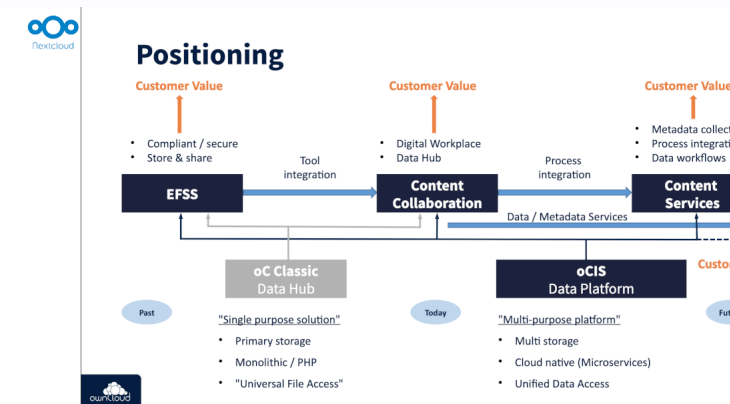
[Read more & participate!](#)

EFSS Platforms

ownCloud, Nextcloud, Seafile



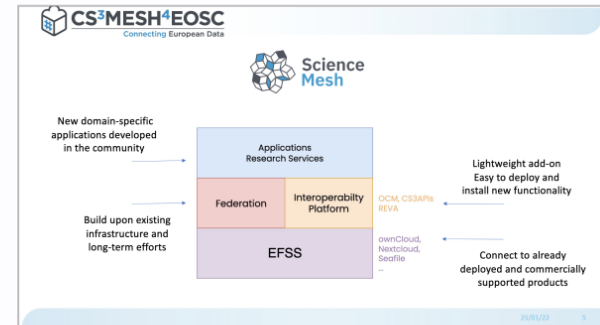
Productivity, Security, On Premises, Open Source



Federated Infrastructures and Clouds



P. Szegedi - Update from the European Commission on Future Evolution of Digital Landscape in Europe



P. Ferreira - ScienceMesh: An Interoperable Federation of EFSS services for European Open Science Cloud

VO FEDERATION IN EFSS

HIFIS Cloud

HELMHOLTZ FEDERATED IT SERVICES

- Making existing IT Services accessible to all of Helmholtz
 - And research partners
 - Users log in to services with their home credentials using Helmholtz AAI

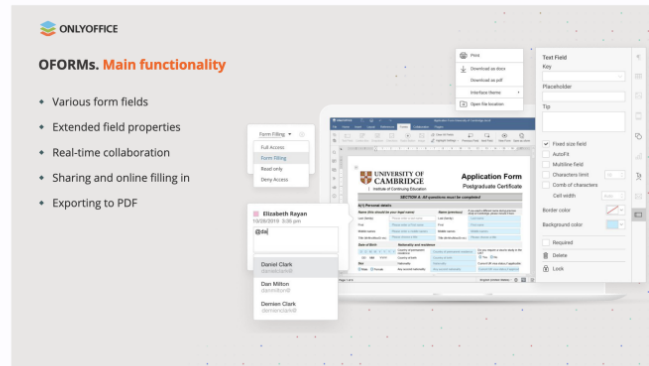
<https://cloud.helmholtz.de>

<https://login.helmholtz.de>

HELMHOLTZ

A. Klotz - HIFIS: VO Federation for EFSS

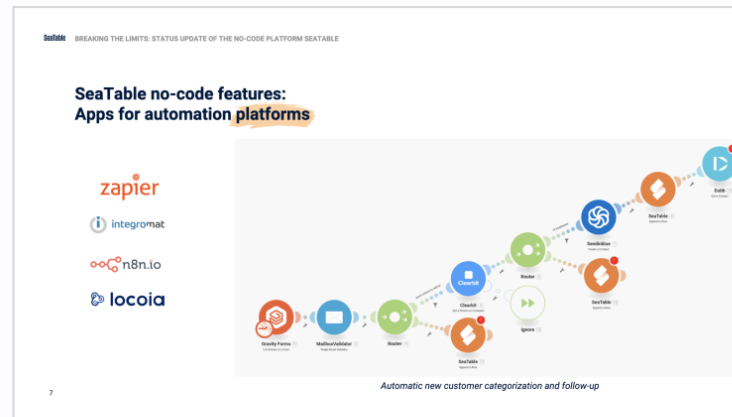
Collaboration Products



G. Goduhina - Introducing ONLYOFFICE Forms

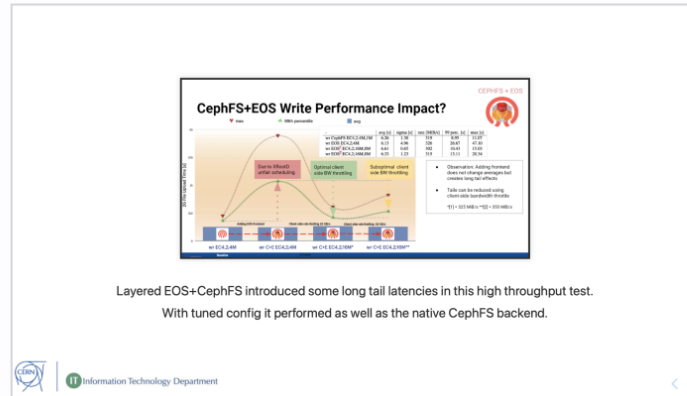


M. Meeks - Collabora Online: easy to deploy and manage document collaboration

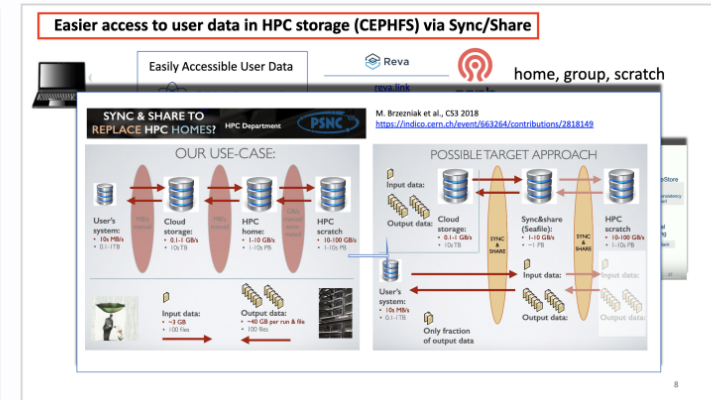


C. Dyllick-Brenzinger - Breaking the limits: short status update of the online spreadsheet solution SeaTable

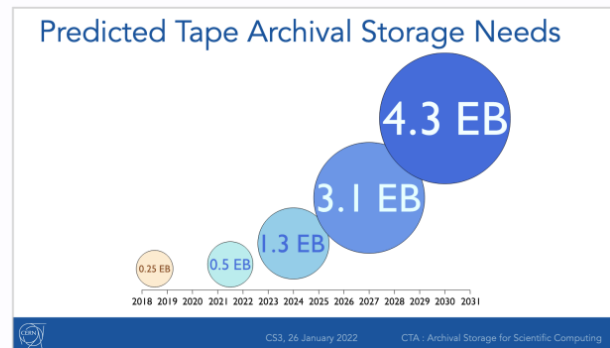
Scalable Storage Backends



R. Valverde - Converging Storage Layers with Virtual CephFS Drives for EOS/CERNBox



D. van der Ster, T. Mouratidis - Sync and Share Access to HPC Resources at CERN



M. Davis - The CERN Tape Archive (CTA): Archival Storage for Scientific Computing

User Stories

The reva storage registry becomes a storage space registry

- Active spaces discovery
- Metadata aggregation and caching
- A "bookmarking" service for spaces

| userid | spaceid | mount path | mtime | address |
|--------|----------|------------|------------|----------------|
| 168481 | 3889144 | juvop/moon | 1643128213 | 127.0.0.1:9157 |
| 168481 | 71468548 | juvop/moon | 1643142213 | 127.0.0.1:9157 |
| 168481 | 44537775 | juvop/moon | 1643128611 | 127.0.0.1:9157 |
| 168481 | 44537775 | juvop/moon | 1643128611 | 127.0.0.1:9157 |
| 168481 | 16124869 | juvop/moon | 1643144213 | 127.0.0.1:9157 |

interview instructions!

a user can change the mount path to be using containers by changing mtime

J. Dreyer - Infinite scale is a design principle

CS³MESH⁴EOSC
Connecting European Data

Distributed Data Science environments

JupyterLab extension (Cs3Api4Lab)

- Integration with ScienceMesh IOP (CS3 APIS)

JupyterLab ScienceMesh: Collaborative Data Science in sync and share environment

01/02/2023 9

M. Sieprawski - JupyterLab+ScienceMesh: Collaborative Data Science

Attendance

- 193 registered users
 - 172 unique users in Zoom
 - Peak of ~90 concurrent users
- 56+ institutions
- 31 speakers
- All CERN departments represented

Background text includes: CERN CH, de, FR, GB, institute, National, NISER, Paris-Saclay, Science, Technical Technology, University

H. Labrador - CERNBox User Forum

iRODS as the Integration Layer

iRODS Clients

Existing Infrastructure

iRODS provides a layer of abstraction which integrates with your pre-existing infrastructure. This flexibility allows your infrastructure to continue to change over time.

T. Russell - iRODS Research Community Requirements Drive Expanded Scale Data Management Features

Thematic Sessions

CS3 Org: Governance Campfire Discussion

- Governance of software products and standards - e.g. CS3APIs, REVA, OCM,...
- Request from industry for a more active role in governance.

Governance


” Repositories under the CS3 GitHub Organization are governed by CERN for the benefit of the community. CERN strives to make the projects under this organization collaborative, open and transparent to ensure that everyone can contribute and have their say on the directions of the projects.

...

The current governance model puts in words how the collaboration currently works in practice today and sets a basic framework for how we collaborate and take decisions in the project. If the nature of the community or contributors changes this governance model may be reviewed and changed if necessary.

<https://github.com/cs3org/.github/blob/master/GOVERNANCE.md>

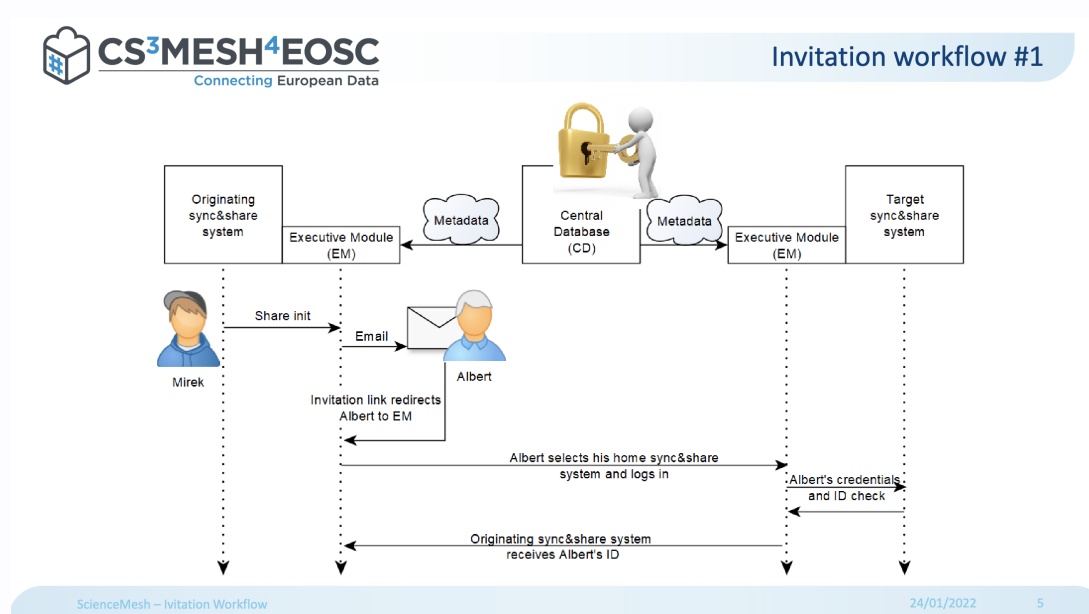
With the increasing maturity of the project we should discuss the governance model.



12

OCM Workshop

- *OpenCloudMesh* - sharing of files between EFSS systems
- Extensions done in the context of ScienceMesh



<https://indico.cern.ch/event/1075584/contributions/4662258/>

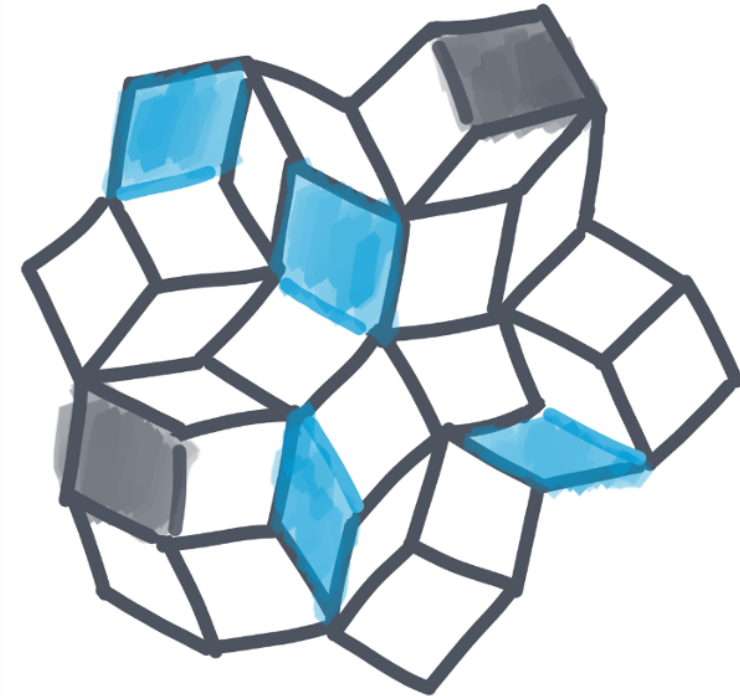
ScienceMesh Workshop



Science
Mesh

ScienceMesh Workshop

- Two parts:
 - "Status updates"
 - Presentations by EOSC Task Forces and Research Initiatives representatives



ScienceMesh Workshop

- Progress validated by community
- New synergies and prospective partners/communities



CS3 2022 Organizing & Programme Committee

Rita Meneses, TRUST-IT



Anna Manou, CERN IT



Ron Trompert, SURF



Hugo Gonzalez, CERN IT



Pedro Ferreira, CERN IT



Tilo. Steiger, ETH Zurich



Kuba Moscicki, CERN IT



Guido Aben, AARNet



It's all online!

<https://indico.cern.ch/e/cs3/2022>

Thank you!

Questions?