

CERNBox: sync, share and science

Diogo Castro
On behalf of the CERNBox team

3 November 2022
HEPiX Autumn 2022 Workshop



Introduction



CERN(Box) user community

- Extremely diverse
 - Around 36K accounts representing 27K people
 - Different work backgrounds
 - Different needs
- Different expectations on how the system should support their daily work

A light gray map of Europe is shown in the background, overlaid with a grid of semi-transparent blue hexagons. Four dark blue rounded rectangular boxes are positioned horizontally across the map, each containing a profession name. The first box, labeled 'Physicists', is highlighted with a red border.

Physicists

Service and
Administration

IT

Engineers



Research community challenges

- **New/improved accelerators in the pipeline**
 - More data: double Run 1,2,3 integrated luminosity in 1y
- **Sync&Share very popular, also for scientific data**
 - e.g. mirror locally output of Grid jobs for fast analysis
 - Plays well with small data formats!
- **HPCs more and more visible in the scientific computing infrastructure**
 - Is it possible to federate HPC/Cloud resources?
 - Or storage as we can federate compute resources?
 - Can we have multiple storage pools behind a single entry-point, with different quality of service?

Scientific Software and Computing in the HL-LHC, EIC, and Future Collider Era

D. Piparo (CERN)
ACAT 2022





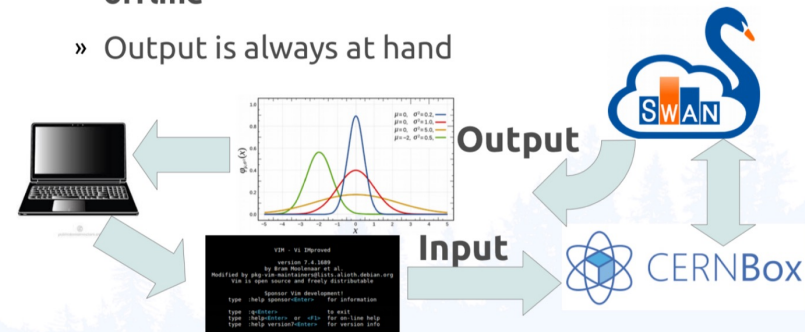
Data Analysis at CERN

- Lxplus/Lxbatch
 - Access to all EOS instances
 - AFS shared FS for submission
- SWAN
 - CERNBox/EOS as home dir
 - Access to all EOS instances
- HPC
 - CephFS for submission
 - Access to all EOS instances
- Spark
 - Access to all EOS instances

EOS provides data and analysis code

How SWAN works for me

- Usage with **cernbox** and **EOS** is great
 - » Develop locally, run your code **both on-line and offline**
 - » Output is always at hand



Sync&Share bridges the gap between local dev and services



CERNBox and HEP



Challenges: from Sync&Share to Integrated Scientific Platform

- **Direct Access to Underlying Storage**
 - Assumption that OC is the single entry point
 - Caching might cause inconsistent views and conflicts
- **Homogeneous Access to Shared Data**
 - Need to sync ACLs to storage
- **Data Ownership**
 - Owner of the data is the user running the server (apache/www-data)
 - To increase security and give users full control of their data, the owner should be the real user
- **Native File Versioning and native Recycle bin**
 - Some storage backend provide these, so the OC layer duplicates the functionality
- **Redundant Expensive Calls and inefficient design**
 - Unique filenames for shares requires checking all existing shares before creating a new one

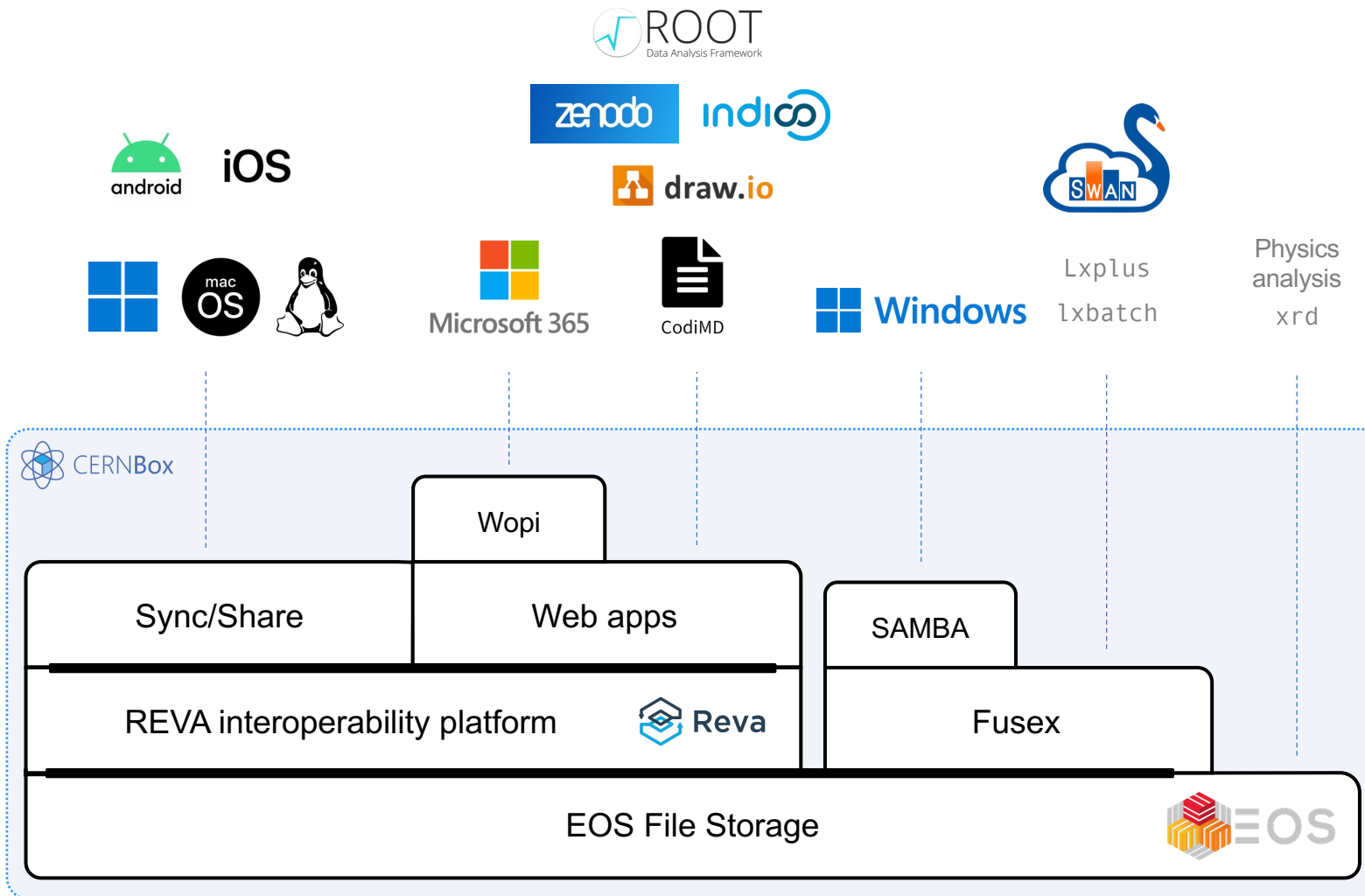


A new solution was required!





Global architecture

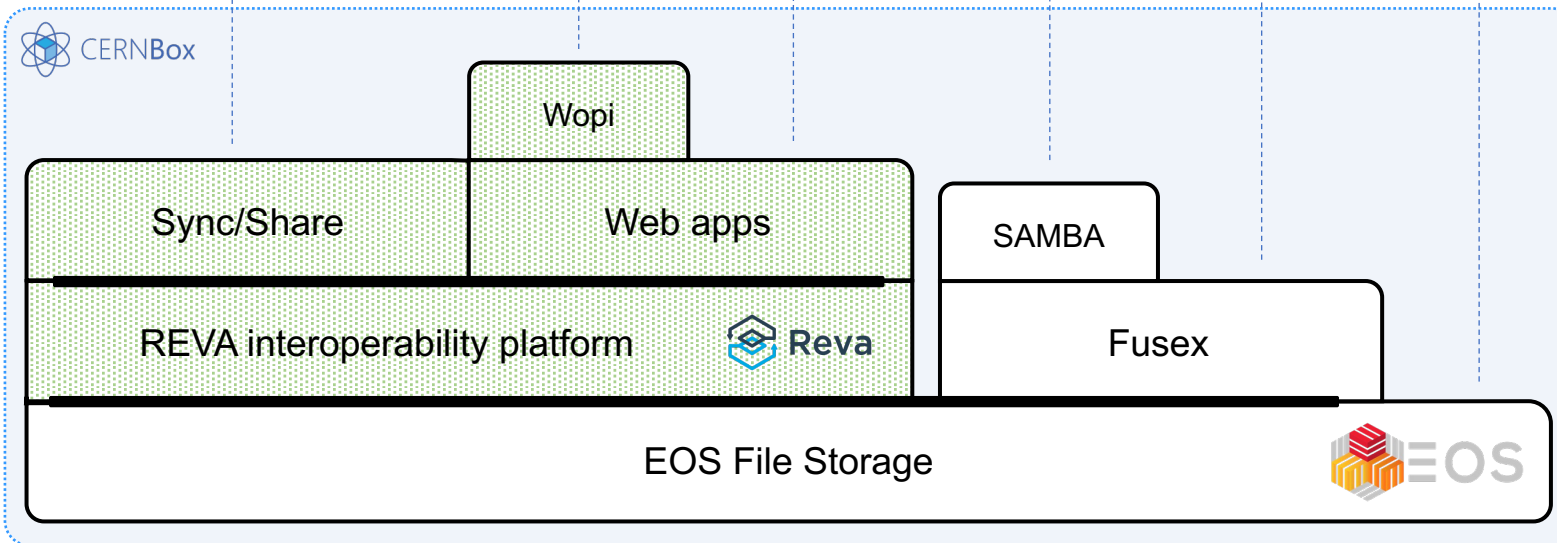


- 36k User accounts
- Sync clients for all major platforms (VFS support on Windows)
- 30k fuse clients
- 15 production clusters
- 18 PB general storage
- 400 PB Physics storage
- <https://reva.link/>





Global architecture



- 36k User accounts
- Sync clients for all major platforms (VFS support on Windows)
- 30k fuse clients
- 15 production clusters
- 18 PB general storage
- 400 PB Physics storage
- <https://reva.link/>
- <https://owncloud.com/infinite-scale/>

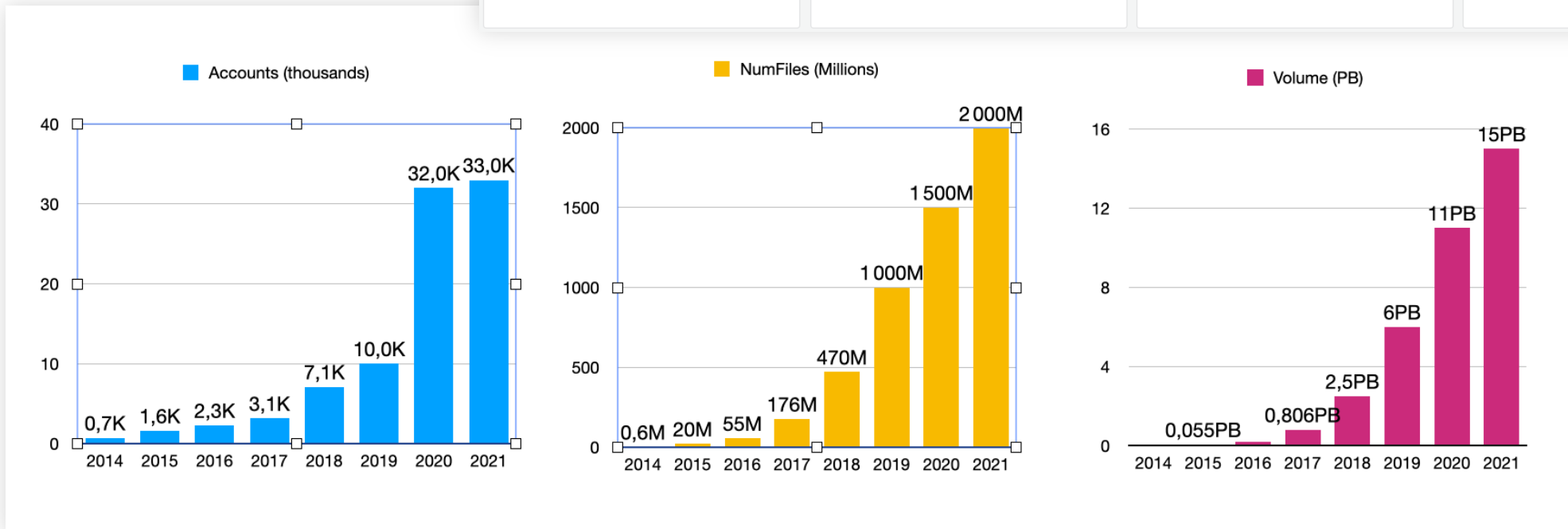




Overall numbers

- 30 LHC and non-LHC experiments home dirs
- Skyrocketing usage

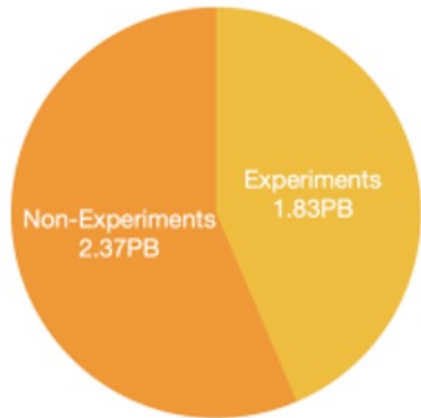
Last 12 months



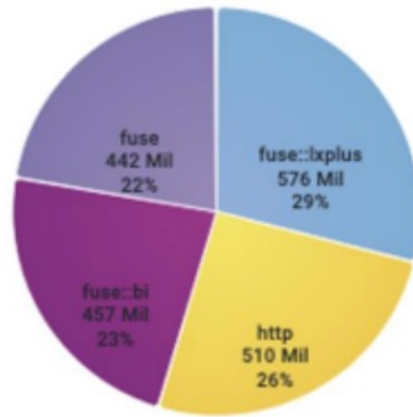


Overall numbers

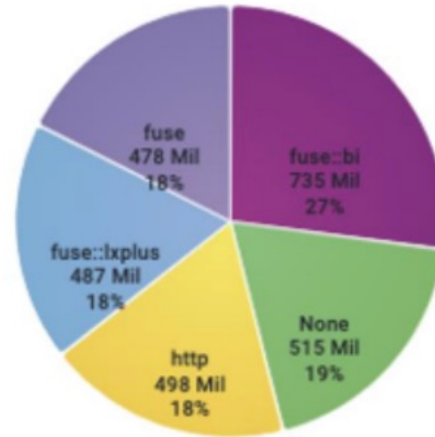
VOLUME DISTRIBUTION



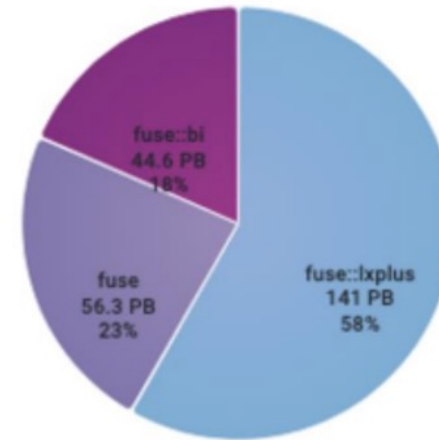
WRITE FILES
74% FUSE



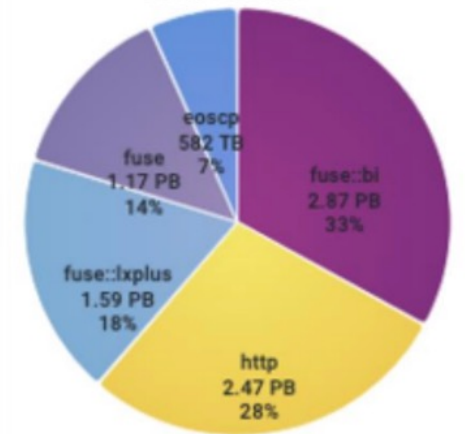
READ FILES
82% FUSE+XROOT



READ VOLUME
100% FUSE



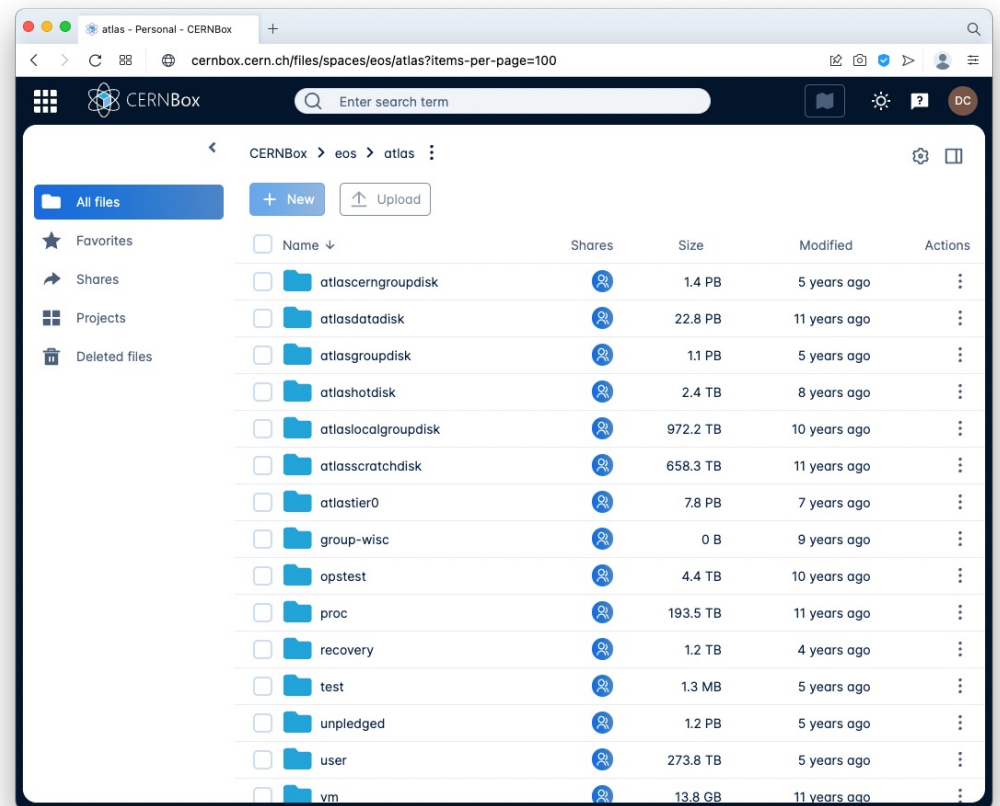
WRITE VOLUME
72% FUSE
28% SYNC





ATLAS Spaces

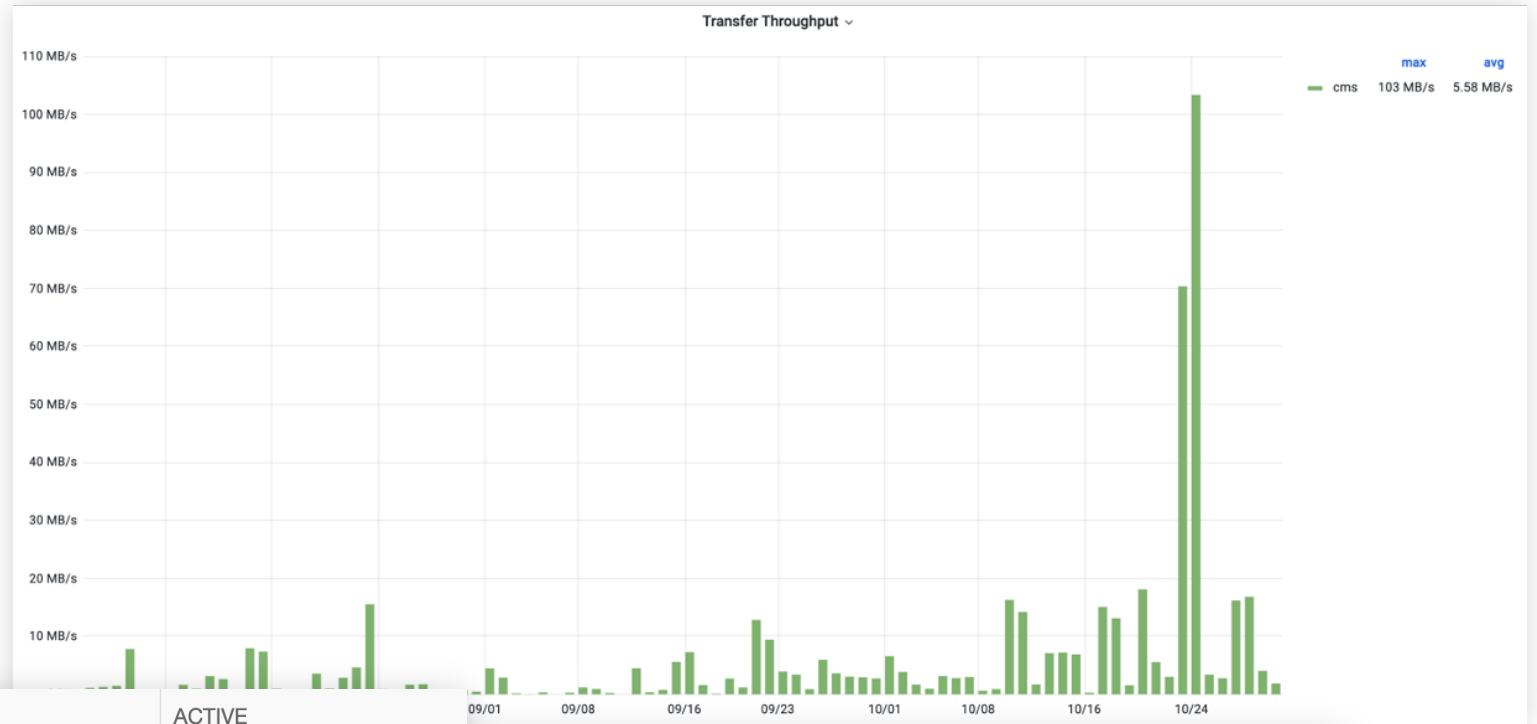
- ATLAS is no longer issuing personal analysis spaces for users
 - They are now in CERNBox
 - 4.2K accounts, 1PB used space
- Spaces now available via Web and Sync Clients
 - They were already exposed through FUSE





CMS Asynchronous Stage Out jobs

- CERNBox is available as an ASO for CRAB jobs
 - Users can run CMSSW and output directly to CERNBox
 - Transfers are orchestrated with FTS from CMS VO using XRootD



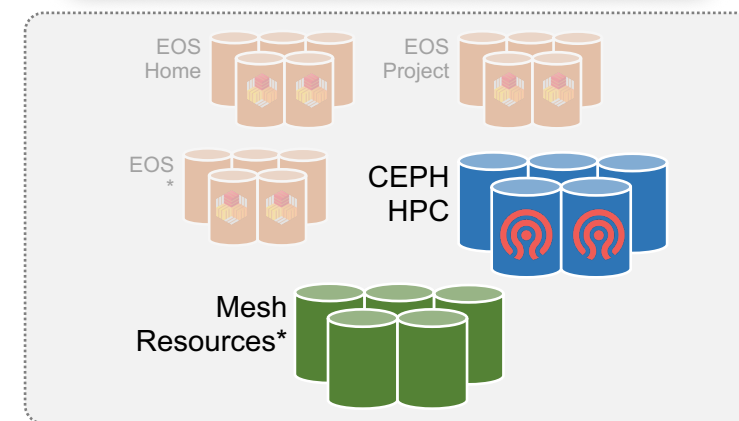
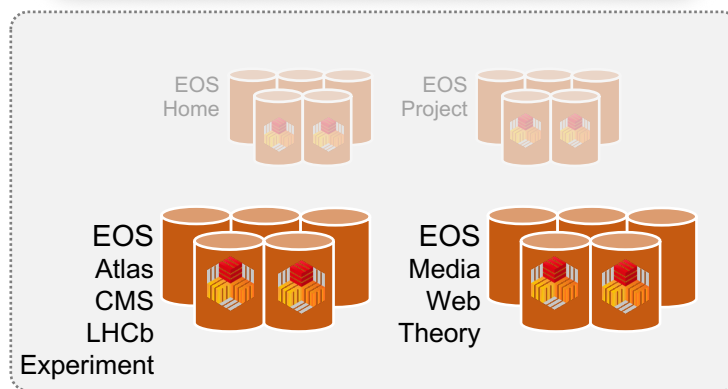
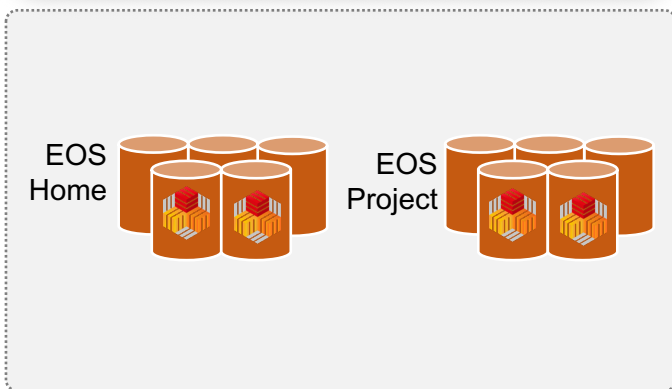
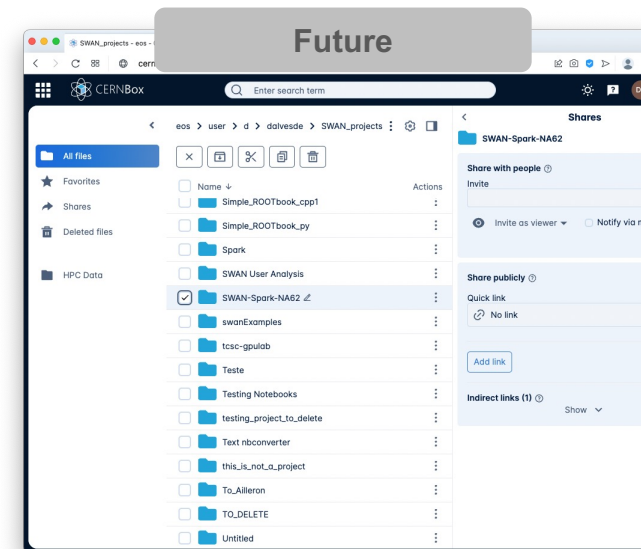
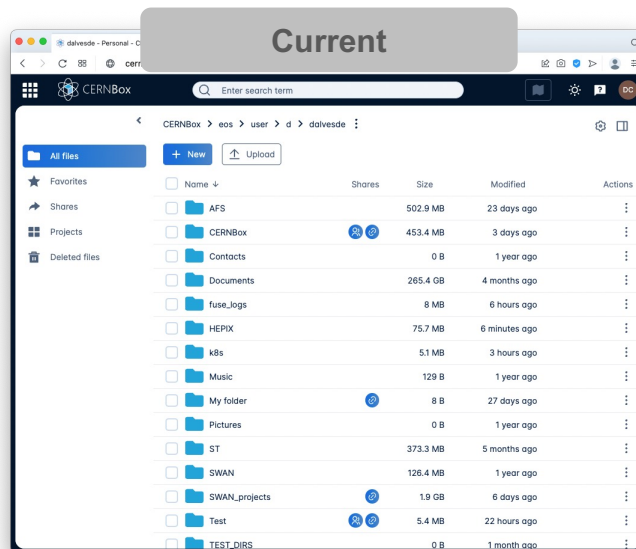
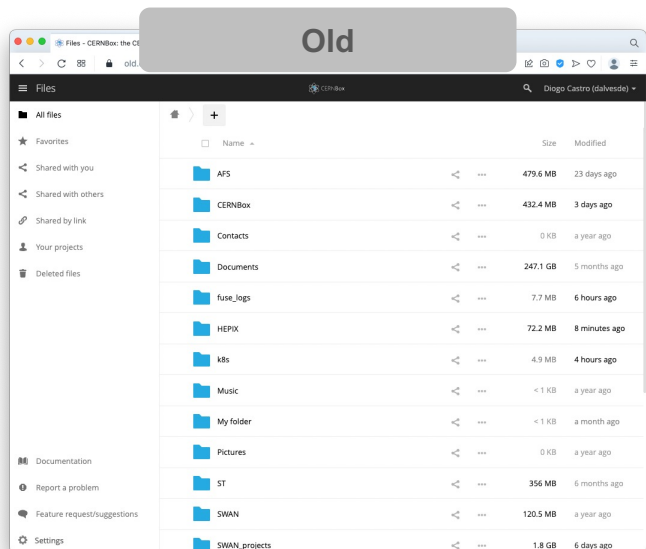
T2_CH_CERNBOX	T2_CH_CERN	ACTIVE
T3_CH_CERNBOX	T3_CH_CERNBOX	ACTIVE



CERNBox future



CERNBox: federation of storages using REVA



*CS3MESH4EOSC: <https://cs3mesh4eosc.eu/>





HPC/CEPH integration

- HPC service at CERN
 - Applications and use cases that do not fit the standard batch HTC model, typically parallel MPI applications.
 - Ex: Computation Fluid Dynamics, Beam simulation, plasma simulation, ...)
 - Uses Ceph FS between submission and work nodes
- CERNBox: give access to user data from HPC service
 - Sync
 - Transform SWAN into a submission node (without having to share Ceph secrets)
- Reva Ceph FS “simplified” storage backend
 - PoC available, some features (sharing/ACLs, snapshotting) not available





Integrating with FTS/Rucio

- The CERNBox backend (Reva) supports HTTP Third-Party copy
 - As defined by the WLCG DOMA working group
 - Demonstrated to work with FTS and Rucio
- Authentication based on OIDC, adopting the ESCAPE IAM for federated ids
 - Aim is to support data transfers across WLCG sites
- Authorization based on token scopes a.k.a. claims
 - Still lots of discussions ongoing around this topic
- GUI is expected to be based on JupyterLab
 - Leveraging the work done within the RUCIO team and ESCAPE project



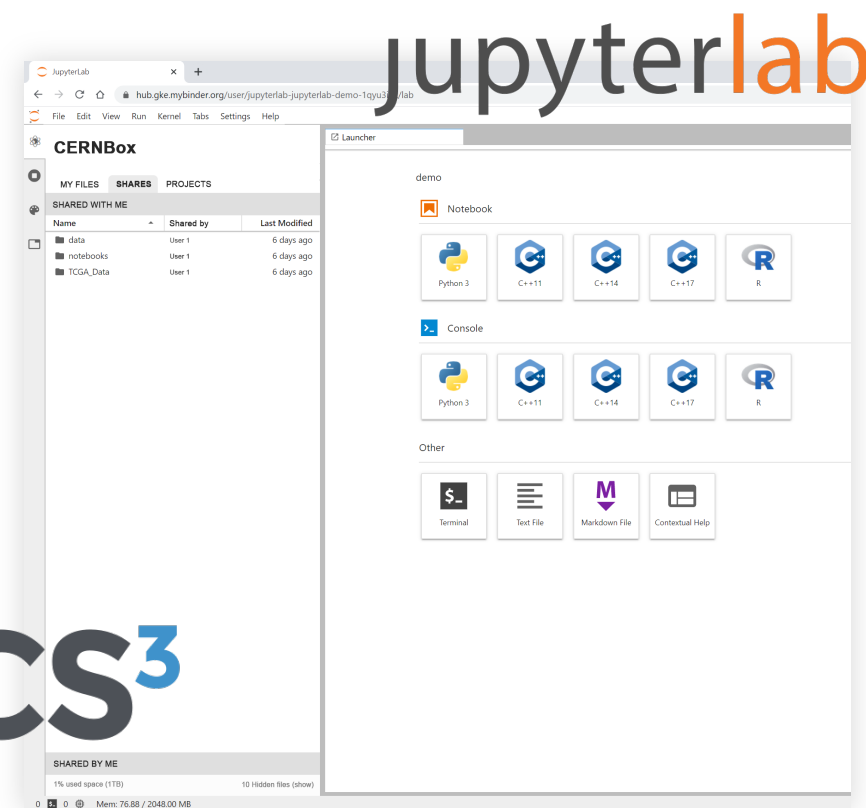


SWAN/Jupyterlab integration

- CERNBox provides the home directory of SWAN
 - FUSEX mounted + sharing API
 - GPU, Spark, HTCondor integrations
- Moving from CERN specific to generic Lab extension
 - Using CS3 APIs
 - Compatible with ownCloud, nextCloud
- Fostering collaborative work
 - Full sharing capabilities
 - Locking protection for concurrent access



CS³



MVP: <https://github.com/sciencemesh/cs3api4lab>

CS3 APIs docs: <https://github.com/cs3org/cs3apis>



Sync&share in HEP

Do you also interface your Sync&Share with the scientific data repositories?
Why not?

Are you interested in joining the Mesh?

Do you have Jupyter services? How do you deploy them?

Do you have interest in some of our integrations?



- Packaged deployment of CERN software
 - Includes CERNBox, EOS and SWAN
 - Helm charts (k8s)
 - Single machine test deployment*
- Pluggable/configurable
 - Can be integrated with already deployed services (i.e auth, EOS)
- Not as complete as the CERN deployment
 - Less integrations but good starting point

* Mboxed: <https://github.com/sciencebox/mboxed>
<https://github.com/sciencebox/sciencebox>





CERNBox : sync, share and science

Thank you

Diogo Castro
diogo.castro@cern.ch

