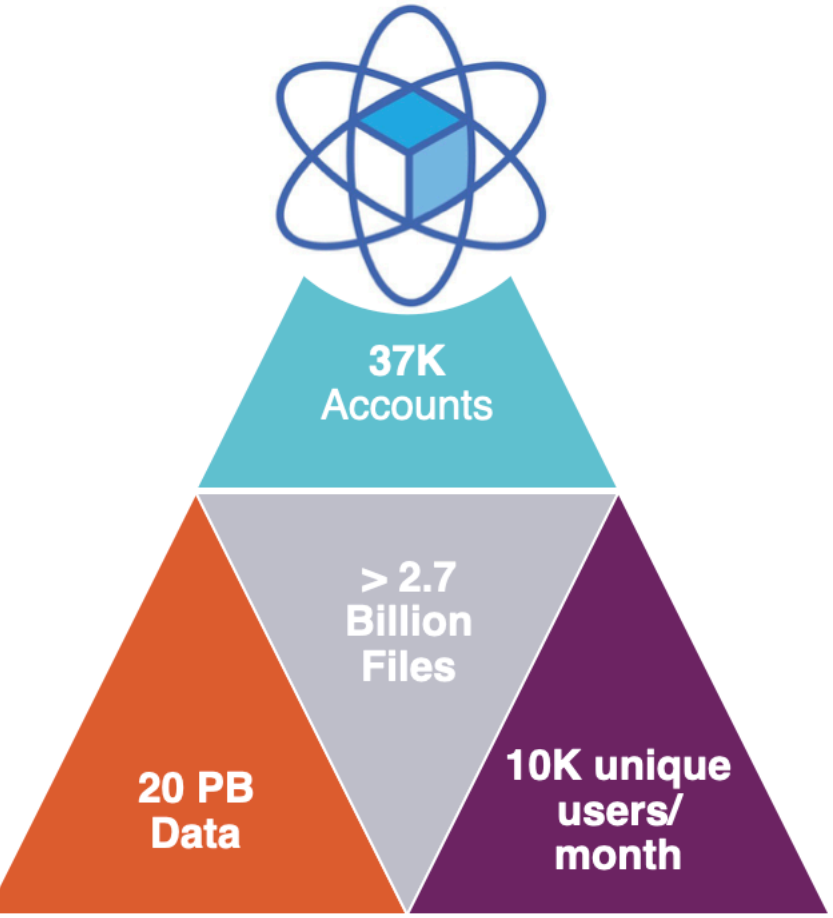
Standing innovation for 10 years CERNBox (2013-2023)

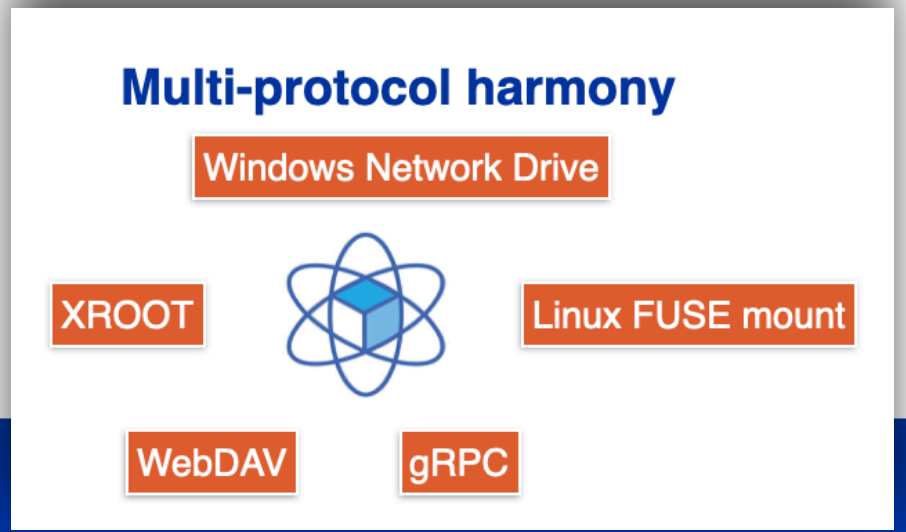
Hugo González Labrador (Project Lead), CS3 Barcelona, 2023

CERNBox today

Sync and Share	• Dropbox-like use-case • Offline access to data • Universal access to data
Web Apps	• Real time collaboration • Office documents
Online filesystem access	• SAMBA (Windows) • FUSEX (Linux)
Physics integration	• ROOT • SWAN • LXPLUS/LXBATCH



Multi-protocol harmony



The tech stack

ownCloud

❤

EOS

10 years ago

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Data & Storage Services

CERN IT Department

Prototyping a file sharing and synchronisation platform with ownCloud

Jakub T. Moscicki
Massimo Lamanna

CERN IT-DSS

CHEP 2013 - Amsterdam

Gateway to the future

- A **unified platform** integrated with physics data storage
- **Federated** "dropbox" service for HEP community
  - … and possibly in wider science
- Novel ways for supporting specialized **scientific workflows**
  - based on a common sharing and syncing platform
- Novel ways of **delivering home directories** in the virtualized IT environment
  - local folder replica lives within the VM snapshot

…however, first we need to positively address the classic Dropbox use-case…

10 years ago

Where are we in 2023?

Data & Storage Services

CERN IT Department

Prototyping a file sharing and synchronisation platform with ownCloud

Jakub T. Moscicki
Massimo Lamanna

CERN IT-DSS

CHEP 2013 - Amsterdam

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Gateway to the future

- A **unified platform** integrated with physics data storage
- **Federated** "dropbox" service for HEP community
  - … and possibly in wider science
- Novel ways for supporting specialized **scientific workflows**
  - based on a common sharing and syncing platform
- Novel ways of **delivering home directories** in the virtualized IT environment
  - local folder replica lives within the VM snapshot

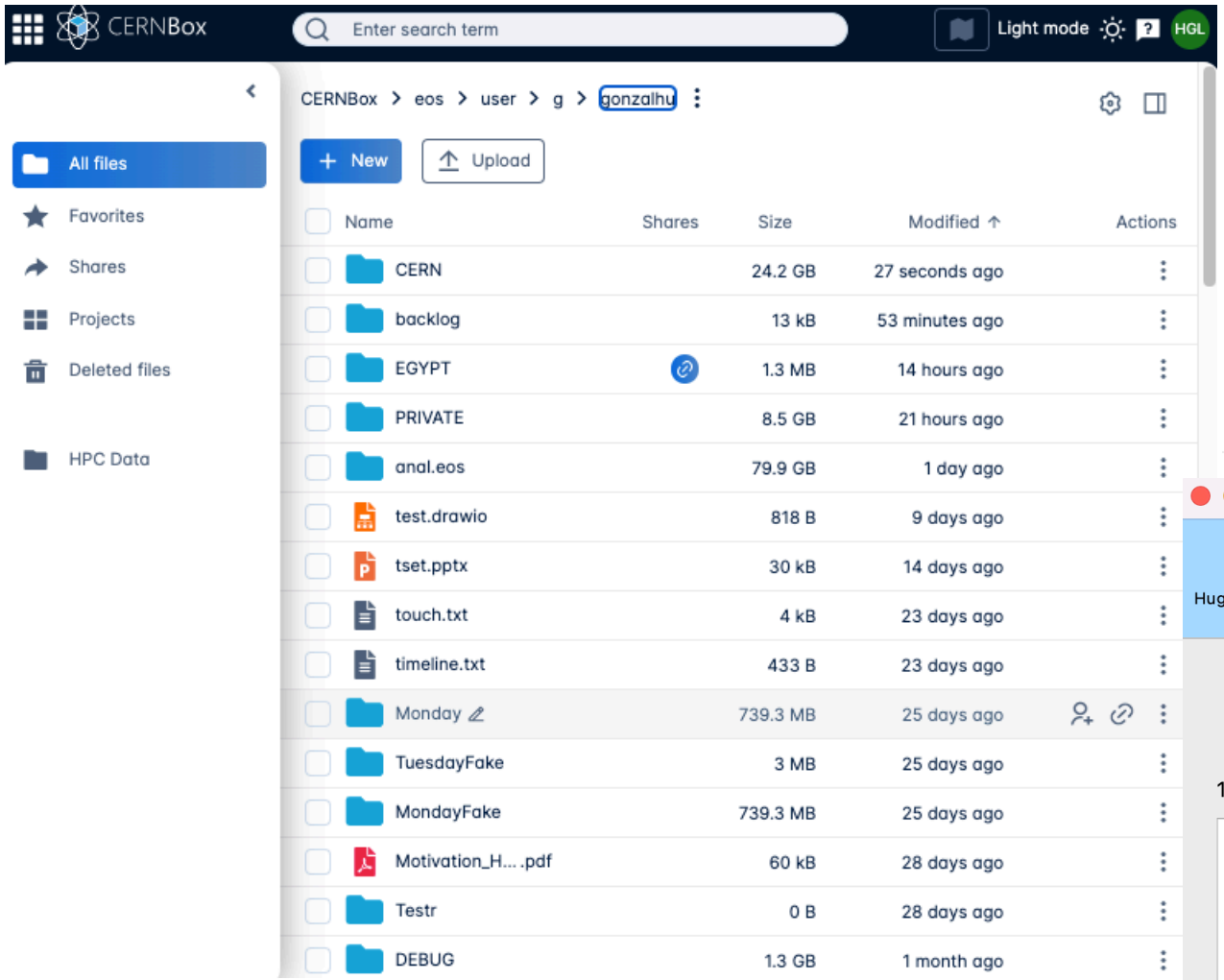
…however, first we need to positively address the classic Dropbox use-case…

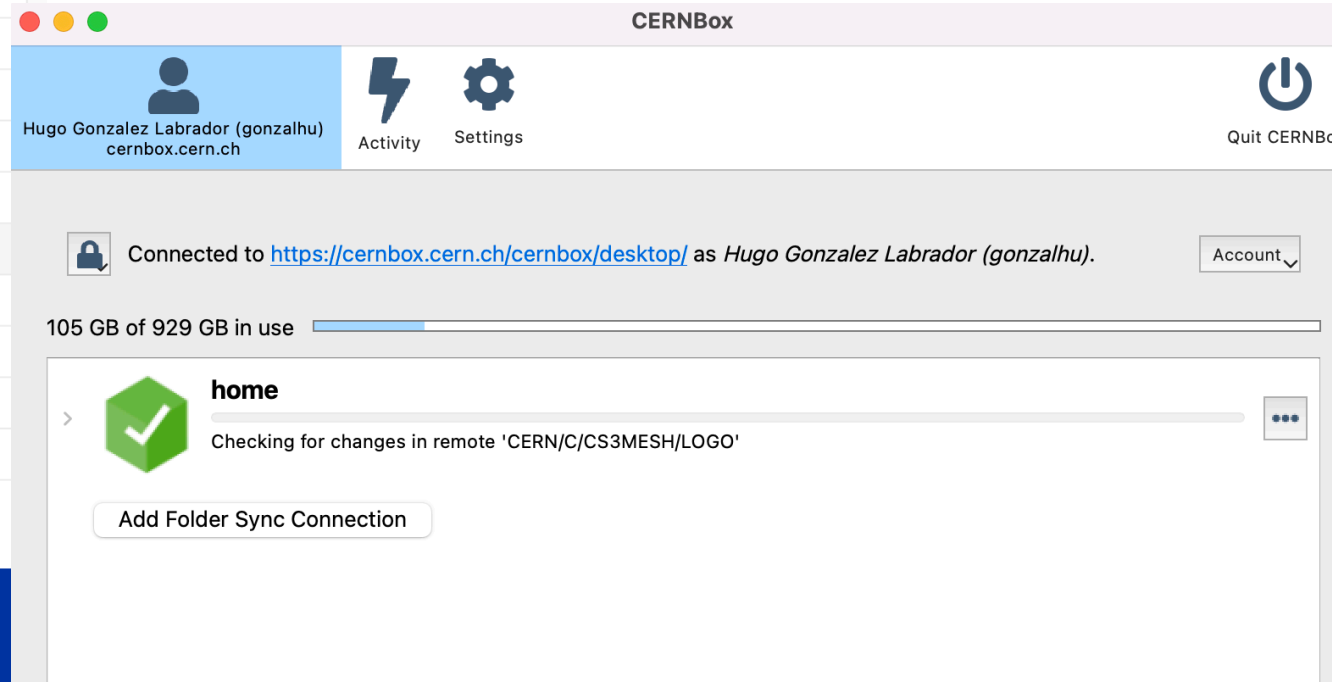
Gateway to the future (as seen in 2013)

A) Classic DropBox use-case
B) Unified platform integrated with physics storage
C) Support scientific workflows
D) Delivering home directories
E) Federated dropbox service for HEP community

A) Classic Dropbox use-case

All major platforms
(Mac OS, Linux, Windows,
Android, IOS)





A) Classic Dropbox use-case

RUN3

The third run of the Large Hadron Collider has successfully started

A round of applause broke out in the CERN Control Centre on 5 July at 4.47 p.m. CEST when the Large Hadron Collider (LHC) detectors started recording high-energy collisions at the unprecedented energy of 13.6 TeV

5 JULY, 2022

Celebrations at the CERN control centre (CCC) to mark the start of LHC Run 3 (Image: CERN)

A round of applause broke out in the CERN Control Centre on 5 July at 4.47 p.m. CEST when the Large Hadron Collider (LHC) detectors switched on all subsystems and started recording high-energy collisions at the unprecedented energy of 13.6 TeV, ushering in a new physics season. This feat was made possible thanks to the operators who had worked around the clock since the restart of the LHC in April to ensure the smooth beginning of these collisions with higher-intensity beams and increased energy.

After over three years of upgrade and maintenance work, the LHC is now set to run for close to four years at the record energy of 13.6 trillion electronvolts (TeV), providing greater precision and discovery potential. Increased collision rates, higher collision energy, upgraded data readout and selection systems, new detector systems and computing infrastructure: all these factors point to a promising physics season that will further expand the already very diverse LHC physics programme!

Pictures of the day are available here.

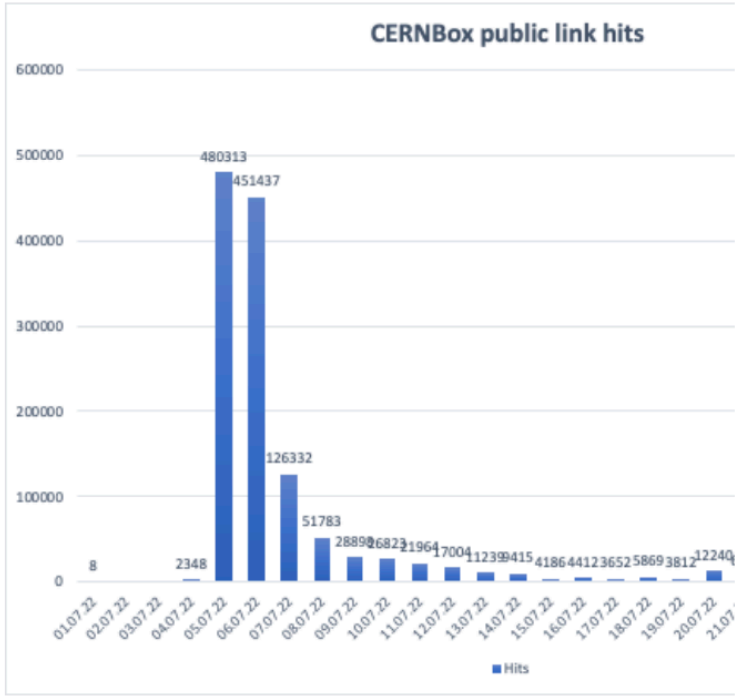
CERNBox

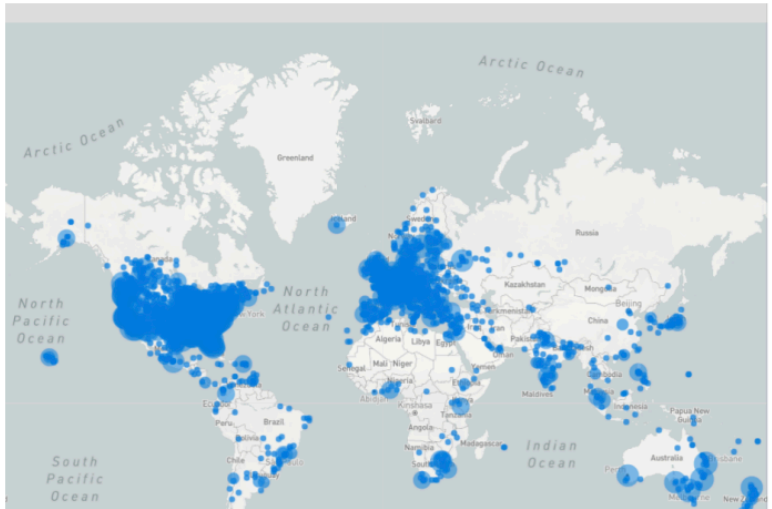
https://cernbox.cern.ch/index.php/s/EacPckkCMFcJ8ya

[redacted]

Name

ALICE_16h02mn_002.jpg
ALICE_16h54mn_015.jpg
ALICE_17h45mn_029.jpg
ALICE_17h50mn_031.jpg
ALICE_17h50mn_032.jpg
ALICE_18h10mn_036.jpg
ATLAS_15h47_015.JPG
ATLAS_16h47_038.JPG
ATLAS_16h47_062.JPG
ATLAS_16h48_043.JPG
ATLAS_16h49_009.JPG
ATLAS_16h50_045.JPG
ATLAS_16h55_047.JPG
ATLAS_16h55_049.JPG





139 COUNTRIES 4830 CITIES

Copy URL New Map

Worldwide	24701 IPs
United States	18323 IPs
Canada	845 IPs
United Kingdom	628 IPs
Germany	541 IPs
Switzerland	525 IPs

B) Unified platform integrated with physics storage

Universal access
Real time collaboration

draw.io
ROOT Data Analysis Framework
ONLYOFFICE
Collabora
CodiMD
jupyter

Offline access
(mobile, desktop, tablet)

ownCloud Sync Clients

POSIX online filesystem

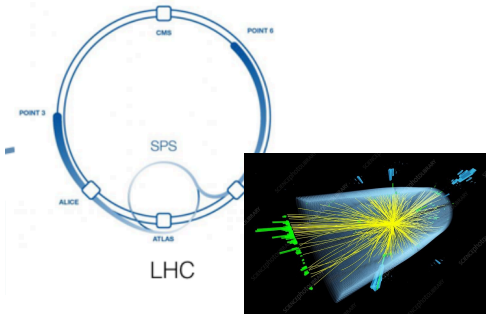
indico
CDS
zenodo
AV
...

Windows online filesystem

Windows

LINUX

Physics analysis



Web Apps

SAMBA

ownCloud Sync & Share

Reva

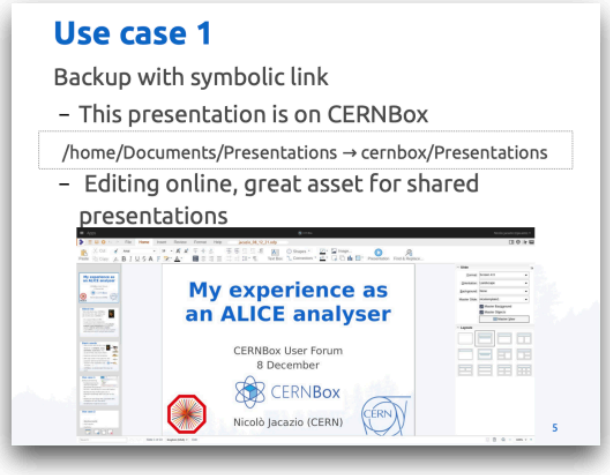
FUSEX

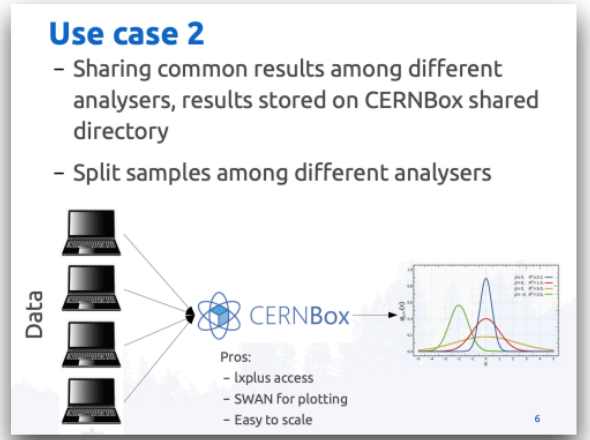
xroot

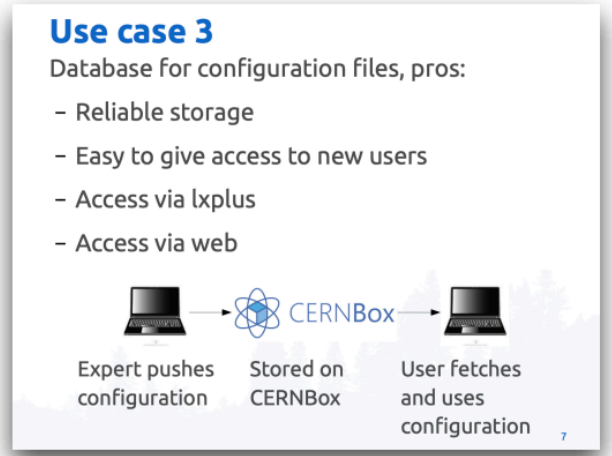
EOS for Home and Project Spaces (9 instances)

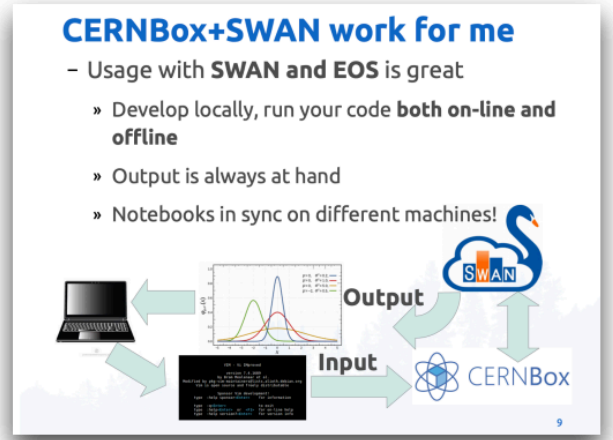
C) Support scientific workflows (SWAN and CERNBox)

Sync & Share storage integration tier — examples







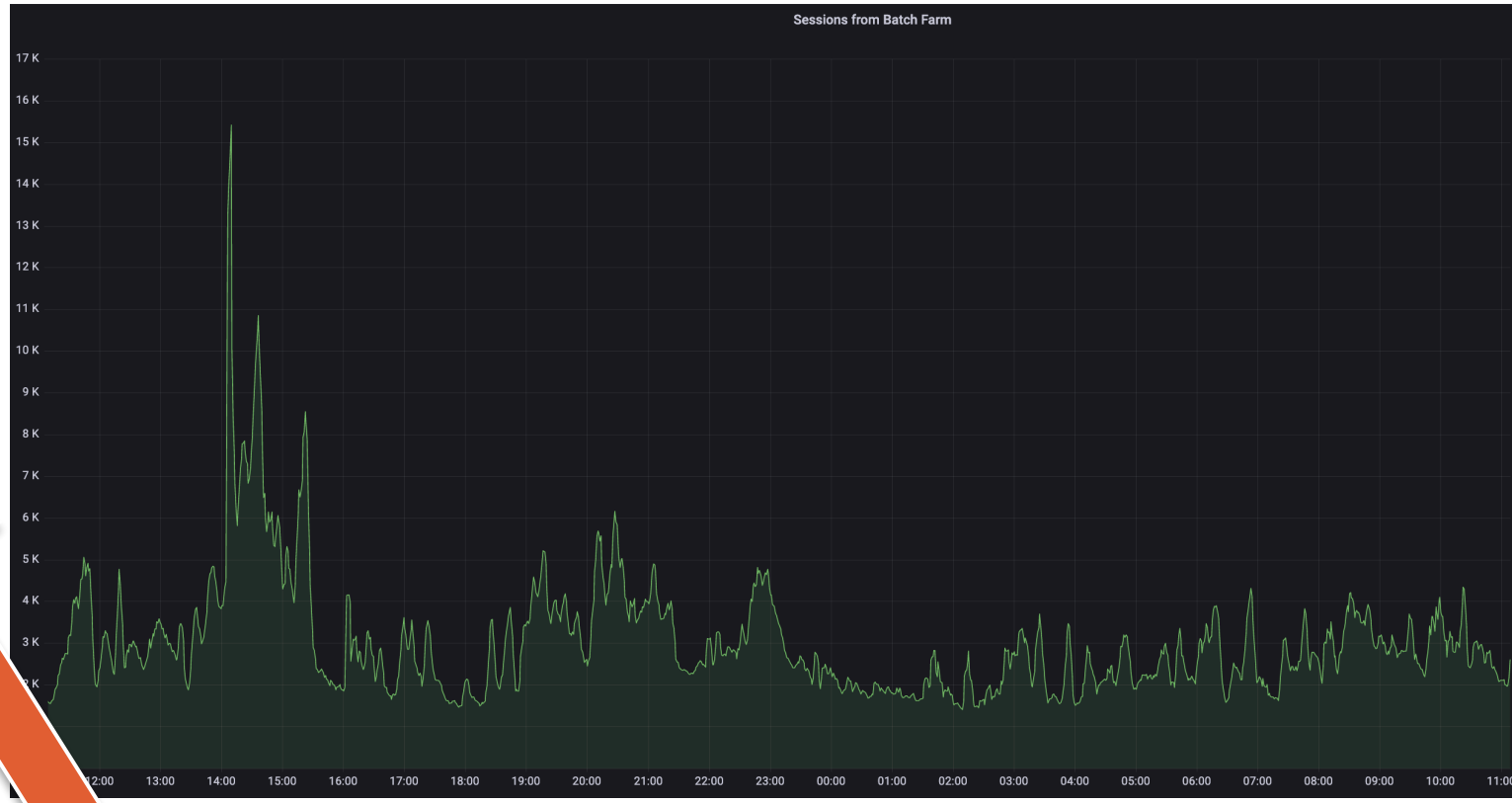


Full presentation

My experience as a ALICE analyser, Nicolo Jacazio, CERNBox User Forum, 2021

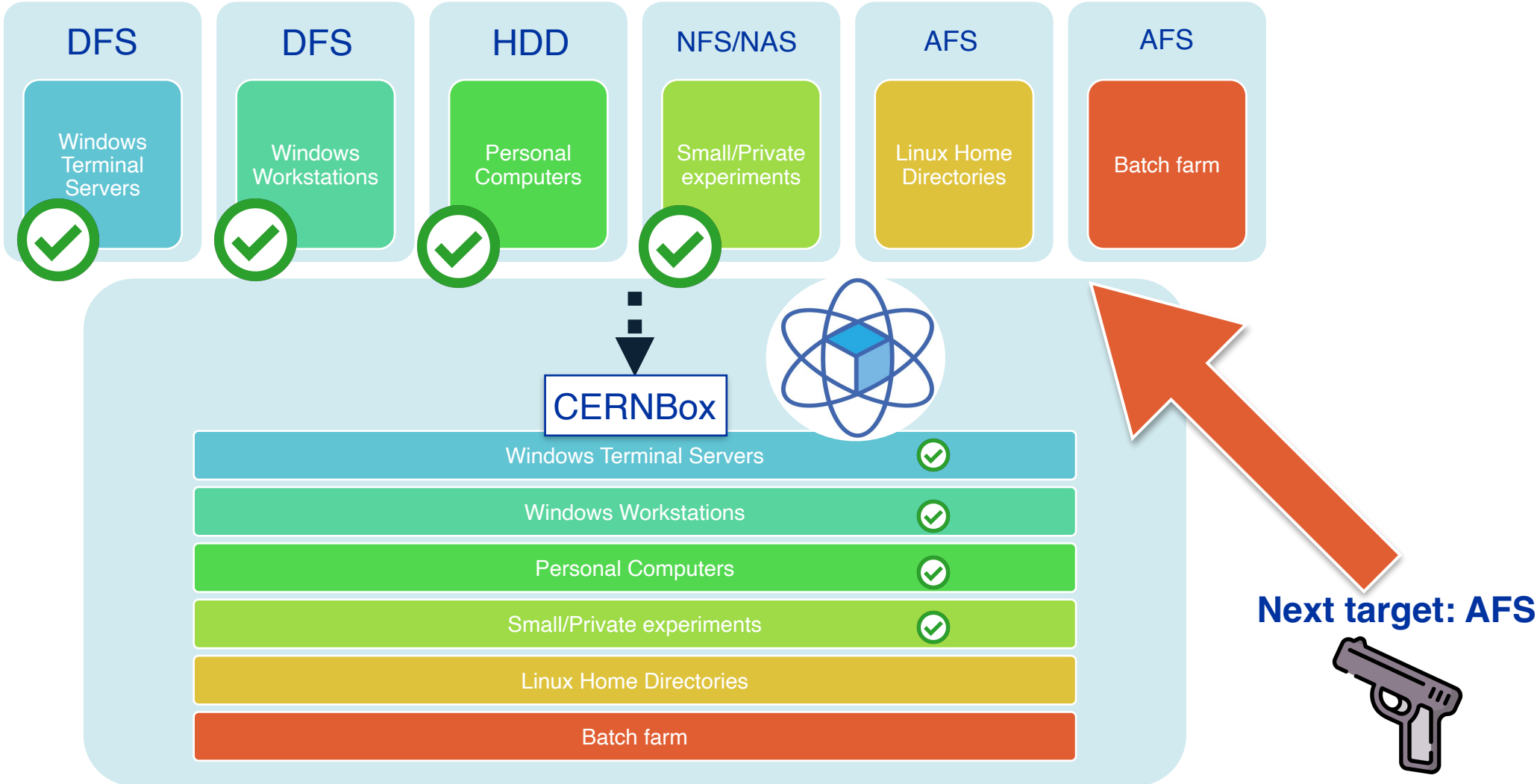
C) Support scientific workflows (Batch computing)

gsiftp://cmseos-gridftp.fnal.gov	22381053
gsiftp://eoscmsftp.cern.ch	13241055
srm://dcache-se-cms.desy.de	12322216
gsiftp://eosuserftp.cern.ch	12097125
gsiftp://cms-gridftp.rcac.purdue.edu	9060208



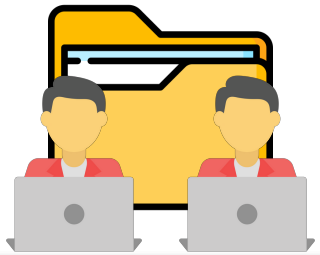
A CMS perspective about CERNBox, Danilo Piparo, CERNBox User Forum, 2021

D) Delivering Home Directories (consolidation)

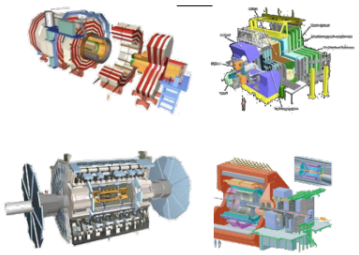


D) Delivering Home Directories (and beyond)

37K user home directories

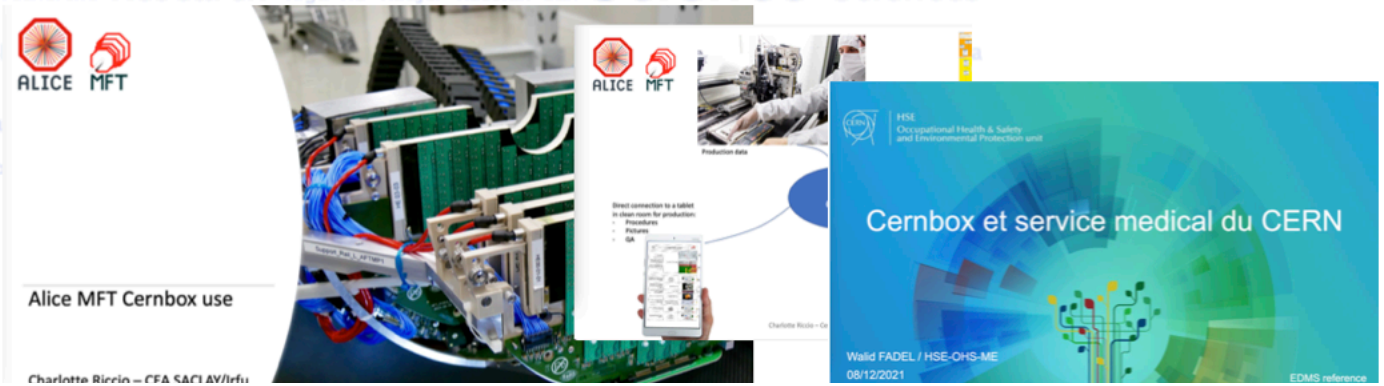


+1K Project Areas

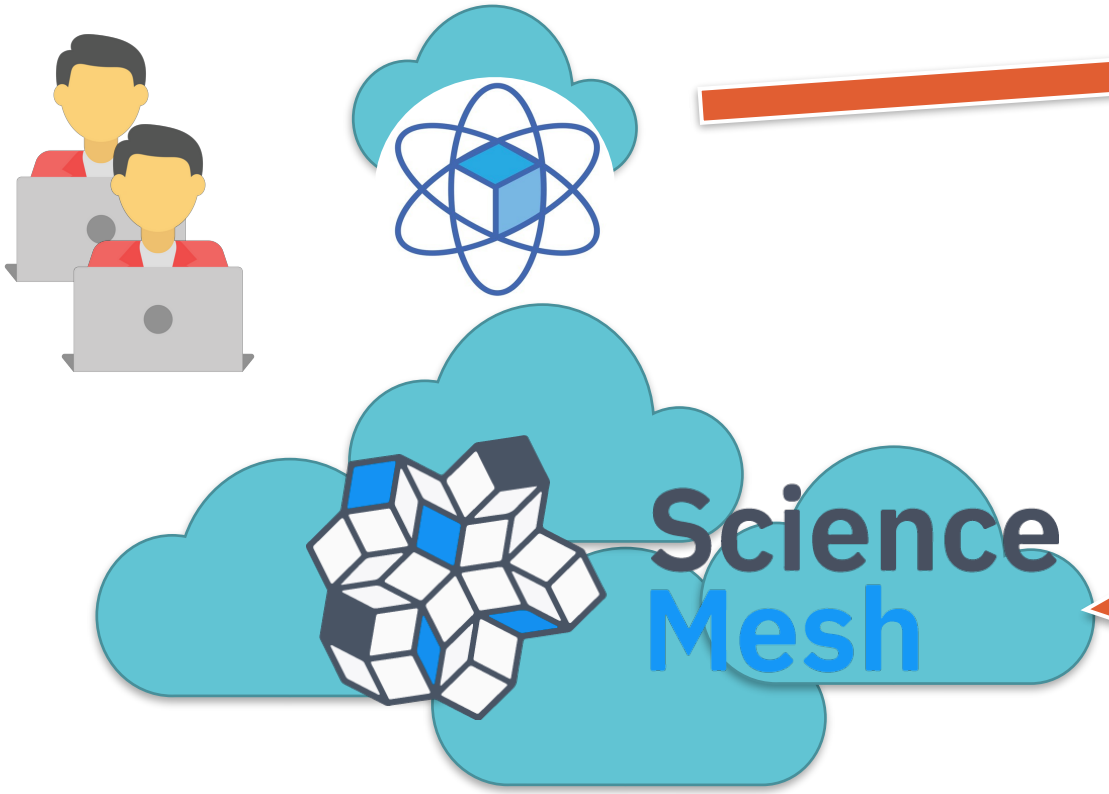


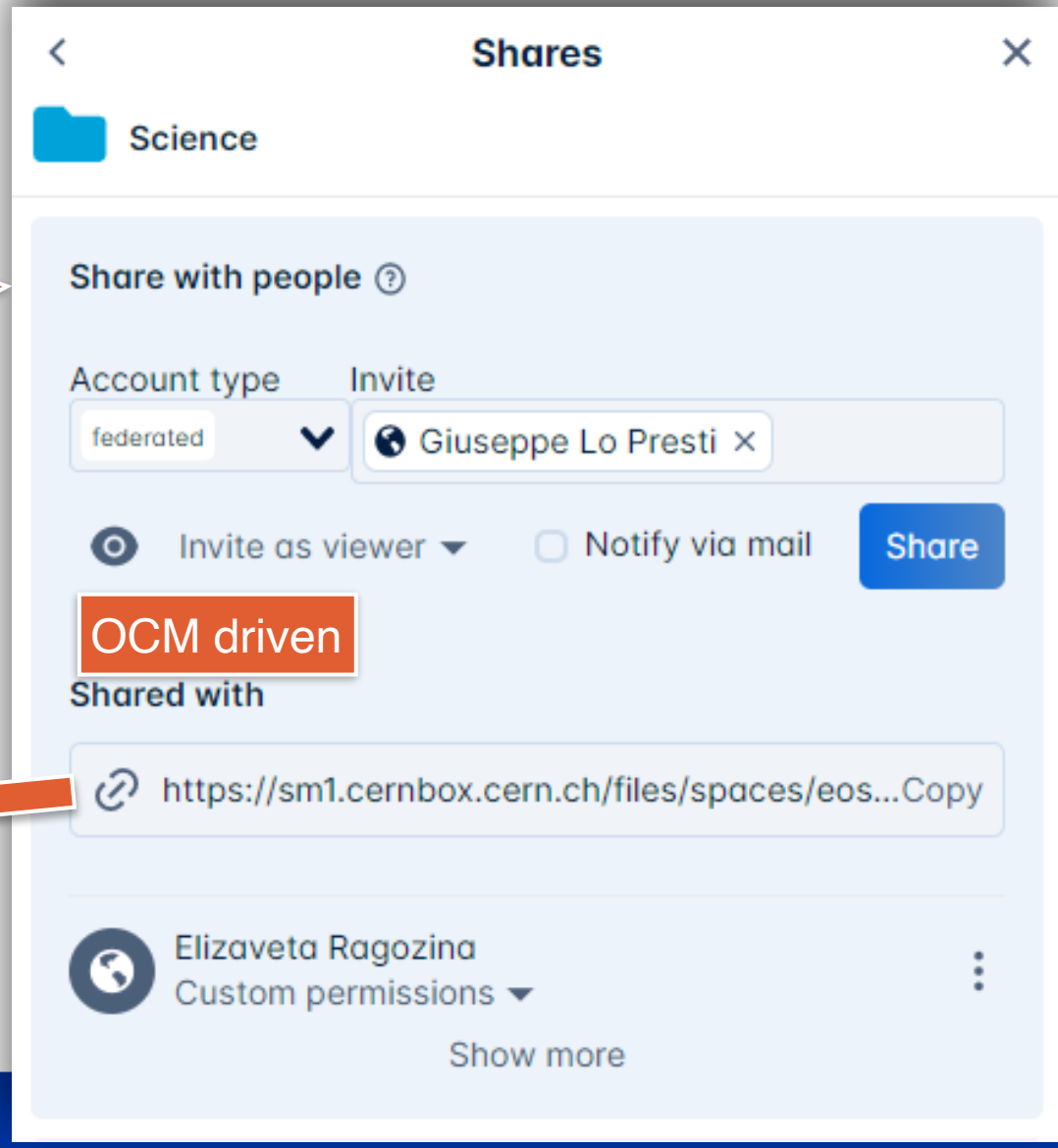
Supporting +31 Experiments

- 1st CERNBox User Forum
  - 193 registered users
    - 172 unique users in Zoom
    - Peak of ~90 concurrent users
  - 56+ institutions
  - 31 speakers
  - All CERN departments represented



D) Federated “dropbox” service for HEP community and beyond

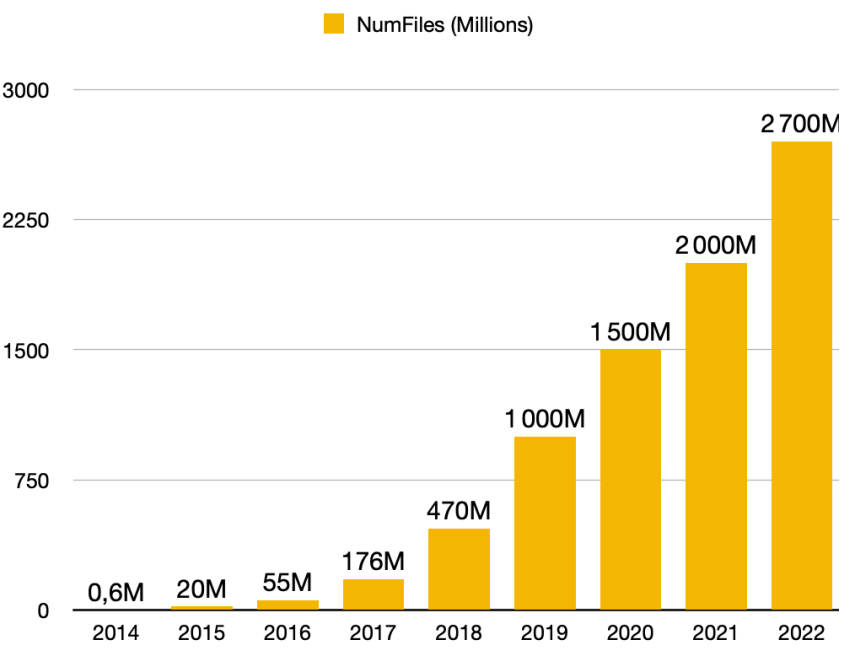


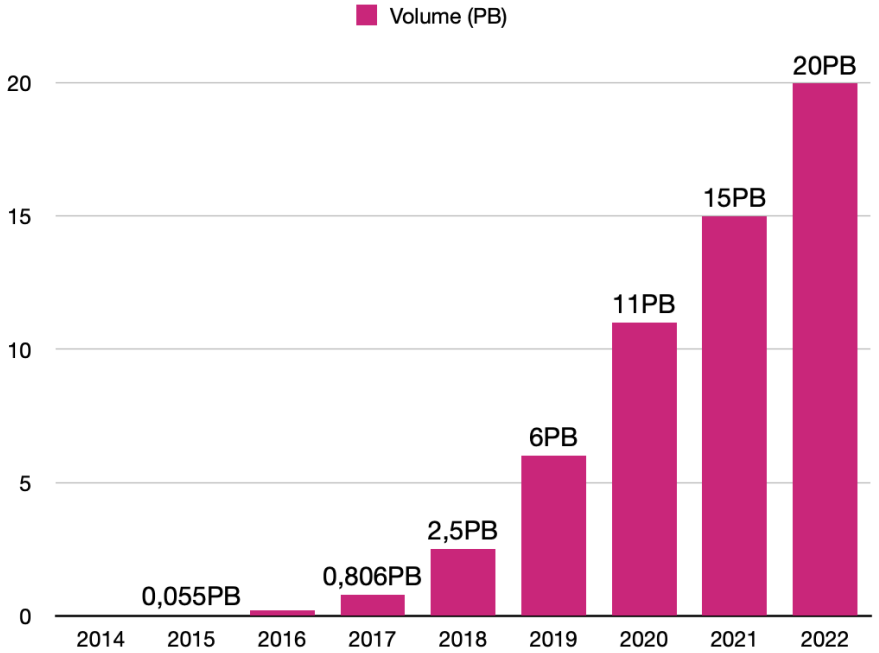


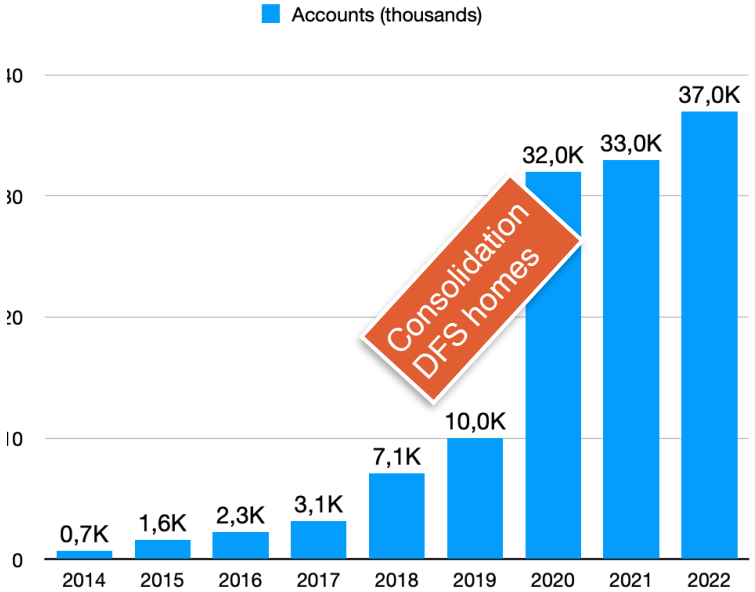
Early steps, huge potential ahead

Was it worth it?

Usage has steadily increased over the years







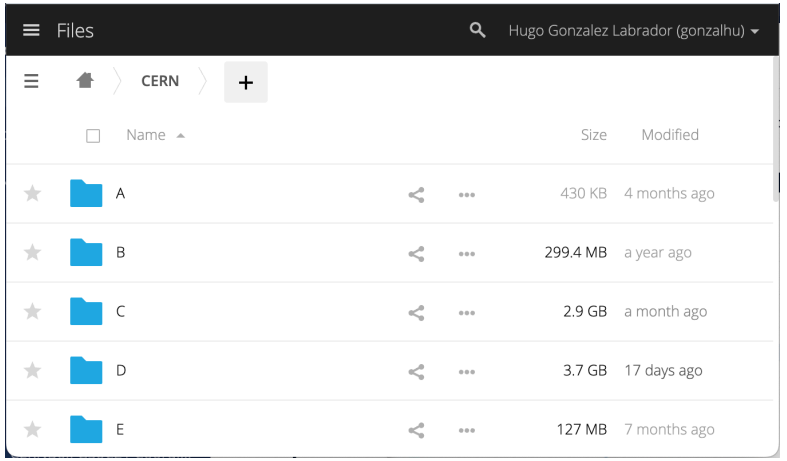
Total amount of files read	Total amount of bytes read	Total amount of files written	Total amount of bytes written
3.96 Bil	222 PB	2.25 Bil	8.85 PB

2022 only

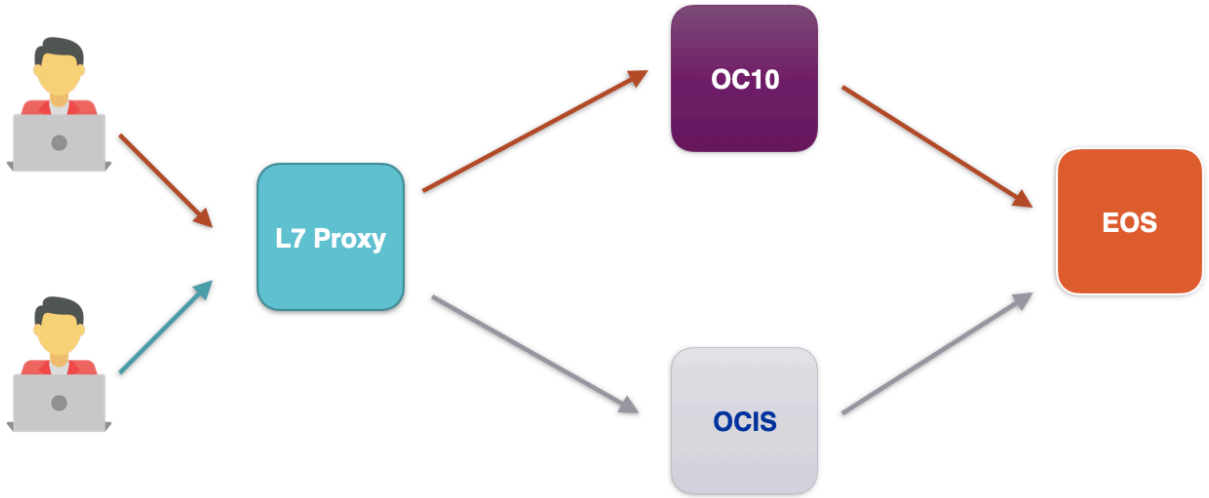
Biggest achievement in 2022?

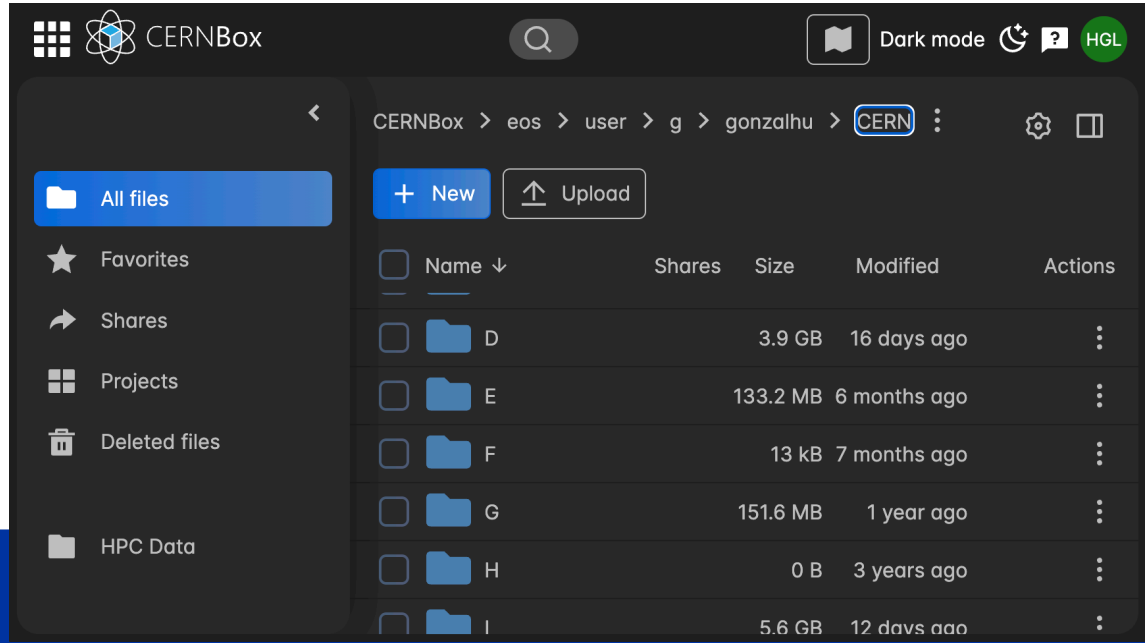
We moved from OC10 to OCIS platform

Transparent migration of 37K user accounts



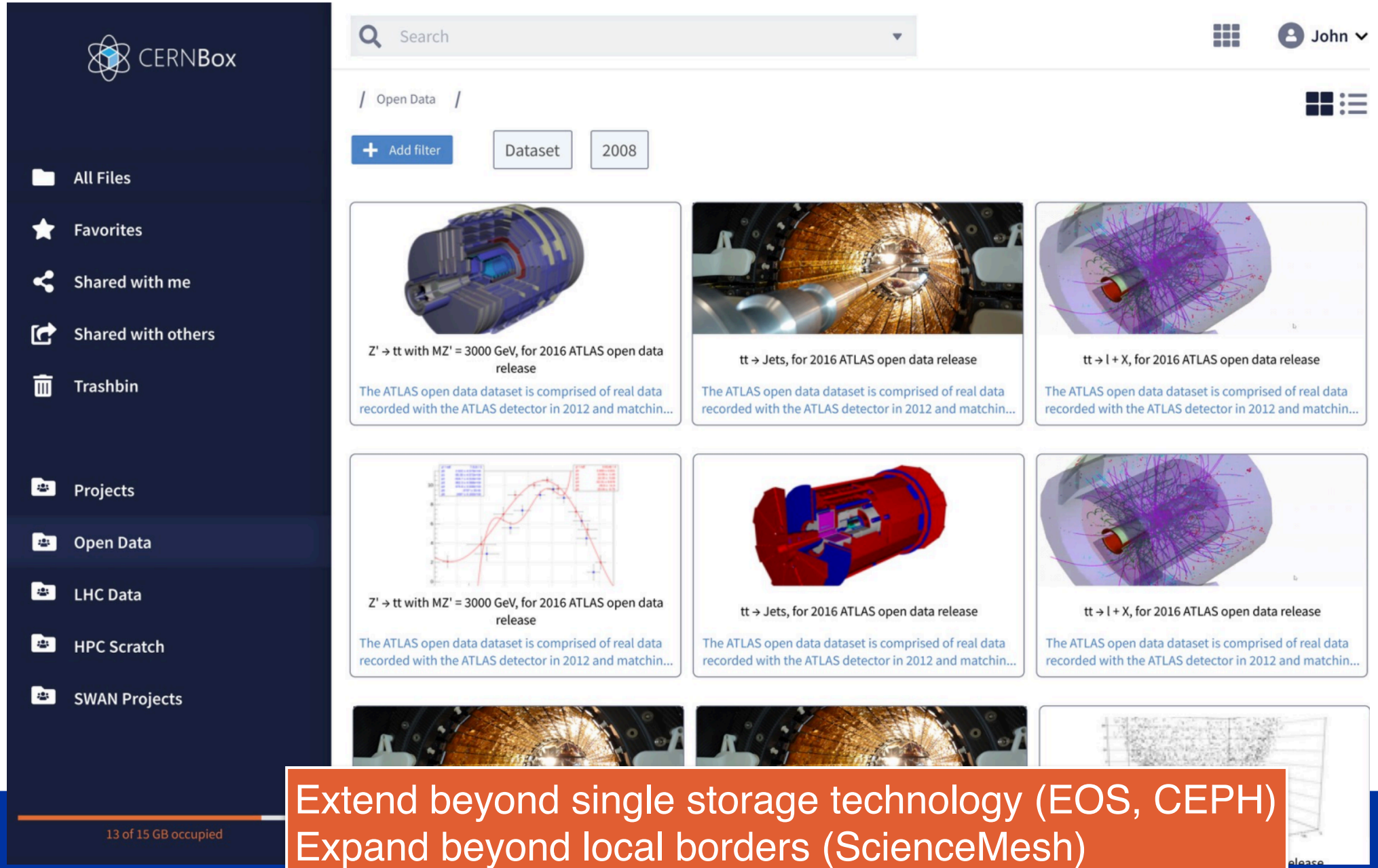
We had to ensure backwards and forwards compatibility of both systems





Next 10 years?

Vision: CERNBox as gateway for scientific environments



Extend beyond single storage technology (EOS, CEPH)
Expand beyond local borders (ScienceMesh)

Evolving SWAN towards an Analysis Facility system, Diogo, Monday @ 11

User-friendly OCM Invitation and Sharing in CERNBox, Elizaveta, Tuesday @ 15:20

Applications integration beyond local clouds with OCM, Giuseppe, Tuesday @ 16:30

C(ERN) BACK(UP), Gianmaria, Wed @ 9:45

Driving the ScienceBox package into the future, Samuel, Wed @ 11:45

Modern notification system for Sync and Share using NATS, Javier, Wed @ 12:15

home.cern