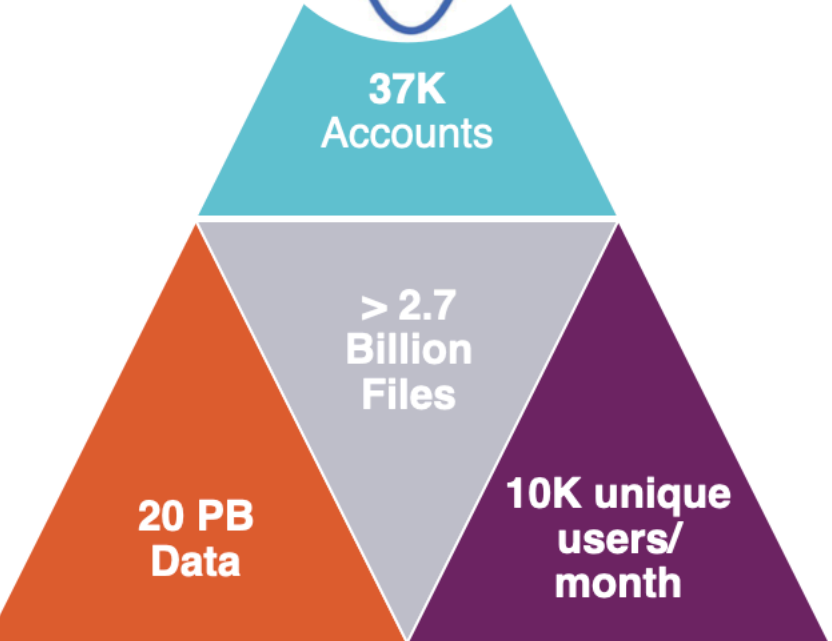
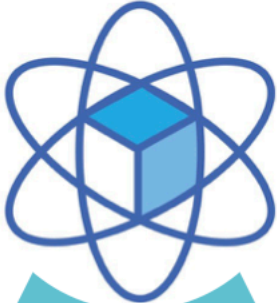


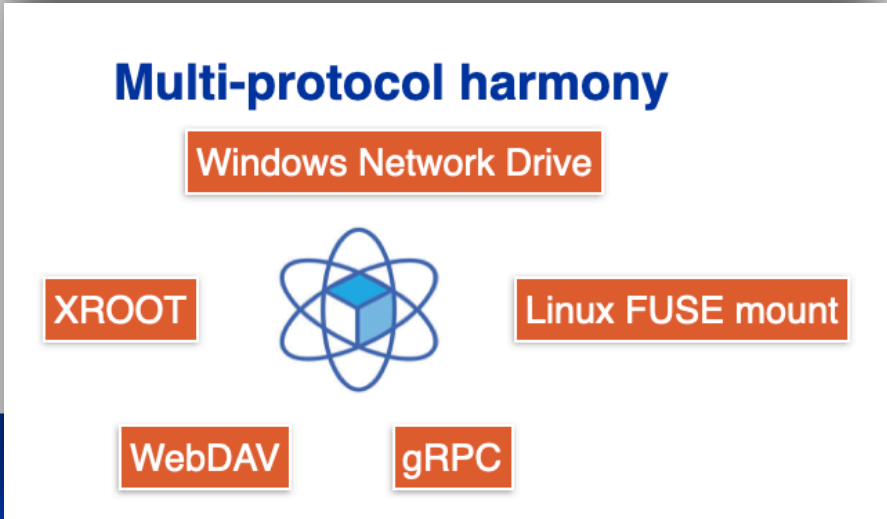
Standing innovation for 10 years
CERNBox
(2013-2023)

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



The tech stack



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)

CERNBox > eos > user > g > gonzahu

Buttons: + New, Upload

Name	Shares	Size	Modified ↑	Actions
CERN		24.2 GB	27 seconds ago	
backlog		13 kB	53 minutes ago	
EGYPT		1.3 MB	14 hours ago	
PRIVATE		8.5 GB	21 hours ago	
ana.eos		79.9 GB	1 day ago	
test.drawio		818 B	9 days ago	
tset.pptx		30 kB	14 days ago	
touch.txt		4 kB	23 days ago	
timeline.txt		433 B	23 days ago	
Monday		739.3 MB	25 days ago	
TuesdayFake		3 MB	25 days ago	
MondayFake		739.3 MB	25 days ago	
Motivation_H... .pdf		60 kB	28 days ago	
Testr		0 B	28 days ago	
DEBUG		1.3 GB	1 month ago	

Windows (v.2.9.2 with VFS) Installation guide

Mac OS Installation guide

Linux Installation guide

Android

IOS

CERNBox

Hugo Gonzalez Labrador (gonzahu) cernbox.cern.ch

Activity Settings

Quit CERNBox

Connected to <https://cernbox.cern.ch/cernbox/desktop/> as *Hugo Gonzalez Labrador (gonzahu)*. Account

105 GB of 929 GB in use

home

Checking for changes in remote 'CERN/C/CS3MESH/LOGO'

Add Folder Sync Connection

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](#).

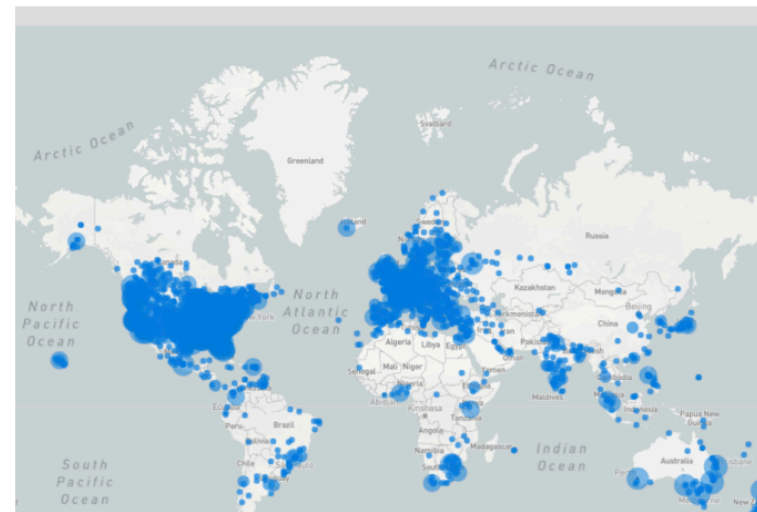
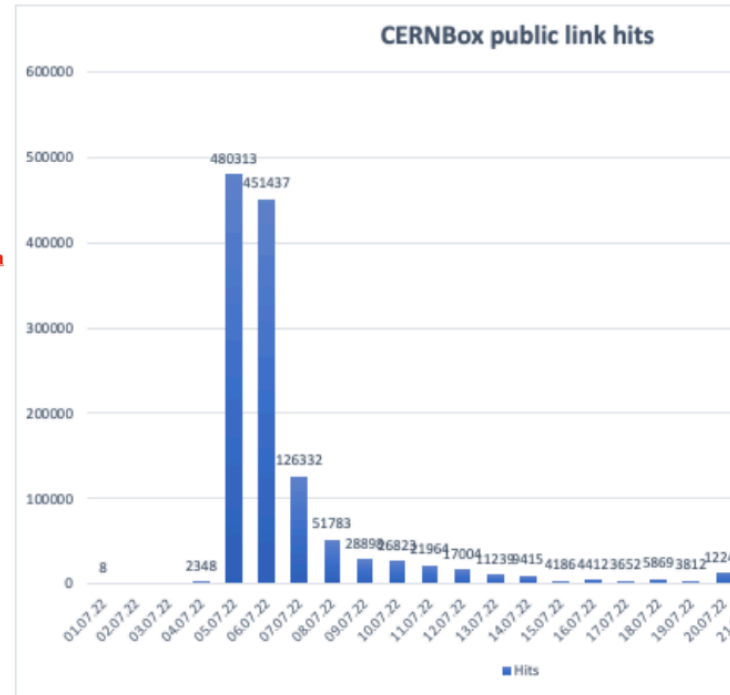


<https://cernbox.cern.ch/index.php/s/EacPckkCMFcJ8ya>

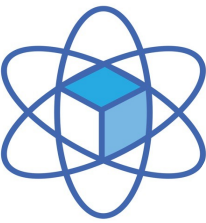


Name ▾

- 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



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



B) Unified platform integrated with physics storage

Universal access
Real time collaboration



Offline access
(mobile, desktop, tablet)



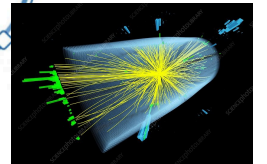
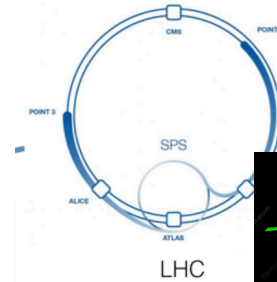
Windows online filesystem



POSIX online filesystem



Physics analysis



C) Support scientific workflows (SWAN and CERNBox)

Sync & Share storage integration tier – examples

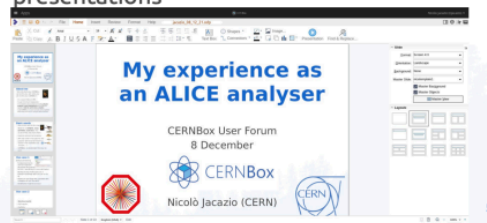
Use case 1

Backup with symbolic link

- This presentation is on CERNBox

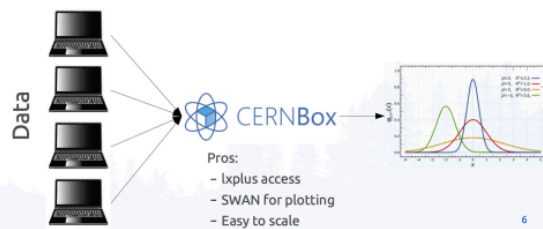
`/home/Documents/Presentations → cernbox/Presentations`

- Editing online, great asset for shared presentations



Use case 2

- Sharing common results among different analysers, results stored on CERNBox shared directory
- Split samples among different analysers



Use case 3

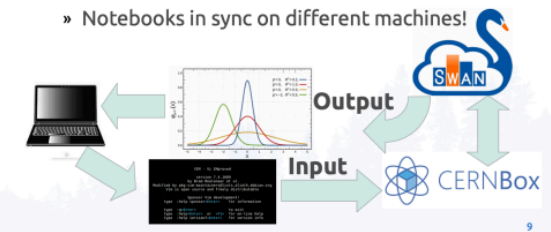
Database for configuration files, pros:

- Reliable storage
- Easy to give access to new users
- Access via lxplus
- Access via web



CERNBox+SWAN work for me

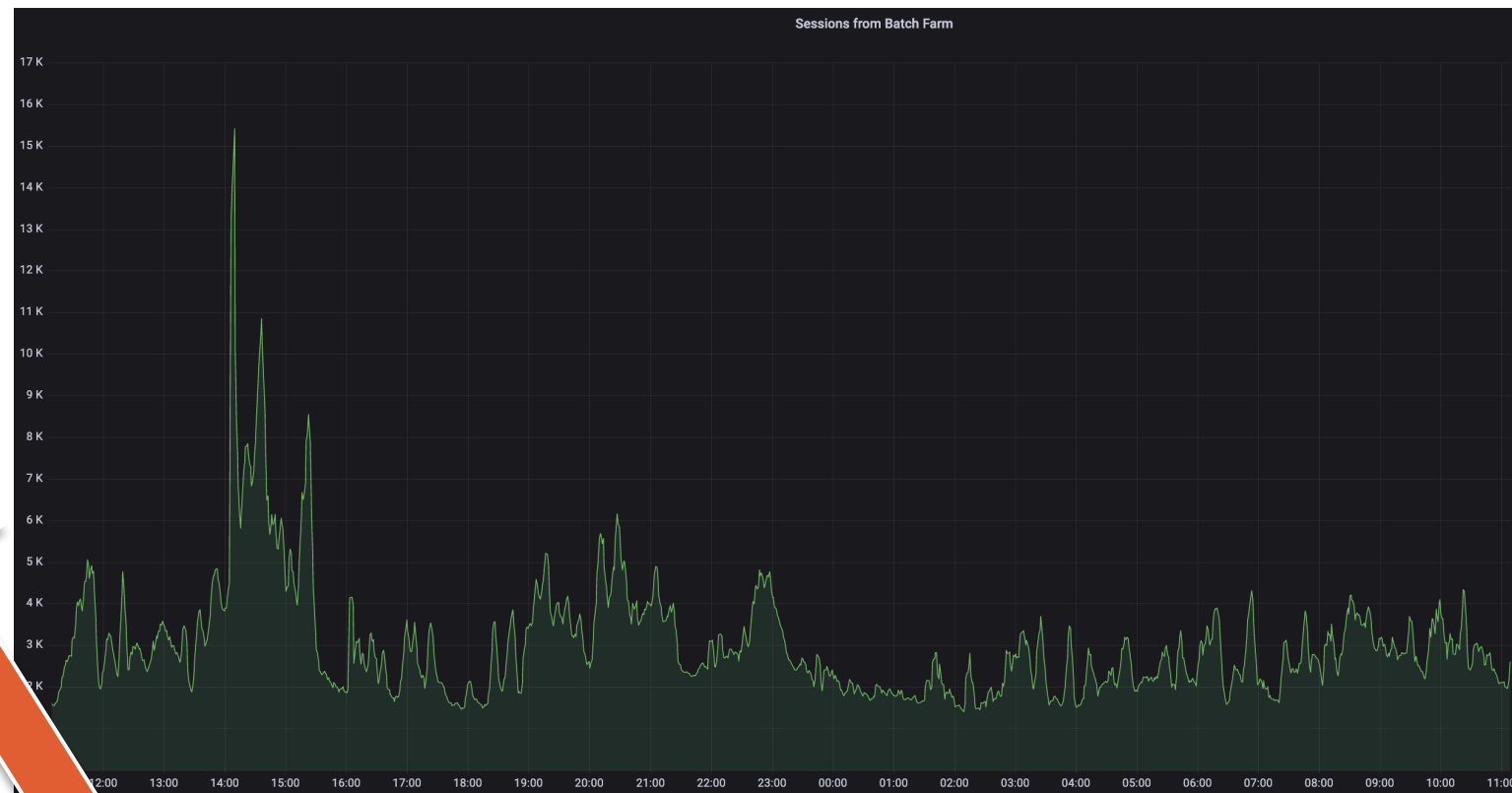
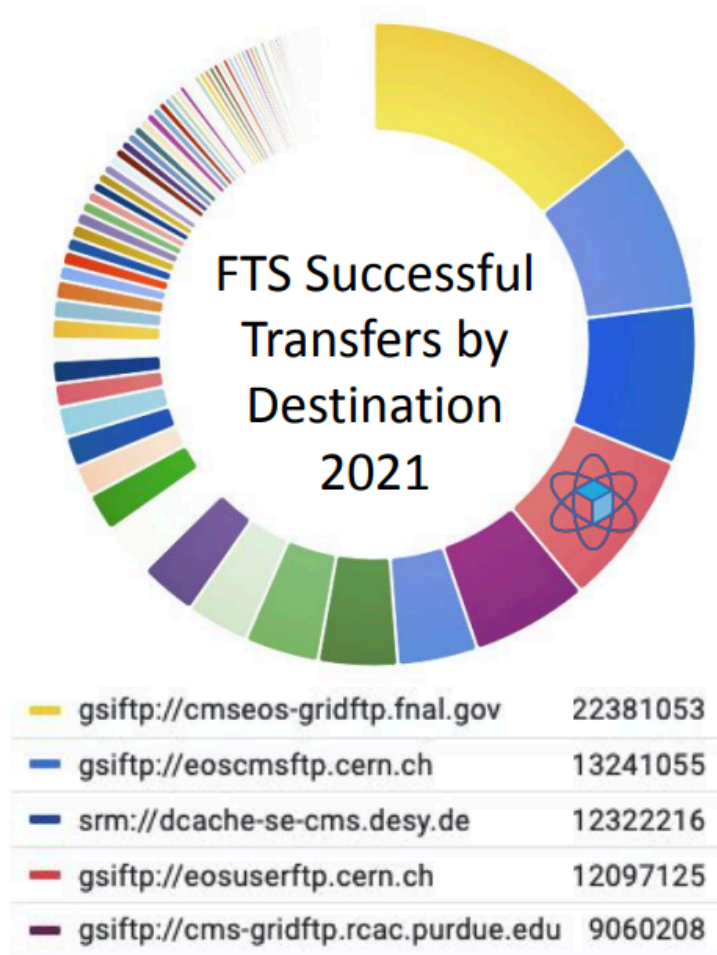
- Usage with **SWAN and EOS** is great
 - » Develop locally, run your code **both on-line and offline**
 - » Output is always at hand
 - » Notebooks in sync on different machines!



[Full presentation](#)

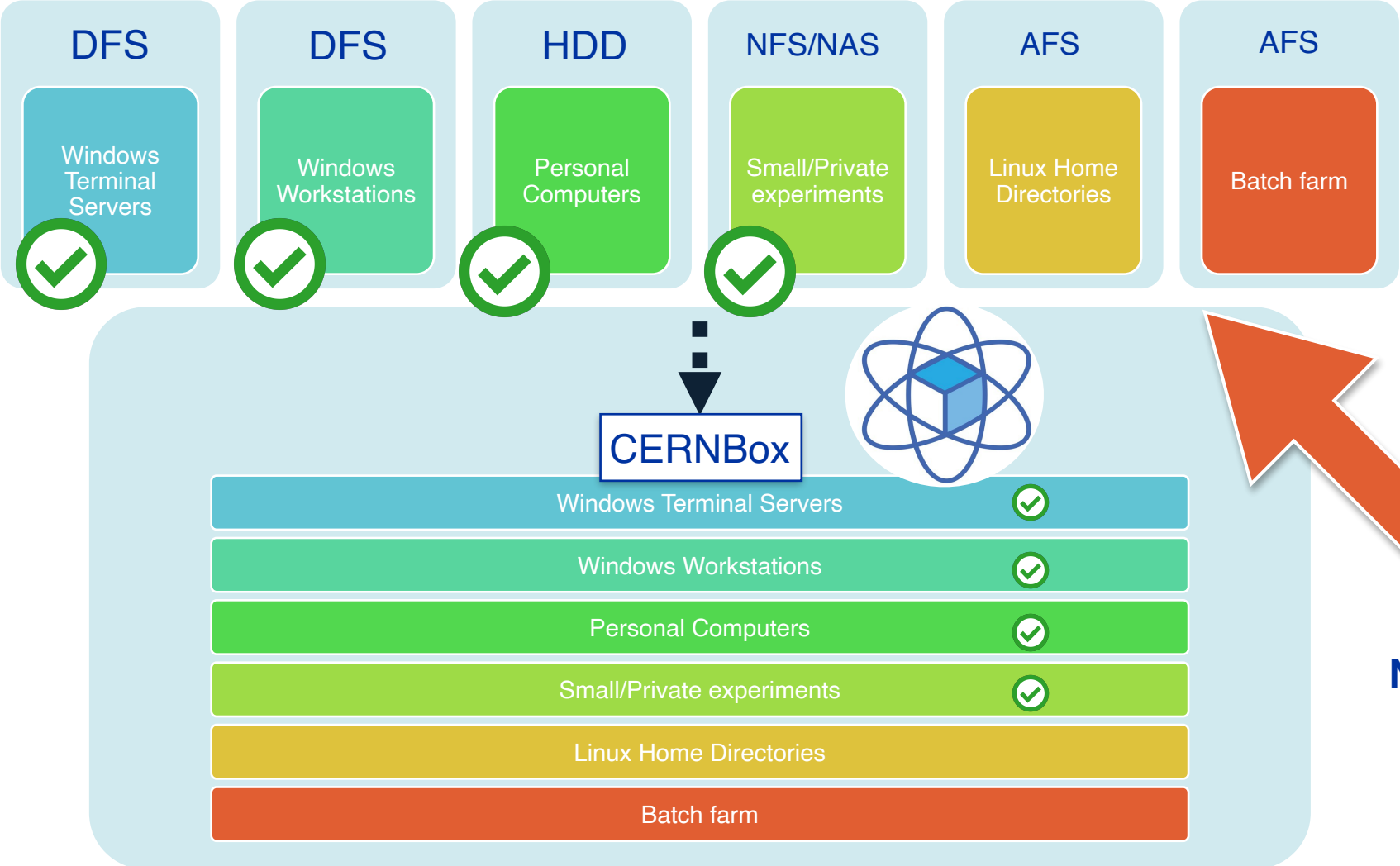
My experience as a ALICE analyser, Nicolo Jacazio, CERNBox User Forum, 2021

C) Support scientific workflows (Batch computing)

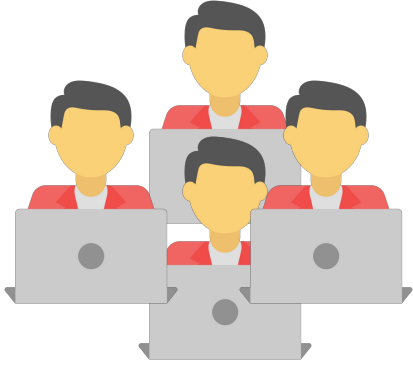


A CMS perspective about CERNBox, Danilo Piparo, CERNBox User Forum, 2021

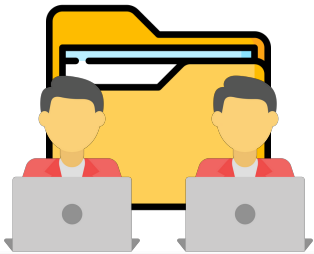
D) Delivering Home Directories (consolidation)



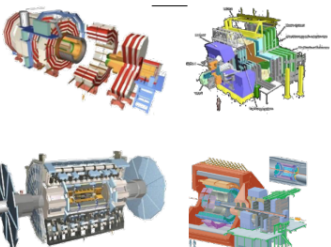
D) Delivering Home Directories (and beyond)



37K user home directories



+1K Project Areas



Supporting +31 Experiments

■ 1st CERNBox User Forum

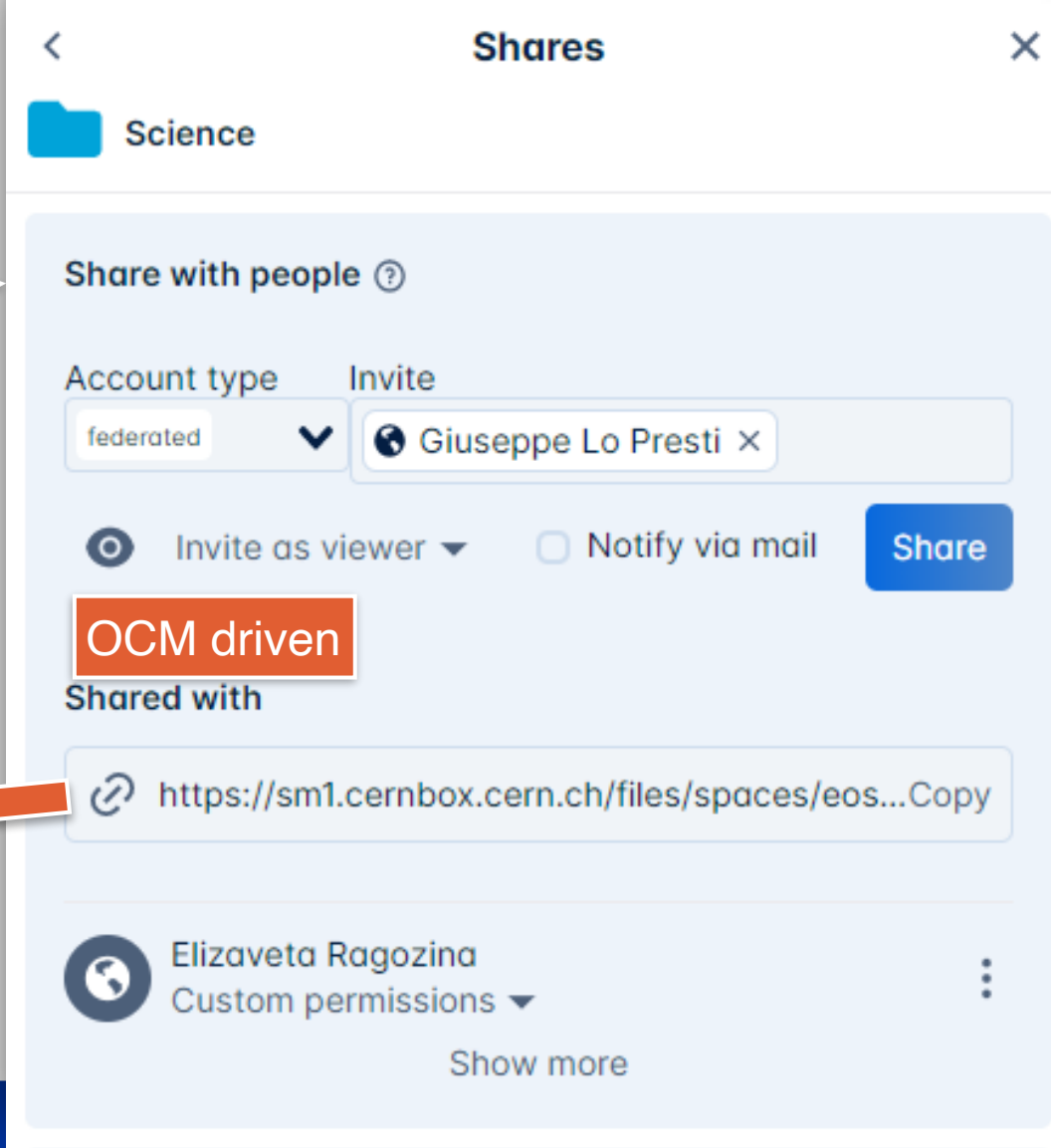
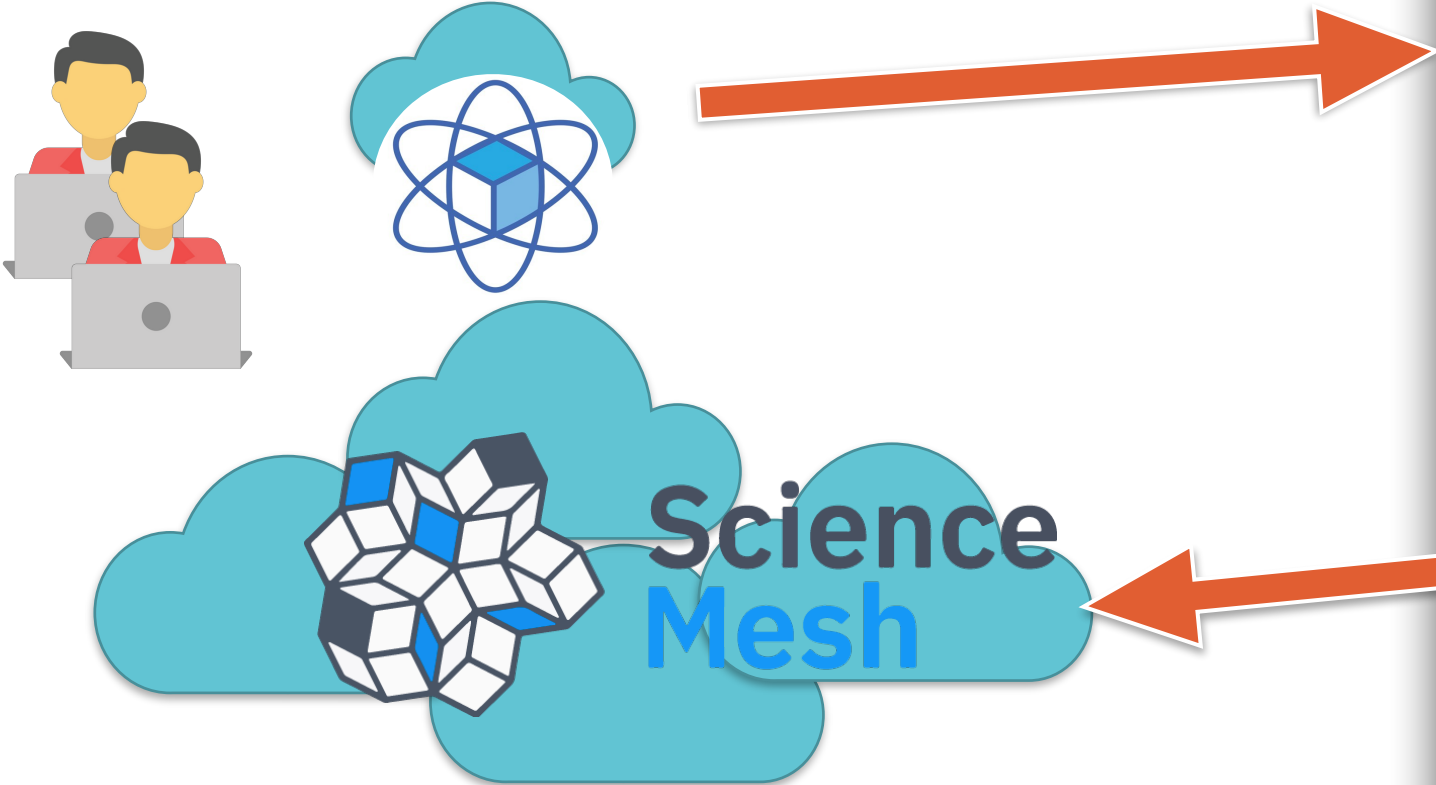
- 193 registered users
 - 172 unique users in Zoom
 - Peak of ~90 concurrent users
- 56+ institutions
- 31 speakers
- All CERN departments represented

CERNBox: User Stories, Proposed Features and Opportunities for Improvements

Piparo (CERN) - December 8, 2021

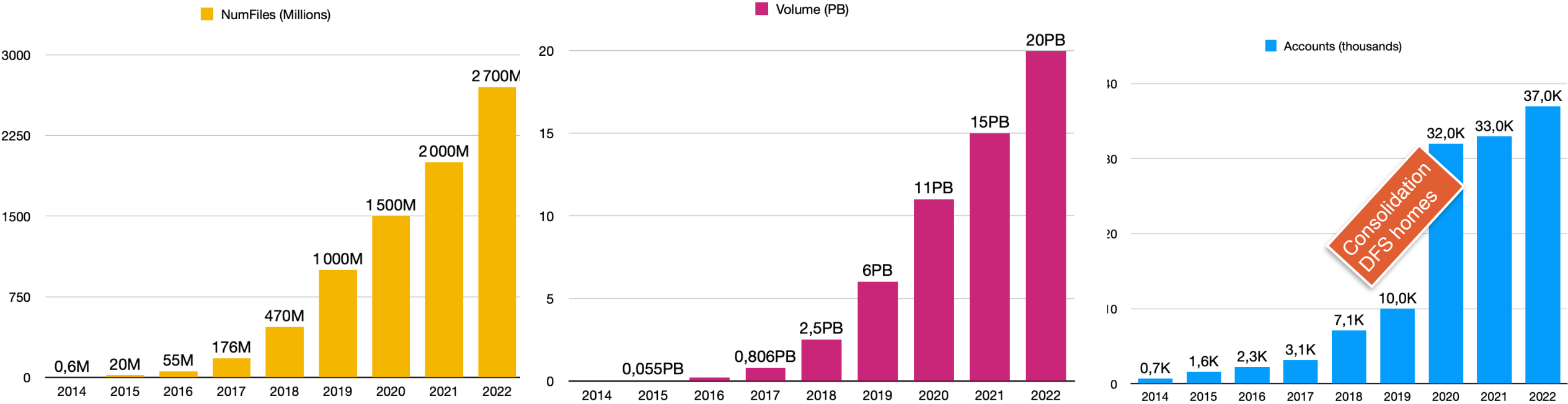


D) Federated “dropbox” service for HEP community and beyond



Was it worth it?

Usage has steadily increased over the years



Total amount of files read

3.96 Bil

Total amount of bytes read

222 PB

Total amount of files written

2.25 Bil

Total amount of bytes written

8.85 PB

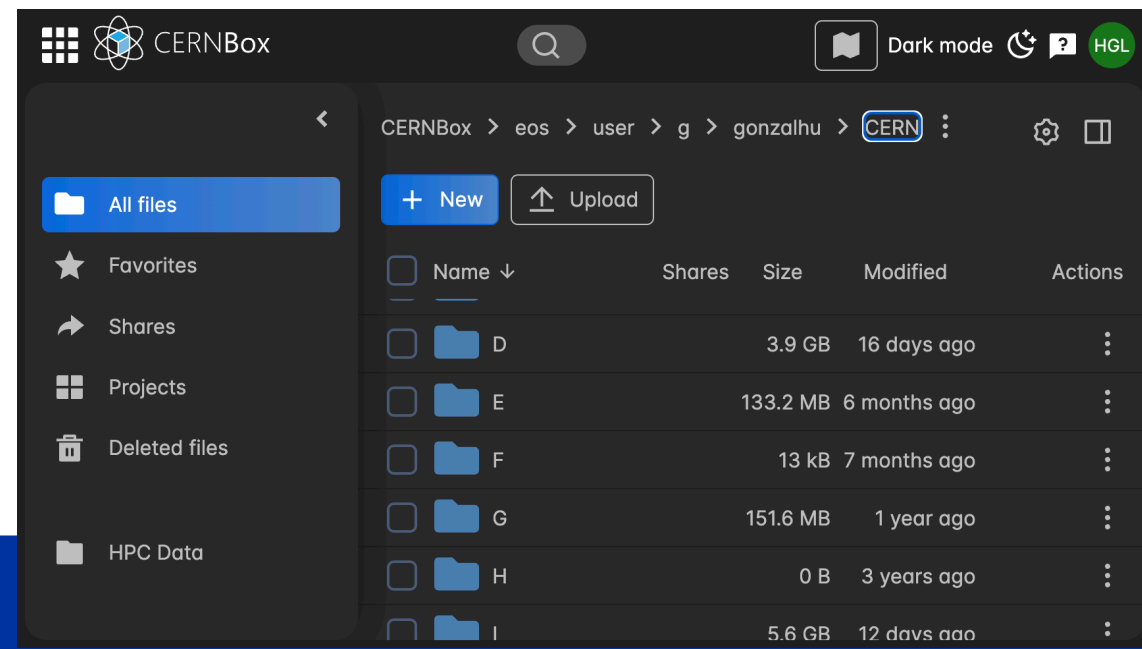
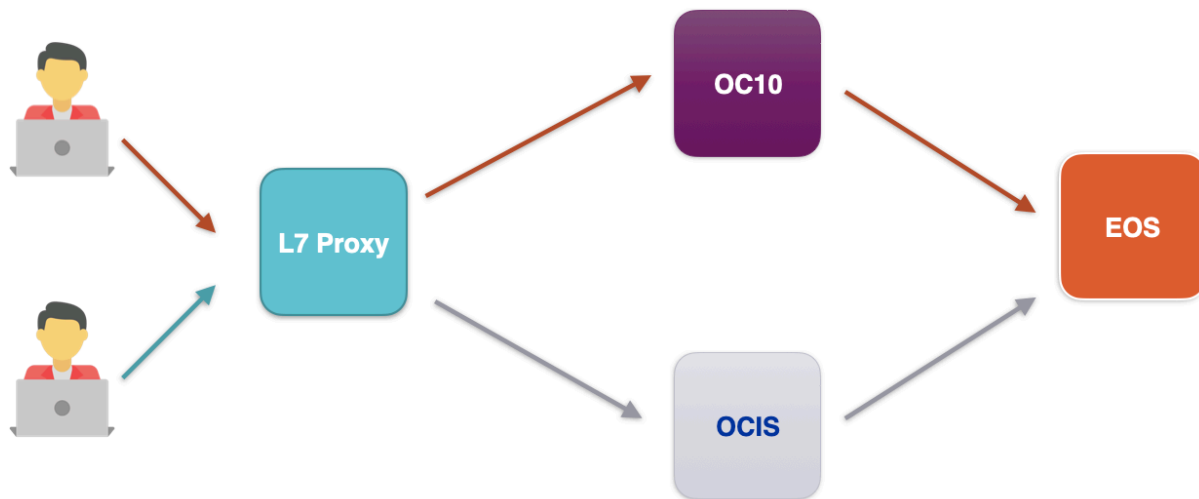
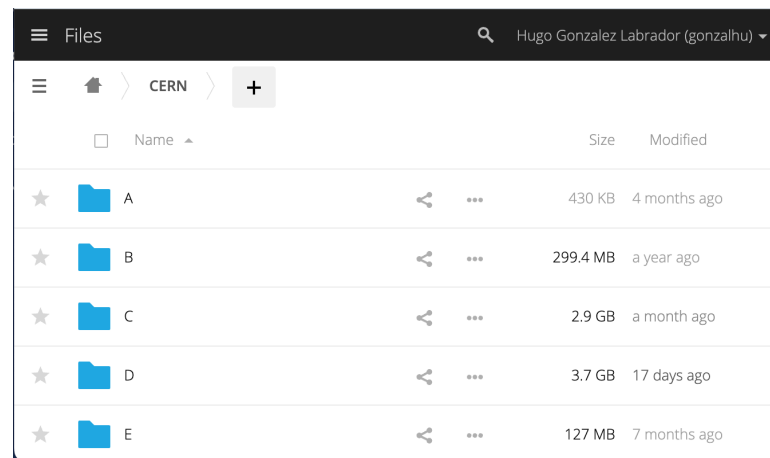
2022 only

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

The screenshot displays the CERNBox web interface. On the left is a dark sidebar with navigation options: All Files, Favorites, Shared with me, Shared with others, Trashbin, Projects, Open Data, LHC Data, HPC Scratch, and SWAN Projects. At the bottom of the sidebar, it shows '13 of 15 GB occupied'. The main content area features a search bar, a user profile for 'John', and a breadcrumb path '/ Open Data /'. Below this are filter buttons for '+ Add filter', 'Dataset', and '2008'. The main area contains a grid of dataset cards. Each card includes a thumbnail image (e.g., ATLAS detector cutaway, particle tracks, or a plot), a title, and a description. The titles for the visible cards are: 'Z' → tt with MZ' = 3000 GeV, for 2016 ATLAS open data release', 'tt → Jets, for 2016 ATLAS open data release', and 'tt → l + X, for 2016 ATLAS open data release'. The descriptions for all cards are: 'The ATLAS open data dataset is comprised of real data recorded with the ATLAS detector in 2012 and matchin...'. At the bottom, an orange banner contains the text: 'Extend beyond single storage technology (EOS, CEPH) Expand beyond local borders (ScienceMesh)'. The CERN logo is in the bottom left corner, and the page number '19' is in the bottom right corner.

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