CERNBox
The CERN Cloud Storage Hub

Hugo González Labrador – Storage Group
On behalf of the CERNBox Team
Description and Architecture
What is CERNBox?

CERNBox provides a cloud synchronization and sharing service

- Available for all CERN Users (1TiB/user)
- Synchronize files (data at CERN) and offline data access
- Easy and convenient way to share data across users and groups
- All major platforms supported
- Based on **ownCloud, powered by EOS**
- Integrated with **SWAN**
What is CERNBox?
What is CERN EOS?

Providing large-scale disk storage at CERN
14:45 - 15:00
https://indico.cern.ch/event/587955/contributions/2936837/

Scaling the EOS namespace
15:00 - 15:15
https://indico.cern.ch/event/587955/contributions/2936873/
CERNBox Architecture

- EOS clients
- XROOTD clients
- SAMBA Gateway
- Synchronization clients
- Browser access
- Multiple micro-services
- ROOT (Data Analysis Framework)
- Office 365
User Community and Service Numbers
## Service Numbers

<table>
<thead>
<tr>
<th></th>
<th>Jan 2016</th>
<th>Jan 2017</th>
<th>Jan 2018</th>
<th>Jul 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Users</strong></td>
<td>4074</td>
<td>8411</td>
<td>12686</td>
<td>14805</td>
</tr>
<tr>
<td><strong># files</strong></td>
<td>55 Million</td>
<td>176 Million</td>
<td>470 Million</td>
<td>604 Million</td>
</tr>
<tr>
<td><strong># dirs</strong></td>
<td>7.2 Million</td>
<td>19 Million</td>
<td>34 Million</td>
<td>41 Million</td>
</tr>
<tr>
<td><strong>Used Raw Space</strong></td>
<td>208 TB</td>
<td>806 Petabytes</td>
<td>2.5 Petabytes</td>
<td>3.47 Petabytes</td>
</tr>
<tr>
<td><strong>Deployed Raw Space</strong></td>
<td>1.3 PB</td>
<td>3.2 Petabytes</td>
<td>5.7 Petabytes</td>
<td>5.9 Petabytes</td>
</tr>
</tbody>
</table>

### Diagram

- **Physicists**
- **Engineers**
- **Services and Administration**
Collaboration

Over 80K shares
New Features and On-going Work
New UI for CERNBox

Petabyte-Scale Cloud Storage File Manager
Collaborative platform

The CERNBox apps ecosystem

- From High Energy Physics...
- Started as an on-premise cloud storage
- Move and more integrating new ways to interact with your data
- Reaching out to Scientific and Cloud Computing
- Microsoft Office integration
- Galleries (pictures as well as ROOT histograms)
- Notebooks (Jupyter)
- more to come
New Unified Storage Vision

AFS
Batch System
Local disk
Personal computers

DFS
Windows Terminal Servers

NFS
Small/Private Experiments
New Unified Storage Vision

- AFS
  - Batch System
- DFS
  - Windows Terminal Servers
- Local disk
  - Personal computers
- NFS
  - Small/Private Experiments

CERNBox

- CERNBox
- Batch System
- Windows Terminal Servers
- Personal computers
- Small/Private Experiments
From EOSUSER to EOSHOME

EOS clients

XROOTD clients

Synchronization clients

SAMBA Gateway

Browser access

EOSUSER

MGM

NS
From EOSUSER to EOSHOME

EOS clients

XROOTD clients

Synchronization clients

SAMBA Gateway

Browser access

Data scale-out

Namespace scale-out
From EOSUSER to EOSHOME

EOS clients

XROOTD clients

Synchronization clients

SAMBA Gateway

Browser access

EOSHOME-A

EOSHOME-B

EOSHOME-C

EOSHOME-Z

EOSHOME REDIRECTOR

CERN

Divide et Impera
From EOSUSER to EOSHOME

EOS clients
XROOTD clients
Synchronization clients
SAMBA Gateway
Browser access

Data scale-out
Namespace scale-out

EOSHOME-A
EOSHOME-B
EOSHOME
Conclusions

• CERNBox plays a key role into the future home directory of CERN.

• From year to year the use of the system is increasing very fast.

• Users demand more and more features and integrations into their workflows.

• We have a nice collaboration with industry to shape the future of sync and share for HEP uses cases.