Hugo Gonzalez Labrador (CERNBox Project Leader)
Elvin Alin Sindrilaru (EOS core developer)
25th March, ICRC Workshop, CERN
Where to store CERN’s User and Scientific data at scale?
CERNBox: Scientific Cloud Storage

- Provides a simple, uniform and turnkey way to access research, administrative and engineering data since 2014.
- Built purely with **open source** components
  - **ownCloud** for the “Dropbox” use-case
  - **EOS** for storage
- Data is **safely stored in CERN campus**, under CERN custody with clear security policies to keep data safe and confidential.
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
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
- Windows Network Drive
- XROOT
- Linux FUSE mount
- WebDAV
- gRPC

37K Accounts

> 2.7 Billion Files

20 PB Data

10K unique users/month
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, 2023

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, opening 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 a large-scale and maintenance work, the LHC is now set to run for close to four years at the record energy of 13.6 TeV (8TeV), providing greater precision and discovery potential. Increased collision rates, higher collision energy, upgraded detector readout and selection systems, new detector systems and computing infrastructure will all play a part in a promising physics season that will further expand the already very diverse LHC physics programme.

Pictures of the day are available [here](#).

Videos of the event are accessible [here](#).
Consolidation of multiple storage offerings

Windows Terminal Servers
Windows Workstations
Personal Computers
Small/Private experiments
Linux Home Directories
Batch farm

CERNBox

Next target: AFS
Sync & Share storage integration tier — examples

**Use case 1**
Backup with symbolic link
- This presentation is on CERNBox
- 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
Support for different communities

- 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
Data safety and confidentiality

Backup

• Based on restic (open source)

• Data stored in an other geographical location (encrypted) in the CERN campus (3km away) in a different technology (CEPH S3)

• WIP: copy to tape (CERN Tape Archive, open source)

GDPR-like (OC11) compliant

• Simple but powerful access control

  • Closed by default, up to the user to open up

  • By sharing data with individuals and groups

  • Making data available publicly
Consolidation
DFS home
Evolution in usage

NumFiles (Millions)
Volume (PB)
Accounts (thousands)
Bits and pieces (abstract)
Bits and pieces (concrete)

Universal access
Real time collaboration

Offline access
(mobile, desktop, tablet)

POSIX online filesystem
SAMBA
FUSEX
xroot

Physics analysis

Windows online filesystem

Web Apps
Sync & Share
Reva

EOS for Home and Project Spaces (9 instances)
**What is EOS?**

**Open-Source** Storage platform designed and developed by CERN IT since 2019

**Distributed storage system**, providing filesystem access and beyond

Elastic, Adaptable and Scalable

Software solution for data recording, user analysis and data processing

Supports thousands of parallel clients

Multi-protocol support (FUSE, HTTPS, WebDAV, XROOT)

Offers a variety of authentication methods (KRB5, X509, SharedSecret, tokens, unix)

---

**Main Data Access Patterns in Physics**

**Data Analysis**
- >100k relatively slow streams reading data (almost) sequentially from 60k HDDs
  - 1-100 MB/s - sometimes forward-seeking
  - “similar to 100k people watching an individual film on Netflix”

**Data Acquisition / Data Taking**
- hundreds of streams possibly as fast as possible
  - 50-250 MB/s with File Replication
  - 400 MB/s-1 GB/s with Erasure Coding
EOS @ CERN

**EOS Physics** for experiment data

**CERNBox** for end-user data and sync&share

- **Total**
  - Files: ~8 Bil
  - # Storage Nodes: ~1300
  - # Disks: ~60000

- **2022 stats**
  - Total amount of files read: 21.8 Bil
  - Total amount of bytes read: 4.08 EB
  - Total amount of files written: 5.73 Bil
  - Total amount of bytes written: 631 PB

Focus on this presentation
EOS architecture

High-available and low latency namespace
- namespace persisted in a distributed key-value store
  - QuarkDB (Raft consensus + Redis API)
- working entries cached in-memory

High available and reliable file storage, based on (cheap) JBODs:
- File replication across independent nodes and disks
- Erasure coding to optimize costs and data durability

MGM: meta data server
MQ: message queue
NS: persistent namespace
FST: file storage server
EOS: beyond CERN

JRC Earth Observation Data and Processing Platform (JEODPP)

**Versatile** platform bringing the users to the data

Examples of exploratory analyses and interactive visualization

- Impact of forest fires on Natura 000 sites
- Deforestation (time lapse)
- CO emissions (time lapse 08/18 vs 08/19)
- Ship traffic (heat maps)
- Deforestation (split map 2000 vs 2017)
- Interactive combination of information layers

Armin Burger,
EOS as storage backend for JRC
What is ownCloud?

- Open-Source Cloud Sync and Share Collaboration Platform
- Early users of it since 2013 (v6.0)
- Provides Web, Desktop and Mobile clients to underlying EOS storage
- Traditional deployment on LAMP stack
Data ownership

Data is not owned by system user (www-data/apache) but owned by end-users.
Data synchronisation

CERNBox does not sync data between MySQL and the storage. The storage is the source of truth
Sharing semantics

ACL information is propagated to the storage
Future CERNBox
Future of CERNBox

- Integrate with other storages, not only EOS.
- POC with CEPHFS
FOSS Opportunities
• Thanks to open source, we created a strong partnership to evolve the platform to satisfy scientific use-cases

• Reva software initially developed at CERN: github.com/cs3org/ lava

• Reva adopted by ownCloud in 2021 as the foundation for their new product: ownCloud Infinite Scale

An important milestone

"We are thrilled to see that CERN’s developed technology, Reva software, has been successfully integrated into the ownCloud’s new product, highlighting CERN's commitment to building further links with industry in terms of knowledge transfer.” - says Hugo Gonzalez Labrador, the Project Leader for CERNBox in CERN’s IT department
Funding opportunities

- CS3 community represents “Sync and Share” solutions in EU and beyond
  - Industry members
  - Research and Education institutions

10M Euros

No borders to Europe’s research environment

Find out what CS3MESH4EOSC is doing for researchers, developers, companies, policy makers and citizens

- Developing an interoperable federation of data, to easily sync, share and deploy applications and software
- Making friction-free collaboration in EUROPE a reality
- Producing high-level services for the EOSC community

Looking forward to meet you in the next CS3 Conference
Trust and neutrality

• Possibility to **audit** the code at any time
• Possibility to **adapt** to our needs at any time (respecting license conditions)
• No **vendor locking** (APIs available)
  • For example, WebDAV
Quick Demo
Try our services with ScienceBox

https://sciencebox.web.cern.ch/

- **EOS**: Storage backbone for LHC + Physics data, and CERNBox
- **CERNBox**: Sync&Share for Personal and Project Files
- **SWAN**: Data Analysis Platform with Interactive Jupyter Notebooks
- **CVMFS**: Software stacks for LHC experiments and scientific analysis

Enrico Bocci, CS3 2022

No need of 10K people, just 1 finger
Thank you

More information?

hugo.gonzalez.labrador@cern.ch
elvin.alin.sindrilaru@cern.ch
Hardware configuration

- EOS
  - Management Node: 64 cores, 384 GB RAM
  - Disk servers: 64 cores, 128 GB Ram
    - Disks: 48 x [6,10] TB
- NGINX and Web Servers:
  - 64 GB RAM, 64 cores
- Network
  - 10 gigabit
<table>
<thead>
<tr>
<th>Community</th>
<th>Commit authors</th>
<th>Issue submitters</th>
<th>Review submitters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>466</td>
<td>87</td>
<td>138</td>
</tr>
<tr>
<td>Total (Last year)</td>
<td>124</td>
<td>36</td>
<td>65</td>
</tr>
</tbody>
</table>

https://cauldron.io/project/7418?from_date=2022-01-01&to_date=2023-03-24&tab=overview
Integrated with computing farms (Batch)

FTS Successful Transfers by Destination 2021

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