ScienceMesh: pan-European collaboration fabric for Science


8th May 2023, CHEP, Norfolk, USA

CS3MESH4EOSC - Interactive and agile/responsive sharing mesh of storage, data and applications for EOSC, has received funding from the European Union’s Horizon 2020 research and innovation programme under Grant Agreement no. 863353.
Outline

• Intro to (some) CERN Scientific services
• Outreach and Education opportunities
• Cross-institutional challenges
• ScienceMesh
• Summary
CERNBox

- CERNBox is CERN’s on-premise cloud collaboration platform.
- Deeply integrated with CERN scientific services portfolio
- Poster session tomorrow
CERNBox user community

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

+1000 Project areas

31 LHC and non-LHC experiments working areas

Skyrocketing usage

13k unique users/month

CERNBox user community
Some Outreach and Educational activities
Helping students to reach their educational goals

Advantages of using CERNBox in educational practice

CERNBox provides an easy way to keep files safe and in the cloud; it also makes it easy for teachers and students to share.

CERNBox provides a plugin for Moodle, allowing teachers to easily integrate it into a course.

Its integration with Moodle and SWAN makes CERNBox a suitable solution in a modern school.

Description

CERNBox is cloud-based storage, synchronization and sharing for science. Users can easily access their files via the Web interface, have them synchronized with the CERNBox desktop client, and share single documents or entire folders with colleagues. CERNBox is also accessible from mobile devices via the CERNBox app with which it is possible to automatically upload pics and videos shot with the on-board camera to the cloud.

More on “ScienceBox 2.0: Evolving the demonstrator package for CERN storage and analysis services” Tuesday @ 11:15, Enrico Bocchi
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 (CERN) preceded 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, following a multi-year upgrade to the large Research Infrastructures (LHCb, ATLAS, CMS, ALICE) and technological developments toward the next generation of LHC operation with higher intensity beams and increased energy.

Initial over three years of upgrade and maintenance work, the LHC is now set to run at the same rate as the record energy of 2.76 TeV in 2012, providing greater precision and discovery potential. An important part of the upgrade has been to increase the computing infrastructure. These efforts point to a promising physics agenda 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.

LHC Run 3
Media Coverage with CERNBox

https://cernbox.cern.ch/index.php/s/8ecPq6kCMF5aJ7yy

CERNBox public link hits

139 COUNTRIES
4830 LINKS

Worldwide
24701 IPs
United States
18323 IPs
Canada
845 IPs
United Kingdom
694 IPs
United Kingdom
541 IPs
Switzerland
542 IPs
France
564 IPs

[Graph showing public link hits]

ALICE_16h102m_002.jpg
ALICE_16h54m_013.jpg
ALICE_17h40m_029.jpg
ALICE_17h50m_031.jpg
ALICE_18h10m_036.jpg
ATLAS_15h47_015.jpg
ATLAS_16h47_038.jpg
ATLAS_16h47_062.jpg
ATLAS_16h48_043.jpg
ATLAS_16h48_049.jpg
CCC_15h48_090.jpg
CERNBox, GSOC, HSF and Open Source

- Sourcing talent while promoting science
SWAN, Education and Outreach

- Frequent requests to use SWAN during schools or workshops
  - Exercise sessions at the CERN School of Computing
  - Tutorials of analysis tools: ROOT workshop for summer students

- Requests that include GPUs for ML are becoming more common
  - “I need XX GPUs to use SWAN during a course about ML”
  - Three examples already in 2023: ATLAS ML workshop, ICSC and Italian Teacher Programme

Outlook

- CERN Open Data portal is constantly growing
  - Data access policies of the ALICE, ATLAS, CMS and LHCb ensure vast amount of new data

- Ongoing effort to publish legacy CMS data in NanoAOD format
  - Reduced data format detached from experiment specific software
  - Suits a wide range of analyses
  - Allows for analyses with simple programming model
    → Bringing students and individuals close to real physics data from the LHC with minimal technical know-how
    → Example: ROOT RDataFrame

- SWAN together with the Open Data portal would be the perfectly suited to bring HEP as close as possible to students and individuals

Outreach

Reach out with SWAN! This section collects a series of outreach efforts involving SWAN.

Particle open data teaching (Huikka/fysikan avion data opetuksessa)

SWAN, the CERN Service for Web based Analysis, is not only made for analysis of scientific data but also the ideal platform for outreach. Pasi Huikka (CMU) put together an introductory course about experimental HEP for future-high school teachers. The result is great: check it out in SWAN!

Esim. pseudoprodukti - millaiseen positiiviin

Lähde tämän esimerkiksi osoitteesta, missä pseudoprodukti voidaan matemaattisesti voidaan havaita CMS-tilastojen mukaan oikein. Kasittelut CMU:n vastaanottaa 2021, ja se katsoo vain joitain 10000 - 10 000 000 sanasta. Tässä esimerkissä on osoitettava positiivinen, jossa sallitaan 2 kohdetta ja riittää esim. joka on konstuluminen, esim. suurimmaksi.
European landscape

- CS3 Site Reports:
  - >30 PB of data
  - > 20 nodes
  - > 600,000 users
  - > 3.5 billion files/directories
  - > 1 million shares

- Great uptake of Cloud Sync and Share and JupyterNotebook

- Satisfying needs of local users
But ...

• Researchers remain isolated on data islands because these services aren’t interconnected
• No common, ratified API among NRENs
• Hard to share add-ons between NRENs
• Not benefiting from economy of scale
Also raised during the pre-CHEP HSF workshop

- Ability to collaborate in a multi-organizational team on a single resource
- Ability to efficiently access collaboration data as well as make intermediate data products available to the team

Abstract

This whitepaper presents the current status of the R&D for analysis facilities and attempts to summarize the current views on the future direction of these facilities. These views have been collected through the High Energy Physics (HEP) Software Foundation’s (HSF) Analysis Facilities forum, established in March 2022, and through the Analysis Ecosystems II workshop that took place in May 2022.

Introduction

In the HL-LHC era (to begin in 2029) LHC analysts will have an order of magnitude more data to analyze than currently. This will require significant changes to analyses techniques and workflows some of which may be better suited to run using cloud technologies and services.
There is an important demand for users to collaborate beyond their local clouds.
Science Mesh
One way to remove barriers
• ScienceMesh is a **trusted federation** of nodes

• Users can **collaborate with remote users** with the same UX as if they were local users

• Built with Open Standards and Open-Source Software

• Allows **interoperability** of clouds
Sharing of data beyond clouds

Local collaborator

Remote collaborator
Increasing value
Building blocks

Open Cloud Mesh data sharing protocol

Interoperability platform
Requirements to join the mesh

• Have a deployment of ownCloud or NextCloud EFSS platforms
• Follow the instructions to join as a node in the trusted federation https://sciencemesh.io
• Get in touch
• Workshop on EGI Conference June 2023
Thank you

Get in touch:
hugo.Gonzalez.Labrador@cern.ch