

ScienceMesh: pan-European collaboration fabric for Science

Hugo Gonzalez Labrador , S. Alfageme , D. Castro, P. Ferreira, G. Del Monte, J. Ferrer, G. Lo Presti, J. Moscicki, E. Ragozina

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

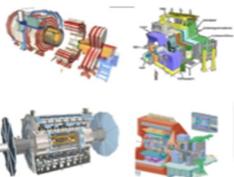
- Intro to (some) CERN Scientific services
- Outreach and Education opportunities
- Cross-institutional challenges
- ScienceMesh
- Summary



37K user home directories



+1K Project Areas



Supporting +31 Experiments

CERNBox user community

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

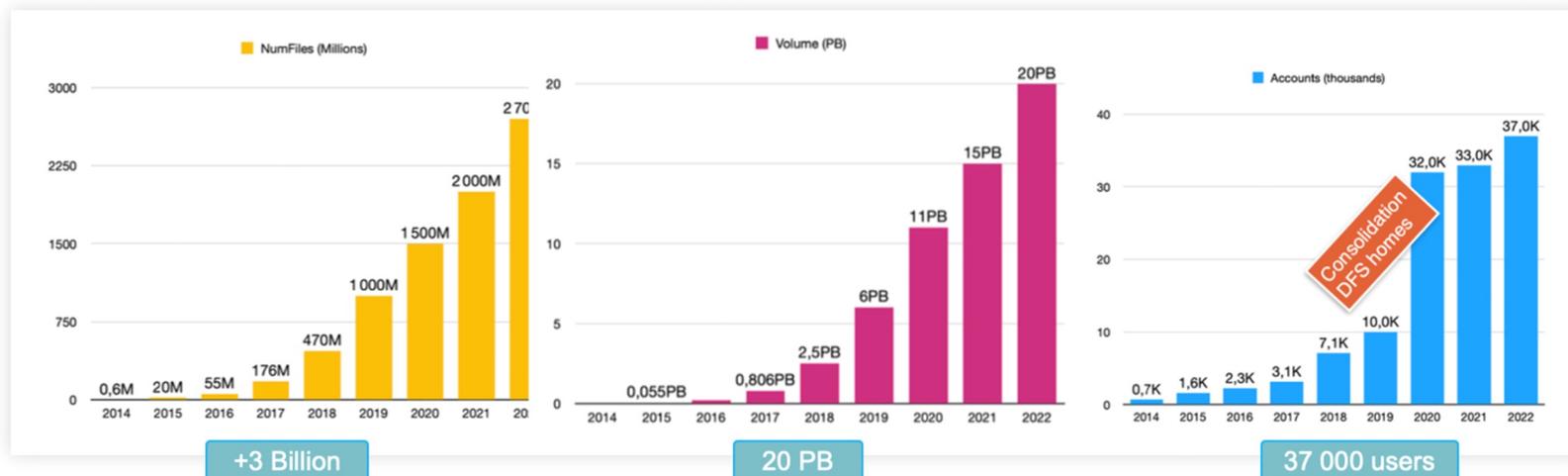
Figaro (CERN) - December 8, 2021



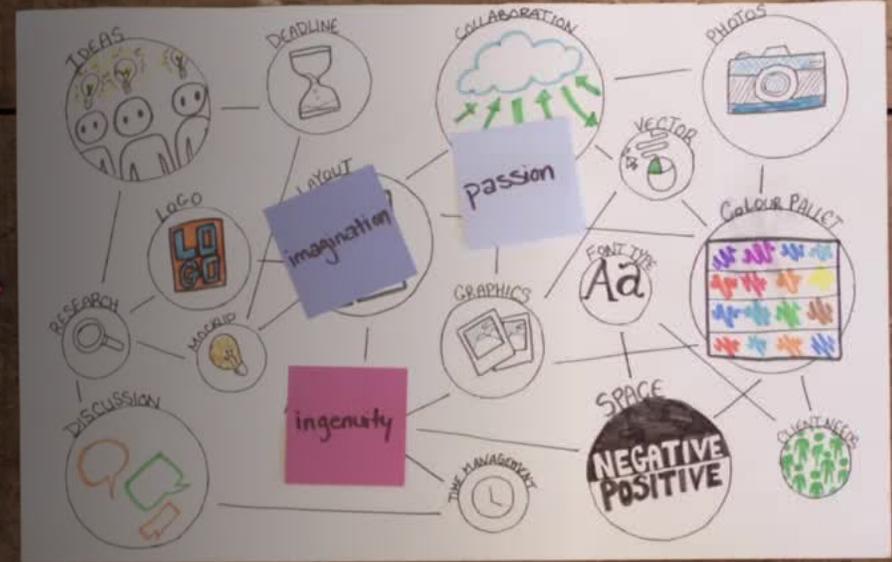
- +1000 Project areas
- 31 LHC and non-LHC experiments working areas

13k unique users/month

Skyrocketing usage



Some Outreach and Educational activities



Helping students to reach their educational goals

UP² Up To University OVERVIEW ▾ NEWS UP2UNIVERSE OPENUP2U LEARNING PLATFORM RESOURCES CONTACT

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 files.

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

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

UP² Up To University OVERVIEW ▾ NEWS UP2UNIVERSE OPENUP2U LEARNING PLATFORM RESOURCES CONTACT

CERNBox

The Sync and Share solution for Science

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”[1](#)”
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 (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!

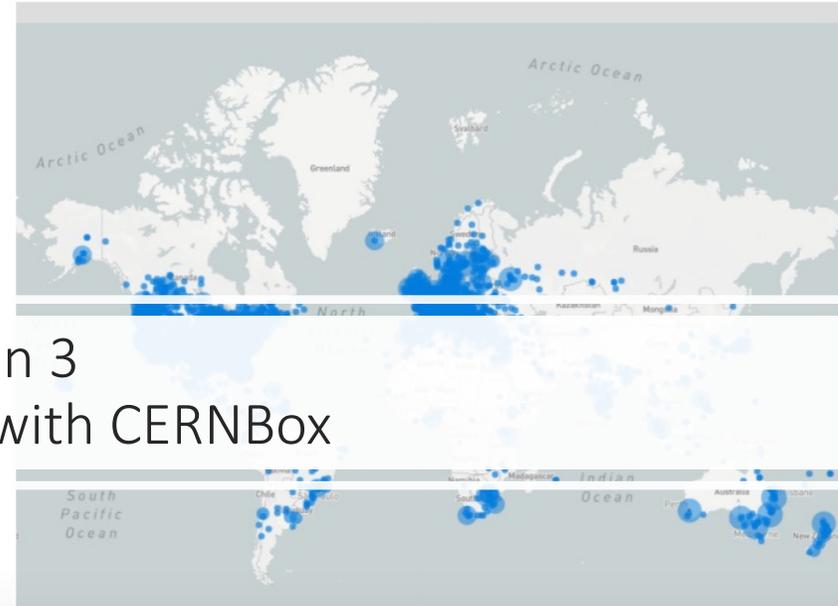
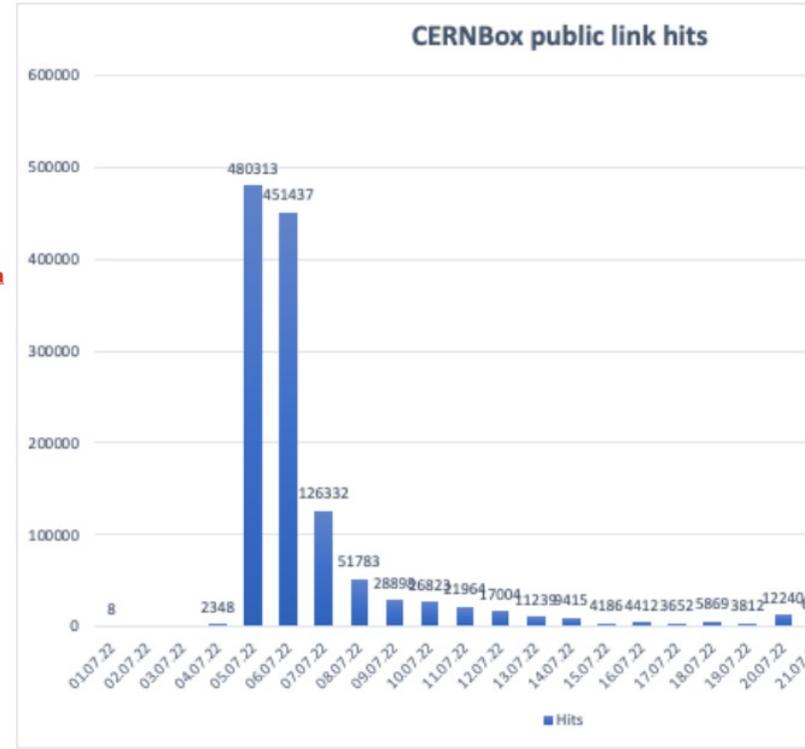
Pictures of the day are available [here](#).

Videos of the event are accessible [here](#).



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

- 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_16h48_089.JPG
- ATLAS_16h50_029.JPG
- ATLAS_16h55_049.JPG
- CCC_15h08_090.jpg



Country	IPs
Worldwide	24701 IPs
United States	18323 IPs
Canada	845 IPs
United Kingdom	628 IPs
Germany	541 IPs
Switzerland	525 IPs
France	504 IPs

LHC Run 3
Media Coverage with CERNBox



CERNBox, GSOC, HSF and Open Source

- Sourcing talent while promoting science

2019 Program | CERN-HSF
Contributor
Mohit Tyagi
View Code
CERNBox: Bring Your Own Application

2018 Program | CERN-HSF
Contributor
Ivantha
View Code
Petabyte-Scale Cloud Storage File Manager

2023 Program | CERN-HSF
Contributor
daniel-rey
Log in
Extend and improve the CERNBox Notifications platform

2021 Program | CERN-HSF
Contributor
Jimil Desai
View Code
Runtime plugin ecosystem support for OCIS

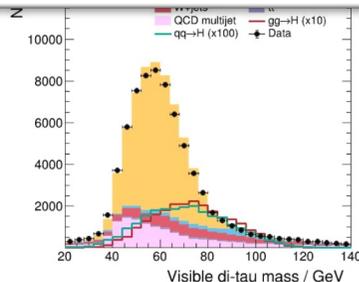
2021 Program | CERN-HSF
Contributor
Rahul Chauhan-1
View Code
Rucio and CS3API to enable data management for the ScienceMesh cloud

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**



Basic Examples
ROOT Primer
Accelerator Complex
Beam Dynamics
Machine Learning
Apache Spark
Outreach
AWAKE

Outreach

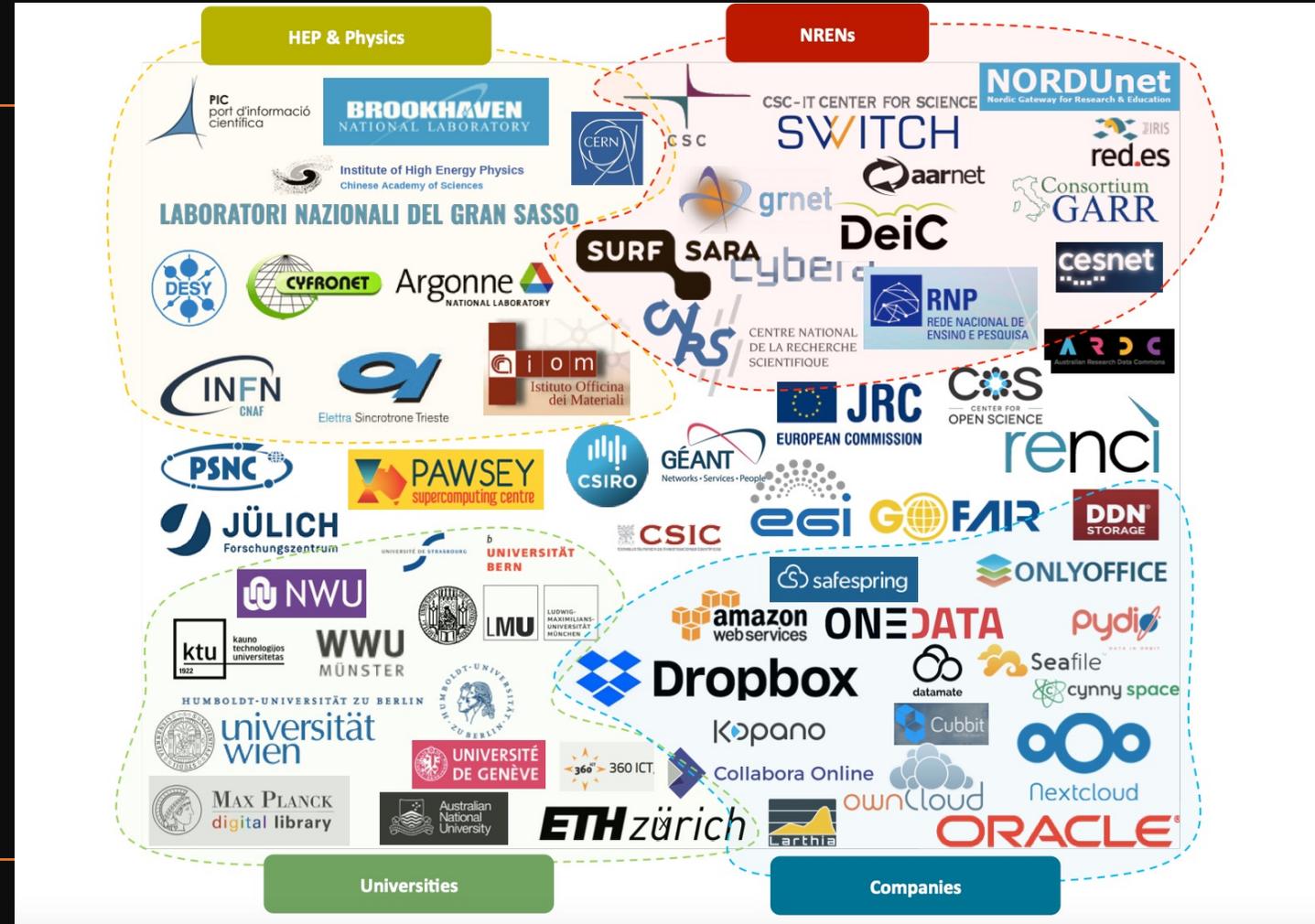
Reach out with SWAN! This section collect a series of outreach efforts involving SWAN.

Particle open data teaching (Hiukkasfysiikan avoin 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. Paavo Rikkilä (CMS) put together an introductory course about experimental HEP for future high school teachers. The result is great: check it out in SWAN!

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

Analysis Facilities White Paper

Diego Ciangottini (INFN Perugia, Editor), Alessandra Forti (University of Manchester, Editor), Lukas Heinrich (Technische Universität München, Editor), Nicola Skidmore (University of Manchester, Editor),

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.

Abstract	1
Introduction	1
Analysis infrastructures from a user perspective	2
Goals	2
Topics for discussion:	7
Analysis infrastructure models	8
DOMA	10
Federated Identity Management	12
GPU Resources	13
Analysis portability and preservation	13
Monitoring and Metrics	14
End user documentation	15
Summary	16

Introduction

In the HL-LHC era (to begin in 2029) LHC analysts will have over 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

There is an important demand for users to collaborate beyond their local clouds





Science Mesh

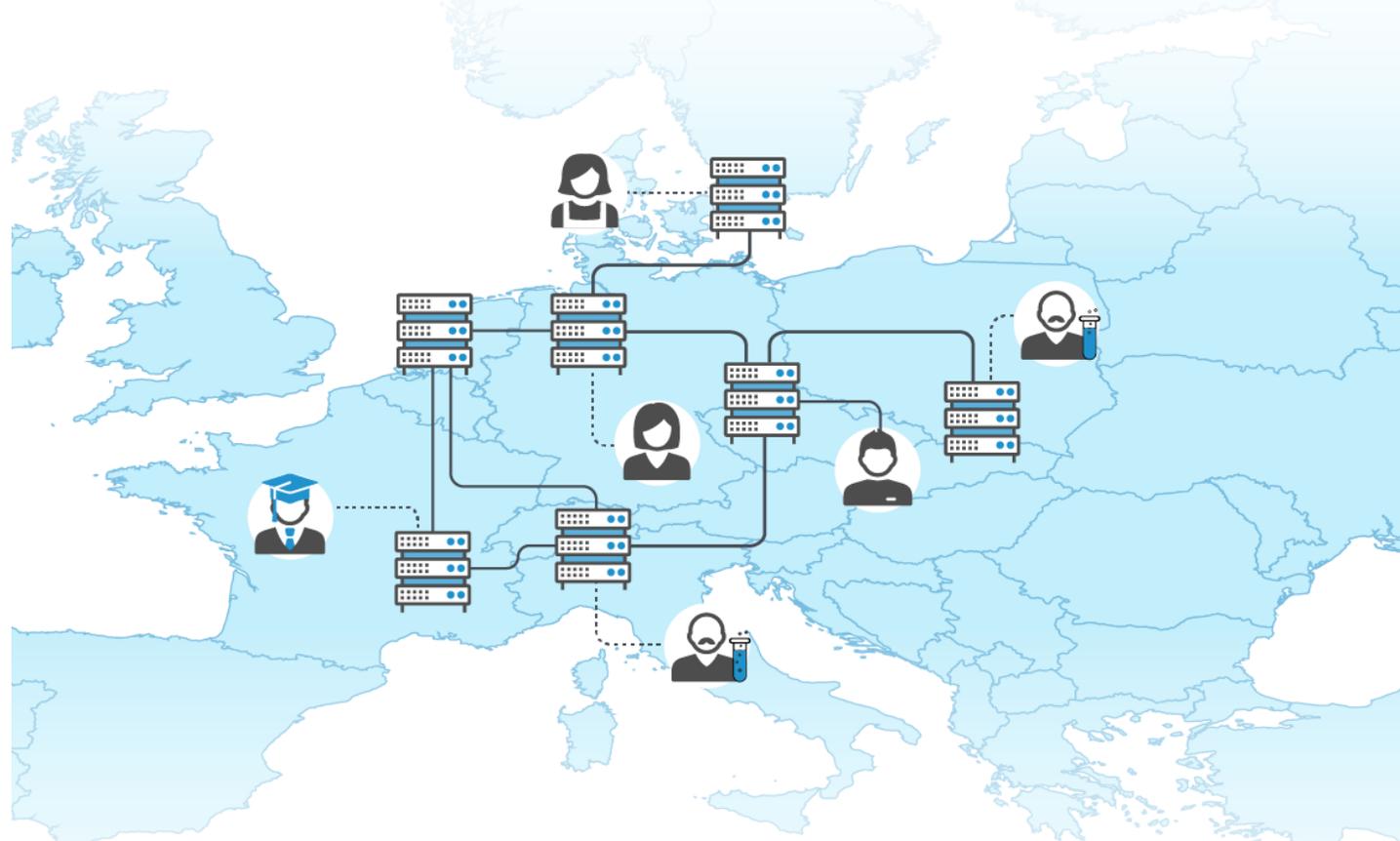
One way to remove barriers



Science Mesh

Uniting European data services with **CS3** APIs
Supporting researchers with **CS³MESH⁴EOSC**

- 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 cloudds

< Shares ×

big-folder

Share with people ?

Account type Invite

standard ▼ GLP Giuseppe Lo Presti ×

Invite as viewer ▼ Notify via mail **Share**

Local collaborator

< Shares ×

big-folder

Share with people ?

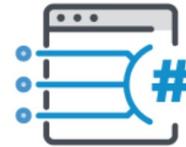
Account type Invite

federated ▼ dbenjamin@bnl.gov

Invite as viewer ▼ Notify via mail **Share**

Remote collaborator

Increasing value



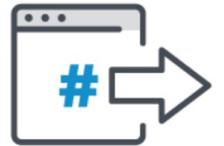
**Data Science
Environments**



Open Data Systems

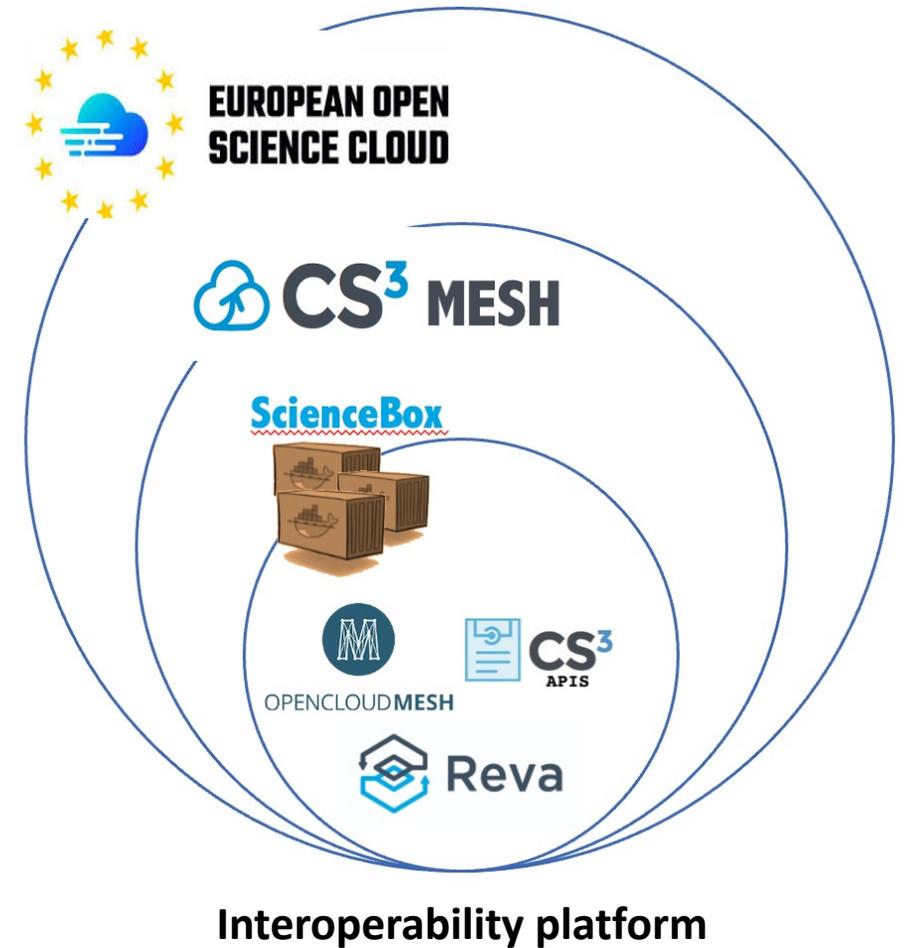
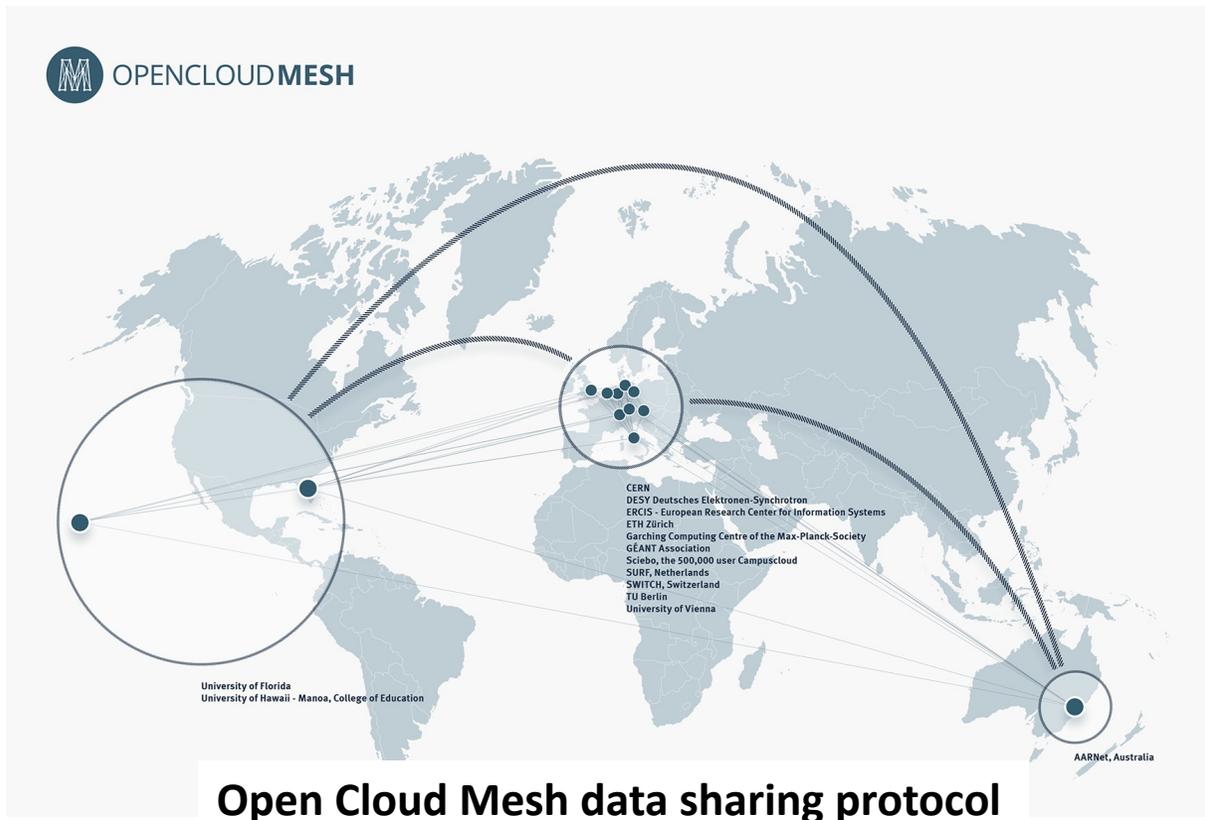


**Collaborative
Documents**



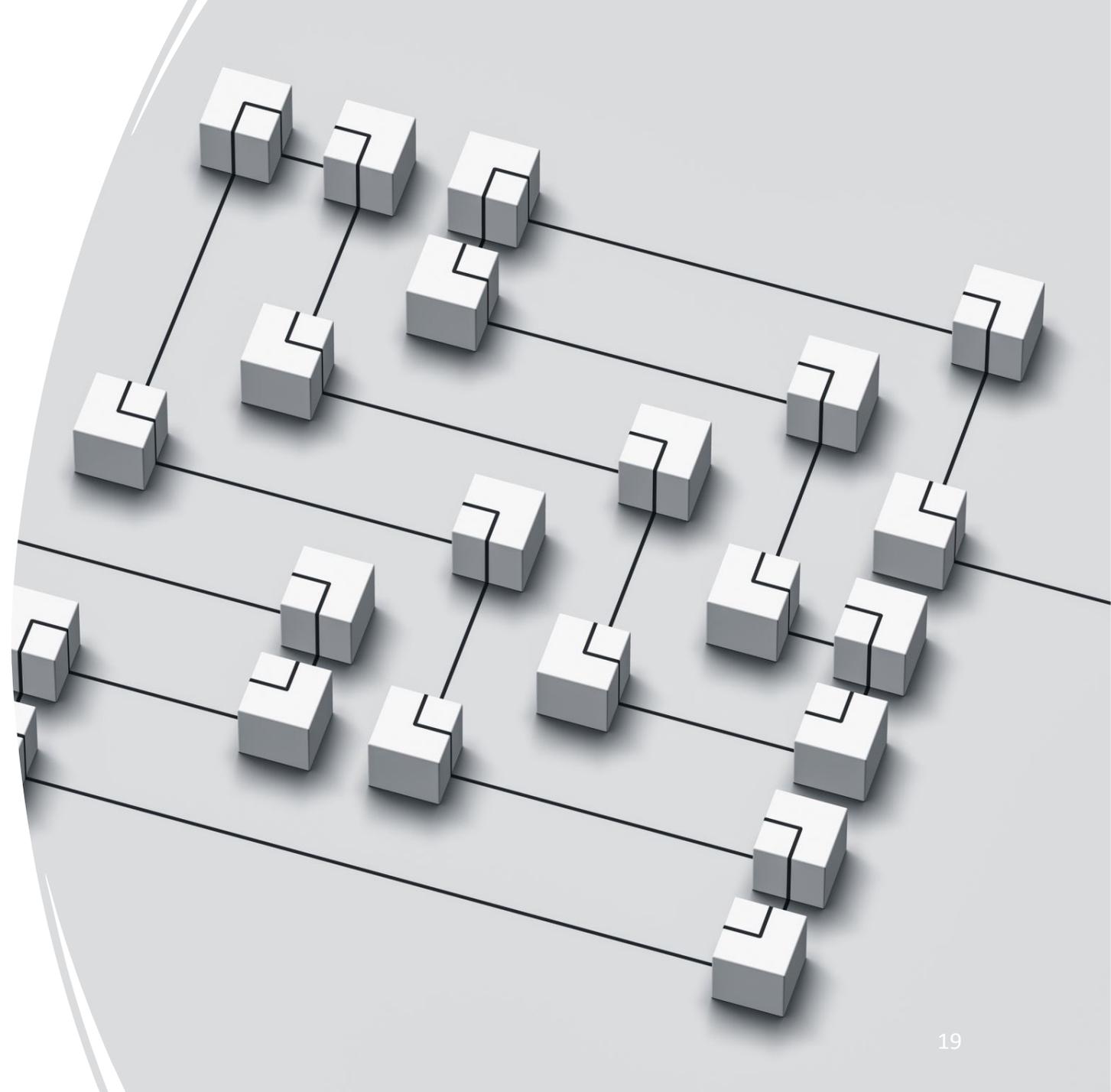
**On demand large
dataset transfer**

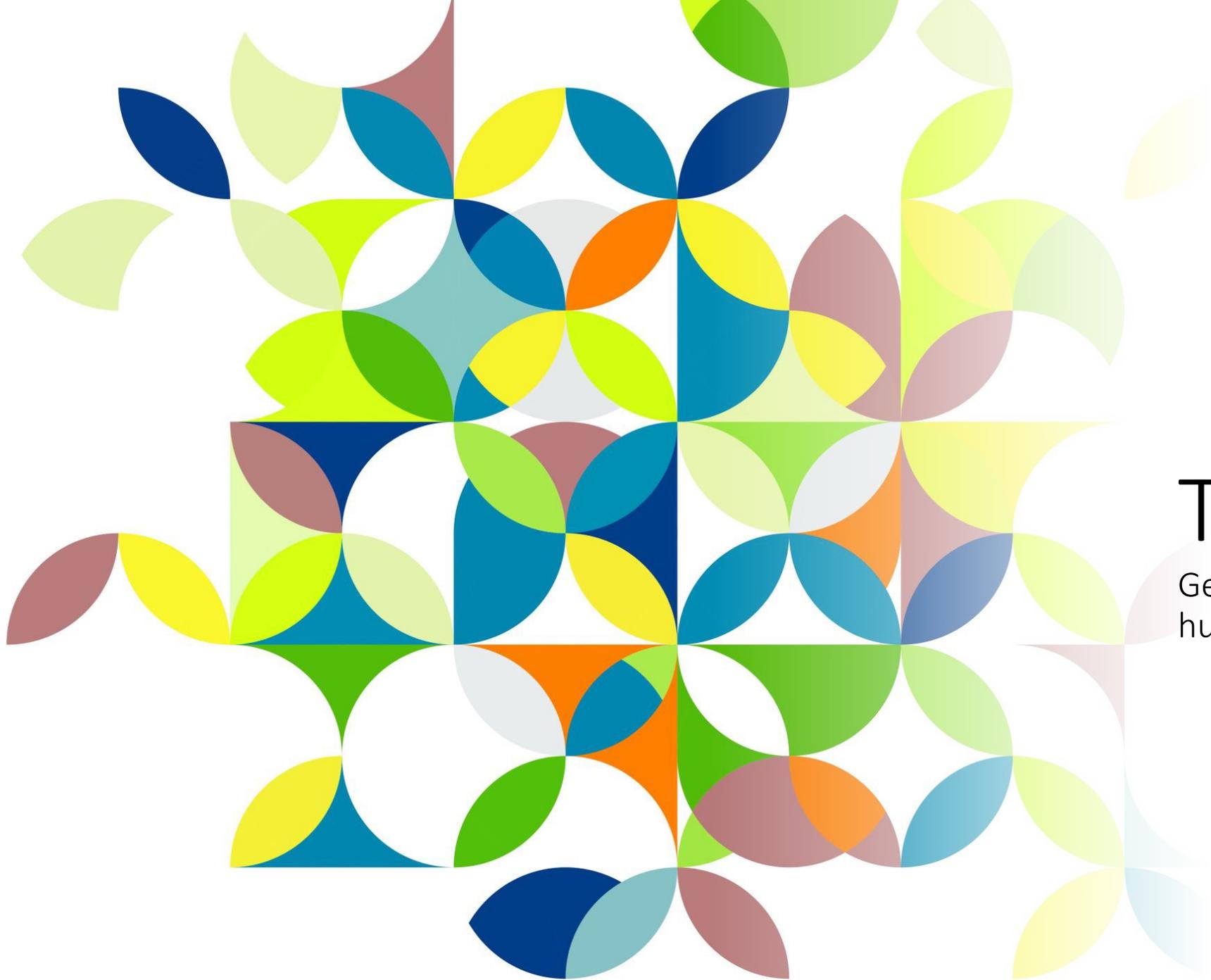
Building blocks



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