OSCARS
Open Science Clusters' Action for Research & Society

Funded Project

MADDEN: Multi-RI Access and Discovery of Data for Experiment Networking

Presenter: Federica Legger, INFN Torino, ORCID: 0000-0003-1400-0709

Implemented by

INFN UCLouvain

Funded by the European Union

What problem(s) did you plan to solve?

- Scientific collaborations often **need to share data**, especially in fields such as Gravitational Wave (GW) research, where the same GW event is detected by several interferometers
- Each Research Infrastructure (RI) typically manages its own data in **isolated** Data Lakes
- We are working to establish **a common data infrastructure** that allows scientists to securely access data across multiple RIs
  - Current interferometers: **LIGO (US), Virgo (Europe), KAGRA (Japan)**
  - Future interferometers: **Einstein Telescope (ET) in Europe, Cosmic Explorer (CE) in US**

What have you done to solve the problem?

- We are using **Rucio** to build a multi-RI Data Lake for the GW community
  - Rucio is an open-source framework for data management, distribution and access, initially developed for the ATLAS experiment at CERN and now used by communities across several domains (HEP, Nuclear, Astro)
  - LIGO and Virgo are using Rucio in production (data management and distribution but not for access), and ET is currently evaluating its use in the design of their computing model
- **Need to extend Rucio functionalities to make it a better fit for the GW community**

What are the key results achieved to date and how have you made them available to the broader community?

- A **multi-RI Data Lake** for ET and CE has been built using Rucio
  - Two independent Rucio instances for ET and CE have been deployed
  - The ET IAM (Identity and Access Management) is used for authz
  - ET users can seamlessly access data from both Rucio instances
- We extended Rucio to **support a Multi-RI setup (PR merged in Rucio 39.2)**
- Participating to the xRIDGE Data Challenge organised by ESCAPE in March 2026

How will make your results sustainable over time - How will the scientific community/-ies further exploit them?

- The Multi-RI Data Lake is being interfaced with the **ET analysis facility** based on the **ESCAPE VRE** (Virtual Research Environment) implemented by the companion **ETAP** project
  - MDC data are already available in the Data Lake
- To increase Rucio usability for scientists we are developing a plugin to provide a **POSIX-like view** of the Rucio catalogue using **CernVM File System (CVMFS)**
- A technology demonstrator for the Virgo collaboration is being set up
- Will provide support for next Mock Data Challenges (MDC) for ET

Who has been doing it?

- N. Avdeev (INFN Torino), ORCID: 0009-0005-0413-633X
- K. Jansson (UCLouvain), ORCID: 0000-0001-8095-7504
- L. Lavezzi (INFN Torino), ORCID: 0000-0002-4928-8151
- A. Tanasijczuk (UCLouvain)