Distributed Data Management with Rucio for the Einstein Telescope

Lia Lavezzi (INFN Torino)

on behalf of the ET Computing Team

The speaker's contract is funded by the project ETIC:

Einstein Telescope Infrastructure Consortium (ETIC - IR0000004)

PNRR MISSIONE 4, COMPONENTE 2, INVESTIMENTO 3.1

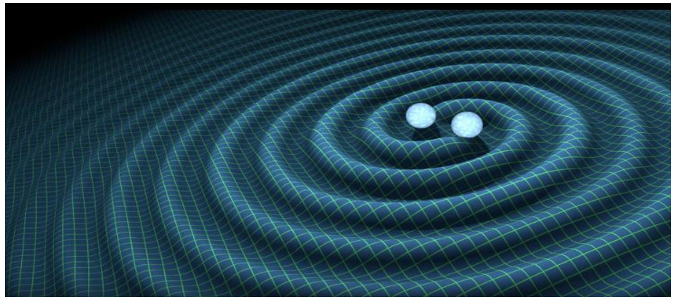
Einstein Telescope

It is a **Third Generation Ground Observatory** (3G) for Gravitational Waves, currently in the preparatory phase, to become operative in about 10 years.

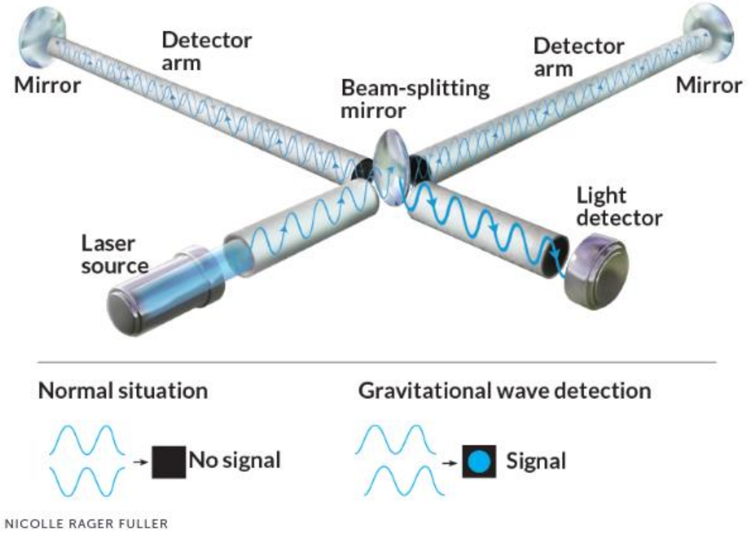
Composed by Michelson interferometers, disposed in a triangular or double-L pattern, with Italy (Sardinia) and EU Regio Rhine-Meuse as candidate sites

Gravitational Waves (GW) are ripples in the space-time fabric, generated by massive celestial bodies or highly energetic cosmic events

- 1915 theoretical prediction (Einstein)
- 2015 first experimental observation (LIGO-Virgo, BBH)
- 2017 first multi-messenger observation (LIGO-Virgo, BNS)



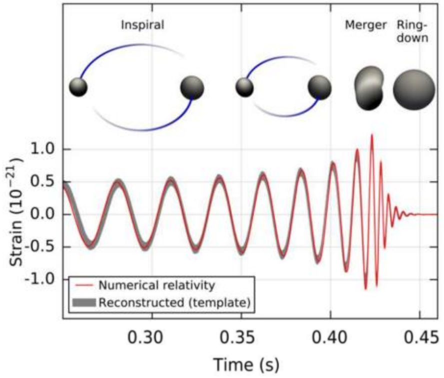
www.ligo.caltech.edu



To detect the passage of a GW the goal is to measure a change in the length of the two arms.

strain = differential arm length

This quantity is sampled in time ⇨ time series



PRL 116, 061102 (2016)

Einstein Telescope

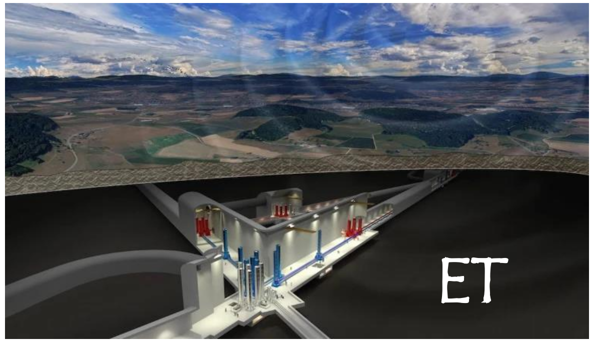
Will improve the sensitivity of more than **one order of magnitude** with respect to current 2G interferometers, will increase the signal-to-noise ratio, suppress all kinds of noise and widen the field of observation.

www.ligo.caltech.edu

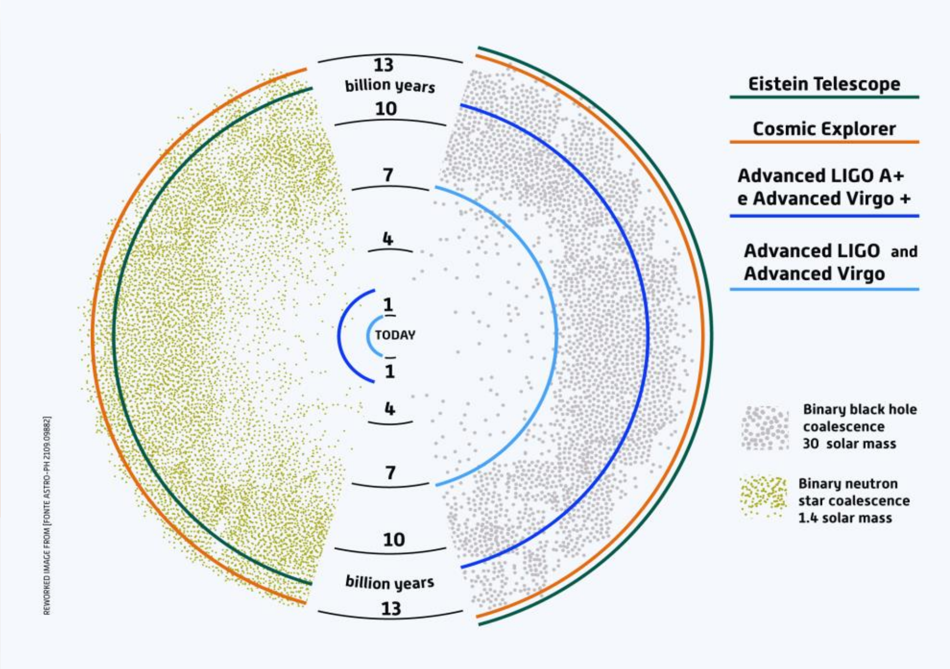
www.virgo-gw.eu

Improvements *w.r.t.* 2G interferometers

- longer arms (10 km)
- underground (**minus** 200-300 m)
- each detector consists of two separate interferometers:
  - cryogenic temperature [low frequency 3–30 Hz]
  - room temperature [high frequency 30–10^3 Hz]



www.nature.com



- **Better sensitivity** means to observe farther in distance, *i.e.* closer in time to the Big Bang
- **Lower frequency** means to observe longer signals (hours) and heavier black holes (M ~ 10^4 $M_\odot$)

The rate of observations will increase

100 obs/year (current) ⇨ 100 thousands obs/year (ET expected)

Gravitational Wave Data Requirements

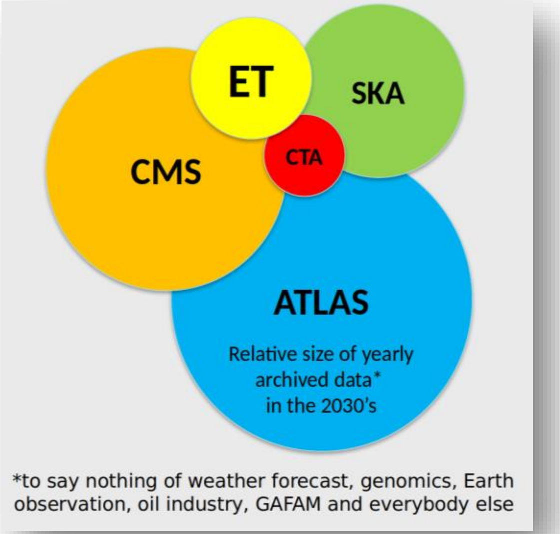
- Two kinds of data:
  - **physics data** are time series with a defined duration of the *strain* variable, do not scale with sensitivity:

    ET expected amount ~10 TB/year
  - **raw data** are collected from all the sensors, do scale with interferometer complexity:

    ET expected amount ~100 PB/year

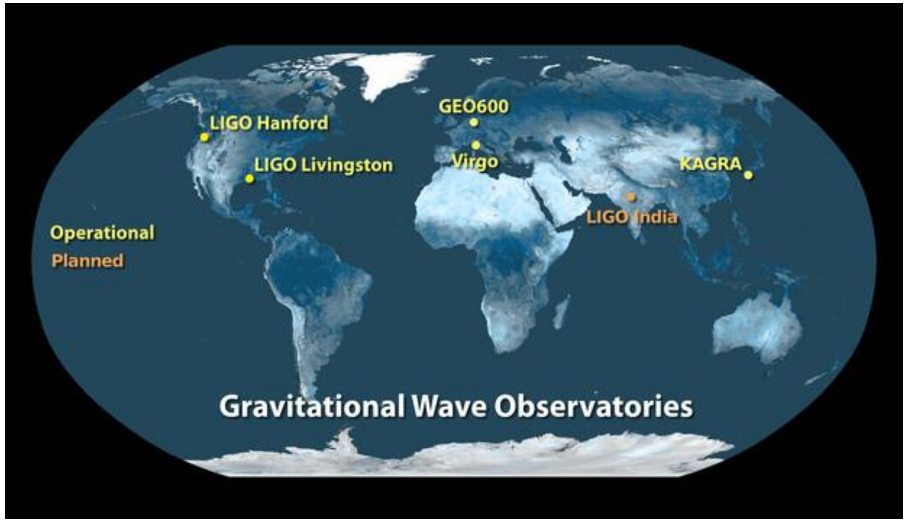
  [See ET Preparatory Phase Project - D8.1 Computing and Data Requirements]

  What grows is the amount of useful scientific information in the data and the needed computing power.



[S. Bagnasco, Erice Science Communication and Journalism International School 2023]

- 2G interferometers LIGO, Virgo and KAGRA participate to the International GW Network (IGWN) and use a common infrastructure.



www.ligo.caltech.edu/

Next generation interferometers will also create a strong **network**, then ET will need to share data with the Cosmic Explorer (3G interferometer in the US) and LISA (ESA mission space interferometer) which will be taking data in the same decades as ET ⇨ we will need a **multi-RI access to the data.**

Gravitational Wave Data Requirements

ET is currenctly evaluating the tools to define its computing model:

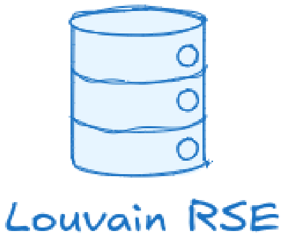
- Data distribution and management via Data Lake using Rucio
- RucioFS to provide a POSIX view of the data
- ESCAPE (*) VRE JupyterLab for interactive development
- Analysis tools as Snakemake (backends: HTCondor grid, Slurm HPC clusters, clouds) and REANA for open science

See Stefano Bagnasco's talk tomorrow for details on ET computing model

IN THIS TALK I WILL SHOW OUR ONGOING WORK RELATED TO RUCIO

(*) European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures projectescape.eu/

Current Rucio Setup in Einstein Telescope



One of the two `XRootD` *origin* data servers federated to OSDF(*) was dedicated to host ET data:

- ~2 TB simulated data
- published also on CVMFS
- **used for Mock Data Challenges**

Louvain OSDF origin is currently publicly accessible.

- authentication: userpass, X.509/proxy
- **used for testing and development**
- hosts:
    - Rucio Server
    - Rucio DB (`Postgres`)

- `XRootD` standalone server
- 260 GB space
- **used for testing and development**

Pictures drawn on excalidraw.com/

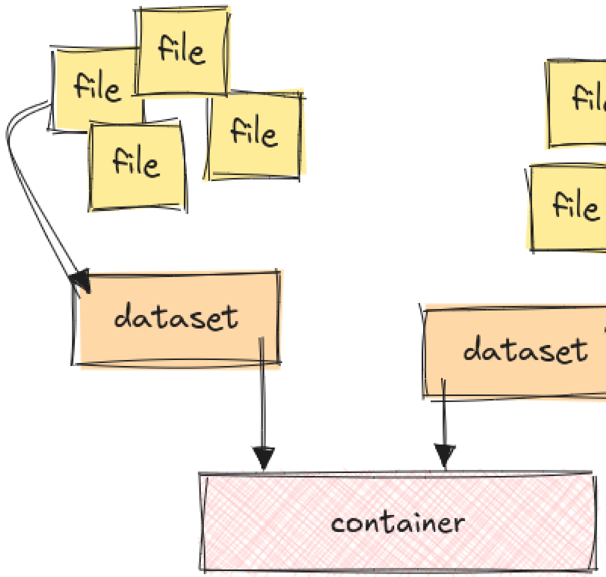
(*) Open Science Data Federation osg-htc.org/services/osdf.html

RucioFS: Rucio Filesystem

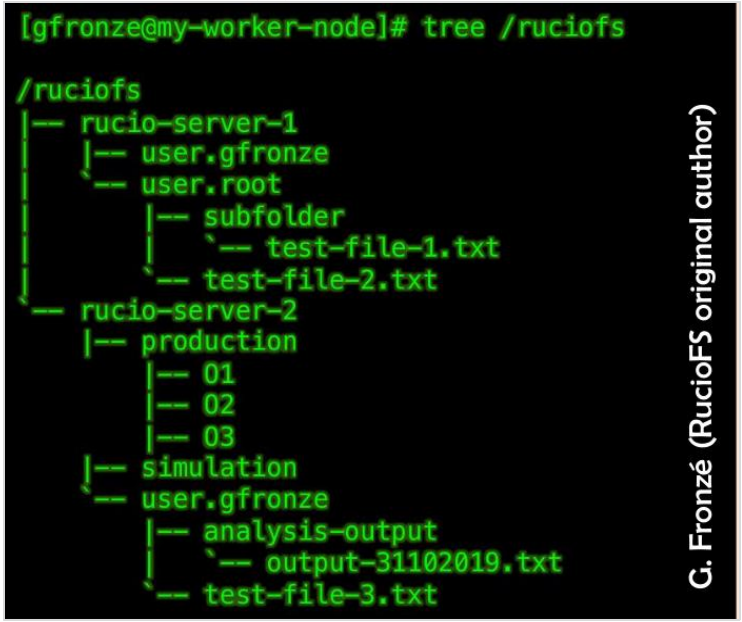
GW users and analysts *love* POSIX-like filesystem, in fact currently all experimental data are published on CVMFS.

RucioFS is a POSIX-like FUSE mount filesystem that transforms Rucio scope/container/dataset/file scheme in a hierarchical structure of directories.

STANDARD RUCIO DATA



RUCIOFS DATA



- read-only filesystem
- wraps around python Rucio client
- supports multiple servers
- tested on CentOS7, Ubuntu 18.04 LTS and AlmaLinux 9
- written in C/C++11
- files downloaded @ first access

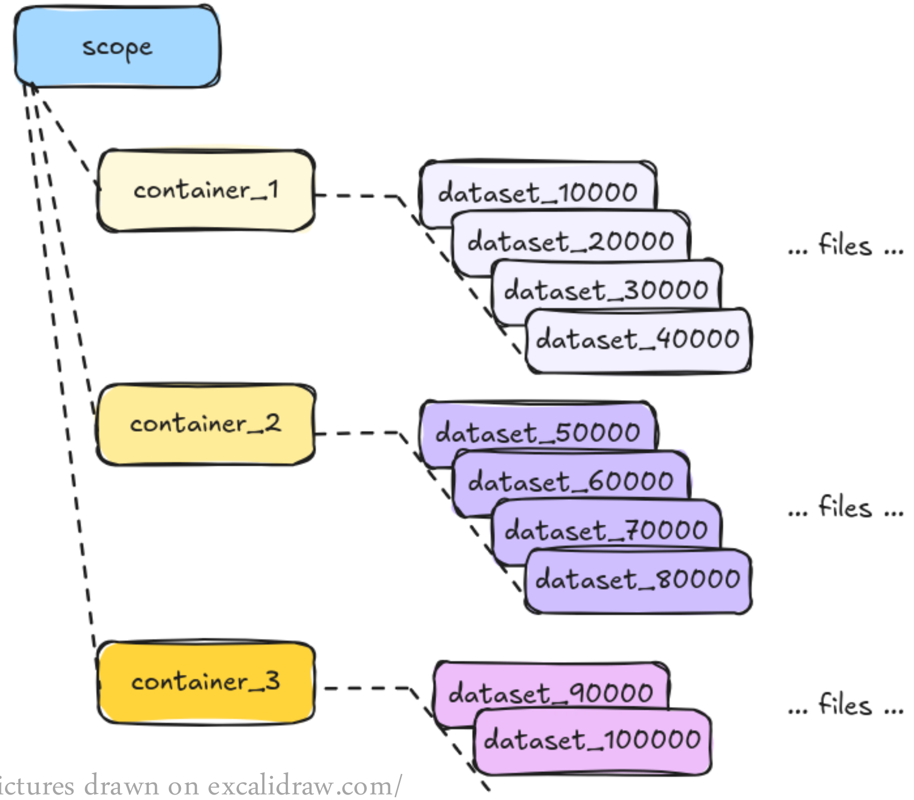
Pictures drawn on excalidraw.com/

RucioFS: Structure

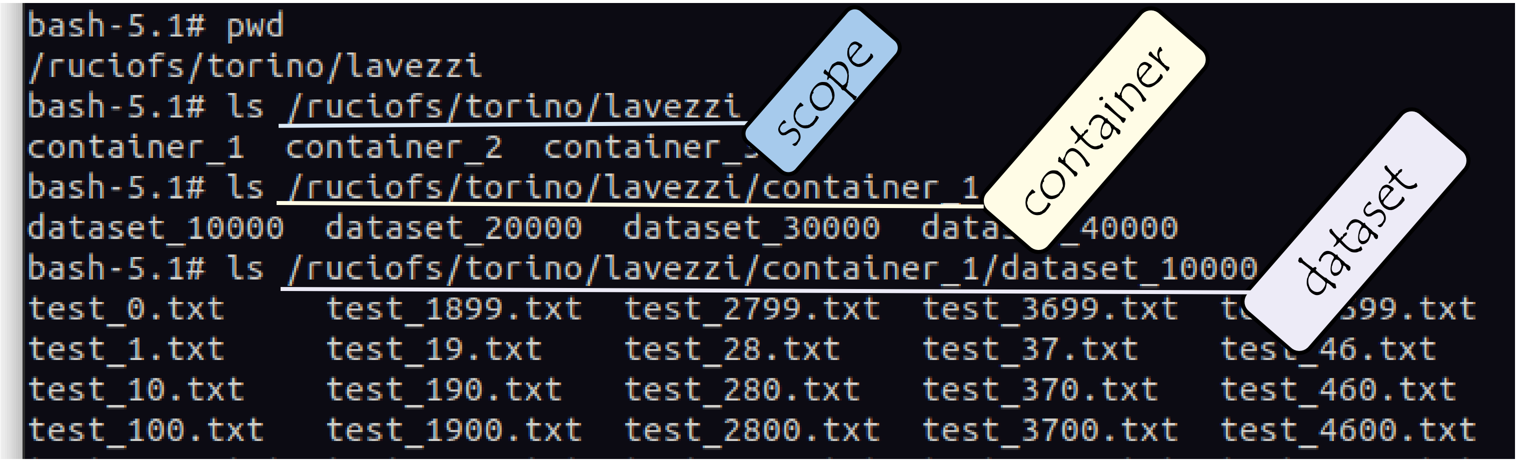
wraps around python Rucio client

- Can be used in a Docker container with the code distributed in `rucio/containers` repo
- Uses the `rucio-client` image and downloads the RucioFS code from its original repo: github.com/rucio/fuse-posix

read-only filesystem



Pictures drawn on excalidraw.com/



Implemented callback functions:

- `rucio_getattr;`
- `rucio_readdir;`
- `rucio_read;`

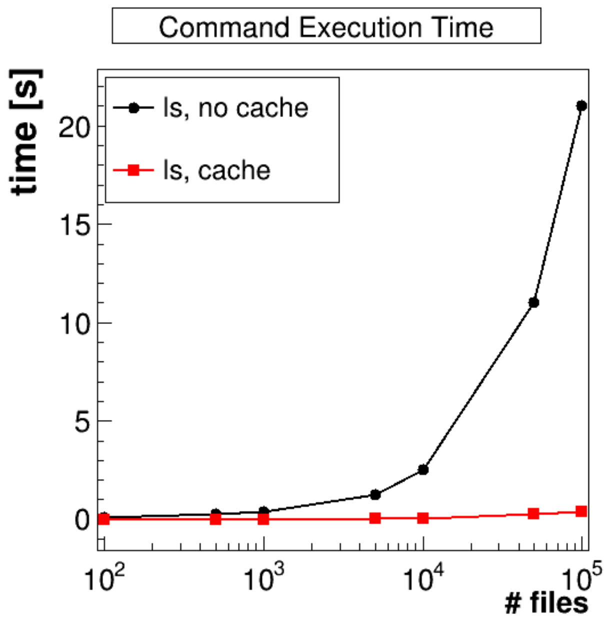
RucioFS: Cache

- The first time the data are accessed, RucioFS gets the information from the Rucio DB via `libcurl`, then it caches it
- Caching system via `std::unordered_map< std::string, std::vector < std::string > >`

- The caching system speeds up the execution time, but a mechanism to automatically refresh the cache is still missing
- The cache is updated only after the `/ruciofs` directory is unmounted and re-mounted (*i.e.* the Docker container is stopped and re-started)

THIS NEEDS TO BE INVESTIGATED & IMPROVED

RucioFS: Scalability Test



Test:

ls in the directory `/ruciofs/torino/lavezzi`, filled with an **increasing number** of files **with/without cache**

- uploaded *n* single files on Torino Rucio RSE
- each one is a text file with a random number written in it
- size ~ 4 kB/file

TIME IS LINEAR WITH THE NUMBER OF FILES

- **with caching** the time stays < 0.5 s even for 10^5 files

⇨ having a caching system is the optimal choice

- with caching, there is the need to **unmount/re-mount** the filesystem, while without caching there is not

⇨ need to implement a way to refresh the cache periodically

ET use case numbers:

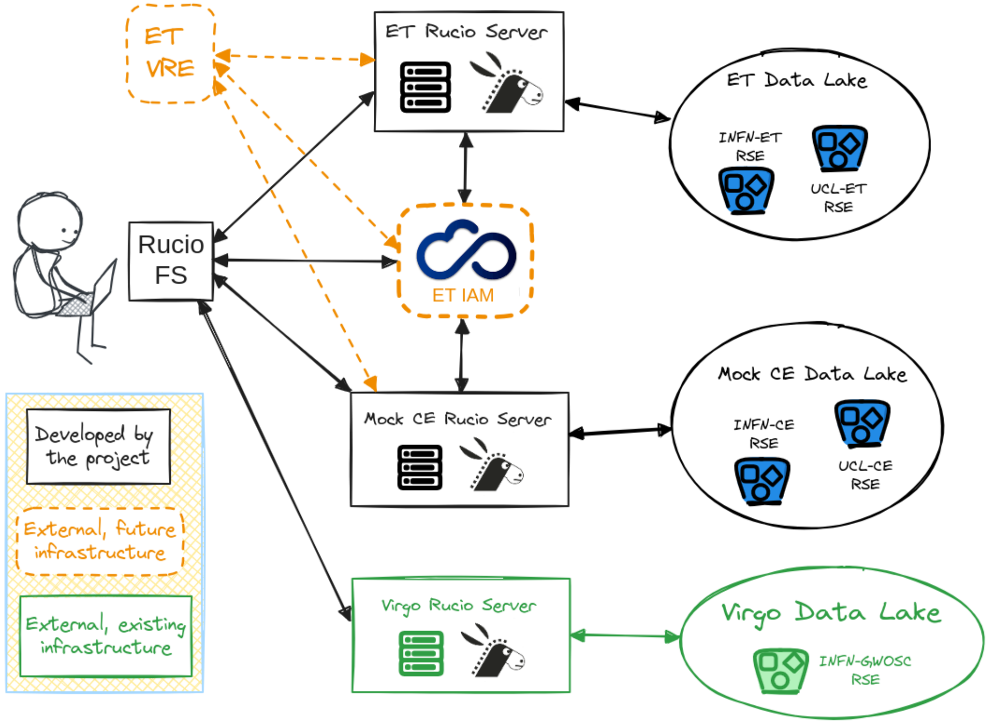
- MDC1 dataset: 1300 files, 155 GB per interferometer (~ 1month data taking)
- 3 interferometers + 1 null channel, noise only + noise&signal ⇨ total 10400 files

Virgo use case numbers

- take O4 as a reference, from Spring 2024 to now
- ~ 57000 files, ~ 3.2 TB (strain + environmental)

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

Federica Legger (INFN Torino)
Andres Tanasijczuk (UCL)



Participating organizations:

- INFN Torino (PI: Federica Legger)
- UC Louvain (coord.: Andres Tanasijczuk)

- Targeted start date: January 2025
- Duration of the project: 24 months
- Overall funding: 210 K€

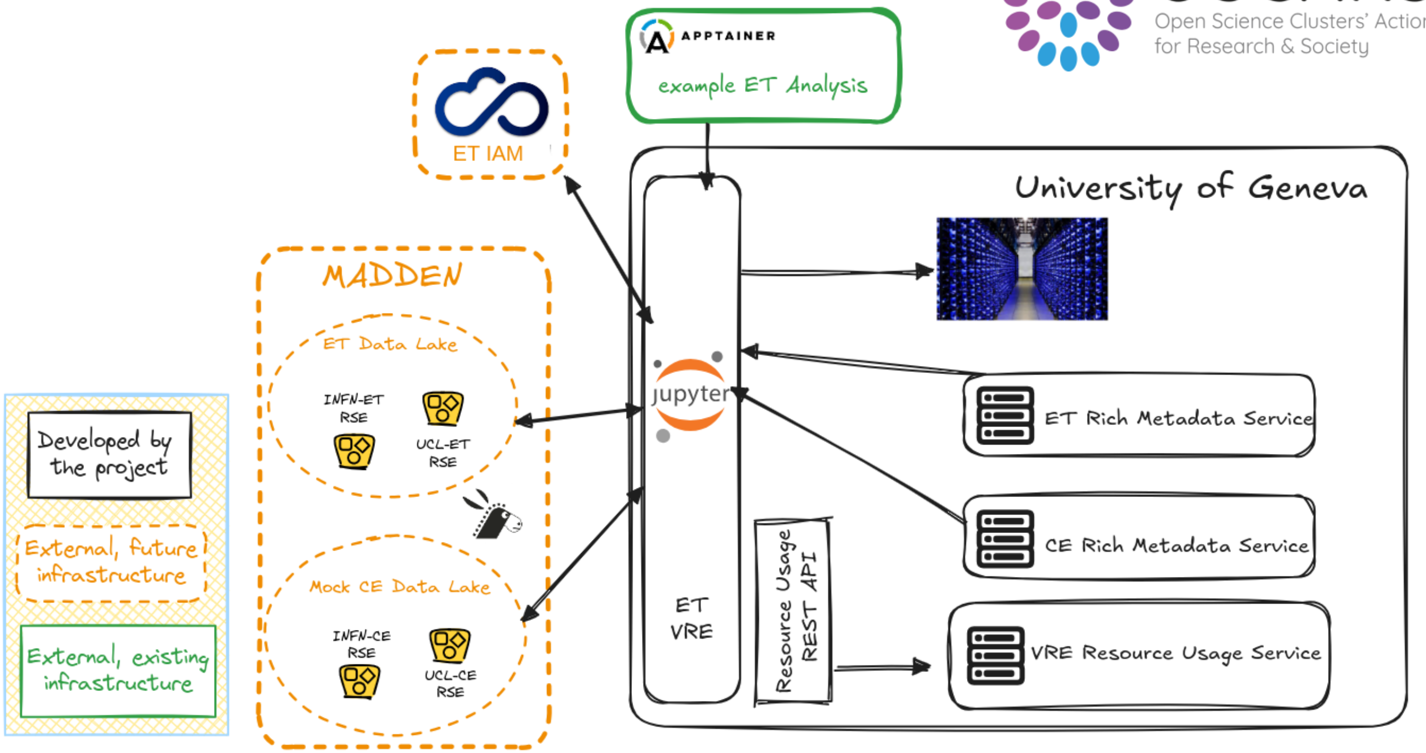
The main objectives of this project are:

- Build a multi-RI Data Lake managed with Rucio.
- Develop and test RucioFS, a tool to provide a POSIX-like view of the Rucio catalogue in a multi-RI environment.
- Extend RucioFS to support advanced querying capabilities using Rucio metadata.

Pictures drawn on excalidraw.com/

ETAP Einstein Telescope Analysis Portal

Paul Laycock (UniGe)



Participating organizations:

- University of Geneva (PI: Paul Laycock)

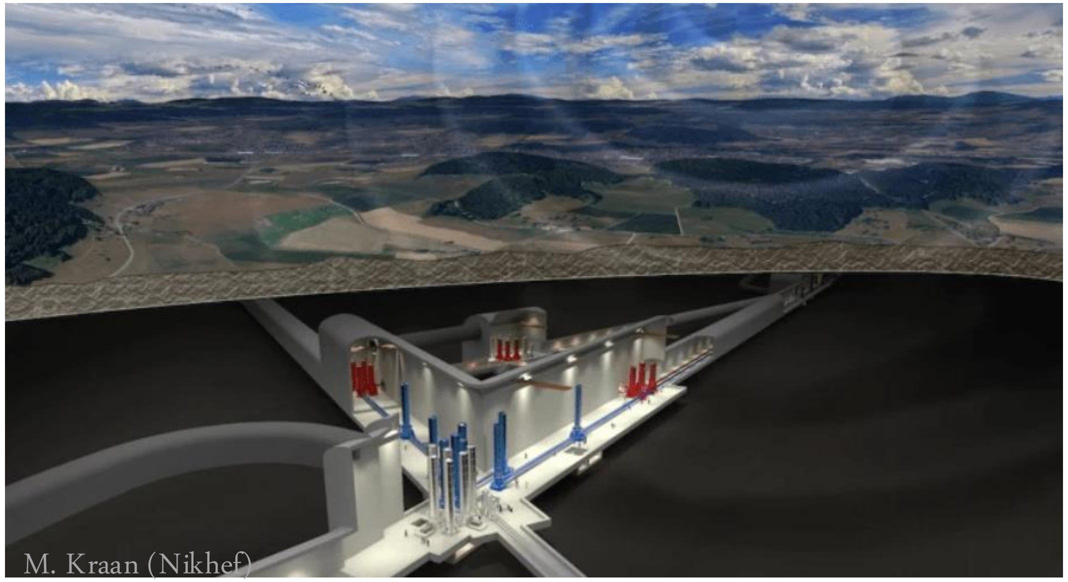
- Targeted start date: January 2025
- Duration of the project: 18 months
- Overall funding: ~250 K€

The main objectives of this project are:

- Deploy the CERN ESCAPE VRE at University of Geneva
- Connect to multi-RI Data Lakes managed by Rucio (MADDEN)
- Deploy multi-RI Metadata services from the HEP Software Foundation (HSF)
- Design a flexible computing resource monitoring service

Pictures drawn on excalidraw.com/

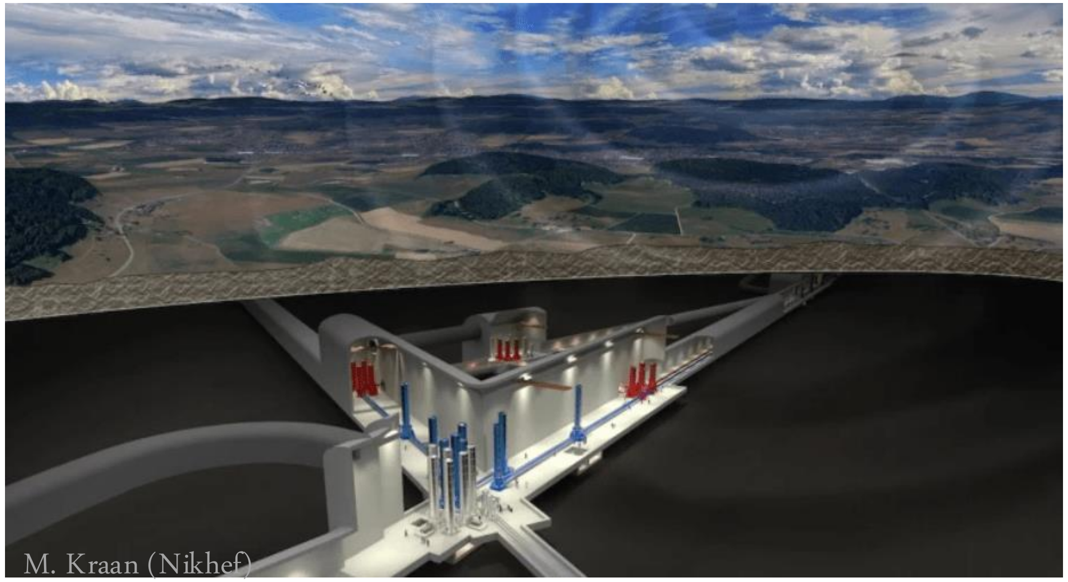
Conclusions



- ET is a 3G ground-based observatory for Gravitational Waves
- It is in preparatory phase that will become operative in about 10 years.
- It is currently defining the computing model and evaluating Rucio for data management.
- Wished features:
  - Support for multi-RI Data Lake
  - POSIX-like view of the file catalogue
  - Extended metadata support

We are actively working on the development of these features through current and future funded EU projects

Conclusions



- ET is a 3G ground-based observatory for Gravitational Waves
- It is in preparatory phase that will become operative in about 10 years.
- It is currently defining the computing model and evaluating Rucio for data management.
- Wished features:
  - Support for multi-RI Data Lake
  - POSIX-like view of the file catalogue
  - Extended metadata support

We are actively working on the development of these features through current and future funded EU projects

Thank you for the attention