Einstein Telescope Mock Data Challenge

ET: Scienza e Tecnologia in Italia, Assisi, 20-23 February 2024

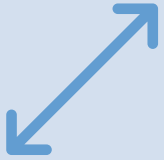
Elena Cuoco
European Gravitational Observatory

ET-0036A-24

Why ET MDC

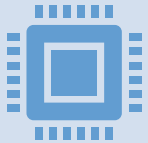
- Training on simulated data to test methods and pipelines
- Find out the limitations of current methods
- Encourage the community to develop new tools
- Provide a common dataset for comparison of analysis methods
- Assess the science potential with ET
- Assess the requirements for computing infrastructure

What's in the ET MDC



Series of MDCs with increasing complexity

First MDC with Gaussian coloured noise and a cosmological population of CBCs (more details in the next slide)



Next MDCs with all type of sources, glitches, correlated noise (need to be implemented in the simulation code)

First ET MDC

1 month duration

Gaussian colored noise (ET-D 10 km, fMin=5Hz), triangle

Tania Regimbau, Thomas Dent, Walter Del Pozzo, Stefanos Giampanis, Tjonnie G. F. Li, Craig Robinson, Chris Van Den Broeck, Duncan Meacher, Carl Rodriguez, B. S. Sathyaprakash, and Katarzyna Wójcik
Phys. Rev. D **86**, 122001 – Published 3 December 2012

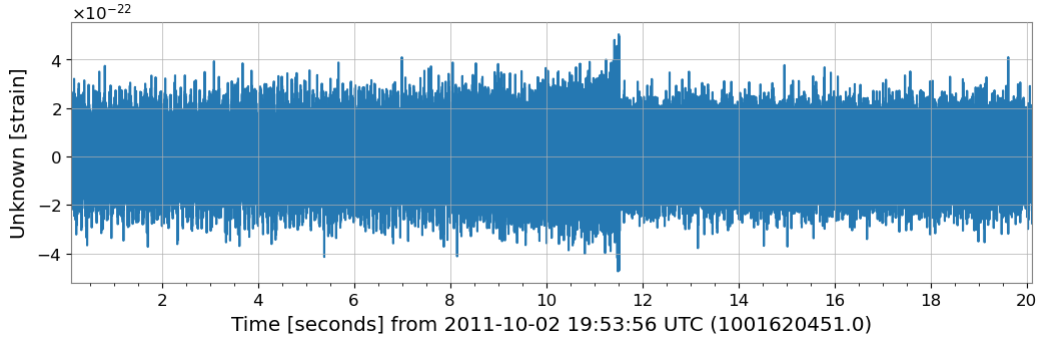
Population of BBHs (10%), BNSs (87%) and NSBHs (3%) (see CoBA) with isotropic distribution in the sky.

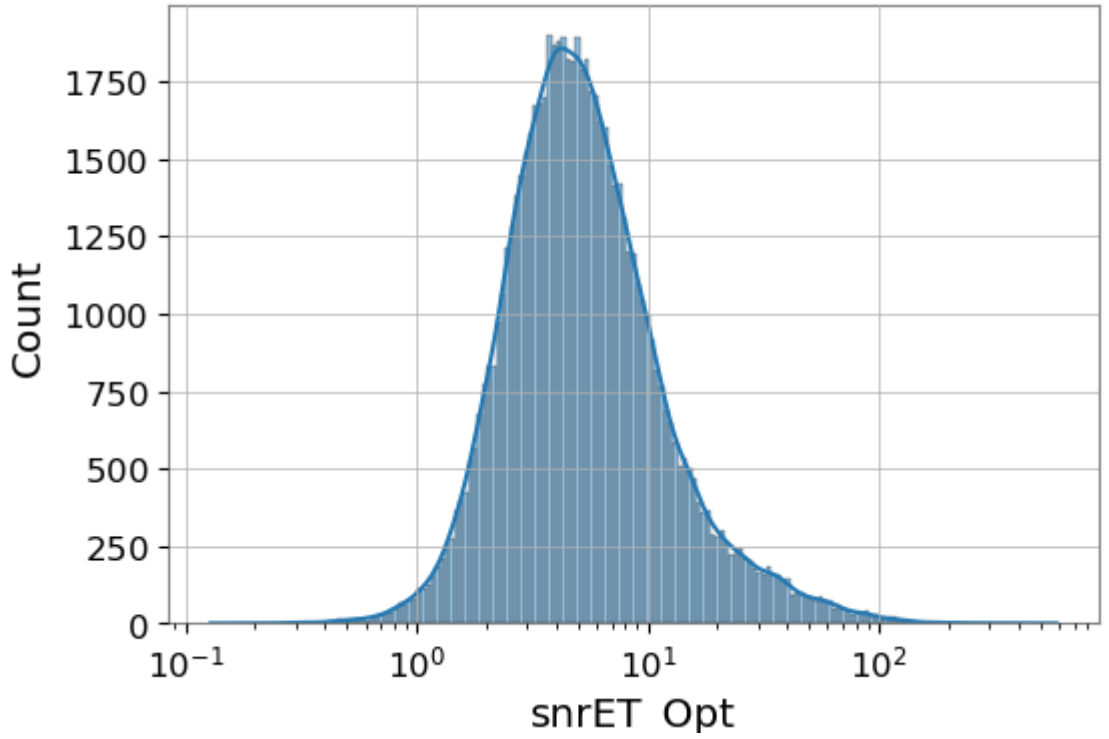
CBC Waveforms: IMRPhenomXPHM for BBHs and BHNSs, and IMRPhenomPv2 with tidal effects NRTidalv2_V for BNSs.

Some Statistics about the injected signals

SNR>8: 11551 BNSs , 537 BHNSs, 6119 BBHs,

SNR>12: 4048 BNSs , 238 BHNSs, 5228 BBHs





Challenges

Beginner

- Recovery of high-SNR signals within given time windows SNR = 597, 386, 383 (BNS), 374, 343, 306

Expert

- Parameter estimation of ultra-high SNR BBH signals
- Long duration binary neutron stars
- Overlapping signals

We should organize more focused Challenges with specific goals and result comparison.

The Data

- /cvmfs/et-gw.osgstorage.org/et-gw/PUBLIC/MDC1
- Instructions here: https://wiki.et-gw.eu/EIB/SoftwareFrameworks/WebHome?validationkey=e2698d03b6eff5856cfab4654d3fbfe5
- Frame files for E1, E2, E3 and E0 (set 0: noise only, set 1: noise+GWs)
- 1300 frames per detector of length 2048s and sampling rate 8192 Hz (1.3 TB) + frames for Cosmic Explorer, CEA and CEB
- Text files with lists with the source parameters and expected SNR

ESCAPE VRE

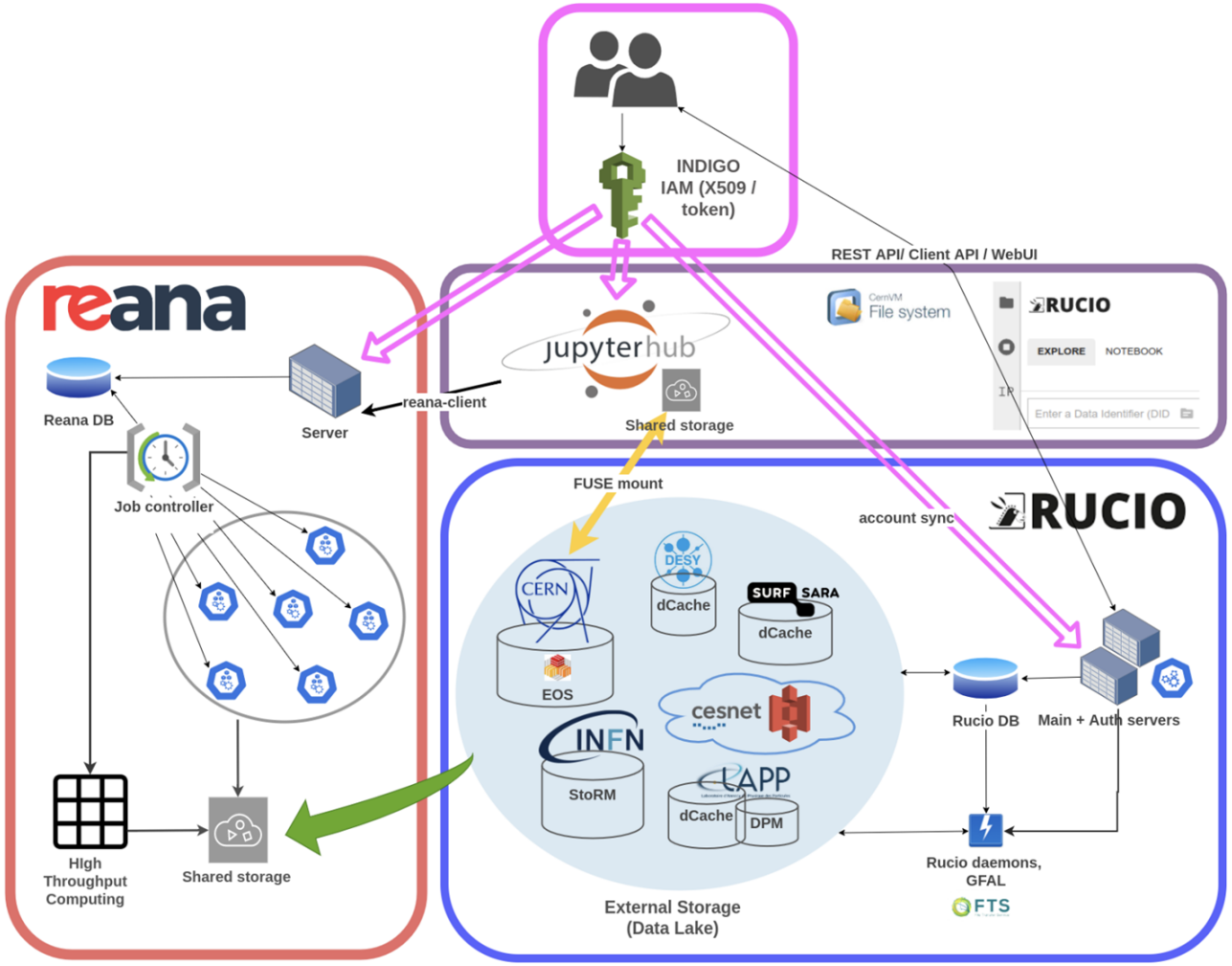
- The European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures - **Virtual Research Environment** aims at fostering collaboration and innovation within the scientific community.
- It seeks to create an open virtual research environment where researchers from diverse fields, including astronomy and particle physics, can collaborate seamlessly, share resources, and access advanced computing tools and data analysis techniques.
- This initiative represents a significant step towards enhancing research efficiency and accelerating scientific discoveries across Europe.

https://vre-hub.github.io/

The CERN-VRE (Virtual Research Environment)

[Link to demo]

https://vre-hub.github.io/



An **analysis facility**, which allows the following services to be connected:

- ESCAPE **Data Lake infrastructure** - A federated distributed storage solution (based on RUCIO).
- Computing cluster powered by **REANA** (analysis platform with special focus on reanalysis).
- Accessible via a **Jupyter** frontend that can
  - load user **software environments,**
  - use a **Notebook** service
  - access to **CVMFS** repositories.
- By a single login to a federated AAI - **ESCAPE INDIGO IAM.**

escape-cern-ops@cern.ch

Accessing the data

- The CernVM File System (CernVM-FS) provides a scalable, reliable and low-maintenance software distribution service (https://cvmfs.readthedocs.io/en/stable/)
- If installed in your environment you can easily access the data as they were on local disks

- ESCAPE AAI: Authentication and Authorization (https://indigo-iam.github.io/escape-docs/)

- ESCAPE Data Lake: consists of several storage services operated by the partner institutes and connected through reliable networks, and it adopts Rucio to orchestrate data management and organisation. (ESCAPE Data Lake - Next-generation management of cross-discipline Exabyte-scale scientific data Riccardo Di Maria, Rizart Dona and on behalf of the ESCAPE project EPJ Web Conf., 251 (2021) 02056 DOI: https://doi.org/10.1051/epjconf/202125102056)

The Virtual Research Environment

- The ESCAPE Virtual Research Environment
  - https://jhub-vre.cern.ch/

- We will find the data under the path /cvmfs/et-gw.osgstorage.org/et-gw/PUBLIC/MDC1

- You can find the notebooks at the github repo:
  - https://github.com/elenacuoco/ET-MDC-Tutorials
  - https://gitlab.et-gw.eu/osb/div10/mdc-tutorial

The MDC1 projects

- https://wiki.et-gw.eu/OSB/DataAnalysisPlatform/MDC
- https://wiki.et-gw.eu/OSB/DataAnalysisPlatform/MDC_Participants

Group	Expertise level	Brief explanation of aims	Software used	Contact person	Remarks
Utrecht University	Experts	Parameter estimation (automated classifier for telling number of overlapped signals), joint parameter estimation, Searches (template bank versus global optimisers, null stream background), Machine-learning	PyCBC, other software developed in UU...	Bhooshan Gadre, Thibear Wouters, Harsh Narola, Justin Janquart, Anuradha Samajdar,	
ICCUB	Medium	PE, searches	cWB, PyCBC	Tomas Andrade, Pablo barneo, Ruxandra Bondarescu	
University of Geneva	Beginners	CBC signals, early-inspiral regime	Not final, machine-learning related	Carlos Moreno Martinez, Sarah Baimukhametova, Steven Schramm	
IJCLab	Experts	Test existing searches based on PySTAMPAS and PyCBC; develop template banks for CBC searches	PySTAMPAS, PyCBC	Tito Dal Canton	

If you wan to join the DIV10 activities write to et-osb-da@ego-gw.it