

TeRABIT – A highway for data in gravitational wave experiments

Alessio Fiori

Virgo-ET Pisa Workshop, May 22 – 23, 2024



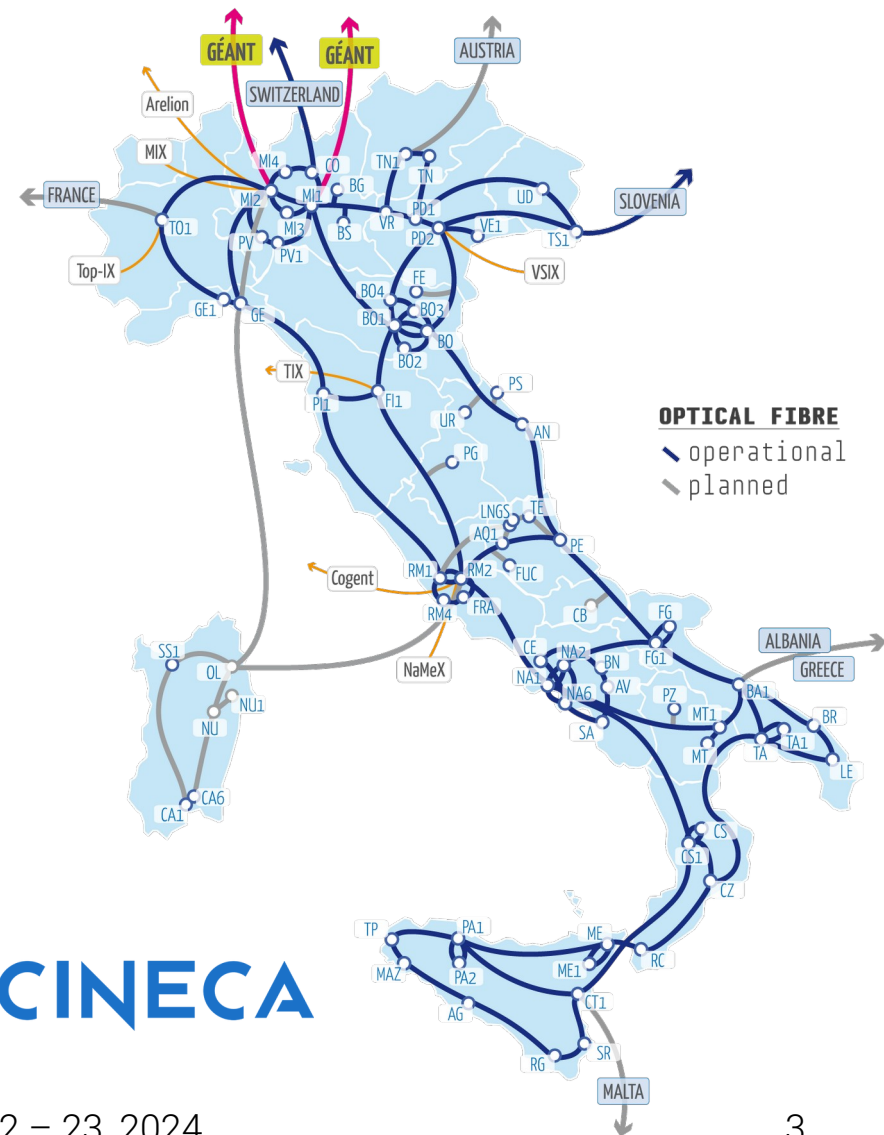
About me

- Member of LVK since 2021, ET since 2022
- Mainly involved in multi-messenger and computing
- **MS degree in physics @ DF Unipi**, Sept 2020
Supervisor: Prof. M. Razzano
Topic: Fermi-LAT pulsar variability analysis
- **PhD in Physics @ DF Unipi**, Nov 2020 – Nov 2023
Supervisor: Prof. M. Razzano
Topic: Bayesian inference for Fermi-LAT pulsars
- **AdR @ INFN Pisa**, Jun 2021 – Jun 2023
Project: Virgo-ESCAPE
Activity: IGWN low-latency alert infrastructure
- **Technologist @ INFN Pisa**, Jun 2023 – ongoing
Project: PNRR TeRABIT
Activity: Servizio calcolo e reti



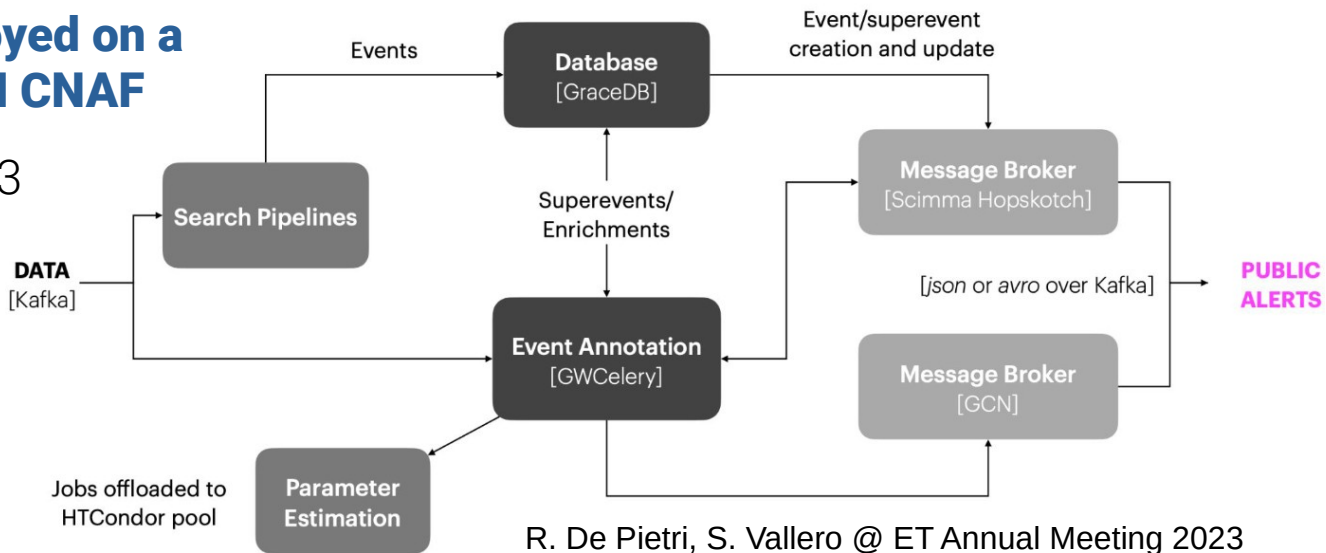
PNRR TeRABIT

- Terabit Network for Research and Academic Big data in Italy
- Integrating and enhancing Italian digital research infrastructures
- **Data transfer via optical fiber up to 1 Tb/s**
- **12M € investment to link research institutes in Sardinia with the Italian network**
- Crucial for ET data distribution from Sos Enattos



The IGWN low-latency alert infrastructure at CNAF

- Started in June 2021 within the Virgo computing group
- **Collaborators** S. Vallero (INFN Torino), R. De Pietri (INFN Parma), R. Poulton (EGO) and others
- **Goal** Testing the performance on modern virtualization technologies
- **Infrastructure services deployed on a Kubernetes provided by INFN CNAF**
- Talk by S. Vallero @ CHEP2023 ([proceedings here](#))



IGWN

The ET sites data repository

- Activity started in June 2023, based on a previous activity @ DF Unipi
- **Collaborators** M. Razzano, F. Fidecaro, E. Mazzoni (INFN Pisa), M. Di Giovanni (Sapienza), C. Giunchi (INGV)
- **Goal** Services for the distribution and analysis of Sos Enattos seismic data (Site Preparation Board)

- **Architecture**

Virtual machine @ INFN vCenter

AlmaLinux 9 OS

8 vCPU , 16 GB memory, scalable

11 TB GPFS disk

- Services deployed via docker



JupyterHub

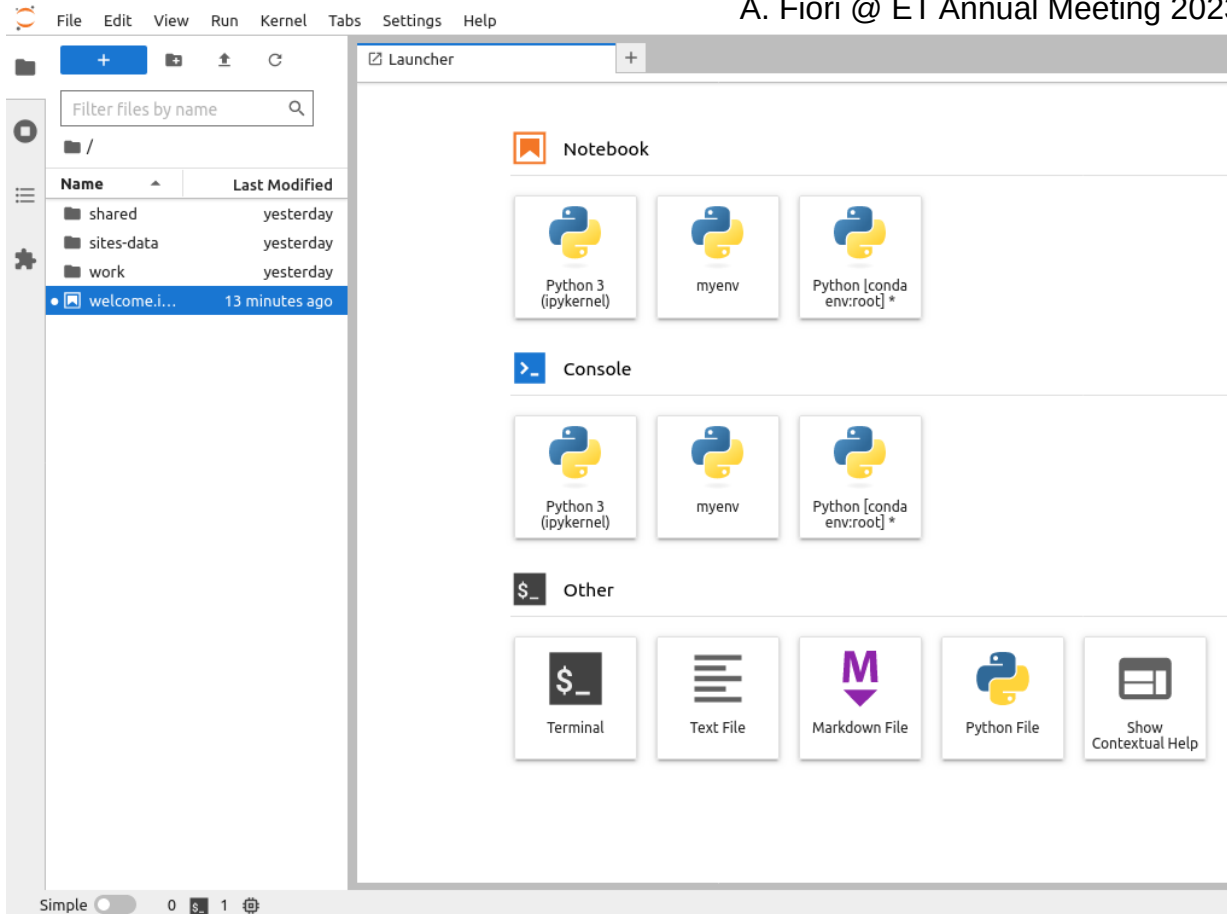


Seedlink

ET Sites Data: JupyterLab



A. Fiori @ ET Annual Meeting 2023



- Conda preinstalled
- Users can create custom envs
- Persisted user work directory
- **ObsPy 1.4.0** already installed
- **Enabled OpenSSH** server for remote login
- [Wiki here](#)

ET Sites Data: SeisComP seedlink



- Python-based modular software for data acquisition and analysis
- Custom docker image for integration in the infrastructure
- Used to transfer **Sos Enattos data from INGV Pisa** to INFN Pisa
- Data archived in **SDS directory structure**

gempa

SeisComP
development
services and trainings

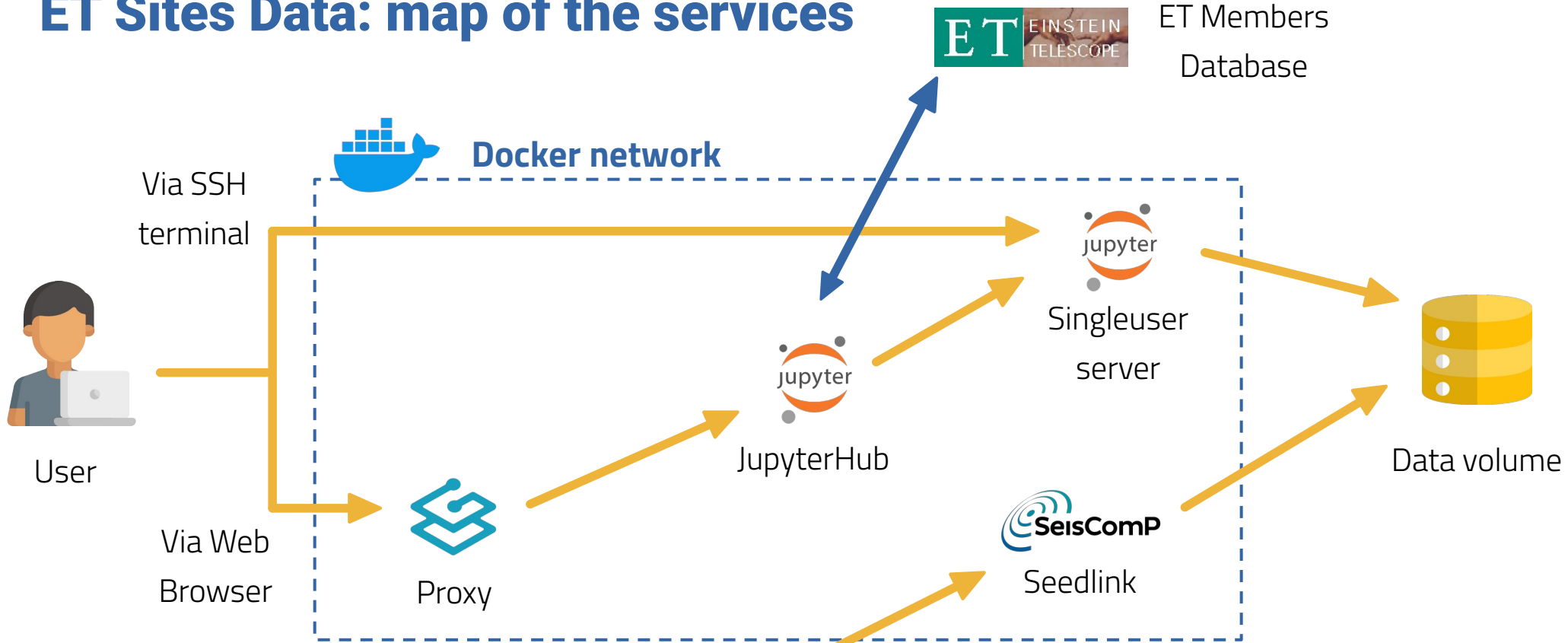
GFZ

Helmholtz-Zentrum
POTS DAM

```
archive
+ year
  + network code
    + station code
      + channel code
        + one file per day and location, e.g. NET.STA.LOC.CHAN.D.YEAR.DOY
```

- Currently about **3 TB of data** collected since 2019 (see [wiki](#))

ET Sites Data: map of the services



Adapted from
A. Fiori @ ET: Scienza e Tecnologia in Italia, 2024

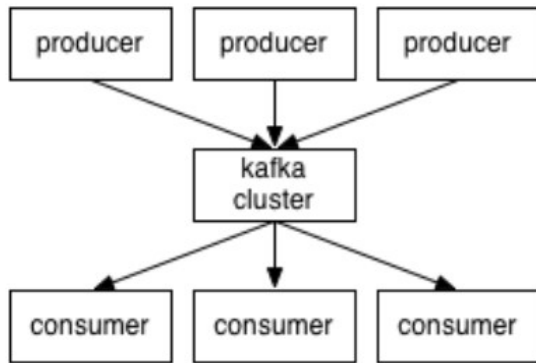


ISTITUTO NAZIONALE
DI GEOFISICA E VULCANOLOGIA

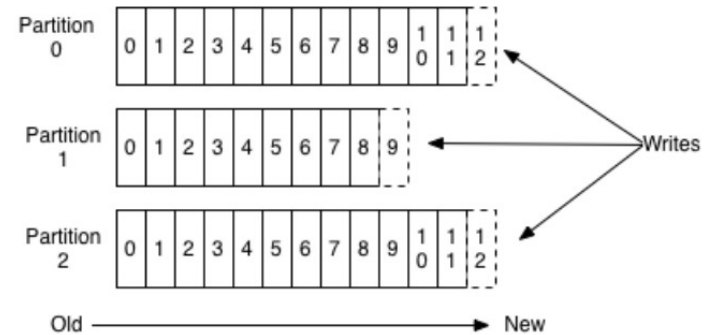
Data streaming with Apache Kafka



- Activity started in April 2024
- **Collaborators** A. Gennai, F. Laruina, E. Mazzoni, P. Prospero, F. Spada (INFN Pisa)
- **Goal** Apache Kafka cluster for low latency transfer and online processing of laboratory data



Anatomy of a Topic



1. To **publish** (write) and **subscribe to** (read) streams of events, including continuous import/export of your data from other systems.
2. To **store** streams of events durably and reliably for as long as you want.
3. To **process** streams of events as they occur or retrospectively.

Apache Kafka: status



CONFLUENT

- Kafka cluster based on Confluent Platform

- **Architecture**

3 broker/controller nodes on VM

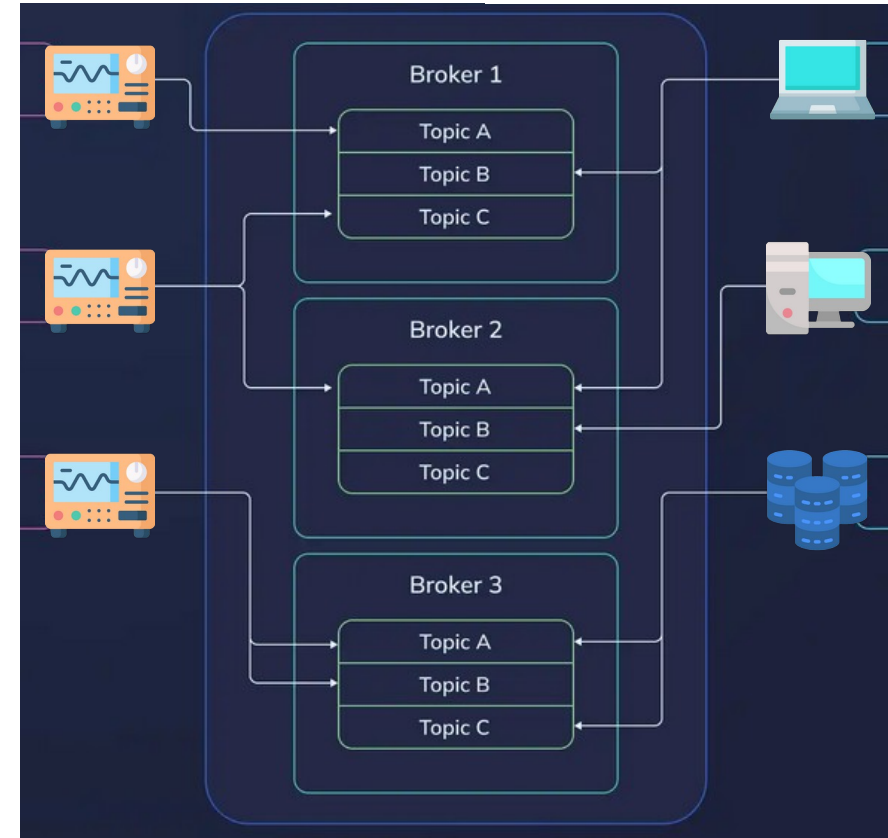
Ubuntu server 24.02

2 vCPUs, 4 GB memory, scalable

50 GB logical XFS volume, scalable

- **Test cluster available for performance tests**

- Currently working on security features



Cloud computing in scientific research

- The research is looking at Cloud as the future computing model



CINECA



- Multiple EU projects directed towards cloud computing



- Gravitational wave experiments will move in the same direction (see [this paper](#) by S. Bagnasco et al.)

OLD MAN YELLS AT CLOUD



1:st g:~no'peri
no:) kf o: ukéve lf:
je: i aud r:ia:em,;
m: istj: / K:u:;
d' hdy:ivr 's:é j:ú
sG: ~ veme us:~: or a
t:shy. ne. r:asow: a
1:st:~:
1:2'
m:~:
~:
1:~' na "0 T:~)

LEGACY INFRASTRUCTURE USER BE LIKE: