SDL – The Simulation Data Lake for managing complex, very large and multi-domain data in Geosciences

Gabriella Scipione (CINECA)

2 September 2024

***Geo-INQUIRE** is funded by the European Union. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.*

CINECA: 50 YEARS OF SUPERCOMPUTERS

TIMELINE OF CINECA’S SUPERCOMPUTERS

1969	1975	1985	1989	1993	1994	1995	1996	1998	1999
CDC 6600	CDC 7600	CRAY X-MP / 48	CRAY Y-MP / 48 IBM 3090-600 VF	CRAY C-90/2128	CRAY T30 64	CRAY T3D 128	CRAY T3D 128	CRAY T3E 256 SGI ORIGIN 16	CRAY T3E 256 SGI ORIGIN 64 IBM SP POWER 3

{ 1° SC IN ITALY }{ 1° VECTOR SC }

{ 1° PARALLEL SYSTEM }

{ 1° MPP SYSTEM }

{ 1 TERAFLOP/S }

{ 10 TERAFLOP/S }{ 100 TERAFLOP/S }{ 2 PETAFLOP/S }

{ PRE-EXASCALE }

2000	2002	2004	2005	2007	2009	2012	2016	2020	2022
VIRTUAL THEATRE	IBM SP4 612	IBM SP5 512	IBM CLX XEON 1024	IBM BCX AMD 5120	IBM SP6	FERMI	MARCONI	MARCONI100	LEONARDO

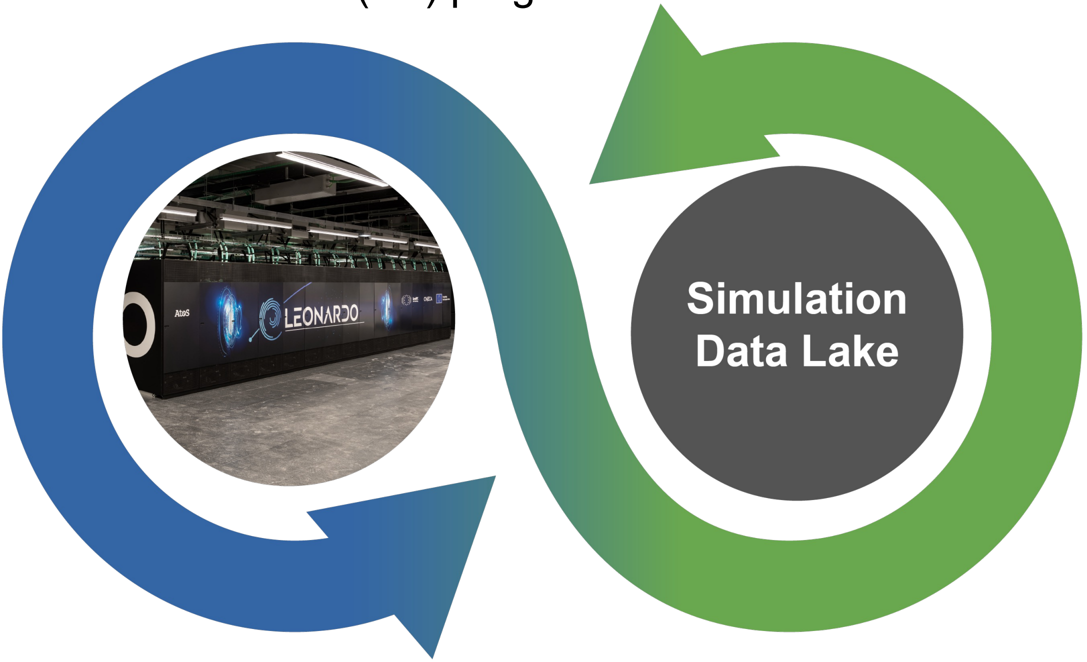
CINECA for Geo-INQUIRE

- focused on advancing scientific knowledge in computational seismology, volcanology, tsunami science, and geo-hazard analysis.

- to improve access to cutting-edge research infrastructures, software, and data through Virtual Access (VA) and Transnational Access (TA) programs

CINECA has twofold role:

HPC support
Provide TA at CINECA HPC infrastructures.



HPC Data Management
Provide a data lake for the storage of input/output data for/from simulations,

LEONARDO: THE FLAGSHIP SYSTEM

EuroHPC Joint Undertaking

#7 TOP500 The List.

PERFORMANCE

10 Exaflops of AI performance

240 Petaflops HPL (scientific)

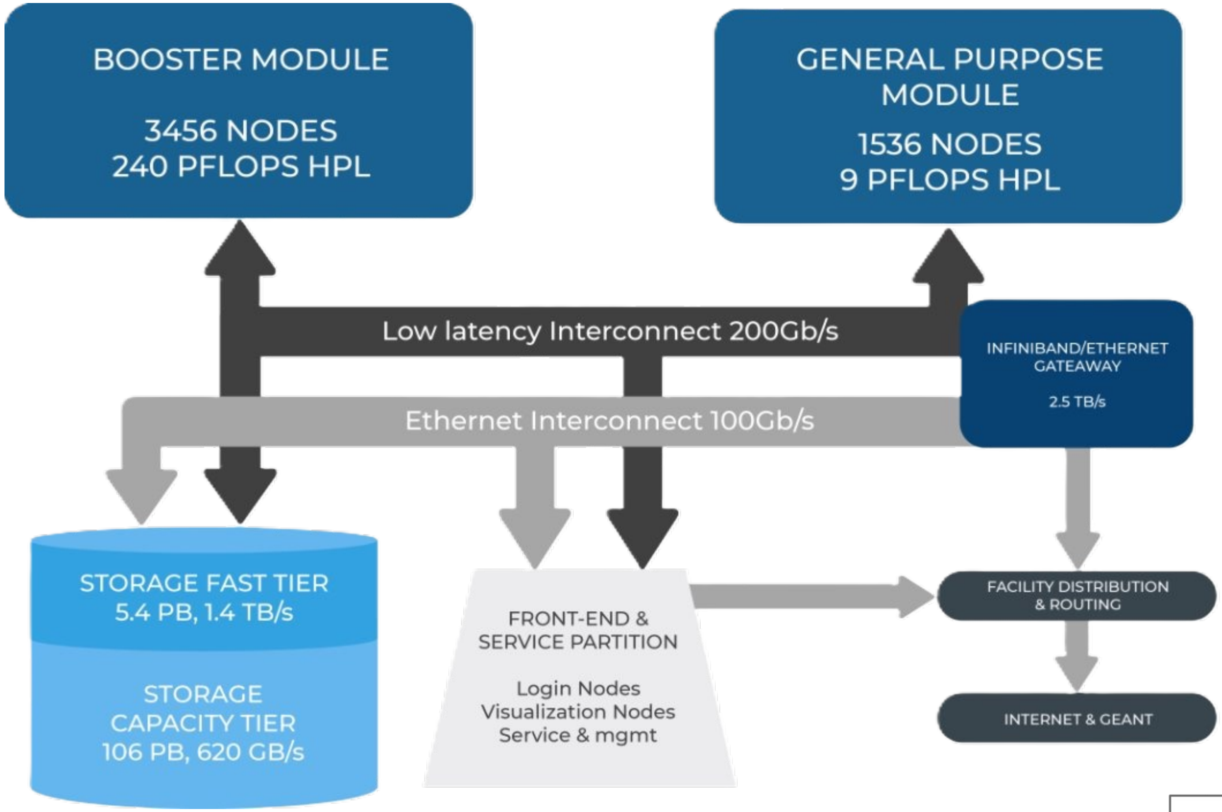
Training ChatGPT 3.5 in 1 day

ARCHITECTURE

14.000 GPUs

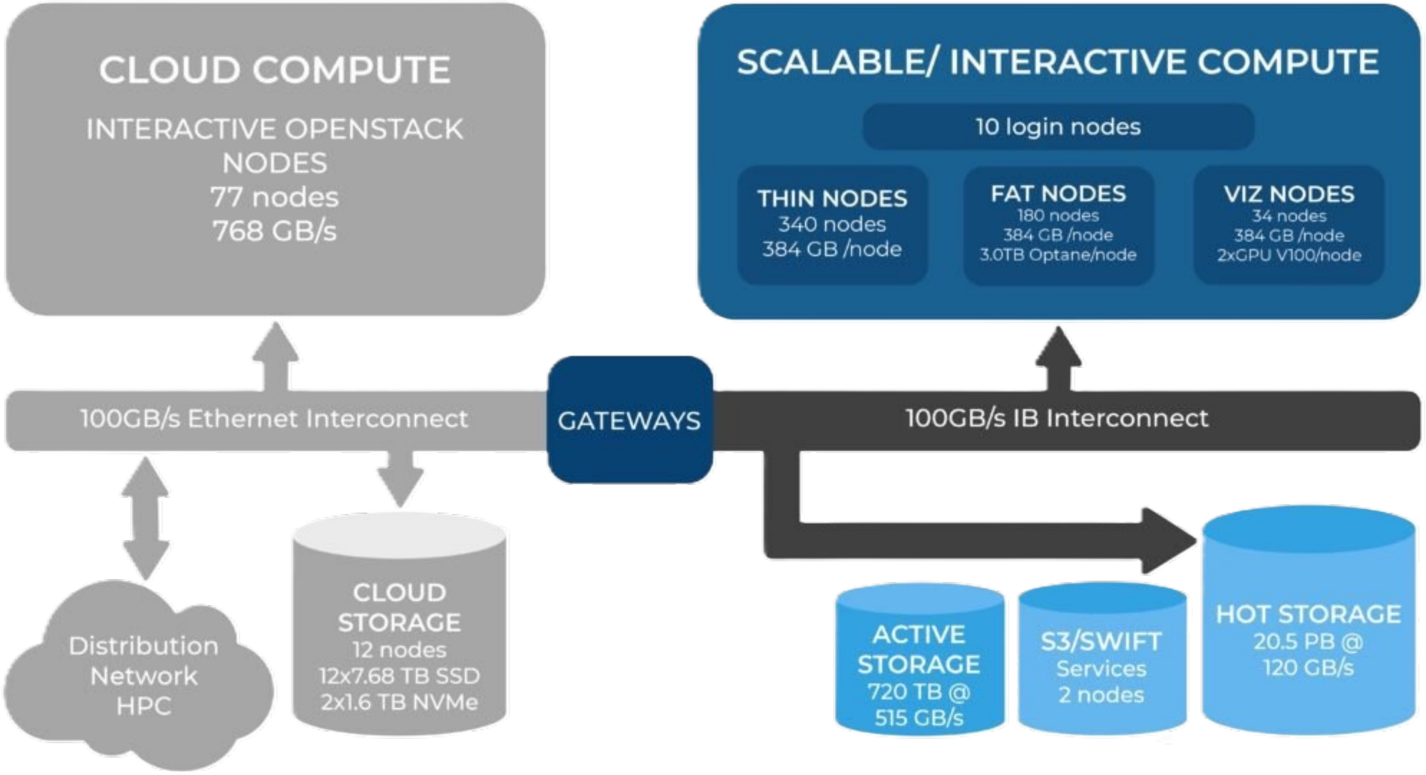
5000 compute nodes

1000 switches (>150km optical fibers)



GALILEO100: an HPC and CLOUD system

- 564 cluster nodes,
-data analysis, interactive computing, HPC



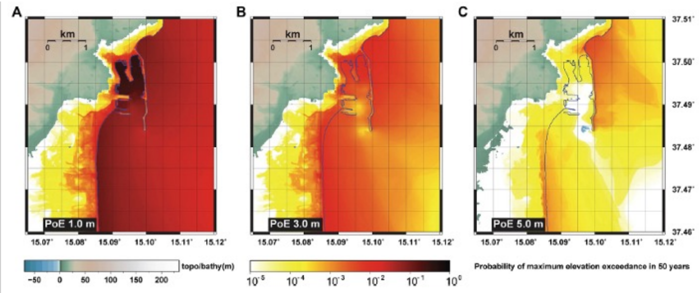
Simulation Data Lake requirements

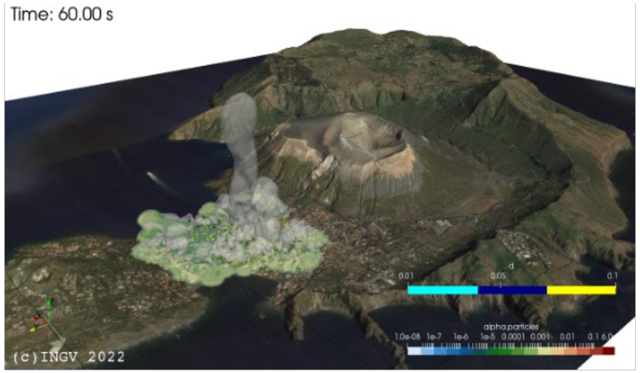
VALUE

- Researchers need a place to store and preserve simulation data for the long term.
- Aim: Setting up a multi-domain **SIMULATION** data lake in Geosciences

VARIETY

- What Data: **Innovative type of data** --> Simulation data
- Manage **complex, very large and multi-domain data in Geosciences**
  - Seismology:
  - Volcanology
  - Tsunami science
  - Geohazard analysis
- Heterogeneous structure and data formats





VOLUME

- Include inputs, outputs and WF descriptions
  - Up to tens of TB (1 experiment).
  - Up to 10K – 100K files.
- Users: open to all
  - Sharing simulation data with the community, to be reused
  - Open to all from September 2025

TECHNICAL REQUIREMENTS

- Maximum **500T** of storage

Simulation Data Lake Challenges

- Definition of the **Data Model**
- Interoperability through the **adoption of a common metadata schema**
- **Proximity of the data to the HPC clusters** for
  - re-use
  - postprocessing
  - comparison with new studies
  - training of AI algorithms

- Synergies with other significant Solid Earth European projects and other initiatives:
  - **EPOS** research infrastructure that facilitates the use of data, data products, and facilities from the solid Earth science community in Europe.(https://www.epos-eu.org/)
  - **DT-GEO** Digital Twin for GEOphysical extremes (https://dtgeo.eu/)
  - **ChEESE** Centre of Excellence (CoE) for Exascale in Solid Earth (https://cheese-coe.eu/)
  - **Destination Earth** (https://destination-earth.eu/)
  - **M@TE** Model Atlas of the Earth (https://mate.science/)
  - **AUScope** Australian Geophysical Observing System (https://www.auscope.org.au/)
  - ....

Simulation Data Lake Dataset (=Experiment)

EXPERIMENT

An experiment is made of 1-to-100K simulation runs
Up to tens of TB

SIMULATION

DATASETS

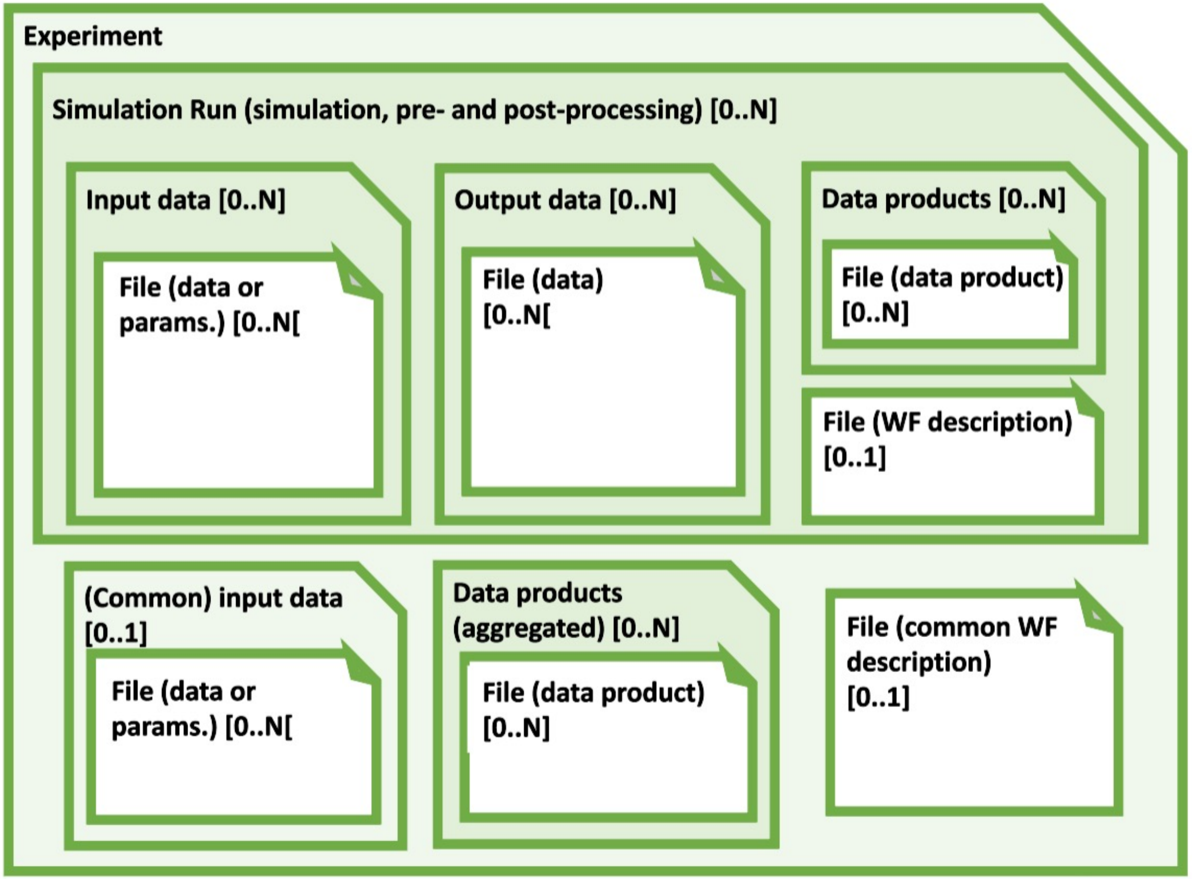
Input data:

- File of parameters (common to all or specific to each simulation run).
- Input data (common to all or specific to each simulation run).

Output data:

- Output data for each simulation run (e.g., netcdf, hdf5).
- Postprocessing
- Scripts
- Data products
- Figures

Workflow description

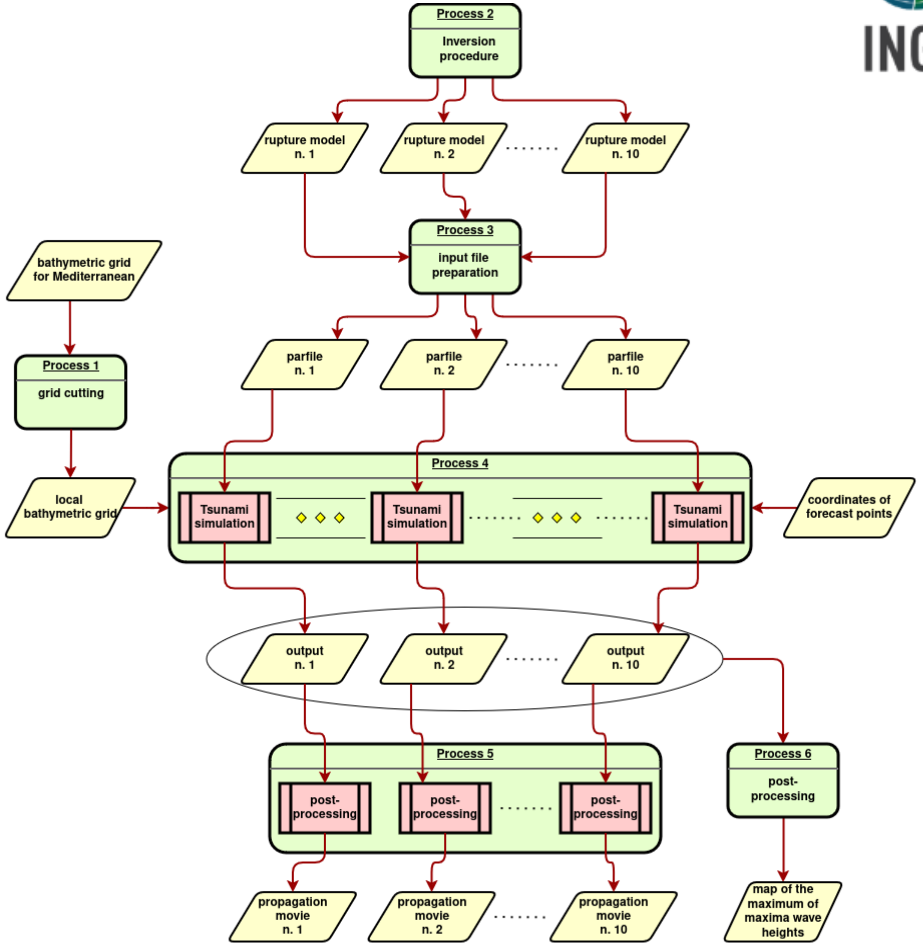

Experiment
Simulation Run (simulation, pre- and post-processing) [0..N]
Input data [0..N]
File (data or params.) [0..N[
Output data [0..N]
File (data) [0..N[
Data products [0..N]
File (data product) [0..N]
File (WF description) [0..1]
(Common) input data [0..1]
File (data or params.) [0..N[
Data products (aggregated) [0..N]
File (data product) [0..N]
File (common WF description) [0..1]

(Common) DATASETS

- Input
- Data product

Experiment example:

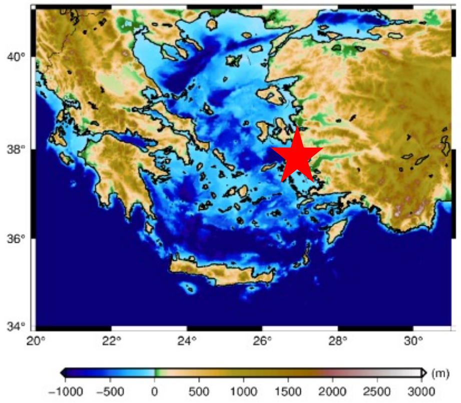
- **Tsunami simulations from INGV**: Experiment_Samos
- The case study is the 30 October 2020 Mw 7.0 Samos earthquake, the largest seismic event in the eastern Aegean Sea.
- 40000 scenarios (simulation runs). A folder for each scenario



Experiment example

- Input is grouped in a folder: **Only the parameter file is specific for each simulation, the other input files are common to all of them.**
- Output: for each simulation two netCDF files are produced
- Subfolders need to be compressed.
- JSON metadata (EUDAT Core Schema)
- User needs to download sim_setup and any single simulation to be able to reproduce it.
- Input / output / Common input

```
Experiment_Samos
├── BS_scenario00001
│   ├── out_ts.nc
│   ├── out_ts_ptf.nc
│   └── parfile.txt
├── BS_scenario00002
│   ├── out_ts.nc
│   ├── out_ts_ptf.nc
│   └── parfile.txt
├── BS_scenario00003
│   ├── out_ts.nc
│   ├── out_ts_ptf.nc
│   └── parfile.txt
├── BS_scenario00004
│   ├── out_ts.nc
│   ├── out_ts_ptf.nc
│   └── parfile.txt
...
├── BS_scenario40000
│   ├── out_ts.nc
│   ├── out_ts_ptf.nc
│   └── parfile.txt
└── sim_setup
    ├── Step1_scenario_list_BS.txt
    ├── Step2_local_domain_2020_1030_samos.grd
    ├── Step2_local_domain_2020_1030_samos_POIs_depth.dat
    └── Step2_ts.dat
```



E. Volpe et al. INVG

FAIRness principles

Findable

Data needs to be **findable at different levels**

- Whole datasets
- Simulation runs
- Single files

Enable to query the SDL in useful ways. Metadata needs to include specific information.

Accessible

Data **accessible at different levels**:

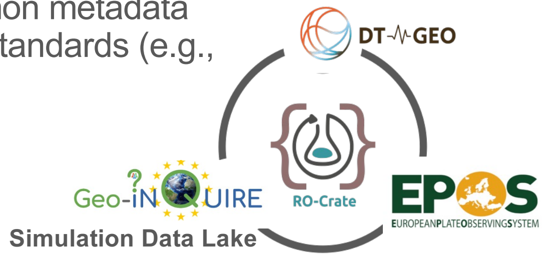
- Whole datasets
- Simulation runs
- Single files

API or/and GUI to access data

Interoperable

Adoption of standards used by the community

Synergies with **DT-GEO** and **EPOS**. Adoption of a common metadata schema and other standards (e.g., **RO-Crate**).



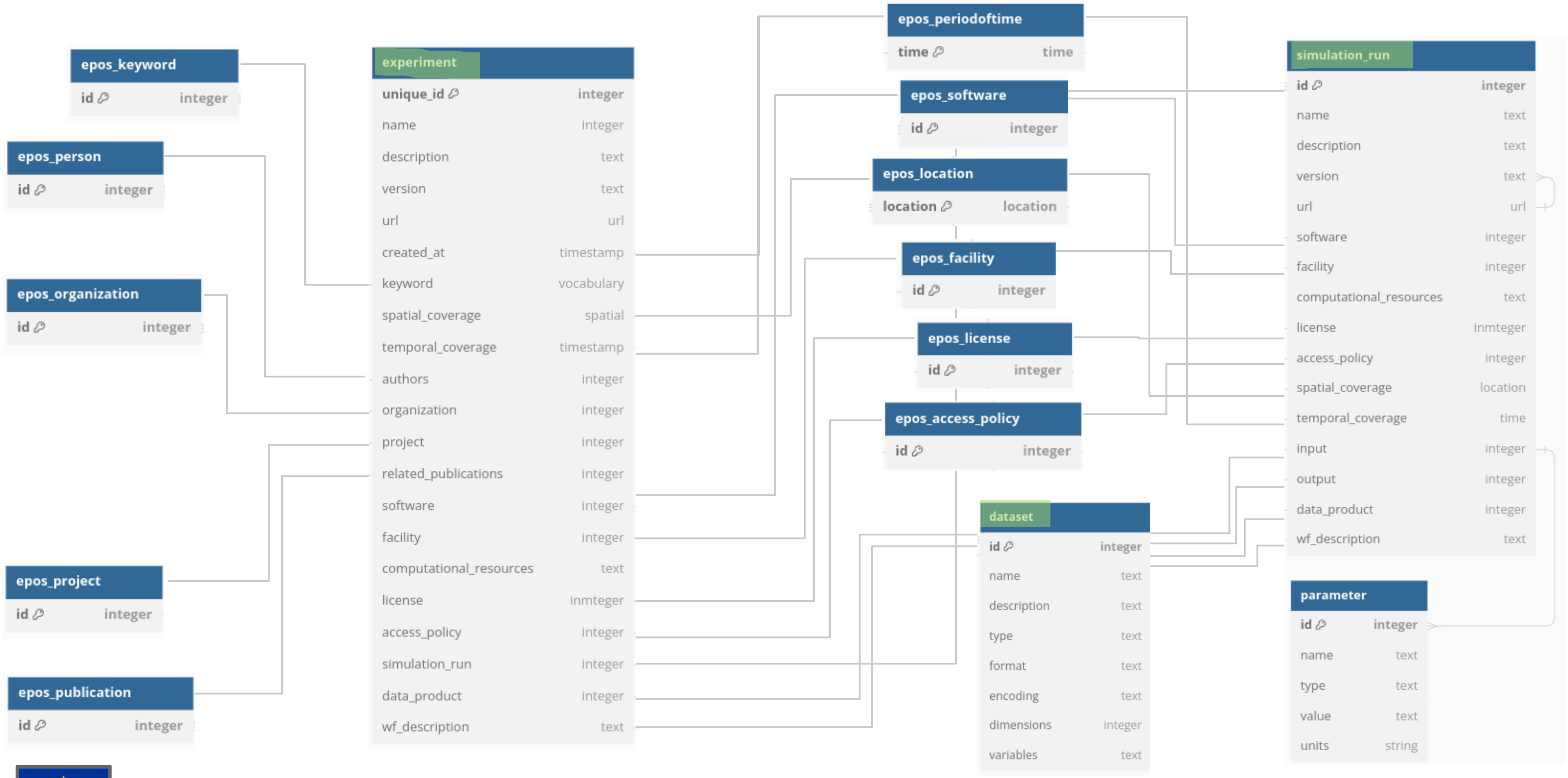
Reusable

Datasets should include **all the data** and the information needed to **re-run simulations**

Datasets need to be well characterised

Geo-INQUIRE Metadata Schema:

EPOS-DCAT-AP is a dialect of DCAT-AP



epos_keyword	
id	integer

epos_person	
id	integer

epos_organization	
id	integer

epos_project	
id	integer

epos_publication	
id	integer

experiment	
unique_id	integer
name	integer
description	text
version	text
url	url
created_at	timestamp
keyword	vocabulary
spatial_coverage	spatial
temporal_coverage	timestamp
authors	integer
organization	integer
project	integer
related_publications	integer
software	integer
facility	integer
computational_resources	text
license	inmteger
access_policy	integer
simulation_run	integer
data_product	integer
wf_description	text

epos_periodoftime	
time	time

epos_software	
id	integer

epos_location	
location	location

epos_facility	
id	integer

epos_license	
id	integer

epos_access_policy	
id	integer

dataset	
id	integer
name	text
description	text
type	text
format	text
encoding	text
dimensions	integer
variables	text

simulation_run	
id	integer
name	text
description	text
version	text
url	url
software	integer
facility	integer
computational_resources	text
license	inmteger
access_policy	integer
spatial_coverage	location
temporal_coverage	time
input	integer
output	integer
data_product	integer
wf_description	text

parameter	
id	integer
name	text
type	text
value	text
units	string

Simulation Data Lake functionalities

- Store/access simulation datasets, promoting data discoverability, reuse, and experiment repeatability
- Functionalities:

Upload and publish datasets: transfer large data to CINECA infrastructures

PID assignment

Metadata assignment compatible with EPOS-DCAT-AP

Search datasets and/or **files**

Download datasets and/or individual files

Data extraction from individual files (~processing)

OGC (Open Geospatial Consortium) Services Integration

EuroHPC Joint Undertaking

Visible and **accessible** from **CINECA HPC** infrastructures

Simulation Data Lake Access

API

- Interact with the system programmatically
- Integration with other services
- For technical users

CLI Command Line Interface

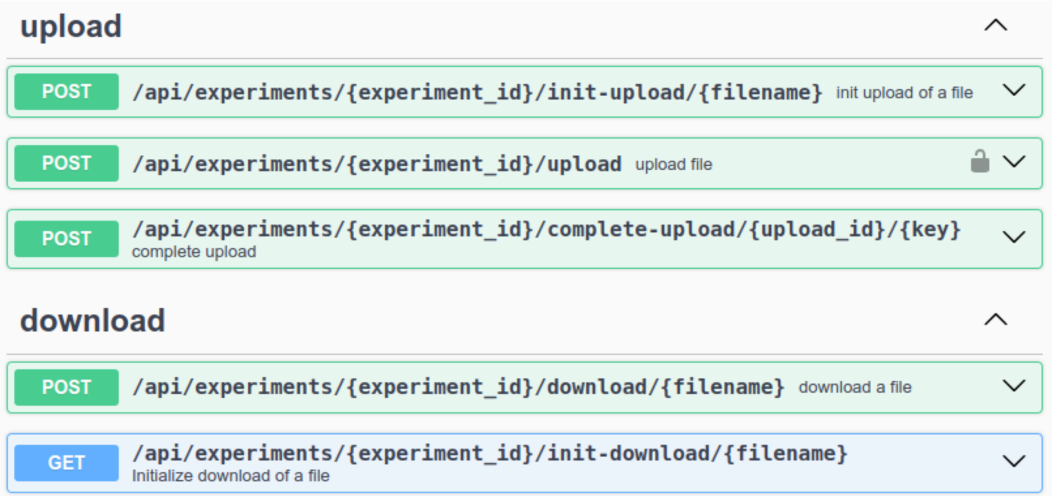
- Ideal for creating, uploading large quantities of experiments, simulations and datasets
- For technical users

WEB PORTAL

- User friendly interface
- Simple access to data
- For non-technical users

API specs

OpenAPI Specification
https://sdl.hpc.cineca.it:7777/



upload

POST /api/experiments/{experiment_id}/init-upload/{filename} init upload of a file

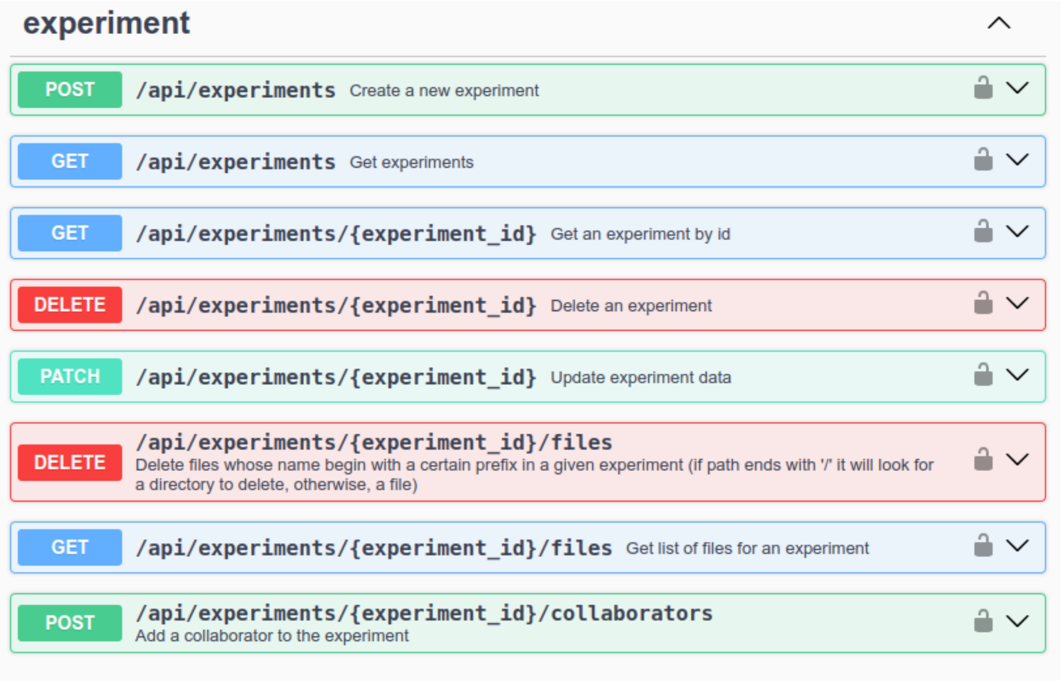
POST /api/experiments/{experiment_id}/upload upload file

POST /api/experiments/{experiment_id}/complete-upload/{upload_id}/{key}
complete upload

download

POST /api/experiments/{experiment_id}/download/{filename} download a file

GET /api/experiments/{experiment_id}/init-download/{filename}
Initialize download of a file



experiment

POST /api/experiments Create a new experiment

GET /api/experiments Get experiments

GET /api/experiments/{experiment_id} Get an experiment by id

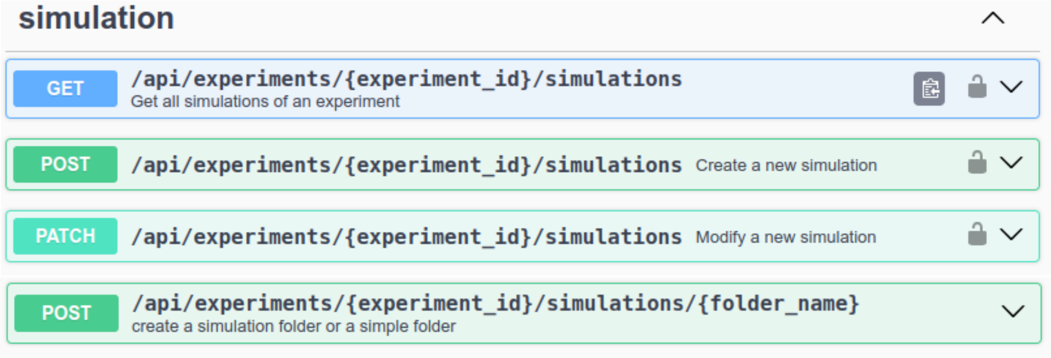
DELETE /api/experiments/{experiment_id} Delete an experiment

PATCH /api/experiments/{experiment_id} Update experiment data

DELETE /api/experiments/{experiment_id}/files
Delete files whose name begin with a certain prefix in a given experiment (if path ends with '/' it will look for a directory to delete, otherwise, a file)

GET /api/experiments/{experiment_id}/files Get list of files for an experiment

POST /api/experiments/{experiment_id}/collaborators
Add a collaborator to the experiment



simulation

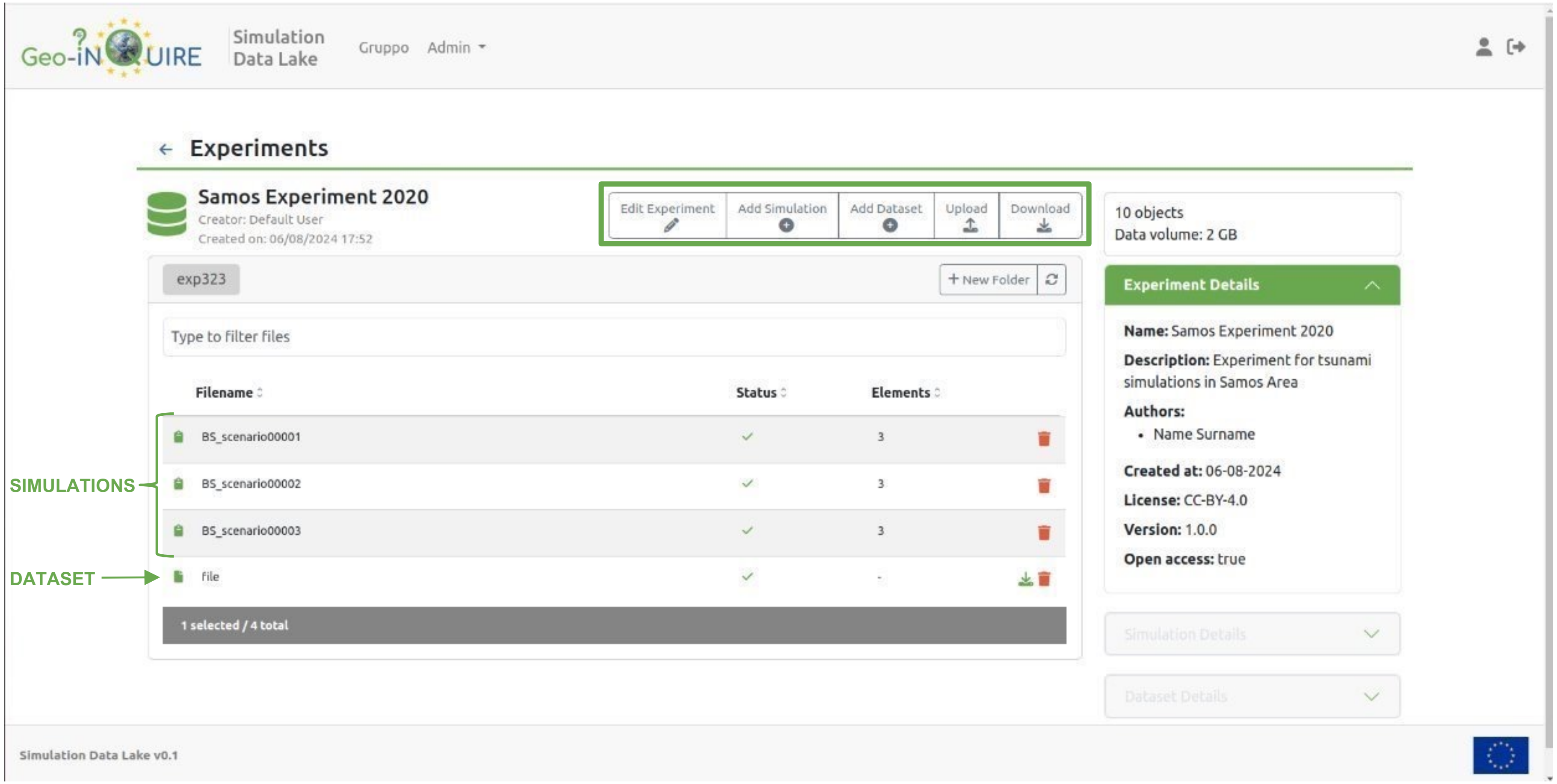
GET /api/experiments/{experiment_id}/simulations
Get all simulations of an experiment

POST /api/experiments/{experiment_id}/simulations Create a new simulation

PATCH /api/experiments/{experiment_id}/simulations Modify a new simulation

POST /api/experiments/{experiment_id}/simulations/{folder_name}
create a simulation folder or a simple folder

Web Portal: Experiment view



Geo-INQUIRE Simulation Data Lake Gruppo Admin

← Experiments

Samos Experiment 2020

Creator: Default User

Created on: 06/08/2024 17:52

10 objects

Data volume: 2 GB

exp323

Type to filter files

Filename	Status	Elements
BS_scenario00001	✓	3
BS_scenario00002	✓	3
BS_scenario00003	✓	3
file	✓	-

1 selected / 4 total

Experiment Details

Name: Samos Experiment 2020

Description: Experiment for tsunami simulations in Samos Area

Authors:

- Name Surname

Created at: 06-08-2024

License: CC-BY-4.0

Version: 1.0.0

Open access: true

Simulation Details

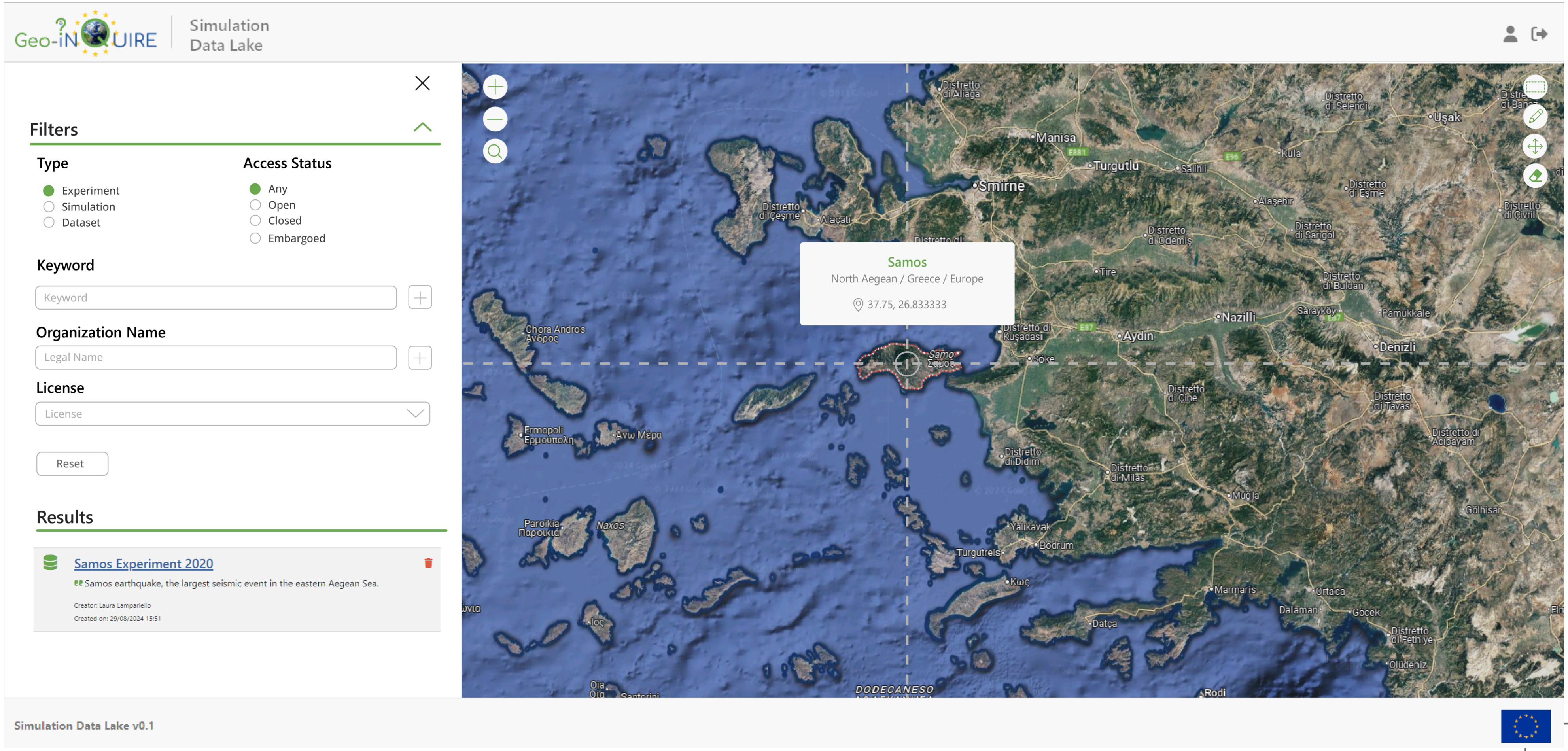
Dataset Details

Simulation Data Lake v0.1

SIMULATIONS

DATASET

Web Portal: Filtering by location



Filters

Type

- ● Experiment
- ○ Simulation
- ○ Dataset

Access Status

- ● Any
- ○ Open
- ○ Closed
- ○ Embargoed

Keyword

Keyword

Organization Name

Legal Name

License

License

Reset

Results

Samos Experiment 2020

Samos earthquake, the largest seismic event in the eastern Aegean Sea.

Creator: Laura Lampariello

Created on: 29/08/2024 15:51

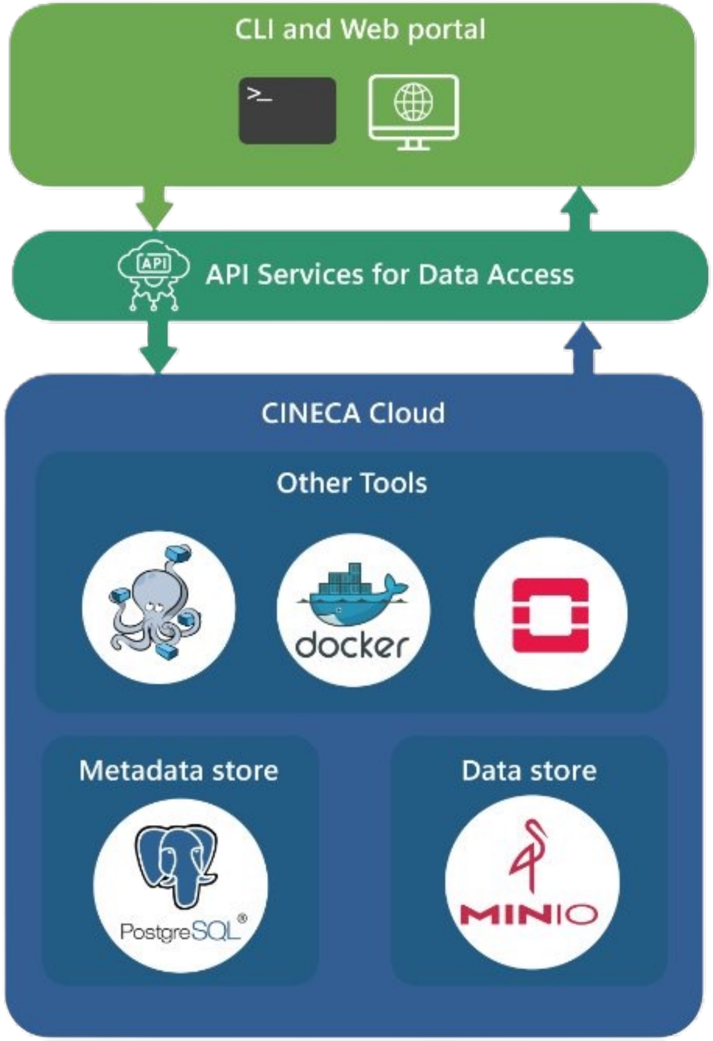
Samos
North Aegean / Greece / Europe
37.75, 26.833333

Simulation Data Lake Architecture

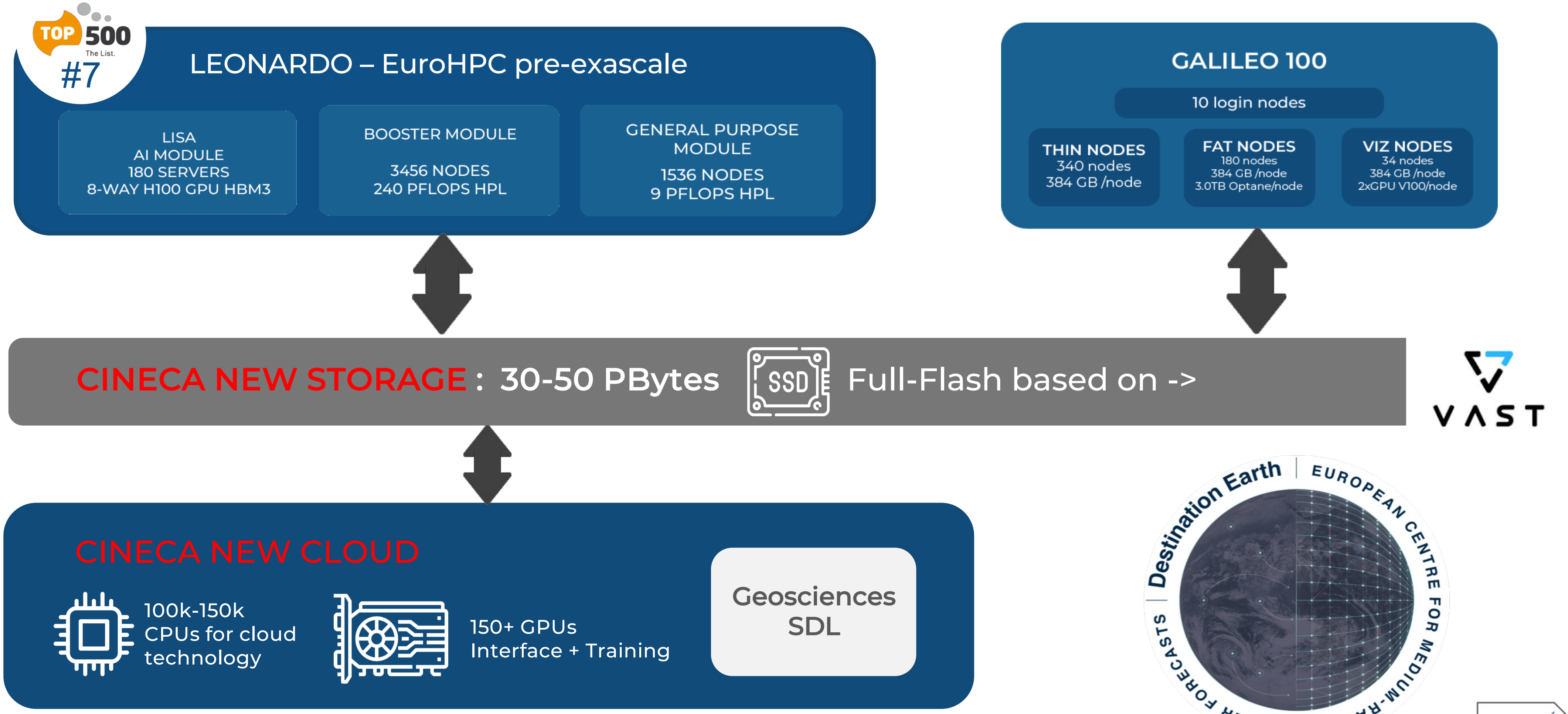
Modular architecture and Object Storage.

Hosted on **CINECA's Cloud Infrastructure.**

- MinIO (will be replaced by VAST Data)
  - Object storage
  - Compatible with S3
- PostgreSQL
  - For storing metadata
  - For data search
- Other tools
  - Docker
  - OpenStack



The Data close to the HPC resources



TOP 500 The List. #7

LEONARDO – EuroHPC pre-exascale

LISA
AI MODULE
180 SERVERS
8-WAY H100 GPU HBM3

BOOSTER MODULE
3456 NODES
240 PFLOPS HPL

GENERAL PURPOSE MODULE
1536 NODES
9 PFLOPS HPL

GALILEO 100

10 login nodes

THIN NODES
340 nodes
384 GB /node

FAT NODES
180 nodes
384 GB /node
3.0TB Optane/node

VIZ NODES
34 nodes
384 GB /node
2xGPU V100/node

CINECA NEW STORAGE : 30-50 PBytes SSD Full-Flash based on -> VAST

CINECA NEW CLOUD

100k-150k CPUs for cloud technology

150+ GPUs Interface + Training

Geosciences SDL

A Digital Twin Use Case from INGV/NGI/UMA – Tsunami inundation

HPC

- An ensemble of 28.000 Tsunami Simulations
- On Marconi GPU-accelarated HPC system of CINECA (resources obtained through a PRACE call)

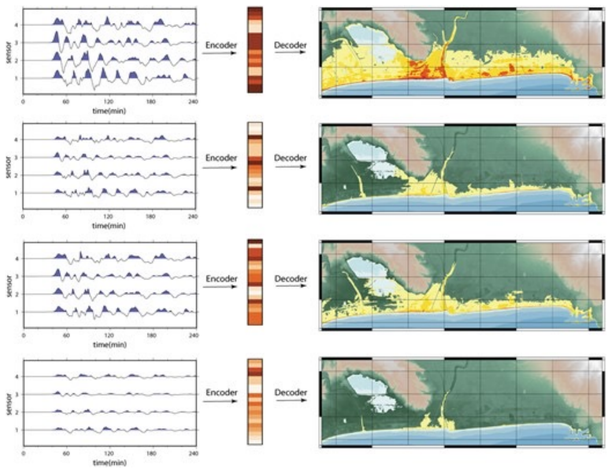
SDL

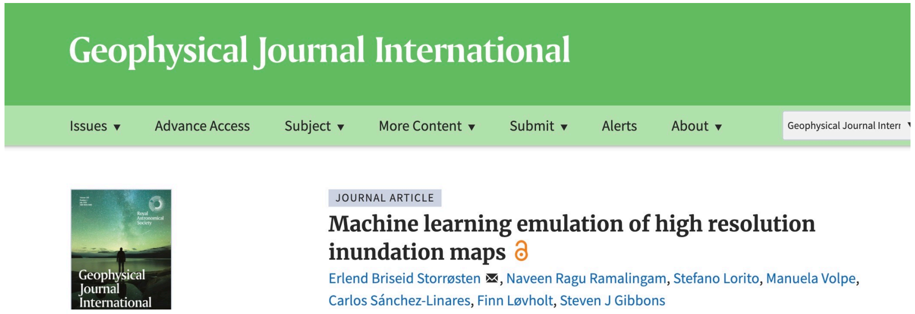
- The dataset stored in the Simulation Dataset to be reused

For AI Training & For Urgent Computing

HPC Cloud

- estimating coastal tsunami impact for early-warning or long-term hazard analysis using the EQ scenarios
- computational demanding in the context of Tsunami Early Warning where strict time constraints apply
- using a ML emulation that predicts inundation maps, trained on full simulations





Geophysical Journal International

Issues ▾ Advance Access Subject ▾ More Content ▾ Submit ▾ Alerts About ▾ Geophysical Journal Intern

JOURNAL ARTICLE

Machine learning emulation of high resolution inundation maps

Erlend Briseid Storrøsten, Naveen Ragu Ramalingam, Stefano Lorito, Manuela Volpe, Carlos Sánchez-Linares, Finn Løvholt, Steven J Gibbons

Simulation Data Lake Status & future prospects

- The SDL just been released! https://sdl.hpc.cineca.it/
- **Under testing** and use of the Geo-Inquire, Cheese and DT-Geo communities
- Stay tuned for the upcoming publication on the SDL!
- Improving the functionalities of the SDL:
  - Enhanced discoverability capabilities
  - Further metadata information
  - Moving towards interoperability

- It will be opened to all the users : **September 2025**
- **Contacts**: sdl@cineca.it
- Training sessions
- Strategic role across multiple projects, by ensuring coordination and alignment at the inter-project level.

Thank you for your attention!
Gabriella Scipione
g.scipione@cineca.it