



Centro Nazionale di Ricerca in HPC,
Big Data and Quantum Computing

Advancing Fundamental Research and Space Economy in Italy through the ICSC large-scale national initiative

Tommaso Diotallevi (University of Bologna / INFN)
Daniele Bonacorsi (University of Bologna / INFN)
Tommaso Boccali (INFN)
Sandra Malvezzi (INFN)



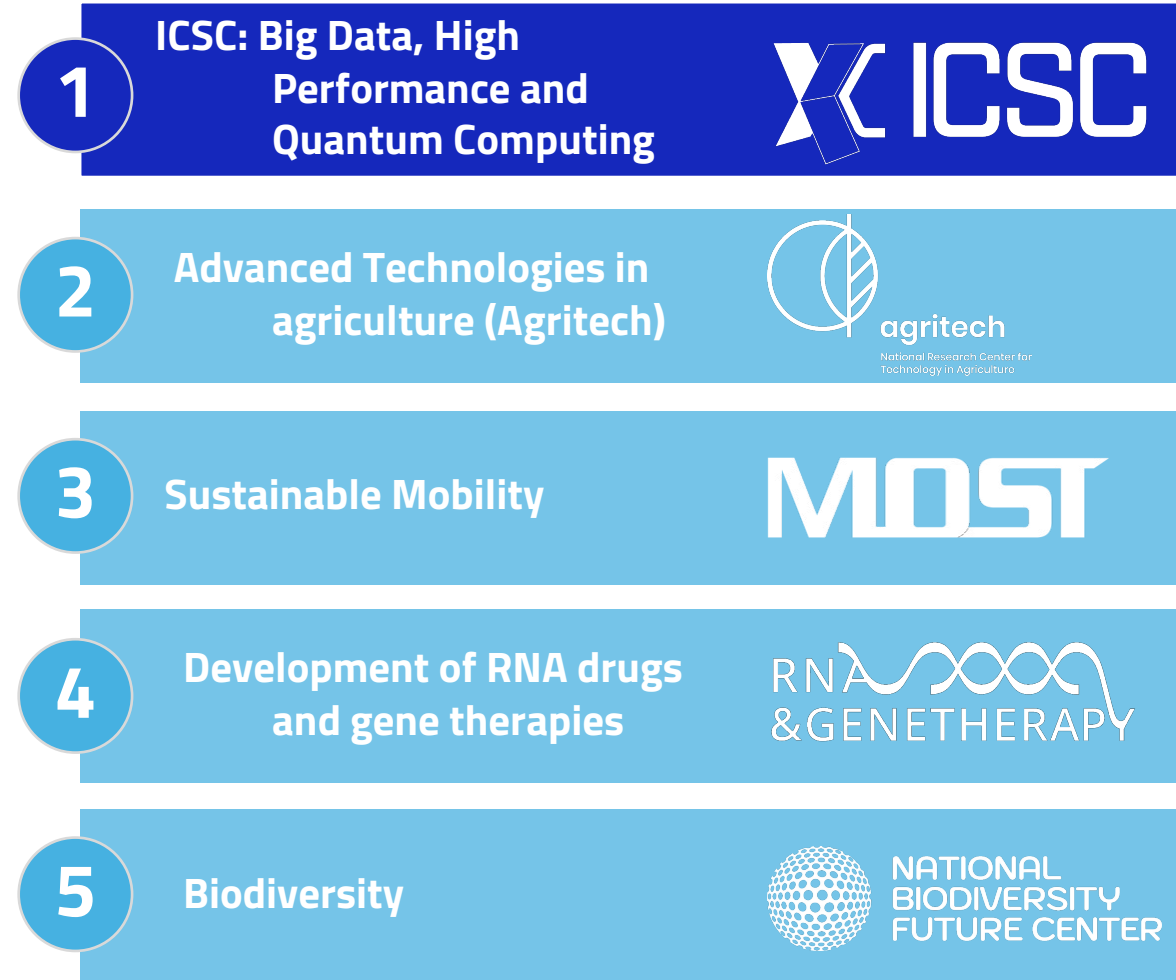
International Symposium on Grids and Clouds (ISGC2026)

18th March 2026

ICSC: The National Center for HPC, Big Data and Quantum Computing

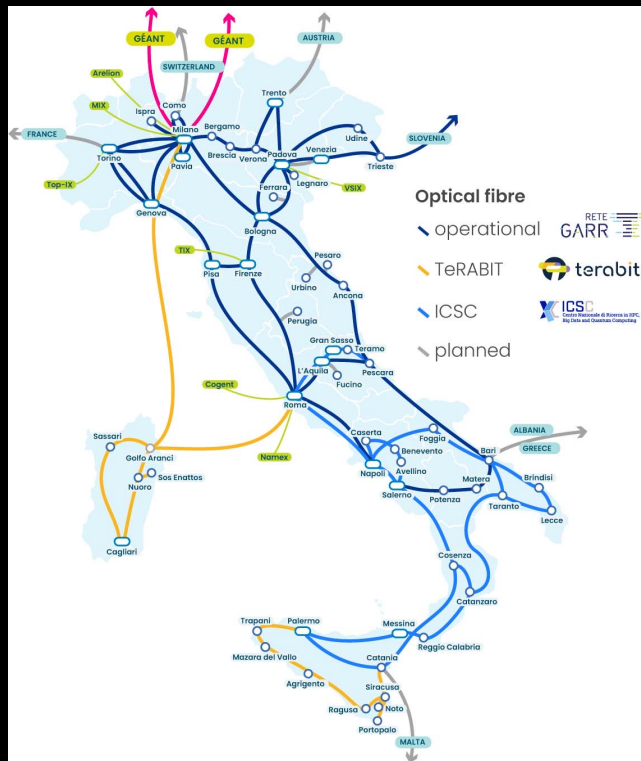
what is it?

- Italy has funded, with NRRP (pandemic recovery) funds, **5 large National Centers**, for a total of **1.6B€** over 3 years, on key future technologies.
- One of them, coordinated by **INFN**, focuses on modern IT technologies, with the final goal of deploying a long-term distributed infrastructure (>> 3y) for national research and industrial development.
- The project started on September 2022, lasting until April 2026.



On the infrastructure: three major pillars

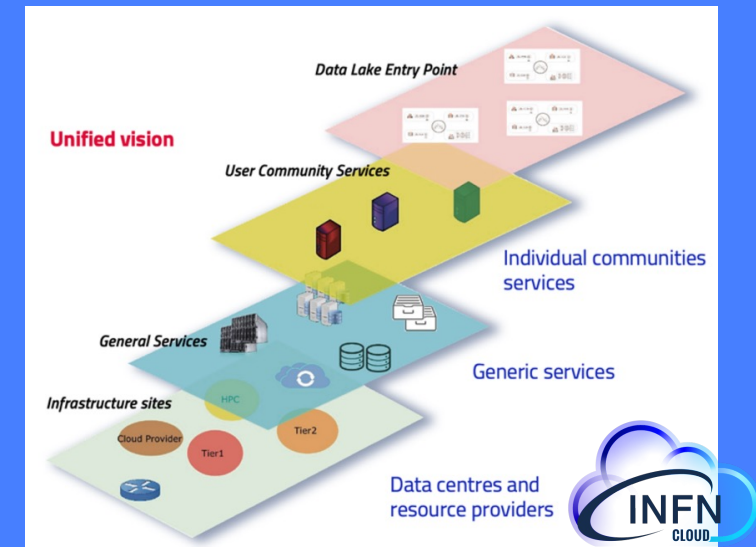
GARR: Tbps-level connectivity between all public data centers.



CINECA: expansion of Leonardo (HPC#8 on top500.org) with Lisa, and deployment of one of the European AI Factories.

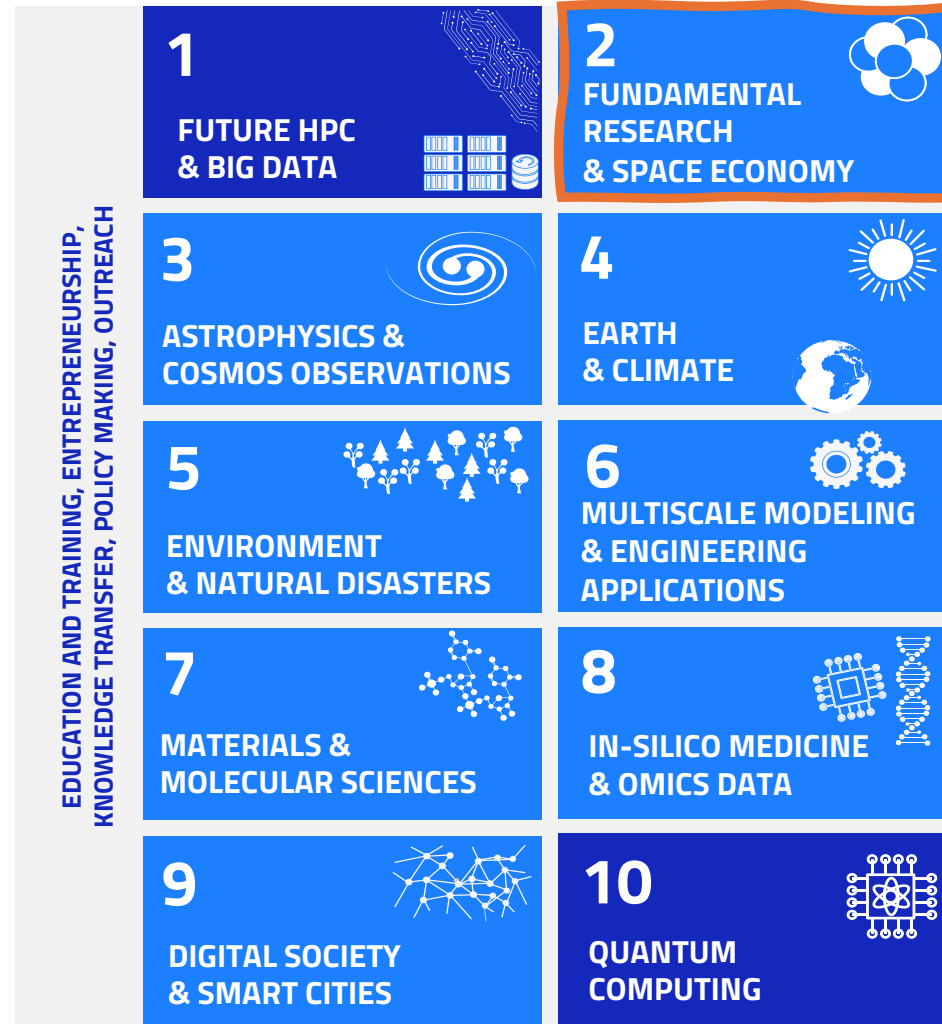
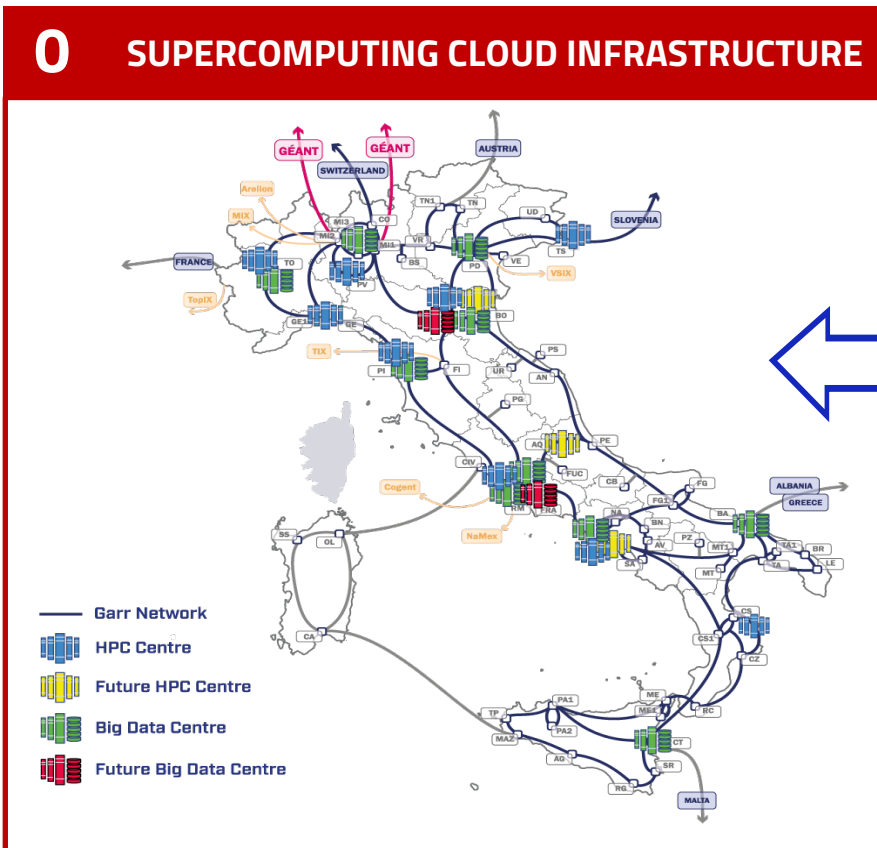


INFN: strengthening of the WLCG infrastructure (1 Tier-1 & 9 Tier-2); acquisition of Cloud resources; implementation of the datalake middleware, based on INFN Cloud.



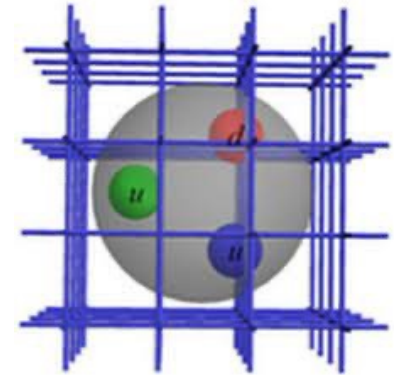
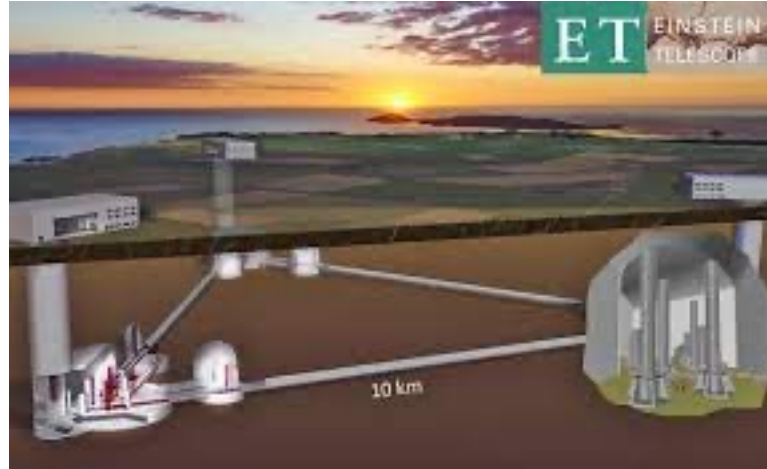
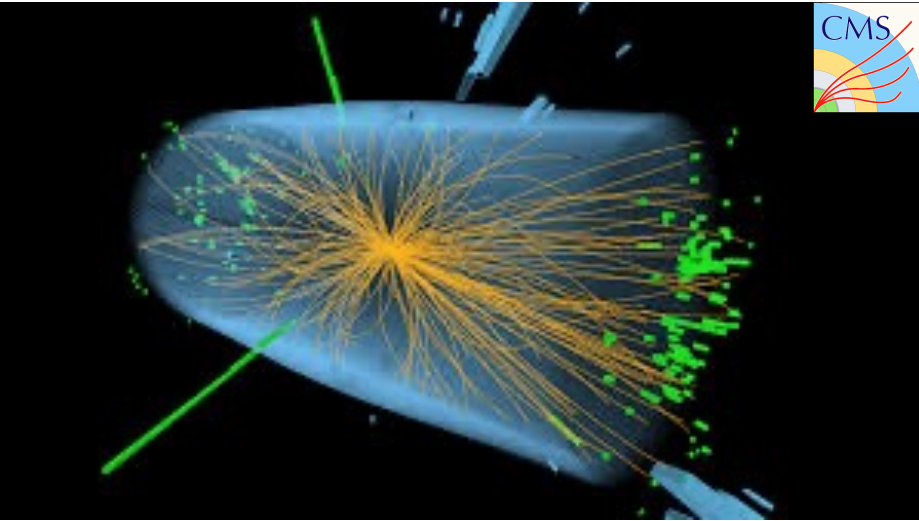
The structure of the ICSC National Center

The ICSC includes:
10 thematic Spokes and **1 Infrastructure Spoke**



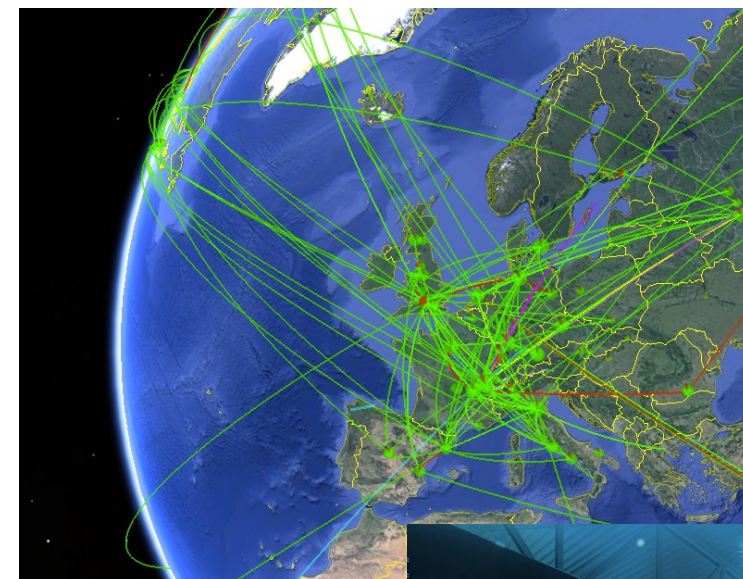
- One Spoke dedicated to building the infrastructure
- Ten thematic Spokes, one of which dedicated to the HEP and Astroparticle research domains.

Spoke 2 – Who are «we»?



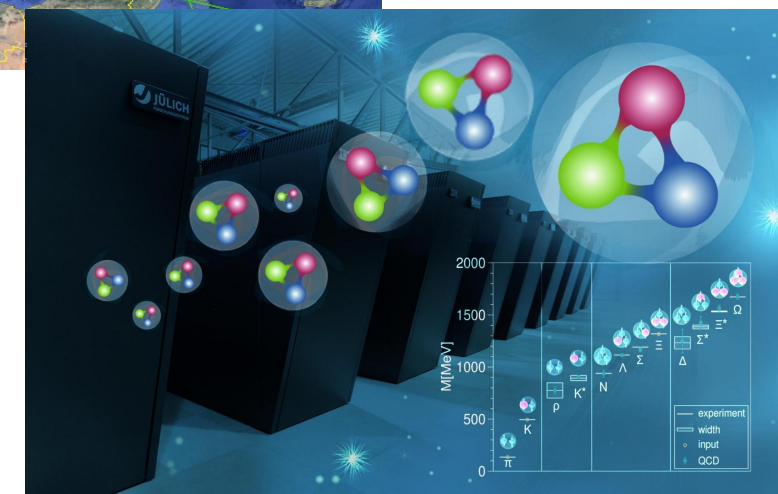
Why our communities need an ICSC?

- Since at least 2 decades, research at the fundamental frontier is heavily “computing bound”
 - Latest (and next) generation experiments collect data at the Exabyte and beyond;
 - Simulation efforts in theoretical and experimental physics are at the Exascale;
 - We have been forced to develop in-house solutions when nothing was available, with a good success. But it is now due time to evolve to a shared infrastructure model
 - The web, the GRID, ...
- Examples:
 - LHC has already surpassed the global scale of several Exabytes of Data, and more than 1 Million CPU cores.
 - Lattice QCD simulations are the main driver and benchmark for HPC systems.



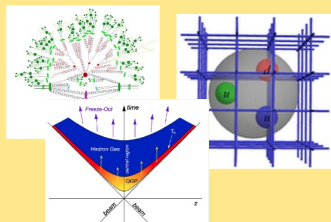
The Worldwide LHC Computing GRID.

Lattice QCD on HPC

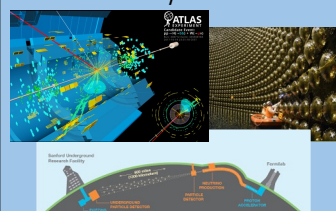


The structure of Spoke 2

WP1: tools and algorithms for Theoretical Physics



WP2: tools and algorithms for Experimental High Energy Physics



Scientific

WP3: tools and algorithms for Experimental Astroparticle Physics and Gravitational waves



ICSC-SPOKE2 Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

WP6: cross domain initiatives + space economy



WP5: Boosting computational performance on the distributed CN infrastructure



WP4: tools for porting/optimization on new architectures (low power, GPU, FPGA, ...)



FUNDAMENTAL RESEARCH & SPACE ECONOMY

Institution leader



Institution co-leader



Institutions and Universities

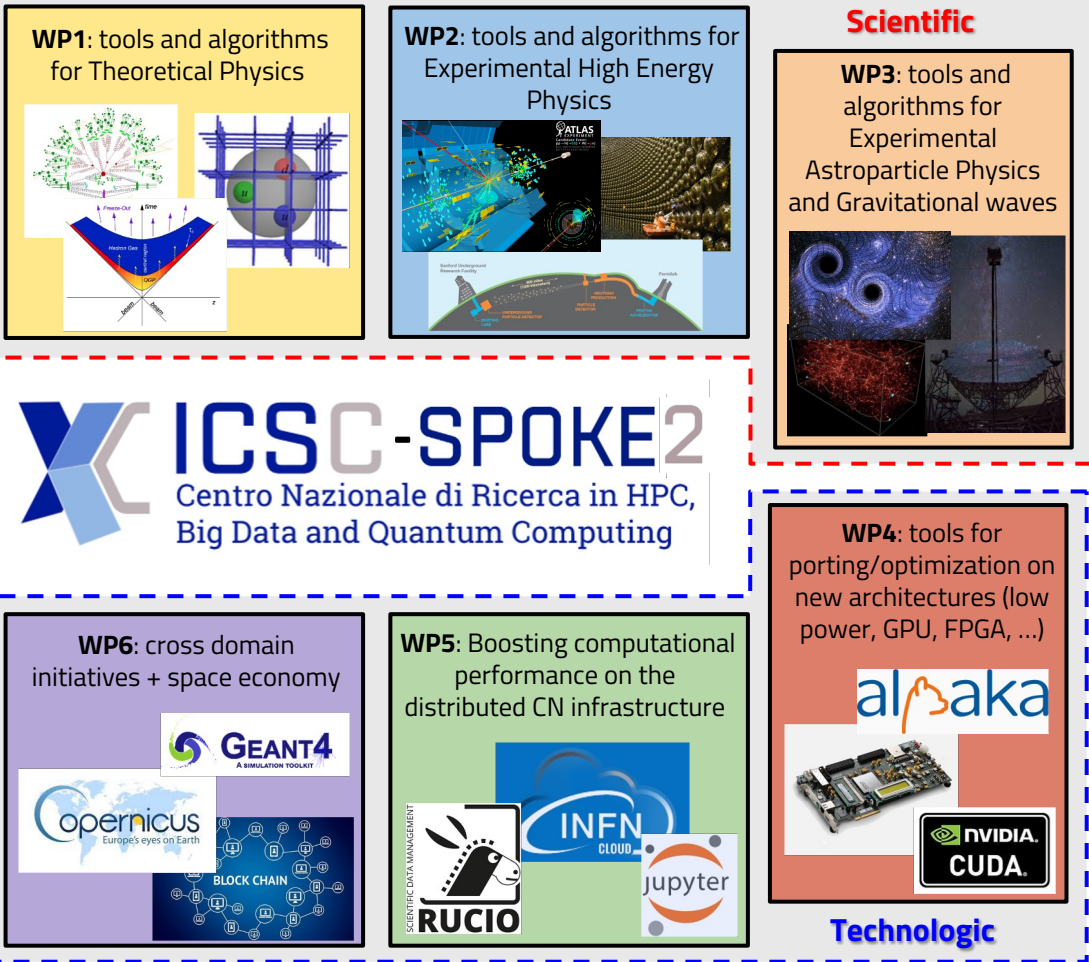


Companies

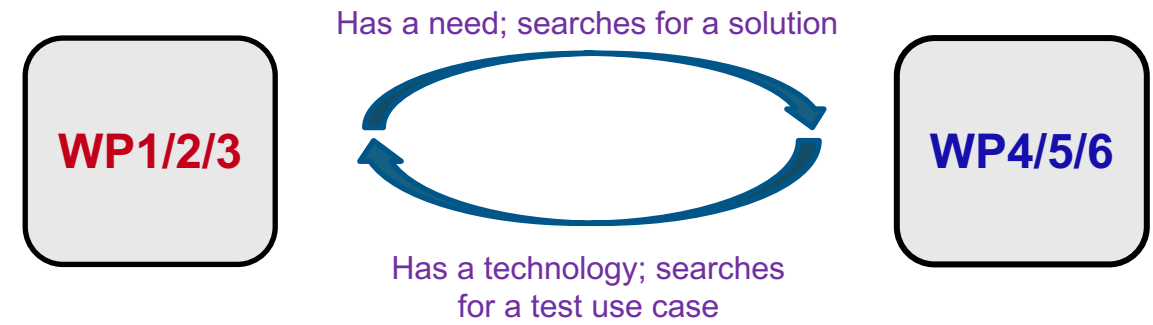


Staff Researchers	195
(kEur)	6333
Recruited researchers	28
(kEur)	5067
Phd positions	25
(kEur)	1992
Budget Innovation Grants (kEur)	1800
Budget Cascade Calls (kEur)	3200
Total Budget (kEur)	18391

The structure of Spoke 2



- We defined 2 types of Work Packages (WP):
 - **"Scientific"** WPs: they analyze the needs of the (sub-) domain, and pose open problems for which advanced computing solutions are needed;
 - **"Technological"** WPs: they harvest/investigate technical solutions in computing, on the infrastructure of the ICSC and beyond, and provide support / training for these; at the same time propose these to a larger audience, including industries.

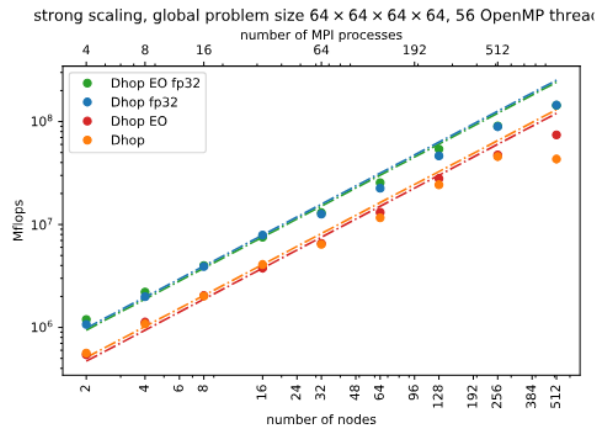


WP1 – Tools and algorithms for theoretical physics

Lattice QCD

- Multilevel Hybrid MC for lattice QCD**

Development, test and validation of fully optimized parallel code with OpenMP+MPI



- QCD under extreme conditions**

QCD equation of state computed in HPC resources, for the first time at high temperatures up to the EW scale.

[Phys. Rev. Lett. 134, 201904](#)

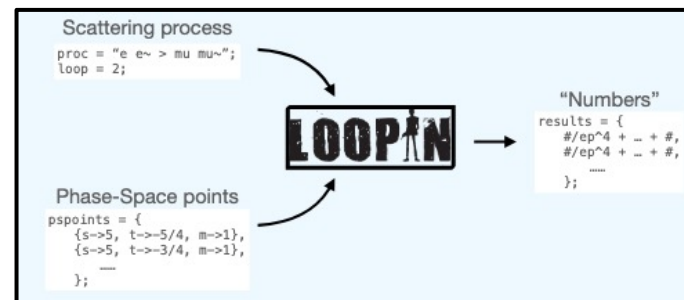
Collider Phenomenology

- Advanced Calculus for Precision Physics**

Developing software and tools for multi-loop scattering amplitudes. New architectures to improve event generation.

LoopIn: Loop Integrals for Virtual Corrections Framework for multi-loop calculations, automated for:

- Diagram generation
- Interferences terms
- Symbolic manipulations
- Integral numerical evaluation



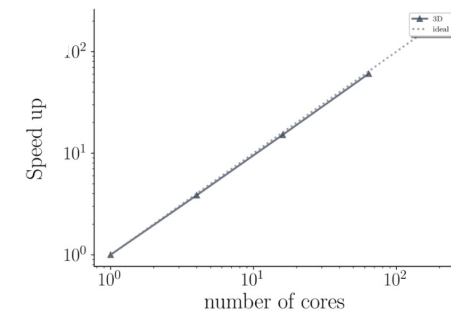
Benchmarked on the Leonardo HPC system.

Physics of Complex Systems

- Large Scale Simulations of Complex Systems**

Development and test of an algorithmic solver for the dynamics of an active liquid-crystal.

- Framework based on the Lattice Boltzmann Method (LBM), developed with MPI.
- Relevant for many applications, including drug delivery.



nature communications

Article <https://doi.org/10.1038/s41467-025-56236-8>

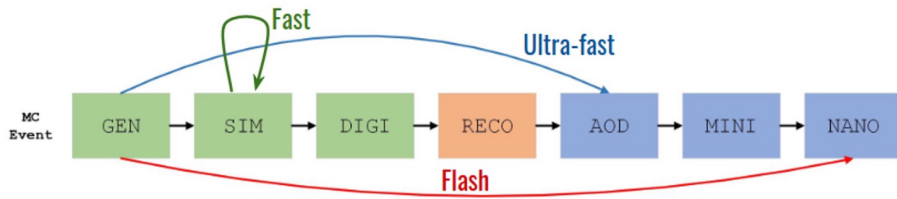
Topology controls flow patterns in active double emulsions

Received: 10 February 2024
Accepted: 10 January 2025
Published online: 06 February 2025

Giuseppe Negro^{1,2}, Louise C. Head^{1,2,3}, Livio N. Carenza⁴, Tyler N. Shandruk⁵, Davide Marenduzzo⁶, Giuseppe Gonnella^{7,8}, Adriano Tribocchi⁸

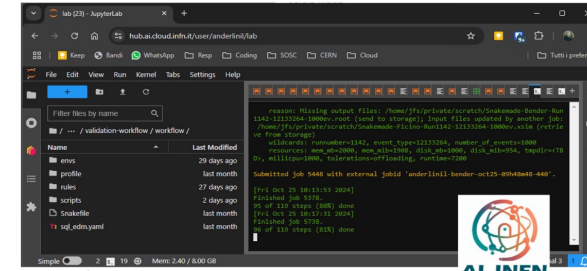
WP2 – Tools and algorithms for experimental HEP

- Advanced ML: flash simulation and other bleeding edge applications

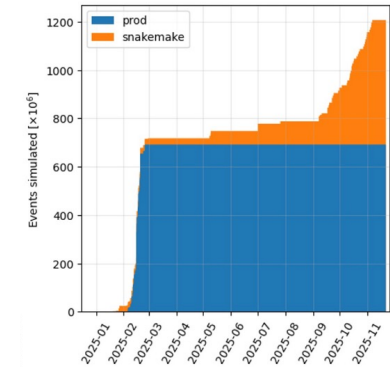


Flash simulation in LHCb as a case study

Express complex workflows (via Snakemake), and enable seamless offloading to heterogeneous resources (HTC, HPC) via Interlink.



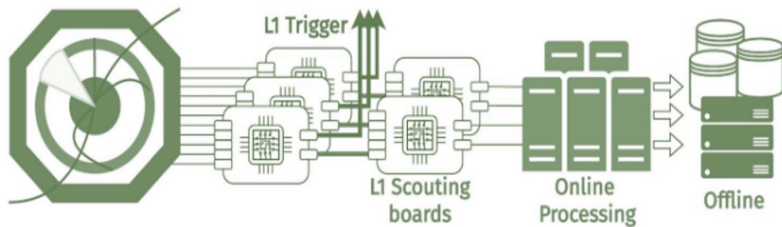
HTC interLink HPC payload



> 1.2 B events produced both via **snakemake** and a **WLCG-like** submission mechanism (for scalability tests)

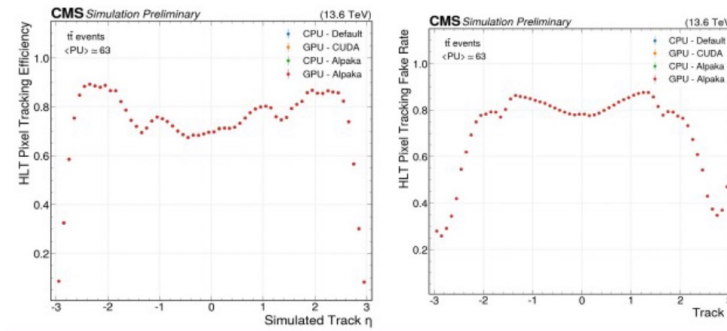
- Development of ultra-fast algorithms on FPGA

- Trigger algorithms in experiments (ATLAS, CMS, ...)
- Development of reconstruction algorithms
- Scouting of data at 40MHz (e.g. CMS)



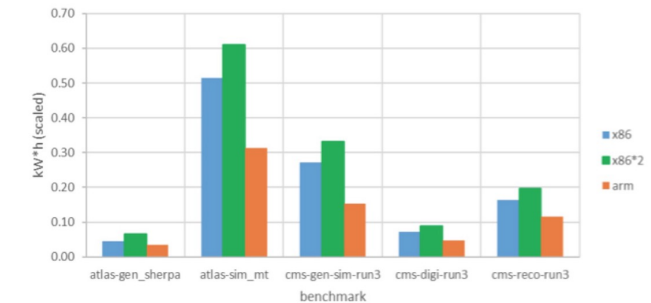
- Porting of algorithms to GPUs

- Port experiment software on GPUs, also using portability tools (e.g. Alpaka)



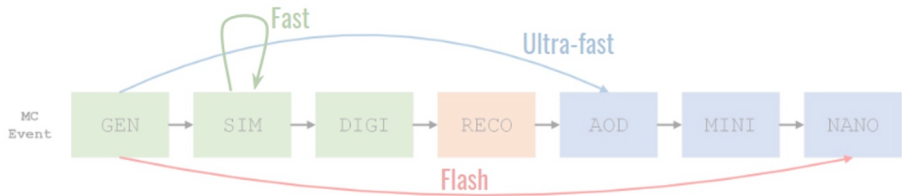
- Physics validation on ARM

- Validation of reconstruction software on ARM architectures.

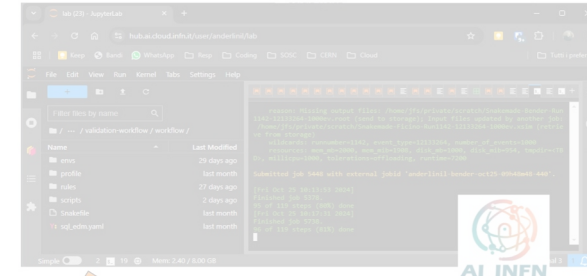


WP2 – Tools and algorithms for experimental HEP

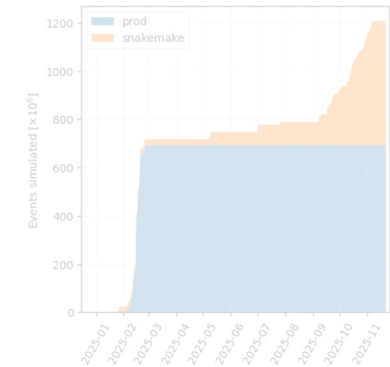
- Advanced ML: flash simulation and other bleeding edge applications



Flash simulation in LHCb as a case study



HTC interLink HPC payload



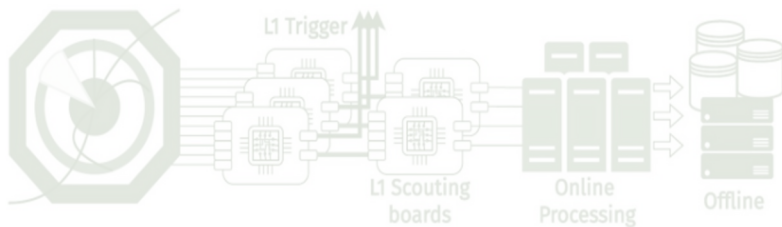
Express complex workflows (to heterogeneous resources (

- Quasi interactive data analysis of big data with high throughput**

More on tomorrow's talk!!

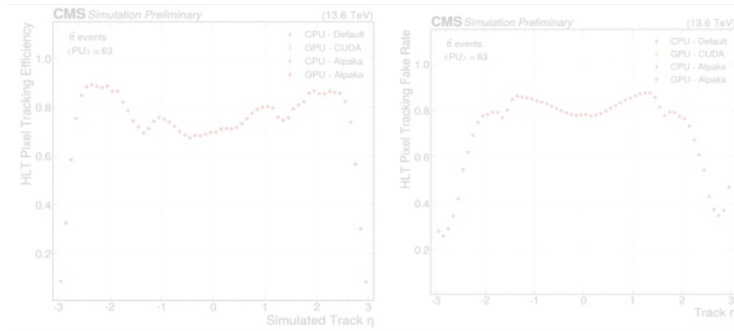
> 1.2 B events produced both via **snakemake** and a **WLCG-like** submission mechanism (for scalability tests)

- Development of ultra-fast algorithms on FPGA
 - Trigger algorithms in experiments (ATLAS, CMS, ...)
 - Development of reconstruction algorithms
 - Scouting of data at 40MHz (e.g. CMS)



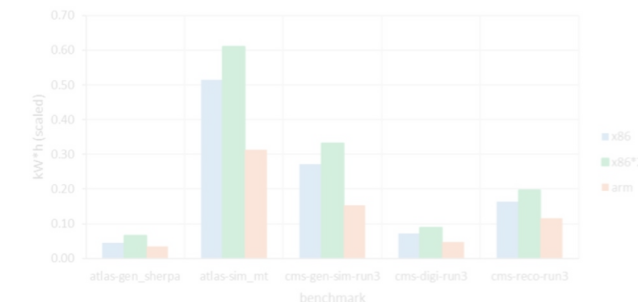
- Porting of algorithms to GPUs

- Port experiment software on GPUs, also using portability tools (e.g. Alpaka)



- Physics validation on ARM

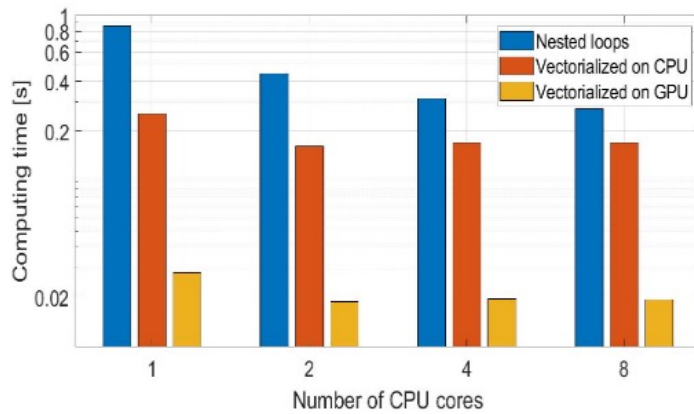
- Validation of reconstruction software on ARM architectures.



WP3 – Tools and algorithms for experimental AP and GW physics

- Frequency Hough Transform on GW sources**

Python implementation of FH transform on PyCUDA

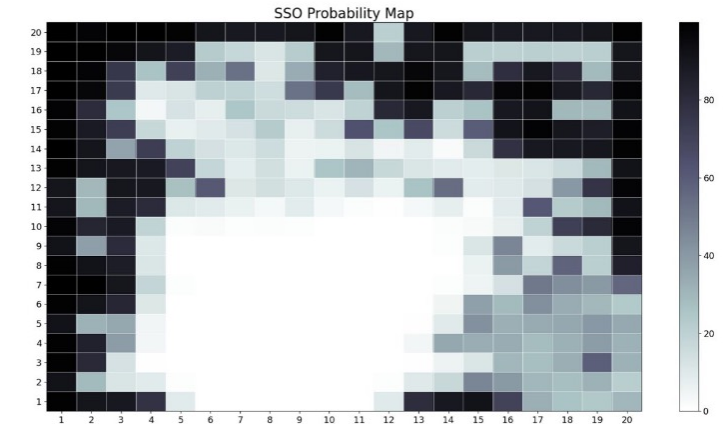


- Efficient use of ML and GPUs in cosmological data analysis**
- Inference of cosmological and astrophysical population properties from GW observations with and without electromagnetic counterparts**
- Innovative analysis techniques using realistic simulations of the upgraded Auger Observatory within the context of a ML environment**

- Detection and classification of SSO in Euclid Simulated data**



[A&A, Volume 694, A116](#)



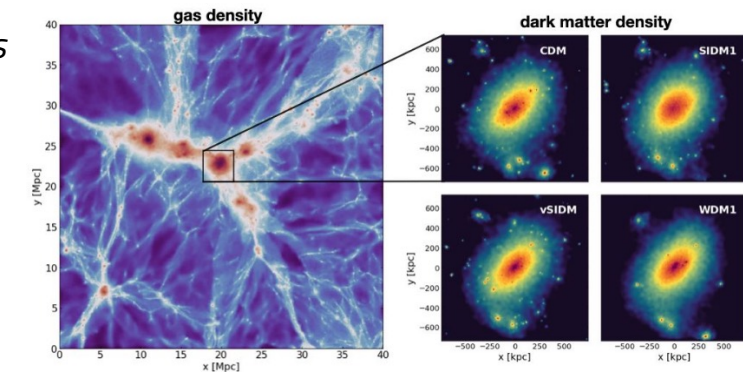
- Pipeline optimization for space and ground based experiments**

Collecting the requirements of different experiments into a common framework based on opportunistic resources offered by ICSC

- Hydrodynamical simulations to test the nature of dark matter**

Analysis of the produced simulations and comparison to observations

[A&A, Volume 697, A213](#)



WP4 – Tools for porting/optimization on new architectures (FPGA, GPU, ...)

• Organization of various courses

First course about the porting on GPUs of code and algorithms

19–21 Jun 2023
Europe/Rome timezone

Introductory course to HLS FPGA programming.

Overview 27–30 Nov 2023

Introductory course to VHDL

4–6 Mar 2024
Europe/Rome timezone

Introductory course to VHDL and HLS FPGA programming

Overview 23–27 Jun 2025
Università degli Studi Milano Bicocca

Second course on porting code and algorithms to GPUs

30 June 2025 to 4 July 2025
Dipartimento di Fisica e Astronomia
Europe/Rome timezone

Advanced course to FPGA programming

Overview 27–29 Oct 2025
INFN Perugia
Europe/Rome timezone

Overview

- Timetable
- Contribution List
- Registration
- Participant List
- Access to GPU Virtual Machines

Overview

- Timetable
- Contribution List
- Registration
- Participant List
- Travel & Accomodation

Starts 27 Oct 2025, 14:00
Ends 29 Oct 2025, 18:00
Europe/Rome

Daniele Spiga
Pasquale Lubrano
Simone Gennai
Stefano Giagu

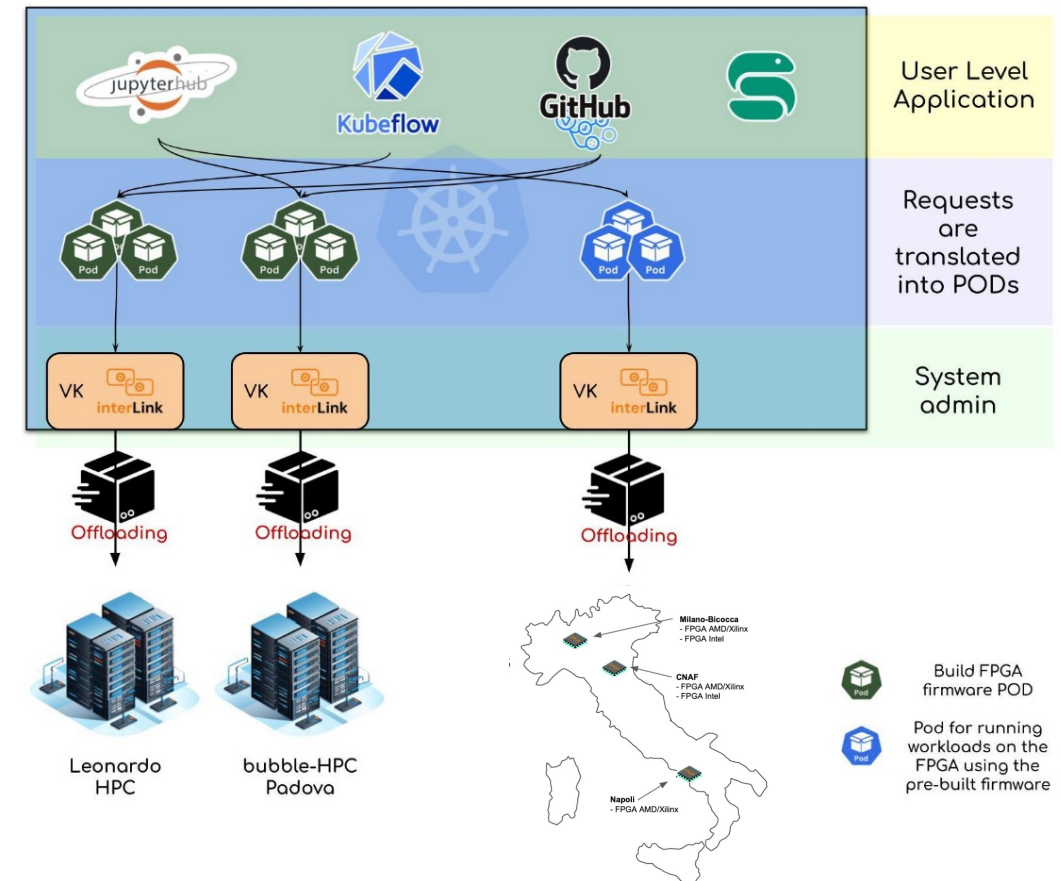
INFN Perugia
Dipartimento di Fisica e Geologia - FISGEO UNIPG
Via Alessandro Pascoli, 06123 Perugia PG
[Go to map](#)

Bulletin Advanced Course on FPGA programmin...

ICSC
Centro Nazionale di Ricerca in HPC,
Big Data and Quantum Computing

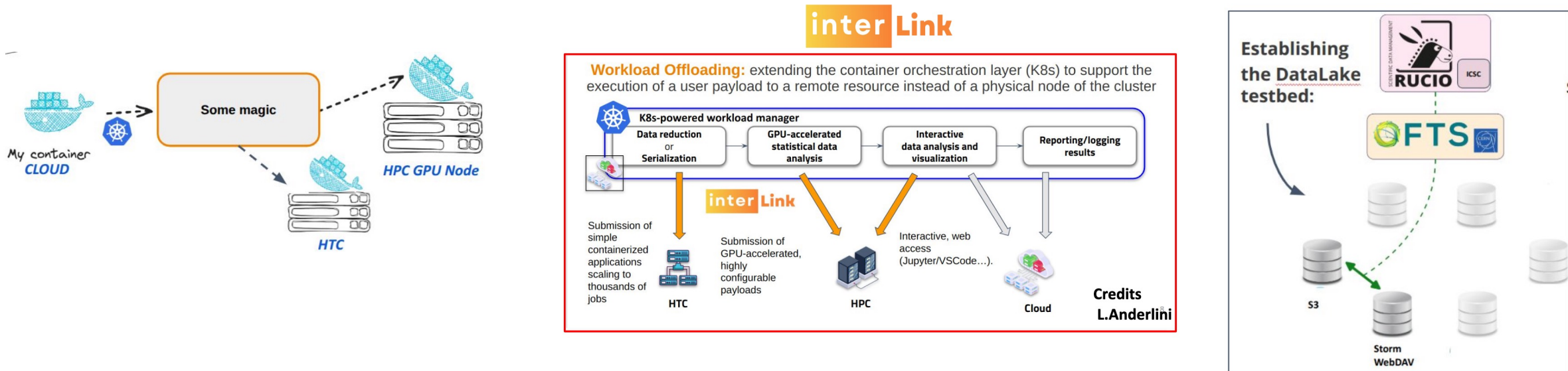
• FPGA bubbles

Creation of a national pool of FPGA opportunistic resources. Mainly Xilinx/AMD (Alveo U55C) and Intel (DE10-Agilex).



WP5 – Computational performance on ICSC resources

- Support the scientific WPs, adapting the existing application on the datalake infrastructure, and design/implement/test the computing models.



- Training activities:

Mini-Workshop on Data Management

5 July 2024
Europe/Rome timezone

Workshop on "Quasi-Interactive Analysis of Big Data with High Throughput"

8–10 Jan 2025
Bologna
Europe/Rome timezone

Enter your search

Workshop on "Quasi-Interactive Analysis of Big Data with High Throughput"

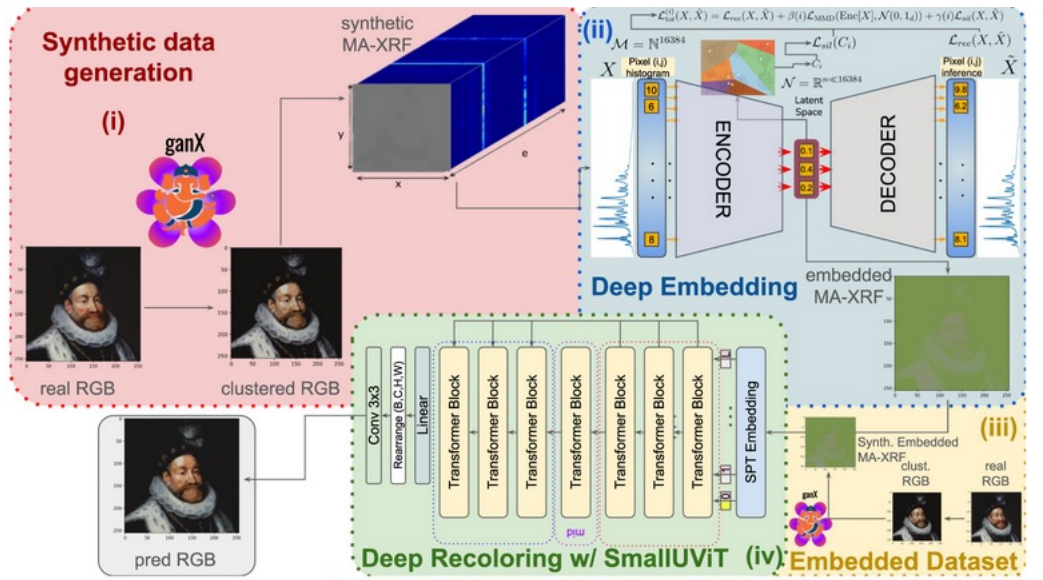
SOSC 2024 Sixth International School on Open Science Cloud

2–6 Dec 2024
University of Bologna, Department of Physics and Astronomy
Europe/Rome timezone

Enter your search

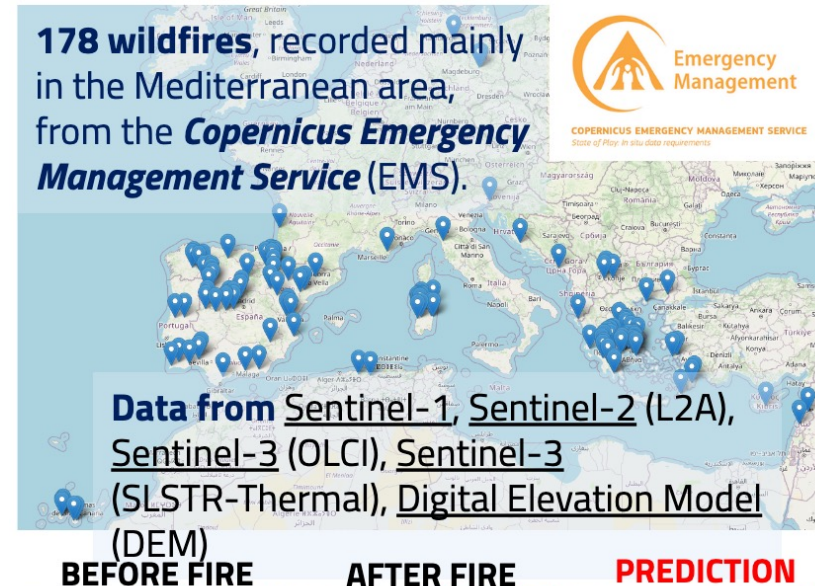
WP6 – Cross domain initiatives + Space Economy

- Extended Computer Vision at high rate
Virtual recolouring of degraded artworks

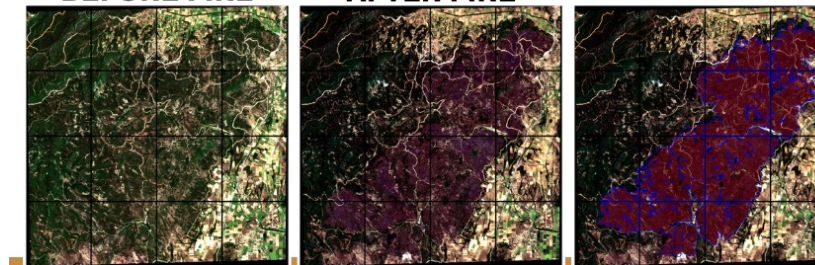
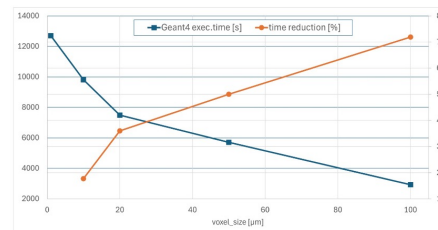
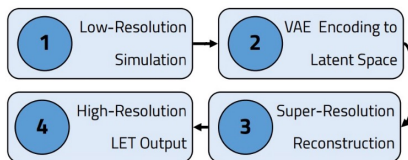


- Satellite imaging

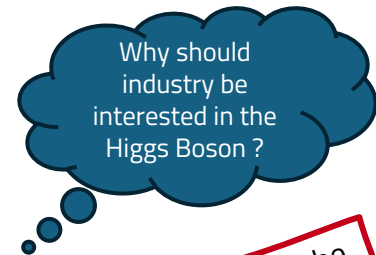
Provide new, powerful deep learning tools for large-scale environmental monitoring and disaster management.



- Enhancing Geant4 Monte Carlo Simulations through Machine Learning Integration



A further step: «From Research to Industry»



Share the platforms, the solutions, the expertise!

- **Eni**
 - Testing HEP workflows on a industrial data center
 - Predictive maintenance on ENI's industrial plants
 - Physics Inspired ML for the exploitation of natural resources
 - **Leonardo + TASI**
 - Interoperable Datalake: Data management + Block Chain pilot for Space Economy data
 - **Unipol**
 - Merging data and AI for improving Seismic Risk assessment and management in populated areas
 - **UnipolSAI + SOGEI**
 - Monitoring of alert and vulnerabilities for fast response and assurance evaluation
 - **Intesa Sanpaolo**
 - Correlations between satellite images and financial situation of agricultural enterprises
 - Fraud detection on streams of financial data
- 2+ M€ on "**Innovation Funds**": projects between the Center and Large Enterprises;
 - 3.2 M€ on "**Cascade Calls**": projects between research and Small-Medium Enterprises.

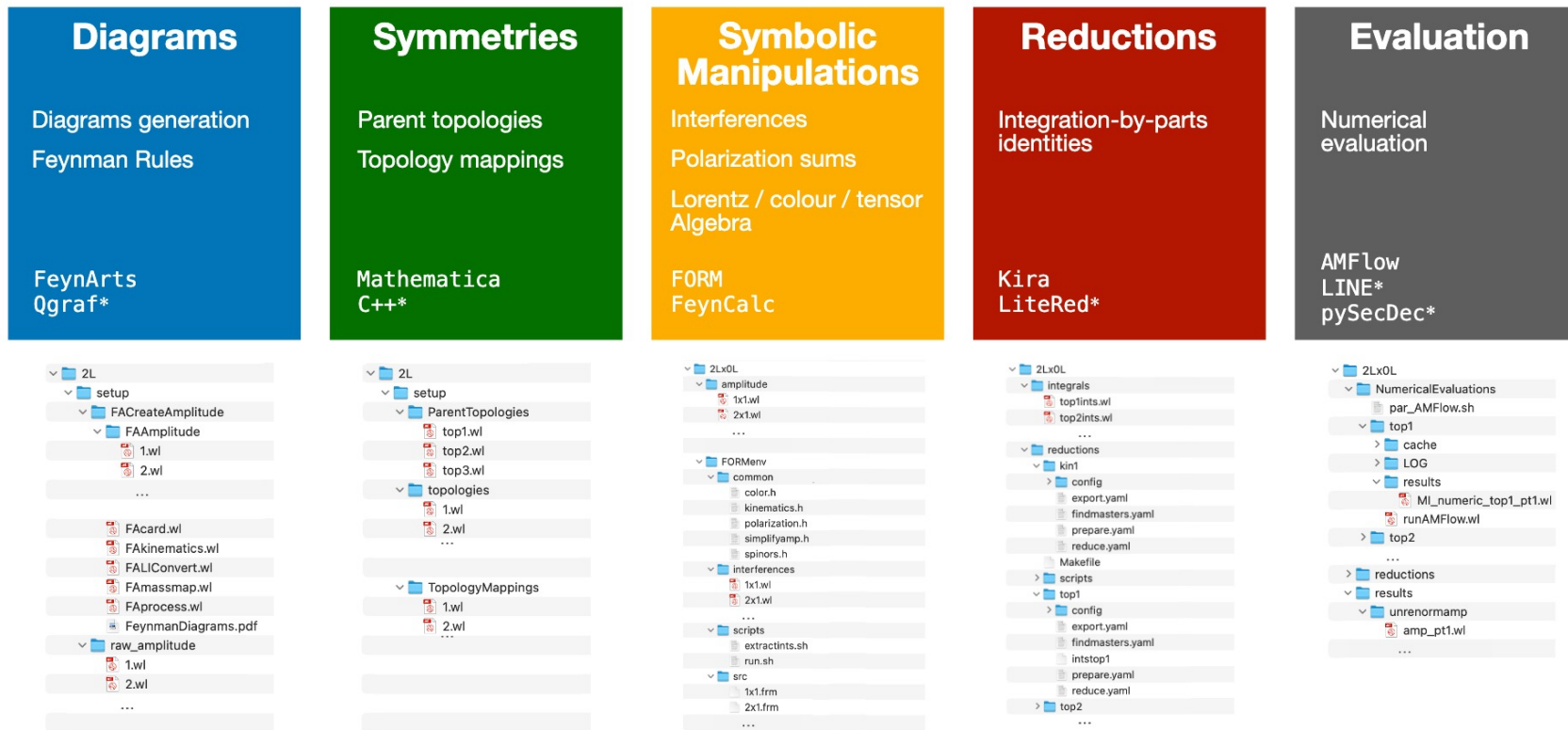
Conclusions

- The National Center for HPC, Big Data and Quantum Computing is a unique opportunity for research in Italy: build a modern infrastructure for research initially, and industry later, and make sustainable beyond the initial fundings
- Spoke 2 activities are aligned with the needs from High Energy and Astroparticle physics
- We are trying to form the next generation of scientists, versed in modern technologies
- While doing that, we want to strengthen the links with Italian industries

BACKUP

LoopIn: Module structure

LoopIn: Module structure

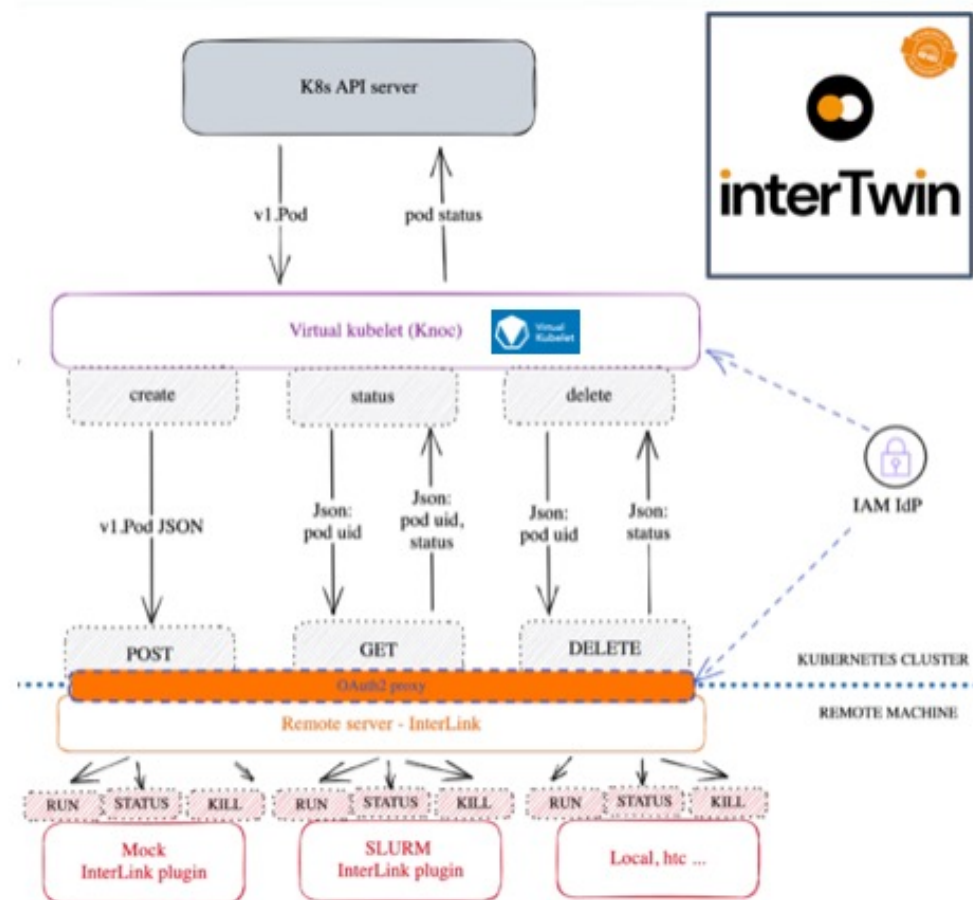


This work is supported by ICSC – Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing, funded by European Union – NextGenerationEU.

Offloading strategy with Interlink

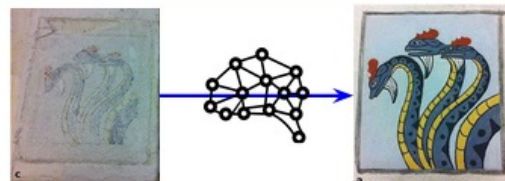
Scheduling worker processes, spawning on multiple remote sites dynamically and transparently:

- Virtual Kubelet** (open-source Kubernetes kubelet implementation that masquerades as a kubelet): registers as a virtual node and pulls work to run;
 - “It takes your pod and executes it wherever”
- InterLink + HTCondor Sidecar (Plugin)**: pods are translated into HTCondor jobs:
 - Translating interlink create/status/delete calls interacting with the proper HTCondor `schedd` via CLI
 - POST /create call -> `condor_submit`
 - GET /status call -> `condor_q`
 - POST /delete call -> `condor_rm`
- In future, this strategy can be also applied to other job scheduling systems (e.g. Slurm/HPC)



Virtual recoloring

Use case: Virtual recoloring



Bombini, A., Bofias, F. G. A., Bracci, C., Ginolfi, M., & Ruberto, C. (2024). Datacube segmentation via deep spectral clustering. *Machine Learning: Science and Technology*, 5(3), 035024. DOI [10.1088/2632-2153/ad622f](https://doi.org/10.1088/2632-2153/ad622f)



The idea - part I: create a *synthetic dataset* ad hoc (i.e., containing the appropriate pigments) in a distributed manner using Dask

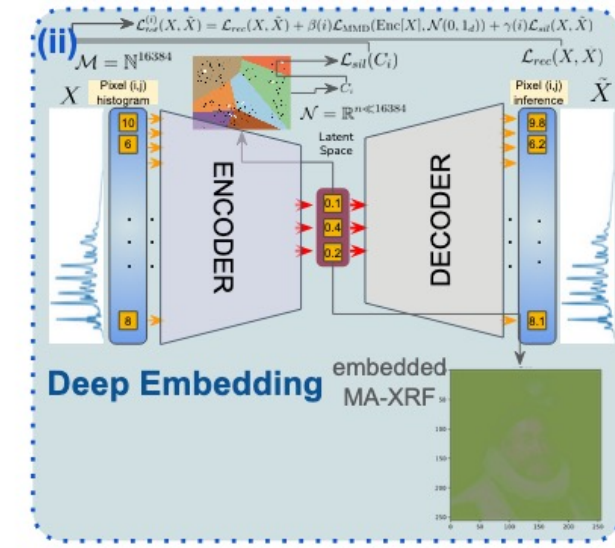
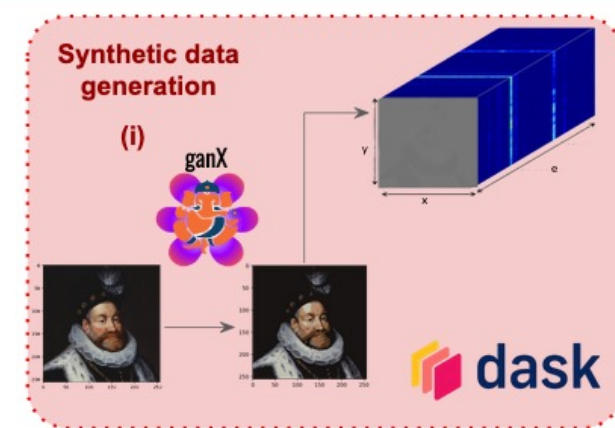
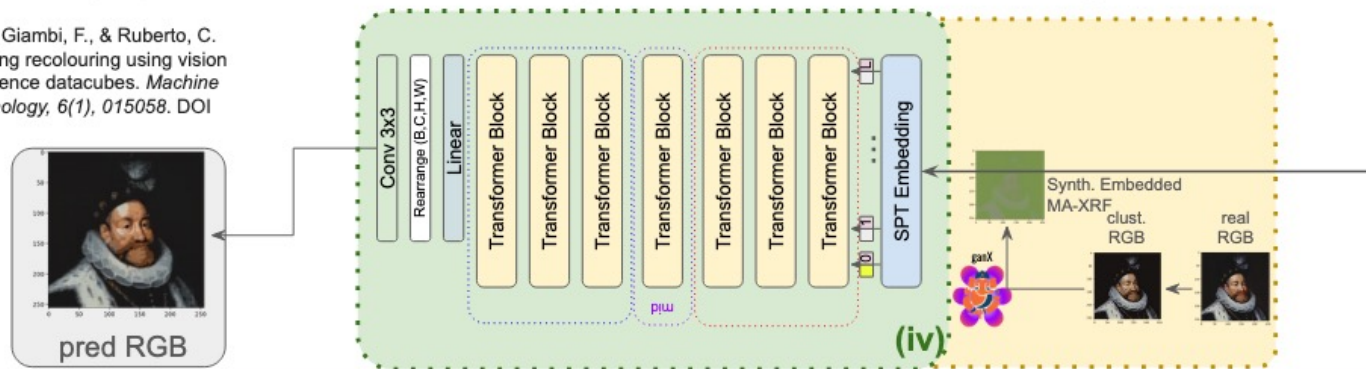


The idea - part II: Train semi-supervisingly a *Deep Variational Embedding*, to embed in lower dimensions the synthetic datacubes



The idea - part III: Train a new model for assigning RGB colors to (embedded) spectra, i.e. achieve *digital restoration* of the degraded surface.

Bombini, A., Bofias, F. G. A., Giambi, F., & Ruberto, C. (2025). Towards virtual painting recoloring using vision transformer on x-ray fluorescence datacubes. *Machine Learning: Science and Technology*, 6(1), 015058. DOI [10.1088/2632-2153/adb937](https://doi.org/10.1088/2632-2153/adb937)



Wildfires segmentation using time-series of images from the Sentinel constellation

Target: provide new, powerful **deep learning** tools for **large-scale environmental monitoring and disaster management**.

Custom Python library employing the SentinelHub API developed to streamline data download, preprocessing and integration.

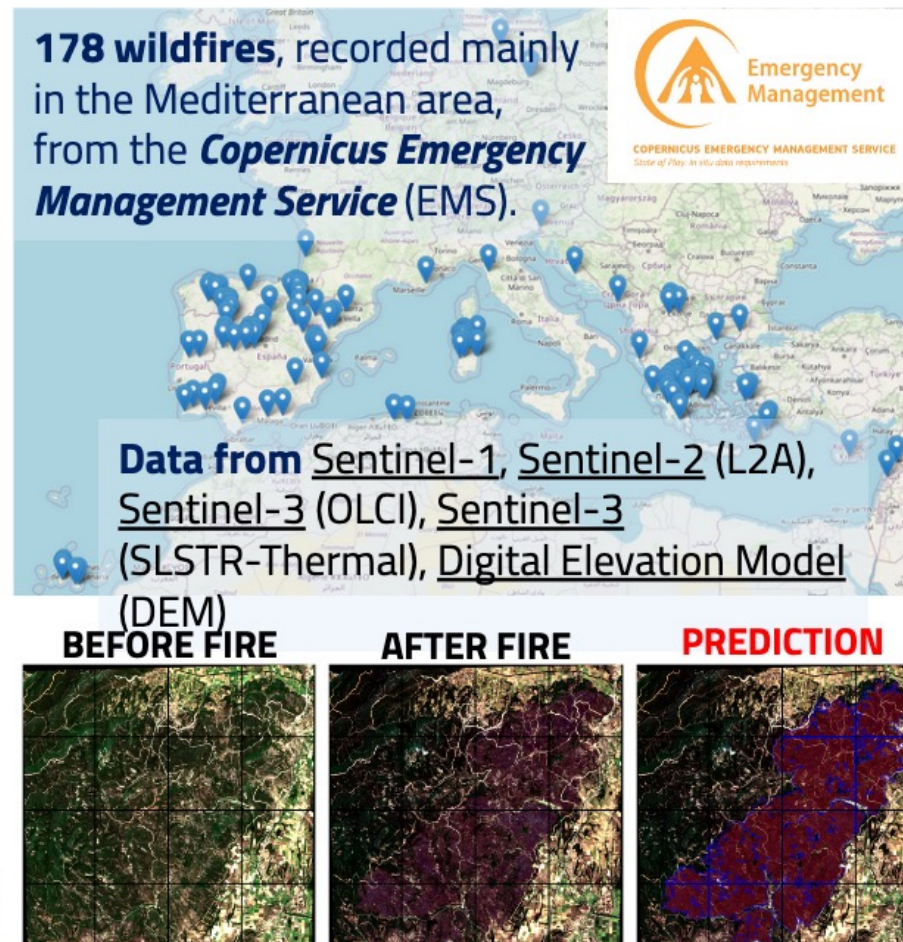
Leonardo Booster system leveraged to run parallel, distributed hyperparameter tuning with **Optuna**, dispatching multiple jobs on different GPUs.

Metrics	
Accuracy	0.97
AUC	0.98
Precision	0.88
Recall	0.92
F1 Score (Dice)	0.90
IoU (Jaccard)	0.82

UNet-like model with 5 encoder heads (for the 5 data sources) and Convolutional Long-Short-Term-Memory layers.

Model trained using 130 events in a **5-fold cross-validation** scheme. Metrics over the 48 test-set events indicate a **solid overlap between predicted and affected regions**.

Copernicus Emergency Management Service (©2022 EU), EMSR605 A0101 available at <https://mapping.emergency.copernicus.eu/activations/EMSR605/>



See also *EGU25-9376*

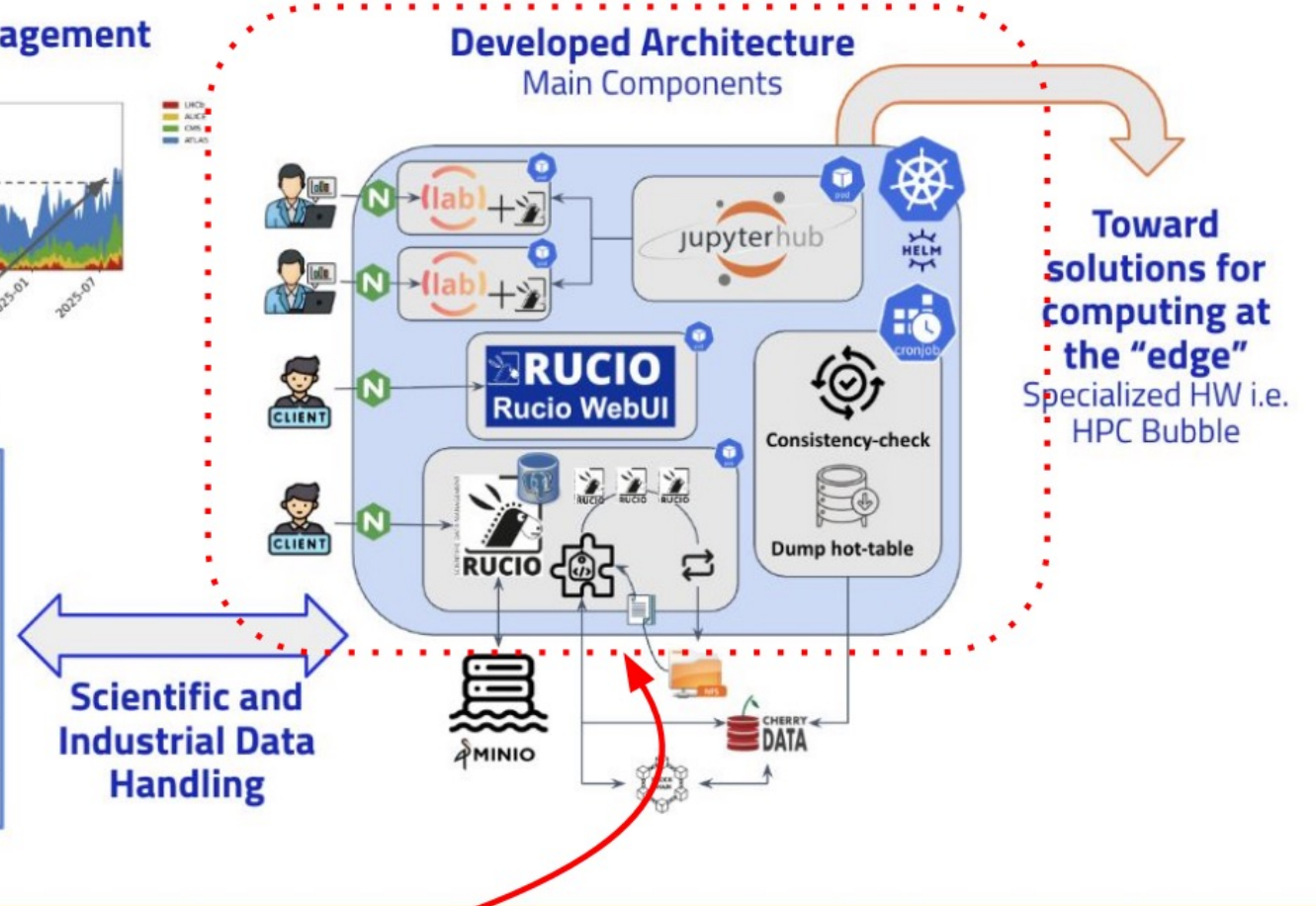
Datalake, data management

Scientific Background: Data Management



~ 1.8 EB/month

Dataset	Status	Custom plugin?	# files	Total time	Mean time (s)
RAW: teres	In progress	Yes	1763	2:23:59.946191	4.900707
RAW: pulhar	In progress	Yes	78	0:43:12.879830	33.242049
RAW: brabas	In progress	Yes	15	0:00:42.692068	2.846138
TRSI	Not started	Yes			
Spacetrack	Completed	Yes	5681116	1:15:22:43	0.07311292
Spacetrack	Completed	No	5681116	94:42:08	0.06001074



Scientific and Industrial Data Handling

Full stack running on ICSC resources hosted in INFN Cloud