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

Quasi interactive analysis of Big Data with high throughput

Tommaso Diotalevi (University of Bologna / INFN)

International Symposium on Grids and Clouds (ISGC2025)
18th March 2025

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 December 2025.

1. **ICSC: Big Data, High Performance and Quantum Computing** — ICSC
2. **Advanced Technologies in agricolture (Agritech)** — agritech, National Research Center for Technology in Agriculture
3. **Sustainable Mobility** — MOST
4. **Development of RNA drugs and gene therapies** — RNA &GENETHERAPY
5. **Biodiversity** — NATIONAL BIODIVERSITY FUTURE CENTER

The participants

The ICSC foundation: public and private members
an initiative spread across Italy

National institutes

INFN Istituto Nazionale di Fisica Nucleare | CINECA | ENEA
Consiglio Nazionale delle Ricerche | INAF ISTITUTO NAZIONALE DI ASTROFISICA | ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA | Consortium GARR

12 Research institutes

HUB

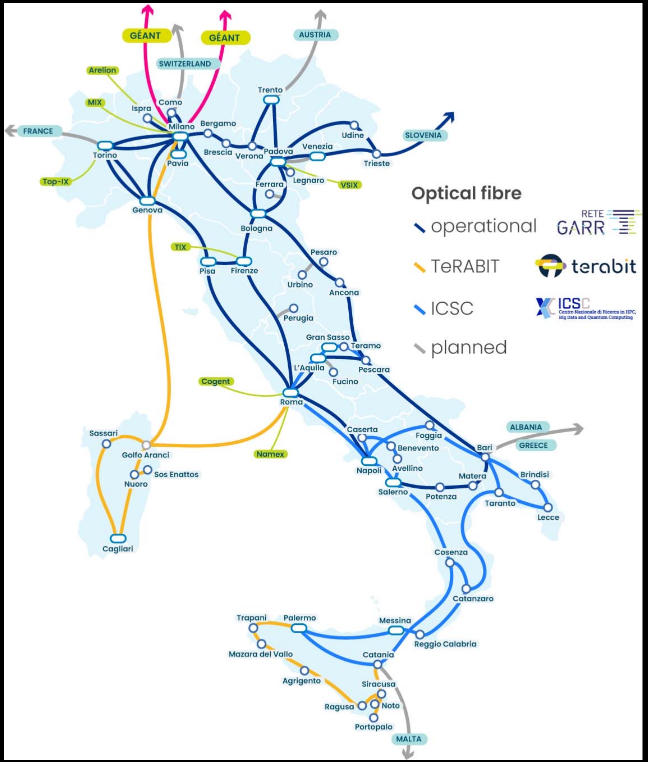
UNIMORE UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA | UNIVERSITÀ DI PARMA | OGS

Privates

enel | ENGINEERING THE DIGITAL TRANSFORMATION COMPANY | eni | FERROVIE DELLO STATO ITALIANE | FINCANTIERI
fondazione innovazione urbana | autostrade per l'italia | HUMANITAS RESEARCH HOSPITAL | iFAB INTERNATIONAL FOUNDATION BIG DATA AND ARTIFICIAL INTELLIGENCE FOR HUMAN DEVELOPMENT | INTESA SANPAOLO
LEONARDO | sogei | ThalesAlenia Space | Terna Driving Energy | UnipolSai ASSICURAZIONI

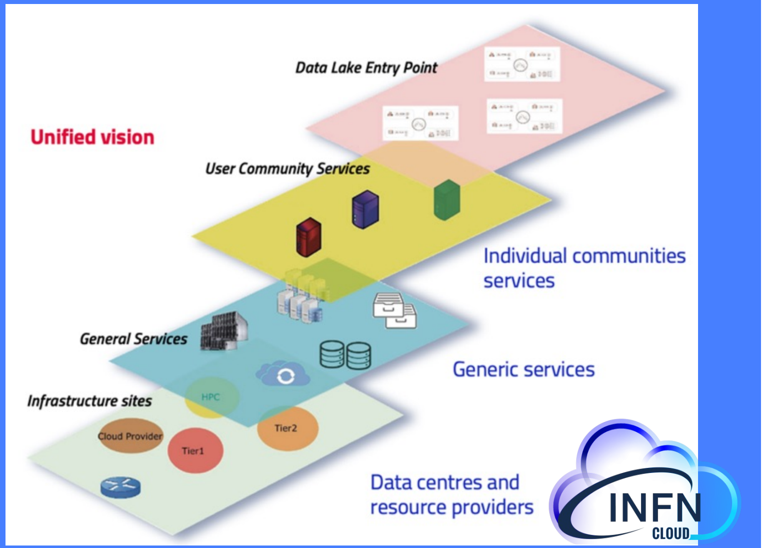
On the infrastructure: three major pillars

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



CINECA: expansion of Leonardo (HPC#6 on top500.org) with Lisa, and deployment of one of the European AI Factories.

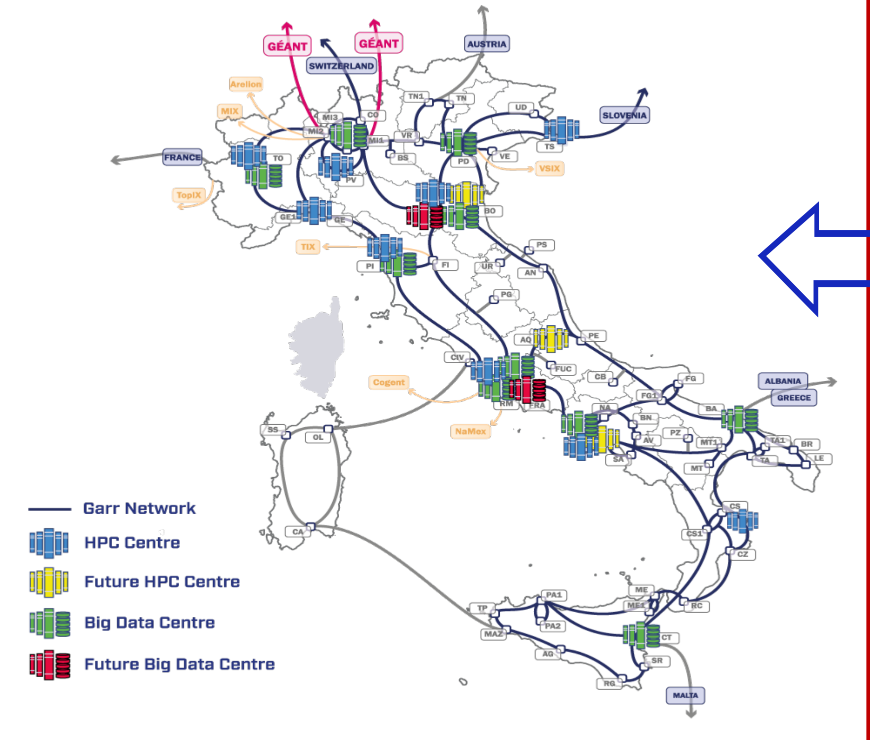
INFN: strengthening of the WLCG infrastructure (1 Tier-1 & 9 Tier-2); acquisition of Cloud resources; implementation of the datalake middleware, based on INFN Cloud.



The structure of the ICSC National Center

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

0 SUPERCOMPUTING CLOUD INFRASTRUCTURE

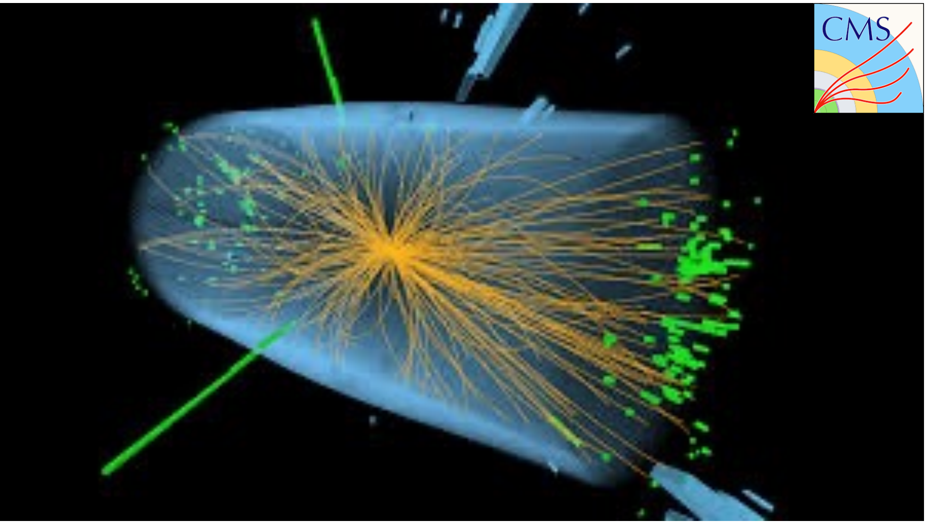


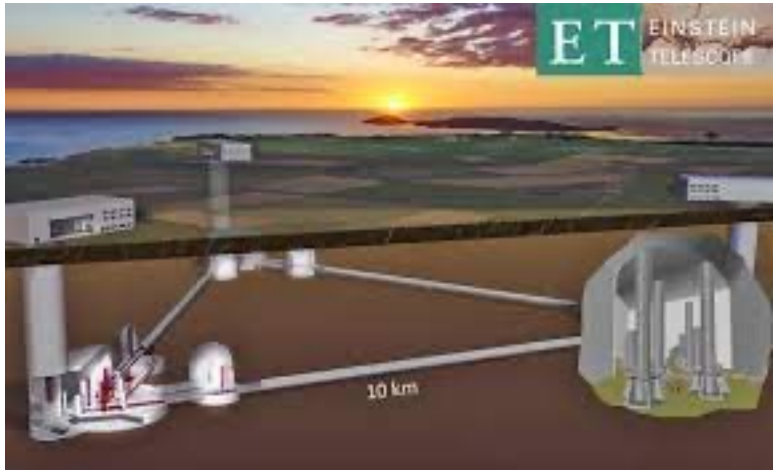
EDUCATION AND TRAINING, ENTREPRENEURSHIP, KNOWLEDGE TRANSFER, POLICY MAKING, OUTREACH

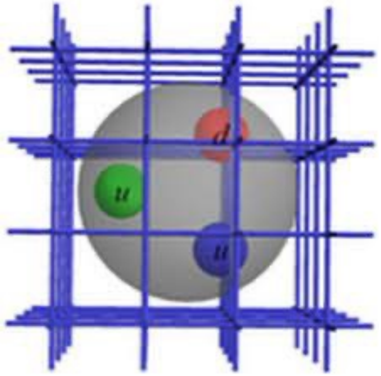
1. FUTURE HPC & BIG DATA
2. FUNDAMENTAL RESEARCH & SPACE ECONOMY
3. ASTROPHYSICS & COSMOS OBSERVATIONS
4. EARTH & CLIMATE
5. ENVIRONMENT & NATURAL DISASTERS
6. MULTISCALE MODELING & ENGINEERING APPLICATIONS
7. MATERIALS & MOLECULAR SCIENCES
8. IN-SILICO MEDICINE & OMICS DATA
9. DIGITAL SOCIETY & SMART CITIES
10. QUANTUM COMPUTING

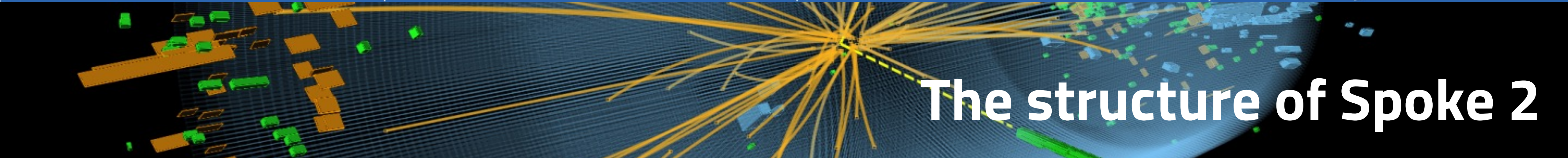
- One Spoke dedicated to building the infrastructure
- Ten thematic Spokes, one of which dedicated to the HEP and Astroparticle research domains.

Spoke 2 – Who are «we»?





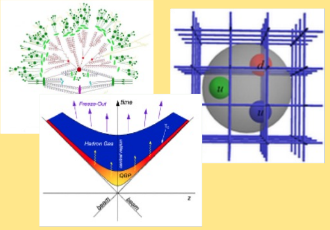




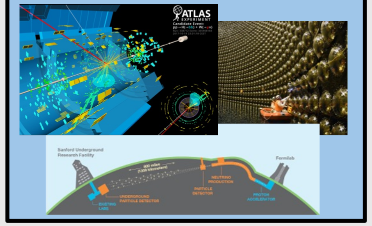
The structure of Spoke 2

Scientific

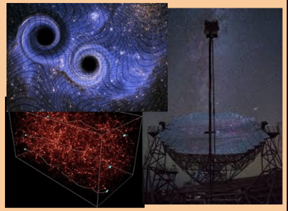
WP1: tools and algorithms for Theoretical Physics



WP2: tools and algorithms for Experimental High Energy Physics



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



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

Technologic

WP6: cross domain initiatives + space economy

GEANT4 A SIMULATION TOOLKIT
copernicus Europe's eyes on Earth

WP5: Boosting computational performance on the distributed CN infrastructure

INFN CLOUD · RUCIO SCIENTIFIC DATA MANAGEMENT · jupyter

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

alpaka

NVIDIA CUDA

FUNDAMENTAL RESEARCH & SPACE ECONOMY

Istitution leader

INFN Istituto Nazionale di Fisica Nucleare

Istitution co-leader

INAF ISTITUTO NAZIONALE DI ASTROFISICA

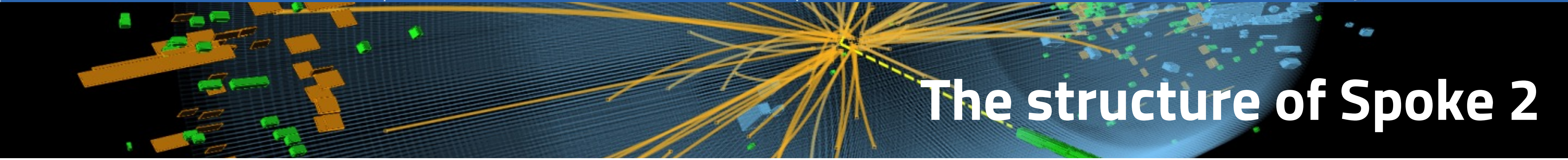
Istitutions and Universities

Politecnico di Bari · Sapienza Università di Roma · Università degli Studi di Bari Aldo Moro · Alma Mater Studiorum Università di Bologna · Università della Calabria · Università di Catania · Università degli Studi di Ferrara · Università degli Studi Firenze · Università degli Studi di Milano Bicocca · Università degli Studi di Napoli Federico II · Università degli Studi di Padova · Università del Salento · Università degli Studi di Trieste

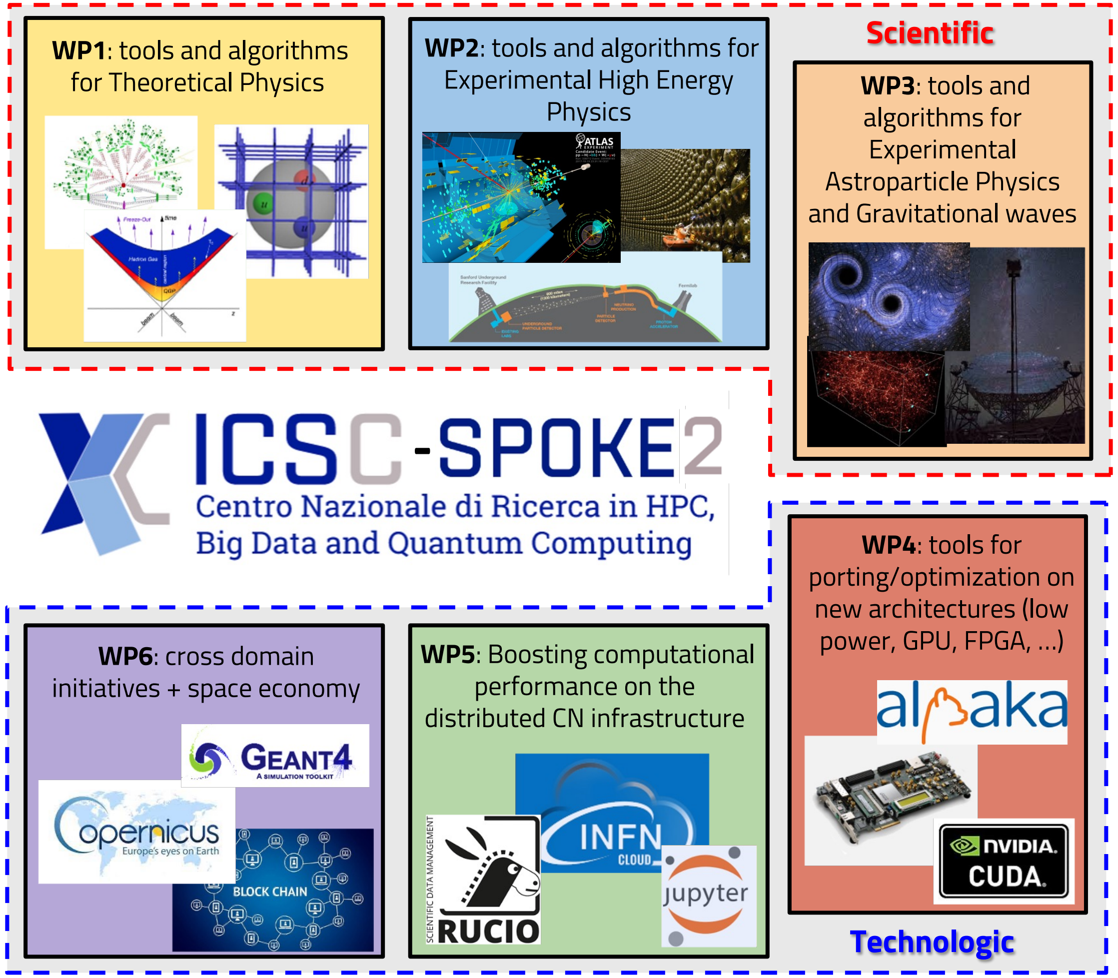
Companies

eni · Intesa Sanpaolo · Leonardo · ThalesAlenia Space · UnipolSai Assicurazioni · iFAB International Foundation Big Data and Artificial Intelligence for Human Development · sogei

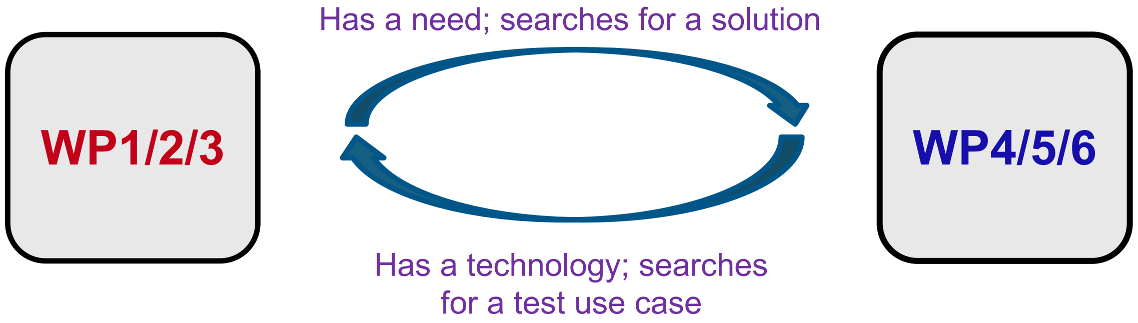
Staff Researchers	195
(kEur)	6333
Rectruited researchers	28
(kEur)	5067
Phd positions	25
(kEur)	1992
Budget Innovation Grants (kEur)	1800
Bugdet Cascade Calls (kEur)	3200
Total Budget (kEur)	**18391**



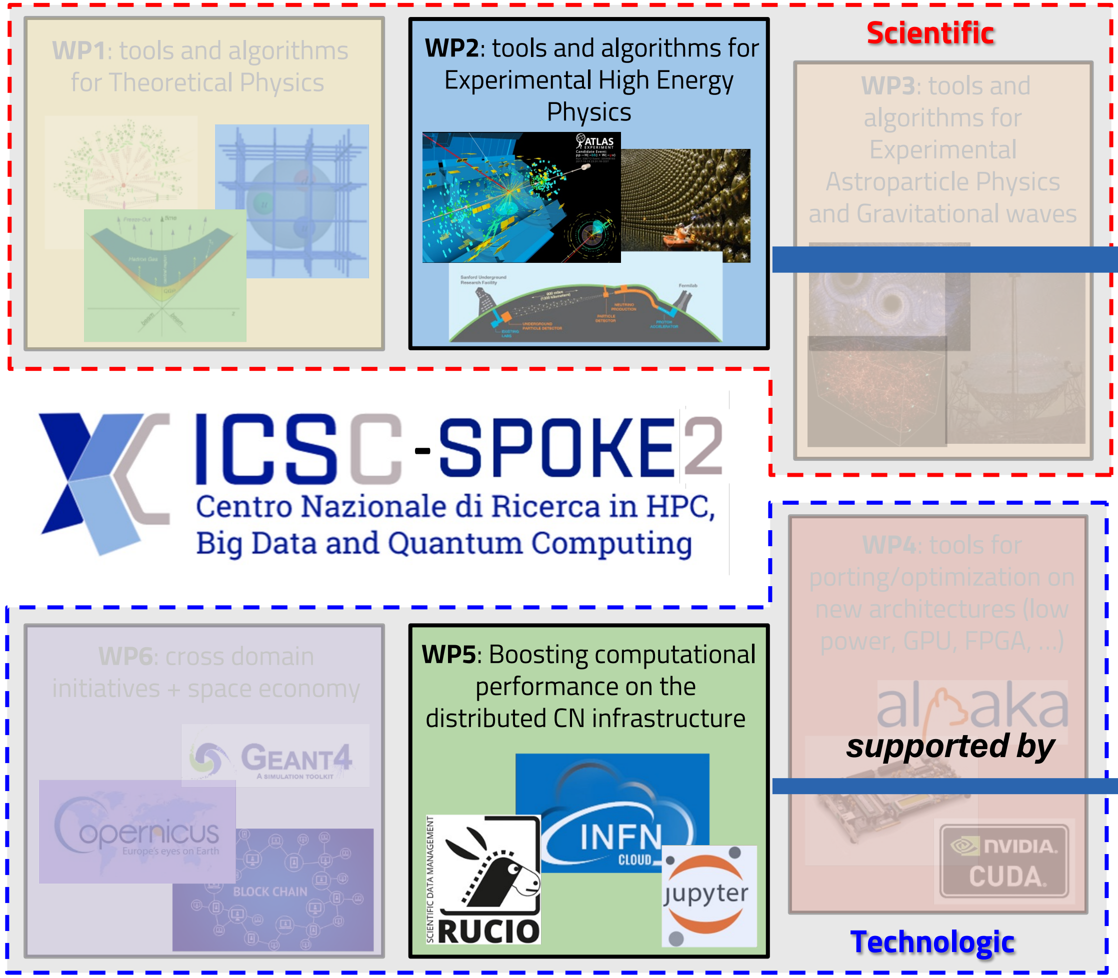
The structure of Spoke 2



- 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.



Quasi interactive analysis of Big Data with high throughput



1 of the 19 Spoke2 <u>flagship use cases:</u>

UC2.2.2

Finanziato dall'Unione europea NextGenerationEU — Ministero dell'Università e della Ricerca — Italiadomani — ICSC

Quasi interactive analysis of big data with high throughput

Spoke	2
WP	2, 5
Use case short name	Quasi interactive analysis of big data with high throughput
Use case ID	UC2.2.2
Expected Completion	31/8/2025

Approval workflow

Status	Version	Date	Submitter	Note	Signature
Draft	1.0	03/07/23	WP Leaders	First version	
Final Version	1.1	1/9/2023	WP Leaders		
Approved by Spoke Leaders	1.1	11/9/2023	Spoke Leaders		

Principal Investigators:

ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA

Tommaso Diotalevi

UNIVERSITÀ DEL SALENTO

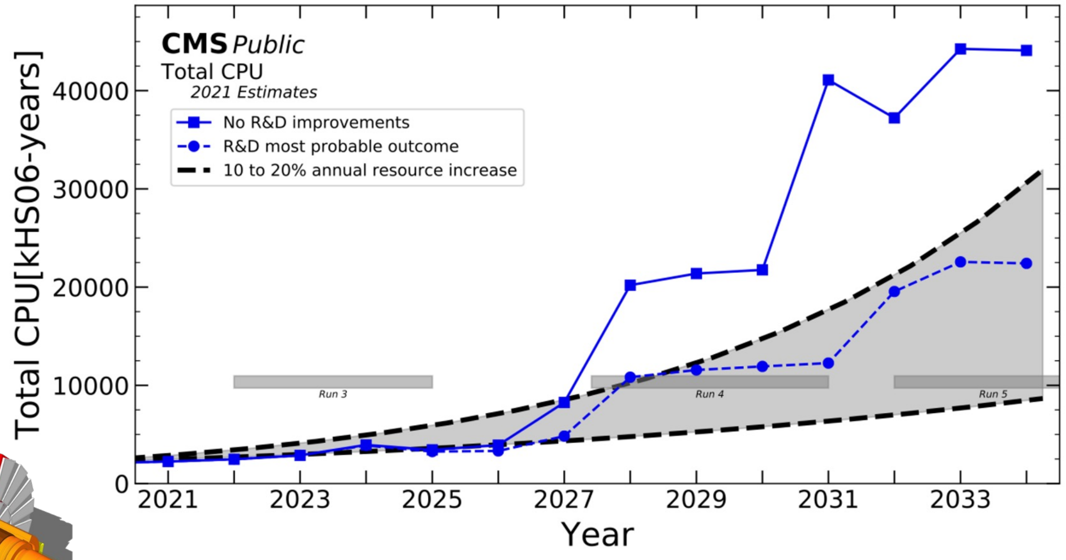
Francesco G. Gravili

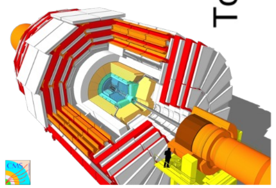
More than 40 people involved in this activity!

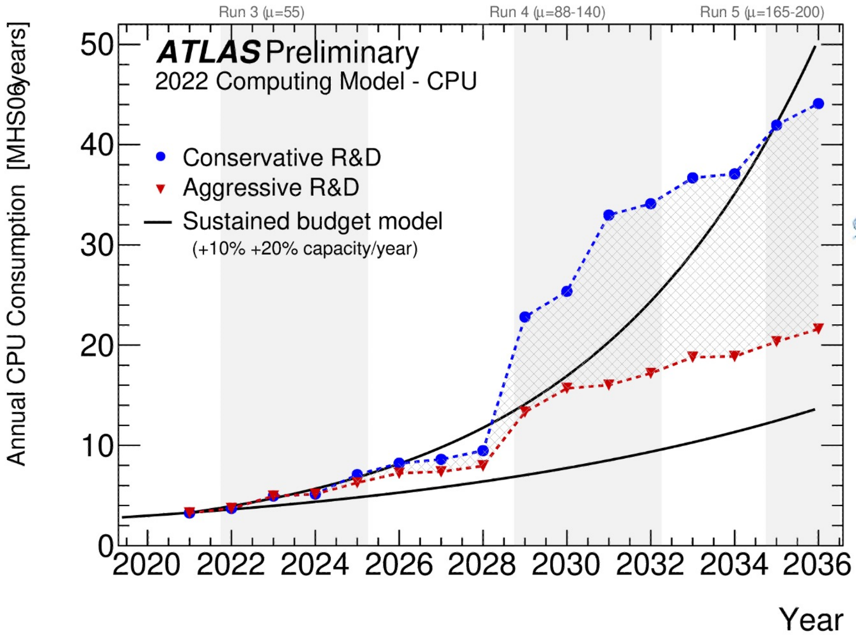
Motivation

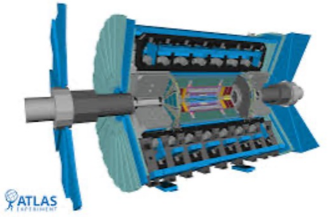
- Analysing large amounts of data efficiently, exploiting the available resources as much as possible, is a common challenge both for research and industry.

- From the beginning, the High Energy Physics (HEP) experiments at CERN, gave much attention to the computing and data management aspects. Nevertheless, the **next phases of the Large Hadron Collider** (HL-LHC) will require an even greater effort.









Motivation

Some estimate for the next 5-10 years of CMS operation:

- ~30 Billion collision events + 30 Billion simulation events;
- Each event: 2-4 kB;
- The last update of the CMS Computing model foresees this throughput:

Name	Length	% of the dataset	Data to process	Event, data rate
"A coffee"	< 5 min	1% (~0.6B evts)	~2 TB	~1.7MHz, ~7GB/s
"A lunch break"	1 hour	10% (~6B evts)	~20 TB	~1.5MHz, ~6GB/s
"A night"	12 hours	100% (60B evts)	~200 TB	~1.2MHz, ~5GB/s

- Difficult to get more than 100 Hz/CPU core → needs efficient distribution on a few tens of machines;

Not only concerning the HEP domain ("Data is data"):

- More and more scientific / industrial / societal domains have or will have soon needs similar to those from LHC:

DUNE DEEP UNDERGROUND NEUTRINO EXPERIMENT — ProtoDune: 2-3GB/s (like CMS); Real Dune: 80x

SKA SQUARE KILOMETRE ARRAY — SKA: up to 2 PB/day;

CTAO — CTA projects: up to 10PB/y

A single genome: ~100GB, a 1M survey=100PB

O(50 TB/y) per sensor; ~10-100 sensors: O(5 PB/y)

New analysis paradigm based on:

- Declarative programming and interactive workflows;
- Distributed computing on geographically separated resources.

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

Motivation

Some estimate for the next 5-10 years of CMS operation:

- ~30 Billion collision events + 30 Billion simulation events;
- Each event: 2-4 kB;
- The last update of the CMS Computing model foresees this throughput:

Name	Length	% of the dataset	Data to process	Event, data rate
"A coffee"	< 5 min	1% (~0.6B evts)	~2 TB	~1.7MHz, ~7GB/s
"A lunch break"	1 hour	10% (~6B evts)	~20 TB	~1.5MHz, ~6GB/s
"A night"	12 hours	100% (60B evts)	~200 TB	~1.2MHz, ~5GB/s

- Difficult to get more than 100 Hz/CPU core → needs efficient distribution on a few tens of machines;

Not only concerning the HEP domain ("Data is data"):

- More and more scientific / industrial / societal domains have or will have soon needs similar to those from LHC:

DUNE DEEP UNDERGROUND NEUTRINO EXPERIMENT — ProtoDune: 2-3GB/s (like CMS); Real Dune: 80x

SKA SQUARE KILOMETRE ARRAY — SKA: up to 2 PB/day;

CTAO — CTA projects: up to 10PB/y

A single genome: ~100GB, a 1M survey=100PB

O(50 TB/y) per sensor; ~10-100 sensors: O(5 PB/y)

New analysis paradigm based on:

- Declarative programming and interactive workflows;
- Distributed computing on geographically separated resources.

High Throughput Platform

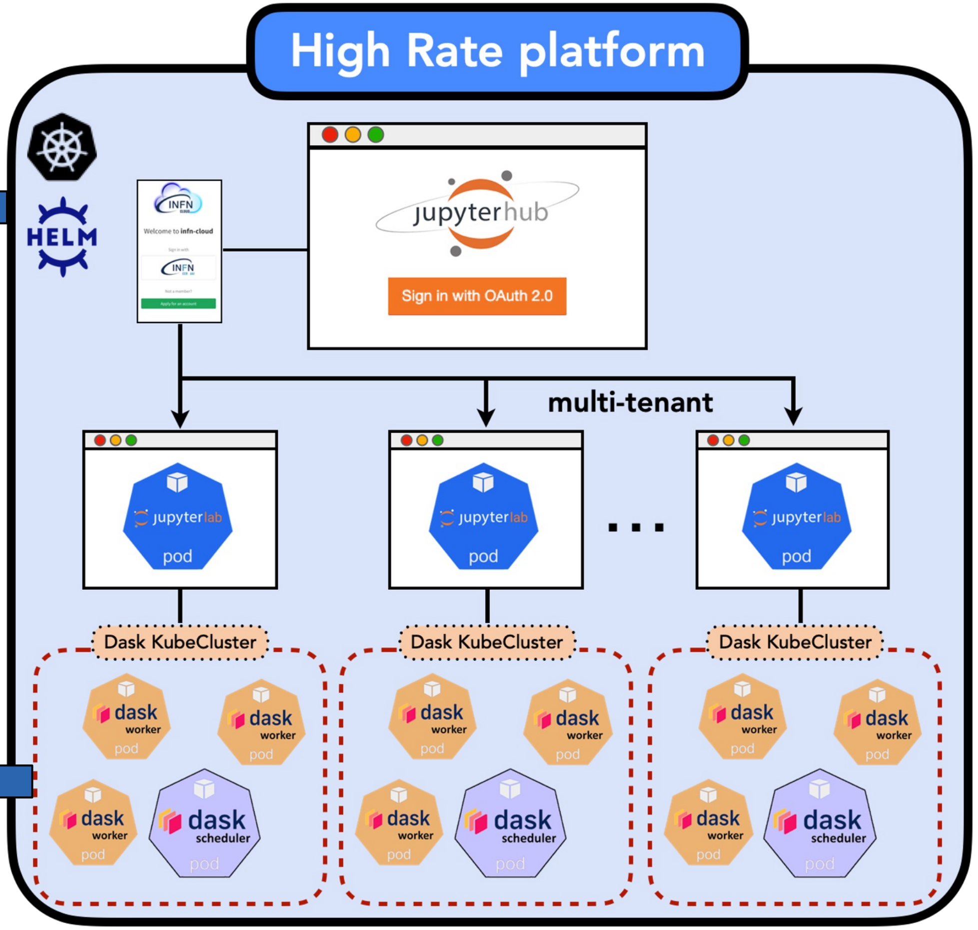
High Throughput platform

Deployment of the Kubernetes resources handled via HELM Charts.

↓

Scalable on available resources

- The execution happens in the Dask Cluster;
- Users choose on how many cores parallelly distribute the analysis.

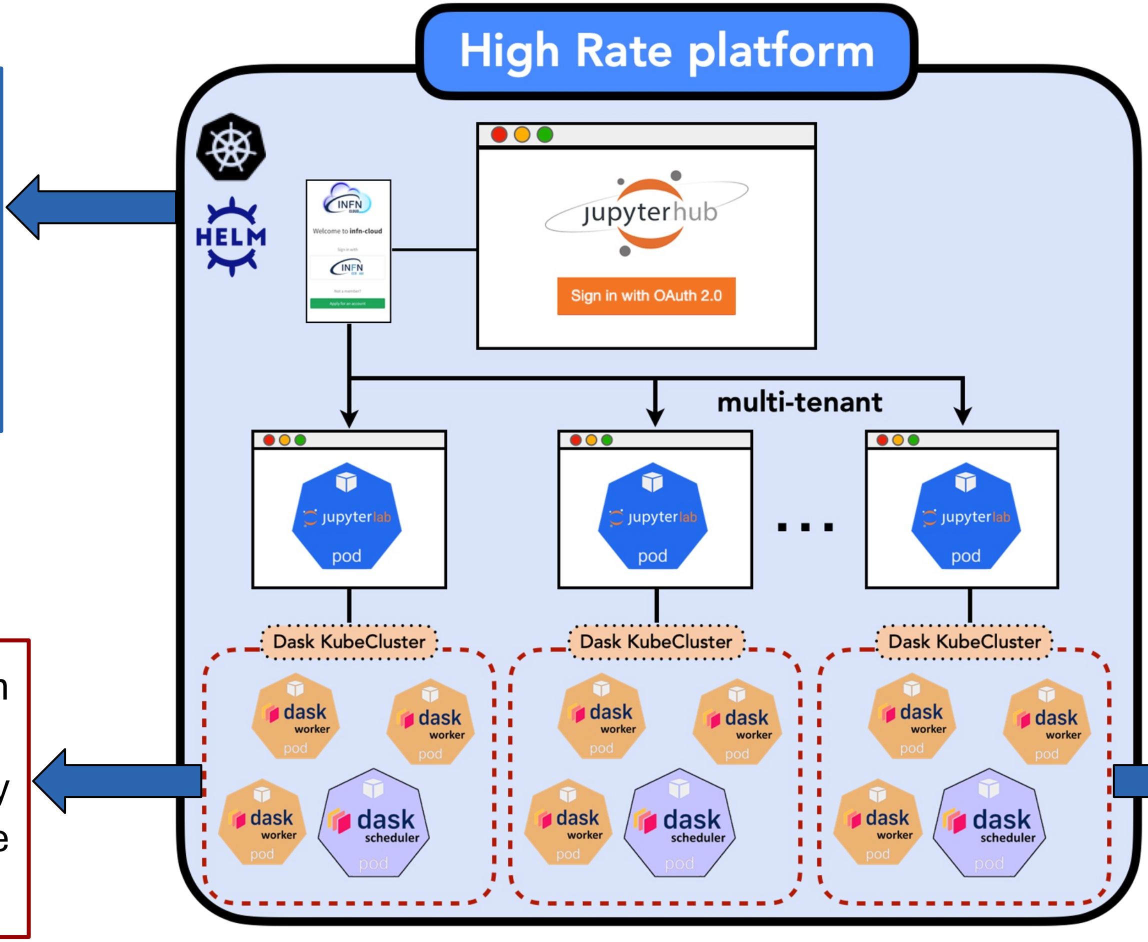


High Throughput platform in ICSC

Deployment of the Kubernetes resources handled via HELM Charts.

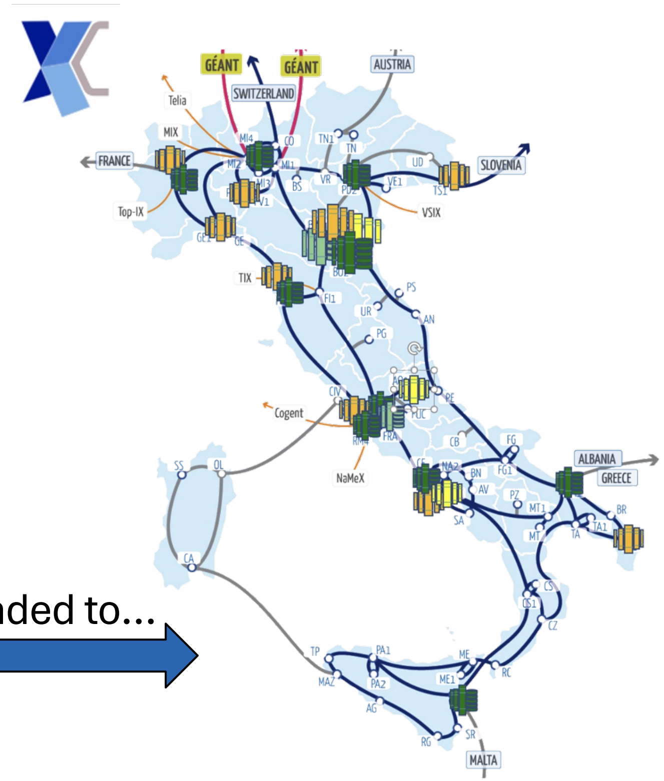
↓

Scalable on available resources



- The execution happens in the Dask Cluster;
- Users choose on how many cores parallelly distribute the analysis.

offloaded to...



Re-thinking the analysis pipeline

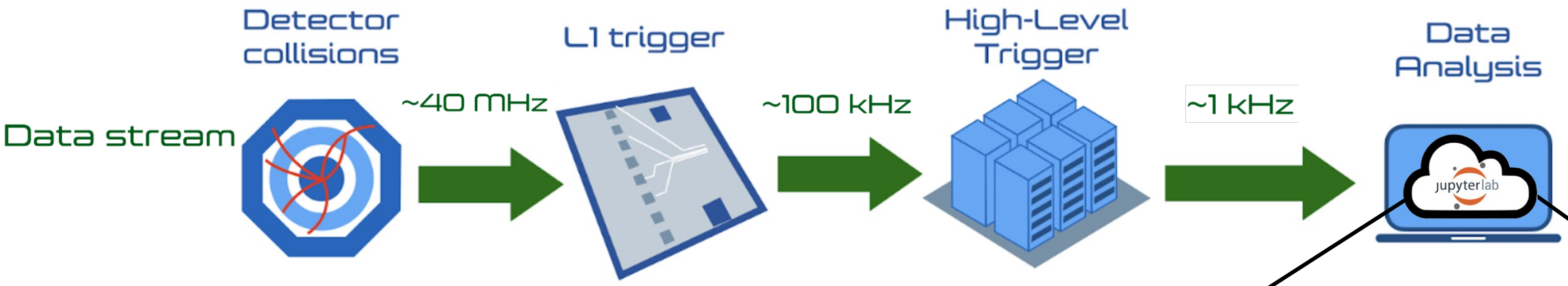
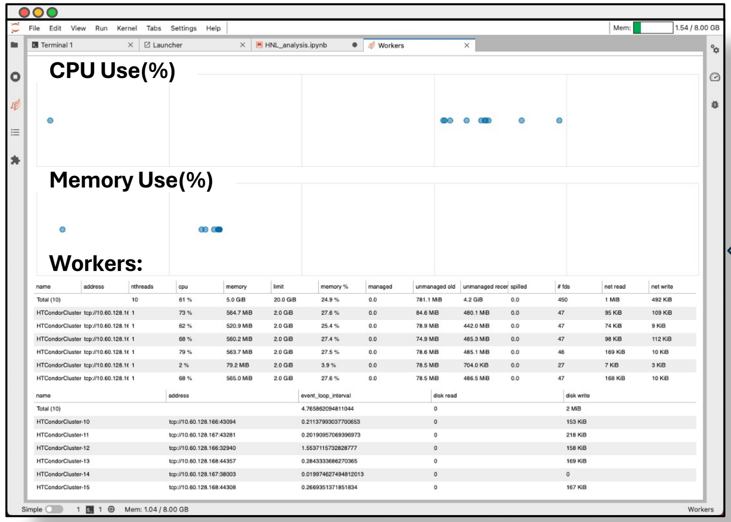
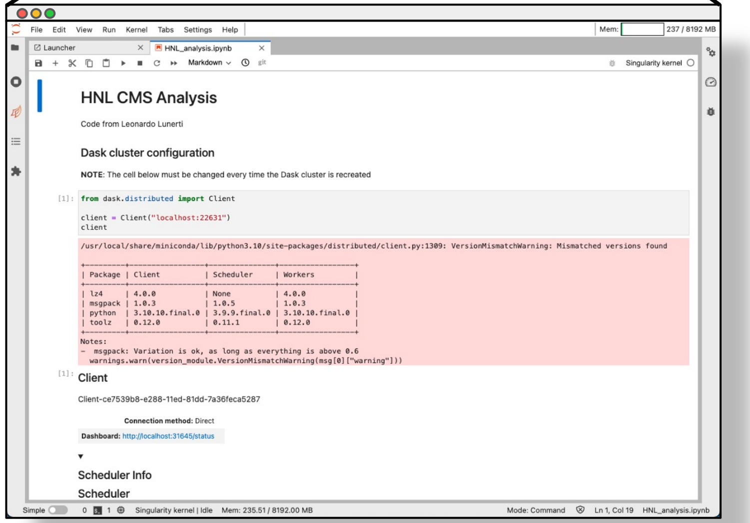


Image taken from: link

Resource monitoring dashboard



Analysis code



Orbiting activities

Vector Boson Scattering ssWW analysis in hadronic tau and light lepton

Heavy Neutral Lepton search on heavy neutrinos in the D_S decays

Muon detector performance analysis

Continuous Integration pipeline, triggering analysis execution on HTP

di-Higgs decaying to two b quarks and two muons

Search of rare events in tau to 3 muons decay

Differential cross section measurement for ttbar inclusive production

Search for new phenomena in events with two opposite-charge leptons, jets and missing transverse momentum

Neural Network hyperparameter optimisation applied to future colliders (FCC-ee)

Declarative paradigms for analysis description and implementation

top quark+MET analysis

Benchmark interactive analysis for future colliders (FCC-ee)

LIVE Next talk by S.Gasperini!

Distributing the Simulated Annealing workload for Quantum Unfolding in HEP

With the infrastructural support of WP5

Results presented @ICHEP2024

Some results

Muon detector performance analysis @ CMS

Typically, Detector Performance Group (DPG) analyses are run on a reduced amount of data (e.g. one run or fill), but processing of large dataset, at once, might be needed:

- To assess/improve systematics of high precision analyses, when they are dominated by the response of a specific detector;
- To reprocess multiple year data, e.g. for detector stability studies (ageing).

Use case:
Porting of a well established Drift Tubes (DT) Tag-and-Probe analysis [CMS-DP-2023-049]

A data sample consisting in a skim of $Z \rightarrow \mu\mu$ decay candidates collected by CMS over 2023, corresponding to ~27fb^{-1} was explored for the study. **Size: 224GB**

- To evaluate the technical performance, the available statistics has been processed 3 times, mimicking roughly a entire year of data taking. **Size**: 224*3 = **672GB**

Technical performance:

- Serial processing (as a single job on HTCondor)

Wall time: ~120 minutes

1 CPU on a AMD EPYC 7302 16-Core Processor, with 2GB memory

→

- Distributed processing on the platform:

Wall time: ~6 minutes

Up to 92 CPUs (46 physical), on two AMD EPYC 7413 24-Core Processor, with 2GB memory per CPU.

Conclusions

- The challenge presented by the next LHC phases requires a strong development effort of new tools, for making data analysis as efficient and as modern as possible;
- The «National Center for HPC, Big Data and Quantum Computing» is a unique opportunity for the creation of a modern infrastructure for research and industry in Italy;
- Aligned with CERN's R&D roadmap, a new **High Throughput Platform** has been developed:
  - Based on *interactive workflows* and on *declarative programming*;
  - Running on distributed resources (and heterogeneous).
- Several analysis from the HEP world are already testing such infrastructure, for performance measurements;
  - The final scale-up and stress tests will happen in the next few months.

Once fully operational, such platform will be used by the **entire ICSC community**, across scientific experiments and including industrial partners.

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