



Finanziato  
dall'Unione europea  
NextGenerationEU



Ministero  
dell'Università  
e della Ricerca



Italiadomani

PIANO NAZIONALE  
DI RIPRESA E RESILIENZA



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



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

## UC2.2.2 - Quasi interactive analysis of big data with high throughput

Tommaso Diotalevi (UniBO), Francesco G. Gravili (UniSalento)

WP2 Meeting - 25 March 2025

## ICSC Workshop on Analysis Facilities

- The workshop was held in Bologna, January 8th - 10th . [Agenda](#)
- A survey was opened (for the workshop participants) to gather feedbacks on the organisation and to collect possible new use-cases.
  - A few people replied to the survey, despite the attendance
  - Please, take some time to reply at [this link](#) (in particular, the survey part)
- We'll circulate a reminder in the upcoming days (on the mailing lists).

## Milestone 10 - Intermediate Report

- Milestone 10 is split in two parts
- The final version of the intermediate report for MS10 is available in the [workflow document](#).
- There are no intermediate KPIs associated to this milestone, since all the remaining objectives have been set for the end of project (MS10 final report).
- Dissemination questionnaire sent

## Conference opportunities for our flagship

- Contribution given at:
  - [ISGC 2025](#):
    - *"Quasi interactive analysis of High Energy Physics big data with high throughput"* (oral [contribution](#)) by *T.Diotalevi* and *F.G.Gravili*
    - *"Distributing the Simulated Annealing workload for Quantum Unfolding in HEP"* (oral [contribution](#)) by *S.Gasperini* (Unibo), *T.Diotalevi* and *M.Lorusso* (Unibo/CNAF) – A future presentation in WP2 will follow!



## KPI - Key Performance Indicator

KPI ID	Description	Acceptance threshold	Status
KPI2.2.2.1	Implementation of $N$ data analyses in the AF	$N \geq 2$	>100%
KPI2.2.2.2	Reference documentation of the AF	$\geq 1$ dedicated web site	100%
KPI2.2.2.3	Hands-on workshops for AF users	$\geq 1$ workshops	100%
KPI2.2.2.4	Scaling up the testbed AF infrastructure, serving $k$ tenants, for a total of $N$ data analyses	$\geq (200 \cdot N)$ cores	32% (1 of 2 batches)
KPI2.2.2.5	Talks at conferences/workshops about AF activities	$\geq 1$ talk	>100%

## KPI - Key Performance Indicator

KPI ID	Description	Acceptance threshold	Status
KPI2.2.2.1	Implementation of $N$ data analyses in the AF	$N \geq 2$	>100%
KPI2.2.2.2	Reference documentation of the AF	$\geq 1$ dedicated web site	100%
KPI2.2.2.3	Hands-on workshops for AF users	$\geq 1$ workshops	100%
KPI2.2.2.4	Scaling up the testbed AF infrastructure, serving $k$ tenants, for a total of $N$ data analyses	$\geq (200 \cdot N)$ cores	32% (1 of 2 batches)
KPI2.2.2.5	Talks at conferences/workshops about AF activities	$\geq 1$ talk	>100%

## Resource extension - Phase 2

- Following our [RAC request](#), the first batch of resources given are:

### 7.1 Resources granted by INFN-CLOUD (PaaS)

vCPU (number of vCores and requested allocation time)

Number	Time	vCORE	Notes
*	*	*	time in hours
4	17520	32	

This resulted in the creation of 8 VM with 16 VCore and 32GB of RAM each.

These resources are cloud provided via INFN Cloud, and handled in IaaS (not via PaaS, for now).

Storage: 50TB of volumes for users quotas.

## Resource extension - Phase 2

- Following our [RAC request](#), the second and last batch of resources will be:

### 7.2 Resources granted by INFN-GRID (Batch processing)

CPU (number of Cores and requested allocation time)

Number	Time	vCORE	Notes
*	*	*	time in hours
21	17520	32	

The total amount (672 equivalent grid/HTC cores) is not allocated statically, i.e. all at once.

After a first misunderstanding, an initial test will be done in Naples, using (and configuring) a few ( O(20) ) cores of the ICSC resources.

The strategy for the final up-scaling will be decided internally among sites (between Naples and CNAF), according to interested parties, being totally transparent for our use-case.



Proposal

## Summer Data Challenge

- During summer, we will start benchmarking the final deployment of the platform:
  - A **first phase** (June-July) will happen on a more technical aspect, performing tests on standard candle use cases and figuring out how to harvest metrics (resource and task usage, execution time, ...)  
In this phase, together with WP5, the platform will be finalised (including all missing aspects e.g. cvmfs, metrics collection, ...)
  - Then, a **second phase** (~ September) will follow, where a call for action will be issued for all flagship (or WP/other) users.  
Final goal: multiuser pressure tests (scalability and stability), monitoring resource consumption, heterogeneous benchmarking (on all the analysis “flavours”).

ideas?



# Thank you

[cn1-spoke2-wp2-analysisfacility@lists.infn.it](mailto:cn1-spoke2-wp2-analysisfacility@lists.infn.it)

(click [here](#) to subscribe or get in touch in case of issues)