



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

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

WP2 Meeting - 13 February 2024

Activities ongoing

In expansion...

CMS:

- Vector Boson Scattering ssWW analysis in hadronic tau and light lepton (*INFN e Univ. Perugia*);
- Heavy Neutral Lepton search on heavy neutrinos in the D_s decays (*INFN e Univ. Bologna*);
 - Studies for performance comparison speed-up, using in-site metrics and RDF monitoring, [talk WP5 \(3 Nov\)](#);
- Search of rare events in $\tau \rightarrow 3\mu$ (*INFN e Polit. Bari*). Explorative phase with RDF, currently porting the old ntupliser to flat data format.
- Continuous Integration pipeline, triggering analysis execution on AF (*INFN, Firenze*)
- top quark+MET analysis (*INFN e Univ. Napoli*). Still in an explorative phase, [talk WP5 \(24 May\)](#);

} [Joint WP2 meeting \(2 May\)](#)

ATLAS:

- Search for new phenomena in events with two opposite-charge leptons, jets and missing transverse momentum (*UniSalento e INFN Lecce*). Currently in a joint effort with *INFN e Univ. Napoli*.
- *UniCal*: possible applications within several subjects: top physics, flavour tagging, trigger analysis.

Future colliders:

- Feasibility studies started on FCCee pseudo-data (*INFN e Univ. Napoli*), exploiting Naples facility. Approach already adopting *Dask, Jupyter and ROOT RDataFrame*. talk [WP5 \(13 Oct\)](#).

LHCb:

- Survey still ongoing (planning to contact more institutions). If someone is interested to the UC activities, please contact us!

Milestone report documents

Milestone 7 (deadline 26 Feb):

- Section 2 TAR2.6 Intermediate report is almost complete. It is included in our [workflow document](#). Some final bullets still TBD, mainly concerning our today's agenda.

Milestones 8, 9, 10 (deadline 15 Feb):

- [Milestone 8](#) (Mar - Jun): confirmed the deliverable "prototypal central infrastructure on available resources".
- [Milestone 9](#) (Jul - Oct): confirming no deliverables.
- [Milestone 10](#) (Nov 24- Aug 25): confirming KPI deliverables.

News from WP5

- Before the holiday break, we left with a first facility deployment effort ongoing in WP5.
 - Configuration available on the official [ICSC-Spoke2](#) repository (maintained by WP5).
- An early-access version is now available, and users are starting to play and help developers with feedbacks and bug reports. [Talk WP5 \(2 Feb 2024\)](#)
 - **Entrypoint:** <https://hub.192.135.24.49.myip.cloud.infn.it>
- The creation of the documentation is now starting.

REPORT TODAY

In the meanwhile:

- First RAC allocated resources will be made available in the upcoming weeks. RAC details: [Talk WP2 \(5 Dec 2023\)](#)

Flagship Working Meetings

We started to organize working meetings, for better harmonizing the analysis efforts and define a common strategy for the performance evaluation.

- Dedicated agendas [growing here](#).

The first meeting (23 Jan) was focused on two main aspects:

- Showing a working demo of the newborn high rate platform ([link](#))
- Gathering ideas for benchmark metrics ([link](#))

Conference opportunities for our flagship

- We submitted an abstract for the [ACAT](#) workshop.
 - The contribution has been approved as a [poster contribution](#).
- A suggestion was given to participate also in the [IFAE](#) conference.
- Keep in mind that:
 - ICHEP [abstract submission](#) deadline is the 16th of February (if they don't extend it)
 - LHCP [abstract submission](#) deadline is the 15th of February (if they don't extend it - but only for poster session)

For every submission, please contact us for keeping track and also fill the [WP2 Tracking](#) document !

Ideas for 1st Workshop on UC2.2.2

- The planning of a “Quasi interactive analysis of Big Data with high throughput” workshop will begin in the next weeks.
 - Such an event is required, as a KPI (minimum one event);
 - Time window: second half of 2024 (max: beginning of 2025).
- Remote or in presence?
 - In presence events require a room booking in advance (especially in university buildings, due to exams and classes).
 - Non-trivial aspects regarding fundings (PNRR or local), if the event is hosted.

Any suggestions?



Thank you

cn1-spoke2-wp2-analysisfacility@lists.infn.it

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

KPI - Key Performance Indicator

These parameters, which are measurable objects, will reflect the final success of the entire use case.

KPI ID	Description	Acceptance threshold
KPI2.2.2.1	Implementation of N data analyses in the AF	$N \geq 2$
KPI2.2.2.2	Reference documentation of the AF	≥ 1 dedicated web site
KPI2.2.2.3	Hands-on workshops for AF users	≥ 1 workshops
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
KPI2.2.2.5	Talks at conferences/workshops about AF activities	≥ 1 talk