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

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

**WP2 Meeting - 23 April 2024**

ICSC Italian Research Center on High-Performance Computing, Big Data and Quantum Computing

Missione 4 • **Istruzione e Ricerca**

# Activities ongoing

**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*);

  } [Joint WP2 meeting (2 May '23)](#)
  - Studies for performance comparison speed-up, using in-site metrics and RDF monitoring, [talk WP5 (3 Nov '23)](#);
- Search of rare events in τ→3μ (*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 '23)](#);

**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,* [update talk WP5 (27 Mar '24)](#)
- *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 '23)](#).

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

# News from WP5

- 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 '24)

  - **Entrypoint**: https://hub.192.135.24.49.myip.cloud.infn.it

- The creation of the documentation already started, talk WP2 (13 Feb '24)

- Need to start discussing the Data Management activities/services. Get in touch if interested to contribute!

In the meanwhile:

- From RAC, our request has been approved! We are in contact now with INFN Cloud for access policies and entrypoints.

# Conference opportunities for our flagship

- Submitted abstracts:
    - [ACAT](#) workshop, approved as a [poster contribution](#) and already done.

    - [IFAE](#) conference, accepted as an [oral contribution](#) and already done.

    - [ICHEP](#) conference:

        - *"Benchmarking distributed-interactive HEP analysis workflows on the new Italian National Centre analysis infrastructure"*: accepted as poster;
        - *"Enhancing CMS data analyses using a distributed high throughput platform"*: accepted as oral.

    - [CHEP](#) conference:

        - *"The ICSC Analysis Facility for Interactive Computing in ATLAS SUSY searches"*: proposed to ATLAS Computing committee, awaiting confirmation.
        - CHEP abstract submission **deadline is 10th of May** (if they don't extend it), but be aware of exp. deadlines (e.g. CMS 28th April, ATLAS already expired);
        - More under discussion…

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

LATEST NEWS

# 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.

  - <u>Such an event is required</u>, as a KPI (minimum one event);

  - Time window: second half of 2024 (max: beginning of 2025).

  - Asked Bologna University for fundings. Waiting for funding decision.

    - If there are no objections, we can start proposing a time window to book the room.

  - Other suggestions or proposals are welcome!

# KPI - Key Performance Indicator

| KPI ID | Description | Acceptance threshold | Status |
|---|---|---|---|
| **KPI2.2.2.1** | Implementation of $N$ data analyses in the AF | $N \geq 2$ | >50% |
| **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 | 0% (in definition) |
| **KPI2.2.2.4** | Scaling up the testbed AF infrastructure, serving $k$ tenants, for a total of $N$ data analyses | $\geq (200 \cdot M)$ cores | 0% (waiting for resources) |
| **KPI2.2.2.5** | Talks at conferences/workshops about AF activities | $\geq 1$ talk | 100% (and even more!) |

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

# 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 <u>measurable</u> 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 | $\geq 1$ dedicated web site |
| **KPI2.2.2.3** | Hands-on workshops for AF users | $\geq 1$ workshops |
| **KPI2.2.2.4** | Scaling up the testbed AF infrastructure, serving $k$ tenants, for a total of $N$ data analyses | $\geq (200 \cdot M)$ cores |
| **KPI2.2.2.5** | Talks at conferences/workshops about AF activities | $\geq 1$ talk |