PROCEEDINGS OF SCIENCE

Quasi interactive analysis of High Energy Physics big data with high throughput

Tommaso Diotalevi[a,*] **and Francesco Giuseppe Gravili**[b]

[a] *University of Bologna / INFN*
[b] *University of Salento / INFN*
E-mail: tommaso.diotalevi@unibo.it,
francesco.giuseppe.gravili@unisalento.it

The ability to ingest, process, and analyze large datasets within minimal timeframes is a milestone of big data applications. In the realm of High Energy Physics (HEP) at CERN, this capability is especially critical as the upcoming high-luminosity phase of the LHC will generate vast amounts of data, reaching scales of approximately 100 PB/year. Recent advancements in resource management and software development have enabled more flexible and dynamic data access, alongside the integration with open-source tools like Jupyter, Dask, and HTCondor. These advancements facilitate a shift from a traditional "batch-like" processing to an interactive, high-throughput platform that utilizes a distributed, parallel back-end architecture. This approach is further supported by the DataLake model developed by the Italian National Center for "High-Performance Computing, Big Data, and Quantum Computing Research Centre" (ICSC). This contribution highlights the transition of various data analysis applications, from legacy batch processing to a more interactive, declarative paradigm using tools like ROOT RDataFrame. These applications are executed on the aforementioned cloud-based infrastructure, with workflows distributed across multiple worker nodes and results consolidated into a unified interface. Additionally, the performance of this approach is evaluated through speed-up benchmarks and scalability tests using distributed resources. Such analyses could help identify potential bottlenecks or limitations of the high-throughput interactive model, providing insights that will guide its further development and implementation within the Italian National Center.

International Symposium on Grids and Clouds (ISGC2025),
18th March 2025, Academia Sinica, Taipei, Taiwan

*Speaker

ISSN 1824-8039. Published by SISSA Medialab.

PoS(ISGC2025)027

1. Introduction

The exponential growth of data in scientific domains has introduced significant challenges in data processing and analysis. High Energy Physics (HEP) is at the forefront of this transformation, with experiments such as those at the CERN Large Hadron Collider (LHC) producing billions of collision and simulation events annually. With the advent of the High-Luminosity LHC (HL-LHC) [1], data volumes are expected to increase dramatically, reaching 100 PB/year [2]. Traditional computing models, primarily batch-oriented, face growing limitations in handling such workloads. This contribution focuses on a quasi-interactive analysis paradigm, enabled by high-throughput computing (HTC) platforms designed to enhance resource utilization and reduce latency in analysis cycles. The approach combines modern interactive tools, distributed infrastructures, and scalable software architectures to redefine how researchers interact with data. This work is presented within the broader context of the Italian National Center for High Performance Computing, Big Data and Quantum Computing (ICSC) [3], a flagship initiative funded through the European Union's NextGenerationEU program.

2. The ICSC Infrastructure

The ICSC is one of five major national centers established under Italy's National Recovery and Resilience Plan (NRRP) to invest in strategic technological domains. Coordinated by INFN and involving several institutions (such as the University of Bologna), the center is structured around three primary technological pillars:

- **GARR**: The Italian consortium responsible for high-speed connectivity for research and education among public data centers.
- **CINECA**: Home to the Leonardo supercomputer (ranked 9th in November 2024 Top500 list) and the newly appointed AI Factory IT4LIA (under the EuroHPC joint initiative).
- **INFN**: Enhancing the Worldwide LHC Computing Grid (WLCG) Tier-1 and Tier-2 infrastructures [4], acquiring cloud resources, and implementing DataLake middleware based on INFN Cloud [5].

The ICSC is based on a hub-spoke model: an Infrastructure hub and ten thematic spokes, each focused on different scientific and industrial application areas, ranging from quantum computing to smart cities. This contribution focuses on *Spoke 2*, entitled "Fundamental Research and Space Economy", addresses experimental, theoretical HEP and astro-particle physics, with an industrial involvement from the aerospace and satellite technology sectors.

2.1 Spoke 2: Organization and Objectives

Spoke 2 is jointly led by INFN (Italian Institute for Nuclear Physics) and INAF (Italian Institute for Astrophysics). It is internally organized into six Work Packages (WPs):

- WP1: Tools for theoretical physics;

- WP2: Tools for experimental HEP (focus of this contribution);
- WP3: Astroparticle and gravitational wave computing;
- WP4: Porting/optimization on novel architectures (GPU, FPGA);
- WP5: Performance optimization on national infrastructure;
- WP6: Cross-domain initiatives including industrial applications.

While the scientific work packages (WP1–WP3) present significant computational challenges, the technological work packages (WP4–WP6) offer targeted solutions, often contributing innovations that benefit a wider community, including industrial stakeholders.
One of the flagship initiatives in WP2 is the "Quasi-interactive analysis of High Energy Physics big data with high throughput", which also represents the focus of the present contribution: a collaborative, cross-experimental effort involving around 40 members and guided by its own set of deliverables and milestones.

3. Motivation and context

HEP computing has long been at the cutting edge of distributed computing. However, projections for CPU demand in the HL-LHC era reveal a pressing need to improve computational efficiency to stay within budgetary constraints. The CMS experiment [6], for example, anticipates handling over 60 billion events (30B collision + 30B simulated) over the next 5–10 years, and according to the CMS Computing Model [7] the following throughput levels are expected:

- 1% of the dataset (≈ 0.6B events, ≈ 3 TB) over a coffee time (<5 minutes);
- 10% of the dataset (≈ 6B events, ≈ 30 TB) over a lunch break (about 1 hour);
- Full dataset (≈ 60B events, ≈ 300 TB) over a night (about 12 hours).

These demands are not unique to HEP. Similar data throughput and processing issues are arising, e.g. in *ProtoDUNE* (2–3 GB/s) [8], *Square Kilometer Array* (up to 2 PB/day) [9], *Cherenkov Telescope Array* (up to 10 PB/year) [10].
This necessitates a shift to analysis paradigms rooted in declarative programming and distributed computing over geographically separated resources.

4. High Throughput Platform: Architecture and Workflow

To address these requirements, a scalable high-throughput platform has been developed within WP2 and WP5 of Spoke 2. Figure 1 presents a schematic overview of the proposed platform architecture. Users access the platform via a designated entrypoint URL, which directs them to a JupyterHub instance [11]. Authentication and authorization are managed through INDIGO-IAM [12], after which computing resources are dynamically allocated to establish the user's workspace. Users are then redirected to a JupyterLab interface [13], where they can develop and manage analysis

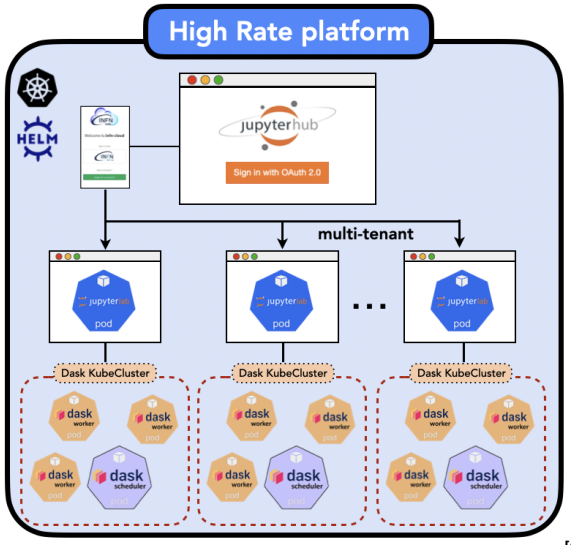

Figure 1: A sketched diagram of the proposed platform.

code. The environment is fully customizable via container-based technologies such as Docker [14] and Singularity/Apptainer [15]. Leveraging centralized services like CVMFS [16], users can store and retrieve container images on demand, enabling the use of domain-specific software (e.g., ROOT [17], plotting libraries, and statistical toolkits). The platform is deployed on a Kubernetes (K8s) cluster [18], with configuration provided through a HELM chart [19], ensuring flexible, scalable, and fault-tolerant operation with minimal administrative overhead.

On the back-end, workloads are transparently distributed across Dask clusters [20], with each worker deployed as a Kubernetes pod. This model offers near-zero configuration for end users and supports standard protocols like XRootD [21] and WebDAV [22] for data access.

4.1 From batch-based to interactive workflows

A key focus of this work is the migration of existing HEP data analysis workflows from the legacy batch-oriented paradigm (e.g HTCondor [23]) — where jobs are submitted, queued, and results retrieved only after long execution cycles — to an interactive high-throughput model. To achieve true interactivity, two aspects are fundamental: (i) an infrastructure capable of running user notebooks connected to a complex back-end based on distributed resources, and (ii) the adoption of modern software frameworks designed to exploit such infrastructures, such as ROOT RDataFrame [24], Coffea [25], and, more generally, high-performance libraries like NumPy and Dask Distributed [26]. From the user side, porting an analysis involves two main steps: first, restructuring the workflow into logical steps executable as Jupyter cells, which is relatively straightforward; second, rewriting selected parts of the logic into modular blocks that will run in parallel across the back-end cluster, which instead requires an initial learning phase (depending on the library used). While this adaptation introduces some overhead in early development, the benefits — faster re-execution, reduced execution times, and improved readability — quickly outweigh the initial effort. This process also encourages a shift towards a declarative style of analysis, where high-level data

transformations are specified by the user (in a *columnar* way, e.g. Pandas DataFrame [27]) while the execution details and optimization are handled transparently by the back-end.

5. Use Cases: From Detector Studies to Future Colliders

Several use cases are validating the platform [28–30], including both present and future experiments. These analyses are either already migrated or in the process of transitioning to such platform. An example is presented in the following subsection, as part of a personal contribution.

5.1 Muon Detector Performance Analysis

This use case focuses on a specific detector performance study: the Tag-and-Probe analysis [31] of the Drift Tubes (DT) muon sub-detector [32]. The dataset comprises a skim of $Z \to \mu\mu$ decay candidates collected by the CMS experiment in 2023, corresponding to an integrated luminosity of 27 fb^{-1}. Unlike standard physics analysis datasets, this sample also includes low-level detector information, resulting in a total size of approximately 224 GB. The original analysis code was implemented primarily in C++, where DT segments were reconstructed and efficiencies calculated. To run this workflow on the platform, the code was ported to Python and integrated into a Jupyter notebook using the ROOT RDataFrame framework. The overall structure of the legacy workflow is preserved: custom libraries and functions are defined in a dedicated C++ header, and objects are manipulated using ROOT vector library (in a . This enables a RDataFrame-based analysis that can leverage Dask as a back-end, allowing the entire workflow to be executed in a distributed manner, across multiple ICSC sites.

To assess the technical performance (presented in [33]), the available dataset was processed three times, simulating an integrated luminosity of approximately 82 fb^{-1} and resulting in a total of around 77 million events. This brought the dataset size to roughly 672 GB.
In the legacy setup, the analysis was executed as a single HTCondor job on one core of an AMD EPYC 7302 16-core processor, with 2 GB of allocated memory. This setup required a walltime of approximately 120 minutes.
In contrast, the distributed processing on the high-throughput platform was carried out on two AMD EPYC 7413 24-core processors[1]. The number of CPUs was progressively increased up to 92, each with 2 GB of memory, reducing the walltime to a minimum of about 6 minutes.

6. Outlook and Future Integration

As the platform matures, final stress tests and full-scale deployment are scheduled before the ICSC project concludes at the end of 2025. A central goal is to make this infrastructure widely available across academic and industrial sectors. While a comprehensive study of the potential bottlenecks of the high-throughput interactive model is still in progress and will be the focus of a future contribution, preliminary observations already indicate that network performance is a critical factor. In particular, the intensive I/O generated by workflows accessing datasets from remote WLCG sites has shown that bandwidth is often the most stressed component of the

[1]Resources provided by the Italian CMS Tier-2 center at Legnaro, with performance monitored via on-site metrics stored in a database.

infrastructure. This is especially evident in scenarios where large data volumes must be streamed from geographically distributed storage endpoints to multiple worker nodes in parallel. Such findings reinforce the importance of the ICSC's recent and planned investments in high-speed interconnects, improved network fabrics between sites, and optimized data access protocols.
At the same time, there are active efforts on integrating machine learning and AI applications, including Large Language Models (LLMs). While there is no direct link to the European AI Factory initiative at this stage, synergistic developments are ongoing.
The long-term vision includes extending support to heterogeneous resources and expanding use cases beyond HEP, into climate science, genomics, and industrial monitoring.

7. Conclusions

The upcoming HL-LHC phase presents unprecedented challenges in data processing. An high-throughput interactive analysis paradigm offers a viable, modern solution. It enables faster analysis, reproducibility, and better integration with cloud-native technologies.
With support from the ICSC national-scale infrastructure and in alignment with CERN's R&D roadmap, the platform is poised to serve a wide range of scientific and industrial users. The preliminary benchmarks and use cases already demonstrate its transformative potential.

Acknowledgements

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

References

[1] Apollinari, G. et al. "High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1". https://doi.org/10.23731/CYRM-2017-004

[2] Elsen, E. "A Roadmap for HEP Software and Computing R&D for the 2020s". https://doi.org/10.1007/s41781-019-0031-6

[3] ICSC main page. https://www.supercomputing-icsc.it/

[4] Bird, I. "Computing for the Large Hadron Collider". https://doi.org/10.1146/annurev-nucl-102010-130059

[5] INFN Cloud webpage: https://www.cloud.infn.it/

[6] "The CMS experiment at the CERN LHC". https://doi.org/10.1088/1748-0221/3/08/s08004

[7] "CMS Phase-2 Computing Model: Update Document". https://cds.cern.ch/record/2815292.

[8] Acciarri, R. et al. "Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE): Conceptual Design Report, Volume 2: The Physics Program for DUNE at LBNF". https://arxiv.org/abs/1512.06148

[9] Weltman, A. et al. "Fundamental physics with the Square Kilometre Array". https://doi.org/10.1017/pasa.2019.42

[10] Actis, M. et al. "Design concepts for the Cherenkov Telescope Array CTA: An advanced facility for ground-based high-energy gamma-ray astronomy". https://doi.org/10.1007/s10686-011-9247-0

[11] JupyterHub main page. https://jupyter.org/hub

[12] Indigo-IAM main page. https://indigo-iam.github.io/v/current/

[13] JupyterLab main page. https://jupyter.org/lab

[14] Docker main page. https://www.docker.com/

[15] Apptainer/Singularity main page. https://apptainer.org/

[16] Buncic, P. et al., "CernVM: A virtual software appliance for LHC applications". https://doi.org/10.1088/1742-6596/219/4/042003

[17] Brun, R. and Rademakers, F., "ROOT: An object oriented data analysis framework". https://doi.org/10.1016/S0168-9002(97)00048-X

[18] Kubernetes main page. https://kubernetes.io/

[19] Helm main page. https://helm.sh/

[20] Dask: Library for dynamic task scheduling. http://dask.pydata.org

[21] XRootD protocol main page. https://xrootd.slac.stanford.edu/

[22] Web Distributed Authoring and Versioning (WebDAV) Redirect Reference Resources. https://doi.org/10.17487/RFC4437

[23] HTCondor main page. https://htcondor.org

[24] ROOT RDataFrame main page. https://root.cern/doc/master/classROOT_1_1RDataFrame.html

[25] Coffea: Smith, N. et al. "Columnar Object Framework For Effective Analysis". https://cds.cern.ch/record/2798133

[26] Dask distributed main page. https://distributed.dask.org/en/stable/

[27] Pandas main page. https://pandas.pydata.org/docs/index.html

[28] Tedeschi, T. et al. "Prototyping a ROOT-based distributed analysis workflow for HL-LHC: The CMS use case". https://doi.org/10.1016/j.cpc.2023.108965

[29] Bartolini, M. et al. "Quasi interactive high throughput analysis of high energy physics data". https://doi.org/10.1393/ncc/i2025-25107-1

[30] D'Onofrio, A. et al. "Benchmarking distributed-interactive HEP analysis workflows on the new Italian National Centre analysis infrastructure". https://doi.org/10.22323/1.476.1043

[31] "Drift Tube Performance in 2023". https://cds.cern.ch/record/2868786.

[32] CMS Collaboration. "The CMS muon project: Technical Design Report". https://cds.cern.ch/record/343814

[33] Diotalevi, T. et al. "Enhancing CMS data analyses using a distributed high throughput platform". https://doi.org/10.22323/1.476.1007

PoS(ISGC2025)027