EPJ Web of Conferences **337**, 01035 (2025) https://doi.org/10.1051/epjconf/202533701035
*CHEP 2024*

# Leveraging distributed resources through high throughput analysis platforms for enhancing HEP data analyses

*Adelina* d'Onofrio[1,*], *Tommaso* Diotalevi[1,4], *Francesco G.* Gravili[1,5], *Salvatore* Loffredo[1,3], *Elvira* Rossi[1,3], *Federica M.* Simone[1,2,**], and *Bernardino* Spisso[1] on behalf of the ATLAS and CMS collaborations

[1]INFN
[2]Polytechnic Bari
[3]University Federico II
[4]University of Bologna
[5]Università del Salento

**Abstract.** The analysis of data collected by the ATLAS and CMS experiments at CERN, ahead of the next phase of high-luminosity at the LHC, requires flexible and dynamic access to big amounts of data, as well as an environment capable of dynamically accessing distributed resources. An interactive high throughput platform, based on a parallel and geographically distributed backend, has been developed in the framework of the "Italian Research Center for High Performance Computing, Big Data, and Quantum Computing" (ICSC), providing experiment-agnostic resources. Starting from container technology and orchestrated via Kubernetes, the platform provides analysis tools via the Jupyter interface and Dask scheduling system, masking complexity for frontend users and rendering cloud resources flexibly. An overview of the technologies involved and the results on benchmark use cases will be provided, with suitable metrics to evaluate the preliminary performance of the workflow. The comparison between the legacy analysis workflows and the interactive and distributed approach will be provided based on several metrics, from event throughput to resource consumption. The use cases include the search for direct pair production of supersymmetric particles and for dark matter in events with two opposite-charge leptons, jets and missing transverse momentum using data collected by the ATLAS detector in Run 2, and searches for rare flavor decays at the CMS experiment in Run 3 using large datasets collected by high-rate dimuon triggers.

## 1 Introduction

The High-Luminosity Large Hadron Collider (HL-LHC) will deliver unprecedented data volumes, posing significant challenges for data processing and analysis at the ATLAS and CMS experiments. With a dramatic increase in event rates and data size, scalable high-throughput analysis facilities are essential for handling petabyte-scale datasets. Efficient data processing, storage, and access will be crucial for timely physics results, requiring a shift toward distributed computing infrastructures, heterogeneous architectures, and scalable analysis frameworks.

*e-mail: adelina.d'onofrio@cern.ch
**e-mail: federica.maria.simone@cern.ch

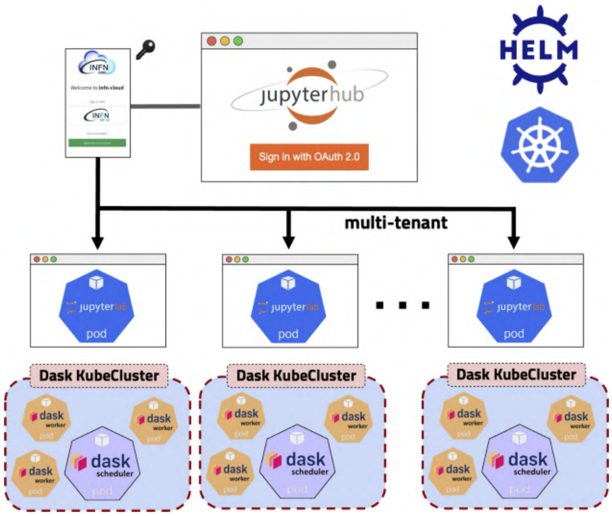


**Figure 1.** Schematic representation of the High Rate Data Analysis Platform

To address these challenges, the "Italian Research Center for High Performance Computing, Big Data, and Quantum Computing" (ICSC), funded by the European Union through NextGenerationEU, has developed a High-Rate Data Analysis Platform. This platform integrates cloud resources with advanced data solutions, enabling physicists to analyze large datasets efficiently using parallel workflows on distributed clusters. This contribution provides a technical overview of the platform, along with two use cases from the ATLAS [1] and CMS [2] collaborations.

## 2 Description of the Analysis Platform

The High-Rate Data Analysis Platform, illustrated in Figure 1, is built on a Kubernetes-based infrastructure [3], providing a scalable, multi-tenant environment for high-performance data analysis.

Users access the platform through JupyterHub [4], with authentication managed via Indigo-IAM [5] on the INFN Cloud [7]. Once logged in, each user is assigned an isolated JupyterLab [6] instance within a Kubernetes pod, ensuring resource separation and scalability.

The JupyterLab interface supports a wide range of plugins, including Dask [8] for real-time monitoring and parallel task execution. The environment is fully customizable through Docker [9] containers, allowing users to deploy experiment-specific software and dependencies. Infrastructure management is handled via Helm [10], making the platform adaptable to future HL-LHC demands.

The computational tasks are offloaded within the same Kubernetes cluster using Dask KubeCluster [11], ensuring seamless integration and efficient task distribution. Current efforts focus on expanding the platform's capabilities to support multiple remote sites, integrating heterogeneous resources from High-Throughput Computing (HTC), HPC, and cloud environments.

# 3 Analysis use cases

In order to evaluate the performance of the platform, two analysis use cases from the ATLAS and CMS are chosen, sharing the need for interactively analyzing big amounts of data.

## 3.1 The ATLAS use case: a search for new Physics with the ATLAS experiment

Preliminary studies on ATLAS Run 2 data have demonstrated the platform's potential to reduce data processing time compared to traditional serial workflows. As a benchmark, previously published ATLAS search [12] is used, focusing on top squark pair production and dark matter in events with two opposite-charge leptons, jets, and missing transverse momentum, using 139 $\text{fb}^{-1}$ of integrated luminosity from proton-proton collisions at $\sqrt{s} = 13$ TeV.

Figure 2 (left) shows one of the signal models of interest, featuring the four-body top squark decays, which has been selected for testing the analysis workflow on the platform. The 95% CL exclusion limit contour for a simplified super-symmetric model assuming the four-body top squark decays is shown in Figure 2 (right) [12].

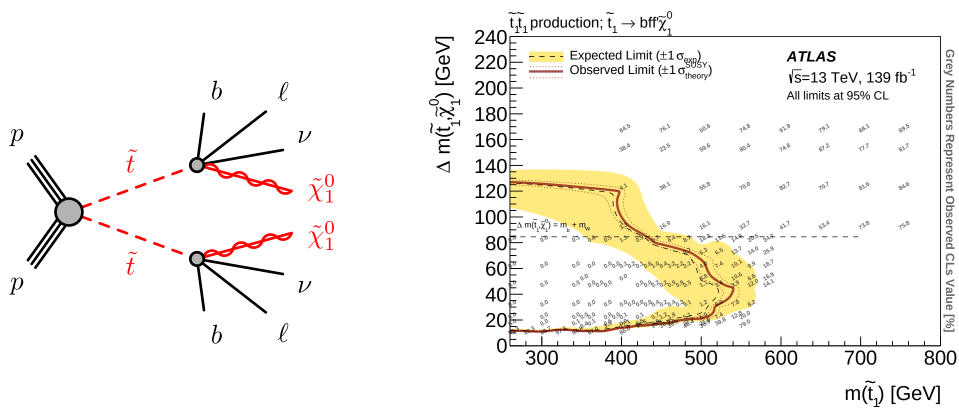


**Figure 2.** Feynman diagram representing one of the signal models targeted by the ATLAS SUSY search (left) and the results obtained [12] (right).

### 3.1.1 Analysis workflow

The analysis workflow begins with standard ATLAS raw data formats for both data and Monte Carlo simulations, that undergo several reprocessing and event reconstruction steps up to an analysis object data format available to the users. The process outlined above is centralized and represents a standard procedure for all ATLAS analyses.

Custom event selections are applied to retain only the physically interesting events, reducing the dataset size from the order of $O$(PB) to the order of $O$(GB). These final steps of the data workflow are summarized in Figure 3, where the analysis platform plays a key role in the rightmost boxes.

The Event Selection step was based on the simple cut&count approach and it was the core of our feasibility studies: ROOT RDataFrame [13, 14] and Dask were applied to a $W$-boson plus top quarks background sample, corresponding to ∼ 1.8 GB of data. According to the SUSY search requirements, several systematic variations were applied to test the code's overall execution time over hundreds of iterations of the event selection step.

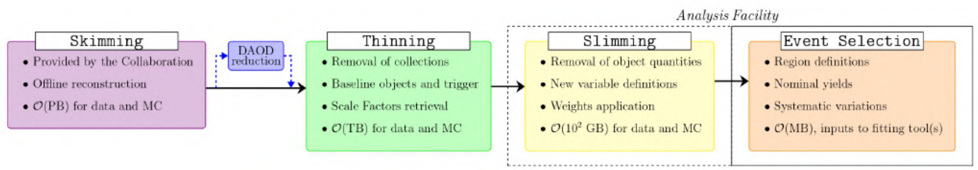


**Figure 3.** Final steps of the ATLAS data workflow, where DAOD stands for derivation-level analysis object data that include event selections already specific for the target signal process.

### *3.1.2 Performance on the analysis facility*

Preliminary results obtained running the same analysis workflow and considering the overall execution time as a metric are shown in Figure 4, highlighting a performance improvement in the distributed approach with respect to the standard/serial approach. Exploiting the distributed approach, the total execution time improves by a factor of 5 with respect to the standard/serial approach when iterating over a significative number of systematic variations. Moreover, other parameters, such as the event throughput or the resource consumption, were studied to assess the reliability of the workflow using resources hosted at the ICSC distributed analysis facility.

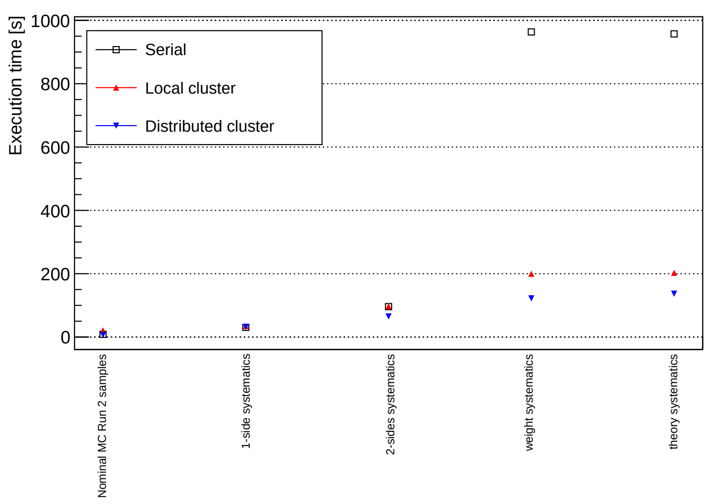


**Figure 4.** Execution time of the analysis workflow, including several systematic variations (to be noticed that each step in the x-axis includes the contributions from the previous bins). Black squares are the result of the standard/serial approach, while red and blue triangles refer to the parallel approach. *LocalCluster*: Dask multi-thread execution on the local machine (max 8 cores, 16 GB). *Distributed*: Dask distributed execution on remote workers.

Considering the Dask distributed execution, the CPU consumption per working node is of the order of 5%, due to the easy cut&count operations performed, opening up the possibility of remarkable improvements when including all other background sources in future iterations.

### 3.2 Benchmark flavor physics analysis at the CMS experiment

The high rate platform is also tested on a CMS physics data analysis, specifically in the field of the search for lepton flavor violating $\tau^+ \to \mu^+\mu^-\mu^+$ decays, where $\tau$ leptons are produced in the decays of D and B mesons, featuring final-state muons with low transverse momentum ($p_T$). The sensitivity of the CMS experiment to $\tau \to 3\mu$ decays is driven by the available statistics. While this search has been successfully performed using Run 2 data [15], the effort continues on Run 3 data benefiting from innovative trigger strategies including inclusive low-$p_T$ muon triggers [16]. The increased statistics come with heavier datasets, thus calling for a new implementation of the analysis to exploit the ICSC analysis facility. To evaluate the performance of the platform, $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ decays are used, serving as control and normalization channel in the $\tau \to 3\mu$ analysis.

#### *3.2.1 Analysis workflow*

Events collected by dedicated triggers requiring two low-$p_T$ muons and one charged track are used to reconstruct $D_s^+ \to \phi(\mu^+\mu^-)\pi^+$ decays in pp collision data. The analysis uses the "Mini-AOD" analysis object format, further compressed into ROOT flat ntuples after selecting events containing $\mu^+\mu^-\pi^+$ candidates with an invariant mass compatible with the $D_s^+$ meson.

Additional event selections, object calibrations and corrections are applied. If more than one $\mu^+\mu^-\pi^+$ candidate per event is found, the ambiguity is resolved by selecting the candidate with the higher quality common vertex. The $\mu^+\mu^-\pi^+$ invariant mass is then used to measure the yield of produced $D_s^+$ in data and simulation. The mass distribution in data from the Run 2 analysis is shown in Figure 5 for reference.

In the legacy analysis workflow, the ROOT ntuples were analyzed by splitting the computation into batches of input files, each run as an HTCondor job. The redesigned analysis reads the whole set of ntuples as a RDataFrame object, where all the selections and analysis steps are implemented without explicitly looping over the events. The RDataframe interface to Dask is used to distribute the computation in a transparent way for the user.

#### *3.2.2 Performance on the analysis facility*

The performance of the platform for this analysis use case is measured in terms of CPU time, first comparing the "non-parallel" execution with the execution on Dask (Figure 6 left).

The facility allows resources to be scaled dynamically. Therefore the total execution time is measured for different numbers of workers (or virtual cores), each being assigned 2 GB of memory. Moreover, the size of the input dataset is varied, ranging from about 10 to 130 GB. Figure 6 (right) shows stable performance, linearly scaling with the input dataset size, indicating full exploitation of the allocated resources and opening up for future tests on heavier datasets.

## 4 Conclusions

The High-Rate Data Analysis Platform developed within the ICSC framework provides a scalable, flexible environment for distributed data analysis. Preliminary results from ATLAS and CMS use cases show significant performance improvements compared to traditional

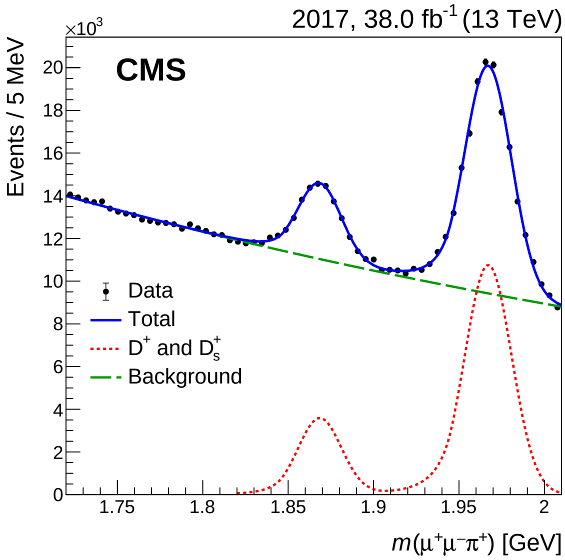


**Figure 5.** The $\mu^+\mu^-\pi^+$ invariant mass distribution with the fits to the sum of the $D^+$ (1.870 GeV) and $D_s^+$ (1.968 GeV) resonances and the background in 2017 data [16].

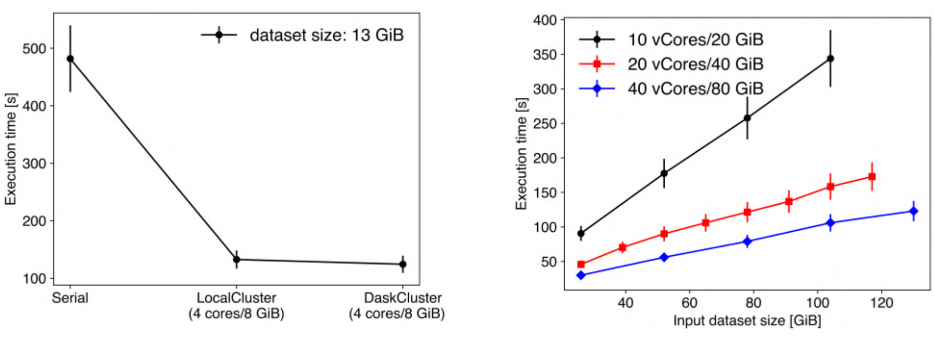


**Figure 6.** Execution time of the analysis workflow. Left: Comparison between the standard serial execution and the distributed execution using Dask across local and remote resources. Right: Execution time on a Dask cluster as a function of input dataset size, evaluated for different resource allocations.

workflows. Future developments will focus on integrating heterogeneous resources and expanding the platform's capabilities to meet HL-LHC demands.

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] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," JINST **3** (2008), S08003, doi:10.1088/1748-0221/3/08/S08003

[2] CMS Collaboration, "The CMS Experiment at the CERN LHC," JINST **3** (2008), S08004, doi:10.1088/1748-0221/3/08/S08004

[3] The Kubernetes Authors, *Kubernetes: Production-Grade Container Orchestration*, 2014. Available at: https://kubernetes.io.

[4] Jupyter Development Team, *JupyterHub: Multi-user Jupyter Notebooks*, 2015. Available at: https://jupyter.org/hub.

[5] A. Ceccanti, E. Vianello and F. Giacomini, "Beyond X.509: Token-based authentication and authorization in practice," EPJ Web Conf. **245** (2020), 03021, doi:10.1051/epjconf/202024503021

[6] Jupyter Development Team, *JupyterLab: The Next Generation Interface for Project Jupyter*, 2018. Available at: https://jupyter.org.

[7] INFN, *INFN Cloud services*, 2021. Available at: https://www.cloud.infn.it/#.

[8] Dask Development Team, *Dask: Parallel Computing with Task Scheduling*, 2016. Available at: https://dask.org.

[9] Dirk Merkel, *Docker: lightweight Linux containers for consistent development and deployment*, Linux J. 2014, 239, Article 2 (March 2014).

[10] The Helm Authors, *Helm: The Kubernetes Package Manager*, 2015. Available at: https://helm.sh.

[11] Dask Development Team, "Dask Kubernetes Operator Documentation," 2025. Available at: https://kubernetes.dask.org/.

[12] ATLAS Collaboration, "Search for new phenomena in events with two opposite-charge leptons, jets and missing transverse momentum in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector," JHEP, 04 (2021) 165, doi:10.1007/JHEP04(2021)165

[13] R. Brun and F. Rademakers, *ROOT: A Data Analysis Framework*, Nucl. Instrum. Meth. A, vol. 389, pp. 81-86, 1997. doi:10.1016/S0168-9002(97)00048-X.

[14] T. Tedeschi, V. E. Padulano, D. Spiga, D. Ciangottini, M. Tracolli, E. Tejedor Saavedra, E. Guiraud and M. Biasotto, "Prototyping a ROOT-based distributed analysis workflow for HL-LHC: The CMS use case," Comput. Phys. Commun. **295** (2024), 108965, doi:10.1016/j.cpc.2023.108965

[15] CMS Collaboration, "Search for the lepton flavor violating $\tau \rightarrow 3\mu$ decay in proton-proton collisions at s=13TeV," Phys. Lett. B **853** (2024), 138633, doi:10.1016/j.physletb.2024.138633

[16] CMS Collaboration, "Enriching the Physics Program of the CMS Experiment via Data Scouting and Data Parking," [arXiv:2403.16134 [hep-ex]].