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

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

Adelina D'Onofrio[1], Tommaso Diotalevi[1,3], Francesco Giuseppe Gravili[1,5], Salvatore Loffredo[1,2], Elvira Rossi[1,2], Federica Maria Simone[1,4], Bernardino Spisso[1]

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

On behalf od ATLAS and CMS Collaborations
CHEP2024, 19-15 Oct 2024, Krakow

Talk info

https://indico.cern.ch/event/1338689/contributions/6010647/

CMS+ATLAS talk

15 min + questions

ICSC is mentioned in the abstract and slides have the official template: anything else? Acknowledgments?

Many thanks to Adele for preparing the presentation draft!!

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

23 Oct 2024, 16:51
18m
Room 5

Talk | Track 9 - Analysis fa... | Parallel (Track 9)

Speakers

Adelina D'Onofrio (INFN Napoli (IT))
Federica Maria Simone (Universita e INFN, Bari (IT))

Description

The analysis of data collected by the CMS and ATLAS experiments at CERN, ahead of the next phase of high-luminosity at the LHC, requires a 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 back-end, has been developed in the framework of the "HPC, Big Data e Quantum Computing Research Centre" Italian National Center (ICSC), providing experiment-agnostic resources. Starting from container technology and orchestrated via Kubernetes, the platform provides analysis tools via 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 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 (JHEP 04 (2021) 165), and searches for rare flavor decays at the CMS experiment in Run-3 using large datasets collected by high-rate dimuon triggers.

Primary authors

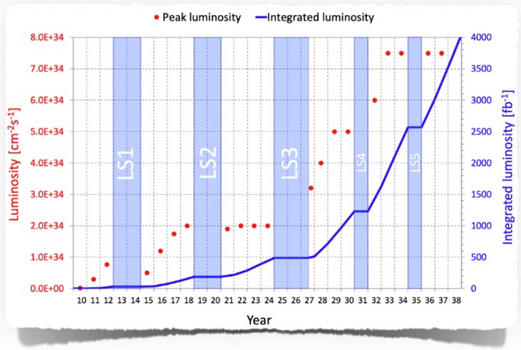
ATLAS Collaboration
Adelina D'Onofrio (INFN Napoli (IT))
Bernardino Spisso (Universita Federico II e INFN Sezione di Napoli (IT))
CMS Collaboration
Elvira Rossi (Universita Federico II e INFN Sezione di Napoli (IT))
Federica Maria Simone (Universita e INFN, Bari (IT))
Francesco Giuseppe Gravili (INFN Lecce e Universita del Salento (IT))
Salvatore Loffredo (INFN - National Institute for Nuclear Physics)
Tommaso Diotalevi (Universita e INFN, Bologna (IT))

Outline

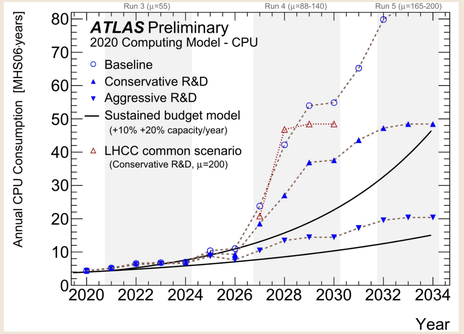
- Motivations
- Test infrastructure
  - ATLAS Experiment use case
  - CMS Experiment use case
    - Preliminary scalability results
- Conclusions

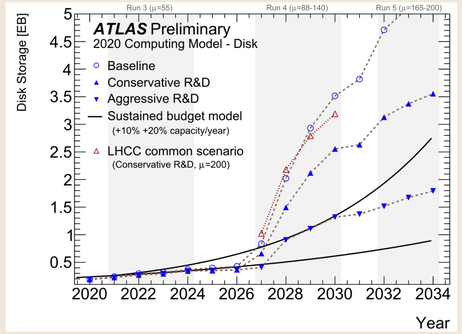
Motivations

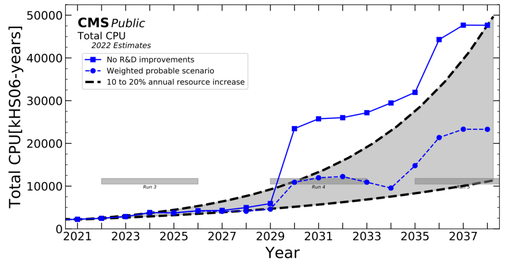
- **Challenges of LHC, and HL-LHC** are pushing to **re-think the HEP computing models**
  - Impact on several aspects, from software to the computing infrastructure

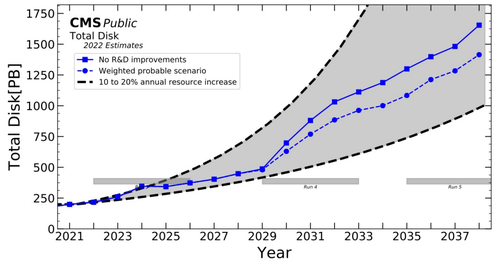


Similar trends for ATLAS and CMS HL-LHC projections







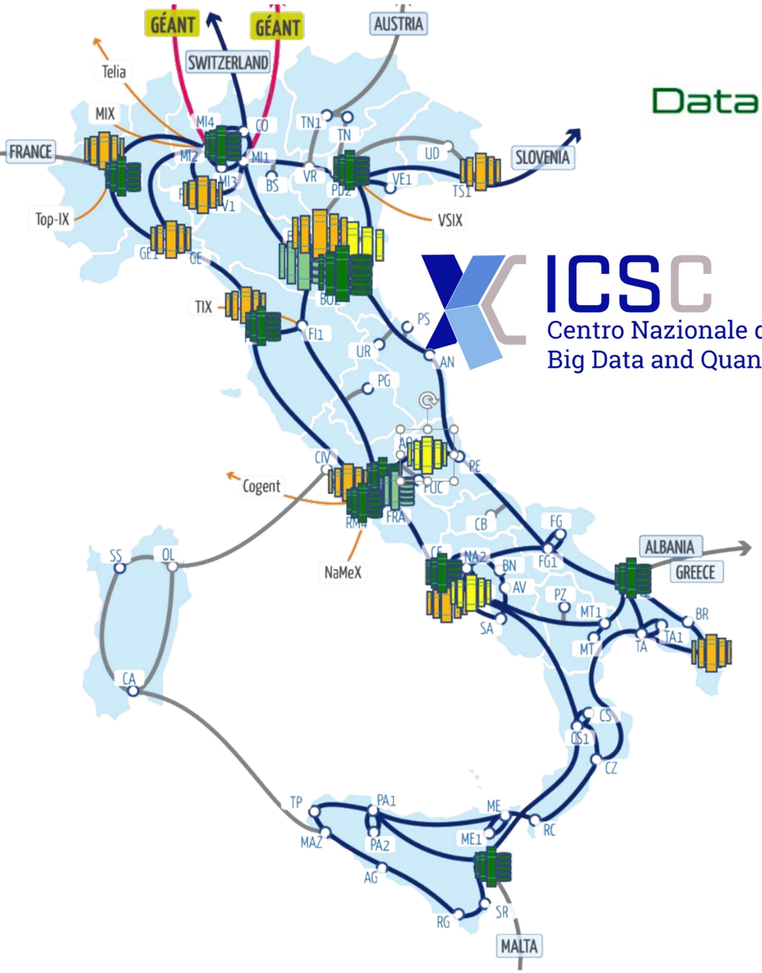


Higher rates of collision events → **Higher demand for computing and storage resources**

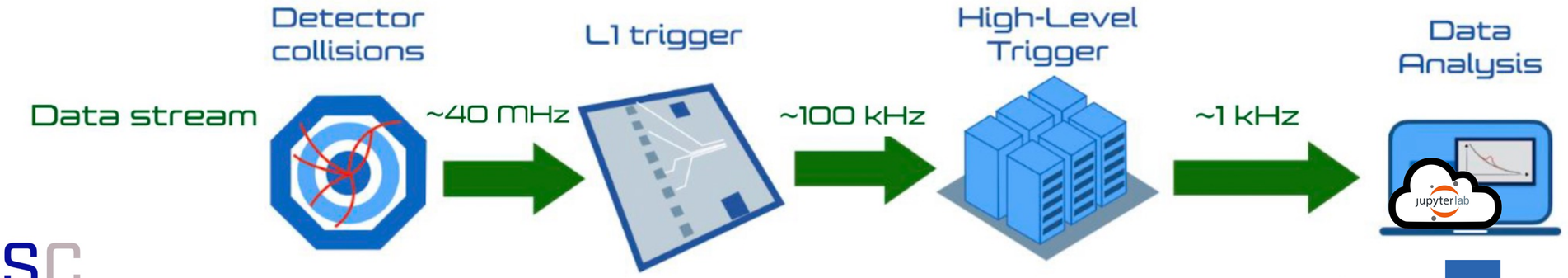
To better analyse this increasing amount of Big Data:

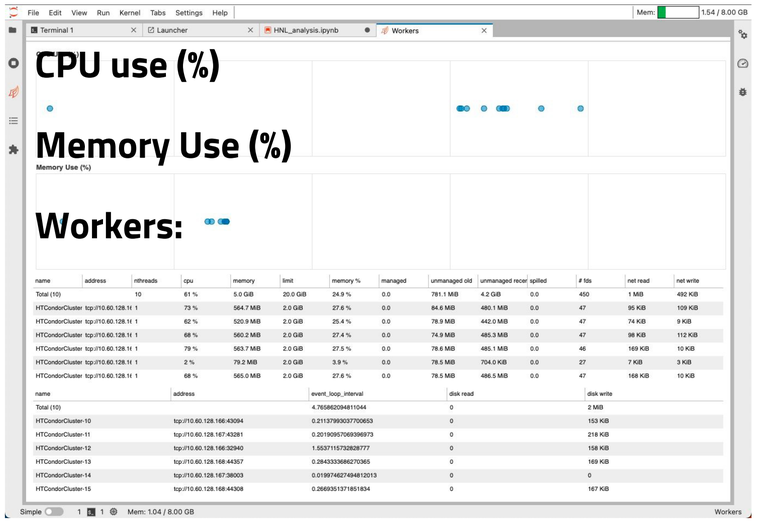
- Optimize the usage of CPU and storage
- Promote the usage of better data formats
- **Develop new analysis paradigms**!
- New software based on declarative programming and interactive workflows
- Distribute on geographically separated resources

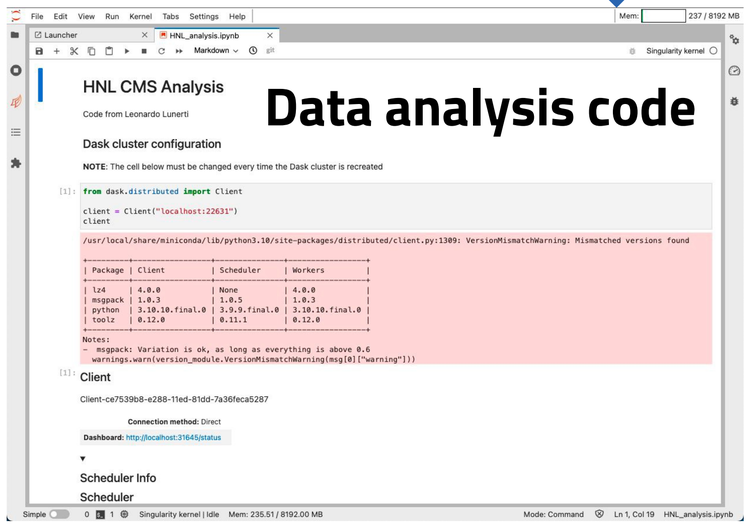
HEP data analysis with ICSC



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







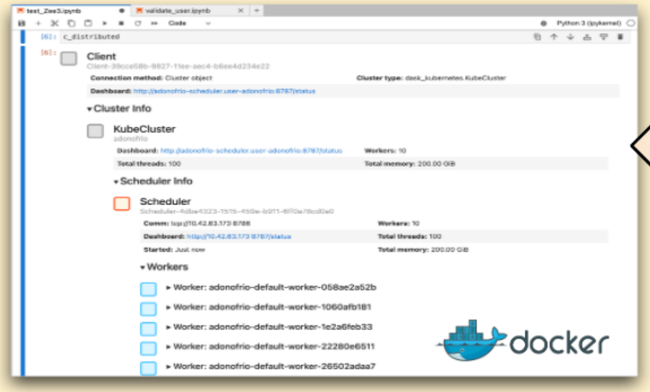
High throughput data analysis platform

Access and security

After connecting to an entrypoint URL, the user reaches a **Jupyterhub** instance that, after authentication and authorization via INDIGO-IAM, allocates the required resources for the user's working area

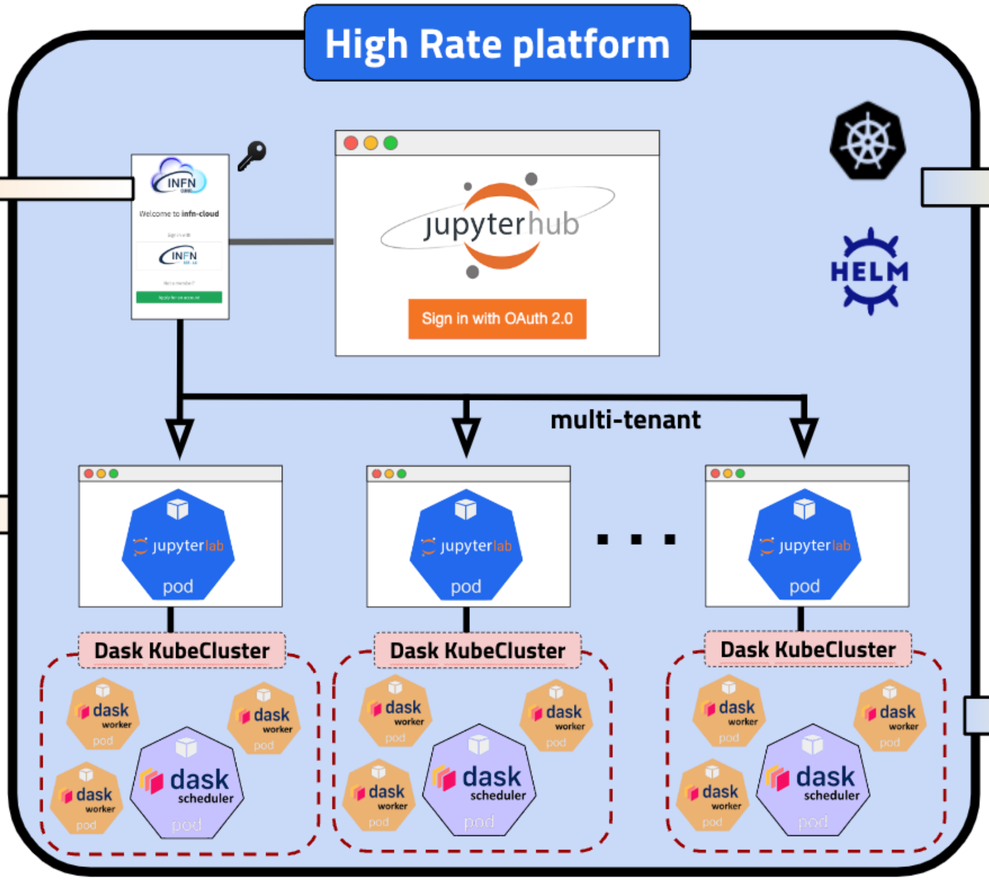
User Interface

The user interface is based on **Jupyterlab**, customised with specific plugins for specific purposes (e.g. Dask).



The working environment is highly customizable, using tailored Docker containers. This is important when analyses require specific software (collaboration-wise)

High Rate platform



Deployment

The deployment of the **Kubernetes** resources needed for the spawning of this platform, is handled via **HELM charts** available in the GitHub organization.

Check the docs!

This allows a seamless, flexible, scalable and fault-tolerant deployment on the available resources, with a limited impact on the admin's work time

Software

From the software perspective, interactive/quasi interactive analysis is a promising paradigm

- User-friendly environment
- Adopting open-source industry standards: *Dask, Jupyter Notebooks* and *HTCondor*
- Validating new frameworks (e.g. *ROOT RDataFrame* with multi-threading)

High throughput data analysis platform

Inputs needed from WP5 experts:

- I see some novelties in Tommaso's backup slides https://agenda.infn.it/event/43423/#8-high-throughput-analyses-fla
  - Using InterLink to dynamically deploy heterogeneous resources -> should I mention this in the plans? Is it under development?

- Should I add more technical details or should I assume people at CHEP already now about Kubernetes and HELM?
  Feedback by Francesco Gravili: add Kubernetes, while HELM can go to backup

- Extra slide on the dask daskboard?

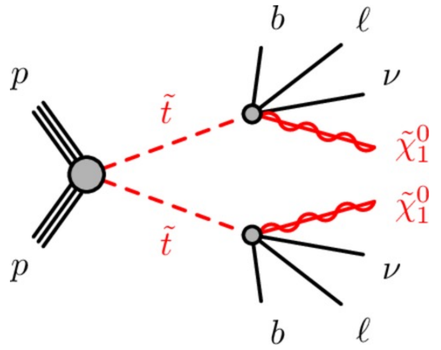
Benchmark interactive analyses

ATLAS use-case

SUperSYmmetry: Beyond Standard Model (BSM) theory

JHEP
Published for SISSA by Springer
Received: February 3, 2021
Accepted: March 3, 2021
Published: April 16, 2021

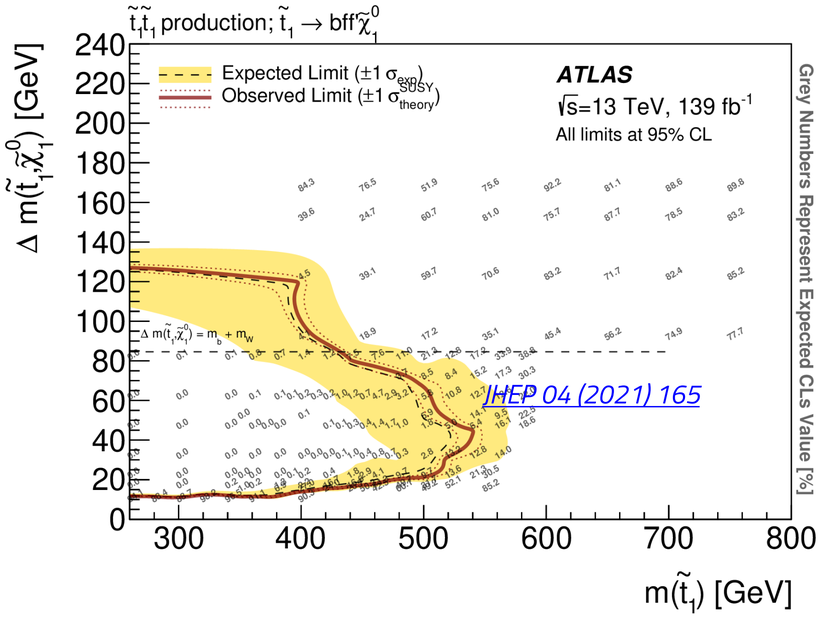
Search for new phenomena in events with two opposite-charge leptons, jets and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector



Soft leptons coming from a virtual W* boson decay

Compressed mass spectra:
$\Delta m < m_W + m_b$

- Three different analysis in the ***Run 2 paper***, already published, according to mass splitting between *stop* ($\tilde{t}_1$) and *neutralino* ($\tilde{\chi}^0_1$), allowing different decay modes:
  - 2 body ⟶ $\Delta m > m_t$
  - 3 body ⟶ $m_W + m_b < \Delta m < m_t$
  - 4 body, the one picked up ⟶ $\Delta m < m_W + m_b$
- Common final state signature: 2 OS leptons (electrons/muons), jets and missing transverse energy
- Cut & Count based approach



JHEP 04 (2021) 165

4-body search workflow

Analysis Facility

Skimming
- Provided by the Collaboration
- Offline reconstruction
- $\mathcal{O}$(PB) for data and MC

DAOD reduction

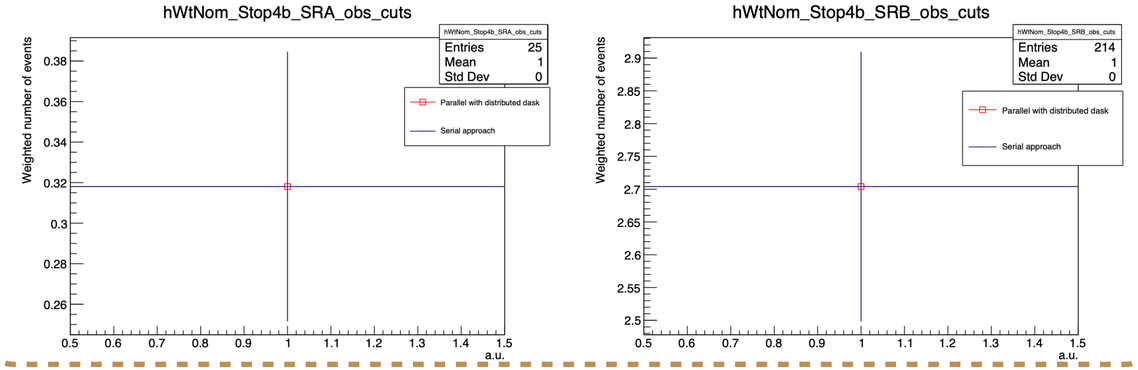
Thinning
- Removal of collections
- Baseline objects and trigger
- Scale Factors retrieval
- $\mathcal{O}$(TB) for data and MC

Slimming
- Removal of object quantities
- New variable definitions
- Weights application
- $\mathcal{O}(10^2$ GB) for data and MC

Event Selection
- Region definitions
- Nominal yields
- Systematic variations
- $\mathcal{O}$(MB), inputs to fitting tool(s)

Sanity check

Weighted number of events in the Wt background sample, after the event selection cuts in signal regions A and B, nominal case



Slimming

ATLAS slimming code already in RDataFrame, but entirely written and compiled in C++ —> NO dask distributed approach

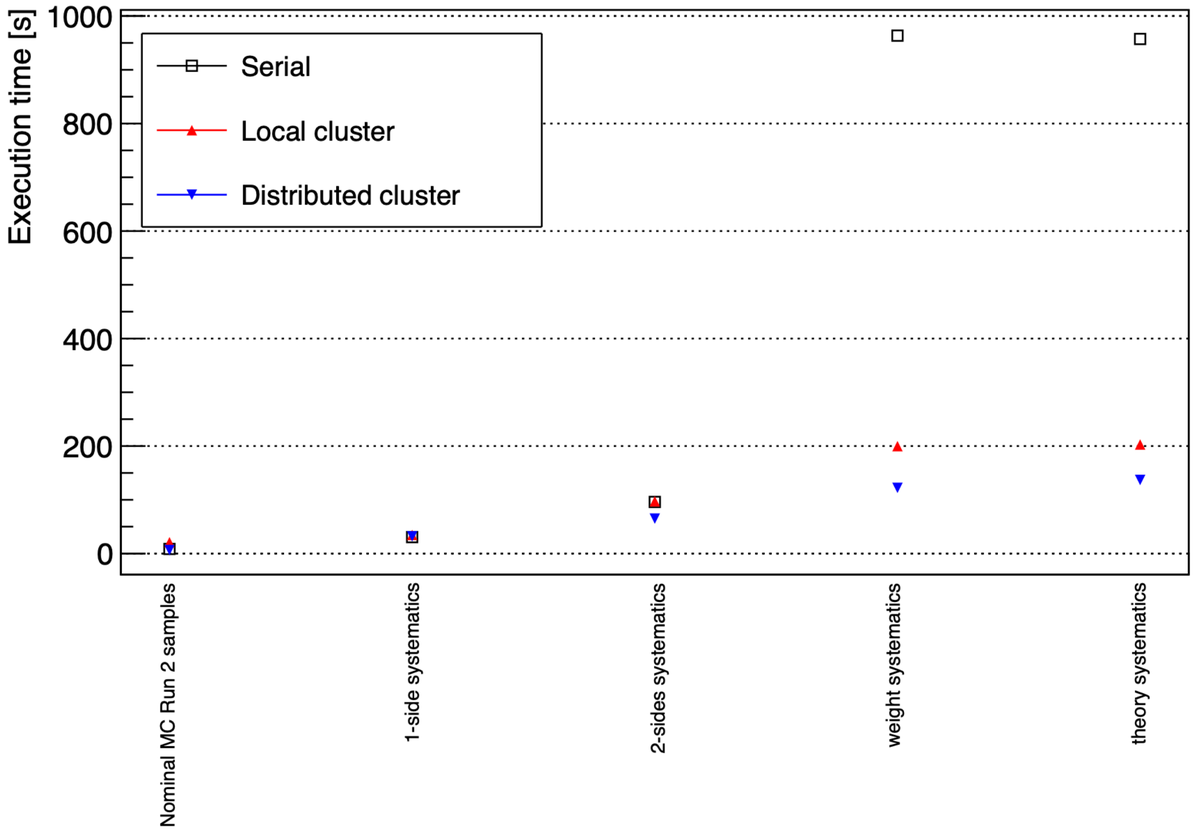
Event Selection

- Event selection for fitting tools
- RDataFrame + Dask applied to Wt bkg sample ~ 1.8 GB copied to the INFN workspace
- Playing with syst. variations
- Code ready to play with other backgrounds

Preliminary results

Defined Metric	
Overall execution time	Time elapsed from the start of the execution (execution triggered) to the end of execution

- Exploiting the distributed approach, the execution time improves *wrt* the standard/serial approach if we iterate over a significative number of systematic variations (each step in the x-axis includes previous contributions)



Scheduler and Working Nodes Reports

Distributed approach

Dask Performance Report

Select different tabs on the top for additional information

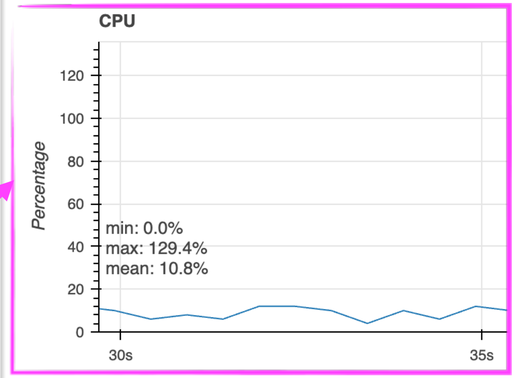
Duration: 252.87 s

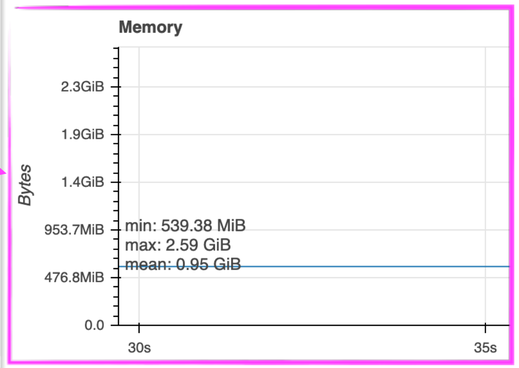
Tasks Information

- number of tasks: 621
- compute time: 118.06 s
- deserialize time: 2.39 s

Scheduler Information

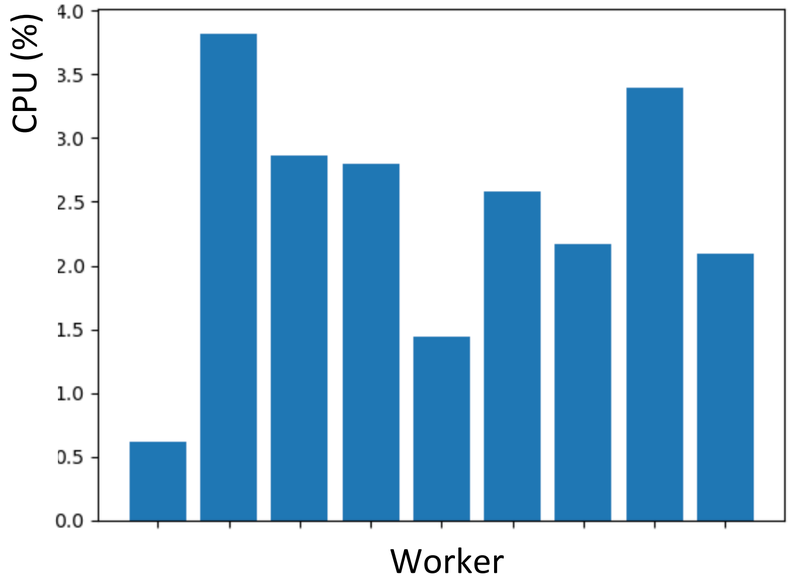
- Address: tcp://127.0.0.1:43821
- Workers: 2
- Threads: 2
- Memory: 4.39 GiB
- Dask Version: 2022.11.0
- Dask.Distributed Version: 2022.11.0



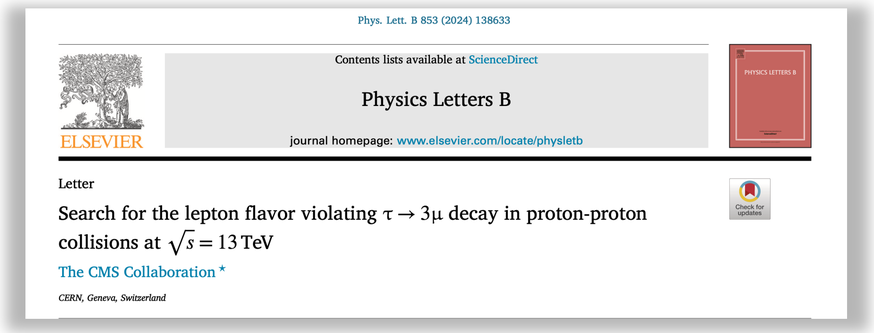


Connecting to working nodes

- Out of 9 worker nodes, we get up to 4% CPU occupancy on each worker node
- Limited CPU consumption due to the easy cut&count operations



CMS use-case


Phys. Lett. B 853 (2024) 138633

Contents lists available at ScienceDirect

Physics Letters B

journal homepage: www.elsevier.com/locate/physletb

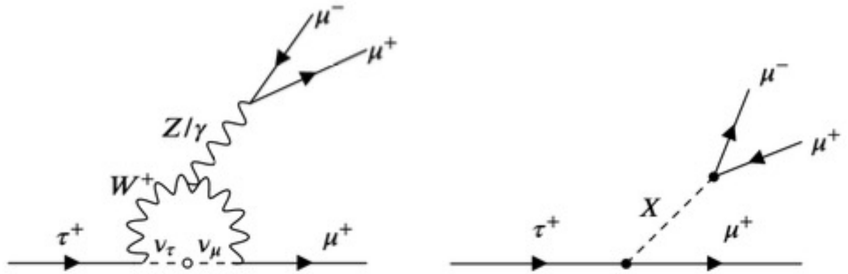
Letter

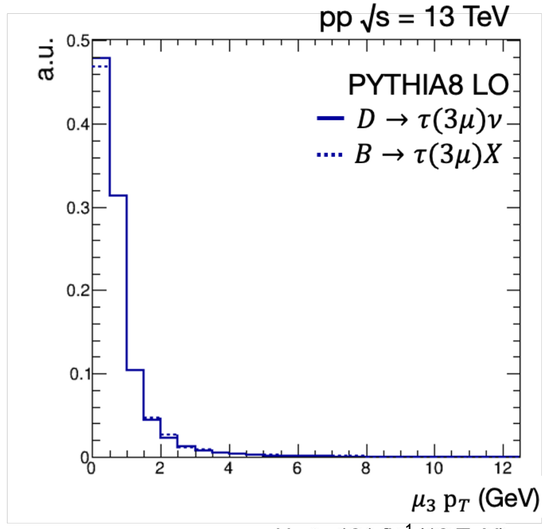
Search for the lepton flavor violating $\tau \to 3\mu$ decay in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration *

CERN, Geneva, Switzerland

Lepton Flavor Violation in the charged sector: $\tau \to 3\mu$

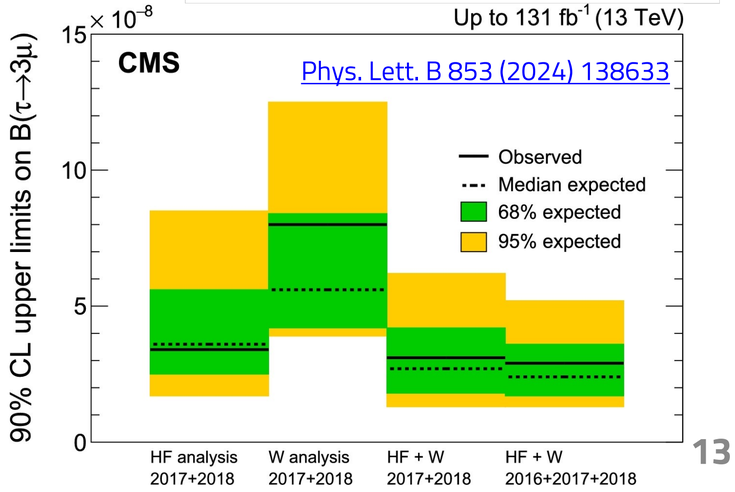




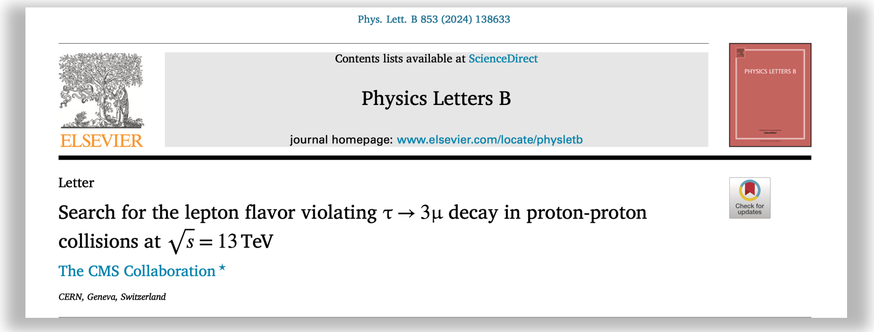
Search for $\boldsymbol{\tau \to 3\mu}$ decays, which have very small SM branching fractions $\mathrm{BR_{SM}} \sim \mathcal{O}(10^{-55})$, while being predicted with sizable BR in several BSM scenarios $\mathrm{BR_{BSM}} \sim \mathcal{O}(10^{-10} \div 10^{-8})$

- $\boldsymbol{\tau}$ leptons produced in D and B meson decays provide large statistics at LHC experiments, but are only accessible with **low-p_T muon triggers**

- Analysis of Run 2 data recently published, **stat. limited**
  - → benefitting from inclusive low-p_T muon L1 trigger in **Run 3**
  - → technical challenge: **new datasets are x3 times heavier**



CMS use-case


Phys. Lett. B 853 (2024) 138633

Contents lists available at ScienceDirect

Physics Letters B

journal homepage: www.elsevier.com/locate/physletb

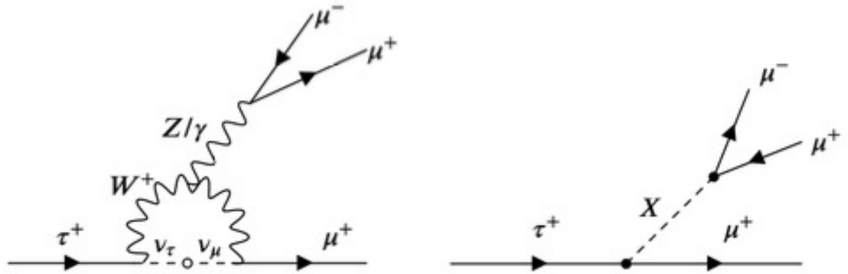
Letter

Search for the lepton flavor violating $\tau \to 3\mu$ decay in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration *

CERN, Geneva, Switzerland

Lepton Flavor Violation in the charged sector: $\tau \to 3\mu$

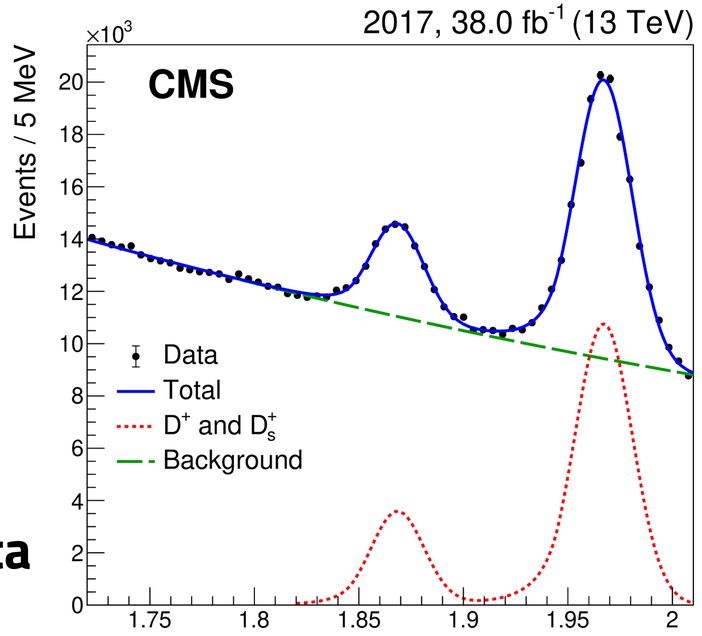


Search for $\boldsymbol{\tau \to 3\mu}$ decays, which have very small SM branching fractions $\mathrm{BR_{SM}} \sim \mathcal{O}(10^{-55})$, while being predicted with sizable BR in several BSM scenarios $\mathrm{BR_{BSM}} \sim \mathcal{O}(10^{-10} \div 10^{-8})$

- $\boldsymbol{\tau}$ leptons produced in D and B meson decays provide large statistics at LHC experiments, but are only accessible with **low-p_T muon triggers**

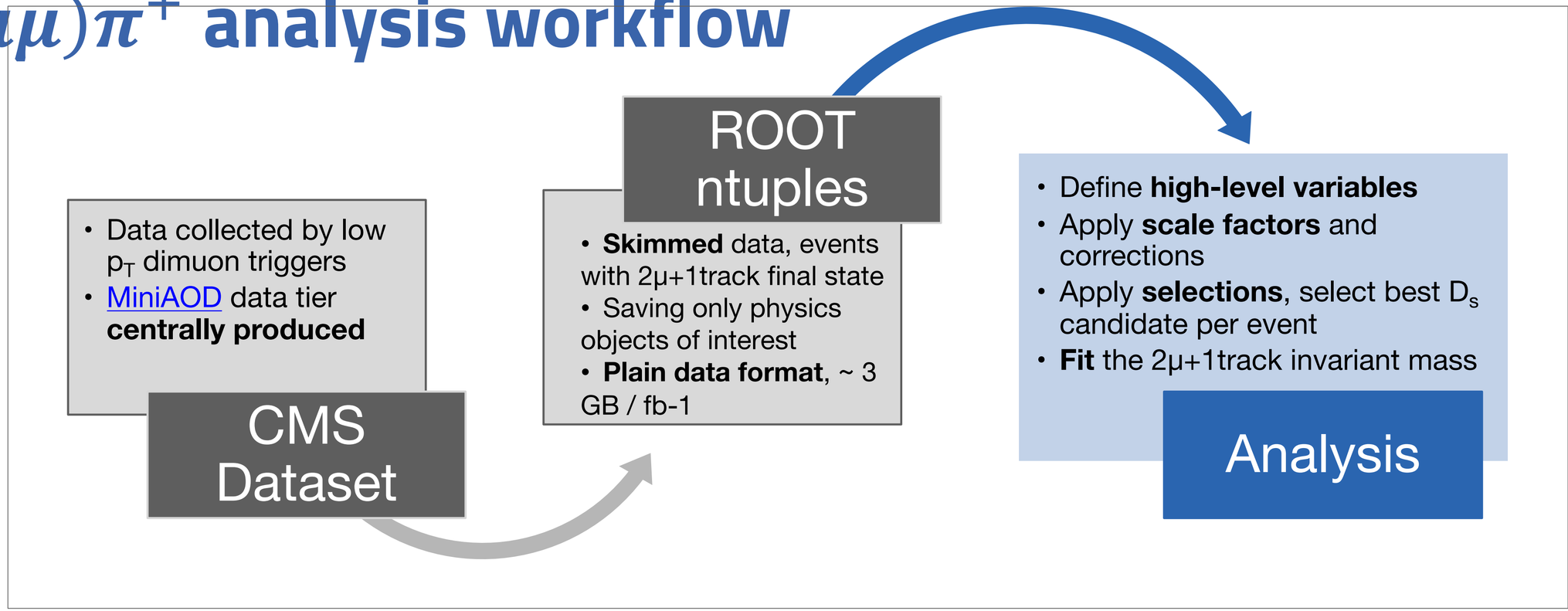
- The normalisation channel used as a benchmark: $D_s^+ \to \phi(\mu\mu)\pi^+$

  → **cut-based analysis + mass fit for measuring the D_s^+ yield in data**



Phys. Lett. B 853 (2024) 138633

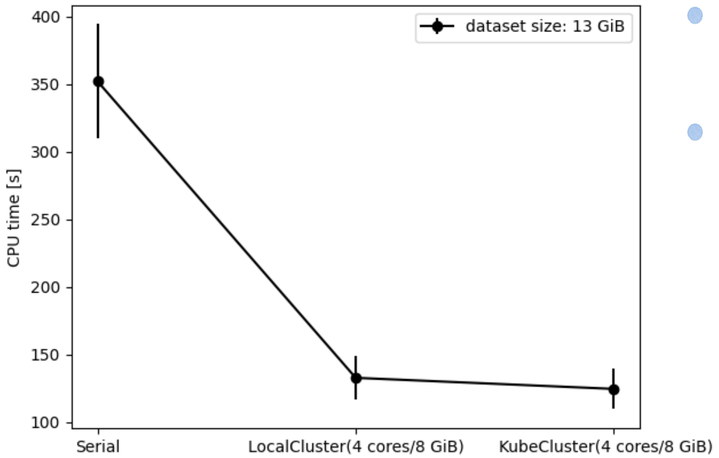
$D_s^+ \to \phi(\mu\mu)\pi^+$ analysis workflow



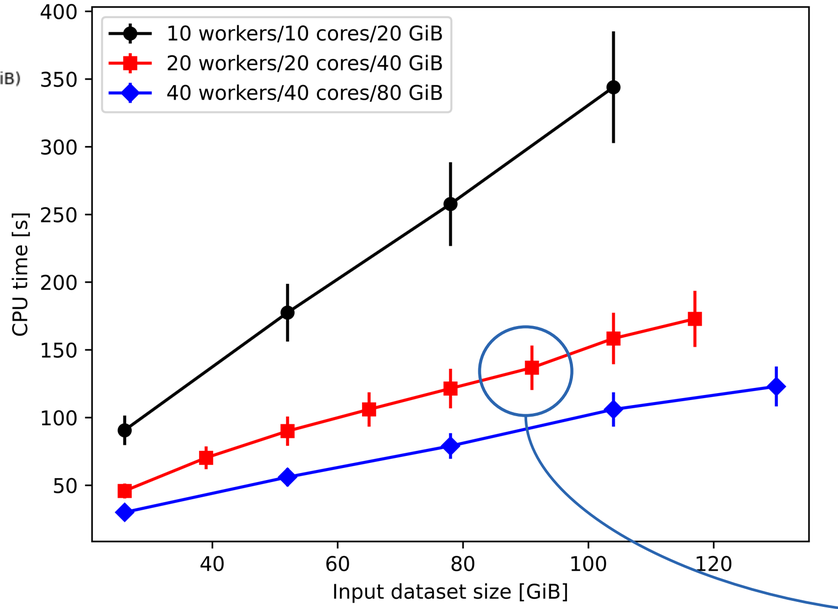
- **Legacy: approach** Loop-based analysis implemented using ROOT TTree:MakeClass
  - split computation in batches of input files, run separately as HTCondor jobs, gather the output rootfiles

- **New**: Ntuples read as RDataFrame, almost all operations "lazy" → no loop triggered till the end
  - go distributed using Dask

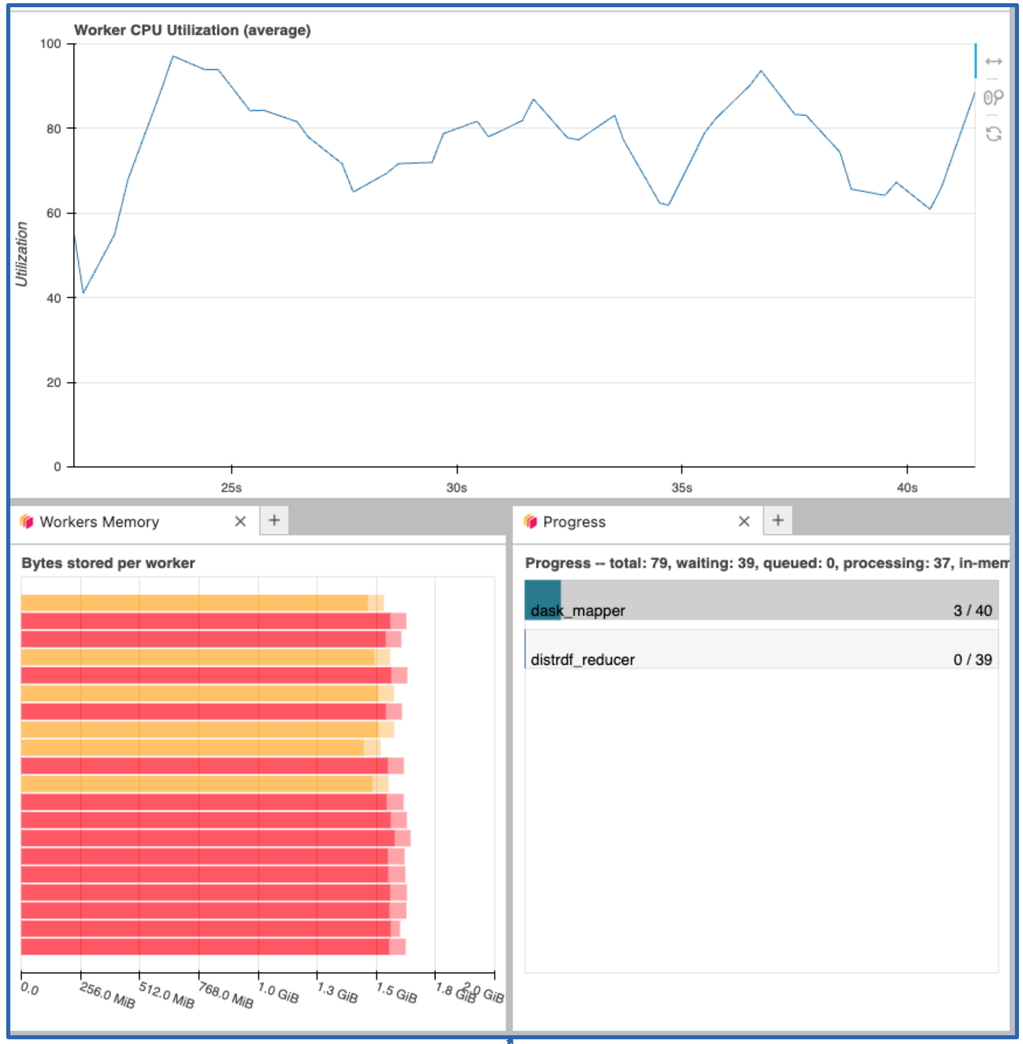
Preliminary results



- Significant improvement in execution time *wrt* the standard/serial approach
- The facility allows for dynamically scaling the resources, here testing the performance at fixed #cores and memory, varying the dataset size



- Stress test at high CPU and memory occupancy
- Stable performance, linearly scaling with the input dataset size
- Dataset size ~ 100 GiB is representative of ~15 /fb of Run3 data for this specific analysis



Conclusions & Next Steps

- HL-LHC poses significant challenges to HEP experiments in terms of storage and computing resources
- An interactive high throughput platform has been developed in the framework of the "HPC, Big Data e Quantum Computing Research Centre" Italian National Center (ICSC)
  - offers users a modern interactive web interface based on JupyterLab
  - experiment-agnostic resources
  - based on a parallel and geographically distributed back-end

- Interactive analyses feasibility studies on INFN cloud succeeded
  - Performance evaluated using the high-rate platform
  - HEP analysis use-case explored from the CMS and ATLAS Collaborations

Medium-long term goals: Towards a data-lake model (FIXME WP5 input here)

*This work is **(partially)** supported by ICSC – Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing, funded by European Union – NextGenerationEU*

Thank you!

Back-up

Run3 CMS Luminosity