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

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

CMS

CMS Detector Performance Analysis using the Analysis Facility

Tommaso Diotalevi (UniBO)

AF Workshop 2025 - 9 January 2025

Introduction

Typically, **Detector Performance Group** (DPG) analyses are run on a reduced amount of data (e.g. one run or fill), but processing of large dataset, at once, might be needed:

- To assess/improve systematics of high precision analyses, when they are dominated by the response of a specific detector;
- To reprocess multiple year data, e.g. for detector stability studies (ageing).

Use Case:

Porting of a well established Drift Tubes (DT) Tag-and-Probe analysis [CMS-DP-2023-049]

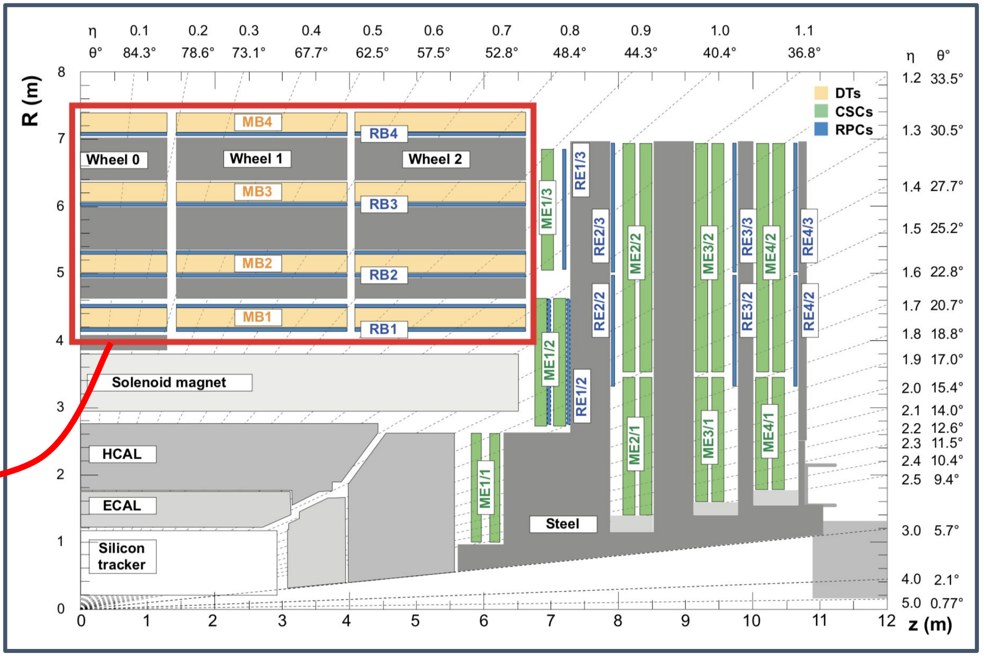
A **data sample** consisting in a skim of Z→μμ decay candidates collected by CMS over 2023, corresponding to ~27fb-1 was explored for the study. Size: 224GB. The format is a NanoAOD-like dataset, tailored for muon DPG studies.

- The original code running mainly on **C++**, for the base histograms and computing the segment efficiencies.
  - The ported code is running on **Jupyter notebook** (in Python), using **ROOT RDataFrame**. The Tag-and-Probe libraries are stored in a dedicated header file.

DT Tag and Probe method - in a nutshell

- Tag-and-Probe method [CMS-DP-2023-049]:
  - Two oppositely-charged well-reconstructed tracker muons;
  - **Tag muon**: p_T > 27GeV passing HLT for isolated muons. TightID criteria in the muon detector reconstruction;
  - **Probe muon**: track with segment matching in at least a chamber other than the one under study, and p_T > 20 GeV.

- A DT chamber **is efficient** if the reconstructed segment is near the extrapolated probe muon track.



Muon system in the barrel region, where the DTs are located.

Let’s take a look at the code!

The simplified version of the analysis is available here!

Technical performance on the full analysis (so far…)

To evaluate the technical performance, **the available statistics has been processed 3 times**, mimicking an integrated luminosity of **~82fb^{-1}** , consisting of **~77M events** in total. Size: 224*3 = **672GB**

- Serial processing (as a single job on HTCondor)

  Wall time: ~120 minutes

  1 CPU on a AMD EPYC 7302 16-Core Processor, with 2GB memory

- Distributed processing on the platform:

  Wall time: ~6 minutes

  Up to 92 CPUs (46 physical), on two AMD EPYC 7413 24-Core Processor, with 2GB memory per CPU. Resources hosted at T2_IT_LNL.

Quasi-interactivity is now reached:

- Every time a re-execution of the analysis is needed (e.g. tweaking some thresholds or using different selection criteria), running a few Jupyter Notebook cells will do the trick (transparently accessing more resources)!!
- This can result in a **great improvement** for any detector performance analysis application.

Thank you!

tommaso.diotalevi@unibo.it