



Contribution ID: 70

Type: **Presentation**

CMS experience at INFN with distRDF over HTCondor

Monday, 9 May 2022 11:15 (20 minutes)

In the context of R&D activities for the evolution of the analysis computing model for the CMS experiment, one of the focus is the capability to leverage an (quasi-)interactive and declarative approach to enhance both the user experience and the analysis throughput (meant as the end to end result delivery time). Another key point is how to make use of both grid resources and opportunistic ones in a coherent and efficient way.

In this talk we will show how distributed RDataFrame has been tested importing a CMS analysis to RDF framework and executing it on a prototype analysis-facility infrastructure at INFN. The presented scenario allows the user to login in a central JupyterHUB instance (or directly via ssh) and to schedule the RDataFrame payload on a remote Dask cluster instantiated via HTCondor. It will also be shown how RDataFrame allowed the transition from legacy code with a minimal effort.

Summary

Primary authors: SPIGA, Daniele (INFN Perugia); CIANGOTTINI, Diego (INFN, Perugia (IT)); TEDESCHI, Tommaso

Presenter: CIANGOTTINI, Diego (INFN, Perugia (IT))

Session Classification: First Session

Track Classification: The View of Physicists