

Python and ROOT: Effective and Interactive Analysis of Big Data

Monday, May 30, 2016 4:45 PM (30 minutes)

For a long time C++ was virtually the only language of HEP data analysis. This has certainly changed in the past few years: Python became a cornerstone of everyday work of Physicists and this is due to a large extent thanks to ROOT.

In this contribution we discuss the technical aspects enabling this innovation.

ROOT is a modular scientific software framework which provides all the functionalities needed to deal with big data processing, statistical analysis, visualisation and storage. One of the reasons for its success is its Python interface, PyROOT.

The programming model offered by PyROOT and the way it complements the ROOT C++ one in everyday analysis is characterised. Concrete examples are given of its usage in the software stacks of LHC experiments. The integration of PyROOT with Jupyter notebooks is described as well as glimpses of its potential for interactive data mining of experiments' data and non scientific data such as logs or machine instrumentation output. Elements of R&D activities are also outlined such as the integration of ROOT with Apache Spark with PyROOT and PySpark.

The CERN Service for Web based ANalysis, SWAN, is introduced describing its potential for delivering ROOT, Python and other analysis ecosystems.

Live demos are provided whenever possible to grant the audience an enhanced experience.

Talk Length

30 minutes

Primary author: PIPARO, Danilo (CERN)

Presenter: PIPARO, Danilo (CERN)