



Contribution ID: 203

Type: Oral

First performance measurements with the Analysis Grand Challenge

Tuesday 25 October 2022 17:20 (20 minutes)

The IRIS-HEP Analysis Grand Challenge (AGC) is designed to be a realistic environment for investigating how analysis methods scale to the demands of the HL-LHC. The analysis task is based on publicly available Open Data and allows for comparing usability and performance of different approaches and implementations. It includes all relevant workflow aspects from data delivery to statistical inference.

The reference implementation for the AGC analysis task is heavily based on tools from the HEP Python ecosystem. It makes use of novel pieces of cyberinfrastructure and modern analysis facilities in order to address the data processing challenges of the HL-LHC.

This contribution compares multiple different analysis implementations and studies their performance. Differences between the implementations include the use of multiple data delivery mechanisms and caching setups for the analysis facilities under investigation.

Significance

This presentation shows for the first time quantitative performance results for various workflows envisioned at the HL-LHC in a realistic analysis environment. Novel aspects are the evaluation of a full analysis pipeline and the use of an analysis task that captures the relevant workflow aspects and scale physicists require in practice.

References

ICHEP presentation, targeted at introducing the overall project at a general level: <https://agenda.infn.it/event/28874/contributions/169204/>
demonstration of analysis task and several implementations at a dedicated workshop: <https://indico.cern.ch/event/1126109/contributions/>
Both of these presentations did not target quantitative performance evaluation.

Experiment context, if any

ATLAS / CMS in particular for the analysis task, to a lesser extent LHCb and other experiments with similar data processing pipelines

Authors: HELD, Alexander (University of Wisconsin Madison (US)); SHADURA, Oksana (University of Nebraska Lincoln (US))

Presenter: SHADURA, Oksana (University of Nebraska Lincoln (US))

Session Classification: Track 2: Data Analysis - Algorithms and Tools

Track Classification: Track 2: Data Analysis - Algorithms and Tools