



Contribution ID: 94

Type: Talk

Building a Columnar Analysis Demonstrator for ATLAS PHYSLITE Open Data using the Python Ecosystem

Monday 21 October 2024 14:42 (18 minutes)

The ATLAS experiment is in the process of developing a columnar analysis demonstrator, which takes advantage of the Python ecosystem of data science tools. This project is inspired by the analysis demonstrator from IRIS-HEP.

The demonstrator employs PHYSLITE OpenData from the ATLAS collaboration, the new Run 3 compact ATLAS analysis data format. The tight integration of ROOT features within PHYSLITE presents unique challenges when integrating with the Python analysis ecosystem. The demonstrator is constructed from ATLAS PHYSLITE OpenData, ensuring the accessibility and reproducibility of the analysis.

The analysis pipeline of the demonstrator incorporates a comprehensive suite of tools and libraries. These include uproot for data reading, awkward-array for data manipulation, Dask for parallel computing, and hist for histogram processing. For the purpose of statistical analysis, the pipeline integrates cabinetry and pyhf, providing a robust toolkit for analysis. A significant component of this project is the custom application of corrections, scale factors, and systematic errors using ATLAS software. Therefore for this component we conduct a comparative analysis of event processing throughput across both the event-loop and columnar analysis environments. The infrastructure and methodology for these applications will be discussed in detail during the presentation, underscoring the adaptability of the Python ecosystem for high-energy physics analysis.

Authors: HELD, Alexander (University of Wisconsin Madison (US)); Dr STARK, Gordon Holtsberg (University of California, Santa Cruz (US)); WATTS, Gordon (University of Washington (US)); CHOI, Kyungeon (University of Texas at Austin (US)); HEINRICH, Lukas Alexander (Technische Universitat Munchen (DE)); FEICKERT, Matthew (University of Wisconsin Madison (US)); VIGL, Matthias (Technische Universitat Munchen (DE)); HARTMANN, Nikolai (Ludwig Maximilians Universitat (DE)); KRUMNACK, Nils Erik (Iowa State University (US)); KOURLITIS, Vangelis (Technische Universitat Munchen (DE))

Presenter: FEICKERT, Matthew (University of Wisconsin Madison (US))

Session Classification: Parallel (Track 5)

Track Classification: Track 5 - Simulation and analysis tools