Contribution ID: **100** Type: **Poster**

Using Legacy ATLAS C++ Calibration Tools in Modern Columnar Analysis Environments

Thursday, 14 March 2024 16:10 (30 minutes)

The ATLAS experiment at the LHC relies on crucial tools written in C++ to calibrate physics objects and estimate systematic uncertainties in the event-loop analysis environment. However, these tools face compatibility challenges with the columnar analysis paradigm that operates on many events at once in Python/Awkward or RDataFrame environments. Those challenges arise due to the intricate nature of certain tools, as a result of years of continuous development, and the necessity to support a diverse range of compute environments. In this contribution, we present the ATLAS R&D efforts to adapt these legacy tools to be used in both event-loop and columnar environments with minimal code modifications. This approach enables on-the-fly calibration and uncertainties calculations, minimizing the reliance on intermediate data storage. We demonstrate the functionality and performance of this approach in a Python Jupyter notebook that reproduces a toy Z-boson-peak analysis.

Significance

This presentation introduces an innovative strategy for incorporating legacy C++ code into the modern data science ecosystem. This approach enables the on-the-fly computation of corrections and uncertainties, consequently diminishing the requirement for intermediary data files. This initiative aligns with the broader goals of the HEP community to curtail reliance on disk storage, especially in preparation for the HL-LHC era and beyond. While columnar analysis using ATLAS lightweight data formats (PHYSLITE) has been demonstrated previously, this marks the first instance where corrections and uncertainties can be computed during data reading. Traditionally, these values were pre-calculated and stored in intermediary data files. ATLAS aims to extend such methodologies to a majority of tools essential for physics analysis, indicating a transformative shift in the HEP data analysis workflow.

References

1. Columnar analysis and on-the-fly analysis corrections at ATLAS https://indico.jlab.org/event/459/contributions/11583/
2. Columnar data analysis with ATLAS analysis formats http://dx.doi.org/10.1051/epjconf/202125103001
3. PHYSLITE - A new reduced common data format for ATLAS https://cds.cern.ch/record/2870350/

Experiment context, if any

ATLAS

Primary authors: Dr STARK, Giordon Holtsberg (University of California,Santa Cruz (US)); HEINRICH, Lukas Alexander (Technische Universitat Munchen (DE)); FEICKERT, Matthew (University of Wisconsin Madison (US)); VIGL, Matthias (Technische Universitat Munchen (DE)); KRUMNACK, Nils Erik (Iowa State University (US)); KOURLITIS, Vangelis (Technische Universitat Munchen (DE))

Co-authors: HELD, Alexander (University of Wisconsin Madison (US)); WATTS, Gordon (University of Washington (US)); HARTMANN, Nikolai (Ludwig Maximilians Universitat (DE))

Presenter: VIGL, Matthias (Technische Universitat Munchen (DE))

Session Classification: Poster session with coffee break

Track Classification: Track 1: Computing Technology for Physics Research