Contribution ID: **658** Contribution code: **contribution ID 658**   Type: **Poster**

RDataFrame enhancements for HEP analyses: systematics, metadata and collection operations

In recent years, RDataFrame, ROOT's high-level interface for data analysis and processing, has seen widespread adoption on the part of HEP physicists. Much of this success is due to RDataFrame's ergonomic programming model that enables the implementation of common analysis tasks more easily than previous APIs, without compromising on application performance. Nonetheless, RDataFrame's interfaces have been further improved by the recent addition of several major HEP-oriented features: in this contribution we will introduce for instance a dedicated syntax to define systematic variations, per-data-sample call-backs useful to define quantities that vary on a per-sample basis, simplifications of collection operations and the injection of just-in-time-compiled Python functions in the optimized C++ event loop.

Significance

RDataFrame sees widespread usage in the HEP community, and this is only expected to increase in the future as more analyses move away from legacy interfaces [1] [2].

The novel interfaces that we will present, for the first time, in this contribution will significantly change the way users write analysis code with RDataFrame.

This will not only impact our many direct users, but also frameworks that build on top of RDataFrame such as bamboo [3], CROWN [4] and TIMBER [5], further increasing the relevance of this work for the HEP community.

[1] https://indico.cern.ch/event/941278/contributions/4084839/
[2] https://indico.cern.ch/event/1028381/#7-root
[3] https://gitlab.cern.ch/cp3-cms/bamboo
[4] https://github.com/KIT-CMS/CROWN
[5] https://github.com/lcorcodilos/TIMBER

References

Speaker time zone

No preference

Primary authors: NAUMANN, Axel (CERN); TEJEDOR SAAVEDRA, Enric (CERN); GUIRAUD, Enrico (EP-SFT, CERN); BLOMER, Jakob (CERN); WUNSCH, Stefan (KIT - Karlsruhe Institute of Technology (DE)); HAGE-BOECK, Stephan (CERN); PADULANO, Vincenzo Eduardo (Valencia Polytechnic University (ES))

Presenter: GUIRAUD, Enrico (EP-SFT, CERN)

Session Classification: Posters: Orange

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