



Contribution ID: 261

Type: **Plenary**

Simpler, faster and bigger: HEP analysis in the LHC Run 3 era

Thursday, 27 October 2022 10:30 (30 minutes)

The production, validation and revision of data analysis applications is an iterative process that occupies a large fraction of a researcher's time-to-publication.

Providing interfaces that are simpler to use correctly and more performant out-of-the-box not only reduces the community's average time-to-insight but it also unlocks completely novel approaches that were previously impractically slow or complex.

All of the above becomes especially true at the unprecedented integrated luminosity that will be achieved during LHC Run 3 and beyond, which further motivates the fast-paced evolution that has been taking place in the HEP analysis software ecosystem in recent years.

This talk analyzes the trends and challenges that characterize this evolution.

In particular we focus on the emerging pattern of strongly decoupling end-user analysis logic from low-level I/O and work scheduling by interposing high-level interfaces that gather semantic information on the particular analysis application.

We show how this pattern brings benefits to analysis ergonomics and reproducibility, as well as opportunities for performance optimizations.

We highlight potential issues in terms of extensibility and debugging experience, together with possible mitigations.

Finally, we explore the consequences of this convergent evolution towards smart, HEP-aware "middle-man analysis software" in the context of future analysis facilities and data formats:

both will have to support a bazaar of high-level solutions while optimizing for typical low-level data structures and access patterns.

Our goal is to provide novel insights useful to boost the ever-ongoing, stimulating conversation that, since always, characterizes the HEP software community.

Experiment context, if any

References

Significance

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

Session Classification: Plenary