



Contribution ID: 256

Type: **Plenary**

Adapting C++ for Data Science

Friday, 28 October 2022 10:30 (30 minutes)

Over the last decade the C++ programming language has evolved significantly into safer, easier to learn and better supported by tools general purpose programming language capable of extracting the last bit of performance from bare metal. The emergence of technologies such as LLVM and Clang have advanced tooling support for C++ and its ecosystem grew qualitatively. C++ has an important role in the field of scientific computing as the language design principles promote efficiency, reliability and backward compatibility - a vital tripod for any long-lived codebase. Other ecosystems such as Python have prioritized better usability and safety while making some tradeoffs on efficiency and backward compatibility. That has led developers to believe that there is a binary choice between performance and usability.

In this talk we would like to present the advancements in the C++ ecosystem; its relevance for scientific computing and beyond; and foreseen challenges. The talk introduces three major components for data science – interpreted C++; automatic language bindings; and differentiable programming. We outline how these components help Python and C++ ecosystems interoperate making a little compromise on either performance or usability. We elaborate on a future hybrid Python/C++ differentiable programming analysis framework which might accelerate science discovery in HEP by amplifying the power and physics sensitivity of data analyses into end-to-end differentiable pipelines.

Experiment context, if any

References

Significance

Presenter: VASILEV, Vassil (Princeton University (US))

Session Classification: Plenary