

## hep\_tables: Heterogeneous Array Programming for HEP

*Wednesday, 19 May 2021 18:45 (13 minutes)*

Array operations are one of the most concise ways of expressing common filtering and simple aggregation operations that is the hallmark of the first step of a particle physics analysis: selection, filtering, basic vector operations, and filling histograms. The High Luminosity run of the Large Hadron Collider (HL-LHC), scheduled to start in 2026, will require physicists to regularly skim datasets that are over a PB in size, and repeatedly run over datasets that are 100's of TB's –too big to fit in memory. Declarative programming techniques are a way of separating the intent of the physicist from the mechanics of finding the data, processing the data, and using distributed computing to process it efficiently that is required to extract the plot or data desired in a timely fashion. This paper describes a prototype library that provides a framework for different sub-systems to cooperate in producing this data, using an array-programming declarative interface. This prototype has a servicex data-delivery sub-system and an awkward array sub-system cooperating to generate requested data. The ServiceX system runs against ATLAS xAOD data.

**Primary author:** WATTS, Gordon (University of Washington (US))

**Presenter:** WATTS, Gordon (University of Washington (US))

**Session Classification:** Software

**Track Classification:** Offline Computing