

Efficient Workflow Management in High-Energy Physics

Wednesday 26 March 2025 11:30 (1 hour)

In high-energy physics (HEP), efficient workflow management is crucial for processing large datasets, running simulations, and managing computational jobs across distributed environments. This lecture introduces Luigi, a workflow management tool originally developed at Spotify that helps automate and scale complex task pipelines, ensuring dependency resolution and fault tolerance.

Building on Luigi, Law (Luigi analysis workflow) provides additional abstractions for HEP workflows by incorporating diverse batch job submission systems, like HTCondor, and different execution environments. Additionally, the automatic management of accessing distributed storage locations using the standard WLCG transfer protocols (e.g., WebDAV and XRootD) enables seamless execution of tasks across remote computing resources and ensures efficient resource utilization, simplifying the life of physicists a lot.

The lecture will provide both conceptual insights and practical approaches for managing large-scale workflows in HEP, leveraging modern tools and distributed computing infrastructure to enhance scientific computing. Specifically, it will cover how Luigi and law can be used to construct robust and scalable workflows for HEP applications, from running jobs on distributed batch systems to automatically managing remote data access to conduct complex analyses.

Number of lecture hours

1

Number of exercise hours

1

Attended school

tCSC 2024 (Belgrade)

Author: VERSTEGE, Cedric (KIT - Karlsruhe Institute of Technology (DE))

Presenter: VERSTEGE, Cedric (KIT - Karlsruhe Institute of Technology (DE))

Track Classification: Physics Computing and Data analysis