Contribution ID: **76** Type: **Oral**

Scaling Diffuse Scattering Workflows with Hybrid HPC Workflows

Monday 19 September 2022 17:40 (20 minutes)

Data analysis pipelines for diffuse scattering workflows consist of various steps with differing requirements for computation time and user interaction. The NXrefine workflow system is a semi-automated Python GUI toolkit based around NeXpy and the NeXus data format for diffuse x-ray scattering and other applications. The user is able to orchestrate many analysis pipelines on different datasets concurrently, distributing work to typical Linux clusters. Its most computationally intensive component is a coordinate transform implemented by the previously developed CCTW application. In practical, live data collection/analysis use cases, a user team may end up with a large backlog of CCTW work capable of exploiting 32K threads for tens of minutes. Executing this workload immediately is important to inspect data quality and other application goals during data collection, but it is difficult to gain access to adequate computational resources to perform this analysis. In this presentation, we describe a new hybrid HPC component which distributes this work to an MPI-enabled workflow system running on Theta or other HPC systems. We describe the portability and scalability of this component with respect to the diffuse scattering application as well as more general workloads.

This work was supported by the U.S. Department of Energy,
Office of Science, Advanced Scientific Computing Research,
under contract number DE-AC02-06CH11357 and the Office of Basic Energy Sciences, Division of Materials Sciences and Engineering.

Email address of presenting author

woz@anl.gov

Authors: Mr JENNINGS, Guy (ANL); WOZNIAK, Justin (Argonne National Laboratory); Dr KROGSTAD, Matthew (ANL); Dr OSBORN, Ray (ANL); Dr ROSENKRANZ, Stephan (ANL)

Presenter: WOZNIAK, Justin (Argonne National Laboratory)

Track Classification: NOBUGS 2022