Accelerated Data Analysis Workflows for Ptychography using Remote Leadership Computing Resources*

Tekin Bicer
Data Science and Learning Division
Argonne National Accelerator (ANL)

Abstract

High data rate experiments at synchrotron light sources can easily generate data in volumes that can take days (or even weeks) to process using locally available resources. To address this challenge, we present a system that unifies leadership computing and experimental facilities by enabling the automated establishment of data analysis pipelines that extend from edge data acquisition systems at beamlines to remote computing facilities; under the covers, our system uses Globus Auth authentication to minimize user interaction, funcX to run user-defined functions on supercomputers, Globus Flows to define and execute workflows, and Tike toolkit for accelerated analysis of data. We describe the application of this system to ptychography. When deployed on the DGX A100 ThetaGPU cluster at the Argonne Leadership Computing Facility and a microscopy beamline at the Advanced Photon Source, our system performs analysis as an experiment progresses to provide timely feedback.

* Title and Abstract of an Invited Presentation at the Workshop: AI/ML for Particle Accelerator, X-Ray Beamlines and Electron Microscopy at Argonne National Laboratory, November 1-3, 2021.