Real-Time Data Pipeline and analysis using spot and HIPGISAXS

Alexander Hexemer
Lawrence Berkeley Laboratory

Abstract

Real-time feedback to scientists during light source and neutron source beamtimes is a capability critically needed by many facility users, yet unobtainable for very large data sets and/or for datasets requiring HPC resources to analyse. Scattering methods like SAXS and GISAXS (Grazing Incidence Small Angel X-Ray Scattering) generates reciprocal space data that cannot be directly analysed for the underlying material structure. Rather, reverse Monte Carlo and other fitting methods are employed to reverse engineer the sample material. HipGISAXS (High Performance GISAXS) has been developed to run scattering simulations on massively parallel platforms such as Titan (OLCF), scalable to thousands of GPUs. Further, with inverse modelling algorithms available in HipGISAXS, such as particle swarm optimization, it can perform a large number of simulations simultaneously during the structure fitting process. In September of 2014 HipGISAXS was used in a real time demonstration that marries the 7.3.3 WAXS/SAXS beamline at the ALS using a high-speed Pilatus 2M camera with the data handling and processing capabilities at NERSC, and simulation capabilities of running at-scale simulations on Titan at OLCF. To accomplish the goal of real time data analysis, we coupled the data management and workflow SPOT Suite infrastructure running at NERSC, the data handling and processing capabilities in CADES at ORNL, and the high-performance data transfer capabilities of Globus Online. The demo involved a slot-die printer installed at beamline 7.3.3. Over the span of 3 days many different organic photovoltaics were printed at the beamline and the crystal structure evolution during drying was recorded using GIWAXS. Real-time GIWAXS fitting during the experiments was attempted on the second fastest computer in the world TITAN in Oak Ridge.