

Developing a shared computing and networking infrastructure for the ALS-ENABLE structural biology program at the Advanced Light Source.

Scott Classen^a, James Holton^a, Corie Ralston^a, Greg Hura^a, Jay Nix^a, Paul Adams^a

^aAdvanced Light Source, Lawrence Berkeley National Lab, One Cyclotron Road, Berkeley, CA 94720, USA, sclassen@lbl.gov

ALS-ENABLE is a new NIH-funded program to support the combined operation of MX and SAXS beamlines at the Advanced Light Source¹. Historically the MX and SAXS beamlines at the ALS have been run by several different organizational entities, including the Berkeley Center for Structural Biology, the Molecular Biology Consortium, SIBYLS, and the University of California. ALS-ENABLE is a first step to bring these resources together under a common banner to offer synchrotron structural biology resources to the broader research community. I will present a short history of structural biology beamlines at the ALS and the challenges that lay ahead for structural biology as we prepare for the impending ALS upgrade². I will also present details of our plans for a shared computing and networking infrastructure going forward.

As part of the ALS-ENABLE effort to consolidate MX and SAXS operations we have begun implementation of a shared computing and networking infrastructure. Initial efforts include shared data storage, data processing, and user access resources. To this end we have purchased an enterprise level NetApp storage system with several hundred TB of usable disk space. We are also upgrading the data network connecting the individual beamlines to the central server room from 1GbE to 10GbE. This will enable us to maximally service the high rates of data output from newer high speed Pixel Array and CMOS detectors.

To facilitate realtime data processing we have acquired several new CPU and GPU compute systems. These include an XDS “dream machine” with 8 Intel Xeon Platinum 8280L CPUs each having 28 hyper-threaded cores for a total of 448 cores. XDS has been shown to perform significantly faster with more cores (rather than higher Hz)³. We are in the process of commissioning this CPU compute system for realtime XDS analysis of MX data. We are also exploring the usefulness of GPU computing for various MX and SAXS data processing packages and are working with software developers to port their programs to CUDA. For these efforts we have acquired a GPU machine with 8 NVIDIA Titan V GP100 cards.

To lower the barrier for users to access their raw data, retrieve archived data, and reprocess data we are also exploring the implementation of the ISPyB⁴ backend database and the SynchWeb⁵ frontend user interface that have been under continual development since 2003 by a consortium of European synchrotrons. My presentation will highlight various technical details of our computing/networking, the challenges we are facing, and plans for the future.

1. <https://als-enable.lbl.gov>
2. <https://als.lbl.gov/als-u/als-u-approach/>
3. <https://bl831.als.lbl.gov/~jamesh/benchmarks/>
4. Delageniere S, Brenchereau P, Launer L, Ashton AW, Leal R, Veyrier S, et al. ISPyB: an information management system for synchrotron macromolecular crystallography. *Bioinformatics*. 2011 Nov 15;27(22):3186–92.
5. Fisher SJ, Levik KE, Williams MA, Ashton AW, McAuley KE. SynchWeb: a modern interface for ISPyB. *J Appl Crystallogr. International Union of Crystallography*; 2015 Jun 1;48(Pt 3):927–32.