

Solving the BES Facilities Data Challenge Using ASCR Facilities and SPOT

Craig E. Tull
Lawrence Berkeley National Laboratory

For decades, synchrotron light sources have been operating with a manual grab-and-go data management model: users run experiments, download data to an external hard drive, and process or analyze it later on their computers - provided they have the local resources, software, and expertise. The recent deluge of data brought on by faster detectors and brighter light sources has made this impossible for many beamlines, and the challenge will only increase as light sources pursue upgrades. To address this challenge, over the last few years we have developed new algorithms, software infrastructure, and tailored user interfaces that allow users to take advantage of high performance computing resources and to remotely visualize and interact with their data. Last year, ASCR issued a challenge to demonstrate the capabilities of our system. We set out to prove two points:

1. It is possible to use our system to analyze cutting-edge in situ experimental data at an ASCR facilities thousands of miles away and provide feedback fast enough to guide experiments as they happen.
2. It is possible for multiple BES facilities to share computing resources and algorithms located at one ASCR facility, in this case again to provide real-time feedback to users.

I will present the results of our data challenge, our end-to-end solution utilizing SPOT Suite, NERSC, ESnet, OLCF, CAMERA, TomoPy, and Globus Online, and present lessons learned. I will discuss our view of the future of big data at BES facilities and how the partnership between BES and ASCR can help address these challenges, producing higher science output of greater impact.