NERSC-In-The-Loop: Supporting Experimental Facilities at The National Energy Research Scientific Computing Center

Dr Bjoern Enders[1]
[1]*Lawrence Berkeley Natonal Lab*
benders@lbl.gov

Large-scale analysis of experimental data is an increasingly important type of workload at supercomputing facilities. The National Energy Scientific Computing Center (NERSC) is the mission HPC and data center for the United States Department of Energy (DOE), and has supported such experimental computing workloads since its inception as a general-purpose supercomputing facility. As these workloads at experimental facilities begin to exceed the capability of their on-site HPC clusters and storage [1], and with the advent of AI/ML in analysis and decision making at experimental facilities, the resources available at supercomputing facilities like NERSC become more attractive. Interfacing with experimental and observational facilities is a key part of the Superfacility concept [2], which envisions science teams at experiment facilities orchestrating automated data analysis pipelines which move data from the instrument to the computing site, perform analysis, and disseminate results - all without any human in the loop.
This presentation will cover the basis of the Superfacility project but focuses primarily on the key engagements and projects with high-data rate facilities like the Advanced Light Source, the National Center for Electron Microscopy, or the Linac Coherent Light Source and the tools [3] developed at NERSC to support them.

References
{1} Physics Today, "Synchrotrons face a data deluge", 2020, https://physicstoday.scitation.org/do/10.1063/PT.6.2.20200925a/full/, doi: 10.1063/PT.6.2.20200925a
{2} B. Enders et al. (2020), "Cross-facility science with the Superfacility Project at LBNL", 2020 IEEE/ACM 2nd Annual Workshop on Extreme-scale Experiment-in-the-Loop Computing (XLOOP), GA, USA, pp. 1-7, doi: 10.1109/XLOOP51963.2020.00006.
{3} D.J. Bard. et al. (2021). "Automation for Data-Driven Research with the NERSC Superfacility API". High Performance {4} Computing. ISC High Performance 2021. Lecture Notes in Computer Science, vol 12761. Springer, Cham. https://doi.org/10.1007/978-3-030- 90539-2_22