Contribution ID: **2**                                                                                                            Type: **not specified**

# Bursty Data Analytics on HPC

*Monday, 24 September 2018 10:15 (20 minutes)*

Abstract: We present a general overview of the challenges and opportunities in marshaling computational intensity around bursts of data generated in a (mostly) scheduled manner. Where computation "fits" in the data analytic pipeline (between detector and actionable knowledge) is an important architectural concern for advanced instruments with bursty data. Design boundary conditions include instrument duty-cycle, experiment predictability, the stubborn constancy of the speed of light, and a variety of data reduction opportunities and constraints. NERSC systems aim to capably capture the most intense computational peaks in these workflows. Opportunistically upstreaming computation in the analytic pipeline has significant promise in mitigating the "data deluge" through HPC-informed DAQ design, using NERSC systems to develop algorithms which can be back-ported to the DAQ system. Examples from LCLS and NCEM are presented with the intent of gathering future needs of DAQ designers.

**Primary author:**   Dr SKINNER, David (LBNL)

**Presenter:**   Dr SKINNER, David (LBNL)

**Session Classification:**   Computing and Storage

**Track Classification:**   Computing & Storage