

Computing (including Analysis)

Policies

LCLS users are responsible for complying with the data management and curation policies of their home institutions and funding agents and authorities. To enhance the scientific productivity of the LCLS user community, LCLS supplies on-site disk, tape and compute resources for prompt analysis of LCLS data, and software to access those resources consistent with the published [data retention policy](#). Compute resources are preferentially allocated to recent and running experiments.

Data Management

LCLS provides space for all your experiment's data at no cost for you. This includes the raw data from the detectors as well as the data derived from your analysis. Your raw data are available as XTC files or, on demand, as HDF5 files. The tools for managing files are described [here](#).

Running the Analysis

The analysis framework is documented in the [Data Analysis](#) page for the LCLS-I/HXR systems and [psana](#) for the LCLS-II (SXR&UED) systems. This section describes the resources available for running the analysis. The following figure shows a logic diagram of the LCLS data flow and indicates the different stages where data analysis can be performed in LCLS:

1. **Data acquisition** - The online monitoring nodes get detector data over the network and place it in shared memory on each node. There is a set of monitoring nodes allocated for each instrument. The detector data are received over the network by snooping the multicast traffic between the DAQ readout nodes and the data cache layer. Analysis performed at this stage provides < 1 s feedback capabilities. The methods for doing (quasi) real time analysis are described in the [Prompt Analysis](#) page. Users should be aware of the different possibilities and choose the approach that works best for their experiment.
2. **Fast feedback** - The processing nodes in the FFB system read the detector data from a dedicated, flash-based file system. It is possible to read the data as they are written to the FFB storage layer by the DAQ without waiting for the run to end. Analysis performed at this stage provides < 5 min feedback capabilities. The resources reserved for this stage are described in the [Fast Feedback System](#) page.
3. **Offline** - The offline nodes read the detector data from disk. These systems include both interactive nodes and batch queues and are the main resource for doing data analysis. We currently support sending the data to three offline systems: [psana](#), [S3DF](#) and [NERSC](#). The psana system is the default offline system and your data will end up in psana unless you arrange a different destination with your experiment POC. The psana system is also relatively old and it will be retired when more storage becomes available in the S3DF system. Please consider running at [NERSC](#) if you expect to have intensive computing requirements (> 1 PFLOPS).

