# Towards a BES Light Source Wide Event-triggered Tomography Data Analysis Pipeline Using a Sustainable Software Stack

Harinarayan Krishnan[1], Tom Caswell[2] and Daniel Allan[2]

[1]Lawrence Berkeley National Laboratory, Richmond, California, United States, [2]Brookhaven National Lab, Upton, New York, United States

Scientific and technological advances in instrumentation at Basic Energy Sciences (BES) facilities are expected to result in petabytes of data per month, the consequence of these advancements have resulted in an increase in data hungry algorithms running at beamlines nationwide to study never before seen phenomena with ever greater richness in detail. Upgrades of ALS-U, APS-U, and LCLS-II are expected to come online resulting in exponential growth rate in data volume outstripping any foreseeable technological innovation in storage capacities making data retention an extremely challenging task.

Therefore, the current mode of siloed data storage methodologies and custom analysis pipelines driven is no longer a tenable option. Additionally, the decentralized process of past efforts have also resulted in data that is not well described, either semantically or structurally which makes writing anything but one-off implementations difficult. Data from the light sources is uniquely varied (many scientific techniques) and quickly changing (compared to say, climate satellites or telescopes where the structure of the raw data is fixed and rigorously specified). Moving from post-facto analysis (when all of the relevant data is available) to streaming analysis (when only some fraction of the data is available and you only see it once) requires changes to algorithms and mind-set. Therefore, a comprehensive DOE wide data life cycle solution is essential for long term operational stability.

Data reduction is one such critical solution to keep data generation and data storage in parity. This requires tight cohesion and fast iterative feedback to determine between data critical to the experiment and noise and to drive the experiment to make the best use of beam time. However, the complexity can range from threshold-based on the fly compression to heuristics and adaptive data analysis pipeline incorporating the holistic experiment.

In response to this need, the DOE commissioned developers from each of the 5 light source facilities to build a long term strategy and solution designated as the BES Data Pilot project. The goal of the project is to develop a coordinated approach that leaves room for each facility to separately innovate and meet its own particular needs and deadlines while working on shareable core that can be co-maintained. A key step to this approach is to identify and work on development of deployment process on key techniques significant across all DOE facilities.

## Real time event-driven nano-tomography processing pipeline

One such key technique for BES is tomography, and addressing its data pipeline and processing requires understanding the intelligent data acquisition modalities can provide a foundational blueprint affecting all 5 light sources. The rest of the article delves into understanding the complexities of Tomography pipeline, specifically at the FXI beamline at NSLS-II and identifying elements unique to beamline operations and ones general across all light sources.

The goal: Run prompt tomographic reconstruction, piping data and necessary context to an external computational resource including secondary measurements and metadata about hardware configuration necessary to interpret the images and metadata about user intent.

The acquisition system at FXI is controlled by the Bluesky ecosystem (NSLS-II), a library for experiment control and collection of scientific data and metadata. The companion project Databroker provides a unified interface for data sources. The Bluesky system was extended to publish metadata events in real time. During development, performance considerations included handling race conditions such as SWMR, where simultaneous reads affect both quality and performance of the data.

In order to get from acquisition to compute, the data flow over the network at FXI had to travel across SSH gateway nodes performing multiple hops. This clear bottleneck in the pipeline was highlighted in feedback provided to NSLS-II staff that needs to be resolved.

A hybrid TomoPy and LTT MPI-based GPU reconstruction code base was run during the analysis phases of the pipeline. The major issues addressed during development required proper allocation of resources during selection of reconstruction modes as certain parameters and configurations could easily overwhelm the resources allocated on the system.

Finally, feedback included visualization of the reconstruction results to ensure successful execution of the entire end-to-end pipeline.

To ensure general reuse the development team also provide ways to leverage the existing pipeline to run in batch processing mode. This allowed staff and users to reconstruct their entire catalog automatically.

**Next Steps**

* Couple resource requests (e.g. auto-scaling).

* Perform full streaming: where reconstruction begins with partial data sets and provides incremental results during acquisition.

* Use inferences from the reconstruction to steer the acquisition process in a fully automated or semi-automated fashion.

* Add Xi-CAM visualization support

**From prototype to production**

In developing a prototype production ready pipeline for the FXI Nano-Tomography beamline at NSLS-II. The development team successfully developed and deployed an event triggered system starting from Bluesky producing unfilled document streams to executing MPI-aware parallel tomography reconstruction on remote computing hardware.

As part of the development process the pilot project identified race conditions in disk access during acquisition, dealing with back pressure during data transfers over the network, handling firewalls between different computational resources, and dealing with dynamic resource allocation for running large scale real time Tomography reconstruction. The next major phase is to understand and optimize the performance behavior in streaming metadata & data simultaneously, reducing network congestion, allocating resources to keep up with real time data acquisition.
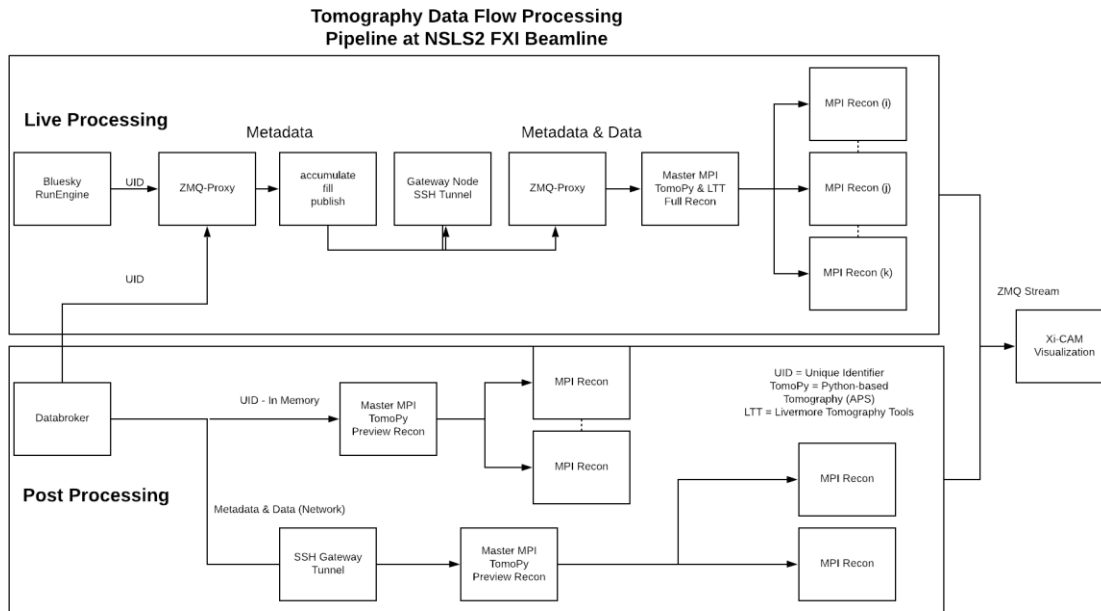
**Figure 1.** Prototype Real time data processing pipeline with for live and batch modes at the FXI Tomography beamline at NSLS-2

References
Allan, Daniel & Caswell, Thomas & Campbell, Stuart & Rakitin, Maksim. (2019). Bluesky's Ahead: A Multi-Facility Collaboration for an a la Carte Software Project for Data Acquisition and Management. Synchrotron Radiation News. 32. 19-22. 10.1080/08940886.2019.1608121.
Gürsoy D, De Carlo F, Xiao X, and Jacobsen C. Tomopy: a framework for the analysis of synchrotron tomographic data. *Journal of Synchrotron Radiation*, 21(5):1188–1193, 2014.
Pandolfi, R. J., Allan, D. B., Arenholz, E., Barroso-Luque, L., Campbell, S. I., Caswell, T. A., Blair, A., De Carlo, F., Fackler, S., Fournier, A. P., Freychet, G., Fukuto, M., Gursoy, D., Jiang, Z., Krishnan, H., Kumar, D., Kline, R. J., Li, R., Liman, C., Marchesini, S., Mehta, A., N'Diaye, A. T., Parkinson, D. Y., Parks, H., Pellouchoud, L. A., Perciano, T., Ren, F., Sahoo, S., Strzalka, J., Sunday, D., Tassone, C. J., Ushizima, D., Venkatakrishnan, S., Yager, K. G., Zwart, P., Sethian, J. A. & Hexemer, A. (2018). J. Synchrotron Rad. 25, 1261-1270.
Livermore Tomography Tools (LTT): https://nci.llnl.gov/content/assets/docs/ChampleyALSmeeting.pdf