Dataflow at the COSMIC Beamline - Stream Processing and Supercomputing

Bjoern Enders[1,2,*], Kasra Nowrouzi[1,2], Harinarayan Krishnan[1], Stefano Marchesini[1], Jungjin Park[1], Young-Sang Yu[1] and David A. Shapiro[1]

[1.] Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA.
[2.] University of California at Berkeley, Berkeley, CA 94720, USA.
[*] Corresponding author, benders@lbl.gov

During the last decade, ptychography has emerged as an invaluable method for X-ray microscopy at synchrotron facilities. It combines scanning microscopy with coherent diffractive imaging, delivering quantitative images of extended specimens at diffraction-limited resolution, surpassing the limits of focusing optics. Hence, ptychographic microscopes position themselves as the natural successor of conventional scanning transmission X-ray microscopy (STXM), since future upgrades at synchrotrons promise almost entirely coherent sources. Then again, probing the sample both in real and in reciprocal space (i.e. 4D) to obtain its goal, ptychography demands for its phase retrieval algorithms to operate on very large data sets which, in turn, poses high demands to the computational hardware.

The Nanosurveyor 2 endstation at the COherent Scattering and MICroscopy (COSMIC) beamline 7.0.1.2 of the Advanced Light Source (ALS) [1] pairs unparalleled stability with a large piezo scan range of 100 μm while its geometry allows for resolution elements in the single digit nm regime. Even without using the coarse translation stage (as in ref. 2), it harbors the potential for microscope images with >100M resolution elements, computed from >1G diffraction images. In this presentation, we discuss latest developments at COSMIC to facilitate ptychographic image retrieval that is quasi-synchronous with a high-bandwidth data stream from its FCCD detector [3] and strategies to deal with big data sets anticipated to be more common in the future.

The data processing pipeline for ptychography at COSMIC consists of three distinct steps. First, the raw diffraction data frames need to be grabbed from the detector and bundled with the scan information made available from the scan control software. A second preprocessing step involves a diverse set of actions: background correction, digitizer equilibration, noise suppressing measures, assembling high dynamic range images from a series of exposures, cropping around the optical axis and binning/interpolation. In the third, last, step the microscopic image is reconstructed from ptychography software like SHARP [4] or PtyPy [5]. Ideally, all steps are integrated in one scalable quasi-synchronous application as conceptualized in ref. 6 where simultaneous image reconstruction for a simulated beamline was demonstrated. However, to replace an entire data processing pipeline in a production environment, a modular and granular approach is favorable, as it allows for piecewise upgrades, and thus a smooth transition with extensive testing.

For the first step, to replace the former monolithic and non-extensible frame grabber binary, the current software splits the task in several pieces (see Figure **2**). An intermediate frameserver buffers the data stream from the camera interface node into a ZeroMQ publisher socket. Both a viewer app for instant visual feedback and the main data control app subscribe to the frameserver asynchronously. The control app uses a single continuous thread to receive the buffered frames, while it launches one-off write threads to save the raw data and metadata to a storage backend. As the saving thread runs independently at the pace of the storage backend, stress from the burst-like data acquisition is decoupled from the continuous but latency-prone process of saving data. Furthermore, the software has been designed with

enough flexibility to bypass the storage backend altogether, once the average data rate outruns the storage backend's I/O bandwidth as it is expected with future detector upgrades.
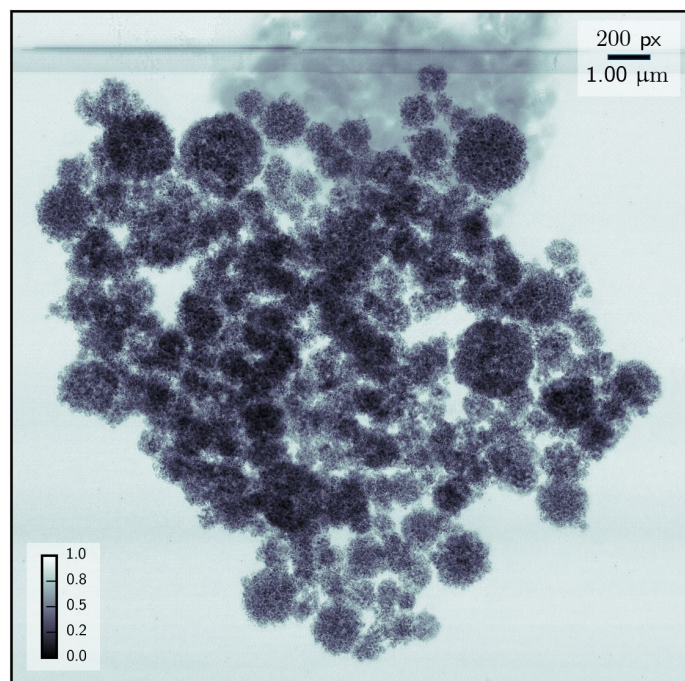

Figure. 1. ~4k×4k ptychographic image reconstructed at 5 nm pixel size from 170k diffraction images. Specimen of Fe_3O_4 primary and secondary particles as a battery anode material.
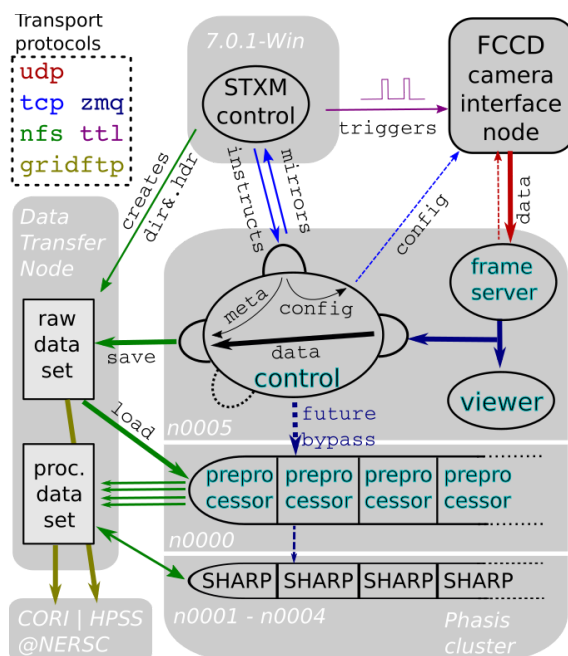

Figure. 2. Schematic of the firmware currently in place at COSMIC. Gray areas correspond to computing resources. Turquoise letters mark new or updated software. Elliptic boundaries correspond to an application or process.

The second step has also seen fundamental upgrades in that the new software is now able to preprocess the data to similar quality as previous scripts, but it is able to do so in smaller contiguous chunks, allowing the preprocessor to be launched at the same time as the scan begins and with the ability to push to the experimental streaming interface of SHARP and/or write to the open-ended data format of PtyPy, allowing for quasi simultaneous ptychographic reconstruction in both cases.

While the two previously mentioned improvements cater to an ever increasing data rate, latest developments aim to support datasets of unprecedented volume. For example, the current storage backend used at COSMIC constitutes a dedicated data transfer node, configured to transfer via the Energy Sciences Network to other DOE facilities at rates of up to 10 Gb/s . If a large dataset outgrows the compute resources of the on-site cluster, the CORI supercomputer at NERSC can now be used to reconstruct the ptychographic image using PtyPy. Figure **1** displays such an example reconstruction carried out at CORI, using 32 nodes at 16 cores per node, running at ~3 seconds per iteration.

References:
[1] R. Celestre, *et. al.,* J. Phys.: Conf. Ser. **849** (2017) 012047.
[2] M. Guizar-Sicairos, *et.al.*, Opt. Express **22** (2014) 14859-14870.
[3] P. Denes *et. al.,* Rev. Sci. Instrum. **80 (**2009) 083302.
[4] S. Marchesini, *et al.*, J. Appl. Cryst. **49** (2016) 1245-1252.
[5] B. Enders, P. Thibault, Proc. R. Soc. A **472** (2016) 064, http://dx.doi.org/10.1098/rspa.2016.064.
[6] B. J. Daurer *et al.*, Adv. Struct. and Chem. Imag. **3** (2017) 7.