

High data rate processing – a puzzle of metadata, compression and software

M. Müller and A. Förster

DECTRIS Ltd., Täfernhof 1, 5405 Baden-Dättwil, Switzerland

marcus.mueller@dectris.com

Latest-generation Hybrid Photon Counting detectors (HPC) like the new EIGER X series are necessary to make full use of the most brilliant synchrotron light sources. Continuous read-out at frame rates of hundreds of hertz enables synchrotron serial crystallography and 2D scanning of even large grids with beams a few micrometers across. Following some recent examples of serial crystallography applications, pressing questions of data handling and processing will be discussed.

At full speed, EIGER produces data at an internal rate of 40 Gbit/s. To minimize data transfer, network load and storage costs, effective compression is necessary. Experiments performed at numerous synchrotrons show that LZ4 compression [1] preceded by bit-shuffling [2] guarantees small files that can be decompressed without adding appreciably to the read time. To process these data efficiently, i.e. automatically, accurate metadata have to be written. EIGER follows the standards set by the NeXus International Advisory Committee [3], which will be presented briefly.

Normally, data processing programs are responsible for reading the data and processing them. We propose separating these two tasks and demonstrate the power of this approach with a plugin [4] that can be used by XDS [5] to read HDF5 data in a parallel way. This specialization ensures a significantly improved read performance and considerably accelerated processing.

When data, metadata, compression, and software are skillfully combined, robust automated data processing routines can be established even at high data rates.

1. <http://lz4.github.io/lz4/>
2. K. Masui *et al.*, arXiv:1503.00638 [astro-ph.IM]
3. M. Könnecke *et al.*, J Appl Crystallogr 48, (2015), 301-305
4. <https://github.com/dectris/neggia>
5. W. Kabsch, Acta Cryst D66, (2010), 125-32