



Contribution ID: 781

Type: **Contributed talk**

## Imaging and Real-Time Data Processing at the Terabyte Scale

*Friday 30 August 2024 12:30 (15 minutes)*

The gradual shift to 4th-generation synchrotron sources has been boosting data production rates that nowadays easily exceed 1-10 TB of produced raw image data per day. This has not only created a big data bottleneck in terms of data storage, management, processing, and visualization, but also requires a revised approach in utilizing state-of-the-art data processing techniques. In the recent past, several open-source efforts have been established, ranging from reconstruction algorithms [1-3], volume data analysis and visualization tools [4] to storage and metadata handling frameworks [5]. However, these can be considered “silo” solutions that not necessarily interface with each other smoothly. Moreover, in many cases the transition from single proof-of-concept-based code toward making software available to a wider (big) imaging community is still far from being established.

Both at TOMCAT and at LNLS, we have independently been developing and utilizing state-of-the-art processing tools. As one of the biggest data producers of the Swiss Light Source, the TOMCAT beamline has guaranteed smooth user operation for over a decade now thanks to an efficient data pipeline [6]. More recently, real-time reconstruction capabilities [7] and TB-sized volume analysis tools [8] have been added to ease the data analysis challenge. At LNLS, we have pioneered novel GPU-enhanced tomographic reconstruction algorithms [9], including phase recovery filters in both directions [10]. Furthermore, we have advanced in-memory processing capabilities, allowing visualization with optimized and fast rendering using the Nvidia/Index API [11] as well as on-the-fly segmentation [12], both using the RDMA protocol. In summary, all these strategies represent cornerstones that can now be integrated into a generalized platform as well as conventional graphical interfaces that will further allow the development of user-defined plugins and facilitate rapid exchange.

In the present work, we devise and describe a general architectural concept for data flow in typical tomographic imaging experiments, which involves both the underlying application stack with its “gluing” components as well as a full operating model in a standardized HPC environment. We present and discuss the feasibility of tomography scans with processing times of several seconds up to a minute, for volumes of several hundreds of GBs. To achieve stable operation, we leverage recent developments in IT architectural frameworks and apply industry-standard best practices for modularity, virtualization, and CI/CD. We show how our architecture drastically improves the user experience. Finally, we discuss different implementations of GPU/CPU communication and present benchmarks of reconstruction and visualization tools using different hardware.

### References:

- [1] D. Gürsoy, F. De Carlo, X. Xiao et al., *J. of Synchrotron Radiat* 21, 1188, 2014.
- [2] V. Nikitin, *J. Synchrotron Radiat* 30, 179, 2023.
- [3] W. van Aarle et al., *Opt. Express* 24(22), 25129, 2016.
- [4] A. Aboulhassan et al., *J. Imaging* 8(7), 187, 2022.
- [5] Moore, J., Allan, C., Besson et al., *Nat. Methods* 18(12), 1496, 2021.
- [6] F. Marone, A. Studer, H. Billich et al., *Adv. Struct. Chem. Imaging*. 3(1), 1, 2017.
- [7] J.-W. Buurlage et al., *Sci. Rep.* 9(1), 18379, 2019.
- [8] A. Miettinen, I. V. Oikonomidis, A. Bonnin et al., *Bioinformatics* 35(24), 5290, 2019.
- [9] E. X. Miqueles et al., *PPSC 2020*. <https://doi.org/10.1137/1.9781611976137.3>
- [10] E. X. Miqueles, P. Guerrero, *Results Appl. Math.* 6, 100088, 2020.
- [11] T. V. Spina et al., *JACoW 2021*, <https://doi.org/10.18429/JACoW-ICALEPCS2021-FRBL05>
- [12] A. Pinto et al., *Synchrotron Radiat. News* 35(4), 36, 2022.

## **I plan to submit also conference proceedings**

No

**Primary authors:** Mr STUDER, Alain (Paul Scherrer Institut); Dr PEIXINHO, Alan (Brazilian Synchrotron Light Laboratory, CNPEM, Campinas, Brazil); MIQUELES, Eduardo (Brazilian Synchrotron Light Lab. (LNL-S-Sirius)Eur.XFEL (European XFEL)); MARONE, Federica (Paul Scherrer Institut); LOVRIC, Goran (Swiss Light Source); SALA, Leonardo (Paul Scherrer Institut); Prof. STAMPANONI, Marco (ETH Zürich - Paul Scherrer Institut); FERRAZ, Paola (Brazilian Synchrotron Light Lab. (LNLS-Sirius)Eur.XFEL (European XFEL))

**Presenter:** LOVRIC, Goran (Swiss Light Source)

**Session Classification:** Mikrosymposium 3/3: Data, Automation and the Use of AI

**Track Classification:** 3. Data, Automation and the Use of AI