

# Towards end-to-end data-management for large scale x-ray facilities

*Monday, 9 December 2019 13:40 (30 minutes)*

Large scale scientific facilities, including x-ray facilities, face an extreme growth in data from instruments. With x-ray instruments data grow exponentially with the increased size of detectors, another exponential factor from the frequency one many sample with and finally x-ray sources are not robust enough that a large set of experiments can be automated, bringing a large increase in the number of experiments an instrument can perform in a session. Thus storing the data alone is a challenge.

The challenges are furthered from the fact that detector-size has grown to a resolution where samples, at least tomograms, cannot fit in the memory of a PC for data-analysis, and thus must be moved onto server-class computers with sufficient memory to hold a raw-data sample and a processed version as well. The increase in data-rate and number of experiments also mean that running through all samples manually easily becomes unfeasible and some means of batch processing must be introduced.

A final challenge is that users of x-ray facilities is widening and many of the new users are not comfortable with data-analysis and need to work with others in that part. This means that large communities, typically geographically distributed, need to collaborate on these, very large, datasets.

The talk will presents our ideas for an integrated solution to the above problems, include status and plans, and also introduce – for discussion – an idea for such an integrated system to help fight scientific fraud.

The work is supported in parts by H2020 European training network MUMMERING.

**Presenter:** Prof. VINTER, Brian (Copenhagen University)