PAUL SCHERRER INSTITUT

PSI

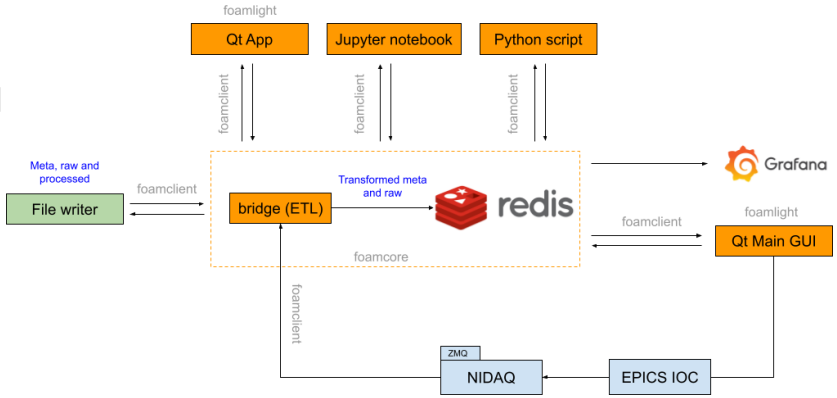

Markus Janousch :: AWI :: Paul Scherrer Institut

# Update of the group “Data Processing Development and Consulting (7902)”

AWI Department Update Meeting, March 06, 2023

# SLS 2.0

- BEC (Ivan)
  - Work on development environment and GUI-frameworks for the BEC. **Daiquiri UI** (https://ui.gitlab-pages.esrf.fr/daiquiri-ui/) and Flint (https://bliss.gitlab-pages.esrf.fr/bliss/master/flint/flint_data_plotting.html) as candidates.
- Debye (Jun)
  - Discussed and agreed upon the interface between the online data processing pipleline and the NIDAQ with Adam and Alvin.
  - Further development of the online data processing pipeline (chained data processing, file writing, etc.).
  - Started to test the NIDAQ implementation with the hardware and helping Alvin implement the data stream interface.



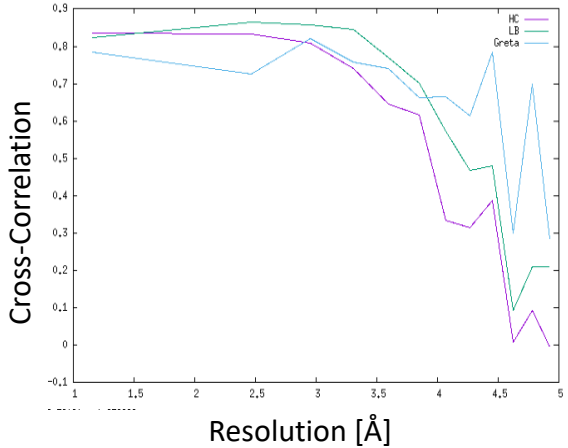
# SLS 2.0 / SLS

- TOMCAT (Jun and Alain)
  - Replacement of the original primitive implementation of data serialization with **Protobuf**, in order to use gRPC internally and exchange data with other data analysis tools written in different languages.
  - Tensor Tomography for TOMCAT
    - Focus lies on a precise mathematical formulation of the tensor tomography problem
    - Implementing / Testing simulation code to prepare for analysis of real data
    - https://gitlab.psi.ch/studer_a1/tomcat/-/tree/master/TensorTomo
  - Rewinding the TOMCAT Reconstruction Pipeline
    - Goal: Going back from derived data (reconstructions) to raw data (projections)
    - Motivation: Simulate ideal raw data (i.e. an ideal Beamline)
    - https://gitlab.psi.ch/studer_a1/tomcat/-/tree/master/RewindTPL
    - https://jira.psi.ch/browse/TCRP-113

# RED-ML (Piero and Hans-Christian)

- **Wrapping up the REDML project** and unifying codes and results from all the partners involved
- Writing a **paper on real-time data reduction pipelines** for serial crystallography with collection rate above 1 KHz
- Realising an **open-source, easy-to-use code supporting REDML's** results dissemination, as well as code for **production runs @SLS serial crystallography experiments**
- Helped **debugging and reimplementing** existing processing pipelines for MX's serial crystallography **data analysis based on CrystFEL**
- Brute force sampling on GPU, refinement so far on CPU.
- Speed ~1.4ms on my Laptop per file, for 10 files provided by Filip, with refinement dominating. Probably needs improvement.
- C++ API, Python and Spack (Elsa) modules.

Indexing Lysozyme Data



# SwissFEL and SINQ (Ivan)

- Photon Diagnostics: Added spectrum intensity correlation and spectrum number of peaks statistics for PSSS camera data stream (to be used / tested during the next week at SwissFEL).
- pyzebra: A new set of processing and displaying options for (m)hkl scan positions and peaks in the corresponding data files (ccl, hdf) after multiple reviews is finally "graduated" from test to prod server (after this point pyzebra is in a "bugfix only" state from our side)

# LEAPS-INNOV

- Annual Meeting end of Mach at DESY (https://indico.desy.de/event/37196/).
- Blosc (https://www.blosc.org/) proposal
  - Faster read / write access to compressed HDF5 datasets
  - Optimized synchrotron data compression pipelines
  - jpeg2000 support

# Questions?

**Many thanks to my group members!**