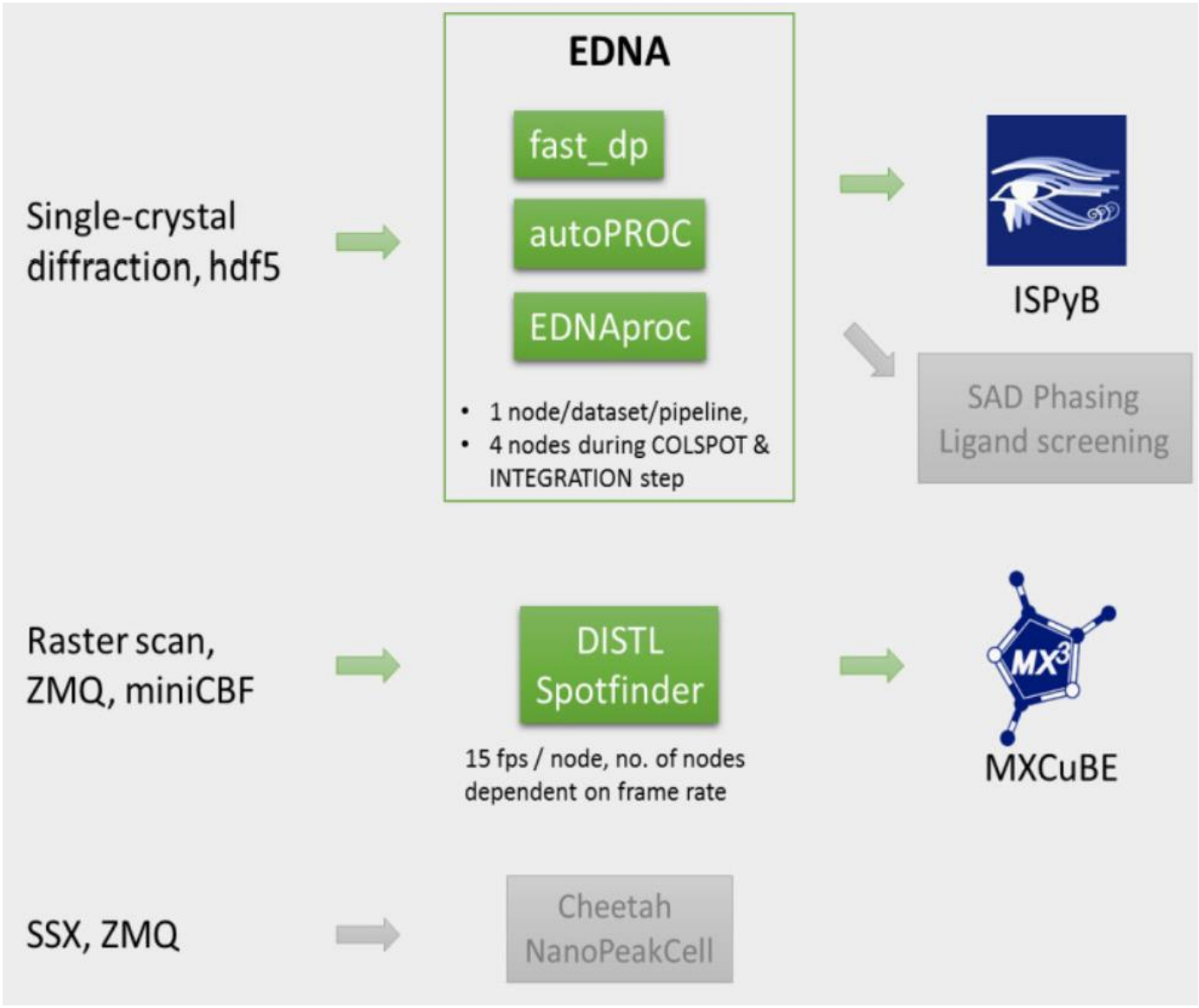
Autoprocessing at Synchrotrons

- MAX IV
- SOLEIL
- DESY
- ALBA
- HZB
- EMBL-HH
- ELETTRA

Data processing workflow at BioMAX



Ongoing

Three pipelines
- autoPROC with staraniso
- Fast_dp or EDNAproc
- Dials/Xia2

- Optimization usage of the HPC nodes

- Improvement of DISTL & Labelit
- Dozor, cheetah!?

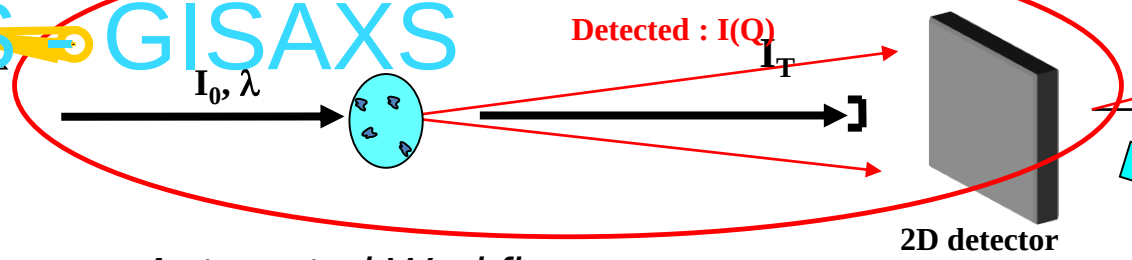
AUTOBIOSAXS

SAXS data reduction and analysis process automation

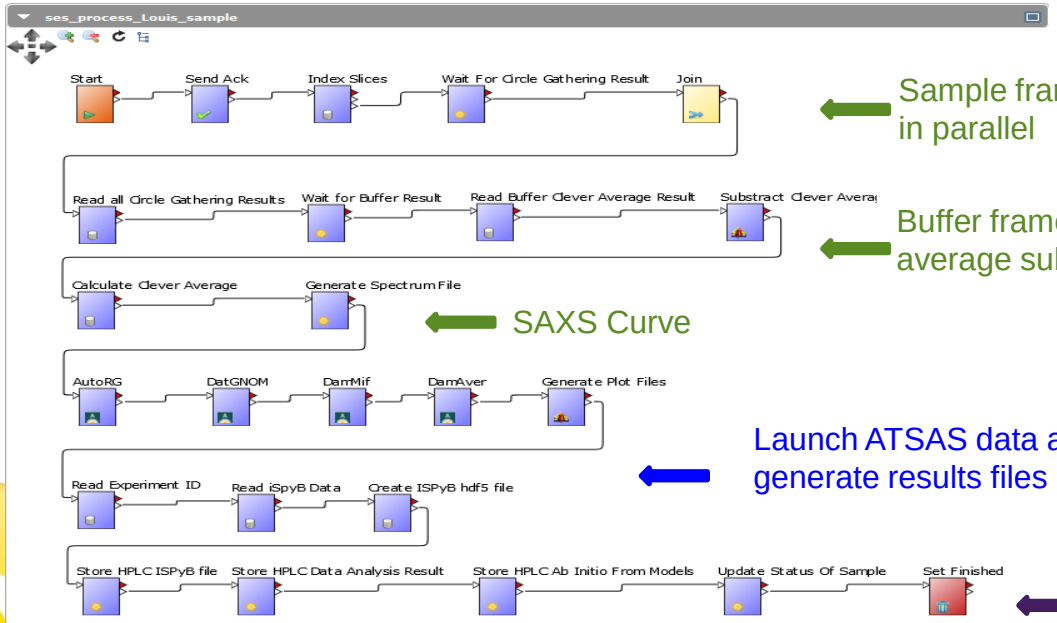
Data Acquisition

WAXS - GISAXS

SWING



Automated Workflow
Deployed on SES Passerelle web server



Sample frames are all reduced in parallel

Buffer frames are all reduced in parallel and average substracted from Sample frames

SAXS Curve

Launch ATSAS data analysis steps and generate results files

Push results to ISPyB portal

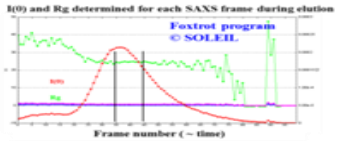
Data Reduction

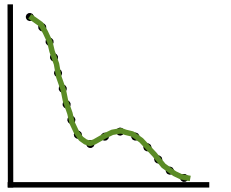
Radial average + buffer subtraction

Foxtrot program

xenocs

SOLEIL SYNCHROTRON

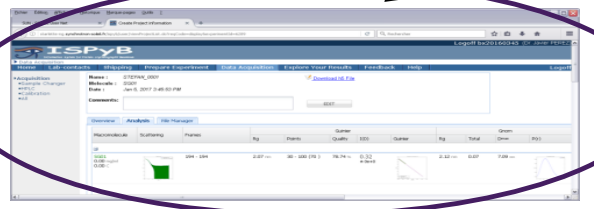




Data Analysis

(ATSAS suite of programs)

EMBL



Web Access to results

- **3 Automated workflows**
  - Dozor
  - AutoPROC
  - XDSME (made by P. Legrand, based onn XDS)

- **Implementation (using haskell language)**
  - Web Server
  - Job submission
  - Job Execution

- **Triggering**
  - Via MXCube

- **Foreseeing**
  - Reprocessing functionalities (How to link to IspyB?)

Autoprocessing

Status of Beamline P11

Standard Collection

- xdsapp is started on a separate machine via ssh.
- Two jobs are started, pre and full.
- presenterd keeps track of the progress, compiles a result HTML file and links it from an index file in the beamtimes root folder.

Characterisation

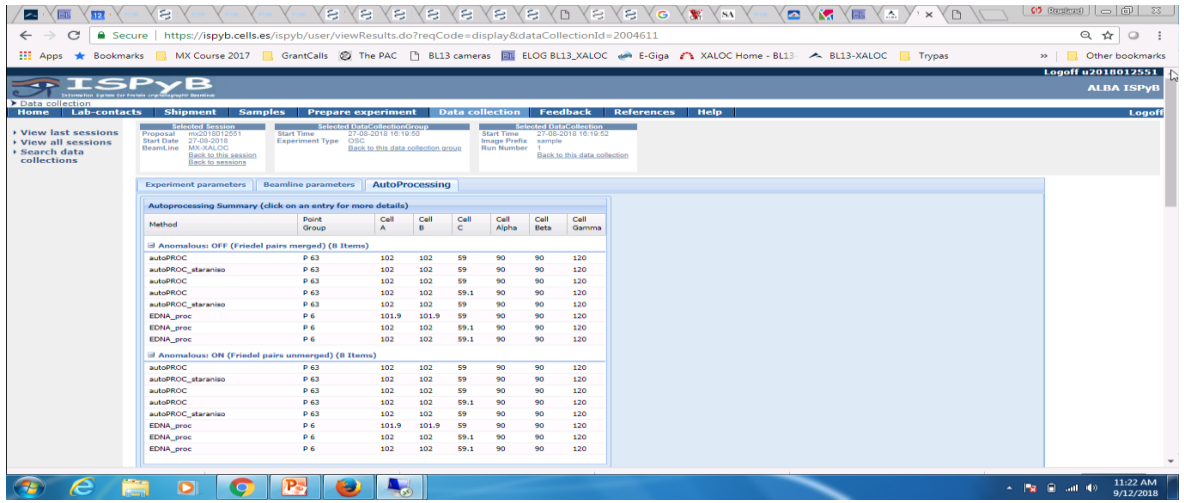
- mosflm is started on a separate machine via ssh.
- presenterd keeps track of the progress, compiles a result HTML file and links it from an index file in the beamtimes root folder.

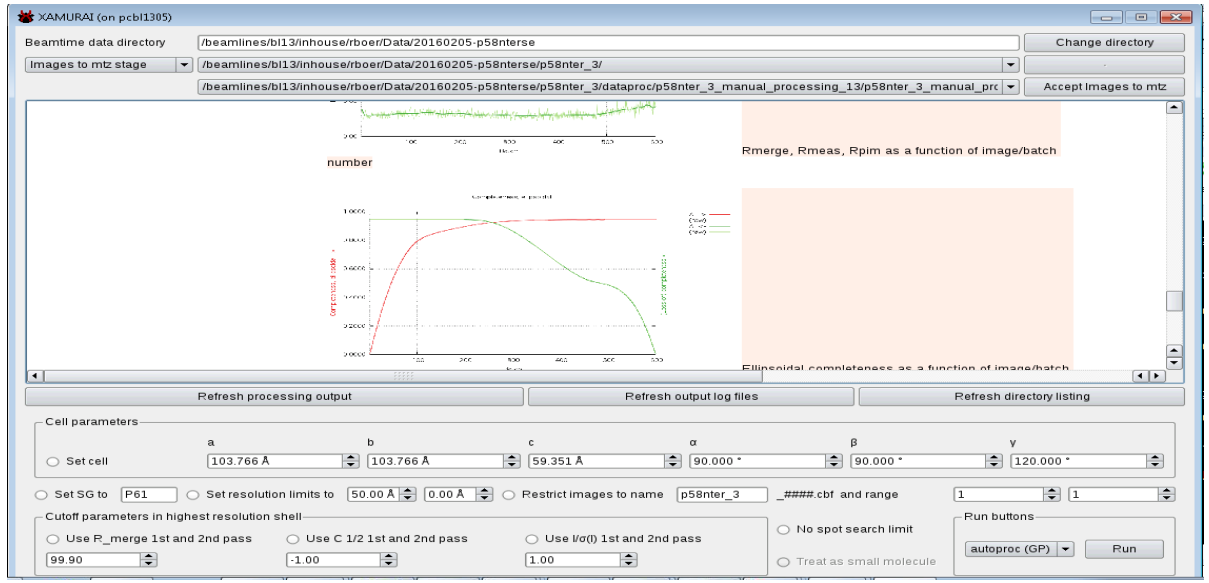
Grid Scan

- spotfinder from CCP4 is called via HTTP.
- Number of Bragg spots is visualised in a heat map overlay.

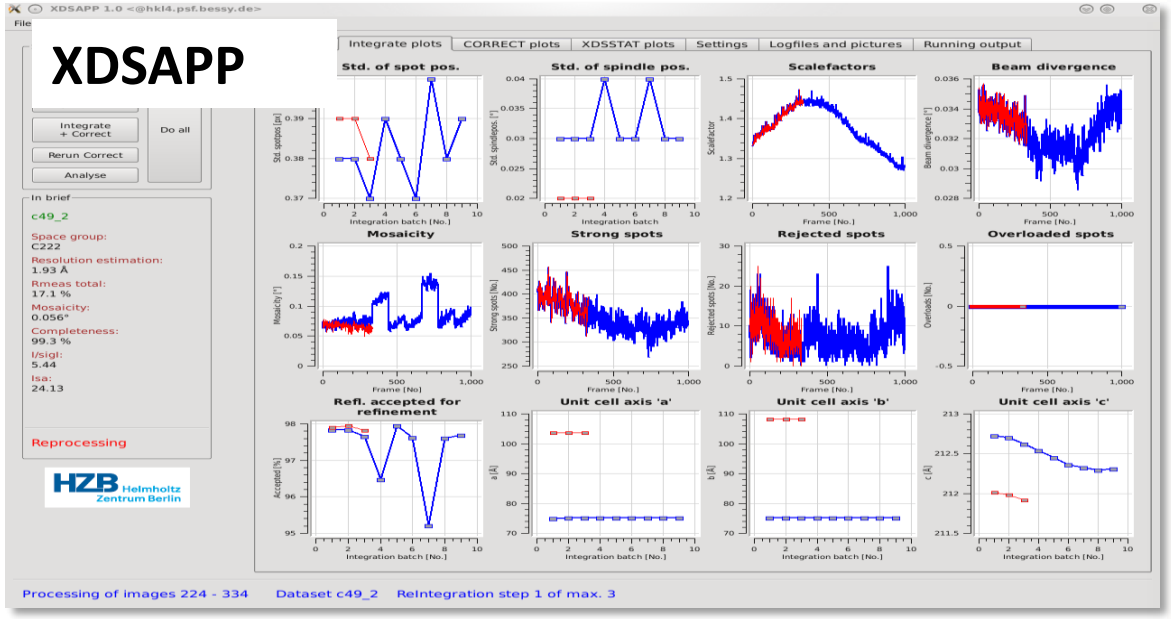
Data processing at XALOC (ALBA)

- Characterization using EDNA
- Default data processing (default parameters) using EDNAProc (fully automated in MxCube and ISpyB) and AutoPROC (script to start autoprocessing to be implemented, results are published in ISpyB)
- Manual processing using Xamurai (publishing in ISpyB pending). Common adjustments such as different resolution cutoffs, image ranges etc are implemented





Auto-processing @ HZB



M. Krug *et al.* (2012). *J. Appl. Cryst.* **45**, 568-572.
K. Sparta *et al*. (2016). *J. Appl. Cryst.* **49**, 1085-1092.

Currently: 445 users, 283 institutes, 36 countries

http://www.helmholtz-berlin.de/xdsapp

Program features

- Python QT-GUI
- Live graphical data analysis and display
- Automatic decision making
  - Smart reintegration cycles
  - Resolution range definition
  - Twinning and pseudo-translation analysis
- Alternative manual control of important processing parameters
- Live mode during data collection
- Data conversion to hkl, mtz and cv formats
- Command line version for fully automatic processing of all measured data sets in one directory tree

Main metrics used

- Isa (monitoring)
- R_{meas} (space group determination)
- CC(1/2)* (resolution cut-off)

Current situation/problem

- K. Röwer left 12/2017
- no development

Auto processing at EMBL-HH

Beamline P13 – PILATUS2 6M

EDNAProc on 40 cores (fast)

AutoPROC(GΦL) and **XIA-DIALS** on 12 cores (slow)

All processed data in ISPYB

Re-processing trigger (**EDNAProc**) in ISPYB

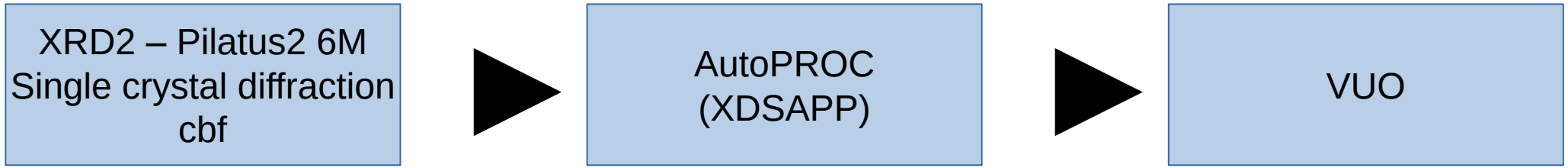
Beamline P14 – Eiger 16M

EDNAProc on 216 cores (still not fast enough)

AutoPROC and XIA-DIALS ***currently not run***, we are short on CPUs

All pipelines run via official EDNA plugins
All results are deposited in ISPYB, presented via ISPYB/EXI
(including e.g. StarANISO etc.)
There is no "decision making" involved

Data processing workflow at Elettra



Pipelines on Pilatus PPU but will be extended to cluster of 16 servers (256 CPU, 128 GB each, ...)