



| The European Synchrotron

REMOTE USER OPERATION OF ESRF EXPERIMENTS

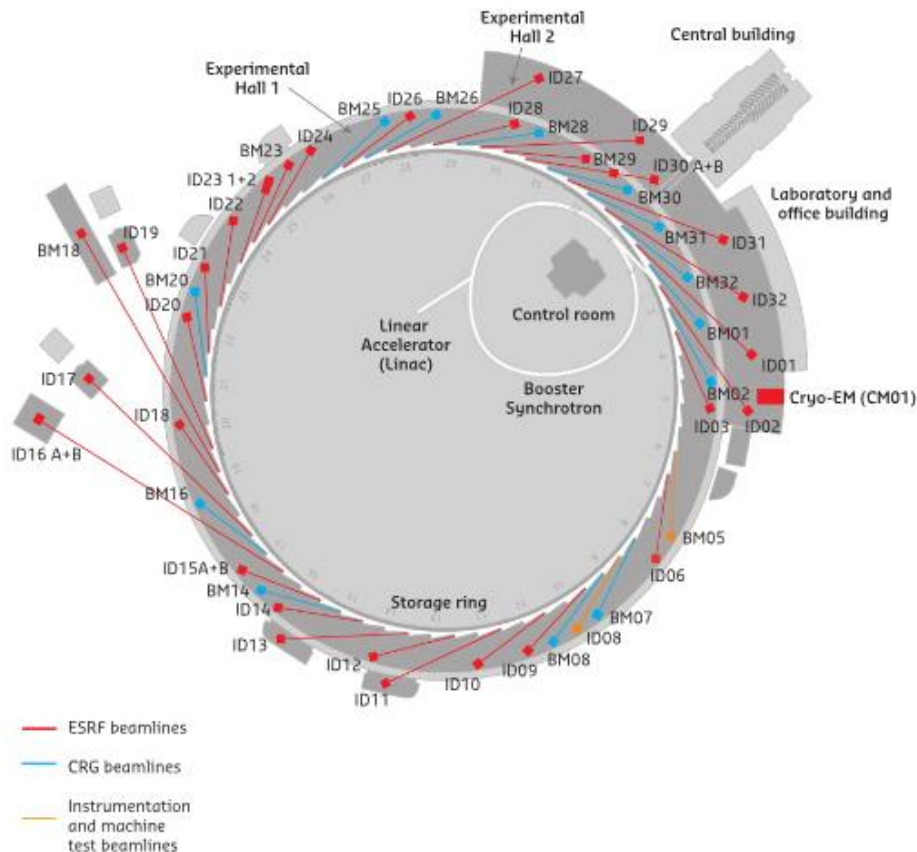


Outline

- **A Global view**
- **Remote Experiments Step by Step**
- **On-going work**

Jens Meyer on behalf of the ESRF Software Group

ESRF Beamlines and Experiment Types



| SOURCE POSITION | INDEPENDENT END-STATIONS | FIELD OF RESEARCH | STATUS | DATE |
|-----------------|--------------------------|---|--------------|-------------|
| ID01 | 1 | Microdiffraction imaging | Operational | Since 12/14 |
| ID02 | 1 | Time-resolved ultrasmall-angle X-ray scattering | Operational | Since 07/14 |
| ID03 | 1 | Hard X-ray diffraction microscopy | Construction | |
| ID06 | 1 | Large volume press /Hard X-ray diffraction microscopy | Operational | Since 10/13 |
| ID09 | 1 | Time-resolved structural dynamics | Operational | Since 09/94 |
| ID10 | 1 | Soft interfaces and coherent scattering | Operational | Since 06/12 |
| ID11 | 1 | Materials science | Operational | Since 09/94 |
| ID12 | 1 | Polarisation-dependent X-ray spectroscopy | Operational | Since 01/95 |
| ID13 | 1 | Microfocus | Operational | Since 09/94 |
| ID14 | 1 | Nuclear scattering | Construction | |
| ID15A | 0.85 | Materials chemistry and engineering | Operational | Since 11/16 |
| ID15B | 0.5 | High-pressure diffraction | Operational | Since 11/16 |
| ID16A | 1 | Nano-imaging | Operational | Since 05/14 |
| ID16B | 1 | Nano-analysis | Operational | Since 04/14 |
| ID17 | 1 | Medical | Operational | Since 05/97 |
| ID18 | 1 | Nuclear scattering | Operational | Since 01/96 |
| ID19 | 1 | Microtomography | Operational | Since 06/96 |
| ID20 | 1 | Inelastic X-ray scattering | Operational | Since 06/13 |
| ID21 | 1 | X-ray microscopy / IR spectroscopy | Operational | Since 12/97 |
| ID22 | 1 | High resolution powder diffraction | Operational | Since 05/14 |
| ID23 | 2 | Macromolecular crystallography MAD | Operational | Since 06/04 |
| | | Macromolecular crystallography microfocus | Operational | Since 09/05 |
| ID24 | 0.5 | Dispersive EXAFS | Operational | Since 12/21 |
| ID26 | 1 | X-ray absorption and emission | Operational | Since 11/97 |
| ID27 | 1 | High pressure | Operational | Since 11/21 |
| ID28 | 1 | X-ray scattering II | Operational | Since 12/98 |
| ID29 | 1 | Multiwavelength anomalous diffraction | Closed | Since 08/20 |
| ID30A | 2 | Macromolecular crystallography | Operational | Since 07/14 |
| ID30B | 1 | Macromolecular crystallography | Operational | Since 04/15 |
| ID31 | 1 | Interfaces and materials processing | Operational | Since 11/15 |
| ID32 | 1 | Soft X-ray spectroscopy | Operational | Since 11/14 |
| BM18 | 1 | Hierarchical tomography | Construction | Since 12/14 |
| BM23 | 1 | X-ray absorption spectroscopy | Operational | Since 01/21 |
| BM29 | 1 | Bio SAXS | Operational | Since 06/12 |
| CM01 | 1 | Cryo-EM | Operational | Since 11/17 |

Operation until 2018:

Users come to the ESRF and bring their samples

Users execute their experiments on-site and intervene in the hutches

Users transfer small amounts of data via the network or bring their own disks for large data volumes

Only structural biology beamlines are automated for remote operation

Operation since 2020

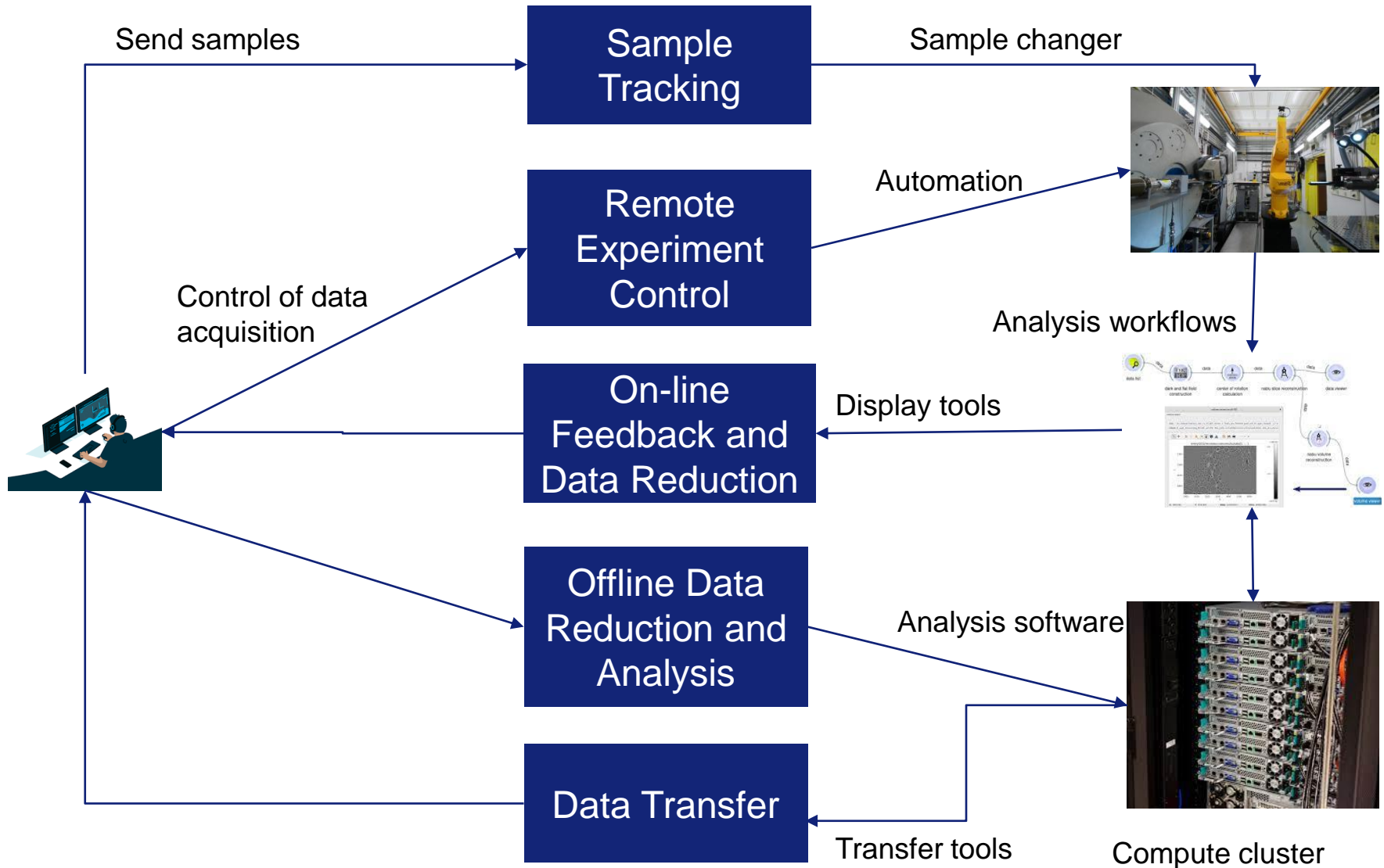
Commissioning of the new storage ring (EBS)

- More photons -> faster experiments and more raw data
- Data analysis needs more compute power
- Large raw data sets are more difficult to transfer (up to 100TB)

COVID restrictions -> No users on site

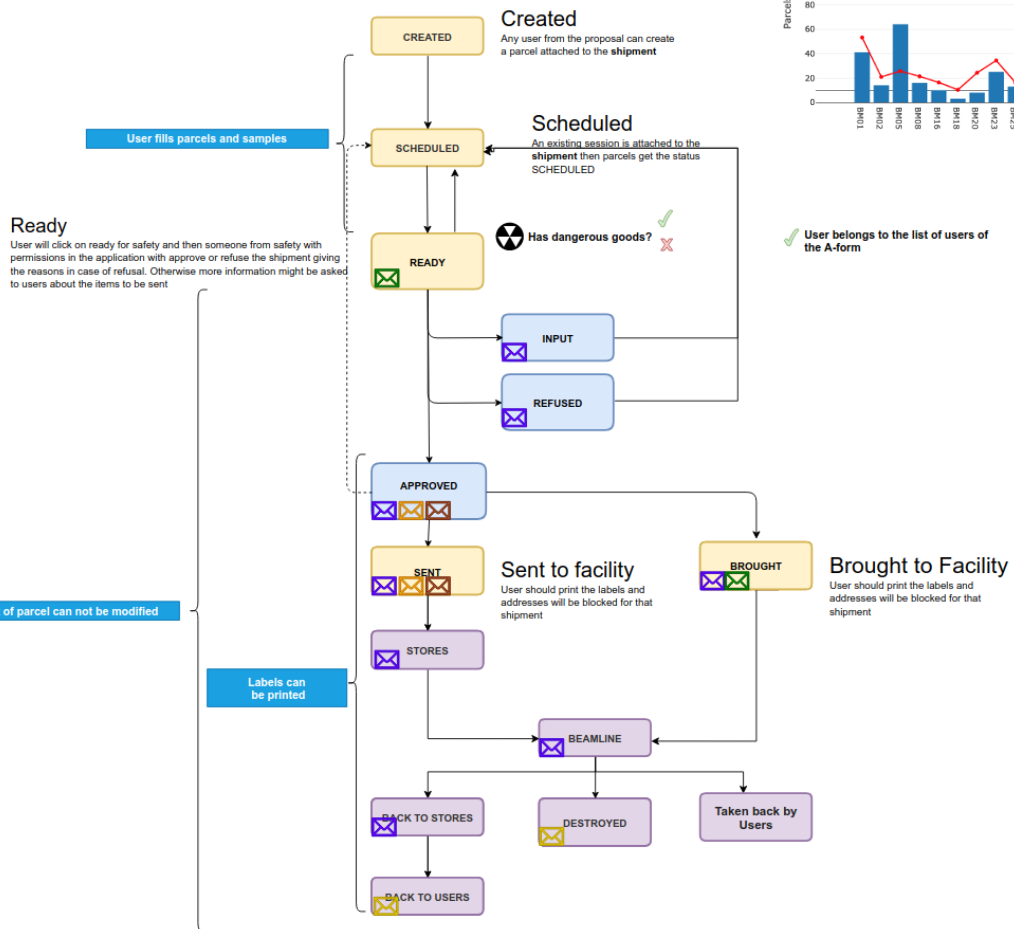
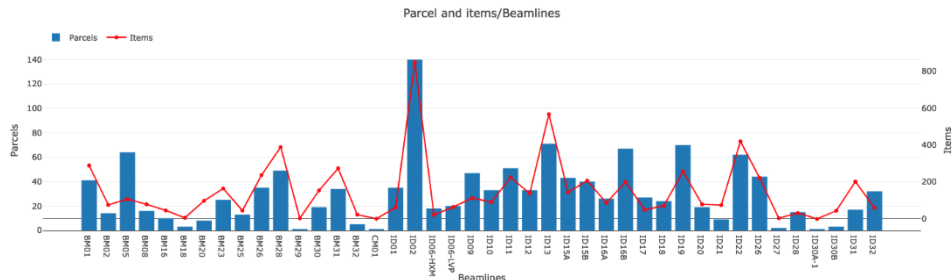
- Remote operation of all beamlines to run the ESRF
- Remote operation with different levels of automation
- Operation with limited local staff
- A fast solution for remote user operation was requested

A GLOBAL VIEW



SAMPLE TRACKING

Known from Structural Biology beamlines
 Sample tracking available for all beamlines



LEGEND

- Notification to expsaf@esrf.fr if contains dangerous goods
- Notification to safety group (expsaf@esrf.fr)
- Notification to users
- Notification to users, local contact and BDO (industry@esrf.eu) in case of IM/IN
- Notification to esrtransport@esrf.fr
- Notification to users and BDO (industry@esrf.eu) in case of IM/IN
- Action done by users
- Action done by safety
- Action done by staff (stores or local contacts)

Automation

- Sample changers (local staff cannot always be available to intervene in the hutch)
- High level data acquisition sequences
- Flexibility to allow users or local staff to adapt the data acquisition
- Standardized data format and data organization
- Experiment registration with meta data and electronic logbook (web applications)

The screenshot shows the ESRF Data Portal interface. At the top, there are navigation tabs: Data Portal, My Data, Open Data, Closed Data, Shipping, My Beamlines, and Manager. Below this is a search bar and a navigation menu with options like Dataset List, Logbook, Shipping, and Proposal. A prominent orange banner contains a message for users wanting to download large volumes of experimental data. The main content area displays a dataset summary for 'BlaC-D164_opt1_CS_2'. The summary includes details such as Name, Resolution (2 Å), Wavelength (0.966 Å), Exposure Time (0.1 s), Flux start/end, X and Y beam sizes, and Transmission (100%). There are three thumbnail images: a diffraction pattern, a 2D intensity map, and a photograph of the sample. Below the summary is a table listing files associated with the dataset, including file names, definitions, sizes, and download links.

| Date | Sample | Dataset | Definition | Files | Size | Download |
|------------------|---------------------|-------------------------------|------------|-------|--------|----------|
| 17:58 9 May 2018 | BlaC-D164_opt1_CS_2 | BlaC-D164_opt1_CS_2_1_2295406 | | 900 | 2.1 GB | Restore |

| Date | Sample | Dataset | Definition | Files | Size | Download |
|------------------|---------------------|------------------------------------|------------|-------|----------|----------|
| 17:57 9 May 2018 | BlaC-D164_opt1_CS_2 | ref-BlaC-D164_opt1_CS_2_4_2295405 | | 4 | 9.5 MB | Restore |
| 17:56 9 May 2018 | BlaC-D164_opt1_CS_2 | line-BlaC-D164_opt1_CS_2_4_2295404 | | 100 | 236.4 MB | Restore |
| 17:55 9 May 2018 | BlaC-D164_opt1_CS_2 | line-BlaC-D164_opt1_CS_2_3_2295403 | | 100 | 236.4 MB | Restore |
| 17:54 9 May 2018 | BlaC-D164_opt1_CS_2 | line-BlaC-D164_opt1_CS_2_2_2295401 | | 100 | 236.4 MB | Restore |
| 17:53 9 May 2018 | BlaC-D164_opt1_CS_2 | mesh-BlaC-D164_opt1_CS_2_1_2295398 | | 51 | 120.5 MB | Restore |



The screenshot shows the ESRF Data Portal interface displaying a list of events. The events are listed with timestamps and descriptions of dataset storage and archiving. Below the event list, there are several data plots. The first plot is a 2D intensity map with a color scale from -5 to 5. The second plot is a 1D intensity profile with a color scale from 0.0000 to 0.0007. The third plot is a 1D intensity profile with a color scale from 0.0000 to 0.0007. The fourth plot is a 1D intensity profile with a color scale from 0.0000 to 0.0007. The plots show data for different datasets and samples.

| Date | Sample | Dataset | Definition | Files | Size | Download |
|----------|--------|---|------------|-------|------|----------|
| 06:45:31 | | Dataset roi3290_3132 (416_LD_Zr8_incl1) | | | | Restore |
| 06:45:31 | | Dataset roi3290_3132 (416_LD_Zr8_incl1) | | | | Restore |
| 06:44:57 | | 416-LD-Zr8_incl1 | | | | Restore |
| 06:38:16 | | Dataset roi3289_3131 (416_LD_Zr8_incl1) | | | | Restore |
| 06:38:16 | | Dataset roi3289_3131 (416_LD_Zr8_incl1) | | | | Restore |

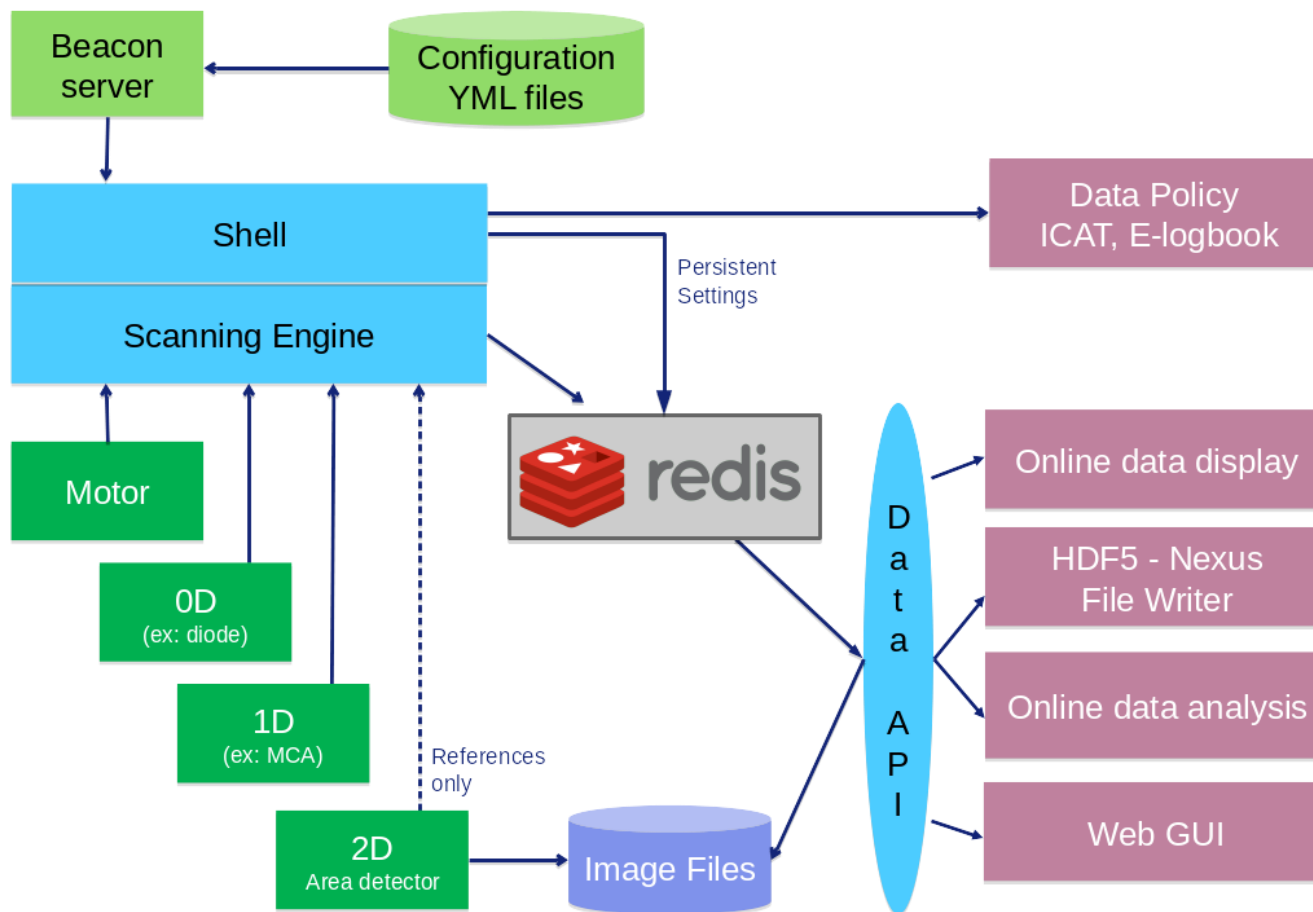
Bliss architecture and data flow



Experiment sequences written in pure python

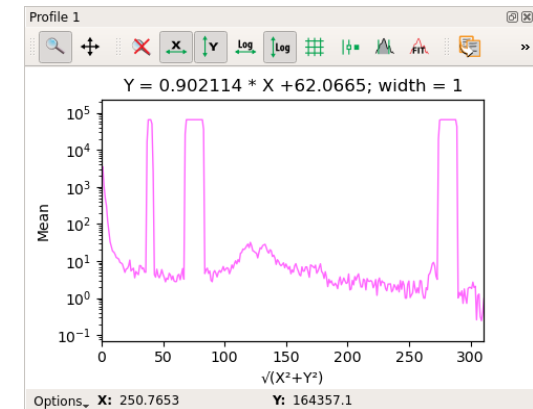
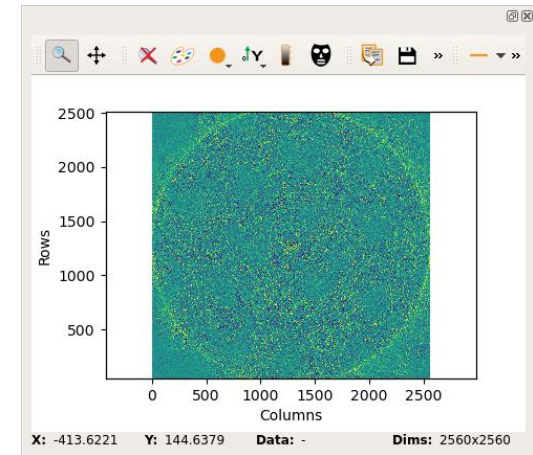
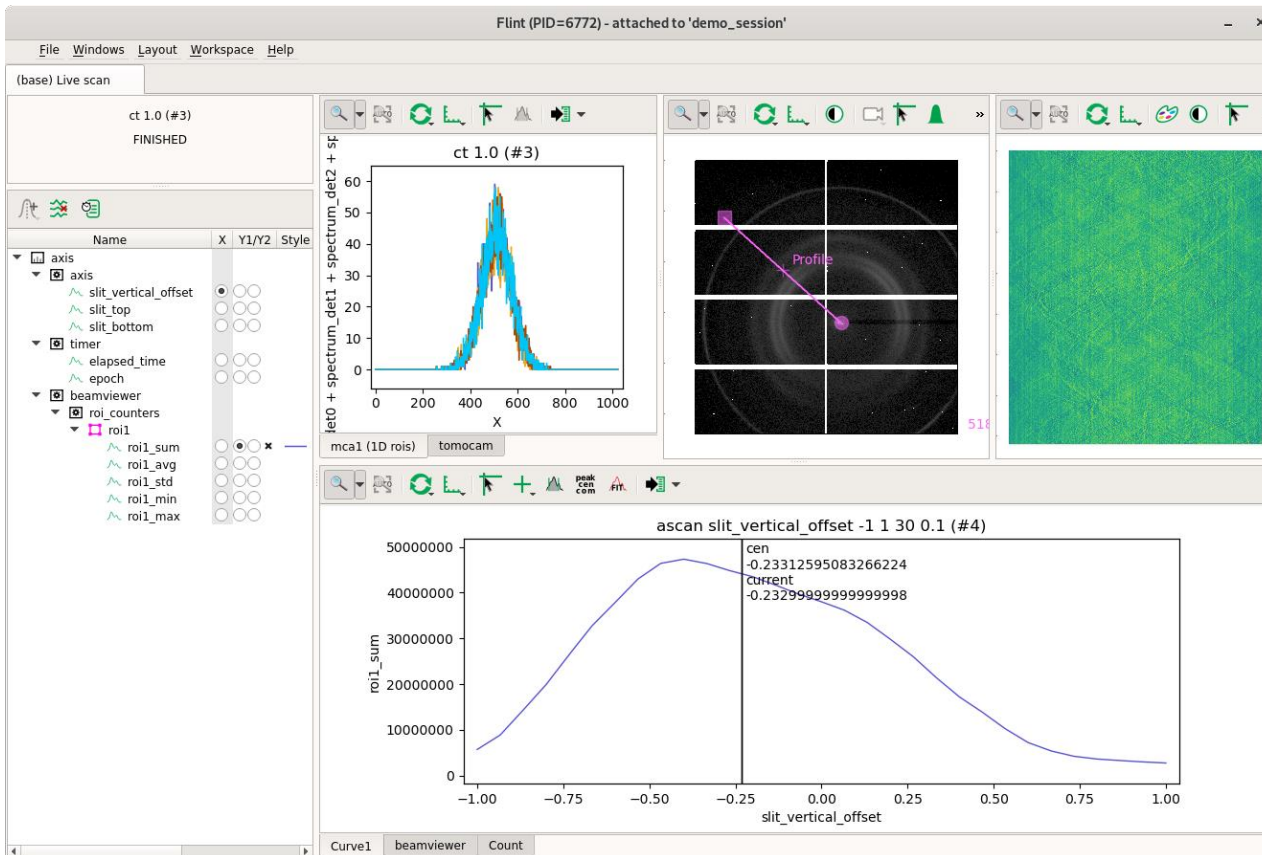
Unified scanning engine for step and continuous scans

Scan data and meta data saving in HDF5 file format following the NeXus definitions



Bliss architecture and data flow

Live data display



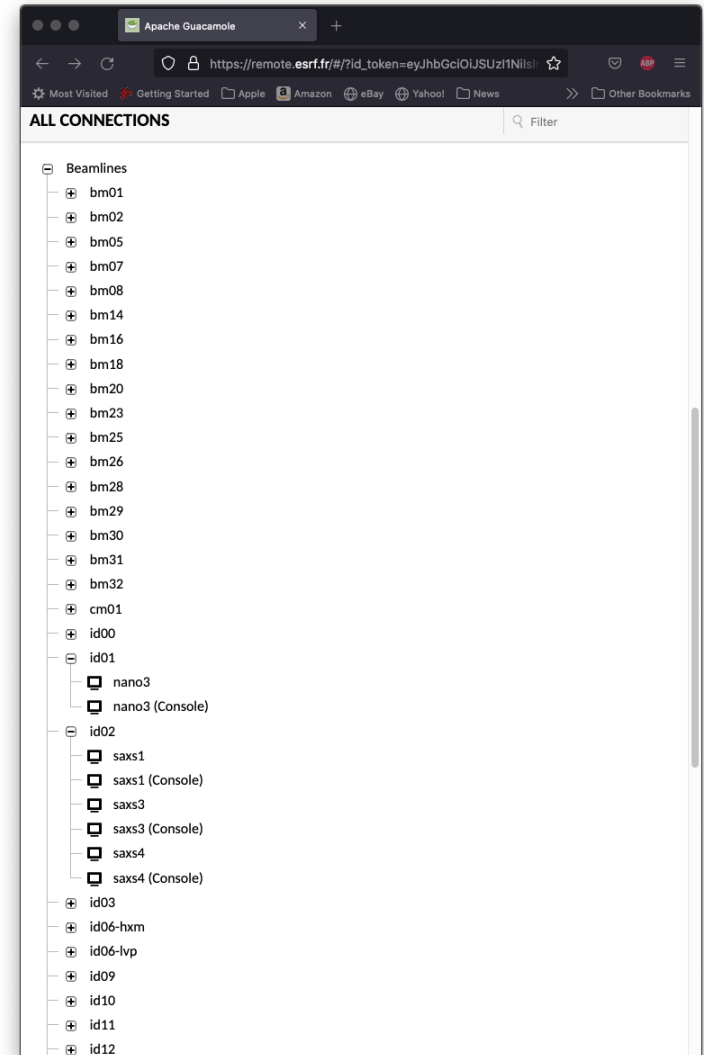
Remote operation for all beamlines

Not all GUIs are web application

Heavy use of the command line interface

Remote desktop to the beamline control workstations

- Guacamole for desktop sharing
 - > *Multiple screens*
 - > *Multi connections*
- Zoom for direct communication



Remote operation for fully automated experiments

Web GUIs

- Experiment control
- On-line feedback
- Pre-analysis for data validation
- Data reduction

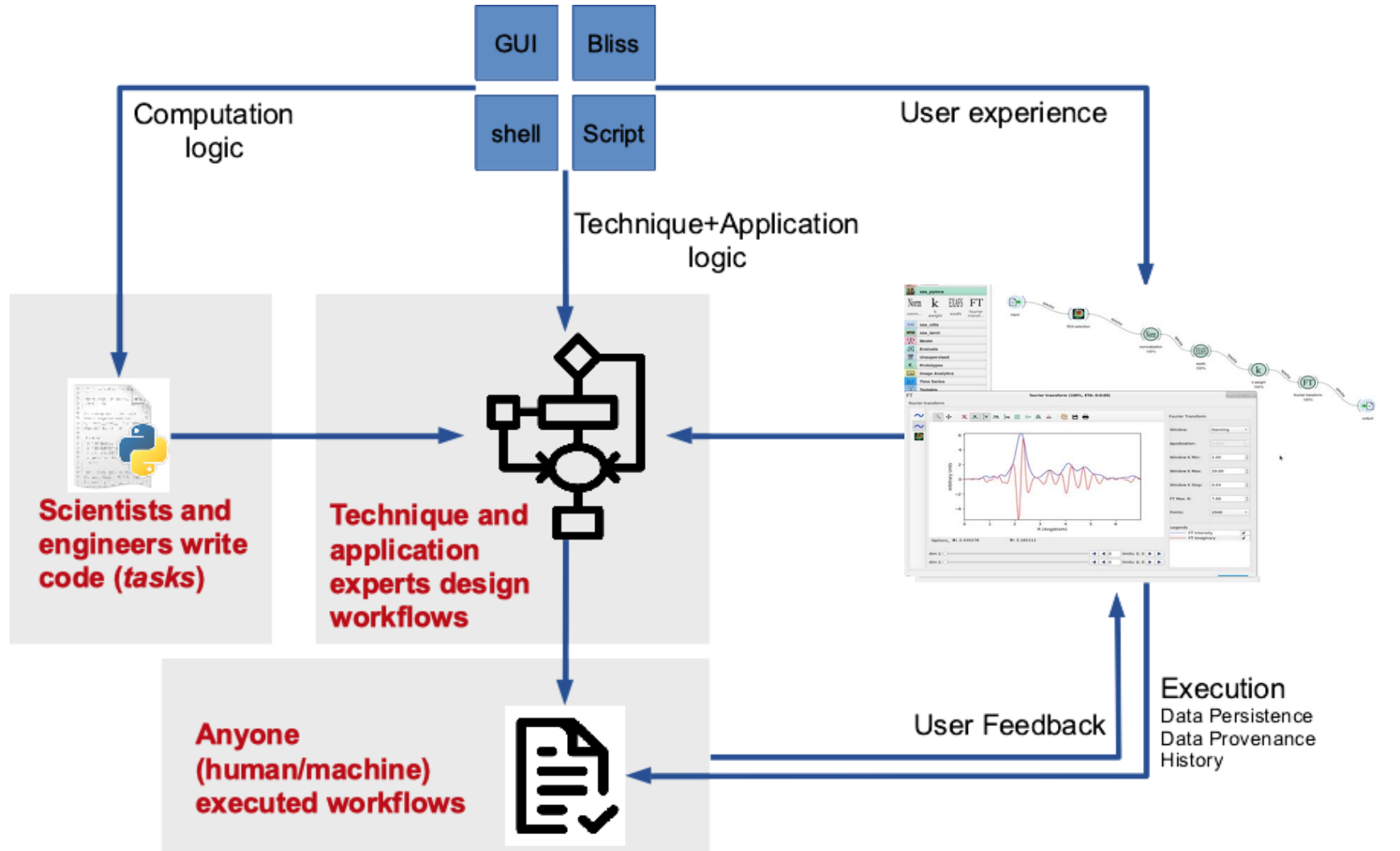
} Data processing workflows



| # | Type | Size |
|------|------|----------|
| 1625 | SC | 35x40 µm |
| 1626 | SC | 90x95 µm |
| 1627 | SC | 45x45 µm |
| 1628 | PC | |
| 1629 | PC | |
| 1630 | PC | |
| 1631 | PC | |
| 1632 | PC | |
| 1633 | PC | |
| 1634 | PC | |
| 1635 | PC | |

| Id | DC | ROI | Px | Py |
|--------------------------|-----|------|--------|-------|
| <input type="checkbox"/> | 779 | 1512 | S-Kal | 45 45 |
| <input type="checkbox"/> | 780 | 1512 | P-Kal | 45 45 |
| <input type="checkbox"/> | 781 | 1512 | Si-Kal | 45 45 |

Data Processing Workflows



Data Processing Workflows

Ewoks

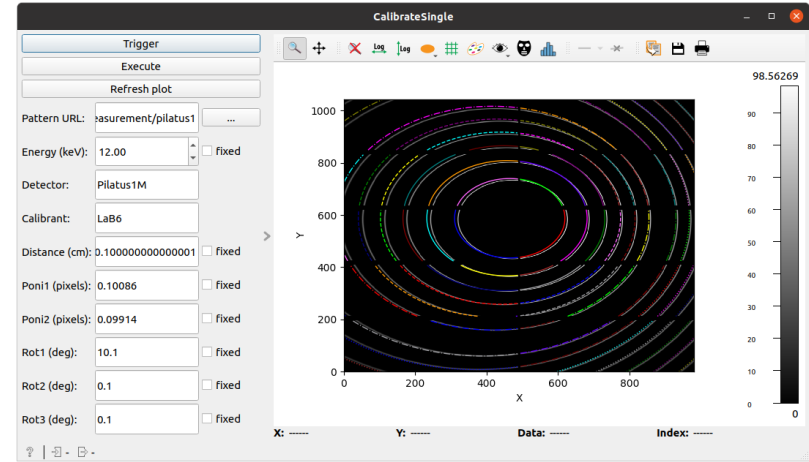
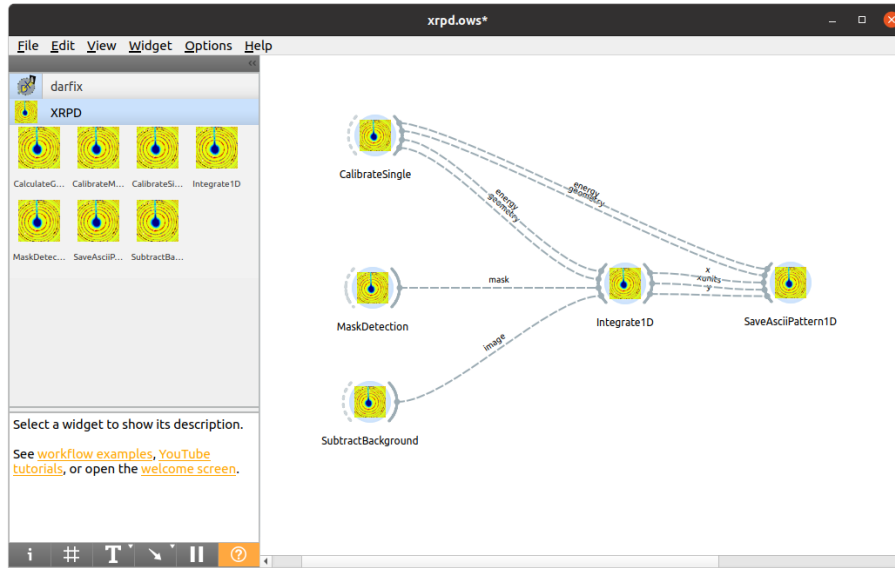
- Common workflow definition
- Common task definition
- Distributed execution (Dask, Slurm)
- Loops and conditions (pypushflow)
- Web GUI (Ewoksweb)
- Data persistence (HDF5, Redis)
- Execution events for client feedback

Projects

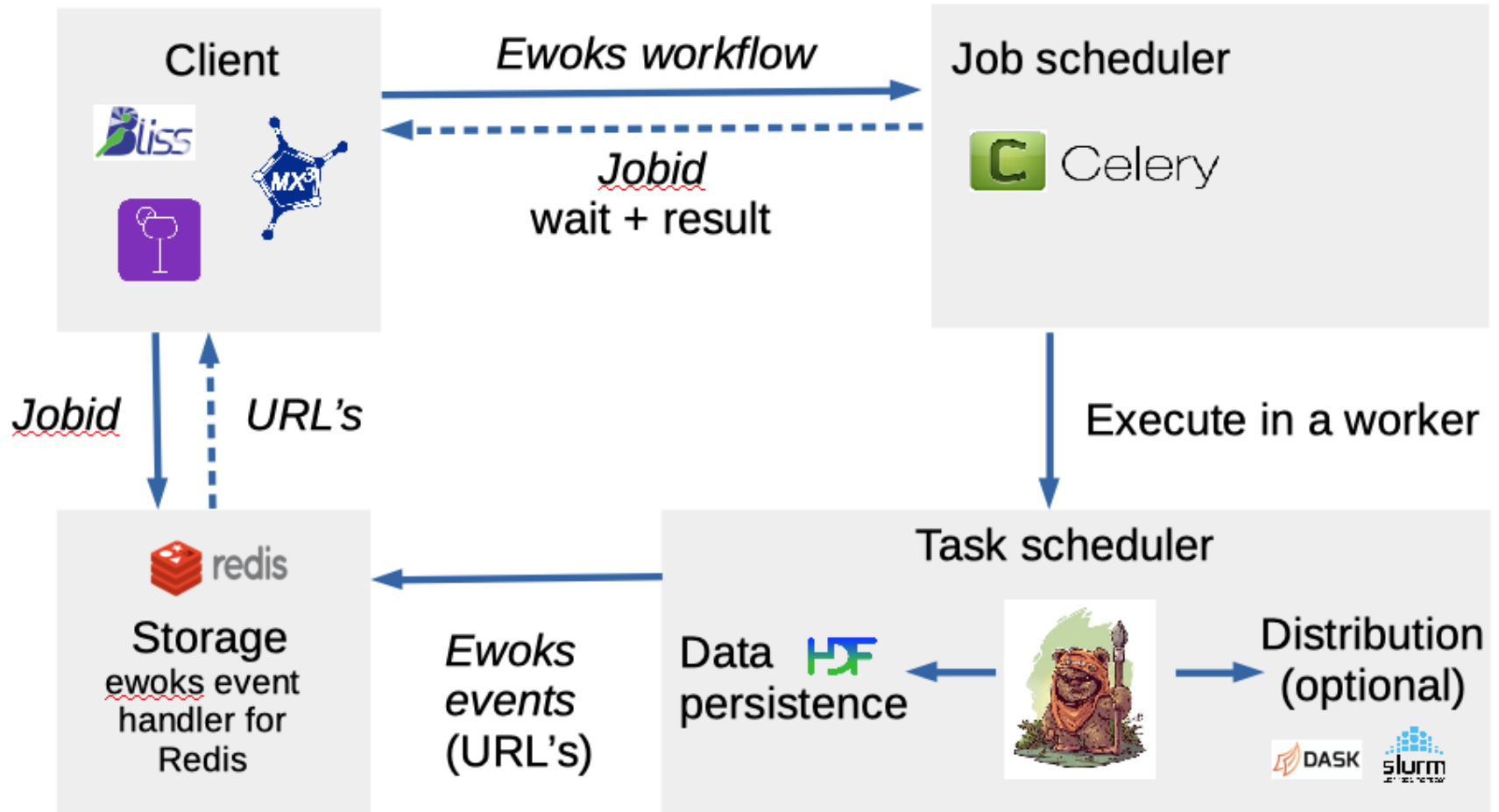
- Tomography (Tomwer)
- Darkfield microscopy (Darfix)
- MX (MASSIF-1)
- Spectroscopy (Est)
- EASI-STRESS (stress-strain analysis)
- STREAMLINE High-Throughput XRPD



DATA REDUCTION AND ONLINE ANALYSIS

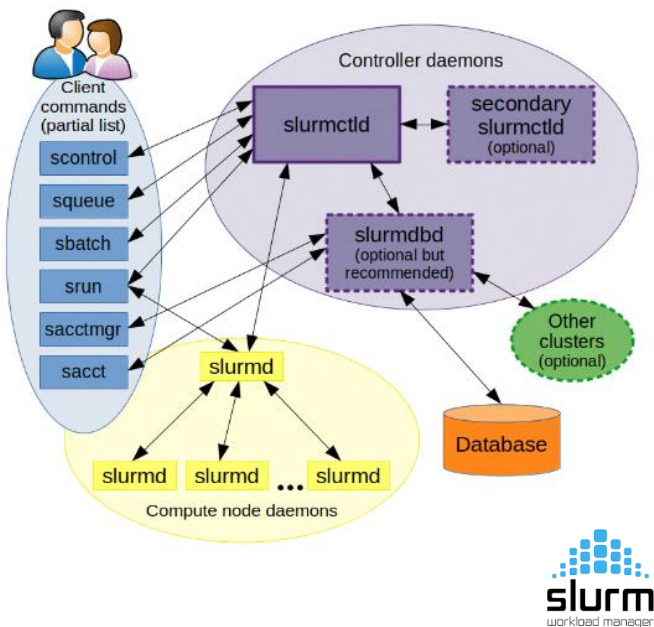


Woks for Online Data Processing



Remote data analysis at the ESRF

- Large data sets are difficult to transfer
- Need of optimized analysis software
- Need of a powerfull computing infrastructure
- Remote data analysis needs specific infrastructure to handle users and resources
 - > *Jupyter notebooks on Slurm*



Server Options

Simple Advanced

Partition

| | |
|---|---|
| Intel Xeon (x86_B6) Partition: jupyter-nice | IBM Power9 (ppc64le) Partition: jupyter-p9gpu |
|---|---|

CPUs

| Minimum | Quarter node | Half node | Entire node |
|---------|--------------|-----------|-------------|
| 1 core | 32 cores | 64 cores | 128 cores |

GPUs

| | | |
|---|---|---|
| 0 | 1 | 2 |
|---|---|---|

Options

Jupyter environment: Operating system (default) ▾

Launch JupyterLab:

Job duration: 1 hour ▾

List of available resources:

| Current Status | | |
|----------------|---------|---------|
| Partition | # nodes | # avail |
| jupyter-nice | 28 | 10 |
| jupyter-p9gpu | 8 | 4 |

Start

Higher degree of automation in data acquisition and analysis

- Need more standardization of the different experiment types
 - > *Definition of the experimental sequence, the options and the parameters*
 - > *Better standardization of the data structure produced*
 - > *Adapted data analysis software*
 - > *Agreement on data processing workflows*
- Efficient data reduction and pre-analysis
 - > *Users should keep only the reduced data*
 - > *Be sure data contains useful information*
- HPC computing for analysis at ESRF
 - > *Enlarge the infrastructure for remote access*
 - > *Make more data analysis software available*
 - > *Increase the computing resources*



Acknowledgements to

- The members of the software group for the development of all the different software tools
- Our colleagues from the Technical Infrastructure Division for their work on the network and computing infrastructure
- All the beamline scientists for their patience and problem reporting since the EBS startup.