Online calibration and analysis

Steve Aplin
Data Department EuXFEL

Material Provided by Philipp Schmidt (Data Analysis Group)

Time structure of data analysis

Freshness ⟷ Completeness

Online analysis

In real time during the experiment, delay of up to a few seconds

- Immediate feedback to run the experiment
- Performance critical
- Inflexible once prepared
- up to 15 GiB/s for AGIPD detector

Near-online analysis

Manually performed during the experiment within minutes of taking data

- Guides experimental decisions
- “Quick’n’dirty”
- Offers some degree of exploration
- Adapted quickly
- Typical run up to 1 TiB

Offline analysis

Publication-quality analysis, may take from hours to months

- Often “after the fact”
- All freedom and flexibility in the world
- Take all measured data into account, up to several PiB

Time structure of data analysis

Freshness ⟷ Completeness

Online analysis

In real time during the experiment, delay of up to a few seconds

- Immediate feedback to run the experiment
- Performance critical
- Inflexible once prepared
- up to 15 GiB/s for AGIPD detector

Near-online analysis

Manually performed during the experiment within minutes of taking data

- Guides experimental decisions
- “Quick’n’dirty”
- Offers some degree of exploration
- Adapted quickly
- Typical run up to 1 TiB

Offline analysis

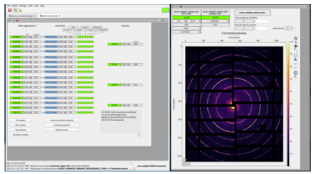
Publication-quality analysis, may take from hours to months

- Often “after the fact”
- All freedom and flexibility in the world
- Take all measured data into account, up to several PiB

Solutions to serve online analysis at EuXFEL

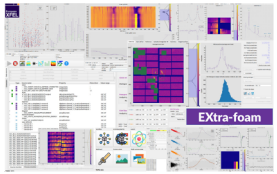
accessible ←

karabo-gui



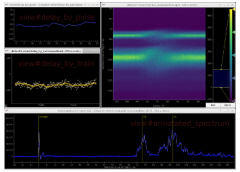
Data preview for diagnostics and monitoring

EXtra-foam



Rich GUI with common analysis features

EXtra-metro



Programmable in realtime to perform any analysis

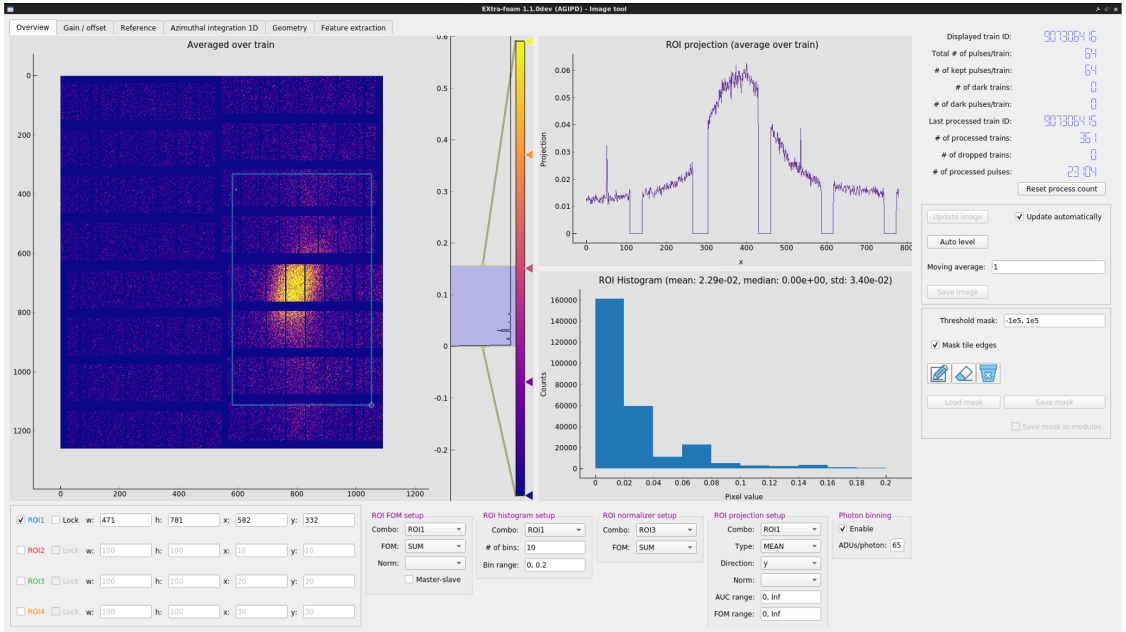
karabo-bridge

Stream data to an external application

→ **flexible**

Graphical analysis in EXtra-foam

- Rich GUI for 2D imaging experiments
- Provides many commonly used features ready-to-go
- Small *special suites* for graphical analysis of special experiments
- Analysis functionality exposed as Python library, core implemented in C++

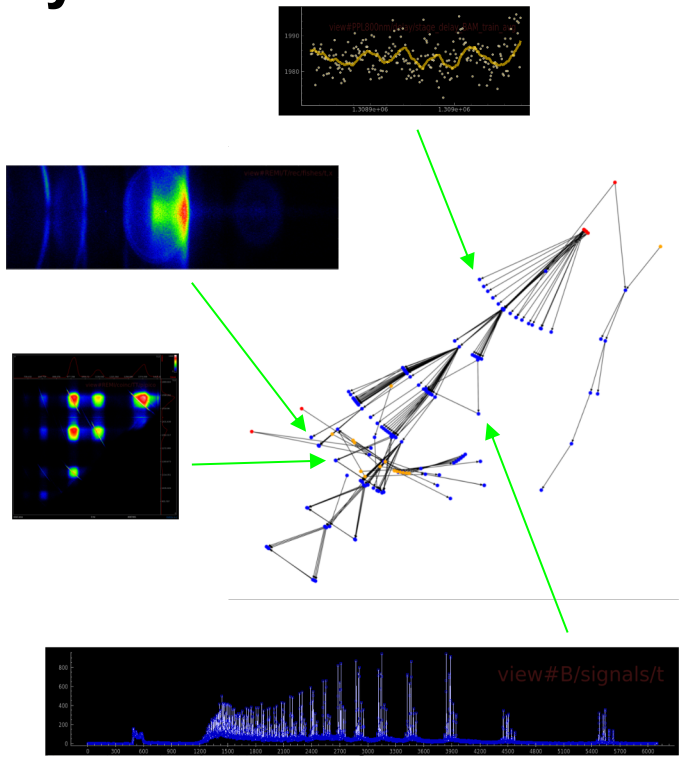


EXtra-metro for programmable online analysis

- Make writing online analysis easy
- Express analysis as a series of small steps, each result may be visualized
- Written in code that may be changed and re-injected at runtime
- Concentrate on the physics not the boilerplate, and let the framework take care of I/O or parallelization

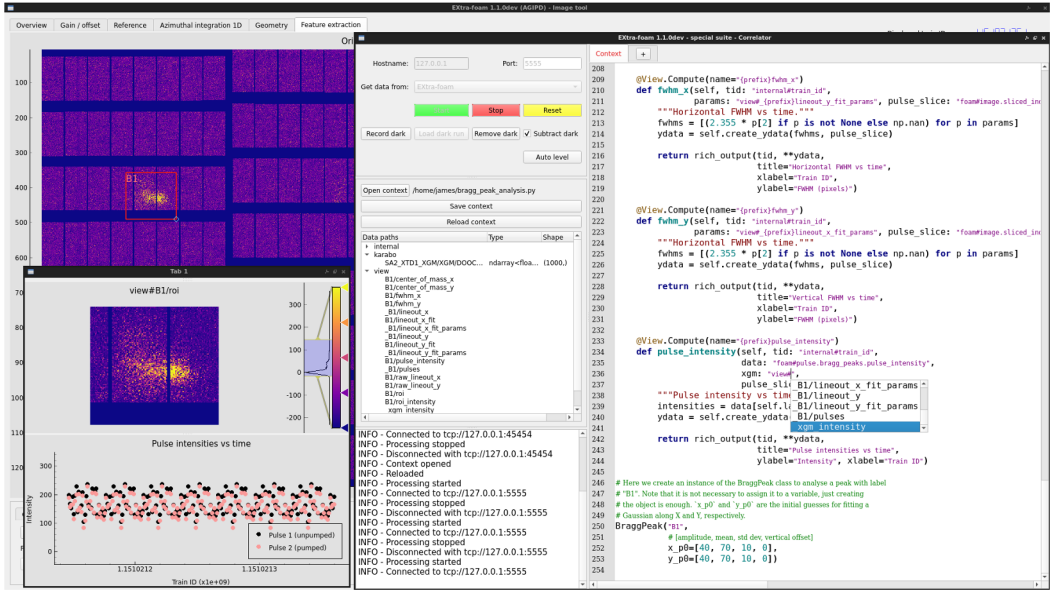
```
@View.Vector
def xgm_by_pulse(pulses: 'SA3_XTD10_XGM/XGM/DOOCS:output.data.intensitySa3TD'):
    """Pulse-resolved XGM intensity."""
    return pulses

@View.Scalar
def xgm_train_avg(pulses: 'xgm_by_pulse'):
    """Train-averaged XGM intensity."""
    return pulses.mean(axis=0)
```



The best of both worlds: EXtra-foam + EXtra-metro

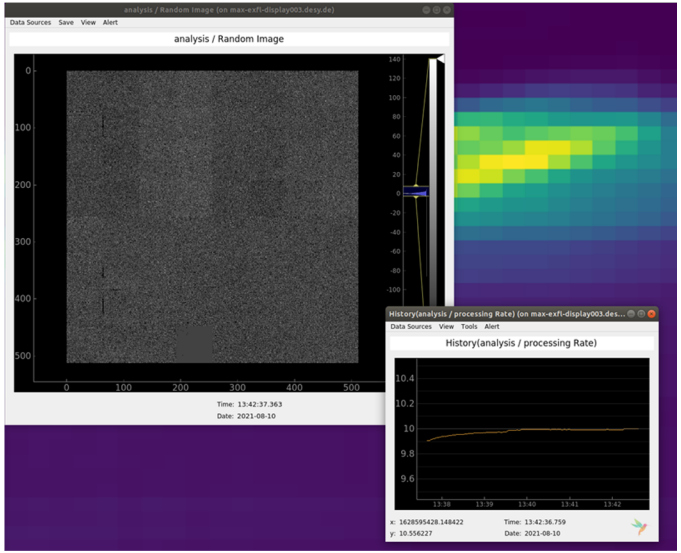
- Combine accessible GUI capabilities of EXtra-foam with optional programmability with EXtra-metro code
- Access to all the direct data from Karabo and precomputed EXtra-foam results
- Adapt now and without an expert:
  The best ideas happen right during a beamtime



Streaming data to users via karabo-bridge

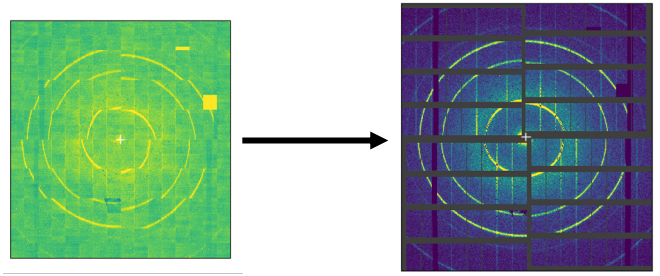
- Stream any data via open protocol, client library available in Python, C++
- Allows integration of externally developed tools and used for in-house developed analysis GUIs

```
>>> from karabo_bridge import Client
>>> krb_client = Client('tcp://server-host-name:12345')
>>> data, metadata = krb_client.next()
>>> data.keys()
dict_keys(['source1', 'source2', 'source3'])
>>> data['source1'].keys()
dict_keys(['param1', 'param2'])
>>> metadata['source1']
{'source1': {'source': 'source1',
  'timestamp': 1528476983.744877,
  'timestamp.frac': '744877000000000000',
  'timestamp.sec': '1528476983',
  'timestamp.tid': 10000000073}}
```

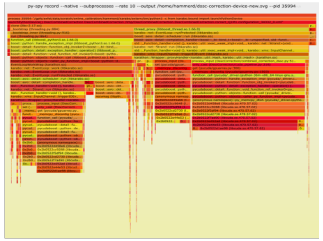


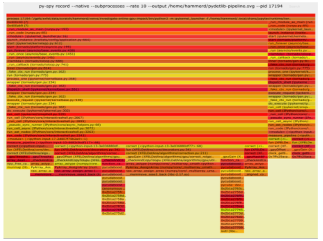
Detector corrections in real time

- For monitoring detector health and serious online analysis, detector data needs to be corrected live
- Previously only possible for a fraction of data
  - up to 1 Hz of trains with all pulses for *fast* detectors, even less streamed to users for analysis
  - closer to 10 Hz for *slow* detectors like JF or pnCCD
- Bottleneck is less computationally and most more I/O, moving around up to 21 GiB/s is a challenge on general purpose hardware



Scale online corrections to MHz

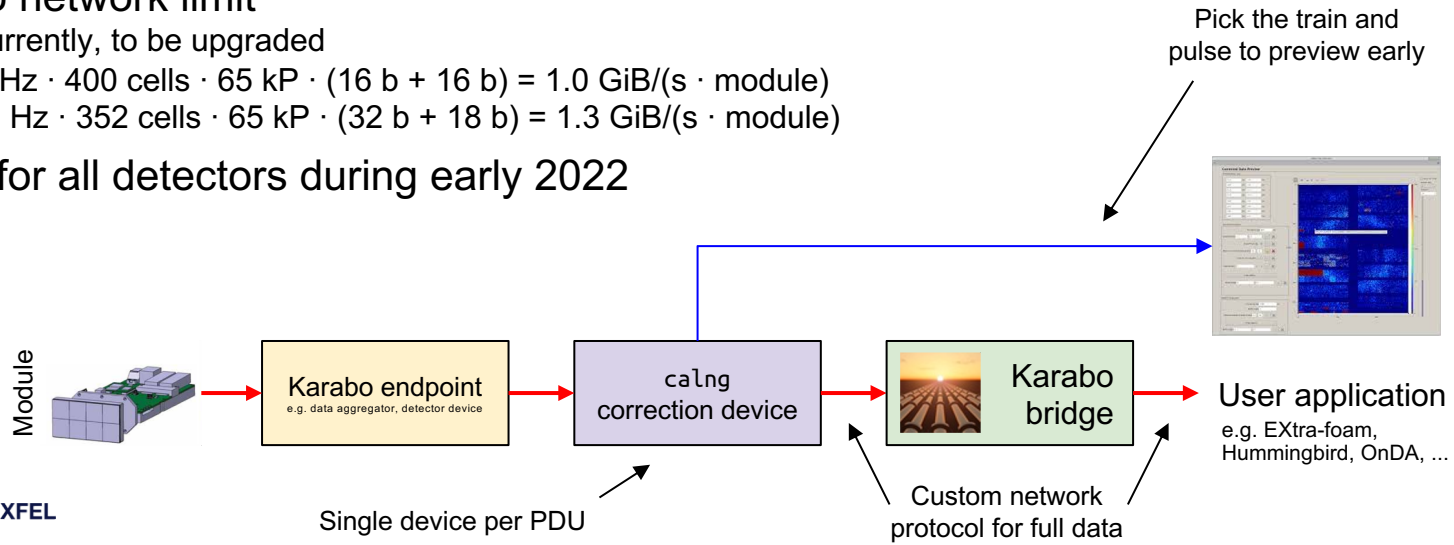




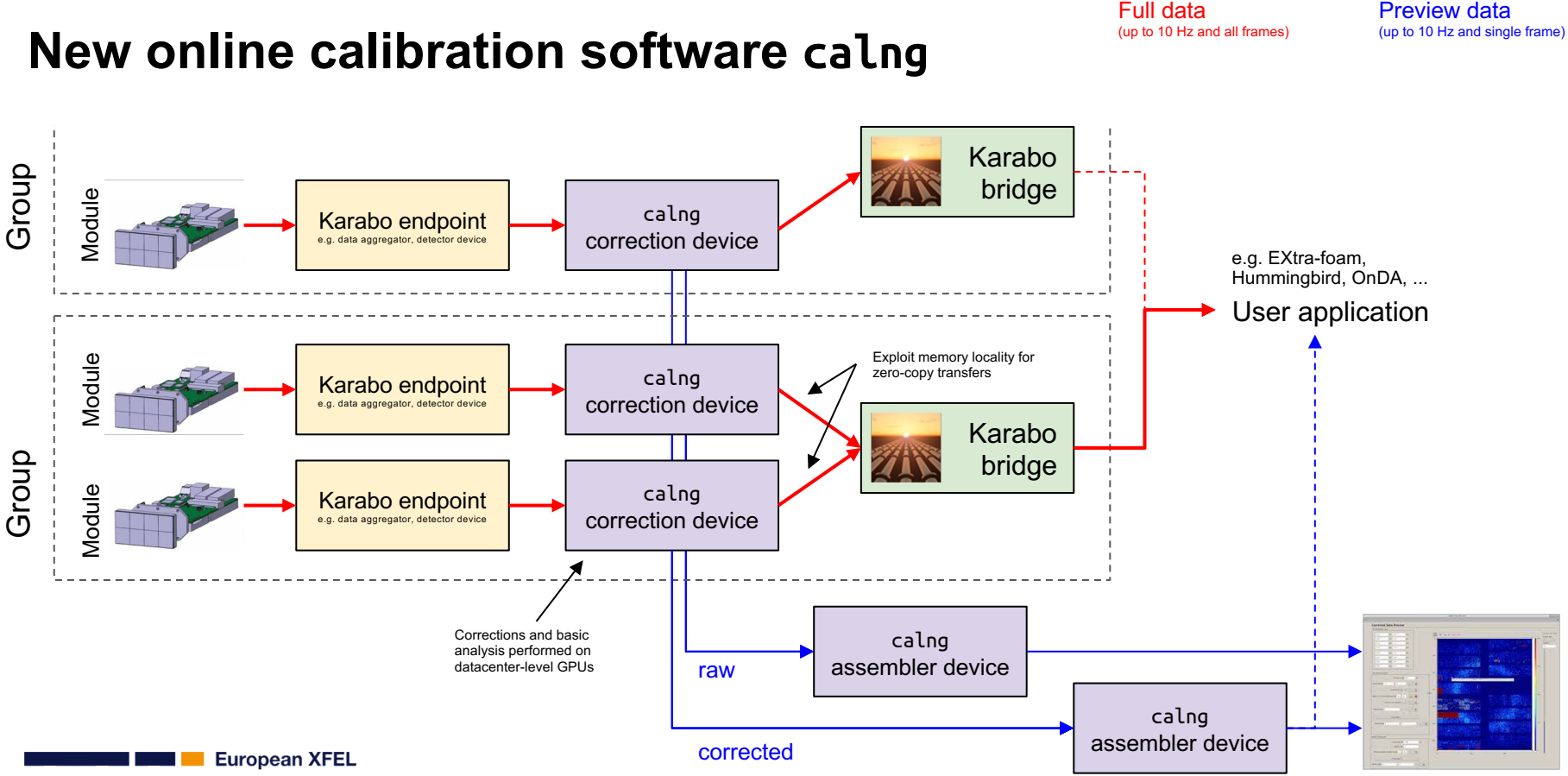
- Investigation into performance limits beginning of this year
  - With our current knowledge of how to correct an AGIPD and a clean state, what can be achieved?
- Promising results after a few weeks of profiling and prototyping:
  - Standalone, correcting data at up to 40 Hz is possible
  - Using GPUs for fast detectors is sound but requires special attention
- Scaffolded prototype for early field test in collaboration with users & instrument
  - SQS-DSSC in September
    - Stable operation of 10 Hz and 800 pulses
    - User application able to ingest all pulses of inner modules for fast feedback
    - After detector move used regularly at SCS
  - SPB-AGIPD in November
    - Achieved ≥ 3 Hz and 352 pulses
    - Investigation into new performance issues ongoing

New online calibration software `calng`

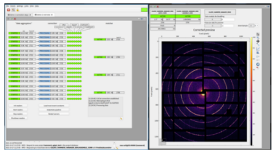
- Design goals from past operation experience
  - Reliable low latency karabo-gui preview
  - Enable online analysis with as much data as possible
  - Configurable and robust: Never run silently
- Scaling up to network limit
  - 7 GiB/s currently, to be upgraded
  - DSSC 10 Hz · 400 cells · 65 kP · (16 b + 16 b) = 1.0 GiB/(s · module)
    AGIPD 10 Hz · 352 cells · 65 kP · (32 b + 18 b) = 1.3 GiB/(s · module)
- Introduction for all detectors during early 2022



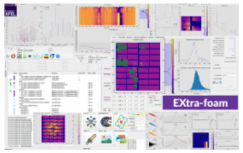
New online calibration software `calng`



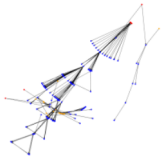
Online calibration and analysis

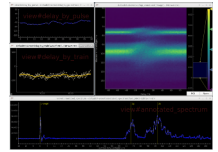


- Covering diverse requirements for online analysis with a set of niche solutions



- “Online analysis” baseline via Karabo control system
- Standalone GUI EXtra-foam for online analysis, covering lowest common denominator of experimental methods (Fast 2D imaging)
- Programmable analysis EXtra-metro for ultimate flexibility
- Streaming data via karabo-bridge to integrate community tools





- Progress on performing full detector calibration online for MHz detectors and delivering data to online analysis