# Control, data acquisition, management and analysis

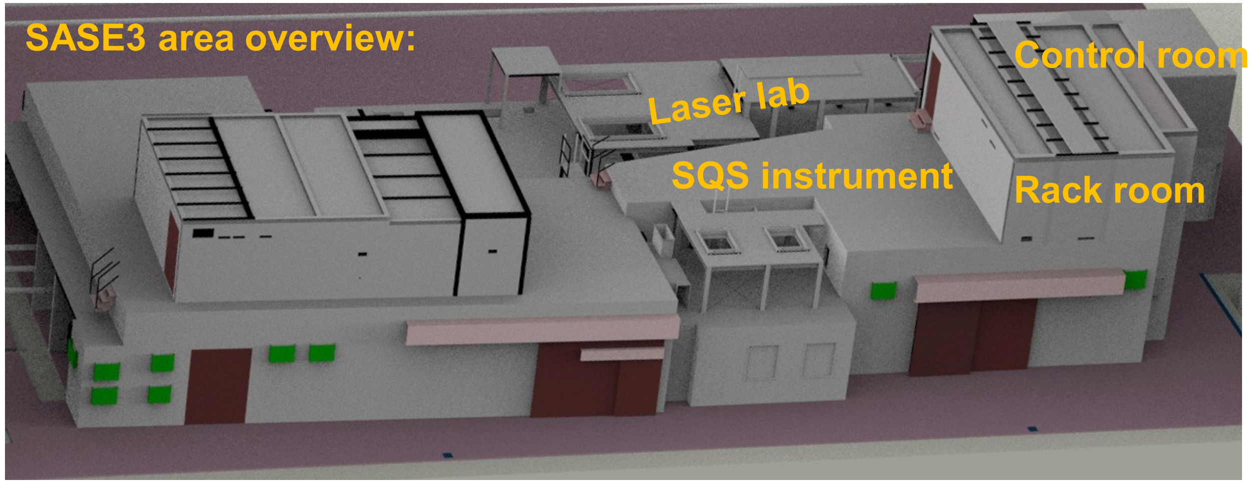
Thomas M. Baumann
Scientific Instrument SQS

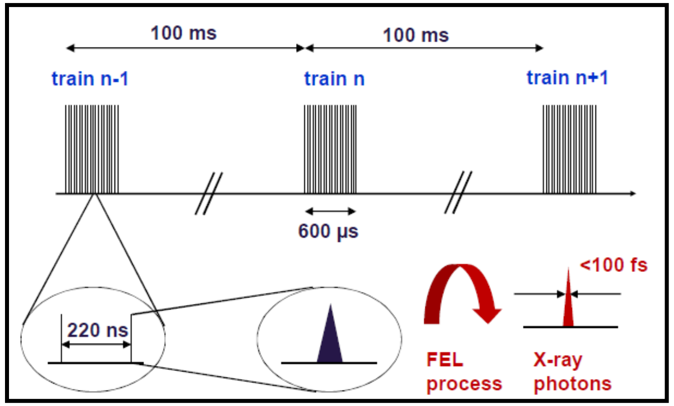
Hans Fangohr
Control & Analysis Software

SQS Early User Workshop
Schenefeld, 12.02.2018

# Outline:
# Experiment control challenges

- Remote control and monitoring of all hardware
- Large number of devices
- Fast data readout
  - 10 Hz train
  - 4.5 MHz pulses
- Large data volume
- Karabo software for control, DAQ, and analysis, developed by European XFEL




| Detector type | Sampling | Data/train | Data/sec |
|---|---|---|---|
| 1 channel digitizer | 10 GS/s | ~12 MB | ~120 MB |
| 4 Mpxl 2D camera | 10 Hz | ~8 MB | ~80 MB |
| 1 Mpxl 2D camera | 4.5 MHz | ~1 GB | ~10 GB |

# Hardware overview

**Beamline Scientists and Users**

**Hardware**
(Motors, pumps, valves,…)

**Electrodes**
(Charged particle optics)

**Detectors**
(MCP, Photodiode,…)

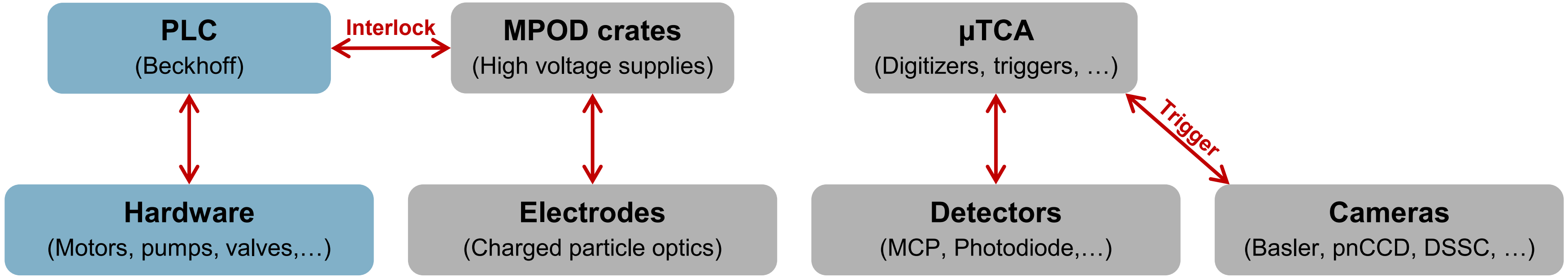
**Cameras**
(Basler, pnCCD, DSSC, …)

# Hardware overview

**Beamline Scientists and Users**

**PLC (Programmable Logic Controller)**

- Low level hardware control
- Interlock system
- Beckhoff and European XFEL






European XFEL Advanced Electronics Group (AE): P. Gessler, *et al.*

# Hardware overview

**Beamline Scientists and Users**

**PLC** (Beckhoff) ↔ Interlock ↔ **MPOD crates** (High voltage supplies)

**PLC** (Beckhoff) ↔ **Hardware** (Motors, pumps, valves,…)

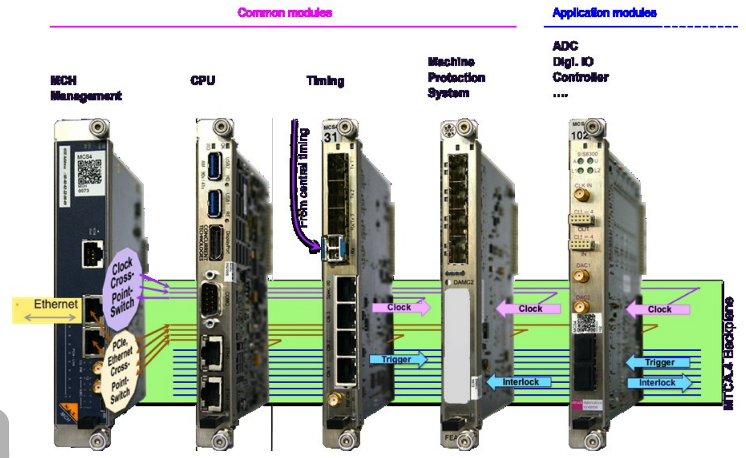
**MPOD crates** (High voltage supplies) ↔ **Electrodes** (Charged particle optics)

**µTCA** (Digitizers, triggers, …) ↔ **Detectors** (MCP, Photodiode,…)

**µTCA** (Digitizers, triggers, …) ↔ Trigger ↔ **Cameras** (Basler, pnCCD, DSSC, …)

## Micro TCA

- Digitizers
- Timing system: clocks and triggers
- Interlock (MPS, DESY)
- FPGA data processing



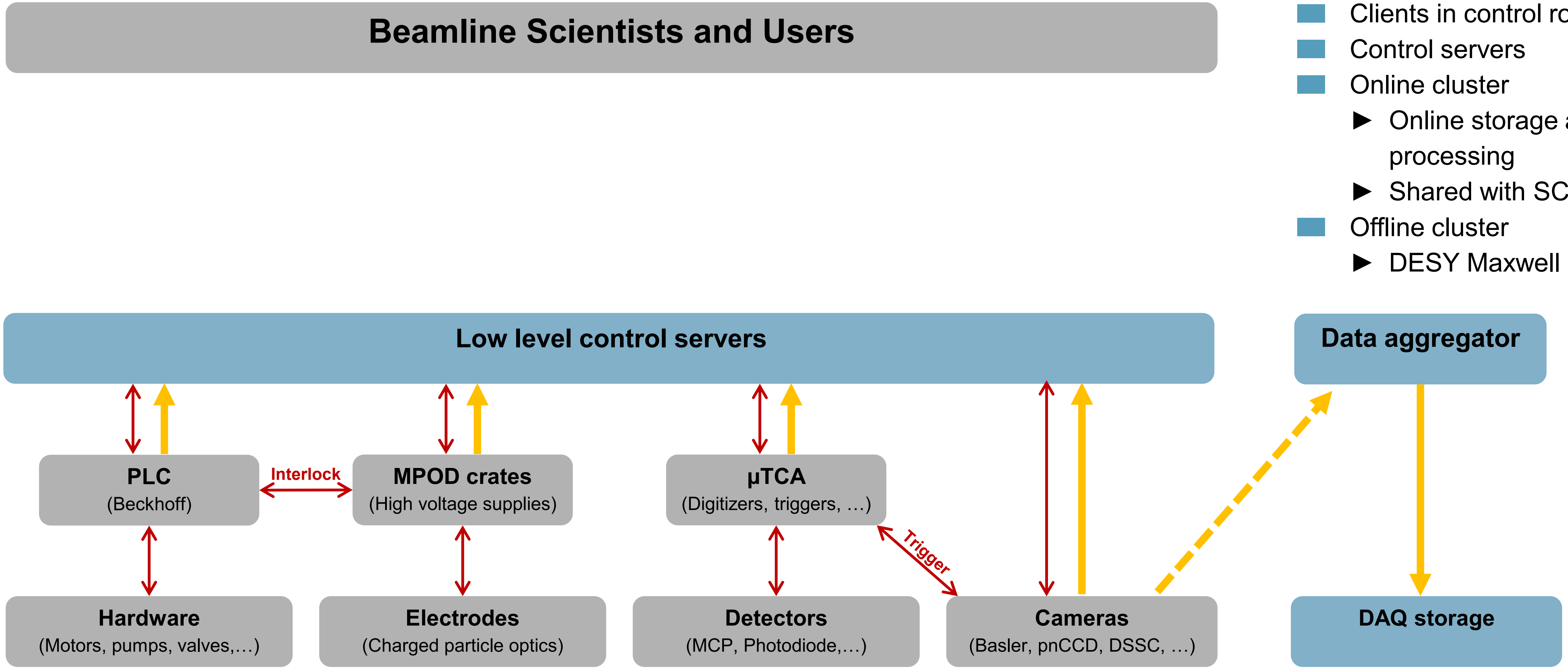



European XFEL Advanced Electronics Group (AE): P. Gessler, *et al.*

# Hardware overview

## Computing infrastructure

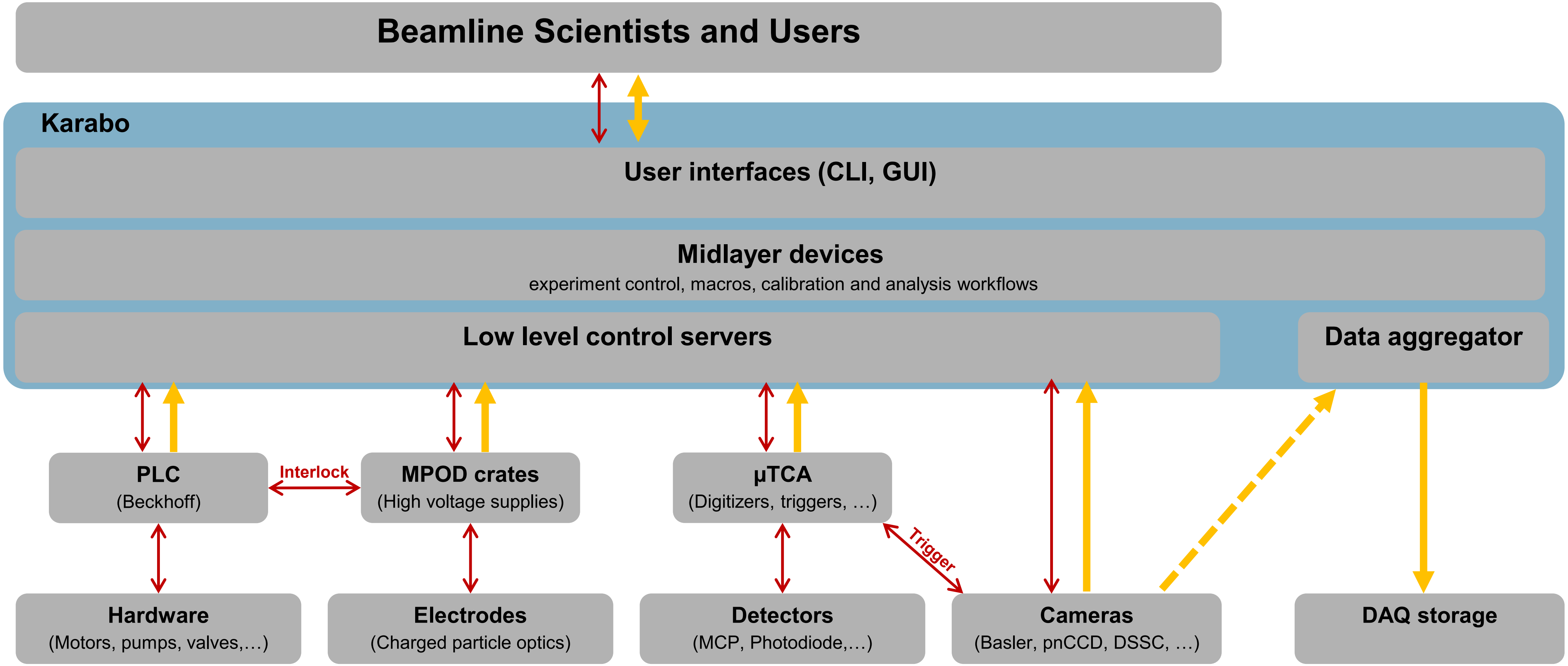
- Clients in control room
- Control servers
- Online cluster
  - Online storage and processing
  - Shared with SCS
- Offline cluster
  - DESY Maxwell cluster






European XFEL IT & Data Management (ITDM): K. Wrona, *et al.*

# Hardware overview





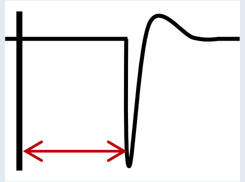

European XFEL Control & Analysis Software Group (CAS): S. Brockhauser, *et al.*

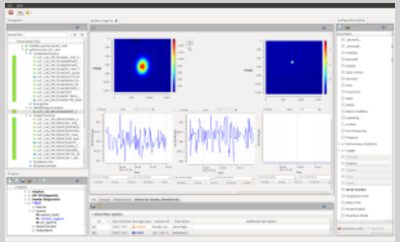
# Fast data acquisition system: Example of a digitizer

**Trigger**
- ▶ FEL timing
- ▶ Adjustable delays

## MicroTCA crate
- ▶ CPUs
- ▶ Control system
- ▶ Clock and trigger

## Computing infrastructure
- ▶ Online cluster and storage
- ▶ Karabo



**Detector**
- ▶ e.g. MCP
- ▶ Analog signal

**Digitizer**
- ▶ SPDevice
- ▶ 8 to 14 bit
- ▶ up to 3 GHz
- ▶ 1.8 to 10 GS/s

**Fast data processing**
- ▶ FPGA based
- ▶ Zero suppression
- ▶ Peak detection
- ▶ Pulse integration

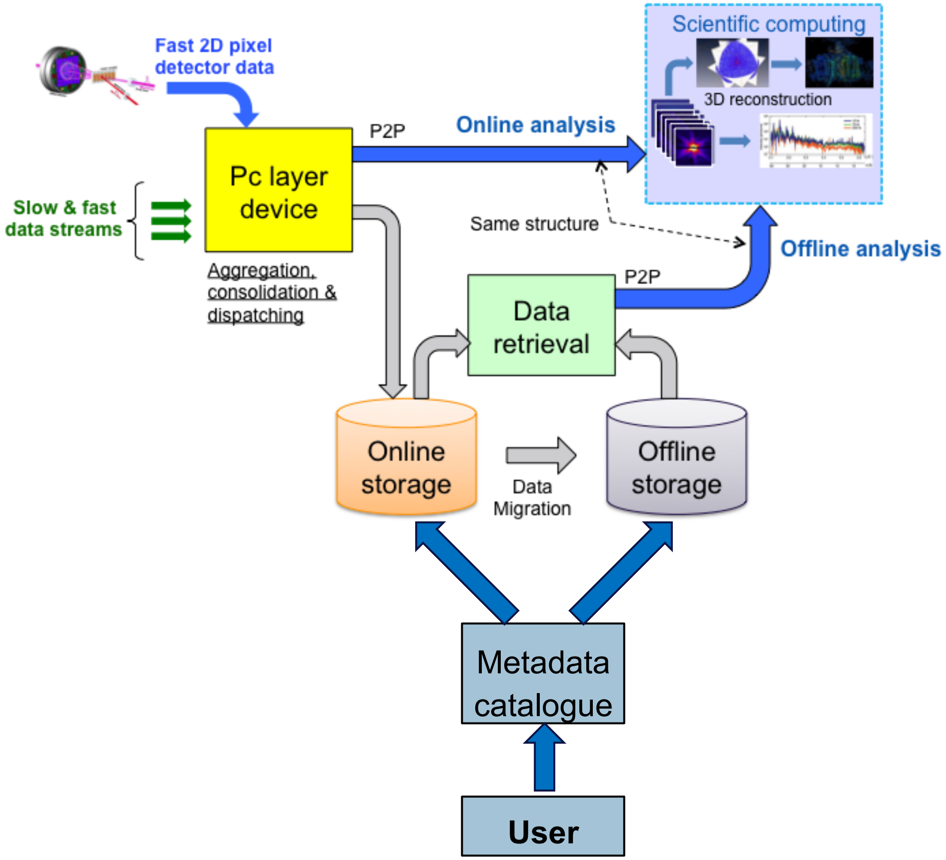
**Data Management**
- ▶ Store raw data
- ▶ Store control data
- ▶ Write metadata

**Pipeline process**
- ▶ Midlayer devices
- ▶ Data processing
- ▶ Online analysis
- ▶ User defined script

**Output**
- ▶ Life scenes
- ▶ Store processed data

# European XFEL data policy

- Raw data and metadata is stored, curated, and archived by European XFEL.
- Access to data through searchable metadata catalogue.
- Raw and metadata will be open access after embargo period.
- Electronic logbook is provided for documentation.
- Access to metadata catalogue and computing infrastructure through UPEX account.

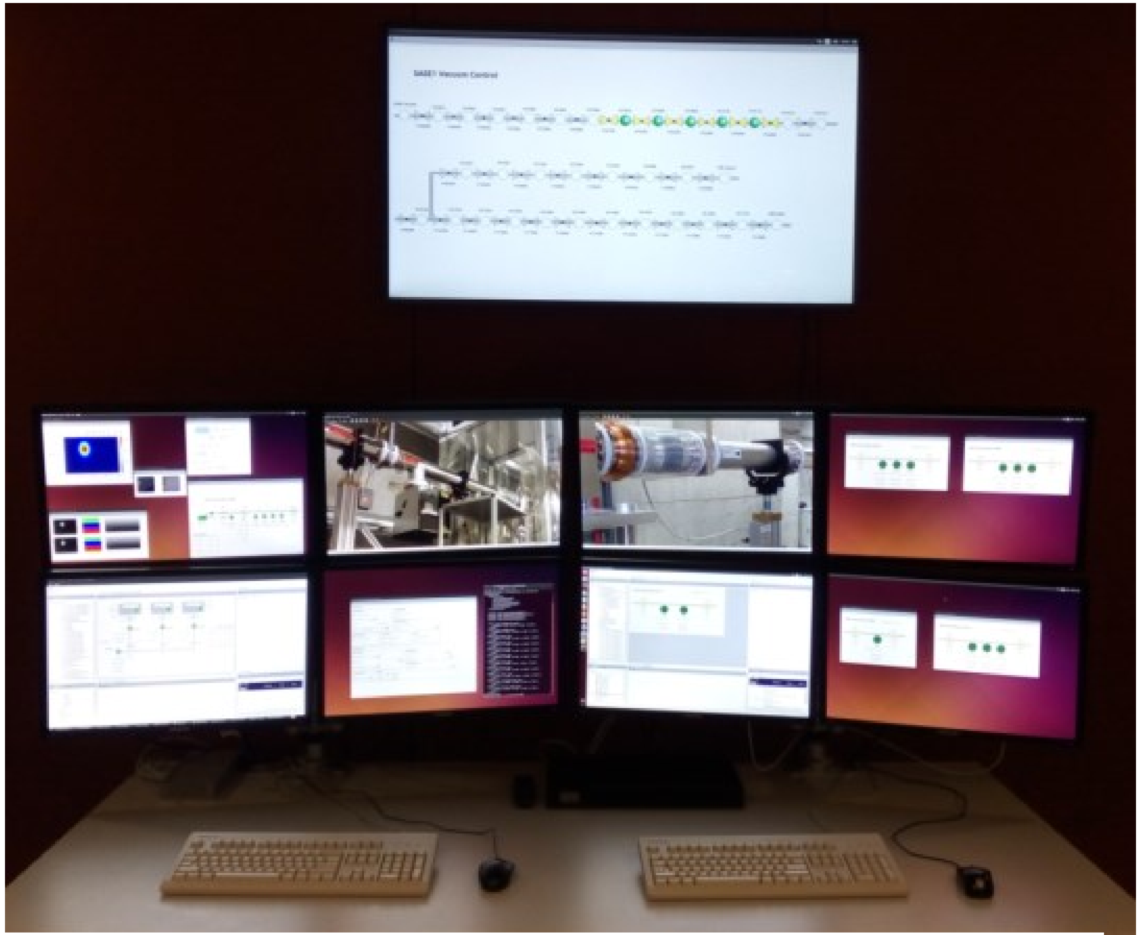



- All details see: www.xfel.eu/users/experiment_support/policies/scientific_data_policy/

# Karabo

(Hans Fangohr on behalf of Control and Analysis Software)

- Software [1] for
  - Hardware control
  - Data acquisition
  - Data management
  - Data analysis and scientific computing
- Multiple interfaces, including Python
- Developed by European XFEL
- User interface to experiment control and data

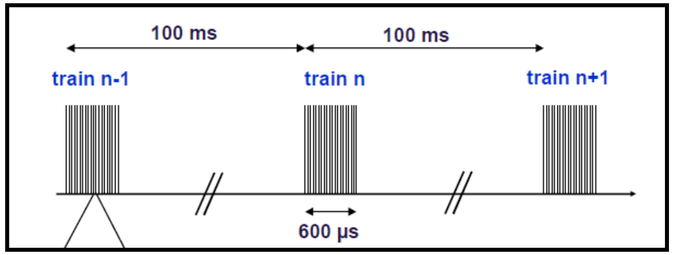


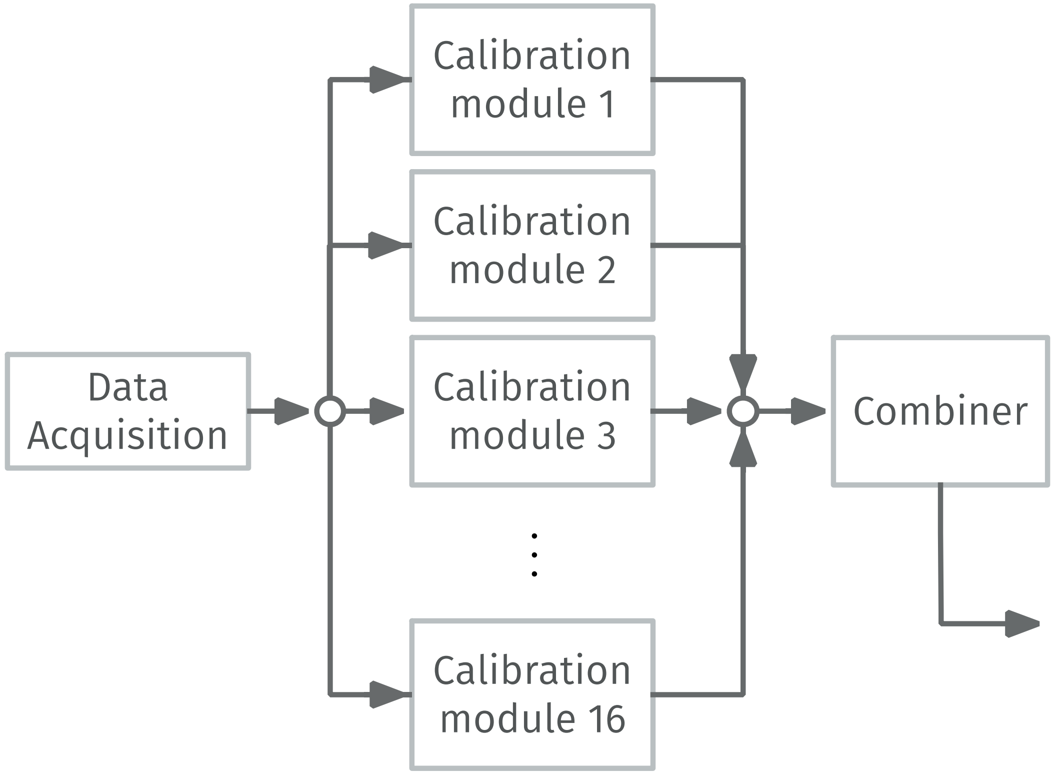
[1] B. Heisen et al: “Karabo: An integrated software framework combining control, data management, and scientific computing tasks,” in14th ICALEPCS2013. San Francisco, CA, 2013.

# Karabo data pipeline example

- Karabo is framework for control *and data*
  - Data tokens pass through pipeline
  - Processing units called “devices”
  - Devices can be distributed over hardware
  - Simplified example in figure: calibration for detector modules carried out in parallel

- Data tokens in pipelines contain data for one train at a time
  - Tagged with “train-id”
    - ► Within train: “pulse-id”

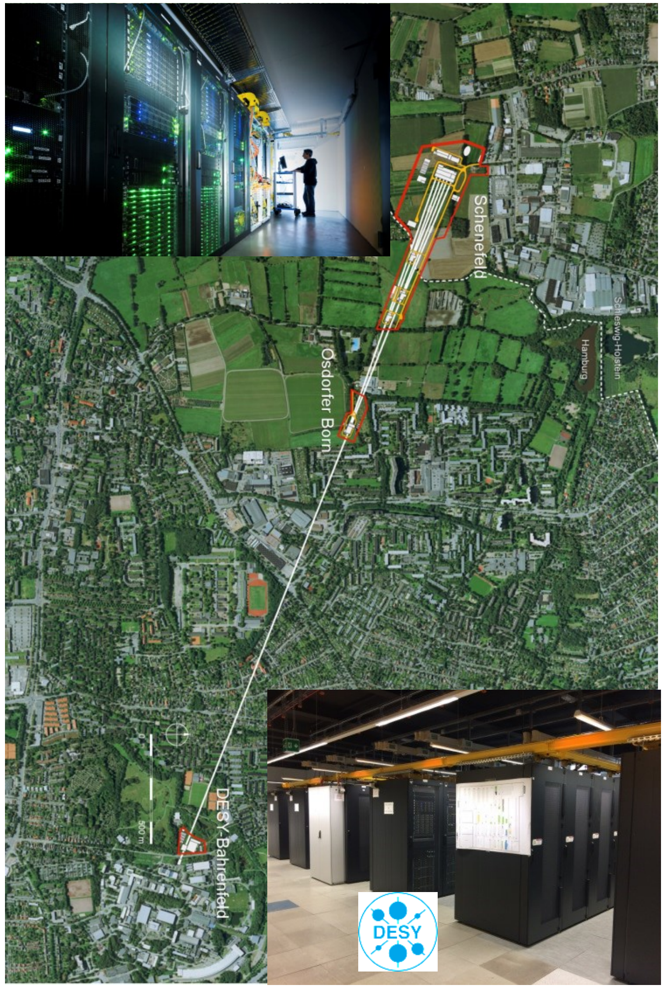


# Data analysis infrastructure

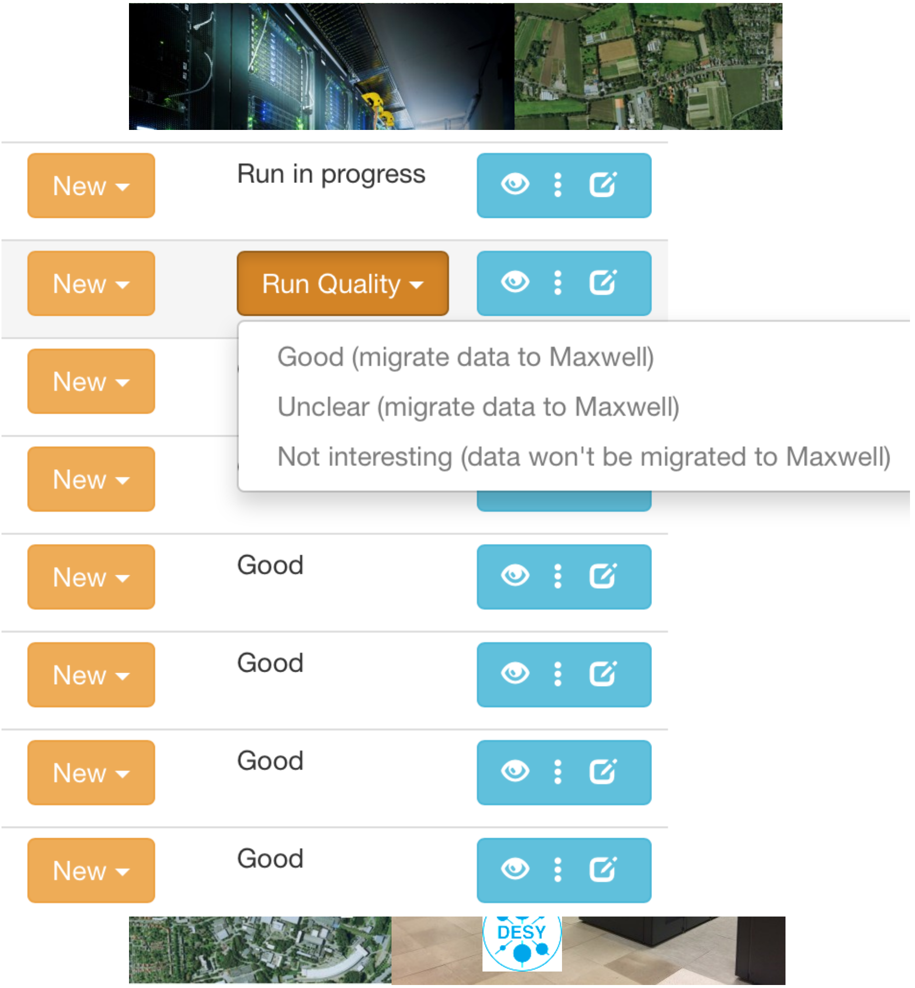
- Hardware: “Online cluster”,
  - 8 nodes x (20 cores, 256GB RAM) dedicated to users
  - Additional nodes for control and XFEL provided calibration and processing

- Hardware: “Offline cluster” = Maxwell cluster (DESY)
  - 80 nodes/3200 cores (Intel Xeon E5-2698v4)
  - ~112 TFlops
  - 512GB RAM each node
  - +20 nodes with other spec
  - 7 GPU nodes available

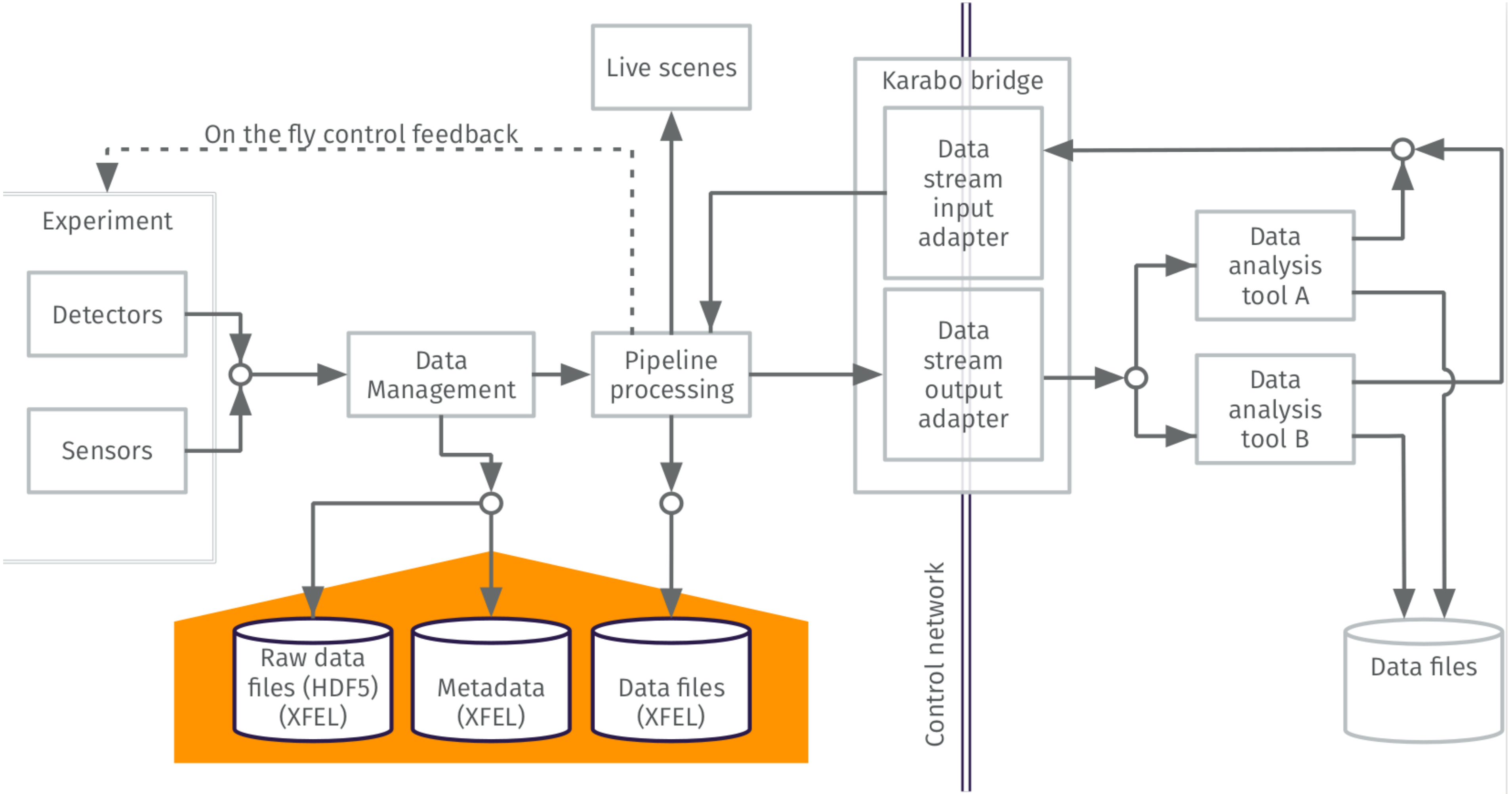


# Data management online -> offline

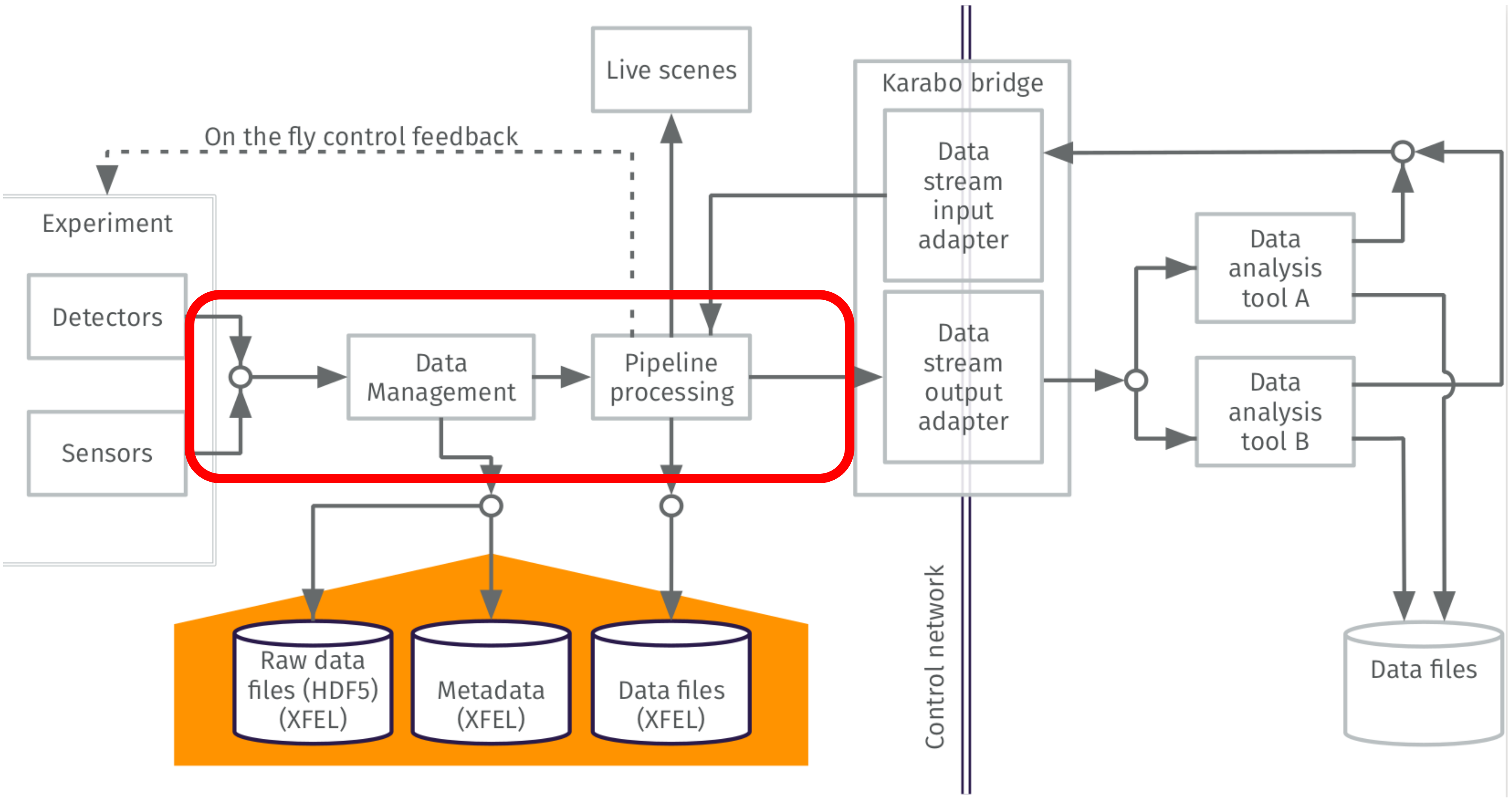
- During measurement (run)
  - Calibrated and raw data available in hutch (GUI, online)
- Data migration after each run
  - After each run, data manager decides on quality of the data: “good”, “unclear”, “not interesting”
  - “good” and “unclear” data transferred to “Offline cluster”
  - Migration triggers computation of calibrated data at online cluster
- After experiment
  - Raw and calibrated data available
  - Analysis on ”Offline cluster” (Maxwell @ DESY)



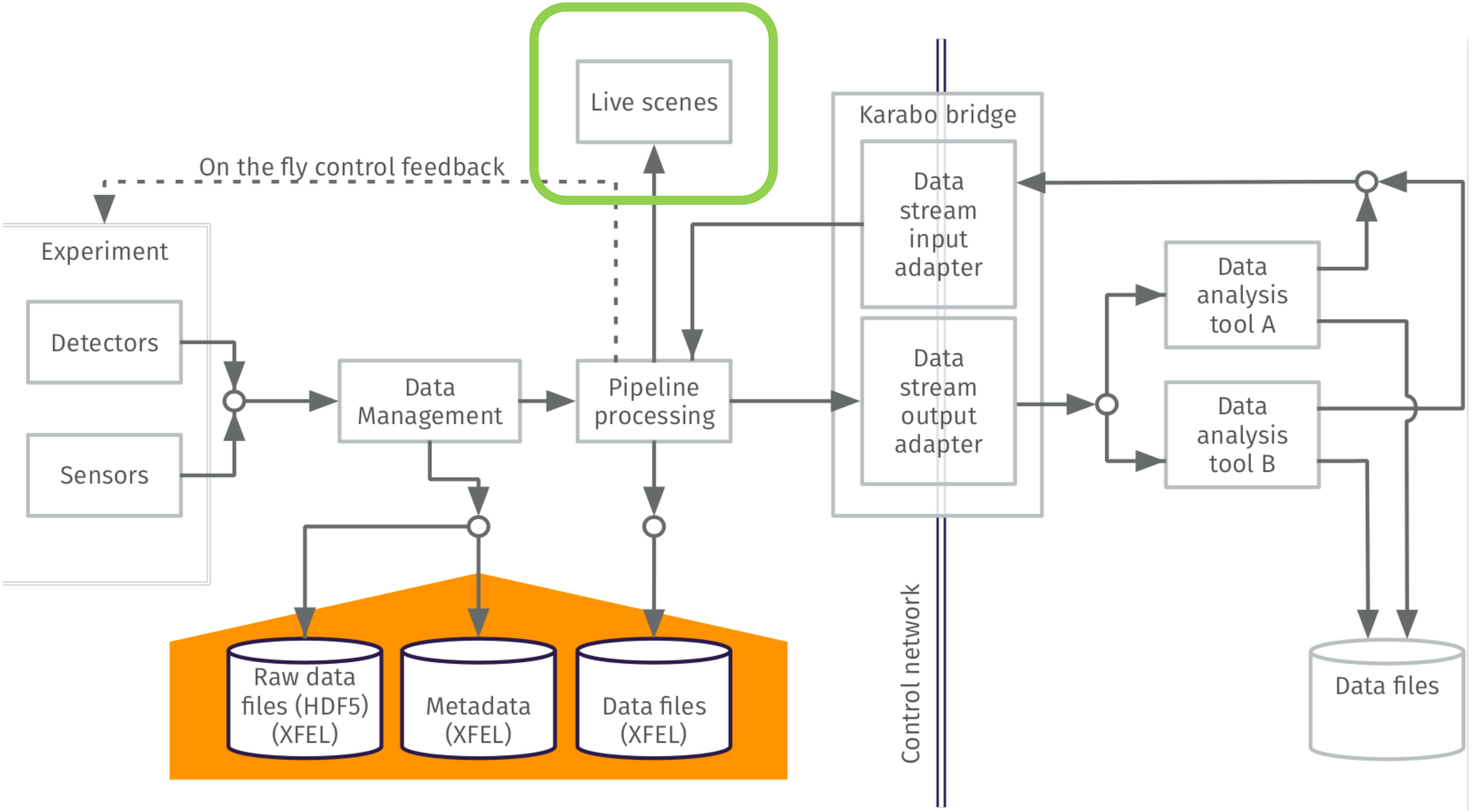
# Online data flow and data analysis



# Karabo data pipeline

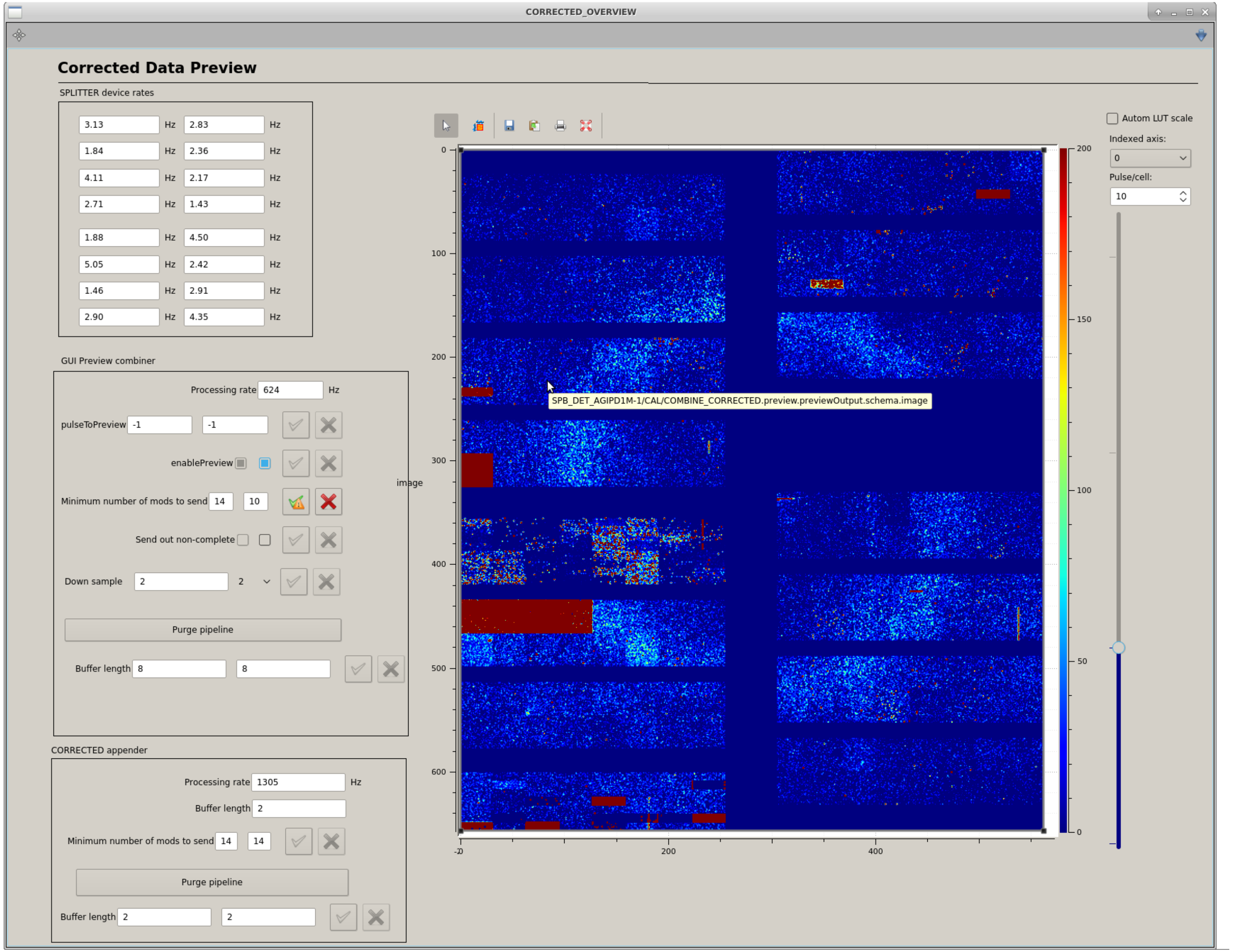


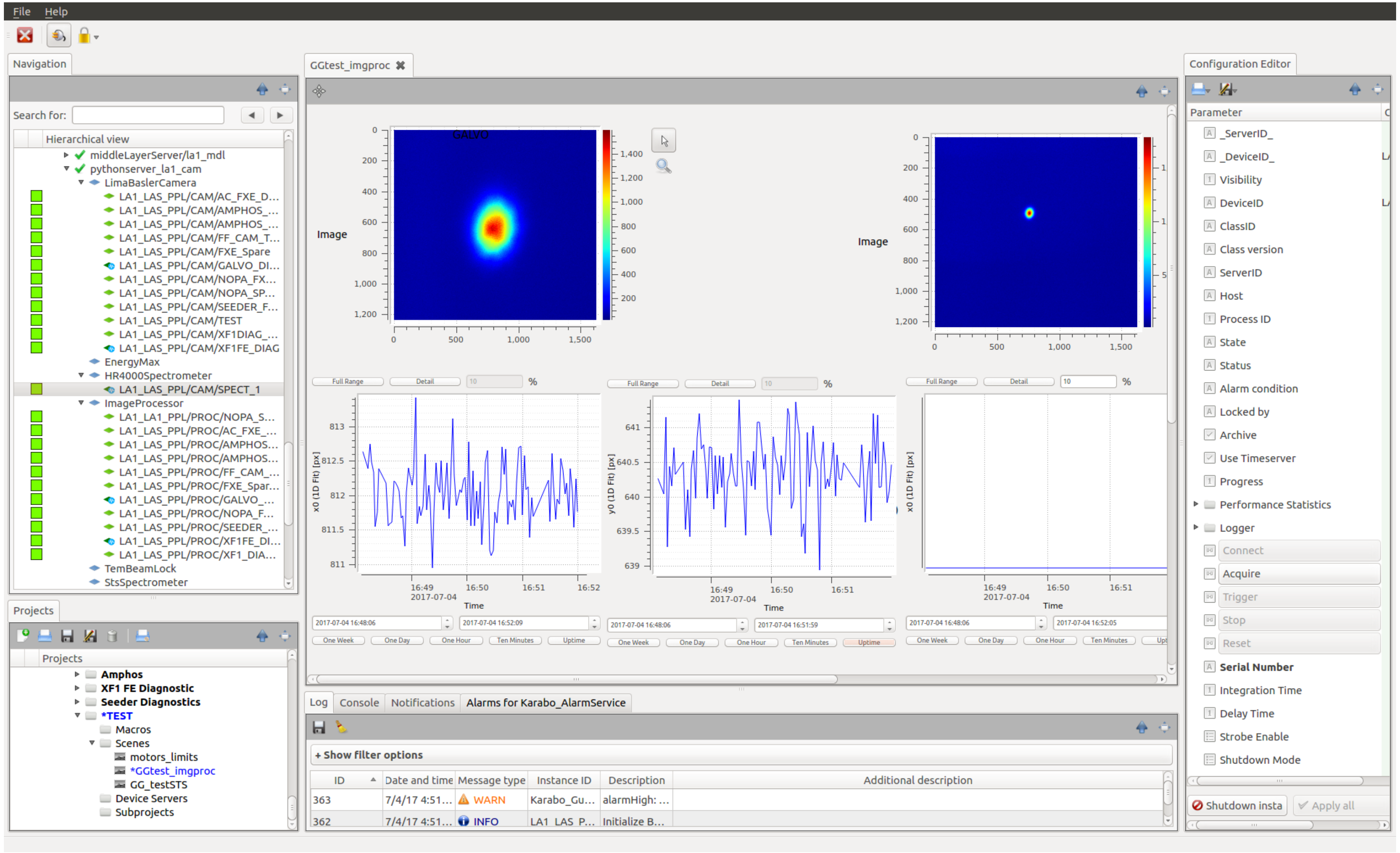
# GUI ‘scenes’


Live scenes
On the fly control feedback
Experiment
Detectors
Sensors
Data Management
Pipeline processing
Karabo bridge
Data stream input adapter
Data stream output adapter
Data analysis tool A
Data analysis tool B
Raw data files (HDF5) (XFEL)
Metadata (XFEL)
Data files (XFEL)
Control network
Data files

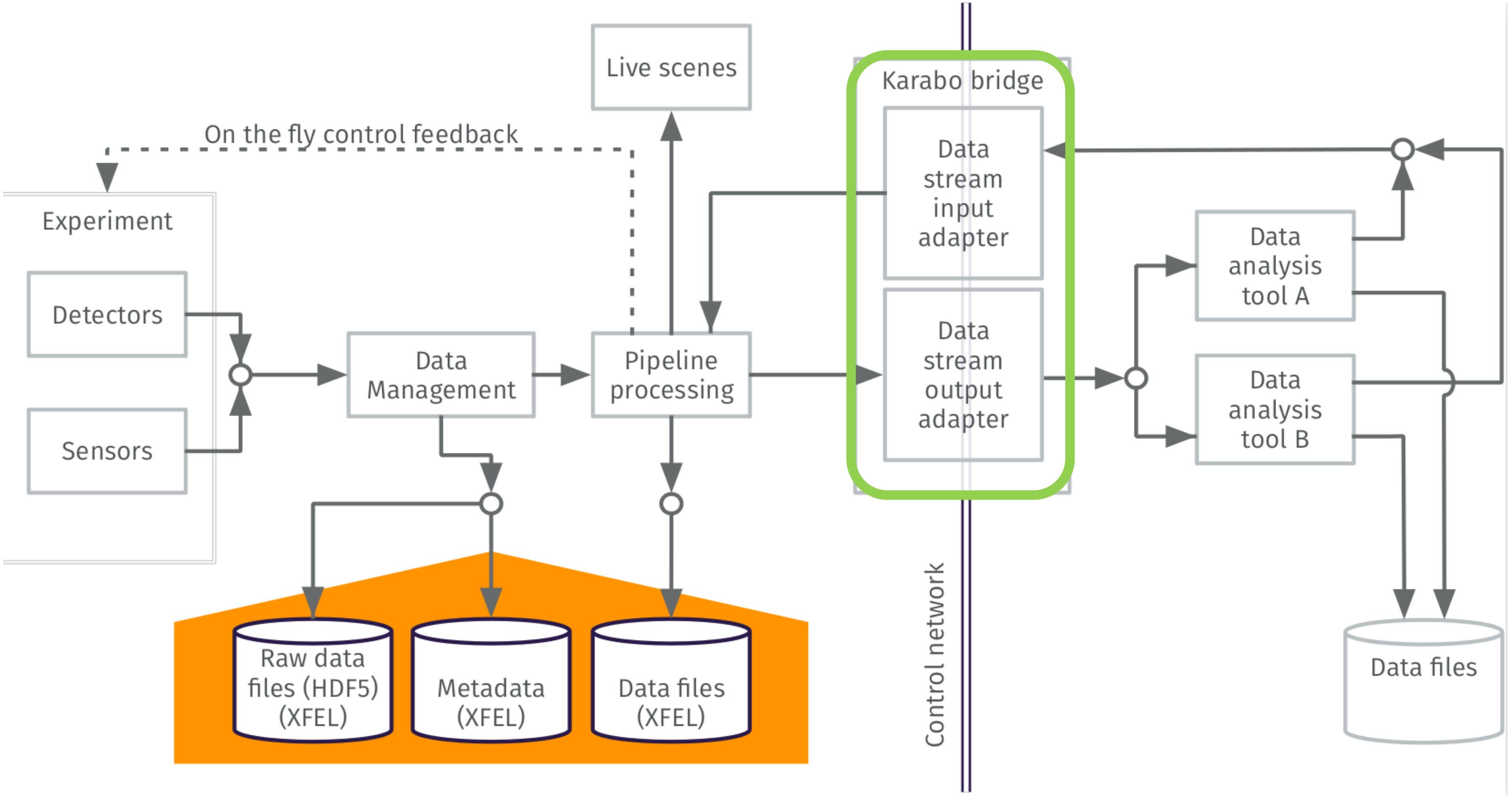
# Online data analysis:

# Rapid feedback through GUI




File
Help
Navigation
Search for:
Hierarchical view
middleLayerServer/la1_mdl
pythonserver_la1_cam
LimaBaslerCamera
LA1_LAS_PPL/CAM/AC_FXE_D...
LA1_LAS_PPL/CAM/AMPHOS_...
LA1_LAS_PPL/CAM/AMPHOS_...
LA1_LAS_PPL/CAM/FF_CAM_T...
LA1_LAS_PPL/CAM/FXE_Spare
LA1_LAS_PPL/CAM/GALVO_DI...
LA1_LAS_PPL/CAM/NOPA_FX...
LA1_LAS_PPL/CAM/NOPA_SP...
LA1_LAS_PPL/CAM/SEEDER_F...
LA1_LAS_PPL/CAM/TEST
LA1_LAS_PPL/CAM/XF1DIAG_...
LA1_LAS_PPL/CAM/XF1FE_DIAG
EnergyMax
HR4000Spectrometer
LA1_LAS_PPL/CAM/SPECT_1
ImageProcessor
LA1_LA1_PPL/PROC/NOPA_S...
LA1_LAS_PPL/PROC/AC_FXE_...
LA1_LAS_PPL/PROC/AMPHOS...
LA1_LAS_PPL/PROC/AMPHOS...
LA1_LAS_PPL/PROC/FF_CAM_...
LA1_LAS_PPL/PROC/FXE_Spar...
LA1_LAS_PPL/PROC/GALVO_...
LA1_LAS_PPL/PROC/NOPA_F...
LA1_LAS_PPL/PROC/SEEDER_...
LA1_LAS_PPL/PROC/XF1FE_DI...
LA1_LAS_PPL/PROC/XF1_DIA...
TemBeamLock
StsSpectrometer
Projects
Projects
Amphos
XF1 FE Diagnostic
Seeder Diagnostics
*TEST
Macros
Scenes
motors_limits
*GGtest_imgproc
GG_testSTS
Device Servers
Subprojects
GGtest_imgproc
Image
Image
Full Range
Detail
%
x0 (1D Fit) [px]
y0 (1D Fit) [px]
Time
2017-07-04
One Week
One Day
One Hour
Ten Minutes
Uptime
Log
Console
Notifications
Alarms for Karabo_AlarmService
+ Show filter options
ID
Date and time
Message type
Instance ID
Description
Additional description
363
7/4/17 4:51...
WARN
Karabo_Gu...
alarmHigh: ...
362
7/4/17 4:51...
INFO
LA1_LAS_P...
Initialize B...
Configuration Editor
Parameter
_ServerID_
_DeviceID_
Visibility
DeviceID
ClassID
Class version
ServerID
Host
Process ID
State
Status
Alarm condition
Locked by
Archive
Use Timeserver
Progress
Performance Statistics
Logger
Connect
Acquire
Trigger
Stop
Reset
Serial Number
Integration Time
Delay Time
Strobe Enable
Shutdown Mode
Shutdown insta
Apply all

# Export data pipeline – Karabo bridge

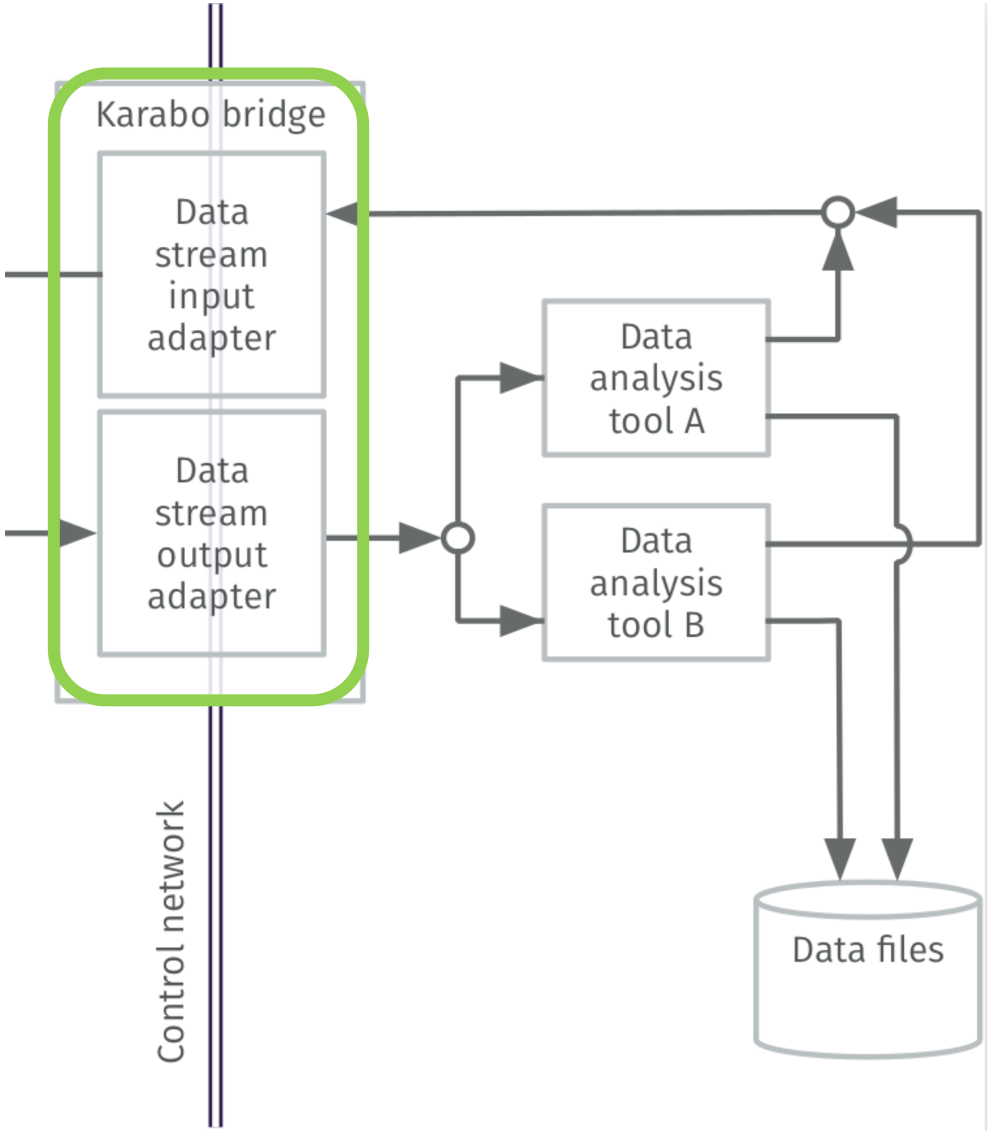


# Export data pipeline – Karabo bridge

- We provide an interface to listen to Karabo pipelines
  - Allows to **integrate** existing (complex) user provided tools
  - Quick (dirty) **specific** scripts to use during an experiment

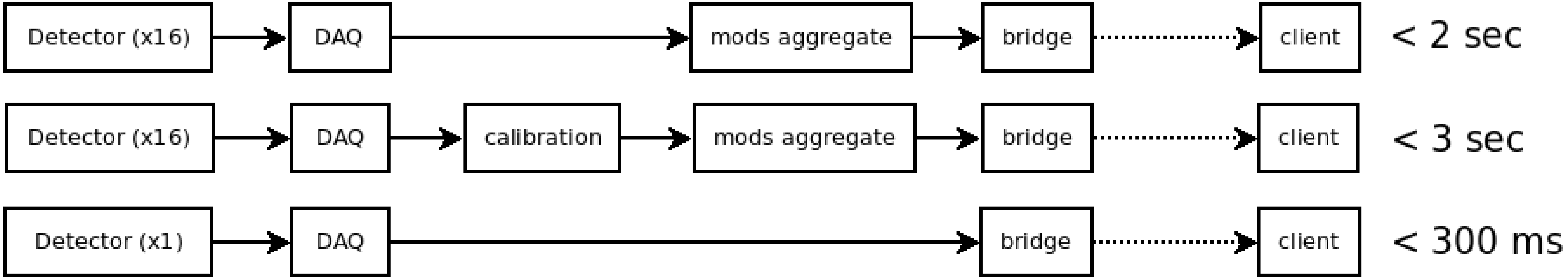
- Karabo Bridge requirements
  - Low latency
  - Loosely coupled **Interface** between Karabo and external programs
  - Export data in a **generic** container
  - Using **straightforward** network interface

- Development in collaboration with CFEL Chapman Group (S. Aplin, A. Barty, M. Kuhn, V. Mariani)



# Export data pipeline – Karabo bridge

- Client implementation and simulator
  - Python: https://github.com/European-XFEL/karabo-bridge-py
  - C++: implementation existing
- Successful use during first early user experiments
  - OnDA, Hummingbird, CASS, custom
- Latency (SPB experiment, AGIPD detector)



# Offline data treatment and analysis

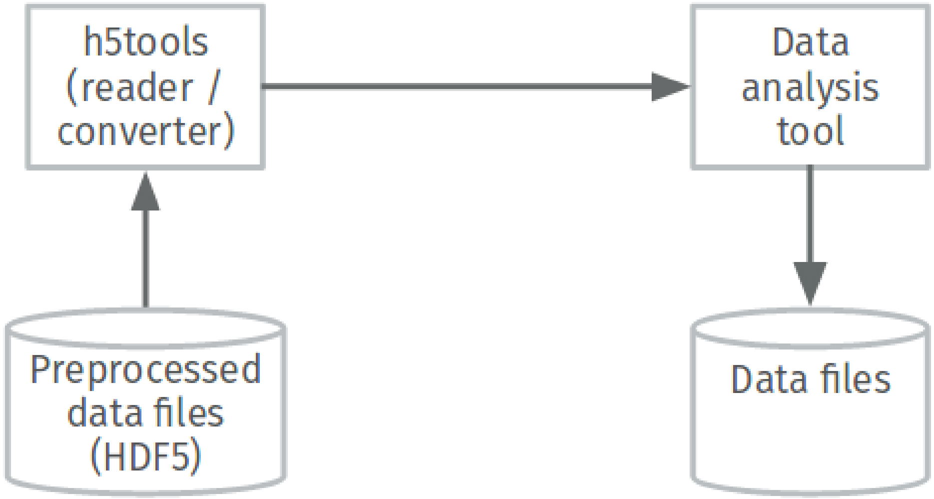
- Offline means after beamtime
- But can also be used 'during' experiments (after data has been migrated)
- DESY's Maxwell cluster
  - Reserved partition ("upex") for XFEL users
  - Accessible a few weeks ahead of beamtime
  - During, and after beamtime.
- HDF5 Files

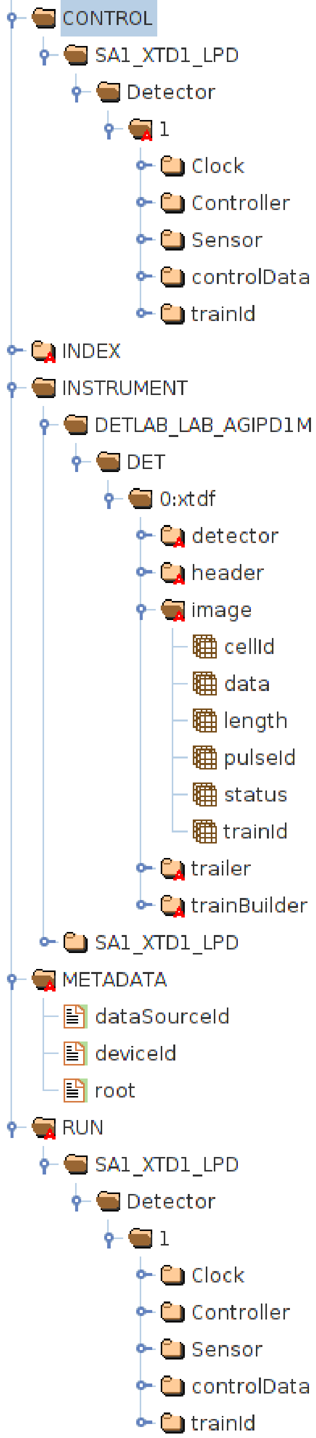
```
[haufs@max-exfl014]/gpfs/exfel/exp/SPB/201701/p002038/proc% ls r0039
CORR-R0039-AGIPD00-S00000.h5  CORR-R0039-AGIPD05-S00002.h5  CORR-R0039
CORR-R0039-AGIPD00-S00001.h5  CORR-R0039-AGIPD05-S00003.h5  CORR-R0039
CORR-R0039-AGIPD00-S00002.h5  CORR-R0039-AGIPD06-S00000.h5  CORR-R0039
CORR-R0039-AGIPD00-S00003.h5  CORR-R0039-AGIPD06-S00001.h5  CORR-R0039
CORR-R0039-AGIPD01-S00000.h5  CORR-R0039-AGIPD06-S00002.h5  CORR-R0039
CORR-R0039-AGIPD01-S00001.h5  CORR-R0039-AGIPD06-S00003.h5  CORR-R0039
CORR-R0039-AGIPD01-S00002.h5  CORR-R0039-AGIPD07-S00000.h5  CORR-R0039
CORR-R0039-AGIPD01-S00003.h5  CORR-R0039-AGIPD07-S00001.h5  CORR-R0039
```

# Data Files

- For each experiment **proposal**
  - Data are stored in **run** folders
  - **Runs** contain a collection of **HDF5** files
  - HDF5 files are structured in EuXFEL specific format

- Euxfel-h5tools
  - Command line mode: quick overview
  - Library to read run data more conveniently
  - https://github.com/European-XFEL/h5tools-py



# Summary

- Outlined data acquisition, management, and analysis
  - Data policy
  - Online and offline analysis
- Early stages of the facility
  - Outlined current state and plans
- Keen to work with users
  - Get in touch
  - Help us prioritise resources

# Thank you for your attention!

## SQS Instrument

**M. Meyer**
A. Achner
T.M. Baumann
R. Boll
A. De Fanis
S. Deinert
P. Grychtol
M. Ilchen
T. Mazza
J. Montaño
Y. Ovcharenko
N. Rennhack
R. Wagner
P. Ziołkowski

## Advanced Electronics

**P. Gessler**
A. Ahmed
F. Babies
K.-E. Ballak
B. Baranašić
N. Coppola
J. Eilers
D. Emes
B. Fernandes
T. Freyermuth
S. Huynh
N. Jardon
M. U. Jidda
M. Meyer
H. Sotoudi Namin
J. Tolkiehn
J. Zach

## IT & Data Management

**K. Wrona**
N. Al-Qudami
D. Boukhelef
I. Derevianko
J. Elizondo
K. Filippakopoulos
M. Knaak
S. D. Kujala
L. Maia
M. Manetti
R. Nazari
B. Poljancewicz
G. Previtali
E. Stoica
J. Szuba

## Control & Analysis Software

**S. Brockhauser**
M. Beg
V. Bondar
C. Danilevski
Y. Kirienko
W. Ehsan
S. Esenov
H. Fangohr
G. Flucke
G. Giovanetti
D. Göries
A. Klimovskaia
T. Kluyver
L. Mekinda
T. Michelat
A. Parenti
H. Santos
A. Silenzi
M. Teichmann
C. Xu
C. Youngman

## References

- Data Analysis focused documentation: http://bit.ly/xfel-da-docs
- European XFEL data policy: www.xfel.eu/users/experiment_support/policies/scientific_data_policy
- H. Fangohr et al, Data Analysis support in Karabo at European XFEL, ICALEPSC 2017: http://icalepcs2017.vrws.de/papers/tucpa01.pdf