PAUL SCHERRER INSTITUT

Filip Leonarski :: Postdoc :: Paul Scherrer Institut

Coping with increased data rates from the JUNGFRAU detector

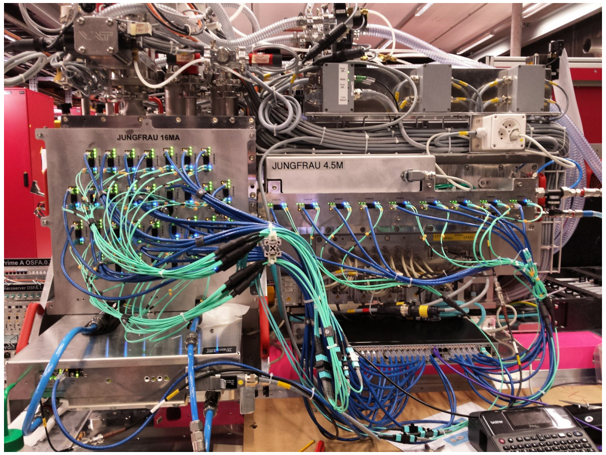
2019 HDR MX Meeting, Didcot, UK

Why JUNGFRAU?

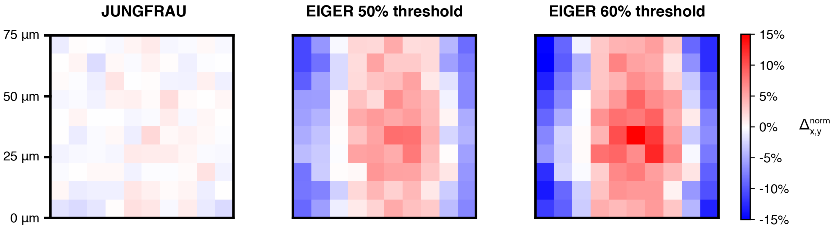
- Count rate limitation makes photon counters not suitable for XFELs → new interest in integrating detectors

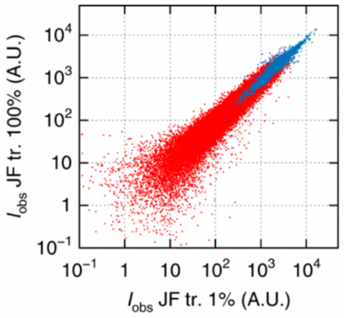
- Upgrade to diffraction limited storage rings and usage of multilayer monochromators will result in problems with count rate at synchrotrons

- Photon counters operate with reduced performance at low energy (≤ 6 keV)



JF 16M and JF 4.5M at SwissFEL operating up to 100 Hz





F. Leonarski, S. Redford, A. Mozzanica, … , M. Wang *Nat. Methods*, **15**, 799-804 (2018)

JUNGFRAU: Charge Integrating Detector

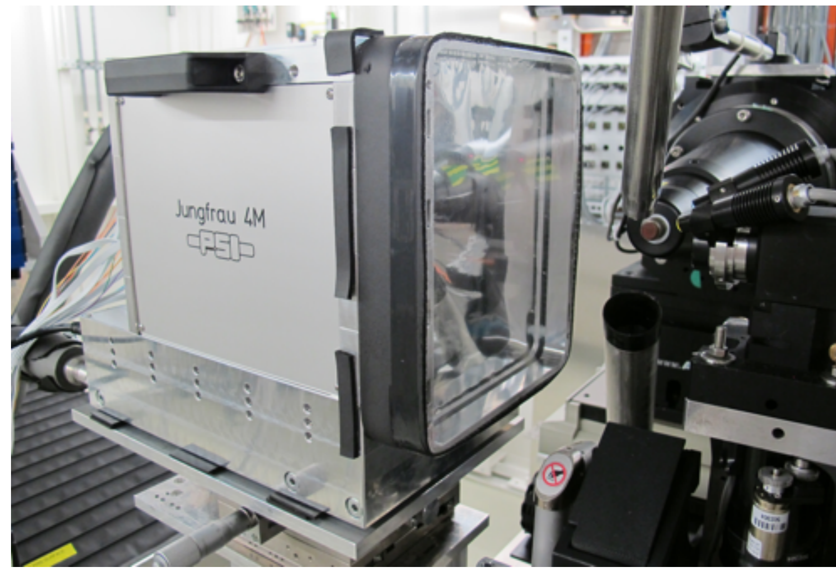
- Developed at Paul Scherrer Institute

- 75 µm x 75 µm pixel Si sensor
- Max framerate 2.2 kHz

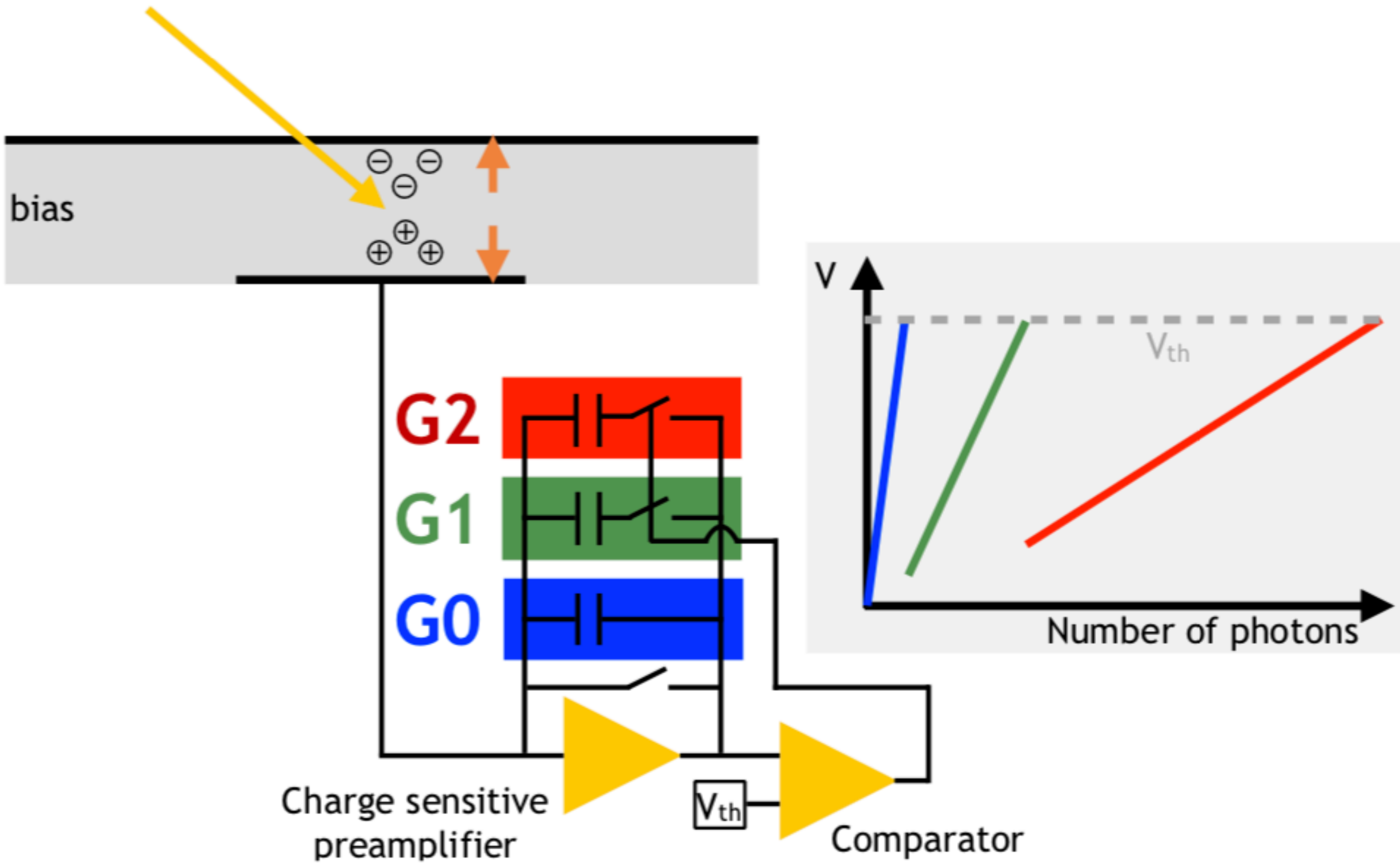
- In operation at SwissFEL

- Integrating technology has potential:
  - High throughput crystallography at diffraction limited storage rings
  - Time resolved with multilayer monochromators
  - Low energy native-SAD

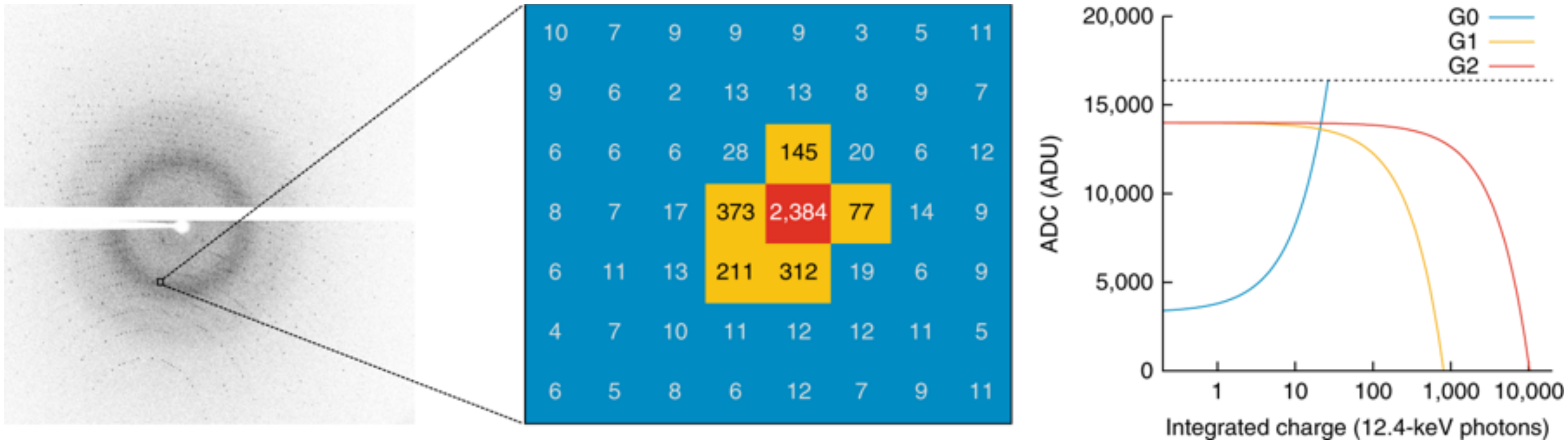
- Planned for X06SA and/or X06DA beamlines at SLS



JUNGFRAU: Adaptive Gain

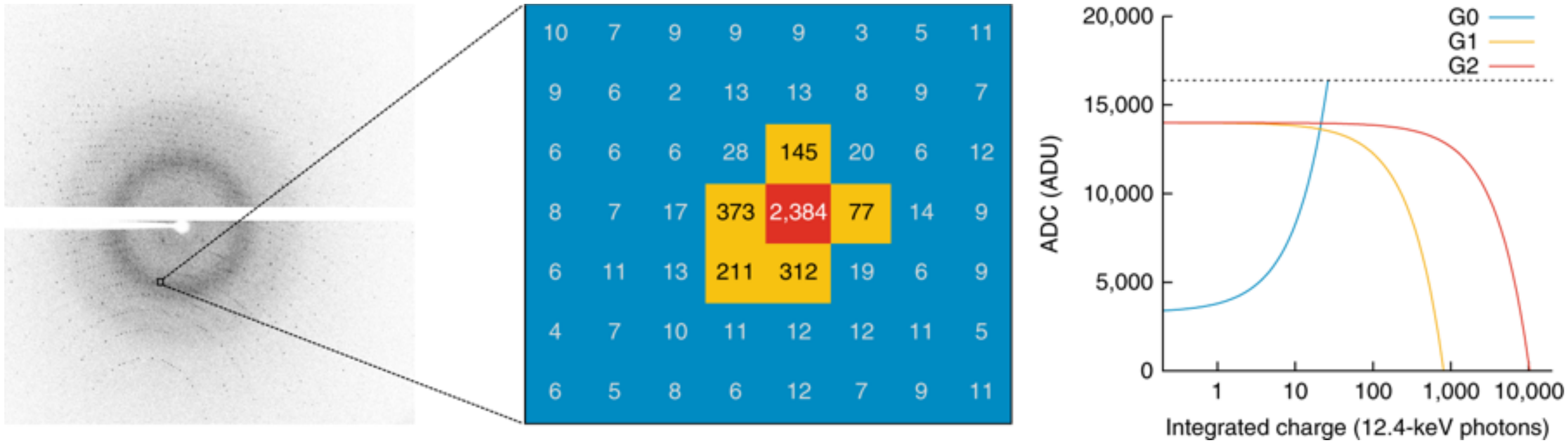


JUNGFRAU: Adaptive Gain



F. Leonarski, S. Redford, A. Mozzanica, … , M. Wang *Nat. Methods*, **15**, 799-804 (2018)

JUNGFRAU: Pixel Readout



F. Leonarski, S. Redford, A. Mozzanica, … , M. Wang *Nat. Methods*, **15**, 799-804 (2018)

Pixel output in JF: 0001010111110011

Gain: 0001010111110011 00:G0 01:G1 11:G2

ADC value: 0001010111110011

Dedicated dark run

Photon number: $= \frac{ADC - pedestal}{gain * photon\ energy}$

Gain and pedestal factors are specific for pixel and gain setting

Prior calibration

Considerations of JUNGFRAU Output

Pixel output in JF:	0001010111110011
Gain:	0001010111110011 00:G0 01:G1 11:G2
ADC value:	0001010111110011

- Raw data
  - Gain+ADU 16-bit
  - Useful for detector characteristics
  - Cannot be added together (5 in G0 + 11 in G1 = ?)
  - Doesn’t compress well
  - Currently useless for data processing
  - Potentially could improve data processing, as this is the most direct output

Considerations of JUNGFRAU Output

Pixel output in JF: 0001010111110011

Gain: 0001010111110011 00:G0 01:G1 11:G2

ADC value: 0001010111110011

Result: $= \frac{ADC - pedestal}{gain}$

- Energy
  - Units of eV
  - Results of floating point operation, so in principle result is float
  - Negative number possible (if ADC < pedestal due to pedestal RMS)
  - Preferred at XFELs
  - Contains good description of background, charge sharing – spot position
  - Doesn't compress well

Considerations of JUNGFRAU Output

Pixel output in JF: 0001010111110011

Gain: 0001010111110011 00:G0 01:G1 11:G2

ADC value: 0001010111110011

Result: $= \left\lfloor \frac{ADC - pedestal}{gain * photon\ energy} + 0.5 \right\rfloor$

- Photon counting
  - Units of photons
  - Rounded to integers
  - Equivalent to EIGER result
  - Possibly small loss of information, but we do care about photons at the end
  - Good compression (roughly 6x vs. 16-bit vs. 10x for EIGER with bshuf/lz4)
  - Results obtained that way give expected data quality (EIGER as reference)
  - Differences are small

Considerations of JUNGFRAU Output

Pixel output in JF: 0001010111110011

Gain: 0001010111110011 00:G0 01:G1 11:G2

ADC value: 0001010111110011

Result: $= \left\lfloor \frac{ADC - pedestal}{gain * photon\ energy * n} + 0.5 \right\rfloor$

- Rounding to multiply of photons
  - Units of n * photons
  - Rounded to integers
  - Adds flexibility
    - Reducing n leads to better description of fractional counts
    - Increasing n reduces size of compressed file

Considerations of JUNGFRAU Output

Overload

Pixel output in JF: 1100000000000000

Bad pixel

Pixel output in JF: 1111111111111111

- Special values
  - Overload is dependent on gain(pixel), so need to assign special value INT32_MAX
  - Bad (dynamically) pixel needs to be also marked, best with different number as overload, INT32_MIN (?)

- Mask is calculated every time pedestal (dark current) is collected

PAUL SCHERRER INSTITUT

Increase in Frame Rates is a Challenge

2006 - PILATUS 6M 12.5 Hz

300 MB/s

2015 - EIGER 16M 133 Hz

3 GB/s

2021 - JUNGFRAU 10M 2.2 kHz

46 GB/s

(uncompressed data at full speed)

Increase in Frame Rates is a Challenge

2006 - PILATUS **6M** 12.5 Hz

300 MB/s

2015 - EIGER **16M** 133 Hz

3 GB/s

2021 - JUNGFRAU **10M** 2.2 kHz

Detector size is relatively constant over time

46 GB/s

(uncompressed data at full speed)

Increase in Frame Rates is a Challenge

2006 - PILATUS 6M **12.5 Hz**

300 MB/s

2015 - EIGER 16M **133 Hz**

3 GB/s

2021 - JUNGFRAU 10M **2.2 kHz**

46 GB/s

Frame rate is constantly improving and makes operating new detectors challenging

(uncompressed data at full speed)

Increase in Performance of Intel CPUs

2006 **Intel Xeon 5160**
3.0 GHz 2-core
21 GB/s memory (4 x DDR2-533 FB-DIMM)
4 GB/s interface (PCI Express Gen1 x16)

2019 **Intel Xeon Platinum 8280**
2.7 GHz 28-core
130 GB/s memory (6 x DDR4-2933)
16 GB/s interface (PCI Express Gen3 x16)

Fastest CPU from Intel

Over 13 years:

SLS MX detector data rate: 160x

Memory bandwidth: 6x

Peripheral interface: 4x

Data rate increase is much faster than improvements in computing

All Steps in the Data Acquisition Pipeline Become a Challenge

- Detector sends the data in asynchronous manner (UDP/IP) – frames not received in time are lost forever
  - With strongest server available from HPE (4 socket Xeon, 1.5 TB RAM) and standard code and kernel network I/O we can receive data for JUNGFRAU 4M 1.1 kHz (10 GB/s) to RAM disk
  - With online conversion and/or writing the data to remote storage capacity is around half (5 GB/s)

- If isolated, the server above can do frame conversion at max ~ 20 GB/s – limitation comes from need to load conversion constants from RAM memory

- Storage – both in terms of performance (currently 5-20 GB/s) and size (PBs)

All Steps in the Data Acquisition Pipeline Become a Challenge

- Detector sends the data in asynchronous manner (UDP/IP) – frames not received in time are lost forever
  - With strongest server available from HPE (4 socket Xeon, 1.5 TB RAM) and standard code and kernel network I/O we can receive data for JUNGFRAU 4M 1.1 kHz (10 GB/s) to RAM disk
  - With online conversion and/or writing the data to remote storage capacity is around half (5 GB/s)

- If isolated, the server above can do frame conversion at max ~ 20 GB/s – limitation comes from need to load conversion constants from RAM memory

- Storage – both in terms of performance (currently 5-20 GB/s) and size (PBs)

- How to handle 50 GB/s for JUNGFRAU 10M 2.3 kHz?
  - **Parallelize with standard components**
  - **Use specialized components**

Moving Data is a Challenge for JUNGFRAU Data Acquisition and Online Processing

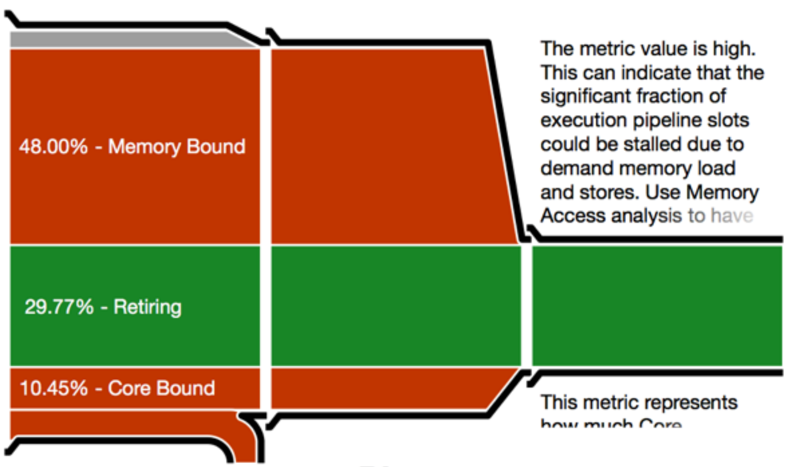
Computing for MX beamline online operation depends mainly on ability to move data fast – computing capability is usually less of a concern.

E.g. conversion of JUNGFRAU 10M (2.2 kHz) images from raw to photon requires handling extreme data rates (46 GB/s = 40 x 10G fiber optics), but only ~300 GFLOPS computing (iPhone 7)

fs.com

apple.com

Moving Data is a Challenge for JUNGFRAU Data Acquisition and Online Processing



μPipe

This diagram represents inefficiencies in CPU usage. Treat it as a pipe with an output flow equal to the "pipe efficiency" ratio: (Actual Instructions Retired)/(Maximum Possible Instruction Retired). If there are pipeline stalls decreasing the pipe efficiency, the pipe shape gets more narrow.

Elapsed Time: 41.725s

SP GFLOPS: 73.079

Effective CPU Utilization: 61.7%

Average Effective CPU Utilization: 29.628 out of 48

Effective CPU Utilization Histogram

Memory Bound: 48.0% of Pipeline Slots

Cache Bound: 21.0% of Clockticks

DRAM Bound: 20.8% of Clockticks

NUMA: % of Remote Accesses: 0.0%

FPU Utilization: 1.3%

SP FLOPs per Cycle: 0.803 Out of 64

Vector Capacity Usage: 97.8%

FP Instruction Mix:

% of Packed FP Instr.: 97.7%

% of 128-bit: 0.1%

% of 256-bit: 0.0%

% of 512-bit: 97.6%

% of Scalar FP Instr.: 2.3%

FP Arith/Mem Wr Instr. Ratio: 0.417

Solutions for Increasing Data Rates

- Continue business as usual
  - Dennard's scaling has ended – no more faster CPUs
  - Moore's law is soon to hit the limit – core count and CPU complexity reach limits
  - Only practical solution is to parallelize into more physical computing systems → instead of one large detector, get data from modules

- Find new ways
  - Some workloads (AI) will still get improvement, but general computing won't
  - New specialized tools available

Learning from High Performance Computing

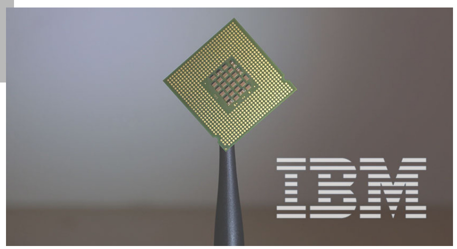
ornl.gov

Summit, ORNL (top supercomputer on Top500)

IBM AC922: POWER9 with Nvidia Volta GPUs

Specialized for AI – doing small computations on large volumes of data

Heterogenous Computing for Data Acquisition



IBM POWER9 CPU

CPU with faster memory and input/output as compared to Intel + single virtual address space for all accelerators

Xilinx FPGA

FPGAs are perfect for real time applications, FPGA boards can offer receiving, conversion and compression in a single device

Nvidia GPU

GPUs are suited for pixel processing - very efficient for tasks like image conversion or spot finding

Mellanox Connect-X network card

User space Ethernet and RDMA – significantly reduces overhead for communication

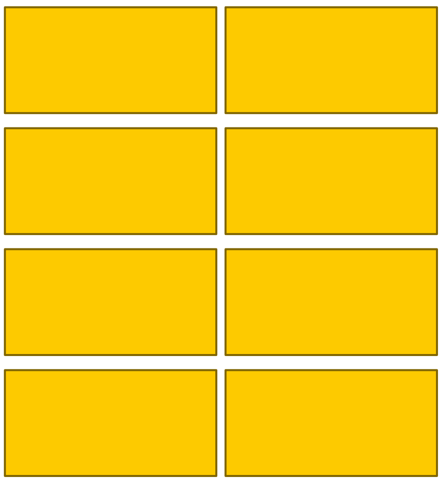
Single 2U to Handle JUNGFRAU 10M 2.2 kHz

ibm.com

IBM AC922

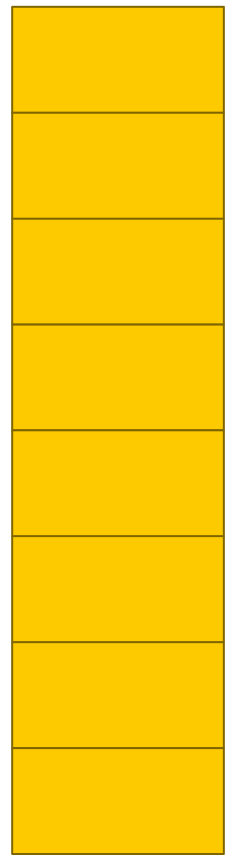
PAUL SCHERRER INSTITUT

Image Representation



“Synchrotron” format

- Modules are placed in the image according to their real position (+/- half pixel)
- Simple for data analysis programs
- One line corresponds to few modules → all of them need to be known for compression
- Virtual datasets possible in HDF5, but performance is a problem



“XFEL” format

- Modules are placed sequentially
- Needs instructions to rebuild real image for analysis and visualization
- Good for non-planar detectors
- Supported by newer data analysis programs (XDS, DIALS, CrystFEL), but might be less intuitive, esp. for XDS users
- Very nice for data acquisition

NXmx Comaptibility

It is important that JUNGFRAU produces NXmx “gold standard” files from 1st day of user operation

HDF5 1.10

- We currently use HDF5 1.10.5 for writing, but without 1.10 specific features (no VDS, etc.)
- Neggia has no issue reading files written with 1.10.5
- The only instability observed are occasional problems with closing the HDF5 library
- Problem with thread-safety – intensive use of direct chunk writer for parallel compression
- Direct chunk writer is part of main HDF5 library in 1.10.5 – while in high-level API in 1.8

NXmx Wish List

What is currently missing to properly describe JF dataset:

- Conversion factor between pixel unit and photons or energy
  - Important for counting statistics model
- Pedestal (dark_current) and gain factors for each gain level
- Pedestal G0 RMS (dark_current_rms) – this is probably the most important piece of information for data processing software
- Pedestal drift correction applied/not applied
  - or even better “list of applied corrections”, which would be the most universal solution
- Trigger timing (pedestal -> delay -> trigger –> measurement)
- Lowest value (underload?)

Nice to have:

- Spot positions list (if HW/ML spot finder is used, this information will be generated together with images) → consistency with CXI format
- Spot finding algorithm

- Gain switching statistics – array with number of G1/G2 switched pixels per frame

Jungfrau detector…

… will enable new science for macromolecular crystallography beamlines at bright X-ray sources.

… needs new methods to handle ever increasing data rates.

Wir schaffen Wissen – heute für morgen

My thanks go to

- SLS MX Group
- SLS Detector Group
- PSI Science IT

- Photon Factory, KEK
- CERN LHCb
- IBM
- InnoBoost AG
- Dectris AG

- PAL XFEL
- LBNL
- DLS