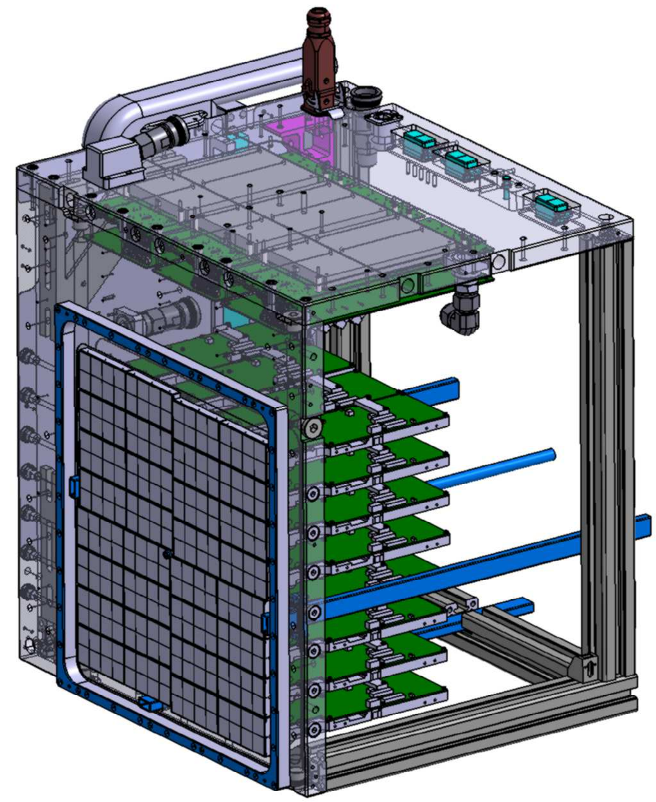
Carlos Lopez-Cuenca :: Paul Scherrer Institut

Online data processing for high data rate detectors

New concepts in ultra fast data acquisition
10-11/04/2018

Why online data processing?

- **Conversion to number of photons**
  - 16Mpx Jungfrau detector @ 100Hz for SwissFEL
    - 3.4GBps of raw data
- **Live data analysis**
  - 4Mpx Jungfrau detector @ 1kHz for a 5-day experiment in Japan
    - 8.4GBps of raw data
    - 250TB of data stored after the experiment
- **Data reduction**
  - 10Mpx Jungfrau detector @ 2kHz for SLS
    - 42GBps of raw data

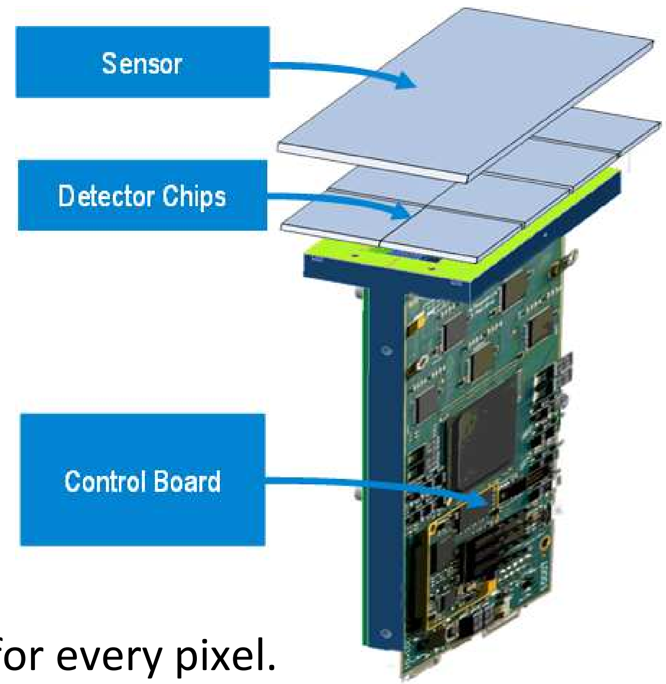


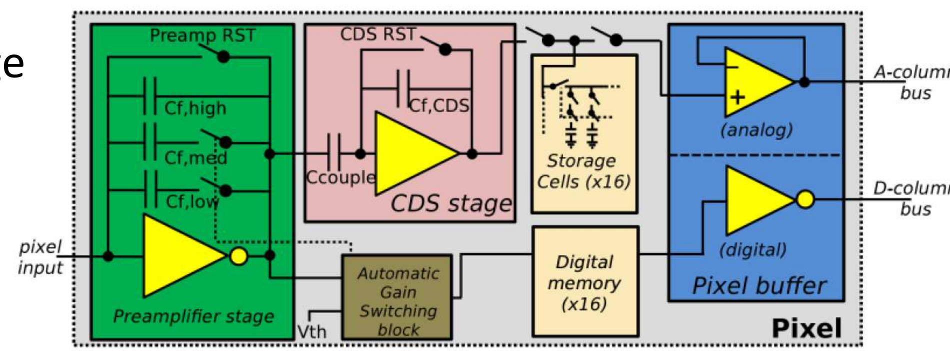
16M Jungfrau detector for the Bernina Experimental Station @ SwissFEL

- **So, why online data processing?**
  - Provide a live (reduced) data analysis enough to monitor the validity of the experiment
  - Reduce the throughput and the amount of data to be stored

PSI detectors

- **Hybrid detectors**
  - Sensor + Detector chip + Control electronics
- **In 2 different families**
  - Single photon counting detectors
    - Counts the photons that hit every pixel during the exposure.
  - Charge integrating detectors
    - Integrate the charge collected by the sensor for every pixel.
    - Automatic gain switching
      - To expand the dynamic range





2 Families – 2 Processing pipelines

- **Single photon counting detectors**
  - Output → Number of photons that hit each pixel

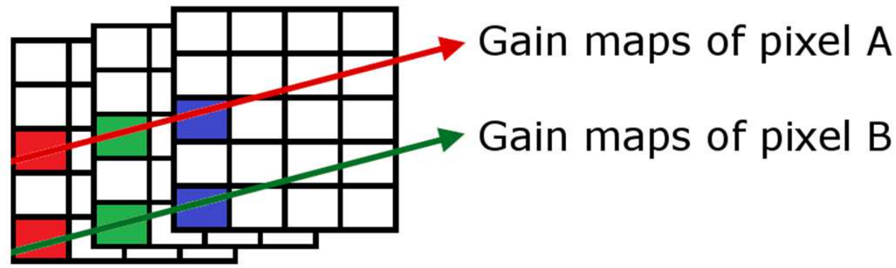
- **Charge integrating detectors**
  - Output → Amplification stage + Amplified charge
  - **Needs to be converted to the number of photons**
    - Q1: which gain are we in?
    - Q2: how far above pedestal are we?
    - Q3: what energy caused that?
    - Q4: how many photons does that mean?

01	00001011001000
Gain	ADC value
$00 = G_0$ $01 = G_1$ $11 = G_2$	0..16383
Amplification used	Amplified charge

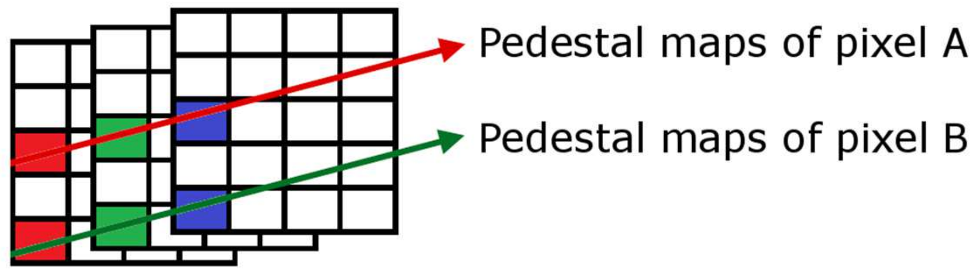
$$N_\gamma = \frac{|\mathrm{ADC} - \mathrm{pede}|\,[\mathrm{ADU}] \times \mathrm{gain}\,[\mathrm{keV}/\mathrm{ADU}]}{E_{\mathrm{beam}}\,[\mathrm{keV}]}$$

Conversion to Number of Photons

- **Every pixel is characterized with:**
  - 3 gain constants
  - 3 pedestal variables

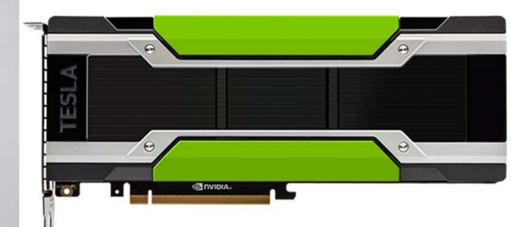


- **For a 0.5Mpx Jungfrau detector:**
  - **768Mb** (16-bit) of memory
    - 16 storage cells
  - For 2.4kHz rates
    - DDR3 controller @ **>800MHz**



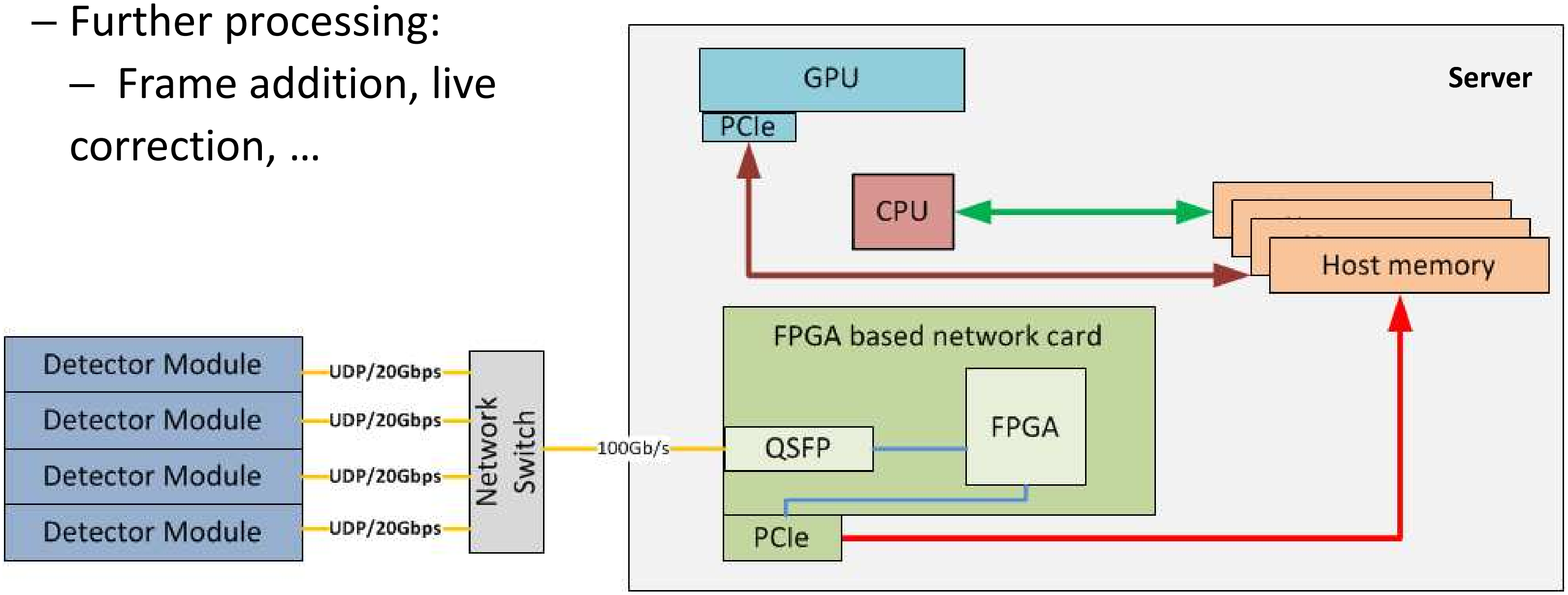
- **System checklist**
  - Computational power
  - Memory bandwidth
  - Scalability
  - Flexibility
  - Development time
  - Cost

FPGA + GPU



Online data processing: FPGA & GPU

- **System proposal**
  - Module control board
    - Controls the acquisition and readout of the detector chips
    - Streams out images via GbE
  - FPGA –based network card
    - Assembles full images in the host memory
  - GPU-based application
    - Computes the number of photons for every pixel
    - Further processing:
      - Frame addition, live correction, …



GPU-based processing application

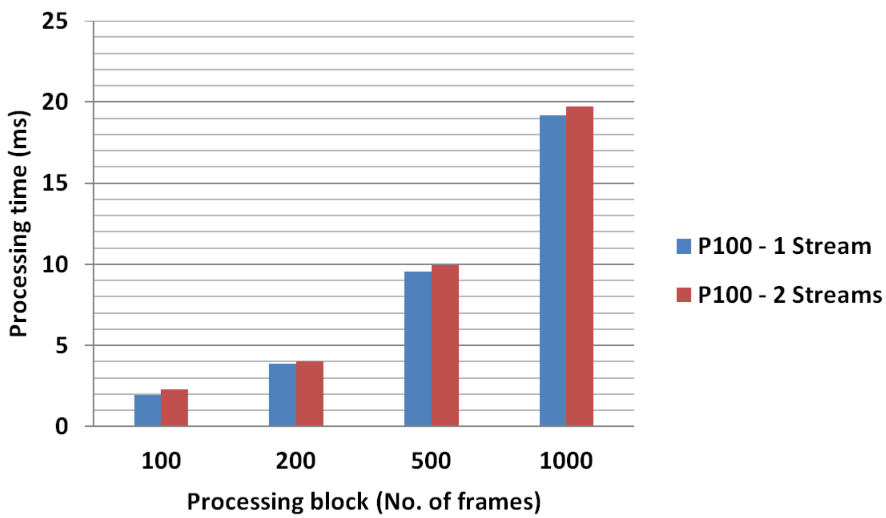
- **Scalable application for charge integrating detectors**
  – Performs the conversion to number of photons
  – Updates pixel pedestal values as they drift over time
  – Sums a configurable number of consecutive images
- **Open collaboration with the HZDR Computational Radiation Physics group**
- **Application integrated @ PSI to process data from the detector**
  – GPU: Nvidia Tesla P100
    – Core clock: 1.3GHz
    – Memory Bandwidth > 500GBps

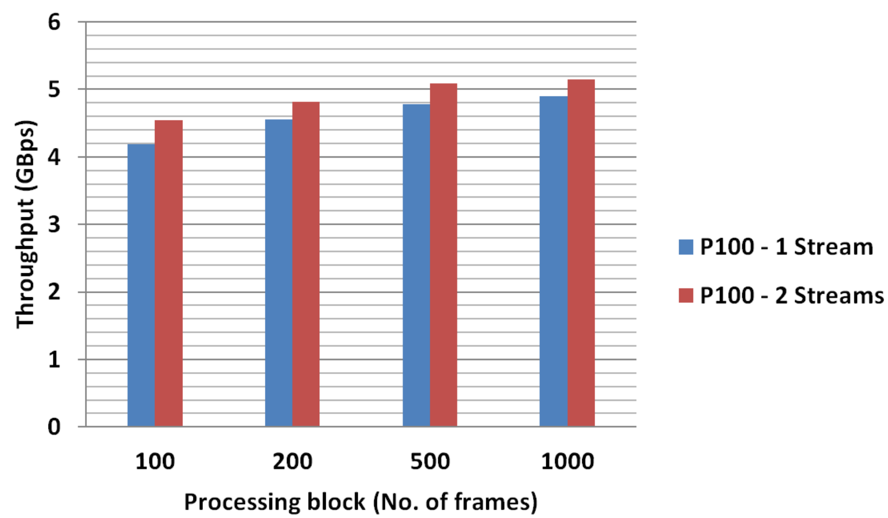
Features	P100
Compute capability	6.0
Number of Multiprocessors	56
Number of Streaming Processors	3584
Streaming Processors/Multiprocessor	64
Threads/Warp	32
Max Warps/Multiprocessor	64
32-bit Registers/Multiprocessor	65536
Memory Controller	8x512b HBM2

	Compute Capability						
	2.0	2.1	3.0, 3.2	3.5, 3.7	5.0, 5.2	5.3	6.0
16-bit floating-point add, multiply, multiply-add	N/A	N/A	N/A	N/A	N/A	256	128
32-bit floating-point add, multiply, multiply-add	32	48	192	192	128	128	64
64-bit floating-point add, multiply, multiply-add	16[1]	4	8	64[2]	4	4	32
32-bit floating-point reciprocal, reciprocal square root, base-2 logarithm (__log2f), base 2 exponential (exp2f), sine (__sinf), cosine (__cosf)	4	8	32	32	32	32	16
32-bit integer add, extended-precision add, subtract, extended-precision subtract	32	48	160	160	128	128	64

Performance: 0.5Mpx Jungfrau GPU processing

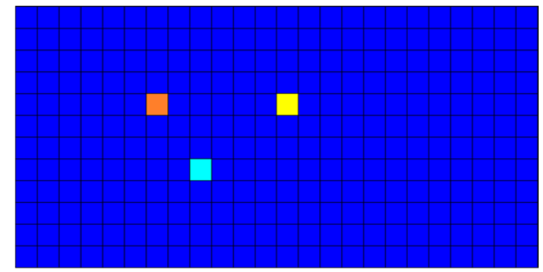
- **Tested the performance of the GPU based application with**
  - Only **20ms** are required to convert 1000 frames per module
  - Online processing application is able to process 1000 frames (1GB) in **198ms**
  - Main bottleneck is data transfer between host memory and GPU
    - Additional processing could be carried out without performance penalty
  - Maximum throughput should be ≈ 11GBps for 2 processing streams
    - GPU application needs to be adapted to PASCAL architecture

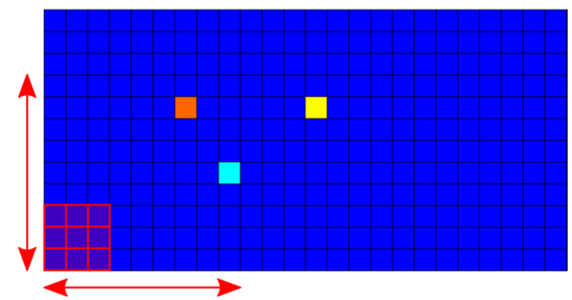




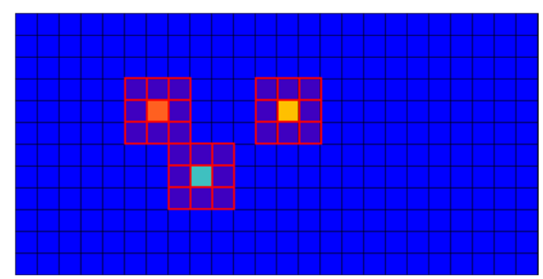
Further processing: Cluster Photon Finding

- **Cluster Finding Algorithm**
  - Extracts the region of interest for each photon detected
  - Input:
    - Raw data from the detector
  - Output:
    - Number of NxN clusters: (3x3), (5x5)
    - Coordinates of the cluster center
    - Signal value of each pixel in the cluster
    - Noise value of each pixel in the cluster



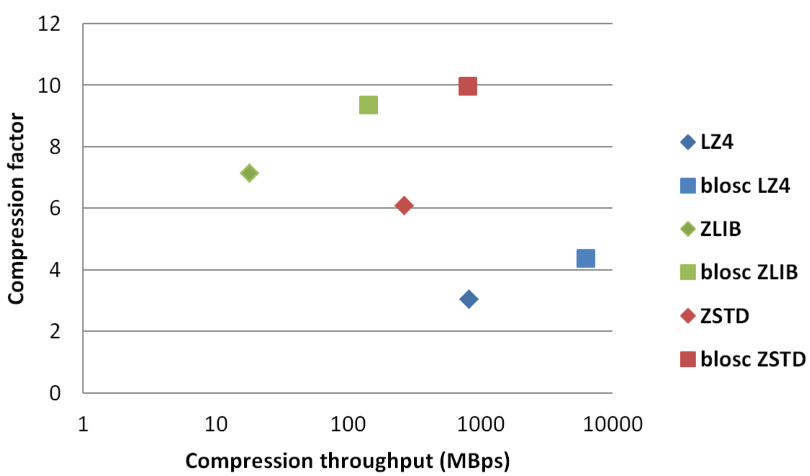


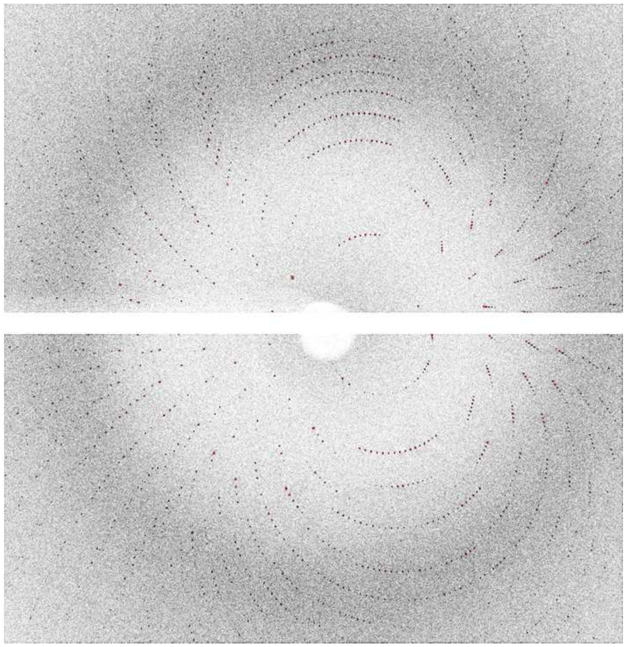
- **Data reduction**
  - Further analysis proceeds with a reduced set of data



Further processing: Data compression

- **Converted images are compressible**
  - Data sets are very different for each experiment
  - Best results obtained with BLOSC library
    - Meta-compressor: Compressor + pre-conditioners
  - Compression factor: x4 - x10 for single frames

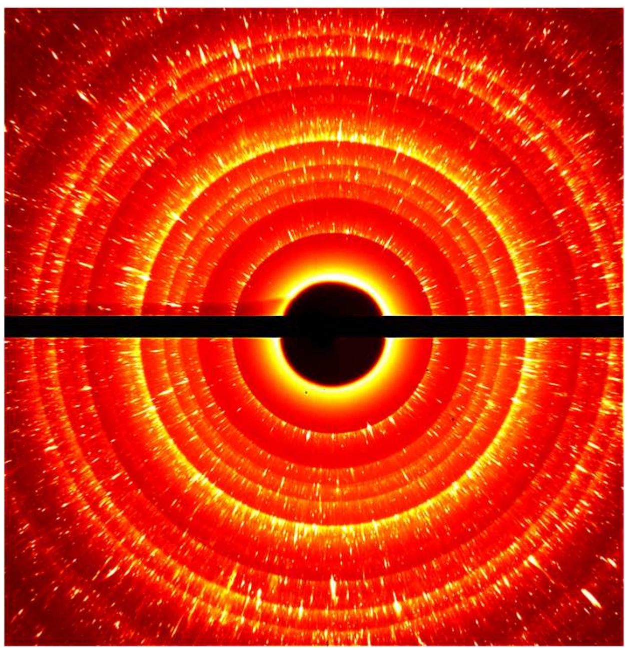




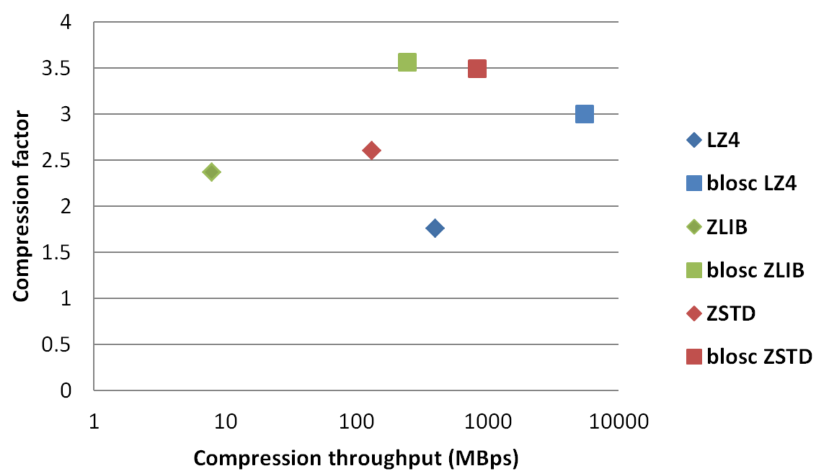
Single frame 1Mpx Jungfrau detector
Single Thaumatin crystal

Further processing: Data compression

- **Converted images are compressible**
  – Data sets are very different for each experiment
  – Best results obtained with BLOSC library
    – Meta-compressor: Compressor + pre-conditioners
  – Compression factor: x3 for 1000 consecutive added frames



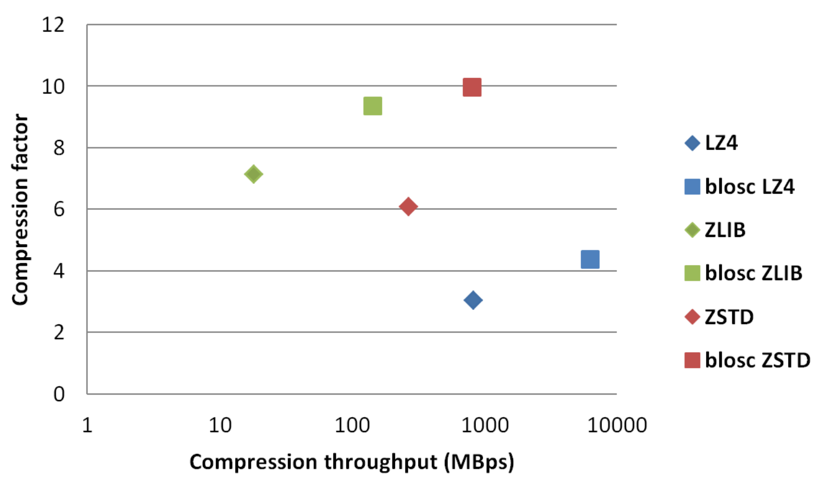
1000 added frames 1Mpx Jungfrau detector
Vitamin-C powder diffraction



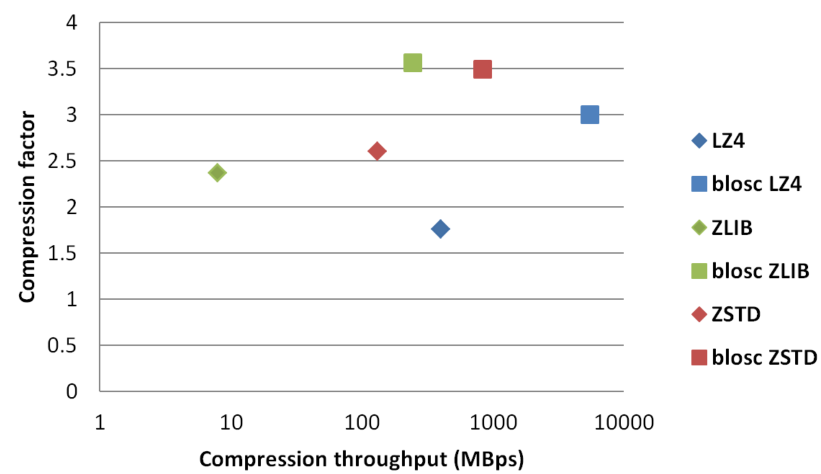
PAUL SCHERRER INSTITUT

Further processing: Data compression

- **Ultimate goal is to reduce the data volume to a point where the maximum throughput is ≈4GBps**
  – Compression factor: x4 - x10 for single frames
  – Compression factor: x3 for 1000 consecutive added frames



Single frame 1Mpx Jungfrau detector
Single Thaumatin crystal



1000 added frames 1Mpx Jungfrau detector
Vitamin-C powder diffraction

Summary

- **Data volumes and throughput of high data rate detectors threatens the capacity of current data back-end systems**
  - Processing (1 CPU core per 10GbE interface)
  - Storage space (max of 4GBps)
- **We have presented a system solution that should offload computational load at the data back-end and reduce storage requirements**
  - FPGA –based network card
    - Offloads CPU from emptying UDP queues
  - GPU based application
    - Performs online conversion to number of photons for charge integrating detectors
    - Addition consecutive frames
- **Further steps**
  - Implement the Cluster Photon Finding algorithm in the GPU -based application
  - Reduce data volume running a high-throughput lossless compression algorithm

Wir schaffen Wissen – heute für morgen