DAQ system at SACLA and future plan for SPring-8-II

Takaki Hatsui
T. Kameshima, Nakajima
T. Abe, T. Sugimoto
Y. Joti, M.Yamaga
RIKEN SPring-8 Center

Evolution of Computing infrastructure from SPring-8, SACLA and SPring-8-II

- SPring-8: 1997-
  - “Edge”-only computing

~60 beamlines
Now, some beamline group owns computation resource

Evolution of Computing infrastructure from SPring-8, SACLA and SPring-8-II

- SPring-8: 1997-
  - Edge-only computing

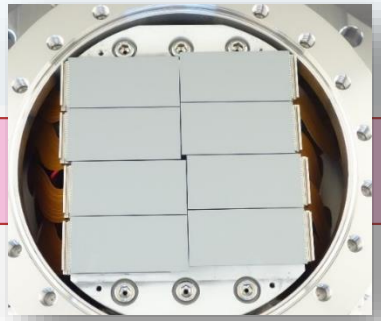
~60 beamlines
Now, some beamline group owns computation resource

- SACLA: 2009-
  - Centralized Computing
  - off-site K Computer

Computation Infrastructure
125 TFLOPS
7-28 Pbyte tage archive
10,000 TFLOPS

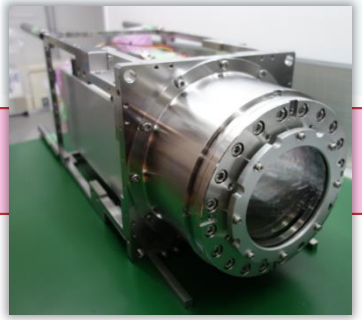
SACLA: Data Flow

Sensor



Detection

Camera Head



Digitization

Data acquisition (DAQ)

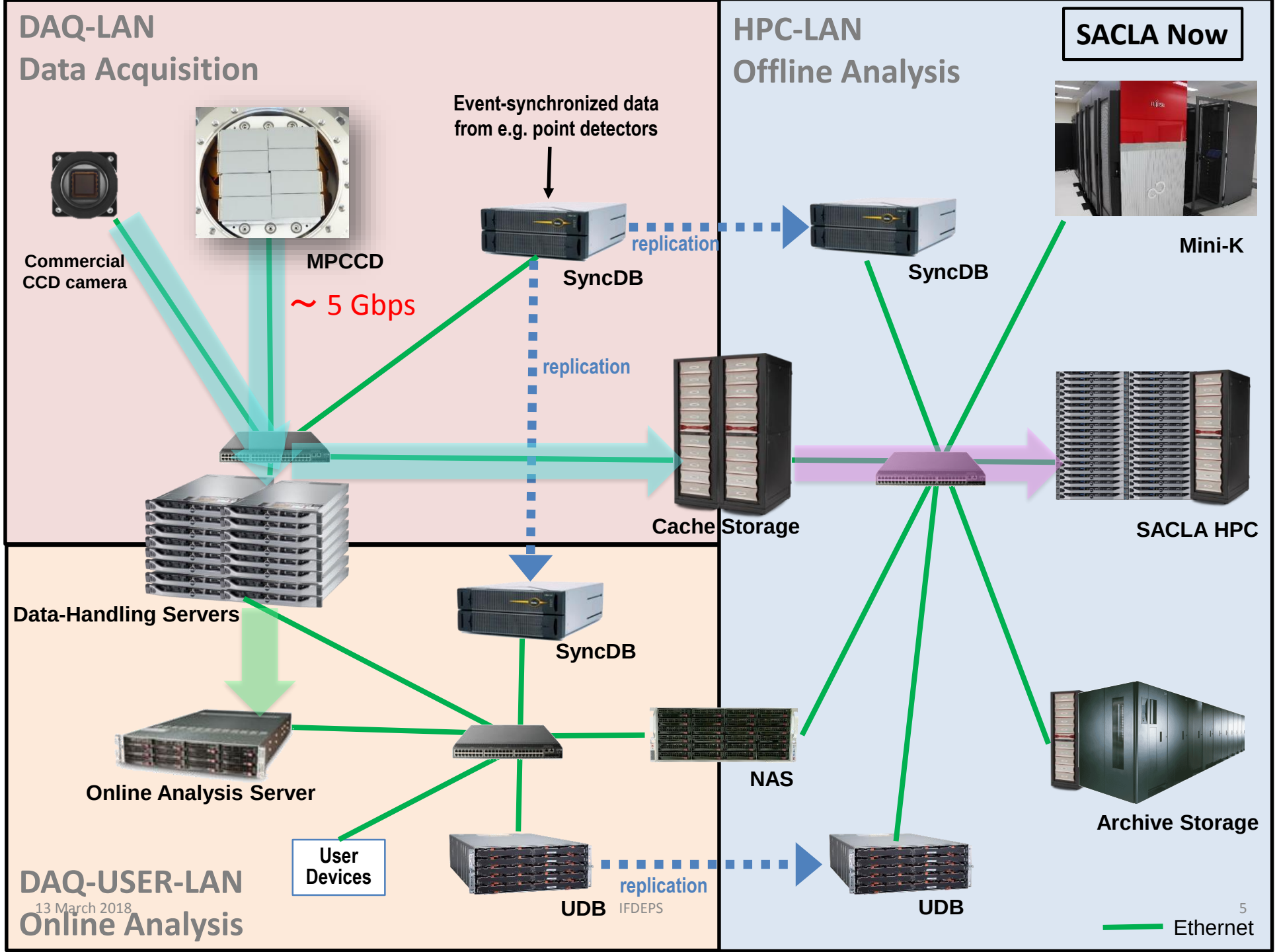
Storage

Supercomputer

Analysis

K-computer

Analysis


DAQ-LAN
Data Acquisition
HPC-LAN
Offline Analysis
SACLA Now
Event-synchronized data
from e.g. point detectors
Commercial
CCD camera
MPCCD
~ 5 Gbps
SyncDB
replication
SyncDB
Mini-K
replication
Cache Storage
SACLA HPC
Data-Handling Servers
SyncDB
Online Analysis Server
NAS
Archive Storage
User
Devices
UDB
replication
UDB
DAQ-USER-LAN
Online Analysis
Ethernet

SACLA Cache and Storage

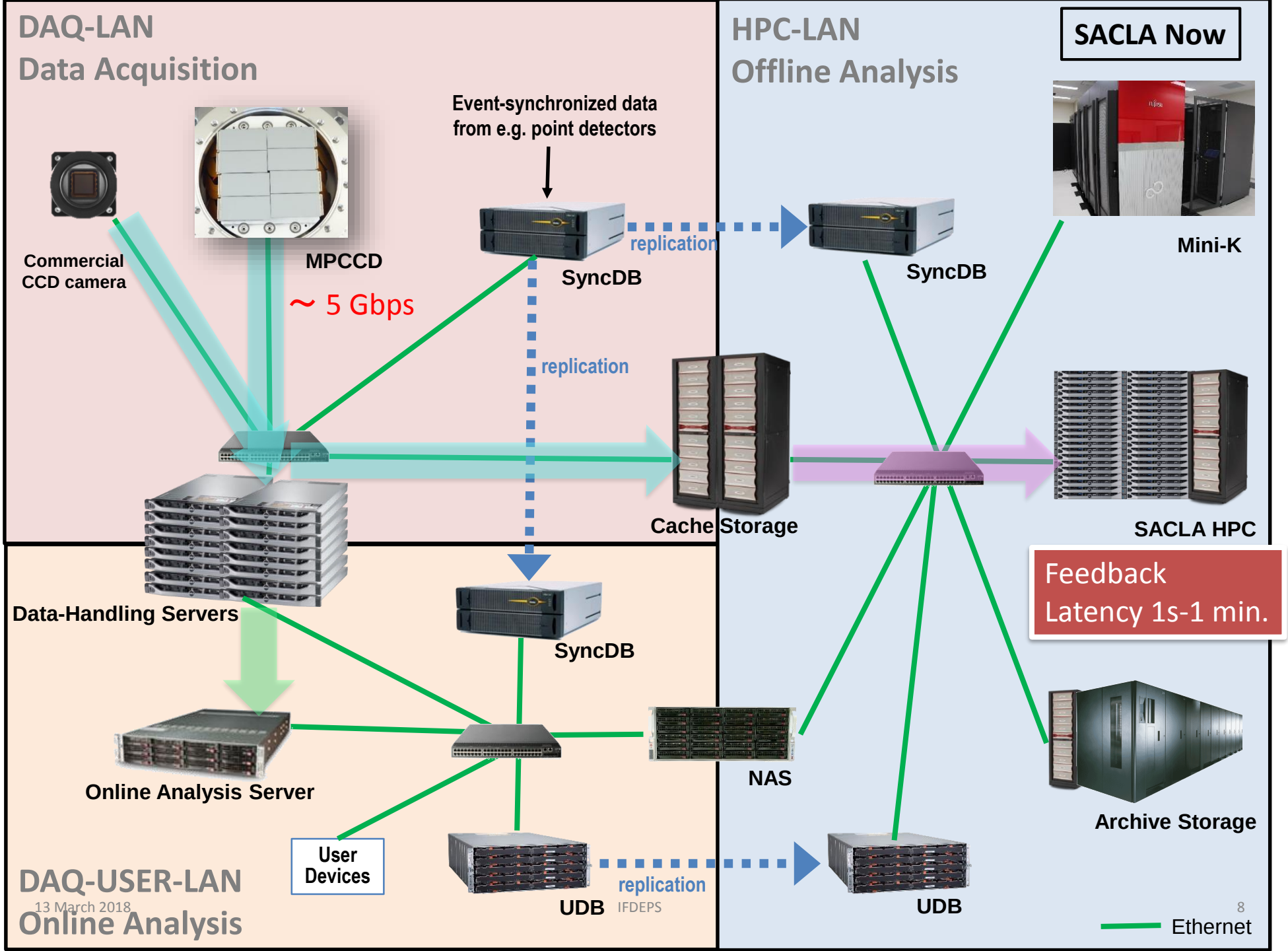
	Specifications
Cache storage system	[System for BL1 experiments] Hardware: DDN SFA7700X Storage capacity: ~100 TB File system: GPFS (I/O ~2 GB/sec) [System for BL2 experiments] Hardware: DDN SFA7700X Storage capacity: ~380 TB File system: GPFS (I/O ~7 GB/sec) [System for BL3 experiments] Hardware: DDN SFA10K Storage capacity: ~250 TB File system: GPFS (I/O ~5 GB/sec)
Archive storage system	Hardware: Disk: DDN SFA10K, Tape: IBM System Storage TS3500 & TS1140 (5 drives) Storage capacity: Disk: 1 PB & Tape: 7 PB (can be increased up to 28 PB) File system of the disk part: GPFS (I/O > 4 GB/sec) Hierarchical management: Tivoli Storage Manager (I/O ~200 MB/sec)

Y. Joti et.al., Synchrotron radiation news" (2017) 30(1) p.16.

Computing

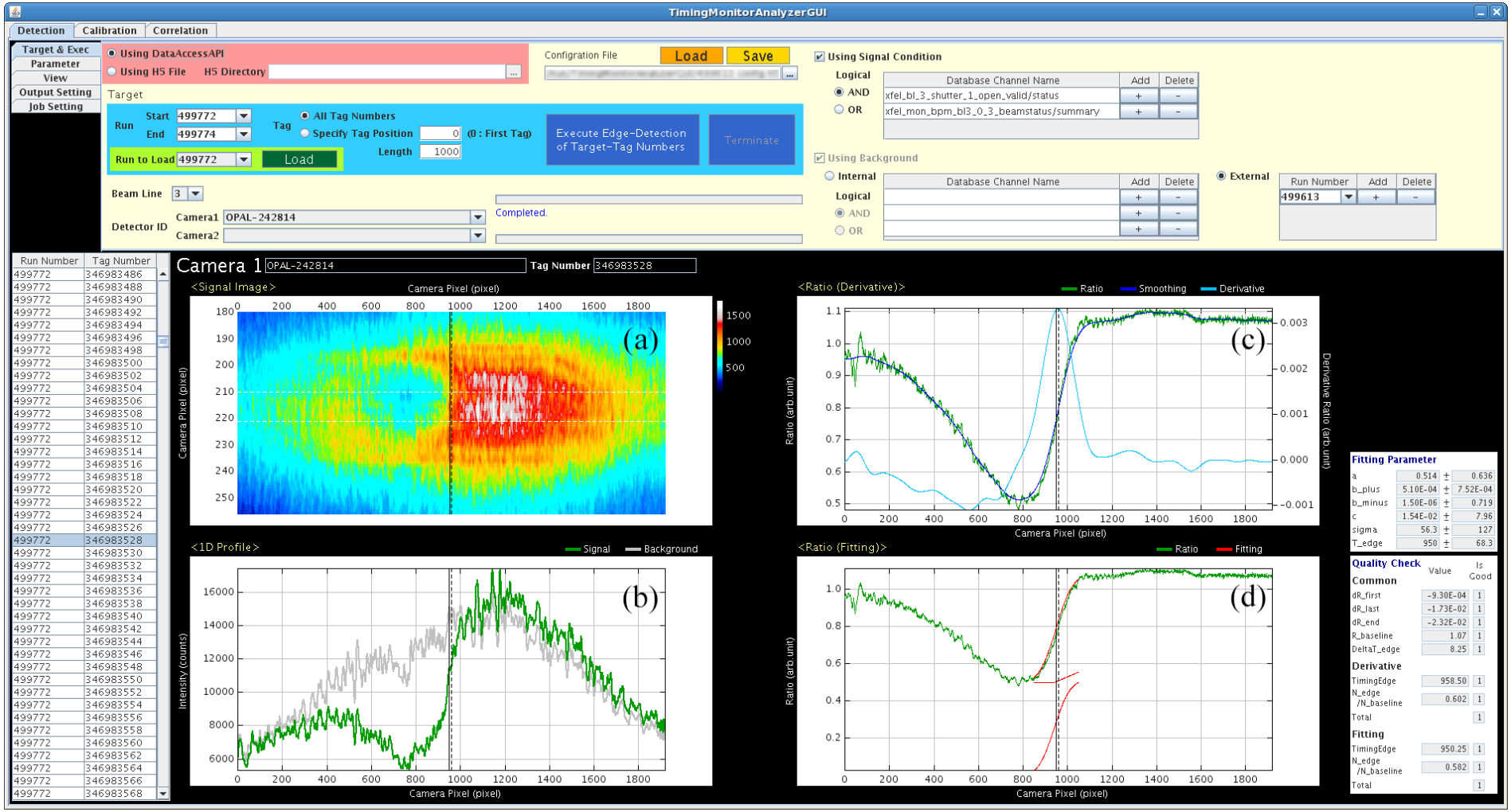
	Specifications
HPC system	Hardware: Compute nodes: DELL PowerEdge M630(40 nodes) Storage: DDN SFA12K Theoretical peak performance: **~35 TFLOPS** ([intel E5-2660 v4×2]×38 + [Intel E5-2699 v4×2]×2) Total memory capacity: ~3.5 TB (64 GB×38 + 512 GB×2) Interconnect: Infiniband FDR Storage capacity: ~2 PB File system: Lustre (I/O ~10 GB/sec)
Mini-K system	[Supercomputer] Hardware: Fujitsu PRIMEHPC FX10 4 rack model Theoretical peak performance: **~90 TFLOPS** (SPARC64 IXfx×384) Total memory capacity: ~12 TB (32 GB×384) Interconnect: Tofu 6D Mesh/Torus architecture Storage capacity: ~600 TB (global: 500 TB, local: 100 TB) File system: FEFS (I/O global: ~5 GB/sec, local:~10 GB/sec) [Storage for efficient data transfer] Hardware: DDN SFA12K Storage capacity: ~1 PB File system: Lustre (I/O ~10 GB/sec)

Y. Joti et.al., Synchrotron radiation news" (2017) 30(1) p.16.

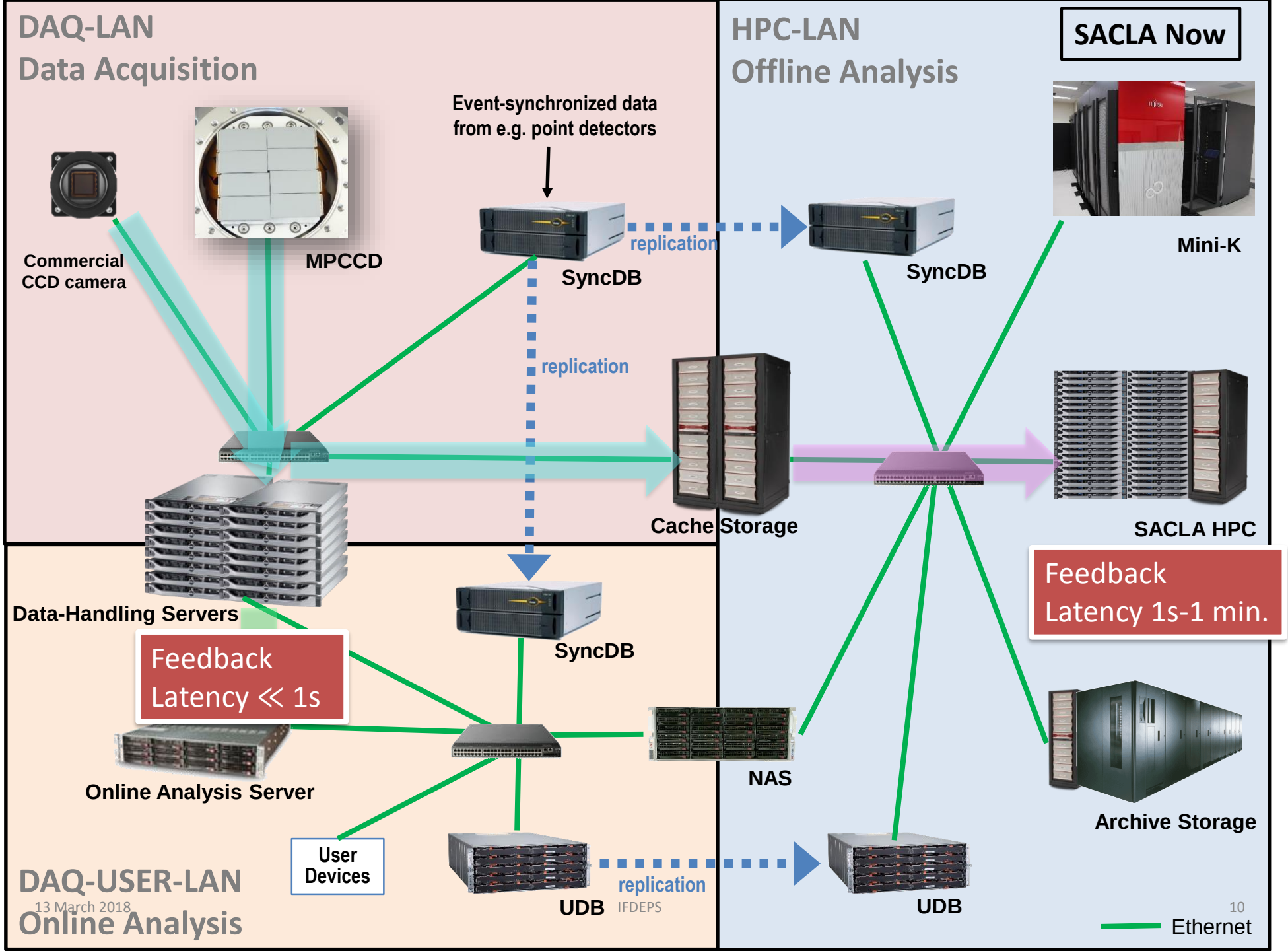

DAQ-LAN
Data Acquisition
HPC-LAN
Offline Analysis
SACLA Now
Event-synchronized data
from e.g. point detectors
Commercial
CCD camera
MPCCD
~ 5 Gbps
SyncDB
replication
SyncDB
Mini-K
replication
Cache Storage
SACLA HPC
Data-Handling Servers
SyncDB
Feedback
Latency 1s-1 min.
Online Analysis Server
NAS
Archive Storage
User
Devices
UDB
replication
UDB
DAQ-USER-LAN
Online Analysis
Ethernet

Offline Feedback: Timing,

K. Nakajima et.al., J. Synchrotron Rad. (2018). **25**, 592–603



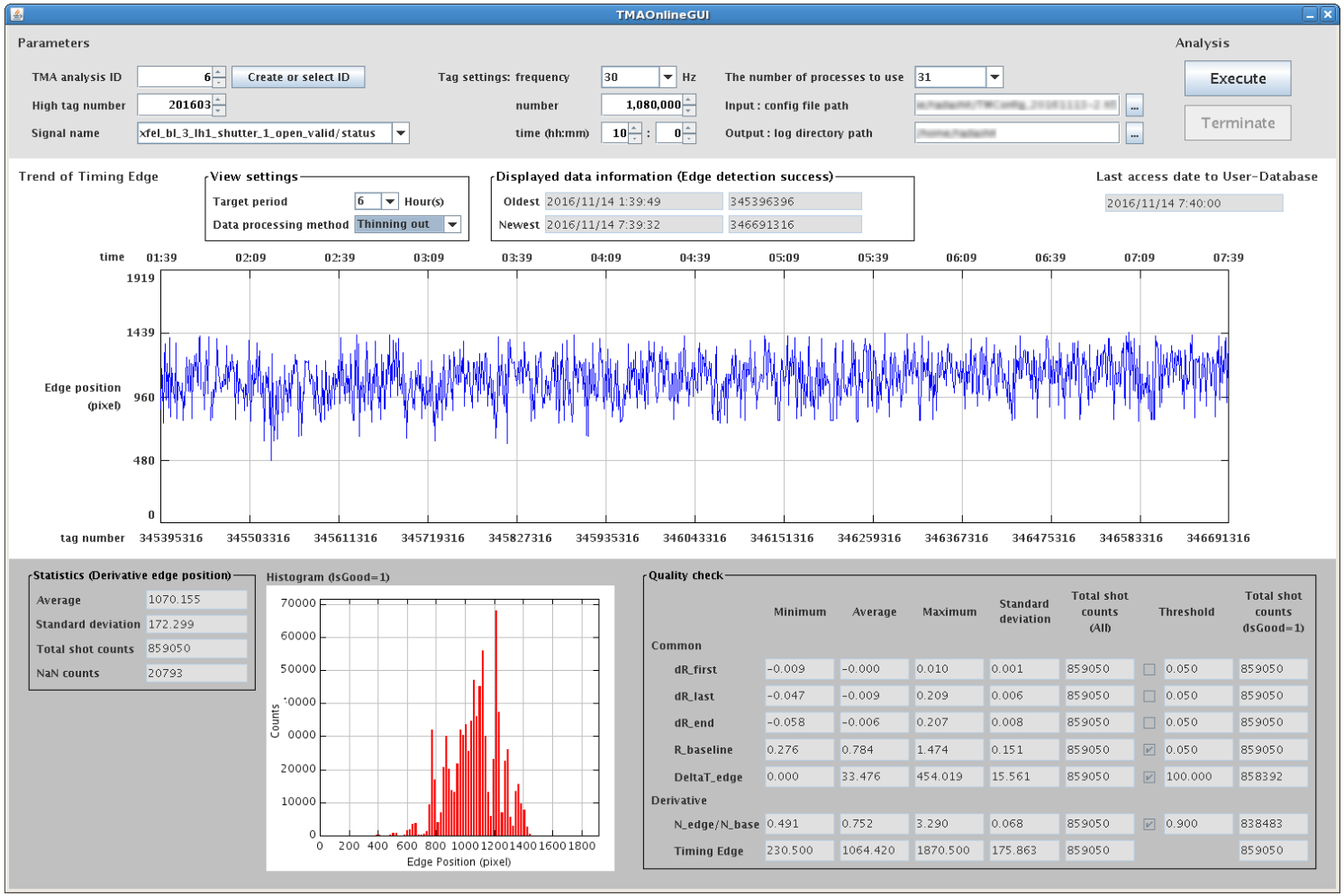
Other example: sample hit rate, indexing statistics of SFX


DAQ-LAN
Data Acquisition
HPC-LAN
Offline Analysis
SACLA Now
Event-synchronized data
from e.g. point detectors
Commercial
CCD camera
MPCCD
SyncDB
replication
SyncDB
Mini-K
replication
Cache Storage
SACLA HPC
Data-Handling Servers
SyncDB
Feedback
Latency ≪ 1s
Feedback
Latency 1s-1 min.
Online Analysis Server
NAS
Archive Storage
User
Devices
UDB
replication
UDB
DAQ-USER-LAN
Online Analysis
Ethernet

Online Feedback: Timing

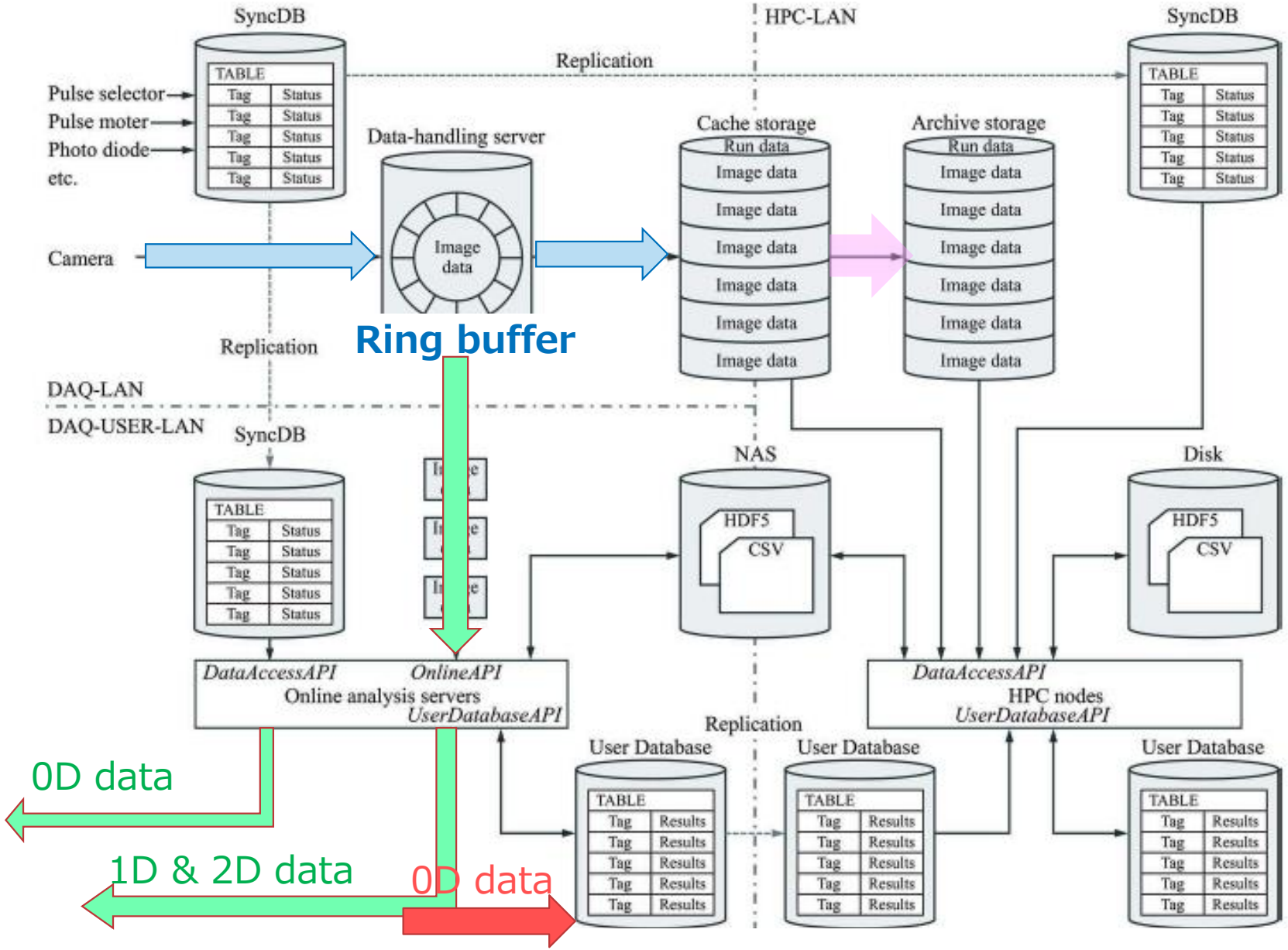
K. Nakajima et.al., J. Synchrotron Rad. (2018). **25**, 592–603

Data format: Value/pulse



A screen image of the GUI of online Timing monitor analyzer executed on the online analysis server.

On-line («1 s)　　Off-line (> 1s)



Lessons learned by operating SACLA DAQ/analysis

- *Now, we do not have phone call in the night. But we had many issues so far.*
- *Some of the critical ones are*
- *All-in-one DAQ/Analysis*
  - *Testing of new experimental setup prior to the beamtime is difficult.*
  - *Learning cost of new comers/ students is very high.*

    I do want to do this only, but why I should learn all of the DAQ software

- *Toward Spring-8-II,*
- *"Divide & Conquer" Policy*

Evolution of Computing infrastructure from SPring-8, SACLA and SPring-8-II

- SPring-8: 1997-
  - Edge-only computing

～60 beamlines
Now, some beamline group owns computation resource

- SACLA: 2009-
  - Centralized Computing
  - off-site K Computer

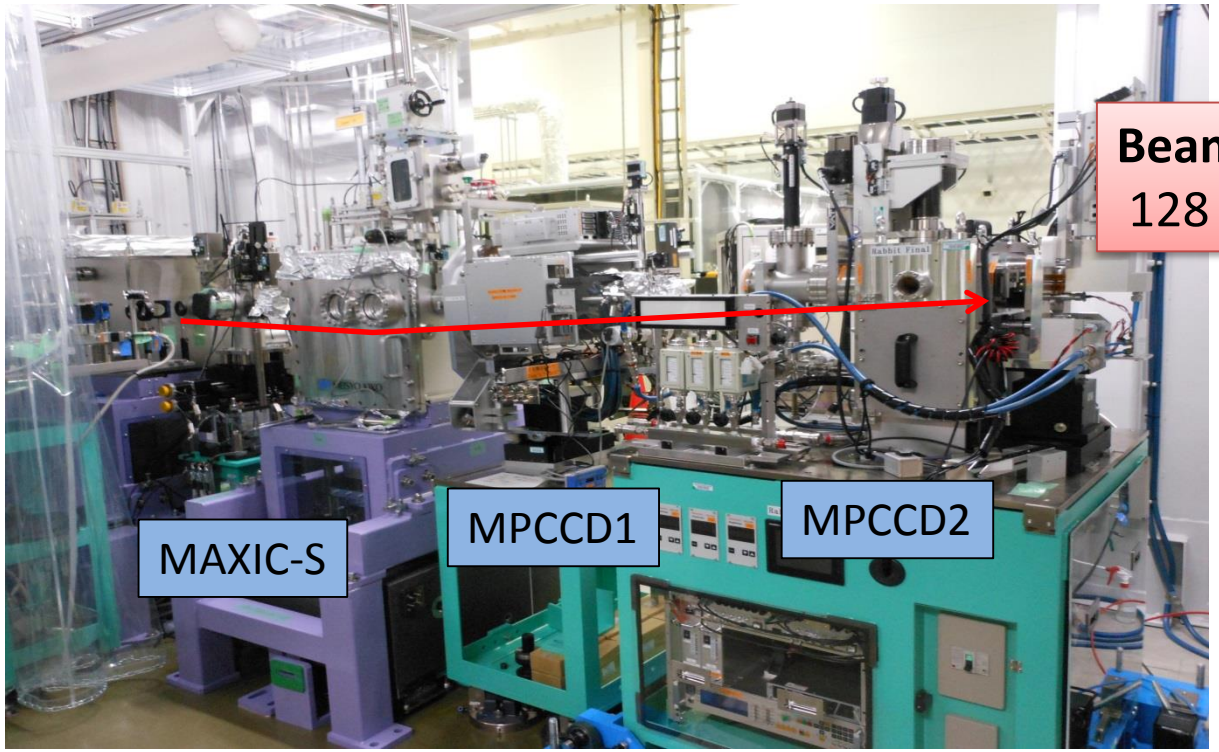
Computation Infrastructure
125 TFLOPS
7-28 Pbyte tage archive
10,000 TFLOPS

- SPring-8-II: 202X-
  - Edge-Heavy, Center-Heavy & Cloud Computing

Nano-beam CDI experiment at BL2

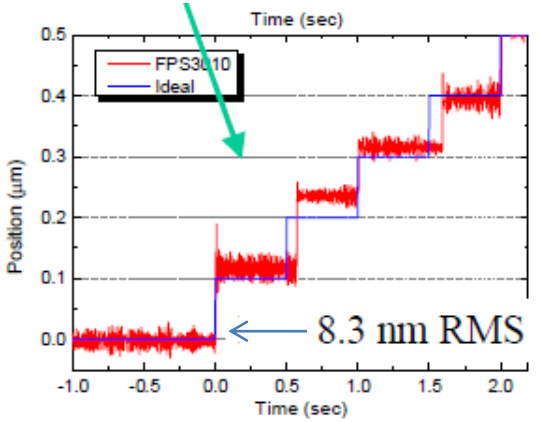

Kimura, Suzuki, Nishino
(Hokkaido U.) et al.

BL2-EH3



Beam size:
128 nm (H) x 70 nm (V) (FWHM)

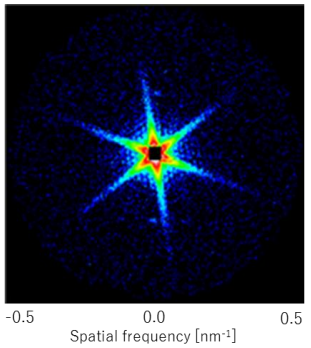
- Fast & precise sample scanning.

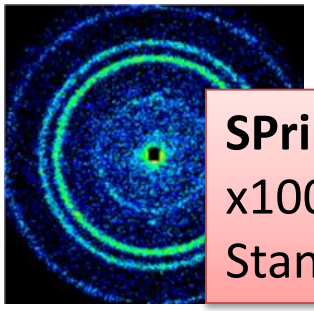


Metal nano triangle

Protein 2D crystal

First diffraction patterns
(MPCCD 1, Sept. 2017)



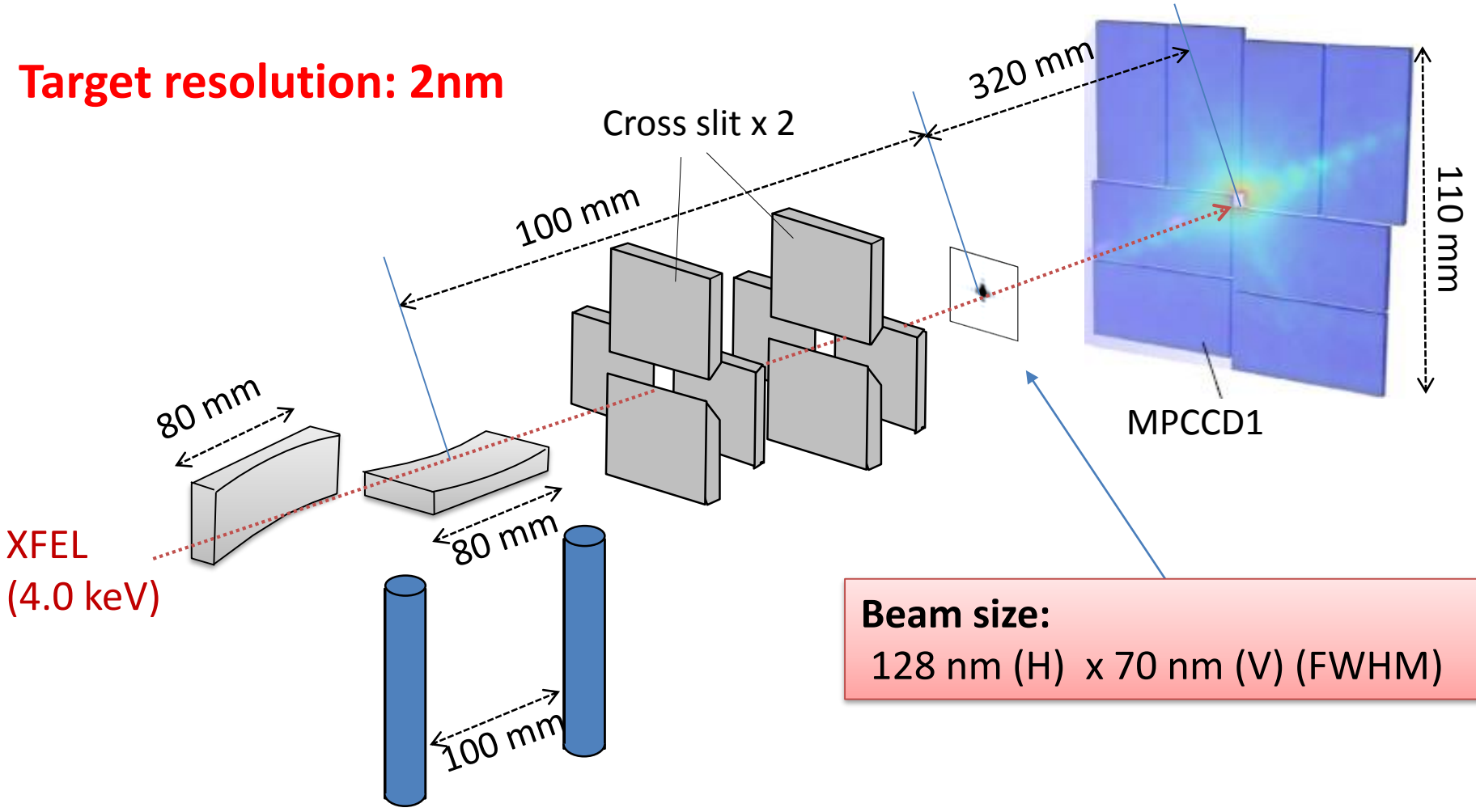


SPring-8-II
x1000 increase in flux for 100 nmϕ
Standard beamline

SACLA: MAXIC-S for nano-beam CDI

Yumoto, Koyama, Ohashi (SACLA & SPring-8) et al.

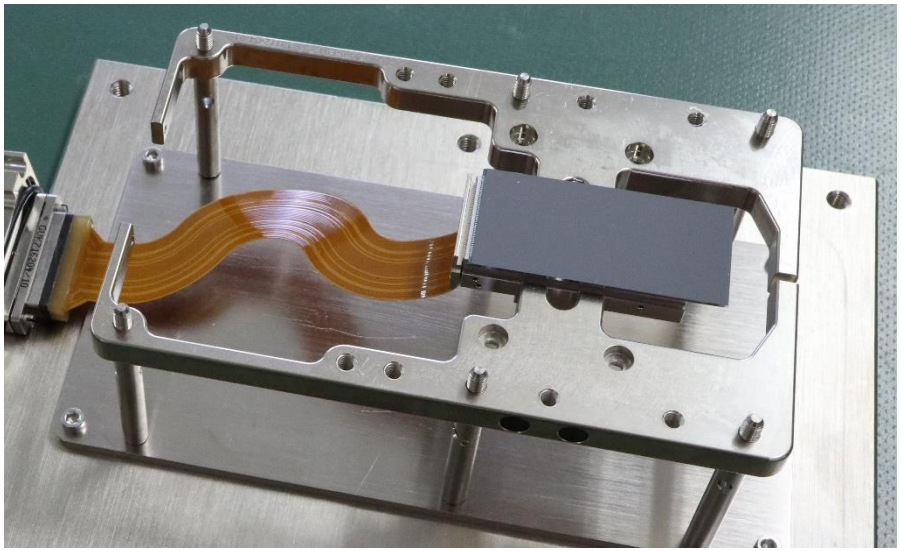
Target resolution: 2nm



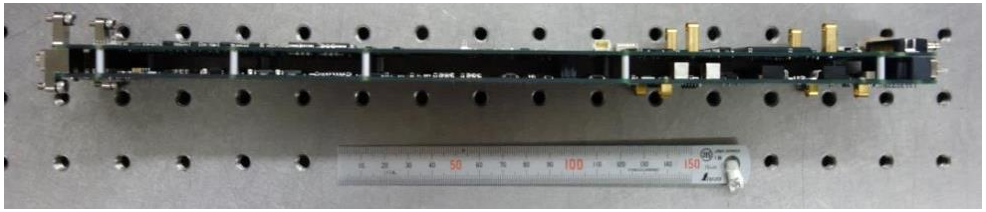
Beam size:
128 nm (H) x 70 nm (V) (FWHM)

10 nm drift @ ΔT = 1 mK

CITIUS: Minimize Heat Dissipation



MPCCD sensor unit
CITIUS package will be similar to MPCCD

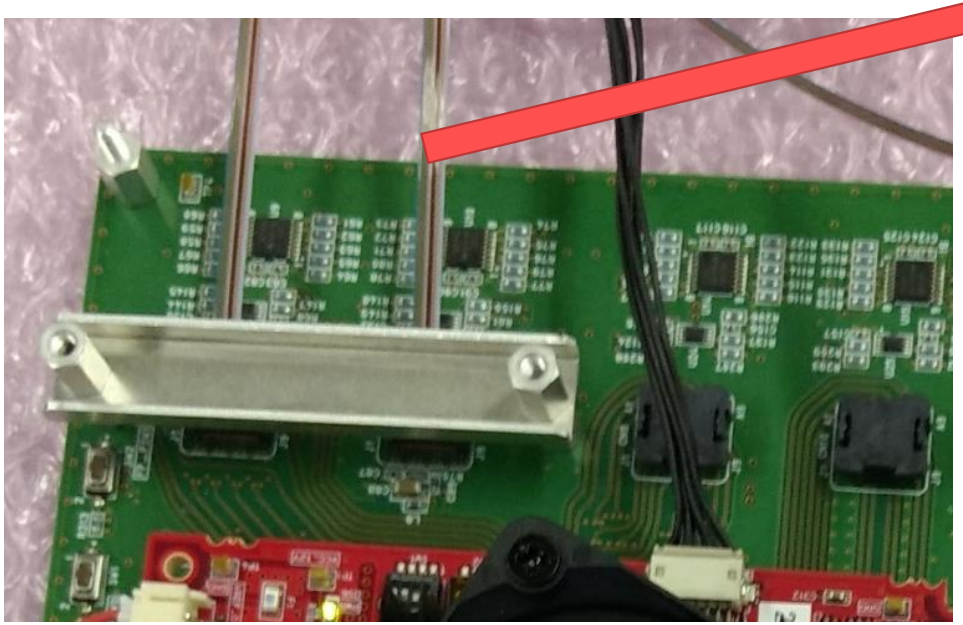


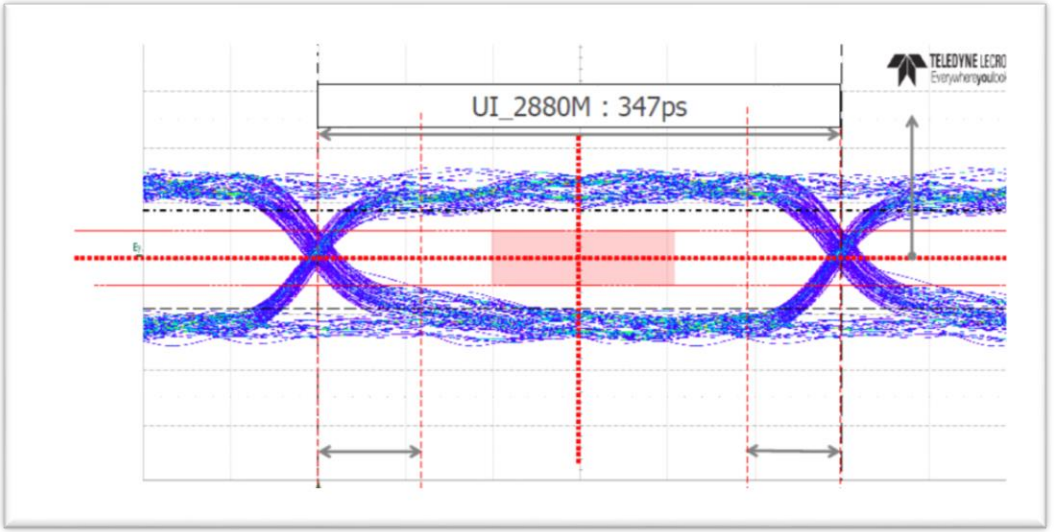
- CITIUS
  - 17 kfps
  - 0.3 Mpixel
  - 5.1 Gpixels/s
- Sensor
  - Digital serializer
  - 140 Gbps/sensor
    - Data:100 Gbps
    - Metadata + ECC: 40 Gbps
  - 8 W/sensor design confirmed
- Proximity board
  - Target 15 W
  - Without digital data handling

Sensor-Proximity Board
Direct transfer by EO transceivers

Toshinori Abe (JASRI/RIKEN)

Single optical cable assembly
gives 35 Gbps data transmission





Calibration

- 80 G operations/s/sensor
  - 2 gains in pixel
  - 2 gains in ADC
  - Dark subtraction
  - Average signal calculation

- 21.6 Mpixel system
  - 5.6 T operations/s/system
  - 1/7 of our current PC cluster resource @ SACLA.

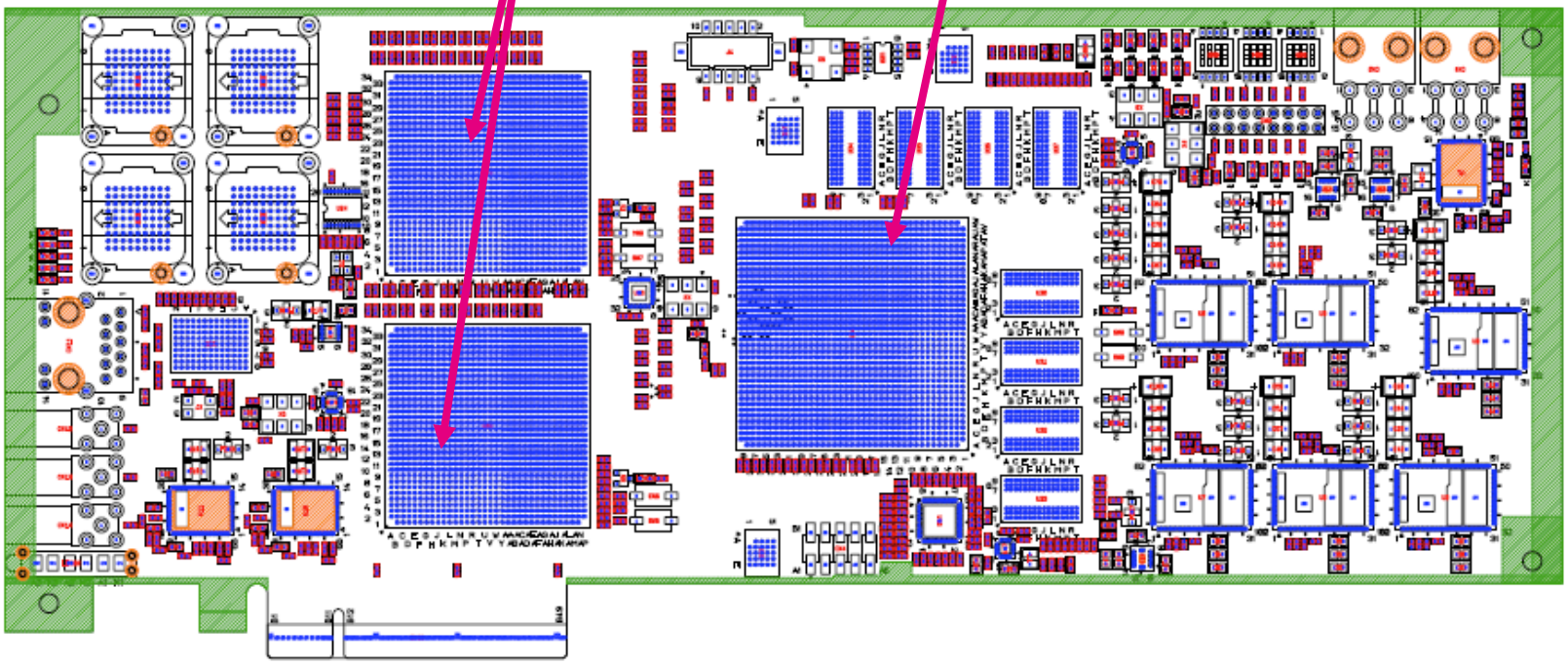
Data Framing Board outside Hutch

Calibration (Pixel wise)

Calibration (cross-talk)
Application specific computation

Optical cable

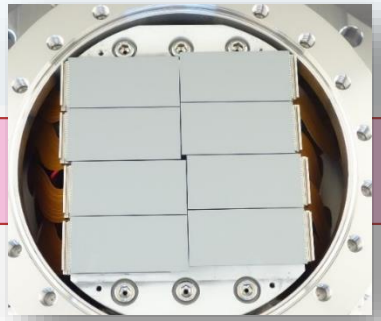
Max. 100 Gbps
(140 Gbps)



PCIe Gen3 x8
320 Gbps (40 GB/s)
Upgrade to Gen4 x8 planned

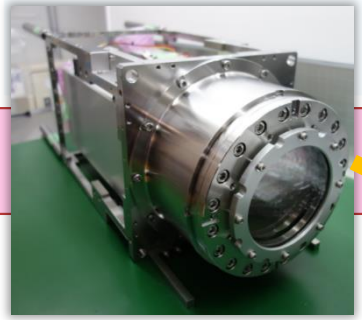
Edge-Heavy Computing Data Flow

Sensor



Detection

Camera Head



Digitization

Optical

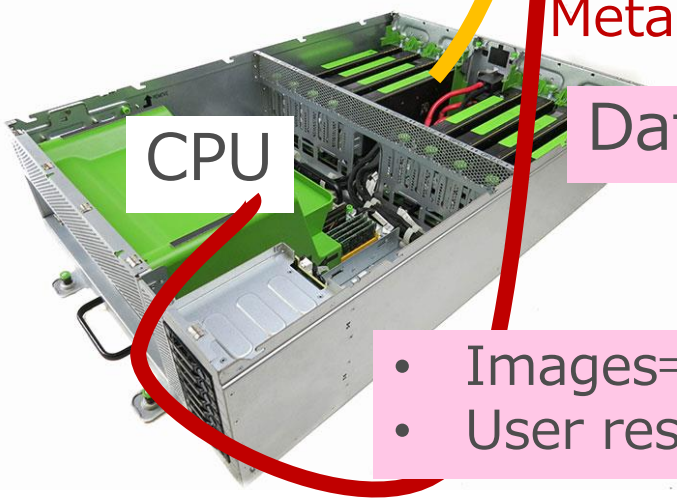
Data acquisition (DAQ)

Storage

Supercomputer

Analysis

Metal



Data framing board

- Images⇒pre-analyzed data
- User resource available

Evolution of Computing infrastructure from SPring-8, SACLA and SPring-8-II

- SPring-8: 1997-
  - Edge-only computing

~60 beamlines
Now, some beamline group owns computation resource

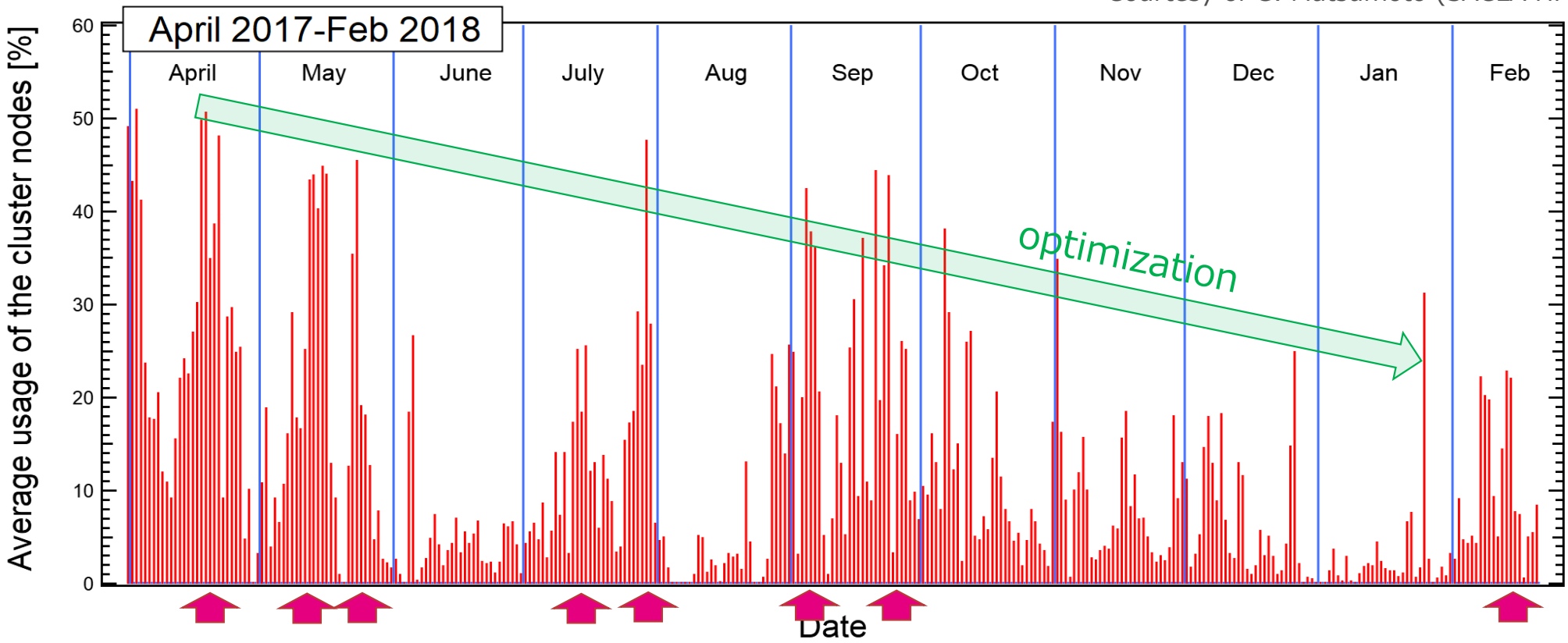
- SACLA: 2009-
  - Centralized Computing
  - off-site K Computer

Computation Infrastructure
125 TFLOPS
7-28 Pbyte tage archive
10,000 TFLOPS

- SPring-8-II: 202X-
  - Edge-Heavy, Center-Heavy & Cloud Computing

Average Node Usage of PC Cluster at SACLA

Courtesy of G. Matsumoto (SACLA HPC)



- HPCs are used
  - during experiments for monitoring the data quality
  - off-line analysis for publication data
- Peaks are during SFX experiments -> node availability is predictable.
- For off-line analysis at SPring-8-II, we will possibly implement the computation offloading to
  - academic HPCI & commercial cloud resources.

Evolution of Computing infrastructure from SPring-8, SACLA and SPring-8-II

- SPring-8: 1997-
  - Edge-only computing

~60 beamlines
Now, some beamline group owns computation resource

- SACLA: 2009-
  - Centralized Computing
  - off-site K Computer

Computation Infrastructure
125 TFLOPS
7-28 Pbyte tage archive
10,000 TFLOPS

- SPring-8-II: 202X-
  - Edge-Heavy, Center-Heavy & Cloud Computing