Belle II PXD DAQ: system, testing, FPGA implementation, performance in RUN 1 (Jennifer2 project report)

Jens Sören Lange, Johannes Bilk, Matthäus Krein, Dmytro Meleshko, Simon Reiter
(Justus-Liebig-Universität Giessen), and former work by alumni referenced hereafter

WORKSHOP ON FAST REALTIME SYSTEMS AND REALTIME MACHINE LEARNING
Justus–Liebig–Universität Giessen
08.04.–11.04.2024

Jennifer2 workpackage 5.2
Online data acquisition and remote controls

From Jennifer2 Kick-Off Workshop, 09/2019, Vienna

- new hardware technologies for high bandwidth data transfer
  - optical technologies (from 6 Gbps to 16 Gbps) THIS TALK
  - From 1G to 10G ethernet THIS TALK

- intelligent realtime algorithms for online data reduction
  - Belle II – background rejection on FPGAs THIS TALK (ROI selection)
  - HyperK – TALK by Benjamin Richards

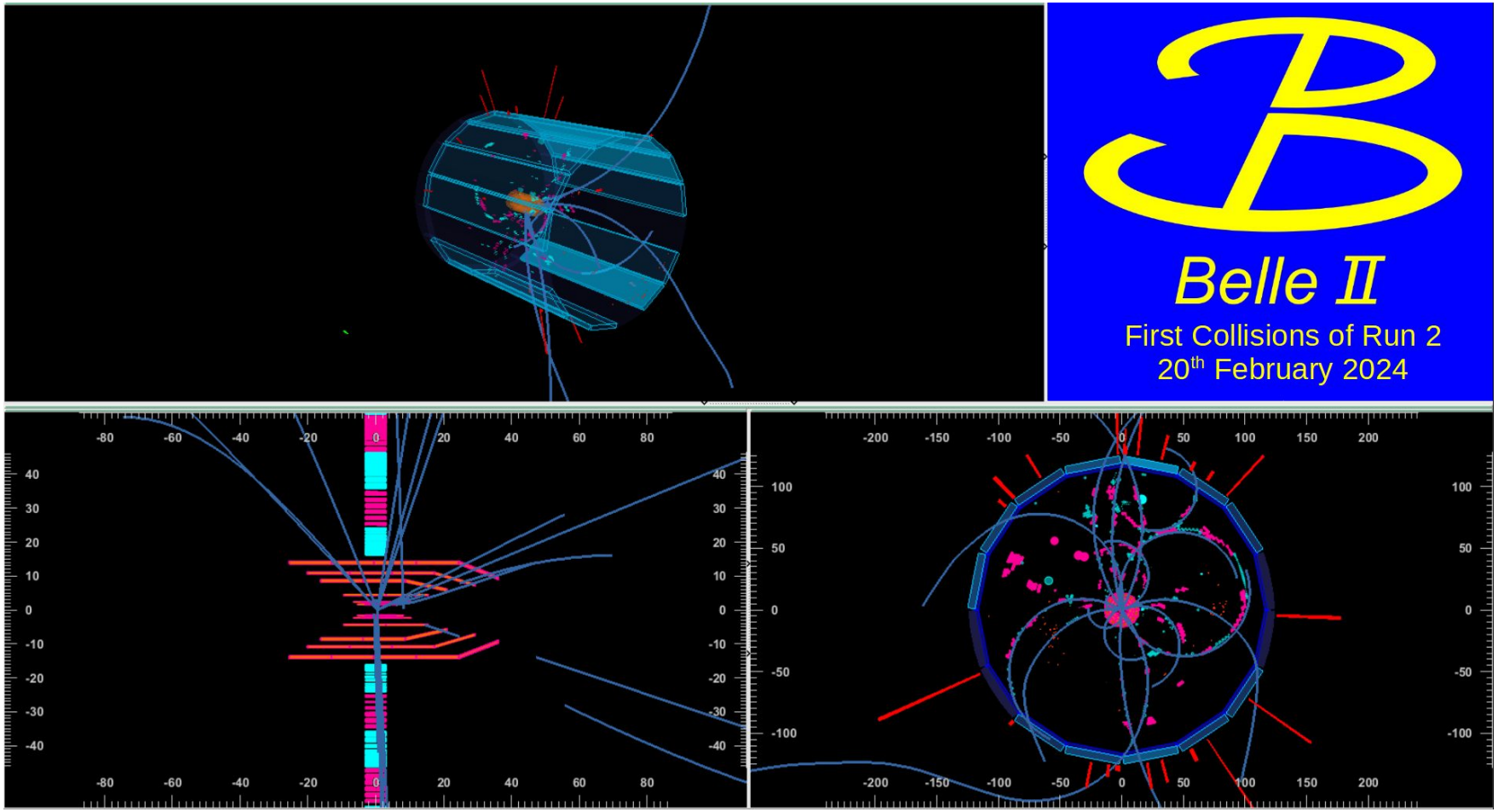
- novel programming and DAQ software techniques
  - parallelisation on both FPGAs or GPUs THIS TALK
  - methods of artificial intelligence e.g. for trigger decisions TALKS April 10 &11
  - monitoring, fault tolerance, dynamic routing and remote control THIS TALK

Belle II Collaboration
6 Tage ·

We are posting live from the Belle II control room, where after 18 months of upgrades and maintenance to both the Belle II detector and the #SuperKEKB accelerator, we have just recorded our first collisions of Run 2, at 22:12 JST.

Run 2 is our second data taking period, after Run 1, which lasted from 2019-2022. This is a very exciting time for Belle II, as we embark on this new adventure, collecting and analysing a much larger dataset, with an early goal of surpassing the size of the dataset recorded by our predecessor, the Belle experiment. We look forward to sharing more news of Run 2 in the days, weeks, and months to come, stay tuned.

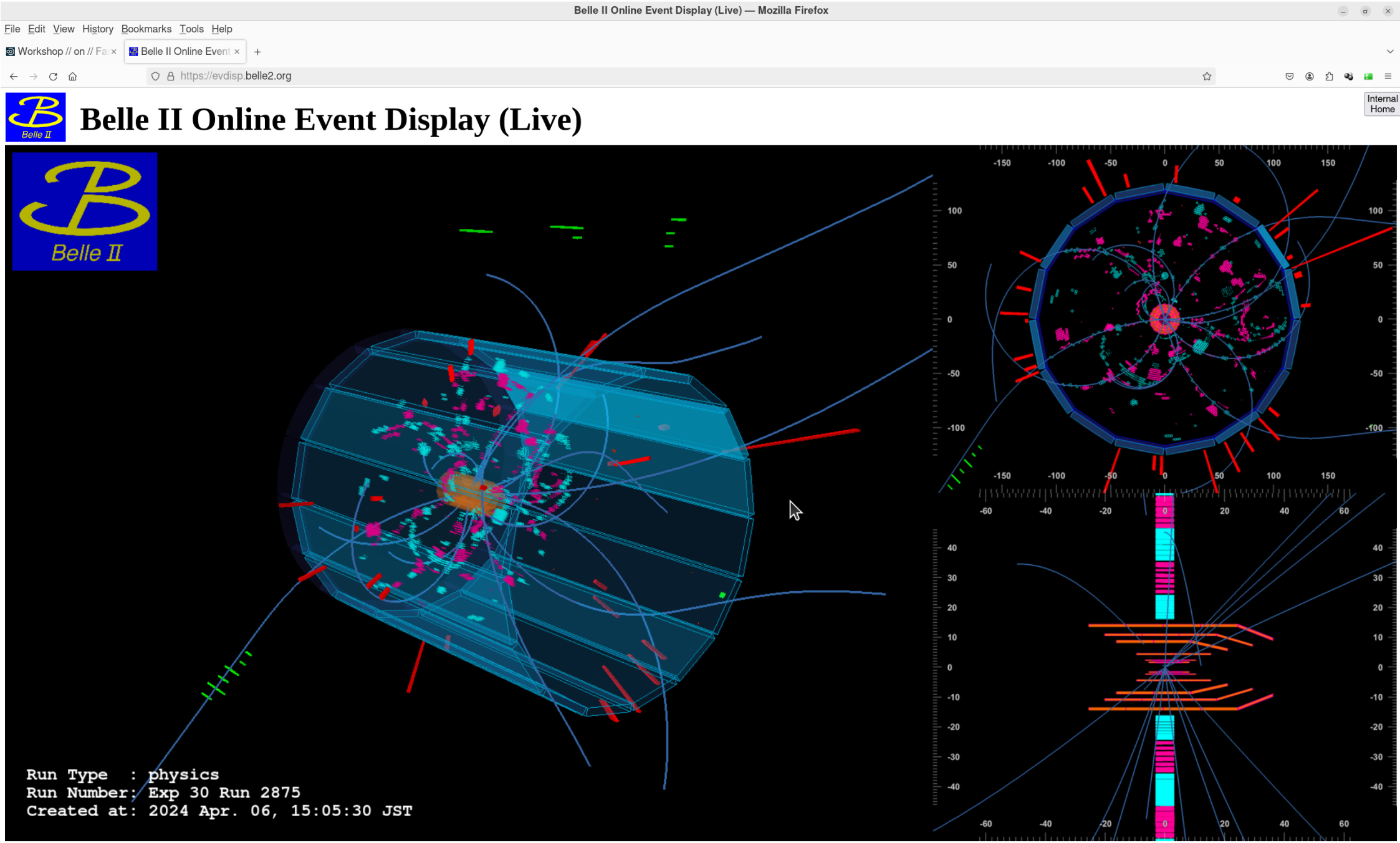
Belle II experiment, e+e– collisions



First collisions after long shutdown (~18 months)

New pixel detector (now 8 Million pixels, factor 2 more FPGA hardware)

Live event display is enabled for run 2

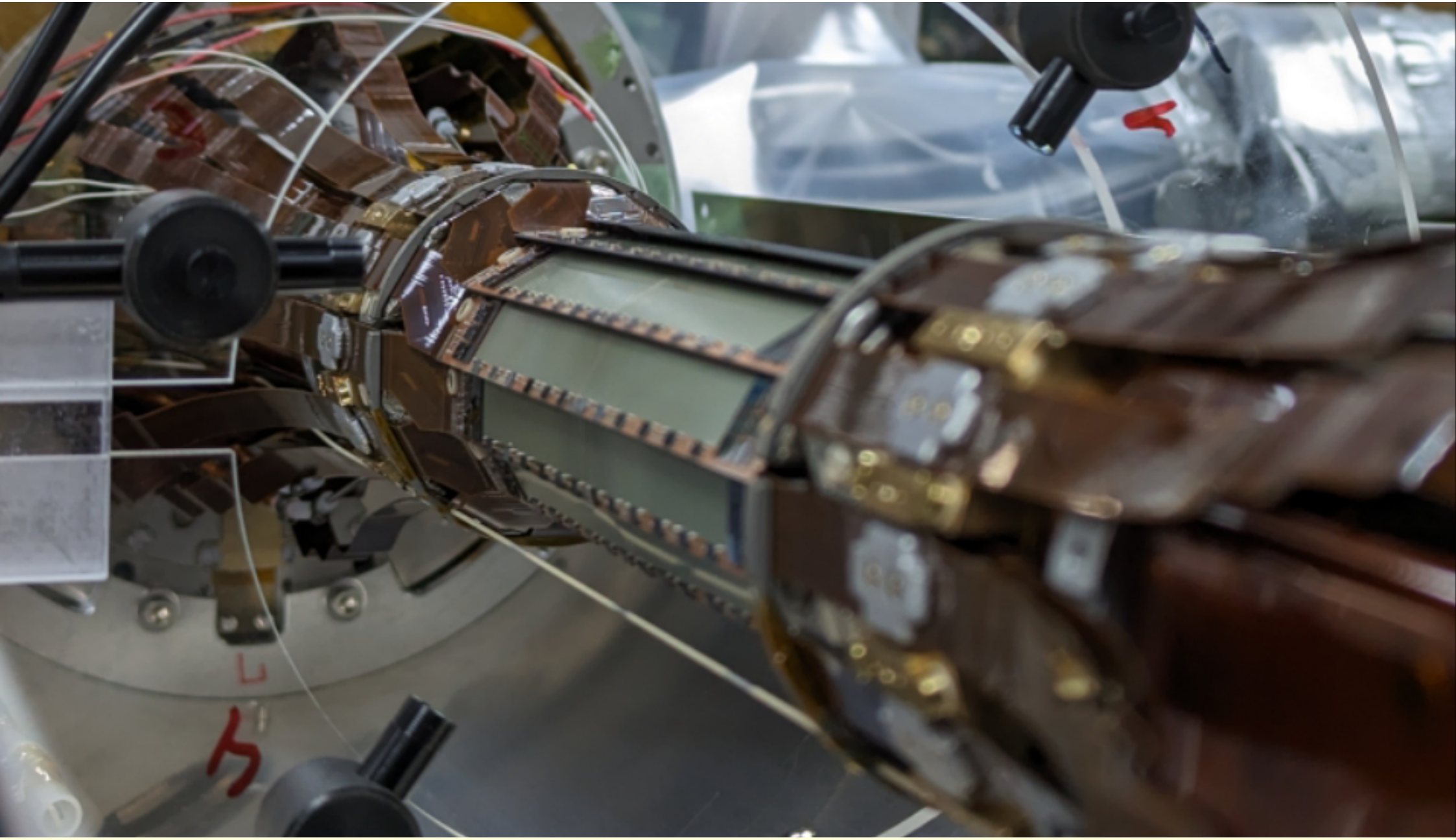


Belle II Silicon Pixel Detector

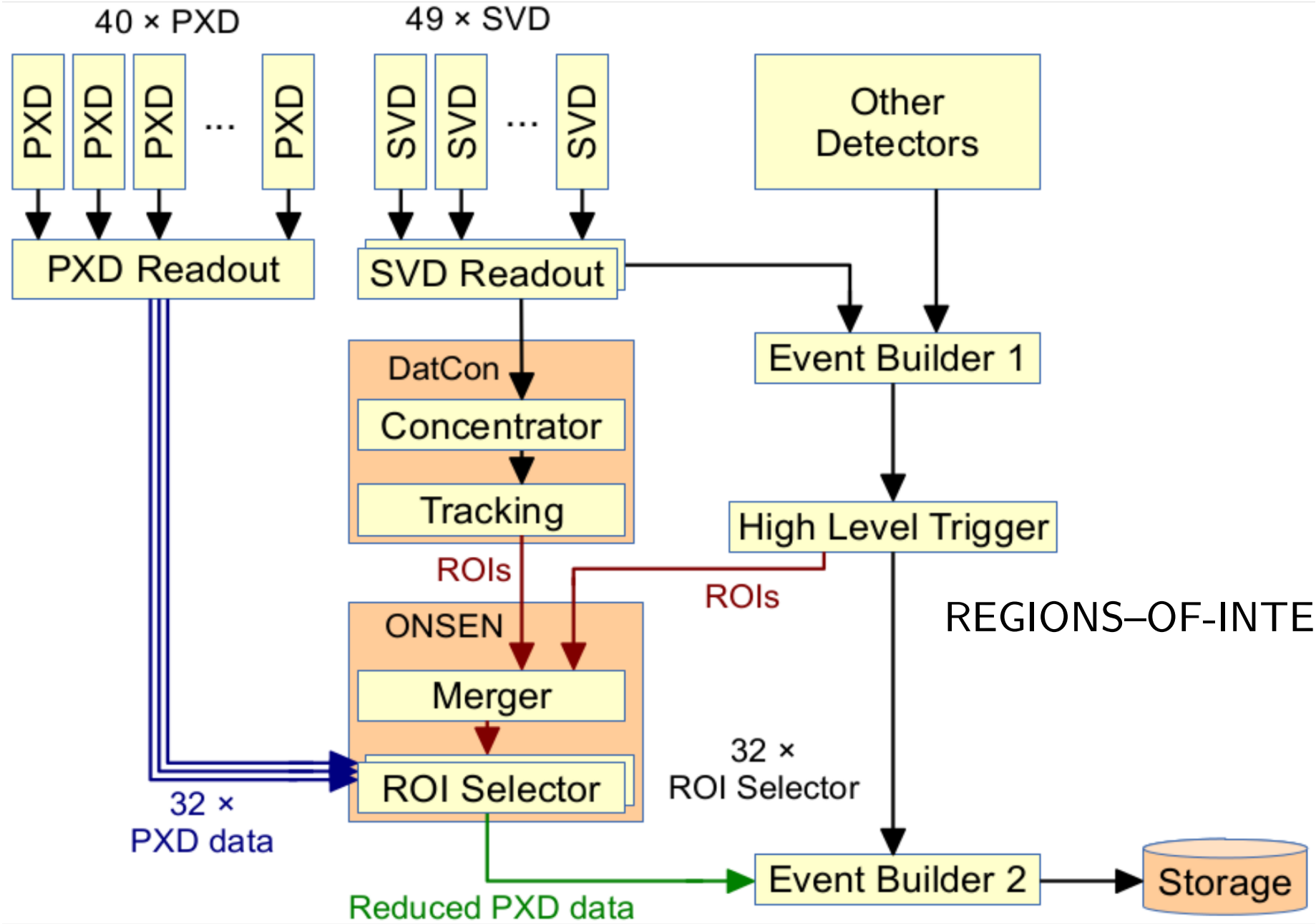
Univ. Bonn, DESY, Univ. Giessen, Univ. Göttingen, Univ. Heidelberg, Univ. Mainz,
KIT Karlsruhe, HLL München, MPI München, LMU München, TU München

An MSCA-RISE project funded by European Union under grant n.644294

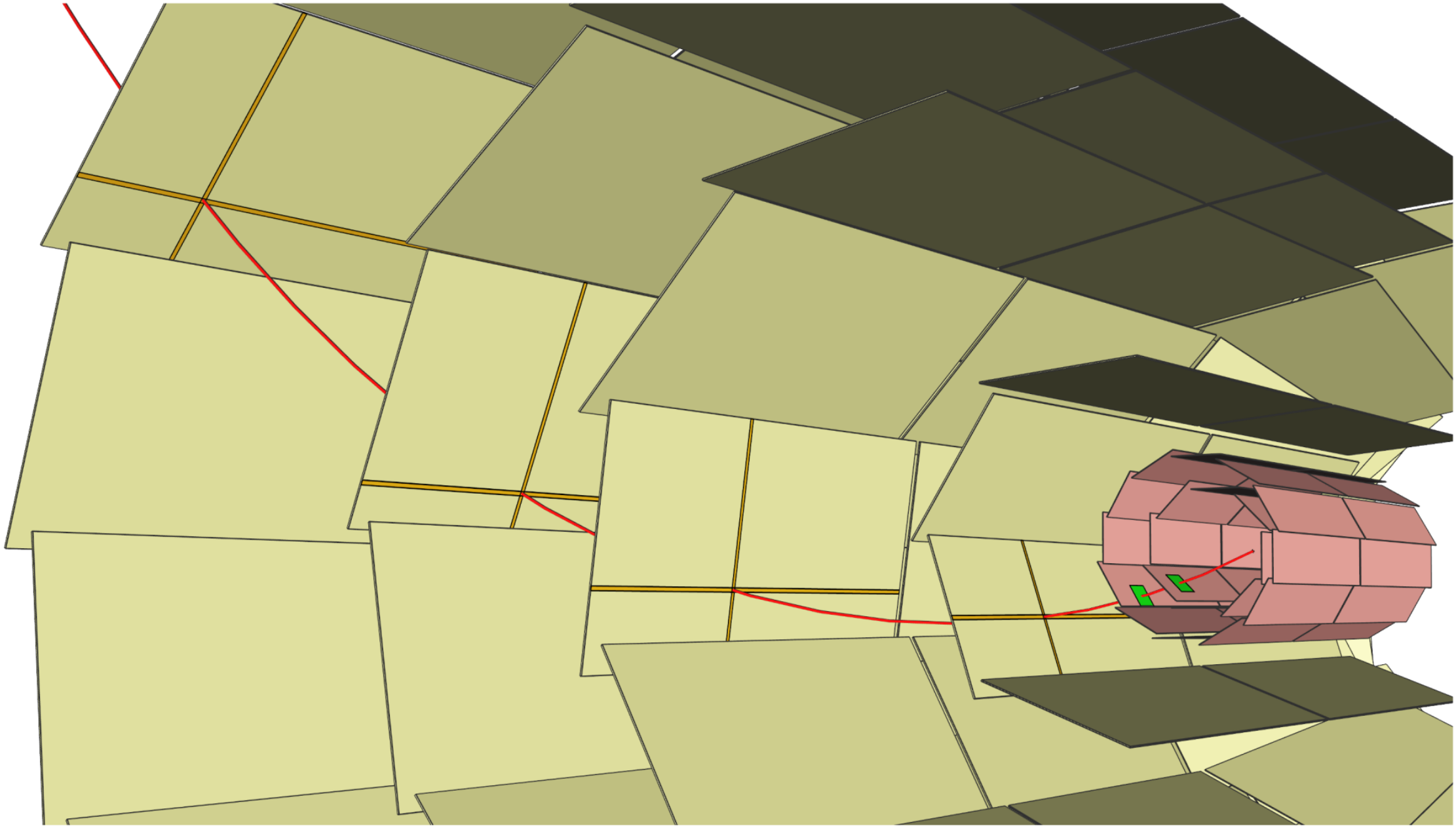
Belle II PXD2, 2 layers, 40 modules, 8 M pixels (installed 2023)



PXD data acquisition and data reduction system



Region–of-Interest (ROI selection)
(HLT tracks, extrapolated to PXD)



PXD DAQ: requirements and solutions

- PXD frame rate (fixed) 50 kHz,
  L1 trigger rate up to 30 kHz (factor 75 more than Belle)
- PXD generates about 10x more data than rest of Belle II
  up to 20 GB/s at 3% occupancy
  solution: optical links, 6 Gbps
- Buffer PXD data until HLT decision, up to 5 seconds
  solution: 4 GB RAM on FPGA board
  buffer management (based upon 6 bits of trigger number)
- Apply ROI selection, ROIs are calculated by online tracking on HLT
  solution:logic on FPGA
  - ROI distribution (ROIs only sent to the one FPGA, which processes related data)
  - coordinate transformation
  - parallised up to 32 ROIs per module, compared to emulator,
    factor $\leq$355 faster than single core PC (Intel i7, 3.4 Ghz)
  - reduce data by factor $>$10
- Send out data to event builder by 32 x Gigabit Ethernet links
- Stable data taking about 10 months per year, 24/7 operation

  (FPGA firmware finds frontend data format errors, avoid crash)

PXD DAQ AMC Card (IHEP Beijing, Giessen)

ONSEN AMC card
v4.0 (final)
Virtex-5 FX70T
2 optical links (6.25 Gbps)
GbE

2 configurations and firmwares

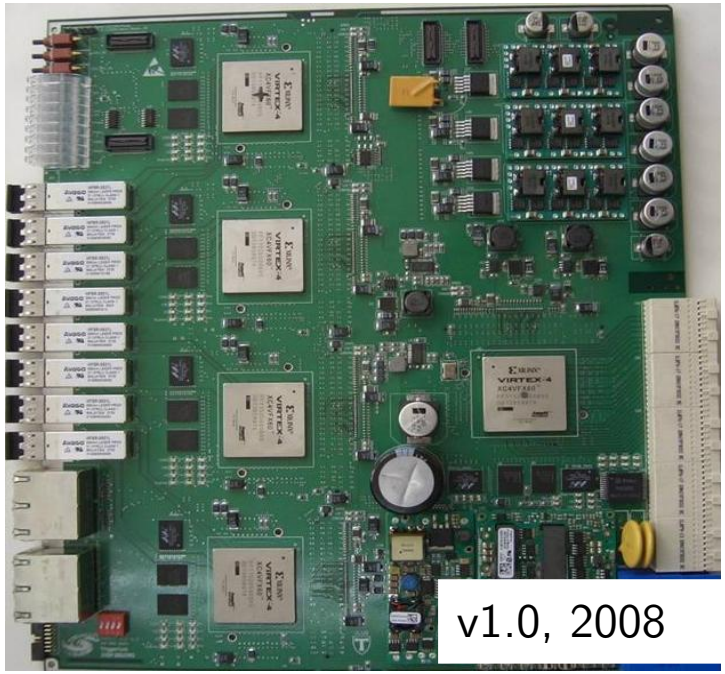
- selector AMC
- merger AMC

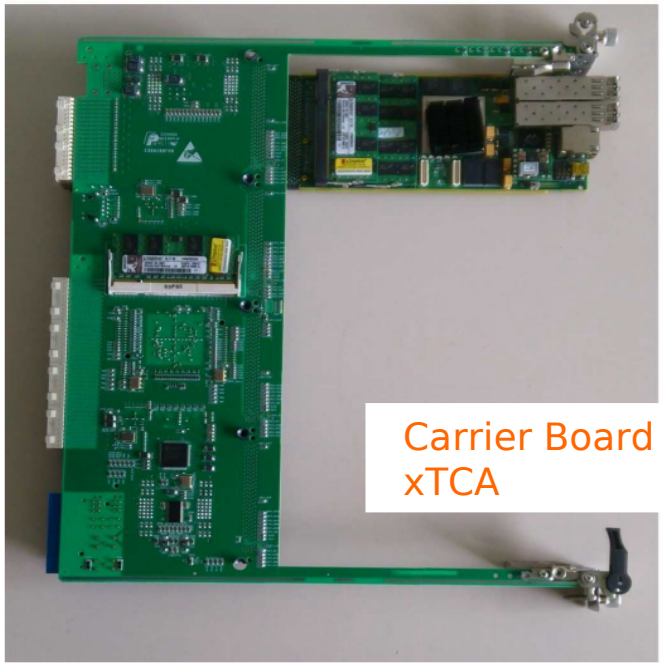
slow control / monitoring:
IPMI add-on boards (Mainz)

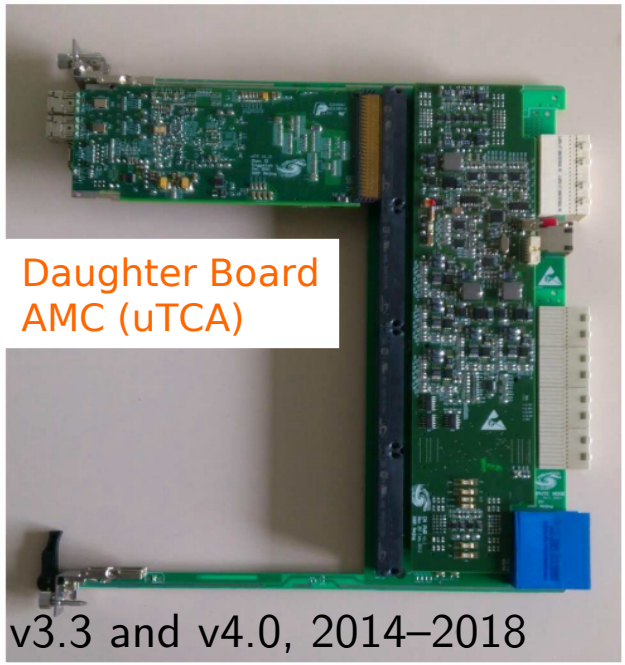
PXD DAQ carrier card (IHEP Beijing, Giessen)

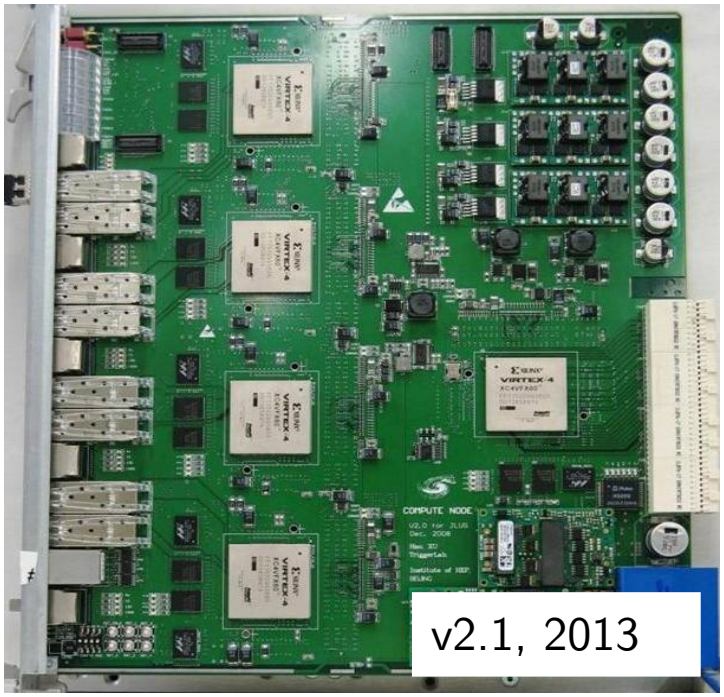
ONSEN xTCA carrier card
v3.3 (final)
Virtex-4 FX60 (switcher to ATCA backplane)
GbE
add-on: RTM board power supply board

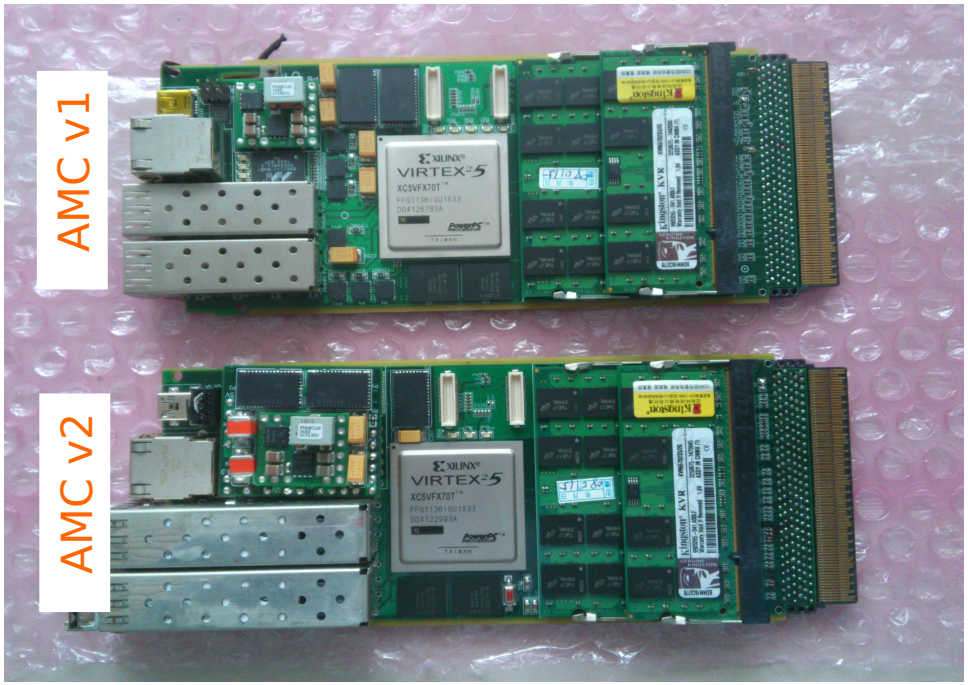
Belle II PXD readout hardware (IHEP Beijing, Giessen)



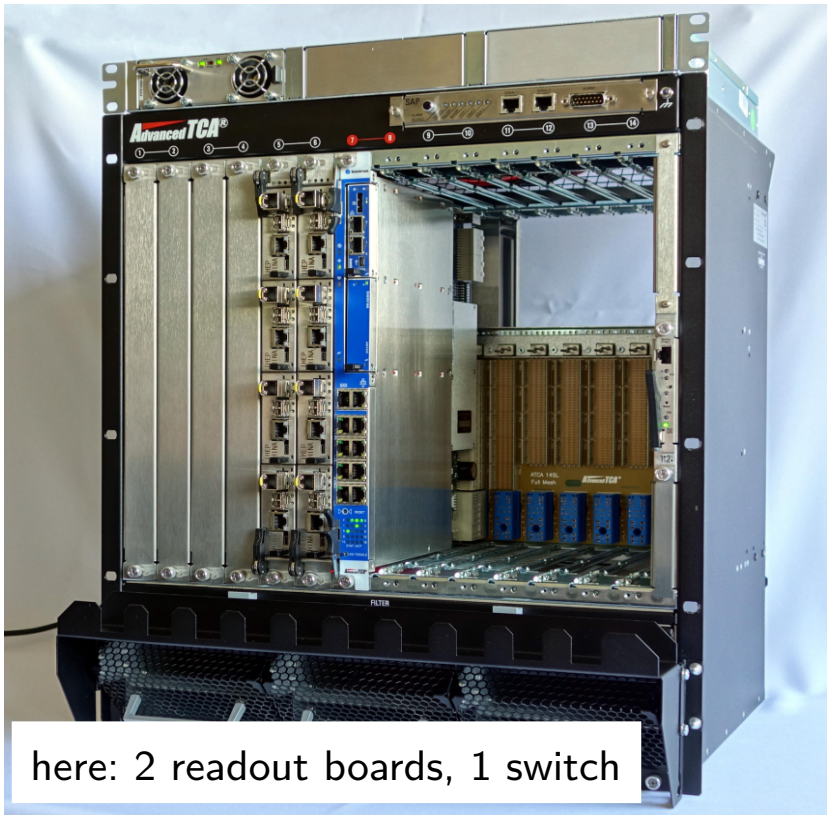




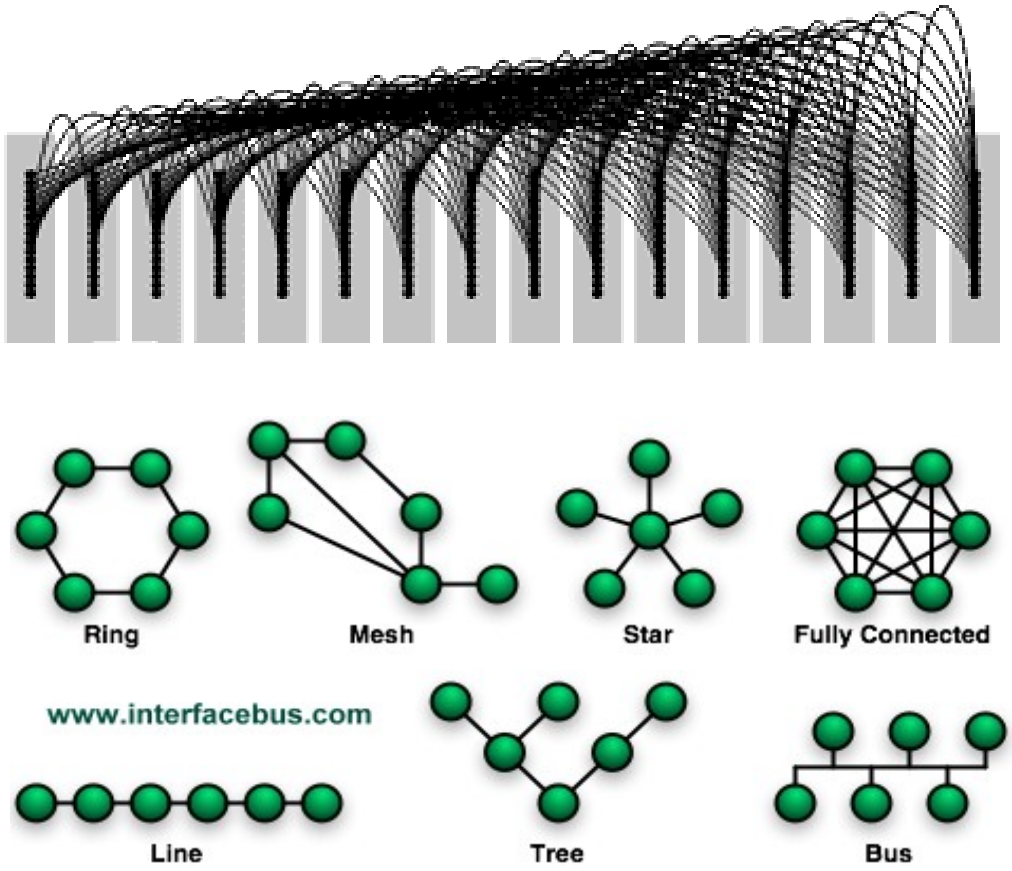




ATCA (Advanced Telecommunications Architecture)



here: 2 readout boards, 1 switch



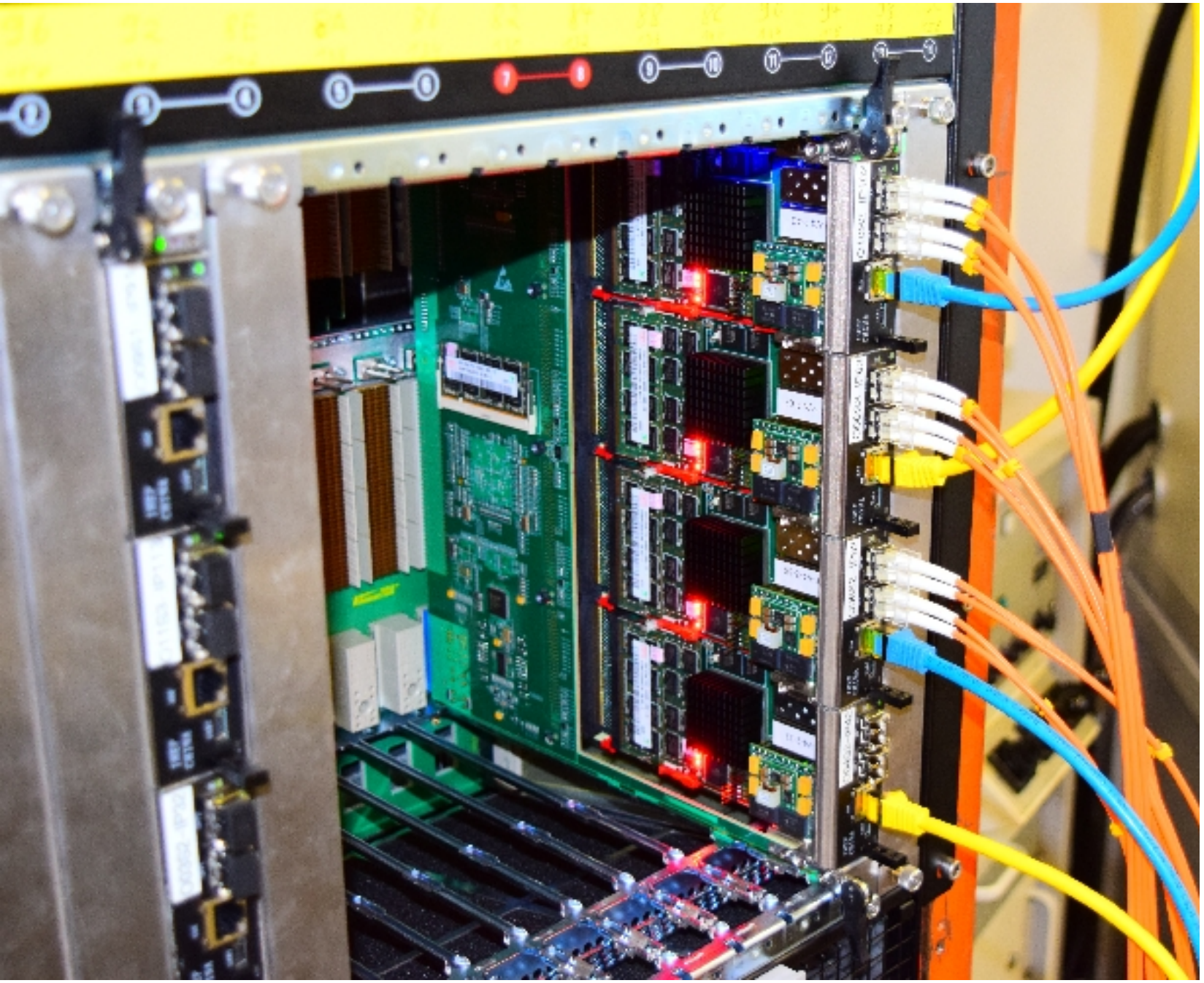
Full-mesh backplane

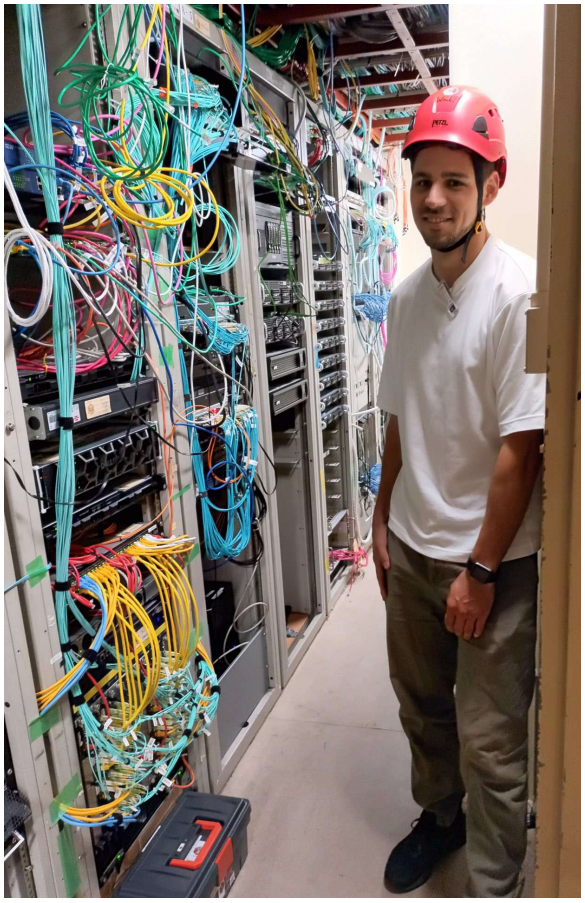
No bus system (different from CAMAC, VME, FASTBUS)

14 slots, all point-to-point connections

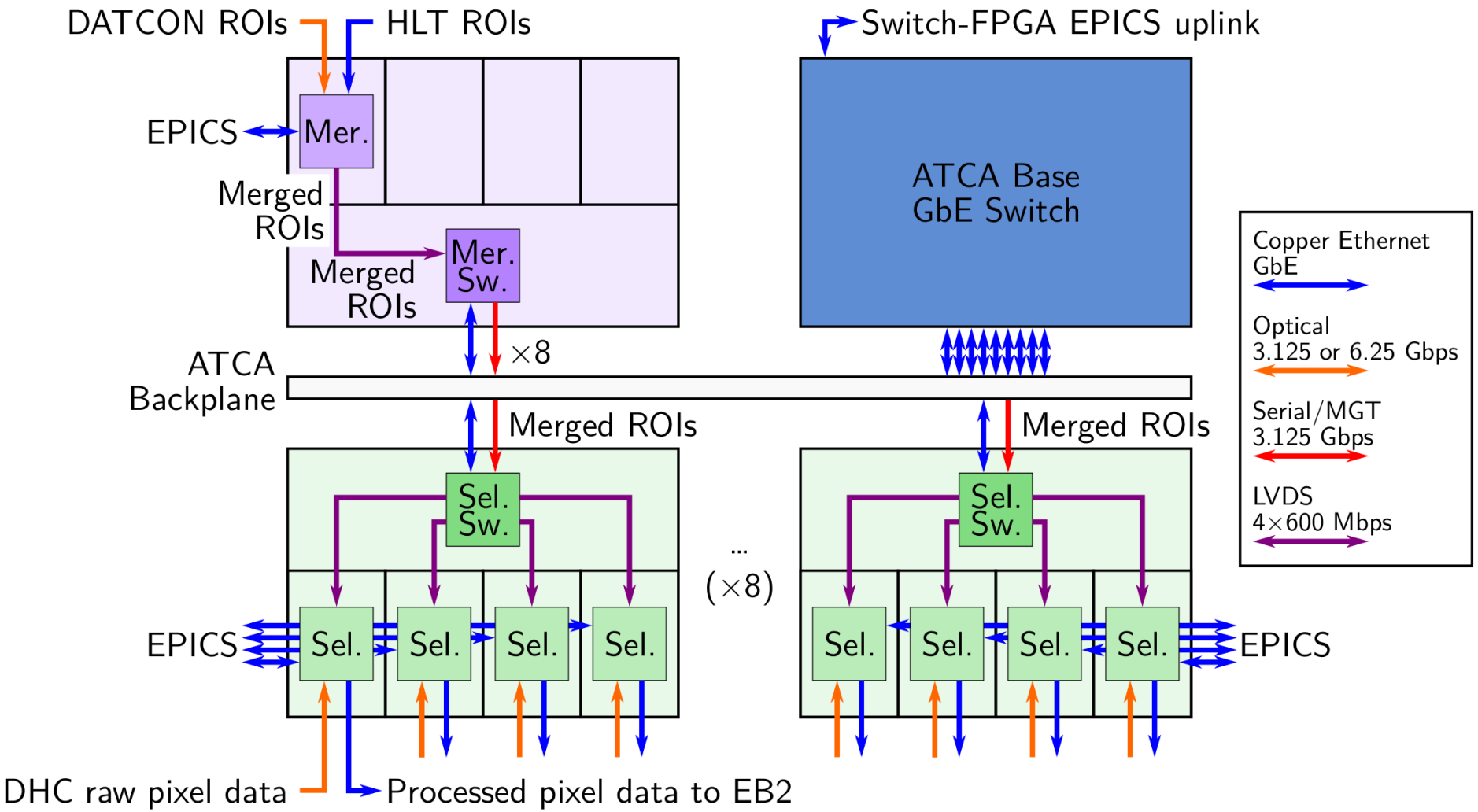
in our case 14 x 13 x 6 Gbps = 1092 Gbps = 109 Gbyte/s
(factor 50 safety margin for Belle II PXD, using 8B/10B encoding)

1 Carrier board with 4 AMC boards in ATCA shelf





PXD DAQ: data links



GbE implementation on FPGA (SiTCP) by BeeBeans Ltd.

≥100 MB/s reached per link, support of “pause frames” (resubmission)

ONSEN AMC board firmware: merger

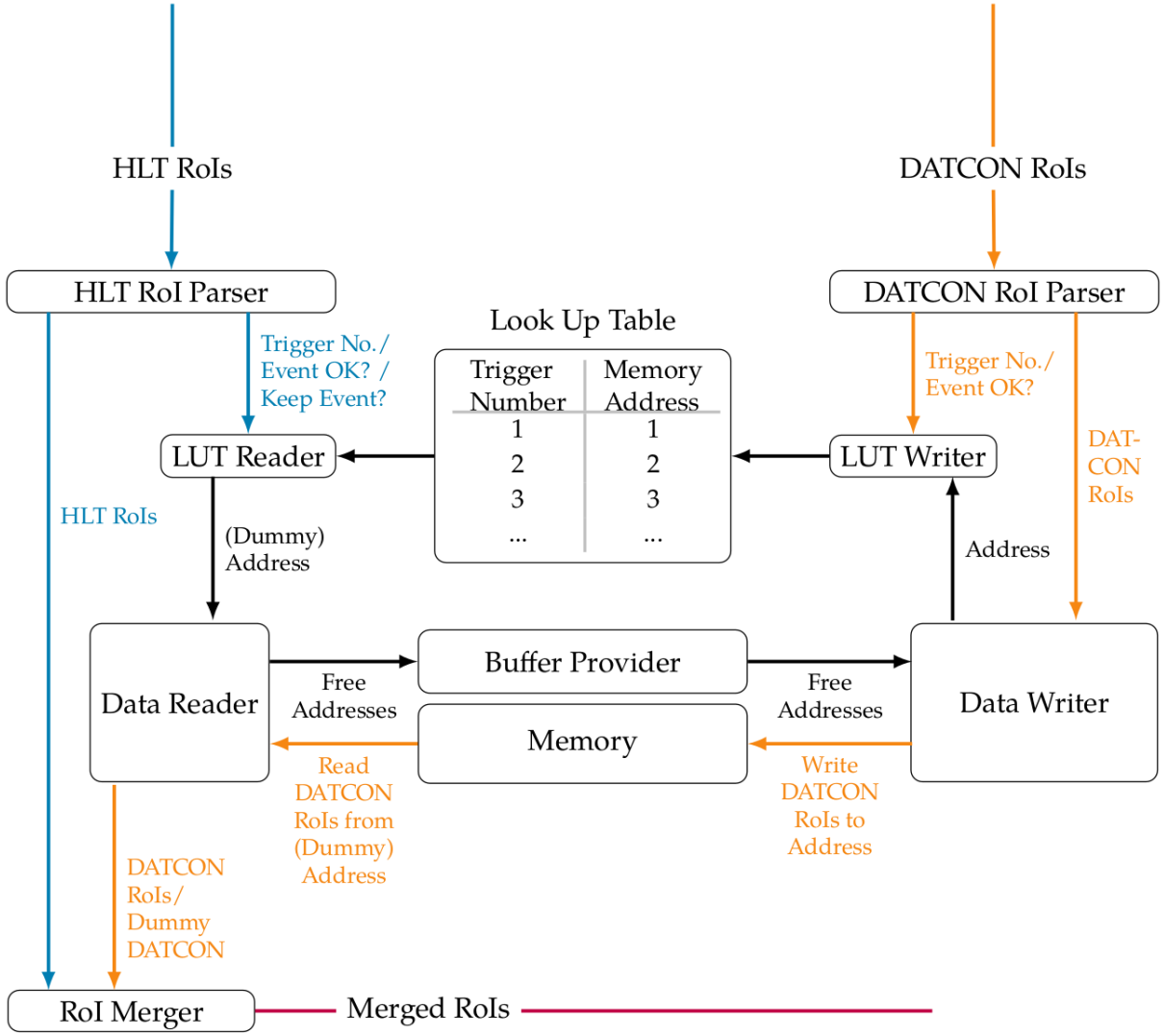


Figure 4.5: Simplified overview of data flow in ONSEN Merger.

Dennis Getzkow, Ph. D. thesis, 2023

ONSEN AMC board firmware: selector

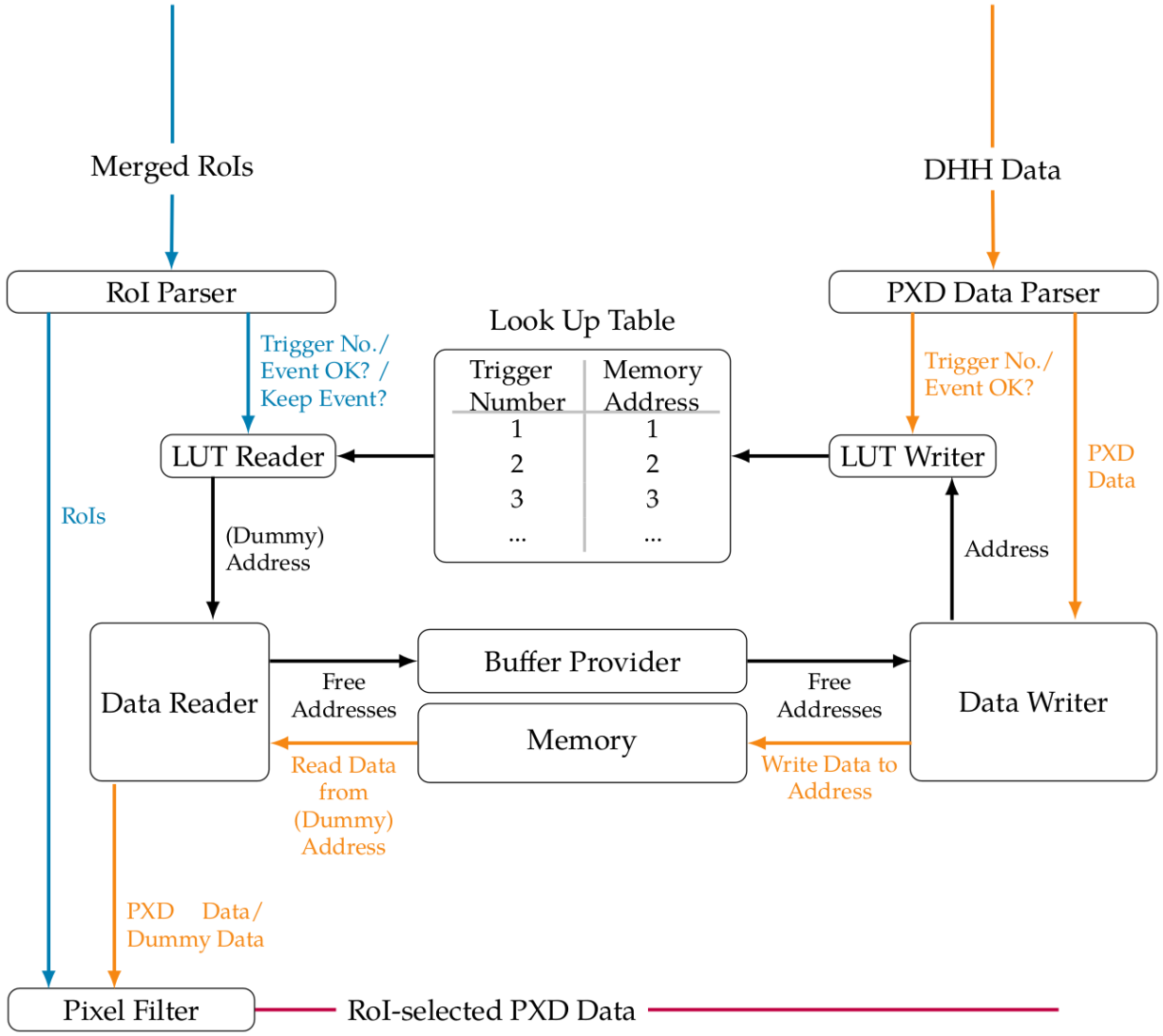
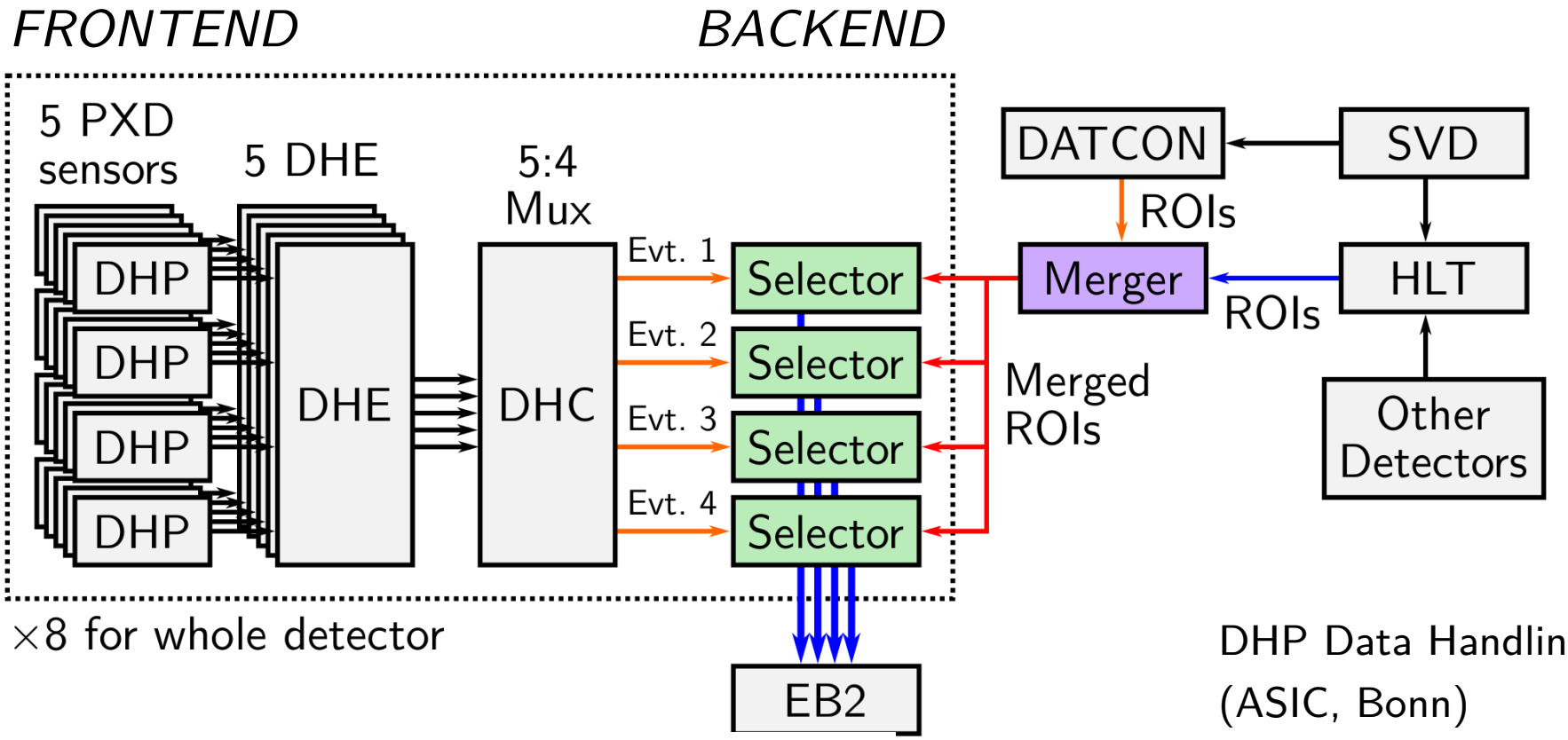
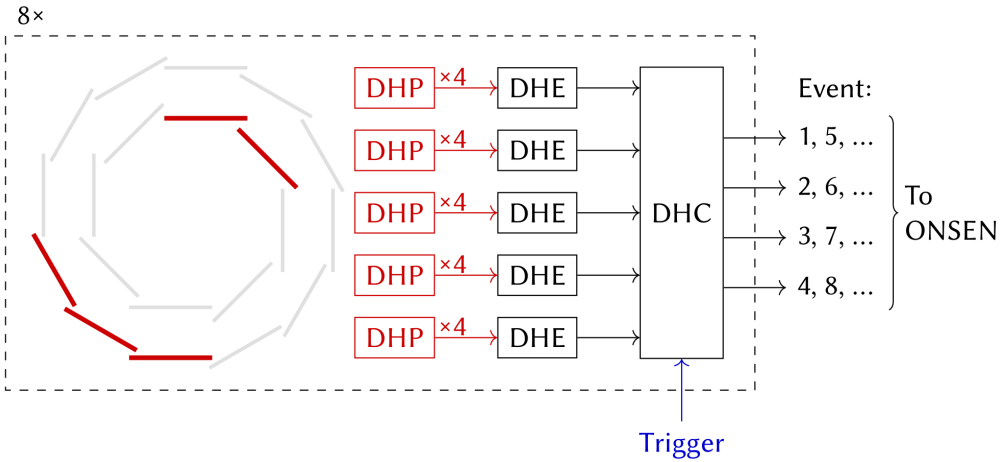


Figure 4.6: Simplified data flow in ONSEN Selector.

Dennis Getzkow, Ph. D. thesis, 2023

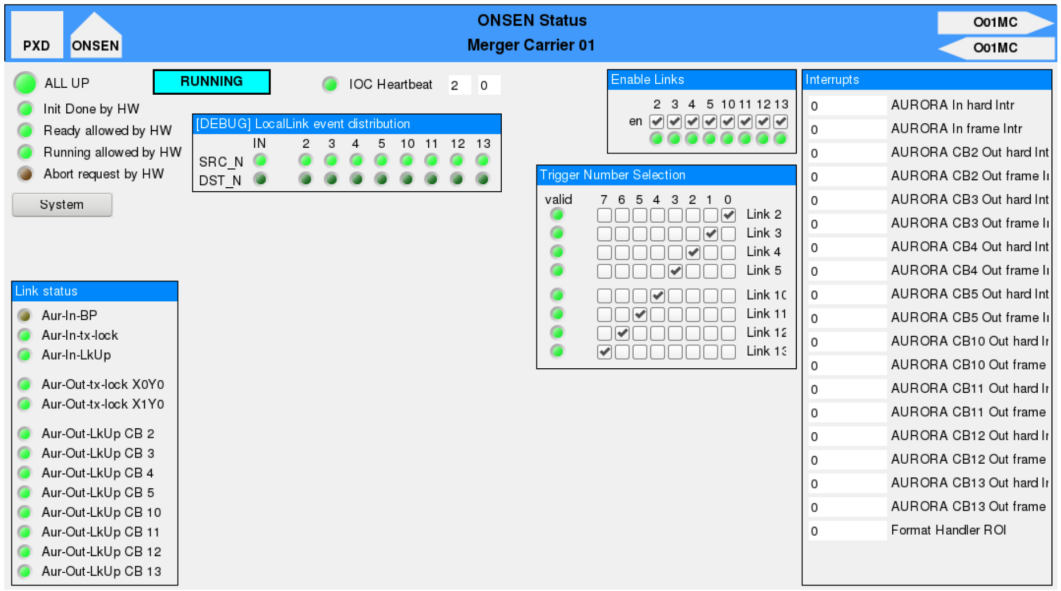
PXD DAQ: multiplexing and load balancing



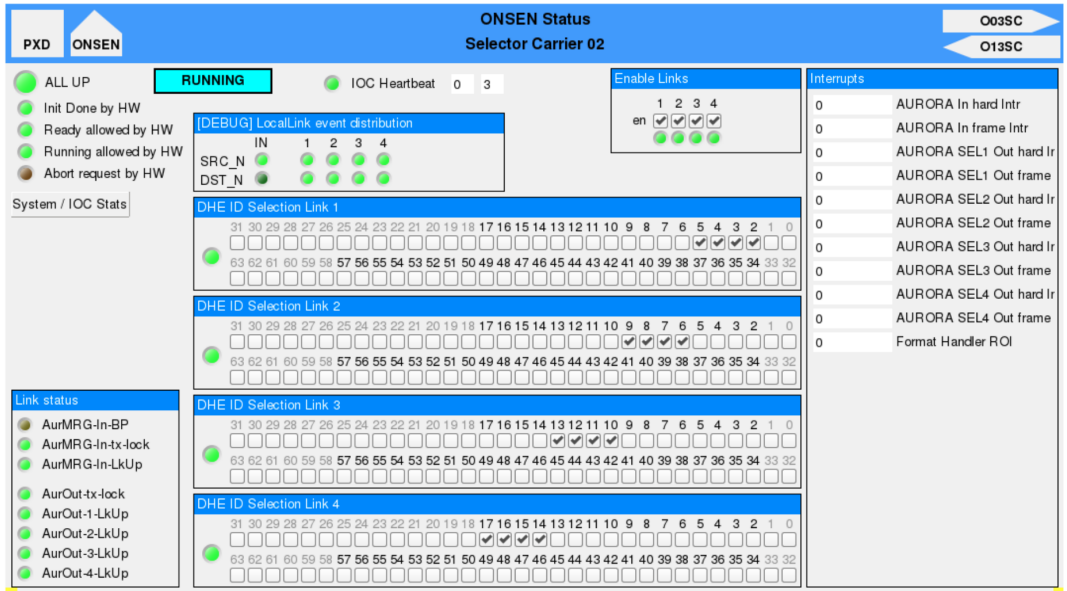


DHP Data Handling Processor
(ASIC, Bonn)
DHC Data Handling Concentrator
DHE Data Handling Engine
(all FPGA, TU Munich)
ONSEN Online Selector Nodes
(FPGA, Giessen & IHEP Beijing)

Load balancing: implementation, configurable



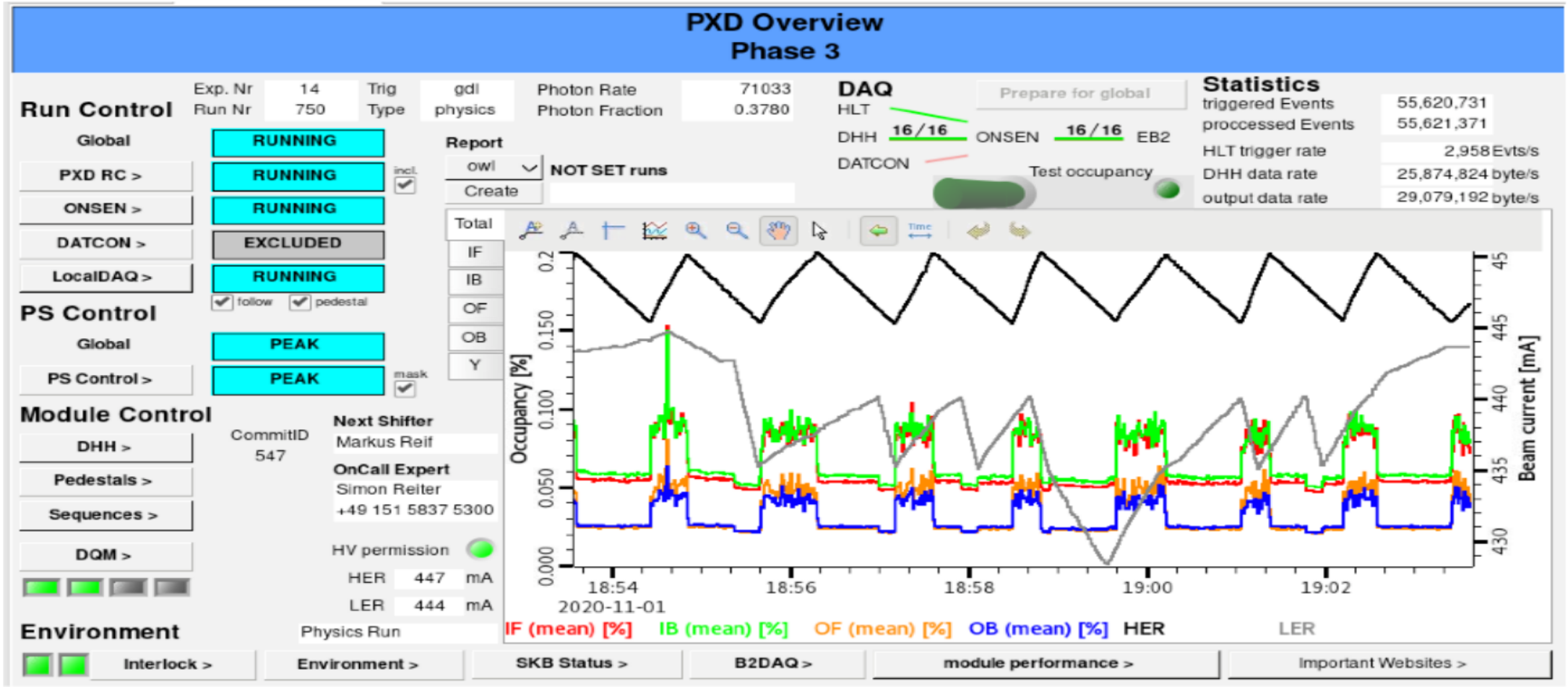
(A) CSS user interface of the ONSEN Merger CNCB. The check-boxes on the top right control the links to the Selector CNCBs (RoI Fork). The check-boxes below control the trigger number selection.



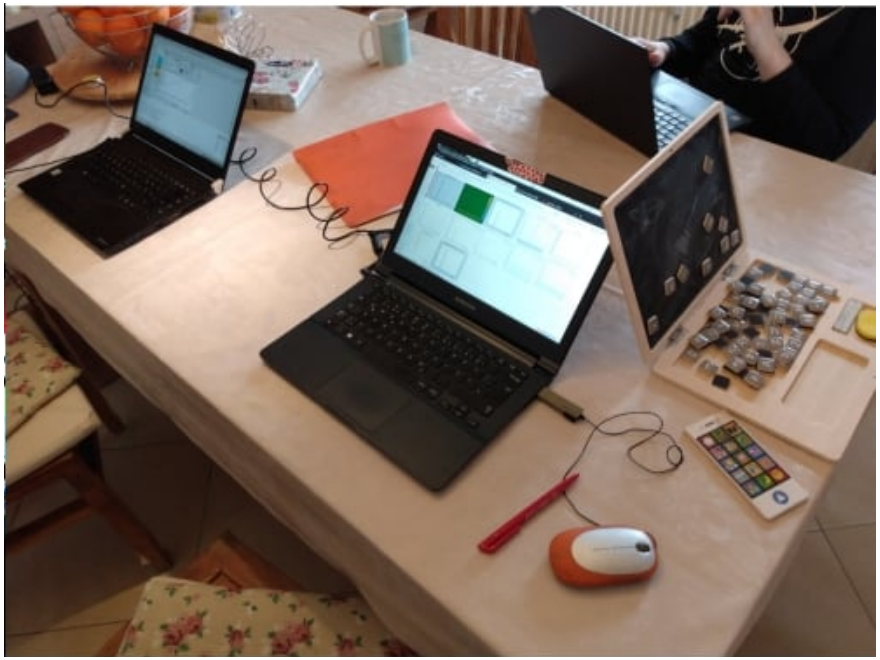
(B) CSS user interface of the ONSEN Selector CNCB. The check-boxes on the top right control the links to the Selector AMCs (RoI Fork). The check-boxes below control the DHE selection.

Dennis Getzkow, Ph. D. thesis, JLU Giessen, 2023

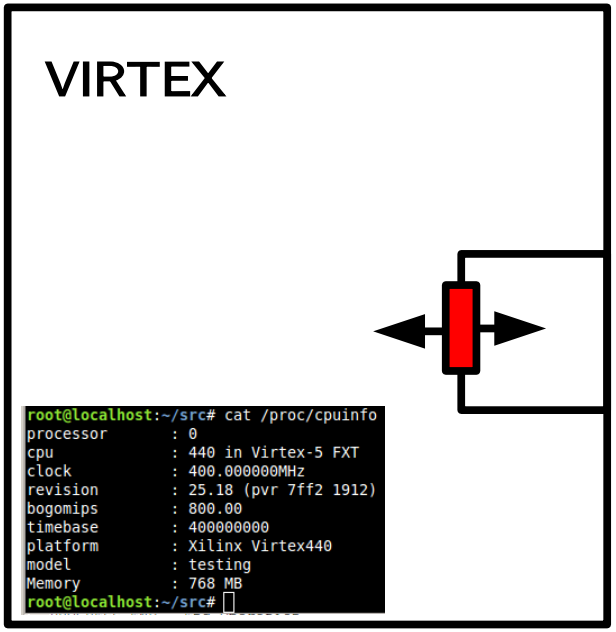
PXD DAQ: monitoring (shift interface and GUI)



PXD DAQ: monitoring



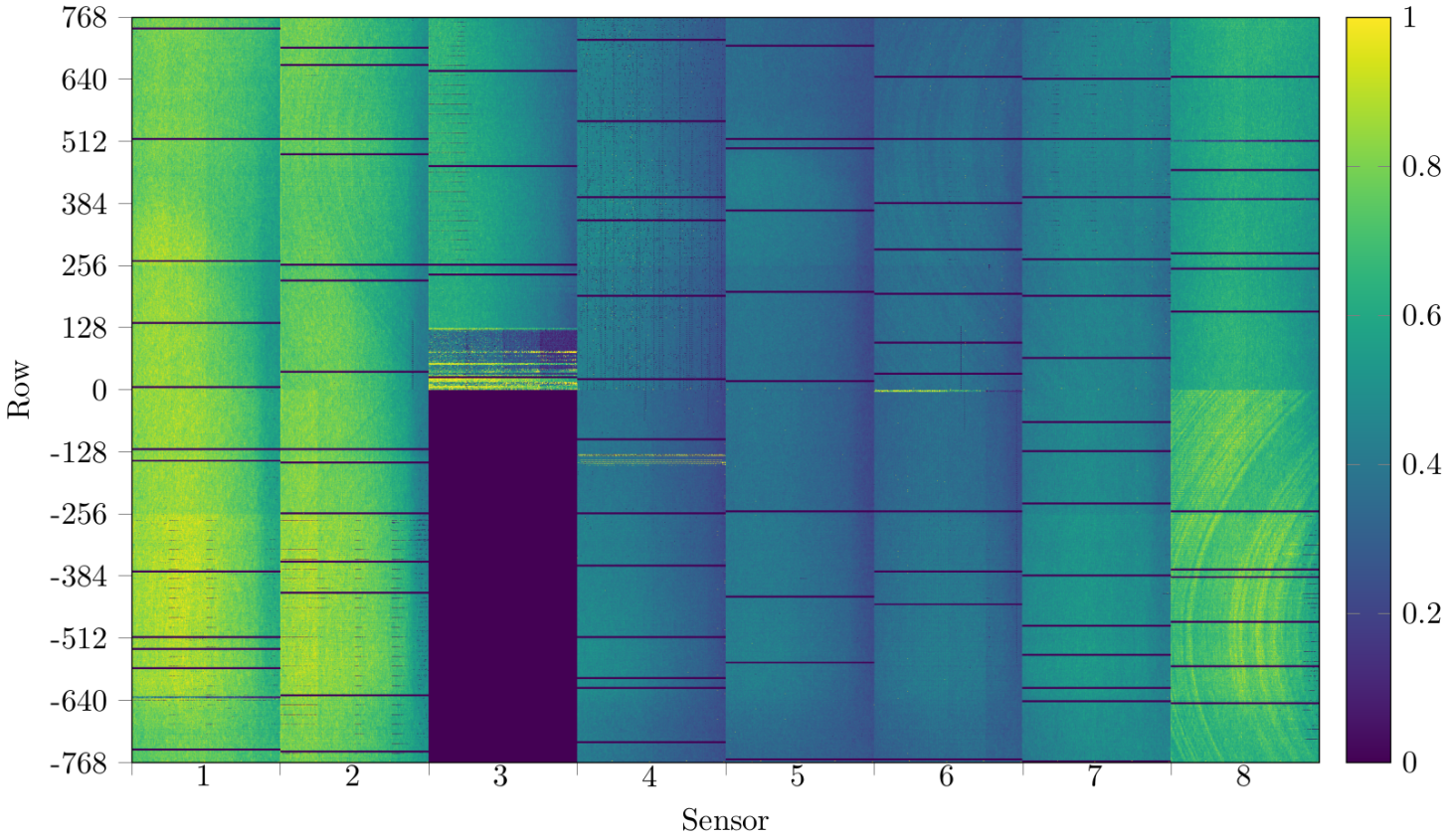
- In a Virtex, part of the FPGA is a hard-core processor (PowerPC 440) running Linux
- Registers read/write from FPGA side and from Linux side (''dual-ported''), with an EPICS input/output controller (IOC) running here and process variables then displayed by Control System Studio (CSS), Eclipse-based GUI
- PXD shifts are remote (ever since 2020, due to Corona)



RUN1 OPERATION RESULTS

Operation results of PXD DAQ

- RUN 1 from 10/2019 to 06/2022 (followed by long shutdown)
- 300 TB zero-suppressed data recorded, 400+ Million events recorded
- Triggerrate up to 8 kHz (design 30 kHz, short tests run up to 35 kHz)
- Hitmap of the PXD (here ~3.2 Mill. Pixels)
  in a single data run in 2022 (~2 hours)



Operation results of PXD DAQ

- RUN 1 from 10/2019 to 06/2022
- Belle II collected data 428 fb^{-1}
  362 fb^{-1} for B mesons
  (~50% of Belle, ~85% of BaBar)
  1 fb^{-1} is about 1.1 Mill. $B\overline{B}$ pairs
- Peak luminosity reached
  4.7×10^{34} cm^{-2} s^{-1}
  - Factor 2+ higher than previous world record by KEKB
  - Factor 4+ higher than KEKB design luminosity
- 89.5% data taking efficiency during pandemic situation
- 300 TB zero-suppressed data recorded
- 400+ Million events recorded
- Triggerrate up to 8 kHz (design 30 kHz, short tests run up to 35 kHz)

PXD DAQ data rates

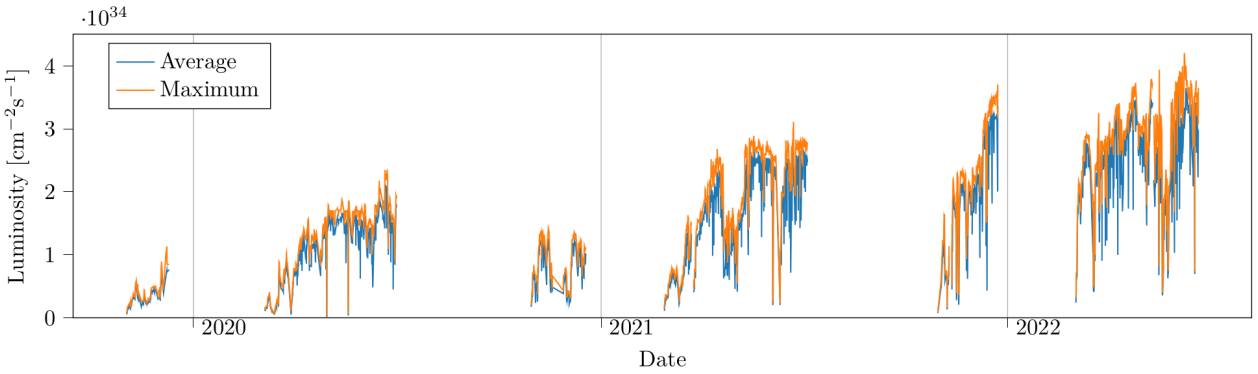


Figure 6.1: Luminosity of each longer run in Phase 3 detected by ECL. In June 2022 the instantaneous luminosity record of 4.71×10^{34} cm^{-2} s^{-1} was reached [23].

	Number of runs	Run time [h]
total	11 695	9 342
physics	9 453	7 454
cosmic	901	1 172

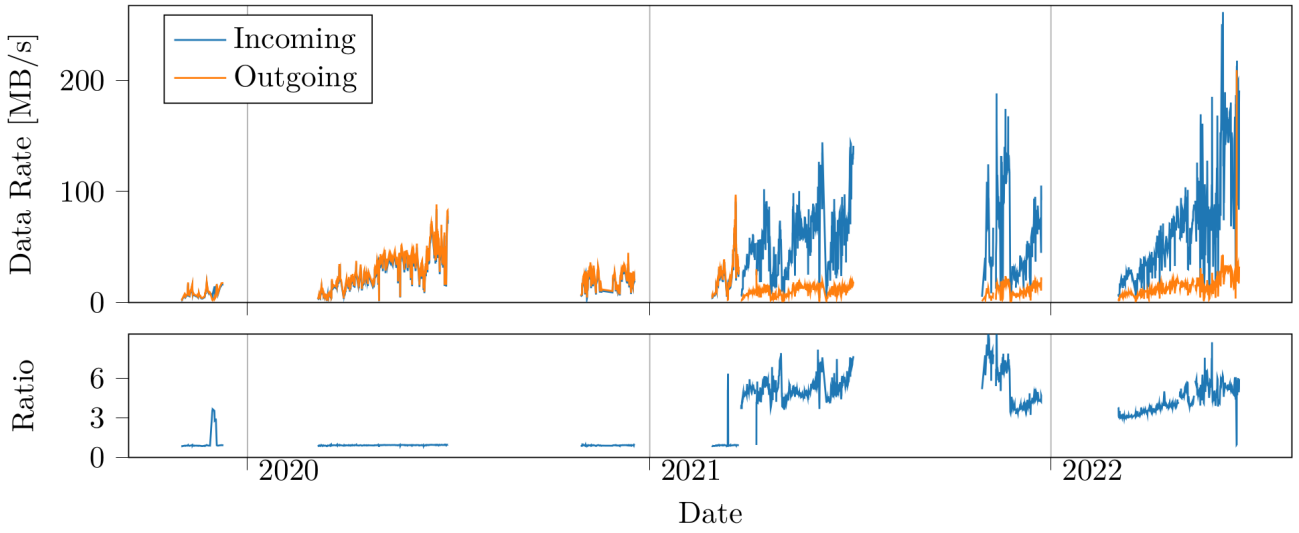


Figure 6.2: Comparison between input and output data rates of ONSEN averaged for each physics run. The highest input data rate was reached near the end at peak luminosity in 2022 with over 250 MB/s. In the lower part the fraction is shown. During 2021 the *event filtering* was enabled, which results in a permanent increase of the ratio.

Simon Reiter, Ph. D. thesis, 2022

PXD occupancy

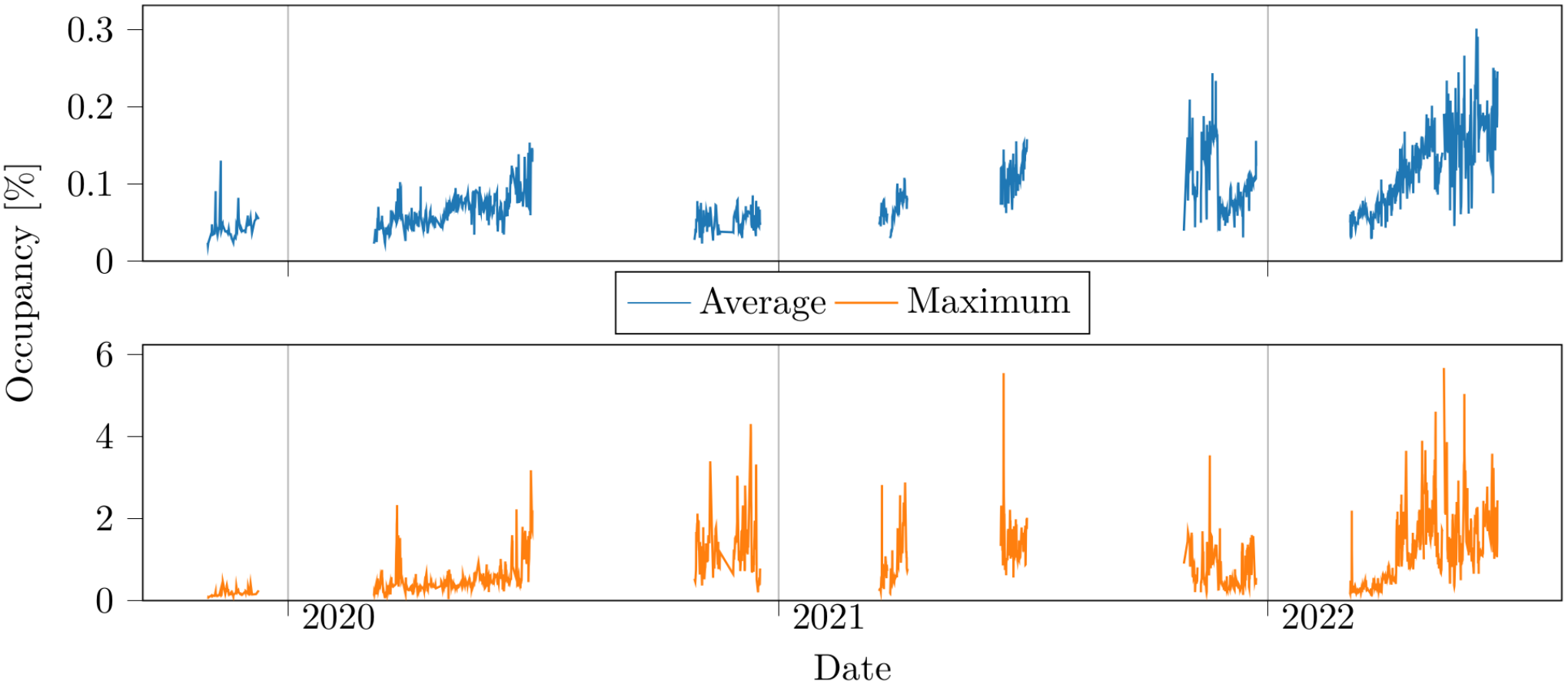


Figure 6.7: Average and maximum occupancy for each run in Phase 3. The PXD DAQ is designed to handle up to 2.5 % average occupancy.

PXD occupancy in run I one order of magnitude below design limit of PXD DAQ.

Simon Reiter, Ph. D. thesis, 2022

PXD DAQ truncation of data

- During injection, occupancy exceeds 10% or more (“burst events”)
- HLT can not keep up with processing, PXD data backpressure
- Detected by common mode value of 63 (“full FIFO” on DHP)
- 2.472.019 truncated events have been detected

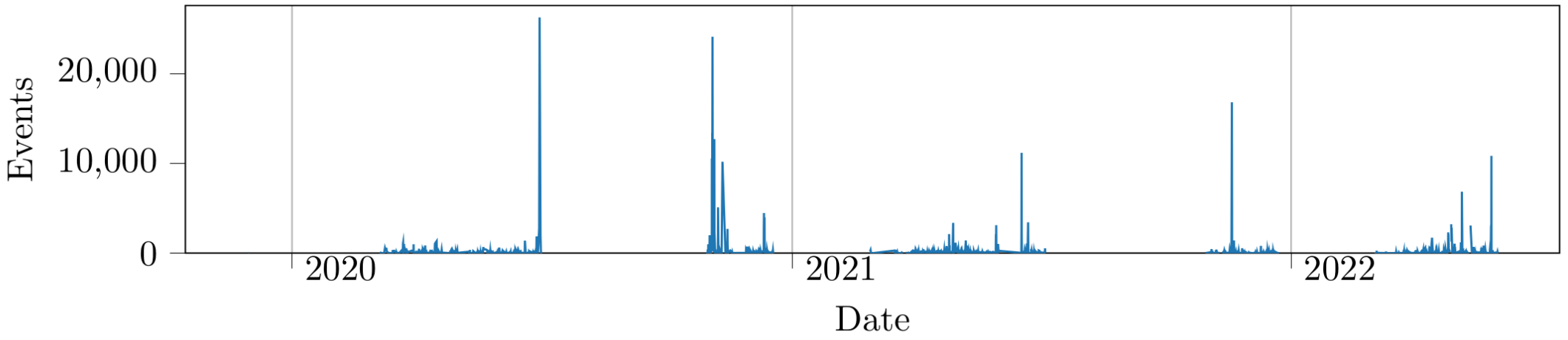


Figure 6.5: Truncated events for each run in Phase 3. In the last month of 2022 a slight upwards trend is visible, when the accelerator was pushed to its limits.

Simon Reiter, Ph. D. thesis, 2022

PXD DAQ slow control

- Stable data taking had and has highest priority
  automatic „recovery mechanism“ during shift used 250 x
- Automated pedestal calibration (between run stop and restart)
- Single Event Upsets (SEU)
  - ASIC registers protected against single event by triple redundancy; SEU unlikely, but possible (and observed)
  - SEU monitor: reading back the DHP (frontend) registers via JTAG periodically
- 35 SEU detected

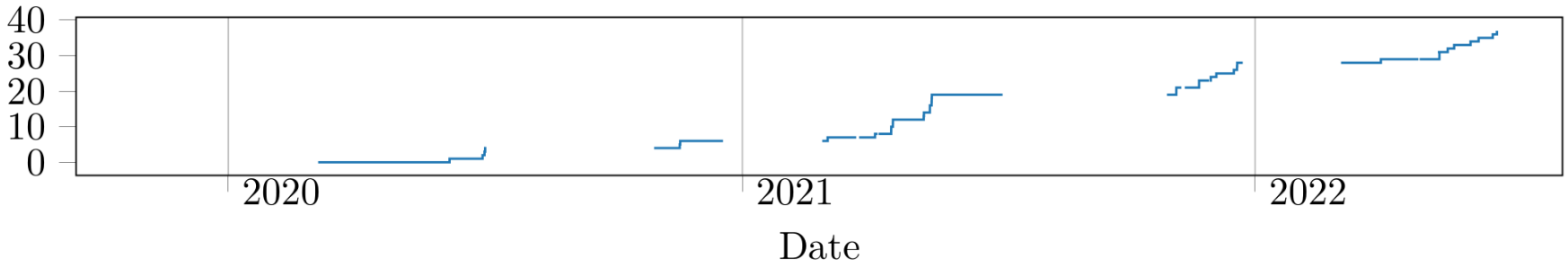


Figure 6.9: Accumulated detected SEUs in DHPs summed up over all modules.

Simon Reiter, Ph. D. thesis, 2022

PXD DAQ ROI (Region–of–Interest) filtering

- Tested, but not enabled yet (low luminosity, bandwidth from PXD DAQ to event builder not saturated yet, plan around 2026)

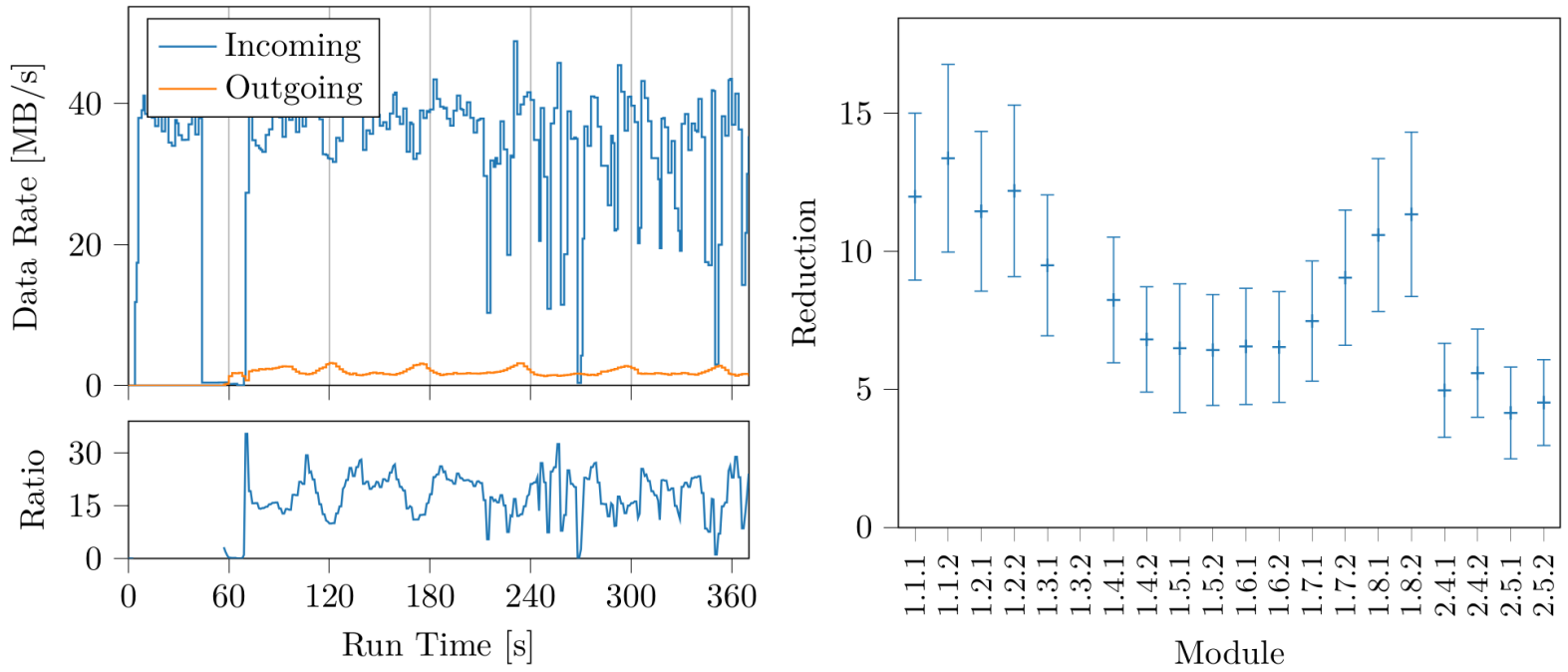


Figure 6.3: Single run with ROI filtering enabled. In the lower part the fraction is shown. The high fluctuations arise from the level-1 trigger, e.g. when a subsystem is busy and additional triggers are rejected. The right plot shows the total reduction ratio for ROI filtering only for each module individually calculated by DQM. The error bars represent the standard deviation.

Simon Reiter, Ph. D. thesis, 2022

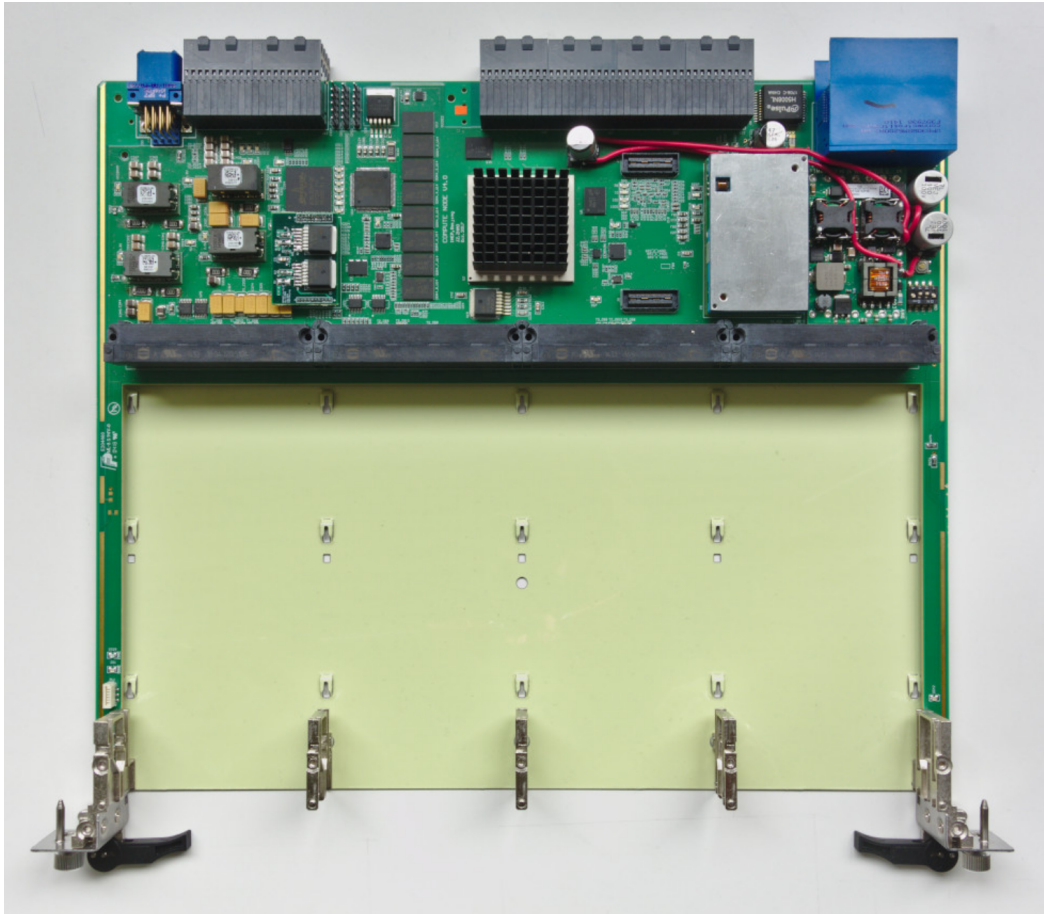
New hardware for high bandwidth data transfer (Jennifer2 WP 5.2 topic)

New Kintex carrier board (IHEP Beijiing, Giessen)

- Board also serves as spare for existing hardware
- Goal: AMC cards can be plugged and run identical firmware
- Kintex firmware modifications
  - New link protocol: AXIstream
  - No hard-core CPU anymore,

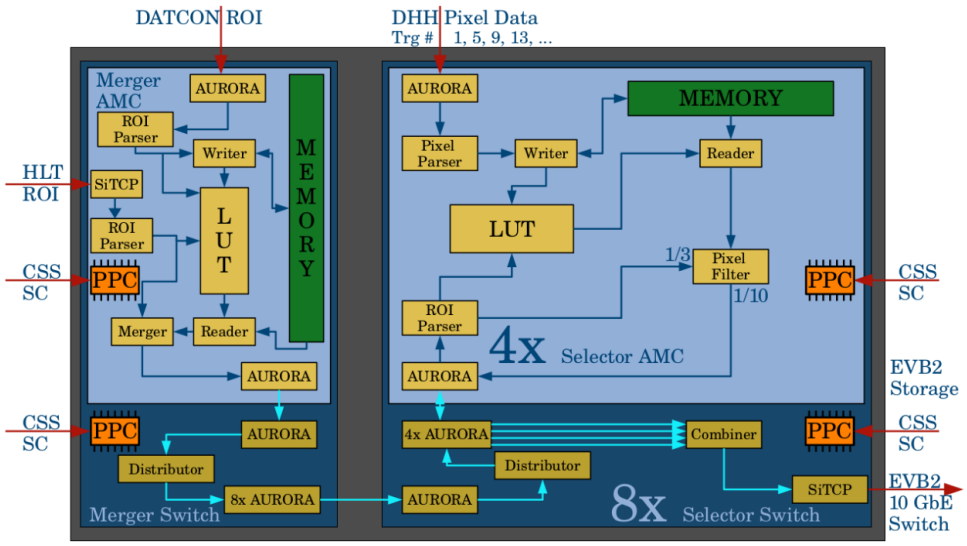
    requires soft-core (e.g. Microblaze) for slow control (EPICS)
  - Kintex requires programming by Vivado (while we used planAhead for Virtex)
- See talk by Matthäus Krein

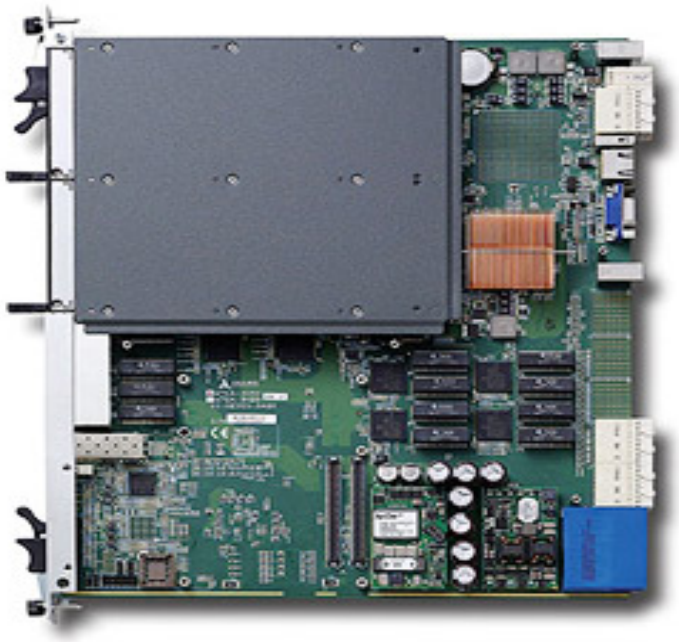
	Virtex-4 FX60 (CNCB)	Virtex-5 FX70T (xFP)	Kintex UltraScale 060 (Upgrade)
Registers	50k	44k	663k
LUTs	50k × 4-input	44k × 6-input	332k × 6-input
DSP Slices	128	128	2760
BRAM	4 Mb	5 Mb	38 Mb
MGT	16 × 6.5 Gbps	16 × 6.5 Gbps	32 × 16.3 Gbps
CPU	PPC405	PPC440	-



Test of ATCA 10G Switch

- Ansatz: send data to event builder
  - not by 32 GbE links
  - but to the ATCA backplane,
    then to a ATCA switch,
    and then by only one 10G link



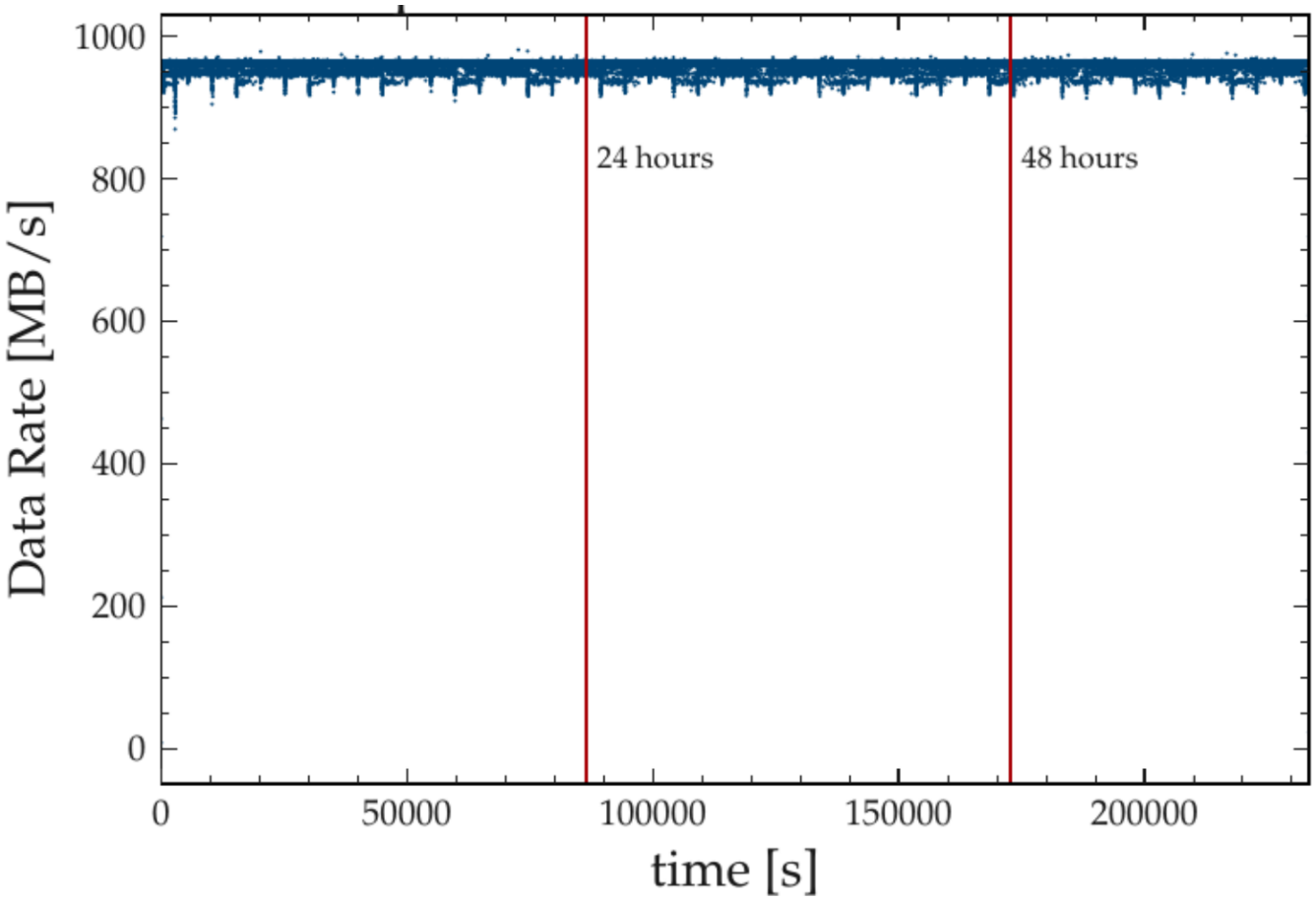


ATCA ADLINK aTCA-3150
10G Uplink

ATCA ADLINK aTCA-3710
40G Uplink

Klemens Lautenbach, Ph. D. thesis, 2020

Test of ATCA 10G Switch (Output data rate of full ONSEN)



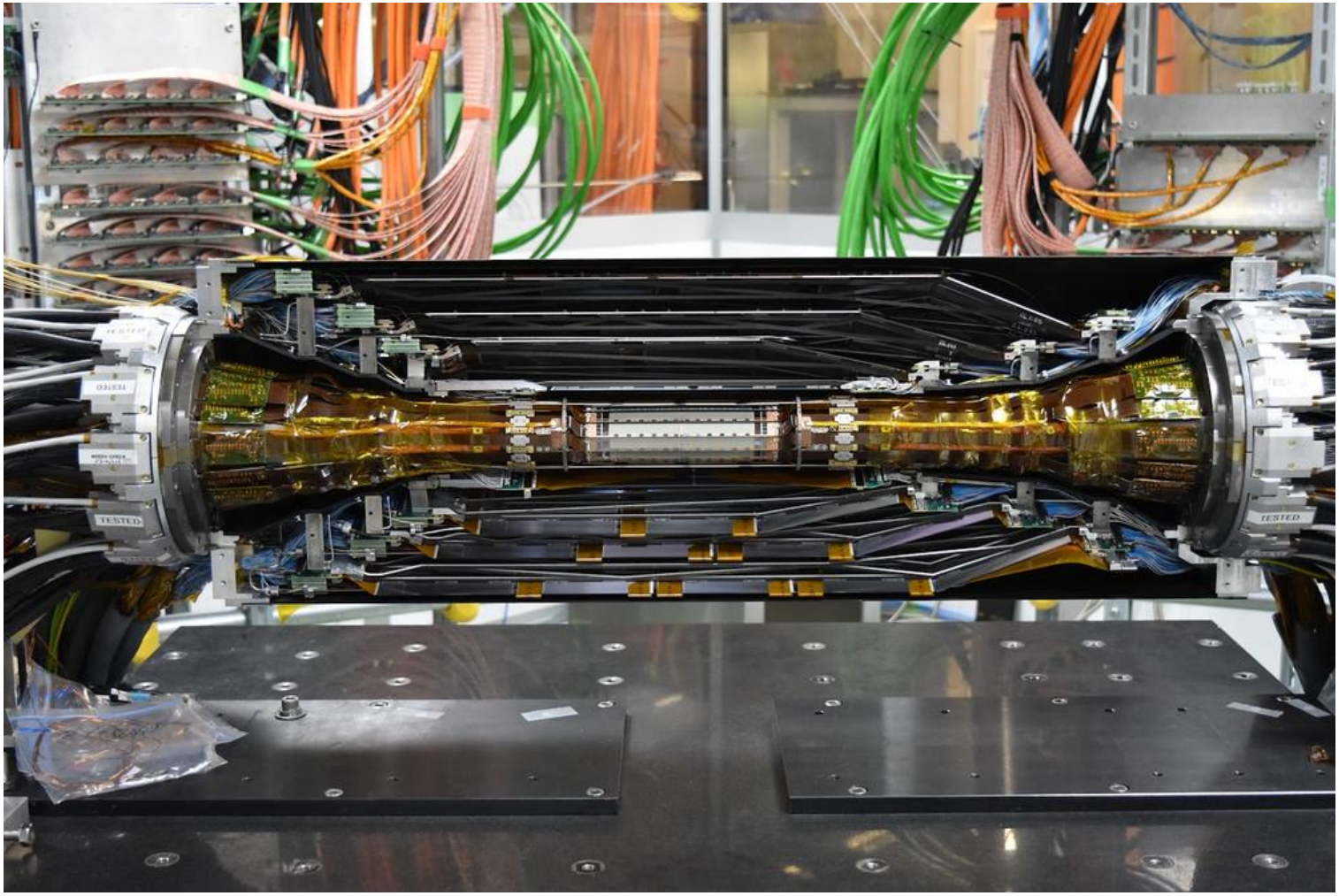
Long term test, almost 60 hours (3 days)

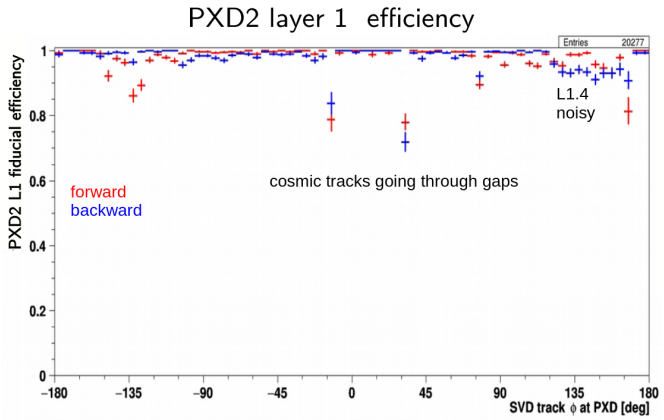
Bandwidth 950 MB/s achieved, close to theoretical 10 GbE limit

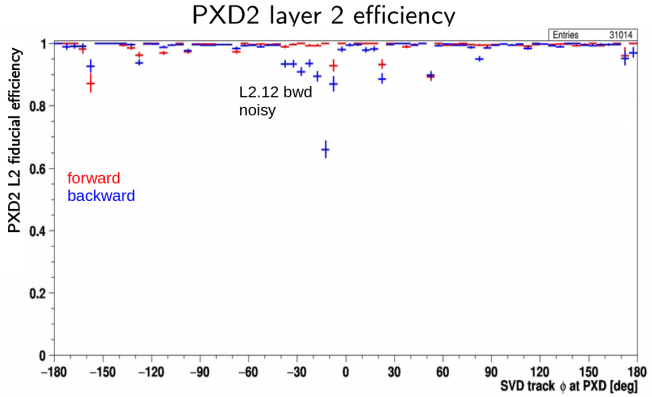
Klemens Lautenbach, Ph. D. thesis, 2020

PXD2 (2023)

Both hit efficiency and hit purity ~20% higher than PXD1.

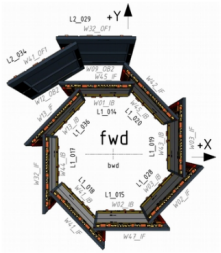






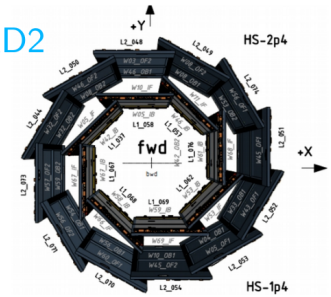
Efficiency
from cosmic data taking
in 09/2023

PXD1



LS1

PXD2

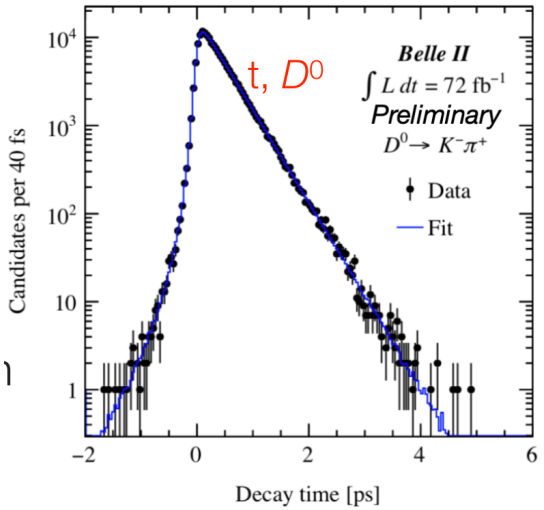


~8 Million pixels
(factor 2 more than PXD1)

Summary and conclusion

- PXD1 stable data taking in Run I finished
  400+ Million events recorded
  among the physics results: precise lifetime measurements of hadrons

Particle	Measured lifetime	Reference
Λ_c^+	203.2 ± 0.9 ± 0.8 fs	Phys. Rev. Lett 130 (2023) 071802
Ω_c^0	243 ± 48 ± 11 fs	Phys. Rev. D 107 (2023) L031103
D^0	410.5 ± 1.1 ± 0.8 fs	Phys. Rev. Lett. 127 (2021) 211801
D^+	1030.4 ± 4.7 ± 3.1 fs	Phys. Rev. Lett. 127 (2021) 211801
D_s^+	499.5 ± 1.7 ± 0.9 fs	Phys. Rev. Lett. 131 (2023) 171803
B^0	1499 ± 13 ± 8 fs	Phys. Rev. D 107 (2023) L091102



- PXD2 installed in 2023, data taking ongoing
- New challenges in RUN 2, plan for 2024:
  - Higher luminosity 10^{35} cm^{-2} s^{-1} (corresponding to ~1 Belle or ~2 BaBar per year)
  - Increase beam currents x factor 2 (higher background, maybe noisier injection)
  - ROI selection will be switched on when link to event builder is satured