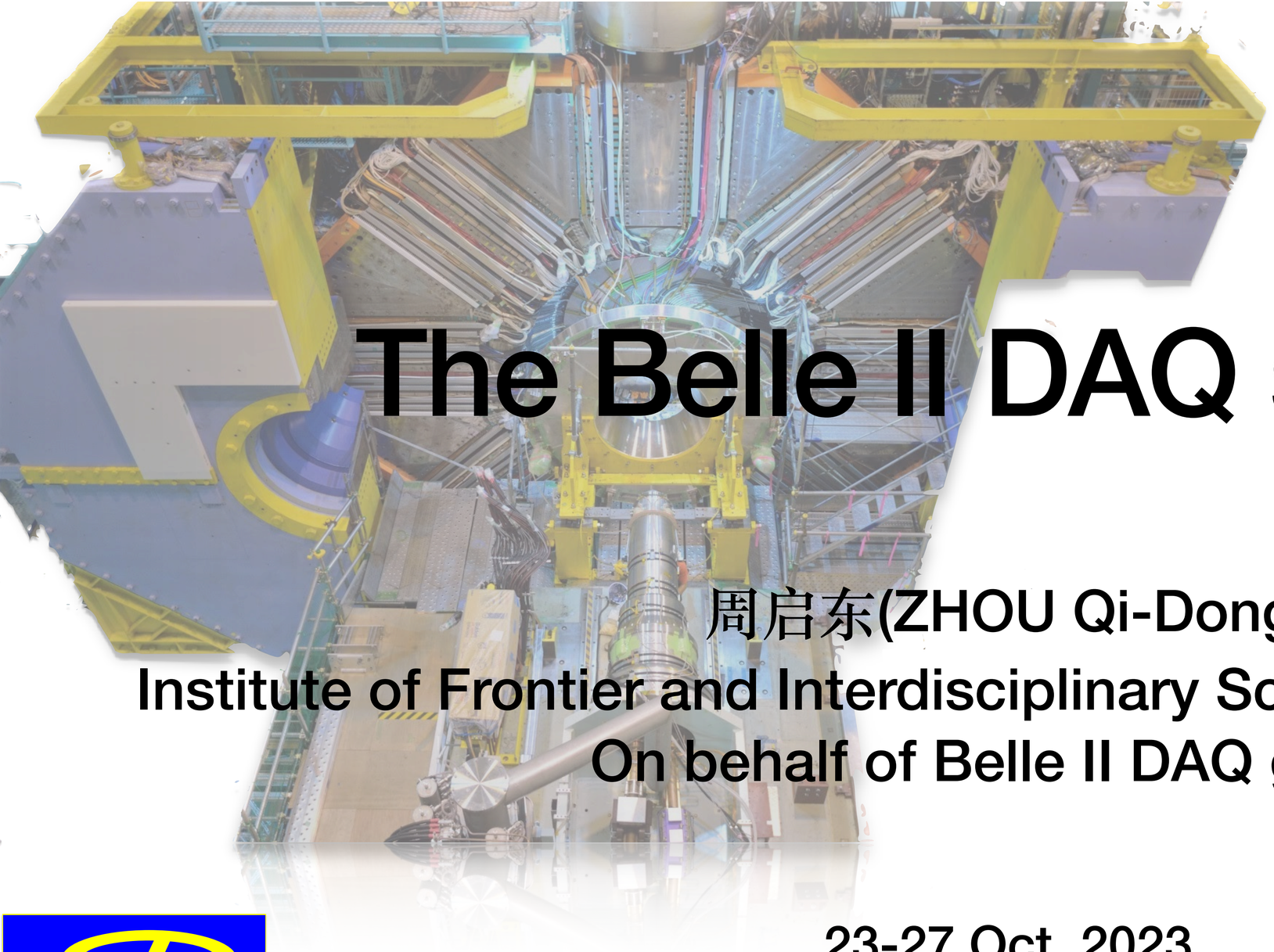


The Belle II DAQ system

周启东(ZHOU Qi-Dong)

Institute of Frontier and Interdisciplinary Science, Shandong Univ.

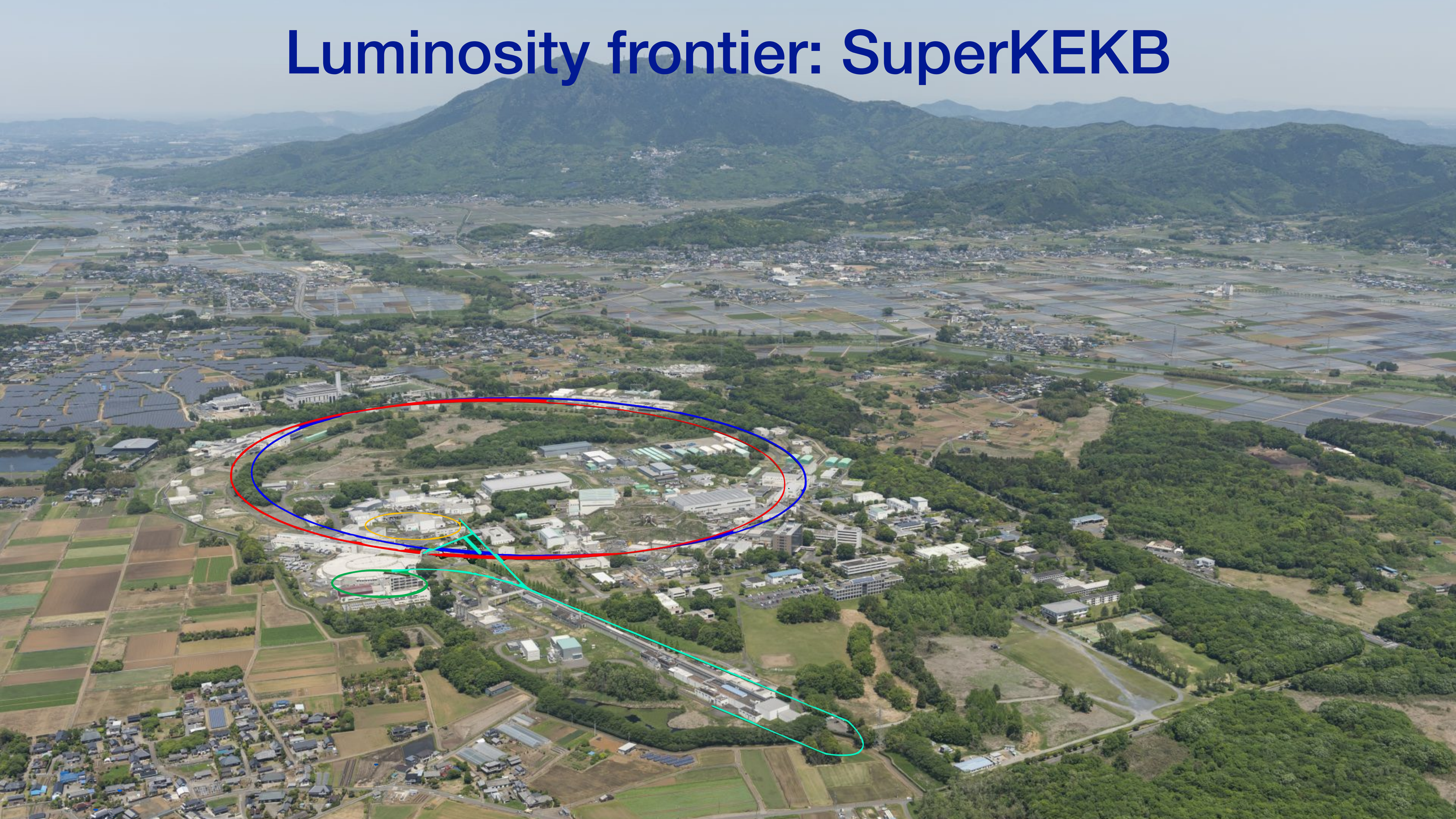
On behalf of Belle II DAQ group

23-27 Oct. 2023

The 2023 international Workshop on

the High Energy Circular Election Positron Collider

Luminosity frontier: SuperKEKB



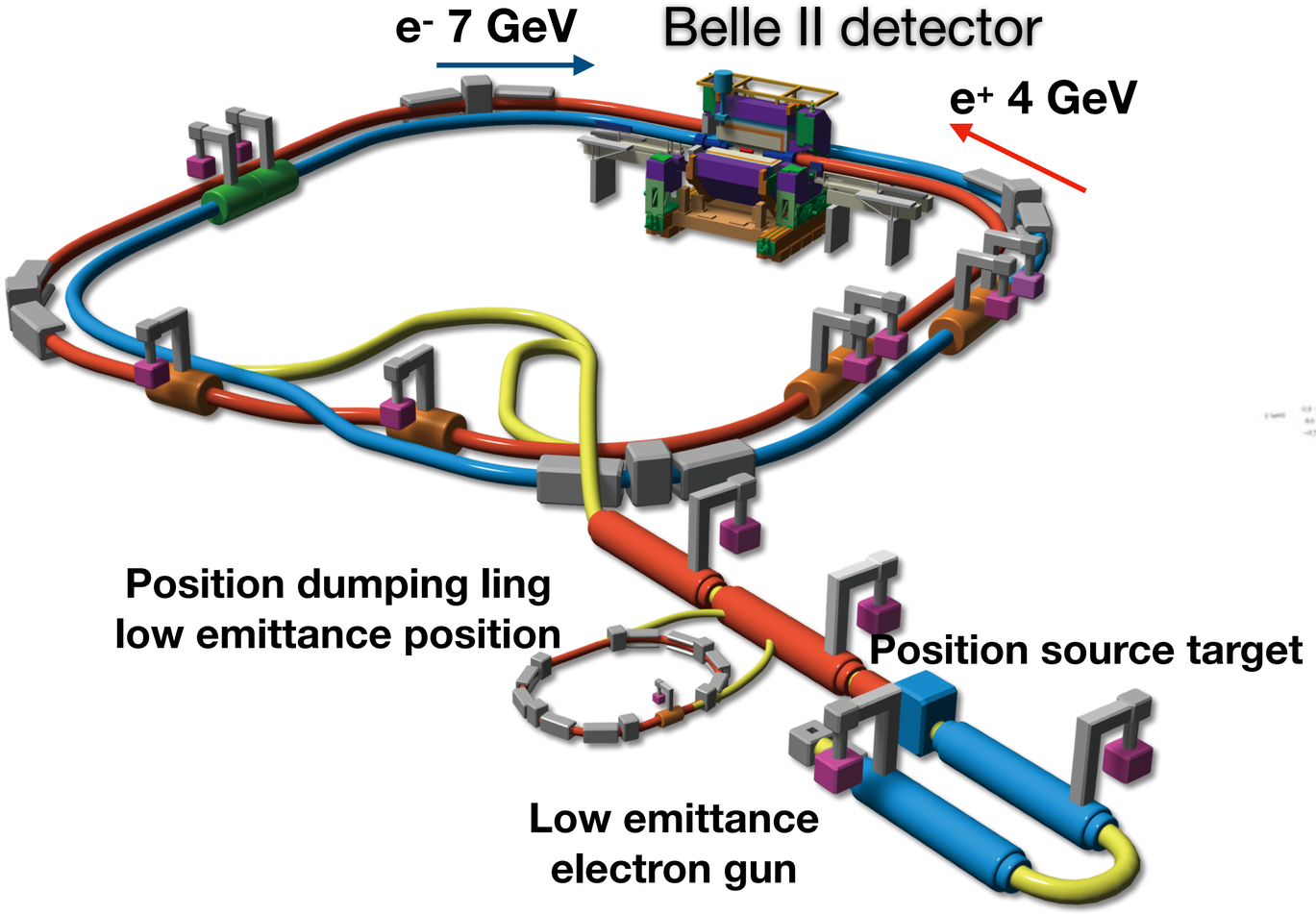
Luminosity frontier: SuperKEKB

- Asymmetric e^+e^- collider
  - $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$
  - very clean and well-known initial state

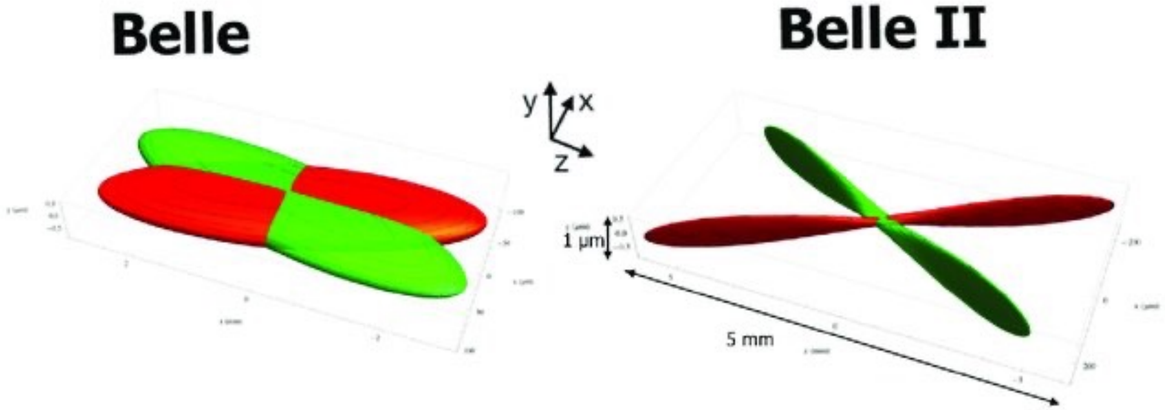
Beam current: KEKB x ~1.5

$$L = \frac{\gamma_\pm}{2er_e}\left(1+\frac{\sigma_y^*}{\sigma_x^*}\right)\frac{I_\pm \xi_{\pm y}}{\beta_y^*}\left(\frac{R_L}{R_y}\right)$$

Beam squeeze: KEKB / ~20



Nano beam scheme



Target: L = 60 x 10^{34} cm^{-2} s^{-1}

Achieved : 4.7 x 10^{34} cm^{-2} s^{-1} (Record)

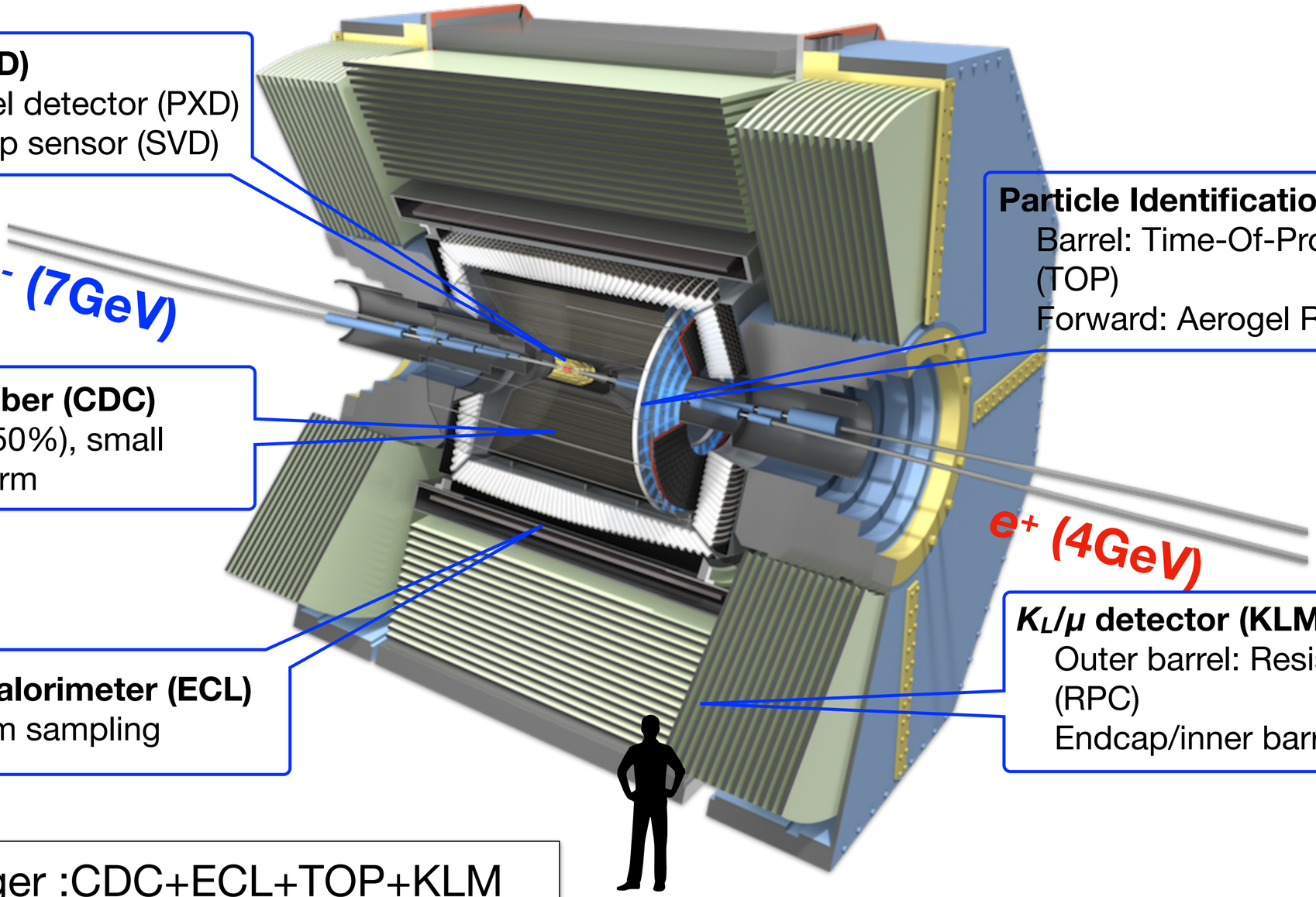
- Data at $\Upsilon(4S)$:
  - 362 fb^{-1} (Belle II) <-> 711 fb^{-1} (Belle)

The Belle II detector

Vertex detector (VXD)
Inner 2 layers: pixel detector (PXD)
Outer 4 layers: strip sensor (SVD)

Central Drift Chamber (CDC)
He (50%), C_2H_6 (50%), small cells, long lever arm

ElectroMagnetic Calorimeter (ECL)
CsI(Tl) + waveform sampling



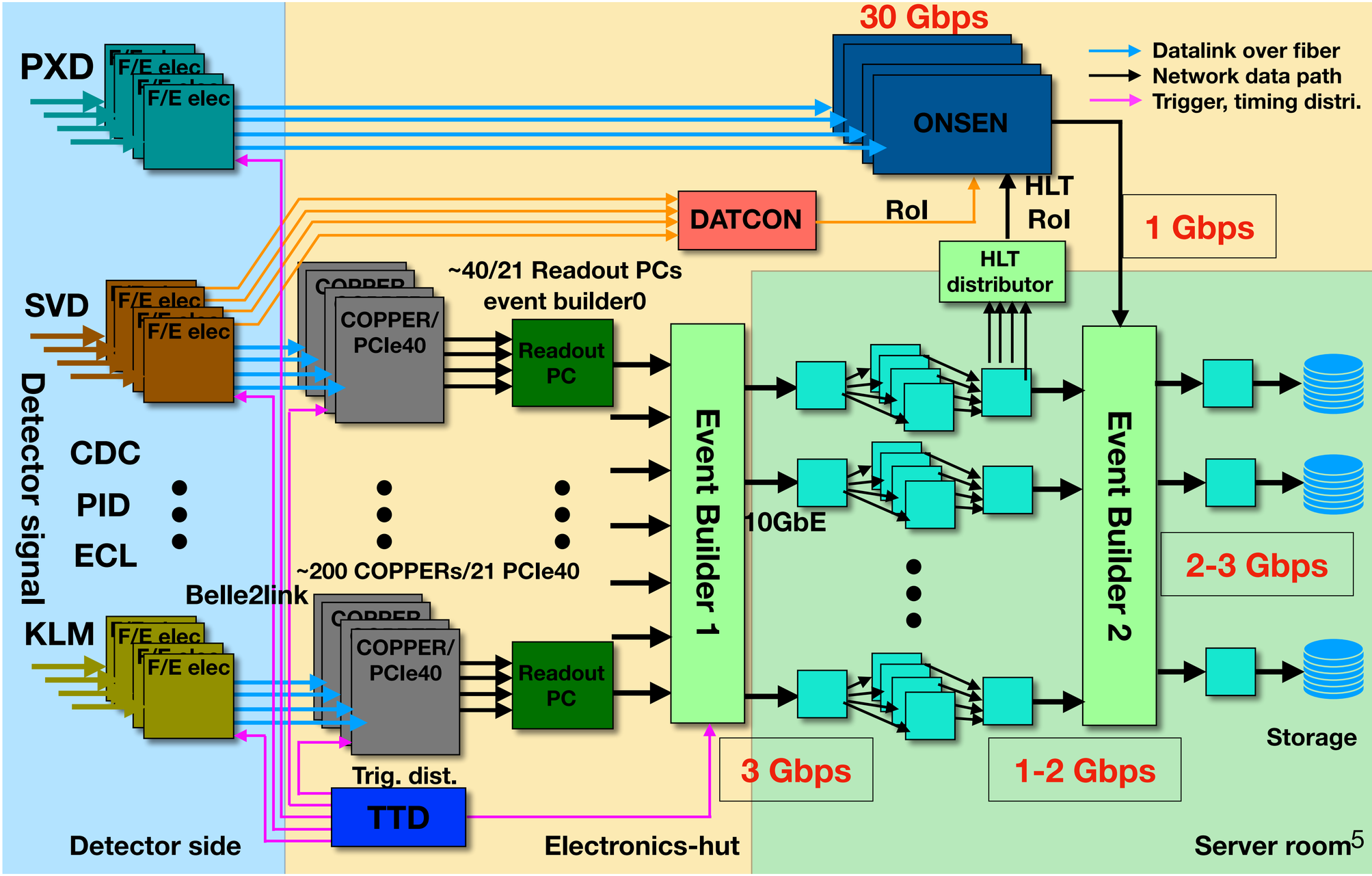
Particle Identification
Barrel: Time-Of-Propagation counters (TOP)
Forward: Aerogel RICH (ARICH)

***K_L/μ* detector (KLM)**
Outer barrel: Resistive Plate Counter (RPC)
Endcap/inner barrel: Scintillator

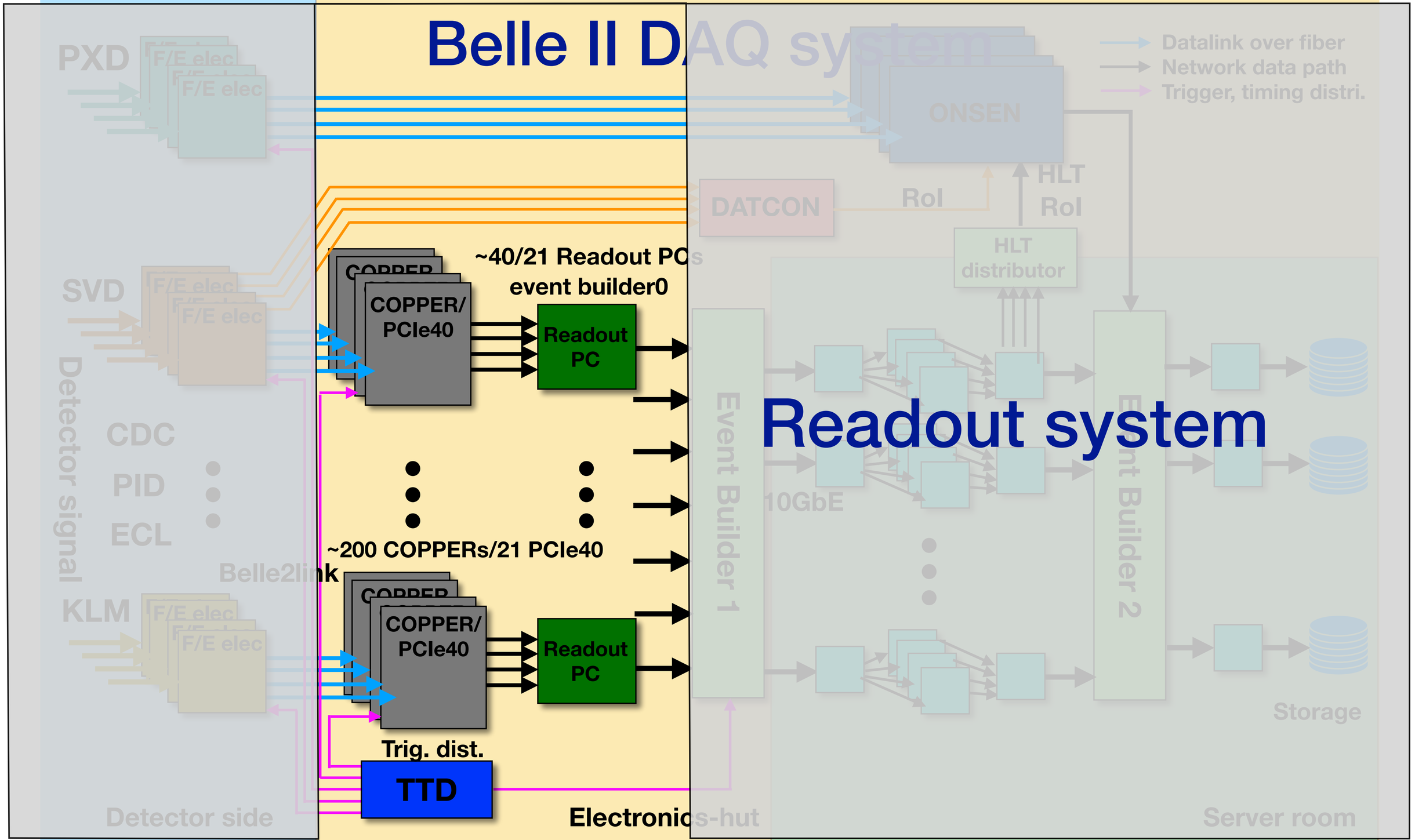
Level-1(L1) trigger :CDC+ECL+TOP+KLM
DAQ: Maximum 30 kHz L1 trigger

Belle II DAQ system

- Unified common readout system (except for PXD)
- Unified timing and trigger distribution (TTD) system
- A pipeline readout
- To handle 30 kHz level 1 trigger with ~1% dead time under raw event size of 1 MB



Belle II DAQ system



Readout system

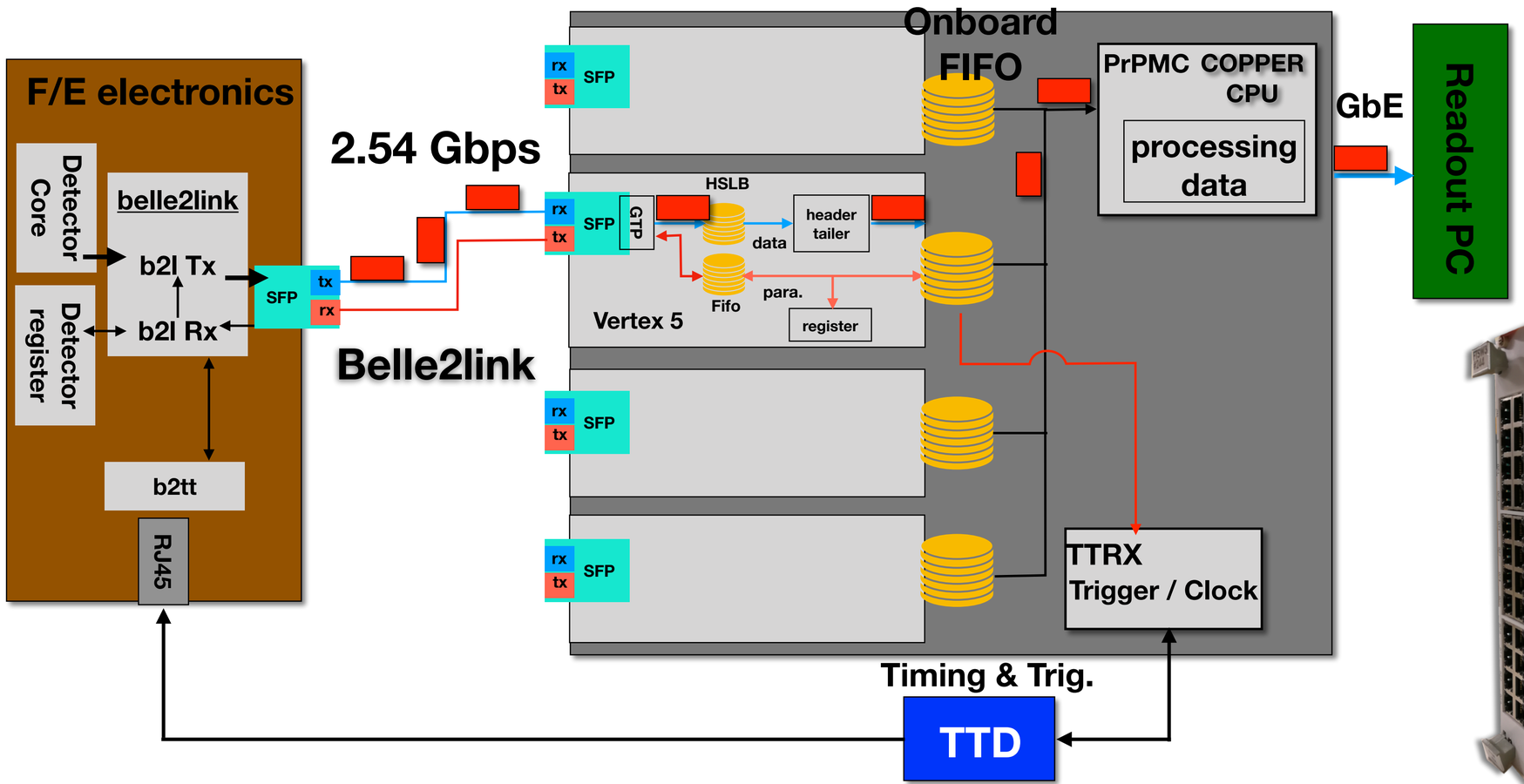
Belle II DAQ readout system

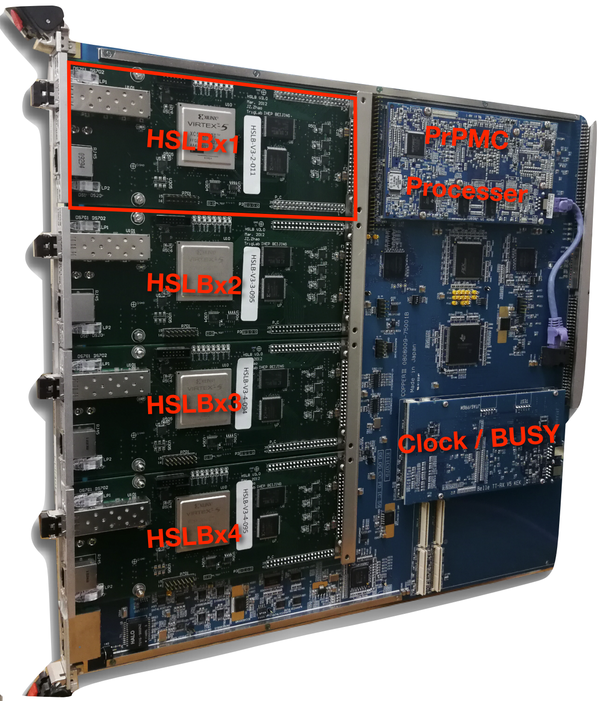
Belle2link:

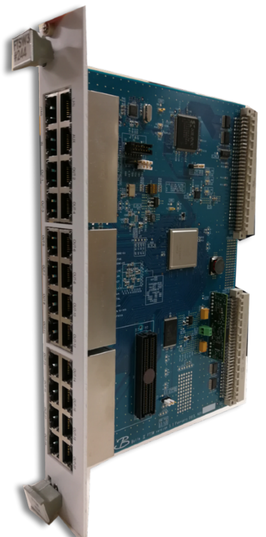
Unified high speed optical link (2.54Gbps) connected Front-End Electronics and DAQ readout board, data transmission based on Rocket I/O.

Functionalities of readout system

- Belle2link,
- TTD interface,
- slow control
- pre event-building, GbE
- Data-formatting
- Data-check

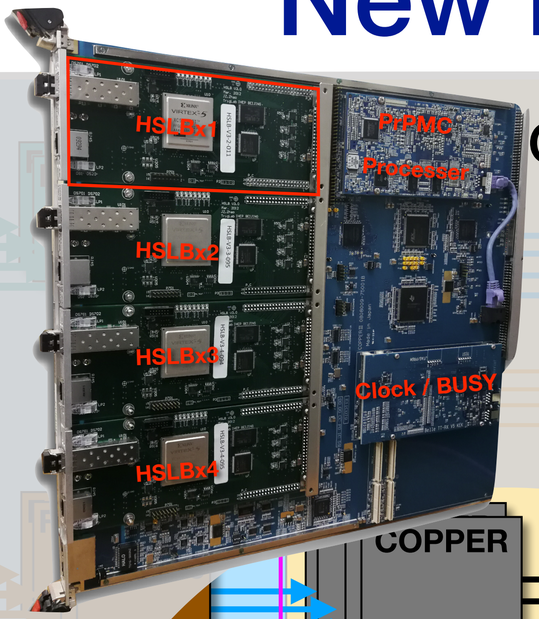






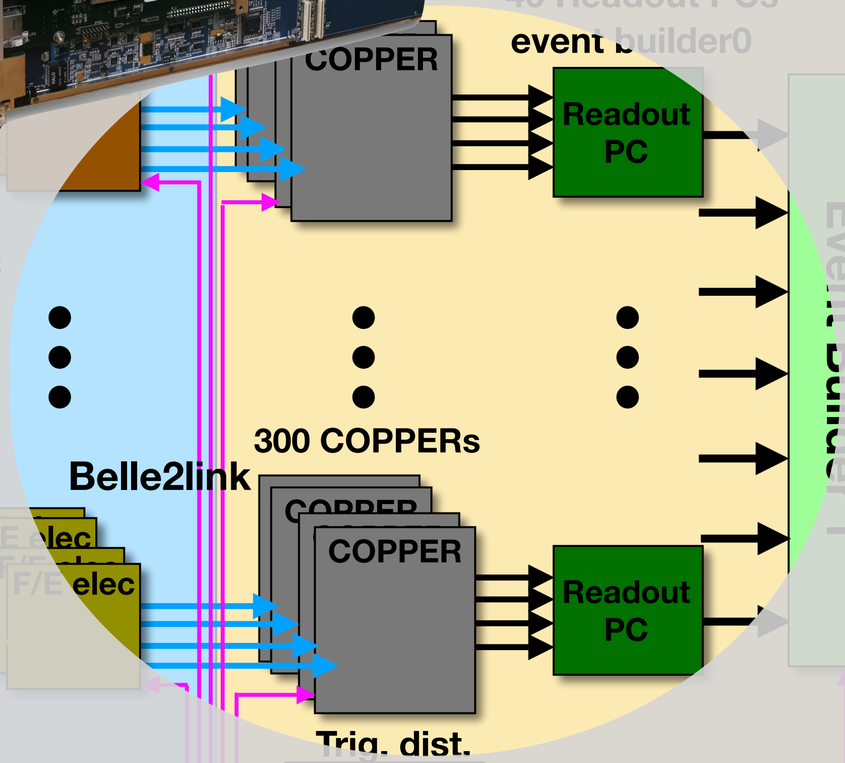
FTSW module for TTD

New readout system using PCIe40



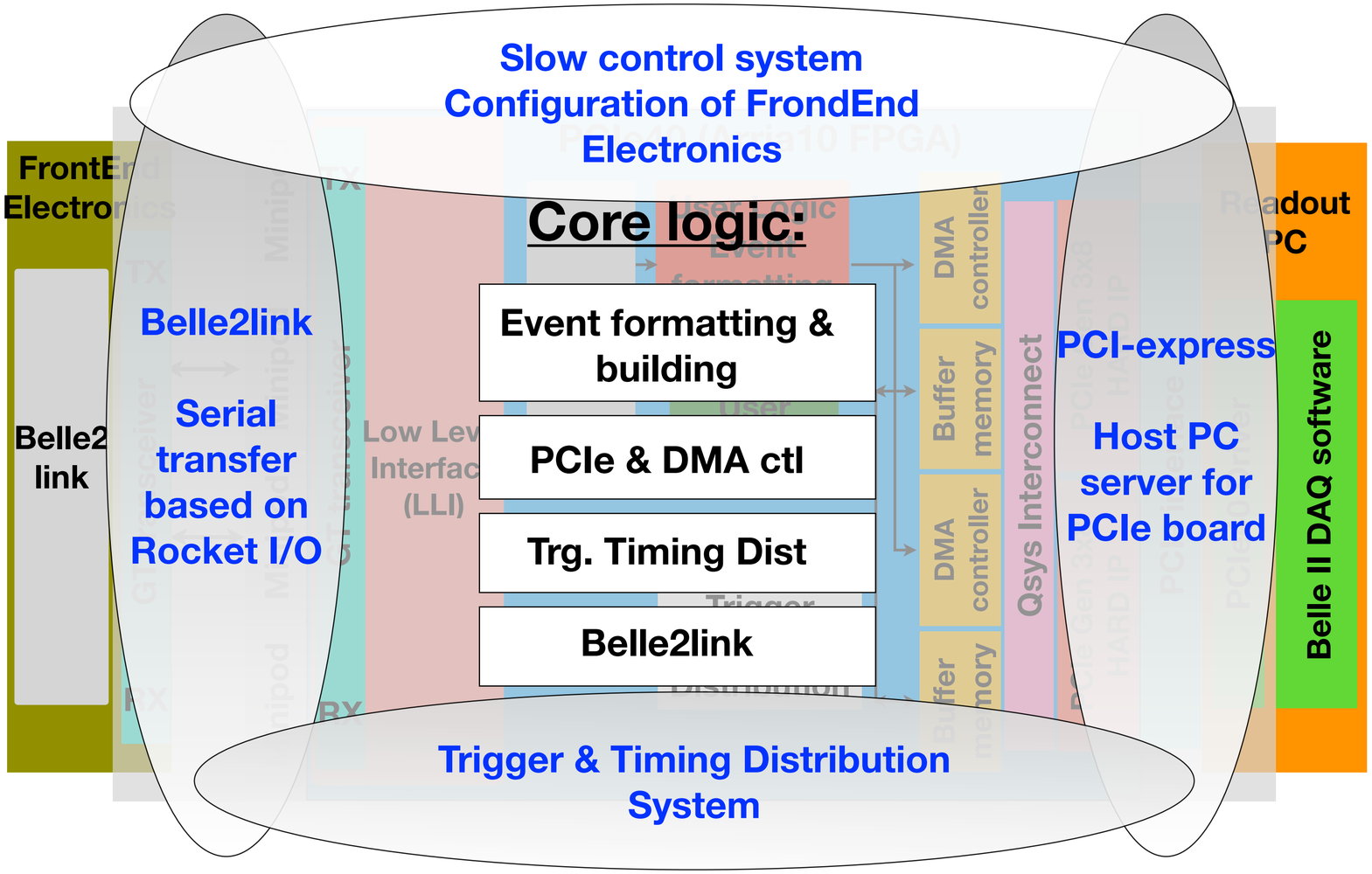
COPPER

PCIe40
(LHCb, ALICE)



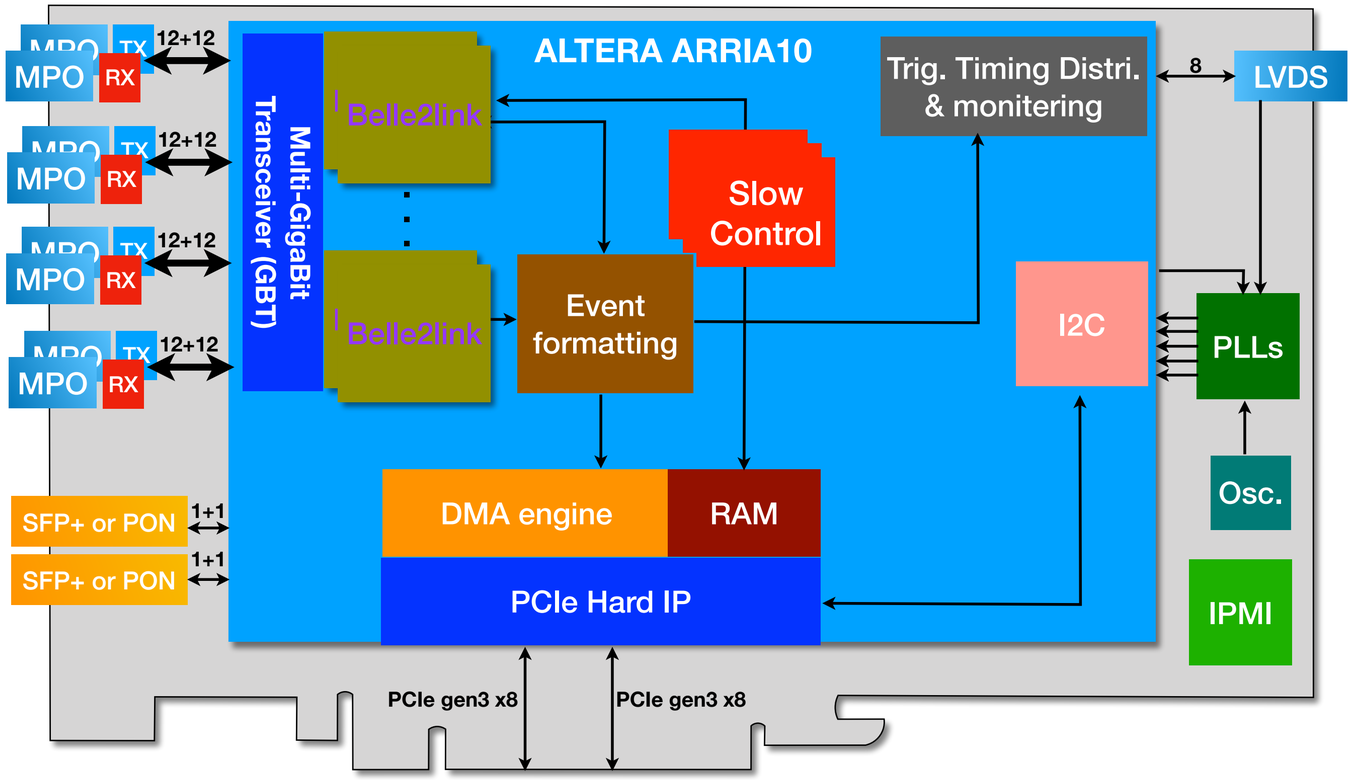
- ✦Difficult to maintain during the entire Belle II operation period, COPPER has been used for Belle experiment
  - Four different boards
  - Broken parts are increasing (PrPMC CPU, chipset,…)
- ✦Limitation to improve DAQ performance
  - Bottlenecks of COPPER
    - CPU usage
    - Data transfer speed (1 Gbps)
  - Bottleneck of network output of ROPC (1 Gbps)
- ✦Future possibility
  - Luminosity & background situation changed
  - For a trigger-less DAQ

PCIe40 based new readout system for Belle II



- A upgrade of current COPPER based readout system with PCIe40 board
- Upgrade of readout system will keep the modification as less as possible, for the system connected.
- No major modification required for hardware and firmware of the sub-detectors' systems.
- Software for slow control and data readout need some upgrades

Features of Belle II PCIe40 readout system

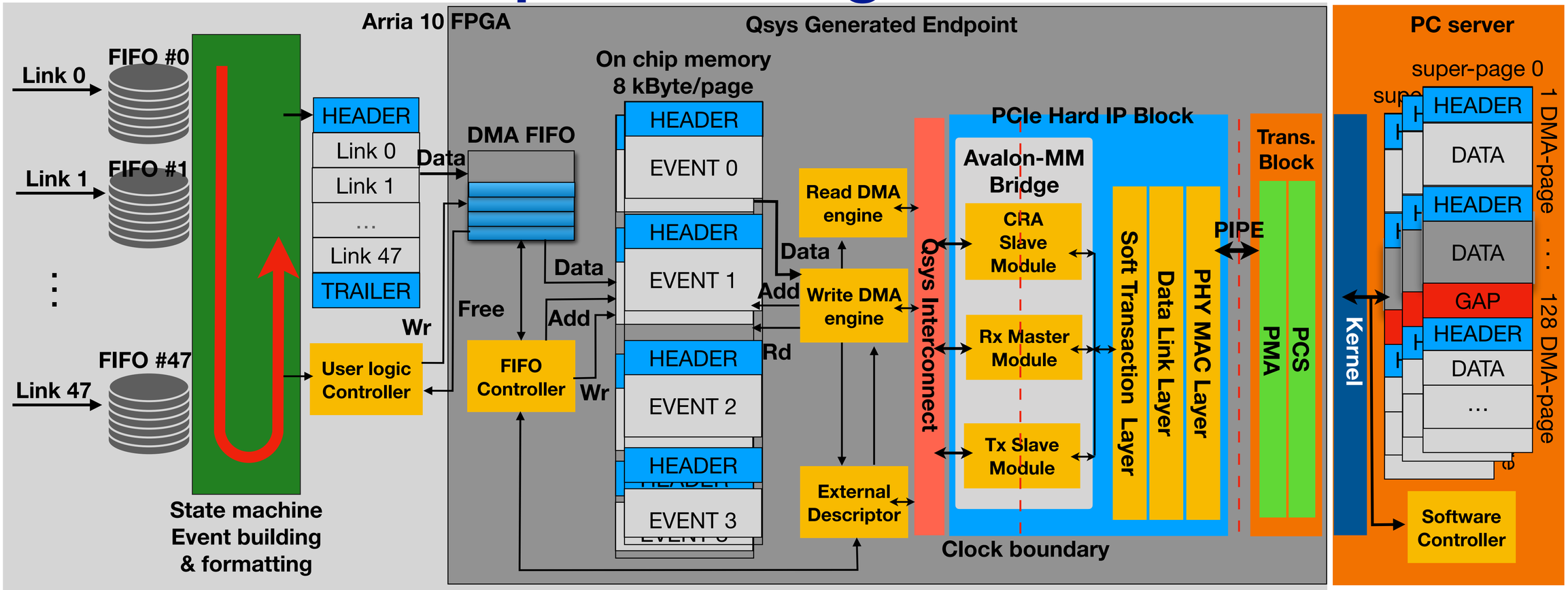


Board	FPGA family	Optical links	PC interface	Experiment
PCIe40	ALTERA Arria10	48	2 PCIe Gen3 x 8	LHCb, ALICE

21 PCIe40 boards to replace 203 COPPERs
A compact system can be achieved

- Belle2link protocol: kept as same functionality, but from Xilinx FPGA port to Intel Arria10 FPGA
- Event building and formatting: newly added based on the FPGA logic (on board CPU for COPPER)
- Slow control: protocol part moved to software
- b2tt link(connect to TTD): new design to handle 48 links
- PCIe based DMA: external DMA descriptor controller apart from DMA engine (based on Qsys).

Data processing with DMA



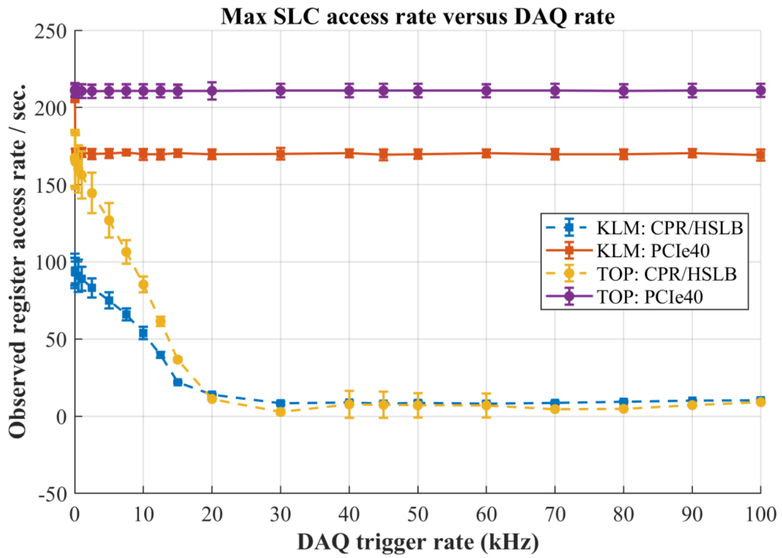
- Event-building
  - Reduction of header and trailer info of each link
  - Data check
    - CRC calculation, mismatch headers among different links
  - Add error-bit flag to the builded event
  - Data transmission rate: 61 Gbits/s <—>

Theoretical maximum data rate is 50 Gb/s
can eventually be increased to 100 Gb/s

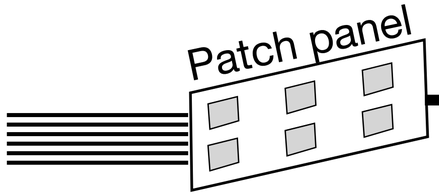
Performance of slow control system

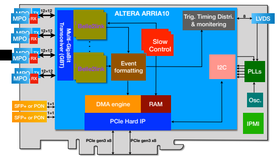
- Belle2link was kept the same as COPPER-HSLB system
- 3 slow control access methods for PCIe40 were implemented and tested
  - A7D8 and A16D32 kept the same features as HSLB
  - Streaming file method separated based on packet size;
    KLM (6 words / pocket), ARICH (100 words / pocket)
- A16D32 access:
  - 83 us / access <—>1 ms / access for HSLB
- Streaming file:
  - 360 KBps (KLM) <—> 350 KBps for HSLB
  - 1-2 sec downloading ARICH firmware <—> 1-2 sec for HSLB
- Parallel access of slow control + data acquisition with multiple links is working well
  - It takes the same time for the access w/ and w/o parallel access
- Slow control configuration for FEEs of KLM, TOP, ARICH has been tested and working fine.

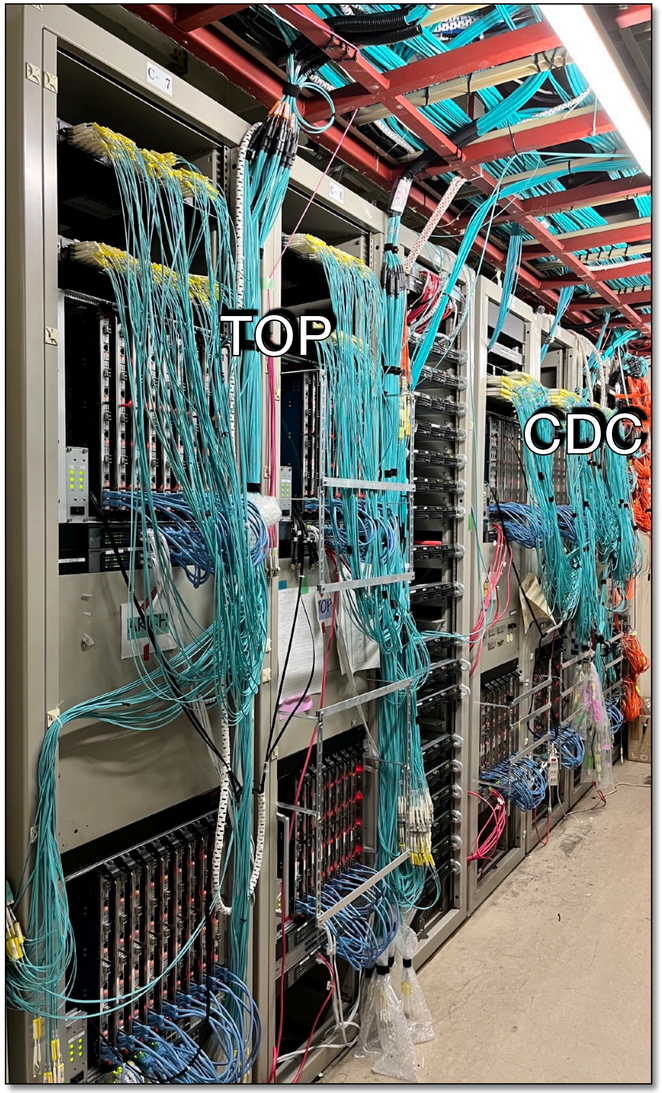
Detector	A7D8	A16D32	byte stream
SVD	○		
CDC		○	
TOP		○	
ARICH		○	○ (~3MB)
ECL	○	○	
KLM		○	○



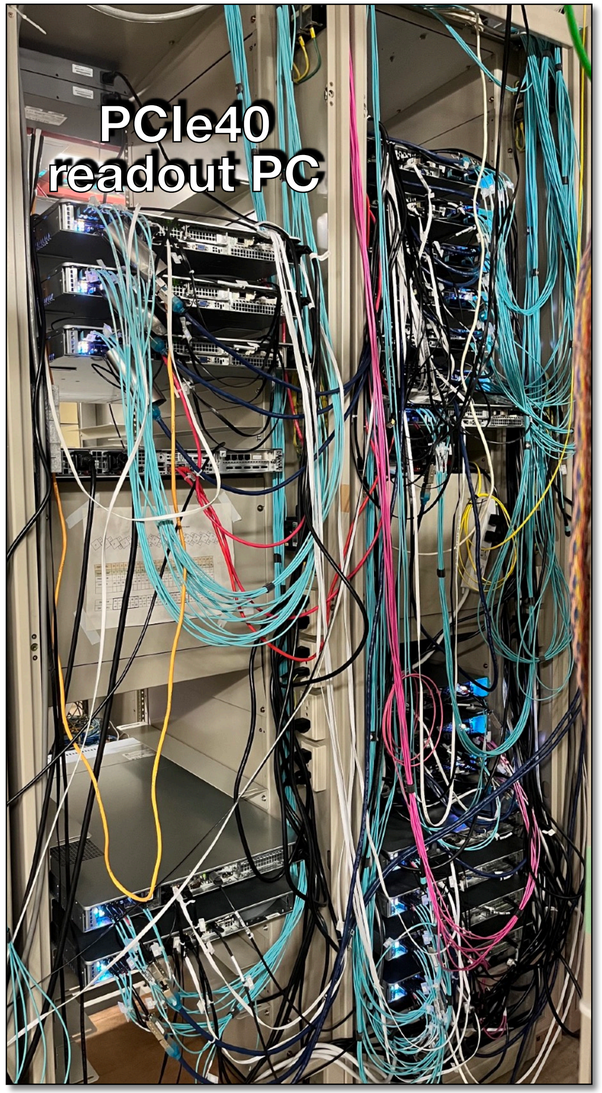
Installation of new readout system



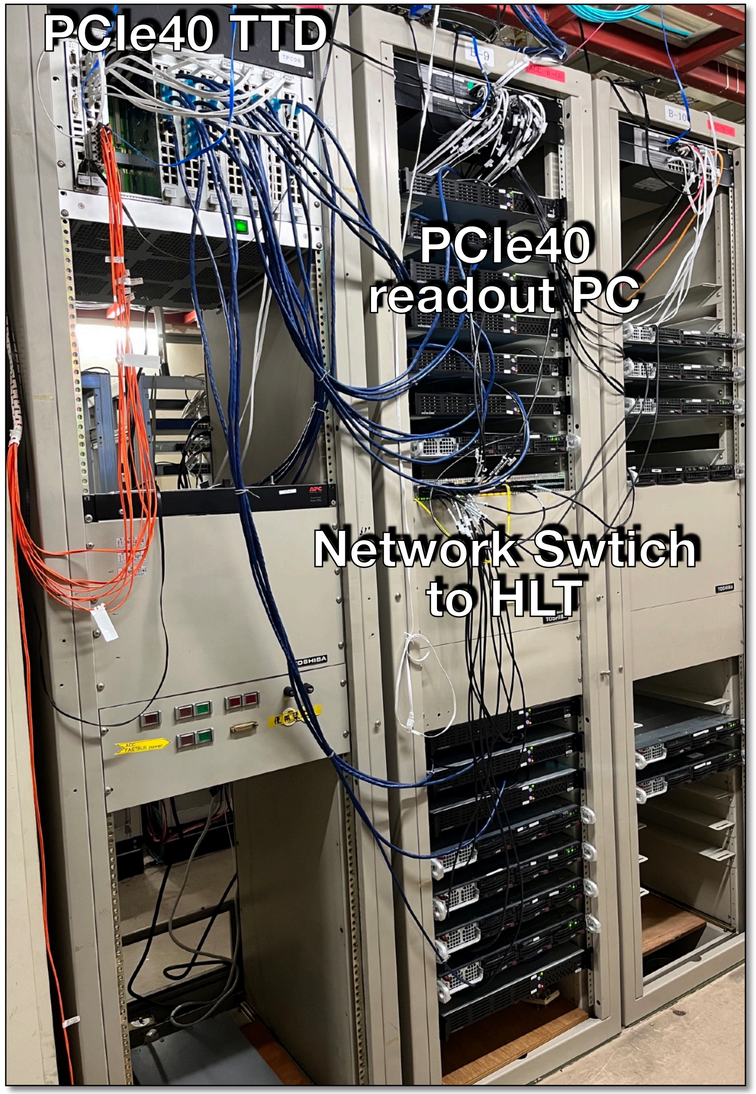


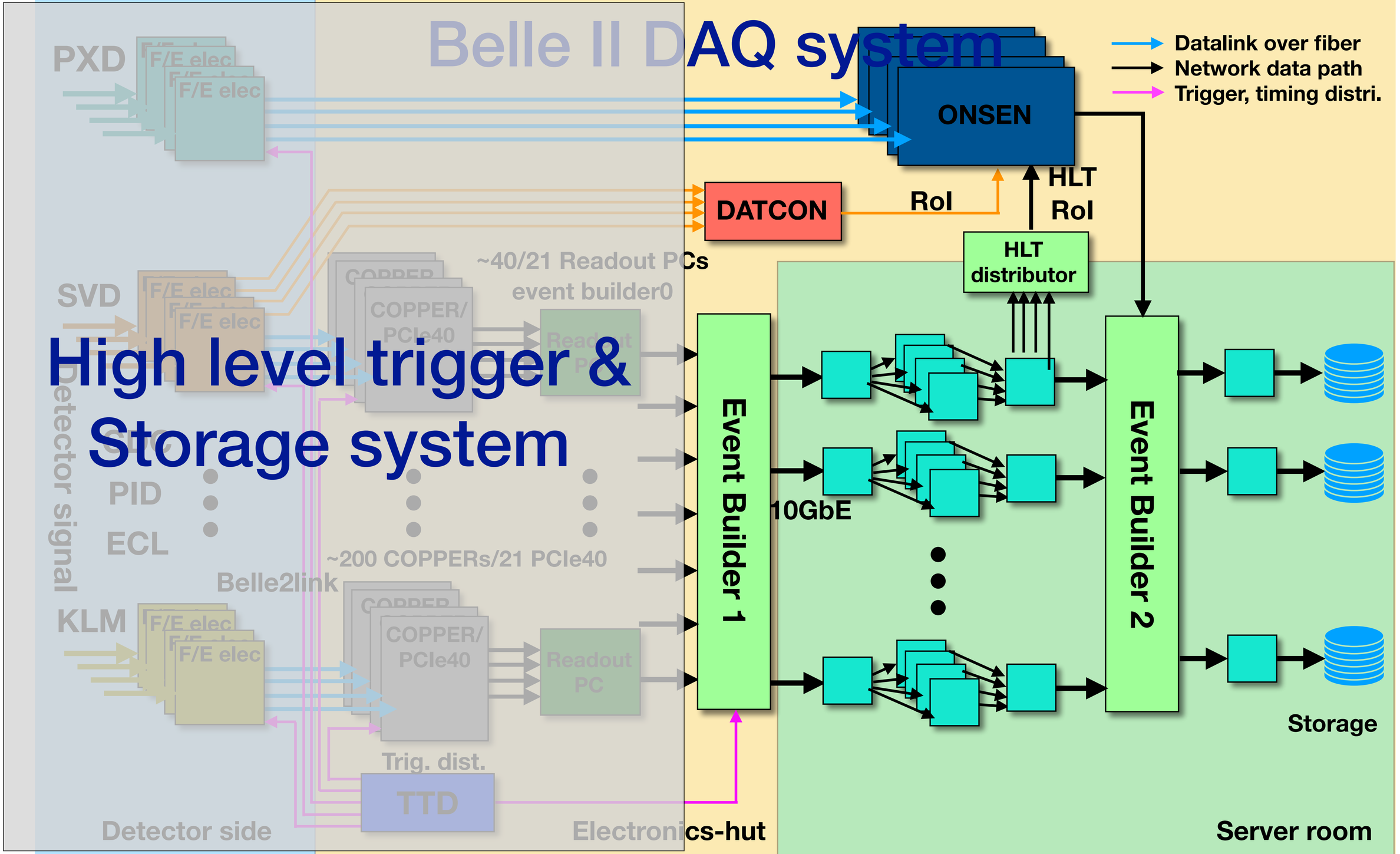


Back view



Front view

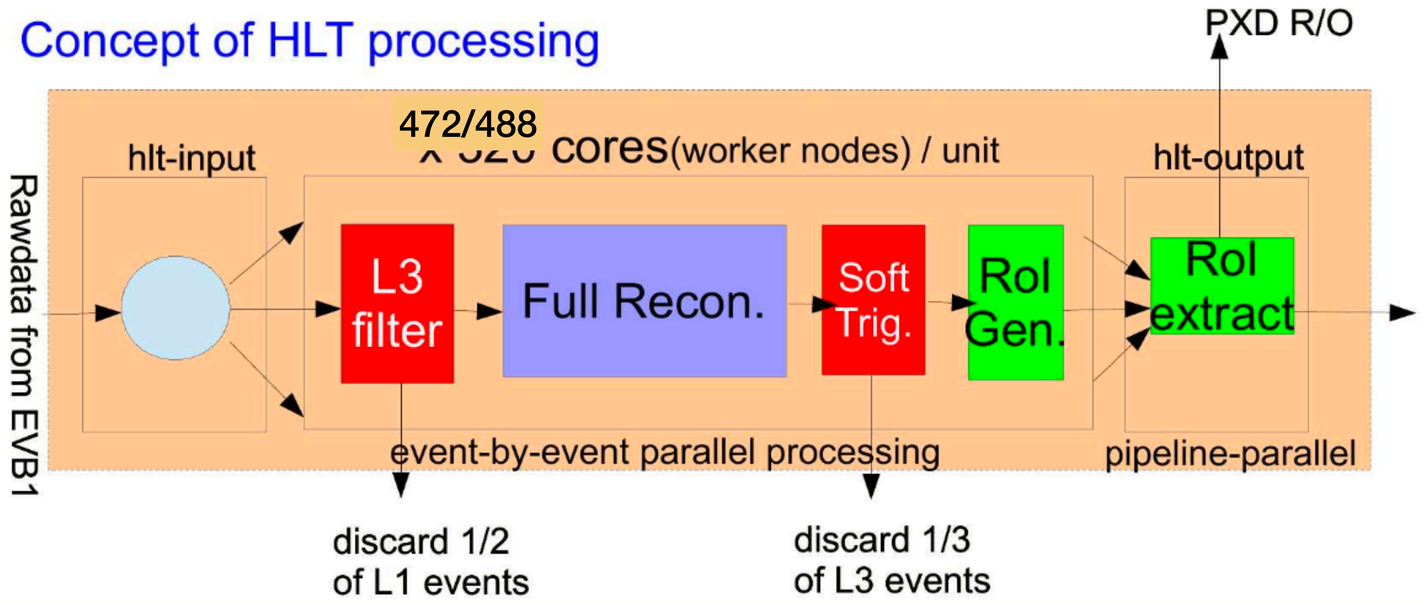


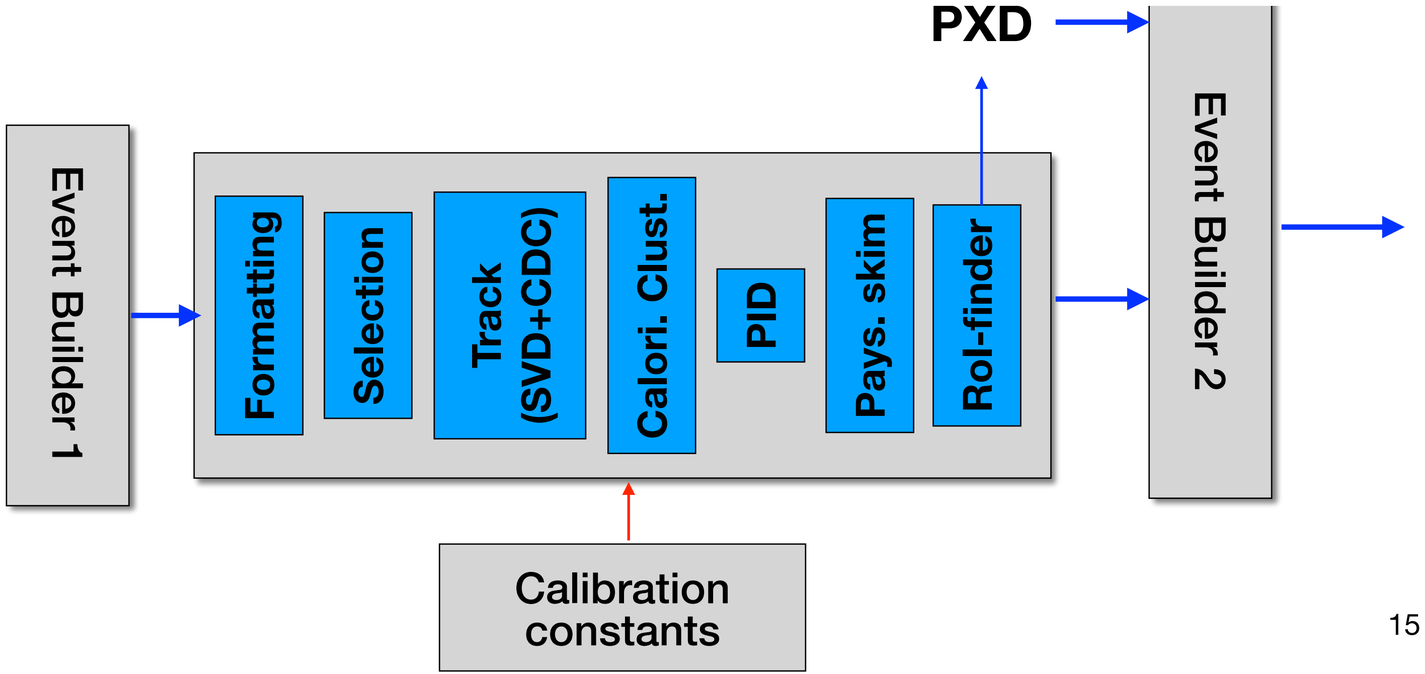


High level trigger & Storage system

Overview of high level trigger system

- Full event reconstruction
- Crude calibration constant
- Factor 8 rate reduction
- 13 HLT units, in total 6212 cores (design: 6400 cores)
- Data processing: ~ 2.1kHz/ HLT unit w/ hyper-threading
- Event size at HLT in the last run period: ~150 kB/event



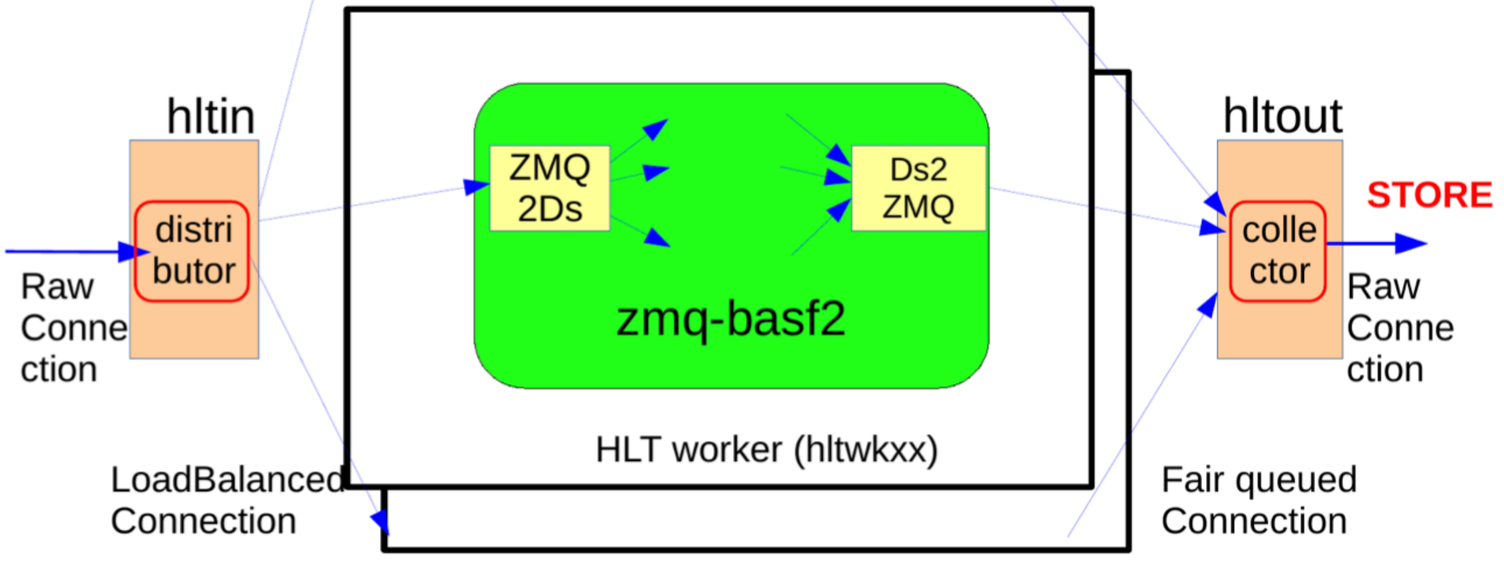


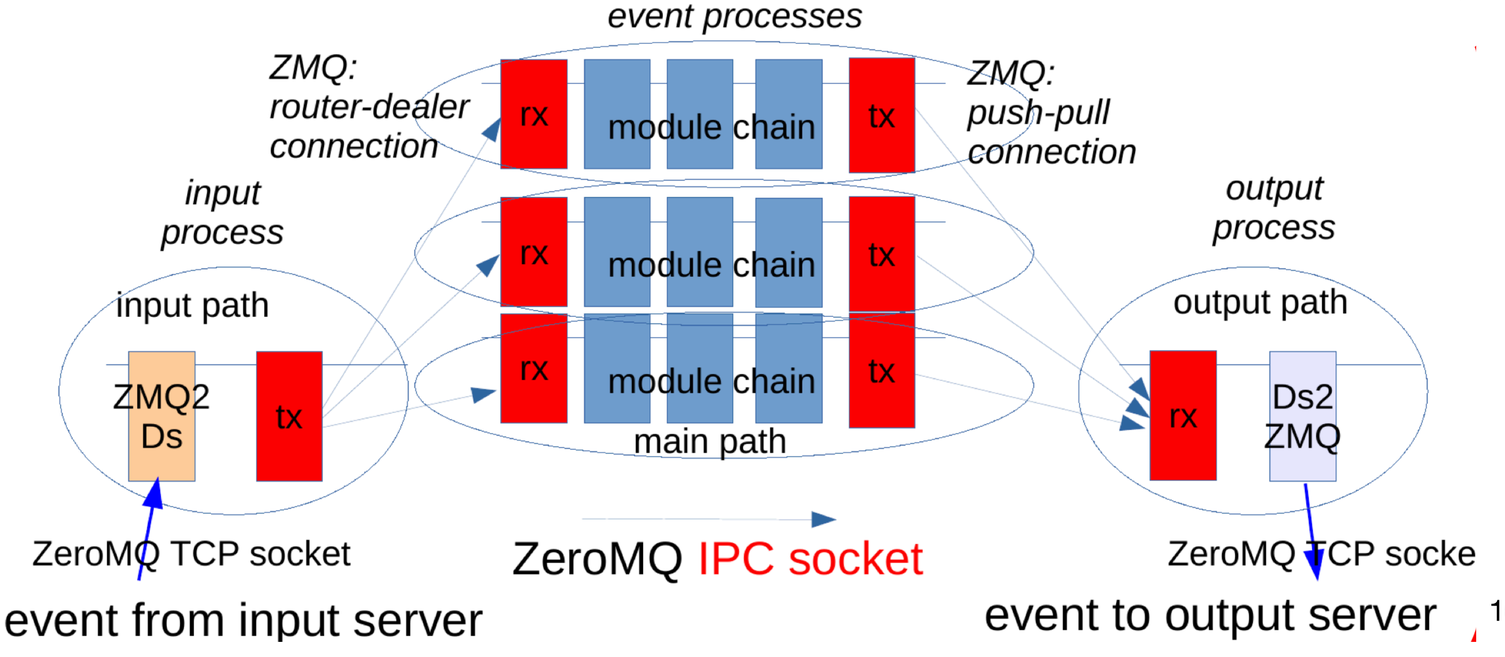
Features of high level trigger system

- An open-source package for the general message passing
- It supports N-to-1 and 1-to-N connection with a variety of connection style including load-balanced pipeline

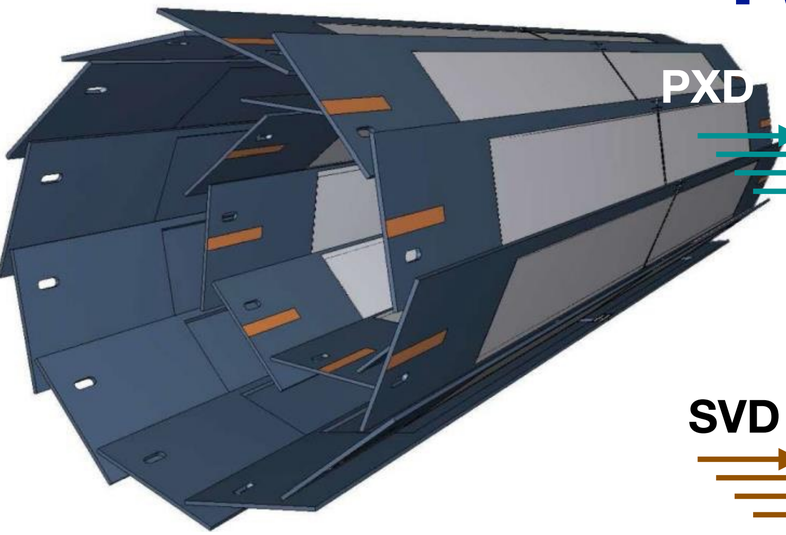
ØMQ

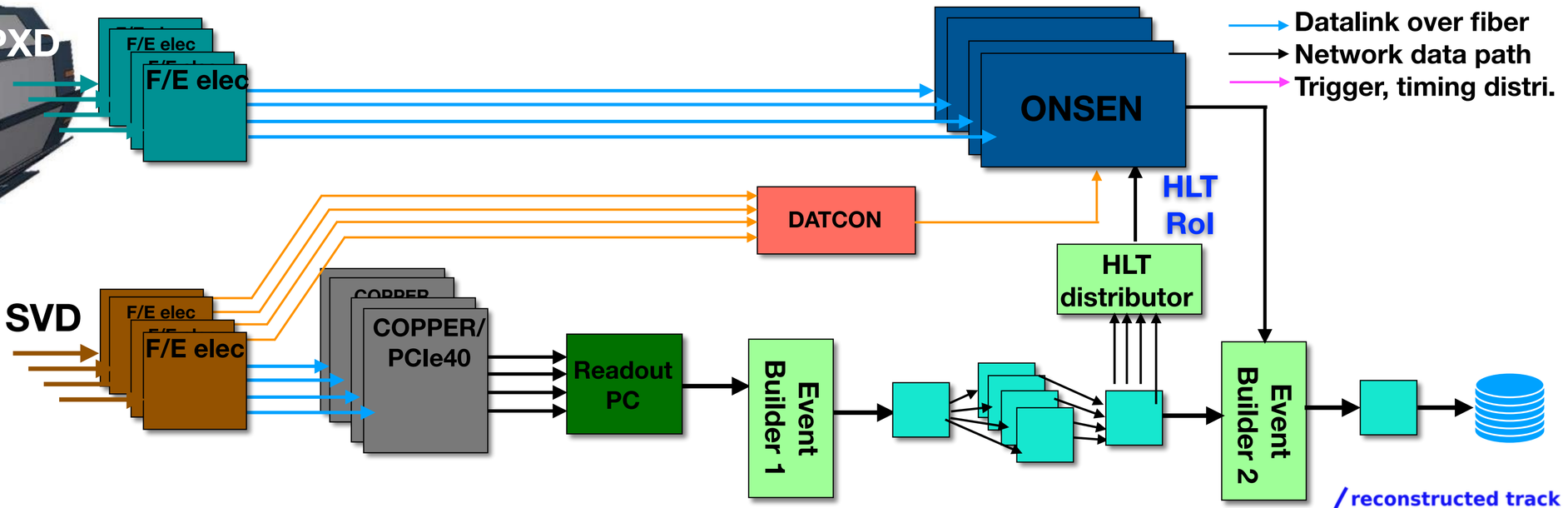
- Based on the event-by-event parallel processing implemented in Belle2 Analysis Framework (basf2)

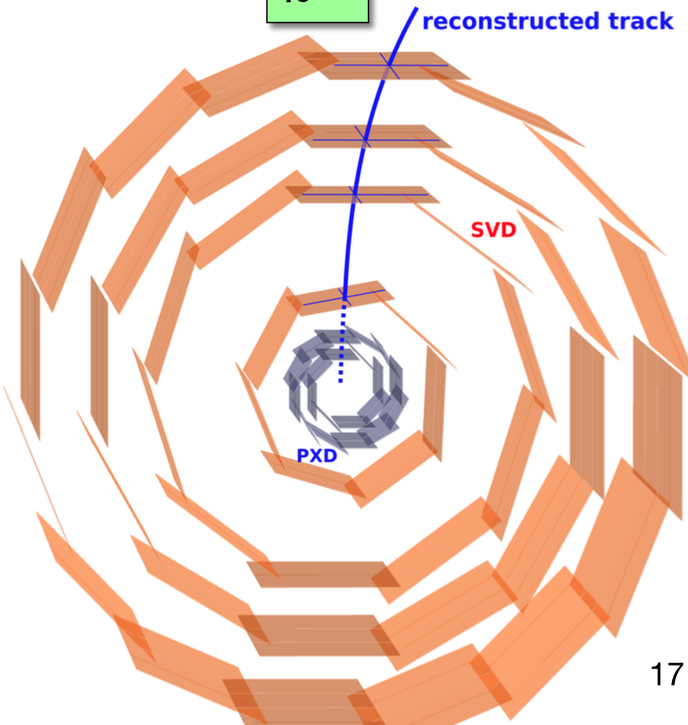




Region Of Interest (ROI)



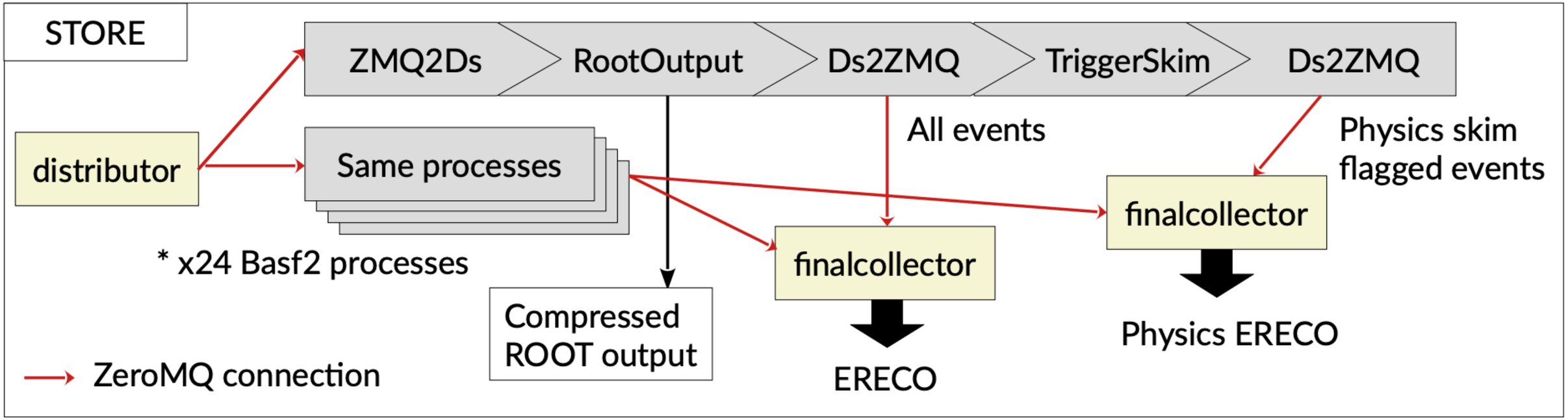




- PXD event size = 1MB/event, 10 times larger than the rest of detectors
- Region of interest method is effective to reduce the data size
- ROI
  - Tracking software running on HLT nodes
- PXD event data size reduced by 1/10 with ROI
  - In addition, trigger rate reduced by 1/3 with HLT ROI

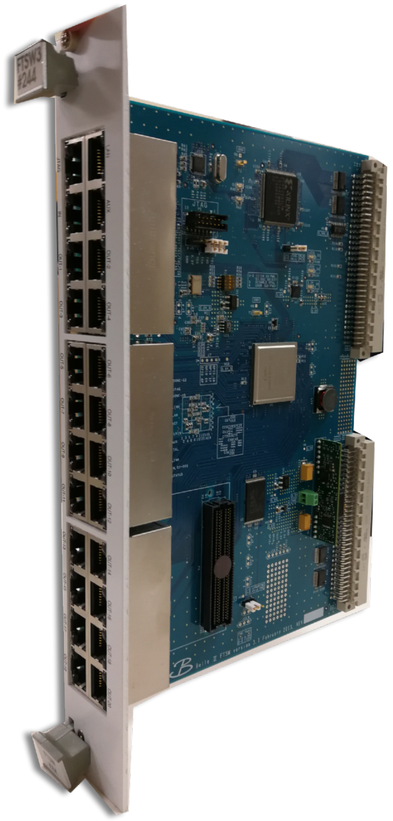
Storage system

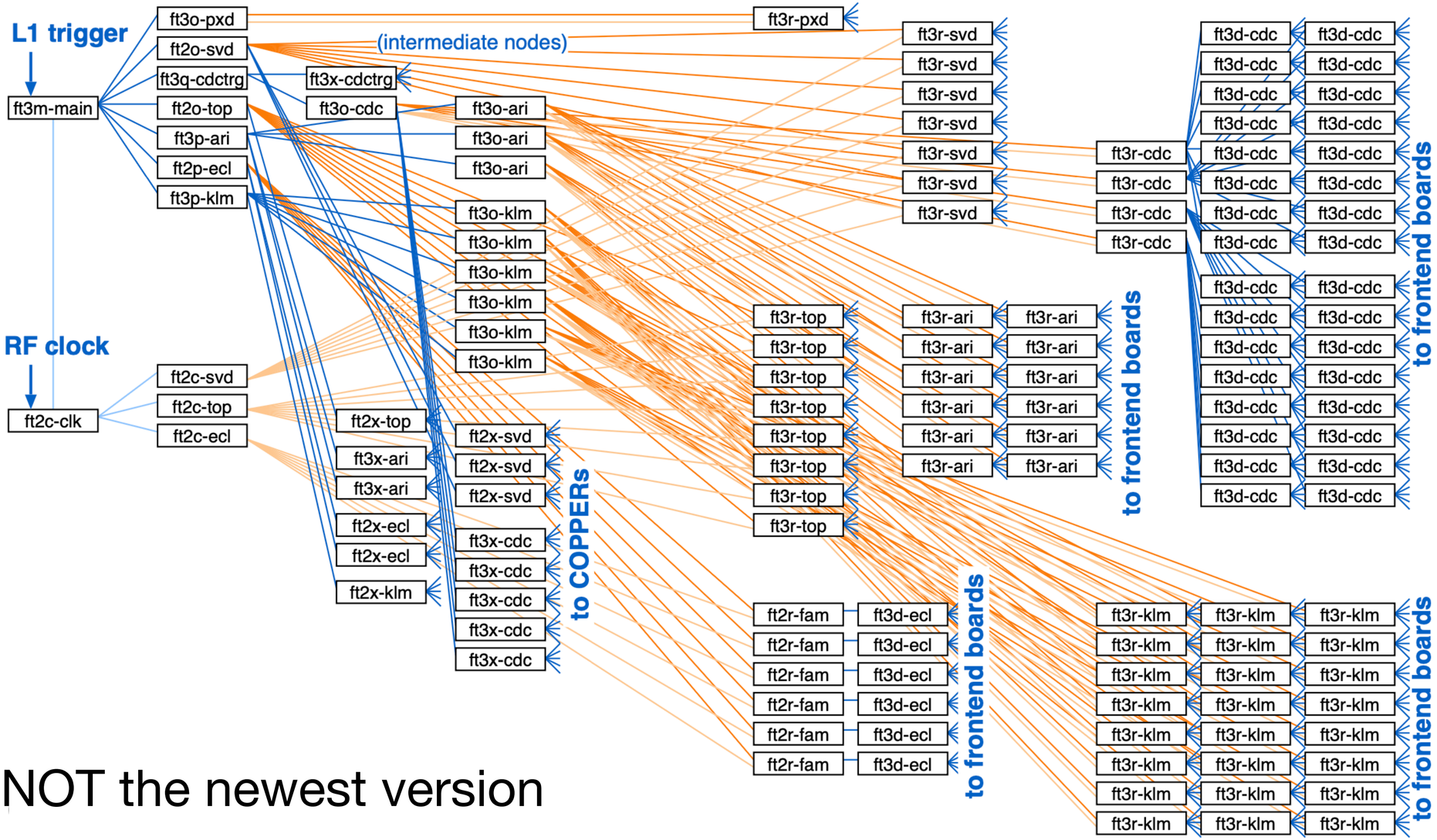
- STORE: online rawdata storage; a part of HLT farm
  - 32~48 threads CPU with three 40TB RAID units
- ERECO: express reconstruction system for online data quality monitoring (DQM), especially for vertex detectors and physics features
- ZeroMQ connections with HLT skim flags



Trigger Timing Distribution system

- TTD construct based on the Frontend Timing SWitch (FTSW) module (Virtex 5 FPGA)
- Serial protocol called b2tt (Belle II Trigger Timing) via LAN cable (254 Mbps)



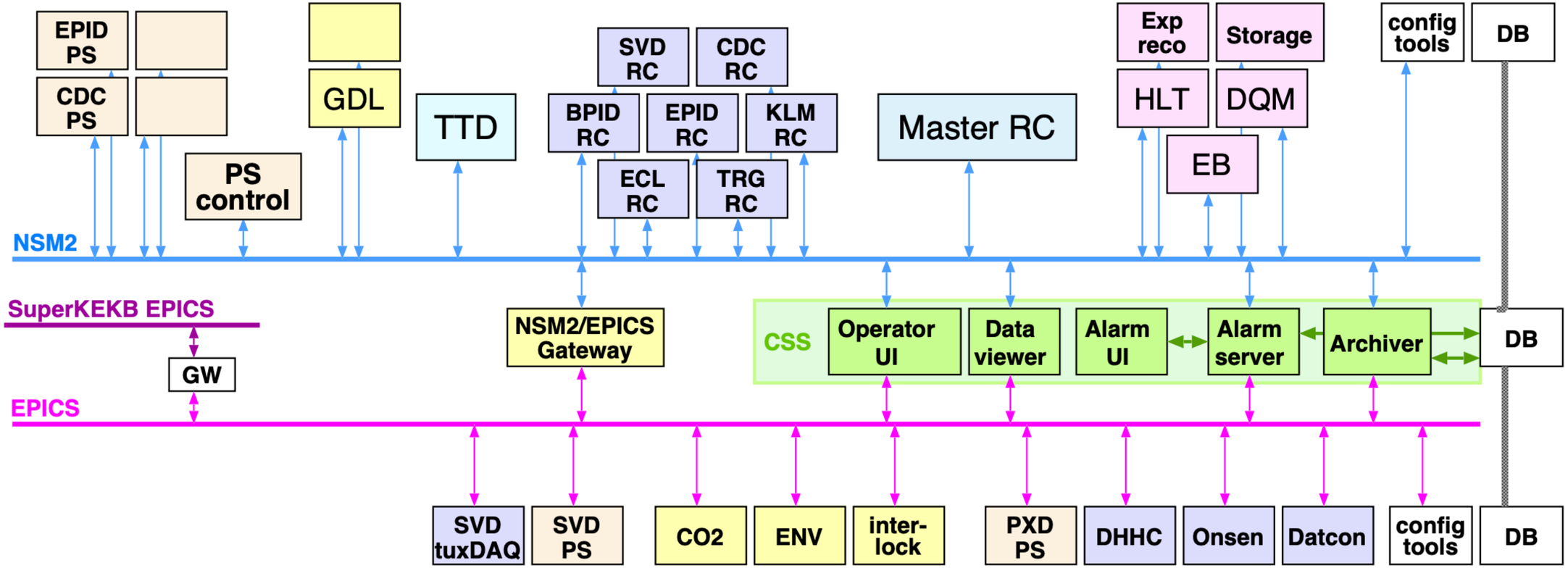


NOT the newest version

Functionally of TTD system

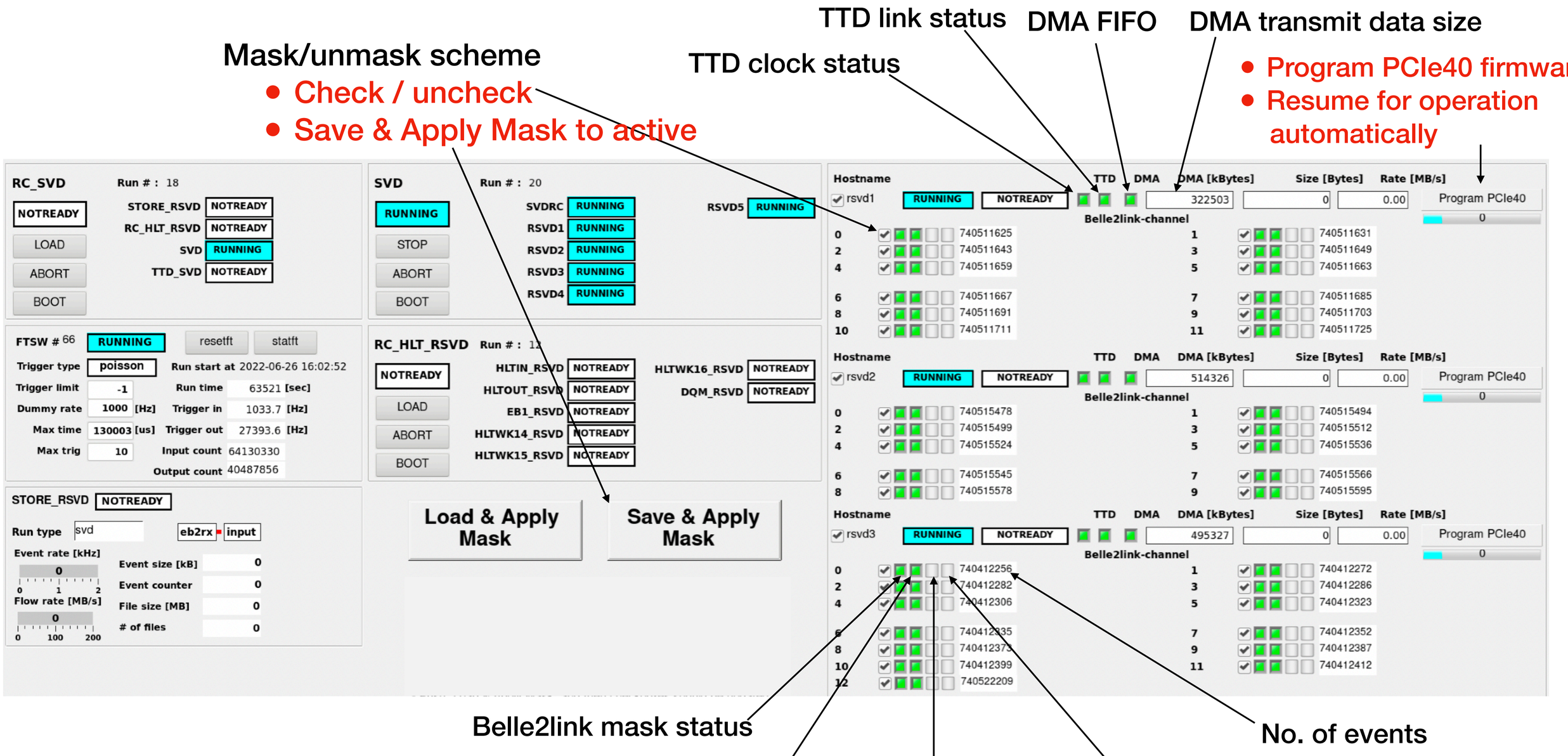
- Distribute **clock** to FEE, COPPER and trigger systems
- Distribute **utime**, **ctime**, **exp/run/subrun/event numbers** to FEE
- Distribute **trigger**, **trigger type** and **time-stamp** to FEE
- Distribute **revolution signal** and **injection veto** to FEE
- Various **reset** to **all** or **single** FEE
- **Deadtime generator** for flow-control to minimize deadtime
- Collect **back pressure** and **error** from FEE and COPPER
- Drive a **global** run, or drive multiple **local runs** in parallel
- **Mask** some of the connected FEEs, COPPERs and HSLBs
- **JTAG** programming of FPGAs on FEE
- General **event data** to EB, in addition to distributed event info

Slow control system



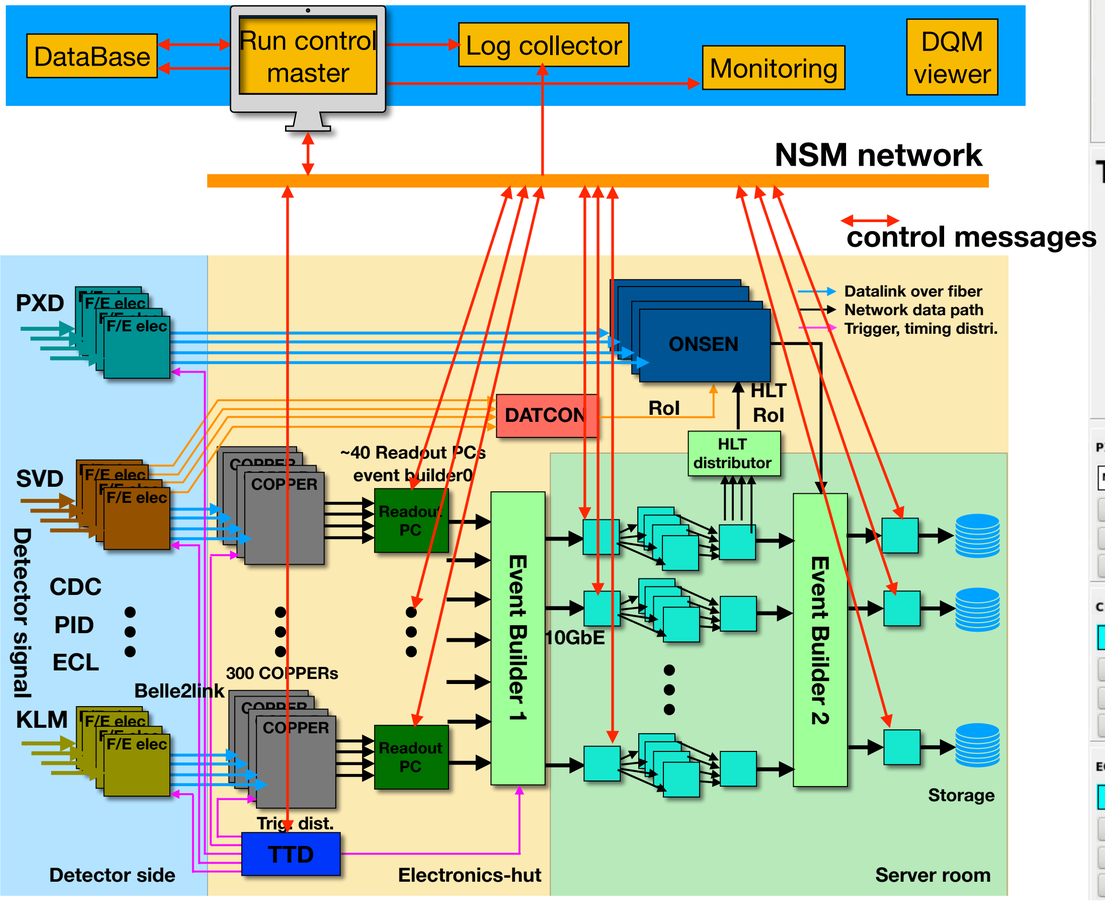
- NSM2 (network shared memory version 2) custom made package
- EPICS: widely used in SuperKEKB, etc.
  - **IOC** (input output controller) to provide **PVs** (process variables) and channel access, which is the way to access PVs
- GUI: previously designed w/ CS-Studio, just moved to Phoebus

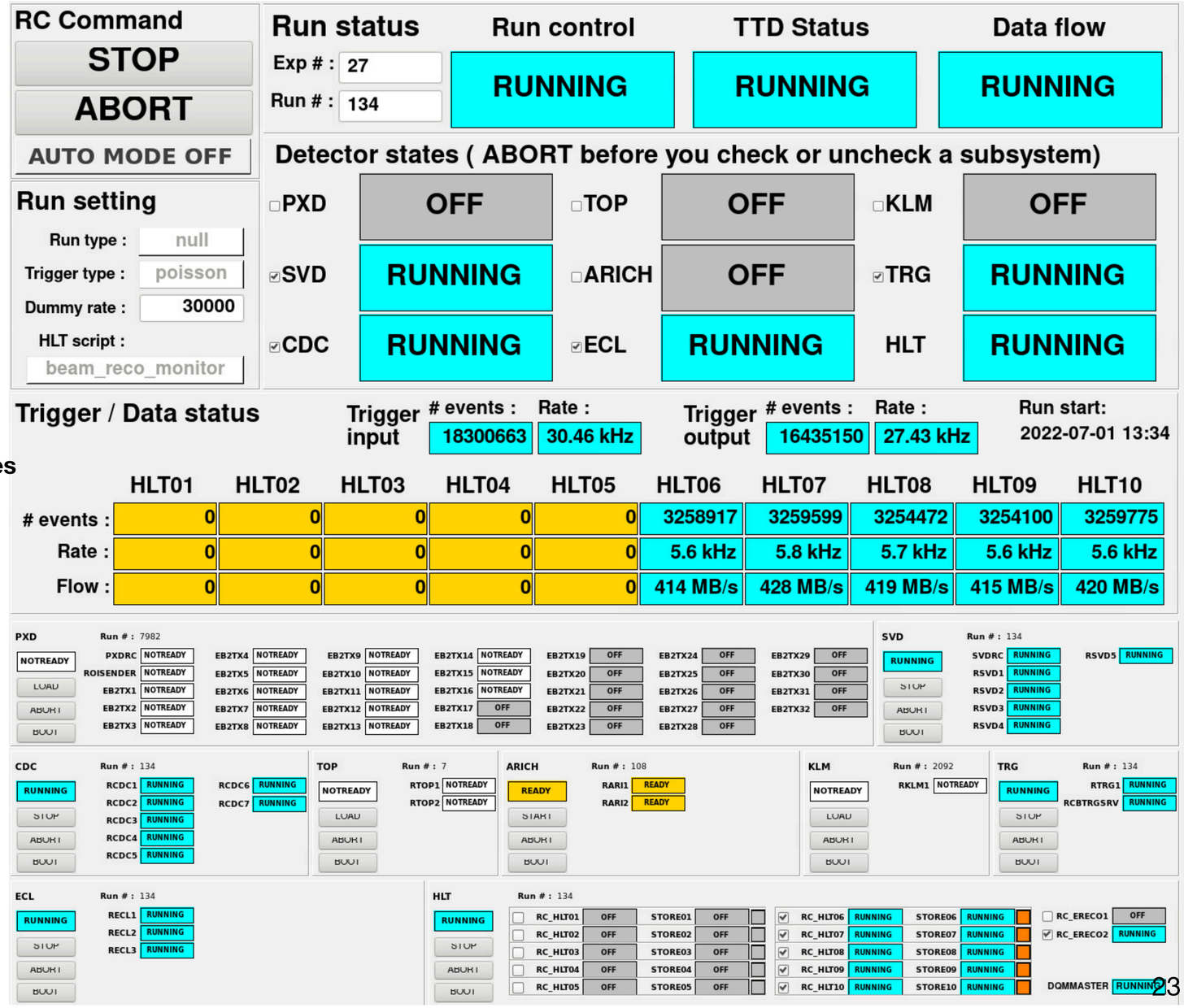
Operation panel for sub-detector



Global run control

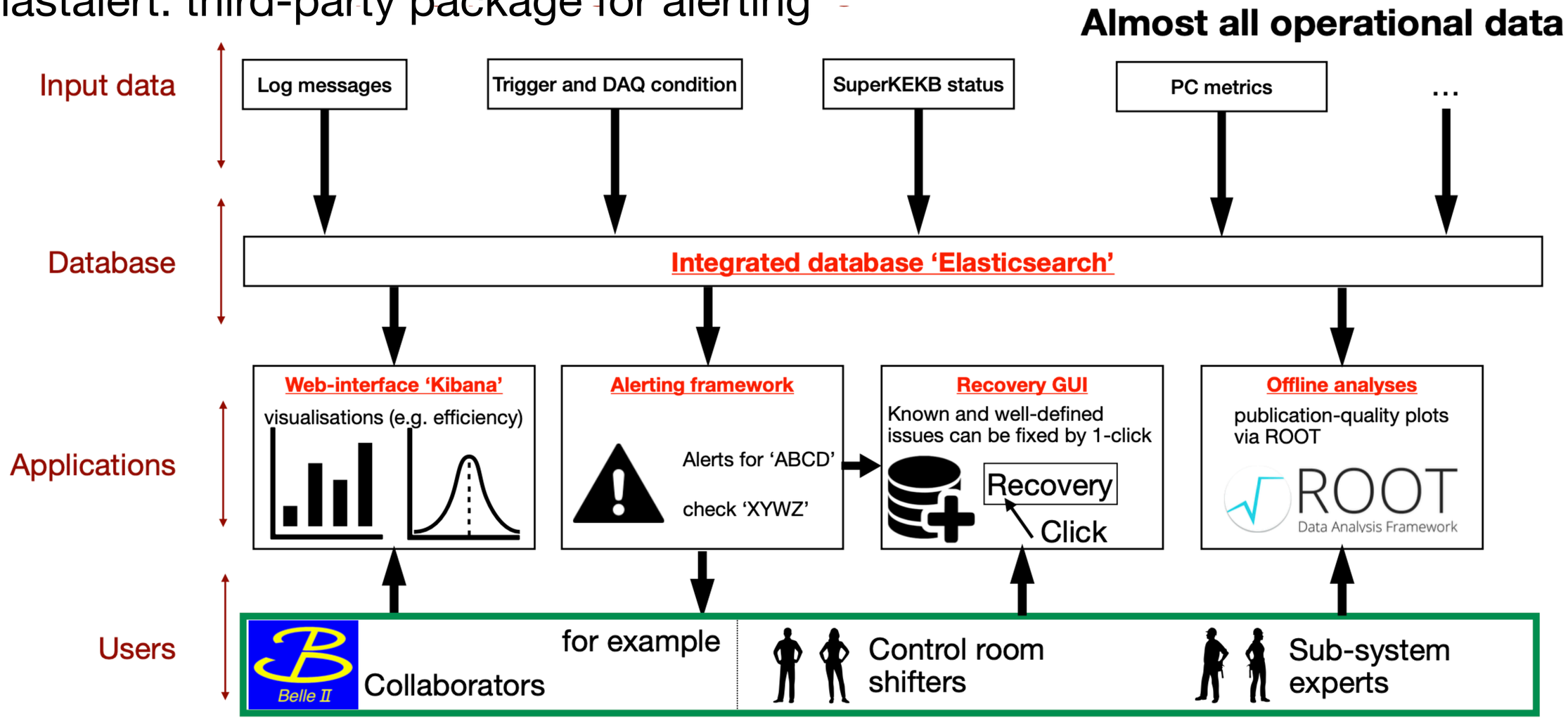
- Single master: centralization of all actions taken during the run
- TTD→PCIe40→ROPC→EB1→HLT→storage





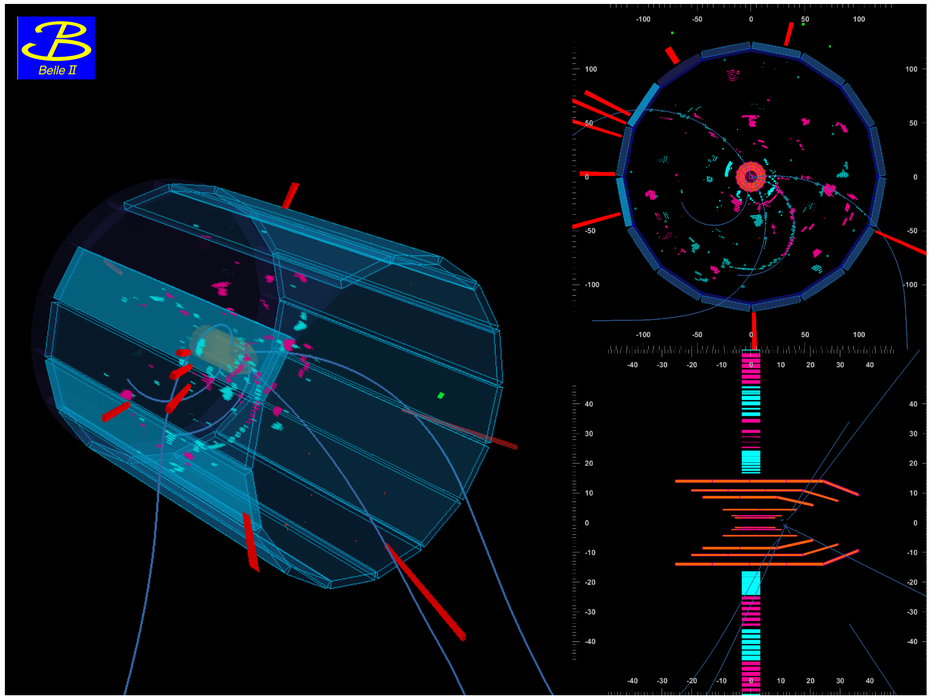
Monitoring system

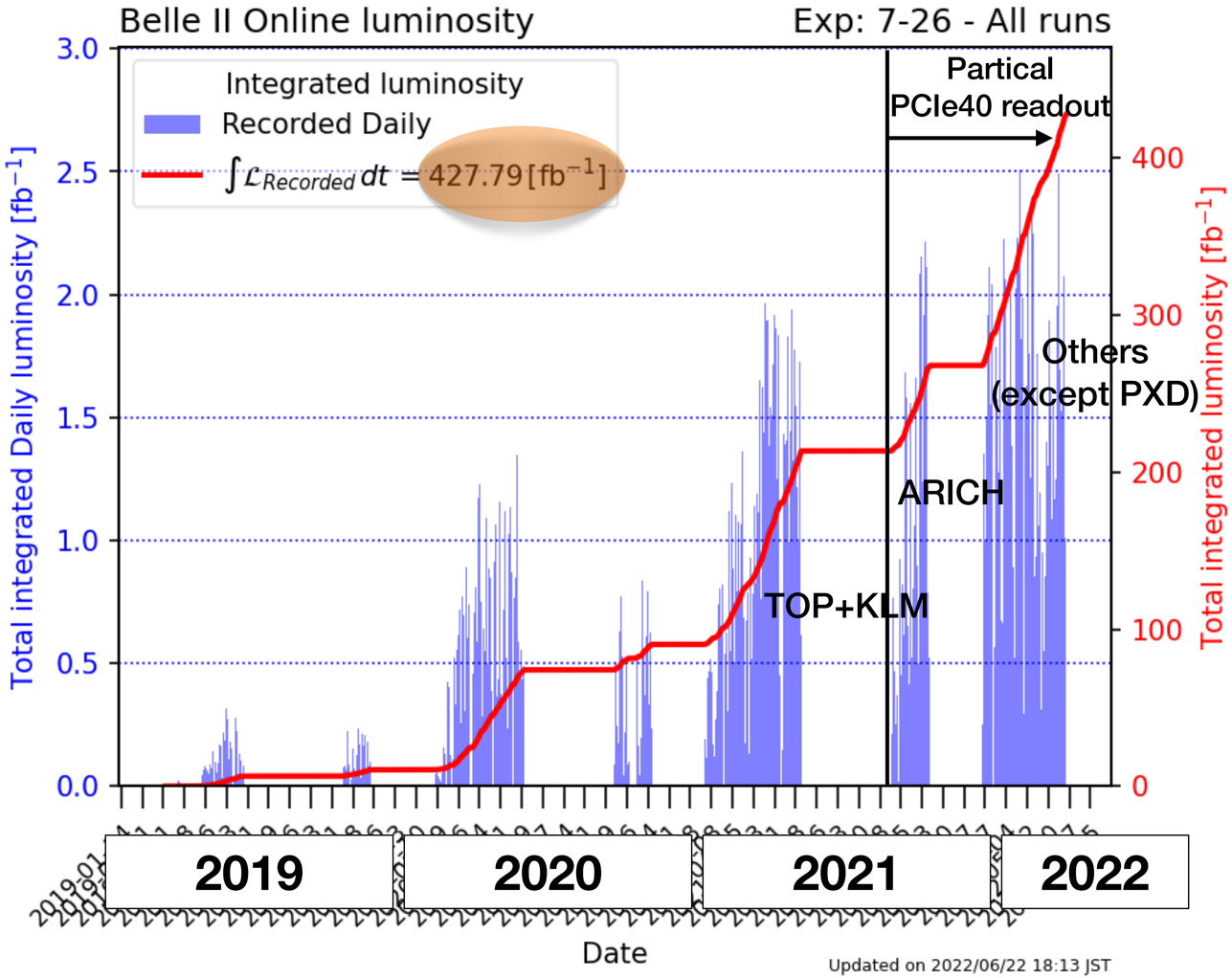
- Elasticsearch: distributed, JSON-based, and fast search and analysis engine
- Logstash and beats: data shipper/converter
- Kibana: web-interface for visualisation
- Elastalert: third-party package for alerting



Control room

Operation status and integrated luminosity





- DAQ upgrade succeed

Belle II data taking efficiency ~90%

- ~427 fb^{-1} till now (Belle: 1 ab^{-1})
- Long shutdown 1 (LS1: 2022-2023), resume data-taking from 2024

Summary

- Belle II DAQ system was designed to handle 30 kHz level 1 trigger
- A unified readout system was designed based on COPPER/PCIe40 module for common readout of sub-detectors (except PXD)
- A unified trigger timing distribution for a pipeline trigger control, etc.
- Region Of Interest (ROI) is used to reduce PXD data (1/10)
- 13 HLT units (6212 cores) prepared for data processing, based on the ZeroMQ frame
- Storage and express reconstruction build also ZeroMQ frame
- Efficient monitoring system developed via Elastic stack
- Data-taking was running stably with new PCIe40 based readout system
- Belle II data-taking efficiency close to 90%

Backup