PCI-express based high-speed readout for the Belle II DAQ upgrade

Qi-Dong ZHOU (IAR/KMI, Nagoya Univ.)
S. Yamada, P. Robbe, D. Charlet, R. Itoh, M. Nakao,
S.Y. Suzuki, T. Kunigo, E. Jules, E. Plaige, M. Taurigna,
H. Purwar, D. Biswas, Y.-T. Lai, O. Hartbrich, M. Bessner

22nd Virtual IEEE Real Time Conference
12-23 Oct. 2020

Contents

- Introduction of Belle II DAQ and DAQ readout system
- Motivation of upgrading Belle II readout system
- PCIe40 module
- Firmware and software development
- Performance of new readout system
- Schedule and plan of DAQ upgrade

Luminosity frontier: SuperKEKB

Beam current: KEKB x 1.7

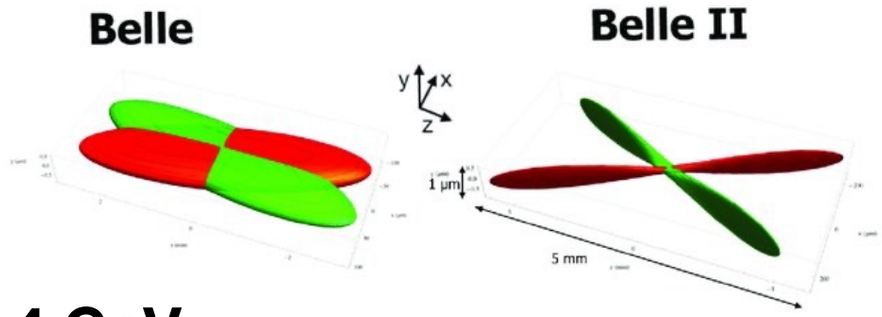
$\mathcal{L}$ = 6.5 x 10^{35} cm^{-2} s^{-1}

Luminosity

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

Beam squeeze: KEKB / 20

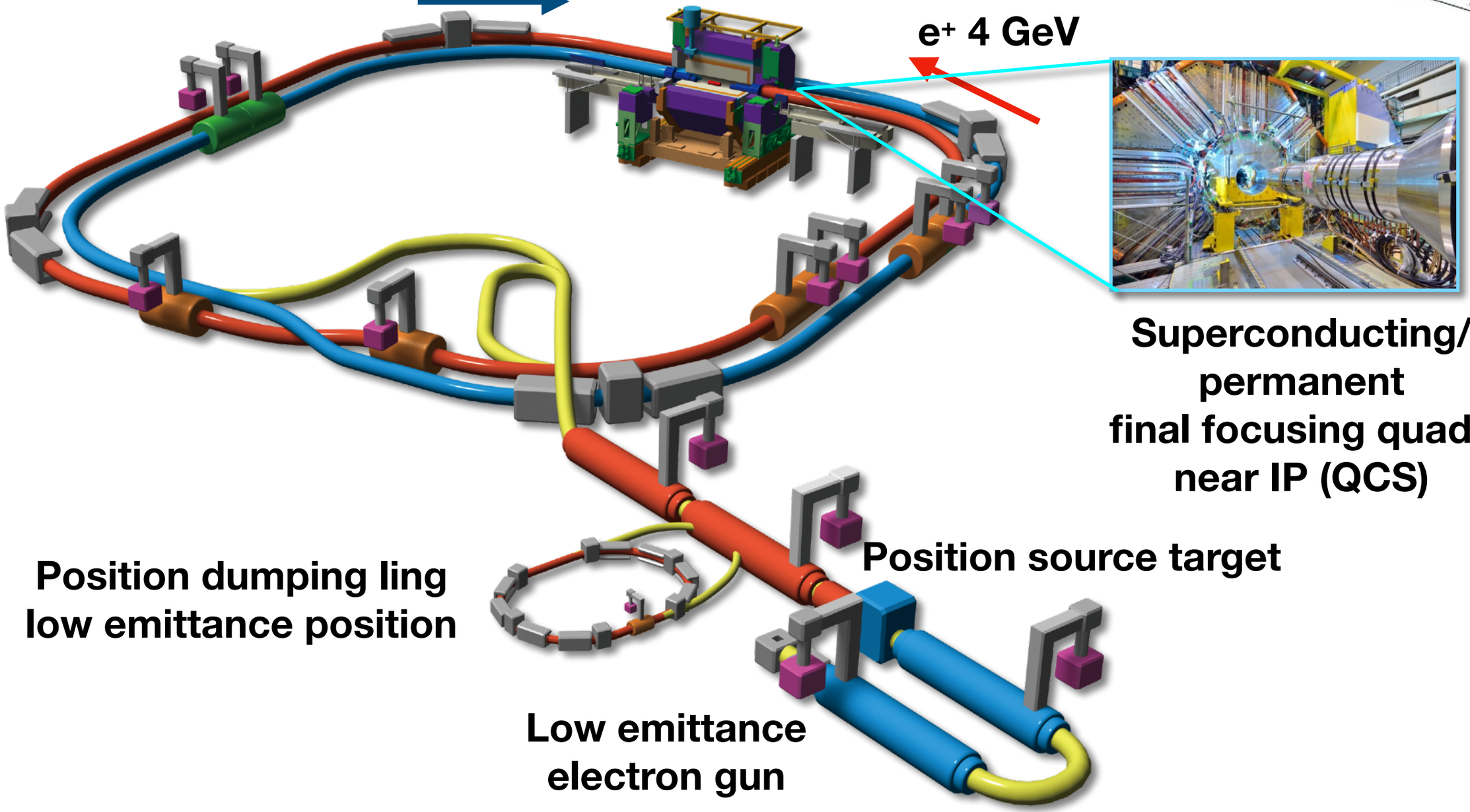
Nano beam scheme



e⁻ 7 GeV

Belle II detector

e⁺ 4 GeV



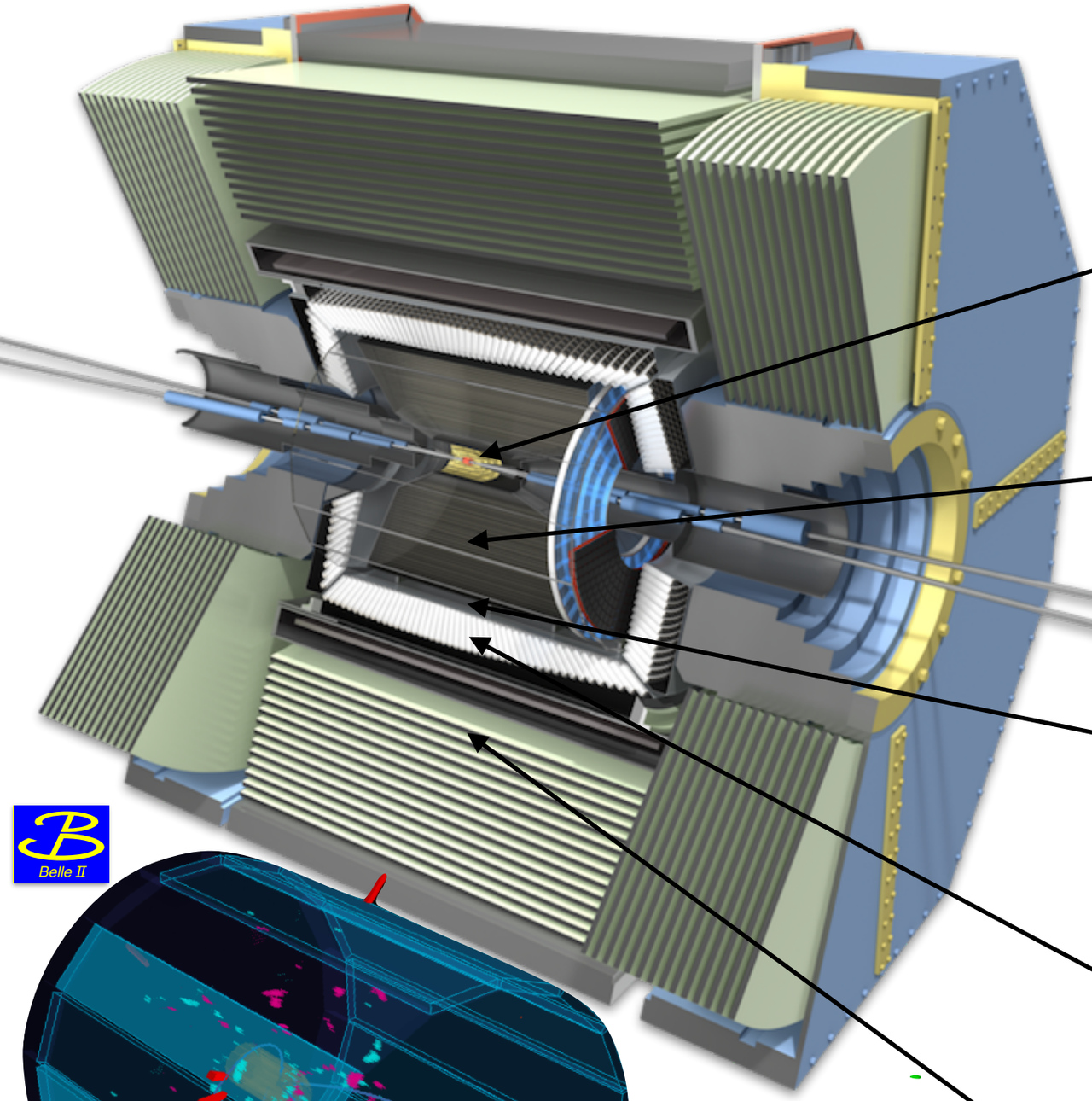
Superconducting/ permanent final focusing quads near IP (QCS)

Position dumping ling low emittance position

Position source target

Low emittance electron gun

Belle II detector



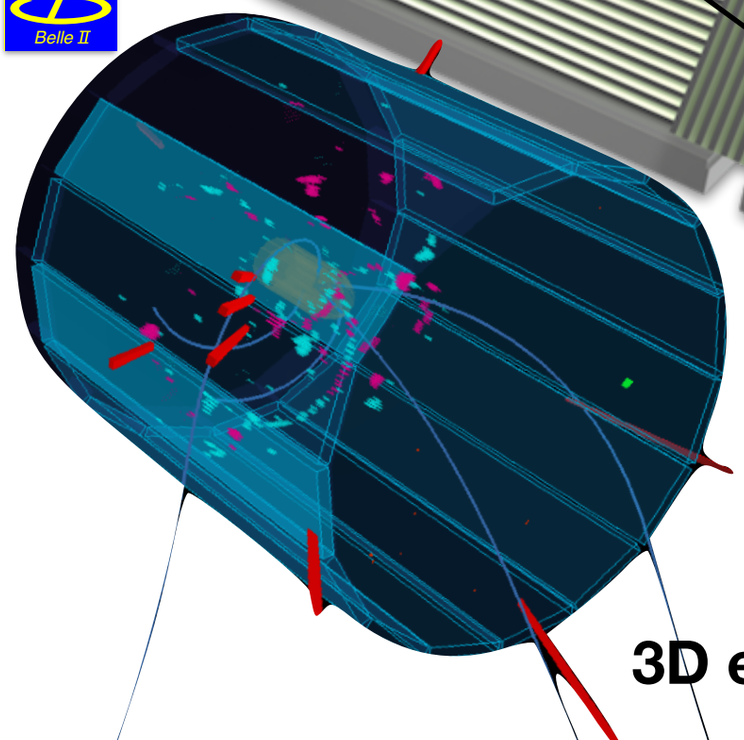
Vertex detector (VXD)
- Inner 2 layers: pixel detector (PXD)
  - 2nd layer not fully installed
- Outer 4 layers: strip sensor (SVD)

Central Drift Chamber (CDC)
- He (50%), C_2H_6 (50%), small cells, fast electrics

Particle ID detector
- Barrel: Time-Of-Propagation counters (TOP)
- Endcap: Aerogel RICH (ARICH)

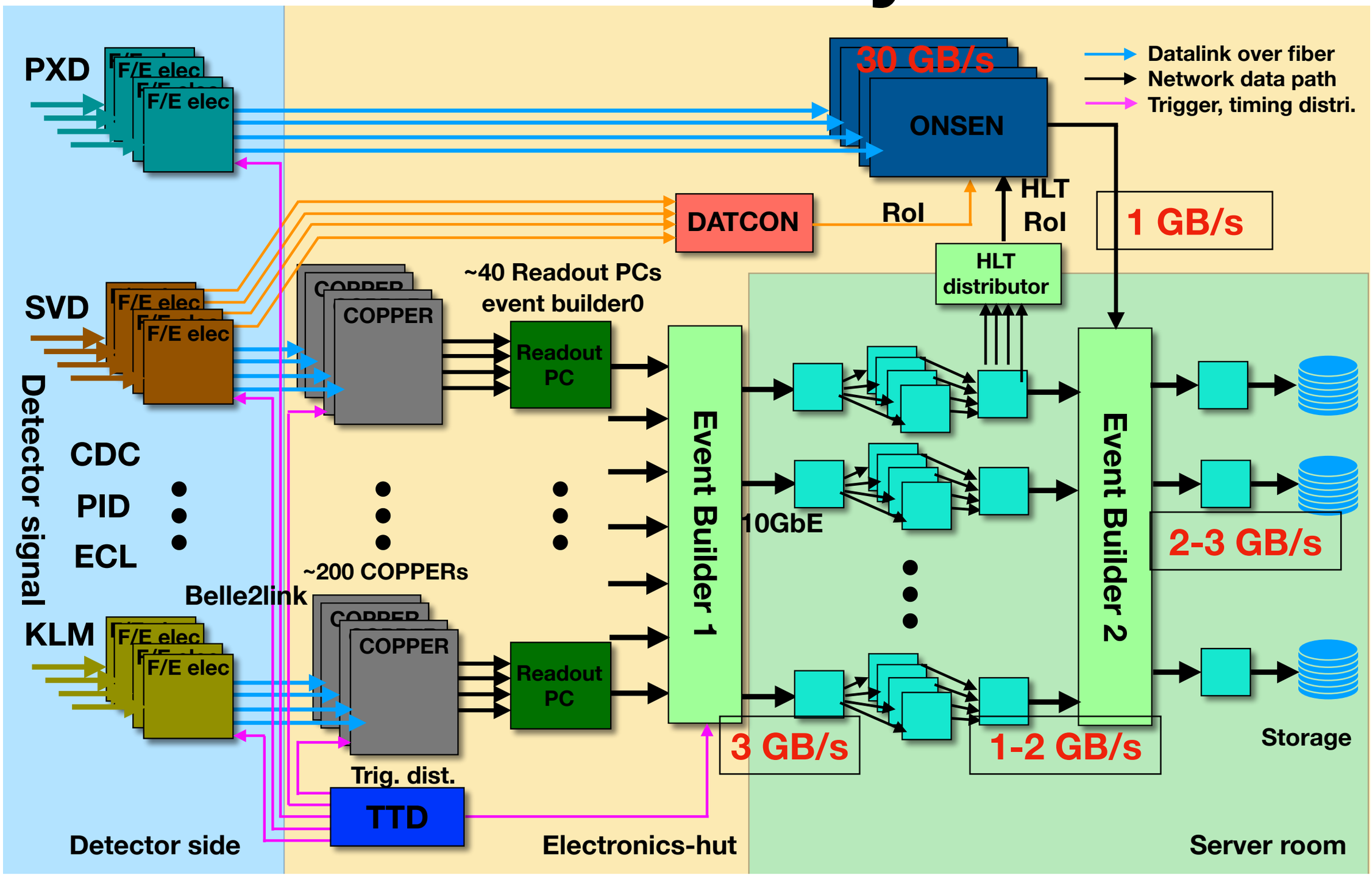
ElectroMagnetic Calorimeter (ECL)
- CsI(Tl) + waveform sampling

K_L/μ detector (KLM)
- Outer barrel: Resistive Plate Counter (RPC)
- Endcap/inner barrel: Sci.



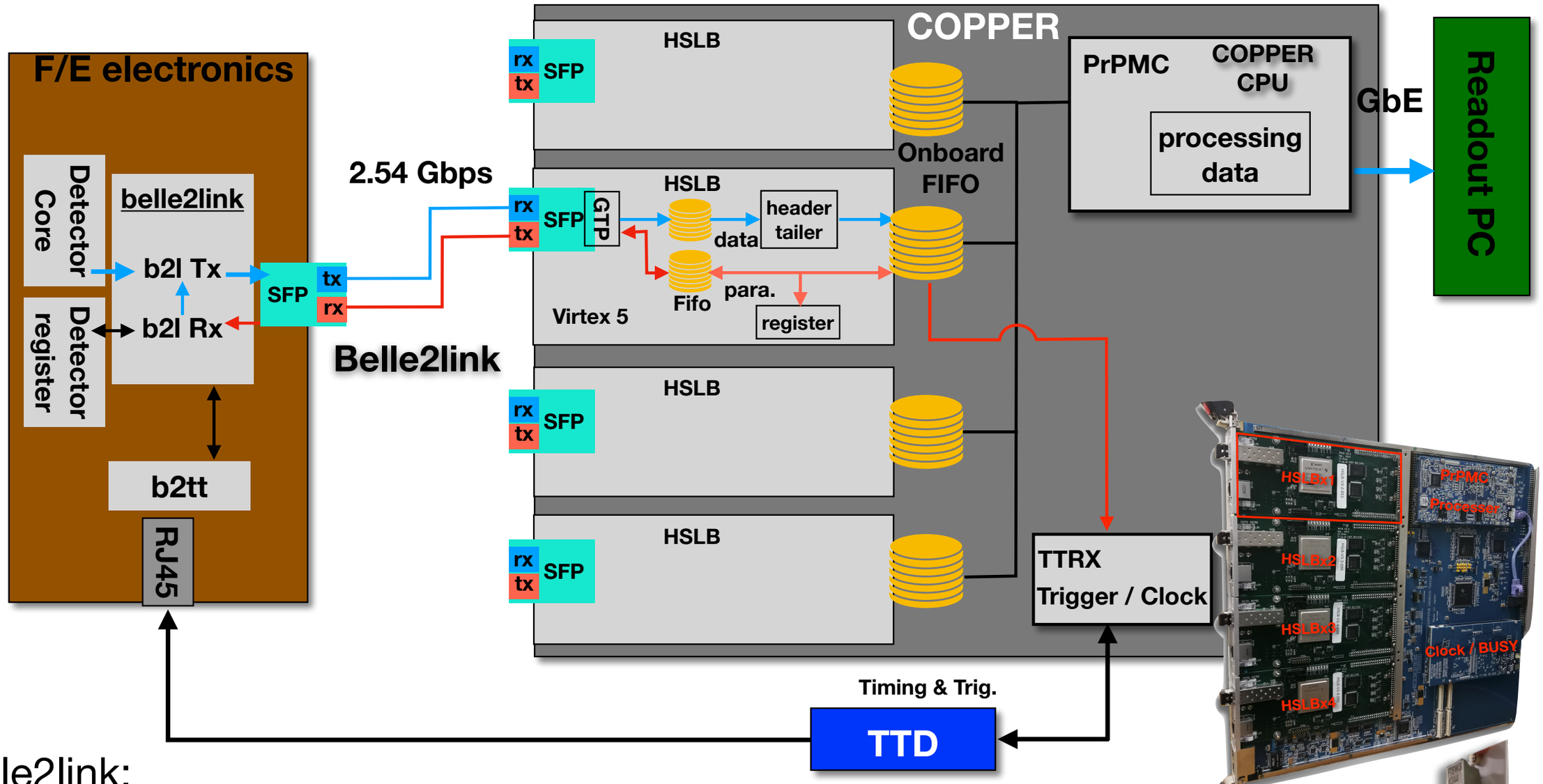
3D event display

Belle II DAQ system



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

Readout system



Belle2link:

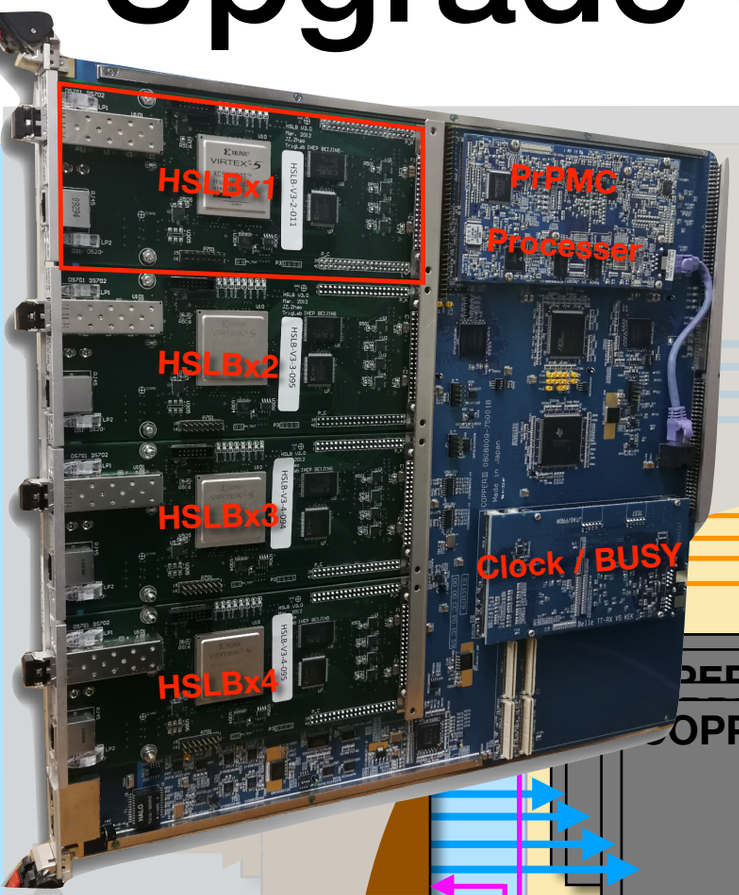
Unified high speed optical link (2.54Gbps) connected Front-End Electronics and DAQ readout board (COPPER-HSLB), data transmission based on Rocket I/O.

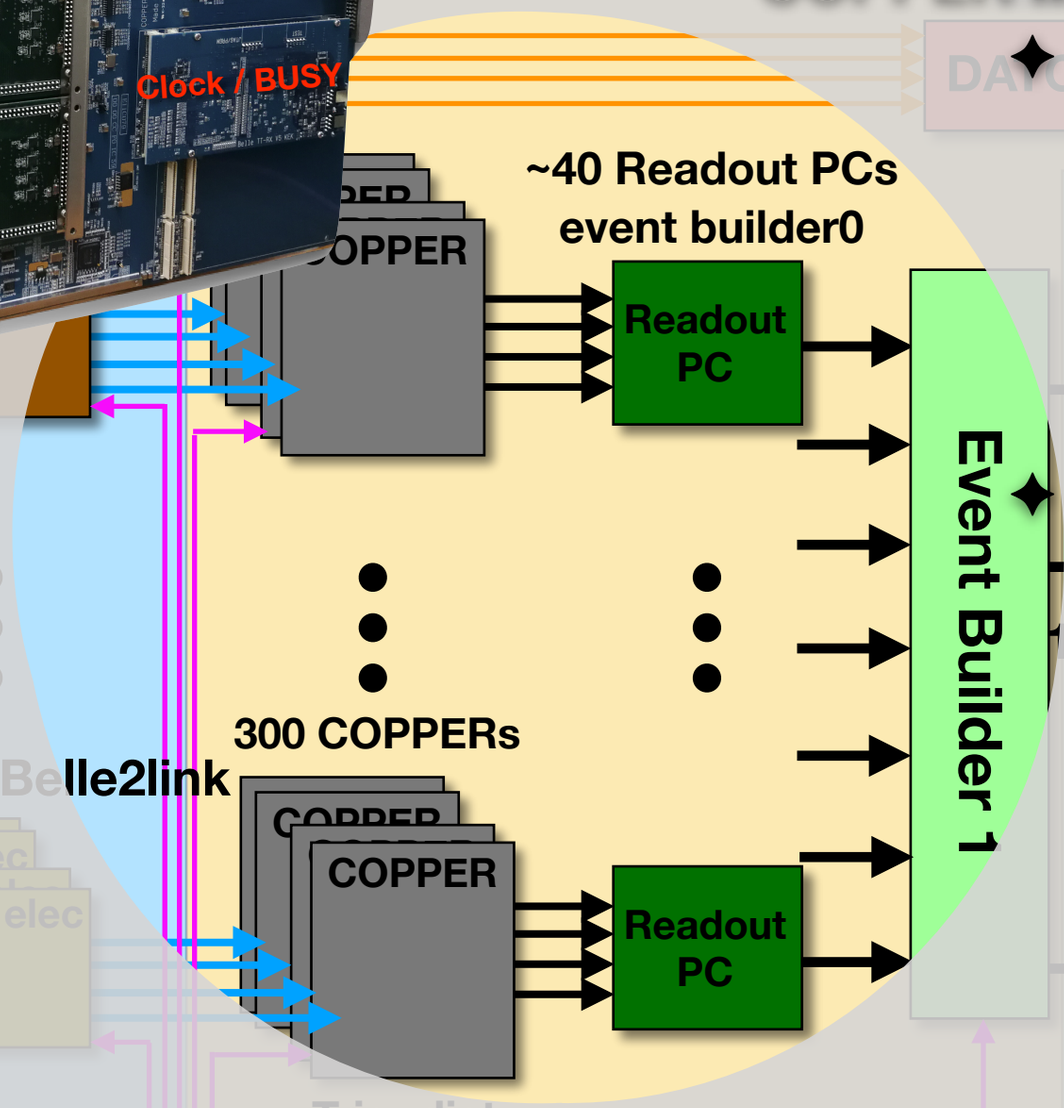
Functionalities of readout system

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

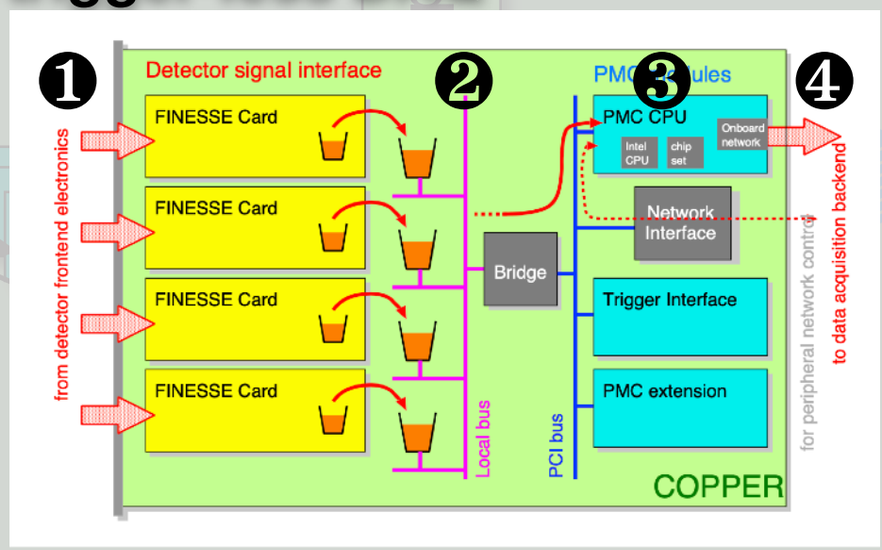
FTSW module for TTD

Upgrade of Belle II readout system





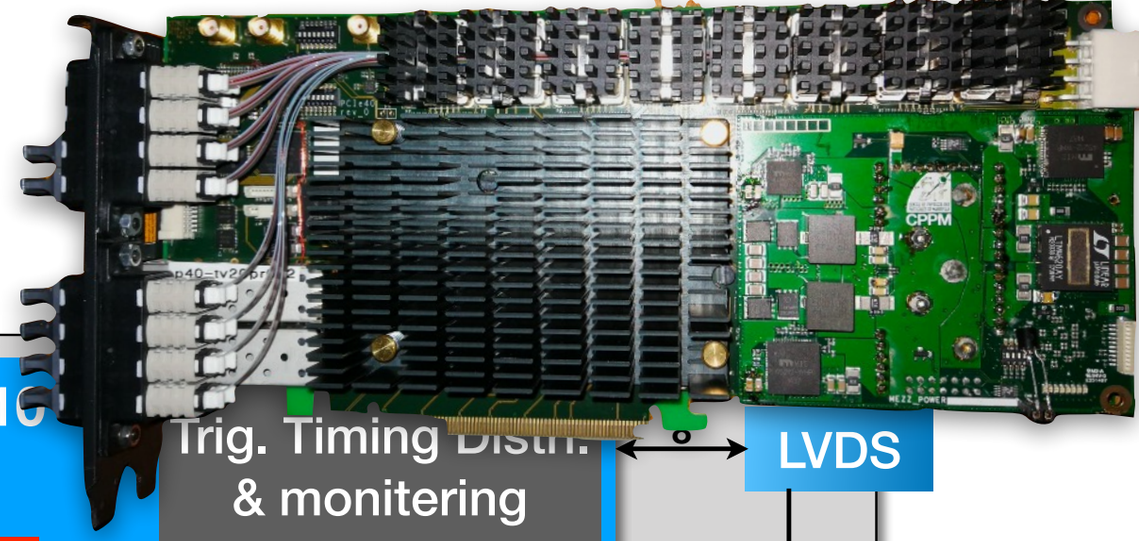
- **Difficult to maintain during the entire Belle II operation period (10 years)**
  - **Four different boards**
  - **Broken parts are increasing (PrPMC CPU, chipset, COPPER III LAN …)**
- **Limitation to improve DAQ performance**
  - **Bottlenecks of COPPER**
    - **CPU usage**
    - **Data transfer speed (1 Gbps)**
  - **Bottleneck of network output of ROPC (1 Gbps)**
- **Future possibility**
  - **Luminosity & background situation changed**
  - **For a trigger-less DAQ**

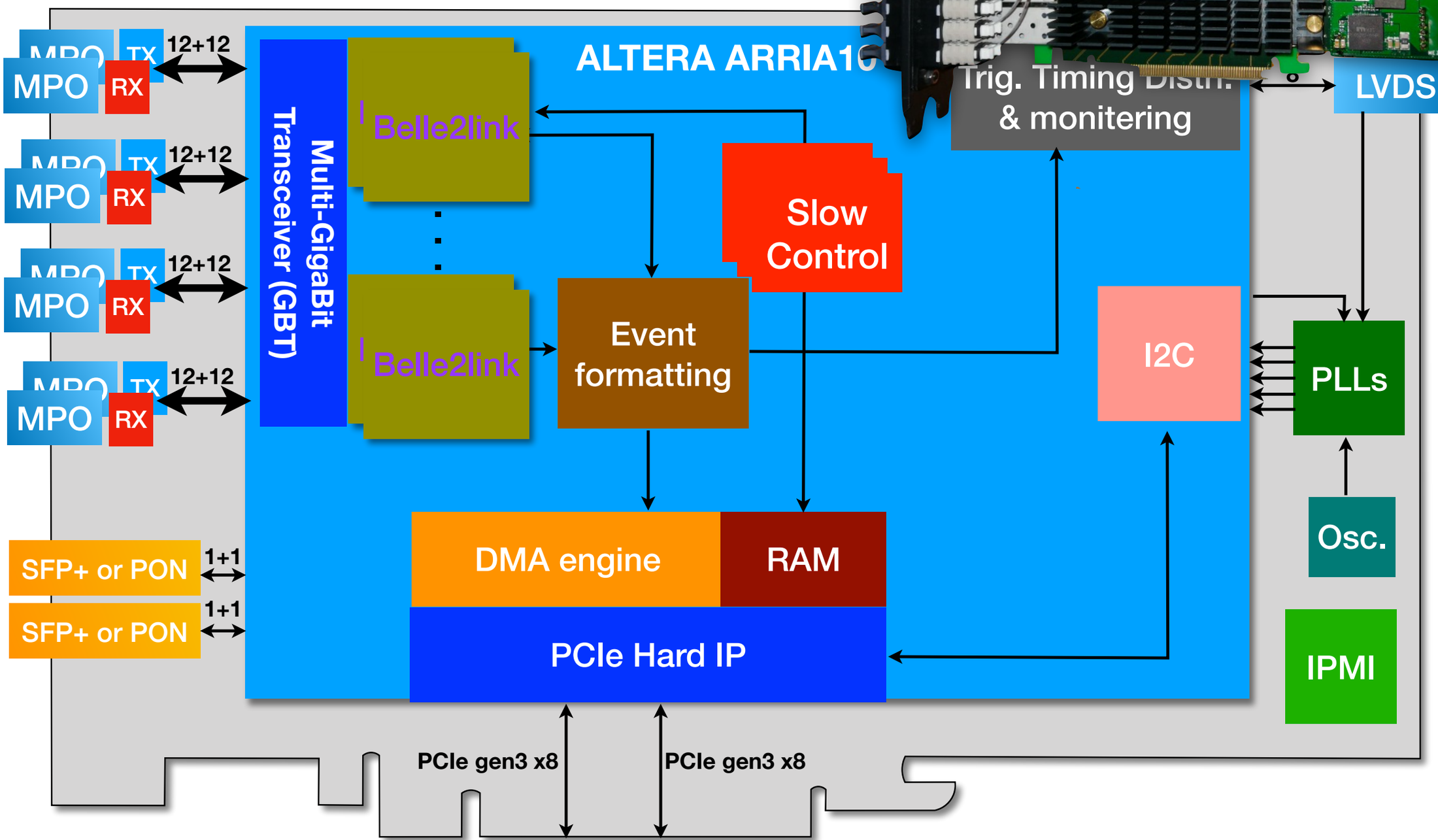


❶ ❷ ❹

2.54 Gbps > 1 Gbps ~ 1 Gbps

PCIe40 module

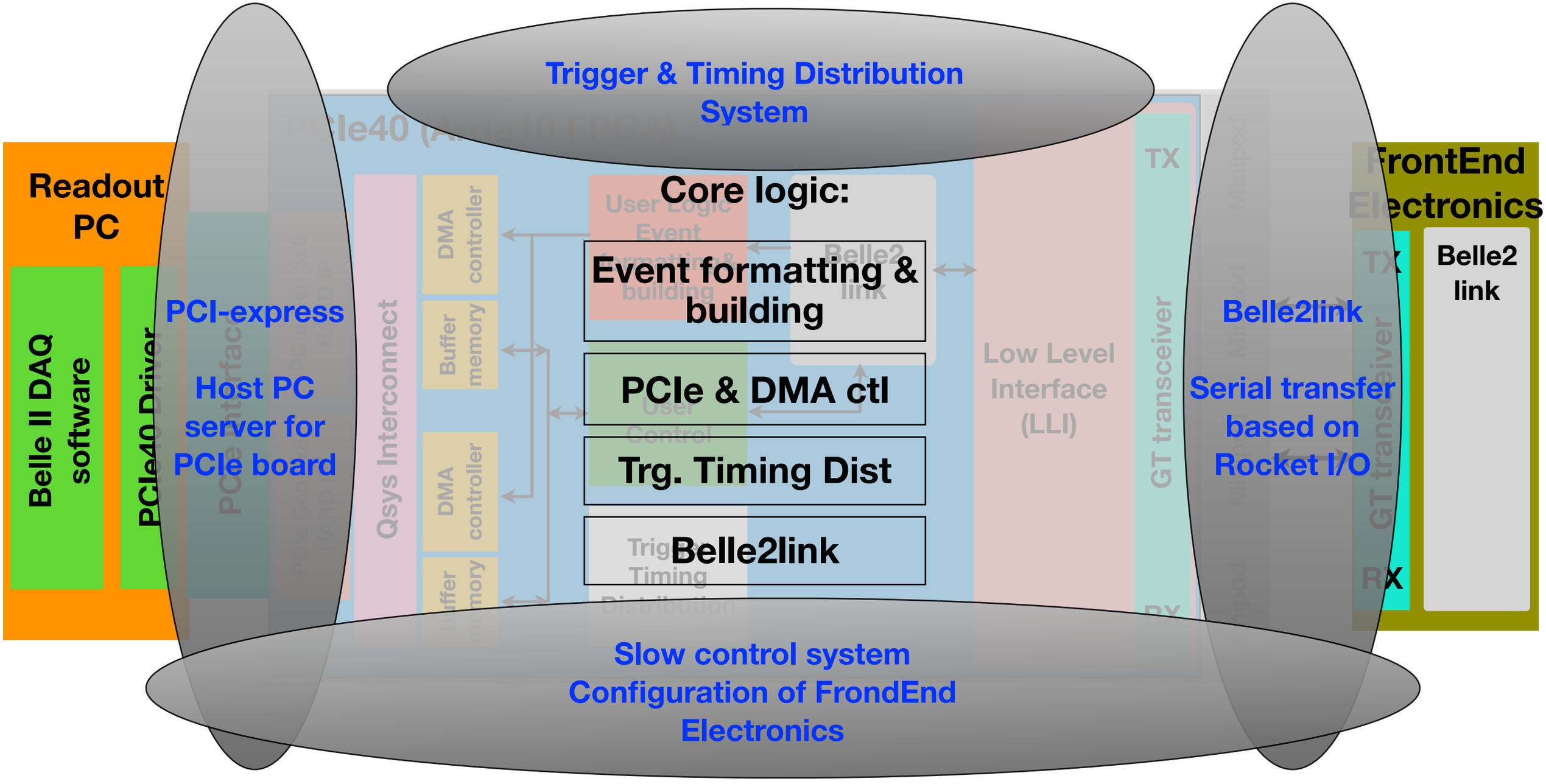




19 PCIe40 boards to replace 203 COPPERs

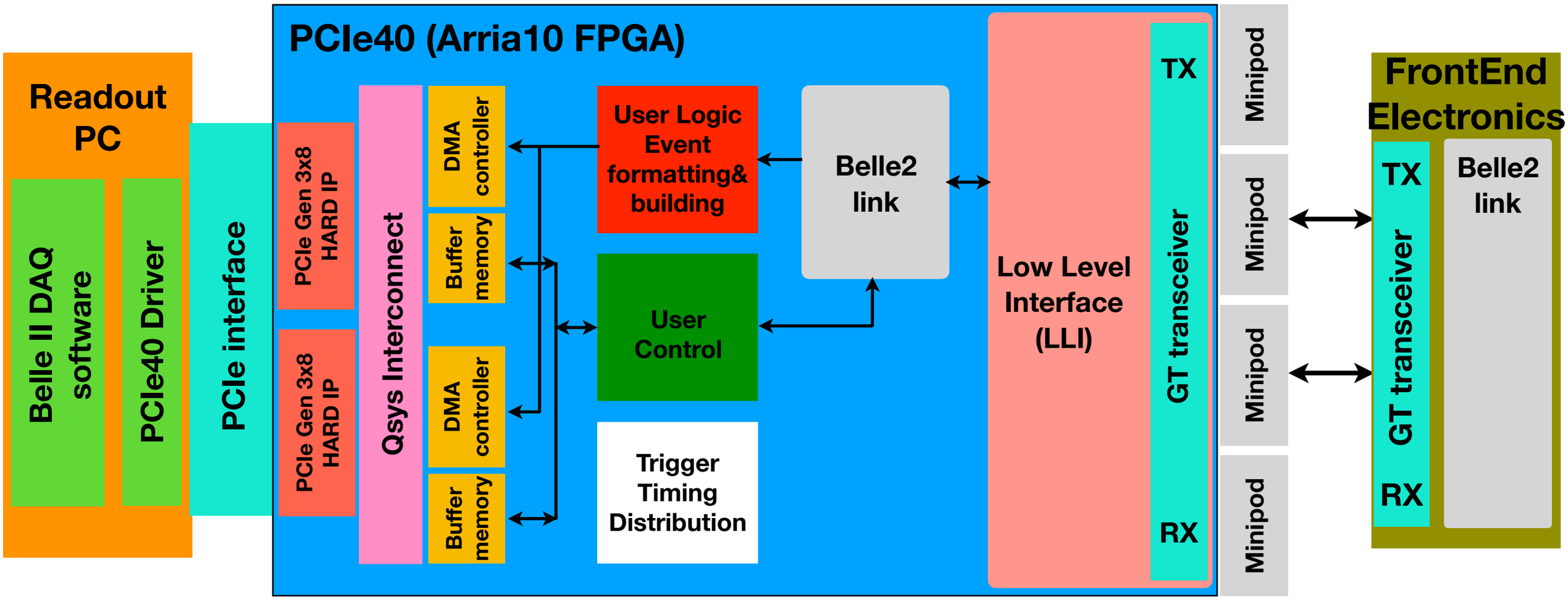
Board	FPGA family	Optical links	PC interface	Experiment
PCIe40	ALTERA Arria10	24 pairs x2	2 PCIe Gen3 x 8	LHCb, ALICE

Upgrade for new readout system



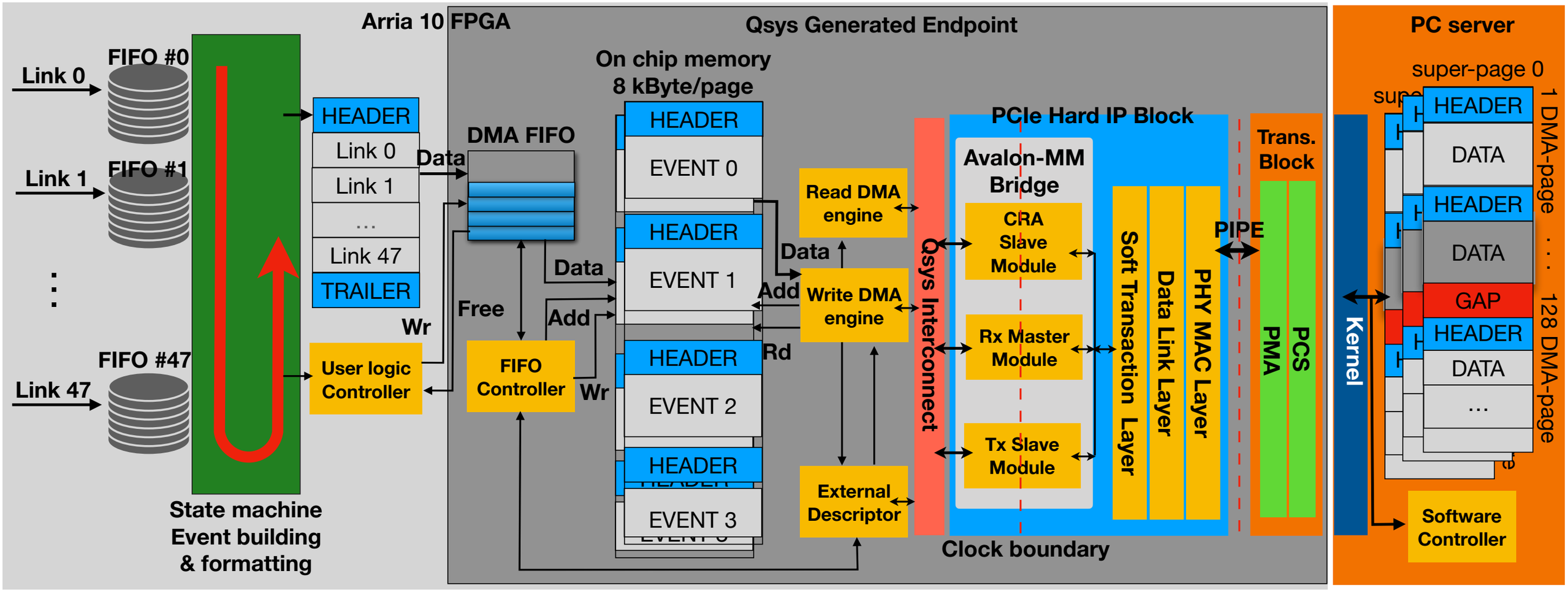
- Upgrade of readout system will keep the modification as small as possible, for the system connected.
- No major modification required for hardware and firmware of the sub-detectors' systems.
- Software for slow control and data readout need some upgrades

Firmware development



- Low Level Interface was designed for user's control and initialization proposes, enables a generic management of functions required by the PCIe40 board
- Belle2link protocol was basically kept as same functionality, but redesigned
- Event building and formatting was newly added based on the FPGA logic, it was done by COPPER on board CPU
- Slow control logic was kept, but moved most of the parts controlled by software
- TTD kept using b2tt protocol, new design to handle 48 links
- PCIe based DMA architecture was designed based on the Qsys with an external DMA descriptor controller apart from DMA engine.

Data processing



- 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
- Pulse trigger rate: 470 kHz (times 8 kBytes)
  - Data transmission rate: 39 Gbits/s <—> Theoretical maximum data rate is 50 Gb/s can eventually be increased to 100 Gb/s
  - 10 % of event detect back-pressure.
- Pulse trigger rate: 260 kHz, 21 Gbits/s, no event lost.

Test bench

b3ropc02, host server

topslc02,
control server

COPPER
dummy FEE

FTSW

Fully instrumented
TOP module

Low Voltage
Power Supply (Manual)

CDC FEE

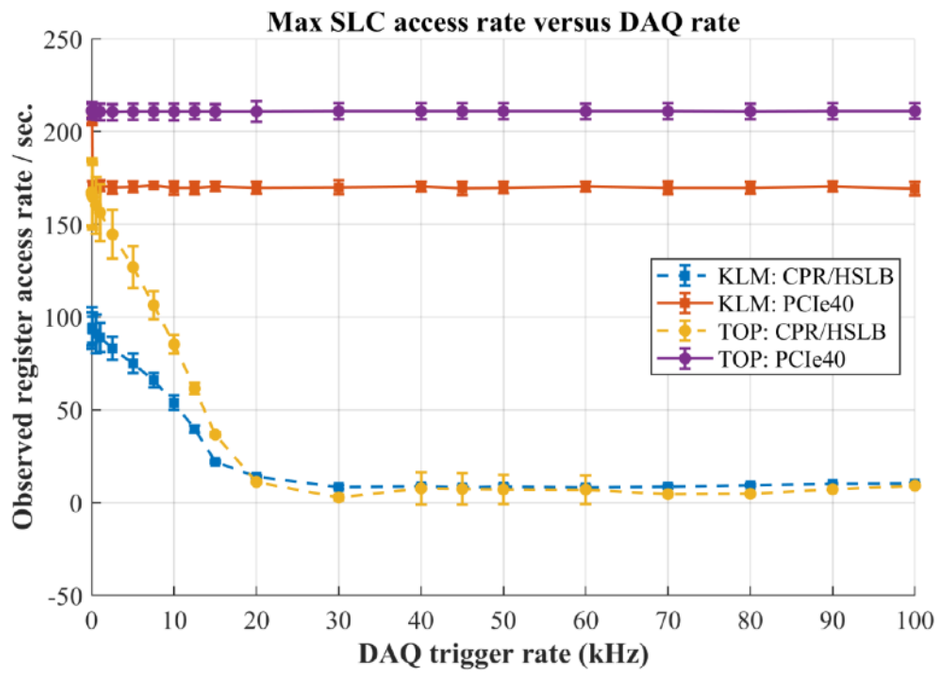
SVD FEE

ARICH FEE

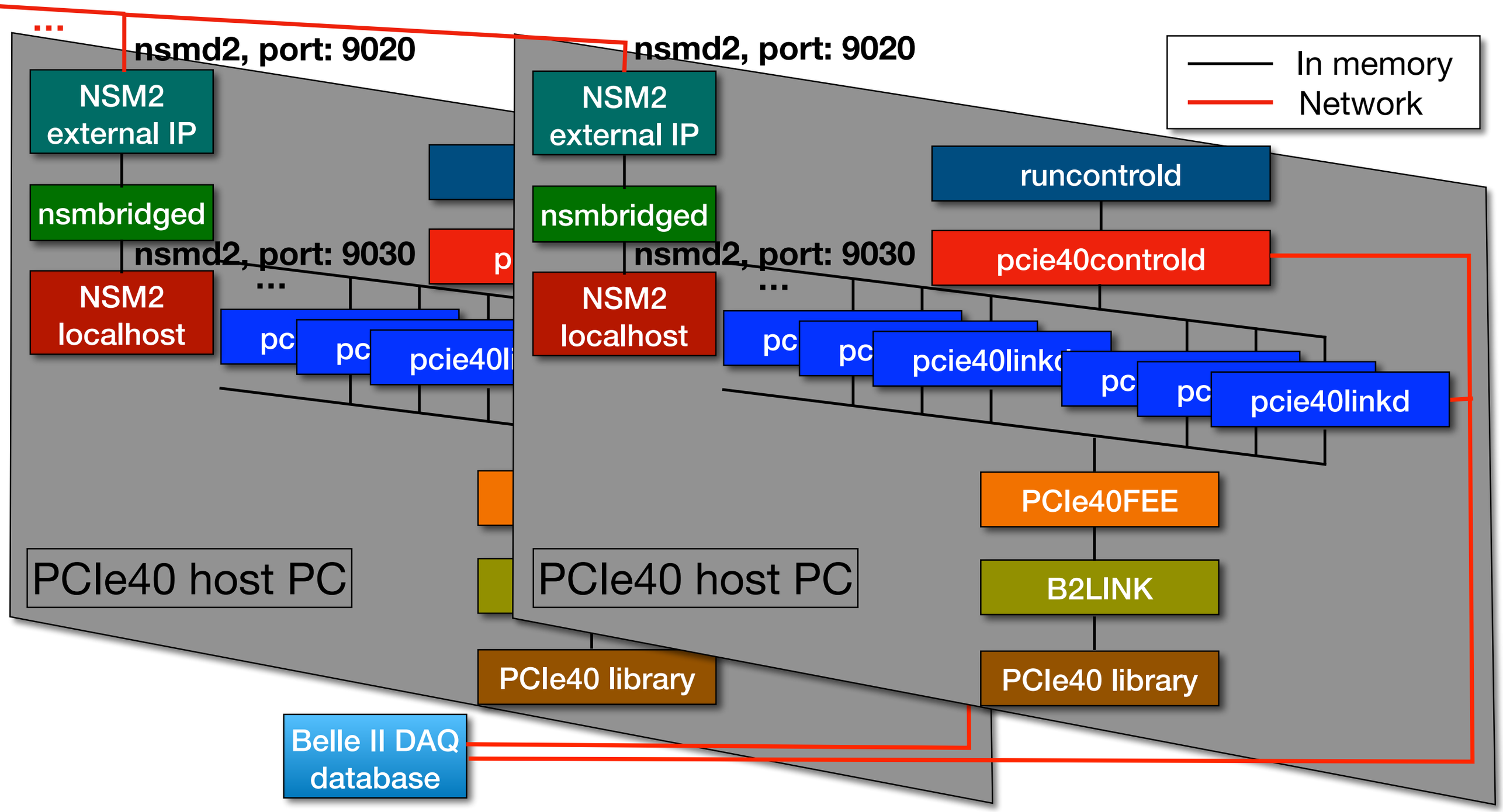
Slow control performance

- Belle2link was kept as most the same as COPPER-HSLB system
- 3 SLC 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 as shown in bottom plot
  - It takes the same time for the access w/ and w/o parallel access
- SLC configuration for FEEs of TOP (64 links) and KLM (32 links) has been tested and working fine.

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

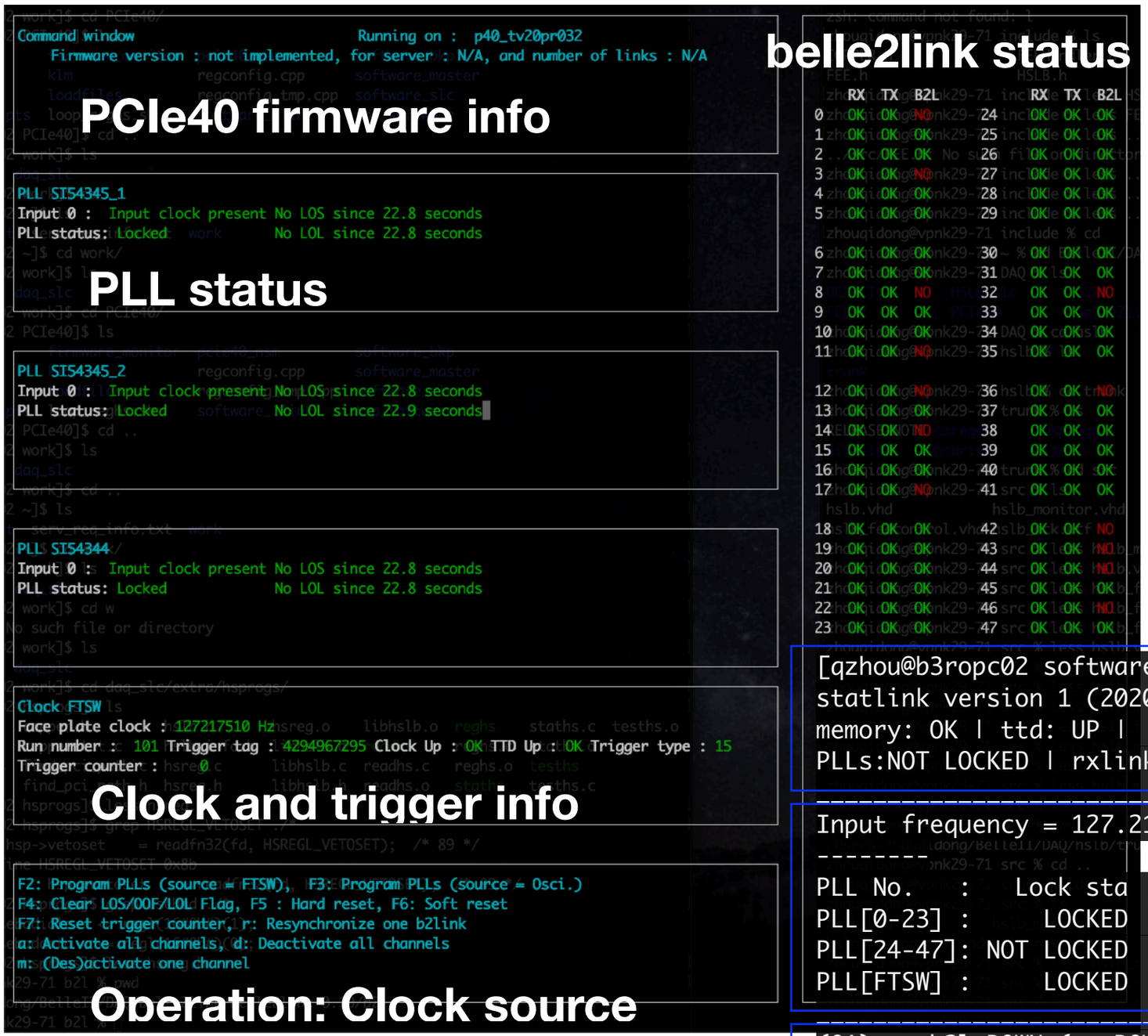


Upgrade of slow control system



- pcie40controld: one node pre board, monitoring PCIE40 board status, manage pcie40linkd
- pcie40linkd: one node pre link, monitoring link status, initialize and monitoring corresponding FEE.

Monitoring system



- **Monitoring and operation GUI**
  - **Select clock source**
  - **Mask links**
  - **Resets**

- **C++ monitoring tools**
  - **all functionality in GUI**
  - **Detailed info about errors**
  - **Enable to integrated to Belle II logging system**

```
[qzhou@b3ropc02 software_monitor]$ ./Pcie40Applications/statlink --info --ch 1 2
statlink version 1 (20200428) / PCIE40 firmware version 9.0
memory: OK | ttd: UP |  ttd clk: UP | run=: 0 | trg: 0 | trg type: 15
PLLs:NOT LOCKED | rxlink: NOT READY | txlink: READY | txlink: SYNCED | rxdata: OK
```

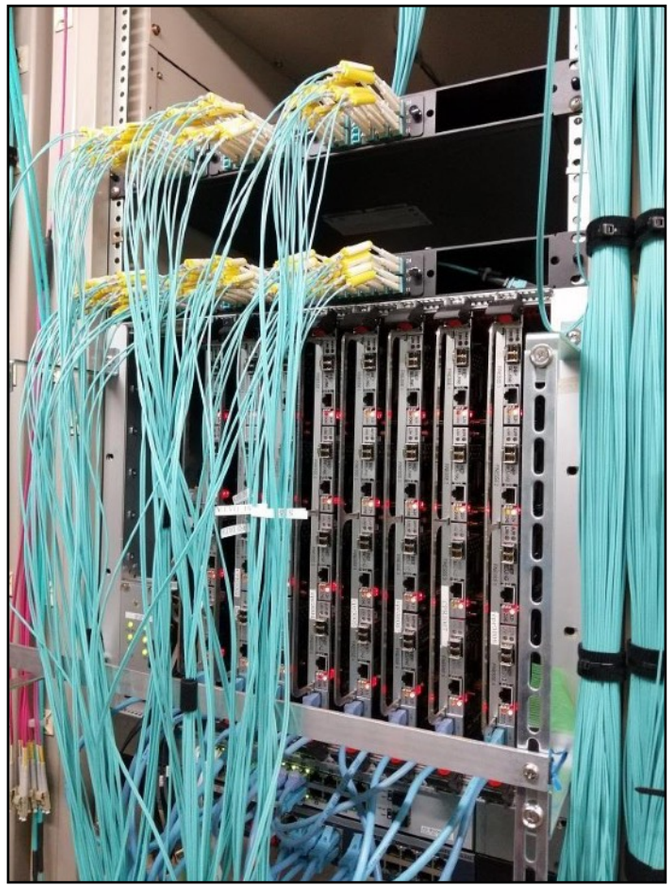
```
Input frequency = 127.217259 [MHz]
--------
PLL No.    :   Lock sta | Input sta | Fre. sta | Lock cnt | Input. cnt | Fre. cnt
PLL[0-23] :      LOCKED | BAD | BAD | LOL | LOS | OOF
PLL[24-47]: NOT LOCKED | OK  | OK  | LOL | LOS | OOF
PLL[FTSW] :      LOCKED | BAD | BAD | LOL | LOS | OOF
```

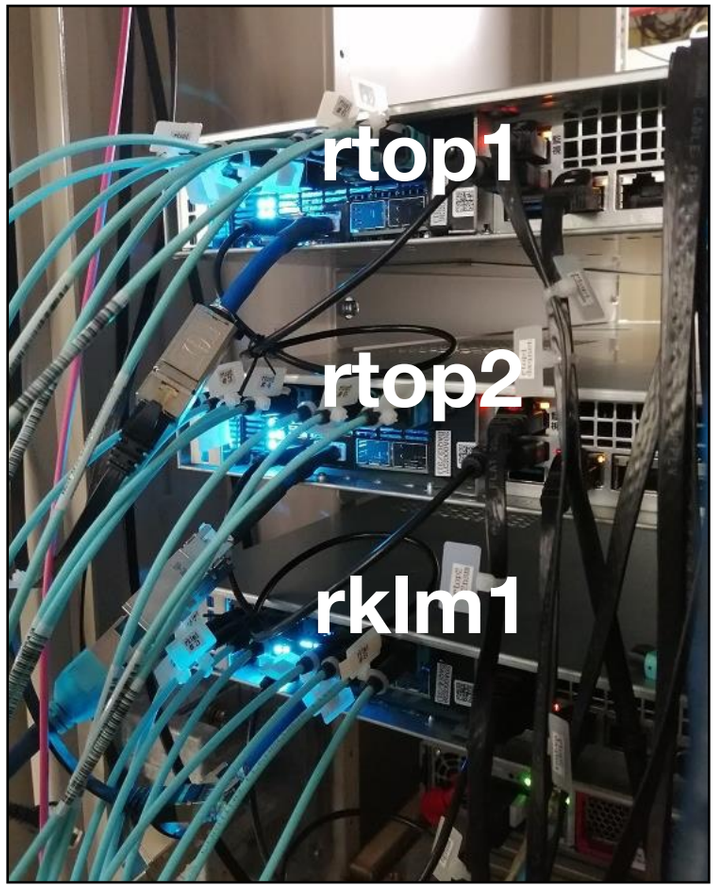
```
(01)    b2l=DOWN (rx=DOWN tx=DOWN rxsta=NOT READY txsta=NOT READY mask=MASK)
(01)    rxdata=0000 rxlinkdown=0 rxcrcerr=0 feecrcerr=0
(01)    event=0 total=0kB
(01)    no b2link error correction
```

```
(02)    b2l=DOWN (rx=DOWN tx=UP rxsta=NOT READY txsta=NOT READY mask=UNMASK)
(02)    rxdata=0000 rxlinkdown=0 rxcrcerr=0 feecrcerr=0
(02)    event=0 total=0kB
(02)    b2link error 49(i) 47(-) D00 Dc7 D00 D00 D00 D2f
```

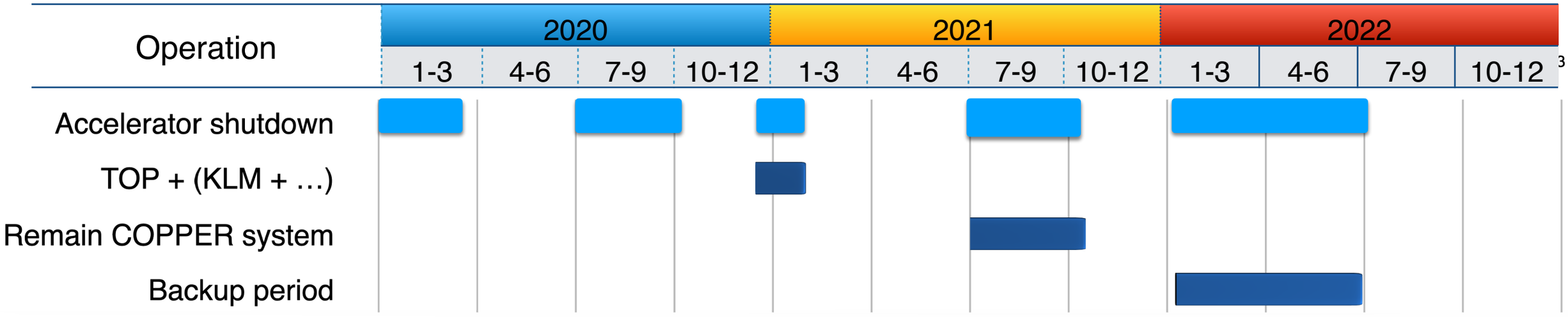
Preparation for replacement

- 31 boards has been produced, then tested at CERN, they already arrived at KEK in Sep. 2020
- 3 boards were installed on 3 servers for replacement of TOP and KLM
- 10 GbE NIC installed on ROPC, 40 GbE network switch <—> HLT server room installed





Full replacement timescale



Current COPPER system will be on standby for a while after the installation,
In case of a serious trouble, we can rollback to the COPPER system quickly

Summary

- Belle II DAQ system was designed to handle 30 kHz L1 triggers within 1% dead time.
- A COPPER based readout system has been the bottleneck to maintain and improve the performance of Belle II DAQ system
- PCIe40 module based upgrade proposal has been adopted for the upgrade of Belle II DAQ readout system
- Upgrade will keep the main features of current readout system to reduce the modification of sub-detectors as much as possible.
- Firmware and software development are almost done
- Full slow control functionalities has been tested with TOP and KLM on-site detector
- A stress test of data readout system at test bench confirmed that no event lost with 260 kHz trigger rate and 21 Gb/s
- Replacement is scheduled in this winter for TOP and KLM, the others during summer shutdown 2021.