The Performance of Belle II High Level Trigger in the First Physics Run

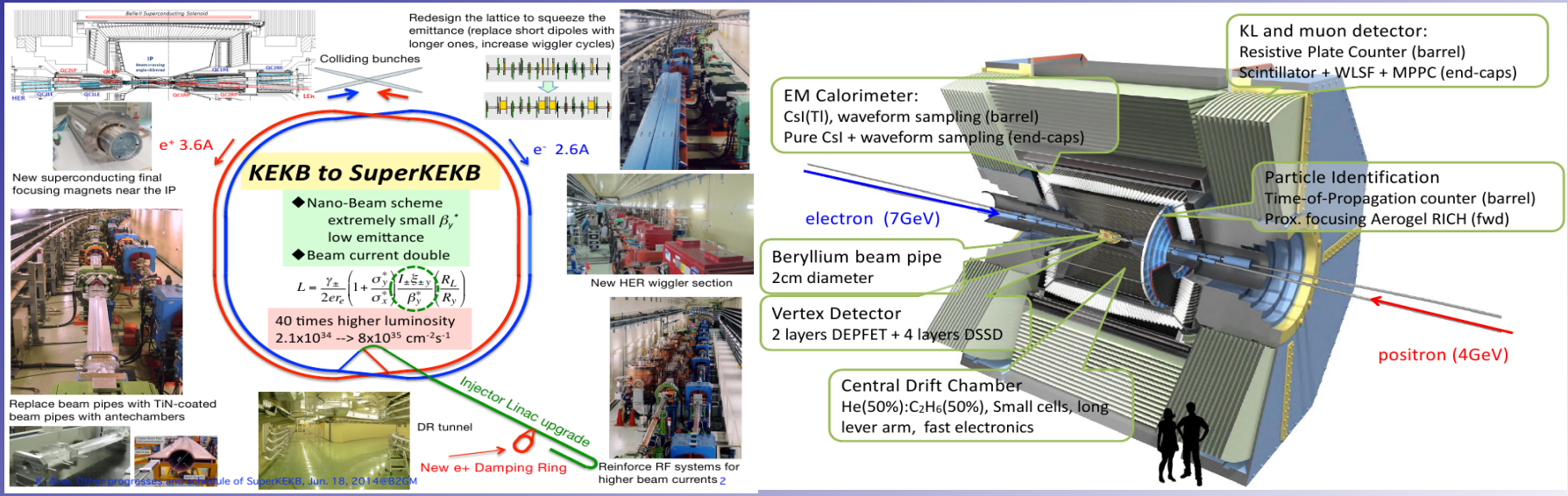
Ryosuke Itoh (KEK),
Nils Braun (KIT), and Chunhua Li (LNNU)
with
Belle II DAQ group

CHEP 2019 Track 1, Nov.5, 2019, Adelaide

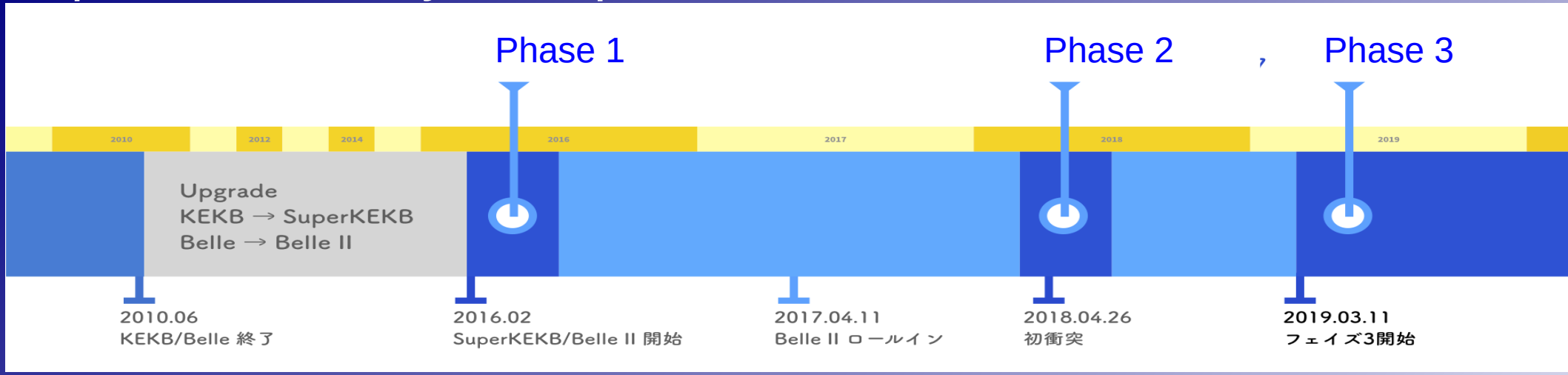
The Belle II experiment

- The Belle II experiment is a new generation B-factory experiment at KEK in Japan aiming at the search for New Physics(NP).

- A huge number of B mesons are produced in e^+e^- collisions in the SuperKEKB accelerator and NP is searched for in B meson decays by the Belle II detector complex.



Operation history of SuperKEKB/Belle II



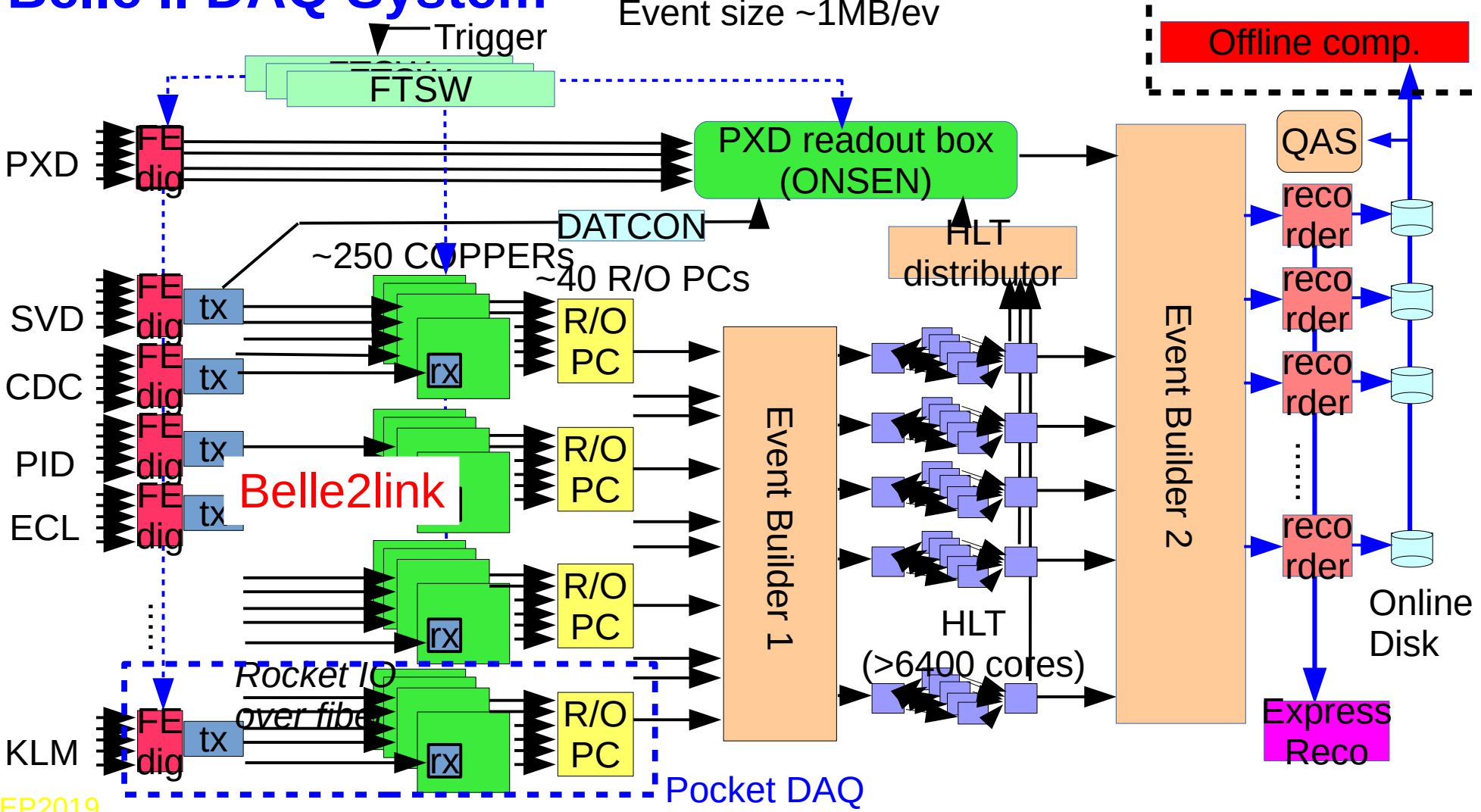
- Phase 1: Commissioning of the accelerator. Vacuum scrubbing in beam pipes.

- Phase 2: The first collision with focused beams. Only the outer detectors were installed and the pilot data taking performed.

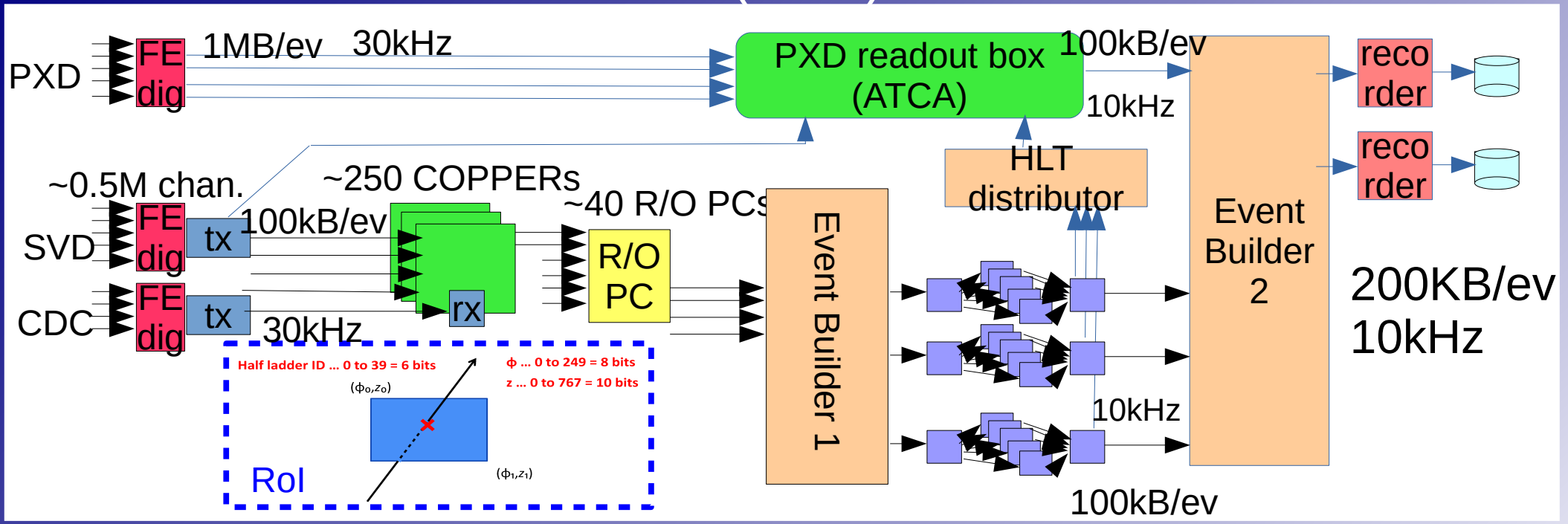
- Phase 3: Physics Run with the vertex (PXD+SVD) + outer detectors and full DAQ started from Mar.11, 2019.

Belle II DAQ System

Maximum design rate = 30kHz
Event size ~1MB/ev



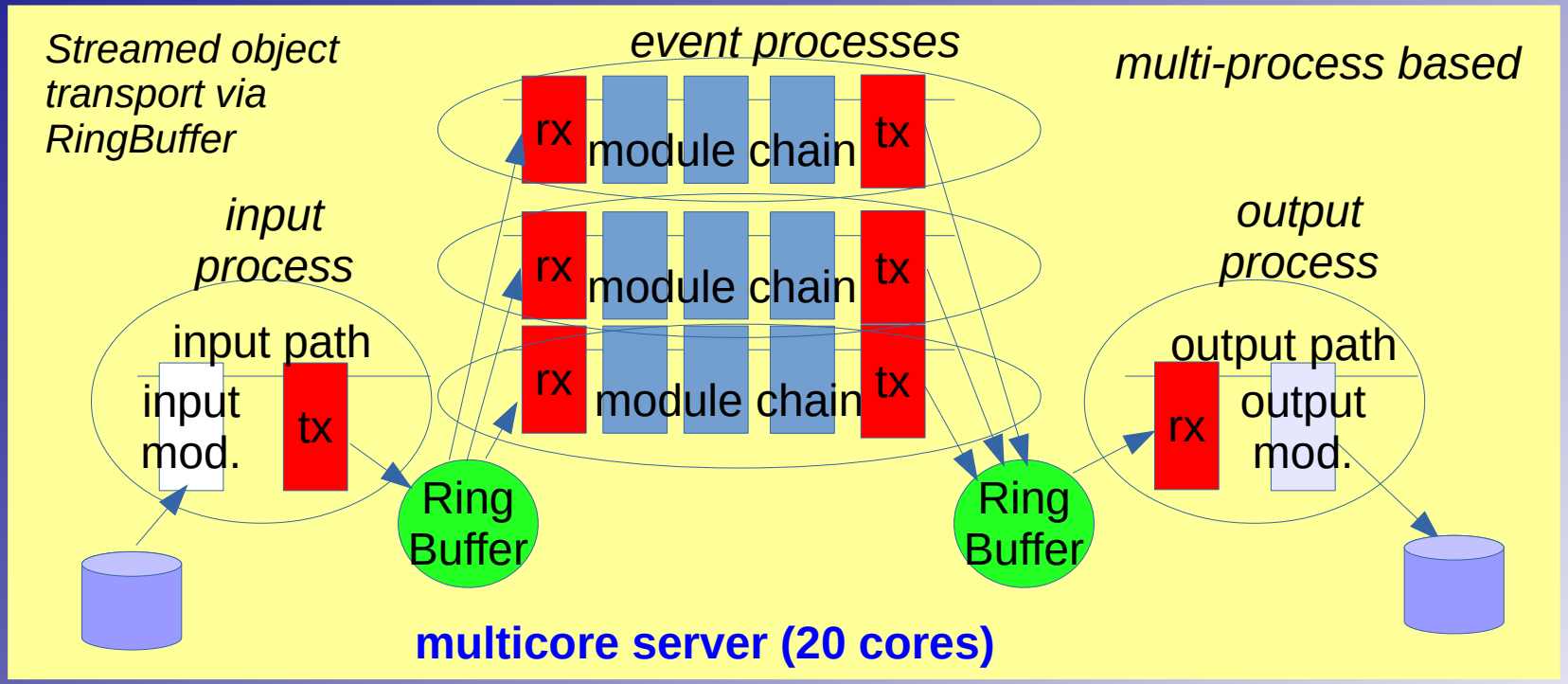
Readout for Pixel Detector (PXD)



- PXD yields a large event sized data when occupancy is high (>1MB) and it cannot be processed by COPPERs, nor recorded without event reduction.
- Data size reduction by 1) extrapolate HLT-reconstructed tracks to the surface of PXD sensors (Region of Interest), 2) send the RoIs to PXD readout box, and 3) discard hits not in RoIs. -> 1/10 reduction is expected.
- RoIs are sent only for HLT-selected events, and the rate reduction is also applied.

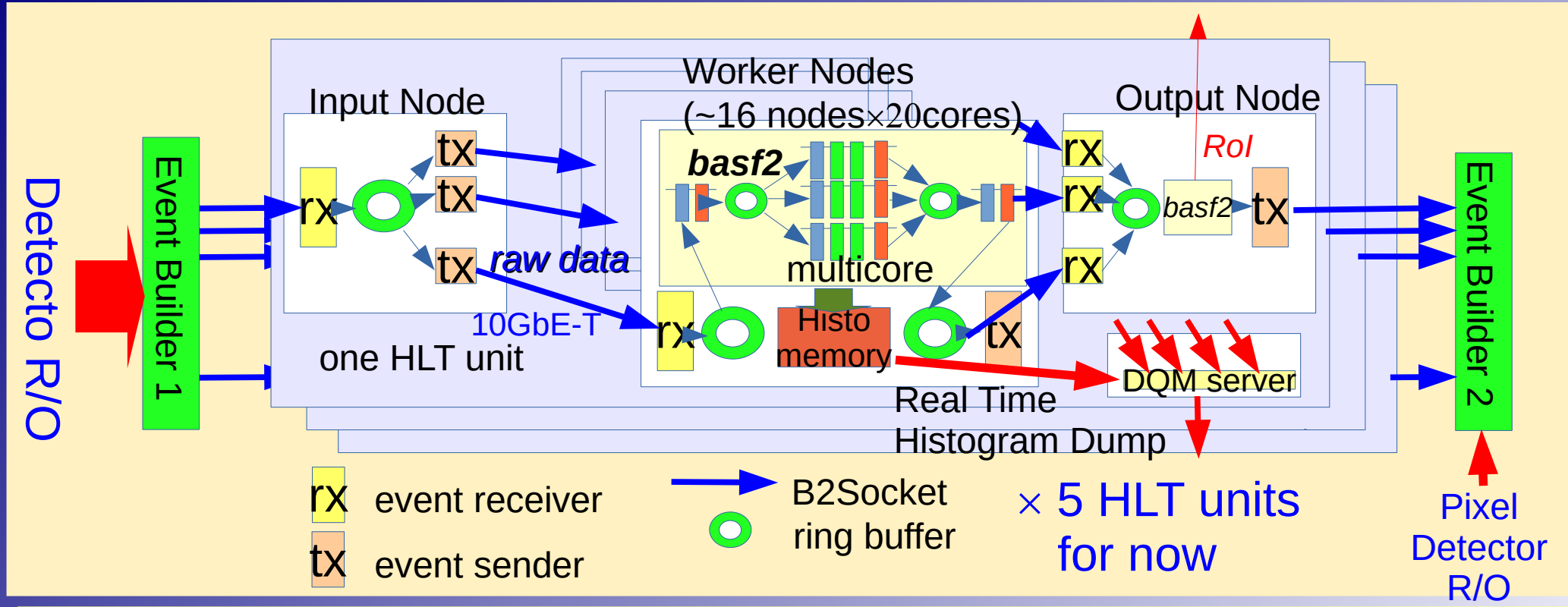
Parallel Processing in Belle II High Level Trigger

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



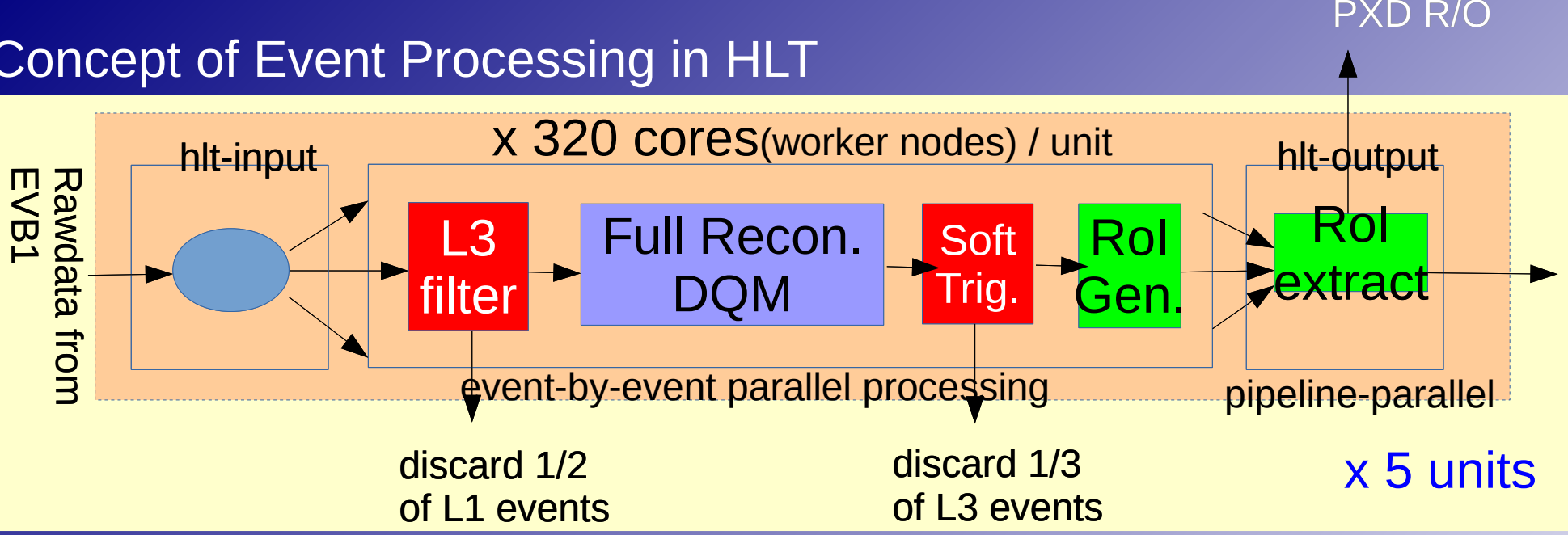
- Extended to PC cluster using the same RingBuffer + network socket connection for the streamed object transport.

Belle II High Level Trigger (HLT)



- Unit structure. One unit houses 320 cores. Current : 5 units = 1600 cores (of 6400)
- System control is based on “NSM2”, which is a home-grown slow control framework used in Belle II DAQ.

Concept of Event Processing in HLT



* Level 3 filter before full event reconstruction.
* The same offline full event reconstruction is performed for accepted events.
* Physics event skim (hadronic/tau event selection....) is used as “software trigger” (with scaled calibration events added)
* “RoI”s are collected on the output node and sent to PXD readout.

DAQ Status in Phase 3 : Short Summary

- Belle II DAQ was working in the first Physics Run w/o serious troubles.
- Nominal L1 rate was still up to 3.5 kHz since the accelerator luminosity was low (Maximum was L = 1.2 x 10^{34}, which is only 1/50 of design).
- The nominal data taking efficiency during the whole run period (~4 months) was 80-85% including the debugging period.
- In the stable physics run, the efficiency was more than 90%.

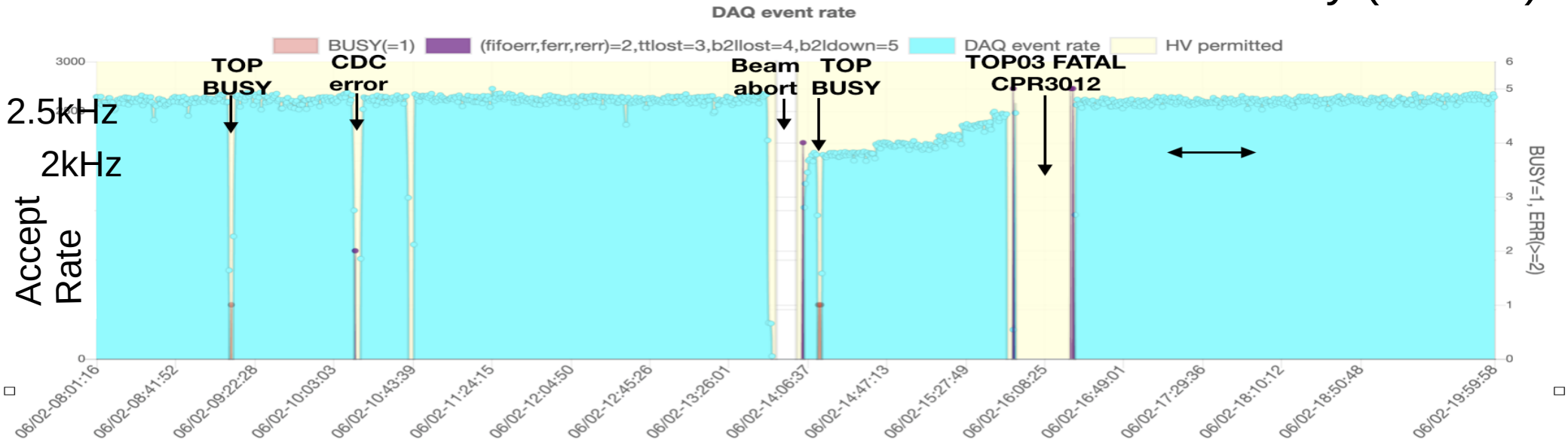
Major sources of DAQ dead time

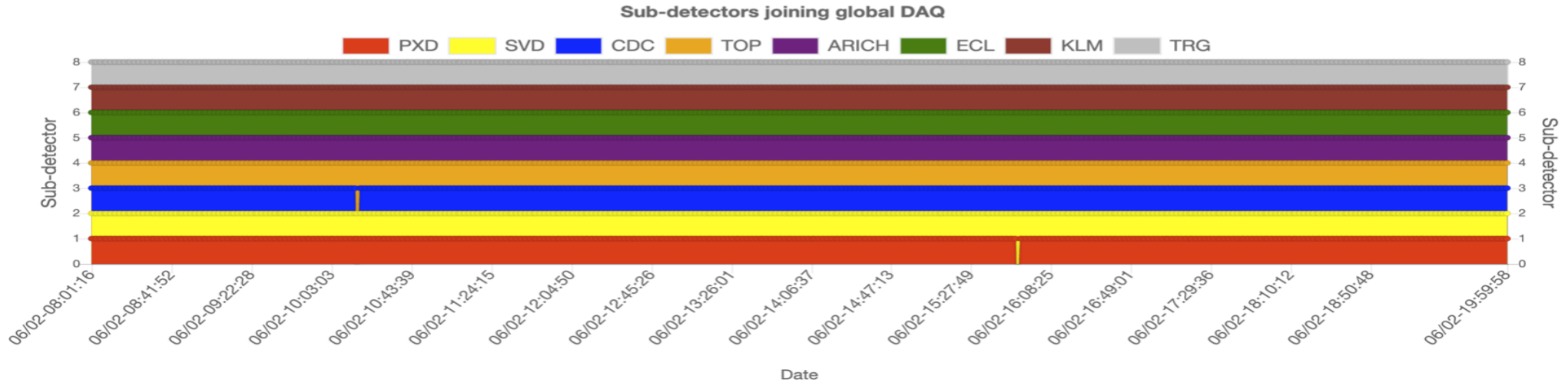
1) “ttlost”
   - Lost the connection / synchronization in trigger/clock link
2) “b2llost”/ “b2ldown”
   - Lost the connection in Belle2link
3) Hang up of COPPER CPU
4) HLT hang up at STOP and ABORT.
5) The entire breakdown of slow control system

Links : last 24hrs last 7days all period VXD test

DAQ Efficiency = 93.4%

Good day (June 2)





Operation experience of Belle II HLT

- A full event reconstruction was performed on HLT using the same offline software and was running stably.

- In the earlier operation period, the HLT selection was not turned on and all events were recorded in the storage for the offline validation.
  -> After the validation, the HLT selection was turned on.

- “Level 3” trigger was not turned on and the processing rate with 5 units was limited up to 7kHz (cf. typical L1 rate was up to 3.5kHz).

The HLT reduction factor in the period was observed to be about 1/8.

Problems in HLT operation

1. Sometimes HLT stuck at run stop and restart.

- When stuck occurred, it was found that IPCs were not properly cleaned up.
  * HLT framework relies on the home-grown RingBuffer for the parallel processing, which uses Linux IPC (SharedMem, Semaphore).
  * The HLT stop is implemented using Linux signal and the problem was found to be caused by the incomplete signal handling.

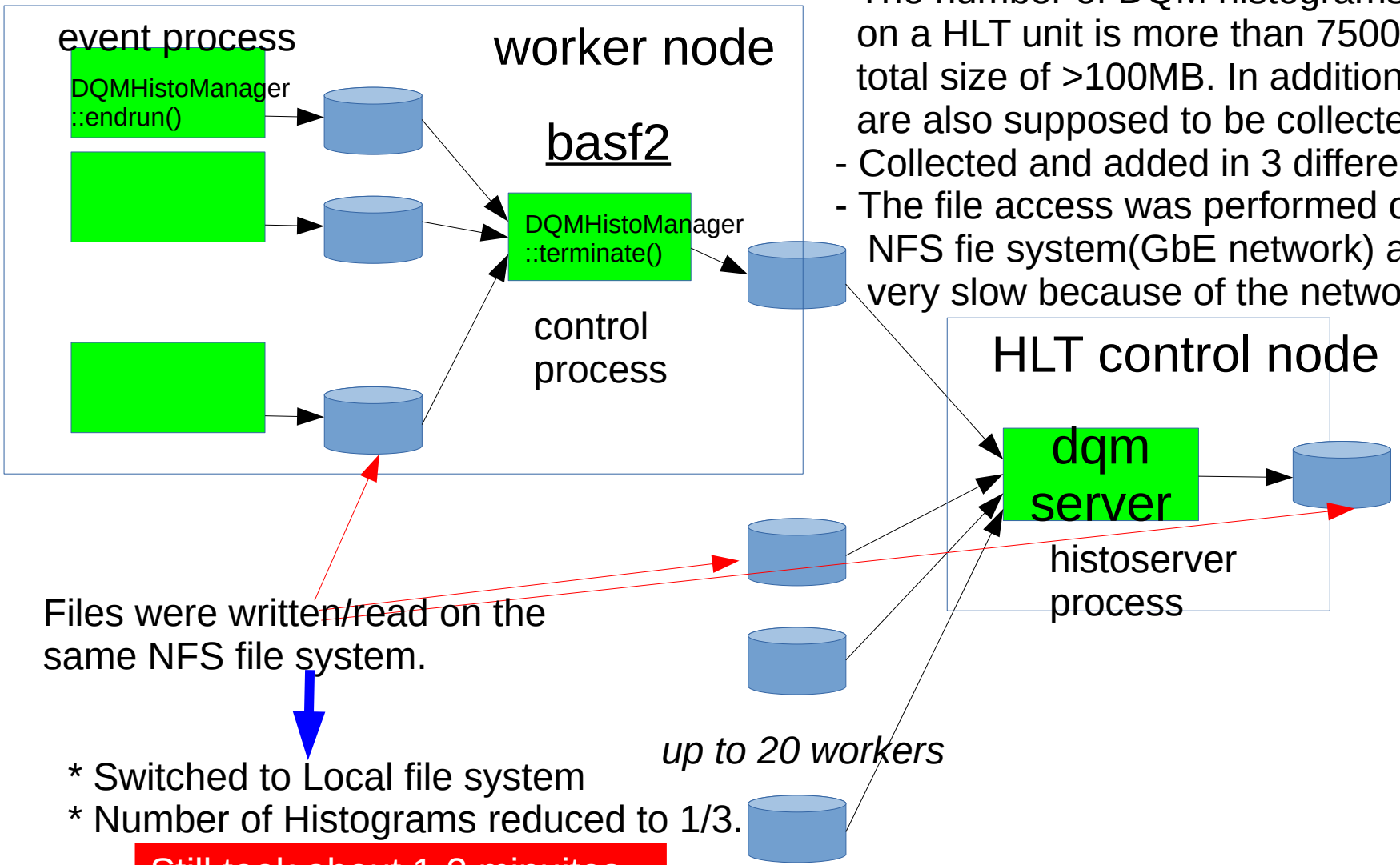
  -> Fixed by improving the signal handler and the operation stabilized.

2. The time to stop HLT took very long up to 5 minutes!

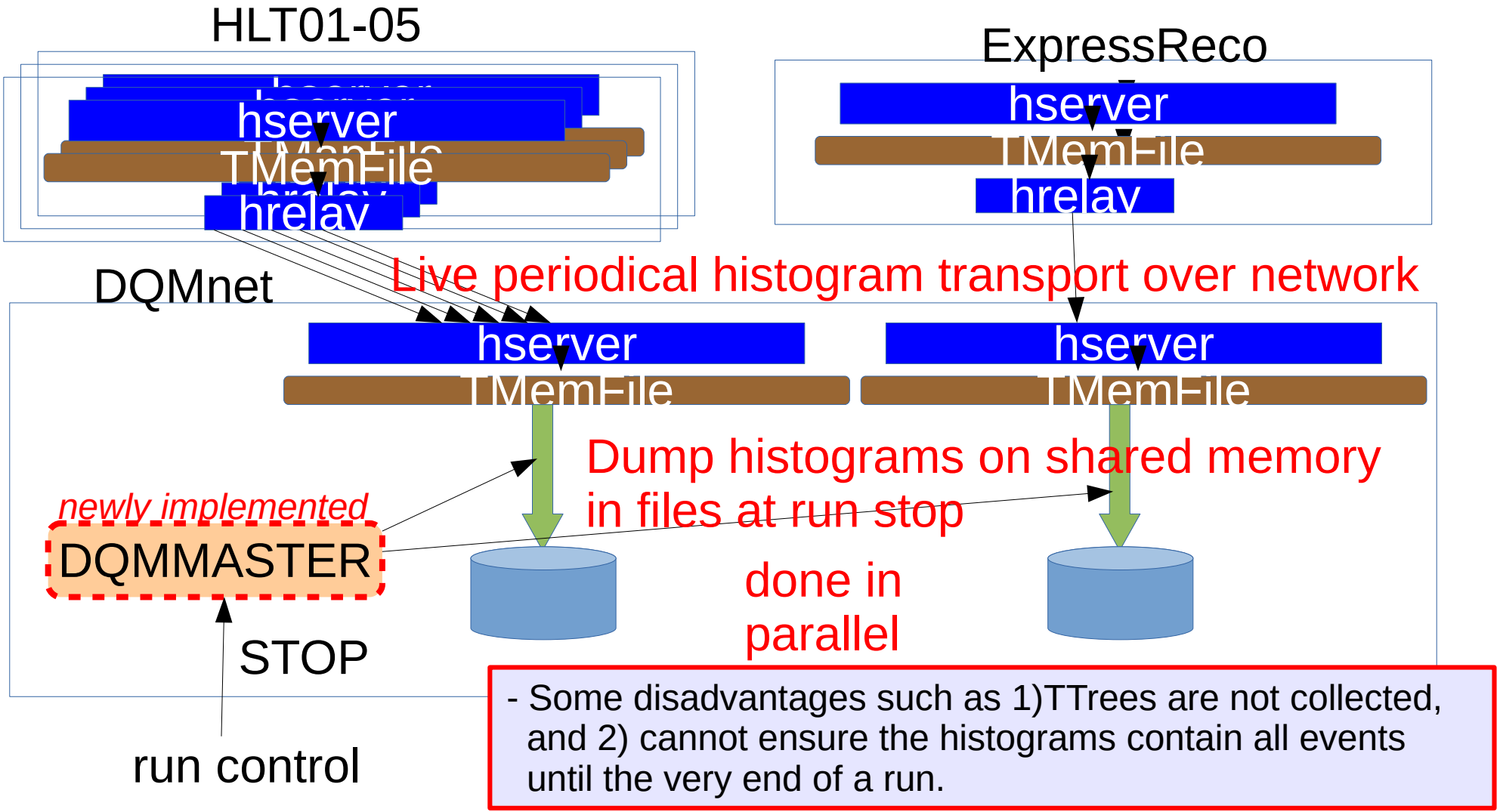
- Time was consumed to collect and store DQM histograms in files at run end.
  * It is done separately from the real-time collection for monitoring.
  * Need to collect/add >7500 ROOT histograms/TTrees from 1600 cores in three steps, every step takes ~ 1 minute. Shared NFS file access made the performance worse.

  -> Finally the histogram file storage was switched to the real-time collection based, the stopping time was shortened to 30sec.

- The number of DQM histograms collected on a HLT unit is more than 7500 with the total size of >100MB. In addition, TTrees are also supposed to be collected.
- Collected and added in 3 different steps.
- The file access was performed on the same NFS fie system(GbE network) and it was very slow because of the network bottleneck

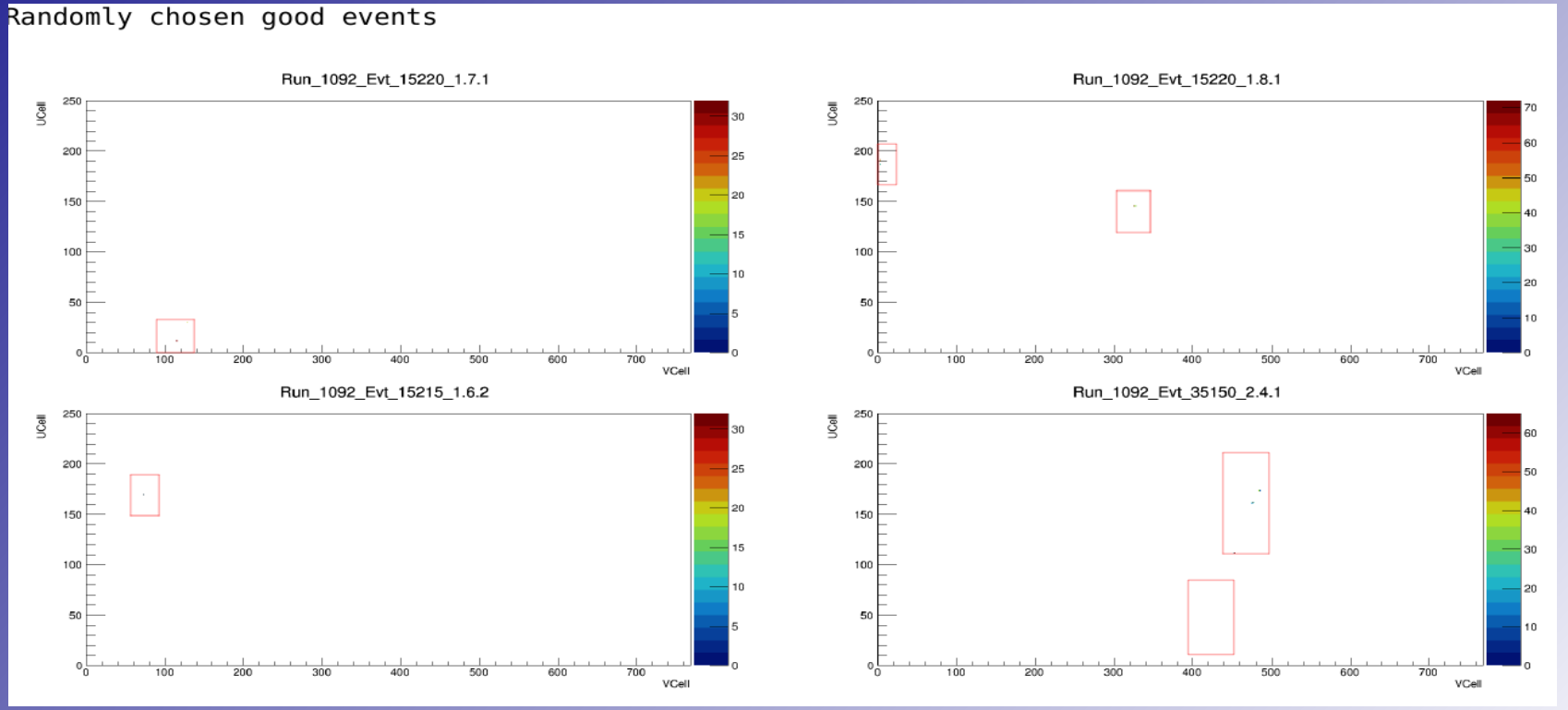


Real Time Histogram Transport and New Histogram Store



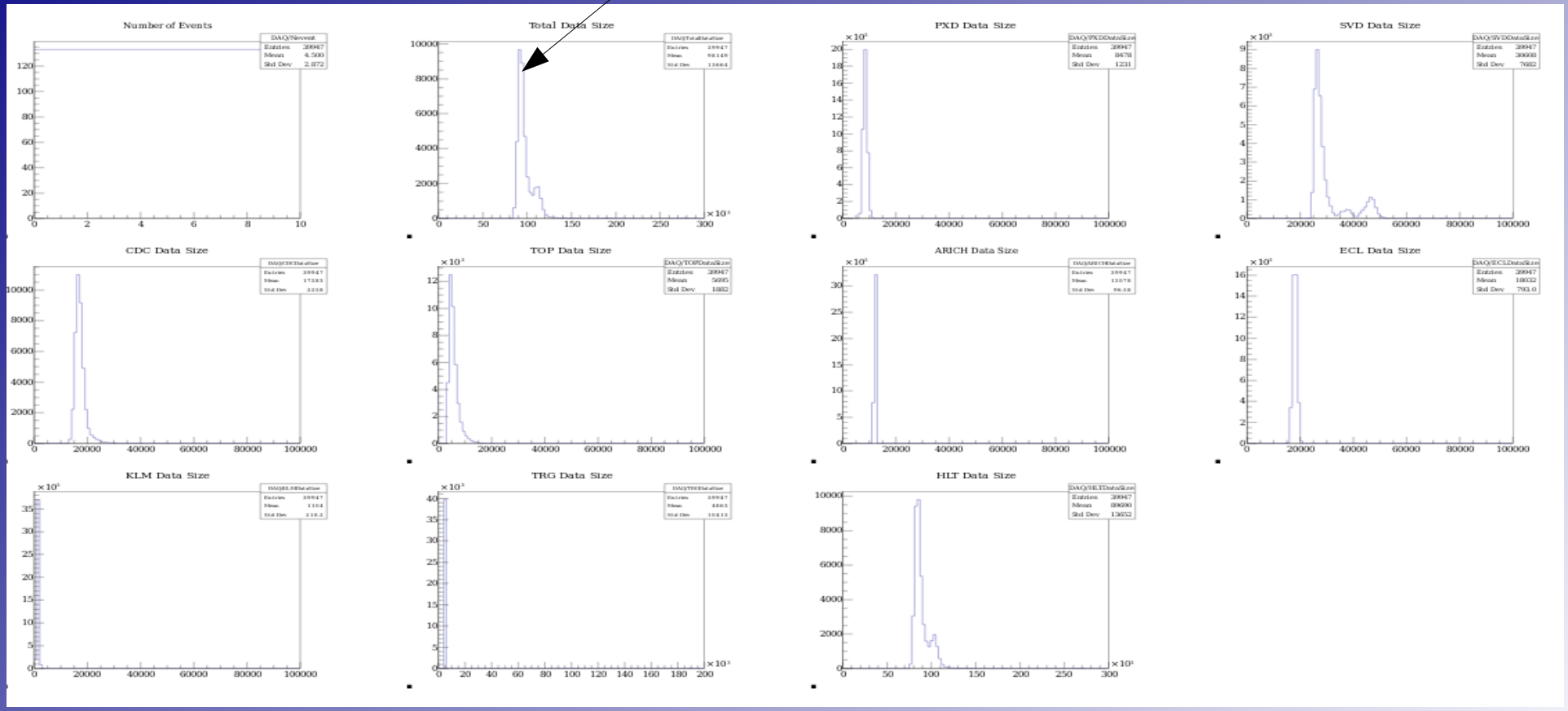
RoI feedback from HLT to PXD readout

- RoI feedback and the data reduction in PXD R/O system was confirmed to work. Clusters seem to be centered in RoI.
- However, there was a problem in cable connection in some part of PXD FEE and the real time reduction was not yet used.



Event size monitored at ExpressReco

Current total event size is ~100kB/ev



Summary

- Belle II HLT worked well in the first physics run for 4 months.

- The full event reconstruction using the same offline software was performed on HLT and the physics-skim based triggering was used.

- The observed reduction rate was about 1/8 and confirmed the expected performance.

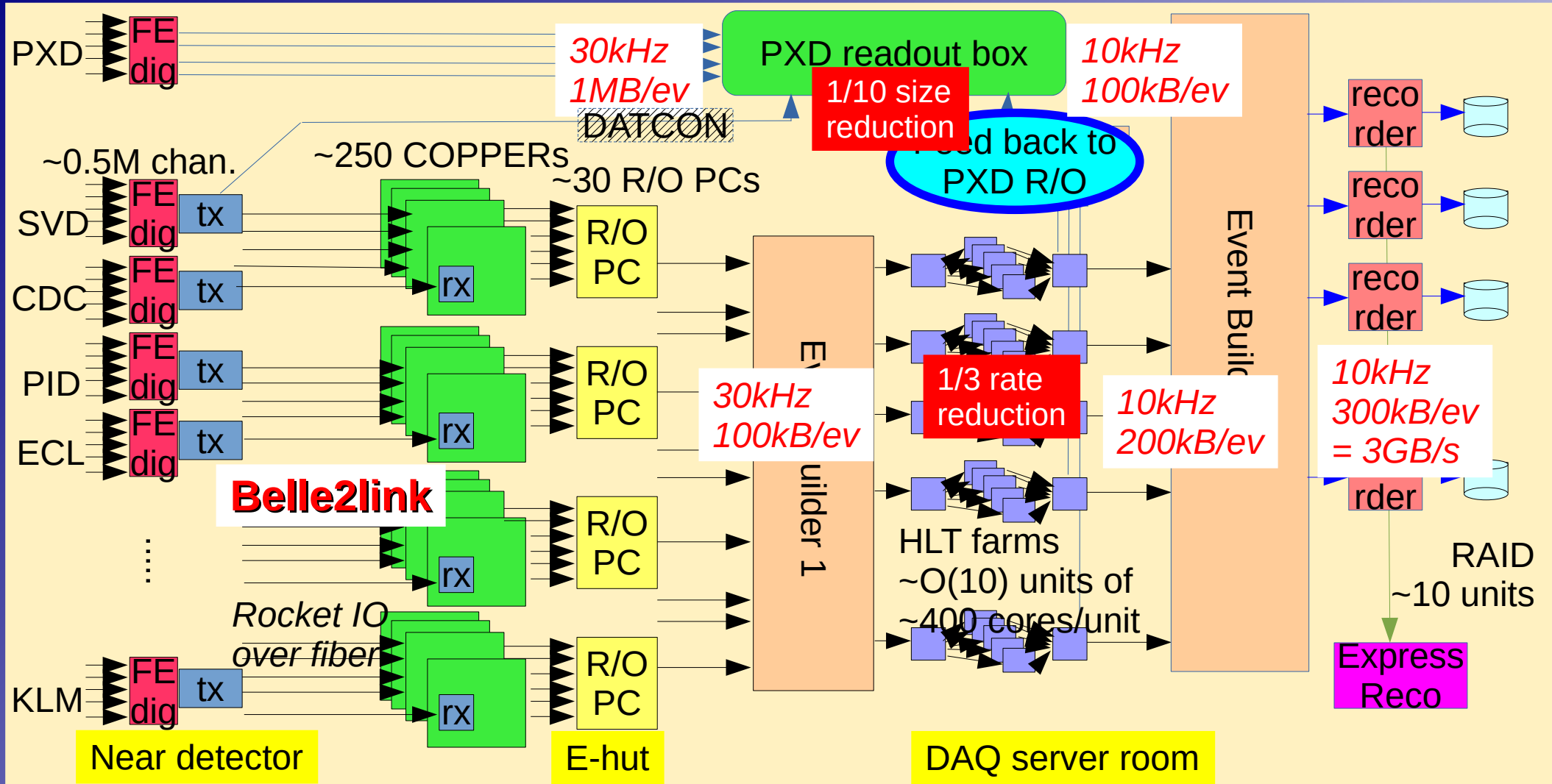
- Encountered several troubles.
  * IPC resources used in RingBuffer were not properly removed.
  * It took too long to store DQM histograms in files at run end.

  They were fixed in a timely manner and the operation was stabilized.

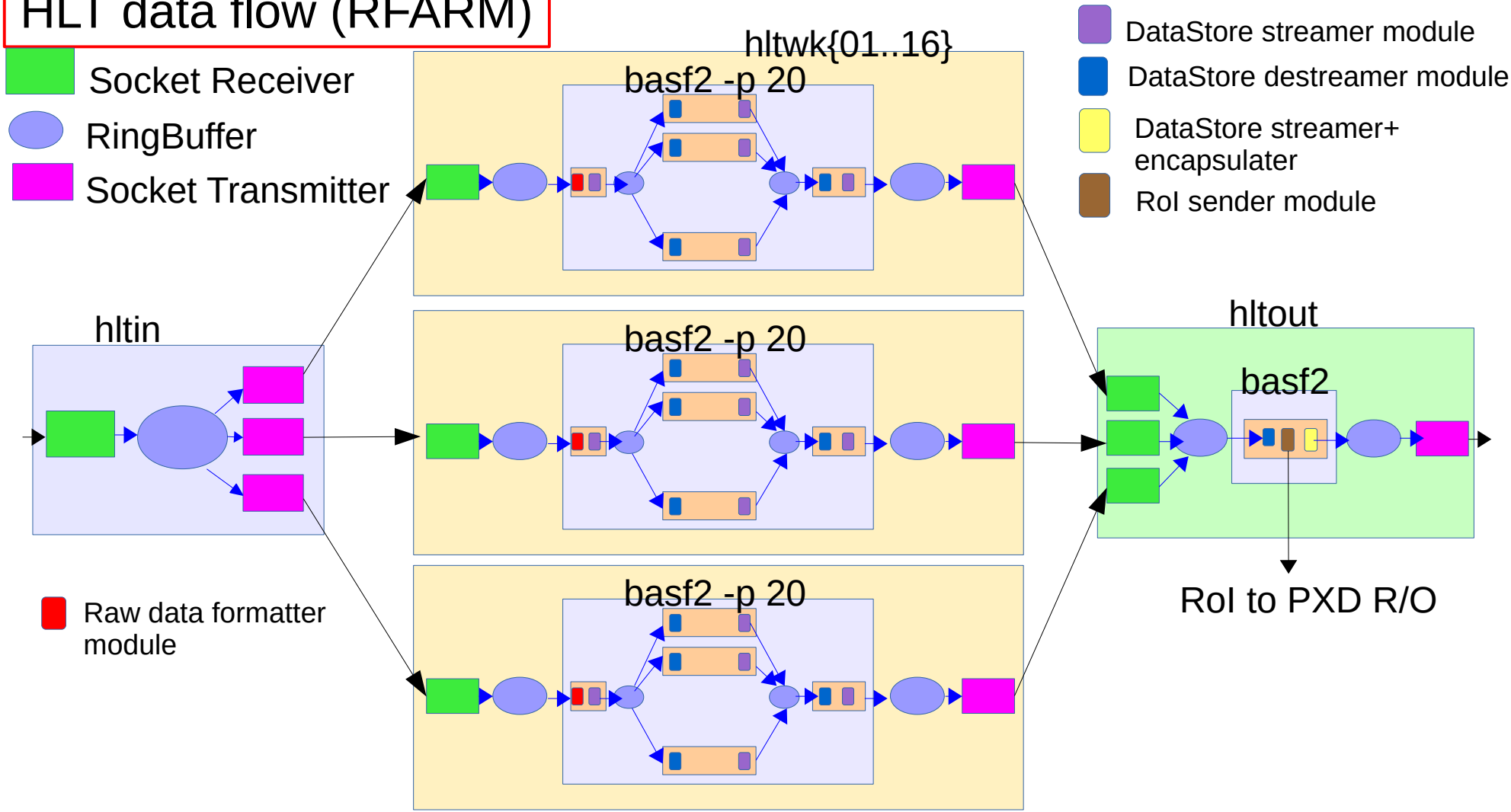
- Upgrade project is going on to replace RingBuffer with ZeroMQ.
  -> being tested in on-going autumn run.

Backup Slides

Data Flow in Belle II DAQ

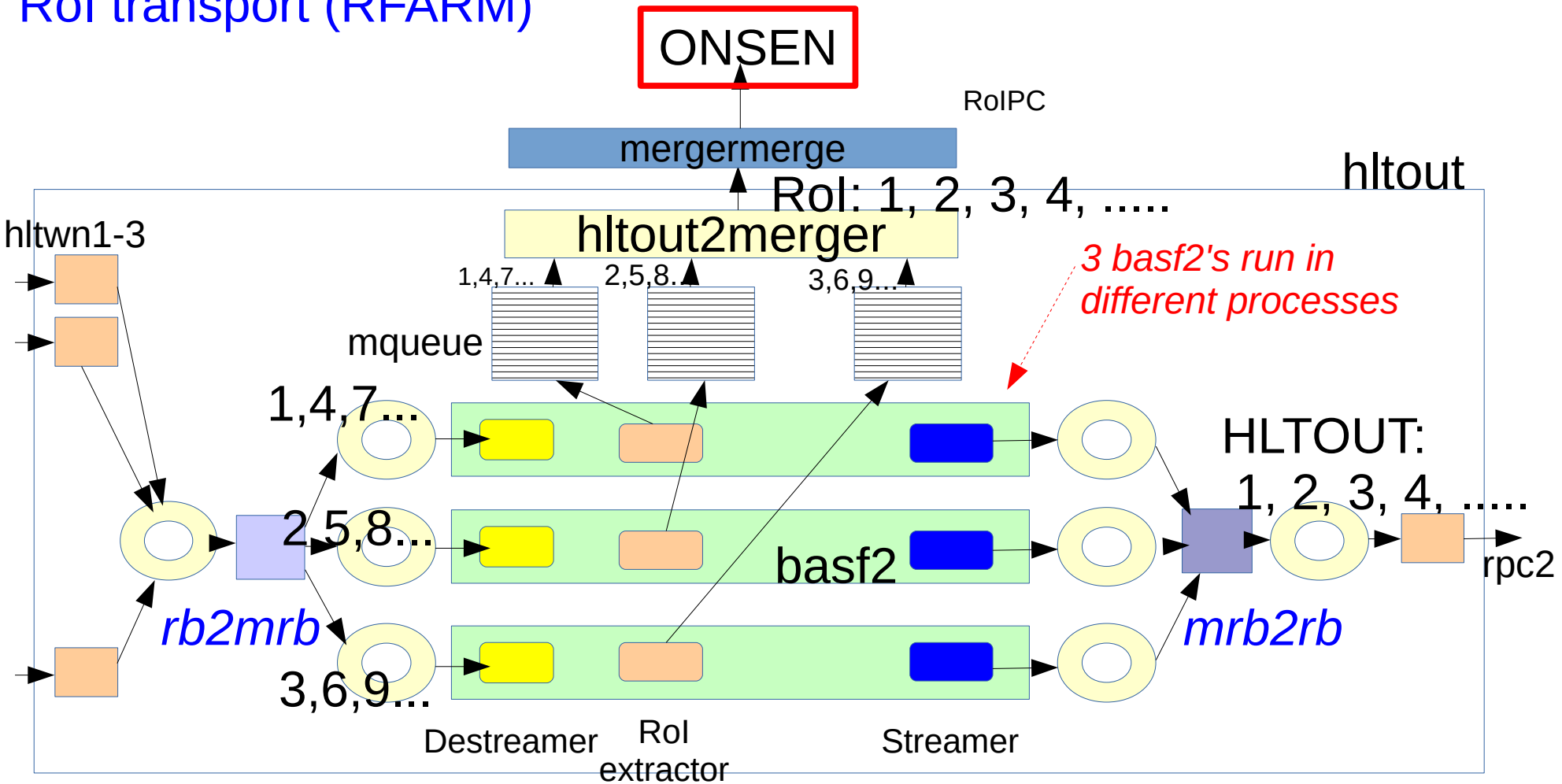


HLT data flow (RFARM)



20 cores (3GHz) x 16 PC servers

RoI transport (RFARM)



* rb2mrb, mrb2rb, and hltout2merger distribute/pick up records in turn to/from ringbuffers/mqueues in the same order.

Data Quality Monitoring