

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

R. Itoh^{1,*}, N. Braun², C. Li³, M. Nakao¹, S. Yamada¹, S. Y. Suzuki¹, Q. Zhou¹, T. Konno⁴, Z-A. Liu⁵, J. Zhao⁵, O. Hartbrich⁶, SH. Park⁷, Y. Guan⁸, K. Lautenbach⁹, S. Reiter⁹, and B. Spruck¹⁰

¹High Energy Accelerator Research Organization (KEK), Tsukuba

²Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

³LiaoNing Normal University (LNNU), Dalian, China

⁴Kitasato University, Sagamihara, Japan

⁵IHEP, Beijing, China

⁶University of Hawaii, Honolulu, U.S.A.

⁷Yonsei University, Seoul, Korea

⁸University of Cincinnati, Cincinnati, U.S.A.

⁹University of Giessen, Giessen, Germany

¹⁰University of Mainz, Mainz, Germany

Abstract. The Belle II experiment is a new generation B-factory experiment at KEK in Japan aiming at the search for New Physics in a huge sample of B-meson decays. The commissioning of the accelerator and the detector for the first physics run has started from March this year. The Belle II High Level Trigger (HLT) is fully working in the beam run. The HLT is now operated with 1600 cores clusterized in 5 units, which is 1/4 of the full configuration. The software trigger is performed using the same offline reconstruction code, and events are classified into a set of physics categories. Only the events in the categories of interest are finally sent out to the storage. Live data quality monitoring is also performed on HLT. For the selected events, the reconstructed tracks are extrapolated to the surface of the pixel detector (PXD) and quickly fed back to the readout electronics for the real time data reduction by sending only the associated hits. The maximum trigger rate in the first physics run was 3.5kHz, and the Belle II data acquisition system was stably operated. There were several problems in the HLT operation, but they have successfully been fixed during the data taking period. The HLT reduction factor is measured to be 8 which is still higher than the design because of the high background environment.

1 Introduction

The Belle II experiment[1], which is a new generation B-factory experiment, has just started the physics data taking. The Belle II detector at the SuperKEKB[2] accelerator is shown in Figure 1. The data acquisition system (DAQ) for the experiment is required to manage a Level 1 trigger rate up to 30 kHz with an event size of more than 1 MB, and the high-speed readout and data reduction are the key of the DAQ.

*e-mail: ryosuke.itoh@kek.jp

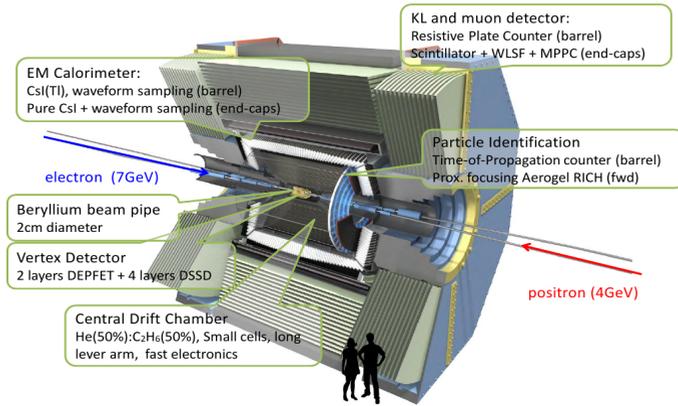


Figure 1. The Belle II detector. From the inside, the pixel detector (PXD), the silicon vertex detector (SVD), and the central drift chamber (CDC) are implemented for the precision particle tracking. The time of propagation counter (TOP) and the aerogel Cherenkov counter (ARICH) are placed outside for particle identification. The detectors are surrounded by the electromagnetic calorimeter (ECL) and the KL and muon detector (KLM).

The commissioning of the SuperKEKB accelerator started from February, 2016 including the accelerator tuning and vacuum scrubbing (Phase 1). The first electron-positron collision was observed in April 2018 and the pilot run was performed only with the outer detectors (Phase 2). The physics run with all detectors with the vertex detectors installed started from March 2019 and the data taking with a full DAQ configuration is being performed (Phase 3).

2 Belle II DAQ System

Fig. 2 shows a schematic view of the Belle II DAQ system.

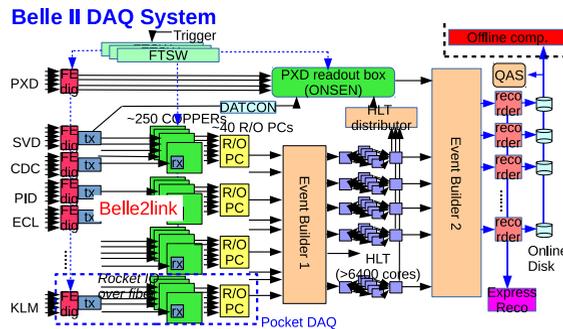


Figure 2. The Belle II Data Acquisition System. The data from detectors are transferred to the common readout cards (COPPER) via the unified optical link (Belle2link). Event building is performed in multiple steps by the combination of the readout PCs and the event builder switch and the event data are fed into multiple HLT units. The data from PXD are merged with the HLT output and recorded in storage units.

The Belle II DAQ system is a conventional trigger driven data acquisition system. The Level 1 trigger generated by the global decision logic is distributed to the detector front-end by the multi-layered distribution logic (FTSW) through both optical and metal links[3].

The detector signals are digitized at the detector front-end located near the detector and transferred to the common readout cards (COPPER)[4] via the unified optical link (Belle2link)[5]. The Belle2link is a bi-directional link and it is also used to download slow control parameters to the detector front-end.

On each COPPER board, a Linux-operated CPU card is mounted as a daughter card where the data formatting and the 1st level data reduction are performed. The CPU card is equipped with an ATOM processor with 512MB of memory and a network-booted instance of Scientific Linux 5.

The processed data are sent to readout PCs via the GbE network. The event building is done in two steps. The first step is done on the readout PC which collects the event fragments from COPPERs and formats them into one event fragment of each subsystem. It is then sent to the network switch complex and distributed to one of High Level Trigger (HLT) units. A full event is built at the input node of a HLT unit.

The readout of the pixel detector (PXD) is handled in a separate way, since the event size is huge (up to 1 MB) and cannot be handled with the COPPER based readout. The data from PXD are fed into a special readout system called ONSEN[7]. The tracks of charged particles reconstructed by the HLT are extrapolated to the surface of PXD sensors, and the regions of interest (RoI) on the sensors are defined. ONSEN receives the RoIs from HLT through the RoI merger and only the hits inside RoIs are sent to the 2nd event builder switch and merged with the output from HLT to be recorded on the online storage as shown in Fig. 3. The data size is expected to be reduced by a factor of 10.

The slow control of the system is implemented based on two different frameworks. One is NSM2[8] which is a home grown framework used in the main DAQ component, while the other is the industrial standard EPICS[9] used in some of detector subsystems and in the SuperKEKB accelerator. A gateway between them is developed and a transparent environment is implemented. The user interface (GUI) is constructed using Control System Studio[10].

3 Architecture of Belle II HLT

Fig. 4 shows the global design of the Belle II High Level Trigger (HLT). The HLT consists of multiple units of PC clusters. One unit consists of an event distributor node(**hltin**), an output collector node(**hltout**), and up to 20 event processing nodes(**hltevp**). Each **hltevp** server houses multiple cores (2 of Intel Xeon E5-2660 or similar) and a total of 320 cores are equipped in one HLT unit for the event processing.

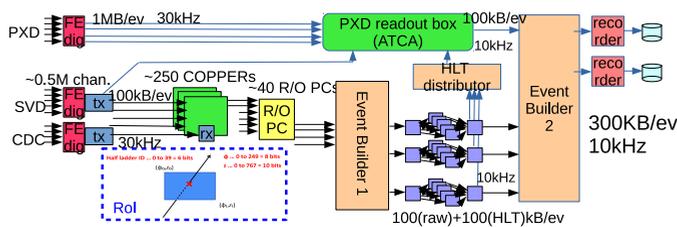


Figure 3. The RoI feedback to PXD readout. The track information obtained by HLT processing (Region of Interest: RoI) is fed back to the PXD readout box and only the hits associated to the tracks are sent.

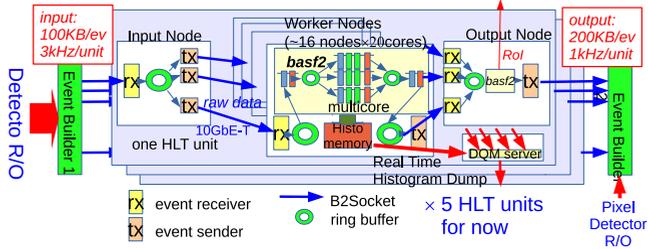


Figure 4. The Belle II High Level Trigger. It is implemented in multiple units of PC farms. One unit consists of one input and output node and up to 16 worker nodes equipped with 20 cores. The data flow between nodes is implemented using a socket connection combined with the ring buffers.

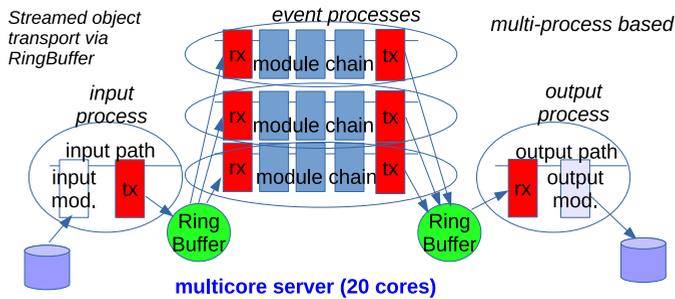


Figure 5. The parallel processing mechanism in basf2. The streamed ROOT objects are transferred among multiple processes over the ring buffers. Each event process receives the data of different events and the event reconstruction and selection is done in parallel.

The raw event data from the event builder are first delivered to the **hltin** of one of HLT units via the socket connection and placed in a ring buffer. The ring buffer is implemented using Linux IPC shared memory and semaphore. The event is fetched by multiple transmitters and sent to one of the **hltvnp** nodes via the socket connection. The ring buffer ensures load balancing.

In each **hltvnp** node, the raw event data are passed to the offline analysis framework (**basf2**) through the ring buffer and processed. Multiple cores in the server are utilized by the parallel processing mechanism implemented in **basf2** based on process forking and the ring buffer[6] as shown in Fig. 5. The processed data, packed as streamed ROOT objects are collected in an **hltout** node in the reverse way and the RoIs are extracted to be sent to ONSSEN.

For the real time monitoring of the data quality, the ROOT histograms accumulated on each core are periodically collected and added up on a shared memory (ROOT::TMemFile) through the socket connection. The memory contents are then relayed to the DQM server where all monitoring histograms are collected and checked. The scheme is shown in Fig. 6.

During the phase 3 run, five HLT units containing a total of 1600 cores are used.

The design of event processing on each processing core is portrayed in Fig. 7. The raw data from the detector are first processed by the Level 3 filter before feeding them into the full event reconstruction chain. Fast CDC tracking and calorimeter clustering are performed

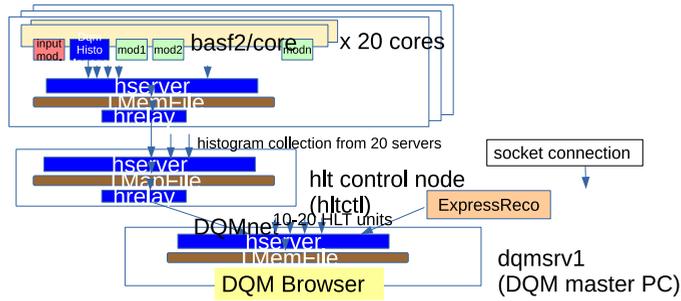


Figure 6. The real time histogram transport for data quality monitoring. The ROOT histograms accumulated in each event process on basf2 are sampled and transferred via the sockets, and added to the shared memory. By repeating the procedure in multiple steps, the histograms are finally collected on the DQM master node.

and a rough selection is done to discard the background events. The expected reduction is about 1/2.

Then the full event reconstruction is performed using the same code as offline and the final event selection is done based on the physics event skim such as the hadronic event selection with the scaled calibration events added. The final overall reduction factor is expected to be more than 3.

4 Status of DAQ in Physics Run

Belle II DAQ has been working without serious trouble in the first physics run. The nominal Level 1 rate is up to 3.5 kHz since the accelerator luminosity is still low, i.e. maximum of $L = 1.2 \times 10^{34} \text{cm}^{-2}\text{sec}^{-1}$, which is only 1/50 of the design value. The nominal data taking efficiency during the whole run period of about 4 months is 80 to 85% which includes the various debugging periods. In the stable physics run, the efficiency is more than 90%. Fig. 8 shows the DAQ accepted rate as a function of time with the descriptions of various troubles. The upper plot shows the case for a “good” day, while bottom is for a “bad” day.

Here is the list of DAQ troubles experienced during the first physics run.

1. **ttlost** The link for the trigger timing distribution from the master timing unit to the detector front end is broken. This occurs frequently, up to once every few hours in the worst case. Mostly it is caused by the incomplete implementation of the interface

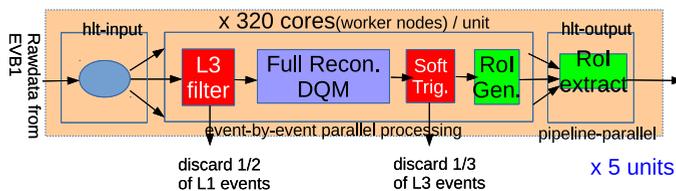


Figure 7. The HLT event processing software chain. The event data are first fed into the Level3 filter for noise reduction. The full event reconstruction follows and the software trigger decision is made. The RoIs for PXD data reduction are calculated using the processing result.

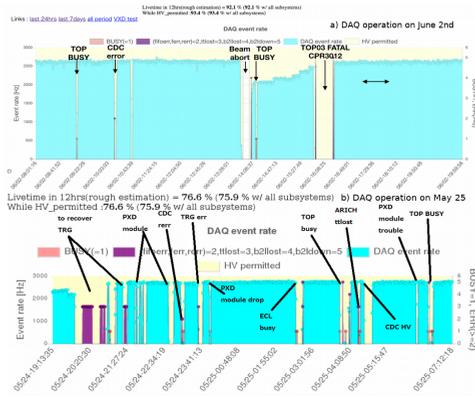


Figure 8. DAQ accepted rate as a function of time. The upper figure shows the operation on a good day, where DAQ efficiency is 93.4%. The lower figure shows that on a bad day, where the efficiency is 76.6%.

firmware in the detector front-end. The debugging of firmware is still in progress. This problem is recoverable just by resetting the link.

2. **b2llst** The data transfer link between the detector front-end and the COPPERs is broken. This is also caused by the incomplete programming in the detector-side firmware and is fixed by a reset. In some cases, the full reprogramming of the firmware is required. This may be caused by single event upsets from the beam radiation and an investigation is in progress.
3. **Hang up of COPPER CPU** The hang up of CPU on a COPPER board sometimes occurs. Restoring functionality requires power cycling the crate, thus restarting of all of COPPERs it contains.
4. **Entire breakdown of slow control system** The entire breakdown of slow control occurred several times. It is mostly caused by a malfunction in the daemon of NSM2 framework due to flooding messages generated by a particular node. An improvement in the NSM2 daemon was made and the frequency of the problem was reduced.

5 Operation of High Level Trigger

During the phase 3 running, the HLT is operated without the Level 3 filtering and the full event reconstruction is performed for all taken events. Using the reconstruction results, the events are classified in the categories of hadronic, Bhabha, $\mu\mu$, $\tau\tau$, $\gamma\gamma$, cosmic, other scaled calibration events including random triggered events, and other background events. Based on the classification, the software triggering is performed.

Under the nominal accelerator condition, the overall reduction factor by the HLT trigger is measured to be about 8, which is more than the design value, because of the high background conditions.

There are two major problems in the operation of HLT. One is the stuck HLT at the run stop and restart. The problem is caused by the incomplete clean up of Linux IPC resources used in the ring buffer. They are supposed to be cleaned up at stop or restart. The HLT stop and restart sequence is implemented using the Linux signal to abort the execution of the

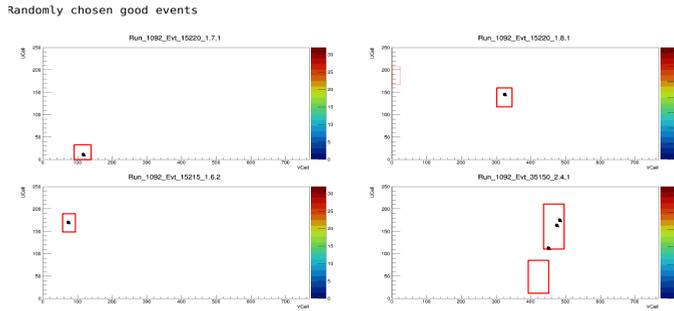


Figure 9. The matching of RoI and PXD hits. The reconstructed RoI boxes are shown with the hits in PXD for four randomly selected events.

processing framework (**basf2**), however, the signal handling is found to be incomplete. After improving the signal handler, the operation stabilized.

The other problem is that it takes a very long time of up to 5 minutes to stop the HLT. It was found that the time is consumed to collect and store ROOT histograms in files at run end used for the data quality monitoring. The histogram collection is done independently of that for the live monitoring in order to store not only histograms but also ROOT TTrees. The total number of histograms is more than 7500, and all of them have to be collected from 1600 cores. The collection is done in three steps and every step takes more than 1 minute. The intermediate files are placed on the same NFS file system which deteriorates the I/O performance. Finally the histogram file storage has since been moved to the live transport basis, and the time was shortened to less than 30 seconds.

6 RoI feedback to Pixel Readout

The feedback of RoIs to the PXD readout system is tested in the beam run. The RoIs are obtained using the tracks reconstructed by the real time HLT processing and sent to the PXD readout system. Fig. 9 shows the hit map of PXD for some events. The rectangular boxes show the RoIs obtained by the HLT processing. As seen, the PXD hits are observed inside the RoI boxes and the RoI feedback is confirmed to work as expected.

7 Summary

The Belle II DAQ system was operated in the first physics run for 4 month. The typical Level 1 trigger rate was up to 3.5 kHz and the overall data taking efficiency including the debugging period is 80 to 85 %. In the stable beam run, the efficiency is more than 90%.

The High Level Trigger system is operated with 5 units equipped with a total of 1600 cores. The same offline event reconstruction is performed in real time, and the software trigger decision is made based on the physics event skimming. The operation is stable after fixing a couple of troubles. The average reduction factor of the software trigger is measured to be about 8.

The feedback of RoI to the PXD readout system for the data reduction is tested in the physics run and its operation is confirmed.

References

- [1] Z.Dolezal and S.Uno (ed.), “Belle II Technical Design Report”, KEK Report 2010-1 (2010).
- [2] Y.Ohnishi, *et al.*, “Accelerator design at SuperKEKB”, Prog. Theor. Exp. Phys. 2013, 03A001 (2013).
- [3] M.Nakao, “Timing Distribution for the Belle II Data Acquisition System”, JINST 7, C01028 (2011).
- [4] T.Higuchi, *et al.*, “Modular Pipeline Readout Electronics For The Superbelle Drift Chamber”, IEEE Trans. Nucl. Sci. **52**, 1912 (2005).
- [5] D.Sun, Z-A.Liu, J.Zhao, H.Xu, “Belle2Link: a Global data Readout and Transmission for Belle II Experiment at KEK”, Phys. Procedia, 37, 1933-1939 (2012).
- [6] R.Itoh, S.Lee, *et al.*, “Implementation of parallel processing in basf2 framework for Belle II”, J. Phys. Conf. Ser. **396**, 022026 (2012).
- [7] T.Gessler, W.Kuhn, *et al.*, “The ONSEN Data Reduction System for the Belle II Pixel Detector”, IEEE Trans. Nucl. Sci. **62:3**, 1149-1154 (2015).
- [8] M.Nakao, S.Y.Suzuki, “Network shared memory framework for the Belle data acquisition control system”, IEEE Trans. Nucl. Sci. **47:2**, 267-271, 2000.
- [9] A.J. Kozubal *et al.*, “Run-time environment and application tools for the ground test accelerator control system”. Nucl. Instr. Meth. A, 293(1-2), 288-291 (1990), <https://epics.anl.gov/index.php> .
- [10] J. Hatje *et al.*, “Control System Studio (CSS)”, Proc. of ICALEPCS2007, 37-39, Knoxville, USA, Oct. 2007, <http://controlsystemstudio.org> .