Machine-Learning-Based Waveform Discrimination in the Front-End Electronics of the Belle II Central Drift Chamber for Cross-Talk Noise Reduction

Yun-Tsung Lai*, Taichiro Koga, Yu Nakazawa, Nanae Taniguchi

KEK/SOKENDAI

ytlai@post.kek.jp

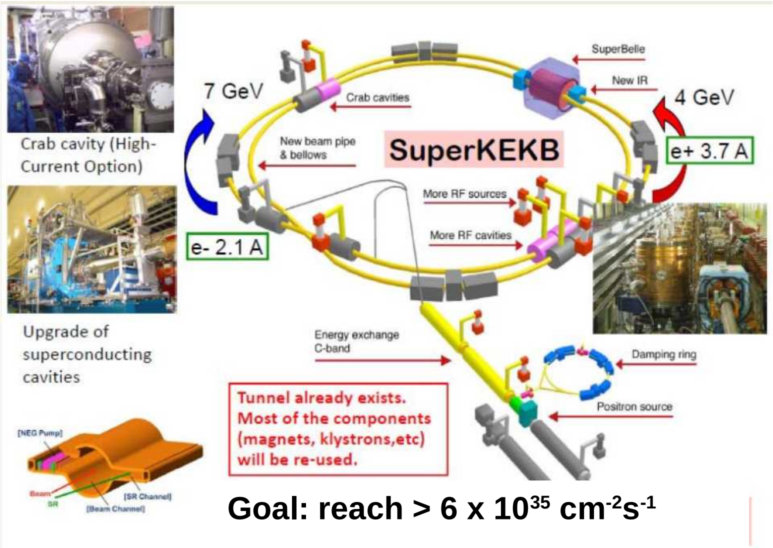
25th IEEE Real Time Conference 2026

29th, May, 2026

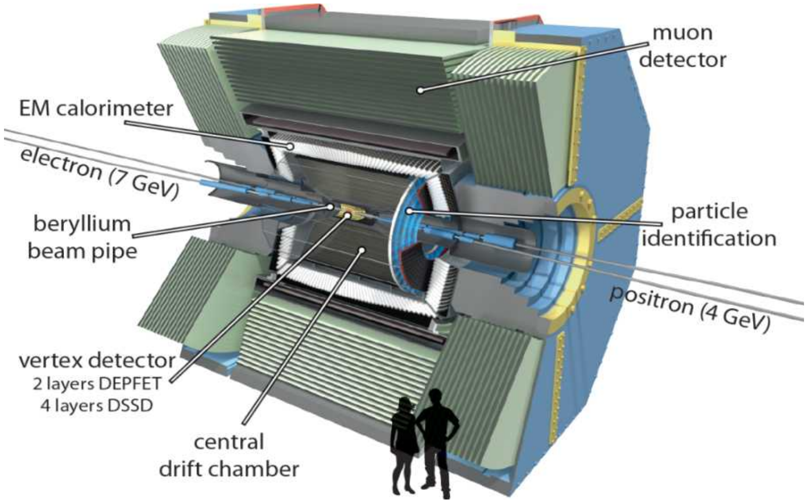
SuperKEKB & Belle II

- **SuperKEKB**: Upgrade from KEKB.
  - More than 30 times larger luminosity of KEKB with nano beam scheme.
  - Asymmetric energy collider:
    - 7.0 GeV e^- and 4.0 GeV e^+ for $\Upsilon(4S) \rightarrow B\overline{B}$.
  - L_{int} = ~820 fb^{-1} up to May 2026.

- **Belle II**: General purpose spectrometer
  - Physics target of Belle II:
    - Rare B, τ, charm physics, Dark Matter search, CP Violation.
  - Requirements with high luminosity for DAQ:
    - High trigger rate (at least 30 kHz) for L1 trigger and an event size up to 1 MB.

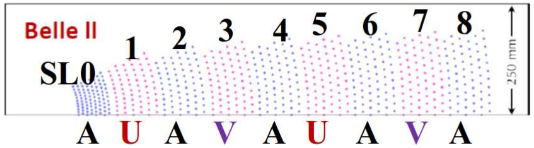


Goal: reach > 6 x 10^{35} $cm^{-2}s^{-1}$

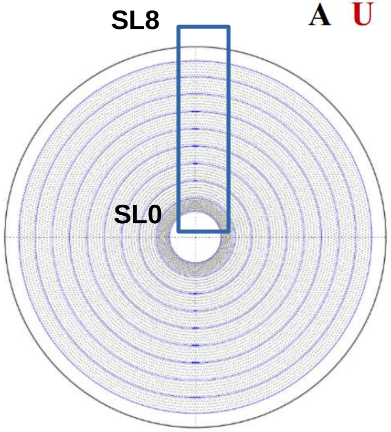


Belle II Central Drift Chamber

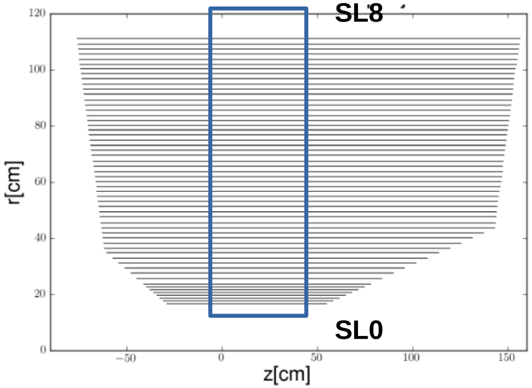
- **Central Drift Chamber (CDC)**: One of the major charged tracking devices in Belle II.
  - CDC data are used for both offline tracking software and real-time tracking in Level-1 trigger system.
  - 14336 anode wires, 56 wire layers into 9 Super Layers (SL).
  - Gas: mixture of Helium and ethane
  - Consists of axial and stereo wire layers



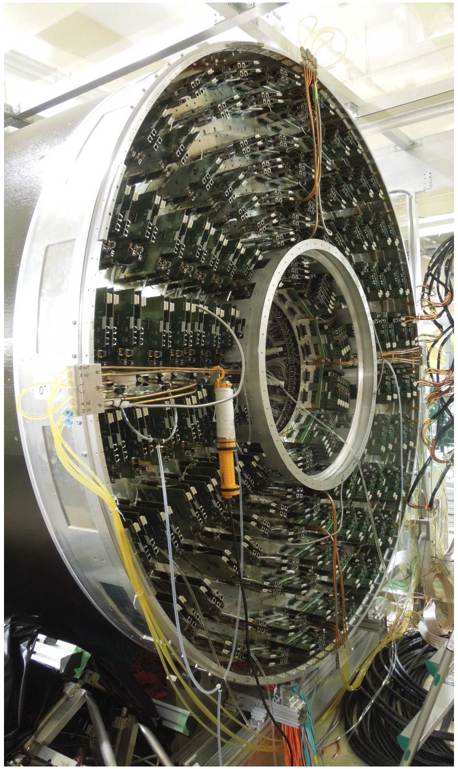
- **A**: axial wires
- **U**, **V**: wires with two small stereo angles.



Transverse cut view



Longitudinal cut view



CDC at an endcap with Front-end Electronics

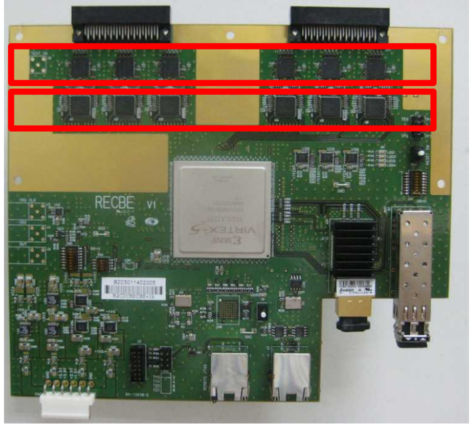
Cross-talk noise in CDC Front-End Electronics

- **CDC Front-End Electornic (FEE)**: based on Xilinx Virtex-5 FPGA with custom ASICs for Analog-to-Digital Converter (ADC), Amplifier-Shaper-Discriminator (ASD).
  - 48 anode wires per FEE, where 8 wires for 1 ASIC for ADC or ASD
- **Cross-talk noise**: Multiple hits (defined as TDC signals where a pulse exceeds the threshold) occurring in localized regions due to large energy deposits
  - Almost all channels within an ASIC
  - Possible reason: The ASD is based on single-ended design. It is suspected that a large charge injected into one channel may draw charge from the decoupling capacitors of neighbors.
  - Also related to Belle II electornics system's environment. Dependence on SuperKEKB beam condition is also observed.

Coverage of 1 FEE: 48 wires

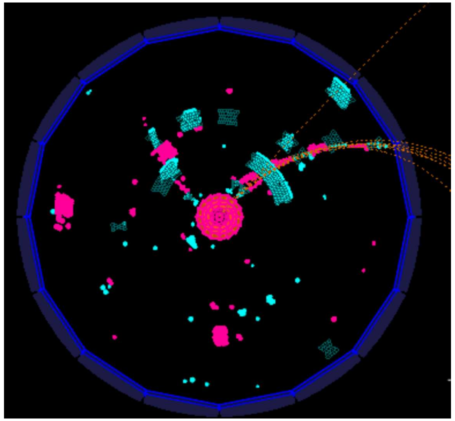
6 ASICs for ASD

6 ASICs for ADC



S. Shimazaki et al., NIM, Sect. A 735 192 (2013)

CDC FEE board featured with Xilinx Virtex-5 FPGA



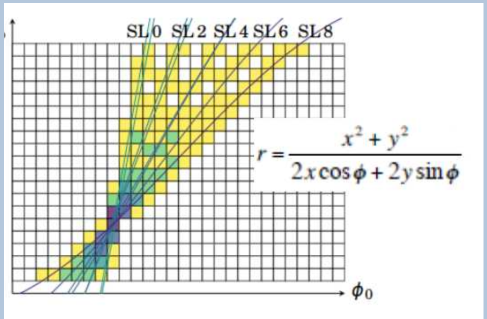
A Belle II event display to demonstrate cross-talk noise in CDC

Cyan point: Axial wire hits
Magenta point: Stereo wire hits

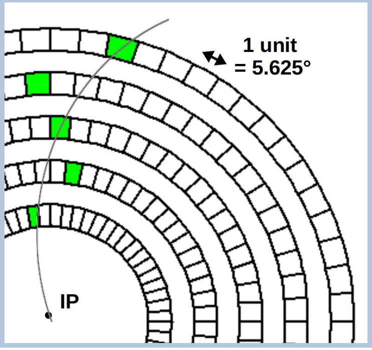
Tracking in Level-1 trigger

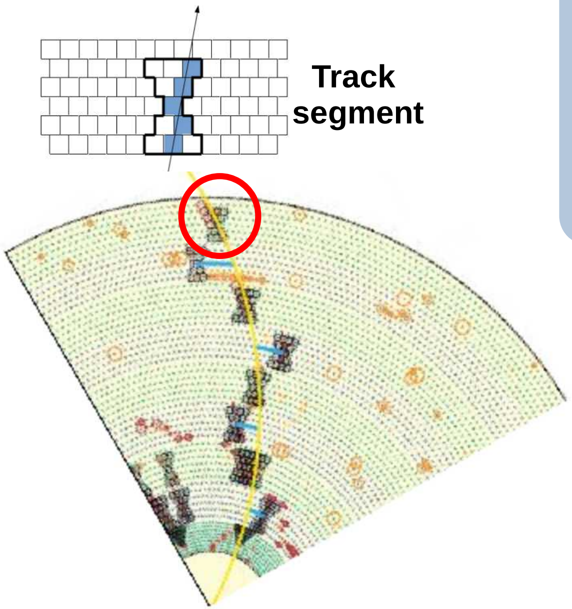
- **Tracking in L1**: Based on Track Segments ("TS", grouped wire hits of specific shapes) in each SL

2D Finder: Hough Finder with 5 axial SLs

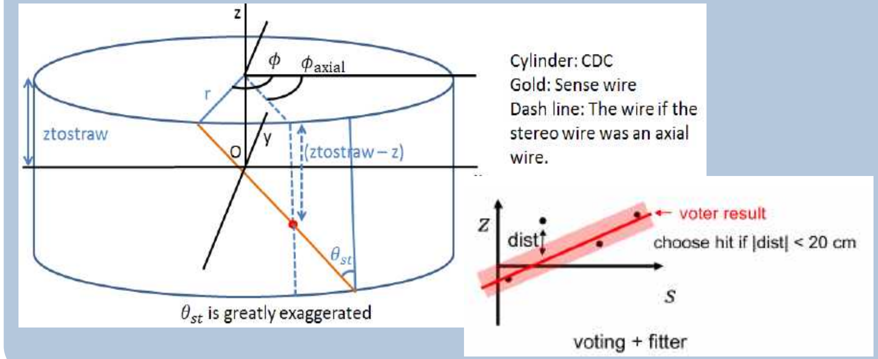


Short tracker: Pattern recognition with inner 5 SL

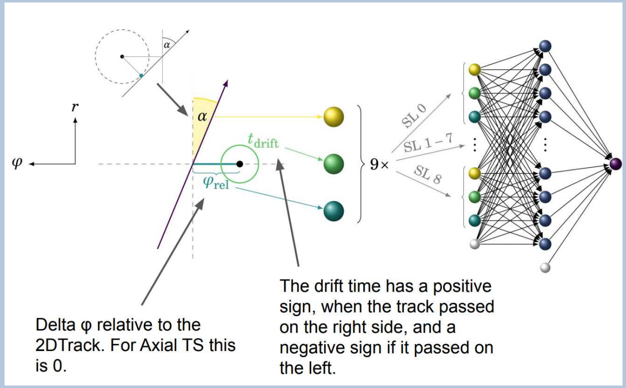




3D Fitter: 2D track with 4 stereo SLs

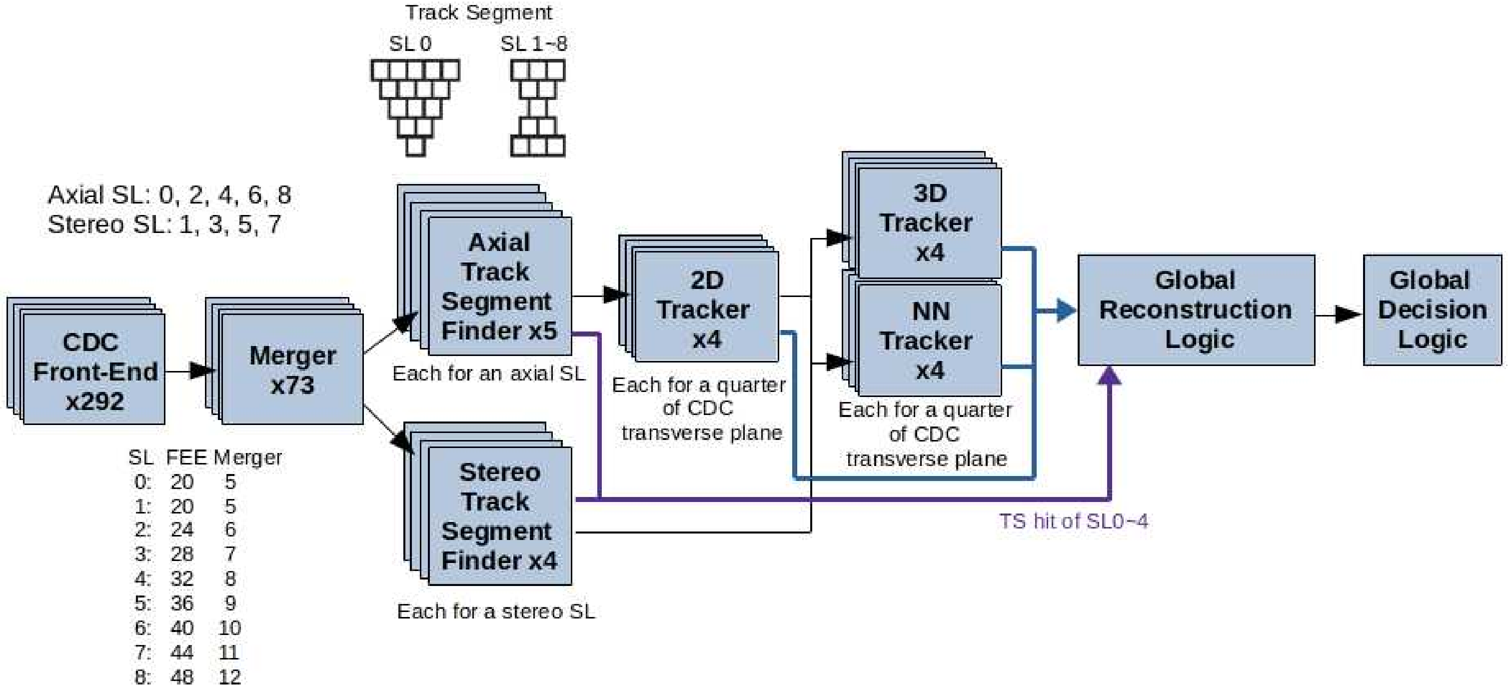


Neural 3D Fitter: 2D track with 4 stereo SLs



Tracking in Level-1 trigger (cont'd)

- **Tracking in L1**: Based on Track Segments ("TS", grouped wire hits of specific shapes) in each SL
  - Major trigger device: Universal Trigger board "UT4" based on Xilinx Virtex UltraScale FPGA
  - Overall latency limit: 4.4 µs
    - Which is one of the major consideration for any extension on the trigger system



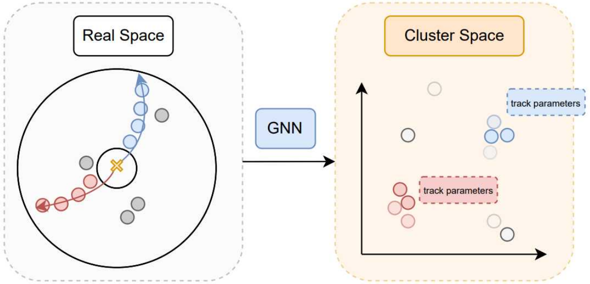
Impact on tracking in Level-1 trigger from the cross-talk noise

- **Impact on L1 tracking**: Segment-based tracking approach relies on a reduced input information, it is highly sensitive to cross-talk noise.
  - Finder-based: Fake tracks would be found due to segments induced by the noise.
  - Fitter-based: Fake segments at the stereo SL would affect the fitting result and hence affect the z0 resolution.

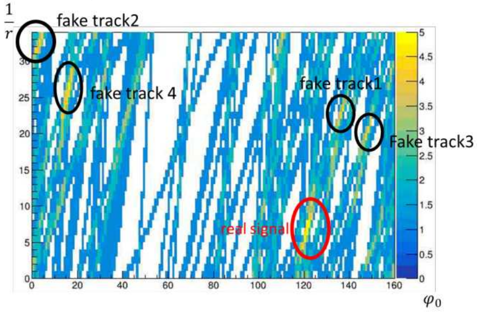
- Fake track trigger rate up to O(10 kHz) ~ O(100 kHz) is possible as luminosity and beam background increase in future SuperKEKB operation. Ping Ni, master thesis, 2022

- Present efforts in L1 Trigger:
  - 2D Hough: Clustering on the nearby peaks in the conformal plane
  - 2D Hough: Segment-based → Wire-based
  - GNN noise filter and tracking

Lea Reuter, Conntecting the Dots 2025



Ping Ni, master thesis, 2022



- In L1 back-end, we have relatively high-end FPGA devices to improve tracking for this noise issue.
  - **How about the FEE? Is there any additional information which can be utilized only in FEE?**

ML in detector FEE

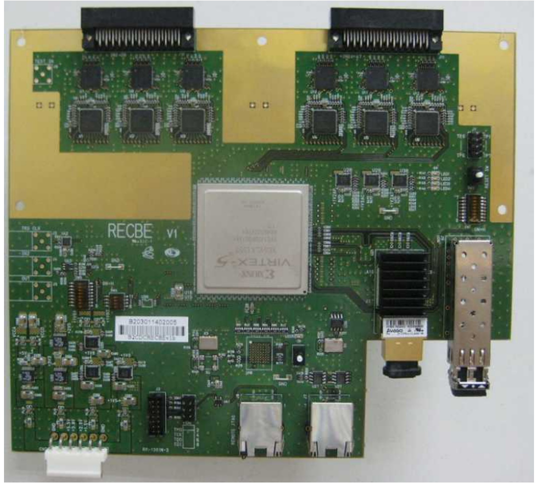
- ML-based reconstruction within a global scope of detector in real-time trigger backend system has been a popular application.
- **For ML in detector FEE**:
  - Technical difficulties such as small coverage over detector and small resource from the FPGA chip.
  - There is also unique information which could be utilized.

- For this study: The CDC FEE boards transmit wire hits (TDC) and timing of the hits to the L1 trigger system.
  - The ADC waveform is not utilized present trigger tracking algorithms.
  - The ADC waveform information is the unique information within CDC FEE.

- Such kind of application could be helpful in terms of post-hardware-level improvement on the detection system.
  - Hardware-level calibration, or more physics-wise purpose.

ML-based noise filter in FEE by waveform discrimination

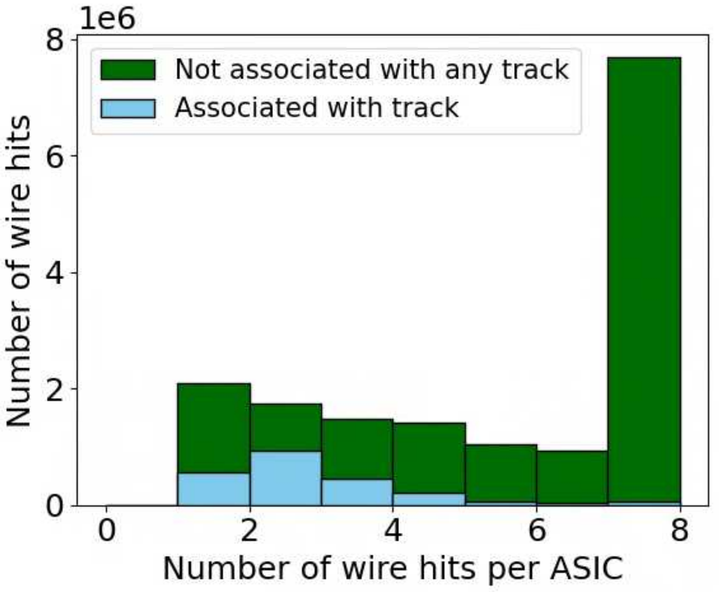
- This study: **Use ML-based methods on the ADC waveform in CDC FEE to distinguish the cross-talk noise**
  - Virtex-5 FPGA of CDC FEE: 97,280 Look-Up Tables (LUT), 97,280 Flip-Flops (FF), and 128 DSP slices
    - A small chip.
  - Has to **minimize the processing latency**: limit from L1 processing.
  - ML module will be implemented **channel-by-channel**: 48 channels per FEE.
    - Model for each wire channel can be trained individually
    - Latency can be much smaller: without any dead time in the pipeline.
    - But resource occupation is a big issue, so the RTL module has to be pretty **compact**.



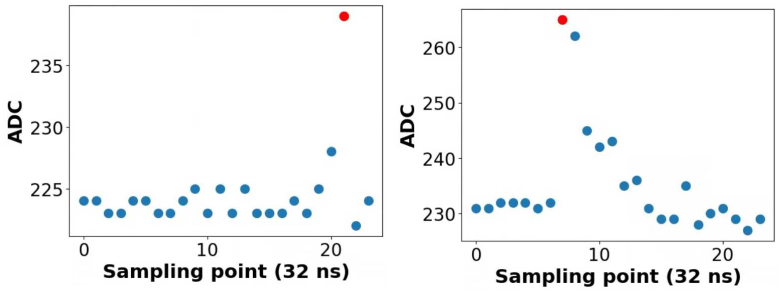
- We are expecting that this pioneering work can demonstrate the technical feasibility of such **ML in FEE** approach, and be further promoted for other detectors' application.

Investigation on the ADC waveform data

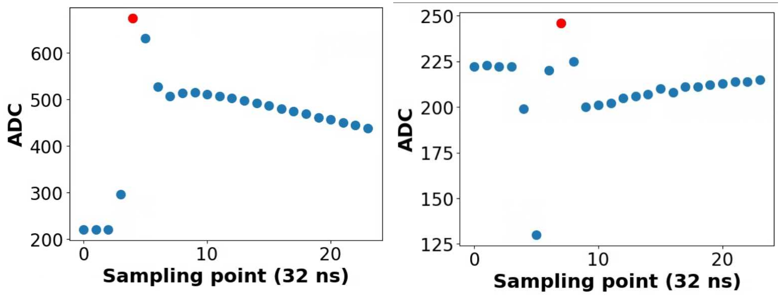
- The ADC waveform information: 31.4 Msps (~32 ns per sampling point).
- With the help from Belle II offline tracking software, we can isolate the wire hits associated to any tracks.
  - For the background hits (not associated to any tracks), the fraction where all the eight wire are fired within an ASIC occupies the majority.
  - Cross-talk noise: defined as those background wire hits from ASICs with more than six hits.



Red point at peak: The ADC point corresponding to TDC



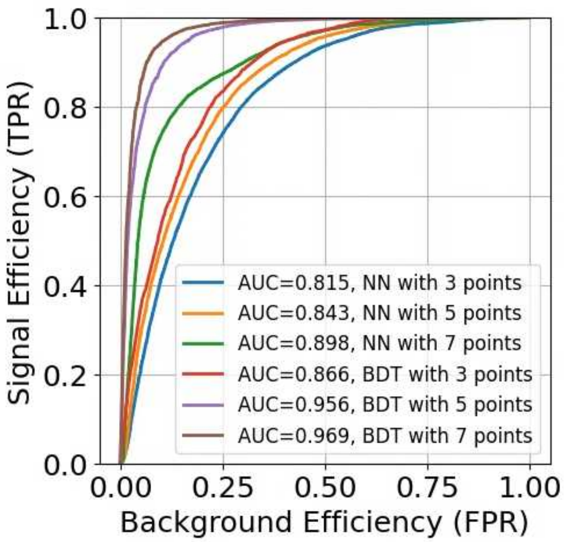
ADC waveform of signal (from charged track)

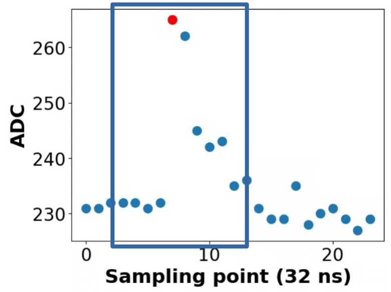


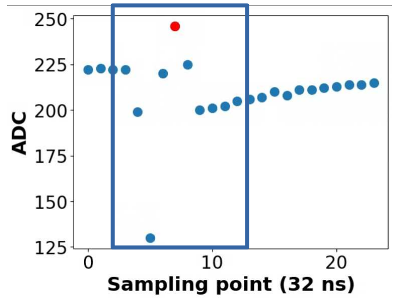
ADC waveform of "cross-talk noise"

ML Model building

- We first explored on various models for Neural Network (NN) and Boosted Decision Tree (BDT)
  - Input features: the ADC values of a few sampling points, where the middle sample corresponds to the TDC timing point.
  - NN: 2 hidden layer.
  - BDT: < 10 estimator, depth < 5





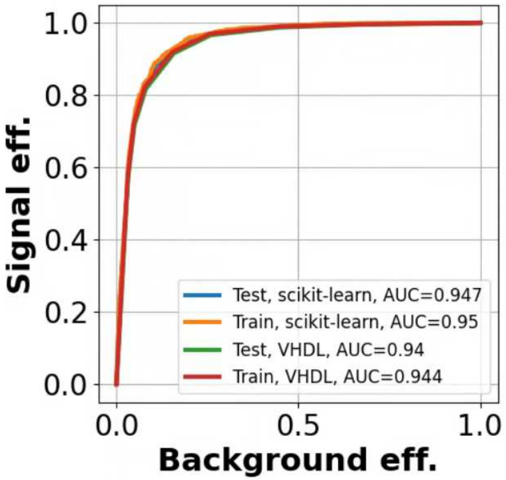


- With a very quick look, BDT models perform better the NN models in most of the cases.
  - Considering the points of latency and FPGA resource usage, BDT is the best selection especially for channel-by-channel implementation.

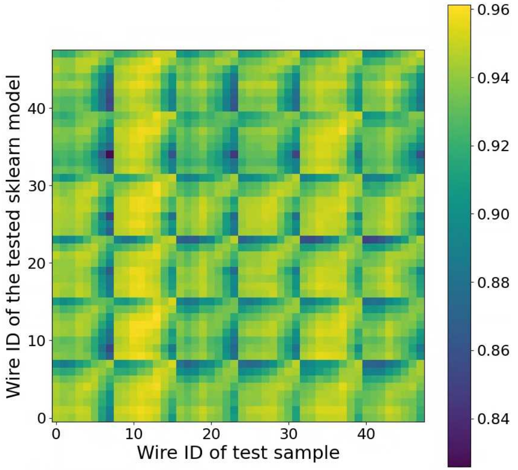
BDT model building

- **Training data selection**:
  - Among all the waveform data, the 48 channels' data are separated and trained individualy
    - Considering the geometry effect within FEE board.
  - Input features: ADC of 5 sampling points, where the middle one corresponds to TDC timing point.
    - Only 1 TDC point within the selected window is allowed.
  - 10,000 samples each of signal and background were selected
    - Signal:
      - Associated with track
      - From ASICs with fewer than 4 hits
    - Background:
      - Not associated with track
      - From ASICs with more than 6 hits
- **BDT model**:
  - Four estimators with a max depth of four
  - GBDT from scikit-learn.
  - Quantization by Conifer

S. Summers et al., JINST 15 P05026 (2020)



ROC curve from one of the 48 wire channel data

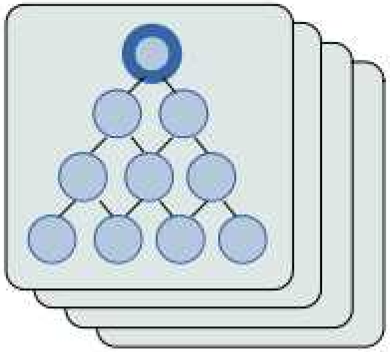


Geometry effect among the 48 wire channel data

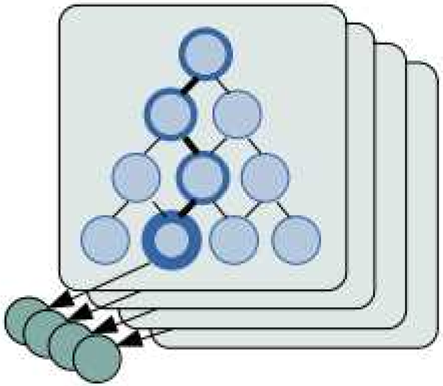
Inference to Virtex-5 FPGA

- The Conifer package supports not only quantization but also transferring the BDT model in python into High-Level-Synthesis (HLS) or VHDL modules.
  - Xilinx HLS software tool does not support Virtex-5 series.
  - We wrote our own BDT VHDL module to further reduce latency and resource usage:

1. List each logic condition corresponding to each node:
1 clock-cycle,
$X*(2^0 + 2^1 + ... + 2^{Y-1})$ registers



2. List each condition corresponding to each output node with AND:
1 clock-cycle,
$X * 2^{Y-1}$ registers



3. Sum over all the outputs:
0 clock-cycle

	FF	LUT	Latency (clock-cycle)
Conifer HLS (based on Kintex-7)	104	1876	7
Conifer VHDL	274	293	9
Home-made VHDL	**96**	**90**	**2**
48X fraction for Virtex-5	**4608 (4.7%)**	**4320 (4.4%)**	

- **We achieved a very short latency and a very small resource usage!**

Performance

- We first evaluated the wire-level performance of the BDT models using the offline ADC waveform data.
- Cut level definition: Among the quantized BDT classifiers, we align all the 48 models with a cut threshold with the same signal efficiency of ~99% as "cut level 1".
- The performance for inner SL (SL0) and others (SL1 - 8) are evaluated separately, since the wiring configuration (granularity) is different.
- Efficiency and rejection rate (= 1 - efficiency) of the three categories:
  - Signal: Associated with tracks
  - Background: Not associated with tracks
  - Cross-talk noise: Not associated with tracks & From ASICs with more than 6 hits

SL0

Cut level	Signal efficiency	Background rejection	Cross-talk rejection
1	0.9926	0.2712	0.4406
2	0.9811	0.4158	0.5990
3	0.9594	0.5432	0.7281
4	0.9130	0.6746	0.8229

SL1 - 8

Cut level	Signal efficiency	Background rejection	Cross-talk rejection
1	0.9954	0.1494	0.2327
2	0.9874	0.3245	0.4709
3	0.9728	0.5134	0.6978
4	0.9507	0.6332	0.8317

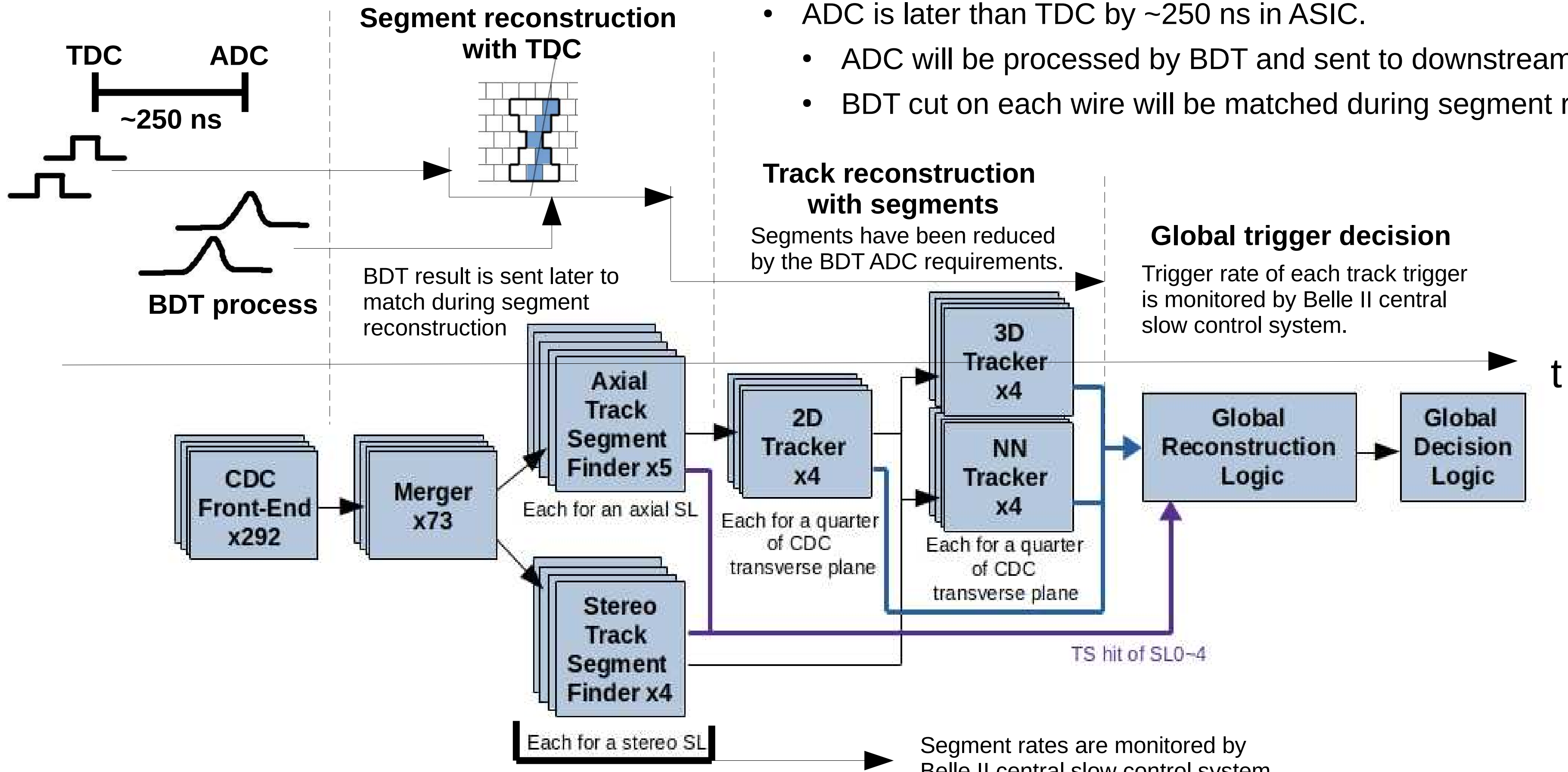
- **With a loss on signal wire of 1%~2% , cross-talk noise can be reduced up to 50%~60%.**

Experimental evaluation: setup

- The performance was also evaluated in realistic experimental environment.
- SuperKEKB single beam run: with only Low Energy Ring (LER), 4.0 GeV electron, under a beam current of 1A
- CDC FEE firmwares of the Virtex-5 FPGA were prepared for different BDT cut thresholds.
- For each condition, we took data in ~20 min operation.
  - Data were taken for the default condition, and BDT cut level 1 and level 2.
- Real-time monitoring by Belle II central slow control system: direct response from the BDT filter
  - L1 trigger rate of the track triggers
  - Track segment rate: representing wire hit rate
- Offline study:
  - Track trigger effieciency over physics condition: to check if the signal wire hits are retained.

Experimental evaluation: setup (cont'd)

- In FEE firmware, TDC is sent to segment finder first.
- ADC is later than TDC by ~250 ns in ASIC.
  - ADC will be processed by BDT and sent to downstream later.
  - BDT cut on each wire will be matched during segment reconstruction.



Experimental evaluation: result on L1 track trigger rate

- L1 track trigger rate is the direct and straightforward check on the effect of CDC wire noise reduction.

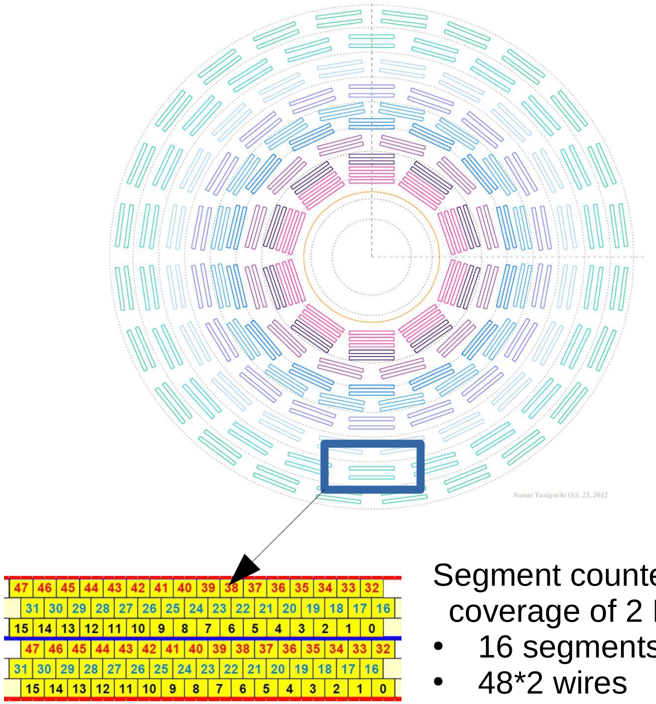
- Three types of track triggers are checked:
  - 2D full track: Hough finder using axial SLs (0, 2, 4, 6, 8)
  - Neural full track: Using all the 9 SLs (0 - 8)

Condition	2D full track (kHz)	Neural full track (kHz)
No requirement	82.6 ± 1.8	5.3 ± 0.1
BDT ADC cut level 1	43.1 ± 0.9	4.1 ± 0.1
BDT ADC cut level 2	23.3± 1.1	3.2 ± 0.1

- **L1 Track trigger rates can be reduced up to more than 50%.**

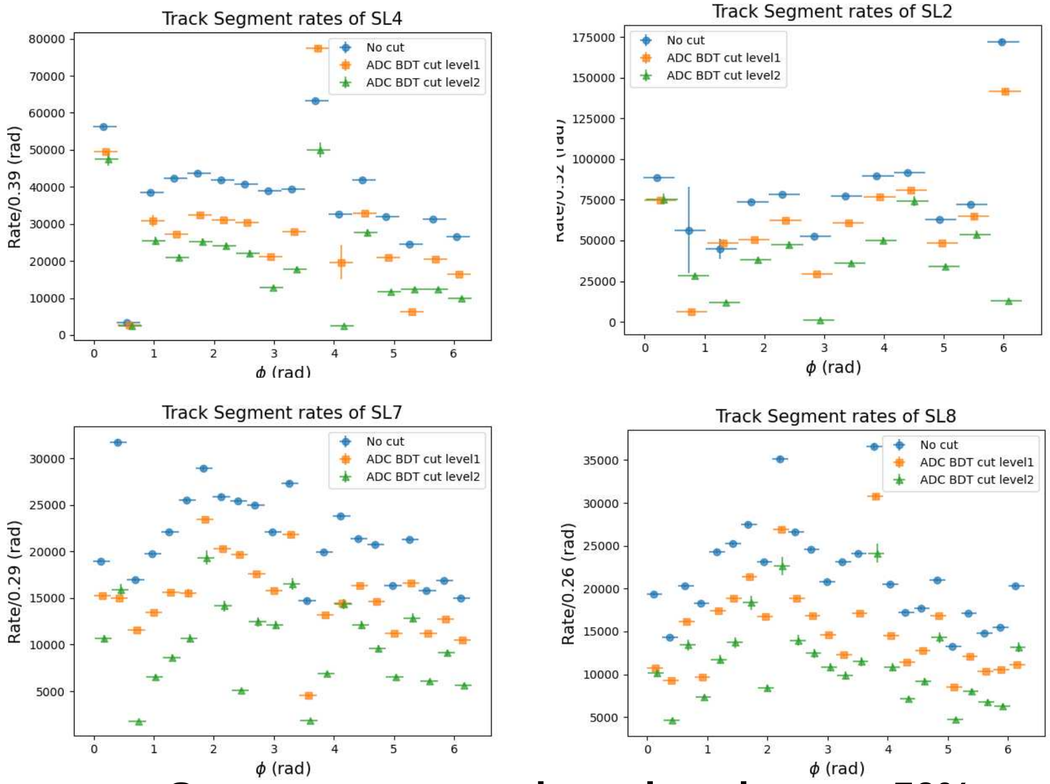
Experimental evaluation: result on track segment hit rate

- In the present trigger system, there is no monitor on the wire hit rate.
  - Instead, there is monitor on segment hit rate over a 2-FEE covergae: 16 segments, 48*2 wires.
  - Also a direct response from the noise reduction effect.



Segment counter over coverage of 2 FEE:
- 16 segments
- 48*2 wires

Hit rate in some of the SLs



- **Segment rates can be reduced up to ~50%.**

Experimental evaluation: result on L1 track trigger efficiency

- To check if the signal wire hits are well retained, we checked on the L1 track triggers' efficiency over selected events with specific physics condition.
  - Calorimeter trigger is used as reference.
  - Track condition is obtained by offline tracking software:
    - Event with tracks from IP: |dr| < 10 cm && |dz| < 10 cm
    - Event with tracks from IP in barrel: |dr| < 10 cm && |dz| < 10 cm && passing through entire CDC
    - Event with no track at all

Experimental evaluation: result on L1 track trigger efficiency (cont'd)

Events with tracks from IP

Condition	2D full track	Neural full track
No requirement	0.736 ± 0.019	0.484 ± 0.014
BDT ADC cut level 1	0.706 ± 0.024	0.473 ± 0.018
BDT ADC cut level 2	0.632 ± 0.015	0.390 ± 0.011

Events with tracks from IP in barrel

Condition	2D full track	Neural full track
No requirement	0.944 ± 0.032	0.790 ± 0.032
BDT ADC cut level 1	0.935 ± 0.046	0.744 ± 0.039
BDT ADC cut level 2	0.896 ± 0.032	0.702 ± 0.027

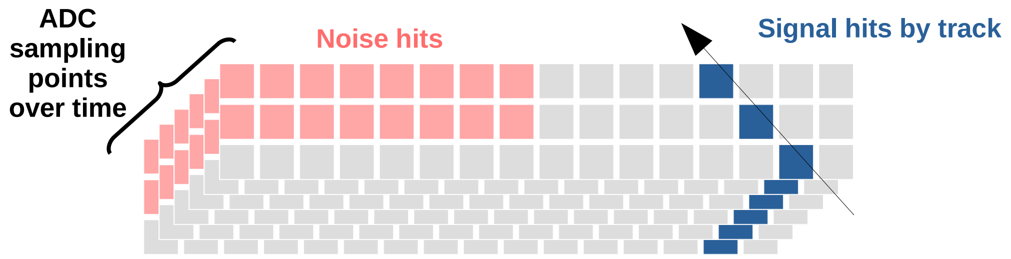
Events with no track at all

Condition	2D full track	Neural full track
No requirement	0.277 ± 0.001	0.044 ± 0.001
BDT ADC cut level 1	0.182 ± 0.001	0.028 ± 0.001
BDT ADC cut level 2	0.173 ± 0.001	0.028 ± 0.001

- **Signal event efficiency is kept with a loss up to ~10%.**
- **Background event (no track) efficiency is reduced by ~40%.**

Prospect for this study, and in Belle II

- In this study, BDT shows good performance and feasibility for channel-by-channel implementation.
- In a FEE, contouring/clutering among the 2D image of the wire hit map could be another potential approach.
  - As each wire contains the ADC waveform points, it is a 3D image processing:
    - Suitable for CNN or GNN



- **For other detectors in Belle II**, we are also expecting more similar application of ML implementation in FEE for pre-processing purpose before the reconstruction in trigger backend.
  - Calorimeter: hadron and e/gamma separation based on waveform
  - Time-Of-Propagation counter: Feature extration
  - KLong-Muon chamber: Photo-electron counting in scintillator strip.

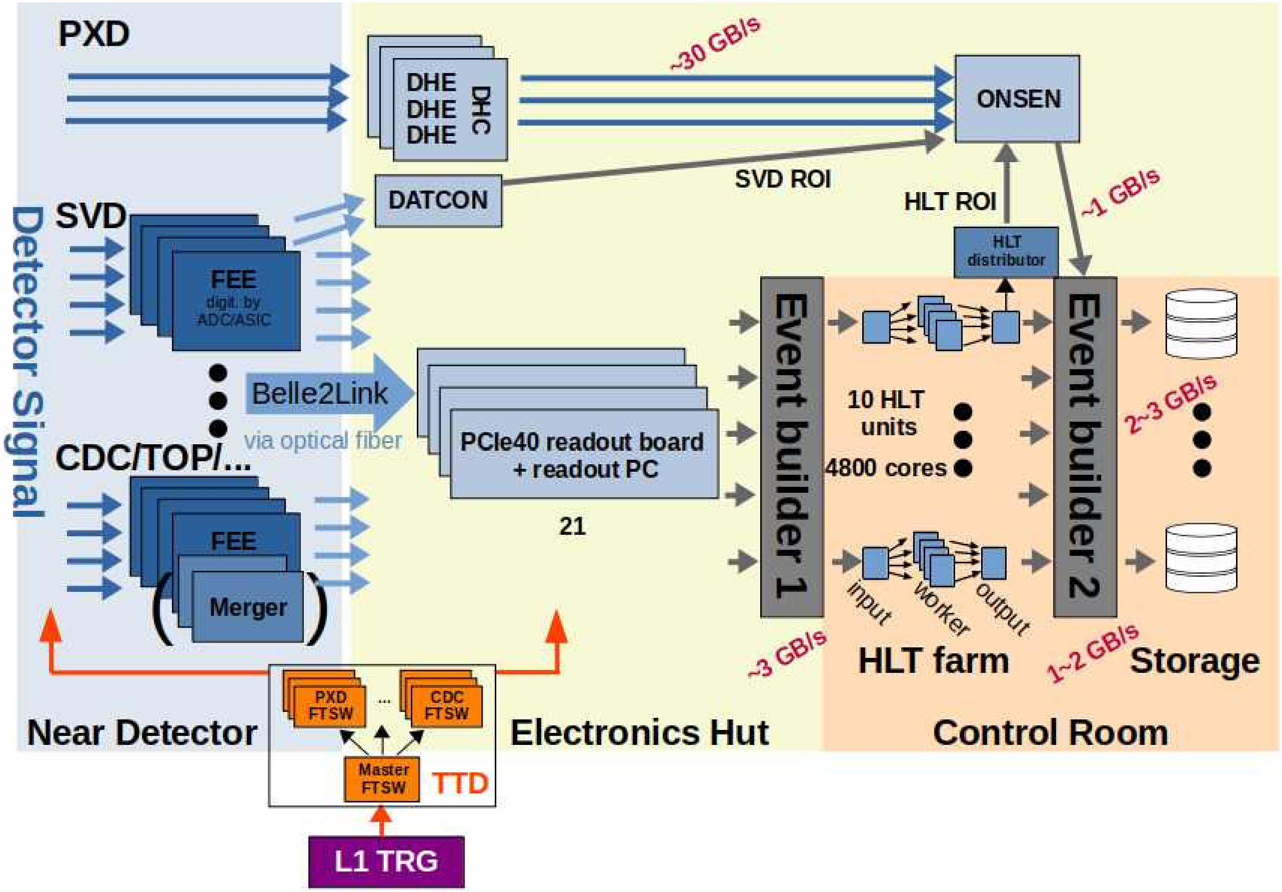
Summary

- We performed a study of ML-based noise filter in Belle II CDC FEE by processing the ADC waveforms.
  - The cross-talk noise is a long-term issue in the Belle II CDC operation, since the noise would introduce direct impact to the real-time tracking in L1 trigger system.
- We chose to use BDT with a channel-by-channel implementation in the Virtex-5 FPGA of CDC FEE. In the deployment within FPGA firmware, both the processing latency and resource usage are minimized.
- Performance has been evaluated using offline waveform data and real experimental beam operation.
- We are also expecting that this pioneering study could demonstrate prospect for other similar ML in FEE application in other detector systems, with other ML methods, for other use cases in near future.

Backup

Belle II DAQ system

- Target of performance: 30 kHz trigger rate from L1, ~1% of dead time, and a raw event size of 1 MB.



Belle II L1 trigger system

- Using four sub-detectors.
- Overall latency limit: 4.4 µs
- Most of the major components are based on a Universal Trigger devices of Xilinx Virtex UltraScale.