Discussion for Beam Background Merging

T. Cao
August 12, 2020
Refer to Kyle's talk in a software meeting

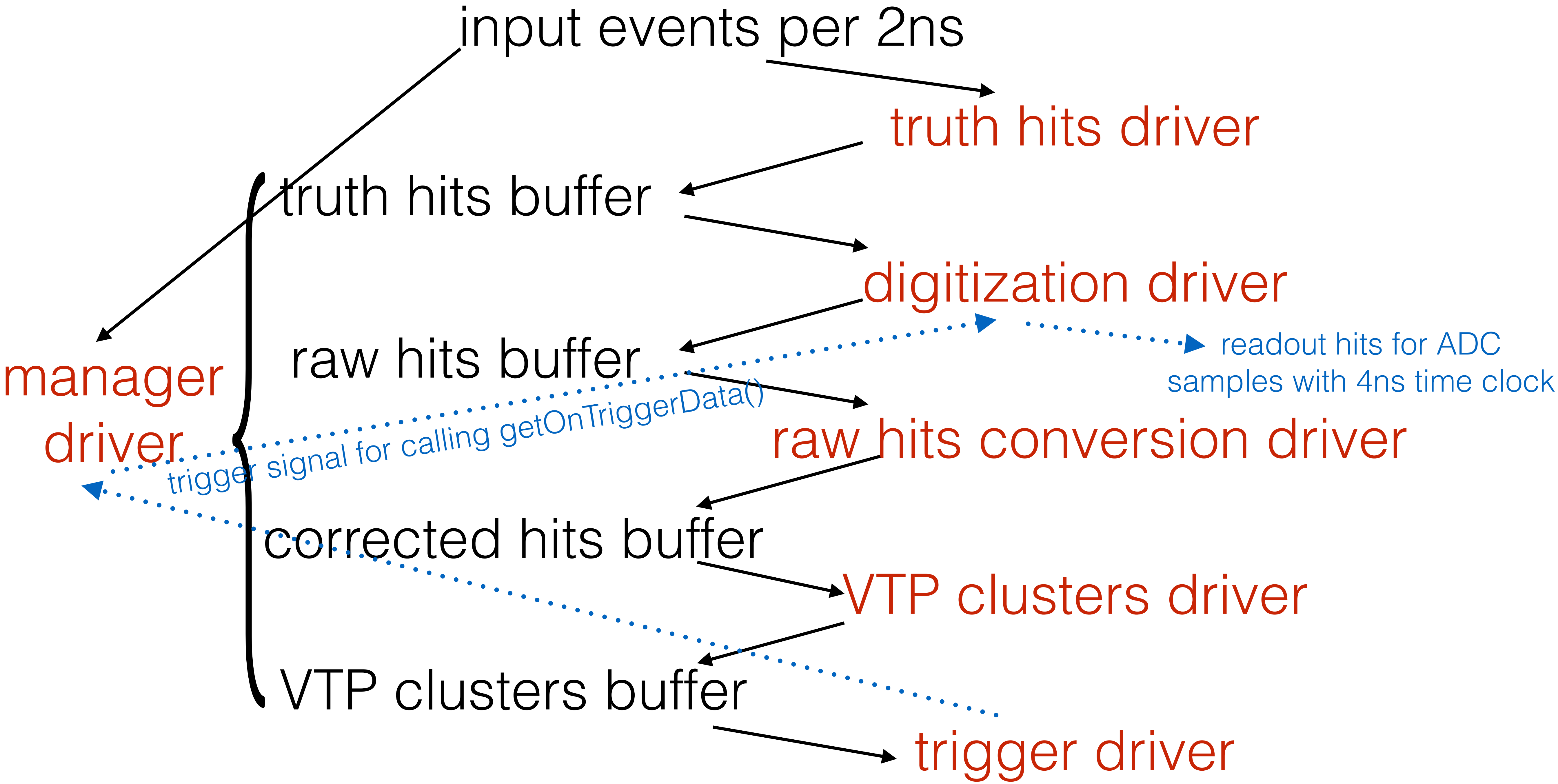
Introduction

- To take into account effects of beam background in MC, signal events (rad, tritrig, wab, ap) are required to be merged with beam background.
- For current MC processing,
    - Beam events are produced by EGS5+SLIC. The produced beam is biased, without wab and tritrig physics. The beam production is time consuming.
    - Signal events are produced by MadGraph+SLIC
    - Produced signal events are filtered and spaced (interval between each two signal events = 250) firstly, and then merged with beam event by event.
- To significantly speed up MC production and improve quality of MC samples, we consider to merge real background data with MC signal samples.
- To make a plan for merging, firstly we need to understand real data and the readout system.

Real Data

- What real data should be applied?
  - Data from good random/pulser runs with the same luminosity as MC (beam current and target thickness are the same between them)
  - Pulser events must be placed in a random order so that each MC run with merging of pulser data does not have the same background signature.
- Raw real data are ADC samples for crystals of Ecal, holes of hodoscope, and strips of SVT. Their time-window offset and width in DAQ are not the same.

Processing in the Readout System (Ecal as example)



input events per 2ns

truth hits driver

truth hits buffer

digitization driver

readout hits for ADC samples with 4ns time clock

manager driver

raw hits buffer

trigger signal for calling getOnTriggerData()

raw hits conversion driver

corrected hits buffer

VTP clusters driver

VTP clusters buffer

trigger driver

Digitization Processing

- Truth hits are obtained from a truth-hits buffer managed by manager
- For a hit belong to a crystal, hit energy is smeared by adding noise
- According to hit energy after smearing, a voltage pulse is emulated and sampled with a fixed length
- Each of samples for voltage pulses belong to the same crystal is added into an internal voltage buffer labeled by the same crystal ID
- The digitization internal clock (4ns time clock) keeps track of which point in the voltage buffer is the “current” time. The “current” voltage buffer cell is read and converted into a ADC sample, and the ADC sample is written into a ADC buffer labeled by the same crystal ID
- According to the ADC buffer, raw hits are reconstructed
- Once a trigger signal is sent to manager, a function getOnTriggerData() is called, and readout hits for a window of ADC buffer of each crystal is output. The output is used by reconstruction, like reconstruction for real data.

Three Ways for Merging

- “Merging” could occur at three different levels:
  - Way 1 (occurs before the readout system): raw real data is reconstructed to be hits, and then reconstructed hits are merged with MC truth hits before the readout system. It means we following everything of existing processing in MC chain, except changing simulated truth hits into real hits for background.
  - Way 2 (occurs in the readout system): raw real data with ADC samples merges with ADC buffer in the digitization driver; to do this, we need to build continuous ADC samples using raw data, and add the continuous ADC samples into ADC stream for simulated signal hits in the digitization driver.
  - Way 3 (occurs after the readout system): readout hits for simulated signal hits are produced by the readout system firstly, and then output ADC samples are merged with raw real data event by event.

Way 1

- The only thing we need to do is convert raw real data into hits. However, data after conversion is significantly biased, which is not expected.
- Therefore, the way 1 is denied.

Way 2

- The way 2 is the most accurate. But we need to use pulser events to build continuous ADC samples for merging with ADC stream in the digitization driver.

- To build continuous ADC samples, a natural way is to stitch events back to back. But obviously, ADC samples around the stitching points might be discontinuous.

- The issue happens since a pulse occurs near the beginning or end of a readout window (called as edge effect). We might build algorithm to identify if such case happens. If so, the event is skipped. We need to keep in mind that there are plenty of channels for Ecal, hodoscope and SVT. If such case happens for any channel, the event should be skipped.

Way 3

- The way avoids to stitch events, so the edge effect is avoided. Meanwhile, since it do not need to touch the complicated readout system, the code development is much easier than way 2.

- However, since “merging” takes place after the readout system, pulser data have no effects for trigger.

- A study need be taken to get estimation for the “trigger stealing” effect. If the effect is small, way 3 is a better option than way 2.

Discussion 1

- MC events are produced based on ideal detector configuration, but real data is based on real experimental facility with misalignment. It causes issues. For example, ADC samples for a SVT strip from a simulated event may be merged into ADC samples for a neighbored SVT strip from real data.

- To avoid such issue, can we build detector configuration with misalignment for MC production?

- How about setup of Ecal in the detector configuration of the most recent version? If it is consistent with the experimental facility?

Discussion 2

- ADC samples of pulser data have pedestals. Therefore, when building ADC sample for truth hits, we should not add pedestal.

- In the current readout system, pedestal from setup in DAQ configuration is applied for truth hits. Pedestal for ADC samples of pulser data is realistic, so it is a benefit for application of pulser data.

Discussion 3

- An event includes ADC samples for hits of Ecal, hodoscope and SVT. During DAQ, time-window offsets in the DAQ configuration are set to offset latencies of detectors. But does it perfectly align time among hits of each detector from the same event?

- From 2016 data, we extract offset of time difference between track and cluster is 55 ns. Does it mean time shift for ADC samples between SVT and Ecal?