

Software Readiness for 2019 Data

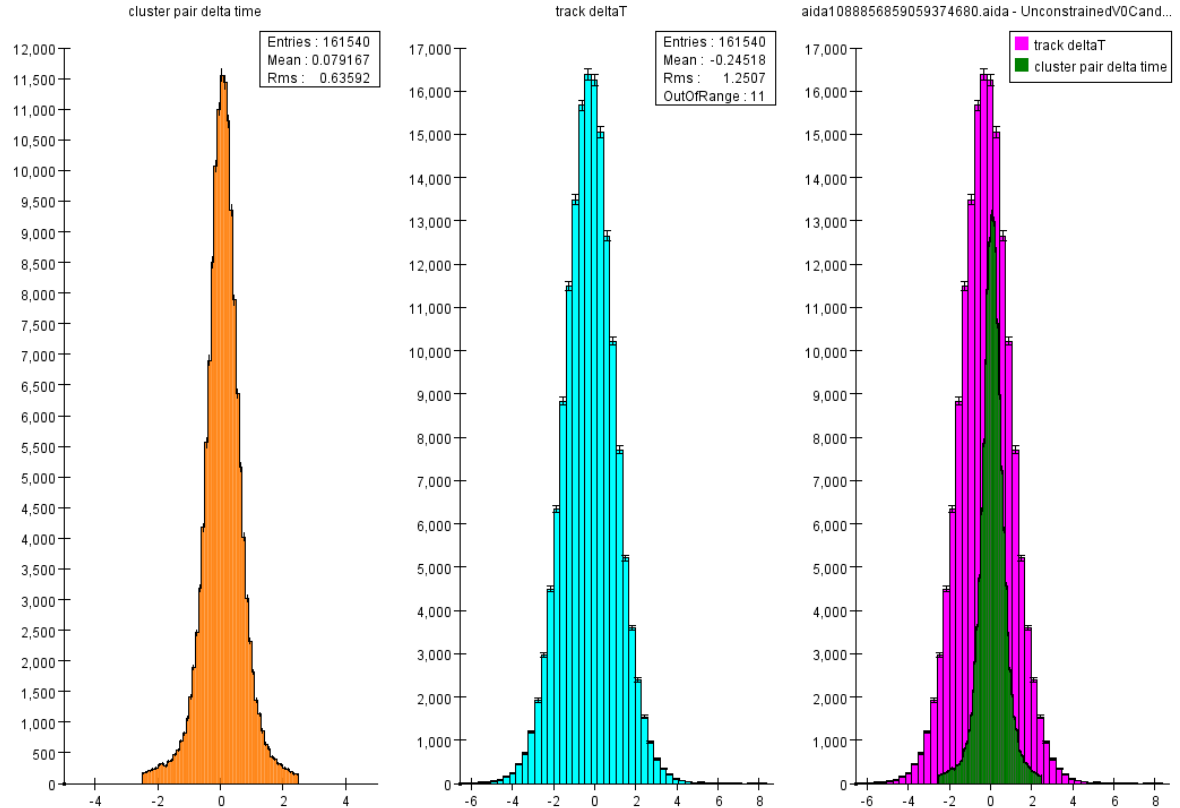
Norman Graf (SLAC)
HPS Software Meeting
September 23, 2020

Data Reconstruction Software Update

- ECal finishing up gains, sampling fractions and timing.
- Hodoscope software OK?
- SVT APV25 waveform fitting
 - Is the current fitting sufficient for our track timing?
 - replacing simplex with migrad improves fitting, gives uncertainties, but takes more time.
 - Need to study this ASAP, as we plan to drop raw data from output.
- **SVT trigger phase needs to be fixed for certain runs.**
- SVT actively working on alignment/calibration
 - PF has either ported or provided bindings to the C++ version of GBL code used to impose constraints on the alignment.
- Tracking group actively improving CPU performance
 - PF has replaced lcsim matrices and vectors with ejml
 - Robert actively developing Kalman Filter
 - **Need characterization and performance evaluation**
- Need a 2019 Event Flag Filter to remove obviously bad events
 - skip “monster” SVT events, wrong SVT position, wrong SVT voltage, etc.
- Output lcio files are bloated with extraneous data.
 - **Remove extraneous Drivers**
 - **Need to prune our data tree and remove unnecessary collections from lcio output**
- Memory footprint needs to be below 1GB to be efficient at JLab.

Track Timing vs Ecal Timing

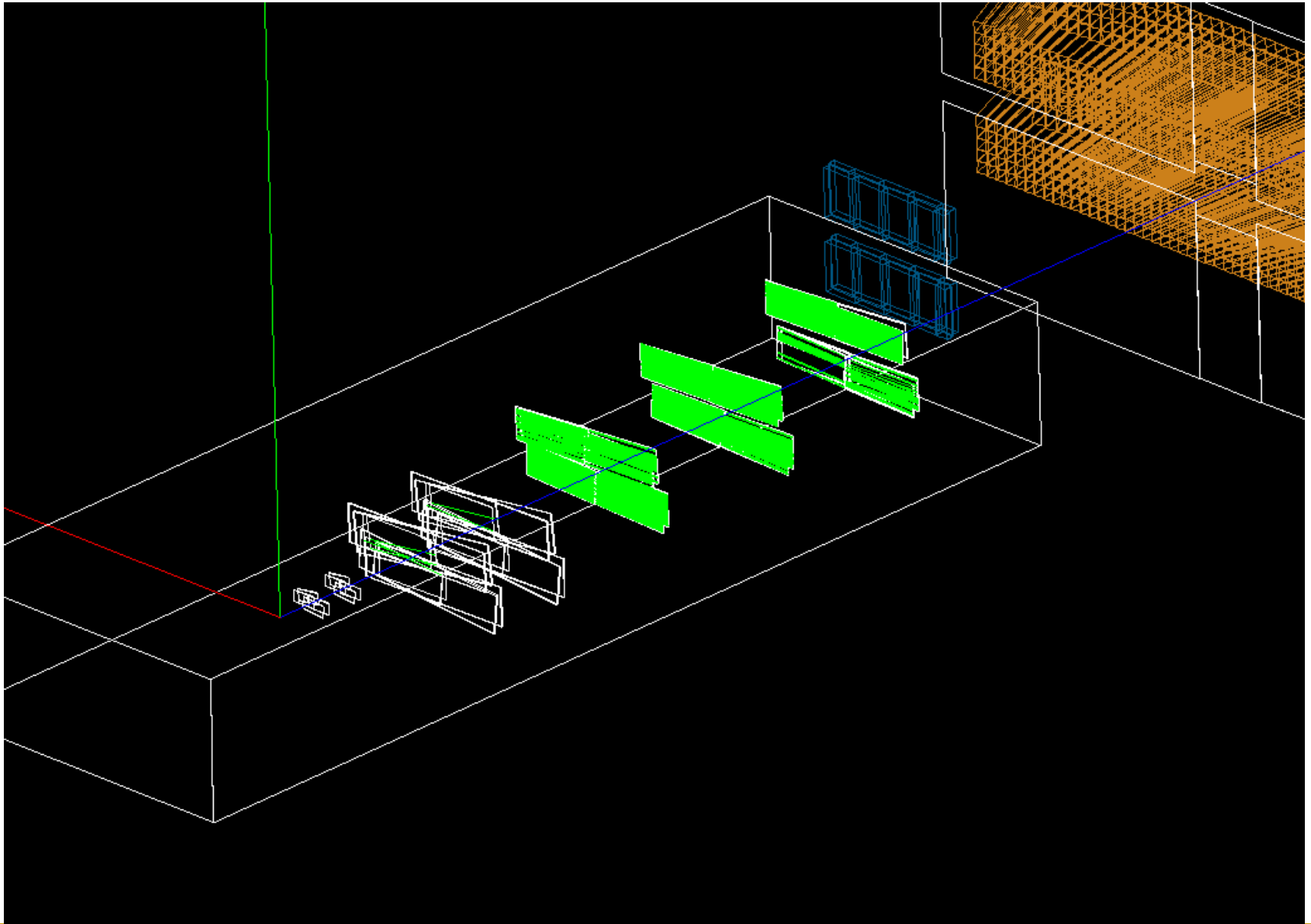
- Select V0 candidates with Ecal Clusters associated with each track.
- Track timing resolution a factor of two worse.
- Good enough?



SVT “Monster” Events

- Plan is to identify and then **skip both processing and writing out the event.**
 - This is new behavior as in the past we simply flagged such events.
- A skim of events containing more than 250 SVtRawTrackerHits is available to characterize the issues, develop the algorithms and test the efficiency of the cuts.
- git issue [iss731](#) addresses this.

SVT “Monster” Event



Logistics

- We need good estimates of our CPU needs to process the full 2019 “good” data sample
 - ~50 Billion events
 - Goal is better than 10Hz with a memory footprint of less than 1 GB
 - Recent work by PF to replace the freehep matrix and vector classes with ejml has shown impressive speedup in the tracking.
- We need good estimates of the amount of computing power we can rely on.
 - will be competing with CLAS for processing resources
- We need good estimates of our storage needs
 - ~600TB of evio data
 - Will tape access be an issue?
- Will inform the overall HPS data processing plan
 - e.g. do we start MC generation with our existing detector geometry?
 - e.g. do we “pre-process” the SVT data now while we wait for recon improvements?