Offline Reconstruction Readiness

Veronique Ziegler
JLab

CLAS12 First Experiment Workshop
Jefferson Lab
October 3, 2017

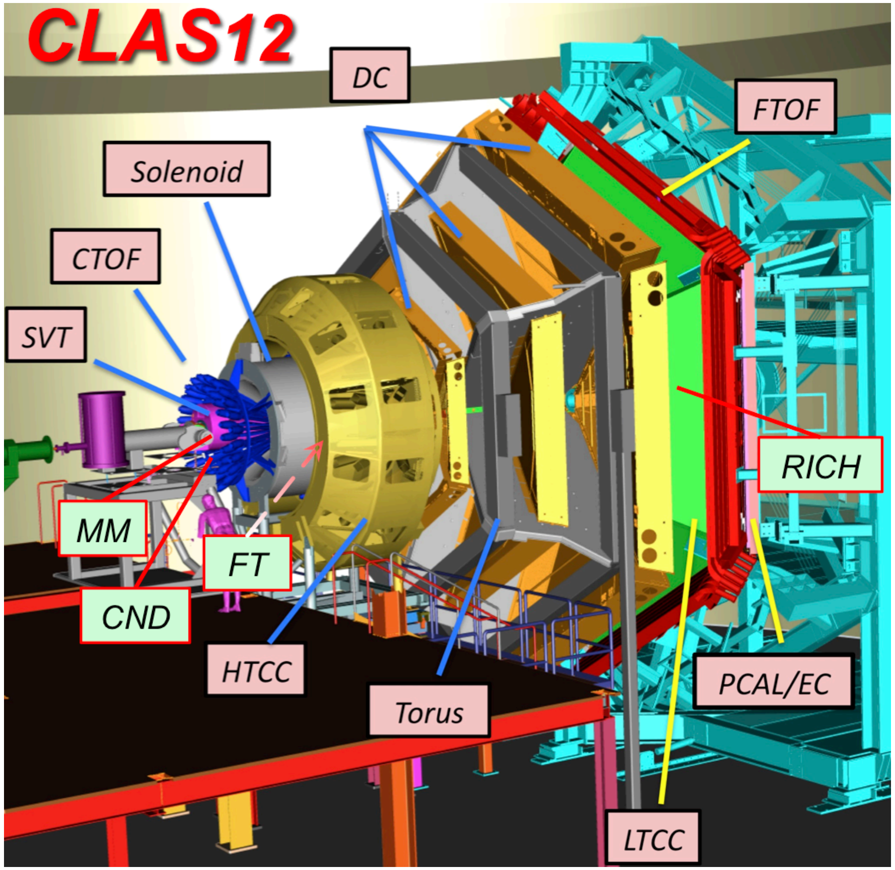
Event Reconstruction Status

Central Detector

- **S**ilicon **V**ertex **T**racker+**C**entral **M**icromegas **T**racker → tracking uses Kalman Filter fitting method
- **C**entral **T**ime **O**f **F**light → β (from path length) for PID
- **C**entral **N**eutron **D**etector →track β for neutron ID

Forward Detector

- **D**rift **C**hambers *Hit-Based & Time-Based* **T**racking → Kalman Filter fitting method to reconstruct tracks
  - **F**orward **M**icromegas **T**racker → refit DC tracks with FMT hits (resolution improvement)
- **F**orward **T**ime **O**f **F**light → β (track path length)
- **F**orward **T**agger calorimeter and hodoscope →id low angle electrons and reconstruct π^0's
- **E**lectromagnetic **C**alorimeter/**P**reshower **CAL**orimeter → detector responses for PID, reconstruction of neutrals
- **H**igh/**L**ow **T**hreshold **C**herenkov **C**ounter→ detector responses for PID, e- tagging using HTCC
- **RICH** detector (reconstruction in development) → detector response for PID

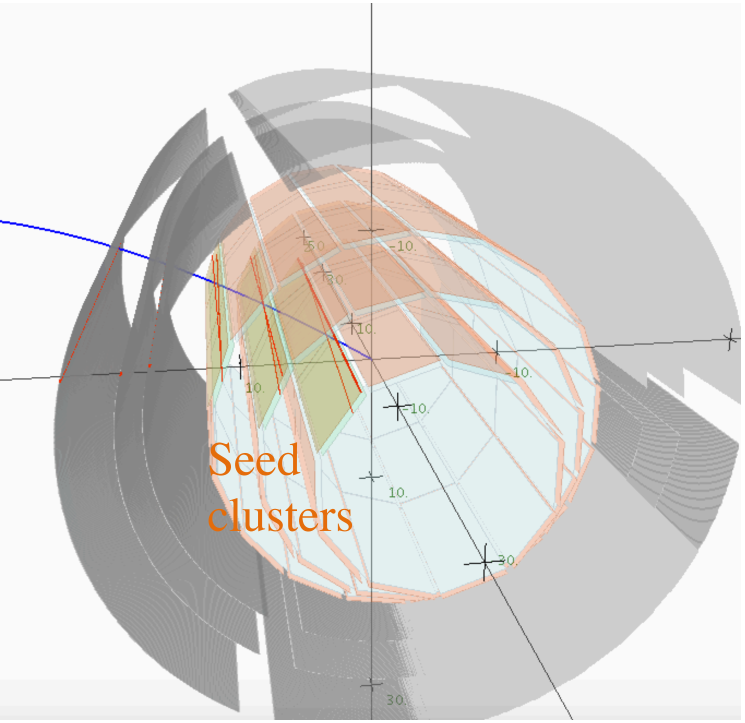


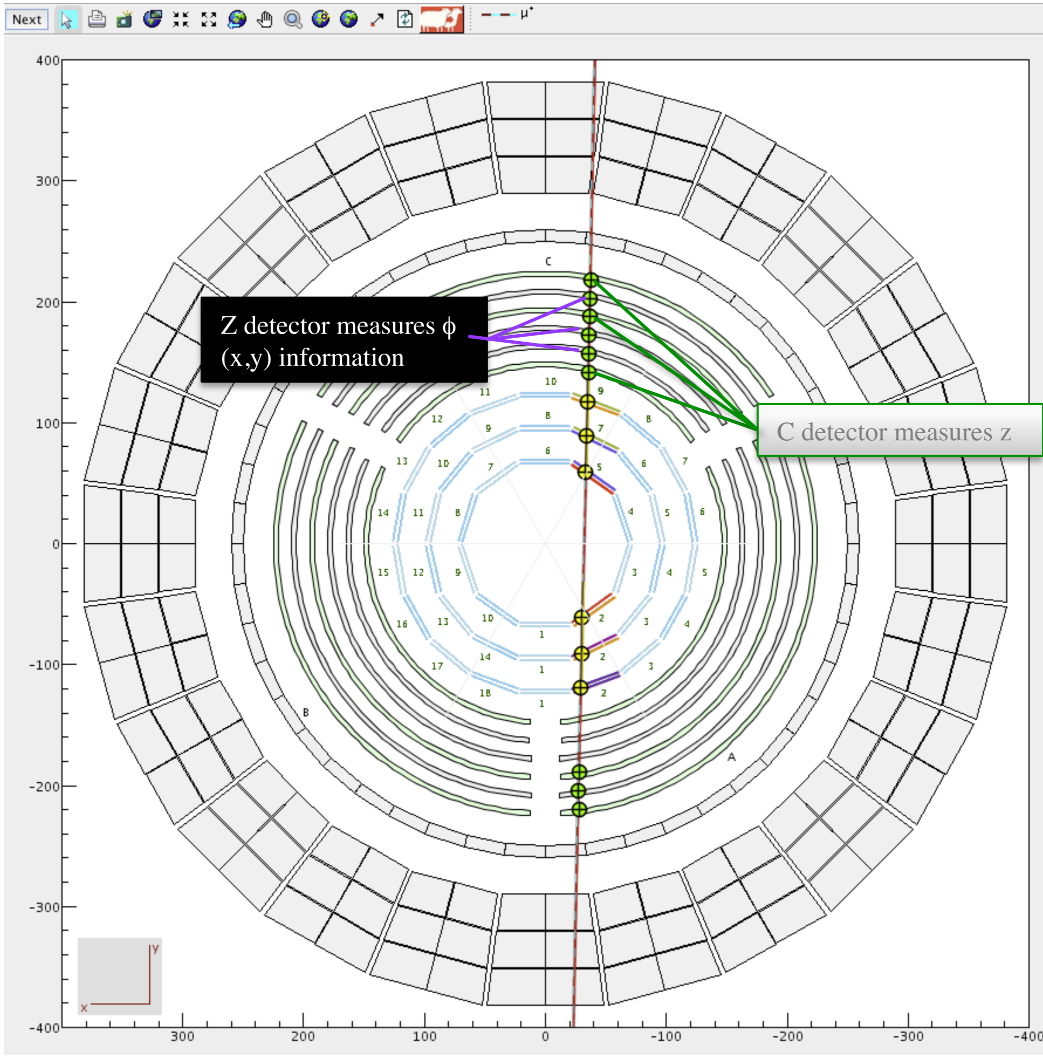
Event Builder

- Matches track to outer detectors, uses TOF, Calorimeters and Cherenkov detector responses for PID

CVT (SVT+BMT) Reconstruction

- Track ***seed***, contains SVT clusters + BMT pseudo-crosses
- Extract helical track parameters from Kalman Fit
- Validated on MC: cosmics & helical tracks





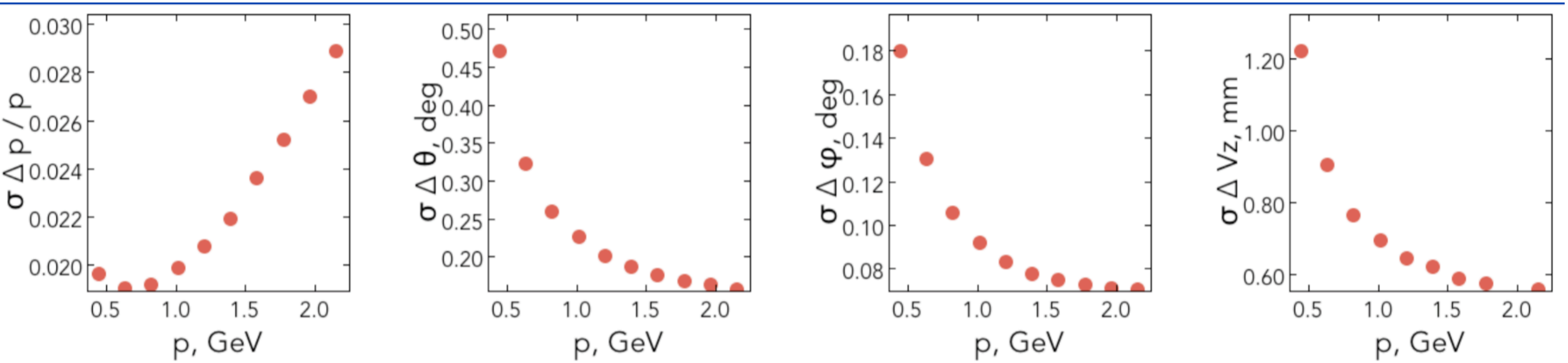
Reconstruction Readiness (Central Detector)

Test of the reconstruction :

1. Validations on MC:
   - Use calibration challenge data sample and kinematic-specific samples.
   - Verify reconstruction resolutions and efficiencies.

Tracking efficiency ~ 96%

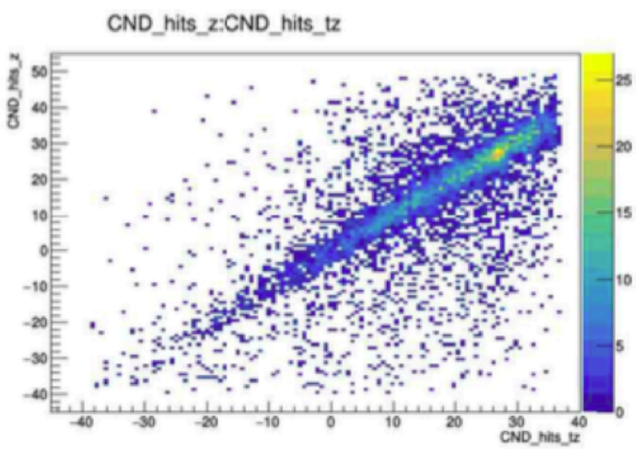
Central tracking resolutions better than specs $\sigma(\Delta p/p) <\sim 6\%$ @ 1 GeV, $\sigma(\theta) < 0.6$ deg., $\sigma(\phi) < 0.6$ deg.



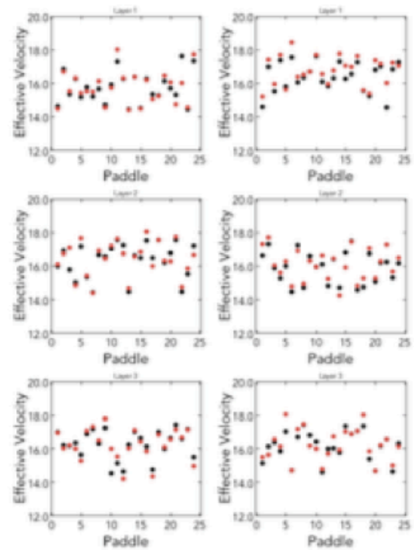
CND Reconstruction

P. Chatagnon (Orsay)

- Reconstruction of hit time, position and energy implemented
- Matching between SVT tracks and CND hits implemented
- Path length computed and included into the output bank



Z from CND reconstruction VS Z extrapolated from SVT



Calibrated effective velocities. Black dots are input values, red are extracted values.

Calibration process uses matching from CVT and gives good results.

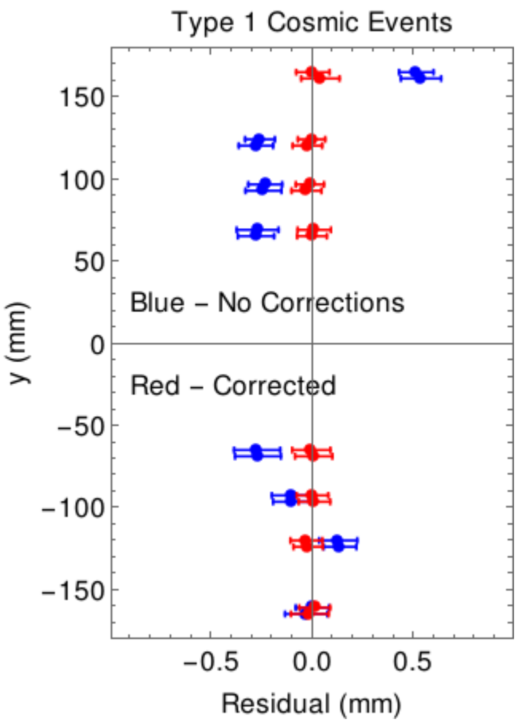
To do :

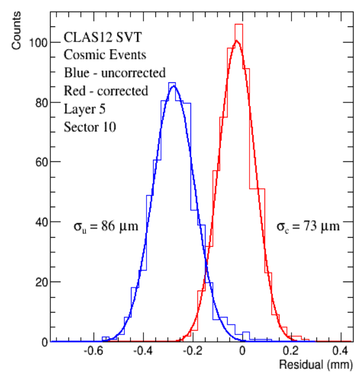
- Bug in the pathlength to be fixed (to take into account upstream edges of the paddles)
- Cross check energy reconstruction and energy calibration

Alignment of the SVT using Millepede

J. Gilfoyle

Reconstruction of cosmic events with millepede misalignments incorporated

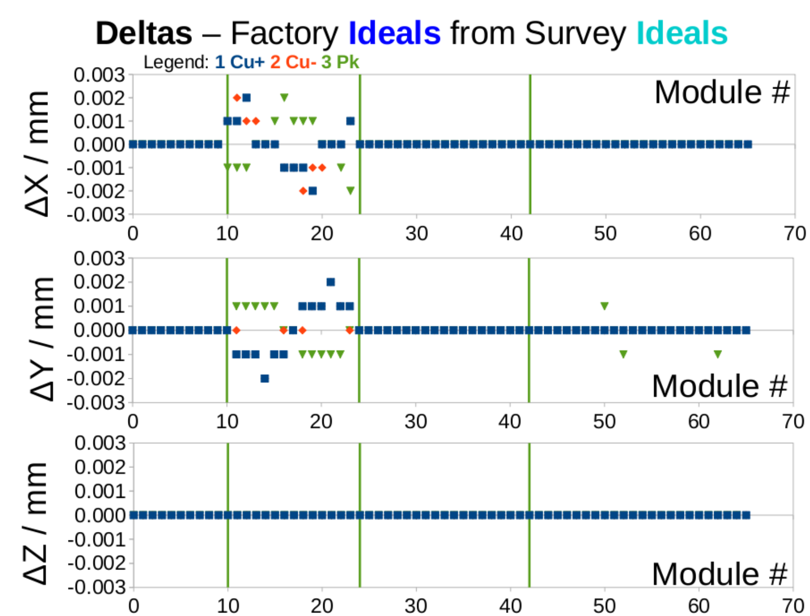




Fit residual and resolution improve

Validations

- Ideal Geometry Validation and Testing.
  - Corrected differences between engineering drawings and ideal geometry – 100 μm down to 3 μm.
  - Developing API for reconstruction – completed one for gemc.
  - **Platt (Surrey masters), Johnston (ANL postdoc).**



Jefferson Lab

DC Software Development

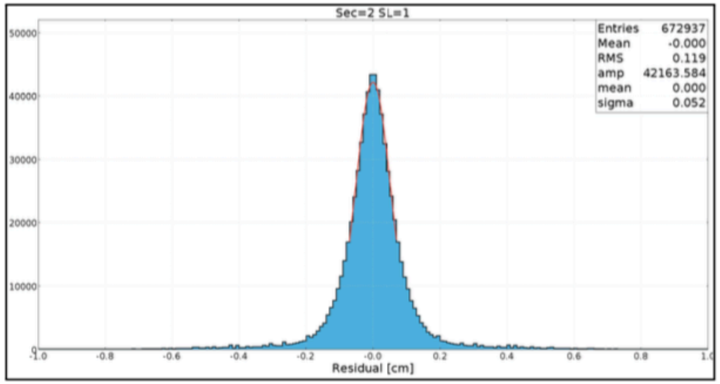
- **Monitoring** (Dilini [ODU student])
  - Standard plots, exploratory package (ntuple)
- **Calibration** (Latiful Kabir)
  - Fit time as function of (doca, beta, B, local angle)
    - Write calibration constants to CCDB
    - **Same function** used for **reconstruction** and **simulation**
- **Simulation**
  - Distance to time (Daniel Lersch)
    - non-linear function, time walk correction, random walk smearing
  - Efficiency (Daniel Lersch, Michael Kunkel)
    - intrinsic inefficiency, background inefficiency, malfunction-related inefficiency
- **Corrections**
  - Time-of-flight, signal propagation, alignment, wire sag, endplate bowing (ongoing)
- **Torus Mapping** (Joseph Newton) (ongoing)
  - Compare sector to sector, measurement to model
  - Fit to individual misplacement, distortion

DC Reconstruction Readiness
Time to Distance Function

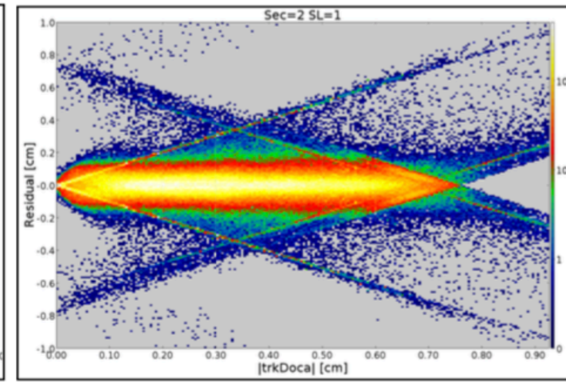
KPP Data Calibration for Sector = 2 Superlayer = 1

Latiful Kabir,
U.Miss

Before
$\sigma = 0.052$ cm

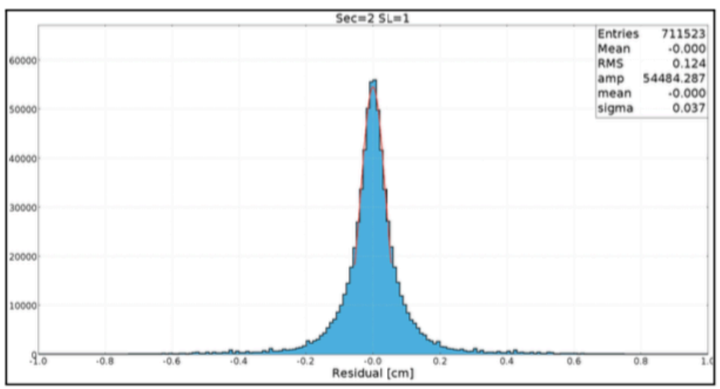


(a) Residual distribution

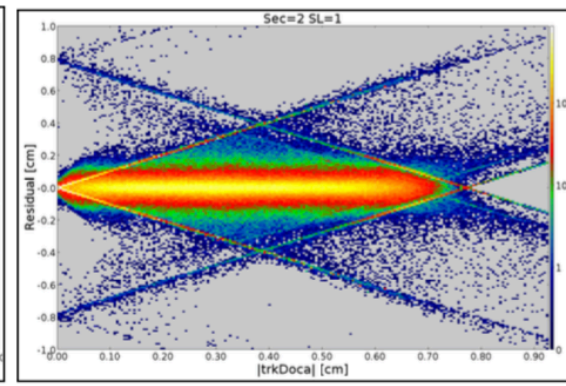


(b) Residual vs trkDoca

After
$\sigma = 0.037$ cm



(a) Residual distribution



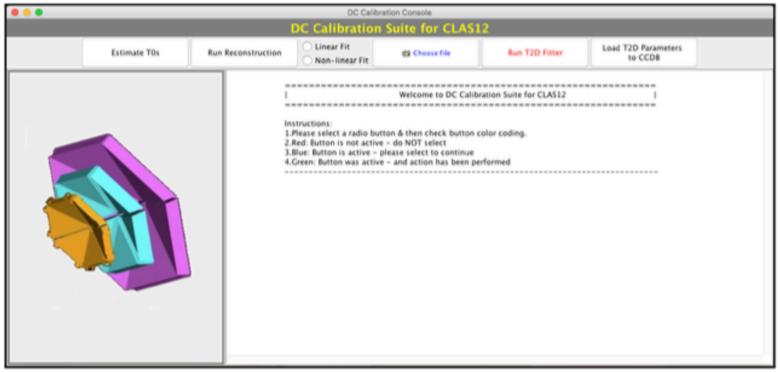
(b) Residual vs trkDoca

DC Reconstruction Readiness
Time to Distance Function

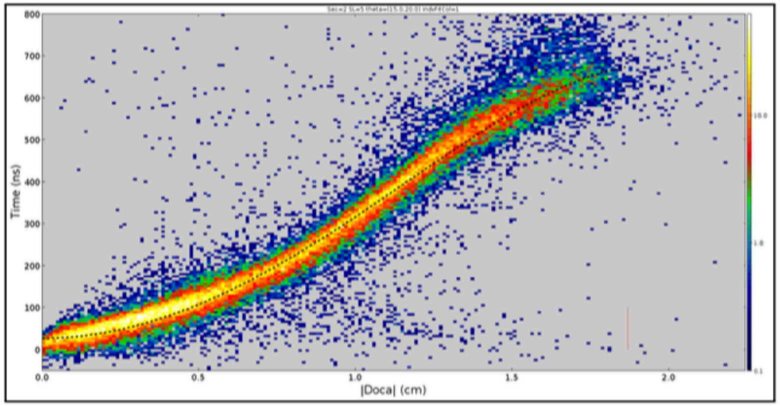
Krishna Adhikari, Latiful Kabir

DC Calibration Suite

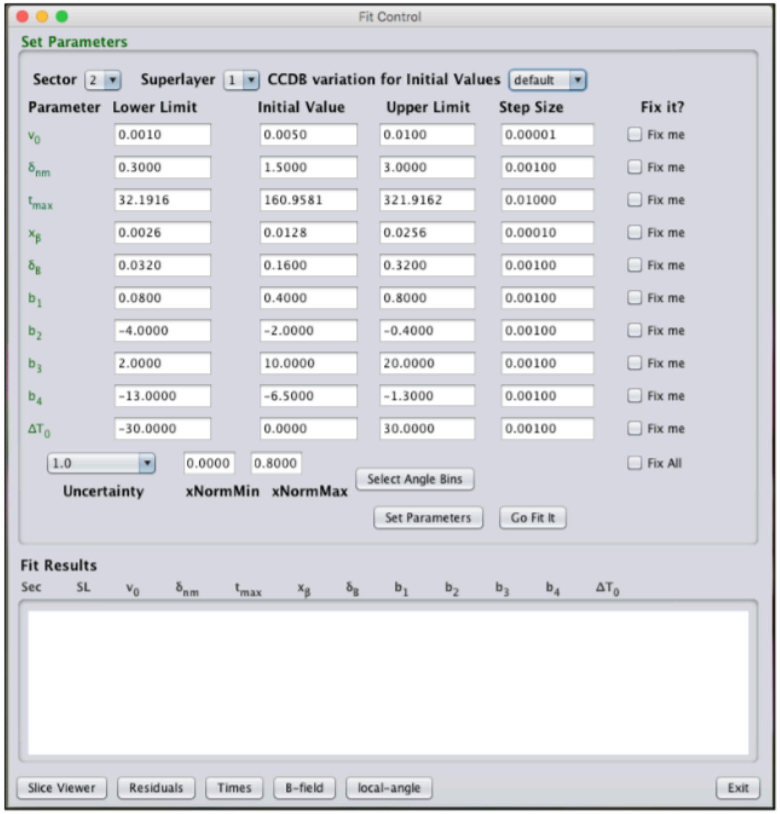
- GUI driven complete calibration suite for DC
- Tested with KPP data for convergence, stability and improved resolution



(DC calibration main GUI)



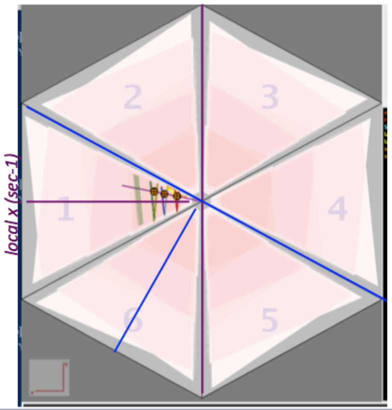
(A typical fit for time to distance function from SL = 5)



(DC calibration fit control panel)

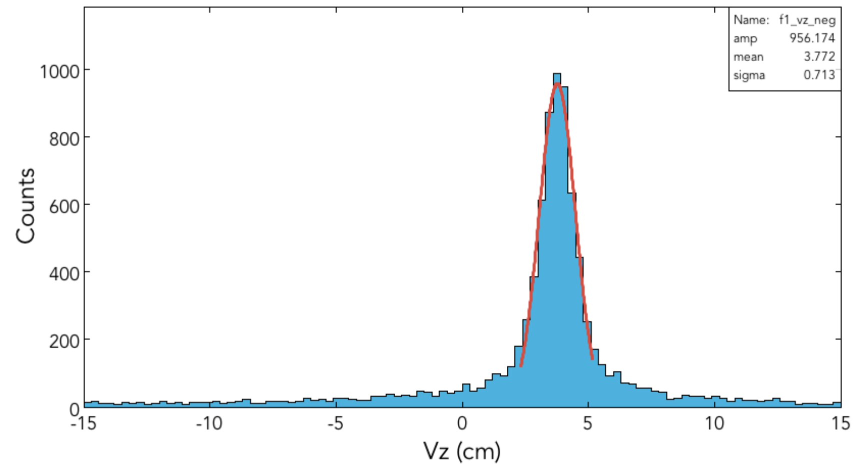
DC Reconstruction Updates

- New swim to plane algorithm in place and validated on MC
- DC service reads beta from TOF banks → requires second pass DCTB after FTOF
- Require minimum of 5 superlayers to fit a track, report superlayer inefficiencies
- Ministagger in GEMC

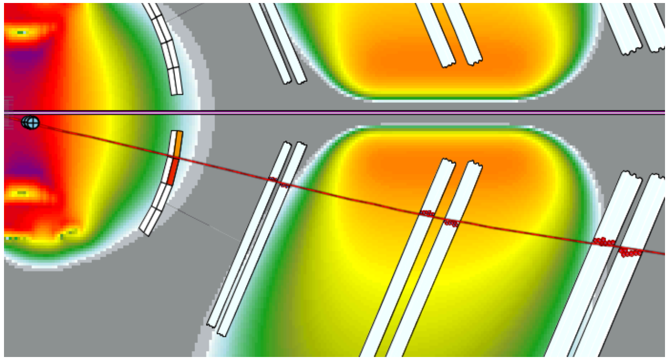


-- Swim to plane perpendicular to the sector local x axis
-- Save track parameters at that plane
-- Plane stopper will subsequently be useful for the trajectory bank (FVT package)

Reconstructed Vertex from Pass 6 KPP data cooking



FMT Reconstruction



- New package FVT in development
  - FMT hits, cluster, crosses reconstruction
  - matching to DC
  - determination of *hits on track*
  - Kalman Fit using DC track + FMT *hits on track*
  - Reads ccdb to get material budget
  - Creates Trajectory bank using the track refit using FMT hits
    - id, detector identifier, x,y,z, ux,uy,uz, pathlength to surface
- In debugging phase…

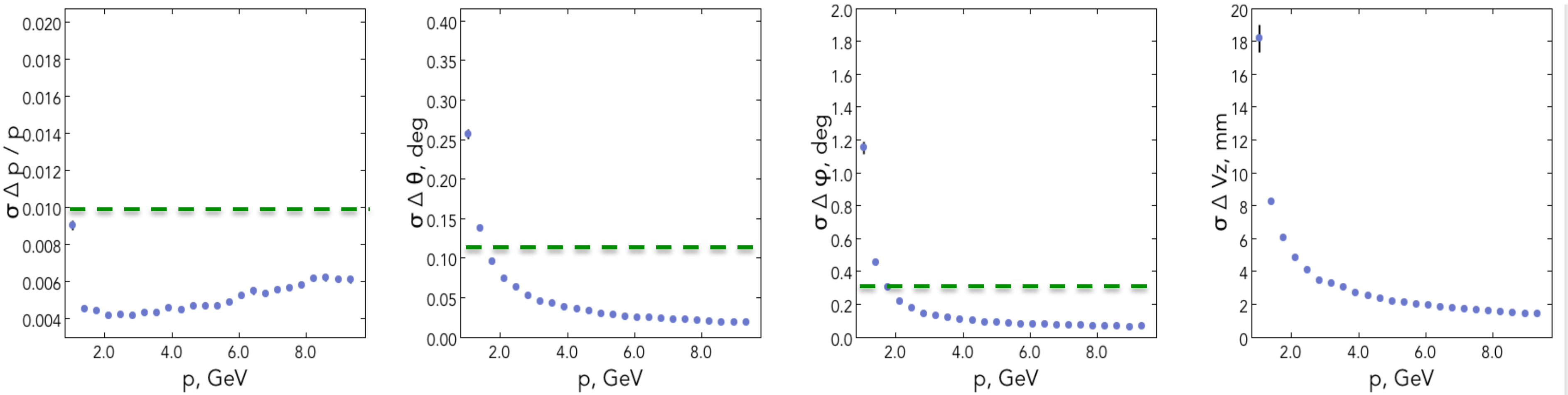
Reconstruction Readiness (Forward Detector)

Test of the reconstruction :

2. Validations on MC (no FMT):
   - Use calibration challenge data sample and kinematic-specific samples.
   - Verify reconstruction resolutions and efficiencies.

Tracking efficiency ~ 97%

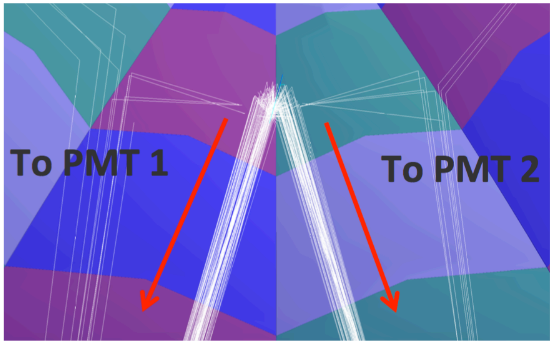
Forward tracking average resolutions well within specs for DC only



Jefferson Lab

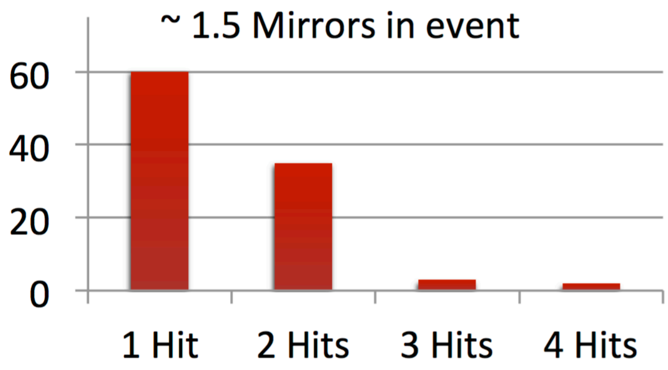
HTCC Reconstruction

N. Markov [U. Conn]

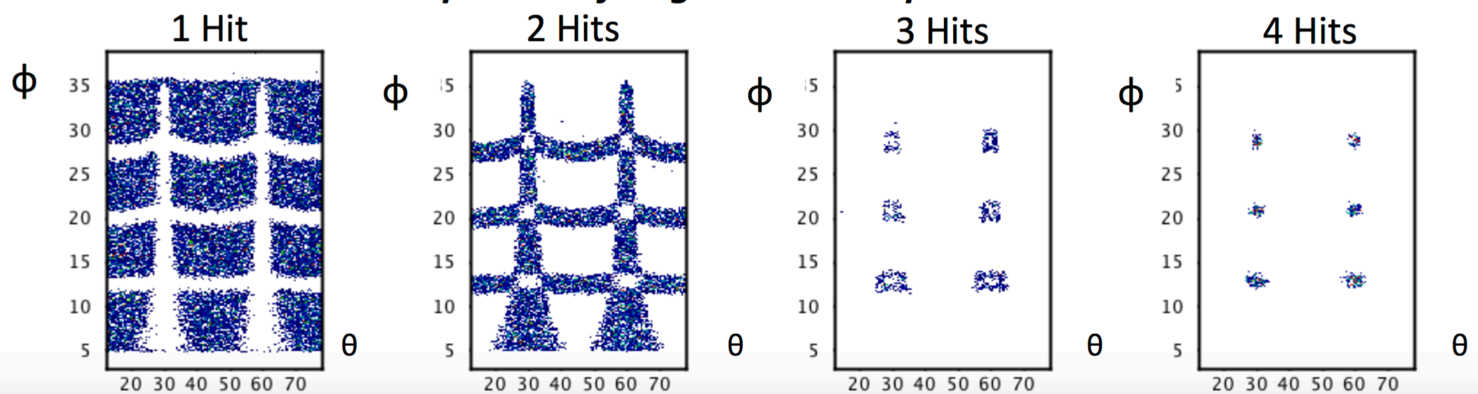


Cerenkov radiation from single electron may split between mirrors and is collected by different PMTs

- Cluster reconstruction
  Events with 1, 2, 3, or 4 hits



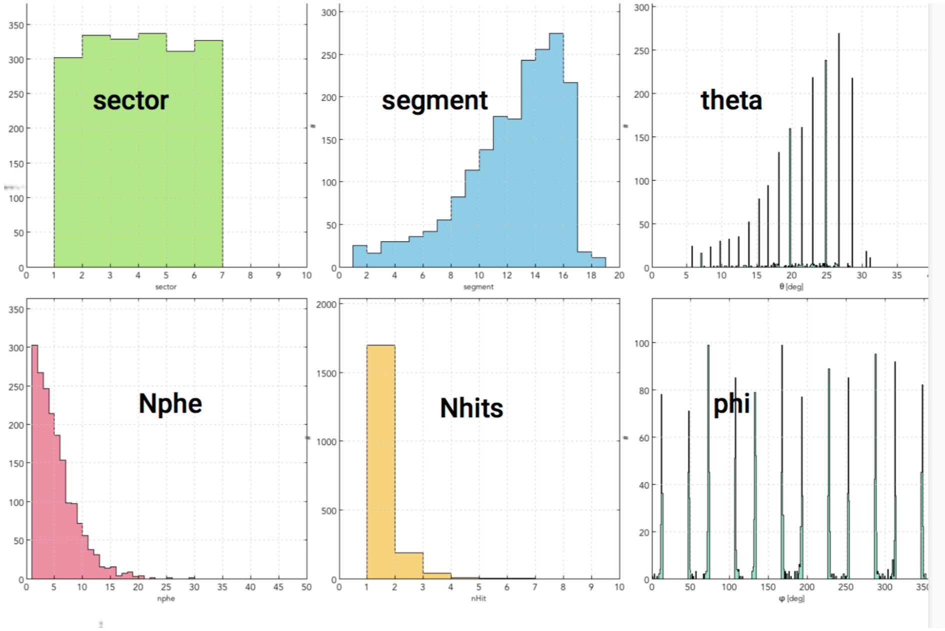
Geometrical pattern of single- and multiple hit events:



LTCC Reconstruction

- LTCC clustering algorithm in place
- Reads calibration data
- Used in service chain

Temple University



GEMC simulations

TOF reconstruction

TOF reconstruction code determines:

- hit times (t_L, t_R, <t>)

$$t_{L,R} = (\mathcal{C}_{TDC} \cdot TDC_{L,R}) - t^{walk}_{L,R} \pm \frac{C_{L,R}}{2} + \mathcal{C}_{p2p}$$

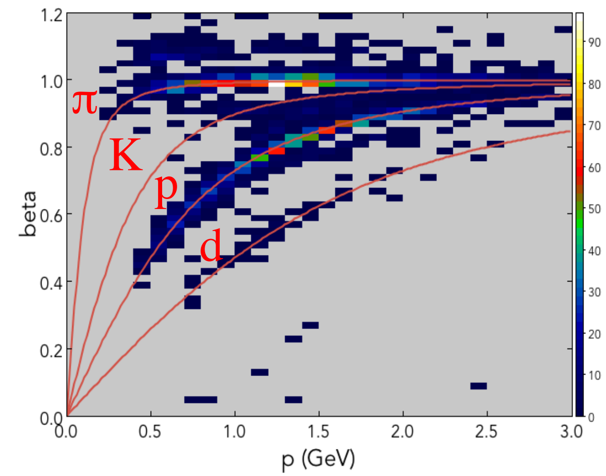
- hit coordinates (x) $x = \frac{v_{eff}}{2}(t_L - t_R)$

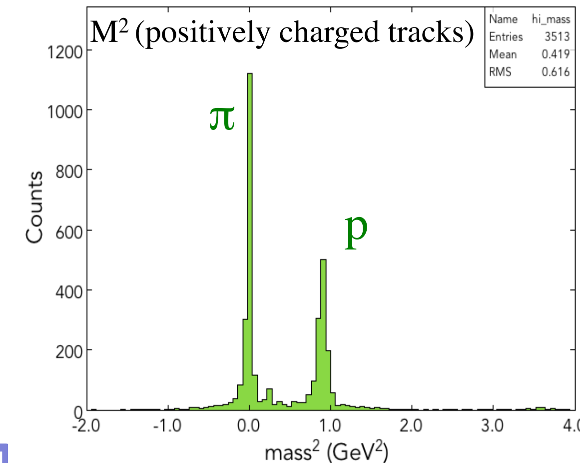
- deposited energies (E_L, E_R, <E_{dep}>)

$$E_{L,R} = (ADC_{L,R} - PED_{L,R}) \left[\frac{(\frac{dE}{dx})_{MIP} \cdot t}{ADC_{MIP}} \right]$$

- associated time, coordinate, and energy uncertainties
- performs hit clustering and matching
- combines hit times from panel-1a and panel-1b

PID from FTOF (KPP data):
Beta vs. p (positively charged tracks)



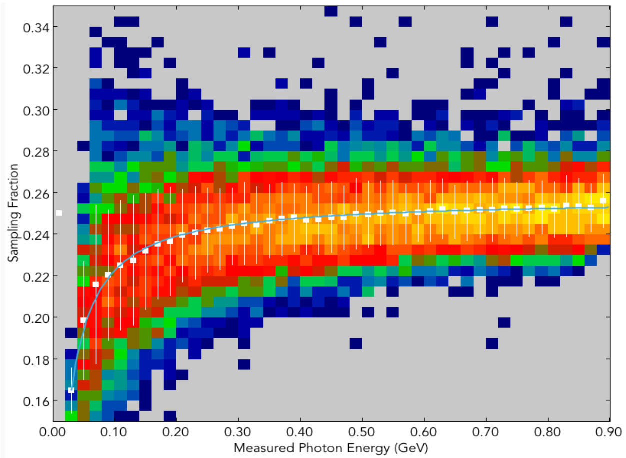


- Code designed to function for all “allowable” hardware conditions

ECAL Reconstruction

- Simulated 2 GeV π^0 events

C. Smith (UVA)

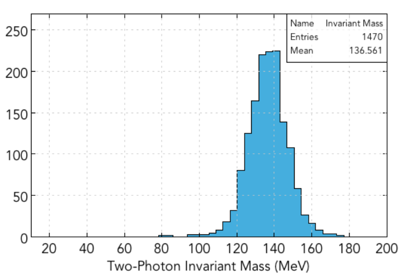


$$S.F. = 0.257 \times (1.0 - 0.0146 / E + 0.000117 / E^2)$$

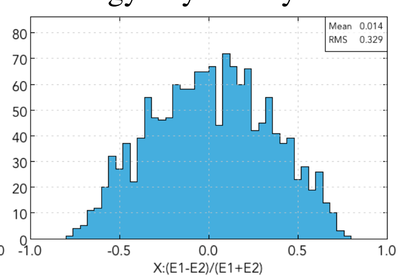
Offline/online energy cluster reconstruction

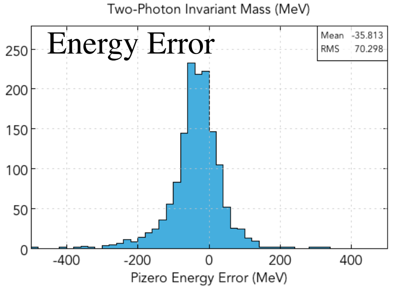
- strips, peaks, clusters reconstruction
- peak splitting and energy sharing for π^0 reconstruction

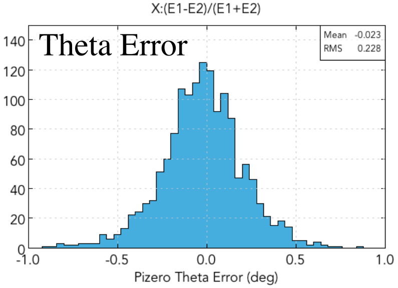
Invariant Mass

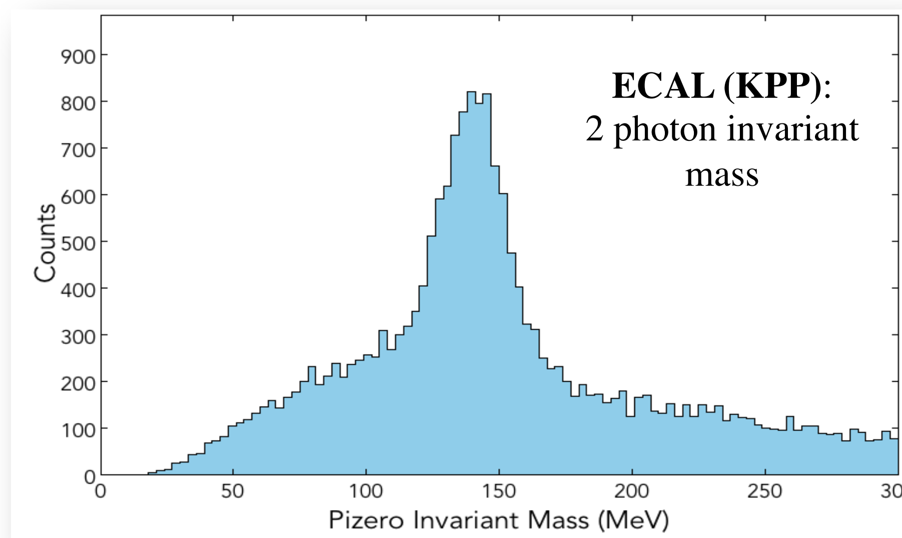


Energy Asymmetry



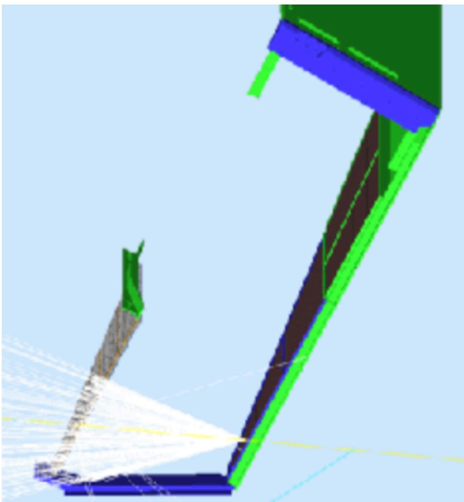






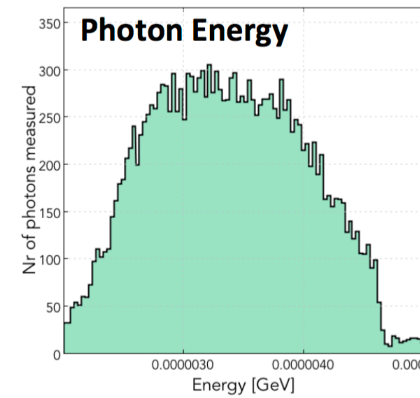
RICH Simulation and Reconstruction

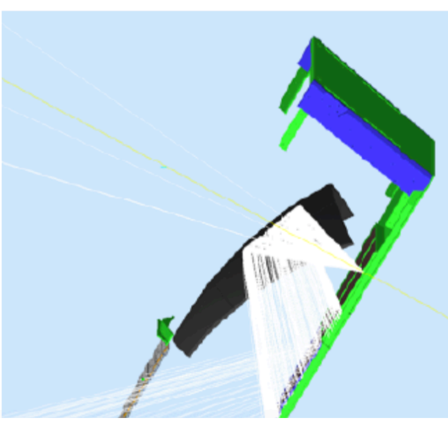
M. Contalbrigo
INFN, Ferrera



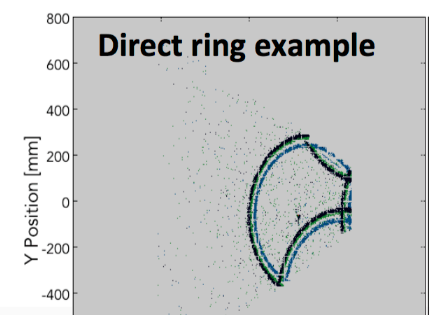
- a detailed, consistent and updated description of the detector can be obtained
- simulation and reconstruction shared the same database

Digitization of the MAPMT response:
- calculate the pixel ID
- interface to CCDB
- apply efficiency
- simulated ADC and TDC spectra





Event reconstruction started
- match with DC information in coat-java
- photon tracing algorithm
  (tested with prototype and cosmic runs)
- event display

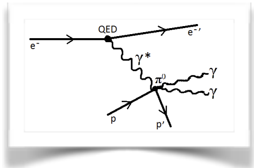


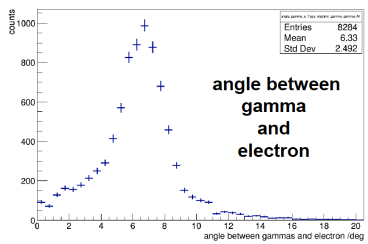
Strong crew: Matteo, Ilaria, Marco, Giovanni,

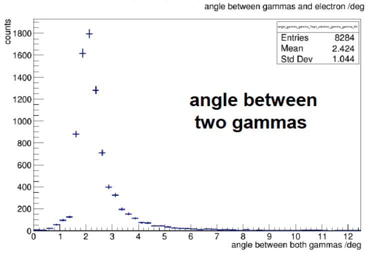
FT Reconstruction Status

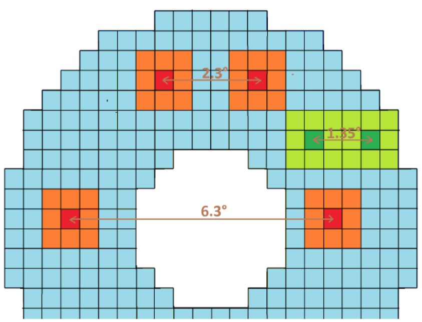
e p → e' p π0 (γ p → p π^0)

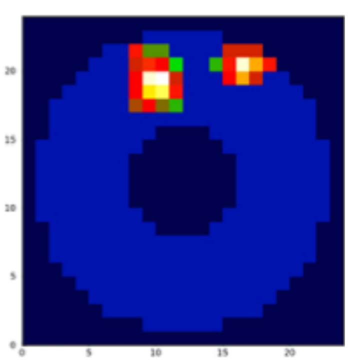
- **S.Diehl (U Giessen)**
- Full CLAS12 (+FT) GEANT4 sim/rec
- JPAC e-production amplitudes (V.Mathieu)
- AMPTOOLS

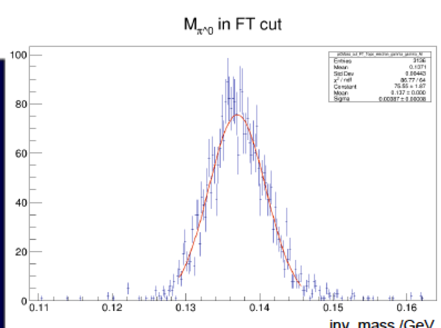












137 MeV/c² reconstrcuted (lit.: 134,98 MeV/c²)

Resolution: 3.9 MeV/c²

M. Battaglieri
INFN, Genoa

FT-Cal:
- Read raw nits from hipo bank
- Read calibration constants from DB
- Create hits, converting from digitized info to E and T
- Reconstruct cluster and determine cluster E, T and pos

FT-Hodo:
- Read raw hit from evio bank
- Read calibration constants from DB
- Create hits, converting from digit info to E &T
- Match hits in the hodoscope layers

FT-Track:
- started based on algorithm developed by G. Charles

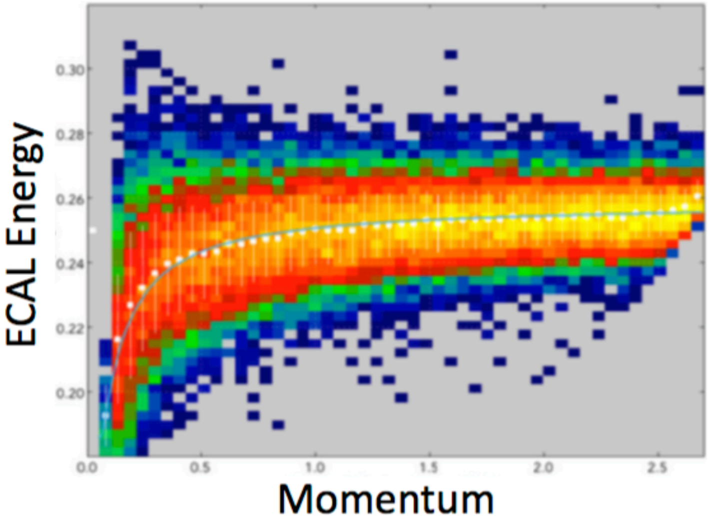
FT-Match:
- Match reconstructed clusters with hits in hodoscope
- Output of final reconstructed particles

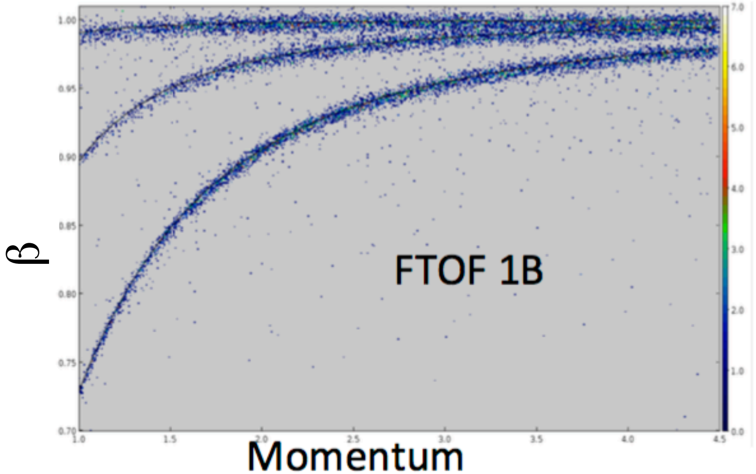
Event Builder

Event Builder framework status

N. Baltzell

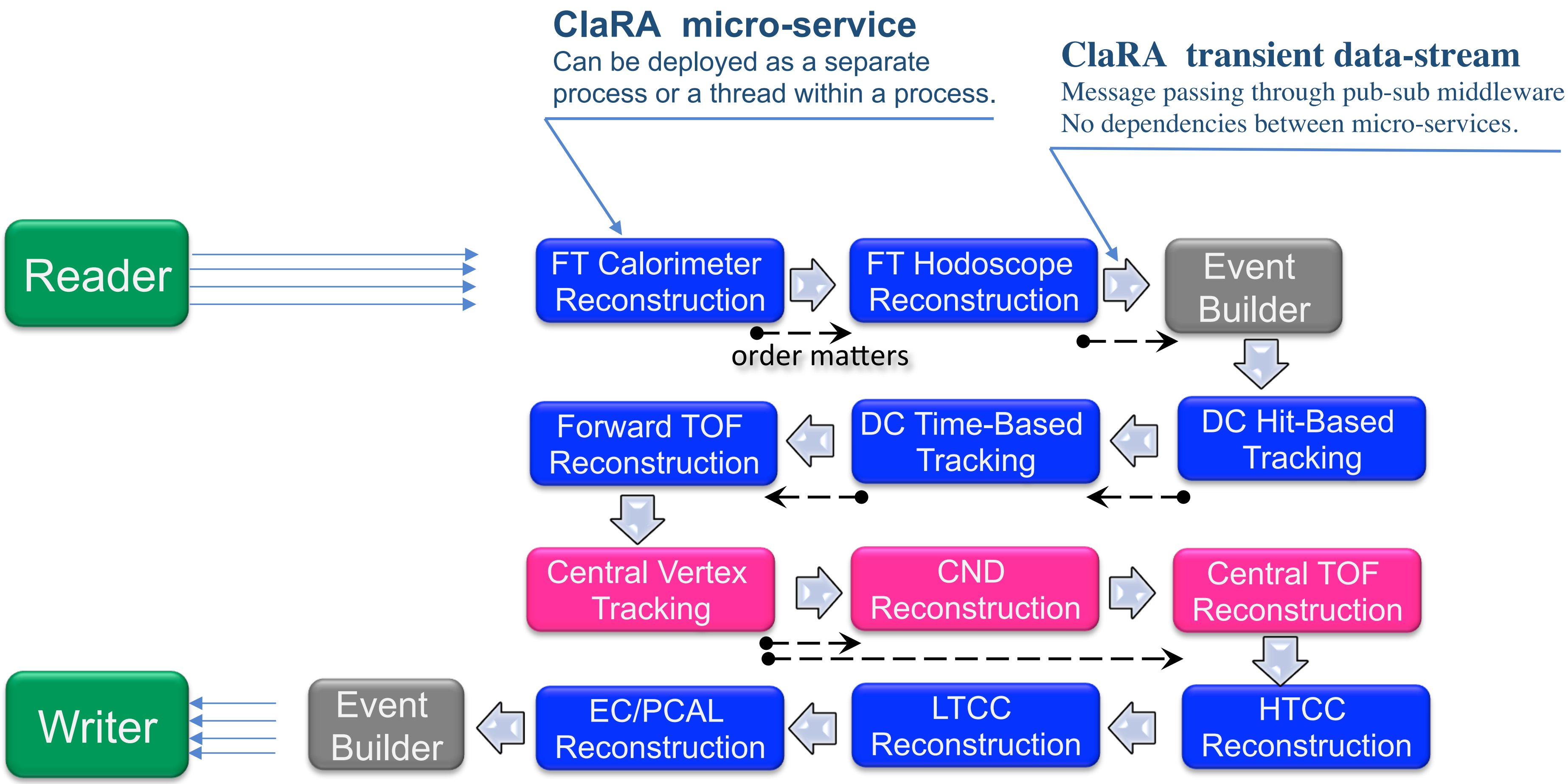
- Trigger Particle / Start Time
  - currently electron-focused, to be extended to π/γ
- Particle Identification
  - currently simple cut-based algorithm with single PID
  - FD algorithms developed and well-tested
  - Forward Tagger incorporated, Central Detector partial
- Validation/testing "suite" in use
  - checking yields, efficiencies/misidentification, data sanity checks
- CCDB structure and codes for Event Builder exists, adoption in progress
- Output bank structure defined and mostly populated
- Devs: Joseph Newton (ODU), Nathan, Raffaella
  - Framework design by Gagik Gavalian





Event Reconstruction Service Composition

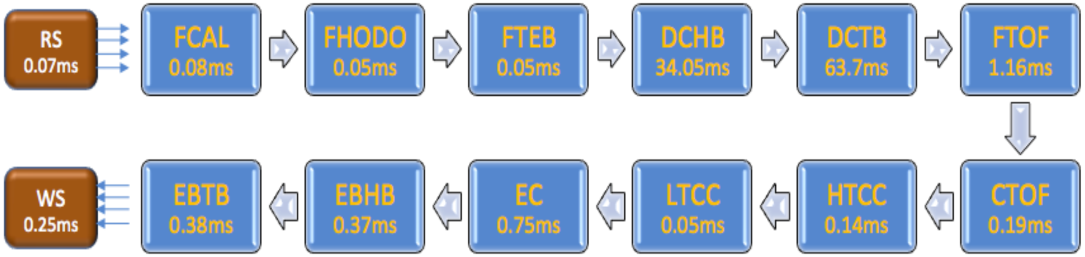
- Event building services (EB) combine info from individual services output banks to reconstruct particle candidate.

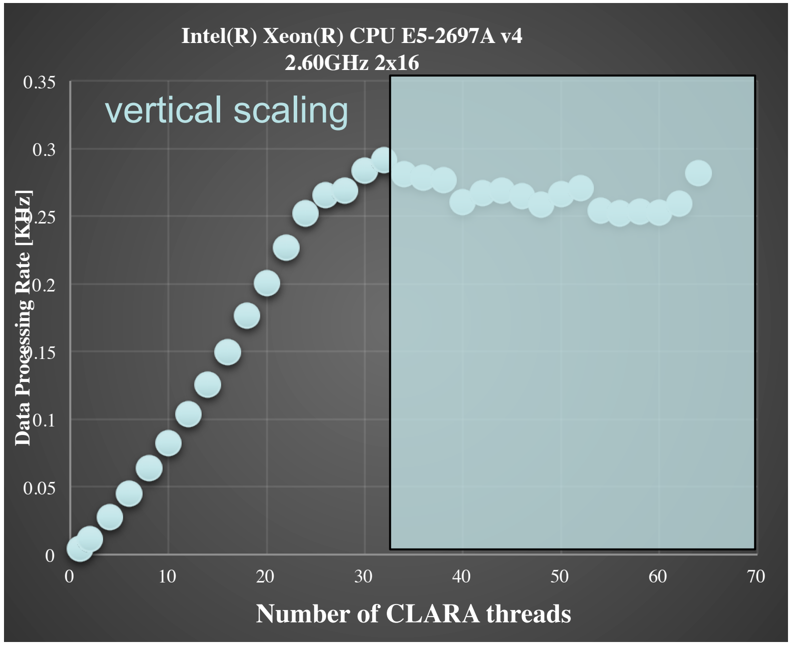


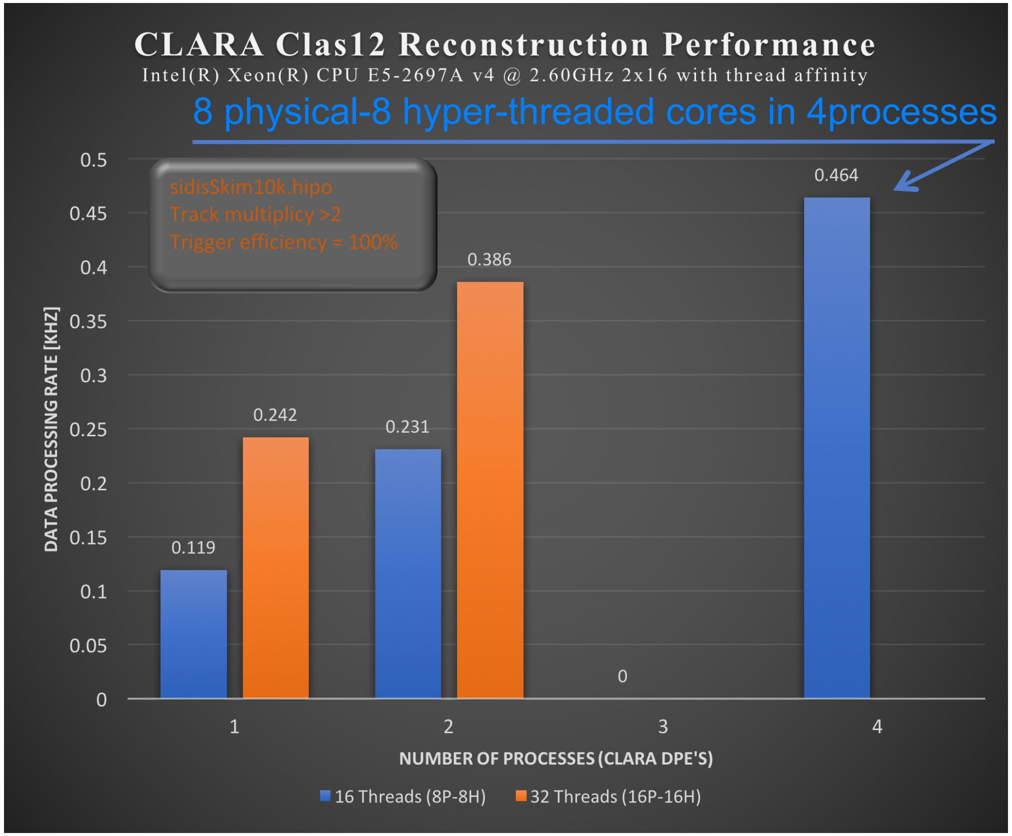
Reconstruction Readiness

Reconstruction framework stable.
Framework performance studies:

- Scaling studies using MC data*
  - Vertical scaling (multi-threading within the same node) ✔
  - Horizontal scaling (across nodes) ✔
- Ongoing optimization (reco. rates)



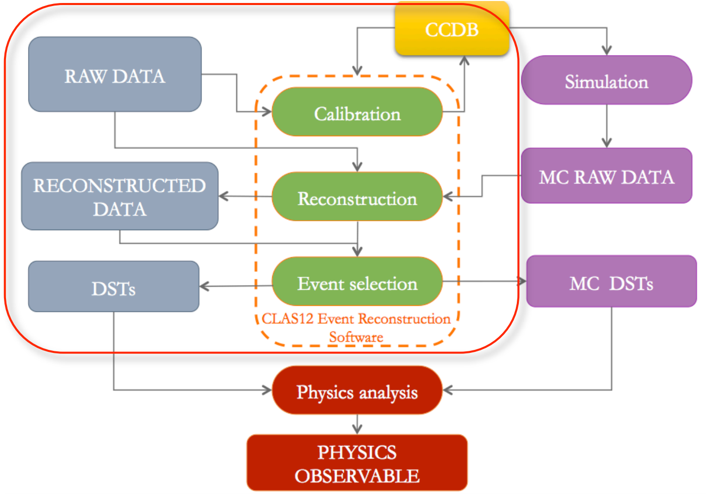




* Trigger efficiency = 100%
Sidis events, Track multiplicity >=2, No background
Node = Intel(R) Xeon(R) CPU E5-2697A v4 @ 2.60GHz 2x16

Overall Offline Reconstruction Readiness

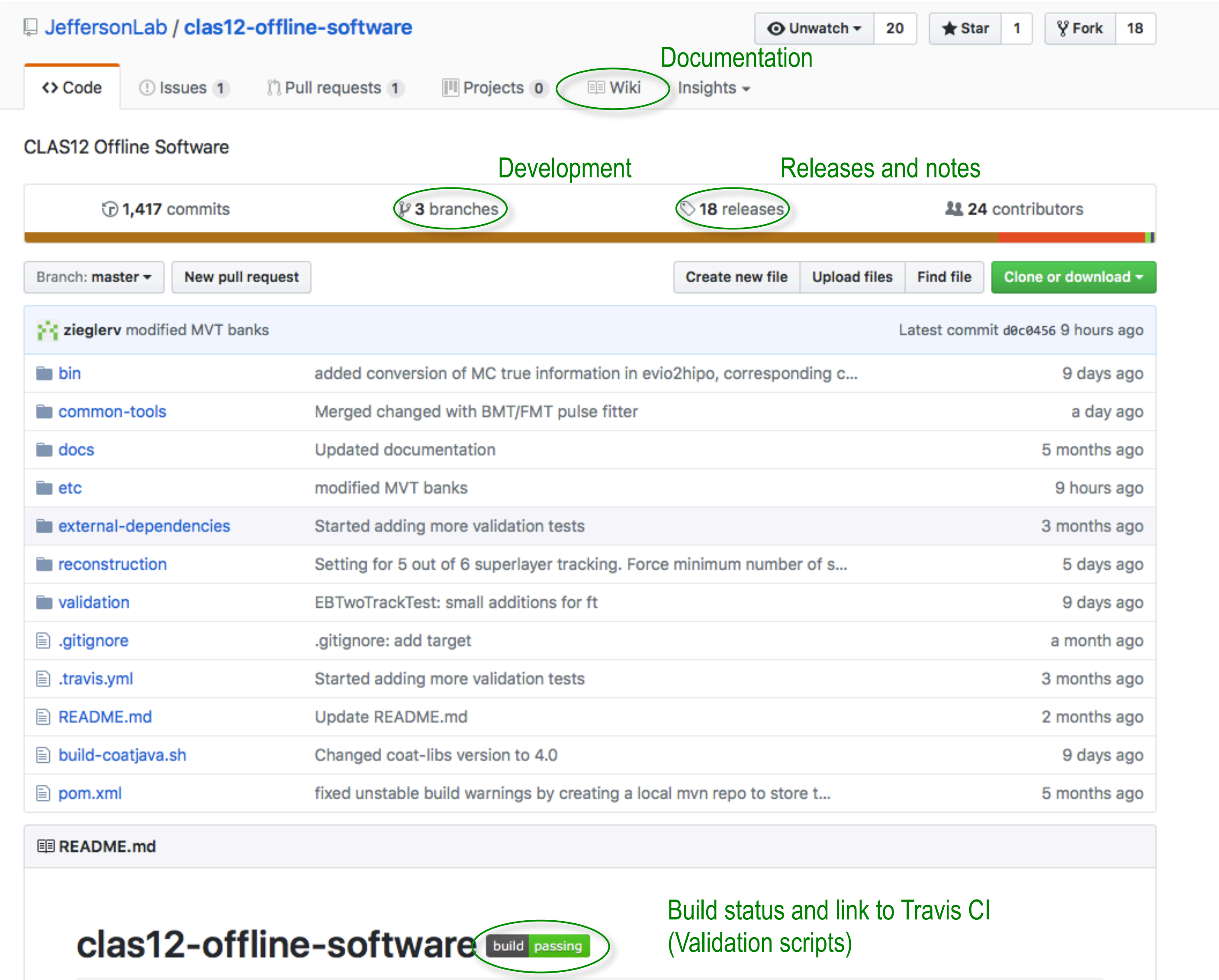
- Data *Cooking*
  - Subsystem calibrations
    - Suites extensively tested on MC & data (Feb. 2017 run: KPP)
  - Calibration quality monitoring tools
    - Validated and successfully used during KPP run and calibration challenges
  - Momentum and angle corrections
    - In development
  - Data output file format and structures
    - Defined and in use for analyses
  - Particle identification schemes
    - Algorithms developed (cut-based, ANNs-based in development)
  - Geometry offsets and field distortions
    - Alignment code using Millepede package in development for SVT
    - Distortion due to DC end-plate bowing and wire sag in development (started)
    - Alignment algorithm for DC just started
    - Field mapping done (inclusion of distortions from mapping in MagField package to be done)
  - Reconstruction code monitoring
    - Suite in use for validation studies
  - Calibration & Geometry Constants
    - Tables defined, code to access constants ready
  - Raw data decoding ready for all subsystems



Software Management

Code management

- github
- release tagging
- code validation (validation suites, Travis tests)



Summary

- Framework
  - Stable, linear vertical and horizontal scaling, can run multiple processes in one node to optimize reconstruction rates
  - Ran successfully with no errors during KPP data cooking and calibration challenges
  - High compressibility data format suitable for DST and data distribution
- Event Reconstruction
  - Full chain of services ready to reconstruct data
  - Raw data decoding ready for all detectors
  - Stable release available for physics reactions studies (RfS)
  - Code management in place

BACK-UPS

CLAS12 Requirements

DAQ data **number** of events per day of running

= event rate in kHz x 1000 x 60 x 60 x 24 x 24-hr duty cycle in % / 100

= 10 x 1000 x 60 x 60 x 24 x 60%/100 **= 518.4×10^6**

DAQ data volume in **TB** per day of running

= DAQ data number of events per day of running x event size in kB x 10^{-9}

= 518.4×10^6 x 80 x 10^{-9} **= 41.5**

Data reconstruction **CPU time in sec** per day of running

= DAQ nb of evts per day of running x rec-processing time per evt per core in ms/1000 x nb rec-passes

= 518.4×10^6 x 120/1000 x 2 **= 124.4×10^6**

Simulated **number** of event per day of running

= DAQ data number of events per day of running x analyzed fraction in % / 100

x e- trigger fraction in % / 100 x ratio of simulated to data events

= 518.4×10^6 x 50% / 100 x 50 % / 100 x 6 = **777.6×10^6**

Simulated evio data volume in **TB** per day of running

= Simulation number of events per day of running x MC event size in kB x 10^{-9}

= 777.6×10^6 x 20 x 10^{-9} **= 15.55**

Simulation **CPU time in sec** per day of running

= Simulation nb of evts per day x simulation time per evt per core in ms/1000 x nb MC-passes

= 777.6×10^6 x 500/1000 x 1.5 **= 583.2×10^6**

MC reconstruction **CPU time in sec** per day of running

= Simulated nb of evts per day x rec-processing time per evt per core in ms/1000 x nb MCrec-passes

= 777.6×10^6 x 120/1000 x 1.5 **= 140.0×10^6**

Data calibration **CPU time in sec** per day of running

= DAQ nb of evts per day x fraction of data for calib in % / 100 x cal-processing time per evt per core in ms/1000 x nb cal-passes

= 518.4×10^6 x 5 / 100 x 50/1000 x 5 **= 6.5×10^6**

CLAS12 Core Requirements

Data reconstruction **CPU time in sec** per day of running = 518.4×10^6 x 120/1000 x 2 **= 124.4×10^6**

Simulation **CPU time in sec** per day of running = 777.6×10^6 x 500/1000 x 1.5 **= 583.2×10^6**

MC reconstruction **CPU time in sec** per day of running = 777.6×10^6 x 120/1000 x 1.5 **= 140.0×10^6**

Data calibration **CPU time in sec** per day of running= 518.4×10^6 x 5 / 100 x 50/1000 x 5 **= 6.5×10^6**

Number of Cores per day of running

= **CPU time in sec** per day of running / (24x60x60 x core efficiency in % /100)

= **CPU time in sec** per day of running / (24x60x60 x 90% /100)

➔ **Data reconstruction : 124.4×10^6** / (24x60x60 x 90% /100) **= 1600**

➔ **Simulations : 583.2×10^6** / (24x60x60 x 90% /100) **= 7500**

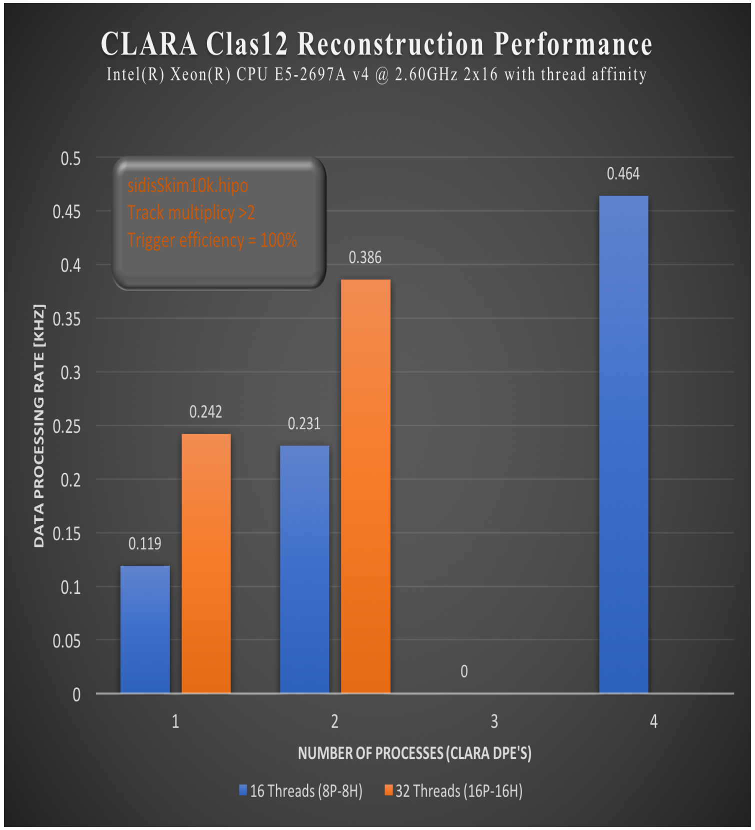
➔ **MC Reconstruction : 140.0×10^6** / (24x60x60 x 90% /100) **= 1800**

➔ **Data Calibration : 6.5×10^6** / (24x60x60 x 90% /100) **= 83**

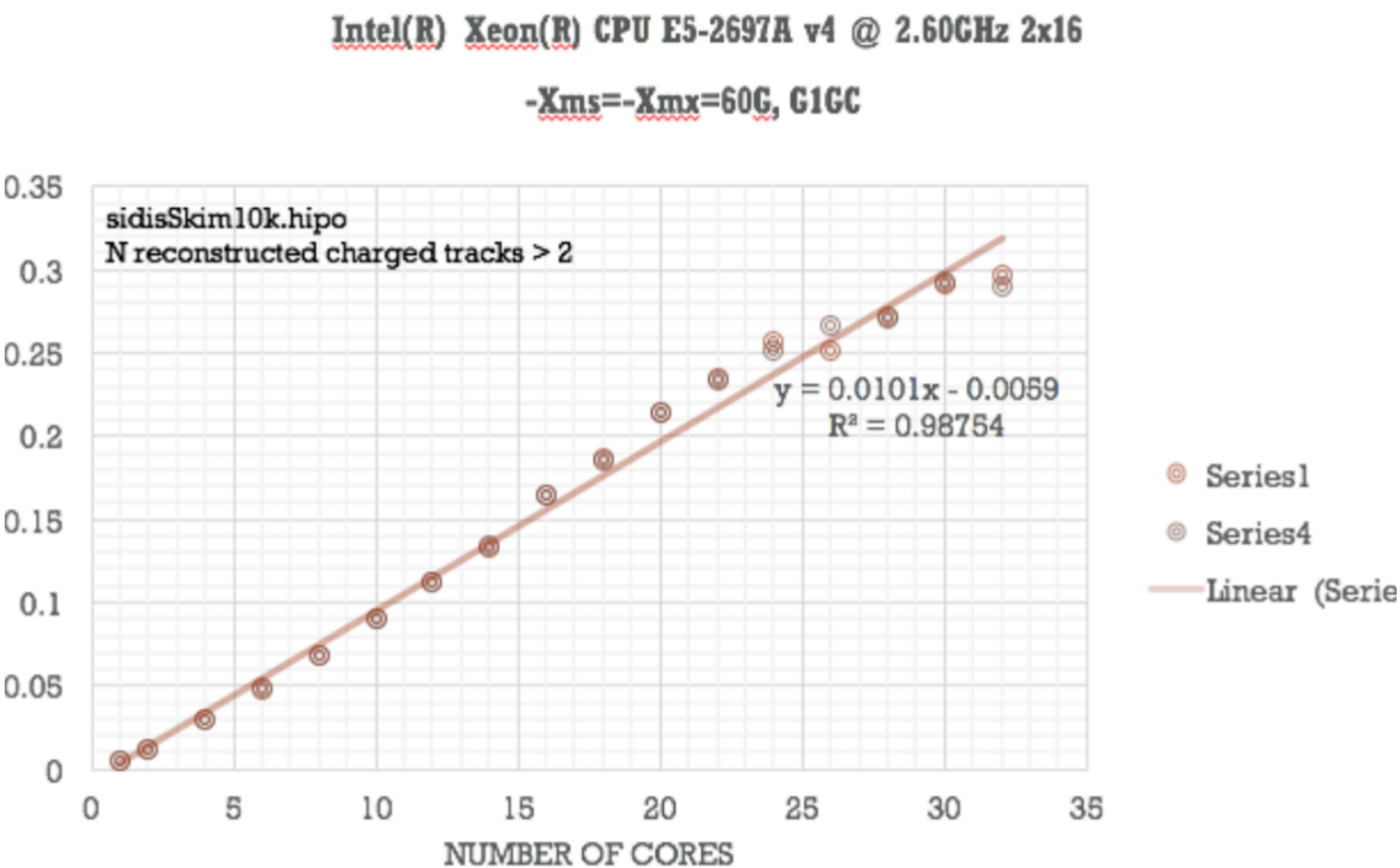
Number of Cores Days for Data Cooking	1600
Number of Cores Days for Simulation	7500
Number of Cores Days for MC reco	1800
Number of Cores Days for Calibration	83
Total for 1 day	**10983**

NUMA Studies (Intel® Xeon® CPU E5-2697A V4 2.60GHz 2x16)

Measurement	1 Process		2 Processes		4 processes	
	8P_8H ms	16P_16H ms	8P_8H ms (N proc.)	16P_16H ms (N proc.)	8P_8H ms (N proc)	
1	8.50	4.10	8.13(1)	5.18(1)	8.51(1)	9.34(3)
			9.30(2)	5.22(2)	7.99(2)	9.20(4)
2	8.32	4.13	8.19(1)	5.16(1)	8.48(1)	8.90(3)
			9.31(2)	5.22(2)	8.05(2)	9.19(4)
3	8.33	4.16	8.14(1)	5.19(1)	8.55(1)	8.90(3)
			9.17(2)	5.10(2)	8.10(2)	8.42(4)
Aver. proc. 1	8.38	4.13	8.15	5.18	8.51	
Aver. proc. 2			9.26	5.18	8.06	
Aver. proc. 3					9.05	
Aver. proc. 4					8.88	
Rate in KHz	0.119	0.242	0.123+0.108=0.231	0.193+0.193=0.386	0.117+0.124+0.110+0.113=0.464	



CLARA Horizontal Scaling

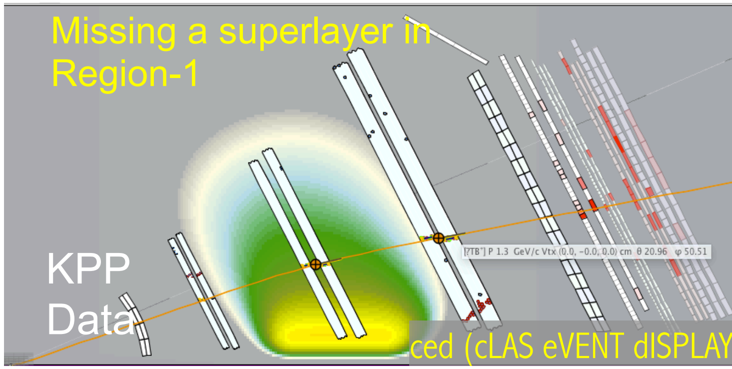


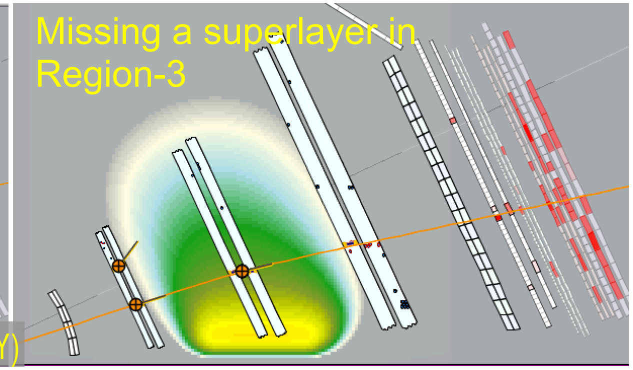
Jefferson Lab

Reconstruction Readiness

Tracking in background & with inefficient chambers

- 5-Out-Of-6 superlayers tracking





MC Studies

Losing a superlayer has a minimal effect on tracking resolutions

Inefficiencies due to missing SL: 5% for SL1, 10% for SL2, less than 3% for all other SLs

- Noise rejection algorithms validation on KPP data