Poster ID #172

Online Event Classification in JUNO

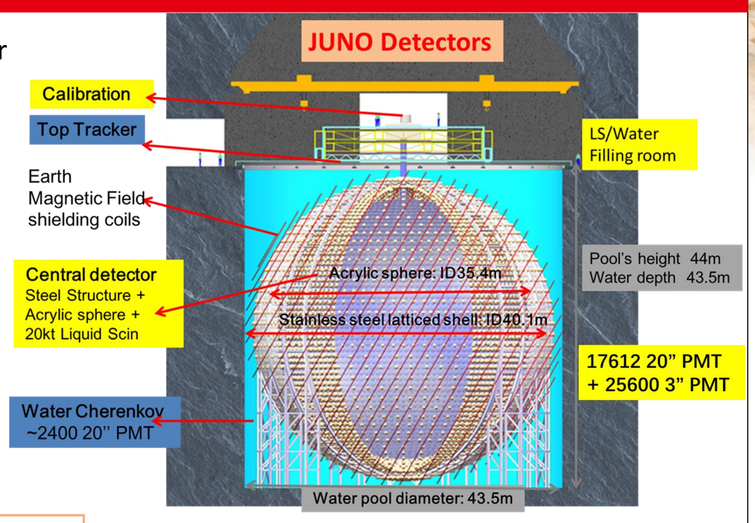
Yongpeng Zhang (ypzhang1991@ihep.ac.cn)
on behalf of the JUNO collaboration
Institute of high energy physics, Beijing, China

I. Motivation

Jiangmen Underground Neutrino Observatory (JUNO), which is under construction at a depth of 700 meters underground in China, is the largest liquid scintillator experiment with many important topics in neutrino and astro-particle physics.

A multi-purpose observatory: *"JUNO Physics and Detector experiment" arXiv 2104.02565*

- Neutrino mass ordering;
- Precision measurement of oscillation parameters;
- Supernova/Solar/Atmospheric/Geo-neutrinos
- Exotic searching



Data volume

There are two types of data generated by the front-end electronics: waveform and TQ

Detector	Channel	Data Size (Byte)	Rate	Data Volume
CD LPMT	17612	2032	1 kHz	35.8 GB/s
CD LPMT-T/Q	17612	16	30 kHz	8.5 GB/s
CD SPMT	25600	30	500 Hz	375 MB/s
CD Calibration	17612	2032	200 Hz	7.2GB/s
WP LPMT	2400	2032	205 Hz	984 MB/s
WP Calibration	2400	2032	200 Hz	960 MB/s
TT				1 MB/s

Baseline requirement

Considering the limitations of transmission bandwidth and the resources for offline data processing, JUNO needs to convert the raw data stream with full waveform (@1kHz) to the about 60 MB/s (with 100% resources redundancy)

The data volume needs to be compressed online by a factor of ~ 750.

(Challenge!)

II. Online Event Classification

Two steps to compress the data volume:

1) Based on the physics requirement and the fast reconstruction/selection decide save waveform or not

2) Compress the filtered data stream with fast lossless compression method

The first step will be carried out by the Online Event Classification (OEC) software running on the DAQ.

Low-Level Event Classification (LEC)

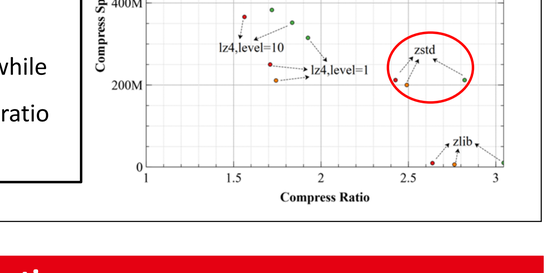
Single Event Classification at distributed nodes

- Input: Waveforms
- Processing
  - FAST waveforms recon.
  - FAST energy, vertex, track recon
  - Single event classification
- Output: Light-weighted reconstructed data and initial Tags

High-Level Event Classification (HEC)

Correlated Event Classification

- Input: Light-weighted reconstructed data objects and initial Tags
- Processing: Correlated event classification using offline algorithms
- Output:
  - Final Tags



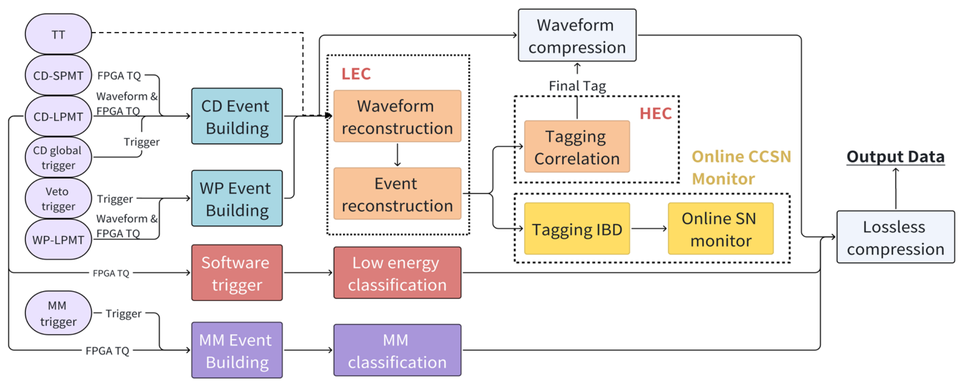
The second step will be carried out by the online lossless data compression technology

- ✓ Different online lossless data compression technology has been studied
- ✓ zstd uses fewer resources while having a high compression ratio (~2.5)

III. Processing of different types of data in DAQ

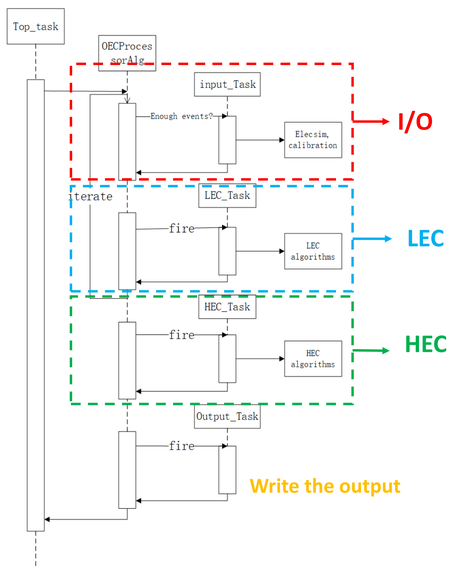
Event Stream: (online processing)

- ➢ Waveform event (hardware global trigger)
  - ✓ CD/WP event assemble based on trigger timestamp
  - ✓ Assemble the waveforms and TQ that belong to the same readout window (1 us).
- ➢ Trigger-less T/Q (software trigger)
  - ✓ T/Q process flow for low energy event
- ➢ Multi-Message data stream (hardware trigger)

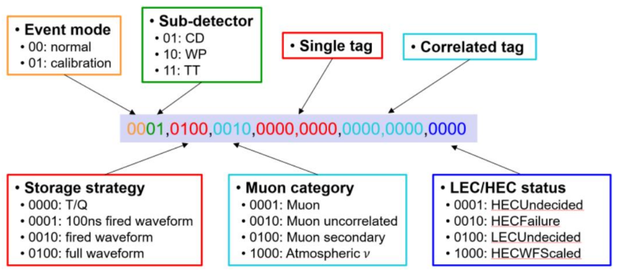


The online Supernova (SN) monitor will use the IBD events tagged by OEC to give pre-SN or SN alerts

IV. OEC offline software

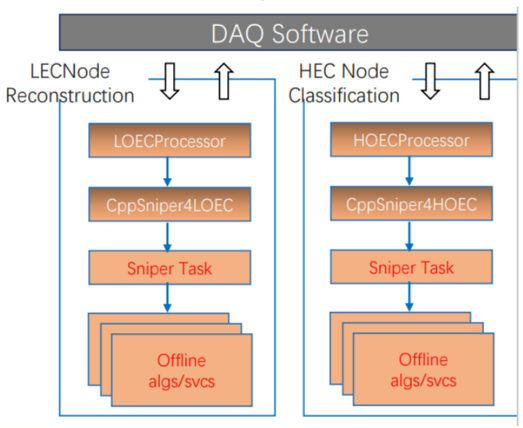


- For physicists, NO need to worry about the DAQ interfaces
- Optimize the performance of fast reconstruction and classification algorithms in offline independently
- Define the events tags (32-bits)
- Decide the waveform-saving strategy
- Test the performance using Mock Data with different physics events mixed



V. Online/Offline Integration

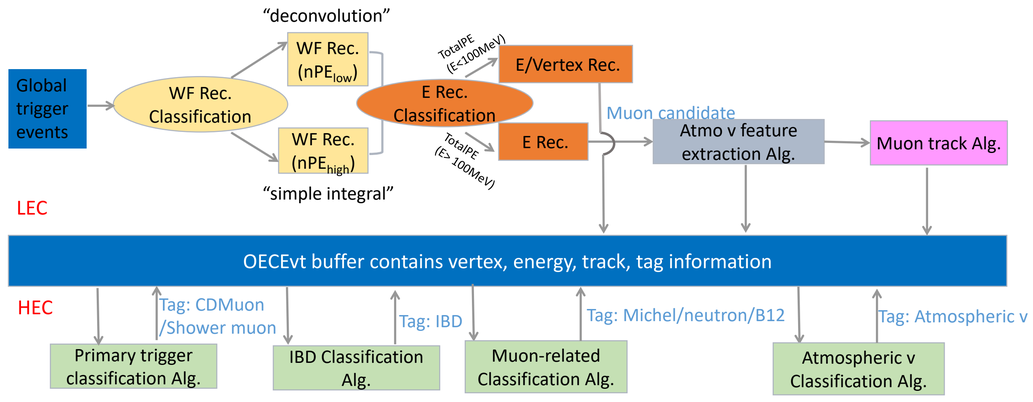
Build data connection between the DAQ and the OEC offline software



[L/H]OECProcessor : driven by data (scheduling Sniper task)

CppSniper4[L/H]OEC : configure, load and invoke offline software modules

VI. Data processing flow in OEC



Singles @LOEC: run on parallel nodes	
Type/Energy range	**Waveform of fired**
< 0.8 MeV	no
(0.8, 2) MeV	R<15 m
(2, 5) MeV	R<16.5 m
(5, 20) MeV	always
(20, 100) MeV	no
SN	As much as possible
Calibration	Non-moving source
Unable to decide	Save all WF
Option	Save WF randomly

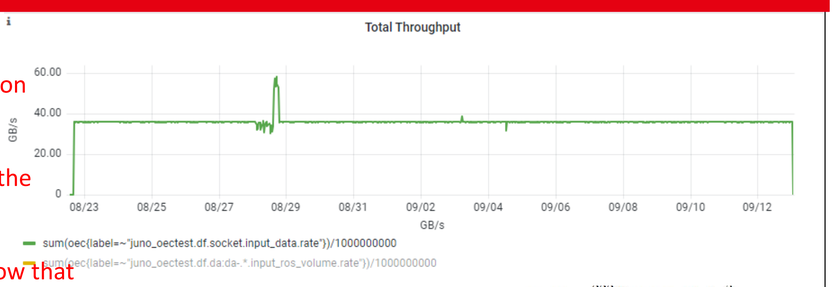
Pairs @HOEC: operate based on LOEC inputs						
Type	Ep, MeV	Ed, MeV	ΔR, m	Δt	WF prompt	WF delayed
IBD	0.5−100	1−12	<3.5	<2 ms	All PMTs	All PMTs
Michel e	>100	12−60	_	<20 us	Risied edge (fired)	Fired PMTs (Δt>7 μs)
Spallation n	>100	1−12	_	<2 ms		Fired PMTs (Δt>8 μs)
^{12}B	>100	1-20	_	2ms~1s	Risied edge (fired)	Fired PMTs (Δt>8 μs)

Tracks	Energy	**Waveform of fired**
Non-shower muon	(100 MeV, 10 GeV)	Rising edge
Shower muon	>10 GeV	Yes
Afterpulses		No
Atmospheric ν	>100 MeV	Yes

The events that do not save waveforms can randomly save waveforms based on event type and energy/vertex

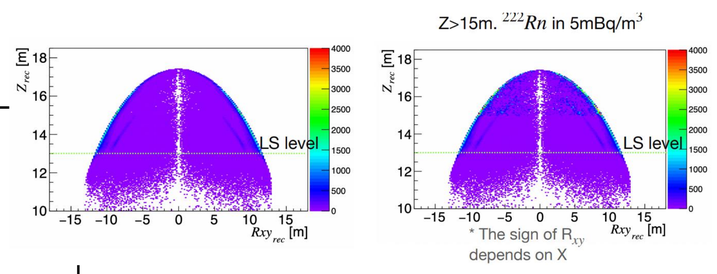
VII. Performance

- ➢ Test based on 1 kHz Mock data shows that OEC can continuously run for more than 21 days without error on DAQ server
- ➢ The OEC software run successfully on the real data of the light-off test done at JUNO site
- ➢ Using Mock data (@1 kHz), preliminary test results show that after OEC and lossless compression, the data volume is approximately 80MB/s, which still needs further optimization.



Online data quality monitoring

- ➢ Quickly provide display plots based on fast reconstruction and classification results.
  - ✓ Check status during the calibration data taking
  - ✓ Check for any background contamination or leakage
  - ✓ Check if there are any anomalies in the detector performance



Once there is radon contamination, it can be clearly seen in the vertex distribution

VIII. Summary

- ✓ The data volume from JUNO detector is enormous, requiring online event classification to compress the data volume.
- ✓ The full-channel test on DAQ with OEC software can steady run for more than 21 days using mock data (@1 kHz).
- ✓ The implementation for all the technical requirements has been met and the physical performance needs to be further optimized.

Looking forward to the data taking at the end of this year for the largest liquid scintillator detector!