



Contribution ID: 88

Type: Sectional reports

## Event index based correlation analysis for the JUNO experiment

*Tuesday 6 July 2021 13:30 (15 minutes)*

The Jiangmen Underground Neutrino Observatory (JUNO) experiment is mainly designed to determine the neutrino mass hierarchy and precisely measure oscillation parameters by detecting reactor anti-neutrinos. The total event rate from DAQ is about 1 kHz and the estimated volume of raw data is about 2 PB/year. But the event rate of reactor anti-neutrino is only about 60/day. So one of challenges for data analysis is to select sparse physics signal events in a very large amount of data, whose volume can not be reduced by using the traditional data streaming method. In order to improve the speed of data analysis, a new correlated data analysis method has been implemented based on event's index data. The index data contain the address of events in the original data files as well as all the information needed by event selection, which are produced in event pre-processing using the JUNO's Sniper-based offline software. The index data are subsequently selected by using refined selection criteria with Spark so that the volume of index data is further reduced. At the final stage of data analysis, only the events within the time window are loaded according to the event address in the index data. A performance study shows that this method achieves a 15-fold speedup compared to correlation analysis by reading all the events. This contribution will introduce detailed software design for event index based correlation analysis and present performance measured with a prototype system.

### Summary

**Authors:** LIN, Tao (IHEP); Dr LIU, Yan (IHEP); Prof. LI, Weidong (IHEP); Dr ZOU, Jiaheng (IHEP)

**Presenter:** LIN, Tao (IHEP)

**Session Classification:** Big data Analytics and Machine learning.

**Track Classification:** 9. Big data Analytics and Machine learning