Research of Readout and Event Building Scheme of Juno DAQ Data Flow

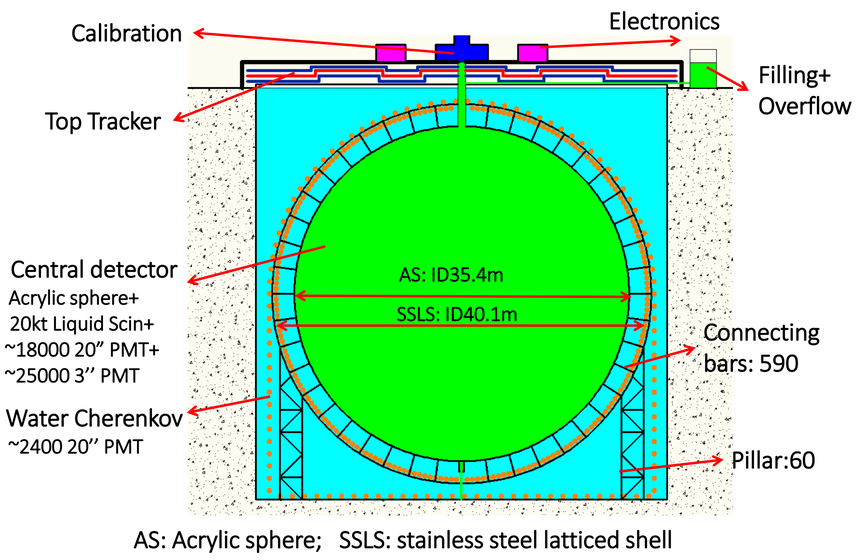
Tingxuan Zeng[1,2], Fei Li[1] , Kejun Zhu[1,2] , On behalf of the JUNO collaboration

[1] State Key Laboratory of Particle Detection and Electronics, Institute of High Energy Physics, CAS, Beijing 100049

[2] University of Chinese Academy of Sciences, Beijing, 100049, China (Email: lifei@ihep.ac.cn; zengtx@ihep.ac.cn)

UCAS IHEP

1. Introduction



AS: Acrylic sphere; SSLS: stainless steel latticed shell

In the JUNO experiment, The Data Acquisition(DAQ) system is an important part to read out and process data from the front end electronics(FEE).The performance requirements of DAQ depends on the design of detector and FEE system.

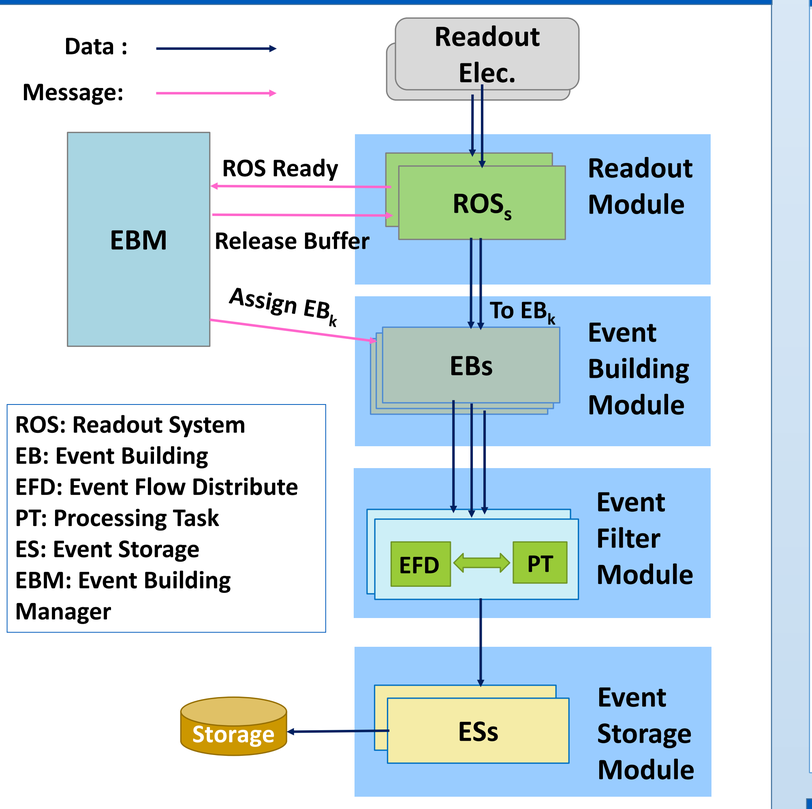
- Electronic scheme is not yet fully determined, we choose the most demanding one for DAQ data flow.
  - Electronic scheme: BX solution.
  - DAQ and electronic connections = number of PMTs (for 20" PMT)
- ~20K PMT(~20K connections), 12 bit 1Gsps FADC waveform sampling , 1us sampling window
  - ✓ CD 18K 20" PMT, Trigger Rate ~ 1KHz
  - ✓ WP 2400 20" PMT, Trigger Rate < 1KHz

2. Objective

Number of Connections	Event Rate	Packet Length	Date Rate
20K	1KHz	2KB	40GB/s

- Design a Data flow schema to read out and process FEE Data
- Verify the feasibility
- Mainly Focus on Readout System and Event Building System

3. Data Flow Architecture



- Based on ATLAS TDAQ and BESIII Data Flow
  - ✓ Mature
  - ✓ Distributed network readout
  - ✓ Complete functionality
  - ✓ Good scalability
- Need verification for a large number of connections and small data packets

4. Research and Test Platform

Hardware

	Lenovo Blade Server
Number of Lenovo Blades	28
CPU	Intel(R) Xeon(R) CPU E5-2620 v2 @ 2.10GHz
CPU cores	12
Hyper-threading	On
Operating System	Scientific Linux CERN SLC release 6.6
Kernel version	2.6.32-504.el6.x86_64

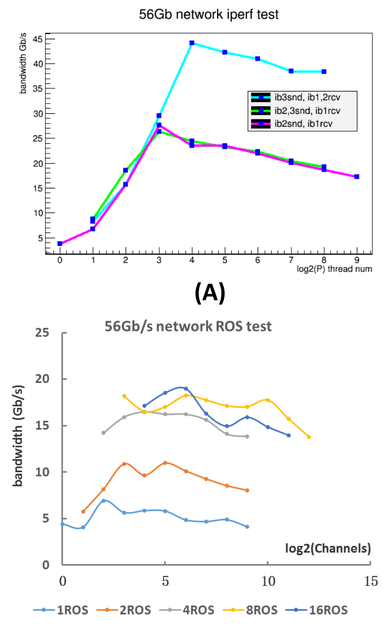
- 28 Lenovo blades(nodes)
- 56 Gb/s NIC

Software

- SoftWare Simulated FEE (send data as fast as possible) + DAQ Data Flow Software
- Iperf as a reference

5. Readout Performance

（1）Single Node TCP Readout Performance



- Fig (A)， iperf send nodes vs recv nodes: 1 vs 2, 2 vs 1 , 1 vs 1
- 1 vs 1 , 8 threads reach max bandwidth 27Gbps@2KB
- 1 vs 2 , 16 threads reach max bandwith 44Gbps, recv node is the bottleneck.

Fig (B), 1 readout thread/ROS (1~16ROS), 1~4096 channels.

- 8ROS reach limitation 18Gbps, support for 1024 channels readout(16Mbps/ch),
- Performance doesn't increase when arranged more than 1024 channels, 128chs/thread.

with single recv buffer

（2）Mutiple Node TCP Readout Performance

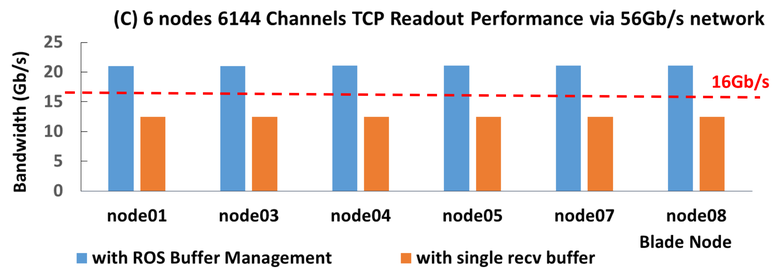


Fig (C)， extend the single node deploy schema : 8ROS * 128 Channels per node to 6 nodes.

- With single recv buffer Management, 21Gbps per node > 16Gb/s
- With ROS Buffer Management, 12.4Gbps per node < 16Gb/s, can't meet the designed data rate(16Mbps per Channel).

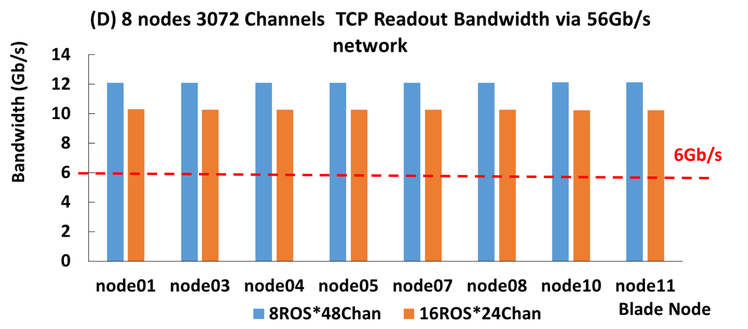


Fig (D), Decrease the number of channels processed by each readout node, based on extensive testing, two deploy schema can meet the designed requirements with enough margin:

- Each node deploy 8ROS * 48 Channels, 12Gbps per node > 6Gb/s
- Each node deploy 16ROS * 24 Channels, 10Gbps per node > 6Gb/s

6. Event Building Performance

（1）Single Readout Node vs Single Event Build Node

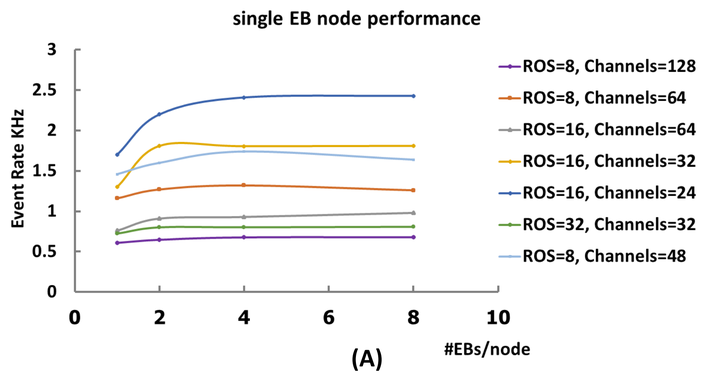


Fig (A)， Software emulated PMT data by ROS self, 1 readout node vs 1 event building node, different deloyment on each node. With Readout node 16ROS*24Channels and Event Building node 8EBs reaches event rate 2.4KHz, keeping enough event rate margin.

（2）Multiple Readout Node vs Multiple Event Build Node

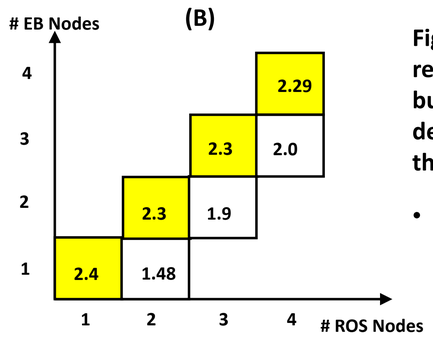


Fig (B), increase the number of readout nodes and event building nodes, the event rate decreases when the ratio of them becomes less than 1.

- To get enough event rate margin , 2.4KHz, readout nodes : event building nodes should be 1:1.

（3）Readout Node + Event Building performance

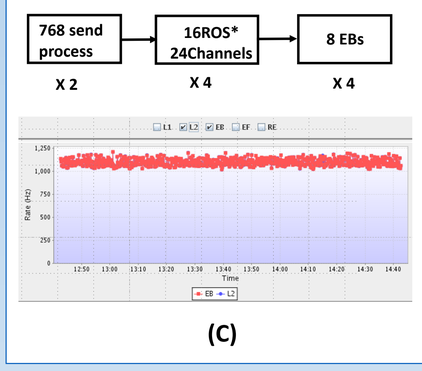


Fig (C), Integrate the readout mudule and the event building module, with each module occupies 4 nodes, reaches 1.06KHz event rate, which meets the designed event rate 1KHz.

7. Overall integration performance

Integrate all the modules to get the overall performance

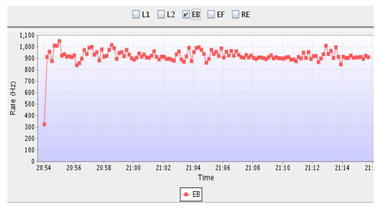


Fig (A)， process a total of 1536 Channels, deploy as follows:

- Simulated FEE: 2nodes*768 send process
- Readout : 4 nodes * 16Chs/node
- Event building: 4 node*8EBs/node
- Other modules: 6 nodes
- Achieved event rate : 926Hz, almost meet the 1KHz requirement

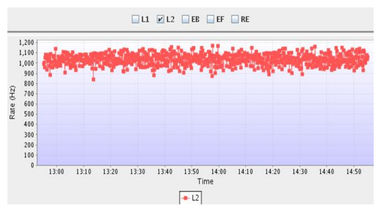
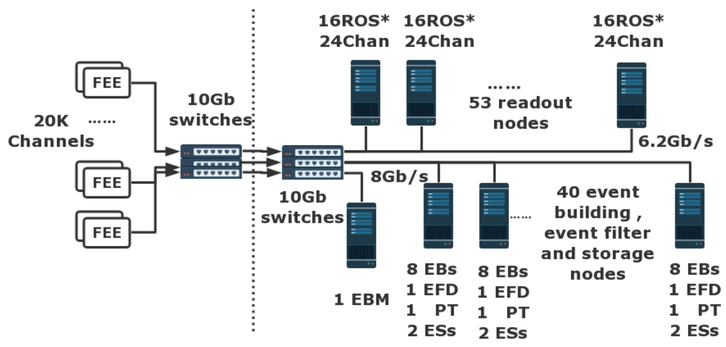


Fig (B)， process a total of 3072 Channels, deploy as folows:

- Simulated FEE: 6nodes*512 send process
- Readout : 8 nodes * 16Chs/node
- Event building + other modules: 6 nodes, to eliminate the back-end network transmission.
- Achieved event rate : 1.03KHz, meet the 1KHz requirement

8. Summary and Outlook



- With dummy event filter algorithm, and Xeon(R) CPU E5-2620 v2 @ 2.10GHz
- To process 20K channels, need:
  - ✓ 53 readout nodes , with 6.2Gb/s per node.
  - ✓ Event building+event filter+storage 40 nodes , with 8Gb/s each node.
  - ✓ 10Gb/s NIC can meet the requirement.

9. Reference

[1]JUNO Collaboration. JUNO CDR.Beijing, China. March 14, 2015

[2] Xiaoshan,Jiang. Electronics Scheme[R],2016.

[3] Xiaoshan,Jiang. One electronics Box for Three PMTs(1F3)[R],2017.

[4] Fei Li, JUNO DAQ Progress[R],2017.7