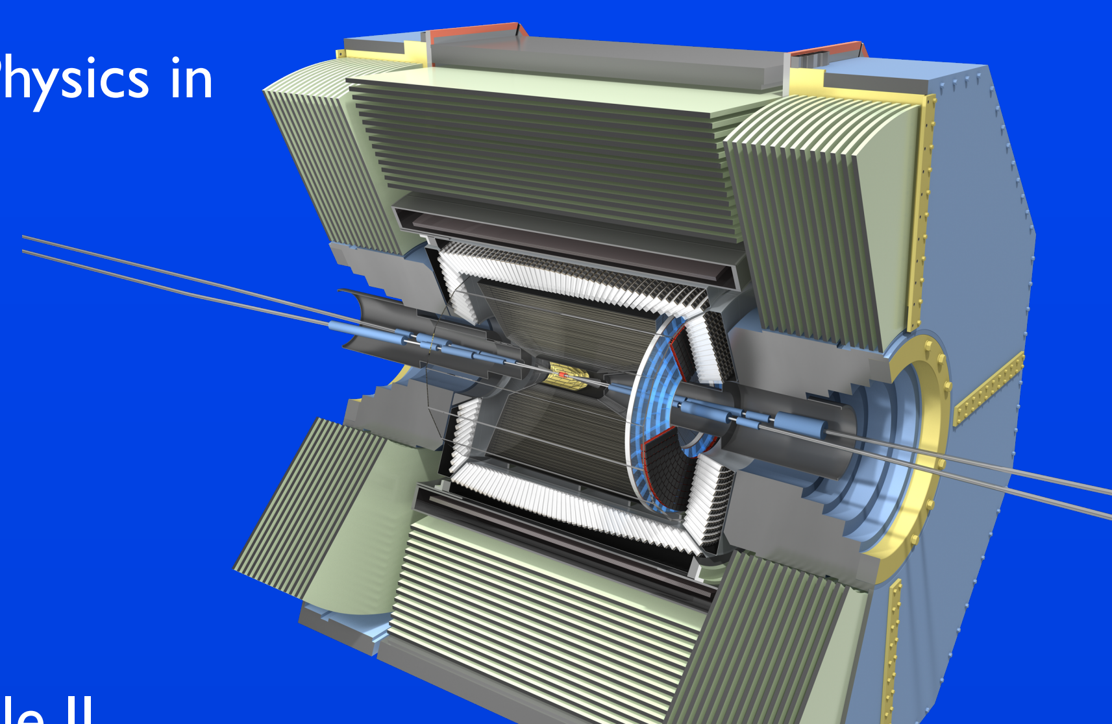
# Belle II High Level Trigger at SuperKEKB

S. Lee[1], R. Itoh[2], T. Higuchi[2], M. Nakao[2], S.Y. Suzuki[2], and E. Won[1] on behalf of the Belle II Collaboration
[1]Korea University, [2]High Energy Accelerator Research Organization (KEK)

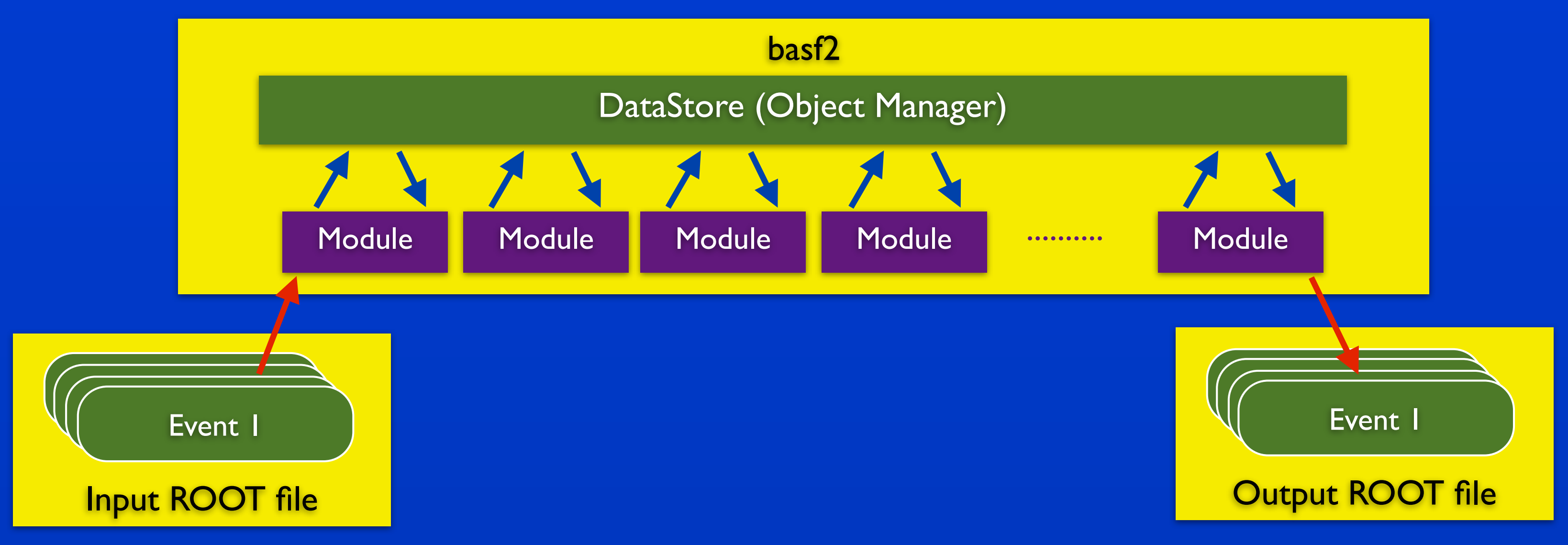
## SuperKEKB and Belle II



- Belle II at SuperKEKB is a new generation B-factory experiment with 40 times larger luminosity than Belle at KEKB
- The main purpose is to search for New Physics in the quantum effect in B meson decay
- Online throughput is estimated to be over 1 GB/s so that the online data processing is a computational challenge

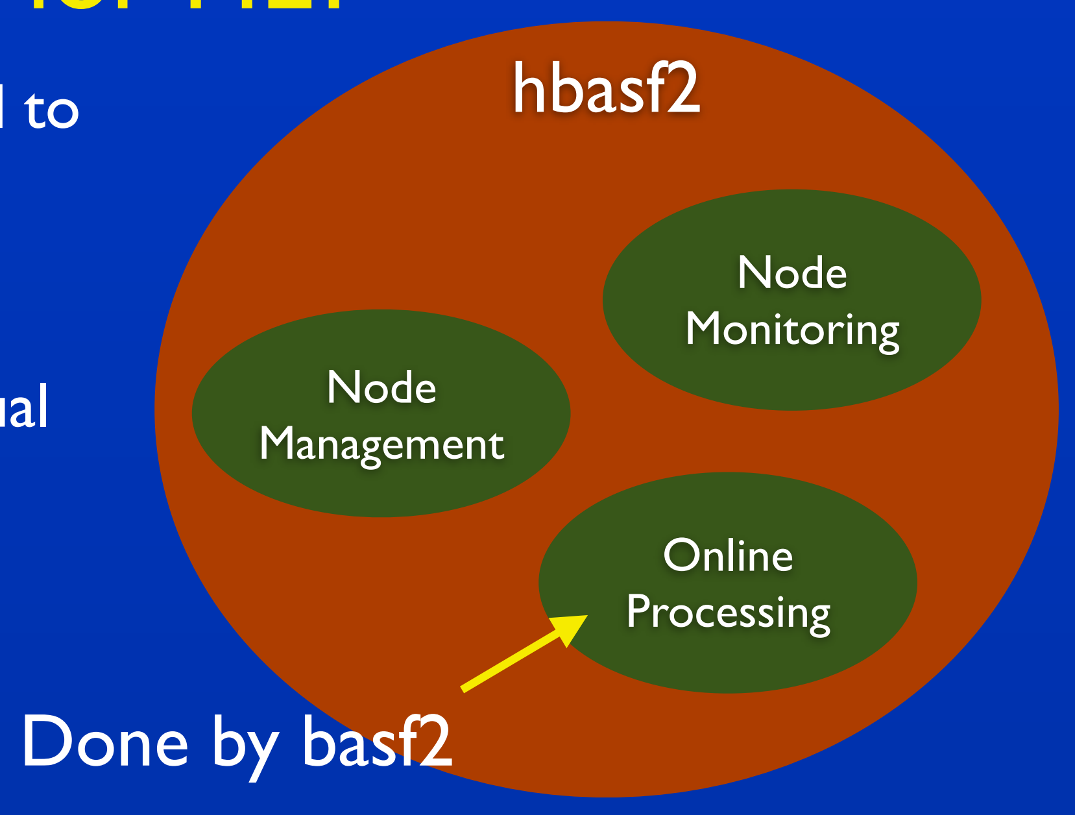
## basf2 framework

- basf2 is a software framework for the Belle II experiment which is supposed to be used in both online and offline processing
- basf2 has modular structure (software pipeline architecture) so that a large scale application can be implemented by combining small modules for specific purpose
- The sub-packages of basf2 extent its functionalities, i. e. networking as a part of HLT package (hbasf2) and grid (gbasf2)
- Parallel processing is also available as:
  - ✓Multi-process based (R. Itoh, Poster session #155)
  - ✓Multi-node (network) based (This contribution)



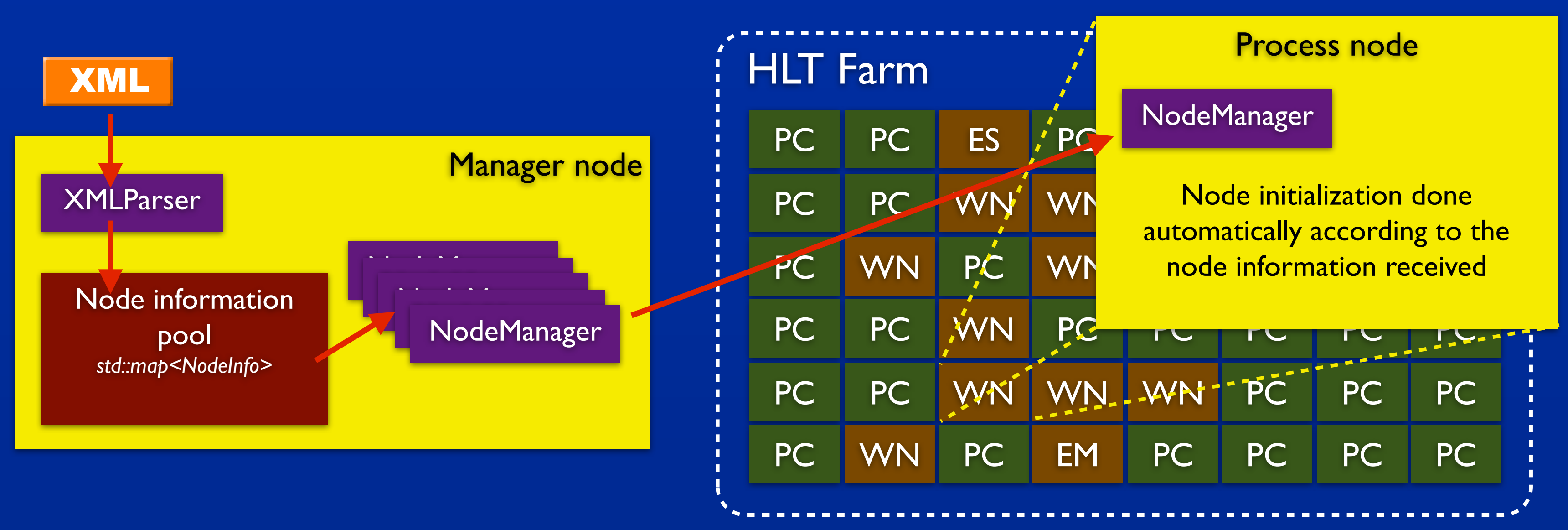
## hbasf2: Super-framework for HLT

- hbasf2 is a super-framework specialized to HLT and it supports the data transfer between nodes over TCP socket, node management and monitoring
- basf2 is called inside hbasf2 for the actual event processing so that the modular structure of the framework is still kept



## Node Management/Monitoring

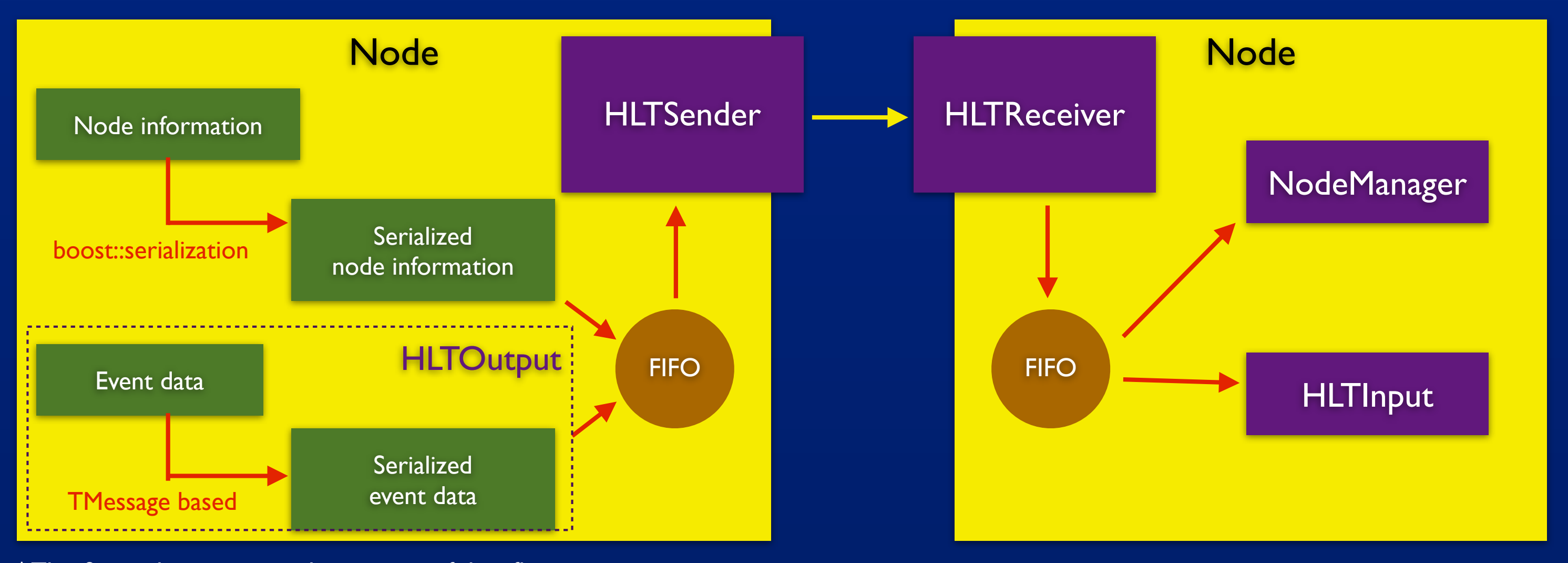
- HLT nodes are controlled by a special node called manager node
- Node information is provided from outside using XML format and manager node provides the individual node information to other nodes
- Process nodes receive the node information and initialize themselves



Node initialization done automatically according to the node information received

## Object Serialization and Networking

- Event data objects are serialized based on TMessage of ROOT and passed to other nodes event by event
- Standard TCP socket based implementation (B2Socket) is used for networking
- Node information distribution is also done by using B2Socket
- Inherited classes named HLTSender/HLTReceiver play roles for data sending/receiving, respectively
- HLTSender/HLTReceiver are independent processes and shared memory based FIFO is used for interprocess communication

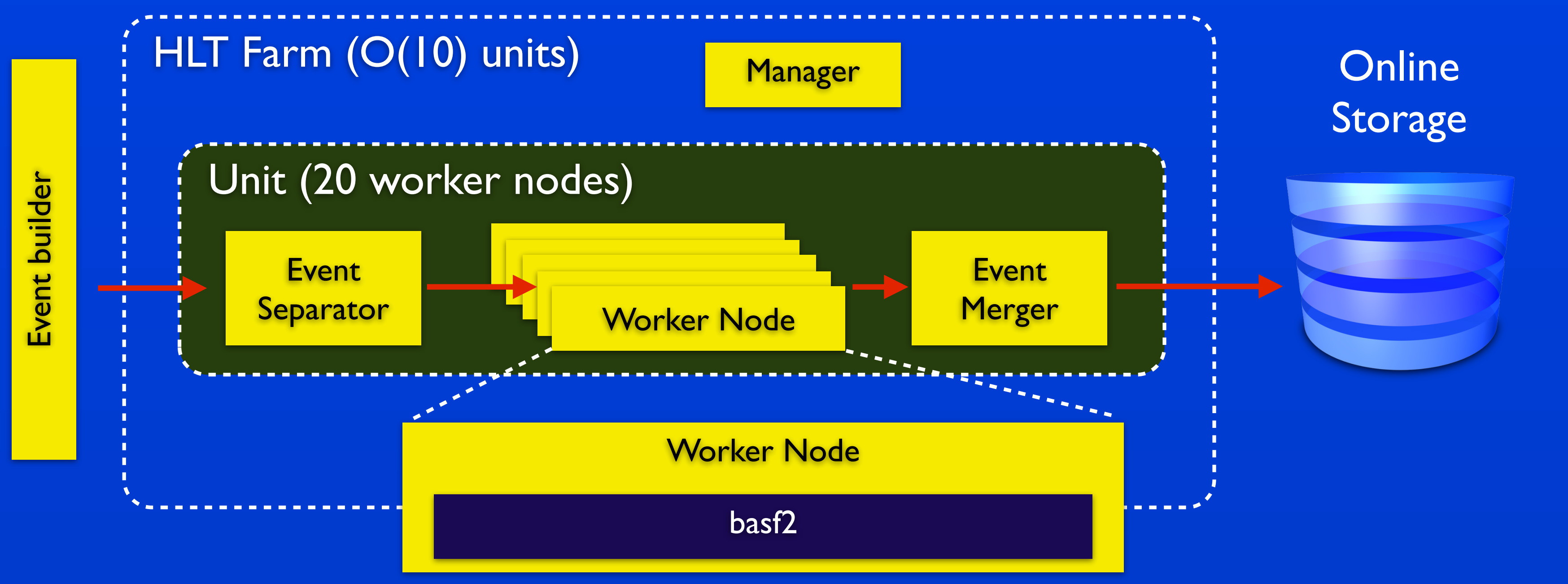


* This figure demonstrates the concept of data flow
Node information and event data don't share the FIFO in reality

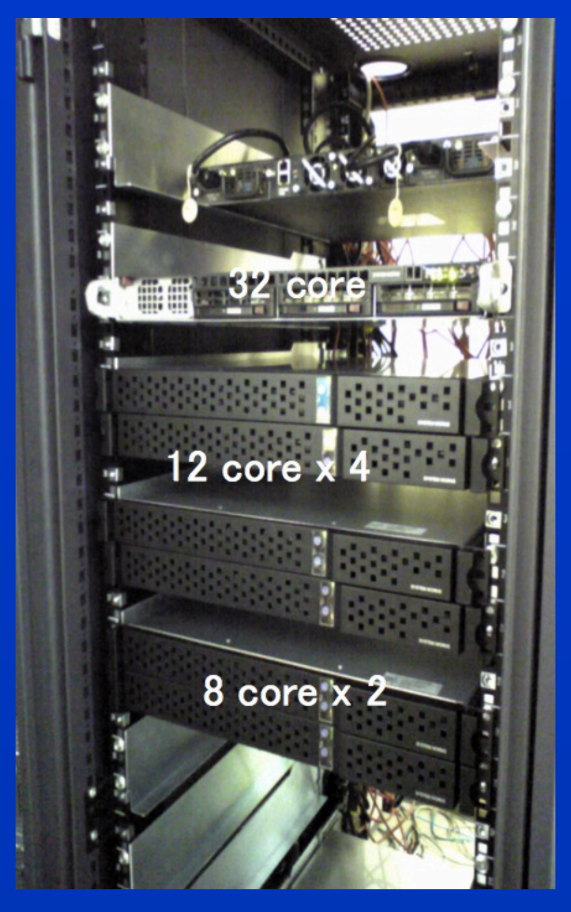
* HLTInput/HLTOutput are modules placed on the module chain of basf2 and provide interface to FIFO

## Belle II High Level Trigger (HLT) system

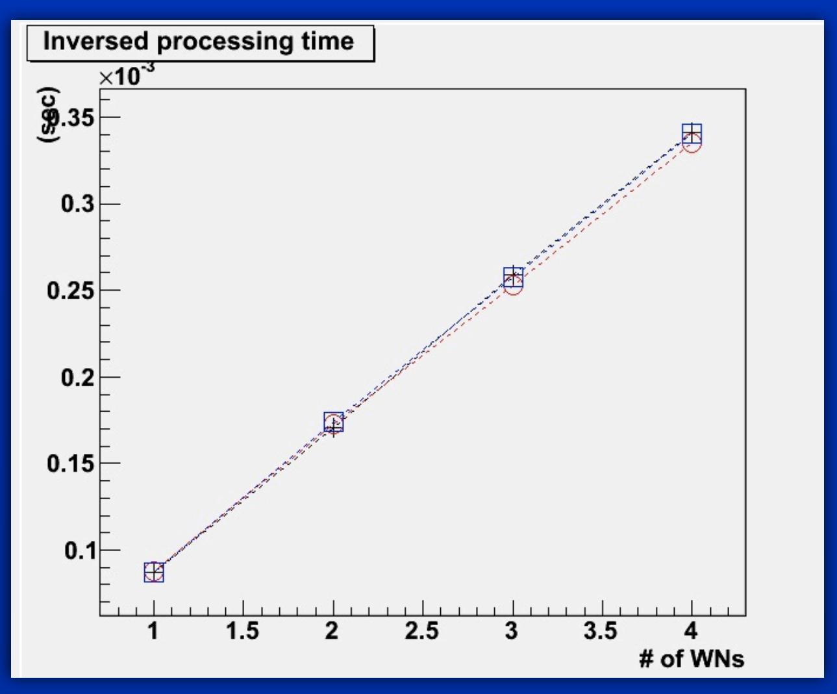
- The Belle II HLT system triggers the event data taken from detector in a software way
- Full event reconstruction using the same offline software is performed
- Full physics skim (i. e. hadronic event selection) is performed and uninterested events are discarded as the actual software trigger
- Major role: Information of selected events are passed to pixel detector processor for reduction of both data rate and size
  - Expected input rate to HLT: 100 kB/event * 30 kHz = 3 GB/s
  - Expected reduction factor at HLT: 1/3 to 1/5 (< 1 GB/s output)
- HLT system consists of units which are basic blocks of event processing
- HLT system = O(10) units with 20 worker nodes of many CPU cores
  - ✓~ 2,000 cores @ 3 GHz (at the beginning of the experiment)



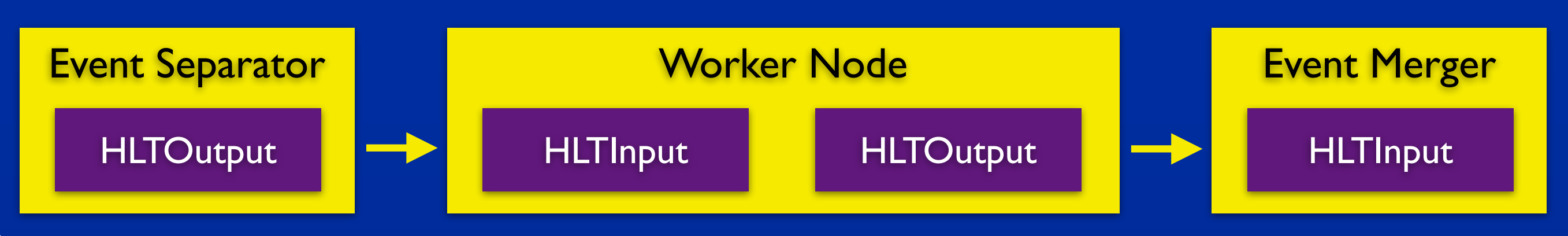
## Performance Tests



- Test environment
  - ✓Intel Xeon CPU / 24GB RAM / 10 GbE network
  - ✓Scientific Linux 5.5 (kernel 2.6.18, x86_64)
- B2Socket unit test
  - ✓Transfer rate = 392.4 MB/s (for 1.4 GB data)
- FIFO unit test (5.7 GB data)
  - ✓1 process writes, 1 process reads = 2.34 GB/s
  - ✓1 process writes, 4 processes read = 2.21 GB/s
  - ✓4 processes write, 1 process reads = 1.60 GB/s
- File I/O unit test (specialized to HLT purpose)
  - ✓Read rate = 110.1 MB/s (for 461 MB data, debug version is used)
- Linearity test for # of worker nodes
  - ✓Event generation and detector simulation are all included (The test configuration does not represent the reality of HLT at all)
  - ✓The linearity of the framework is confirmed up to 4 worker nodes



- Maximum performance test
  - ✓Ideal case (no file I/O and object serialization)



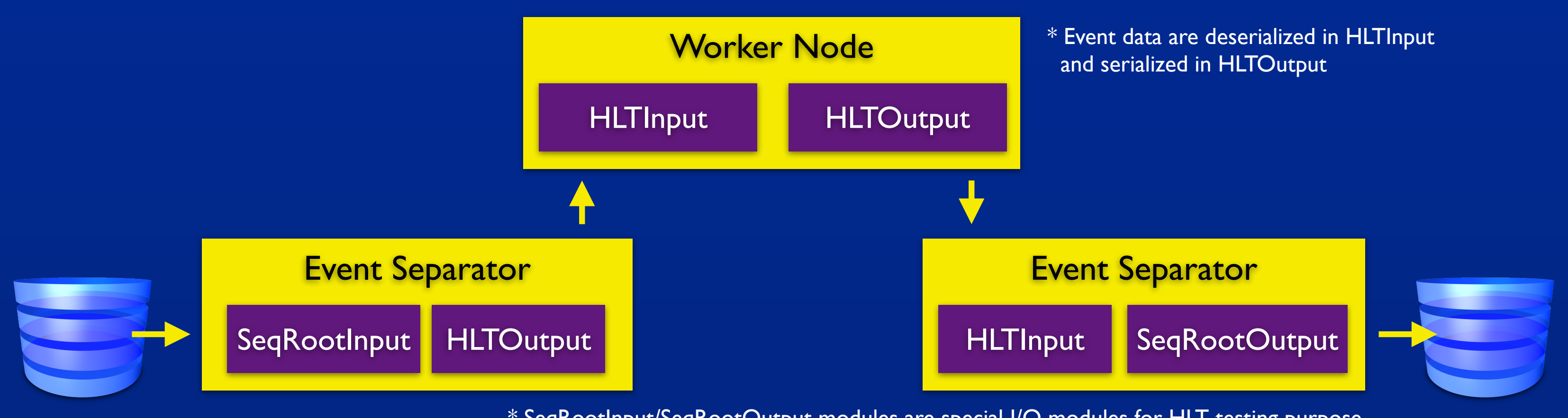
* Event separator generates a dummy data (size = 5 MB) and sends it to worker nodes 1,000 times

* Event data are directly passed without (de)serialization

* Event merger accepts events and just discard them

➡ Estimated performance = 192.3 MB/s

  - ✓Realistic case (object serialization implied, minimum file I/O)



* Event data are deserialized in HLTInput and serialized in HLTOutput

* Pre-generated/simulated input data (5 kB/event)

* SeqRootInput/SeqRootOutput modules are special I/O modules for HLT testing purpose
In real case, input and output data come in and go out over network

➡ Estimated performance = 40.7 MB/s (debug version is used)
➡ The performance drops because of overhead from data serialization

## Further Works

- Optimization of data serialization from DataStore is in progress and good improvement of performance is expected with optimized version
- Performance test combined with multi-process based parallel processing (R. Itoh, Poster Session #155) will be done

Computing in High Energy and Nuclear Physics 2012, May 21-25 2012, New York University, New York, NY, US