Data preservation at BESIII/IHEP

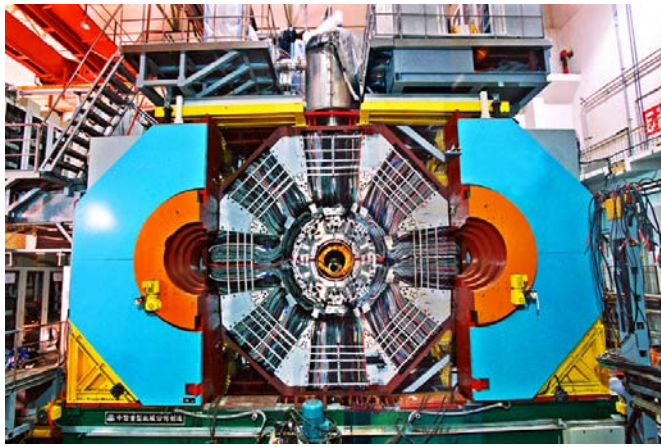
Gang Chen, IHEP/CAS

Third Workshop on Data Preservation
and Long Term Analysis in HEP

Dec. 07, 2009

BESIII/BECPII

- BEPC: Beijing Electron Positron Collider
  - Started in 1989, upgraded to BEPCII since 2004
  - Dual-Ring, 2~5GeV/C
  - Luminosity (3~10)$\times 10^{32}$ $cm^{-2}s^{-1}$
- BES: Beijing Spectrometer
  - Upgraded to BESIII with BEPCII
  - Started to collect data in May of this year



Data Volume of BESIII

Data Type	Volume (TB)	Media
Raw	960	Tape
Reconstructed	2880	Tape & disk
DST	80	Disk
MC-Raw	480	Tape
MC-Rec.	1440	Tape
MC-DST	80	Disk
Total	5760	

Data Analysis Model

- **RAW data:** delivered by DAQ for reconstruction in byte stream format
- **MC Event Data:** contain digits, hits and other MC information in ROOT format.
- **REC data:** written in ROOT format.
- **DST Data:** a reduced event representation suitable for analysis in ROOT format.

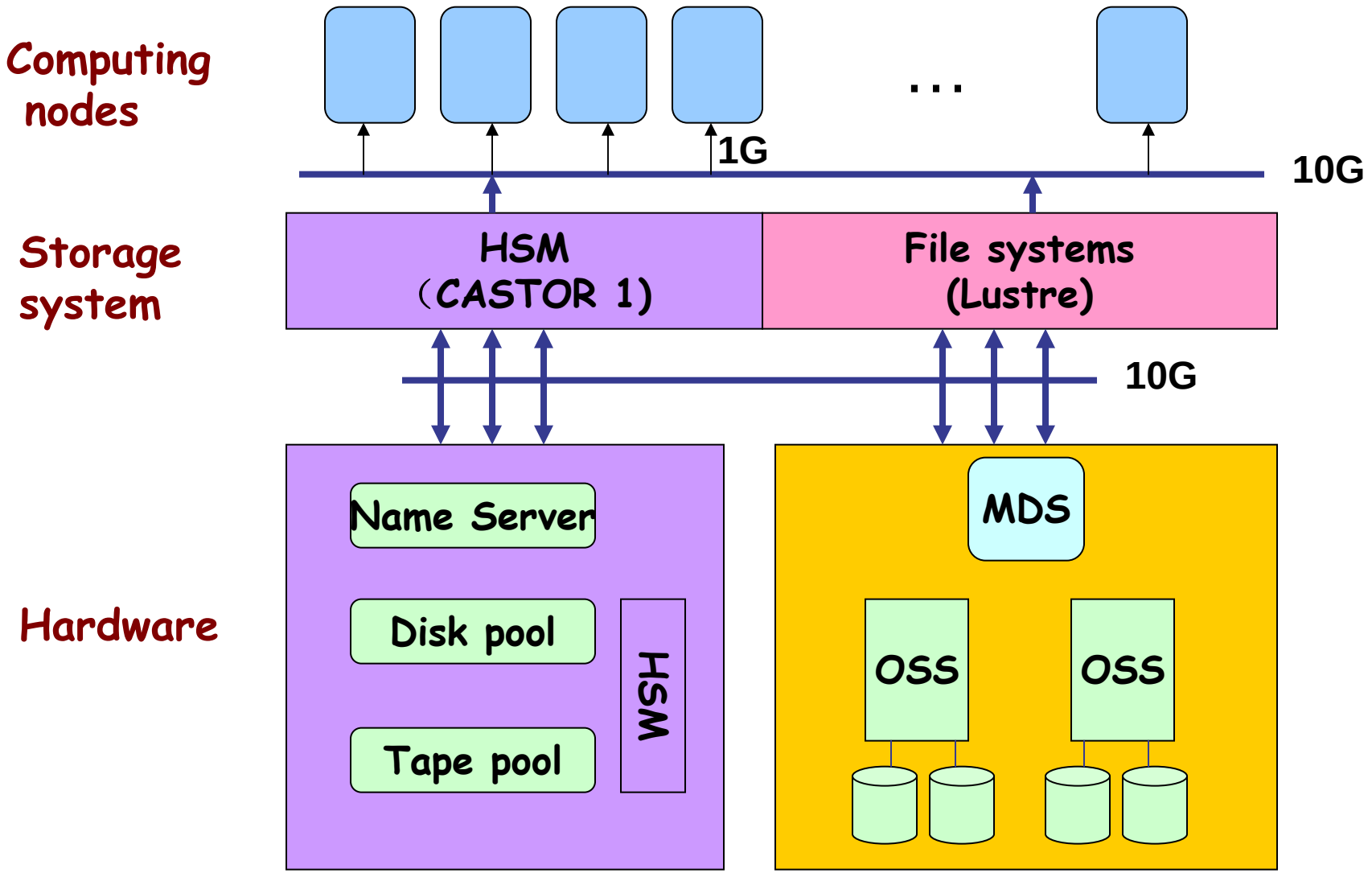
Data Analysis Model, contd.

- **Size of current real event:**
  - 12 KB for raw, 35 KB for REC and 5 KB for DST
- **Size of MC event:**
  - 8 KB for raw, 40 KB for REC and 13KB for DST
- **All the data are copied to disk for processing or physics analysis**
- **Information about job and processing history are recorded in a book-keeping system.**

Computing Resources

- **PC cluster: for reconstruction and physics analyses**
  - 2650 cores, Xeon/3.0GHz, E5140, E5430, E5520
  - 1GB-2GB memory/core
  - IBM/HP/Dell blade systems with 10Gb uplink to Core Switch.
- **GPGPU cluster: for partial wave analyses**
  - 31 nodes
  - 69 ATI/Radeon-4870×2 cards
  - 32 NVIDIA/Tesla C1060 cards
  - Developed on Brook+ , mainly for computing intensive jobs
  - ~10 times faster with double precision or ~100 times faster with combined precision, than CPU

Storage System



Storage System

- **HSM based on CASTOR v1.**
  - No Oracle at IHEP, and scope is not large as CERN
  - IBM3584 and LTO-4 as tape library, 5300 slots
  - Stage system re-written
  - File reservation function added.
- **Lustre as disk file system**
  - Infortrend and Huawei used as hardware platform
  - 750 TB, with above 10 GB/s throughput
  - Possibility to replace CASTOR disk pool being investigated

Data Preservation Prospects

- **Not yet clearly defined, but the followings should be preserved:**
  - Data about the experimental conditions and various parameters like calibration constants, detector geometry data etc.
  - Raw data and DST data should be conserved when the experiment system becomes stable..
- **Lifespan of preserved data is expected to be about 15 years**
- **Data will be kept in disk/tape and software source codes in CVS at IHEP**
- **Under investigation: Replica of raw data is supposed to be conserved at the Scientific Data Center of CAS (Chinese Academy of Sciences)**
  - 15 km from IHEP
  - Same tape library, TSM+GPFS as the data management

Software: BOSS

- **Framework:** GAUDI
- **External Libs:** CERNLIB, ROOT, CLHEP, Geant4,...
- **Developing language:** C++, some Fortran, and Java for web applications
- **Database:** MySQL
- **Configuration management tool:** CMT
- **Operation system and compiler:** SLC4/gcc3.4.6
  - **Computing nodes partly in 32-bit, others in 64-bit**
  - **Moving all to 64-bit system**

Moving software to 64-bit

- It is time to do the porting since all machines are 64-bit
  - Not too late , but definitely not too early to get going
- All BESIII storage systems are already in 64-bit mode
  - Lustre etc. is more stable on the 64-bit operating system
- Three modes
  - Legacy mode: 32-bit compilation running on 32-bit operating system (64-bit hardware)
  - Compatibility mode: 32-bit compilation on 64-bit system
  - Native mode: 64-bit compilation on 64-bit system
- Compatibility mode has been fully tested, BESIII software has been working
- Native mode is been validated, and we hope to migrate to native mode in next year.

Conclusion

- BESIII just started to be operational, data preservation may not be the urgent issue, true or not.
- But the data preservation has been started to think about.
- Software is still continuously upgraded.
- Not just meaningful to HEP, but also attracts interests from other communities of CAS
- No funding scheme was defined, but support from CAS will be pursued.

Thanks!