BESIII Data Management System

Qiumei Ma
Institute of High Energy Physics, CAS
12 July, 2018, Sofia,
Outline

• The BEPCII & BESIII Experiment

• BESIII Offline Data Management System

• Summary
Beijing Electron and Positron Collider (BEPCCII)

beam energy: 1.0 – 2.3 GeV

2004: started BEPCII upgrade, BESIII construction
2008: test run
2009 - now: BESIII physics run

- 1989-2004 (BEPC):
  \[ L_{\text{peak}} = 1.0 \times 10^{31} / \text{cm}^2 \text{s} \]
- 2009-now (BEPCII):
  \[ L_{\text{peak}} = 1.0 \times 10^{33} / \text{cm}^2 \text{s} \]
The BESIII Detector

<table>
<thead>
<tr>
<th>Sub-detectors</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDC</td>
<td>Momentum resolution</td>
</tr>
<tr>
<td></td>
<td>dE/dx resolution</td>
</tr>
<tr>
<td>EMC</td>
<td>Energy resolution</td>
</tr>
<tr>
<td></td>
<td>Spatial resolution</td>
</tr>
<tr>
<td>TOF</td>
<td>Time resolution</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>MUC</td>
<td>9 layers RPC, 8 layers for endcap</td>
</tr>
</tbody>
</table>

Trigger rate: ~ 4 KHz
Data rate: ~ 40 MB/s
Types of BESIII data

• Stored in hard Disk & Tape
  ➢ RawData
  ➢ Simulation Data
  ➢ Reconstruction Data
    • RawData
    • Simulation data

• Stored in MySQL Database
  ➢ Data from Daq and SC’s DataBase (about one million records)
  ➢ Calibration Constant data (about ten thousand records)
  ➢ Data from offline or physics software (about two hundred thousand records)
  ➢ BookKeeping data (about four million records)
How to move RawData to offline hard disk

• The Original raw data from DAQ are written in hard disk of DAQ server
• Raw data will be written in the tape ASAP
• Offline people will copy these data from tape to disk by the application bemp.

<table>
<thead>
<tr>
<th>Raw Data Files</th>
<th>Raw Data Size</th>
<th>Dst Data Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>psip</td>
<td>32000</td>
<td>80T</td>
</tr>
<tr>
<td>jpsi</td>
<td>135600</td>
<td>255T</td>
</tr>
<tr>
<td>Psipp</td>
<td>90000</td>
<td>170T</td>
</tr>
<tr>
<td>psipscan</td>
<td>10000</td>
<td>18T</td>
</tr>
<tr>
<td>XYZdata</td>
<td>130000</td>
<td>210T</td>
</tr>
<tr>
<td>Rscan</td>
<td>43000</td>
<td>80T</td>
</tr>
<tr>
<td>tauscanscan</td>
<td>4000</td>
<td>12T</td>
</tr>
<tr>
<td>2175</td>
<td>12000</td>
<td>22T</td>
</tr>
<tr>
<td>4180</td>
<td>45000</td>
<td>80T</td>
</tr>
<tr>
<td>chic1data</td>
<td>12000</td>
<td>30T</td>
</tr>
<tr>
<td>Sum</td>
<td>513600</td>
<td>957T</td>
</tr>
</tbody>
</table>

6DAQ server Hard Disk

DAQ server

Hard Disk

BEMP

Hard Disk (offline)
BESIII Central DB System

1. Code
2. Web
3. MySQL Client

API

Central DB (Master)

selection &
extraction
API

Online DB

Slave of Online DB

Slave of Slow Control DB

Slow Control DB

Slaves DB
@JINR

Slave DB
@SDU

Slave DB
@IHEP

replication

replication

replication

replication

Write

Write

Write

Read

Read

Read

Read

Applications
(BEMP, phpMyAdmin)
Manage data from SC and Online DB

- Stored BEPC & BESIII Detector’s status
- For safety's sake, SC and Online DB server located in Internal Network, they can not be accessed outside.
- Replicated the SC and Online DB data to the server outside, so users can access these data by browser/client/code.
- API extracts the data which users are interested from replication DBs into the Central DB.
Manage Calibration Constant Data

1. Every BESIII detector has its own Calibration Constant data
2. New Raw data files taken, create new Calibration Constant data
3. Store in Hard disk
4. Verify
5. Insert into central DB by the application bemp
6. At the same time, replicated to all slaves
7. Users can read constants from slave (local server)
8. For security reasons database records are never deleted.

<table>
<thead>
<tr>
<th>CalParVer</th>
<th>char(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status</td>
<td>enum('OK', 'TEST', 'ABORT')</td>
</tr>
</tbody>
</table>
BEMP Application (Yao Zhang)

BesIII experiment management platform (BEMP) keeps track of data handling steps and provides facilities to access meta data for physics analysis.

- Typical use cases
  - Physicist Users
    - Submit jobs with input of datasets
    - Retrieve information about a dataset
    - User’s own data production
  - Data Production Manager (collaboration wide data reprocessing)
    - Create datasets in the database
    - Data transfer by dataset
    - Retrieve datasets as input for reconstruction
  - Calibrators
    - Calibration constants management

- Web interface
  [http://bes3db.ihep.ac.cn:8080/bemp/mainMenu.html](http://bes3db.ihep.ac.cn:8080/bemp/mainMenu.html)
The components of Bookkeeping

More please see: https://indico.ihep.ac.cn/event/370/contribution/23
Replication

• Why select MySQL replication pattern
  1. Less write
  2. More read
  3. Data distribution
  4. Load balancing
  5. High availability
  6. Simple to use
  7. Backup

• Results
  • This pattern has satisfied BESIII’s need successfully.

Central DB
Master
Slave @ IHEP
Slave @ SDU
......
slave @ JINR

BESIII DB Replication’s pattern
Access Database Data

- MySQL Client
- API(C++, shell)
  - Framework
    - DatabaseSvc
    - CalibMySQLCnvSvc
- Web
  - http://bes3db.ihep.ac.cn
  - http://bes3db.ihep.ac.cn/phpMyAdmin/
  - http://bes3db.ihep.ac.cn:8080/bemp
Maintenance of data

• Consistency and integrality
  • Monitor the difference between data stored in Master and Slaves
  • Monitor the status of the slaves

• Security
  • Limit specific IP
  • Different users have different privileges
  • Log enable
  • Backup

• Recoverability
  • Replication
  • Backup to tapes
Monitor Application

• Monitor the server’s status
  • Got errors will email me and send SMS information to my phone

• Monitor the Slave replication status
  • Get replication errors will email to me
Summary

• A brief introduction about BEPCII and BESIII Experiment
• Details about BESIII Offline data management.
The End.

Thanks for your attention!