

April 27, 2010

Belle II Computing meeting, KEK, Japan

# Belle II Data Handling Report

Kihyeon Cho

# Contents

- Minutes of Meeting
  - April 8, 2010: Belle II DH KISTI meeting
  - April 15, 2010: Belle II DH meeting
- Current Status
- Discussions

# Report on Belle II DH Group

- Kihyeon reported time schedule and each group's activities including KISTI, Polish, Melbourne and computing groups.
- We also reported to do list and schedule.
- Hara San commented that the bidding procedure for the KEK resources will start this summer. Final specifications are needed next summer.

# Data Handling Process

- The whole process:

KISTI

Melbourne

- Data Transfer => Extraction => Scalability => Replication => LFC => Grid farm

- KISTI will take care for up to LFC. Then, Melbourne will get LFC file and do Grid test.

- The current status is on the processing of extraction. The extracted files had been already sent to Tom for LFC.

- The procedure for metadata is as follows:

- Metadata => LFC => SE => Physical Data

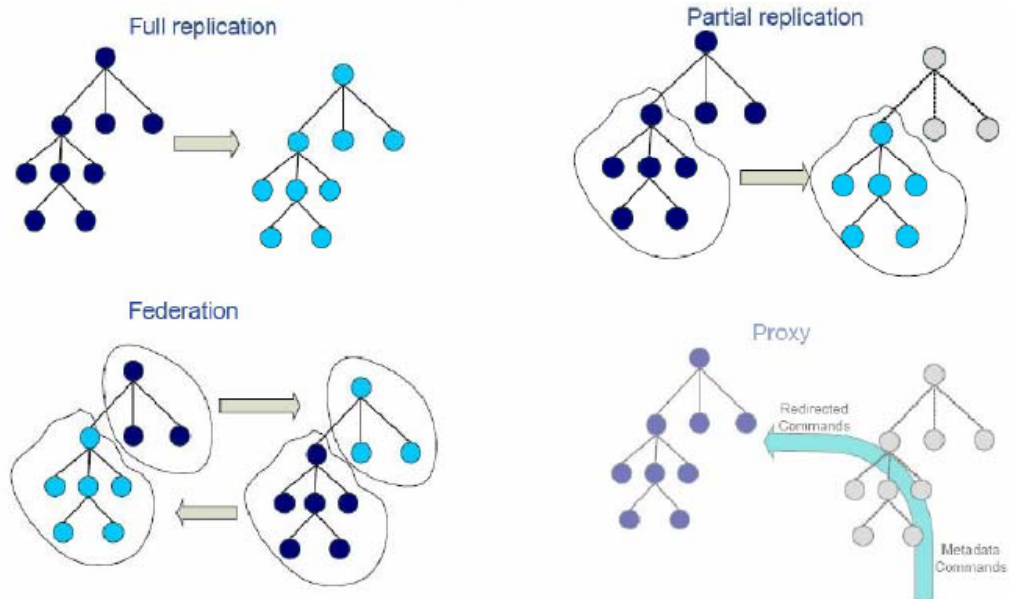
- For this processing, we need to install Grid SE.

# Polish group

- Maciej Siarz (April 8)
- Test for AMGA replication
- Draft plan
  1. Setup one system with AMGA (Virtual machine)
  2. Test this AMGA server
  3. Copy that virtual machine and setup about 4 different servers
  4. Test the replication (master <-> master) (?)  
=> Sunil will give talk about AMGA structure.

- **Replication**

- Master-Slave & Asynchronous communication model
- Full & Partial replications allowed
- Proxy (redirection) will be implemented soon !



# Structure of AMGA (Sunil Ahn)

- Sunil explained the structure of AMGA. There is only one master node is needed for one full data set. The main structure is master-slave nodes. Therefore, a comment is that the plan of master-master test at Polish group will have to be changed as master-slave test.
- Possible slave size of disk space at Poland is 5TB.

# Status of Belle II DH system test (JungHyun Kim)

- Junghyun reported the estimating of size of metadata, current status of meta data extraction and future plan of grid embedment.
- The comment is that from the test, we need the back data for plots of achievements, accessing and processing time, scalability, performance , quality, catalog size and etc.  
=> Today's Junghyun's talk
- His plan is as follows:
  - Extraction => Right now
  - Test metadata => Right now
  - Grid => Right now
  - DH test using Belle data at KEK => This summer






- The software is complete
- We've started generation of 100Million MC events on the cloud,
  - powered by DIRAC
- We're staging out 3TB of data to gridKa
- 80 cores right now.
- On the weekend we go to 2000

# Exciting news about the cloud project (Tom Fifield)

- Tom explained that the cloud management solution based DIRAC.
- He generated 100 Million MC events on the cloud using 80 cores. He will use 1,000 cores this weekend (April 17).
- He will do two different stages.
  - The first stage is to collect data.
  - The second stage is data transfer.

# Belle LSDH Test & NSDC/KISTI



Belle Data Transfer	<ul style="list-style-type: none"> <li>• KEK → KISTI</li> <li>• Currently, 142TB data are archived</li> </ul>	<ul style="list-style-type: none"> <li>• 90% of planned done</li> </ul>
Metadata Extraction		
Scalability Test		
Metadata Replication		
Grid-awaring of Belle data	<ul style="list-style-type: none"> <li>• SRMDoor &amp; Disk Pool (SEDPM)</li> <li>• File registration to SE without copy</li> <li>• Registration to LFC</li> <li>• Co-test with SE &amp; LFC</li> </ul>	<ul style="list-style-type: none"> <li>• SRM Door &amp; Disk Pool → Done</li> <li>• File Registration to SE → Done</li> <li>• Registration to LFC → in progress</li> <li>• Test with SE &amp; LFC → partially done</li> </ul>
Test on Grid Env.	<ul style="list-style-type: none"> <li>• Job submission to Grid ( with access to AMGA, SE, LFC )</li> </ul>	<ul style="list-style-type: none"> <li>• Every Testers should have grid cert. and Belle VO membership</li> </ul>
What Next ?	<ul style="list-style-type: none"> <li>• ???</li> </ul>	

# Progress of making local Belle data grid-aware (Christophe Bonnaud)

- Christophe reported on the status of file transfers and the hardware configuration for the large scale data handling test. The transfers are done to around 90%. The following items are discussed.
- Which LFC?  
=> We will use LFC at KEK Belle. If necessary, please contact Go.
- CE?  
=> First we will use LCG-CE at KEK. Then, later we may use CREAM-CE.
- Who will provide LFC service? Who will put the data into LFC?  
=> Tom and Christophe (Beobkyun) will care for it.  
=> We may use Belle VO at KEK.

# Current status of NSDC (4/26)

- Belle data download : is in progress (as we planned), but no problem for LSDH
- The gridification of local Belle data : done (construction of dpns )
- Internal test of LFC registration : done
- Official registration to LFC (kek2-lfc.cc.kek.jp) : started
  - \* expected lfc path : /grid/belle/mcprod/dat/e0000\*\*\*
  - \* expected time for this registration : not sure(internal test consumed several hours in multiple stream. but LFC @ KEK might have network latency and unexpected interruption.)

# Discussions

1. We will get the results of plots and figures by next Belle II computing meeting.  
=> Today's Junghyun's talk
2. Tom has the converter from Logical to Physical file catalog. Therefore, Tom will care for LFC with Christophe (Beob Kyun) when Junghyun gives LFC of metadata.
3. Next Belle II Data Handling Working Group meeting  
- on Thursday, May 6, 17:00 (KEK time).