

Report on Large Scale DH Test

J.H Kim¹ & S. I Ahn² & K. Cho¹

¹High Energy Physics Team
²e-Science Grid IT Team
KISTI, Daejeon, Korea

Belle II DH group meeting, 2010.04.15

Overview

- 1 After B2GM, Estimating the size of meta-data
- 2 Current status
- 3 Summary and next plan

Estimating the size of meta-data at Belle II

- Assumption
 - ▶ We suppose that it be 4GB/file.
 - ▶ We suppose that skims be around 30 kinds.
- Question: How can we control the skims?
 - ▶ We didn't decide how to control the skims at last B2GM.

Table: Estimation of the meta-data size for Belle II

	# of files (Belle)	# of files	Size for file level	Size for event-level
Raw data		100.0M	60.0GB	
mDST	24k	4.3M	2.6GB	5.6TB
skims	720k	12.9M	7.8GB	
MC	240k	12.5M	7.5GB	17.0TB
skims	7,200k	37.5M	22.5GB	

The large Data Handling test

Current status

- Extraction:
 - Extract meta-data from the Belle Data.
 - on the 1st processing (exp07 ~ exp27)
 - done by 4.08
 - on the 2nd processing(exp31 ~ exp39)
 - updated by 4.12
 - on the 3rd processing(exp41 ~ exp55)
 - from 4.12 to current
- Generating meta-data: by Dr. S.I Ahn
 - We generated the meta data from exp07 to exp39.
- We are waiting the next procedure which converts the LFN to PFN.

- Relication:
 - ▶ Master(KISTI) : 150.183.246.196
 - ▶ Slave(Melbourne) : 150.183.246.196
 - ▶ The slaves are updating automatically.
- After the replication, we search the interesting files from Melbourne and KISTI system.
We get same result between KISTI and Melbourne.
- By Tom's e-mail :
The Polish group has expressed interested in running another slave AMGA server for the purposes of the test.
- Send the IP address for the slave server.

On grid

- We try to work on grid for the meta-system.
 - ▶ KISTI group are trying to get VO
- We will set the meta system on Grid.
We will get the host certification.
- Meta system will be included in grid.

Summary and Next plan

- 1 The extraction and generating meta-data will finish in this week (exp07-exp55).
- 2 We will test the meta-data.
- 3 We will include the meta system in grid.
- 4 After large scale data DH test, we will do the same procedure with KEK data.