

# Report on Large Scale DH Test

J.H Kim<sup>1</sup> & S. I Ahn<sup>2</sup> & K. Cho<sup>1</sup>

<sup>1</sup>High Energy Physics Team  
<sup>2</sup>e-Science Grid IT Team  
KISTI, Daejeon, Korea

Belle II DH group meeting, 2010.04.26

## Overview

- 1 Large Scale data DH test
- 2 Summary and next plan

# The large Data Handling test

## Current status for Belle data

- Extraction:
  - Extract meta-data from the Belle Data.
  - on the 1<sup>st</sup> processing (exp07 ~ exp27)
  - done by 4.08
  - on the 2<sup>nd</sup> processing(exp31 ~ exp39)
  - updated by 4.12
  - on the 3<sup>rd</sup> processing(exp41 ~ exp49)
  - from 4.12 to 4.15
- Generating meta-data: by Dr. S.I Ahn
  - We generated the meta data from exp07 to exp49.
- We finished these works.

- Relication:
  - ▶ Master(KISTI) : 150.183.246.196
  - ▶ Slave(Melbourne) : 192.231.127.47
  - ▶ 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.

## Meta-system test for Large Scale data.

- 1 Size : 31MB, 21 experiments(exp07-exp49), on\_resonance, stream 0,1,2
- 2 extraction time : 1.8min - 18min/file
- 2 generating time : 400files/sec
- 2 Performance :
  - UI : [hep2.kisti.re.kr](http://hep2.kisti.re.kr)
  - Meta system : Melbourne slave( for global network environment )
  - Prototype : `belle_amga_access`

## Performance: searching time for full data of each experiment.

Table: searching time for full data of each experiment

exp#	07	09	11	13	15	17	19	21	23	25	27
spending time (sec)	37	74	70	73	58	54	51	42	53	91	70

  

exp#	31	33	35	37	39	41	43	45	47	49
spending time (sec)	67	50	51	59	60	60	40	40	32	29

- Total spending time (sequential searching for all meta-data): 1161 sec
- Average spending time: 55 sec

## Summary and Next plan

- 1 The extraction and generating meta-data will finish in this week (exp07-exp55).
- 2 We are testing the meta-system.
- 3 multi-query test :  
1<sup>st</sup> : Same queries are performed in meta-system such as multi-query.:done  
(need to make a plot )  
2<sup>nd</sup> : Random queries are performed in meta-system.: will do.
- 4 We are generating the random meta-data for scalability.: doing
- 5 Same test will be performed such as large scale data. : will do.