

May 12, 2011

Belle II Data Handling Group Meeting, KEK, Japan

Status of Belle II Data Handling Group

Kihyeon Cho

High Energy Physics Team

KISTI (Korea Institute of Science and Technology Information)

Contents



- Meetings
- Manpower
- Standalone data registration tool kit
- Drafts
- Issues

Minutes of Computing meeting

(April 19)

2) Status of Data Handling group [Kihyeon Cho : KISTI]

- AMGA was plugged in the GRID farm and test was done. (test with AMGA server at KEK is delayed).
Availability of DIRAC from the GRID farm ?
- Snapshot-like method for user-created data was presented, which is alternative of query-based method. But it easily increases the meta-data size, so the DH group is working on conceptual design for this. In order to assess this idea, we need test results.
- A data set registration tool is required for the real Belle data analysis on the distributed computing.



● KISTI DH meeting

- Date: 2010. 4. 27 11:00–13:00
- Participants: Kihyeon Cho, Junghyun Kim, Sunil Ahn, Taesang Huh, Soonwook Hwang, etc.

● Korea Belle/Belle II Collaboration Meeting

- Date: 2011. 4. 30 (Saturday)
- Place: Hanyang University, Seoul, Korea
- Report on Belle II DH Group
 - Kihyeon Cho

Upcoming meetings



● The 4th Belle II computing workshop

- Date: May 23–25, 2011
- Place: Slovenia

⇒ Three talks by KISTI

- Kihyeon Cho (In person)
- Junghyun Kim (EVO)
- Taesang Huh (EVO)

Upcoming meetings (cont'd)



● Korea Belle/Belle II Collaboration meeting

- Date: End of August, 2011
- Place: KISTI, Daejeon, Korea

● BAM/BAS

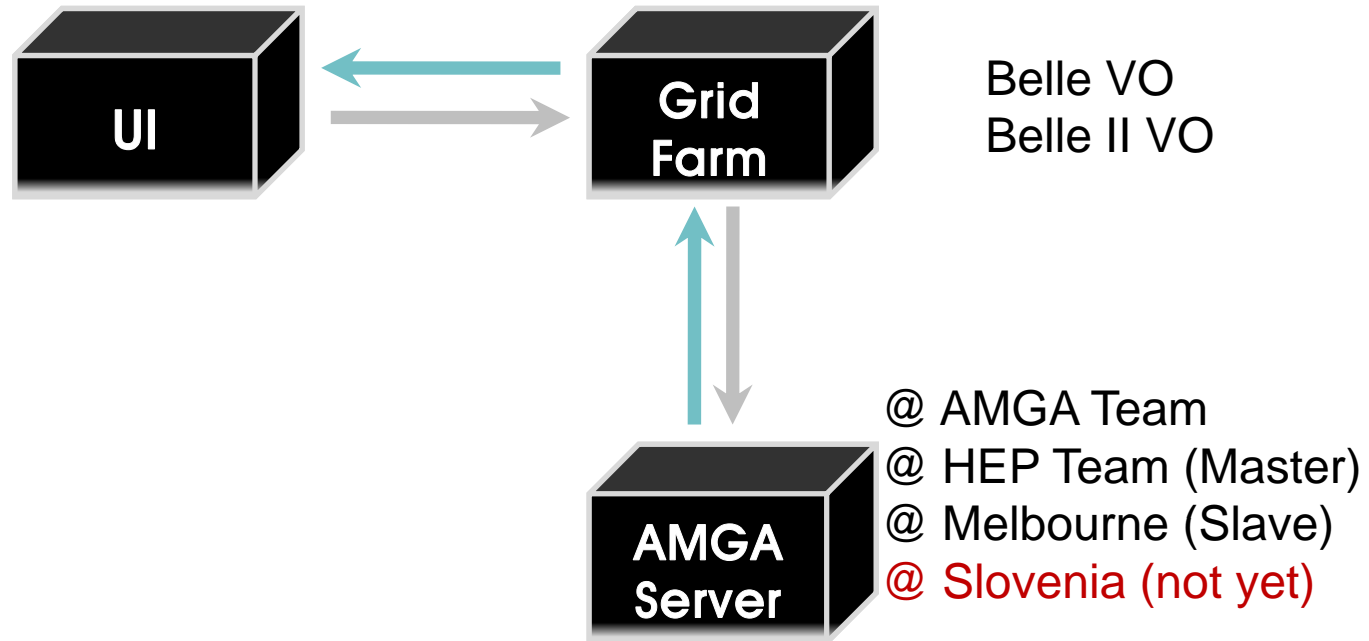
- Date: Sep. 26–27, 2011 (BAM)
Sep. 28–30, 2011 (BAS)
- Place: Jeju Island, Korea
- Host: Korea U., Yonsei U. SNU, KISTI, etc.

Man Power @ AMGA Team



- Sunil Ahn, Taesang Huh, Soonwook Hwang (P.I), etc.
- Sunil Ahn will be at USC, USA from June 1 for one year. He will work there.
- However, Taesang Huh is mainly in charge of AMGA stuffs.
 - Will Give talks about his activities at today's meeting and computing workshop

AMGA test at Grid Farm



Farm	Belle VO	Belle II VO
KISTI Grid farm (sdfarm.kr)	OK (By Sunil)	N/A
KEK Grid farm (kek2-cd05.cc.kek.jp)	OK (By Tom)	OK (By Sunil)

- Sunil has written documentation of how to submit grid jobs.
- We had tested AMGA Sever at HEP team(150.183.246.196).
- AMGA Server at Melbourne is in Grid.

Kihyeon Cho

⇒ Draft for JKPS by this month

#196 Standalone dataset registration tool kit

- Junghyun and Taesang Huh
- C++ module for Belle II

1st step. Gbasf2 @ KISTI

- To check Logfile message and meta data
⇒ **Working on Progress**

2nd step C++ Module

- Open metadata system
 - New metadata system with registration
 - Close metadata system
- A couple of month

Draft for JKPS

(Journal of Korean Physical Society)

- The embedment of metadata system at grid farm at the Belle II Experiment
 - Working on progress by this month

- User created metadata set
 - This summer

● Snap shot vs. Query

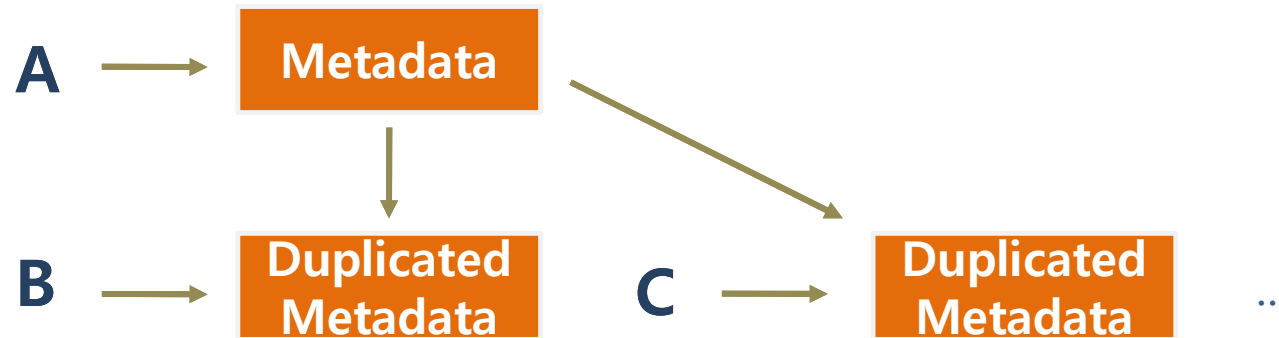
1. Snap shot
2. Query
3. Snap-shot like method

● For Snapshot

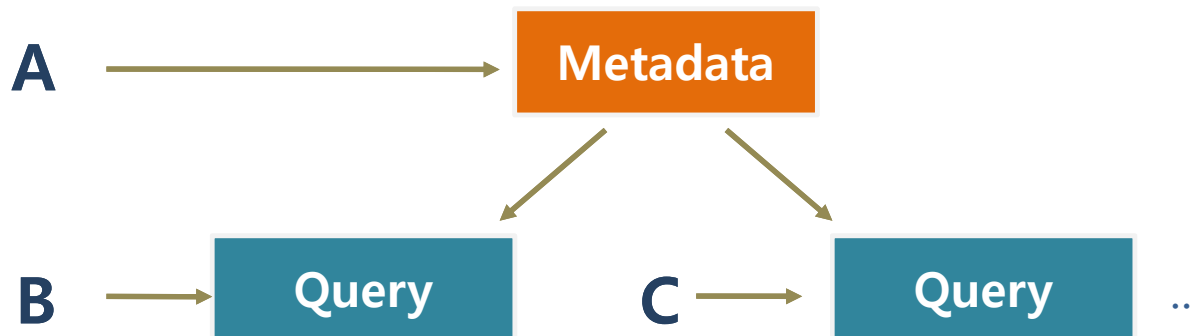
1. Official data → 1.5TB
2. User Created data
 - ① User file (signal) → 70 MB
 - ② Subset of official data → Up to 1.5 TB/user (huge)

Snap shot vs. Query

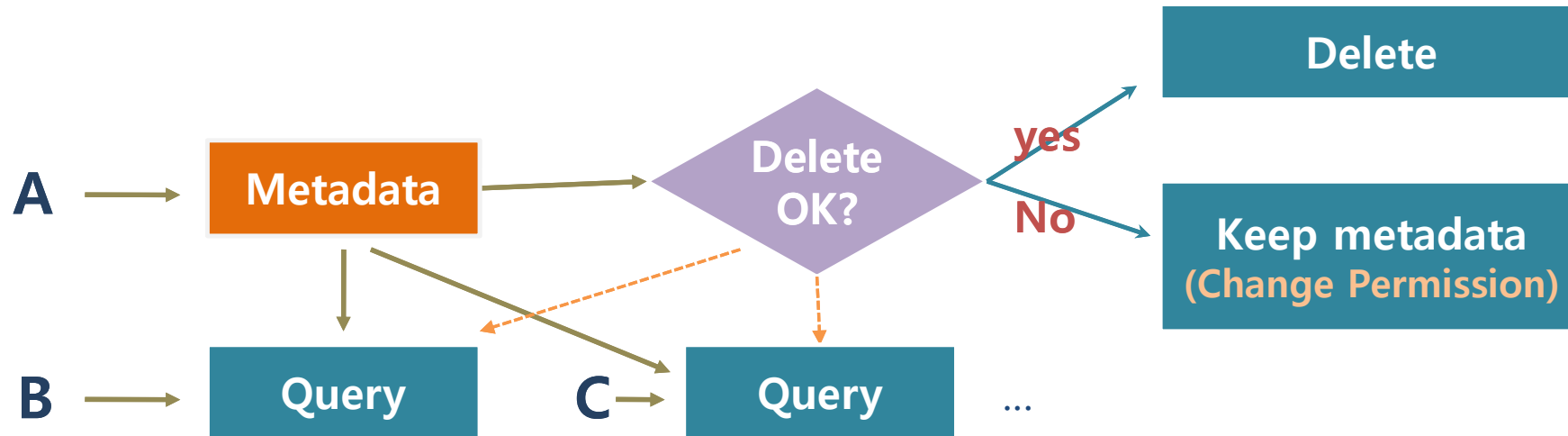
1. Snap shot



2. Query

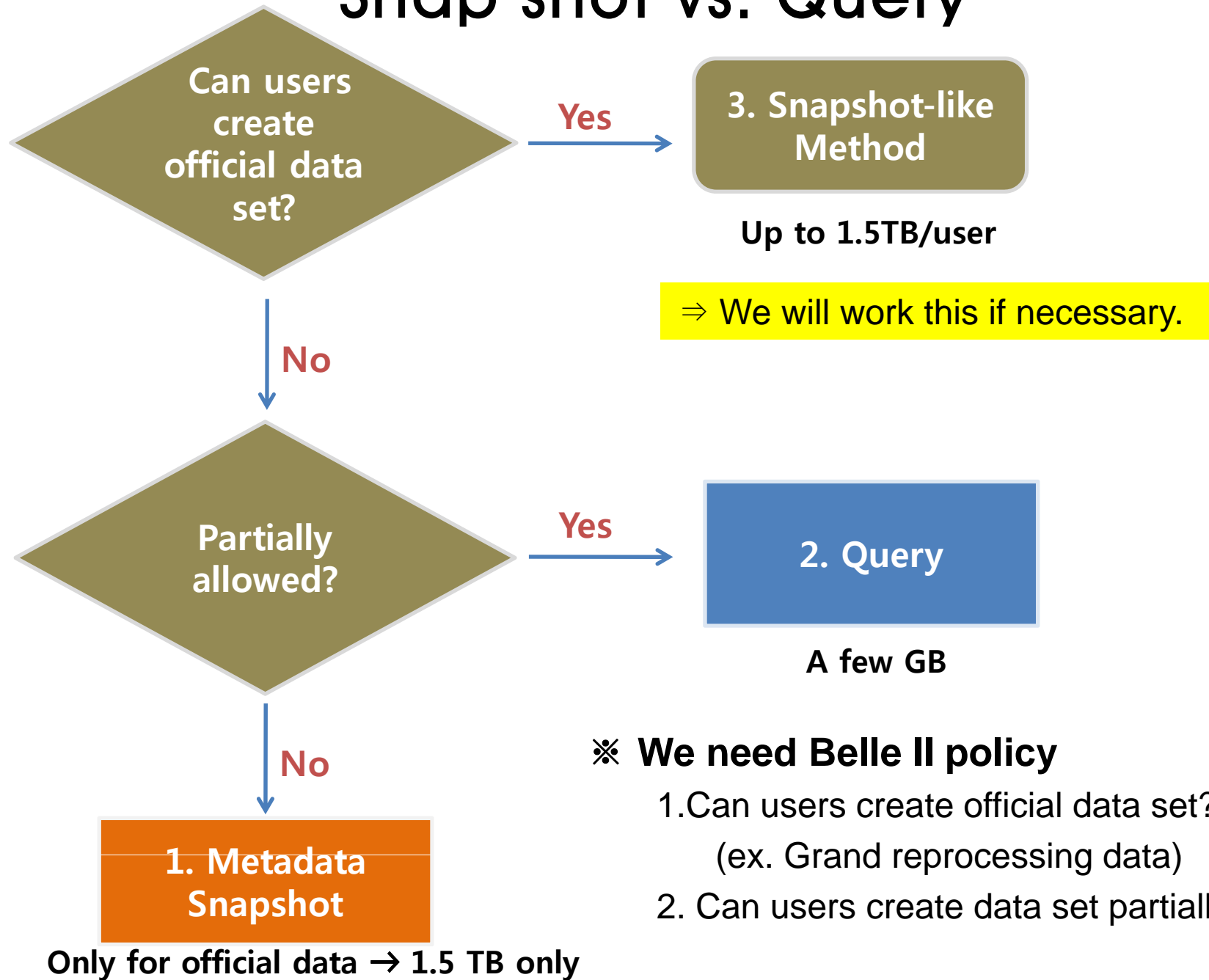


3. Snapshot-like Method



- Snap-shot style \Rightarrow snap-shot like style for user created data
- To store queries which are not duplicated
- To keep the metadata for old data

Snap shot vs. Query



※ **We need Belle II policy**

- 1. Can users create official data set?
(ex. Grand reprocessing data)
- 2. Can users create data set partially?

Thank you.