Belle/Belle-II Activities at KISTI

J.H Kim[1] &

[1]High Energy Physics Team
KISTI, Daejeon, Korea

KSC 2009 Conference, 2009.10.13

Overview

1. The status of Data Handling system
   - Current status in Belle Data Handling system
   - What is AMGA?
   - The Data Handling Scenario
   - The progress of Belle/Belle-II Data Handling system
   - Summary & Plan
2. $B^+ \rightarrow \rho^0 K^{*+}$ decay and research progress
   - The motivation of $B^+ \rightarrow \rho^0 K^{*+}$ decay
   - Strategy
   - Future plan

- There is the Data Handling system based on postgreSQL, PHP and HTTP in Belle
  - To find the data, the Belle users use the Belle File Searching Engine.
  - check_process_url is used for analysis based on postgreSQL.
  - The skim tool is in basf.
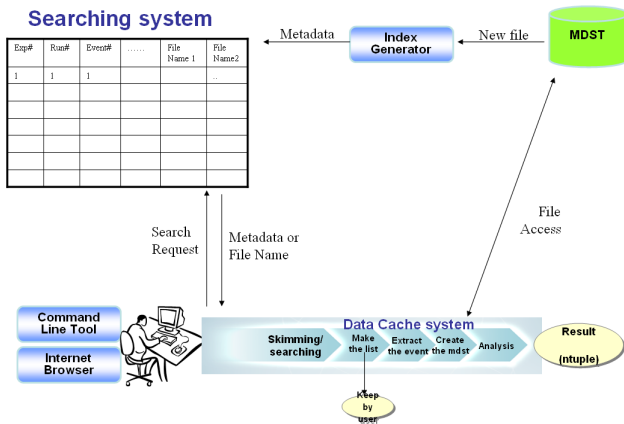 - There has been already the Data Handling system in Belle.

- However, there are some problems for Belle II.
  - The postgresSQL do not have the capacity so that it may control Belle data size $\times$ 50-60.
  - We consider that the system should be adapted to GRID.
  - We propose AMGA to solve the problems

- Belle File Searching Engine
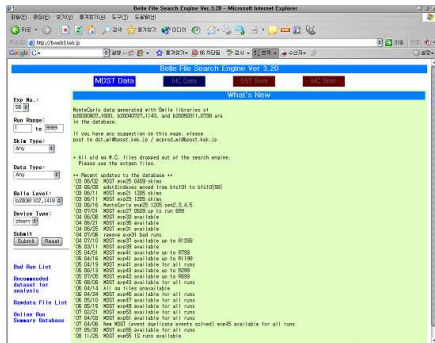
Data Handling Scenario

- We should construct the meta-data catalog of MC/data for Belle II.
- Belle II users search the data/MC from the meta-data catalog.
- We need the command line and Internet browser.

Starting Point

- We started to make the meta-data catalog using Belle.
- We will move the system to Belle II framework.
- Belle library:

    Belle Library : b20090127_0910
- Configuration of the AMGA : 150.183.250.215
    - It is based on the postgreSQL.
- We extracted the parameters(attributes) from Belle library.

To extract the attributes from the Belle data

- Extraction Interface : belle_amga_extract LFN filename
    - Based on Belle framework
    - Used on Belle data
- Output:

    mcprod-on_resonance-evtgen-uds-00-all-e000007r002865-b20030807_1600.list
    : for Logical filename

    mcprod-on_resonance-evtgen-uds-00-all-e000007r002865-b20030807_1600.log
   - file level: 1) exp number, 2) run number, 3) event number
    - event level: 1) # of $+$ charged track, 2) # of $-$ charged track, 3) # of K_S, 4) # of K_L, 5) R2
- Flexible to attributes suggested by collaboration

The architecture in AMGA

- AMGA support the directory structure.
- The metadata sheets can be seperated with the directory structures.
- Scalability will be imppoved in AMGA

AMGA: support directory structure
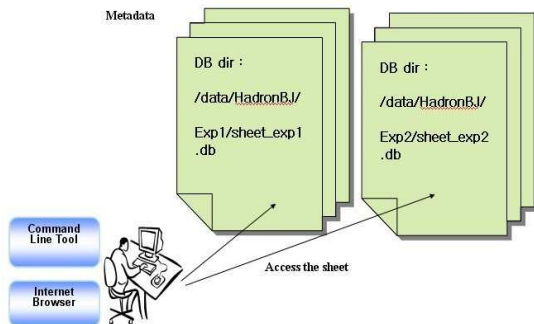
The definition of the attributes for both MC and real

- ex) sheet_real_expXX.db & sheet_MC_expXX.db

- Logical file location
- skim type : HardronBJ, fullrec, and so on
- stream
- exp number
- event number
- start run
- end run
- data type: on_resonance, off_resonance, $\Upsilon(5S)$ and so on
- type: evtgen-uds, evtgen-charm, evtgen-charged, evtgen-mixed, real
- library version

The definition of the additional attributes for events

- This is a new idea for Data Handling. $\rightarrow$ by J.H Kim
  ; We publish the abstract at Computing Physics Conference.

- number of + charged tracks
- number of - charged tracks
- number of K_S^0s
- number of K_L^0s
- R2 : 0.0 - 1.0
- the request of the sub analysis group

- The attributes can compose flexible by users' requests.

- Data
    - The biggest skimming data contain all information for events.
    - The others skimming data are the subsets of the biggest one.
    - Therefore, we don't need to make additional sheets for all skimming.
    - To get the additional information, the meta-data size will be increased a little bit.

- MC
    - We expect the massive size for MC, but we can improve the scalability using hierarchical table structure and replicaton of meta-data in AMGA.

- Command Line Interface
    - belle_amga_access (...)

- Programming API
    - belle_amga_connect (host,port,dir)
    - belle_amga_search (condition)
    - belle_amga_eot ()
    - belle_amga_fetch (variable)
    - belle_amga_write (...)
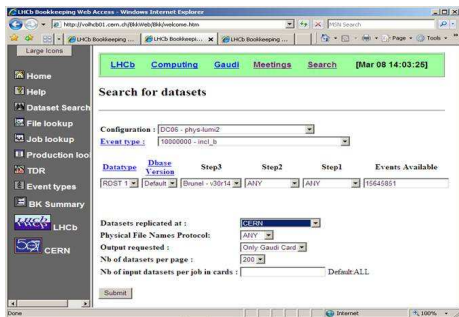   - belle_amga_close ()

- Web Interface:

Figure 1

- Command Line Interface
    - belle_amga_access:
      [amgadev]/home/jhkim > ./belle_amga_access
      Usage : belle_access_amga [options] exp_type stream_or_skim exp_number

−h	help message	
−d	debug	
−s	amga_server	
−p	port	
− − start_run	#start_run	
− − sr	#start_run	
− − end_run	#end_run	
− − er	#end_run	
− − data_type	on_resonance	off_resonance	
− − dt	on_resonance	off_resonance	
− − type	uds |charm |charged	mixed	
− − lib_version	#version	
− − lv	#version	

EX) belle_access_amga − −start_run 1000 − −end_run 1200 − −type uds m 0 7

The usage in command line :

- What is benefit to use it?
    - We can control the event level in the meta-data catalog.
    - We can choose either the file level searching or the events level searching alternatively.
      If we don't describe for event level, the answer will be file level.
    - We can use it at remote network with strong security (Grid-Proxy certificates, VOMS).
    - The command tool have simple question for user's convenience.
      We don't need to describe as "any" or "legacy" of check_process_url.
    - We can use it based on Grid.

The usage in command line; under control in file level

- Server:150.183.250.215, client:150.183.250.215
- We ask the question without the event level.

  [amgadev] /home/jhkim > ./belle_amga_access –start_run 1000 –end_run 1002 –type uds m 0 7

  process_event bfss06:/bdata/mcprod/dat/e000007/evtgen/uds/00/all/0807/on_resonance/10/ evtgen-uds-00-all-e000007r001000-b20030807_1600.mdst 0

  process_event bfss06:/bdata/mcprod/dat/e000007/evtgen/uds/00/all/0807/on_resonance/10/ evtgen-uds-00-all-e000007r001001-b20030807_1600.mdst 0

  process_event bfss06:/bdata/mcprod/dat/e000007/evtgen/uds/00/all/0807/on_resonance/10/ evtgen-uds-00-all-e000007r001002-b20030807_1600.mdst 0

  [amgadev] /home/jhkim >

- We ask for the question by adding options..

  [amgadev] /home/jhkim > ./belle_amga_access –start_run 100 –end_run 100 –type uds –ks=5 –kl=4
  –r2==0.5 m 0 7
  Exp 7 Run 2865 Events 91
  Total number of events 91
  7 2865 0 25
  7 2865 0 120
  7 2865 0 127
  7 2865 0 147
  7 2865 0 169
  7 2865 0 252
  7 2865 0 262
  7 2865 0 314

- "mkdb" is not working well for MC.
- We are developing the skimming tool without database.
  $\rightarrow$ We will contact Itoh san for this development.

The optimization of the meta-data

- We suppose that the experiment is from 07 to 55.
- We suppose that there are 10 streams for MC.
- The data type is uds, charm, charged and mixed.
- There are 30 kind of skimming type.
- The total data size of Belle II will be $\times 60$ than that of Belle.
- We consider only the on_resonance because the off_resonance is very small.

Table: Reference

Sapce Occupation per file in DB	600bytes
Average number of events in a file	111,190
Sapce Occupation per event in DB	12bytes
Multiples in Belle II	60

Table: The evaluation of meta-data

	# of files	Size for file level	Size for event level	Size in Belle II for events
number of run in Belle	24,000	14 MB	125 GB	
number of skim types	30			
total number of real files	720,000	412MB		1.8TB
number of MC streams	10			
total number of MC files	240,000	137MB	1,988GB	
number of MC skim types	30			
total number of MC files	7,200,000	4120MB		17.4TB

- We composed the value of the parameters as the bin for the event-level.
  - ex) R2 : 0.0-0.1, 0.1-0.2, 0.2-0.3 .... 0.9-1.0
    number of + charged tracks : 1, 2, 3, 4, 5, 6, 6 < # of track.
    number of - charged tracks : 1, 2, 3, 4, 5, 6, 6 < # of track.
  We compose the 10 bins for R2 value.
- We can reduce about 85% of the meta-data size from the way

The data cache system.

- At the beginning of the Belle experiment, every skimmed data is stored in "mdst" format.
- However, These files occupied so much disk space, Then, We have started suffering from the lack of disk space...
- Index file has basically information of the exp #/run #/ evt # of events passed the selection.
- Therefore the file size can be reduced much.
- From exp #/run #/(etc) written in "index file", "basf" can specify the location of the file body by using DB(Belle), in where this information has been resistered in advance. Then "basf" can access(random access) the file body.

- We developed the the simple data cache system. → The first version takes so so longer time than that of Belle.
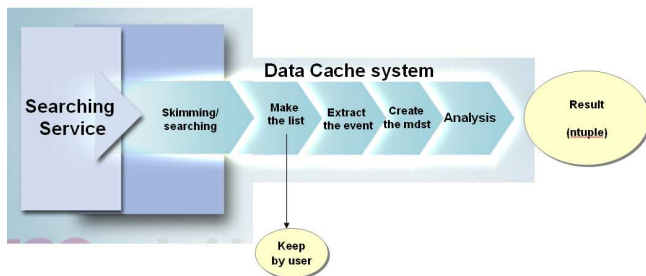
	Belle	Belle II
type	PostgreSQL	text
size	8byte/event	3.7byte/event
Grid	X	O

- We considered the sites, KISTI(master) and Melbourne(slave), for AMGA system.

    - Melbourne-KISTI cooperated to make the master-slave for the replication of the meta-data catalog.
    - Master node : 150.183.246.196(KISTI)
    - Slave node: 192.231.127.47(Melbourne)
    - We use the mata-data suggested by the computing group.
    - The meta-data catalog is the file-level.
    - The meta-data catalog was composed of exp7(on_resonance{uds,charm,charged,mixed}, stream 00).

- It was successful for the replication.
- We wrote the document for the meta-data catalog.
  http://b2comp.kek.jp/ twiki/bin/view/Computing/DataHandling

Releasing the command tool

- We released the first version of the command tool.
  - The command tool is such as check_process_url.
  - It is based on AMGA client-2.0.
  - We need to decide on whether standalone or dependence on glite.
  - We need to discuss what kind of the security is allowed.
  - We found a minor problem:ex) dummy message from AMGA.
- We will improve the command tool with the collaborations suggestions.

Summary and Next step

1. The replication from KISTI to Melbourne worked well.
2. We released the document for DH with AMGA.
3. We released the first version of the command tool.
4. We reduced about 85% for the meta-data size.
5. We are making the skimming tool without the database.
6. We receive CCP 2009 acceptance notice for poster presentation.
   "The advandced data searching system with AMGA at the Belle II experiment"
   the abstract has been accepted for inclusion in the Conference Abstract Book.

- LFC system will be developed by Melbourne group.
- We consider the central DH system;
  - on-line: file-level and manual update
  - off-line: event-level and manual update→ our idea!!
  - end-user: file-level and automatic update.

- The helicity measurement of B meson have a relation with beyond SM.
  - $B^+ \rightarrow \rho^0 K^*(892)^+$ decay; $B \rightarrow VV$
  - From $B \rightarrow VV$ decays, the angular correlation measurements are sensitive to beyond SM.
  - $A_0 >> A_T$ in SM, where A_0 is the longitudinal, A_T is transverse for polarization fraction.; Belle and BaBar measured $A_0 \sim A_T$
- There are Tree and Penguin phase in $B^+ \rightarrow \rho^0 K^*(892)^+$ decay; Direct CP measurement.

- $B \rightarrow VV$ decays

Decays	$\mathfrak{B}(10^{-6})$	Exp	Year
$B^+ \rightarrow \omega K^+$	6.7 ± 0.8	BaBar	07
$B^+ \rightarrow \omega K^{*+}$	< 3.4	BaBar	06
$B^+ \rightarrow \rho^0 K^+$	4.2 ± 0.5	Belle	06
$B^+ \rightarrow \rho^0 K^{*+}$	< 6.1	BaBar	06
$B^+ \rightarrow \rho^+ K^0$	8.0 ± 1.5	BaBar	07
$B^+ \rightarrow \rho^+ K^{*0}$	9.2 ± 1.5	BaBar	06
$B^+ \rightarrow \phi K^{*+}$	10.5 ± 1.5	BaBar	07
$B^0 \rightarrow \omega K^0$	5.0 ± 0.6	BaBar	07
$B^0 \rightarrow \omega K^{*0}$	1.8 ± 0.7	Belle	08
$B^0 \rightarrow \rho^0 K^0$	5.4 ± 0.9	Belle	07
$B^0 \rightarrow \rho^0 K^{*0}$	5.6 ± 1.6	BaBar	06
$B^0 \rightarrow \rho^- K^+$	< 12.0	BaBar	06
$B^0 \rightarrow \rho^- K^{*+}$	9.2 ± 1.5	Belle	04
$B^0 \rightarrow \phi K^{*0}$	9.5 ± 0.8	BaBar	07
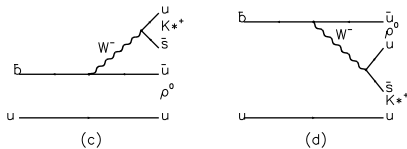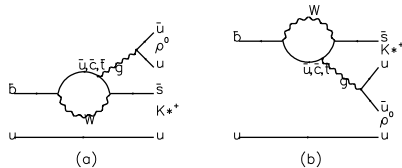
Feynman Diagram of $B^+ \rightarrow \rho^0 K^*(892)^+$ decay

- $B^+ \to \rho^0 K^{*+}$
    $\rho^0 \to \pi^+ \pi^-$
        $K^{*+} \to K_S^0 \pi^+$
        $K^{*+} \to K \pi^0$

- *BaBar*(06) found the upper limit $< 6.1 \times 10^{-6}$
- We consider to search the decay mode.
- We will make the model of $B^+ \to \rho^0 K^{*+}$ with 4-dim extended likelihood fit.

- Signal MC with considering the longitudinal for polarization fraction; done
- Using the $q\bar{q}$ MC of real data $\times 3$, $b \to c$ MC of real data $\times 6$, rareB MC of real data $\times 50$;
  skimming MC(from exp07 to exp65: 800M events):doing

- Particle Identification
- $q\bar{q}$ study with KSFW, $cos\theta^B*$, ΔZ and $cos\theta_H$
- B background & Veto
- PDF modeling & Fitter check
- Control sample
- To estimate the peak
- To estimate the branch fracion: if we can observe (O,X)
- If we have the enough yields, to measure the direct CP

Future plan

1. $B \rightarrow \phi\pi$ will be published with "World best limit" nearby future.
   - I keep the topic at Belle II, until the measurement.
   - I will try to study the Dalitz analysis and upgrade $2D \rightarrow 3D$ fit.

2. $B^+ \rightarrow \rho^0 K^{*+}$ is studing
   - 1st goal is to measure the branching fraction.
   - 2nd goal is to measure the CP.
   - 3rd goal is to measure the polarization fraction.

3. I have a plan to study $B \rightarrow \phi\rho$:
   The behavior is same as $B \rightarrow \phi\pi$ and is more suppressed decay.
   Also, The mode is $B \rightarrow VV$ decays