

# AMGA Metadata Catalog for Belle II

J.H Kim<sup>1</sup> & D. J Kong<sup>1</sup> & K. Cho<sup>1</sup>

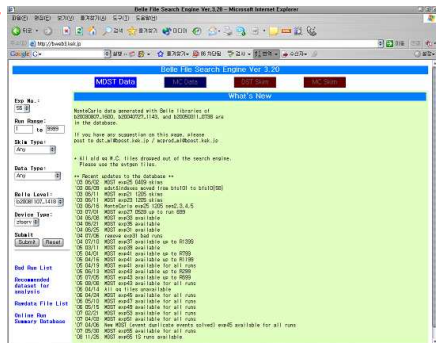
<sup>1</sup>High Energy Physics Team  
KISTI, Daejeon, Korea

Belle 2 General Meeting, 2009.07.07

## Overview

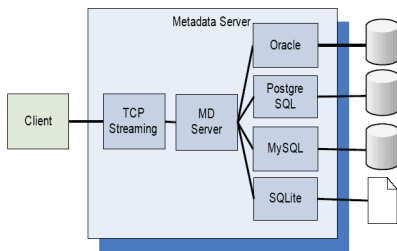
- 1 Current status in Belle
- 2 What is AMGA?
- 3 Benefit from using the AMGA in Belle II
- 4 The Data Handling scenario with AMGA
- 5 Starting
- 6 To extract the attributes from Belle data
- 7 The architecture in AMGA
- 8 The definition and control of the attributes
- 9 How to access AMGA
- 10 The Usage in command line
- 11 Summary and next 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 a Data Handling system in Belle.
- However, there is some problems for Belle II.
- The postgresSQL don't have a capacity so that it control Belle data size  $\times$  50-60.
  - We consider that the system should be adapted to GRID.
  - We proposes the way to solve the problems with AMGA.
- Belle File Searching Engine
- 
- A screenshot of the Belle File Searching Engine web interface running in a Microsoft Internet Explorer browser. The page title is "Belle File Search Engine Ver.3.20 - Microsoft Internet Explorer". The address bar shows "http://www.belle.jp". The main content area has tabs for "MDST Data", "ALL Data", "CDT Data", and "MC Data". Under the "MDST Data" tab, there's a section titled "What's New" which displays a list of recent updates to the database, including file names like "820900007\_1000\_833047027\_1143\_aid 833050111\_0708.ais" and "820900007\_1000\_833047027\_1143\_aid 833050111\_0708.ais". There are also search filters for "Exp No.", "Run Range", "Data Type", and "Data Level".



# What is AMGA?: (Reference:www.eu-egee.org)

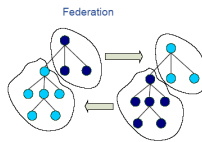
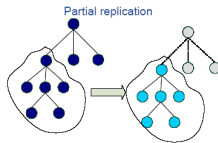
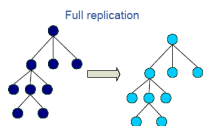
- AMGA is the Meta-data catalog of EGEE's gLite 3.1 Middle-ware.
  - Meta-data is relationally structured data for Grid jobs stored in databases.
- The AMGA service provides access to relational Data bases on the Grid, taking into account:
- The AMGA features:



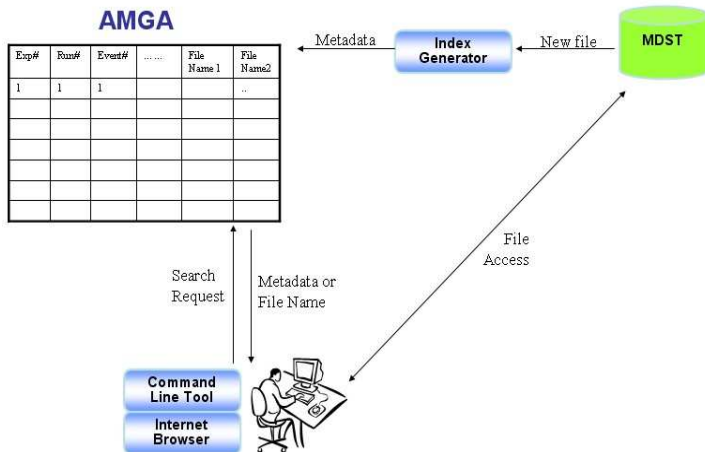
- Authentication(Grid-Proxy certificates, VOMS)
- Logging, tracing
- DB connection pooling
- Replication of Data
- Use of hierarchical table structure
- ..the Grid idea.

# Benefit from using the AMGA in Belle II.

- ❶ AMGA can improve the scalability for the Data Handling system.
  - AMGA integrates replication of meta-data.
    - Asynchronous replication
    - DBs are consistent(transactions supported)
    - However, Not all DBs necessarily in same state.
  - Replication makes use of hierarchical table structure.
  - Good Performance and Scalability.
- ❷ Belle II will support the Grid → the AMGA is included in [grid middle-ware,gLite 3.1](#).
- ❸ We can make the new Data Handling system of Belle II in Grid.
  - AMGA replication makes use of hierarchical concept:



# The Data Handling Scenario with AMGA



- 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.

- We started to make the meta-data catalog with Belle.
- We will move the system to Belle II framework.
- Belle lib:  
    Belle Lib : b20090127\_0910
- Configuration of the AMGA : kenobi.kisti.re.kr
  - It is based on the postgresSQL.
- We extracted the parameters(attributes) from Belle lib.

## 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

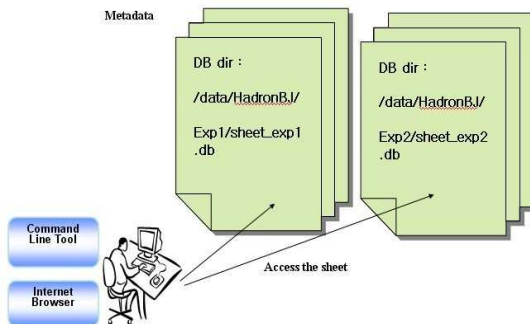
`mcprod-on_resonance-evtgen-uds-00-all-e000007r002865-b20030807_1600.log`

- 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 separated with the directory structures.
- Scalability will be improved 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
- type: `evtgen-uds`, `evtgen-charm`, `evtgen-charged`, `evtgen-mixed`, `real`
- lib version

## The definition of the additional attributes for events

- number of + charged tracks
- number of - charged tracks
- number of  $K_S^0$ s
- number of  $K_L^0$ s
- R2 : 0.0 - 1.0

- 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.

# How to access AMGA: made by AMGA Team(Mr. SunIl An)

## • 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:

The screenshot shows the LHCb Bookkeeping Web Access interface. The browser address bar displays the URL: `http://volhc01.cern.ch/llweb/llweb/welcome.htm`. The page has a green header bar with navigation links: LHCb, Computing, Gaudi, Meetings, Search, and a timestamp [Mar 08 14:03:26]. Below the header, there is a section titled "Search for datasets". The search configuration includes a dropdown for "Configuration" (set to "DC05 - phys-lum2") and a text input for "Event type" (set to "10000000 - incl\_b"). Below this is a table with columns: Datatype, Phase Version, Step3, Step2, Step1, and Events Available. The table contains one row of data: ROST 1, Default, Brunel - v30r14, ANY, ANY, and 1564581. Below the table, there are several form fields: "Datasets replicated at:" (set to CERN), "Physical File Names Protocol:" (set to ANY), "Output requested:" (set to Only Gaudi Card), "Nb of datasets per page:" (set to 200), and "Nb of input datasets per job in cards:" (set to Default:ALL). A "Submit" button is located at the bottom of the form.

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

## Case 1st: The usage in command line; under control in file level

- Server:150.183.250.215, client:150.183.250.215

- Output : belle\_amga\_access:default [same with Belle]

```
[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 >
```

- If you want to use it, send to e-mail([hyun@kisti.re.kr](mailto:hyun@kisti.re.kr);[siahn@kisti.re.kr](mailto:siahn@kisti.re.kr)).

## Case 2nd: The usage in command line; under control in event level

- We can ask more property of events.
- Output : `belle_amga_access:optional [additional attributes]`

```
[amgadev] /home/jhkim > ./belle_amga_access -start_run 1000 -end_run 1000 -type uds -ks=5 -kl=4  
-r2==0.5 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 7041
```

```
process_event bfss06:/bdata/mcprod/dat/e000007/evtgen/uds/00/all/0807/on_resonance/10/  
evtgen-uds-00-all-e000007r001000-b20030807_1600.mdst 2796
```

```
process_event bfss06:/bdata/mcprod/dat/e000007/evtgen/uds/00/all/0807/on_resonance/10/  
evtgen-uds-00-all-e000007r001000-b20030807_1600.mdst 5582
```

```
process_event bfss06:/bdata/mcprod/dat/e000007/evtgen/uds/00/all/0807/on_resonance/10/  
evtgen-uds-00-all-e000007r001000-b20030807_1600.mdst 4075
```

```
process_event bfss06:/bdata/mcprod/dat/e000007/evtgen/uds/00/all/0807/on_resonance/10/  
evtgen-uds-00-all-e000007r001000-b20030807_1600.mdst 514
```

```
process_event bfss06:/bdata/mcprod/dat/e000007/evtgen/uds/00/all/0807/on_resonance/10/  
evtgen-uds-00-all-e000007r001000-b20030807_1600.mdst 8747
```

- We should understand the File format conversion of the BASF.
- File format conversion : `bfcf`, `dblist`, `dbls`, `mkdb`
- We will control event I/O within the BASF.

- 1 Defined the attributes
- 2 constructed the architecture in AMGA
- 3 We made the meta-data based on Belle.
- 4 Is there any missing for the usage?
- 5 We will make fully the meta-data for Belle to test the performance.
- 6 We will apply the LFC to find the best location for GRID-[KEK-KISTI].
- 7 We will serve it for the purpose of the test and evaluation at KISTI.  
We should discuss to service with each team in KISTI.

## Acknowledgement

- e-Science Grid technology Development Team:
  - AMGA software development
  - Technical Support
  - Tutorial
  - sun-Il Ahn, Soonwook Hwang

