

## Current status of the activity on the BK

Participants: Philippe, Zoltan, Elisa

### Synchronization of the 2 DB

The procedure is the following:

- first, we will start to feed in parallel the 2 DB. A script will catch the XML files sent to the volhcb05 and will translate them in the new format (the format to insert data in the new BK). This XML file will be sent to the volhcb07. Then, the BK Agent will be switched on and the data should be inserted into the new schema. During these operations, the BKManager on volhcb05 will always be on.

This has been done. The 2 schemas have been populated in parallel since (+ o -) 10 July 08

In order to have the migration completed, it is necessary also to migrate the productions altered after February 2008.

The following productions have been altered after February:

- 1856 (now finished) all the jobs are in the new schema (1658 jobs and 4974 files) ok
- 1896 migrated on July 24. (All the jobs and files have been copied, but the logfile is full of errors: multiple input files issue...).
- 1932
- 1959

Philippe should provide a list of the productions that are still ongoing. The productions still running will need some special checks in order to ensure that no file/job has been lost during the process of synchronization.

Mon Aug 18 2008: the procedure I'll follow will be:

- 1) identify the new data inserted after Feb 1 2008 (they are about 273000 jobs spread over 103 productions) No! the list of files has to be retrieved on the basis of the JOBID, not the date, because some old jobs (of 2006) have just been inserted in the last months after Feb 2008. Done
  - 2) delete these productions from the AMGA tables in the new schema. Done
  - 3) copy them again from the production DB to the AMGA tables in the new schema  
Finished at 17:50 Aug 20. A lot of errors due to the problem of multiple input files. Done
  - 4) correct again the production number for prods 201 and 200. This should be done on the Oracle tables, after that Zoltan has copied again them from the AMGA tables.
  - 5) recover again the files and jobs with null prod id that have been deleted. Prods:00001354, 00001501, 00001503. Done
- Done by 28 Aug 08

### Removal of bad productions

Philippe notices that: this list of event types: [11102003, 11102013, 11102402, 11104103, 11144103, 13102002, 13102013] in phys-lumi2 and phys-lumi5 have a bug and should be removed. Some of them are 1501, 1512, 1514, 1601 (lumi2) and 1523, 1526, 1628 and 1610 (lumi5). But Marianne probably knows the whole list. Marianne said that she will provide the list in the end of August. Here are the productions which correspond to these config versions and event type: "" "00001502" "00001512" "00001514" "00001523" "00001526" "00001528" "00001601" "00001610"

More in general, there a list (see **badProductions.txt** file in the attachment) that contains all the productions which have to be ignored (for different reasons). I keep this list updated every time that it comes out that a production is bad.

Done. 25 Aug: Provided to Zoltan the list of bad prods

### Post synchronization checks

Some consistency checks after the synchronization will be done, to ensure that we get the same results from the 2 (old and new) schemas.

This has to be done, but after the item **Synchronization of the 2 DB**, see above.

### BKK client documentation

In this twiki a new link has been added with the documentation of the Bkk Client methods developed by Zoltan.

This has to be done.

### How to deal with the 2000 and 2001 productions

These 2 productions present a problem for the computation of the processing pass because 2 different versions of the application have been used (Brunelv30 and Brunelv31) within THE SAME production.

The following solution has been proposed: to alter the *production* field in the *jobs* table for those jobs which have one programVersion (i.e. v30) and set it to a new value (of course taking care of using a production number that will never be used!). The other jobs, those which have the other version (i.e. v31), will be left unchanged.

Proposal for the algorithm:

if production==2000 and programVersion==v30 then set production to 200

if production==2001 and programVersion==v30 then set production to 201

(action for Elisa).

This has been fixed with the following SQL statements (1 July 08):

```
update jobs set Production=201 where (jobs.production=2001 and jobs.programVersion='v30r17' and j
```

```
update jobs set Production=200 where (jobs.production=2000 and jobs.programVersion='v30r17' and j
```

This has been fixed on 1 July 08

Later it was done also for the AMGA tables in the new schema with the statement:

```
update dir3 set "user:Production"=201
```

```
where (dir3."user:Production"=2001 and dir3."user:ProgramVersion"='v30r17' and dir3."user:Program
```

```
update dir3 set "user:Production"=200
```

```
where (dir3."user:Production"=2000 and dir3."user:ProgramVersion"='v30r17' and dir3."user:Program
```

This has been done on 24 Sept 08

### Populating the Simulation Conditions table for the DC06 data

Finally, the attributes of the *simulation conditions* table are:

SIMID NUMBER

SIMDESCRIPTION VARCHAR2(256)

BEAMCOND VARCHAR2(256)

BEAMENERGY VARCHAR2(256)

GENERATOR VARCHAR2(256)

MAGNETICFIELD VARCHAR2(256)

DETECTORCOND VARCHAR2(256)

## LUMINOSITY VARCHAR2(256)

and the table has been populated with 15 rows, defining all the possible simulation conditions of DC06 simulated data.

This has been done

Now the attribute DAQId in the jobs table has to be updated with the corresponding *SimId*.

This has been done on July 24

For more details see the twiki page for the simulation conditions.

Some checks to do:

we have to check that a job has the same simulation conditions of the input jobs.

This has to be done (Elisa).

### **Problem of multiple input files**

Data have been migrated to the new schema and here we will have to fix this problem. Philippe should provide some ascii files containing the information to delete some of the entries in the `_inputFiles_tables` of the problematic productions.

**This has to be fixed**

### **Files with no entry in jobs table**

During the migration of the files with null production number, some files have been found with no entry in the jobs table. The files are 29 in total, spread over 5 productions ( 00001323, 00001324, 00001325, 00001326, 00001327). It has been decided to ignore these files, since we don't have the necessary information about them to register them in the new schema, and also because they are only few files.

This is not an issue, just a warning so I put it in green 😊

### **Productions with ONLY SIM files**

During the computation of the processing files, some productions have been found to have only SIM files. This is not foreseen in the normal production workflow. The files of these production do not appear in the *InputFiles* table, this means that there's no job in the BK which has taken in input those files. Then we have 3 possible cases:

- 1- The jobs and their output (DIGI) files which have processed these SIM files have been removed from the BK
- 2- Or they have never been registered in the BK
- 3- Or these SIM files were not interesting and were not processed.

the productions affected by this problem are 92: 1329, 1330, 1331 1332 1333 1334 1335 1336 1358 1359 1360 1361 1362 1363 1364 1365 1375 1376 1377 1378 1384 1385 1386 1387 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1462 1463 1470 1471 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496

These productions are done with a buggy version of brunel. They can safely be removed. I have added them to the list in the badProductions.txt file.

On 5 of Aug 08 it has been decided that: we remove them from the oracle tables in the new schema. For completeness sake, we keep them in the AMGA tables in the new schema. In the computation of the processing pass these prods are skipped.

### **Modifications of the schema**

The new schema has been modified. A new table has been created: *CONFIGURATIONS* with the *ConfigName* and *ConfigVersion*. As a consequence, all the stored procedures have to be modified.

- stored procedure to compute and insert the simulation conditions: updated
- stored procedure to compute the processing pass: updated (Aug 5 2008)

### Computation of Processing Pass

It has been decided to restructure the *processing\_pass* table and *pass\_index* table and to add a *PASS\_GROUP* table (Mon Sept 1 2008).

Moreover, after the meeting of Sept 22 some modification has been applied to the *PROCESSING\_PASS* table:  
**1** - the total processing pass includes all the steps, from the simulation. So we have to apply the following changes: for the processing pass DC06-Stripping\_v31 + DC06-Stripping\_v30 (this affects only productions 200 and 201)

```
update processing_pass
set totalprocpass='DC06-Sim + DC06-Reco_v30 + DC06-Stripping_v31 + DC06-Stripping_v30' where tota
```

and the proc pass DC06-Reco\_v30 + DC06-Stripping\_v31 (it affects only prod 2000):

```
update processing_pass set totalprocpass='DC06-Sim + DC06-Reco_v30 + DC06-Stripping_v31' where to
```

Done 24 Sept 08

**2** - The totalprocpass attribute in the *PROCESSING\_PASS* table will be expressed on the basis of the numeric group ID instead of the group descriptions strings. For example the entry: 'DC06-Sim + DC06-Reco\_v30 + DC06-Stripping\_v31 + DC06-Stripping\_v30' will be replaced by: '1<3<5<7'.

This table shows the total processing pass old and new format:

Old Format	New Format
DC06-Sim	1
DC06-Sim-Reco_v30	2
DC06-Sim-Reco_v31	9
DC06-Sim + DC06-Reco_v30	1<3
DC06-Sim + DC06-Reco_v32	1<6
DC06-Sim + DC06-Recon-L0-v1-lumi2	1<8
DC06-Sim-Reco_v30 + DC06-Stripping_v31	2<5
DC06-Sim-Reco_v31 + DC06-Stripping_v31	9<5
DC06-Sim + DC06-Reco_v30 + DC06-Stripping_v31	1<3<5
DC06-Sim + DC06-Reco_v30 + DC06-Stripping_v31 + DC06-Stripping_v30	1<3<5<7

Done 25 Sept 08

**3** - The simulation conditions ID will be added as a new column in the *PROCESSING\_PASS* table.

Done 29 Sept 08

More details are given in the processing pass page.

-- ElisaLanciotti - 23 Jun 2008

This topic: LHCb > Meeting080623

Topic revision: r36 - 2011-02-15 - ZoltanMathe1



Copyright &© 2008-2019 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback