Offline computation of the Processing Pass

The Processing Pass (PP) is the set of application version and condition DB tags which have been used to produce a file and its ancestors.

The processing pass is common for all the files of the same production.

By consequence, a processing pass corresponds to a set of productions. A complementary table has been introduced in the new schema, with the information relative to each production, its processing pass and, if it is the case, the processing pass of the input productions.

Computation of the Processing Pass during the migration of DC06 data

The computation of the Processing Pass for already existing data has been done on the basis of the information available in the BK.

- Make a first loop on the DC06 productions (about 600 prods.). For each of them compute the Pass Index. This is defined as the set of application name and version and cond db tag used in the production (and only in that production, regardless of the input data!).

In the current production system, many different work flow can be followed to produce files. See the picture *Production workflow diagram* for a diagram of the possible work flows and a classification of the production types in there.

- Set the Processing Pass Group on the basis of the CONFIGVERSION and of the event type, and following some additional indications provided by Philippe (this is some oral culture that is not possible to extract from the BK database or from any other source). Merging this information, finally we managed to assign a PP to every production following these rules.
- Store the Pass Index in the *Pass_Index* table. More in detail, the content of the *Pass_Index* table is:

PASSID NUMBER
DESCRIPTION VARCHAR2(50)
GROUPID NUMBER
STEP0 VARCHAR2(50)
STEP1 VARCHAR2(50)
STEP2 VARCHAR2(50)
STEP3 VARCHAR2(50)
STEP4 VARCHAR2(50)
STEP5 VARCHAR2(50)
STEP6 VARCHAR2(50)

the content of the *Pass_index* table is shown in the attached file *pass_index.html* below. As the table shows, many different Pass Indexes belong to the same group. This means that they can be mixed at processing stage.

- Add a new line in the *Prod2Pass* table with: production number and *PassId*. This table is simply a dictionary to associate to each production its Pass Index.

Once finished the loop, the *Pass_index* and the *prod2pass* tables have been populated. In these tables we have stored the information about the production itself, but not about the input data, so the story of the files is not complete. What we need now is a further table to store the Processing Pass of the production and also of the input production: we will call it the Total Processing Pass and it will be stored in the *Processing_Pass* table.

- Make a second loop on the DC06 productions. For each production go to the *Prod2Pass* table and pick up the corresponding *PassId*, then go to the *Pass_Index* table and retrieve the information about the Pass Index and the Group. If the production has any input data, repeat the same for the input production. Compute the Total Processing Pass as a concatenation of the Processing Pass (this is the Group Description field) of the input production and the one of the current production. Of course, if there is no input, the Total Processing Pass is the same as the Group Description. Store all this info in the *Processing_Pass* table. See the attached file *processing_pass.html* which displays the content of this table. In the meeting of 2/07/08 we said that it would be useful to add the simulation conditions. (It has been added now).

- Finally, all the available Processing Passes are obtained making a query like: select distinct processing_pass.totalprocpass from processing_pass. This is the list of available processing passes that will be shown to users to make their query.

Some remarks about the computation of the processing pass

Special cases:

- Production 1959

The production 1959: this is the only production which runs Brunel on DIGI files to produce DST files. It is taken into account in the creation of the processing pass (methods: *populatePass_Index@procPass.py*) This is not a problem, just a remark

- Productions 2000 and 2001

these productions are a special case because in the same production different versions of the application have been used. We have changed the production number for those jobs with Brunel v30. This is fixed

- Productions with only SIM files

There are many productions which only have SIM files. Some of them are: 1329, 1330, 1331, 1332, 1333, 1334, 1335, 1336,1358, 1359, 1360, 1361, 1362, 1363, 1364, 1365, 1368. They are leftovers of buggy productions. To be ignored.

-- ElisaLanciotti - 14 Oct 2008

This topic: LHCb > OfflinePPass
Topic revision: r3 - 2008-10-15 - ElisaLanciotti