

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

Working Group productions: mini tutorial

Antonio Falabella

INFN - CNAF (Bologna)
9th LHCb Computing Workshop

May 16, 2017

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

- 1 Working group centralized productions
- 2 LHCbDirac Production Concepts
- 3 Step creation
- 4 New Request setup
- 5 Summary

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

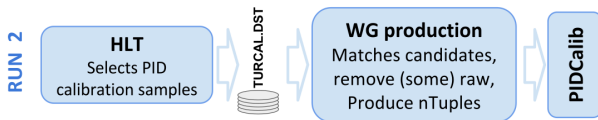
Step creation

New Request setup

Summary

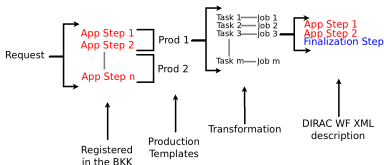
- A problem faced by the LHCbDirac developers in 2013 to handle different workflows with respect to the usual Processing/Stripping/MC ones
- In practice the main use case was *Swimming* and a few working groups
- In general a typical use case for WG includes some of the following requirements:
 - Data creation(i.e. Ntuples) where the code used by the WG is quite stable but require several releases for a data taking period
 - the number of files to be processed is too high for a single user
 - Data need to be registered in the bkk

- Additional development has been triggered by the PID group
 - For Run 2 the PID WG has developed a Stripping-like step through the TurboCalib stream



- Given the high number of files to handle we profited of Centralized WG Productions
- The output consists of μ DST and ntuples
- The internal note can be found here:
<https://cds.cern.ch/record/2161754>
- The work for PID triggered also the implementation of ntuple merger

- LHCbDirac is a DIRAC extension to support LHCb specific workflows.



Some definition:

- Request*: A request is a set of steps setup up by the user.
- Step*: A step is the definition of an execution task. It requires the definition of input data types, output data types, and the application.
- Production*: The *production template* creates your productions starting from the request. The system create *tasks* an abstract concept for a job.
- BKK query*: The data on which the workflow will run.

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

LHCbDirac submits to the Grid jobs executing LHCb code. So in order to run the code of a working group you have to:

- Prepare the gaudi configurables and put them under control version in an existing package or in a new one;
- Ask explicitly for Grid deployment when you release your code (code must be in cvmfs)
- Be aware of the correct environment on which your code will run. Basically LHCbDirac uses `lbrun` to create your particular environment and run your application through `gaudirun.py`

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

Example of a production MC job run:

- ```

lb-run --use-grid -c x86_64-slc6-gcc49-opt --use="AppConfig
v3r304" --use="ProdConf"Boole/v30r2 gaudirun.py -T
$APPCONFIGOPTS/Boole/Default.py
$APPCONFIGOPTS/Boole/EnableSpillover.py
$APPCONFIGOPTS/Boole/DataType-2015.py
$APPCONFIGOPTS/Boole/Boole-SetOdinRndTrigger.py
$APPCONFIGOPTS/Persistency/CompressionZLIB1.py
prodConf_Boole_00061064_00012093_1.py

```
- ```

--use="ProdConf" : Enable the use of the ProdConf configuration
package (See in the next slides)

```
- ```

prodConf_Boole_00061064_00012093_1.py : ProdConf configuration file
generated automatically

```

Working Group  
productions: mini  
tutorial

Antonio Falabella

Working group  
centralized  
productions

LHCbDirac  
Production  
Concepts

Step creation

New Request setup

Summary

- ProdConf has been developed to act as an interface between LHCbDirac and LHCb applications. The package is under /DBASE/ProdConf
- The documentation about the ProdConf can be found here <http://svnweb.cern.ch/world/wsvn/lhcb/DBASE/trunk/ProdConf/doc/description.pdf>;
- The ProdConf package takes care of hiding some configurations that are used by the Gaudi application module;
- It gathers information from the definition of your application *step*;
- When testing a specific package, to run it on the Grid you should remember that some parameters are set by the ProdConf package
- A tutorial on step test has been prepared by Federico :  
<https://twiki.cern.ch/twiki/bin/view/LHCb/ProdReqPreLaunchTest>



Working Group  
 productions: mini  
 tutorial

Antonio Falabella

Working group  
 centralized  
 productions

LHCbDirac  
 Production  
 Concepts

Step creation

New Request setup

Summary

Let's suppose that you are going to run a DaVinci application. This is an example of the parameters set up automatically by ProdConf:

- ```

from ProdConf import ProdConf
ProdConf(
    TCK='0xYXY',
    NOEvents=-1,
    DDBTag='head-YXYXY',
    CondDBTag='head-YXYXY',
    AppVersion='vYrZ',
    OptionFormat='merge',
    Application='DaVinci',
    OutputFilePrefix='00017857_00000157_1',
    RunNumber=102790,
    InputFiles=['input list'],
    OutputFileTypes=['output.mdst'],
)

```

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

You can reach the DIRAC Web portal from this link:

- <https://lhcb-portal-dirac.cern.ch>
- Before accessing this page you need to load in you web browser a valid grid certificate.
- You also need the `lhcb_tech` rights. These rights can be granted by the LHCbDirac production team.

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

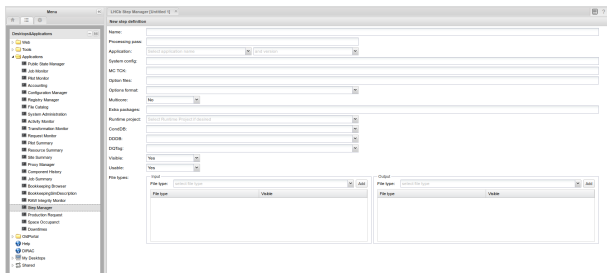
- From the Dirac web portal you can choose the Step Manager page

LHCb Step Manager [Untitled 3]

Selectors: Application: DaVinci Visible: Usable: Processing pass: Registered after (YYYYmm.dd):

	id #	Name	Processing pass	Application	Version	Visible	Usable
<input type="checkbox"/>	130467	Merge14 for Stripping21r1 B2OC F...	Merge14	DaVinci	v41r3	No	Yes
<input type="checkbox"/>	130466	Merge14 for Stripping21 B2OC Filter...	Merge14	DaVinci	v41r3	No	Yes
<input type="checkbox"/>	130465	Merge14 for Stripping26 RDWG Fil...	Merge14	DaVinci	v41r3	No	Yes
<input type="checkbox"/>	130464	Merge14 for Reconstructible RDWG...	Merge14	DaVinci	v41r3	No	Yes
<input type="checkbox"/>	130460	Stripping26-NoPrescalingFlagged fo...	Stripping26NoPrescalingF...	DaVinci	v41r2p2	Yes	Yes
<input type="checkbox"/>	130458	Turbo lines (MC), Turbo 2016 (gA) ...	Turbo03	DaVinci	v41r2p1	Yes	Not ready
<input type="checkbox"/>	130454	Merge14 for Stripping21 RDWG Fil...	Merge14	DaVinci	v41r3	No	Yes
<input type="checkbox"/>	130453	Stripping21FFiltered for RDWG (Lept...	Stripping21FFiltered	DaVinci	v30r1p4	Yes	Yes
<input type="checkbox"/>	130451	Stripping21r1p1Filtered for RDWG (...)	Stripping21r1p1Filtered	DaVinci	v30r1p1	Yes	Yes
<input type="checkbox"/>	130404	Stripping25-NoPrescalingFlagged - ...	Stripping25Flagged	DaVinci	v40r4	Yes	Yes
<input type="checkbox"/>	130397	Turbo lines (MC), Turbo 2016 - L2M...	Turbo03	DaVinci	v41r2p1	Yes	Yes
<input type="checkbox"/>	130396	Stripping24Filtered - B2OC - B2dID	Stripping24Filtered	DaVinci	v38r1p1	Yes	Yes
<input type="checkbox"/>	130395	Merge14 for Stripping24 B2OC Filter...	Merge14	DaVinci	v41r3	No	Yes
<input type="checkbox"/>	130394	Merge14 for Stripping26 B2OC Filter...	Merge14	DaVinci	v41r3	No	Yes
<input type="checkbox"/>	130393	Stripping26FFiltered - B2OC - B2dID	Stripping26FFiltered	DaVinci	v41r2	Yes	Yes
<input type="checkbox"/>	130387	Stripping25-Flagged - Sem09 - pA (fl...	Stripping25Flagged	DaVinci	v40r4	Yes	Obsolete
<input type="checkbox"/>	130386	Stripping24-NoPrescalingFlagged fo...	Stripping24NoPrescalingF...	DaVinci	v38r1p1	Yes	Yes
<input type="checkbox"/>	130384	Merge14 for Stripping21 2015 B2OC...	Merge14	DaVinci	v41r3	No	Yes
<input type="checkbox"/>	130383	Merge14 for Stripping21 2015 B2OC...	Merge14	DaVinci	v41r3	No	Yes
<input type="checkbox"/>	130382	Stripping21FFiltered for 2016 B2OC (...)	Stripping21FFiltered	DaVinci	v30r1p3	Yes	Yes
<input type="checkbox"/>	130381	Stripping21FFiltered for 2015 B2OC (...)	Stripping21FFiltered	DaVinci	v30r1p3	Yes	Yes
<input type="checkbox"/>	130380	ReconstructibleFFiltered - RDWG - M...	ReconstructibleFFiltered	DaVinci	v41r2	Yes	Yes
<input type="checkbox"/>	130379	ReconstructibleFFiltered - RDWG - M...	ReconstructibleFFiltered	DaVinci	v41r2	Yes	Yes
<input type="checkbox"/>	130378	ReconstructibleFFiltered - RDWG - M...	ReconstructibleFFiltered	DaVinci	v41r2	Yes	Yes
<input type="checkbox"/>	130377	Merge14 for Reconstructible RDWG...	Merge14	DaVinci	v42r1	No	Obsolete

- Click on **New** to be prompted on this page.



- After filling with appropriate values you save the step and it will be then ready for use

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

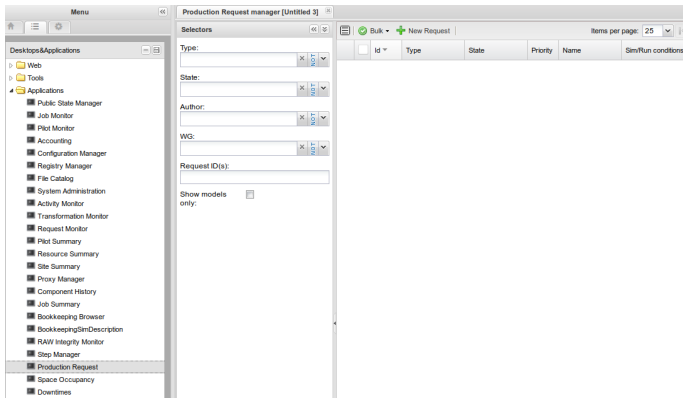
LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

- Now that you have submitted your steps you are ready to set up your production request. From the dirac web portal choose *Production* → *Request Manager as lhcb_user*.



The screenshot shows the Dirac Production Request Manager interface. On the left, a sidebar lists various desktops and applications, with 'Production Request' highlighted. The main window is titled 'Production Request manager [Untitled 3]' and contains a 'Selectors' form. The form has the following fields:

- Type: []
- State: []
- Author: []
- WG: []
- Request ID(s): []
- Show models only:

Below the form is a table with the following columns: Id, Type, State, Priority, Name, and Sim/Run conditions. The table is currently empty. The interface also includes a 'Bulk' button and a '+ New Request' button at the top right.

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

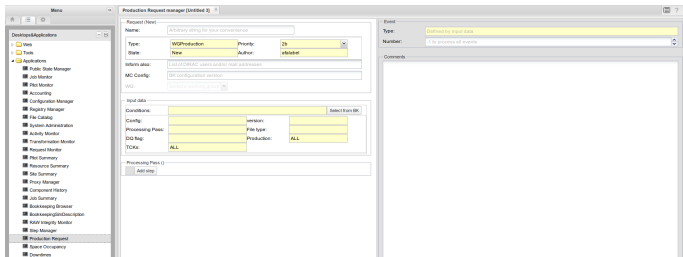
LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

- In the top left box you can define the name of your production. This will be the name that will appear in the dirac web portal.



The screenshot shows the 'Production Request Manager (DIRAC 2)' interface. The main window is titled 'Request (New)' and contains the following fields and sections:

- Name:** A text input field with the placeholder 'Arbitrary string for your convenience'.
- Type:** A dropdown menu set to 'WProduction'.
- Priority:** A dropdown menu set to '2h'.
- State:** A dropdown menu set to 'New'.
- Author:** A dropdown menu set to 'afalabel'.
- Info:** A section with the text 'List of DIRAC users and their addresses' and a table for 'MC Config'.
- MC Config:** A table with columns for 'MC Config' and 'Version'.
- Input data:** A section with a 'Select from DB' button.
- Conventions:** A section with a table for 'Config' and 'Version'.
- Processing Pools:** A section with a table for 'Pool type' and 'Production'.
- Job type:** A dropdown menu set to 'ALL'.
- Production:** A dropdown menu set to 'ALL'.
- Processing Pools:** A section with an 'Add new' button.
- Event:** A section with 'Type' set to 'Defined by input data' and 'Number' set to '1 to process all events'.
- Comments:** A large text area for adding comments.

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

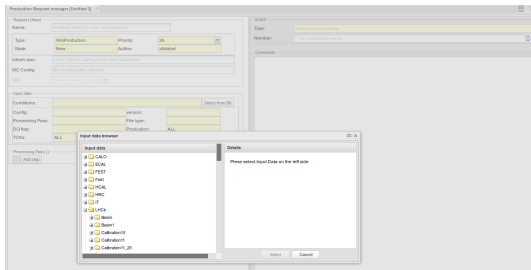
LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

- In the second box you can select the BKK query on which run your production.



- Click on *Select from BKK* a BKK selection dialog window pops up has been chosen.

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

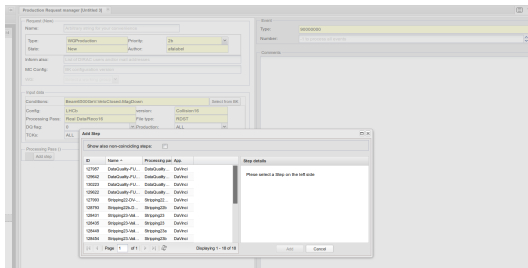
LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

- The third box on the left side is the place you can add your step/steps.
- When completed you can *Submit to the production team*



Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

- When you submit a request each production will be notified.
- If the application fails you will be notified and the production manager will give you all the details to understand what went wrong. But no debug will be tempted.
- You can also ask for a special kind of production called *Validation production*. This is used when you want produce a small bunch of files to see if they fulfill your expectations
- A validation production will put your data in the BKK under the *VALIDATION* folder so this test data will not be mixed with the real production

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

- After some initial development made few years WG has been used successfully by the PID WG (<https://cds.cern.ch/record/2161754>)
- This trigger some additional development on the LHCbDirac side that make the WG production a viable way to generate data centrally by the WGs
- To profit of this possibility the code have to be tested carefully
- The advantages are enormous from the point of view of users as well as for the Grid infrastructure
- <https://twiki.cern.ch/twiki/bin/view/LHCb/WGProductions>

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

- Name: A string meaningful for you that identify your step. It will be identified also by an ID. Although it's not compulsory but just a good habit I would advise to use dashes or underscores instead of blank spaces.
- Processing pass: This string will be used to build the BKK path. You can control this with the visibility flag.
- Application: Your main application and version. You can choose from a list of grid deployed projects and versions.
- Option Files: Your gaudi configurable.
- Option Format: This is related to ProdConf options handling. Refer to the documentation.

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

- Extra packages: Additional packages to enrich your environment. Usually this is the package under which you released your code.
- Runtime project: Another additional project if you need.
- CondDB, Dddb: CondDB and Dddb tags.

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

- DQTag: Not yet available.
- Visibility: The visibility flag if set to *Y* will trigger the attachment of the processing pass to the BKK path of the output files. This needs to be checked carefully. If you have multiple steps It would be a nice idea to set the visibility flag to *Y* only for one of the steps (the last one) to avoid a lengthy BKK path with redundant informations. On the other hand let's suppose that you have two requests. Both of them with three steps i.e. *Selection* → *Intermediate step* → *Merge*. Let's again suppose that there are two similar selections with a different set of cuts and the remaining steps are exactly the same. This is a use case where it would be nice to separate the two output in the BKK otherwise they will be mixed, and distinguish them can be not that easy.

Working Group
productions: mini
tutorial

Antonio Falabella

Working group
centralized
productions

LHCbDirac
Production
Concepts

Step creation

New Request setup

Summary

- Input File Types: The input files of your step.
- Output File Types: The output file of your step. If you are going to use a particular filename you can use the *LHCbDIRAC CLI* to enter a new filename in the BKK.

```
[lplus430] $dirac-bookkeeping-filetypes-insert  
Filetype: <You new type>
```

Or you can do it directly from the portal taking care of not inserting blanks.