



ILCDIRAC

A grid solution for the LC community

S. Poss

CERN, LAPP

Jan 29, 2013

Outline



Introduction

Job Management

- Using the python API

- Using the web portal

- Performance

DIRAC File Catalog

- Details

- Web interface

- Performance

Production Management

ILCDIRAC Support

Future plans

Conclusion

What is ILCDIRAC ?



It's a **DIRAC client** specific to the Linear Collider community:

- ▶ DIRAC is a **grid solution** initially developed for LHCb
- ▶ Extended for **other communitites**: Belle 2, BES, ILC, FERMI-LAT, several biomed applications, etc.
- ▶ **Complete solution**: Job Management, File catalog, etc.

ILCDIRAC:

- ▶ Enables use of **ALL ILC software** on the grid in a **unified manner**
- ▶ Use of **any available resource** (WLCG, OSG, local batch farms, etc.)
- ▶ Comes with a **web portal** that allows for most operations

Job Management with the python API



Idea: **Users need to care what to run, not how.**

- ▶ All ILC applications are configured uniformly thanks to a common interface
 - ▶ “Applications” share logical properties like SteeringFile, Version, etc.
 - ▶ Application specific things (Model, DST file name, etc.) have dedicated setters: they are only available to applications for which they make sense
- ▶ Data can be accessed from the catalog using meta data queries directly during job submission
 - ▶ DIRAC File catalog has meta data information (and replica info), see later
- ▶ Steering files used for Production are available directly to the users
 - ▶ Only the file names are needed
- ▶ Output files are stored in a user configurable directory, without need of Storage Element specification (cloud-like)
- ▶ Failover mechanism: **produced files cannot be lost**
 - ▶ copied to another storage, and a replication request is issued for later treatment by dedicated agent

Simple job example: Marlin



```
from DIRAC.Core.Base import Script
Script.parseCommandLine()
from ILCDIRAC.Interfaces.API.NewInterface.UserJob import UserJob
from ILCDIRAC.Interfaces.API.NewInterface.Applications import Marlin
from ILCDIRAC.Interfaces.API.DiraclLC import DiraclLC
d = DiraclLC() # Provides job checking utilities
j = UserJob() # You are running a user job
m = Marlin() # Get an application instance
m.setVersion("v0116") # Define the version to use
m.setSteeringFile("cllc_ild_cdr_steering.xml") #What the app should do
m.setInputFile("LFN:/ilc/prod/cllc/3tev/ee_h_bb/ILD/DST/00000375/000/\
ee_h_bb_dst_375_999.slcio") # Add some input
m.setGearFile("cllc_ild_cdr.gear") # Application specific field
res = j.append(m) # Add the application to the job
if not res['OK']:
    print res['Message'] # Catch any error
j.submit(d) # Submit the job
#Not shown here: metadata queries, chaining of applications
```

Complete examples and documentation in

<http://www.cern.ch/lcd-data/doc/HeadFirstTalk.pdf>

Job Management with the Web portal



- Possibility to submit small jobs through the portal directly

The screenshot displays the Job Management Web portal interface. The main window shows a table of jobs with columns for JobID, Status, MinorStatus, ApplicatorStatus, Site, JobName, LastUpdate [UTC], LastSignOffLife, SubmissionTime, and Owner. A modal window titled "Proxy Status: Valid" is open, showing job configuration details for "DIRAC_spose_369127", including Executable, Arguments, OutputSandbox, InputData, and OutputSE. The modal also has checkboxes for "InputData", "OutputData", "OutputSE", "OutputPath", "CPUTime", "Site", "BannedSite", "Platform", "Priority", "StdError", "StdOutput", "Parameters", "ParameterStart", and "ParameterStep". The main interface includes various filters on the left and a "Submit" button at the bottom left.

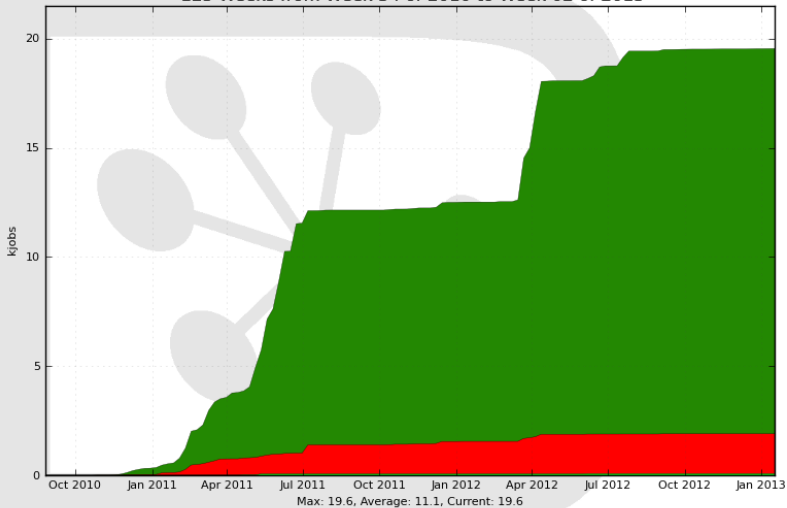
- Interface to monitor the statuses, apply some operations
- Small sandboxes can be obtained from the portal, big ones are stored on the GRID

Performance: User jobs



Jobs for 1 user

125 Weeks from Week 34 of 2010 to Week 02 of 2013



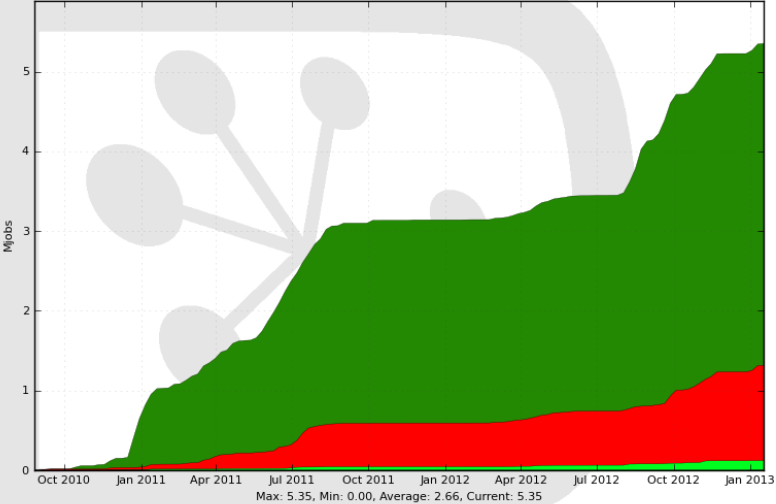
Generated on 2013-01-23 10:39:10 UTC

Performance: Production jobs



Production jobs

125 Weeks from Week 34 of 2010 to Week 02 of 2013



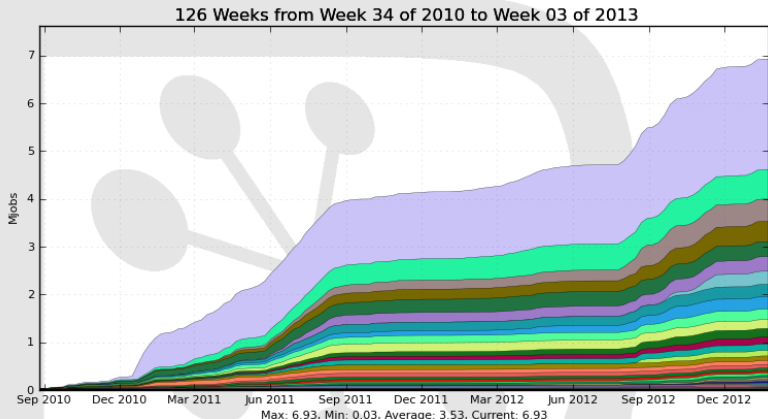
■ Done 4.0 ■ Failed 1.2 ■ Completed 0.1 ■ Rescheduled 0.0

Generated on 2013-01-23 14:01:41 UTC

Performance: Total jobs



All jobs



LCG.CERN.ch	2.3	LCG.IN2P3-IRES.fr	0.2
LCG.DESY-HH.de	0.6	LCG.DESYZN.de	0.2
LCG.RAL-LCG2.uk	0.5	LCG.Bristol.uk	0.2
LCG.Manchester.uk	0.4	LCG.UKI-NORTHGRID-LIV-HEP.uk	0.1
LCG.IN2P3-CC.fr	0.3	LCG.Brunel.uk	0.1
LCG.QMUL.uk	0.3	LCG.UKI-SOUTHGRID-RALPP.uk	0.1
LCG.FNAL_GPGRID_1.us	0.3	LCG.POLGRID.fr	0.1
LCG.UKI-LT2-IC-HEP.uk	0.3	LCG.DESY.de	0.1
LCG.GRIF.fr	0.3	... plus 27 more	

Generated on 2013-01-25 08:48:01 UTC



- ▶ Meta data catalog: **Directory and file level meta tags**
 - ▶ Searchable (like Event type) and non searchable tags (like number of events)
 - ▶ **Ancestor/daughter relationships**

- ▶ **Replica catalog**: Where the files are physically, with a path depending only on Storage Element logical name
 - ▶ SE logical name is defined in the Configuration Service: **if an end-point changes its name, the files are not lost**

File **Register operations** in ILCDIRAC now done **in the DFC and the Lcg File Catalog** simultaneously (failover applies here too).

ILC DBD files in the LFC are synchronised in the DFC.

File catalog web interface



System Jobs Data Views Tools

MetadataCatalog Metadata Query

Path to start from: /lc/prodclrc

EvType: hh_nunu
Datatype: gen
ProdID: 116

Select All Select None

File Name
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_1_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_10_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_11_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_12_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_13_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_14_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_15_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_16_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_17_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_18_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_19_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_2_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_20_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_21_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_22_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_23_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_24_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_25_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_26_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_27_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_28_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_29_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_3_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_30_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_31_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_32_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_33_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_34_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_35_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_36_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_37_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_38_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_39_stdhep
<input type="checkbox"/> /lc/prodclrc/3tevhh_nunu/gen/00001103/000hh_nunu_gen_1103_4_stdhep

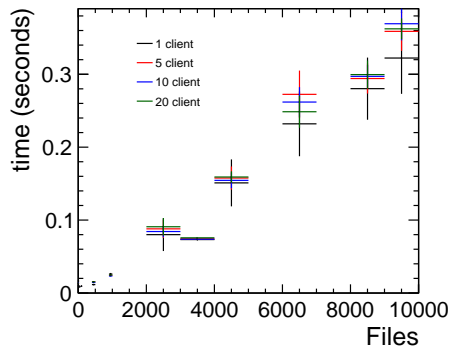
Refresh Submit Reset

File catalog performance

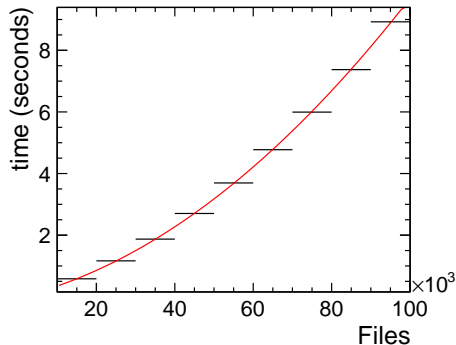


Timing for replica resolution:

Many concurrent clients:



One client:



Meta data search times are compatible with those of the AMGA system

Idea: **Apply a set of operations to a set of files automatically**

- ▶ **Generation** of a given channel at a given machine (with WHIZARD 1.95)
- ▶ **Simulation, Reconstruction** of a given channel (ILD and SiD reconstruction chains)
- ▶ **Transfer**: replication from one site to another
- ▶ Adding new tasks at the Generation steps automatically implies the creation of Simulation/Reconstruction of the new files: **data driven process**

All failing tasks that analyse files are reset: **99.9% of the data produced was simulated/reconstructed** during the CLIC_CDR production



- ▶ **Mailing lists:**
 - ▶ for registering users: ilcdirc-register@cern.ch
 - ▶ for registered users: ilc-dirac@cern.ch
- ▶ **JIRA bug tracking:** <https://its.cern.ch/jira/browse/ILCDIRAC>
 - ▶ CERN IT provides this service among others
- ▶ **Growing community** of users sharing tools
- ▶ **Large support** from the DIRAC developers



Short term:

- ▶ **Wiki** documentation: flexibility
- ▶ **WHIZARD 2** and **PHYSSIM** supported in ILCDIRAC: completeness

Longer term:

- ▶ **WHIZARD 2 process factory**: no precompiled binary needed, only process definition
- ▶ **Configure** and **Run the applications locally**, no installation necessary, let DIRAC do it. **Ideal for testing.**

ILCDIRAC is:

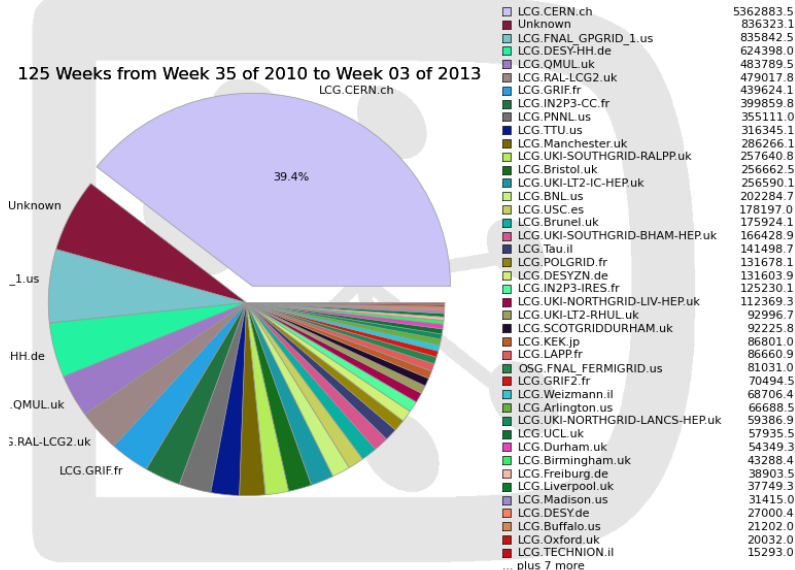
- ▶ A **fully functional grid solution** for the Linear Collider community
 - ▶ Used for **CLIC CDR** and **ILC DBD** activities
- ▶ Coming with an **intuitive/user friendly** interface
- ▶ **Easily connected with any new resources** (Site, storage, etc.)
- ▶ **Open source**
- ▶ **Available** for any ILC VO member



Backup Slides

Grid jobs per site

125 Weeks from Week 35 of 2010 to Week 03 of 2013



Generated on 2013-01-25 10:08:08 UTC