



iLCDirac: Status and Plans

André Sailer

CERN-EP-LCD

Asian Linear Collider Workshop
May 29, 2018

Table of Contents



1 Introduction

2 Status

3 Data Management

4 Computing Resources

5 Support

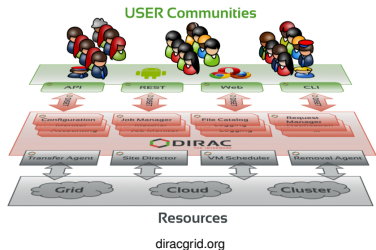
6 Plans

iLCDirac in a Nutshell



iLCDirac is based on the DIRAC interware originally developed for LHCb

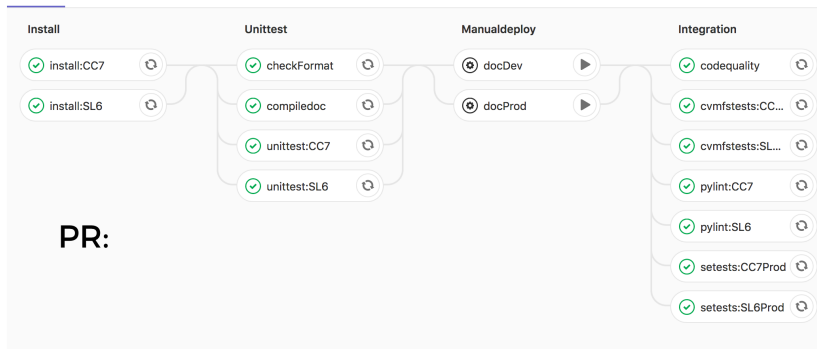
- Dirac (Distributed Infrastructure with Remote Agent Control): High level interface between users and distributed resources
- iLCDirac: Additional functionality to provide simple interface for the users to the LC Software (Whizard, Whizard2, Marlin, Mokka, org.lcsim, SLIC, ROOT, ddsim)
- Central system for large scale productions



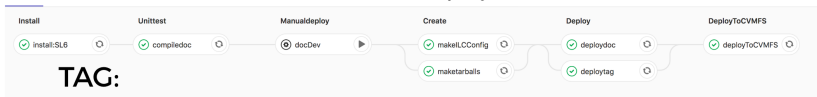
```
from DIRAC.Core.Base import Script
Script.parseCommandLine()
import UserJob, Marlin, DiracILC
d = DiracILC()
j = UserJob()
j.setOutputData("recEvents.slcio")
m = Marlin()
m.setVersion("ILCSoft-01-17-09")
m.setSteeringFile("Steering.xml")
m.setInputFile("SimEvents.slcio")
j.append(m)
j.submit(d)
```

- iLCDirac version v28r0p5, based on DIRAC v6r19p5
 - ▶ DIRAC v6r20 was released last week
- Same setup for iLCDirac servers:
 - ▶ Total of 100 Cores and 200 GB of Ram, SLC6 Virtual Machines, 2 × 3 Servers running Agents and Services: 8 Cores, 16 GB RAM; Split by DIRAC-System
 - ▶ 3 DIRAC DIP-Storage SEs: 4 Cores, 8 GB RAM, 1 TB Volume
 - ▶ All databases in CERN DB on Demand service
 - ▶ Web interface, CI, development, spares
- Unit test coverage of 62%; including tests running jobs and file upload/download/removal
 - ▶ Increased code-base due to new developments, but constant code coverage

Continuous Integration



Continuous Deployment



A number of new developments under the hood to ease operational workload

- **MonitoringAgents**: active monitoring of *Agents* and *Executors* to restart them in case of stall
 - Jobs should no get stuck in checking status any more
- **JobResetAgent**: Reset requests for jobs with waiting requests for set the job status to be finished
- **FileStatusTransformationAgent**: treats tasks for *Replication* transformations

Job Splitting



- Job Splitting: Not completely new, but probably rarely used
- Quickly and efficiently create a larger number of jobs
 - ▶ Split jobs by events: set number of jobs and number of events per job
 - ▶ Split files by job: use N files for each job
- See <http://lcd-data.web.cern.ch/lcd-data/doc/ilcdiracdoc/DOC/Files/UserGuide/examplejobs.html#automatic-job-splitting>
- Increase the Job Splitting Variants
 - ▶ Split a file into many jobs, skipping events

```
# [...]
dIlc = DiracILC(); job = UserJob()
job.setOutputSandbox("*.log")
job.setCLICConfig("ILCSOFT-2017-07-27")

# creates 10 jobs with 100 events each
job.setSplitEvents(eventsPerJob=100, numberOfJobs=10)

# output data name is automatically changed to, e.g., ddsimout.5.slcio
job.setOutputData("ddsimout.slcio", outputPath="sim1")

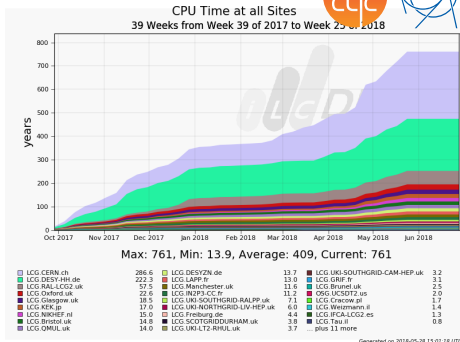
ddsim = DDSim()
ddsim.setVersion("ILCSOFT-2017-07-27_gcc62")
ddsim.setDetectorModel("CLIC_o3_v13")
ddsim.setExtraCLIArguments("_--enableGun--gun.particle=mu-")
ddsim.setNumberOfEvents(100)
ddsim.setSteeringFile("cllc_steer.py")
ddsim.setOutputFile("ddsimout.slcio")
myJob.append(ddsim)
myJob.submit(dIlc)
```

- Required further testing of FTS3 system inside DIRAC, will only be deployed with DIRAC v6r20 in the coming weeks
- Mostly improvement for replication transformations
- Tested replication between the all combinations of CERN-SRM, DESY-SRM, CERN-DST-EOS, RAL-SRM, KEK-SRM, IN2P3-SRM
 - ▶ IN2P3-SRM not working yesterday

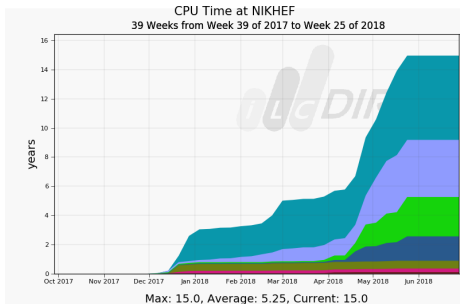
Source	Destination	Size	Progress	Speed	Time	Status
root://fts3prod01.cern.ch	root://fts3prod01.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod02.cern.ch	root://fts3prod02.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod03.cern.ch	root://fts3prod03.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod04.cern.ch	root://fts3prod04.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod05.cern.ch	root://fts3prod05.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod06.cern.ch	root://fts3prod06.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod07.cern.ch	root://fts3prod07.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod08.cern.ch	root://fts3prod08.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod09.cern.ch	root://fts3prod09.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod10.cern.ch	root://fts3prod10.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod11.cern.ch	root://fts3prod11.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod12.cern.ch	root://fts3prod12.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod13.cern.ch	root://fts3prod13.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod14.cern.ch	root://fts3prod14.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod15.cern.ch	root://fts3prod15.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod16.cern.ch	root://fts3prod16.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod17.cern.ch	root://fts3prod17.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod18.cern.ch	root://fts3prod18.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod19.cern.ch	root://fts3prod19.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod20.cern.ch	root://fts3prod20.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod21.cern.ch	root://fts3prod21.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod22.cern.ch	root://fts3prod22.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod23.cern.ch	root://fts3prod23.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod24.cern.ch	root://fts3prod24.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod25.cern.ch	root://fts3prod25.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod26.cern.ch	root://fts3prod26.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod27.cern.ch	root://fts3prod27.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod28.cern.ch	root://fts3prod28.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod29.cern.ch	root://fts3prod29.cern.ch	100 MB	100%	10 MB/s	10s	OK
root://fts3prod30.cern.ch	root://fts3prod30.cern.ch	100 MB	100%	10 MB/s	10s	OK

- Resources are currently discovered via the LDAP BDII system with the GLUE v1 schema
- This system is currently undergoing an overhaul
 - ▶ Many sites are unhappy with maintaining the information stored in the system and hope for a simpler solution.
 - ▶ There is a (totally over-engineered) GLUE v2 schema which is supposed to replace the v1 schema for 10 years.
 - ▶ CERN resources are now only published in GLUE v2
 - ▶ OSG resources are only published in *condor collectors*
 - ▶ CREAM, ARC and HTCondorCE information is not published the same way
 - ▶ Some Middleware developers (ARC) want to replace the LDAP system by a static json file
 - ▶ Information between v1 and v2 can be inconsistent
- We have to keep up with these changes and to access as many resources as are available for the ILC VO

Resources II



- Short discussion with NIKHEF to allow ILC VO into opportunistic resources, already in top 10 of sites





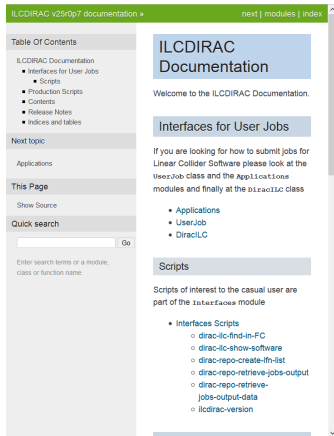
- More and more resources moving to CentOS7
- Access to some storage elements *might not* work on these worker nodes
- Have not had sufficient time for dedicated tests
- There will be a new way of providing the *lcg-bundles* (storage plugins) with DIRAC soon, which should definitively solve this problem

■ In case of fire:

- 1 Consult documentation:
<http://lcd-data.web.cern.ch/lcd-data/doc/ilcdiracdoc/>
- 2 Before submitting a ticket, see:
<http://lcd-data.web.cern.ch/lcd-data/doc/ilcdiracdoc/DOC/Files/UserGuide/support.html>
- 3 Submit a ticket to the issue tracker
<https://its.cern.ch/jira/browse/ILCDIRAC>
 - ★ See also “Report a Problem” buttons in web portal and documentation
- 4 Email: ilcdirac-support@cern.ch

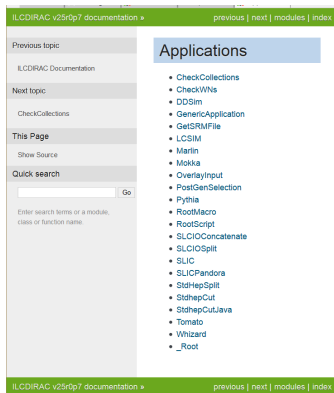


- <http://lcd-data.web.cern.ch/lcd-data/doc/ilcdiracdoc/>
- Information about commands (scripts) including options
- API, examples for all applications



The screenshot shows the ILCDIRAC v25R0p7 documentation website. The page has a green header with the text "ILCDIRAC v25R0p7 documentation" and navigation links for "next", "modules", and "index". On the left side, there is a "Table Of Contents" section with a tree view: "ILCDIRAC Documentation" (expanded) containing "Interfaces for User Jobs" (expanded) with sub-items "Scripts", "Production Scripts", "Contents", "Release Notes", and "Indices and tables". Below this are sections for "Next topic", "Applications", "This Page" (with "Show Source" link), and "Quick search" (with a search input field and "Go" button). The main content area on the right features a blue header "ILCDIRAC Documentation", a welcome message, a blue header "Interfaces for User Jobs", a paragraph about submitting jobs, a list of links for "Applications", "UserJob", and "DiracILC", a grey header "Scripts", and a list of "Interfaces Scripts" including "dirac-lic-find-in-FC", "dirac-lic-show-software", "dirac-repo-create-lfn-list", "dirac-repo-retrieve-jobs-output", "dirac-repo-retrieve-jobs-output-data", and "ilcdirac-version".

- <http://lcd-data.web.cern.ch/lcd-data/doc/ilcdiracdoc/>
- Information about commands (scripts) including options
- API, examples for all applications



ILCDIRAC v25r0p7 documentation » previous | next | modules | index

Previous topic
ILCDIRAC Documentation

Next topic
CheckCollections

This Page
Show Source

Quick search
 Go

Enter search terms or a module, class or function name.

Applications

- CheckCollections
- CheckWNs
- DDSim
- GenericApplication
- GetSRMFile
- LCSIM
- Marlin
- Mokka
- OverlayInput
- PostGenSelection
- Pythia
- RootMacro
- RootScript
- SLICConcatenate
- SLICOSplit
- SLIC
- SLICPandora
- StdHepSplit
- StdHepCut
- StdHepCutJava
- Tomato
- Whizard
- _Root

ILCDIRAC v25r0p7 documentation » previous | next | modules | index

- Move to DIRAC v6r20 in the next few weeks
 - FTS3 for replication transformations
- Follow-up changes of resource discovery systems
 - Integrate OSG condor collectors into resource discovery
- Add further options for automatic *JobSplitting*
- Continue to provide support and high-available service for users