iLCDirac: Status and Plans

André Sailer

CERN-EP-LCD

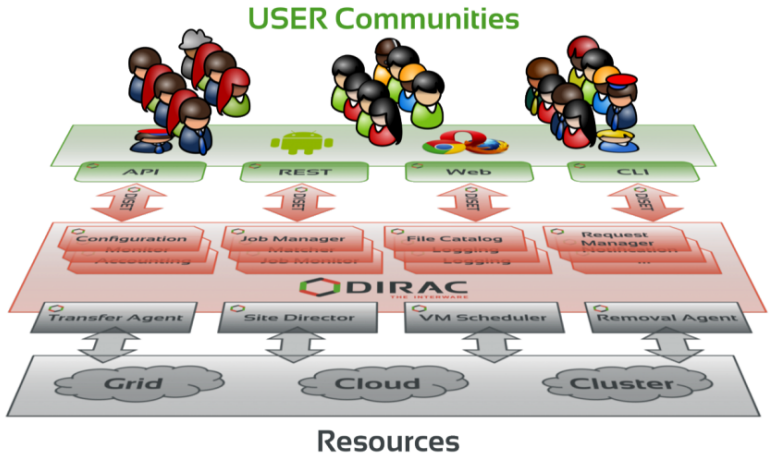
Asian Linear Collider Workshop
May 29, 2018

Table of Contents

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



diracgrid.org

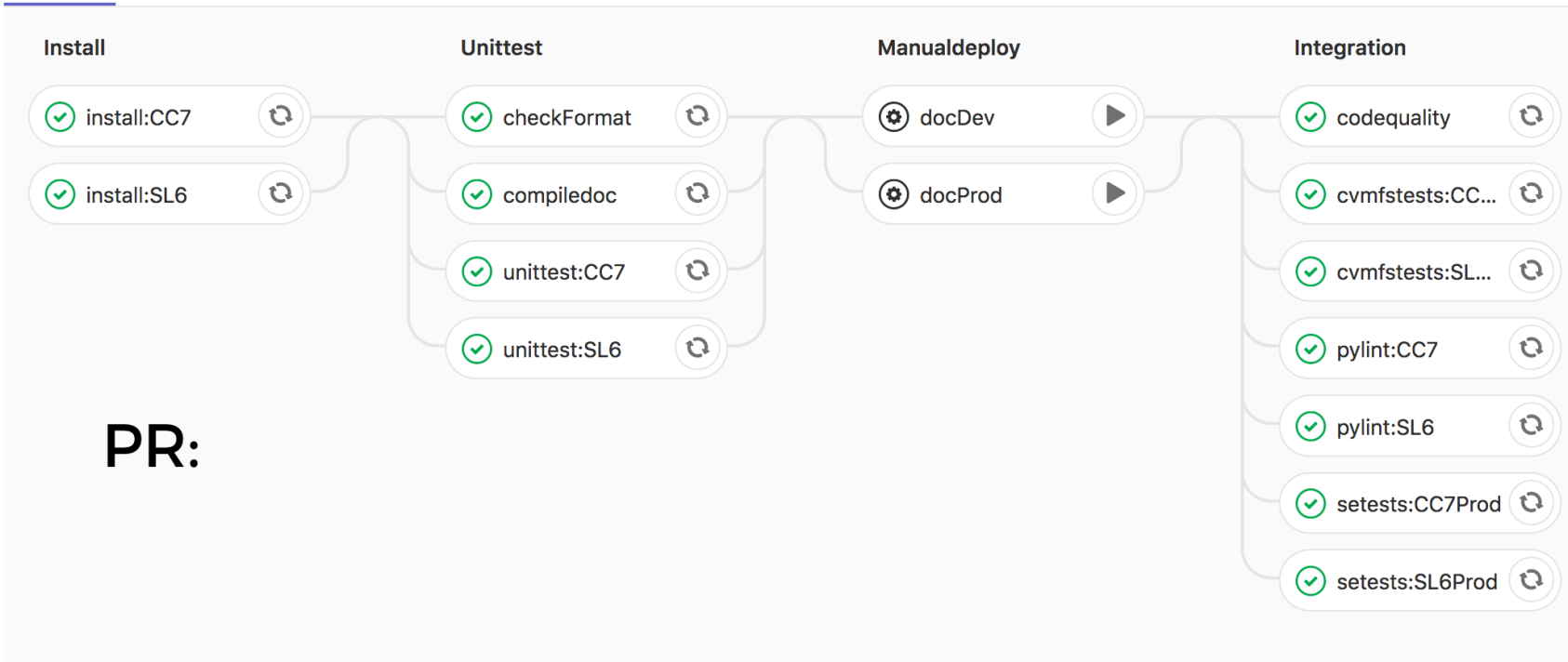
```
from DIRAC.Core.Base import Script
Script.parseCommandLine()
import UserJob, Marlin, DiracILC
d = DiracILC()
j = UserJob()
j.setOutputData("recEvents.slcio")
m = Marlin()
m.setVersion("ILCSoft-01-17-09")
m.setSteeringFile("Steering.xml")
m.setInputFile("SimEvents.slcio")
j.append(m)
j.submit(d)
```

Status

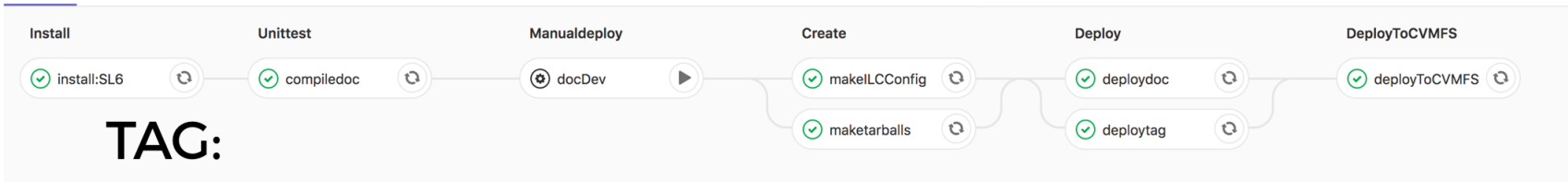
- iLCDirac version v28r0p5, based on DIRAC v6r19p5
  - DIRAC v6r20 was released last week
- Same setup for iLCDirac servers:
  - Total of 100 Cores and 200 GB of Ram, SLC6 Virtual Machines, 2×3 Servers running Agents and Services: 8 Cores, 16 GB RAM; Split by DIRAC-System
  - 3 DIRAC DIP-Storage SEs: 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

CI/CD

Continuous Integration



Continuous Deployment



New Developments

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

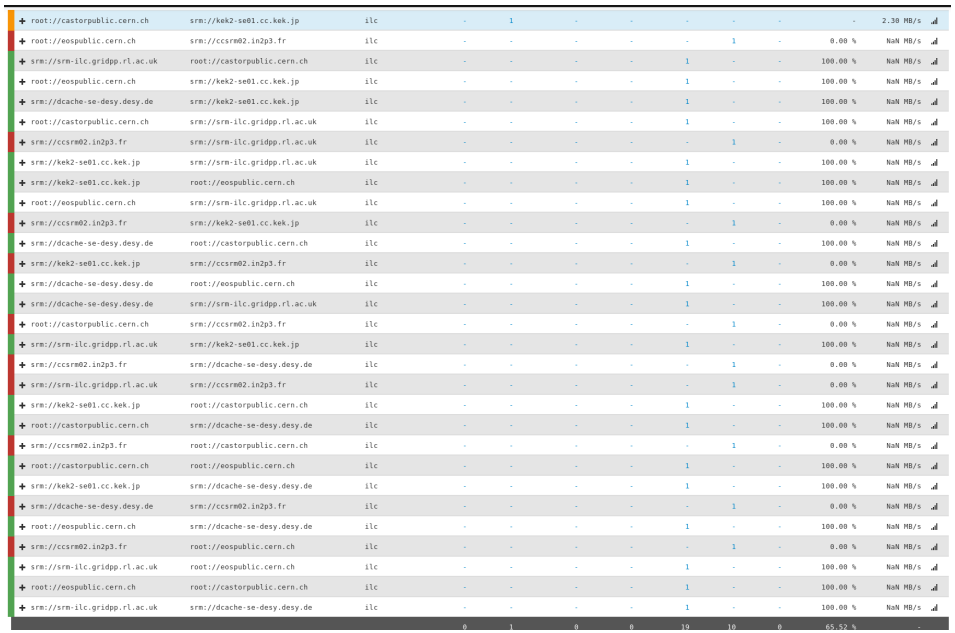
Job Splitting

- 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("clic_steer.py")
ddsim.setOutputFile("ddsimout.slcio")
myJob.append(ddsim)
myJob.submit(dIlc)
```

FTS3

- 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



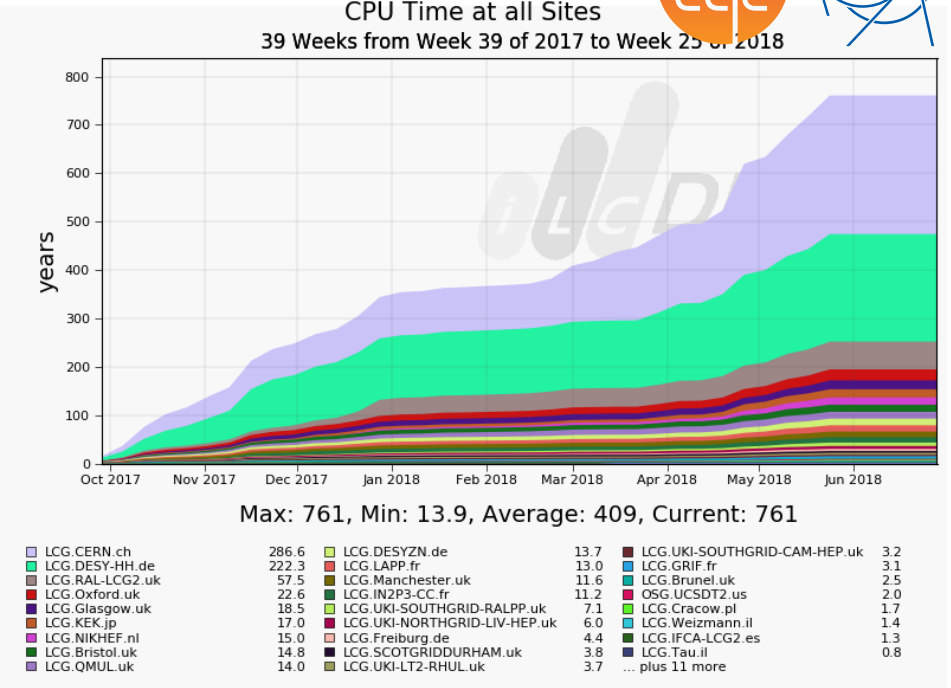
root://castorpublic.cern.ch	srm://kek2-se01.cc.kek.jp	ilc	-	1	-	-	-	-	-	-	2.30 MB/s
root://eospublic.cern.ch	srm://ccsrm02.in2p3.fr	ilc	-	-	-	-	-	1	-	0.00 %	NaN MB/s
srm://srm-ilc.gridpp.rl.ac.uk	root://castorpublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
root://eospublic.cern.ch	srm://kek2-se01.cc.kek.jp	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
srm://dcache-se-desy.desy.de	srm://kek2-se01.cc.kek.jp	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
root://castorpublic.cern.ch	srm://srm-ilc.gridpp.rl.ac.uk	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
srm://ccsrm02.in2p3.fr	srm://srm-ilc.gridpp.rl.ac.uk	ilc	-	-	-	-	-	1	-	0.00 %	NaN MB/s
srm://kek2-se01.cc.kek.jp	srm://srm-ilc.gridpp.rl.ac.uk	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
srm://kek2-se01.cc.kek.jp	root://eospublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
root://eospublic.cern.ch	srm://srm-ilc.gridpp.rl.ac.uk	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
srm://ccsrm02.in2p3.fr	srm://kek2-se01.cc.kek.jp	ilc	-	-	-	-	-	1	-	0.00 %	NaN MB/s
srm://dcache-se-desy.desy.de	root://castorpublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
srm://kek2-se01.cc.kek.jp	srm://ccsrm02.in2p3.fr	ilc	-	-	-	-	-	1	-	0.00 %	NaN MB/s
srm://dcache-se-desy.desy.de	root://eospublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
srm://dcache-se-desy.desy.de	srm://srm-ilc.gridpp.rl.ac.uk	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
root://castorpublic.cern.ch	srm://ccsrm02.in2p3.fr	ilc	-	-	-	-	-	1	-	0.00 %	NaN MB/s
srm://srm-ilc.gridpp.rl.ac.uk	srm://kek2-se01.cc.kek.jp	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
srm://ccsrm02.in2p3.fr	srm://dcache-se-desy.desy.de	ilc	-	-	-	-	-	1	-	0.00 %	NaN MB/s
srm://srm-ilc.gridpp.rl.ac.uk	srm://ccsrm02.in2p3.fr	ilc	-	-	-	-	-	1	-	0.00 %	NaN MB/s
srm://kek2-se01.cc.kek.jp	root://castorpublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
root://castorpublic.cern.ch	srm://dcache-se-desy.desy.de	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
srm://ccsrm02.in2p3.fr	root://castorpublic.cern.ch	ilc	-	-	-	-	-	1	-	0.00 %	NaN MB/s
root://castorpublic.cern.ch	root://eospublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
srm://kek2-se01.cc.kek.jp	srm://dcache-se-desy.desy.de	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
srm://dcache-se-desy.desy.de	srm://ccsrm02.in2p3.fr	ilc	-	-	-	-	-	1	-	0.00 %	NaN MB/s
root://eospublic.cern.ch	srm://dcache-se-desy.desy.de	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
srm://ccsrm02.in2p3.fr	root://eospublic.cern.ch	ilc	-	-	-	-	-	1	-	0.00 %	NaN MB/s
srm://srm-ilc.gridpp.rl.ac.uk	root://eospublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
root://eospublic.cern.ch	root://castorpublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
srm://srm-ilc.gridpp.rl.ac.uk	srm://dcache-se-desy.desy.de	ilc	-	-	-	-	1	-	-	100.00 %	NaN MB/s
			0	1	0	0	19	10	0	65.52 %	-

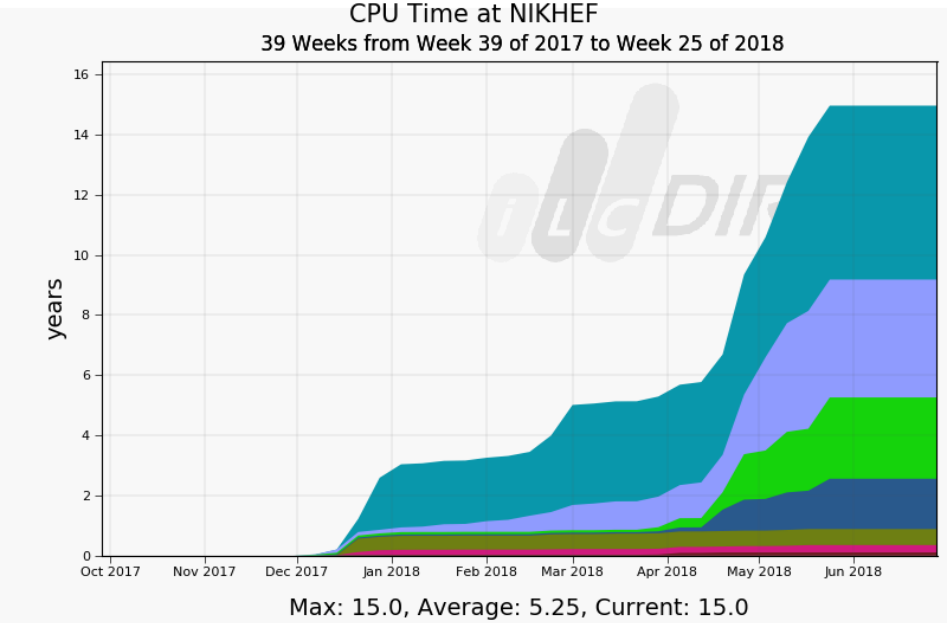
Resources I

- 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





Resources III

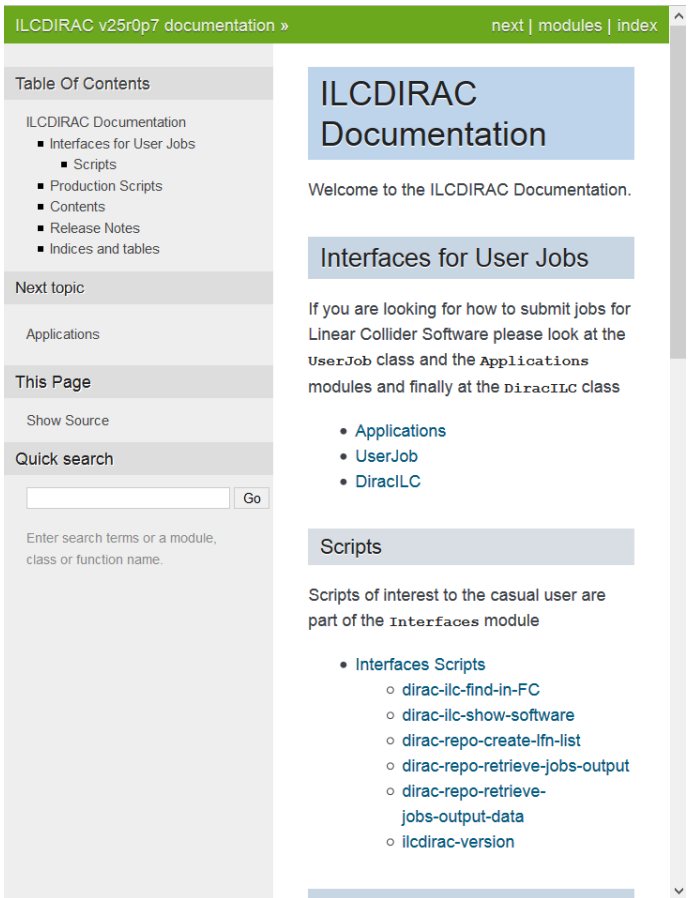
- More and more resources moving to CentOS7
- Access to some storage elements *might not* work on these worker nodes
- Have not had sufficient time for dedicated tests
- There will be a new way of providing the *lcg-bundles* (storage plugins) with DIRAC soon, which should definitively solve this problem

Support

- 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

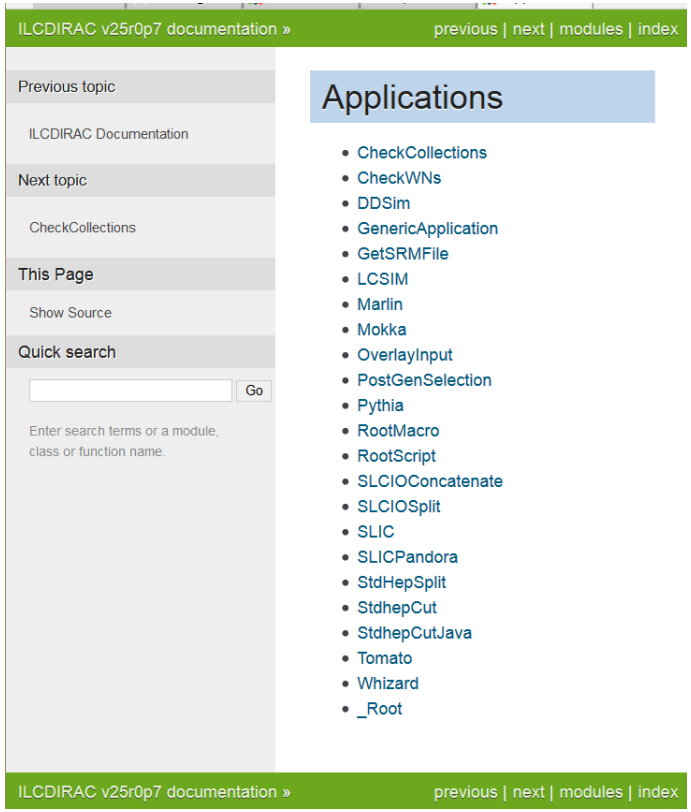
Documentation

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



Documentation

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



Plans

- 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