STATUS AND PLANS

CLICdp Collaboration Meeting

Marko Petrič

On behalf of the CLICdp collaboration

Genève, 29 August 2018

iLCDirac Use Case

- ILC VO: virtual organization for linear colliders
- iLCDirac is an extension of the DIRAC system for the ILC VO
  - Workflow Modules for LC Software, Overlay System
  - JPCS. ILCDirac, a DIRAC extension for the Linear Collider community. Proceedings of CHEP2013. 513 CLICdp-Conf-2013-003
  - JPCS. Using OSG Computing Resources with (iLC)DIRAC. Proceedings of CHEP2016. CLICdp-Conf-2017-003
- Centralized MC Production (Event Generation, Sim and Rec)
- User jobs (Generation, Simulation, Reconstruction, Analyses)

Capacity:

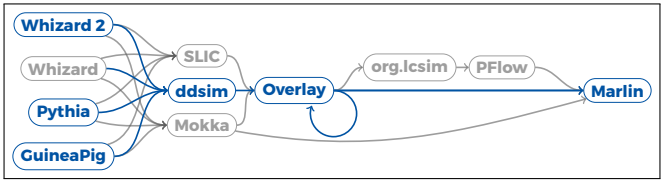
- Using WLCG and OSG resources (CREAM, Arc, HTCondorCE)
  - Mostly opportunistic, some dedicated
  - Around 15k to 20k job slots available at best of times

Code: https://gitlab.cern.ch/CLICdp/ILCDIRAC

API, Workflow

- Define application payload via interfaces
- Chain applications (append one after the other)

```
from DIRAC.Core.Base import Script
Script.parseCommandLine()
import UserJob
import Marlin
import 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 v29r0p1, based on DIRAC v6r20p7
  - Latest DIRAC version
  - Major new version in preparation
- 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

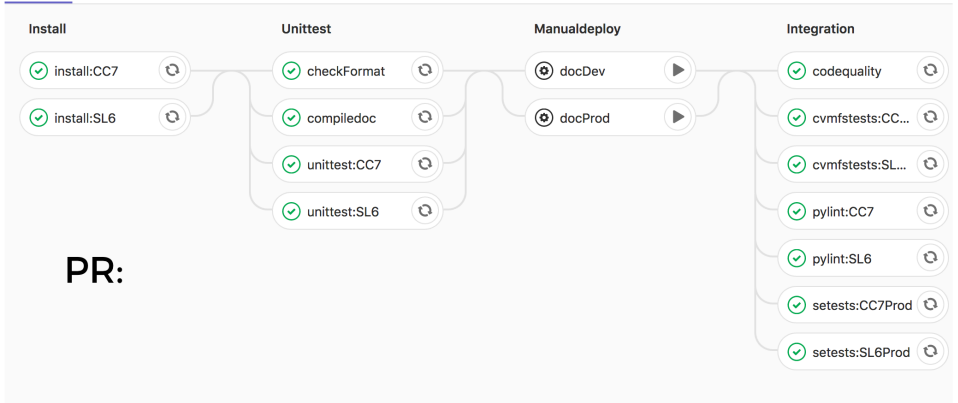
Testing iLCDirac evolution

- Extended tests since last Collaboration Meeting

test type	August 2017	August 2018
Workload	9	12
SE	3	5
unit	1322	1620

- Constant coverage 61% to 62%
- From 9435/16390 relevant lines to 11090/18401 relevant lines covered
- Direct deployment of client to cvmfs and tarballs and documentation to webserver via GitLab-CI

Testing iLCDirac pipelines



New Developments

New developments under the hood to streamline operations

- 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 and set the job status to finished
- FileStatusTransformationAgent: treats tasks for Replication transformations
- Automation of the production system
  - Ulrike Schnoor production manager for new transformations
  - First productions with Whizard 2

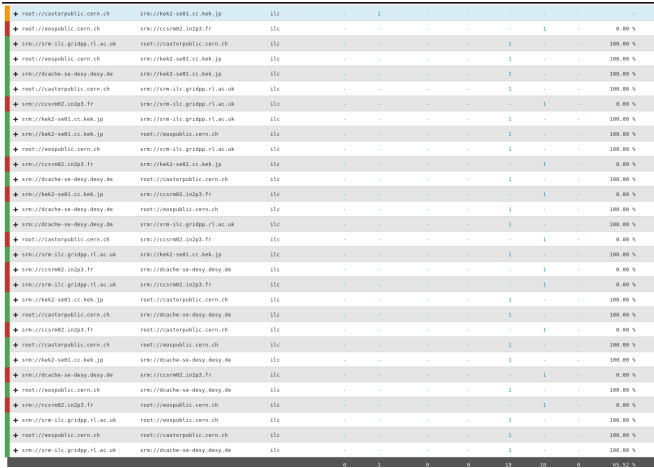
Job Splitting

- Job Splitting: Not completely new, but probably rarely used
- Quickly and efficiently create a larger number of jobs
  - Split jobs by evts: set N jobs and N events per job
  - Split files by job: use N files for each job
- Very fast submission (e.g 4k jobs in 3min, limit 10k)
- See web documentation
- Increase the Job Splitting Variants
  - Split a file into many jobs, skipping events

```
dIlc = DiracILC()
job = UserJob()
job.setOutputSandbox("*.log")

# 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

- FTS3 system inside DIRAC deployed with DIRAC v6r20
- Mostly improvement for replication transformations
- Tested replication of data between all combinations of CERN-SRM, DESY-SRM, CERN-DST-EOS, RAL-SRM, KEK-SRM, IN2P3-SRM



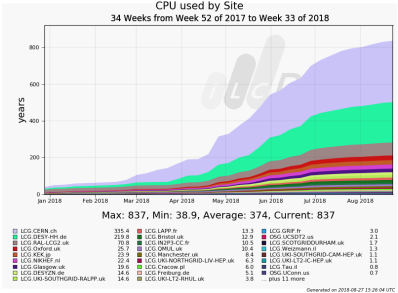
root://castorpublic.cern.ch	srm://kek2-se01.cc.kek.jp	ilc	-	1	-	-	-	-	-	-
root://eospublic.cern.ch	srm://ccsrm02.in2p3.fr	ilc	-	-	-	-	-	1	-	0.00 %
srm://srm-ilc.gridpp.rl.ac.uk	root://castorpublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %
root://eospublic.cern.ch	srm://kek2-se01.cc.kek.jp	ilc	-	-	-	-	1	-	-	100.00 %
srm://dcache-se-desy.desy.de	srm://kek2-se01.cc.kek.jp	ilc	-	-	-	-	1	-	-	100.00 %
root://castorpublic.cern.ch	srm://srm-ilc.gridpp.rl.ac.uk	ilc	-	-	-	-	1	-	-	100.00 %
srm://ccsrm02.in2p3.fr	srm://srm-ilc.gridpp.rl.ac.uk	ilc	-	-	-	-	-	1	-	0.00 %
srm://kek2-se01.cc.kek.jp	srm://srm-ilc.gridpp.rl.ac.uk	ilc	-	-	-	-	1	-	-	100.00 %
srm://kek2-se01.cc.kek.jp	root://eospublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %
root://eospublic.cern.ch	srm://srm-ilc.gridpp.rl.ac.uk	ilc	-	-	-	-	1	-	-	100.00 %
srm://ccsrm02.in2p3.fr	srm://kek2-se01.cc.kek.jp	ilc	-	-	-	-	-	1	-	0.00 %
srm://dcache-se-desy.desy.de	root://castorpublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %
srm://kek2-se01.cc.kek.jp	srm://ccsrm02.in2p3.fr	ilc	-	-	-	-	-	1	-	0.00 %
srm://dcache-se-desy.desy.de	root://eospublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %
srm://dcache-se-desy.desy.de	srm://srm-ilc.gridpp.rl.ac.uk	ilc	-	-	-	-	1	-	-	100.00 %
root://castorpublic.cern.ch	srm://ccsrm02.in2p3.fr	ilc	-	-	-	-	-	1	-	0.00 %
srm://srm-ilc.gridpp.rl.ac.uk	srm://kek2-se01.cc.kek.jp	ilc	-	-	-	-	1	-	-	100.00 %
srm://ccsrm02.in2p3.fr	srm://dcache-se-desy.desy.de	ilc	-	-	-	-	-	1	-	0.00 %
srm://srm-ilc.gridpp.rl.ac.uk	srm://ccsrm02.in2p3.fr	ilc	-	-	-	-	-	1	-	0.00 %
srm://kek2-se01.cc.kek.jp	root://castorpublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %
root://castorpublic.cern.ch	srm://dcache-se-desy.desy.de	ilc	-	-	-	-	1	-	-	100.00 %
srm://ccsrm02.in2p3.fr	root://castorpublic.cern.ch	ilc	-	-	-	-	-	1	-	0.00 %
root://castorpublic.cern.ch	root://eospublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %
srm://kek2-se01.cc.kek.jp	srm://dcache-se-desy.desy.de	ilc	-	-	-	-	1	-	-	100.00 %
srm://dcache-se-desy.desy.de	srm://ccsrm02.in2p3.fr	ilc	-	-	-	-	-	1	-	0.00 %
root://eospublic.cern.ch	srm://dcache-se-desy.desy.de	ilc	-	-	-	-	1	-	-	100.00 %
srm://ccsrm02.in2p3.fr	root://eospublic.cern.ch	ilc	-	-	-	-	-	1	-	0.00 %
srm://srm-ilc.gridpp.rl.ac.uk	root://eospublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %
root://eospublic.cern.ch	root://castorpublic.cern.ch	ilc	-	-	-	-	1	-	-	100.00 %
srm://srm-ilc.gridpp.rl.ac.uk	srm://dcache-se-desy.desy.de	ilc	-	-	-	-	1	-	-	100.00 %
			0	1	0	0	19	10	0	69.52 %

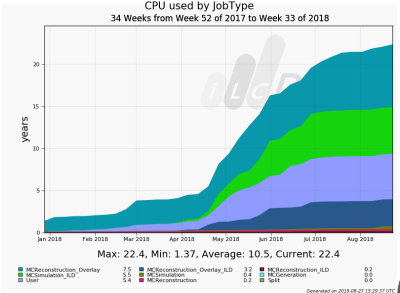
Resources

- Resources are currently resolved via the BDII system via GLUE1
- This system is currently undergoing an overhaul
  - Many sites are unhappy with maintaining the information
  - There is a GLUE2 successor
  - CERN resources are now only published in GLUE2
  - 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

New Resources

- Short discussion with NIKHEF to allow ILC VO into opportunistic resources, already in top 10 of sites
- In discussion with new sites for additional resources:
  - JINR should be available soon
  - CNAF and Poland in discussion





Changes in Resources

- The grid is migrating to CentOS7
- Might cause access problems to storage elements on these 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
  - New major release of DIRAC in the working, with support for CentOS7 and possible other platforms
- Work to reduce user's use of Castor (CERN-SRM)
  - Development of new special type SE for users to use with automatic archiving or removing

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