



# ILCDirac: A DIRAC Extension for the Linear Collider Community

André Sailer

CERN-PH-LCD

5th DIRAC User Workshop  
Ferrara, Italy  
May 27, 2015

# Table of Contents



- 1 What is ILCDirac?
- 2 Systems in Use
- 3 Systems not in Use
- 4 Other Things
- 5 Summary

# What Do We Do With DIRAC



- ILC VO: virtual organisation for linear colliders (ILC and CLIC)
- ILCDirac is an extension of the DIRAC system for the ILC VO
  - ▶ Workflow Modules for LC Software, Overlay System
  - ▶ J. Phys.: Conf. Ser. ILCDirac, a DIRAC extension for the Linear Collider community. Proceedings of CHEP2013. 513 [CLICdp-Conf-2013-003](#)
- Centralized MC Production (Event Generation, Geant4 Simulation, Reconstruction)
- User jobs (Generation, Simulation, Reconstruction, Analyses)

## Capacity:

- Using WLCG resources
  - ▶ Mostly opportunistic, some dedicated
  - ▶ Around 10k job slots available at best of times
- Working on including OSG resources

Total of 50 Cores and 100 GB of Ram, SLC6 Virtual Machines

- 3 Servers running Agents and Services: 8 Cores, 16 GB RAM; Split by DIRAC-System
  - 1 Framework, Transformation, DataManagement, Configuration
  - 2 StorageManagement, WorkloadManagement
  - 3 RequestManagement, Accounting, (ResourceStatus)
- 3 DIRAC SEs: 4 Cores, 8 GB Ram, 1 TB Volume
  - ▶ DIP-SE, Log-SE, SB-SE
- Web Server 2 Cores, 4 GB RAM
- 2 Dev/Test instances 8/4 Cores, 16/8 GB RAM
- DBs hosted on CERN DB on Demand (ilcdirac, ilcacdb (accounting DB), ilcdtest)

- Using v6r11, since October last year
  - ▶ Troubles with moving to new version: we have a different setup than what DIRAC is being tested on (e.g.: no FTSManger (yet))
  - ▶ Generally fast response for fixes
  - ▶ Some issues still open (Web-interface, StorageManagement::RequestCleaningAgent)
- Started looking into v6r12, i.e., using the BDII2CS agent
- I understand I should move to v6r13 as soon as possible

# (Some of the) Systems In Use



- Accounting

- ▶ Filling large database, creating monitoring plots

- Transformation

- StorageManagement

- ▶ Some issues in the past when the inputs were non-normalized LFNs

Using the DFC in production for many years

- 11 million files, 13 million replicas, 2.3 PB total size
- Metadata for production files (Energy, Software, Detector, Machine, physics process, Number of Events, . . .)
- Metadata search used for transformations

Using RMS for mostly failover, some replication transformations

- RequestTypes: Failover, Removal after failed jobs, Replication Transformations, LogUploadFailover (adapted from LHCb), DISSETForward,
- No FTS yet



## Using the old webportal

- Mostly looking at: JobMonitor, ProductionMonitor, SiteSummary, PilotMonitor, PilotSummary, Presenter, SystemMonitor, Configuration(Browser, Manager), SytemAdministrator

When there is some time will look into the new webportal

- SiteDirectors for CREAM, ARC (and Globus CEs)
  - ▶ wrote `GlobusComputingElement` to do direct submission to OSG Globus computing elements (e.g., PNNL)  
<https://github.com/andresailer/DIRAC/commit/bdc67e2ea39fd3216882910adb54efeab3dfa1f2>
  - ▶ Works with `SiteDirector` `CETType=Globus` at the moment
  - ▶ Will adapt this for v6r13 `LocalComputingElement`
- `TaskQueueDirector` to submit to KEK via grid WMS

- Tried to install this for *Development* instance only
- Also impacted *Production* part (e.g.: could not download sandboxes from Webportal any longer)
- Gave up on it, do not know the benefit of RSS



- Will look into this when there is some time
- Cannot find “Site Summary” in this instance: <https://dirac.ub.es/ILC/>
- Resource Summary requires Resource Status

- Grid is dynamic, need to keep track of failures and changes: Site troubles, CEs retiring (dirac-admin-remove-site?), external software updates breaking things (SSLv3), ...
- BDII2CS agent gets confused by some sites
- More automated way of controlling resources would be great
  - ▶ How much of that can the RSS do?
- Better (automated) monitoring (some of which already developed by Belle) would be great

Updating to new releases is a lengthy process

- Need to adapt our code for possible interface changes
  - Some can only be seen at runtime (e.g., return values)
- Need to update DB schemas, which is hard to do when one is just supposed to look at diffs and impossible if DB changes only hidden in commits
  - Please think twice (or more times) before changing interfaces or DBs and then immediately write this down in the migration documentation
- Then there are still a few bugs that need to be found and fixed before deploying releases to production

- It feels like there is more documentation than a year ago
- But still missing documentation for Accounting, StorageManagement, WorkloadManagement (LocalComputingElement (which is totally a misnomer!) configuration)
- Release notes are very confusing:
  - ▶ What are the major changes (improvements) for new releases (v6r12, v6r13)?
  - ▶ What are the plans for new releases?
  - ▶ What are the lifetimes of the different releases?

# Thanks



- Great support from DIRAC Developers and Community
  - Subscribe to the forum with your email address:  
`diracgrid-forum+subscribe@googlegroups.com`
- Generally fast answers and solutions



# Summary



- Have used (ILC)DIRAC for many years with great success
- Will see more usage in ILC VO

# Backup Slides

