GRID PRODUCTION TOOLS

The 2015 International Workshop on Future Linear Colliders

Marko Petrič

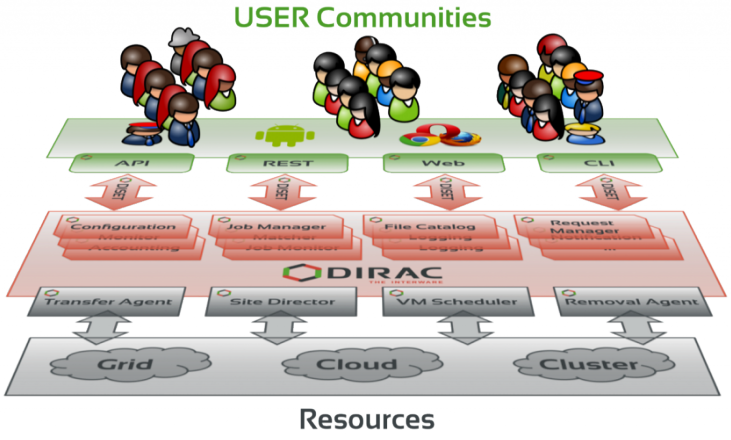
on behalf of the CLICdp Collaboration

3 November 2015

ILCDirac in a Nutshell

ILCDirac is based on the DIRAC interware originally developed for LHCb

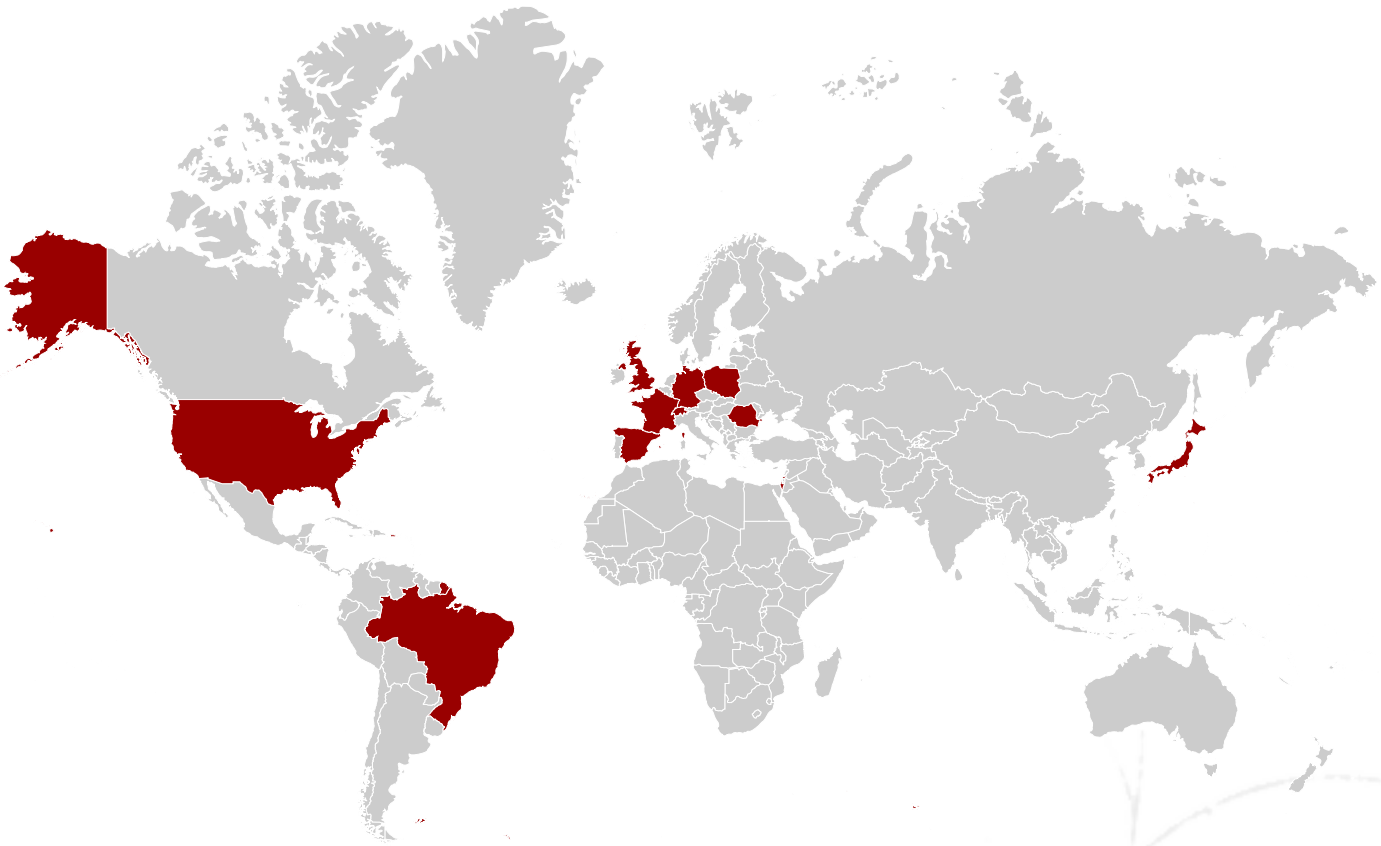
- Dirac (Distributed Infrastructure with Remote Agent Control): High level interface between users and distributed resources
- ILCDirac: Additional functionality to provide simple interface for users to the LC Software (Whizard, Marlin, Mokka, org.lcsim, SLIC, ROOT)



diracgrid.org

```
from DIRAC.Core.Base import Script
Script.parseCommandLine()
import UserJob
import Marlin
import DiracILC
d = DiracILC()
j = UserJob()
j.setOutputSandbox("recEvents.slcio")
m = Marlin()
m.setVersion("0116")
m.setSteeringFile("Steering.xml")
m.setInputFile("SimEvents.slcio")
j.append(m)
j.submit(d)
```

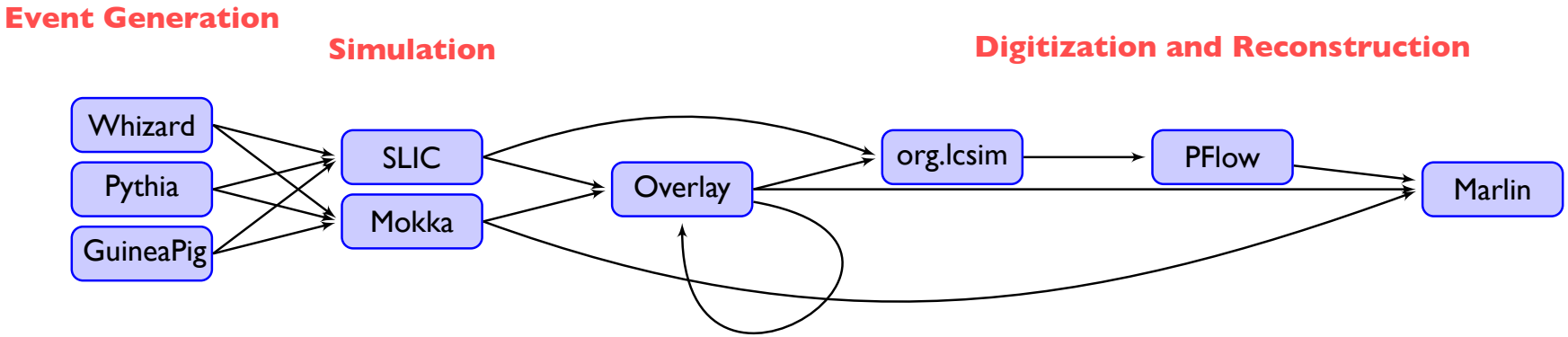
Resources



- Running on all sites (41) that support ILC VO and CALICE VO
- 4 different middlewares – Cream, Arc, Globus, HTCondorCE
- running everywhere is a major achievement
- At peak running 12 000 jobs in parallel

Production System

We offer centralised MC Production



- Need to chain many different applications for linear collider detector studies
- Software packages interfaced through common IO format
- Allow for easy extension:
  - User `Jobs`: input and output provided directly by users
  - Production `Jobs`: submission delegated to the `Transformation System`, so no input sandbox allowed
  - Derived classes take care of specialised metadata and path conventions

Transformation System

Transform one set of files into another

- Generation, Simulation, Reconstruction, Replication, …
  - Example Whizard generation:

```
pwh = ProductionJob()
pwh.setOutputSE("CERN-SRM")
pwh.setProdType("MCGeneration")
...
pwh.setWorkflowName(wname)
pwh.setProdGroup(analysisname)
res = pwh.append(wh)
pwh.setDescription(descrp)
res = pwh.createProduction()
pwh.addMetadataToFinalFiles({"":})
res = pwh.finalizeProd()
meta = pwh.getMetadata()
```

- One production is created
- Use metadata from this production for connection to next step

Transformation System

Continue transformation flow:

- Simulation: Define detector model, input production ID, …

```
mo = Mokka()
mo.setVersion('0706P08')
mo.setSteeringFile("ild_00.steer")

pmo = ProductionJob()
pmo.setProdType('MCSimulation')
pmo.setInputDataQuery(meta)
pmo.setOutputSE("CERN-SRM")
res = pmo.append(mo)
pmo.setDescription(descrp)
res = pmo.createProduction()
pmo.addMetadataToFinalFiles("...")
res = pmo.finalizeProd()
meta = pmo.getMetadata()
```

Transformation System

Continue transformation flow:

- Append overlay, reconstruction, …
- Number of jobs dictated by first production

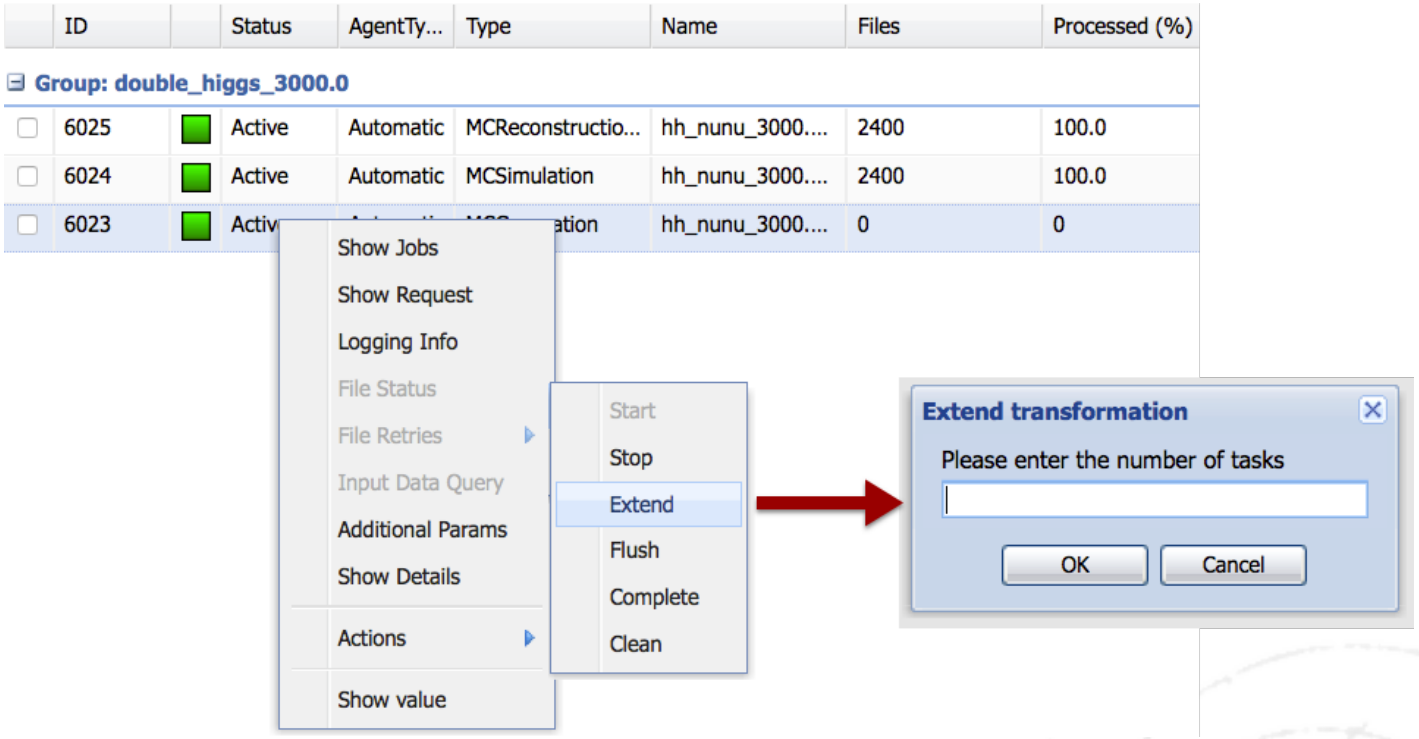
Select All Select None Start

ID	Status	AgentType	Type	Name	Files	Processed (%)
Group: double_higgs_3000.0						
6025	Active	Automatic	MCReconstruction_Overlay	hh_nunu_3000.0_ild_rec_overlay_beamrecoil_e...	2400	100.0
6024	Active	Automatic	MCSimulation	hh_nunu_3000.0_ild_sim_beamrecoil_e1F_E1F	2400	100.0
6023	Active	Automatic	MCGeneration	hh_nunu_3000.0_beamrecoil_e1F_E1F	0	0

- Transformation system checks every ~10 min based on metadata for need to process new files
- If sim or reco job fails, a new job is created, almost no interventions necessary
- The system takes care that each file is processed exactly once

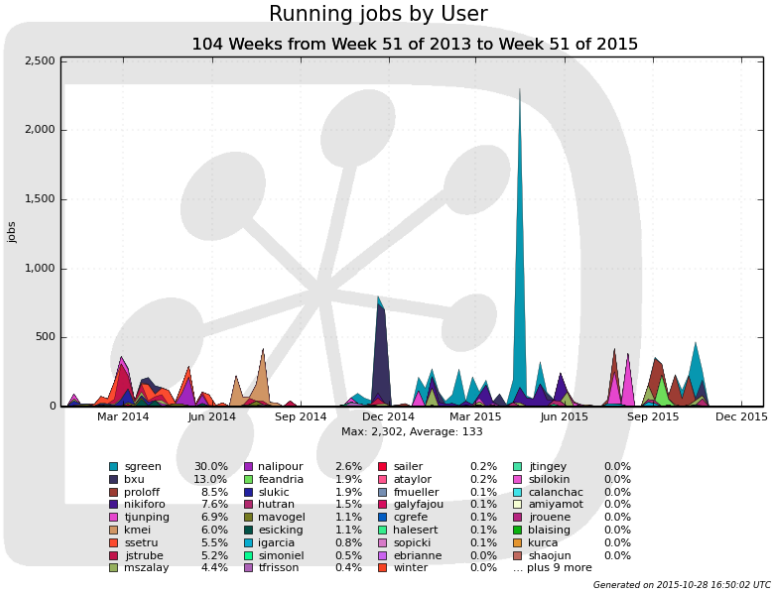
Transformation System

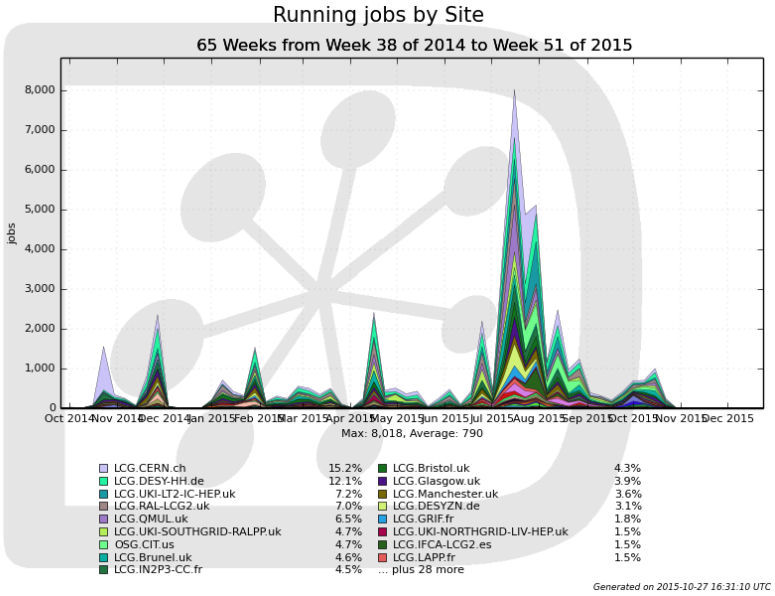
- Easy scalable productions



- You don't have enough Monte Carlo → no problem
- Productions stay in the system and can be re-activated and increased anytime

Performance



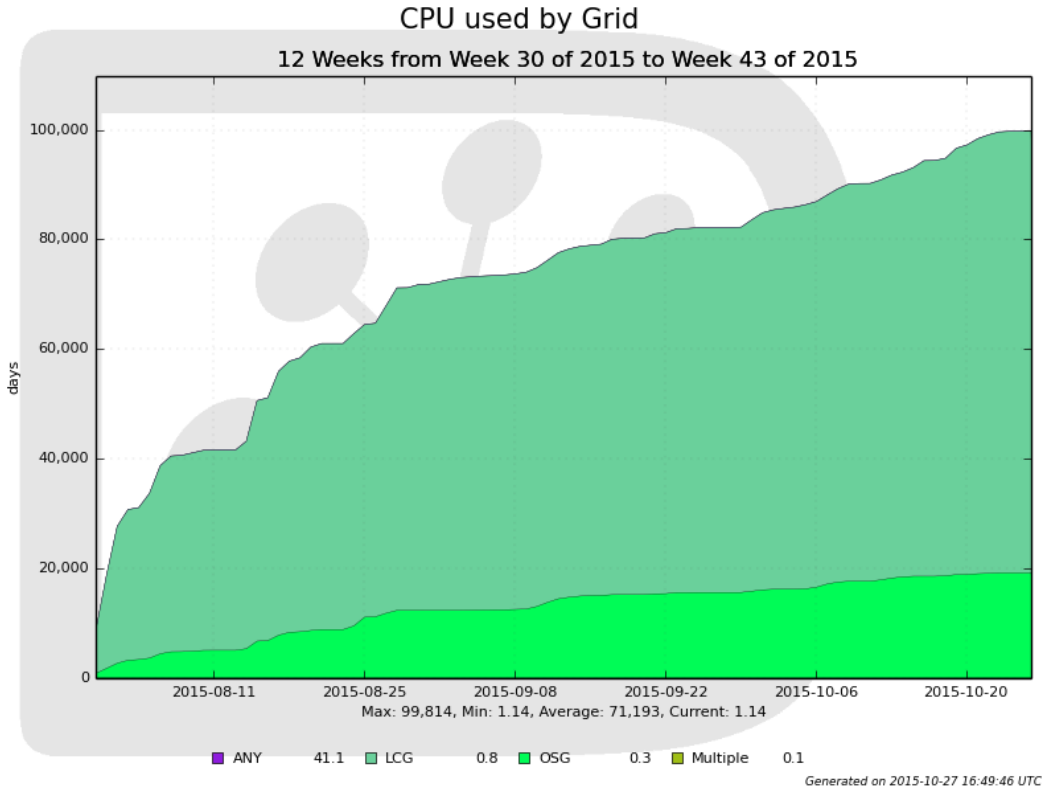


- Widely adopted by CLICdp, ILD, SiD and Calice Collaborations
  - For productions and analysis
- In the last 2 years 40+ users

Site	Contribution
CERN	15.2%
DESY	12.1%
Imperial	7.2%
RAL	7.0%
QMUL	6.5%
…	

Conquest of the New World

- From August 2015 we are running on Open Science Grid

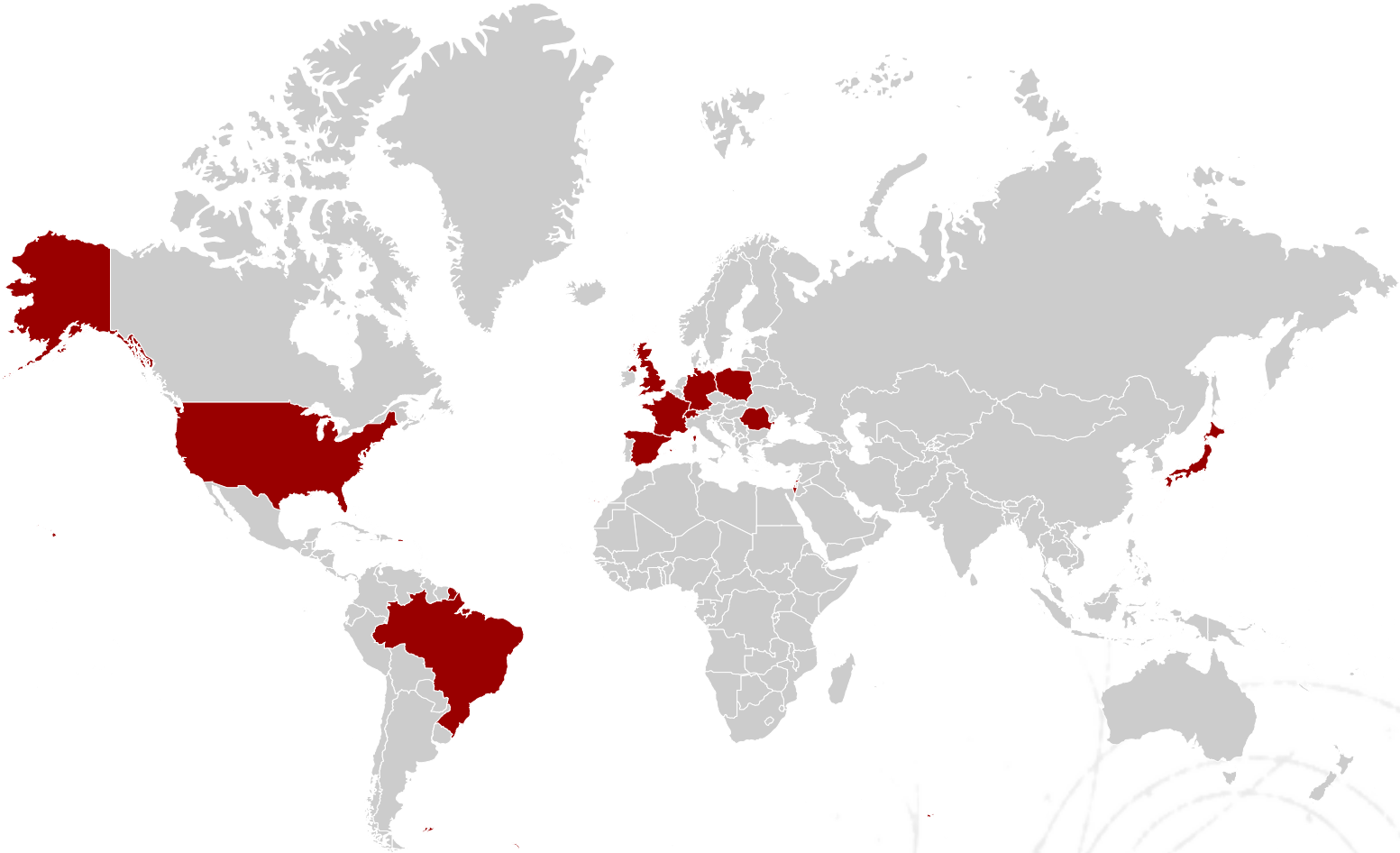


- OSG sites provide 20% of ILC VO and CALICE VO resources now
- At moment not running over cvmfs as on LCG, but deployment should follow soon

Summary

- ILCDirac is offering an easy interface for users to run jobs on the GRID
- Adopted by all detector concepts and enables centralised production of MC
- Easy to use for individual simulation, reconstruction or analyses
- No major changes to user interface foreseen in the future
  - If case of fire:
    1. `twiki.cern.ch/twiki/bin/view/CLIC/DiracForUsers`
    2. Submit a ticket to the issue tracker `https://its.cern.ch/jira/browse/ILCDIRAC`
    3. Email: ilcdirac-support@cern.ch (no e-mails to persons directly)

Do you have something to spare?



- Is your country not coloured in red but you want it to be?
- If you have resources and are willing to share contact me.