Status of GridPP DIRAC

Daniela Bauer & Simon Fayer

Imperial College London

GridPP
UK Computing for Particle Physics

Vogons

This talk bears a resemblance to Vogon poetry as in “it’s not very exciting”.
That’s a good thing:

- It’s a production level service
- We work hard to keep it that way

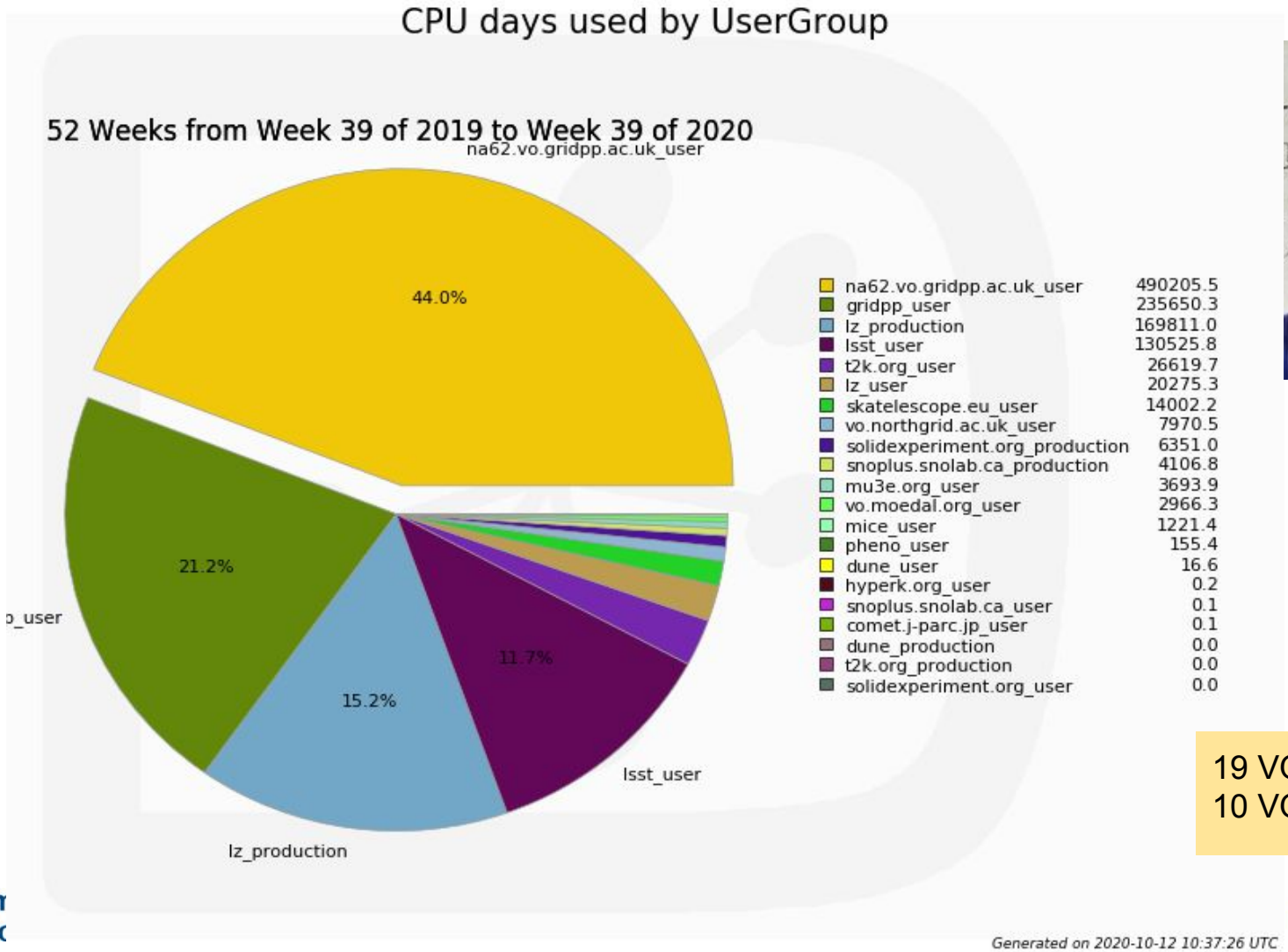
In this talk:

- What is DIRAC (for the new people)
- Usage: Last year in Overview
- DIRAC news
- GridPP DIRAC developments

What is DIRAC ?

- DIRAC is a software originally developed by LHCb that comprises of:
  - **Workload Management System** (“Global batch system”)
    - pilot model: send pilot job to resource/start as part of contextualisation in a VM, if pilot starts successfully pull in real work
  - **File Catalog**
  - Workflow Management System
  - Documentation: https://dirac.readthedocs.io/en/latest/
- Provides a standardized user interface to multiple compute and storage resources
- Written in Python (for Linux)
  - OpenSource: https://github.com/DIRACGrid/DIRAC
- It is used by a number of communities to manage the various aspects of their data processing:
  - **Experiment specific**: e.g. Belle2, ILC, Cherenkov Telescope Array, NICA (JINR), BES (Beijing), biomed
  - **Multi-Community:** e.g GridPP, France-Grilles, EGI

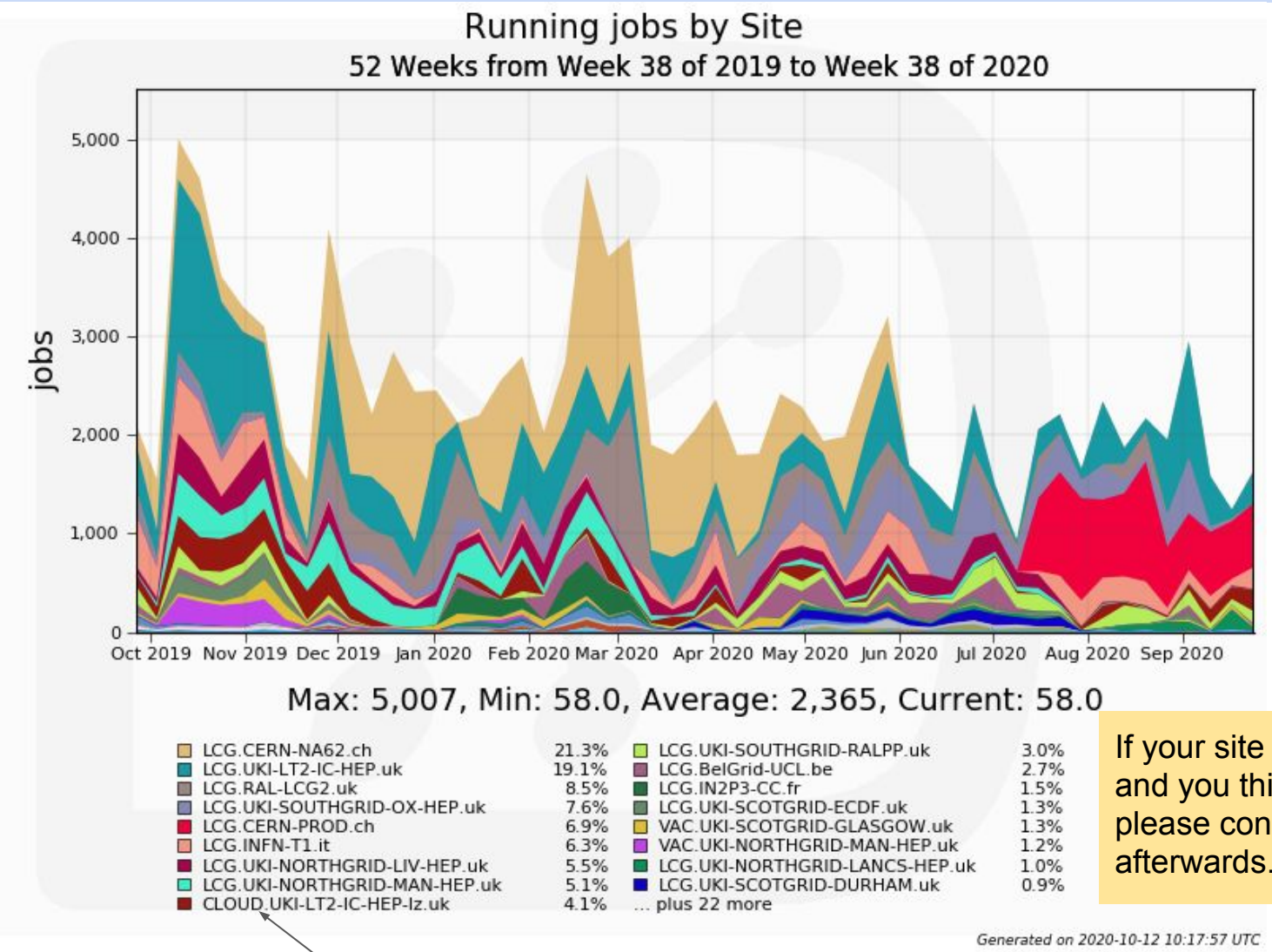
GridPP DIRAC: Oct 2019 - Oct 2020: Usage by Community



19 VOs configured
10 VOs active users

GridPP
UK Computing for Particle Physics

Oct 2019 - Oct 2020 Usage by Site



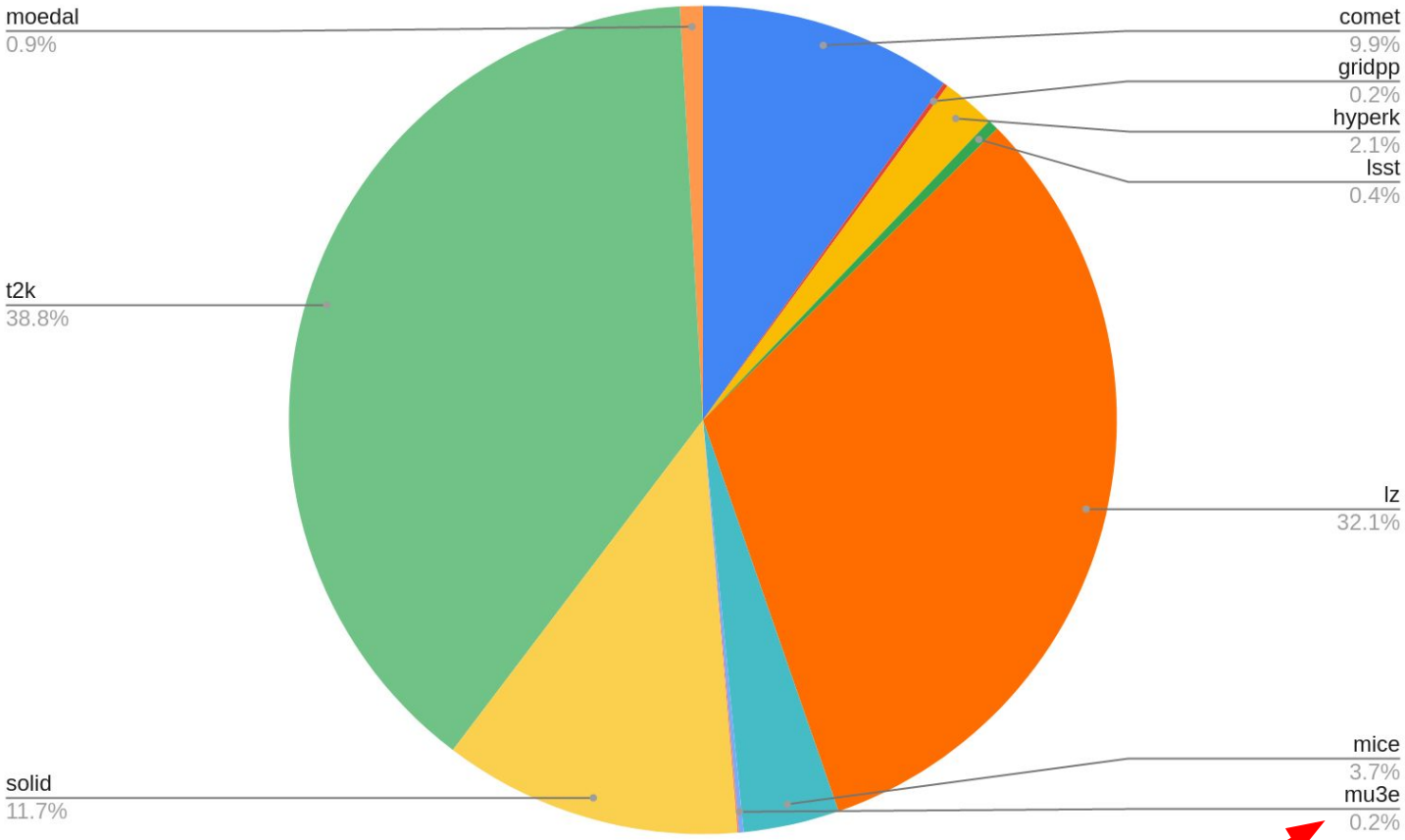
If your site is missing on this plot and you think it should be there, please contact Simon F and me afterwards.

IRIS

Imperial College London

GridPP UK Computing for Particle Physics

File Catalogue: 18 million files and counting.



Our newest community.

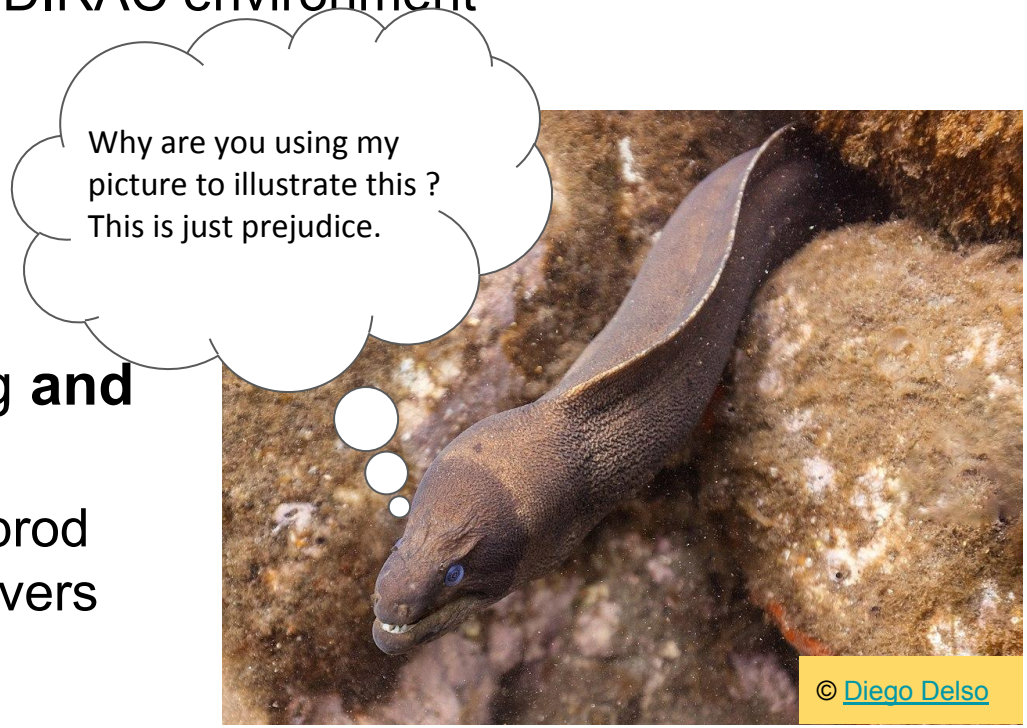
Imperial College London

GridPP
UK Computing for Particle Physics

Things that lurk beneath the surface

DIRAC is under active development:

- pilot2 to pilot3 (“how we configure pilots”)
- lcgBundle to diracos (“how we make DIRAC less dependent on the underlying operating system”)
- CREAM to HTCondorCE/ARC (“learning the finer points of condor”)
- glue1 to glue2 (“where we get our configs from”)
- better separation of user and DIRAC environment
- IRIS digital asset:
  - cloud usage
  - multi-VO metadata
  - multi-VO resource status
  - DIRAC-Rucio
- So, to keep DIRAC running **and** working, we have:
  - one production, one pre-prod and two development servers
    - and we need them all



© Diego Delso

Upgrades

- Major upgrades happen about once a year.
  - Downtimes generally last a couple of hours.
  - Intermittent minor upgrades are done ‘live’.
- The next major upgrade (incorporating the previous slide and more) of GridPP DIRAC is scheduled for November 16th.
  - Extensively tested and fixed.
  - This upgrade is not fully backwards compatible, so we expect some teething problems.

IRIS Digital Assets: Shaking the curse of the single VO DIRAC

- Multi-VO Metadata (Janusz M):
  - The DIRAC File Catalogue now not only keeps the entries for the different VOs separate, but also the metadata.
  - This would allow VO specific automated data management, if requested.

- Resource Status System (Janusz M/work in progress)

Items per page: 100 | Page 1 of 2 | Updated: 2020-10-16 14:44 [UTC](0 00:04)

Name	ResourceT…	Stat…	Status	Reason	DateEffective	LastCheck…	TokenOwner	TokenExpir…
arc-2-kit.gri…	Computing…	all	Banned	110301G0 …	Fri Oct 16 …	Fri Oct 16 …	rs_svc	Fri Dec 31 …
IN2P3-SE	StorageEle…	4 el…	Degraded	Not comple…				
UKI-LT2-IC…	StorageEle…	4 el…	Degraded	Not comple…				

Name	ResourceType	StatusType	Sta…	Reason	DateEffec…	LastChec…	TokenOw…	TokenEx…
UKI-LT2-IC-…	StorageElem…	CheckAccess	Act…	No Down…	Thu Oct …	Fri Oct 1…	rs_svc	Fri Dec 3…
UKI-LT2-IC-…	StorageElem…	ReadAccess	Act…	No Down…	Thu Oct …	Fri Oct 1…	rs_svc	Fri Dec 3…
UKI-LT2-IC-…	StorageElem…	RemoveAcc…	Act…	No Down…	Thu Oct …	Fri Oct 1…	rs_svc	Fri Dec 3…
UKI-LT2-IC-…	StorageElem…	WriteAccess	Error	FreeDisk…	Fri Oct 1…	Fri Oct 1…	rs_svc	Fri Dec 3…

jenkins-mp…	Computing…	all	Active	No DownTi…	Fri Oct 16 …	Fri Oct 16 …	rs_svc	Fri Dec 31 …
TSCatalog	Catalog	all	Active	No DownTi…	Mon Mar 0…	Fri Oct 16 …	rs_svc	Fri Dec 31 …
jenkins-mp…	Computing…	all	Active	No DownTi…	Tue May 1…	Fri Oct 16 …	rs_svc	Fri Dec 31 …
diracci.cer…	Computing…	all	Active	No DownTi…	Tue Mar 03…	Fri Oct 16 …	rs_svc	Fri Dec 31 …
pps-arc-1-k…	Computing…	all	Active	No DownTi…	Fri Oct 16 …	Fri Oct 16 …	rs_svc	Fri Dec 31 …
pps-htcond…	Computing…	all	Active	No DownTi…	Fri Oct 16 …	Fri Oct 16 …	rs_svc	Fri Dec 31 …
ceprod02.g…	Computing…	all	Active	No DownTi…	Thu Oct 15…	Fri Oct 16 …	rs_svc	Fri Dec 31 …

non working CE automatically banned

perfectly working SE wrongly accused of not having any space left

any number of happily working CEs etc

The same component can be working for one VO, but not another, the resource status system needs to take this into account.

Shaking the curse of the single-VO DIRAC

In related news:

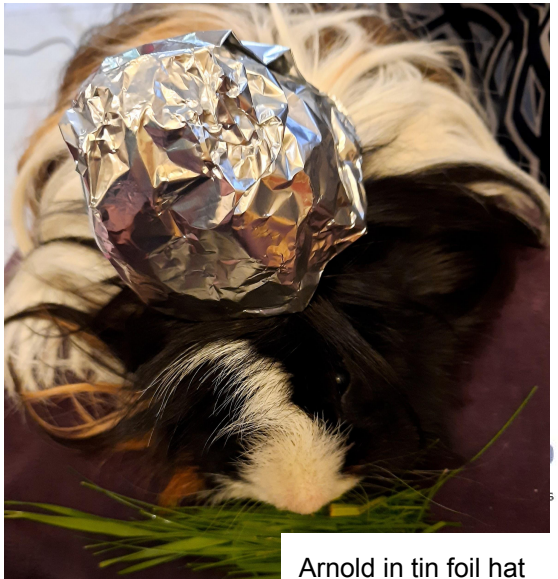
- Core DIRAC has now agreed to make their certification instance multi-VO.
- This will go a long way to make the Core DIRAC multi-VO by default.
- As previous slide has shown, there is still some way to go.

multi^{multi}: multi-VO DIRAC to multi-VO RUCIO interface
(Ditigal Asset/Janusz M)

- Mapping DIRAC objects (SEs, users) to RUCIO objects not straightforward
- Due to delays in multi-VO RUCIO deployment, development against single VO server.
- Proof of concept is working:
  - upload/download/list files: These are the main user interactions on DIRAC
  - But: depends on RUCIO library, replace with requests (HTTP) library wrapper
- DIRAC interface is generic, i.e. can work with any RUCIO server

IRIS Digital Asset: DIRAC and Openstack

- Starting point: Shiny new IRIS Openstack Cloud and VO already using DIRAC
  - added cloud-init (“the industry standard multi-distribution method for cross-platform cloud instance initialization” -- their words, not mine) support in DIRAC
  - added Singularity support to be able to run a container in the cloud instances (like cernvm, but with better payload isolation)



Arnold in tin foil hat

DIRAC: Outlook and Plans

- Keep DIRAC working: Maintenance/Upgrades/Certification/User support.

- Work with developers to ensure upgrades are compatible with GridPP DIRAC
  - Next big update: python2 to python3
  - diracos2: dirac install by conda ?

- Most interesting for users: https based RPC interface in development
  - Replaces proprietary diset protocol
  - enables easier integration with other (experiment) software

Conclusions

- GridPP DIRAC has enabled a large number communities to effectively access grid and cloud resources in the UK and elsewhere.
- IRIS Digital Assets used to develop features in response to user requirements.
- multi-VO DIRAC becoming the norm
- Operations/Upgrades/User support keeps us busy

- Site admin questions ? lcg-site-admin@imperial.ac.uk
- User questions ? gridpp-dirac-users@imperial.ac.uk