



**EUROPEAN OPEN  
SCIENCE CLOUD**



# DIRAC EGI service

*A. Tsaregorodtsev*

*CPPM-IN2P3-CNRS, France,*

*1DIRAC-Rucio Workshop, 8 October,*

*KEK, Tsukuba*

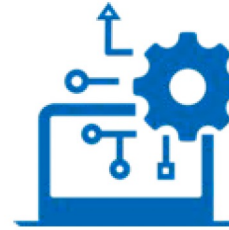
## ▶ EGI-ACE Project:

Implement the Compute Platform of the European Open Science Cloud and contribute to the EOSC Data Commons by delivering integrated computing, platforms, data spaces and tools as an integrated solution.



▶ **EGI Workload Manager service hosting at CC/IN2P3**

- ▶ Openstack VMs
- ▶ 10 production servers:
  - ▶ 8 cores, 16GB RAM
- ▶ 2 development servers
- ▶ 1 Alma Linux VM for the DIRAC certification tests
- ▶ MariaDB server
- ▶ Elasticsearch server



▶ **DIRAC 8.0.24 installed**

- ▶ Python3
- ▶ DIPS services
- ▶ Tornado/HTTPS services: only left AuthServer and TokenManager

- ▶ The EGI DIRAC services are maintained by the members of the DIRAC@IN2P3 Project
  - ▶ CPPM/Marseille, LUPM/Montpellier, CC/IN2P3/Lyon
  
- ▶ DIRAC services
  - ▶ WMS services
    - ▶ Transformation service is enabled but not much used (demos, training)
  - ▶ DMS services
    - ▶ Multiple FileCatalogs (General, Biomed, Eiscat, HESS, Auger, ...)
    - ▶ FTS service was recently requested – will be enabled shortly
  - ▶ RMS
  - ▶ Accounting
  - ▶ Monitoring
    - ▶ Elasticsearch for WMSHistory only

- ▶ Main user communities

- ▶ WeNMR

- ▶ App portal for structural biology

- ▶ Biomed

- ▶ Life sciences, image analysis
    - ▶ Dedicated catalog, specific services

- ▶ Eiscat 3D

- ▶ Climatology
    - ▶ Dedicated catalog with extension for access policies

- ▶ vo.complex-systems.eu

- ▶ Anything complex ;) (multiple users)

- ▶ Pierre Auger Observatory, HESS

- ▶ Astrophysics

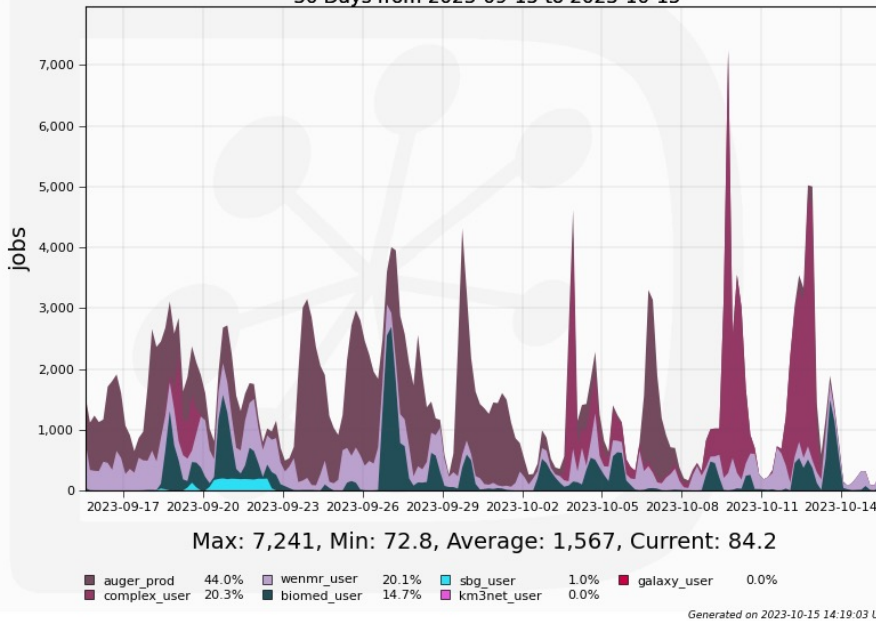


- ▶ New user communities
  - ▶ KM3NeT
    - ▶ Underwater neutrino observatory is enabled recently
    - ▶ Expresses interest in using Rucio for data management
  - ▶ Galaxy application portal (<https://usegalaxy.org>)
    - ▶ Starting to use DIRAC as executor of its workflows
    - ▶ Development of a custom SiteDirector is envisaged
- ▶ Regional communities
  - ▶ E.g. Strasbourg University - [vo.sbg.in2p3.fr](http://vo.sbg.in2p3.fr)
- ▶ In total ~20 communities with over 700 registered users



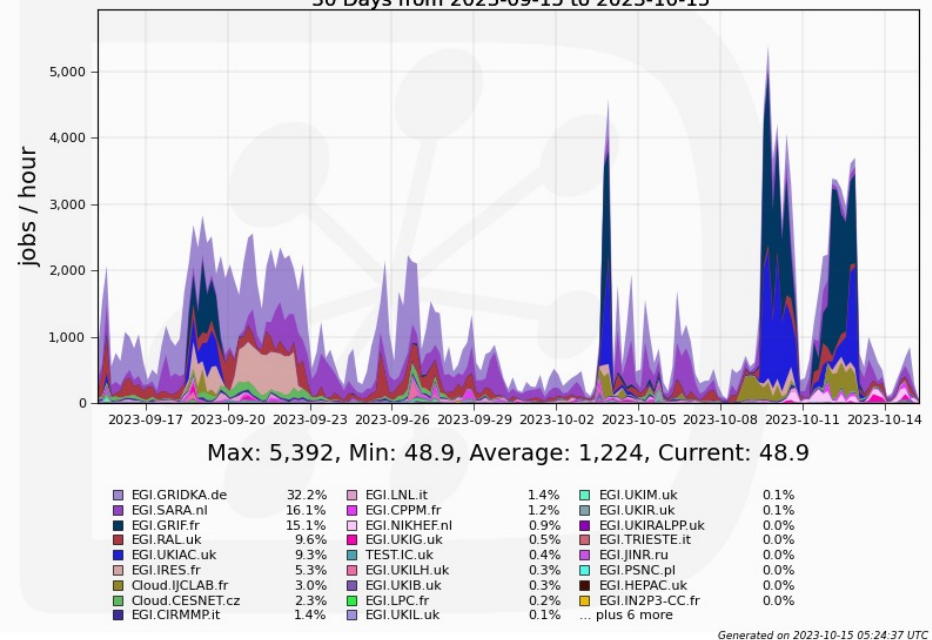
Running Jobs by User Group

30 Days from 2023-09-15 to 2023-10-15



Jobs by Site

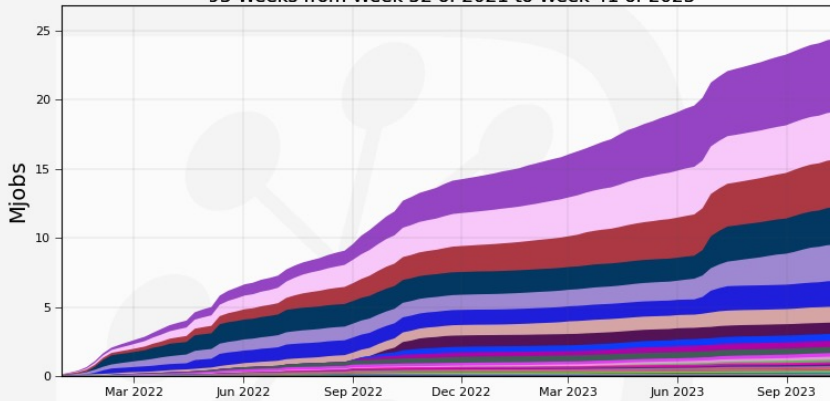
30 Days from 2023-09-15 to 2023-10-15



- ▶ Many small jobs
- ▶ E.g. vo.complex-systems.eu

Cumulative Jobs by Site

93 Weeks from Week 52 of 2021 to Week 41 of 2023



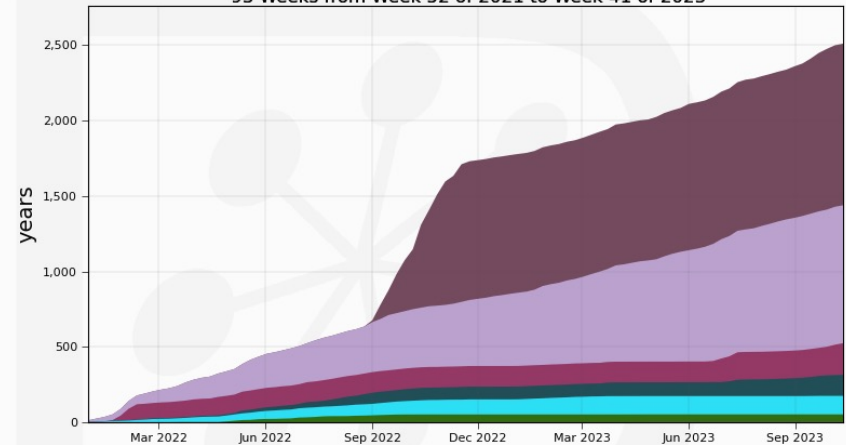
Max: 24.4, Min: 0.13, Average: 12.7, Current: 24.4

EGI.SARA.nl	5.3	EGI.CNAF.it	0.4	Cloud.PADOVA.it	0.1	EGI.UKIM.uk	0.0
EGI.NIKHEF.nl	3.5	EGI.CBPF.br	0.4	EGI.UKILH.uk	0.0	OSG.OSG.us	0.0
EGI.RAL.uk	3.4	EGI.CPPM.fr	0.3	EGI.UKIB.uk	0.0	TEST.CC.fr	0.0
EGI.GRIF.fr	2.7	EGI.LNL.it	0.2	Cloud.JICLAB.fr	0.0	EGI.UKIR.uk	0.0
EGI.GRIDKA.de	2.6	Cloud.CESNET.cz	0.2	EGI.ROMA3.it	0.0	TEST.OSG.us	0.0
EGI.UKIAC.uk	1.9	EGI.UKIG.uk	0.1	EGI.IN2P3-CC.fr	0.0	EGI.DESYZN.de	0.0
EGI.IRES.fr	1.1	EGI.CIRMMMP.it	0.1	TEST.IC.uk	0.0	EGI.UKIL.uk	0.0
EGI.PRAGUE.cz	0.8	EGI.INGRID.pt	0.1	EGI.LPC.fr	0.0	EGI.JINR.ru	0.0
EGI.CESNET.cz	0.5	EGI.TRIESTE.it	0.1	EGI.LAPP.fr	0.0	... plus 19 more	0.0

Generated on 2023-10-13 18:44:32 UTC

Cumulative wall time by UserGroup

93 Weeks from Week 52 of 2021 to Week 41 of 2023



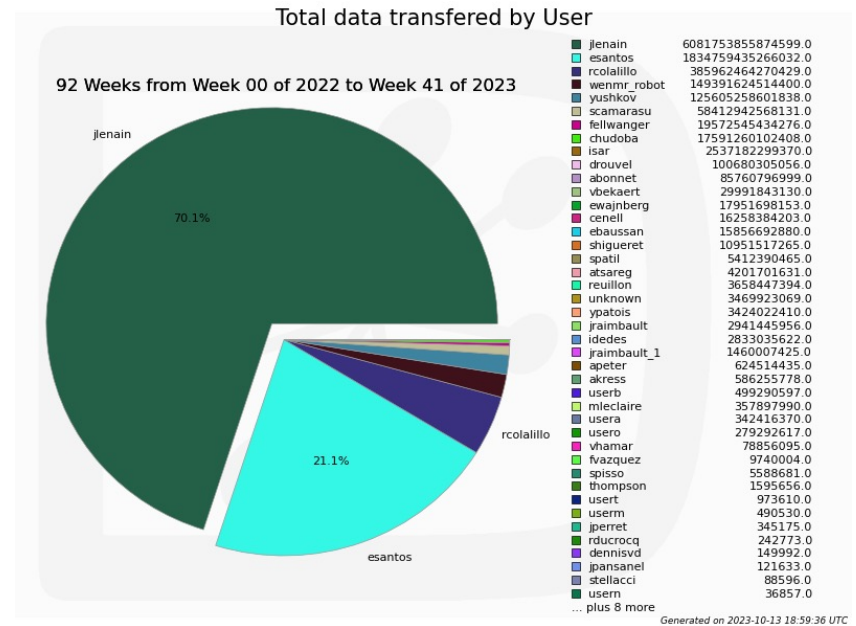
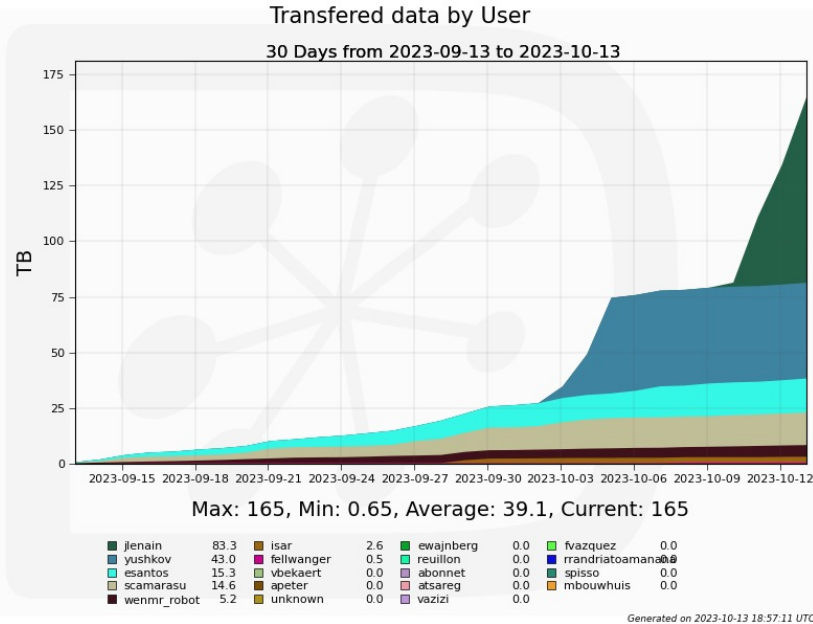
Max: 2,509, Min: 12.9, Average: 1,316, Current: 2,509

auger_prod	1071.0	hess_user	51.6	dirac_tutorial	0.0	galaxy_user	0.0
wenmf_user	914.3	sbg_fluka	0.3	dirac_user	0.0	dirac_test	0.0
complex_user	209.8	km3net_user	0.2	grandest_user	0.0		
biomed_user	138.9	frangillies_user	0.1	escat_common	0.0		
sbg_user	122.7	auger_user	0.1	dteam_user	0.0		

Generated on 2023-10-13 18:48:38 UTC

- ▶ 24M jobs executed since 1.01.2022
  - ▶ 10M jobs in 2023
- ▶ 2.5K years wall time since 1.01.2022
  - ▶ 0.8k years in 2023
- ▶ ~60 different sites (HTCondorCE, ARC)
  - ▶ 15 EGI Fedcloud sites configured, ~8 sites actively used (Openstack)





- ▶ ~9PB data transfers since 1.01.2022
  - ▶ 165TB in the last month
  - ▶ No FTS involved
- ▶ ~15 SE configured (dCache, DPM, EOS)

- ▶ The EGI VO's can define special `eduperson_entitlement` claims that define the VO (as a top level group) and its groups (as subgroups) in the Check-In authorization service
  - ▶ The user tokens can be mapped on to the DIRAC groups
  - ▶ These tokens can be used to communicate with DIRAC services via the Tornado/HTTP protocol
- ▶ The full user job management with tokens was demonstrated
  - ▶ Job submission, sandbox upload, job monitoring, sandbox download
- ▶ This work is suspended as Tornado/HTTP protocol services are not supposed to be further developed/supported.
- ▶ Working with the Check-In developers to ensure access with tokens to the computing resources
  - ▶ Adding `compute.*` scopes to tokens obtained with `client_credentials` flow
  - ▶ Adding `eduperson_entitlement` information to the client token scope to define user's VO and group to allow single DIRAC client to represent multiple VO's
  - ▶ This the work in progress

- ▶ The CVMFS dirac.egi.eu repository is maintained
  - ▶ Admins: Andrei, Daniela
- ▶ The repository is updated automatically nightly:
  - ▶ DIRAC v7r3 releases
  - ▶ DIRACOS2 installers
  - ▶ Pilot files
- ▶ The DIRAC clients for 8.\* releases are installed manually – to be automated
- ▶ Configuration bashrc\_XXX scripts are defined for several DIRAC installations
  - ▶ E.g., bashrc\_egi, bashrc\_gridpp, bashrc\_jinr, etc

# DIRAC in the EGI Jupyter notebooks

---

- ▶ EGI Jupyter service provides notebooks with CVMFS mounted

- ▶ Allows to enable the DIRAC client with a simple:

```
source /cvmfs/dirac.egi.eu/dirac/bashrc_egi
```

- ▶ **dirac-login** command allows to obtain X.509 certificate proxy in the notebook using **device\_flow** Check-In authentication and the DIRAC TokenManager to deliver certificate proxy instead of a token:

```
dirac-login -g biomed_user
```

- ▶ This requires the user's proxy uploaded to the DIRAC ProxyManager service
    - ▶ This mechanism allows to use DIRAC without the need to install the user's certificate in the notebook
    - ▶ Used in several demos, training sessions

- ▶ EGI Workload Manager service was supported by the EGI-ACE project
  - ▶ Finished in June 2023
  
- ▶ It continues on the best effort basis
  - ▶ Service management
  - ▶ Service hosting at CC/IN2P3
  
- ▶ The service will be supported by the GreenDIGIT grant
  - ▶ Starting 1st April 2024
  - ▶ Infra TECH-01-01 call - technologies and practices for sustained RI computing environments

- ▶ Main frustration
  - ▶ Nothing particularly serious
  - ▶ Managing multiple services/agents deployed on a dozen of hosts is rather laborious – looking forward for single entry-point installation with easy logging/monitoring
    - ▶ Fluentbit based solution for logging (Bertrand) can help nicely
  
- ▶ Incidents
  - ▶ Opensearch client compatibility problems when talking to Elasticsearch service still pending
  - ▶ Accounting database cleanup needed to avoid too much unnecessary volume growth
  
- ▶ Need clarity in the upcoming evolution of DIRAC for collaboration with EGI developers
  - ▶ E.g., Check-In usage with the new DIRAC Security Model – part of DiracX
  - ▶ Need well defined storage access protocols and policies accepted by AAI and Storage services providers

- ▶ DIRAC Workload Manager is one of the EGI core services
- ▶ The service demonstrates stable operation in the last years
- ▶ Several developments are ongoing to accommodate the needs of the EGI communities (tokens, custom pilot factories)
- ▶ Expecting multiple advantages from DiracX generation for the multi-VO EGI workload Manager service

- This work is co-funded by the EOSC-hub project (Horizon 2020) under Grant number 777536
- EGI-ACE receives funding from the European Union's Horizon 2020 research and Innovation programme under grant agreement no. 101017567