

# DIRAC Project status

*A. Tsaregorodtsev,  
CPPM-IN2P3-CNRS, Marseille,  
DIRAC@IN2P3 meeting,  
22 May 2023*



- ▶ Status of releases
- ▶ Developments
- ▶ EGI DIRAC service
- ▶ DiracX
- ▶ Conferences

- ▶ 4 patch releases since last meeting
  - ▶ v7r3p37 – no changes
  - ▶ 8.0.23
  - ▶ Minor fixes
    - ▶ Fixes for AREX CE
    - ▶ Fixes in RemoteRunner/PushJobAgent
    - ▶ TornadoSandboxStore()

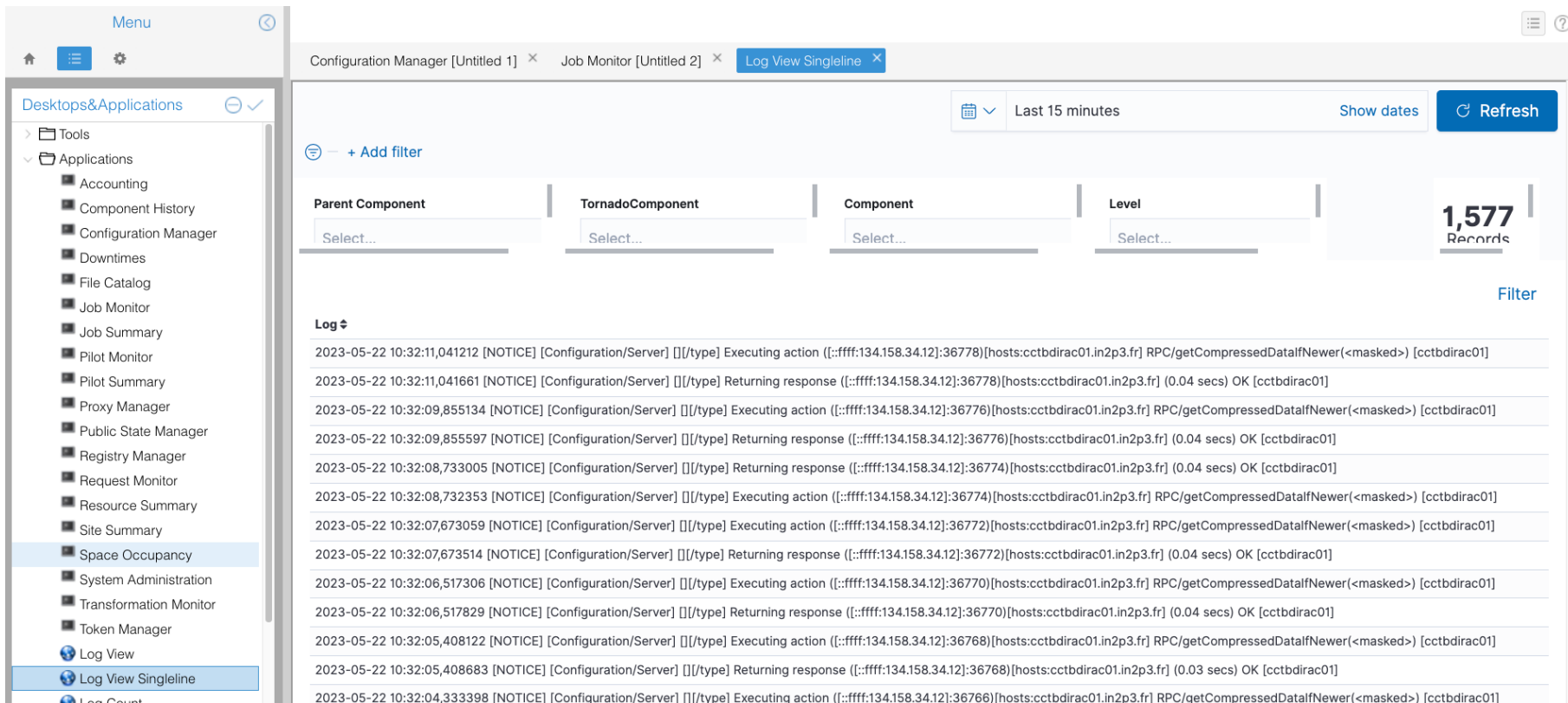
- ▶ 8.1.0
  - ▶ 8.1.0a14 is compiled
    - ▶ Next certification hackathon 6<sup>th</sup> July

## OAuth/OIDC AuthN/AuthZ

- ▶ Biomed use case:
  - ▶ biomed\_user, biomed\_pilot groups defined with corresponding entitlements in Check-In
  - ▶ Submitting jobs with Check-In tokens
    - ▶ OK for HTCondorCE with Check-In validation script
    - ▶ KO at CC without Check-In validation
    - ▶ In the test ARC7: UIOCloud site
- ▶ Other VO's are being examined
  - ▶ Auger (Perun community management)
  - ▶ WeNMR: review together with new applications
  - ▶ Galaxy

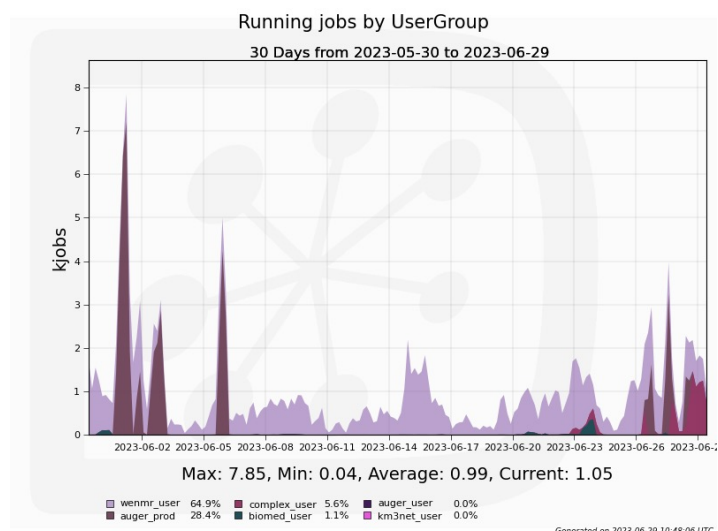
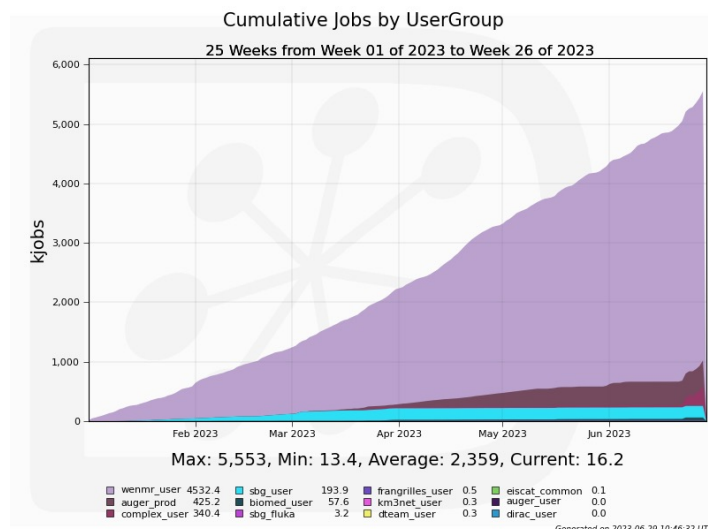
- ▶ WMS with tokens
  - ▶ All the WMS services are at HTTPS at the CC dev testbed
    - ▶ cctbdirac01,02
    - ▶ Including SandboxStore (released in 8.0.23)
  - ▶ All user WMS operations with tokens are demonstrated
    - ▶ Job/sandbox submission
    - ▶ Job monitoring, getting results as sandboxes
  - ▶ To be done
    - ▶ Getting user token in the pilots

- ▶ FluentD/bit logging advancing (Bertrand)
  - ▶ Kibana dashboards
  - ▶ Local log files per service (!)



The screenshot displays the DIRAC Log View Singleline dashboard. On the left, a sidebar menu lists various components under 'Desktops&Applications', with 'Log View Singleline' selected. The main panel shows a log view for the 'Configuration Manager' service. It includes filters for 'Parent Component', 'TornadoComponent', 'Component', and 'Level', all set to 'Select...'. A 'Log' section displays a list of log entries with timestamps, log levels (NOTICE), and messages. A 'Filter' button is visible on the right. The top right corner shows '1,577 Records' and a 'Refresh' button. The top navigation bar includes tabs for 'Configuration Manager [Untitled 1]', 'Job Monitor [Untitled 2]', and 'Log View Singleline'.

- ▶ Running more or less smoothly
  - ▶ Since beginning 2023
    - 5.5M jobs
    - Up to 7k concurrent jobs
      - ~1000 on average last month
- ▶ Servers installation
  - ▶ Smooth server running
  - ▶ Servers are all at 8.0
    - Plus some hacks for tokens support
  - ▶ ES Monitoring not working
    - To be fixed
- ▶ Service maintenance after EGI-ACE end
  - ▶ CC's confirmation for hosting





- ▶ Infra TECH-01-01 call (~5M) **GreenDIGIT** – technologies and practices for sustained RI computing environments
  - ▶ Rumours are that it is likely to be accepted
    - ▶ 5 proposals left for 5 slots
- ▶ PEPR STEEL accepted
  - ▶ DIRAC participation on the level of 15PMs for 3 years
  - ▶ Data placement and data-aware job scheduling

- ▶ ESCAPE

- ▶ NTR

- ▶ Galaxy

- ▶ Advances in discussion during the EGI Conference
  - ▶ They are willing to deploy ARC7 on their HPC's
  - ▶ Willing to make DIRAC as one of the backends
    - ▶ Got instructions on how to set up Pilot Factory and customize pilot submission and job matching
  - ▶ Let's see

- ▶ DIRAC Community workshop at the EGI Conference 2023 (suggested by EGI)
  - ▶ 18-23 June 2023
  - ▶ Demo during a coffee break
  - ▶ 90 mins training session
    - ▶ Demo using Jupyter terminal
    - ▶ 2 persons but motivated (WeNMR, EGI)
  - ▶ GDB report
    - ▶ Check-In tokens for pilot submission
    - ▶ Check-In tokens to interact with the DIRAC services
  - ▶ Multiple discussions
    - ▶ Data access, AuthZ, community support

- ▶ ISGC'23 : paper submitted
  - ▶ DIRAC: OIDC/OAuth2 based security framework
    - ▶ Waiting for reviewer's report
  
- ▶ CHEP
  - ▶ Production system (Alice)
  - ▶ Paper status ?
  
- ▶ DiracX Hackathon (see below)
  - ▶ 4-5 July at CERN
  
- ▶ Joint DIRAC + Rucio Workshop
  - ▶ 16-20 October 2023
  - ▶ KEK

- ▶ Consortium Agreement Extension Amendment goes on
  - ▶ Signing procedure started: already signed by CNRS, IC, CERN (!)

- ▶ Proposal for a complete revision of the DIRAC implementation
  - ▶ See presentations at <https://indico.cern.ch/event/1297397>, <https://indico.cern.ch/event/1292287/> and <https://indico.cern.ch/event/1285339/>
- ▶ Résumé:
  - ▶ DIRAC is a successful project, but
  - ▶ It is at the end of its technology cycle
  - ▶ To continue the DIRAC's success story we have to update the underlying technology

- complex, with high entrance bar
  - got better dropping python2 compatibility
- somewhat cumbersome deployment
  - got better dropping python2 compatibility
- late on “standards”
  - http services
  - tokens
  - monitoring
- “old”-ish design (RPC, “cron” agents...)
- not very developer-friendly
  - rather un-appealing/confusing, especially for new (and young) developers
- multi-VO, but was not designed to do so since the beginning
- no clear interface to a running DIRAC instance

- **Done:** Python 3
  - py3 clients supported since version 7.2 (pip installable)
  - py3 server supported since version 7.3
  - py2 support ended with 8.0 (released last week)
    - with some obvious exceptions of part of pilots code
- **Done:** ES/kibana/grafana dashboards
- **Ongoing/advanced:** dips:// → https://
  - dips: DIRAC proprietary protocol for RPC calls
  - http: based on [tornado](#)
  - most DIRAC services already available using HTTP
    - we said that http would be the default for all the DIRAC services from version 8.1
- **Ongoing:** token support, and IdP (IdM, Check-in)
- **Ongoing:** running on kubernetes (goal: define a *helm* chart)
- **Started:** using celery and RabbitMQ (retiring executors)

## issues with HTTPS in DIRAC

- started (too) long time ago, not yet finished and even less adopted
- the horse (tornado) is feeling old-ish
- there's still an RPC call to be made at the end, no routes
- we are using our own tornado fork because of GSI support



## The list can go on

- the WebApp is highly custom, and somewhat un-maintainable
  - with an intermediate python layer
- runsv is “dead”, we create the RPMs...
- DIRAC's plotting is “old-ish”
- Moving to JSON serialization quite painful
- Upgrades are not always easy, and sometime scary
- ...
- we have been accumulating problems for years
- out there the world evolved in different directions (e.g. REST APIs)



- ▶ 2 hackathons to get developers onboard
  - ▶ 1-2 June, CERN
    - ▶ First shock introduction
  - ▶ 4-5 July, CERN
    - ▶ Getting hands dirty
    - ▶ Precising the migration plan
    - ▶ Taking decision to go ahead (or no)
    - ▶ Distributing tasks
  
- ▶ DIRAC@IN2P3 role: we have to see who, where and how to contribute
  - ▶ Data Management
  - ▶ Transformation/Production
  - ▶ WebApp
  - ▶ AuthZ
  - ▶ Overall testing, getting EGI service ready

- ▶ Who is attending:
  - ▶ Andrei, Vanessa, Bertrand, ...
  
- ▶ Getting ready
  - ▶ See requirements at the hackathon page  
<https://indico.cern.ch/event/1292289/>
  
- ▶ DiracX repo in Github
  - ▶ <https://github.com/DIRACGrid/diracx>
  
- ▶ We should quickly set up a dev environment
  - ▶ Common
  - ▶ Individual