DIRAC

Current, upcoming and planned capabilities and technologies

CHEP 2023

May 9th 2023

Alexandre F. Boyer

alexandre.boyer@cern.ch

Federico Stagni

federico.stagni@cern.ch

European Organization for Nuclear Research
Meyrin, Switzerland

DIRAC
THE INTERWARE

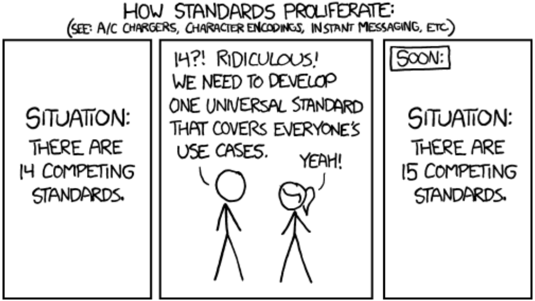
Introduction

Have access to remote computing/storage resources but don't know how to use them?

There exist many paradigms and implementations to interact with shared computing and storage resources.

- Batch systems (BS), Computing Elements (CE), Storage Elements (SE), communication protocols...

DIRAC aims at providing an abstraction layer between different user communities and these different resources.

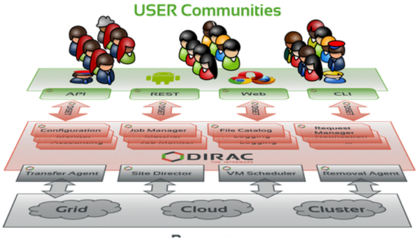

HOW STANDARDS PROLIFERATE:
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC)
SITUATION: THERE ARE 14 COMPETING STANDARDS.
14?! RIDICULOUS! WE NEED TO DEVELOP ONE UNIVERSAL STANDARD THAT COVERS EVERYONE'S USE CASES.
YEAH!
SOON:
SITUATION: THERE ARE 15 COMPETING STANDARDS.

DIRAC

Brief description of DIRAC

An open source middleware for distributed computing

- Started as an LHCb project.
- Experiment-agnostic since 2009.
- Developed by communities, for communities.
    - GitHub hosted
    - Publicly documented, active assistance forum, yearly users workshops, open developers meetings and hackathons.



Installations and communities

LHCb
BESIII
JUNO
CEPC
FUTURE CIRCULAR COLLIDER
ilc
CLIC
cta cherenkov telescope array
PROJECT 8
Belle II
GridPP UK Computing for Particle Physics
EGI-ACE
DutchGrid Large-scale Distributed Computing in The Netherlands
EISCAT 3D
T2K
SNO+
PIERRE AUGER OBSERVATORY
JENNIFER 2 EU grant n.822070
Fermi Gamma-ray Space Telescope
NA62
we-nmr
NICA
SKA SQUARE KILOMETRE ARRAY
OpenMOLE
LSST

A framework shared by multiple experiments/projects, both inside HEP, astronomy, and life science: Experiment agnostic, Extensible, Flexible.

You can find us in CHEP2023

Getting the big picture

A few presentations/posters about DIRAC:

- Standardizing DIRAC's Cloud Interfaces
- Improved Pilot Logging in DIRAC
- Analysis Productions: A declarative approach to ntupling
- LbMCSubmit: A new flexible and scalable request submission system for LHCb simulation
- Integrating LHCb workflows on Supercomputers: State of Practice
- ARC and the EuroScienceGateway project
- Migration to WebDAV in Belle II Experiment
- The Cherenkov Telescope Array Observatory workflow management system

SUPPORTED RESOURCES

Computing Resources

Where to run the jobs

DIRAC embeds plugins to interact with various computing resources:

- Clusters (Plugin: `SSHCE` + `BatchSystem` interface): Orchestrated by a BS. Generally accessed through an SSH/GSISSH tunnel.
- Grid Sites (Plugin: `HTCondorCE`, `ARCCE/AREXCE`): Clusters with specific policies, accessed through a CE.
- HPC Sites: Clusters with additional constraints. More details in this presentation: Integrating LHCb workflows on Supercomputers: State of Practice
- Cloud resources (Plugin: `CloudCE`): More details on this poster: Standardizing DIRAC's Cloud Interfaces.
- Volunteering resources (Plugin: `BOINCCE`): BOINC Volunteer resources.
- Locally (Plugin: `LocalCE`).

Storage Resources

Where to store the results

DIRAC also support various protocols to interact with storage resources:

- S3: e.g. AWS and CEPH.
- SRM, XROOT, HTTPS, GSIFTP: using GFAL2.
- RFIO protocol (deprecated)
- "File": abstraction of the local storage as an SE.
- DIP: DIRAC Custom Protocol.

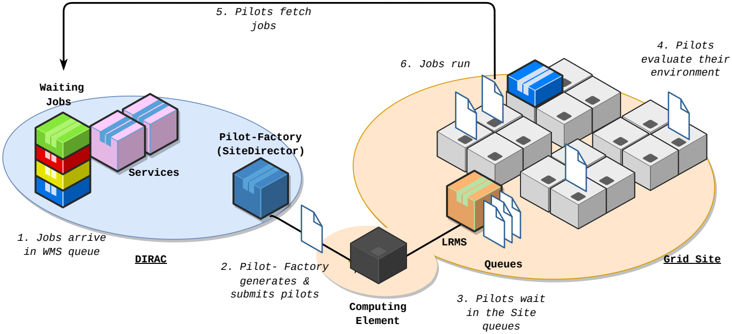
Several abstractions of the same physical endpoint are possible (Multi-protocol). Storage occupancy information can be fetched from BDII or WLCG Accounting.

SYSTEMS

Workload Management System (WMS): Transferring jobs to computing interfaces

Basics of DIRAC WMS

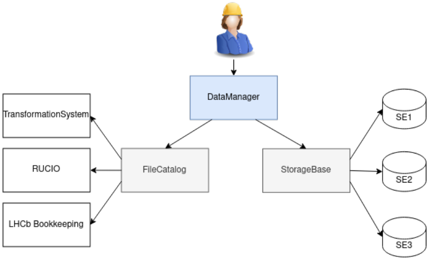
- Push model: Error-prone, but reintroduced to exploit HPCs with no external connectivity.
- Pull model: Pilot-Job paradigm is the most used way of submitting jobs.
- Vacuum model: (HLT Farm, VAC).


5. Pilots fetch jobs
Waiting Jobs
6. Jobs run
4. Pilots evaluate their environment
Pilot-Factory (SiteDirector)
Services
1. Jobs arrive in WMS queue
DIRAC
2. Pilot- Factory generates & submits pilots
Computing Element
LRMS
Queues
Grid Site
3. Pilots wait in the Site queues

Data Management System (DMS): Transferring data to storage interfaces

Basics of DIRAC DMS

- LFNs (Logical File Name): unique identifier of a file within DIRAC.
- LFNs may have physical replicas, stored in SEs.
- LFNs are registered in catalog(s). There exist multiple implementations of catalogs. Several of them can live in parallel:
  - DIRAC File Catalog: full replica and metadata catalog.
  - Plugins for DIRAC TS, LHCb Bookkeeping, RUCIO.
- DMS integrates FTS3 to schedule and monitor efficient transfer of large amounts of data between SEs.



DIRAC+Rucio

DIRAC & Rucio

Some VOs using DIRAC would like to use Rucio as DMS (and maybe some VOs using Rucio would like to use DIRAC WMS).

- Discussions started at the 8th DIRAC workshop (May 2018).
- Few developments done on both sides: integration of (multi-VO) DIRAC with (multi-VO) Rucio. In Progress
- Since January 2021 Belle2 uses DIRAC and Rucio (from LCG file catalog to Rucio FC).

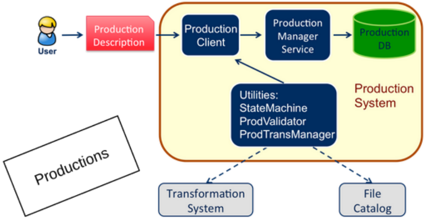
If interested, we will held a DIRAC & Rucio Workshop, details here: **https://indico.cern.ch/e/DR23**

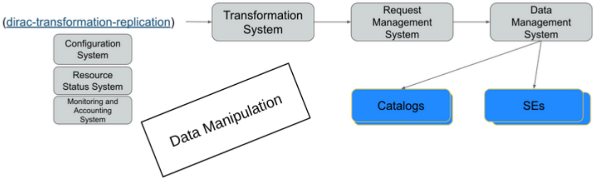
Transformation System: Job productions and datasets management

Used to automate common tasks related to production activities

- Production: "Data Processing" transformation (e.g. Simulation, Merge, DataReconstruction...). It ends up creating jobs in the WMS.
- Data Manipulation: transformation to replicate, or remove data from storage elements. It ends up creating requests in the RMS (Request Management System), which feeds the DMS.

The TransformationSystem is finely tuned and can manage millions of jobs and files daily





Accounting & Monitoring

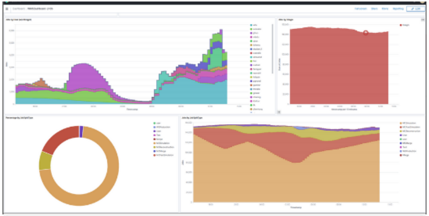
Accounting

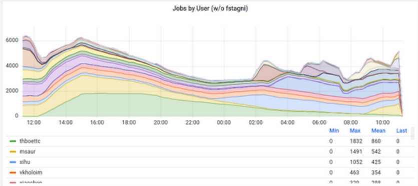
- For historic data: jobs, pilots, data operations, storage
- MySQL backend, visualized in DIRAC WebApp

Monitoring

- Real Time monitoring and not only
- OpenSearch backend, visualized in Kibana, Grafana and (partially) DIRAC WebApp New: v8.0

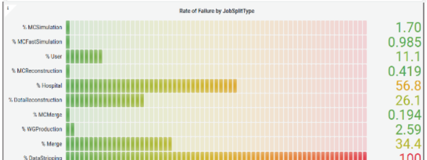
Settling with Grafana as next-gen visualization tool, also for historic data In-Progress





Jobs by User (w/o fstagni)

	Min	Max	Mean	Last
thboettc	0	1832	860	0
msaur	0	1491	542	0
xihu	0	1052	425	0
vkholoim	0	463	354	0



Rate of Failure by JobSplitType

Authentication & Authorization

From X509 certificates to OIDC Tokens

DIRAC v8 rationalizes many aspects to AuthN, AuthZ, Tokens and OAuth2 support.

- Support new Identity Providers (IAM & EGI CheckIn) New: v8.0 .
- Use tokens (and/or proxies) to submit jobs/pilot-jobs New: v8.0
- Use tokens (and/or proxies) to interact with storage resources (v8.x)

Framework

Highlights of the latest developments

- Transitioning service communications from DIPS (DIRAC in-house protocol) to HTTPS. In Progress
- Replacing the in-house task queue system (Executors) by Celery, a widely used task queue system. In Progress
- Centralizing logs coming from Pilot-Jobs. More details on this poster: Improved Pilot Logging in DIRAC In Progress

DEVELOPMENT

Development, testing and deployment

Development

~5 FTE as core developers, a dozen contributing developers.

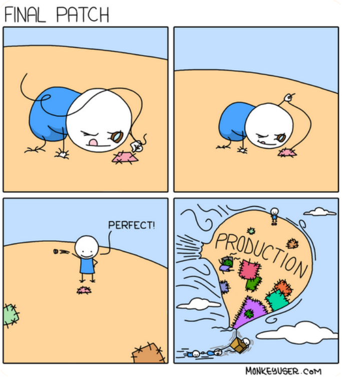
Testing

Tests, certification, integration process is a daily work.

- We use GitHub Actions, and Jenkins for some bits.
- We run certification hackathons every two weeks.

Deployment

Puppet profiles used for long time, helm chart available on request.



CONCLUSION

Last details

Additional resources

- Documentation: dirac.readthedocs.io.
- Code documentation: here
- Dev and DevOps issues: on GitHub
- Ops and general questions: GitHub discussions
- Bi-weekly developers meetings (and/or hackathons): BILD

DIRAC & Rucio Workshop

- 16-20 October 2023 in KEK, Japan.
- Registration and details: **https://indico.cern.ch/e/DR23**

Questions ? Comments ?

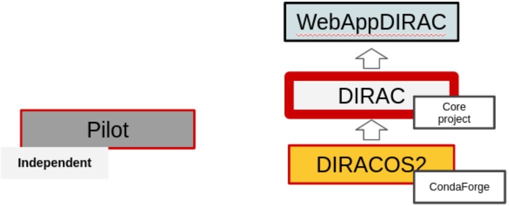
BACKUP

DIRAC extensions

"Horizontal" extensibility

-

For specific requirements



WebAppDIRAC

DIRAC

Core project

DIRACOS2

CondaForge

Pilot

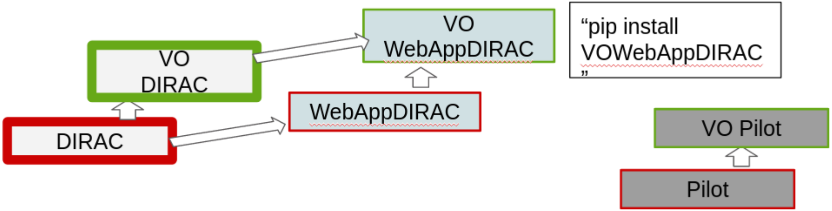
Independent

Each project is independently versioned

"Vertical" extensibility

-

Community driven



VO DIRAC

DIRAC

VO WebAppDIRAC

WebAppDIRAC

"pip install VOWebAppDIRAC"

VO Pilot

Pilot

Resource Status System

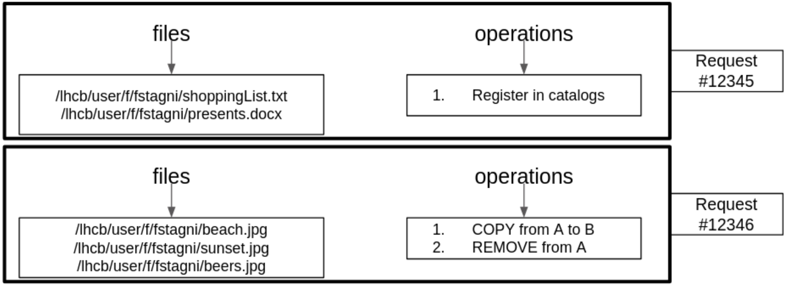
Presentation

- Stores info on the status of Resources (e.g. SEs)
- An autonomic computing tool evaluates a few policies to determine the status of the resources. E.g.:
  - Space left < threshold → ban for writing
  - Endpoint in downtime in GocDB → ban r/w
- DIRAC SEs states are sync-ed from DIRAC RSS to Rucio via a DIRAC agent

Request Management System

RSS

A generic system, which can be used for queueing (also) DMS operations. Operation types: ReplicateAndRegister (e.g. using FTS), RemoveFile/RemoveReplica, ...


files
/lhcb/user/f/fstagni/shoppingList.txt
/lhcb/user/f/fstagni/presents.docx
operations
1. Register in catalogs
Request #12345
files
/lhcb/user/f/fstagni/beach.jpg
/lhcb/user/f/fstagni/sunset.jpg
/lhcb/user/f/fstagni/beers.jpg
operations
1. COPY from A to B
2. REMOVE from A
Request #12346

and the Request Executing Agent will… execute the requests.