DIRAC framework for distributed computing

L. Arrabito\textsuperscript{1}, J. Bregeon\textsuperscript{1}, P. Gay\textsuperscript{2}, V. Hamar\textsuperscript{3}, F. Hernandez\textsuperscript{3}, S. Pop\textsuperscript{4}, A. Tsaregorodtsev\textsuperscript{5}

\textsuperscript{1}LUPM-IN2P3/CNRS, Montpellier
\textsuperscript{2}Université de Bordeaux
\textsuperscript{3}CC-IN2P3/CNRS, Lyon
\textsuperscript{4}Creatis CNRS, Lyon
\textsuperscript{5}CPPM-IN2P3-CNRS, Marseille

Journées SUCCESS 2017, 16\textsuperscript{th}-17\textsuperscript{th} October, Grenoble
Plan

- DIRAC overview
- Managing computing and storage resources
- Users communities
- DIRAC as a service
- Conclusions
The DIRAC interware

- A software framework for **distributed computing**
- Intermediates between **users and resources**
- Allows **interoperability** simplifying **interfaces**
The project

- Originally developed to support the production activities of the LHCb experiment at CERN (~10 years ago)
- Today is a general purpose software, targeting several large scientific communities
- Open source project developed by user communities for users
  - [http://github.com/DIRACGrid/DIRAC](http://github.com/DIRACGrid/DIRAC)
  - Publicly documented, active assistance forum, yearly users workshops, open developers meetings
- The DIRAC consortium as representing body
  - CERN, CNRS, University of Barcelona, KEK, IHEP, University of Montpellier as members
- In 2017 started as IN2P3 project to support further software generalization
  - CPPM, CC-IN2P3, LUPM, Creatis/CNRS, University of Bordeaux
Main DIRAC components

- Workload Management System (WMS)
  - Job brokering with Pilot Jobs
  - Interoperability with different types of computing resources

- High level workflows Management System (Transformation System)
  - Support for automated massive data production and processing

- Data Management System
  - Storage management (access to various storage systems)
  - Data bookkeeping (File Catalog)

- These components should be installed on dedicated services to which clients connect
Command line tools, Python API, RESTful interface

Web portal
Implementation of Pilot Jobs
- Introduced by the DIRAC project, now widely used in various WMS
- Efficient usage of resources
- Transparent access to different types of resources
  - grids, clouds, clusters, etc.
Conceived for **Production Managers**

Enables automated *workflow execution* handling *large datasets*, *e.g.*:

- Data Processing -> Merging -> Data Analysis -> Data Replication

Workflows as chains of ‘data transformations’

Transformation = input data filter + recipe to create tasks (jobs or data operations)

Transformations are created by the Production Manager

Tasks are automatically created through as soon as input data are registered in the system

- Fully data-driven system
Resources available via the DIRAC system
DIRAC was initially developed with the focus on accessing conventional Grid computing resources

- It fully supports *multiple grid middleware and infrastructures*
  - EGI, WLCG, OSG, NorduGRID, etc
- Other types of grids can be supported
  - As requested by users

**Standalone clusters**

- Access through SSH/GSI/SSH tunnel
- Batch systems supported: LSF, SGE, PBS/Torque, Condor, OAR, SLURM

**BOINC Volunteer resources**
VMDIRAC module

Allows transparent access to various private or public cloud resources
- Apache-libcloud
- Rocci
- EC2
- Others are in the works
  - OCCI, Google, Azur, IBM, ...
  - Preferring RESTful interfaces

Manages the whole VMs life cycle
- Creation, Monitoring, Discarding
Multiple HPC centers are available for large scientific communities
  - E.g., HEP experiments started to have access to a number of HPC centers
    - Using traditional HTC applications
    - Filling in the gaps of empty slots
    - Including HPC into their data production systems

Advantages of federating HPC centers
  - More users and applications for each center - better efficiency of usage
  - Elastic usage: users can have more resources for a limited time period

Collaboration with HPC centers to integrate them into a common framework under the DIRAC WMS control
  - France: Aix Marseille University
  - Russia: Dubna, NNGU, others
  - China: IHEP HPC center, Beijing
Unlike grid sites, HPC centers are not uniform
- Different access protocols
  - Different user authentication methods
- Different batch systems
- Different connectivity to outside world

DIRAC work in progress to overcome these differences
- Support HPC access protocols (SSH, GSISSH, ARC, OAR, SLURM)
- Methods for remote control of user payloads and data at the HPC centers
  - Site proxy/gateway services
  - User data import and export
High Energy Physics

- LHcb
- RHIC
- Belle II
- CLIC
- BES III

Astrophysics

- CTA

And many others

- FRANCE GRILLES
- eGI
- Fermi
- T2K
- LSST
- JUNO
- Geant4
- CERN@school
- CERN & Society
- NA62
- Large Synoptic Survey Telescope
- Life Science Grid Community
More than 100K concurrent jobs in ~120 distinct sites

- This is not the limited by the system capacity, but by the available resources
- In needed, further optimizations to increase the capacity are possible
  - Hardware, database optimizations, service load balancing, etc.
DIRAC dedicated installations

- Belle II Collaboration, KEK
  - First use of clouds (Amazon) for data production
- ILC/CLIC detector Collaboration, Calice VO
  - DIRAC File Catalog was developed to meet the ILC/CLIC requirements
- BES III, IHEP, China
  - Dataset management developed for the needs of BES III
- CTA
  - Contributing to Transformation System development enabling data-driven workflows
- Geant4
  - Validation of MC simulation software releases
- DIRAC evaluations by other experiments
  - LSST, Auger, TREND, Daya Bay, Juno, ELI, NICA, …
  - Evaluations can be done with general purpose DIRAC services
5 out of Top-10 EGI communities used heavily DIRAC for their payload management in the last year

- lhcb, belle, biomed, ilc, vo.cta.in2p3.fr
- compchem will likely join the club
Getting started

- For a full evaluation
  - Install a server instance dedicated to your community

- The easiest
  - Connect to a running **DIRAC service** (see next slides)
    - Just need to register and install the DIRAC client on your laptop

- Doc/Support
  - Documentation
    - [http://diracgrid.org/](http://diracgrid.org/)
  - User Forum
    - [https://groups.google.com/forum/?hl=en#!forum/diracgrid-forum](https://groups.google.com/forum/?hl=en#!forum/diracgrid-forum)
DIRAC as a service
DIRAC as a service

- DIRAC framework was updated to support multi-VO installations
  - Allows to provide better support for several small user communities with a single DIRAC instance
  - Keeps the costs of operating DIRAC under control
- Several services provided by national grid infrastructure
  - FG-DIRAC, France
  - GridPP, UK
    - [https://www.gridpp.ac.uk/services/gridppdirac/](https://www.gridpp.ac.uk/services/gridppdirac/)
- DIRAC4EGI
  - [http://dirac.egi.eu/DIRAC/](http://dirac.egi.eu/DIRAC/)
  - Starting from 2018 DIRAC becomes a Core Service of EGI
  - Serving both Grid and FedCloud resources
Joint effort to provide FG-DIRAC service
- Hosted by CC-IN2P3
- Distributed team of service administrators
  - 5 participating institutes (CC-IN2P3, CPPM, Creatis, U. Bordeaux, LUPM)

Usage
- 21 Virtual Organizations
  - robot users
    - VIP/GateLab Biomed
- About 12 million jobs processed every year
Conclusions

- Distributed computing is no more something exotic, it is used in a daily work by users in various scientific domains.

- DIRAC provides a framework for building distributed computing systems aggregating multiple types of computing and storage resources.

- Several large scientific collaborations adopted DIRAC for their production systems. Multiple evaluations are ongoing.

- Increasing number of projects providing ‘DIRAC as a service’ (multi-community).

- In 2018 DIRAC will become an EGI core service, replacing gLite WMS.

- DIRAC can help users to get started in the world of distributed computing and discover its full potential.

http://diracgrid.org
Backup
Storage plugins

- Storage element abstraction with a client implementation for each access protocol
  - DIPS, SRM, XROOTD, RFIO, etc
  - gfal2 based plugin gives access to all protocols supported by the library
    - HTTP, DCAP, WebDAV, S3, …

- Each SE is seen by the clients as a logical entity
  - With some specific operational properties
  - SE’s can be configured with multiple protocols
Managing VM life cycle

- **VM creation** through a *CloudDirector* (similar to grid jobs)
  - Based on Task Queue status
    - If there are waiting user payloads
    - VM properties corresponding to payload requirements
- **VM contextualization**
  - On the fly installation of DIRAC, CVMFS, …
  - Starting as many pilots as they are cores (single core jobs)
- **Starting the VM Monitor Agent**
  - Monitor and report the VM state, VM heartbeats
  - Halt the VM in case of no activity
  - Getting instructions from the central service, e.g. to halt the VM
- **VM Scheduler** orchestrates spawning and halting virtual machines depending on the Task Queue status, Accounting history
  - Necessary for fair sharing of cloud resources
  - Work in progress
Accessing HPC resources

- Pilot submitted to the batch system through an (GSI)SSH tunnel
- Pilot communicates with the DIRAC service through the Gateway proxy service
- Output upload to the target SE through the SE proxy
HPC resources allow a rich description with respect to traditional grid resources

- DIRAC work in progress to develop a more elaborated model of their description with the corresponding payload matching mechanisms

- Worker node micro-management
  - Single-core and multi-core applications
  - Multi-processor, multi-node applications are in the works
Computing Resources: HPC

- Multi-core job scheduling
- Pilots with partitionable internal slots
  - M-core Pilots pull N-core jobs (N\leq M) until internal slots used up
  - Pilot is standard-size, can be whole-node, 4-node, 8-node....
  - Optimizing CPU efficiency
In “best effort” production since 2014

Partners
- Operated by EGI
- Hosted by CYFRONET
- DIRAC Project providing software, consultancy

10 Virtual Organizations
- enmr.eu, vlemed, eiscat.se
- fedcloud.egi.eu
- training.egi.eu

Usage
- Workload Management solution
  - > 6 million jobs processed in the last year
- Data Management solution
  - E.g. Eiscat 3D

Starting from 2018 DIRAC becomes a Core Service of EGI
- WMS replacement
- Serving both Grid and FedCloud resources