DIRAC Services



A. Tsaregorodtsev, CPPM-IN2P3-CNRS

EGI Community Forum, Bari, 11 November 2015



- DIRAC Project brief overview
- DIRAC services
- DIRAC4EGI service
- Conclusions



DIRAC provides all the necessary components to build ad-hoc grid infrastructures interconnecting computing resources of different types, allowing interoperability and simplifying interfaces. This allows to speak about the DIRAC interware.



Interware: WMS

- Pilot based Workload
 Management
 - High user job efficiency
 - Suitable for usage with heterogeneous resources
 - Allowing application of community policies
- Pilot 2.0 framework
 - Modular for easy customizations for different environments, communities, etc







Pilots 2.0 framework

- Modularity: pilots are designed as a configurable sequence of commands
- Extensions: each community can provide custom commands for specific operations
 - Environment checks
 - Software installation
 - Monitoring reports



LHCb example: "pilots to fly in all the sky's"



ISGC2015, F.Stagni and Ph.Charpentier, CERN



Computing resources available via pilots

Grids

- gLite/EMI: EGI (CREAMComputingElement)
- VDT: OSG (GlobusComputingElement)
- ARC: NDGF (ARCComputingElement)
- Standalone clusters
 - Using SSH/GSISSH tunnel
 - LSF, BQS, SGE, PBS/Torque, Condor(G), ...
 - E.g. Yandex computing farm
 - HTCondor ComputingElement is in the test
- HPC centers:
 - OAR, SLURM, using DIRAC proxy servers to move data in and out
- Clouds
 - VM scheduler for EC2, OCCI, Nova, libcloud
 - Amazon, OpenNebula, OpenStack, CloudStack, Stratuslab
 - VAC, Vcycle
- BOINC
 - IDGF
 - Standalone, e.g. LHCb@HOME

. .

VM Scheduler developed for Belle Data Production System

- Dynamic VM spawning driven by the user workload
- Discarding VMs automatically when no more needed
- Support for various cloud technologies
 - Amazon, OCCI, OpenNebula, OpenStack, CloudStack, Stratuslab
- DIRAC is an active participant of the EGI FedCloud activities
 - FedCloud sites are accessible through the DIRAC4EGI service





Clouds



PUSP Proxies

Support for the Per User SubProxies (PUSP)

- To be in line with the EGI developments
- Used for the DIRAC tutorial
- Can be used for scientific gateway portal users

Each PUSP user is registered in person in DIRAC

- Per User policies (group membership)
 - E.g. File Catalog access rights
- Per User Accounting





- Command line clients
- REST interface
 - Useful to use as a WMS engine for Application portals
 - VIP, WS-PGrade, WeNMR. Etc
- Web Interface
 - Web Portal development framework
 - Tornado CMS, ExtJS GUI
 - Secured with X509 certificates
 - Desktop GUI paradigm
 - Natural for non-expert users
 - Support of most of the user tasks (jobs, data, monitoring, management)



D

Web Portal: example interfaces

🗲 \rightarrow C 🔒 https://dirac.ub.edu/CTA/s:CTA/g:cta_user/?theme=Grey&url_state=0 DIRAC.ConfigurationManager.classes.ConfigurationManager::431:352:386:269:0:0,1, 🏠 🚍			
🏥 Apps 🗋 Apple 🗋 Yahoo! 👌 Google Maps 🗈 YouTube 🗋 Wikipedia 🧰 News 🧰 Popular 🥅 Views 🧰 Personal 🛑 DIRAC 🧰 CTA 🥅 UB 🛑 Belle 🛑 Fundación BBVA			
Selectors	Items per page: 100 🗸 🖂	I of 13006 ▶ ▶ Displaying topics 1 - 100 of 1300594	Updated: 2013-10-16 14:49 [UTC]
Sitema and a second sec		Site JobNar LastUpdate [UTC] LastSignOfLife [UTC]	SubmissionTime [UTC] Own
Selected Statistics :: Status (GMT+0200 (CEST))	Wed Oct 16 2013 20:22:59	LCG.CIEMAT.es Sta 2013-10-16 14:21:54 2013-10-16 14:21:54	2013-10-16 14:21:54 th
S Selected Statistics	Completed	LCG.CIEMAT.es Sta 2013-10-16 14:02:06 2013-10-16 14:02:06	2013-10-16 13:55:38 th
Status	Failed	LCG.CIEMAT.es Sta 2013-10-16 14:02:04 2013-10-16 14:02:04	2013-10-16 13:55:28 tł
Completed 18.1%	Other	LCG.DESY-ZEUT Unk 2013-10-16 14:01:08 2013-10-16 14:01:08	2013-10-16 12:33:16 tł
		LCG.CAMK.pl Unk 2013-10-16 12:29:59 20 Proxy Upload	
Failed		LCG.DESY-ZEUT Ast 2013-10-16 10:03:22 20	-
Killed		LCG.DE	×
		Proxy Status: Valid + Add Parameters	either your private key nor
Waiting 81.7%		Predefined Sets of Launchpad Values	our service. While we try to ure as possible by using SSL
	001444	Available Sets	vith your credentials when it
Bunning jobs by Site	CSV data		anually convert and upload
41 Weeks from Week 53 of 2012 to Wee	K		
5,000 -	Direc CTA [2012 10 16 14:29:50 20]	Executable: mandelbrot	4E.p12
4.000	Dirac-C TA [2013-10-16 14:36:59.302	JobName: Mandelbrot_%j	GROUP_NAME
	🕀 🧰 Systems	Arguments: -W 600 -H 600 -X -0.46490 -Y -0.56480 -P 0.	Prouto
	Website	OutputSandbox: *.bmp	browse
2.000 -	Registry	StdError: %j.err	
1.000 -		CPUTime: 3600	d 3 David
	🕀 🧰 SiteLocalSEMapping	StdOutput: %j.out	ld C Reset
jan 2013 Feb 2013 Mar 2013 Apr 2013 May 2013 Jun 2013 Jul 2013 Max: 5,143, Min: 0.00, Average: 600, Current:	3. 🕀 🧰 Shifter		
LCG.CYPRONETp1 46.6% ■ LCG.MFFG.fr 2.3% ■ LCG. LCG.GRIF.fr 42.5% ■ LCG.MFG.Tr 2.3% ■ LCG. LCG.DESY.ZEUTHEN.de 12.0% ■ LCG.MFN.TORINO.It 1.3% ■ LCG.DESY.ZEUTHEN.de 12.0% ■ LCG.MFN.TORINO.It 1.3% ■ LCG.MFN.TORINO.It 1.3% ■ LCG.MFN.TORINO		Browse	Ă
LCG INTERC Fr 39% LCG ONERAFY 04% DAW	Generated on 2013-10-16 14:48:15 UTC	Submit 2 Reset	
🗘 📰 Configuration Man 🔦 Proxy Upload 📰 Account	ing 📰 Job Monitor	📰 Job Monitor 📰 Job Launchpad Theme Gro	ey • ricardo@ cta_user • CTA •



- DIRAC is aiming at providing an abstraction of a single computer for massive computational and data operations from the user perspective
 - Logical Computing and Storage elements (Hardware)
 - Global logical name space (File System)
 - Desktop-like GUI

Running jobs by UserGroup 72 Weeks from 19 of 2015 MC production: stable operations for > 5 months 5,000 User analysis: constant over 4,000 all the year • high I/O á 3,000 2,000 **BESIII Experiment** 1,000 Feb 2014 May 2014 Aug 2014 Nov 2014 Feb 2015 May 2015 Max: 5,434, Average: 1,238, Current: 159

Generated on 2015-05-19 08:41:06 UTC

LHCb, Belle II, BES III, ILC, CTA

cta_prod 85.5% cta_ifae cta_user 14.4% cta_sgm

0.1% 📒 user

0.0%

0.0%

Dedicated installations for the large user communities

CTA MC Production run











National services

- DIRAC services are provided by several National Grid Initiatives: France, Spain, Italy, UK, China, Romania, ...
 - Support for small communities
 - Heavily used for training and evaluation purposes
- Example: France-Grilles DIRAC service
 - Hosted by the CC/IN2P3, Lyon
 - Distributed administrator team
 - 5 participating universities
 - 15 VOs, ~100 registered users
 - In production since May 2012
 - >12M jobs executed in the last year
 - □ At ~90 distinct sites







- In production since 2014
- Partners
 - Operated by EGI
 - Hosted by CYFRONET
 - DIRAC Project providing software consultancy
- IO Virtual Organizations
 - enmr.eu
 - Vlemed
 - fedcloud.egi.eu
 - training.egi.eu
 - • • •
- Usage
 - > 6 million jobs processed in the last year

DIRAC4EGI activity snapshot





Details of Cyfronet instance

- DIRAC services split over 4 VMs + DB server, all with regular backups.
- 1 TB disk space for I/O sandboxes
- · Web Portal address: dirac.egi.eu
- Contact for interested groups : dirac@mailman.egi.eu





- There is a clear need for services like DIRAC for an increasing number of communities with a low expertise in (distributed) computing and with high demands for computing resources
- Important goal is to lower the threshold for scientists of these communities
 - Training is one of the main purposes of all the infrastructures deploying DIRAC services
 - Examples of training projects
 - vo.formations.idgrilles.fr (FG-DIRAC)
 - training.egi.eu (DIRAC4EGI)
 - □ The next DIRAC tutorial is at the EGI Community Forum, November
 - Distributed computing student courses
 - CERN@school project using the GridPP DIRAC service
 - Assistance in porting application to the (DIRAC) grids



Conclusions

- DIRAC is providing a framework to bring together various services and computing resources in a single coherent system
- From the user perspective the whole system is seen as a single computer with an intuitive (graphical) interface
- Multi-VO DIRAC services is an excellent way to open access to distributed computing resources for non-expert user communities.

