DIRAC Services for Grid and Cloud Infrastructures

A.Tsaregorodtsev,
CPPM-IN2P3-CNRS, Marseille,
Plekhanov University of Economics, Moscow
NEC'17, 27 September 2017, Budva

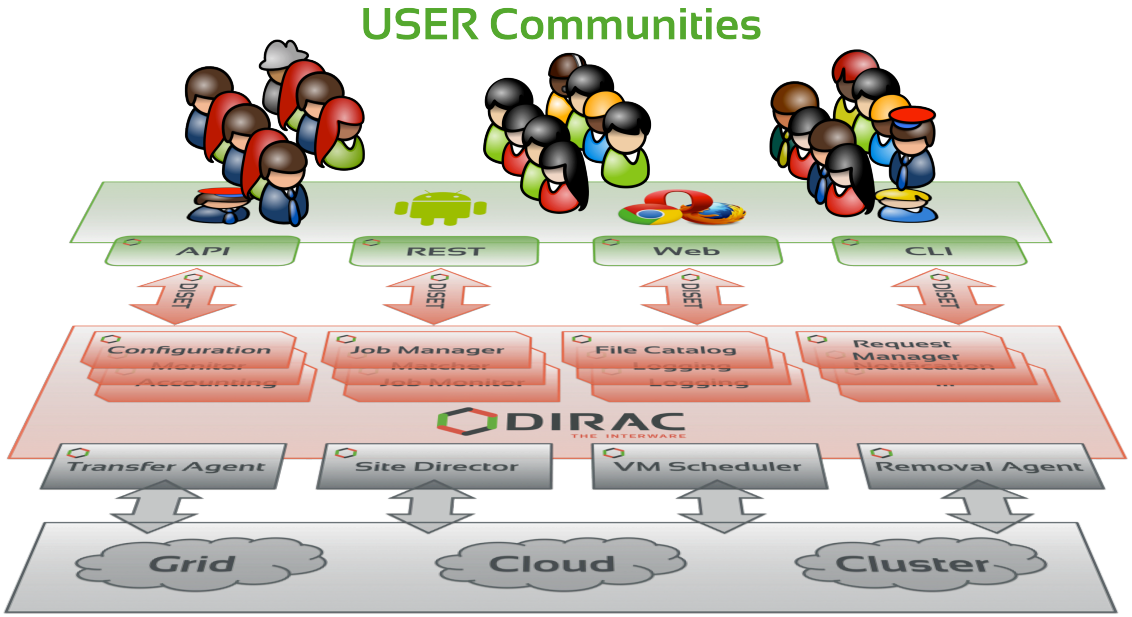
DIRAC
THE INTERWARE

Plan

- DIRAC in a nutshell
- DIRAC communities
- Services for multi-community installations
- Conclusions

Interware

- 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***.



LHCb Computing System

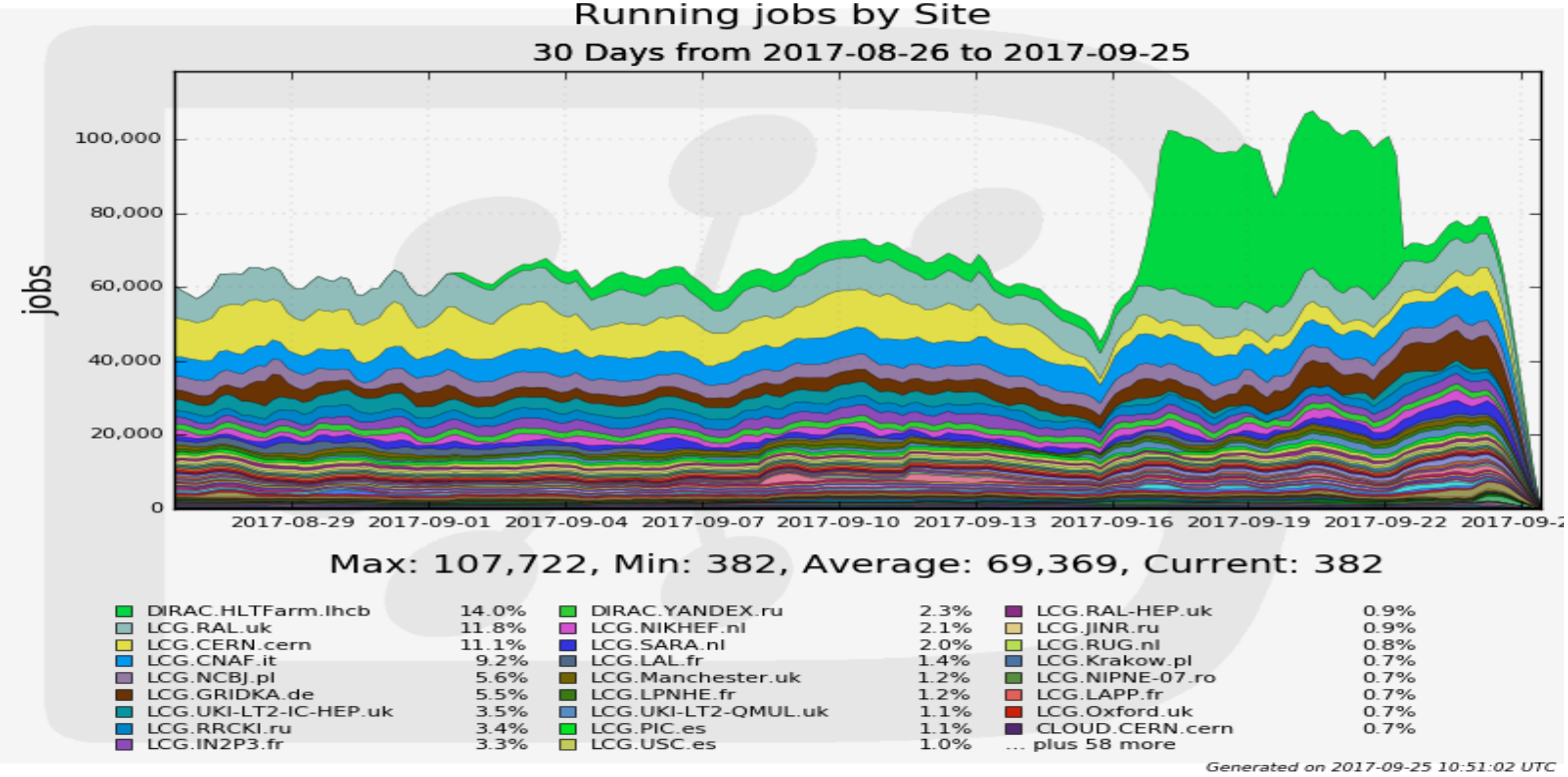
- The project was started in 2003 for the LHCb experiment at CERN
  - First focusing on the workload management for the LHCb data production system
- Now DIRAC is the basis for almost all the aspects of the LHCb distributed computing
  - Workload Management
  - Data Management
  - High level production services
  - Monitoring of resources, activities and services
  - Accounting
  - Interfaces

DIRAC
THE INTERWARE

LHCb Computing

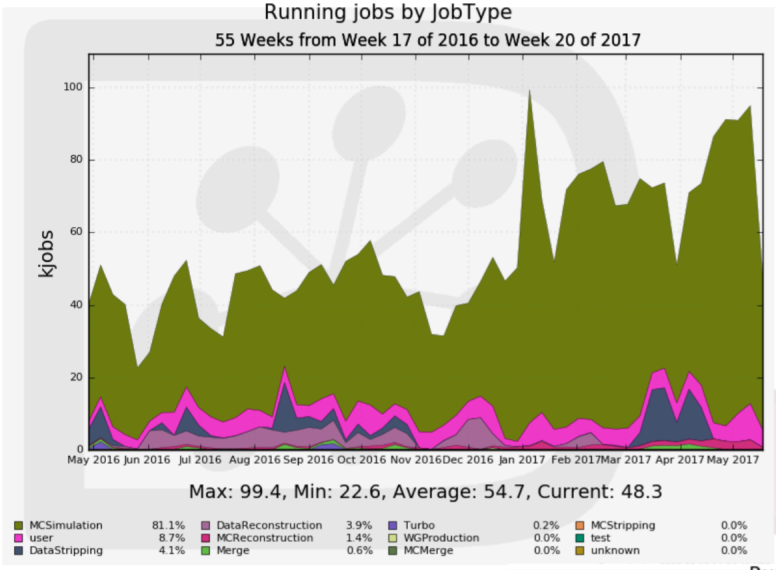
- Supporting various activities
  - MC production
  - Real data processing
  - User analysis
- Accessing all the resources available to LHCb
  - HTC/WLCG
  - Cloud (running DIRAC pilots via Vac/VCycle)
  - BOINC
  - Standalone, e.g.
    - Online HLT farm
    - Yandex (container based) cloud
    - Non-CE clusters

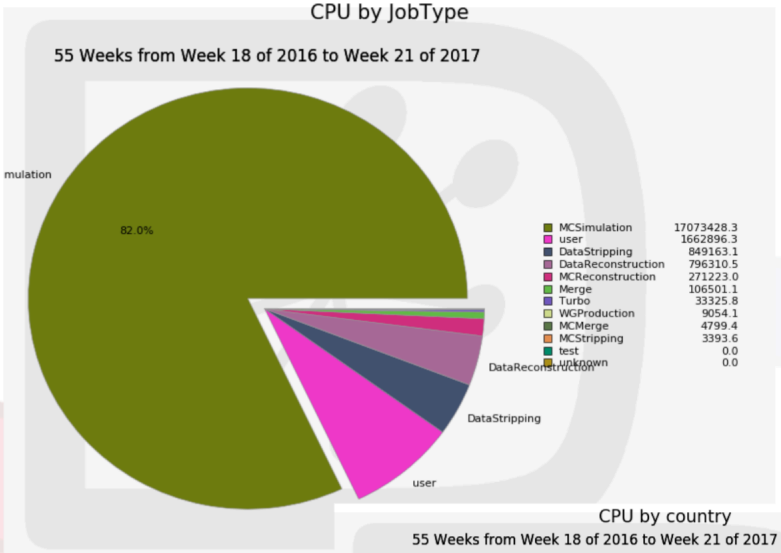
LHCb Collaboration

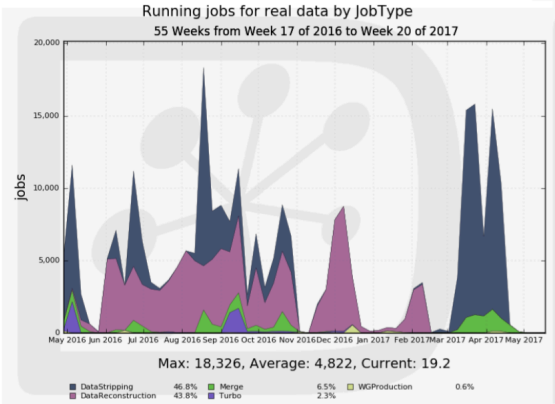


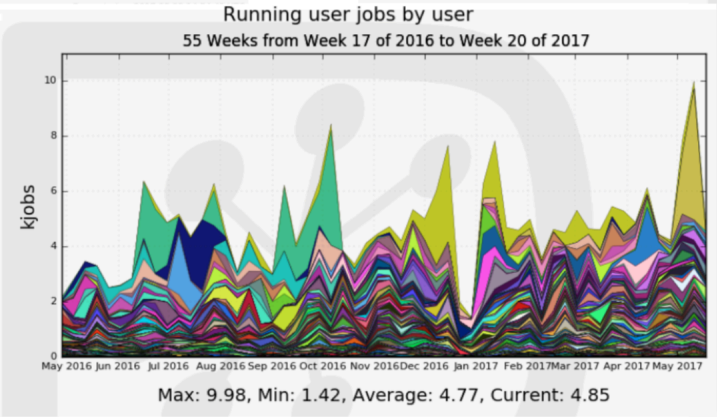
- More than 100K concurrent jobs in ~120 distinct sites
  - Limited by available resources, not by the system capacity
- Further optimizations to increase the capacity are possible
  - Hardware, database optimizations, service load balancing, etc

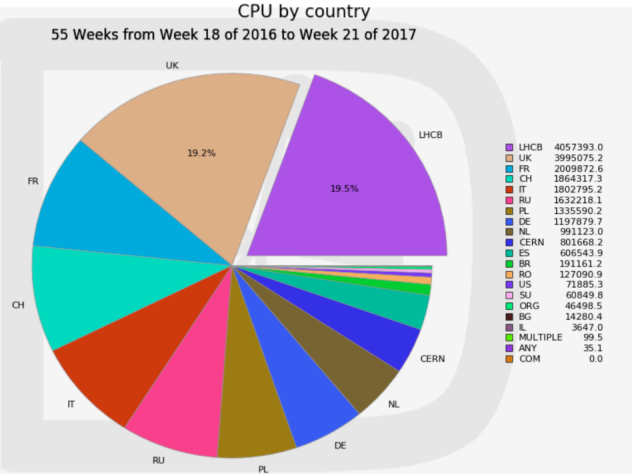
LHCb activities











Core framework

- The experience collected with a production grid system of a large HEP experiment proved to be very valuable
  - Several new experiments expressed interest in using this software relying on its proven in practice utility
- In 2009 the core DIRAC development team decided to generalize the software to make it suitable for any user community.
  - Split the generic functionality from the LHCb specific features
  - Consortium to develop, maintain and promote the DIRAC software was created in 2014
    - Current members: CERN, CNRS, University of Barcelona, IHEP, KEK, University of Monpellier, PNNL
- The results of this work allow to offer DIRAC as a general purpose distributed computing framework

Community installations

- Belle II Collaboration, KEK
  - First use of clouds (Amazon) for data production
- ILC/CLIC detector Collaboration, Calice VO
  - Dedicated installation at CERN, 10 servers, DB-OD MySQL server
  - MC simulations
  - DIRAC File Catalog was developed to meet the ILC/CLIC requirements
- BES III, IHEP, China
  - Using DIRAC DMS: File Replica and Metadata Catalog, Transfer services
  - Dataset management developed for the needs of BES III
- CTA
  - CTA started as France-Grilles DIRAC service customer
  - Now is using a dedicated installation at PIC, Barcelona
  - Using complex workflows
- Geant4
  - Dedicated installation at CERN
  - 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

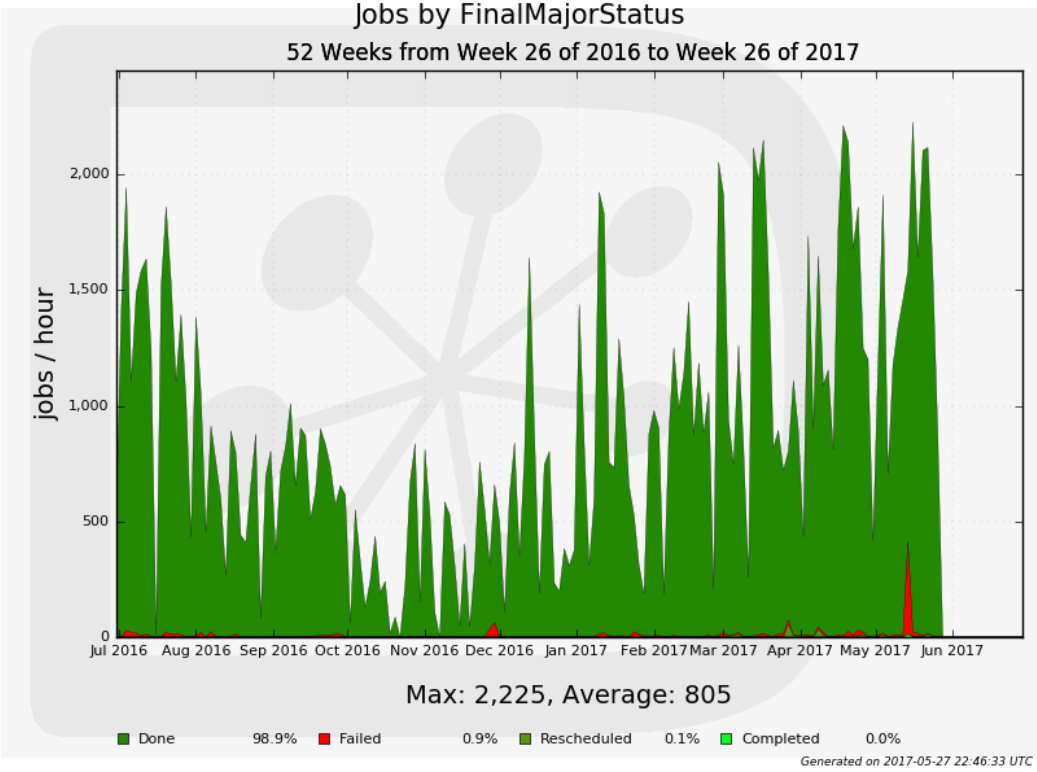
Multi-community services

- Maintaining DIRAC services for small communities is not affordable
  - Need for multi-VO installations
  - DIRAC framework was updated to support this kind of installations
- Several services provided by grid infrastructure projects
  - FG-DIRAC, France
  - GridPP, UK
  - DIRAC4EGI
- Some dedicated installations evolved into multi-community services
  - CERN: ILC, CALICE
  - IHEP: BES III, Juno, CEPC
- Recently added services
  - PNNL: Belle II, Project8, MiniCLEAN, SuperCDMS, nEXO
  - DIRAC@JINR: NICA, Dubna University

DIRAC4EGI service

- 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

DIRAC4EGI activity snapshot



DIRAC
THE INTERWARE

EGI ACCOUNTING PORTAL

Normalised CPU time [units 1K.SI2K.Hours] by DATE and VO												
DATE	alice	atlas	belle	biomed	cms	compchem	ilc	lhcb	virgo	vo.cta.in2p3.fr	Total	%
Nov 2015	83,043,071	213,187,021	29,633,040	2,992,249	107,998,028	812,409	3,051,240	44,495,710	365,193	5,203,790	**490,781,751**	**8.60%**
Dec 2015	81,681,064	167,642,164	30,755,315	2,771,463	81,200,999	1,197,402	10,250,775	42,772,247	4,370	9,643,804	**427,919,603**	**7.50%**
Jan 2016	100,472,899	212,596,116	8,254,706	2,221,994	99,768,667	2,869,544	3,904,455	32,614,451	329,113	8,746,790	**471,778,735**	**8.27%**
Feb 2016	80,340,391	202,531,157	48,965	1,312,309	100,330,129	1,220,127	2,704,948	44,547,976	1,962,465	5,563,528	**440,561,995**	**7.72%**
Mar 2016	108,810,699	172,663,251	3,412,262	2,286,939	75,113,354	1,623,540	2,049,130	83,154,401	1,917,611	1,539,919	**452,571,106**	**7.93%**
Apr 2016	111,707,745	211,516,946	496,969	1,622,314	67,855,621	1,970,394	3,051,624	78,821,567	3,517,152	3,079,316	**483,639,648**	**8.47%**
May 2016	88,434,699	229,055,135	457,771	3,055,283	64,161,648	3,990,478	4,366,309	70,550,242	11,311,493	669,299	**476,052,357**	**8.34%**
Jun 2016	91,963,895	220,222,321	10,039,317	1,375,916	104,040,606	1,755,334	2,097,169	66,545,602	2,558,741	1,103,183	**501,702,084**	**8.79%**
Jul 2016	113,408,142	187,198,001	3,614,046	2,152,445	104,373,741	1,614,892	1,596,155	65,898,735	8,005,698	7,794,153	**495,656,008**	**8.69%**
Aug 2016	88,278,412	212,942,846	34,225	6,500,219	51,366,225	3,474,177	5,538,912	72,803,805	2,919,127	5,410,036	**449,267,984**	**7.87%**
Sep 2016	88,164,653	309,040,532	7,314,602	514,897	90,018,815	2,602,763	3,297,430	106,365,999	1,770,213	6,487,567	**615,577,471**	**10.79%**
Oct 2016	68,902,764	167,532,717	1,528,430	467,733	82,329,281	1,301,416	5,324,702	71,019,670	2,752,272	104,325	**401,263,310**	**7.03%**
Total	**1,105,208,434**	**2,506,128,207**	**95,589,648**	**27,273,761**	**1,028,557,114**	**24,432,476**	**47,232,849**	**779,590,405**	**37,413,448**	**55,345,710**	**5,706,772,052**	
Percentage	**19.37%**	**43.91%**	**1.68%**	**0.48%**	**18.02%**	**0.43%**	**0.83%**	**13.66%**	**0.66%**	**0.97%**		

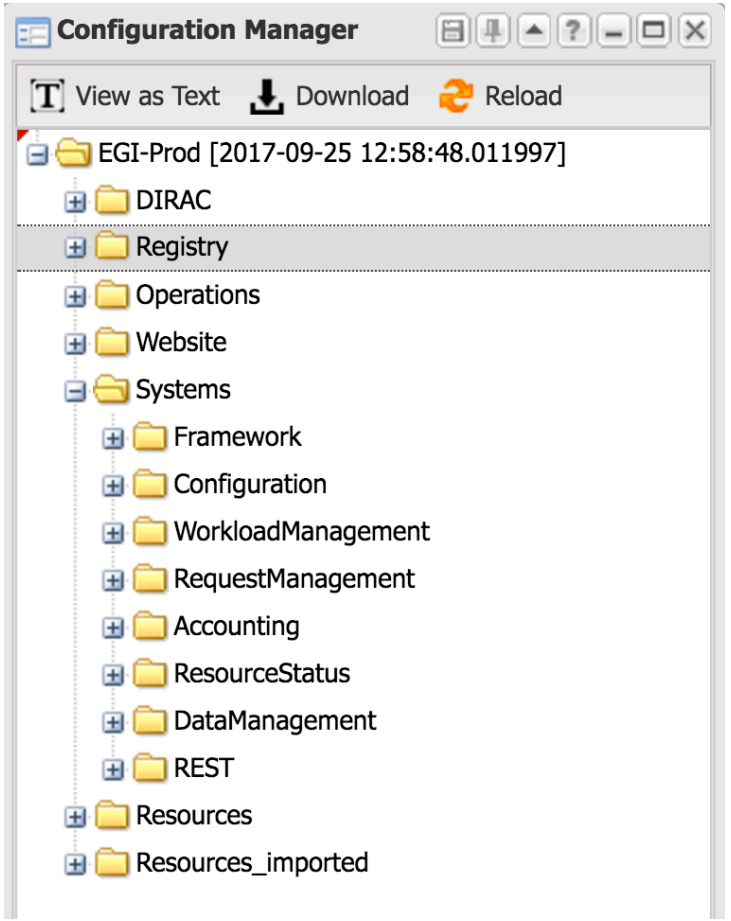
- 5 out of Top-10 EGI communities used heavily DIRAC for their payload management in the last year
  - 4 out of 6 top communities excluding LHC experiments
    - belle, biomed, ilc, vo.cta.in2p3.fr
    - compchem will likely join the club soon

DIRAC
THE INTERWARE

DIRAC Services for EGI

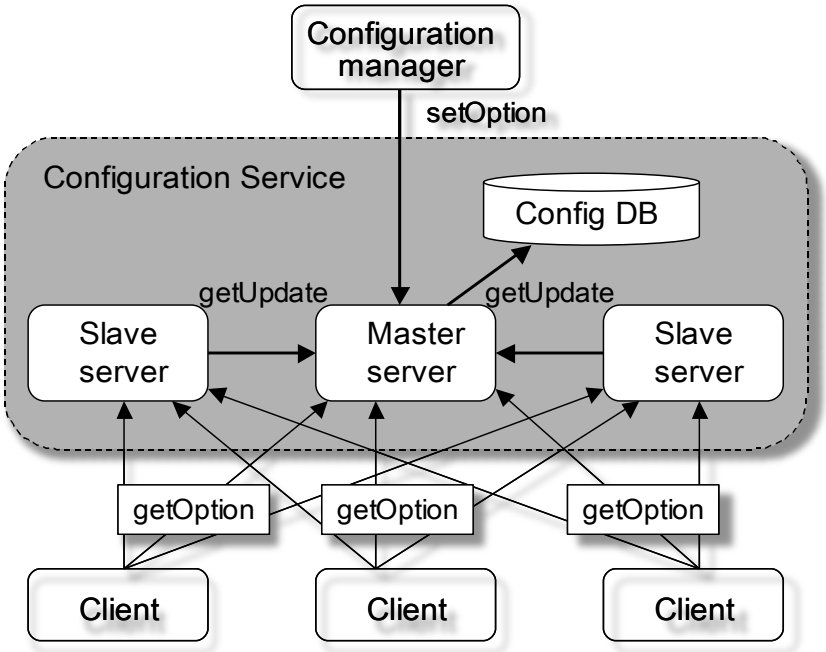
Configuration Service

- This is the DIRAC information index:
  - All the static configuration information
  - Services description for configuration and discovery
  - Resources description
    - Computing
    - Storage
    - Third party services (data transfer, catalogs, message queues, etc)
  - User registry
    - Including VO and group membership
  - Operational parameters
    - For various activities
    - For various communities



Configuration service

- Redundant highly available service
  - Multiple distributed slaves synchronized with the master



- Work in progress:
  - Fine grained access control
    - E.g., VO administrator privileges to manage relevant sections
  - Synchronous slave updates
    - Increase information consistency

Security infrastructure

- Using X509 certificates for all the client/server communications
- Custom implementation of the GSI
  - Python API
- Proxy Manager service
  - Similar to the MyProxy service
    - E.g. providing payload owner proxies to the pilots

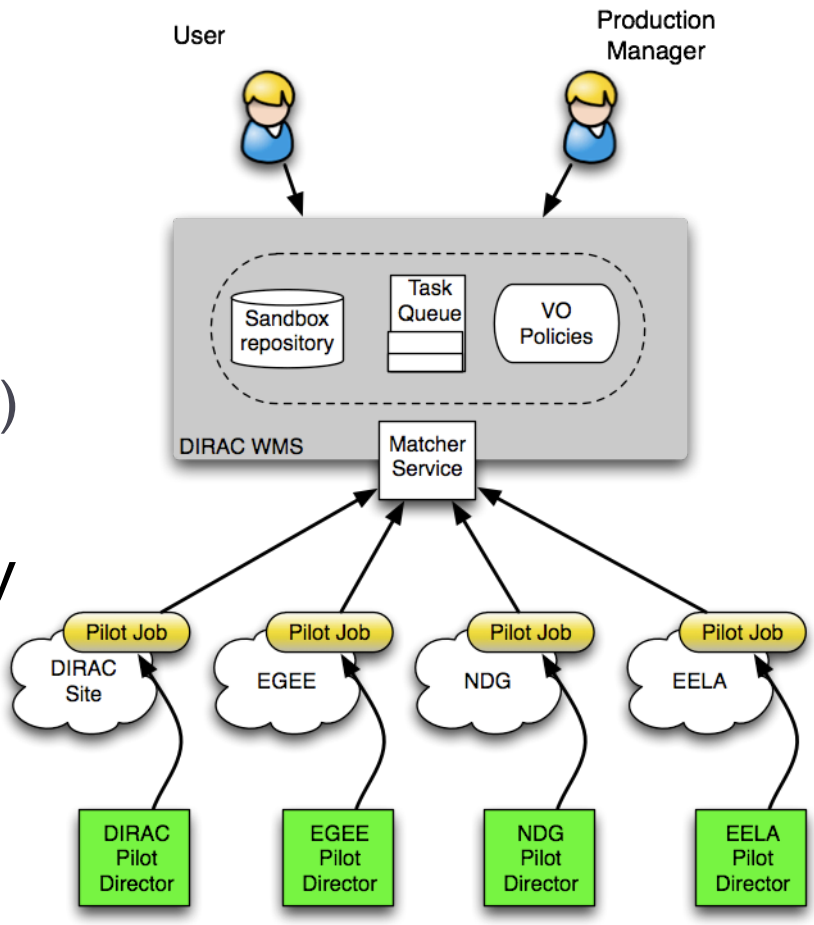
- Work in progress:
  - Getting ready for eventual AAI solutions
    - Enhanced User Profile DB to store security tokens, e.g. login/password
    - Using PUSP proxies (EGI solution)
    - Using DIRAC CA for internal system communications together with not-X509 authentication

Resources Management

- Automatic synchronization of resources description with external information services (BDII or …)
  - Resource access information
  - VO access rights
- Resource Monitoring Service
  - A framework to define and run various probes and update the resource status appropriately
    - E.g. downtimes announced in the GocDB, testing access to storage and computing resources, VOMS servers, etc
- Resource Status Service
  - Serve resource status information to interested clients
- Work in progress
  - Per VO probes and resource status information

Workload Management

- Pilot based workload management
- Targeting various computing resources
  - HTC sites (CREAM,ARC,HTCondor)
  - EGI FedCloud sites
- Possibility to define community specific ad hoc resources
  - E.g. local cluster accessible through (GSI)SSH/VPN tunnel

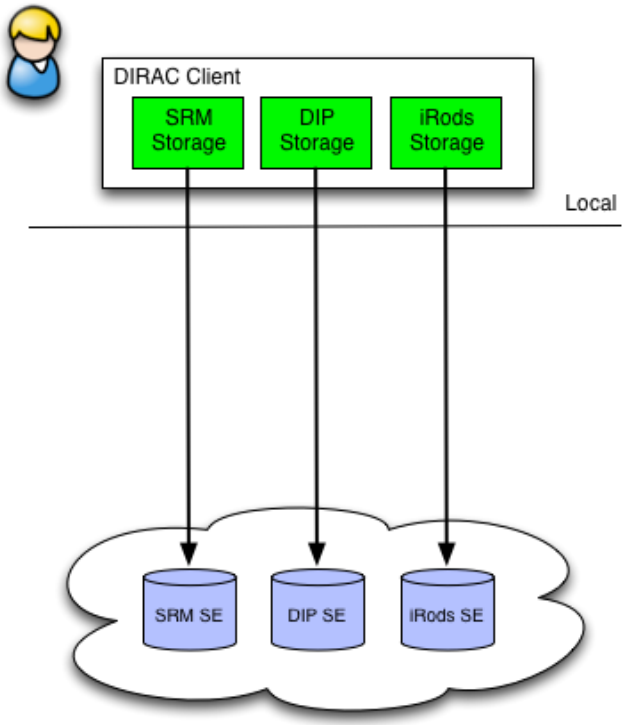


Workload Management

- Work in progress (still a lot)
  - More flexible Pilot framework as a separate DIRAC independent set of modules
    - To run in various ad hoc environments: clouds, containers, BOINC, ...
  - Accessing HPC resources
    - Managing multi/single core jobs
    - Managing HPC special features for efficient job matching
    - Managing limited outbound connectivity of HPC nodes
    - ...
  - VM scheduling for cloud resources
    - Intelligent scheduler for fair sharing of common resources, optimization of the resource usage cost, etc.
  - Transactional bulk job submission
    - The existing bulk submission (up to 50Hz) is not secure enough for the production level system

Data Management

- 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, …



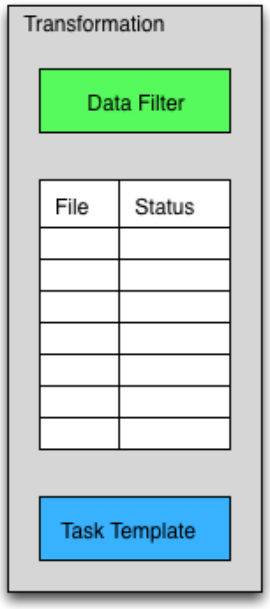
- Central File Catalog
  - DIRAC replica and metadata catalog
    - Dataset management
    - Storage usage reports
  - Possibility to use LFC
    - Tools for migration from LFC to DFC
  - Per VO File Catalog service
    - Possibility to plugin VO-specific modules, e.g. ACLs, metadata or dataset engine

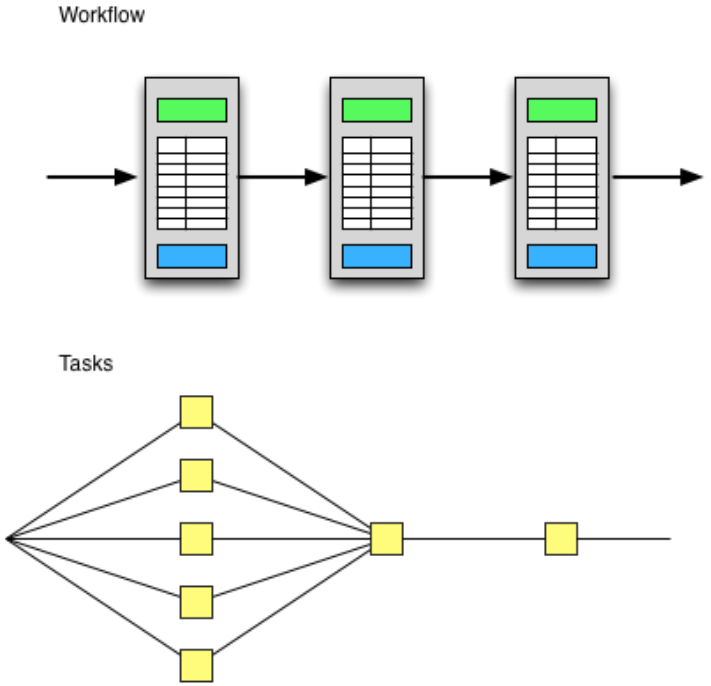
Data Management

- Using bulk data operations
  - Replication, removal, etc
  - Using DIRAC agents for asynchronous operations with retries and validation
  - Using external data transfer services, e.g. FTS3
- Work in progress
  - File Catalog Web interface
    - E.g. coupling data selection with the job submission application
  - Enhanced dataset management

High level services

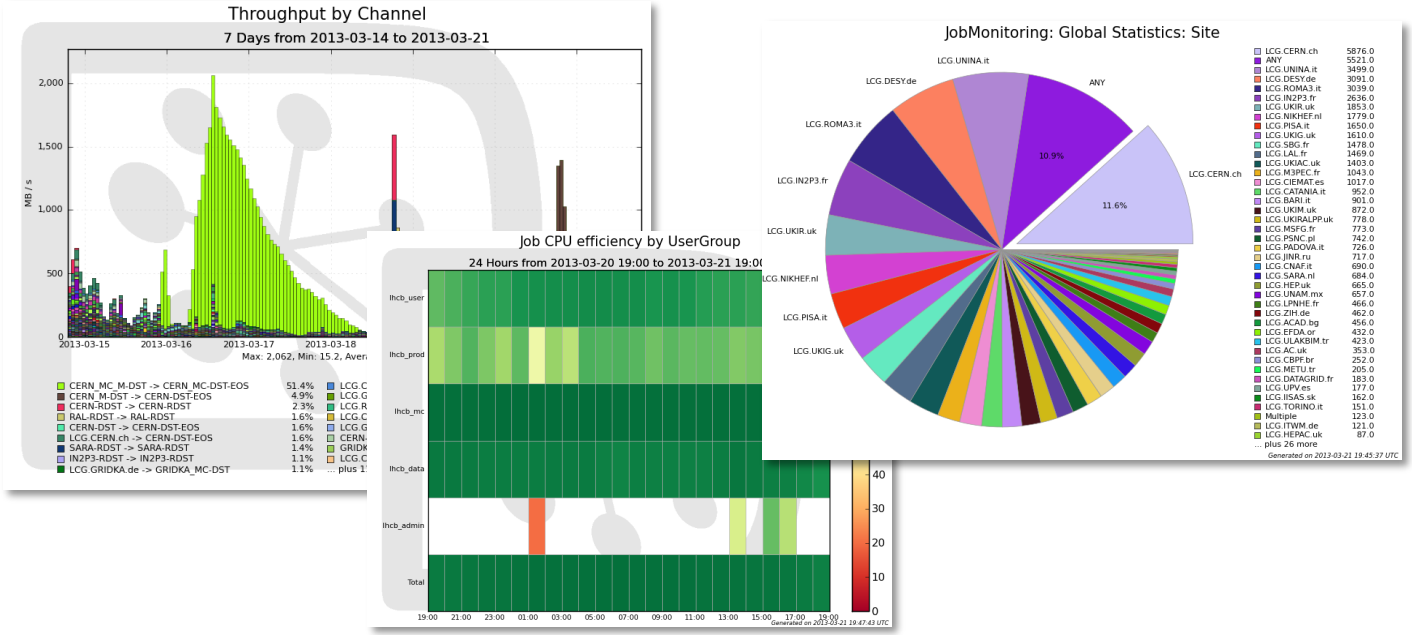
- Data driven workflows as chains of data transformations
  - Transformation: input data filter + recipe to create tasks
  - Tasks are created as soon as data with required properties is registered into the system
  - Tasks: jobs, data operations, etc
- Automatizing community production pipelines
  - Plugins for custom operations, transformation validation
- Work in progress:
  - Production System as a set of tools to help defining complex workflows by chaining multiple transformations





Accounting

- Comprehensive accounting of all the operations



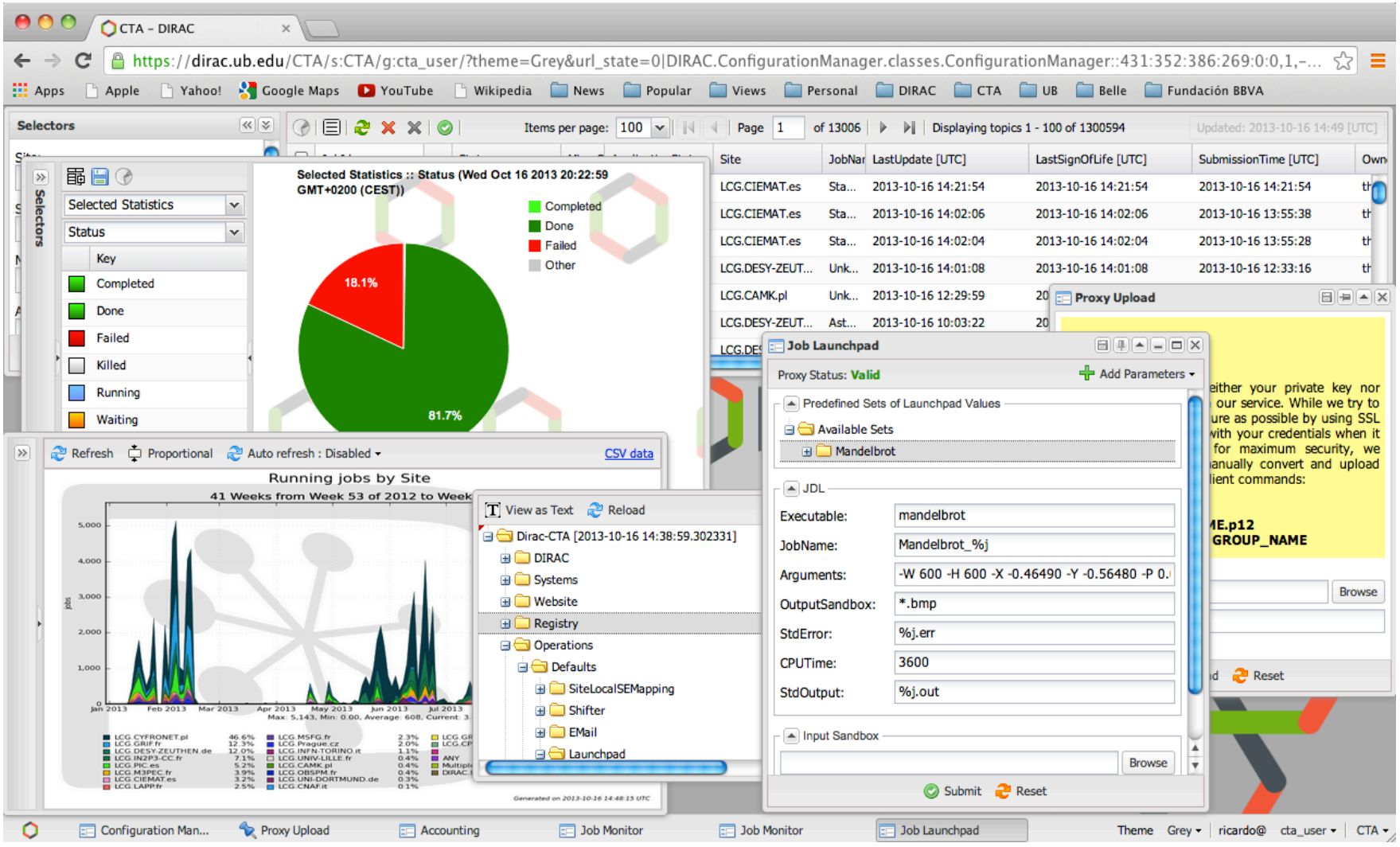
- Using MySQL backend and custom plotting
- Work in progress:
  - Using ElasticSearch/Kibana set of tools
  - Adding more Accounting data types

Interfaces

- Command line for all the operations
  - "gLite-style"
    - *dirac-wms-job-submit job.jdl*
  - COMDIRAC style
    - *dsub echo Hello World !*
- Web Portal for the most common user and administrator operations
- Python API
- REST service interface for third party developments using DIRAC services

- Work in progress
  - More functionality to be exposed through COMDIRAC and REST interfaces
  - Focus on usability of the Web Portal as the main user interface
  - Support for community custom applications built in the DIRAC Web Portal framework

Web Portal examples



Conclusions

- DIRAC provides a framework for building distributed computing systems aggregating multiple types of computing and storage resources
- The list of services available for users of multi-community DIRAC installations provided by grid and cloud infrastructure projects includes basic framework services, resources, workload and data management.
- High level services as well as customized services can be added on demand by interested user communities
- Several developments are still necessary to follow evolution of the available computing and storage resources, AAI frameworks, etc

http://diracgrid.org

DIRAC
THE INTERWARE

Backup slides

VM management