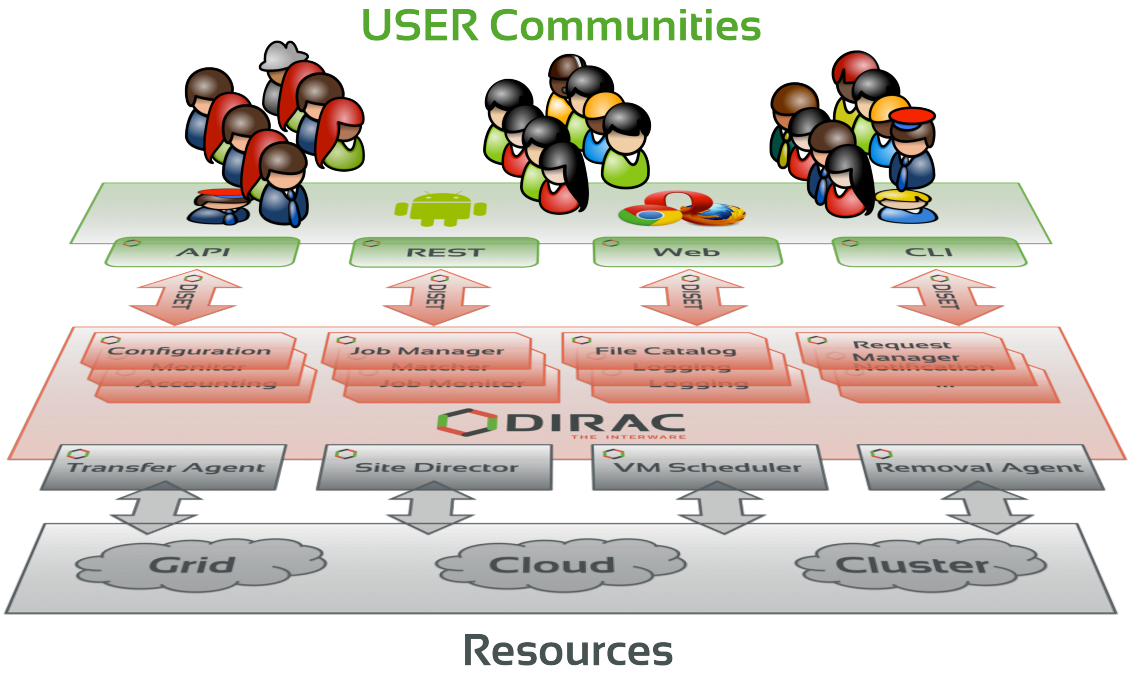
DIRAC Distributed Computing Services

A. Tsaregorodtsev,
CPPM-IN2P3-CNRS

CHEP 2013, Amsterdam, 17 October 2013

Interware

- DIRAC has all the necessary components to build ad-hoc grid infrastructures interconnecting computing resources of different types. This allows to speak about the DIRAC ***interware***.



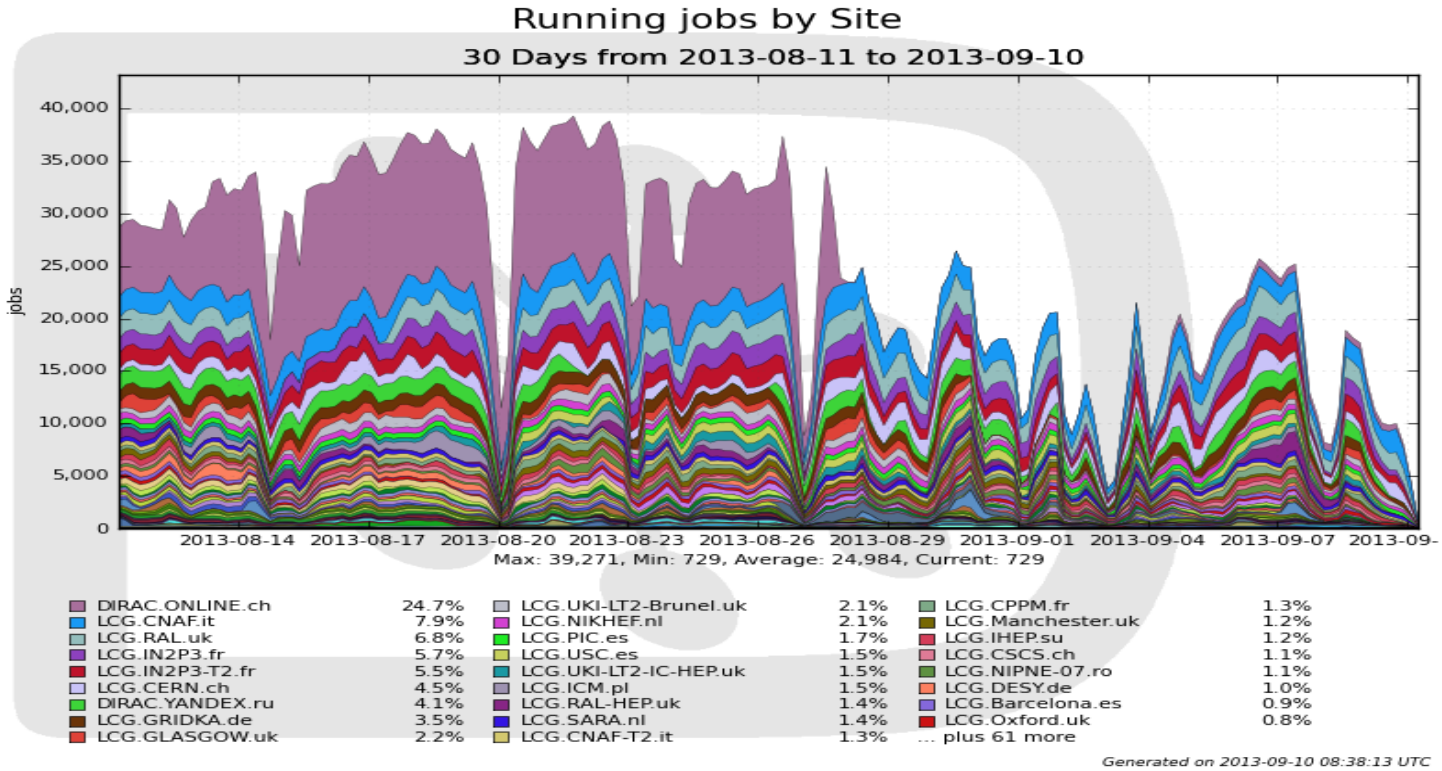
Towards general purpose middleware

- The experience collected with a production grid system of a large HEP experiment is 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.
  - Separate LHCb specific functionality into a set of extensions to the generic core libraries
  - Introduce new services to make it a complete solution
  - Support for multiple small groups by a single DIRAC installation
  - General refurbishing of the code, code management, deployment, documentation, etc

DIRAC
THE INTERWARE

DIRAC Community Installations

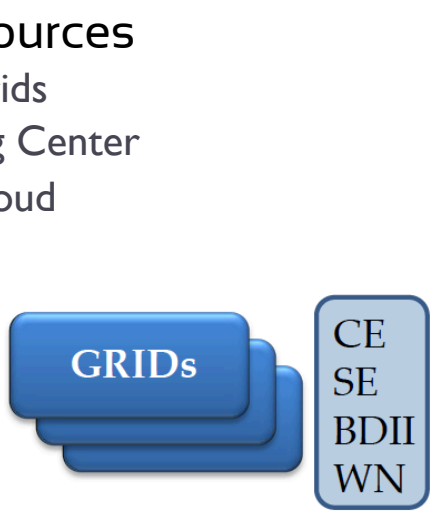
LHCb Collaboration

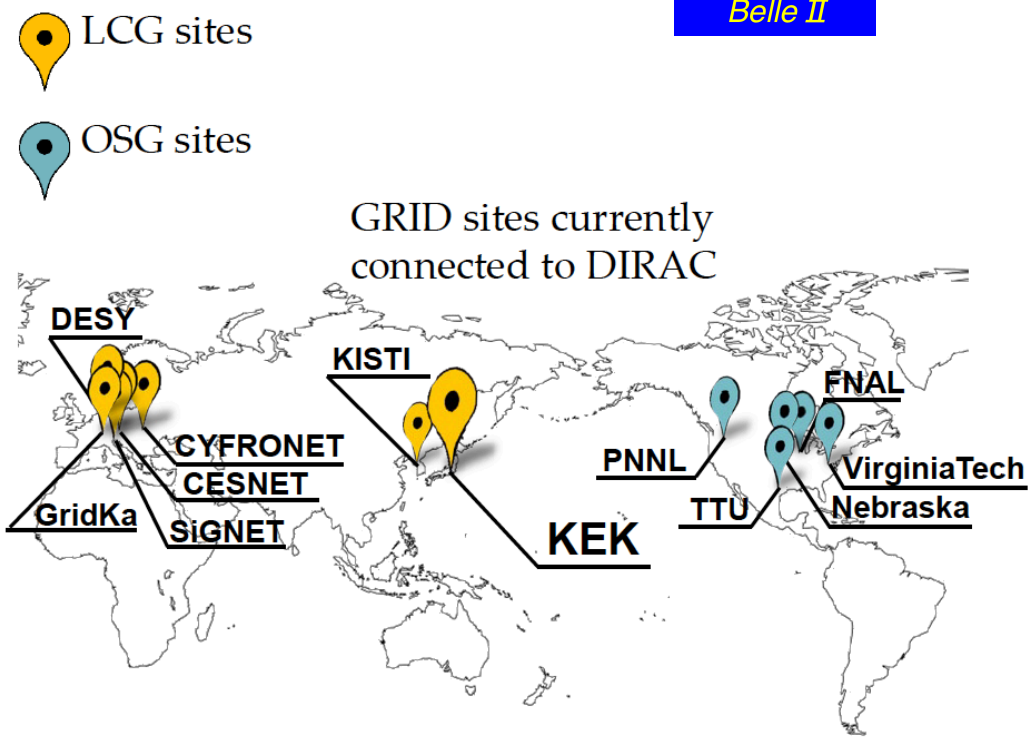


- Up to 50K concurrent jobs in ~120 distinct sites
  - Limited by the resources available to LHCb
- 10 mid-range servers hosting DIRAC central services
- Further optimizations to increase the capacity are possible
  - Hardware, database optimizations, service load balancing, etc

Belle II

- Belle II, KEK, Japan
  - DIRAC is chosen as the basis of Computing Model for phase II of the experiment
  - 2GB/s DAQ rate
  - Combination of the non-grid, grid sites and (commercial) clouds is a requirement
- Belle II grid resources
  - WLCG, OSG grids
  - KEK Computing Center
  - Amazon EC2 cloud





Hideki Miyake, KEK

DIRAC dedicated installations

- ILC/CLIC detector Collaboration
  - Base production system on DIRAC
  - MC simulations
  - DIRAC File Catalog was developed to meet the ILC/CLIC requirements

- BES III, IHEP, China
  - DIRAC is chosen for the phase III
  - Using DIRAC DMS: File Catalog, Transfer services

- CTA
  - CTA started as FG-DIRAC customer for DIRAC evaluation
  - Now is using a dedicated installation at PIC, Barcelona
  - Using complex workflows

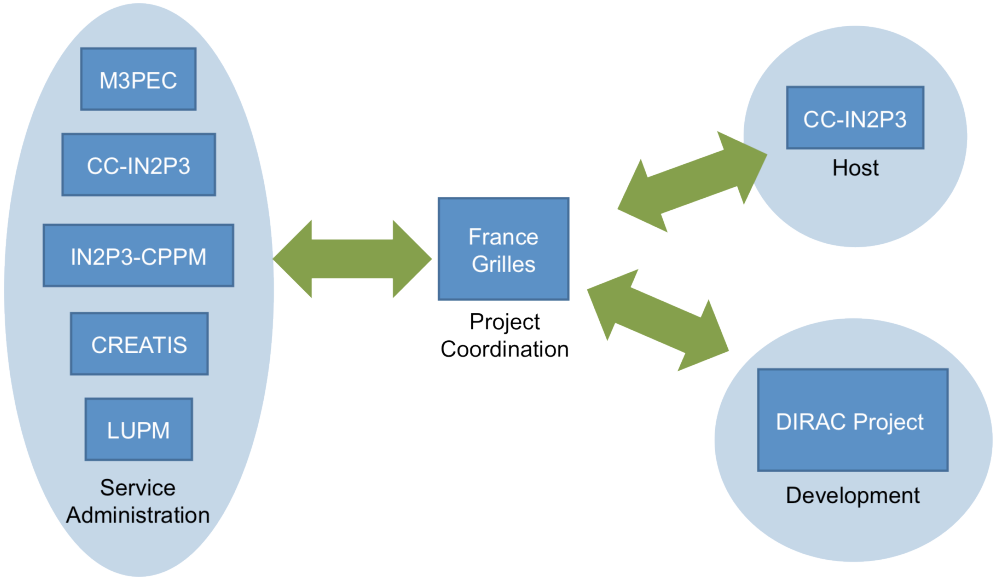
- DIRAC evaluations by other experiments
  - LSST, Auger, TREND, …

DIRAC
THE INTERWARE

DIRAC as a Service

France-Grid DIRAC service

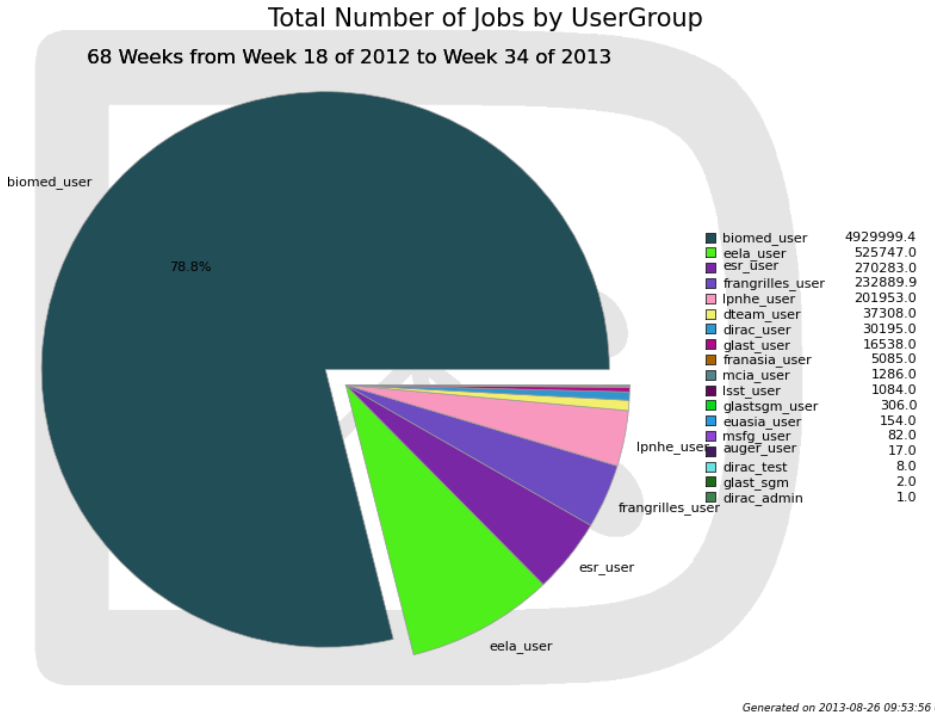
- Several regional and university campus installations in France
  - Complex maintenance
- Joint effort to provide France-Grid DIRAC service
  - Hosted by the CC/IN2P3, Lyon, T1 center
    - 6 virtual servers, MySQL server
  - Distributed team of service administrators
    - 5 participating universities



http://dirac.france-grilles.fr

FG-DIRAC users

- France-Grilles users
  - 15 VOs, 88users registered
    - astro, auger, biomed, esr, euasia, gilda, glast.org, prod.vo.eu-eela.eu, vo.cta.in2p3.fr, vo.formation.idgrilles.fr, vo.france-asia.org, vo.france-grilles.fr, vo.msfg.fr, vo.mcia.org
    - 1 robot user VIP/GateLab Biomed
    - More VO's and users can be added as necessary
  - In production since May 2012
    - First ~7 millions jobs went through the system
      - Mostly biomed applications



DIRAC
THE INTERWARE

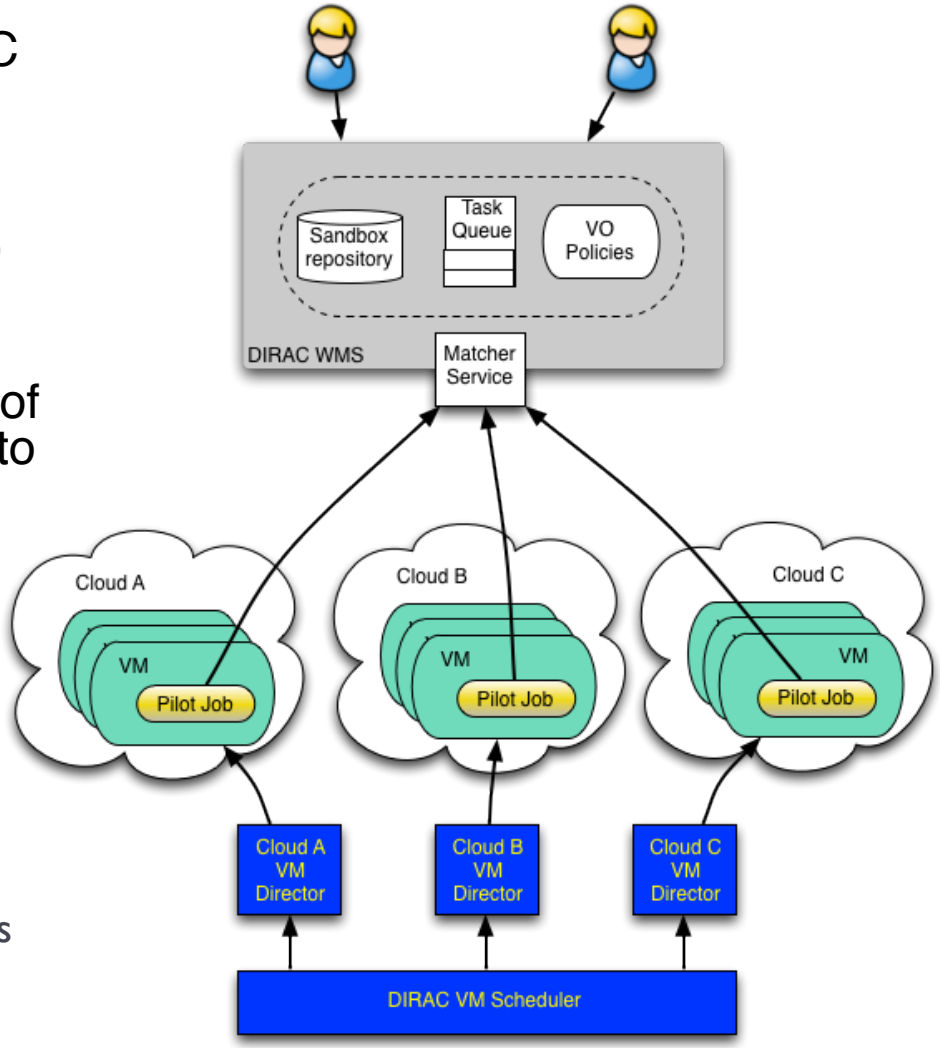
Resources Available via DIRAC service

Computing Grids

- DIRAC was initially developed with the focus on accessing conventional Grid computing resources
  - WLCG grid resources for the LHCb Collaboration
- It fully supports gLite middleware based grids
  - EGI, GISELA, etc
    - Using gLite WMS or accessing CE's directly
  - OSG
- Support for ARC middleware based services
  - NorduGrid, RAL
- Other types of grids can be supported
  - As long we have customers needing that

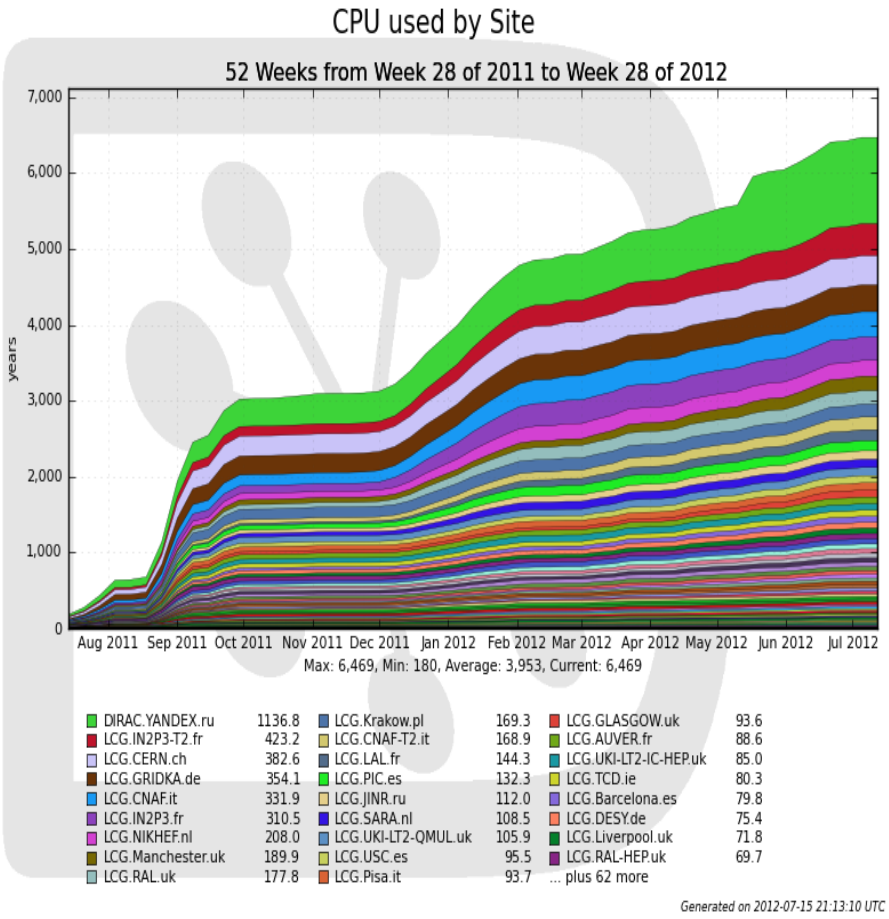
Clouds

- VM scheduler developed for Belle MC production system
  - Dynamic VM spawning taking into account the Task Queue state
  - Discarding VMs automatically when no more needed
- The DIRAC VM scheduler by means of dedicated VM Directors is interfaced to
  - OCCI compliant clouds:
    - OpenStack, OpenNebula
  - CloudStack
  - Amazon EC2
- Intensive development now
  - different access methods, VM contextualization, VM scheduling policies
  - part of the EGI Cloud Task Force activities



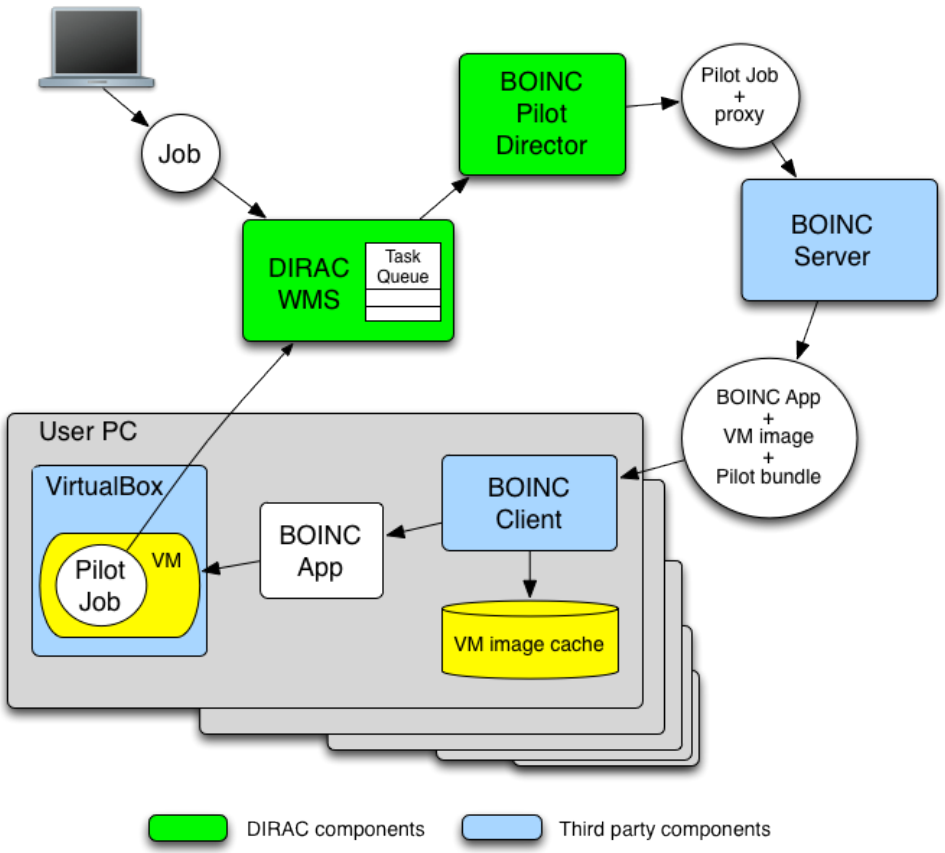
Standalone computing clusters

- Access through SSH tunnel
  - No grid middleware installation needed on site
- Examples:
  - DIRAC.Yandex.ru
    - 1800 cores
    - Torque batch system, no grid middleware, access by SSH
    - Second largest LHCb MC production site
  - LRZ Computing Center, Munich
    - SLURM batch system, GRAM5 CE service
    - Gateway access by GSISSH
    - Considerable resources for biomed community (work in progress)
  - Mesocentre Aix-Marseille University
    - OAR batch system, no grid middleware, access by SSH
    - Open to multiple communities (work in progress)



BOINC Desktop Grids

- On the client PC the third party components are installed:
  - VirtualBox hypervisor
  - Standard BOINC client
- A special BOINC application
  - Starts a requested VM within the VirtualBox
  - Passes the Pilot Job to the VM and starts it
- Once the Pilot Job starts in the VM, the user PC becomes a normal DIRAC Worker Node
- Possibility to use the MarketPlace repository of VM images
- Interfacing DIRAC to EDGI resources
  - Using EDGI provided special CREAM CE service



Data Management components

- Storage Elements
  - gLite/EGI Storage Elements
    - Standard SRM interface
    - Gridftp data transfer protocol
      - Need Globus libraries, limited number of platforms
  - DIRAC Storage Elements
    - DIPS (Dirac Secure Protocol) data transfers
    - Possibility to synchronize ACLs with the DIRAC File Catalog
  - More Storage Elements plug-ins can be included
    - (F,SF,HT,BBF)TP servers
    - iRods

DIRAC
THE INTERWARE

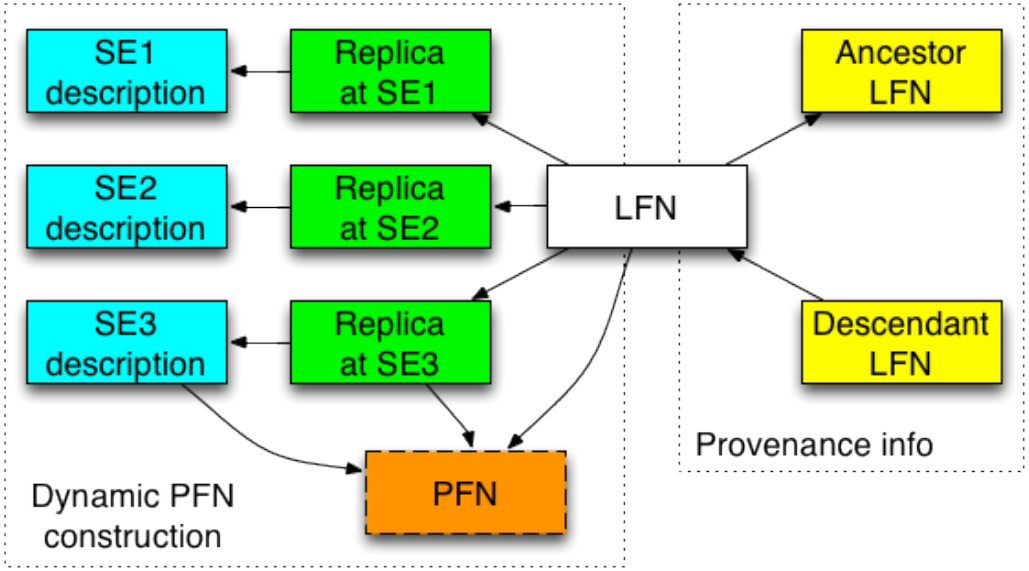
Services

Services in CC/Lyon

- Basic DIRAC services
  - Resources description and monitoring
  - WMS – pilot based management of user jobs
    - Job submission, monitoring, retrieval
    - Accounting of the resources consumed
  - DMS – managing user data basic tasks
    - Access to standard Grid Storage Elements
      - SRM, DIRAC SEs
    - Replicating data between SEs
    - Providing Simple Storage Element in Lyon
    - DIRAC File Replica Catalog
    - DIRAC File Metadata Catalog
    - Several LFC services configured in DIRAC DMS
    - Accounting of data transfer operations

Replica Catalog

- Standard Replica Catalog functionality
  - Optimized for bulk queries
- On the fly PFN construction
  - Small database footprint
  - Pattern used in LHCb
- Ancestor-descendent relations
  - Basic provenance information
  - Possibility to select ancestors in a given generations



DIRAC
THE INTERWARE

Storage Usage

- Efficient Storage Usage reports
  - Necessary for quota policy management
- Using special prefilled tables
  - Updated at each new file or replica insertion
    - More efficient with bulk insertion
  - Instant reports for any directory
  - Possibility of instant "*du*" command

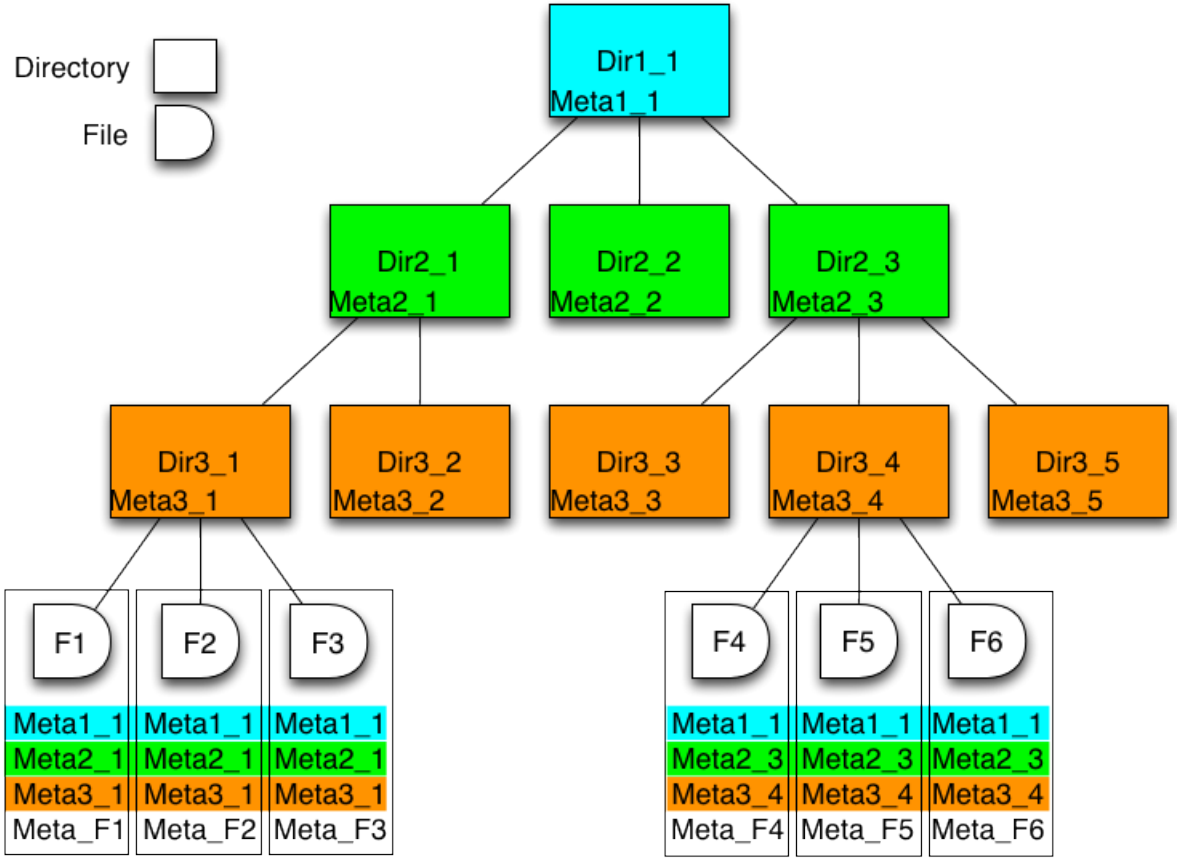
Storage Usage

```
FC:/> size -l /lhcb/user/a/atsareg/1
directory: /lhcb/user/a/atsareg/1
Logical Size: 134,756,846 Files: 498 Directories: 500

    StorageElement  Size          Replicas
-------------------------------------------
  1 IN2P3-USER      20,254,050    75
  2 CNAF-USER       18,363,672    68
  3 RAL-USER        16,473,294    61
  4 CERN-USER       19,443,888    72
  5 GRIDKA-USER     21,064,212    78
  6 SARA-USER       20,254,050    75
  7 PIC-USER        18,903,780    70
-------------------------------------------
    Total           134,756,946   499
Query time 0.98 sec
```

- Report of storage usage for any directory
  - Whole community data
  - Per user data
  - “Logical” storage
    - LFNs, sum of the LFN sizes
  - “Physical” storage
    - Physical replicas, total volume per Storage Element

DFC Metadata



File Catalog Metadata

- Similar functionality with the AMGA metadata service
  - But coupled with the replica catalog to boost efficiency
- Metadata can be associated with each directory as key:value pairs to describe its contents
  - Int, Float, String, DateTime value types
- Some metadata variables can be declared indices
  - Those can be used for data selections
- Subdirectories are inheriting the metadata of their parents
- Data selection with metadata queries. Example:
  - `find . Year=2010 Ver=v1r0,v1r1 SE=CERN-disk LastAccess>10-10-2013`
- File metadata can also be defined

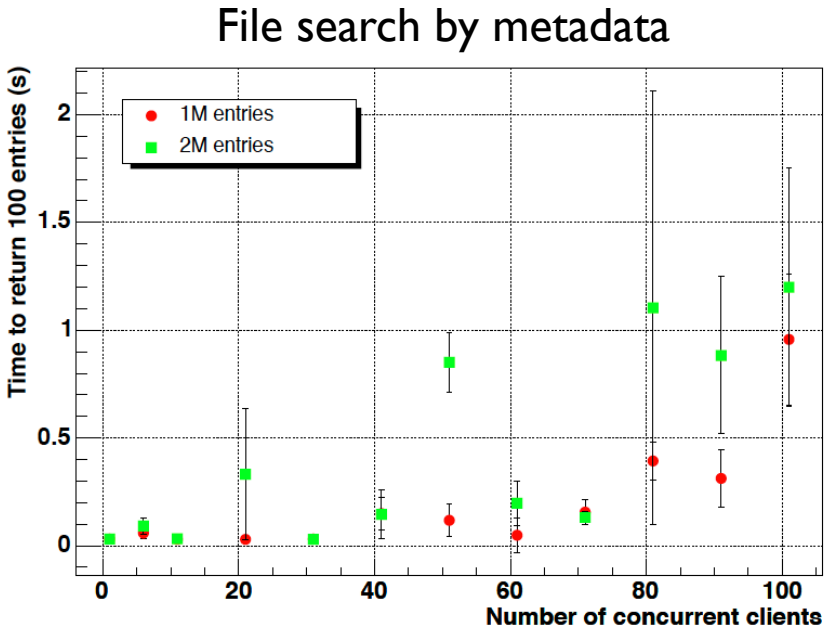
DIRAC
THE INTERWARE

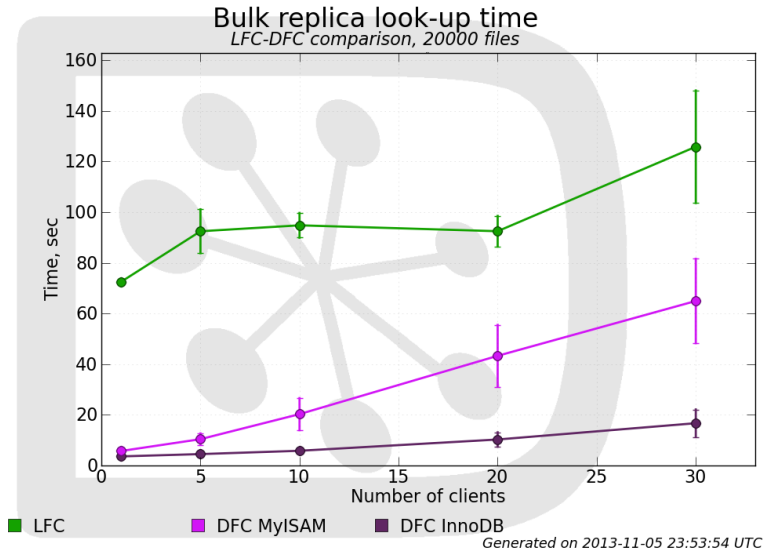
Architecture

- Modular architecture
  - Directory hierarchy plug-ins
  - Security management plug-ins
    - POSIX like ACLs, per directory ACLs, global read access, etc
  - Metadata engines
  - Dataset engines

DIRAC File Catalog evaluation

- Tests with Auger data
  - ~30M files
  - Identical LFC and DFC server hardware





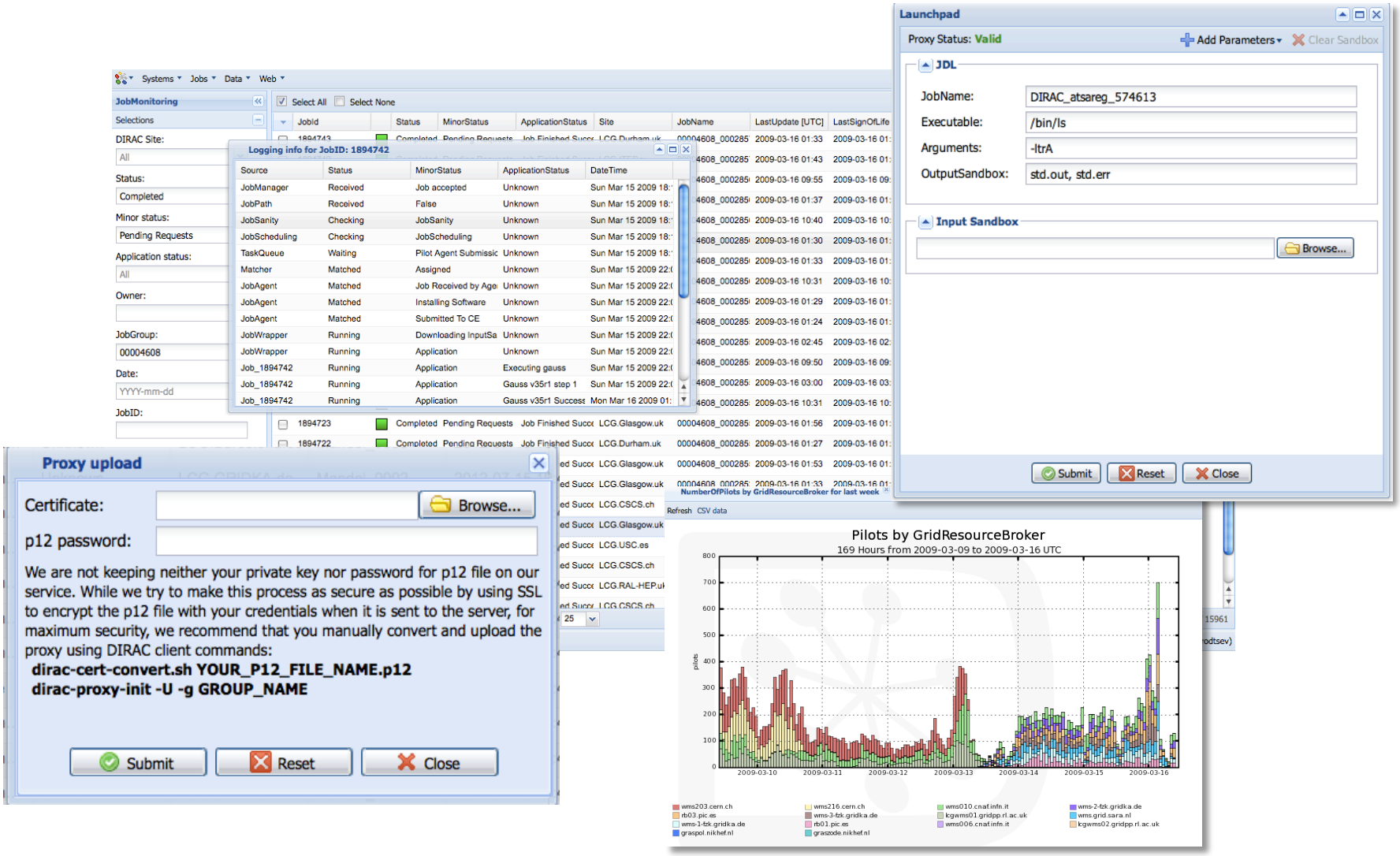
- BES Collaboration made a thorough comparison of DFC vs AMGA
  - Similar performance
  - More suitable functionality

DIRAC
THE INTERWARE

Web Portal

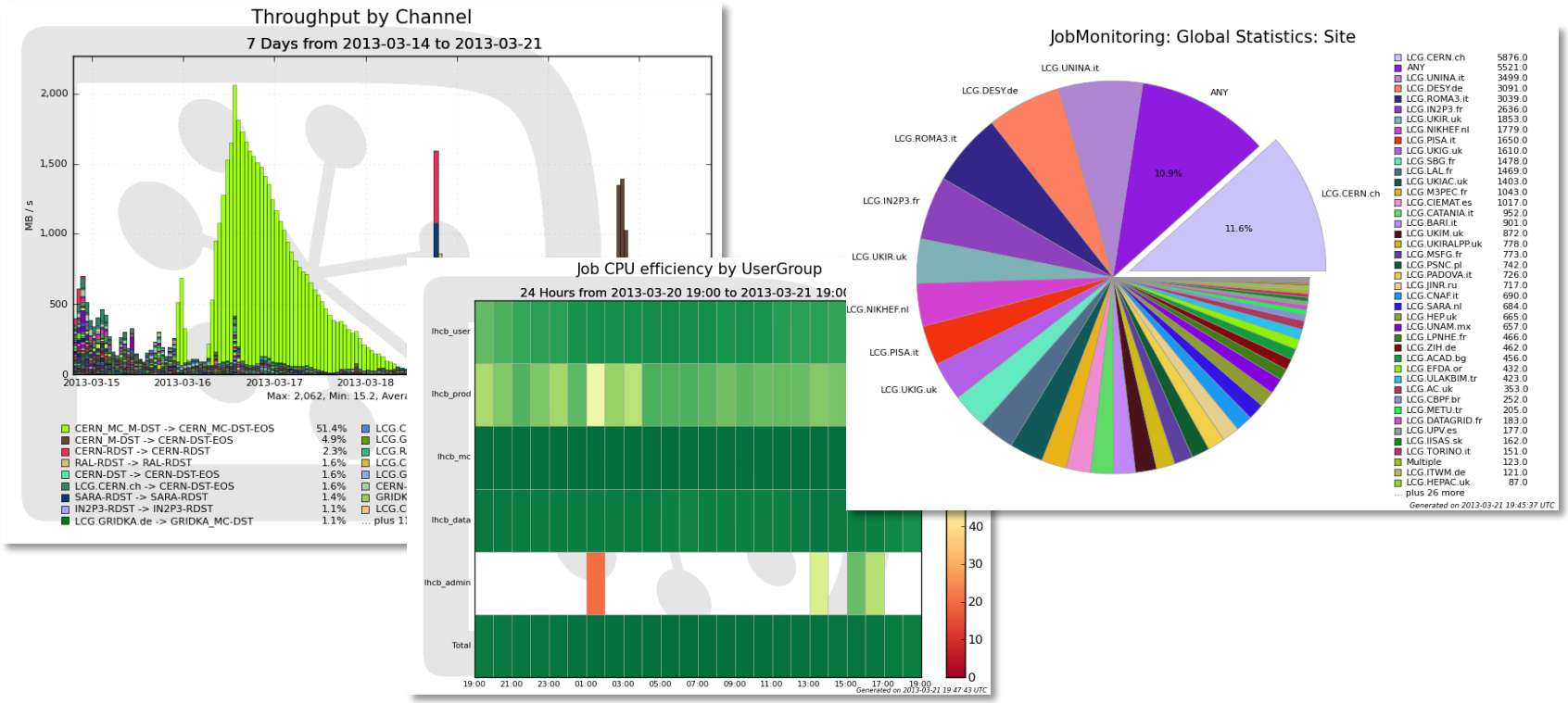
- Web Portal
  - Support of most of the user tasks (jobs, data)
  - Secure with X509 certificates
- Specific application portals can be built in the DIRAC Web Portal framework
  - Community Application Servers
- DIRAC RESTful interface
  - Language neutral
  - Suitable to use with portals written in Java, PHP, etc
- Other interfaces include
  - Extensive Python API
    - E.g. used by GANGA user front-end
  - A rich set of command line tools (>200 commands)

Web Portal: example interfaces



Accounting

- Comprehensive accounting of all the operations



- Publication ready quality of the plots
  - Plotting service can be used by users for there own data

Advanced services

- More advanced services can be made available in CC Lyon
  - Following the user demands
  - Transformation Service (automated job submission)
  - Replication Service (automated data replication)
  - Data integrity inspection
  - User storage consumption accounting
  - Support for MPI jobs
  - …
- Hosting Community DIRAC services
  - Specific services developed for particular communities can be hosted in the same infrastructure

Conclusions

- The computational grids are no more something exotic, they are used in a daily work for various applications
- Rich experience with using computational grids in the LHC experiments, as well as the developed tools, can now be shared with users in other experiments and in other scientific domains
- DIRAC is providing a framework for building distributed computing systems and a rich set of ready to use services. This is used now in a number of DIRAC service projects on a regional and national levels
- Services based on DIRAC technologies can help users to get started in the world of distributed computations and reveal its full potential

http://diracgrid.org