



- the new generation of the
LHCb grid software

*A. Tsaregorodtsev,
CPPM-IN2P3-CNRS, Marseille*

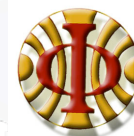
For the LHCb DIRAC team



26 March, CHEP 2009, Prague



DIRAC



DIRAC development team

A. Tsaregorodtsev, N. Brook, A. Casajus Ramo, Ph. Charpentier, J. Closier, G. Cowan, R. Graciani Diaz, E. Lanciotti, Z. Mathe, R. Nandakumar, S. Paterson, V. Romanovsky, R. Santinelli, A.C. Smith, M. Seco Miguez, A. Zhelezov



Science & Technology Facilities Council
Rutherford Appleton Laboratory



University of
BRISTOL

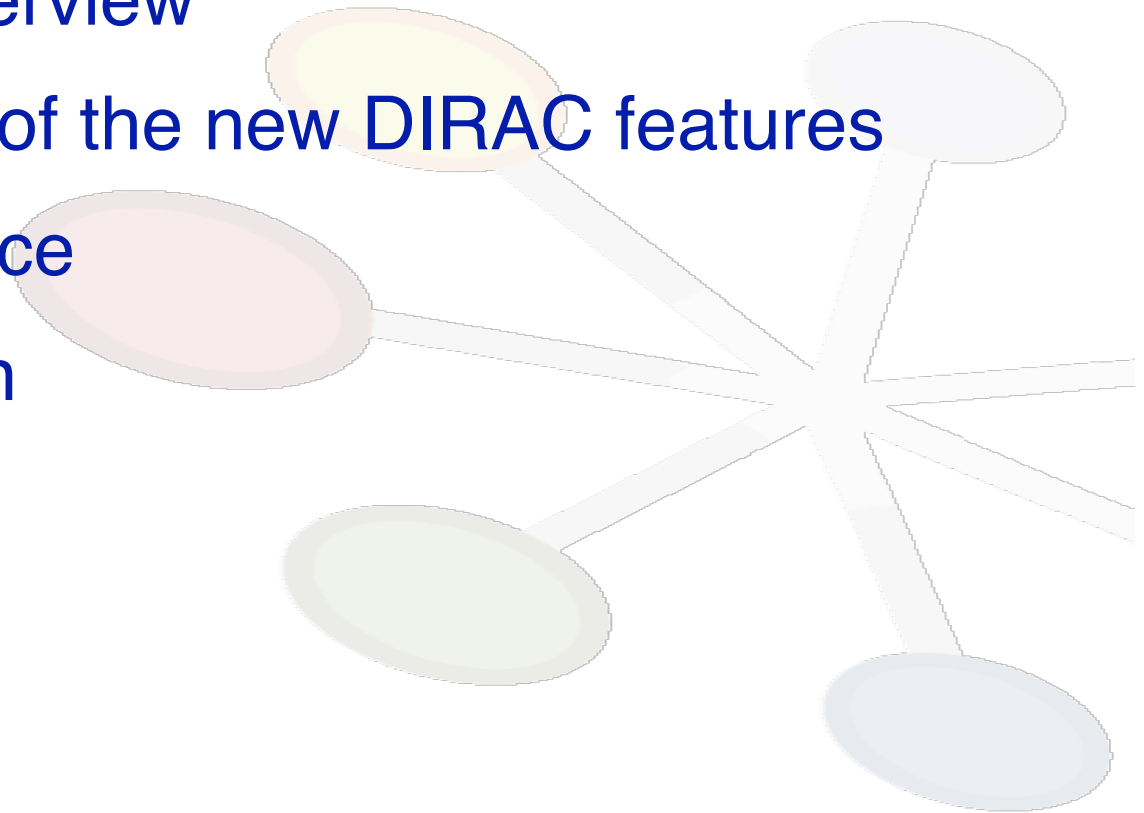




Outline

DIRAC

- ◆ DIRAC overview
- ◆ Highlights of the new DIRAC features
- ◆ Performance
- ◆ Conclusion

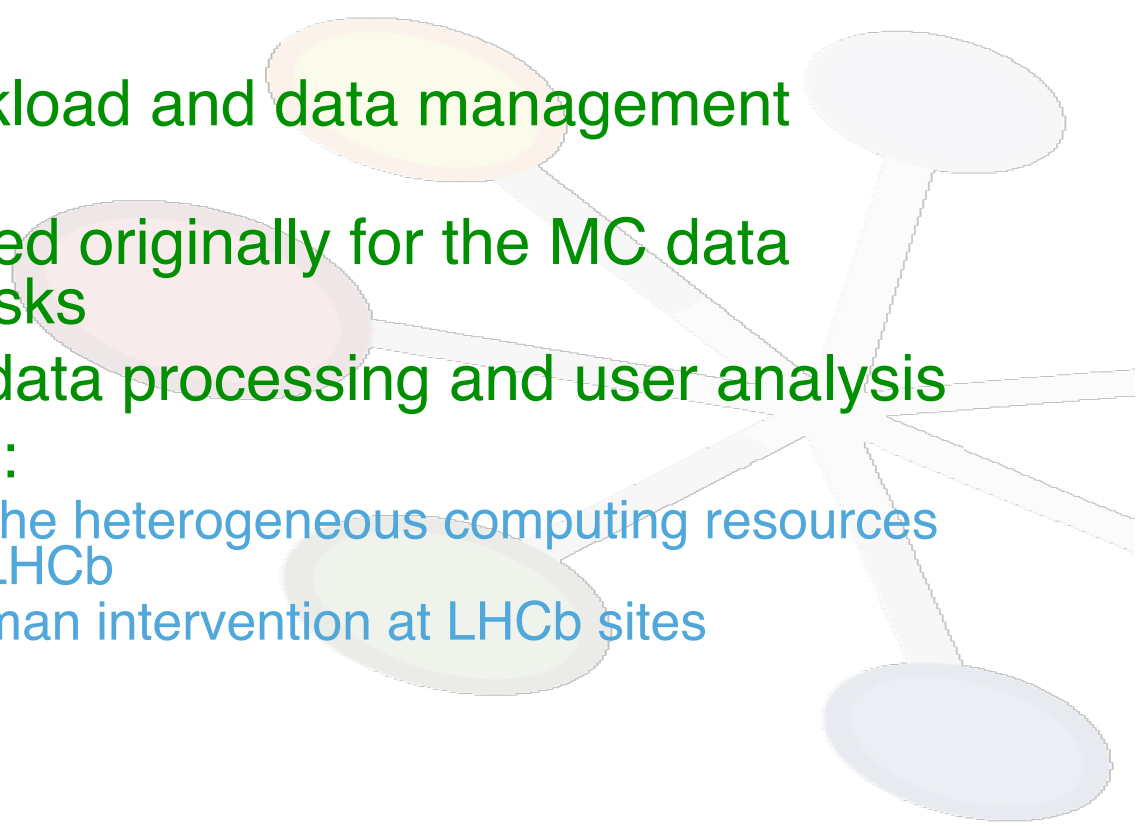




DIRAC overview

DIRAC

- ◆ DIRAC is a distributed data production and analysis system used by the LHCb experiment
 - ✦ Includes workload and data management components
 - ✦ Was developed originally for the MC data production tasks
 - ✦ Extended to data processing and user analysis
 - ✦ The goal was:
 - integrate all the heterogeneous computing resources available to LHCb
 - Minimize human intervention at LHCb sites





DIRAC project review

DIRAC

- ◆ In 2005-2006 the design and implementation of various DIRAC systems was reviewed:
 - ✦ Independent experts
 - Non DIRAC
 - Non LHCb
 - ✦ Multiple recommendations were collected in the final 35 pages review report

DIRAC review report



LHCb Computing Note

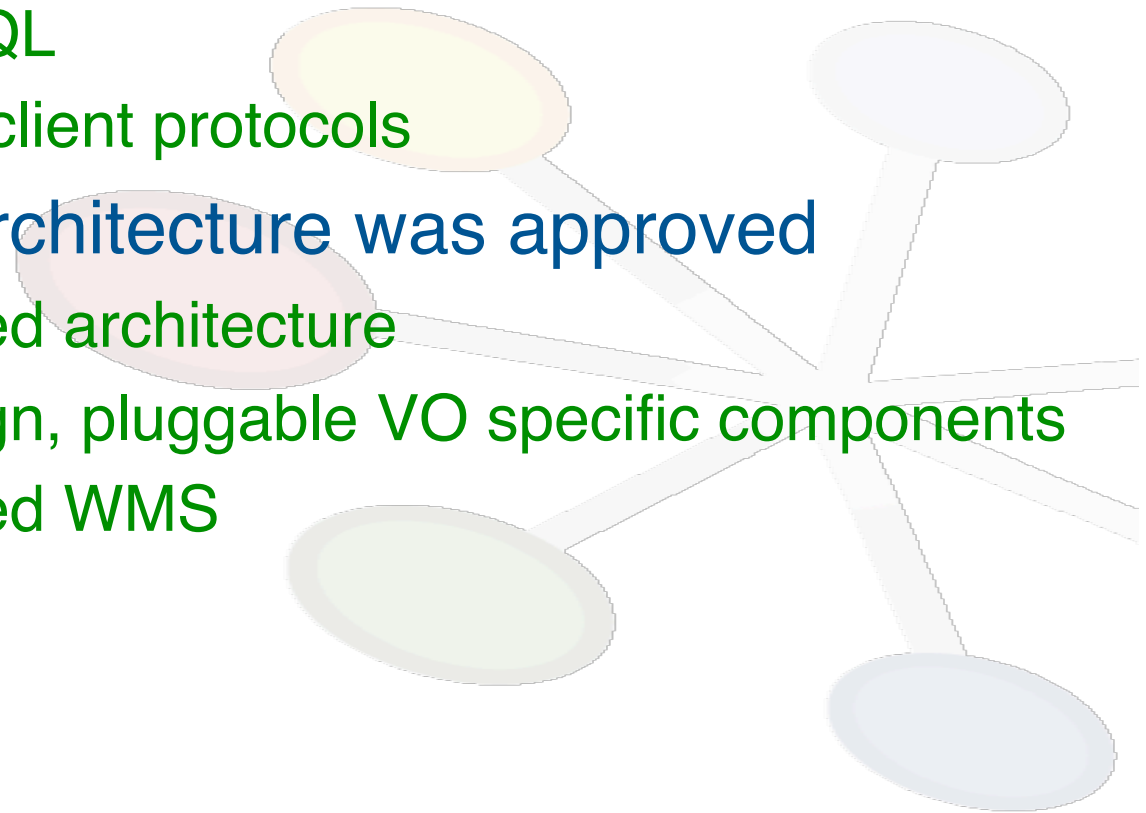
Issue:	Final
Revision:	1.5
Reference:	LHCb-2006-04 COMP
Created:	26 January 2006
Last modified:	16 March 2009
Prepared By:	J.-P. Baud, Ph. Charpentier (Editor), J. Closier, R. Graciani, A. Maier, P. Mato-Vila



Review outcomes

DIRAC

- ◆ Base technologies are OK
 - ✦ Python, MySQL
 - ✦ Light service/client protocols
- ◆ The general architecture was approved
 - ✦ Services based architecture
 - ✦ Modular design, pluggable VO specific components
 - ✦ Pilot Job based WMS





Review outcomes

DIRAC

◆ Main recommendations

- ✦ DIRAC security model to be complete and following grid standards
- ✦ Possibility to apply LHCb VO policies
- ✦ Build in more redundancy to cope with the grid resources instabilities
- ✦ Comprehensive monitoring of production and user activities
- ✦ Improve project management, separate generic and VO specific parts

◆ The outcome of the review process was the new generation DIRAC3 system

- ✦ Essentially rewritten from scratch to meet the review recommendations
- ✦ Took long as the same team was supporting ongoing operations



DIRAC: complete chain

DIRAC

- ◆ DIRAC is covering all the LHCb needs in the distributed data processing
 - ✦ Data export from the experiment pit to CERN off-line storage
 - ✦ Automatic data distribution to Tier-1 centers
 - ✦ Automatic creation and submission of the data reconstruction jobs
 - ✦ Automatic distribution of the analysis data
 - ✦ Full management of the MC data production
 - ✦ Full support for the user analysis jobs
- ◆ Different subsystems built in the same framework
 - ✦ Reuse of technical solutions in different subsystems
 - ✦ A concerted team of developers sharing experience



DISET Secure framework

DIRAC

- ◆ All the communications between the distributed components in the DIRAC framework (*DISET*) are secure
 - ✦ X509, GSI security standards
 - ✦ Fine grained authorization rules
 - Per individual user FQAN
 - Per service interface method
 - Per job
- ◆ Full featured proxy management system
 - ✦ Proxy storage and extension
 - ✦ Support for multiuser pilot jobs
 - Limited proxies
 - Special tokens to limit the number of served proxies

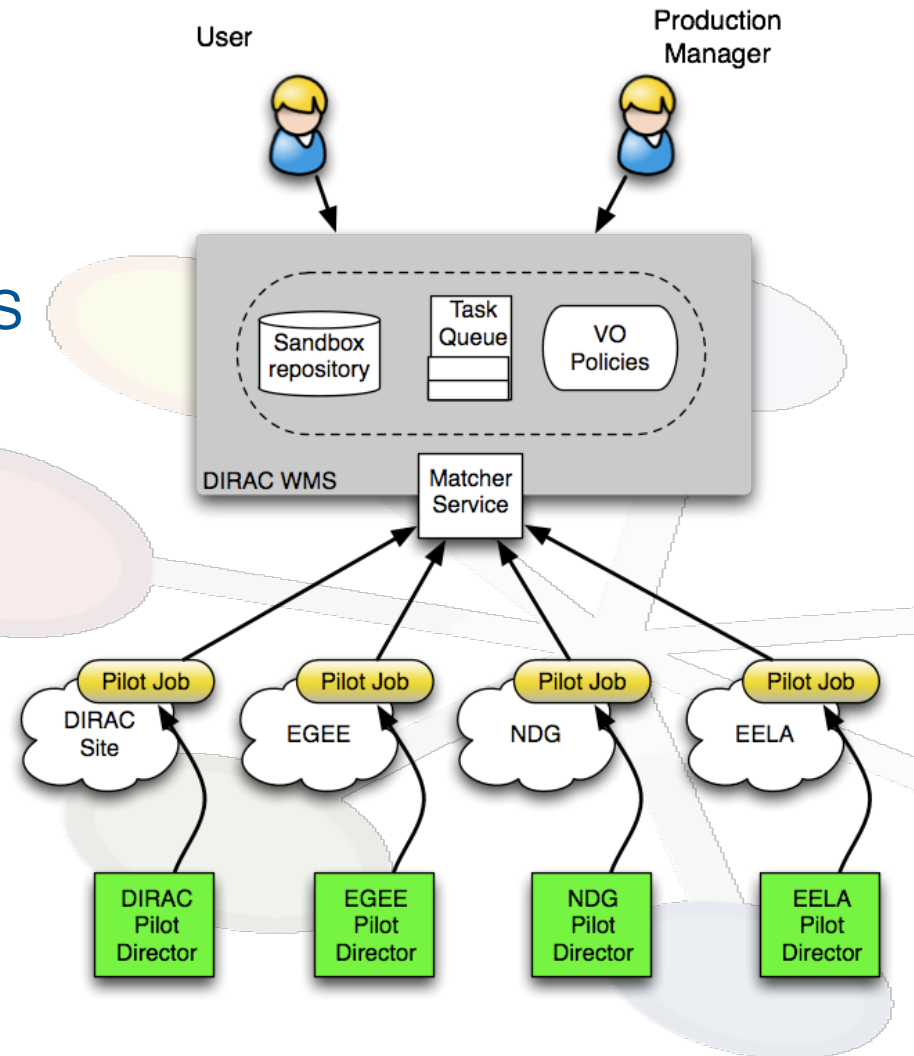
see [90] – R.Graciani ,A.Casajus



WMS: Pilot Jobs

- ♦ Jobs are submitted with credentials of their owner (VOMS proxy)
- ♦ Pilot Jobs are submitted by specific Directors to a Grid WMS
- ♦ The Pilot Job fetches the User Job and proxy
- ♦ The User Job is executed with its owner's proxy used to access SE, catalogs, etc

see [108] – R.Graciani et al

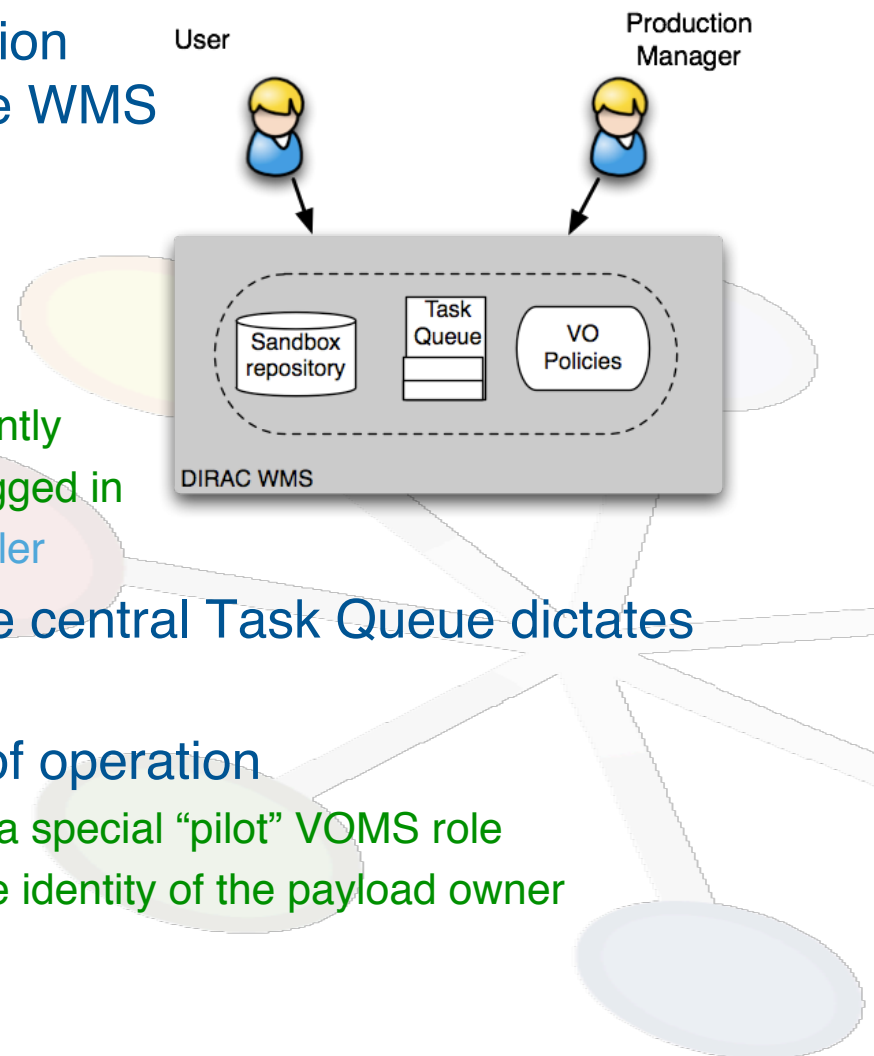




WMS: applying VO policies

DIRAC

- ◆ In DIRAC both User and Production jobs are now treated by the same WMS
- ◆ This allows to apply efficiently policies for the whole VO
 - ✦ Assigning Job Priorities for different groups and activities
 - ✦ Static group priorities are used currently
 - ✦ More powerful scheduler can be plugged in
 - demonstrated with MAUI scheduler
- ◆ The VO policies application in the central Task Queue dictates the use of Multiuser Pilot Agents
- ◆ DIRAC fully supports this mode of operation
 - ✦ Multiuser Pilots Jobs submitted with a special “pilot” VOMS role
 - ✦ Using glxexec on the WNs to track the identity of the payload owner

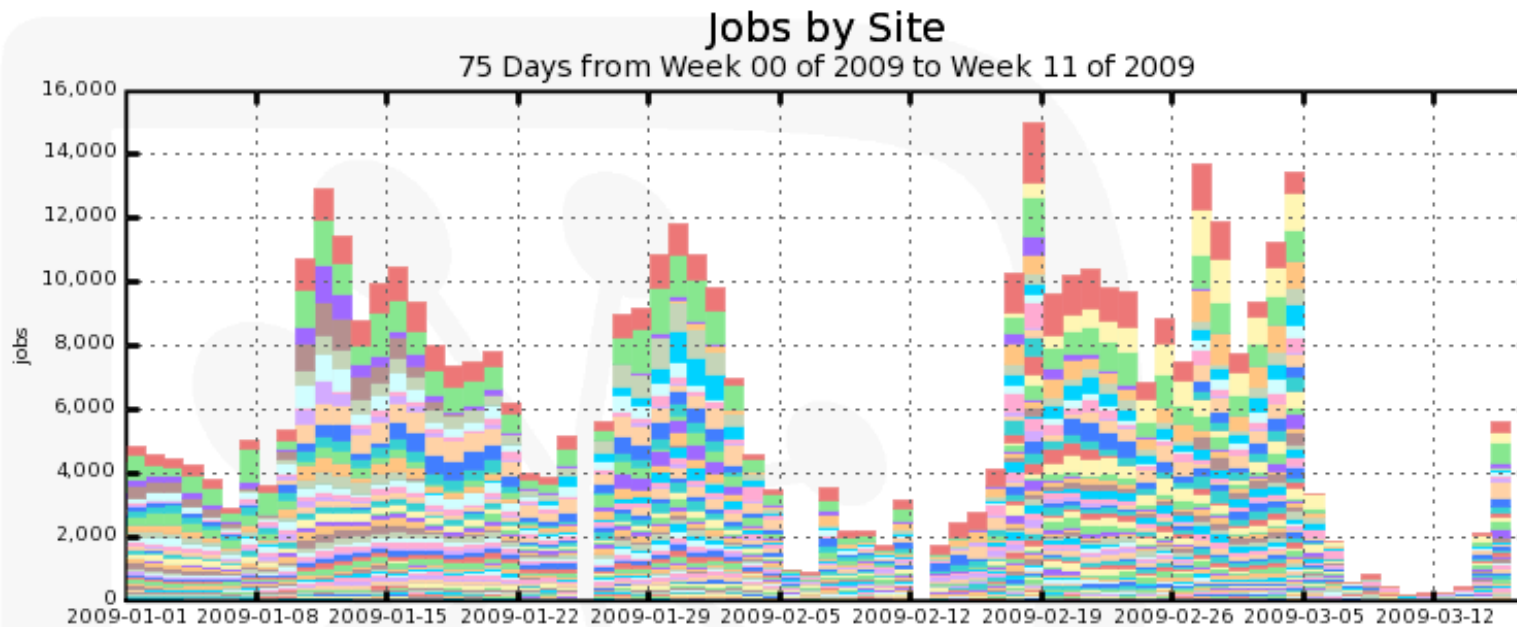


see [439] – S.Paterson et al
[278] – M.Cattaneo et al



Performance

DIRAC



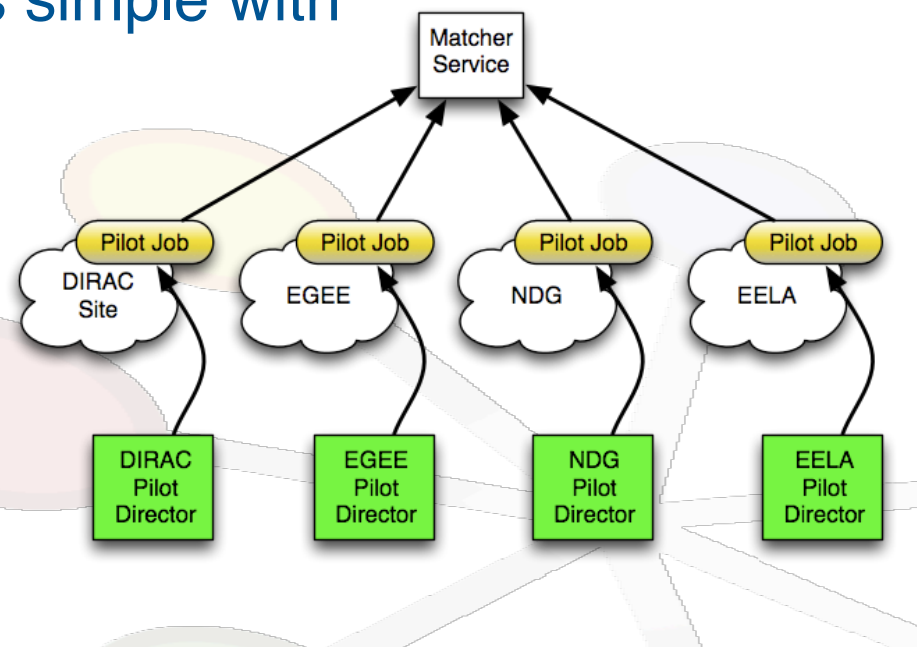
- ◆ DIRAC performance measured in the recent production and FEST'09 runs
 - ✦ Up to 15K concurrent jobs in ~120 distinct sites
 - ✦ One mid-range central server hosting DIRAC services
 - ✦ Further optimizations to increase capacity are possible
 - Hardware, database optimizations, service load balancing, etc



WMS: using heterogeneous resources

- ◆ Including resources in different grids and standalone clusters is simple with Pilot Jobs

- ★ Needs a specialized Pilot Director per resource type
- ★ Demonstrated with NDG and EELA grid sites
- ★ Users just see new sites appearing in the job monitoring

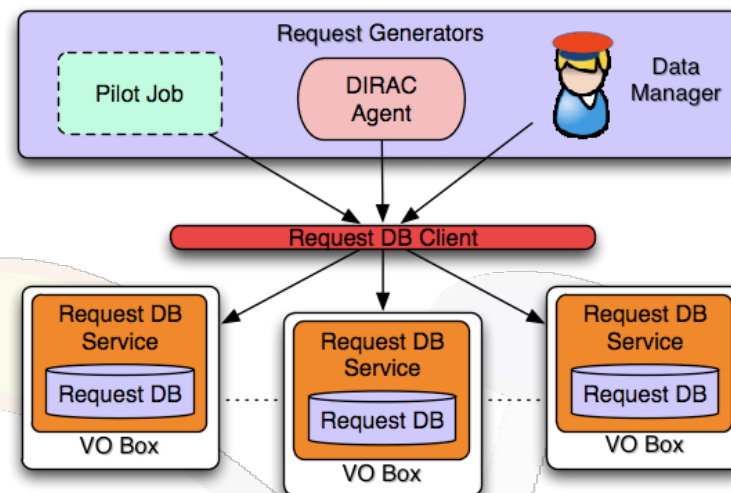


JobId	Status	MinorStatus	ApplicationStatus	Site	JobName	LastUpdate [UTC]	LastSignOffLife [UT]	SubmissionTime [L]	Owner
6824	Done	Execution Complet	Unknown	LCG.CERN.ch	Unknown	2009-02-21 23:51	2009-02-21 23:51	2009-02-21 23:45	atsareg
6823	Done	Execution Complet	Unknown	LCG.CERN.ch	Unknown	2009-02-21 23:52	2009-02-21 23:52	2009-02-21 23:45	atsareg
6822	Done	Execution Complet	Unknown	EELA.CIEMAT.es	Unknown	2009-02-21 22:38	2009-02-21 22:38	2009-02-21 22:16	atsareg
6821	Done	Execution Complet	Unknown	LCG.TCD.ie	Unknown	2009-02-21 22:24	2009-02-21 22:24	2009-02-21 22:16	atsareg
6820	Done	Execution Complet	Unknown	LCG.TCD.ie	Unknown	2009-02-21 22:18	2009-02-21 22:18	2009-02-21 22:16	atsareg
6819	Done	Execution Complet	Unknown	EELA.CIEMAT.es	Unknown	2009-02-21 22:29	2009-02-21 22:29	2009-02-21 21:43	atsareg
6818	Done	Execution Complet	Unknown	LCG.TCD.ie	Unknown	2009-02-21 22:02	2009-02-21 22:02	2009-02-21 21:43	atsareg
6817	Done	Execution Complet	Unknown	LCG.TCD.ie	Unknown	2009-02-21 22:20	2009-02-21 22:20	2009-02-21 21:43	atsareg



Request Management system

- ◆ A new Request Management System (RMS) to accept and execute asynchronously any kind of operation that can fail
 - ★ Data upload and registration
 - ★ Job status and parameter reports
- ◆ Request are collected by RMS instances on VO-boxes at 7 Tier-1 sites
 - ★ Extra redundancy in VO-box availability
- ◆ Requests are forwarded to the central Request Database
 - ★ For keeping track of the pending requests
 - ★ For efficient bulk request execution

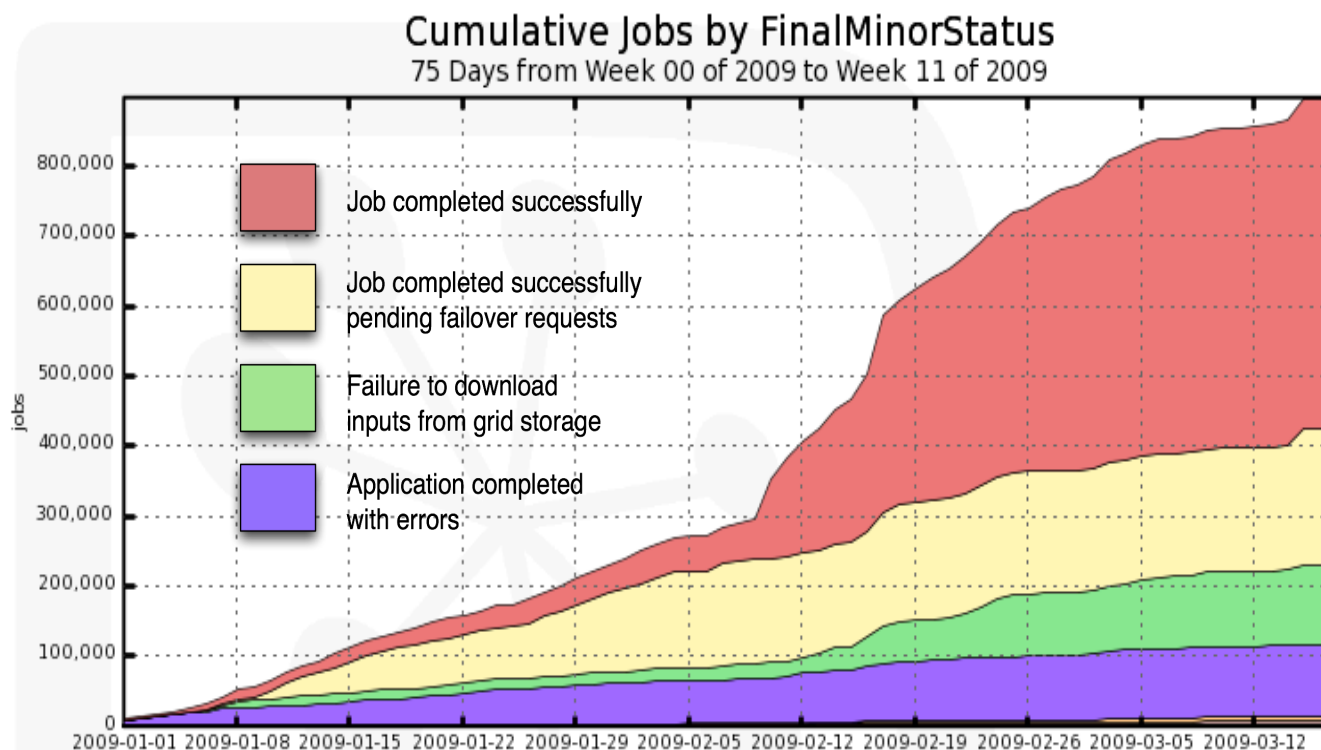


DIRAC



Failover redundancy

DIRAC



- ◆ Using failover system allows to recover almost all the successful jobs with incomplete operations
 - ✦ up to 20% of the total number of jobs
- ◆ Applied initially for production jobs
 - ✦ Now is being extended also for the user jobs

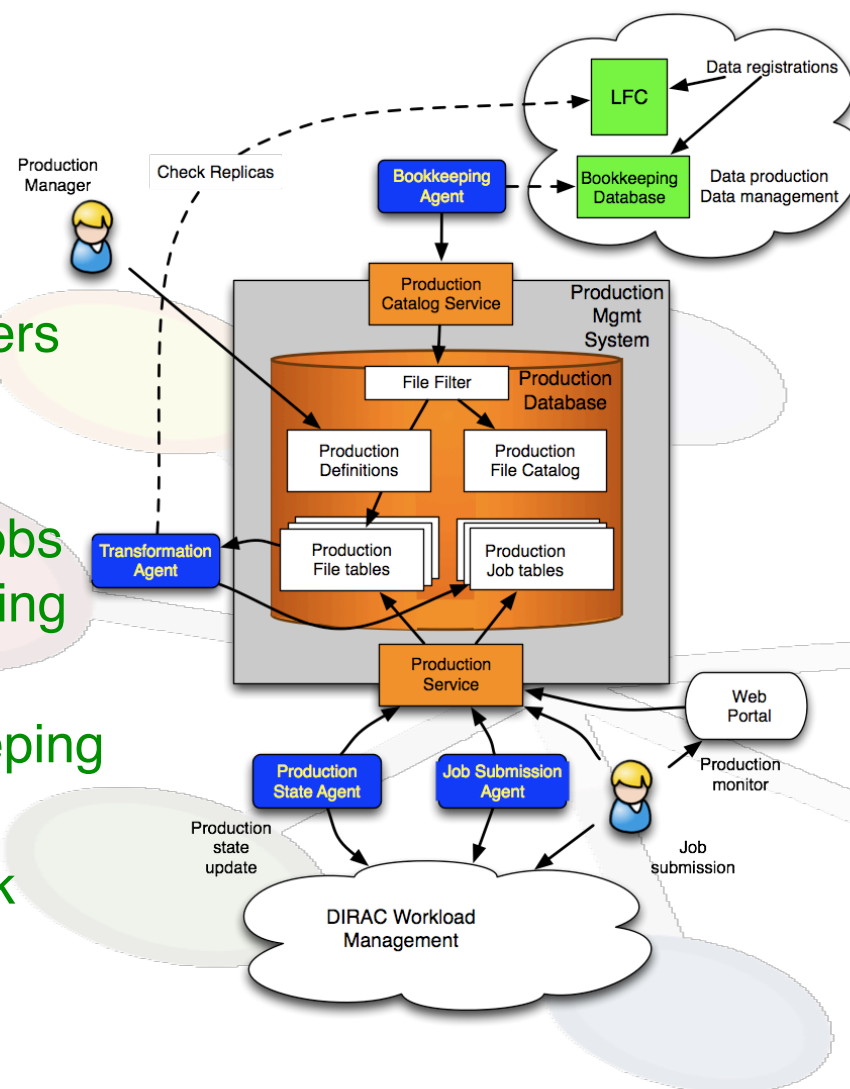


DIRAC: High level services

◆ Production Management built on top of the DIRAC WMS and DMS

- ★ Data requests formulated by users are processed and monitored using Web based tools
- ★ Automatic data reconstruction jobs creation and submission according to predefined scenarios
- ★ Interfaced to the LHCb Bookkeeping (metadata) Database
- ★ Built using the DSET framework

see [419] – A.Zhelezov et al

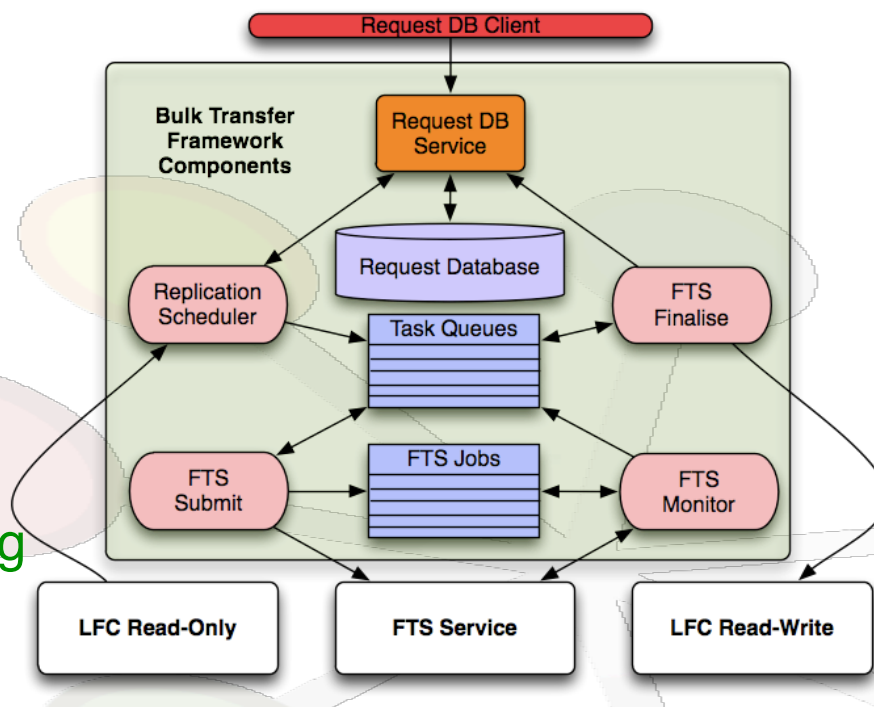




Data Management System

DIRAC

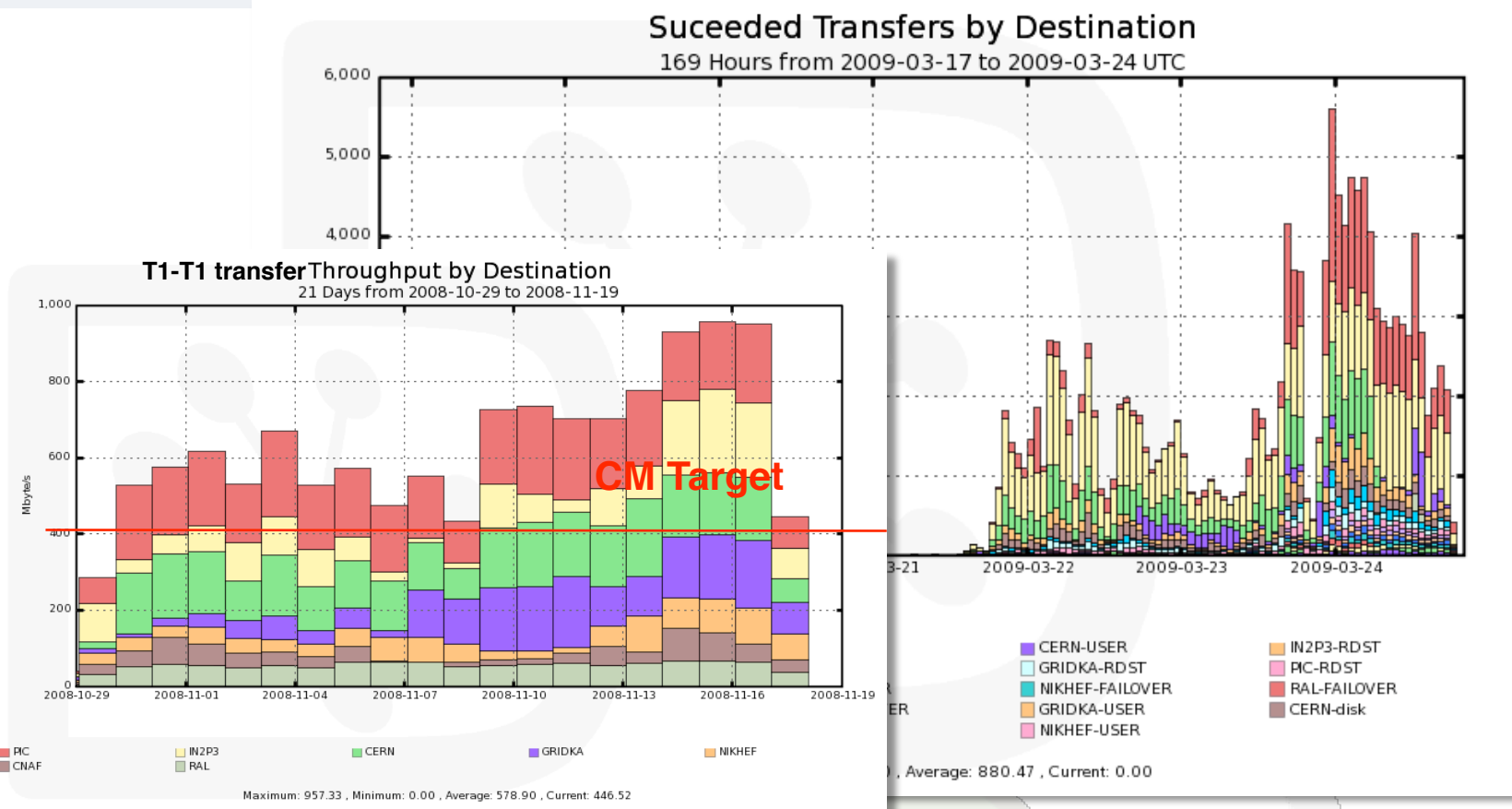
- ◆ All the Data Distribution operations
 - ✦ Pit to CERN transfers
 - ✦ T0-T1 transfers
 - ✦ T1-T1 transfers
- ◆ Based on the Request Management System
 - ✦ Automatic transfer scheduling
 - ✦ Full monitoring of ongoing operations
- ◆ Using FTS for bulk data transfers
 - ✦ Full failure recovery
- ◆ Comprehensive checks of data integrity in SEs and File Catalogs





Performance

DIRAC



- ◆ Extensively tested in a series of tests (CCRC, FEST'09)
 - ★ Proven to support the LHCb Computing Model targets

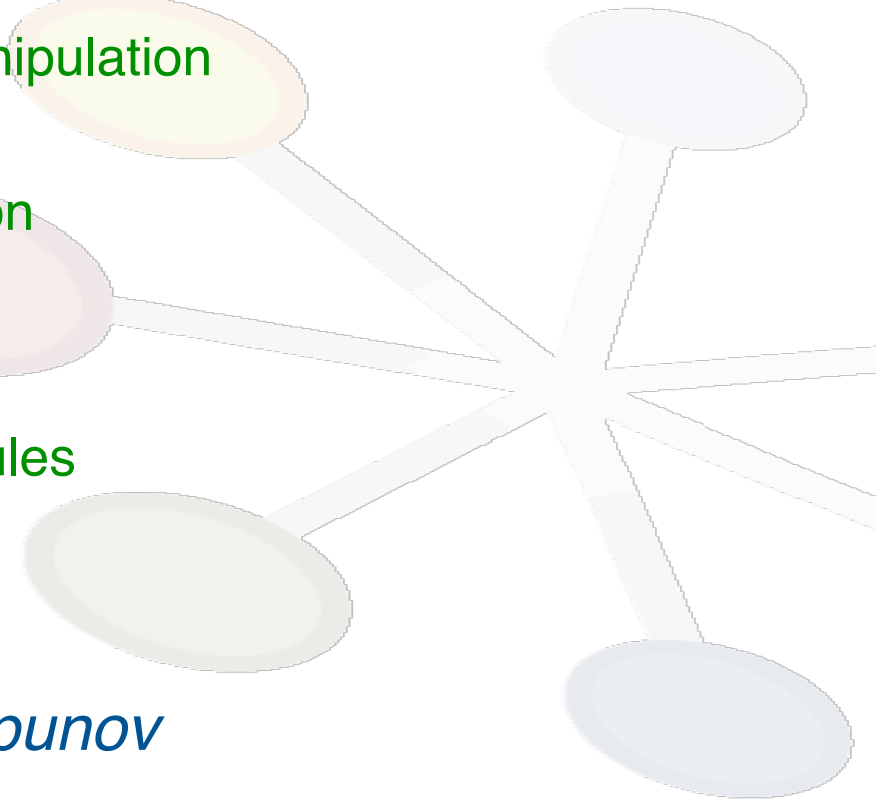


DIRAC: Secure Web Portal

DIRAC

- ◆ Web portal with intuitive desktop application like interface
- ◆ Monitoring and control of all activities
 - ✦ User job monitoring and manipulation
 - ✦ Data production controls
 - ✦ DIRAC Systems configuration
- ◆ Secure access
 - ✦ Standard grid certificates
 - ✦ Fine grained authorization rules
- ◆ This work is in progress

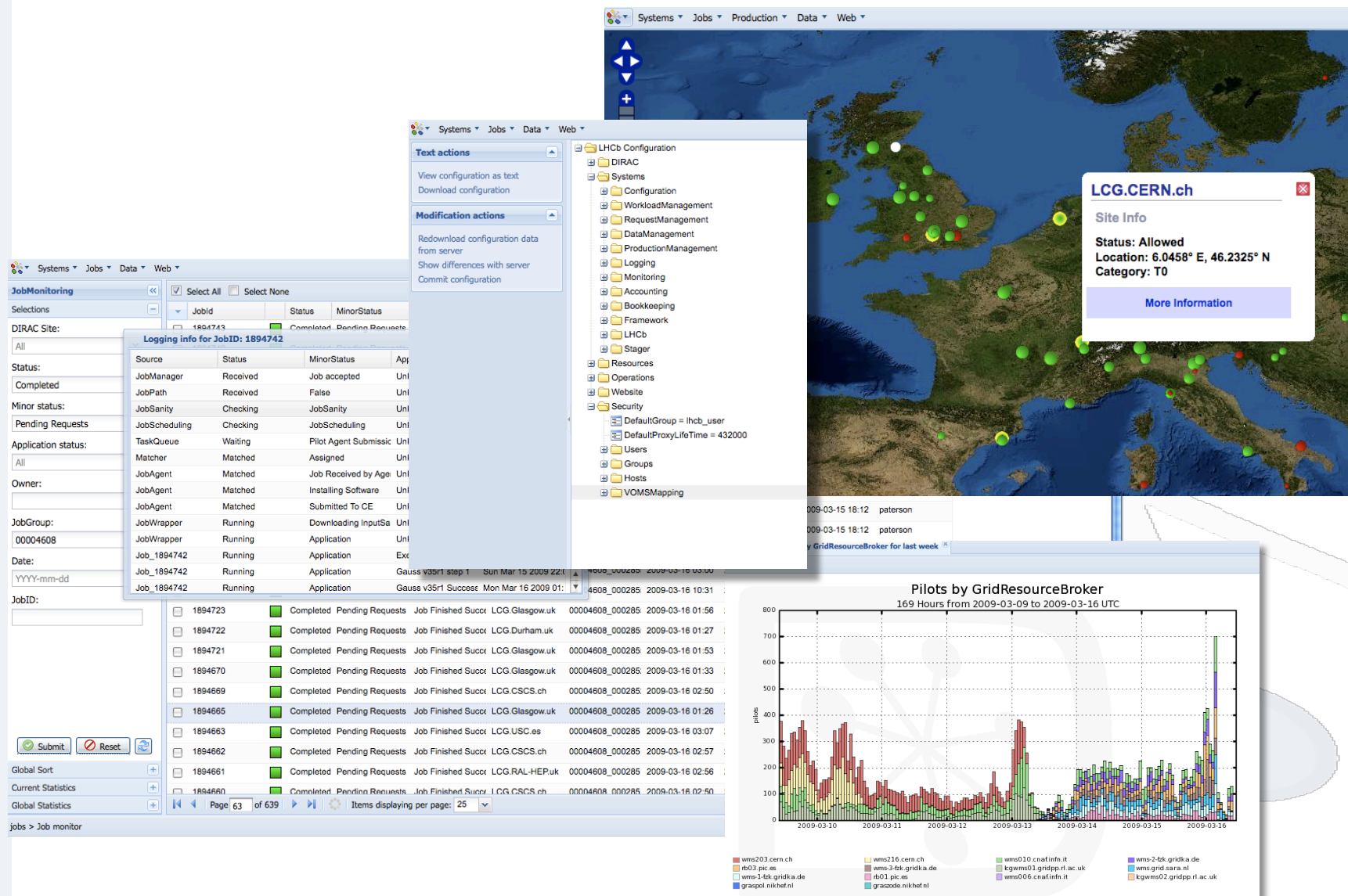
see [369] – A.Casajus, M.Sapunov





Web Portal: example interfaces

DIRAC





Other new services

- ◆ Bookkeeping Service
 - ✦ Moved into the DIRAC framework
 - ✦ *see [449] – Z.Mathe, E.Lanciotti*
- ◆ Monitoring Service
 - ✦ Monitoring DIRAC components
- ◆ System Logging Service
 - ✦ Collecting error reports from all the components
 - ✦ Error analysis and alarm setting alarms
- ◆ Notification Service
 - ✦ Notifying users and experts on predefined events

DIRAC



Conclusions

DIRAC

- ◆ DIRAC time proven general architecture is now realized on a solid base technology
- ◆ It covers all the LHCb tasks for the distributed data processing
- ◆ DIRAC demonstrates the performance adequate to the LHCb needs for real Data Taking
- ◆ DIRAC has evolved to a general purpose grid system, is considered to be applied for other grids and application domains

