



# *DIRAC status update*

*A. Tsaregorodtsev,  
CPPM-IN2P3-CNRS, Marseille  
For the DIRAC team*

6 October 2009, LHCb Software Week, CERN

# Outline

- ◆ DIRAC subsystem highlights
  - ✦ Workload management
  - ✦ Production Management
  - ✦ Data Management
- ◆ Web portal
- ◆ Hardware configuration
- ◆ Code management, release procedures
- ◆ Conclusions

# WMS update: Priorities

- ◆ Recently introduced priorities based on user and group shares
  - ✦ Statically defined first
  - ✦ Allowed to run concurrently user and production jobs with fair shares
- ◆ Turned out to be insufficient
  - ✦ The accounting for the history of the CPU usage is necessary
  - ✦ The appropriate corrections are implemented
    - More tuning needed

# WMS: New Sandbox service

- ◆ The old database based service started to be a bottle neck with the increased load
- ◆ New Sandbox service
  - ✦ Use any Storage Element as the backend storage
    - Sandboxes stored as compressed tar files
  - ✦ Is a Storage Element itself
  - ✦ Avoids duplication of the same sandboxes
    - Will be necessary for the bulk job submission
- ◆ Migration of the clients to use the new service
  - ✦ DIRAC version prior to v4r17 is deprecated
    - Version of the 22th July
  - ✦ With the next DIRAC release old clients will get error messages while job submission
  - ✦ Old clients will not be able to retrieve output sandboxes

# WMS: further developments

- ◆ Slow entry job control
  - ✦ Single threaded optimizers that check the job sanity
  - ✦ Will be made multithreaded
- ◆ Bulk job submission
  - ✦ Bunches of jobs sharing sandboxes can be submitted in one go
  - ✦ Sandbox service is already adapted

# Prospects: new resources

- ◆ More computing resources will become available
  - ✦ DIRAC sites resurrected
    - Using pilot jobs
    - Following security rules
- ◆ OSG grid
  - ✦ If we will get access to these resources
- ◆ LHCb Online farm
  - ✦ Potentially can double the LHCb available computing resources
  - ✦ Not a standard batch system
  - ✦ Should be included seamlessly into the DIRAC production system
    - Data reconstruction and MC simulation

# Production management

- ◆ Much effort spent on the creation, monitoring and validating productions
  - ✦ Production Request is prepared by the physics coordinators
  - ✦ Productions are run and monitored by the production managers
  - ✦ Monitoring tools are available to all the users

# Production Requests

- ◆ Production Requests are defined via a Web interface
  - ✦ Sim conditions, processing path, input data, event types, etc
  - ✦ Subrequests can be defined to quickly create similar productions
- ◆ The Requests are validated by physics coordinators and production experts
- ◆ Productions are created automatically afterwards
- ◆ The Production Request progress can be followed in the Web Portal
- ◆ The recent productions were already defined using this functionality
  - ✦ The new workflows are defined (stripping, merging)

# Production Request

The screenshot displays the LHCb Production Request web interface. The main window is titled 'Registered Production Requests' and shows a 'New request' form. The form includes the following fields and values:

- Name:** Arbitrary string for your convenience
- Type:** Reconstruction
- Priority:** 2b
- State:** New
- Author:** atsareg

The 'Input data' section includes fields for Conditions, Config, Processing Pass, and DQ flag. The 'Processing Pass (not registered yet)' section includes a Description field and a 'Step 1' section with fields for Application (Brunel), Version, CondDB, Option files, DDDDB, and Extra packages.

On the right side, a table displays a list of registered production requests:

Id	Type	State	Priority	Name	Sim/Run conditions	Proc. pass	Event type	Events requested
								30,000,00
30000000								-
30000000								-
91000000								-
								100,00
								400,00
								100,00
								400,00
								4,000,00
30000000								41,000,00
								10,000,00
								5,000,00
91000000								-
								2,000,00
								350,00
								47,800,00
46000014								50,00
								200,00
								200,00
30000000								180,000,00

The interface also includes a 'Comments' section and a 'New request' button. The bottom status bar shows the user's email address (atsareg@lhcb) and the current setup (LHCb-Production).

# Production Life Cycle Mgmt

- ◆ More elaborated production life cycle is introduced
- ◆ Extensive validation checks are performed after a production is completed:
  - ✦ BK information vs LFC
  - ✦ LFC vs BK, LFC vs SE
  - ✦ SE vs LFC
- ◆ Once production is validated, it is archived
  - ✦ Production definition is retained
- ◆ All the checks are fully automated

# Data Mgmt: Banning faulty SEs

- ◆ Storage Elements can be unavailable
  - ✦ Failures, scheduled or unscheduled shutdowns
- ◆ This should be taken into account
  - ✦ While job scheduling
    - User and production jobs
- ◆ SEs now can be declared as unavailable
  - ✦ This is equivalent to banning sites for jobs needing input data on these sites
  - ✦ For jobs without input data, the sites are still available

# Data Mgmt: Storage monitoring

The top screenshot shows a 'Storage Directory Summary' window with a table of directory paths and their sizes. The 'Replicas' and 'Size' columns are circled in red. The 'Directory' field is also circled in red.

Directory Path	Replicas	Size
/lhcb/user/a/acsmith/B+2DStar-D0_0002	1	169.9 MB
/lhcb/user/a/acsmith/B+2DStar-D0_0004/3129/3129082	1	4.4 kB
/lhcb/user/a/acsmith/B+2DStar-D0_0004/3129/3129083	1	4.4 kB
/lhcb/user/a/acsmith/B+2DStar-D0_0004/3129/3129084	1	4.4 kB
/lhcb/user/a/acsmith/B+2DStar-D0_0004/3129/3129085	1	4.4 kB
/lhcb/user/a/acsmith/B+2DStar-D0_0004/3129/3129087	1	4.4 kB
/lhcb/user/a/acsmith/B+2DStar-D0_0004/3129/3129088	1	4.4 kB
/lhcb/user/a/acsmith/B+2DStar-D0_0004/3129/3129089	1	4.4 kB

The bottom screenshot shows a 'Storage Directory Summary' window with a table of directory paths and their sizes. The 'SE Usage' table below it is also circled in red. The 'Directory' field is also circled in red.

Directory Path	Replicas	Size
/lhcb/MC/MC09/DST/00004838/0001	29607	30.3 TB
/lhcb/MC/MC09/DST/00004838/0000	26580	27.2 TB
/lhcb/MC/MC09/DST/00005113/0000	4847	22.1 TB
/lhcb/MC/MC09/DST/00005018/0000	4269	20.1 TB
/lhcb/MC/MC09/DST/00005016/0000	4257	20 TB
/lhcb/MC/MC09/DST/00004838/0002	19207	19.7 TB
/lhcb/MC/MC09/DST/00005015/0000	3660	13.4 TB
/lhcb/MC/MC09/DST/00005013/0000	2424	11.3 TB
/lhcb/MC/MC09/DST/00005103/0000	1863	8.4 TB
/lhcb/MC/MC09/DST/00005071/0000	1410	6.4 TB
/lhcb/MC/MC09/DST/00005017/0000	1383	6.4 TB
/lhcb/MC/MC09/DST/00005112/0000	1128	4.6 TB
/lhcb/MC/MC09/DST/00004952/0000	1010	4.6 TB
/lhcb/MC/MC09/DST/00004953/0000	997	4.5 TB
/lhcb/MC/MC09/DST/00005102/0000	939	4.2 TB
/lhcb/MC/MC09/DST/00005005/0000	813	3.8 TB
/lhcb/MC/MC09/DST/00004987/0000	846	3.8 TB
/lhcb/MC/MC09/DST/00004827/0000	765	3.2 TB
/lhcb/MC/MC09/DST/00005014/0000	699	3.2 TB
/lhcb/MC/MC09/DST/00004951/0000	686	3.1 TB
/lhcb/MC/MC09/DST/00004981/0000	429	1.9 TB
/lhcb/MC/MC09/DST/00005009/0000	402	1.8 TB
/lhcb/MC/MC09/DST/00005075/0000	381	1.7 TB
/lhcb/MC/MC09/DST/00004954/0000	378	1.7 TB
/lhcb/MC/MC09/DST/00005073/0000	375	1.7 TB

SE Usage	Replicas	Size
CERN_MC_M-DST	54780	92.7 TB
CNAF_MC-DST	12924	15.6 TB
PIC_MC_M-DST	11704	18.6 TB
NIKHEF_MC-DST	10165	16.3 TB
RAL_MC-DST	10155	16.6 TB
PIC_MC-DST	8345	14.7 TB
IN2P3_MC-DST	8172	14.3 TB
IN2P3_MC_M-DST	8137	17.1 TB
GRIDKA_MC-DST	7868	14.3 TB
RAL_MC_M-DST	7627	11.1 TB
GRIDKA_MC_M-DST	7541	17.4 TB
CNAF_MC_M-DST	7180	14.7 TB
NIKHEF_MC_M-DST	6588	11 TB
RAL-FAILOVER	17	1.5 GB
NIKHEF-FAILOVER	15	1.4 GB
IN2P3-FAILOVER	14	1.3 GB
GRIDKA-FAILOVER	14	5.4 GB
CNAF-FAILOVER	13	1 GB
PIC-FAILOVER	8	632.6 MB
CERN-FAILOVER	4	4.8 GB
CNAF-DST	1	4.8 GB

# DMS: User Storage quotas

- ◆ Storage space on the grid is not unlimited
  - ✦ Users are supposed to clean their spaces but rarely do
    - Unless they are notified about exceeding quotas
- ◆ The user storage consumption is periodically checked by a dedicated agent
  - ✦ The results are available to users
  - ✦ Currently they can be just consulted
    - Command line and API tools available
  - ✦ Eventually the user space will be locked for writing if the quotas are exceeded
    - The quotas are defined in the CS per user

# DIRAC: Secure Web Portal

- ◆ Web portal with intuitive desktop application like interface
  - ✦ Ajax, Pylons, ExtJS Javascript library
- ◆ 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

# Web Portal: example interfaces

The screenshot displays a web portal interface for job monitoring and configuration. The main window is titled "JobMonitoring" and shows a table of jobs with columns for JobID, Status, and MinorStatus. A "Logging info for JobID: 1894742" popup is visible, showing a table of job details. A "Text actions" popup is also present, listing actions like "View configuration as text" and "Download configuration".

On the right, a map of Europe shows the location of the LCG.CERN.ch site. A popup window titled "LCG.CERN.ch" provides site information:

**LCG.CERN.ch**  
 Site Info  
 Status: Allowed  
 Location: 6.0458° E, 46.2325° N  
 Category: T0  
 More Information

Below the map, a "Pilots by GridResourceBroker" chart shows the number of pilots over time from 2009-03-09 to 2009-03-16 UTC. The chart is a stacked bar chart with a legend at the bottom listing various sites and their pilot counts.

JobID	Status	MinorStatus	App
1894743	Completed	Pending Requests	Job Finished Succ
1894723	Completed	Pending Requests	Job Finished Succ
1894722	Completed	Pending Requests	Job Finished Succ
1894721	Completed	Pending Requests	Job Finished Succ
1894670	Completed	Pending Requests	Job Finished Succ
1894669	Completed	Pending Requests	Job Finished Succ
1894665	Completed	Pending Requests	Job Finished Succ
1894663	Completed	Pending Requests	Job Finished Succ
1894662	Completed	Pending Requests	Job Finished Succ
1894661	Completed	Pending Requests	Job Finished Succ
1894660	Completed	Pending Requests	Job Finished Succ

# Web Portal: new features

- ◆ Multiple developments to improve current views and to add new ones
  - ✦ Production requests, monitoring
  - ✦ Pilot monitoring
  - ✦ Etc
- ◆ Site Map resurrected
  - ✦ Mostly PR, but can be a convenient tool to navigate to the relevant site info
- ◆ Bookkeeping query page
  - ✦ Close to the Qt/GUI interface and functionality

# DIRAC: Bookkeeping page

The screenshot displays the DIRAC Bookkeeping interface. On the left is a 'DIRAC SideBar' with a 'BK browser' showing a tree view of data folders. The main area is a table with columns: #, File Name, Event Stat, File Size, Job Start, and Job End. The table lists 25 entries, all with 'Event Stat' of 9000 and 'Job Start' of 'Wed Sep 17 2008 00:'. On the right, 'Bookkeeping info' shows configuration details for 'LHCb' and 'Cosmics', including simulation conditions (1019), processing pass (First Beam), event type (30000000), and file type (RAW). A 'Statistics' section shows 258 files, 2 322 000 events, and a total size of 411.2 GB. The status bar at the bottom indicates 'Displaying 1 - 25 of 258' and 'Page 1 of 11'.

#	File Name	Event Stat	File Size	Job Start	Job End
1	/lhcb/data/2008/RAW/	9000	1887561036	Wed Sep 17 2008 00:	Wed Sep 17 2008
2	/lhcb/data/2008/RAW/	9000	1887537100	Wed Sep 17 2008 00:	Wed Sep 17 2008
3	/lhcb/data/2008/RAW/	9000	1887489532	Wed Sep 17 2008 00:	Wed Sep 17 2008
4	/lhcb/data/2008/RAW/	9000	1887564324	Wed Sep 17 2008 00:	Wed Sep 17 2008
5	/lhcb/data/2008/RAW/	9000	1887511908	Wed Sep 17 2008 00:	Wed Sep 17 2008
6	/lhcb/data/2008/RAW/	9000	1887596536	Wed Sep 17 2008 00:	Wed Sep 17 2008
7	/lhcb/data/2008/RAW/	9000	1887637596	Wed Sep 17 2008 00:	Wed Sep 17 2008
8	/lhcb/data/2008/RAW/	9000	1887517372	Wed Sep 17 2008 00:	Wed Sep 17 2008
9	/lhcb/data/2008/RAW/	9000	1887502168	Wed Sep 17 2008 00:	Wed Sep 17 2008
10	/lhcb/data/2008/RAW/	9000	1887563448	Wed Sep 17 2008 00:	Wed Sep 17 2008
11	/lhcb/data/2008/RAW/	9000	1887645424	Wed Sep 17 2008 00:	Wed Sep 17 2008
12	/lhcb/data/2008/RAW/	9000	1887609956	Wed Sep 17 2008 00:	Wed Sep 17 2008
13	/lhcb/data/2008/RAW/	9000	1887728000	Wed Sep 17 2008 00:	Wed Sep 17 2008
14	/lhcb/data/2008/RAW/	9000	1887695240	Wed Sep 17 2008 00:	Wed Sep 17 2008
15	/lhcb/data/2008/RAW/	9000	1887518476	Wed Sep 17 2008 00:	Wed Sep 17 2008
16	/lhcb/data/2008/RAW/	9000	1887458084	Wed Sep 17 2008 00:	Wed Sep 17 2008
17	/lhcb/data/2008/RAW/	9000	1887765984	Wed Sep 17 2008 00:	Wed Sep 17 2008
18	/lhcb/data/2008/RAW/	9000	1887477456	Wed Sep 17 2008 00:	Wed Sep 17 2008
19	/lhcb/data/2008/RAW/	9000	1887462724	Wed Sep 17 2008 00:	Wed Sep 17 2008
20	/lhcb/data/2008/RAW/	9000	1887758020	Wed Sep 17 2008 00:	Wed Sep 17 2008
21	/lhcb/data/2008/RAW/	9000	1887546524	Wed Sep 17 2008 00:	Wed Sep 17 2008
22	/lhcb/data/2008/RAW/	9000	1887541640	Wed Sep 17 2008 00:	Wed Sep 17 2008
23	/lhcb/data/2008/RAW/	9000	1887600160	Wed Sep 17 2008 00:	Wed Sep 17 2008
24	/lhcb/data/2008/RAW/	9000	1887695240	Wed Sep 17 2008 00:	Wed Sep 17 2008
25	/lhcb/data/2008/RAW/	9000	1887494220	Wed Sep 17 2008 00:	Wed Sep 17 2008

◆ In production this week

✦ Last minute bug in the BK service communication

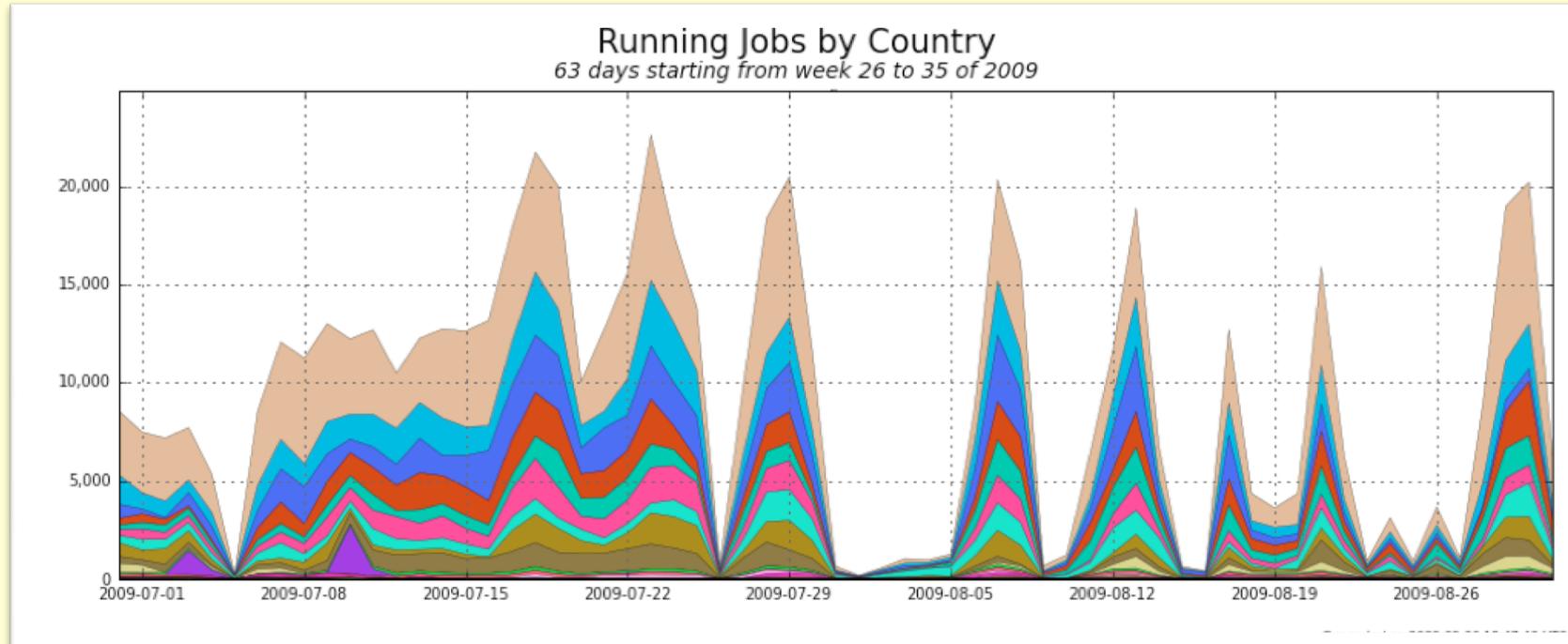
# Web Portal: further developments

- ◆ Web portal development will continue
- ◆ New views
  - ✦ Summaries of the site, resources, user info
- ◆ User profiles
  - ✦ Per user customizations
- ◆ Improved graphics
- ◆ More interactivity
  - ✦ Simple job submission
  - ✦ Access to the job sandboxes

# Hardware update

- ◆ The single central DIRAC WMS/Production server was split in 3:
  - ✦ volhcb04 - security sensitive services
    - Proxy repository, Security logging
  - ✦ volhcb13 - WMS components
    - More powerful with two HDDs
  - ✦ volhcb09 - Production Mgmt components
- ◆ The overall system capacity increased by a factor  $\sim 2$ .

# Performance



- ◆ DIRAC performance measured in the recent production and FEST'09 runs
  - ✦ Up to 25K concurrent jobs in ~120 distinct sites
    - Up to 75K jobs per day
  - ✦ Further optimizations to increase capacity are possible
    - Hardware, database optimizations, service load balancing, etc

# DIRAC Central Servers

## ◆ Front-end DIRAC servers

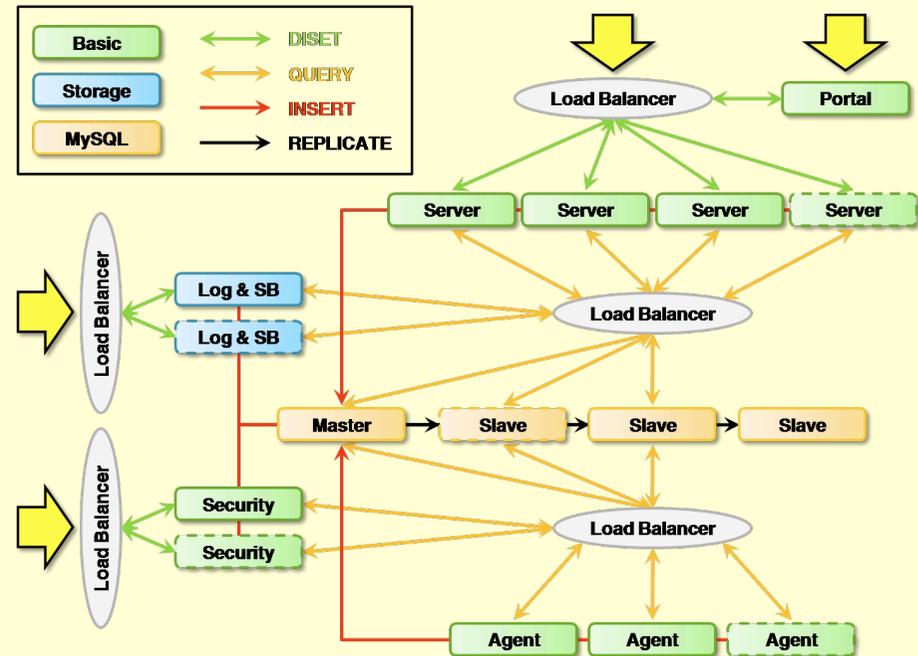
- ◆ DNS or client level load balancing

## ◆ Database

- ◆ Multiple MySQL databases
- ◆ Real time mirroring of Master to Slave databases
- ◆ Redundant disk arrays
- ◆ Load balancing on the database client level
- ◆ ORACLE solution is considered as well

## ◆ Expected operations advantages

- ◆ Large capacity margin ( factor 2-3 )
- ◆ Failure of any server will reduce but not disrupt operations, hot swapping of any server



# Distribution, installation

- ◆ DIRAC is now fully available on the SLC5 platform
  - ✦ Shipped grid middleware binaries
- ◆ Use of Python 2.5 is still subject to the availability of the python grid middleware bindings
  - ✦ Memory leaks reported
- ◆ CA, CRL renewal automation
  - ✦ Command line tools
    - Useful for the user client installation
  - ✦ Automatic agent
    - For service installations

# Certification DIRAC instance

- ◆ DIRAC releases are now first tested on the Certification setup
  - ✦ Separate servers for the WMS services and for the Web Portal
  - ✦ Other systems will be added as well
- ◆ The quality of the releases should be much improved

# DIRAC v5r0 release

- ◆ A new major DIRAC v5r0 release is being prepared
- ◆ Updated binary packages
  - ✦ GSI
  - ✦ MySQL
  - ✦ Matplotlib graphics library
- ◆ Discontinued old Sandbox service
- ◆ No backward compatibility with the deprecated versions of DIRAC
  - ✦ Prior to v4r17

# Prospects: code consolidation

- ◆ Many small improvements to fix problems and improve performance and stability
- ◆ New hardware configuration will be supported in the code
  - ✦ Multiple service end-points
  - ✦ Multiple backend database instances
    - Read/write and read only
- ◆ More Database optimizations are envisaged

# Moving to SVN repository

- ◆ The DIRAC code is stored in the CVS repository
  - ✦ Discontinued in the fall 2009
- ◆ The code will be moved to the SVN repository
- ◆ The code will be restructured
  - ✦ Several independent projects with separate release cycle:
    - Core DIRAC, LHCbSystem, Web Portal
- ◆ The new repository structure is being prototyped now

# Conclusions

- ◆ DIRAC covers all the LHCb tasks for the distributed data processing implemented in a single coherent framework
- ◆ Many efforts going into the system polishing and code consolidation
- ◆ DIRAC demonstrates the performance adequate to the LHCb needs for real Data Taking

*<http://dirac.cern.ch>*