Click to edit Master title style

DIRAC in LHCb and beyond

Philippe Charpentier (LHCb-CERN)
Slides courtesy of A.Tsaregorodtsev
BigPanda Workshop, October 21 2013

DIRAC Grid Solution

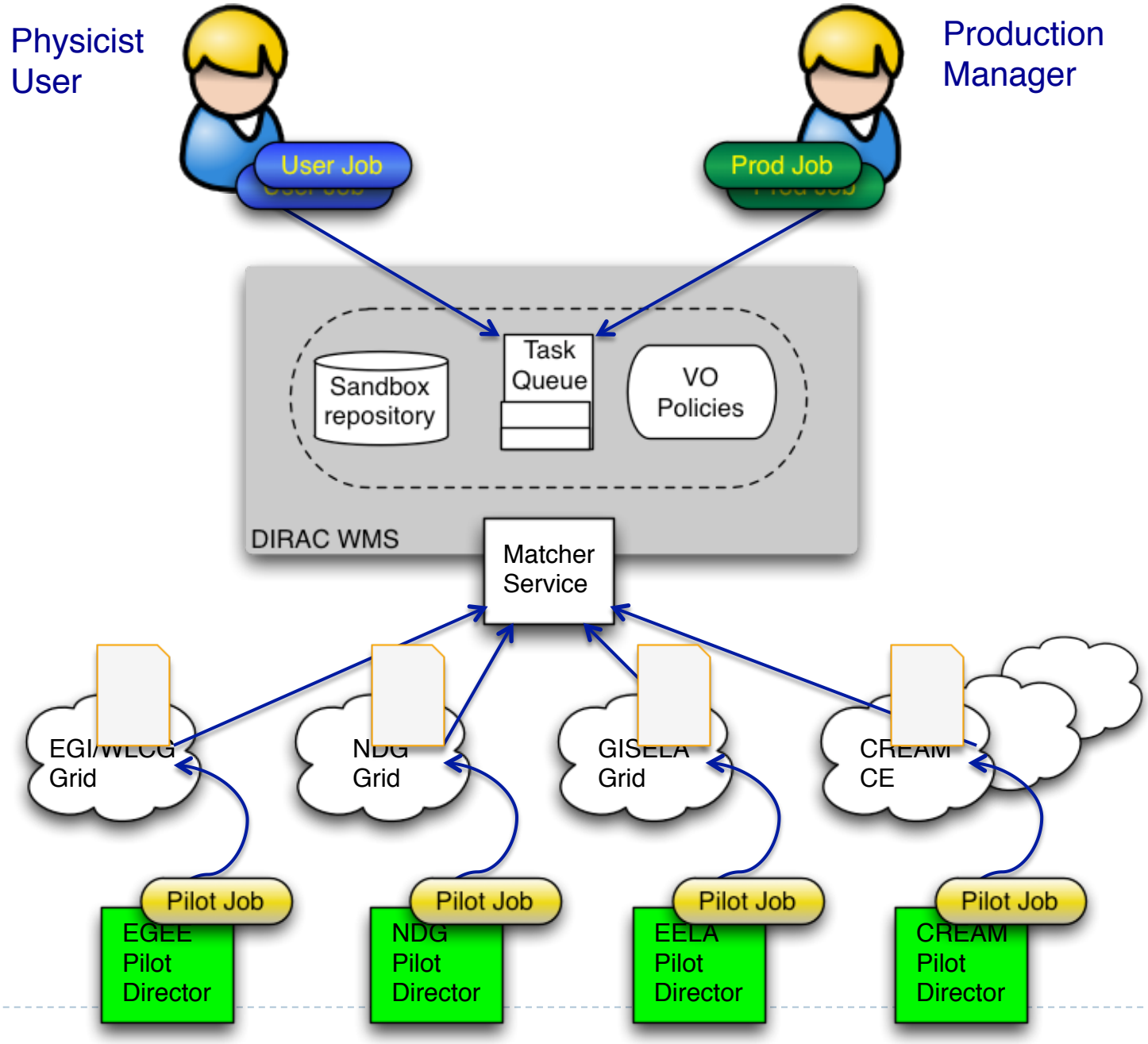
- LHC experiments, all developed their own middleware to address the above problems
  - PanDA, AliEn, glideIn WMS, PhEDEx, …
- DIRAC is developed originally for the LHCb experiment with the goals:
  - Integrate all the heterogeneous computing resources available to the community
  - Provide solution for both WMS and DMS tasks
  - Minimize human intervention at sites providers of resources
  - Make the grid convenient for the users:
    - Simpler intuitive interfaces
    - Fault tolerance, quicker turnaround of user jobs
    - Enabling Community policies

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
- The results of this work are presented in the following

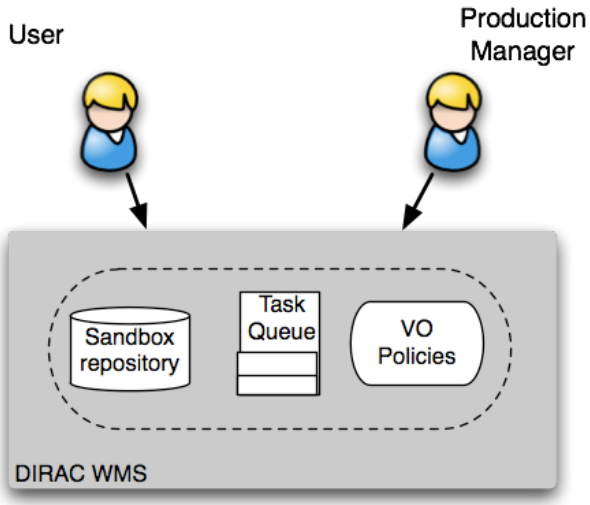
DIRAC
THE INTERWARE

DIRAC Workload Management


Physicist
User
Production
Manager
User Job
Prod Job
Task
Queue
Sandbox
repository
VO
Policies
DIRAC WMS
Matcher
Service
Grid
Grid
Grid
CE
Pilot Job
EGEE
Pilot
Director
Pilot Job
NDG
Pilot
Director
Pilot Job
EELA
Pilot
Director
Pilot Job
CREAM
Pilot
Director

WMS: applying VO policies

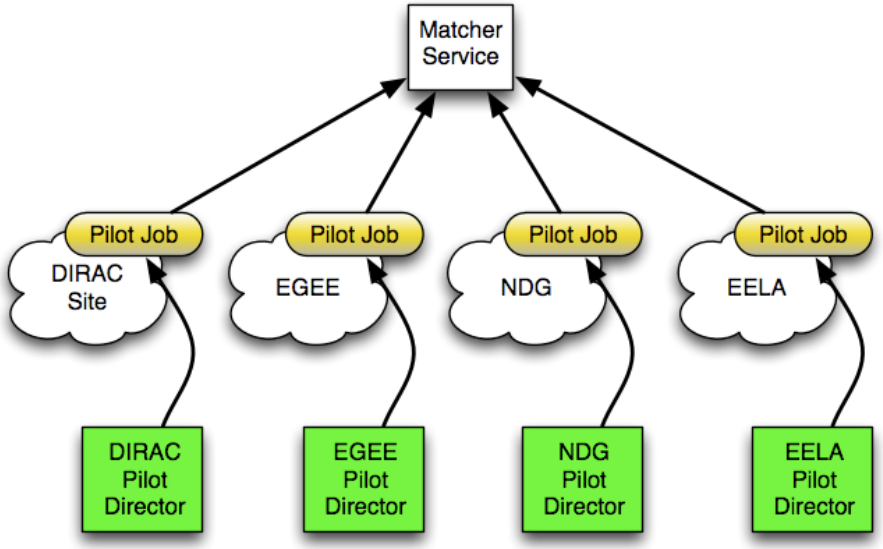
- In DIRAC both User and Production jobs are treated by the same WMS
  - Same Task Queue
- This allows to apply efficiently policies for the whole VO
  - Assigning Job Priorities for different groups and activities
  - Static group priorities are used currently
  - More powerful scheduler can be plugged in
    - demonstrated with MAUI scheduler



- The VO policies application in the central Task Queue dictates the use of Multiuser Pilot Agents
  - Do not know apriori whose job has the highest priority at the moment of the user job matching
- DIRAC fully supports this mode of operation
  - Multiuser Pilots Jobs submitted with a special “pilot” VOMS role
  - Using glexec on the WNs to track the identity of the payload owner

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
  - Users just see new sites appearing in the job monitoring



DIRAC
THE INTERWARE

DIRAC as a resource manager

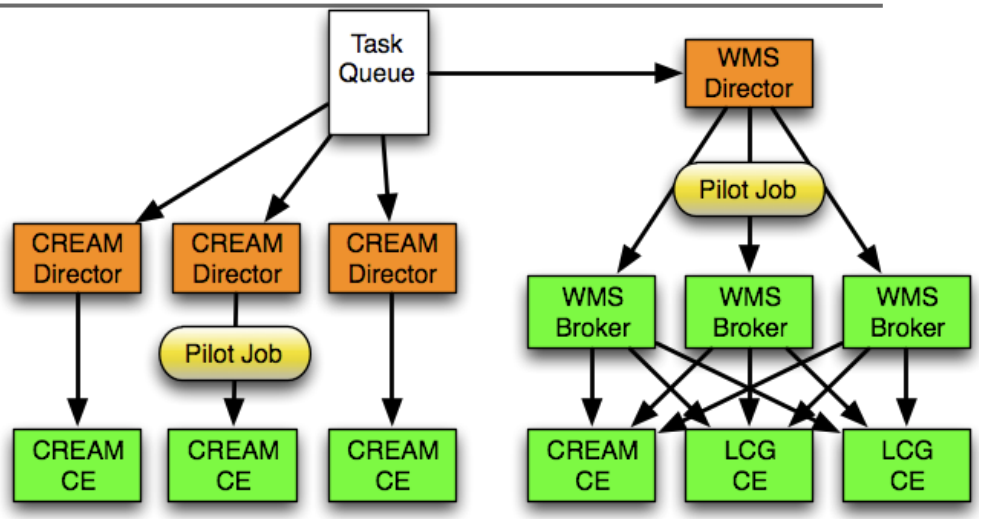
DIRAC
THE INTERWARE

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
- The work is in progress to support ARC middleware based grids
  - NorduGrid
  - A successful demonstration was already done
- Other types of grids can be supported
  - As long we have customers needing that

Direct submission to CEs

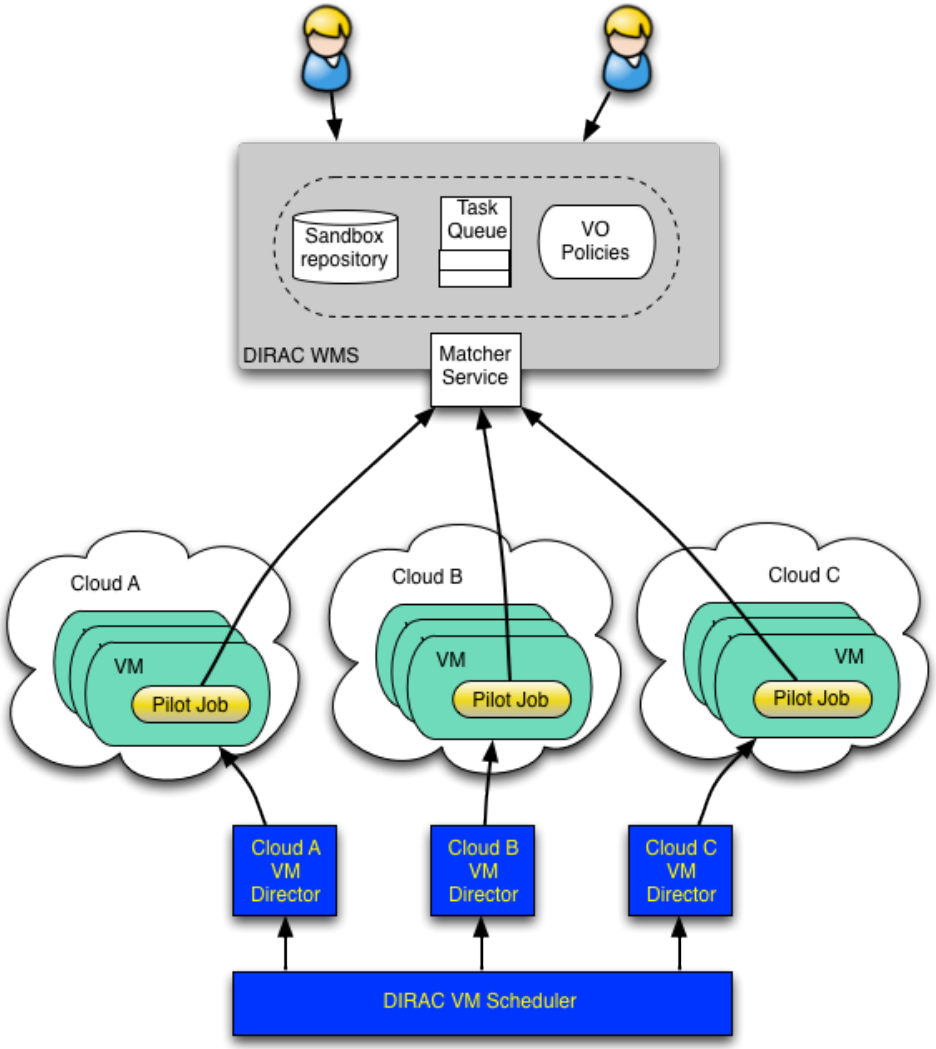
- Using gLite WMS now just as a pilot deployment mechanism
  - Limited use of brokering features
    - For jobs with input data the destination site is already chosen
  - Have to use multiple Resource Brokers because of ***scalability*** problems



- DIRAC is supporting direct submission to CEs
  - CREAM CEs or batch clusters through SSH tunnel
  - Can apply individual site policy
    - Site chooses how much load it can take (*Pull* vs *Push* paradigm)
  - Direct measurement of the site state watching the pilot status info
- This is a general trend
  - All the LHC experiments declared abandoning eventually gLite WMS

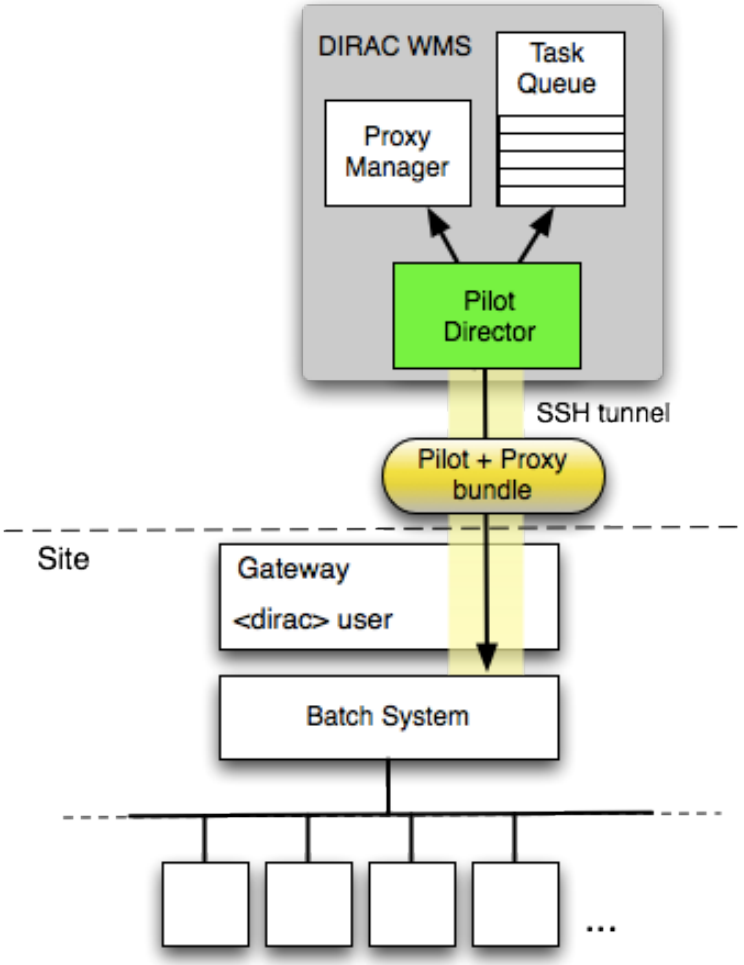
Clouds

- VM scheduler developed for Belle MC production system
  - Dynamic VM spawning taking Amazon EC2 spot prices and Task Queue state into account
  - 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



Standalone computing clusters

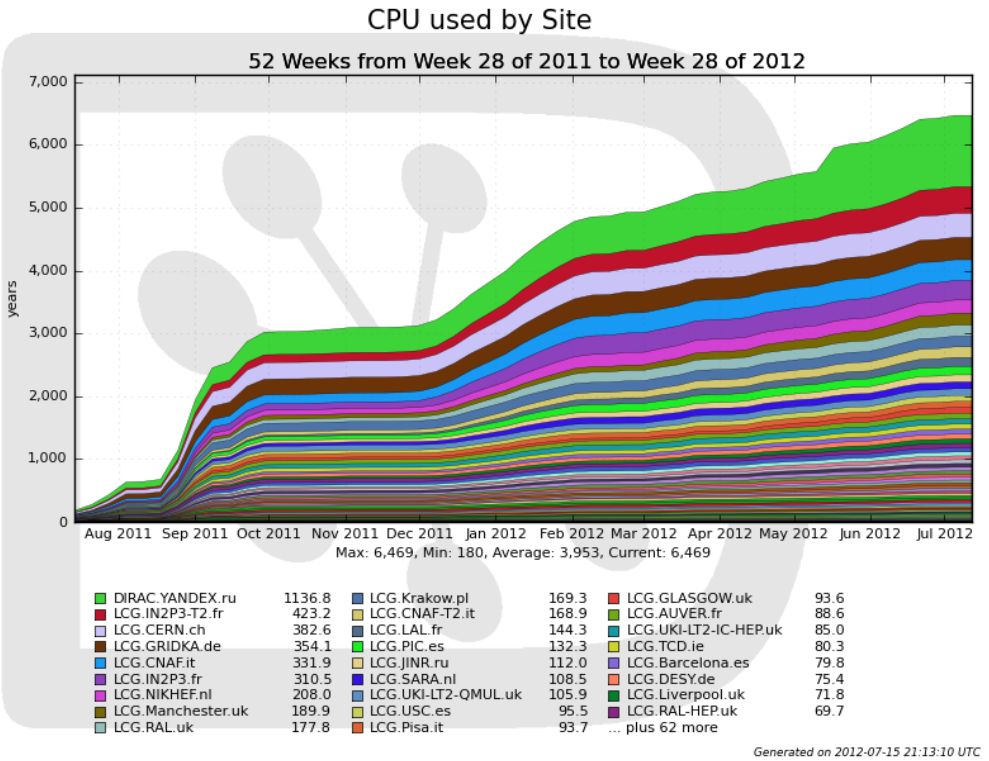
- Dedicated Pilot Director per group of sites
- Off-site Director
  - Site delegates control to the central service
  - Site must only define a dedicated local user account
  - The payload submission through the SSH tunnel
- The site can be a single computer or a cluster with a batch system
  - LSF, BQS, SGE, PBS/Torque, Condor
  - More to come:
    - OAR, SLURM, LoadLeveler. etc
- The user payload is executed with the owner credentials
  - No security compromises with respect to external services



Standalone computing clusters

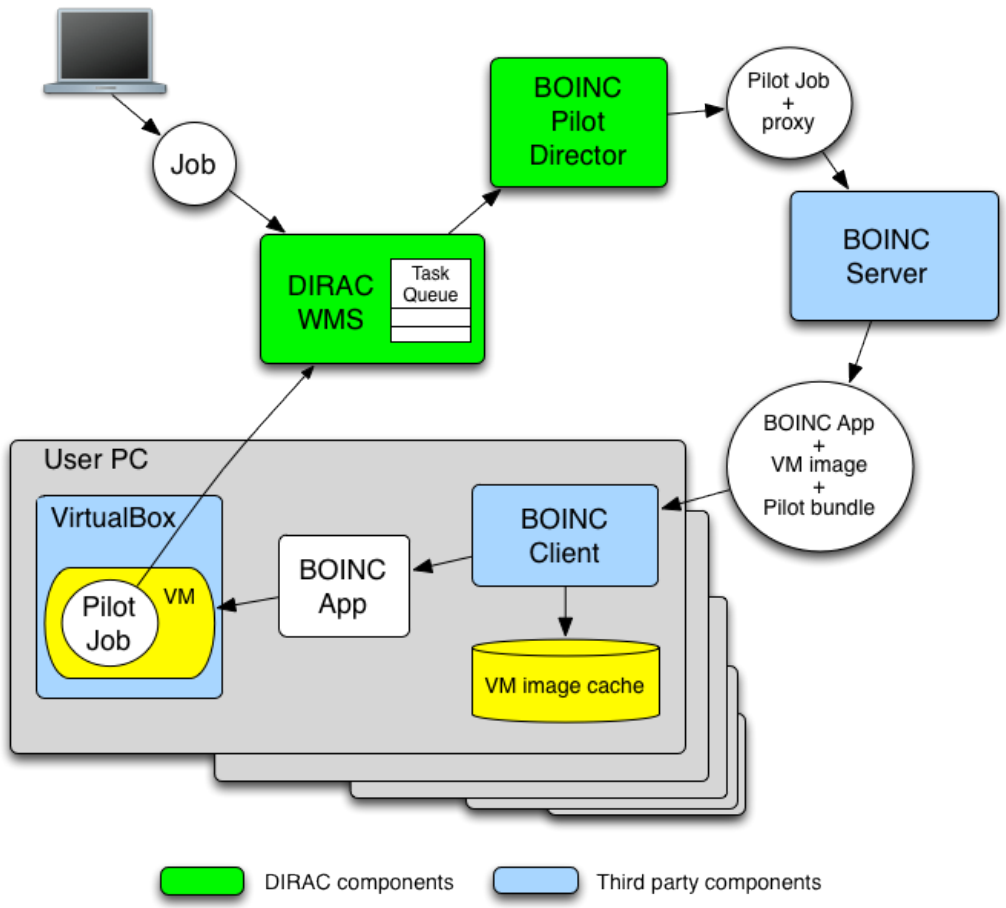
Examples:

- 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)



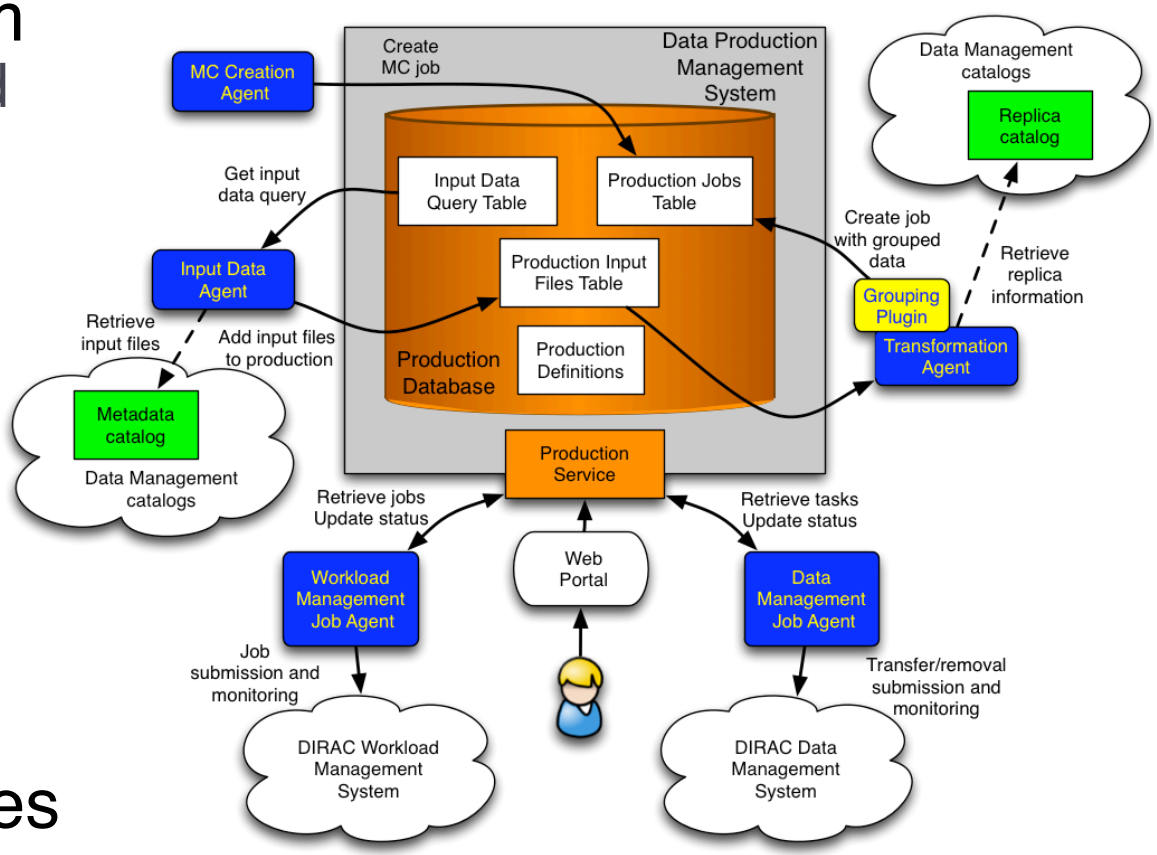
BOINC Desktop Grids

- 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
- Work on interfacing DIRAC to EDGI resources is in progress



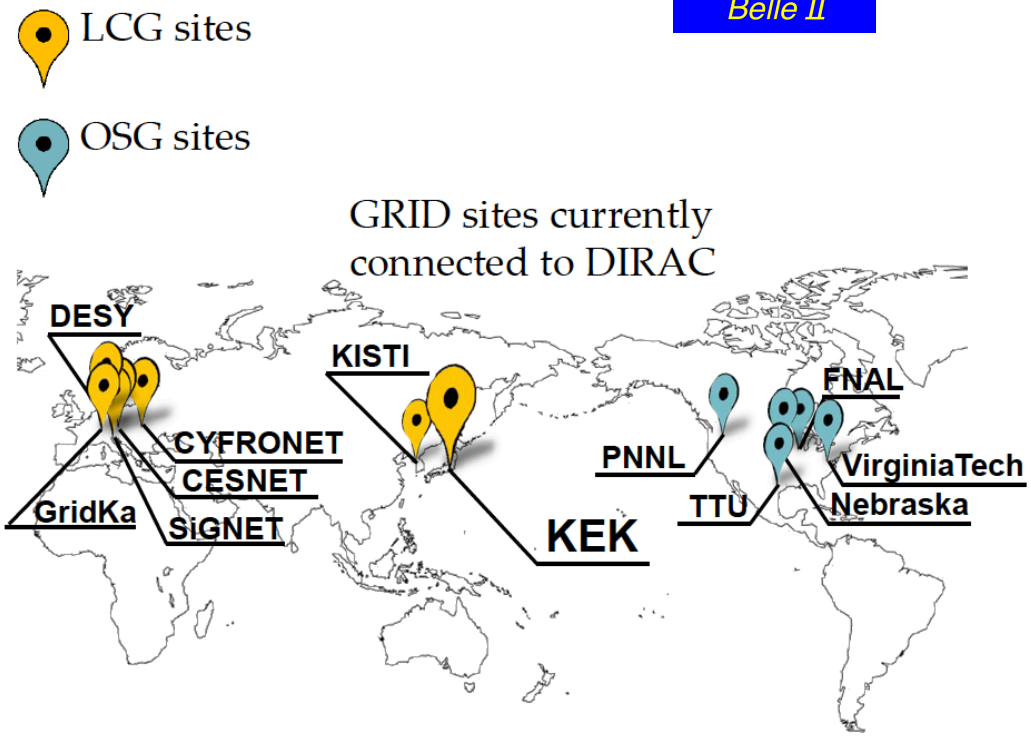
LHCb Production system

- Based on the DIRAC Transformation System
  - Multiple extensions and custom plugins
- Data driven payload generation based on templates
- Generating data processing and replication tasks
- LHCb specific templates and catalogs



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

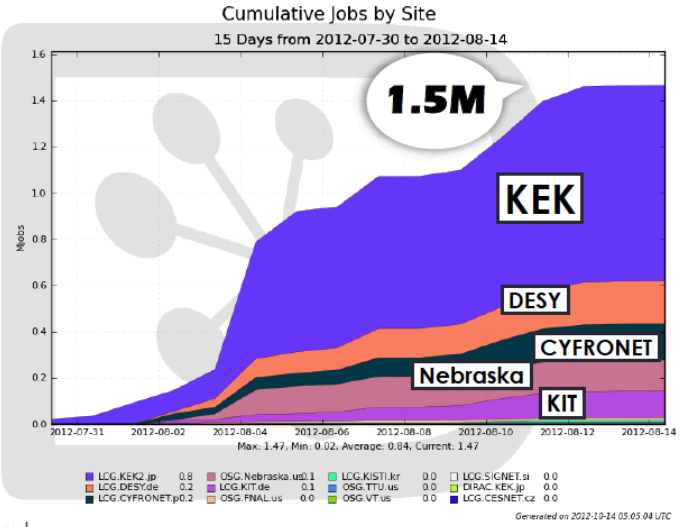


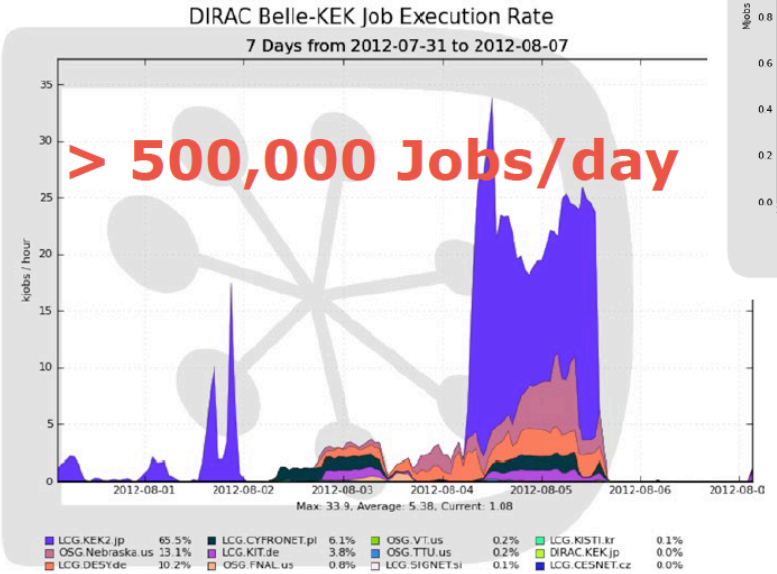
More resources to be incorporated (including clouds and local clusters)

Hideki Miyake, KEK

DIRAC Scalability tests

- Random number generation (500/job) or just filling pilot job
  →no SE/AMGA used
- Good performance
  - Even saturated KEKCC GRID
- DIRAC itself was stable





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 as a service

- DIRAC client is easy to install
  - Part of a usual tutorial
- DIRAC services are easy to install but
  - Needs dedicated hardware for hosting
  - Configuration, maintenance needs expert manpower
  - Monitoring computing resources
- Small user communities can not afford maintaining dedicated DIRAC services
  - Still need easy grid access
- Large grid infrastructures can provide DIRAC services for their users.

France-Grid DIRAC service

- Started as a support for user tutorials

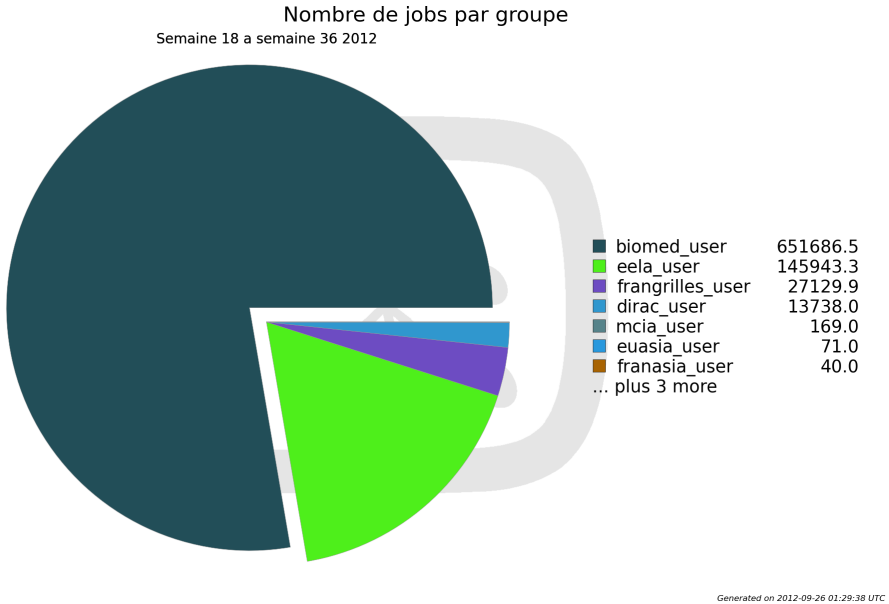
- Several regional and university campus installations
  - Complex maintenance

- 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 ~3 millions jobs went through the system
      - Mostly biomed applications



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