



Enabling Grids for E-scienceE

Application Porting using DRMAA and Globus GridWay on the EGEE Infrastructure

J.L. Vázquez-Poletti (UCM)

EGEE08 (Istambul)

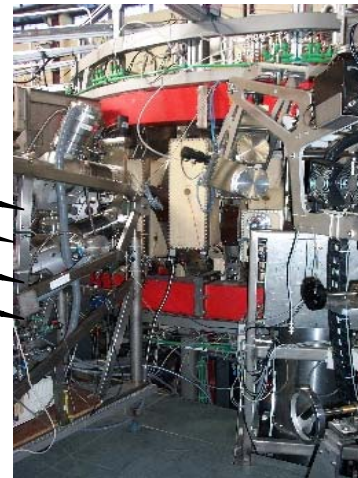
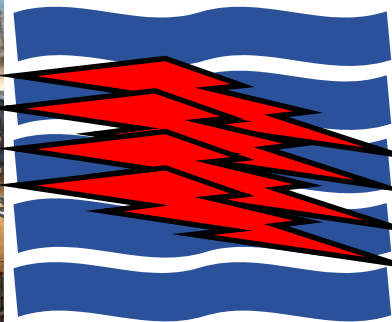
www.eu-egee.org

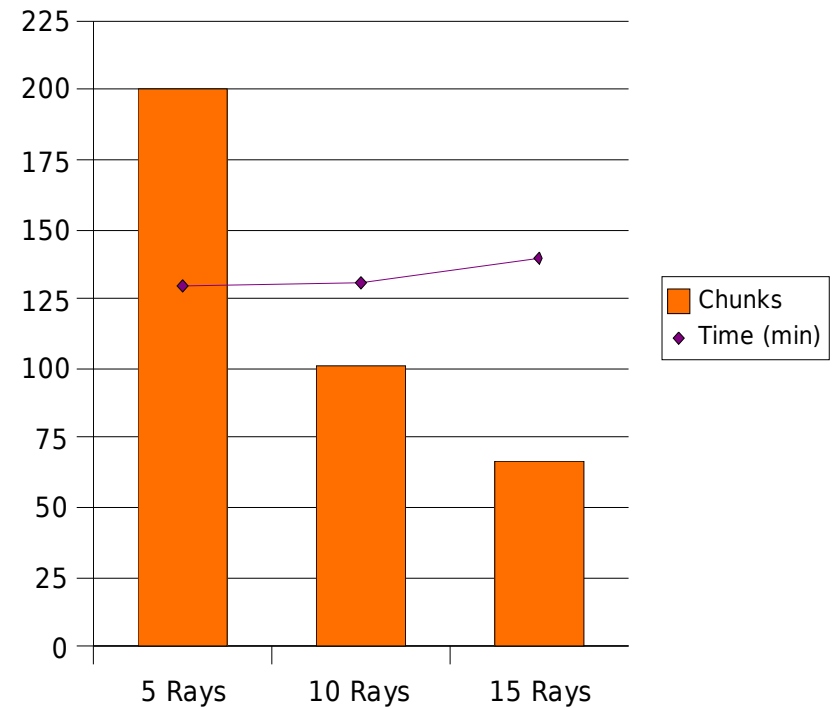
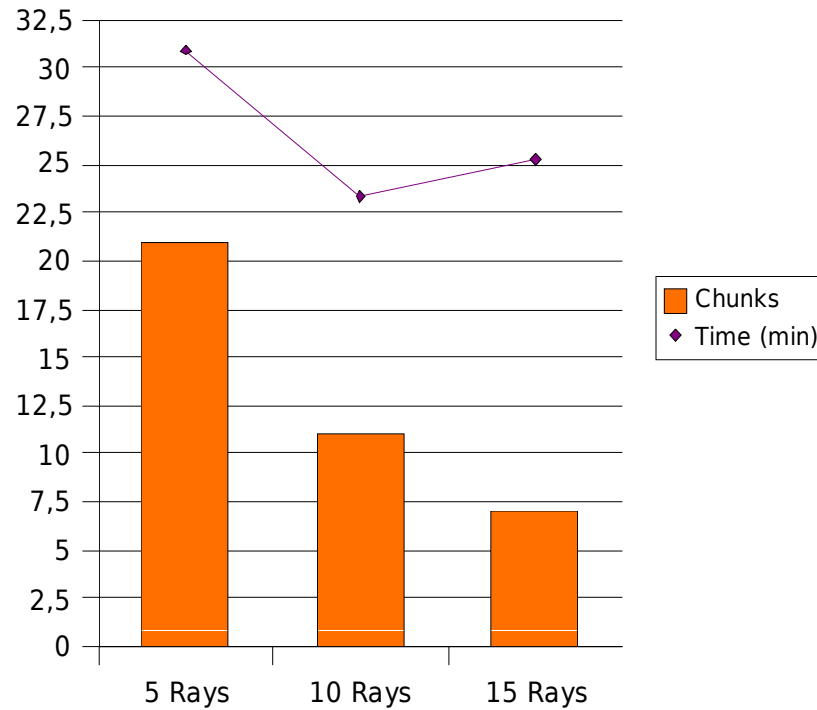
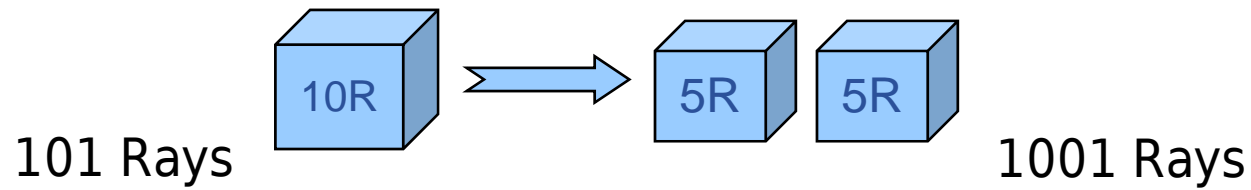
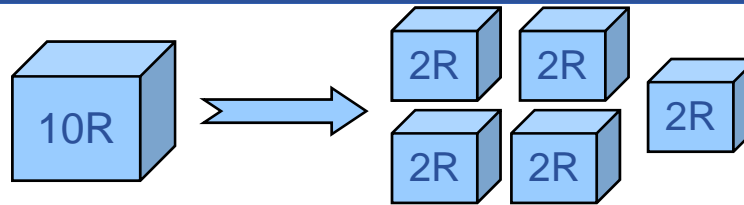


Information Society
and Media

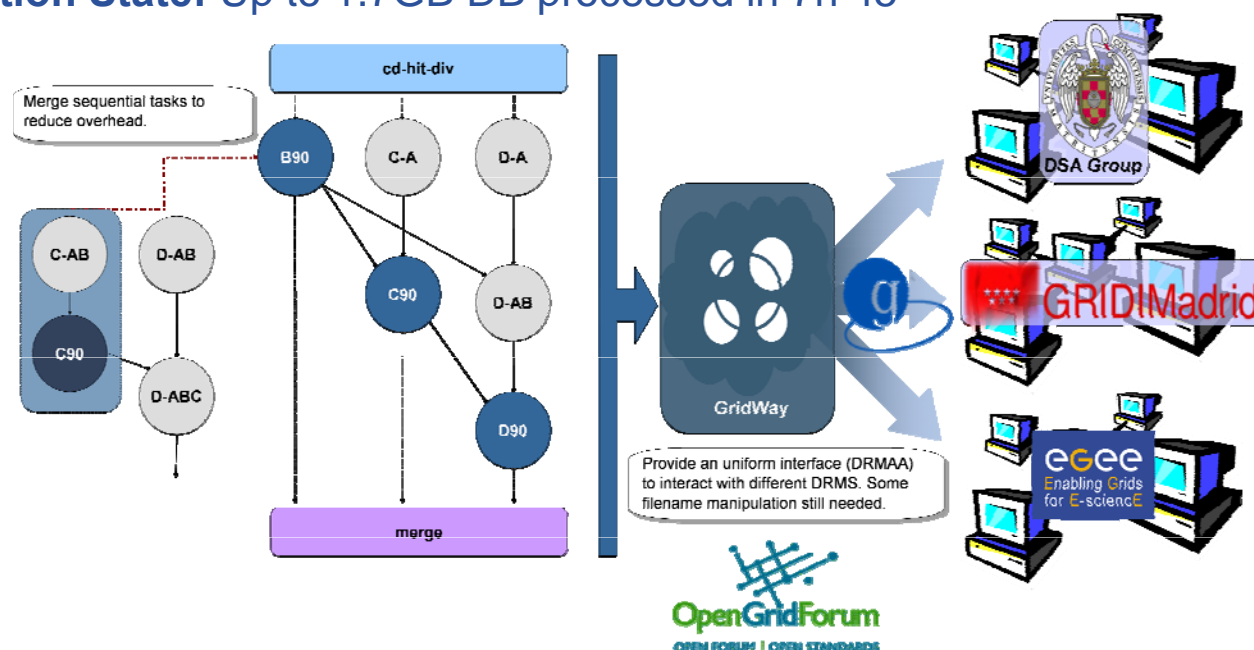


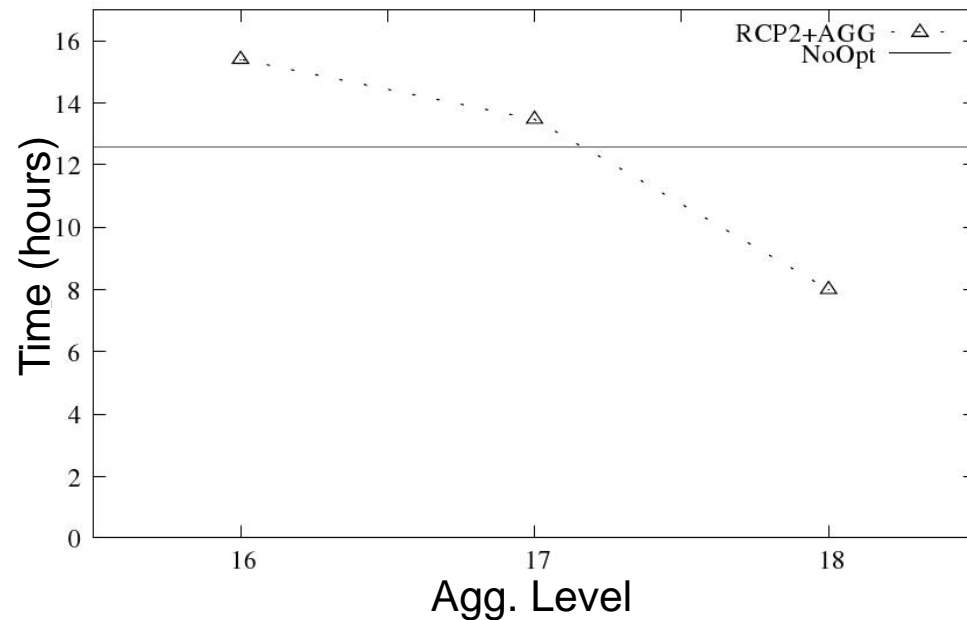
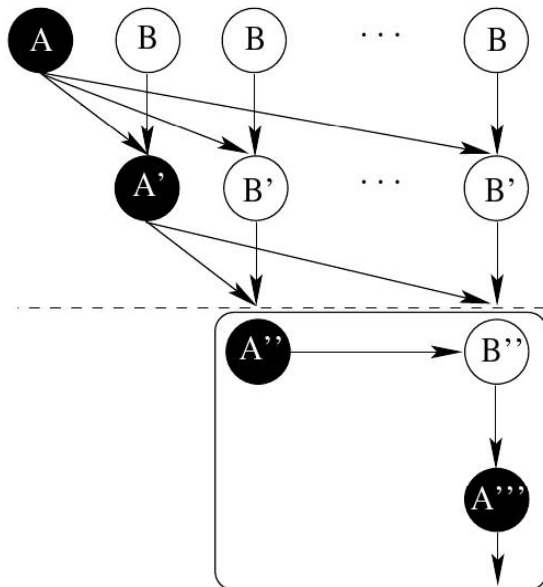
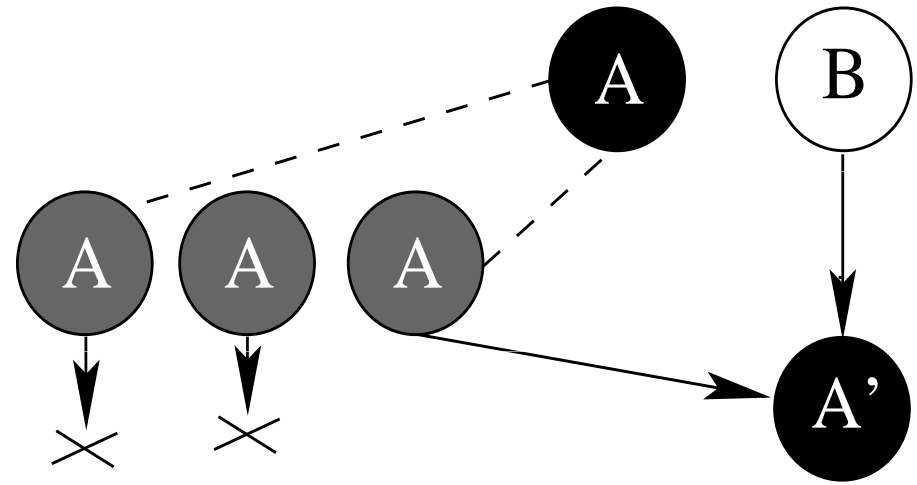
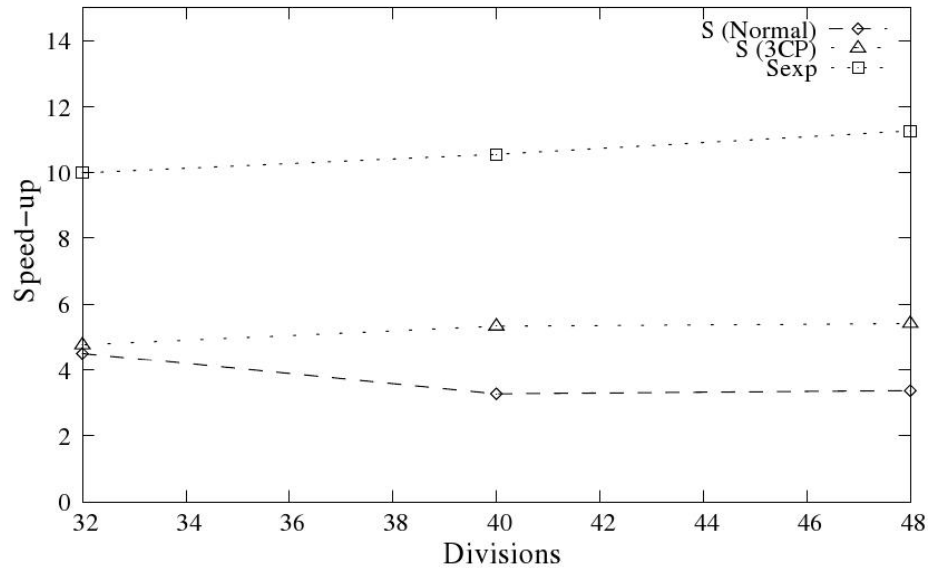
- **Massive Ray Tracing in Fusion Plasmas**
 - Traces each ray contained in a microwave beam which heats fusion plasma
 - Finds best “shooting” position
 - **Needs to be executed as much as possible in less time**
- **Fusion Physics (Fusion VO) – Proposed by CIEMAT (Spain)**
- Parameter Sweep (with pre and postprocess)
- 1 ray = 1 job = 9’ on a Pentium 4
 - **Upgrade:** Using *chunks* reduces transfer and queue overhead
- CLI developed using OGF DRMAA C API provided by GridWay
 - lcg2.1.69 User Interface C++ API was tested at the beginning (couldn’t handle big collections of jobs)
- **Production Status:** up to 1001 Rays in 2h 10’

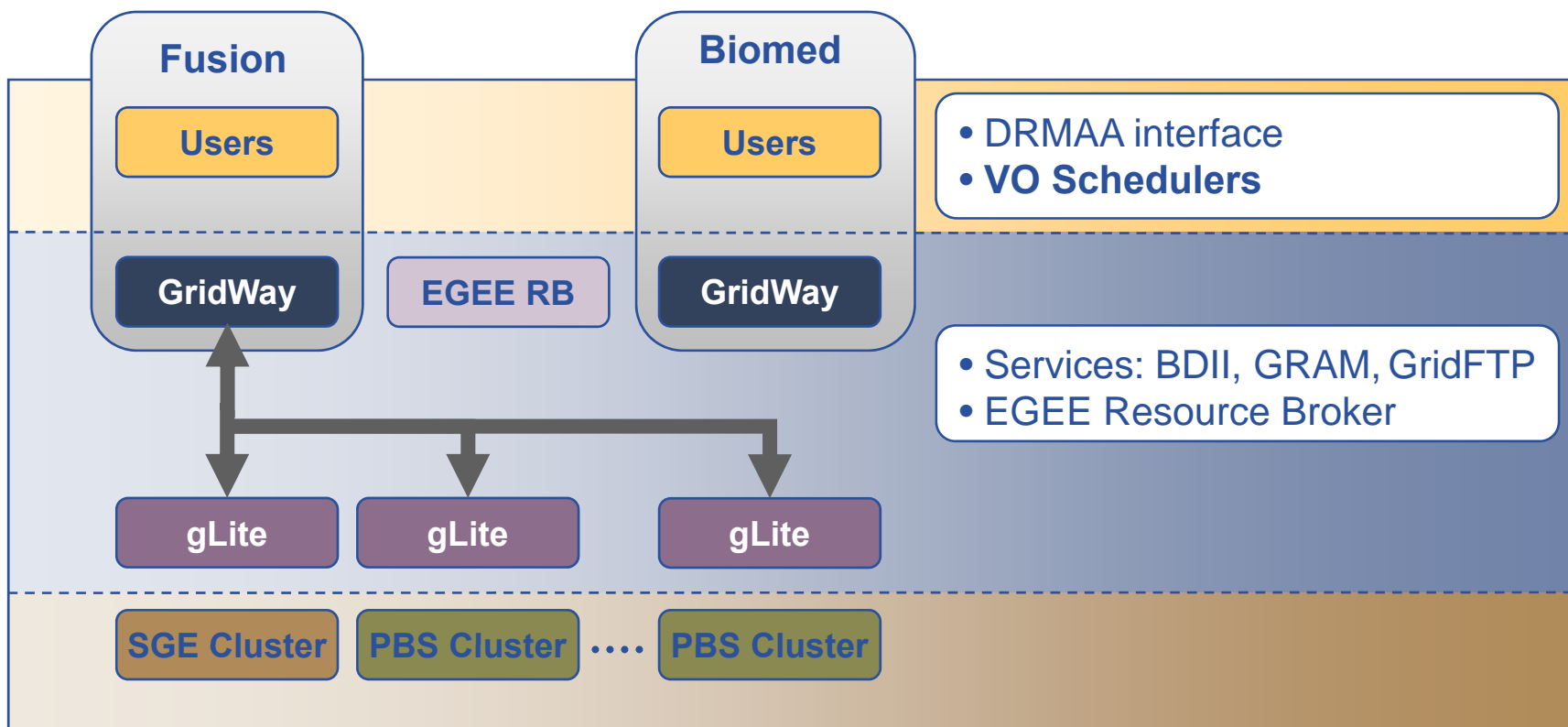
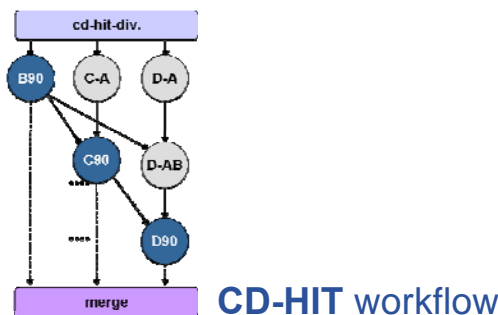
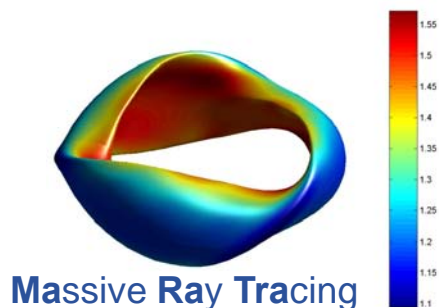


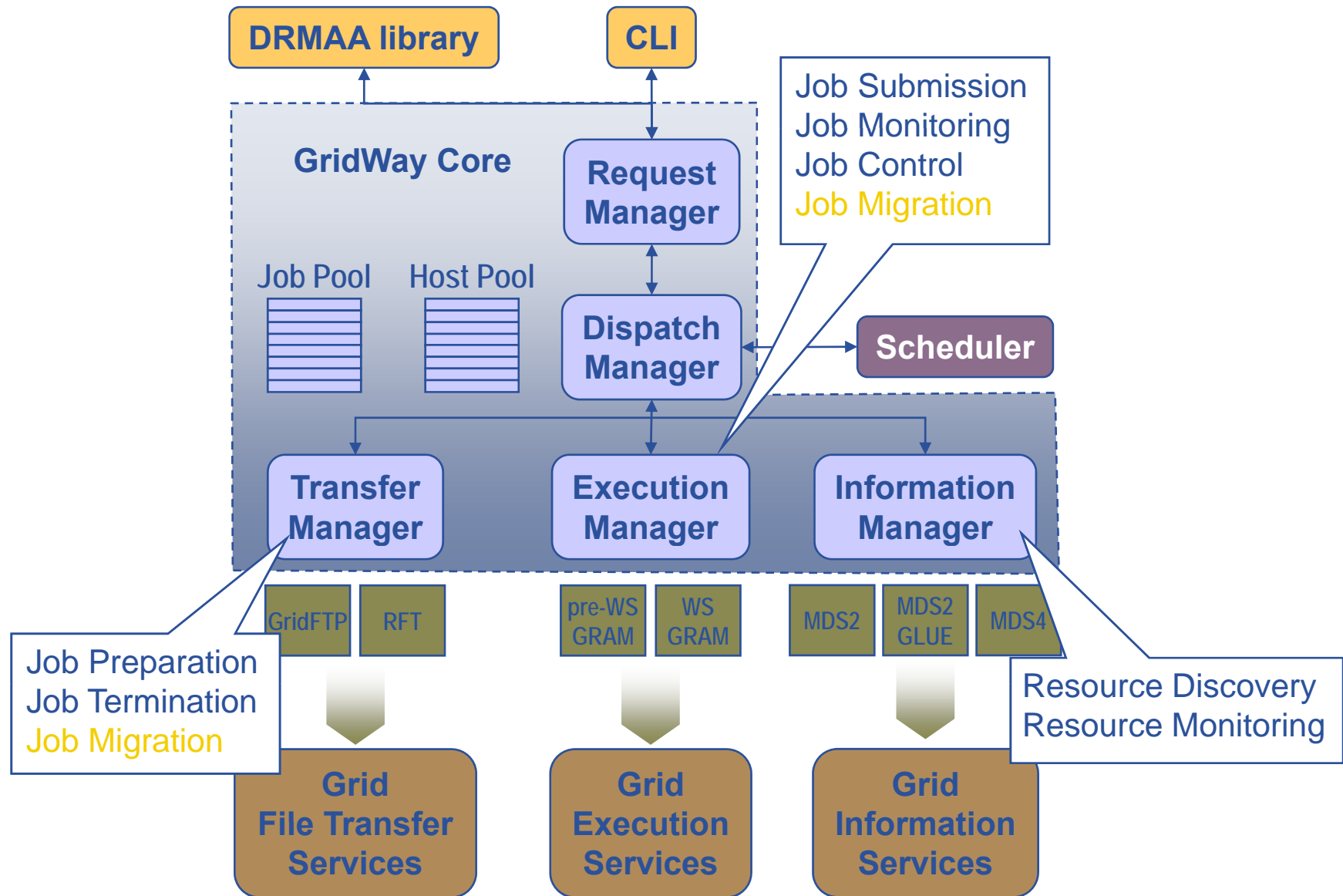


- **CD-HIT: Cluster Database at High Identity with Tolerance**
 - Compares protein DB entries and eliminates redundancies
 - **Example:** UniRef data sets generated by UniProt
 - **Infeasible to be executed on a single machine** (memory requirements and total execution time)
- **Bioinformatics (BioMed VO) – Proposed by CNIO (Spain)**
- Intree Workflow (with pre and postprocess)
 - Grain can be adjusted
 - **Upgrade:** optimization heuristics for efficient execution have been added
- CLI developed using OGF DRMAA C API provided by GridWay
- **Production State:** Up to 1.7GB DB processed in 7h 45'









What is DRMAA?

- Distributed Resource Management Application API
 - <http://www.drmaa.org/>
- Open Grid Forum Standard
- Homogeneous interface to different Distributed Resource Managers (DRM):



- SGE
- Condor
- PBS/Torque
- GridWay
 - C
 - JAVA
 - Perl
 - Ruby
 - Python





Enabling Grids for E-science

The GridWay Metascheduler

Development Process

- **Community** – Open Source Project. Globus Development Philosophy
- **Development Infrastructure** (thanks to Globus Project!)
 - Mailing Lists
 - Bugzilla
 - CVS
- **You are very welcome to contribute:**
 - Reporting Bugs (gridway-user@globus.org)
 - Making feature requests for the next GridWay release (gridway-user@globus.org)
 - Contributing your own developments (bug fixes, new features, documentation)
- **Detailed Roadmap:**
 - **GridWay Campaigns** at bugzilla.mcs.anl.gov/globus/query.cgi
 - www-unix.mcs.anl.gov/~bacon/cgi-bin/big-roadmap.cgi#Gridway

- **Chemistry**
 - **GAMESS and GAUSSIAN** at group QCyCAR (UCLM)
 - **Chemical Reactor Design** in BE14 (“New Product & Process Development”) of BEinGRID project
- **Life Sciences**
 - **Dynamic BLAST** at UABGrid
 - **Multi-Resolution Docking** at Centro de Investigaciones Biológicas (CSIC)
 - **Verification and Optimization of Radiotherapy Treatments** in BE25 (“BEinEIMRT: Business Experiment in Enhanced IMRT planning using Grid services on demand with SLAs”) of BEinGRID project
- **Astronomy**
 - **XMM-Newton Data Processing** at ESAC (ESA)
 - **Use of Special Hardware for N-Body Dynamics in Galactic Nuclei** at AstroGrid-D
- **Climate**
 - **Global Atmospheric Climate Simulations** at EELA
 - **Weather Research and Forecasting** at LAGrid
- **Geology**
 - **Petroleum Reservoir Simulations** at TIGRE

Teşekkür ederim!
Questions?