

## Services for advanced workflow programming on gLite with WS-PGRADE portal

Gergely Sipos  
MTA SZTAKI

[sipos@sztaki.hu](mailto:sipos@sztaki.hu)

[www.lpds.sztaki.hu/gasuc](http://www.lpds.sztaki.hu/gasuc)



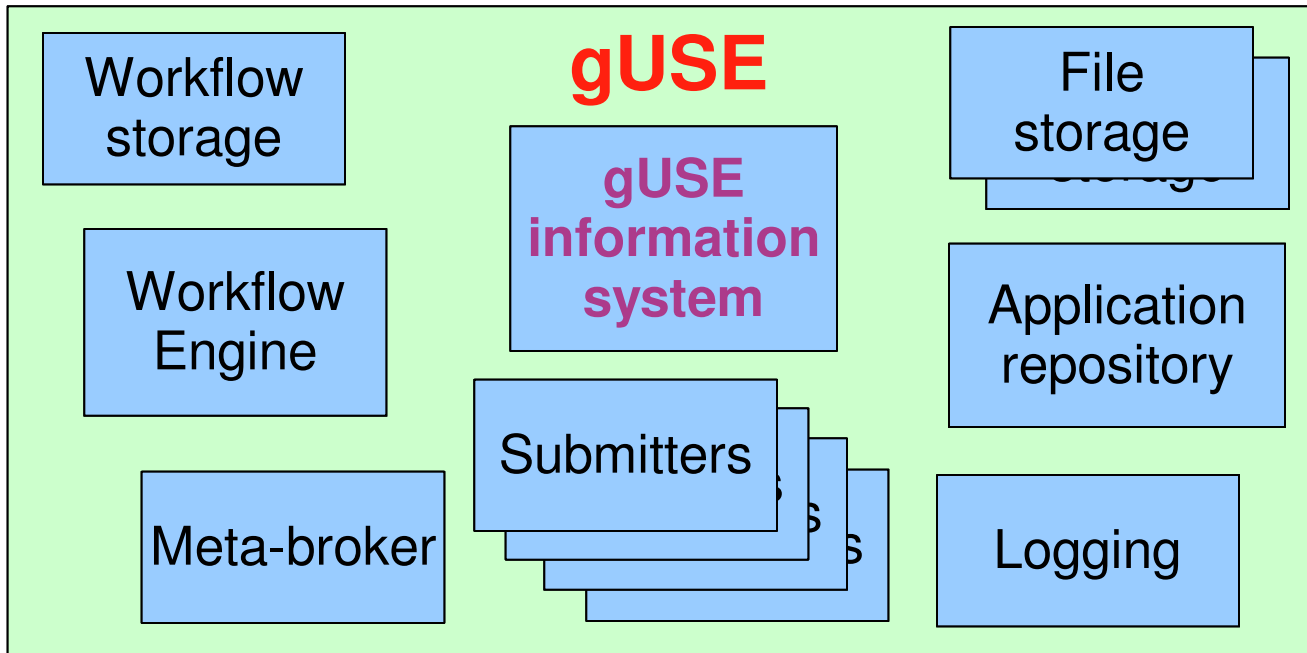
- **WSPGRADE in context**
  - P-GRADE Portal, GEMICA P-GRADE Portal, WS-PGRADE
- **WS-PGRADE features**
  - Scalable architecture
  - Seamless access to various types of resources
- **Advanced data-flows**
- **Comfort features**
  - Separated views for different roles, repository
- **Users and applications**
- **Next steps and conclusions**

- **P-GRADE portal**
  - Creating (basic) workflows and parameter sweeps for clusters, service grids, desktop grids
  - [www.portal.p-grade.hu](http://www.portal.p-grade.hu)
  - *Over 1700 Sourceforge downloads*
- **P-GRADE/GEMLCA portal (University of Westminster)**
  - To wrap legacy applications into Grid Services
  - To add legacy code services to P-GRADE Portal workflows
  - <http://www.cpc.wmin.ac.uk/cpcsite/gemlca>
- **WS-PGRADE**
  - Creating complex workflow and parameter sweeps for clusters, service grids, desktop grids, databases
  - Creating complex applications using embedded workflows, legacy codes and community components from workflow repository
  - [www.wspgrade.hu](http://www.wspgrade.hu)

- **To overcome (most of) the limitations of P-GRADE portal:**
  - Provide better modularity → you can replace any service
  - Improve scalability → to millions of jobs
  - Enable advanced dataflow patterns
  - Interface with wider range of resources
  - Separate Application Developer view from Application User view
- **WS-PGRADE (Web Services Parallel Grid Runtime and Developer Environment)**  
**and**  
**gUSE (Grid User Support Environment) architecture**

Graphical User Interface: **WS-PGRADE**

*Gridsphere portlets*

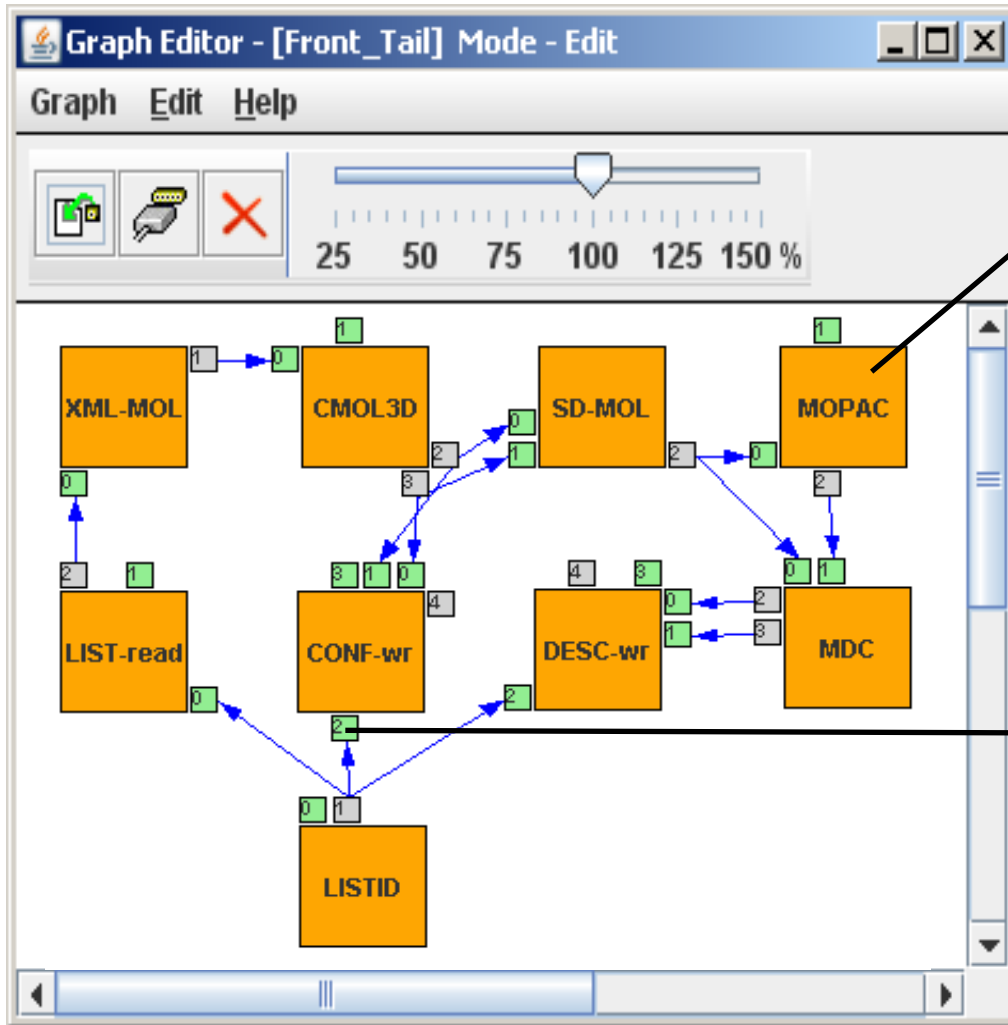


*Autonomous Services: high level middleware service layer*

Local resources, service grid VOs,  
Desktop Grid resources, Web services, Databases

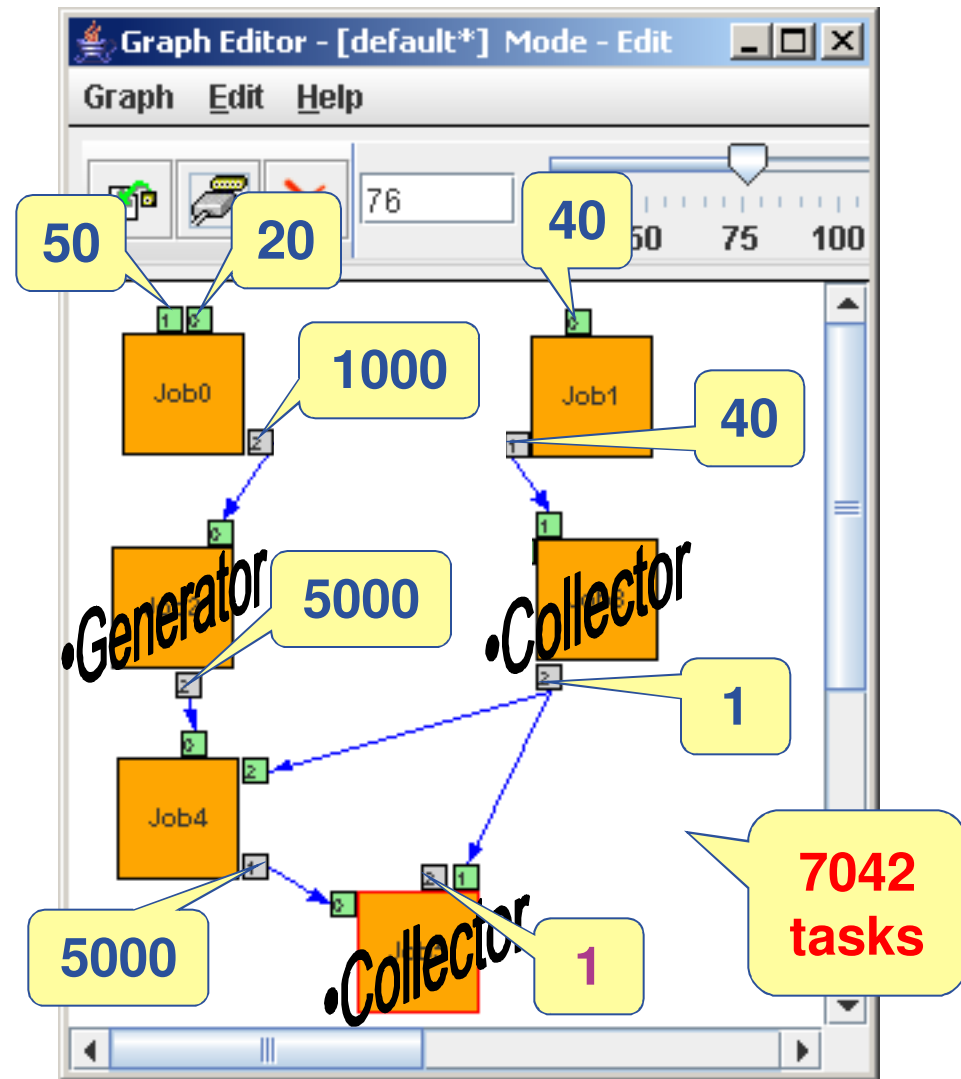
*Resources: middleware service layer*

- **Define workflow structure**
- **Configure workflow**
  - Define the meaning of computational tasks
- **Run a test**
  - Use local resources, Web services, Databases
- **Scale workflow for large simulations**
  - Use batch systems, use cluster grids, use desktop grids
- **Fix some parameters, leave some open**
  - Application specific science gateway for end users



- Job to run on dedicated machine
  - Job to run in a gLite VO
  - Job to run in a Globus 2 VO
  - Job to run in a Globus 4 VO
  - Task to run in a BOINC Grid
  - Web service invocation
  - Database operation (R / W)
- 
- File from the client host
  - File from a GridFTP site
  - File from an LFC catalog
  - Input string for a task or service
  - Result of a Database query

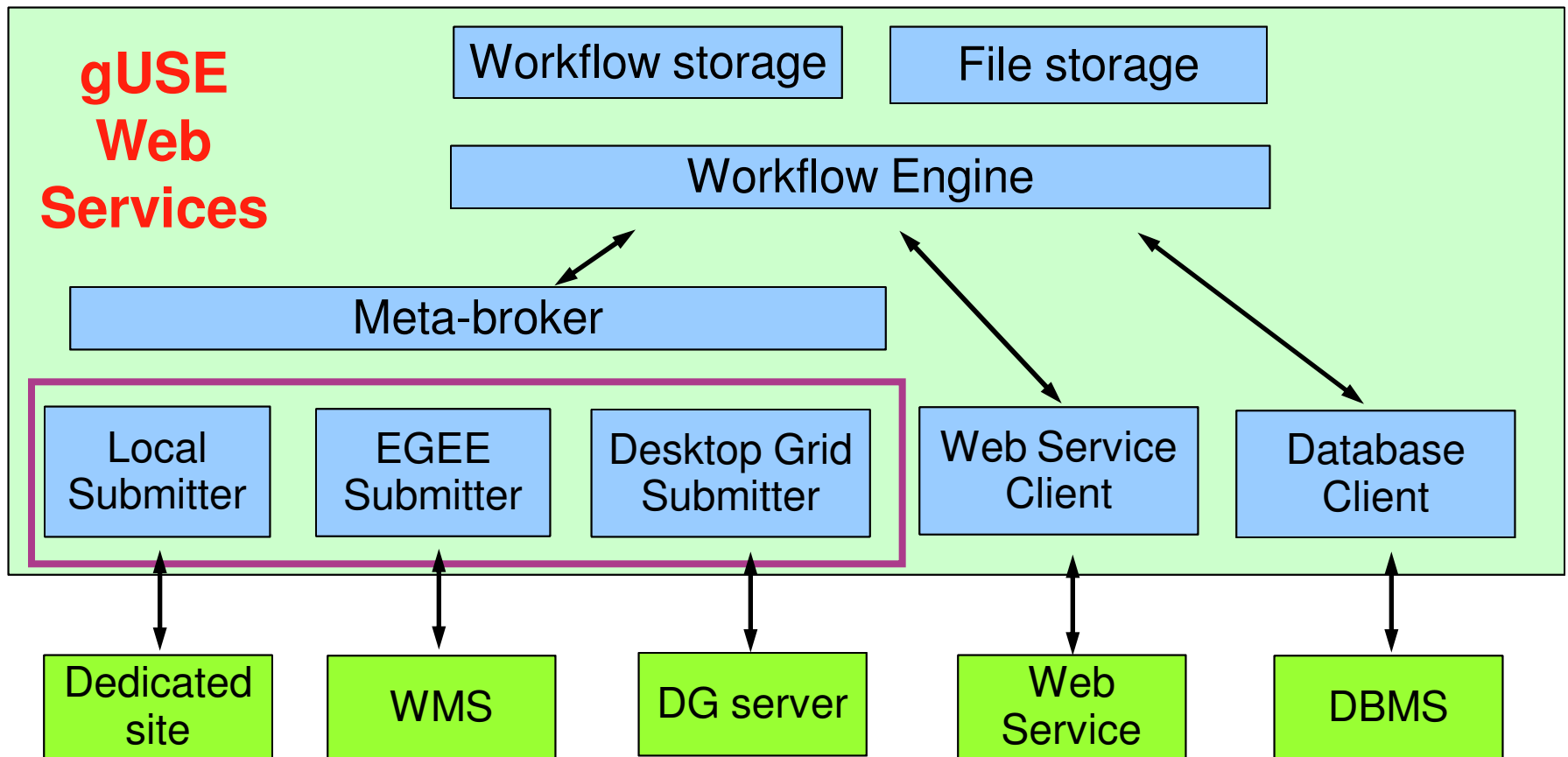
- **Separate application logic from data**
- **Cross & dot product data-pairing**
  - Concept from Taverna
  - All-to-all vs. one-to-one pairing of data items
- **Generator components:** to produce *many output files from 1 input file*
- **Collector components:** to produce *1 output file from many input files*
- **Any component** can be generator or collector
- **Conditional execution based on equality of data**
- Nesting, cycle, recursion

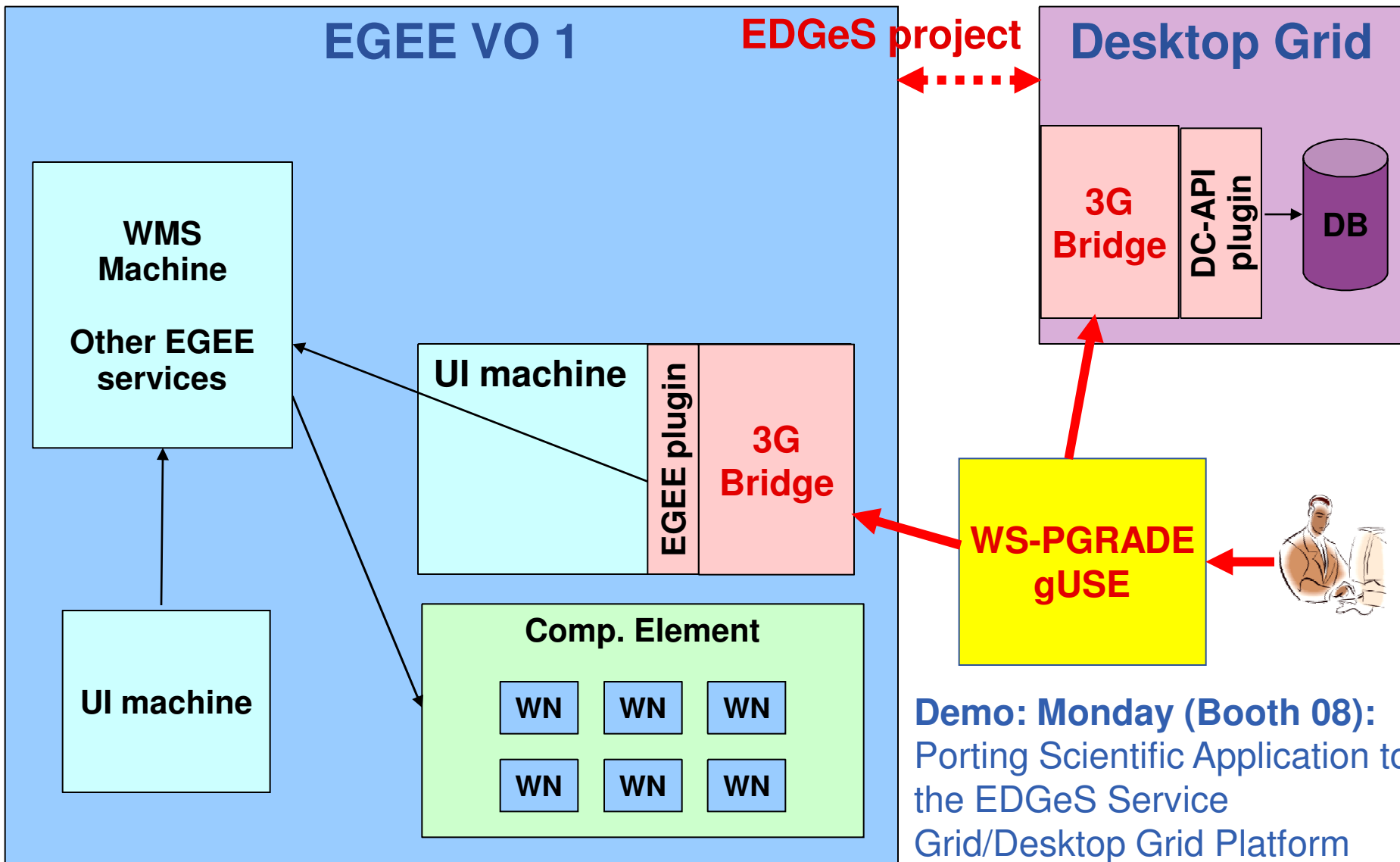




User action, external event or time triggering

**WS-PGRADE**





- **Users can be grid application developers or end-users.**
- **Application developers design sophisticated dataflow graphs**
  - embedding into any depth, recursive invocations, conditional structures, generators and collectors at any position
  - Publish applications in the **repository** at certain stages of work
    - Applications
    - Projects
    - Concrete
    - Templates
    - Graphs
- **End-users see gUSE as a science gateway**
  - List of ready to use applications in repository
  - Import and execute application without knowledge of programming, dataflow or grid

PGrade Grid portal - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://n47.hpcc.sztaki.hu:9080/gridsphere/gridsphere?cid=117&gs\_mode=view&gs\_...

Workflow name: MathOpCascade

Note: 2008-6-25 command line parameter control

Workflow Graph: Front\_Tail --|-- Optional selection of a

Workflow Template:

Job's name: Front

optional note: Front has two free inputs

Job Executable: Job Inputs and Outputs

Job execution model:

- Interpretation of job as Workflow
- Interpretation of job as Service
- Interpretation of job as Binary

Type:

- gemlca
- YY1
- LCG-2
- Local
- YYY
- Hungrid\_BDII
- GLITE
- gUSE
- GT-4
- CancerGrid\_DB
- GT-2
- s

Grid: hungrid

Resource/Broker:

JobManager:

Replicate settings in all Jobs:

Kind of binary: Sequential

MPI Node Number:

Executable code of binary: Recently stored : null

Parameter:

Save.. Quit

PGrade Grid portal - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://n47.hpcc.sztaki.hu:9080/gridsphere/gridsphere?cid=cwork2

gUSE RELEASE 3.1

## WS-PGRADE

portal

Logout  
Welcome, hermann

Welcome | New Features Settings | AMRI database | Workflow | Help | Certificates | Settings | File Management | Information System

Graph Create Concrete Concrete Applications Template Timing Remoting Storage Upload Import

Real Workflows

Submit All Refresh

Names of Workflows Submitted Running Finished Error Actions

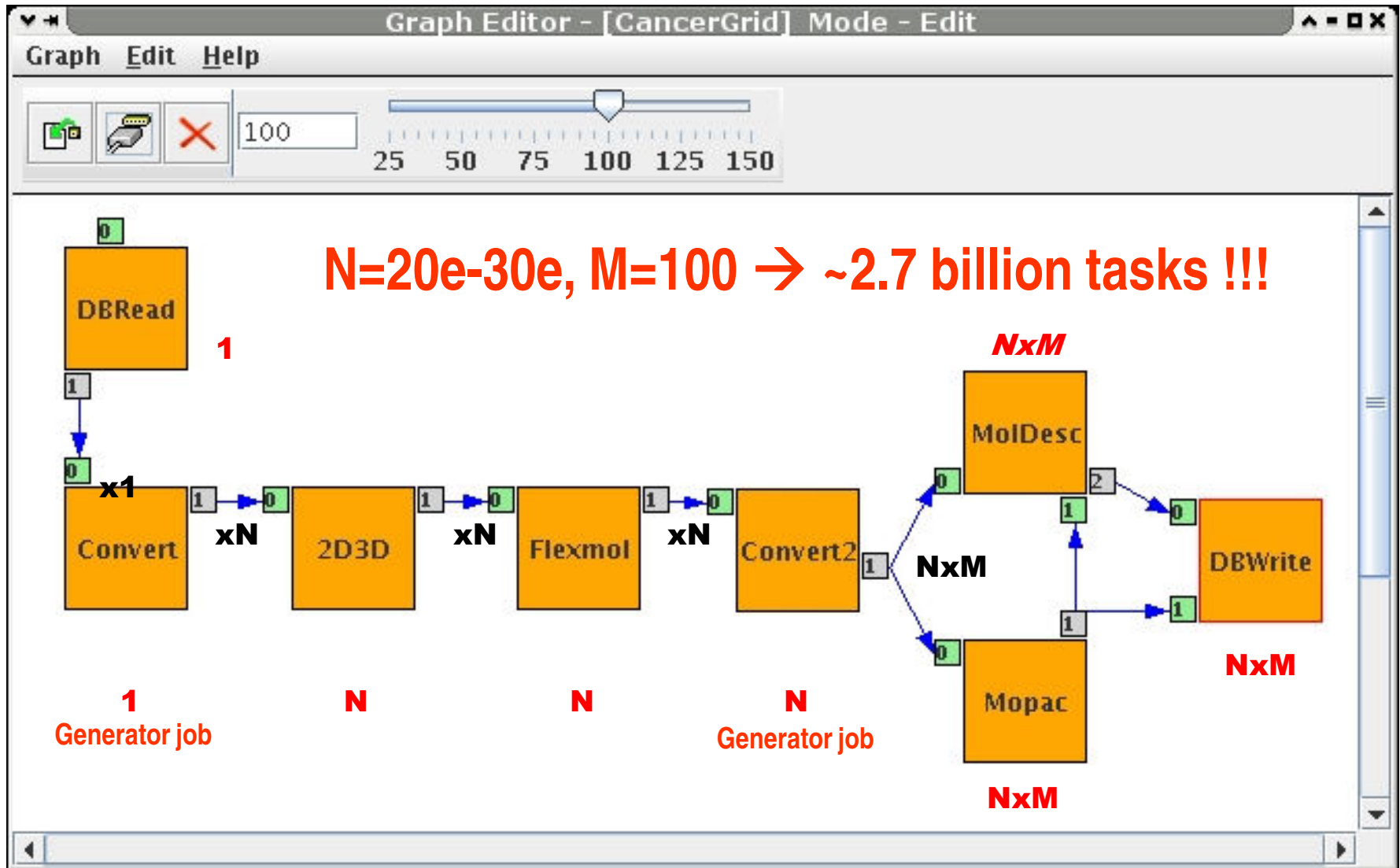
**MathOpCascade** 0 0 0 0 Configure Info Details Submit Delete Export

2008-6-25 command line parameter control

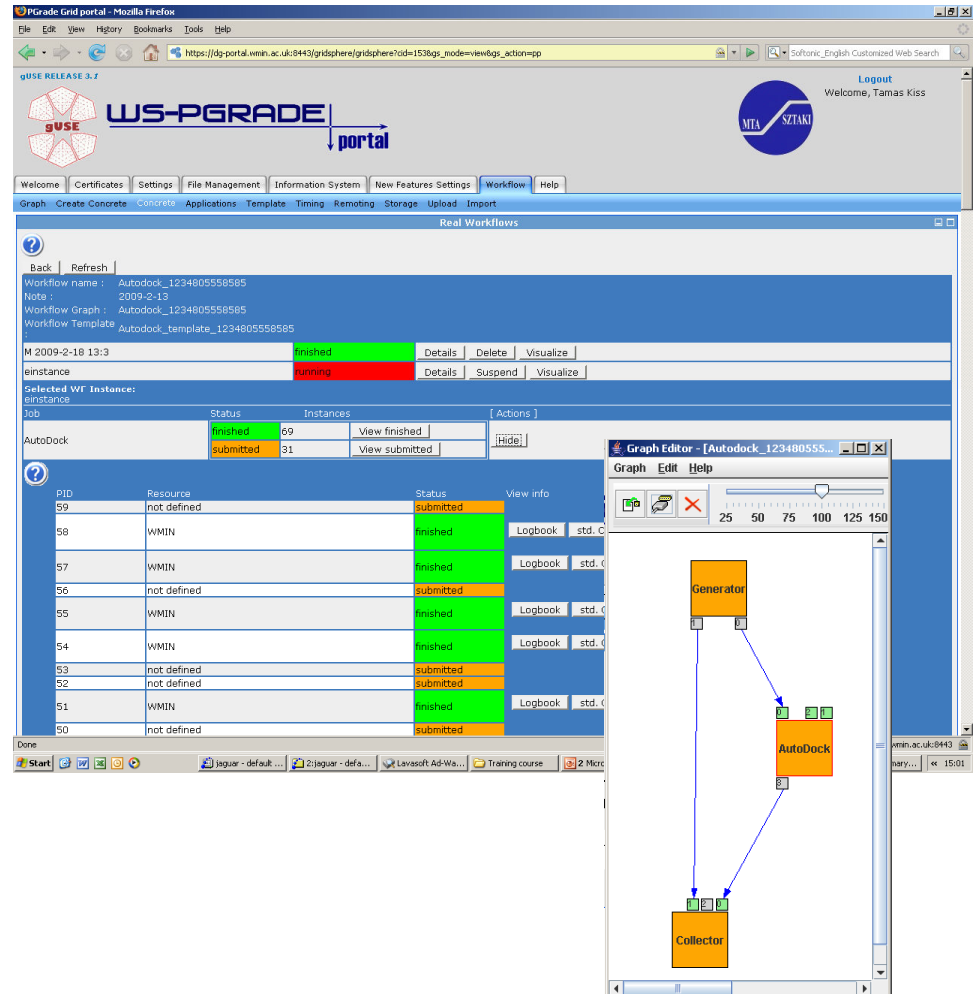
Message :

June 25, 2008

Done



- **CancerGrid project**
  - Predicting various properties of molecules to find anti-cancer leads
  - Creating science gateway for chemists
- **EDGeS project (Enabling Desktop Grids for e-Science)**
  - Integrating EGEE with BOINC and XtremWeb technologies
  - User interfaces and tools
- **ProSim project**
  - In silico simulation of intermolecular recognition
  - JISC ENGAGE program
  - **Demo on Wednesday**
- **University of Westminster Desktop Grid**
  - Using AutoDock on institutional PCs



The screenshot displays the PGrade Grid portal interface. The top navigation bar includes 'Welcome', 'Certificates', 'Settings', 'File Management', 'Information System', 'New Features Settings', 'Workflow', and 'Help'. Below this, there are tabs for 'Graph', 'Create Concrete', 'Concrete', 'Applications', 'Template', 'Timing', 'Remoteing', 'Storage', 'Upload', and 'Import'. The main content area shows 'Real Workflows' with a table of workflow instances. A 'Graph Editor' window is open, showing a workflow graph with nodes for 'Generator', 'AutoDock', and 'Collector' connected by arrows.

Workflow name	Workflow Graph	Workflow Template
Autodock_123480558585	Autodock_123480558585	Autodock_template_123480558585

Instance	Status	Instances	Actions
M 2009-2-18 13:3	finished		Details   Delete   Visualize
einstance	running		Details   Suspend   Visualize

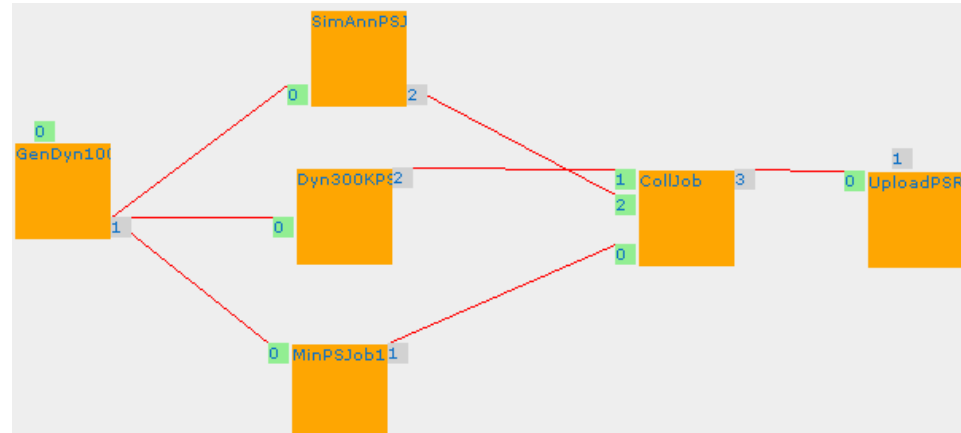
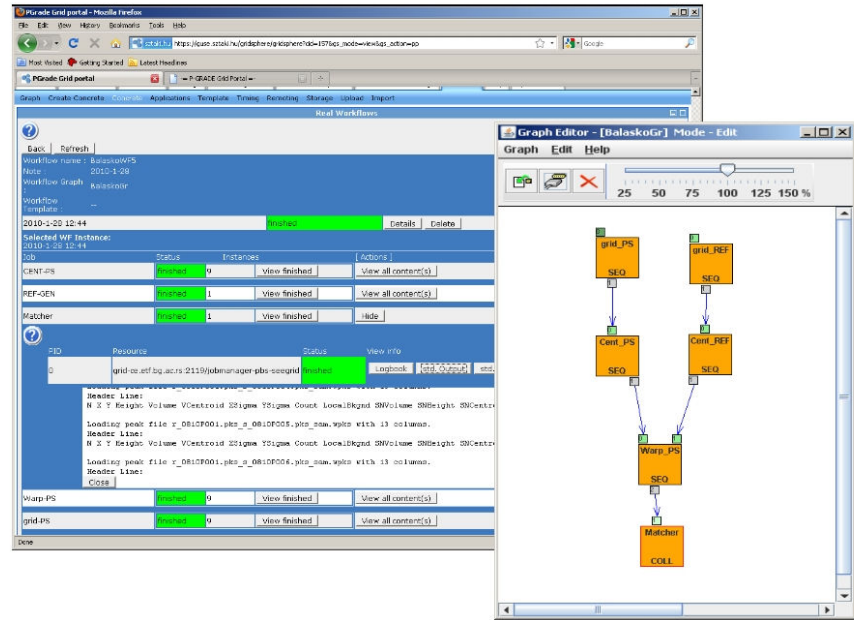
  

Selected WF Instance:	Status	Instances	Actions
Autodock	finished	69	View finished
einstance	submitted	31	View submitted

PID	Resource	Status	View info
59	not defined	submitted	
58	WMIN	finished	Logbook   std. C
57	WMIN	finished	Logbook   std. C
56	not defined	submitted	
55	WMIN	finished	Logbook   std. C
54	WMIN	finished	Logbook   std. C
53	not defined	submitted	
52	not defined	submitted	
51	WMIN	finished	Logbook   std. C
50	not defined	submitted	

- EGEE Application Porting Support Group
  - [www.lpds.sztaki.hu/gasuc](http://www.lpds.sztaki.hu/gasuc)
- Current applications with WS-PGRADE (Presentation on Monday)
  - Proteomics analysis for biomarker discovery
    - <http://www.lpds.sztaki.hu/gasuc/index.php?m=6&r=15>
  - TINKER Conformer Generator
    - <http://www.lpds.sztaki.hu/gasuc/index.php?m=6&r=12>



- **P-GRADE Portal is still supported and open source**
  - Features can serve most grid scenarios
  - Open source project on Sourceforge
  
- **WS-PGRADE provide more advanced services**
  - Implemented on top of scalable, WS based gUSE architecture
  - More expressive dataflow patterns
  - Transparent access to
    - Local resources
    - Service Grids
    - Desktop Grids
    - Databases
    - Web services
  - Application repository
    - Service for collaboration of developers and end-users
  - **Create science gateways with few clicks**





The screenshot shows a Windows Internet Explorer browser window displaying the gUSE website. The browser's address bar shows the URL <http://www.guse.hu/>. The website header features a logo consisting of a red star-like shape with the text 'gUSE' in the center, followed by the title 'grid User Support Environment' in a large, bold, black font. Below the header, the page is divided into two main sections. On the left is a navigation menu with the following items: Home, Architecture, Features and Releases, Portal installations, Documents, For grid users (with sub-items: Client requirements, How to get access), For grid administrators (with sub-item: Install the portal), and Problem reporting (with sub-item: Report Problems). On the right is the main content area, which starts with a 'Home' heading, followed by a sub-heading 'gUSE 3.1 released' and a paragraph of text: 'SZTAKI is glad to announce the release of gUSE, version 3.1.'. At the bottom of the page, there is a copyright notice: '© Copyright 2008, MTA-SZTAKI LPDS, Hungary. All rights reserved.' and an email address: [webmaster@lpds.sztaki.hu](mailto:webmaster@lpds.sztaki.hu). The browser's status bar at the bottom indicates 'Internet' and '100%' zoom.

User manual

Request a user account



**Free event!**  
**Register now!**

## Home

In the last two years P-GRADE portal became popular and many Grids and VOs selected it as their science gateway for their user communities (see <http://portal.p-grade.hu/?m=installations&s=0>). Due to the increased interest and number of user communities, the developers would like to provide stronger support and faster response to the requirements of the user communities. In order to achieve this goal we organize the 1st P-GRADE Portal User Community Workshop. The major goal is to share experience of using P-GRADE Portal among the various user communities and portal developers. The program is discussion-oriented. Every presentation will be followed by 15 minutes discussion time in order to give opportunity for the users to express their P-GRADE experience and for the developers to better understand the problems.

The presentations and discussions will be organized in the following sessions:

1. P-GRADE portal installation, administration and maintenance This session is for system admins who manage P-GRADE portal installations. Here we would like to discuss problems they encountered and improvements they recommend.
2. Applications developed by P-GRADE portal This session is for application developers who develop applications using the portal. Here we would like to discuss what they like and what they do not like or miss in the portal.
3. End-user experience with P-GRADE portal Representatives of end-user communities are welcome to share their experience with the portal.
4. Further development of P-GRADE portal Portal developers will present future plans on the further development of the portal. User communities are welcome to give presentation on their additional needs to improve the portal.

A half-day WS-PGRADE portal tutorial will be also part of the program. WS-PGRADE is the second generation P-GRADE portal that will be presented and demonstrated during the tutorial. Two application-specific portals developed during the workshop will also be presented: the generic purpose WS-PGRADE portal will also be presented: ProSim portal for protein folding simulation and CancerGrid portal for drug design.

The whole event is free of charge. We would like to share experience with the P-GRADE portal user community in order to improve the portal for their sake and not to make profit from this event.

**WS-PGRADE**  
**tutorial**

- Home
- Program
- Tutorial
- Program committee
- Organizing committee
- Call for presentations
- Important dates
- Contact

# Thank you

[www.wspgrade.hu](http://www.wspgrade.hu)

