P-GRADE Portal: a workflow-oriented generic application development portal

Peter Kacsuk

MTA SZTAKI, Hungary

Univ. of Westminster, UK

User concerns of Grid systems

- How to cope with the variety of Grid systems? **(How to move from LCG-2 to gLite?)**
- How to develop/create new Grid applications?
  - To use workflows (DAG)
  - To use MPI codes
- How to execute Grid applications in a fault-tolerant way?
- How to observe (and debug) the application execution in the Grid?
- How to tackle performance issues?
- How to port legacy applications
  - to Grid systems
  - between Grid systems?
- How to interoperate among Grids, how to execute Grid applications over several Grids in a transparent way? **(see Earth Science slides)**

- **The goal of this talk is to show how a portal, like P-GRADE, can solve all these problems at a high abstraction level**
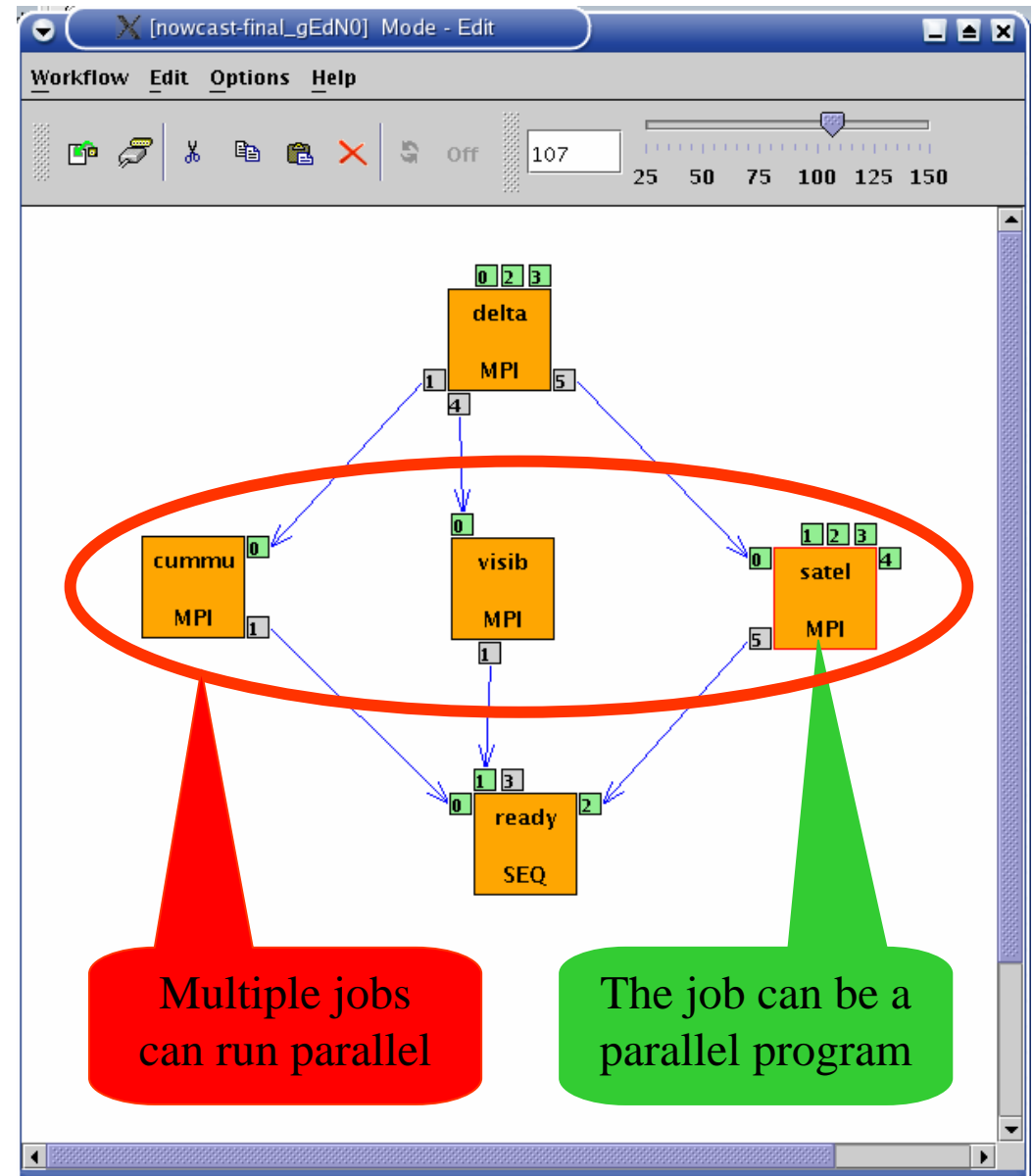
Properties of the P-GRADE Portal

- **General purpose**, **graphical**, **workflow-oriented Grid portal**
- Supports the development and execution of workflow-based Grid applications
- Components of the workflows can be
  - Sequential jobs
  - Parallel jobs (MPI, PVM)
  - Legacy code (GEMLCA) services
- Enables the exploitation of two levels of parallelism
- Solves the interoperability of Grids at the workflow level

Two levels of parallelism by the P-GRADE workflow

- Semantics of the workflow enables two levels of parallelism:
  - Parallel execution inside a workflow node
  - Parallel execution among workflow nodes

- The P-GRADE Portal workflow concept enables the efficient parallelization of complex problems

Multiple jobs can run parallel

The job can be a parallel program

Solving Grid interoperability by P-GRADE Portal
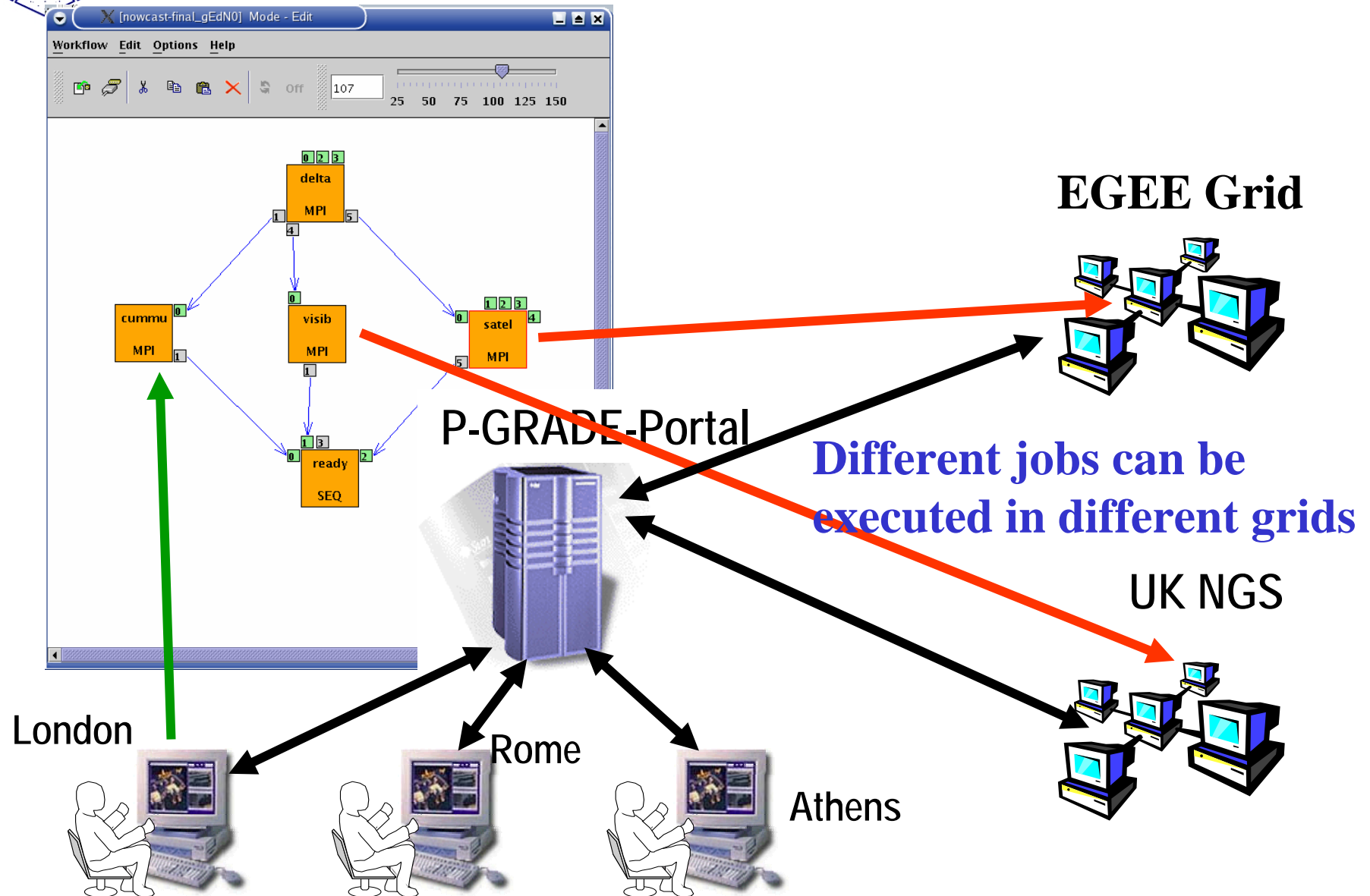
Properties of P-GRADE Grid Portal

- Grid services supported by the portal:
  - **MyProxy** – proxy credential management
  - **GridFTP** – file transfer
  - **GT2/GT3/GT4 GRAM** – job execution
  - **Mercury** – parallel job monitoring
  - **PROVE** – workflow & job execution visualization
  - **BDII and MDS** – obtain information about resources
  - **LCG-2 broker** – resource selection
  - **GEMLCA** – invoke legacy codes
- **GridSphere** based
  - Easy to expand with new portlets
  - Easy to tailor to end-user needs
- Support for **grid interoperability** at workflow level

P-GRADE portal in a nutshell

Proxy management

Definition of Grid resources

Workflow creation

Job mapping to Grid resources

Workflow management and execution visualization

User concerns of Grid systems

- How to cope with the variety of Grid systems?
- How to develop/create new Grid applications?
- How to execute Grid applications in a fault-tolerant way?
- How to observe the application execution in the Grid?
- How to tackle performance issues?
- How to port legacy applications
  - to Grid systems
  - between Grid systems?
- How to execute Grid applications over several Grids in a transparent way?

Workflow creation at the client machine

Certificate server

CERTIFICATE
(download)

WORKFLOW MANAGER
(submit)

CERTIFICATE
(upload)

**WF EDITOR
(save / upload)**

Portal server

Remote Clusters to be controlled

**WF EDITOR
(open)**

Workflow
(result)

WORKFLOW MANAGER
(output)

9
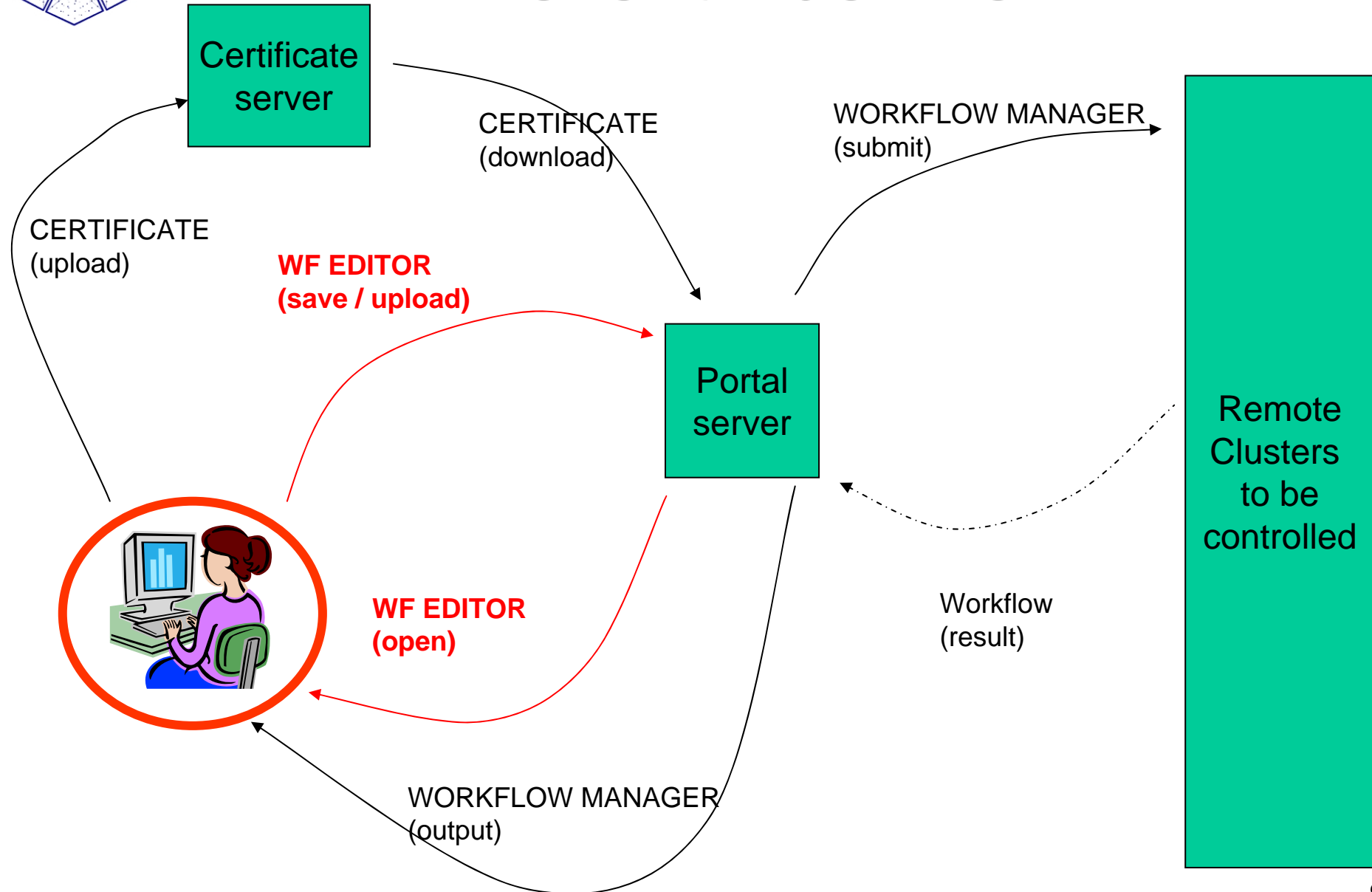
Workflow Editor: Grid aware workflow mapping

Visualization of monitoring system information

Non-Grid aware (abstract) workflow mapping

Can be used if the selected Grid (eg. LCG-2) has a broker:

1. Describe the requirements of the job

2. Select a Grid with broker

3. The workflow manager will contact the broker to find the best resource for your job

Workflow Editor extension with JDL

Select an LCG-2 based Grid!

Don't select any resource!

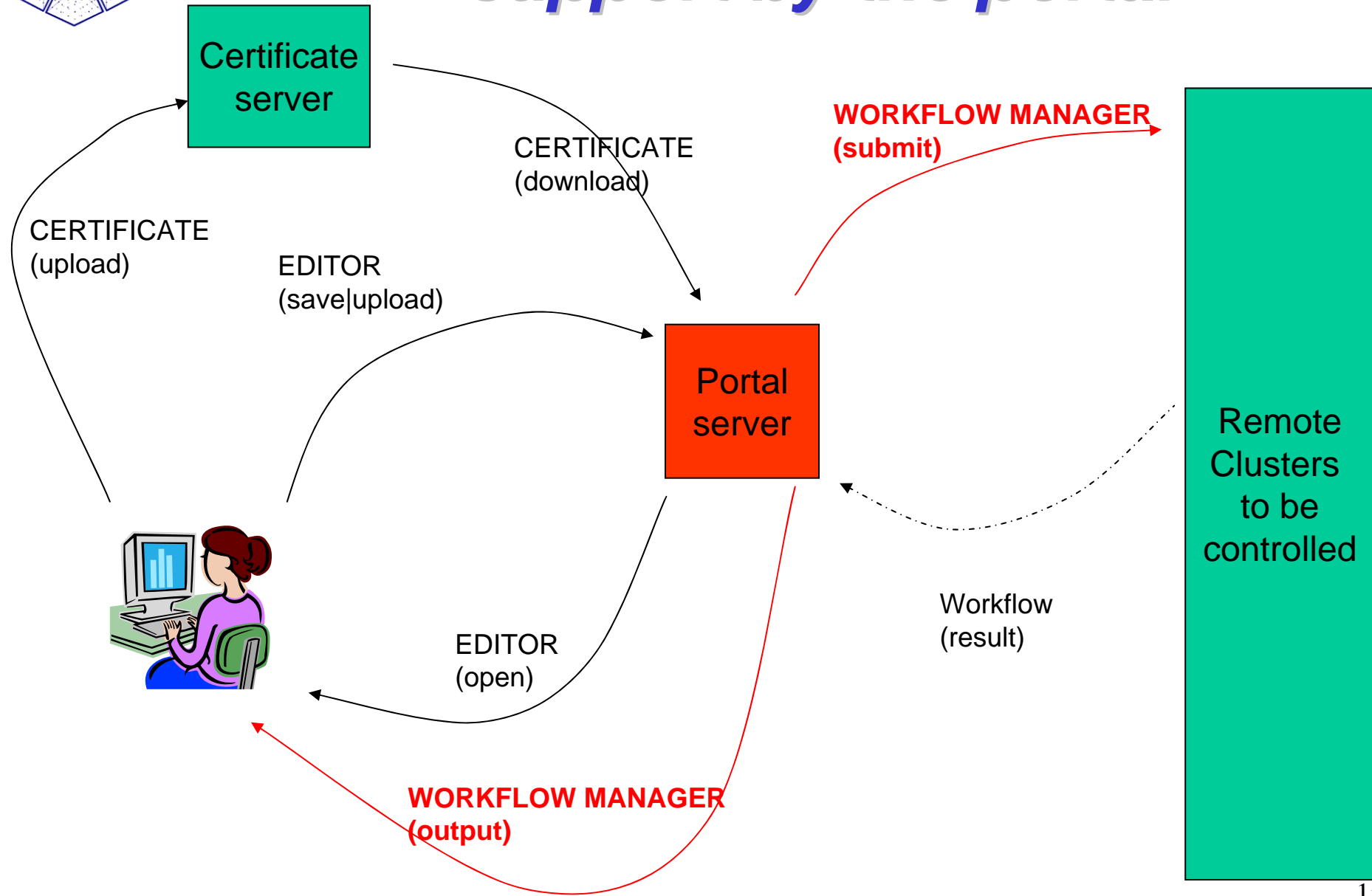**Define job requirements with the "Job Description Language"!**

User concerns of Grid systems

- How to cope with the variety of Grid systems?
- How to develop new Grid applications?
- How to execute Grid applications in a fault-tolerant way?
- How to observe the application execution in the Grid?
- How to tackle performance issues?
- How to port legacy applications
  - to Grid systems
  - between Grid systems?
- How to execute Grid applications over several Grids in a transparent way?

Workflow execution support by the portal

Certificate server

CERTIFICATE (download)

WORKFLOW MANAGER (submit)

CERTIFICATE (upload)

EDITOR (save|upload)

Portal server

Remote Clusters to be controlled

Workflow (result)

EDITOR (open)

WORKFLOW MANAGER (output)

Workflow Manager Portlet

- **Displays the list of jobs and their status**

- **The current status of the jobs are represented by colors**

- **Provides access to their logs and outputs, and visualizes them**

White/Red/Green color means the job is initialised/running/finished

16

Rescuing a failed workflow

Rescuing a failed workflow

User concerns of Grid systems

- How to cope with the variety of Grid systems?
- How to develop new Grid applications?
- How to execute Grid applications in a fault-tolerant way?

➡ How to observe the application execution in the Grid?

➡ How to tackle performance issues?

- How to port legacy applications
  - to Grid systems
  - between Grid systems?
- How to execute Grid applications over several Grids in a transparent way?

On-Line Monitoring

Job execution visualization

User concerns of Grid systems

- How to cope with the variety of Grid systems?
- How to develop new Grid applications?
- How to execute Grid applications in a fault-tolerant way?
- How to observe the application execution in the Grid?
- How to tackle performance issues?
- How to port legacy applications
  - to Grid systems
  - between Grid systems?
- How to execute Grid applications over several Grids in a transparent way?

GEMLCA legacy code architecture objectives

- **To deploy legacy code applications as Grid services without reengineering the original code and minimal user effort** — **GEMLCA**

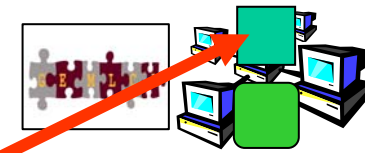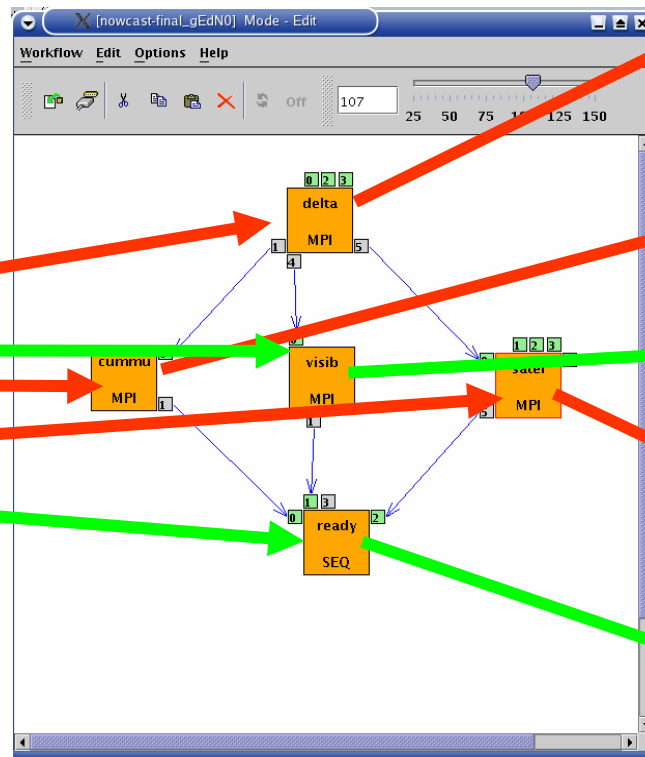- **To create Grid workflows where components can also be legacy code applications**

- **To make these functions available from a Grid Portal**
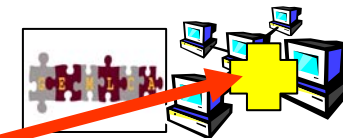
GEMLCA & P-GRADE Portal Integration

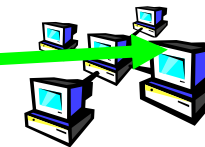# Combining legacy and non-legacy components

Combine legacy codes
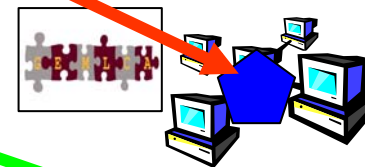**with new codes**
inside the same workflow!

Code invocation

Code invocation

Job submission

Code invocation

Job submission

Use successfully executed new code as legacy code

Legacy code registration page

GEMLCA client in a nutshell: Traffic Simulation Workflow

Workflow creation

Definition of legacy code service properties

Monitoring and execution visualization

References of P-GRADE portal

- **Official portal of**
  - SEE-GRID infrastructure

  - HUNGRID infrastructure

- **P-GRADE portal is available as service for:**
  - VOCE
  - UK National Grid Service
- **Under preparation for**
  - Croatian Grid
  - EGRID (Italy)
  - GridIreland

How to access P-GRADE Portal?

- If you are interested in using P-GRADE Portal:
  - Take a look at **www.lpds.sztaki.hu/pgportal**
  - **If you are a user,** get an account for one of its production installations:
    - **HUNGrid** portal – SZTAKI
    - **VOCE** portal – SZTAKI
    - **SEEGRID** portal – SZTAKI
    - **UK NGS** portal – University of Westminster
  - **If you are the administrator** of a VO or Grid, ask SZTAKI to install P-GRADE Portal for your VO or Grid:
    - **EGRID VO** portal – ICTP
    - **Croatian Grid** portal – Boskovic Institute

New, planned features

- **Parameter study support at**
  - Job level
  - Workflow level
- **New types of parallelism** (collaboration with Johan Montagnat):
  - Pipeline parallelism (e.g. Planck VO and biomed community need it)
  - multi-thread parallelism
- **Collaborative portal** (collaboration with Univ. of Reading)
- **Automatic testing of Grid services and resources from the portal** (collaboration with Univ. of Westminster)
  - Intelligent brokering
  - **Intelligent error messages**
  - **Automatic handling of error situations**
- **New application-oriented portlets** (collaboration with E-Grid and Croatian Grid)
- **Interactive workflow development and debugging support**

Final conclusions

- **Users should access any Grids transparently by Grid portals**
- **Every Grid should be accessed via different portals in order to provide a choice for the users**
  - **In the case of EGEE:**
    - **Genius Portal**
    - **P-GRADE Portal**
  - **In the case of UK NGS:**
    - **Daresbury Portal**
    - **P-GRADE Portal**
- **P-GRADE Portal provides the following principles:**
  - **Learn once, use everywhere**
  - **Develop once, execute anywhere**
- We are ready to collaborate with any team and support any EGEE application by the portal, and extend it with the special needs of the application

www.lpds.sztaki.hu/pgportal/

Grid-Enabling Legacy Applications and Supporting End Users Workshop

within the framework of the
15th IEEE International Symposium on
High Performance Distributed
Computing
HPDC'15
Paris, France
June 19-23, 2006

- **IMPORTANT DATES**

- Paper Abstract Submission: February 06, 2006

- Full Papers Submission: February 13, 2006

- Notice of Acceptance: March 27, 2006

- Final Manuscript Due: April 24, 2006

- Workshop: June 20, 2006