Exposing Application as Grid Services

P-GRADE portal

Grid Execution Management for Legacy Code Applications

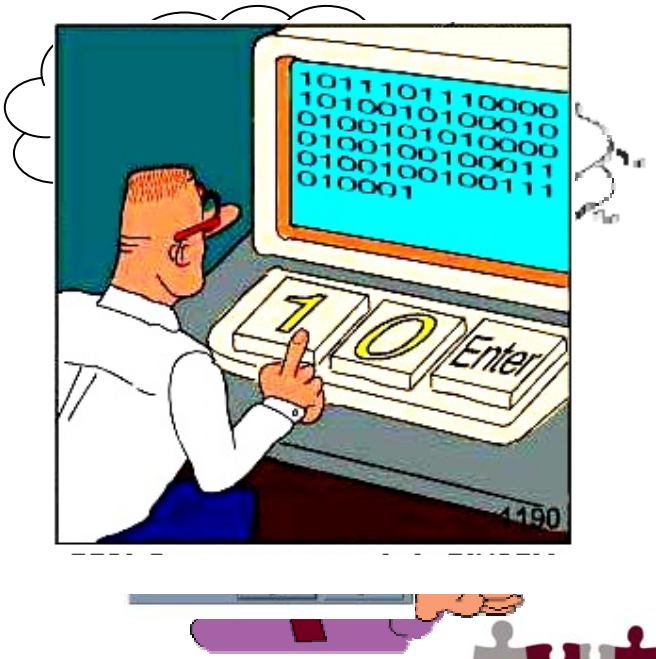
Porto, Portugal, 23 January 2007

Legacy Applications

- Code from the past, maintained because it works
- Often supports business critical functions
- Not Grid enabled

What to do with legacy codes when utilising the Grid?

- Bin them and implement Grid enabled applications
- Reengineer them
- Port them onto the Grid with minimum user effort



GEMLCA – Grid Execution Management for Legacy Code Architecture

Objectives

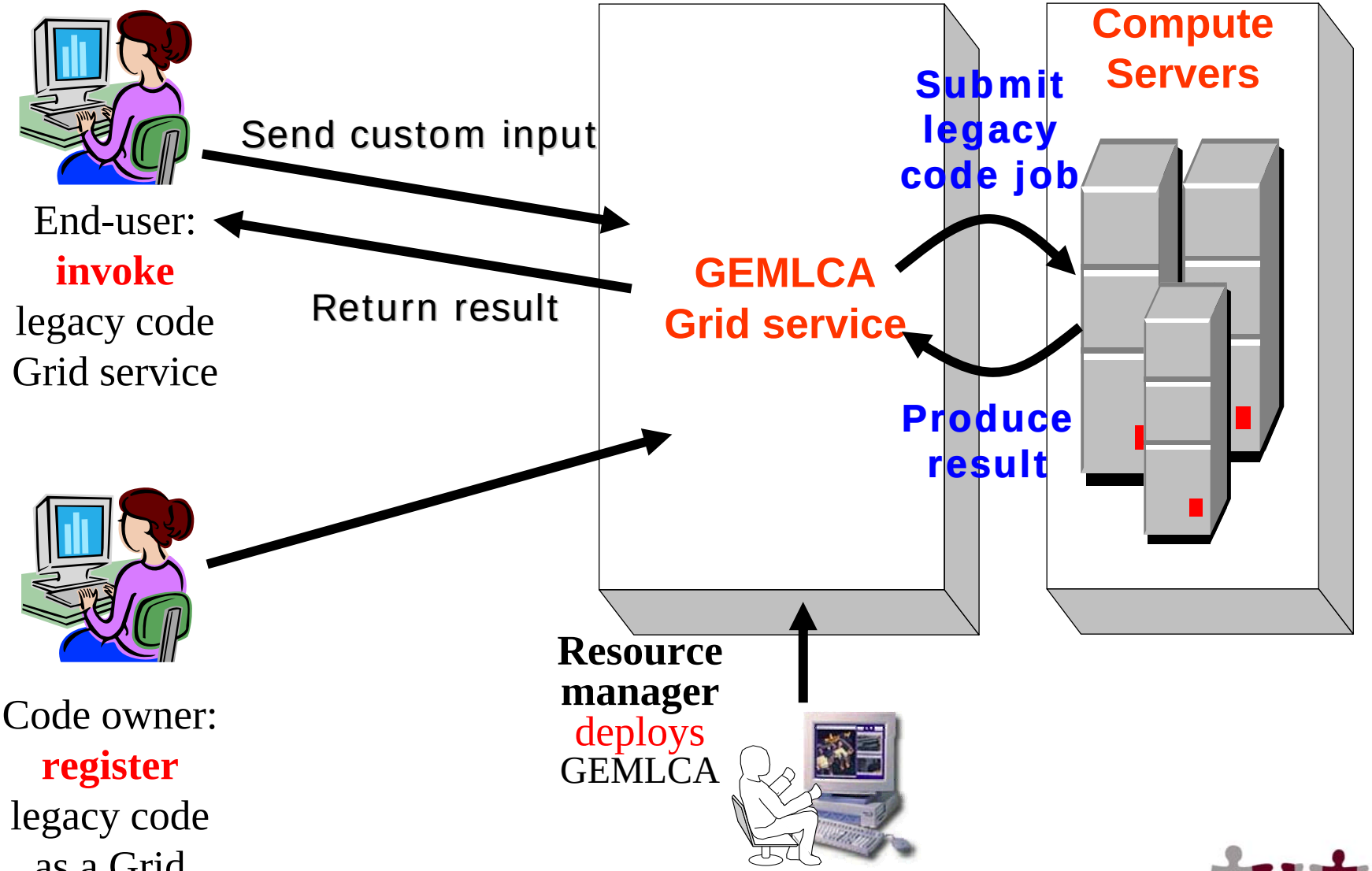
- To deploy legacy code applications as Grid services without reengineering the original code and minimal user effort

GEMLCA

- To create complex Grid workflows where components are legacy code applications
- To make these functions available from a Grid Portal

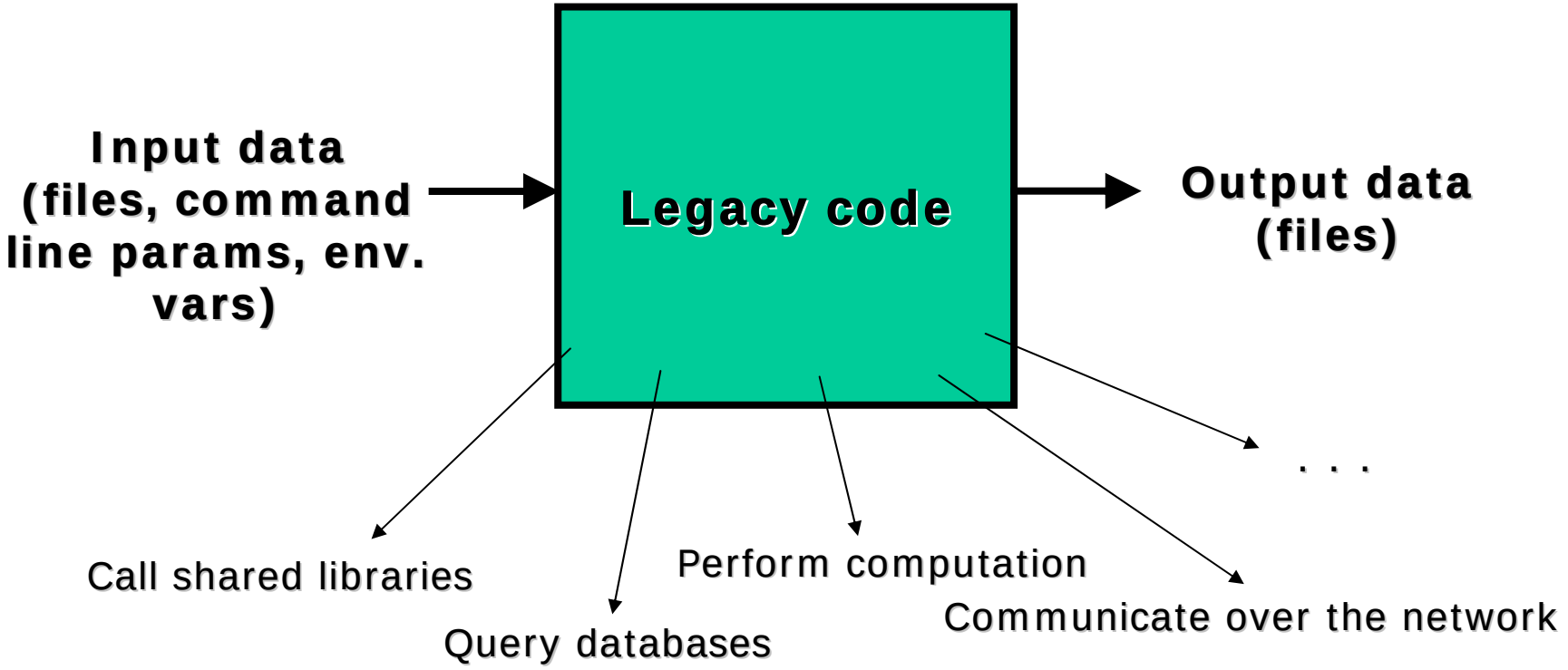
GEMLCA PGPortal Integration

GEMLCA Concept



The GEMLCA-view of a legacy code

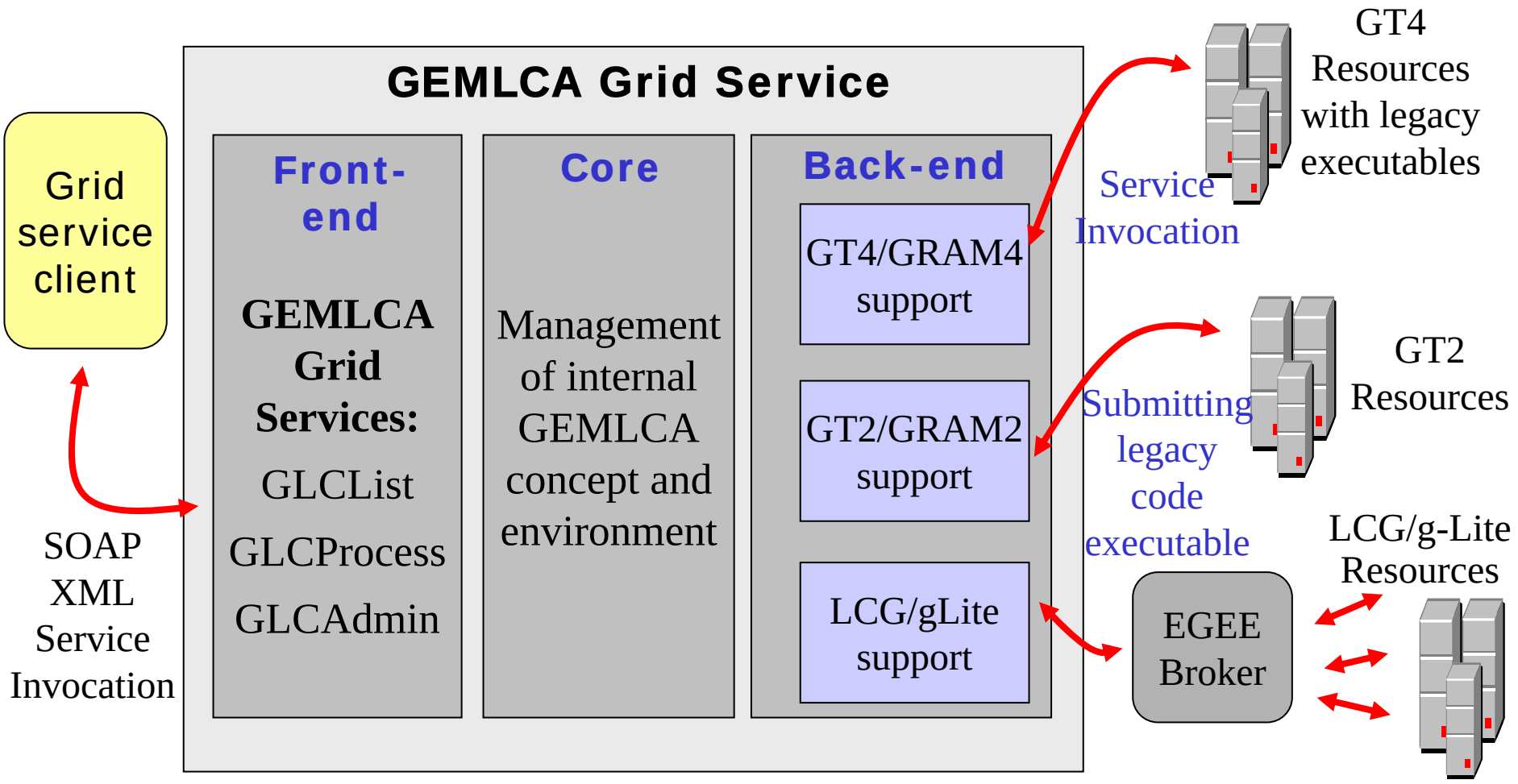
- Any code that correspond to the following model can be exposed as Grid service by GEMLCA:



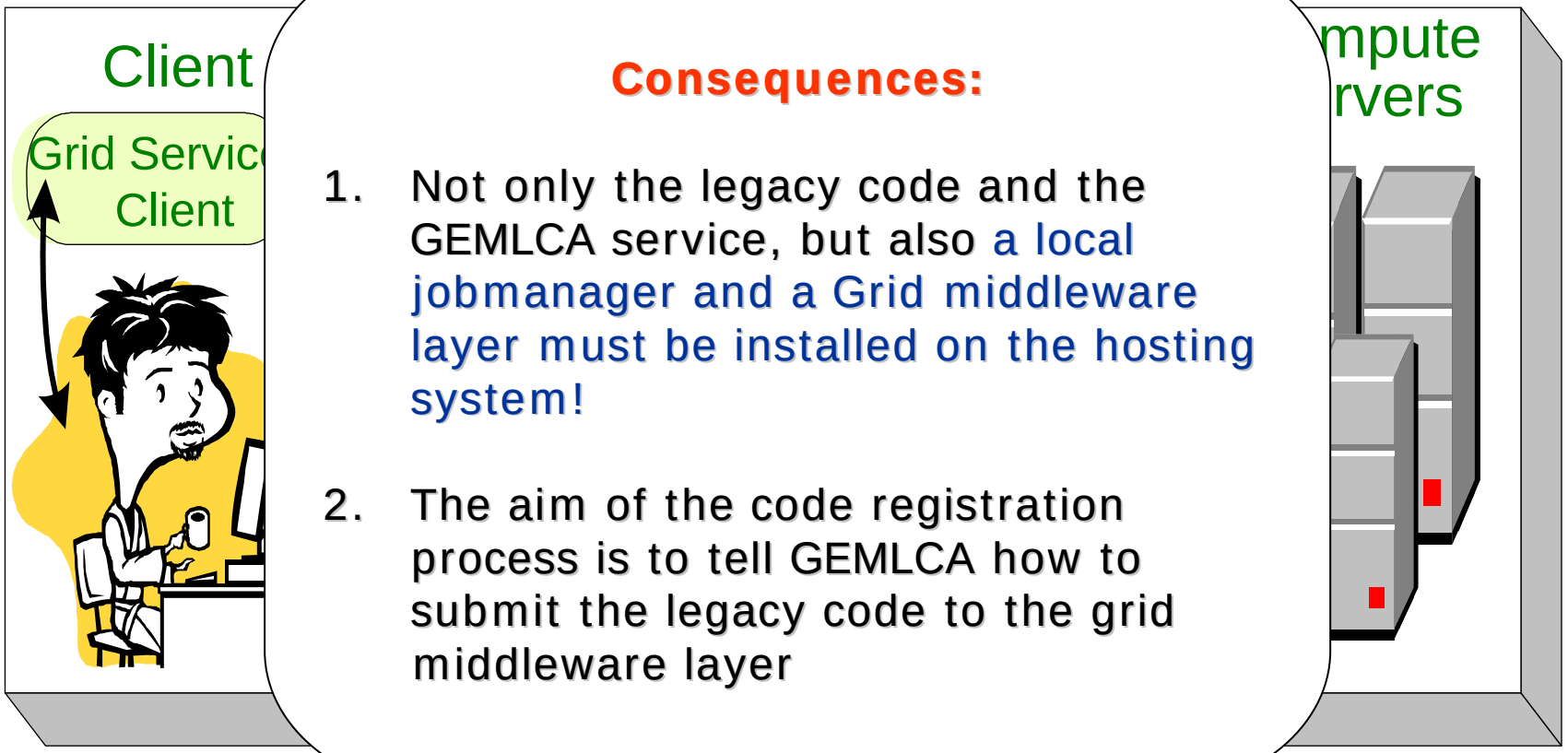
Implementing the concept

- The GEMLCA service can be implemented with any grid/service-oriented technology E.g:
  - Globus (3 or) 4 → **currently available implementations**
  - Jini
  - Web services
  - …
- GEMLCA service could invoke legacy codes in many different ways. Current implementation:
  - **Submit the legacy code as a batch job to a local job manager (e.g. Condor or PBS) through a Grid middleware layer (e.g. GT2/3/4, LCG/g-Lite)**

Implementing the GEMLCA concept



What's behind the GEMLCA service…



Consequences:

1. Not only the legacy code and the GEMLCA service, but also a local jobmanager and a Grid middleware layer must be installed on the hosting system!

2. The aim of the code registration process is to tell GEMLCA how to submit the legacy code to the grid middleware layer

What's the point?

- **Heterogeneous codes can be hidden behind the same interface** (the programming interface of the GEMLCA service)
  - Different programs can be invoked in the same way
- **Extend non grid-aware programs with security infrastructure** (access enabled through a Grid service)
  - Share your codes with your colleagues or partner institutes
  - Expose business logic to your employees or customers
- **Create and browse repositories of legacy applications**
- **Build customized GEMLCA clients** (such as the GEMLCA P-GRADE Portal)
  - Compose complex processes by connecting multiple legacy code grid services together

The GEMLCA P-GRADE Portal

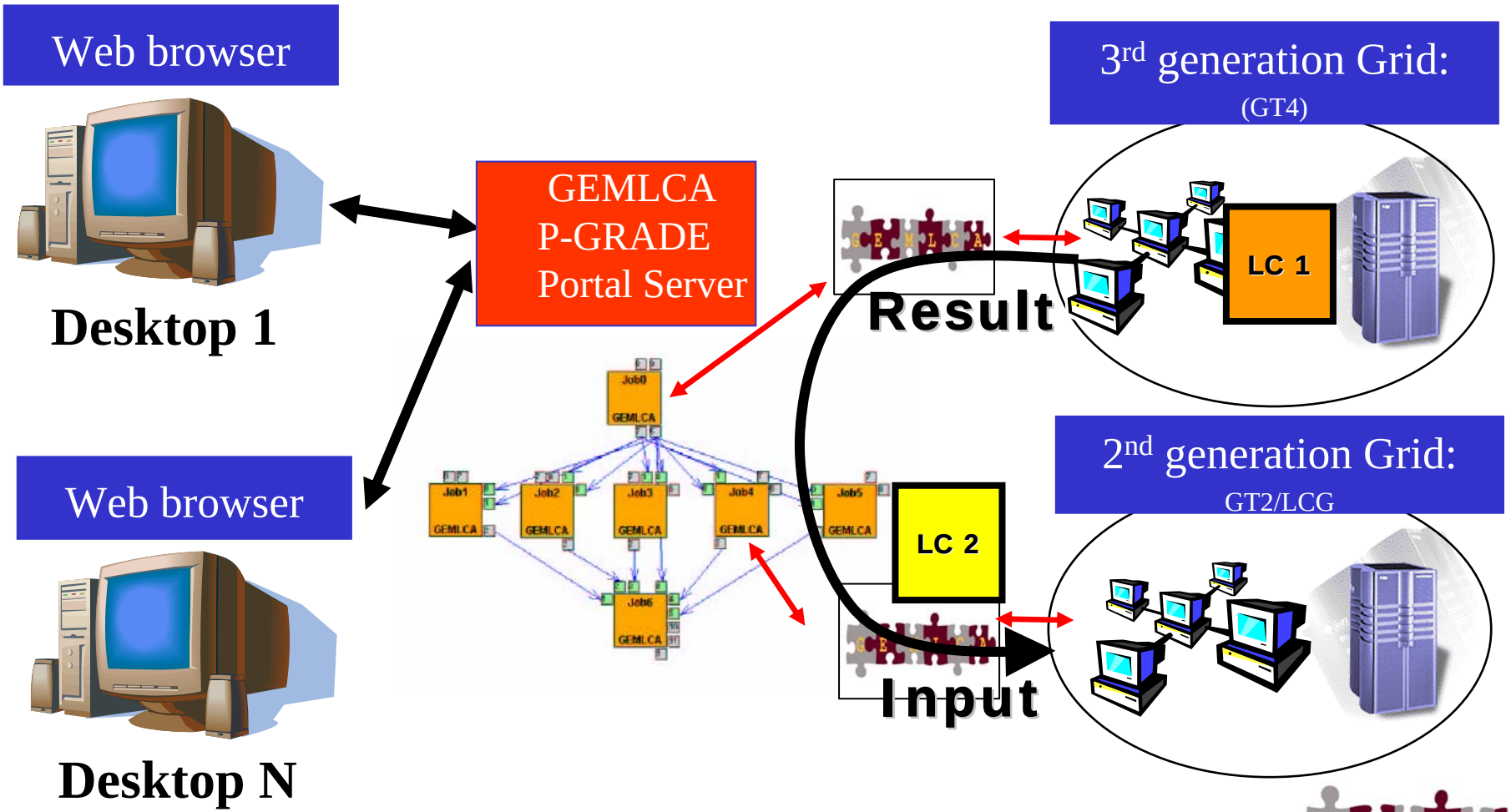
A Web-based GEMLCA client environment…

University of Westminster, London
MTA SZTAKI, Budapest

The aim of the GEMLCA P-GRADE Portal

- To provide graphical clients to GEMLCA with a portal-based solution
- To enable the integration of legacy code grid services into workflows

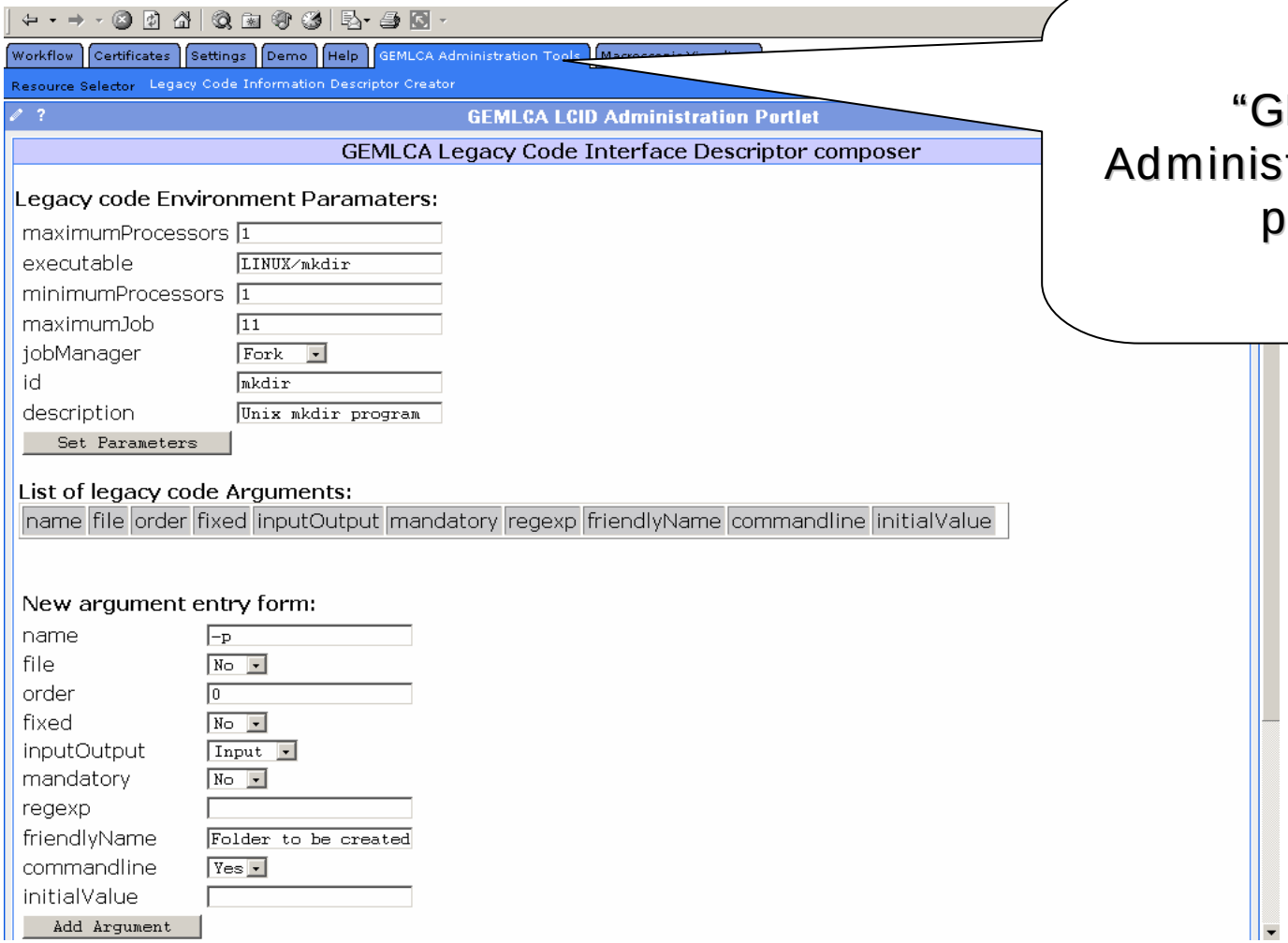
GEMLCA in the P-GRADE Portal



The GEMLCA-specific version of the P-GRADE Portal is different from the original P-GRADE Portal!

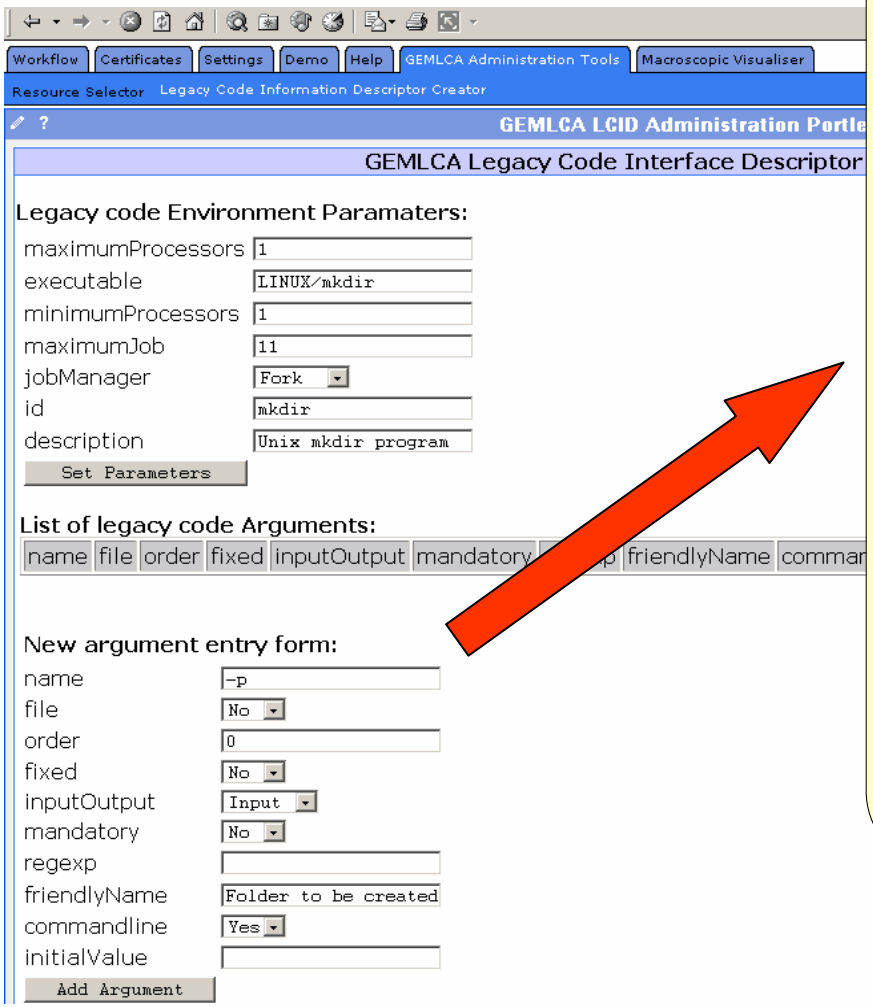
- It contains a web page to register legacy codes as grid services
- It contains a GEMLCA-specific workflow editor
  - Workflow components can be “legacy code grid services” (not only batch jobs)
- It contains a GEMLCA-specific workflow manager subsystem
  - It can invoke GEMLCA services (not only submitting jobs)

Legacy code registration page



"GEMLCA Administration Tool" portlet

Legacy code registration page



Mkdir Legacy Code exposed as a Grid Service

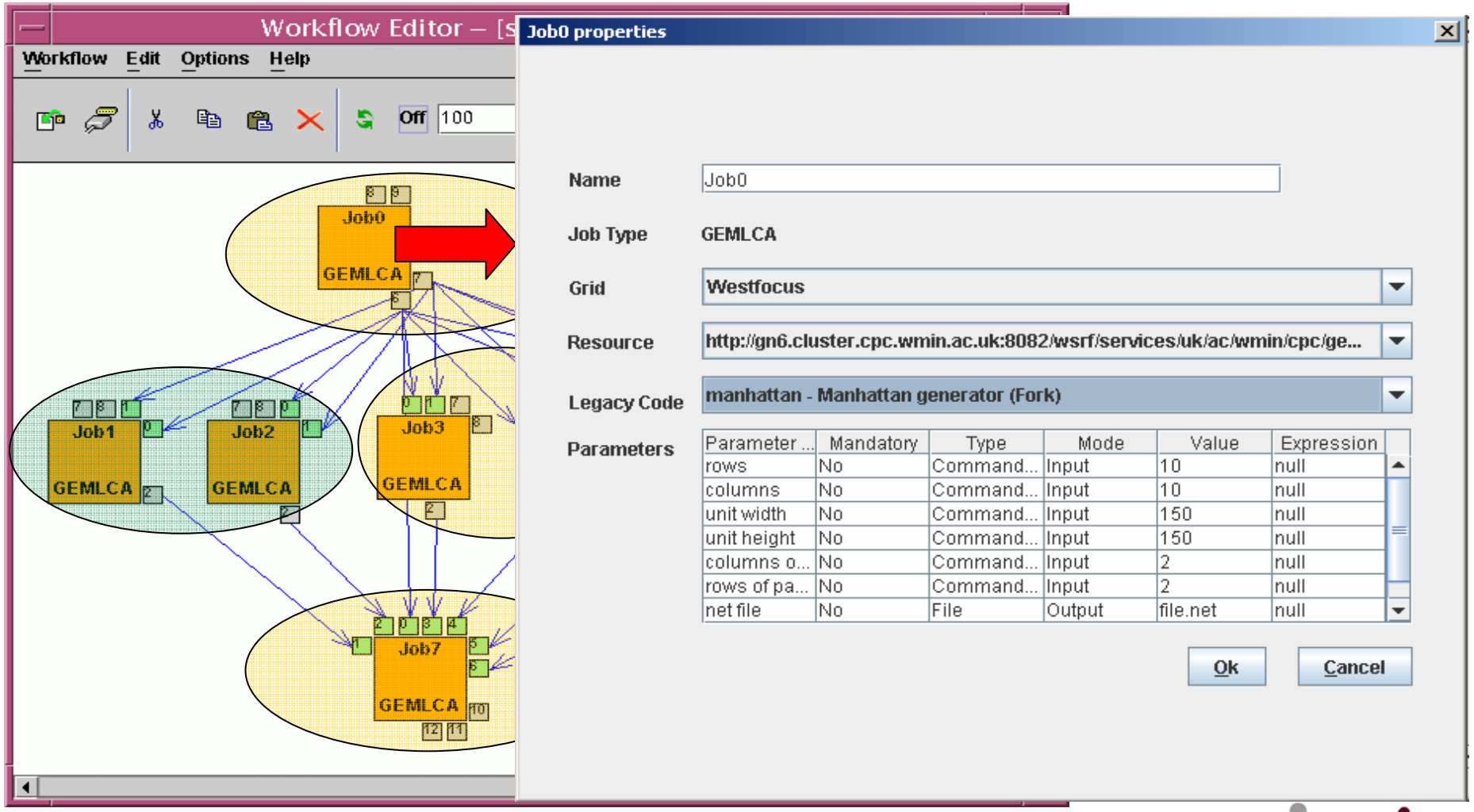
Folder : /../.gemlca/legacycodes/mkdir

Content : i) mkdir binary or link ii) config.xml

Legacy Code Interface Description File: config.xml

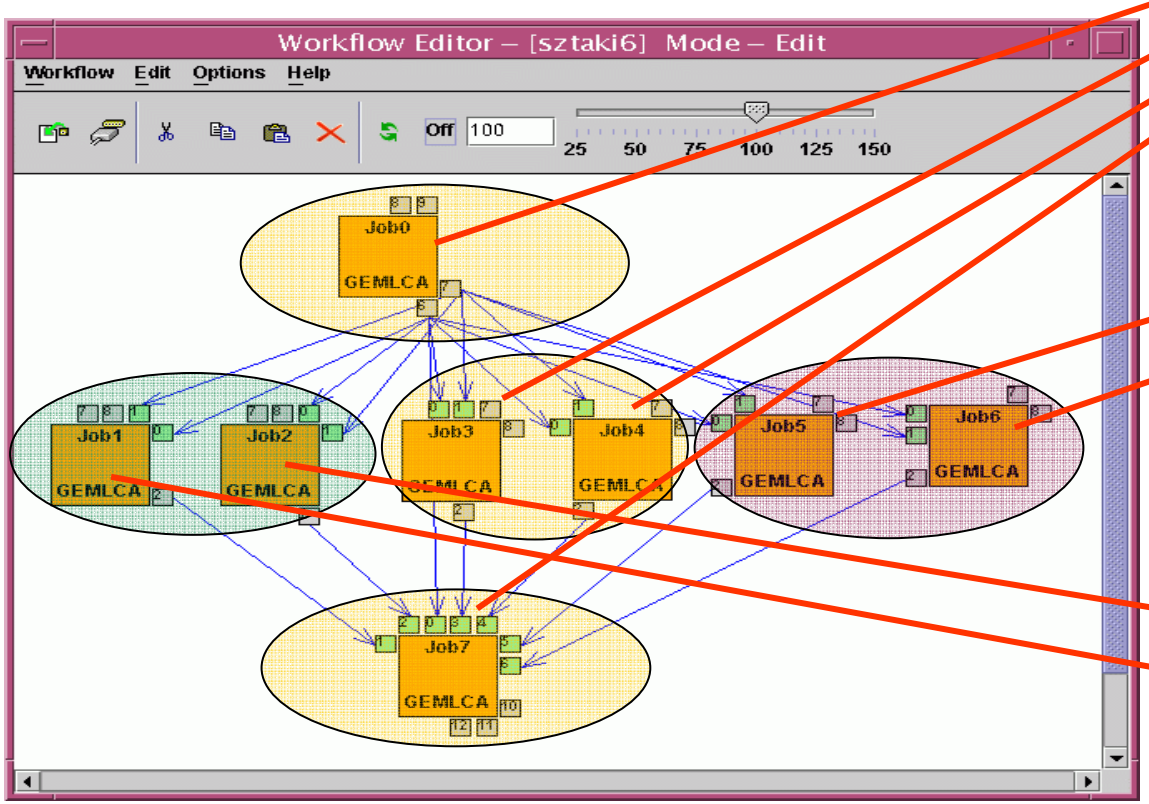
```
<?xml version="1.0"?>
<!DOCTYPE GLCEnvironment "gemlcaconfig.dtd">
<GLCEnvironment
    id="mkdir" executable="LINUX/mkdir" jobManager="Fork"
    maximumJob="11" minimumProcessors="1"
    maximumProcessors="1" universe="PVM"
>
<Description>Unix mkdir program</Description>
 <GLCParameters>
   <Parameter  name="-p" friendlyName="Folder to be created"
          fixed="No" inputOutput="Input" order="0"
          mandatory="No" fileCommandline="Commandline">
       <initialValue> </initialValue>
   </Parameter>
  </GLCParameters>
</GLCEnvironment>
```

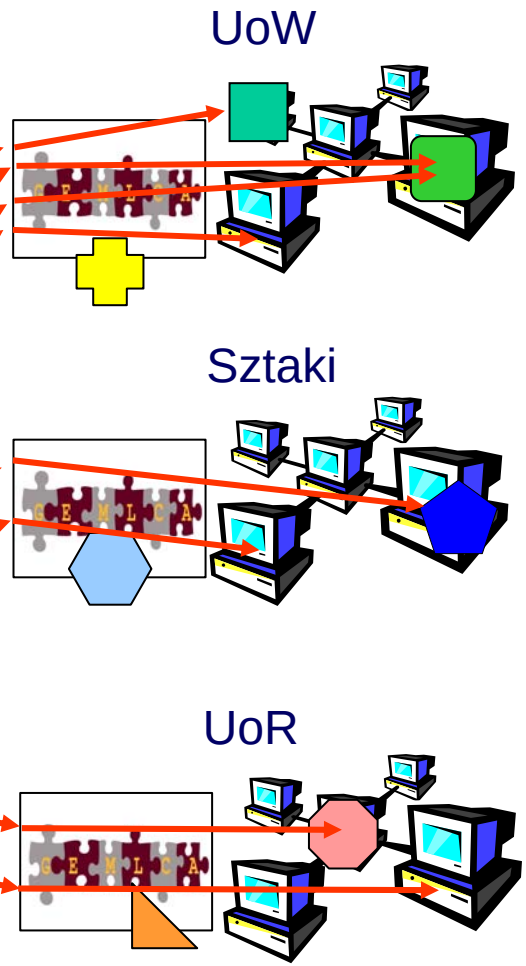
GEMLCA Specific Workflow editor



GEMLCA workflow editor in a nutshell

Workflow Creation





Batch components vs. GEMLCA components in P-GRADE Portal workflows

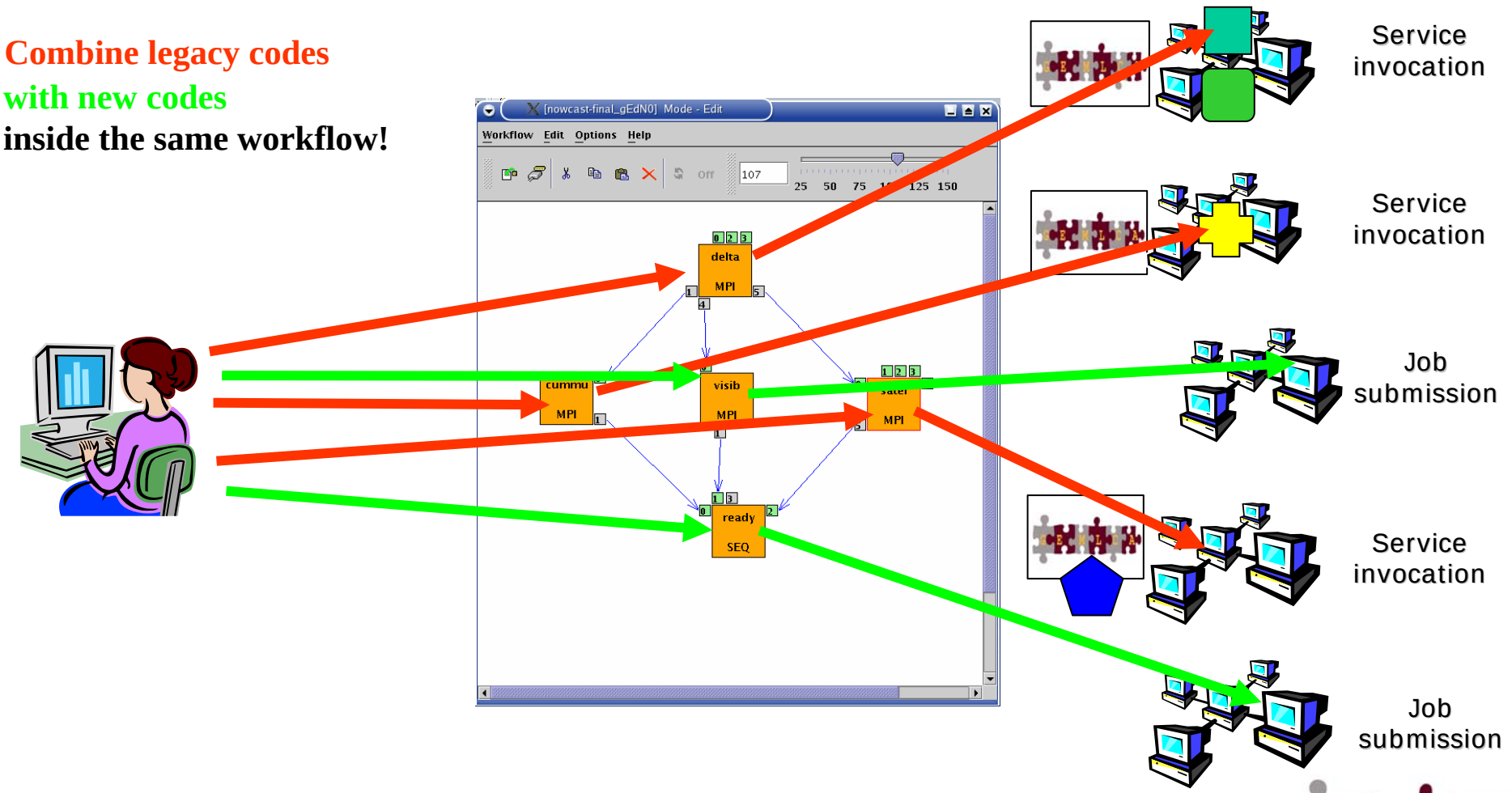
Batch component GEMLCA component

- Workflow components must be defined in different ways
- Input files represented by ports
- Output files represented by ports

Ports guarantee compatibility → batch and GEMLCA components can mutually produce data to each other!

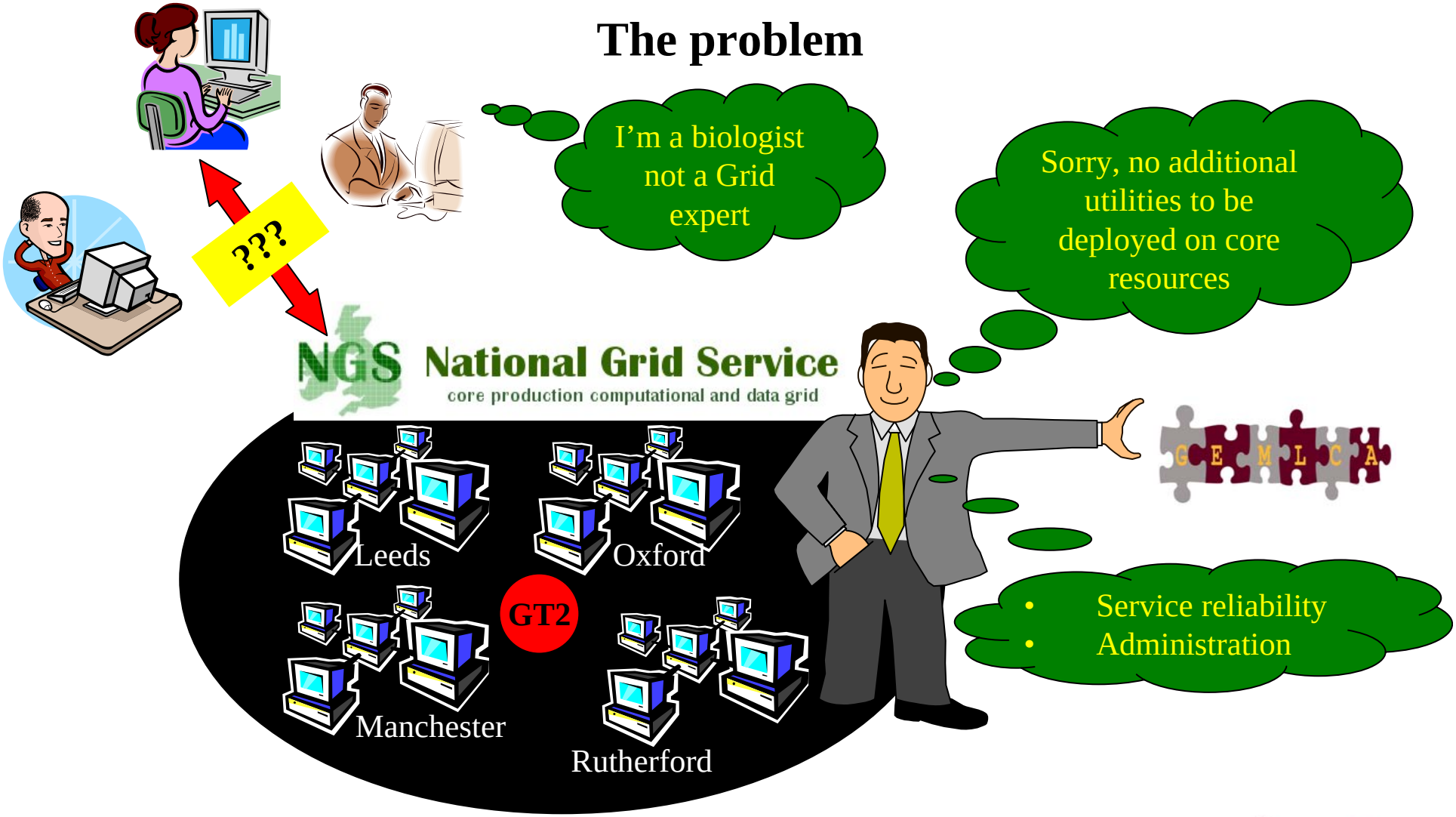
Combining legacy and non-legacy components

Combine legacy codes
with new codes
inside the same workflow!

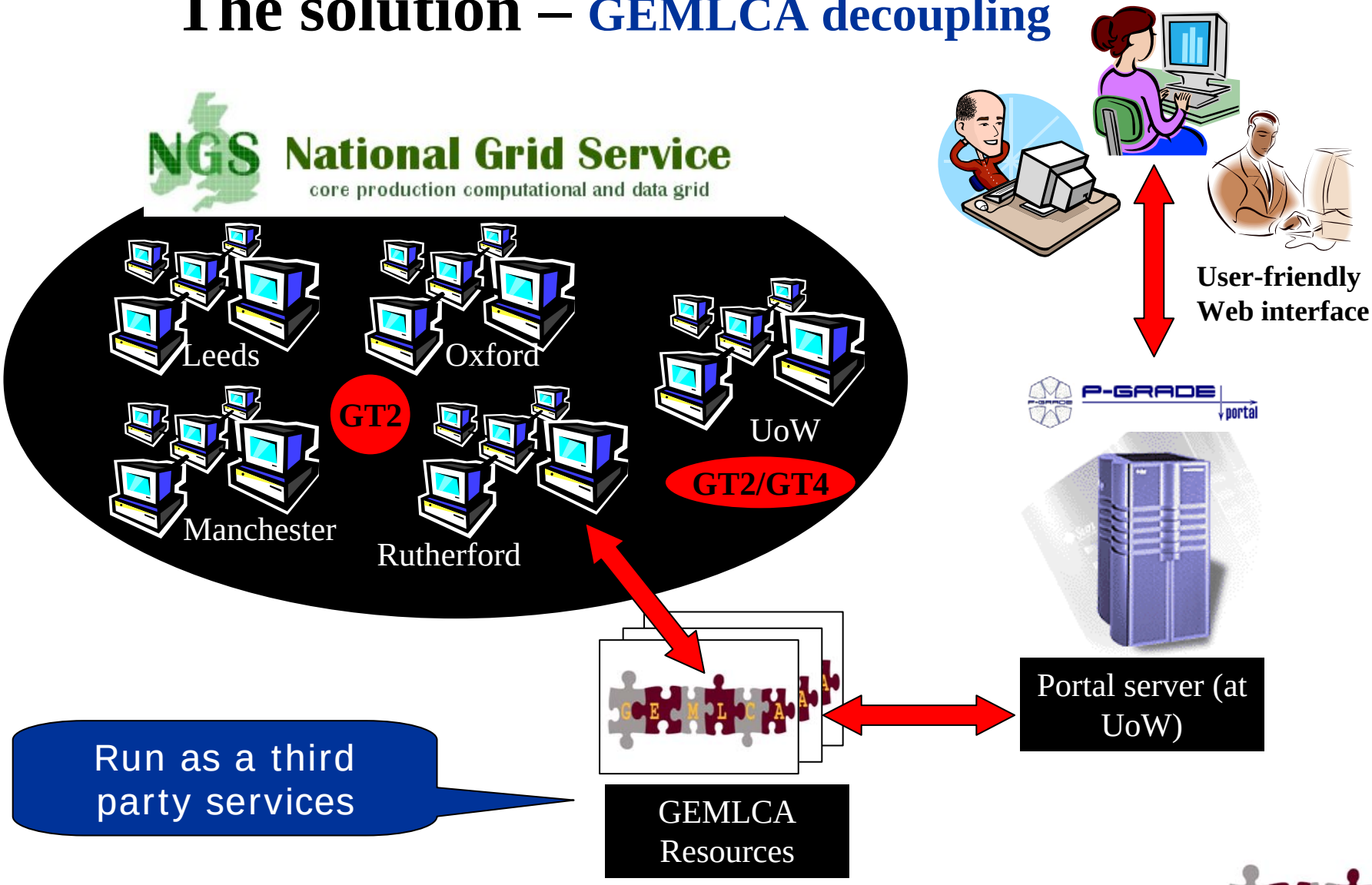


GEMLCA and Production Grids

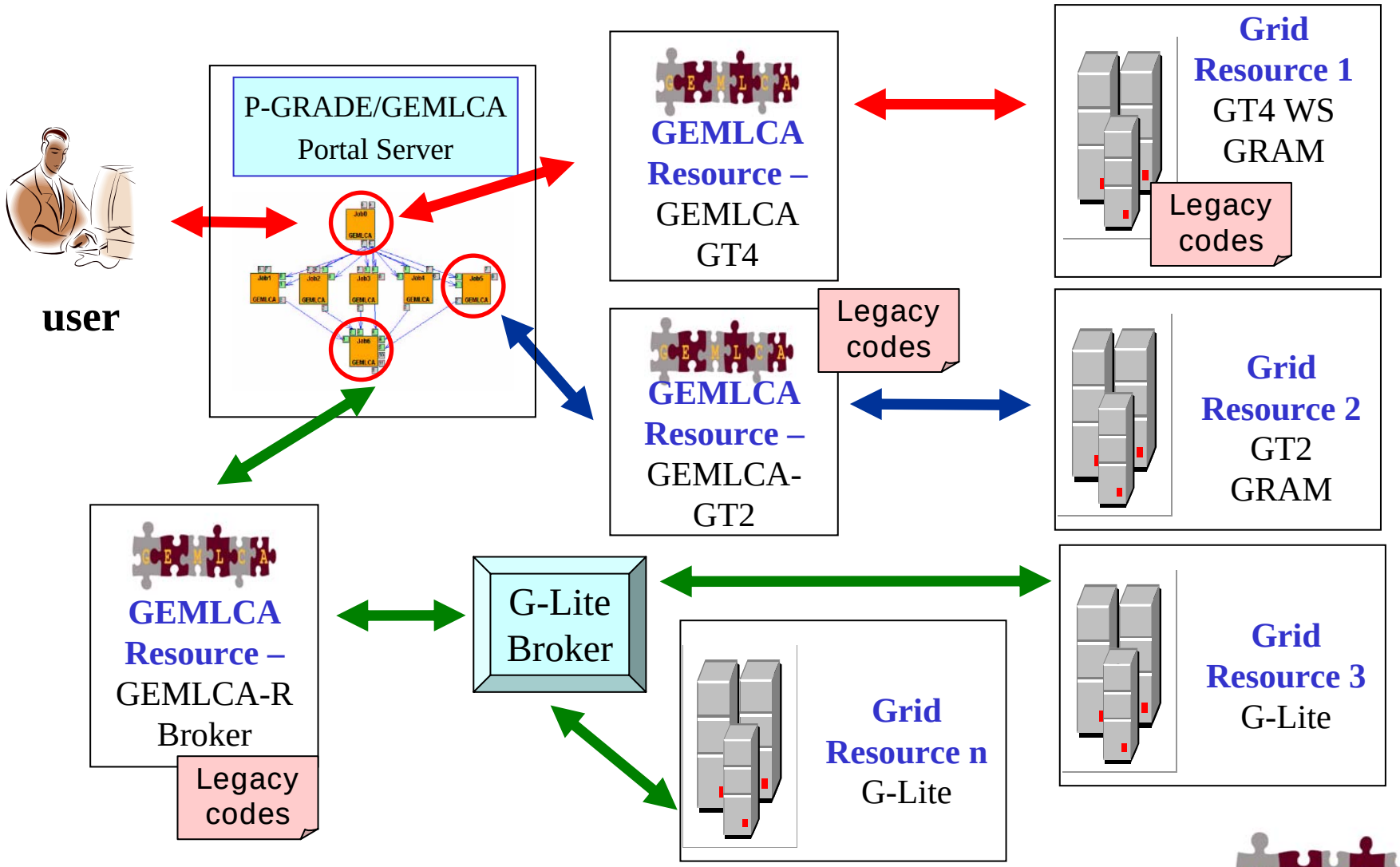
The problem



The solution – GEMLCA decoupling



P-GRADE GEMLCA portal for different Grids



GEMLCA on the UK NGS

The P-GRADE NGS GEMLCA Portal

- **Portal Website:** http://www.cpc.wmin.ac.uk/ngsportal/
- Runs both GT4 and GT2 GEMLCA

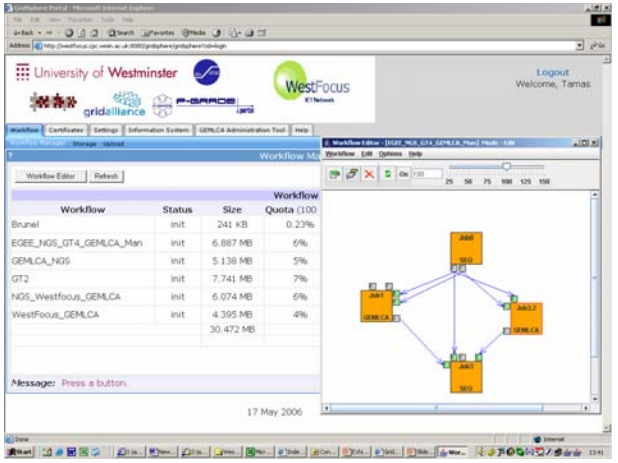
NGS National Grid Service
core production computational and data grid

GEMLCA on the WestFocus GridAlliance Grid

- GT4 testbed for industry and academia
- Connects two 32 machine clusters at Westminster and one at Brunel University
- Runs the P-GRADE Grid portal and GEMLCA
- Connected to and interoperable with the UK NGS

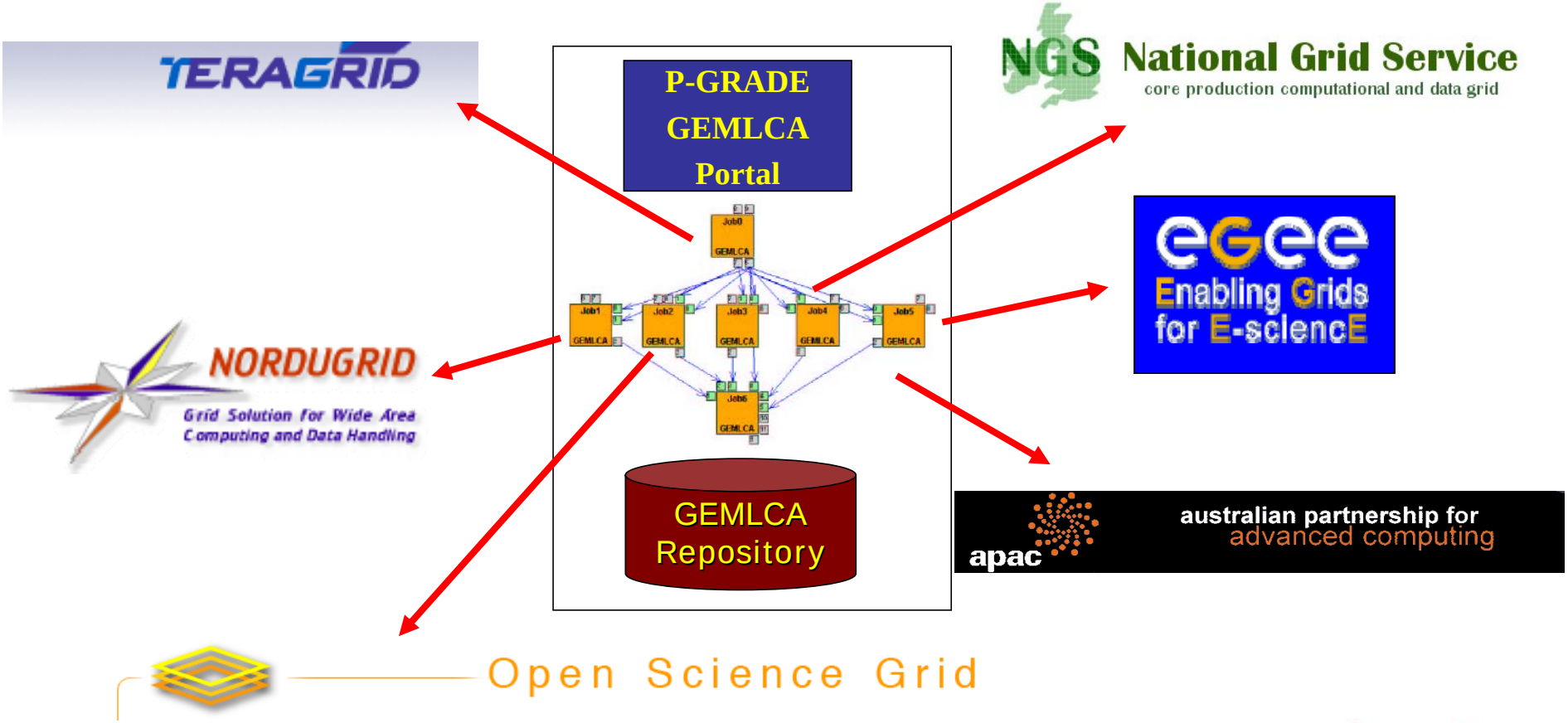
WestFocus ICT Network

gridalliance

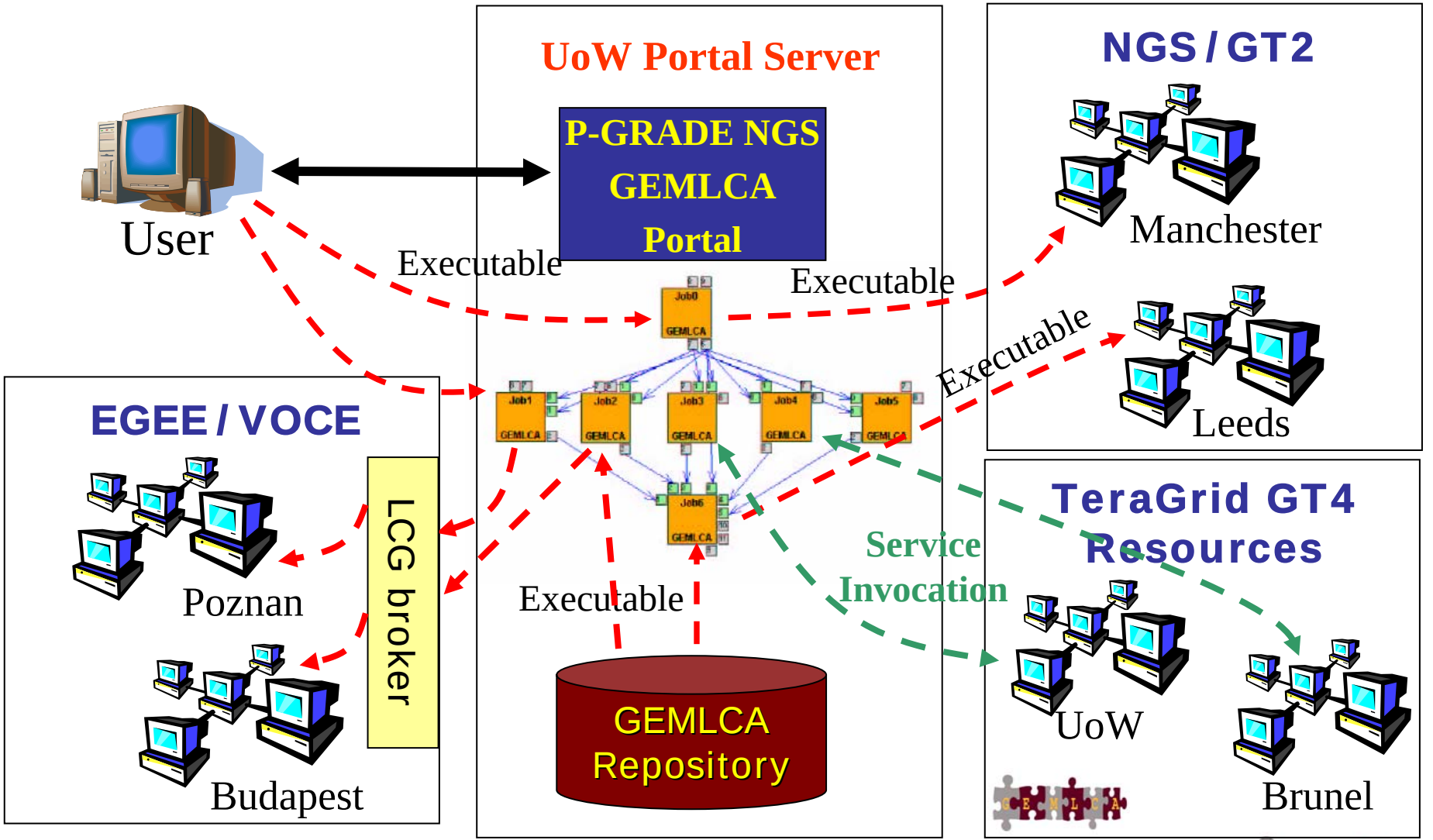


The GIN Resource Testing portal

Portal service to demonstrate workflow level interoperability between major production Grids and monitor GIN resources



Connecting GT2, GT4 and LCG/g-Lite based Grids



Traffic simulation on multiple Grids

GT2 Job submission to NGS (Oxford)

LCG Job submission to EGEE/GIN (IC)

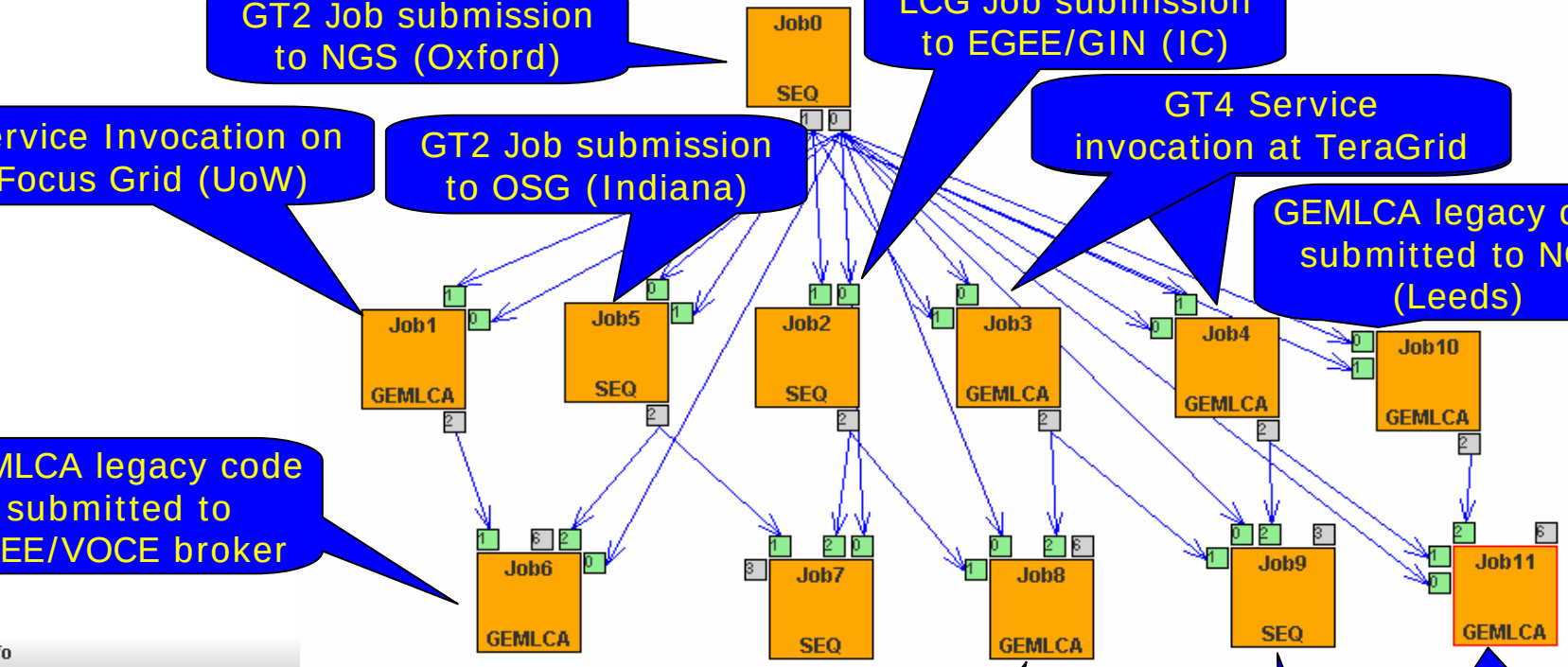
GT4 Service Invocation on WestFocus Grid (UoW)

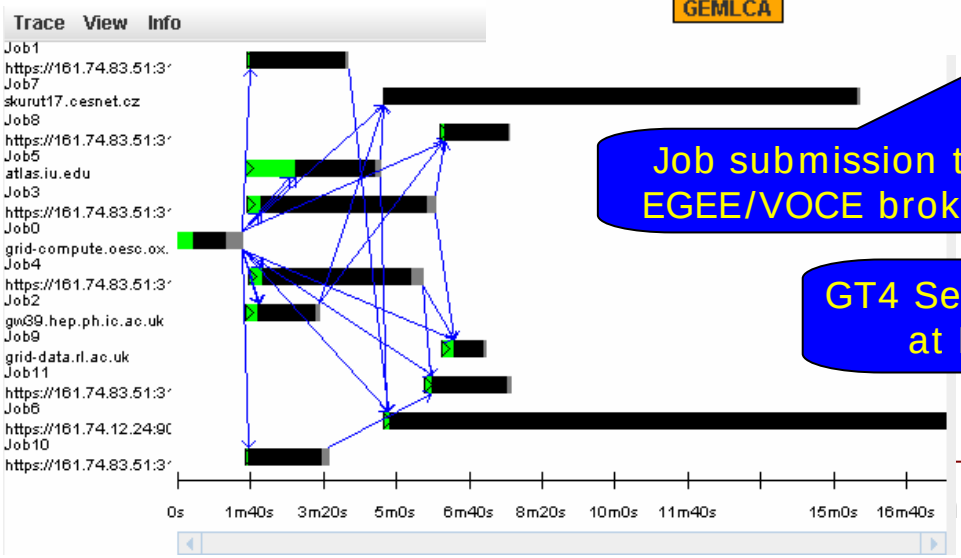
GT2 Job submission to OSG (Indiana)

GT4 Service invocation at TeraGrid

GEMLCA legacy code submitted to NGS (Leeds)

GEMLCA legacy code submitted to EGEE/VOCE broker





Job submission to EGEE/VOCE broker

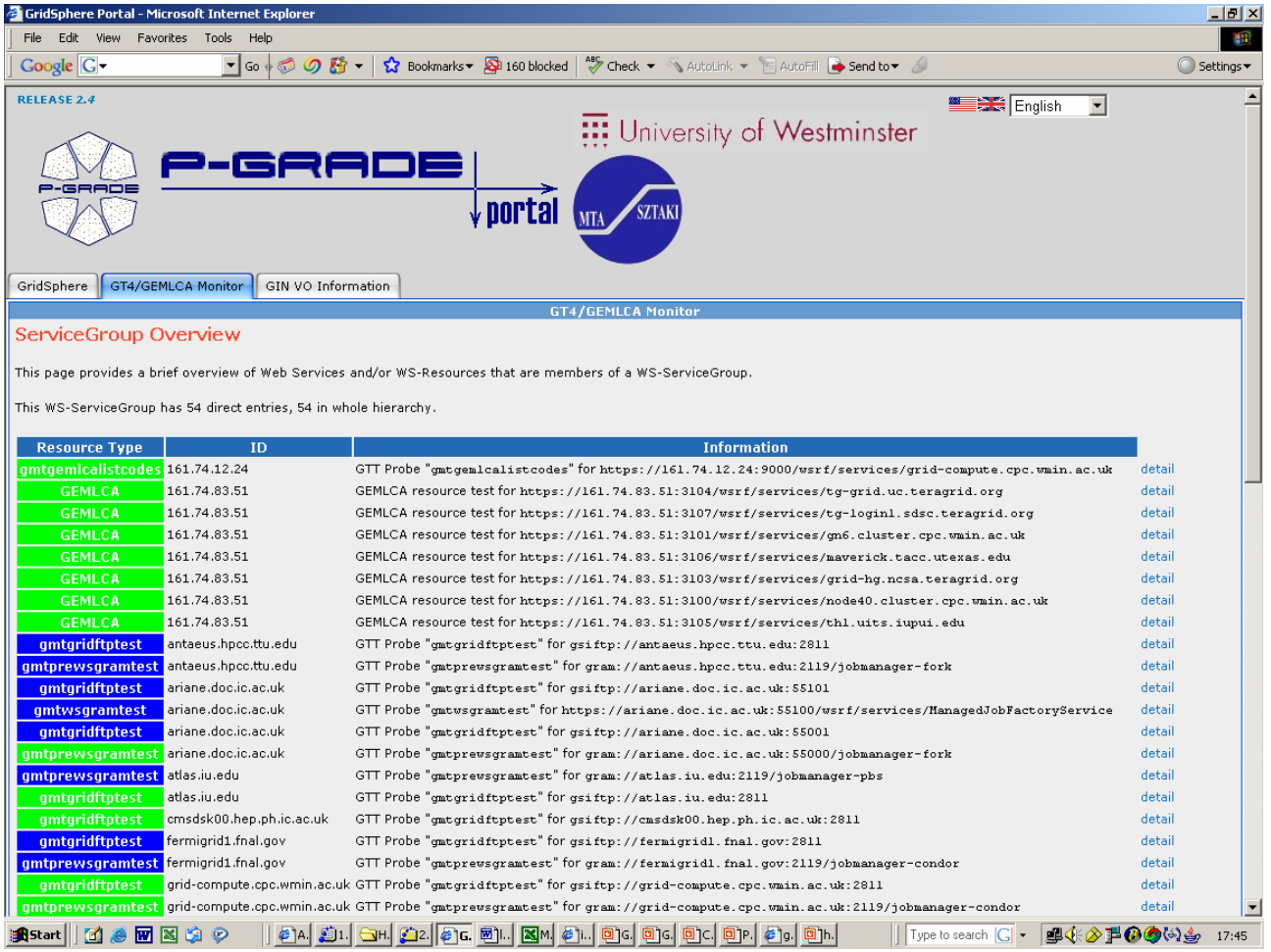
GT4 Service invocation at NGS (UoW)

GT4 Service Invocation on WestFocus Grid (UoW)

GT2 Job submission to NGS (Rutherford)

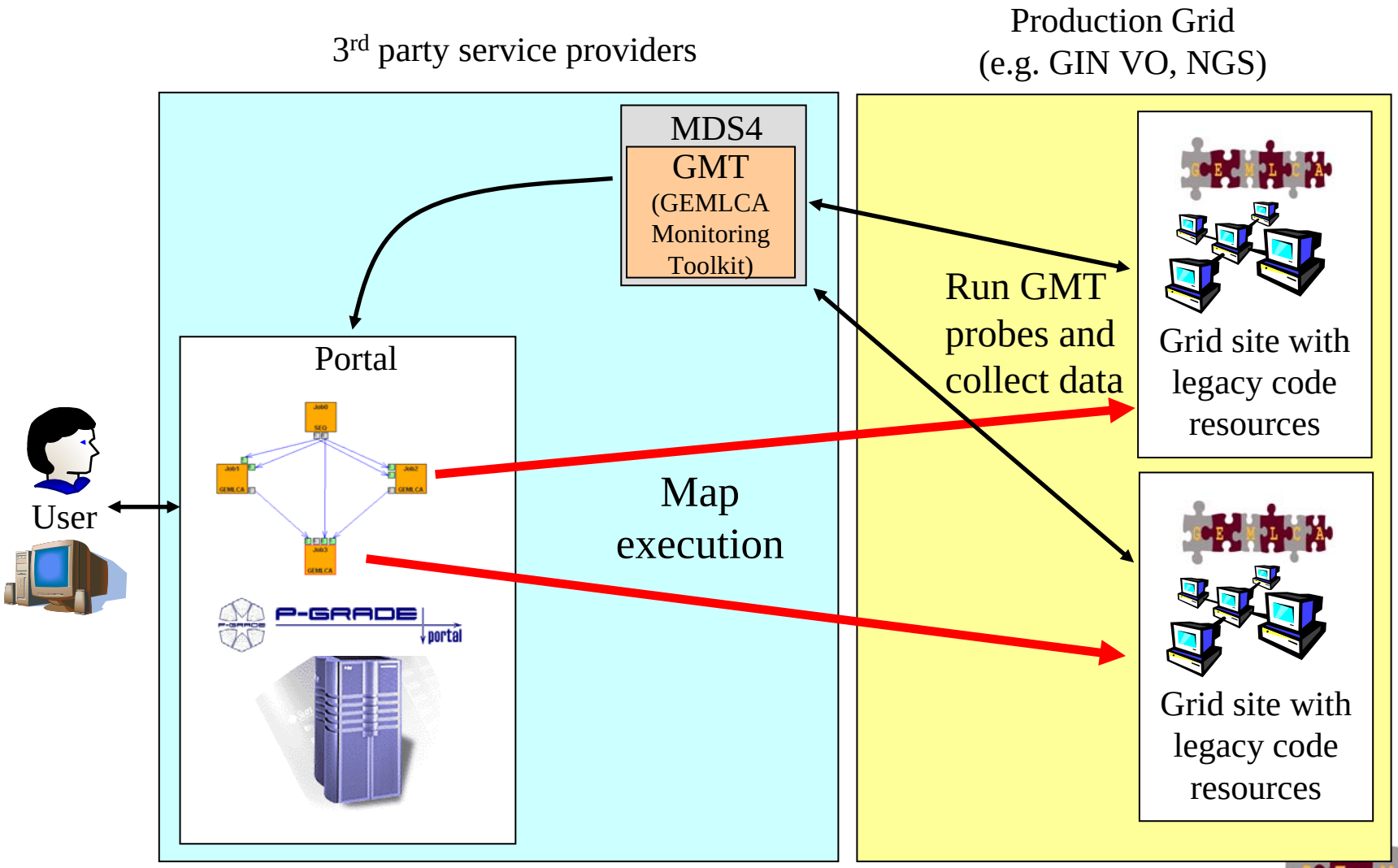
GMT – GEMLCA Monitoring Toolkit

- to test resource availability
- implementation is based on MDS4
- probes are implemented as scripts and their outputs are displayed in a monitoring portlet
- Runs on the NGS and GIN portals



Resource Type	ID	Information	
gmtgemlcalistcodes	161.74.12.24	GTT Probe "gmtgemlcalistcodes" for https://161.74.12.24:9000/wsrf/services/grid-compute.cpc.wmin.ac.uk	detail
GEMLCA	161.74.83.51	GEMLCA resource test for https://161.74.83.51:3104/wsrf/services/tg-grid.uc.teragrid.org	detail
GEMLCA	161.74.83.51	GEMLCA resource test for https://161.74.83.51:3107/wsrf/services/tg-loginl.sdsc.teragrid.org	detail
GEMLCA	161.74.83.51	GEMLCA resource test for https://161.74.83.51:3101/wsrf/services/gm6.cluster.cpc.wmin.ac.uk	detail
GEMLCA	161.74.83.51	GEMLCA resource test for https://161.74.83.51:3106/wsrf/services/maverick.tacc.utexas.edu	detail
GEMLCA	161.74.83.51	GEMLCA resource test for https://161.74.83.51:3103/wsrf/services/grid-hg.ncsa.teragrid.org	detail
GEMLCA	161.74.83.51	GEMLCA resource test for https://161.74.83.51:3100/wsrf/services/node40.cluster.cpc.wmin.ac.uk	detail
GEMLCA	161.74.83.51	GEMLCA resource test for https://161.74.83.51:3105/wsrf/services/thl.uits.iupui.edu	detail
gmtgridftptest	antaeus.hpcc.ttu.edu	GTT Probe "gmtgridftptest" for gsiftp://antaeus.hpcc.ttu.edu:2811	detail
gmtprewsgramtest	antaeus.hpcc.ttu.edu	GTT Probe "gmtprewsgramtest" for gram://antaeus.hpcc.ttu.edu:2119/jobmanager-fork	detail
gmtgridftptest	ariane.doc.ic.ac.uk	GTT Probe "gmtgridftptest" for gsiftp://ariane.doc.ic.ac.uk:55101	detail
gmtwsgramtest	ariane.doc.ic.ac.uk	GTT Probe "gmtwsgramtest" for https://ariane.doc.ic.ac.uk:55100/wsrf/services/ManagedJobFactoryService	detail
gmtgridftptest	ariane.doc.ic.ac.uk	GTT Probe "gmtgridftptest" for gsiftp://ariane.doc.ic.ac.uk:55001	detail
gmtprewsgramtest	ariane.doc.ic.ac.uk	GTT Probe "gmtprewsgramtest" for gram://ariane.doc.ic.ac.uk:55000/jobmanager-fork	detail
gmtprewsgramtest	atlas.iu.edu	GTT Probe "gmtprewsgramtest" for gram://atlas.iu.edu:2119/jobmanager-pbs	detail
gmtgridftptest	atlas.iu.edu	GTT Probe "gmtgridftptest" for gsiftp://atlas.iu.edu:2811	detail
gmtgridftptest	cmsdsk00.hep.ph.ic.ac.uk	GTT Probe "gmtgridftptest" for gsiftp://cmsdsk00.hep.ph.ic.ac.uk:2811	detail
gmtgridftptest	fermigrid1.fnal.gov	GTT Probe "gmtgridftptest" for gsiftp://fermigrid1.fnal.gov:2811	detail
gmtprewsgramtest	fermigrid1.fnal.gov	GTT Probe "gmtprewsgramtest" for gram://fermigrid1.fnal.gov:2119/jobmanager-condor	detail
gmtgridftptest	grid-compute.cpc.wmin.ac.uk	GTT Probe "gmtgridftptest" for gsiftp://grid-compute.cpc.wmin.ac.uk:2811	detail
gmtprewsgramtest	grid-compute.cpc.wmin.ac.uk	GTT Probe "gmtprewsgramtest" for gram://grid-compute.cpc.wmin.ac.uk:2119/jobmanager-condor	detail

GMT architecture



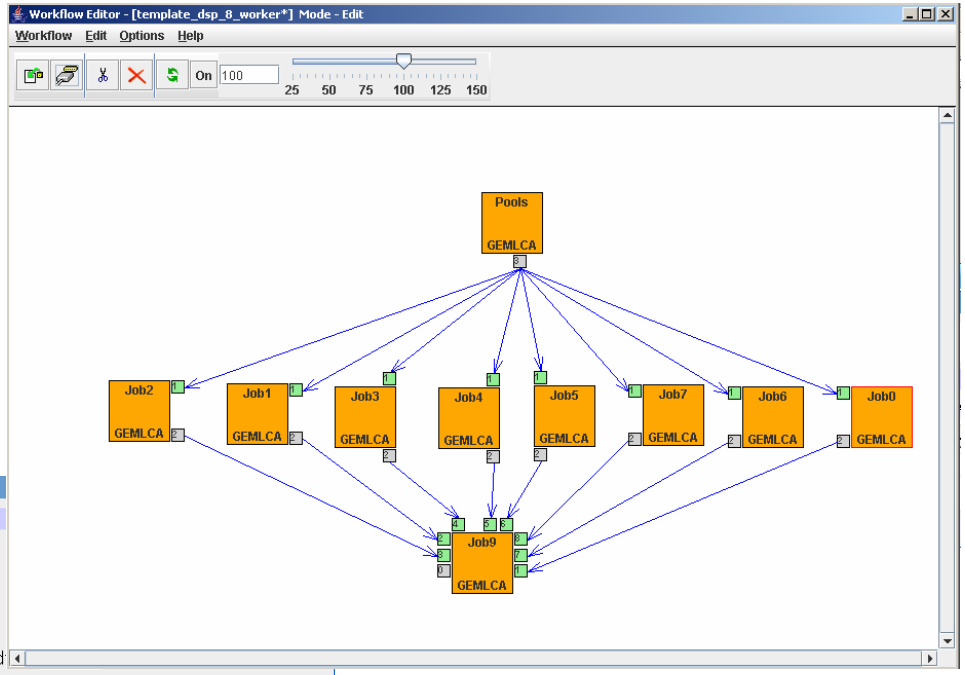
What does GMT test? (just examples)

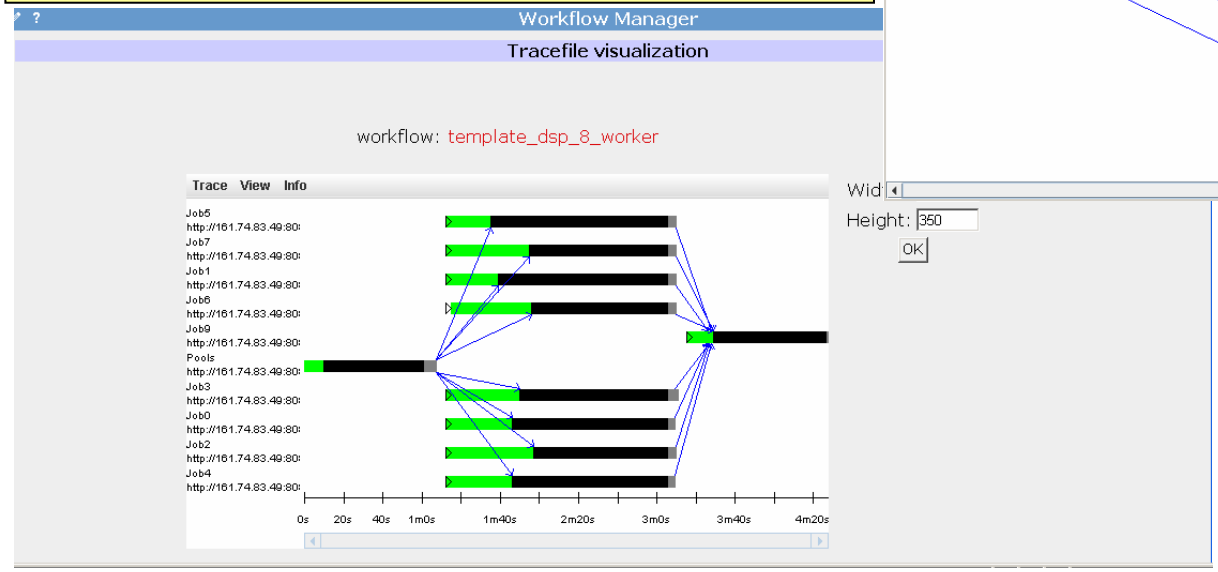
- Basic network connectivity of a remote host
- Remote MyProxy server is running and accepting requests
- Remote host GridFTP server is accepting file transfer requests
- Test Globus job submission (WS-GRAM)
- Verify the availability of the local information system (MDS service)
- Test local job manager (Condor, PBS, SGE etc.)
- Check GEMLCA services: GLCAdmin, GLCList, GLCProcess

Application examples

DSP-Designing Optimal Periodic Nonuniform Sampling Sequences

T Factor	Sequential	GEMLCA
18	~19min	~8min
20	~3h 33min	35min
22	~41h 53min	~7h 23min

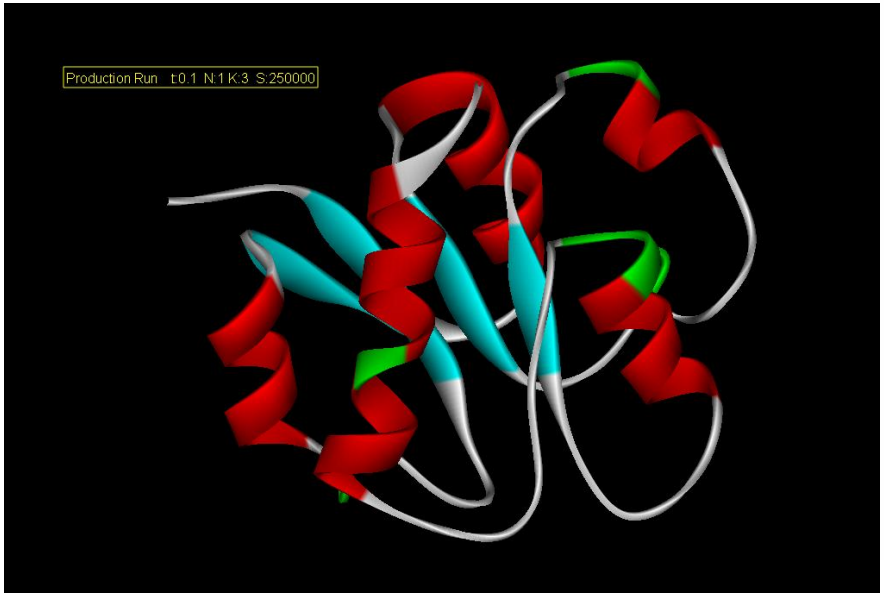


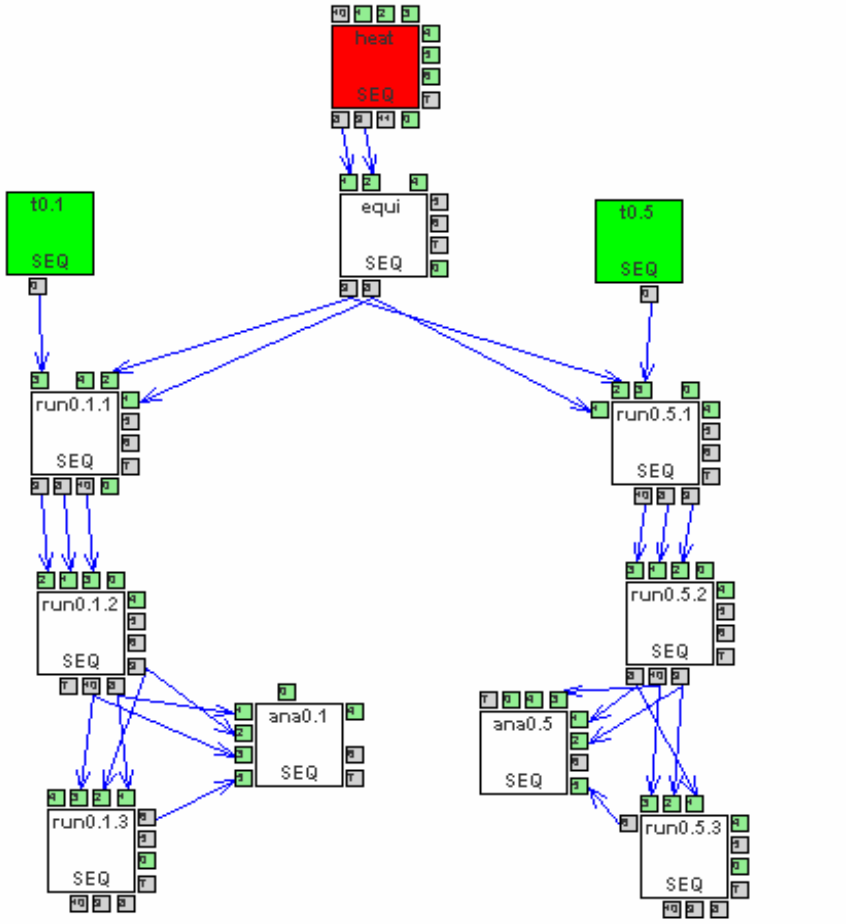


Application examples

Molecular Dynamics Study of Water Penetration in Staphylococcal Nuclease using CHARMm

- Analysis of several production runs with different parameters following a common heating and equilibrium phase





Conclusions

- GEMLCA enables the deployment of legacy code applications as Grid services without any real user effort.

- GEMLCA is integrated with the P-GRADE portal to offer user-friendly development and execution environment.

- The integrated GEMLCA P-GRADE solution is available
  - for the UK NGS as a service!
    www.cpc.wmin.ac.uk/ngsportal
  - for the GIN VO as a resource test service

https://gin-portal.cpc.wmin.ac.uk:8080/gridsphere/gridsphere

Thank you for your attention!

http://www.cpc.wmin.ac.uk/gemlca

gemlca-discuss@cpc.wmin.ac.uk