



Grid Enabling Legacy Applications

Grid Execution Management for Legacy Code Applications



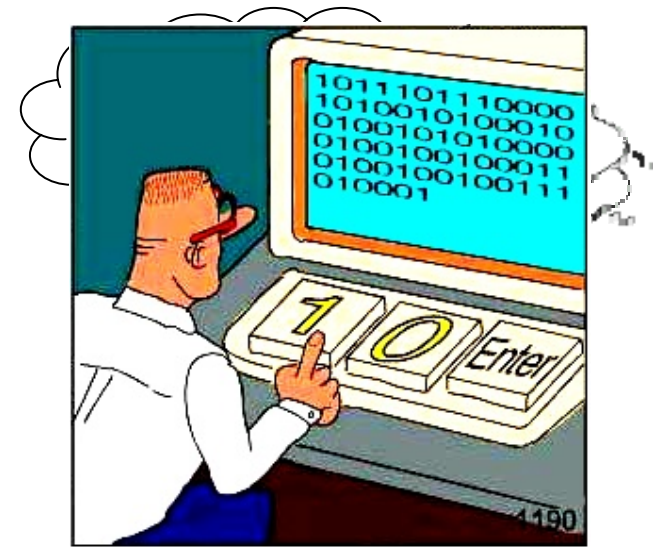


Legacy Applications

- 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





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

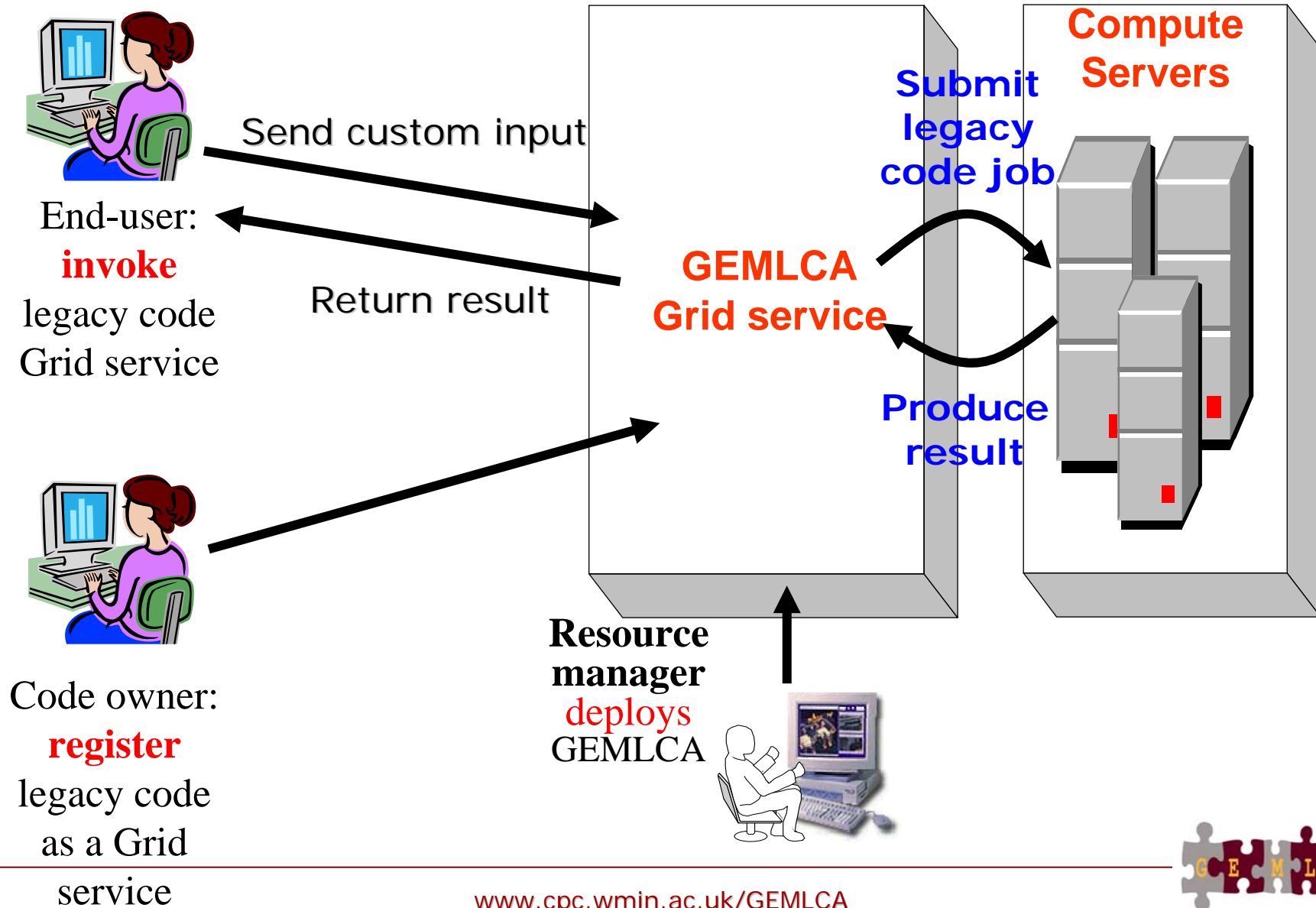
GEMMLCA

**GEMMLCA
PGPortal
Integration**





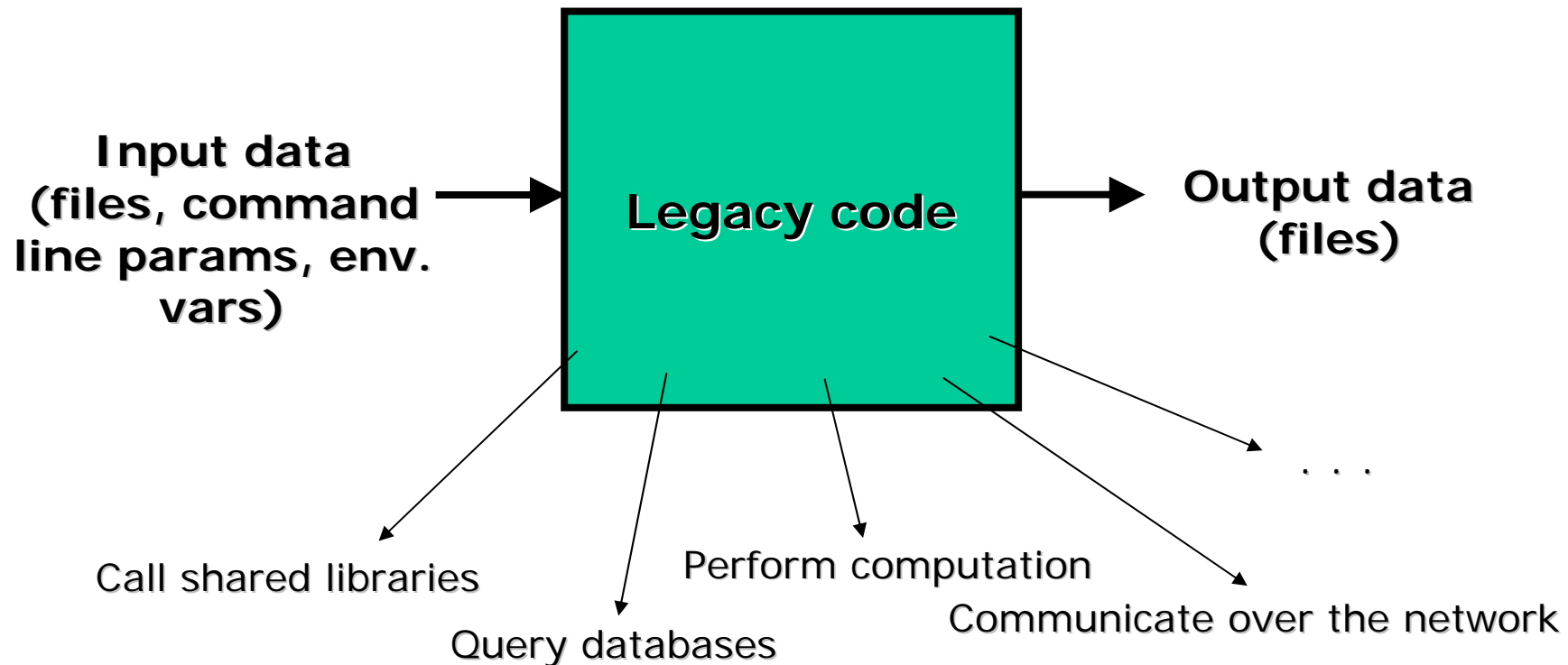
GEMMLCA Concept





The GEMMLCA-view of a legacy code

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





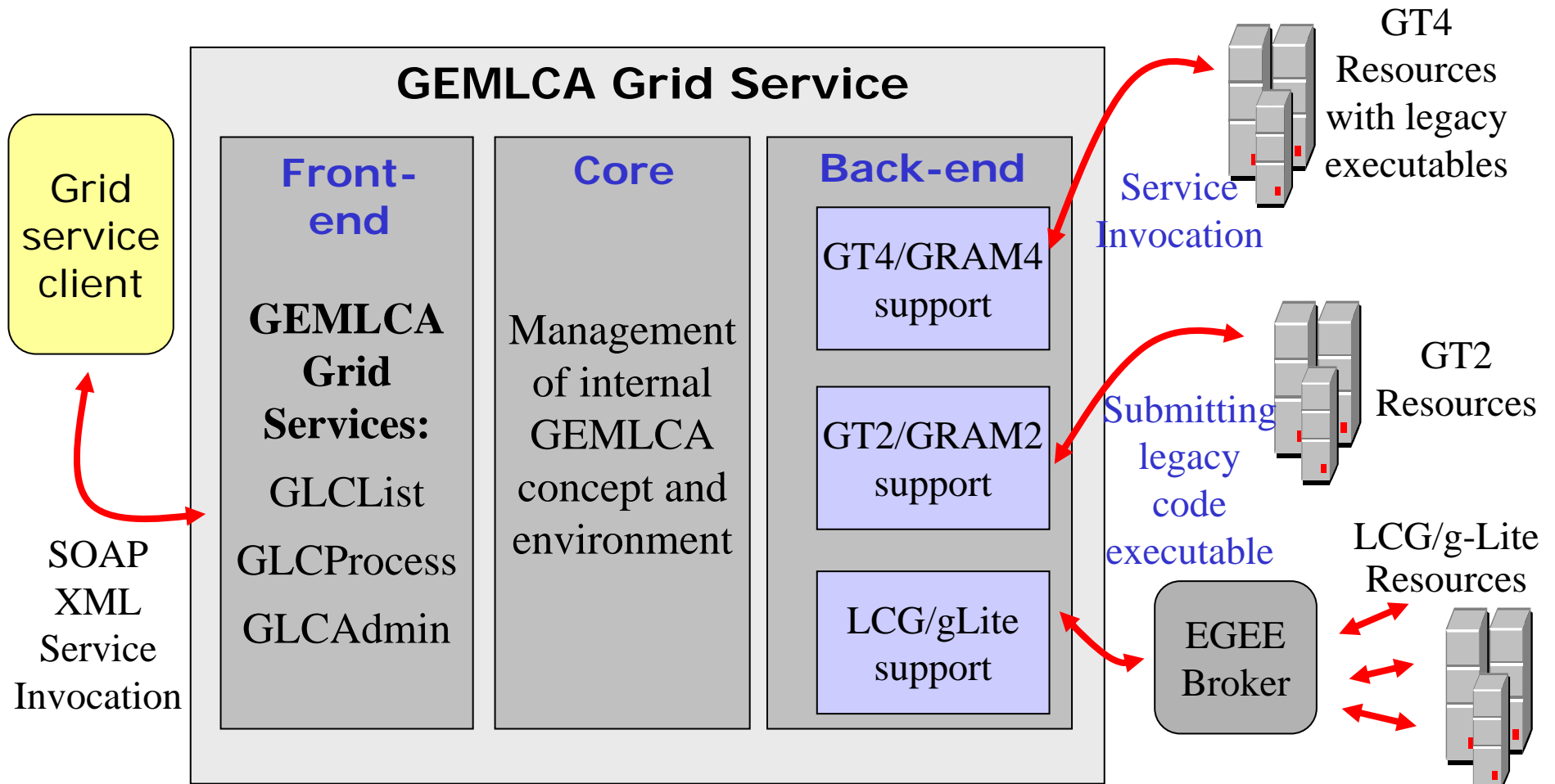
Implementing the concept

- The GEMMLCA service can be implemented with any grid/service-oriented technology E.g:
 - Globus (3 or) 4 → **currently available implementations**
 - Jini
 - Web services
 - ...
- GEMMLCA 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)**



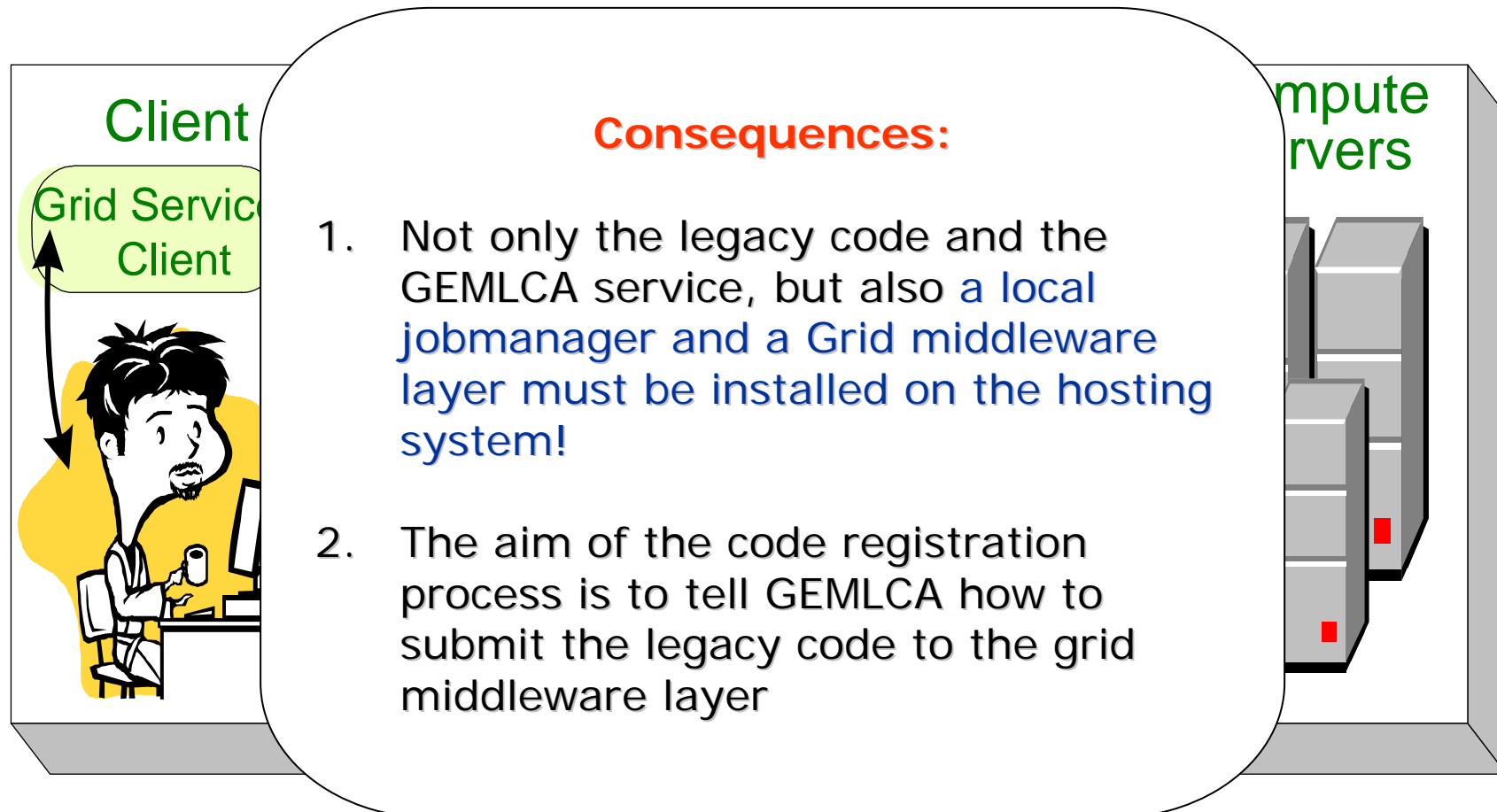


Implementing the GEMMLCA concept





What's behind the GEMMLCA service...





What's the point?

- **Heterogeneous codes can be hidden behind the same interface** (the programming interface of the GEMMLCA 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 GEMMLCA clients** (such as the GEMMLCA P-GRADE Portal)
 - Compose complex processes by connecting multiple legacy code grid services together





The GEMMLCA P-GRADE Portal

A Web-based GEMMLCA client environment...

University of Westminster, London
MTA SZTAKI, Budapest





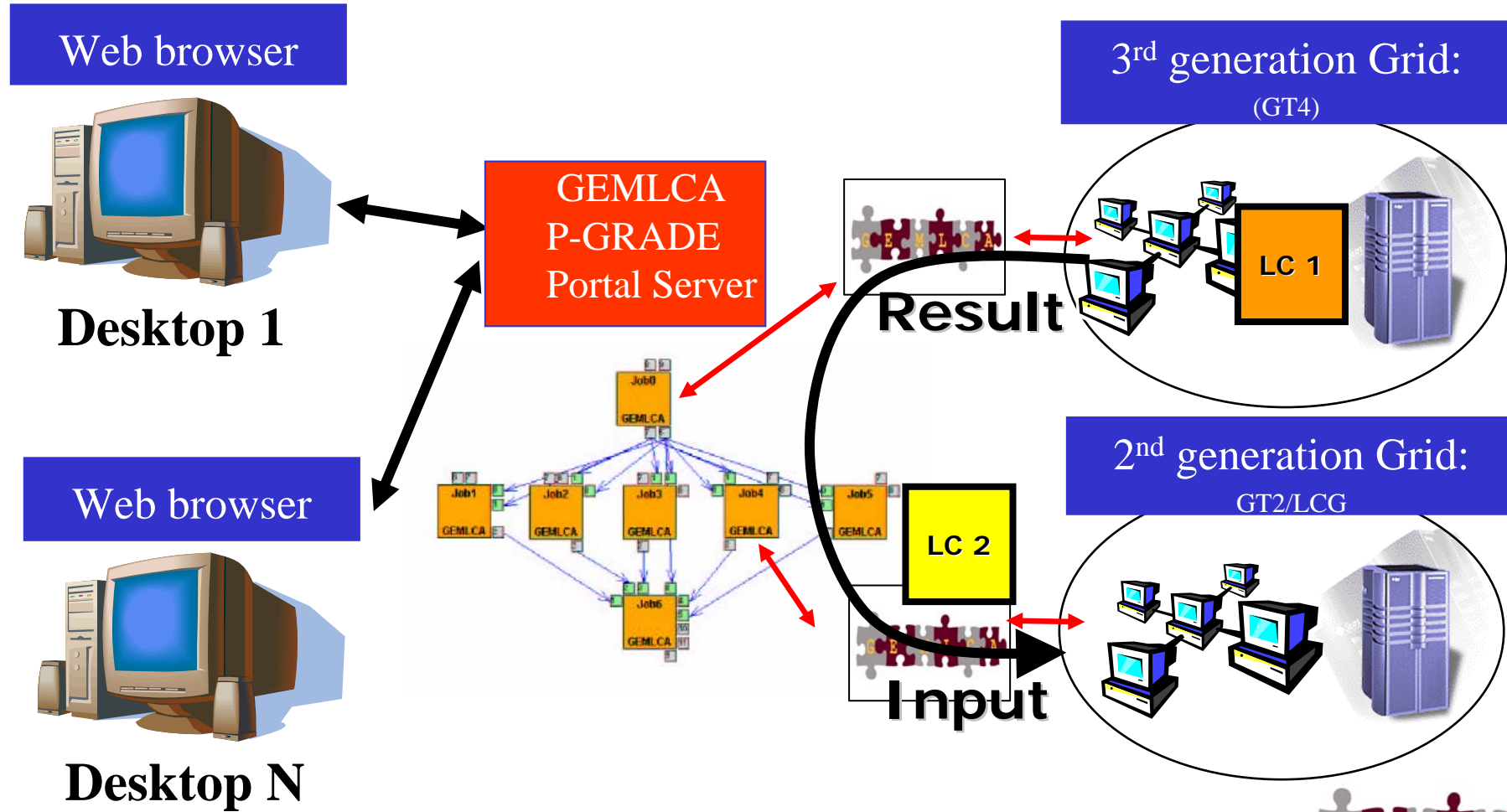
The aim of the GEMMLCA P-GRADE Portal

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





GEMMLCA in the P-GRADE Portal





The GEMMLCA-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 GEMMLCA-specific workflow editor
 - Workflow components can be “legacy code grid services” (not only batch jobs)
- It contains a GEMMLCA-specific workflow manager subsystem
 - It can invoke GEMMLCA services (not only submitting jobs)





Legacy code registration page

Workflow Certificates Settings Demo Help GEMMLCA Administration Tools Macroprocessor...

Resource Selector Legacy Code Information Descriptor Creator

GEMMLCA LCID Administration Portlet

GEMMLCA Legacy Code Interface Descriptor composer

Legacy code Environment Paramaters:

maximumProcessors

executable

minimumProcessors

maximumJob

jobManager

id

description

List of legacy code Arguments:

name	file	order	fixed	inputOutput	mandatory	regexp	friendlyName	commandline	initialValue
------	------	-------	-------	-------------	-----------	--------	--------------	-------------	--------------

New argument entry form:

name

file

order

fixed

inputOutput

mandatory

regexp

friendlyName

commandline

initialValue

"GEMMLCA Administration Tool" portlet





Legacy code registration page

Workflow Certificates Settings Demo Help GEMLCA Administration Tools Macroscopic Visualiser

Resource Selector Legacy Code Information Descriptor Creator

GEMLCA LCID Administration Portal

GEMLCA Legacy Code Interface Descriptor

Legacy code Environment Parameters:

maximumProcessors

executable

minimumProcessors

maximumJob

jobManager

id

description

List of legacy code Arguments:

name	file	order	fixed	inputOutput	mandatory	friendlyName	commandline
-p	No	0	No	Input	No	Folder to be created	Yes

New argument entry form:

name

file

order

fixed

inputOutput

mandatory

regex

friendlyName

commandline

initialValue

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>
```





GEMMLCA Specific Workflow editor

Workflow Editor – [s] Job0 properties

Workflow Edit Options Help

Off 100

Job0 properties

Name: Job0

Job Type: GEMMLCA

Grid: Westfocus

Resource: <http://gn6.cluster.cpc.wmin.ac.uk:8082/wsrf/services/uk/ac/wmin/cpc/ge...>

Legacy Code: manhattan - Manhattan generator (Fork)

Parameters

Parameter ...	Mandatory	Type	Mode	Value	Expression
rows	No	Command...	Input	10	null
columns	No	Command...	Input	10	null
unit width	No	Command...	Input	150	null
unit height	No	Command...	Input	150	null
columns o...	No	Command...	Input	2	null
rows of pa...	No	Command...	Input	2	null
net file	No	File	Output	file.net	null

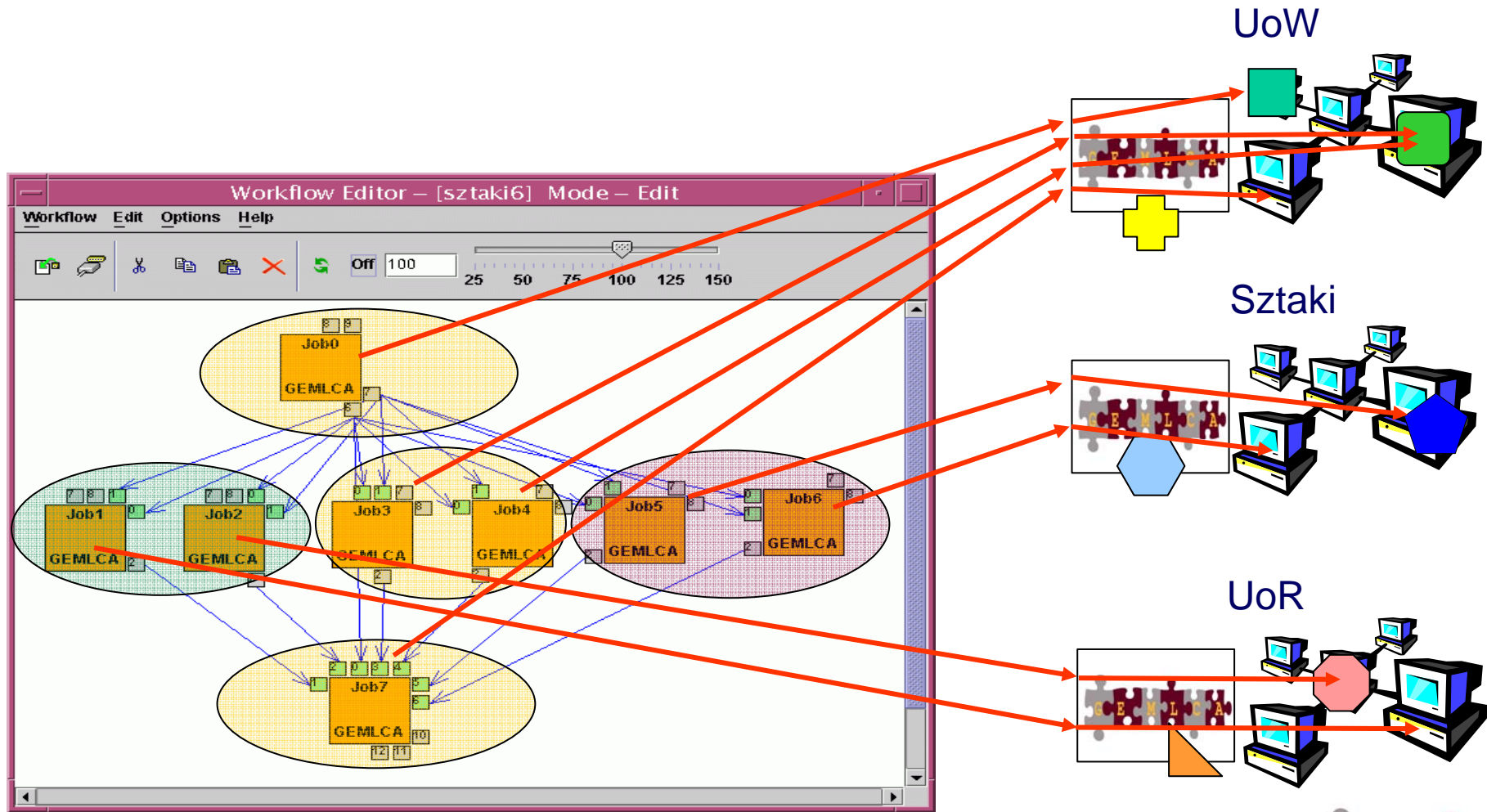
Ok Cancel





GEMMLCA workflow editor in a nutshell

Workflow Creation





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

Batch component

GEMMLCA component

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

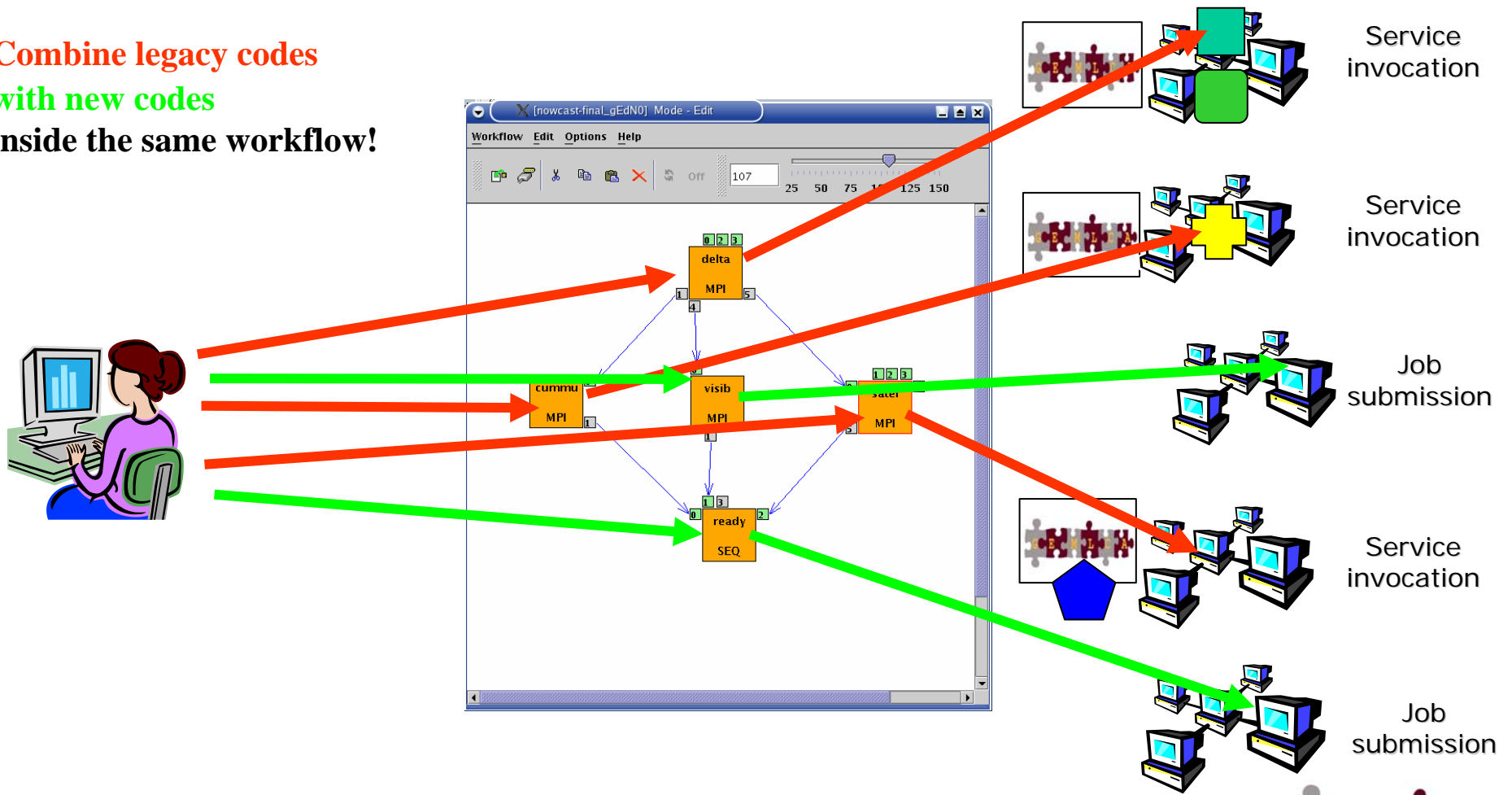
Ports guarantee compatibility → batch and GEMMLCA components can mutually produce data to each other!





Combining legacy and non-legacy components

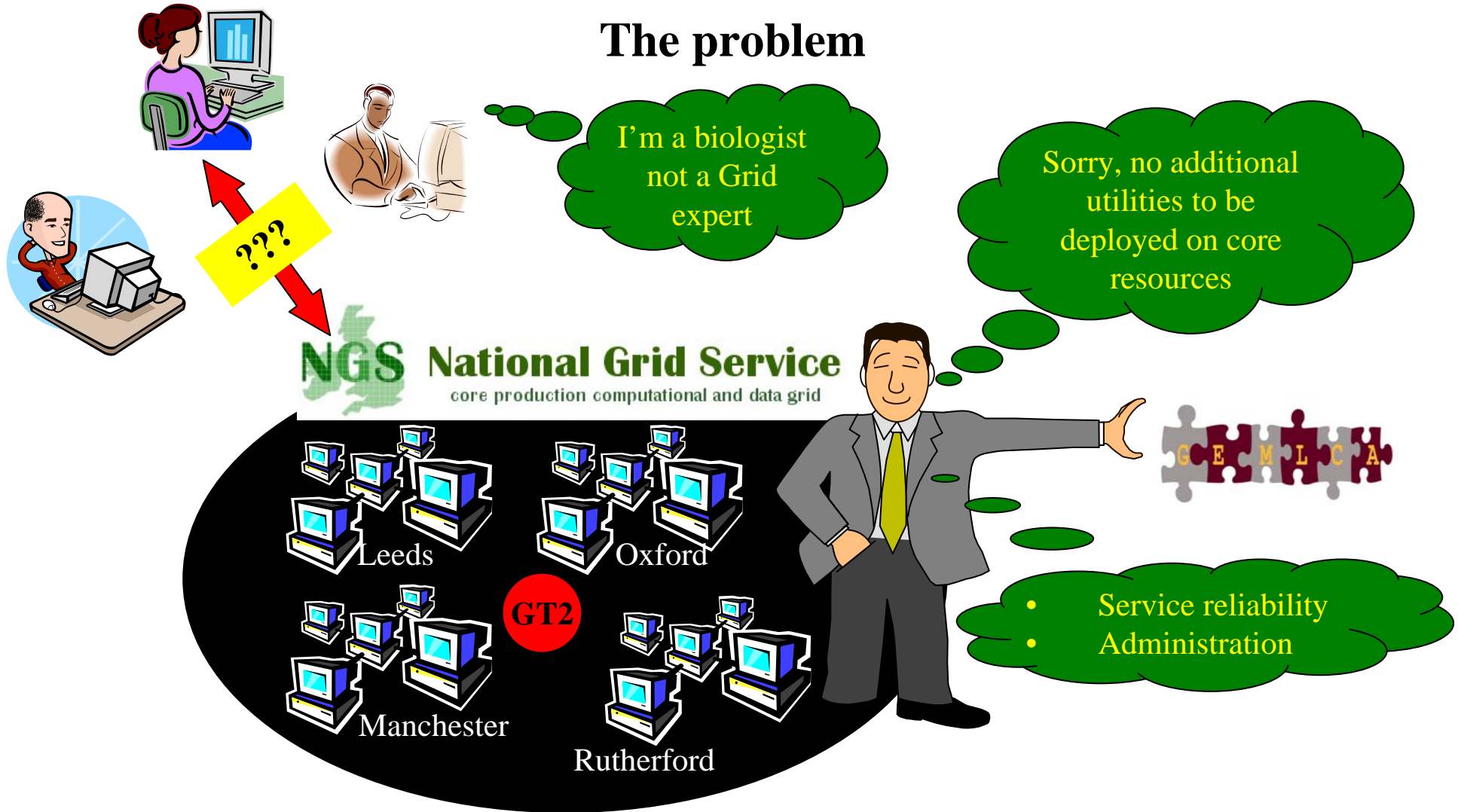
Combine legacy codes
with new codes
inside the same workflow!





GEMMLCA and Production Grids

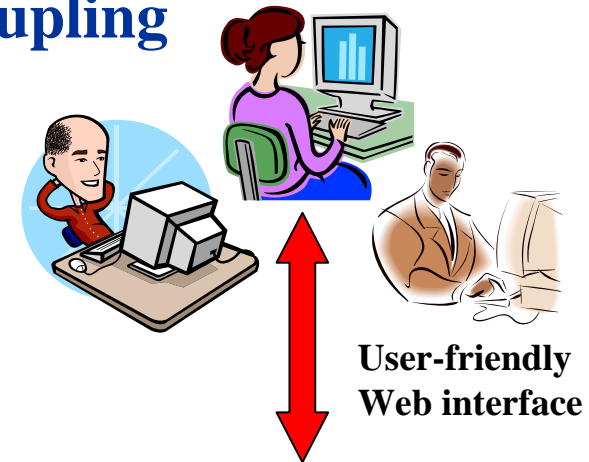
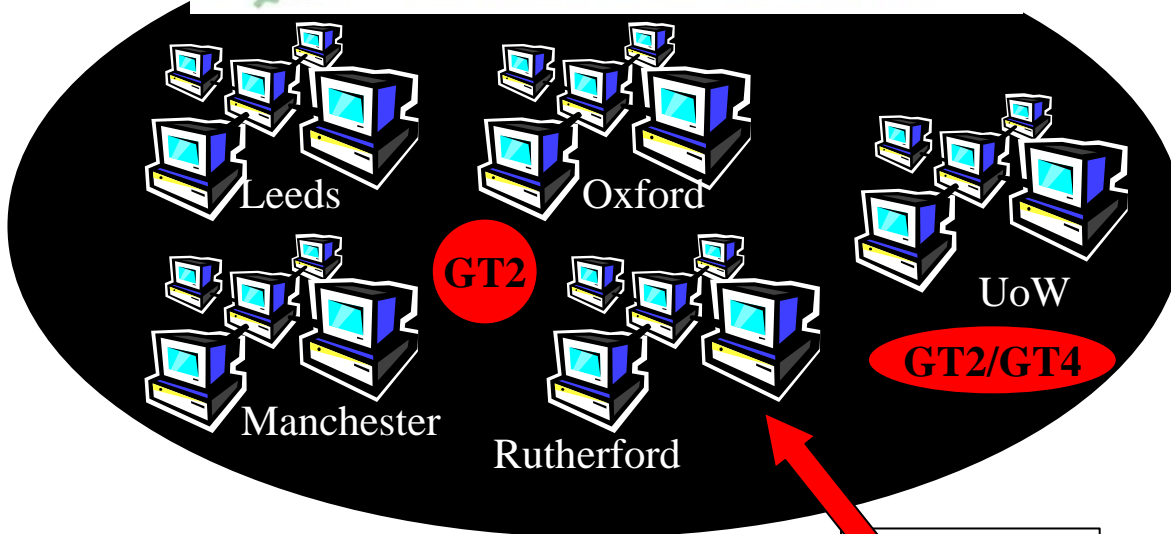
The problem





The solution – GEMMLCA decoupling

NGS National Grid Service
core production computational and data grid

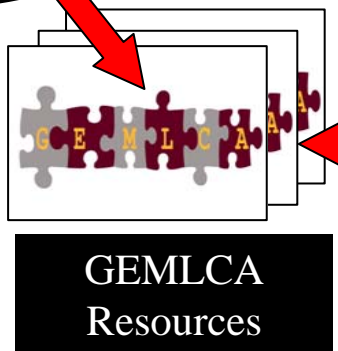


P-GRADE portal



Portal server (at UoW)

Run as a third party services





GEMMLCA on the UK NGS

The P-GRADE NGS GEMMLCA Portal

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





GEMMLCA 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 GEMMLCA
- Connected to and interoperable with the UK NGS



The screenshot displays the WestFocus Grid portal interface. At the top, there are logos for the University of Westminster, WestFocus, and GridAlliance. Below the navigation menu, there is a 'Workflow Editor' section with a table of workflows and a diagram of a workflow graph.

Workflow	Status	Size	Quota (100)
Brunel	init	241 KB	0.23%
EGEE_NGS_GT4_GEMMLCA_Man	init	6.887 MB	6%
GEMMLCA_NGS	init	5.138 MB	5%
GT2	init	7.741 MB	7%
NGS_Westfocus_GEMMLCA	init	6.074 MB	6%
WestFocus_GEMMLCA	init	4.395 MB	4%
		30.472 MB	

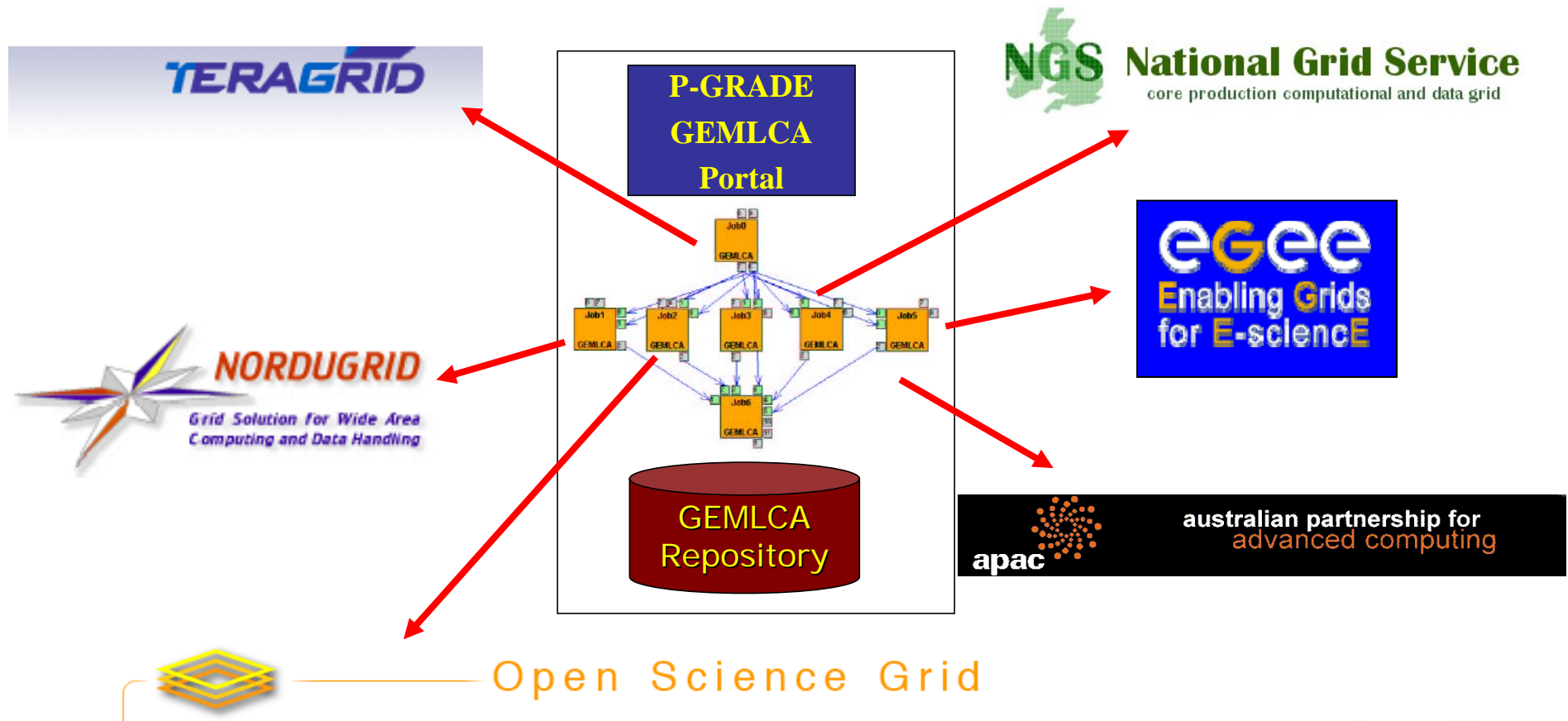
Message: Press a button.
17 May 2006





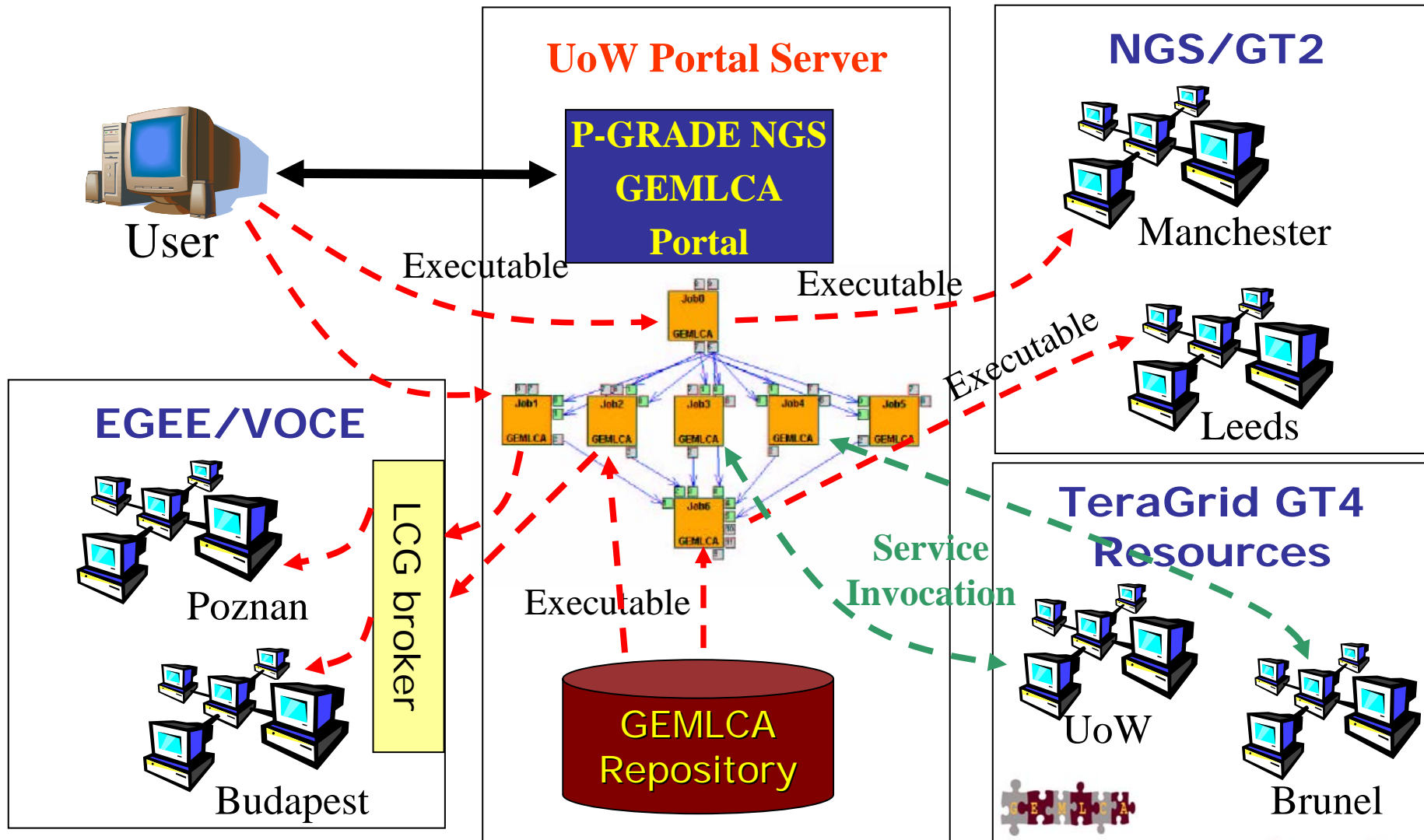
The GIN Resource Testing portal

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



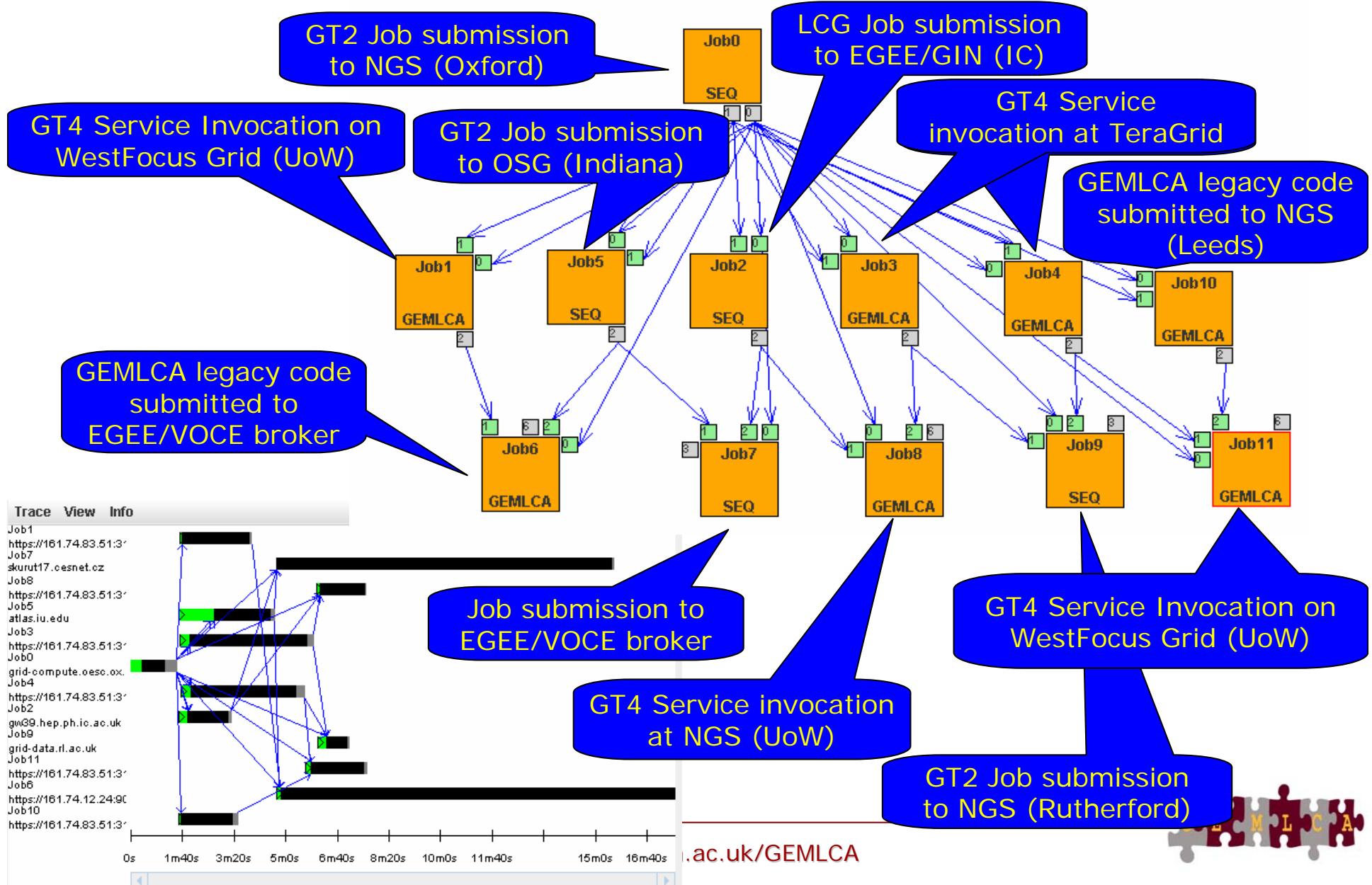


Connecting GT2, GT4 and LCG/g-Lite based Grids



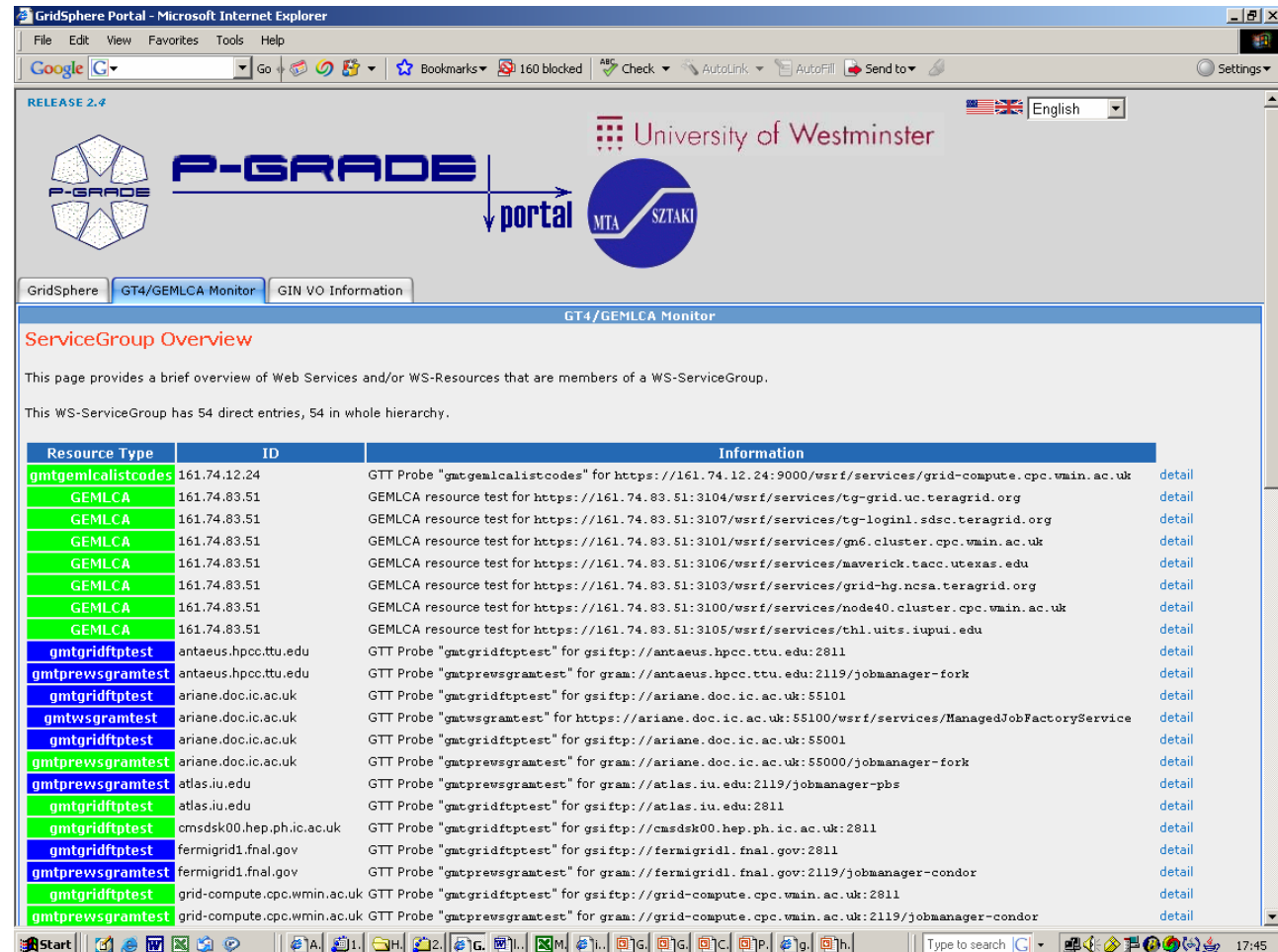


Traffic simulation on multiple Grids



GMT – GEMMLCA 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



GridSphere Portal - Microsoft Internet Explorer

RELEASE 2.4

University of Westminster

GridSphere | **GT4/GEMMLCA Monitor** | GIN VO Information

GT4/GEMMLCA Monitor

ServiceGroup Overview

This page provides a brief overview of Web Services and/or WS-Resources that are members of a WS-ServiceGroup.

This WS-ServiceGroup has 54 direct entries, 54 in whole hierarchy.

Resource Type	ID	Information	
gmtgemcalistcodes	161.74.12.24	GTT Probe "gmtgemcalistcodes" for https://161.74.12.24:9000/wsrf/services/grid-compute.cpc.wmin.ac.uk	detail
GEMMLCA	161.74.83.51	GEMMLCA resource test for https://161.74.83.51:3104/wsrf/services/tg-grid.uc.teragrid.org	detail
GEMMLCA	161.74.83.51	GEMMLCA resource test for https://161.74.83.51:3107/wsrf/services/tg-login1.sdsc.teragrid.org	detail
GEMMLCA	161.74.83.51	GEMMLCA resource test for https://161.74.83.51:3101/wsrf/services/gm6.cluster.cpc.wmin.ac.uk	detail
GEMMLCA	161.74.83.51	GEMMLCA resource test for https://161.74.83.51:3106/wsrf/services/maverick.tacc.utexas.edu	detail
GEMMLCA	161.74.83.51	GEMMLCA resource test for https://161.74.83.51:3103/wsrf/services/grid-hg.ncsa.teragrid.org	detail
GEMMLCA	161.74.83.51	GEMMLCA resource test for https://161.74.83.51:3100/wsrf/services/node40.cluster.cpc.wmin.ac.uk	detail
GEMMLCA	161.74.83.51	GEMMLCA 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

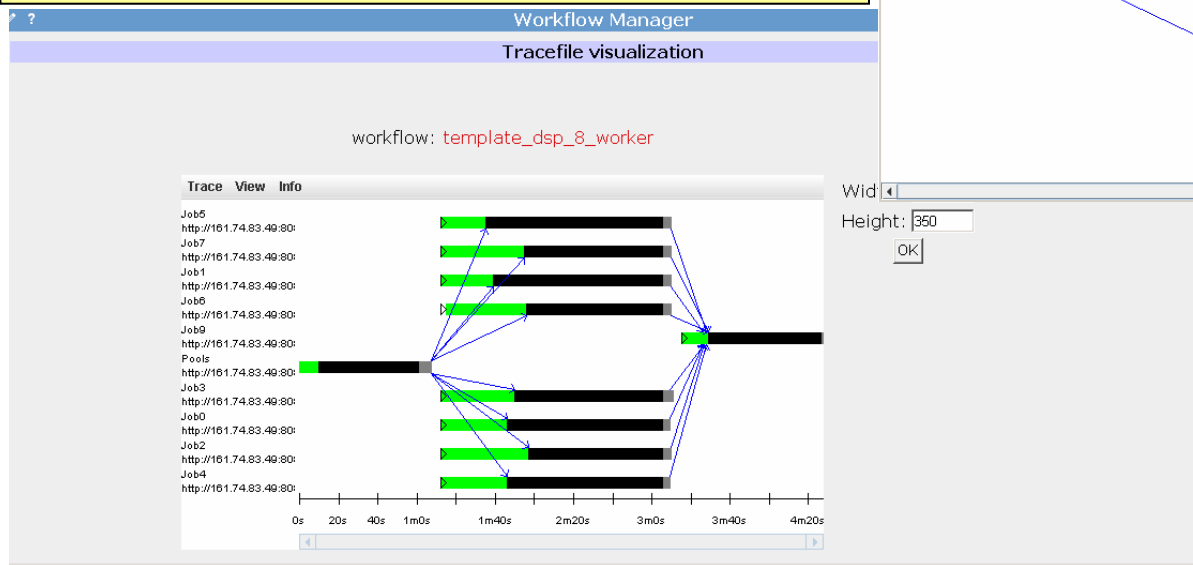
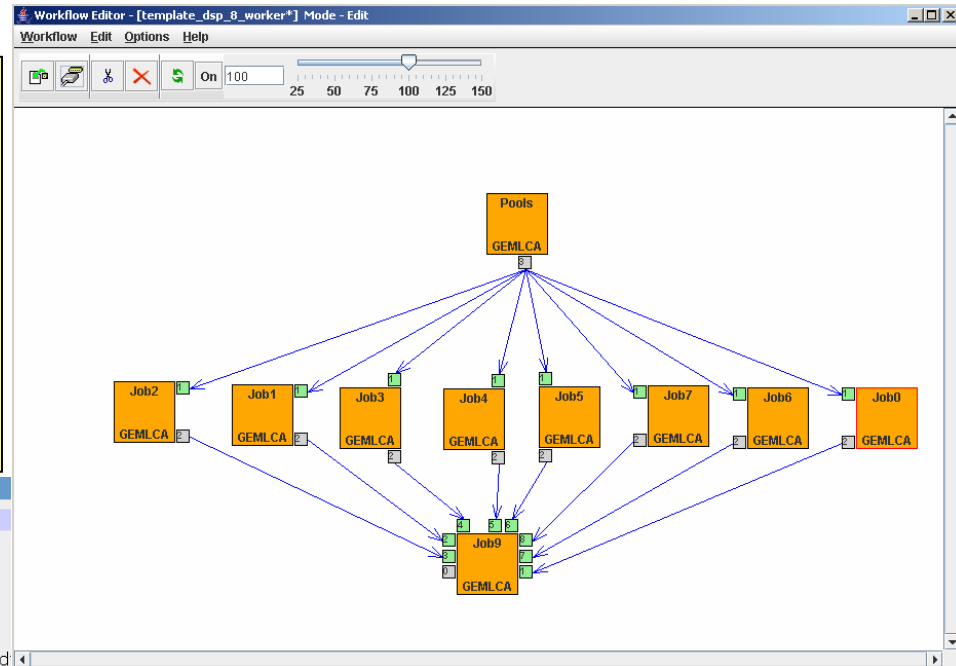




Application examples

DSP-Designing Optimal Periodic Nonuniform Sampling Sequences

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

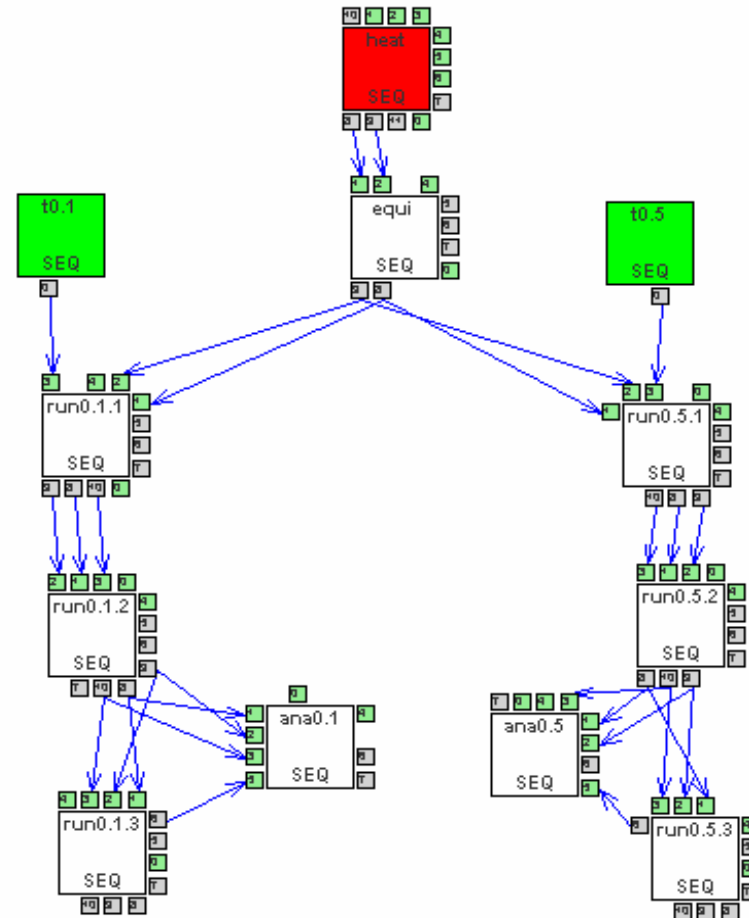
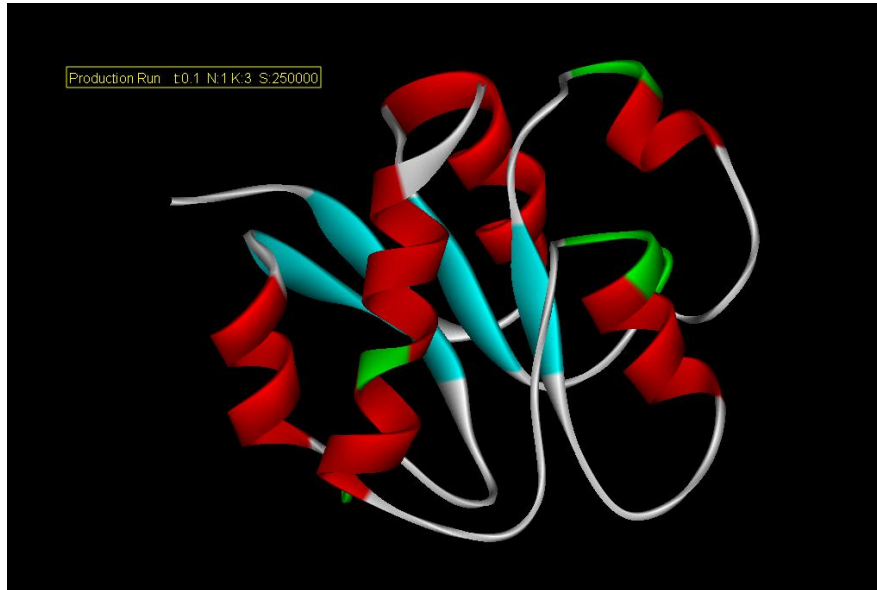




Application examples

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





Thank you for your attention!

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

gemlca-discuss@cpc.wmin.ac.uk

