



Exposing Legacy Applications through a Generic Grid Service Interface



Grid Execution Management for Legacy Code Applications

Tamas Kiss
kisst@wmin.ac.uk

Joint EGGE/SEE-Grid Summer School
Budapest, Hungary, 28 June 2007



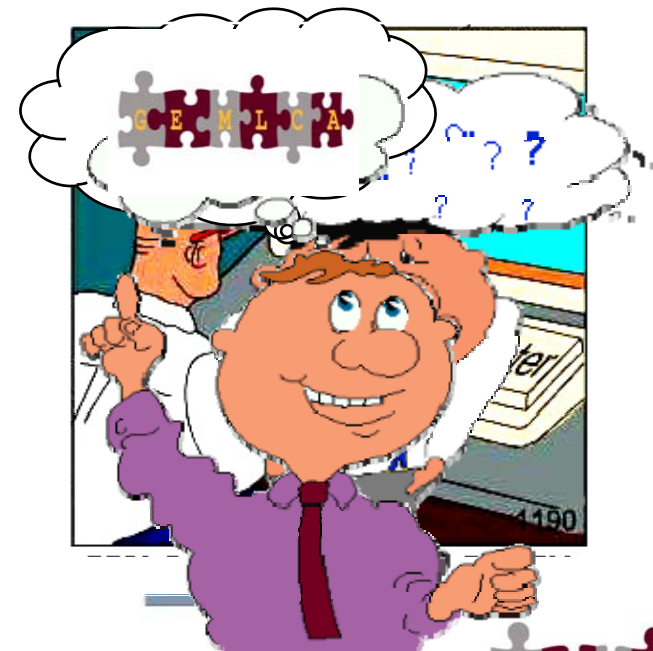


Legacy Applications

- 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





Grid Execution Management for Legacy Code Applications

University of Westminster, London





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

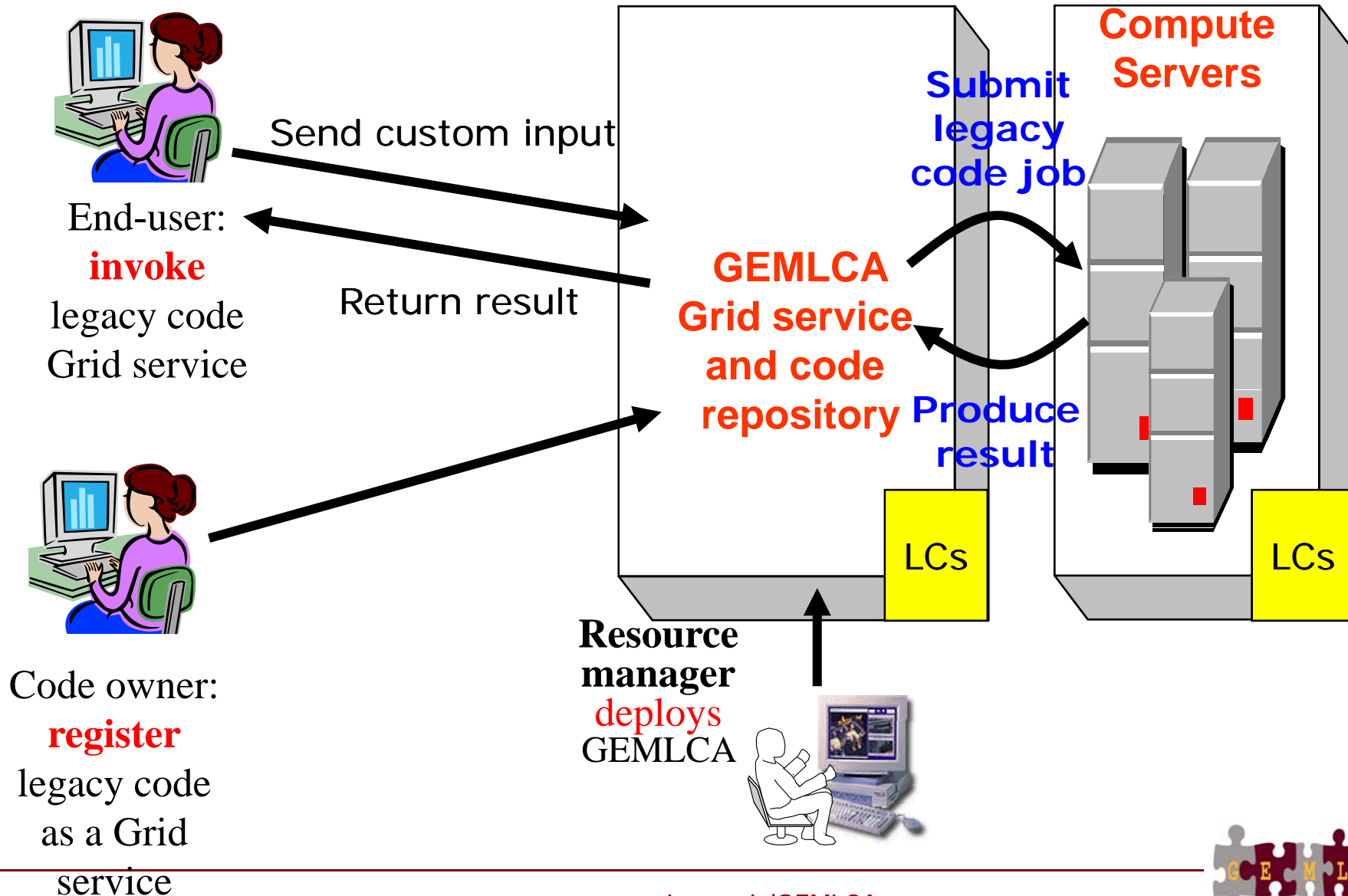
GEMMLCA

**GEMMLCA
PGPortal
Integration**





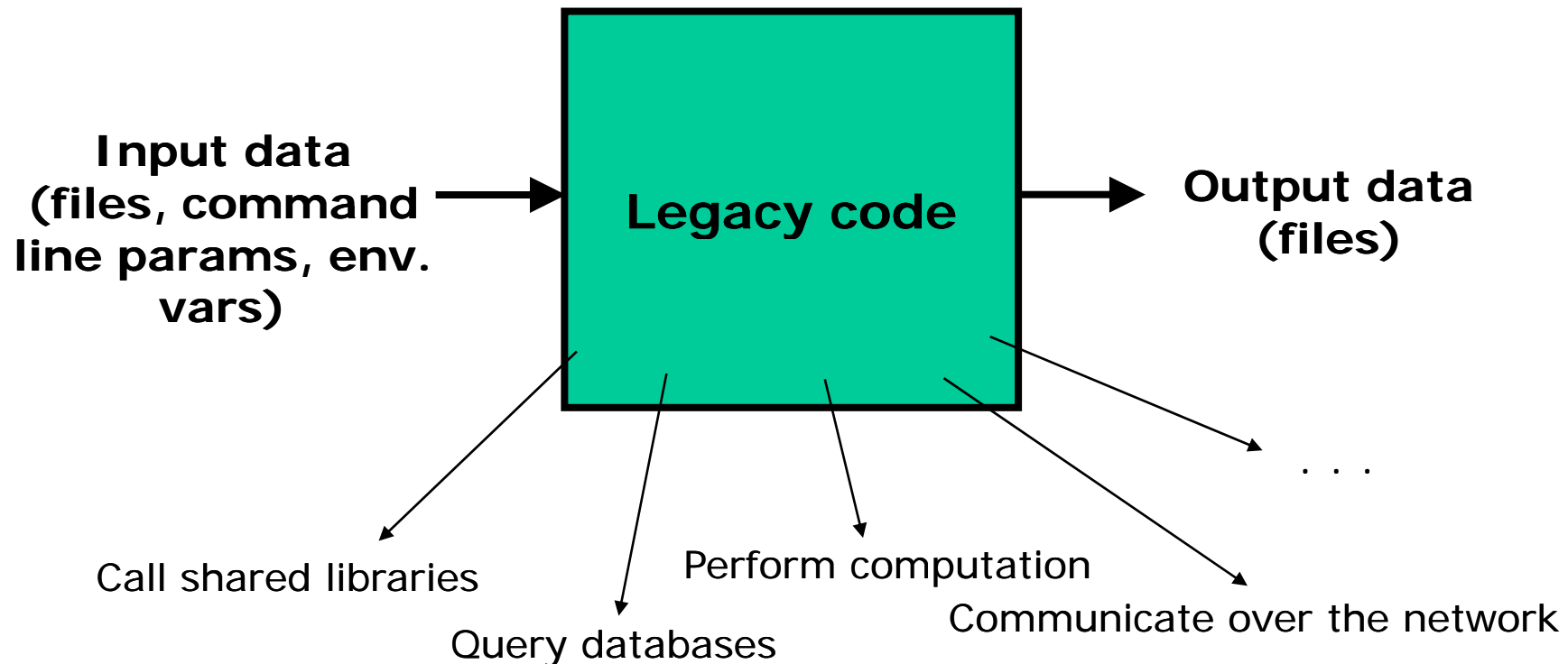
GEMMLCA Concept





The GEMMLCA-view of a legacy code

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





Implementing the concept

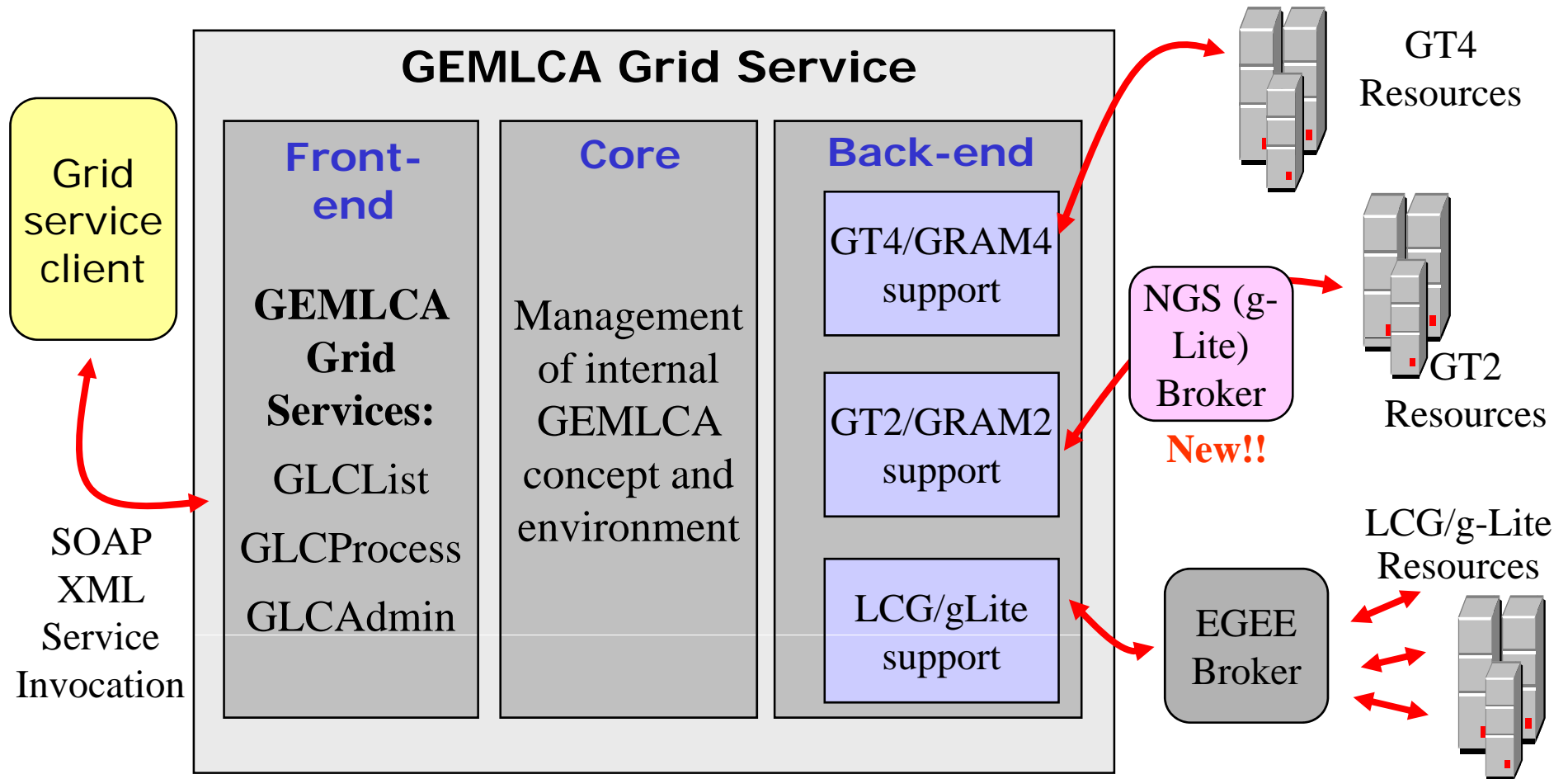
- The GEMMLCA service can be implemented with any grid/service-oriented technology E.g:
 - Globus (3 or) 4 → **currently available implementations**
 - Jini
 - Web services
 - ...
- GEMMLCA service could invoke legacy codes in many different ways. Current implementation:
 - **Submit the legacy code as 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)**





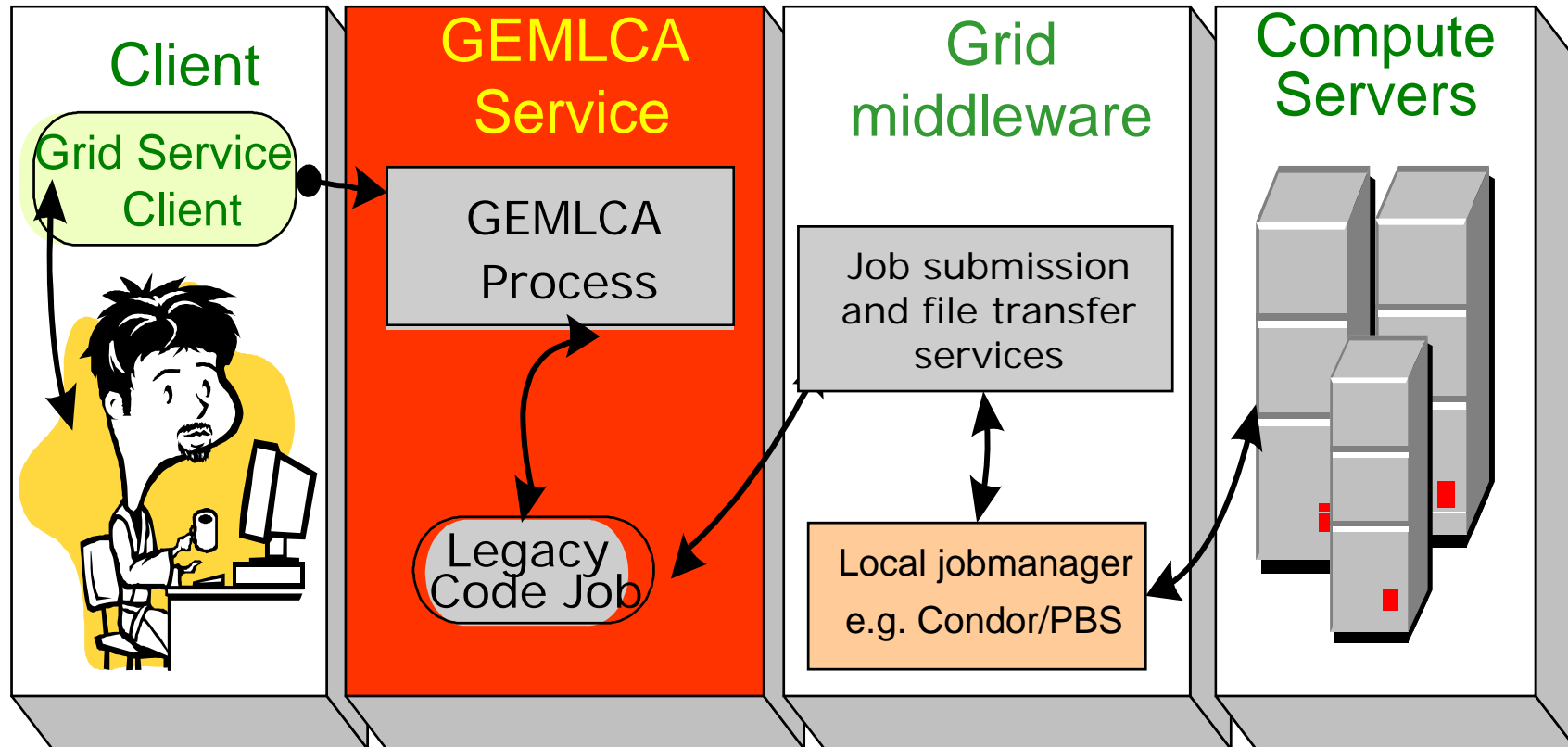
Implementing the GEMMLCA concept

The centralised and decoupled GEMMLCA architecture





How GEMMLCA works?





What's the point?

- **Heterogeneous codes can be hidden behind the same interface**
 - Different programs can be invoked in the same way using SOAP/XML communication only
- **Extend non grid-aware programs with security infrastructure** (access enabled through a Grid service)
- **Create and browse repositories of legacy applications**
 - Share your codes with your colleagues or partner institutes
 - Expose business logic to your employees or customers
- **Build customized GEMLCA clients** (such as the GEMLCA 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





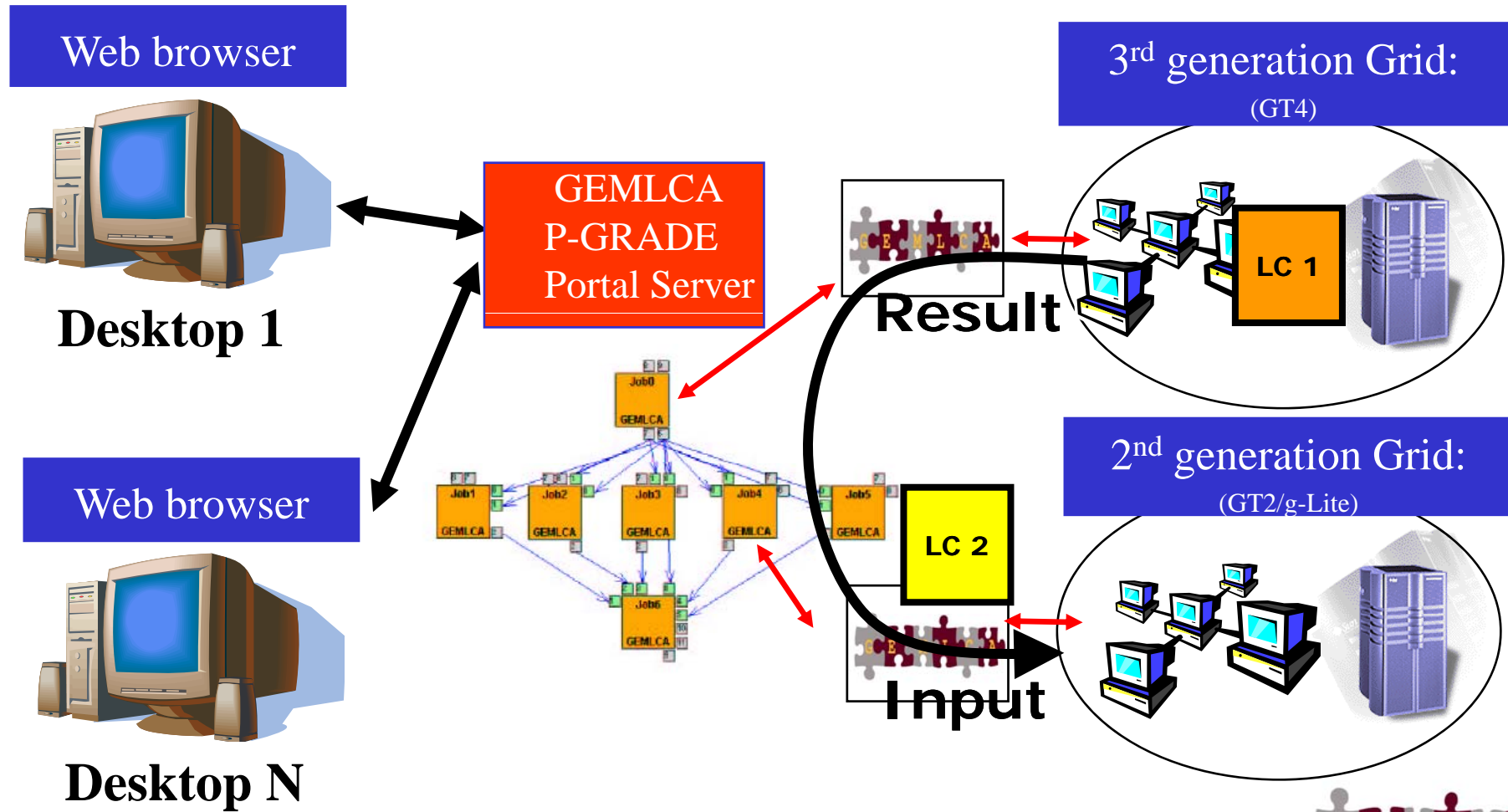
The aims of the GEMMLCA P-GRADE Portal

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





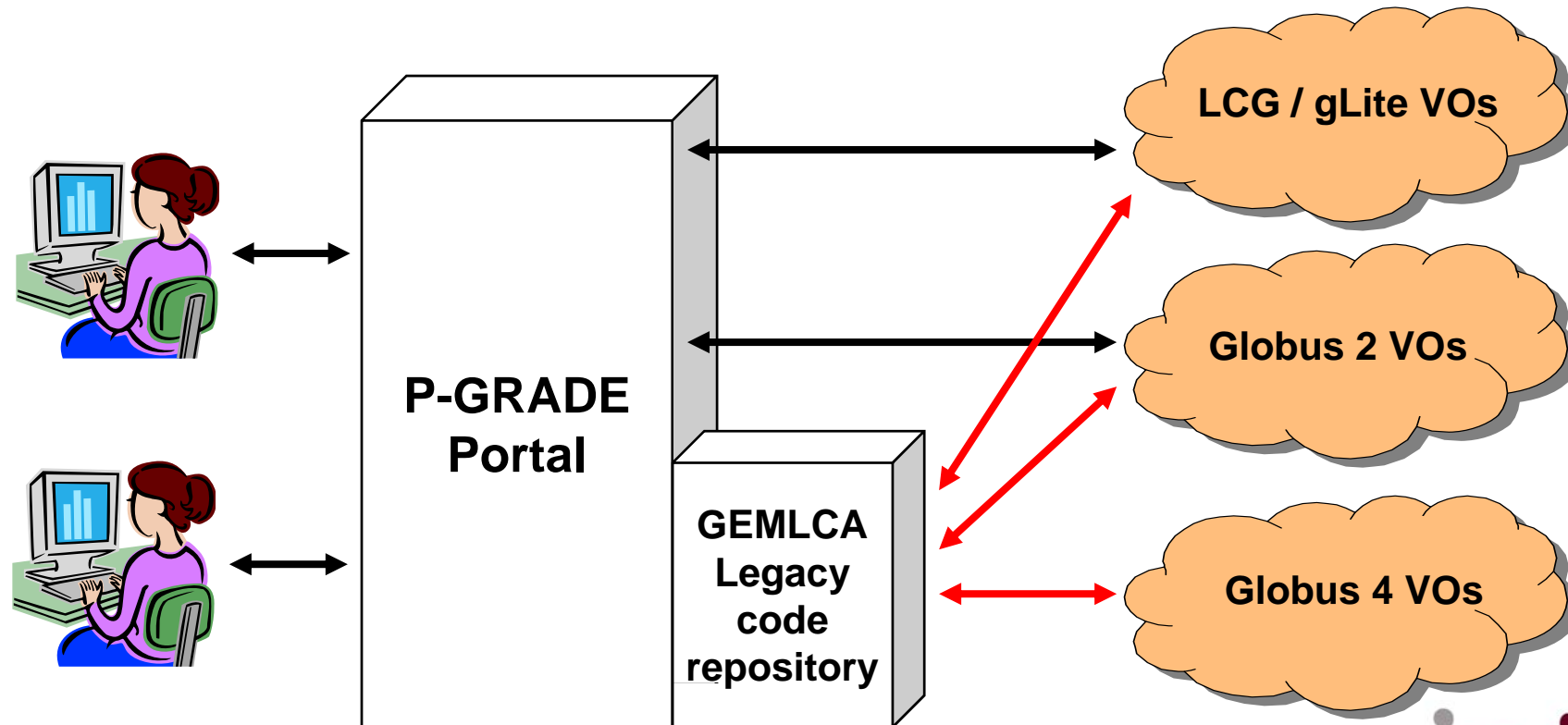
GEMMLCA in the P-GRADE Portal





How GEMMLCA extends the P-GRADE Portal

- P-GRADE Portal extended with GEMMLCA back-end
 - Sharing jobs and legacy codes as workflow components
- A step towards collaborative e-Science
- Support for Globus 4 grids (besides GT2 and EGEE)





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

The screenshot shows a web browser window with the following content:

- Navigation tabs: Workflow, Certificates, Settings, Demo, Help, GEMMLCA Administration Tools, Message...
- Breadcrumbs: Resource Selector > Legacy Code Information Descriptor Creator
- Page Title: GEMMLCA LCID Administration Portlet
- Section: GEMMLCA Legacy Code Interface Descriptor composer
- Section: Legacy code Environment Paramaters:
 - maximumProcessors: 1
 - executable: LINUX/mkdir
 - minimumProcessors: 1
 - maximumJob: 11
 - jobManager: Fork
 - id: mkdir
 - description: Unix mkdir program
 - Button: Set Parameters
- Section: List of legacy code Arguments:

name	file	order	fixed	inputOutput	mandatory	regexp	friendlyName	commandline	initialValue
------	------	-------	-------	-------------	-----------	--------	--------------	-------------	--------------
- Section: New argument entry form:
 - name: -p
 - file: No
 - order: 0
 - fixed: No
 - inputOutput: Input
 - mandatory: No
 - regexp: (empty)
 - friendlyName: Folder to be created
 - commandline: Yes
 - initialValue: (empty)
 - Button: Add Argument

"GEMMLCA Administration Tool" portlet





Legacy code registration page

Legacy code Environment Parameters:

maximumProcessors
 executable
 minimumProcessors
 maximumJob
 jobManager
 id
 description

List of legacy code Arguments:

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

New argument entry form:

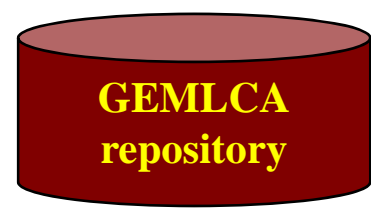
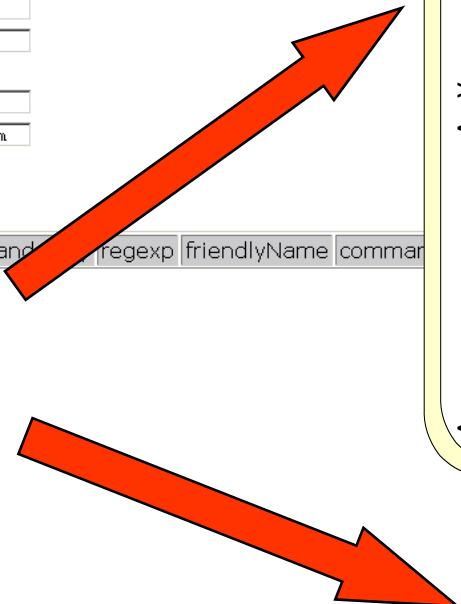
name
 file
 order
 fixed
 inputOutput
 mandatory
 regex
 friendlyName
 commandline
 initialValue

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/wsrp/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

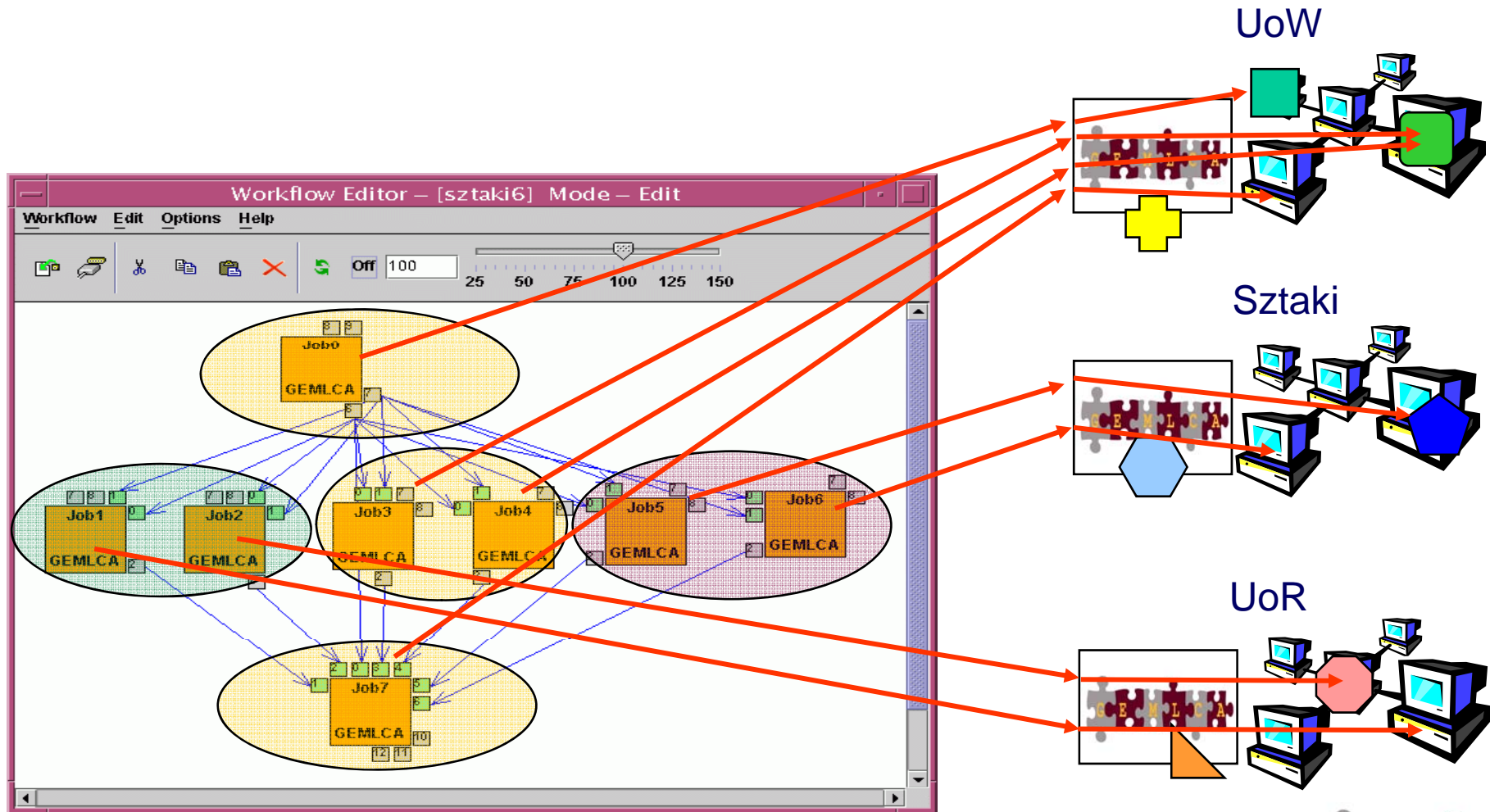
Ok Cancel





GEMMLCA workflow editor in a nutshell

Workflow Creation





Batch components vs. 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

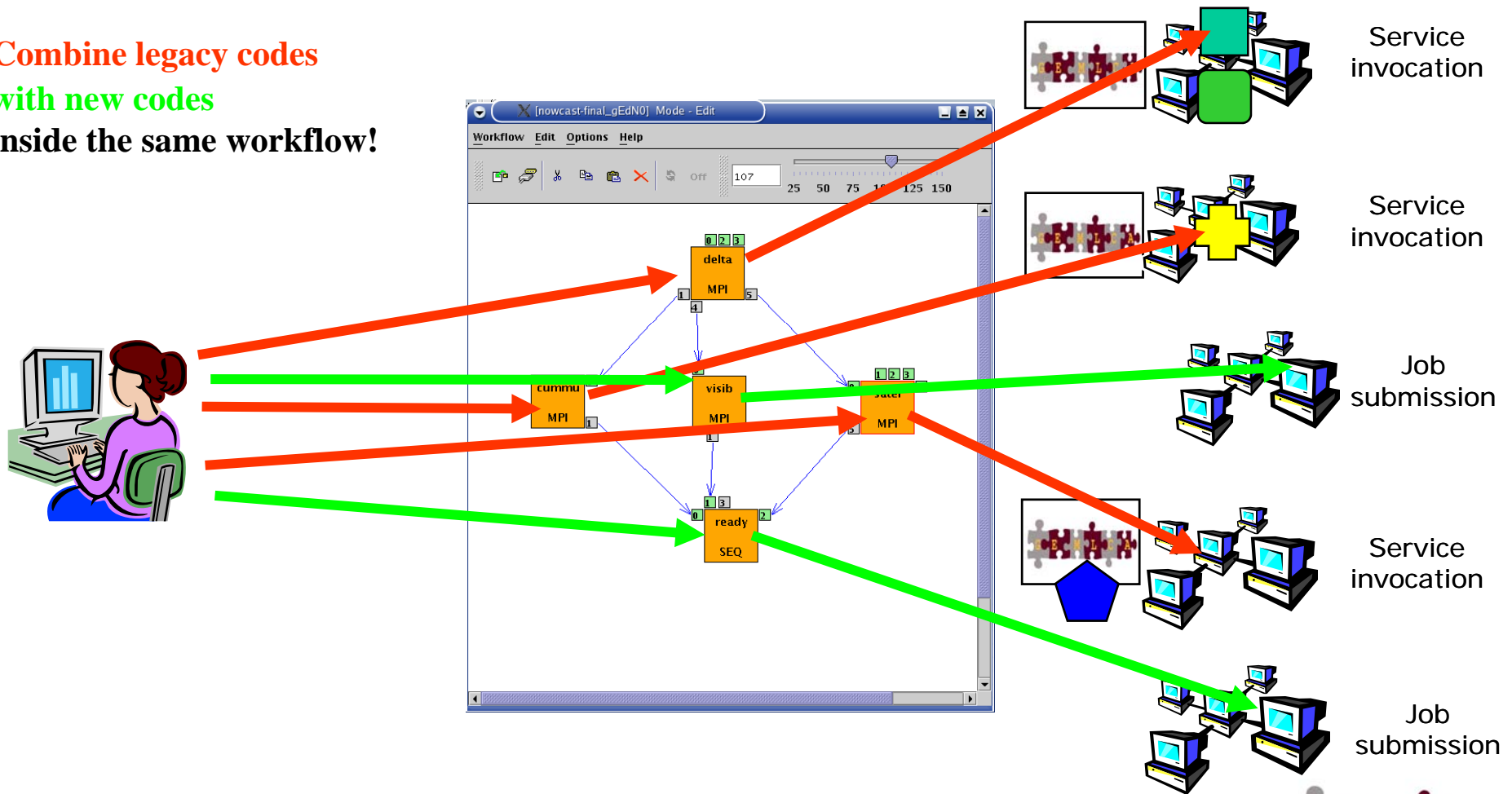
Ports guarantee compatibility → batch and GEMLCA components can mutually produce data to each other!





Combining legacy and non-legacy components

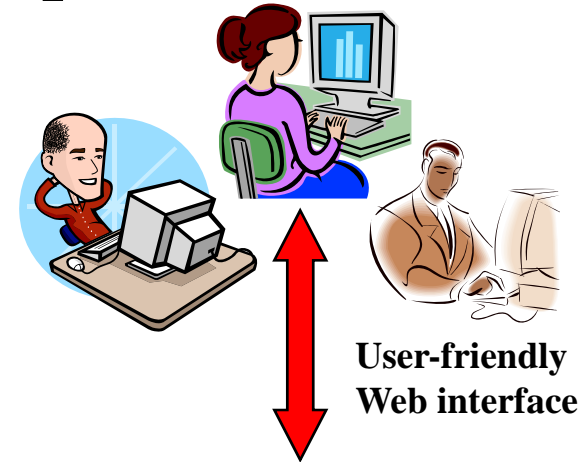
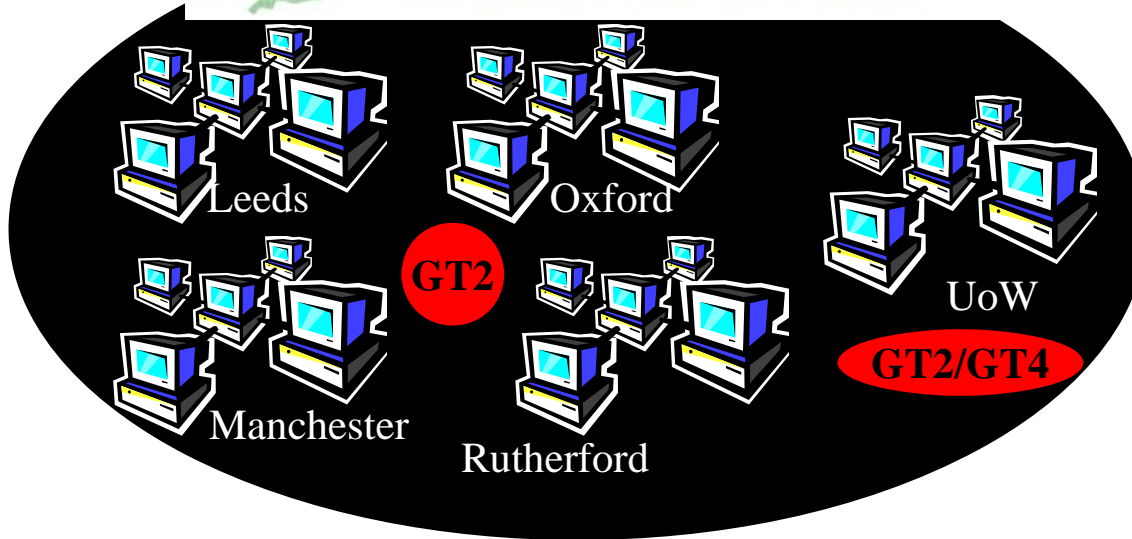
Combine legacy codes
with new codes
inside the same workflow!





The GEMMLCA P-GRADE portal on production Grids

NGS National Grid Service
core production computational and data grid



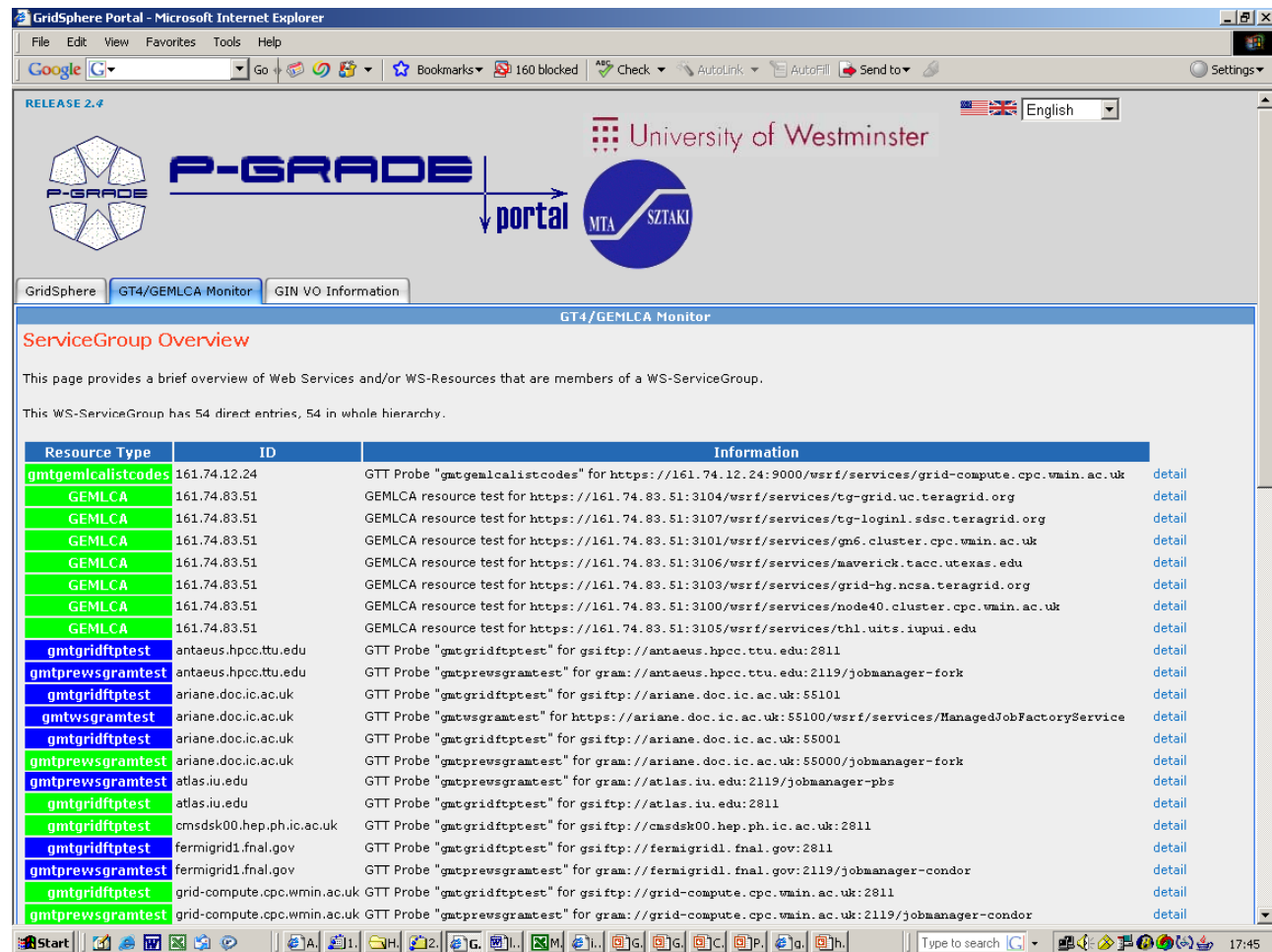
Portal server (at UoW)

Run as third party services
Decoupled from portal and Grid middleware

GMT – GEMMLCA Monitoring Toolkit

Resources need to be constantly monitored in production environments

- 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-loginl.sdsc.teragrid.org	detail
GEMMLCA	161.74.83.51	GEMMLCA resource test for https://161.74.83.51:3101/wsrf/services/gn6.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.iu.uts.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



NGS P-GRADE GEMMLCA Portal

- **portal website:**
<https://gngs-portal.cpc.wmin.ac.uk:8080/gridsphere/gridsphere>
- Interface for NGS GT2 sites
- Interface for GT4 Westminster site
- Interface for EGEE GILDA sites
- Connected to the NGS and the GILDA Resource Brokers

NGS P-Grade GEMMLCA portal

Welcome to NGS P-Grade Portal!

This Portal provides access to the UK National Grid Service. It is a workflow-oriented Grid portal that enables the creation, execution and monitoring of computational workflows in Grid environments through user-friendly graphical Web interfaces.

To get an account for this Portal and find further information and tutorials please visit <http://www.cpc.wmin.ac.uk/ngsportal/241/>.

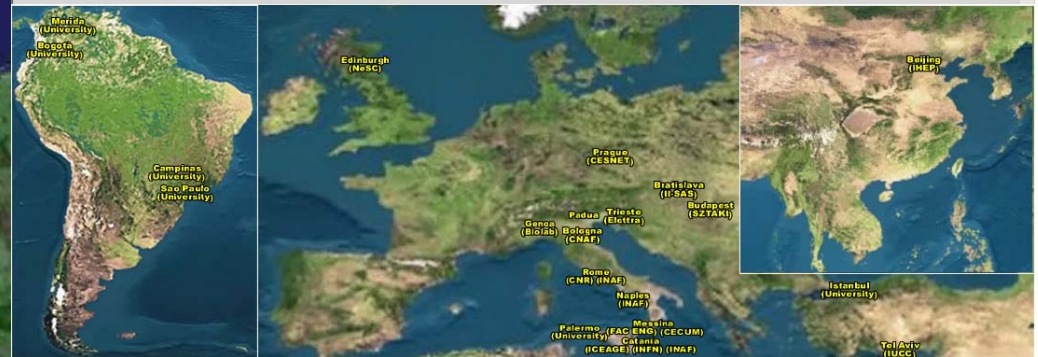
If you need any assistance please e-mail to the support team: noam@cpc.wmin.ac.uk

This portal works under an academic license and hence:
"You can use this portal under the condition that you only perform non-commercial research and education activities. By using your account on this portal you hereby warrant that you understand that contravention of this clause may result in a liability arising for historic, current and future license payments under the terms of the prevailing commercial license that was in effect during the time period of the contravention. If you are in any doubt please contact portalreq@ipis.sztaki.hu."

This portal is based on P-Grade Grid Portal technology, which was developed by the Laboratory of Parallel and Distributed Systems at MTA-SZTAKI, Hungary. The P-Grade Portal is a workflow-oriented Grid portal that enables the creation, execution and monitoring of computational workflows in Grid environments through high-level, graphical Web interfaces. The P-Grade Portal serves various Grid Communities in research and industry.

This site is built with P-Grade Portal version 2.4.1 software. This Portal has been tested with Mozilla 1.6, Netscape (4.x, 6 and higher) and Internet Explorer (5 or 6 - but may behave strangely with IE7) with JRE 1.4.2_x and 1.5.0_06 Java plug in. Using other versions of Web browsers or Java Virtual Machines may lead to visualization problems.

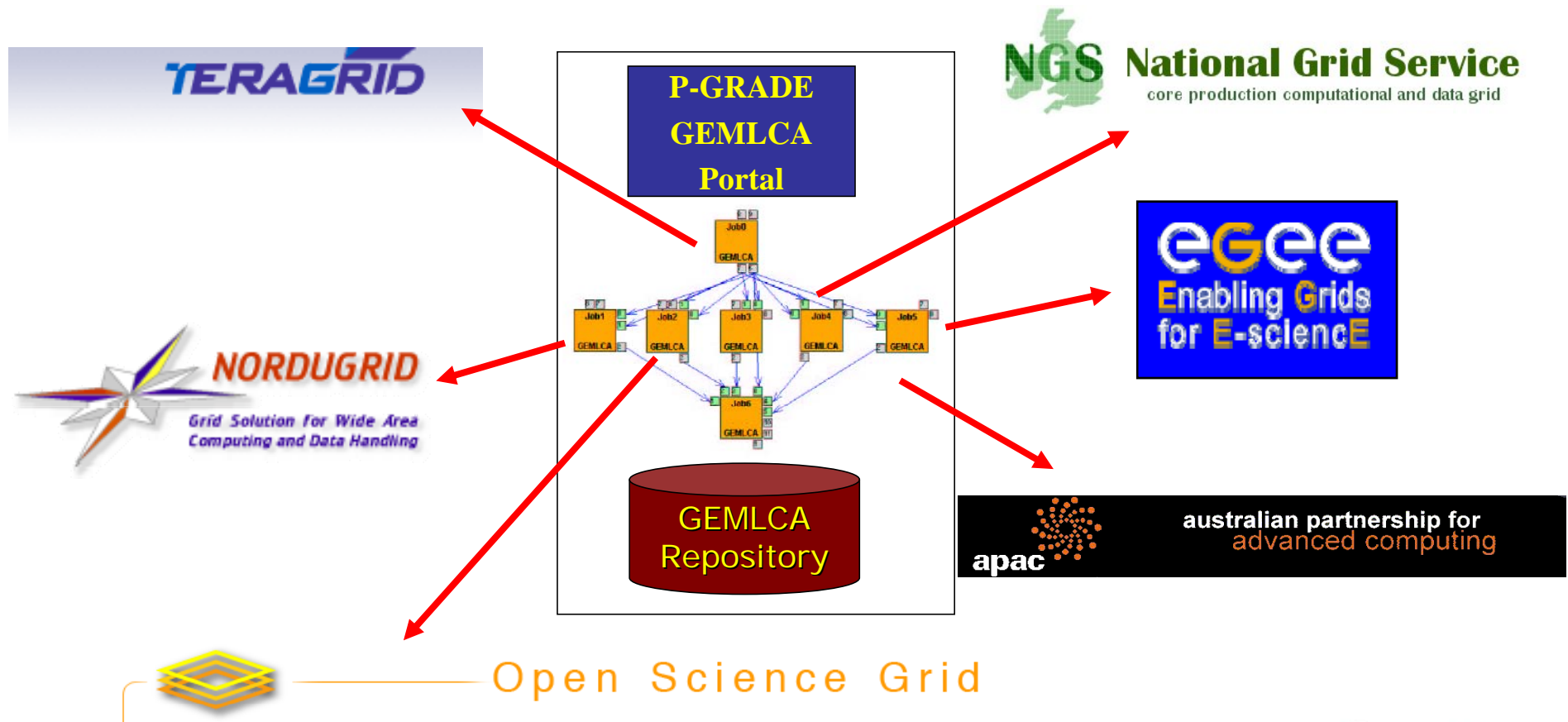
The portal has been continuously developed and hence any feed-back, comment advice, request concerning the portal features is highly appreciated. Please, send your comments to noam@cpc.wmin.ac.uk





The GIN Resource Testing portal

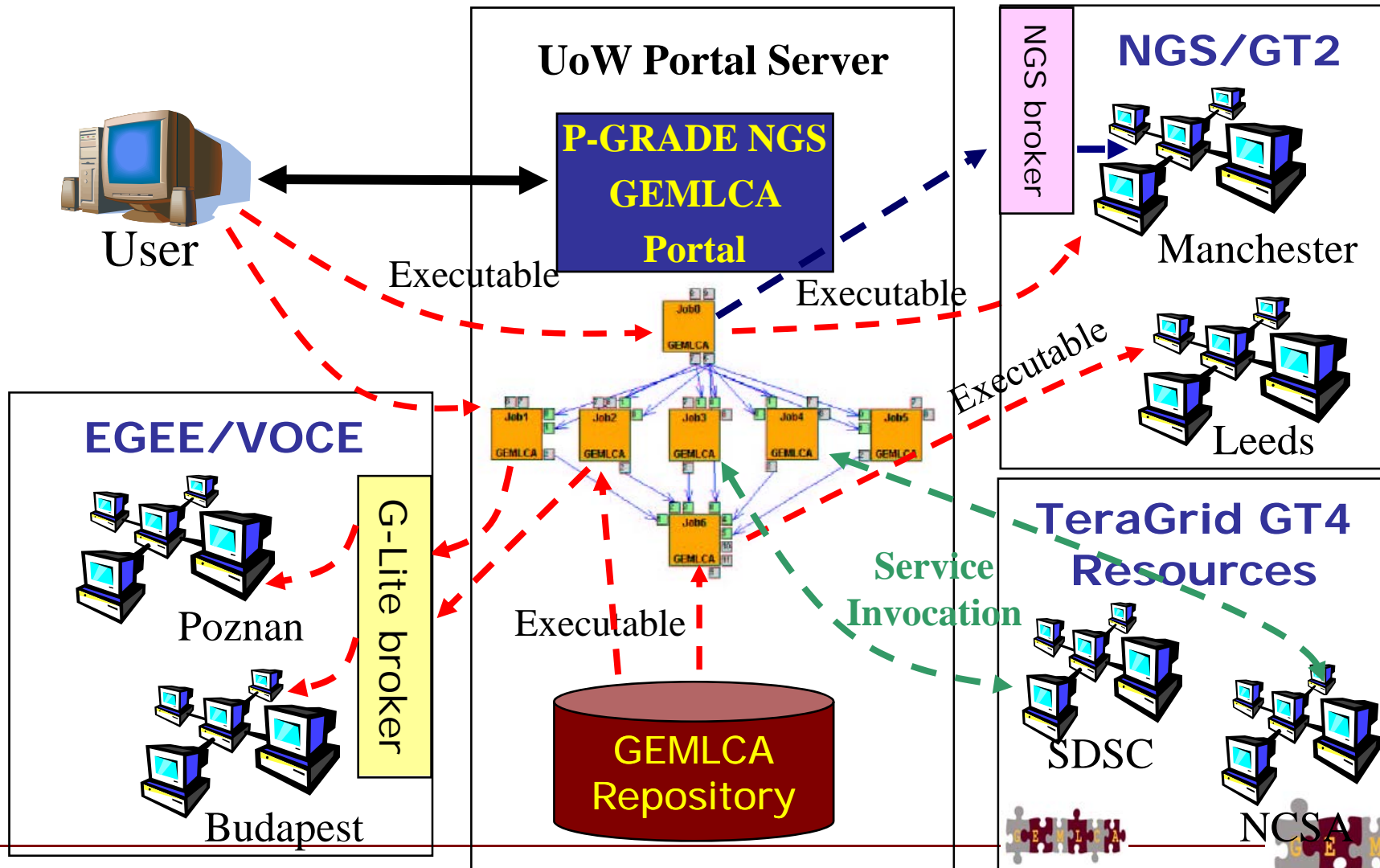
OGF effort to demonstrate workflow level interoperability between major production Grids and monitor GIN resources





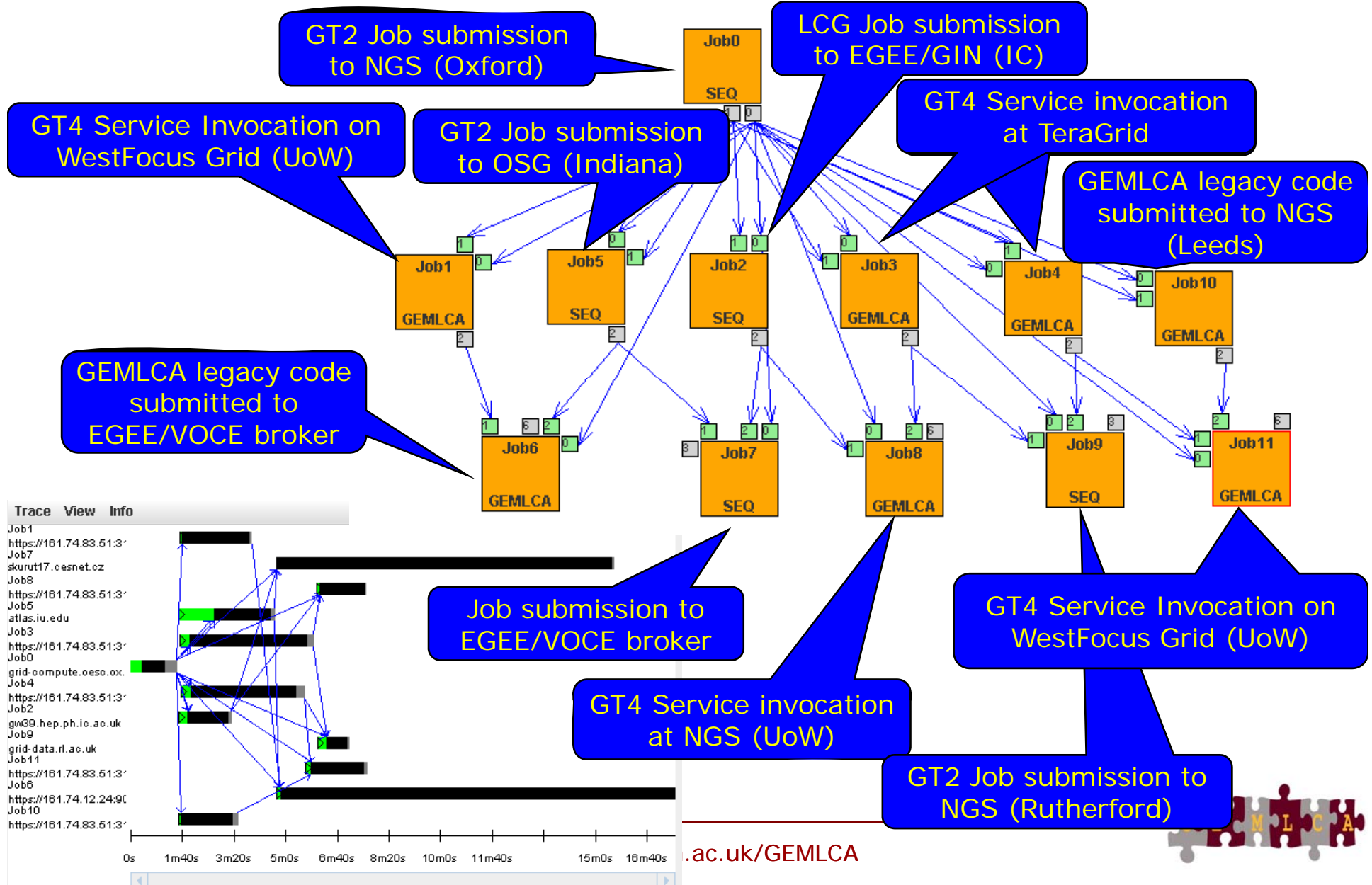
Workflow level interoperability of Grid systems

New!!





Traffic simulation on multiple Grids

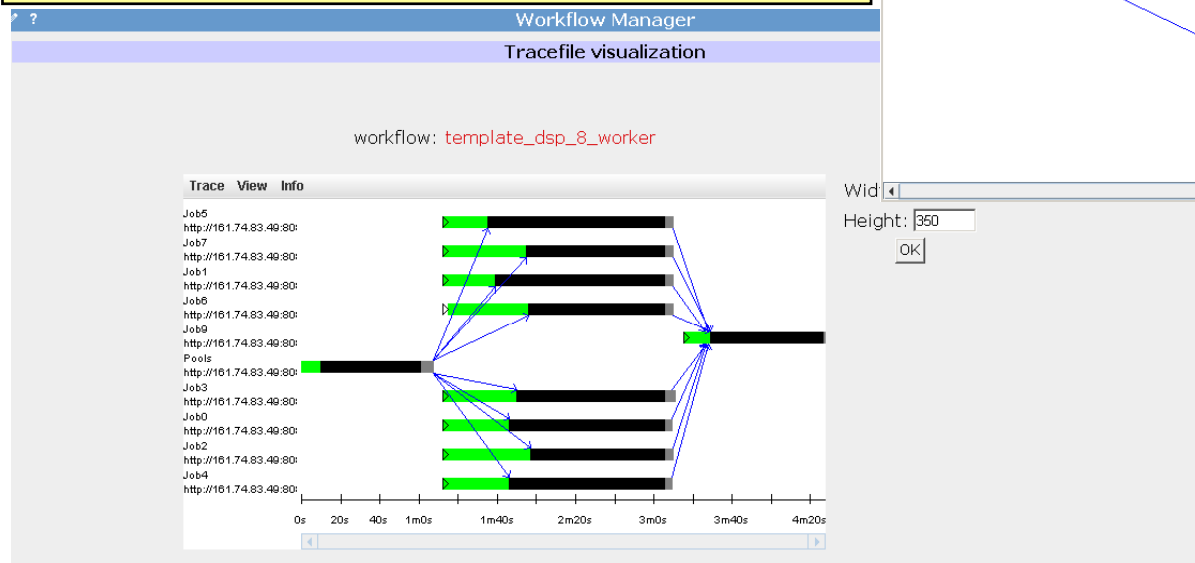
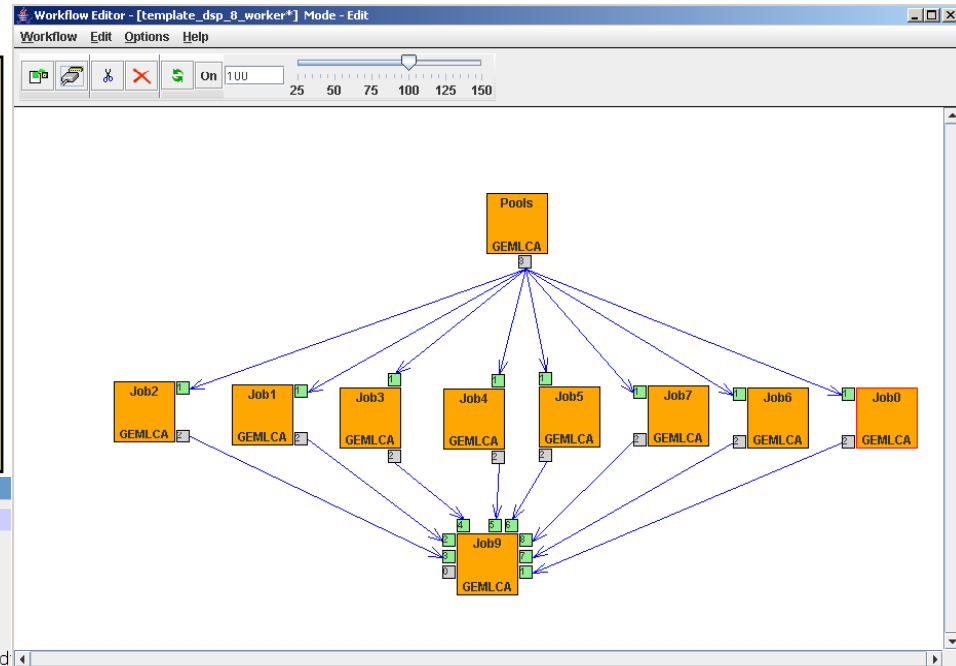




Application examples

DSP-Designing Optimal Periodic Nonuniform Sampling Sequences

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

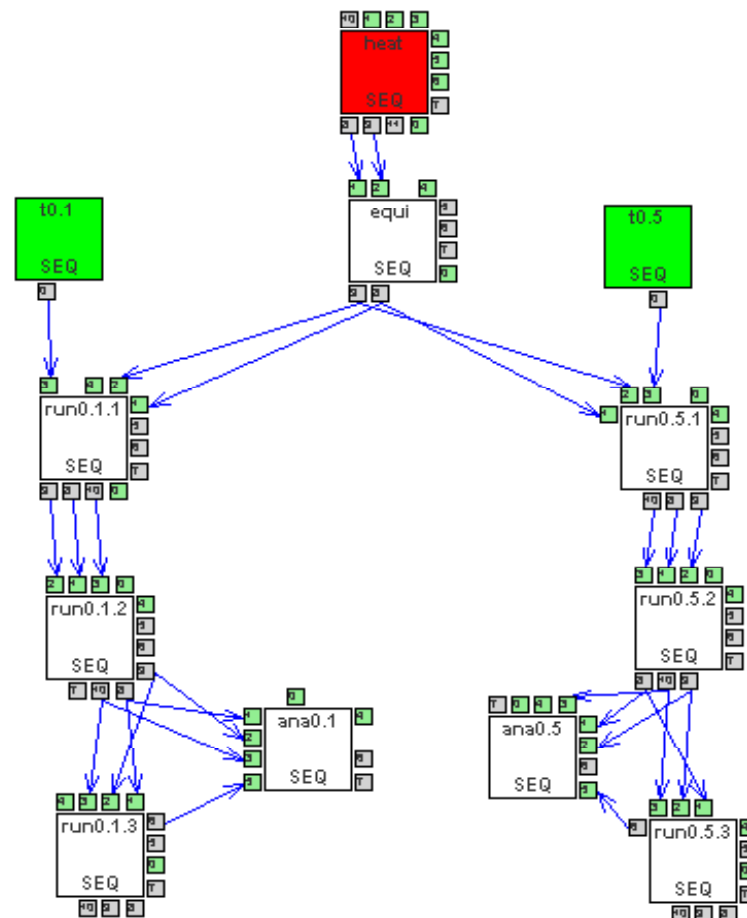
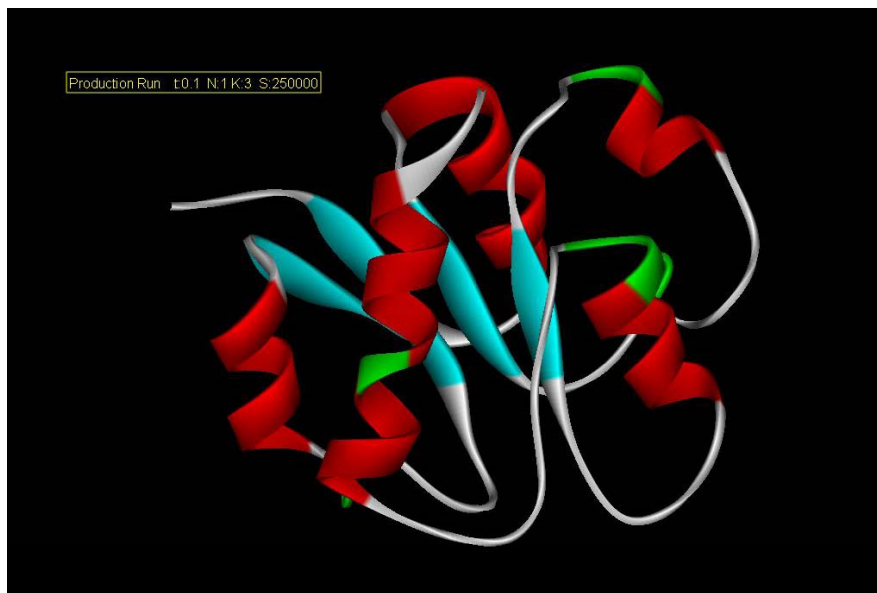




Application examples

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

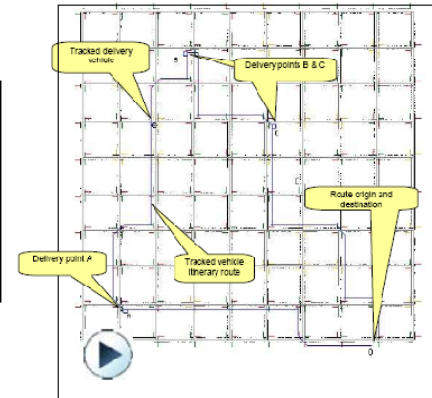




Application Specific Portals - The Traffic Portal

Only essential/related portlets displayed

Workflow creation and structure is hidden from the user



Application specific visualisation is embedded

Specifically tailored for one specific application - transport and logistics studies with a traffic simulator





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

