



Grid Enabling Legacy Applications

Grid Execution Management for Legacy Code Applications



Tamas Kiss
Centre for Parallel Computing
University of Westminster
London, UK



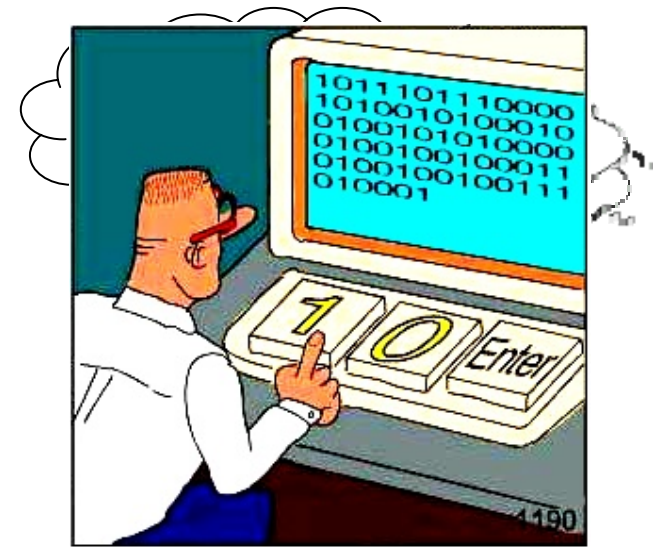


Legacy Applications

- 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

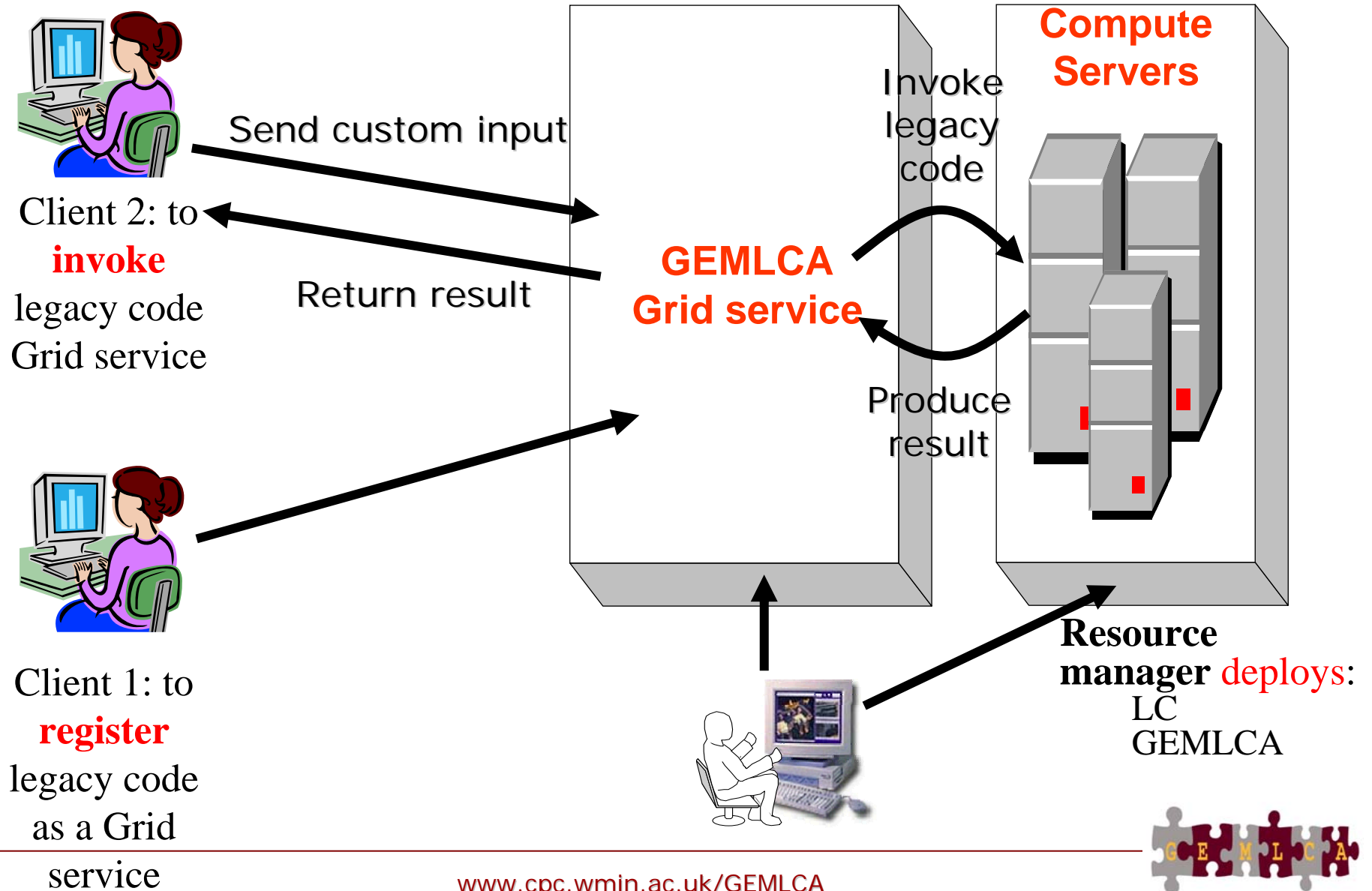
GEMMLCA

**GEMMLCA
PGPortal
Integration**





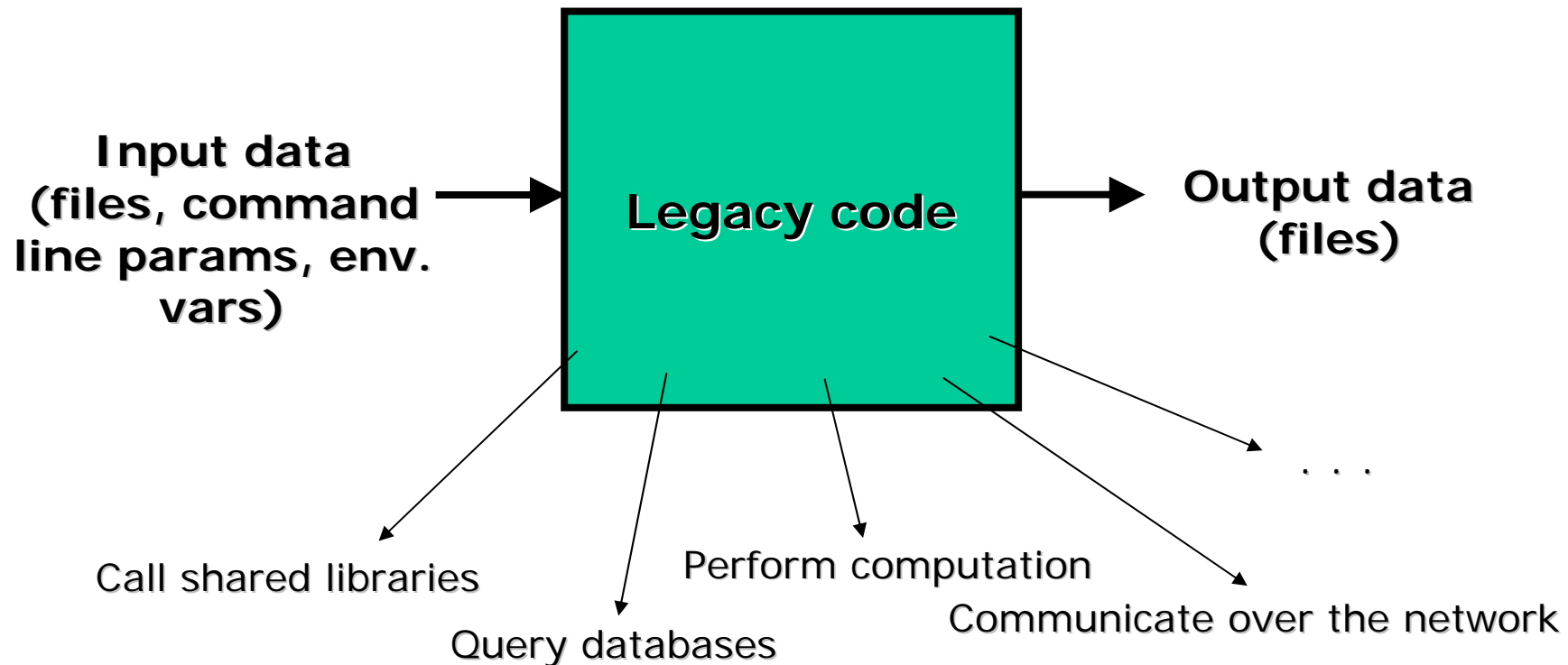
GEMMLCA Concept





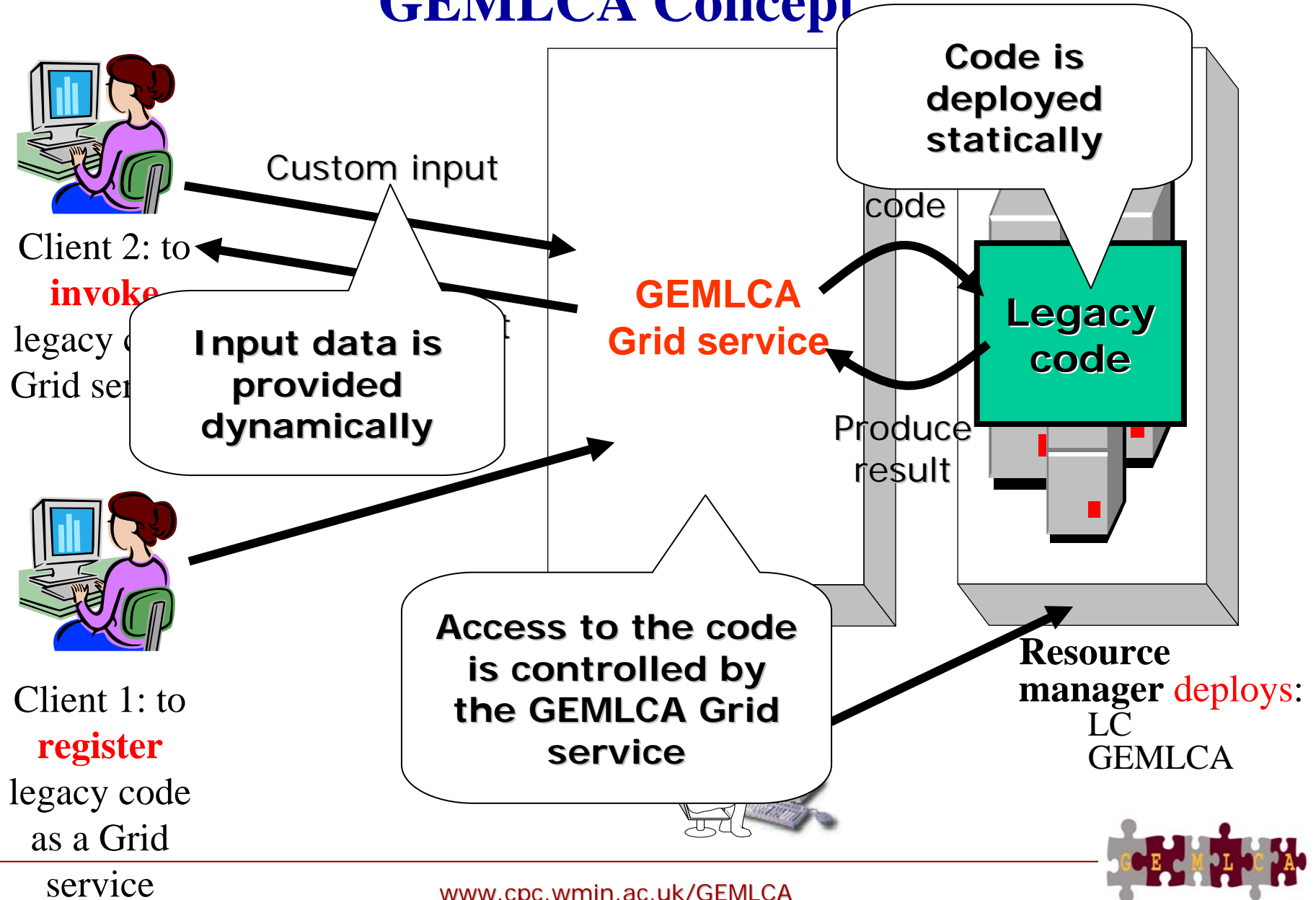
The GEMMLCA-view of a legacy code

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





GEMMLCA Concept





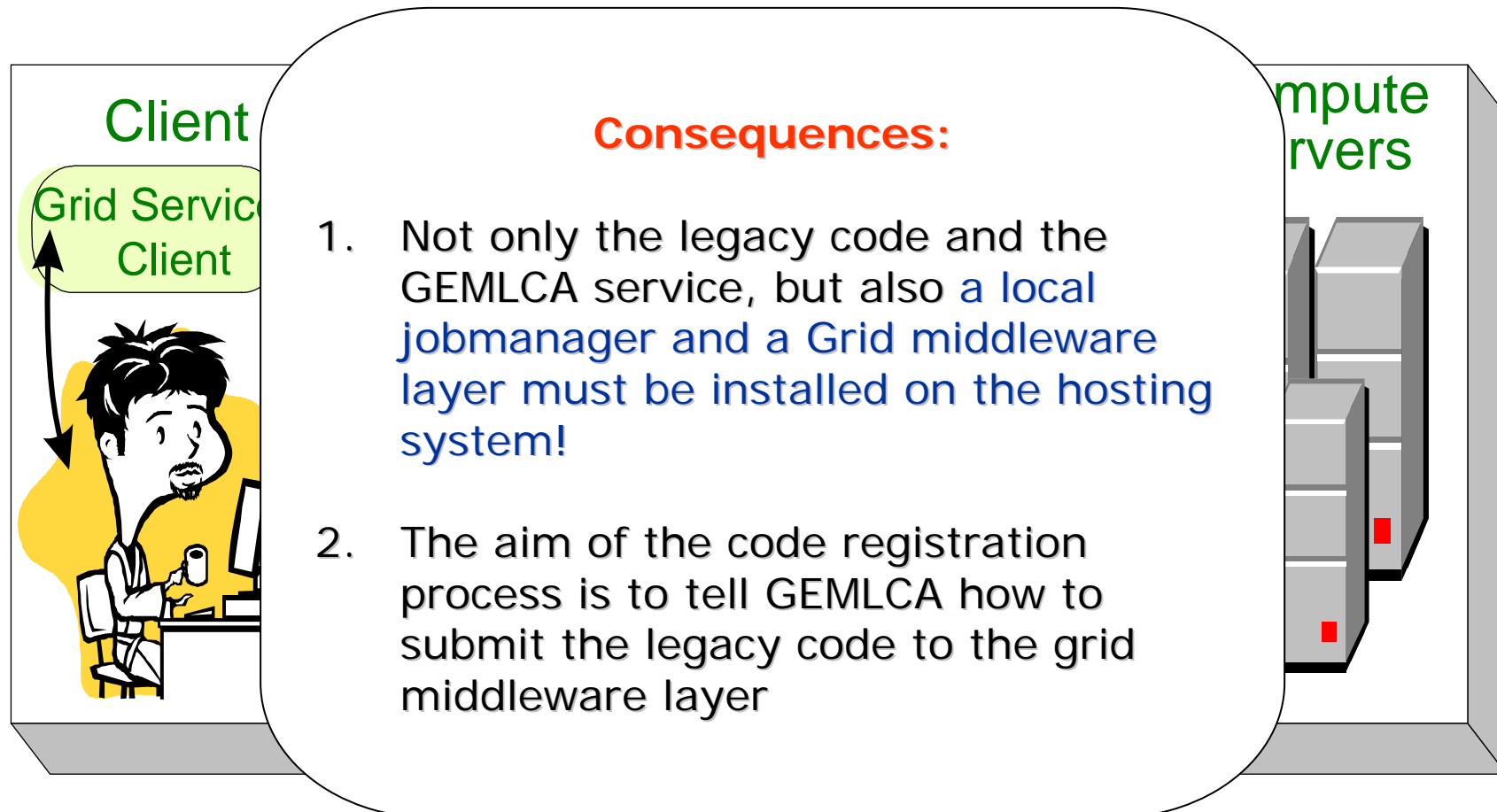
Implementing the concept

- 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 (GT3/4)**





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
- **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





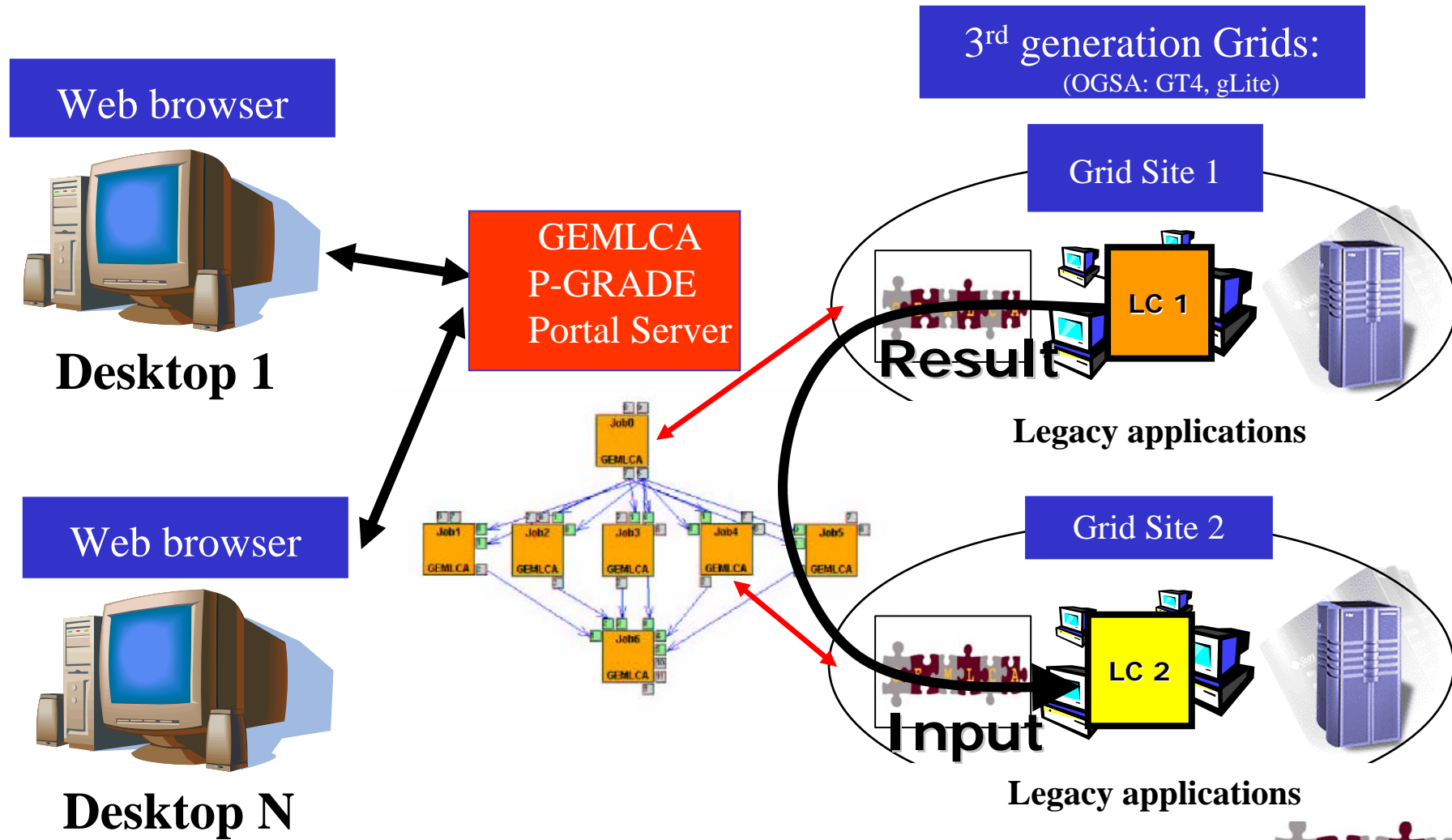
The aim of the GEMMLCA P-GRADE Portal

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





The GEMMLCA 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

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

- Navigation tabs: Workflow, Certificates, Settings, Demo, Help, GEMMLCA Administration Tools, Macroprocessor...
- 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:
 - friendlyName: Folder to be created
 - commandline: Yes
 - initialValue:
 - Button: Add Argument

"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

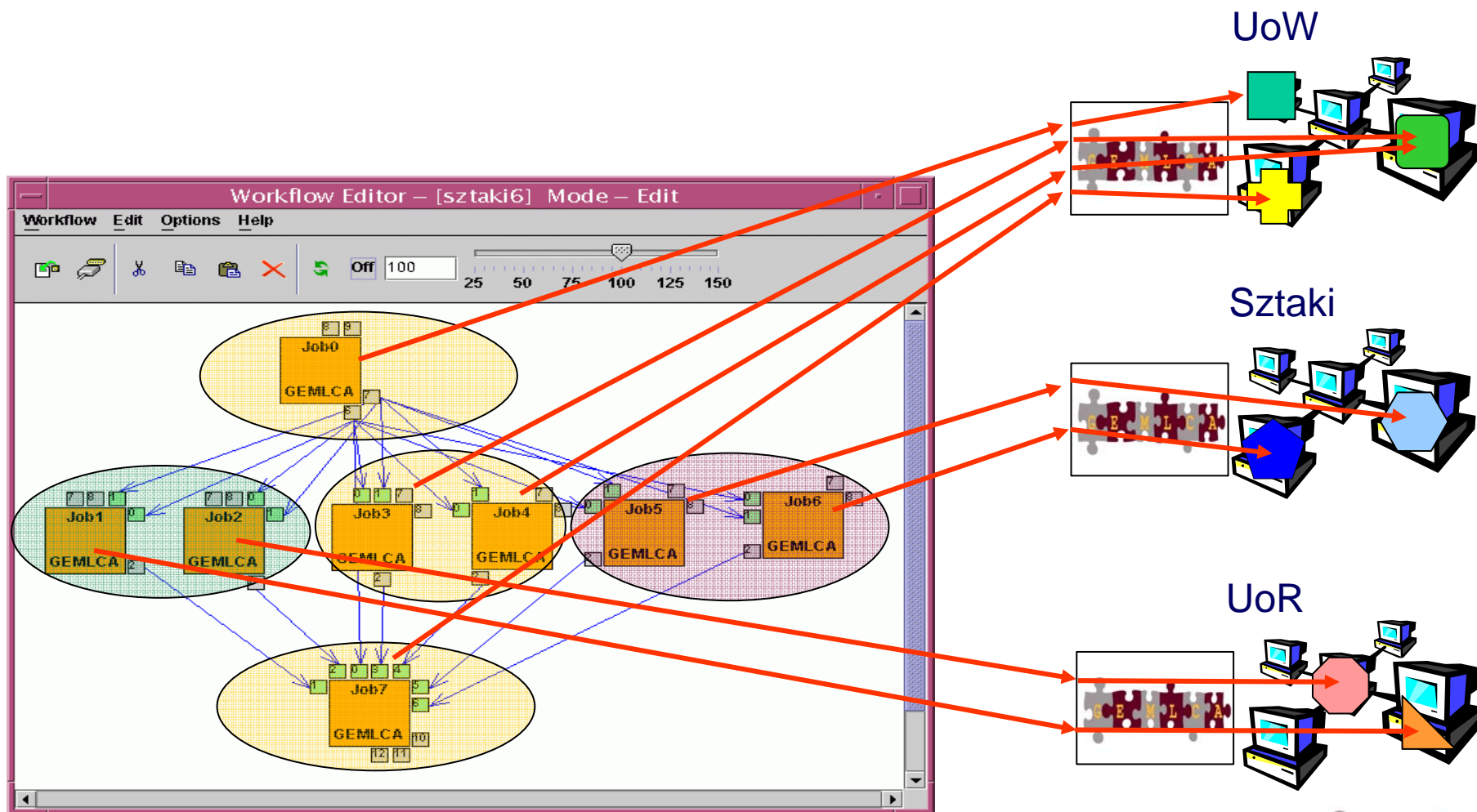
Ok Cancel





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

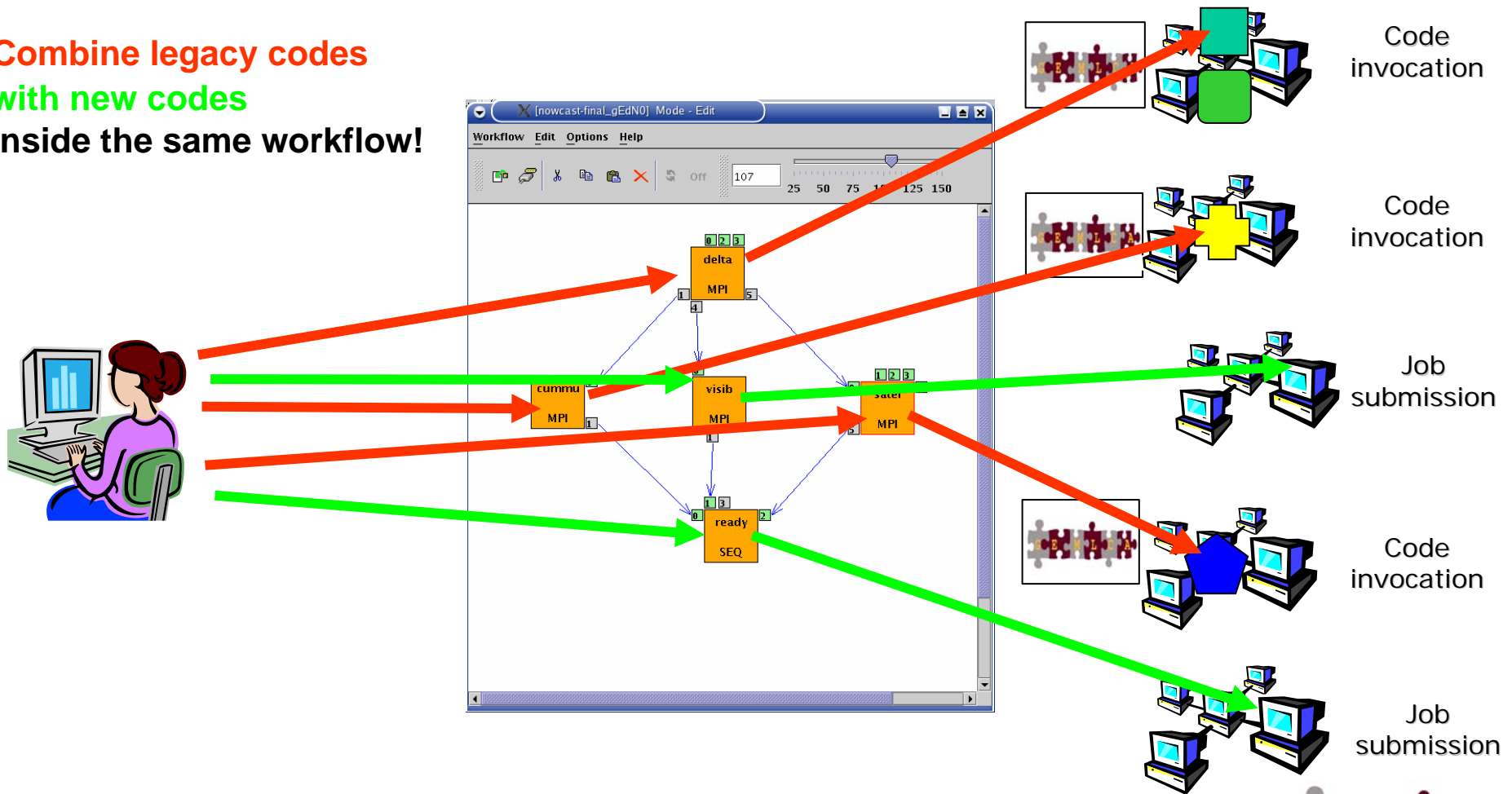
Ports guarantee compatibility → batch and GEMMLCA components can mutually produce data to each other!





Combining legacy and non-legacy components

Combine legacy codes
with new codes
inside the same workflow!

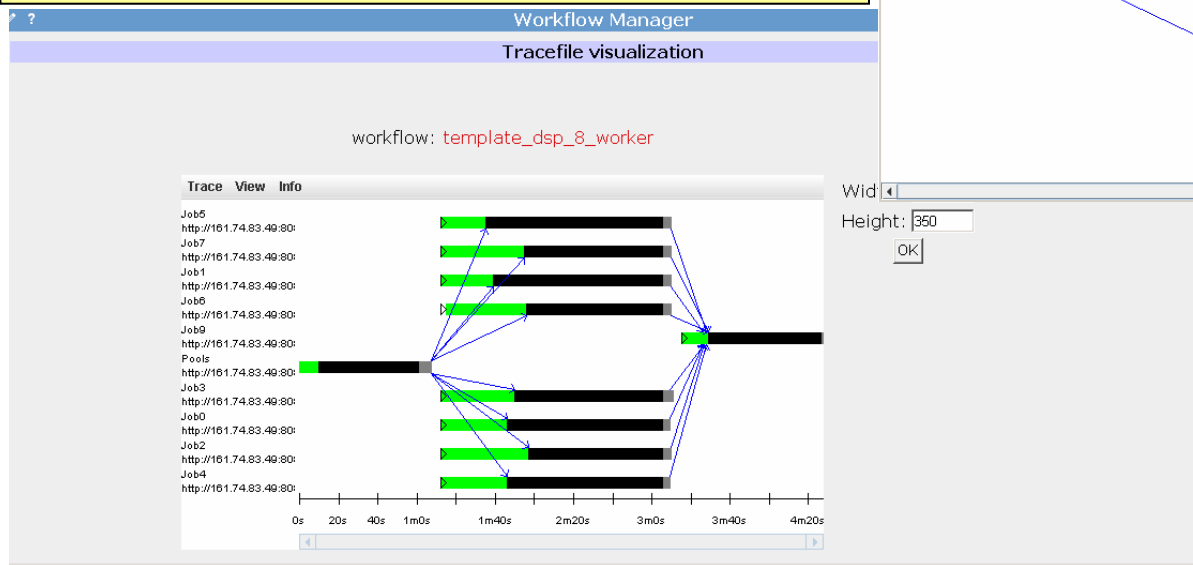
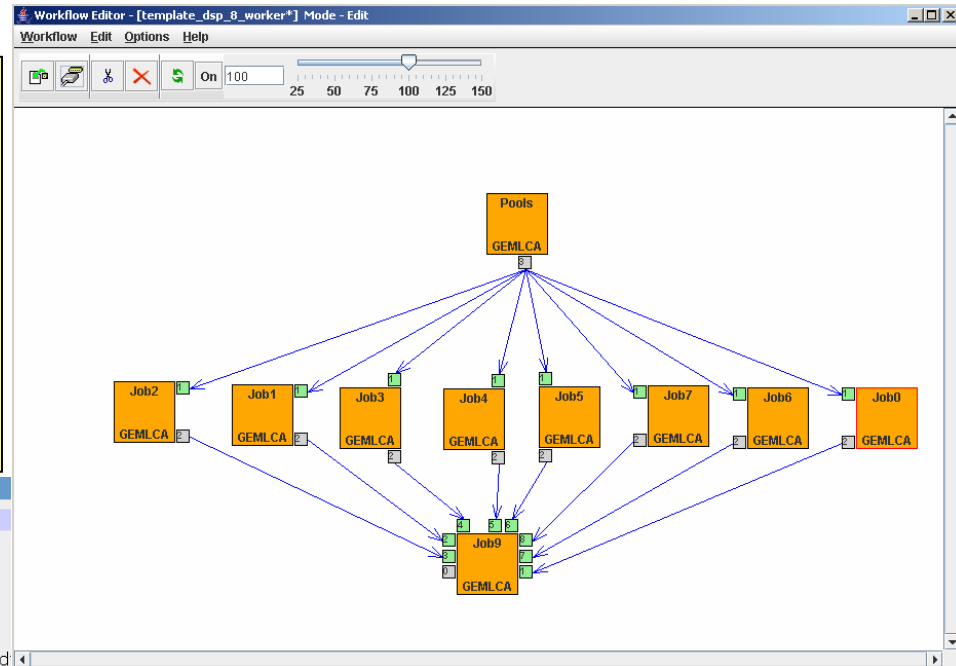




Application example

Designing Optimal Periodic Nonuniform Sampling Sequences

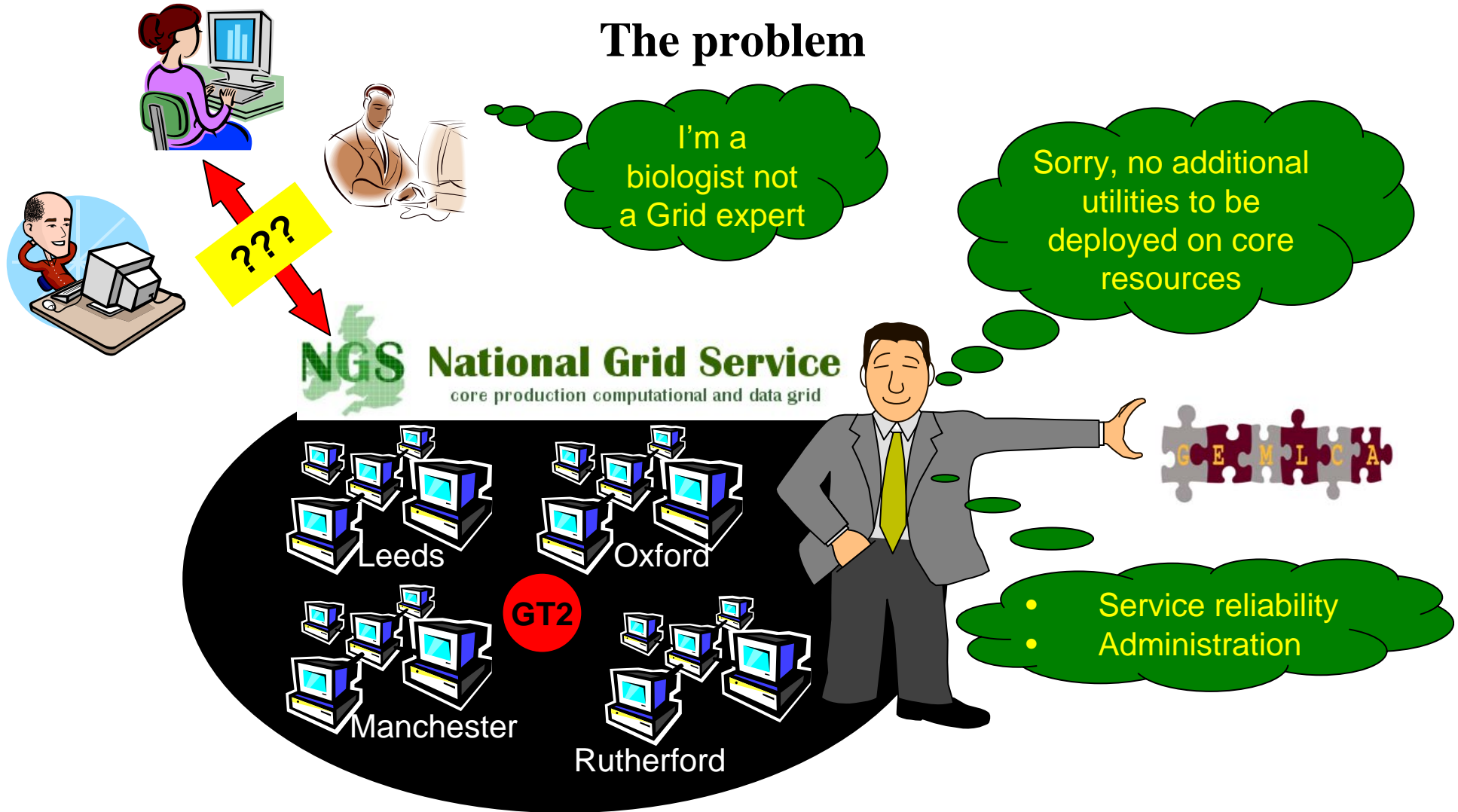
T Factor	Sequential	GEMMLCA
18	~19min	~8min
20	~3h 33min	35min
22	~41h 53min	~7h 23min





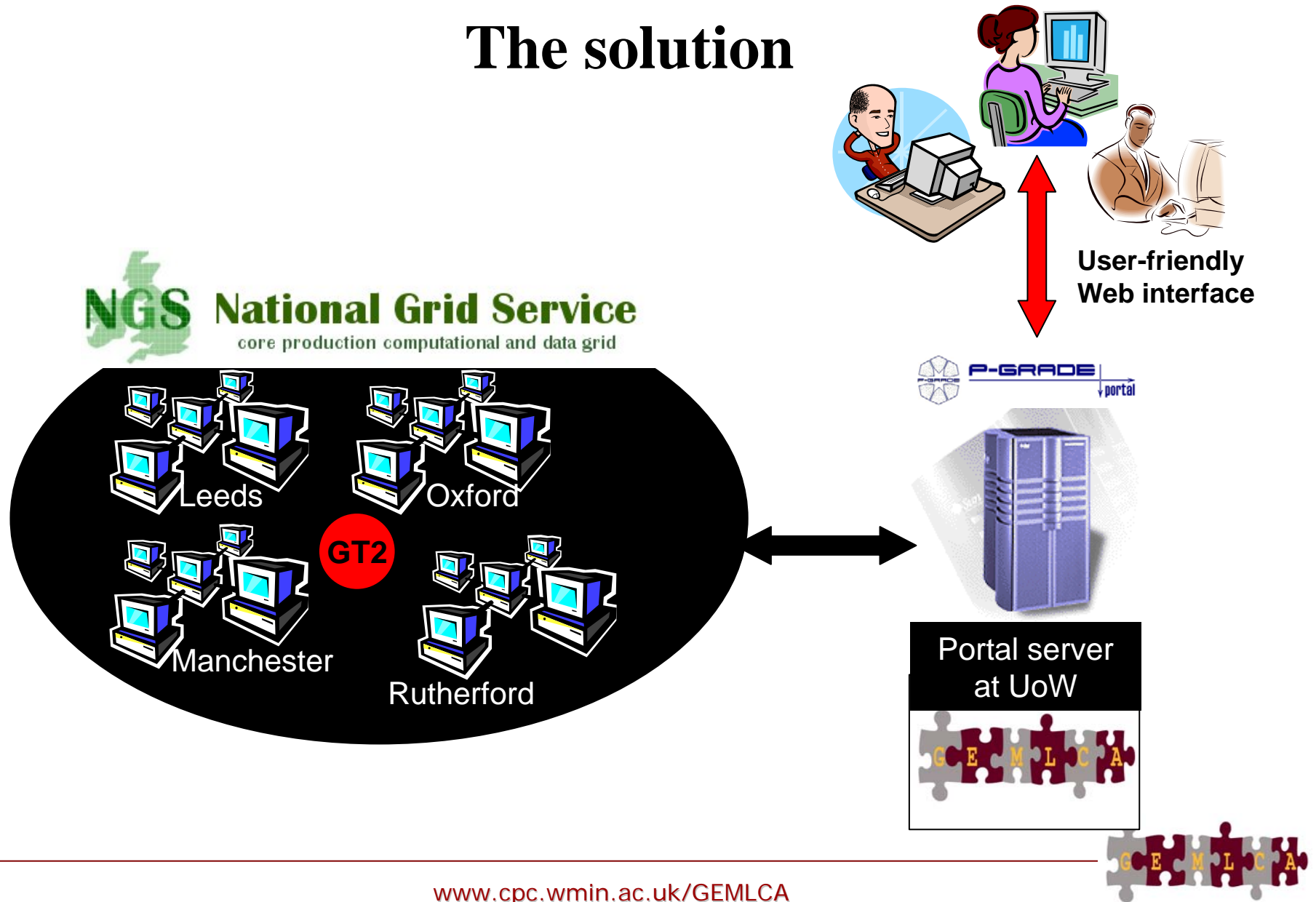
GEMMLCA and Production Grids

The problem



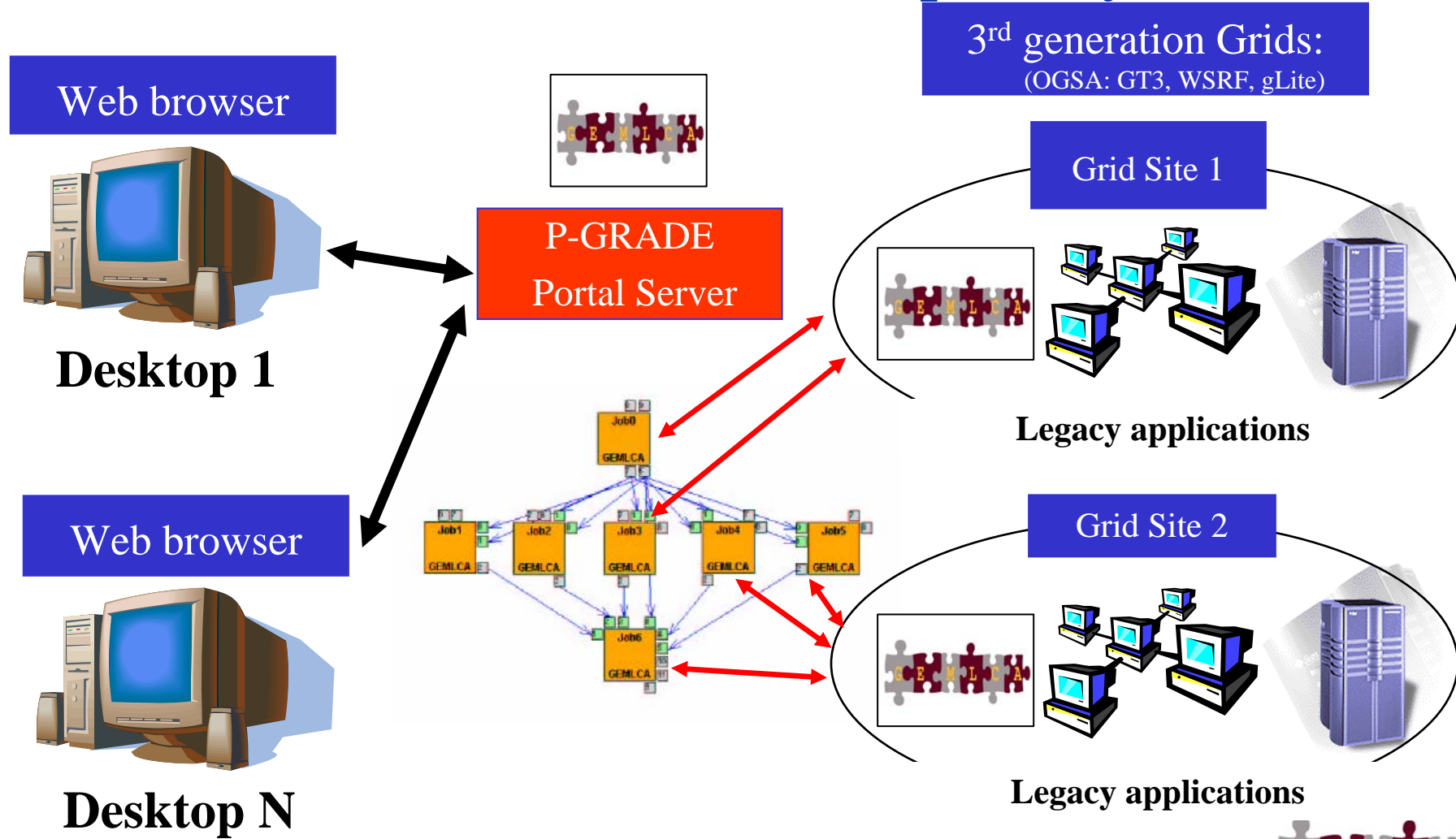


The solution



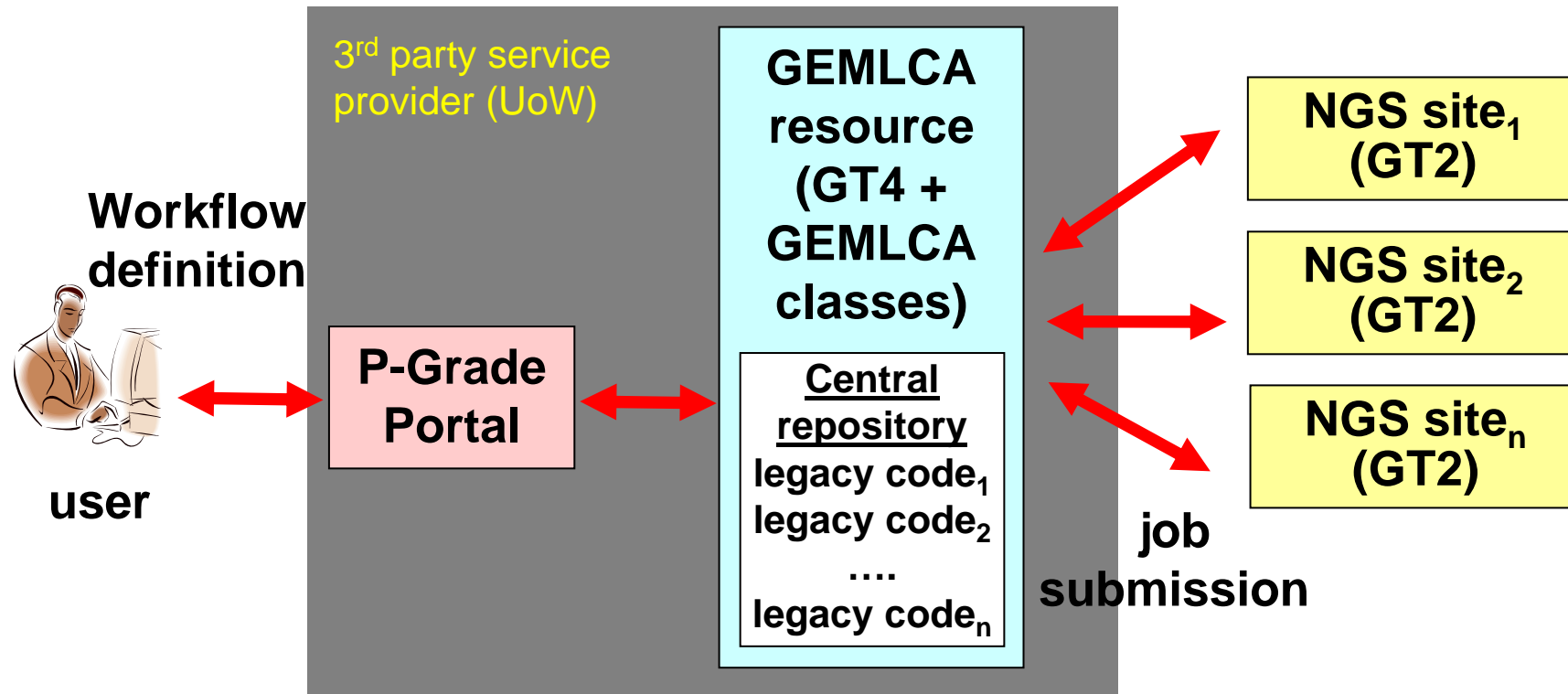


Centralised GEMMLCA solution – The GEMMLCA repository





GEMMLCA legacy code repository





Advantages of the GEMMLCA legacy code repository

- legacy codes can be uploaded into a central repository and made available for authorised users through a Grid portal
- extends the usability of NGS as users utilise others' legacy codes stored in the repository
- **No support needed at the NGS sites**





GEMMLCA on the UK NGS

The P-GRADE NGS GEMMLCA Portal

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





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 GEMMLCA web interface. At the top, there are logos for the University of Westminster, WestFocus, and gridalliance. Below the navigation menu, a table lists workflows with their status, size, and quota. To the right of the table is a network diagram showing connections between various nodes.

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	





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





Live demonstration of the P-GRADE GEMMLCA portal

P-Grade Portal - Microsoft Internet Explorer
Address: <http://ngs-portal.cpc.wmin.ac.uk:8080/gridsphere/gridsphere>

P-GRADE NGS GEMMLCA portal
Log in to the portal

Login
User Name:
Password:

Job2 properties

Name:
Job Type: GEMMLCA
Grid: NGS
Resource:
Legacy Code: Please select a legacy code:
Parameters: Please select a legacy code:
compare - Compare macroscopic traces (Fork)
gamess1 - Gamess description
liles - Description
liles2 - Description
lp - lp lc
madcitcity - Sequential traffic simulator (Fork/micro/macro/turns file)

Create application workflow

Map execution to available resources

Execute workflow

Tracefile visualization

workflow: demo
Width:
Height:

Visualise execution and download results

Job0 grid-compute.leeds.ac.uk
Job1 grid-compute.oeso.ox
Job2 grid-data.man.ae.uk
Job3 grid-data.rl.ac.uk

