



GEMLCA / P-GRADE: A workflow-oriented portal and application hosting environment

A short introduction

Tamas Kiss

University Of Westminster

kisst@wmin.ac.uk



www.portal.p-grade.hu
www.cpc.wmin.ac.uk/gemlca



E-scientists' concerns



- How to concentrate own **my own research** if the tool I would like to use is in continuous change?
- How can I learn and understand **the usage of the Grid**?
- How can I **develop Grid applications**?
- How can I **execute grid applications**?
- How to **tackle performance issues**?
- How to **use several Grids at the same time**?
- How to **migrate my application** from one grid to another?
- How can I utilise **legacy applications**?
- How can I **collaborate with fellow researchers**?

The GEMMLCA / P-GRADE Portal gives you the answers!



P-GRADE Portal in a nutshell



- **General purpose, workflow-oriented computational Grid portal.** Supports the development and execution of workflow-based Grid applications – **a tool for Grid orchestration**
- Based on **GridSphere-2**
 - Easy to expand with new portlets (e.g. application-specific portlets)
 - Easy to tailor to end-user needs
- Developed by SZTAKI
- **Grid services** supported by the portal:

Service	EGEE grids (LCG/gLite)	Globus 2 grids
Job execution	Computing Element	GRAM
File storage	Storage Element	GridFTP server
Certificate management	MyProxy	
Information system	BDII	MDS-2, MDS-4
Brokering	Workload Management System	(GTbroker)
Job monitoring	Mercury	
Workflow & job visualization	PROVE	

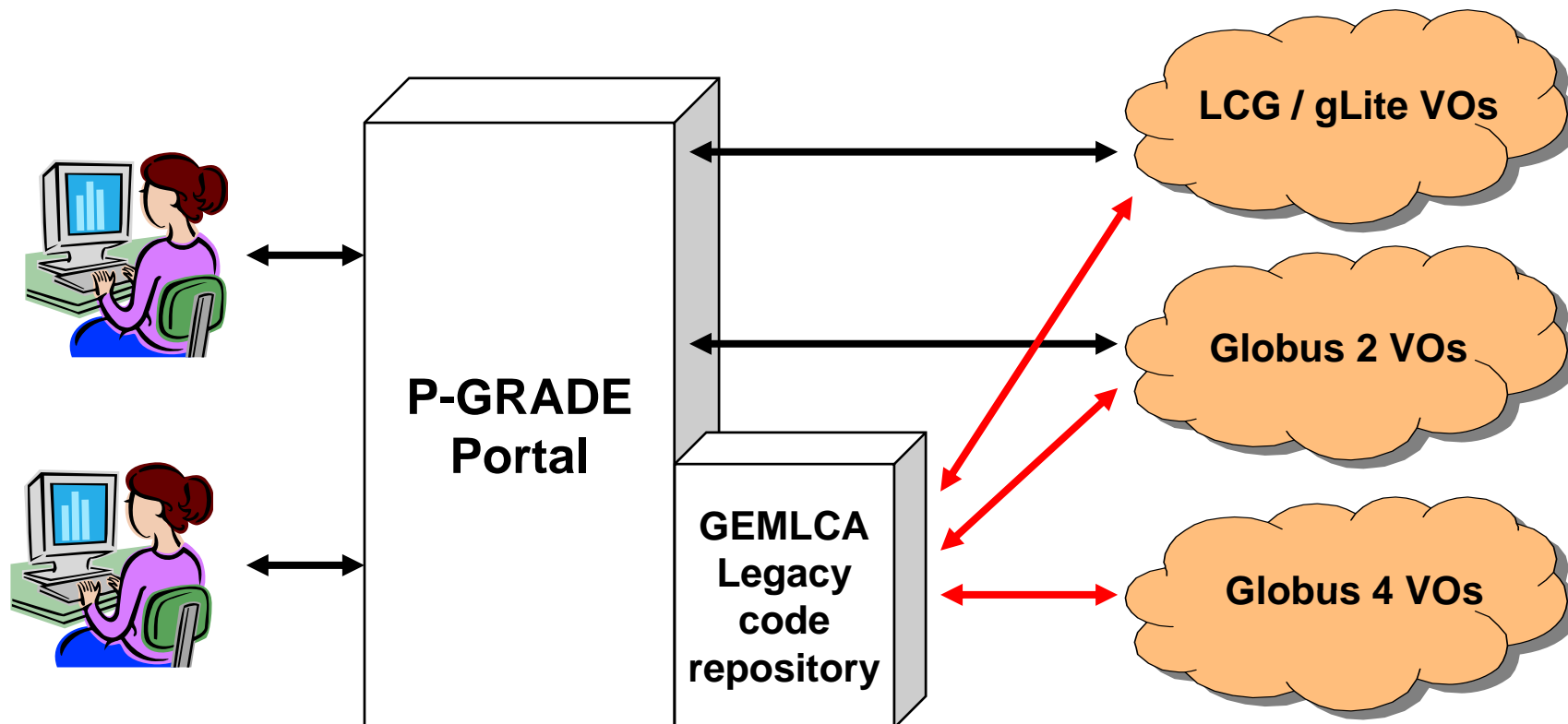
Solves Grid interoperability problem at the workflow level



GEMLCA / P-GRADE Portal ***in a nutshell***



- **P-GRADE Portal extended with GEMLCA back-end**
 - Sharing jobs and legacy codes as workflow components
 - GEMLCA is a grid service implemented by UoW
- **A step towards collaborative e-Science**
- **Support for Globus 4 grids (besides GT2 and EGEE)**

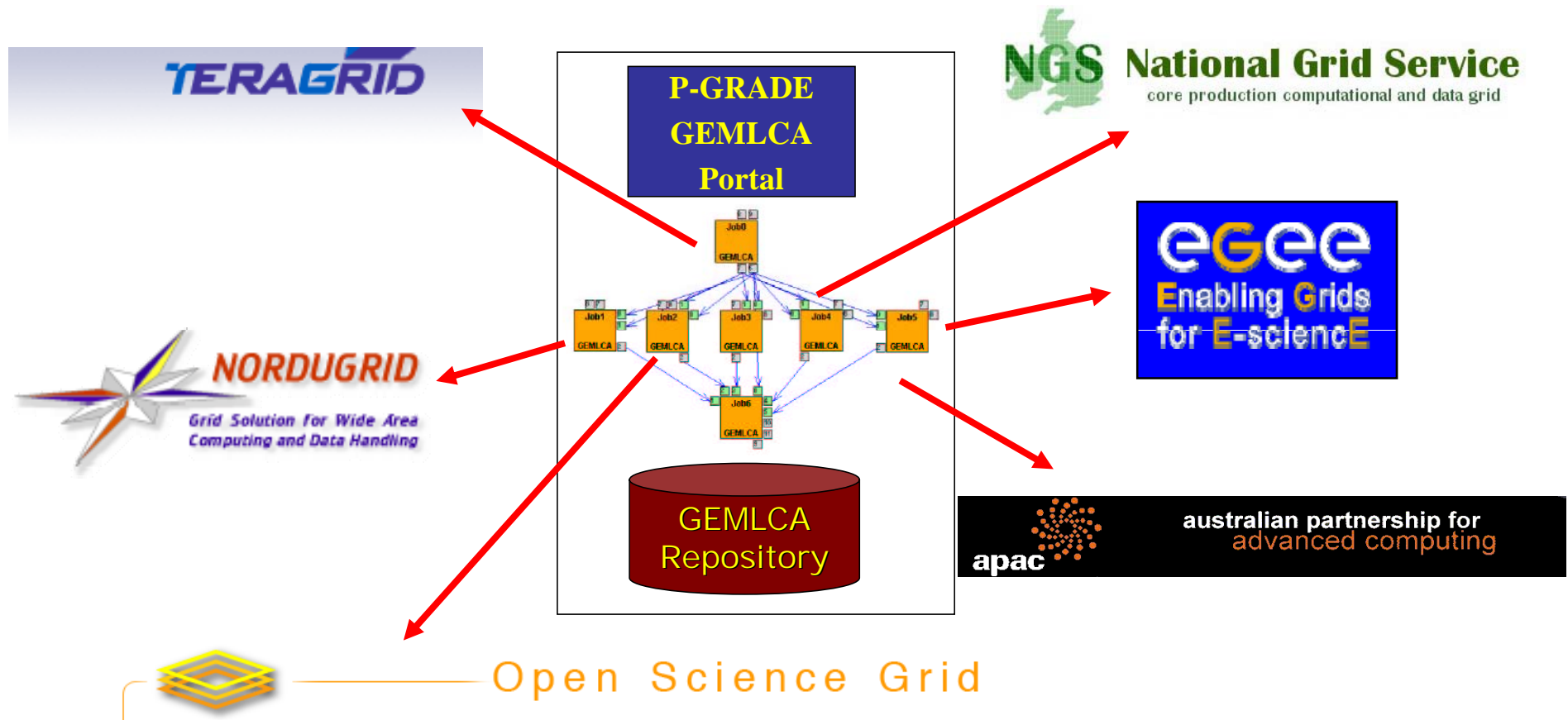




Workflow-level Grid interoperability: The GIN Resource Testing portal

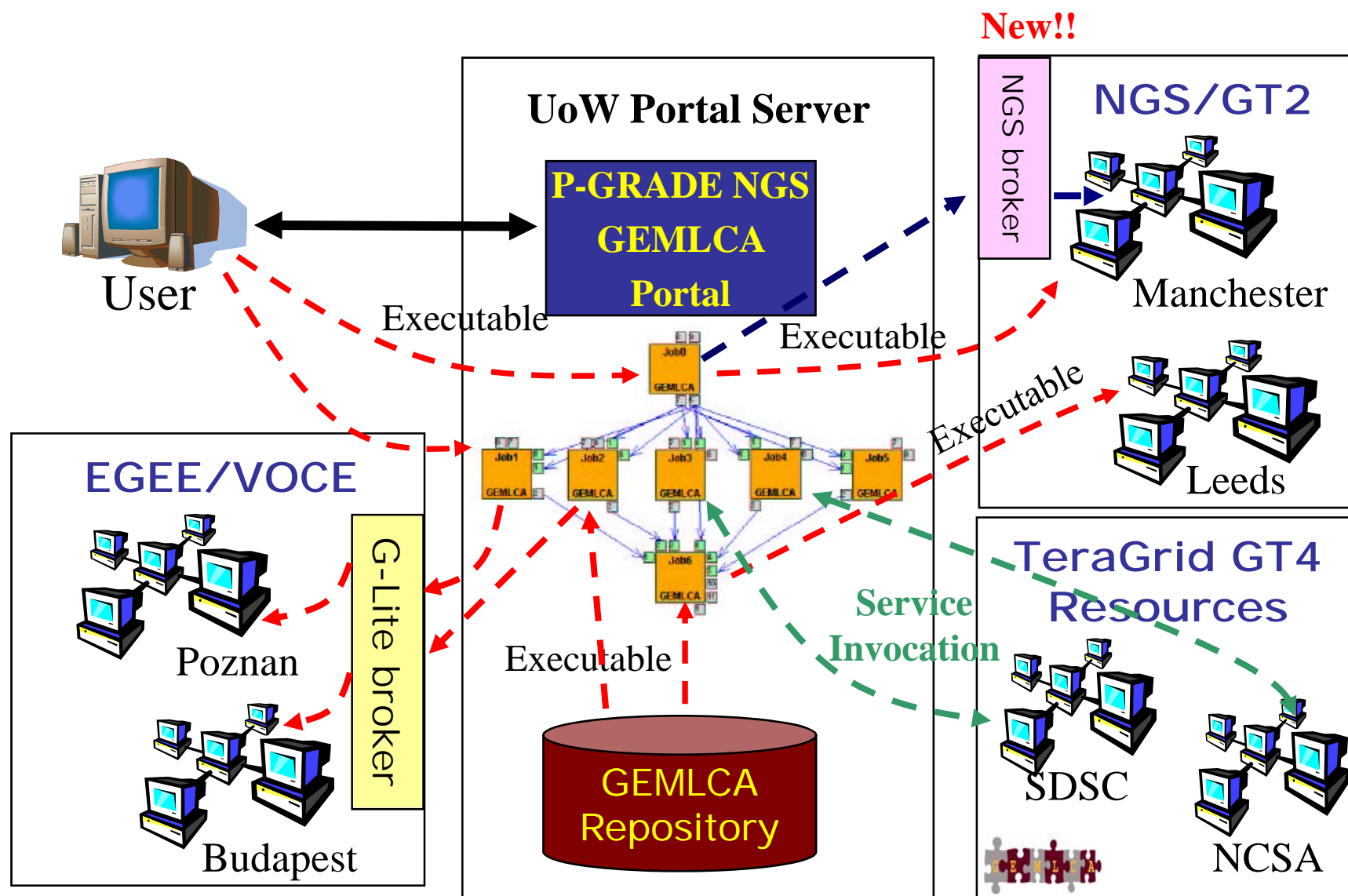


OGF effort to demonstrate workflow level grid interoperability between major production Grids and to monitor OGF GIN VO resources





Workflow level interoperability of Grid systems



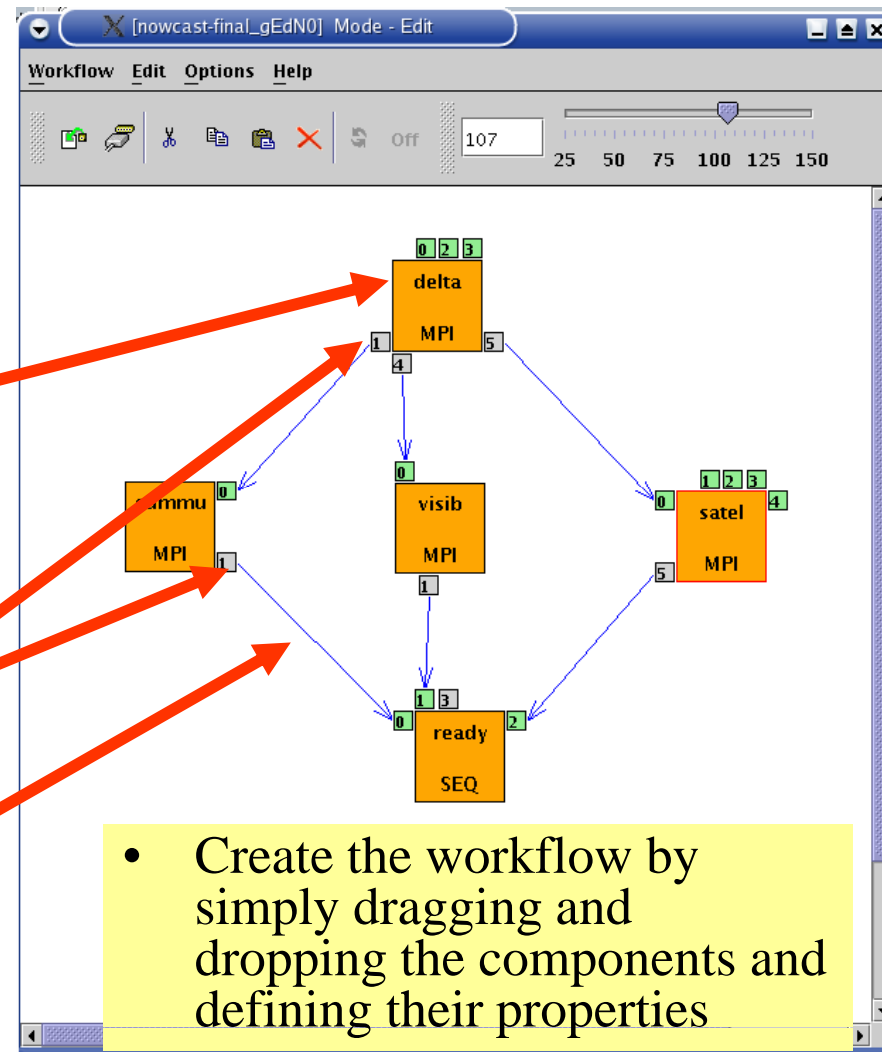


What is a GEMLCA / P-GRADE Portal workflow?



- A directed acyclic graph where:

- Nodes represent jobs - either sequential or parallel programs
- Ports represent input/output files the jobs expect/produce
- Arcs represent file transfer between the jobs

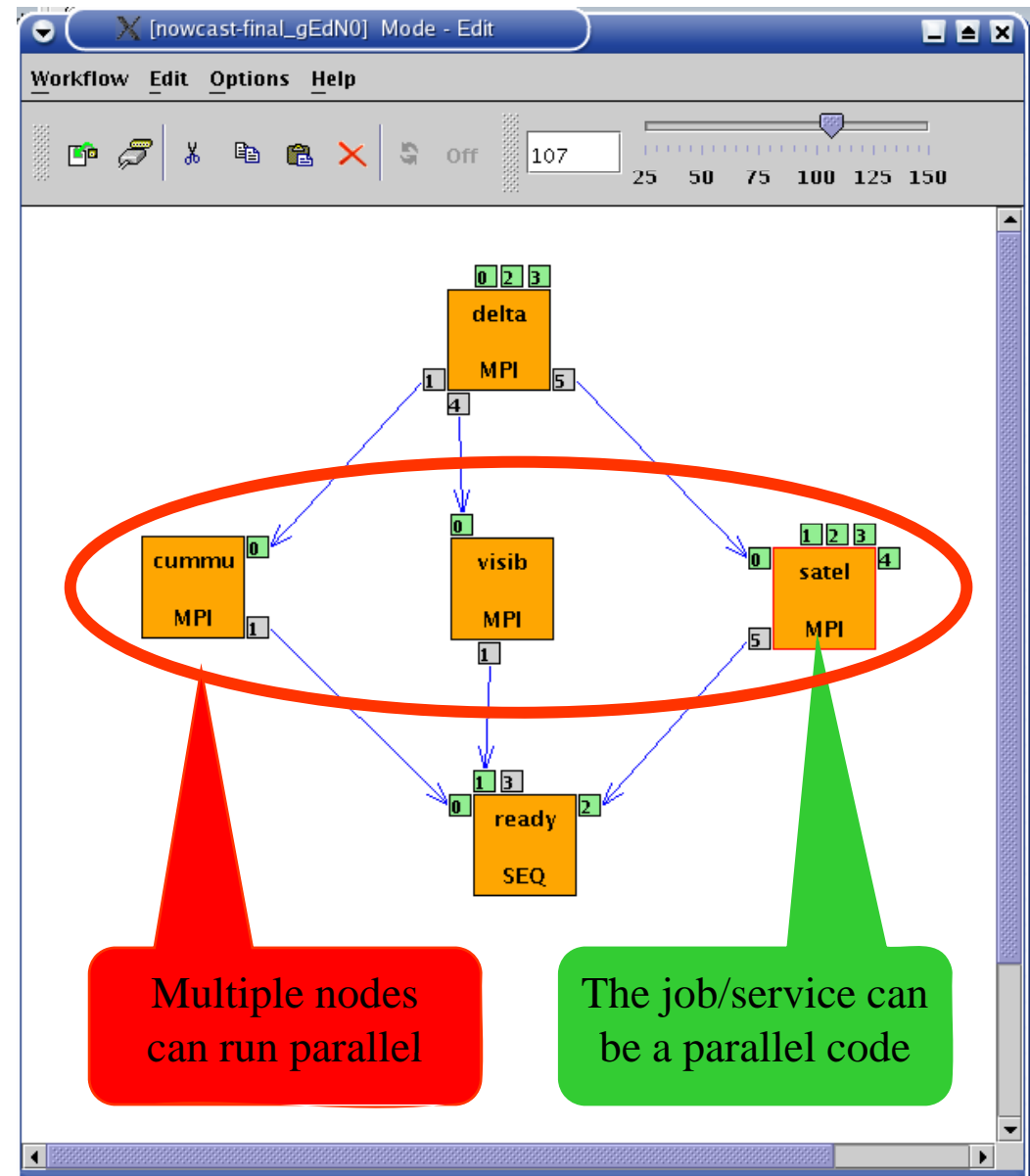




Two levels of parallelism within a workflow

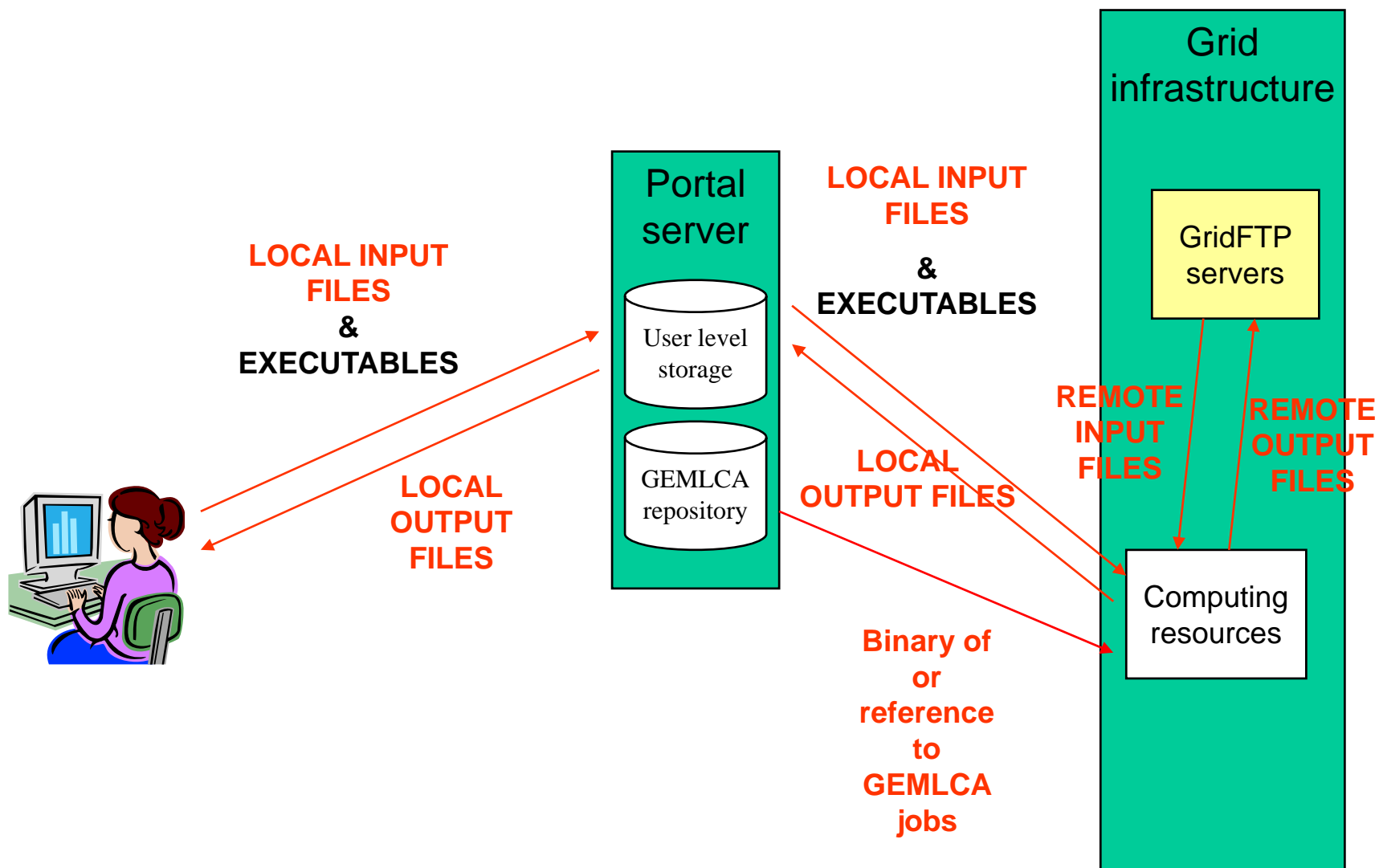


- The workflow concept of the GEMMLCA/ P-GRADE Portal enables the efficient parallelization of complex problems
- Semantics of the workflow enables two levels of parallelism:
 - Parallel execution inside a workflow node
 - Parallel execution among workflow nodes



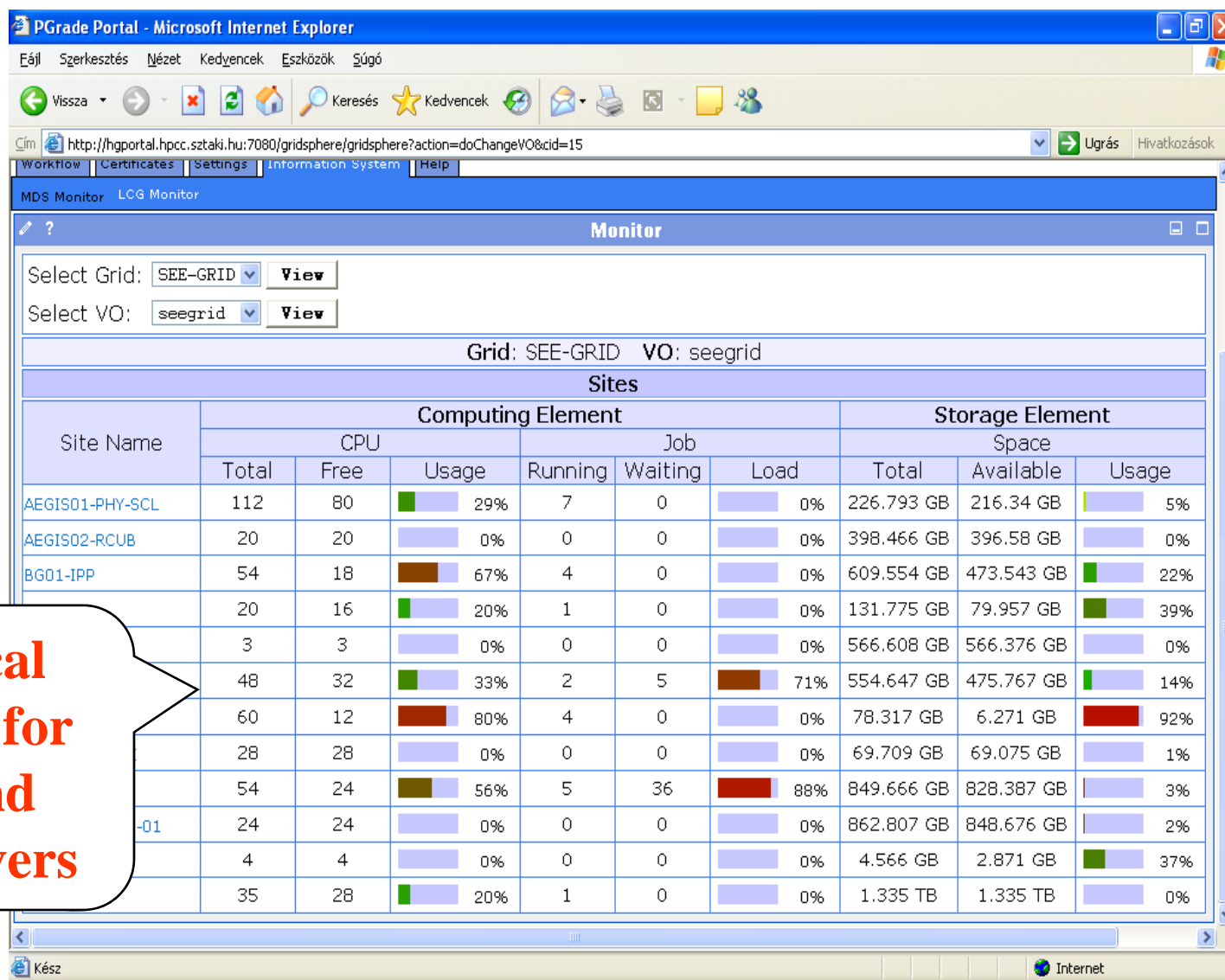


Workflow level file transfer





Browsing computing resources by the information system portlet



Graphical
interface for
GIIS and
BDII servers



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

GridSphere Portal - Microsoft Internet Explorer

RELEASE 2.4

P-Grade University of Westminster

MTA SZTAKI

GridSphere GT4/GEMLCA Monitor GIN VO Information

GT4/GEMLCA 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
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-login1.sdsc.teragrid.org	detail
GEMLCA	161.74.83.51	GEMLCA resource test for https://161.74.83.51:3101/wsrf/services/gn6.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/th1.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	fermigrd1.fnal.gov	GTT Probe "gmtgridftptest" for gsiftp://fermigrd1.fnal.gov:2811	detail
gmtprewsgramtest	fermigrd1.fnal.gov	GTT Probe "gmtprewsgramtest" for gram://fermigrd1.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



Certificate Manager

Multi-grid → Multi-proxy



PGrade Portal - Microsoft Internet Explorer

http://hgportal.hpcc.sztaki.hu:7080/gridsphere/gridsphere?action=doMapProxy&cid=5

P-GRADE portal

Workflow Certificates Settings Information System Help

Certificate Manager

Issuer	Set for Grids	Time left	[Actions]
DC=ORG,DC=SEE-GRID,O=People,O=SZTAKI,CN=Jozsef Patvarczki,CN=proxy	SEE-GRID	99:50:24	Details Set for Grid Delete
C=HU,O=KFKI RMKI CA,OU=SZTAKI,CN=Patvarczki Jozsef,CN=proxy	HUNGRID	99:57:25	Details Set for Grid Delete

[Refresh](#)

[Download](#) (Download certificate from MyProxy server.) [Upload](#) (Upload authentication data to MyProxy server.)

Message: Certificate successfully set for HUNGRID.

Multiple proxies can be available on the portal server at the same time!

SEE-GRID CEs and SEs

HUNGRID CEs and SEs



Workflow Execution

(observation by the workflow portlet)



PGrade Portal - Microsoft Internet Explorer

http://hgportal.hpcc.sztaki.hu:7080/gridsphere/gridsphere?action=doGotoPage&cid=2

Workflow Manager

Refresh Back

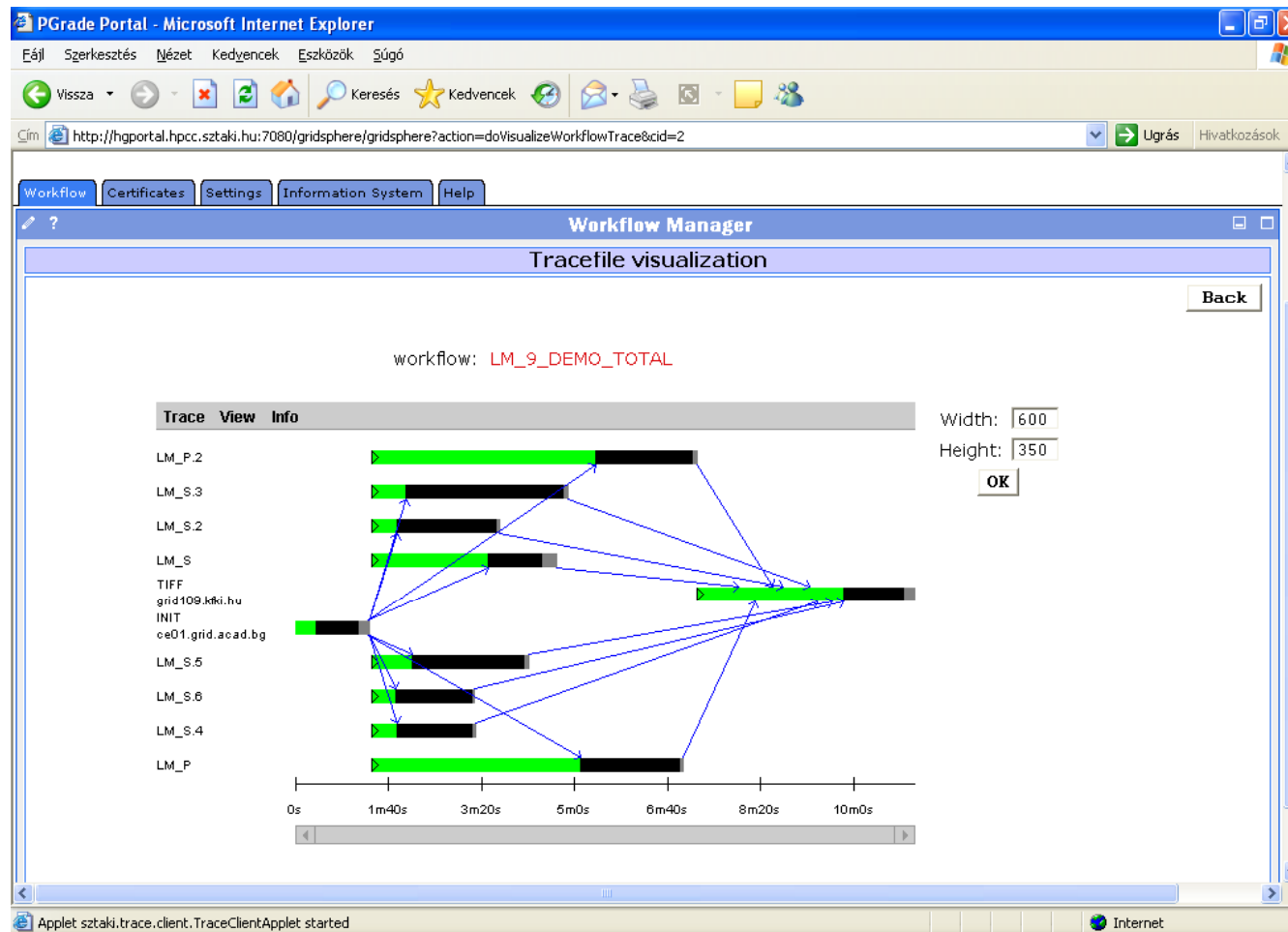
Workflow	Job	Gridname	Hostname	Status	[Logs]	[Output]	[Visualization]	
LM_9_DEMO_TOTAL				running	-	N/A	Visualize	All Abort
	INIT	SEE-GRID	ce01.grid.acad.bg	finished	- -		-	
	LM_P	SEE-GRID	n40.hpcc.sztaki.hu	init	- -		-	
	LM_P.2	SEE-GRID	n40.hpcc.sztaki.hu	init	- -		-	
	LM_S	SEE-GRID	grid-ce.ii.edu.mk	running	- -		-	
	LM_S.2	SEE-GRID	grid1.irb.hr	finished	Out -		-	
	LM_S.3	SEE-GRID	grid1.netmode.ece.ntua.gr	running	Out -		-	
	LM_S.4	SEE-GRID	grid1.irb.hr	finished	Out -		-	
	LM_S.5	SEE-GRID	testbed001.grid.ici.ro	running	Out -		-	
	LM_S.6	HUNGRID	chemgrid3.chemres.hu	finished	Out -		-	
	TIFF	HUNGRID	grid109.kfki.hu	init	- -		-	

Message: Job list refreshed.

White/Red/Green color means the job is initial/running/finished state



On-Line Monitoring both at the workflow and job levels *(workflow portlet)*



- The portal monitors and visualizes workflow progress



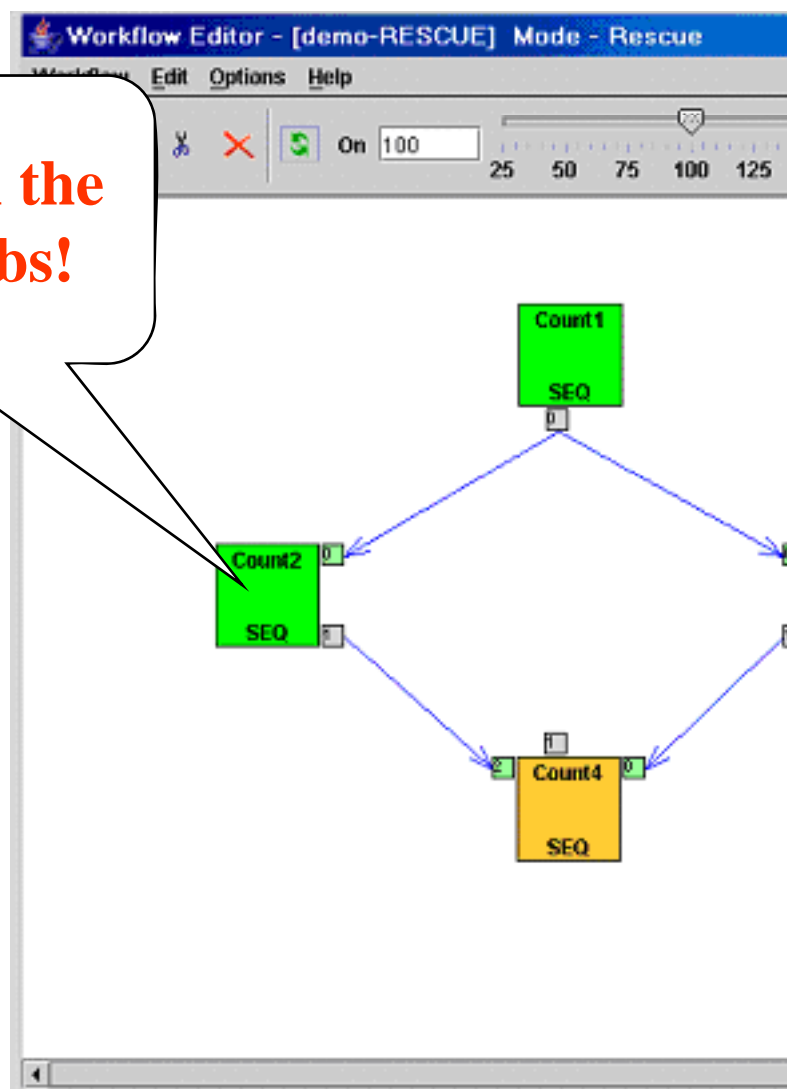
Rescuing a failed workflow



Don't touch the finished jobs!

Map the failed job onto a different resource or download a new proxy for it

The execution can continue from the point of failure





Downloading the results...



PGrade Portal - Mozilla

File Edit View Go Bookmarks Tools Window Help

http://fn2.hpcc.sztaki.hu:9080/gridsphere/gridsphere?action=doGotoPage&cid=2

Home Bookmarks The Mozilla Or... Latest Builds

P-GRADE portal

Logout

Welcome, Nemeth Csaba

Workflow Credentials Settings Demo Help

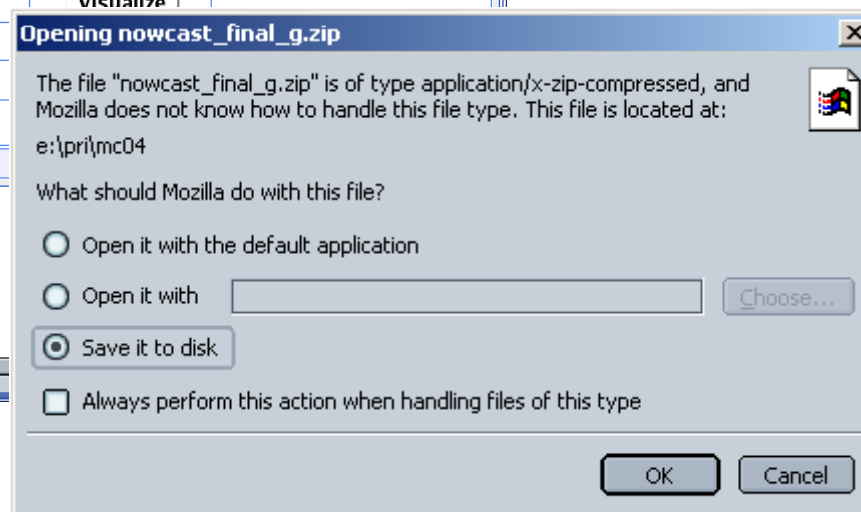
Workflow Manager

Refresh Back

Workflow	Job	Hostname	Status	[Logs]	[Output]	[Visualization]	[Action]
nowcast-final-g_SGE			finished		✓	Visualize All	Subm Attach Delete
	cummu	n0.hpcc.sztaki.hu	finished	--		Visualize	
	delta	n0.hpcc.sztaki.hu	finished	--		Visualize	
	ready	n0.hpcc.sztaki.hu	finished	--		Visualize	
	satel	n0.hpcc.sztaki.hu	finished	--			
	visib	n0.hpcc.sztaki.hu	finished	--			

Message: Job list refreshed.

Transferring data from fn2.hpcc.sztaki.hu...





Putting a successfully finished job into the GEMLCA repository



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 1

executable LINUX/mkdir

minimumProcessors 1

maximumJob 11

jobManager Fork

id mkdir

description Unix mkdir program

Set Parameters

List of legacy code Arguments:

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

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

Add Argument

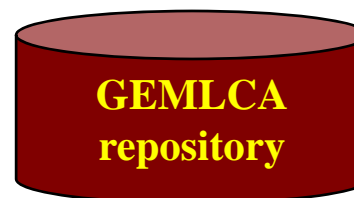
Mkdir Legacy Code exposed as a Grid Service

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>
```



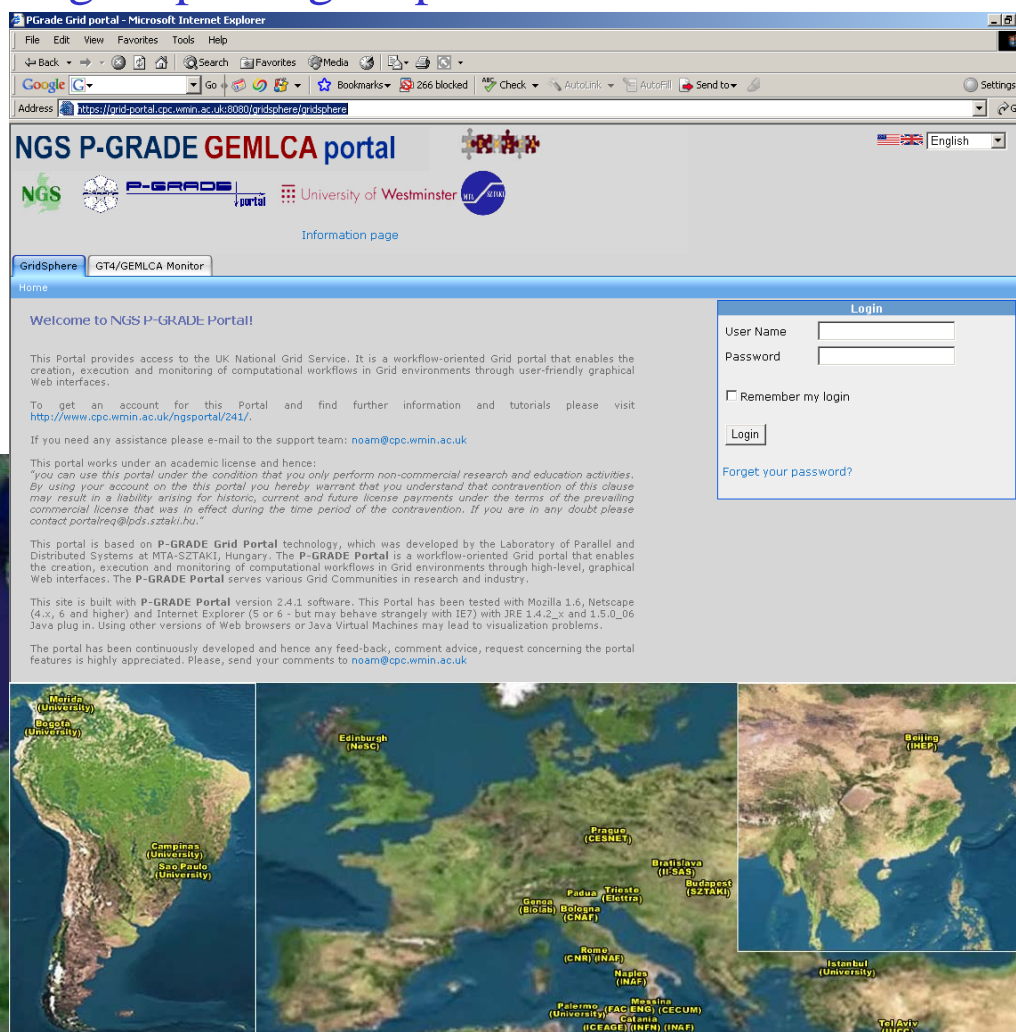


GEMLCA / P-GRADE on the UK NGS:

NGS P-GRADE GEMLCA Portal



- portal website:
<https://ngs-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





Other P-GRADE Portal installations

- P-GRADE Portal service is available for
 - SEE-GRID infrastructure
 - Central European VO of EGEE
 - GILDA: Training VO of EGEE
 - US Open Science Grid, TeraGrid
 - Economy-Grid, Swiss BioGrid, Bio and Biomed EGEE VOs, BioInfoGrid, BalticGrid
 - **OGF GIN (also connected to NGS)**





***Thank you for your
attention!***

***Hands-on session with the P-
GRADE/GEMMLCA portal will now
follow.***