



Workflow and parameter study management by P-GRADE Portal

Gergely Sipos
sipos@sztaki.hu

MTA SZTAKI
Hungarian Academy of Sciences



portal.p-grade.hu



Outline

- **P-GRADE Portal and Developer Alliance**
- **Capabilities of P-GRADE Portal 2.5**
 - Workflow support
 - Parametric study support
- **Other portal versions**
 - GEMLCA P-GRADE Portal
 - Application specific portals based on P-GRADE
- **Hand-on exercises**
- **How to get access**
- **Roadmap**
 - P-GRADE grows into gUSE
- **Summary**



Motivations to P-GRADE Portal

- Fast evolution of Grid systems and middleware:
 - **GT2, OGSA, GT3 (OGSI), GT4 (WSRF), LCG-2, gLite, ...**
- Many production Grid systems are built with them
 - EGEE (**LCG-2 → gLite WMS → WMPProxy**), UK NGS (**GT2**), Open Science Grid (**GT2 → GT4**), NorduGrid (**~GT2**)
- Although **the same set of core services** are available everywhere, they **are implemented in different ways**
 - Data services (file management)
 - Computation services (job submission)
 - Security services (proxy based single sign-on)
 - Brokers (not in every middleware, but e.g. in gLite - WMS)
- **Provide an easy to use environment for the management of grid applications and services**



P-GRADE Portal in a nutshell

- General purpose grid user environment
- Based on GridSphere web portal framework
- Development started in 2003
- Tool that helps you develop and execute workflows and workflow based parametric studies
- A Grid orchestration environment. Supported services:

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

Solves Grid interoperability problem at the workflow level



Portal Developer Alliance, Portal versions

- **Core development by MTA SZTAKI, Budapest:**
P-GRADE Portal 2.5
- Support for legacy codes and code repositories by
University of Westminster, London:
GEMICA P-GRADE Portal 2.4.1
- File manager and credential manager portlet by **Middle East Technical University, Ankara**
Portlets used in Turkish National Grid Portal
- Alternative job scheduling component by **Eötvös Loránd University, Budapest**
Under development
- Monte Carlo simulation support library portlet by **Rudjer Boskovic Institute, Zagreb**
Under development



Related projects

The development and support services of P-GRADE are founded by the following projects:

- **SEE-GRID** www.see-grid.eu
portal development
grid application development
- **Coregrid** www.coregrid.net
Research & development
- **EGEE** www.eu-egee.org
grid end user training
grid application development
- **ICEAGE** www.iceage-eu.org
Grid end user training, conceptual training





EGEE'07 Announcement: ***P-GRADE Portal goes open source!***

- P-GRADE Portal Alliance – Booth #10
 - Demo sessions on Monday and Tuesday

- P-GRADE Portal and Developer Alliance
 - Conference session on Wednesday
11:00-12:30



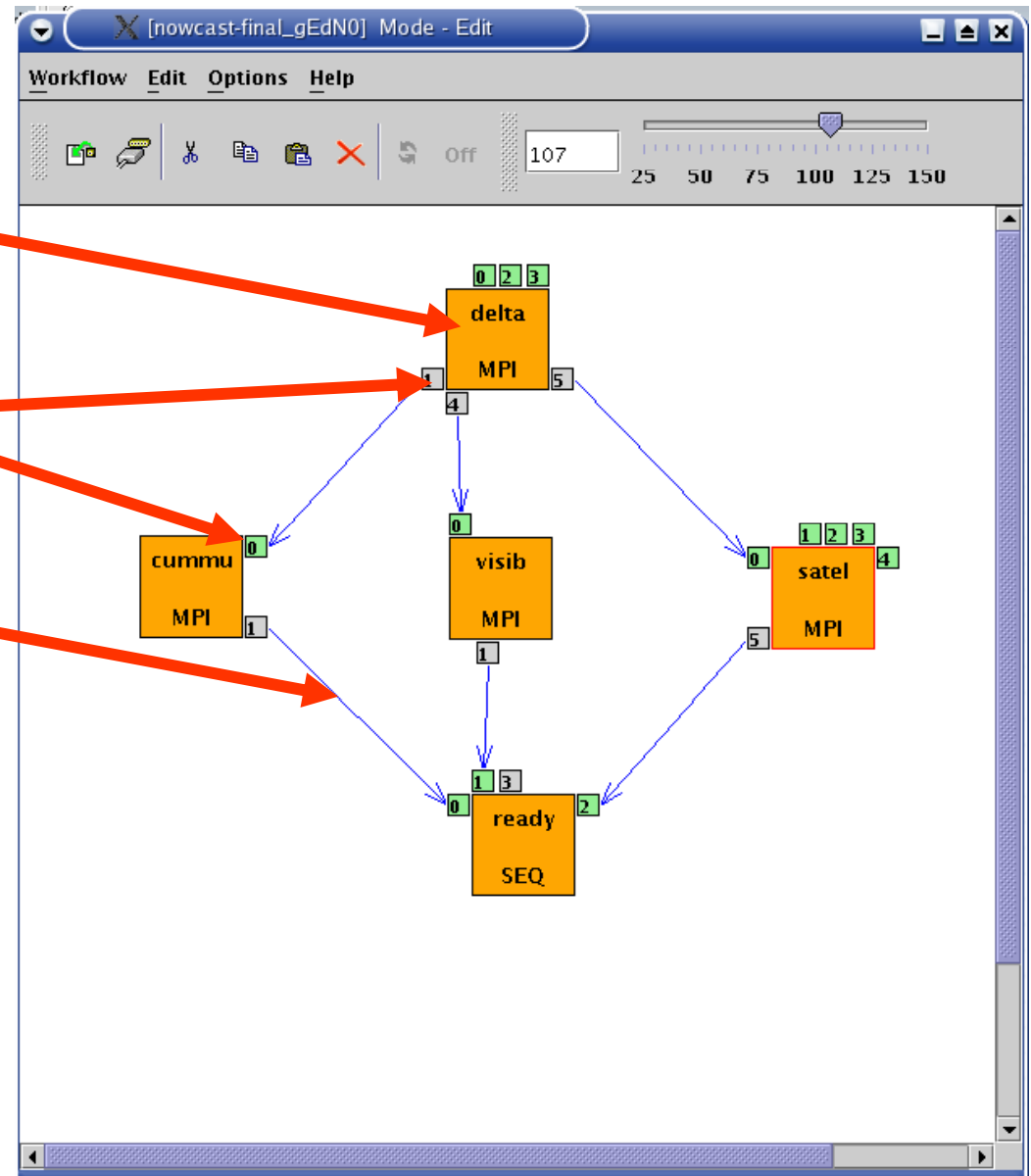
Outline

- **P-GRADE Portal and Developer Alliance**
- **Capabilities of P-GRADE Portal 2.5**
 - Workflow support
 - Parametric study support
- **Other portal versions**
 - GEMLCA P-GRADE Portal
 - Application specific portals based on P-GRADE
- **Hand-on exercises**
- **How to get access**
- **Roadmap**
 - P-GRADE grows into gUSE
- **Summary**



Elements of a P-GRADE Portal 2.5 application

- A directed acyclic graph where
 - Nodes represent batch jobs to be submitted from the client side to a job queue
 - Ports represent input/output files the jobs require or produce
 - Arcs represent file transfer operations and dependencies among jobs
- semantics of the workflow:
 - A job can be executed if all of its input files are available
 - Responsibility of the built-in workflow manager





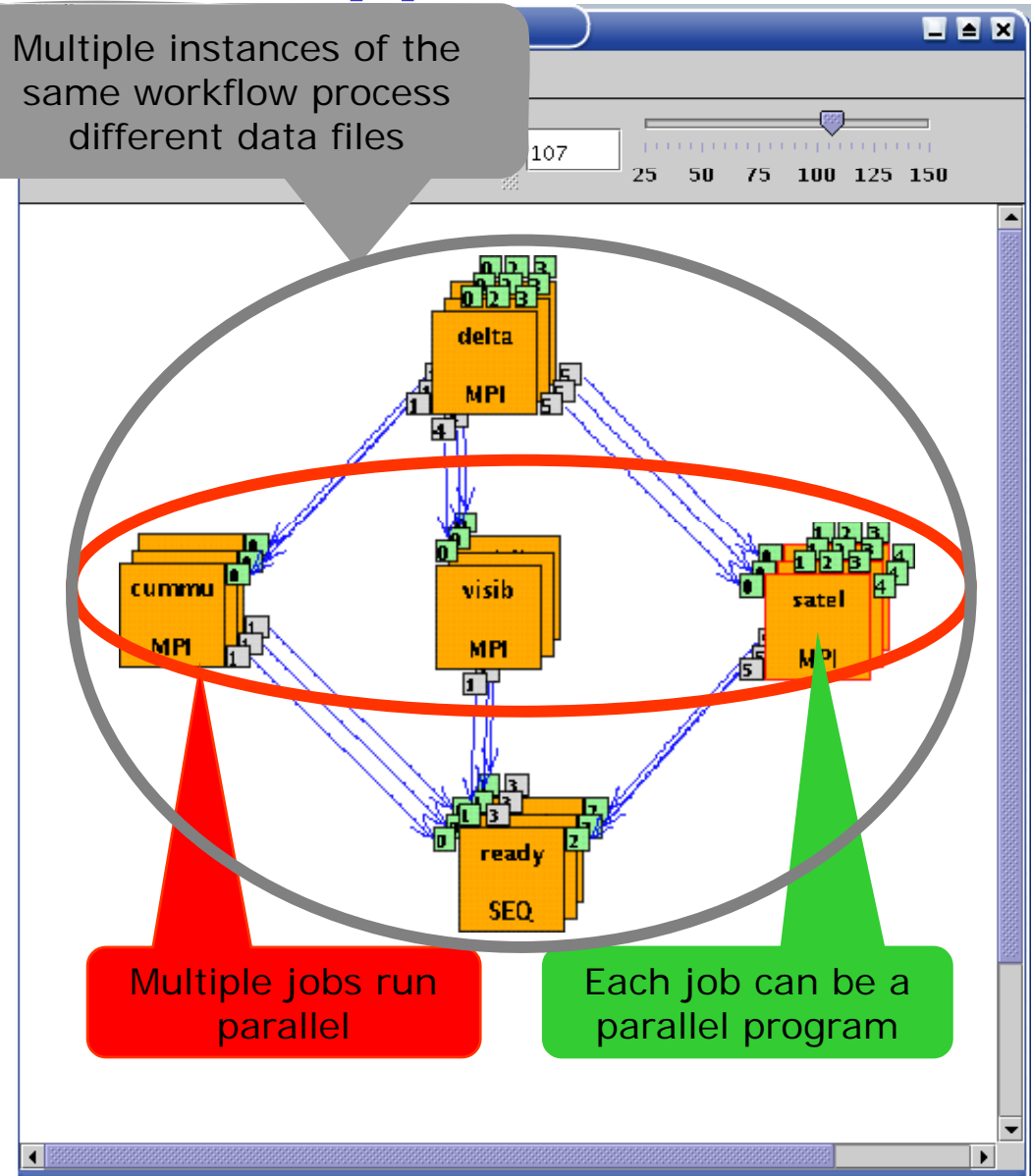
Parallel execution by a P-GRADE Portal application

– Parallel execution inside
a workflow node
(MPI job as workflow
component)

– Parallel execution among
workflow nodes
(different jobs on different
clusters)

– Parameter study execution
of the workflow (Single
instruction Multiple Data)

Multiple instances of the
same workflow process
different data files

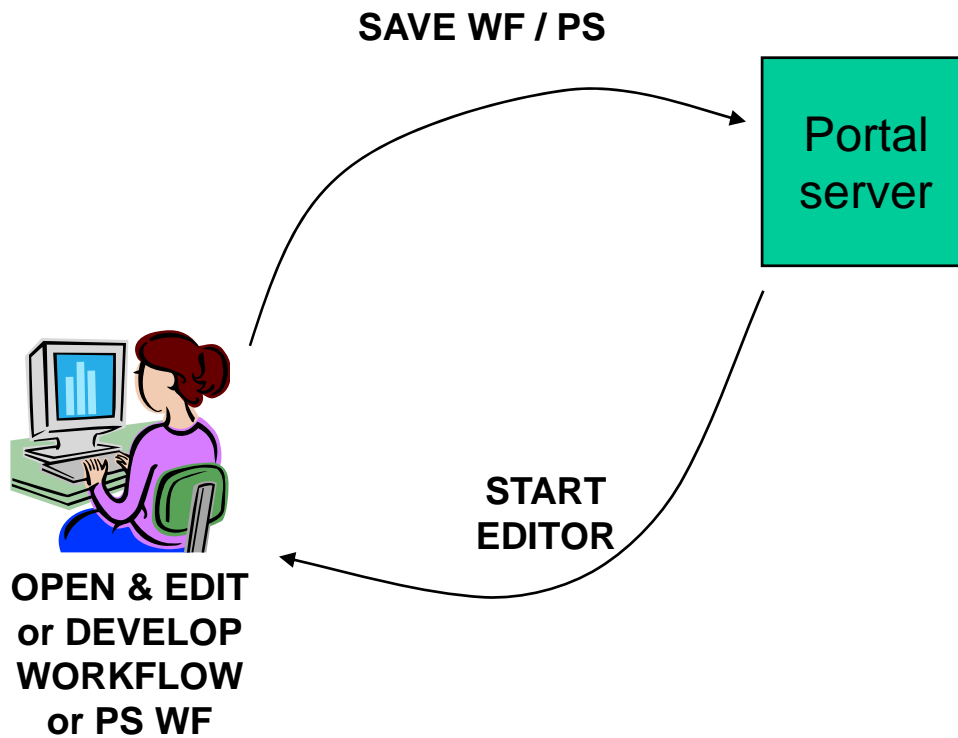




The typical user scenario

Part 1 - development phase

MyProxy
servers

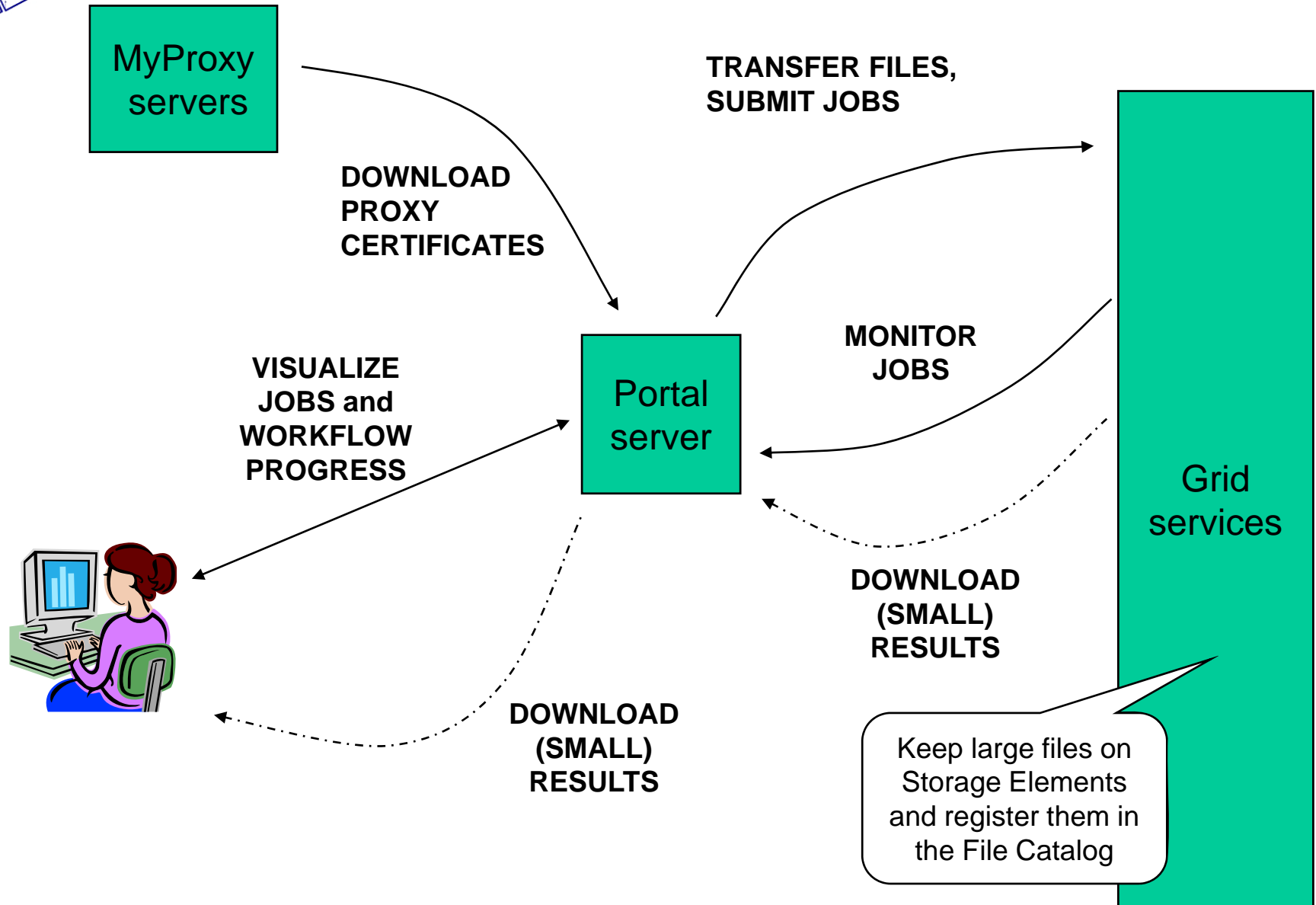


Grid
services



The typical user scenario

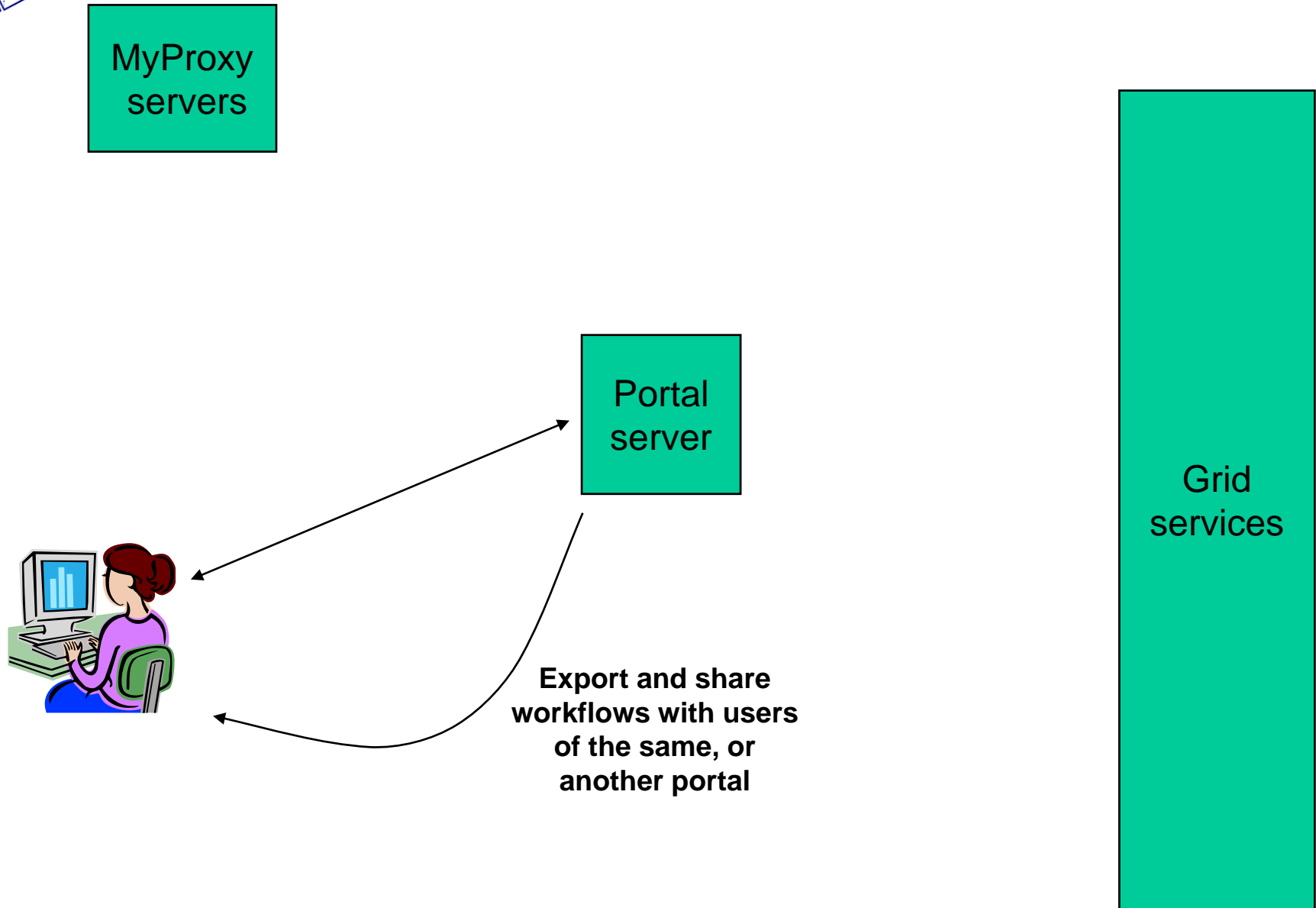
Part 2 - execution phase

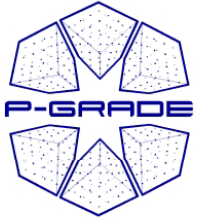




The typical user scenario

Part 3 - collaborative phase





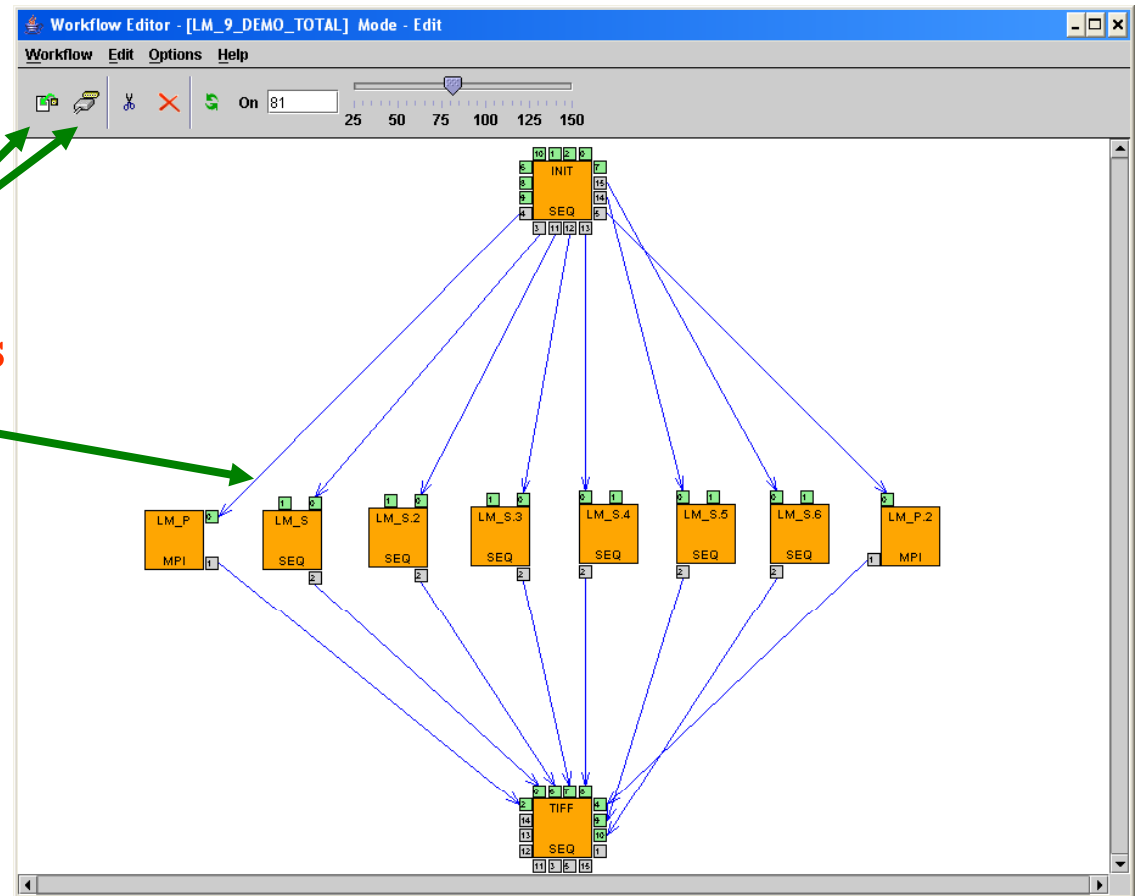
Defining a grid application

Define a Directed Acyclic Graph (*DAG*) of jobs:

1. **Drag & drop components:**
nodes and ports
2. **Define component properties**
3. **Connect ports by channels**
(no cycles, no loops, no conditions...)

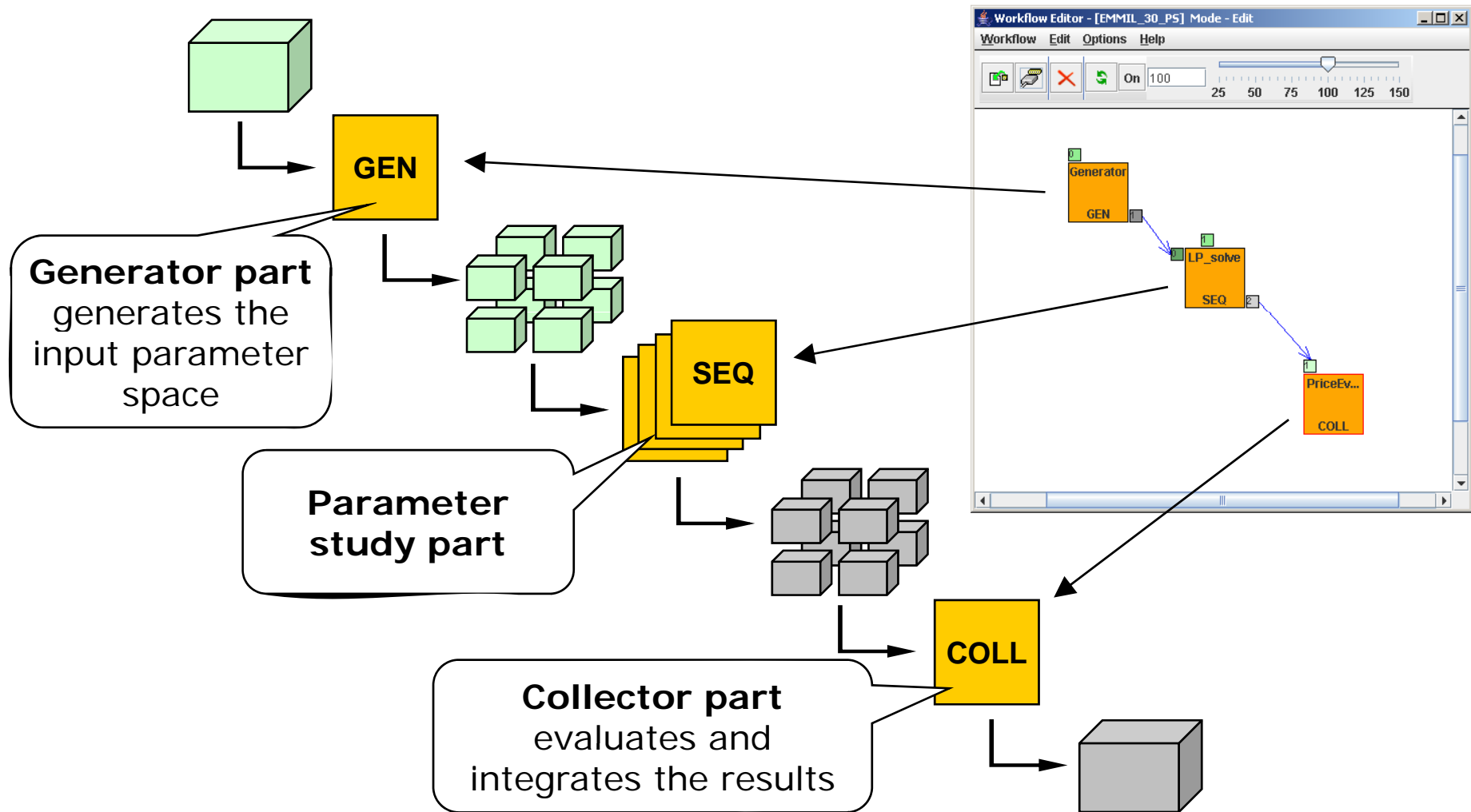
Extend the workflow into a parameter study:

1. **Add generator components**
2. **Add collector components**





Concept of parameter study workflows





Graphical User Interfaces in a nutshell

Proxy management

Workflow Certificates Settings Demo Help

Certificate Manager

Used certificate

Downloaded from:	albert.lpds.sztaki.hu
Issued by:	O=HunGrid,O=SZTAKI HPCC,OU=lpds.sztaki.hu,CN=Hermann Gabor,CN=proxy
Timeleft:	40:0:56
Description:	proxy for the demo

Certificate list

Issuer	Time Left	Status	[Actions]
O=HunGrid,O=SZTAKI HPCC,OU=lpds.sztaki.hu,CN=Hermann Gabor,CN=proxy	40:0:56	[used]	<input type="button" value="Details"/> <input type="button" value="Delete"/>

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

Message: Download successful.



Graphical User Interfaces in a nutshell

Proxy management
Grid, VO and Grid
resource management

Globus resources

URL	Job manager	[Actions]
fs0.das2.cs.vu.nl	jobmanager-fork	Delete
hitcross.lrz.muenchen.de	jobmanager-fork	Delete
litchi.zib.de	jobmanager-fork	Delete
parsifal.cpc.wmin.ac.uk	jobmanager-fork	Delete
peyote.aei.mpq.de	jobmanager-fork	Delete
skirit.ics.muni.cz	jobmanager-fork	Delete
sr8000.lrz-muenchen.de	jobmanager-fork	Delete

Contact string: Job manager:

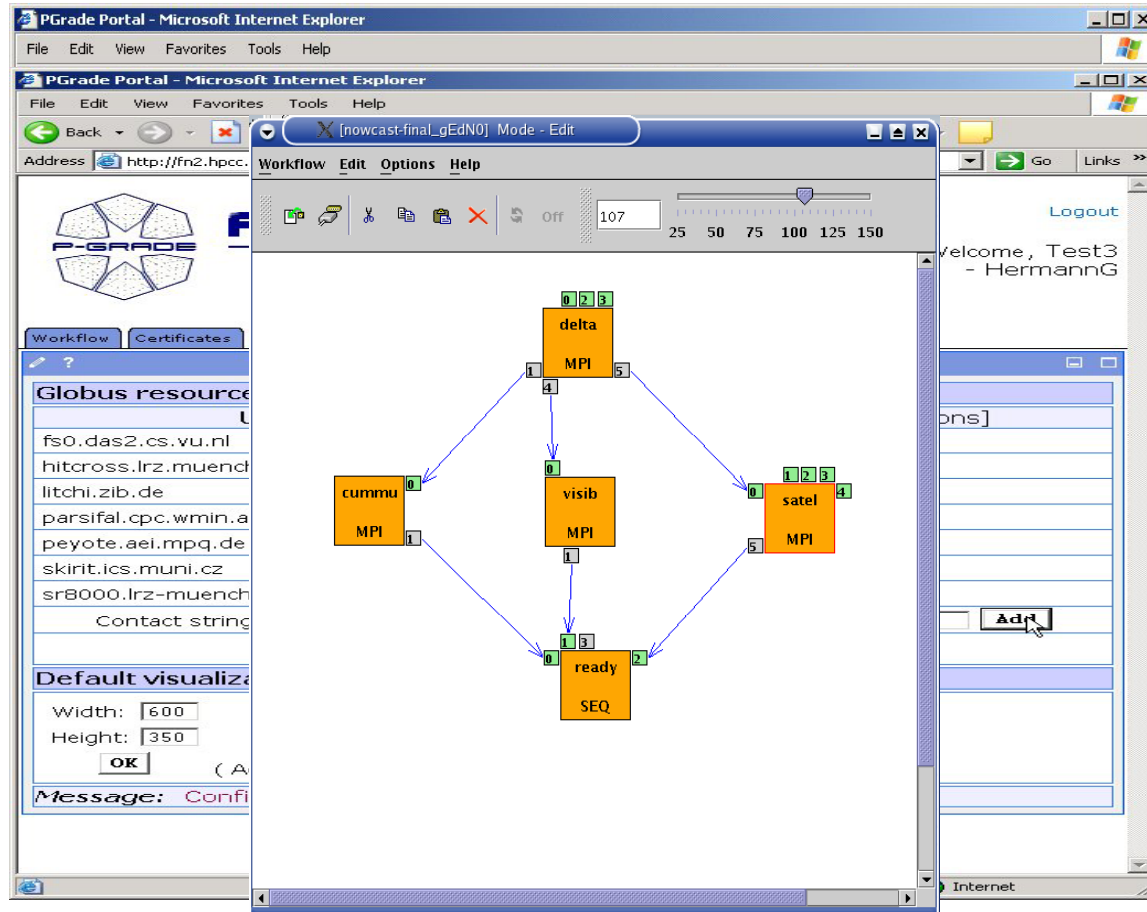
Default visualization size

Width:
Height:
 (Accept values between 150-1000.)

Message: Configuration successfully deleted.



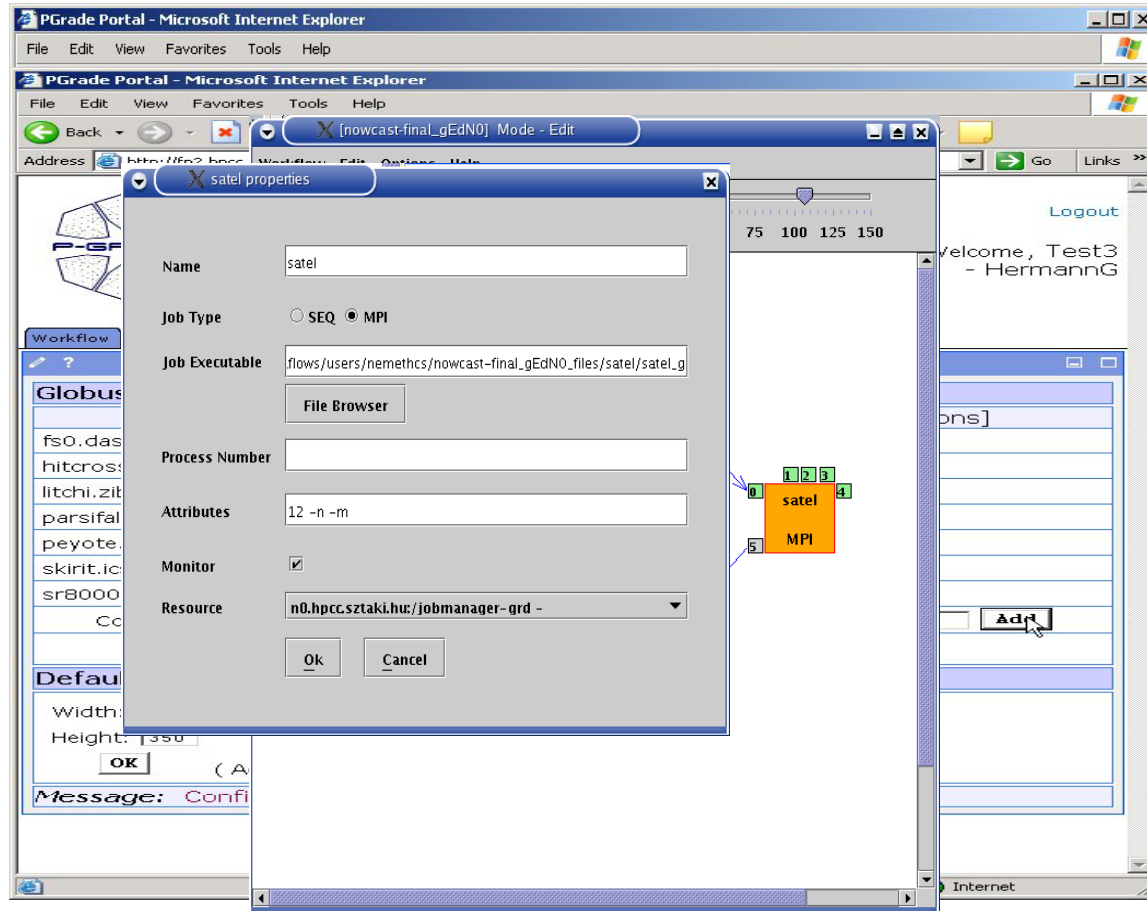
Graphical User Interfaces in a nutshell



- Proxy management
- Grid, VO and Grid resource management
- Graphical editor for workflow creation



Graphical User Interfaces in a nutshell



Proxy management

Grid, VO and Grid resource management

Graphical editor for workflow creation

Mapping job to Grids and Grid resources



Graphical User Interfaces in a nutshell

PS workflow details of 'Ax_EQUAL_B_seeGRID_broker__PS'

Statistics						Logs
Total	Init	Submitted	Rescue	Error	Finished	-
6	2	4	0	0	0	

eWorkflow list				
Workflow	Status	[Output]	[View]	[Action]
Ax_EQUAL_B_seeGRID_broker__PS.1	submitted	N/A	Details	Abort
Ax_EQUAL_B_seeGRID_broker__PS.2	submitted	N/A	Details	Abort
Ax_EQUAL_B_seeGRID_broker__PS.3	submitted	N/A	Details	Abort
Ax_EQUAL_B_seeGRID_broker__PS.4	submitted	N/A	Details	Abort
Ax_EQUAL_B_seeGRID_broker__PS.5	init	N/A	Details	Abort

Message: The list refreshed.

September 9, 2006

Proxy management

Grid, VO and Grid resource management

Graphical editor for workflow creation

Mapping job to Grids and Grid resources

Workflow management



Graphical User Interfaces in a nutshell

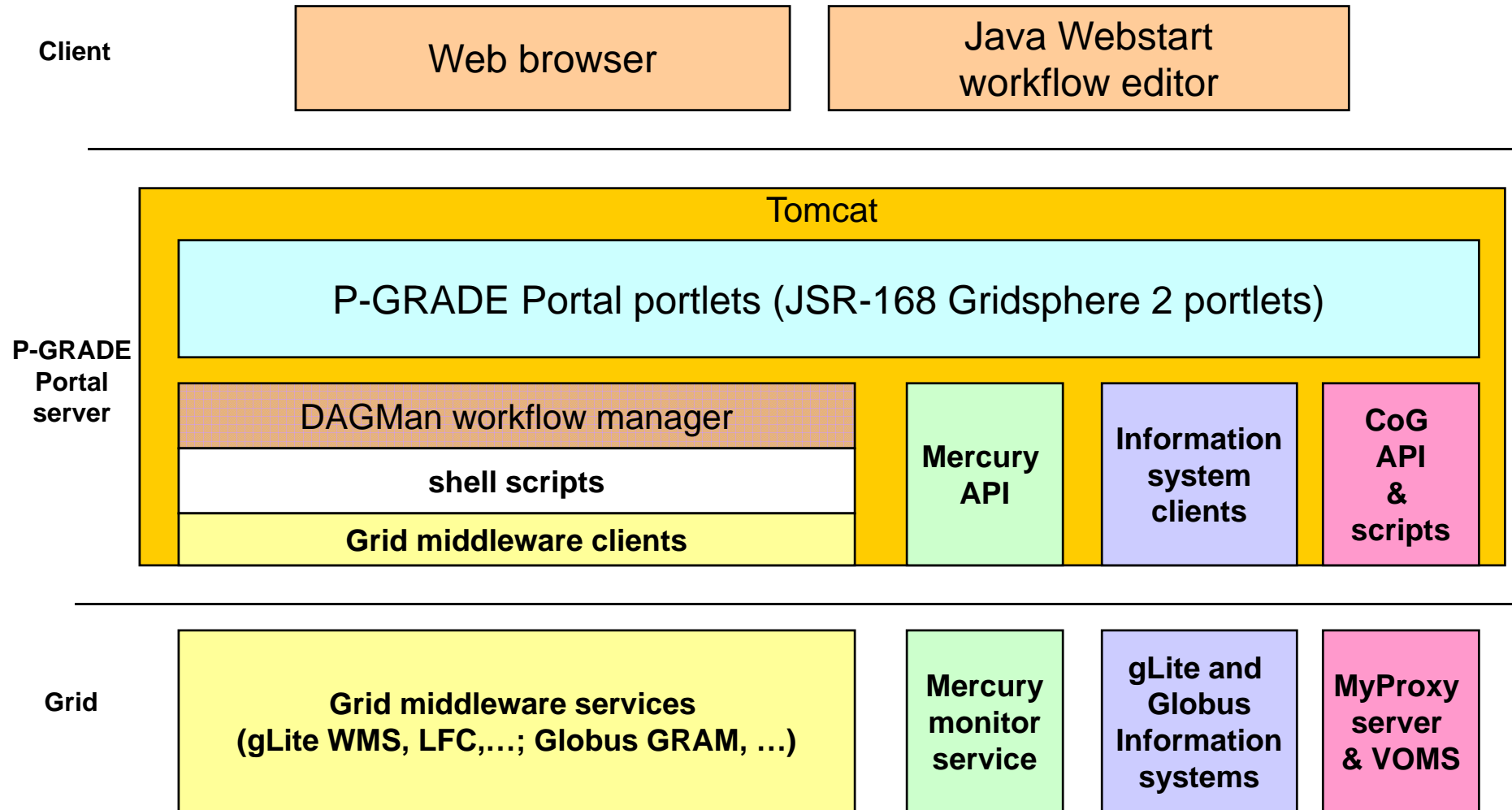
Workflow	Job	Hostname	Status	[Logs]	[Output]	[Action]
nowcast-final_gEdN0_b	cummu	n0.hpcc.sztaki.hu	3_finished	Out	-	Attach Delete
	delta	n0.hpcc.sztaki.hu	3_finished	Out	Err	
	ready	n0.hpcc.sztaki.hu	3_finished	-	Err	
	satel	n0.hpcc.sztaki.hu	3_finished	-	-	
	visib	n0.hpcc.sztaki.hu	3_finished	Out	-	

The 'Tracefile visualization' section shows a workflow diagram for 'nowcast-final_gEdN0_b' with a timeline from 0s to 6m0s. The diagram includes nodes for 'visib', 'delta', 'ready', 'satel', and 'cummu' on the n0.hpcc.sztaki.hu host, connected by arrows indicating the flow of the workflow.

- Proxy management
- Grid, VO and Grid resource management
- Graphical editor for workflow creation
- Mapping job to Grids and Grid resources
- Workflow management
- Execution visualization



Implementation overview





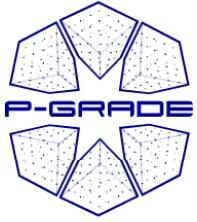
Some applications gridified with P-GRADE 2.5 by SZTAKI

- Ultra short range weather forecast (MEANDER): **workflow** that integrates 4 meteorological algorithms and one visualizer component
- Road traffic simulation: predict the density of cars on the roads of Manchester. **Workflow** that integrates 4 simulator components
- Minimizing operational cost of factories and logistic service providers (EMMIL): **Parametric workflow** resulting thousands of short running jobs
- Molecular Dynamics Study of Water Penetration (CHARMM): **Parametric workflow** resulting hundreds of long running jobs
- Studying oscillons and magnetic monopole configurations: **Parametric workflow** resulting hundreds of short running jobs



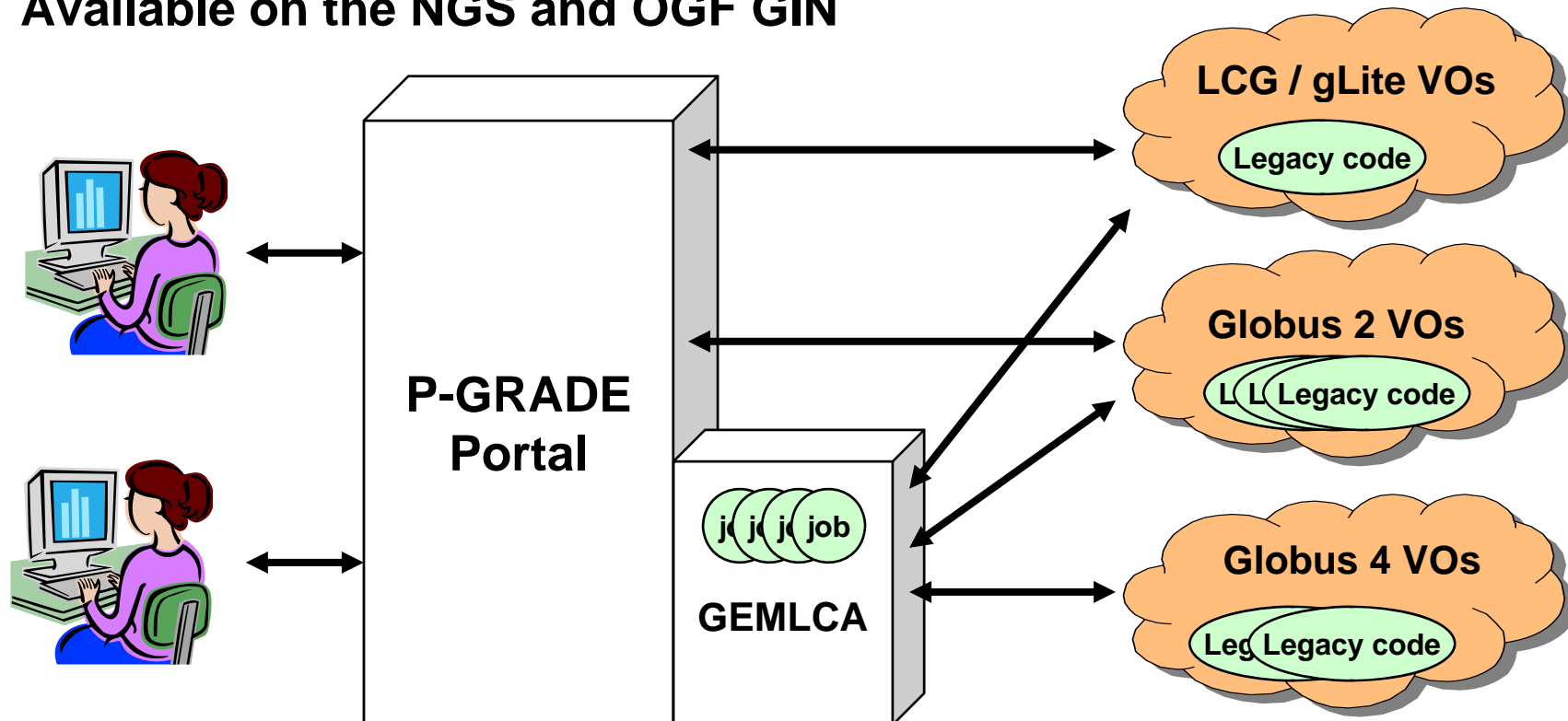
Outline

- **P-GRADE Portal and Developer Alliance**
- **Capabilities of P-GRADE Portal 2.5**
 - Workflow support
 - Parametric study support
- **Other portal versions**
 - GEMLCA P-GRADE Portal
 - Application specific portals based on P-GRADE
- **Hand-on exercises**
- **How to get access**
- **Roadmap**
 - P-GRADE grows into gUSE
- **Summary**



GEMMLCA P-GRADE Portal 2.4.1

- P-GRADE Portal extended with GEMMLCA Grid service back-end
 - To share jobs and legacy codes as application components with others
 - A step towards collaborative e-Science
- Developed by the University of Westminster (London)
- Support for Globus 4 grids (besides GT2 and EGEE)
- Available on the NGS and OGF GIN





Application Specific Portals

“Make the grid disappear”

The Rendering Portal

RELEASE 2.5



Rendering

Create a new rendering job by uploading file to be rendered

New rendering job | Manage rendering jobs | Help

Name:

Type:

Start frame:

End frame:

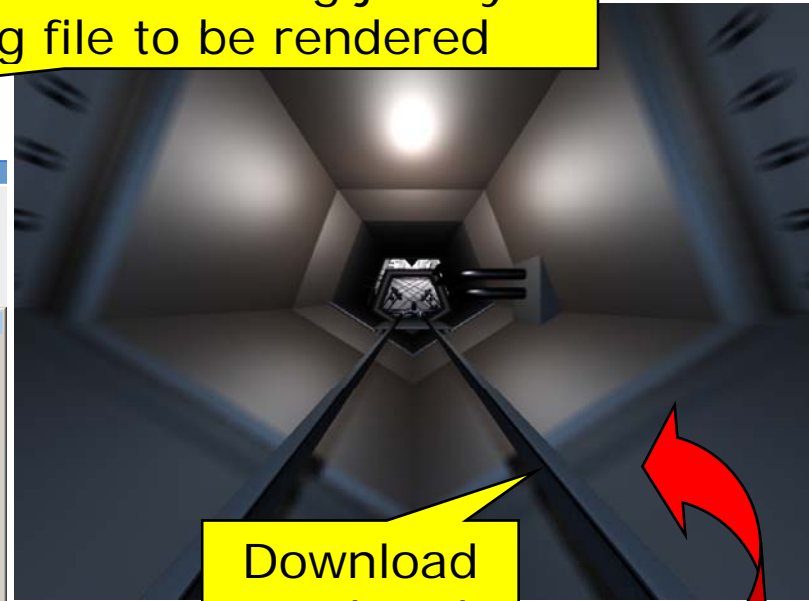
Step size:

Upload file from local pc:

File Upload

Look in: EGEE 07 demo

- ~\$E07 UoW demos
- Document Scrap 'http__mut_cpc_w...'
- EGEE07 UoW demos
- GEMLCA_EGEE2007Conf
- knight.blend
- knight-fast.blend
- longer.blend
- lostride.blend



Download rendered images

Workflow Manager

Workflow list					
Workflow	Status	Size	Quota (100 Mb)	[Output]	[View]
Ball	finished	220 KB	0.22%	<input checked="" type="checkbox"/>	<input type="button" value="Details"/>
Knight	finished	1.17 MB	1.17%	<input checked="" type="checkbox"/>	<input type="button" value="Details"/>
Lostride	finished	11.175 MB	11.18%	<input checked="" type="checkbox"/>	<input type="button" value="Details"/>
Overall used quota:		12.561 MB	12.56%		

Submit workflow



Application Specific Portals

“Make the grid disappear”

Supplier chain portal

Welcome Emmil

Emmil 3.0

Download finished Auction's results

Welcome Emmil

Emmil 3.0

ID

MESSAGE

Available Auctions for Suppliers

ID	Auction Adviser	Date/Time	Remaining Time	Number of Bids	Set Sup
0	Buyer_01	Fri Sep 28 13:49:51 CEST 2007	0 weeks, 6 days, 5 hours, 58minutes, 12 seconds	1	

Available Auctions for 2D1

ID	Auction Adviser	Date/Time	Remaining Time

Closed

ID	Auction Adviser	Date/Time

Microsoft Excel -

Product #	Quantity	Unit Price	Value of order	Delivery Variable Cost	Price
0	25	140.65	3516.26	209.34	3725.59523
3	70	19.14	1339.54	586.15	1926.68952
5	50	18.65	932.28	418.68	1350.952805
7	5	939.68	4698.39	41.87	4740.255005
8	50	39.15	1957.69	418.68	2376.370132
1	31	32.49	1007.12	259.58	1266.702419
3	130	19.14	2487.65	1088.55	3576.20147
4	100	25.95	2594.84	837.35	3432.187057
Summ Quantity	461			Delivery Fix Cost per container	373.14
Container Number	5			Delivery Fix Cost	1865.70
				Price for Supplier 4	22767.09
				TOTAL PRICE	153088.56

Flow Editor - [EMMIL_30_P5] Mode - Edit

Generator GEN

LP_solve SEO

PriceEv... COLL

Parameter study optimization workflow is automatically generated and executed

Set up the most optimal supplier chain

Log in as buyer supplier, or logistic service provider. Create auctions or take your bids

Best buyer-seller-logistic service provider pairs discovered, result presented in an Excel file



Outline

- **P-GRADE Portal and Developer Alliance**
- **Capabilities of P-GRADE Portal 2.5**
 - Workflow support
 - Parametric study support
- **Other portal versions**
 - GEMLCA P-GRADE Portal
 - Application specific portals based on P-GRADE
- **Hand-on exercises**
- **How to get access**
- **Roadmap**
 - P-GRADE grows into gUSE
- **Summary**

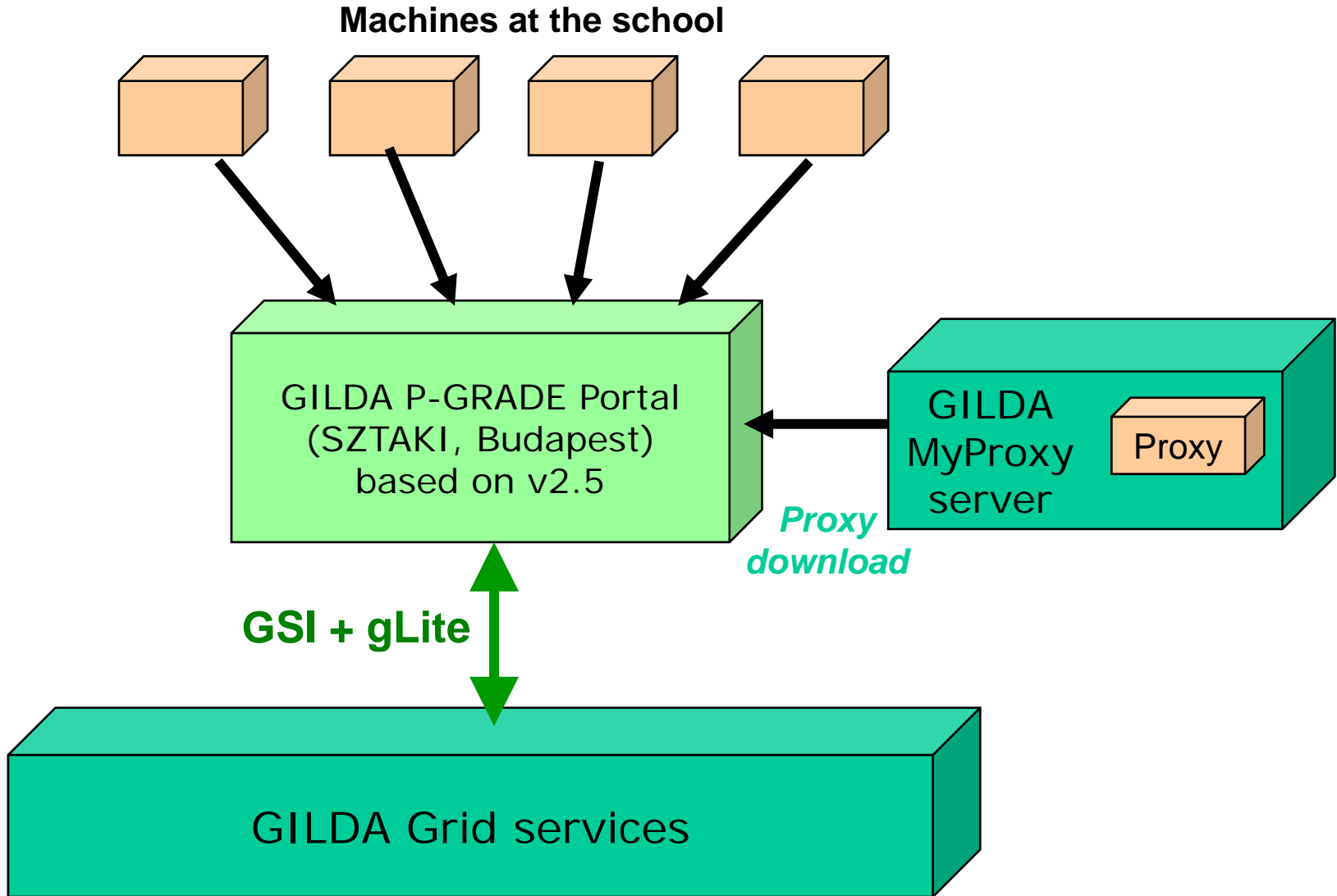


Hands-on

- Aim is to learn the basics of P-GRADE Portal workflows
 - Understand the concept of data driven workflow applications
 - Go through the typical workflow application development cycle
- Grid to be used:
 - GILDA (gLite)



Infrastructure for the hands-on

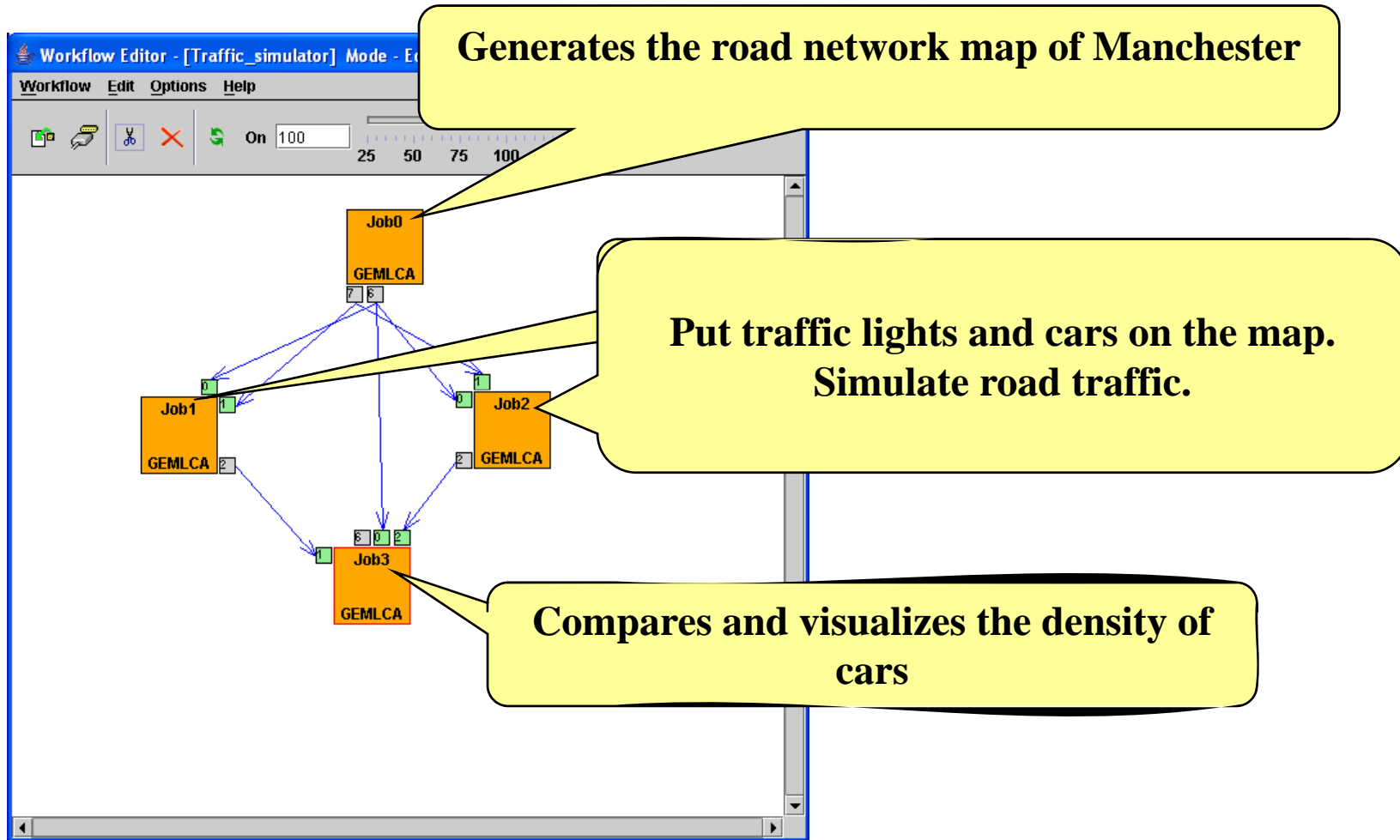




Exercise 1

Import and execute a pre-defined application

Traffic simulation

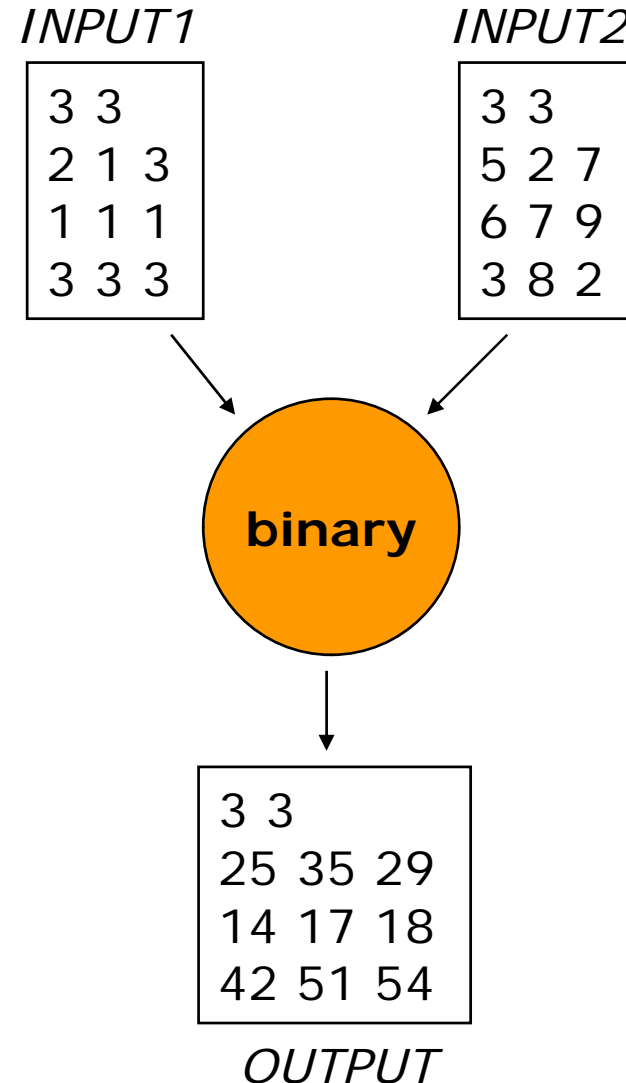




Exercise 2

Execute a matrix multiplication job on GILDA

- C code – Familiar from Monday
- Compiled on GILDA UI
- Reads matrixes from two files called **INPUT1** and **INPUT2**
- Writes result matrix into file called **OUTPUT**
- Controllable by command line parameters
- Local execution:
`./multiply M V`





Exercise 3

Save the multiplication **OUTPUT** on a SE

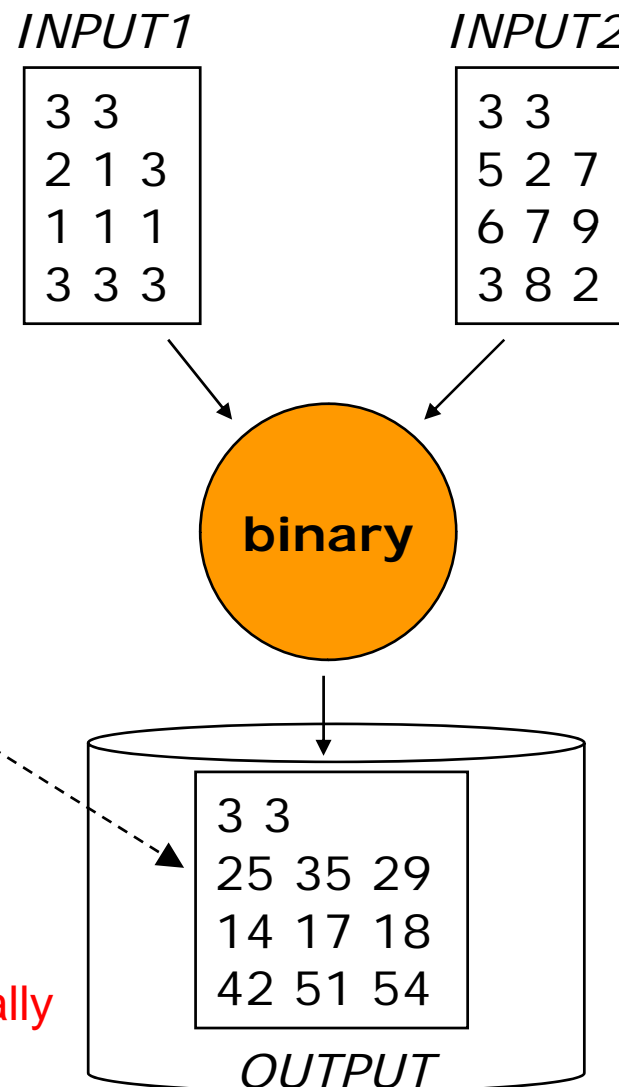
- Modify output file type from “Local” to “Remote”
- Specify a logical file name as target location:

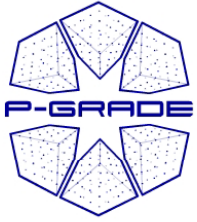
`lfn:/grid/gilda/. . .`

`lfn:/grid/gilda/budapest12/...`

Logical location defined by the user
Registered in LFC automatically

Storage Element selected automatically

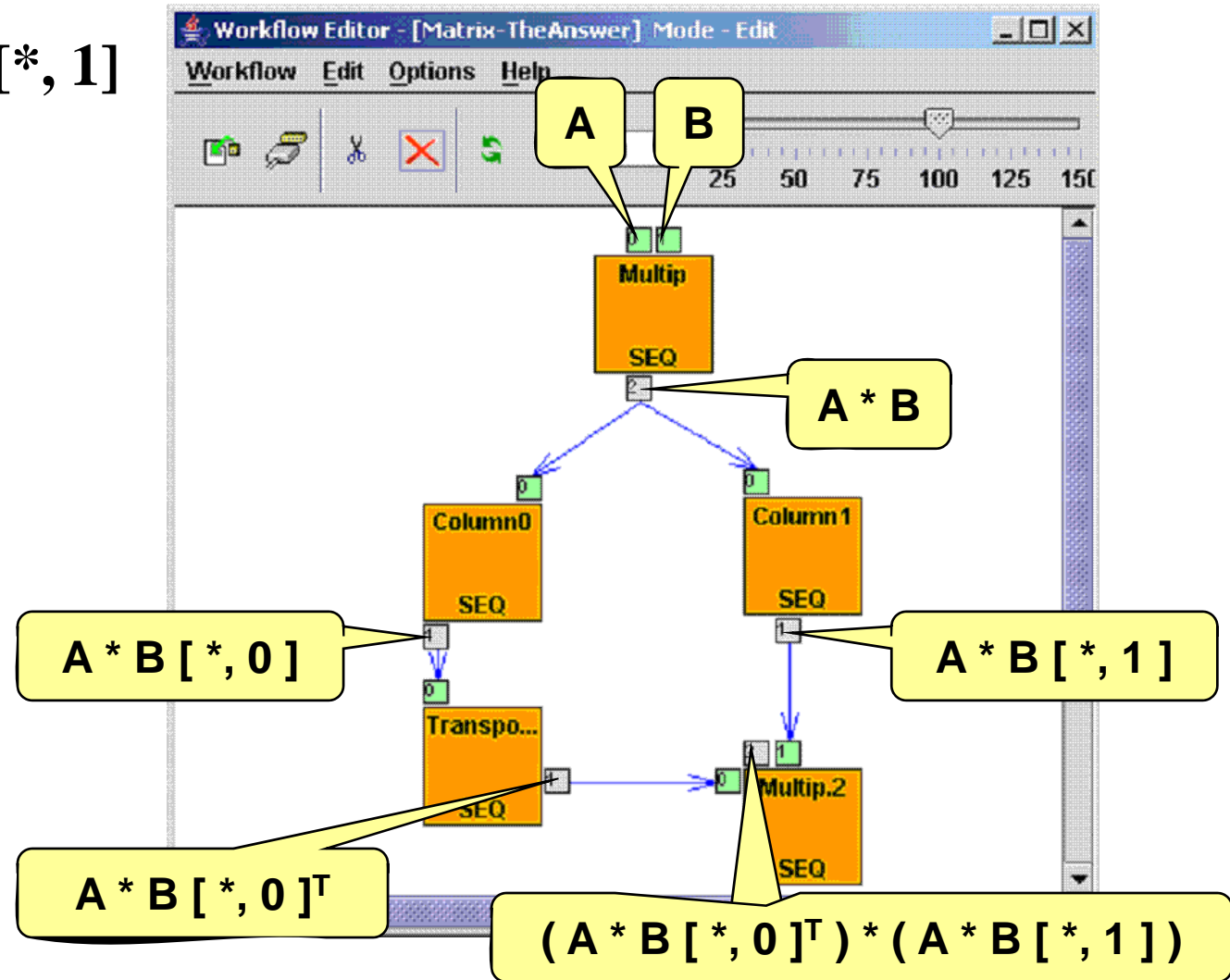




Exercise 4

Define a complex workflow using the MatrixOperations job

- $AB[* , 0]^T * AB[* , 1]$





Tips and tricks

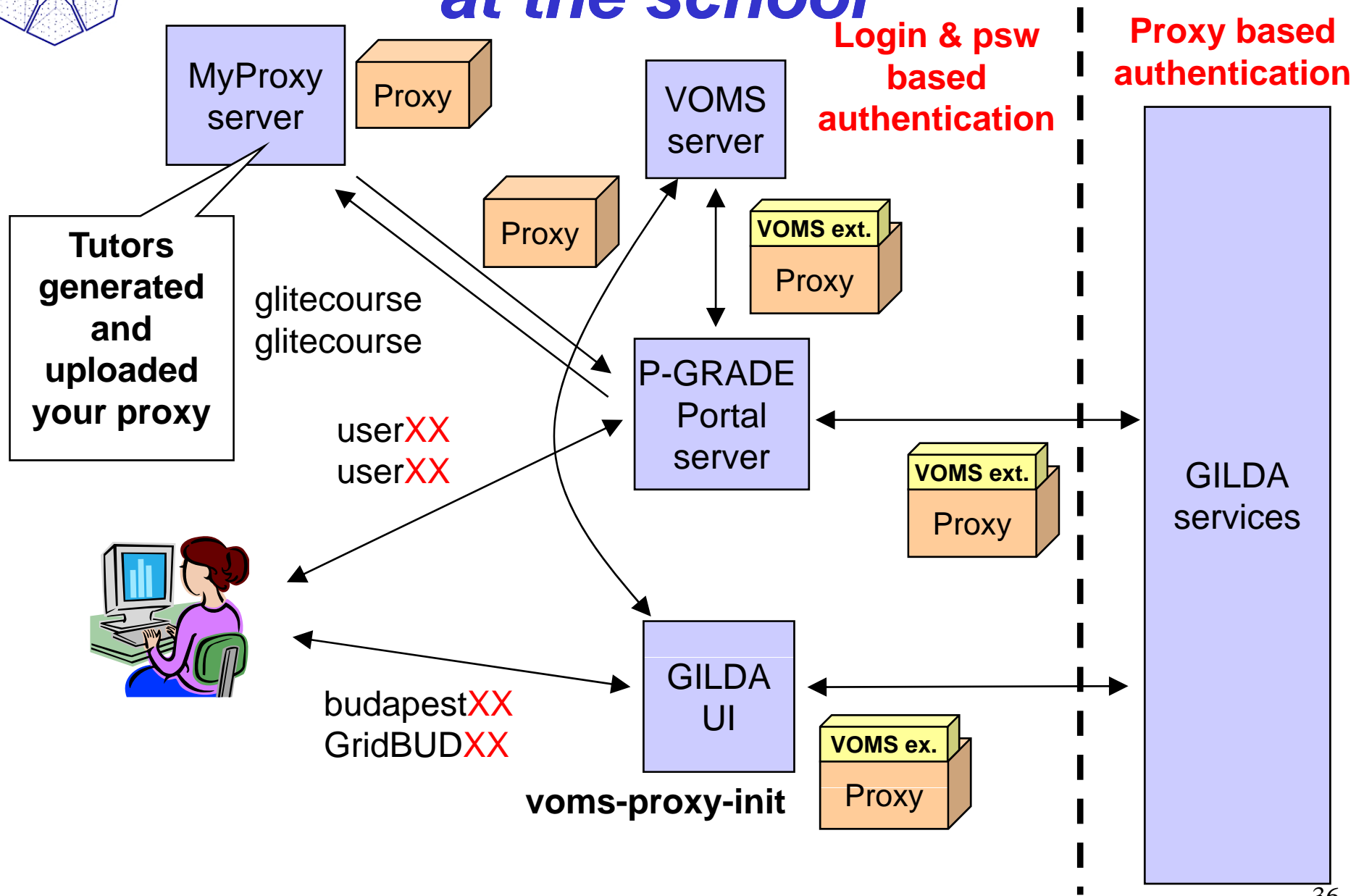
- **Exclude an erroneous site from the brokering process:**
 1. Open the JDL editor of the job
 2. Write

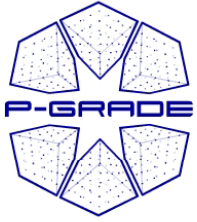
```
other.GlueCEInfoHostname != "hostname"
```

into the “Ranks & requirements” window
- **Automatic job resubmission:**
 1. Open the JDL editor of the job
 2. Set **Retry count** to a higher value



User authentication at the school





Open the tutorial

Skip the “Upload your certificate...” section

Portal and Myproxy accounts:

Portal accounts: user01 – user30

Portal passwords: user01 – user30

MyProxy download:

Hostname: n40.hpcc.sztaki.hu

Account: glitecourse

Password: glitecourse



Outline

- **P-GRADE Portal and Developer Alliance**
- **Capabilities of P-GRADE Portal 2.5**
 - Workflow support
 - Parametric study support
- **Other portal versions**
 - GEMLCA P-GRADE Portal
 - Application specific portals based on P-GRADE
- **Hand-on exercises**
- **How to get access**
- **Roadmap**
 - P-GRADE grows into gUSE
- **Summary**



Portal installations

- **P-GRADE Portal services:**

- **SEE-GRID infrastructure**
- **Central European VO of EGEE**
- **GILDA: Training VO of EGEE**
- **Many national Grids (UK National Grid Service, HunGrid, Turkish Grid, etc.)**
- **US Open Science Grid, TeraGrid**
- **Economy-Grid, Swiss BioGrid, Bio and Biomed EGEE VOs, BalticGrid**
- **OGF Grid Interoperability Now (GIN) VO**

Portal services and account request:

- portal.p-grade.hu/index.php?m=5&s=0





Support services

<http://portal.p-grade.hu>

- Training:
 - Training events, trainer community:
email list: pgrade-training@lpds.sztaki.hu
 - Test portal based on the GILDA VO
 - Online tutorials
- User forum
- Installation and operational support team

www.lpds.sztaki.hu/gasuc

- Application Development and Porting Support:



Roadmap

From P-GRADE to gUSE

- P-GRADE:
portal interface + high level middleware services
- **gUSE: Grid User Support Environment**
 - ~ **P-GRADE services exposed as Web Services**
 - Scalable
 - Flexible
 - Extendable
 - Autonomous
 - Replaceable
 - **New features**
 - Loops at workflow level
 - If-then-else structures at workflow level
 - Nested workflows
- gUSE release for Supercomputing'07:
November 10



Lessons learnt

- Workflows need parametric study support ✓
- Portals must be easily customizable for applications ✓
- P-GRADE must be open source **Started**
- Workflows need loops, if-then-else structures **Next release**
- Job failure rate can be high, grid error messages can be rude: failure management layer required **Future work**



Summary & conclusions

- **Benefits of being member of the user community**
 - Short learning curve → Swift uptake of grid technology
 - Graphical access → Protection against cmd line and API changes
 - High level, abstract tools → easy to perform complex operations (e.g. file transfer + LFC update)
 - Support services
- **Benefits of being member of the developer community**
 - Customizable to certain user communities
 - Customizable to certain applications
 - Customizable to certain middleware releases
- **Benefits of keeping your eyes on P-GRADE activities**
 - Regular releases with remarkable new features
 - Broadening user community → new applications
 - Broadening developer community



Learn once, use everywhere
Develop once, execute anywhere

Questions?

portal.p-grade.hu
pgportal@lpds.sztaki.hu