

A Grid programozása felhasználóbarát módon



Peter Kacsuk

MTA SZTAKI

www.lpds.sztaki.hu

User concerns of Grid systems

- Fast evolution of Grid systems and middleware:
 - **EGEE, NorduGrid, TeraGrid, Grid2003, UK NGS, etc.**
 - **GT1, GT2, OGSA, OGSF, GT3, WSRF, GT4, ...**
- How to cope with the variety of these Grid systems?
- How to develop/create new Grid applications?
- How to execute Grid applications?
- How to observe the application execution in the Grid?
- How to tackle performance issues?
- How to port legacy applications
 - to Grid systems
 - between Grid systems?
- How to execute Grid applications over several Grids in a transparent way?

Solution

- A high-level program development and execution environment for the Grid
- Supports the development and execution of
 - Parallel Grid applications
 - Workflow applications
- Hides the low level details of the Grid access mechanism
- Hides the variety of Grid systems, i.e., **learn once and use everywhere**
- Supports the transparent migration of applications among different Grid sites (fault-tolerance, load-balance)
- Supports the monitoring and visualization of Grid application execution
- **P-GRADE** for Grid program development
- **P-GRADE portal** for Grid program execution



P-GRADE

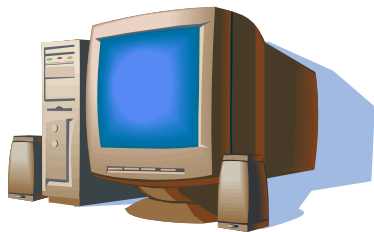
**Parallel Grid Run-time and
Application Development
Environment**

THE BIG ISSUE in HPC GRIDS



Sequential
components
(code)

C, C++
Fortran



Parallelization?



Debugging?



Performance?



Execution?



Parallel
components
(code)

?



P-GRADE: Unified Solution for Distributed supercomputing

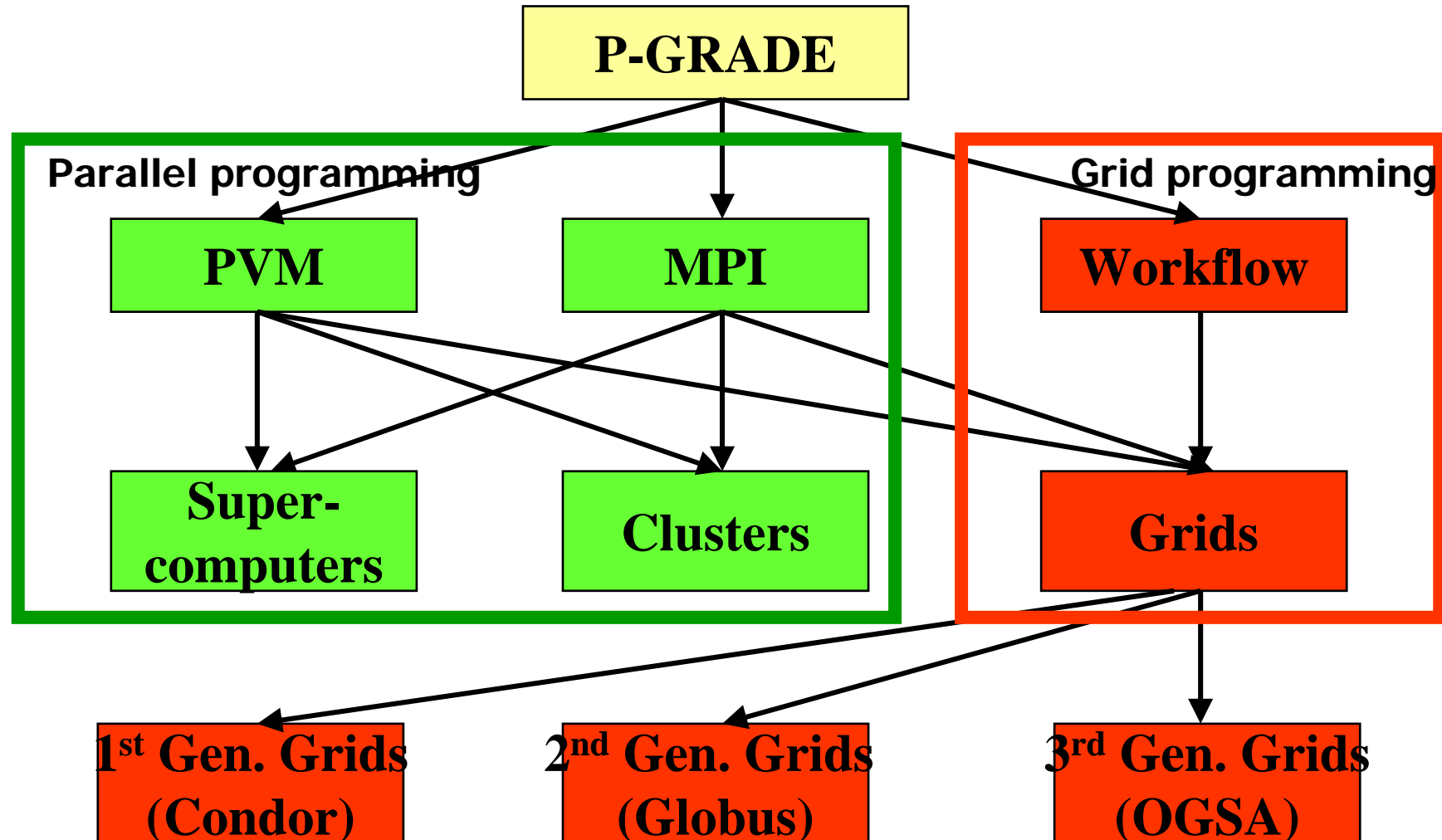
- **P-GRADE**

Parallel **G**rid **R**un-time and
Application **D**evelopment **E**nvironment

- A highly integrated parallel application development and execution system for clusters and Grids
- Provides:
 - Parallel, supercomputing, Grid programming
 - Fast and efficient development of Grid programs
 - Observation and visualization of Grid execution
 - Fault and performance analysis support for

SUPERCOMPUTERS
CLUSTERS
GRIDS

Goal: Use P-GRADE for the whole range of parallel/Grid systems



Parallel programming paradigms and P-GRADE support



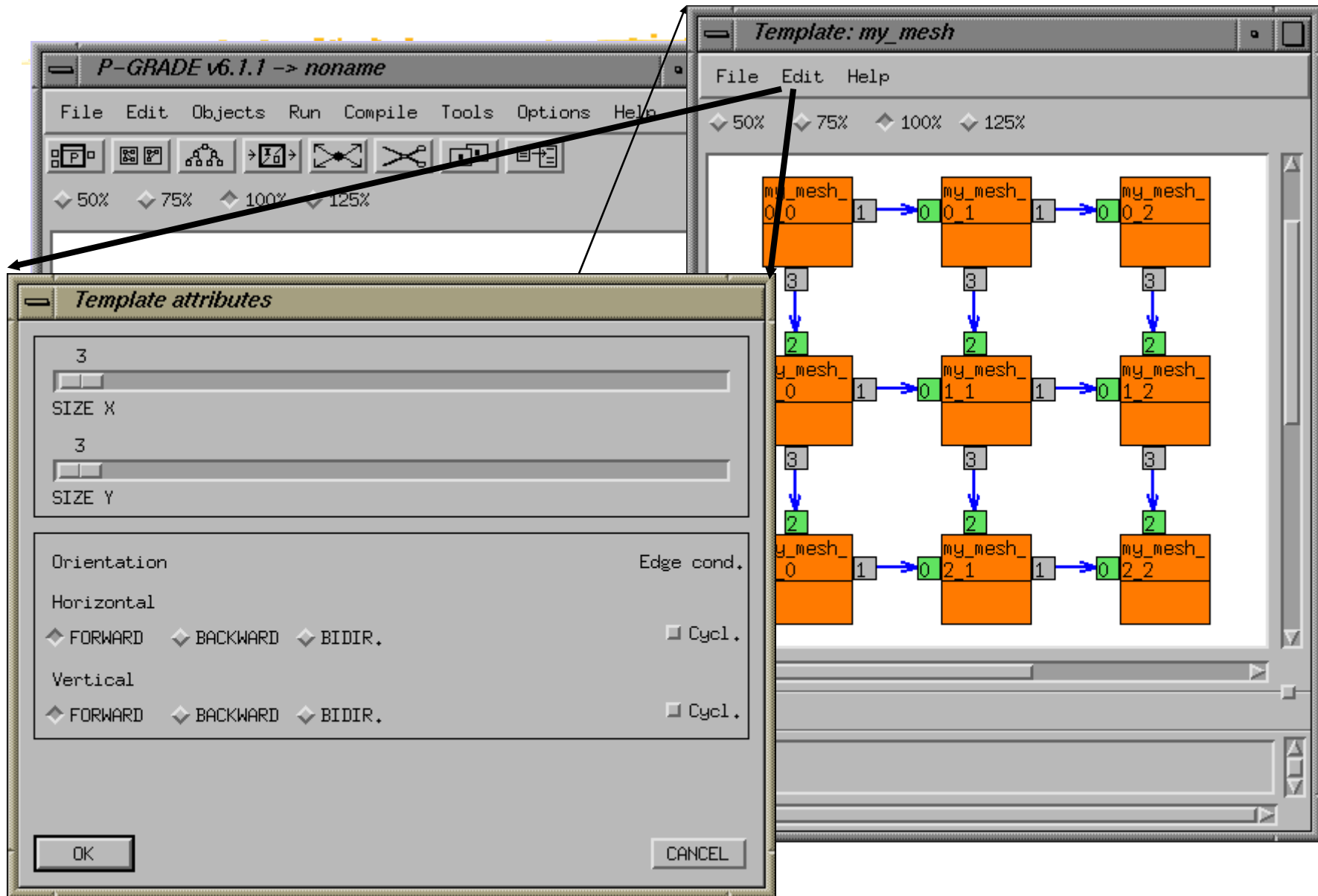
Paradigm

- Master/worker
(parameter study)
- Pipe-line programming
- Neighborhood-oriented,
Cellular programming

P-GRADE template

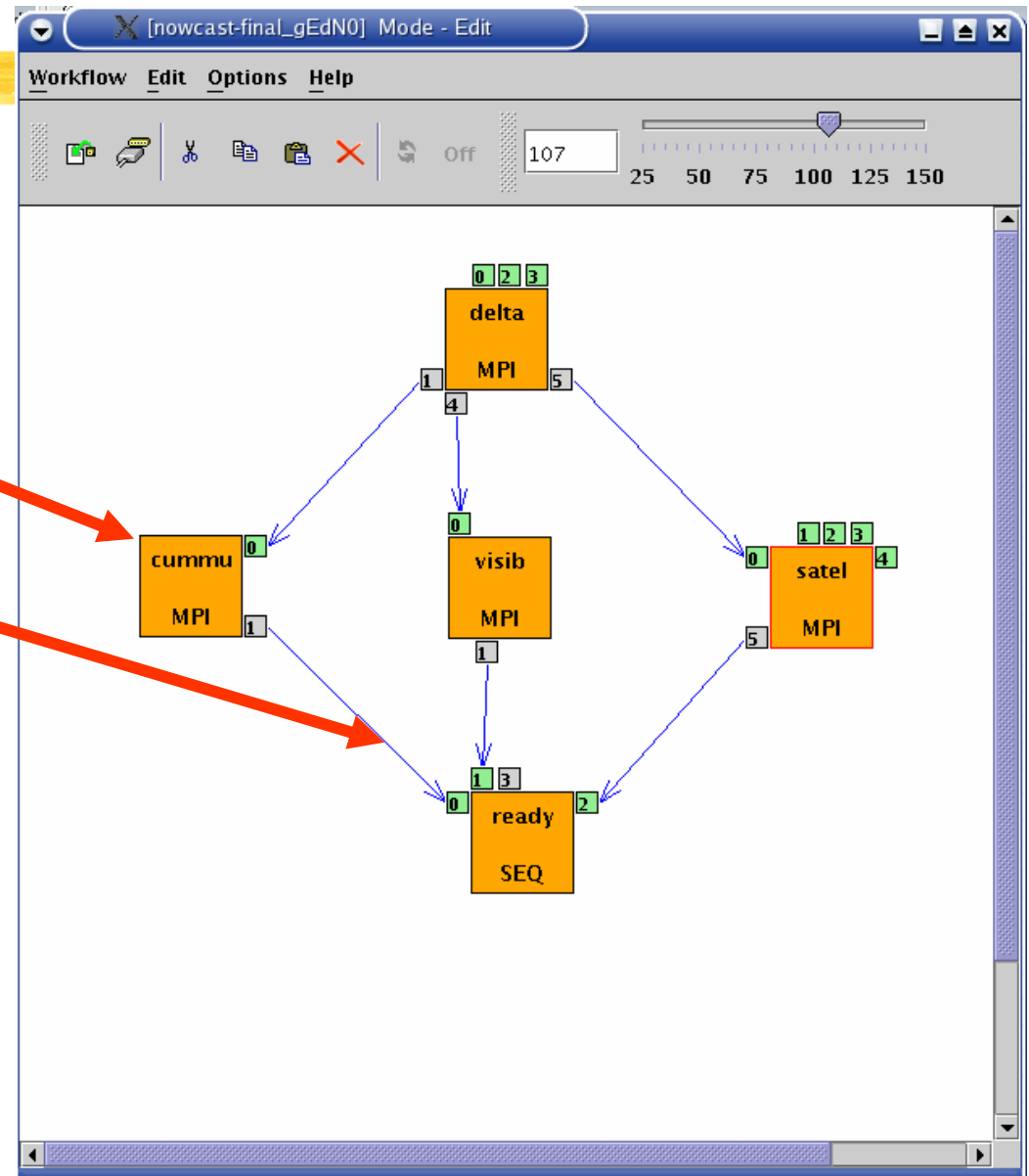
- Processor farm
- Pipe
- Mesh

Mesh Template



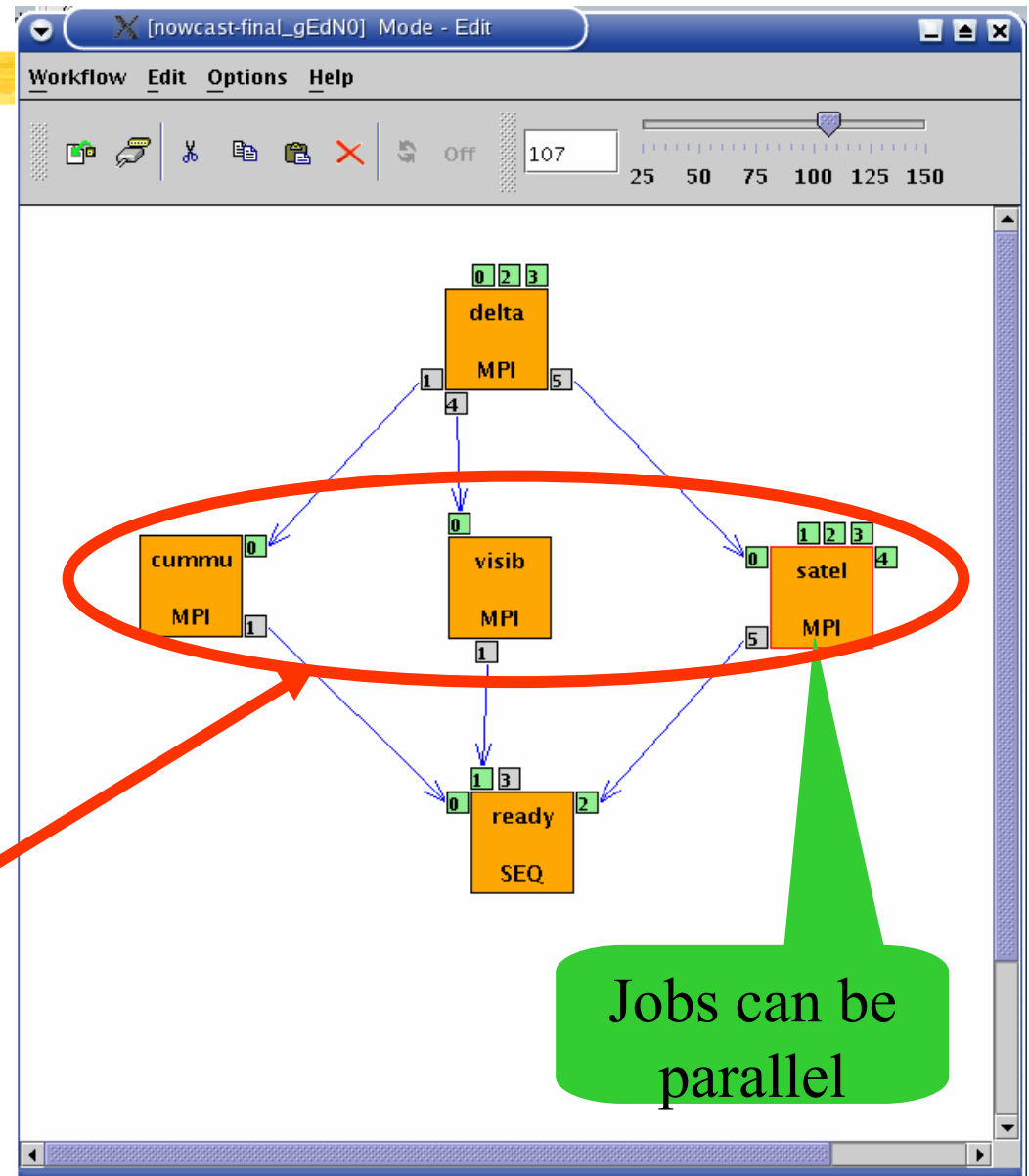
Workflow support in P-GRADE

- The workflow is a graph where
 - Nodes are jobs (or services)
 - Arcs represent file transfer between the jobs (services)
- Semantics of the workflow:
 - Job can be executed if all the necessary file transfers represented by the arcs are completed

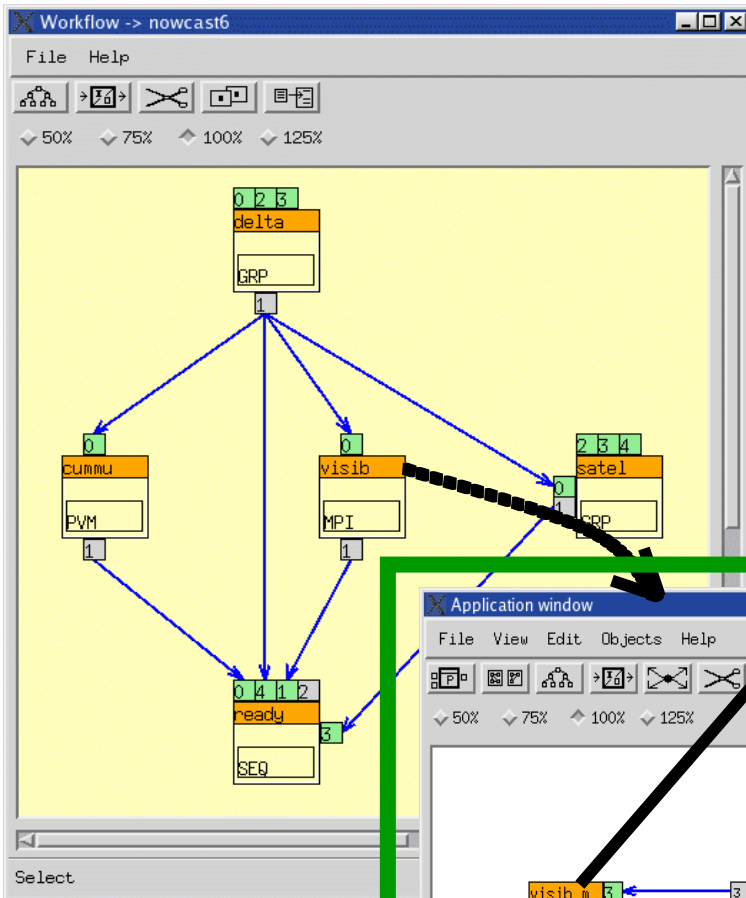


Two level parallelism by a workflow

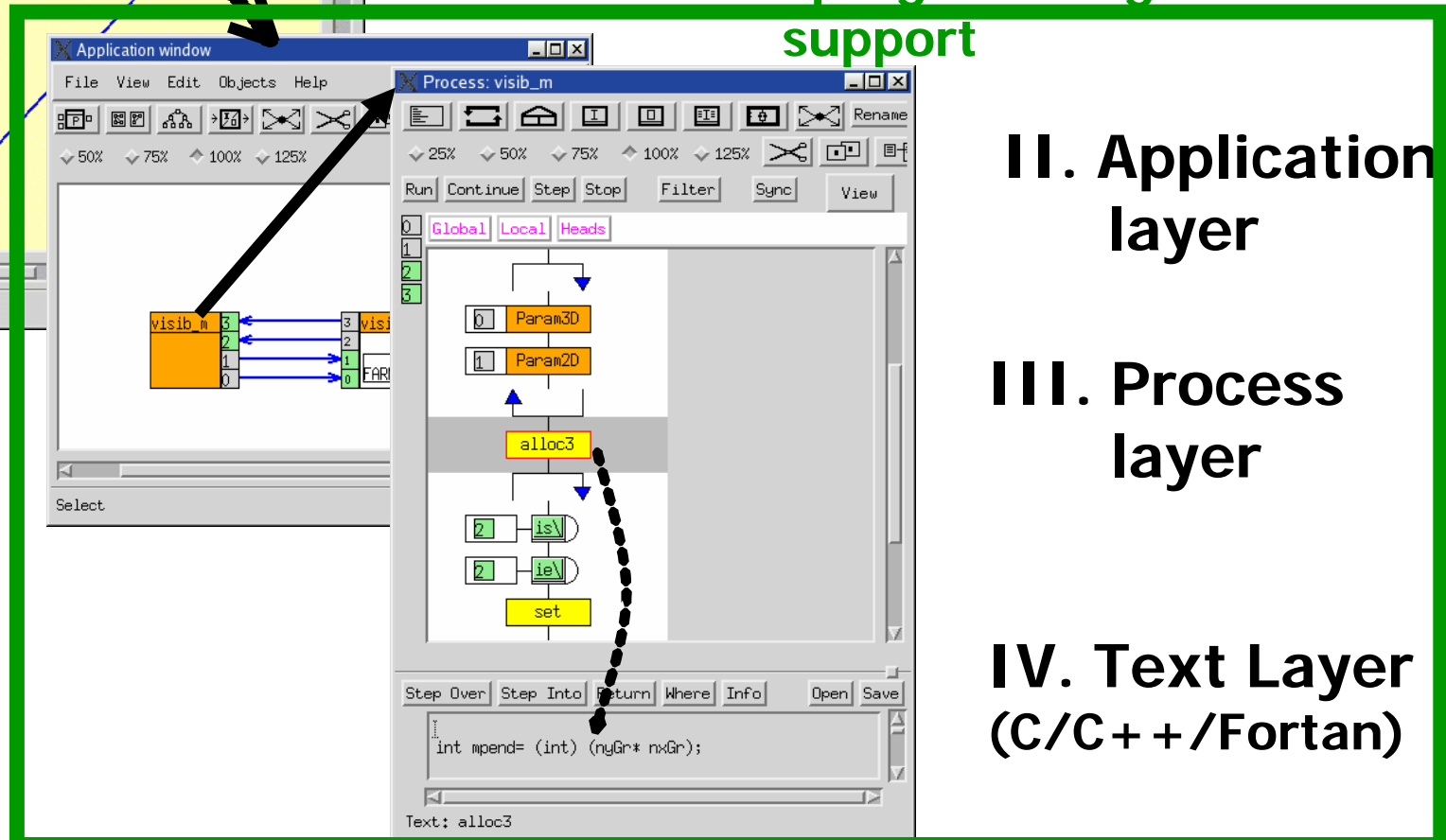
- The workflow concept enables the efficient solution of complex problems in a distributed environment like Grid
- Semantics of the workflow enables **two levels of parallelism**:
 - Parallel execution inside a workflow node
 - Parallel execution among workflow nodes



RADE in a nutshell



Parallel programming
support




I. Workflow
layer

II. Application
layer

III. Process
layer

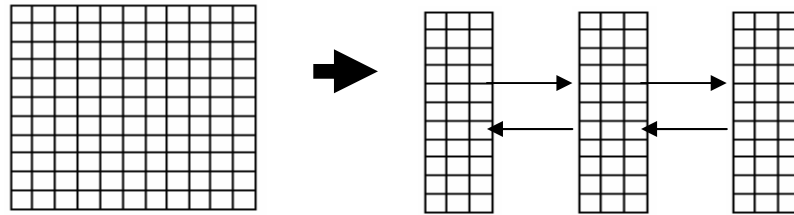
IV. Text Layer
(C/C++/Fortan)



Levegőminőségi (LM) alkalmazás

Alapok

- Cél: reakció-diffúzió-advekción egyenletek megoldása
- Alkalmazás: passzív nyomanyagok (pl. radioaktív nuklidok) terjedésének szimulációjára



- Párhuzamosítás folyamata
 - terület felosztása az egyes processzorok között
 - szimulációs lépésenként a határfeltételek kölcsönös kicserélésével

Párhuzamos alkalmazás P-GRADE -ben

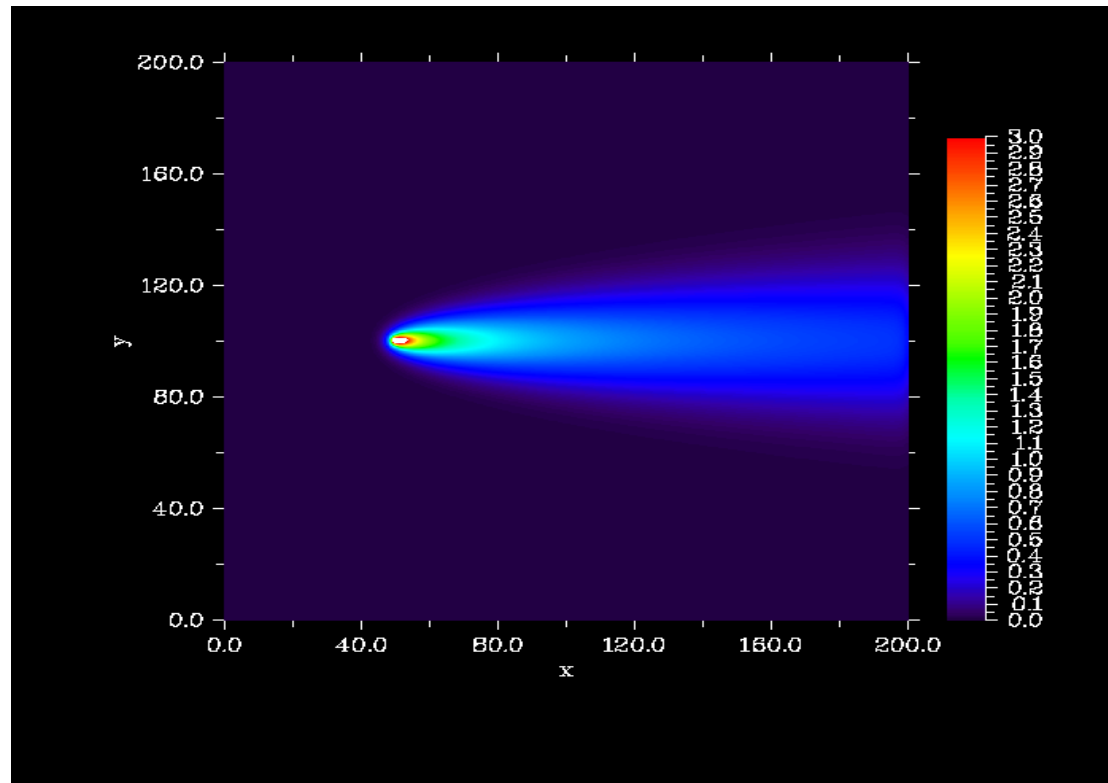
The image displays the P-Grade software interface for a parallel application. The main window shows a process graph with **LM_MAIN** and **LM_SIM** components connected via a **PIPE**. Below this, the **Process: LM_MAIN** window shows a control flow graph with nodes like **init cond**, **co**, **re**, and **result**. The **Process: LM_SIM -> LM_SIM_1** window provides a detailed view of the execution flow, including nodes for **init_1**, **init_2**, **RD_solver**, and **init3**. The **Prove** window at the bottom shows a trace of the execution, with a timeline from 5.4s to 6.5s and a list of processes: **LM_SIM_2 (2)**, **LM_SIM_1 (1)**, **LM_SIM_0 (0)**, and **LM_MAIN (0)**. The **I/O Files Dialog** window shows the file names and abstract names for the application. The **Text: init cond** window displays the source code for the initialization condition.

```
fscanf(fin, "%s %s", filename, fileid);
for(k=0;k<ncom;k++) fscanf(fin, "%s", &fileid[k]);
grp_printf("deltak: ");
for(k=0;k<ncom;k++) printf("%s ", fileid[k]);
fscanf(fin, "%s %lf", &h);
```

```
for(i=1;i<=imax;i++) {
    for(k=0;k<ncom;k++) {
        c[old][k][i][0]=c[old][k][i][jmax];
    }
}
```

Eredmény: szennyeződés csóvák

- $h = 1000$ m (forrás magassága)
- $\Delta t = 900$ s (időlépcső)
- $D = 50$ m²/s (kibocsátás)
- $u = 5$ m/s (szélmező vízszintes komponense)
- $v = 0$ m/s (szélmező függőleges komponense)





P-GRADE portal



P-GRADE portal in a nutshell

The screenshot displays the P-Grade portal interface across multiple browser windows. The main window shows a workflow editor with a central node labeled 'delta' and 'MPI'. Below it, a 'Job list' table provides details on job execution. A 'Tracefile visualization' window at the bottom shows a timeline of job execution across various hosts.

Workflow	Job	Hostname	Status	Logs	Output	Action
nowcast-final_gEdN0_b	summu	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	

Proxy management

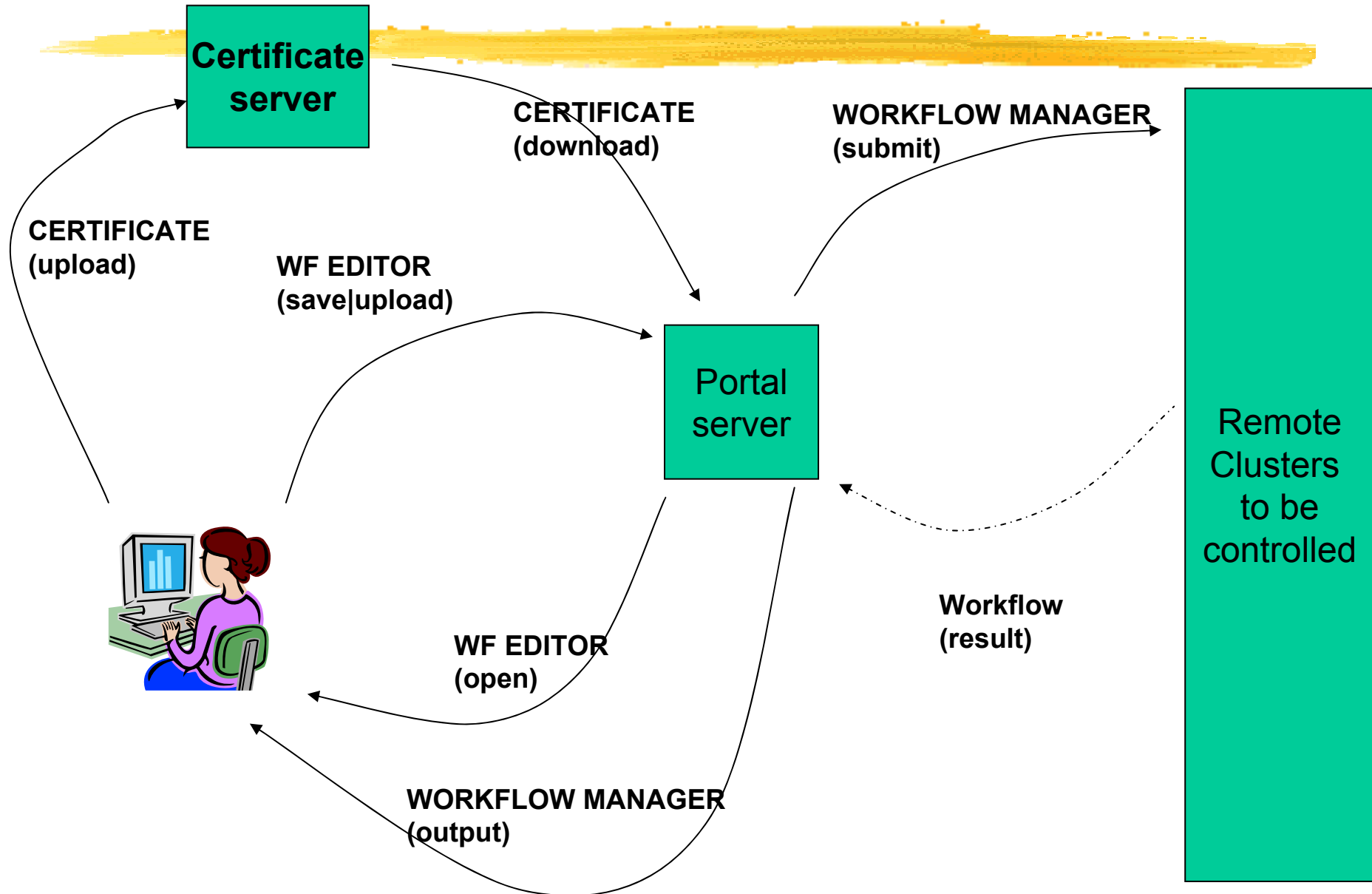
Definition of Grid resources

Workflow creation

Job mapping to Grid resources

Workflow management and execution visualization

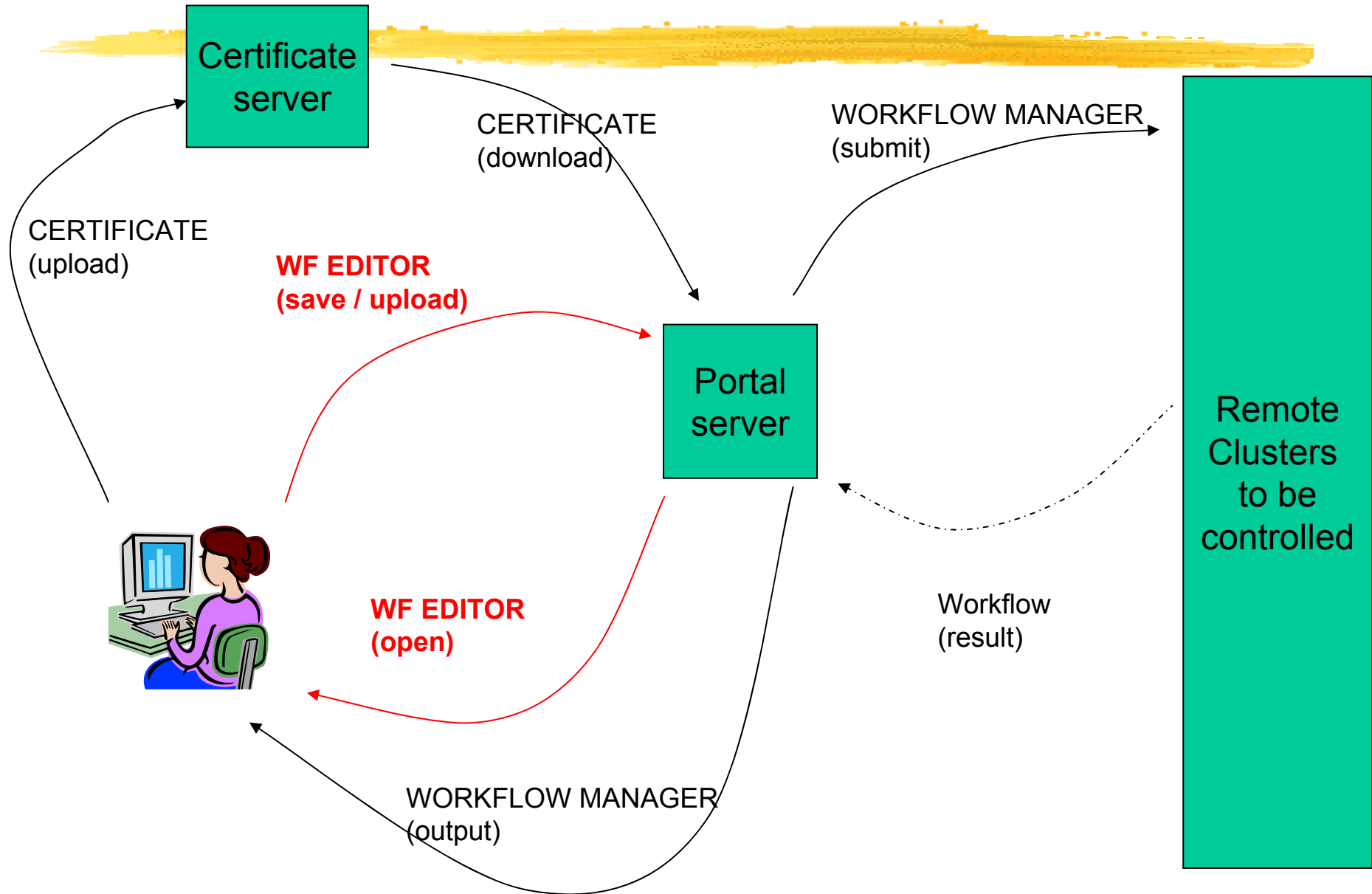
Principles of the P-GRADE portal



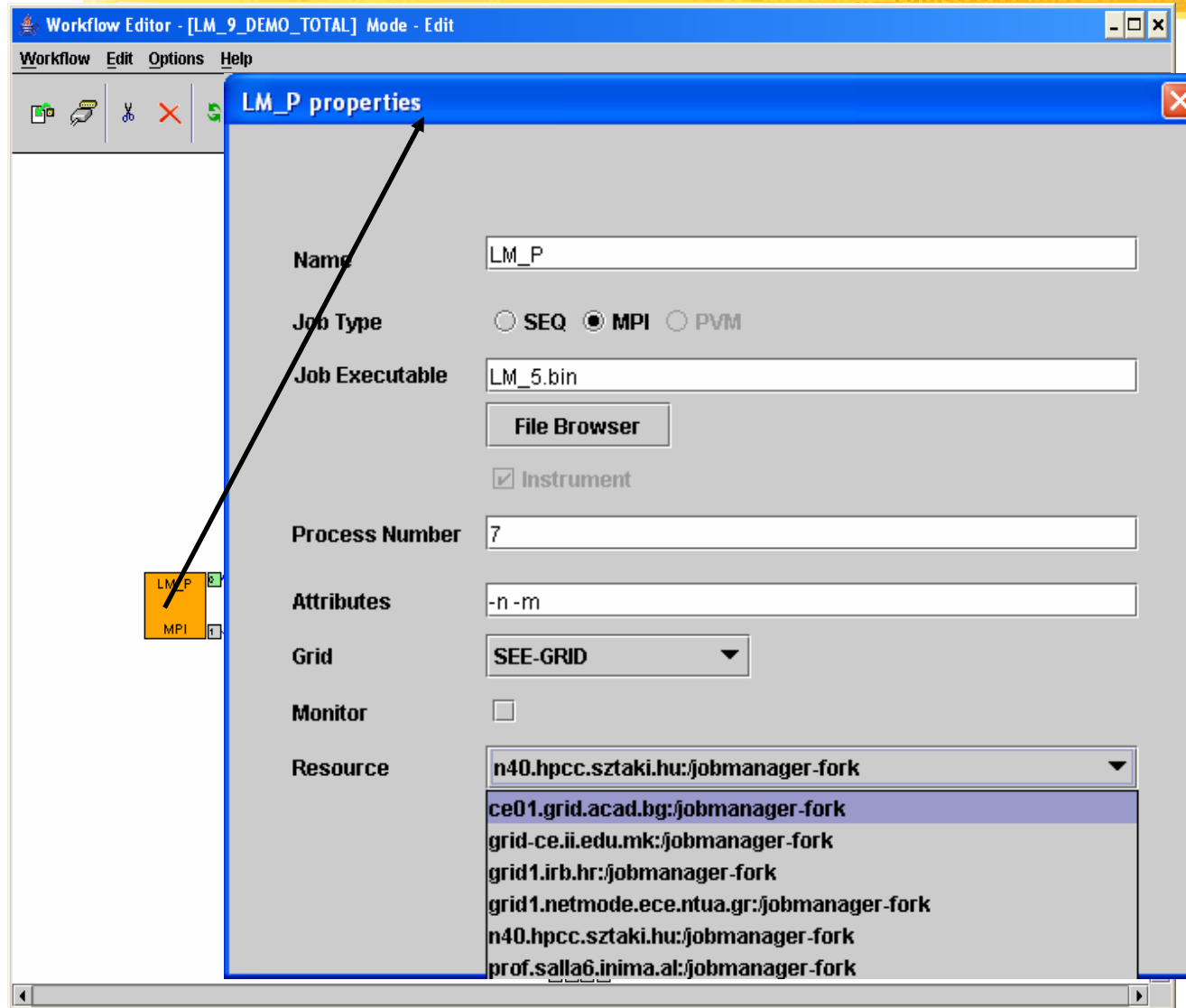
User concerns of Grid systems

- How to cope with the variety of these Grid systems?
- ➔ How to develop/create new Grid applications?
- How to execute Grid applications?
- How to observe the application execution in the Grid?
- How to tackle performance issues?
- How to port legacy applications
 - to Grid systems
 - between Grid systems?
- How to execute Grid applications over several Grids in a transparent way?

Main interactions 2.



Workflow Editor: Grid aware workflow mapping



- All jobs have properties window

-This window contains the most important information about the job eg. type, name, required process number, Grid and resource name for execution

Monitoring System

MDS Monitor LCG Monitor

Monitor

Select Grid: SEE-GRID View

Select VO: seegrid View

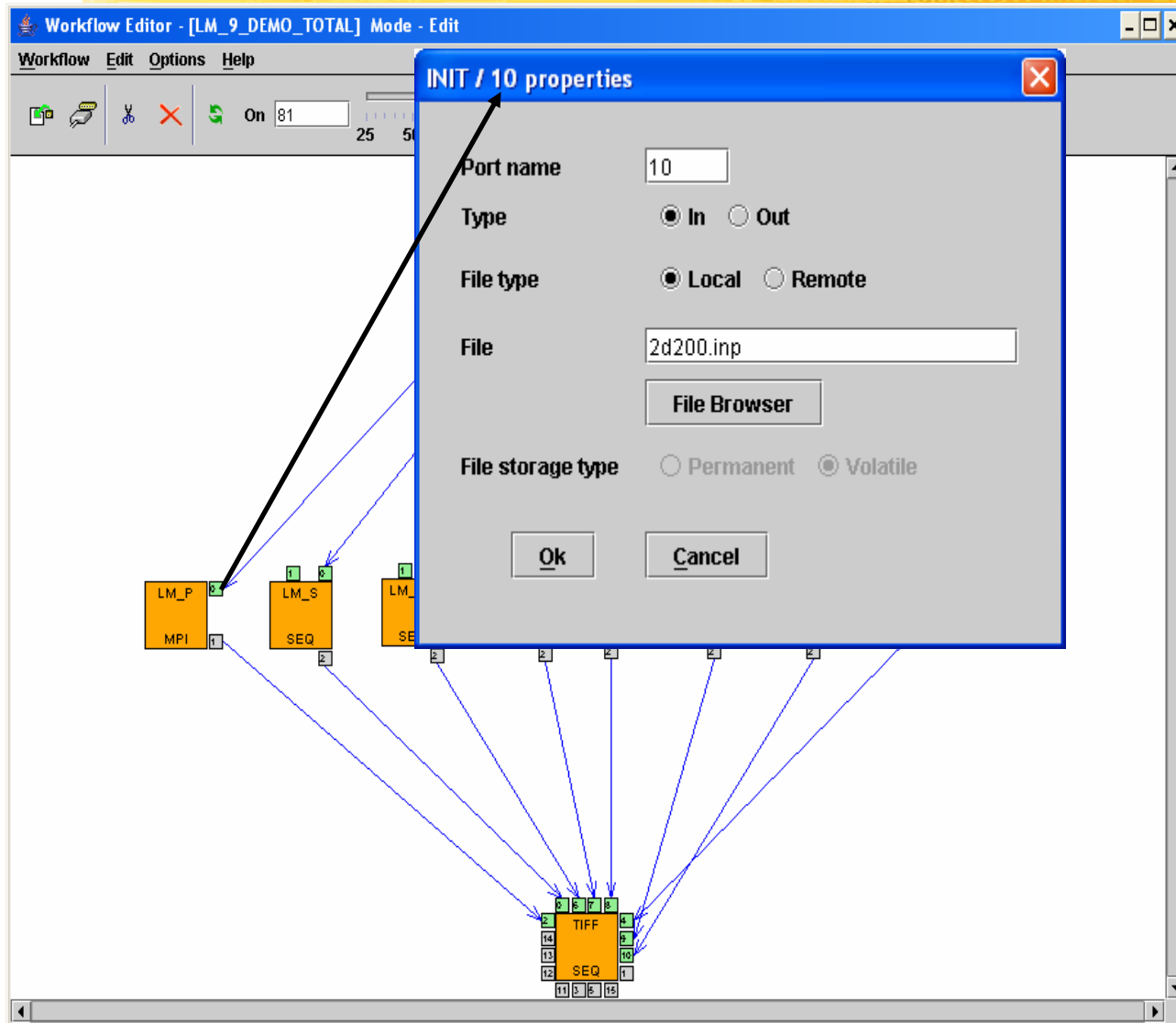
Grid: SEE-GRID VO: seegrid

Site Name	Computing Element						Storage Element		
	CPU			Job			Space		
	Total	Free	Usage	Running	Waiting	Load	Total	Available	Usage
AEGIS01-PHY-SCL	112	80	29%	7	0	0%	226.793 GB	216.34 GB	5%
AEGIS02-RCUB	20	20	0%	0	0	0%	398.466 GB	396.58 GB	0%
BG01-IPP	54	18	67%	4	0	0%	609.554 GB	473.543 GB	22%
BG02-IM	20	16	20%	1	0	0%	131.775 GB	79.957 GB	39%
BG03-IPP-N	3	3	0%	0	0	0%	566.608 GB	566.376 GB	0%
BG04-ACAD	48	32	33%	2	5	71%	554.647 GB	475.767 GB	14%
HR-01-RBI	60	12	80%	4	0	0%	78.317 GB	6.271 GB	92%
MK-01-UKIM_II	28	28	0%	0	0	0%	69.709 GB	69.075 GB	1%
RO-01-ICI	54	24	56%	5	36	88%	849.666 GB	828.387 GB	3%
ROGRID-NIPNE-01	24	24	0%	0	0	0%	862.807 GB	848.676 GB	2%
SZTAKI	4	4	0%	0	0	0%	4.566 GB	2.871 GB	37%
tubitakcg2	35	28	20%	1	0	0%	1.335 TB	1.335 TB	0%

- The portal uses MDS-2 and LCG-2 information systems

- Users can select a Grid and a Virtual Organization to be displayed

Workflow Editor



-Every job has ports

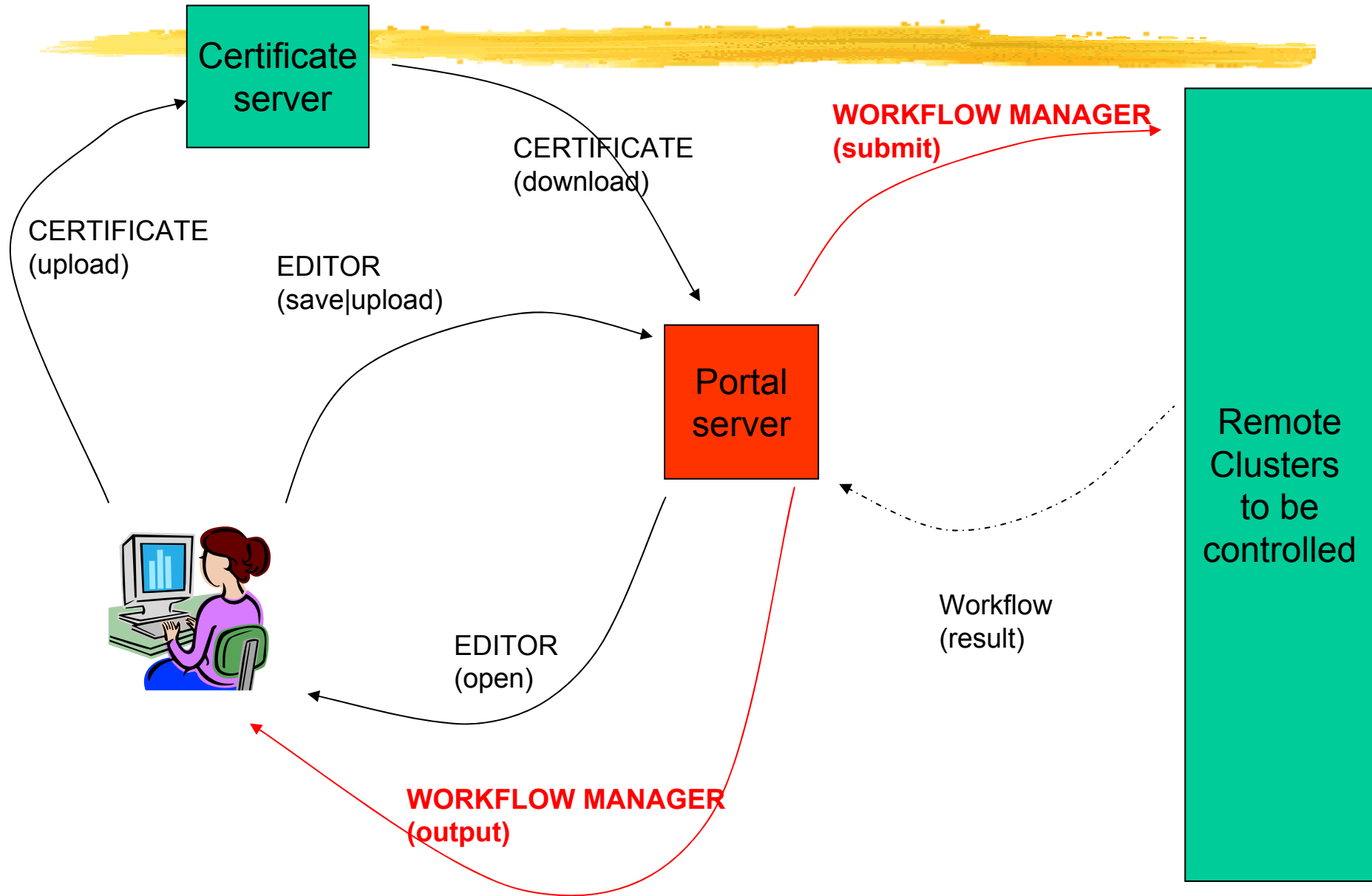
- The box in Green color represents the input port

-Every port has a Properties window to define the file settings

User concerns of Grid systems

- How to cope with the variety of these Grid systems?
- How to develop new Grid applications?
- ➔ How to execute Grid applications?
- How to observe the application execution in the Grid?
- How to tackle performance issues?
- How to port legacy applications
 - to Grid systems
 - between Grid systems?
- How to execute Grid applications over several Grids in a transparent way?

Main interactions 3.



Workflow Execution

The screenshot shows the PGrade Portal interface in Microsoft Internet Explorer. The browser title is "PGrade Portal - Microsoft Internet Explorer". The address bar shows the URL: <http://hgportal.hpcc.sztaki.hu:7080/gridsphere/gridsphere?action=doGotoPage&cid=2>. The page content includes a navigation menu with "Workflow", "Certificates", "Settings", "Information System", and "Help". Below the menu is the "Workflow Manager" section, which contains a "Refresh" button and a "Back" button. The main content is a table titled "Job list" with the following columns: Workflow, Job, Gridname, Hostname, Status, [Logs], [Output], and [Visualization]. The table contains 11 rows of job data. The status of each job is indicated by a colored box: white for "init", red for "running", and green for "finished". A message at the bottom of the table reads "Message: Job list refreshed.".

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	running	Out	-	Visualize
	LM_P.2	SEE-GRID	n40.hpcc.sztaki.hu	running	Out	-	Visualize
	LM_S	SEE-GRID	grid-ce.ii.edu.mk	finished	Out	-	-
	LM_S.2	SEE-GRID	grid1.irb.hr	finished	Out	-	-
	LM_S.3	SEE-GRID	grid1.netmode.ece.ntua.gr	finished	Out	-	-
	LM_S.4	SEE-GRID	grid1.irb.hr	finished	Out	-	-
	LM_S.5	SEE-GRID	testbed001.grid.icl.ro	finished	Out	-	-
	LM_S.6	HUNGRID	chemgrid3.chemres.hu	finished	Out	-	-
	TIFF	HUNGRID	grid109.kfki.hu	init	-	-	-

- It displays the list of jobs and their status

- The current status of the jobs are represented by colors

- It also provides access to their logs and outputs, and visualizes them

White/Red/Green color means the job is initialised/running/finished

Downloading the results

The screenshot shows the P-Grade Portal interface in a Mozilla browser window. The page title is "P-Grade Portal - Mozilla" and the URL is "http://fn2.hpcc.sztaki.hu:9080/gridsphere/gridsphere?action=doGotoPage&cid=2". The page features the P-Grade logo and a "portal" link. A "Workflow Manager" section displays a "Job list" table with columns for Workflow, Job, Hostname, Status, Logs, Output, Visualization, and Action. The first row, "nowcast-final-g_SGE", is highlighted, and a red arrow points from its "Output" column to a "Visualize" button. A second red arrow points from this button to a "Save it to disk" option in a dialog box titled "Opening nowcast_final_g.zip". The dialog box provides information about the file type and offers options for how to handle it.

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.

Opening nowcast_final_g.zip

The file "nowcast_final_g.zip" is of type application/x-zip-compressed, and Mozilla does not know how to handle this file type. This file is located at: e:\pri\mc04

What should Mozilla do with this file?

- Open it with the default application
- Open it with Choose...
- Save it to disk
- Always perform this action when handling files of this type

OK Cancel

User concerns of Grid systems

- How to cope with the variety of these Grid systems?
- How to develop new Grid applications?
- How to execute Grid applications?
- How to observe the application execution in the Grid?
- How to tackle performance issues?
- How to port legacy applications
 - to Grid systems
 - between Grid systems?
- How to execute Grid applications over several Grids in a transparent way?

On-Line Monitoring

PGrade Portal - Microsoft Internet Explorer

Fájl Szerkesztés Nézet Kedvencek Eszközök Súgó

Vissza Keresés Kedvencek

Cím <http://hgportal.hpcc.sztaki.hu:7080/gridsphere/gridsphere?action=doVisualizeWorkflowTrace&cid=2> Ugrás Hivatkozások

Workflow Certificates Settings Information System Help

Workflow Manager

Tracefile visualization

Back

workflow: LM_9_DEMO_TOTAL

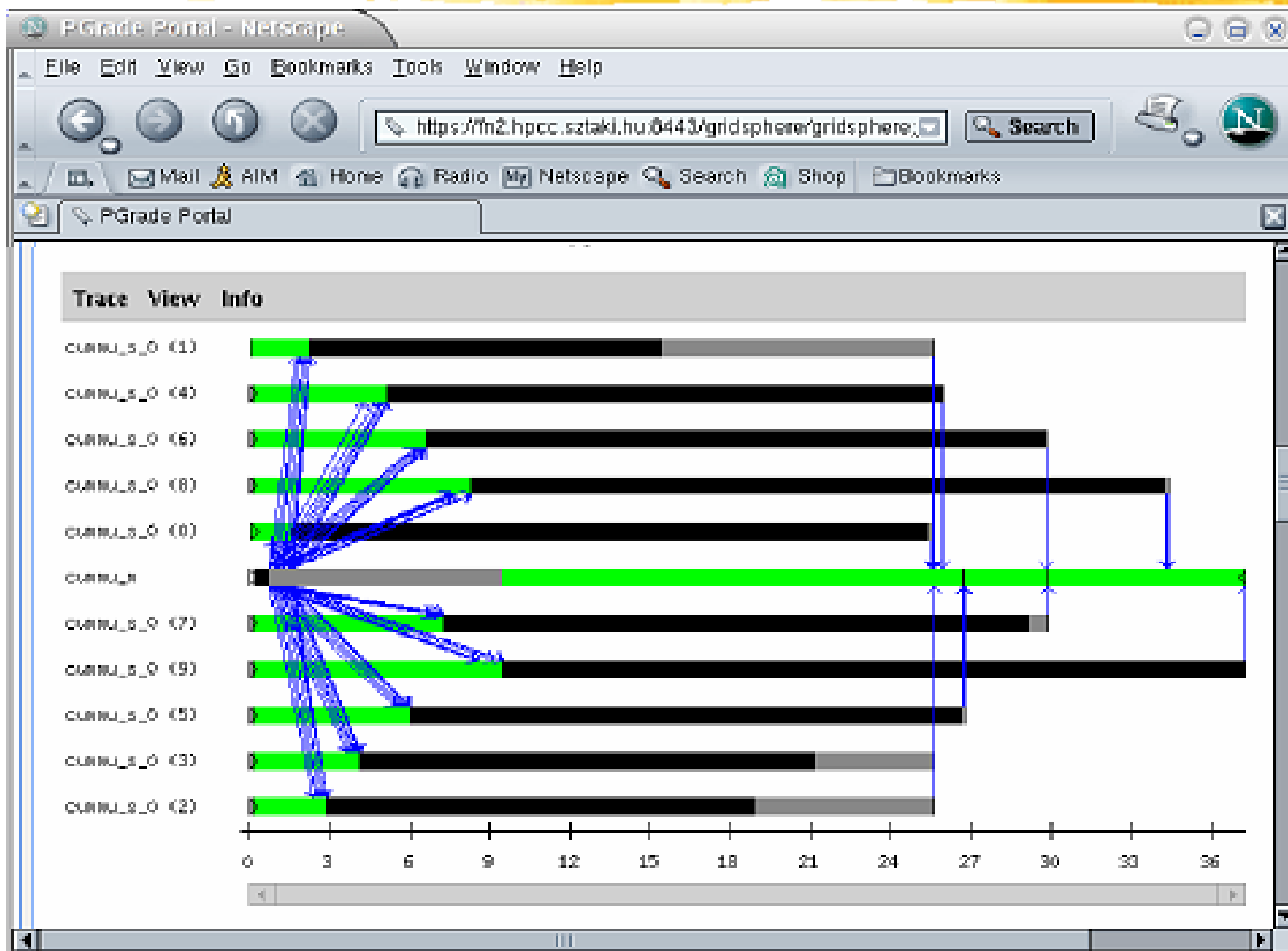
Trace View Info	Width:	Height:
LM_P.2	600	350
LM_S.3		
LM_S.2		
LM_S		
TIFF grid109.kfki.hu INIT ce01.grid.acad.bg		
LM_S.5		
LM_S.6		
LM_S.4		
LM_P		

0s 1m40s 3m20s 5m0s 6m40s 8m20s 10m0s

Applet sztaki.trace.client.TraceClientApplet started

Internet

Job execution visualization

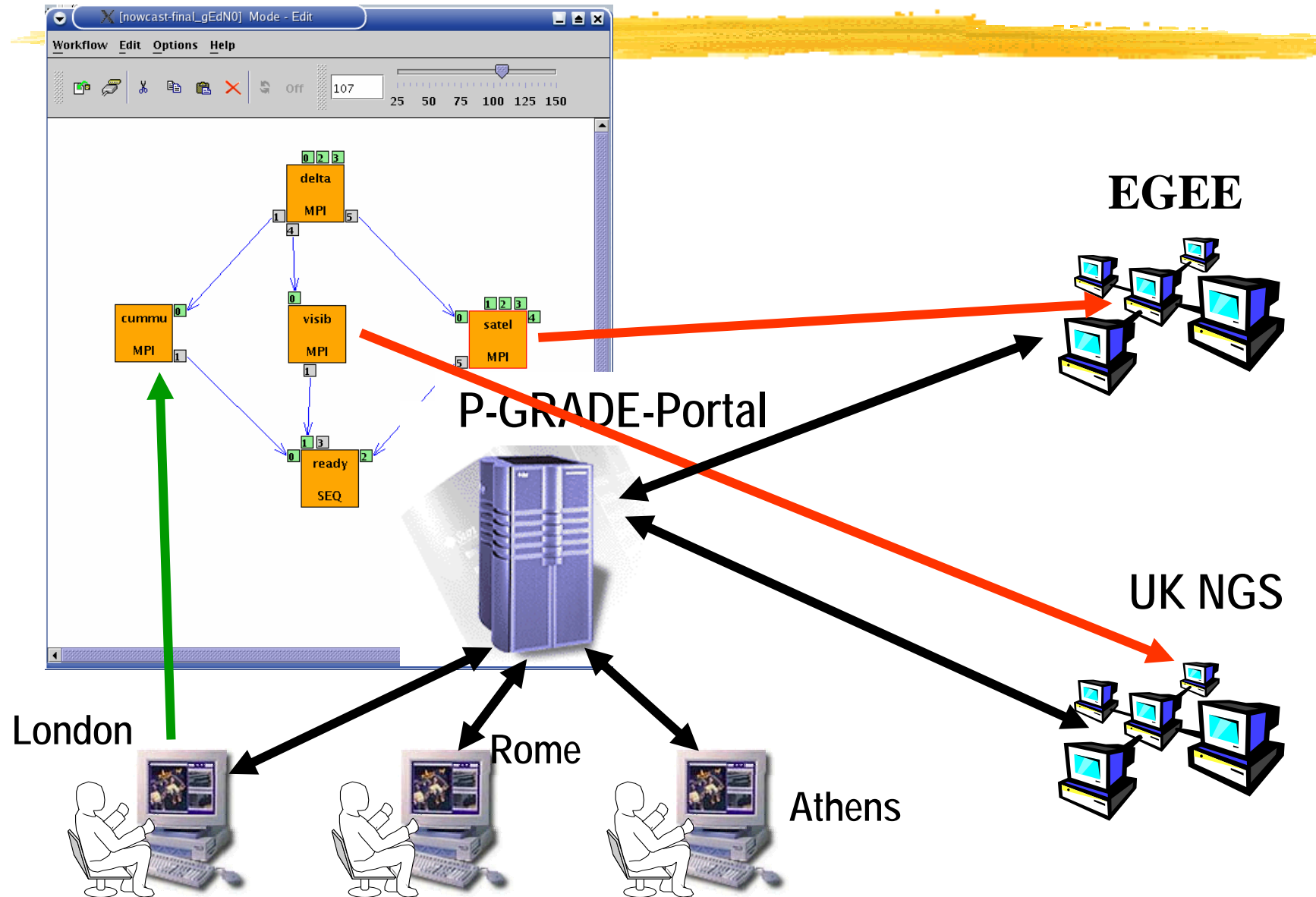



User concerns of Grid systems

- How to cope with the variety of these Grid systems?
- How to develop new Grid applications?
- How to execute Grid applications?
- How to observe the application execution in the Grid?
- How to tackle performance issues?
- How to port legacy applications
 - to Grid systems
 - between Grid systems?
- How to execute Grid applications over several Grids in a transparent way?

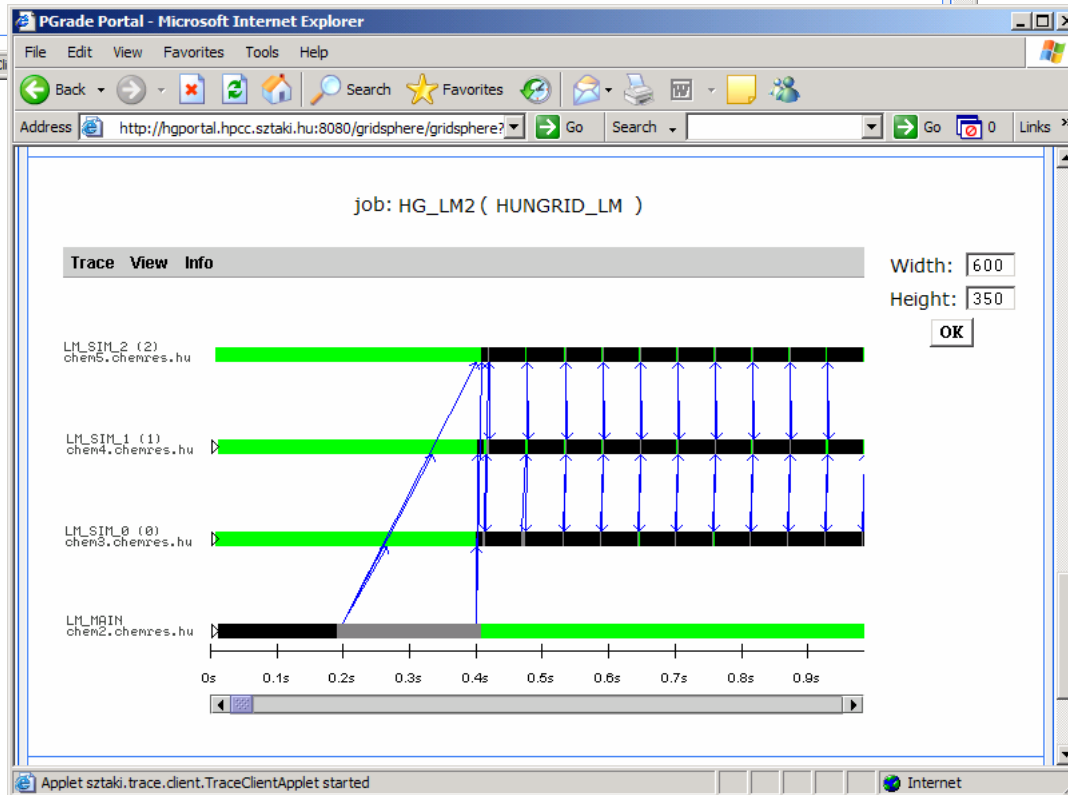
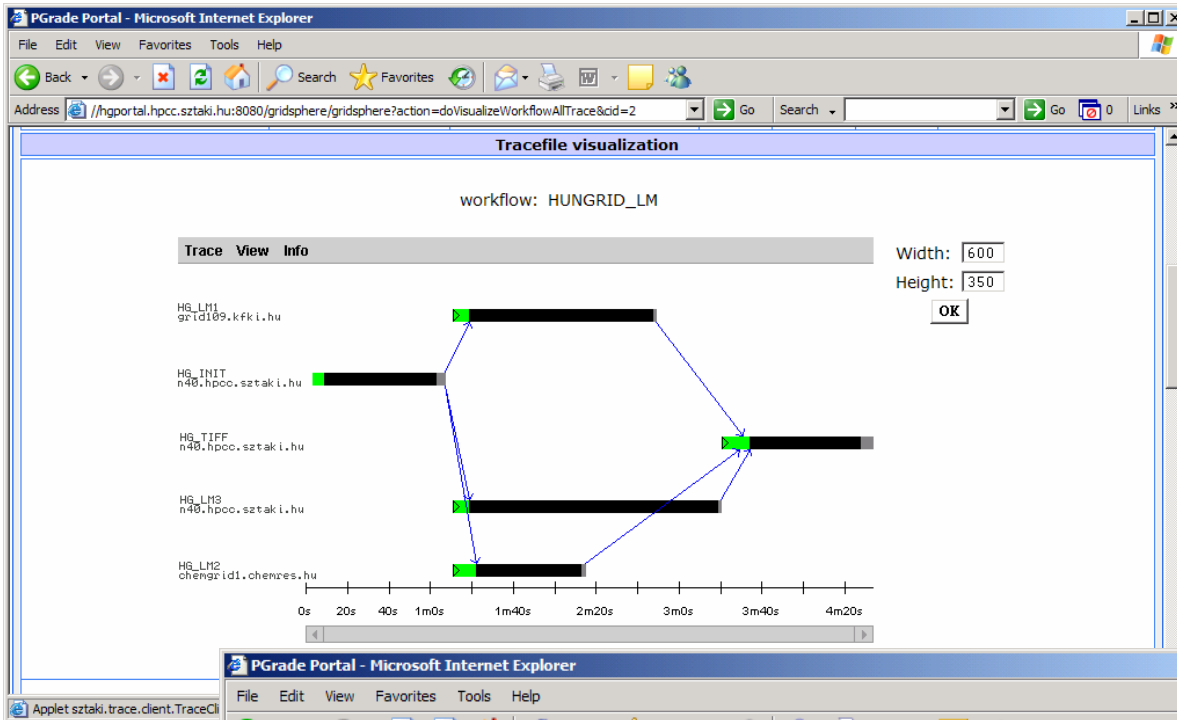


Multi-Grid portal



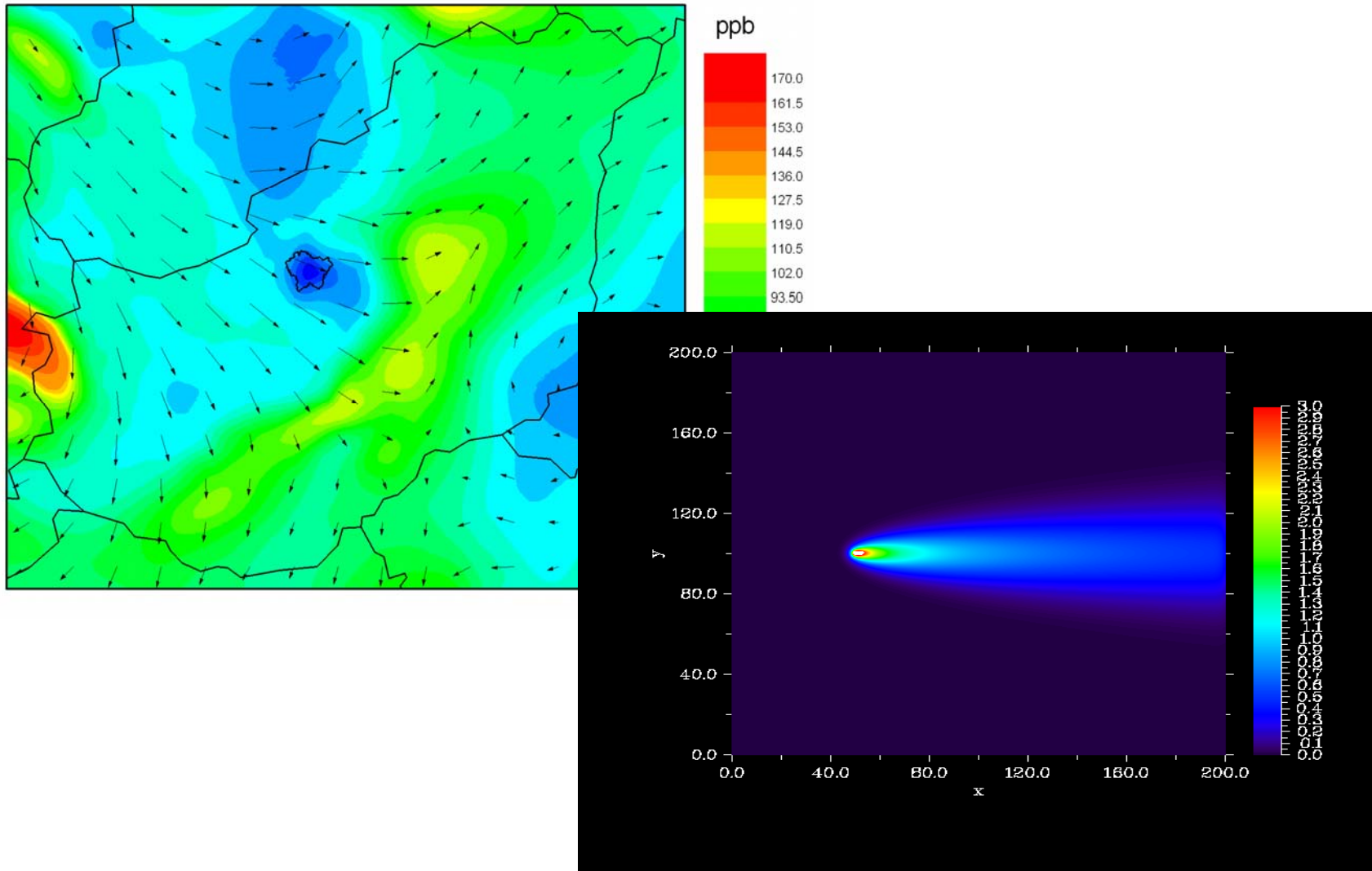


Levegőminőségi (LM) alkalmazás



Futás- és teljesítmény vizualizáció a HUNGRID-en

Results: Air pollution forecast



Relationship between P-GRADE and P-GRADE Portal

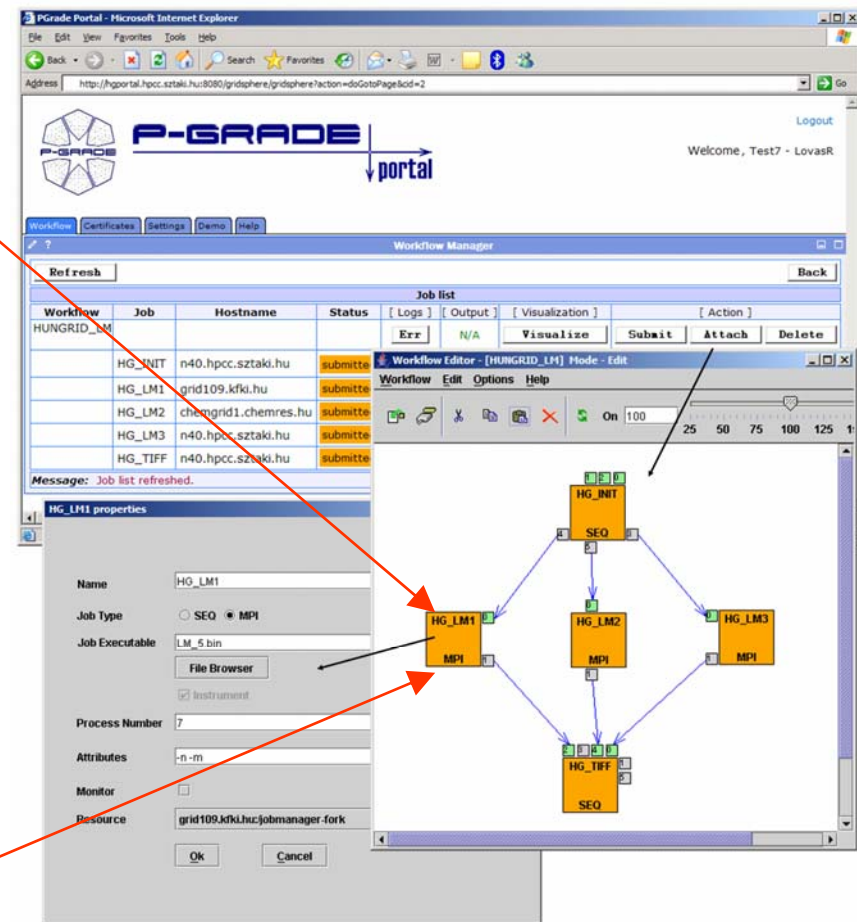
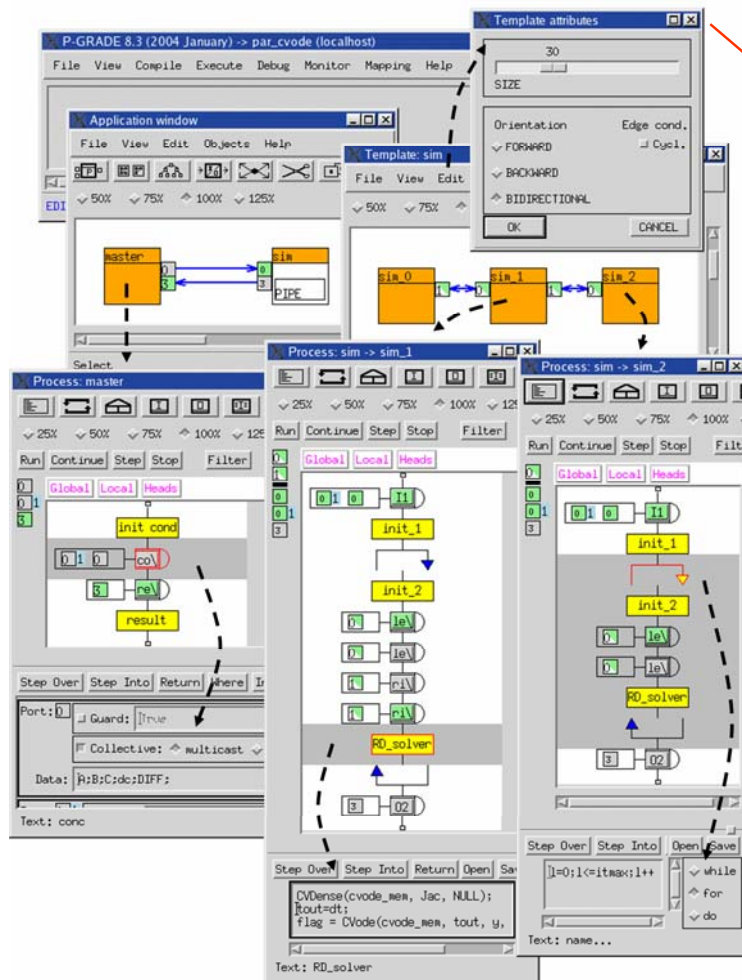
- The workflow editor of the portal is compatible with the P-GRADE workflow editor



- MPI programs and workflows developed in P-GRADE can be executed in various Grids by the P-GRADE portal
- They can be
 - Submitted to the Grid by the portal
 - Workflows can be modified by the workflow editor of the portal
- P-GRADE is the **development environment** and the portal is the Grid **execution environment**

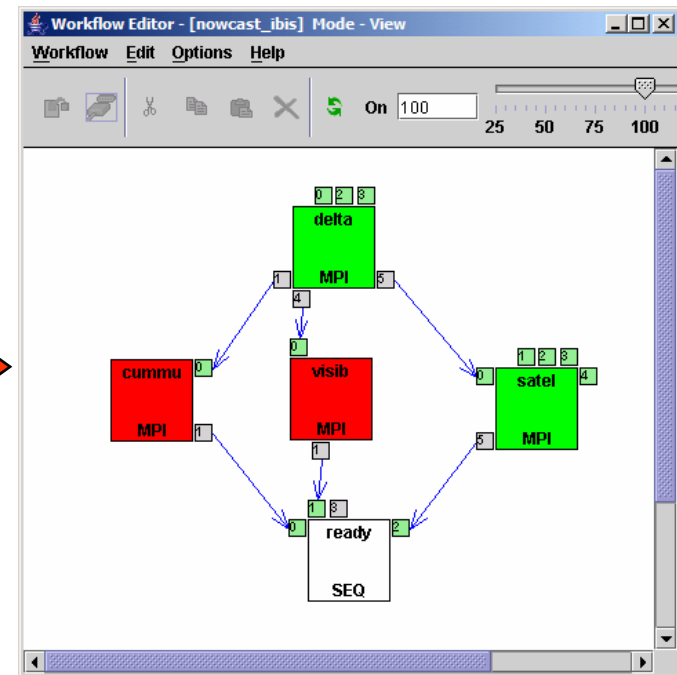
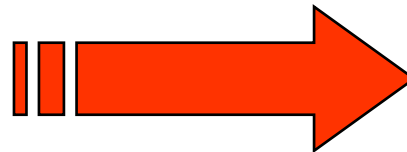
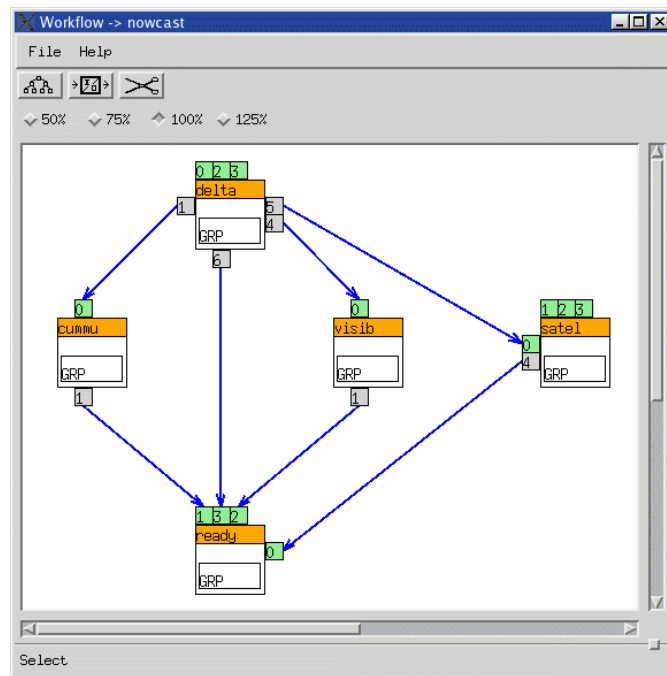
P-GRADE → EGEE

Application → MPI Job of a workflow & P-GRADE portal



P-GRADE → EGEE

P-GRADE workflow → “Export” function → P-GRADE portal



Final conclusions



- **P-GRADE and P-GRADE portal** provide a user-friendly, high-level Grid programming and execution environment
- **Portal technology like the P-GRADE portal helps the end-users in many ways:**
 - Easy-to-use workflow concept for solving complex problems
 - Execution visualization support to observe Grid execution
 - Switching between Grid technologies will be transparent to the end-user
 - Interoperability between different Grid systems can be solved
 - Simultaneous use of different Grid systems can be solved

Learn once, use everywhere

Develop once, run everywhere

- More information at:

www.lpds.sztaki.hu/pgrade/
www.lpds.sztaki.hu/pgportal/