



GRID activities at MTA SZTAKI

Peter Kacsuk

MTA SZTAKI

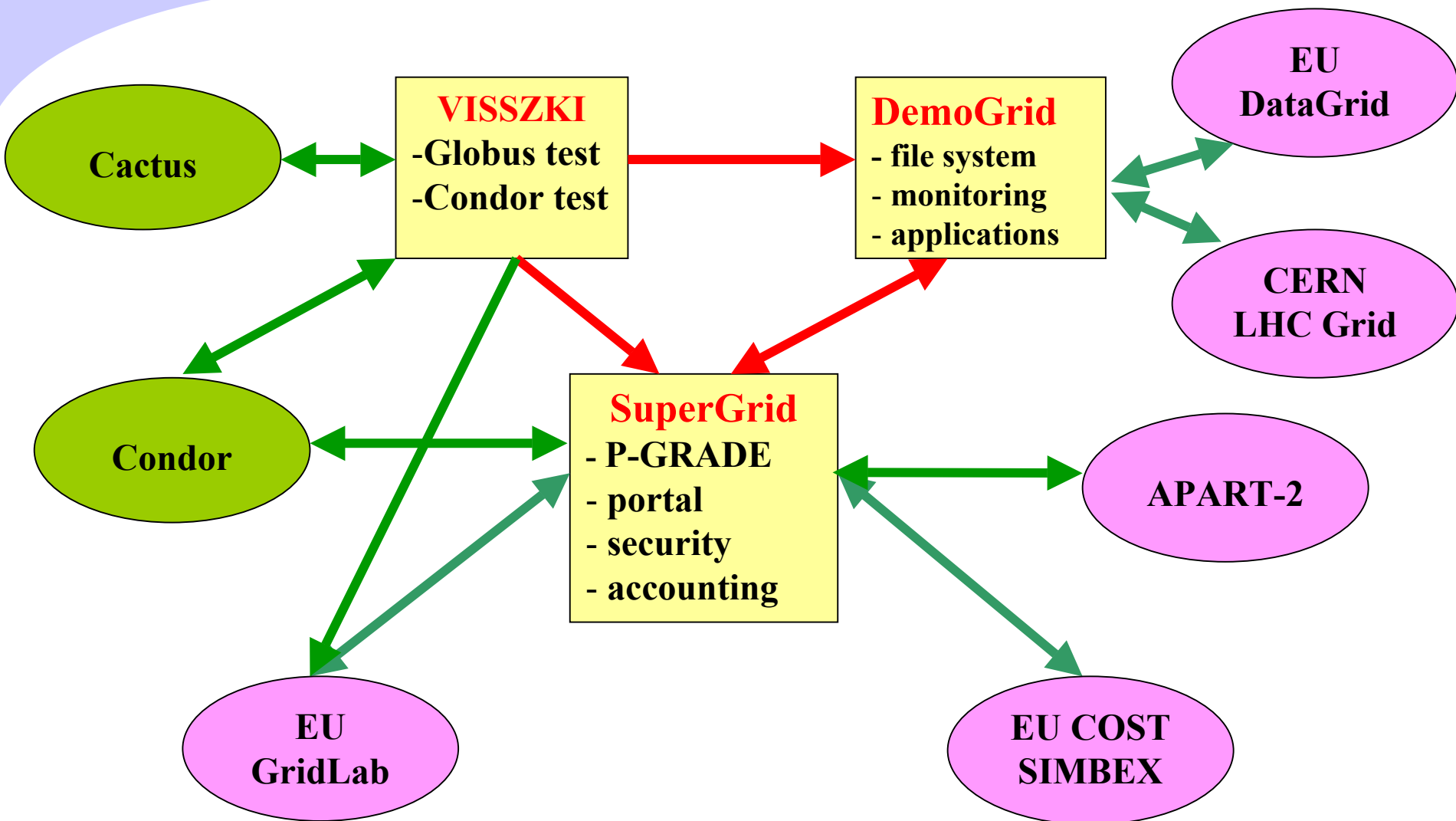
Laboratory of Parallel and Distributed Systems

www.lpds.sztaki.hu

Contents

- SZTAKI participation in EU and Hungarian Grid projects
- **P-GRADE** (Parallel Grid Run-time and Application Development Environment)
- Integration of P-GRADE and Condor
- **TotalGrid**
- Meteorology application by TotalGrid
- Future plans

Hungarian and international GRID projects





EU Grid projects of SZTAKI

- **DataGrid** - application performance monitoring and visualization
- **GridLab** - grid monitoring and information system
- **APART-2** - leading the Grid performance analysis WP
- **SIMBEX** - developing a European metacomputing system for chemists based on P-GRADE



Hungarian Grid projects of SZTAKI

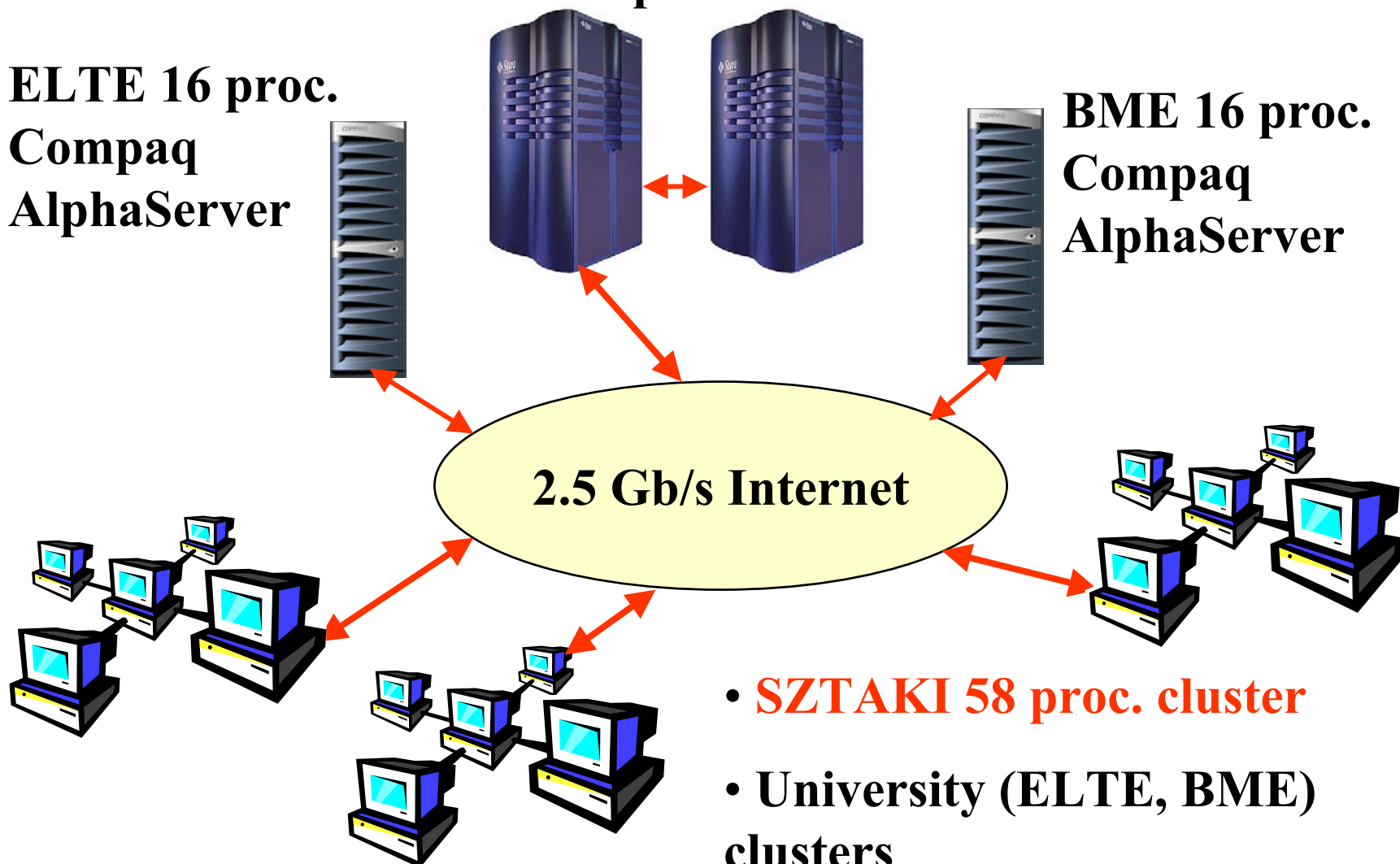
- **VISSZKI**
 - explore and adopt Globus and Condor
- **DemoGrid**
 - grid and application performance monitoring and visualization
- **SuperGrid (Hungarian Supercomputing Grid)**
 - integrating P-GRADE with Condor and Globus in order to provide a high-level program development environment for the Grid
- **ChemistryGrid (Hungarian Chemistry Grid)**
 - Developing chemistry Grid applications in P-GRADE
- **JiniGrid (Hungarian Jini Grid)**
 - Combining P-GRADE with Jini and Java
 - Creation the OGSA version of P-GRADE
- **Hungarian Cluster Grid Initiative**
 - To provide a nation-wide cluster Grid for universities

Structure of the Hungarian Supercomputing Grid

NIIFI 2*64 proc. Sun E10000

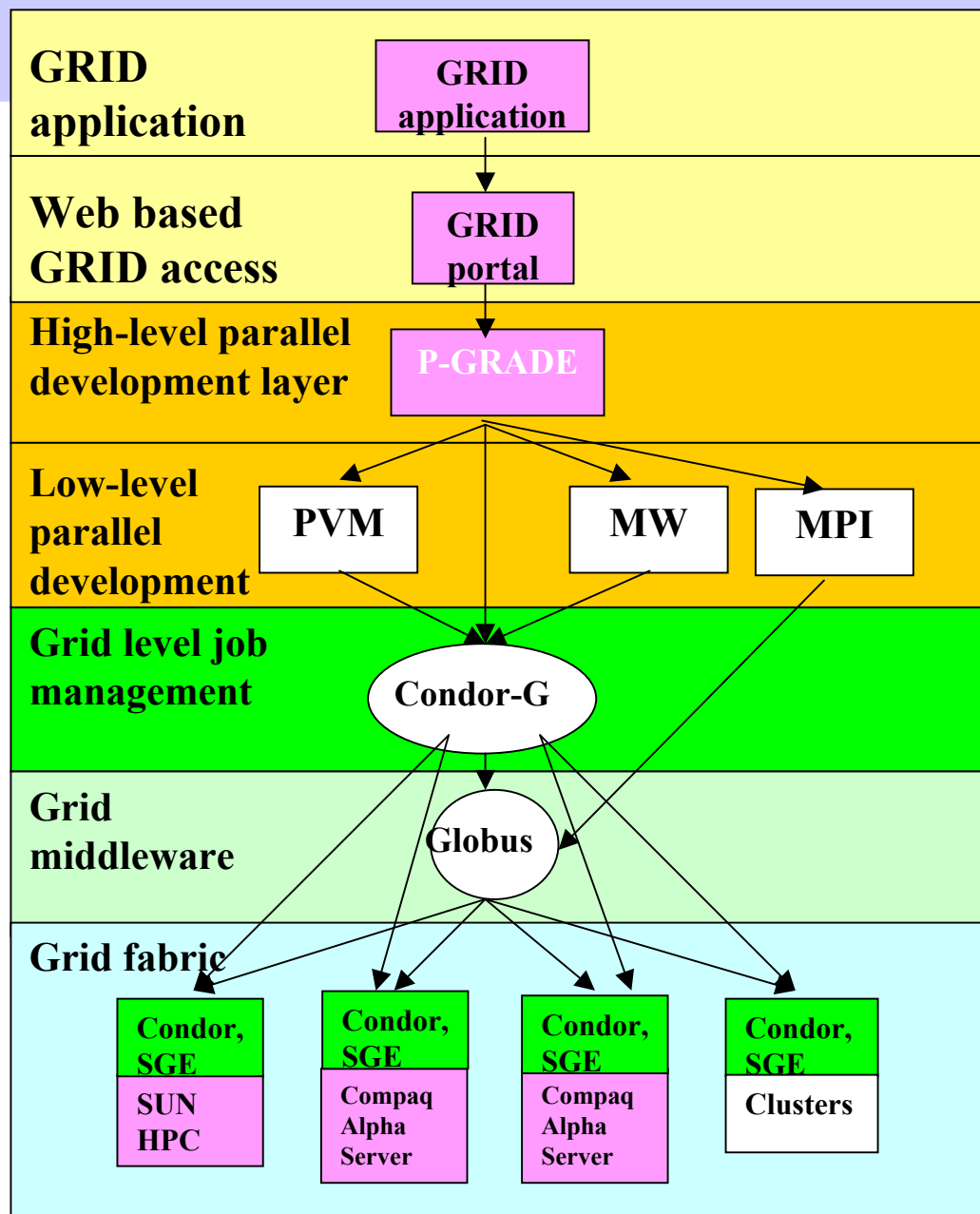
**ELTE 16 proc.
Compaq
AlphaServer**

**BME 16 proc.
Compaq
AlphaServer**





The Hungarian Supercomputing GRID project



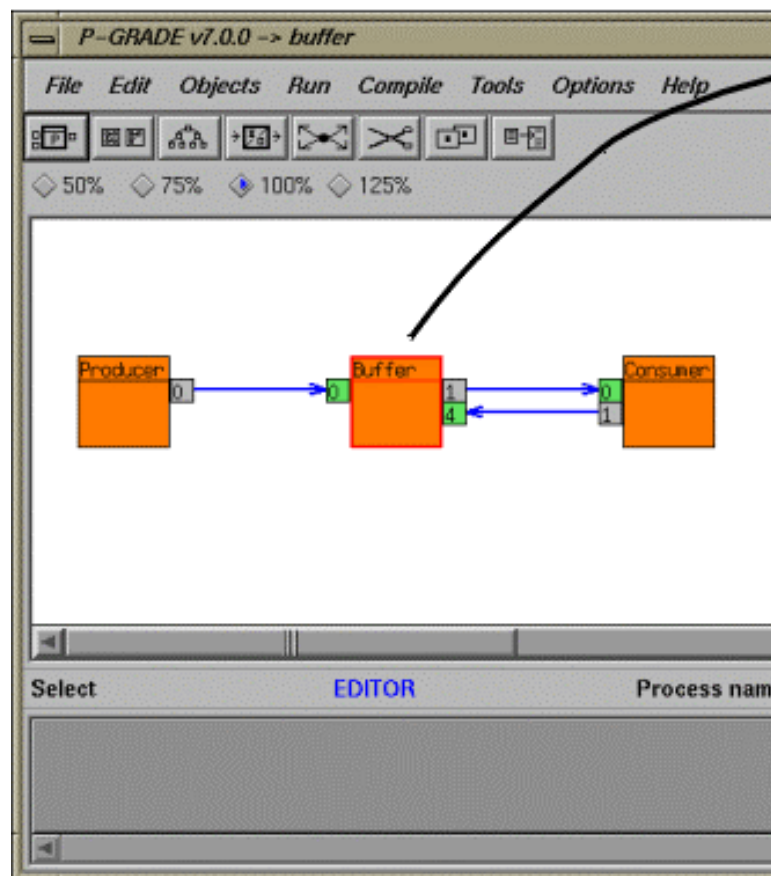


Distributed supercomputing: P-GRADE

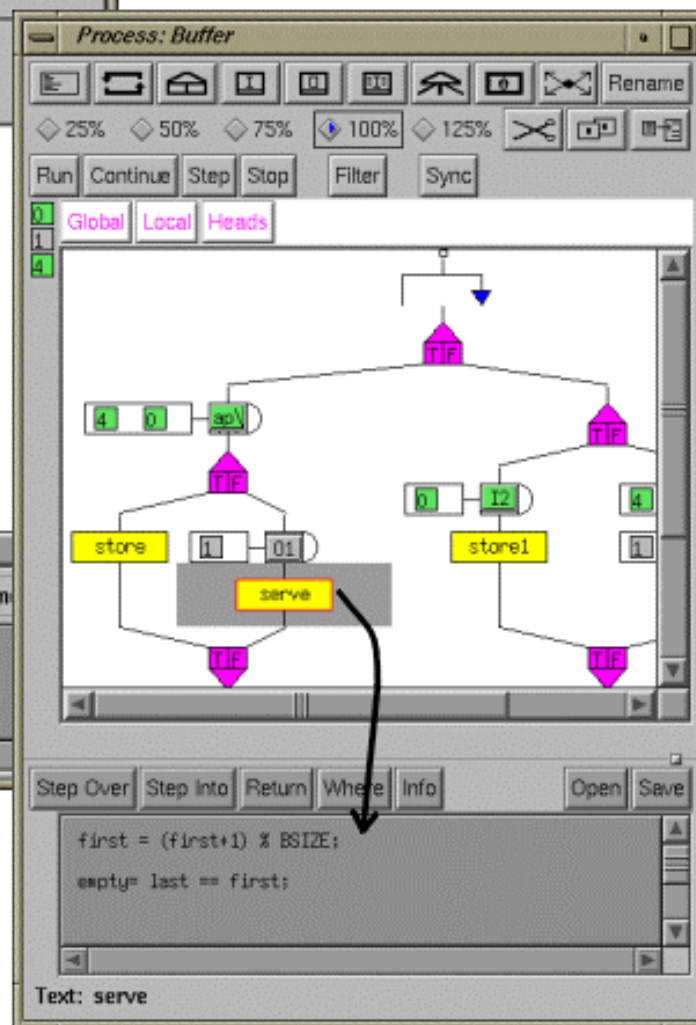
- **P-GRADE** (**P**arallel **G**rid **R**un-time and **A**pplication **D**evelopment **E**nvironment)
- A highly integrated parallel Grid application development system
- Provides:
 - Parallel, supercomputing programming for the Grid
 - Fast and efficient development of Grid programs
 - Observation and visualization of Grid programs
 - Fault and performance analysis of Grid programs
- Further development in the:
 - **Hungarian Supercomputing Grid** project
 - **Hungarian Chemistry Grid** project
 - **Hungarian Jini Grid** project

Three layers of GRAPNEL

Application window



Process window



Text window

Communication Templates

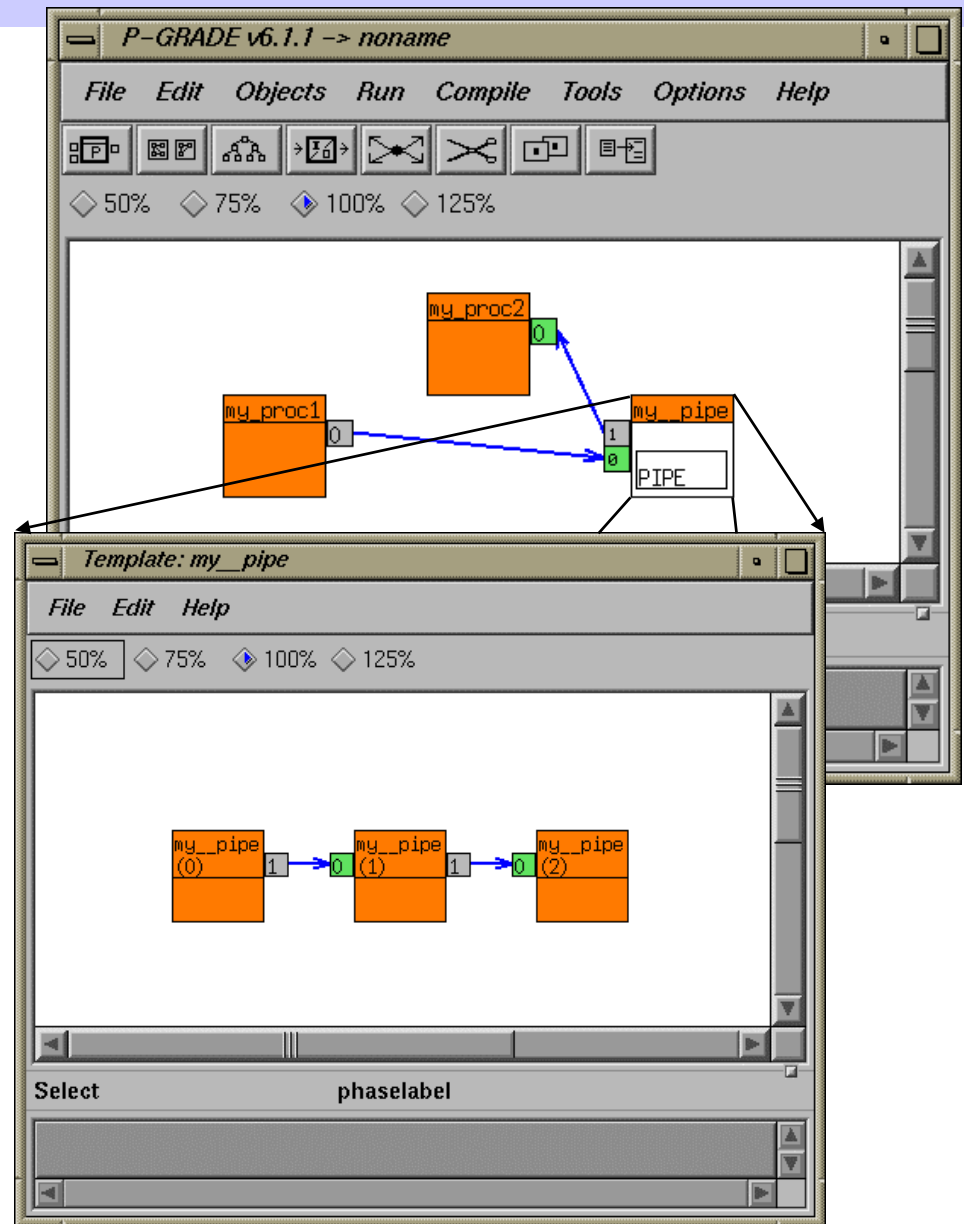
- Pre-defined regular process topologies

- process farm
- pipeline
- 2D mesh

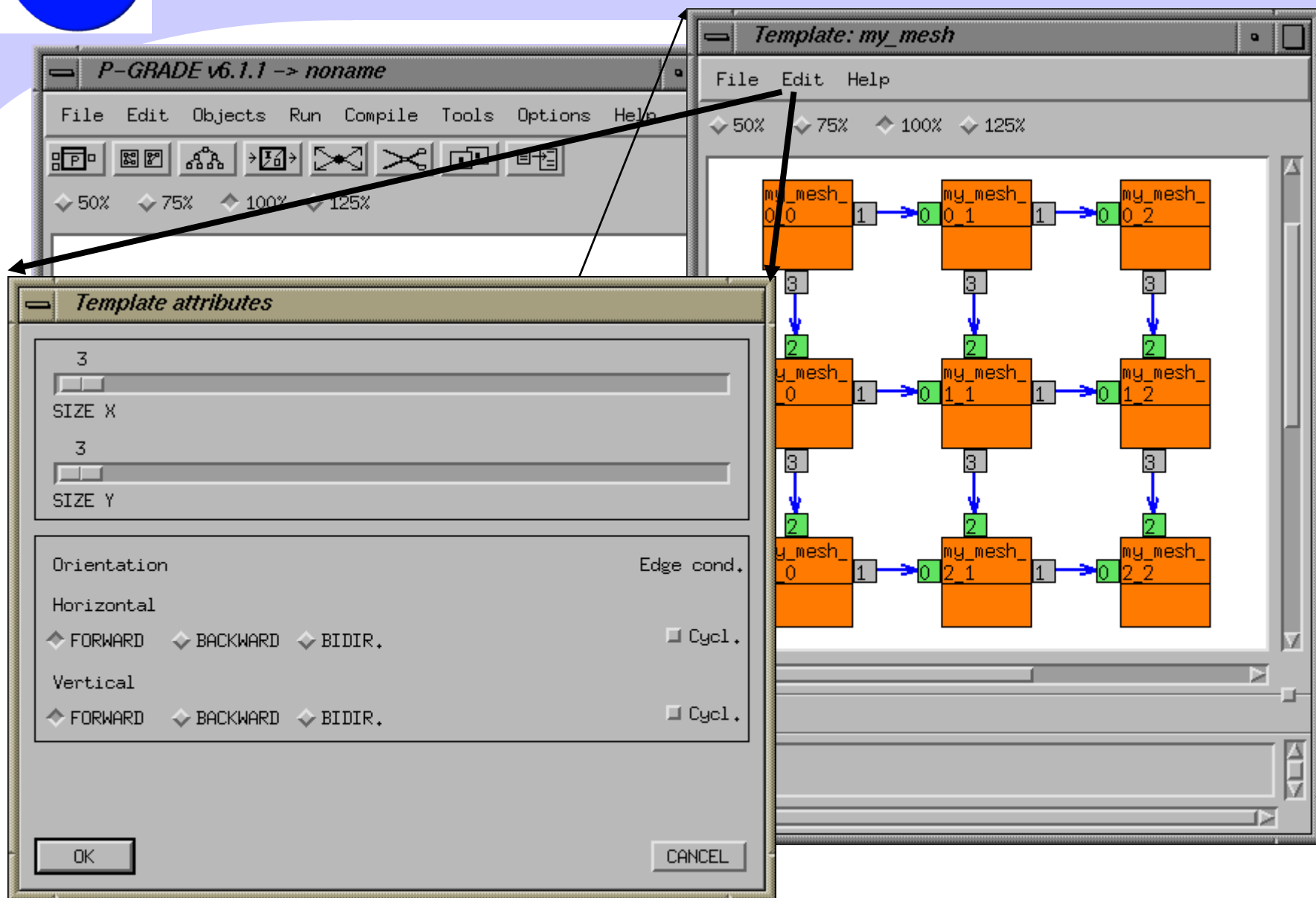
- User defines:

- representative processes
- actual size

- Automatic scaling

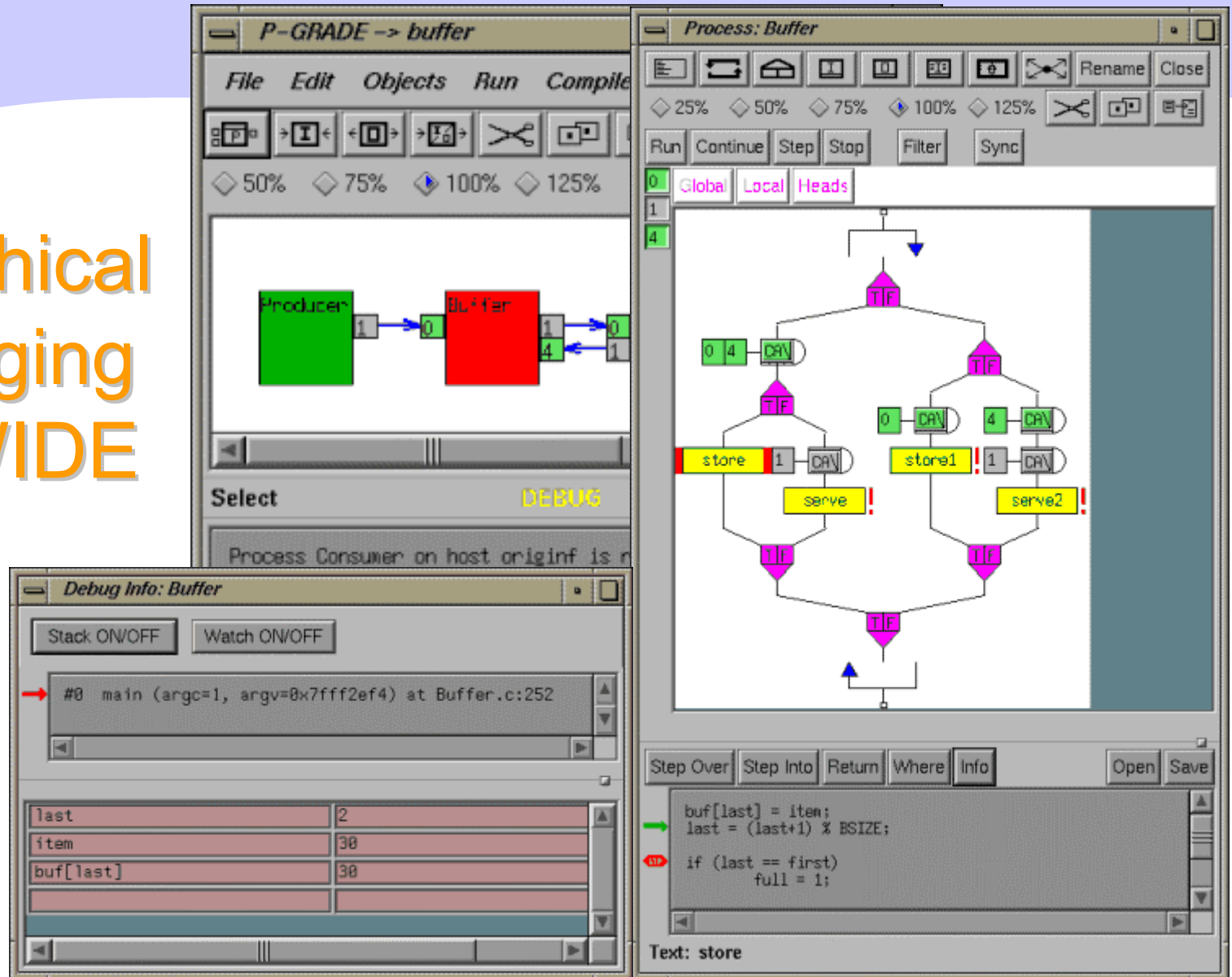


Mesh Template



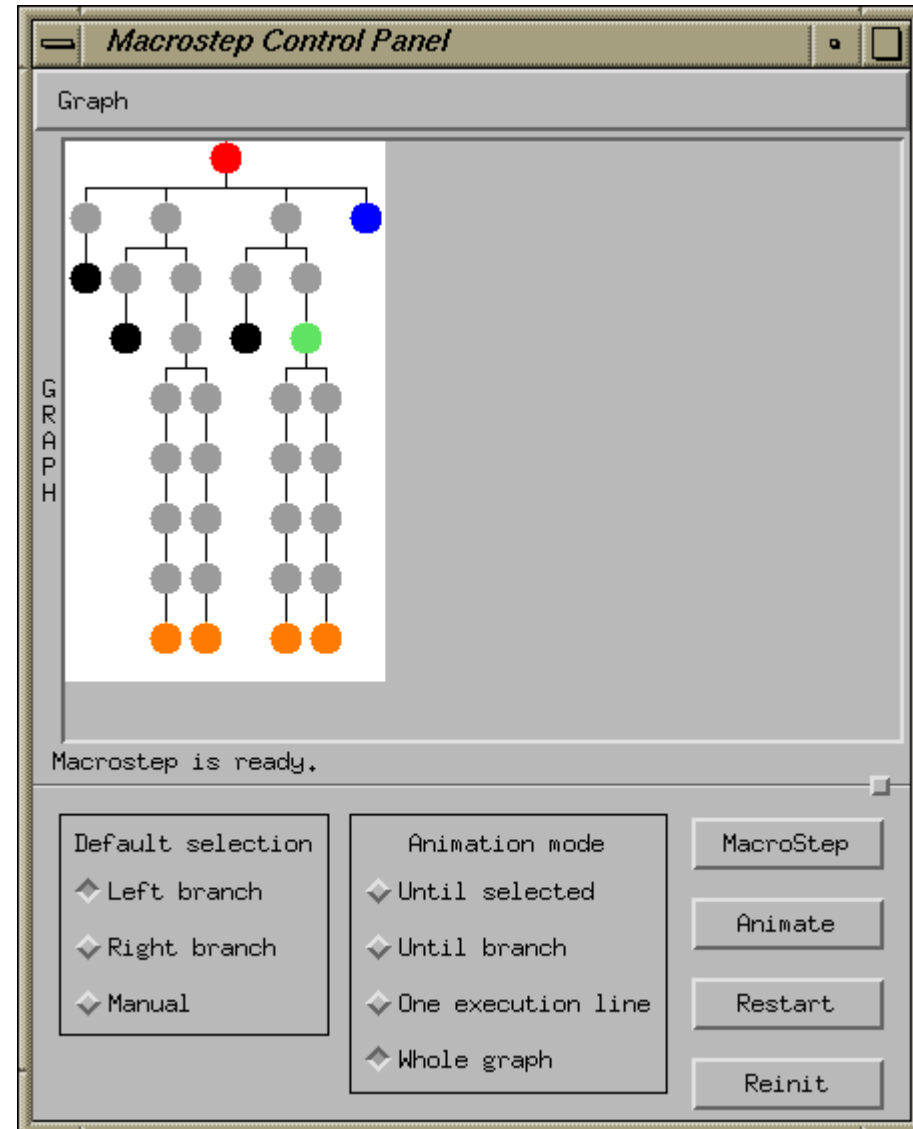


Hierarchical Debugging by DIWIDE



Macrostep Debugging

- Support for **systematic debugging** to handle non-deterministic behaviour of parallel applications
- **Automatic dead-lock detection**
- **Replay technique** with collective breakpoints
- Systematic and automatic generation of **Execution Trees**
- Testing parallel programs **for every time condition**



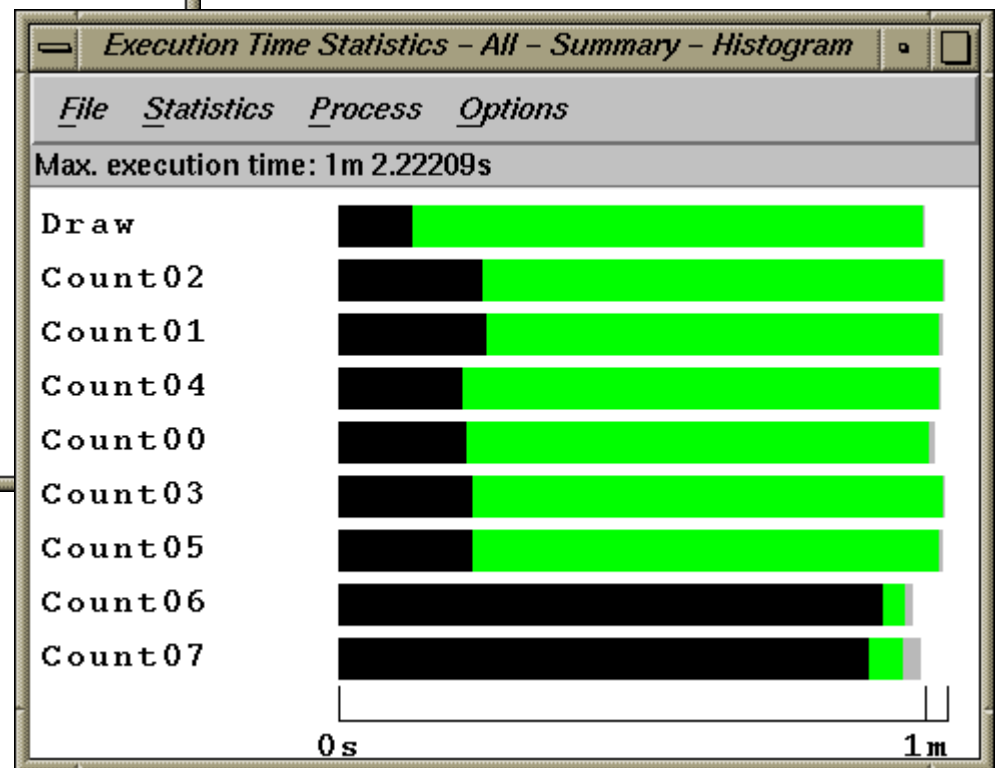
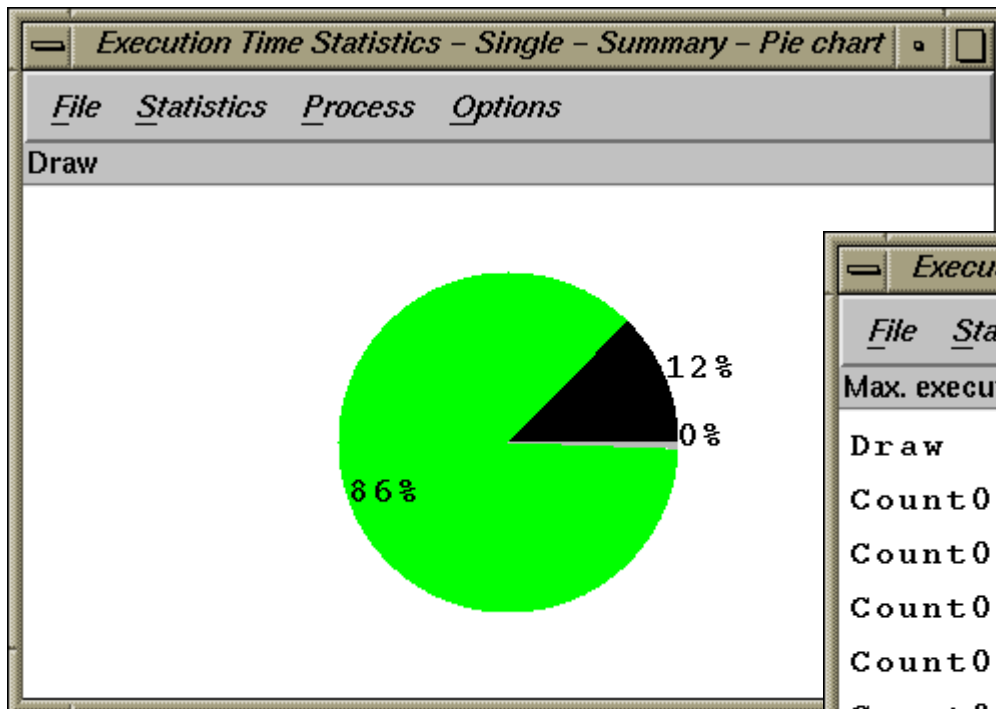


GRM semi-on-line monitor

- Monitoring and visualising parallel programs **at GRAPNEL level**.
- Evaluation of **long-running programs** based on **semi-on-line** trace collection
- Support for debugger in P-GRADE by **execution visualisation**
- Collection of both **statistics** and **event traces**
- Application monitoring and visualization in the **Grid**
- **No lost of trace data at program abortion**. The execution can be visualised to the point of abortion.

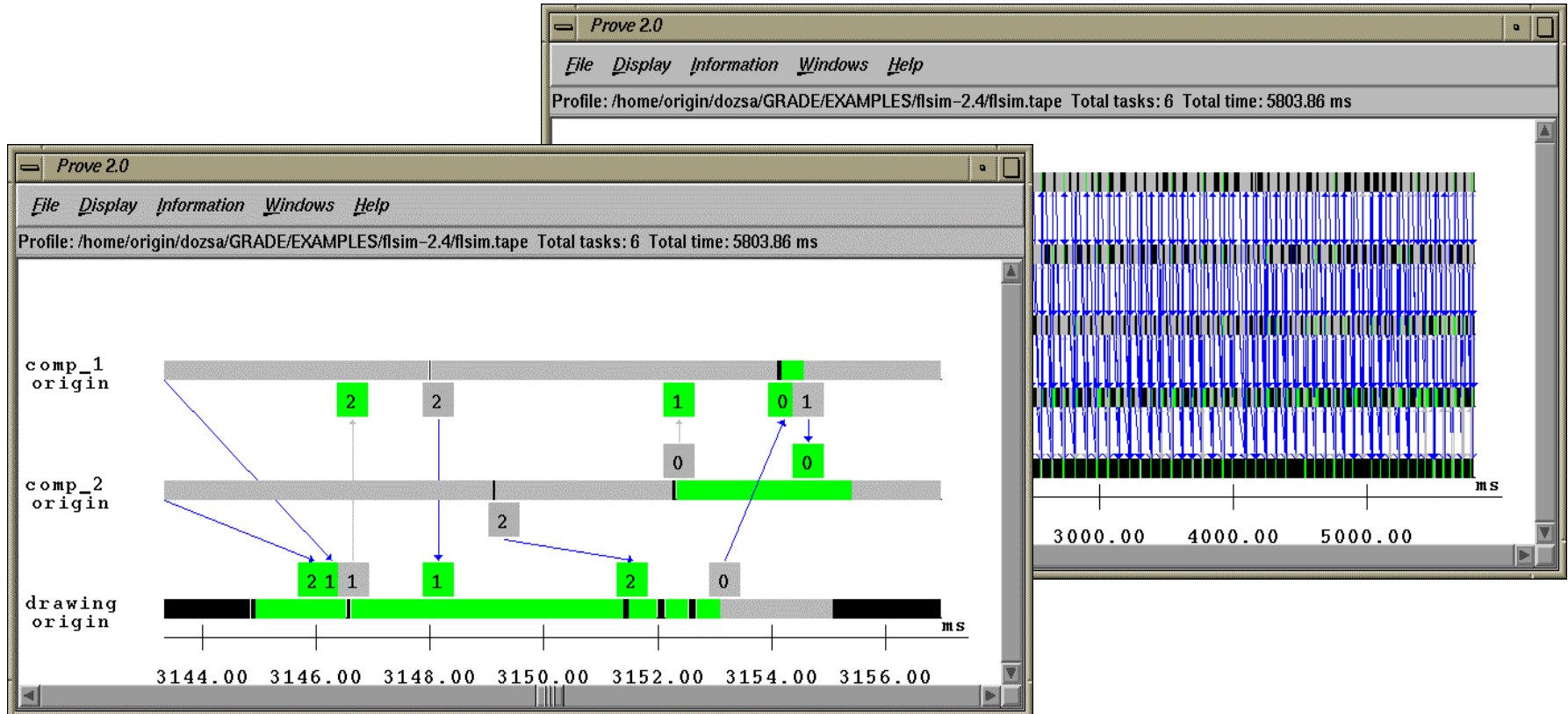
PROVE Statistics Windows

- Profiling based on counters
- Analysis of very long running programs is enabled



PROVE: Visualization of Event Traces

- User controlled focus on processors, processes and messages
- Scrolling visualization windows forward and backwards





Integration of Macrostep Debugging and PROVE

P-GRADE -> Dinning_Philosophers

File Edit Objects Run Compile Tools Options Help

50% 75% 100% 125%

Select **DEBUG**

Collect trace by DIWIDE!
Monitor: Trace collection started
Monitor: Trace collection finished

Macrostep Control Panel

Graph

Macrostep is ready.

Default selection
☒ Left branch
☐ Right branch
☐ Manual

Animation mode
☒ Until selected
☐ Until branch
☐ One execution line
☐ Whole graph

MacroStep
Animate
Restart
Reinit

Prove 3

File Display Information Statistics Help

Profile: /tmp/Dinning_Philosophers.1119.trace Total tasks: 4 Total time: 1m 6s

Philosopher_0
originf

Chopstick_1
originf

Philosopher_1
originf

Chopstick_2
originf

55s 1m 1m5s

Number of Events: 97, Behaviour list length: 86

GRAPNEL Application Console

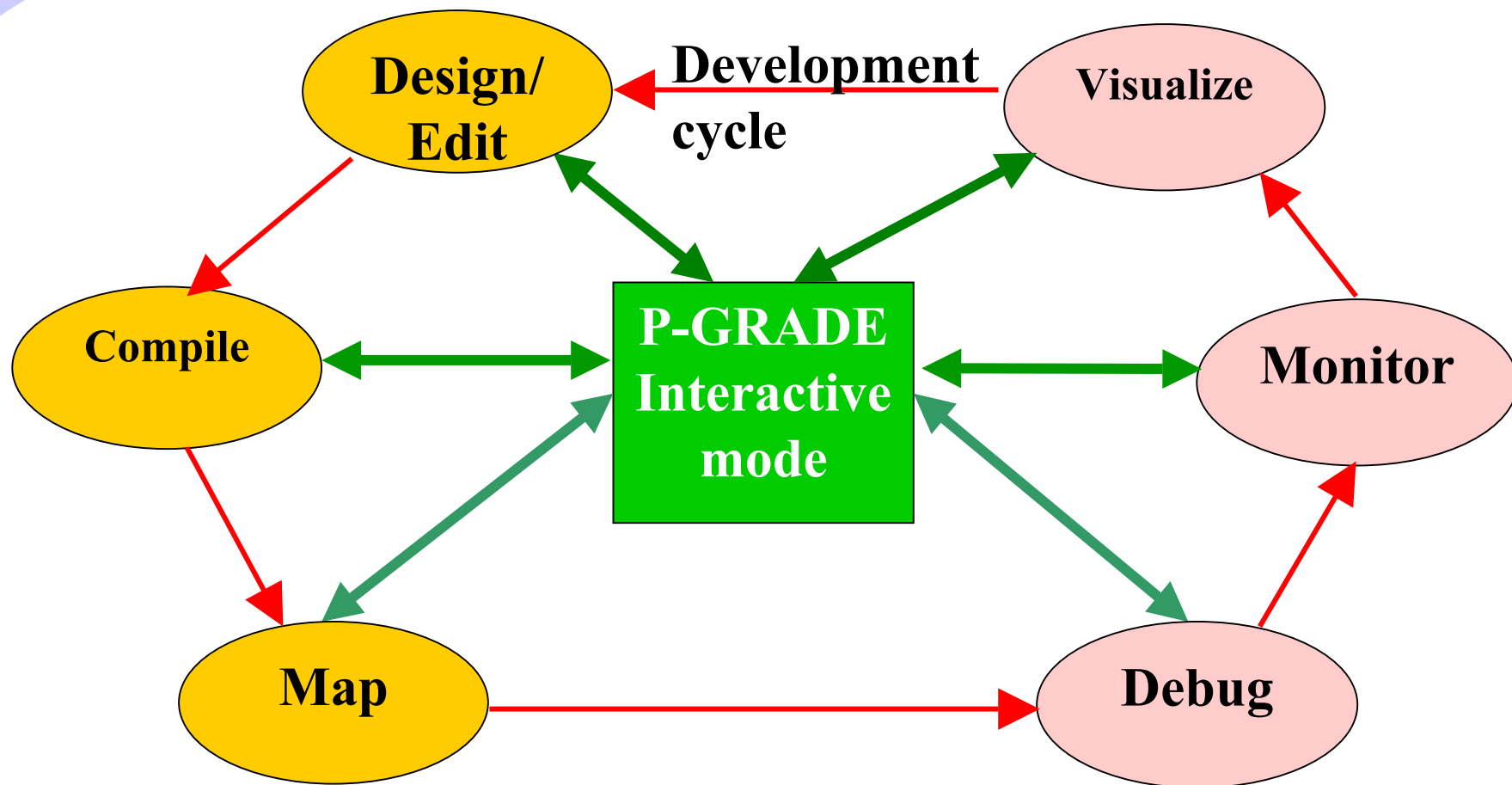
Application server is running!

1st philosopher is eating...
1st chopstick is used by 1st ph
2nd chopstick is released
1st philosopher EXITED!
1st chopstick is used by 2nd ph
2nd chopstick is used by 2nd ph
2nd philosopher is eating...

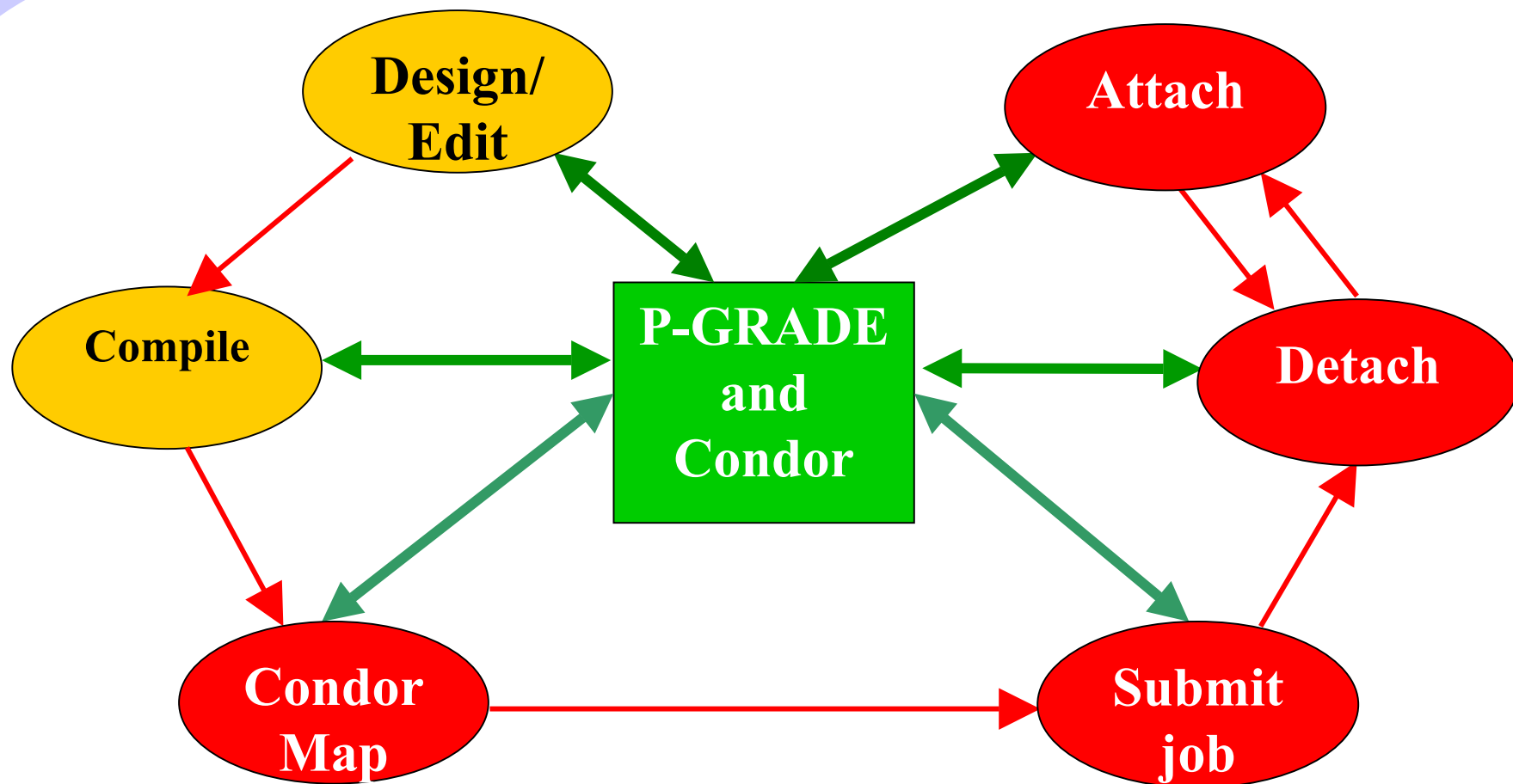


Features of P-GRADE

- Designed for **non-specialist programmers**
- Enables **fast reengineering** of sequential programs for parallel computers and Grid systems
- **Unified graphical support** in program design, debugging and performance analysis
- **Portability** on
 - supercomputers
 - heterogeneous clusters
 - components of the Grid
- **Two execution modes:**
 - **Interactive mode**
 - **Job mode**



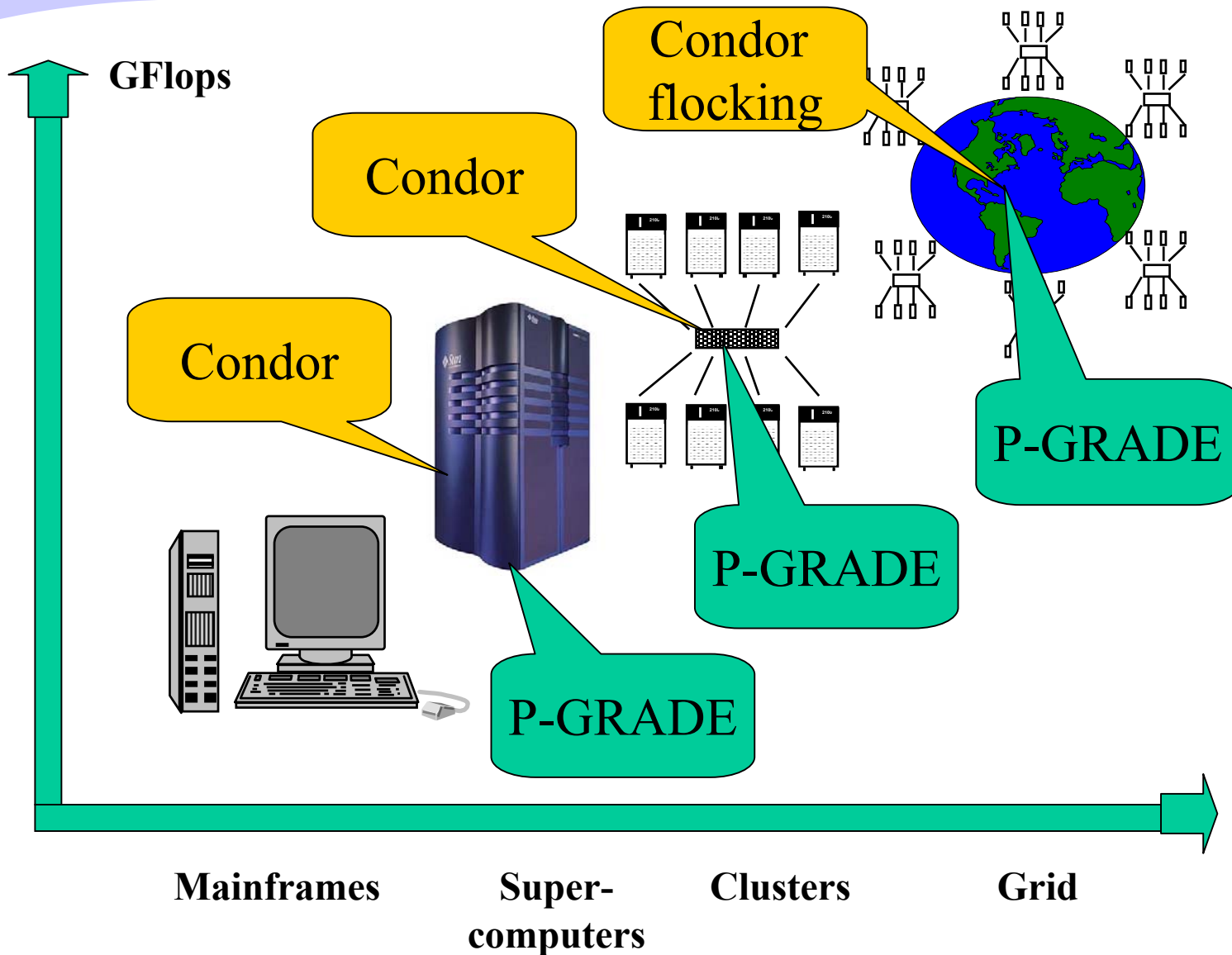
P-GRADE Job Mode with Condor



Typical usage on clusters or in the Grid

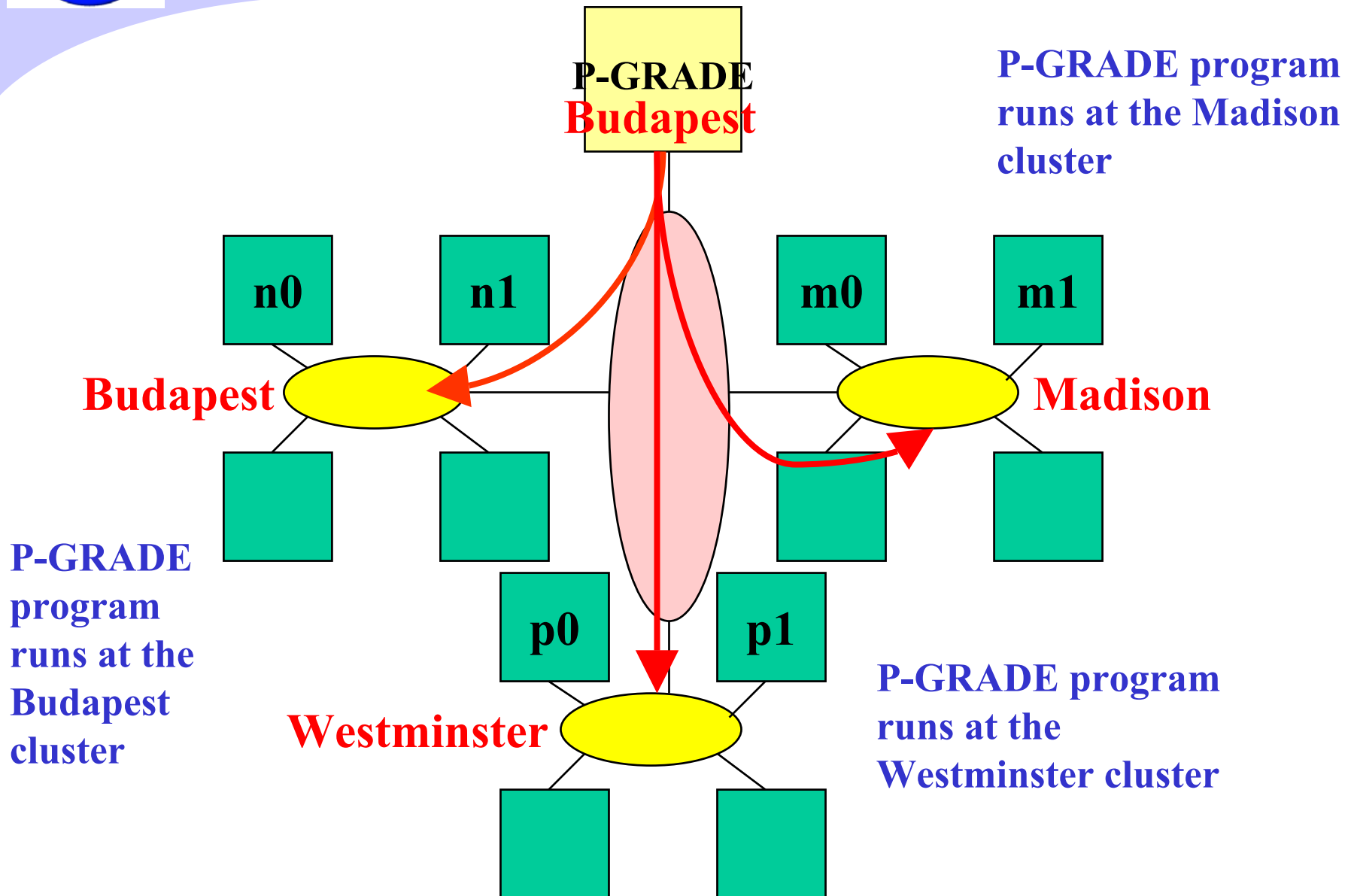


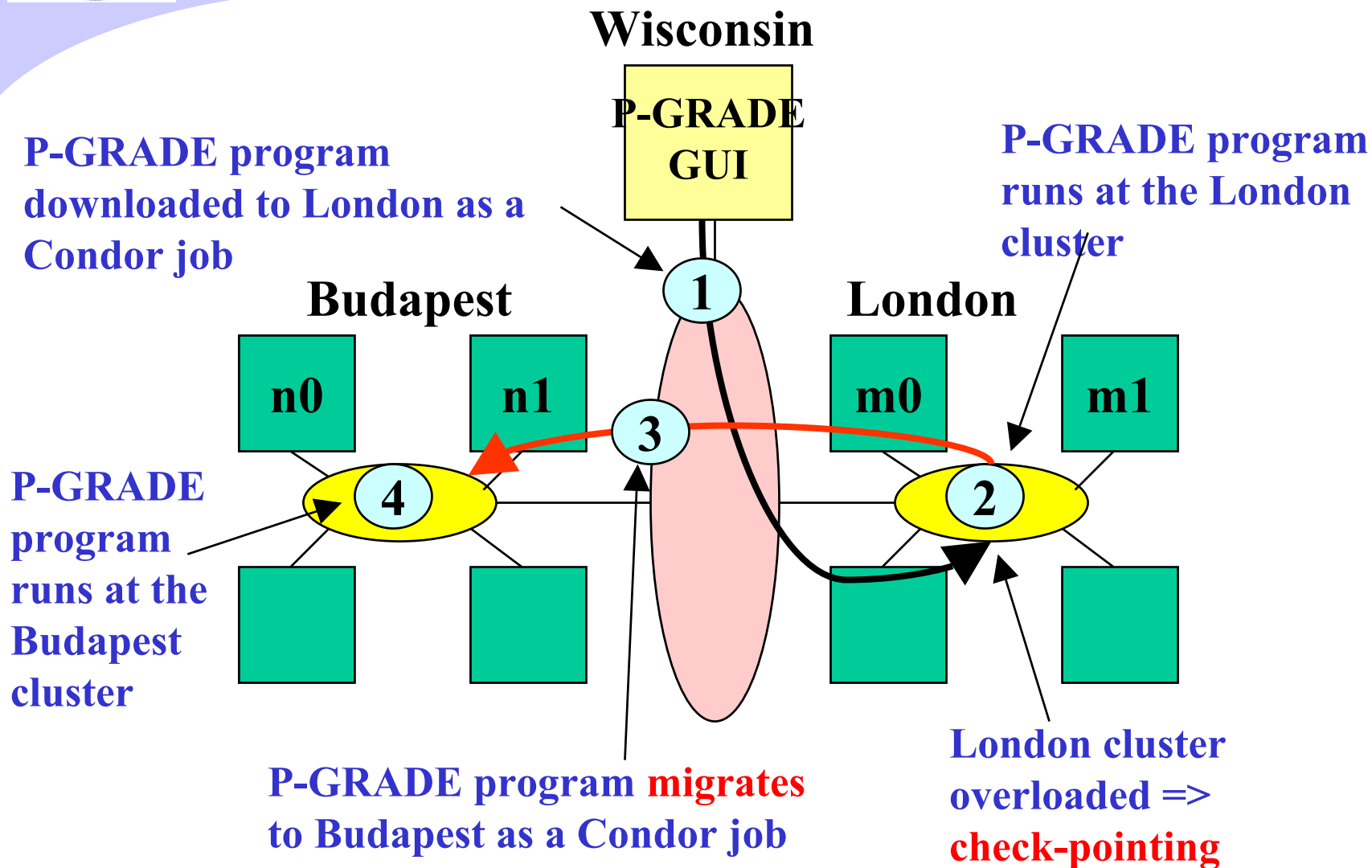
Condor/P-GRADE on the whole range of parallel and distributed systems





Berlin CCGrid Grid Demo workshop: Flocking of P-GRADE programs by Condor



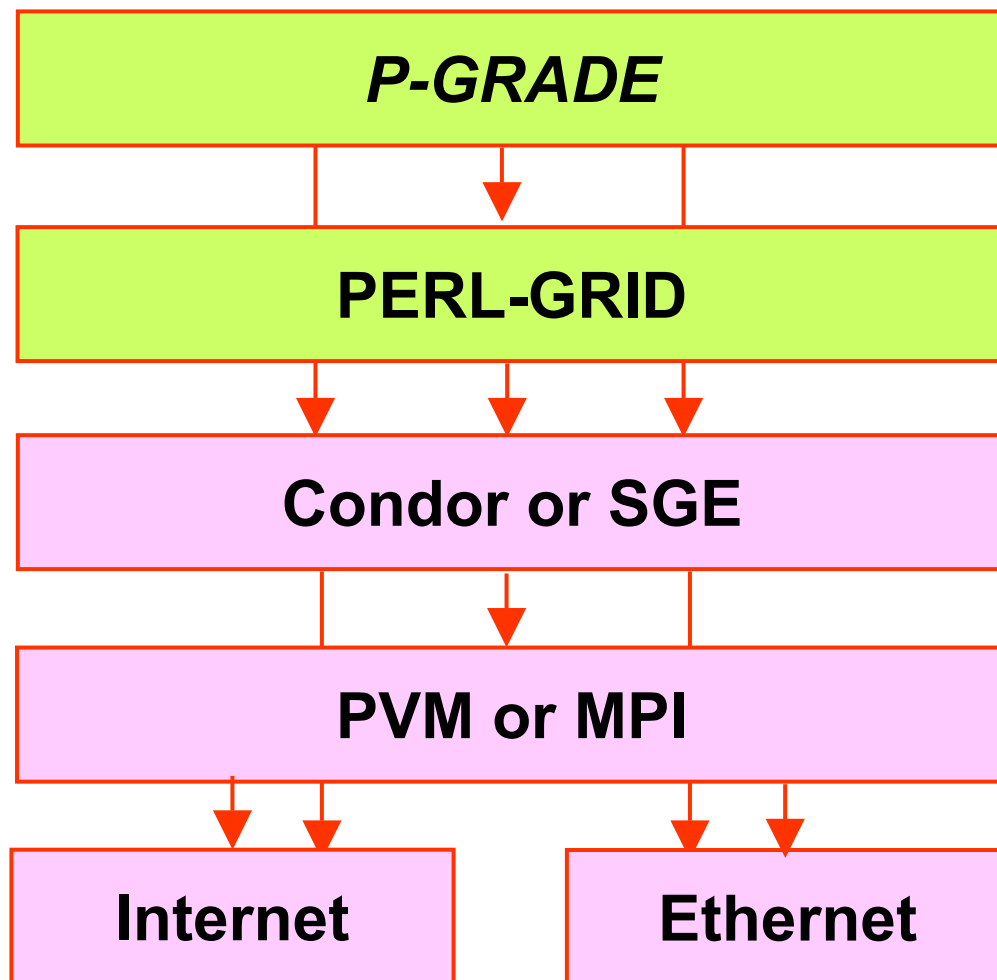




Further development: TotalGrid

- TotalGrid is a total Grid solution that integrates the different software layers of a Grid (see next slide) and provides for **companies** and **universities**
 - exploitation of free cycles of desktop machines in a Grid environment after the working/labor hours
 - **achieving supercomputer capacity** using the actual desktops of the institution without further investments
 - **Development and test** of Grid programs

Layers of TotalGrid



- A thin layer for
 - Grid level job management between P-GRADE and various local job managers like
 - Condor
 - SGE, etc.
 - file staging
- Application in the Hungarian Cluster Grid

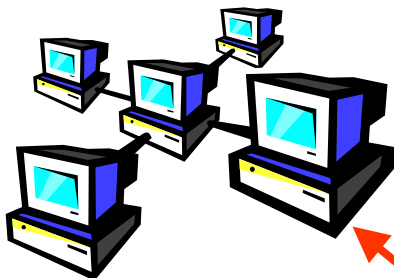


Hungarian Cluster Grid Initiative

- **Goal:** To connect the **99** new clusters of the Hungarian higher education institutions into a Grid
- Each cluster contains **20 PCs** and a network server PC.
 - **Day-time:** the components of the clusters are used for education
 - **At night:** all the clusters are connected to the Hungarian Grid by the Hungarian Academic network (**2.5 Gbit/sec**)
 - Total Grid capacity by the end of 2003: **2079 PCs**
- Current status:
 - About 400 PCs are already connected at 8 universities
 - Condor-based Grid system
 - VPN (Virtual Private Network)
- **Open Grid:** other clusters can join at any time

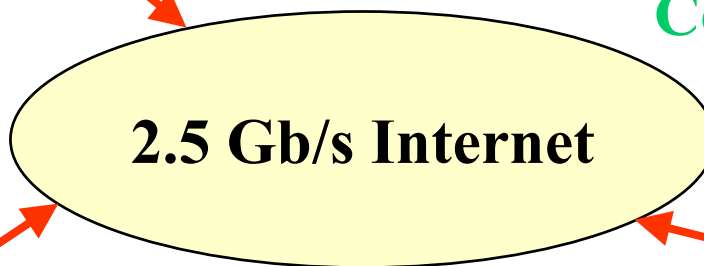
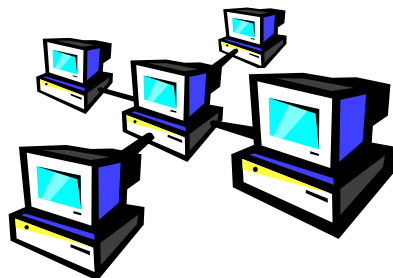
Structure of the Hungarian Cluster Grid

Condor => TotalGrid

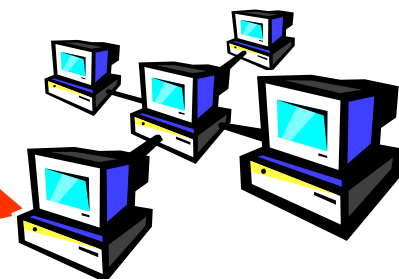


2003: 99*21 PC Linux clusters, total 2079 PCs

Condor => TotalGrid



Condor => TotalGrid

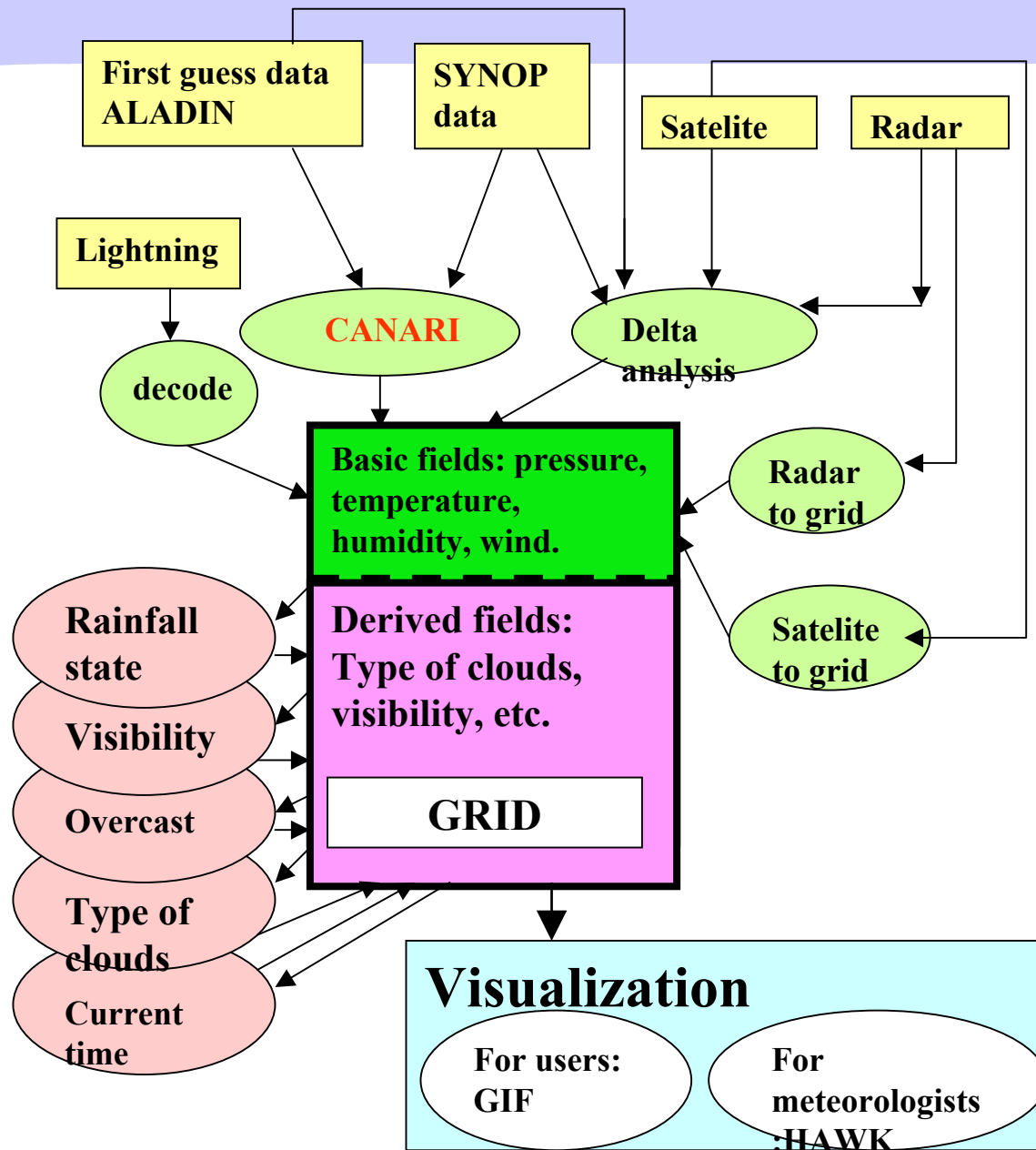




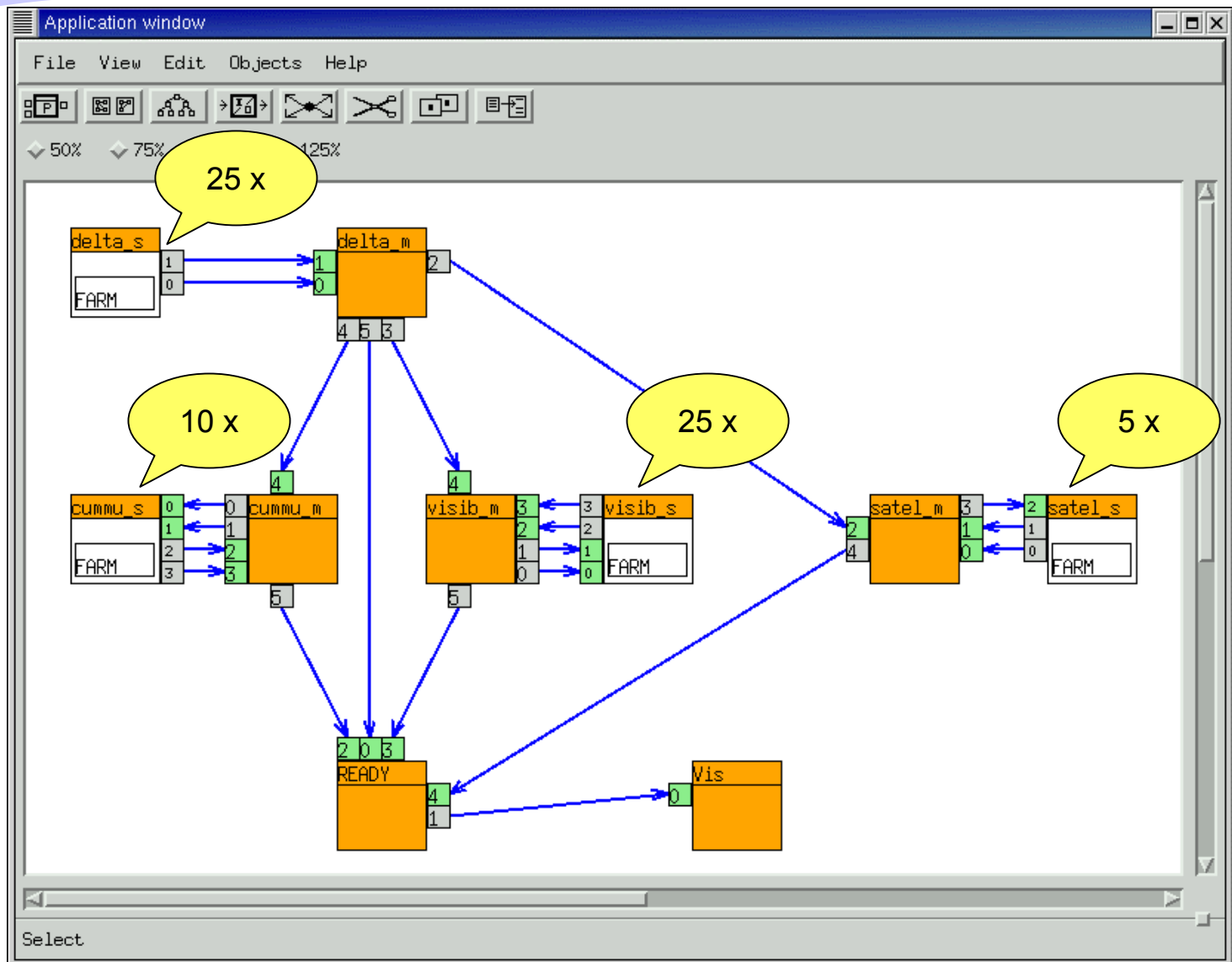
Live demonstration of TotalGrid

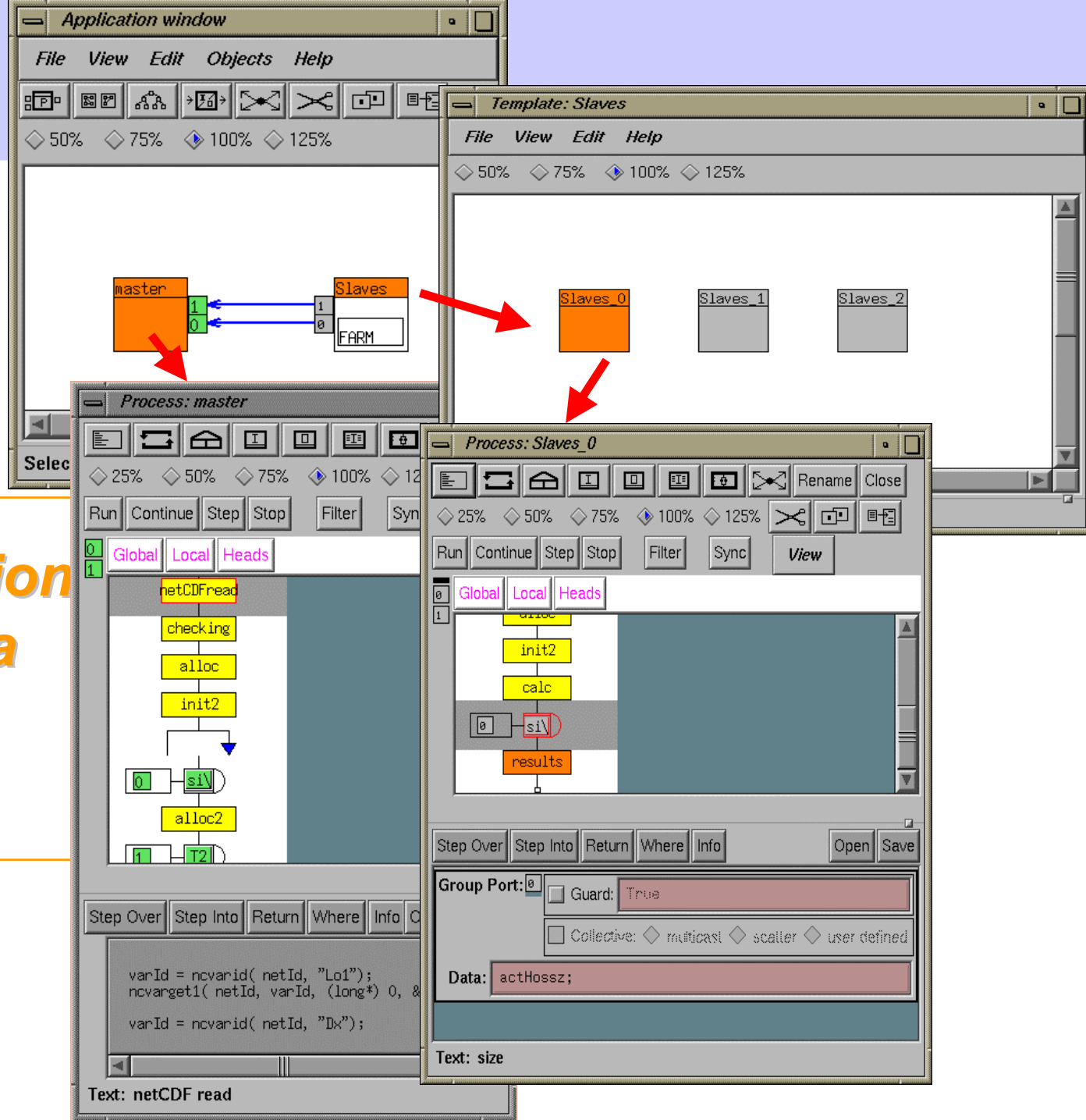
- MEANDER Nowcast Program Package:
 - **Goal:** Ultra-short forecasting (30 mins) of dangerous weather situations (storms, fog, etc.)
 - **Method:** Analysis of all the available meteorology information for producing parameters on a regular mesh (10km->1km)
- Collaborative partners:
 - OMSZ (Hungarian Meteorology Service)
 - MTA SZTAKI

Structure of MEANDER



P-GRADE version of MEANDER

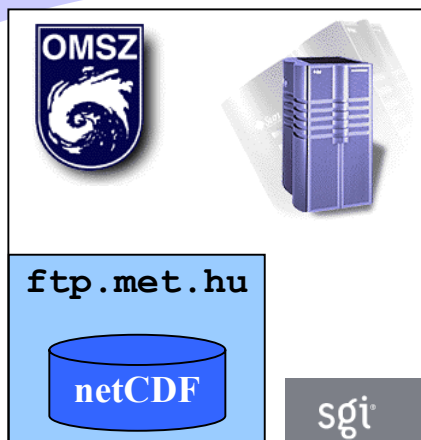




Implementation of the Delta method in P-GRADE



Live demo of MEANDER based on TotalGrid



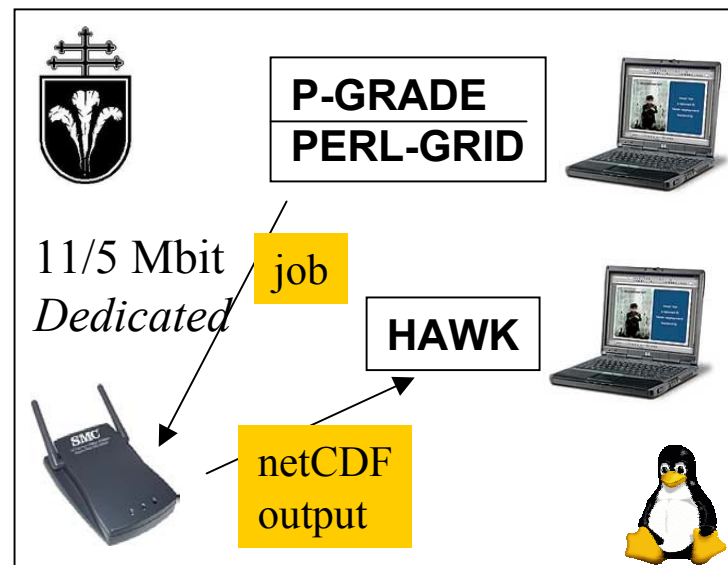
netCDF
input

512 kbit
Shared

34 Mbit
Shared

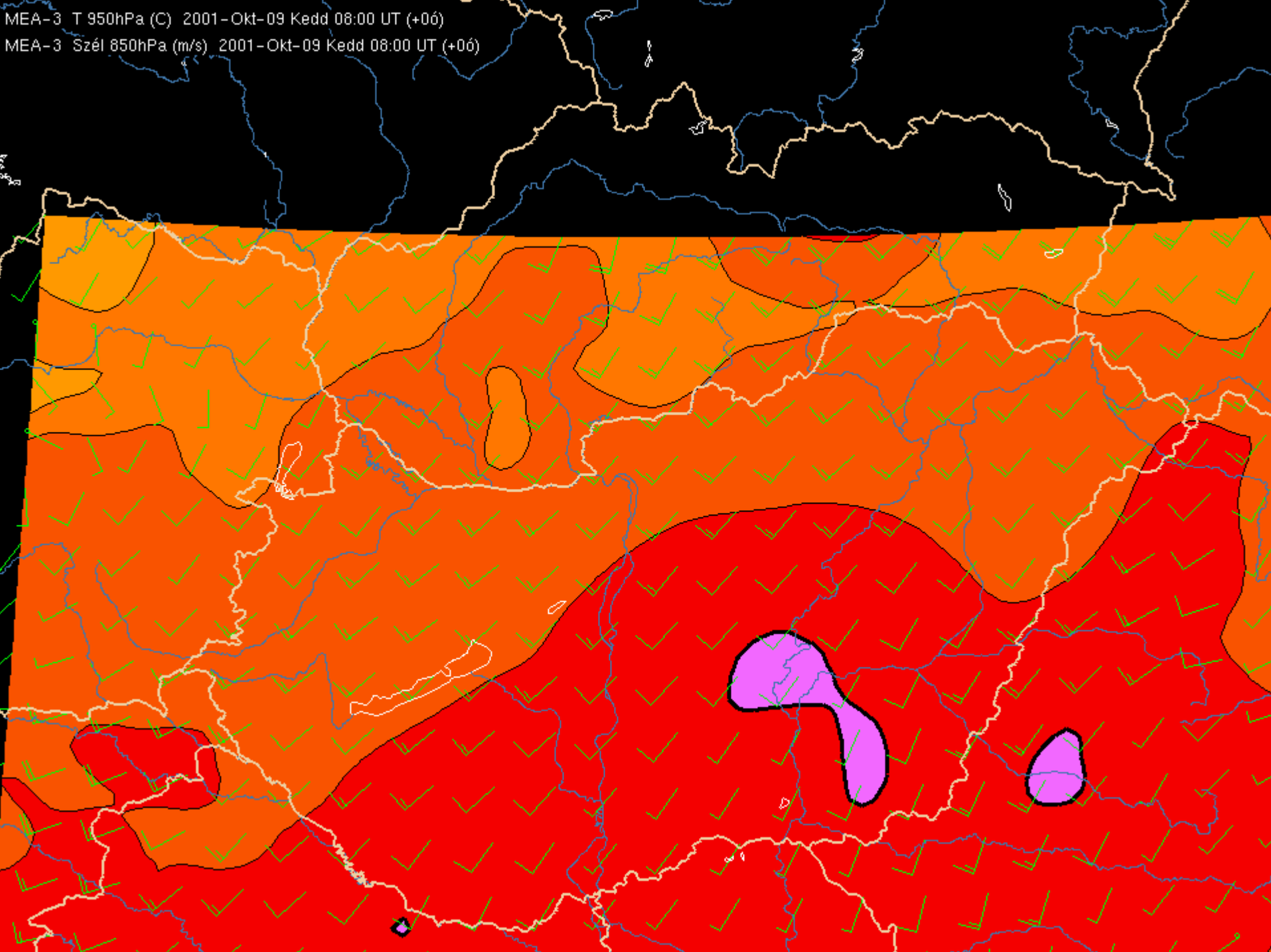
job

netCDF
output



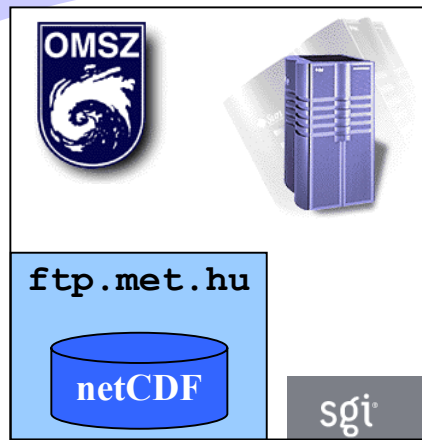
MEA-3 T 950hPa (C) 2001-Okt-09 Kedd 08:00 UT (+00)

MEA-3 Szél 850hPa (m/s) 2001-Okt-09 Kedd 08:00 UT (+00)





On-line Performance Visualization in TotalGrid



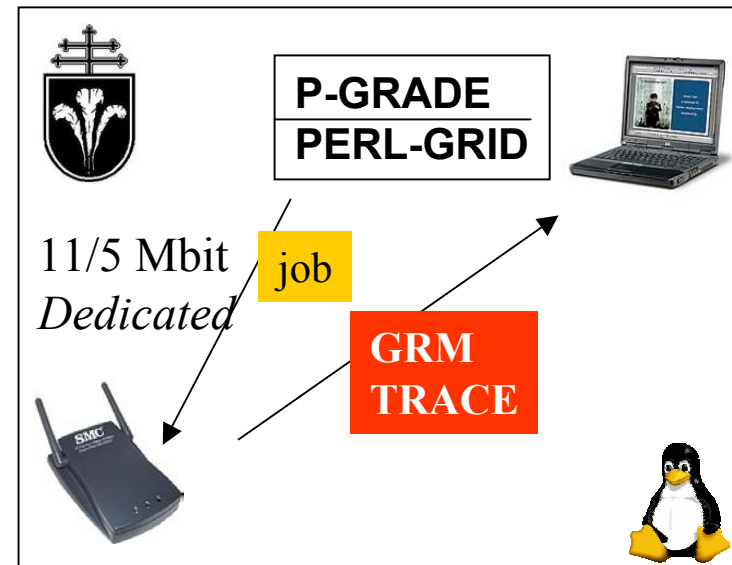
netCDF
input

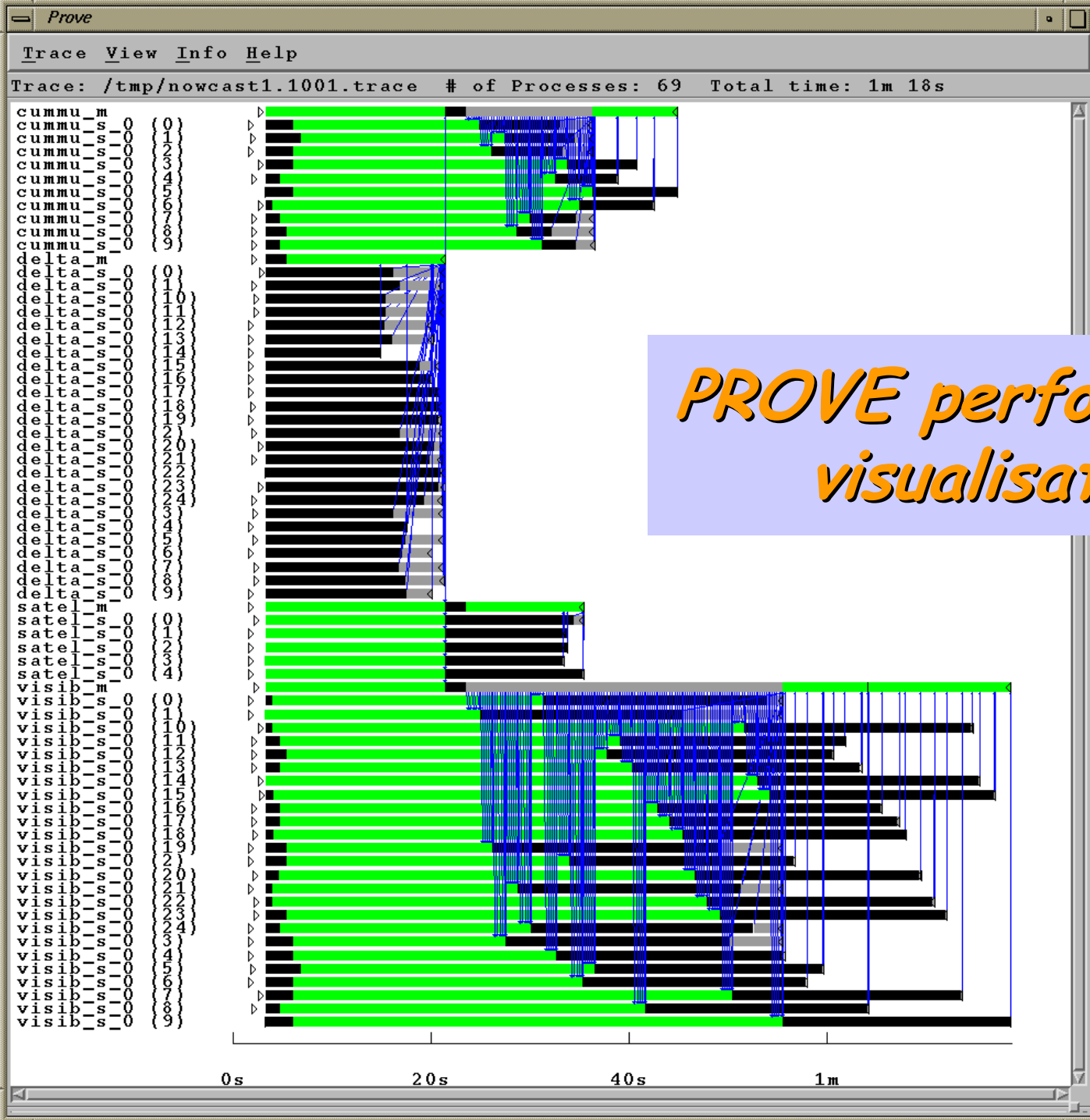
512 kbit
Shared

34 Mbit
Shared

job

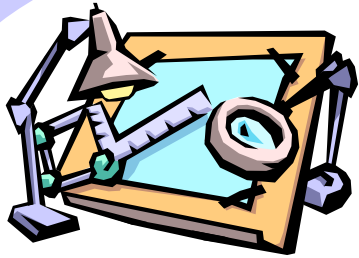
GRM
TRACE



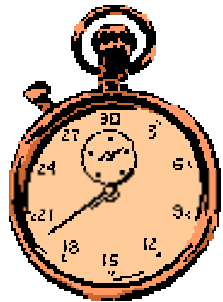
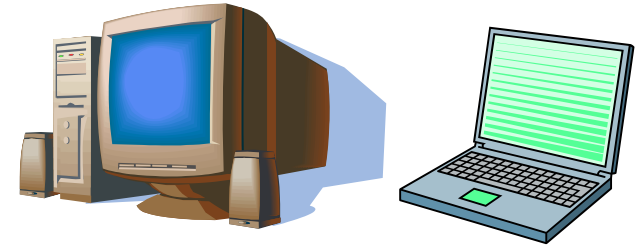




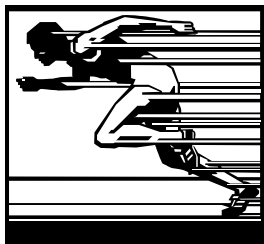
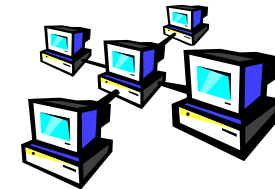
P-GRADE: Software Development and Execution



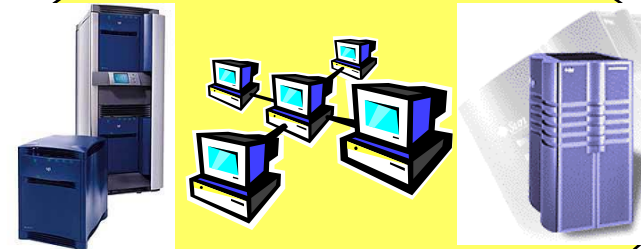
**Edit,
debugging**



**Performance-
analysis**



Execution



Grid



Applications in P-GRADE

Completed applications

- Meteorology: Nowcast package (Hungarian Meteorology Service)
- Urban traffic simulation (Univ. of Westminster)

Applications under development

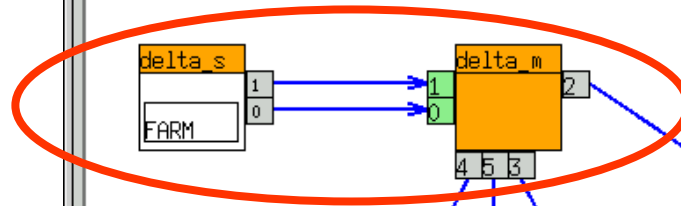
- Chemistry applications
- Smog forecast system
- Analysis of smog alarm strategies

Further extensions of P-GRADE

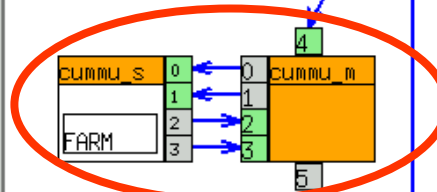
- Automatic check-pointing of parallel applications inside a cluster (already prototyped)
 - Dynamic load-balancing at
 - Fault-tolerant execution mechanism
- Automatic check-pointing of parallel applications in the Grid (under development)
 - Automatic application migration in the Grid
 - Fault-tolerant execution mechanism in the Grid
 - Saving unfinished parallel jobs of the Cluster Grid
- Extensions under design
 - Parameter study support
 - Connecting P-GRADE with GAT
 - Workflow layer for complex Grid applications

Workflow interpretation of MEANDER

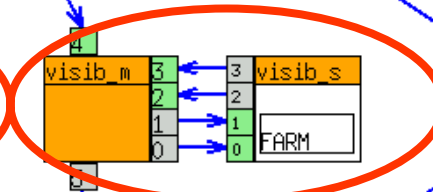
1st job



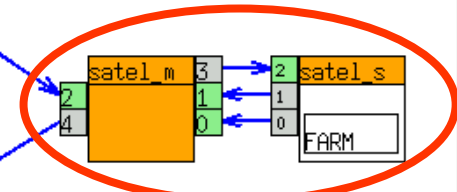
2nd job



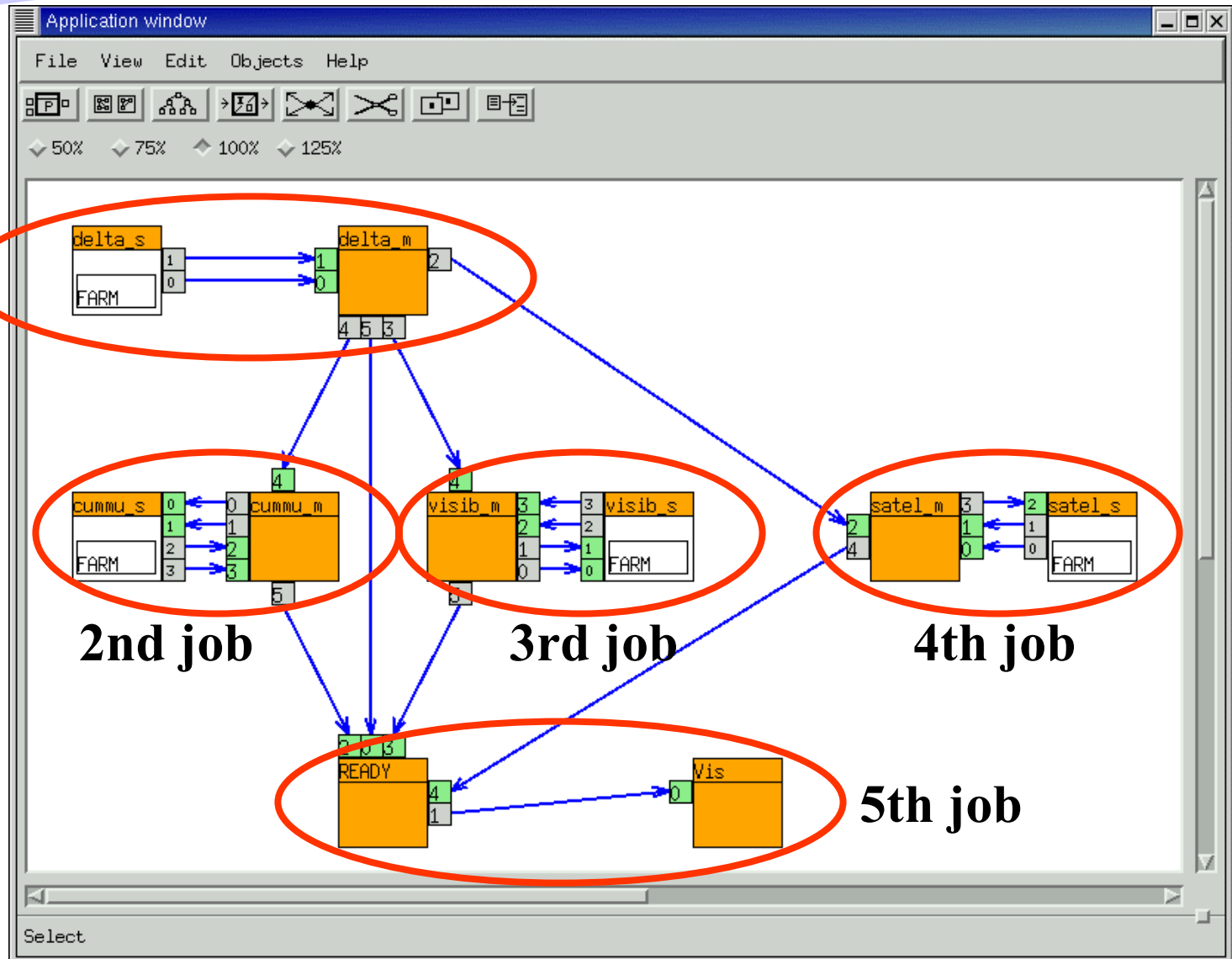
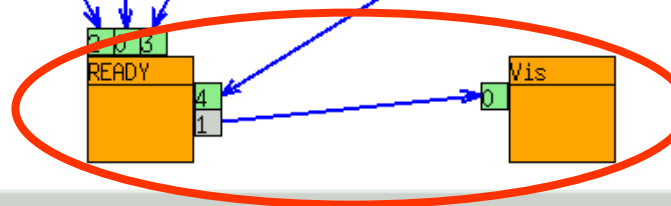
3rd job



4th job



5th job



Conclusions

- SZTAKI participates in the largest EU Grid projects and in all the Hungarian Grid projects
- **Main results:**
 - **P-GRADE** (SuperGrid project)
 - Integration of P-GRADE and Condor (SuperGrid)
 - demo at Berlin CCGrid
 - **TotalGrid** (Hungarian Cluster Grid)
 - Meteorology application in the Grid based on the P-GRADE and TotalGrid approaches
 - demo at the 5th EU DataGrid conference

- **Access of P-GRADE 8.2.2:**

www.lpds.sztaki.hu

Thanks for your attention



Further information: www.lpds.sztaki.hu