

Plugin development – extensibility of UNICORE

Piotr Bała

ICM Warsaw University

bala@icm.edu.pl.pl



Faculty of Mathematics and Computer Science
N. Copernicus University, Toruń



1st UNICORE Summit 11-12.10.2005

UNICORE projects at ICM

- **EUROGRID** (Nov. 2000 – Nov 2003)
 - ICM develops middleware for biomolecular applications,
 - ICM operates European grid for molecular biology
- **GRIP** (2002 – 2003)
 - ICM deployed interoperable biomolecular applications
- **UNIGRIDS** (2004 – 2006)
 - ICM develops high level services (visualization, database access, access to remote instruments)

Unicore plugins

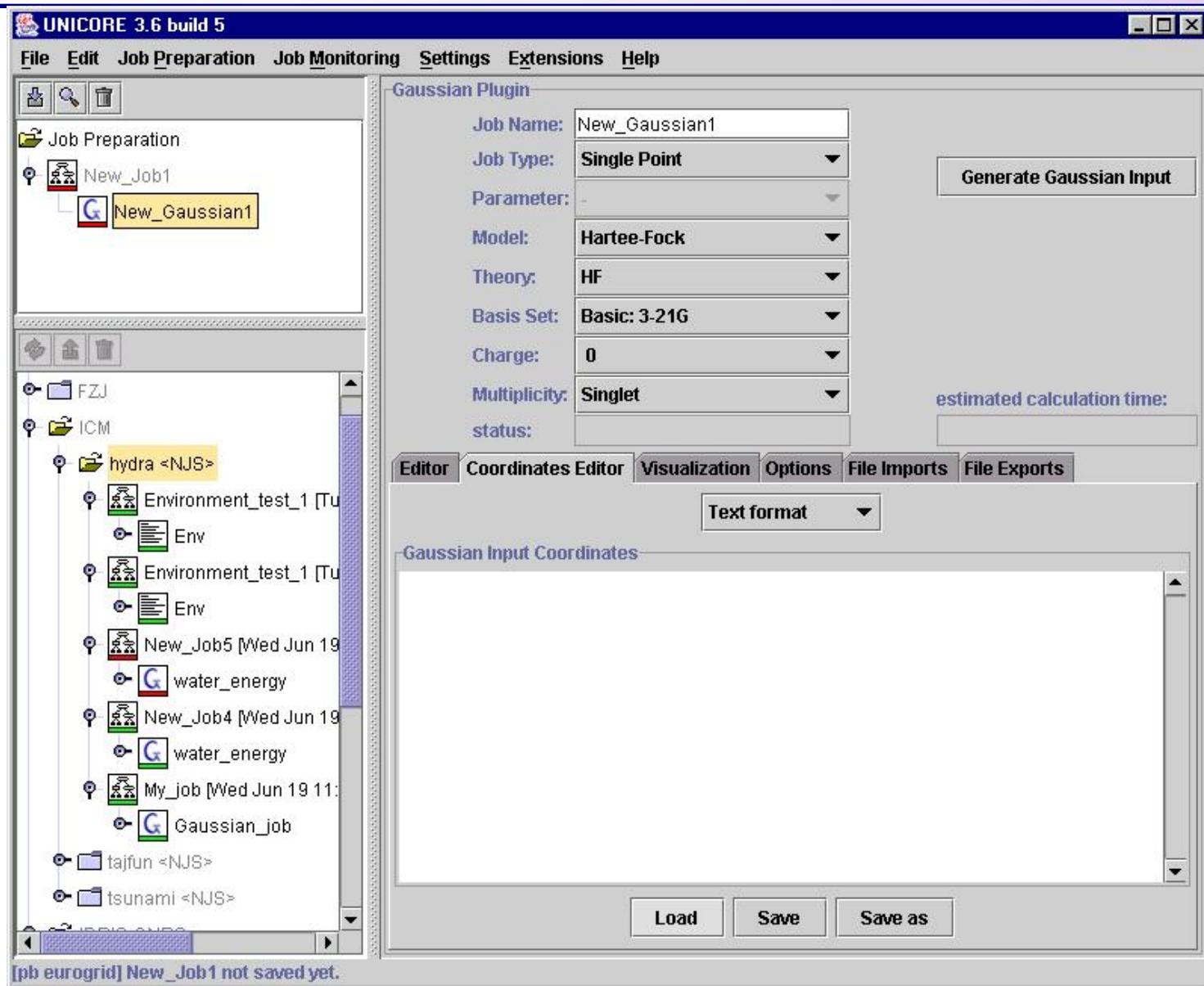
- UNICORE Client (5.3, 5.6)
- Example tasks:
 - **Quantum Chemistry:** Gaussian, TBDFT
 - **Molecular Dynamics:** Amber, Gromos
- Plugins:
 - **Quantum Chemistry:** Gaussian, Gamess, CPMD (V. Huber FZJ)
 - **Molecular Dynamics:** AMBER
 - **BioInformatics:** BLAST (S. Haubold, TU Dresden)
 - **DataBase Access:** PDB Search, DBAccess
 - **Visualization:** SapphireVis, IADemo, IVis
 - **Broker:** Resource Broker (UoM)
 - **Other:** ListAllJobs, PluginLoaderPlugin

Gaussian98 plugin

J. Pytliński

Input preparation
reads existing input,
recognizes keywords

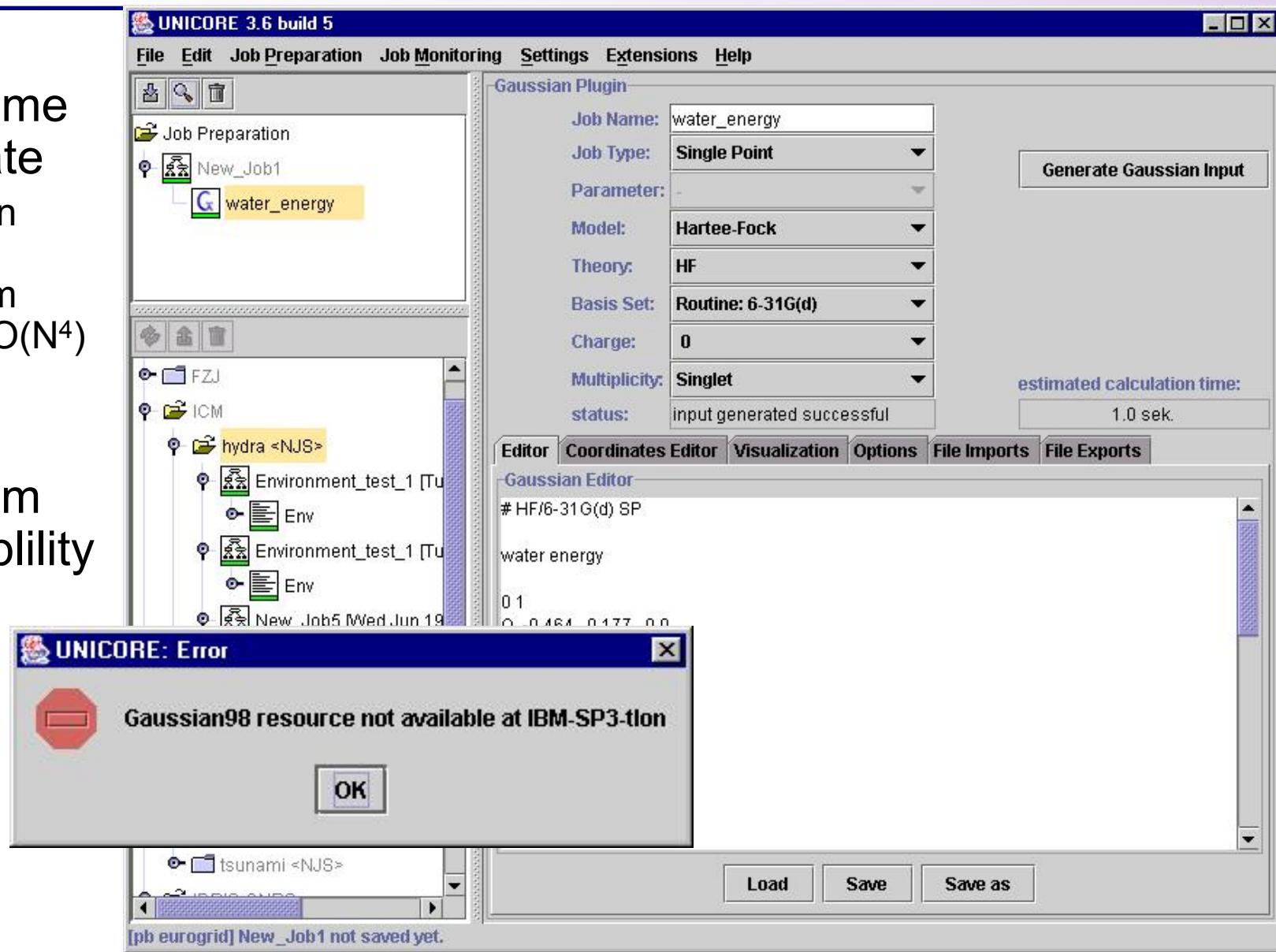
Molecule coordinate editor
formats:
•XYZ
•Z-matrix
•text



Gaussian98 plugin

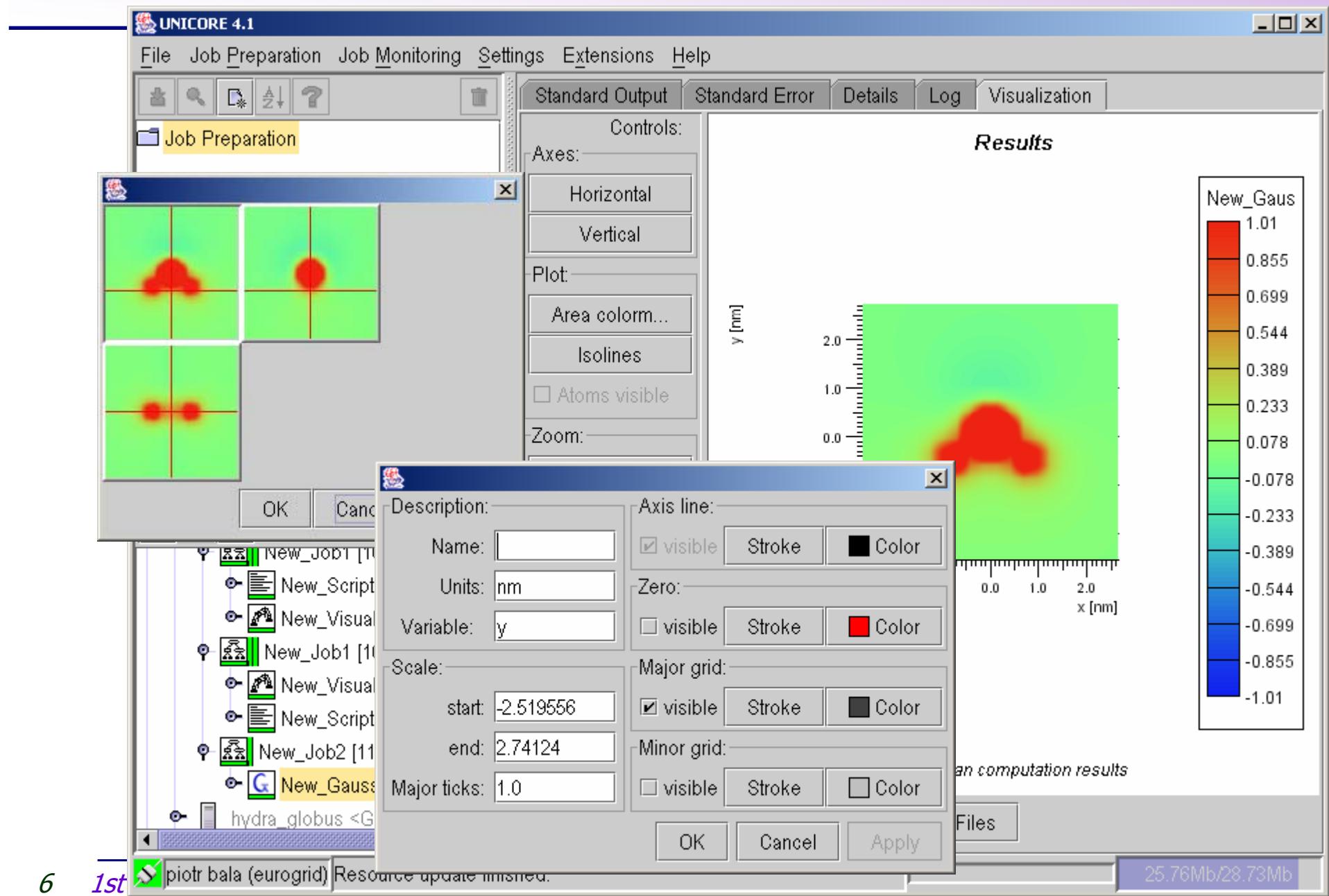
CPU time estimate based on known algorithm scaling $O(N^4)$

Check program availability



Gaussian Plugin

J. Pytliński, A. Nowiński



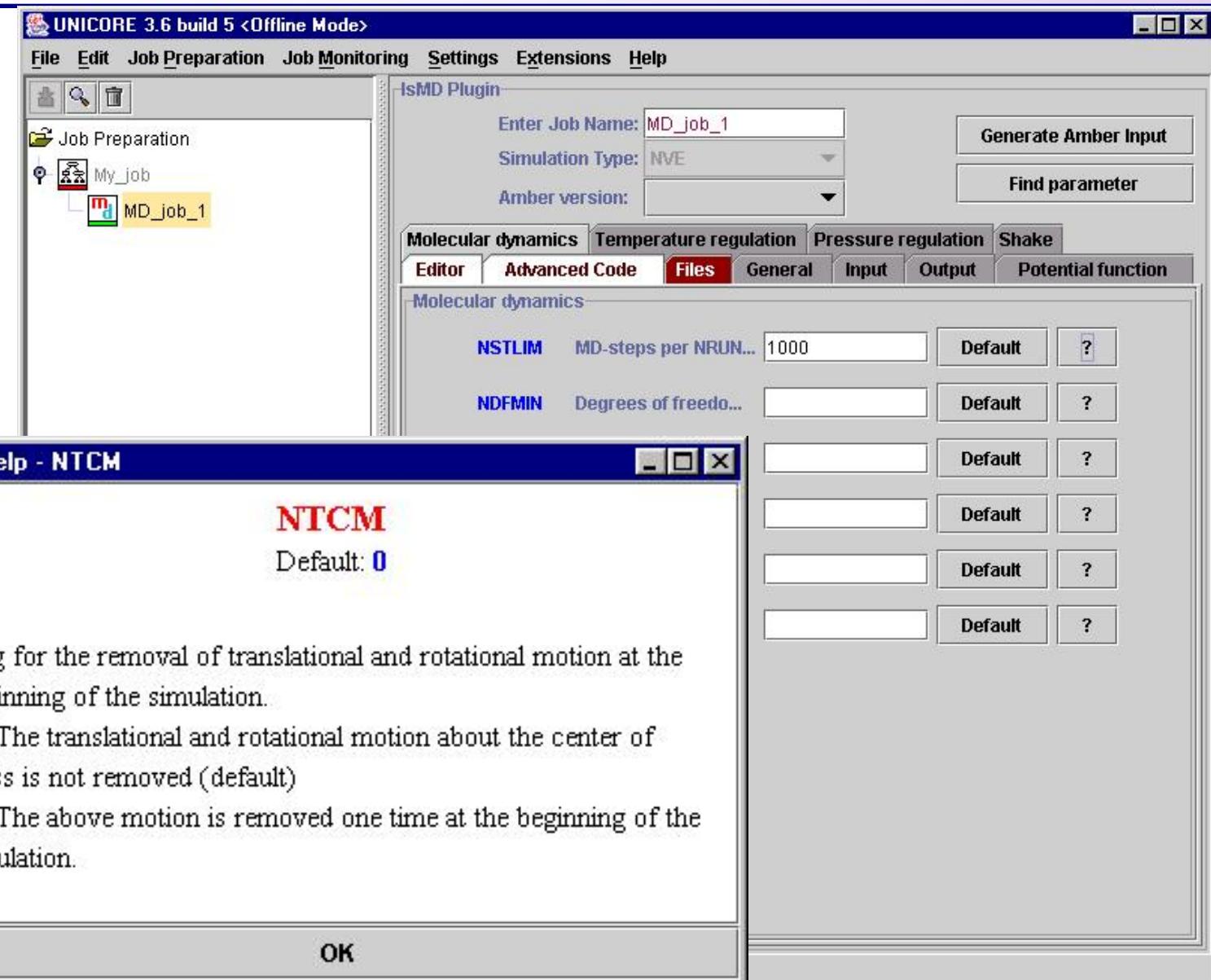
Amber plugin

Ł. Skorwider, J. Wypychowski

Panel for each group of options

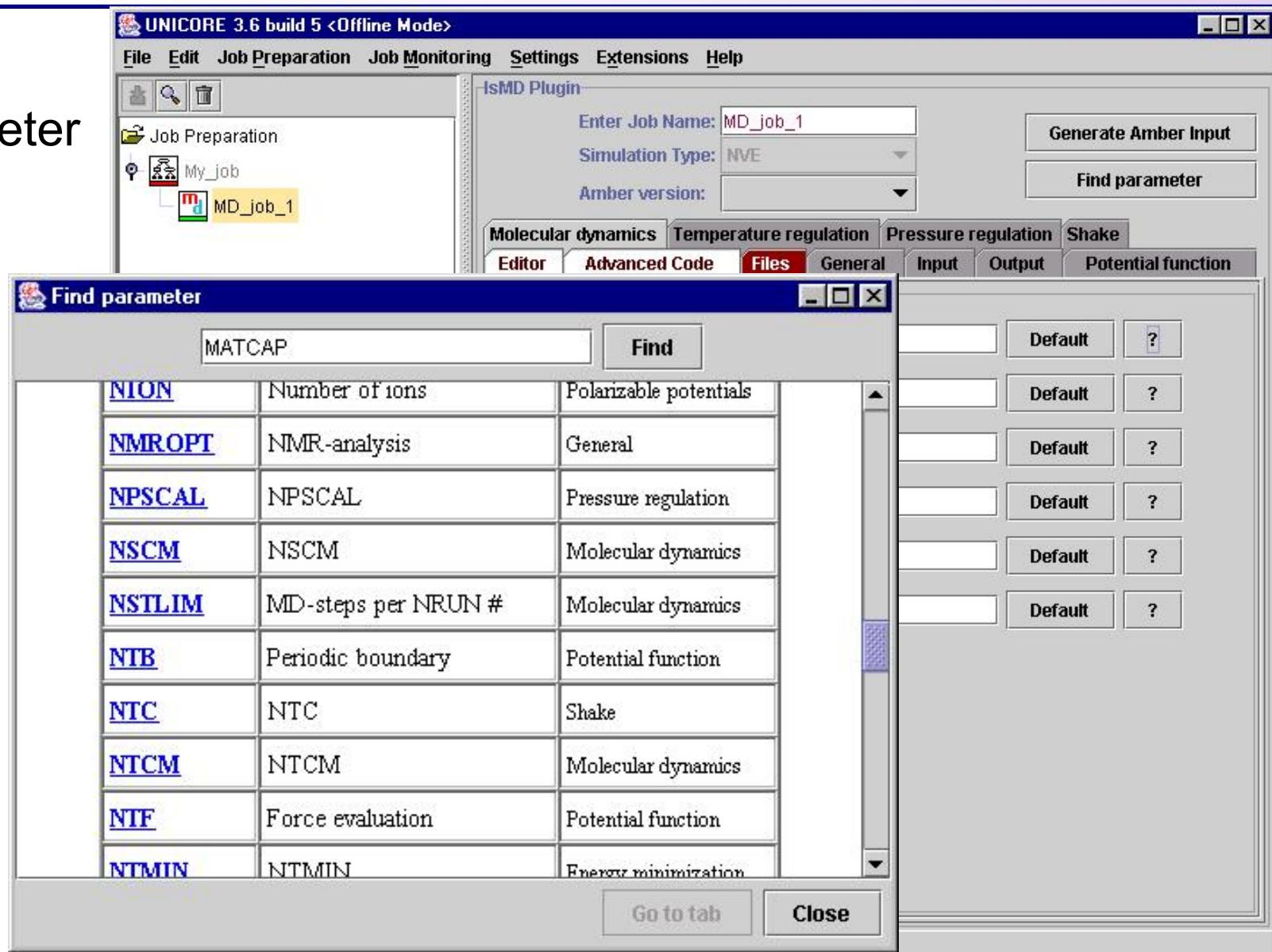
Default options

Help window



Amber plugin

Find parameter button



List All Jobs Plugin

K. Benedyczak

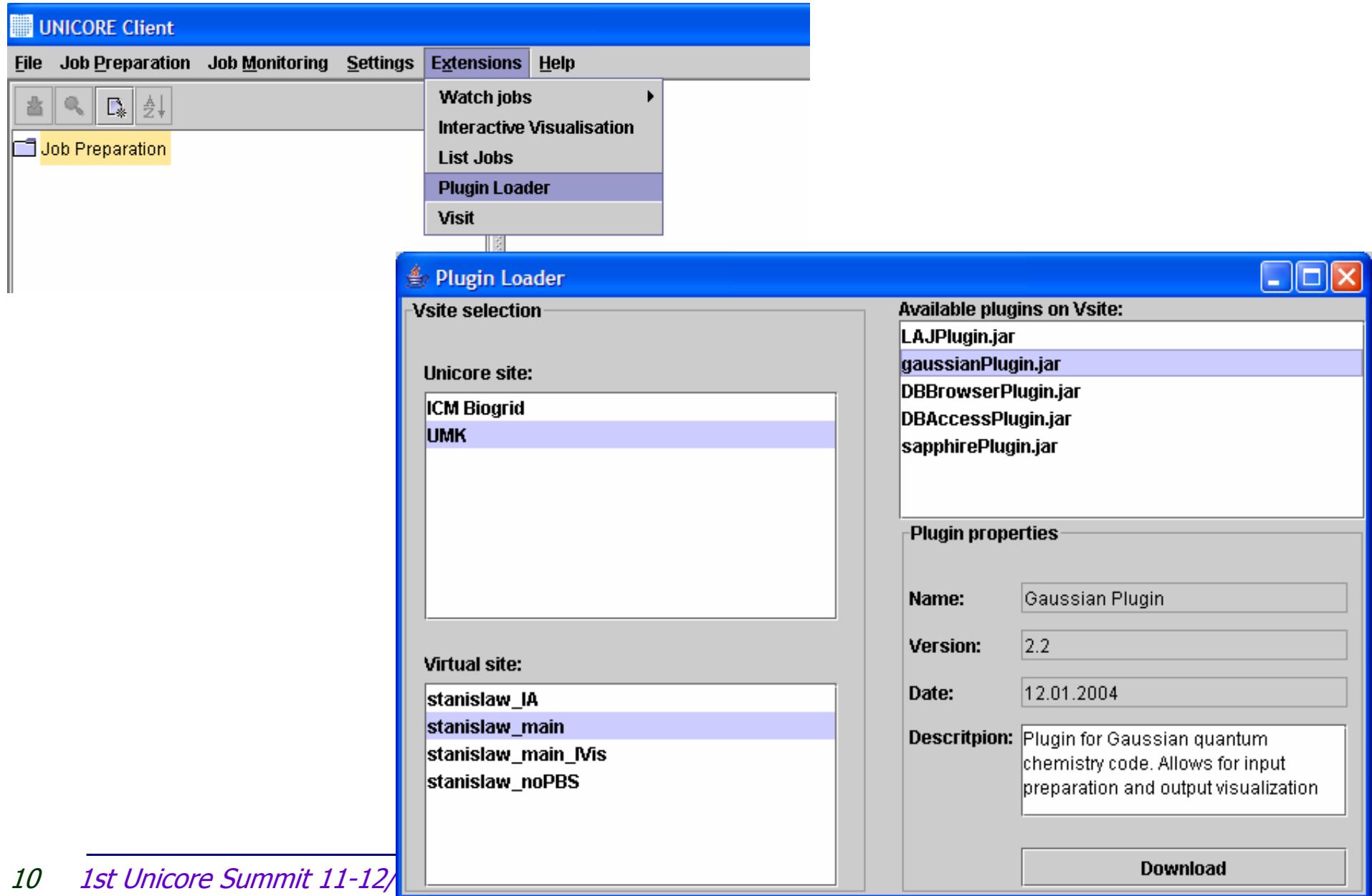
The screenshot shows two windows of the 'LAJ Window' application. The top window is a header bar with columns: Usite, Vsite, Job name, Status, Date, and ID. The bottom window is a detailed view of job entries:

Usite	Vsite	Job name	Status	Date	ID
UoM	O300_wren_aft	UNAVAILABLE	UNAVAILABLE	UNAVAILABLE	UNAVAILABLE
UoM	PS2_frik	UNAVAILABLE	UNAVAILABLE	UNAVAILABLE	UNAVAILABLE
UoM	solttest	UNAVAILABLE	UNAVAILABLE	UNAVAILABLE	UNAVAILABLE
Parallab	SGI_Origin2000	Environment_test_1	SUCCESSFUL	11-05 00:02:22	161d71f3
Parallab	SGI_Origin2000	Gaussian98_job	SUCCESSFUL	10-25 02:19:12	8fe7ffff
IDRIS-CNRS	IBM-SP3-tlon	New_Job1	NOT_SUCCESSFUL	11-01 20:10:02	8b3fe892
Parallab	SGI_Onyx2	New_Job1	NOT_SUCCESSFUL	11-01 20:08:59	8b3fe83c
FZJ	IBM44P-170_aft	UNAVAILABLE	UNAVAILABLE	UNAVAILABLE	UNAVAILABLE
ICM	tajfun	TOO SLOW	TOO SLOW	TOO SLOW	TOO SLOW
ICM	hydra	New_Job1	NOT_SUCCESSFUL	11-20 06:19:54	4543fb9f

Buttons at the bottom include 'Running threads: 3', 'Last refresh time: Mon Nov 25 17:35:29 CET 2002', ' Show also control jobs', 'Refresh', 'Close', and 'Show in JMC'.

PluginLoader Plugin

K. Benedyczak



ResourcesList plug-in

Michał Wroński

The screenshot shows a window titled "ResourceList" with two main sections: "Hardware resources" and "Software resources".

Hardware resources:

Usite	Vsite	Nodes	Proc/Node	Speed	Disk	Memory
ICM	tsunami	12	1	200.0 Me...	100.0 Me...	108.0 Me...
ICM	tajfun	1	32	1200.0 M...	100.0 Me...	16384.0 ...
ICM	hydra	4	1	200.0 Me...	100.0 Me...	128.0 Me...
Parallab	Parallab	1	4	200.0 Me...	100.0 Me...	2056.0 M...
UoM	T3E_turing	512	1	1200.0 M...	100.0 Me...	256.0 Me...

Software resources:

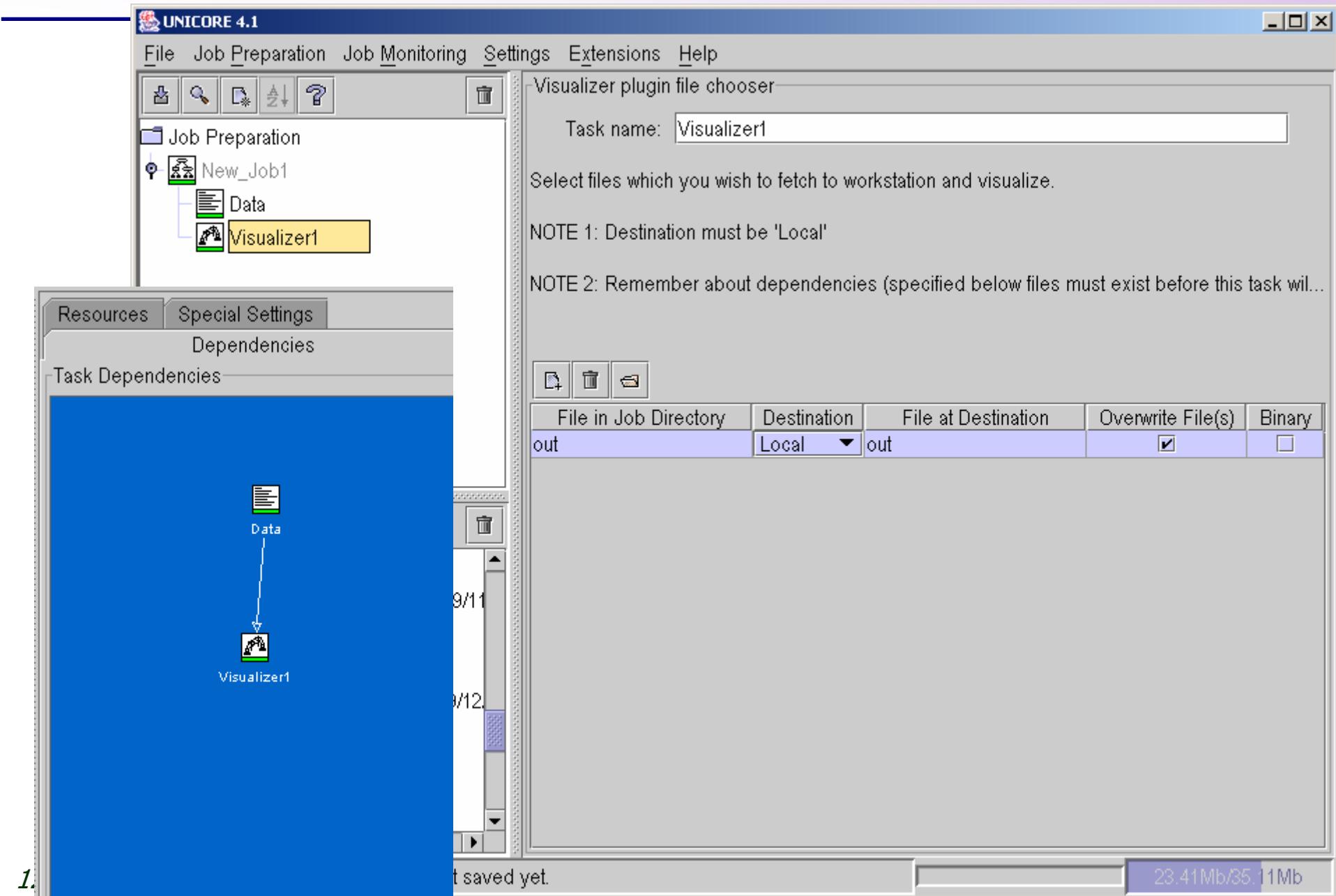
Usite	Vsite	Software
ICM	hydra	Gaussian98
ICM	hydra	Amber
ICM	hydra	Charmm
ICM	hydra	Gromos96
ICM	tajfun	Gaussian98
ICM	tajfun	Gromos96
ICM	tsunami	Gaussian98

At the bottom of the window are buttons for "Save as XML" and "OK".

- List of resources
- Software resources
- Software resources can be filter

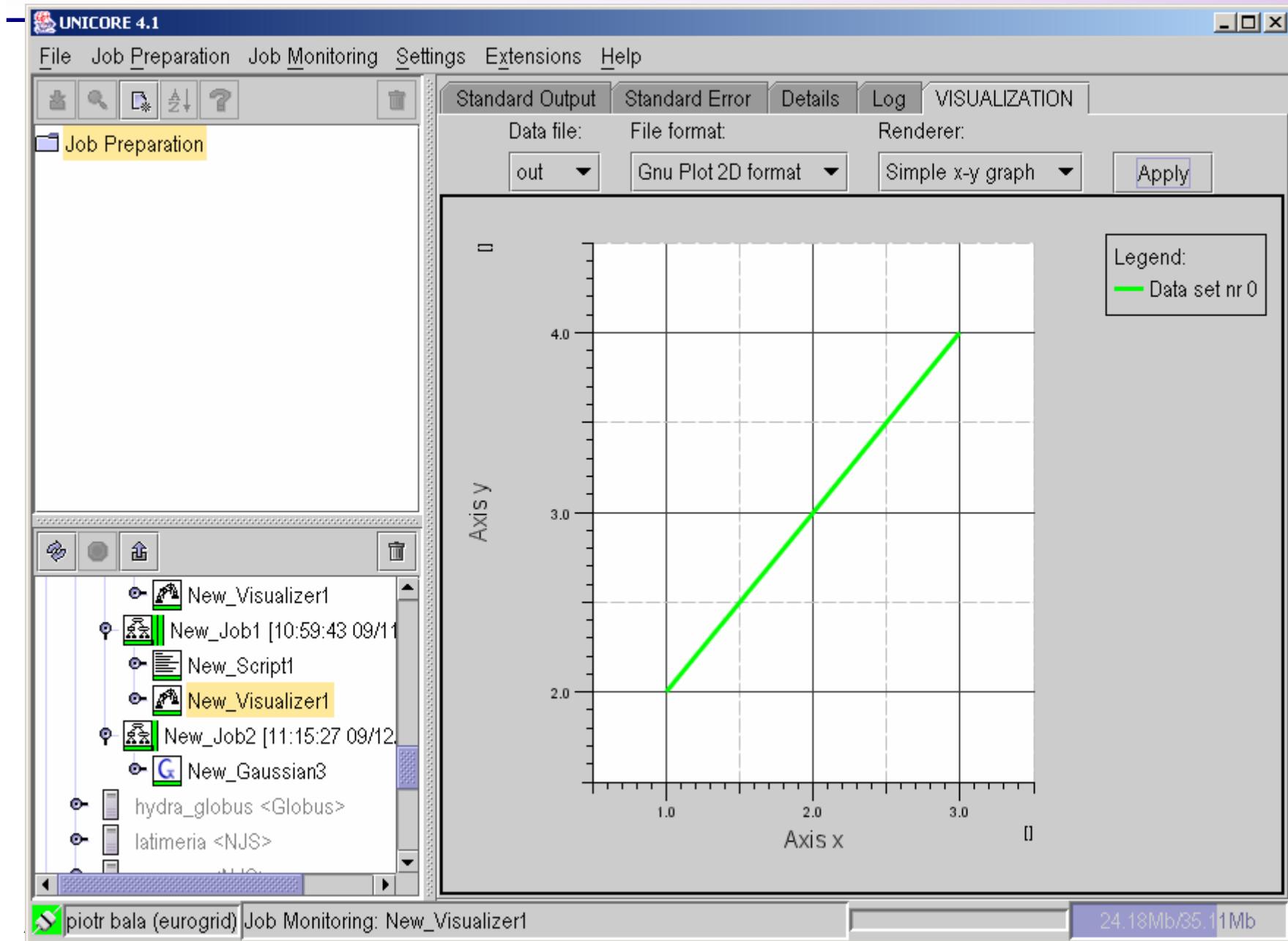
Visualizer Plugin

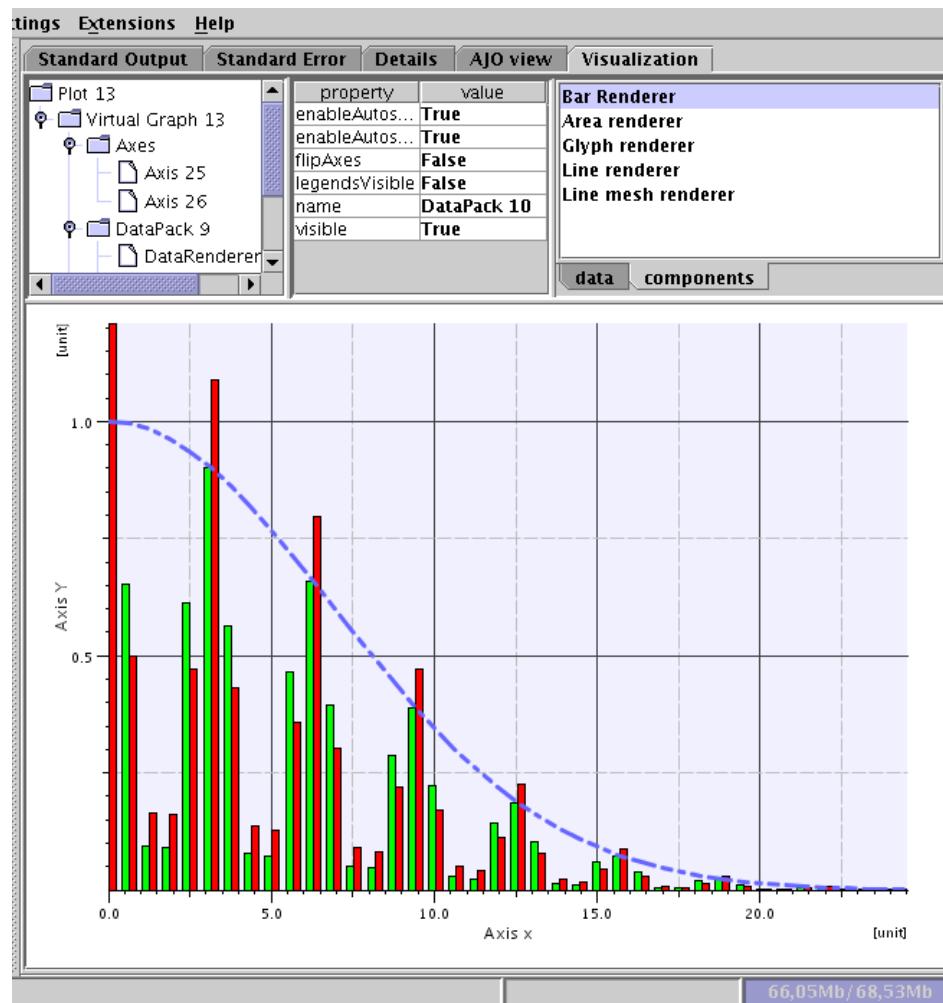
K. Benedyczak, A. Nowiński



Visualizer Plugin

K. Benedyczak, A. Nowiński





- Provides quick visualization for results files
- Allows quite advanced 2D visualization
- Easy in use
- Simple sample of sapphire library power

PDB search plugin

M. Wroński

Simple and advanced search

User can use different mirrors

The best one is selected (fastest answer)

The screenshot shows the 'PDB Search' plugin window. At the top, there's a menu bar with tabs: 'SearchLite' (selected), 'Adv. Search', 'Display Options', 'Query Results', 'Details', and 'Visualization'. Below the menu, there are several search fields and options:

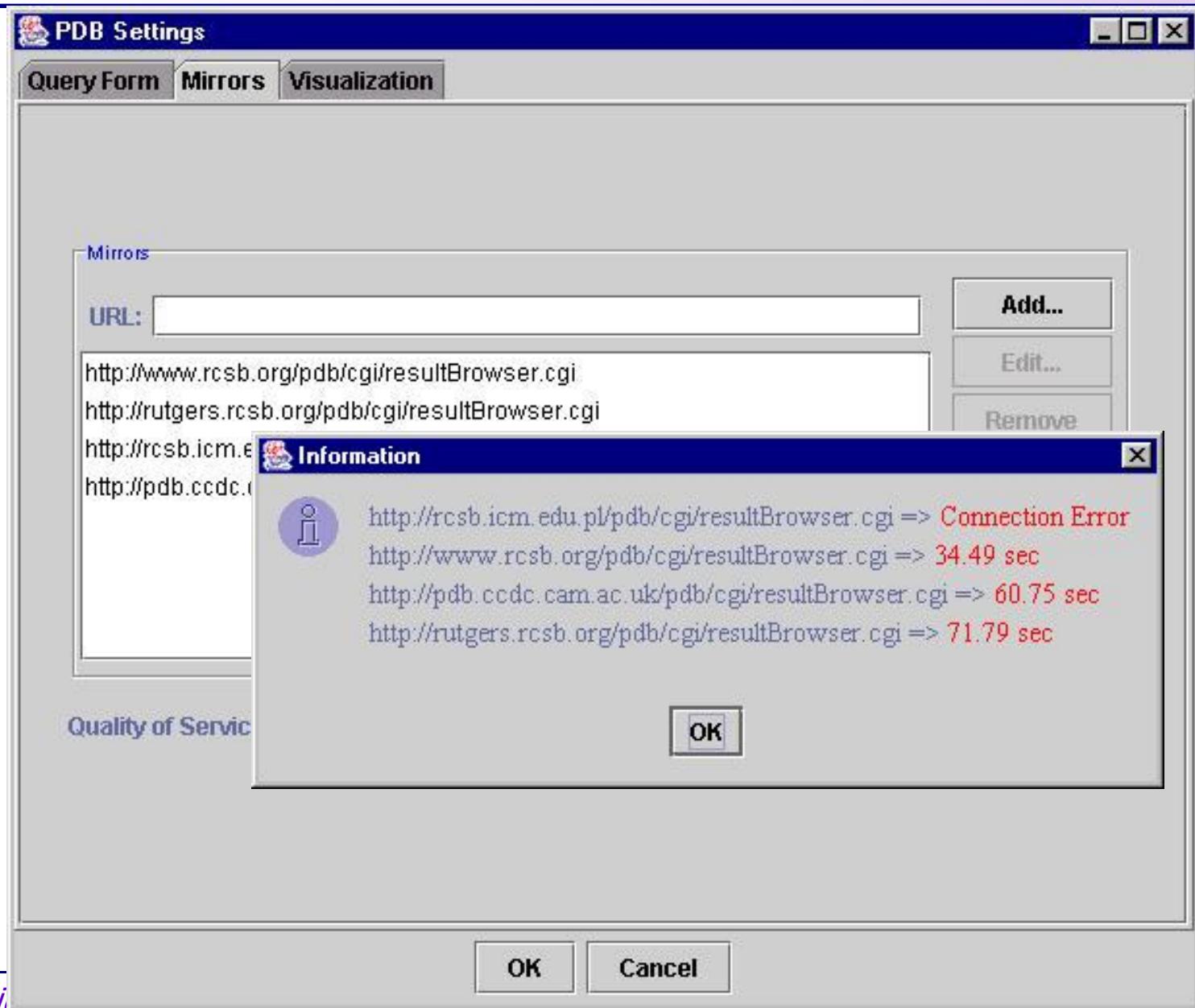
- PDB Identifier:** A text input field.
- Citation Author:** A text input field with a checkbox below it: Authors of primary citation only!
- Contains Chain Type:** A section with two rows of radio buttons:

Protein:	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	DNA:	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Enzyme:	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	RNA:	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Glycoprotein:	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	DNA/RNA hybrid:	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Carbohydrate:	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>				
- Compound Information:** A text input field.
- Title:** A text input field.
- PDB HEADER:** A text input field.
- Exp. Technique:** A dropdown menu with the option -- select one --.
- Text Search:** A text input field.
- Selecting Options:** A section with a checkbox: Select a subset of results with sequence homologues removed, followed by a dropdown menu set at 90% sequence identity.

At the bottom of the window are three buttons: 'Adv. Search', 'Reset Form', and 'Cancel'.

User can use different mirrors

The best one is selected (fastest answer)



Search results

The screenshot shows a Windows application window titled "PDB Search". The menu bar includes "File", "Edit", "View", "Search", "Help", and "About". Below the menu is a toolbar with icons for "SearchLite", "Adv. Search", "Display Options", "Query Results", "Details", and "Visualization". The main content area displays the following search results:

Title: Crystal Structure Of Human Tyrosine-Protein Kinase C-Src, In Complex With Amp-Pnp

Compound: Mol_Id: 1; Molecule: Tyrosine-Protein Kinase Src; Chain: Null; Fragment: Residues 86-836, Containing Sh2, Sh3, Kinase 2 Domains and C-Terminal Tail; Synonym: C-Src, P60-Src; Ec: 2.7.1.112; Engineered: Yes; Biological Unit: Monomer

Authors: W. Xu, A. Doshi, M. Lei, M. J. Eck, S. C. Harrison

Exp. Method: X-ray Diffraction

Classification: Tyrosine-Protein Kinase

EC Number: 2.7.1.112

Source: Homo sapiens

Primary Citation: Xu, W., Doshi, A., Lei, M., Eck, M. J., Harrison, S. C.: Crystal Structures of C-Src Reveal New Features of its Autoinhibitory Mechanism *Molecular Cell* pp.629 (1999) [[Medline](#)]

Deposition Date: 29-Dec-1998 **Release Date:** 22-Jul-1999

Save

PDB search plugin

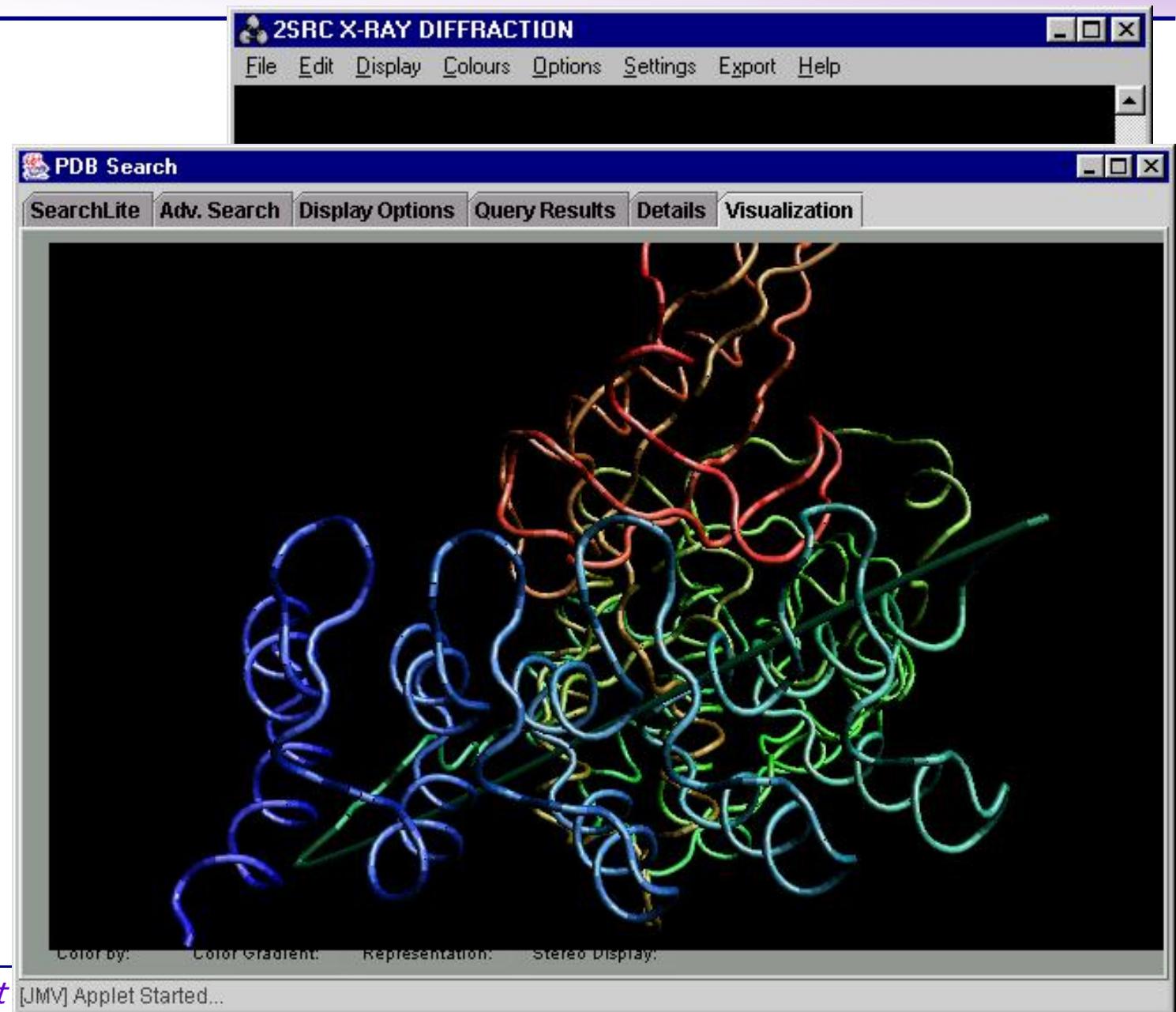
Visualization
with external
packages

Rasmol

JMV

Jmol

User-defined



Sequence search plugin

M. Wroński

Search interface

Nucleotide Search

QueryForm Query Results Details Visualization

Search Nucleotide for: leukemia

Limited To:

All Fields

exclude ESTs exclude STSs
 exclude GSS exclude working draft
 exclude patents
 exclude all of the above

Molecule Gene Location

Segmented Sequences Only from

Modification Date

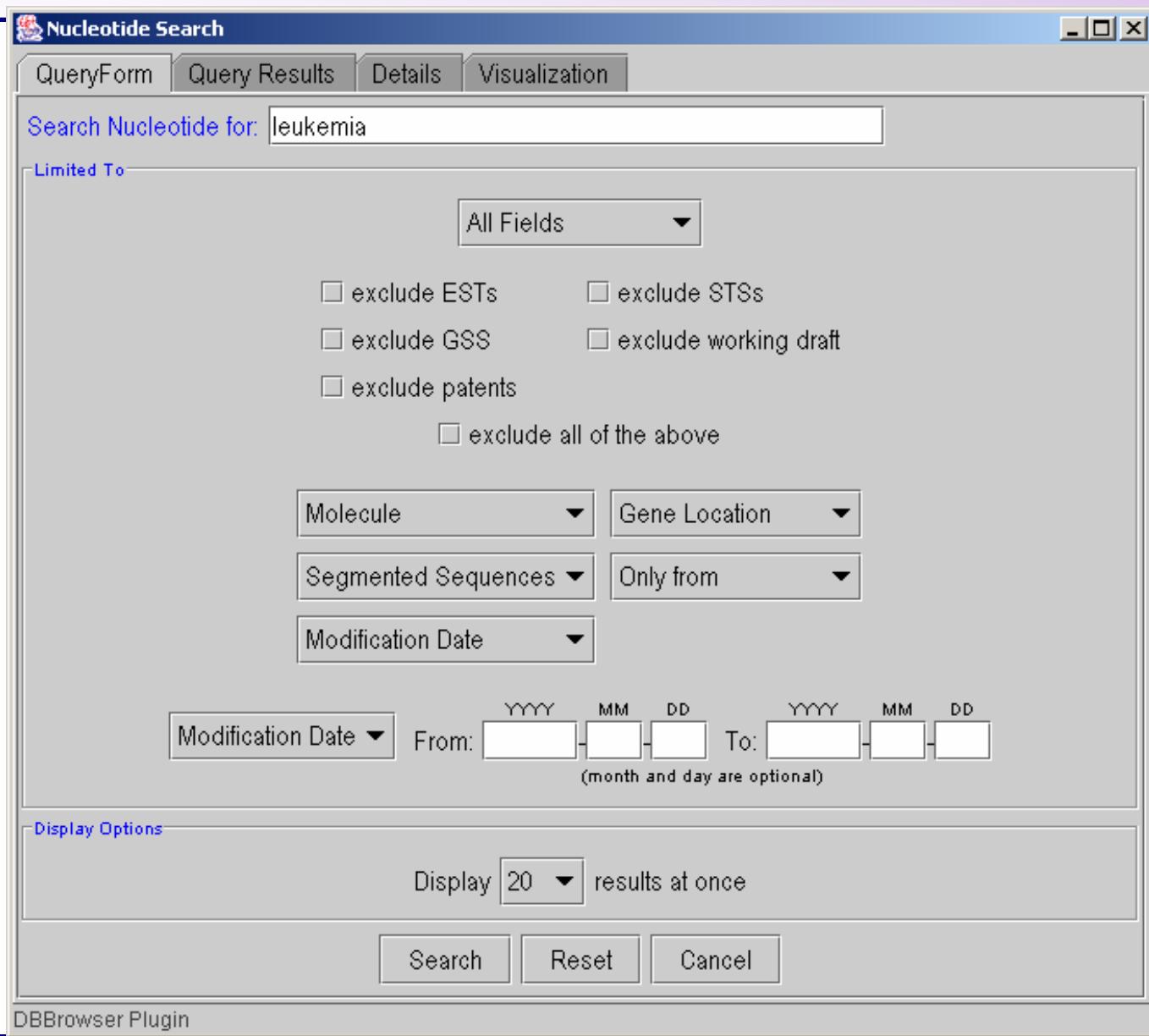
Modification Date From: YYYY MM DD To: YYYY MM DD
(month and day are optional)

Display Options

Display 20 results at once

Search Reset Cancel

DBBrowser Plugin



Sequence search plugin

Detailed view

Nucleotide Search

QueryForm Query Results Details Visualization

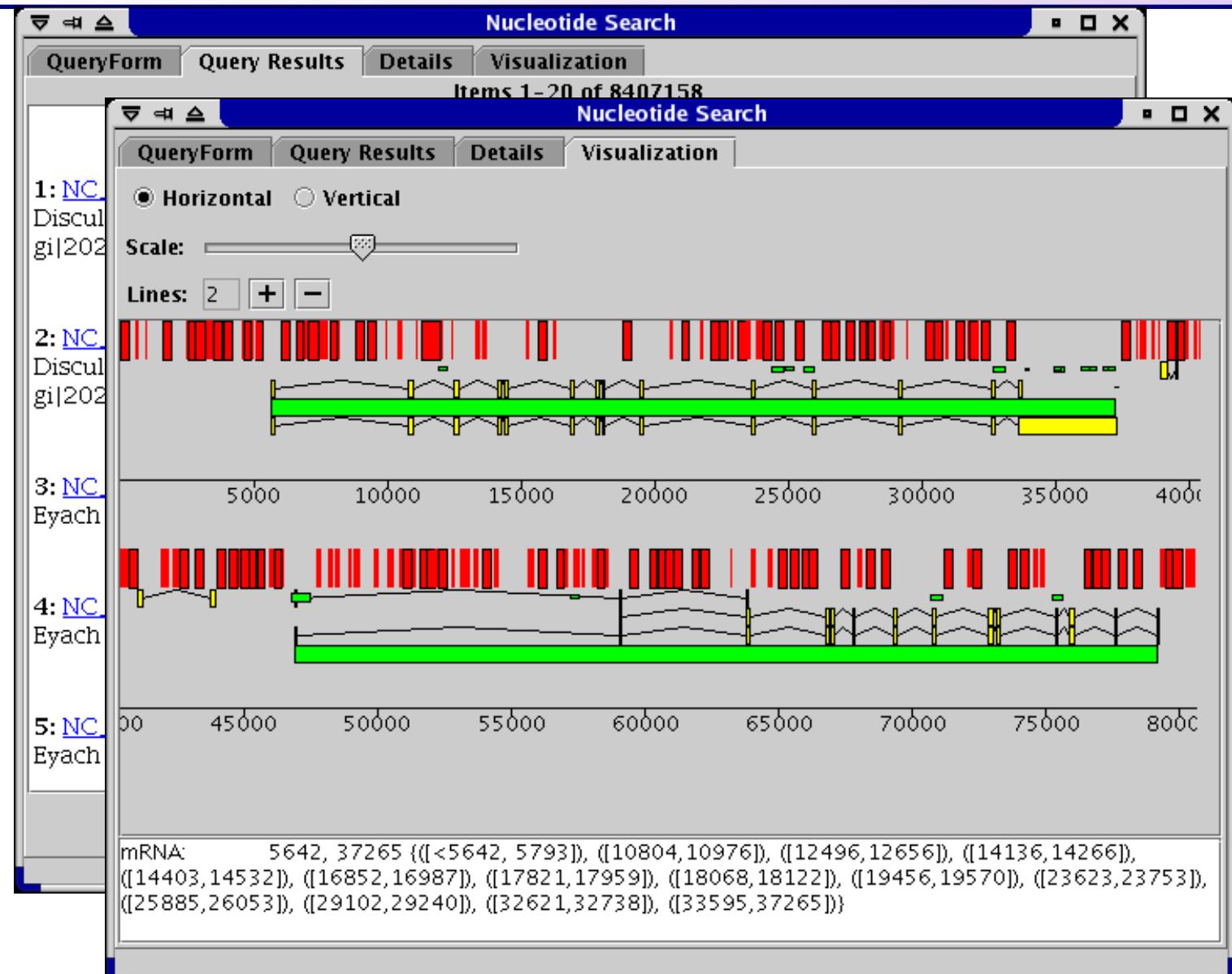
Summary Information

LOCUS	XM_292775	951 bp	mRNA	linear	PRI	17-OCT-2003
DEFINITION	Homo sapiens similar to ABO histo-blood group B transferase (LOC342885), mRNA.					
ACCESSION	XM_292775					
VERSION	XM_292775.2	GI:37552060				
KEYWORDS	.					
SOURCE	Homo sapiens (human)					
ORGANISM	<u>Homo sapiens</u> Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
COMMENT	MODEL <u>REFSEQ</u> : This record is predicted by automated computational analysis. This record is derived from an annotated genomic sequence (<u>NT_011109</u>) using gene prediction method: GNOMON. Also see: Documentation of NCBI's Annotation Process					
On Oct 7, 2003 this sequence version replaced gi: 29741532 .						
FEATURES	Location/Qualifiers					
source	1..951 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606"					

Save

Sequence search plugin

Simple
visualization
based on
BIOJava



DBAccess Plugin

- Access to the SQL databases

- MySQL
- Postgress
- SRB

- Results can be saved in the file

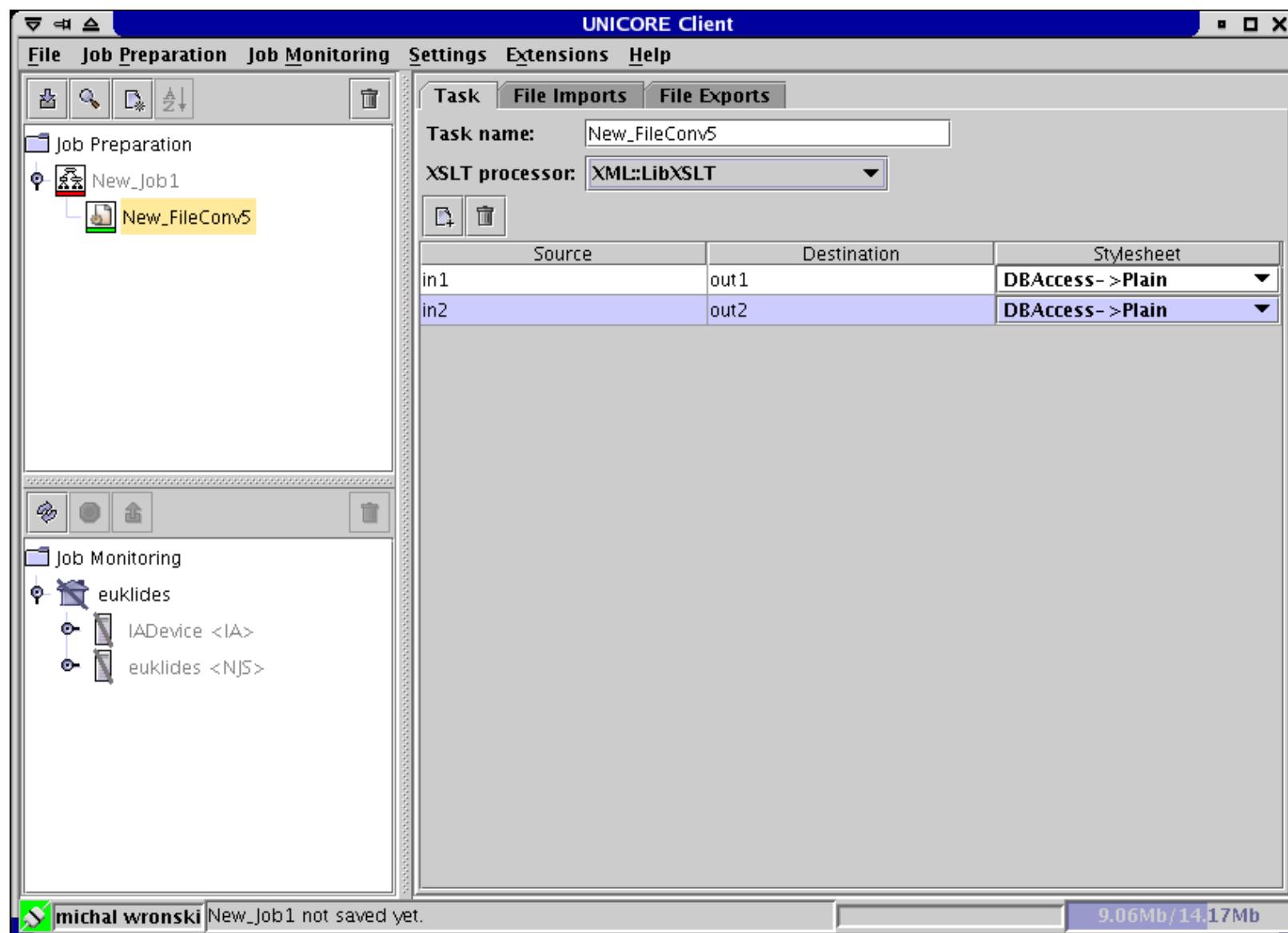
- XML output

The screenshot shows the DBAccess Plugin interface. At the top, there's a 'Task Name' field containing 'DBAccessScript2' and a tab bar with 'Settings', 'Batch Mode', 'File Imports', and 'File Exports'. Below this, under 'SQL Server', 'PostgreSQL' is selected over 'MySQL' and 'Other'. The 'Host' field is set to 'ultra60' and the 'Port' field is empty. In the 'Login' section, 'User' is 'bala', 'Database' is 'test', and 'Password' is empty. In the 'Query' section, the SQL command 'SELECT * FROM tab;' is entered. The bottom part of the window displays the query results in a table titled 'Table 1'. The table has columns 'id', 'name', 'surname', and 'birthdate'. It contains two rows: one for 'Jan Kowalski' (birthdate 1965-06-15) and one for 'Michał Wronski' (birthdate 1980-09-02). Below the table, there's an 'Output Format' dropdown set to 'XML' and a checkbox for 'Put output into file:' which is unchecked.

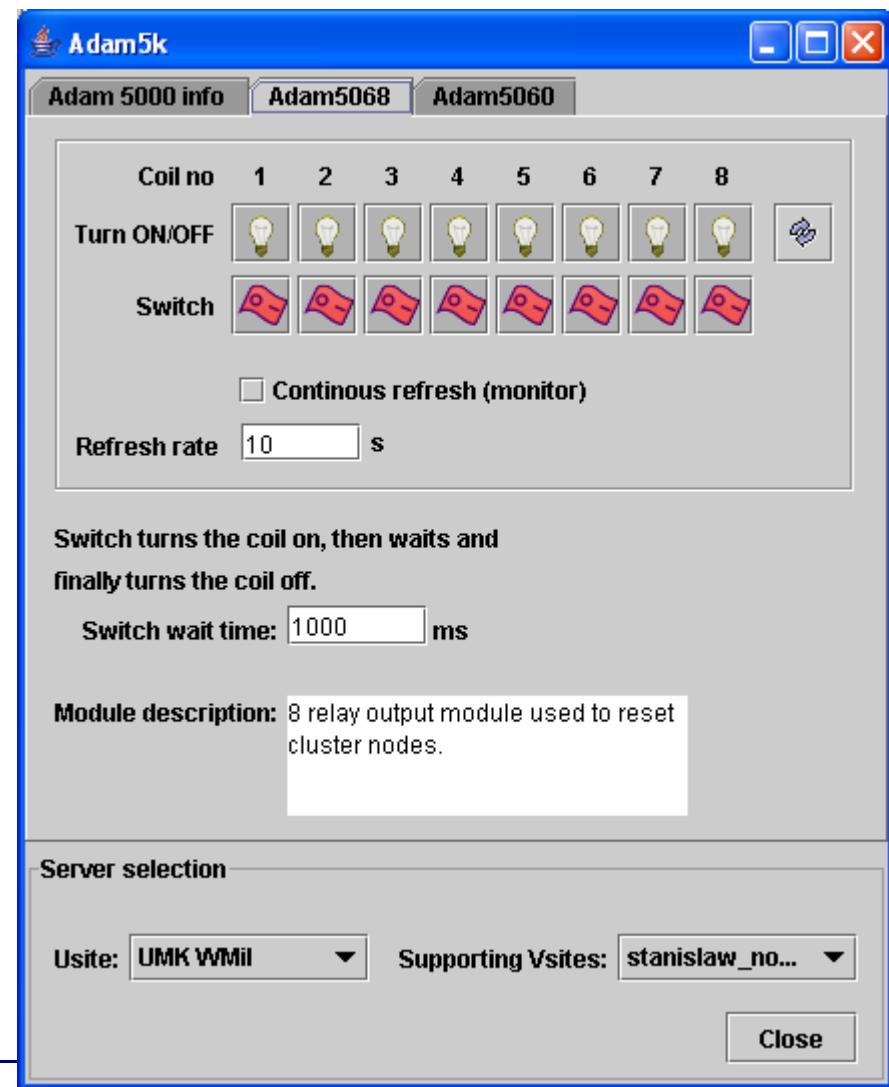
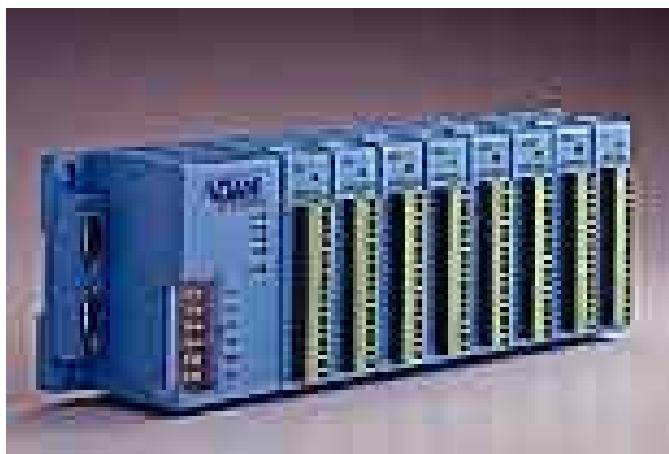
id	name		
2	Jan	Kowalski	1965-06-15
1	Michał	Wronski	1980-09-02

FileConv Plugin

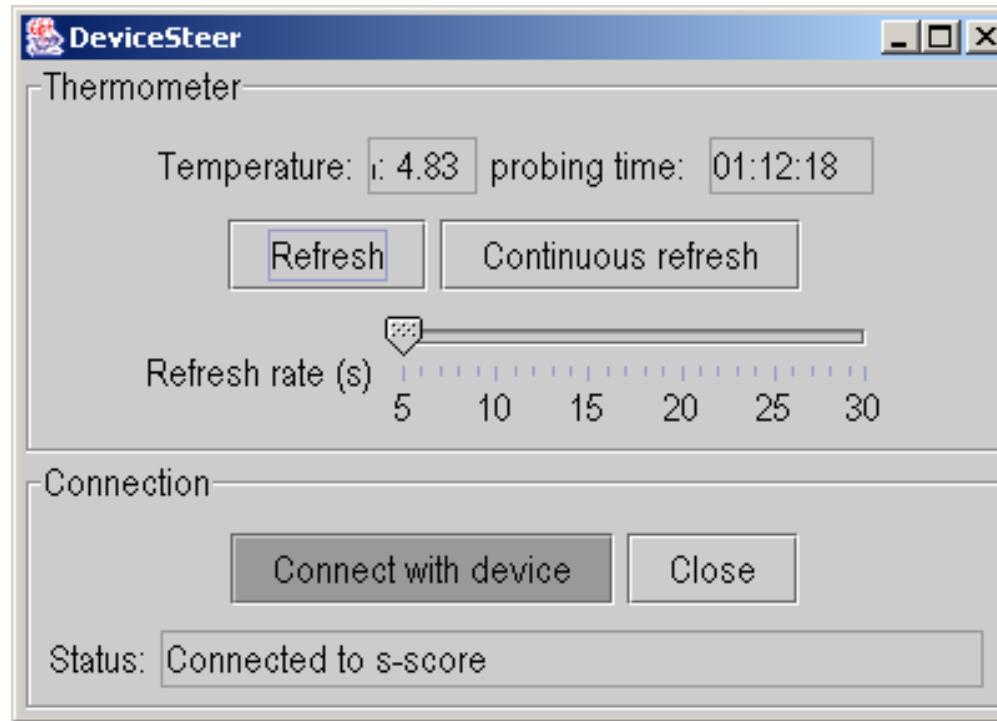
M. Wroński



- Access to the external devices
 - Advantech ADAM5000
 - Daemon for the device running on the target system
 - **Management of the cluster hardware**

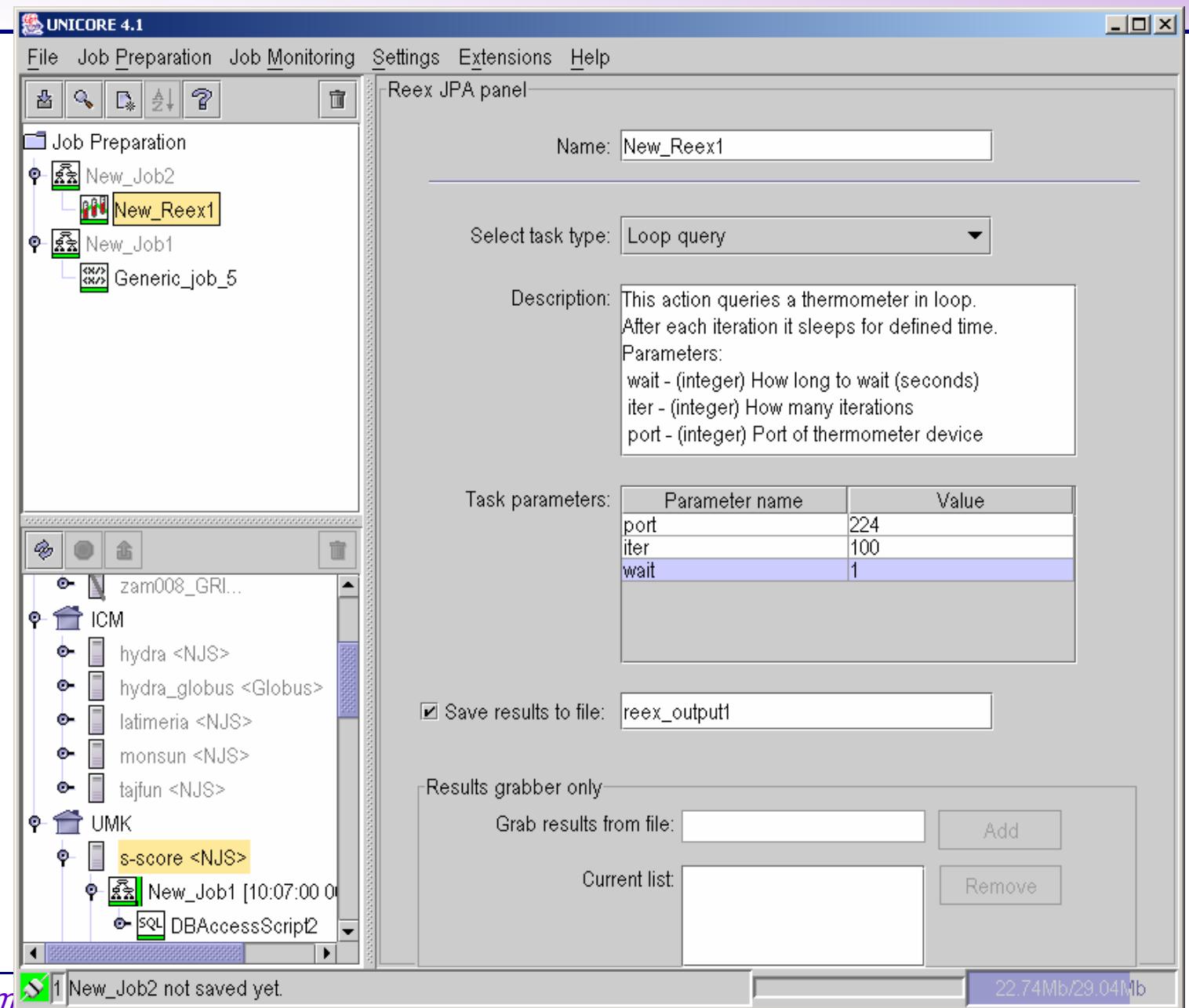


- Access to the external devices
 - Thermometer connected through serial port to TSI server
 - Daemon for the device running on the target system
 - DeviceSteer Plugin



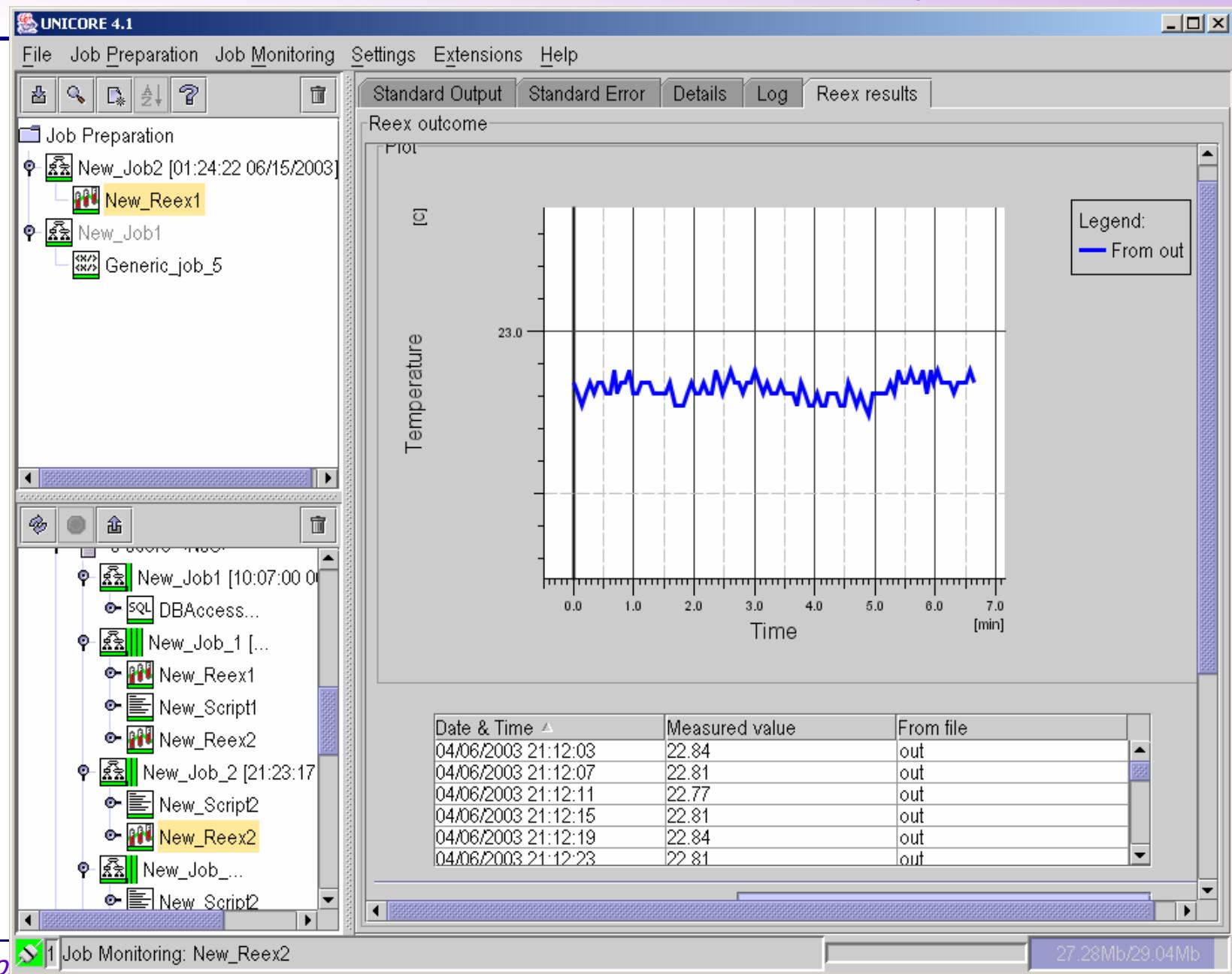
DeviceSteer

K. Benedyczak



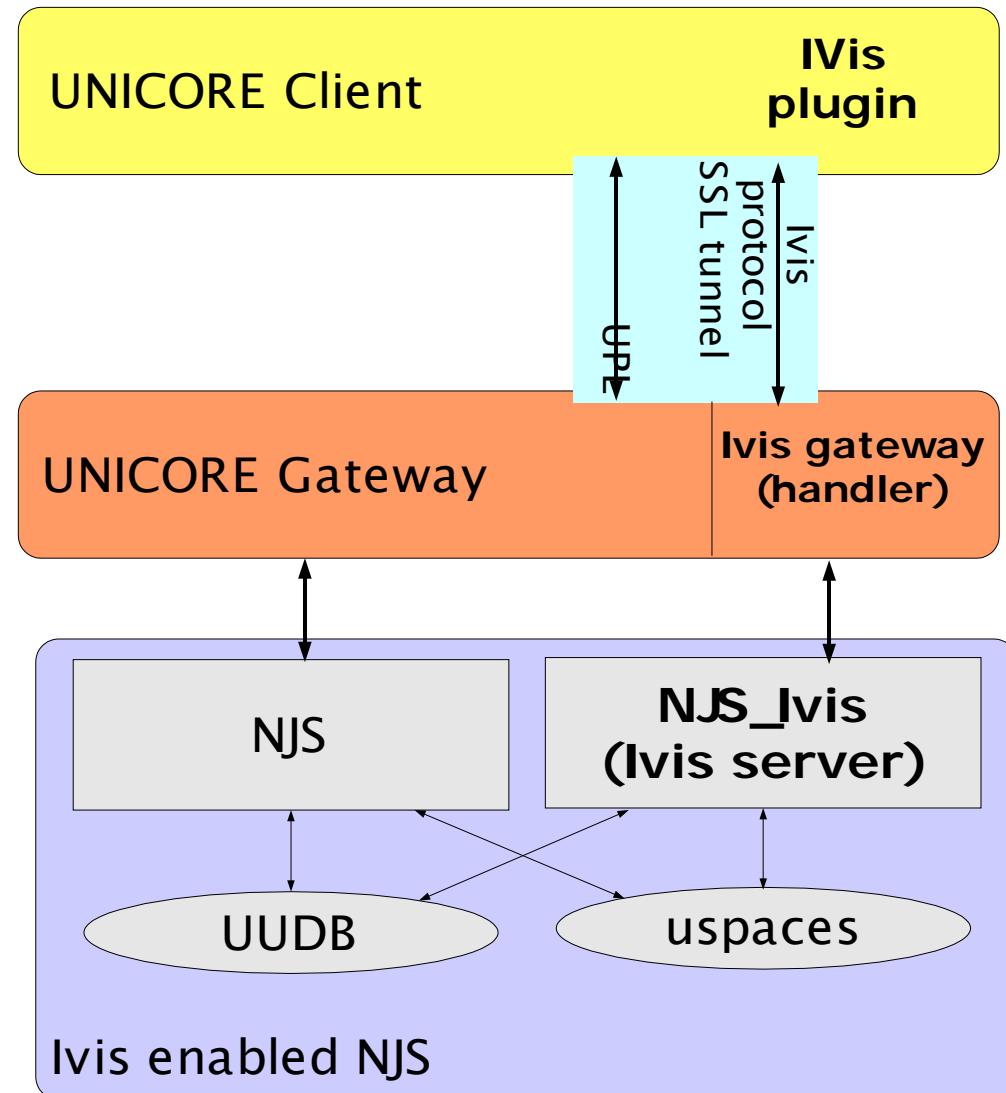
DeviceSteer

K. Benedyczak

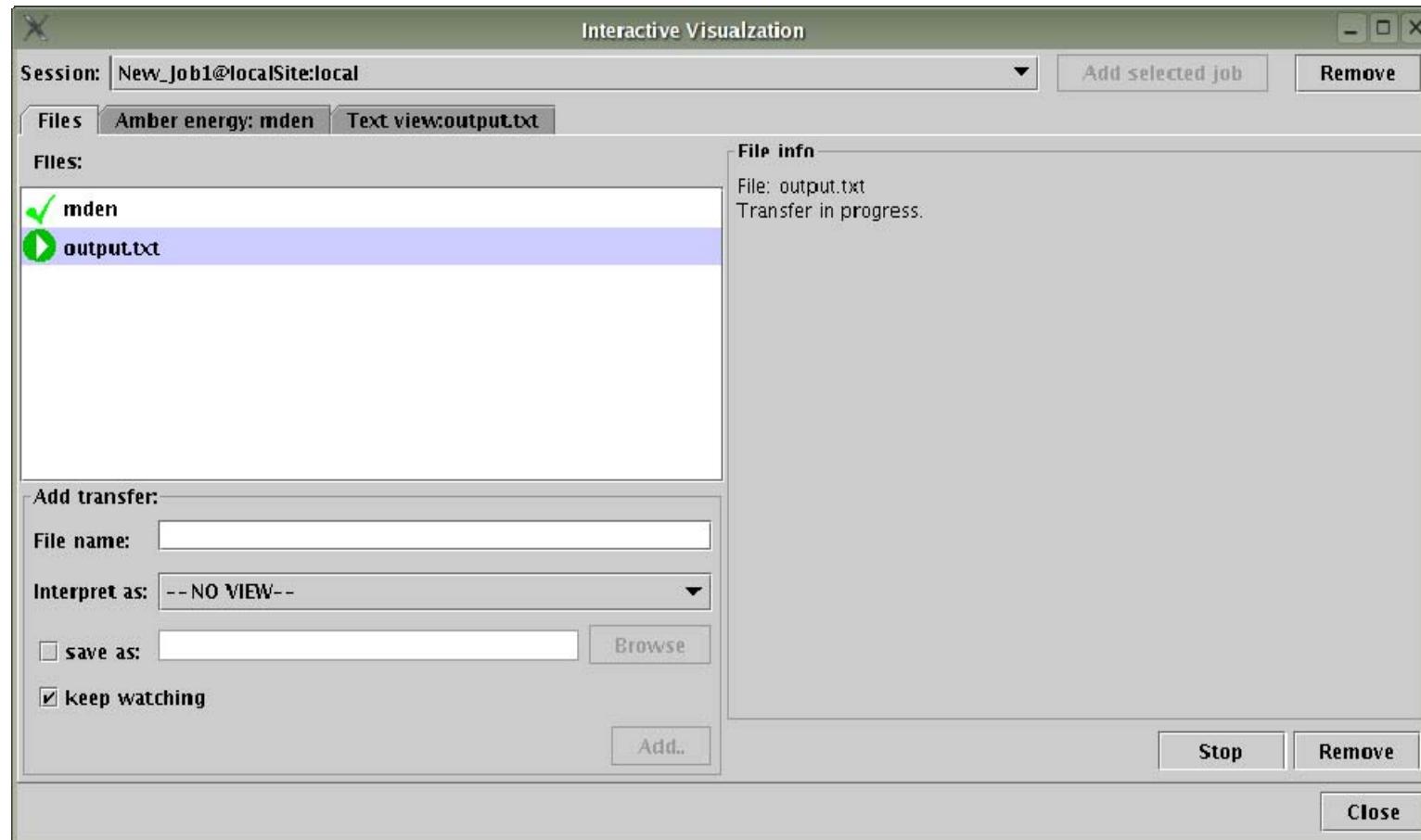


- **IVis concept**

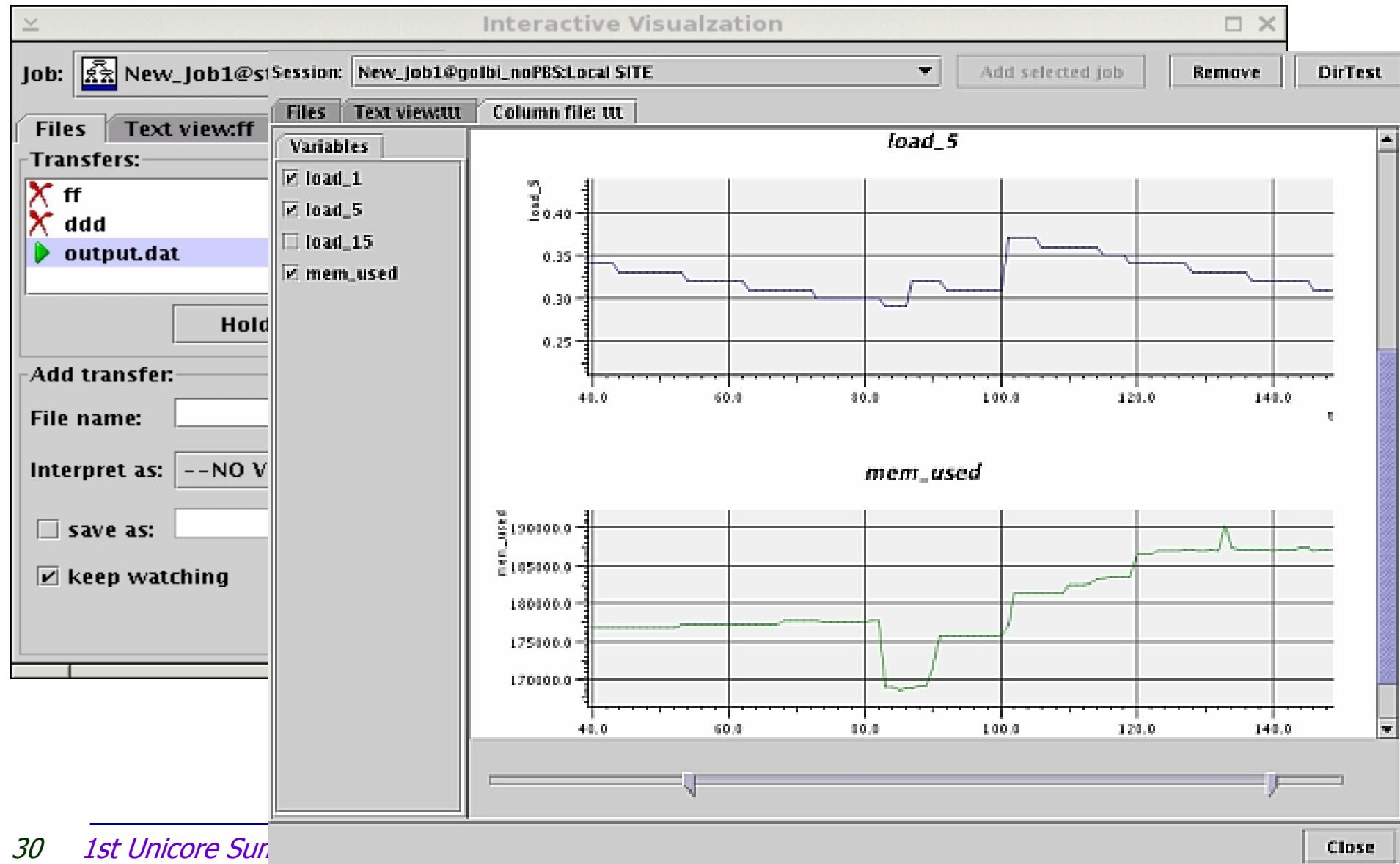
- **data stream**
- **performance**
- **firewalls**
- Stream data through UNICORE Gateway using Gateway plugin
- Simple authorization
- Multiple clients can connect to the single visualization
- IVis is available at SourceForge



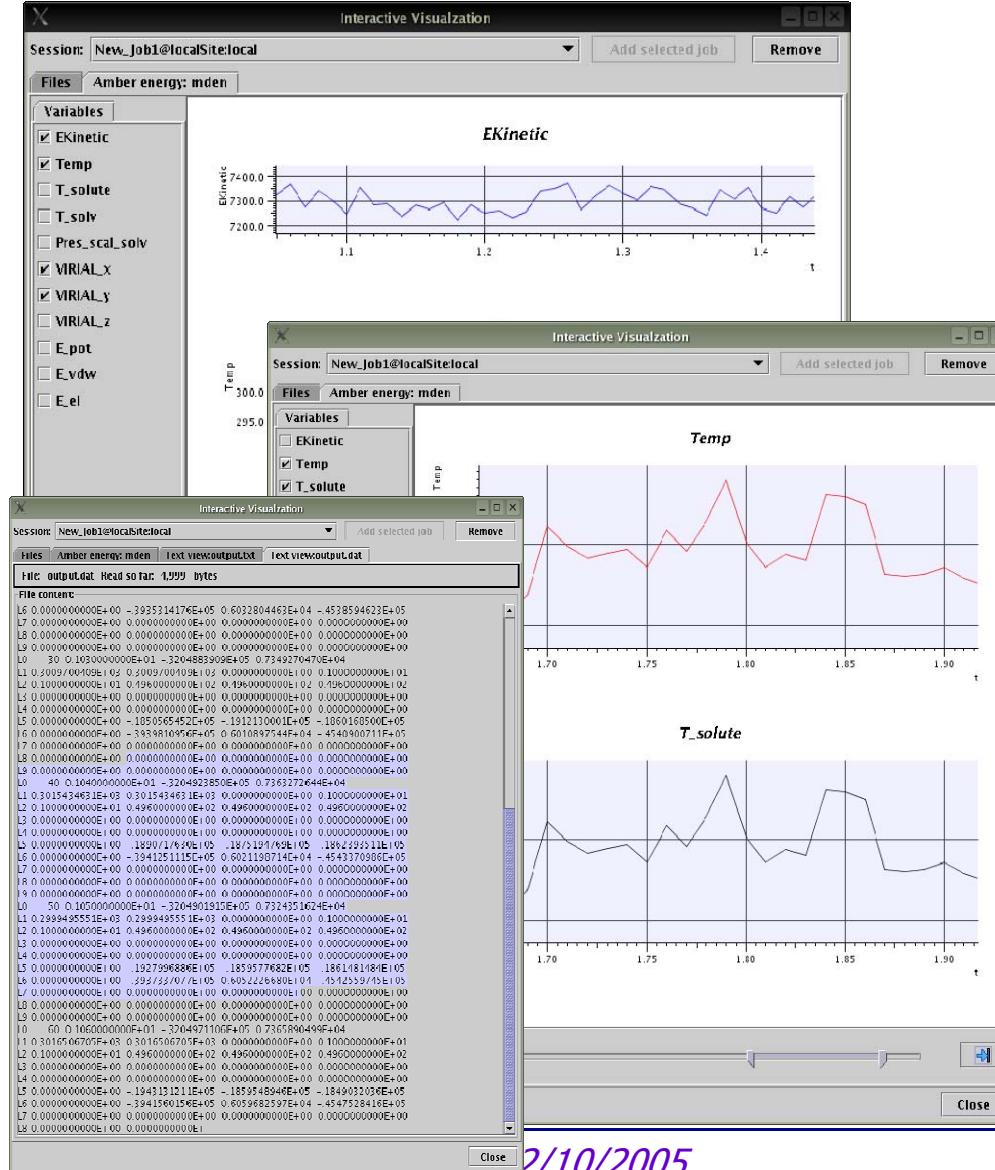
Plugin allows for easy files download from running job directory and visualization (or export to disk)



IVis – build-in visualization



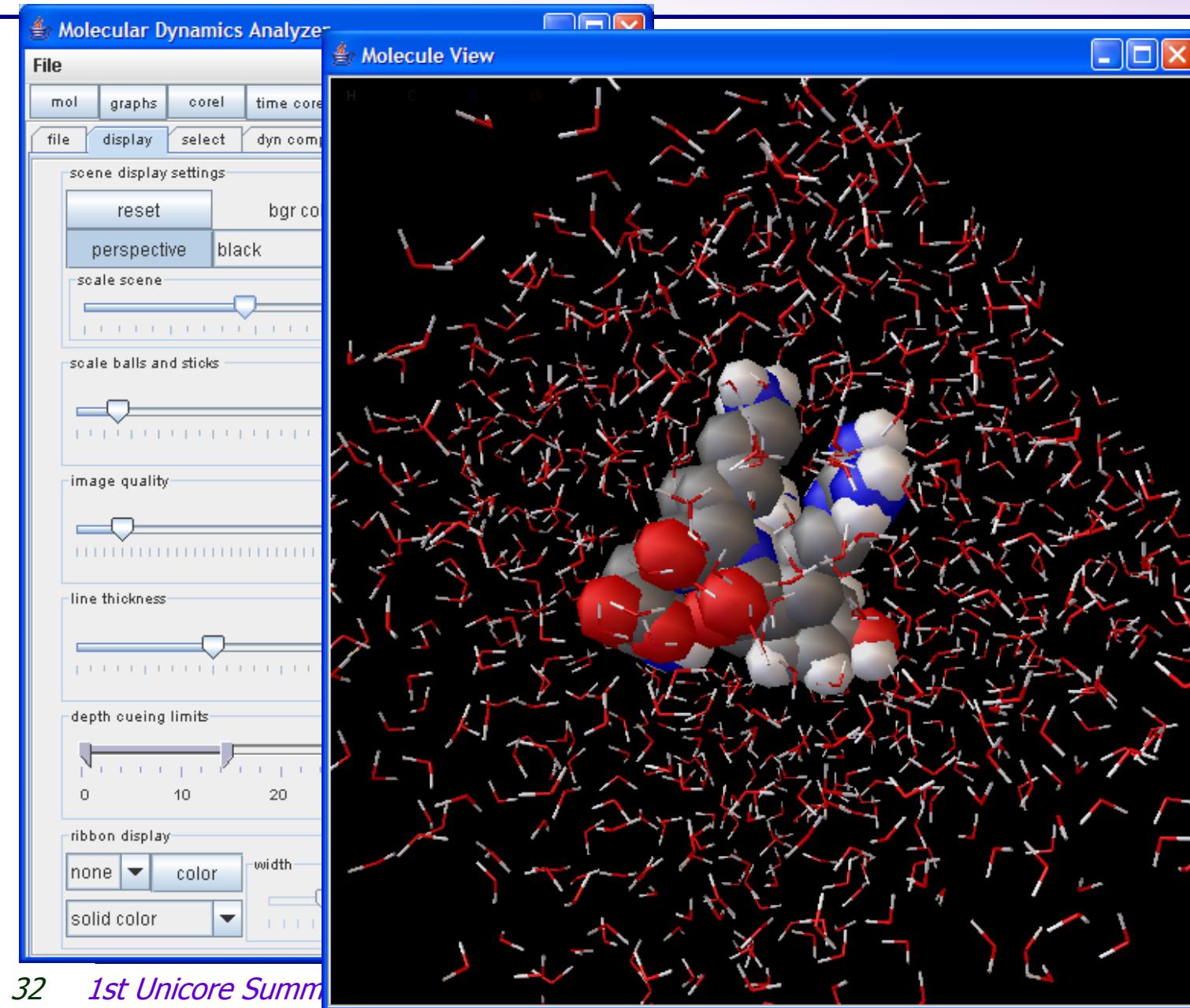
IVis – build-in visualization



- Currently Amber job monitor is provided
- Plain text file preview is available
- Any ASCII data file can be interpreted
- Other visualizations can be easily added, depending on requirements

IVis – build-in visualization

K. Nowiński



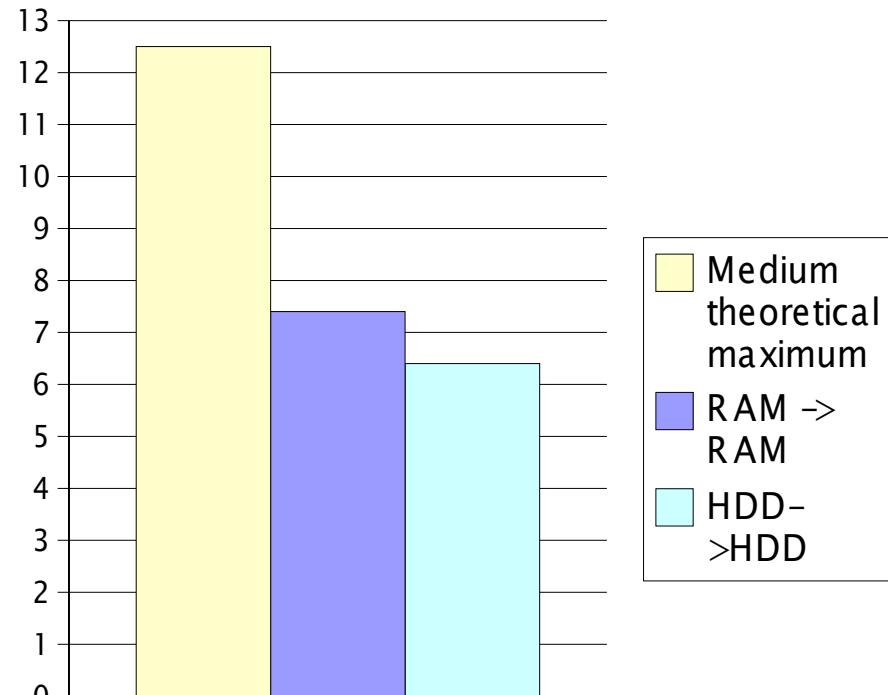
MolDyAna
visualization

Reads
streamed
data

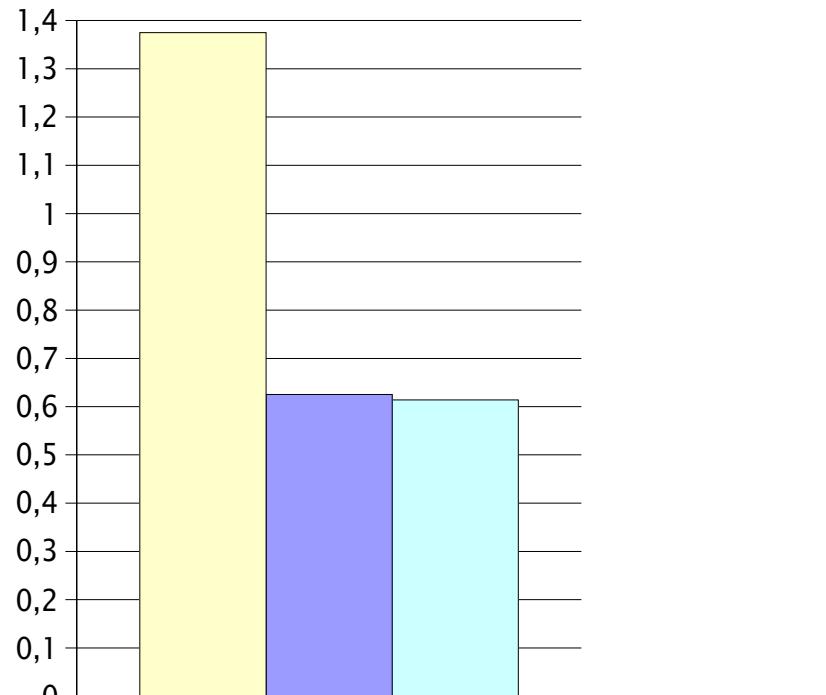
IVis performance

- **IVis tests**
 - Throughput is well acceptable
 - 2D and 3D visualization
 - Works with Amber, Gromos

100Mbit LAN, no routers



11Mbit Wi-Fi + LAN with 3 ro



uPortal 2.4 - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://tarawa.icm.edu.pl:8080/uPortal/tag.590d3917f84e97cd.render.userLayoutRootNode.targe

Fedora Linux Mozilla Firefox Help Mozilla Firefox Discuss... Plug-in FAQ

BrowseSites

Change USite

tarawa biogrid.icm.edu.pl 5533

Warsaw Set as current

AJO testScript1 Set as current

AJO TestScript7 Set as current

AJO testScript3 Set as current

AJO testScript4 Set as current

AJO Test5Script Set as current

AJO testScript2 Set as current

AJO testScript5 Set as current

AJO TestScript6 Set as current

AJO testScript3 Set as current

Torun Set as current

Warsaw

Warsaw

NJS information : NJS start time: Fri Nov 19 10:36:20 CET

Done

TestShellRunner

Wybierz typ skryptu: bash

Podaj nazwe skryptu: MyScript8

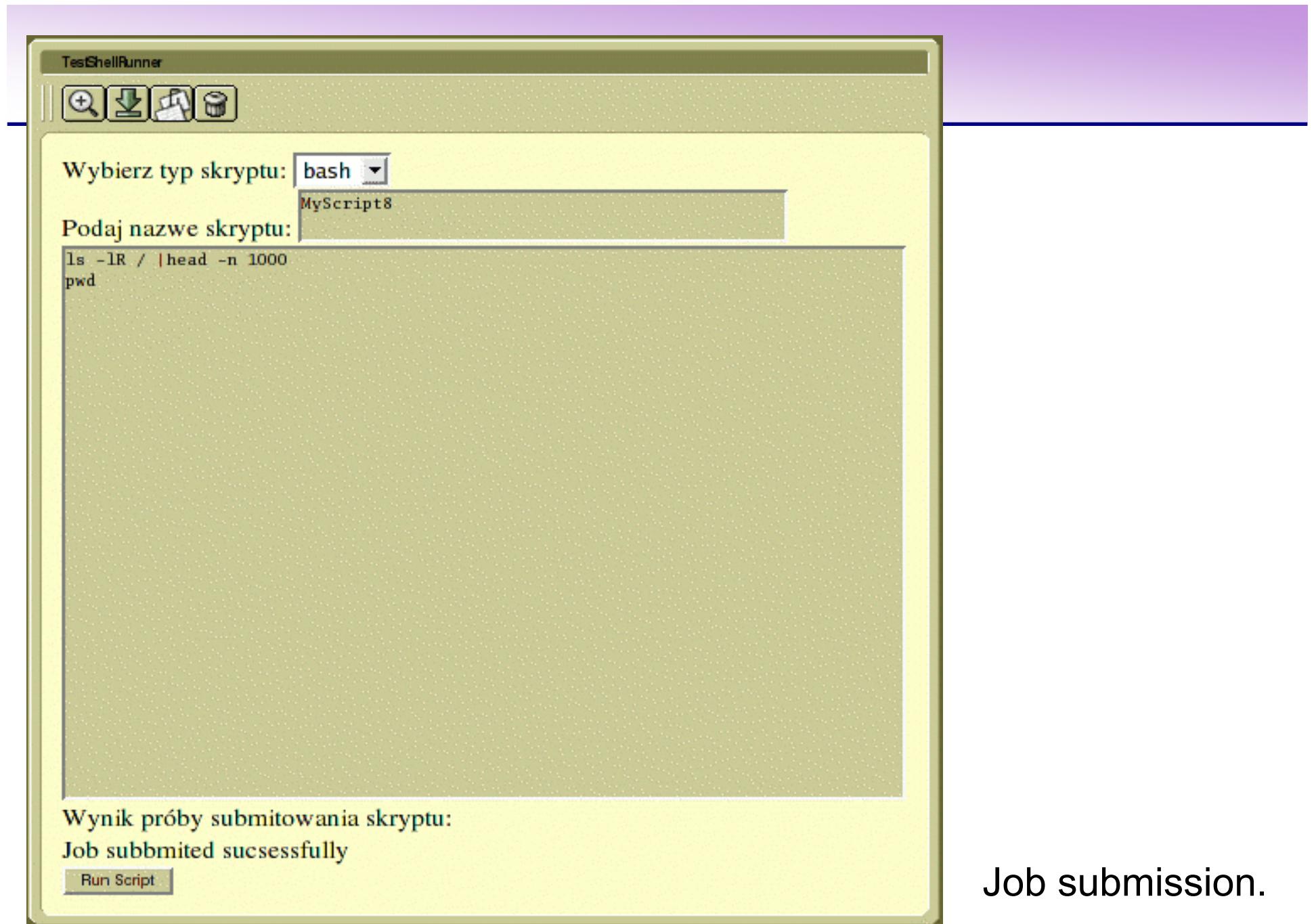
ls -lR / | head -n 1000
pwd

Run Script

GetOutcome

1st ORIONC Summit 11-12/10/2005

11. Dział (ICN/ONIK)



Job submission.

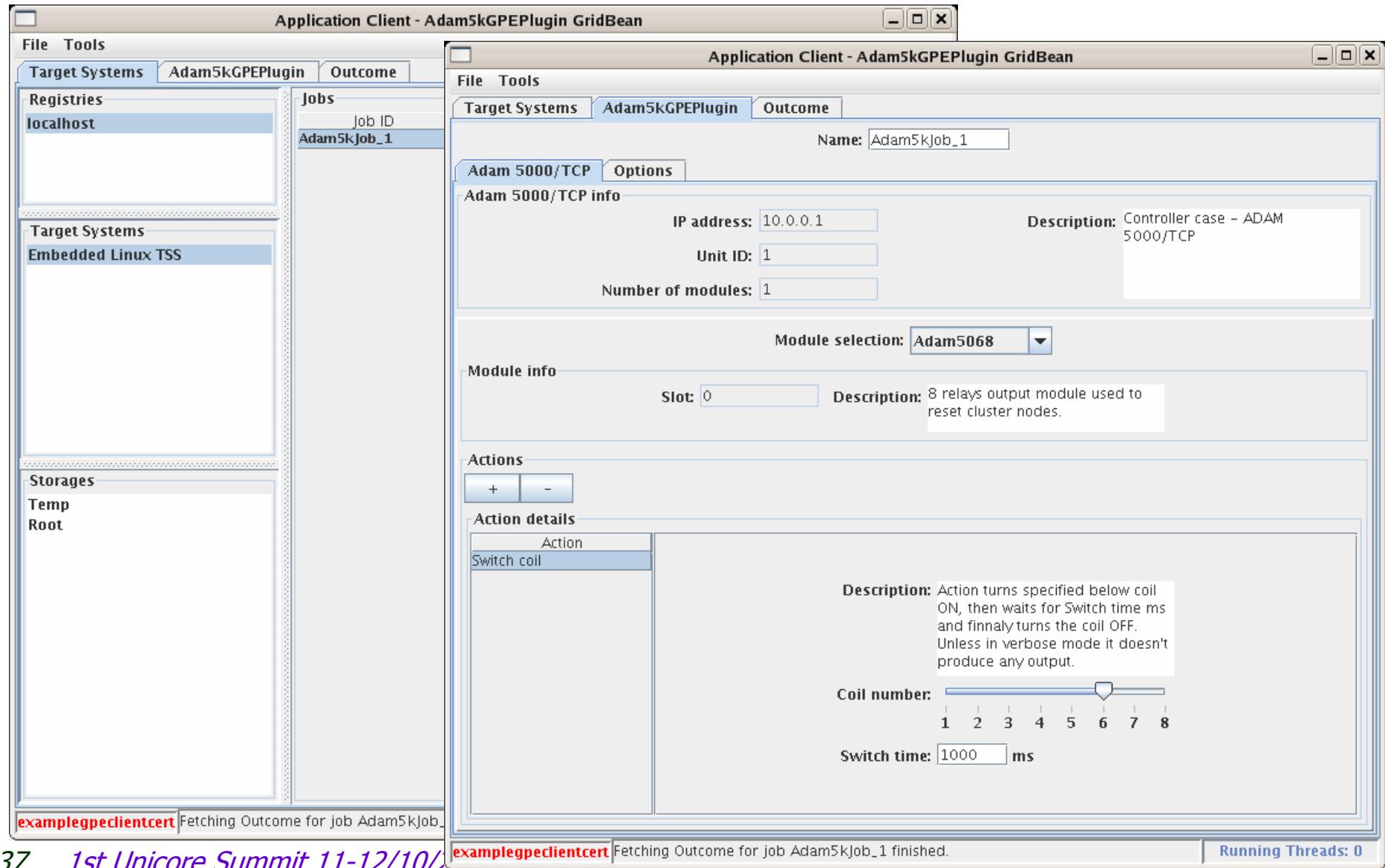
GetOutcome

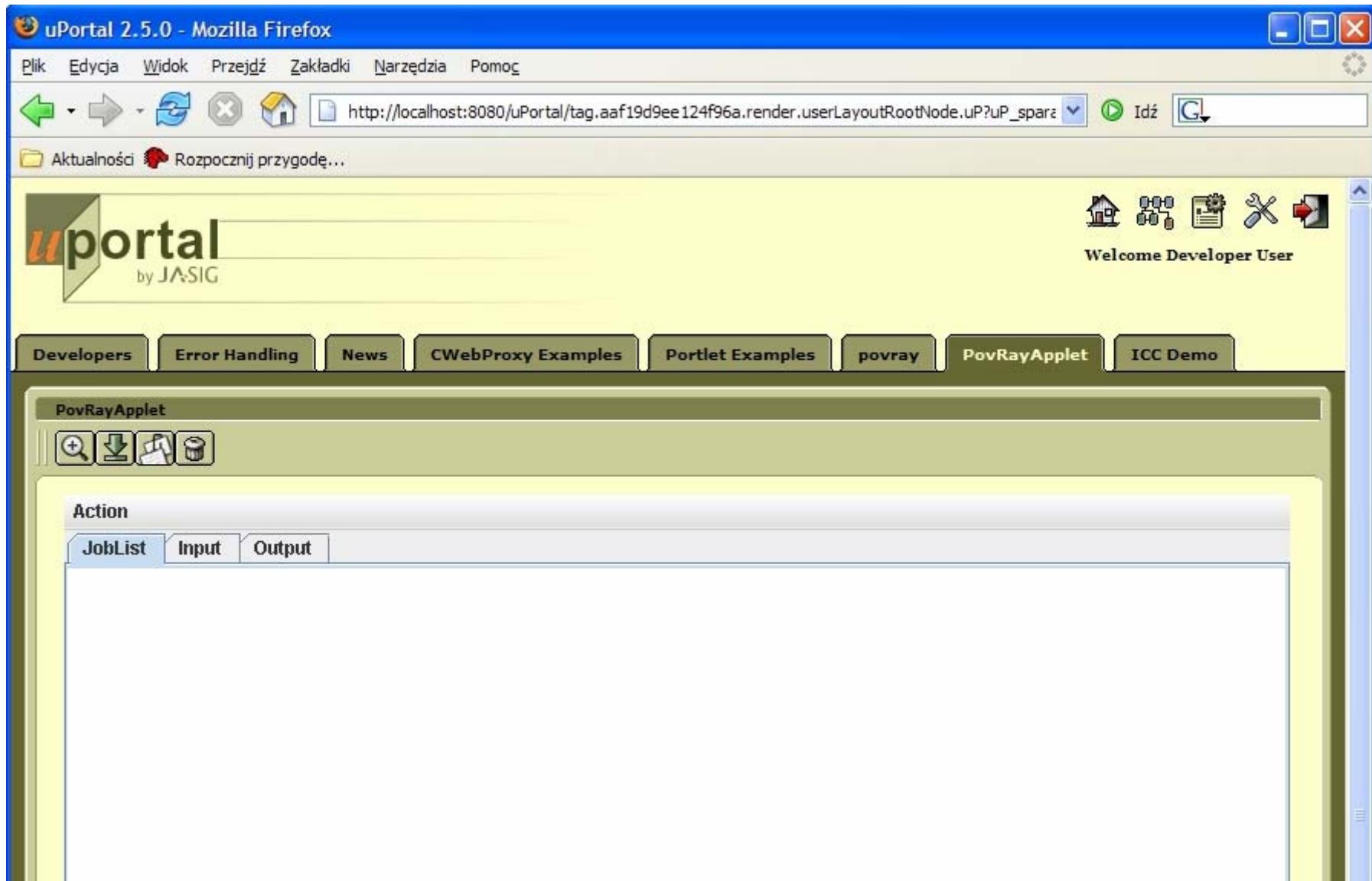
For the job: MyScript8 on site: Warsaw get:

stdout	stderr	log
<pre>This script was created and executed by UNICORE - start of user output /: total 208 drwxr-xr-x 2 root root 40 drwxr-xr-x 3 root root 40 lrwxrwxrwx 1 root root 40 drwxr-xr-x 22 root root 118 drwxr-xr-x 97 root root 8 drwxr-xr-x 7 root root 40 drwxr-xr-x 2 root root 40 drwxr-xr-x 10 root root 40 drwxr----- 2 root root 16 drwxr-xr-x 2 root root 40 drwxr-xr-x 4 root root 40 drwxr-xr-x 3 root root 40 dr-xr-xr-x 60 root root 40 drwxr-x--- 19 root root 40 drwxr-xr-x 2 root root 8 drwxr-xr-x 3 root root 40 drwxrwxrwt 11 root root 40 drwxr-xr-x 2 root root 40 drwxr-xr-x 30 root root 40 drwxr-xr-x 26 root root 40</pre>		<pre>15:44:15 09/12 S Initial user incarnation is <unicore_demo/> Y 15:44:15 09/12 C Arrived at Warsaw 15:44:15 09/12 S Status is now READY 15:44:15 09/12 S Status is now EXECUTING. Message: Start processing of AJO 15:44:15 09/12 T Starting to execute 15:44:15 09/12 T Incarnating a Uspace 15:44:15 09/12 T AJO and Uspace initialised. Starting execution 15:44:15 09/12 T Starting child action <MyScript8_ACTIONGROUP_1102603455028_17/596756d> 15:45:19 09/12 T Child action done <MyScript8_ACTIONGROUP_1102603455028_17/596756d> with status <SUCCESSFUL> 15:45:19 09/12 T AJO execution finished, now clearing Uspace incarnation 15:45:19 09/12 S Status is now SUCCESSFUL 15:45:19 09/12 T Uspace cleared. Finishing 15:48:29 09/12 S RetrieveOutcome request from <EMAILADDRESS=kura@icm.edu.pl, CN=kura, OU=ICM, O=UW, L=Wa, ST=MA, C=PL> 15:48:29 09/12 T Complete Outcome (including incarnated data) will be sent to client 15:50:02 09/12 S RetrieveOutcome request from <EMAILADDRESS=kura@icm.edu.pl, CN=kura, OU=ICM, O=UW, L=Wa, ST=MA, C=PL> 15:50:02 09/12 T Complete Outcome (including incarnated data) will be sent to client 15:57:05 09/12 S RetrieveOutcome request from <EMAILADDRESS=kura@icm.edu.pl, CN=kura, OU=ICM, O=UW, L=Wa, ST=MA, C=PL> 15:57:05 09/12 T Complete Outcome (including incarnated data) will be sent to client</pre>

Adam2k GridBean

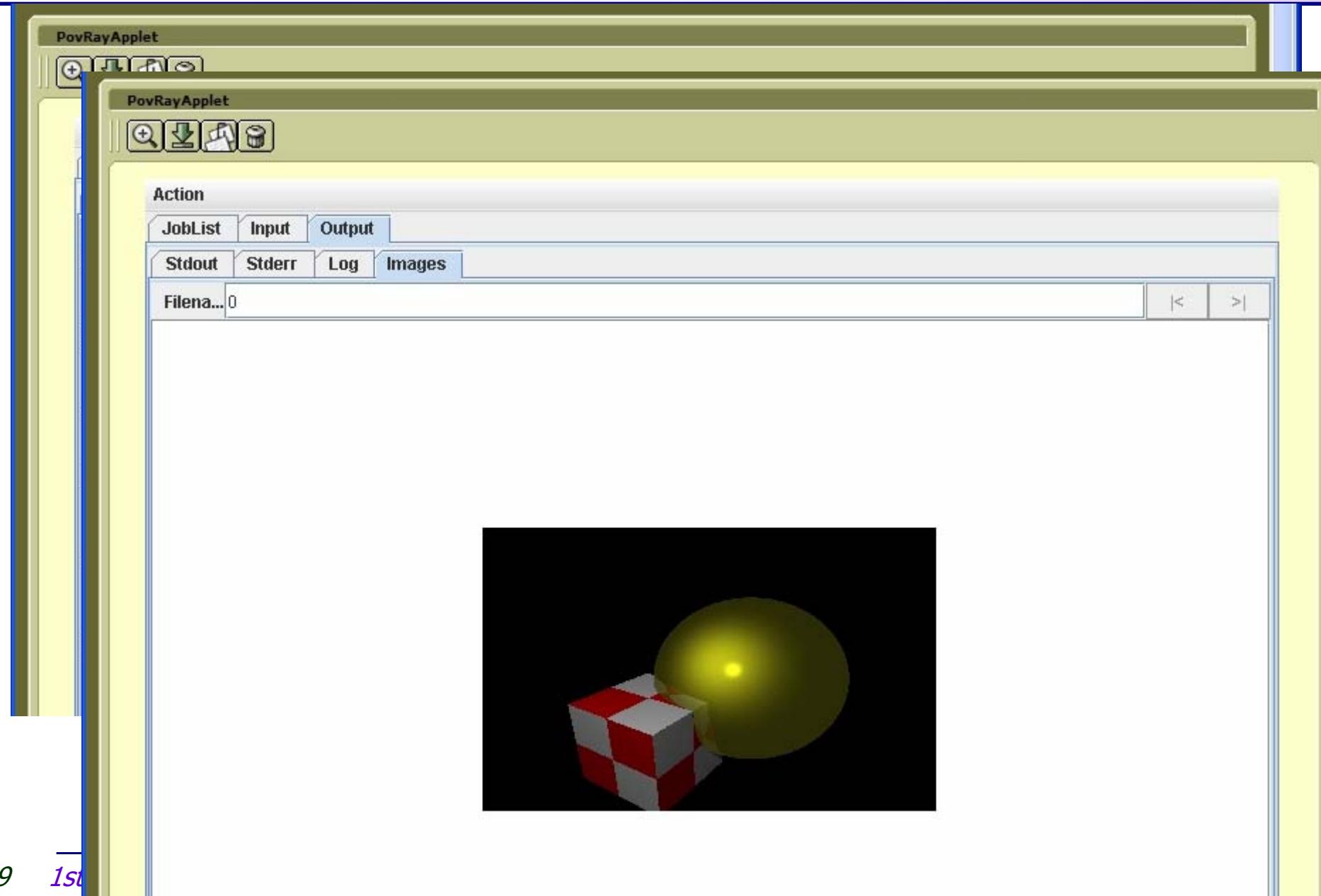
K. Benedyczak

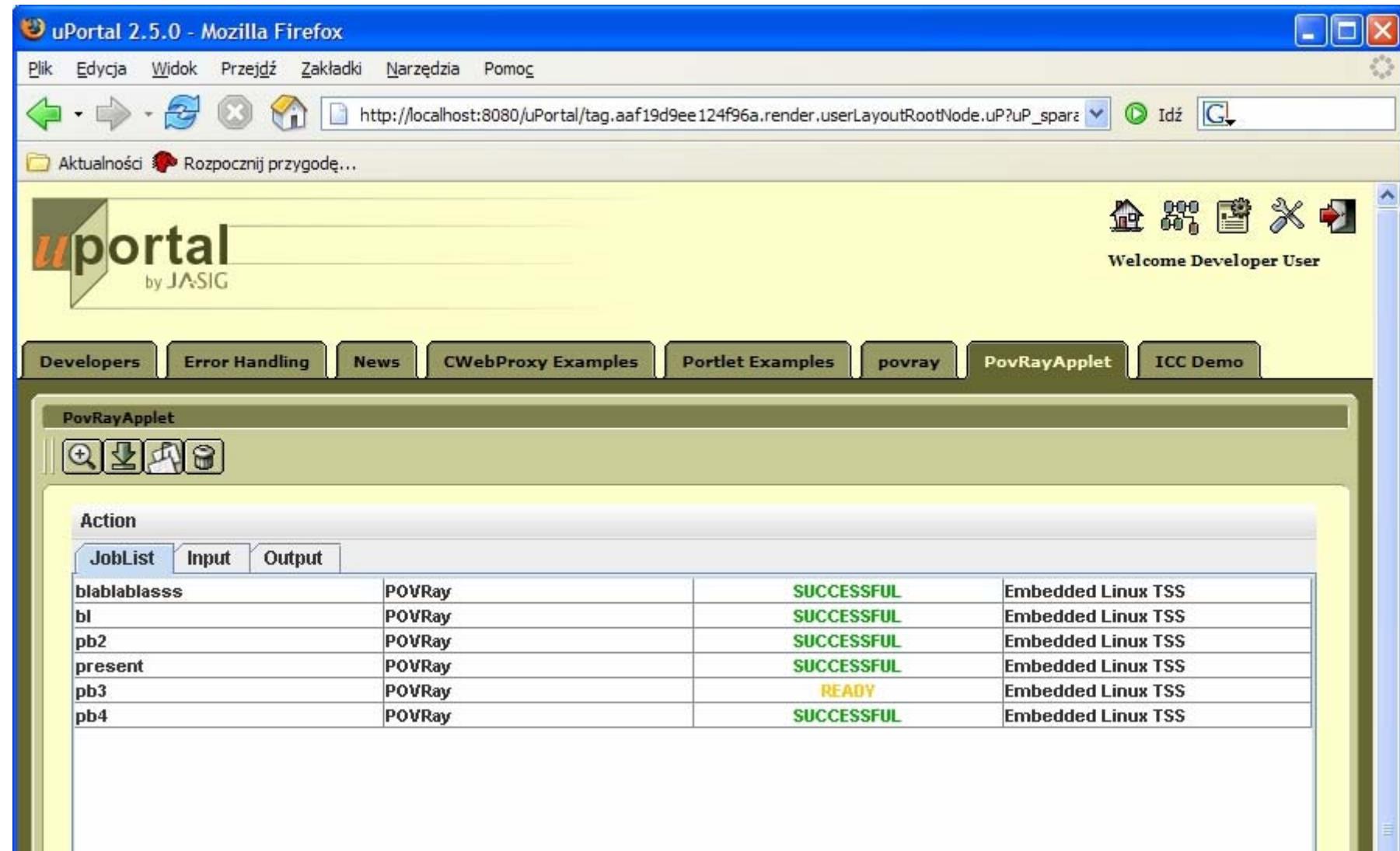




GPE AppletClient

J. Jurkiewicz





Lesson learned

- Plugin is a very good concept
- However....
 - development must be easier
 - configure rather than program
 - Amber, GenericPlugin
 - plugins too much depend on the UNICORE Client
 - UNICORE is a moving target
 - plugins are difficult to maintain
 - need for modularity and internal workflows
 - MetaPlugin
 - applications are changing

Plugin development Team

Piotr Bała

bala@icm.edu.pl

ICM Warsaw University

Faculty of Mathematics and Computer Science
N. Copernicus University, Toruń



- **Krzysztof Benedyczak**
- **Łukasz Korzybski**
- **Jarosław Pytliński**
- **Łukasz Skorwider**
- **Michał Wroński**
- **dr Krzysztof Nowiński**
- **Aleksander Nowiński**
- **Jakub Jurkiewicz**
- **Jarosław Wypychowski**