



Enabling Grids for E-science

## ***GENIUS Grid Portal for e-Science community: a new experience for managing workflow in a SOA***

**Livia Torterolo** <[livia@nice-italy.com](mailto:livia@nice-italy.com)>

**NICE srl**

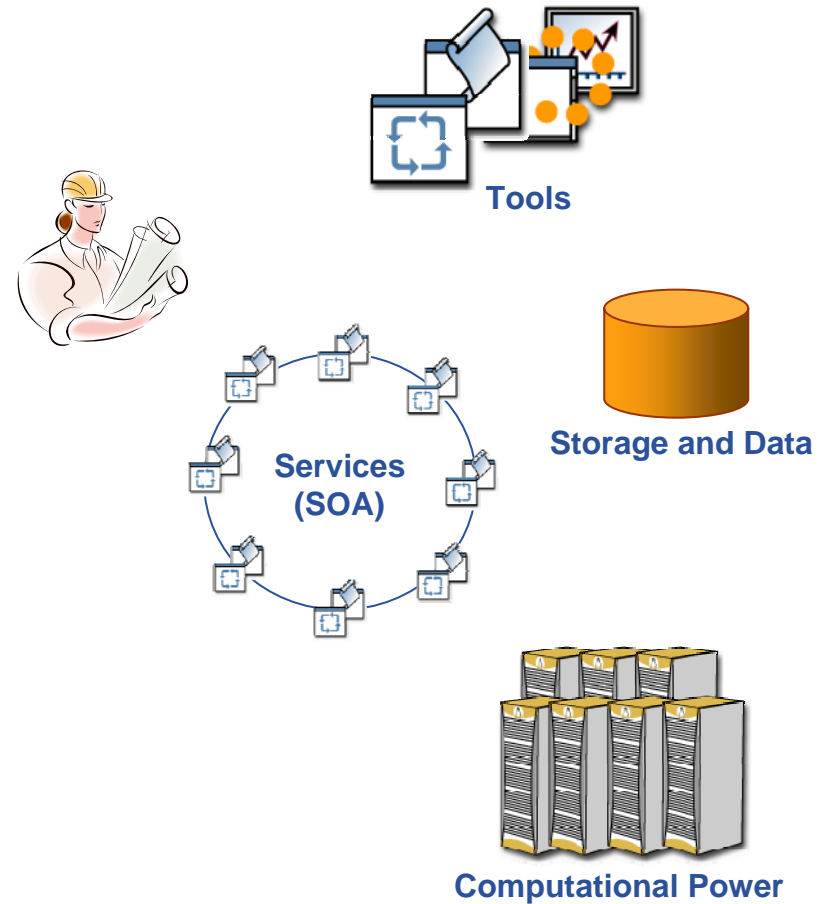
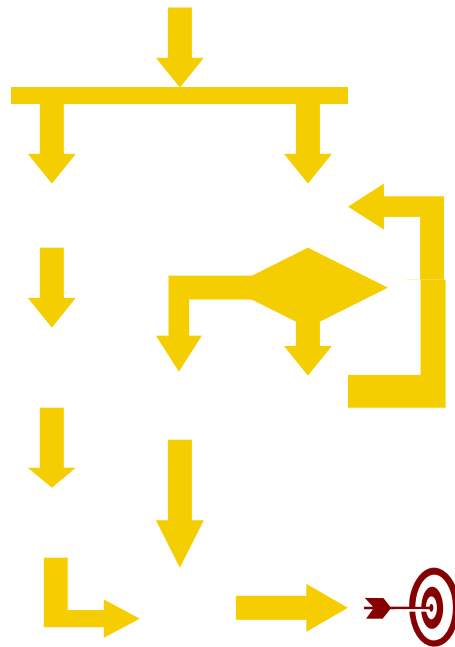
**EGEE'06**

**September 25-29, 2006, Geneva, Switzerland**

[www.eu-egee.org](http://www.eu-egee.org)



- Scientists have to execute complex tasks
- There is the need to orchestrate these services in workflows



## Slide 2

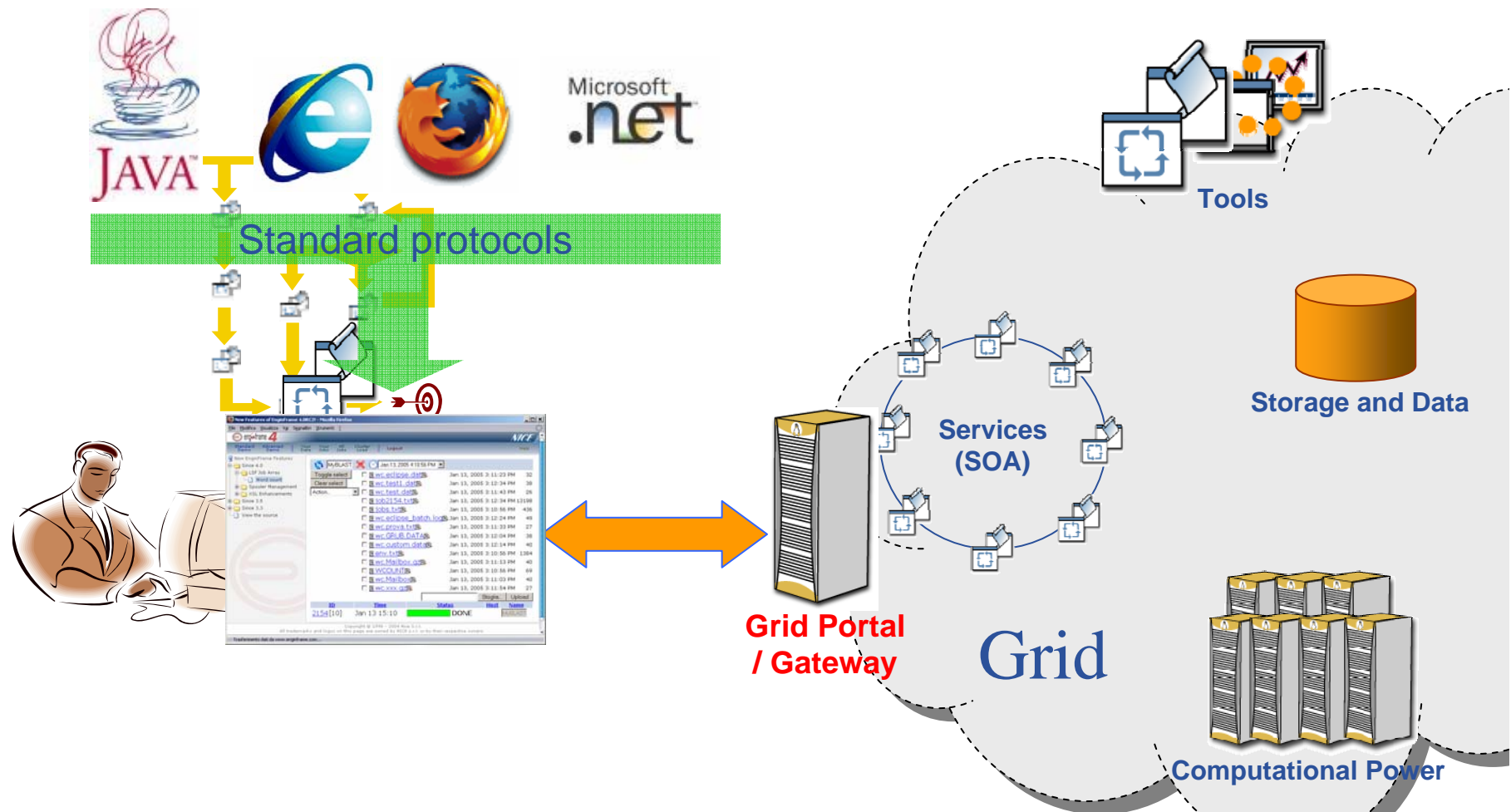
---

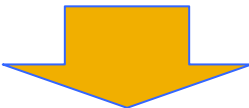
**M1**

Too much written, it would be great to have PICTURES describing such statements

Mauri, 7/7/2006

- Workflows based on Grid services
- Providing optimisation, scalability, reliability, fault tolerance, ...



- Goal: proposal of a **Grid-based solution** able to provide the basic building blocks for composition and enactment of workflows for e-Science community
  - Tools:
    - EnginFrame or Genius:
      - As Grid services provider and Web interface for managing workflow
    - Taverna:
      - As Workflow designer
    - Moteur:
      - As Workflows enactor
- 
- **Users can**:
    - build workflows on top of exposed Grid services
    - run and monitor workflows via a standard Web browser
    - exploit in a transparent way the computational power and data access capabilities provided by the backend Grid infrastructure...

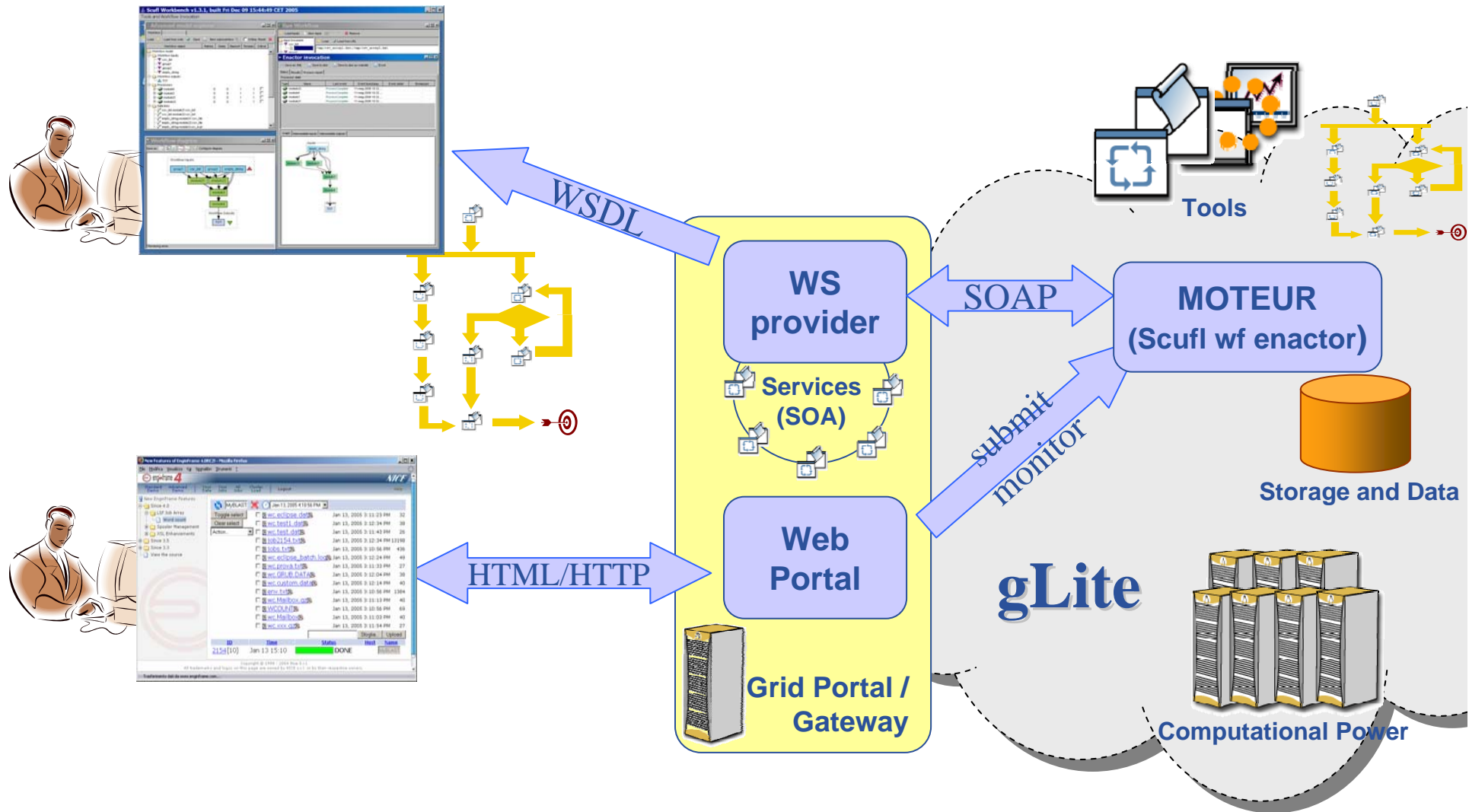
## Slide 4

---

M2

A picture also for representing the impact of our proposak

Mauri, 7/7/2006



The service interface implemented provides the following functionalities:

- *Upload of Scufi workflow and related inputs*
- *Submission by Moteur*
- *Monitoring of workflow*
- *Results visualization*



- **Implementation of a bioinformatics application workflow exploiting the proposed architecture**
- **Based on DChip, one of the most complete and diffuse free software for the microarray data analysis**
- **Composed of different modules:**
  1. data set opening and normalization
  2. model based gene expression
  3. extraction of differentially expressed genes
  4. clustering

Scufl Workbench v1.3.1, built Fri Dec 09 15:44:49 CET 2005

Tools and Workflow Invocation

### Advanced model explorer

Workflow: Object properties

Load Load from web Save New subworkflow Offline Reset

Workflow object	Retries	Delay	Backoff	Threads	Critical
Workflow model					
Workflow inputs					
csv_list					
group1					
group2					
empty_string					
Workflow outputs					
TOT					
Processors					
module4	0	0	1	1	
module3	0	0	1	1	
module21	0	0	1	1	
module22	0	0	1	1	
Data links					
csv_list-module21:csv_list					
csv_list-module22:csv_list					
empty_string-module21:csv_file					
empty_string-module22:csv_file					
empty_string-module3:csv_expr					

### Run Workflow

Load Inputs New Input New List Remove

Input Document Load Load from URL

csv\_list /tmp/csv\_array1.dat:/tmp/csv\_array2.dat

around

### Enactor invocation

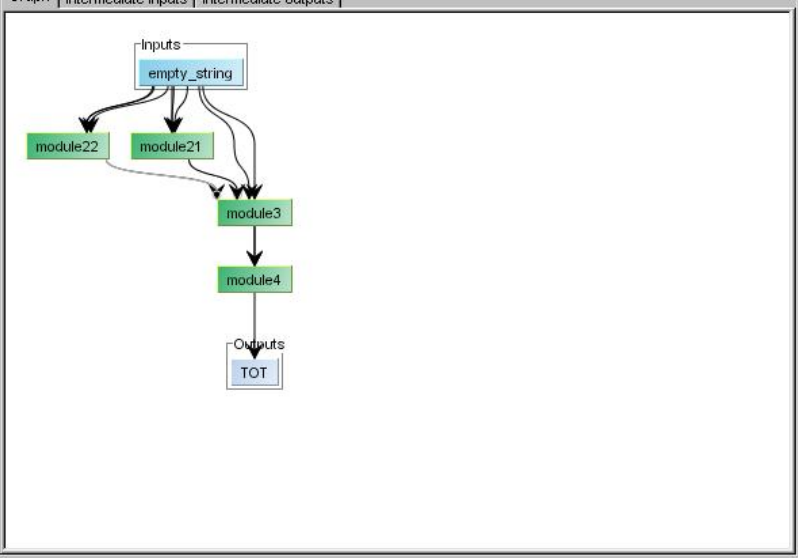
<> Save as XML Save to disk Save to disk as website Excel

Status Results Process report

Processor status

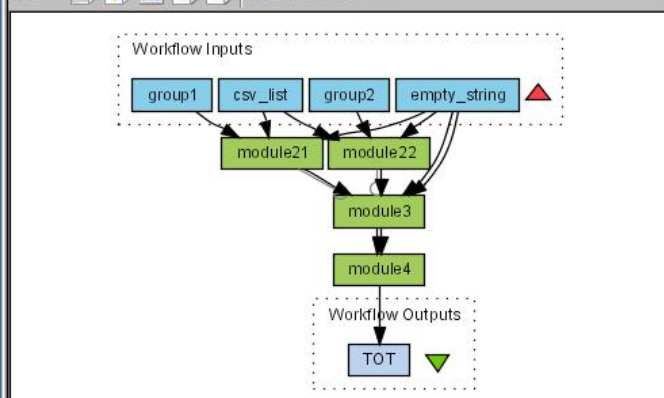
Type	Name	Last event	Event timestamp	Event detail	Breakpoint
module22	module22	ProcessComplete	11-mag-2006 19.32....		
module4	module4	ProcessComplete	11-mag-2006 19.32....		
module3	module3	ProcessComplete	11-mag-2006 19.32....		
module21	module21	ProcessComplete	11-mag-2006 19.32....		

Graph Intermediate inputs Intermediate outputs

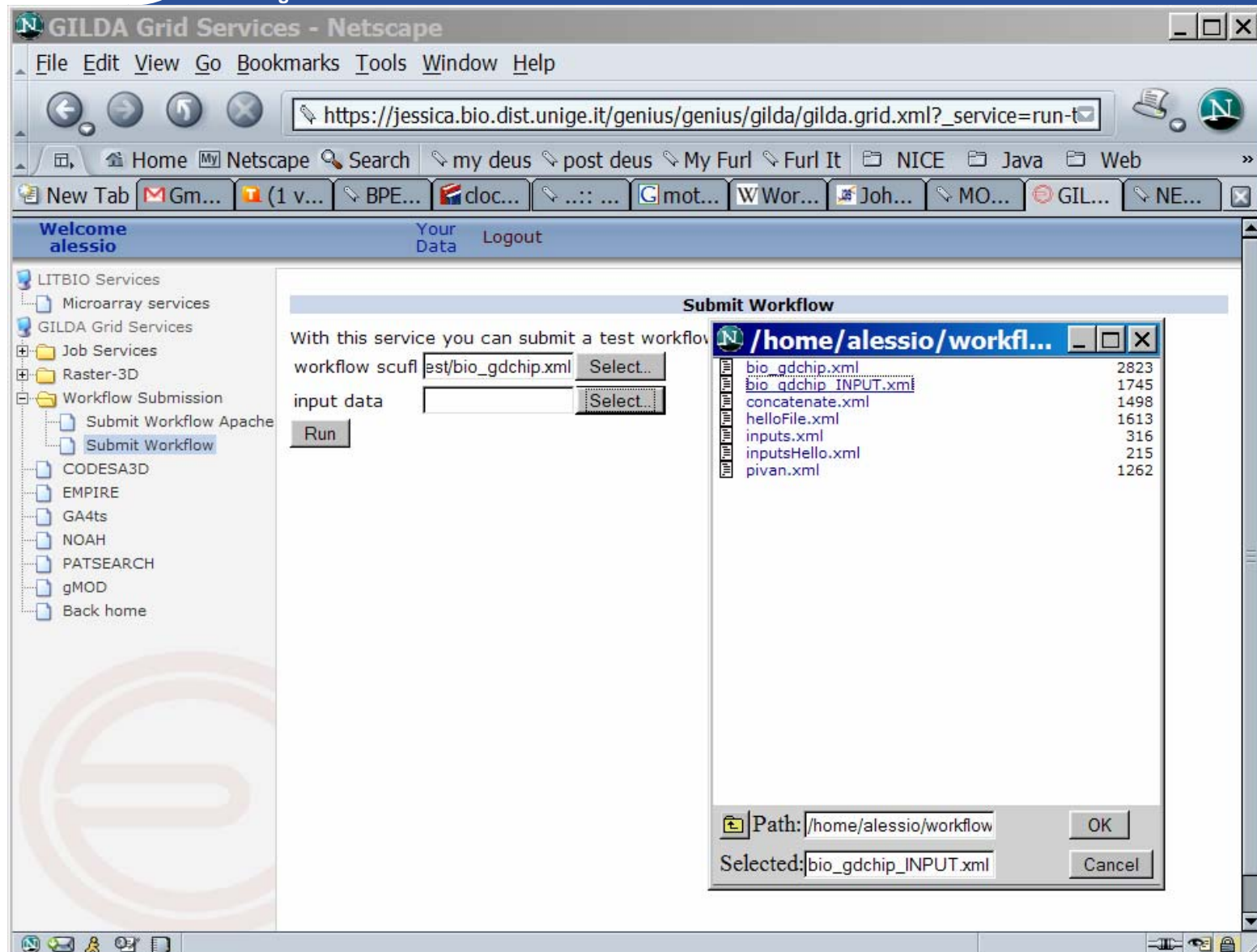


### Workflow diagram

Save as Configure diagram



Rendering done.



**GILDA Grid Services - Netscape**

File Edit View Go Bookmarks Tools Window Help

https://jessica.bio.dist.unige.it/genius/genius/gilda/gilda.grid.xml?\_service=run-t

Home My Netscape Search my deus post deus My Furl Furl It NICE Java Web

New Tab Gm... (1 v... BPE... doc... ...: ... mot... Wor... Joh... MO... GIL... NE...

Welcome alessio Your Data Logout

**LITBIO Services**

- Microarray services
- GILDA Grid Services**
  - Job Services
  - Raster-3D
  - Workflow Submission
    - Submit Workflow Apache
    - Submit Workflow**
  - CODESA3D
  - EMPIRE
  - GA4ts
  - NOAH
  - PATSEARCH
  - gMOD
  - Back home

**Submit Workflow**

With this service you can submit a test workflow

workflow scufl

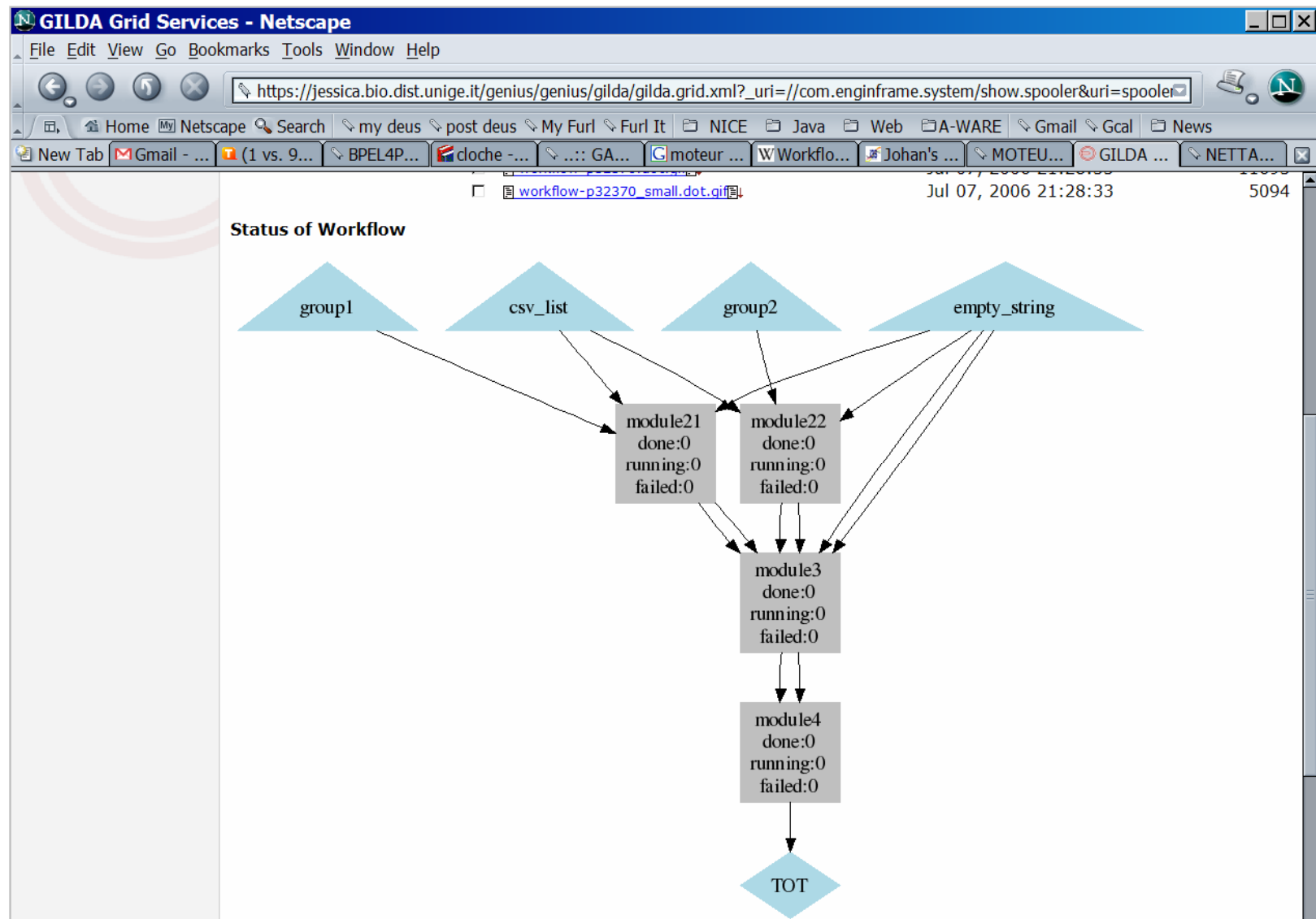
input data

**/home/alessio/workfl...**

bio_gdchip.xml	2823
bio_gdchip_INPUT.xml	1745
concatenate.xml	1498
helloFile.xml	1613
inputs.xml	316
inputsHello.xml	215
pivan.xml	1262

Path: /home/alessio/workflow

Selected: bio\_gdchip\_INPUT.xml



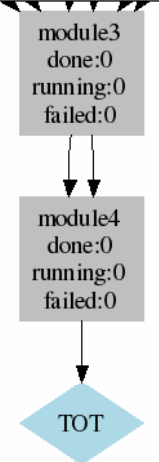
**GILDA Grid Services - Netscape**

File Edit View Go Bookmarks Tools Window Help

https://jessica.bio.dist.unige.it/genius/genius/gilda/gilda.grid.xml?\_uri=//com.engineframe.system/show.spooler&uri=spooler

Home My Netscape Search my deus post deus My Furl Furl It NICE Java Web A-WARE Gmail Gcal News

New Tab Gmail - ... (1 vs. 9... BPEL4P... cloche -... :: GA... moteur ... Workflow... Johan's ... MOTEU... Loading... NETTA...



```

graph TD
    module3["module3  
done:0  
running:0  
failed:0"] --> module4["module4  
done:0  
running:0  
failed:0"]
    module4 --> TOT{"TOT"}
  
```

ID	Time	Status	Host	Name
Jul 07 21:25	DONE	jessica	/opt/genius/jdk15/bin/java Moteur --egee --	

**Details** Resource usage: CPU: 00:00:00s.

**Submission:**

- Job manager: os
- Time: 2006 Jul 07 21:25
- Host: jessica.bio.dist.unige.it
- Directory: /opt/genius/engineframe/spoolers/alessio/tmp64720.ef
- Command: /opt/genius/jdk15/bin/java Moteur --egee --noBoring -w /opt/genius/engineframe/spoolers/alessio/tmp64720.ef/scufl.32367.xml -d /home/alessio/workflow/test/bio\_gdchip\_INPUT.xml -id workflow-p32370 1> workflow-p32370.out 2>workflow-p32370.err

**Execution:**

- Time: 2006 Jul 07, 21:25
- Host: jessica
- Directory: /opt/genius/engineframe/spoolers/alessio/tmp64720.ef

**Termination:**

- Time: 2006 Jul 07, 21:25
- Total cpu usage: 00:00:00
- Exit code: 0

**GILDA Grid Services - Netscape**

File Edit View Go Bookmarks Tools Window Help

Address bar: [https://jessica.bio.dist.unige.it/genius/genius/gilda/gilda.grid.xml?\\_uri=//com.engineframe.system/show.spooler&uri=spooler](https://jessica.bio.dist.unige.it/genius/genius/gilda/gilda.grid.xml?_uri=//com.engineframe.system/show.spooler&uri=spooler)

Navigation: Home My Netscape Search my deus post deus My Furl Furl It NICE Java Web A-WARE Gmail Gcal News

Tab bar: New Tab Gmail - ... (1 vs. 9... BPEL4P... cloche -... :: GA... moteur ... W Workflow... Johan's ... MOTEU... GILDA ... NETTA...

Welcome alessio Your Data Logout

Left sidebar:

- LITBIO Services
  - Microarray services
- GILDA Grid Services
  - Job Services
  - Raster-3D
  - Workflow Submission
    - Submit Workflow Apache
    - Submit Workflow
  - CODESA3D
  - EMPIRE
  - GA4ts
  - NOAH
  - PATSEARCH
  - gMOD
  - Back home

Content area:

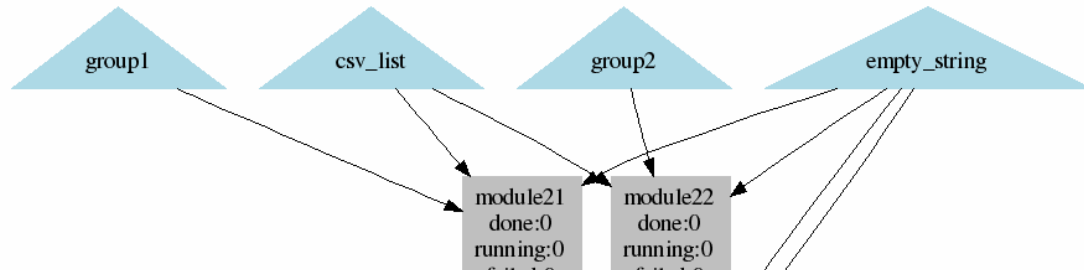
Workflow: tmp64720.ef/workflow\_runs/workflow-p32370 Jul 08, 2006 21:25:01

[ Top ] > workflow\_runs > workflow-p32370

Toggle select Clear select Action...

File Name	Date/Time	Size
<a href="#">workflow-p32370.out</a>	Jul 07, 2006 21:25:03	2688
<a href="#">workflow-p32370.err</a>	Jul 07, 2006 21:25:03	225
<a href="#">workflow-p32370.status</a>	Jul 07, 2006 21:25:03	247
<a href="#">workflow-p32370.dot</a>	Jul 07, 2006 21:25:03	931
<a href="#">workflow-p32370-status.html</a>	Jul 07, 2006 21:25:03	320
<a href="#">workflow-p32370.html</a>	Jul 07, 2006 21:25:03	1427
<a href="#">workflow-p32370-module21.html</a>	Jul 07, 2006 21:25:03	628
<a href="#">workflow-p32370-module22.html</a>	Jul 07, 2006 21:25:03	628
<a href="#">workflow-p32370-module4.html</a>	Jul 07, 2006 21:25:03	626
<a href="#">workflow-p32370-module3.html</a>	Jul 07, 2006 21:25:03	654
<a href="#">workflow-p32370-csv_list.html</a>	Jul 07, 2006 21:25:03	470
<a href="#">workflow-p32370-group1.html</a>	Jul 07, 2006 21:25:03	466
<a href="#">workflow-p32370-group2.html</a>	Jul 07, 2006 21:25:03	466
<a href="#">workflow-p32370-empty_string.html</a>	Jul 07, 2006 21:25:03	478
<a href="#">workflow-p32370-TOT.html</a>	Jul 07, 2006 21:25:03	465
<a href="#">workflow-p32370-pivan.err</a>	Jul 07, 2006 21:26:25	0
<a href="#">workflow-p32370.dot.gif</a>	Jul 07, 2006 21:26:25	11693
<a href="#">workflow-p32370_small.dot.gif</a>	Jul 07, 2006 21:26:25	5094

**Status of Workflow**



```

graph TD
    group1 --> module21
    csv_list --> module21
    csv_list --> module22
    group2 --> module22
    empty_string --> module22
    module21["module21  
done:0  
running:0"]
    module22["module22  
done:0  
running:0"]
  
```



- **Advantages provided by the introduction of a server-side component**
- **Limits: Moteur appears as a performing enactor tool, but it imposes some limitations about processors and data inputs**
- **Future steps: the proposed infrastructure is still a prototype.**
  - Stay tuned on the evolution of MOTEUR...
  - Improve Web interface to provide users with dynamically generated forms for workflow inputs definition

- **Project goals**
  - Simplify users'
    - life (focus on problems)
    - way of perceiving the GRID
  - Fill an existing gap
    - between middleware and portals
  - common interface to access gLite, Unicore/GS (DEISA project) and GT4 (TeraGrid project)
  - EnginFrame + A-WARE to be a completely integrated solution



**An easy Way to  
Access GRID  
Resources**





## Q&A

