

Porting application to Grid

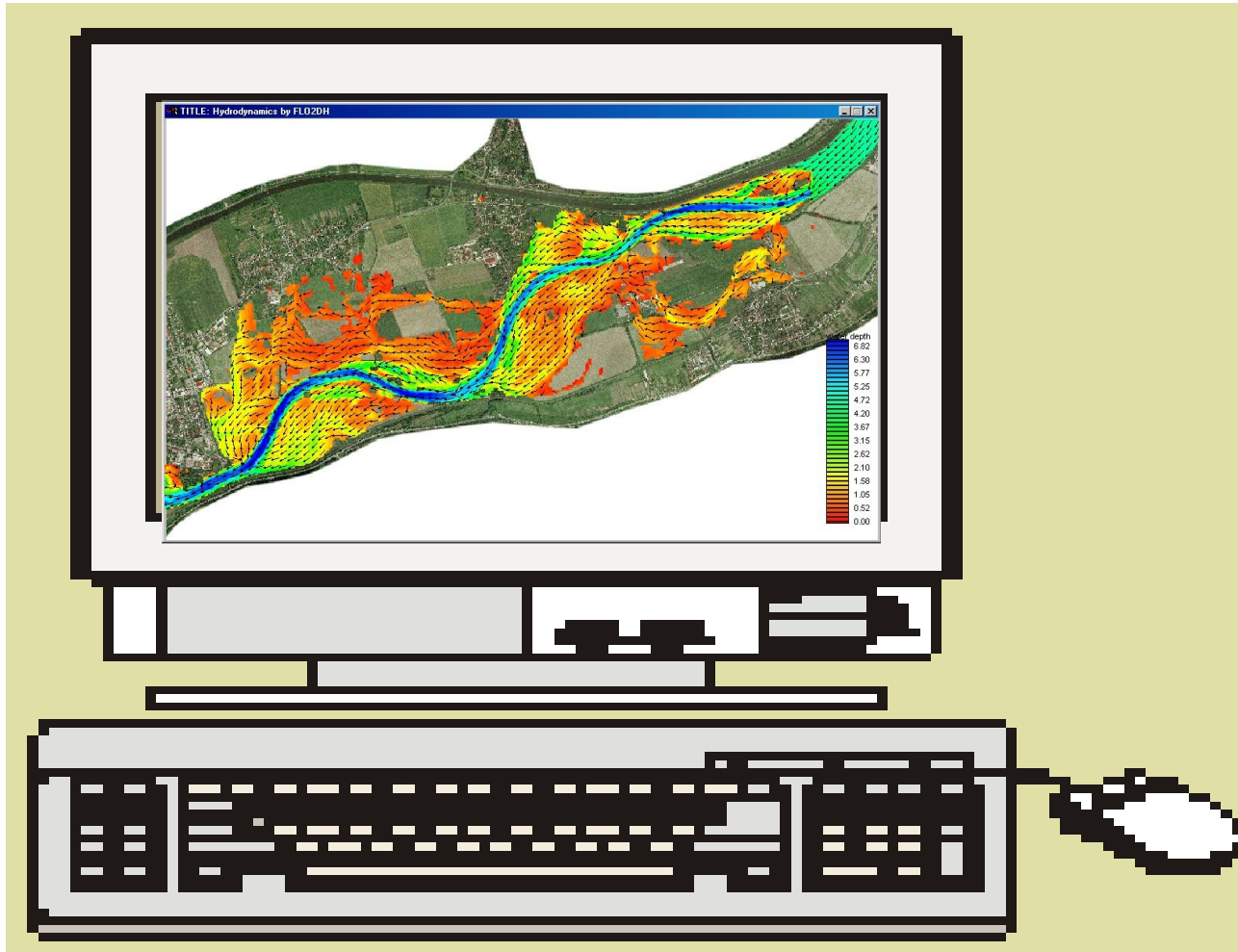
Viet D. Tran

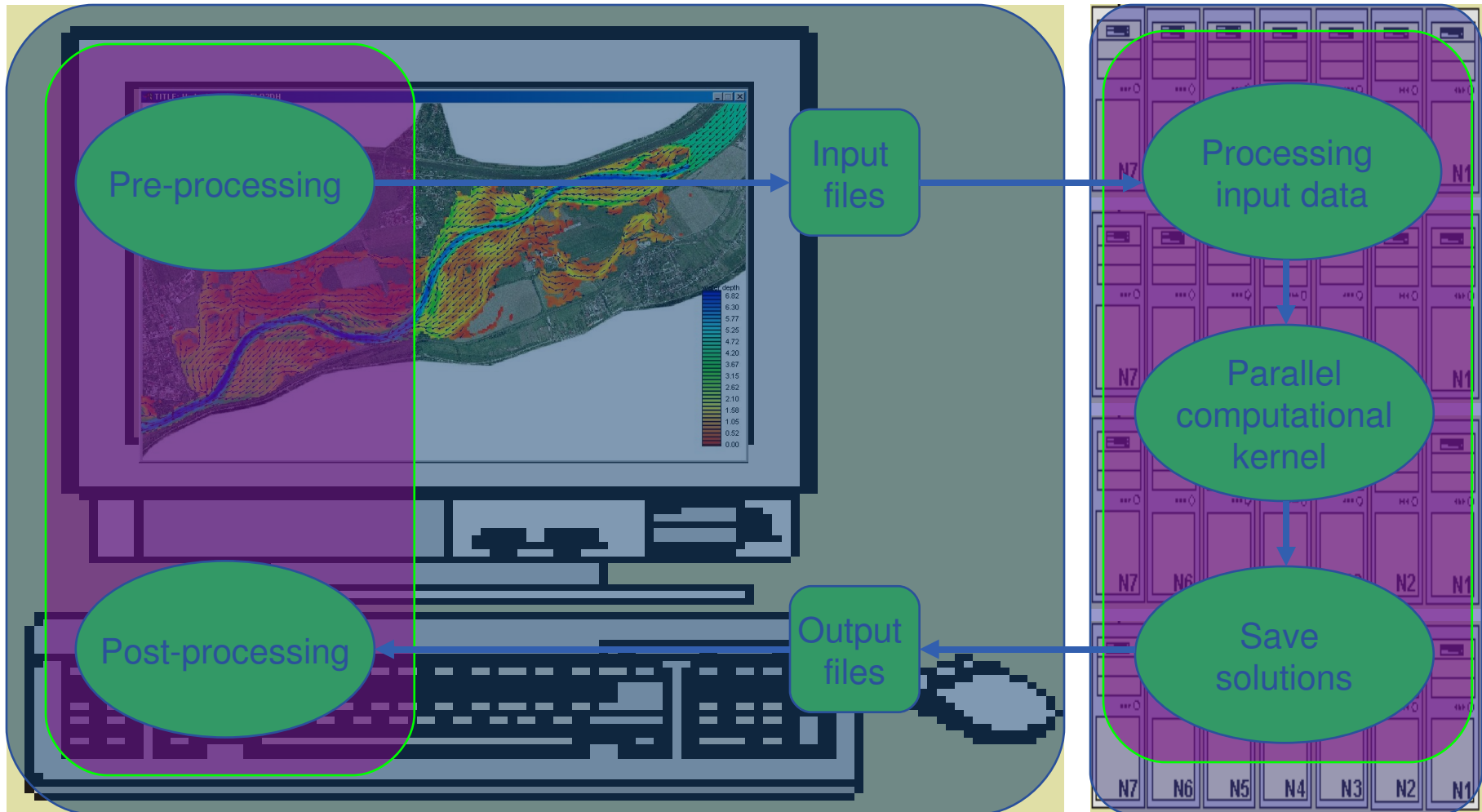
Institute of Informatics, Slovakia

- 1. Preparation of your application**
- 2. Joining Grid**
- 3. First job**
- 4. Full computation**
- 5. Advanced Grid tools**
 - Portal
 - Data management
 - Workflow

- **Preparation of your application**
- **Joining Grid**
- **First job**
- **Full computation**
- **Advanced Grid tools**
 - Portal
 - Data management
 - Workflow

- **Application should run in batch mode:**
 - Separate computations from graphical interfaces
 - Input and output data from/to files
- **Porting application to Linux**
 - Compile, link and test your application in Linux
 - For MPI parallel program, testing on Linux cluster
 - Non-standard libraries (e.g. LAPACK) should be linked statically
 - TEST: try to copy your application to a new, clean Linux machine and test it there





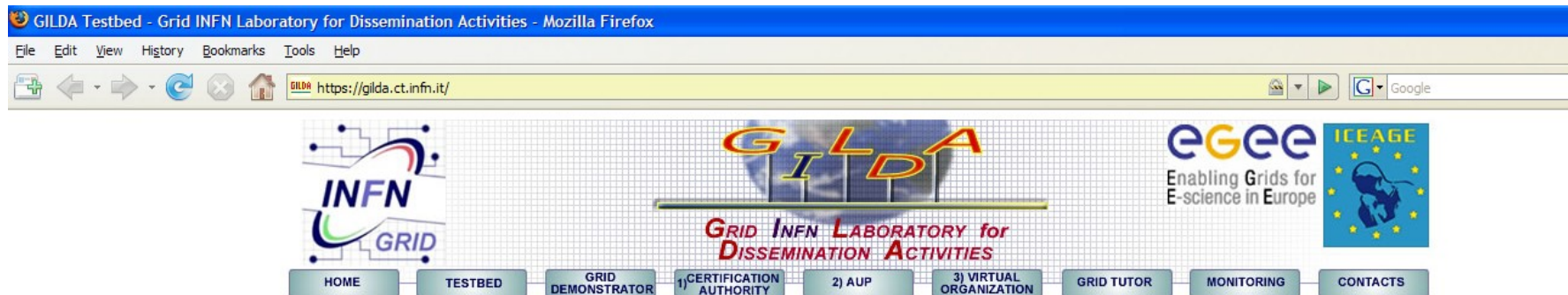
1. Preparation of your application

Results: Applications running in Linux in batch mode

- **Joining Grid**
- **First job**
- **Full computation**
- **Advanced Grid tools**
 - Portal
 - Data management
 - Workflow

- **Choosing the infrastructure for porting**
 - GILDA testing infrastructure
 - Simple registration, available for every ones
 - Suitable for testing, learning and demonstration
 - Should not be used for stress testing (too long or too many jobs), especially during some courses or demonstrations
 - EGEE production infrastructure
 - Suitable for running application in production

- **GILDA main web page:** <https://gilda.ct.infn.it/> with many tutorials, documentations and training materials
- **Register and get a GILDA certificate**
<https://gilda.ct.infn.it/CA/>
- **Get an account on User Interface (UI) machines**
 - GILDA automatically create an account on `glite*.gilda.infn.it` during registration but the UI machines have firewalls for SSH
 - Contact with us viet.ui@savba.sk to have an account on our UI `dgt03.ui.sav.sk`.
- **Start to learn grid commands and test them**



- News ^{NEW}
- List of Tutorials
- Request a Tutorial
- Posters
- Video Tutorials
- User Interface PnP
- Virtual Services
- Instructions for Users
- Instructions for Sites
- Training Material (wiki)
- GILDA Forge ^{NEW}
- Support System
- Useful Links
- Sponsors
- Usage Statistics
- Old Usage Statistics

^{NEW} [New GILDA usage rule](#) ^{NEW}

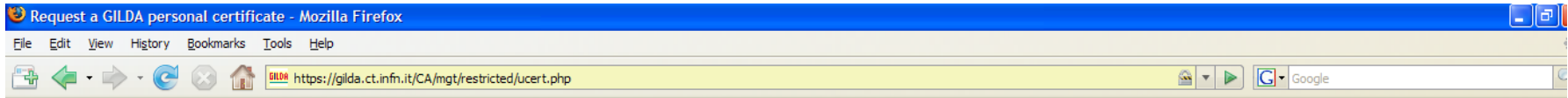
GILDA (G rid I nfn L aboratory for D issemination A ctivities)

is a virtual laboratory to demonstrate/disseminate the strong capabilities of grid computing.

GILDA consists of the following elements:

- ♦ [the GILDA Testbed](#): a series of sites and services (Resource Broker, Information Index, Data Managers, Monitoring tool, Computing Elements, and Storage Elements) spread all over Italy and the rest of the world on which the latest version of the [INFN Grid](#) middle-ware (fully compatible with [gLite](#)) is installed;
- ♦ [the Grid Demonstrator](#): a customized version of the full [GENIUS web portal](#), jointly developed by INFN and [NICE](#), from where **everybody** can submit a pre-defined set of applications to the GILDA Testbed;
- ♦ [the GILDA Certification Authority](#): a fully functional Certification Authority which issues 14-days X.509 certificates to everybody wanting to experience grid computing on the GILDA Testbed;
- ♦ [the GILDA Virtual Organization](#): a Virtual Organization gathering all people wanting to experience grid computing on the GILDA Testbed; the GILDA Virtual Organization is based on the Virtual Organization Membership Service (VOMS) developed by INFN;
- ♦ [the Grid Tutor](#): based on a full version of the [GENIUS web portal](#), to be used only during [grid tutorials](#);
- ♦ [the GILDA P-GRADE Portal](#): a Grid workflow developer and executor environment based on the latest release of the [P-GRADE Grid Portal](#) provided by [MTA SZTAKI](#);
- ♦ [the monitoring systems](#): several tools monitoring the status of the GILDA Testbed and the availability of its grid services;
- ♦ [the GILDA mailing lists](#): gilda-sitemanagers@ct.infn.it, for system administrators (both lists are restricted to members only). A web form to subscribe/unsubscribe to/from the lists is available [here](#).

GILDA is an activity of the Italian [Istituto Nazionale di Fisica Nucleare \(INFN\)](#) carried on in the context of the following Projects:



The GILDA Certification Authority

➤ General information

➤ GILDA CA certificate

➤ Request a personal certificate

➤ Request an account

➤ Request a host certificate

➤ Renew a certificate

➤ Check a personal certificate

➤ Certificate Revocation List

Request a GILDA personal certificate

If you did not do it already, please [download the GILDA CA certificate](#) first.

In order to correctly generate a request it is mandatory to fill **all** fields in the form below. Please, double check the correctness of the e-mail address that you are going to provide since **no verification** will be performed by the server.

The password you are prompted about in the form below is the password of your personal account on the **GENIUS Portal** from where you will access the GILDA Testbed and it is **NOT** the passphrase of your personal certificate.

When the certificate will be signed by the GILDA CA manager you will be notified by e-mail with the instructions to download your GILDA CA personal certificate and access the GILDA Testbed.

Institute/University/Company:	<input type="text"/>
First name and last name:	<input type="text"/>
Account username (max 8 characters; only not-accented letters and digits are allowed, both lowercase and uppercase):	<input type="text"/>
Account password (only not-accented letters and digits are allowed, both lowercase and uppercase):	<input type="password"/>
Confirm account password (only not-accented letters and digits are allowed, both lowercase and uppercase):	<input type="password"/>
E-mail:	<input type="text"/>
KeySize:	2048 (High Grade) <input type="button" value="v"/>

- **Register and get an EGEE certificate**
 - Contact with Miroslav Dobrucky (dobrucky.ui@savba.sk)
- **Join to a Virtual Organization (VO)**
 - If your application belongs to areas of some existing VO (e.g. biomedicine, chemistry, earth science, ...) you can contact with the VO manager and ask him to add you to his VO
 - Send us short description of your application, we can give contact of the corresponding VO manager
 - If your application does not belong to any VO, you can still join the generic VOCE VO (Central European VO)
<https://voce-register.farm.particle.cz/voce>

VOCE REGISTRATION - Mozilla Firefox

File Edit View History Bookmarks Tools Help

https://voce-register.farm.particle.cz/voce/

Google

VOCE REGISTRATION

For detailed information see [VOCE homepage](#). If you encounter problem during registration please contact voce@cesnet.cz

For access to the **VOCE** resources, you must agree to the **VOCE Usage Rules** and register with the Virtual Organization (VO). Please fill out all fields in the form below and click on the appropriate button at the bottom. After submitting the request a confirmation email will be sent to your address. In order to finish the registration process please follow instructions from the mail.

IMPORTANT: By submitting this information you agree that it may be distributed to and stored by VOCE and site administrators, that action may be taken to confirm the information you provide is correct, that it may be used for the purpose of controlling access to VOCE resources and that it may be used to contact you in relation to this activity.

Family name:

Name:

Institute:

Phone number:

Address:

Email:

Certificate DN: /C=SK/O=SlovakGrid/O=IISAS/CN=Viet Tran

Done

voce-register.farm.particle.cz

1. Preparation of your application

Results: Applications running in Linux in batch mode

2. Joining Grid

Results: Receive Grid certificate and join VO

- **First job**
- **Full computation**
- **Advanced Grid tools**
 - Portal
 - Data management
 - Workflow

- **Copy your executable program (myprogram.bin) and input data (myinput.dat) to User Interface (UI) machine** (dgt03.ui.sav.sk)
- **Log in to the UI** (ssh dgt03.ui.sav.sk)
- **Create a simple JDL file** (vi myfirstjob.jdl)
Executable= "myprogram.bin";
Arguments= "myinput.dat";
InputSandbox = {"myprogram.bin","myinput.dat"};
StdOutput = "std.out";
StdError = "std.err";
OutputSandbox = {"myoutput.dat", "std.out","std.err"};
- **Create proxy certificate:** voms-proxy-init
- **And submit the program:** glite-wms-job-submit -a myfirstjob.jdl
- **That is applicable only for smaller data and binary (< 2MB)**

- If you have large data and binary, first upload them to storage elements (storage.ui.sav.sk)
- Then create a simple script to download the binary, input data and upload the output data (myscript.sh)

```
globus-url-copy gsiftp://storage.ui.sav.sk/data/myprogram.bin \  
file://$PWD/myprogram.bin
```

```
chmod a+x myprogram.bin
```

```
globus-url-copy gsiftp://storage.ui.sav.sk/data/myinput.dat \  
file://$PWD/myinput.dat
```

```
./myprogram.bin myinput.dat
```

```
globus-url-copy file://$PWD/myoutput.dat \  
gsiftp://storage.ui.sav.skdata/myoutput.dat
```

- Create a JDL to submit a job with the script

```
Executable= "myscript.sh";
```

```
InputSandbox = {"myscript.sh"};
```


1. Preparation of your application

Results: Applications running in Linux in batch mode

2. Joining Grid

Results: Receive Grid certificate and join VO

3. First job

Results: Application is running in Grid with simple testing data

- **Full computation**
- **Advanced Grid tools**
 - Portal
 - Data management
 - Workflow

- **For parametric study, first create the JDL (param.jdl):**

```
JobType = "Parametric";
```

```
Executable = "myprogram.bin";
```

```
Arguments = "myinput_PARAM_.dat";
```

```
StdOutput = "std_PARAM_.out";
```

```
StdError = "std_PARAM_.err";
```

```
Parameters = 100;
```

```
ParameterStart = 1;
```

```
ParameterStep = 1;
```

```
InputSandbox = {"myprogram.bin", "myinput_PARAM_.dat";
```

```
OutputSandbox = {"myoutput_PARAM.dat", std_PARAM_.out",  
  "std_PARAM_.err"};
```

- **Then submit the JDL to Grid:** `glite-wms-job-submit param.jdl`

- **For MPI jobs, just add to the JDL**

```
JobType = "MPICH";
```

```
NodeNumber = 16;
```

```
Executable= "myprogram.bin";
```

```
Arguments= "myinput.dat";
```

```
InputSandbox = {"myprogram.bin","myinput.dat"};
```

```
StdOutput = "std.out";
```

```
StdError = "std.err";
```

```
OutputSandbox = {"myoutput.dat", "std.out","std.err"};
```

1. Preparation of your application

Results: Applications running in Linux in batch mode

2. Joining Grid

Results: Receive Grid certificate and join VO

3. First job

Results: Application is running in Grid with simple testing data

4. Full computation

Results: Application is running in Grid with full data

5. Advanced Grid tools

- Portal
- Data management
- Workflow

- **Command line interface** (glite-wms-job-* commands) **is powerful but not user-friendly**
- **Portal is useful because:**
 - Accessible from any where
 - Hides complexity: the actual users do not have to learn about gLite or similar things
 - More secure
 - Useful for demonstration, can be combined with visualization
- **Recommended Grid portal framework: Gridsphere**
 - <http://www.gridisphere.org/>
 - Supporting portlet (JSR168)
 - Easy to use and extend
- **Ready to use portlets:**
 - Certificate management
 - Job submission
 - Data browser

New Credential

This credential can be retrieved from **myproxy.gridlab.org**.

User Name: (Your credential repository username)

Credential Name: (Name you assigned to credential in repository)

Credential Label: (Label to display for credential in portal)

Passphrase: (Your credential repository password)

Use Portal Credential: ☒ (Leave checked to use the portal credential)

Single Sign-On: ☒ (Leave checked to sign-on to the portal)

File Browser Portlet

Physical Files Logical Files

File Browser 1

<Select Resource>

Portal
Peyote
Helix
Venus
Skrit
FS0
NO
Rage1
Litchi
HitCross

File Browser 2

<Select Resource>

Portal
Peyote
Helix
Venus
Skrit
FS0
NO
Rage1
Litchi
HitCross

Job creation

Select template:

Job

Job name:

Factory URI:

Job configuration

daveFOutputFile:

visualizationOutputFile:

User job list

Job: **DaveF2D visualization** State: CREATED Creation date: 2006-08-23 15:43:35.472

Job configuration

daveFOutputFile: davef-out-2005-06-24

visualizationOutputFile: davef-2d-visualization-2005-06-24

Job: **HSPF experiment 1** State: CREATED Creation date: 2006-08-23 15:43:07.27

Job: **DaveF 2D visualization** State: CREATED Creation date: 2006-08-23 15:42:56.399

Data Resource Explorer

Welcome Grid Job services Data Services Visualization

Metadata Services VFS Browser Data Resource Explorer

Target host

Host:

Connected: https://gaia.ui.savba.sk:8443/wsrf/services/medigrid/dataservices/DataResourcesManager

Host browser

All resources #fixed_data#

Content	Attributes
[..]	
[test_data]	
ahoj.meta	View Attributes
ahoj	View Attributes
metatest.txt.meta	View Attributes
ahoj.meta~	View Attributes
metatest.txt	View Attributes
rftInx.tmp	View Attributes
rftInx.tmp.meta	View Attributes
testdata1	View Attributes

Attributes

Owners	Users	Hosts
/O=MediGrid/O=IISAS/CN=ROOT		
/O=MediGrid/O=ALGO/CN=Milos Kokkosoulis	RW	
/C=SK/O=SlovakGrid/O=IISAS/CN=Martin Maliska	RW	
/C=SK/O=SlovakGrid/O=IISAS/CN=Viet Tran	R	
/O=MediGrid/O=IISAS/CN=Branislav Simo	RW	
/O=MediGrid/O=IISAS/CN=Martin Maliska	RW	
/O=MediGrid/O=IISAS/CN=Marek Ciglan	RW	

Add new attribute

New identified as:


☒ Read ☐ Write

- **Big applications usually also have big data sets**
- **Some data management tools in Grid**
 - Replica management
 - Each piece of data may have multiple copies (files) to improve the data access
 - Can provide some optimization
 - Metadata catalogs
 - Describe the data in Grid
 - Searching, browsing
 - Access to databases (OGSA-DAI)
 - Access to databases (e.g. Oracle, MySQL) from your program running on the Grid

Grid Workflow :: Workflow Engines - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.gridworkflow.org/snips/gridworkflow/space/Workflow+Engines



The Grid Workflow Forum [start | index | login]

start > Workflow Engines

Workflow Engines

Created by [bassheide](#). Last edited by [bassheide](#), 131 days ago. Viewed 4,194 times. #18

[diff] [history] [edit] [rdf]

WORKFLOW ENGINES

- BioPipe
- BizTalk
- BPWS4J
- DAGMan
- GridAnt
- Grid Job Handler
- GRMS (GridLab Resource Management System)
- GWFE (Gridbus Workflow Engine)
- GWES (Grid Workflow Execution Service)
- IT Innovation Enactment Engine
- JIGSA
- JOpera
- Kepler
- Karajan
- OSWorkflow
- Pegasus (uses DAGMan)
- Platform Process Manager
- ScyFLOW
- SDSC Matrix
- SHOP2
- Taverna

labels

attachments

Search

Events/Calls for Papers
Glossary
Projects
Research Fields
Researchers
Workflow Composition Tools
Workflow Description Languages
Workflow Engines
Links
Imprint
Contact

Help
For hints about formatting text
see [snipsnap-help](#).

Logged in Users: (0)
... and 3 Guests.

Recently Changed

- 👤 j.chen
- 📅 Events
- 📅 Calls for Papers
- 👤 o.ezenwoye
- 📅 2007-10-08 #1
- 📅 IEEE Transactions on Automation Science and Engineering
- 📅 snipsnap-portlet-1
- 👤 j.chen (l.cancan)
- 👤 e.zimeo
- 👤 e.elmroth
- 👤 j.chen (j.chen)

Done

1. Preparation of your application

- Applications running in Linux in batch mode

2. Joining Grid

- Receive Grid certificate and join VO

3. First job

- Application is running in Grid with simple testing data

4. Full computation

- Application is running in Grid with full data

5. Advanced Grid tools

- User-friendly GUI for applications

- **Porting application to Linux in batch mode**
- **Getting Grid certificates and joining some VO**
- **Testing application to Grid with simple testing data**
- **Full computation**
- **Portal and data/workflow management**

Please contact to our team for additional helps