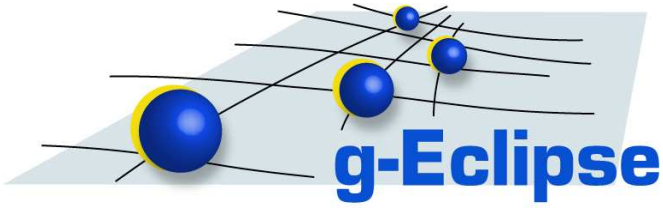

g-Eclipse hands-on @ EGEE09

Handout

Document Filename:	**g-Eclipse-EGEE09-handout.pdf**
Partner(s):	**FZK, PSNC, JKU, UCY, INO, RUR, IT Innovation, NLE-IT**
Presenter(s):	**S. Girtelschmid, S. Mueller, A. García**

Abstract: This document gives a short introduction into the most important features of g-Eclipse.

Contents

1 Installation **3**

2 Setup **4**

3 Creating a Project **5**

4 Data Management **6**

4.1 Mounting a Storage Element . . . 6

4.2 Manually mounting a location . . . 6

4.3 Manually mounting a remote location . . . 7

4.4 Transferring and Managing Data . . . 7

5 Job Submission **9**

5.1 Creating and editing a Job Description File . . . 9

5.2 Submitting a Job Description File . . . 9

5.3 Managing already submitted jobs . . . 10

5.4 Creating and submitting a Workflow . . . 10

6 Visualisation **11**

6.1 Creating a VTK pipeline . . . 11

6.2 Rendering a VTK pipeline . . . 11

7 Application Development **12**

7.1 Creating a HelloWorld Application . . . 12

7.2 Building your application remotely . . . 12

7.3 Debugging your application on a remote machine . . . 13

8 Site Administration **14**

8.1 Managing Batch Services . . . 14

8.2 Remote Command Console . . . 14

A Additional Information **15**

1 Installation

There are basically two ways of installing g-Eclipse[1].

- If you are installing Eclipse from scratch you can simply download and install one of the g-Eclipse EPP-packages.
  1. Download one of the EPP-packages at http://www.geclipse.eu/download/ or if you want to use a nightly build from http://iwr-geclipse.fzk.de:8443/luntbuild/app.do (Latest build →EPP-packages).
  2. Extract the package.
  3. g-Eclipse can now be started from the extracted executable (*eclipse.exe* or *eclipse*).
- If you already have a running Eclipse 3.4 instance on your machine you may want to use the Eclipse update manager to install g-Eclipse.
  1. Check your JVM memory settings in *eclipse.ini*; use at least *-Xmx512m* as maximum Java heap size and *-XX:MaxPermSize=256m* to avoid OutOfMemoryError exceptions.
  2. Start your Eclipse 3.4 instance and open the Eclipse Update Manager with **Help →Software Updates...**.
  3. In the Software Updates dialog choose **Add Site...**.
  4. Add the g-Eclipse Update Site http://www.geclipse.eu/update/1.0milestones/.
  5. After scanning the different update sites you will get a list of new features that are available. Select all g-Eclipse features.
  6. Press **Install...** and accept the terms in the license agreement. Follow the wizard that will download and install all necessary features and plug-ins. Depending on your Internet connection this may take some time.
  7. After all features have been installed Eclipse must be restarted.

[1] We assume, that a Java Runtime Environment (JRE) version 1.5 or higher is already installed on your machine.

2 Setup

In this section we will guide you through the process of setting up g-Eclipse for the first time.

- Starting Eclipse/g-Eclipse
  1. When starting Eclipse/g-Eclipse for the first time you are asked for the location of your workspace. Your workspace is a folder on your hard disk where Eclipse saves all projects and related settings. So make your choice and press **Ok**.
  2. Eclipse does now start with the Welcome screen. There you have the possibility to explore Eclipse and also g-Eclipse on your own by looking at available tutorials or documentation. At the moment we are skipping this welcome screen by pressing the button with the curved arrow which will bring you to the Eclipse workbench.
  3. There are four different perspectives in g-Eclipse: User, Developer, Operator, and Grid Exploring (changable by the buttons in the upper right corner of Eclipse). The first tree mentioned perspectives reflect the three Grid actor groups and the last perspective is there for file management. First we will work in the User's perspective.
- Setting up the most important g-Eclipse parameters[1].
  - Importing the CA certificates:
    1. Go to **Window →Preferences...**.
    2. In the tree view on the left expand the **g-Eclipse** node.
    3. Select **Security** and press **Import...**.
    4. In the upcoming wizard choose the **European Policy Management Authority →accredited** and press **Next**.
    5. On the last wizard page press **Select All** and afterwards Finish.
  - Importing additional certificates such as the Gilda CA certificate[2][3]:
    1. Go to `https://gilda.ct.infn.it/CA/mgt/getCA.php`, choose *PEM* as format and press **Download/Visualise certificate**.
    2. Now copy and paste everything between and including `-----BEGIN CERTIFICATE-----` and `-----END CERTIFICATE-----` to a local text file called *gildaCA.pem*.
    3. From the g-Eclipse Security preference page press **Add...**
    4. Select the file that you just created and press **Ok**.
  - Setting up a VO:[4]
    1. Go to **Window →Preferences...**.
    2. In the tree view on the left expand the **g-Eclipse** node.
    3. Select **VO-Declarations** and press **Import...**.
    4. In the upcoming wizard page leave the repository URL as it is and press **Next**.
    5. Select the VO you would like to import and press **Finish**.
    6. Sometimes due to errors in the published parameters you may have to change some settings by hand after importing your VO. Therefore select the VO and press **Edit...**. Verify and if necessary modify your VO settings and press **Finish** when done.

[1] If you are behind a proxy server do not forget to adjust your network settings accordingly at **General →Network Connections** before starting with the setup

[2] Only necessary if your certificate was issued by for example the Gilda CA

[3] For importing the Gilda CA, you may also use the Import button to download the CA directly (i.e. not having to follow the following steps)

[4] You may also postpone the VO setup until you start creating your Grid project

3 Creating a Project

In this section you will create your first Grid project.

1. If you are not in the g-Eclipse User Perspective, please switch to it now. Select **Window →Open Perspective →Other...**. In the upcoming dialog select **g-Eclipse (User)** and press **Ok**.
2. Right-click on the View titled **Grid Projects** and select **New →Grid Project**.
3. Choose a project name (*MyFirstGridProject*) and press **Next**.
4. Select a VO for this project by activating the corresponding checkbox and press **Finish**. If no VO is present at the VO list you have to define a VO first. Therefore press **New VO...** and set up a VO as described in Section 2.

After creating your first Grid project try to get familiar with the structure of this project. There are several standard folders within a Grid project. Most of them are optional except for the VO folder. The VO folder is the last folder in your Grid project. It is a so called virtual folder, i. e. a folder that has no corresponding files in your workspace. The VO folders present your personalised Grid. Expand it to see nodes for your available computing and storage resources and the services that you may use to access your Grid infrastructure.

Once you have created at least one Grid project you may also want to have a look at the Glue Information Viewer. This viewer is not in the User Perspective by default, but you can open it if you go to **Window →Show View →Other...** and select **g-Eclipse →Glue Information Viewer**. This view gives you a lot of details about your available infrastructure. Feel free to explore your Grid by browsing the Glue Information Viewer.

A very impressive view to your Grid can be achieved with the World Wind View. Therefore go to **Window →Show View →Other...**. Now select **g-Eclipse →World Wind**. This will open the World Wind View showing all computing sites on an interactive 3-dimensional globe that are currently available from the information system.

4 Data Management

In this section we give an introduction to the data management capabilities of g-Eclipse. Therefore we will transfer data from your local hard drive to a remote location and vice versa. In order to access any data you have to create a so called Grid connection. This can be either done by mounting a storage element, a data service like LFC or by manually mounting a remote location.

4.1 Mounting a Storage Element

1. Expand the VO folder of your Grid project.
2. Expand the Storage folder of your Grid project.
3. Right-click a Storage Element.
4. From the context menu select **mount**. This mount menu may have several sub items depending on the protocols that are registered for this storage element and the protocols that are supported by g-Eclipse. Most likely there is at least a *gsiftp:2811*-entry.
5. Select the *gsiftp*-entry.
6. If you are now opening the **Connections** folder of your Grid project you will realise that there is a new connection according to your mounted Storage Element.
7. To access this connection simply expand and browse it as it would be a local folder. If you need to authenticate in order to access the Storage Element g-Eclipse will pop up an authentication dialog where you have to enter your authentication details.

4.2 Manually mounting a location

1. Right-click the **Connections** folder of your Grid project and select **New →Grid Connection**.
2. Choose a name for your connection (*MyLocalConnection*) and press **Next**.
3. In the next wizard page you have to specify the scheme of the protocol that you would like to use in order to access the mounted location. We want to mount a local directory first. Therefore, choose *file* in the **Scheme** combo box. You will notice that the *URI* field has changed to *Local Directory*. If you choose another scheme the input fields will change in order to reflect the type of parameters you have to enter for that specific scheme.
4. Now enter a local path you would like to mount. In the easiest case just enter */* to mount your root directory (or *C:/* on Windows).
5. Furthermore, you can now press **create a temporary connection** in order to browse the specified location for a dedicated directory that will be mounted afterwards.
6. If your are done press **Finish**. The newly mounted connection will now appear in the **Connections** folder and can be browsed like any other connection.

4.3 Manually mounting a remote location

1. Right-click the **Connections** folder of your Grid project and select **New →Grid Connection**.
2. Choose a name for your connection (*myRemoteConnection*) and press **Next**.
3. In the next wizard page you have to specify the scheme of the protocol that you would like to use in order to access the mounted location. We want to mount your remote account directory (which has been created for you only for the purpose of this exercise) on hydra.gup.uni-linz.ac.at. Therefore, choose *gsiftp* in the **Scheme** combo box. You will notice that the wizard fields have changed to *GridFTP Server*, *GridFTP Port*, and *Remote Directory*. (If you choose another scheme the input fields will change in order to reflect the type of parameters you have to enter for that specific scheme.)
4. Enter into the fields the following in order: *hydra.gup.uni-linz.ac.at*, *2811*, and */home/local/bnc/*.
5. You can now click on **create a temporary connection** in order to browse the specified location for a dedicated directory that will be mounted afterwards.
6. If your are done press **Finish**. The newly mounted connection will now appear in the **Connections** folder and can be browsed like any other connection. Your personal folder under the */home/local/bcn/* directory corresponds to the number of your Gilda issued certificate. Also, under the *exampleFiles* folder, you will find the files we will use the exercises.

4.4 Transferring and Managing Data

Transferring data is straightforward if you have created a Grid connection. Simply drag and drop files or directories from your workspace to the connection or vice versa. You can also copy and paste data by making use of the context menus. The best way to transfer data is to make use of the Grid Exploring Perspective.

1. Switch to the Grid Exploring Perspective. Therefore select **Window →Open Perspective →Other...**. In the upcoming dialog select **Grid exploring** and press **Ok**.
2. The Grid exploring perspective consists of two so called connection views located side by side and a centered editor area.
3. To transfer data simply drag and drop the data from one connection to the other or respectively from one connection view to the other.
4. To open a file simply double-click that file. Notice that opening a file may take some time if the file is not local. Be careful to not open the wrong file (Opening a several Gigabyte binary would be a bad idea). But you may have a try with different file types like GIFs or PDFs.
5. After you opened a file g-Eclipse will show an appropriate editor for this file. In the simplest case this will be the normal text editor.
6. Now, try to edit the text file and save it back. Note that you just edited a file remotely! g-Eclipse does NOT create a local copy of remote files. It simply uses the input and output streams to read from and write to files. Therefore there is no difference in accessing local or remote files except that you may have to authenticate when accessing remote files.
7. You can also create new files or directories in a Grid connection, at least if the underlying access protocol does allow such operations. If you mounted a GSIFTP connection just right-click a remote folder and select **New →Other...**. To create a directory choose **General →Folder** or select

General →File to create a file. In the upcoming wizard specify the name of the file or folder and press **Finish**. As we will need an empty file later on to have our job output written to it, create such a file, somewhere, where you will remember, called stdout. You may also like to create an empty file for error messages to be written to and call it, for example, stderr (very handy!). The newly created files will be immediately opened in the editor after their creation.

8. To delete existing files and folders just right-click them and select **Delete**. The deletion will take place only after your confirmation.

5 Job Submission

This section will guide you through the process of job submission with g-Eclipse. [1]

5.1 Creating and editing a Job Description File

1. Right-click into the Grid Projects view and select **New →Job Description**.
2. Enter a name for your job description file (*MyFirstJob*) and press **Next**.
3. Enter */bin/bash* as the Executable file and *simple.sh* as one of the arguments.
4. Click the browse button on the right of the *Standard input file* input field. In the upcoming file dialog browse to your local or your hydra connection and navigate to the simple.sh script. Press **Ok**. This script is working on a .pov file, therefore, we need to stage in the dna.pov file as well, but this can be done on the next wizard page.
5. Click the next browse button on the right of the *Standard output file* input field. Again, navigate to one of your formerly created GSIFTP-connections and select the file stdout you created for output. Press **Ok**. If you have also created a file dedicated for error messages, follow the above procedure for the *Standard error file*. Click **Next**.
6. On the next page of the wizard, add the stage in file dna.pov and a stage out file called dna.tga, which is produced by the script.
7. Press **Finish**. Now a JSDL-file will be created and the JSDL-Editor will be opened.
8. The JSDL-Editor enables you to easily edit your JSDL-file within different editor pages. For example, under the Resources tab, you may specify which CEs you want to be allowed to execute your job. This might be helpful, for example, when you know that not all available CE have a specific program that your job needs to complete successfully. To have a look at the raw JSDL file, select the *MyFirstJob.jsdl* tab at the bottom of the editor.

5.2 Submitting a Job Description File

1. Once you have created and edited your JSDL you are ready to submit it to the Grid. Right-click the JSDL file in your Grid project and select **Submit Job...**.
2. Choose a directory where a refernece to the job status information is stored (usually the Jobs folder in your project) and press **Next**.
3. Choose one of the available WMS-services and press **Finish**. If you did not already create a VOMS-proxy you are asked now to do so.
4. After the job was submitted you will notice that it appears in the Jobs folder of your Grid project. The small decorator on the bottom left of the job icon gives you some basic information about the job's status.

[1] For the duration of this tutorial, the files refered to in these examples can be found in your hydra.gup.uni-linz.ac.at mounted directory under the exampleFiles folder or can be downloaded from http://www2.gup.jku.at/ sgirtel/sgirtel.html.

5.3 Managing already submitted jobs

1. Switch to the Jobs View either by selecting it at the bottom of the Grid User Perspective or by directly opening it with **Window →Show View →Others →g-Eclipse →Jobs**.
2. The job view contains all submitted jobs. It gives you some basic information about the job's id, its status and when the last update took place.
3. A Grid job is represented as a folder. If you expand the job folder you will see a copy of the job description file from which the job was created. Furthermore if you have specified input and output files these files are directly accessible as Grid connections and can easily be opened or transfered by double-clicking or dragging and dropping them.
4. To get more details about your job right-click on it and select **Open**. This will open the Job Details View showing you all available details of your job.
5. From the Job Details View, you may download the Logging and Bookkeeping information to trace any Grid related failures during job execution. [2]
6. You can see the results of a successfully executed job by openning the files in the Output Files folder of you Job. You can also drag and drop these files onto your locally mounted directory to have them available for viewing offline.
7. You can also delete old jobs by right-clicking them and selecting **Delete**. A job can be deleted from both the Job View and the Grid project.

5.4 Creating and submitting a Workflow

1. Right-click the Job Descriptions folder of your Grid project and select **New →Workflow Diagram**.
2. Give your workflow a name (*MyFirstWorkflow.workflow*) and press **Finish**. The workflow will now be created and the Workflow Editor will be started.
3. Add jobs to the workflow by dragging and dropping JSDL files into the workflow editor (one that generates an output file and one that takes an input file). The boxes that represent the jobs will have input and output ports which represent the data files that are staged in and out for the jobs.
4. You can connect an output port of a job with an input port of another job by drawing a link between both.
5. After saving your workflow you can submit it by right-clicking on the workflow file in your Grid project and selecting **Submit Job...**.
6. Choose one of the available WMS-services and press **Finish**. If you did not already create a VOMS-proxy you are asked now to do so.
7. After the workflow was submitted you may realise that it appears in the Jobs folder of your Grid project. The small decorator on the bottom left of the job icon gives you some basic information about the workflow's status.

[2] If your job has been aborted, you will find a hint about the reason in yuor stderr or in the Logging and Bookkeeping file. In our example, you may have run into trouble if the CE to which your job was send to does not support povray, which the simple.sh script invokes. To overcome this problem, limit the execution to a specific host by adding the egee-ce1.gup.uni-linz.ac.at CE to the Candidate Hosts in your jsdl file - this option is editable from under the Resource tab of the JSDL Editor.

6 Visualisation

Visualisation within g-Eclipse is done with the help of VTK. This section will show you how to set up a basic VTK pipeline and how to render this pipeline.

6.1 Creating a VTK pipeline

1. Right-click your Grid project and select **New →Other...**. Select **g-Eclipse →VTK Pipeline** and press **Next**.
2. Enter a filename (*MyFirstVisualisation.vtkpipeline*) and press **Next**.
3. Give your pipeline a name and press **Next**.
4. Now you can choose a predefined vtkpipeline by just checking the **Examples** box (leaving everything else empty) and clicking **Next**.
5. On the next page, choose the type of example set by selecting one of the radio buttons and then select the example from the drop down menu. Most examples require an input data and to find which exactly, just hover over the input data field to see a tooltip giving you the necessary hint. After doing so, press **Finish**.
6. The VTK pipeline file will be created and the VTK Pipeline Editor will open. Within the editor you have the chance to also edit your pipeline.

6.2 Rendering a VTK pipeline

1. Right-click your VTK pipeline file and select **Render →locally (VTK)**.
2. The rendering process starts and the Visualisation View will open. Within this view you can interact with your visualisation by left-/right-clicking and dragging the mouse. You can rotate, translate and zoom the visualisation.

7 Application Development

In this section we give a brief overview of features like remote compiling and debugging. Note that the remote compiling/debugging is using glogin which has to be available on the remote node.

7.1 Creating a HelloWorld Application

1. Switch to the g-Eclipse Developer Perspective by simply selecting **Window →Open Perspective →Other...**. In the upcoming dialog select **g-Eclipse (Developer)** and press **Ok**.
2. Right-click on the Project Explorer and select **New →Other...**. Select **C++ →C++ Project** and press **Next**.
3. Give your project a name (*MyHelloWorld*), select **Makefile →Hello World C++ Project** from the Project types and the **Linux GCC** from the Toolchain list and press **Finish**.[1]
4. Make minor modifications to the generated files: in the Makefile remove the -O2 flag and, in the C++ file, add some code, for example as done in the loop bellow. Then, place a breakpoint somewhere in the main. Note, that if you have a gcc compiler on your machine, the project will be build automatically. If that is the case, just remove these files that were generated by the local build.
   Example change in the C++ file:

```
int main(void) {
  for (int i = 0; i < 10; i++) {
    printf("hello world %i\n", i);
    fflush(stdout);
  }
  return EXIT_SUCCESS;
}
```

7.2 Building your application remotely

1. Locally built executables may not run on a remote machine due to different architectures. Therefore g-Eclipse offers the possibility to compile your application remotely. To do so right-click your C++ project and select **Properties**.
2. From the list on left, select **Remote Builder**.
3. Check the **Remote Builder for this Project** check box to activate the Remote Builder.
4. In the Contact String field, enter *hydra.gup.uni-linz.ac.at* .
5. *Leave the glogin executable filed empty (only needed if the glogin is not installed in the default location) and check the **Use Handshake File** check box.*
6. *For the Executable press **Search Project...** and select the MyHelloWorld executable.*

[1] If the Linux GCC does not show up under the available Toolchains, uncheck the ¿¿Show project types and toolchains only if they are supported on the platform¡¡.

7. *For the **URI**, enter* gsiftp://hydra.gup.uni-linz.ac.at:2811/home/local/bcn/bcnXX as your remote home directory, where the XX is the 2 digit number of your certicate.
8. For the executable, enter the same string as the name you gave to your C++ project. Press **Apply**.
9. In the list on the left, choose **Builders** and uncheck the CDT Builder. Your newly created Remote Builder should appear in the list and should be checked. Press **OK**

7.3 Debugging your application on a remote machine

1. Select **Debug Configurations...**.
2. Right-click **C/C++ Grid Application** and select **New**.
3. At the Main page press **Browse...** to select a project and select the formerly created MyHelloWorld project.
4. For the Executable press **Search Project...** and select the MyHelloWorld executable.
5. At the Grid page enter a remote server *hydra.gup.uni-linz.ac.at as the contact string and leave the field to specify the remote glogin executable empty again.*
6. *Press **Debug**. This will transfer your executable to the remote machine and launch it there. All output is fetched and retargeted to the local console.*
7. *You can now debug your application. The code is executed remotely and the debug information is available locally. You can step into your code, set breakpoints and do all the fancy stuff you could do if you would debug your application locally.*

8 Site Administration

In this section we give a brief overview of how to administrate a computing site.

8.1 Managing Batch Services

1. Switch to the g-Eclipse Operator Perspective. Therefore select **Window →Open Perspective →Other...**. In the upcoming dialog select **g-Eclipse (Operator)** and press **Ok**.
2. Right-click your Grid project and select **New →Batch Service**.
3. Give your batch service a name (*MyBatchService*) and press **Next**. Enter the hostname of the machine where the batch service is located and specify your user name for this machine.
4. Press **Finish**. Now you will be connected to the remote host. Therefore you'll be asked for your SSH passphrase. After entering the passphrase the Batch Editor will show up.
5. Within the Batch Editor you can see the computing element and its associated queues and worker nodes.
6. By selecting one of the elements in the Batch Editor the Properties View will show further information for these elements while the Batch Jobs View will show you further information about their currently associated batch jobs.
7. By right-clicking the elements you will get access to further actions corresponding to the element. From within the context menus you can create new queues, drain, stop or delete existing queues and enable or disable worker nodes. Of course you need to have administration rights to access one of these actions.

8.2 Remote Command Console

1. The Operator Perspective offers by default a command console for accessing remote hosts either by SSH or by glogin. This console is called the Terminal View.
2. If not already visible open the Terminal View by selecting **Window →Show View →Other...** and **g-Eclipse →Terminal**.
3. In the Terminal View press the New button on the top right.
4. Select either the SSH type or the glogin type for logging in. In the subsequent wizard page you have to specify type-specific login data. After doing so a terminal session to the remote host will be created.

A Additional Information

g-Eclipse Homepages
EU-Home: http://www.geclipse.eu/
Eclipse-Home: http://www.eclipse.org/geclipse/

Java Download
Internet: http://java.sun.com/javase/downloads/index.jsp

Eclipse Download
Internet: http://www.eclipse.org/downloads/

Eclipse Update Site
Internet: Hardcoded in the Eclipse Update Manager

g-Eclipse Update Site
Internet: `http://www.geclipse.eu/update/1.0milestones/`

g-Eclipse EPP Packages
Internet: http://www.geclipse.eu/index.php?id=downloads

g-Eclipse Subversion Coordinates
Connection String: `svn://dev.eclipse.org/svnroot/technology/eu.geclipse`

g-Eclipse CVS Coordinates
Host: `cvs.fzk.de`
Repository Path: `/data/cvs/fzk/geclipse`
User: anoncvs
Connection Type: extssh

Gilda Certificates

Gilda CA Certificate:	https://gilda.ct.infn.it/CA/mgt/getCA.php
Gilda Certificates:	gilda/BARCELONA*
Gilda Certificate Passwords	BARCELONA

Gilda VO Settings

Name:	geclipsetutor
Host:	dgrid-voms.fzk.de
Port:	15007
Host-DN:	/O=GermanGrid/OU=FZK/CN=host/dgrid-voms.fzk.de
Info system end-point:	ldap://iwrbdii.fzk.de:2170

Sample Data

VTK Data Files:	http://www.vtk.org/get-software.php

Batch System Access

Login Information:	geclipse020@dgrid-ce.fzk.de (only user rights)
SSH Key:	ssh/id_rsa