Hands-on with the P-GRADE/GEMLCA Portal

Part 1: GEMLCA components vs. Client-side binaries

1. **Login to the NGS Portal**

   1.1 Go to the URL:
   https://gngs-portal.cpc.wmin.ac.uk:8080/gridsphere/gridsphere
   1.2 Use your account and password to login (user**xx** / user**xx**)

2. **Download a short-term proxy credential from the MyProxy server into the Portal server. The proxy will be used by the workflow manager to access UK NGS resources.**

   2.1 Hit the "Certificates" tab
   2.2 Click on the "Download" button
   2.3 Submit the download form with the following data:
   Hostname: leave as default
   Port: 7512
   Login: bpuser**xx**
   Password: bpuser**xx**
   Lifetime: 10
   Description: <*optional*>
   2.4 Set your certificate to be used with the NGS Grid.

3. **Create a traffic simulation workflow combining standard and GEMLCA jobs**

   The aim of this exercise is to demonstrate the difference between standard and GEMLCA jobs. The first component of your workflow will be a standard job where you have to define the executable and parameter format. The rest of the workflow will be composed of GEMLCA jobs selected from the repository. The "manhattan" executable can be downloaded from the Summer School's Agenda page. A description of the different components of the traffic simulation workflow can be found here:
   http://portal.p-grade.hu/tutorials/induction/Traffic_simulation_workflow_description.htm

   3.1. Load default resources

   Hit the "Settings" tab, then the "Resources" button and load the default resources of the NGS Grid.

   3.2. Create a standard job using the "Manhattan" executable

   3.2.1 Create a new standard job in the workflow editor. Go into properties by right-clicking on the job and set job type as "Standard job". Define the job with the following parameters:
   Name: manhattan
   Job type: SEQ
   Job executable: <*path of "manhattan" executable*>
   Attributes: -r 10 -c 10 -w 150 -h 150 -C 2 -R 5 -n file.net -t file.trn
   Grid: NGS
   Resource: Any NGS resource

   3.2.2 Define a port for the job with the following parameters:
   Port Name: 0
   Type: out
   Internal File Name: file.net

   3.2.3 Define another port to the job with the following parameters:
   Port Name: 1
   Type: out

Internal File Name: file.trn

3.2.4 Save and submit your workflow. Go back to the browser, click "Refresh" then "Submit" on the "Workflow manager" panel. Monitor and visualize the progress of execution.

3.2.5 Download and unzip the result file.

3.3 Extend your workflow with 3 additional GEMLCA jobs

In this exerci you will utilize already published legacy codes in your workflow from the GEMLCA repository. The executables of these programs are already uploaded, you only have to specify actual parameter values.

3.3.1 Open your previous workflow in the workflow editor and create a new GEMLCA job. Go into properties by right-clicking on the job and set job type as "GEMLCA job". Double click on the job and go into job properties. Define a job with the following parameters:
   Name: Sim1
   Grid: NGS
   Resource: Any GEMLCA resource
   Legacy code: MadCity traffic simulator
   Parameters with default values and ports are created automatically. Keep default parameters for this job.

3.3.2 Define another job with the following parameters:
   Name: Sim2
   Grid: NGS
   Resource: Any GEMLCA resource
   Legacy code: MadCity traffic simulator
   Change the parameter value "number of cars per lane" from 14 to another value.

3.3.3 Define your last job with the following parameters:
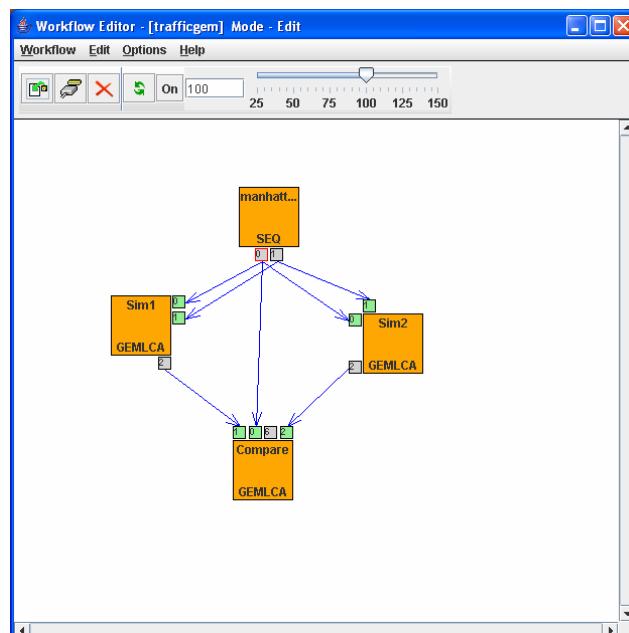   **Name:** Compare
   Grid: NGS
   Resource: Any GEMLCA resource
   Legacy code: Trace file compare
   Keep default parameter values.

3.3.4 Draw the workflow graph as follows:

3.3.5 Save the workflow, submit it, visualize execution, and download results.

Part 2: Workflows on EGEE; processing storage files

4. **Login to the GILDA P-GRADE Portal**

    1.1 Go to the URL:
    http://portal.p-grade.hu/gilda
    1.2 Use your account and password to login (user**xx** / user**xx**)

5. **Download a short-term proxy credential from the MyProxy server into the Portal server. This proxy will be used by the workflow manager to access GILDA resources.**

    2.1 Hit the "Certificates" tab
    2.2 Click on the "Download" button
    2.3 Submit the download form with the following data:
    Hostname: grid001.ct.infn.it
    Port: 7512
    Login: bpuser**xx**
    Password: bpuser**xx**
    Lifetime: 10
    Description: *<optional>*
    2.4 Set your certificate to be used with the **GILDA** Grid. (gilda_GLITE_BROKER)

6. **Define a workflow using a client-side binary: Matrix multiplication**
    Note: The "Matrix operations" program will be used as the core of the workflow. Please read the description of this program and download the binary and the two sample input files. (See http://portal.p-grade.hu/tutorials/induction/Matrix_operations_program_description.htm for details!)

    Open a new workflow (Workflow menu) and define a new job with the following parameters:
    ```
    Name:           <any>
    Job type:       SEQ
    Job executable: <path of the matrix_operations file>
    Attributes:     M V
    Grid:           gilda_GLITE_BROKER
    ```
    Define a port for the job with the following parameters:
    ```
    Port Name:      0
    Type:           In
    File type:      Local
    File:           <path of INPUT1 file>
    Internal File Name (case sensitive): INPUT1
    ```
    Define another port to the job with the following parameters:
    ```
    Port Name:      1
    Type:           In
    File type:      Local
    File:           <path of INPUT2 file>
    Internal File Name (case sensitive): INPUT2
    ```
    Define a third port to the job with the following parameters:
    ```
    Port Name:      2
    Type:           Out
    File type:      Local
    Internal File Name (case sensitive): OUTPUT
    File storage type: Permanent
    ```

    Save your workflow as Multiply*, go back to the browser, click "Refresh" then "Submit" on the "Workflow manager" panel.*

Monitor the execution from the workflow editor and from the browser
Download and unzip the result file

7. Save the result of matrix multiplication into a Grid file

Save the Multiply workflow as Multiply_remoteout *("Save as" in the "Workflow" menu)*
Modify port 2 *to a remote file:*
- Double click on port 2
- Set the "File type" parameter to **Remote**
- Set the "File" field to
  **lfn:/grid/gilda/budapestXY/result_matrix
  (substitute XY with your GILDA user number!)**
Open the Property window of the job and click on the "JLD Editor" button.
On the "Environment" tab set the following environment variables:
- **LCG_CATALOG_TYPE = lfc**
- **LFC_HOST = lfc-gilda.ct.infn.it**
- **LCG_GFAL_INFOSYS= grid004.ct.infn.it:2170**
Save your workflow
Submit your workflow and monitor its execution
Note that after the execution there is no result file to download!
Check the existence of the result file in a cmd line window using the lfc-ls command

8. Define a workflow which computes the following expression: $AB[*,0]^T * AB[*,1]$
(A and B represents the INPUT1 and INPUT2 input matrixes)
Hint: The "Matrix operations" program reads and produces files in the same format.
Add the matrix_operation program 4 times to the Multiply workflow. The new
workflow is shown in the Figure below.

Set the jobs to compute the following operations:
- Multip – $A^.B$ → parameter: M
- Column0 - $A^.B[*, 0]$ → parameter: C 0
- Column1 - $A^.B[*, 1]$ → parameter: C 1
- Transpose – $A^.B[*, 0]^T$ → parameter: T
- Multip.2 - $A^.B[*,0]^T * A^.B[*,1]$ → parameter: M
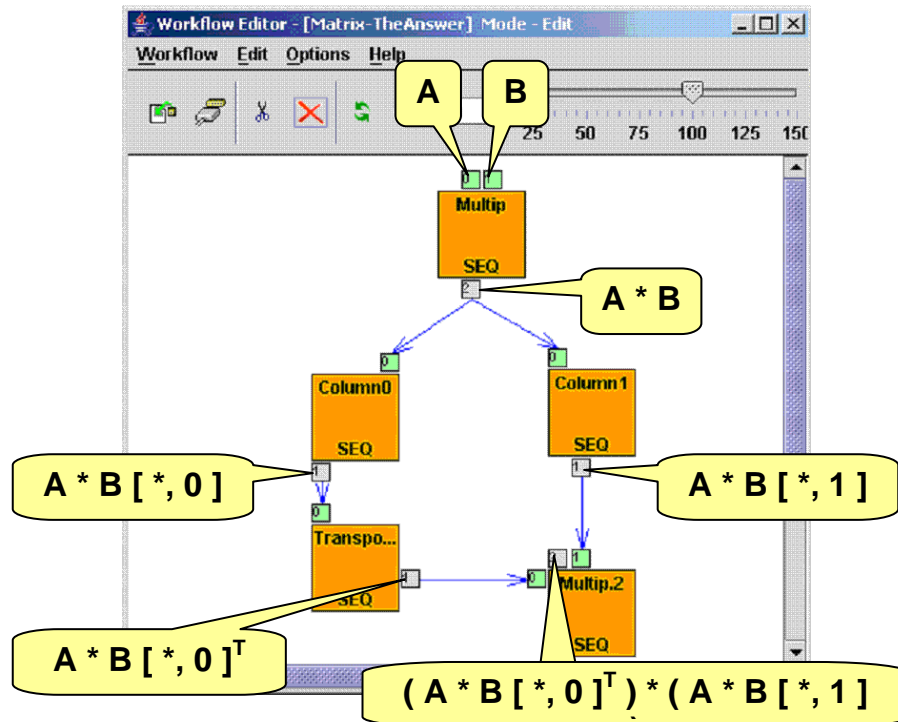
Figure 1. Matrix workflow to compute AB[*,0]^T * AB[*,1]

The full workflow can be downloaded from:
http://jfe.lpds.sztaki.hu/~sipos/p-grade/matrix-full/Matrix-full.tar
Do not forget to reallocate the jobs to the GILDA VO!