



# Globus Job Submission in DEISA

J. Laitinen, H. Heller, G. Pringle (ed.)

(c) 2004 - 2010 DEISA



DEISA is funded by the European Commission in FP7 under grant agreement RI-222919



# Table of contents

1 General.....	1
1.1 Overview of Globus Toolkit.....	1
1.2 Required certificates.....	2
1.3 Globus installations at DEISA.....	2
2 Job submission in DEISA.....	3
2.1 Job submission within DEISA network.....	3
2.2 Job submission from the Internet.....	3
3 Local resource management systems.....	5
3.1 Fork job.....	5
3.2 Job for batch scheduling system.....	5
4 Options for globusrun-ws.....	7
4.1 Re-direction of output.....	7
4.2 Fire and forget job submission.....	7
5 Globus job script.....	9
5.1 RSL schema.....	9
5.2 Extension tag.....	10
5.3 Job script with file transfer.....	10
5.4 MPI Globus job script.....	11



# 1 General

Globus WS-GRAM enables users to submit jobs to different machines by using the same job script for any of the DEISA machines. Naturally, the executable must be pre-compiled for each target architecture.

## 1.1 Overview of Globus Toolkit

The Globus Toolkit grid middleware makes the user's life easier by providing a uniform interface, hiding the idiosyncrasies of a wide variety of computing resources. Further, it provides a single sign-on mechanism: the user types the passphrase only once, even if many jobs are submitted to different resources. Besides DEISA, Globus is also being used world-wide by many flagship Grid projects such as TeraGrid.

In this chapter is presented instructions for job submission with Globus. There are separate DEISA User Guides for

- Data Transfer with GridFTP[1]
- Interactive Access to HPC Resources[2]
- MyProxy - Proxy Certificate Storage Service[3]
- certificate usage within DEISA[4]

There are basically two slightly different possibilities to use Globus: (i) from a DEISA computer, (ii) from your home computer.

(i) If you submit jobs from a DEISA computer, then the Globus environment is pre-configured and is ready to be used. Chose this mode of operation for first experiments or if you are new to Grid computing. In the following guide, we will focus on this mode of operation. However, this is not the normal Grid-like way to do things, especially if you also use resources outside of DEISA.

(ii) The more Grid-like way is to use Globus commands on your local machine and to submit jobs to DEISA and other Grids directly from there; however, you must set-up Globus on your local machine. Such instructions are beyond the scope of this guide. Please consult your local system administrator or look at the Globus web site[5] or at its European counterpart[6].

We will now describe in a step-by-step how to submit computational jobs (including MPI parallel production jobs) through WS-GRAM to DEISA HPC resources.

- 
1. <http://www.deisa.eu/usersupport/user-documentation/data-transfer-with-GridFTP>
  2. <http://www.deisa.eu/usersupport/user-documentation/interactive-access>
  3. <http://www.deisa.eu/usersupport/user-documentation/deisa-myproxy-service>
  4. <http://www.deisa.eu/usersupport/user-documentation/faq/CertificatesFAQ>
  5. <http://www.globus.org>
  6. <http://www.ige-project.eu/>

### 1.2 Required certificates

For the single sign-on processes to work, Globus commands require certificate files. There are two types of certificate files: your private X.509 certificate and certificate files of those Certificate Authorities (CAs) which you trust. If you are using Globus commands at DEISA sites or the GSISsh-Term tool provided by DEISA then the CA certificates are set up for you automatically. If you submit Globus commands from your local machine then please refer to the certificates guide within the DEISA FAQ on how to set-up the CA certificates on your machine at <http://www.deisa.eu/usersupport/user-documentation/faq/CertificatesFAQ>.

### 1.3 Globus installations at DEISA

The easiest way to start using Globus commands is to use DEISA's Java GSISsh-Term which permits users to login to DEISA HPC systems via the public Internet.

Currently, DEISA sites provide Globus 4.0 WS-GRAM service as an Optional Service, meaning that it may not be available at every site by default. If you need Globus WS-GRAM at a site where it is not available, please submit a request to DEISA[7].

---

7. <http://tts.deisa.eu/UserSupport>

## 2 Job submission in DEISA

### 2.1 Job submission within DEISA network

Several DEISA sites offer the Globus 4.0 job submission service through the DEISA private network. Further, LRZ operates a WS-GRAM door node to submit WS-GRAM jobs to DEISA sites from the public Internet.

To use command line Globus commands on DEISA machines, you have to load the globus modulefile:

```
module load deisa
module load globus
```

Users do not need to know a particular address and port number if they are using the Globus command directly on the command-line of a DEISA machine since users can use the `deisa_service` script (See DEISA FAQ[1]). For example, the command

```
globusrun-ws -submit -F `deisa_service -i -g bsc` -c /bin/date
```

runs the command `/bin/date` on BSC on the login node of MareNostrum. Notice that the DEISA machine at BSC has been referred to as simply 'bsc'. This is a machine abbreviation. Users can determine each machine abbreviation, along with some help text, by simply invoking the `deisa_service` command without any parameters.

### 2.2 Job submission from the Internet

Of all the DEISA sites, only LRZ provides WS-GRAM access from the public internet, where LRZ's address is `a01.hlr2.lrz-muenchen.de`, i.e. the following command runs `/bin/date` on the front end of the LRZ execution site

```
globusrun-ws -submit -F a01.hlr2.lrz-muenchen.de -Ft PBS -c /bin/date
```

NB, to use this host, users must register their IP address by submitting a request to DEISA[2].

There is a possibility to send jobs directly from your workstation, over the public Internet, to other DEISA systems via LRZ's gateway host. Again, if that is needed, please create a DEISA trouble ticket.

- 
1. <http://www.deisa.eu/usersupport/user-documentation/faq/CertificatesFAQ>
  2. <http://tts.deisa.eu/UserSupport>



## 3 Local resource management systems

### 3.1 Fork job

The switch `-g` runs jobs as a so-called Fork job, which means the job is to be run immediately, on the machine which runs the WS-GRAM service without going to the compute nodes via the batch queueing system.

Using Fork is recommended to run only those jobs which are needed to run on the front-end/login node. Since fork jobs are executed immediately, this method is well suited to debugging job scripts quickly, thus avoiding long batch queue waiting times. No production jobs may be submitted to the Fork scheduler! NB it is the **default** method, so `-Ft Fork` can also be omitted.

### 3.2 Job for batch scheduling system

Production jobs have to be submitted to the batch scheduling system of the target site. That is done using the `deisa_service` script with the `-b` switch. Recall that this script is available at DEISA machines and it will automatically set up the correct batch system identifier. The following example performs a job submission on a DEISA machine to Huygens at SARA:

```
globusrun-ws -submit -F `deisa_service -i -b sara` -c /bin/date
```



## 4 Options for globusrun-ws

### 4.1 Re-direction of output

To stream command output to the screen, use the `-s` switch, i.e.,

```
globusrun-ws -submit -F `deisa_service -i -g lrz` -s -c /bin/date
```

Without the `-s` flag, the previous command would not show the result of the `date` command. To write the standard output and the standard error to files, use the `-so` and `-se` flags, i.e.,

```
globusrun-ws -submit -F `deisa_service -i -g lrz` \  
-so stdout_file.txt -se stderr_file.txt -c /bin/date
```

### 4.2 Fire and forget job submission

The commands given above block the command line until the remote job finishes. For longer running jobs, it is better to use `-batch` switch to send the job to the target server and then the job submission command returns to the prompt. Thus there is no need for a continuous network connection and it is not necessary to keep a shell open.

The `-batch` switch does not mean that the job is sent to the batch scheduling system, which is specified with `-Ft` switch. This switch is added if you employ the `deisa_service` script.

Later, to access the job output, to query its state, or to kill the job, a reference for the job will be written to a file defined with the `-o` switch:

```
globusrun-ws -submit -batch -o epr.xml -S -F `deisa_service -i -b lrz` -f myscript.rsl
```

When you have obtained the EPR file `epr.xml`, `globusrun-ws` can be used to kill the job or to monitor its status:

```
globusrun-ws -kill -j epr.xml globusrun-ws -status -j epr.xml
```



# 5 Globus job script

## 5.1 RSL schema

For non-trivial jobs, a Globus job script is needed. It is written in RSL (Resource Specification Language), which is a Globus-specific XML schema. The simplest example of RSL script is following three line script. More examples can be found below.

```
<job>
  <executable>/bin/date</executable>
</job>
```

That script saved in a file `myjob.rsl` can be submitted with the command

```
globusrun-ws -submit -S `deisa_service -i -b lrz` -f myjob.rsl
```

Commonly used tags are:

- `<job>` : every RSL file begins with this tag
- `<executable>` : executable with absolute path to be run
- `<directory>` : default home directory is by default `$GLOBUS_USER_HOME` which is usually `$HOME`
- `<count>` : number of processes
- `<maxWallTime>` : length of the job in minutes
- `<jobType>` : use `mpi` for MPI jobs
- `<fileStageIn>` : Contains one or more `<transfer>` entries, which contain `<sourceUrl>` addresses, which needs to be in `gsiftp://` address syntax. Target is specified with `<destinationUrl>`. It can be `file:///path/target.txt` address i.e. to the directory where the file is needed on that host.
- `<fileStageOut>` : contains also `<sourceUrl>` and `<destinationUrl>`, but in this direction `sourceUrl` can be with `file:///` syntax and `<destinationUrl>` with `gsiftp://` syntax.
- To remove files such as input files, it is possible to specify files in `<fileCleanup>` tags. Each file is surrounded with `<file>` tag and referred with `file:///` syntax.

The only mandatory field inside `<job>` tags is `<executable>`, all other fields are optional. A full list of possible tags is available on the Globus web site[1].

The `globusrun-ws` switch `-f` specifies the location and name of the RSL file. Note that `-s` is needed if the script contains `<fileCleanup>` tags. Example job submission command:

```
globusrun-ws -submit -S `deisa_service -i -b lrz` -f myjob.rsl
```

You cannot use `$DEISA_HOME` and other DEISA environmental variables in Globus job script. However, in the executable script you can use the DEISA variables provided that you include `module load deisa` in it or in `<preamble>` as shown in next section.

---

1. [http://www.globus.org/toolkit/docs/4.0/execution/wsgram/schemas/gram\\_job\\_description.html](http://www.globus.org/toolkit/docs/4.0/execution/wsgram/schemas/gram_job_description.html)

## Globus job script

You have to use absolute paths for the executable. The default directory to copy files and to create output files (stdout and stderr) is user's home directory. It can be referred with `${GLOBUS_USER_HOME}`. You can see example usage from next RSL example below.

### 5.2 Extension tag

DEISA Globus installations support a special extension which, for example, can be used to load a module file or to set environment variables (indeed any UNIX command can be used here). It is a must if you want to use pre-installed parallel programs from the DEISA CPE jobs, as you have to load the respective module file (e.g., cpmd).

```
<extensions>
  <preample>module load deisa cpmd</preample>
</extensions>
```

### 5.3 Job script with file transfer

It is often the case that a compute job needs input files and produces output files. These files should be transferred automatically to/from the computing resource before and after the job. This can be done with the staging commands using GridFTP.

To find out the names and ports for GridFTP servers, you can call the command line, use `deisa_service -i -f <site>`. In the following example there is a file transfer to the execution site (input file) and back to the sending host. There is also a clean-up step of the files at target site. The following example transfers `text_input.txt` from LRZ's GridFTP server to the target system into user's home directory at the target site.

Absolute paths need to be used with `gsiftp://` addresses syntax.

```
<job>
  <executable>/bin/cat</executable>
  <directory>${GLOBUS_USER_HOME}</directory>
  <argument>text_input.txt</argument>
  <stdout>test.stdout</stdout>
  <stderr>test.stderr</stderr>
  <fileStageIn>
    <transfer>
      <sourceUrl>
```

## Globus job script

```
gsiftp://a01-deisa.hlrb2.lrz-muenchen.de:2813/home/deisa/lrz/lrz00001/lrz04jex/text_input.txt
  </sourceUrl>
  <destinationUrl>file:///${GLOBUS_USER_HOME}/text_input.txt</destinationUrl>
</transfer>
</fileStageIn>
<fileStageOut>
  <transfer>
    <sourceUrl>file:///${GLOBUS_USER_HOME}/test.stdout</sourceUrl>
    <destinationUrl>

gsiftp://a01-deisa.hlrb2.lrz-muenchen.de:2813/home/deisa/lrz/lrz00001/lrz04jex/test.stdout
  </destinationUrl>
</transfer>
<transfer>
  <sourceUrl>file:///${GLOBUS_USER_HOME}/test.stderr</sourceUrl>
  <destinationUrl>

gsiftp://a01-deisa.hlrb2.lrz-muenchen.de:2813/home/deisa/lrz/lrz00001/lrz04jex/test.stderr
  </destinationUrl>
</transfer>
</fileStageOut>
<fileCleanUp>
  <deletion>
    <file>file:///${GLOBUS_USER_HOME}/test.stdout</file>
  </deletion>
  <deletion>
    <file>file:///${GLOBUS_USER_HOME}/test.stderr</file>
  </deletion>
</fileCleanUp>
</job>
```

If the above script is saved to a file called `cat.rsl` an example job submission command could be:

```
globusrun-ws -submit -F `deisa_service -i -g sara` -S -f cat.rsl
```

## 5.4 MPI Globus job script

MPI jobs need the `<jobType>mpi</jobType>` tag. Also, the number of MPI processes has to be defined with the `<count>` tag. In the example below, we also specified the maximum amount of wall clock time in minutes.

## Globus job script

Because the MPI executable must have been compiled on the target machine before running it, this example assumes that the binary exists in the user's home directory.

```
<job>
  <executable>mpi_test</executable>
  <stdout>mpitest.stdout</stdout>
  <stderr>mpitest.stderr</stderr>
  <count>4</count>
  <maxWallTime>10</maxWallTime>
  <jobType>mpi</jobType>
  <fileStageOut>
    <transfer>
      <sourceUrl>file:///${GLOBUS_USER_HOME}/mpitest.stdout</sourceUrl>
      <destinationUrl>
        gsiftp://a01-deisa.hlrb2.lrz-muenchen.de:2813/home/deisa/lrz/lrz00001/lrz04jex/mpi_test.stdout
      </destinationUrl>
    </transfer>
    <transfer>
      <sourceUrl>file:///${GLOBUS_USER_HOME}/mpitest.stderr</sourceUrl>
      <destinationUrl>
        gsiftp://a01-deisa.hlrb2.lrz-muenchen.de:2813/home/deisa/lrz/lrz00001/lrz04jex/mpi_test.stderr
      </destinationUrl>
    </transfer>
  </fileStageOut>
  <fileCleanUp>
    <deletion>
      <file>file:///${GLOBUS_USER_HOME}/mpitest.stdout</file>
    </deletion>
    <deletion>
      <file>file:///${GLOBUS_USER_HOME}/mpitest.stderr</file>
    </deletion>
  </fileCleanUp>
</job>
```

Job submission example using fire & forget style:

```
globusrun-ws -submit -F `deisa_service -i -b lrz` -S -b -o mpijob.epr -f mpi.rsl
```

Later, one can query the status of the job:

```
globusrun-ws -status -j mpijob.epr
```

or, if there is need to kill it

```
globusrun-ws -kill -j mpijob.epr
```