



Enabling Grids for E-science

Training event:
Application development course
10.March 2005, IISAS Bratislava

Job submission in Grid testbed

Miroslav Dobrucký
Institute of Informatics SAS

www.eu-egee.org

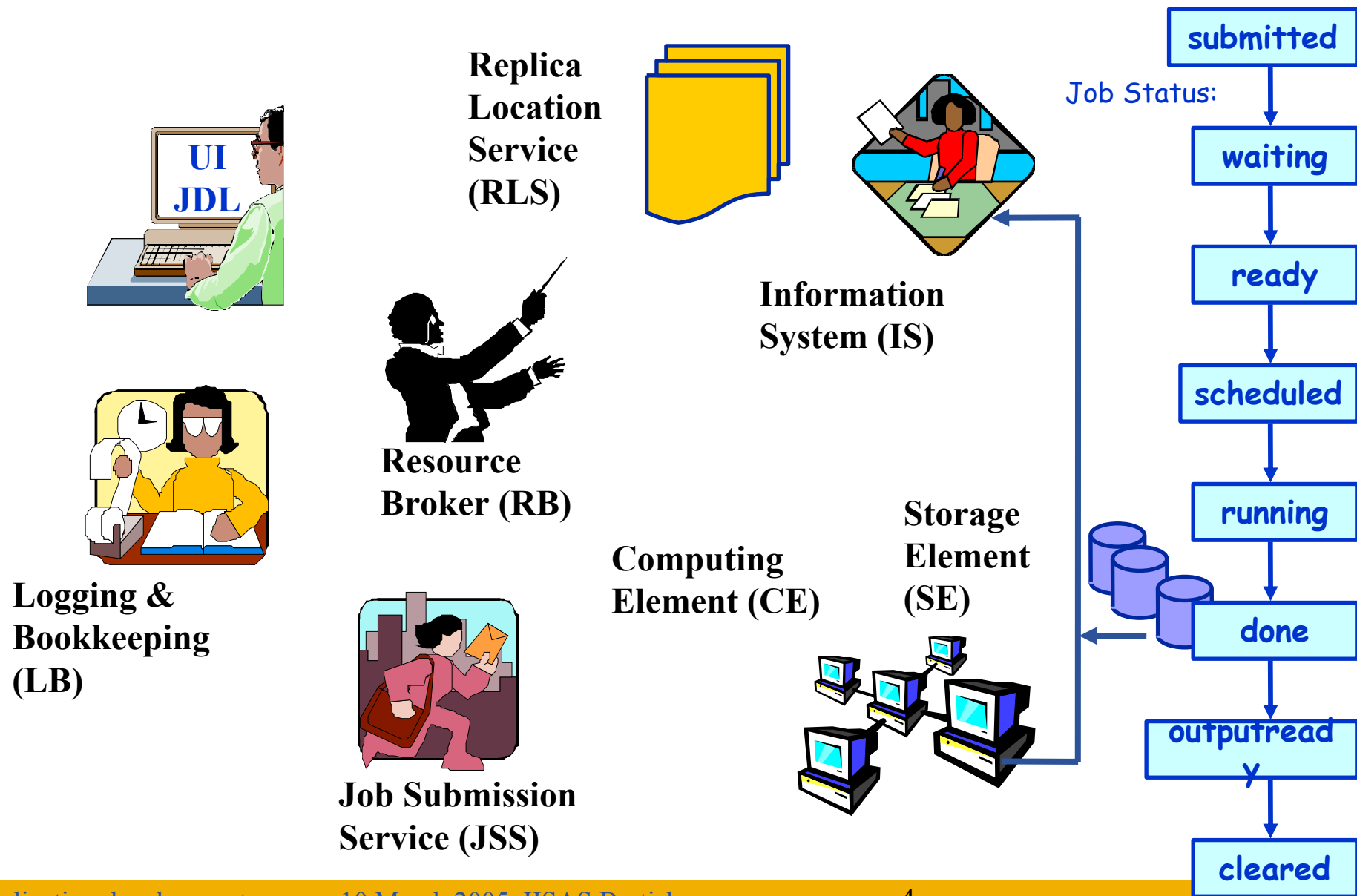


EGEE is a project funded by the European Union under contract INFSO-RI-508833

- **Grid middleware**
- **Job submission**
- **File transfer: GridFTP**
- **Application (DaveF)**
- **Other specifications in JDL**
- **Bypassing resource broker**
- **MyProxy: working through a portal**
- **Other useful commands**
- **Job types**



- **The matchmaker has the goal to find the best suitable CE where to execute the job**
 - optimising resource consumption
 - get the shortest job run-time
- **The logging and bookkeeping service takes trace of scheduled, running and finished jobs**
- **The replica manager knows all replicas of data files**
- ...



- User have to create a file describing the submitted job in **Job Description Language (JDL)**
 - What program (submitted into insecure world)
 - What data files (replicas?)
 - What requirements on OS, libraries, ...
- User submits jobs to **Resource Broker (RB)**

```
[miro@cluster2 miro]$ cat hostname.jdl
```

```
Executable = "/bin/hostname";  
StdOutput = "std.out";  
StdError = "std.err";  
OutputSandbox = {"std.out", "std.err"};
```

Before doing any Grid operation, user must create his proxy:

```
[miro@cluster2 miro]$ grid-proxy-init
```

```
Your identity: /C=SK/O=SlovakGrid/O=IISAS/CN=Miroslav Dobrucky
```

```
Enter GRID pass phrase for this identity:
```

```
Creating proxy ..... Done
```

```
Your proxy is valid until: Fri Nov 12 12:37:05 2004
```

edg-job-submit command

```
[miro@cluster2 miro]$ edg-job-submit hostname.jdl
```

```
Selected Virtual Organisation name (from UI conf file): cg
```

```
Connecting to host rb01.lip.pt, port 7772
```

```
Logging to host rb01.lip.pt, port 9002
```

```
*****
```

JOB SUBMIT OUTCOME

```
The job has been successfully submitted to the Network Server.
```

```
Use edg-job-status command to check job current status. Your job  
identifier (edg_jobId) is:
```

```
- https://rb01.lip.pt:9000/vrTuD4Tm-rQZvtdMma0a7w
```

```
*****
```

```
[miro@cluster2 miro]$ edg-job-submit hostname.jdl -o id_file
```

edg-job-status command

```
[miro@cluster2 miro]$ edg-job-status
https://rb01.lip.pt:9000/vrTuD4Tm-rQZvtdMma0a7w
```

```
*****
```

```
BOOKKEEPING INFORMATION:
```

```
Status info for the Job : https://rb01.lip.pt:9000/vrTuD4Tm-rQZvtdMma0a7w
```

```
Current Status:      Done (Success)
```

```
Exit code:           0
```

```
Status Reason:       Job terminated successfully
```

```
Destination:         ce.grid.cesga.es:2119/jobmanager-pbs-infinite
```

```
reached on:          Fri Nov 12 12:48:53 2004
```

```
*****
```

```
miro@cluster2 miro]$ edg-job-status -i id_file
```


edg-job-get-output command

```
[miro@cluster2 miro]$ edg-job-get-output https://rb01.lip.pt:9000/vrTuD4Tm-rQZvtdMma0a7w
```

```
Retrieving files from host: rb01.lip.pt ( for
https://rb01.lip.pt:9000/vrTuD4Tm-rQZvtdMma0a7w )
```

```
*****
```

```
JOB GET OUTPUT OUTCOME
```

```
Output sandbox files for the job:
```

```
- https://rb01.lip.pt:9000/vrTuD4Tm-rQZvtdMma0a7w
```

```
have been successfully retrieved and stored in the directory:
```

```
/tmp/miro_vrTuD4Tm-rQZvtdMma0a7w
```

```
*****
```

```
[miro@cluster2 miro]$ ls -lA /tmp/miro_vrTuD4Tm-rQZvtdMma0a7w
```

```
total 4
```

```
-rw-rw-r-- 1 miro      miro      0 Nov 12 13:53 std.err
```

```
-rw-rw-r-- 1 miro      miro     21 Nov 12 13:53 std.out
```

```
[miro@cluster2 miro]$ cat /tmp/miro_vrTuD4Tm-rQZvtdMma0a7w/std.out
grid01.grid.cesga.es
```

globus-url-copy command

```
[miro@cluster2 miro]$ globus-url-copy \  
  gsiftp://storage.ui.sav.sk/home/miro/grid-data/Malpaset.2dm \  
  file://$PWD/Malpaset.2dm
```

Other useful commands:

- edg-gridftp-ls
- edg-gridftp-exist
- edg-gridftp-mkdir
- edg-gridftp-rmdir
- edg-gridftp-rm
- edg-gridftp-rename

```
[miro@cluster2 testgrid]$ cat davef.jdl
```

```
Executable      = "davef.sh";
StdOutput       = "davef.out";
StdError        = "davef.err";
InputSandbox    = {"davef.sh"};
OutputSandbox   = {"davef.out", "davef.err"};
```

Potential error

```
[miro@cluster2 testgrid]$ cat davef.sh
```

```
globus-url-copy gsiftp://storage.ui.sav.sk/~grid-data/davef \
  file://$PWD/davef
chmod a+x davef
```

```
globus-url-copy gsiftp://storage.ui.sav.sk/home/miro/grid-
  data/Malpasset.prj file://$PWD/Malpasset.prj
```

```
./davef Malpasset.prj
```

```
globus-url-copy file://$PWD/Malpasset.flo \
  gsiftp://storage.ui.sav.sk/home/miro/grid-data/Malpasset.flo
```

Edg-job-list-match command

```
[miro@cluster2 testgrid]$ edg-job-list-match davef.jdl
```

```
Selected Virtual Organisation name (from UI conf file): cg
Connecting to host rb01.lip.pt, port 7772
```

```
*****
```

```
COMPUTING ELEMENT IDs LIST
```

```
The following CE(s) matching your job requirements have been found:
```

```
*CEId*
```

```
ce.grid.cesga.es:2119/jobmanager-pbs-infinite
ce.grid.cesga.es:2119/jobmanager-pbs-long
ce.grid.cesga.es:2119/jobmanager-pbs-short
ce001.grid.ucy.ac.cy:2119/jobmanager-pbs-infinite
ce001.grid.ucy.ac.cy:2119/jobmanager-pbs-long
ce001.grid.ucy.ac.cy:2119/jobmanager-pbs-short
cluster.ui.sav.sk:2119/jobmanager-pbs-workq
cms.fuw.edu.pl:2119/jobmanager-pbs-infinite
cms.fuw.edu.pl:2119/jobmanager-pbs-long
cms.fuw.edu.pl:2119/jobmanager-pbs-short
grid01.physics.auth.gr:2119/jobmanager-pbs-infinite
```

```
[miro@cluster2 testgrid]$ cat davef-client-cluster.jdl
```

```
Executable      = "davef_loader_script.sh";
Arguments       = "davef_param_sk.cfg";
StdOutput       = "davef.out";
StdError        = "davef.err";
InputSandbox    = {"davef_loader_script.sh", "davef_param_sk.cfg"};
OutputSandbox   = {"davef.out", "davef.err"};
JobType         = "MPICH";
NodeNumber      = 16;
Requirements    = other.GlueCEUniqueID ==
  "cluster.ui.sav.sk:2119/jobmanager-pbs-workq"
```

- **Environment variables**
- **Input data (PFN and/or LFN)**
 - What data access protocol
- **Output data (to which SE, registering)**
- **Rank (preference: freeCPU, response time, random)**
- **Requirements** (maxCPUtime, maxWallClock, *doména, testbed, estimated response time, freeCPU>x, totalCPU>x, ...)
- ...

globus-job-run command

```
[miro@cluster2 testgrid]$ globus-job-run cluster.ui.sav.sk /bin/ls
8queen_speedup
Desktop
bin
davef_param.cfg
davef_vis.cfg
encrypt.pl
flo2dh_310_src.zip
gr
gram_job_mgr_11873.log
gram_job_mgr_11904.log
```

- **Myproxy server is useful for:**
 - Very long jobs (ordinary proxy would expire)
 - Proxy is automatically renewed in the job runtime
 - Getting proxy on other machines than UI (typical for portals)

myproxy-init command

```
[miro@cluster2 testgrid]$ myproxy-init -s kentauros.rtd.algo.com.gr
Your identity: /C=SK/O=SlovakGrid/O=IISAS/CN=Miroslav Dobrucky
Enter GRID pass phrase for this identity:
Creating proxy ..... Done
Proxy Verify OK
Your proxy is valid until: Fri Nov 12 15:25:40 2004
Enter MyProxy pass phrase:
Verifying password - Enter MyProxy pass phrase:
A proxy valid for 168 hours
```

myproxy-get-delegation command

```
[miro@cluster miro]$ myproxy-get-delegation -s  
kentauros.rtd.algo.com.gr
```

```
Enter MyProxy pass phrase:
```

```
A proxy has been received for user miro in /tmp/x509up_u1001
```

- **grid-proxy-info**
- **grid-proxy-destroy**
- **edg-job-cancel**
- **edg-job-get-logging-info**
- **myproxy-info**
- **myproxy-destroy**
- **globus-job-submit**
- **globus-job-status**
- **globus-job-cancel**
- **globus-job-get-output**

- **Simple (sequential)**
- **MPICH (P4, G2)**
- **Interactive**
 - StdOut & StdErr on-line delivered
 - X-windows
 - Other (MD, VNC)
- **Checkpointable (in the future)**

In short: job submission

- **grid-proxy-init**
- **edg-job-submit job.jdl**
- **edg-job-status**
- **edg-job-get-output**

```
Executable    = "davef.sh";  
StdOutput     = "davef.out";  
StdError      = "davef.err";  
InputSandbox  = {"davef.sh"};  
OutputSandbox =  
{"davef.out","davef.err"};
```

Thank you.

<http://public.eu-egee.org>