



Job Management

DIRAC Project



Overview

- ▶ DIRAC JDL
- ▶ DIRAC Commands
- ▶ Tutorial Exercises
- ▶ What do you have learned?



Job Description Language (JDL)

- ▶ JDL is the language used to specify the characteristics and resources required by the job.
- ▶ For example:
 - ▶ Executable to run and its parameters.
 - ▶ Files to be moved to and from the Worker Node
 - ▶ Input files
 - ▶ Grid files
 - ▶ Job requirements.



Differences with gLite JDLs

- ▶ DIRAC is not using JDLs internally, it transform the user JDL to DIRAC internal job representation.
- ▶ Not all the standard gLite JDLs attributes are allowed in DIRAC.
 - ▶ In particular, there is no Requirements JDL attribute.
- ▶ Some attributes are interpreted as requirements
 - ▶ E.g. Site, CPUTime

- ▶ The JDL syntax consists of statements ended by a semicolon, like:

`attribute = value;`

- ▶ Literal strings (for values) are enclosed in double quotes.

`Site = "LCG.CPPM.fr";`



Example Simple JDL

```
JobName = "Simple_Job";  
Executable = "/bin/ls"; *  
Arguments = "-ltr";  
StdOutput = "StdOut"; *  
StdError = "StdErr"; *  
OutputSandbox = {"StdOut", "StdErr"}; *
```

* Mandatory attributes



Example Input and Output Sandbox

▶ Simple.jdl

```
Executable = "testJob.sh";  
StdOutput = "StdOut";  
StdError = "StdErr";  
InputSandbox = {"testJob.sh"};  
OutputSandbox = {"StdOut", "StdErr"};
```

▶ testJob.sh

```
#!/bin/bash  
/bin/hostname
```

- ▶ Sandbox files are stored in the DIRAC Workload Management Service and are only available for the time of life of the corresponding job
 - ▶ Volatile storage
 - ▶ Some grace period, typically 1 week, for keeping sandboxes to allow their retrieval
- ▶ Data are stored in the Grid Storage
 - ▶ Permanent storage
- ▶ The size of the Sandboxes is limited to ~10MB per job
 - ▶ Submitting a job with a large sandbox will result in an error
 - ▶ Large output sandboxes are stored in the Grid Storage
 - ▶ This is the user responsibility to clean up this storage



Example

Input and Output Data

▶ testJob.jdl

```
Executable = "testJob.sh";
StdOutput  = "StdOut";
StdError   = "StdErr";
InputSandbox = {"testJob.sh", "LFN:/
    vo.formation.idgrilles.fr/user/v/vhamar/test.txt"};
OutputSandbox = {"StdOut", "StdErr"};
OutputSE     = "M3PEC-disk";
```

▶ testJob.sh

```
#!/bin/bash
/bin/hostname
cat test.txt
```



DIRAC Commands

- ▶ In the directory *scripts* of your DIRAC installation there are a lot of commands that can be used for different purposes:
 - ▶ Job Management
 - ▶ Data Management
 - ▶ Proxy Management
 - ▶ Administration of the system



DIRAC Commands

DIRAC WMS

dirac-wms-job-attributes
dirac-wms-job-delete
dirac-wms-job-get-input
dirac-wms-job-get-jdl
dirac-wms-job-get-output
dirac-wms-job-get-output-data
dirac-wms-job-kill
dirac-wms-job-logging-info
dirac-wms-job-parameters
dirac-wms-job-peek
dirac-wms-job-reschedule
dirac-wms-job-status
dirac-wms-job-submit
dirac-wms-jobs-select-output-search
dirac-wms-select-jobs

DIRAC ADMIN

dirac-admin-add-group
dirac-admin-add-host
dirac-admin-add-site
dirac-admin-add-user
dirac-admin-allow-catalog
dirac-admin-allow-se
dirac-admin-allow-site
dirac-admin-ban-catalog
dirac-admin-ban-se
dirac-admin-ban-site
dirac-admin-bdii-ce
dirac-admin-bdii-site
dirac-admin-ce-info
dirac-admin-delete-user
dirac-admin-externals-versions



DIRAC Commands

DIRAC DMS

dirac-dms-catalog-metadata
dirac-dms-change-replica-status
dirac-dms-check-file-integrity
dirac-dms-clean-directory
dirac-dms-data-size
dirac-dms-filecatalog-cli
dirac-dms-fts-monitor
dirac-dms-fts-submit
dirac-dms-get-file
dirac-dms-get-storage-directories
dirac-dms-lfn-accessURL
dirac-dms-lfn-logging-info
dirac-dms-lfn-metadata
dirac-dms-lfn-replicas
dirac-dms-pfn-accessURL
dirac-dms-pfn-metadata
dirac-dms-remove-catalog-files

DIRAC PROXY

dirac-proxy-get-uploaded-info
dirac-proxy-info
dirac-proxy-init
dirac-proxy-upload

DIRAC UTILITIES

dirac-platform
dirac-rss-reassign-token
dirac-rss-renew-token
dirac-service



DIRAC Commands: Sending Jobs

- ▶ The command used to submit jobs using DIRAC is:

```
dirac-wms-job-submit <JDL>
```

```
$ dirac-wms-job-submit Simple.jdl
```

```
JobID = 11
```



Job ID



Dirac Commands: Job Status

- ▶ The command used to get the job status is:

```
dirac-wms-job-status <Job_ID>
```

```
$ dirac-wms-job-status 11
```

```
JobID=11 Status=Waiting;
```

```
MinorStatus=Pilot Agent Submission;
```

```
Site=ANY;
```



Dirac Commands: Job Retrieval

- ▶ To retrieve job outputs use the command:

```
dirac-wms-job-get-output [ --dir output_dir] <Job_ID> [<Job_ID>]
```

```
$ dirac-wms-job-get-output --dir this_dir 470  
469 468
```

```
$ ls -la this_dir/
```

```
total 20
```

```
drwxr-xr-x 5 hamar dirac 4096 Oct 26 19:02 .  
drwxr-xr-x 3 hamar dirac 4096 Oct 26 19:02 ..  
drwxr-xr-x 2 hamar dirac 4096 Oct 26 19:02 468  
drwxr-xr-x 2 hamar dirac 4096 Oct 26 19:02 469  
drwxr-xr-x 2 hamar dirac 4096 Oct 26 19:02 470
```



User exercises

- ▶ Exercises can be found at:

[http://marela.in2p3.fr:9200/dirac/wiki/DIRAC/Tutorials/
JobManagement](http://marela.in2p3.fr:9200/dirac/wiki/DIRAC/Tutorials/JobManagement)

Note: All the files than you need are available as attachment of the wiki page!!!



What we have learned?

- ▶ Write JDLs
- ▶ Differences between Input Sandbox, Output Sandbox and Input Data and Output Data
- ▶ DIRAC Commands to submit, get status, and retrieve output.

