The GridFactory web services
- technical specifications

Frederik Orellana
Niels Bohr Institute
University of Copenhagen

1 Job directives

The file “job” in the job directory can either be:

- a single directive pointing to the executable job script (containing the in-lined job directives)

```
#GRIDFACTORY -x [executable script]
```

- the executable job script itself – containing the in-lined directives

```
#GRIDFACTORY -n [job name]
#GRIDFACTORY -i [input file 1] [input file 2] ...
#GRIDFACTORY -e [executable 1] [executable 2] ...
#GRIDFACTORY -t [running time in seconds]
#GRIDFACTORY -m [required memory in megabytes]
#GRIDFACTORY -z [-1|0|1] (virtualization indifferent|no|yes)
#GRIDFACTORY -o [output files]
#GRIDFACTORY -r [runtime environment 1] [runtime environment 2] ..
#GRIDFACTORY -y [operating system]
#GRIDFACTORY -s [distinguished name identifying the submitter]
#GRIDFACTORY -v [allowed virtual organization 1]
                [allowed virtual organization 2]
#GRIDFACTORY -u [unique identifier of the form
                 https://[my.server]/gridfactory/jobs/[identifier]]
```

List-arguments, like the list of input files, can be given either like a space-separated list of strings (not containing spaces) or by repeated use of the directive.

2 Database schemas

Field	Type	Description
identifier	varchar(255)	Unique ID.
name	varchar(255)	Mnemonic name.
csStatus	varchar(255)	Execution status.
userInfo	varchar(255)	DN of the submitter.
inputFileURLs	text	URLs of input files.
outFileMapping	text	Output file names and destination URLs in the form: file1 https://server/dir/file1 file2 https://server/dir/file2 ...
providerInfo	varchar(255)	DN of the worker node.
created	datetime	Creation time stamp.
lastModified	datetime	Modification time stamp.
outTmp	varchar(255)	Internal use: path to stdout file on the worker node.
errTmp	varchar(255)	Internal use: path to stderr file on the worker node.
jobID	varchar(255)	Internal use: process ID on the worker node.
metaData	text	Internal use.
host	varchar(255)	IP name or address of the worker node.
runningSeconds	int(16)	Requested running time in seconds on the worker node.
ramMb	int(16)	Requested memory in MB.
executable	varchar(255)	Name of executable file.
executables	text	Names of files to be set executable.
opSys	varchar(255)	Requested operating system.
runtimeEnvironments	text	Space-separated list of requested software packages.
allowedVOs	text	Space-separated list of virtual organizations or DNs that are allowed to run the job.
virtualize	tinyint(4)	1: Demand that the job be run inside a virtual machine, 0: Demand that the job not be run inside a virtual machine, -1: Indifferent.
stdoutDest	varchar(255)	Final destination of stdout file.
stderrDest	varchar(255)	Final destination of stderr file.

Table 1: Schema of the job definition table.

Field	Type	Description
csStatusHistory	text	Newline-separated list of lines of the form “[status] [date]”.

Table 2: Additional field of the schema of the job history table as compared to that of the job definition table.

Field	Type	Description
identifier	varchar(255)	Hardware address of the first network card found.
host	varchar(255)	IP name or address.
maxJobs	int(16)	Maximum number of concurrently running jobs.
allowedVOs	text	Space-separated list of virtual organizations or DNs, the jobs of whom will be picked up.
virtualize	tinyint(4)	1: Jobs will be run inside a virtual machine, 0: Jobs will not be run inside a virtual machine, -1: Indifferent.
hypervisors	text	Space-separated list of installed hypervisors.
maxRamMbPerJob	int(16)	Max RAM (in MB) each job is allowed to consume.
inPorts	varchar(255)	Space-separated list of open inbound ports.
outPorts	varchar(255)	Space-separated list of open outbound ports.
providerInfo	varchar(255)	DN of the active certificate.
created	datetime	Creation time stamp.
lastModified	datetime	Modification time stamp.

Table 3: Schema of the node information table.

3 Mime types

Mimetype	Description
gridfactory/jobrecords+text	A list of job records in the form of an SQL response (with newline as record separator and tab as field separator)
gridfactory/jobrecord+text	A job newline-separated job record with lines of the form [key]:[value]
gridfactory/noderecords+text	A list of job records in the form of an SQL response (with newline as record separator and tab as field separator)
gridfactory/noderecord+text	A job newline-separated job record with lines of the form [key]:[value]

Table 4: Mime types for request bodies and response documents.

A job record is represented by a tab-separated line of values corresponding to the fields of table 1 plus a value corresponding to an extra field, 'dbUrl', added by mod_gridfactory. This last field, 'dbUrl', is the URL representing a single job as seen from the worker node, i.e. the URL on the server where information on a specific job can be obtained. Such a job may have been submitted directly to the server or the server may have pulled it from another server. Thus 'dbUrl' is not to be confused with 'identifier', which is the URL to which the job was originally submitted; even in the first case, depending on the set-up, they may or may not agree.

4 Web service interface

4.1 Submitting a job and uploading input files

The string [identifier] is chosen by the submitter. The submission utility that comes with GridFactory generates a time-based UUID [UUID], but in principle, any string can be used[1].

URL	https://[my.server]/db/jobs/[identifier]
Method	MKCOL
Returns	201 Created
	401 Access Denied
	403 Forbidden
	405 Method Not Allowed

Table 5: Creating a job directory.

Example:

```
MKCOL http://lx08/db/jobs/3b16dcdd-2d5f-11dd-80f2-c3b981785945
- - - - - - - - - - - - - - - - - - - - - - -
201 Created
```

URL	https://[my.server]/db/jobs/[identifier]/job
Method	PUT
Returns	201 Created
	200 OK
	401 Access Denied
	403 Forbidden

[1]Uniqueness of the full URL is guaranteed by the fact that two different directories cannot have the same path.

	405 Method Not Allowed

Table 6: Uploading a job file.

Example:

PUT http://lx08/db/jobs/3b16dcdd-2d5f-11dd-80f2-c3b981785945/job

- -

201 Created

4.2 Getting job definition records

URL	https://[my.server]/db/jobs/	
Method	GET	
Query string	csStatus=	Filter by status
	userInfo=	Filter by DN of submitter
	providerInfo=	Filter by DN of provider
	start=	The number of the first job to return - 0 is the number of the first record
	end=	The number of the last record to return
Returns	200 OK & TEXT (gridfactory/jobrecords+text)	
	401 Unauthorized	
	404 Not Found	

Table 7: Getting multiple job definition records.

Example:

GET https://gridfactory.nbi.dk/gridfactory/db/jobs/?status=ready&start=0&end=1

- -

200 OK

identifier name csStatus userInfo inputFileURLs outFileMapping providerInfo stdoutDest stderrDest created lastModified outTmp errTmp jobID metaData host runningSeconds ramMb executable executables opSys runtimeEnvironments allowedVOs virtualize dbUrl

https://orellana.nbi.dk/gridfactory/jobs/3a86aacc-2d5f-11dd-80f2-c3b981785945 job ready https://gridfactory.nbi.dk/gridfactory/jobs/3a86aacc-2d5f-11dd-80f2-c3b981785945/job https://gridfactory.nbi.dk/gridfactory/jobs/3a86aacc-2d5f-11dd-80f2-c3b981785945/stdout https://gridfactory.nbi.dk/gridfactory/jobs/3a86aacc-2d5f-11dd-80f2-c3b981785945/stderr 2008-05-29 11:11:32 2008-05-29 11:17:23

-1 -1 job 0 https://lx08/db/jobs/3a86aacc-2d5f-11dd-80f2-c3b981785945

https://gridfactory.nbi.dk/gridfactory/jobs/3b16dcdd-2d5f-11dd-80f2-c3b981785945 job ready https://gridfactory.nbi.dk/gridfactory/jobs/3b16dcdd-2d5f-11dd-80f2-c3b981785945/job https://gridfactory.nbi.dk/gridfactory/jobs/3b16dcdd-2d5f-11dd-80f2-c3b981785945/stdout https://gridfactory.nbi.dk/gridfactory/jobs/3b16dcdd-2d5f-11dd-80f2-c3b981785945/stderr2008-05-29 11:18:10 2008-05-29 11:24:56
-1 -1 job 0 https://lx08/db/jobs/3a86aacc-2d5f-11dd-80f2-c3b981785945

4.3 Getting a specific job definition record

URL	https://[my.server]/db/jobs/[identifier]
Method	GET
Returns	200 OK & TEXT (gridfactory/jobrecord+text)
	401 Unauthorized
	404 Not Found

Table 8: Getting a single job definition record.

Example:

GET http://lx08/db/jobs/3b16dcdd-2d5f-11dd-80f2-c3b981785945

- -

200 OK

identifier: https://orellana.nbi.dk/gridfactory/jobs/3b16dcdd-2d5f-11dd-80f2-c3b981785945

name: job

csStatus: ready

userInfo:

inputFileURLs: https://gridfactory.nbi.dk/gridfactory/jobs/3b16dcdd-2d5f-11dd-80f2-c3b981785945/job

outFileMapping:

providerInfo:

stdoutDest: https://gridfactory.nbi.dk/gridfactory/jobs/3b16dcdd-2d5f-11dd-80f2-c3b981785945/stdout

stderrDest: https://gridfactory.nbi.dk/gridfactory/jobs/3b16dcdd-2d5f-11dd-80f2-c3b981785945/stderr

created: 2008-05-29 11:11:32

lastModified: 2008-05-29 11:11:32

outTmp:

errTmp:

```
jobID:
metaData:
host:
runningSeconds:-1
ramMb:-1
executable:job
executables:
opSys:
runtimeEnvironments:
allowedVOs:
virtualize:0
```

Notice that the identifier contains a host name different from the one of the host running the service. This indicates that the host running the service has pulled the job from somewhere else. The host name in the identifier is the name of the host to which job was originally submitted.

4.4 Modifying the status of a job

To modify an existing job, we'll PUT a text representation of a job to the URL of our job.

URL	https://[my.server]/db/jobs/[identifier]
Method	PUT
Request body	TEXT (gridfactory/jobrecord+text)
Returns	201 Created & Location
	401 Unauthorized
	404 Not Found
	415 Unsupported Media Type

Table 9: Modifying a single job definition record.

Since we don't want people to create new job records using PUT, only update existing job records, we'll return a '404 Not' Found error if the resource doesn't already exist.

Example:

```
PUT http://lx08/db/jobs/3b16dcdd-2d5f-11dd-80f2-c3b981785945

csStatus: ready
- - - - - - - - - - - - - - - - - - - - - - -
201 Created
```

Notice all values of a job record can be changed, except for those corresponding to the keys 'created' and 'lastModified'. These are set by the server. The value of 'lastModified' is changed on each PUT, regardless of whether the job record is actually changed.

4.5 Downloading output files

URL	https://[my.server]/db/jobs/[identifier]/[file]
Method	GET
Returns	200 OK & BODY
	401 Unauthorized
	404 Not Found

Table 10: Getting a single job output file.

Example:

GET http://lx08/db/jobs/3b16dcdd-2d5f-11dd-80f2-c3b981785945/output_file_1.data

\- -

200 OK

File body

4.6 Getting job history records

URL	https://[my.server]/db/history/	
Method	GET	
Query string	* created=	Filter by created. Format: “YYYYMMDDHHMMSS”. This string may be prepended by “<” or “>”
	* lastModified=	Filter by lastModified
	csStatus=	Filter by status
	userInfo=	Filter by DN of submitter
	providerInfo=	Filter by DN of provider
	start=	The number of the first job to return - 0 is the number of the first record
	end=	The number of the last record to return
Returns	200 OK & TEXT (gridfactory/jobrecords+text)	

	401 Unauthorized
	404 Not Found

Table 11: Getting multiple job history records. (*) not yet implemented

Example:

```
GET https://gridfactory.nbi.dk/gridfactory/db/history/?
status=done&start=0&end=1&created=>29052008100000&created=<29052008160000
- - - - - - - - - - - - - - - - - - - - - - - -
200 OK


...
```

4.7 Getting a specific job history record

URL	https://[my.server]/db/history/[identifier]
Method	GET
Returns	200 OK & TEXT (gridfactory/jobrecord+text)
	401 Unauthorized
	404 Not Found

Table 12: Getting a single job history record.

Example:

```
GET http://lx08/db/history/3b16dcdd-2d5f-11dd-80f2-c3b981785945
- - - - - - - - - - - - - - - - - - - - - - - -
200 OK


...
```

4.8 Getting node information records

URL	https://[my.server]/db/nodes/	
Method	GET	
Query string	maxJobs=	Filter by max jobs - may be prepended by “<” or “>”

	maxRamMbPerJob=	Filter by memory - may be prepended by “<” or “>”
	maxRunningSecondsPerJob=	Filter by memory - may be prepended by “<” or “>”
	providerInfo=	Filter by DN of provider
	start=	The number of the first record to return - 0 is the number of the first record
	end=	The number of the last record to return
Returns	200 OK & TEXT (gridfactory/noderecords+text)	
	401 Unauthorized	
	404 Not Found	

Table 13: Getting multiple node information records.

Example:

GET https://gridfactory.nbi.dk/gridfactory/db/nodes/?start=0&end=1

- -

200 OK

identifier host maxJobs allowedVOs virtualize hypervisors maxRamMbPerJob maxRunningSecondsPerJob opSys providerInfo created lastModified dbUrl

https://orellana.nbi.dk/gridfactory/nodes/3a86aacc-2d5f-11dd-80f2-c3b981785946 gridworker01 2 https://gridfactory.nbi.dk/vos/atlas.txt 0 VM/KQEMU, VMVirtualBox 1024 -1 Linux /C=DK/ST=Aarhus/L=Aarhus/O=CABO/CN=test_user 2008-05-29 10:07:28 2008-05-29 11:19:12 https://orellana.nbi.dk/db/nodes/3a86aacc-2d5f-11dd-80f2-c3b981785946

https://orellana.nbi.dk/gridfactory/nodes/3a86aacc-2d5f-11dd-80f2-c3b981785947 gridworker02 2 https://gridfactory.nbi.dk/vos/atlas.txt 0 VM/KQEMU, VMVirtualBox 1024 -1 Linux /C=DK/ST=Aarhus/L=Aarhus/O=CABO/CN=test_user 2008-05-29 10:07:33 2008-05-29 11:19:27 https://orellana.nbi.dk/db/nodes/3a86aacc-2d5f-11dd-80f2-c3b981785947

4.9 Getting a specific node information record

URL	https://[my.server]/db/nodes/[identifier]
Method	GET
Returns	200 OK & TEXT (gridfactory/noderecord+text)

	401 Unauthorized
	404 Not Found

Table 14: Getting a single node information record.

Example:

GET http://lx08/db/nodes/3b16dcdd-2d5f-11dd-80f2-c3b981785946

- -

200 OK

identifier: 08-00-27-44-BD-FA

host: gridworker01

maxJobs: 2

opSys: Linux

virtualize: 0

maxRamMbPerJob: 1024

providerInfo: /C=DK/ST=Aarhus/L=Aarhus/O=CABO/CN=test_user

created: 2008-05-29 10:07:28

lastModified: 2008-05-29 11:19:12

Bibliography

REST, RESTful web services are described in e.g.
http://www.peej.co.uk/articles/restfully-delicious.html

UUID, International Telecommunication Union, “Information technology - Open Systems Interconnection - Procedures for the operation of OSI Registration Authorities: Generation and registration of Universally Unique Identifiers (UUIDs) and their use as ASN.1 Object Identifier components “, http://www.itu.int/ITU-T/studygroups/com17/oid.html