**eGee**

Enabling Grids for E-sciencE

# Job Workflow

*Emidio Giorgio*
*INFN*
*NA4 Generic Applications Meeting*
*10 January 2006*

**www.eu-egee.org**

Information Society

- **Job workflow overview**

- **Features available with WM proxy**

- **Examples**
  - DAG job
  - Job collection
  - Parametric jobs

**Enabling Grids for E-sciencE**

- **DAG job is a set of jobs where the input, output, or execution of one or more jobs depends on one or more other ones**

- **Dependencies are represented through Directed Acyclic Graphs, where the nodes are graphs, and the edges identify the dependencies**

**eGee**

**Enabling Grids for E-sciencE**
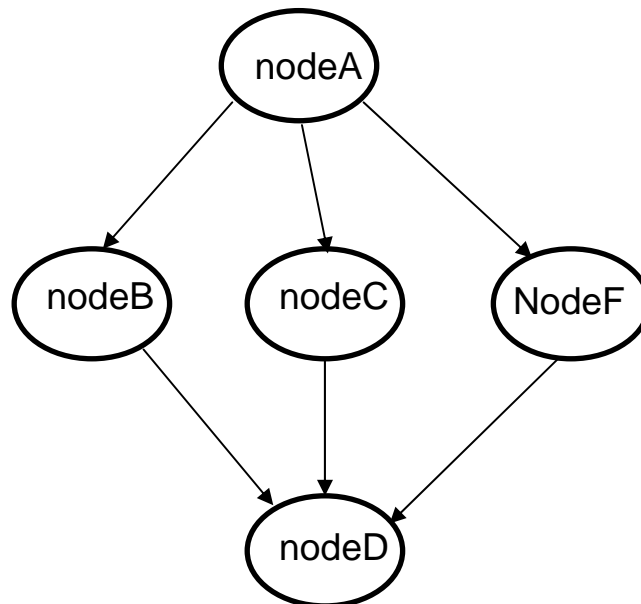
```
[
  type = "dag";
  max_nodes_running = 4;
  nodes = [
    nodeA = [
      file ="nodes/nodeA.jdl" ;
    ];
    nodeB = [
      file ="nodes/nodeB.jdl" ;
    ];
    nodeC = [
      file ="nodes/nodeC.jdl" ;
    ];
    nodeF = [
      file ="nodes/nodeF.jdl";
    ];
    dependencies = {
      {nodeA, nodeB},
      {nodeA, nodeC},{nodeA, nodeF},
      { {nodeB,nodeC,nodeF}, nodeD }
    }
  ];
;
]
```

Node description could be done also here, instead of using separate file

- **In order to submit job with WMProxy, it's mandatatory credentials delegation**

```
glite-wms-job-delegate-proxy -d del_ID_01
```

- **The submission/monitoring commands are slightly different, but the most of the "old" options are supported**

```
glite-wms-job-submit -d del_ID_01 collection.jdl

glite-wms-job-output \
https://glite-rb.ct.infn.it:9000/LHIIGaCVdl7Olm
sz0jpI_g
```

**Enabling Grids for E-sciencE**

- **Job collection is a set of independent jobs that user wants to submit and monitor as a single request**
- **Jobs of a collection are submitted as DAG nodes, without dependencies**
- **JDL is a list of classad, which describes the subjobs**

```
[
        Type = "collection";
        VirtualOrganisation = "gilda";
        nodes = {
                [  <job descr 1 >],
                [  <job descr 2 >],
                 …
                };
]
```

# Job collection examples

```
[
 Type = "collection";
 InputSandbox = {"start_hostname.sh"};
 RetryCount = 0;
 nodes={ [
        Executable = "/bin/sh";
        StdOutput = "host.out";
        StdError = "host.err";
        InputSandbox = root.InputSandbox;
        OutputSandbox = {"host.err","host.out"};
  OutputSandboxURI={"gsiftp://glite-rb.ct.infn.it:2811/tmp/host.out",
   "host.err"};
        Arguments = "start_hostname.sh";
   ],[
        Executable = "/bin/sh";
        StdOutput = "test.out";
        StdError = "test.err";
 InputSandbox={"starter.sh","gsiftp://glite-
   rb.ct.infn.it:2811/tmp/t01.txt"};
        OutputSandbox = {"test.err","test.out"};
        Arguments = "starter.sh";
    ],[
      file = "hostname.jdl";
   ] };
]
```

**eGee**

Enabling Grids for E-sciencE

- **A parametric job is a job where one or more of its attributes are parametrized**

- **Value of attributes vary according to parameter**

```
[
        JobType = "Parametric";
        Executable = "/bin/echo";
          Arguments = "PARAM";
        #InputSandbox = "input_PARAM_.txt";
        StdOutput = "myoutput_PARAM_.txt";
        StdError = "myerror_PARAM_.txt";
        Parameters = 2500;
        ParameterStep = 100;
        ParameterStart = 1000;
          OutputSandbox  =
{"myoutput_PARAM_.txt"};
```

- **Job monitoring / managing is always done through an unique jobID, as if the job was single (see submission of collection**

**egee**

Enabling Grids for E-sciencE

- **Parameter can be also a list of string**
- **InputSandbox (if present) has to be coherent with parameters**

```
[ui-test] /home/giorgio/param > cat param2.jdl
[
        JobType = "Parametric";
        Executable = "/bin/cat";
         Arguments = "input_PARAM_.txt";
        InputSandbox = "input_PARAM_.txt";
        StdOutput = "myoutput_PARAM_.txt";
        StdError = "myerror_PARAM_.txt";
        Parameters = {earth,moon,mars};
        OutputSandbox  = {"myoutput_PARAM_.txt"};
]
[ui-test] /home/giorgio/param > ls
inputEARTH.txt   inputMARS.txt   inputMOON.txt
param2.jdl
```

- **JDL attributes specification for WM proxy**
  - `https://edms.cern.ch/document/590869/1`

- **WMProxy quickstart**
  - `http://egee-jra1-wm.mi.infn.it/egee-jra1-wm/wmproxy_client_quickstart.shtml`

- **WMS user guides**
  - https://edms.cern.ch/document/572489/1