Grid Computing

Time: 2:00h

In the mulitple choice questions, there may be more than one correct alternative. Explain your choices.

1. What is the function of the JDL below?

```
[
   JobType = "Parametric";
   Executable = "example.sh";
   Arguments = "2ptl_PARAM_.tar.gz '' _PARAM_";

   StdOutput = "std_PARAM_.out";
   StdError = "std_PARAM_.err";

   InputSandbox = {"example.sh","2ptl_PARAM_.tar.gz"};
   OutputSandbox = {"std_PARAM_.out","std_PARAM_.err",
                    "aa2ptl000_PARAM_.pdb","2ptl_idl_PARAM_.pdb",
                    "timing","aa2ptl000_PARAM_.png","2ptl_idl_PARAM_.png",
                    "aa2ptl000_PARAM_.wrl","2ptl_idl_PARAM_.wrl"};

   Parameters = 3;
   ParameterStep = 1;
   ParameterStart = 1;

   Requirements =
    Member("RASTER3D",other.GlueHostApplicationSoftwareRunTimeEnvironment)
    &&
    Member("MOLSCRIPT-1.0.2",
           other.GlueHostApplicationSoftwareRunTimeEnvironment);
]
```

2. What is the difference between a "web service" and a "grid service"? (Enumerate advantages and disadvantages of grid services)

3. What is High Throughput Computing (HTC)?

   a. it is the same as High Performance Computing (HPC)

   b. efficiency to execute each application

   c. efficiency to execute all applications

4. What is a meta-scheduler?

a. a hierarchical scheduler

b. a resource broker

c. an orchestrator of resource managers

5. Describe briefly the main layers of the OGSA architecture.

6. How can grid applications be classified?

7. How can grid infrastructures be classified?

8. The gLite middleware is based on the OGSA. Are there any alternative architectural models for the middleware and its mode of utilization? Give examples of middlewares that do not follow the same model as gLite.

9. Given the description below, what does Condor do with this application? Is there an equivalent form of executing the same application in gLite? What should be the equivalent JDL?

```
Universe   = standard
Executable = my_job
Log        = my_job.log
Arguments  = -arg1 -arg2
InitialDir = run_$(Process)
Requirements = Memory >= 256 && Disk > 10000
Rank = (KFLOPS*10000) + Memory
Queue 600
```

10. What is the importance of the scheduling strategy in grid environments?

a. more efficient execution of each application

b. implementation of fault tolerance mechanisms

c. efficient execution of all applications

d. all above

11. Assume you have a bag-of-tasks application that runs on a cluster and you want to run this same application in a grid environment. Enumerate the steps needed to be able to access a grid insfrastructure, submit and collect results of your application using the gLite middleware.

12. Given the following example, what do you think it is going to be executed?

```
Type = "Job";
JobType = "mpich";
NodeNumber = 8;
Executable = "mpi.sh";
Arguments = "passtonext 8";
StdOutput = "output.file";
StdError = "error.file";
InputSandbox = {"passtonext","mpi.sh"};
OutputSandbox = {"error.file","output.file","program.out"};
Requirements = (other.GLUECEInfoLRMSType == "PBS") ||
               (other.GLUECEInfoLRMSType == "LFS")
```

13. What is the function of attribute InputData in the JDL below?

```
Executable = "glite_wrapper.sh";
Arguments = "myapp wf_12221605883_756000_760000.tgz
             guid:b0591574-0cfa-4537-85cd-b67a724c2d00
             \"./bio_env_grid.sh word_filter_task.sh
             wf_12221605883_756000_760000.i
             wf_12221605883_756000_760000.o  47\"
             wf_12221605883_756000_760000.o.tgz
           \"wf_12221605883_756000_760000.o\"";
StdOutput = "wf_12221605883_756000_760000.out";
StdError =   "wf_12221605883_756000_760000.err";
InputData = {"guid:b0591574-0cfa-4537-85cd-b67a724c2d00",
                      "guid:9a8d8024-a1f3-4921-a6ff-3d6f1f1f1ae0"};
DataAccessProtocol = {"rfio","gsiftp","gsidcap"};
InputSandbox = {"glite_wrapper.sh",
               "myapp/ubijobs/wf_12221605883_756000_760000.tgz"};
OutputSandbox = {"wf_12221605883_756000_760000.out",
        "wf_12221605883_756000_760000.err",
                          "wf_12221605883_756000_760000.o.tgz"};
```

14. What is the matchmaking process and how does it work?

15. The following questions refer to installation and configuration of gLite services.

  15.1 What are the preliminary steps to be able to correctly install the gLite middleware?

  15.2 What is the function of the LFC gLite service?

  15.3 What are the main gLite components that need to be installed in order to have a minimum set of services?

16. The following questions refer to the assignment 2

16.1 Write a pseudo-code to solve the floorplan design problem.

16.2 What do you think it would be the best method to "gridify" this application to run on an infrastructure that uses the gLite middleware?

16.3 In what conditions this code would not be suitable to run on a grid?

17. The following questions refer to the assignment 3

17.1 Eucalyptus is a cloud system ideal to test small programs before committing to the use of actual cloud services to put these programs in production. What are the main characteristics of Eucalyptus? Describe its architecture?

17.2 Amazon EC2 and S3 are what kind of services?

17.3 What are the main differences between Clouds and Grids?