Contribution ID: **137**   Type: **oral presentation**

Workflow level parametric study support by the P-GRADE portal

Thursday 10 May 2007 10:00 (20 minutes)

Describe the scientific/technical community and the scientific/technical activity using (planning to use) the EGEE infrastructure. A high-level description is needed (neither a detailed specialist report nor a list of references).

The P-GRADE portal is already used by several EGEE and EGEE-related
communities. It is used by specific, scientific communities like
the EGRID
(Economics Grid) and BioInfoGrid as well as by regional
communities like VOCE,
SEEGRID and BalticGrid. Besides, many national Grid installed it
as their Grid
portal: Bulgarian Grid, CroGrid, HunGrid, SwissGrid, Turkish
Grid, UK NGS, etc.
Recently, the GIN VO of OGF uses the portal as their resource
testing portal.

Report on the experience (or the proposed activity). It would be very important to mention key services which are essential for the success of your activity on the EGEE infrastructure.

A parametric study application, particularly if it has to execute
a workflow as
many times as many input parameter sets are available, requires the
orchestration of large number of grid resources. Therefore the
grid broker plays
a key role in such a service. The robustness of the broker is
particularly
important when thousands of jobs should be handled in a reliable
way. The
current middleware is not robust enough to handle thousands of
jobs without
errors and hence we extended the portal with the following features:

1. The user can suspend (and resume) the execution of any
   workflow. An often
   required feature when - for example - the selected Grid resource
   seems too
   slow to perform the needed job, and rescue would be advisable.

2. From now on the user does not have to wait for getting partial
   results until
   the last job of his workflow terminates. The results of
   terminated jobs can be
   downloaded immediately. So, workflows can be aborted if partial
   results make
   them useless.

With a forward look to future evolution, discuss the issues you have encountered (or that you expect) in using the EGEE infrastructure. Wherever possible, point out the experience limitations (both in terms of existing services or missing functionality)

P-GRADE portal enables grid interoperability at the job, workflow and parameter
study level. It means that jobs of a parameter study workflow can be
simultaneously executed in different grids (LCG-2, glite, GT2, GT4) connected by
the portal. However, to achieve load-balancing between grid resources and
grids we need a meta-broker that can connect brokers of different grids. We
are working on the development of such a meta-broker. It will be first
connected to glite WMS and GTBroker.

Describe the added value of the Grid for the scientific/technical activity you (plan to) do on the Grid. This should include the scale of the activity and of the potential user community and the relevance for other scientific or business applications

Large workflow applications often require to run the same workflow with many
different input parameter sets. This was a clear message from our user
communities. Therefore we have decided to extend the portal with the
capability of supporting this kind of applications.

We introduced a new workflow concept in P-GRADE where the workflow is
partitioned into three parts that are executed in three separate phases. The
first partition consists of Generator jobs (either user written or special portal
provided ones) that run in the 1st phase to generate the necessary input files
for the core workflow part. The 2nd partition consists of the core workflow that
is executed with all the input parameter sets generated in the 1st phase. The
third partition contains Collector jobs that collects and processes the results of
phase 2.

The generated input and output files are stored in SEs and the portal uses the
broker services of glite in order to orchestrate the parallel execution of
workflows.

Author: Prof. KACSUK, Peter (MTA SZTAKI)

Co-author: Mr FARKAS, Zoltan (MTA SZTAKI)

Presenter: Prof. KACSUK, Peter (MTA SZTAKI)

Session Classification: Workflow

Track Classification: Workflow