Contribution ID: **68**   Type: **Training session**

Computing on a big scale with OpenMOLE

Tuesday, 9 April 2013 11:00 (1h 30m)

Impact

In this training session we propose to teach how to design OpenMOLE workflows. Attendees will learn how to:
- embed an external application in OpenMOLE,
- compose task with each-other to design a workflow,
- design large scale workflows generating several thousand of jobs,
- gather / aggregate / store produced data,
- delegate the computing load to EGI,
- create an ad-hoc desktop grid.

At the end of the session attendees should all be able to design a large scale distributed application. During the session people will be provided with formation certificates or use there own certificate. They will work on the VO of their choice.

The example applications will mostly concern model exploration since OpenMOLE has been primarily designed to explore simulation models. However this training can be beneficial for enabling distributed computing for other purposes. For instance the bioemergences (http://bioemergences.eu) platform that process 4D morphogenesis images on EGI is based on OpenMOLE.

Summary

OpenMOLE is based on a blackbox approach to embbed application based on very different technologies: java / scala / C / C++ / fortran / scilab / octave / netlogo... Once embedded in the platform, OpenMOLE automatically distributes numerous executions of the application on a distributed environment specified by the user, such as a cluster or even on the 100,000+ cores available on the European computing grid. The platform deals with software installations, file transferts, job failures, rendering the distributed execution entierly transparent for the user.

URL

www.openmole.org

Description

The workflow paradigm is particularly well suited to describe distributed executions. Several platforms leverage this representation to fil the gap between scientific processes and high performance computing environments. In this training session we propose to advertise OpenMOLE (Open MOdel Experiment, http://www.openmole.org/), a workflow platform that exposes several advantages when compared to other mainstream worflow systems:
- it delegates transparently the computing load from the user computer directly to the remote execution environment following a zero-deployment approach and makes it convenient to embed third party software components in a workflow,
- OpenMOLE workflows are environment agnostic and the workload can be submited on EGI as well as on other distributed computing environment such as clusters, remote servers or local desktop grid,
- the dataflow is statically typed and formally checked prior to the runtime, minimizing the risk of error when desining complex workflows.

This platform has already been used to execute huge amount of computation and is now scalable, reliable and mature. In November and Decembre 2012 the major part of the computing time provided by the biomed VO has been orchestrated by OpenMOLE workflows.

Primary author: REUILLON, Romain (CNRS)

Co-author: Mr LECLAIRE, Mathieu (ISC-PIF)

Presenters: Mr LECLAIRE, Mathieu (ISC-PIF); REUILLON, Romain (CNRS)

Session Classification: Training

Track Classification: Virtual Research Environments (Track Lead: G Sipos and N Ferreira)

Primary author: REUILLON, Romain (CNRS)

Co-author: Mr LECLAIRE, Mathieu (ISC-PIF)

Presenters: Mr LECLAIRE, Mathieu (ISC-PIF); REUILLON, Romain (CNRS)

Session Classification: Training

Track Classification: Virtual Research Environments (Track Lead: G Sipos and N Ferreira)