Contribution ID: **77** Type: **Oral Presentation**

SOMA2 –Gateway to Grid Enabled Molecular Modelling Workflows in WWW-Browser

Monday, 11 April 2011 15:00 (30 minutes)

Overview

SOMA2 gateway is a molecular modeling workflow environment developed and deployed by CSC - IT Center for Science Ltd. The SOMA2 environment is used with WWW-browser and it allows users to combine scientific applications into unique application workflows, which are automatically executed in the underlying computing system. SOMA2 offers a flexible framework for integrating and executing molecular modeling applications, and facilitates automated molecular data exchange. SOMA2 source code is distributed under the GPL open source license.

For end users, SOMA2 offers a secure, personalized and easy to use environment for utilizing computing infrastructure. In SOMA2, scientific applications are presented and configured via web forms, which guide users to correctly configure a program by supplying default values, thresholds, runtime help and validation. For experts, SOMA2 offers a framework to make virtually any molecular modeling application accessible to the end users.

Impact

SOMA2 gateway provides an easy to use and intuitive single user interface to scientific applications and it hides all technicalities from end users. The system automates repeating tasks and eliminates redundant work so that end users can focus in the actual scientific task instead of dealing with technical issues requiring manual work.

With the new DCI-support features of SOMA2, we can really say that SOMA2 does not only integrate applications but also different computing infrastructures making complex computing environments reachable for all users. We think that technical complexity is still a major bottleneck, which prevents users to fully benefit from the distributed resources. SOMA2 system can alleviate this problem a lot.

The framework for describing scientific applications in SOMA2 facilitates transfer of technical know-how from experts to service users so that everything remains in machine readable form. In addition, core SOMA2 software is separated from the program descriptions and basically no real programming skills are required to create a SOMA2 capsule. Flexibility, application oriented approach for reusable workflows and open source license make SOMA2 system very unique in its domain.

Description of the work

SOMA2 enables communication and data exchange between applications by employing a common data exchange format, CML (Chemical Markup Language). The common data format is crucial for seamless integration of different applications within SOMA2 system. SOMA2 web application handles user's initial molecular

data by converting the data to CML format and preparing it for the application workflow. With the web application user is able to setup the workflow, configure chosen applications, submit project for computation, review status of the project and access the results.

In SOMA2, scientific applications and their execution is described in pluggable capsules, which make use of interfaces to manage the internal data. A capsule consists of an XML description, used e.g. to generate an application web form, and scripts & file templates to enable scientific application execution and processing of the program output.

SOMA2 system includes a workflow manager program which is responsible for internal data transfers and controlling execution of the capsules.

Recently we have added DCI-support in SOMA2. In SOMA2 web application and its utilities, this includes handling of users' X509 certificates, requesting available resources from the grid and generating proxy certificates upon project submission. In capsules, we make use of grid middleware to submit the jobs. This work has been conducted as part of EGI-InSPIRE project's WP6-SA3. Currently, we have support for Nordugrid Arc middleware but other middleware could be used as well.

URL

http://www.csc.fi/soma

Conclusions

SOMA2 source code is distributed openly for all interested parties. At CSC, SOMA2 is also available as a service for CSC's academic users providing access to 14 different molecular modeling applications, which are seamlessly integrated within the system. System is fully integrated with the local computing infrastructure.

Within EGI-InSPIRE WP6-SA3, we plan to extend the current SOMA2 service to include DCI-enabled applications. This service will be first introduced to FGI. Later on, we plan to make the service available for other communities in EGI.

In addition to common enhancements in the web application, SOMA2 development plans include investigating other middleware to be used in SOMA2 DCI-integration and new applications to be integrated in SOMA2.

Primary author: Dr KINNUNEN, Tapani (CSC)

Presenter: Dr KINNUNEN, Tapani (CSC)

Session Classification: User Environments

Track Classification: User Environments - Portals