Contribution ID: **152** Type: **not specified**

SOMA2 –Application Oriented Molecular Modelling Workflows in WWW-Browser

Thursday, 29 March 2012 11:20 (20 minutes)

Overview (For the conference guide)

SOMA2 gateway is a WWW-browser operated molecular modeling workflow environment developed by CSC - IT Center for Science Ltd. The SOMA2 environment 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. Since version 1.4 SOMA2 has support to make use of grid resources with suitable middleware software. Grid related development work has been part of EGI-InSPIRE project's WP6-SA3.

Description of the Work

For end users, SOMA2 is made very easy to use. SOMA2 tackles common technical problems related to utilizing complex 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. SOMA2 focuses on describing the scientific applications, their parameters and services that the applications provide, and therefore SOMA2 is not only an interface to technical computing environment. For service providers, SOMA2 offers a framework to make virtually any molecular modeling application accessible to the end users.

Application workflows of SOMA2 are based on seamless data exchange between different scientific applications. This is achieved by employing a common data format, CML (Chemical Markup Language). SOMA2 web application handles user's initial molecular data by converting the data to CML format and preparing it for the application workflow. SOMA2 web application allows user 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. A capsule consists of an XML description, which is used to generate an application web form. In addition a capsule contains scripts & templates to enable scientific application execution and processing of the program output.

Since version 1.4, SOMA2 supports integration with grid resources. SOMA2 handles users' X509 certificates, resolves users'privileges to a grid enabled application and generates proxy certificates upon project submission. In job submission, we make use of suitable grid middleware. Currently, we have support for Nordugrid Arc middleware but other middleware could be used as well. Our first grid enabled application is AutoDock version 4 performing virtual screening for computational drug discovery.

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 large set of 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 more DCI-enabled applications and make the service available for other communities in EGI as well.

Impact

SOMA2 gateway provides an easy to use and intuitive single user interface to scientific applications and it hides all technicalities from end users. A concept of application workflow allows users to focus on the task at hand without a need to know technical details of format conversions, data copying and job submission, for example. The system automates repeating tasks and eliminates redundant work so that end users can focus in use of the scientific application instead of solving technical issues, which require additional skills and work.

With the grid related features, SOMA2 does not only integrate applications but also different computing infrastructures making complex computing environments reachable for the users. We think that technical complexity is still a major bottleneck, which prevents users to fully benefit from available resources. SOMA2 system alleviates this problem a lot. Flexibility, application oriented approach for reusable workflows and open source license make SOMA2 system very unique in its domain.

URL

http://www.csc.fi/soma

Primary author: Dr KINNUNEN, Tapani (CSC)

Co-authors: Mr MATTILA, Kimmo (CSC); Ms KOSKINEN, Pinja (CSC)

Presenter: Dr KINNUNEN, Tapani (CSC)

Session Classification: Portals & Gateways

Track Classification: Software services for users and communities