Contribution ID: **202** Type: **Oral Presentation**

Coarse Grained Interoperability of Heterogeneous Grid Workflow Systems

Tuesday, 12 April 2011 17:00 (30 minutes)

Conclusions

The experience gained so far in this project has indicated that the reusability of workflows can be improved by following simple guidelines to reduce dependencies on the specific environment adopted by the user community that develops and runs them. However some dependencies on the description language, workflow system, credential handling and file access protocols cannot be controlled at the workflow level, so they need to be solved at the SHIWA platform level.
The SHIWA project is currently completing the implementation of SSP including coarse-grained interoperability mechanism. Future work consists of enacting the pilot workflows in different workflow systems using the SSP and preparing other (more complex) applications for publication on the SSP.

Overview

Several Grid workflow management systems (WfMS) emerged in the last decade. These systems were developed by different scientific communities for various purposes, and many applications have been ported to the grid using them. Reusing workflows in different WfMS and building new experiments by utilising existing workflows as building blocks is currently not supported. WfMS typically have their own user interfaces/APIs, description languages, provenance strategies, and enactment engines, which are not standard and do not interoperate.
The European SHIWA project aims to leverage existing solutions and enable cross-workflow and inter-workflow exploitation of DCIs by applying coarse- and fine-grained strategies. The coarse-grained approach treats workflow engines as distributed black box systems, while the fine-grained approach addresses language interoperability by defining an intermediate representation. This contribution concentrates on results of the coarse-grained approach.

Impact

Coarse-grained workflow interoperability enables the coordinated execution of workflows created in different workflow systems, from a host workflow. The created “meta-workflows”may span multiple heterogeneous grid infrastructures.
Workflow interoperability allows workflow sharing to support and foster the adoption of common research methodologies, improves efficiency and reliability of research by reusing these common methodologies, increases the lifetime of workflows, and reduces development time for new workflows.
The SHIWA project develops a coarse-grained interoperability solution that enables the publishing, searching and sharing of workflow engines and workflows, in a repository. Using the repository, users can select already existing workflows or workflow components and embed them into their native workflow. The developed solution supports the above functionalities without requiring users to understand the nature and technical details

of the workflow system that hosts the embedded workflow. The embedded workflow is considered as a black-box, and represented by its input and output parameters and the executing workflow engine only.
The new services developed and deployed by SHIWA allow different scientific communities to exchange applications, workflows, and data resources regardless of the workflow system in which they are used. Therefore, it enables the development of such inter-disciplinary and inter-organizational workflow applications that were not possible before.
The results of the project are directly utilised by selected user communities from the areas of medical imaging, bioinformatics and chemistry. The communities provide pilot application workflows that serve as benchmarks for the developed platform and form the basis of a larger user community that will utilise SHIWA services in the future.

Description of the work

The coarse-grained interoperability work inside the SHIWA project refers to the nesting of different workflow systems to achieve interoperability of execution frameworks. The solution integrates different workflow engines to a grid application repository and submitter service, called GEMLCA, and allows grid based workflow systems to access non-native workflow engines if they have to execute embedded workflows.
The current approach facilitates the embedding of Kepler, Taverna, Triana, GWES and Moteur workflows into P-GRADE workflows. GEMLCA is capable to either invoke pre-deployed workflow engines through a variety of grid middleware solutions (including GT2, GT4 and g-Lite), or to submit and execute the engines on resources of infrastructures built with these types of middleware.
The first version of the SHIWA Simulation Platform (SSP) uses the NGS P-GRADE portal and resources of the Westfocus VO of the UK NGS to execute embedded workflows. The platform is currently being tested and utilised by the SHIWA user community. The first considered workflows implement pilot applications from different areas: neuroimaging (Charité Berlin, D-Grid, GWES); DNA sequence alignment (AMC Amsterdam, DutchGrid, MOTEUR); simulation of chemical reactions (Charité Berlin, EGI, MOTEUR) and medical simulation (CNRS Creatis, EGI, MOTEUR). The existing workflows had to be adapted to increase their portability respectively to credentials needed to access the workflow description and DCI resources for enactment; flexibility and neutrality of invocation and management of input data and output results; data location and transfer protocols; description language; and implicit functionality of the workflow system; etc. These workflows are now available at the SHIWA repository and can be used to compose more complex "meta-workflows".

URL

http://www.shiwa-workflow.eu/

Primary authors: KREFTING, Dagmar (Charité –Universitätsmedizin Berlin); KECSKEMETI, Gabor (MTA SZTAKI); TERSTYANSZKY, Gabor (University of Westminster); MONTAGNAT, Johan (CNRS I3S Lab, UMR 6070); MAHESHWARI, Ketan (CNRS I3S Lab, UMR 6070); WEINGARTEN, Noam (University of Westminster); KACSUK, Peter (MTA SZTAKI); D. OLABARRIAGA, Silvia (Academic Medical Center of the University of Amsterdam); KISS, Tamas (University of Westminster); KUKLA, Tamas (University of Westminster); GLATARD, Tristan (CREATIS CNRS UMR 5220); KORKHOV, Vladimir (Academic Medical Center of the University of Amsterdam)

Session Classification: Technologies for Distributed Computing

Track Classification: User Environments - Portals