Geophysical Research Abstracts
Vol. 15, EGU2013-10856, 2013
EGU General Assembly 2013

Flexible workflow sharing and execution services for e-scientists

Péter Kacsuk (1), Gábor Terstyanszky (2), Tamas Kiss (2), and Gergely Sipos (3)

(1) MTA SZTAKI, LPDS, Budapest, Hungary, (2) University of Westminster, Centre for Parallel Computing, London, UK, (3) EGI.eu, Amsterdam, Netherlands

The sequence of computational and data manipulation steps required to perform a specific scientific analysis is called a workflow. Workflows that orchestrate data and/or compute intensive applications on Distributed Computing Infrastructures (DCIs) recently became standard tools in e-science. At the same time the broad and fragmented landscape of workflows and DCIs slows down the uptake of workflow-based work. The development, sharing, integration and execution of workflows is still a challenge for many scientists.
The FP7 "Sharing Interoperable Workflow for Large-Scale Scientific Simulation on Available DCIs" (SHIWA) project significantly improved the situation, with a simulation platform that connects different workflow systems, different workflow languages, different DCIs and workflows into a single, interoperable unit. The SHIWA Simulation Platform is a service package, already used by various scientific communities, and used as a tool by the recently started ER-flow FP7 project to expand the use of workflows among European scientists. The presentation will introduce the SHIWA Simulation Platform and the services that ER-flow provides based on the platform to space and earth science researchers.

The SHIWA Simulation Platform includes:
1. SHIWA Repository: A database where workflows and meta-data about workflows can be stored. The database is a central repository to discover and share workflows within and among communities .
2. SHIWA Portal: A web portal that is integrated with the SHIWA Repository and includes a workflow executor engine that can orchestrate various types of workflows on various grid and cloud platforms.
3. SHIWA Desktop: A desktop environment that provides similar access capabilities than the SHIWA Portal, however it runs on the users' desktops/laptops instead of a portal server.
4. Workflow engines: the ASKALON, Galaxy, GWES, Kepler, LONI Pipeline, MOTEUR, Pegasus, P-GRADE, ProActive, Triana, Taverna and WS-PGRADE workflow engines are already integrated with the execution engine of the SHIWA Portal. Other engines can be added when required.

Through the SHIWA Portal one can define and run simulations on the SHIWA Virtual Organisation, an e-infrastructure that gathers computing and data resources from various DCIs, including the European Grid Infrastructure. The Portal via third party workflow engines provides support for the most widely used academic workflow engines and it can be extended with other engines on demand. Such extensions translate between workflow languages and facilitate the nesting of workflows into larger workflows even when those are written in different languages and require different interpreters for execution.
Through the workflow repository and the portal lonely scientists and scientific collaborations can share and offer workflows for reuse and execution. Given the integrated nature of the SHIWA Simulation Platform the shared workflows can be executed online, without installing any special client environment and downloading workflows.
The FP7 "Building a European Research Community through Interoperable Workflows and Data" (ER-flow) project disseminates the achievements of the SHIWA project and use these achievements to build workflow user communities across Europe. ER-flow provides application supports to research communities within and beyond the project consortium to develop, share and run workflows with the SHIWA Simulation Platform.