Contribution ID: **82** Type: **Presentations**

Experiences from workflow sharing using the SHIWA Workflow Repository for application porting to DCI

Thursday, 11 April 2013 14:20 (20 minutes)

Impact

Porting applications to DCIs is a complex task that requires collaborations between computer scientists and domain scientists. By using workflow technology, porting of new applications and data processing pipelines can be simplified. Reusing workflows that are published in the SHIWA Repository simplifies this process even further.
The effort of porting the workflow to the AMC supported infrastructure was minor if compared with a complete application porting from scratch.
A new workflow implementation is now available, which will be also published in the SHIWA Workflow Repository.
By sharing workflows between London and Amsterdam, we have also implicitly initiated collaboration between two duo's of domain and computer scientists at our respective institutes.
Currently we have done the initial cycle of importing the original workflow and modifying that to our needs. We will be exploring how the loop can be completed and how the original author and users can benefit from our additions through the SHIWA Workflow Repository.
Although still in the early stage of collaboration, the scaling up of the experiments has already triggered the people from the Medical Biochemistry department to think of bigger and more frequent experiments. This will impact the research at that department similarly to what we have seen on other groups at the AMC.
The user interface to be provided for these users will be a very important factor to determine whether this new workflow will be adopted in practice. We will create a new portlet for the WS-PGRADE portal at the AMC that will act as a customized interface to the workflow for our user. Plans are to include visualizing capabilities and rudimentary statistical analysis for result selection.
Similar efforts are being undertaken at UoW and we will be evaluating how we can collaborate on that level while still maintaining the ability to create a customized environment for the specific (and possibly different) needs of our users. This activity will take place in the ER-Flow support action project that was established by the European Commission to further extend the user community behind the SSP.
While creating a workflow is not the definite step in creating a dedicated "service"for our end-users, it certainly is an important component, and the ability to share and collaboratively improve it enables us to focus on the customization.

Summary

AutoDock Vina is an open-source program for drug discovery, molecular docking and virtual screening and experiments with this tool demand large computing power.
Computer Scientists at the University of Westminster composed an Autodock Vina workflow for scientists from the School of Life Sciences.
At the Academic Medical Center in Amsterdam, scientists of the department of Medical Biochemistry also use Autodock Vina for virtual screening experiments. They contacted the e-BioScience group to run this application on the Dutch Grid.
Although the workflow developed by Westminster was designed to run on a different infrastructure, modifying it to meet the local requirements was a reasonable task.
Our experience shows that workflows provide a suitable level of abstraction for describing complex experiments on DCIs and that a workflow repository can effectively support the sharing and re-using of scientific data processing pipelines between different communities and infrastructures.

Description

AutoDock Vina is an open-source program for drug discovery, molecular docking and virtual screening. It offers multi-core capability, high performance, enhanced accuracy and ease of use.
Scientific Workflows provide a well-established abstraction for describing and executing computationally intensive science experiments. Computer Scientists at the University of Westminster composed an Autodock Vina workflow, using the WS-PGRADE workflow system and science gateway framework, for scientists from the School of Life Sciences. The resulting workflow enabled the researchers to execute large scale virtual screening experiments from a high level science gateway interface.
The workflow was originally supported to run on the BOINC-based local desktop grid of the University of Westminster, and later ported to the public volunteer computing resources of EDGeS@home.
The workflow has been published in the SHIWA Workflow Repository enabling other researchers to execute the workflow directly on the SHIWA Simulation Platform, or to download and customize the workflow according to their requirements.
At the Academic Medical Center (AMC) in Amsterdam, scientists in the department of Medical Biochemistry also use Autodock Vina for virtual screening experiments.
Virtual screens of large databases are used as a starting point to identify small molecules that can interact with proteins and modulate activity. Potential targets are further evaluated using conventional biochemical assays.
With the growth of databases, running their experiments on their local infrastructure became more and more challenging due to long computation times. They approached the e-BioScience team down the hall, who had experience with running Next Generation Sequencing and Medical Imaging applications on the Dutch Grid infrastructure through their e-BioInfra Gateway. The original Autodock Vina workflow was imported from the SHIWA Repository into the AMC environment, and modified for the local usage scenario.

Primary author: SANTCROOS, Mark (Academic Medical Center of the University of Amsterdam)

Co-authors: BLEIJLEVENS, Boris (Academic Medical Center of the University of Amsterdam); TERSTYANSZKY, Gabor (University of Westminster); GREENWELL, Pamela (University of Westminster); OLABARRIAGA, Silvia (Academic Medical Center of the University of Amsterdam); KISS, Tamas (University of Westminster, London, UK)

Presenter: SANTCROOS, Mark (Academic Medical Center of the University of Amsterdam)

Session Classification: Workflow Collaboration Workshop

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