Contribution ID: **53**   Type: **Presentations**

BioVeL: Taverna Workflows on distributed grid computing for Biodiversity

Thursday, 11 April 2013 14:40 (25 minutes)

Impact

Thanks to the activities carried on in BioVeL project the researchers involved in the biodiversity studies are able to exploit several different computing infrastructure, like: EGI grid, Desktop Grid, local cluster, cloud resources provided by EGI Cloud Task Force, dedicated servers or public available services. All those facilities are easily and seamless available to the end users and could be used to build complex data processing workflows allowing also non expert users to process big dataset of data without the need to learn the complexity of each computing infrastructure. The choice to use standard interface like a web services, between the computing environments and the Taverna workflow manager give the possibility to re-use this kind on infrastructure also in other context with different user communities and different tools. Moreover the work done at the level of user interface has greatly improved the users experience reducing the time needed to achieve the analysis results

URL

http://www.biovel.eu/

Summary

BioVeL is a virtual e-laboratory that supports research on biodiversity issues using large amounts of data from cross-disciplinary sources.BioVeL offers the possibility to use computational workflows to process data, be that from one's own research and/or from existing sources.A researcher can build his own workflow by selecting and applying successive services (data processing) registered in BiodiversityCatalogue.org, or he can re-use existing workflows available from BioVeL's library held in biovel.myexperiment.org.This e-laboratory cuts down research time and overhead expenses. BioVeL also provides access to a world-wide network of experts and support through a Specific Support Centre.We will present a solution for exploit large and complex computing and storage facilities, behind simple and seamless web services that are able to provide access to a specific application or software executed in a specific computing infrastructure accordingly to the characteristics of the application

Description

BioVeL users are interested in building complex workflow by means of the Taverna Workbench, and re-use those workflow in a simple web interface such as Taverna Lite or through their own and others applications/portals.Workflows execute on a Taverna Server.We will describe the activities carried on within the BioVeL project in order to build robust web services that could handle the users'requests both in terms of application execution and data management.Using standard REST and SOAP interfaces guarantee the users to be able to exploit in the same workflow public available services(e.g. GBIF, Google Refine, OpenModeller)or dedicated services that could be deployed on: dedicated server such as GPU enabled server, local batch farm, EGI grid,AWS cloud resources,BOINC enabled desktop, etc.We can easily handle different kinds of application from a fast and simple file parsing to a large campaign of comparing two big dataset of biological data,or a long-running MPI multi CPU application.Within BioVeL project we have implemented the data services that are able to provide to the final users the capabilities to handle the data needed by the application.This service provides to the end user a REST based web services interface, but give also the possibility to mount as local file-system the storage area on the users computer so that is easy for the end users to move input and

output.In case of EGI infrastructure the standard EGI data-management layer is exploited in all the functionalities.Future alternative is to exploit EUDAT services as soon as they will be available.We will show also the development done in order to ensure a simple and user friendly interface to the final users providing the possibility to easily exploit already developed data processing workflows with small intervention useful only to customize the single execution.The framework is able to deal with automatic job submission,monitoring and resubmission to hide the complexity of computing infrastructures

Primary author: Dr DONVITO, Giacinto (INFN)

Co-authors: Dr WILLIAMS, Alex (University of Manchester); Dr MAROSI, Attila (Laboratory of Parallel and Distributed Systems Computer and Automation Research Institute, Hungarian Academy of Sciences); Dr BALECH, Bachir (CNR-IBB); Dr GOBLE, Carole (University of Manchester); Dr HORVÁTH, Ferenc (Institute of Ecology and Botany Centre for Ecological Research, Hungarian Academy of Sciences); Dr BACALL, Finn (University of Manchester); Dr ALEX, Hardisty (Cardiff University); Dr MORRISON, Norman (University of Manchester); Dr NOTARANGELO, Pasquale (INFN-Bari); Dr ITTZÉS, Peter (Institute of Ecology and Botany Centre for Ecological Research, Hungarian Academy of Sciences); HAINES, Robert (University of Manchester); Dr VICARIO, Saverio (CNR-ITB); Dr WITHERS, david (University of Manchester)

Presenter: Dr DONVITO, Giacinto (INFN)

Session Classification: Workflow Collaboration Workshop

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