VIP ANR-09-COSI-03
Virtual Imaging Platform

D2.3.4: VIP client to semantic and execution services

Rafael Ferreira da Silva, Sorina Camarasu-Pop, Frédéric Cervenansky,
Germain Forestier, Bernard Gibaud, Tristan Glatard

Abstract

This document describes the VIP portal providing interface to the VIP execution service, and to semantic repositories to share simulation object models. A model repository with annotation features is now available. Tools and interfaces used for system administration, general-purpose usage, simulation execution on the European Grid Infrastructure (EGI), and interaction with repository of simulation object models are described. The VIP portal is deployed by the project and freely accessible at http://vip.creatis.insa-lyon.fr. According to EGI data, its robot certificate has the largest number of users in Europe. In terms of activity, the EGI accounting system[a] reports that 379 normalized CPU years have been consumed by the VIP portal from January 2011 to August 2012, which makes it the 2nd user of the whole biomed VO.

[a] http://accounting.egi.eu

Contents

1 Introduction

This document describes the features of the VIP portal available in production at http://vip.creatis.insa-lyon.fr. It allows users to submit, monitor and manage simulations on distributed computing resources of the European Grid Infrastructure[1]. Users can also transfer files between their local machine and grid storage by using a file transfer tool. Interfaces to semantic repositories are also available for simulation object model sharing.

Special care has been devoted to make the interface minimalistic, easy to access, easy to administrate, and easily extensible. The VIP portal implements a client-server architecture based on servlets implemented using the Google Web Toolkit[2].

Applications are organized in classes that users can access based on their group membership. When the user selects an application, the portal generates a form from its Gwendia workflow description [2]. This form is filled in with input files and parameters and the workflow is submitted to the VIP execution service (see VIP D2.2.1). The portal provides statistics of workflow executions, jobs statuses and execution logs. Active workflows can be killed and inactive ones can be cleaned-up and purged.

Section 2 details the tools used for system management, Section 3 explains the general-purpose tools available to the users, Section 4 focuses on interfaces to semantic services for model sharing, Section 5 describes the launching and monitoring of simulations, and Section 6 concludes with some usage statistics.

2 System management

The following tools are devoted to VIP system administrators.

Users. VIP users only require an email address to get an account. Account creation is automatically validated after email validation. Authentication with X509 certificates was rapidly abandoned due to their unwieldiness. Instead, a single (robot) certifi-

[1] http://www.egi.eu

[2] https://developers.google.com/web-toolkit/

cate is used by the portal for all grid operations. Three user levels are available: *beginners* can only launch one simulation at a time, and cannot write in shared directories; *advanced users* have extended rights, but they have to have a valid X509 certificate registered in the biomed EGI VO; *administrators* have all rights. 250 users from 25 countries are currently registered.

Groups. VIP users may belong to one or several groups. Groups may be public. Public groups can be joined without moderation, while private group membership requires validation by an admin.

Applications. VIP applications are only defined by the LFN of a Gwendia worfklow. A new application can therefore be added very quickly, without any particular development. Applications are deployed on the biomed VO independently from the portal. No application file is kept by the portal.

Applications classes. VIP applications are grouped in classes. Classes may include different versions of an application, or related applications. For instance, "Simulation - open" contains publicly accessible simulators. Application classes are associated to user groups. The resulting user↔group↔class↔application association is very flexible to define access rights to applications in the platform (Figure 1).

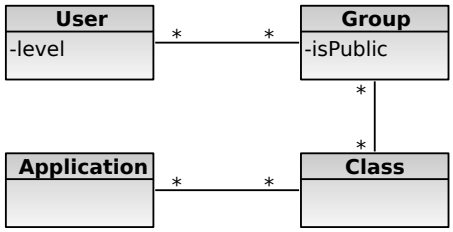

Figure 1: Association among users, groups, classes and applications.

File operations. This tool allows to monitor and manage the status of all data transfers between EGI and the portal. For instance, transfer errors can quickly be spotted, and the transfer queue can be inspected. It gives a view of the database used by the transfer agent GRIDA used for the transfers[3].

[3] http://kingkong.grid.creatis.insa-lyon.fr:9002/projects/vletagent

Cached files. This tool shows the files cached in the platform, their size, and hit rate. It gives a good indication of file popularity. Typically, application workflows are the most popular files since they are accessed each time a simulation is launched.

Zombie files. Files in EGI are stored on storage elements (SE), and registered in a logical file catalog (LFC). Unfortunately, no mechanism exists to ensure the consistency between file entries on SEs and on the LFC. This tool records inconsistencies detected by the platform so that administrators can curate these files. 10,400 zombie files have been identified between March and August 2012.

3 General features

The following general-purpose tools are provided to all users.

Account management. Users can modify their personal information (name, institution, phone, country) and password. They can also see all existing groups, and join public groups. Users can also delete their account, which deletes all their personal data, files stored on the platform, and anonymizes their simulations.

Messages. A messaging system can send emails to users and groups. It is used to announce system downtimes, updates, and other news. It avoids maintaining and synchronizing a user mailing list separately from the portal.

Documentation. Basic documentation is available about how to launch simulations and transfer files.

Gallery. An image gallery shows some simulation results obtained with the platform.

File Transfer. The file transfer tool enables data migration between local user machines and grid storage. The upload process consists in (i) uploading the file from the local machine to the portal and (ii) transferring the file to the grid through an

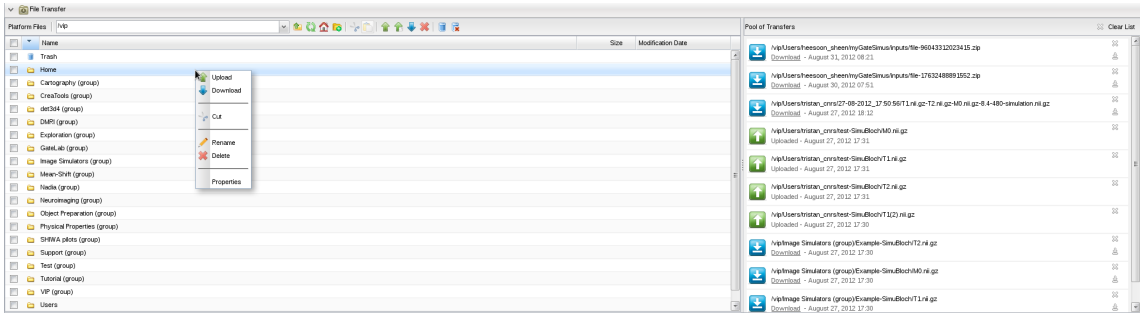

Figure 2: Screenshot of the file transfer tool.

asynchronous pool of transfers processed sequentially. Download is performed similarly, in the opposite direction. This two-step process avoids connectivity issues between user machines and distributed grid hosts. The transfer pool manages the load of concurrent transfers performed by the server to avoid network clogging on the portal machine. It also replicates files to ensure availability, and it caches a local copy of the transferred files. An optimized file browser is available to interact with the transfer service. It caches grid directory content and uses native catalog commands to ensure fast browsing. File permissions are also enforced by the browser: users have a private folder readable only by themselves, and each group has a shared folder. A screenshot of the file transfer tool is shown on Fig. 2.

4 Interfaces to semantic repositories for simulation object model sharing

The following tools are provided to users of group “Image simulators”.

Model repository with online annotation of object models. A semantic repository is available for sharing simulation object models. This repository allows to import models, browse existing ones using semantic queries, inspect semantically annotated model content, and annotate new models online. Model import, browsing, and annotation inspection were described in milestone M2.3.2. In addition, a specific interface is now available to create empty models, annotate and update them with new components. This interface provides the usual actions (add, remove, duplicate, rename, etc) for the different layers, timepoints and instants of a model. The OntoVIP ontology (see D1.1.1) can be searched to associate anatomical objects with semantic definitions. Semantic queries, made available by the model repository (see D1.2.1), are performed to infer statements such as the affiliation of an object part to a particular layer (e.g. pathological). Inference rules are also evaluated to check if models are ready to be used in simulations of particular imaging modalities. A save command allows to complete the import in model repository. To guarantee long-term tracability of the models, model modifications are always performed on a copy of the initial model.

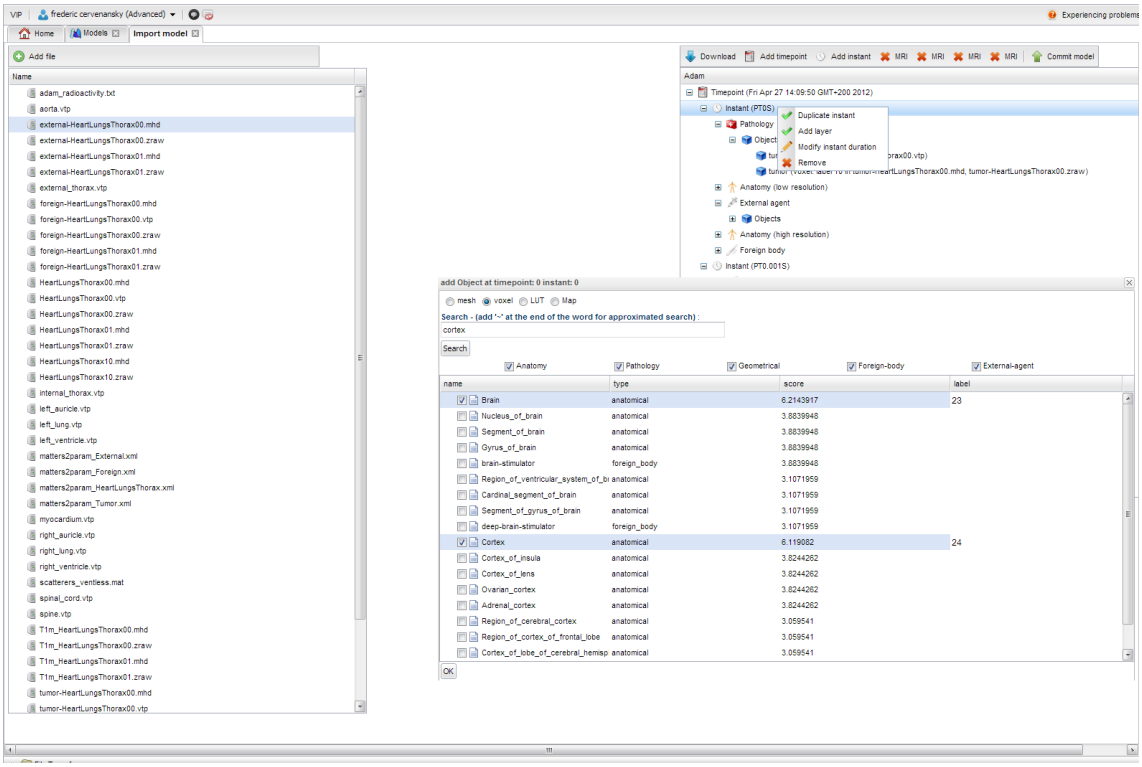

Figure 3: Model annotation interface

Simulation scene editor. A simulation scene editor allows to visualize in 3D a model from the repository (using Web-GL technology[4]), to position imaging scanners w.r.t this model, and to launch simulations. It was described in M2.3.2.

5 Applications

The following tools are provided to users belonging to a group mapped to at least one application class.

Application launcher. An application launcher is automatically created for each VIP application. Using this launcher, users can:

- specify simulation input parameters;
- save/load parameters for later re-use;
- save/load parameters as examples available platform-wide;

[4] http://www.khronos.org/webgl

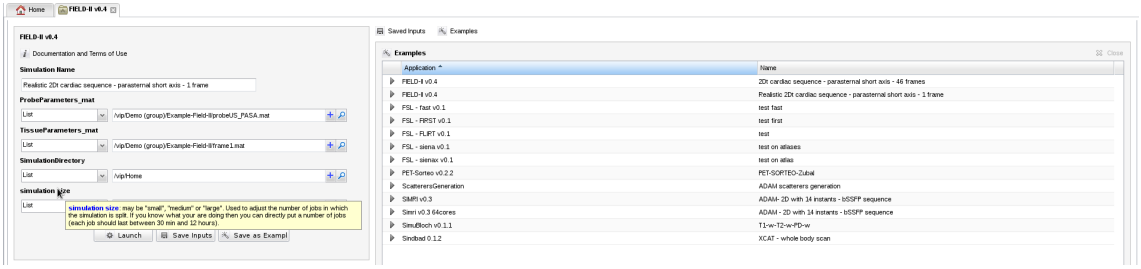

Figure 4: FIELD-II application launcher.

- launch simulations.

Basic documentation about the application and its inputs is also available, extracted from the workflow description. Fig. 4 shows a screenshot of the application launcher, where a FIELD-II example has been loaded and is ready for launching.

Simulation monitor. The simulation monitor provides a list of active and past simulations. Users can get information about the input and output data involved in the simulation, about detailed task statuses and performance statistics collected by the VIP execution service (see D2.2.1).

GATE Lab. A specific interface is provided to GATE users. It was developed in ANR project COSINUS hGATE (2009-2012). It is detailed, e.g., in [1].

6 Usage statistics

Fig. 5 shows the number of users per robot certificate as reported by EGI. According to this data, the production instance of the VIP portal deployed by the project has the highest number of users in Europe, which is an indication of its usability and of the robustness of the underlying execution service. In terms of activity, the EGI accounting system[5] reports that 379 normalized CPU years have been consumed by VIP from January 2011 to August 2012. It makes VIP the 2nd

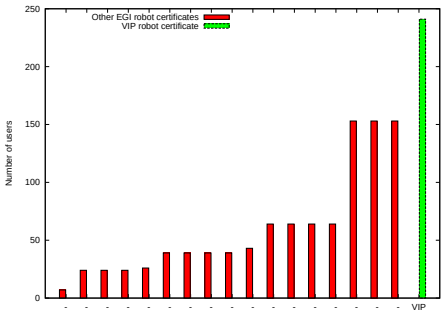

Figure 5: Users of portal with robot certificates in EGI in August 2012. Extracted from https://wiki.egi.eu/wiki/EGI_robot_certificate_users

[5] http://accounting.egi.eu

user of the whole biomed VO. Since January 2011, 6204 simulations have been launched on VIP, which represent about a million of tasks processed.

7 Conclusion

The VIP portal provides an interface to the VIP execution service, and to semantic repositories to share simulation object models. It is available for download at http://vip.creatis.insa-lyon.fr:9002/projects/platform/wiki, with installation and configuration instructions. The current release is 0.8.3. It has been tailored for ease of access (no complicated login mechanism), ease of use (minimalistic interface), and ease of administration (system management tools). According to EGI data, the robot certificate used by the production instance deployed by the project has the largest number of users in Europe, which is an indicator of its quality. The model repository is now available with annotation features, but it is still in an early release and it was not largely advertised yet. Based on its functionalities, it will be enriched with a wider range of models in the reminder of the project.

References

[1] S. Camarasu-Pop, T. Glatard, H. Benoit-Cattin, and D. Sarrut. *Enabling Grids for GATE Monte-Carlo Radiation Therapy Simulations with the GATE-Lab*. 2011.

[2] J. Montagnat, T. Glatard, and D. Lingrand. Data composition patterns in service-based workflows. In *Workshop on Workflows in Support of Large-Scale Science (WORKS'06)*, Paris, France, jun 2006.