Composition and migration to the cloud of the @neurIST workflow in VPH-Share

Maria-Cruz Villa-Uriol[1], Gerhard Engelbrecht[1], Héctor Fernandez[1], Xavier Planes[1], Luigi Carotenuto[1], Susheel Varma[2], Rod Hose[2], Alejandro Frangi[2]

[1] Universitat Pompeu Fabra, [2] University of Sheffield

Correspondence: cruz.villa@upf.edu, +34935422237, c/Roc Boronat 138, 08018 Barcelona, Spain

1. Introduction

The Virtual Physiological Human (VPH) intends to provide an integrative vision of physiological and computational knowledge. In general, Physiome projects face the challenge of encountering a solution to their common infrastructure needs, able to facilitate integration and prediction, and to describe the *Homo sapiens in silico*[1,2]. Along these lines, the goal of the Virtual Physiological Human Initiative (VPH-I) is to provide a systematic framework for understanding physiological processes in the human body in terms of anatomical structure and biophysical mechanisms across multiple length and time scales. In the long term it should transform the delivery of European healthcare into a more personalised, predictive, and integrative process, with significant impact on healthcare and on disease prevention.

The European project VPH-Share[4] was motivated by the needs of the whole VPH community to harness ICT technology to improve health services for the individual. The main objective of VPH-Share is to provide the organisational fabric (the infostructure), realised as a series of services, offered in an integrated framework, to expose and to manage data, information and tools, to enable the composition and operation of new VPH workflows and to facilitate collaborations between the members of the VPH community.

To prove that the proposed solution is not only conceptually elegant, computationally efficient and robust, but also usable by VPH researchers, VPH-Share chose four flagship computational workflows from ICT projects in the 6th and 7th Framework Programmes, namely @neurIST, euHeart, VPHOP and Virolab. Together they encapsulate the breadth of challenges presented to the VPH researcher.

This paper outlines a novelty of composing a complex VPH-workflow using a scientific workflow management system and migrating the entire workflow in the Cloud. The exemplar VPH workflow considered is taken from the @neurIST project and its complex information processing tool chain[3], which aims for the physical characterization of aneurysms to find candidate risk factors associated with aneurysm rupture.

2. Workflow composition

In VPH-Share we define a computational workflow as a sequence of operations to access data or information, to operate upon it and to analyse or to summarise the results of these operations. In this context, normally clinical data or information about the individual patient is to be accessed. The original @neurIST workflow comprised morphological, haemodynamic and structural characterisations of intracranial aneurysms appearing in a medical image. In VPH-Share we concentrated on the morphological and haemodynamic characterisations. A first step was to express the @neurIST workflow as a composition of the operations outlined in Figure 1.

The process starts from a medical image and obtains a 3D mesh representing the vascular geometry to be analysed. Once the segmented surface mesh is available, it is manually manipulated to remove some of the artefacts not belonging to the cerebral vasculature or not relevant for the subsequent analyses. As a result of the morphological characterisation, a collection of morphological descriptors of various complexities is computed. For the haemodynamic characterisation, a set of qualitative and quantitative descriptors (candidate risk factors) are extracted from the flow solution automatically, except for a few qualitative ones such as flow stability and intra-aneurysmal flow type requiring from expert judgment.

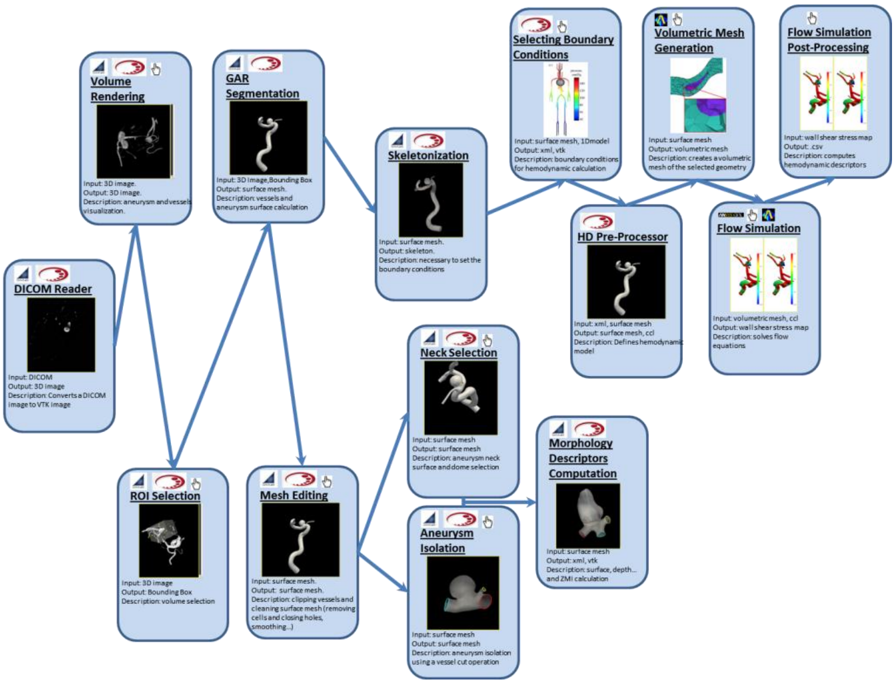

Figure 1. Atomization of services in the @neurIST workflow

3. Workflow migration to the cloud

The atomization of the @neurIST tool-chain comprised a decomposition of the entire chain of activities into separate operations, each performed with an atomic service. Each atomic service comes with a detailed definition of its inputs and outputs, as well as an internal application that actually provides the desired functionality. In addition the details of a potential user interaction are also defined if a particular service requires user inputs to continue the task.

The overall orchestration of these individual services eventually facilitated a workflow-enabled version of the @neurIST tool-chain. The actual structure and sequence of activities with all associated service invocations has been defined and stored as a well-defined workflow. In VPH-Share the workflow management system Taverna[5] is used for this purpose. Taverna is a popular solution for the definition and execution of scientific workflows.

In terms of deployment, each atomic service is embedded in a virtual machine template, which can be instantiated on demand in a Cloud environment such as the Atmosphere Cloud computing platform. This scalable approach provides a separate instance of the workflow for each user and consequently all users can interact with the @neurIST applications in parallel without interfering with each other.

Figure 2 depicts a snapshot of the workflow execution (left) and the interaction with a particular atomic service (right). In more detail the @neurIST workflow is executed using the Taverna workbench on the local machine or on the Taverna Server installed in the VPH-Share cloud. By using the Taverna workbench the current status is updated whenever progress is made. In case of a service, which requires user interaction, the user connects to the remote machine using a Virtual Network Computing (VNC) client and performs the interaction required (e.g. like mesh editing). Once the interaction is finished, the workflow execution continues to the next step.

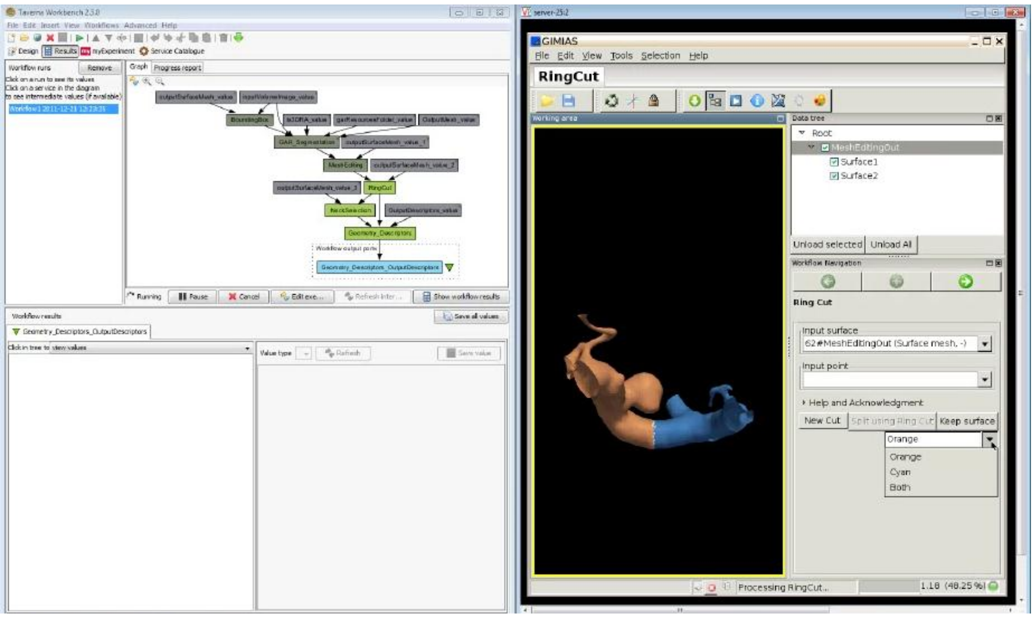

Figure 2. @neurIST workflow execution example.

4. Conclusion

From a user's perspective, ease of use is the major achievement of the composition and migration of the @neurIST workflow to the cloud. Executing the @neurIST workflows does not need any complex installation or configuration of scientific software, only standard software components such as a Web browser or an arbitrary VNC client (for interactive services) is required. From the system's perspective Cloud-enabling the @neurIST workflow with Taverna comes with many advantages. For instance, the scientific maintenance is improved by the use of standardized and re-usable workflow-definitions. Another example is real scalability, which is achieved by instantiating the workflow on a per user basis.

Although the initial experiences of executing this workflow in the Cloud have been promising and experts provided positive feedback, some potential for improvement has also been identified. Concerning the usability an issue with respect to latency has been observed, which hampers interactions and requires an in depth investigation to understand the actual cause of the latency (network, virtual machine resources, graphics processing, etc.). However, more user-feedback, which is expected to increase after integration of the VPH-Share security model, will help to improve the @neurIST workflow in the Cloud in future.

Acknowledgments

The authors would like to thank the VPH-Share (FP7-ICT-2009-6-269978) consortium partners for their inputs to the work presented in this paper.

References

1. Fenner JW, Brook B, Clapworthy G, Coveney PV, Feipel V, Gregersen H et al. The EuroPhysiome, STEP and a roadmap for the virtual physiological human. Phil. Trans. R. Soc. A. 2008; 366(1878):2979-2999. DOI: 10.1098/rsta.2008.0089
2. Hunter P, Coveney PV, de Bono B, Diaz V, Fenner J, Frangi AF et al. A vision and strategy for the virtual physiological human in 2010 and beyond. Phil. Trans. R. Soc. A. 2010; 368(1920):2595-2614. DOI: 10.1098/rsta.2010.0048
3. Villa-Uriol MC, Berti G, Hose DR, Marzo A, Chiarini A, Penrose J et al. @neurIST complex information processing toolchain for the integrated management of cerebral aneurysms. Interface Focus . 2011; 1:308–319. DOI: 10.1098/rsfs.2010.0033

4. VPH-Share, Virtual Physiological Human: Sharing for Healthcare. [Internet] 2012 [cited 2012 June 28]. Available from: http://www.vph-share.eu
5. Taverna - open source and domain independent Workflow Management System [Internet] 2012 [cited 2012 June 28]. Available from:http://www.taverna.org.uk