A Conception Process for Abstract Workflows: An Example on Deep Water Oil Exploitation Domain[i]

Wallace Pereira[1]
Eduardo Ogasawara[1]
Daniel de Oliveira[1]
[1]*Federal University of Rio de Janeiro Computer Science Dep.*

Fernando Chirigati[1]
Fabrício Correa[2]
Breno Jacob[2]
[2]*Federal University of Rio de Janeiro Civil Engineering Dep.*

Ismael Santos[3]
Guilherme H. Travassos[1]
Marta Mattoso[1]
[3]*Petrobras Brazilian Oil Company*

Experimentation is one of the ways used to support theories based on a scientific method. In-silico experiments are highly dependent of massive use of computational resources to execute their simulations [1,2].

One way to support in-silico experiments is through the use of scientific workflows. It is a model that represents the flow of programs, services and data [3] usually orchestrated to support a simulation. Scientific workflows are executed in engines called Scientific Workflow Management Systems (SWfMS), which are responsible for enacting, controlling and monitoring the workflow [1]. Each one of the scientific workflows within an experiment follows specific phases regarding composition, execution and analysis. Usually, when conducting a scientific experiment, the first phase to be considered is called *Composition* [3]. One important sub-phase is the *Conception*, which is responsible for setting up the experiment.

Many initiatives suggest that composition should be done by abstracting the several concepts usually used in the scientific workflows [4, 5]. An abstract workflow is the chaining of activities that describes what should be done, but without saying how. In the next abstraction level, when a scientific workflow is ready to be executed, it is called a concrete workflow [1]. In this context, scientific experiments based on computational simulations introduced new concerns for the scientists about their prior organization, modelling and specification. In general, the scientific experiments are specified as scientific workflows directly bound to the computational resources required for their implementation. So, the experiments composition directly in SWfMS becomes difficult for the scientists and may result in risks to their experiments [4]. However, most of the scientists are not computer experts and they may have some difficulties to deal with infra-structure issues and so composition phase may become a barrier to build more sophisticated models.

To deal with these issues, we propose an approach for the conception sub-phase. This work foresees that such approach may aid scientists to facilitate the modeling process of the scientific experiment. It is also expected that the effective usage of a conception process based on standard software engineering approach may bring many benefits, such as the documentation quality increase and decrease of the incidence of problems. Our approach uses the UML activity diagram and tailored *requirement forms*. The activity diagram has the advantage of being a standard model, frequently used in software industry and academy. We have successfully applied this approach to real workflow focused on evaluating risers fatigue in oil platforms.

[1] E. Deelman, D. Gannon, M. Shields, and I. Taylor, 2008, Workflows and e-Science: An overview of workflow system features and capabilities, *Future Generation Computer Systems (Jul.)*.

[2] G.H. Travassos and M.O. Barros, 2003, Contributions of In Virtuo and In Silico Experiments for the Future of Empirical Studies in Software Engineering, In: *Proc. of 2nd Workshop on Empirical Software Engineering the Future of Empirical Studies in Software Engineering, Roma*.

[3] E. Ogasawara, C. Paulino, L. Murta, C. Werner, and M. Mattoso, 2009, Experiment Line: Software Reuse in Scientific Workflows, In: *Proceedings of the 21th international conference on Scientific and Statistical Database Management*, p. 264–272, New Orleans, LA.

[4] C.B. Medeiros, J. Perez-Alcazar, L. Digiampietri, J. G. Z. Pastorello, A. Santanche, R.S. Torres, E. Madeira, and E. Bacarin, 2005, WOODSS and the Web: annotating and reusing scientific workflows, *SIGMOD Record*, v. 34, n. 3, p. 18-23.

[5]M.C. Cavalcanti, R. Targino, F. Baião, S.C. Rössle, P.M. Bisch, P.F. Pires, M.L.M. Campos, and M. Mattoso, "Managing structural genomic workflows using web services," Data Knowl. Eng., vol. 53, 2005, pp. 45-74.

[i] Partially funded by CNPq and Petrobras