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

Wallace Pereira[1], Eduardo Ogasawara[1], Daniel de Oliveira[1], Fernando Chirigati[1], Fabrício Correa[1], Breno Jacob[1], Ismael Santos[2], Guilherme H. Travassos[1], Marta Mattoso[1]
[1] Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
[2] Petrobras - Brazilian Oil Company, Brazil

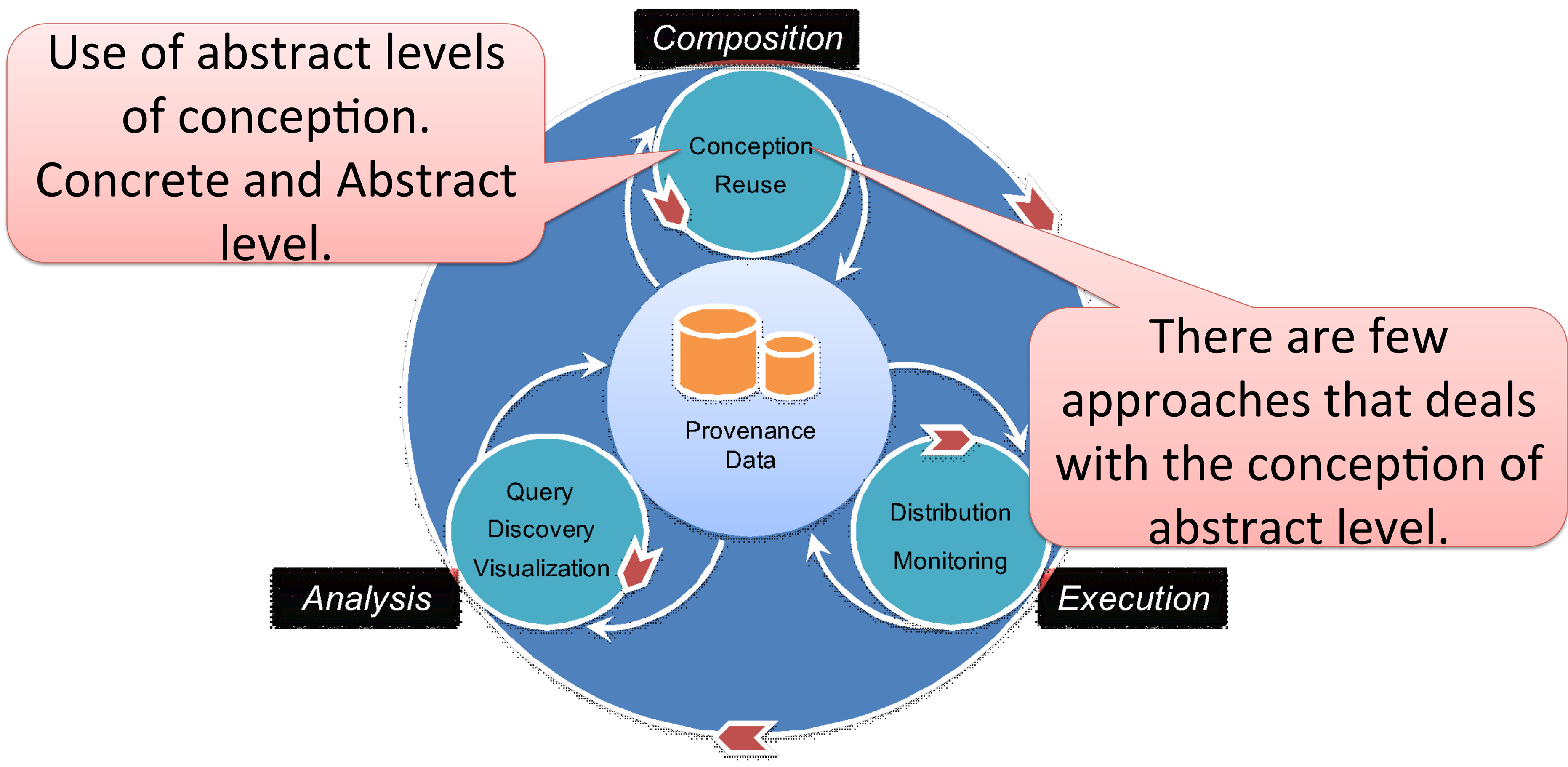
UFRJ COPPE

Federal University of Rio de Janeiro

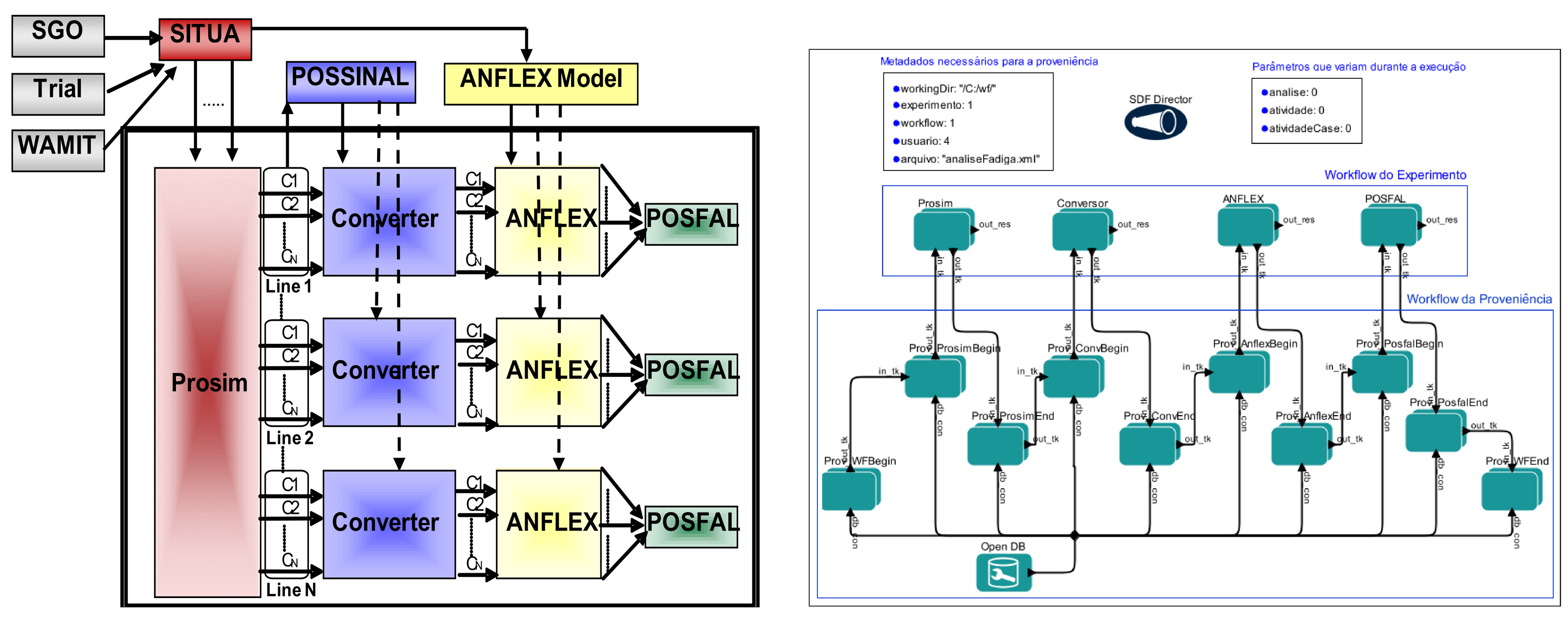
Introduction

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, which may be modeled by scientific workflows. Scientific Workflows are executed in engines called Scientific Workflow Management Systems (SWfMS), which are responsible for their control and monitoring. Usually, when conducting a scientific experiment, the first phase to be considered is called Composition. One important sub-phase of the Composition is the Conception, which is responsible for setting up the experiment. Today, scientists have to deal with a huge amount of information and constraints, and without a systematic approach this phase may become a barrier for scientists in modeling scientific experiments.

Experiment Life Cycle



Current Approaches

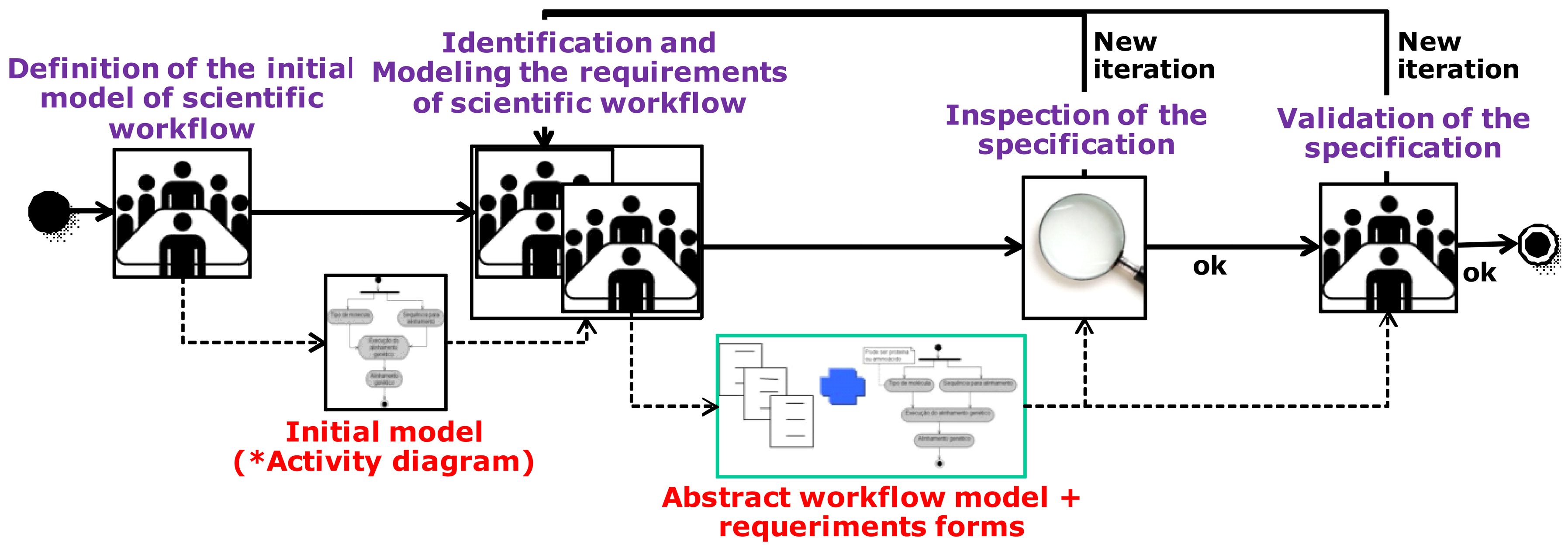


(a) (b)

(a) Models without patterns: It may lead to lower expressiveness and possible inconsistencies between abstract and concrete level.

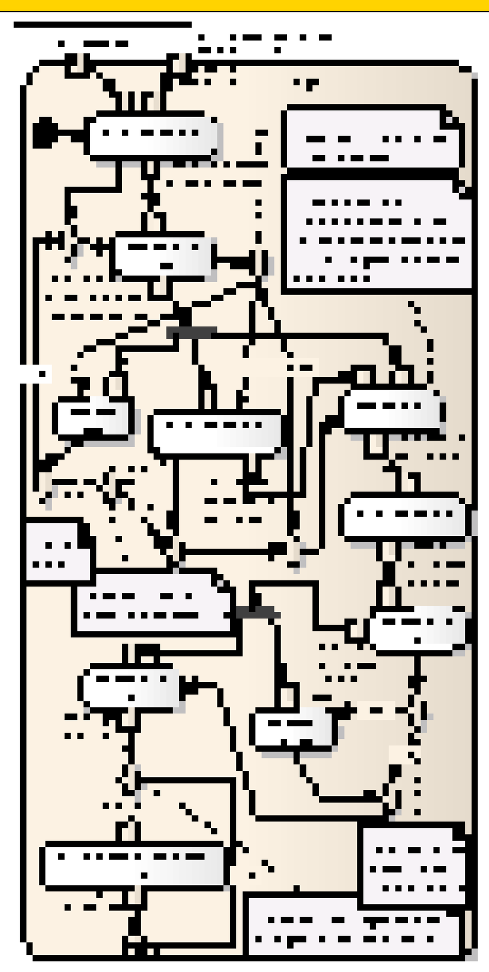
(b) Models directly at concrete level: Too much information of concrete level hides the notion of the experiment and it may turn difficult the replication in another infra-structure.

Process for requirements elicitation



We propose an approach for the conception sub-phase, where it aid scientists in 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 increase of the documentation quality and the decrease of the incidence of problems.

The instruments

(a) 

(b)

Name	[Name of the activity.]
Description	[Detailed description of the activity, including steps for execution.]
Type of activity	[Automated, Semi-automated or Manual.]
Tool	[List of the tools that support the execution.]
Input	[List of all artifacts used in the execution.]
Output	[List of all artifacts generated in the execution.]
Pre-condition	[List of the necessary pre-conditions to start the execution of the activity.]
Post-Condition	[List of the post-conditions that are generated at the end of the execution of activity.]
Mandatory	[Mandatory or Optional. List of the cases where it is optional.]
Roles	[List of the roles played by one or more persons to carry out the activity.]
Pre-activities	[List of the activities that have been completed in order to execute it.]
Risks	[List of the risks associated with the activity, which may led to loss of information or reliability of the activity]

(a) The activity diagram (UML 2.0) has the advantage of being a standard model, frequently used in software industry and academy. It is used for represent the **initial model** and the **abstract workflow model**.

(b) The requirement forms that represents 3 elements of scientific workflow: Artifact, Activity and Tool. Each one of these are composed by an individual set of fields that represents them.

Conclusions

We proposed a systematic process for conception of scientific abstract workflows. By using this approach it is expected that most of the requirements and relevant information about the in silico experiment can be identified. After specifying the abstract model, the scientists should be able to develop concrete workflows based on the abstract workflow definition, considering the variant and optional activities defined in the forms. This process, when executed manually, is usually laborious and error prone. Our research group at Federal University of Rio de Janeiro is already developing tools and methodologies to facilitate this process for the scientists, using an innovative approach known as experiment lines [1]. Once the requirement forms and the activity diagram capture the tacit knowledge of the scientists to an explicit model, it can be used as the input to model an experiment line. We have successfully applied this approach to real workflow focused on evaluating risers fatigue in oil platforms.

[1] E. Ogasawara, C. Paulino, L. Murta, C. Werner, and M. Mattoso, “Experiment Line: Software Reuse in Scientific Workflows,” SSDBM 2009, New Orleans: Springer-Verlag, 2009, pp. 264–272.

This work was partially sponsored by CNPq through grant MCT/CNPq/CT-INFO n° 007/2007 process no. 551013/2007-9 and Petrobras

CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico

IEEE

Visit our site : http://gexp.nacad.ufrj.br/

BR PETROBRAS

e-science 2009