11th World Congress on Computational Mechanics (WCCM XI)
5th European Conference on Computational Mechanics (ECCM V)
6th European Conference on Computational Fluid Dynamics (ECFD VI)
July 20 - 25, 2014, Barcelona, Spain

USER INTERACTION IN UNCERTAINTY QUANTIFICATION ANALYSIS WORKFLOWS

Jonas Dias[1], Gabriel Guerra[2], Fernando Rochinha[2], Alvaro L. G. A. Coutinho[3], Patrick Valduriez[4] and Marta Mattoso[1]

[1] Computer Science, COPPE/Federal University of Rio de Janeiro, PO Box 68511, Rio de Janeiro, RJ Brazil, {jonas, marta@cos.ufrj.br}, www.cos.ufrj.br
[2] Mechanical Engineering, COPPE/Federal University of Rio de Janeiro, PO Box 68503, Rio de Janeiro, RJ Brazil, {gguerra, faro@mecanica.ufrj.br}, www.mecanica.ufrj.br
[3] High Performance Computing Center and Civil Engineering, COPPE/Federal University of Rio de Janeiro, PO Box 68506, Rio de Janeiro, RJ Brazil, alvaro@nacad.ufrj.br}, www.nacad.ufrj.br
[4]Antenne INRIA, LIRMM, 95 rue de la Galera, 34095 Montpellier Cedex 5, France, patrick.valduriez@inria.fr, http://www-sop.inria.fr/members/Patrick.Valduriez/

Key Words: *Uncertainty quantification, Scientific workflow management, User interaction, E-science.*

Uncertainty quantification (UQ) enables the definition of confidence intervals in predictions by stressing the numerical model with varying inputs to perform very large data exploration. The choice of which slice of the input parameter space to explore in a UQ workflow impacts the amount of data produced, which can be very large [1]. A single data configuration exploration workflow execution currently may produce Terabytes of data and can easily take more than a week to run in a parallel computer, e.g. a cluster with several hundreds (thousands) of multicore execution nodes. In order to avoid wasting time and storage, UQ scientists usually start the workflow with a modest configuration. If the outcome is below a given quality criteria, they change the exploration and run the workflow again. Thus, scientists resubmit the execution of a UQ workflow by changing the input configuration until the result meets their expectations. The problem of manually managing the iterative process is that the execution restarts over and over and the UQ analyst may also lose track of what has already been explored and how the UQ workflow evolved. To improve this iterative experimental process, the user should be able to analyze partial results during execution to dynamically interfere in the next steps of the workflow, instead of interrupting and resubmiting the workflow [2]. To tackle the iterative nature of uncertainty quantification analysis, dynamic workflows are necessary. However, dynamic behavior has been identified as an open challenge as workflows are subject to continuous adaptation and improvement [3]. In particular, they require the ability of adapting a scientific workflow, at runtime, based on external events such as human interaction and dynamic steering. A typical data-centric runtime change would be adjusting the data input configuration such as filters or an interpolation level limit, based on the runtime data behavior of the workflow execution. In this paper, we explore workflow algebraic operators to support data-centric iteration in dynamic workflows. We evaluate our dynamic execution model for these operators using a real large-scale UQ workflow running on a 640-core cluster. By allowing partial workflow result analysis, the scientists could monitor several execution parameters like converged error, average and variance of kinetic energy. By querying this information, they were able to conclude whether the simulation converged (before the predefined limit) or stopped after a different threshold. The results show execution time savings from 2.5 to 24 days when compared to a traditional non-iterative workflow execution. We also perform complex queries for partial result analysis along the iterations and we assess the max overhead introduced by our iterative model as 3.63% of execution time. The performance of our proposed steering algorithms run in less than 1 millisecond in the worst-case scenario we measured.

REFERENCES

[1] G. Guerra, F. A. Rochinha, R. Elias, D. de Oliveira, E. Ogasawara, J. Dias, M. Mattoso, A. L. G. A. Coutinho, Uncertainty Quantification in Computational Predictive Models for Fluid Dynamics Using Workflow Management Engine. *International Journal for Uncertainty Quantification*, vol. **2** (1), pp. 53–71, 2012.

[2] M. Mattoso, J. Dias, D. Oliveira, K. Ocaña, E. Ogasawara, F. Costa, F. Horta, V. Silva, I. Araújo, User-Steering of HPC Workflows: State of the Art and Future Directions. *Proceeding of the 2nd International Workshop on Scalable Workflow Enactment Engines and Technologies*, New York, NY, USA, 2013.

[3] Y. Gil, E. Deelman, M. Ellisman, T. Fahringer, G. Fox, D. Gannon, C. Goble, M. Livny, L. Moreau, J. Myers, Examining the Challenges of Scientific Workflows. *Computer*, vol. **40** (12), pp. 24–32, 2007.