Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A.A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

INTEGRATING SCIENTIFIC WORKFLOW TO SCIENCE GATEWAY: AN EXPLORATION OF SCIENCE GATEWAY ARCHITECTURE FOR DISTRIBUTED SIMULATION

Chukwudi Nwogu
Simon J. E. Taylor
Anastasia Anagnostou

Modelling & Simulation Group
Brunel University
London, UB8 3PH, UK

ABSTRACT

The modelling and simulation (M&S) community is faced with the complications of developing distributed simulation (DS) and gaining access to distributed computing infrastructure (DCI) for speeding up simulation experimentations and manipulating the resultant data output. There is also the challenge of widely disseminating DS modules or their outputs to support decision-making and collaboration within operational research and management science (OR/MS). Science gateways (SGs) could alleviate these challenges. Therefore, this work will review the architectures of integrating scientific workflows (SWF) to SGs, as a means of adopting a suitable architecture for SG for DS.

1 INTRODUCTION

Distributed Simulation (DS) is the approach of performing multiple simulations on multiple systems that are interconnected through a network (Anagnostou and Taylor 2017). DS is faced with many challenges, including access to distributed computing infrastructures (DCIs). Science gateways (SGs) could be employed to give seamless access to DCIs, such as cloud, grid, etc., by providing a graphical user interface (GUI) which hides the underlying complexity of the DCI technologies (Gesing et al. 2018). Glatard et al. (2017) note that access to SGs are via a web interface that enables user manage access rights, and execution of tasks. The authors report that most SGs consist of one or more workflow engines; while Zhao et al. (2014) suggest that in many cases, large-scale simulations are implemented as SWFs that run on DCIs. Therefore, the architectures for integrating SWF to SG will be reviewed, in the next section, to determine a suitable architecture for SG for DS. SWF and workflows will be used interchangeably in this work.

2 RELATED WORK

There is limited literature on architectures of SG for DS. Yu and Buyya (2006) propose a generic architecture for grid workflow, but not in relation to an SG. Conversely, the SG canvas—a business reference model—proposed by Shahand et al. (2015) does not deal with the architecture of SG for DS. Glatard et al. (2017) summarize the architectures to integrate workflow engines with SG, as follows: tight integration; service invocation; task encapsulation; pool model; nested workflows; and workflow conversation. Glatard et al. (2017) conclude that task encapsulation, and workflow conversion performed better than other architectures; while nested workflow performed worse. The authors also state that nested architecture is the only architecture that has the capability to support distributed workflow (DW) execution, and also supports meta-workflow—one of the main features of the nested architecture. The same study notes that DW engines consist of meta-workflows and have the ability of executing a single workflow among different engine instances. It refers to DW as a workflow that has tasks that are loosely-coupled and can be executed on geographically distributed DCIs.

3 ARCHITECTURE FOR SCIENCE GATEWAY FOR DISTRIBUTED SIMULATION—CHOICE & JUSTIFICATION

From the literature review, nested architecture appears to be the most suitable architecture for SG for DS. In order to justify this choice, we will put some of the key points from the literature into perspective. Firstly, we chose to evaluate the workflow architectures for SG, since most SGs consist of one or more workflow engines, on which they rely to execute tasks. Secondly, large-scale simulations are implemented as scientific workflows that run on DCIs, and DS is the decomposition of large-scale simulation into smaller simulation modules that are interconnected in a loosely-coupled manner. The same principle applies to DW, which are found to consist of loosely-coupled meta-workflows. Therefore, we conclude that DS could be implemented as DWs on SGs. Thirdly, nested architecture is the only workflow architecture that supports DW and also supports meta-workflow.

4 METHODOLOGY

This research will adopt an empirical approach. DS experimentation artefacts, that are based on a case study, will be developed in JaamSim (a Java-based open source simulation package), and integrated to an SG using the nested architecture. Data from the simulation runs will be collected and analyzed, and the results will be validated by comparing with the results from the literature, in line with (Wieringa, 2014).

5 CONCLUSION & FUTURE WORK

We have inferred that DS could be integrated to SG in the same way as SWF. Further work will be undertaken to integrate DS experimentation artefacts with SG and result validated against real-life data.

REFERENCES

Anagnostou, A., and S. J. E. Taylor. 2017. "A Distributed Simulation Methodological Framework for OR/MS Applications". *Simulation Modelling Practice and Theory* 70:101–117.

Gesing, S., R. Dooley, M. Pierce, J. Krüger, R. Grunzke, S. Herres-Pawlis, and A. Hoffmann. 2018. "Gathering Requirements for Advancing Simulations in HPC Infrastructures via Science Gateways". *Future Generation Computer Systems* 82:544–554.

Glatard, T., M.-É. Rousseau, S. Camarasu-Pop, R. Adalat, N. Beck, S. Das, … and A. C. Evans. 2017. "Software Architectures to Integrate Workflow Engines in Science Gateways". *Future Generation Computer Systems* 75:239–255.

Shahand, S., A. H. C. van Kampen, and S. D. Olabarriaga. 2015. "Science Gateway Canvas". In *Proceedings of the 1st Workshop on The Science of Cyberinfrastructure Research, Experience, Applications and Models*, 45–52. New York, NY: ACM Press.

Wieringa, R. 2014. "Empirical Research Methods for Technology Validation: Scaling Up to Practice". *Journal of Systems and Software* 95:19–31.

Yu, J., and R. Buyya. 2006. "A Taxonomy of Workflow Management Systems for Grid Computing". *Journal of Grid Computing* 3–4(3):171–200.

Zhao, Y., Y. Li, I. Raicu, S. Lu, and X. Zhang. 2014. "Architecting Cloud Workflow: Theory and Practice". In *Proceedings of the 2014 IEEE International Conference on Computer and Information Technology*, 466–473. IEEE.