THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

SE4HPCS'15

OPEN ACCESS

SE4HPCS'15: The 2015 International Workshop on Software Engineering for High Performance Computing in Science

Jeffrey C. Carver*, Neil Chue Hong†, Paolo Ciancarini‡

*University of Alabama, USA, carver@cs.ua.edu

†Software Sustainability Institute, University of Edinburgh, UK, N.ChueHong@software.ac.uk

‡University of Bologna, Italy, paolo.ciancarini@unibo.it

Abstract—HPC software is developed and used in a wide variety of scientific domains including nuclear physics, computational chemistry, crash simulation, satellite data processing, fluid dynamics, climate modeling, bioinformatics, and vehicle development. The increase in the importance of this software motivates the need to identify and understand appropriate software engineering (SE) practices for HPC architectures. Because of the variety of the scientific domains addressed using HPC, existing SE tools and techniques developed for the business/IT community are often not efficient or effective. Appropriate SE solutions must account for the salient characteristics of the HPC, research-oriented development environment. This situation creates a need for members of the SE community to interact with members of the scientific and HPC communities to address this need. This workshop facilitates that collaboration by bringing together members of the SE, the scientific, and the HPC communities to share perspectives and present findings relevant to research, practice, and education. A significant portion of the workshop is devoted to focused interaction among the participants with the goal of generating a research agenda to improve tools, techniques, and experimental methods regarding SE for HPC science.

Index Terms—Software Engineering, Computational Science, Computational Engineering, High Performance Computing

I. INTRODUCTION

This workshop is concerned with identifying and understanding the unique aspects of Software Engineering (SE) for the development of scientific software that exploits High Performance Computing (HPC) architectures. This class of software applications includes software developed to support various scientific endeavors which would be difficult or impossible to perform experimentally. This type of software development has not received enough attention from the SE community. Members of these communities need to better understand the strengths for the others and identify areas of common interest and need which can be pursued.

Despite its importance, the development of scientific HPC software historically has attracted less attention from the SE community than other subdomains have. Indeed, the development of this type of software is significantly different than the development of business information systems, from which many of the SE best practices, tools and techniques have been drawn. Therefore, in order to identify and develop appropriate methods, tools and techniques for scientific HPC software, members of the SE community must interact with members of the HPC and science communities to understand the differences and determine the most appropriate SE tools, methods and techniques.

New researchers are coming into this line of research and are often unaware of each other's work. There is no one preferred journal for publication or other readily found source for researchers with this common interest. The SE4HPCS workshop represents the kind of interdisciplinary work that is critical to the continued vitality of the SE discipline. One of the main objectives is to provide an opportunity for these scattered communities of researchers to coalesce into a single community. The intersection of SE, computational science and HPC is particularly interesting because of the unique characteristics of the research environment which impact the choice of SE practices. Those characteristics include:

- The exploration of unknown science implies that many of the requirements (beyond the laws of nature) cannot be known *a priori* and must emerge over time.
- Investigation of new scientific results implies that the software's expected output is sometimes unknown, making it difficult or impossible to define test oracles and to use traditional software testing approaches.
- Development of software for HPC platforms introduces an additional level of complexity.
- Project lifecycles are likely to differ from business/IT software, both in terms of the steps involved as well as the duration.
- Many scientific software developers are domain experts, but have little formal training in SE methods, tools and techniques.

The problem of matching SE tools and techniques to HPC architectures for scientific purposes is of great interest and importance. In addition, there is an increasing amount of attention being given to this effort including a US Department of Energy workshop on Software Productivity for Extreme Scale Science[1], a US Department of Energy Funding Opportunity[2], and the European PRACE research infrastructure[3].

---

[1] http://www.orau.gov/swproductivity2014/

[2] http://science.energy.gov/~/media/grants/pdf/lab-announcements/2014/LAB_14-0001.pdf

[3] http://www.prace-ri.eu

II. Workshop Goals

Most conference and journal venues focus either on SE, on Computational Science, or on HPC, but rarely on the intersection of the three domains. Specifically, within the science and HPC communities, there are few places to publish results related to the unique SE challenges faced by developers adressing scientific challenges using HPC platforms. The goal of this workshop is to provide a useful venue for researchers from SE, science and HPC to discuss issues relevant to the intersection of these three fields. The workshop provides an opportunity for members of these three groups to interact when they normally do not have such opportunities. By bringing these groups together, our goal is to support the building of a common research agenda to deal with the complex software development issues typical of scientific HPC architectures.

There are very few consistent venues in which to publish this type of important research. Recent endeavors to bring the SE and CSE communities together include two special issues of *IEEE Software* [5], [7], four special issues of *Computing in Science and Engineering* [1]–[4], a special issue of the *Journal of Organizational and End User Computing* [6] and an upcoming special issue of *Scientific Programming*. In addition, this workshop series (http://SE4Science.org/workshops) brings together computational scientists, software engineering researchers and software developers to explore issues such as:

- Those characteristics that distinguish scientific HPC software from general business/IT software;
- The different contexts in which scientific HPC developments take place;
- The quality goals of scientific HPC software and software- or data-intensive research;
- How the perceived chasm between the scientific and the SE communities might be bridged.

III. Plans for Discussion at the Workshop

After a thorough review process, we accepted 6 *full papers* (8 pages) and 4 *position papers* (5 pages). The 2015 workshop is organized like the previous workshops. In the morning, paper authors will present their work to help motivate and frame the afternoon discussions. Full papers, which will have longer presentations, report on more mature research. Position papers, which will have shorter presentations, will describe early stage research or interesting new perspectives. The following papers are included in the workshop:

- *Development of Scientific Software for HPC Architectures using OpenACC: The Case of LQCD*
- *A Scientific Function Test Framework for Modular Environmental Model Development: Application to the Community Land Model*
- *Commit Quality in Five High Performance Computing Projects*
- *Regression Testing of GPU/MIC Systems for HPCC*
- *SIMPL: A Patter Language for Writing Efficient Kernels on GPGPU*
- *Supporting Scientists in Re-engineering Sequential Programs to Parallel Using Model-driven Engineering*

- *POSITION PAPER: Ensuring an Effective User Experience When Managing and Running Scientific HPC Software*
- *POSITION PAPER: Using Software Engineering Methodologies to Port a Scientific Code to GPUs*
- *POSITION PAPER: Towards an Engineering Methodology for Multi-Model Scientific Simulations*
- *POSITION PAPER: Computation for Genomics Knowledge Discovery*

With these presentations as context, the afternoon will be spent in small and large group discussions focused on the important themes that emerge during the presentations. Historically, the group discussions are the most interesting and important aspect of the workshop because they allow members of the different communities to interact on a more direct basis. In constructing the discussion groups, we will seek to balance the representation from the SE, HPC and scientific communities to provide for more balanced and fruitful discussions.

The 2015 workshop will build upon the themes from previous workshops and identify new topics of interest. Some of the discussion themes that emerged during previous workshops include: 1) The characteristics of research software that affect its development choices; 2) The appropriate context dimensions to describe research software; 3) The major software quality goals for research software to run on HPC architectures; 4) Crossing the communication chasm between SE and Science when experiments are software or data-intensive; 5) Effectively involving scientists in software development and training; 6) Measuring the impact of SE on scientific productivity; 7) SE tools and methods needed by the scientific community; and 8) How to effectively test scientific HPC software.

Acknowledgments

References

[1] J. C. Carver, "First international workshop on software engineering for computational science & engineering," *Computing in Science Engineering*, vol. 11, no. 2, pp. 7–11, 2009.

[2] ——, "Report: The second international workshop on software engineering for cse," *Computing in Science Engineering*, vol. 11, no. 6, pp. 14–19, 2009.

[3] ——, "Software engineering for computational science and engineering," *Computing in Science & Engineering*, vol. 14, no. 2, pp. 8–11, 2011.

[4] J. C. Carver and T. Epperly, "Software engineering for computational science and engineering [guest editors' introduction]," *Computing in Science and Engineering*, vol. 16, no. 3, pp. 6–9, 2014.

[5] J. Segal and C. Morris, "Guest editors' introduction: Developing scientific software, part 2," *IEEE Software*, vol. 26, no. 1, pp. 79–79, 2009.

[6] J. Segal, D. Kelly, and J. C. Carver, "Special issue on scientific end user computing," *Journal of Organizational and End-User Computing*, vol. 23, pp. i–iv, 2009.

[7] J. Segal and C. Morris, "Developing scientific software," *IEEE Software*, vol. 25, no. 4, pp. 18–20, 2008.