Sustainable & Productive: How CSE Funding Agencies and Publishers Can Help

Michael A. Heroux, Sandia National Laboratories, St. John's University
maherou@sandia.gov

Introduction

Funding agencies and publishers play an essential role in determining expected behavior from the computational science and engineering (CSE) research community. Publishers use a peer review process to ensure that published content is accurate, complete and relevant. Funding agencies place expectations on how research is conducted, requiring, for example, that research teams perform work within certain ethical norms.

Recently, there have been a number of initiatives to place increased expectations on CSE software activities. A few examples are:

- The ACM Transactions on Mathematical Software (TOMS) initiated a Replicated Computational Results (RCR) review process during which a dedicated reviewer is tasked with replicating the computational results of a submitted manuscript.
- The US National Science Foundation (NSF) has a relatively new program, SI2, which emphasizes sustainable software development.
- NSF and the US Department of Energy have recently required that proposals have a data management plan (DMP), for the purposes of assuring that software and data are preserved from loss and openly available.

Development of further CSE software requirements from these clients can provide natural incentives for CSE software teams to improve sustainability and productivity, and spur innovation in these areas.

Creating Incentives for Sustainability and Productivity

Expectations from CSE software clients provide a means to give intrinsic value to virtuous behaviors that are otherwise difficult to reward. CSE software teams certainly want to do a good job of being productive and producing sustainable software tools. But the drive to produce science results on a short time scale often leads to shortcuts and lack of investment up front that is necessary for sustainable software and long-term productivity. Given the demands to provide a computational answer as soon as possible, especially in a competitive environment, other activities such as careful design, efficient implementation and thorough software testing environments are often compromised.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced

The connection between these clients (publishers and funding agencies) and science teams who use computation as part of their scientific endeavors can be further leveraged to provide intrinsic value to other activities that can improve sustainability and productivity.

A Time of Disruptive Change

Presently the CSE community is in an era of disruptive changes. System architecture changes, driven by the stagnation of hardware clock speeds, are resulting in exponential increases in available concurrency such that performance sensitive applications must expose proportional levels of concurrency in order to stay on the new commodity curves of thread scalability and vector/SIMT availability.

Furthermore, the tremendous growth in performance potential leads to increased opportunities for multi-scale and multi-physics coupling. In turn, coupling demands require increased collaboration across historical separate efforts. Projects will become more inter-disciplinary, distributed (since areas of expertise are typically not all in one group or location), and larger.

Taking advantage of these trends, CSE projects will benefit from more deliberate efforts to improve software sustainability and productivity, and we can accelerate exploration of effective approaches by putting in place explicit requirements from publishers and funding agencies.

Toward Sustainable and Productive

Numerous studies have shown that activities such as code review and test driven development are universally valuable. Furthermore, explicit life cycle models, including models for refactoring legacy codes, are also extremely beneficial, especially for larger teams. Productivity models, which attempt to characterize and measure the value and cost attributes of a project, can also be extremely helpful, especially in the face of disruptive changes in computing platforms and application requirements. At the same time, these productivity improvements are often not adopted by CSE software teams because other priorities are ranked higher.

In a similar way to the NSF and DOE requirements for a DMP, funding agencies can require that proposals contain a description of their testing plan, software lifecycle or productivity model. The requirement that a proposal contains these elements—without being specific about how to satisfy the requirement—will spur innovation. CSE software teams will have a natural incentive to effectively and efficiently satisfy these requirements because doing so will improve their competitive advantage.

The CSE software community can benefit greatly from an improved focus on sustainability and productivity, but for many teams the incentive for investment is not strong enough when compared to other demands. Funding agencies and publishers can play a key and truly unique role by increasing quality expectations that will naturally provide incentives for improved sustainability and productivity.