What US DOE Labs Look for in Trained Computational Scientists

Originally presented as a position paper for, SIAM-EESI workshop on
"Future Directions in CSE Education and Research"

Carol S. Woodward and Jeffrey A. F. Hittinger, Lawrence Livermore National Laboratory

The United States Department of Energy Laboratories conduct large numbers of simulations in numerous scientific areas. To support this effort, these laboratories have a strong need for trained computational scientists with strong multidisciplinary training in applied mathematics, computer science, and scientific or engineering domain science. In this paper, we overview attributes that we look when we hire for new graduates for postdoctoral and staff research positions. While we rarely find candidates who exhibit all of these attributes, we do judge candidates on all of them. Many of these attributes are not attained through traditional coursework. However, many of these skills can be developed through course project requirements, and we recommend that these skills be incorporated into the CSE curriculum.

Deep knowledge in a specific area: Researchers should show ability to think though all details of a problem, including theory and computational work. The point here is that once a researcher has thought through details of a given topic, they are more able to do so for new topics.

Science knowledge: Researchers should have a working knowledge of 2-3 science areas. This knowledge should include the scientific motivation for computational modeling and simulation. In addition, researchers must understand computational needs and use patterns of each application area for algorithms and computing architectures. A working knowledge of 2-3 areas ensures that researchers have some understanding of common issues that arise in simulation and the different vocabulary that fields employ for similar concepts.

Methods knowledge: Researchers should have a strong understanding of the convergence and accuracy of numerical methods used in the application areas in which they work. Their goal should be to ensure that algorithms balance the needs of performance, robustness, and accuracy. In addition, as a community, we need to ensure efficient use of our large-scale machines, and well-designed algorithms greatly aid in that endeavor.

Coding knowledge: Because we often work with high performance computers and need to maximize performance of these machines to solve problems of interest, researchers at DOE labs must be able to write code in compiled languages where performance optimizations are more feasible and where there is less reliance on outside vendors who may not have products for these machines.

Software Engineering: Since the DOE makes a large investment in software, it is common that once a code is used successfully, it will become part of a long-term code base. As a result, we prefer people who understand software engineering concepts, particularly as they relate to high performance computing. This understanding includes algorithms and data structures (including complexity), computer architecture, software architecting and design (including concepts of modularity and paradigms such as procedural, object-oriented, and functional), parallel programming models and languages, and software quality assurance concepts. This skillset should include the ability to write portable and well-documented software that other researchers can understand.

Computer architecture: DOE labs often are the first purchasers of new HPC architectures. As a result, our researchers need to be able to translate implementations of applications and algorithms to produce performance on a given architecture with multiple processor nodes, a deep memory hierarchy, and specific interconnect properties. Computational scientists at the lab need at least a basic understanding of how computer hardware and compilers work so that they can develop highly performant implementations.

Tools: Researchers must be versatile and able to exploit tools in their computational environment, including build systems, source code control systems, visualization systems, performance analysis and memory checking tools, scripting languages, debuggers, and efficient input/output formats.

Social skills: By its nature, CSE is a collaborative discipline, and we expect researchers to be able to interact productively with other researchers from a variety of backgrounds. Multidisciplinary training can be very helpful in this regard because it facilitates better communication between disciplines with different cultures and languages.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC. LLNL-AR-657870.