by Steven Gottlieb
Indiana University

Whither the Future of NSF Advanced Computing Infrastructure?

This issue of *Computing in Science & Engineering* is devoted to heterogeneous high-performance computing (HPC) systems, which are becoming more common at National Science Foundation (NSF) and Department of Energy (DOE) centers, and might represent the only viable future for HPC. To deal with rapid changes in technology and computing needs, a Committee on Future Directions for NSF Advanced Computing Infrastructure to Support US Science in 2017-2020 was recently established. This committee is co-chaired by William Gropp (University of Illinois) and Robert Harrison (Stony Brook), and appointed through the Computer Science and Telecommunications Board of the National Research Council (NRC). The committee has issued an interim report detailing the issues it's studying and inviting community input (www.nap.edu/catalog/18972/future-directions-for-nsf-advanced-computing-infrastructure-to-support-us-science-and-engineering-in-2017-2020). The final report is expected in mid-2015. Unfortunately, this editorial will appear after the committee's deadline of 31 January for giving full consideration to outside input, but interested readers are still encouraged to submit feedback by email to sciencecomputing@nas.edu or via www.nas.edu/sciencecomputing.

The NSF supercomputing centers have a long and productive history. The "Report of the Panel on Large Scale Computing in Science and Engineering" laid the foundation for the centers in 1982 (www.pnl.gov/scales/docs/lax_report1982.pdf). I was reminded of this report in November 2013, when I attended the Kenneth Wilson Memorial Symposium—Wilson was an influential member of the panel. The report noted that "Important segments of the research and defense community lack effective access to supercomputers; and students are neither familiar with their special capabilities nor trained in their use." Of course, we've come a long way since 1982, but the challenges remain.

The NSF provides computing for a *broad* range of science, independent of specific mission. Blue Waters, the NSF flagship supercomputer project, cost over $200 million. The XSEDE program, which is a distributed array of computers and storage systems of varying architecture, now sees demand that greatly exceeds capacity. At the December 2014 allocation meeting, requests for time totaled over a factor of three greater than what was available. The peer review process didn't result in the suggestion that everyone be awarded their requested amount, but the recommended allocations exceeded the available resources by 180 million service units, and even highly recommended NSF-funded proposals were cut substantially. The architecture of advanced systems is getting more complicated, with accelerators such as GPUs and Intel Xeon Phi processors requiring students and older researchers to take further training in the effective use of this new hardware. The NSF has been interested in engaging new communities in HPC, so the problems of capacity and training are only going to grow, requiring significant investments of time and money.

In addition, almost every area of science and engineering now depends on advanced digital resources. Technological progress and economic growth can also be linked to advances in computation. Other countries have already recognized this and are increasing investments in HPC and cyberinfrastructure. However, the NSF budget has basically remained flat.

Remember all the rhetoric in Congress about doubling the NSF budget? Let me refresh your memory as I had to mine from an item by the American Physical

 1521-9615/15/$31.00 © 2015 IEEE Copublished by the IEEE CS and the AIP

Topics on Which the NRC Committee Seeks Input

The National Research Council Committee is seeking comments on the following topics:

1. how to create an advanced computing infrastructure that enables integrated discovery involving experiments, observations, analysis, theory, and simulation;
2. technical challenges to building more capable advanced computing systems and how the National Science Foundation might best respond to them;
3. the committee will review data from the NSF and the advanced computing programs it supports and seeks input, especially quantitative data, on the computing needs of individual research areas;
4. the match between resources and demand for the full spectrum of systems, for both compute- and data-intensive applications, and the impacts on the research community if the NSF can no longer provide state-of-the-art computing facilities;
5. the role that private industry and other federal agencies can play in providing advanced computing infrastructure, including opportunities, costs, issues, and service models, as well as balancing the different costs and making trade-offs in accessibly (for example, guaranteeing on-demand access is more costly than providing best-effort access);
6. the challenges facing researchers in obtaining allocations of computing resources and suggestions for improving the allocation and review processes that make advanced computing resources available to the research community;
7. whether wider collection and more frequent updating of requirements for advanced computing could help inform strategic planning, priority setting, and resource allocation, and if so, how these requirements might be used, developed, collected, aggregated, and analyzed;
8. the tension between the benefits of competition and the need for continuity as well as alternative models that might more clearly delineate the distinction between performance review and accountability and organizational continuity and service capabilities; and
9. how the NSF might best coordinate and set overall strategy for advanced computing-related activities and investments as well as the relative merits of both formal, top-down coordination and enhanced, bottom-up processes.

Society entitled, "Senate Bill to Double NSF Budget" (www.aps.org/publications/apsnews/200210/senate.cfm). The year was 2002, and the bill was introduced by Senator Edward Kennedy (D.-Mass.) just before an August Congressional recess. The House had already passed a similar bill, H.R. 4664. The NSF budget was $4,789.2 million in FY2002, and the Senate bill would have increased funding to $9,839.3 million for FY2007:

> The House bill would reauthorize NSF for fiscal years 2003 to 2005, putting the foundation's budget on track to double in five years by calling for 15% increases in each of the years authorized. The Senate bill, known as the 'National Science Foundation Doubling Act,' is co-sponsored by Senators Ernest Hollings (D-SC), Barbara Mikulski (D-MD), and Christopher Bond (R-MO). The Senate bill would reauthorize NSF through FY 2007 and recommends annual increases of approximately 15.5% in each of these years, more than doubling the foundation's budget by FY 2007.

Fast forward to last April, when the American Institute of Physics reported that "Senators have until this Thursday, April 10 to sign a letter supporting a $7.5 billion budget for the National Science Foundation in FY 2015"′ (www.aip.org/fyi/2014/senators-urged-sign-letter-supporting-75-billion-budget-nsf). The FY2014 NSF budget was $7,171.9 million, and the administration was requesting a 1.2 percent increase. Thus, there was a move in the Senate to increase the NSF budget by 4.6 percent, which is certainly more than the average compounded 3.4 percent annual increase from FY2002 to FY2014. A total of 21 Senators signed the letter. Did Congress pass a budget?

The NRC Committee is investigating several issues detailed in the interim report I mentioned earlier (see the sidebar). I had an opportunity to address the panel producing the report during its 15 December meeting at the Computer History Museum in Mountain View, California. There's insufficient space to recount all that I presented to the panel, but I want to mention one issue that I don't think will make it into the report.

The Division of Advanced Cyberinfrastructure (ACI) is now part of the NSF's Directorate of Computer & Information Science & Engineering (CISE). Previously, the Office of Cyberinfrastructure reported directly to the NSF director. The Office was established after much community input to build the vision for its creation, culminating in the report, "Revolutionizing Science and Engineering through Cyberinfrastructure" (www.nsf.gov/cise/sci/reports/atkins.pdf). I think the former organization better enabled the head of the Office to reach out to other directors to assess what hardware, networking, and services were required to support the research needs of their directorates. This kind of planning is essential to ensure that needs of a broad range of scientists are met. With ACI part of CISE, there's a danger that competing demands from other activities within CISE will result in insufficient funding for the necessary cyberinfrastructure. Of course, the main problem for the NSF is lack of financial support by Congress. One NRC Committee member pointed out how US investments in research and development have fallen, while other nations are increasing such investments.

I hope you will look at the NRC Committee's interim report and offer comments. In the longer term, please contact your Representatives and Senators to let them know how critical it is to increase funding for the NSF. Finally, enjoy this issue on heterogeneous HPC, which is both an opportunity and a challenge. ■

Steven Gottlieb is a distinguished professor of physics at Indiana University, where he directs the PhD minor in scientific computing. He's also an associate editor in chief of *CiSE*. Gottlieb's research is in lattice quantum chromodynamics, and he has a PhD in physics from Princeton University. Contact him at sg@indiana.edu.