ASCAC WORKFORCE SUBCOMMITTEE LETTER

Executive Summary

Simulation and computing are essential to much of the research conducted at the DOE national laboratories. Experts in the ASCR-relevant Computing Sciences, which encompass a range of disciplines including Computer Science, Applied Mathematics, Statistics and domain sciences, are an essential element of the workforce in nearly all of the DOE national laboratories. This report seeks to identify the gaps and challenges facing DOE with respect to this workforce.

The DOE laboratories provided the committee with information on disciplines in which they experienced workforce gaps. For the larger laboratories, the majority of the cited workforce gaps were in the Computing Sciences. Since this category spans multiple disciplines, it was difficult to obtain comprehensive information on workforce gaps in the available timeframe. Nevertheless, five multi-purpose laboratories provided additional relevant data on recent hiring and retention.

Data on academic coursework was reviewed. Studies on multidisciplinary education in Computational Science and Engineering (CS&E) revealed that, while the number of CS&E courses offered is growing, the overall availability is low and the coursework fails to provide skills for applying CS&E to real-world applications. The number of graduates in different fields within Computer Science (CS) and Computer Engineering (CE) was also reviewed, which confirmed that specialization in DOE areas of interest is less common than in many other areas.

Projections of industry needs and employment figures (mostly for CS and CE) were examined. They indicate a high and increasing demand for graduates in all areas of computing, with little unemployment. This situation will be exacerbated by large numbers of retirees in the coming decade. Further, relatively few US students study toward higher degrees in the Computing Sciences, and those who do are predominantly white and male. As a result of this demographic imbalance, foreign nationals are an increasing fraction of the graduate population and we fail to benefit from including women and underrepresented minorities.

There is already a program that supports graduate education that is tailored to the needs of the DOE laboratories. The Computational Science Graduate Fellowship (CSGF) enables graduates to pursue a multidisciplinary program of education that is coupled with practical experience at the laboratories. It has been demonstrated to be highly effective in both its educational goals and in its ability to supply talent to the laboratories. However, its current size and scope are too limited to solve the workforce problems identified. The committee felt strongly that this proven program should be extended to increase its ability to support the DOE mission.

Since no single program can eliminate the workforce gap, existing recruitment efforts by the laboratories were examined. It was found that the laboratories already make considerable effort to recruit in this area. Although some challenges, such as the inability to match industry compensation, cannot be directly addressed, DOE could develop a roadmap to increase the impact of individual laboratory efforts, to enhance the suitability of existing educational opportunities, to increase the attractiveness of the laboratories, and to attract and sustain a full spectrum of human talent, which includes women and underrepresented minorities.

1. Introduction

This letter is ASCAC's response to the charge of February 19, 2014 to identify disciplines in which significantly greater emphasis in workforce training at the graduate or postdoctoral levels is necessary to address workforce gaps in current and future Office of Science mission needs.

ASCAC formed a subcommittee to respond to the charge. It included participants engaged in DOE workforce training efforts, researchers and managers at DOE laboratories, academics in disciplines related to Advanced Scientific Computing, an industry leader in HPC, and a leader of graduate education and training in a major NSF HPC Center. They are as follows:

Barbara Chapman (Chair), University of Houston	Scott A. Lathrop, NCSA, University of Illinois Urbana-Champaign
Henri Calandra, Total SA	Vivek Sarkar, Rice University
Silvia Crivelli, Lawrence Berkeley National Laboratory and University of California Davis	Eric Stahlberg, Advanced Biomedical Computing Center
Jack Dongarra, University of Tennessee	Jeffrey S. Vetter, Oak Ridge National Laboratory
Jeffrey Hittinger, Lawrence Livermore National Laboratory	Dean Williams, Lawrence Livermore National Laboratory
Chris Johnson, University of Utah	

Approach Taken: The committee discussed the charge with Roscoe Giles and Barbara Helland. They solicited information from the DOE national laboratories on their workforce gaps that are significant for the DOE mission and reviewed their responses. They also reviewed reports from many sources, including the National Academies, National Science Foundation, and Computing Research Association. The committee prepared its report in teleconference calls and by email.

Interpretation of the Charge: We focus our attention on **Computing Sciences** disciplines. We use this term throughout the document to cover multiple areas of importance to DOE including, but not limited to, Computational Science and Engineering. It includes fields such as Algorithms (both numerical and non-numerical); Applied Mathematics; Data Analysis, Management and Visualization; Cybersecurity; Software Engineering and High Performance Software Environments; and High Performance Computer Systems.

In addition to specific recommendations for programs at the graduate and postdoctoral level that can address workforce development needs, the committee feels that a much broader perspective must be taken if the identified workforce challenges are to be resolved in the long term. Accordingly, we propose a number of additional, complementary recommendations.

Limitations of this Report: The fields covered by the term Computing Sciences, as defined above, do not fit clearly within the traditional boundaries of academic disciplines and, in some instances, are inherently multidisciplinary. As a result, individual studies and surveys typically do not address the entire breadth of fields in the Computing Sciences nor is Human Resources data readily available for these sciences. Much of the material reviewed during the preparation of this report focused on subsets of the disciplines under consideration. Although we have made an effort to assimilate data in the area of interest, it remains difficult to make definitive statements.

2. Summary of Findings

Under the auspices of the Office of Science, a significant fraction of the research and development undertaken at DOE national laboratories is centered on scientific computation. Therefore, maintaining a sufficient workforce in Computing Sciences is critical if the investment in ASCR facilities is to be realized. This report evaluates the Computing Sciences workforce critical for the Office of Science to meet its scientific mission.

Results of data analyzed are that the Computing Sciences workforce recruitment and retention activities are below the level necessary to sustain ASCR facilities and maintain DOE's high standards of excellence for innovative research and development. In particular, the findings reveal that:

- All large DOE national laboratories face workforce recruitment and retention challenges in the fields within Computing Sciences that are relevant to their mission (termed *ASCR-related Computing Sciences* in the following findings and the recommendations), including Algorithms (both numerical and non-numerical); Applied Mathematics; Data Analysis, Management and Visualization; Cybersecurity; Software Engineering and High Performance Software Environments; and High Performance Computer Systems.
- Insufficient educational opportunities are available at academic institutions in the ASCR-related Computing Sciences that are most relevant to the DOE mission.
- There is a growing national demand for graduates in ASCR-related Computing Sciences that far exceeds the supply from academic institutions. Future projections indicate an increasing workforce gap and a continued underrepresentation of minorities and females in the workforce unless there is an intervention.
- The exemplary DOE Computational Science Graduate Fellowship (CSGF) program, deemed highly effective in every one of multiple reviews, is uniquely structured and positioned to help provide the future workforce with the interdisciplinary knowledge, motivation, and experiences necessary for contributing to the DOE mission.
- The DOE laboratories have individually developed measures to help recruitment and retention, yet more can be done at the national level to amplify and extend the effectiveness of their locally developed programs.

The subcommittee recommends:

- Preserve and increase investment in the DOE CSGF program while developing new fellowship programs modeled after the CSGF program to increase opportunities for more high-quality students, particularly students from underrepresented populations and demographics, in the computing sciences.
- Develop a recruiting and retention program that increases DOE's visibility on university and college campuses, establish uniform measures across DOE laboratories to improve

the attractiveness of careers in DOE laboratories, and examine the laboratory funding model and its impact on recruiting and retention.

- Establish a DOE-supported computing leadership graduate curriculum advisory group to identify and raise visibility of graduate level curricular competencies specifically required to fulfill DOE's Computing Sciences workforce needs.
- Expand support for local laboratory programs, collect workforce data pertaining to the ASCR-related Computing Sciences, and encourage greater inter-laboratory sharing of information about locally successful programs and workforce related data.
- Working with other agencies, develop a strategic plan with programs and incentives to pro-actively recruit, mentor and increase the involvement of significantly more women, minorities, people with disabilities, and other underrepresented populations into active participation in CS&E careers.

The Findings and Recommendations are further elaborated in Section 8.

3. Workforce Challenges at DOE's National Laboratories

In a joint letter with the corresponding HEPAC and NSAC subcommittees, this committee invited the DOE laboratories to identify disciplines where they experience difficulties in recruiting and retention as a result either of an insufficient number of trained students in the U.S. at the graduate student and postdoctoral level or of high national and international demand, The information the labs provided is summarized in Table 1 in the Appendix along with the number of FTEs in each of the Office of Science laboratories (including the number of postdoctoral and graduate students) in Table 2. The acronyms used to denote the individual laboratories are also provided in the Appendix.

Recruitment and retention challenges in the Computing Sciences were among the most frequently cited, and difficulties were reported in this area by all DOE national laboratories with a workforce greater than 2000 FTEs. An overview of the competencies in the Computing Sciences that were reported as causing difficulties with respect to recruitment and retention are given in Table 1. Although some of the categories overlap, we reproduce them as reported. Several laboratories noted that they make significant efforts to recruit in these areas.

The responses also indicate that, in technical areas, some laboratories increasingly rely on foreign nationals to fill positions and that many recent openings had very small applicant pools. For example, TJNAL reported that, for 1 postdoc position in computational science, there was just 1 applicant, a foreign national. ORNL noted: "Over 75% of qualified applicants for a currently posted staff position in computational biology are foreign nationals. Over 75% of qualified applicants for currently posted postdoc positions in computer vision and machine learning are foreign nationals. For example, in a currently posted staff position in this area, where ability to obtain a Q clearance is preferred, 50% of the applicants are foreign nationals."

The subcommittee solicited further details about hiring and retention from the seven multidisciplinary labs; the data from these inquiries are provided in Table 3 in the Appendix.

From the five responses, the number of open positions suggests strong demand for M.S. and Ph.D.-level employees in the Computing Sciences, particularly at LBNL and ORNL, where the number of open positions represented roughly 25% of the total staff in the Computing Sciences. The laboratories appear to have difficulty filling these positions – it takes nominally 100 days at an Office of Science lab to fill a position and more than twice this time at an NNSA lab. Compare this to industry (Rothwell 2014): an average of 48 and 50 days for M.S. and Ph.D. across STEM fields, respectively, and an average of 39 days for positions in Computer and Mathematical Occupations (all degrees). Both LBNL and ORNL reported high percentages (38%) of foreign nationals in their Computing Sciences workforce. This number is below 10% for the NNSA labs (across all disciplines), where security clearances are often required, and it correlates with the longer time to fill open positions at NNSA labs. Of offers made, four labs report offer acceptance rates above 80%; PNNL has a lower rate of 68%, with location being the main reason. These figures appear to contradict the laboratories' perceptions that they lose candidates to competing offers; however, industry tends to make offers more quickly than the labs, so these numbers may not count qualified candidates who accepted competing offers before a lab offer could be made. Finally, in terms of attrition rates, the NNSA labs reported rates close to 5%, while the Office of Science labs had rates near or above 8% (although the rates for PNNL do not correct for a recent incentivized voluntary separation program). These rates are lower than the roughly 10% average annual (4/2013-4/2014) voluntary turnover rate in US technology companies of similar size (Radford, 2014). Again, this seems to contradict the laboratories' perceptions of high attrition. The attrition rates for the laboratories may be increasing above historic norms *for the labs*, leading to this perception. Furthermore, the laboratories generally cannot support "deep benches" of similar skill sets, so the loss of one or two employees in a particular area could represent a catastrophic loss of expertise. The mismatch of perceptions and data in acceptance and attrition rates suggests that better data and further investigation is needed.

4. Status of Related Disciplines in Academia

Several of the Computing Sciences disciplines are part of the curriculum in many academic institutions throughout the nation, most often as part of a Computer Science, Information Science or Computer Engineering program. In contrast, Computational Science and Engineering (CS&E) is an emerging field of study that is truly interdisciplinary, with participating faculty from Mathematics, Computer/Information Science, and many domain sciences across the curriculum that have embraced computationally intensive methods. Surveys (SIAM 2014, Krell 2010, Dongarra 2012) identify graduate programs of various composition that have recently started in CS&E in the USA and internationally. Programs can be found that have a course concentration, an emphasis, a certificate, a minor, or a degree in CS&E.

CS&E is considered to encompass applications in science/engineering (domain sciences), applied mathematics, numerical analysis, and computer science. Going from application area to computational results requires domain expertise, problem formulation, mathematical modeling, numerical analysis, model discretization, algorithm development, software implementation, program execution, analysis, uncertainty quantification, validation, data analysis and visualization of results. CS&E focuses on the integration of problem-solving methodologies and tools for solving scientific and engineering problems and involves all of these activities.

Two models for the organization of graduate education programs in CS&E have emerged. In the first, a graduate degree in CS&E is awarded by an existing department, usually mathematics or computer science. Examples are Georgia Tech and Rice University. In the second model, graduate degrees are awarded in one or more of the scientific disciplines chemistry, physics, mathematics, biology, computer science, science and engineering, with an area of specialization of CS&E. Examples can be found at the University of Illinois U/C and University of Tennessee. The CS&E programs residing in different departments usually share a core curriculum and standards for graduation.

The need for more such programs and the difficulties of establishing them in compartmentalized academic environments were already stated in the PITAC report (PITAC 2005). We observe that although progress has been made since then, the number of students graduating remains low. Another observation is that Computing areas important to DOE are being studied at a lower rate than more traditional areas of Computing, such as Artificial Intelligence and Networking. As Table 4 in the Appendix shows, the number of PhDs granted over the past four years in areas such as High Performance Computing, Scientific and Numerical Computing, and Visualization are significantly smaller than in other areas of Computing.

Graduate CS&E programs typically do not provide exposure to real-world applications and hence are not able to impart some of the complexities of the field. An NSF taskforce (NSF-ACCI 2011) concluded that universities are not adequately preparing students with the right skills to become tomorrow's computational scientists and engineers able to harness powerful new supercomputers for scientific application and innovation. They also state that educational programs do not teach the skills essential for applying CS&E in modern scientific and technological enterprises. Almost no universities have, or are likely to develop, a curriculum that focuses on topics associated with petascale and exascale science. Indeed, many of the issues in exascale science are not yet understood. However, it is important to start soon to lay the educational foundations for thinking about the emerging computational challenges.

While CS&E has long focused on compute-intensive applications, in the past few years, there is a growing awareness of importance of data-intensive applications. In the 2013 DOE ASCAC Report on Synergistic Challenges in Data-Intensive Science and Exascale Computing, it was noted that: *"Data-intensive research activities are increasing in all domains of science, and exascale computing is a key enabler of these activities… Over the past decade, a new paradigm for scientific discovery is emerging due to the availability of exponentially increasing volumes of data from large instruments such as telescopes, colliders, and light sources, as well as the proliferation of sensors and high-throughput analysis devices."*

Data science, which overlaps with CS&E, is also interdisciplinary, bringing together applied mathematics, statistics, computer science, and domain knowledge; as such, it requires training that crosses traditional academic disciplines. According to a recent McKinsey report (McKinsey

2011), there is a growing shortage of people who have the proper training to deal with big data, so-call Data Scientists. While universities are responding to this growing demand with multiple new M.S. level programs (IAA 2014), there will be significant competition with industry for people with Data Science training.

5. Impact of National Demand and Other Workforce Factors

The disciplines within the Computing Sciences are in extremely high demand nationally and internationally. Consequently, workforce challenges have been widely reported (HEC-IWG 2013, IDC 2010). The Council on Competitiveness studied the role of these technologies in driving private-sector competitiveness and concluded: “HPC is a proven game-changing technology” (Council 2008). The expected job growth across all of Computer Science nationwide has been variously articulated. The US Bureau of Labor Statistics projects that there will be over 1 million jobs available for people trained in this discipline during 2010-2020, compared to under 100,000 jobs for “non-computer” electrical and electronics engineers. Bill Gates’ testimony (Gates 2008) to the House Science Committee on the COMPETES bill put this into perspective: "According to the Bureau of Labor Statistics, we are adding over 100,000 new computer-related jobs each year. But only 15,000 students earned bachelor's degrees in computer science and engineering in 2006 and that number continues to drop."

As described in the previous section, growth in degrees earned in the Computing Sciences remains modest despite the increasing reliance of US industry on computing for its business products. One favorable trend in this data is the growth in doctoral degrees earned in HPC (Table 4). The Taulbee survey (Zweben 2014), which focuses primarily on Computer Science and Computer Engineering, reports a very low unemployment rate (0.8%) in this area, with only 8.2% of PhD graduates from US institutions taking positions outside North America in the fields of interest. Both the unemployment figure and the number departing the US declined over the previous year, indicating a high level of job availability. The same document reports that 70% of HPC graduates went to industry and only 13% to government jobs. For Scientific and Numerical Computing, 55% of PhD graduates entered industry and 7% took government jobs.

The shortage is likely to grow. In 2003, the National Science Board projected a large increase in retirements among Science and Engineering degreed workers over a period of two decades (National Science Board 2003). NSF data (NSF 2002) indicated that retirements in this sector of the workforce were expected to increase dramatically over the next 20 years. They stated that more than half of these workers were age 40 or older and that the 40-44 age group was nearly 4 times as large as the 60-64 age group. Roughly 10 years later, this age group is in the early to mid-fifties; a large number of retirees may be expected during the coming decade.

Several laboratories stated that a major challenge with respect to recruitment and retention in the Computing Sciences was their inability to match industry compensation. PNNL reported: “In the areas of big data, high performance computing, cyber security and biotech, we experience recruiting and retention challenges due to strong competition for candidates and more specifically, the willingness/ability of our competitors to provide compensation packages that we, as a national lab, cannot compete with.” Other impediments include the current constraints on conference travel, since such venues are ideal places to find promising candidates, and an observed general lack of awareness among graduate students of career opportunities at the DOE laboratories.

Many sectors of the population are significantly underrepresented in the Computing Sciences. According to the Taulbee data, in 2014 women comprise a low and declining percentage of computing graduates, with 17.2% of Computer Science and 18% of all computing doctorates. Less than 2% of computational science doctorates are awarded to Hispanic or African-American students. The fraction of degrees awarded to non-US citizens continues to climb, reaching over 58% of all Computing Science doctoral degrees (Table 5 in the Appendix gives examples). Similar demographic data at the career level reveals a workforce that is mostly male and mostly white. Within the DOE, LBNL is about to release its demographic data for the first time (see Table 6 in the Appendix). Their data reveals a specific retention challenge, since there is a decrease in the percentage of women from postdoctoral (15.6%) to career level. The fraction of underrepresented minorities (4.5%) does not change from the postdoctoral to the career level, indicating a problem of recruitment not retention. Similar data can be observed at the other DOE laboratories. While this is a broader societal problem (CEOSE 2013), an increase in diversity would not only greatly increase the pool of US talent but also provide the intrinsic benefits of a diverse workforce to the DOE. We discuss this finding further in the Appendix.

6. Impact of Existing DOE-Wide Workforce Training

To address the need for doctorate-level computational scientists and engineers, in 1991, ASCR established the **DOE Computational Science Graduate Fellowship (CSGF) program.** This cornerstone of DOE CS&E workforce development is jointly funded by ASCR and NNSA at roughly $6-7M per year, which supports roughly 20 new fellows each year.

CSGF is a unique program with four principal objectives (McNeely 2012): a) to help ensure an adequate supply of scientists and engineers appropriately trained to meet national workforce needs in computational sciences including those of the DOE; b) to provide fellows with cross-disciplinary experience opportunities in highly productive work teams by making national DOE laboratories available for practical work experiences; c) to strengthen collaborative ties between the national academic community and DOE laboratories; and d) to raise the visibility of careers in the computational sciences and to encourage talented students to pursue such careers, thus building the next generation of leaders in computational science.

Four effective features that enable the CSGF program to achieve these goals include:

i. An interdisciplinary **Program of Study (POS)** that provides a broad foundation for further CS&E training and practice. Fellows are required to propose and to follow a POS that includes courses in science, applied mathematics, and computer science. The intent is to provide the comprehensive set of mathematical, scientific, and computational skills and techniques needed to tackle the challenges in government, academic, and industrial sectors.

ii. A **research practicum** that opens the doors of DOE laboratories to the fellows and provides them with the opportunity to work with laboratory experts on real-life, large-scale, multidisciplinary research projects. The practicum exposes fellows to the power of HPC as well as the advantages of collaboration, and it is usually a defining point in their careers.

iii. An **annual program review** where past and current fellows, DOE staff, faculty, and other members of the CS&E community come together to share ideas, support one another, and learn about DOE research and employment opportunities. This meeting also provides further opportunities for training in HPC and development of leadership skills.

iv. A **selection process** that is crafted to evaluate individuals and their potential as future computational scientists and engineers. Selection is not solely based on numerical metrics, such as grades or GRE scores, but is the result of a careful assessment by a committee of technical experts who understand the program goals.

CSGF effectively lowers the barriers that separate the different scientific disciplines and exposes fellows to knowledge, experiences, and tools that alter their single-faceted view of science. Another key outcome is that this program allows fellows to build collaborative networks across generational and discipline lines. Hence, because of the deeply invested pedagogical intervention that is at the core of the program, CSGF creates a new kind of scientist that does not naturally result from traditional academic environments or from other fellowship programs.

Reviews of the CSGF program (Kerman 2006, Manteuffel 2011, McNeely 2012) indicate that it has been highly successful within its size and scope. The 2011 ASCAC review of the program (Manteuffel 2011), based on data from 102 alumni from 2001-2009, states that “a large percentage of fellows spend a portion of their early career in the DOE laboratories and an even larger portion continue interaction with the DOE laboratories as they pursue their careers in academia and industry.” Data from this report indicate that roughly one third were employed at DOE and other government laboratories, roughly one quarter were employed in industry, and the remainder predominantly held academic positions. Nine of these alumni received prestigious postdoctoral fellowships. Both the 2006 and 2011 reviews recommend that the program be increased in size and that its scope either be expanded or that similar fellowships in enabling sciences, such as computational mathematics and computer science, be established.

The 2012 longitudinal study of the CSGF program (McNeely 2012), based on surveys of 236 of the 344 alumni and fellows in 2011, further indicates success. Of 155 respondents with completed degrees, it found that 28% of worked in government, 38% in education, and 34% in industry; 89% reported CS&E-related employment. Of 142 alumni for whom CVs were available, the median number of peer-reviewed publications ranged from 18 to 33 across five subgroups based on years since degree; 96 fellows reported publications in top-ranked journals. Other data indicated that fellows and alumni are engaged in service activity and are assuming professional leadership positions. Fully 98% of the fellows indicated that the fellowship directly impacted both the methods used to pursue their own research and their scientific research foci.

7. Role of DOE Laboratories in Workforce Development and Retention

The DOE laboratories will continue to have an important role to play in workforce development. Training in areas that are vital to the DOE mission, and, in particular, in pushing the limits of HPC research, complex systems modeling, and the intersection of these two areas, is especially important and is unlikely to be provided elsewhere.

The laboratories actively address the workforce challenges identified in the previous sections, each pursuing their own strategies for attracting, training, and retaining workers in the Computing Sciences. The labs must continue to engage in workforce development in these disciplines to address specific laboratory needs and to develop and maintain a workforce pipeline that spans staff from graduation throughout their career. Unfortunately, these strategies are not coordinated.

Many laboratories actively engage graduate students and postdocs through a variety of programs that serve both to introduce the students to the laboratories as potential places of future employment, and to begin to build skills not often taught in academia, especially in the context of HPC. Most laboratories provide summer internship programs that are funded directly from laboratory staff's research projects. At several DOE laboratories, it is typical for each to host hundreds of student interns every summer. In addition, some principal investigators fund university subcontracts that support graduate students and academic postdocs, not sited at the laboratory.

A better approach, in the sense that it provides more direct exposure to the laboratory environment, are graduate degree programs in which laboratory researchers jointly serve as the student's Ph.D. advisor and where the student conducts a large portion of their research at the laboratory. Examples include the Higher Education Research Experiences at ORNL program, the Graduate Research Assistant Program at LANL, the Livermore Graduate Scholar program at LLNL, a variety of graduate research appointments at ANL, and the Master's Fellowships at SNL Similarly, the recently instituted Office of Science Graduate Student Research (SCGSR) program supports graduate students to conduct part of their thesis research at a DOE laboratory. Postdoctoral programs are a key component to the hiring pipeline; in fact, many staff are hired from the ranks of the postdoctoral researchers. In particular, named postdoctoral fellowships, such as ANL's Wilkinson and ORNL's Householder, are invaluable in attracting top new talent to the DOE laboratories. A common theme identified by many labs is that early exposure to the laboratory environment can attract better-qualified students into permanent laboratory roles. And, once at the laboratory, programs, such as LBNL's Tuition Reimbursement Program and LLNL's Education Assistance Program, provide tuition assistance to allow existing employees to pursue lab-relevant academic coursework and earn graduate degrees while working.

With respect to mid-career challenges, the DOE Office of Science and the Laboratories should reexamine the career paths they offer to staff to ensure that they are competitive career choices when benchmarked against their peers. In this regard, the committee identified two possible areas for further exploration.

First, recent changes in funding models for computing research at the DOE Office of Science Laboratories have introduced several workforce challenges for engineers and scientists at all levels. For several decades, the DOE Office of Science provided stewardship to the laboratories and in this role, it provided large grants to senior investigators and hierarchical teams (e.g., multiple FTEs) at the laboratories for long-term periods to work on important long-term scientific challenges. This model has been replaced by one where DOE SC provides smaller amounts of funding (less than one FTE) to one or more investigators at laboratories via competitively selected solicitations, which are openly competed frequently (e.g. every 2-3 years). While this new approach enables broad portfolios that address many topics, this cadence and funding level is particularly intimidating for early career laboratory researchers who have no other funding options, must constantly assess their funding situation, and spend an inordinate amount of time writing proposals. This laboratory model sits in stark contrast to that of academic tenure-track researchers who derive most, or all, of their salary from their host university, and is one of the primary complaints of departing research staff members. DOE and the laboratories need to develop more targeted programs to ensure that staff members see the career choice of a laboratory as a sustainable career option.

Second, specific attention should be targeted toward early and mid-career funding opportunities. These opportunities are one way to allow young engineers and scientists to establish a research program at the labs without dividing their time among multiple projects. Two examples of such programs are early-career LDRDs and DOE's early career solicitation.

Indeed, in the realm of retention, the laboratories still have much work to do to find approaches that ensure the workforce maintains a healthy pipeline of workers throughout their careers. New attitudes, incentives (e.g., portable 401k's versus pensions), and 'flat world' accessibility allow talented employees to investigate fresh challenges and opportunities frequently; however, laboratory cultures have traditionally valued longer-term commitments in order to tackle grand-challenge, interdisciplinary scientific problems. In this area, DOE and the laboratories need to find ways to become more agile, to accommodate employee's desire for a greater variety of opportunities throughout employee's careers, and expecting employees to change roles and organizations much more frequently than in the past decades. The need for such opportunities may be met through rotations that allow employees to gain exposure and experience in a variety of areas and training opportunities, short-term and longer-term sabbaticals that allow researchers to explore fresh areas and ideas, and detail assignments, provided that government-imposed rules and barriers can be removed (with the help of DOE). Other measures that may help with the retention problem include providing family care resources to address work/life balance issues, such as childcare and backup care services, community child resources, elder care, and lactation programs. Of course, creating an environment that values and eliminates barriers to productive scientific inquiry, such as stable funding, less restrictive travel policies, and open intellectual property policies, will also be necessary to attract and retain world-class researchers.

As in industry, DOE and the laboratories need to maintain active contacts with employees that have left the laboratories, and try to re-recruit these individuals as they enter mid-career, when their priorities have changed and they are seeking a more stable work-life balance, or are, perhaps, re-entering the workforce, from military service or raising a family, for example.

Finally, with a more dynamic workforce, laboratories will need to reevaluate the priority placed on "training up" new employees if they are unlikely to stay for an extended period of time. Thus, DOE will need academia to provide new employees who are ready to hit the ground running. To address the educational shortcomings identified in Section 4, the DOE Laboratories should proactively work with academic institutions to clearly define suitable curricula that address the workforce needs in computing science. One example of such engagement by DOE laboratories in graduate education is the Nuclear Science and Security Consortium (NSSC) partnership, which includes 7 major universities, 5 minority-serving universities, and 4 DOE labs (LBNL, LLNL, LANL, and SNL). The NSSC's five-year program to train a new generation of nuclear scientists is funded by the National Nuclear Security Administration and has supported over 100 students so far. In fact, in today's reformulation of education, with online education (e.g., massive open online courses) and distance learning becoming common, the laboratories could aggressively partner with universities to offer and teach in-depth classes in topics of interest to the laboratories and in areas where they hold world-class expertise and specialties.

8. Specific Recommendations for Graduate / Postdoctoral Workforce Development

Area 1. Existing programs addressing DOE workforce needs.
The DOE national laboratories face workforce recruitment and retention challenges in all areas of the Computing Sciences that are relevant to their mission (termed *ASCR-related Computing Sciences* in the following findings and the recommendations), including Algorithms (both numerical and non-numerical); Applied Mathematics; Data Analysis, Management and Visualization; Cybersecurity; Software Engineering and High Performance Software Environments; and High Performance Computer Systems. The demand for graduates in Computational Sciences and Engineering far exceeds the supply from academic institutions. The individual labs have developed a range of measures to help recruitment, yet there is no effective sharing of programs, program-related successes, or workforce-related data among the DOE laboratories. The CSGF program is an exemplary DOE agency level program for preparing students with the unique interdisciplinary knowledge, motivation and experiences necessary for successful contributions to the DOE mission. As described in Section 6, the CSGF program is also a very effective program to help meet DOE workforce needs in computational science. While effective, the CSGF program is presently too limited in scope and in size to overcome the workforce recruitment and retention challenges in the DOE national laboratories across the breadth of workforce needs in the computing sciences.

Recommendations:

- **Preserve and increase investment in the DOE CSGF program to increase opportunities for more high-quality students, particularly students from underrepresented populations and demographics.**
- **Establish new fellowship programs, modeled after the CSGF program, for research opportunities in enabling technologies in the computing sciences, including computer science for HPC, large-scale data science, and computational mathematics.**
- **Expand support for local laboratory programs, collect workforce data pertaining to the ASCR-related Computing Sciences, and encourage greater inter-laboratory sharing of information about locally successful programs and workforce related data.**

Area 2. Ability of existing academic programs to meet workforce demand

The successful execution of the DOE mission at present and into the future will require dramatic improvements in multi-disciplinary scientific understanding, where current and future methodologies will need to be coupled with advanced computational modeling and data analytics on state-of-the-art computers. Positions in exascale and extreme computing will create an even greater need for students with a multidisciplinary background that also includes high-performance computing, yet academic programs preparing students in these key areas remains extremely limited. While there are a number of successful CS&E certificate and degree programs across the country, the number of programs and the number of graduates need to be substantially increased over the next 10 years to meet the growing academic, industrial, and government sector career openings. Relatively little education is available at academic institutions in areas of Computing Science that are most relevant to the DOE mission, and even where relevant educational courses are available, the content does not fully meet DOE requirements.

Recommendations:

- **Establish a DOE-supported Computing Leadership graduate curriculum advisory group to annually publish competencies of DOE need at the graduate and undergraduate level in order to influence curriculum development efforts such as those within ACM, CRA and NSF.**
- **In collaboration with the relevant educational organizations such as ACM, SIGHPC, NSF, provide a rich repository of DOE mission-oriented learning materials and engagement opportunities to attract and guide individuals towards careers in areas of DOE need.**
- **Support the development of certificate programs to address the need for competency certification in key ASCR-related Computing Sciences. Work with other agencies and organizations to support the implementation of curricular programs, particularly online programs that benefit DOE, industry and university interests.**

Area 3. Attracting and retaining a DOE mission oriented workforce

There is a national demand for graduates in Computing Sciences and in Computer Science that far outpaces the graduates produced in the US academic institutions with the resulting workforce gap expected to grow over the next decade. There is a large industry demand for students with Master’s level education, which drains the number of students pursuing advanced degrees, and this competition is expected to increase as industry needs more closely parallel those of DOE. The sense of the committee is that there exists a general lack of awareness among college students (both graduate and undergraduate) of the rewarding career opportunities offered by the DOE national laboratories. The shortage of graduates educated in the Computing Sciences is exacerbated by a general shortage of U.S. students in the STEM disciplines. Moreover, there is a profound lack of diversity among the students who graduate in STEM areas, ultimately limiting the diversity of the future DOE workforce even as demographics of the broader population continue to change. While the DOE laboratory culture continues to propagate an emphasis on longer-term commitment, employee attitudes and workplace incentives encourage a more mobile workforce with talented employees seeking out new challenges and opportunities. Lab funding models appear to be less attractive than those of academia and industry.

Recommendations:

- **Examine the laboratory funding model with respect to its implications for workforce recruitment and retention. Implement uniform measures across the DOE laboratories to improve the attractiveness of careers in the DOE laboratories, as well as ensuring long-term commitment, e.g. via continued relocation assistance, ongoing professional development in DOE strategic areas and position rotation, facilitating a sabbatical program for DOE employees.**
- **Increase awareness of DOE opportunities by working with multiple universities to develop a DOE lab presence on campus and increase support for lab employees to visit campuses to promote opportunities within DOE.**
- **Working with other agencies, develop a strategic plan with programs and incentives to pro-actively recruit, mentor and increase the involvement of significantly more women, minorities, people with disabilities, and other underrepresented populations through the completion of their PhD program and their active participation in CS&E careers.**

Citations

(CEOSE 2013) Committee on Equal Opportunities in Science and Engineering, 2011-2012 Biennial Report to Congress. Broadening Participation in America's STEM Workforce. CEOSE 13-01, July 2013

(Chen 2013) J. Chen, A. Choudhary, S. Feldman, B. Hendrickson, C.R. Johnson, R. Mount, V. Sarkar, V. White, D. Williams. "Synergistic Challenges in Data-Intensive Science and Exascale Computing," DOE ASCAC Data Subcommittee Report, Department of Energy Office of Science, March, 2013.

(Council 2008) Council on Competitiveness. Advance: Benchmarking Industrial Use of High Performance Computing for Innovation, May 2008

(Dongarra 2012) Dongarra J, Petzold L and Voevodin V, Survey of Computational Science and Engineering Graduate Programs, December 2012. Online at http://web.eecs.utk.edu/~dongarra/Summary-of-CSE-Survey.pdf

(Gates 2008). Online at: http://www.microsoft.com/enus/news/billg/speeches/2008/congress.aspx

(HEC-IWG, 2013) High End Computing Interagency Working Group (HEC-IWG), Education and Workforce Development in the High End Computing Community, National Coordination Office for Networking and Information Technology Research and Development.

(IAA 2014) Institute for Advanced Analysis Survey on Degree Programs in Data Science, 2014, http://analytics.ncsu.edu/?page_id=4184

(IDC 2010) iDC HPC User Forum, A Study of the Talent and Skill Set Issues Impacting HPC Data Centers

(Kerman 2006) Kerman AK, et al., Review of the DOE Computational Science Graduate Fellowship Program, convened by DOE Office of Science and National Nuclear Security Administration, 2006.

(Krell 2010) Krell Institute, Survey of Undergraduate and Graduate Programs in Computation Science and Engineering, 2010. Online athttp://www.krellinst.org/doecsgf/CS&E_survey/all.php

(Manteuffel 2011) Manteuffel T, et al., Review of DOE Computational Science Graduate Fellowship. Advanced Scientific Computing Advisory Committee Report, 2011. Online at http://science.energy.gov/~/media/ascr/ascac/pdf/reports/ASCAC_CSGF_Report_2011-Final.pdf

(McKinsey 2011) Manyika J et al., Big data: The next frontier for innovation, competition, and productivity, McKinsey & Company, 2011. Online at http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation

(McNeely 2012) McNeely CL The U.S. Department of Energy Computational Science Graduate Fellowship 1991-2011: Recipient Outcomes and Programmatic Impacts. Report of the Krell Institute. 2012. Online at http://www.krellinst.org/doecsgf/DOE_CSGF_Longitudinal_Study_Final_2012.pdf

(National Academy 2011) National Academy of Sciences: Expanding Underrepresented Minority Participation, Washington DC, National Academies Press, 2011

(National Academy 2007) Committee on Prospering in the Global Economy of the 21st Century, Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future, The National Academies Press, 2007

(National Science Board 2003) National Science Board, The Science and Engineering Workforce. Realizing America's Potential. 03-69, National Science Foundation, Arlington VA, August 2003

(NSF 2004) National Science Foundation, Science and Engineering Indicators 2004, May 2004

(NSF 2002) National Science Foundation, Science and Engineering Indicators 2002, May 2002

(NSF-ACCI 2011) NSF Advisory Committee for Cyberinfrastructure (NSF-ACCI) Task Force on Grand Challenges, Final Report, National Science Foundation, March 2011

(PITAC 2005) President's Information Technology Advisory Committee (PITAC), Report to the President: Computational Science: Ensuring America's Competitiveness, June, 2005

(Radford 2014) Radford Trends Report – Q2 2014, US Supplemental Salary Increase and Turnover Report, 1 April 2013 to 1 April 2014, Aon Corporation.

(Rothwell 2014) Rothwell J, "Still Searching: Job Vacancies and STEM Skills," Metropolitan Policy Program, Brookings Institution, 2014. Online at http://www.brookings.edu/research/interactives/2014/job-vacancies-and-stem-skills

(SIAM 2014) SIAMs List of CSE programs, May 2014. Online at http://www.siam.org/students/resources/report.php

(Zweben et al, 2014) Zweben, S and Bizot, B: 2013 Taulbee Survey, Computing Research News, Vol 26 (5), Computing Research Association (CRA), May 2014

(Zweben 2014) Zweben, Stuart. Computing Degree and Enrollment Trends, From the 2012-2013 CRA Taulbee Survey, Computing Research Association (CRA), Washington, May 2014

Appendix

A: Charge to Committee
B. List of acronyms used to identify individual Department of Energy National Laboratories
C. Tables
D. Demographics of Computing Science graduates

February 19, 2014

To: Chairs of the Office of Science Federal Advisory Committee:
Professor Roscoe C. Giles, ASCAC
Professor John C. Hemminger, BESAC
Professor Gary Stacey, BERAC
Professor Mark Koepke, FESAC
Professor Andrew J Lankford, HEPAP
Dr. Donald Geesaman, NSAC

From: Patricia M. Dehmer
Acting Director, Office of Science

Charge: Assessment of workforce development needs in Office of Science research disciplines

The Office of Science research programs have a long history of training graduate students and postdocs in disciplines important to our mission needs as part of sponsored research activities at universities and DOE national laboratories. In addition, the Office of Workforce Development for Teachers and Scientists supports undergraduate internships, graduate thesis research, and visiting faculty programs at the DOE national laboratories.

We are asking the assistance of each of the Office of Science Federal Advisory Committees to help us identify disciplines in which significantly greater emphasis in workforce training at the graduate student or postdoc levels is necessary to address gaps in current and future Office of Science mission needs. As part of your expert assessment, please consider:

· Disciplines not well represented in academic curricula;
· Disciplines in high demand, nationally and/or internationally , resulting in difficulties in recruitment and retention at U.S. universities and at the DOE national laboratories;
· Disciplines identified in the previous two bullets for which the DOE national laboratories may play a role in providing needed workforce development; and
· Specific recommendations for programs at the graduate student or postdoc levels that can address discipline-specific workforce development needs.

Please submit to me, no later than June 30, 2014, a letter report describing your findings and recommendations. These results will be used to help guide future activities and investments.

If you would like to discuss this charge, please do not hesitate to contact me (patricia.dehmer@science.doe.gov). Thank you very much for your help with this important task.

B. Acronyms used to identify individual Department of Energy National Laboratories

Ames	Ames Laboratory, SC
ANL	Argonne National Lab, SC
BNL	Brookhaven National Laboratory, SC
FNAL	Fermi National Accelerator Laboratory, SC
INL	Idaho National Laboratory, NE
LANL	Los Alamos National Laboratory, NNSA
LBNL	Lawrence Berkeley National Laboratory, SC
LLNL	Lawrence Livermore National Laboratory, NNSA
NSCL	National Superconducting Cyclotron Laboratory, NSF
ORNL	Oak Ridge National Laboratory, SC
PNNL	Pacific Northwest National Laboratory, SC
SNL	Sandia National Laboratory, NNSA
TJNAL	Thomas Jefferson National Accelerator Facility, SC

C. Tables

Table 1. Computing Science competencies at national laboratories that are not well represented in the academic curricula and/or have challenges in recruiting and retaining staff.

Competencies	National Laboratory
Advanced computing architectures	LBNL, ORNL
Applied mathematics (Numerical PDEs/high-order discretization modeling and methods, advanced methods for coupled hybrid physical systems; advanced techniques for inverse problems)	ANL, LBNL, LLNL, ORNL
Computational sciences/simulation; scientific software	ANL, BNL, INL, LBNL, LLNL, ORNL, SNL
Cyber security	INL, LLNL, ORNL, SNL
Data acquisition software	FNAL, ORNL
Data informatics, including data mining, machine learning; big data, including statistical techniques	ANL, LBNL, LLNL, ORNL, PNNL
Dynamic mesh algorithms	LLNL
High-performance/ Extreme/Exascale computing	ANL, INL, LANL, LBNL, LLNL, ORNL, PNNL, SNL
Performance analysis of HPC applications	LBNL, ORNL
Software quality assurance	LLNL, ORNL
Solvers	LBNL, LLNL
Storage systems	LBNL, ORNL
Uncertainty quantification	LLNL, ORNL
Visualization and scientific data analysis	LLNL, ORNL

Table 2. Size of workforce at Office of Science national laboratories. Includes number of graduate students and postdocs.

FY2013	FTEs	Joint Faculty	Postdocs	Grad Students	Undergrads	Facility Users	Visiting Scientists
Ames	308	84	42	73	43	0	3
ANL	3460	178	285	689	80	6547	991
BNL	2882	25	162	201	279	4134	1377
FNAL	1720	9	55	0	0	2097	19
LBNL	3396	228	496	308	146	9320	1565
ORNL	4586	164	358	204	518	3215	1888
PNNL	4344	3	200	167	183	1733	66
PPPL	429	4	14	40	0	0	300
SLAC	1596	23	92	121	0	4474	26
TJNAF	729	22	20	34	6	1261	1095

Table 3. Recruitment and retention of graduates at multi-purpose DOE labs May 2013-May 2014.

	Open Positions	Avg. Time to Fill Position (days)	Total Number of Technical Staff	Total Number of Technical Staff in Computing	Proportion of Foreign National Technical Staff	Declined Job Offers	Attrition Rate
LANL	148*	263*	1903*		5.4%*	21/173*	4.9%*
LBNL[1]	56	112		206	38.4%	2/39	8.0%
LLNL[2]	146	311	2094*		7.4%*	7/36	4.8%*
ORNL[3]	87	110		379	38%	11/73	7.6%
PNNL[4]	44	107	1113*		16%*	16/50	8-9%**

* Data for all scientific and engineering disciplines, M.S. and Ph.D. level

** Data for all scientific and engineering disciplines, all degree levels

[1] LBNL data for "all scientists and engineers on the Computer Science curve"

[2] LLNL data based on best attempt to identify positions in the Computing Sciences; time-to-fill may be skewed by indefinite postings; attrition rate corrected for voluntary separation program

[3] ORNL data for "lab-wide computing/computational science" positions; attrition rate corrected to account for voluntary separation program (37% of terminations)

[4] PNNL attrition rate is uncorrected for voluntary separation program; historical rate is 4-5%; total number of job offers is estimated.

The seven DOE multi-purpose laboratories were contacted to provide further data on hiring in the Computing Sciences; five provided responses. For the purposes of the request, the population and time frame were defined as "M.S., Ph.D., and post-doctoral positions over the last year (May 2013 to May 2014) in computer science, statistics, applied mathematics, computational science and engineering (e.g., computational physics, computational chemistry, etc.), data analytics, machine learning, and related fields." Requested data were: (1) Total number of vacancies in the population; (2) Average time to fill a position (posting to acceptance); (3) Average time taken to fill a position that requires US citizenship; (4) Average number of M.S.- and Ph.D.-level scientists and engineers on staff across all disciplines; (5) Average percentage of M.S.- and Ph.D.-level scientific and engineering staff that are foreign nationals across all disciplines; (6) Number / percentage of declined offers in the population and top reasons, where known (e.g. pay, location, better offer); and (7) Attrition rate from the MS and Ph.D. scientists and engineers on the technical staff. Because laboratory job classifications do not consistently line up with the Computing Sciences disciplines, data on this category were difficult to reconstruct for some labs. In addition, LLNL, ORNL, and PNNL had incentivized voluntary separation plans during the period, for which the data are only partially corrected.

Table 4. Taulbee Survey Number of Ph.D.'s in CS and Comp. Eng. granted (Zweben 2014)

PhD Specialty Area	2010	2011	2012	2013	Total
Artificial Intelligence	181	193	203	171	748
Computer-Supported Cooperative Work	19	6	8	17	50
Databases/Information Retrieval	99	106	122	125	452
Graphics/Visualization	87	111	99	99	396
Hardware/Architecture	78	70	92	91	329
Human-Computer Interaction	49	69	80	81	279
High-Performance Computing	29	37	49	60	175
Informatics: Biomedical/Other Science	75	68	97	74	314
Information Assurance/Security	70	82	69	77	298
Information Science	20	30	57	45	452
Information Systems	36	39	49	30	154
Networks	150	140	147	152	589
Operating Systems	59	55	66	55	235
Programming Languages/Compilers	65	48	64	58	435
Robotics/Vision	65	60	78	76	279
Scientific/Numerical Computing	33	27	32	29	121
Social Computing/Social Informatics	28	31	20	25	104
Software Engineering	126	147	149	140	562

Table 5. Taulbee Survey Number of Ph.D.'s in CS and Comp. Eng. granted to US Citizens or Permanent Residents (Zweben 2014)

PhD Specialty Area	Count	% of Specialty
Artificial Intelligence	439	58.7%
Computer-Supported Cooperative Work	31	62.0%
Databases/Information Retrieval	203	44.9%
Graphics/Visualization	228	57.6%
Hardware/Architecture	147	44.7%
Human-Computer Interaction	196	70.3%
High-Performance Computing	78	44.6%
Informatics: Biomedical/Other Science	183	58.3%
Information Assurance/Security	147	49.3%
Information Science	106	69.7%
Information Systems	85	55.2%
Networks	205	34.8%
Operating Systems	108	46.0%
Programming Languages/Compilers	151	64.3%
Robotics/Vision	147	52.7%
Scientific/Numerical Computing	78	64.5%
Social Computing/Social Informatics	66	63.5%
Software Engineering	328	58.4%
Theory and Algorithms	214	46.9%
Other/Unknown	863	55.8%

Table 6. Overall Demographics: Berkeley Technical Workforce, October 1, 2013

Types of Jobs at Berkeley Lab	TTL	Women	%	URM	%	OPC	%
Scientists and Engineers (Conducting research)	640	100	15.6%	29	4.5%	131	20.5%
Postdoctoral Scientists	486	133	27.4%	26	5.3%	209	43.0%
Engineers (Information, Mechanical, and Electrical)	483	102	21.1%%	51	10.6%	118	24.4%
Research Support (Non S&Es in programmatic divisions)	907	390	43.0%	145	16.0%	207	22.8%
Ops Support (Non S&Es in operational divisions)	677	324	47.9%	161	23.8%	117	17.3%
Totals	3193	1049	32.9%	412	12.9%	782	24.5%

Career, term, and postdoctoral employees only
URM=Underrepresented Minorities (African American/Black, Hispanic/Latino, and American Indian/Alaskan Native)
OPC= Other People of Color (Asian/Asian American, Middle Eastern/Southwest Asian/North African and Pacific Islanders)

D. Demographics of Computing Science graduates

The demographics of graduates in the Computing Sciences are noteworthy. If the workforce gap identified in this report is to be overcome, then they will need to change. The Taulbee survey (Zweben et al, 2014) reported that the fraction of degrees awarded to non-US citizens continues to climb, reaching over 58% of all Computing Science doctoral degrees (see also Table 5 in the Appendix). The number of US citizens among new enrolments in PhD programs dropped by over 12% from 2012 to 2013. Indeed, extensive data shows that comparatively few US students graduate in science and engineering degree programs and the fraction of foreign nationals has steadily grown (National Academy 2007, National Science Board 2003, NSF 2004). According to (National Academy 2007), a major contributing factor is the large number of dropouts at the undergraduate level; a disproportionate number of dropouts are female or non-white.

Data from different sources indicates that many sectors of the population are significantly underrepresented in the Computing Sciences. The Taulbee report (Zweben et al, 2014) shows that the number of females enrolled at all levels of study in the Computing Sciences is low and declining. According to its data, in 2014 women comprise 17.2% of CS doctoral graduates and 18% of all computing doctorates, which is below the corresponding figures from the previous year. Foreign nationals are a larger fraction of the female PhDs than of the male PhDs. Participation by Hispanic and African-American students is fairly constant and low over MSc and PhD studies (Zweben et al, 2014). Less than 2% of the computational science doctorates are Hispanic or African-American students. The (CEOSE 2013) report points out some of the socio-economic factors contributing to this lack of diversity, including lack of educational preparation, lack of role models and mentors, poorly equipped schools and systemic biases (see also (National Academy 2011) on this topic).

Similar demographic data is being reported at the career level, revealing a workforce that is mostly male and mostly white. Google recently released its diversity numbers for the first time, followed by Linkedin, Yahoo, and Facebook. Google's tech staff is 1% African American, 2% Hispanic, and 34% Asian. Men account for 83% of the tech staff while women occupy only 21% of leadership positions and 17% of tech jobs at the company. Google has no female executive officers, and only one woman on its senior leadership team. Within the DOE, LBNL is about to release its demographic data for the first time. According to a preview of this data (see Table 6 in the Appendix), women comprise 15.6% of all scientists and engineers conducting research at LBNL and underrepresented minorities (including African American, Hispanic and American Indian) make up 4.5%. Other ethnicities (including Asian) are 20.5% of the total. The table shows that 27.4% of postdoctoral scientists at LBNL are female, which is substantially higher than the percentage of female career employees. Similar numbers can be observed in the other DOE labs. This decrease in the percentage of women from postdoctoral to career level reveals a specific retention challenge. The percentage of underrepresented minorities does not change from the postdoctoral to the career level, indicating a recruitment, but not a retention, problem.

One of the reasons companies, such as Google and LinkedIn, are releasing their diversity data is because they are convinced that the only way to meet their demands for technical talent is to increase the talent pool. A more diverse workforce brings additional benefits. As the DOE moves toward exascale computing, software engineering and modeling will require skilled multidisciplinary teams with highly creative individuals. Greater workforce diversity will provide a richer variety of points of view; therefore, it has the potential to yield more innovative solutions. Organizations that create and execute a roadmap to develop a diverse workforce are likely to attract more talent in the long run. The DOE already has elements of such a plan. The Office of Science's Office of Workforce Development for Teachers and Scientists (WDTS) offers internship programs for undergraduate students in collaboration with the DOE labs. They encourage the students to pursue STEM careers before they select a major, and reach out to women and underrepresented minorities in particular. An expansion of this effort might help create a larger and more diverse population of graduates in Computing Sciences.