ASCAC WORKFORCE SUBCOMMITTEE REPORT TO ASCAC COMMITTEE

Prepared by

ASCAC Workforce Subcommittee

July 15, 2014

The Charge

identify disciplines that need more workforce training at grad student or postdoc level for DOE Office of Science mission needs.

Please consider:

- Disciplines not well represented in academic curricula;
- Disciplines in high demand, nationally and/or internationally, resulting in difficulties in recruitment and retention;
- Disciplines where the DOE labs may play a role in providing needed workforce development;
- Specific recommendations for programs at the graduate student or postdoc levels to address discipline-specific workforce development needs.

The Subcommittee

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

Role of Advanced Scientific Computing

- Scientific computation is at the core of a large fraction of the R&D undertaken under the auspices of Office of Science.
- ASCR facilities and experts in their utilization form the underpinning of successful research programs in all areas of scientific activity in the DOE national laboratories.
- With the growing importance of the pre- and post-processing of vast amounts of data, the breadth of expertise required to effectively deploy available tools and to use them to accomplish scientific goals is also increasing.

- Maintaining a sufficient workforce is this area is critical if research outcomes are to be produced and if the significant investment in ASCR facilities is to be amortized

Interpretation of Charge

- ASCAC disciplines needed for DOE mission do not fit neatly into traditional academic categories
  - Some positions require multidisciplinary training
- We coined term **Computing Sciences** for this report
  - Includes: Algorithms; Applied Mathematics; Data Analysis, Management and Visualization; Cybersecurity; Software Engineering and High Performance Software Environments; and High Performance Computer Systems.
- Committee decided to take a broader perspective
  - in order to address workforce challenges identified, more than a program is needed

Limitations of Report

- Information on the disciplines (and multidisciplinary expertise) needed to support DOE mission is not readily available
- Data and information obtained usually covers some part of the Computing Sciences
- HR data typically not available for this range of subjects
  - E.g. some positions are categorized in area of domain expertise
- Hence our findings cannot be conclusive

Summary of Findings

- The multidisciplinary national labs face workforce recruiting and retention challenges in Computing Sciences
- Insufficient educational opportunities are available at academic institutions in areas of the Computing Sciences most relevant to the DOE mission.
- There is a growing demand for graduates in Computing Sciences that far exceeds the supply. A larger workforce gap and continued underrepresentation of minorities and females are expected.
- The exemplary DOE CSGF program, deemed highly effective in multiple reviews, is uniquely structured and positioned to 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 local programs.

Summarized Recommendations

- Leverage and strengthen the successful DOE CSGF program by doubling its funding and expanding its scope to include research in HPC-enabling sciences such as computational mathematics, computer science, and data analytics.
- Establish a DOE-funded computing leadership graduate curriculum advisory group for establishing graduate level curricular competencies specifically to fulfill DOE's Computing Sciences workforce needs.
- Develop a recruiting and retention program that increases DOE's visibility on university and college campuses and that provides relocation assistance, travel for recruiting, ongoing professional development, opportunities to take sabbaticals and other non-monetary career incentives.
- Build a diverse workforce that spans demographics and universal accessibility for a broader awareness of career opportunities within DOE. This includes inter-agency coordination, targeted opportunities, and strategic development plans to expand and build a workforce to include non-traditional, underrepresented, and military veterans.
- Expand support for local laboratory programs and encourage greater inter-laboratory sharing of information about locally successful programs as well as workforce related data.

Major Findings

- The multidisciplinary national labs face workforce recruiting and retention challenges in Computing Sciences
- Insufficient educational opportunities are available at academic institutions in areas of the Computing Sciences most relevant to the DOE mission.
- There is a growing demand for graduates in Computing Sciences that far exceeds the supply. A larger workforce gap and continued underrepresentation of minorities and females are expected.
- The exemplary DOE CSGF program, deemed highly effective in multiple reviews, is uniquely structured and positioned to 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 local programs.

Workforce Recruitment/Retention Challenges

- Labs invited to provide information on areas where they experience recruitment / retention difficulties
- Areas in Computing Sciences figured prominently in responses, especially in large labs
- For Computing Sciences, labs reported:
  - Low number of qualified applicants, many of which are foreign nationals
  - Takes a long time to fill positions
  - Labs spend significant effort on recruiting in this area
- Additional information solicited from larger labs to quantify problem
  - No uniform method for recording information; not all labs provided data
  - Takes labs twice as long as industry to fill positions in computing sciences
  - About 4 times as long when security clearance does not permit foreigner
  - Acceptance and retention rates mostly favorable

Competencies	National Laboratories
Advanced computing architectures	LLNL, ORNL
Applied mathematics (including advanced modeling and methods)	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 (data mining, machine learning, big data, statistical techniques)	ANL, LBNL, LLNL, ORNL, PNNL
Dynamic mesh algorithms	LLNL
HPC /extreme-scale/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

Recruitment/Retention at Labs

Lab	# Open Posns	Ave. Time To Fill (Days)	Total Technical Staff	% Foreign Nationals	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.

Major Findings

- The multidisciplinary national labs face workforce recruiting and retention challenges in Computing Sciences
- Insufficient educational opportunities are available at academic institutions in areas of the Computing Sciences most relevant to the DOE mission.
- There is a growing demand for graduates in Computing Sciences that far exceeds the supply. A larger workforce gap and continued underrepresentation of minorities and females are expected.
- The exemplary DOE CSGF program, deemed highly effective in multiple reviews, is uniquely structured and positioned to 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 local programs.

Disciplines in Academia

- Interdisciplinary Computational Science and Engineering studies emerging, but insufficient number of graduates
  - Do not provide full skillset needed by DOE
- Related fields in Computer Science (CS), Info Science, Computer Engineering (CE)
  - Fewer CS and CE graduate degrees in Computing Sciences areas than traditional areas
- Interdisciplinary Data Science education beginning to emerge but not likely to satisfy demand

Number of PhDs in CS and CE Areas

PhD Specialty	2010	2011	2012	2013	Total
Artificial Intelligence	181	193	203	171	748
Databases/Info Retrieval	99	106	122	125	452
Graphics/Vis	87	111	99	99	396
HW/Architecture	78	70	92	91	329
High Performance Computing	29	37	49	60	175
Networks	150	140	147	152	589
Operating Systems	59	55	66	55	235
Scientific/Numerical Computing	33	27	32	29	121
Software Engineering	126	147	149	140	562

Major Findings

- The multidisciplinary national labs face workforce recruiting and retention challenges in Computing Sciences
- Insufficient educational opportunities are available at academic institutions in areas of the Computing Sciences most relevant to the DOE mission.
- There is a growing demand for graduates in Computing Sciences that far exceeds the supply. A larger workforce gap and continued underrepresentation of minorities and females are expected.
- The exemplary DOE CSGF program, deemed highly effective in multiple reviews, is uniquely structured and positioned to 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 local programs.

Workforce Demand

- Huge demand for graduates with computing expertise
  - Taulbee survey reports very low unemployment in US for computing graduates
  - Large majority of graduates enter industry
  - Retirement of current workforce is expected to grow gap over coming decade
- Labs cannot compete with industry compensation
- Awareness of lab careers among graduates low
- Conference travel restrictions impede recruitment

Workforce Demand

- Fraction of foreign nationals in graduate population is growing steadily
  - 58% of graduates in Computing Sciences are now foreign nationals
- Lack of diversity in US graduates in CS and CE is a major contributing factor in national shortage
  - Current US citizens among graduates are mostly white male; same holds for workforce (where data public)
  - Number of African American and Hispanic graduates remains very low (ca. 1% each)
  - Percentage of females among graduates is declining
- Data from DOE labs reflect national demographics
  - Also indicate retention problem for female postdocs
- Lack of diversity in STEM disciplines widely acknowledged but not effectively addressed
  - Appendix includes additional information

US Citizens / Permanent Residents as Percentage of PhDs in CS, CE Areas

PhD Specialty	Citizens, Permanent Residents	% of Total
Artificial Intelligence	439	58.7%
Databases/Info Retrieval	203	44.9%
Graphics/Vis	228	57.6%
HW/Architecture	147	44.7%
High Performance Computing	78	44.6%
Networks	205	34.8%
Operating Systems	108	46.0%
Scientific/Numerical Computing	78	64.5%
Software Engineering	328	58.4%

LBNL Demographics in STEM

Overall Demographics: Berkeley

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%**

Date of Data: October 1, 2013
Career, term, and postdoctoral employees only
URM=Underrepresented Minorities (African American/Black, Hispanic/Latino, and American Indian/Ala
OPC=Other People of Color (Asian/Asian American, Middle Eastern/Southwest Asian/North African and

Major Findings

- The multidisciplinary national labs face workforce recruiting and retention challenges in Computing Sciences
- Insufficient educational opportunities are available at academic institutions in areas of the Computing Sciences most relevant to the DOE mission.
- There is a growing demand for graduates in Computing Sciences that far exceeds the supply. A larger workforce gap and continued underrepresentation of minorities and females are expected.
- The exemplary DOE CSGF program, deemed highly effective in multiple reviews, is uniquely structured and positioned to 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 local programs.

Existing Workforce Training

- DOE CSGF program established 1991; jointly funded by ASCR and NNSA
- Trains graduate students to meet national workforce needs in computational sciences, including those of DOE
- Provides practical work experiences at DOE labs; improves collaboration between labs and academia; raises visibility of careers in computational sciences
- Effective elements
  - Interdisciplinary program of study
  - Research practicum at DOE laboratories
  - Annual review that enables networking
  - Careful selection process

Existing Workforce Training

- Multiple reviews attest to success of program
- Show that it contributes to lab and national workforce in Computational Sciences

- 2011 ASCAC review of the program, 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."
- 2012 longitudinal study of the CSGF program: Of 155 respondents with completed degrees, 28% worked in government, 38% in education, and 34% in industry; 89% reported CS&E-related employment

Major Findings

- The multidisciplinary national labs face workforce recruiting and retention challenges in Computing Sciences
- Insufficient educational opportunities are available at academic institutions in areas of the Computing Sciences most relevant to the DOE mission.
- There is a growing demand for graduates in Computing Sciences that far exceeds the supply. A larger workforce gap and continued underrepresentation of minorities and females are expected.
- The exemplary DOE CSGF program, deemed highly effective in multiple reviews, is uniquely structured and positioned to 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 local programs.

Role of Labs in Workforce Training

- Individual labs have efforts, but not coordinated
  - Measures include summer internships, individual collaborations with universities, joint supervision of graduate students
- Labs also have programs that help employees to obtain degrees
- Postdoc programs help attract new talent to labs
- Labs face many challenges
  - Change in funding models
  - Need to adapt to shorter-term commitment
  - Help universities provide more suitable education

Full Set of Recommendations (1)

- Preserve and increase investment in the DOE CSGF program to 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 and encourage greater inter-laboratory sharing of information about locally successful programs and workforce related data.

Full Set of Recommendations (2)

- Establish a DOE-funded Computing Leadership graduate curriculum advisory group to spearhead participation in efforts within ACM, CRA and NSF to develop and annually publish competencies of DOE need at the graduate and undergraduate level.
- Working with ACM SIGHPC, NSF and other organizations, 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.
- Working with other agencies and organizations, establish certificate programs to address need for competency certification. Work with other agencies to fund implementation of curricular programs, particularly online programs, in the areas of DOE need.

Full Set of Recommendations (3)

- Improve attractiveness of DOE opportunities with continued relocation assistance, ongoing professional development in DOE strategic areas and position rotation, and establish a sabbatical program for DOE employees.
- Increase awareness of DOE opportunities by working with multiple universities to develop campus champions and increase support for DOE 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 sustain 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.