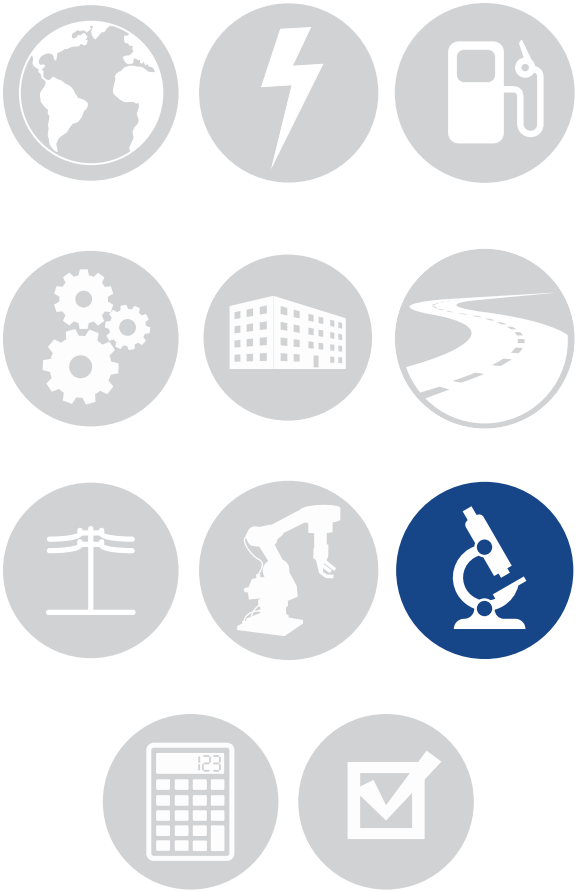
Quadrennial Technology Review 2015

Chapter 9: Enabling Capabilities for Science and Energy

Supplemental Information



A Comparison of Research Center Funding Modalities

High-Performance Computing Capabilities and Allocations

User Facility Statistics

Examples and Case Studies

U.S. DEPARTMENT OF ENERGY

Quadrennial Technology Review 2015

User Facility Statistics

Chapter 9: Enabling Capabilities for Science and Energy

Scientific User Facility Statistics

User facilities[1] – federally sponsored research facilities available for external use to advance scientific or technical knowledge – are a core component of the Department of Energy's (DOE) Office of Science (DOE-SC) mission and an important part of the broader DOE mission. The 34 DOE user facilities[2] provide state-of-the-art experimental and/or computational resources that are prohibitively expensive to develop, build, and operate by a university, private sector, or nonprofit laboratory. The DOE user facility access model (see below) enables the DOE national laboratories that host these facilities[3] to bring thousands of outside researchers on-site every year to both leverage the unique tools and staff expertise for basic science and technology RD&D and lend their technical expertise toward the maintenance, development, and application of these tools in support of the broader scientific community.

The breadth of capabilities offered by the DOE user facilities means that the user community represents an incredible diversity of scientific and engineering disciplines. The broad scope of science and technology R&D enabled by these facilities is clear from the thousands of publications produced each year.[4] Detailed information on publications, including both peer-reviewed articles as well as published patents, for work either partly or wholly conducted at the DOE user facilities, can be found at the respective websites for each facility.

All DOE-SC user facilities are open access, with allocation of time determined by merit-based peer review of user proposals.[5, 6] The submitted proposals are reviewed irrespective of nationality or institutional affiliation, enabling domestic and international scientists from universities, federal laboratories, the private sector, and nonprofits to use the unique capabilities and sophisticated instrumentation. User fees are not charged for nonproprietary research if the user intends to publish the results in the peer-reviewed literature; full cost-recovery is required for proprietary research. In general, the demand from users for access to the DOE-SC user facilities exceeds the available resources; all are oversubscribed. Users that are not allocated facility resources may apply in a subsequent proposal competition cycle.

The DOE-SC collects statistics on the usage of its user facilities, including user's institutional affiliation, institution sector (i.e. university, laboratory, or private sector firm), and number of unique users from that institution (see *Defining and Counting Users*, below). Collection of detailed user statistics facilitates:

- Understanding and documenting of a facility's impact on and/or importance to a scientific community;
- Strategic planning by the facility, laboratory, and federal sponsor; and
- Identifying trends and patterns in usage that reflect the evolution of a facility and/or its user community.

The following brief review presents the FY2014 usage of the user facilities sponsored by the DOE-SC, specifically, the Offices of Advanced Scientific Computing Research (SC-ASCR), Biological and Environmental Research (SC-BER), and Basic Energy Sciences (SC-BES). The complete set of FY2014 user facility statistics, which includes the user facilities sponsored by the Offices of High Energy Physics, Nuclear Physics, and Fusion

Energy Science, is presented via an interactive map hosted at the DOE-SC website.[7] The complete data set of FY2014 user statistics is available for download.[8]

Defining and Counting Users

A user is an individual or a member of a research team who is granted access to resources at a user facility through an approved peer-reviewed proposal.[9] An individual is counted as a user only once for a given facility in a fiscal year. A user who utilizes more than one facility can be counted as a unique user for each facility. Facilities annually report unique users in three categories:

- *On-site user:* an individual who is physically present at the facility at least once during the fiscal year.
- *Remote user:* an individual who remotely accesses the facility at least once during the fiscal year.
- *Data user:* an individual who remotely accesses data from an electronic archive supported by the facility at least once during the fiscal year.

The DOE-SC issues an annual statement of its user statistics collection practices that includes a description of the modes of user interface and the methods of acquiring user statistics. This information is available at the DOE-SC website.[10]

Office of Science General User Statistics for All Facilities

The 28 DOE-SC user facilities hosted 33,671 users in FY2014 (Figure 1). The majority of these users (16,118) are from SC-BES sponsored facilities, which support 12 of the 28 facilities. Domestic users across all sectors made up 81% of the user base, with the remaining 19% originating from the international science and technology community (Figure 2). Within the international user community, the top five countries (by total users) are Germany (710), United Kingdom (657), China (622), Canada (502), and France (388). By sector, the academic community (both domestic and international) is the largest, with nearly 22,000, or 65%, of the total user base. The remaining 35% originate from the DOE national laboratories, other research laboratories (including non-profit laboratories), private sector firms, federal agencies, and the non-academic international community (see below for more information on users by sector).

Figure 1 The total number of users of DOE-SC user facilities, by program, for Fiscal Year 2014. The number of users given for each program represents the total number of users for all facilities stewarded by the program.

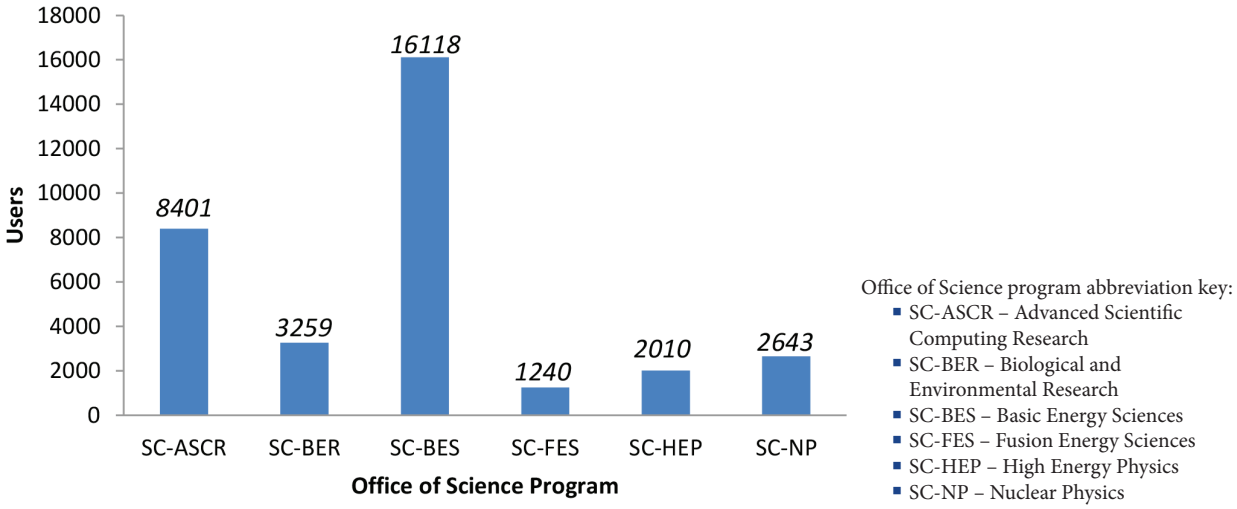
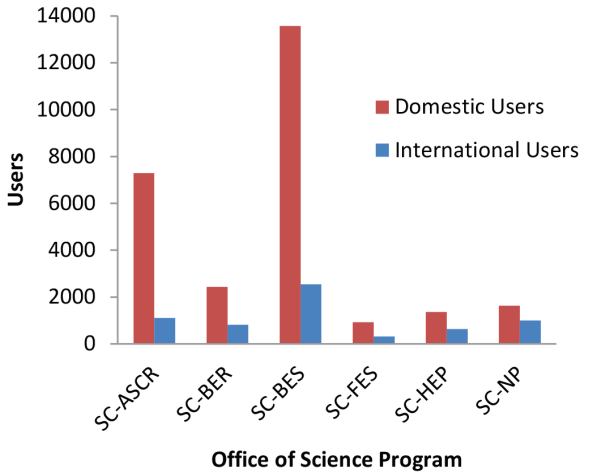


Figure 2 The total number of users of DOE-SC scientific user facilities from domestic and international institutions, by program, for Fiscal Year 2014. The number of users given for each program represents the total number of users for all facilities stewarded by the program.



Program	Total Users	Domestic	International
SC-ASCR	8401	7288	1113
SC-BER	3259	2442	817
SC-BES	16118	13573	2545
SC-FES	1240	925	315
SC-HEP	2010	1369	641
SC-NP	2643	1638	1005

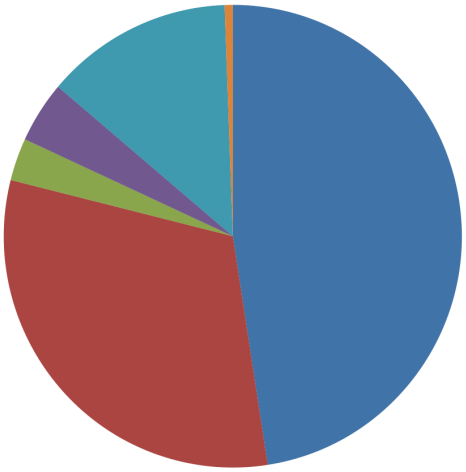
Office of Science, Office of Advanced Scientific Computing Research (SC-ASCR)

The mission of SC-ASCR is to discover, develop, and deploy computational and networking capabilities in service of analysis, modeling, simulation, and prediction of complex phenomena important to the DOE and to the broader science and engineering community. A critical part of this mission is developing emerging computing systems and novel computing architectures for application to modern science and engineering challenges.

SC-ASCR hosts four user facilities, three of which–the Argonne Leadership Computing Facility (ALCF), the Oak Ridge Leadership Computing Facility (OLCF), and the National Energy Research Scientific Computing Center (NERSC)–are high performance computing facilities housing 4 of the top 100 fastest computers in the world and supporting 8349 users in FY2014 (Figure 3). Nearly 50% of the users hosted by SC-ASCR facilities are from domestic academic institutions. Of the remaining 50%, more than half (31% of the total user base) are from DOE national laboratories. Users from private sector firms make up 4% of the user base, and are engaged in projects with strong ties to the energy technologies surveyed in the QTR (see section 9.6.1 and the supplemental information *High Performance Computing Capabilities and Allocations* for more information on private sector supercomputing projects for energy technology R&D).

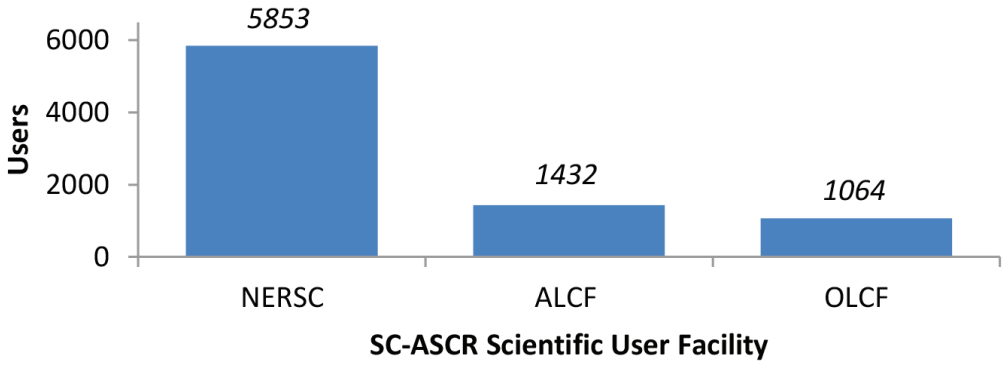
The majority of users of SC-ASCR-supported facilities utilize NERSC's two petaflop-scale computers, Hopper and Edison, which are classified as 'production' machines. Production machines serve the Office of Science research awardees, a very large and scientifically diverse user base, addressing challenges that do not require the computing power of the leadership class machines. The two leadership class machines, Titan and Mira at OLCF and ALCF, respectively, are reserved for a much smaller set of science and engineering problems that would be prohibitively expensive or impossible to solve on less powerful machines. The user statistics reflect these differences with NERSC (5853 users) surpassing ALCF (1432 users) and OLCF (1064 users) (Figure 4). More information about facilities and computers is available in the main report of the Quadrennial Technology Review (QTR) 2015, Chapter 9--*Enabling Capabilities for Science and Energy*, section 9.6.1. A more complete description of the SC-ASCR allocation programs for both leadership class and production class machines is given in the supplemental information (*High-Performance Computing Capabilities and Allocations*).

Figure 3 The number of users by sector of SC-ASCR user facilities. The total users are an aggregation of all users from the four user facilities stewarded by SC-ASCR. The category "other" includes PK12 schools and non-profit research organizations, among others.



Sector	Users	%Total Users
Academia (domestic)	3996	47.6
DOE national laboratories	2633	31.3
Other federal laboratories & agencies	250	3.0
Private sector	361	4.3
International (all sectors)	1113	13.2
Other	48	0.6
Total	**8401**	**100**

Figure 4 The total number of users for the three SC-ASCR-stewarded high performance computing user facilities for FY2014. The number for each facility represents unique users at that facility. Users that utilized more than one facility are counted as unique users for each facility.



Scientific user facility abbreviation key:

- NERSC – National Energy Research Scientific Computing Center (Lawrence Berkeley National Laboratory)
- ALCF – Argonne Leadership Computing Facility (Argonne National Laboratory)
- OLCF – Oak Ridge Leadership Computing Facility (Oak Ridge National Laboratory)

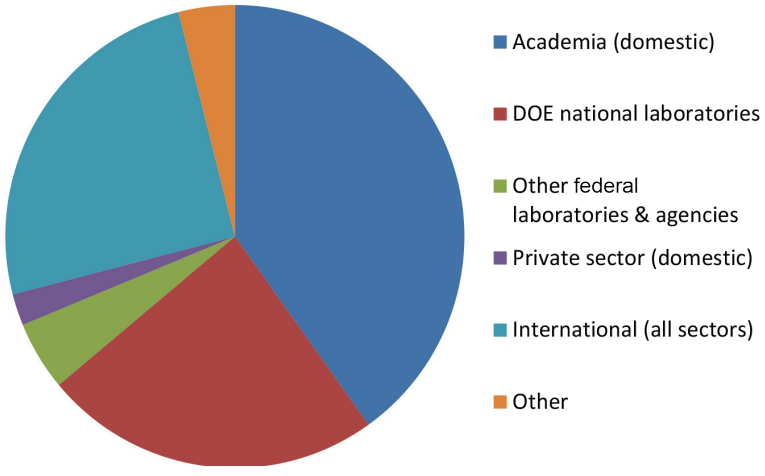
SC-ASCR also supports the Energy Sciences Network (ESNet), a 100 Gbps network that connects 40 DOE research sites, including the entire national laboratory complex, its supercomputing facilities, and its large-scale experimental facilities. ESNet is optimized to support lossless data transfer of very large experimental and computational data sets at very high bandwidth. ESNet is used by thousands of DOE supported researchers each year. Connections to 140 research and commercial networks, as well as selected locations in Europe, including the European Organization for Nuclear Research (CERN), which operates the Large Hadron Collider, enables world-wide scientific collaboration. A component of ESNet operates as a user facility: a national-scale network research testbed that is open to both DOE and non-DOE scientists conducting applied research in a variety of networking topics. In FY2014, the ESNet testbed had 52 unique users representing academia, DOE and other government laboratories, and the private sector. More information on ESNet is available in the QTR 2015 report, section 9.6.2.

Office of Science, Office of Biological and Environmental Research (SC-BER)

The mission of the SC-BER program is to support fundamental research and scientific user facilities to achieve a predictive understanding of complex biological, climatic, and environmental systems for a secure and sustainable energy future. SC-BER-supported research seeks to understand the continuum of biological, biogeochemical, and physical processes from the smallest scales (genomes and metabolic pathways) to the largest scales (ecosystems and atmospheric observation).

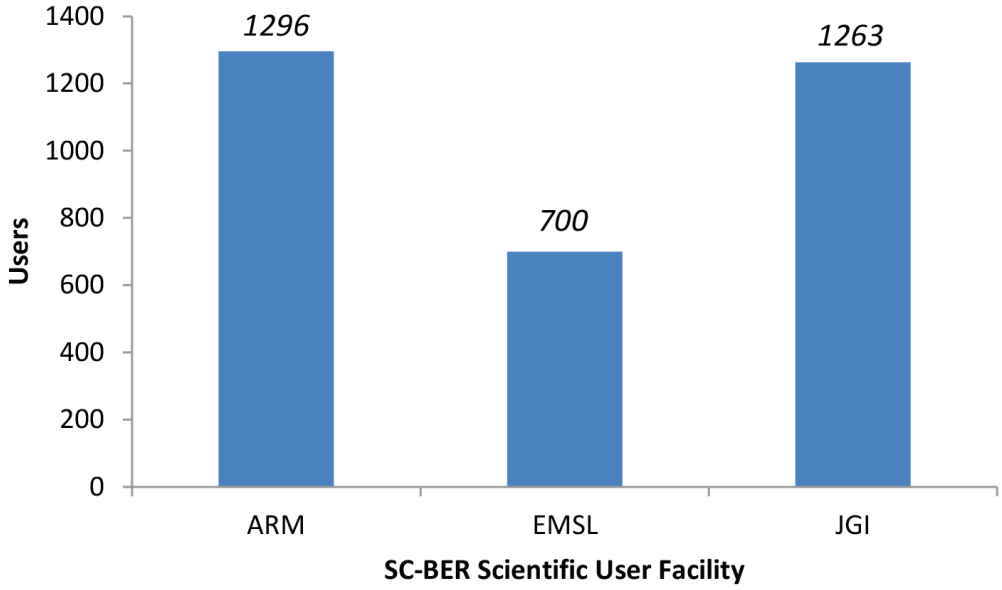
SC-BER hosts three user facilities–the Joint Genome Institute (JGI), the Environmental Molecular Sciences Laboratory (EMSL), and the Atmospheric Radiation Measurement (ARM) Climate Research Facility–that together serve a scientifically diverse user base that includes molecular biologists, chemists, and atmospheric scientists, to name a few. In FY2014, these three facilities hosted 3259 users, with more than a third (1307) originating from domestic academic institutions (Figures 5 and 6). The remaining two thirds of the users originated primarily from DOE national laboratories (776) and the international scientific community (817).

Figure 5 The number of users by sector of SC-BER-sponsored scientific user facilities. The total users are an aggregation of all users from the three facilities stewarded by SC-BER. The category "other" includes non-profit institutions, including laboratories, those affiliated with universities, hospitals, and hospital-affiliated research centers. Users with no institutional affiliation are also included in this category.



Sector	Users	%Total Users
Academia (domestic)	1307	40.1
DOE national laboratories	776	23.8
Other federal laboratories & agencies	158	4.8
Private sector (domestic)	72	2.2
International (all sectors)	817	25.1
Other	129	4.0
Total	**3259**	**100**

Figure 6 The total number of users for the three SC-BER-sponsored scientific user facilities for FY2014. The number for each facility represents unique users at that facility. Users that utilized more than one facility are counted as unique users for each facility.



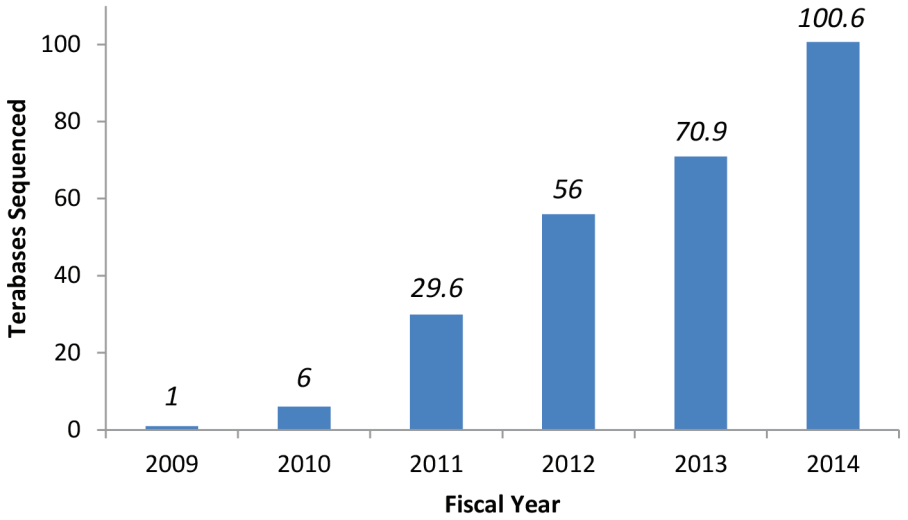
Scientific user facility abbreviation key:

- ARM – Atmospheric Radiation Measurement Climate Research Facility
- EMSL – Environmental Molecular Sciences Laboratory
- JGI – Joint Genome Institute

Technology advancements in high-throughput genome sequencing have enabled an increase in sequencing output by two orders of magnitude since 2009 while operating with a relatively flat budget (Figure 7). This dramatic increase in productivity has been concomitant with the generation of petabytes worth of sequencing data. JGI supports broad access to and analysis of this data by the scientific community through multiple web-based resources as well as a dedicated high-performance computing resource, the 8000+ core computing cluster Genepool located at and supported by NERSC. Additionally, the JGI partnership with NERSC has provided more than 10 million CPU hours on Hopper and Edison for JGI users. The JGI data portals and web-based resources saw more than 1 million unique visits in 2014.[11] To convert this data into knowledge that is useful for increasing understanding pertinent to DOE missions, JGI has started two new programs, the Emerging Technologies Opportunity Program[12] and the JGI-EMSL Collaborative Initiative,[13] to encourage development of new technologies and cross-user facility collaborations linking sequence to function.

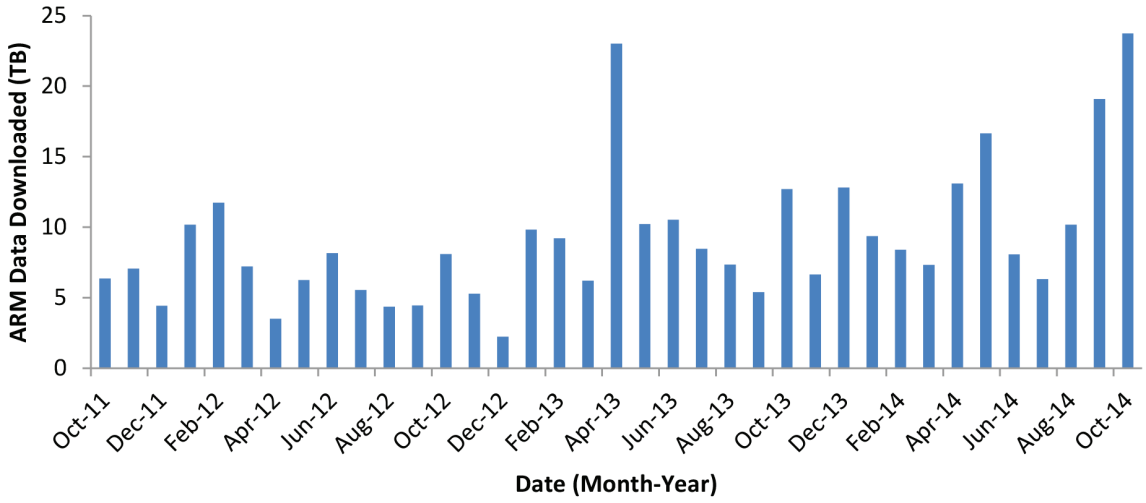
The observational data generated by the suite of fixed and mobile experimental facilities comprising the ARM Climate Research Facility are freely available to registered users through the ARM data archive (see section 9.5.3 of *Enabling Capabilities for Science and Energy*).[15] The data is of significant value to the climate research community, demonstrated, in part, by the incidence of ARM data citations in the Intergovernmental Panel on Climate Change *Fifth Assessment Report*.

Figure 7 The number of terabases (10^9 bases) sequenced at the Joint Genome Institute from FY2009 to 2014. For comparison, the human genome contains approximately 3 terabases.[14]



Technology advancements in remote sensing instrumentation, including advanced radar, lidar, and radiometer systems, are enabling higher spatial and spectral resolution observations of atmospheric clouds and aerosols and their impacts on the surface energy budget. Advances in radar technology, including 3-dimensional scanning cloud and precipitation radars, measurement of polarization, and high resolution Doppler spectra, have enabled scientists to study details of vertical motions, cloud structure, and microphysical properties of clouds and precipitation. Since 2011, data downloads have increased by a factor of 3, from 6.4 TB in October 2011 to 23.7 TB in October 2014 (Figure 8). Much of the increase in data download volume is due to the introduction of these new radar capabilities to the facility.

Figure 8 The number of terabytes (TB) of data downloaded from the Atmospheric Radiation Measurement (ARM) Climate Research Facility data archive from October 2011 through October 2014.



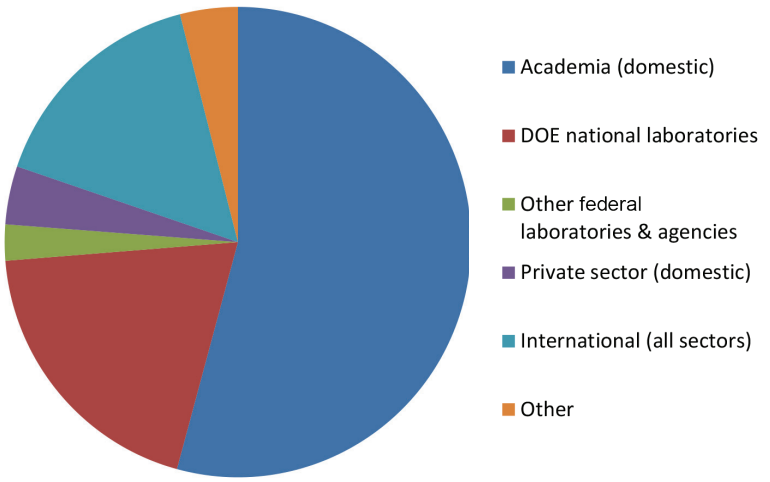
Office of Science, Office of Basic Energy Sciences (SC-BES)

The SC-BES program supports fundamental research to understand, predict, and ultimately control matter and energy at the electronic, atomic, and molecular levels in order to provide the foundations for new energy technologies and to support DOE missions in energy, environment, and national security. SC-BES enables this research through support of a diverse suite of scientific user facilities providing unique capabilities in characterization using X-rays, neutrons, and electrons, as well as in synthesis and fabrication. These facilities make possible leading-edge research that benefits from a merging of ideas and techniques from different disciplines.

SC-BES supports three classes of user facilities–X-ray light sources, neutron sources, and nanoscale science research centers–that together are used by more than 16,000 researchers each year representing many science and technology disciplines (see section 9.4 in *Enabling Capabilities for Science and Energy*). More than half of all users originate from domestic academic institutions (8738) (Figure 9). By virtue of the number (12 of the 28 SC-sponsored user facilities) and the broad scope of available experimental techniques, SC-BES facilities support the largest number of users from the domestic private sector (646, 54% of domestic private sector users across the 28 SC-supported facilities).

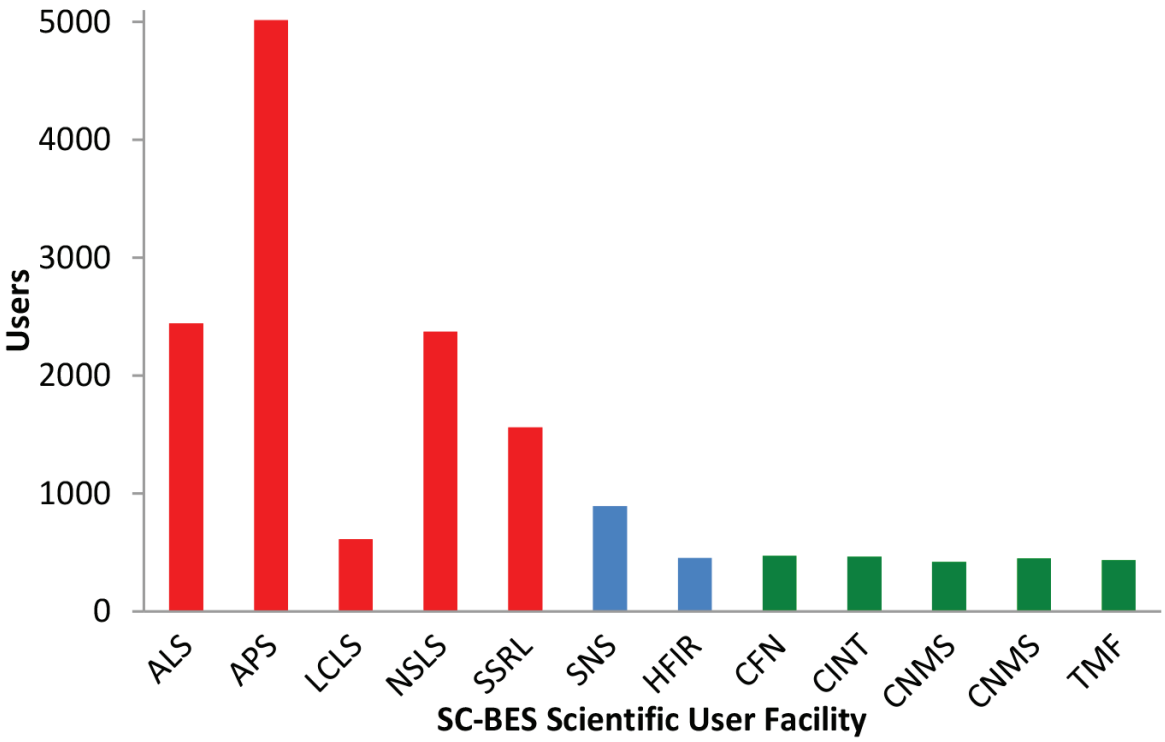
SC-BES's nanoscience research centers were purposefully co-located in proximity to X-ray and neutron sources. For example, Oak Ridge National Laboratory hosts the Spallation Neutron Source, High Flux Isotope Reactor, and Center for Nanophase Materials Science, while each of Argonne, Brookhaven, and Lawrence Berkeley National Laboratories host both a synchrotron light source and a nanoscale science research center. The coupling of these unique resources gives users the opportunity to leverage complimentary techniques to characterize novel materials and solve challenging scientific problems.

Figure 9 The number of users, by sector, of SC-BES scientific user facilities. Total users are an aggregation of all users from the 12 facilities stewarded by SC-BES. The category "other" includes PK12 schools, non-profit organizations (including research laboratories), hospitals and hospital-affiliated research centers, and users with no official affiliation.



Sector	Users	%Total Users
Academia (domestic)	8738	54.2
DOE national laboratories	3146	19.5
Other federal laboratories & agencies	399	2.5
Private sector (domestic)	646	4.0
International (all sectors)	2545	15.8
Other	644	4.0
Total	**16118**	**100**

Figure 10 The total number of users for the twelve SC-BES-sponsored scientific user facilities for FY2014. The number for each facility represents unique users at that facility. Users that utilized more than one facility are counted as unique users for each facility.



	X-ray Light Sources					Neutron Sources		Nanoscale Science Research Centers				
Facility	ALS	APS	LCLS	NSLS	SSRL	SNS	HFIR	CFN	CINT	CNMS	CNM	TMF
Users	2443	5016	612	2372	1559	893	453	473	465	421	451	433

Scientific user facility abbreviation key:

X-ray light sources:
- ALS – Advanced Light Source (Lawrence Berkeley National Laboratory)
- APS – Advanced Photon Source (Argonne National Laboratory)
- LCLS – Linac Coherent Light Source (SLAC National Accelerator Laboratory)
- NSLS – National Synchrotron Light Source (Brookhaven National Laboratory)*
- SSRL – Stanford Synchrotron Radiation Light Source (SLAC National Accelerator Laboratory)

Neutron sources:
- SNS – Spallation Neutron Source (Oak Ridge National Laboratory)
- HFIR – High Flux Isotope Reactor (Oak Ridge National Laboratory)

Nanoscale Science Research Centers:
- CFN – Center for Functional Nanomaterials (Brookhaven National Laboratory)
- CINT – Center for Integrated Nanotechnologies (Los Alamos and Sandia National Laboratories)
- CNMS – Center for Nanophase Materials Science (Oak Ridge National Laboratory)
- CNM – Center for Nanoscale Materials (Argonne National Laboratory)
- TMF – The Molecular Foundry (Lawrence Berkeley National Laboratory)

*The NSLS ceased operations in 2014. It has been replaced by the National Synchrotron Light Source – II (NSLS-II) at Brookhaven National Laboratory, which started operation in 2015.

Endnotes

[1] The Office of Science defines a user facility according to the criteria laid out in a 2012 memorandum, available at http://science.energy.gov/~/media/_/pdf/user-facilities/memoranda/Office_of_Science_User_Facility_Definition_Memo.pdf.

[2] The 34 DOE user facilities are given in table 9.1 of *Enabling Capabilities for Science and Energy*. 28 of the facilities are supported by the Office of Science, while 5 are supported by two of the DOE technology programs (4 by Energy Efficiency and Renewable Energy, 1 by Nuclear Energy). A final facility, the Wireless National User Facility, receives support from, but is not wholly owned by, the DOE. A complete, searchable list of DOE designated user facilities, including Web links, is available at http://energy.gov/technologytransitions/technology-transitions-facilities-database.

[3] 31 of the DOE user facilities are hosted by a DOE national laboratory. The DIII-D Tokamak and Alcator C-mod are hosted at General Atomics and the Massachusetts Institute of Technology, respectively. The Atmospheric Radiation Measurement Climate Research Facility is a suite of facilities hosted at multiple sites in the United States and internationally.

[4] As an example, work conducted at the five SC-BES X-ray light sources led to 3,907 peer-reviewed publications in FY2014. For the three SC-BER user facilities, 735 peer-reviewed papers were published. Users of the SC-ASCR facility NERSC reported 1,808 refereed publications (published or in press) in FY2014.

[5] Each user facility provides prospective users with information on the application and review process specific to that facility. This includes the review process, proposal scoring criteria, allocation procedures, and frequency of proposal calls. Interested readers are encouraged to review the relevant information available at the user facility's website for a more detailed description of their proposal process.

[6] As DOE designated user facilities, research conducted at the 5 user facilities sponsored by DOE-EERE and DOE-NE is free of cost for non-proprietary work; full cost recovery is required for proprietary work. The principles described above and laid out in the 2012 DOE-SC user facility memorandum (http://science.energy.gov/~/media/_/pdf/user-facilities/memoranda/Office_of_Science_User_Facility_Definition_Memo.pdf) have not been formally adopted by all of the DOE programs. However, the user facilities sponsored by DOE-EERE and DOE-NE similarly review proposals based on their scientific merit, the suitability of the facility capabilities for the proposed work, and alignment with DOE objectives. Specific information on becoming a user at these facilities is available through the facility websites and from relevant staff at the facilities.

[7] https://fortress.maptive.com/ver4/doe-sc-user-stats-fy-2014

[8] FY2014 user statistics source data for all DOE-SC scientific user facilities is available at http://science.energy.gov/~/media/_/excel/user-facilities/DOE-SC_User_Statistics_by_Institution_FY2014_rev3.xlsx.

[9] *Defining and Counting Users for the Office of Science User Facilities*, memorandum to the Office of Science Associate Directors from Acting Director/Deputy Director of Science Programs Patricia M. Dehmer, August 23, 2013. http://science.energy.gov/~/media/_/pdf/user-facilities/memoranda/SC_User_Statistics_Memorandum.pdf.

[10] *DOE Office of Science User Facilities User Statistics Collection Practices*, http://science.energy.gov/user-facilities/policies-and-processes/user-statistics-collection-practices.

[11] *2014 Progress Report*, U.S. Department of Energy Joint Genome Institute, http://jgi.doe.gov/wp-content/uploads/2013/11/2014-JGI_Progress_Report.pdf.

[12] JGI supports development of new technologies that both extend and complement the high-throughput sequencing techniques available to users at their main facility through the Emerging Technologies Opportunity Program. Technology development projects supported by this program are critical to the development of high-throughput annotation techniques that will enable researchers to link genome sequence to function. A list of current projects and more information about the program is available at http://jgi.doe.gov/2015-call-for-etop-letters-of-intent/.

[13] The JGI-EMSL Collaborative Science Initiative (JECSI) enables researchers to leverage the unique high-throughput genomic sequencing capabilities of JGI with the molecular characterization techniques of EMSL in a single project proposal. Proposals for this initiative must be high-risk/high-reward in a focused topic area and have a realistic expectation of completion on an accelerated time scale. Further, proposals must leverage a range of capabilities at both facilities and generate data sets beyond the scope of a single facility. Projects from FY2014 and FY2015, as well as more information about the initiative, are available at http://jgi.doe.gov/collaborate-with-jgi/community-science-program/how-to-propose-a-csp-project/emsl/.

[14] *2014 Progress Report*, U.S. Department of Energy Joint Genome Institute, http://jgi.doe.gov/wp-content/uploads/2013/11/2014-JGI_Progress_Report.pdf.

[15] The ARM data archive is located at http://www.archive.arm.gov/armlogin/login.jsp.

Acronyms

ALCF	Argonne Leadership Computing Facility
ALS	Advanced Light Source
APS	Advanced Photon Source
ARM	Atmospheric Radiation Measurement Climate Research Facility
CFN	Center for Functional Nanomaterials
CINT	Center for Integrated Nanotechnologies
CNM	Center for Nanoscale Materials
CNMS	Center for Nanophase Materials Science
DOE	Department of Energy
DOE-SC	Office of Science
EMSL	Environmental Molecular Sciences Laboratory
ESNet	Energy Sciences Network
HFIR	High Flux Isotope Reactor
JGI	Joint Genome Institute
LCLS	Linac Coherent Light Source
NERSC	National Energy Research Scientific Computing Center
NSLS	National Synchrotron Light Source
NSLS – II	National Synchrotron Light Source - II
OLCF	Oak Ridge Leadership Computing Facility
PK12	Pre-kindergarten to 12th Grade education
SC-ASCR	Office of Advanced Scientific Computing Research
SC-BER	Office of Biological and Environmental Research
SC-BES	Office of Basic Energy Sciences
SC-FES	Office of Fusion Energy Science
SC-HEP	Office of High Energy Physics
SC-NP	Office of Nuclear Physics
SNS	Spallation Neutron Source
SSRL	Stanford Synchrotron Radiation Light Source
TB	terabyte
TMF	The Molecular Foundry