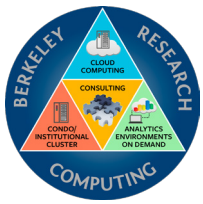
Campus Impact of Berkeley Research Computing

Berkeley Research Computing (BRC) provides consulting and computing resources to campus researchers. It was developed and continues to evolve in response to faculty needs. BRC is supported as a partnership among the AVC-CIO, the VCR, and the Chancellor. This document describes the impact of the program through 2016.



I. How does BRC support Research, Recruitment, and Teaching?

After review with our peers, our sponsors, and the BRC User Advisory Group, we began this past year to standardize our impact metrics, and to gather associated data. A survey at the end of AY2016 yielded data and some interesting narratives, and we continue to update this.

Grants supported

In the survey results in July 2016 BRC users reported **$38,023,970 in awarded grants** that are supported by BRC, and another **$4,289,381 in submitted grants** that will be supported by BRC if proposals are funded. Given a roughly 20% response rate, we conservatively extrapolated that BRC supports $80M to $120M in grants, with another $10M to $15M in submitted grants.

This fall, BRC staff worked with the Research Administration and Compliance (RAC) team to collate the value of all grants awarded to all PI's that are supported with BRC services. Based upon this analysis:

BRC supports ~180 PIs with nearly $600 million in awarded grants
and contracts that were active since BRC launched in 2014.

Publications and Conference Proceedings

As reported in the survey, 96 publications and conference proceedings were supported by BRC in the period 2014-2016. Given a roughly 20% response rate to the survey, we conservatively extrapolate that the program supported **at least 200-300 publications**.

Recruitment and Retention

From previous discussions with faculty partners, and data reported in the survey, we know of **15 faculty recruitment or retention cases in which BRC played a role**, along with two additional postdoc recruitment cases. Some quotes from recently recruited faculty, postdocs, and faculty involved in recruitment:

"I am just joining the Berkeley astronomy department and access to Savio helped with my recruitment."

"It was great to be able to use the language about the Faculty Computing Allowance in [our] offer letter."

"[We are] currently making two offers where Savio is an integral part of the recruitment offer."

"I am aware of two recent cases where the availability of BRC services - in particular the condo cluster - was very helpful in the recruitment of potential faculty."

"High-performance computing at Berkeley was an important factor in deciding to come here for a postdoc."

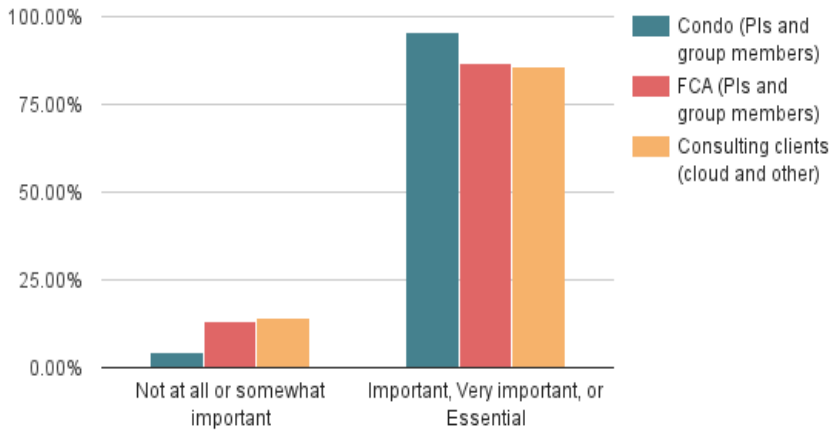
Courses

Survey respondents described **seven current or planned courses** that are or would be supported by the BRC program, in a range of departments/domains: Computer Science, Civil and Environmental Engineering, Physics, Genomics/Computational Biology, Environmental Science Policy Management, and the Social Science D-Lab.

The new **Instructional Computing Allowances** (ICA's) provide a formal means of supporting classes with computational needs. We have two ICA's currently in process, and several more in early discussions.

We are working closely with the Data Science Education program to support their emerging needs as well.

Importance of BRC to research



An **overwhelming majority** of survey respondents reported that the **BRC program is Important, Very Important, or Essential to their research**. The figure below shows this, broken down by the type of engagement the researcher has with BRC.

The broad reach of BRC Programs across campus domains

A major goal for the program is to provide support all across the campus, and not just for the traditional sciences that use high performance computing. BRC has made good progress on this: in addition to supporting the Physical, Life, and Geo Sciences, we serve researchers in Engineering, Computer and Information Sciences, Social, Behavioral, and Economic (SBE) Sciences, and the Arts and Humanities. As Data Science methods and tools are being taken up across the academy, BRC is providing consulting to faculty on how to use computational resources to advance their research.

BRC consultants have worked with researchers in over 85 departments, labs, and centers. Figure 1 summarizes this activity; the full list is detailed in Appendix C.

Note that while the Physical Sciences, Geosciences, and Engineering consume the largest proportions of the Savio capacity (see Figure 2), the Biological Sciences and Social, Behavioral, and Economic Sciences often request more consulting. This likely reflects the more recent shift in those domains to computationally intensive research, and their use of newer, emerging computational resources (such as cloud computing).

In addition to the nearly 800 Savio users, several hundred researchers are supported by the cloud consulting services, and the more recently launched AEoD service will grow to 100 users in the next year. We estimate that represents about ¼ to ⅓ of the faculty and graduate students who may need computing resources for their research, and so we expect the user base (and the associated need for consulting) to continue to grow for some time.

Figure 1 Consulting engagements by domain

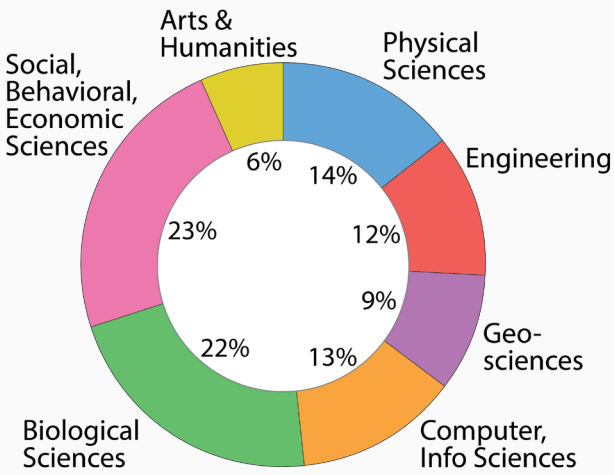
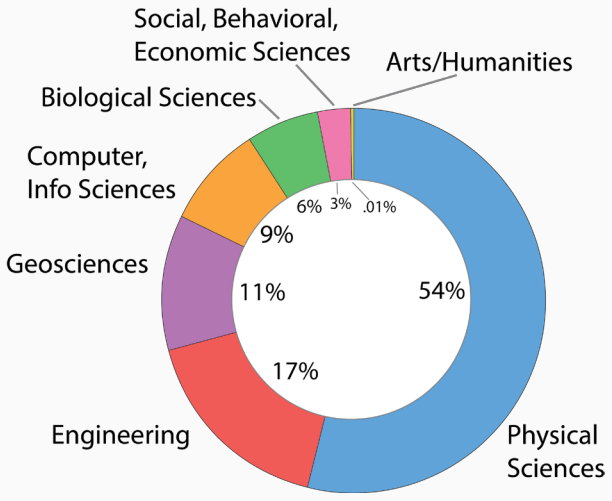


Figure 2 Savio utilization by domain



The new Instructional Computing Allowances will further expand access, and may add many undergraduates to the community of users (especially for upper division (*connector*) courses in Data Science.

Quotes from survey respondents include:

> *"**The BRC services made a world of difference** in my group's ability to access cluster resources. Previously, I had maintained my own cluster and it was a very costly process. **The current setup just works a lot better.**"*
> *- John Chiang, Associate Professor, Dept, of Geography & Berkeley Atmospheric Sciences Center*

> *"**BRC servers have been always very fast and reliable, and the staff has been very helpful.**"*
> *- Hiroshi Nikaido, Professor of Biochemistry and Molecular Biology*

> *"**Very helpful in moving our program forward with AWS services**"*
> *- Brian Peterson, PATH/Institute of Transportation Studies"*

II. Leveraging the BRC program

In addition to the research supported and the impact metrics described above, there are additional ways that the BRC program can be leveraged in support of the campus mission, and to help address campus financial challenges:

1. Leveraging the Faculty Computing Allowance as cost-share. Beginning in 2017, BRC will provide 300,000 SUs (Service Units, roughly equal to core-hours) per year, up from 200,000 in previous years. Using an Indiana University valuation[1] of $0.05/SU, the FCA is worth $15,000/year for each faculty member. A similar valuation can be associated to Instructional Computing Allowances, for training-related grants.
2. The campus investment in the Condo computing program subsidizes infrastructure costs as well as system administration, for condo contributions. For a modest 8-node configuration, this is worth over $6000/yr, or $30,000 over the life of the cluster[2]. For a larger configuration of 24 nodes (not uncommon for research groups) the campus support would be somewhat over $13,000/yr, or $65,000 over the cluster life.

III. Appendices

Appendix A: Departments and Labs supported by BRC

Appendix B: Profiles of researchers supported by BRC

[1] See: "How much money is an XSEDE allocation worth?" https://kb.iu.edu/d/bapq. Accessed 12/9/2016. XSEDE operates at a much larger scale (with associated economies of scale), and so this likely is a low estimate for the value of BRC SUs.

[2] Per node: Colocation: .5 RU * 14$/RU/mo = $84/yr. Racks, cooling doors, network switches, and cables cost $242+$486/node. Current admin costs at LBNL: $300/mo + $14/node/mo. Total for 8 nodes: $5616. This does not include the cost of the network, the parallel filesystem, etc., and so is very conservative. We conservatively assume a 5 year lifetime for the nodes.

Appendix A: Departments and Labs supported by BRC

Agricultural & Resource Economics	Data Science Education Program	Library
Anthropology	Digital Humanities at Berkeley (DH)	Linguistics
Archaeological Research Facility	Dutch Studies	Materials Science and Engineering
Astronomy	Earth and Planetary Science (EPS)	Mathematics
Berkeley Institute for Data Science	Ecology	Mechanical Engineering
Berkeley Institute of Design	Econometrics Laboratory (EML)	Molecular & Cell Biology
Bioanthropology	Economics	Museum of Vertebrate Zoology
Bioengineering	Electrical Engineering and Computer Science/EECS	Near Eastern Studies
Bioinformatics	Energy & Resources Group (ERG)	Neuroscience
Biophysics	Energy Institute at Haas	Nuclear Engineering
Biostatistics	English	Nutritional Sciences & Toxicology
California Institute for Quantitative Biosciences (QB3)	Environmental Science, Policy & Management	Partners for Advanced Transportation Technology
Cancer Research Laboratory	Epidemiology	Physics
Center for Computational Biology	Ethnic Studies	Plant and Microbial Biology
Center for Effective Global Action	Gender and Women's Studies	Political Science
Center for Information Technology in the Interest of Society (CITRIS)	Geography	Psychology
Center for Open Science (COS)	Geospatial Innovation Facility	School of Information
Center for Science, Technology, Medicine, & Society (CSTMS)	Goldman School of Public Policy	School of Law
Chemical Engineering	Graduate School of Education	School of Public Health
Chemistry	Haas School of Business	School of Social Welfare
CIRM/QB3 Shared Stem Cell Facility and High-Throughput Screening Facility	Hearst Museum of Anthropology	Seismology
City & Regional Planning	Helen Wills Neuroscience Institute	Social Science D-Lab
Civil and Environmental Engineering	History	Social Science Matrix
College of Engineering	History of Art	Sociology
College of Environmental Design	Industrial Engineering & Operations Research	Space Sciences Laboratory
College of Letters and Science (L&S)	Institute of Transportation Studies	Statistical Computing Facility
College of Natural Resources (CNR)	Integrated Cancer Research	Statistics
Computational Genomics Research Lab	Integrative Biology	Visual Resources Center
Computer Science		

Appendix B – Profiles of researchers supported by BRC services

BRC interns have been interviewing some of the researchers using BRC services, and documenting their stories on our website. Here are some excerpts and quotes from those profiles:

Eric Neuscamman, Assistant Professor of Chemistry
in: Theoretical chemists use Savio to build molecular computational models

Theoretical chemists use supercomputers to solve complex mathematical equations … that yield predictive models of what molecules will do in chemical reactions. "I can add to the chorus of voices that say we, ***without Savio, would not have a lab. It's like our wet lab****. We would have no lab without a computer to run our calculations on."*

Alexander Tchekhovskoy, Postdoctoral Researcher (and Einstein Fellow), working with Eliot Quataert, Professor of Astronomy and Physics
in: BRC Program supports Astrophysicists' research on black holes

*"****A local resource like Savio is really useful because we need a day-to-day ability to test new ideas****. You may not know what physics effect is important, so you want to experiment, to run many models... If we can try something, and get an answer in a few hours, that is great. At a national center, you may wait a day or more for job to even start.* ***Getting fast***

turnaround is a fantastic capacity that we cannot get elsewhere. *When we are developing new software models, we need a local resource to test and debug our software."*

"Another lifesaving capacity is visualization support. ***The BRC team got our visualization tools to work in the Savio environment. No-one else in other supercomputing centers could get these tools to work. Everyone struggles with these issues at other Universities, but here it just works, which makes a huge difference to my research.****"*

Jennifer Ahern, Associate Professor of Epidemiology (and Chancellor's Professor of Public Health), and Ellie Matthay and Scott Zimmerman, Research Data Analysts in the School of Public Health
in: Epidemiologists develop a computational tool to optimize study design

[Our team] worked closely with BRC consultants on the simulator's software architecture so that it could be run on Amazon AWS and Microsoft Azure servers in the commercial cloud, to take advantage of resource grants those vendors awarded to the Study Simulator team; as well as on Savio, to leverage BRC's Faculty Computing Allowance.… ***Ahern says, "these services are making researchers feel like their projects are supported, and will be able to thrive. The Study Simulator wouldn't be happening without BRC."***

Stephen Floor, Postdoctoral Researcher (and Helen Hay Whitney Fellow), working with Jennifer Doudna, Professor of Biochemistry, Biophysics, and Structural Biology
in: Savio supports Doudna Lab biophysicist's investigation of human protein production

Floor's work ***processing RNA sequencing data ... only needs about 3 hours to finish a task on Savio, compared to the previous 4 days on his own computer****. And because the Doudna Lab's aggregate computing needs fit within the free BRC Faculty Computing Allowance, research funding remains available for other necessary expenditures…. "I've been in science for 10 to 15 years now, and … it's not like every university has this.* ***This is pretty special****."*

Evan Lyall, Biophysics grad student, working with Hillel Adesnik, Asst. Professor of Neurobiology
in: Systems Neuroscientists use BRC services to understand the neural basis of perception

Using the Savio cluster has sped up analysis of "data-dense videos of neural activity by an order of magnitude*, from multiple hours to a matter of minutes. We have recently begun working with Maurice Manning, a consultant at BRC, who is helping us to design a pipeline to analyze our [videos of neural activity] in near real-time." The realm of systems neuroscience is collectively moving to the acquisition and analysis of larger and larger datasets. We* ***will be fully reliant on the services provided by the BRC in our pursuit of understanding*** *how the cortex functions."*

Ron Cohen, Professor of Chemistry and Earth and Planetary Science, graduate students Alexis Shusterman, Josh Laughner, and visiting Harvard graduate student Alex Turner
in: Atmospheric chemists use Savio to build computational gas emission models

Beyond its utility as a modeling workhorse, the Savio cluster serves as a teaching tool within Cohen's research team*, on which new graduate students learn how to do the computation associated with the lab's research goals and pilot their calculations before running even larger tasks on Yellowstone, a massive nationally-supported HPC cluster operated by the National Center for Atmospheric Research.*

Kelly Rowland, graduate student working with Rachel Slaybaugh, Assistant Professor of Nuclear Engineering
in: BRC program supports neutron transport research and advanced nuclear reactor design

*"****This numerical assessment of reactor designs … can only be accomplished efficiently on a high performance computing cluster ... such as Savio*** *[which is] designed to optimize this kind of parallel computing." Rowland is a researcher and a domain consultant embedded in the department of Nuclear Engineering. "****With BRC consulting we are attempting to bridge the clear gap between undergraduate training and graduate expectation for computational research****."*

Smitha Milli, EECS undergraduate, working with David Bamman, Assistant Professor, School of Information
in: Undergraduate student uses Savio to perform Natural Language Processing on Fanfiction

Milli emphasizes that "running this [Natural Language Processing] pipeline takes a lot of compute power and time, so it is essential for us to be able to parallelize our work on the [Savio] cluster." ***BRC's consulting staff "was also incredibly helpful,"*** *Milli continues, "****They're doing a good job of taking in feedback from users, and making the cluster easier and easier to access****."*

*Bamman says, "****There's no question we'll keep using the free Faculty Computing Allowance. That has been essential for this work*** *... Having the ability to redistribute that allowance to other students in my research group is hugely important, and will be a continued strategy to use going forward."*