

Campus Bridging Through Facilitation: ACI-REF, From Project to Consortium

Contents courtesy of ACI-REF project team

KC Wang

Associate Professor - Clemson University

Computing Director - Center of Excellence for Next Gen Computing & Creativity

Networking CTO - CCIT



NSF-Funded Project – ACI-REF

\$5.3M NSF Award supports the project leadership team and 2 Facilitators for each of the 6 partner sites for 2 years.



PI: Jim Bottum, **Clemson**

Project Leadership:

- James Cuff, **Harvard** (PI Chair)
- Maureen Dougherty, **USC**
- Gwen Jacobs, **Hawaii**
- Paul Wilson, **Wisconsin**
- Tom Cheatham, **Utah**
- Barr von Oehsen, **Clemson**

Facilitator Lead: Bob Freeman, **Harvard**

Chief Scientist: Miron Livny, **Wisconsin**

Background and Context

National – HPC as Demand Driver

- Labs, Centers, PACI, TeraGrid, XSEDE, OSG

Campus Computing Demand Growing in Parallel

- MRIs, CRIs, Start-Up Packages
- Condo and Co-lo Approaches
- Big Data Driving New Communities

University based research computing operations thin on people
ESPECIALLY user-facing people

Result

- Training and education gap between resources and researchers – high barrier to entry without human assistance
 - ...and the barriers become higher as we bring in new communities

Takeaway: We Need People!

Campus Bridging Takes People

The need for a new workforce – a new flavor of **mixed science and technology professional** – is emerging. These individuals have expertise in a particular domain science area, as well as considerable expertise in computer science and mathematics. Also needed in this interdisciplinary mix are professionals who are **trained to understand and address the human factors** dimensions of working across disciplines, cultures, and institutions using technology-mediated collaborative tools.”

-2003 NSF Blue Ribbon Advisory Panel on Cyberinfrastructure

- Research enablement takes concentrated, committed effort by campuses and organizations to the right people.
- Typical research problems today are not technology or infrastructure problems – it’s navigating the complexity, which takes people.
- Nationwide gap on campuses of these professionals, while demand grows.

An Approach

Answered Need: The ACI-REF Project (People)

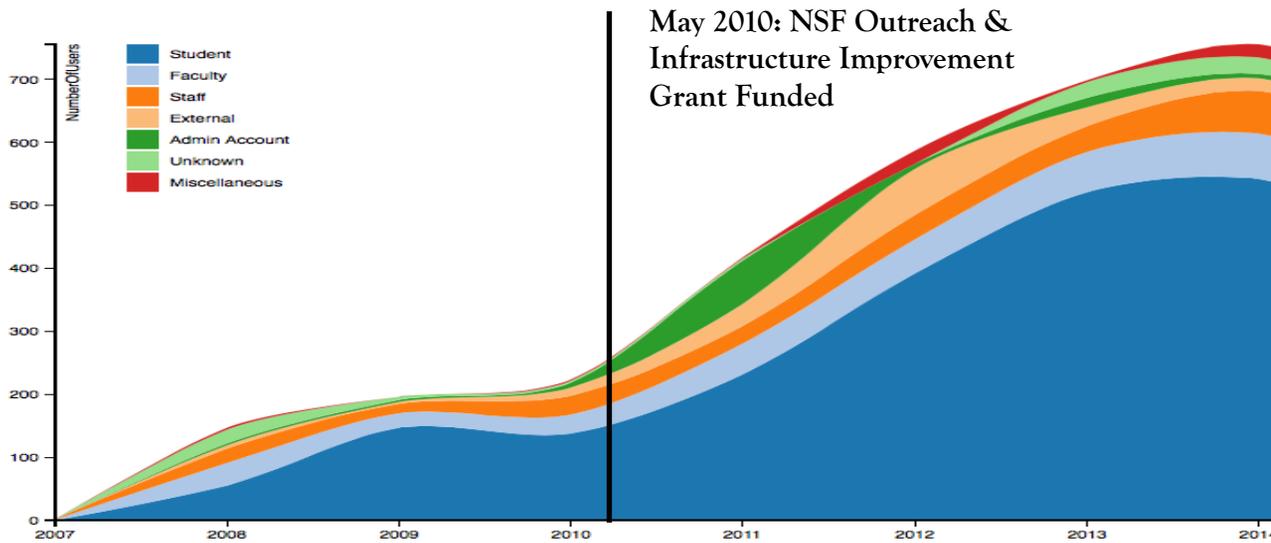
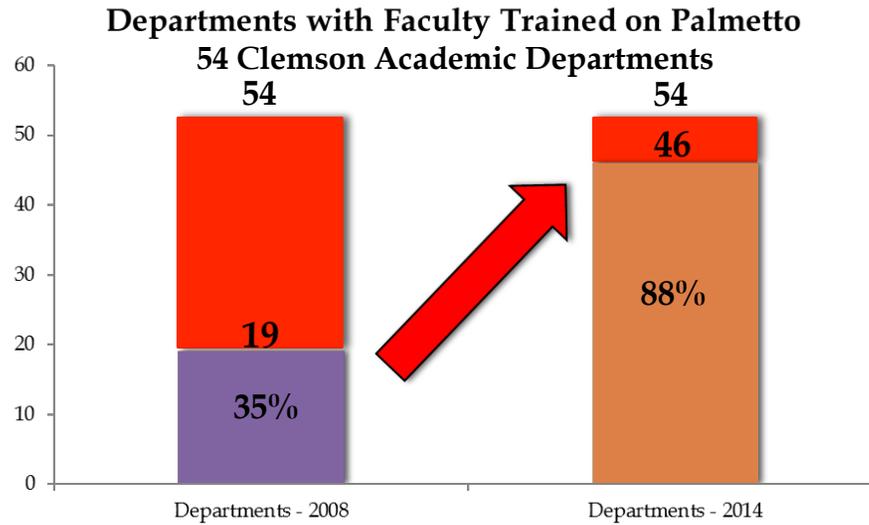
Goal: Seed investments in user-facing people – **facilitators** – at campuses to:

- Assist researchers in taking advantage of advanced computing resource investments, especially at the local campus level; and
- Build inter-institutional collaborative networks of knowledge to share expertise across campuses.



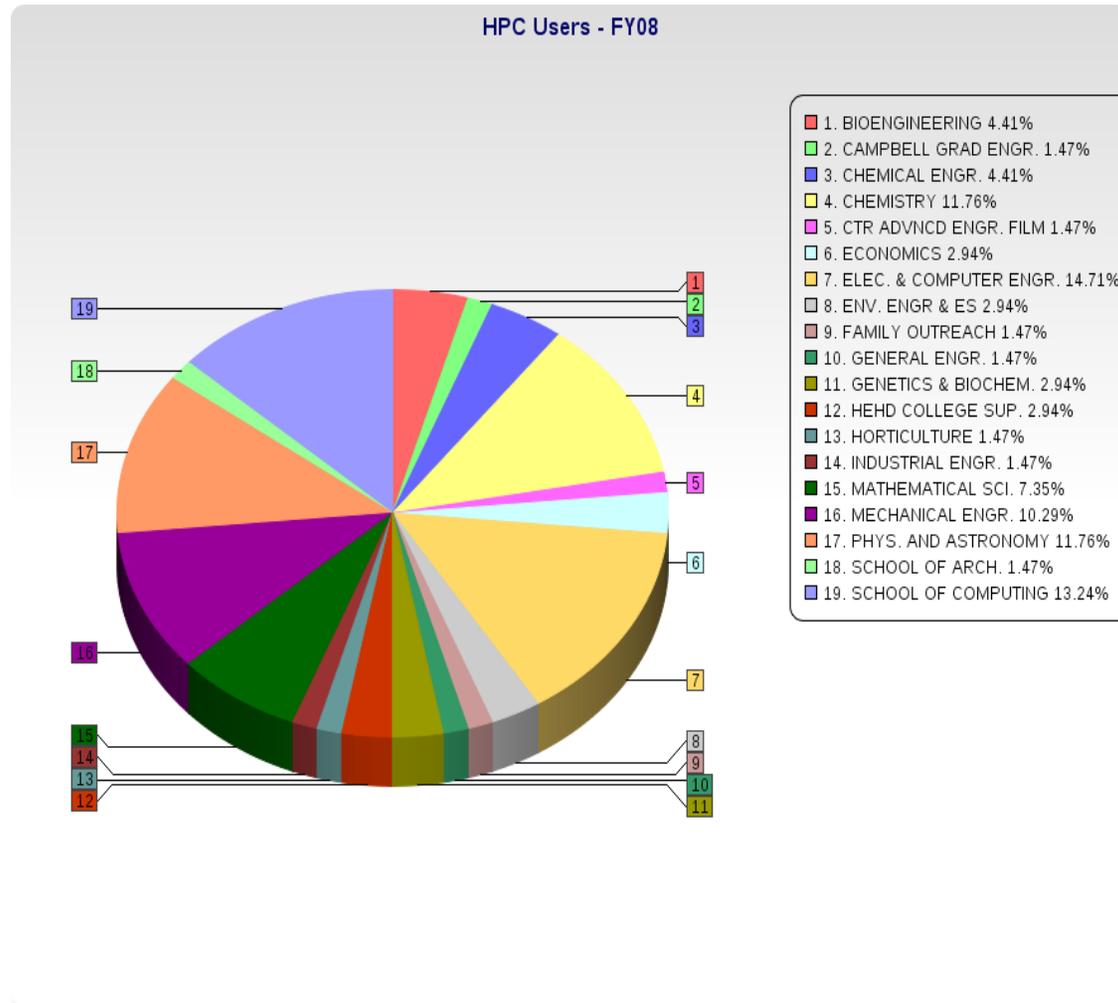
One Campus' Experience

Clemson
 May 2010 –
 first Clemson
 “facilitator”
 funded



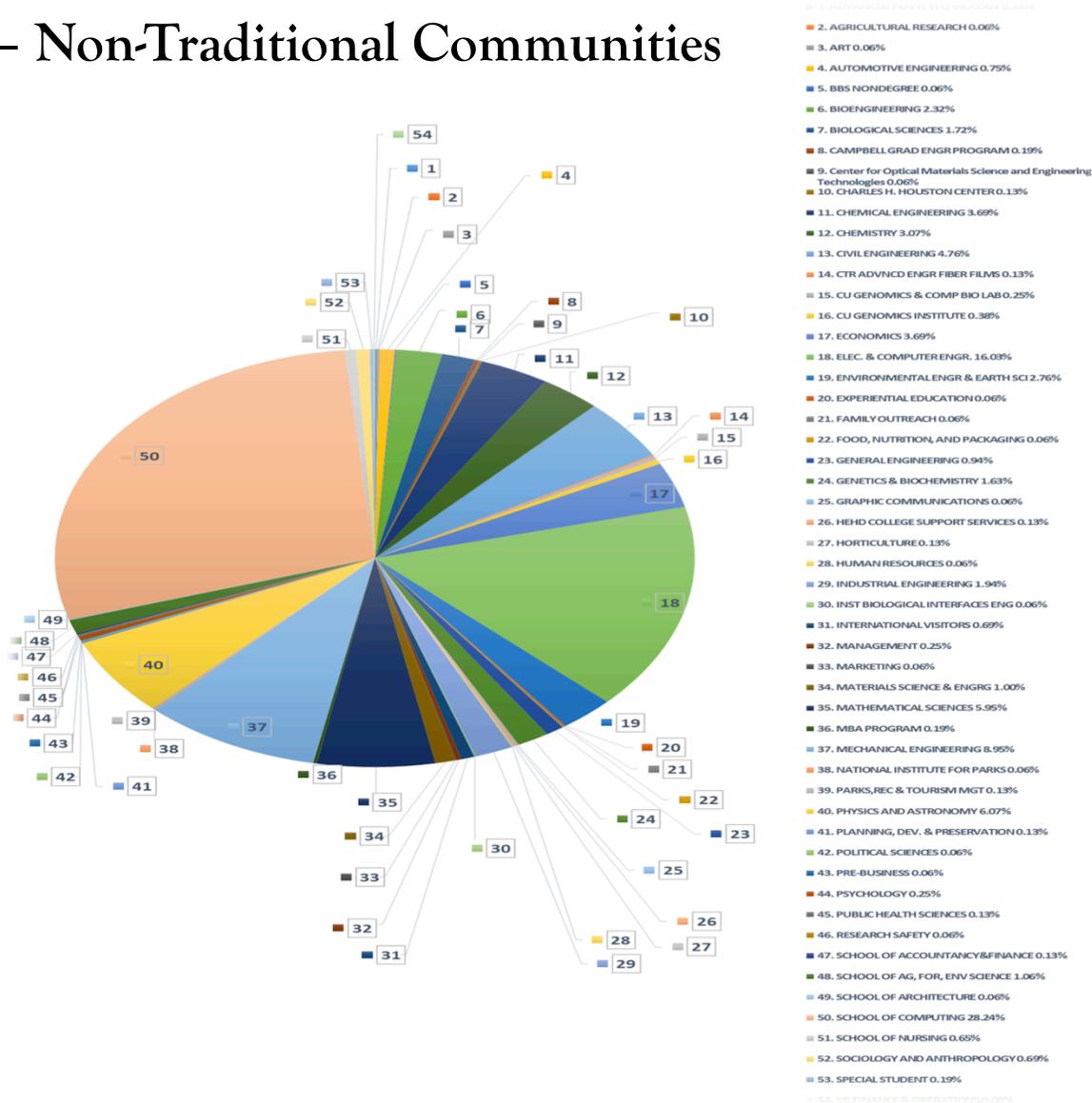
Campus Community Growth

2007/2008 – The Usual Suspects



Campus Community Growth

2014 - Non-Traditional Communities



Non-Traditional Impacts

Case Study: Hadoop In Action

Kevin McKenzie, Clemson Chief Information Security Officer, and his team used the Clemson Hadoop platform as part of a recent security incident response. His team needed to evaluate multi-year volumes of log data from the Clemson network to validate the extent of an incident they were investigating. The team loaded log data into the Hadoop cluster to gain a higher performance of log analysis.

Using Hadoop proved very beneficial, as it eliminated the estimated weeks (if not months) to accomplish on current local systems and the analysis was completed in less than a couple of hours, allowing the team to more quickly determine the extent of the issue.

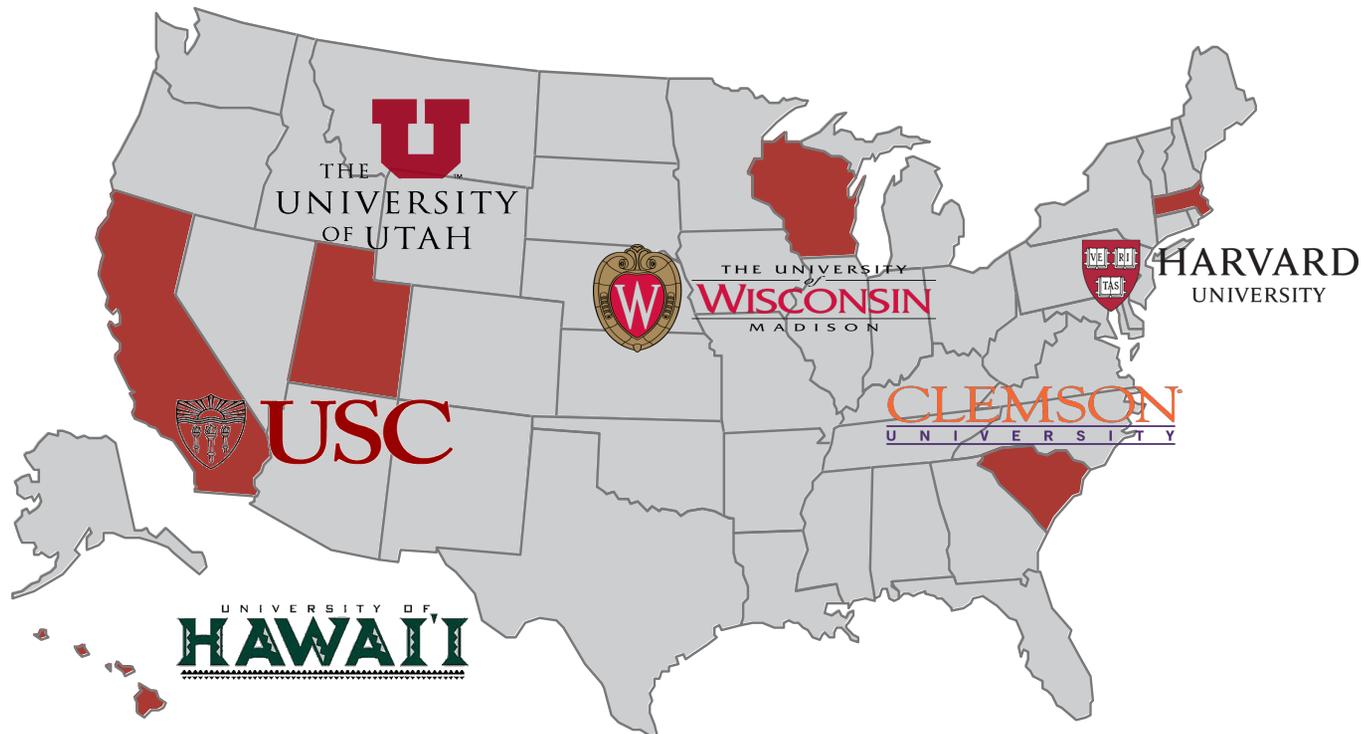
Departments Receiving Hadoop Training

Bioengineering	Management
Chemical Engineering	Mathematical Sciences
Chemistry	Mechanical Engineering
Civil Engineering	Medicaid IT Services
Economics	Physics And Astronomy
Elec. & Computer Engr.	Public Health Sciences
Environmental Engr & Earth Sci	Research Safety
Experiential Education	School of Ag. For. Env Science
General Engineering	School of Computing
Genetics & Biochemistry	Univ Facilities Support Svcs
Industrial Engineering	VP Finance & Operations
International Programs	Information Security & Privacy
Law Enforcement & Safety	



ACI-REF Formation

- Award for NSF-sponsored workshops held in 2012 helped define the needs of the broader community
- Goal: Advance our nation's research & scholarly achievements through the transformation of campus computational capabilities and enhanced coupling to the national infrastructure.



Progress – 2015

August 2015 – ACI-REF Project Halfway Mark

- Participating campuses saw:
 - Estimated 17% Growth in Departments Served
 - Estimated 16% Growth Number of Advanced Computing Users
 - 821 Individual Consultations with ACI-REFs since March 2015
 - 1288 Training Attendees in Sessions by ACI-REFs since March 2015
 - 75 Training Sessions Given by ACI-REFs since March 2015
 - 317M+ Core Hours Delivered by ACI-REF Campuses since March 2015
 - At least 65 Non-Traditional Departments Using ACI Across Campuses
- Other Notables:
 - **Multi-site Support Network** – for example, ACI-REF at USC able to tap into Clemson GIS ACI-REF to help a faculty member with their research
 - **Best Practices Manual** – ACI-REFs have developed a “best practices for facilitation” manual to aid in the onboarding of new facilitators and to formalize the practice
 - **Office Hours** – most ACI-REF campuses are now holding set advertised hours each week where researchers can stop by and get help with any problems or ask questions

Example Successes – Clemson & OSG

The Systems Genetics Lab at Clemson University (F. Alex Feltus, PI; Will Poehlman, PhD student) requires high performance computing (HPC) to build and interpret biomolecule interaction networks (node-edge graphs) to discover gene sets underlying complex traits in plants and animals. Dr. Feltus regularly interacts with Clemson ACI-REFs on his research needs.

“Experiment completion time is highly queue dependent of course, but as an example, we recently broke a single experiment into 12,000 jobs (1GByte RAM each) and launched on OSG which took 19.5 hours to complete. In contrast, the same experiment would have taken 14-21 days to complete on the Palmetto cluster, which is also part of OSG, given the PI’s resource allocations.”

– F. Alex Feltus, Clemson University

Clemson contributed **over 300K core hours** last month to the OSG pool
– in line with the spirit of sharing in the ACI-REF project.

Example Successes – Utah

High-Energy Theoretical Physics – Chris Kelso, University of Utah

*“I work in high energy theoretical particle physics. Specifically, I investigate physics beyond the Standard Model with a focus on dark matter implications. My research often requires scans of models that have very large numbers of parameters. This work could not be completed without the computing resources provided at CHPC. Almost as valuable as the use of the CHPC machines was the extremely helpful assistance I received from **Wim R. Cardoen**. Many of the codes I often use are serial, open source code that has been developed by many physics experts. To try and convert these codes to parallel would be a monumental task. Wim worked very hard to help me to find a solution that allowed this serial code to still utilize the numerous processors available on the CHPC machines. Without this, my projects would take months to finish, rather than a few days.”*

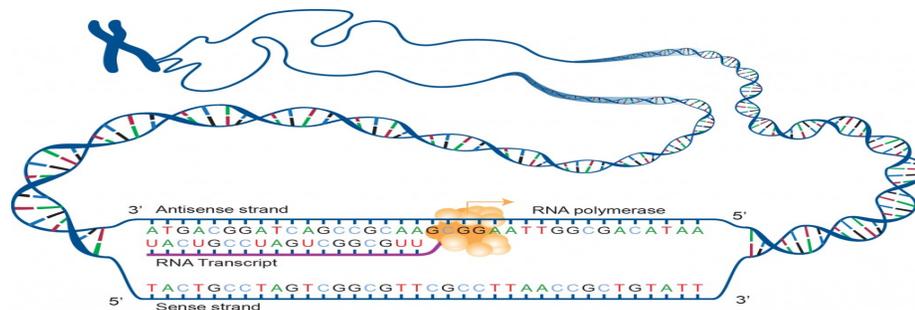
– Chris Kelso, University of Utah PostDoc, on Utah
ACI-REF Wim Cardoen



Example Successes – Harvard

HPC Assistance in Biology Software and Workflow – Zack Lewis, Harvard University

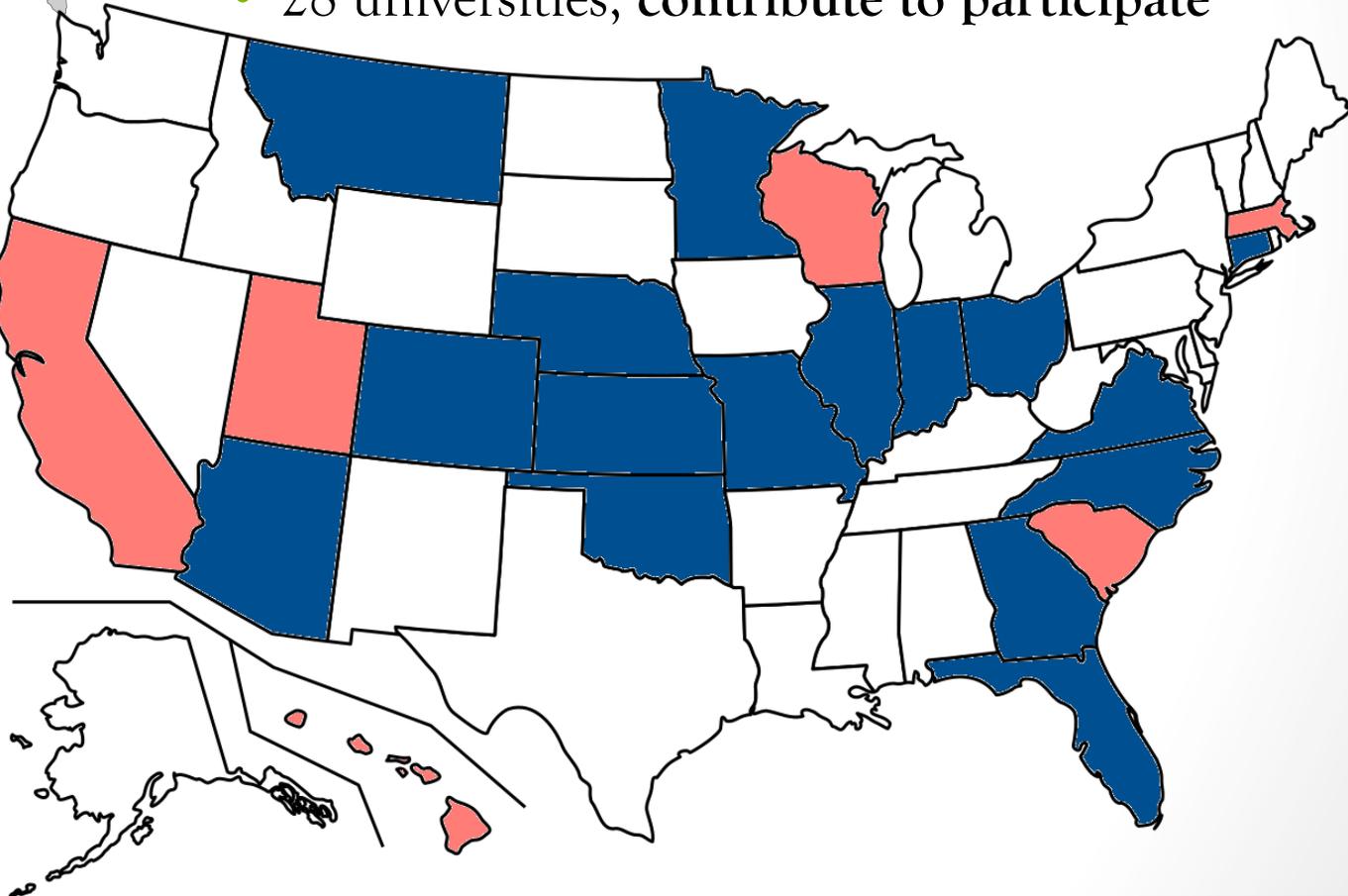
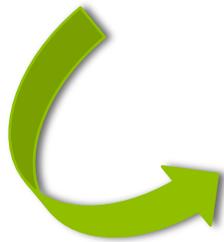
*“I am a sixth year graduate student in the Department of Organismic and Evolutionary Biology. I started a transcriptomics project with little experience in coding and no experience in high powered computing (HPC). Without **Bob Freeman’s** work through ACI-REF I do not think I would have been able to complete my bioinformatics project. I was not aware of ACI-REF at the time I started my HPC bioinformatics work. To my good fortune I happened to connect with Bob Freeman at the weekly Research Computing office hours. Bob has accompanied me nearly every step of the way along my 6 month journey into HPC. Bob’s help has taken the form of instruction on coding, monitoring active jobs, writing and adapting scripts for my project, as well as connecting me with researchers working on similar problems or at similar stages in learning transcriptomics. In particular, building connections with other researchers at Harvard through ACI-REF has been one of the most useful experiences. I now often work through my HPC issues with graduate student and postdoc peers that I have connected with through Bob.” – Zack Lewis, Harvard University PhD Candidate, on Harvard ACI-REF Bob Freeman*



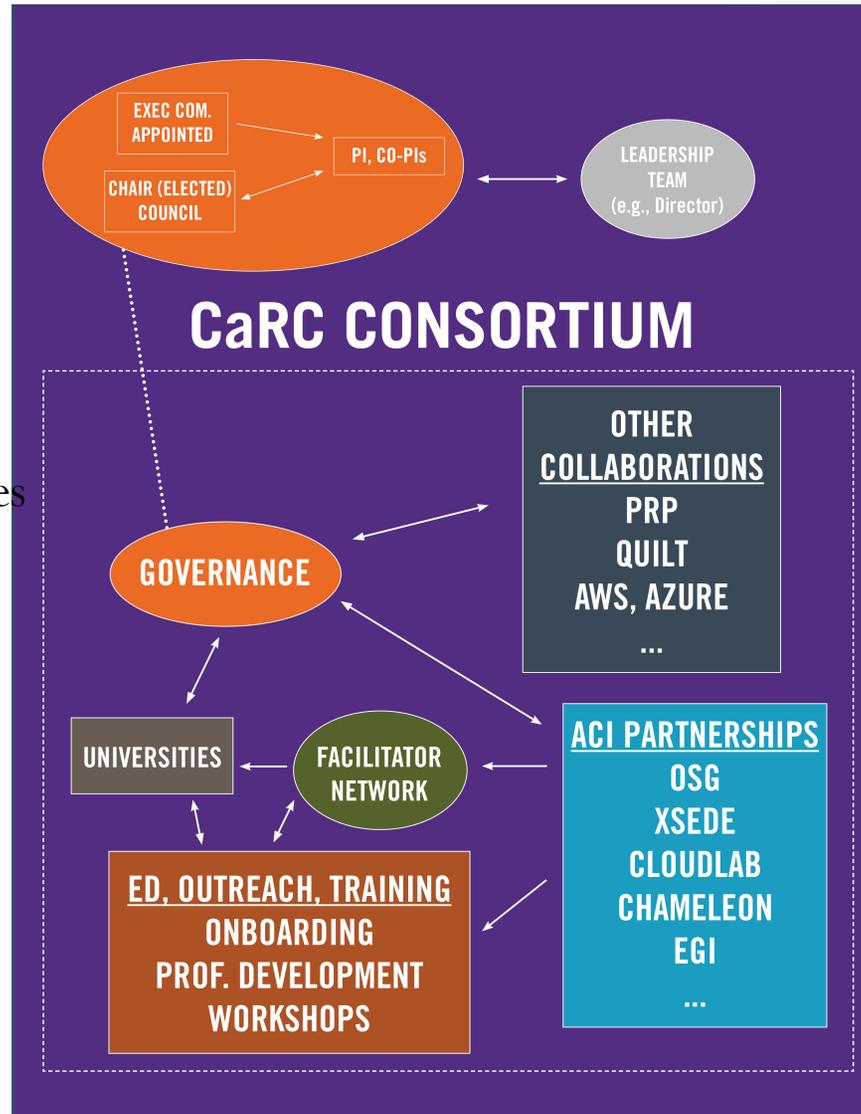
From Project to Consortium



- Consortium seeded by *pending* RCN proposal
- To build and sustain a **national coalition** of research computing support practitioners
- 28 universities, contribute to participate



Consortium Governance



Thoughts re: ACI-REF vs. GENI

	ACI-REF	GENI
Purpose	Bridging ACI to broadened constituents	Enabling large scale advanced testbed for distributed system researchers
User-facing emphasis	Help researchers navigate & use complex computing resources	Help researchers navigate & use bleeding edge technologies
Inter-campus coordination	Sharing support best practice, network of diverse expertise	Engineering of campus hosted infrastructure, central coordination
Campus motive & investment	Value of ACI support network, campus paid user-facing IT staff	GENI researcher needs, in-kind investment so far
Transforming science	Computing in more disciplines & enhanced productivity	Computing innovation in more disciplines & enhanced testbed access
Sustainment	NSF seed, university membership next, others?	NSF seed, then?
Partnership	Government, research projects, industry	Government, research projects, industry?

QUESTIONS?