CASC | Coalition for Academic Scientific Computation

Teaming Up for Impact: Regional Collaborations for Research Computing and Data

Background

Research is not a solo endeavor. Scientists, engineers, and educators understand that people working together results in more impactful research and education, and that high quality research and educational programs benefit from a rich support system. Research Computing and Data (RCD) collaborations are an essential element of that system. RCD centers are hubs of activity at research universities, government labs, and business research and development centers, providing infrastructure, resources, and expertise to support high-end computing, visualization, artificial intelligence, data analysis, simulation, and modeling. Because they are likely to be found at universities that conduct grant-funded research and offer doctoral programs, RCD centers often serve as learning labs for students in computational and data science disciplines. But in today's world of competitive research, RCD centers, like the researchers they serve, need collaborators to win major grants and secure new funding streams. Collaborations mean new sources of funds and new opportunities for innovative research, data storage, software tools, and applications, as well as training and consulting services. By bringing people with a wide range of skills and viewpoints together collaborations make meeting the requirements of government funding opportunities more feasible. For example, a single institution might not easily meet the broader impact requirements of National Science Foundation (NSF) proposals, which look at a grant program's benefits to the broader community and society. Henry Neeman, executive director of research

Contact Us kelley@casc.org | (202) 670-5798 | casc.org

computing at the University of Oklahoma and one of the leaders of OneOklahoma Cyberinfrastructure Initiative (OneOCII) explained it this way: “The advantage of OneOCII is that it costs almost nothing but immediately addresses the broader impacts requirements of NSF proposals.”

Not all collaborations are cost-free to participants, but many allow smaller institutions to compete for research dollars and contribute to the democratization of research infrastructure by enabling them to use some of the world’s best cyberinfrastructure (CI) resources. “There is a huge divide in computing resources and infrastructure between well-established R1 institutions and smaller schools,” said Tabitha Samuel, interim director at the University of Tennessee’s National Institute for Computational Sciences and co-lead of a collaboration formed at Texas A&M University called Building Research Innovation at Community Colleges (BRICCs)-Pathways. “BRICCs and similar projects help narrow the divide.”

The intent of this paper is to share the benefits of RCD collaborations, particularly at the regional level. It shares the experiences of some collaborations and highlights the common elements people need to consider to make a regional collaboration work. By examining the current landscape, it offers guidelines to help people at RCD centers plan, secure funding, and implement regional collaborations.

Below, we delineate categories of collaborations by their major stated functions, while realizing that collaborations often serve multiple purposes and don’t always fit neatly into one category.

Research cyberinfrastructure partnerships. These collaborations typically involve research universities, government laboratories, and sometimes, smaller colleges and universities, funding organizations, government agencies, and vendors. They often form as a way for members to get more – more resources, more ideas, more advanced software and technologies – augmenting their institutional investments or economically sharing RCD CI. These collaborations enable members to come together to compete for grants that they might not have considered on their own. Examples include the Midwest Research Computing and Data Consortium (MW-RCD) [1], the Massachusetts Green High Performance Computing Center (MGHPCC) [2], and at the national level, NSF’s Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) [3].

Regional research and education networks. These collaborations generally form to support the advanced networking and research CI needs of members. They provide the high-performance, low-friction networking needed for advanced research – often more than a single institution can offer. They can serve as on-ramps to a variety of services available through members or the cloud, such as data security and storage, computing resources, and scientific software. Examples include OARnet in Ohio [4] and KeystoneREN in Pennsylvania [5]. In turn, the regional collaboratives often participate in national communities, such as Internet2 [6] and The Quilt [7].

Science-driven collaboratives. These collaborations form when scientists from different institutions form partnerships to further research in a specific domain or related domains. The goals are to make discoveries easier for scientists through shared CI [8], data services, and enhanced

capacity when applying for grants. Examples include the Mid-Atlantic Quantum Alliance, which seeks to build the region's leadership in quantum technologies and research [9], and the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) [10], which provides software, data sets and other resources for the water science community.

Opportunity and Entrepreneurial Initiatives. These collaborations, both regional and national, form to engage new stakeholders, such as community colleges and emerging research institutions (ERIs) in using CI. Opportunity-based collaborations contribute to workforce development, spark entrepreneurial ideas, and expand access to top-tier CI research tools. Examples include Building Research Innovation at Community Colleges (BRICCs), a national collaboration that has launched several regional cyberinfrastructure partnerships [11] and the National Science Foundation's Big Data Innovation Hubs, which formed to advance science and business innovation using big data and data science [12].

Challenges to Successful Collaborations

Multi-institutional and multidisciplinary collaborations come with their own set of challenges, including geographic and institutional barriers, coordinating costs and administrative services, uneven contributions among partners, and potential conflicts about co-authorship credit on papers [13].

How can a new regional RCD collaboration overcome these challenges? Every situation is different, but all collaborations require clear and consistent communication with community members. "It's because we listen to the community and we strive to develop programs to provide them with the value they seek from us," said Florence Hudson, executive director of the Northeast Big Data Hub, when reflecting on the Hub's success. Launched as one of four big data collaborations in 2015, the Northeast Big Data Hub has expanded to include more than 20,715 individuals representing 1,966 institutions in all 50 states and 69 countries as of March 2025.

While listening is a good launchpad for collaborations, other challenges are likely.

Insufficient administrative support. Technical specialists and researchers who want to collaborate will need at least tacit support from managers and administrators. Buy-in is particularly necessary if the collaboration's plans for growth will involve decisions about strategic investments in new organizational infrastructure [14]. Administrators must understand the need for the collaboration, the value it will bring to the organization, and be presented with its broad scope, rather than a more myopic view that addresses only their organization's role.

Insufficient time. With many RCD collaborations, especially when they are

new or involve institutions without large research budgets, the staff involved are not paid for their efforts. They have full-time jobs, and adding the responsibilities of a new collaboration could result in burnout. “Our largest challenge is one of time,” said Leslie Jones, former director of program development for the Minority Serving Cyberinfrastructure Consortium (MS-CC). “Participants in the MS-CC community are typically wearing multiple hats at their institutions.”

Community/organizational differences. Every organization has its own system for managing payments across state lines and to multiple vendors. If the collaboration involves public universities, state law regulates how these systems work. Successful collaborations manage to navigate the different requirements and expectations of members, manage grants that span departments and institutions, and through it all, remain respectful and understanding of long-standing differences in approach.

Communication over distance. Regional collaborations can cover relatively small distances, such as the Mid Atlantic HPC User Group in Washington, DC and nearby states, or they can cover vast distances, such as the Rocky Mountain Advanced Computing Consortium (RMACC), which spans nine states across the American West. Even collaborations within a small geographic area must deal with varied work schedules and endless conference calls. While Zoom is an essential collaboration tool, it does not deliver the value of in-person gatherings. “Our biggest challenges come from being so widely dispersed without a lot of in-person connections,” said Becky Yeager of the RMACC. The RMACC addresses this challenge the way most collaborations do: regular video meetings and two in-person meetings each year.

Insufficient funding/sustainability. For some regional RCD collaborations, acquiring funding might be as simple as charging a membership fee. Other collaborations are seeded with funding from federal agencies, but long-term funding is not always part of that plan. For example, the four Regional Big Data Innovation Hubs received NSF funding for four years. To survive beyond 2020, Hubs needed to secure additional resources. Hudson said the Northeast Data Hub has been able to thrive by fostering relationships among members, demonstrating the value of the Hub to new members and funders, and developing a sustainability plan, which emphasizes finding new partnerships, including industry partnerships, creating measurable value, and maintaining an engaged community.

Loss of expertise and/or champions. Collaborations develop in ways that complement the skills of the experts in leadership roles. Many times, collaborations recruit researchers who benefit from being “champions,” willing to evangelize about the collaboration and its value. These kinds of stakeholders help collaborations succeed; word-of-mouth praise from a colleague often is the most effective way to show value. But what happens when a key expert leaves their job, or when a key champion retires? Losses can leave a vacuum; a sustainability plan can help

collaborations prepare for turnover and develop methods for attracting new talent.

Opportunities/Actions: Seeding Successful Collaborations

The challenges are many, but so are the rewards of successful regional RCD collaborations. For one, garnering funding becomes a much less daunting prospect with partners to share the load of writing, submitting, and administering grants. In addition, funding strengthens a collaboration, giving it resources to work with and visibility that helps build a reputation. “We have received multiple NSF grants to create shared regional supercomputers that can be used by any RMACC member across the region,” said the University of Colorado’s Yeager, of the RMACC. “This has helped support smaller institutions who do not have their own on-prem resources and has in turn helped drive regional research projects and other grants.”

Tabitha Samuel of the University of Tennessee reports that the Tennessee Research and Education Computing Collaborative initiative has led to jointly published research papers, reports, and position papers, as well as awareness of new grant opportunities and the chance to submit grant proposals with partners. “It’s opened eyes for economic and entrepreneurial partnerships regionally and nationally,” she said. The economic impact of such collaborations is exemplified by the success of South Big Data Innovation Hub that turned an initial investment of approximately $10 million from the NSF, which included funding for the South Hub and Spoke projects, into a total research inflow of $141M over its 10-year project duration. As Renata Rawlings-Goss, its executive director, explained: “For every dollar given to the Hub, we increased it by 14 times, either in economic research dollars brought in from other sources or to our partners in our region, or economic development.” Collaborations also create opportunities that no one institution could pursue alone. For example, as part of its mission to provide access to data and data science collaborations, the Northeast Big Data Hub created a Student Data Corps, which now includes more than 12,000 people around the globe.

From Idea to Successful Collaboration

How does an idea become a full-fledged regional RCD collaboration? The pathways to success are as varied as the collaborations themselves, but the following

tips could make the journey less stressful and the collaboration itself more successful.

Recognize Your Priorities. Regional collaborations work best if partners agree on their top priorities. While some collaborations focus on sharing research infrastructure to enable science, others focus on particular types of scientific discovery, or exist as frameworks for educational opportunities, workforce development, or entrepreneurial endeavors. Before putting pen to paper (or, more accurately, fingers to keyboard), know why you're forming a collaboration and have your key goals in mind.

Get Your Plan and People in Place. Before you look for funding or develop a catchy name, it is essential to gather the champions in your collaboration and make sure everyone agrees on your mission, vision, and initial plans for development and deployment. Write everything down, develop a mission statement and a few key goals. Think about actions – how the collaboration will achieve its goals. Put together a leadership team and a clear plan that makes the case for your collaboration and helps everyone explain its value to others.

Win Administrative Support. Administrative support and a general understanding of the value of your collaboration by administrators is essential. With a plan and champions in place, explaining your value to organizational leaders should be simple, and that might be all the buy-in you need. For collaborations that require expenditures or administrative approval, you might need to work a little harder to garner support. Put yourself in the administrator's shoes. What do they care about? If they're worried about expenditures, demonstrate how the collaboration will be cost effective. If they strive for a higher research profile, show how the collaboration enables new and/or breakthrough research. If possible, get others to speak to how the collaboration will help them, such as researchers or students.

Right Size Your Mission. Too many ideas – even if they're good ones – can stretch the collaboration too thin and result in staff burnout. Hudson, of the Northeast Big Data Hub, said the Hub had "eight or nine" focus areas when she took the role as director in 2020. One of her first actions was to cut the focus areas to a more manageable four. A "right sized" mission and plan of action means you have goals that are achievable given the available resources. Find the low-hanging fruit, achieve a few early successes, and promote them to start building your reputation.

Develop a Sustainability Model. Not every regional RCD collaboration needs external funding, but most need at least some source of support to get started and to grow. Funding can be as simple as membership fees or as all-encompassing as multiple funding awards and agreements from federal agencies and vendors. Regardless of how you start, external funding can help grow and sustain a collaboration. Sustainability also means stable leadership and governance, and having ongoing processes to keep focused on priorities and keep members engaged.

Come Together. In a survey of regional RCD collaborations conducted by CASC,

almost all respondents mentioned the value of human interaction in collaborations. While professional organizations offer annual meetings, regional collaborations can bring people together on a smaller scale, often making it feasible for members to drive to workshops and annual meetings. For younger researchers, technologists and students, in-person meetings provide the chance to learn from leaders in their fields and develop relationships with colleagues.

Keep the Lines of Communication Open. Transparency and regular communication are critical. Share information about events and opportunities for funding and professional development. Give members a chance to brag about their accomplishments in a newsletter. Involve a wide range of collaboration participants in planning at least one annual in-person meeting that reinforces your shared vision, addresses challenges, and encourages sharing successes.

As a nationwide organization that advocates for research computing and data centers, CASC can facilitate new regional RCD collaborations by providing opportunities to present new collaborations, explain their value, find sponsors and new members, and share successes.

Key Takeaways

Collaborations bring opportunity. RCD center collaborations can help organizations manage costs and gain access to computing and data resources not available locally. Regional RCD collaborations provide opportunities for collaborative funding proposals, professional development, and new published research results.

Collaborations bring challenges. Despite their value, regional RCD collaborations come with some potential roadblocks.They require getting buy-in from organizational leaders, carving out time to manage the collaboration, breaking down silos, and understanding the differences in organizations that might operate on different schedules, handle payments and expenses differently, and see the value of the collaboration in a new light.

Collaborations need buy-in. Although not all collaborations need active involvement of organizational administrators, management should be informed of the collaboration and should support its overall goals. Determine how much organizational support your collaboration will need from the start – whether it will be ongoing financial support or simply a “thumbs up.”

Collaborations require focus. A focused collaboration is more likely to succeed than one that tries to be everything to all members. Put a plan together that includes a “right sized” mission, scope, and goals. Keep your mission and scope tightly focused when starting out (you can always expand as you grow) and keep your goals attainable. A couple of “wins” on narrowly focused goals will start to build your reputation.

Collaborations thrive on communication and listening. Refocus and recalibrate your regional RCD collaboration as necessary. The needs of your community will change, as will some of the personnel and organizations involved. Be flexible, be transparent, communicate often, and always listen to what the community is saying.

Collaborations are about people. All collaborations share the understanding that people working together results in more impactful research and education. Successful collaborations build communities that break down organizational and domain boundaries.

Work with National Organizations. Many national organizations provide opportunities for networking and idea sharing that often lead to new regional RCD collaborations or strengthen existing ones. Examples include our own organization, the Coalition of Academic Scientific Computation(CASC), EDUCAUSE, the Minority Serving Cyberinfrastructure Consortium (MS-CC), the United States Research Software Engineer Association (US-RSE), and the Campus Research Computing Consortium (CaRCC). CASC is committed to providing a forum where regional RCD collaborations can share successes, lessons learned, and best practices. The Appendix lists contact information for some of these organizations.

CASC gratefully acknowledges the organizations that participated in our survey or interviews for this paper:

Building Research Innovation at Community Colleges
Ecosystem for Research Networking
KeystoneREN
Mid Atlantic HPC User Group
Mid-Atlantic Quantum Alliance
Midwest Research Computing & Data Consortium
Minority Serving-Cyberinfrastructure Consortium
Northeast Big Data Innovation Hub
OneOklahoma Cyberinfrastructure Initiative
Rocky Mountain Advanced Computing Consortium
South Big Data Innovation Hub
Tennessee Research and Education Computing Collaborative
U.S. Research Software Engineer Association

Authors: *Karen Green; Clark Gaylord, George Washington University; Tabitha Samuel, University of Tennessee, Knoxville; Dhruv Chakravorty, Texas A&M University; Richard Gerber, Lawrence Berkeley National Laboratory; Alisa Kang, Georgetown University; Ralph Zottola, University of Alabama at Birmingham; Sandra Gesing, The United States Research Software*

Engineer Association; Dan Majchrzak, University of Oregon; Bill Barnett, University of Massachusetts; Semir Sarajlic, Vanderbilt University; Stratos Efstathiadis, New York University; Erik Deumens, University of Florida; Carolyn Casler, CASC; Kathryn Kelley, CASC

About CASC

The Coalition for Academic Scientific Computation is an educational nonprofit 501(c)(3) organization with 105+ member institutions representing many of the nation's most forward-thinking universities and computing centers. CASC is dedicated to advocating for the use of the most advanced computing technology to accelerate scientific discovery for national competitiveness, global security, and economic success, as well as develop a diverse and well-prepared 21st century workforce. Learn more at http://casc.org.

References

[1] Midwest Research Computing and Data Consortium (MW-RCD). https://midwestresearchcomputing.org/.
[2] Massachusetts Green High Performance Computing Center (MGHPCC). https://www.mghpcc.org/.
[3] Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS). https://access-ci.org/.
[4] OARnet. https://www.oar.net/.
[5] KeystoneREN. https://www.keystoneren.org/.
[6] Internet2. https://internet2.edu/.
[7] The Quilt.https://www.thequilt.net/.
[8] Shah, D., Belshe, R., Speyer, G. et al. Measuring the Impact of Centralized High-Performance Computing with Research Collaboration Networks. SN COMPUT. SCI. 5, 600 (2024). https://doi.org/10.1007/s42979-024-02890-6
[9] Mid-Atlantic Quantum Alliance. https://mqa.umd.edu/.
[10] Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUASHI). https://www.cuahsi.org/.
[11] Building Research Innovation at Community Colleges (BRICCs). https://hprc.tamu.edu/briccs/.
[12] National Science Foundation Big Data Regional Innovation Hubs. https://www.nsf.gov/funding/opportunities/bd-hubs-big-data-regional-innovation-hubs/505185 /nsf18-598.

DOI 10.13140/RG.2.2.22935.07846

[13] D. Ravasi, J. Zhu, W. Wan, S. Dorobantu, and M. Gruber, (2024). What Makes Research Collaborations Successful? Advice from AMJ Authors. AMJ, 67, 583–594, https://doi.org/10.5465/amj.2024.4003.
[14] S. Broude Geva, et al. (2020).. Fostering Collaboration Among Organizations in the Research Computing and Data Ecosystem. In Practice and Experience in Advanced Research Computing (PEARC '20), July 26–30, 2020, Portland, OR, USA. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3311790.3396645.

Appendix

US national organizations that can be resources for regional RCD collaborations: a short, noncomprehensive list

Campus Research Computing Consortium (CaRCC): https://carcc.org/

Coalition for the Advancement of Scientific Computation (CASC): https://casc.org/.

EDUCAUSE: https://www.educause.edu/.

Minority Serving Cyberinfrastructure Consortium (MS-CC): https://ms-cc.org/

United States Research Software Engineers (US-RSE): https://us-rse.org/