ACCESS: Advancing Innovation

NSF's Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support

Timothy J. Boerner
National Center for Supercomputing Applications at the University of Illinois
Urbana, Illinois, USA
tboerner@illinois.edu

Stephen Deems
Pittsburgh Supercomputing Center
Pittsburgh, Pennsylvania, USA
deems@psc.edu

Thomas R. Furlani
Roswell Park Comprehensive Cancer Center and University of Buffalo
Buffalo, New York, USA
furlani@buffalo.edu

Shelley L. Knuth
University of Colorado Boulder
Boulder, Colorado, USA
Shelley.Knuth@colorado.edu

John Towns
National Center for Supercomputing Applications at the University of Illinois
Urbana, Illinois, USA
jtowns@illinois.edu

ABSTRACT

As the National Science Foundation evolves its investments in cyberinfrastructure, it has made a significant investment in the ACCESS (Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support) program instantiating a novel set of services along with a novel governance and management model. Research cyberinfrastructure (CI) is a key catalyst for discovery and innovation and plays a critical role in ensuring U.S. leadership in science and engineering, economic competitiveness, and national security, consistent with NSF's mission. Funding of a set of awards through the ACCESS program has established a suite of CI coordination services targeted at supporting a broad and diverse set of requirements, researchers, and usage modalities spanning all areas of science and engineering research and education complemented by support for the collective and coordinated operation of the overall ACCESS program.

CCS CONCEPTS

• **Computer systems organization** → **Architectures**; • **Information systems** → **Computing platforms**.

KEYWORDS

NSF ACCESS, cyberinfrastructure ecosystems, Research Computing, Federation

ACM Reference Format:
Timothy J. Boerner, Stephen Deems, Thomas R. Furlani, Shelley L. Knuth, and John Towns. 2023. ACCESS: Advancing Innovation: NSF's Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support. In *Practice and Experience in Advanced Research Computing (PEARC '23), July 23–27, 2023, Portland, OR, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3569951.3597559

PEARC '23, July 23–27, 2023, Portland, OR, USA

ACM ISBN 978-1-4503-9985-2/23/07.
https://doi.org/10.1145/3569951.3597559

1 INTRODUCTION

The use of digital technology and research-enabling services is critical to the advancement of scholarly endeavors across nearly all areas of study wherein researchers rely on increasingly diverse and distributed resources. The National Science Foundation (NSF) has for decades funded the coordination and integration of these advanced technologies, services, and professionals—collectively known as cyberinfrastructure (CI)—to expand the horizons of scientific achievement and human understanding. Building on identified needs of science and engineering research and education and, with considerable input from the community via requests for information, invited presentations, community workshops, and experience gained through decades of investment, the NSF issued its "Blueprint" for an agile, integrated, robust, trustworthy and sustainable CI ecosystem.[1] The ACCESS program solicitations reflect and reference this Blueprint, calling for a fundamental shift in how the CI ecosystem is provisioned, supported, and governed. The NSF's ACCESS program builds upon the successes of the 11-year XSEDE [10] project, while also expanding the ecosystem with capabilities for new modes of research and further democratizing participation. The awards made under this program have established a suite of CI services targeted at supporting a broad and diverse set of requirements, researchers and usage modalities complemented by support for the collective and coordinated operation of the overall ACCESS program. Below we provide an overview of this service ecosystem aligned with the programmatic areas of ACCESS discussing Allocations, Support, Operations, Metrics and Monitoring, and the Coordination Office in turn in the following sections.

2 ALLOCATIONS

The ACCESS Allocations Service envisions the NSF-funded CI resources for the nation's science and engineering research and education community being accessible and equitably available for all researchers no matter the size of the institution, the scale of the planned work, the discipline of the research, or the demographics of the requestor. To achieve this vision, the ACCESS Allocations team is creating and continuously improving upon an open, inviting, and

democratized allocations marketplace. The marketplace provides a welcoming gateway for participation across disciplines, institutions, and demographic groups. Diversity, Equity, and Inclusion (DEI) considerations are central to every aspect of the project to incorporate stakeholder views and to ensure that all modifications to the allocations processes contribute to the team's DEI initiatives. Feedback from the community, collected through multiple mediums, combined with effective change management processes are essential to improving the allocations process for current and future investigators while ensuring that key stakeholders are informed throughout the process.

Within the allocations marketplace, the goal of having an efficient, scalable, and simplified request and review framework guides the policies, practices, and review and assessment components. Supporting an evolving portfolio of resources and usage modes, standardized allocation project types allow for multiple points of entry to accommodate a variety of needs ranging from small-scale allocations to support coursework, code development, and exploration, to operating a science gateway or pursuing multi-grant large-scale research initiatives. By making many allocation awards in ACCESS Credits, a universal currency later exchanged for resource-specific units, researchers are empowered to explore the many resource options to identify the systems best suited for their science and engineering objectives. ACCESS Credits allow Resource Providers (RPs) to set their exchange rates for resources in the marketplace. RPs are integral to the allocations process by providing their expertise on requests for their resources.

The Allocations team is building its marketplace and request and review framework on top of a robust, decentralized, and flexible software platform. The eXtensible Resource Allocation Service (XRAS) [5] serves as a foundation for innovation and scalability in the ACCESS Allocations Service. XRAS provides a full suite of interfaces and backend components to manage an allocations process with panel management, central reporting and accounting functions, and distinct interfaces for proposal submission, review, and administration. XRAS is also a software-as-a-service platform, currently used by multiple providers to support independent allocations processes. Through an agile development process, combined with feedback received through Continuous Improvement and DEI [2] engagements, the team extends and enhances XRAS to improve allocations process management, user profile management, publications reporting and tracking, database management, and five proposed Innovative Pilots.

These Innovative Pilots are designed to pursue transformative advances to the allocations marketplace and disrupt the notion of the ACCESS ecosystem being a compute-centered environment. The Variable Marketplace pilot will allow RPs to modify the exchange rates for their resources in response to factors such as supply and demand, time-bounded "sales", or to encourage computing during times where energy costs are lower. Campus On-Ramps will embed the ACCESS allocations interfaces into site-local infrastructure and provide local campus researchers with a virtual datacenter of national-scale CI resources. Innovative pilots targeting Commercial Cloud, CI Resource Workflows, and Sensor Networks and Instruments will also be pursued to fully leverage the flexible practices and policies to support a diversified resource portfolio beyond traditional high-performance computing systems.

3 SUPPORT

The mission of the ACCESS Support team is to provide equitable, scalable support to best enable research on NSF funded CI by democratizing access to advanced CI resources.[7] Our goal is to decrease the time to science by removing the complexity of working with advanced CI and allowing researchers to focus on what is important for them – answering their research questions, securing grants, or writing publications. Our approach is through a scalable, tiered model that allows researchers to gain access to these systems quickly and efficiently and without complexity. By leveraging existing tools in the community that are designed to democratize access, such as Open OnDemand [6] and Pegasus [4], allowing users to "self-serve" and get answers to their questions through a Knowledge Base, or providing certain concierge level services that can provide more directed support we aim to reach our goals my reducing the number of hands-on engagements required for researchers. We have also deployed a community engagement program that expands resources that are available in our Knowledge Base and diversifies our content by considering a multitude of perspectives.

4 OPERATIONS

The ACCESS Operations team's mission is to serve as a source of innovation and support across the NSF-funded CI ecosystem while also seeking to evolve, advance, and expand that ecosystem across the areas of CI operations, resource integration, data and networking, cybersecurity, and the CI professional workforce. It seeks to reduce barriers to conducting and moving research projects among the increasingly diverse services and resources needed to undertake a wide range of scientific research. The various components of the ACCESS Operations project, and their contributions to these stated objectives, are summarized below. Throughout all these activities is community engagement and focused coordination efforts with the key stakeholder for ACCESS Operations: the Resource Provider.

The Operational Support area develops and maintains both the foundational architecture that defines the ACCESS program's core CI and the technical instantiation of that CI. The community engagement, support, and documentation activities undertaken by Operational Support make usable and accessible the architectures, servers, applications, and databases they operate to manifest the many services provided by ACCESS Operations to Resource Providers and other service areas of the ACCESS program. The Operational Support area comprises three key activities throughout which key enterprise services are operated for the program: Integration Roadmaps, Concierge Integration Experts, and an Information-Sharing Platform.

The Network and Data Transfer Services area advances the practice of research networking and data transfer across the NSF-funded CI ecosystem through provisioning of services that support resource providers and by seeking to establish a community of practice around data transfer within the expanding ACCESS ecosystem. Best-known practices will be shared and supported broadly as the team works across the key areas of architecture and service operations, metrics and monitoring, and community engagement.

In a context where domestic and international cybersecurity threats both evolve continuously and appear suddenly, a robust

and comprehensive portfolio of cybersecurity activities is the most critical component for the ACCESS ecosystem when seeking to maintain the integrity and reliability of its services and resources. The ACCESS Operations Cybersecurity Support area is the center of gravity for ACCESS program service areas and Resource Providers when it comes to ensuring a safe, secure, and trustworthy CI ecosystem. The extensive range of activities for this team include governance and oversight, risk management, cybersecurity operations, intelligence sharing, and authentication strategies.

Recognizing the value of a diverse and well-prepared workforce, the ACCESS Operations team runs a tiered training program that focuses on high-impact experiences that train and promote diversity in the next generation workforce. The Student Training and Engagement Program is designed around three levels of progressively in-depth engagement with CI professionals that provide training and mentorship. These levels include a two-week, in-person training program for 15 students; a full-time, in-person summer program that embeds students with CI professionals; and a part-time, remote fall semester program that sees the student working on a larger project with a remote team.

The ACCESS Operations team's focus on essential technical elements that are flexible and adaptive, ensures that the ACCESS program can be agile and responsive to the needs of the growing and diversifying NSF-funded cyberinfrastructure ecosystem that it was created to support.

5 METRICS AND MONITORING

Recognizing the importance of monitoring and measurement services to ensuring a healthy and robust national CI system, the NSF Office of Advanced Cyberinfrastructure has invested in a portfolio of monitoring and measurement development projects over the years, including CI usage and performance metrics through XDMoD [8] [3] and networking-based monitoring and measurement via NetSage (https://www.netsage.global/) [11], perfSONAR (https://www.perfsonar.net/) and CAIDA (https://www.caida.org). The ACCESS Monitoring and Measurement Services (MMS) award builds on this solid foundation to improve the ability of the national CI ecosystem as well as campus-based CI resources to meet the increasing demands of the scientific and engineering communities and provide greater insight into CI use, behavior, and performance to help inform future investments.

ACCESS MMS supports the comprehensive management of the NSF innovative CI resources, as well as for CI systems in general. It does so primarily through the XDMoD and Open XDMoD tools which track operational, performance, and usage data for NSF supported CI resources and local, campus-based CI resources, respectively. ACCESS MMS is designed to meet the following objectives: (1) provide the user community with a tool to more effectively and efficiently use their allocations and optimize their use of CI resources, (2) provide operational staff with the ability to monitor, diagnose, and tune CI system performance as well as audit the performance of all jobs and applications running on their system, (3) provide application developers with the ability to obtain detailed analysis of application performance to aid in optimizing performance, (4) provide stakeholders with a diagnostic tool to facilitate CI planning and analysis, and (5) provide metrics to help measure scientific impact and return on investment of advanced cyberinfrastructure.

In addition to ongoing development of ACCESS XDMoD and Open XDMoD, substantial new capability will be developed under ACCESS MMS. This includes the introduction of a data analytic framework built on Jupyter Notebooks to access the rich and extensive repository of ACCESS XDMoD usage and performance data. Stakeholders will no longer be tethered by XDMoD's built-in analytic capabilities, but rather will be able to leverage their own analytical skills and a wide range of existing data science tools within a popular, widely deployed, and extensible framework. Automated workload analysis functionality will also be incorporated into XDMoD and Open XDMoD to inform future upgrades as well as provide a baseline to measure the impact of system updates and modifications.

A CI simulator is under development that can be used to predict the response of the CI ecosystem to new systems or changes in the operational parameters of existing systems, as well as provide users of ACCESS resources, especially new users, with recommendations on the optimal ACCESS resource to conduct their research.

Under ACCESS MMS we are also developing automated tools with XDMoD to monitor the power consumption of scientific applications running on established and novel CI architectures (such as ARM) and couple this data with the time to solution on each system to provide decision makers with quantifiable data from which to inform deployment of new systems, as well as provide tools for software developers to improve application performance [9]. These are only a few of the many new features under development as part of the ACCESS MMS award.

6 ACCESS COORDINATION OFFICE (ACO)

The National Science Foundation's (NSF) Blueprint for an agile, integrated, robust, trustworthy, and sustainable cyberinfrastructure ecosystem has called for a fundamental shift in how the CI ecosystem is governed, supported, and provisioned. The ACCESS Coordination Office (ACO) has been established to facilitate governance and coordination, partnering with service teams to provide reliable and flexible services to the CI community. The ACO has four primary functions: to establish and facilitate governance, cultivate expert guidance, facilitate and coordinate engagement with the community, and facilitate and coordinate evaluation efforts across the program.

The Executive Council (EC) is the primary management and governance body of ACCESS, facilitated by the ACO, and the formal mechanism by which Service Teams coordinate service delivery and present a coherent CI ecosystem to the community. The EC operates under guiding principles that aim to keep processes and policy lightweight, minimize complexity, build consensus, encourage collaboration, use a data-driven approach, and maintain focus on programmatic innovations.

The External Advisory Board (EAB) serves as a trusted source of advice for the EC, providing expert advice spanning the CI ecosystem, and the science and engineering research and education community. The EAB is a resource to the EC that can help the EC look ahead and provide anticipatory advice, and is selected based on experience and knowledge in areas relevant to ACCESS.

The ACO serves as a primary interface to the broader community, providing a focal point for the flow of public-facing information between the ACCESS program and the science and engineering research and education community. The ACO coordinates with service teams to project a coherent ecosystem of resources and services to the community, and works to engage both established and emerging CI communities, with a strong focus on engaging traditionally underrepresented segments of the community. The ACO further coordinates community-building efforts across the program to expand participation to new and underserved communities, and deepen participation in existing communities. The ACO provides coordination and information sharing across service teams to leverage contributions from one another, remain coherent, and avoid "blind spots" in their approaches. Finally, the ACO facilitates evaluation efforts across the program, working with service teams evaluators to coordinate, strengthen, and maximize the shared use and dissemination of evaluation findings. The ACO works to summarize and aggregate service team-specific evaluation plans into a program-wide evaluation matrix, identifying commonalities, complementarity, redundancies, and gaps in program-wide evaluation efforts, and leading to the development of common metrics that can be used across service teams to reflect compound effectiveness and impact.

Overall, the ACCESS Coordination Office plays a crucial role in coordinating and governing the ACCESS program, facilitating collaboration among service teams, engaging the CI community, and evaluating the program's effectiveness.

REFERENCES

[1] 2019. Transforming Science Through Cyberinfrastructure: NSF's Blueprint for a National Cyberinfrastructure Ecosystem for Science and Engineering in the 21st Century. https://www.nsf.gov/cise/oac/vision/blueprint-2019/

[2] Agbeli Ameko, Stephen Deems, Dave Hart, Laura T. Herriott, and AJ Lauer. 2023. *ACCESS Allocation Services DEI Plan: Advancing an open, inviting, and democratized allocations marketplace for the national research cyberinfrastructure ecosystem.* Report. Carnegie Mellon University. https://doi.org/10.1184/R1/21824850.v1

[3] James C. Browne, Robert L. DeLeon, Abani K. Patra, William L. Barth, John Hammond, Matthew D. Jones, Thomas R. Furlani, Barry I. Schneider, Steven M. Gallo, Amin Ghadersohi, Ryan J. Gentner, Jeffrey T. Palmer, Nikolay Simakov, Martins Innus, Andrew E. Bruno, Joseph P. White, Cynthia D. Cornelius, Thomas Yearke, Kyle Marcus, Gregor von Laszewski, and Fugang Wang. 2014. Comprehensive, open-source resource usage measurement and analysis for HPC systems. *Concurrency and Computation: Practice and Experience* 26, 13 (2014), 2191–2209. https://doi.org/10.1002/cpe.3245

[4] Ewa Deelman, Karan Vahi, Mats Rynge, Rajiv Mayani, Rafael Ferreira da Silva, George Papadimitriou, and Miron Livny. 2019. The Evolution of the Pegasus Workflow Management Software. *Computing in Science and Engineering* 21, 4 (2019), 22–36. https://doi.org/10.1109/mcse.2019.2919690

[5] David L. Hart, Amy Schuele, Ester Soriano, Maytal Dahan, and Matthew Hanlon. 2014. XRAS. , 8 pages. https://doi.org/10.1145/2616498.2616562

[6] Dave Hudak, Doug Johnson, Alan Chalker, Jeremy Nicklas, Eric Franz, Trey Dockendorf, and Brian McMichael. 2018. Open OnDemand: A web-based client portal for HPC centers. *Journal of Open Source Software* 3, 25 (2018). https://doi.org/10.21105/joss.00622

[7] Shelley L. Knuth, Julie Ma, Joel C. Adams, Alan Chalker, Ewa Deelman, Layla Freeborn, Vikram Gazula, John Goodhue, James Griffioen, David Hudak, Andrew Pasquale, Dylan Perkins, Alana Romanella, and Mats Rynge. 2022. The Multi-Tier Assistance, Training, and Computational Help (MATCH) Project, a Track 2 NSF ACCESS Initiative. *The Journal of Computational Science Education* 13, 2 (2022), 17–20. https://doi.org/10.22369/issn.2153-4136/13/2/4

[8] Jeffrey T. Palmer, Steven M. Gallo, Thomas R. Furlani, Matthew D. Jones, Robert L. DeLeon, Joseph P. White, Nikolay Simakov, Abani K. Patra, Jeanette Sperhac, Thomas Yearke, Ryan Rathsam, Martins Innus, Cynthia D. Cornelius, James C. Browne, William L. Barth, and Richard T. Evans. 2015. Open XDMoD: A Tool for the Comprehensive Management of High-Performance Computing Resources. *Computing in Science and Engineering* 17, 4 (2015), 52–62. https://doi.org/10.1109/mcse.2015.68

[9] Nikolay A. Simakov, Robert L. Deleon, Joseph P. White, Matthew D. Jones, Thomas R. Furlani, Eva Siegmann, and Robert J. Harrison. 2023. Are we ready for broader adoption of ARM in the HPC community: Performance and Energy Efficiency Analysis of Benchmarks and Applications Executed on High-End ARM Systems. , 78-86 pages. https://doi.org/10.1145/3581576.3581618

[10] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson, R. Roskies, J. R. Scott, and N. Wilkins-Diehr. 2014. XSEDE: Accelerating Scientific Discovery. *Computing in Science and Engineering* 16, 5 (2014), 62–74. https://doi.org/10.1109/mcse.2014.80

[11] Katrina Turner, Mahesh Khanal, Tyson Seto-Mook, Alberto Gonzalez, Jason Leigh, Andrew Lake, Sartaj Singh Baveha, Samir Faci, Brian Tierney, Daniel Doyle, Lisa Ensman, Jennifer M. Schopf, Douglas Southworth, and Edward Balas. 2020. The NetSage Measurement Framework: Design, Development, and Discoveries. , 45-56 pages. https://doi.org/10.1109/indis51933.2020.00010

Received 21 April 2023; accepted 12 May 2023