NIST RESEARCH COMPUTING ENVIRONMENT BENCHMARKING STUDY REPORT

Kevin M. Gunn
Dr. Christine Harvey
Kanan Shah
Megan Rinard

April 2024
MITRE Product: MP240268

SOLVING PROBLEMS FOR A SAFER WORLD®

This publication was produced as part of contract 1333ND23FNB180289P24003 with the National Institute of Standards and Technology. The contents of this publication do not necessarily reflect the views or policies of the National Institute of Standards and Technology or the US Government.

Executive Summary

The Research Computing Environment Benchmarking Study Report presents an evaluation of the research computing environment (RCE) at the National Institute of Standards and Technology (NIST) and compares it with similar institutions. The study focuses on five key components of research computing: High Performance Computing (HPC), Scientific Software Portfolio (SSP), Research Data Management Services (RDMS), Research Computing Support Staffing (RCSS), and Networking Services.

Overall, NIST compares well to the other study participants. It outperforms the institutions in this study in RDMS and RCSS. There are some opportunities for improvement, however, especially with respect to its HPC infrastructure and its provisioning of scientific software.

Key findings include:

1. **High Performance Computing (HPC)**: NIST operates multiple HPC systems, many of which are specific to certain research groups. NIST's computational capability is reported as lower than all participants, even when scaled by the size and impact of the research program. NIST's HPC spending is lower than that reported by other study participants, with more spent on labor but significantly less on hardware, software, and capital investments.
2. **Scientific Software Portfolio (SSP)**: NIST provides a larger variety of centrally managed scientific software to researchers as part of its RCE than the other participants. The practice common to all participants is that the majority of scientific software is provisioned either on a fee-for-service or a bring-your-own-software model.
3. **Research Data Management Services (RDMS)**: Although data in this area is limited to what the participants could provide, NIST is employing research data management best practices as they apply to federally funded research. More data from additional research institutions would help NIST better gauge its RDM spending and storage capacity.
4. **Research Computing Support Staffing (RCSS)**: NIST invests much more than other participants in Research Computing Support Staffing (RCSS), both in terms of dollars and people. However, NIST’s relatively high number of support staff is inclusive of staff in NIST’s research laboratories – a perspective apparently not measured by other participants.
5. **Networking Services**: Central funding of the network infrastructure is a norm among the study participants. However, the scarcity of data from the study participants makes it difficult to draw any firm conclusions related to networking.

The report concludes that NIST could benefit from a deeper dive into what comparable institutions characterize as RCE as well as how to gauge whether RCE support levels are appropriate to an institution’s research program.

Table of Contents

Table of Figures

Table of Tables

INTRODUCTION

BACKGROUND

Information technology is crucial to the National Institute of Standards and Technology's (NIST) scientific and technical mission. NIST's advancements in areas such as the Creating Helpful Incentives to Produce Semiconductors (CHIPS) Act of 2022, advanced manufacturing, engineering biology, climate change, quantum, and artificial intelligence depend on high performance computing (HPC) resources and staff expertise. The increasing data generated by NIST's instrumentation and modeling/simulation demands more storage, faster transmission, and in–depth analysis. However, an internal survey of NIST researchers revealed that new hires find NIST's current research computing infrastructure and support inferior to their previous institutions. This perceived limitation impacts NIST's ability to attract top–tier researchers who require robust computational resources. NIST desires its research computing infrastructure to match the quality of the scientific and technical research conducted and to operate at a level comparable to similar institutions.

PREVIOUS STUDIES

INTERNAL NIST SURVEY

In 2020, NIST conducted an internal study to evaluate its research computing environment (RCE) from the perspective of the users of the services: the research staff. NIST had the following key findings from that survey: [1]

- Research computing investment needs to be increased.
- Network infrastructure needs to be improved.
- Compute capacity needs to be improved.
- Research data management capabilities need to be improved.
- Software offerings need to be improved.
- Support for effective utilization of research computing resources needs to be available.

EPOC REPORT

One outcome of the survey was NIST's engagement of the Engagement and Performance Operations Center (EPOC) to conduct a "deep dive" to investigate NIST's RCE requirements. EPOC delivered its report to NIST in 2023. "A Deep Dive comprehensively surveys major research stakeholders' plans and processes in order to investigate data management requirements over the next 5–10 years [2]." The key findings of the EPOC report were:

- NIST's organizational approach to research information technology (IT) support is at best federated and at worst disjoint or duplicative.

- NIST storage solutions are a mixture of different technologies and approaches.
- NIST does not have a coherent, organization–wide HPC strategy that provides a low–barrier–to–entry resource for researchers.
- Data mobility in or out of NIST is felt to be the responsibility of the user, and some users can waste time trying to solve technical problems on their own without reaching out for assistance.
- There is a general lack of technical staff that can serve in a "research IT coordination" role that spans NIST.
- Network challenges negatively impact research projects.

STUDY PURPOSE

NIST commissioned this report to present NIST with the information it needs to make its research computing infrastructure match the quality of the scientific and technical research it conducts and make it commensurate with the RCE resources available to researchers at comparable institutions.

STUDY APPROACH

MITRE evaluated the research computing environments at institutions comparable to NIST, including NIST itself. The selected key components of research computing integral to the research mission included: (1) HPC, (2) scientific software portfolio (SSP), (3) research data management services (RDMS), (4) research computing support staffing (RCSS), and (5) networking services.

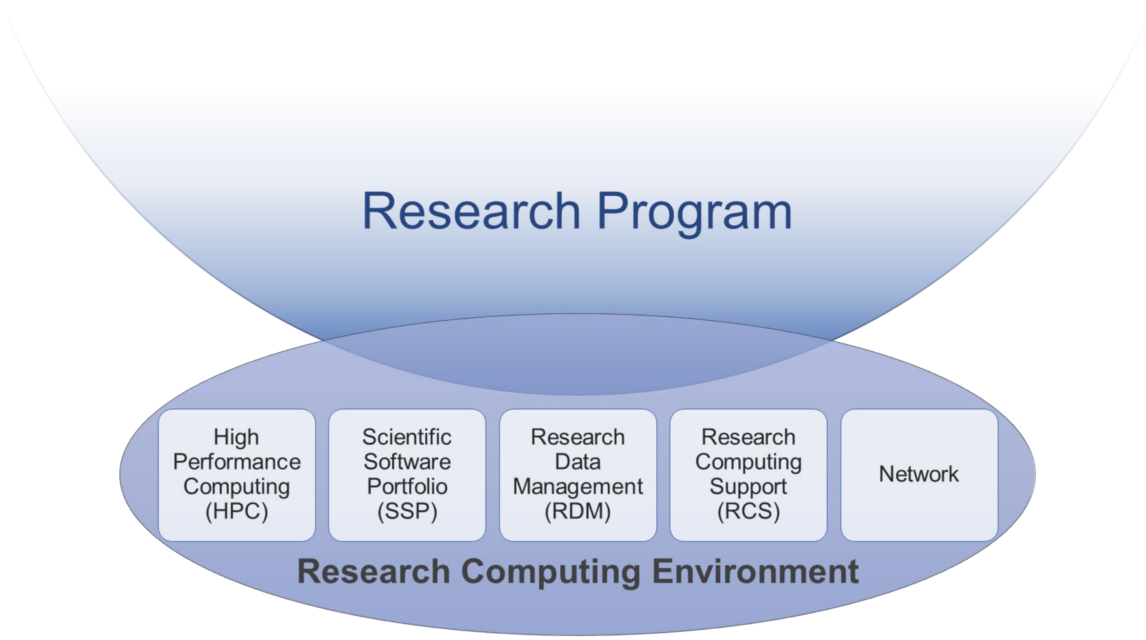

Figure 1. Study Scope: Research Program and Research Computing (Source: MITRE)

MITRE also analyzed the research performed by each participant in the 2013—2022 timeframe (i.e., the "study period") to develop a basis for comparison for the RCE at each organization.[1] For this study, MITRE used the relationships depicted in Figure 1 to distinguish the research program, research computing environment, and the components of the RCE. RCE also contains some elements of research, indicated by the overlap shown in the figure. Generally, in terms of budgets and staffing, the scale of the research conducted is orders of magnitude greater than that of the supporting RCE.

STUDY PARTICIPANTS

To help identify potential improvements, NIST Research Services Office (RSO) selected the candidate research institutions for comparison (i.e., the "study participants") from both the federal government and academia. NIST based the selection on similarity to NIST in terms of research breadth, budget, and user community. MITRE, along with NIST, engaged with six candidates.

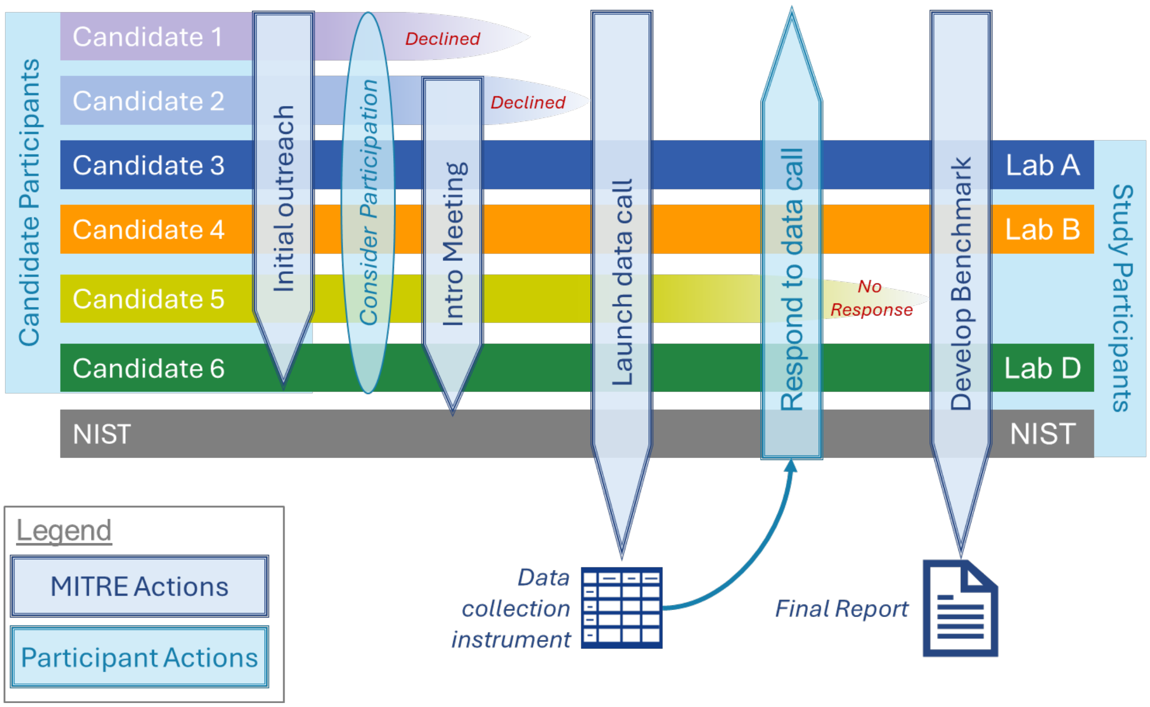

Figure 2. Study Approach and Participants (Source: MITRE)

As illustrated in Figure 2, two candidate organizations declined to participate in the study. Five agreed (including NIST). Among the five initial participants, four provided inputs to the study in the form of a response to a questionnaire that MITRE developed to gather data. One (Lab A) provided responses only to the HPC section of the study questionnaire. MITRE conducted

[1] Throughout this report, the terms "participant", "organization", and "institution" are used interchangeably.

independent research on the participants' research program scope and impact, as well as the budget and staffing numbers for Lab A.[2]

To maintain confidentiality and candor, this report masks the identities of the participants. This report names NIST specifically, but anonymizes the other participants as "Lab A," "Lab B," and "Lab D," respectively.[3] ***MITRE acknowledges the small sample size, so caveats all findings and benchmarks as directional, not decisive.***

Throughout this report, ***findings are highlighted in bold, blue, italicized font***.

[2] Lab A is a government agency with public budget and staffing data available online.
[3] A fifth participant ("Lab C") did not provide the data needed to be included with this report.

FINDINGS: RESEARCH PROGRAMS

This section summarizes the characterization of the research program that the RCE supports. This characterization is provided in terms of the scope (research areas), size (budget, staffing), and impact (publications, patents, etc.). These measures serve as the basis for comparison of the RCEs between the study participants.

RESEARCH SCOPE: SUBJECT AREAS

Typically, institutions that study similar topics will have similar demands for research computing. Scope also helps make the case that the compared institutions are similar. MITRE analyzed the subject areas each of the participants published in during the study period. The details are in Appendix C(Research Scope Analysis). MITRE found, with small variations, that each participant's most-published subject areas indicated a strong subject area correlation between the organizations.

This strong correlation implies ***the demand for RCE resources for each participant will be roughly proportional to the scale and impact of its research*** as discussed in the next two sub–sections**:** Research Scale: Budget and Staffing**,** and Research Impacts**.**

RESEARCH SCALE: BUDGET AND STAFFING

In this study, MITRE related the participants by two measures of its size: the total annual research budget (for fiscal year 2023), and the number of researchers (i.e., "users" of RCE resources). The analysis details are in Appendix C(Research Scale Analysis).

Researchers, who represent the highest demand for RCE resources, comprise over 75 percent of all research staff across the four organizations, and 92 percent at NIST. Figure 3 summarizes the total research budgets and staff for the four study participants. Of note are the significantly higher research budgets at Lab A and Lab D, while NIST has a proportionally larger number of researchers. This naturally leads to NIST's spending significantly less *per Researcher* (about $300,000 vs. Lab A's $3,100,000).

MITRE assumes the number of researchers—and therefore the number of potential users of RCE resources—will have a greater and more direct impact on the RCE needs than the research budgets alone.

The participants reported four types of research staff: researchers, management, support staff, and "other." In particular, the support staff category was defined as "e.g., administration, custodial, security staff" in the data collection questionnaire. Figure 4 summarizes those responses (Lab A did not provide staffing details). As the chart shows, only 5.0 percent of NIST's research organization staff are in the support or "other" category, while Lab B's and Lab D's staff in those categories is two- to three-times higher. This may be significant when discussing research computing support staffing in a later section.

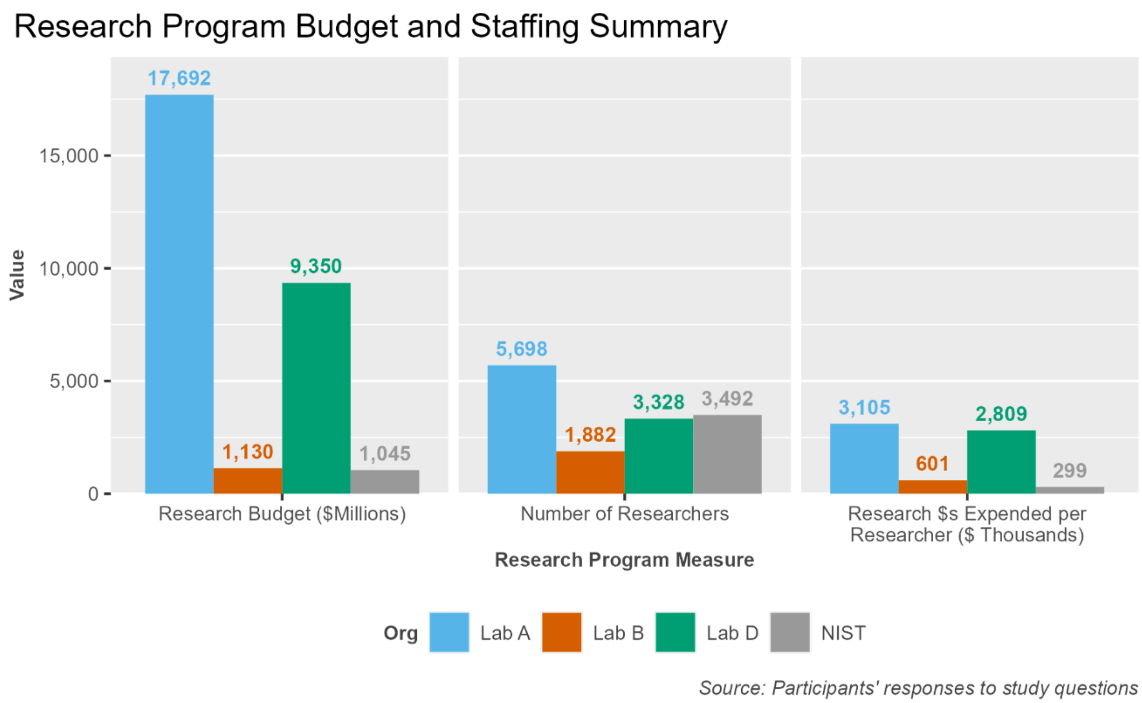

Figure 3. Participants' Research Program Budget and Staffing Summary (Source: MITRE)

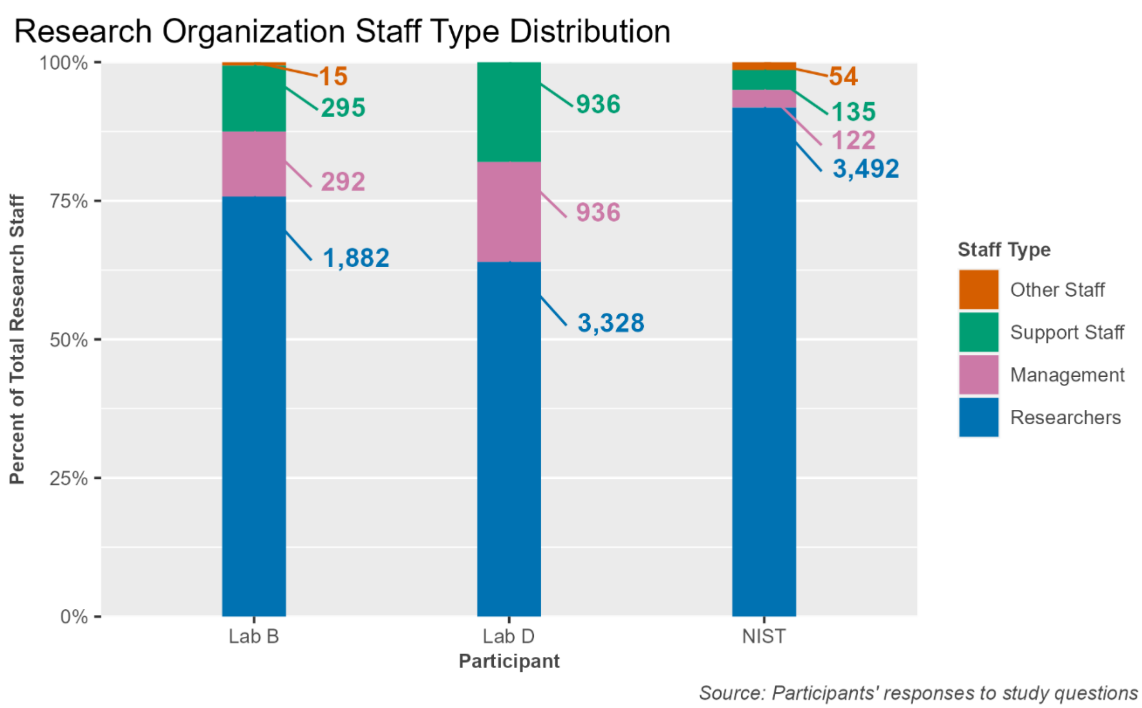

Figure 4. Research Organization Staff Type Distribution (Source: MITRE)

We do not include the non-researcher staff in our benchmarking for three reasons: (1) Researchers comprise the bulk of the staff in the research organizations, (2) non-researchers do not have as much of a direct logical relationship to research computing resource requirements, and (3) the study participants may have had different interpretations of staff categories proposed.

RESEARCH IMPACTS

In addition to the scope of an organization's research, MITRE assumes an organization's research *impact* is positively correlated with the amount of RCE needed to support the research. Note that the *causation* could go either way. That is, a more available and capable RCE can increase the impact of the organization's research. Conversely, more impactful research may drive more investment in the RCE by the organization. MITRE could not arrive at a definitive conclusion from the data collected. Each organization, when applying these benchmarks, should make that determination for itself. *In this study, MITRE assumes only a correlation between the impact of the research and scale of the RCE supporting it.* The research impact analysis is detailed in Appendix C.

PUBLICATIONS AND CITATIONS

To calculate a common measure of research impact—publications—MITRE retrieved publication data for all study participants from *Scopus*, an abstract and citation database launched by Elsevier in 2004.[4] The data used in this study only includes peer–reviewed documents: research articles, review articles, conference proceedings, data papers, and book chapters. If one or more authors of a document were from a participating organization, that document is counted as one document. If two or more participants collaborated on the document, it is counted for both organizations. Details are in Appendix C, page C-4.

While the number of publications is a measure of the *output* of a research program, the *impact* each publication has can be better measured by the number of times an article is cited by other authors. *Where* the document is published (e.g., in which journal or at which conference) is an additional measure of its impact. The *SCImago Journal Rank* (SJR) provides a weighting by the prestige of a journal [3] (see Appendix C, page C-8). Figure 5 illustrates the weighted number of citations[5] per document for each participant over the study period. ***NIST's weighted citations per document is comparable to the other study participants.***

To represent the overall impact of an organization's publications, MITRE calculated the "h–index" for each participant. The h–index combines the number of publications with the number of citations of each to generate a more accurate reflection of an organization's impact. Appendix C(page C-9) explains how the h–index is calculated. Figure 6 shows the results of the calculations for the four study participants. In the figure, "n" equals the number of peer-reviewed documents the organization's authors published during the study period (2013-2022). ***NIST's h–index for the study period is on par with the other participants.***

---

[4] https://www.scopus.com

[5] Calculated by multiplying the number of citations of each article times the SJR score of the journal it was published in, divided by the total number of documents.

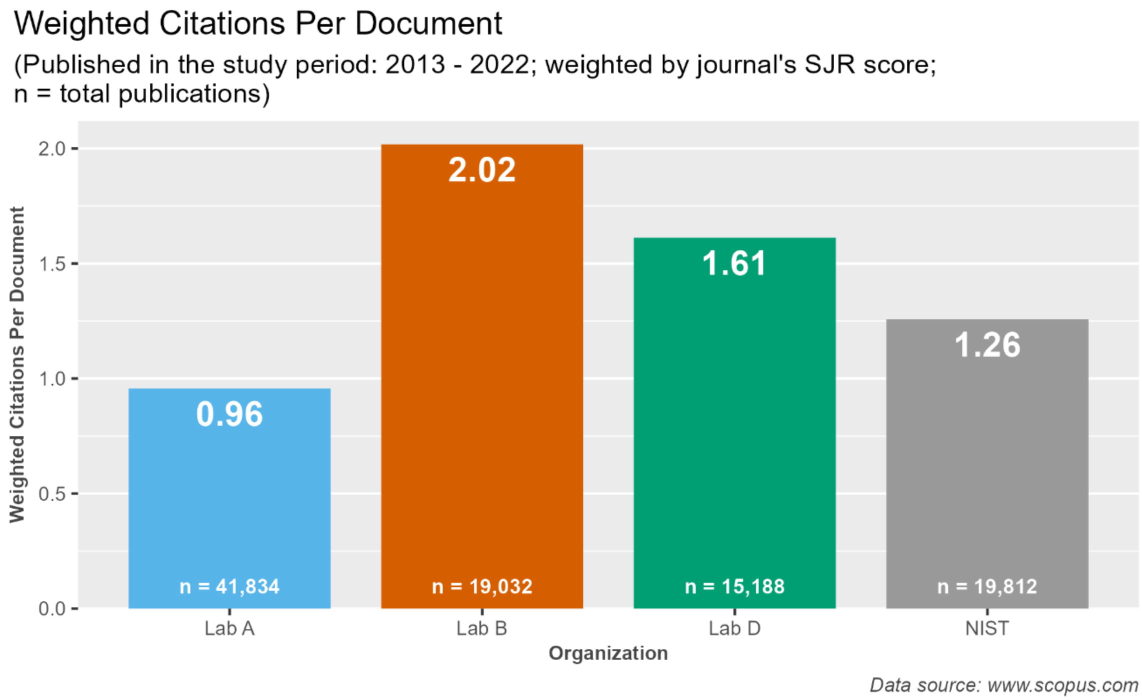

Figure 5. SJR-weighted Citations per Document (Source: MITRE)

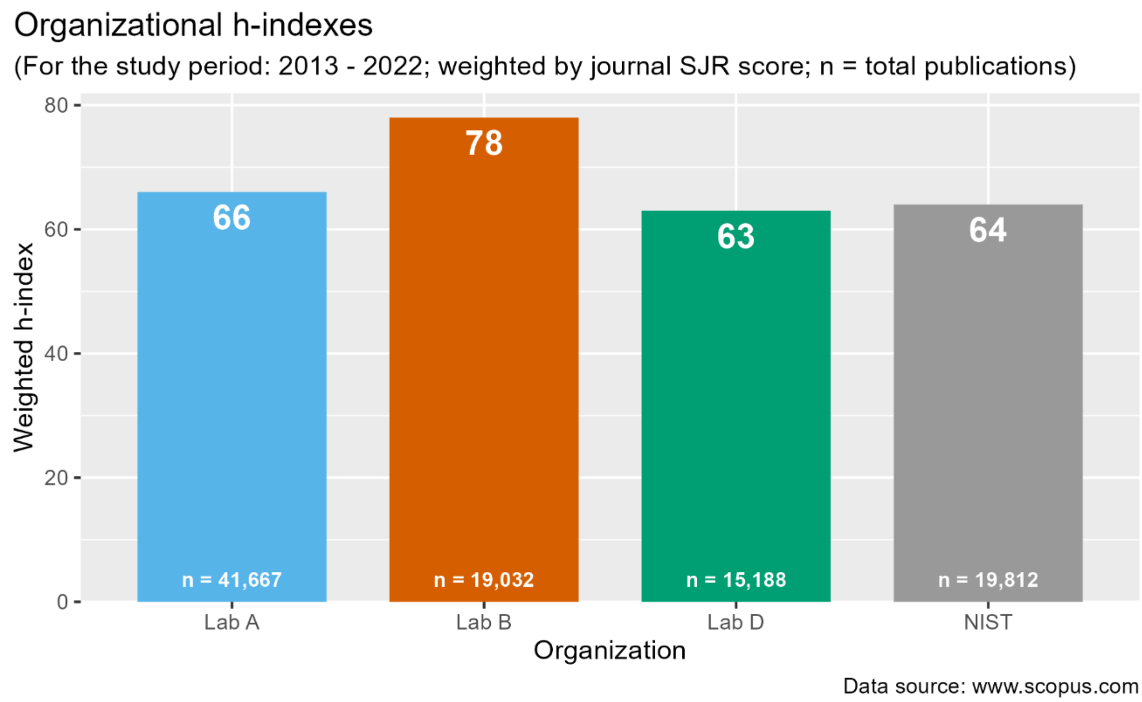

Figure 6. Weighted Organizational h–indexes [4] (Source: MITRE)

RESEARCH PROGRAM MEASURES

To apply the above measures more easily, MITRE computed “relative” measures in each area. Those results are shown in Table A. MITRE selected the four measures based on their intuitive correlation[6] with the level of research computing resources required to support those measures. Conversely, it may be possible to improve the research program measures by increasing the available resources. Until more data is available, MITRE reserves judgement on that relationship.

The scaled values shown in the table are derived from a simple calculation to generate a scaling factor when considering the scope and scale of the RCE components in the next section. Each value was simply divided by NIST’s value for a given measure, thereby setting NIST to one in all cases. This is a technique called *reference scaling* or *relative scaling* [13].

For example, in Figure 3 above, Lab A has 5,698 researchers; NIST has 3,492. So the scaled value for Lab A is 5,698/3,492 = 1.6, and for NIST it is 3,492/3,492 = 1.0. The scaled numbers have the advantage of being unit–less so they can be applied to any measure, regardless of the units of that measure. The overall score in the last row of Table A was calculated by simply taking the average of the three measures. With more data, it would be possible to weigh the average score by the strength of the correlation of each measure. The cells are coded in a “heat map” with the highest scaled values for each row in green and the lowest in yellow.

Table A. Summary of Research Program Factors (scaled to NIST)

Measure	Lab A	Lab B	Lab D	NIST
Research Budget	16.9	1.1	8.9	1.0
Number of Researchers	1.6	0.5	1.0	1.0
Weighted Citations per Document	0.7	1.3	1.2	1.0
Organization h–index	1.0	1.2	1.0	1.0
Overall Research Program “Score”	**5.1**	**1.0**	**3.0**	**1.0**

The overall “score” in the last row of Table A was calculated by simply taking the average of the three measures. With more data, it would be possible to weigh the average score by the strength of the correlation of each measure. The “overall research program score” should be interpreted loosely as the relative scale and impact of the participants’ research program. The scaled research computing environment measures for each organization (next sections), in turn, should be expected to be commensurate with the research programs.

When benchmarking NIST’s RCE against the other study participants, this report will adjust quantitative measures, as appropriate, by these scores. For example, if Lab D has an RDMS budget of $450,000, divide this value by Lab D’s research program score (3.0) to get a scaled budget for Lab D of $150,000. That is, if Lab D’s research program scale and impact was the

[6] The sample size was too small to generate any strong correlations.

same as NIST's, we would expect their RDMS budget to be closer to $150,000, because it supports one-third the research program as NIST's RDMS.

This study's objective is to make a fair comparison between the organizations' respective provisioning of RCE services to researchers. The authors believe this approach will allow for reasonable benchmarking of NIST's services.

FINDINGS: RESEARCH COMPUTING ENVIRONMENT

This section summarizes MITRE's findings with respect to the five components of the study participants' research computing environments: high performance computing, scientific software portfolio, research data management services, research computing support staffing, and networking services.

HIGH PERFORMANCE COMPUTING

In general, high performance computing (HPC) refers to the practice of aggregating computing power in a way that delivers dramatically improved processing speed. It often involves the use of supercomputers and parallel processing techniques. HPC systems are typically employed to solve large problems in science, engineering, or business. These powerful resources are customarily reserved for tasks that require high–speed computations, such as weather forecasting, climate research, quantum mechanics, physical simulations, cryptanalysis, and complex simulations.

PREVIOUS FINDINGS

The NIST Survey [1] found that "[Respondents] felt that improvements to on–site high–performance computing resources would significantly benefit productivity and output." Respondents to the survey also expressed a need for "improvements for system documentation, for system capacity and availability, and for more assistance in optimizing software for the computing resources available locally" [2].

EPOC [2] found that research projects across the NIST would greatly benefit from the use of HPC. However, EPOC identified that NIST's computational support needs to expand in the coming years to meet these needs. At the time of that report, NIST lacked a coherent, organization–wide HPC strategy that provides a low–barrier–to–entry resource for researchers. That lack of strategy had led to reliance on the external supercomputing resources to run complex structural models.

The EPOC report also concluded that the decisions made for IT investments had not always captured the needs of research. It recommended a more transparent process for how technology, including software and hardware, is evaluated by the Office of Information Systems Management (OISM). Finally, the model of relying solely on OISM's overhead–funded support to develop solutions for the Laboratories resulted in gaps that individual Laboratories must address. That approach made it difficult to provide solutions at economies of scale that were affordable to many research projects.

To follow up on these findings, this report provides the following set of findings related to the HPC environments. The four study participants each have different approaches to delivering HPC to research staff. All organizations offer multiple HPC systems, often with varied scope and use cases. This section focuses on general similarities and major differences across the group. All systems discussed in this section are on–premises; the report does not account for projects' use of external resources or of cloud computing services.[7]

CAPACITY

Each organization provided specifications on each of its systems in operation, including the purpose of the system, the audience/user–base, and the number of power, casual, and inactive users. Table B shows the number of each organization's HPC systems and its overall purposes and audiences. Figure 7 shows the total numbers of HPC users across the organizations.

Table B. Research Systems and User Capacity.

Org	Systems	Purpose
Lab A	4	All General Purpose • 1 GPU–Specific
Lab B	6	3 General Purpose • 1 GPU–Specific 3 Specific to Research Group(s)
Lab D	7	4 General Purpose • 1 GPU–Specific • 1 On–Ramp for other System 1 Specific to Climate Modeling 2 Specific to HPC Developers1 GPU–Specific
NIST	11	3 General Purpose • 1 GPU–Specific 8 Specific to Research Group(s)

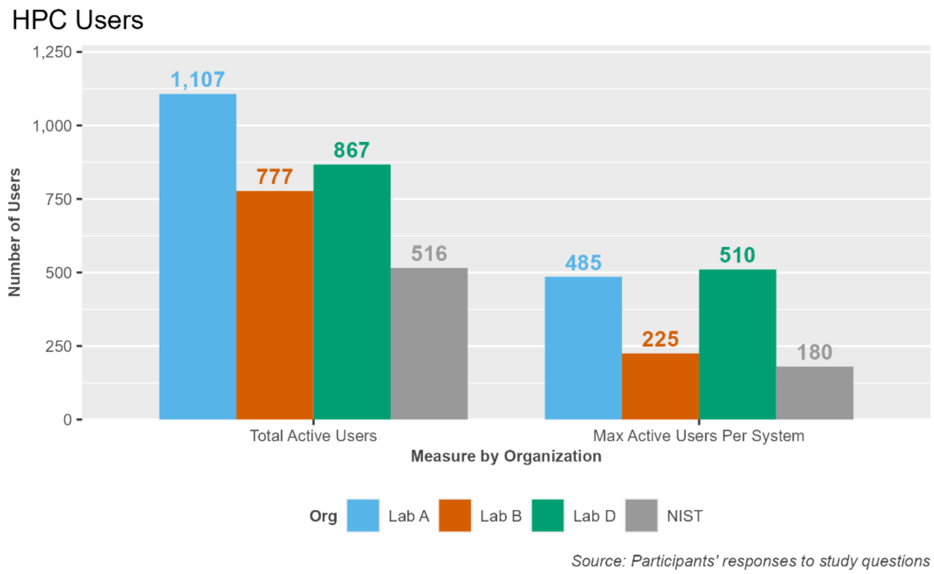

Figure 7. Number of HPC Users (Source: MITRE)

[7] See the discussion in Appendix C , Research Impact Analysis, Collaborations

PURPOSE AND AUDIENCE

All organizations operate multiple HPC systems, although ***NIST reported the largest number of individual HPC systems, most of which are specific to certain research groups.*** Only Lab A reported operating all general-purpose systems within the scope of this study. While Lab B and Lab D both offer systems specific to certain research groups, the quantity of general-purpose systems is similar to the number of research group–specific systems provided. ***Every organization has at least one dedicated general-purpose system focused on GPU computing***.

In Figure 7**,** the total number of active users is defined as the total number of power and casual users across all clusters at the organization. The total numbers do not account for users who may work on multiple systems, meaning that users of more than one system may be accounted for multiple times. NIST has the fewest total users among the participants. Another metric calculated was the biggest system by total users. The total number of power and casual users was calculated for each cluster. The largest system by user count is shown in the “Max Active Users per System” column. Lab A and Lab D have the highest numbers of users on a single system while Lab B and NIST have nearly half the number of maximum users per system.

While *all* Lab A users are on general-purpose systems, the majority (over $^2/_3$) of listed users are on general-purpose systems at Lab B and Lab D. In contrast, ***less than 20 percent of listed users at NIST are on general-purpose systems; the majority are leveraging systems for specific research groups***.

OPERATIONS

Study participants provided information about various operational aspects for the HPC systems they run. Operational data collected included the time frame for system deployment, user access methods, storage, and the top challenges to system stability.

UPDATES AND REPLACEMENTS

Table C shows the two most recent years any system was updated, general details on the replacement system, and strategy and systems for user access across the clusters.

Labs A, B, and D performed system updates or deployed new systems in 2021 and 2023 for the two most recent changes. NIST’s most recent system refresh was in 2022 and was the only participant with hardware over 10 years old. NIST has two systems in place that were originally deployed in 2005 and updated more recently. Lab A’s model is to add new hardware to clusters as the yearly budget allows, often rotating which of the four systems receives upgraded hardware. Older hardware is removed from the system in partial retirement as new servers are added to the clusters. Lab B fully retires systems and funds replacements. At Lab B, the general-purpose systems report a replacement schedule over five years, while the systems owned by specific research groups do not provide a replacement plan. Lab D’s CPU–based general-purpose system was replaced after four years, with the old system remaining online. The replacement plan for the general-purpose GPU system is unknown, and only one of the special–purpose systems is scheduled for replacement after seven years of life.

Table C. Summary of HPC System Replacement Strategies

Org	Most Recent Updates	Replacement	Average System Age
Lab A	2021, 2023	Adds hardware to existing systems periodically as budget allows	6.7
Lab B	2021, 2023	Systems reach end–of–life and are replaced as needed	3.2
Lab D	2021, 2023	Systems reach end of life and are replaced as needed	3.6
NIST	2021, 2022	Two systems have an annual refresh plan, one has a five-year replacement plan, and others have no reported plan	7.0

The process for adding additional hardware across NIST varies greatly by system. Only two of the eleven systems have annual refresh plans to keep hardware updated. The rest of the systems have no explicit refresh plans and only one system reported an expected system lifespan.

SYSTEM RESOURCE MANAGEMENT

High–performance computing environments require a resource management system to allocate computing resources. The resource management system is the user's gateway to access compute resources after using SSH to initially gain access to the system. Resource managers schedule jobs to run on compute nodes and manage the available hardware. They may also handle fault tolerance and recovery, ensuring the system continues to function efficiently even when there are hardware or software failures. The goal is to maximize the performance and efficiency of the HPC system, allowing for complex computations and data–intensive tasks to be completed quickly and accurately.

Nearly all participants use Slurm[8] as a resource manager. NIST has one system currently using the Maui Cluster Scheduler,[9] which has plans to transition to Slurm. Lab A exclusively uses Altair's PBS Professional[10] as the scheduling system, and Lab B has one system leveraging PBS Pro. With the exception of Lab B, each organization is currently using or plans to use a consistent scheduler across the organization.

STORAGE

Table D summarizes the storage capacity for each participant. Each organization and even each cluster within an organization have differing approaches to storage. Lab B provides the most straightforward storage system, in which all clusters use IBM's Storage Scale11 parallel file system and local solid–state drives for storage. Default limits are the same across all systems. At Lab D, all user home directories are stored on a Qumulo12 appliance and individual systems use

[8] https://slurm.schedmd.com/documentation.html
[9] https://sourceforge.net/projects/mauischeduler/
[10] "Portable batch system"; https://altair.com/pbs–professional
[11] https://www.ibm.com/docs/en/storage–scale/
[12] https://qumulo.com/

one or more storage systems from the following types: Qumulo, BeeGFS,[13] VAST,[14] and Lustre.[15] Scratch storage is global on some systems, while others use local scratch space on individual nodes. Base storage is consistent across all systems. Storage strategies across NIST are less cohesive than the other participants, which use network file system (NFS), BeeGFS, ZFS,[16] and Lustre across the organization with each system having stand–alone dedicated storage.

Table D. HPC System Storage Summary

Measure	Lab A	Lab B	Lab D	NIST
Base Storage	2TB	100GB Home/1TB Project	400GB	Varies from 120GB per user to 200TB, some have no quotas
Code and Scratch Storage	Not Provided	IBM Storage Scale/IBM Storage Scale and Local Storage	Home and Project NFS/Mix of Global and Local Scratch Per Node	User home directories, Gitlab (uncommon), /toolbox directory, Varied scratch storage
Total Storage	Not Provided	12PB Globally Shared	~7PB Total Across Multiple Systems	~2.7PB Across Multiple Systems

CHALLENGES TO SYSTEM STABILITY

Across Lab A, the top challenge for the general purpose HPC systems is using shared file systems. For the GPU–specific system, Lab A's top challenge is having multiple users per node. Lab B did not list any top challenges to system stability. Lab D provided a variety of system challenges including old file systems, small files with artificial intelligence and machine learning workloads, system age, and difficulties with specific hardware. Across NIST's eleven systems, the following top challenges were provided: age of compute and storage hardware, lack of support for hardware and operating system, software compatibility on the IBM architecture, heterogeneous hardware, and high power consumption. Heterogeneous hardware as a concern was listed across three systems.

The highly heterogeneous nature of many of the NIST systems would be expected to cause additional overhead for systems-facing staff working on the clusters and to cause challenges for users working with less portable codes. Both Lab D and NIST cited difficulties with working the IBM Power architecture, which would cause additional effort for users of HPC systems. ***System age for the compute and storage hardware was a top stability concern for both Lab D and NIST across multiple systems. Aging hardware may impact the ability of researchers to make innovative developments and will challenge systems-facing staff to keep the clusters and storage systems operational.***

[13] https://www.beegfs.io/c/
[14] https://vastdata.com/
[15] https://www.lustre.org/
[16] https://openzfs.org/

REQUESTING SUPPORT

Across all organizations, NIST has the most variable support model. At Lab A, all accounts are allocated through a specific directorate. Lab B uses an online system or users can reach out directly to owning groups. Lab D also uses a single request form to request account access. Processes at NIST for requesting an account on an HPC system include submitting an email request with pro forma approval,[17] emailing the HPC staff, or applying for an account with special approval from a system owner.

Lab A offers a phone support line or an online ticket submission system where support is immediate. Both Labs B and D have an online email ticketing system for support. At Lab B, requests are usually resolved the same day, Lab D typically resolves internal requests within a day; non–staff support requests may take days to weeks to resolve. The technical support assistance process varies across NIST, with options including emailing a ticketing system, emailing the user community, or directly requesting support from staff. ***Despite the multiple avenues to request support, NIST reports the time to a response is short, with most requests being met within a day.***

PERFORMANCE

A common performance measure for HPC systems is Floating–point Operations Per Second (FLOPS). While some organizations have actual data from running benchmarks to calculate FLOPS, others provided estimates. Table E provides a high–level overview of CPU types and memory across all systems at the organization. Figure 8 depicts FLOPS in terms of Tera–FLOPS (TFLOPS)[18] for readability, along with the number of nodes and cores.

Table E. Participants' HPC System CPUs, GPUs, and Memory

Org	CPU Overview	GPUs	Memory
Lab A	Mostly Intel CPUs, some AMD. Clusters are typically 2–4 types of CPUs	236 V100, 514 A100	~58–478GB RAM/Node
Lab B	Clusters are either all Intel, AMD, or Power9. At most two types of CPU per cluster.	48 NVIDIA A100	64–320GB RAM/Node
Lab D	Clusters are all one type of CPU.	200 GPUs from P100–A100, 4 V100	64–256GB RAM/Node
NIST	Clusters are frequently a mix of multiple types of CPUs. Mixing Intel/AMD and multiple types of Intel or AMD CPU.	88 V100, 24 RTX[1] A5000, 3 RTX 8000, RTX 6000, 4 A100, 6 P100	32GB–1TB RAM/Node

[17] That is, approval is typically just a formality. HPC staff check in with the system owner, and requests are almost always approved.

[18] "…a teraflop refers to a processor's capability to calculate one trillion floating–point operations per second."

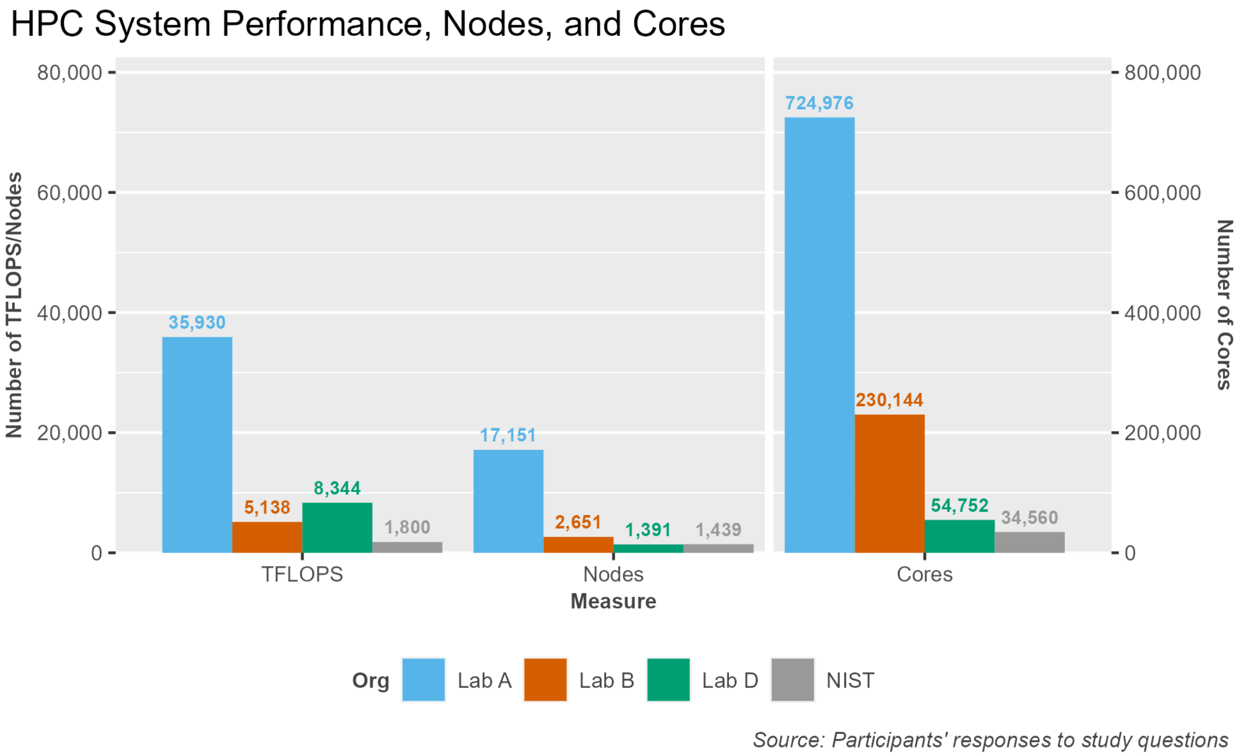

Figure 8. HPC Performance, Nodes, and Cores (Source: MITRE)

COMPUTATIONAL CAPABILITY

Lab A reported the highest computational capability across all organizations with over 35 *Peta*-FLOPS (PFLOPS) of performance (see Figure 8). The systems at Labs B and D are comparable and within a similar range. ***NIST's computational capability is lower than all participants at only 1.8 PFLOPS.***

In looking at the broad, high–level overview of CPUs offered on each system (Table E), the mix of CPU types varies by organization. Lab D has the most homogeneous systems, with nearly all clusters having only a single type of CPU. Lab A and B both have clusters with some mixed hardware; at most, the clusters are a combination of four types of CPUs. ***At NIST, many clusters are a mix of multiple types of CPUs, resulting in a more complex mix of CPUs than any of the participants.*** Three NIST systems have more than six different types of CPU.

All organizations reported GPUs as part of several HPC systems. NIST supports the widest range of GPUs, and any of NIST's listed GPUs are not general-purpose GPUs but are instead from the Nvidia "RTX" lines, specialized for graphics rendering. These types of GPUs are not the recommended hardware for artificial intelligence and machine learning workloads.

The memory for each cluster and individual node types varies on each of the clusters. Labs B and D had similar ranges for memory across the clusters, while Lab A and NIST had wide ranges of memory across systems.

NETWORKING

High performance computing systems heavily rely on networking for data transfer, parallel processing, resource sharing, job scheduling, and system management. A robust internal network infrastructure is critical for overall HPC performance. All organizations reported having at least some clusters with Infiniband networking as the data backplane for systems. Lab A's systems are all Infiniband, and Lab B reported all Infiniband systems except for one using Omnipath, an alternate high–speed backplane used in HPC. Lab D's systems reported a wide range, including Infiniband, HDR Ethernet, Omnipath, 100GB Ethernet, standard Ethernet, and EDR Ethernet. NIST's HPC network backplanes include Infiniband and Ethernet.

BUDGET AND STAFFING

Participants provided information on the processes researchers used for requesting HPC support and the internal staffing to provide support. Lab A did not provide HPC budget figures.

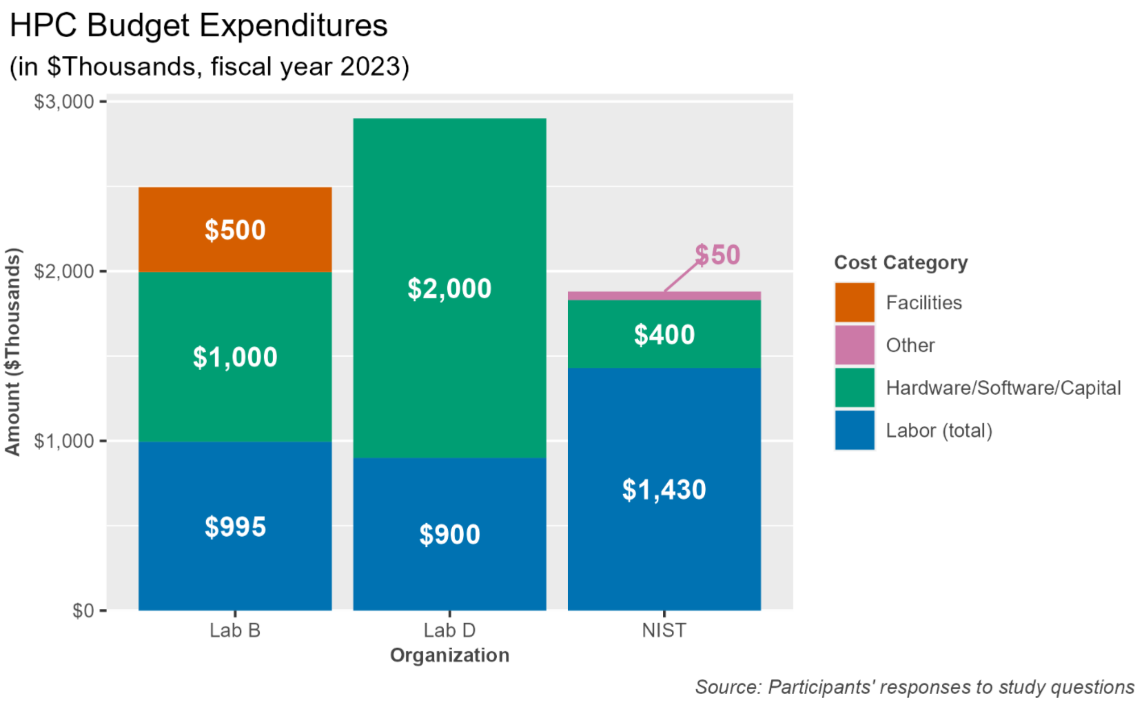

Figure 9. HPC Budget Expenditures (FY2023)[19] *(Source: MITRE)*

HPC BUDGET

The study participants provided the HPC–related budget information shown in Figure 9. As the chart shows, NIST spends more than the others on labor but significantly less than the others on

[19] The green visual indicators in this table span all rows and columns (except the total row) to better see where the higher and lower values are in the data. The total row is visualized on its own.

hardware, software, and capital investments. However, overall NIST's HPC spending is lower than Lab B's and Lab D's.

SUPPORT STAFFING

This study used the Campus Research Computing Consortium (CARCC) "facings" model to report on the types of roles for staff operating research computing systems across the organizations. A brief summary of each of the roles is provided below [5]:

- **Researcher-Facing Roles**. Includes research computing and data staffing, outreach, and advanced support, as well as support in the management of the research lifecycle. Example roles include *Research IT User Support, Research Computing Facilitator, Research Data Consultant.*
- **Data-Facing Roles**. Includes data creation; data discovery and collection; data analysis and visualization; research data curation, storage, backup, preservation, and transfer; and research data policy compliance. Example roles include *Research Data Management specialist, Data Librarian, Data Scientist.*
- **Software-Facing Roles**. Includes software package management; research software development; research software optimization or troubleshooting; workflow engineering; containers and cloud computing; securing access to software; and software associated with physical specimens. Example roles include *Research Software Engineer, Applications Specialist, Data Engineer.*
- **Systems-facing Roles**. Includes infrastructure systems, systems administration and operations, networking engineering, and systems security and compliance. Example roles include: *HPC systems engineer, Storage Engineer, Network Engineer.*
- **Strategy and Policy-Facing Roles**. Includes RCE leadership, institutional alignment, culture for research support, funding, partnerships, and engagement with external communities. Example roles include *Director, Assistant/Associate Director.*

Staffing levels in terms of full time equivalent (FTE) are summarized in Figure 10. Lab A did not provide HPC staffing details. Lab B has the lowest overall staff support for HPC systems, followed by NIST, with Lab D having the highest support staffing levels. Each organization also had different roles as the most demanding on the overall staffing workload. *All organizations had Researcher-Facing Staffing as the highest or second–highest staffing need.* Lab B's staff has the highest concentration in systems-facing and Researcher-Facing roles. Lab D's high levels of staffing are primarily in data and Researcher-Facing roles. NIST's support is primarily in software and Researcher-Facing roles. Within NIST, staffing varied across the owners of the systems as well.

Support funding was difficult to draw conclusions and comparisons from, as the reported labor expenses did not correlate with the FTE reported by the non-NIST participants.

Lab A, Lab B, and NIST all fund labor through various sources of lab overhead, while Lab D draws funding from charge–back of resources. All organizations paid facilities costs from lab overhead budgets.

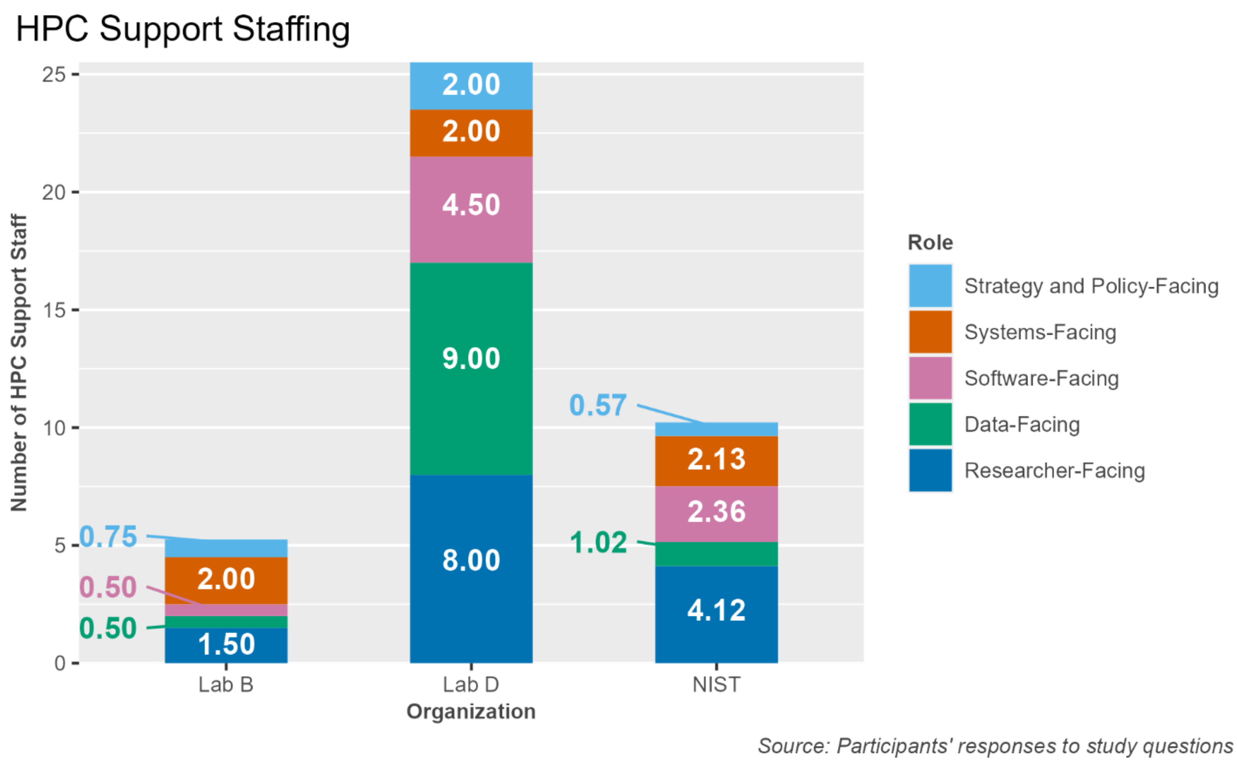

Figure 10. HPC Support Staffing (Source: MITRE)

UTILIZATION

Each organization reported utilization with a variety of metrics. Not all organizations have exact numbers to report for each type of utilization measurement, so broad numbers and descriptors are used to provide comparisons. The five measures reported (summarized in Table F) are: CPU usage, GPU usage, memory usage, oversubscription, and wait time.

Table F. HPC Utilization Summary

Org	CPU Usage	GPU Usage	Memory Usage	Oversubscription	Wait Time
Lab A	67–86%			100%	
Lab B	N/A or 75–85%	50%		"really no oversubscription; we only allocate hours we have; most jobs have no wait time, but larger requests may wait up to a day"	Hours to Days
Lab D	5–100%, several around 80%	5–50%	10–80%, mostly around 50%	Varies: Never, <10%, and one is "constantly oversubscribed"	Hours
NIST	16–100%, many around 20–50%	22–60%	20–125GB	Ranges from rare to several days	Varies from Hours, Days, and Weeks

CPU, GPU, AND MEMORY UTILIZATION

Many systems did not include CPU and/or GPU utilization, therefore the summary table provides a high–level overview of systems that did report utilization. Both Labs A and B reported consistently high utilization across all operational systems, with reported CPU utilization in the 67–86 percent range. ***Lab D and NIST provided wider ranges of utilization across systems with some as low as 5–16 percent utilization and others up to 100%.***

For all organizations with GPU utilization reported, the rates were lower than for CPUs. GPU utilization ranged from 5–50 percent across Lab B, Lab D, and NIST. Labs A and B did not provide memory utilization, and the numbers provided for Lab D and NIST were highly variable.

OVERSUBSCRIPTION

In HPC environments, oversubscription occurs when more resources are in demand than are readily available. While Lab A's systems were reported as constantly oversubscribed, other organizations reported lower or more variable demands across systems. At Lab B, only available hours are allocated for use, which eliminates the ability of users to oversubscribe the systems. Lab D reported some systems as never being oversubscribed, one experiencing oversubscription less than 10 percent of the time, and one system as being constantly oversubscribed. Across NIST's systems, many are rarely oversubscribed and some experience a constant oversubscription.

WAIT TIME

Wait time is the amount of time computational jobs submitted to the HPC clusters spend waiting to run after being submitted to the scheduler. Lab A did not provide details on wait times. Lab D reported the lowest wait time, with most systems having a wait time in the range of hours. Lab B reported more variability, with some systems hosting jobs that take hours to days to run. ***Operations at NIST reported the longest wait times with some systems taking weeks for jobs to begin running while other systems had ranges from hours to days.***

RESEARCH WORKLOADS

Participants provided information about the primary workloads that run on each of its systems, including sharing if any of the workloads dominated utilization on the systems. Understanding the types of workloads and the variety of research run on the systems can provide an understanding of the level of demand placed on support staffing. Table G summarizes the workload utilization responses.

Only Lab A reported a specific domain for each cluster but also did not report any single workload type as dominating the system. The other participants all had varied workloads on each system. Both Lab B and Lab D reported similar workloads on multiple clusters as well as dominant workloads existing on many of the HPC systems. NIST had many systems described as general or having unknown areas of research. While some systems had dominant workloads, many did not, or the dominant workload was unknown.

Table G. Workloads for HPC Clusters

Org	Research Workloads	Dominant Workload
Lab A	Different for each cluster (ocean modeling, aero design, heliophysics)	No dominant workloads
Lab B	Varies per cluster but some repeat cases (Pelegant, Converge, VASP, E3SM)	Yes, often dominant utilization
Lab D	Varies per cluster but some repeat cases (some overlap with B) (MCNP, nwchem, Machine Learning, cp2k, vasp, E3SM, Model Training/Inference, ExaGo)	Mostly dominant utilization
NIST	Lots of general-use systems quantum chemistry, Python, VASP, LAMMPS, custom codes, physics, etc.	Ranges from no, unknown, and yes

COMMUNICATIONS

The section discusses communications to and from the users of the HPC systems at the four participating organizations.

"MARKETING" SERVICES TO RESEARCHERS

Informing researchers about available HPC capabilities is crucial as it allows them to leverage these tools to accelerate research and handle complex computations. It also aids in effective project planning, saves resources by avoiding unnecessary investments, and fosters collaboration and knowledge sharing, leading to a more productive research environment.

Lab A did not provide any data on its marketing activities or how user feedback is collected. Lab B uses web sites, organization–wide announcements, talks, and newsletters to promote HPC services across the organization. At Lab D, new staff are made aware of the research computing services during onboarding. Lab D also has an advocate program where senior staff in the research directorates are responsible for sharing information across organizations. NIST markets the HPC system through word of mouth, documentation pages, and through more recent efforts in HPC training workshops. ***Overall, NIST's approaches to marketing are less holistic and strategic than the study participants.***

COLLECTING USER FEEDBACK

Feedback from users is vital to an HPC system as it helps identify system issues, guides future enhancements, improves usability, and informs strategic decisions about resource allocation and system upgrades. It provides valuable insights into user needs and experiences, ensuring the system effectively meets its goals.

Lab B reported using surveys to gather feedback on the HPC services from users. Lab D uses its advocate program to collect feedback as well as the user ticketing system. NIST has an informal information collection process through management chains. The HPC team also has monthly meetings with primary system users. The Research Computing Advisory Committee (RCAC) has also surveyed staff in the past for its input on existing systems and thoughts on future infrastructure.

HPC BENCHMARKING

Table H (on page 24) summarizes the quantitative measures throughout this section in the form of scaled values based on NIST as the reference value for each measure. In addition, the measures are scaled by the composite research program "score" for each organization as described in Table A on page 9. These numbers form the basis for comparison for the benchmark conclusions in the last section (Summary of RCE).[20]

NIST's RCE capabilities are in line with the other study participants in some areas (e.g., system resource management, base storage, system support). However, the norms among the participants' HPC practices vary from NIST's in a few ways:

- Less than 20 percent of listed users at NIST are on general-purpose systems; the majority are leveraging systems for specific research groups.
- Storage strategies across NIST are less cohesive than the other participants, which use network file system (NFS), BeeGFS, ZFS, and Lustre across the organization with each system having stand–alone dedicated storage.
- Across all organizations, NIST has the most diverse collection of support models.
- NIST's computational capability is reported as far lower than all participants (1.8 PFLOPS). This is low even when factoring the differences between the participants' research program measures (see Table H).
- At NIST, many clusters are a mix of multiple types of CPUs and node architectures, resulting in a more complex mix of CPUs than any of the participants.
- NIST supports the widest range of GPUs, and any of NIST's listed GPUs are not general-purpose GPUs but are instead from the Nvidia "RTX" lines, specialized for graphics rendering.
- NIST spends significantly less than the others on hardware, software, and capital investments.
- Overall NIST's HPC spending is between Lab B's and Lab D's, but much less when taken as a percentage of the overall RCE budget and the amount per researcher.

As the heatmaps in Table H show, ***with the exception of HPC support staffing and spending, NIST appears to be under-provisioning HPC services in most of the quantitative HPC measures.***

[20] The heatmaps in this table represent the distribution of the values in each category (e.g., HPC Storage) and across each **total** row (e.g., Total HPC Support Staffing). They are colored from lowest (yellow) to highest (green). The light gray shaded cells (under Lab A) indicate where no data was submitted for a given measure

Table H. Summary of Scaled HPC Measures (scaled to NIST and research 'score')

Category	Measure	Lab A	Lab B	Lab D	NIST
HPC Storage	Base Storage (TB)	1.0	0.5	0.2	1.0
	Total Storage (PB)	n/a	4.3	0.9	1.0
HPC Performance, Cores, and Nodes	TFLOPS	3.9	2.8	1.5	1.0
	Nodes	2.3	1.8	0.3	1.0
	Cores	4.1	6.4	0.5	1.0
HPC Support Staffing	Researcher-Facing Roles		0.4	0.6	1.0
	Data-Facing Roles		0.5	2.9	1.0
	Software-Facing Roles		0.2	0.6	1.0
	Systems-Facing Roles		0.9	0.3	1.0
	Strategy and Policy-Facing Roles		1.3	1.2	1.0
	Total HPC Support Staffing		**0.5**	**0.8**	**1.0**
	HPC Staff as a percent of RCE Staff		***3.0***	***3.1***	***1.0***
HPC Budget	Labor: Researcher-Facing		0.5	0.0	1.0
	Labor: Data-Facing		0.5	0.2	1.0
	Labor: Software-Facing		0.3	0.3	1.0
	Labor: Systems-Facing		0.9	0.5	1.0
	Labor: Strategy and Policy -Facing		2.2	0.2	1.0
	Labor Costs (total)		**0.7**	**0.2**	**1.0**
	Hardware/Software/Capital Investments		2.4	1.7	1.0
	Other Costs		0.0	0.0	1.0
	Total HPC Costs		**1.3**	**0.5**	**1.0**
	HPC Budget as percent of Total RCE Budget		***7.2***	***2.5***	***1.0***
	HPC $s per Researcher		***2.5***	***1.6***	***1.0***

SCIENTIFIC SOFTWARE PORTFOLIO

A scientific software portfolio refers to a collection of software tools and applications that are used in scientific research and analysis. This can include software for data collection, data analysis, statistical modeling, simulation, visualization, and more.

The portfolio can cover a wide range of scientific disciplines, including physics, chemistry, biology, astronomy, geology, and many others. The software included in the portfolio can be commercial, open–source, or custom–built, and it can run on a variety of platforms, from personal computers to high–performance computing clusters.

PREVIOUS FINDINGS

NIST researchers believe that resource limitations hinder the effective use of current and emerging research technologies and computing architectures. They also see unexploited opportunities for sharing software, solutions, and strategies within the organization [1].

The EPOC study [2] suggests that the research community would benefit from a clearer process for OISM's software evaluation. The study found that software investment decisions don't always reflect research needs. EPOC advises OISM to collaborate with the RCAC and research stakeholders to enhance communication about upcoming software upgrades. They also suggest that OISM and the Associate Director for Laboratory Programs (ADLP) establish a regular update strategy for scientific software, including planning for funding. This would allow both OISM and researchers to better anticipate capability and capacity increases driven by research needs. The RCAC could be the ideal group to tackle these issues.

TYPES OF SCIENTIFIC SOFTWARE

Three of the four participants provided the types of scientific software and software licenses they maintain. For all respondents, most of the SSP consisted of commercial off–the–shelf (COTS) software with annual maintenance cycles. The software portfolios reported are largely focused on modeling and simulation with some lab automation and analysis tools.

Lab A reported that researchers provide its own software, so did not list any particular types. Lab B reported four software packages: Converge, Nek5000, Pelegant, and Vienna Ab initio Simulation Package (VASP). Lab D reported a larger range of packages including Anaconda, Ansys (multiple products), ArcGIS, COMSOL Multiphysics, Dassault Systemes SolidWorks, Gurobi, MathWorks MATLAB, and OmniViz. NIST reported the widest range of scientific software with titles including: Abaqus, a broad Ansys suite, Autodesk Product Design Suite, Intel Composer/OneAPI, L3 Harris Geospatial ENVI/IDL, Maplesoft Maple, MathWorks MATLAB, NI LabVIEW, Ntopology, Scienomics MAPS, Sonnet, and Wolfram Mathematica.

From the information provided, ***NIST provides a broader portfolio of scientific software to researchers as part of its RCE than the other participants.***

FUNDING

Lab A and Lab B both reported that research projects either bring their own software or pay for their own licenses. They have no central funding for scientific software. Lab D and NIST reported that scientific software spending makes up about 7.04 percent and 10.6 percent of total RCE spending, respectively (see Table I). With the exception of some software development management tools, Lab B and Lab D fund researchers' software needs by each project (i.e., “bring your own software”). NIST maintains over 5,000 scientific software licenses at an annual cost of over $4 million in FY2023. However, that number included a non-recurring purchase in FY2023, so that amount is backed out of the “normal” budget for benchmarking purposes.

NIST's scientific software costs are largely covered with a fee-for-service funding model. Fee-for-service project–based SSP spending accounts for the majority of the SSP costs.

Table I. Scientific Software Budget

Org	Budget	percent of RCE Budget	SSP$ per Researcher	Software Funding Model
Lab A	Users bring their own software			
Lab B	Project–provided and funded			
Lab D	$2,490,000	7.04%	$75	Fee-for-service:[21] $2.3M Overhead Funding:[22] $140K
NIST	$3,418,000	13.3%	$978	Fee-for-service

MAINTENANCE & SUPPORT

Nearly all of NIST's scientific software portfolio is under annual maintenance cycles. Other participating institutions report varied maintenance schedules.

SSP BENCHMARKING

While the data do not support any conclusions related to the size or cost of a benchmark SSP, ***the practice common to all participants is that the majority of scientific software is provisioned either on a fee-for-service or a bring-your-own-software model.*** Centrally funded/provisioned software does not appear to be the norm.

[21] In a "fee-for-service" model, we assume that software is centrally purchased and maintained by the research computing support organization and provided as a service to researchers by charging a fee to their project to recover the costs of the software.

[22] In the "overhead funding" model, software is centrally funded, managed by the research computing support organization, and provided to researchers at no cost to their project.

RESEARCH DATA MANAGEMENT SERVICES

Research data management services are a set of professional services that help researchers and scientists manage, organize, store, and preserve data collected by research projects. RDMS aims to ensure the integrity, accessibility, and usability of research data throughout its lifecycle, from creation to preservation.

PREVIOUS FINDINGS

In the 2020 survey [1], NIST found that most respondents found OISM's centrally provided data storage offerings unsuitable for capturing, processing, and archiving research data. Consequently, local solutions are maintained for research data storage, which may not be suitable for long–term data stewardship. Researchers also noted a significant increase in the volume of data to be managed. However, they felt they lacked the tools to effectively monitor this growing data volume for its own future use or for others, and they felt they didn't have the capabilities to intelligently archive data.

The EPOC study [2] found that highly advanced experimental tools are generating vast datasets across various scientific fields. These datasets have the potential to provide new insights with enduring societal impacts. However, scientists can't effectively utilize this data if they can't move, store, and analyze it. As data volumes grow and research collaborations extend beyond NIST boundaries, collaborators need access to model source code and data, currently provided through GitHub. They also found that data needs to be archived, backed up, and easily retrievable by scientists, as the data life cycle can span decades. EPOC reported that NIST's storage solutions consist of a variety of technologies and approaches. They also discovered that users often feel responsible for data mobility in or out of NIST, which can lead to wasted time solving technical problems independently instead of seeking help.

To address these issues, EPOC recommended that NIST prioritize investment in research data storage solutions in the coming years. They suggested developing a set of tools to assist with data mobility and creating a storage architecture that can meet high–level storage needs. They also advocate for improving the ability to integrate laboratory instruments.

CAPABILITIES

Three of the four study participants provided the requested information about four RDMS Capabilities: Research Data Storage, Public Repositories, Research Data Exchange, and Research Metadata Management. The findings are summarized in this section.

RESEARCH DATA STORAGE

Research Data Storage (RDS) refers to the systems and strategies used to store and preserve data collected during research projects. As shown in Figure 11, participating institutions reported a wide range of storage allocations, with most systems having no upper limits for individual project allocations. NIST provides all on–premises storage, and Lab D provides a majority of on–premises storage with some cloud storage offerings. Refresh cycles vary widely across the various repositories employed. Lab B reported just one RDS system, but no amplifying information beyond the average technology refresh rate. Lab A did not provide information on its RDS capabilities.

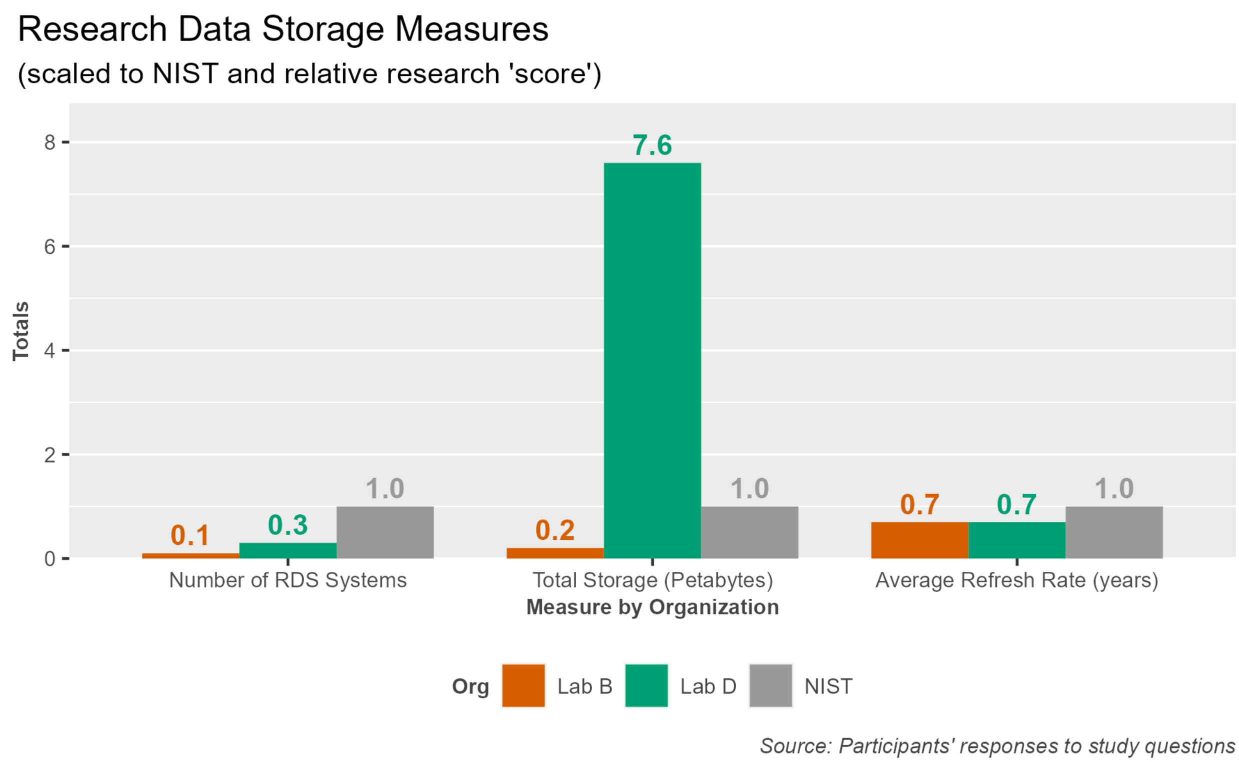

Figure 11. Research Data Storage Measures Summary (Source: MITRE)

PUBLIC REPOSITORIES

A public research data repository is an online database or archive where researchers can store, share, and access research data. These repositories are designed to preserve and disseminate data, making it available for further research, verification of published results, or for educational purposes. Some repositories are subject–specific, focusing on specific areas of research, while others are general and accept data from multiple research fields.

Two participants provided information about its public repositories: NIST and Lab D. Lab D consolidates its public repositories based on the sensitivity of the data, but each of the three levels has unbounded storage capacity and retention. Data refresh rates for NIST's public repositories are dependent on the source project (Lab D did not specify this measure). The NIST systems are on a continuous improvement refresh cycle. Lab D is replacing or upgrading its systems in 2024 after four years of service. While Lab A did not report any public data repositories, MITRE easily discovered a robust capability for Lab A, including visualization of the tens of thousands of datasets available to the public through multiple channels. Lab B, however, did not seem to make its research data easily discoverable or available to the public.

RESEARCH DATA EXCHANGE

Research Data Exchange refers to the process of sharing and transferring research data between researchers, institutions, or organizations. This can be done through various means, such as public research data repositories, direct data transfer, or data sharing platforms. The purpose of research data exchange is to promote transparency, collaboration, and reproducibility in research. It allows other researchers to validate findings, build upon existing research, and conduct new analyses. Research data exchange is often governed by data sharing policies and agreements to ensure ethical use and protect sensitive information.

NIST maintains several industry-standard research data exchanges, such as Globus,[23] Box,[24] Google Drive,[25] and SharePoint,[26] among others. Lab D maintains one industry-standard data exchange (Globus) and other sensitive exchanges for which they could not share details. Lab D's Globus hub provides 100TB of capacity, while NIST's are user dependent. The lack of data reported by the participants (other than NIST) makes it impractical to make any conclusions about research data exchanges.

RESEARCH METADATA MANAGEMENT

Research Metadata Management involves the organization, integration, and control of metadata associated with research data. Metadata is essentially data about data—it provides detailed information about the content, context, quality, structure, and accessibility of the research data. This can include information such as the data's author, date of creation, format, and source.

Effective metadata management is crucial for ensuring that research data can be easily discovered, understood, and reused by other researchers. It can involve creating standards and policies for metadata creation and use; designing and implementing metadata schemas; and using metadata management tools or systems. Research metadata management also plays a key role in data preservation, as it helps to ensure that the data can still be understood and used in the future.

Lab D and NIST provided a list of the systems available to provide inventory, citation, and/or re–use information to users. Lab D's capability is contained within its public repository systems. NIST provides four metadata management systems: the NIST Science Data Portal,[27] NIST Public Data Inventory,[28] the NIST Extensible Resource Data Model (NERDm),[29] and the Management of Institutional Data Assets System (MIDAS).[30]

[23] https://www.globus.org/
[24] https://nist.account.box.com/
[25] https://www.google.com/drive/
[26] https://office.com/
[27] https://data.nist.gov/sdp
[28] https://doi.org/10.18434/M31
[29] https://data.nist.gov/od/dm/nerdm/
[30] https://github.com/usnistgov/oar–midas–portal

RDMS BUDGET

Three of the four participants provided budget information related to its RDM capabilities in three requested categories: (1) provisioning/maintenance/operation/backup of research data storage services; (2) operation and maintenance (O&M) of publicly accessible research data repositories; and (3) O&M of data management systems. However, Lab B could only report figures for one of the three categories of spending. The insufficient data provided by the participants makes it impossible to reach any conclusions with respect to RDMS budgets.

CONFORMANCE TO BEST PRACTICES

The FAIR data principles[31] are guidelines that aim to make data findable, accessible, interoperable, and reusable. These principles emphasize the ability of machines and people to automatically find and use the data, as well as the supporting tools and workflows, with appropriate access and citation. The goal is to ensure that research data can be easily located, accessed, and used, both now and in the future, thereby enhancing the value and usefulness of such data.

The *Desirable Characteristics of Data Repositories for Federally Funded Research* is a guidance document released by the Subcommittee on Open Science of the National Science and Technology Council (NSTC) in May 2022 [6]. The document aims to improve consistency across federal departments and agencies in the instructions they provide to researchers about selecting repositories for data resulting from federally funded research. It includes applying the FAIR data principles to online, public access data repositories, while integrating privacy, security, and other protections. Individual departments and agencies may use these characteristics to guide the development of further instructions for the research communities they support.

NIST reports complying with both the FAIR data principles and the NSTC guidance. MITRE conducted a brief review of NIST's public data repositories and found no evidence to the contrary. Lab A did not indicate compliance with this guidance in its responses, but MITRE's brief scan of their public data offerings did not uncover anything to suggest that they did not comply. Lab B provides the ability for individual researchers to apply the guidelines (e.g., metadata management, public repositories), but does not enforce the practices on them. Lab D reports complying with the NSTC guidelines on its data hubs, but they are still working on implementing the FAIR principles. They do not consider the guidance to be applicable to its file systems, which represent 99 percent of its reported research data storage capacity.

RDMS BENCHMARKING

Although data in this area is limited to what the participants could provide, ***NIST is employing research data management best practices as they apply to federally funded research***. More data from additional research institutions would help NIST better gauge its RDM spending and storage capacity.

[31] https://www.go–fair.org/fair–principles/

RESEARCH COMPUTING SUPPORT STAFFING

Research Computing Support refers to the services provided to assist researchers in utilizing computing technology effectively and efficiently. This can include providing assistance accessing and using high–performance computing systems, and assistance with data storage and management solutions, software and hardware support, training and consultation on various computing tools and techniques, and data analysis and visualization. The goal of research computing support is to enhance the research capabilities and productivity of researchers by helping them leverage advanced computing technologies.

PREVIOUS FINDINGS

The 2020 NIST survey [1] revealed that *most respondents were frustrated with its ability to easily discover available research computing resources*. They were also dissatisfied with the scope, adequacy, and currency of the information they found. Furthermore, *most respondents expressed a strong desire for access to consulting staff who could assist them in using existing research computing resources and in developing software code* to address specific research problems.

The 2023 EPOC report [2] found that *OISM offers fee-for-service solutions that many research projects cannot afford*. This can lead to *projects opting out of these services, creating additional unsupported heterogeneity because the projects develop or seek out one-off solutions*. The report also noted that OISM can be overly bureaucratic, reducing agility for fast-paced research IT needs.

Regarding laboratory equipment, EPOC recommended that ADLP and OISM collaborate to improve the integration of laboratory instruments. This may require expanding the scope, hardware and software support, and team that can provide this service.

EPOC also found that staffing was a challenge, with *difficulties in recruiting new IT experts*, partly due to higher industry pay scales. New recruits often find the *NIST environment challenging due to the culture of allowing many disparate solutions to flourish*. While some IT expertise is distributed across NIST, these community members are often already overloaded.

Communication was another area of concern in the EPOC study, which reported that NIST resource users were often unaware of the full suite of IT services available to them. The use of short–term support can lead to a lack of institutional knowledge, and staff frequently rely on word–of–mouth to learn about services. Better communication with the research community about technology operation expectations and realities would benefit NIST. EPOC recommended that OISM collaborate with the RCAC and other stakeholders to improve communication about upcoming technology upgrades.

BUDGET

Figure 12 summarizes the RCSS budget data provided by three of the four participants. The chart depicts the categories of RCSS spending for each organization. NIST does not track RCSS spending using these categories, so was unable to report spending for Data-Facing costs, Software-Facing costs, and Systems-Facing costs. Therefore, those scaled measures are not included in Figure 12. Lab B and Lab D reported spending in those categories totaling 18.2 percent and 20.6 percent of their total RCSS spending, respectively, as shown in Figure 13. So, the bulk of their budgets are accounted for in Figure 12.

In the three categories reported, NIST's research computing support staff budget is high relative to the other study participants. NIST spends significantly more on operations,[32] but it is a smaller percentage of its overall RCSS budget than Lab B's (see Figure 13).

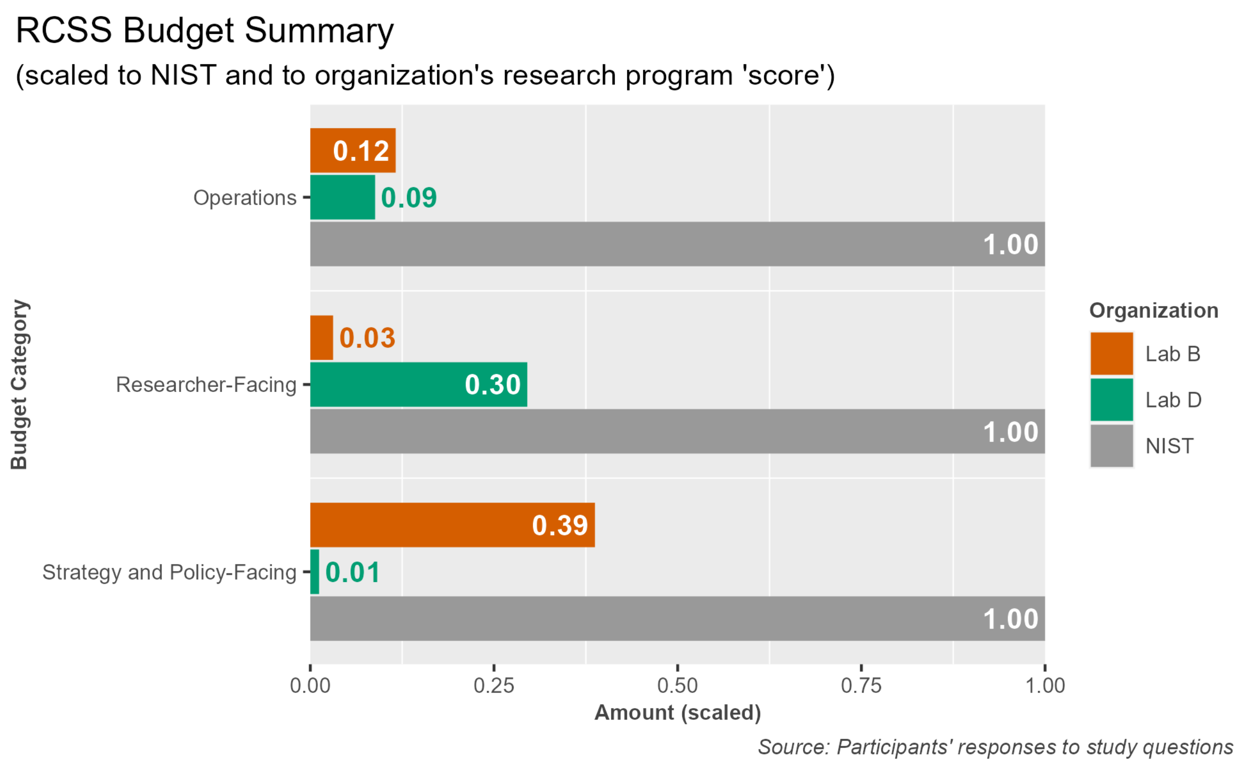

Figure 12. RCSS Budget Summary (Source: MITRE)

[32] None of the participants include facilities in their reported budget numbers since they are centrally funded outside the RCE organization in all case.

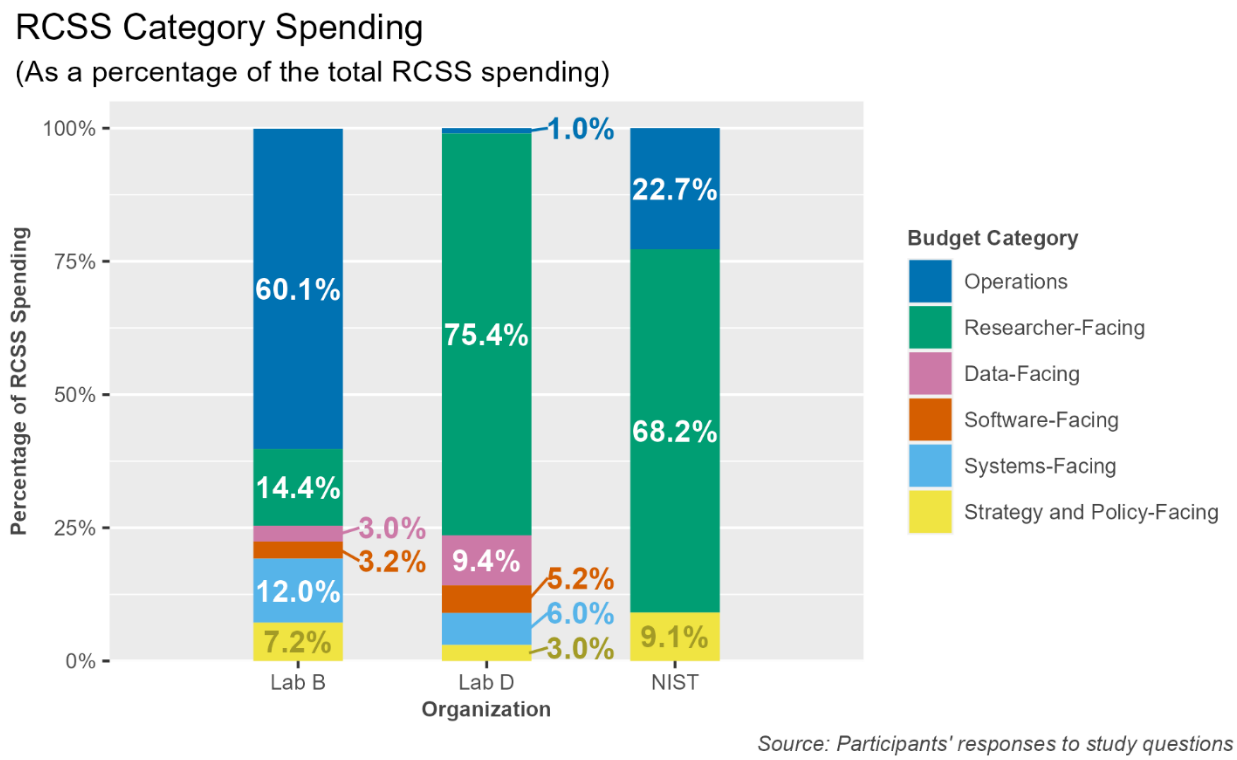

Figure 13. RCSS Category Spending as a percent of Total RCSS Spending (Source: MITRE)

STAFFING

Figure 14 summarizes the RCSS reported by three of the four study participants in the five roles described by the CARCC Facings model (see the High Performance Computing–Budget and Staffing section on page 18 for more details). The staff represented here are in addition to the staff reported in the HPC section. RCS staffing is focused on broader research computing support.

NIST reports maintaining an RCS staff of 52, compared to 25.5 for Lab D and 5.3 for Lab B. NIST's staff is mostly composed of Researcher-Facing roles. In addition, most of these support staff are in the NIST research organizations dedicated to assignments in those labs. Lab B reported over twice as many support staff as NIST in their research organization but nearly half as many researchers (see Figure 3). [33] With the relatively low number of researcher-facing roles reported by Lab B, compared to NIST and Lab D, it is likely that Lab B and NIST interpreted those questions differently.

[33] See also the *Research Scale: Budget and Staffing* discussion starting on page 5.

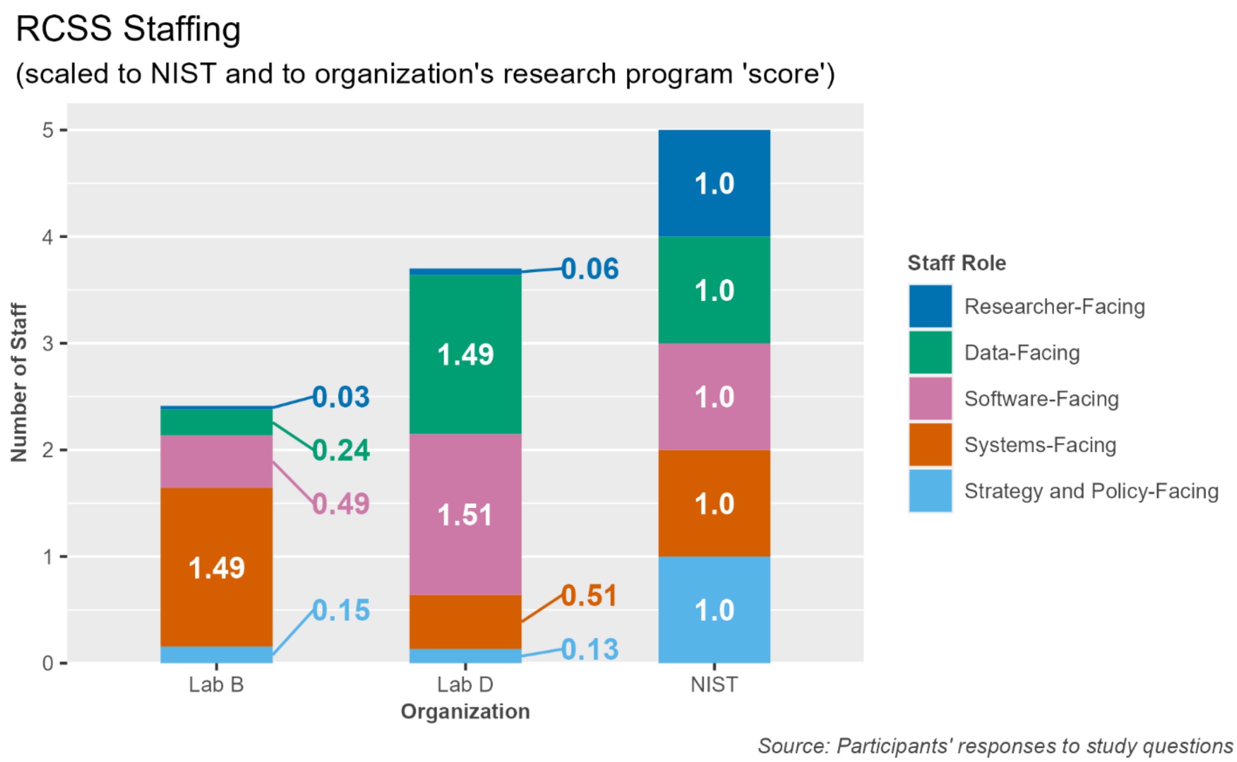

Figure 14. Research Computing Support Staffing (Source: MITRE)

RCSS BENCHMARKING

The data show that NIST invests much more than the other participants in research computing support, both in terms of dollars and people. This may be balanced by the relatively low numbers of support staff in the Research Program, as noted in Figure 4 on page 6.

However, the higher number of RCS staff at NIST seems to be at odds with the earlier findings of the NIST survey and EPOC study discussed above. Specifically:

- A strong desire for access to consulting staff who could assist them in using existing research computing resources and in developing software code (NIST survey).
- OISM offers fee-for-service solutions that many research projects cannot afford…leads to projects opting out of these services…develop or seek out one-off solutions (EPOC Report).
- Difficulties in recruiting new IT experts (EPOC Report).
- The NIST environment is challenging due to the culture of allowing many disparate solutions to flourish (EPOC Report).

This apparent mismatch may be a symptom of the very small sample size available for this study. A deeper dive into the costs and types of research computing support staff may be warranted. This would also include determining how other research laboratories make support discoverable and accessible.

NETWORKING

Networking plays a crucial role in the research program of an institution. The network is critical infrastructure that enables the following for research organizations:

- **Data Sharing**: Research often involves the collection of large amounts of data, which need to be shared among researchers, often in different locations. A robust network allows for the fast, secure, and reliable transfer of this data.
- **Collaboration**: Research is increasingly a collaborative effort, often involving teams spread across different locations, even different countries. Wide area networks enable teams to communicate and collaborate effectively, sharing data, resources, and ideas.
- **Access to Resources**: Many research tasks require access to specialized resources, such as high performance computing clusters, large databases, or scientific instruments. A reliable network allows researchers to access these resources remotely, often in real time.
- **Scalability**: As research projects grow, they may require more computing resources, more data storage, or more bandwidth. A well–designed network can scale to meet these growing needs.
- **Efficiency**: The network can help to automate data collection, data analysis, and reporting, increasing the efficiency of the research process.
- **Security**: Research data can be sensitive and valuable. The network infrastructure helps secure the data by protecting it from unauthorized access, data loss, and other threats.

PREVIOUS FINDINGS

The NIST survey [1] revealed that researchers are experiencing difficulties with almost all aspects of NIST's networking infrastructure that supports its technical work. These issues include the reliability and bandwidth of on–campus wired networking and intercampus bandwidth and remote network access to on–campus IT assets; and . This may constrain researchers access to cloud service providers and connectivity to external collaborators. The survey highlights the need for improvements in the networking infrastructure to support the researchers' work more effectively.

The EPOC Report [2] highlighted that the NIST enterprise network provides default speeds of 1Gbps, with higher speeds available for specific needs. Off–site access requires a VPN, which reduces responsiveness. The report suggests NIST could benefit from a faster network connection, potentially scaling access from 10 to 1,000 users. It also proposes exploring non–traditional networking approaches, like wireless edge and satellite–based networking. The report recommends NIST adopt modern practices from similar institutions, despite the challenge of keeping up with cyberinfrastructure trends.

CAPABILITIES

While MITRE asked for details of its respective network infrastructure, for security reasons, they were able to share very little. There was insufficient information collected to describe or provide a comprehensive comparison of network capability. However, in MITRE's opinion, ***based on the information provided, NIST's and the other participating institutions' network deployments appear to meet industry standards.***

MAINTENANCE AND FUNDING

NIST reported a 5-year target refresh schedule for some components but indicated insufficient funding to maintain that schedule. Other components are refreshed on an as-needed basis. Lab D reported a maintained refresh schedule of 5 years and Lab B refreshes as needed. Lab B and Lab D reported that network funding and maintenance are centrally funded, while NIST's funding model is mixed. This suggests that ***central funding of the network infrastructure is a norm among the study participants.***

PERFORMANCE

NIST expects its ongoing network upgrades to improve network performance levels. MITRE cannot comment further on current network satisfaction levels at NIST because it was unable to evaluate whether recent upgrades have increased performance.

NETWORKING BENCHMARKING

The scarcity of data from the study participants makes it difficult to draw any firm conclusions related to networking. The data collected does seem to suggest that ***central funding of the network infrastructure is a norm among the study participants.***

The previous findings related to the network infrastructure give NIST a potential area to explore, and it has been addressing it as noted above. However, MITRE suggests that NIST investigate whether the perceived performance issue is indeed caused by the network or is perhaps a symptom of *non-network* (i.e., computing hardware and/or software) bottlenecks.

SUMMARY OF RCE BENCHMARKING

This section summarizes and combines the results discussed in the previous sections (Findings: Research Programs and Findings: Research Computing Environment)

RESEARCH PROGRAM SUMMARY

In the Findings: Research Programs section (page 9), the participants' research programs were compared in terms of scope, scale, and impact. Table A summarized those results. That table is reproduced here for convenience as Table J. As discussed above, the overall research program "score" is used in this section to scale the values of the RCE components to define the benchmarks for each measure. That is, to determine whether NIST is providing RCE services at a level commensurate with other participants, we first apply reference scaling to the measure across all organizations by dividing their value by NIST's value for that measure. That results in a value relative to NIST's for each organization. Since each organization's research program has a different scale and impact (as shown in Table J), we then scale the values by the organization's "score." This assumes a correlation between the scale and impact of the research program and the level of RCE capability as expressed by the various measures used in this study.

Table J. Summary of Scaled Research Program Factors

Measure	Lab A	Lab B	Lab D	NIST
Research Budget	16.9	1.1	8.9	1.0
Number of Researchers	1.6	0.5	1.0	1.0
Weighted Citations per Document	0.7	1.3	1.2	1.0
Organization h–index	1.3	1.2	1.1	1.0
Overall Research Program "Score"	**5.1**	**1.0**	**3.0**	**1.0**

RCE SUMMARY

In the Findings: Research Computing Environment section (page 11), the participants' RCE components were measured primarily in terms of budget, staffing, and technical capacity. Each RCE section above summarized the quantitative results in tables and/or figures. Some of those tables are repeated here for convenience and to provide a complete view of the RCE.

BUDGET BENCHMARKS

As Figure 15 shows, the largest portion of each participant's total RCE budget is allocated for RCSS. More data from other organizations would be needed to provide a useful target distribution of RCE funds across the components.

Table K recaps each participant's RCE spending across four of the five components in this study.[34] NIST spends more on all components except HPC and more in total across RCE. NIST's RCE spending also represents a larger percentage of their overall research budget: 2.66%, as compared to 0.45 percent and 0.18 percent for Lab B and Lab D, respectively. In addition, NIST spends more on RCE per researcher than the other participants.

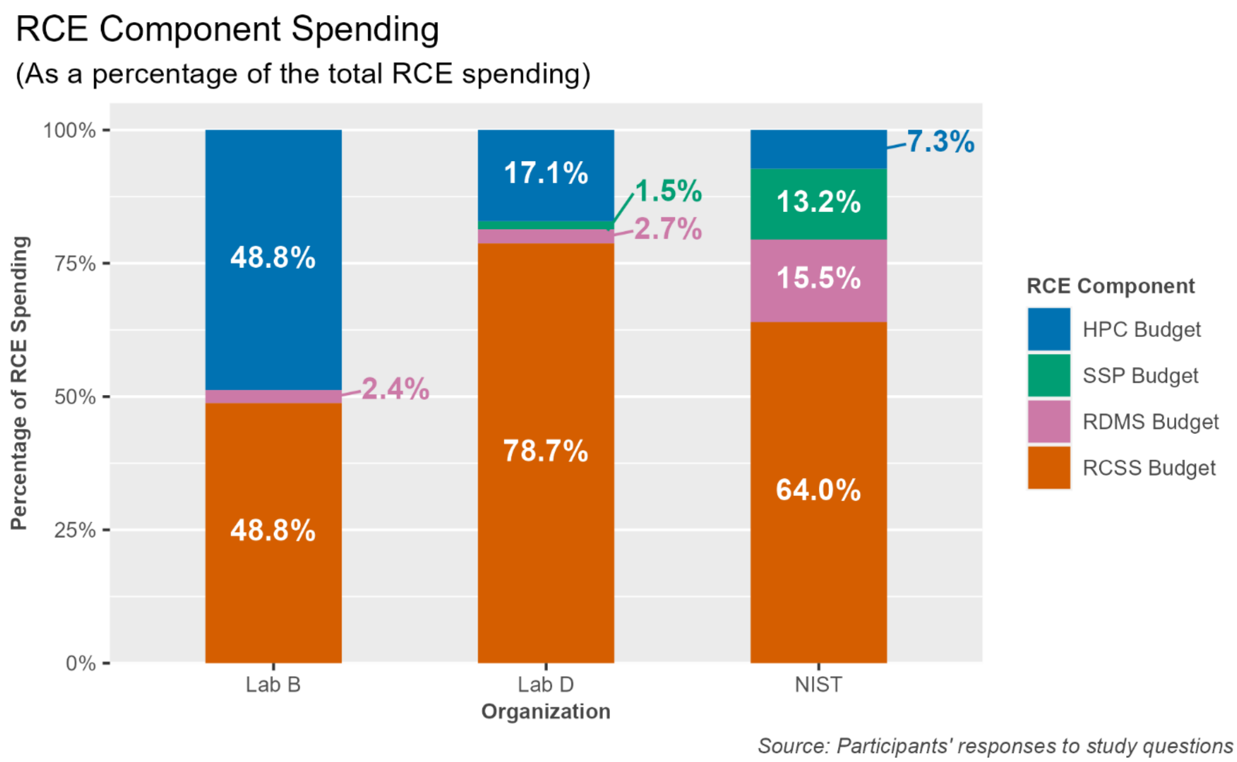

Figure 15. RCE Component Spending as a Percentage of RCE Spending (Source: MITRE)

Weighting[35] each participant's RCE budget by its research program score provides a slightly more nuanced view of the scaled investment in RCE across the study participants. As Figure 16 shows, NIST's investment in HPC, while seemingly low in absolute numbers (see Table K), when adjusting the values based on the scale and impact of the institution's research program, it is actually higher than the other two participants, although not dramatically so. However, Figure 16 does reinforce the disparity between NIST's overall RCE spending (and its spending in SSP, RDMS, and RCSS) and that of the other participants.

[34] Recall that participants did not provide network budget data for this study. In addition, Lab A did not provide any budget or staffing numbers for this study.
[35] Scaled by the research scores in Table A (and Table L) where appropriate.

Table K. Summary RCE Component Budgets

Measure	Lab B	Lab D	NIST
RCE Component			
HPC	$2,495,000	$2,900,000	$1,880,000
SSP	$0	$249,000	$3,418,000
RDMS	$125,000	$450,000	$4,000,000
RCSS	$2,495,000	$13,317,133	$16,500,000
Total RCE Budget	$5,115,000	$16,916,133	$25,798,000
Research Program Measures			
Research Budget ($ Millions)	$1,130	$9,350	$1,045
Number of Researchers	1,882	3,328	3,492
RCE Investment			
RCE Budget as percent of Research Budget	0.45%	0.18%	2.47%
RCE $ per Researcher	$2,718	$5,083	$7,388

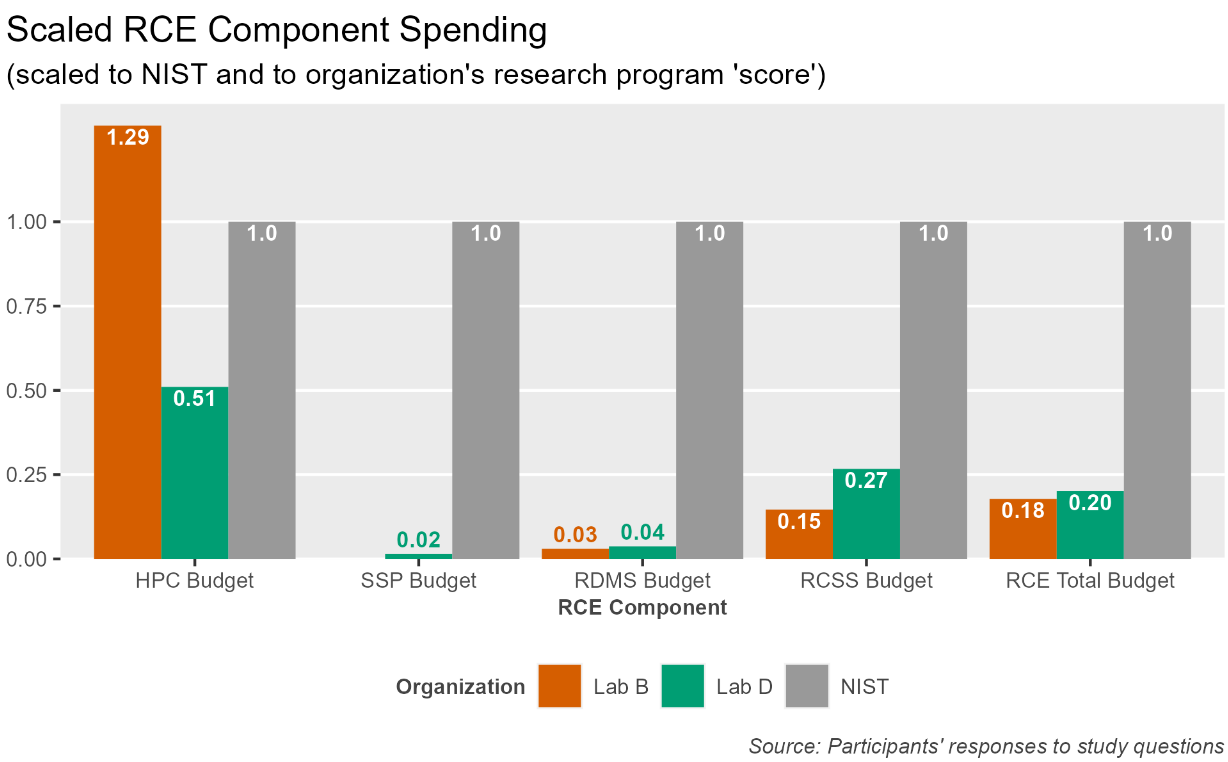

Figure 16. Scaled RCE Component Budgets (Source: MITRE)

Table L further refines the scaled investment NIST makes in RCE against Lab B and Lab D when scaled by their research program scores. While NIST has a slightly higher adjusted research budget than Lab B, it has a significantly lower adjusted research budget than Lab D. However, NIST's scaled total RCE budget is 3–10 times that of the other participants. That ratio holds when considering the RCE budget as a percentage of the research budget and the amount spent on RCE per researcher in each organization.

That said the small sample size of this study makes it imprudent to conclude whether NIST is under- or over-spending on its RCE investments. However, the results imply that more data is needed from other research institutions.

Table L. Scaled RCE Investment

Measure	Lab B	Lab D	NIST
Scaled Research Budget	1.05	2.96	1.0
Scaled Total RCE Budget	0.19	0.22	1.0
Scaled RCE Budget as percent of Research Budget	0.18	0.07	1.0
Scaled RCE$ per Researcher	0.37	0.69	1.0

RCE STAFFING BENCHMARKS

The scaled RCE staffing numbers in Figure 17 indicate that, with the exception of Lab D's HPC staffing, NIST maintains significantly more staff in research computing than the norm. This is consistent with the budget findings above.

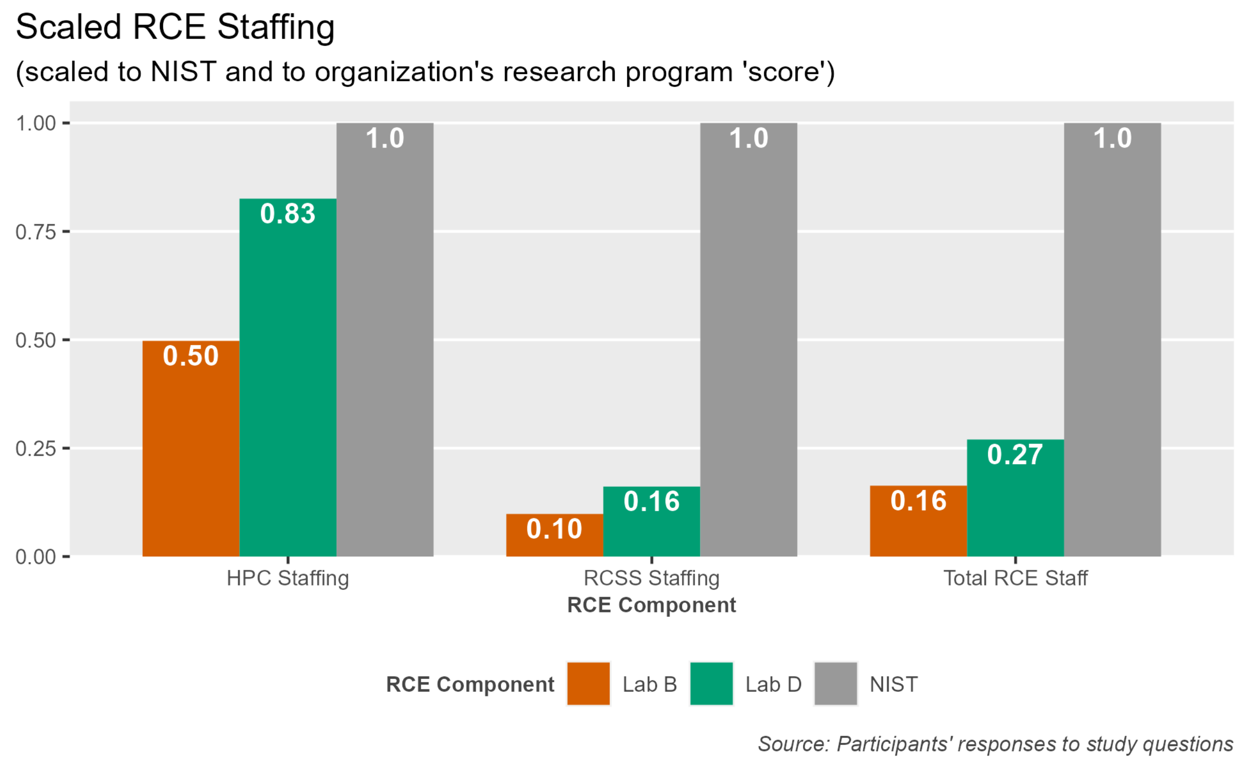

Figure 17. Scaled RCE Staffing (Source: MITRE)

HPC BENCHMARKS

All study participants operated multiple HPC systems with multiple general-purpose systems and at least one with GPU capability. NIST operates the most HPC systems with many more focused on specific research groups than other participants. NIST also had fewer users per system than the others. Overall, NIST had far fewer researchers on general-purpose systems than all other participants with only 20 percent of total users being on general-purpose systems compared to more than 60 percent at other participating institutions. This suggests that:

> *The norm for research organizations is to provide fewer but more powerful general-purpose HPC systems.*

HPC SYSTEM STABILITY CHALLENGES

The responses provided by the study participants regarding the top challenges to system stability further support the conclusion above. While the top challenges to system stability varied across all participants and systems, NIST uniquely reported struggling with the heterogeneous hardware. Heterogeneity across the system was listed as a major challenge for three separate NIST systems, all of which have differing CPU types and memory available across nodes. Seven of NIST's eleven clusters have a mix of multiple types of CPUs, resulting in a more complex mix of CPUs than any of the participants. Furthermore, while all organizations reported several

HPC clusters with GPUs, NIST supports the widest range of GPUs. In addition, many of NIST's listed GPUs are not general-purpose GPUs but are instead specialized for graphics rendering.

The highly heterogeneous nature of many of the NIST systems would be expected to cause additional overhead for systems-facing staff working on the clusters and cause challenges for users working with less portable codes.

The norm among this study's participants appears to be for more homogenous, general-purpose HPC systems.

HPC LIFECYCLE PLANNING

The participants all provided succession and decommission plans for many of its systems. NIST's future planning for system replacement and upgrades was more limited with only two of the eleven systems having upgrade or replacement plans. NIST also supports systems with a much longer lifespan than the other participants. Although Lab D reported a concern with one of its HPC systems' hardware age, it has a replacement plan in place for the older hardware showing the challenges are already being addressed. Aging hardware may impact the ability of researchers to make innovative developments and will challenge systems-facing staff to keep the clusters and storage systems operational.

For on-premises HPC systems, the norm among this study's participants is for planned obsolescence of HPC systems with an expected lifespan of 4–6 years.

HPC STORAGE

NIST provides multiple storage systems with each HPC cluster having a dedicated storage system. This is consistent with the norm supported by the data:

Variety in the types of HPC storage systems employed seems to be the standard practice, including global storage systems used across multiple systems.

HPC UTILIZATION

Regarding utilization, the participants reported higher utilization (around 60–80 percent) across general-purpose HPC clusters with the exception of one general purpose system at Lab D. While NIST operates one high–utilization HPC general-purpose HPC system, two of the systems operate at 19 percent and 27 percent. Despite having lower utilization, NIST reported longer wait times across several systems, sometimes taking weeks to run. Again, this finding suggests that:

Focusing on fewer but more powerful, homogenous systems might alleviate NIST's utilization issues.

HPC USER COMMUNICATION

In terms of service management, feedback, and marketing, other participants have a cohesive strategy for informing research staff about systems and collecting feedback from users. The process at NIST is more varied in terms of promotion and collecting feedback as well as the process to go through to request an account.

HPC CAPABILITIES

As discussed in the Findings: Research Computing Environment section (page 16) and recapped in Table M, NIST's computational capability is reported as lower than all participants, even when scaled by the size and impact of the research program. NIST also provides a more complex mix of CPUs than any of the participants, more non-general-purpose systems, and some GPUs atypical of those ordinarily used in high performance computing. These results, overall, suggest that:

> *NIST has room to improve its HPC capabilities to make them more commensurate with comparable institutions.*

Table M. Summary of Scaled HPC Measures

Measure	Lab A	Lab B	Lab D	NIST
HPC Systems and User Capacity (Table B, p. 11)				
Systems	0.22	1.01	0.67	1.0
Total Active Users	1.31	2.79	1.76	1.0
Max Active Users Per System	1.65	2.32	2.97	1.0
HPC Storage				
Base Storage (TB)	1.00	0.50	0.20	1.0
Total Storage (PB)		4.30	0.86	1.0
HPC Performance, Cores, and Nodes				
TFLOPS	3.93	2.76	1.53	1.0
Nodes	2.32	1.78	0.32	1.0
Cores	4.13	6.45	0.52	1.0

SSP BENCHMARKS

As discussed in the Findings: Research Computing Environment section (page 24), the data collected for this study do not support any conclusions related to the size or cost of a benchmark scientific software portfolio. However, they do indicate that centrally funded provisioned software does not appear to be the norm.

> *The SSP practice common to all participants is that the majority of scientific software is provisioned either on a fee-for-service or a bring-your-own-software model.*

RDMS BENCHMARKS

As discussed in the Findings: Research Computing Environment section (page 27), NIST is employing research data management best practices as they apply to a federally funded research.

As shown in Table N, NIST reports spending more on each category of RDMS than any of the other reporting participants. However, NIST also has more obligations to share data due to its unique mission, likely accounting for the difference. NIST also maintains several solutions in each of the categories of RDMS (research data storage, public repositories, research data exchange, and research metadata management).

This suggests that while more information is needed:

NIST provides a level of RDMS that out-performs comparable research institutions.

Table N. Research Data Management Services Scaled Budget Summary

Measure	Lab B	Lab D	NIST
Budget: Provisioning/O&M/ backup of RDS	0.08	0.01	1.00
Budget: O&M of RDM systems	0.00	0.20	1.00
Budget: O&M of Public Repositories	0.00	0.02	1.00
Total RDMS Costs	**0.03**	**0.04**	**1.00**
RDMS Budget as percent of Total RCE Budget	0.17	0.18	1.00
RDMS $s per Researcher	0.06	0.12	1.00

RCSS BENCHMARKS

As discussed in the Findings: Research Computing Environment section (page 32) The data show that NIST invests much more than the other participants in research computing support, both in terms of dollars and people. This seems to be at odds with the earlier findings of the NIST survey and EPOC study. However, NIST's relatively high number of support staff in the research organizations could balance this discrepancy. NIST's RCSS spending, in terms of dollars-per-researcher, appears to be in line with Lab D, its closest peer, by the measures used in this study. While it was outside the scope of this study, these results suggest that:

NIST could benefit from a deeper dive with Lab D into how they provide RCS services with the aim of improving researchers' perspective on the support provided.

Table O. RCSS Scaled Budget Summary

Measure	Lab B	Lab D	NIST
Operations Costs	0.39	0.01	1.00
Researcher-Facing Costs	0.03	0.30	1.00
Strategy and Policy-Facing Costs	0.12	0.09	1.00
Total RCSS Budget	**0.15**	**0.27**	**1.00**
RCSS Budget as percent of Total RCE Budget	***0.82***	***1.33***	***1.00***
RCSS $s per Researcher	***0.28***	***0.85***	***1.00***

NETWORKING BENCHMARKS

As discussed in the Findings: Research Computing Environment section (page 36), the scarcity of data from the study participants makes it difficult to draw any firm conclusions related to networking. However, the data collected does seem to suggest that:

> *Central funding of the network infrastructure is a norm among the study participants.*

SUGGESTIONS FOR FURTHER RESEARCH

This study was a first attempt to characterize and benchmark research computing environments in the context of the supported research program. As such, along with the benchmarks that it could define, it serves as baseline of approaches that worked and did not work.

The constrained scope and timeframe for this study, together with relatively poor and extended response rates to the study questions asked of the participants, limited our ability to dive deeper into some areas of interest. It also restrained the ability to solicit engagement with additional comparable institutions in time to complete the study.

As mentioned throughout the report, MITRE suggests additional research to better understand potential challenges and solutions. Specifically:

- Revise some of the metrics used in this study and/or devise additional metrics to better characterize both research programs and research computing environments. This may require a more collaborative approach with comparable research institutions.
- Augment the scope of this study by further refining the questions to be asked and cast a broader net with ten or more additional research institutions. The methodology developed for this study allows for a broader set of research institutions to be solicited for input.
- More data from additional research institutions would help NIST better gauge its RDM spending and storage capacity (page 30) as well as its RCSS spending.
- More data from other organizations would be needed to provide a useful target distribution of RCE funds across the components (page 38).

Appendix A Acronyms

Term	Definition
ADLP	Associate Director for Laboratory Programs
ASJC	All Science Journal Classification Codes
CARCC	Campus Research Computing Consortium
CHIPS	Creating Helpful Incentives to Produce Semiconductors
COTS	Commercial Off-the-Shelf
CPU	Central Processing Unit
EPOC	Engagement and Performance Operations Center
FLOPS	Floating point Operations Per Second
GPU	Graphics Processing Unit
HPC	High Performance Computing
IQR	Interquartile Range
IT	Information Technology
MIDAS	Management of Institutional Data Assets System
NERDm	NIST Extensible Resource Data Model
NFS	Network File System
NIST	National Institute of Standards and Technology
NSTC	National Science and Technology Council
O&M	Operations and Maintenance
OISM	Office of Information Systems Management
RCAC	Research Computing Advisory Committee
RCSS	Research Computing Staff Services
RDMS	Research Management Database Services
RSO	Research Services Office
SJR	SCImago Journal Rank
SSP	Scientific Software Portfolio
USPTO	U.S. Patent and Trademark Office

Appendix B References

[1] NIST, "Executive Summary of Research/Technical Computing Survey & Follow-up Meetings with Respondents," 2020.

[2] J. Zurawski and J. Schopf, "National Institute of Standards and Technology Requirements (Analysis Report)," Lawrence Berkeley National Laboratory (LBNL), 2023.

[3] Elsevier, "How is SJR (SCImago Journal Rank) used in Scopus?," 2024. [Online]. Available: https://service.elsevier.com/app/answers/detail/a_id/14883/supporthub/scopus/kw/SJR/.

[4] Elsevier, "Scopus," 2024. [Online]. Available: https://www.scopus.com/. [Accessed 2024].

[5] N. Berente, J. Howison, J. Cutcher-Gershenfeld, J. L. King, S. R. Barley and J. Towns, "Professionalization in Cyberinfrastructure," *SSRN,* 2017.

[6] Subcommittee on Open Science, "Desirable Characteristics of Data Repositories For Federally Funded Research," National Science and Technology Council, 2022.

[7] F. Bordignon, "Dataset of search queries to map scientific publications to the UN sustainable development goals," *Data in Brief,* vol. 34, p. 12, January 2021.

[8] Elsevier, "Scopus," [Online]. Available: https://www.scopus.com/. [Accessed 2023].

[9] USPTO, "Patent Public Search Basic (PPUBS Basic)," 2024. [Online]. Available: https://ppubs.uspto.gov/pubwebapp/static/pages/ppubsbasic.html.

[10] NIST, "Patent and Published Patent Application Database," 2024. [Online]. Available: https://www.nist.gov/patents.

[11] LENS.org, "LENS.org," 2024. [Online]. Available: https://www.lens.org/.

[12] Google, "Google Patents," 2024. [Online]. Available: https://patents.google.com/.

[13] J. Martindale, "What is a teraflop?," digitaltrends.com, 14 June 2021. [Online]. Available: https://www.digitaltrends.com/computing/what-is-a-teraflop/. [Accessed 3 February 2024].

[14] P. V. Kamat, "Absolute, Arbitrary, Relative, or Normalized Scale? How to Get the Scale Right," *ACS Energy Letters,* no. 4, pp. 2005-2006, 2019.

Appendix C Research Area Analysis

This appendix provides detail for the research area analysis MITRE conducted and summarized in the Findings: Research Programs section (page 5).

Research Scope Analysis

All Science Journal Classification Codes (ASJC) provide a reasonable method of defining the scope of each participant's research program. ASJC is a subject classification system developed by Elsevier. It includes 27 fields and 334 subfields [7]. Elsevier, through their Scopus web application, publishes summary data on number of documents by subject area for an organization or author [8].

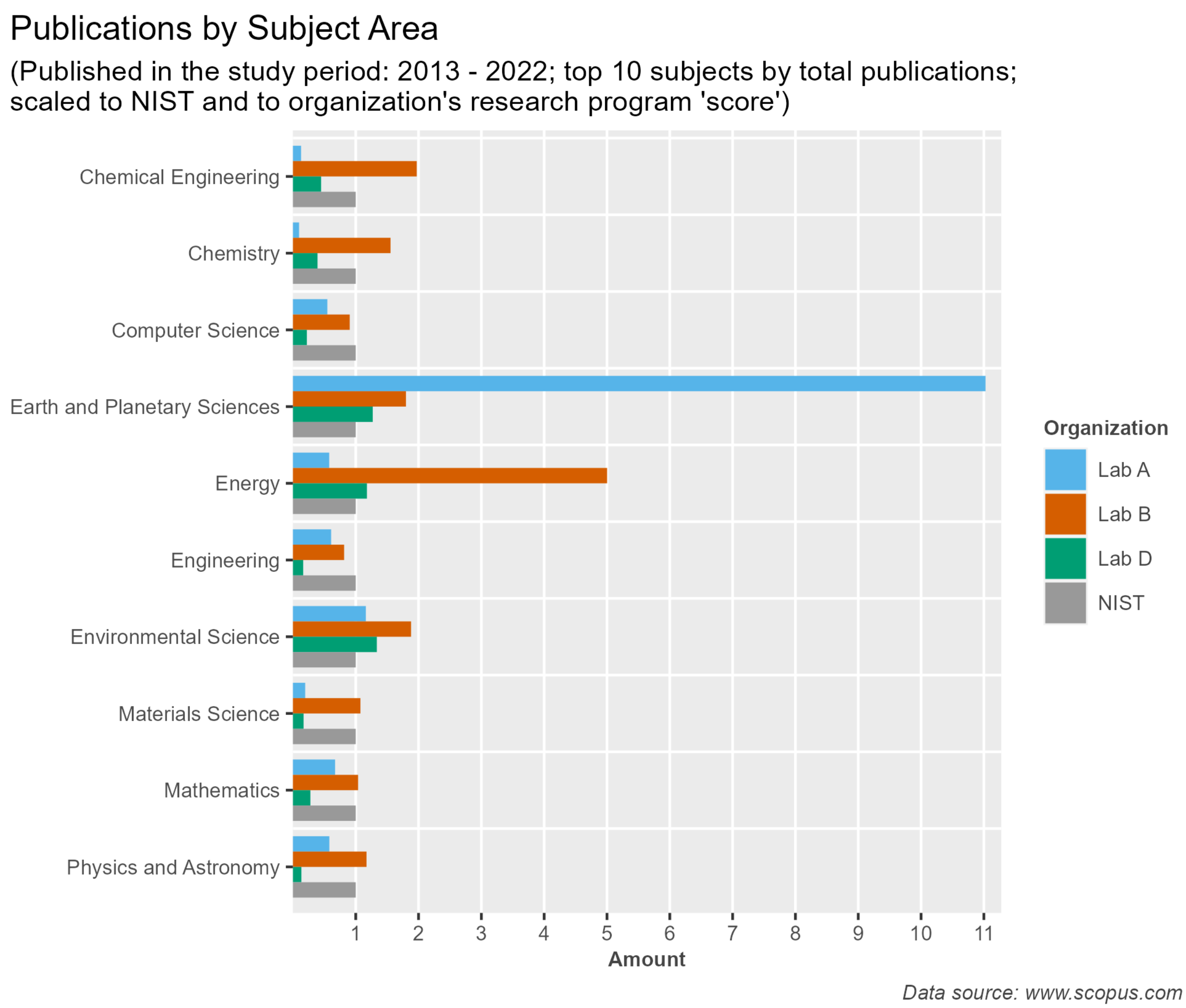

Figure C-1. Study Participants' Publications by Subject Area

Figure depicts the number of documents for each participant in the top 10 AJSC subject areas across all participants (ranked by the total number of documents in each subject area across all study participants). The top 10 subject areas account for over 80 percent of the total documents for all participants across all subject areas. Each participant's top 10 subject areas were close (with two standard deviations) to the ranking of the four participants' subject areas, overall, indicating a relatively strong subject area correlation among the organizations.

When scaled to the research program score and to the number of NIST's publications, NIST outperforms, by a good margin, both Lab A and Lab D. It is most closely correlated with Lab B in this measure (correlation coefficient of 0.92).

Figure C-2 provides another view of the same subject area data, this time showing the distribution of the documents as a percentage of the total number of documents published by each participant in the top 10 subject areas. The notable differences between the participants are Lab A's larger focus on earth and planetary sciences, and its lesser focus than the other study participants on biochemistry, genetic and molecular biology, chemical engineering, chemistry, and materials science.

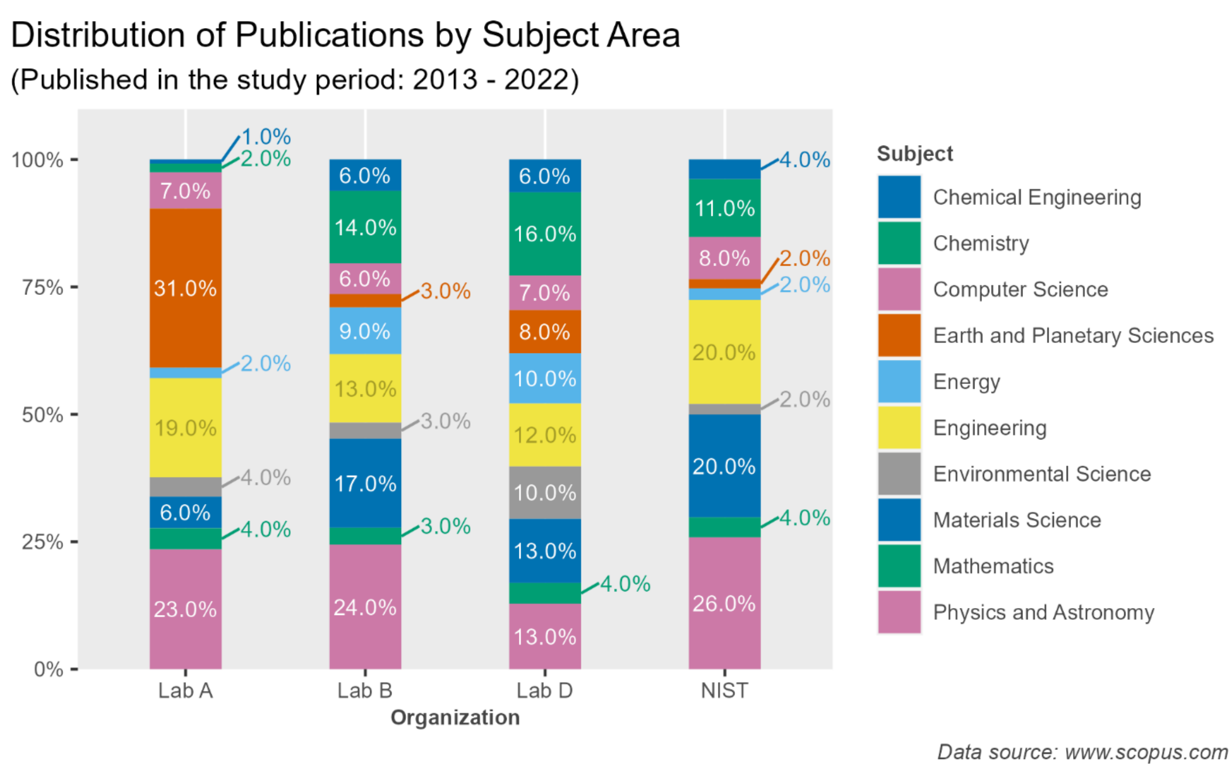

Figure C-2. Distribution of Publications by Subject Area (Source: MITRE)

Research Scale Analysis

In this study, MITRE related the participants by two measures of their size: the total annual research budget (for fiscal year 2023), and the number of researchers (i.e., "users" of RCE resources). Note that while Labs B and D and NIST provided estimated research budget and staffing numbers for this analysis, Lab A did not. However, Lab A is a government agency, so budget and staffing data were publicly available. MITRE derived the totals used in this study from that information.

Figure C-3. depicts the number of staff in the participants' research organizations by type. As the chart indicates, researchers outweigh the management, support, and other staff in those organizations. Because it is the researchers who create demand for and are the users of the RCE, their contribution to the total will be a better indicator of the scale of each participant's research program. Researchers represent over 75 percent of all research staff across the four organizations, and 92 percent at NIST.

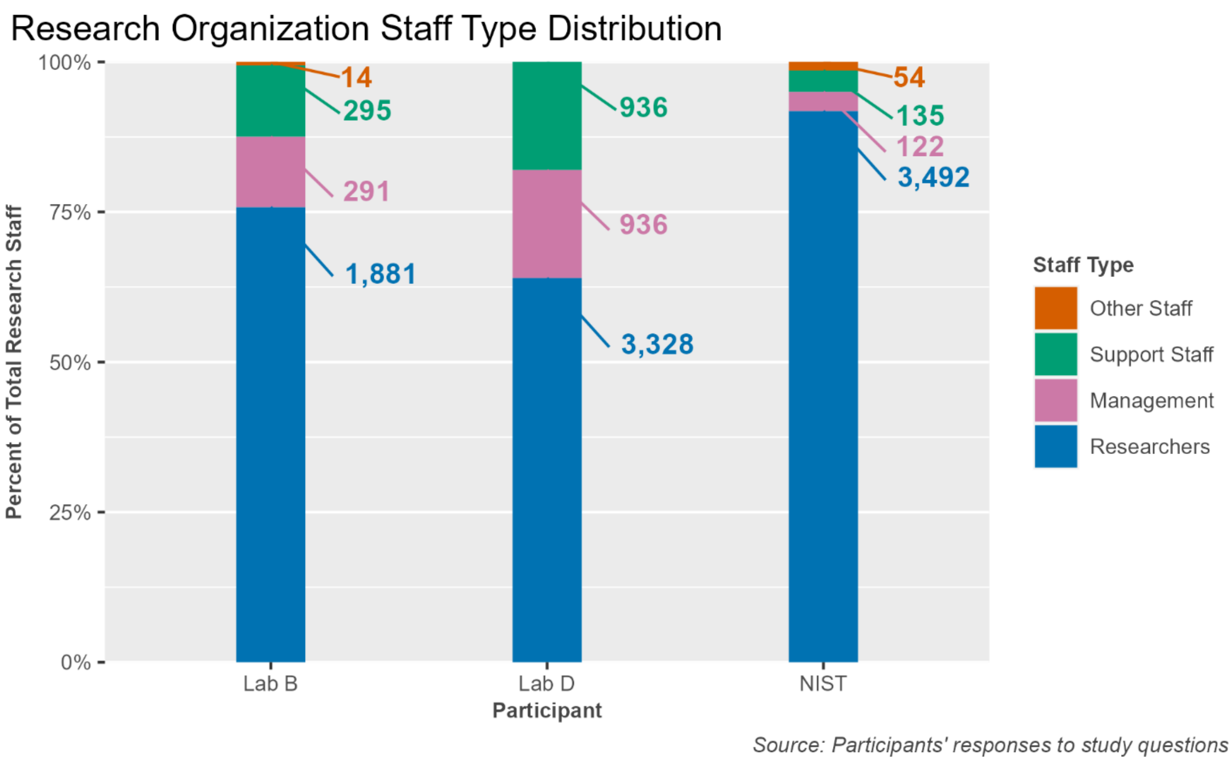

Figure C-3. Research Organization Staff by Participant and Staff Type (Source: MITRE)

Figure C-4. summarizes the FY23 research budgets for the four study participants. MITRE calculated the "cost per researcher" by simply dividing the total research budget by the number of researchers reported by the participants (see Figure C-3. and Figure 4 on page 6).

As the figure shows, Lab A's research budget ($18 billion) is significantly higher than the other participants. Lab B and NIST have similar research budgets ($1 billion). Lab D's budget ($9 billion) is half that of Lab A but nine times that of NIST. These factors are included in the benchmarking as measures to compare the RCE components.

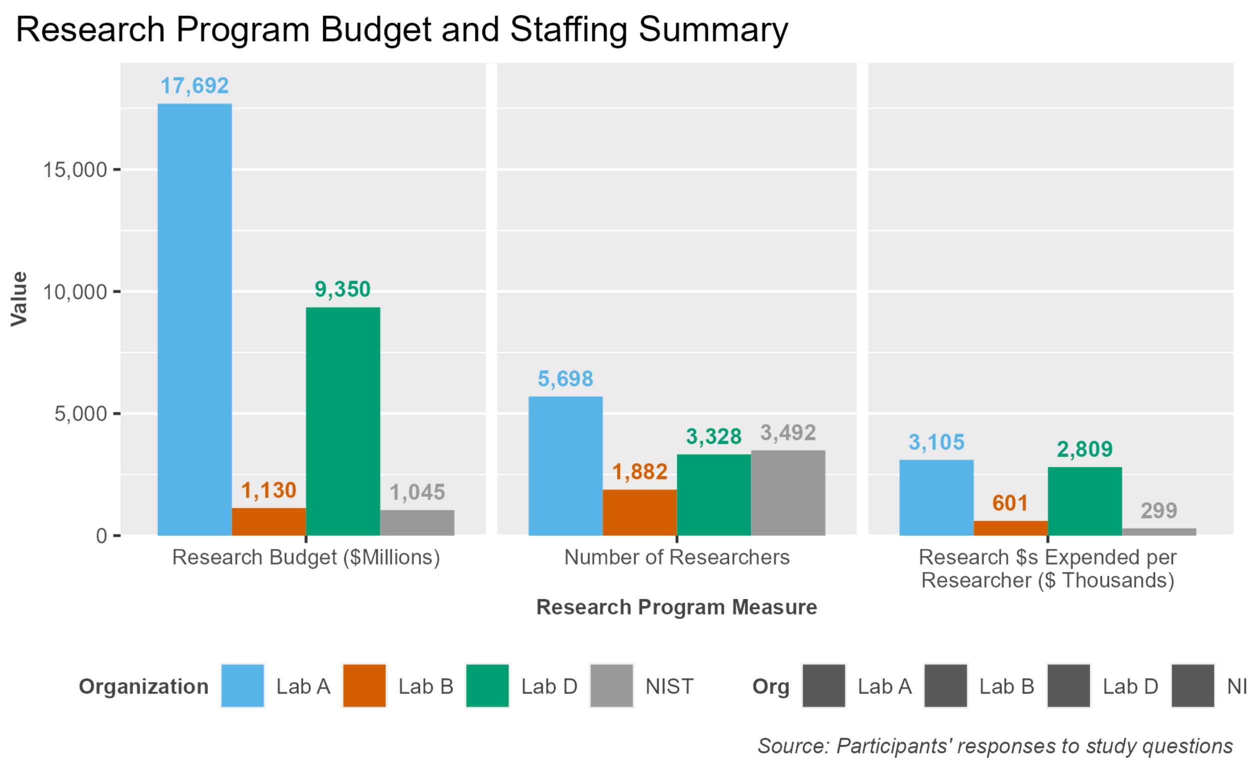

Figure C-4. Total FY23 Research Budgets and Cost per Researcher (Source: MITRE)

Research Impact Analysis

MITRE investigated research impact along three dimensions: (1) publications and citations (i.e., publications cited by other authors), (2) collaborations with other research institutions, and (3) patents applied for or awarded.

Publications & Citations

Figure shows the number of documents by each participant during the study period. These include only peer-reviewed documents: research articles, review articles, conference proceedings, data papers, and book chapters. MITRE retrieved publication data for all study participants from Scopus (www.scopus.com), an abstract and citation database launched by Elsevier in 2004. If one or more authors of a document were from a participating organization, that document is counted as one document. If two or more authors from different organizations collaborated on the document, the document is counted for both organizations.

Publications

Figure C-5 depicts the study participants' number of peer-reviewed documents from 2013–2022. NIST is on par with Lab B and Lab A, and well ahead of Lab D considering the relative sizes and impacts of their research programs.

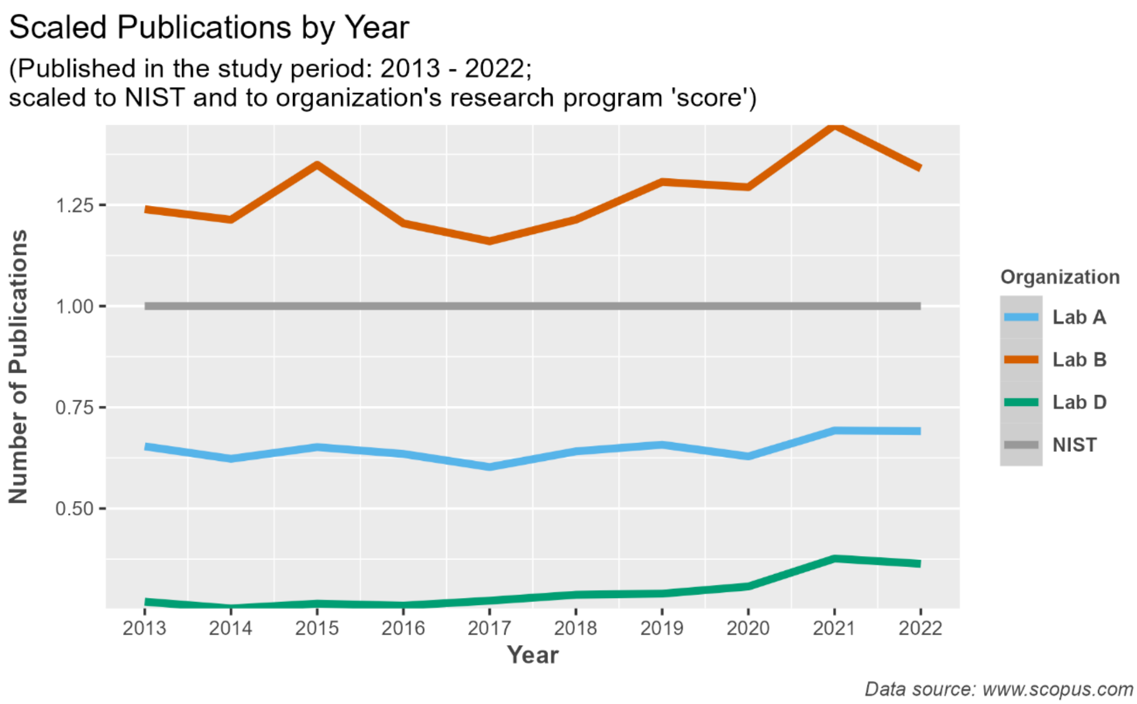

Figure C-5. Scaled Publications by Year (2013–2022) [8] (Source: MITRE)

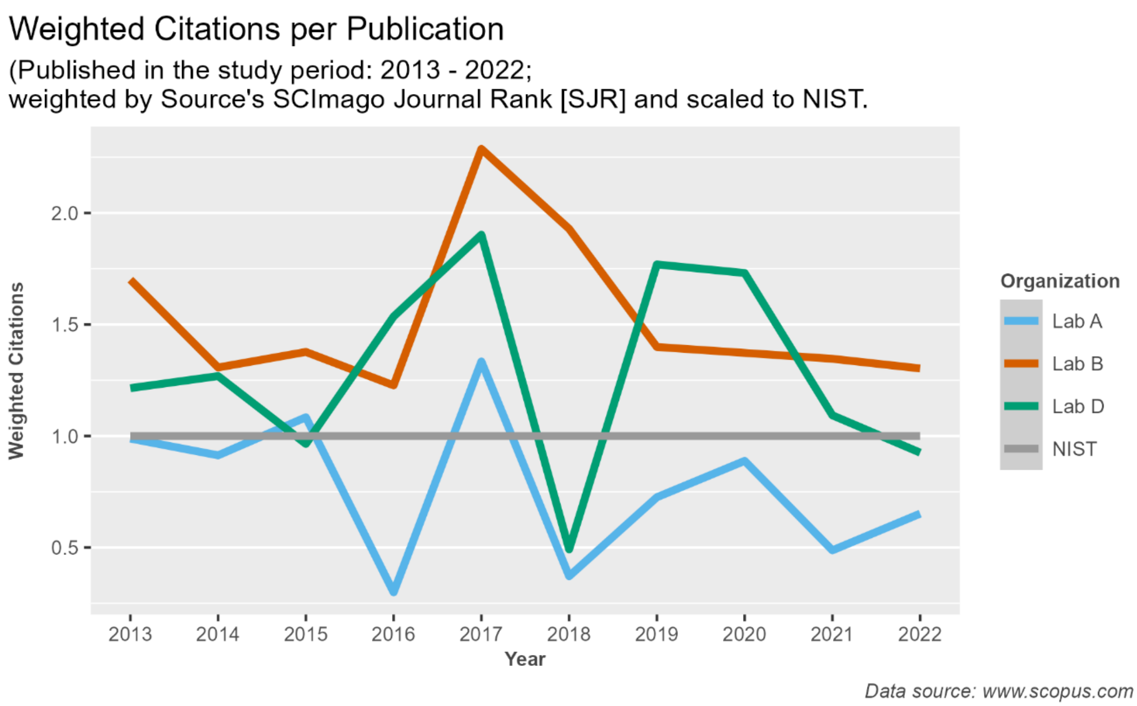

Figure C-6. SJR-weighted [8] Publications Cited by Other Publications (Source: MITRE)

Citations

While the number of peer-reviewed documents for an organization is a measure of *output* of a research program, the *impact* of the program requires more nuanced measures. For a more meaningful impact measure, the scientific community uses measures related to the number of times each of their documents is cited by other authors in other peer-reviewed documents. These data are also available from Scopus. Figure C-6 summarizes the SJR-weighted citations for each organization over the study period. These charts tend to tail off because there is often a two-to-three-year lag between the time of document publication and the time it is cited by another author.

While numbers of documents and citations are useful measures, they do not account for the relationship between the size of the organization and its output and impact. Figure C-7 depicts the distribution of each organization's citations in the form of a box plot. The box plot shows how the number of citations per document are distributed, with the box illustrating the spread between the second and third quartile of values (the "interquartile range (IQR)") and the median of all the documents' citations.

The points above the box represent outliers (i.e., values greater than 1.5 times the IQR above the 75^{th} percentile. Since the number of citations has no natural upper bound and a minimum value of zero, and most documents receive a moderate number of citations (13 or fewer in this sample), the distribution is positively skewed with many outliers on the upper end. The plot also includes a mean value for each to further illustrate the positive skewing of the distributions.

To make the figure more readable, the outliers for each participant that are cited by more than 500 other documents are not shown. They are summarized at the top of each box plot (e.g., Lab A has 909 outliers with more than 5.0 weighted citations, with a maximum outlier of 928 weighted citations).

The relative number and distribution of citations across the four participants show that the impact of their research is similar.

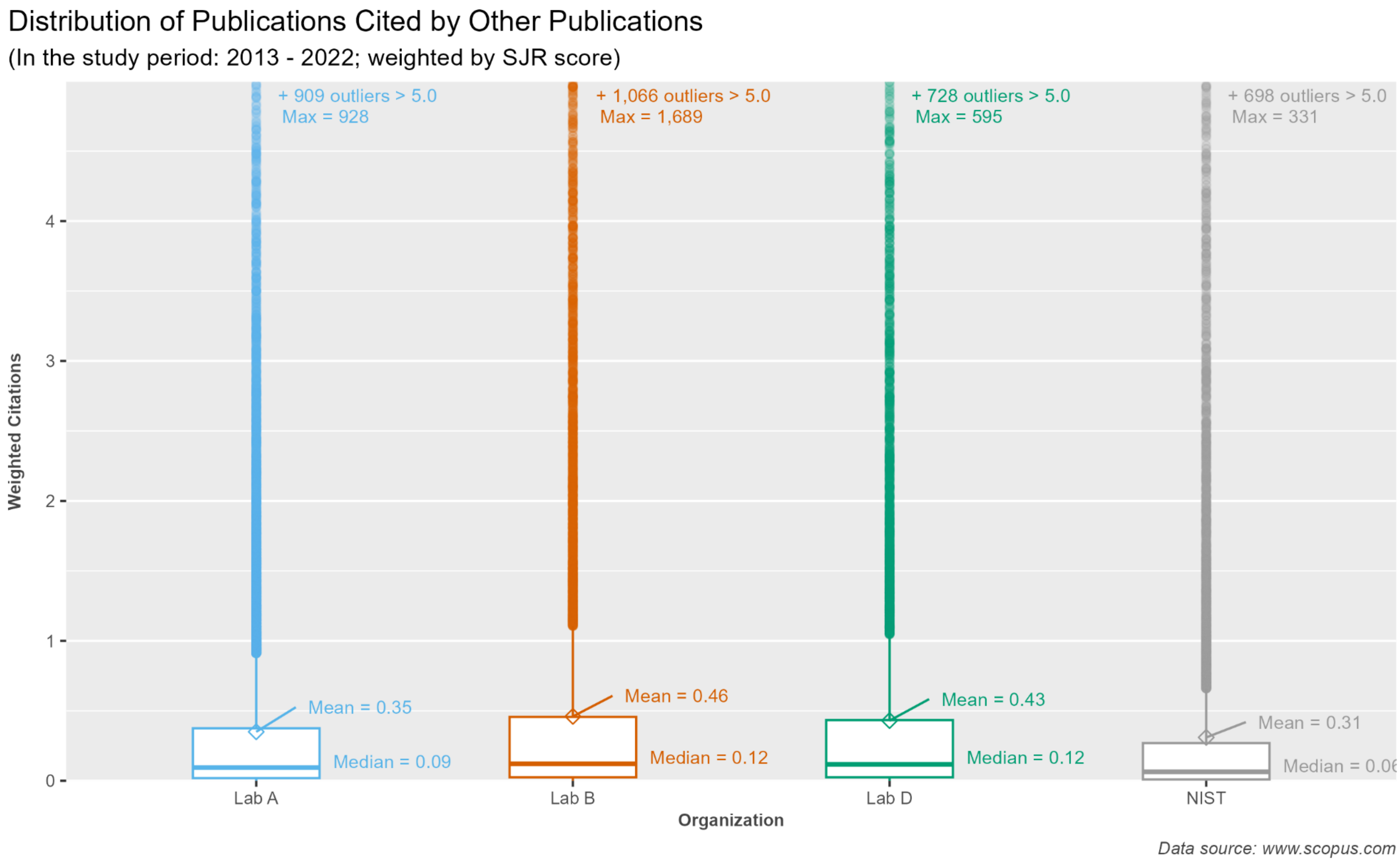

Figure C-7. Distribution of Publications Cited by Other Publications [8] (Source: MITRE)

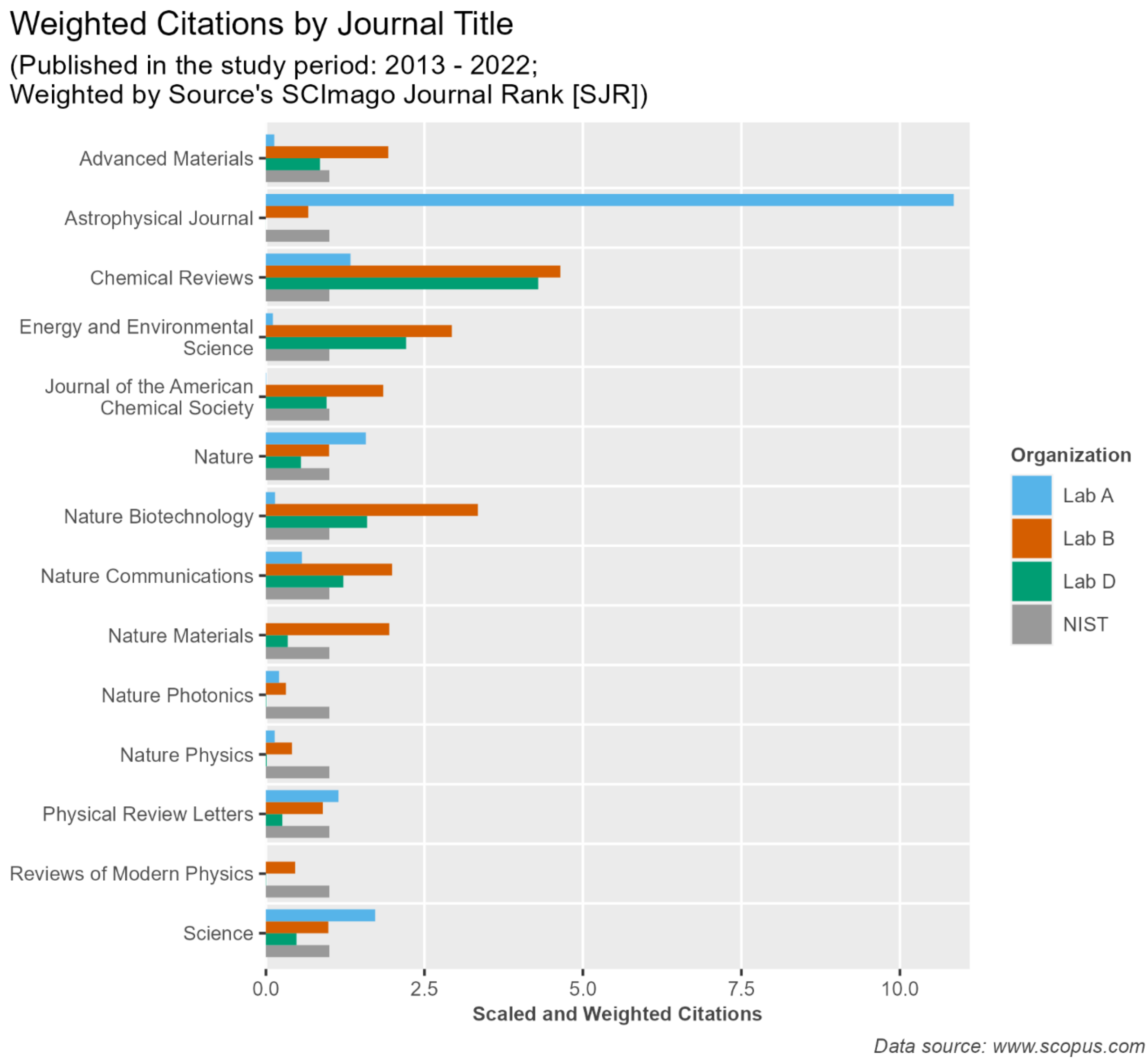

Figure C-8. Publications by Source Title

Weighting Citations by Journal Ranking

To further distinguish between the participants, it is important to look at additional factors of their impact, namely, where each participant's articles are published. Figure C-8 depicts the number of documents each participant published during the study period by the title of the journal.

The *SCImago Journal Rank* (SJR) provides a weighting by the prestige of a journal.

According to Elsevier [3]:

> *SJR assigns scaled scores to all of the sources in a citation network. Its methodology is inspired by the Google PageRank algorithm, in that not all*

> *citations are equal. A source transfers its own prestige, or status, to another source through the act of citing it…. A source's prestige for a particular year is shared equally over all the citations it makes in that year; this is important because it corrects for the fact that typical citation counts vary widely between subject fields…. The result is to even out the differences in citation practice between subject fields and facilitate direct comparisons of sources. SJR emphasizes those sources that are used by prestigious titles.*

Essentially, a "***citation from a source with a relatively high SJR is worth more than a citation from a source with a lower SJR*** [3]."

The "h-index"

The "h-index" is a measure that attempts to evaluate the *productivity* and *impact* of an author's published work. It is based on the set of the author's most cited papers and the number of citations they have received in other publications. The h-index reflects both the number of publications and the number of citations per publication.

For example, an h-index of 20 means that the author has 20 papers that have each received at least 20 citations. This index can also be used to measure the productivity and impact of a group of scientists, such as a department or university or country.

Org	CitedBy	Rank	
NIST	2942	1	
NIST	2362	2	
NIST	1833	3	
NIST	1803	4	
NIST	1430	5	
NIST	1406	6	
NIST	1346	7	
⋮	⋮	⋮	
NIST	248	242	
NIST	248	243	
NIST	247	244	
NIST	246	245	h-index =
NIST	246	246	246
NIST	246	247	
NIST	244	248	
NIST	244	249	
NIST	243	250	
NIST	242	251	
⋮	⋮	⋮	

Figure C-9. H-index Calculation Illustrated (Source: MITRE)

The h-index reflects a balance between the quantity and quality of an author's work. It avoids the bias of considering only the total number of papers or total number of citations, which can be skewed by a single highly cited paper or a large number of rarely cited papers. Figure C-10 summarizes the h-indexes for the study participants. The h-index was calculated by ranking all of each organization's publications in descending order by the number of times a document was cited by other publications. The h-index was the lowest number for which the rank of a document in the list was greater than or equal to the number of citations that document had.

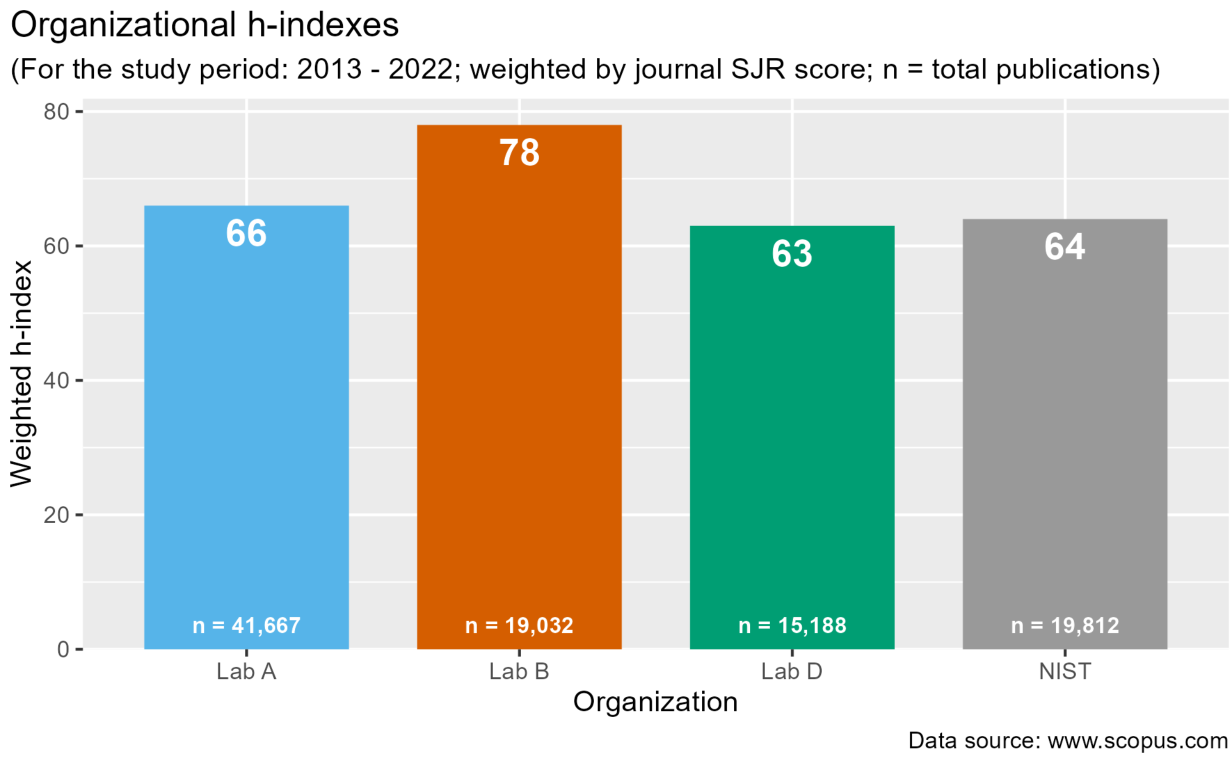

Figure C-10. Study Participants' h-index (Source: MITRE)

Collaborations

Another indicator of the impact of an organization's research is the number of collaborative relationships it establishes with other researchers. This can be measured by the affiliations of the co-authors on papers published by researchers at each organization. Figure C-11 shows the total number of collaborations each participant had on all of its documents during the study period. Notably, Lab B had significantly more collaborations represented in the Scopus data than the other participants. MITRE is uncertain, however, what type of effect this might have on the demand for RCE resources. One could argue that Lab B could be more heavily leveraging its collaborators' RCE. Conversely, the collaborators could be leveraging Lab B's resources. From the data available, it is impossible to draw any conclusions.

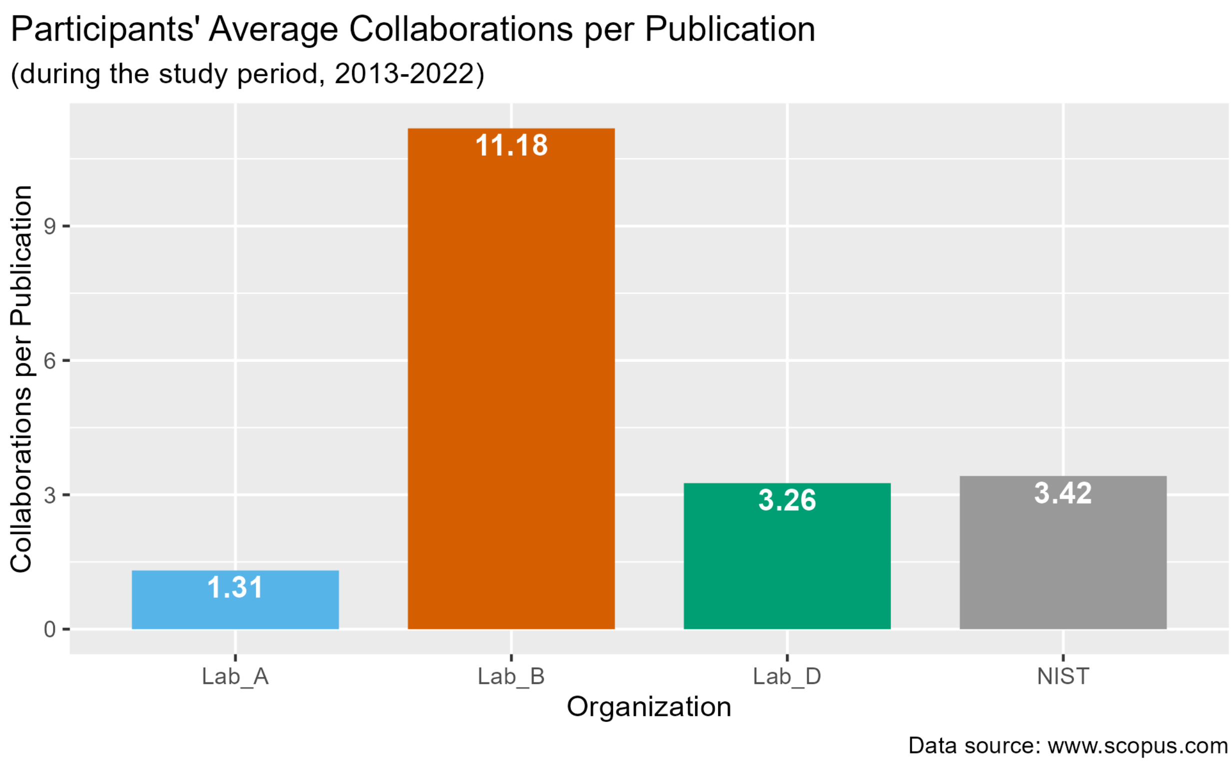

Figure C-11.Participants' Collaborative Publications (Source: MITRE)

Patents

Another measure of the impact of a research program is the total number of patents filed and/or granted. Figure C-12 through Figure C-14 summarize, in varying views, the best information available to MITRE at the time of this report on the number of patent applications filed by the study participants during the study period (2013–2022). MITRE derived this information from a combination of data from the U.S. Patent and Trademark Office (USPTO) [9], NIST's own reporting on patents [10], LENS.org [11], and Google Patents [12]. We scaled the data by each organization's research score, as described on 9.

Patents may take, on average, about three years from the time the application is filed ("file date") to the time the patent is approved by USPTO ("grant date" or "patent date"). It can take much longer for complex, obscure, or hotly-contested inventions. In most cases, the majority of the work that goes into an invention occurs prior to the time the inventor files an application for a patent. This is the period when the innovation, research, experimentation, design, etc. take place to develop a patentable invention. Once the inventor files the patent application, they will usually move on to the next innovation. With the exception of the occassional USPTO "office action" (typically handled by a patent attorney or agent), there is little for the inventor to do concerning the invention under consideration.

While patents are a typical measure of the impact of a research program, NIST has faced varying levels of "enthusiasm" for patent filing over several Administrations, so MITRE does not consider it a useful measure for the purposes of this study.

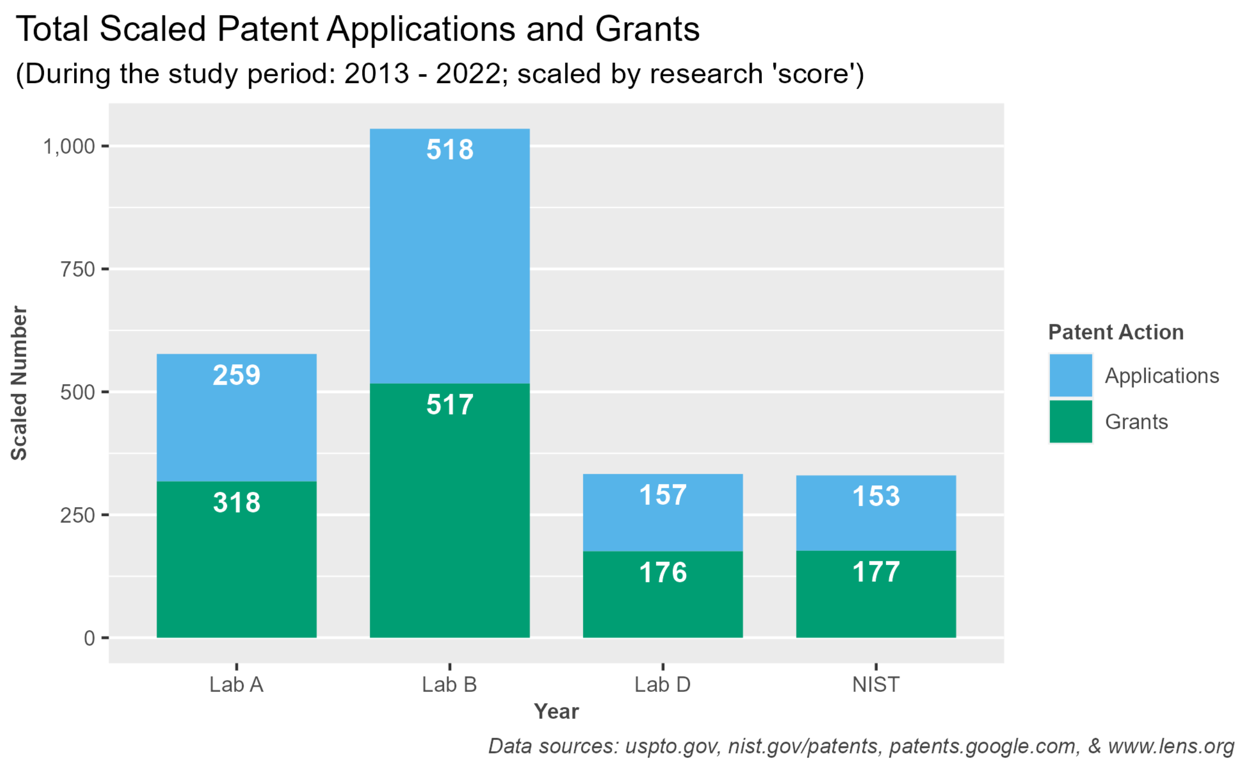

Figure C-12. Total Patent Applications and Grants (2013–2022) (Source: MITRE)

Table C-1. Scaled Patent Application and Grant Totals

Action	Lab A	Lab B	Lab D	NIST
Applications	8.74	3.25	3.09	1.00
Grants	9.11	3.02	2.99	1.00

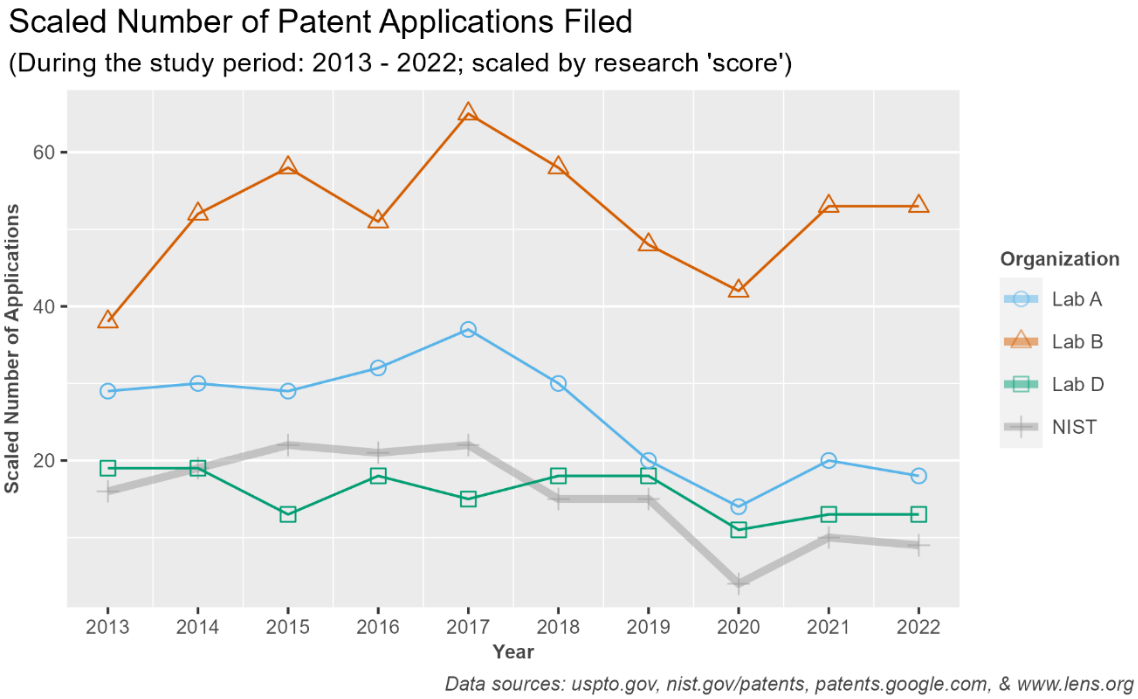

Figure C-13. Patents Filed by Year (Source: MITRE)

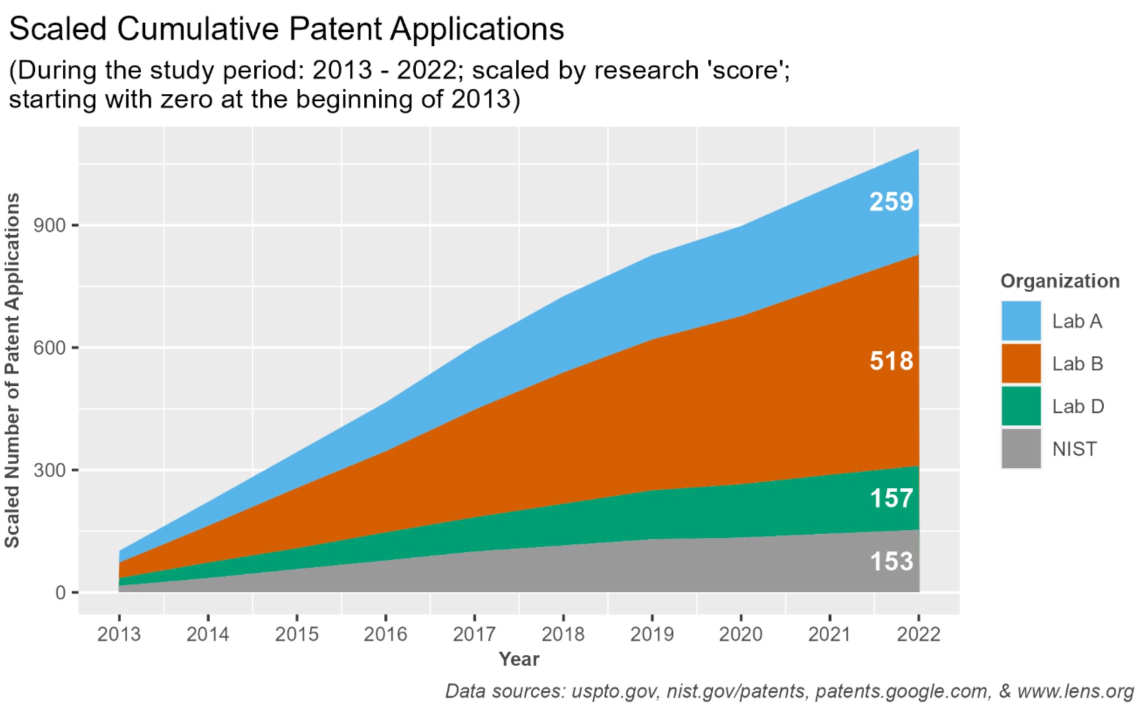

Figure C-14. Cumulative Patent Applications (Source: MITRE)