Contents

Executive Summary 3
Key Findings and Recommendations 4
Introduction 5
Defining Research Computing: More Than Infrastructure 7
Findings 9
Maturity 25
Conclusions 27
Further Exploration 27
Recommendations 28
Methodology 29
Acknowledgments 33
Appendix 34

Author

Jacqueline Bichsel, EDUCAUSE Center for Applied Research

Citation

 Executive Summary

The provision of research computing resources has until recently been considered an undertaking of highly research-intensive institutions in response to data-intensive scientific research. However, institutions that were once more teaching oriented are transitioning to place greater emphasis on research. Researchers who once would have settled only for a position at a “Research 1 University” are being recruited by and are being drawn to less research-intensive institutions. In addition, scholars in the humanities are developing increasingly data-intensive research interests. In conjunction with this changing landscape, the ability to collaborate and partner with others globally has lessened the need to house oneself in a resource-laden facility with local experts.

The provision of research computing services has also become more global, and the information technology (IT) departments at less research-intensive institutions are finding themselves negotiating over which data-intensive research needs should be met, how they should be met, and what IT’s role should be in instituting or sustaining these services. There are also questions about how these services should be funded. IT leaders may also wonder how their institution compares with others of similar research intensity. Are they providing the same kinds of research computing services as similar institutions? Are they organized to support these services similarly? Do they have similar numbers of staff supporting research computing?

The 2012 study on research computing was designed to address these questions. This report provides research computing benchmarks for institutions of low, medium, and high research intensity. The report concludes with an overview of the factors that will make up the forthcoming ECAR Research Computing Maturity Index.
Key Findings and Recommendations

- **Information technology’s primary research computing role is to serve as an enabler of research.** IT should be proactive in uncovering research computing needs and communicating which research computing services are available.

- **Provision of specialized staff**—particularly programmers, analysts, and data-intensive research staff with subject matter expertise—is just as important as the provision of hardware and software resources for research computing.

- **There is room to improve the effectiveness of research computing practices and strategies.** Even institutions of high research intensity did not rate any of their research computing practices/strategies any higher than “somewhat effective” on average.

- **The structure and organization of research computing support relates to the quality of research computing practices and services.** The most effective means of support is provision of an integrated set of IT services.

- **IT departments should leverage the research computing services and infrastructure they provide to partner with researchers on grants.**

- **IT should publicize the research computing services they provide to encourage faculty research and partnerships.**

- **Adopting a funding model that IT assesses as effective is key for successful research computing practice.**
Introduction

ECAR’s last report on research computing, in 2006, compared the then-contemporary research environment with what it had been previously and found that (a) more collaboration and teamwork characterized the contemporary research atmosphere; (b) more interdisciplinary research was occurring; (c) more knowledge systemization was developing; and (d) the use of high-performance computing (HPC) to support all of the above was being instituted.²

A case study on the University of Washington in 2010 outlined some of the IT needs of researchers: (a) IT and data management expertise; (b) data management infrastructure; (c) computing power; (d) communication and collaboration tools; and (e) data analysis and collection assistance. They found that researchers needed not only advanced technical infrastructure to meet these needs but also specialized staff assistance.³

From the survey and focus group data collected in the 2012 study on research computing, we can glean that this research environment has not changed considerably. There remains a focus on enabling collaboration and teamwork, both intra- and interinstitutionally. New ways of organizing and systematizing knowledge and data continue to be developed. Additional federal requirements necessitate more archiving of data for longer periods of time. The use of HPC in conjunction with these efforts continues to enable IT to help researchers by providing an infrastructure and applications with more efficiency; faster processing speeds; greater data storage capacity; and improved collaboration, stability, and security. The need for specialized staffing assistance continues to grow.

A recent Advanced Core Technologies Initiative (ACTI) white paper addressed an investigation into long-term cyberinfrastructure issues by the National Science Foundation’s Advisory Committee for Cyberinfrastructure. ACTI recommended that (a) higher education leaders should invest more in local cyberinfrastructure resources; (b) higher education IT staff should communicate one-on-one with researchers to develop specialized data visualization tools; (c) institutions need to bridge and network hardware resources from local to national centers; and (d) specialized software assistance from dedicated staff is a critical investment in addressing research needs.⁴ These recommendations are echoed to varying extents in this report.
The 2012 study of research computing addresses the effectiveness of research computing strategies and practices in the current higher education IT environment. The objectives of this research study are:

1. Provide a contemporary definition of research computing.
2. Explore the extent to which research computing is a competitive advantage for an institution.
3. Identify the current and optimal roles of IT in research computing support at both research-intensive and non-research-intensive institutions.
4. Develop a maturity index of research computing practice.

To address this study’s objectives, data from the Core Data Survey (CDS) Module 1 (IT Organization, Staffing, and Financing; \( N = 675 \)) and Module 4 (Research Technology Services; \( N = 427 \)) were analyzed. In addition, focus groups were conducted to collect qualitative data that corresponded to these objectives. For details on the questions asked on both the CDS and in the focus groups, refer to the Methodology section at the end of this report.
Defining Research Computing: More Than Infrastructure

Focus group participants for the ECAR 2012 study on research computing were asked what they would include in a definition of research computing. Themes that emerged can be classified as follows:

The resource demands of research computing involve
- Server space/storage
- Networking (intranet and Internet connectivity)
- Security
- Workstations
- Specialized laboratory equipment
- Specialized training
- Specialized scientific applications
- Software/hardware maintenance
- Involvement of dedicated personnel specializing in data visualization, statistics, programming

Research computing may also include
- High-performance computing
- Dedicated budget lines for maintenance and personnel
- Collaborative and outreach efforts
- Integration of computing-intensive research into pedagogy
- Administrative structures for policies and resource allocation

Provided with this feedback, the following definition was developed:

*Research computing* consists of the infrastructure, applications, expert staff, policies, and other resources required to support data-intensive activity related to research.

Although research computing has historically been the purview of fields such as engineering, biology, and computer science, other disciplines—including history and other social science and humanities fields—have ventured into data-intensive research.
Research computing is:

- Any service or infrastructure provided to faculty or research staff for the purpose of performing research at a higher education institution
- Highly resource-intensive; demanding of computing infrastructure, hardware, software, and personnel
The expectation for at least a modicum of research computing support is rising even among smaller and less research-intensive institutions.

“Before [we provided RC services], one of our faculty was trying to build a data cluster inside a janitor’s closet. He spent the whole time figuring out how to cool it, time that he could have spent on his research.”

—Focus Group Participant

Findings

Results discussed in this section were obtained from quantitative analyses of the CDS items and qualitative analyses of the focus group responses. Specific questions analyzed can be found in the Methodology section at the end of this report.

Strong Research Computing Resources Provide a Competitive Advantage

Focus group participants were asked whether research computing impacts faculty recruitment, performance, and retention. Most participants agreed that having research computing resources makes an institution more competitive in recruiting and retaining faculty with research computing needs. Many mentioned that this was particularly true for newer faculty, who—after having recently obtained a graduate degree from a program in which strong research computing support was provided—come to the interview inquiring about hardware/software availability, storage capacity, networking capability, Internet connectivity, computing speed, and IT staff with specialized programming skills. Although these inquiries have become customary at research-intensive institutions, the expectation for at least a modicum of research computing support is rising even among smaller and less research-intensive institutions.

Participants also thought the presence of research computing (RC) services enhanced faculty retention, especially if additional staff support was provided. Provision of specialized staff—particularly programmers, analysts, and scientific software specialists—was viewed as especially beneficial for faculty, particularly in kick-starting research programs that may not yet be funded. Several participants commented that they have noticed boosts in faculty research performance after instituting research computing services, especially in obtaining research grants. RC resources help faculty spend more time on research and less time on programming, analysis, system administration, networking, and data storage and management. A centralized RC infrastructure that is available to all faculty also helps ease the burden individual researchers face in finding funding for research computing for every project.

Research Computing Has Ancillary Benefits

Although faculty derive the primary benefit from research computing, the focus group participants noted benefits for others.

1. The institution—IT in particular—has the potential to benefit from more grant funding by providing RC support. Faculty will often include partial or entire IT staff salaries and equipment expenditures in their grant budgets. In addition, having an RC program already in place may enhance the ability of researchers to
obtain grants if they can make the case to reviewers that critical infrastructure is already in place. More than 50% of CDS Module 4 (Research Technology Service) respondents stated that their IT staff play some sort of role in their institution's current research grant process.

2. Research computing pushes IT to keep current with infrastructure and to develop new systems and processes. It may also provide institutions with a more robust infrastructure for non-RC services and hardware. For example, research computing increases resources for institutional assessment and analytics; institutional researchers and administrators benefit from having increased data storage and communication capacity. The library may also benefit from the increased storage capacity and speed that an RC infrastructure provides, as libraries are increasingly transitioning from paper to digital materials. Libraries will also be better able to support researchers in accessing data sets with enhanced RC infrastructure.

3. The enhanced infrastructure provided by research computing helps form a community of researchers. Data are more easily shared and communicated intra- and interinstitutionally. Collaborations and partnerships with industry and with other institutions are more easily formed with shared RC infrastructure.

4. Research computing helps generate PR and enhances the reputation of the institution both locally and nationally. Institutions with RC infrastructure are able to provide or sell unused computing power. This improves local ties and communication. Such outreach can often help fund scholarships and produce partnerships that benefit both the institution and the community.

5. Research computing enhances research instruction. It broadens the environment in which students learn, and it better prepares them for the constantly evolving infrastructure they are likely to encounter when they enter the academic or nonacademic workforce. This is true for undergraduates as well as for graduate students.

6. Research computing provides for the possibility of metavirtual computing labs (VCLs such as nanoHUB and HUBzero), which can help institutions save on physical space. Meetings, testing, and analysis can be done online rather than in a physical environment.
**Research Intensity**

To assess research intensity (RI), ECAR researchers used an equation that corresponded to the ratio of total research dollars to student FTE. Table 1 displays median and mean research intensity for each Carnegie class designation in the sample.

<table>
<thead>
<tr>
<th>Carnegie Class</th>
<th>N</th>
<th>Median ($)</th>
<th>Mean ($)</th>
<th>Std. Deviation</th>
<th>Minimum ($)</th>
<th>Maximum ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>112</td>
<td>0.00</td>
<td>1.48</td>
<td>11.16</td>
<td>0.00</td>
<td>108.88</td>
</tr>
<tr>
<td>BA Private</td>
<td>128</td>
<td>$140.90</td>
<td>$554.05</td>
<td>956.65</td>
<td>0.00</td>
<td>$4,976.32</td>
</tr>
<tr>
<td>BA Public</td>
<td>19</td>
<td>$157.14</td>
<td>$318.96</td>
<td>687.79</td>
<td>0.00</td>
<td>$3,072.06</td>
</tr>
<tr>
<td>MA Private</td>
<td>104</td>
<td>0.00</td>
<td>$392.90</td>
<td>1,030.31</td>
<td>0.00</td>
<td>$7,397.49</td>
</tr>
<tr>
<td>MA Public</td>
<td>106</td>
<td>$198.51</td>
<td>$519.71</td>
<td>750.34</td>
<td>0.00</td>
<td>$3,786.25</td>
</tr>
<tr>
<td>DR Private</td>
<td>53</td>
<td>$7,417.32</td>
<td>$14,022.25</td>
<td>20,297.05</td>
<td>0.00</td>
<td>$119,768.99</td>
</tr>
<tr>
<td>DR Public</td>
<td>110</td>
<td>$6,755.10</td>
<td>$8,012.30</td>
<td>6,646.99</td>
<td>$107.70</td>
<td>$30,965.73</td>
</tr>
<tr>
<td>Other U.S.</td>
<td>29</td>
<td>0.00</td>
<td>$1,518.72</td>
<td>5,515.23</td>
<td>0.00</td>
<td>$29,359.42</td>
</tr>
<tr>
<td>All Institutions</td>
<td>661</td>
<td>$154.72</td>
<td>$2,786.19</td>
<td>7,774.04</td>
<td>0.00</td>
<td>$119,768.99</td>
</tr>
</tbody>
</table>

For descriptive purposes, three categories of research intensity were derived. These categories were based not on the sample of CDS respondents but on IPEDS data, from which research dollar and FTE figures were obtained. Low research-intensive institutions had an RI of $0. A median split was used to derive the other RI levels. Those spending up to $439 per student FTE were classified as medium research intensity. High-RI institutions had an RI above $439. Figure 1 displays the percentage of low-, medium-, and high-RI institutions in each Carnegie class.

![Figure 1. Percentage of Low, Medium, and High Research Intensity Institutions Within Each Carnegie Class](image-url)
Figure 1 shows—not surprisingly—that two-year institutions are primarily low in research intensity. Doctoral institutions are primarily high in research intensity. These are the institutions that are spending the most research dollars for their size. More high-RI institutions are represented among private than public four-year institutions, whereas more high-RI institutions are represented among public than private master’s institutions. Among master’s institutions there is great disparity between public and private institutions in the percentage of low-RI institutions: low RI is more heavily represented in private master’s institutions. Public institutions have more medium-level research intensity than do both four-year and master’s private institutions.

**Current IT Research Computing Support**

Three-fourths (74.5%) of CDS respondents provide research technology services.\(^\text{10}\) Of these, IT is either primarily responsible or shares responsibility for research technology functions in approximately two-thirds of responding institutions (Figure 2). One-fourth (24.83%) do not provide RC services or have no unit responsible for them.
Number of Staff Dedicated to Research Computing. Of the 2012 CDS Module 1 respondents who provide RC services, the number of FTE staff (including student workers) dedicated to supporting research computing ranged from 0 to 106, with a mean of 1.78 (SE = .33) and a median of 0. The mean percentage of RC staff/Total IT staff was 1.00% (SE = .10). These figures differed by research intensity (Figure 3). High-RI institutions devote more than three times as many staff to research computing than do medium- or low-RI institutions. Interestingly, there was not much difference in the staffing provided by medium- and low-RI institutions. Since research intensity was calculated on the basis of research dollars (standardized by student FTE), this suggests that the research dollars dedicated to research computing in medium-RI institutions are generally not spent on additional staff.

Figure 3. RC Staffing at High, Medium, and Low Research Intensity Institutions

Centralized versus Local RC Support. Respondents to CDS Module 4 were asked how their institution is organized to support research computing (Figure 4). Approximately one-quarter (23%) of institutions provide an integrated set of IT services for research computing, and 11% are planning such services. Almost half (41%) of institutions support research computing through independent units, and approximately 11% provide both central (integrated) and local (independent units) support. Half (51%) stated that they provide minimal RC services. These responses differed by research intensity. High-RI institutions were much more likely to provide RC support regardless of how it was organized. Low- and medium-RI institutions were more likely to report providing minimal support. Also, almost...
10% of medium-RI institutions support research computing via independent units. Institutions providing RC support do so in a variety of ways. Support can be centralized, local, or a combination of both. However, the use of multiple support models is common only at high-RI institutions.

**Figure 4. Organization for Research Computing (May Have Multiple Responses)**

**IT’s Role in Providing Specific Research Computing Services.** CDS Module 4 contains items asking about IT’s role in providing operational and specific RC services in 22 different areas (Figure 5). More than half of IT departments provide at least shared support in the following areas:

- Videoconferencing services
- Provision of data center facilities for academic units to operate their servers
- Review or approval of technical aspects of research
- Consulting/support for storage and data access
- Data management, storage, and curation services
- High-performance network provisioning within the institution
- Management of research servers owned by academic units
- Review or approval of information security plans for research involving sensitive data
- Consulting/support for cloud services

“Institutions need to provide research computing resources if they want to be considered even a player in research.”

—Focus Group Participant
Figure 5. Research Computing Services Provided by IT
Provision of specific services differs by research intensity. Figure 6 shows the percentage of high-, medium-, and low-RI institutions providing each service. Note that provision is by any means: central, local, system-provided, or outsourced.

Figure 6. Provision of Specific Research Services, by Research Intensity
High-RI institutions are more likely to provide all research computing services than are medium-RI institutions, which are more likely to provide them than are low-RI institutions. That finding is not surprising, but this figure provides a way to compare by percentage and by RI category what institutions are providing in the way of RC services. For example, if a low-RI institution wanted to get started on providing RC services, it might be helpful to know that nearly one-third of low-RI institutions provide data center facilities for academic units to operate their servers. In addition, the major differentiator between high-RI institutions and others is in the provisioning of HPC services (including HPC consulting/support and high-throughput computing services). If medium-RI institutions want to take the next step in research computing, they might want to consider the provision of HPC.

**IT Roles in Research Grants.** The enhanced ability to collaborate and communicate effectively online in the past decade has attracted more researchers to institutions that are less research intensive. In addition, low- and medium-RI institutions are seeking to attract more researchers with grant-writing experience to provide opportunities for collaboration and for increasing grant dollars throughout the institution. IT departments in less research intensive institutions have a special opportunity to become involved in research from the ground up by helping early on to provide and promote RC services to new faculty and by making themselves available to serve on grants.

Figure 7 shows that high-RI institutions are more likely to have IT staff serving on grants, regardless of the role they fulfill on those grants. However, approximately one-fifth of low- and medium-RI institutions also have IT staff serving on grants in either a secondary or a contractual role. Low- and medium-RI institutions may be able to help support the RC services they provide by communicating with faculty about (a) the specific services available and (b) the availability of IT staff to serve on grants.
Figure 7. IT Roles in Research Grants
**Research Services to Other Institutions.** IT provides a range of research services to other institutions (Figure 8). High-RI institutions are more likely to make these services available to other institutions, but a substantial percentage of medium-RI institutions and a small percentage of low-RI institutions also provide these services. Provision of these services is often for a fee, so this may be a method for sustaining or developing RC services.

![Research Services IT Provides to Others](image-url)
Funding Models of Research Support

Research computing represents an extraordinary demand for IT services and resources. The most well-endowed and well-funded universities have an RC structure that is complex, layered, and dynamic in meeting researchers’ needs as they arise. The most difficult-to-define funding model for research computing is that of the institution that lies between the extremes of having no research computing and having research computing as an integrated priority. Many variable factors enter into a decision to establish research computing as a priority for an institution where research is expected but not mandatory. These factors are considered here and in the next sections on effectiveness and maturity.

CDS Module 1 respondents were asked which funding sources are used to support research computing (Figure 9). Almost half of the entire sample responded that funding for research computing was not applicable. Among those who reported at least one funding source, the majority reported that research computing was funded through the operating budget. Many also responded that they funded research computing through “other sources.” Some of the other sources specified included grants and fees from the provision of printing, copying, telephone, television, library, and other services.
CDS Module 4 respondents were asked to assess the status of their current funding model for research computing (Figure 10). Slightly more than half (54.3%) stated their funding model was serving them well, whereas nearly one-third (32.3%) reported their funding model was serving them poorly. For comparison, between one-quarter and one-third of respondents in the 2006 ECAR study agreed that their budget model was sustainable for maintaining and evolving services and infrastructure related to research.12
Figure 10. Assessment of Current Research Computing Funding Model

Approximately equal (but low) percentages of high-, medium-, and low-RI institutions gave their funding model the highest ranking (responding that it serves them well and is adaptable/scalable for the foreseeable future). However, a high percentage of high-RI institutions reported that their funding model is serving them poorly, although many of them are actively assessing a reformulation of that model.

Assessment of funding model is predicted by type of support organization. Three organizational factors were used in a regression analysis to predict funding model assessment: whether IT provides an integrated set of services to support research computing, whether research computing is supported by independent units, and whether minimal resources are provided in supporting research computing. Having either an organizational model that supports research computing by independent units or one that provides minimal resources for research computing is significantly associated with negative assessments of the funding model. In other words, these organizational models were more likely to result in assessments that the funding model “serves us poorly.” More satisfaction is reported with a funding model through which integrated IT services are provided for research computing.¹³
Effectiveness of Research Computing Strategies and Practices

CDS Module 4 respondents were asked to indicate the effectiveness of current RC strategies and practices (Figure 11). More than half of respondents stated that none of these strategies were in place. Of those who did employ one or more of these specific strategies, the majority stated they were at least somewhat effective. Effectiveness ratings for each strategy/practice were converted to a numeric scale ranging from 1 (no strategy or practice with no plans for one) to 8 (effective strategy or practice that is adaptable/scalable). Each of these assessments differed significantly by research intensity.

Figure 11. Mean Effectiveness Ratings of RC Strategies and Practices, by Research Intensity
High-RI institutions rated each of these strategies/practices as more effective than did medium- or low-RI institutions. It is noteworthy, however, that even high-RI institutions did not rate any of their RC practices/strategies any higher than “somewhat effective” on average.

**Relation of Funding Model Assessment and Effectiveness Assessments.** The assessment respondents provided for their funding model (serves us well/serves us poorly) was significantly correlated with the assessments of the effectiveness of all but two of the specific RC practices: planning for research computing and collaborative science approach. This indicates that adopting an appropriate funding model is key for effective RC practice in most areas. Since the CDS data indicate that there is little variance in funding models (most institutions fund research computing through the operating budget), the data are not granular enough to recommend specific funding models based on research intensity. Information from the focus groups and the subject matter experts for this study suggests that funding models should be based on input from both faculty and IT. Good communication between these groups may enhance satisfaction with both the funding model and RC practice.

Given the similarity of the effectiveness ratings for the specific RC practices and the small variation by research intensity (Figure 11), these ratings were averaged to provide a single score for RC effectiveness (again ranging from 1 to 8). This score was then used in the following two sections as a criterion in assessing the relative importance of each factor in determining overall RC effectiveness.

**RC Effectiveness Impacted by RC Support Organization.** The same three organizational factors used to predict funding model assessment were used to predict RC effectiveness: whether IT provides an integrated set of services to support research computing, whether research computing is supported by independent units, and whether minimal resources are provided in supporting research computing. The strongest predictor was the provision of an integrated set of services by IT. Support from independent units was also a significant—but weaker—predictor of RC effectiveness. Providing minimal resources was—as would be surmised—a significant negative predictor of RC effectiveness. These results indicate that the best organizational model for research computing is provision of integrated services by IT. When that is not possible, however, providing nonintegrated RC support at the unit level may also increase the effectiveness of research computing.

**Funding Sources Predict RC Effectiveness.** All funding sources—with the exception of student IT fees and system/district-level provisions—predict RC effectiveness to a certain extent. The strongest predictor of RC effectiveness is funding through operating budget appropriation. However, this finding may be spurious, as the operation budget is the primary budget source for most institutions. It appears safe to say that any funding source contributes to the effectiveness of research computing. In fact, the correlation between having any funding source (as opposed to no funding source) and RC effectiveness is .50 ($p = .00$).
Maturity

Insight into RC maturity was provided by focus group discussion, evaluation of CDS data, and input from subject matter experts. What follows is a summary of the factors that will be covered in the ECAR Research Computing Maturity Index, which will be published separately from this report.\textsuperscript{16}

Infrastructure

Maturity in infrastructure consists mainly of the hardware and software that are provided for data-intensive research. It includes bandwidth, storage capacity, processing speed, high-performance computing, and specialized software and scientific apparatuses. Maturity ranges from provision of no infrastructure to provision of a complex, layered infrastructure that meets the needs of the majority—if not all—of the research at the institution. Infrastructure at the low end may have significant and chronic gaps that interfere with data-intensive research. Infrastructure at the high end is robust, sustainable, adaptable, and available, and it is continually upgraded and updated. The difference in levels corresponds to the amount of time researchers actually spend on research rather than on dealing with infrastructure issues specifically.

IT Involvement

Institutions that are low in maturity on IT involvement have IT staff who have little awareness of research computing and its benefits. IT staff may consider research computing to be a burden, or they may be spread too thin to address research computing needs. Institutions high in maturity have dedicated research computing IT staff who understand the complex set of needs involved in research computing across a variety of disciplines. More mature IT departments are proactive rather than reactive in responding to data-intensive research needs. The amount of time researchers spend waiting for research computing needs to be addressed is a variable that plays into maturity on this factor. IT staff at institutions with more-mature RC programs are written into more research grants to provide RC services. They also provide training in new research computing technologies.
Investment

Maturity in investment ranges from providing no resources for research computing to ensuring that all data-intensive research needs are addressed. Institutions at the low end of investment maturity view research computing as an expense, whereas institutions at the high end view research computing as an investment in which benefits outweigh costs. Having no budget for research computing places an institution at the low end of the maturity scale. Institutions at the high end include research computing as a regular part of the budget, regardless of the funding model. The number of specialized staff—or the amount of time regular staff spend on research computing, in institutions of lesser research intensity—is also a variable that marks maturity on this factor.

Culture

Maturity in the area of RC culture at least partially consists of the priority placed on research computing by administration, IT staff, and researchers. Institutions low in maturity place the burden of providing research computing services squarely on researchers. High-maturity institutions provide a large number of research computing services, reducing the burden on researchers and freeing them to do more research. The existence and enforcement of policies (e.g., data security, tech transfer) involved with research computing also play a part in maturity on this factor. The amount and quality of communication between IT and researchers also separates those who are high and low on maturity. High-maturity institutions will also have faculty who have incorporated data-intensive research into the curriculum.
Conclusions

IT’s role in research computing can vary along a number of dimensions, including the breadth of the specific services provided and the depth with which IT is involved—centrally, peripherally, or not at all. The breadth and depth of IT’s involvement may also vary considerably, depending on the research intensity of the institution.

Research computing is highly fractured and perhaps necessarily so; the services and practices of research computing are dictated by the specific needs of the researchers at each institution. Researchers may be skeptical of centralizing RC services, either because they don’t trust IT to make integrated RC services work or because they don’t want to cede power over specific services obtained or foreseen. A variety of models exist for providing RC services either centrally or locally. There is a need for discussion between IT leaders and researchers to determine the best model for each institution. Institutions of high research intensity may depend more on faculty grants to provide infrastructure for research computing at the local level; however, this is not sufficient. Not all researchers are able to obtain grants, even in high-RI institutions. In disciplines or in institutions that have historically been less data-intensive and where grant dollars are not as dependable, more RC infrastructure may need to be provided centrally.

One unifying role of IT is in the enablement of research, regardless of the model decided on. Researchers tend to focus on what they need for their current project. They might not be fully aware of what they could accomplish with more resources. IT staff can show researchers what can be accomplished with better RC infrastructure and services. They can demonstrate the value of including IT staff and services in grant budgets. They can explain how research computing can open up entrepreneurial partnerships with enterprise.

Further Exploration

A number of important questions surrounding research computing could not be explored in this report.

- What are the specific concerns institutions have in initiating or expanding RC services?
- How do researchers from various disciplines (e.g., engineering, physical sciences, biomedical sciences, social sciences, humanities) need to be differentially supported with research computing resources?
- What are the differences in research computing needs among undergraduate students, graduate students, and postdoctoral students? Do they require different levels and/or models of research computing support?
Recommendations

• IT departments should leverage the research computing services and infrastructure they provide to partner with researchers on grants.

• IT should explore providing RC services to other institutions for a fee. Doing so may help fund the maintenance or development of such services. It may also help foster research collaborations and partnerships with other institutions.

• Institutions of medium research intensity should consider spending more of their research computing dollars on additional staffing. RC staffing at medium-RI institutions is no greater than at low-RI institutions. RC personnel (as opposed to other RC resources) are considered especially beneficial when it comes to faculty recruitment and faculty research performance.

• IT should be proactive in uncovering research computing needs and communicating which research computing services are available. One-on-one communication with researchers is the best method of enabling partnerships and addressing needs.

• Institutions that want to advance their research should develop an RC strategy that includes an investment-based funding approach as well as integrated services and infrastructure.
Methodology

Survey Data. Quantitative data for the 2012 Research Computing Study was obtained from Modules 1 and 4 of the 2012 Core Data Survey (CDS). Responses from 675 institutions were analyzed. A total of 427 of these institutions completed Module 4, which is optional and is the portion of the CDS that surveys research technology services. Table A summarizes CDS Module 1 respondents by Carnegie classification and institution size, and Table B provides that information for CDS Module 4 respondents. Data collection for the Core Data Survey occurred between March and August of 2012.

Table A. Summary of Respondents to CDS Module 1, by Carnegie Classification and Institution Size

<table>
<thead>
<tr>
<th>Carnegie Class</th>
<th>Total</th>
<th>FTE ≤2,000</th>
<th>2,001–4,000</th>
<th>4,001–8,000</th>
<th>8,001–15,000</th>
<th>&gt;15,000</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>112</td>
<td>18</td>
<td>37</td>
<td>35</td>
<td>14</td>
<td>8</td>
<td>16.6%</td>
</tr>
<tr>
<td>% of Total</td>
<td>2.7%</td>
<td>5.5%</td>
<td>5.2%</td>
<td>2.1%</td>
<td>1.2%</td>
<td></td>
<td>16.6%</td>
</tr>
<tr>
<td>BA Private</td>
<td>136</td>
<td>78</td>
<td>56</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>20.1%</td>
</tr>
<tr>
<td>% of Total</td>
<td>11.6%</td>
<td>8.3%</td>
<td>.3%</td>
<td>.0%</td>
<td>.0%</td>
<td></td>
<td>20.1%</td>
</tr>
<tr>
<td>BA Public</td>
<td>19</td>
<td>9</td>
<td>7</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2.8%</td>
</tr>
<tr>
<td>% of Total</td>
<td>1.3%</td>
<td>1.0%</td>
<td>.4%</td>
<td>.0%</td>
<td>.0%</td>
<td></td>
<td>2.8%</td>
</tr>
<tr>
<td>MA Private</td>
<td>105</td>
<td>20</td>
<td>48</td>
<td>32</td>
<td>4</td>
<td>1</td>
<td>15.6%</td>
</tr>
<tr>
<td>% of Total</td>
<td>3.0%</td>
<td>7.1%</td>
<td>4.7%</td>
<td>.6%</td>
<td>.1%</td>
<td></td>
<td>15.6%</td>
</tr>
<tr>
<td>MA Public</td>
<td>107</td>
<td>1</td>
<td>14</td>
<td>38</td>
<td>36</td>
<td>18</td>
<td>15.9%</td>
</tr>
<tr>
<td>% of Total</td>
<td>.1%</td>
<td>2.1%</td>
<td>5.6%</td>
<td>5.3%</td>
<td>2.7%</td>
<td></td>
<td>15.9%</td>
</tr>
<tr>
<td>DR Private</td>
<td>54</td>
<td>0</td>
<td>4</td>
<td>12</td>
<td>27</td>
<td>11</td>
<td>8.0%</td>
</tr>
<tr>
<td>% of Total</td>
<td>.0%</td>
<td>.6%</td>
<td>1.8%</td>
<td>4.0%</td>
<td>1.6%</td>
<td></td>
<td>8.0%</td>
</tr>
<tr>
<td>DR Public</td>
<td>110</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>21</td>
<td>82</td>
<td>16.3%</td>
</tr>
<tr>
<td>% of Total</td>
<td>.0%</td>
<td>.0%</td>
<td>1.0%</td>
<td>3.1%</td>
<td>12.1%</td>
<td></td>
<td>16.3%</td>
</tr>
<tr>
<td>Other U.S.</td>
<td>32</td>
<td>19</td>
<td>9</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>4.7%</td>
</tr>
<tr>
<td>% of Total</td>
<td>2.8%</td>
<td>1.3%</td>
<td>.6%</td>
<td>.0%</td>
<td>.0%</td>
<td></td>
<td>4.7%</td>
</tr>
<tr>
<td>Total</td>
<td>675</td>
<td>145</td>
<td>175</td>
<td>133</td>
<td>102</td>
<td>120</td>
<td>100.0%</td>
</tr>
<tr>
<td>% of Total</td>
<td>21.5%</td>
<td>25.9%</td>
<td>19.7%</td>
<td>15.1%</td>
<td>17.8%</td>
<td></td>
<td>100.0%</td>
</tr>
</tbody>
</table>
### Table B. Summary of Respondents to CDS Module 4, by Carnegie Classification and Institution Size

<table>
<thead>
<tr>
<th>Carnegie Class</th>
<th>AA</th>
<th>BA Private</th>
<th>MA Private</th>
<th>MA Public</th>
<th>DR Private</th>
<th>DR Public</th>
<th>Other U.S.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Count</td>
<td>% of Total</td>
<td>Count</td>
<td>% of Total</td>
<td>Count</td>
<td>% of Total</td>
<td>Count</td>
<td>% of Total</td>
</tr>
<tr>
<td>≤2,000</td>
<td>2,001–4,000</td>
<td>4,001–8,000</td>
<td>8,001–15,000</td>
<td>&gt;15,000</td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>6</td>
<td>1.4%</td>
<td>15</td>
<td>3.5%</td>
<td>17</td>
<td>4.0%</td>
<td>5</td>
<td>1.2%</td>
</tr>
<tr>
<td>BA Private</td>
<td>41</td>
<td>9.6%</td>
<td>37</td>
<td>8.7%</td>
<td>0</td>
<td>.0%</td>
<td>0</td>
<td>.0%</td>
</tr>
<tr>
<td>MA Private</td>
<td>5</td>
<td>1.2%</td>
<td>3</td>
<td>.7%</td>
<td>3</td>
<td>.7%</td>
<td>0</td>
<td>.0%</td>
</tr>
<tr>
<td>MA Public</td>
<td>13</td>
<td>3.0%</td>
<td>28</td>
<td>6.6%</td>
<td>13</td>
<td>3.0%</td>
<td>3</td>
<td>.7%</td>
</tr>
<tr>
<td>DR Private</td>
<td>1</td>
<td>.2%</td>
<td>9</td>
<td>2.1%</td>
<td>28</td>
<td>6.6%</td>
<td>28</td>
<td>6.6%</td>
</tr>
<tr>
<td>DR Public</td>
<td>0</td>
<td>.0%</td>
<td>4</td>
<td>.9%</td>
<td>10</td>
<td>2.3%</td>
<td>23</td>
<td>5.4%</td>
</tr>
<tr>
<td>Other U.S.</td>
<td>10</td>
<td>2.3%</td>
<td>5</td>
<td>1.4%</td>
<td>3</td>
<td>1.4%</td>
<td>0</td>
<td>.0%</td>
</tr>
<tr>
<td>Total</td>
<td>76</td>
<td>17.8%</td>
<td>101</td>
<td>23.7%</td>
<td>80</td>
<td>18.7%</td>
<td>77</td>
<td>18.0%</td>
</tr>
</tbody>
</table>

### Table C. Summary of Questions from the CDS Analyzed for This Study

<table>
<thead>
<tr>
<th>Module</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9. Which organizational units are responsible for [research technology services]? Choose one: primarily central IT; shared between central IT and other units; primarily other units; primarily outsourced; N/A.</td>
</tr>
<tr>
<td>1</td>
<td>20. Which funding sources are being used to support [RC]? Check all that apply: operating budget appropriation; capital budget appropriation; student IT fee; cost recovery; resale of products; other sources; provided at system/district level; N/A.</td>
</tr>
<tr>
<td>1</td>
<td>22. How many FTE staff and students are employed in central IT [for research technology services]?</td>
</tr>
</tbody>
</table>

*Cont'd*
<table>
<thead>
<tr>
<th>Module</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1. How is your institution organized to support research that is dependent on IT? Check all that apply: provides an integrated set of IT services; planning an integrated set of IT services; independent units support; minimal resources provided for RC; other.</td>
</tr>
<tr>
<td></td>
<td>2. What IT-related <em>consulting and support</em> services are offered to researchers, and which units provide these services? Check one: primarily central IT; shared between central IT and other units; primarily system or district office; primarily outsourced to non-commercial; primarily outsourced to commercial; N/A.</td>
</tr>
</tbody>
</table>
|        | a. assistance in preparing grant applications  
b. review or approval of information security plans for research involving sensitive data  
c. review or approval of NSF-required data management plans  
d. review or approval of other technical aspects of research  
e. statistical consulting  
f. consulting/support for software development and porting  
g. consulting/support for visualization  
h. consulting/support for HPC  
i. consulting/support for storage and data access  
j. consulting/support for access to federally funded research resources  
k. consulting/support for cloud services  
l. hosting or facilitating events to introduce new technologies |
|        | 3. What *operational* IT services are available to researchers, and which units provide the services? Check one: primarily central IT; shared between central IT and other units; primarily system or district office; primarily outsourced to non-commercial; primarily outsourced to commercial; N/A. |
|        | a. management of research servers owned by academic units  
b. provision of data center facilities for academic units to operate their servers  
c. HPC services  
d. high-throughput computing services  
e. institutional grid computing services  
f. data management, storage, and curation services  
g. videoconferencing services  
h. high-performance network provisioning within the institution  
i. high-performance external network provisioning  
j. an integrated system of research-related services  
k. access to specialized scientific apparatus |
|        | 5. Which statement best describes the status of the current funding model? Choose one: serves us well and is adaptable/scalable for the foreseeable future; serves us well, but needs future revision or overhaul; serves us poorly, and we are actively assessing a reformulation; serves us poorly, and we are not actively assessing a reformulation; other. |
|        | 6. What role does central IT have in your institution’s grant awards? Check all that apply: faculty PI, IT co-PI; faculty PI, IT SI; IT PI, IT PI, faculty co-PI; IT PI, faculty SI; IT contracted; other; N/A. |
|        | 7. Which of the following services does your institution make available to researchers at other institutions? Check all that apply: HPC; storage resources; visualization resources; server management; server hosting; access to specialized apparatus; other; N/A. |
|        | 11. Please indicate the status of the following RC practices. Check one: no strategy or practice; ineffective strategy or practice; somewhat effective strategy or practice; effective strategy or practice; N/A. |
|        | a. data management for RC  
b. central IT staffing for RC  
c. central IT support for RC  
d. funding model for RC  
e. planning for RC  
f. collaborative science approach |
**Focus Groups.** In addition to the CDS data used, data were collected from 14 individuals in online focus groups conducted in September 2012. Participants were selected from a sample of EDUCAUSE primary representatives. The sample was stratified by Carnegie class. Each representative was asked to participate in the focus group or to select a proxy from his or her IT department to participate. Representatives were also asked to solicit the participation of a faculty member at their institution who was involved in research computing. In total, 9 IT professionals and 5 faculty members from 10 different institutions participated in 5 focus groups. Institutions from both research-intensive and non-research-intensive institutions were represented.

An e-mail was sent inviting each participant to sign up for one of the online focus groups. The e-mail also described the purpose of each focus group and contained an attachment with the questions that would be asked (Table D). Instructions for participating online via Vidyo conferencing software were also provided. To maintain anonymity, this study does not identify the participating institutions. An honorarium of $100 was provided to participants.

**Table D. Focus Group Questions**

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>When you think of “research computing,” what comes to mind? What would you include in a definition of research computing?</td>
</tr>
<tr>
<td>Thinking about your institution, how does (or how would) research computing impact faculty recruitment, performance, and retention? Are there other benefits of research computing? Do the benefits outweigh the costs for your institution?</td>
</tr>
<tr>
<td>Describe the ideal set of IT services and staff roles to support research at your institution.</td>
</tr>
<tr>
<td>Let’s say we developed a maturity index on a scale of 1 to 5, where 1 means an institution has no or a barely minimal research computing program and 5 means an institution is one of the most advanced research computing institutions in the country. Let’s say you were in charge of assessing institutions on this scale. What would you look for or what would need to be in place to move an institution up the scale? [concrete indicators]</td>
</tr>
</tbody>
</table>
Acknowledgments

This study resulted from the contributions of several individuals. The subject matter experts on this study, Curt Hillegas (Director of Research Computing, Princeton University), Brian Voss (Vice President and CIO, University of Maryland), and Michele Norin (CIO, University of Arizona), added helpful comments and suggestions for this report. Kathryn Northcut (Missouri University of Science and Technology) provided help with the focus group analyses. My EDUCAUSE colleagues Susan Grajek, Eden Dahlstrom, Pam Arroway, and Gregory Dobbin contributed many insightful suggestions, additions, and changes to this report, and Tyson Anderson created the graphics.
Appendix

A regression table for each regression analysis performed is presented in this section. Beta coefficients indicate the predictive strength of each factor.

### Coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Zero-order</td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td>t</td>
<td>Sig.</td>
</tr>
<tr>
<td></td>
<td>2.976</td>
<td>.141</td>
<td>21.137</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Integrated IT services provided.</td>
<td>.155</td>
<td>.141</td>
<td>.066</td>
<td>1.103</td>
</tr>
<tr>
<td></td>
<td>Supported by independent units.</td>
<td>−.372</td>
<td>.131</td>
<td>−.178</td>
<td>−2.834</td>
</tr>
<tr>
<td></td>
<td>Minimal resources provided.</td>
<td>−.343</td>
<td>.146</td>
<td>−.164</td>
<td>−2.342</td>
</tr>
</tbody>
</table>

Dependent variable: Funding model assessment
R² = .04

### Coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Zero-order</td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td>t</td>
<td>Sig.</td>
</tr>
<tr>
<td></td>
<td>2.392</td>
<td>.189</td>
<td>12.630</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Integrated IT services provided.</td>
<td>2.017</td>
<td>.204</td>
<td>.424</td>
<td>9.868</td>
</tr>
<tr>
<td></td>
<td>Supported by independent units.</td>
<td>.841</td>
<td>.185</td>
<td>.206</td>
<td>4.538</td>
</tr>
<tr>
<td></td>
<td>Minimal resources provided.</td>
<td>−.827</td>
<td>.202</td>
<td>−.206</td>
<td>−4.098</td>
</tr>
</tbody>
</table>

Dependent variable: RC effectiveness
R² = .39
## Coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
<td>Zero-order</td>
</tr>
<tr>
<td>1 (Constant)</td>
<td>1.653</td>
<td>.125</td>
<td></td>
<td>13.194</td>
<td>.000</td>
</tr>
<tr>
<td>Operating budget appropriation.</td>
<td>1.496</td>
<td>.173</td>
<td>.373</td>
<td>8.646</td>
<td>.000</td>
</tr>
<tr>
<td>Capital budget appropriation</td>
<td>.570</td>
<td>.262</td>
<td>.094</td>
<td>2.173</td>
<td>.030</td>
</tr>
<tr>
<td>Student IT fee</td>
<td>−.478</td>
<td>.355</td>
<td>−.057</td>
<td>−1.346</td>
<td>.179</td>
</tr>
<tr>
<td>Cost recovery (chargeback)</td>
<td>1.438</td>
<td>.320</td>
<td>.192</td>
<td>4.501</td>
<td>.000</td>
</tr>
<tr>
<td>Resale of products</td>
<td>2.043</td>
<td>.867</td>
<td>.098</td>
<td>2.357</td>
<td>.019</td>
</tr>
<tr>
<td>Other sources</td>
<td>1.108</td>
<td>.213</td>
<td>.214</td>
<td>5.196</td>
<td>.000</td>
</tr>
<tr>
<td>Provided at system/district level</td>
<td>.792</td>
<td>.762</td>
<td>.043</td>
<td>1.039</td>
<td>.299</td>
</tr>
</tbody>
</table>

Dependent variable: RC effectiveness

$R^2 = .30$
Notes

1. “ECAR Research Computing Maturity Index” (Maturity Index; Louisville, CO: EDUCAUSE Center for Applied Research, forthcoming 2012).


5. CDS Module 4 questions were developed in collaboration with the Research University CIO Conclave.

6. Because Carnegie class was a factor in many analyses for this report, institutions without a Carnegie class designation were excluded from analysis. In addition, institutions classified as medical schools were excluded because (a) they totaled five in number and (b) their research funding data was highly variable, which skewed the interpretation of some analyses.


8. Ibid.

9. This formula for research intensity was used to hold institutional size constant.

10. The phrase “research technology” rather than “research computing” was used in the CDS phrasing for this question.


12. Precise comparisons cannot be made, since the questions asked in the 2006 and 2012 surveys differed significantly in their phrasing.

13. See the Appendix for the regression analysis table, including beta coefficients.

14. Ibid.

15. Ibid.


17. See http://www.educause.edu/coredata.