

# A PetaFLOPS Supercomputer as a Campus Resource: Innovation, Impact, and Models for Locally-Owned High Performance Computing at Research Colleges and Universities

Abhinav Thota  
Pervasive Technology Institute  
Indiana University  
Bloomington, IN, USA  
athota@iu.edu

Robert Henschel  
Pervasive Technology Institute  
Indiana University  
Bloomington, IN, USA  
henschel@iu.edu

Jenett Tillotson  
Pervasive Technology Institute  
Indiana University  
Bloomington, IN, USA  
jtillots@iu.edu

Ben Fulton  
Pervasive Technology Institute  
Indiana University  
Bloomington, IN, USA  
befulton@iu.edu

David Hancock  
Pervasive Technology Institute  
Indiana University  
Bloomington, IN, USA  
dyhancoc@iu.edu

Matt Link  
Pervasive Technology Institute  
Indiana University  
Bloomington, IN, USA  
mrlink@iu.edu

Le Mai Weakley  
Pervasive Technology Institute  
Indiana University  
Bloomington, IN, USA  
llnguyen.edu

Matt Allen  
Pervasive Technology Institute  
Indiana University  
Bloomington, IN, USA  
malallen@iu.edu

Craig A. Stewart  
Pervasive Technology Institute  
Indiana University  
Bloomington, IN, USA  
stewart@iu.edu

## ABSTRACT

In 1997, Indiana University (IU) began a purposeful and steady drive to expand the use of supercomputers and what we now call cyberinfrastructure. In 2001, IU implemented the first 1 TFLOPS supercomputer owned by and operated for a single US University. In 2013, IU made an analogous investment and achievement at the 1 PFLOPS level: Big Red II, a Cray XE6/XK7, was the first supercomputer capable of 1 PFLOPS (theoretical) performance that was a dedicated university resource [2]. IU's high performance computing (HPC) resources have fostered innovation in disciplines from biology to chemistry to medicine. Currently, 185 disciplines and sub disciplines are represented on Big Red II with a wide variety of usage needs. Quantitative data suggest that investment in this supercomputer has been a good value to IU in terms of academic achievement and federal grant income. Here we will discuss how investment in Big Red II has benefited IU, and argue that locally-owned computational resources (scaled appropriately to needs and budgets) may be of benefit to many colleges and universities. We will also discuss software tools under development that will aid others in quantifying the benefit of investment in high performance computing to their campuses.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

SIGUCCS '16, November 06-09, 2016, Denver, CO, USA

© 2016 ACM. ISBN 978-1-4503-4095-3/16/11...\$15.00

DOI: <http://dx.doi.org/10.1145/2974927.2974956>

## CCS Concepts

•**General and reference** → *General conference proceedings*; •**Social and professional topics** → *Professional topics*; *Management of computing and information systems*; *Computing education*; •**Applied computing** → *Physical sciences and engineering*; *Life and medical sciences*; *Arts and humanities*;

## Keywords

supercomputing; HPC; university; innovation; petaflops, value, analytics, impact

## 1. INTRODUCTION

In 1997, Indiana University (IU) began a purposeful and steady drive to expand the use of supercomputers and what we now call cyberinfrastructure. This was one part of the IU Information technology organization response to a directive and challenge from then President Myles Brand for IU to be a “leader in absolute terms in the use and application of IT.” In 2001, IU implemented the first 1 TFLOPS supercomputer owned by and operated for a single US University. In 2013, IU made a similar achievement at the 1 PFLOPS level: Big Red II, a Cray XE6/XK7, was the first supercomputer capable of 1 PFLOPS (theoretical) performance owned by a US university and funded exclusively with university funds. This means that IU was free to use this system entirely based on university priorities to best further the mission of IU. (There were of course prior to this other PFLOPS+ systems located at other universities – but these were funded in significant part with federal funds and were allocated in part as national resources, with federal guidance on how these systems were to be used and allocated). IU's

strategy in high performance computing (HPC) has been to have locally funded resources available to the IU research community to enhance the research and creative activities of the IU community under what we refer to as a principle of abundance in which we make resources available to the research community as a whole, without application processes, with usage allocated on a fair share basis. To a first order approximation, the more people want to compute, the more they can compute.

We previously reported on the resources, support, and impact of HPC resources at IU [34–36]. Big Red II was purchased at a time of fiscal and political uncertainty. As a result, we felt particularly compelled to make strong commitments about the value IU would reap from the acquisition and use of this system. Thus, as part of our justification for funding of the system by the university, we made a commitment that the system would be used by and useful to researchers, clinicians, scholars, and artists representing at least 150 disciplines and sub-disciplines at IU. In this paper, we will describe Big Red II, the system and human resources that support and enable operation of Big Red II, the usage of Big Red II by the IU community, and some of the lessons we have learned in the implementation and use of Big Red II. Most importantly, we argue that regardless of the size of a local HPC resource, a HPC resource owned by and operated in keeping with the research priorities of a college or university can be an important asset to research, creativity, and scholarly accomplishment.

IU’s implementation of advanced information technology services has been guided by faculty input since the very first days of the Research Computing Center at Indiana University. The first director of this center was noted astronomer Marshall C. Wrubel. More recently, the IU Chief Information Officer has been a faculty member and has received guidance from various faculty committees to develop strategic plans for IT for the entire university. Examples include the strategic architectural plan for IT, which was approved by the Trustees of Indiana University [11], and a 2005 report from a blue ribbon panel called the “Indiana University Cyberinfrastructure Research Taskforce” [9]. That report identified as a priority for IU that the central IT organization provide “education and training that is suitable to the particular needs of individuals in particular areas of research, clinical, engineering, and artistic pursuit”. The implementation of Big Red II was one aspect of the central IT organization’s specific responses to this taskforce report.

We will also discuss how, in our experience, it is beneficial for colleges and universities to have local advanced HPC resources or supercomputers with human resources to support the users. We will also discuss some tools that can be used by others to justify investments in cyberinfrastructure.

## 2. BIG RED II AND SUPPORTING SYSTEM AND HUMAN RESOURCES

By supercomputing we mean very large scale parallel computing systems with a low-latency internal network (and recognize that the old saw “if it costs more than \$1M, its probably a supercomputer” has a lot of merit). By high performance computing (HPC) we mean any sort of integrated parallel computing system, including very small computer

clusters that may have an internal network with modest performance characteristics.

### 2.1 System Resources

Big Red II [1] is a Cray XE6/XK7 supercomputer (Figure 1) with a total of 1020 compute nodes: 676 CPU/GPU compute nodes, each containing one CPU, one NVIDIA Tesla K20 GPU accelerator with a single Kepler GK110 GPU and 344 dual-CPU nodes. Having both CPUs and GPUs as part of the architecture allows researchers and students at IU to be able to choose the kind of compute resources that are best for their software to achieve the best performance, thus catering to both CPU and GPU computing users.

Big Red II is the flagship supercomputer of the university; but there are other HPC machines at IU. They are Karst (a high throughput and serial jobs cluster) and Mason (which is designed for high memory jobs). In addition, IU also runs Jetstream [31, 33] and Wrangler [17] that are part of the XSEDE [38] national cyberinfrastructure. These two systems are available to researchers across the country through an allocation system. In addition to the compute resources, we have storage systems that make it possible to use the supercomputers at scale. HPC users produce large amounts of data during their runs that require a huge number of I/O operations. The Data Capacitor II [6] filesystem is 6 PB parallel file system that can support this kind of usage. There is a home file system for more permanent storage for data that is used day to day and a 15 PB tape archive for archival needs.



Figure 1: Big Red II in the IU data center.

### 2.2 Organizational Structure

The central information technology services for IU are provided under the leadership of the Office of Vice President for Information Technology and Chief Information Officer. The areas that UITs supports include Enterprise Systems, Learning Technologies, Client Services, Networks, Clinical Affairs, and Research Technologies (RT). Research Technologies is also a core component of the Indiana University Pervasive Technology Institute (IUPTI) [15], a collaborative organization that encompasses leadership and staff from IU’s School of Informatics, Maurer School of Law, and College of Arts and Sciences. Within IUPTI, we integrate the process of creating, hardening, delivering, and supporting new information technology and cyberinfrastructure services. IUPTI has received funding from various sources – primarily US federal science agencies. The fact that Research Technologies and IUPTI report administratively to the CIO (with strong collaboration with academic units at our university) means that advanced cyberinfrastructure support – including HPC systems – is well integrated with and leverages all of the core services of the central IT organization.

### 2.2.1 Pervasive Technology Institute

IUPTI has two types of affiliated centers: Research Centers, and Service and Cyberinfrastructure Centers. Research Technologies (RT) is a Service and Cyberinfrastructure (CI) Center. RT provides comprehensive HPC services for IU. As such, RT runs and manages the supercomputers and offers consulting and support services to users.

## 2.3 Human Resources

As stated previously, RT offers robust support services for HPC users at IU. There is an online resource called the IU Knowledge Base (KB) [12] that provides answers to hundreds of questions that a user of our HPC systems might have. This resource has been in place for decades, and when it has the answer to a user's question, is a more reliable way to deliver the proper command syntax than reading a command over the phone. More information about the value, cost-effectiveness, and utility of the Knowledge Base at IU is available [29]. While IU uses a system developed at IU (now in its 3rd generation), there are now many tools available and accessible for implementation of a KB at colleges and universities of all sizes. Examples include the hosted service offered by the University of Wisconsin-Madison [21] and the KB service built on Confluence [4] used by the University at Albany[20]. At IU, there are several teams of staff who support users, including:

- A team that does basic application support and also long term extended support for complex issues, including helping with performance tuning.
- A specialized team that does support for data statistical, data analytics, and mathematical applications.
- A visualization team that can assist with application of visual technologies and visualizing complex data.
- A team that focuses promoting and supporting use of supercomputers digital humanities team assists users with using supercomputers in humanities research
- A team that can help the computing center *and* users with automating workflows through convenient web-based workflow systems called science gateways [28].

As we will demonstrate with data in the following sections, and as we have described elsewhere, our experience is that a strong support infrastructure – people – is important to promote and enable effective local use of hardware resources [27]. And even for researchers who need hardware not available locally, expert support can help those researchers obtain allocations on federally funded CI systems like XSEDE or INCITE [10] for example. XSEDE has a Campus Champions [3] program which is a way to keep specific appropriate people at a university informed about the resources offered through XSEDE and the allocation policies that are being followed.

## 3. IMPLEMENTING BIG RED II AS A RESOURCE IMPORTANT TO MANY DISCIPLINES AT IU

### 3.1 Dedication and Early User Phase

In our experience, while events through out the academic year are important for bringing in new users and for keeping existing users updated, the time when a machine is first introduced is critical. We found that there need to be specific outreach efforts at probable users of the system and bringing them in during the early user phase. This way the system can go through its motions and the administrators can even out any issues before the production day. This period usually lasts anywhere from a few months to up to an year for experimental systems.

In addition to having an early user phase, we found that it was valuable to organize a dedication ceremony that can have an impact on the entire university audience. We made sure that the ceremony stressed the importance of the new system to the university community and made everyone notice it. This is a one time opportunity that does not come around again until the next new machine is purchased and dedicated to the university.

### 3.2 Usage

The operating system on Big Red II is the Cray Linux Environment [5] (based on SUSE Linux SLES 11). Users of Big Red II have access to their home directories with a 100 GB quota and a high speed scratch space that is for temporary storage of research data. Access to the system is through any SSH [16] client that users can run on their desktops or laptops. Users connect to one of two login nodes for Big Red II that they are directed to in a round-robin fashion. Login nodes are intended for light interactive tasks such as setting up the (shell) environment by exporting the right variables for use by various applications, compiling applications, or setting up input and output directories for the jobs that run on the compute nodes.

Given the shared nature of the login nodes, all of the computational work that users do happens on the compute nodes. There are 1020 compute nodes on Big Red II, but at any point in time, there is more demand than can be satisfied. These nodes are therefore allocated through a scheduling system. Big Red II uses the TORQUE resource manager [30] and the Moab Workload Manager [13] to manage and schedule jobs. Big Red II uses a fair share scheduling algorithm to set job priorities for users. Administrators set a usage goal for each user and when a user exceeds that goal, that user's jobs are given a lower scheduling priority [8]. There are multiple queues on Big Red II setup for different kinds of usage. There are separate queues for GPU and CPU applications and separate debug queues for people running quick jobs that are in testing phase. Users submit scripts using the "qsub" command and these scripts specify the job requirements like the amount of compute time needed, the number of compute nodes required and the type of compute nodes is specified through the queue type selected.

We strongly encourage and assist users in doing a benchmarking and scaling study of their code(s) to ensure that the applications are being run with the most efficient number of compute nodes. This is especially true for users running large numbers of jobs, as even a few percentage points im-

provement in performance can have a big impact over a few thousand jobs.

## 4. EDUCATION, OUTREACH AND SUPPORT

There are many very positive aspects of supporting students who are “digital natives” in terms of IT support in general. However, to the current generation of students who grow up using touchscreen devices that are highly intuitive, supercomputers are not what they expect. The vast majority of supercomputers run some version of Linux and are still almost always accessed through text terminals with a command line interface. Given this less than user-friendly computing environment, we found that it is important to have a strong education, outreach and support structure available for the users.

### 4.1 Outreach and Training

In this section we will describe all the activities that we think are important in making HPC accessible to the students and faculty on a university campus. If you consider a large public university, thousands of new students and tens of new faculty members come in every year. On top of this, computing hardware gets upgraded every three to five years and software changes happen even more frequently. To develop a new, larger user community and keep the existing HPC user community informed about the available resources, we host more than a dozen outreach events on campus every year. This includes presentations at departmental meetings, workshops for beginner and advanced users, data center tours, on site support for classes making use of HPC resource.

When Big Red II was first dedicated in 2013, we did a pre-launch workshop for the machine’s launch that drew more than 100 people from the campus. We also did an introduction to Big Red II workshop after the machine went into production. Our goal was to get to 150 disciplines using the system, which is not possible with just the traditional HPC user departments like physics and chemistry. We also held information sessions for non-traditional user departments on campus. During the introduction to Big Red II workshops, we noticed that many attendees are not familiar with Linux. Many new comers to the HPC field do not usually have experience with command line terminals. IU started offering an introductory Linux class that is co-located with the introductory HPC workshop to address this problem. We also offer more focused workshops for specific research groups and departments. Since Big Red II was dedicated in the spring of 2013, we held 83 education, outreach, and training events at IU to promote its use. We reached over 2,000 faculty, staff, and students from across all the IU campuses. Coupled with over twenty news releases aimed at drawing attention to the capabilities of Big Red II, this outreach helped us immensely in achieving our goals. Even with these extensive outreach and training opportunities, there are more users who could benefit from our services.

### 4.2 Big Red II User Base

One of the critical challenges for us was to measure, track progress against, and then verify that we had achieved the goal of having researchers and students representing at least 150 disciplines using Big Red II. We began collecting these

Karst supports research, development, scholarship, and creative activities at IU. So that we may better understand its value to you, we need to know how you use Karst. We ask you to do the following:

\* Once a year, send to [researchtechnologies@iu.edu](mailto:researchtechnologies@iu.edu) a list of citations to any intellectual property (e.g., journal articles, reports, presentations) you developed as a result of your use of Karst.

\* Acknowledge your use of IU’s cyberinfrastructure in your published papers (see <https://kb.iu.edu/d/anwt#II> for details).

\* Indicate up to three disciplines and sub-disciplines below that best represent your activities.

(Required values denoted by a \*)

\* I have read and understand my responsibilities with respect to ePHI on Karst. [1]

\* I agree to provide an annual list of my publications based on work done with Karst.

\* I agree to cite the use of Karst and other university resources in my publications. [2]

\* Country of citizenship:

\* Primary Discipline:

\* Primary Sub-discipline:

Secondary Discipline:

Secondary Sub-discipline:

Tertiary Discipline:

Tertiary Sub-discipline:

**Figure 2: Users are asked indicate their discipline and sub-disciplines, if applicable, during the account creation process using this form. The form shown here is for Karst, but the Big Red II form is essentially identical.**

data as we set up the signup system for people to get new accounts on Big Red II. As accounts were requested, users indicated the disciplines and sub-disciplines appropriate to describe their research. An image of the form people see to indicated disciplines is shown in Figure 2. A full listing of the disciplines and sub-disciplines is available online at [7].

### 4.3 Evaluation of the System

One of the primary means by which we evaluate all of the IT services at IU is through an annual survey of user satisfaction. The survey and methodology have been described previously [29]. In short, we contract with an independent survey organization within IU to do a stratified randomized survey sampling undergraduate students, graduate students, staff, and faculty. We measure usage (the percentage of the user community that indicate that they make use of Big Red II and other supercomputers), satisfaction scores (on a standard 5 point Likert scale), and what we refer to as a satisfaction percentage – the percentage of respondents who answer a score of 3 or higher where 5 is “extremely satisfied”).

The survey asks about HPC systems generally, so the results include Big Red II plus our other high performance computing systems, but these results and user comments in the survey tend to focus on the flagship system at any given time – which has been Big Red II since it was put into service. We note that the percentage of people who indicate that they use Big Red II is actually higher than the percentage of people who have accounts and run jobs on the system. Our interpretation (which we have confirmed through interviews) is that this discrepancy comes from group leaders who “use” Big Red II in the sense of making use of it as a tool

Supercomputers and HPC	Average rating (out of 5)	Satisfaction	Usage
2016	4.09	95.32%	16.68%
2015	4.22	96.71%	13.64%

**Table 1:** The table shows the average user ratings of supercomputers and HPC at IU across all campuses for 2015 and 2016. The survey takers are considered to be satisfied if they give at least a 3 rating to the services. The percentage of respondents who said they used the service is displayed under the Usage column.

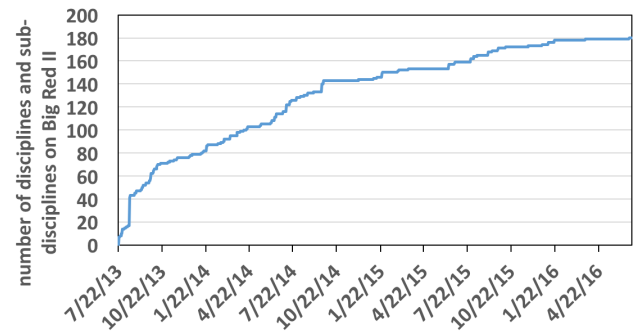


**Figure 3:** Disciplines available for user selection on Big Red II. The Arts (Fine and Performing) section is expanded to show the five sub-disciplines that are represented on Big Red II.

in their labs, but delegate the actual use of the computer and do not personally have accounts on the system. Data for usage and satisfaction of IU high performance computing systems and supercomputers for the last two years are displayed in Table 1. The community surveyed is what we traditionally think of as the research community – faculty, staff, and graduate students. Full details are available online at [19], and contains yearly survey information from the last 20+ years.

## 5. VALUE ASSESSMENT OF HPC INVESTMENTS AT IU

In 2005, the Indiana University Cyberinfrastructure Research Taskforce identified “Providing education and training that is suitable to the particular needs of individuals in particular areas of research, clinical, engineering, and artistic pursuit” [24] as a useful approach to accelerating the use of cyberinfrastructure. To further that goal, Indiana Uni-



**Figure 4:** Number of distinct disciplines and sub-disciplines of users of Big Red II (self-reported) FY 2013-2016

versity promised at the time of dedication of Big Red II to “have such a breadth of impact that Big Red II would matter to at least 150 disciplines and sub-disciplines at IU”, with a special focus on biological and biomedical disciplines, humanities, and the arts [32]. In order to fulfill this promise, administrators needed to better understand what research was actually being performed on the machine. We thus asked users, at the time of account creation request, to self-select up to three disciplines from a total of 381 disciplines. Not restricted to the science disciplines typical of supercomputer users, the disciplines included the sub-disciplines of Fine Arts, Humanities, and Sport Science as well as life science, physics, and informatics (Fig. 3). By the end of FY 2014, IU researchers representing a total of 144 disciplines and sub-disciplines were using Big Red II, and by the end of FY 2016 the number had increased to 180 (Fig. 4).

Another way we have assessed the value of Big Red II (and other parts of IU’s advanced cyberinfrastructure) is through interviews [14]. We contracted with an assessment group at another university to conduct interviews of faculty members who make extensive use of IU’s supercomputers and HPC systems. This report is online and contains a number of anecdotes and analyses about the value of supercomputers such as Big Red II [14].

### 5.1 XDMoD Value Analytics

XDMoD (XD Metrics on Demand) [22] is an NSF-funded open source tool designed to audit and facilitate the utilization of the XSEDE cyberinfrastructure by providing a wide range of metrics on XSEDE resources, including resource utilization, resource performance, and impact on scholarship and research. We are developing novel modules to be added to the existing CI metrics tool to enable assessment of the value of investment in campus-based CI in scientific terms (number of publications) and in financial terms (grant income from researchers who use campus CI as compared to those who do not). IU has over the past several years been developing a set of tools that links financial information – such as grant awards to IU faculty members – with usage of our supercomputers and HPC systems. This allows quantification of financial income to the university in the form of grants and contracts with usage of our supercomputer and HPC systems. IU is now collaborating with The University at Buffalo to add these capabilities to Open XDMoD, a widely used software tool that enables analysis of usage of HPC systems and supercomputers.



XDMoD is already straightforward to install and operate. XDMoD\_VA [23] will allow cyberinfrastructure centers and IT organizations to quantify the scientific and financial value of investments in HPC systems and supercomputers. XDMoD\_VA is being developed so that it can be implemented with or without direct connections to a university or college’s local financial systems. Where that is permitted by policy and practice XDMoD\_VA will enable analysis of all sorts of grants and contracts received by a particular institution. If it is not possible for a college or university’s IT organization to have direct read access from the institution’s financial systems then XDMoD\_VA will provide the capability to download NSF and NIH grant awards from the NIH and NSF grant data and perform an analysis against of data using these two federal funding agencies. We recognize that it may be relatively common that institutional policies restrict access to internal financial management systems, and this capability will enable institutions to work with data from the NIH and NSF which are in many cases the most significant sources of grant income for an academic institution.

## 5.2 Return on Investment

	All IU grants	IU College of Arts and Science grants
FY 2014	\$11.7 million	\$6.5 million
FY 2015	\$24.5 million	\$11.2 million
FY 2016	\$39.8 million	\$14.0 million

**Table 2: The table shows the grant income to PIs/Co-PIs with accounts on Big Red II university-wide at IU and also specifically in the College of Arts and Science (COAS).**

Existing studies of return on investment (ROI) in campus CI show that a steady and significant investment in high performance computing is very likely to lead to increases in publications and grant income [25, 26, 37]. Big Red II exemplifies this, having made its debut in 2013 on the Top 500 [18] most powerful computer systems list at #46. Table 2 shows the IU grant income for FY ’14, ’15 and ’16 according to PI/Co-PI team use of Big Red II. The grant income is separated for the College of Arts and Science given that it is a much more relatable department that is present in many, if not most, institutions of higher education.

Big Red II, personnel, and support will have a total average cost, over the expected 5 year lifespan of the system, of something lower than \$15 million dollars total. At roughly the halfway point of the life of Big Red II, grants to IU researchers that use this system total just under \$40 M. A rough projection might be that over the lifespan of Big Red II the total grant awards brought to IU by PI/ Co-PI teams that use Big Red II will total \$90 M. Of that, roughly one third might come to IU budgeted in grant awards as roughly \$30M in facilities and administration funds. That’s roughly \$6M per year in facilities and administration monies to IU when IU’s investment per year in Big Red II is \$3M per year. In other words, the facilities and administration funds income to IU that comes along with grant awards to people who use Big Red II is twice what Big Red II costs to op-

erate. Factor in the value of increased competitiveness for grant funds overall and impact on total grant income, and the scientific value of the research done with Big Red II, and all together one can make a qualitative but seemingly reasonable argument that Big Red II is a reasonable investment for the University. There are qualitative parts to this argument, and some of it depends implicitly on value judgments (such as “there is significant value to research done with supercomputers that could not be done without them”). But it’s an argument reasonably supported by data and an argument that constitutes a reasonable start for facts-based discussions about investment priorities within a college or university. The ability to have the data that enable such discussions is overall the most important aspect of the data collection activities in which we have engaged and the capabilities that XDMoD\_VA will put into the hands of many colleges and universities.

## 5.3 A Scalable Model for HPC Investment

Many institutions may be able to invest at some level in a local flagship HPC or supercomputer system, but perhaps not one with capacity and capability to meet all local computational needs. IU for example cannot meet all of the local demand for computer resources. Or, some institutions may not be able to invest in a central HPC flagship system at all. In these cases, investment in personnel to enable use of federally-funded HPC and supercomputer systems can be of great benefit to a university or college of most any size.

The basic components of HPC system are compute, a networked file system that is backed up and parallel scratch file system that is not backed up. The size of this system can be flexible and it can be configured to be built up to increase capacity later. The number of employees that are needed to effectively support this system depends on what kind of uptime and response time is expected. This is a difficult number to define as much depends on the usage model of hardware and the service expectations. From what we have observed in the HPC field in academia, about three full time employees could run the hardware described here, but more would be needed to do user support and outreach. This also does not address outages and issues that happen during non-business hours.

If this is not a possibility, having a few people on campus who are interested in HPC and are placed in a position within a department or organization that does IT support be XSEDE Campus Champions [3] is a great way to get inside information on the various HPC resources that are available as part of the national HPC infrastructure. Campus Champions get sample allocations on all the XSEDE resources, which makes it possible to quickly provide access to people on campus who are considering getting an account on one of those resources. Campus Champions can also be the de facto HPC user support person on the campus for both local and XSEDE resources.

For teaching colleges that do not have research as part of their organizational mission, having a small HPC system locally that can be used for teaching purposes or having a Campus Champion on campus that can get training accounts on XSEDE resources might make sense. Moreover, owning local supercomputing and HPC infrastructure does not preclude the institution from foregoing access to the national cyberinfrastructure that is available to researchers through XSEDE and other organizations. In fact, having

an appropriate amount of local resources can act as a catalyst for local users to request and get more resources from national providers. And it is also true that while a university can hope to address 90% of the local needs by having local resources in place, it would not always be financially possible to address 100% of user needs. It might just make more financial sense to guide and support users with large requirements on national resources.

## 6. CONCLUSION

Our experience is that local investments in supercomputing and HPC resources and appropriate human resources foster and enhance innovation and academic achievement. Local resources reduce the hurdles that researchers and scholars need to cross before getting access to HPC resources and this is having a transformative effect on departments. We are able to relate grants to Big Red II users at our university and this suggests that the benefits to the university in terms of grant income is a good value for the investment in supercomputing resources. Universities and colleges should consider investments in HPC at the appropriate scale for their campuses and they can use the XDMoD\_VA tools in development to quantitatively evaluate the benefit of the investments to their campuses. We conclude this in the basis of interviews and on the basis of linkage of use of our HPC and supercomputing systems and grant success for IU researchers.

## 7. ACKNOWLEDGMENTS

This material is based, in part, upon work supported by the National Science Foundation under Grant No. 1053575 and 1566393 and also by the Indiana University Pervasive Technology Institute (founded with funding from Indiana University and the Lilly Endowment, Inc.). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, Indiana University, or the Lilly Endowment. We would like to thank our vendor partners Cray and DDN for their collaboration in the implementation of Big Red II.

## References

- [1] Big Red II at Indiana University. <https://kb.iu.edu/d/bcqt>. Accessed: 2016-08-22.
- [2] Big Red II Dedication. <http://newsinfo.iu.edu/news/page/normal/24183.html>. Accessed: 2016-09-08.
- [3] Campus Champions - Overview. <https://www.xsede.org/campus-champions>. Accessed: 2016-08-22.
- [4] Confluence - Team Collaboration Software. <https://www.atlassian.com/software/confluence>. Accessed: 2016-09-12.
- [5] Cray Linux Environment (CLE) Software Release Overview. <http://docs.cray.com/books/S-2425-52xx/>. Accessed: 2016-08-22.
- [6] Data Capacitor II. <https://kb.iu.edu/d/avvh>. Accessed: 2016-08-22.
- [7] Discipline categories (for cyberinfrastructure) at Indiana University (PTI Technical Report PTI-TR13-008). <http://hdl.handle.net/2022/20985>. Accessed: 2016-09-09.
- [8] Fairshare scheduling at IU. <https://kb.iu.edu/d/avmu#fair>. Accessed: 2016-08-22.
- [9] Final Report of the Indiana University Cyberinfrastructure Research Taskforce. <http://hdl.handle.net/2022/469>. Accessed: 2016-08-30.
- [10] INCITE Program. <http://www.doeleadershipcomputing.org/>. Accessed: 2016-08-22.
- [11] Indiana University Information Technology Strategic Plan: Architecture for the 21st Century. <http://hdl.handle.net/2022/471>. Accessed: 2016-08-30.
- [12] Indiana University Knowledge Base. <https://kb.iu.edu/>. Accessed: 2016-09-08.
- [13] MOAB HPC Suite. <http://www.adaptivecomputing.com/products/hpc-products/moab-hpc-basic-edition/>. Accessed: 2016-08-22.
- [14] Office of the Vice President for IT and CIO at Indiana University Cyberinfrastructure Value Assessment Report. <https://scholarworks.iu.edu/dspace/handle/2022/20568>. Accessed: 2016-09-12.
- [15] Pervasive Technology Institute. <https://pti.iu.edu/>. Accessed: 2016-09-08.
- [16] Secure Shell. [https://en.wikipedia.org/wiki/Secure\\_Shell](https://en.wikipedia.org/wiki/Secure_Shell). Accessed: 2016-08-22.
- [17] TACC Wrangler. <https://portal.xsede.org/tacc-wrangler>. Accessed: 2016-08-22.
- [18] The TOP500 project. <https://www.top500.org>. Accessed: 2016-08-30.
- [19] UITS User Satisfaction Survey. <http://www.indiana.edu/~uitssur/>. Accessed: 2016-09-08.
- [20] University at Albany Information Technology Services askIT. <https://wiki.albany.edu/display/askit>. Accessed: 2016-09-12.
- [21] University of Wisconsin-Madison Knowledge Base. <https://kb.wisc.edu/page.php?id=3>. Accessed: 2016-09-08.
- [22] XDMoD (XD Metrics on Demand). <http://xdmod.sourceforge.net/>. Accessed: 2016-09-09.
- [23] XDMoD (XD Metrics on Demand) Value Analytics Module. <http://indiana.edu/~xdmodva/>. Accessed: 2016-09-09.

- [24] Final Report of the Indiana University Cyberinfrastructure Research Taskforce. 2005. Available at <http://hdl.handle.net/2022/469>.
- [25] A. Apon, S. Ahalt, V. Dantuluri, C. Gurdgiev, M. Limayem, L. Ngo, and M. Stealey. High Performance Computing Instrumentation and Research Productivity in U.S. Universities. *Journal of Information Technology Impact*, 10(2):87–98, 2010.
- [26] A. W. Apon, L. B. Ngo, M. E. Payne, and P. W. Wilson. Assessing the effect of high performance computing capabilities on academic research output. *Empirical Economics*, 48(1):283–312, 2015.
- [27] L. Michael and B. Maas. Research computing facilitators: The missing human link in needs-based research cyberinfrastructure. *Research bulletin. Louisville, CO: ECAR, May*, 16, 2016.
- [28] S. Pamidighantam, S. Nakandala, E. Abeysinghe, C. Wimalasena, S. R. Yodage, S. Marru, and M. Pierce. Community science exemplars in SEAGrid science gateway: Apache airavata based implementation of advanced infrastructure. *Procedia Computer Science*, 80:1927 – 1939, 2016. International Conference on Computational Science 2016, ICCS 2016, 6-8 June 2016, San Diego, California, USA.
- [29] C. Peebles, C. Stewart, B. Voss, and S. Workman. Measuring quality, cost, and value of it services. In *EDUCAUSE Conference Proceedings, Indianapolis IN, SIGUCCS '01*, 2001.
- [30] G. Staples. Torque resource manager. In *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, SC '06, New York, NY, USA, 2006. ACM.
- [31] C. Stewart, D. Hancock, M. Vaughn, J. Fischer, T. Cockerill, L. Liming, N. Merchant, T. Miller, J. Lowe, D. Stanzione, J. Taylor, and E. Skidmore. Jetstream - performance, early experiences, and early results. In *Proceedings of the 2016 XSEDE Conference: Scientific Advancements Enabled by Enhanced Cyberinfrastructure*, XSEDE '16, New York, NY, USA, 2016. ACM.
- [32] C. A. Stewart. Doing science with big red ii as a campus resource. Presented at Cray Customer Briefing, a satellite meeting of XSEDE13 in San Diego, CA, 2013.
- [33] C. A. Stewart, T. M. Cockerill, I. Foster, D. Hancock, N. Merchant, E. Skidmore, D. Stanzione, J. Taylor, S. Tuecke, G. Turner, M. Vaughn, and N. I. Gaffney. Jetstream: A self-provisioned, scalable science and engineering cloud environment. In *Proceedings of the 2015 XSEDE Conference: Scientific Advancements Enabled by Enhanced Cyberinfrastructure*, XSEDE '15, pages 29:1–29:8, New York, NY, USA, 2015. ACM.
- [34] C. A. Stewart, D. Hart, A. Shankar, E. Wernert, R. Repasky, M. Papakhian, A. D. Arenson, and G. Bernbom. Advanced information technology support for life sciences research. In *Proceedings of the 31st Annual ACM SIGUCCS Fall Conference*, SIGUCCS '03, pages 7–9, New York, NY, USA, 2003. ACM.
- [35] C. A. Stewart, M. Link, D. S. McCaulay, G. Rodgers, G. Turner, D. Hancock, P. Wang, F. Saied, M. Pierce, R. Aiken, M. S. Mueller, M. Jurenz, M. Lieber, J. Tillotson, and B. A. Plale. Implementation, performance, and science results from a 30.7 tflops ibm bladecenter cluster. *Concurrency and Computation: Practice and Experience*, 22(2):157–174, 2010.
- [36] C. A. Stewart, C. S. Peebles, M. Papakhian, J. Samuel, D. Hart, and S. Simms. High performance computing: Delivering valuable and valued services at colleges and universities. In *Proceedings of the 29th Annual ACM SIGUCCS Conference on User Services*, SIGUCCS '01, pages 266–269, New York, NY, USA, 2001. ACM.
- [37] C. A. Stewart, R. Roskies, R. Knepper, R. L. Moore, J. Whitt, and T. M. Cockerill. Xsede value added, cost avoidance, and return on investment. In *Proceedings of the 2015 XSEDE Conference: Scientific Advancements Enabled by Enhanced Cyberinfrastructure*, XSEDE '15, pages 23:1–23:8, New York, NY, USA, 2015. ACM.
- [38] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson, R. Roskies, J. R. Scott, and N. Wilkens-Diehr. Xsede: Accelerating scientific discovery. *Computing in Science and Engineering*, 16(5):62–74, 2014.