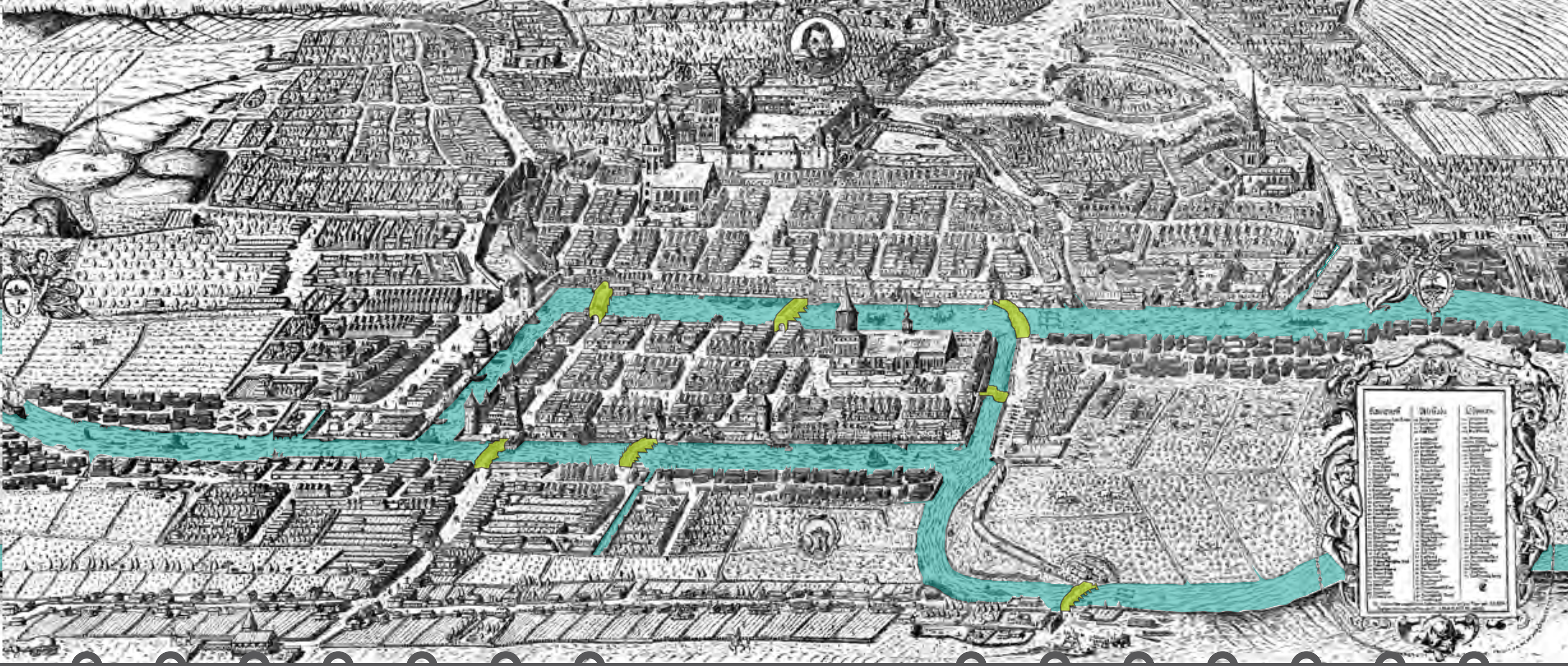


Campus Bridging

Data and Networking Issues Workshop Report

April 7-8, 2010
Indianapolis, Indiana

Editors: Guy T. Almes, David Jent, Craig A. Stewart

Copyright 2011 by The Trustees of Indiana University

This document is released under the Creative Commons Attribution 3.0 Unported license (http:creativecommons.org/licenses/by/3.0/). This license includes the following terms: You are free to share – to copy, distribute and transmit the work and to remix – to adapt the work under the following conditions: attribution – you must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work). For any reuse or distribution, you must make clear to others the license terms of this work.

Please cite as:

Campus Bridging Technologies Workshop: Data and Networking Issues Workshop Report. G.T. Almes, D. Jent and C.A. Stewart, eds., 2011. Available from: http://hdl.handle.net/2022/13200

Acknowledgments

The workshop organizers would like to thank Dale Lantrip (Indiana University Pervasive Technology Institute (PTI)) for handling logistics for the workshop; Matt Zekauskas, (Internet2) and Ray Sheppard (PTI) for taking notes; and Malinda Lingwall, Richard Knepper and Peg Lindenlaub of PTI and Maria Morris of Indiana University Creative Services for their support in the production of the final report document and web pages related to this workshop.

We would like to thank the following individuals and institutions for providing images and/or case study text included in this report: Fran Berman, Rensselaer Polytechnic Institute (Figure 1); Indiana University (Figure 2); Ocean Observatories Initiative Cyberinfrastructure, University of California, San Diego (Figures 3 and 5); The Global Lambda Integrated Facility (Figure 4); Alan Walsh (Figure 6); Guy Almes, Texas A&M University (case study); Akhtar Mahmood, Bellarmine University (case study and Figures A-C); Gary Crane, Southeastern Research Universities Association (case study and Figures D-E); and Eric Wernert, Indiana University (case study and Figures F-H). Alexander Ramírez, Hispanic Association of Colleges and Universities; Karl Barnes, National Association for Equal Opportunity in Higher Education; Al Kuslikis, American Indian Higher Education Consortium; and Geoffrey C. Fox, Indiana University also provided case study text. Unless otherwise noted, each figure and case study text are copyrighted by the person or institution providing the figure and released under the Creative Commons Attribution 3.0 Unported license (http://creativecommons.org/licenses/by/3.0/).

Figure 4 is used by permission from The Global Lambda Integrated Facility. The Global Lambda Integrated Facility (GLIF) Map 2008 visualization was created by Robert Patterson of the Advanced Visualization Laboratory (AVL) at the National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign (UIUC), using an Earth image provided by NASA. Data was compiled by Maxine D. Brown of the Electronic Visualization Laboratory (EVL) at the University of Illinois at Chicago (UIC). Funding was provided by GLIF and US National Science Foundation grants # SCI-04-38712 to NCSA/UIUC and # OCI-0441094 to EVL/UIC. The GLIF map does not represent all the world's Research and Education optical networks, and does not show international capacity that is dedicated to production usage. The GLIF map only illustrates excess capacity that its participants are willing to share with international research teams for applications-driven and computer-system experiments, in full or in part, all or some of the time. GLIF does not provide any network services itself, and researchers should approach individual GLIF network resource providers to obtain lightpath services. For more information on GLIF, see http://www.glif.is/.

The preparation of this report and related documents was supported by several sources, including:

- The National Science Foundation through grant #0948142 (David Jent, Principal Investigator; Kenneth Klingenstein, James Bottum, Jan Odegard, Guy Almes (Co-PIs)
- Indiana University Pervasive Technology Institute (http://pti.iu.edu/) for funding staff providing logistical support of the task force activities, writing and editorial staff, and layout and production of the final report document.
- Texas A&M University (http://www.tamu.edu/) supported this workshop and report by generously providing the time and effort of Guy Almes.

PERVASIVE TECHNOLOGY INSTITUTE
INDIANA UNIVERSITY

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, the Indiana University Pervasive Technology Institute, or Indiana University.

Other materials related to campus bridging may be found at: https://pti.iu.edu/campusbridging/

Workshop participants

The 45 workshop participants and their affiliations are listed below:

Mike Abbiatti, AREON
Guy Almes, Texas A&M University
Alan Blatecky, NSF
Grover Browning, Indiana University
Marianne Chitwood, I-Light
Steve Corbato, University of Utah
Gary Crane, SURA
Matt Davy, Indiana University
Ray Ford, University of Montana
Jim French, University of Virginia
Rob Gardner, University of Chicago
Brandon George, University of Oklahoma
Andrew Grimshaw, University of Virginia
Victor Hazelwood, University of Tennessee Knoxville
Russ Hobby, Internet2
Kay Hunt, Purdue University
Dave Jent, Indiana University
Ken Klingenstein, Internet2
Dale Lantrip, Indiana University
Scott Lathrop, Argonne National Laboratory
Jim Leous, Pennsylvania State University
Akhtar Mahmood, Bellarmine
John McGee, RENCI
Susan Mehringer, Cornell
Greg Monaco, Great Plains Network
Richard Moore, SDSC
Michael R. Mundrane, University of California-Berkeley
Radha Nandkumar, NCSA
Jan Odegard, Rice University
Jim Pepin, Clemson University
Ruth Pordes, Fermilab
Kurt Seiffert, Indiana University
Michael Shoffner, University of North Carolina at Chapel Hill
Stephen Simms, Indiana University
Larry Smarr, Cal-IT2
Craig Stewart, Indiana University
Craig Struble, Marquette
Martin Swany, University of Delaware
David Todd, University of Vermont
John Towns, NCSA
Rob Vietzke, Internet2
James Vincent, University of Vermont
Von Welch, Independent
Matt Zekauskas, Internet2
Tom Zeller, Indiana University

Organizing committee

The workshop organizing committee consisted of David Jent, Kenneth Klingenstein, James Bottum, Jan Odegard, Guy Almes, and Craig Stewart.

Case studies

Executive Summary

The goal of campus bridging is to enable the seamlessly integrated use among: a scientist or engineer's personal cyberinfrastructure (CI); cyberinfrastructure on the scientist's campus; cyberinfrastructure at other campuses; and cyberinfrastructure at the regional, national, and international levels; so that they all function as if they were proximate to the scientist. In April of 2010, Indiana University coordinated a workshop, funded in part by the National Science Foundation (NSF), on the data and networking aspects of cyberinfrastructure and campus bridging.

Today's system of national and international cyberinfrastructure is increasingly complicated and growing. We have a good deal of cyberinfrastructure available relative to data and networks nationally. What is greatly needed, and currently lacking, is a national architecture for cyberinfrastructure that would allow the sort of seamless integration of local, national, and international resources that is the goal described earlier for campus bridging.

The gap between what campus, regional, and national resources are available and what is understood generally by researchers is quite large within US higher education. Given this, a first step in effective campus bridging is to focus on education about availability of these resources. For researchers at institutions of higher education to bridge from where they are to the best facilities as effectively as possible, they must first know what resources are available and appropriate.

Effective, efficient federated identity management and authentication are among the most basic requirements for effective use of distributed cyberinfrastructure. Identity management, namespace management, and authentication remain critical and ongoing challenges in coordination of cyberinfrastructure from the campus level to the national levels. An NSF requirement to employ the InCommon Federation global federated system for identity management for all systems and services it funds, combined with National Institutes of Health adoption of InCommon, should lead the nation to consistent use of a single, interoperable, federated identity system.

Data sets growing in size and complexity present difficult challenges in areas such as storage, networking, metadata, provenance, and security. The size of data sets is growing more rapidly than either the growth of CPU speed or the growth in end-to-end network performance. To support the data moving needs of researchers, every campus of a higher education research institution should have both high-speed networking on campus and a connection to wide-area research networks as a component of its basic infrastructure. Campuses and the US research community generally should lead and challenge the national networking community to build networks connecting the nation's universities so that moving of large datasets is experienced as easy and fast, so that location is not a barrier to data sharing and to access to national CI resources, and so that campus CI and TeraGrid CI can work together flexibly.

There is a significant need for a wide-area federated distributed file system to support data-intensive and data-driven open scientific research in the United States. Such a distributed file system would harness advances in storage, networking, and in distributed file system, scientific data management, and workflow design. It would strengthen our ability to conduct collaborative data-intensive scientific and engineering research and would help bridge campus, regional, and national resources.

Humans are now collecting data that is highly accurate, of potential perpetual interest, and impossible to replace if lost. Examples include weather and climate data, or genome sequence data of species in the wild. In order to be useful over long periods of time, data must be preserved with sufficiently rich metadata that they can be understood by people who were not involved in collecting the data.

Long-term funding continuity for major and important CI projects is critical in part because of its impact on the workforce. If we are as a nation to retain, within the fields of cyberinfrastructure and computational and data-enabled science and engineering, the best and brightest experts then the value proposition they face as individuals must be such that it is rational, and consistent with a good quality of life, to pursue and maintain a career in these fields.

Finally, The cost of the electrical power needed to power (and cool) CI resources will become an increasingly significant issue. Cyberinfrastructure experts should be concerned about their impact on the global environment. Improvements suggested in this report should decrease barriers to full use of campus cyberinfrastructure, thus enabling maximal utility of cyberinfrastructure hardware over the course of its useful life, and supporting breakthrough and practical research enabling the development of human societies in ways that are in harmony with a healthy global environment.

Recommendations

Recommendation 1: The National Science Foundation should lead (and fund) the development of a national architecture for cyberinfrastructure that will enable the seamlessly integrated use among: a scientist or engineer's personal cyberinfrastructure; cyberinfrastructure on the scientist's campus; cyberinfrastructure at other campuses; and cyberinfrastructure at the regional, national, and international levels; so that they all function as if they were proximate to the scientist.

Recommendation 2: The National Science Foundation should strengthen funding for the Campus Champions and similar campus-oriented outreach and education programs.

Recommendation 3: The National Science Foundation must design its cyberinfrastructure programs, including in computational resources, software, networking, storage, and visualization, to incent campus cyberinfrastructure investment. The desired outcome is a balanced and at-least-partially coordinated pattern of investments in campus and national cyberinfrastructure.

Recommendation 4: The National Science Foundation should fund the architecting, implementation, and ongoing maintenance and improvement of a Campus Bridging Software Stack. This should permit use with Unix-based and other operating systems. It should be standards-based rather than implementation-based. It must be simple to use, secure, and enable effective performance of local cyberinfrastructure.

Recommendation 5: As part of a strategy of coherence between the National Science Foundation and campus cyberinfrastructure and reducing reimplementation of multiple authentication systems, the NSF should encourage the use of the InCommon Federation global federated system by using it in the services it deploys and supports, unless there are specific technical or risk management barriers.

Recommendation 6: The National Science Foundation should fund the strengthening of the emerging federated identity, authentication, and authorization infrastructure, with particular regard to improved scalability (including through inter-federation), improved adequacy of authorization in the face of cyberinfrastructure-related applications, and improved security.

Recommendation 7: Campuses should deploy and operate perfSONAR and related tools to systematically measure, debug, record, and display the measured performance.

Recommendation 8: The National Science Foundation should create a new program funding high-speed (currently 10 Gbps) connections from campuses to the nearest landing point for a national network backbone. The design of these connections must include support for dynamic network provisioning services and must be engineered to support rapid movement of large scientific data sets.

Recommendation 9: The National Science Foundation should fund the architecting, implementation, and operations of a wide-area federated distributed file system for use by the US open research community. The resulting system should support federated identity, very high-speed transfer of data (files or blocks) among major repository components of the system, and replication of files to further robustness and performance.

Recommendation 10: The National Science Foundation should fund the architecting of costs-effective ways to archive and preserve data collections, and fund at least some facilities for archiving important data at the national level.

Recommendation 11: The National Science Foundation should fund development of software tools and technology needed for effective remote visualization, and the NSF and institutions of higher education should fund the technology implementation and infrastructure needed for effective remote visualization.

Recommendation 12: The National Science Foundation should encourage and fund the training of more researchers of all types (especially staff) in computational and data-intensive science and engineering.

Recommendation 13: The National Science Foundation should provide more funding for staff supporting use of cyberinfrastructure in research and in particular should provide funding that is more stable and predictable over time.

Findings

Finding 1: New instrumentation (including that installed at the campus lab level) is producing volumes of data that cannot be supported by most current campus networking facilities. There is a critical need to restructure and upgrade local campus networks to meet these demands.

Finding 2: Technological development and implementation of cyberinfrastructure in ways that promote effective campus bridging will also have the natural side effect of enabling cyberinfrastructure use to have the minimal possible impact on the global environment while promoting US research capabilities.

1. Introduction

As laid out in the National Science Foundation's (NSF) "Dear Colleague Letter: Cyberinfrastructure Vision for 21st Century Discovery," [1] cyberinfrastructure (CI) is a key and necessary component to support science and engineering. In the same document, NSF set for itself a vision to lead the development of a comprehensive cyberinfrastructure: "NSF will play a leadership role in the development and support of a comprehensive cyberinfrastructure essential to 21st century advances in science and engineering research and education." In support of this vision, the NSF Advisory Committee on Cyberinfrastructure (ACCI) created a set of six task forces to investigate various aspects of the development of cyberinfrastructure, including the Task Force on Campus Bridging. The Task Force on Campus Bridging has published the following definition [2]:

> *The goal of campus bridging is to enable the seamlessly integrated use among: a scientist or engineer's personal cyberinfrastructure; cyberinfrastructure on the scientist's campus; cyberinfrastructure at other campuses; and cyberinfrastructure at the regional, national, and international levels; so that they all function as if they were proximate to the scientist. When working within the context of a Virtual Organization (VO), the goal of campus bridging is to make the 'virtual' aspect of the organization irrelevant (or helpful) to the work of the VO. Campus bridging is critical to supporting the ever-increasing level of cross-disciplinary and cross- organizational aspects of scientific research, as it enables not just the connection of scientists with CI beyond their campus, but also the connection of scientists with other scientists to support collaboration.*

In April of 2010, Indiana University coordinated a workshop, funded in part by the National Science Foundation, on the data and networking aspects of cyberinfrastructure and campus bridging. The workshop took a broad view of cyberinfrastructure, networking, and data. Specifically, the workshop addressed the following goals related to the general themes of campus bridging:

- Networking
  - From the campus perspective: identify best practices in campus networking and end-to-end computing architecture (where that may mean lab to campus to RON to national backbone to resource hanging off of national backbones)
  - Address the question, "How do you design a network for researchers when you are designing overall for the masses or for campus business operations?"
  - What is the role of IPv6?
- Data
  - Due to trends in data storage and instrumentation, the volumes of data to be moved are increasing in an aggressive exponential fashion – the (end-to-end) network will need to keep up. Discuss networking design within campus and how it interacts with national trends, international trends (including technology trends of different scaling rates for different areas of technology). Specifically, how must the end-to-end network architecture meet the needs (including data access and remote visualization) stemming from Campus Bridging?

- o From a researcher's standpoint, how do I know where my data are and how do I get them? How does the scientific community deal with discovery of data resources generally (in the sense of finding public and/or shareable data resources).
- o Data storage infrastructure: what are the expectations on campuses? What are reasonable expectations regarding facilities that should be provided by the NSF?
- o What are the relationships between data management and the proper safeguarding of campus intellectual property?

Prior to the workshop, the organizing committee solicited position papers on networking and data matters related to campus bridging. A total of 12 such position papers were submitted; they are included in Appendix1. Participants were invited to take part in this workshop through a combination of invitations to US leaders in data and networking, and in part through self-nomination by submission of a position paper. A total of 45 individuals took part in the workshop itself.

There were already many excellent reports regarding networking and data in general terms at the time this workshop was held. A handful of those documents include:

- Developing a Coherent Cyberinfrastructure from Local Campuses to National Facilities: Challenges and Strategies [3]
- Our Cultural Commonwealth [4]
- Cyberinfrastructure Software Sustainability and Reusability Workshop Final Report [5]

Since this workshop concluded, the National Science Foundation Advisory Committee Task Force on Data and Visualization has completed a thorough report on data and visualization matters [6]. The workshop described in this report did not attempt a comprehensive review of all matters data and networking. Rather, we focused specifically on those matters most closely related to data and networking from a campus bridging perspective, with a focus on scientific research carried out in a discipline supported by the National Science Foundation. This workshop was also, chronologically, the first of several workshops related to the general area of campus bridging (see [7] for more information about other workshops and reports on this topic). As a result, much of the discussion at this conference focused on defining what campus bridging means; it was one of the critical steps toward the goal stated for campus bridging at the beginning of this report.

The remainder of this document summarizes the discussions and consensus outcomes of the workshop. Appendix 2 includes images of the slides from presentations given at the workshop. Two-thirds of the workshop participants have voted to endorse this report, with no votes in opposition.

2. General discussion of campus bridging and cyberinfrastructure

It has long been observed that high performance computers can function as a time machine – allowing science and engineering researchers to use IT years before market IT would allow. For example, the NSF supercomputer centers program [8] and the generally small number of supercomputer users on each campus in the mid-1980s drove campus-wide and nationwide change in the way scientific research is pursued. This observation regarding 'time machine' functionality was first made as far as we can tell in the early 1990s by Larry Smarr, and appears in a 1995 report by the National Research Council [9]; it also applies to advanced cyberinfrastructure in general.

Cyberinfrastructure is a much harder concept to grasp than supercomputing and more difficult to use practically. Use of a supercomputer – particularly a vector supercomputer – in the mid-to-late 1980s was relatively straightforward and was driven by machine capabilities and a relatively small number of straightforward programming principles. The structure of the national research IT environment (which we might now retrospectively call cyberinfrastructure) was relatively straightforward – well illustrated by the Branscomb Pyramid, defined in 1993 [10]. The Branscomb Pyramid defines a series of 'vertical strata' between workstations, campuses, large-scale systems, and leadership class supercomputers and makes implications about the relative abundance of systems of each type. For many years, the Branscomb Pyramid has served as a useful heuristic to understanding the structure of the US science and research cyberinfrastructure. In the open science community, this was because, in large part, no entity other than the NSF had the financial capability to fund the systems occupying the pinnacle of the pyramid for use by the national open science community. Figure 1 illustrates the Branscomb Pyramid circa 2006 as depicted in a talk by Dr. Fran Berman [11].

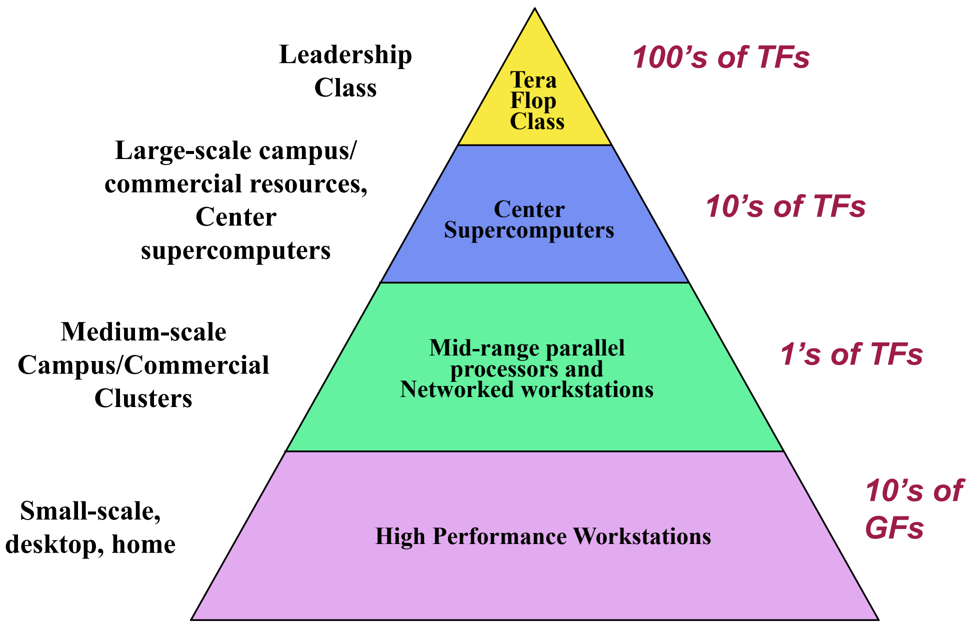


Figure 1. The Branscomb Pyramid, circa 2006, from presentation by Dr. Fran Berman, Director, San Diego Supercomputer Center [11]. (This image by Fran Berman, licensed under the Creative Commons 3.0 unported attribution license [12].)

Today's system of national and international cyberinfrastructure is much more complicated and of much greater scale than when the Branscomb Pyramid was first set out in 1993, as is discussed in detail in the NSF Advisory Committee for Cyberinfrastructure Task Force on Campus Bridging Final Report [2]. Figures 2 through 5 provide examples of two specific cyberinfrastructure installations – the TeraGrid and the Ocean Observatories Initiative. They depict a much more complex cyberinfrastructure than captured by the Branscomb Pyramid.

Figure 2. A map of the TeraGrid showing major computational and storage resources as of April 2011. Image courtesy of Indiana University, based on illustration by Nicolle Rager Fuller, National Science Foundation.

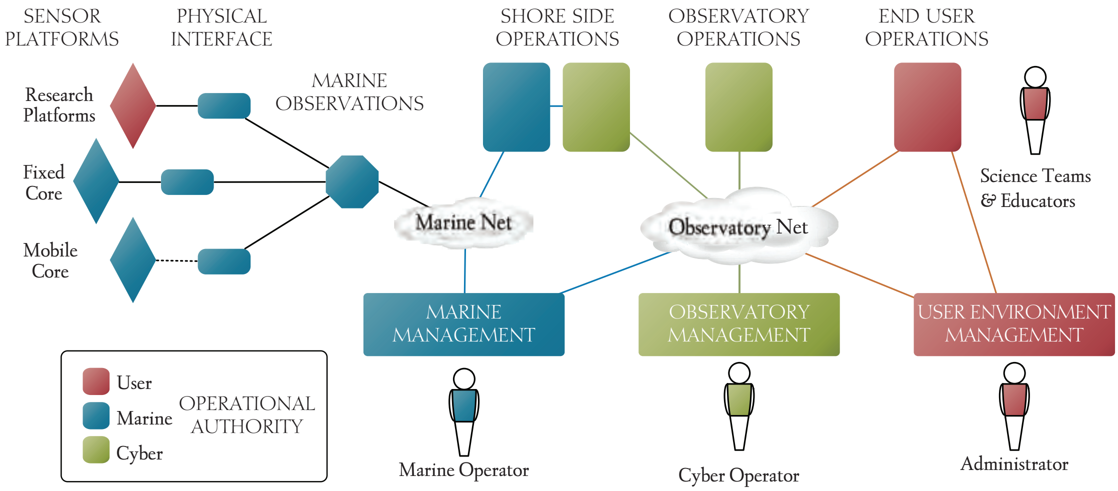


Figure 3. Diagram of the observatory topology of the Ocean Observatories Initiative. Image provided by Ocean Observatories Initiative Cyberinfrastructure, University of California, San Diego.

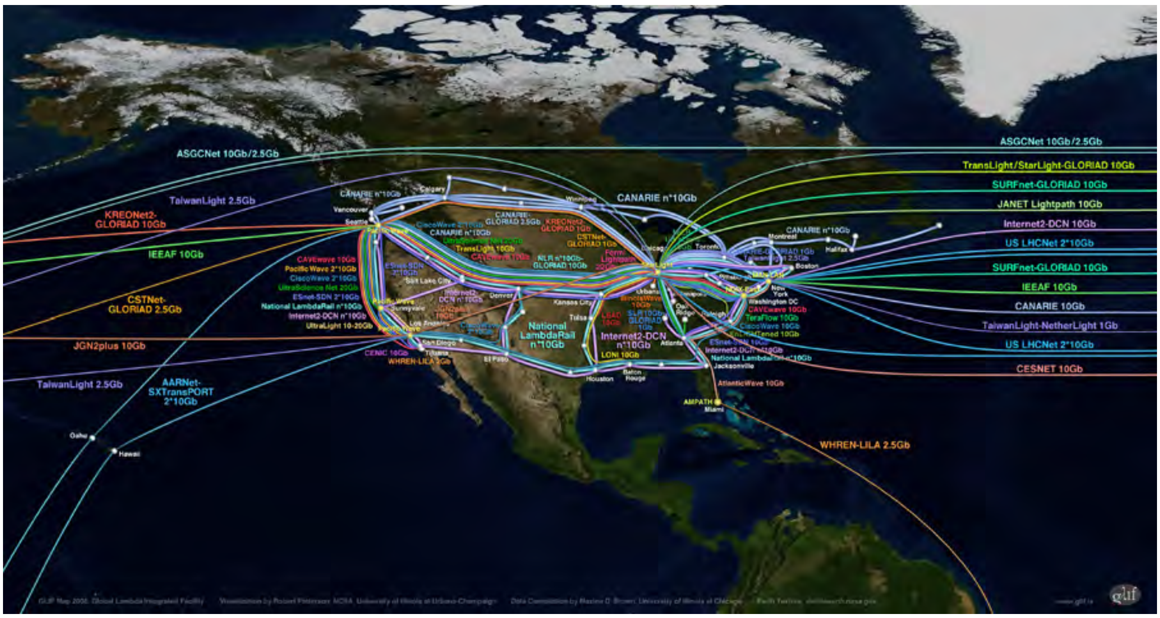


Figure 4. Research and education network bandwidth made available for scheduled application and middleware research experiments as of May 2008. Map used with permission from the Global Lambda Integrated Facility (GLIF) [13].

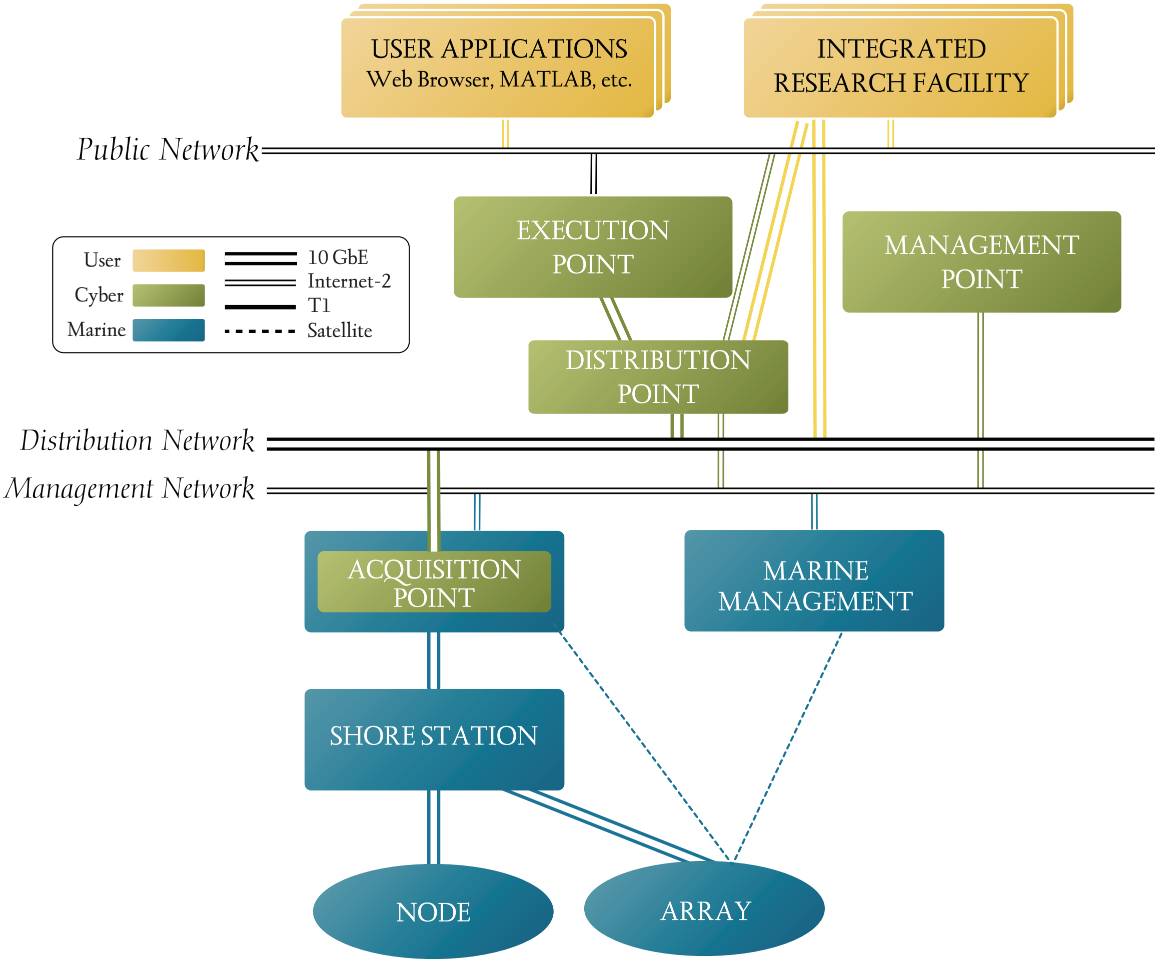


Figure 5. Diagram of the OOI cyberinfrastructure. Image provided by Ocean Observatories Initiative Cyberinfrastructure, University of California, San Diego.

In 2010, we have a much larger and growing number of components within the national cyberinfrastructure, including:

- National and regional networks, including Internet2 [14], National Lambda Rail [15], and a host of regional optical networks.
- Computationally-oriented national CI such as the TeraGrid [16] and Open Science Grid [17].
- Regional computationally-oriented CI such as SURAgrid [18].

- Nationally accessible visualization and collaboration facilities such as the OptIPuter [19], the TeraGrid visualization facilities Longhorn [20] and Nautilus [21]
- National and international data facilities such as DataONE [22], Data Conservancy [23], HathiTrust [24], and the International Virtual Observatory Alliance [25].
- Important storage and management facilities operated by individual universities, such as the University of North Carolina's DigCCurr [26], and Illinois's CIRSS [27], and the IU Data Capacitor [28].

With the growing capacity and capability of these and other very successful projects we have a good deal of cyberinfrastructure available relative to data and networks nationally. What we do not yet have is a national cyberinfrastructure architecture. In thinking about CI, it is important to keep the holistic nature of CI firmly in mind. Thus, CI is not just about HPC and data and visualization, but also about how they combine and the people that make it work. Thus, were we to have a national cyberinfrastructure architecture, these components, working together, would be a much more powerful instrument in the hands of the nation's scientists and engineers.

The NSF's 2007 Cyberinfrastructure Vision for 21st Century Discovery [1] was excellent in many regards, but somewhat unbalanced (reflecting concerns and needs of the time) in that it has a much greater focus on high performance computing than on any other area of cyberinfrastructure. Some of the oldest of the current national cyberinfrastructure facilities, the TeraGrid and Open Science Grid, are both more successful at organizing compute cores than at organizing data.

The Internet we enjoy today – research networks as well as commodity networks that have revolutionized business and communications – are the direct result of NSF leadership in networking in the 1980s. The NSF then exerted a tremendous unifying force in networking standards by prudently selecting TCP/IP as the single networking protocol for NSFNET [29]. This decision aligned and organized millions of dollars in research and development investment by public and private sectors, leading to the Internet revolution and the billions of dollars of revenue created in the process. Intelligently mandating appropriate architecture can be of very great value. This was relatively easy in the given example, but still crucial. It was critical in enabling the NSFnet community, including contributors at the campus, regional, and backbone layers, to "think globally and act locally." This is more difficult, both technically and politically, in the present environment, but no less important. [See "NSFnet as a valuable cyberinfrastructure precedent," page 20.]

There is no exact analogy between networking protocols and the challenges of creating an effective national cyberinfrastructure through campus bridging, but the NSF has a tremendous opportunity to influence and direct with strategic investments from its own budget a total investment by public and private sectors much greater than the actual budgets under control of the NSF. What is greatly needed, and currently lacking, is a national architecture for cyberinfrastructure that would allow the sort of seamless integration of local, national, and international resources that is the goal described earlier for campus bridging.

The Open Science Grid is a good example of campus bridging, leveraging campus CI resources, with limited federal funding, and serving as a national CI resource. The Open Science Grid also demonstrates the value of providing open access for researchers to resources and not letting existing CI limitations or unnecessary policy restrictions get in the way.

NSF defines "Track 3" computing resources in its taxonomy, but provides little funding for them in general. This has been a defined part of the NSF funding strategy [30], and represents a generally reasonable approach to the empirical fact that the NSF budget is insufficient to supply the nation's cyberinfrastructure needs. The one program that routinely funds Tier 3 resources – the Major Research Infrastructure program – was viewed by a strong majority of workshop attendees as not having had as much emphasis or success in the past in handling cyberinfrastructure software in ways that facilitate campus bridging and optimal use of the aggregate national cyberinfrastructure.

This leads to the first recommendation stemming from this workshop:

> ***Recommendation 1. The National Science Foundation should lead (and fund) the development of a national architecture for cyberinfrastructure that will enable the seamlessly integrated use among: a scientist or engineer's personal cyberinfrastructure; cyberinfrastructure on the scientist's campus; cyberinfrastructure at other campuses; and cyberinfrastructure at the regional, national, and international levels; so that they all function as if they were proximate to the scientist.***

A strong consensus among the participants of this workshop was that this recommendation should be carried out in the context of National Science Foundation plans for a Cyberinfrastructure Framework for 21st Century Science and Engineering (now known as CIF21 [31]). An effective software architecture is the key to success in effective campus bridging and for accomplishing the NSF vision for CIF21.

NSFnet as a valuable cyberinfrastructure precedent

The NSFnet program, spanning roughly from 1986 to 1995, offers several lessons for Campus Bridging.

At a technical level, the early decision to focus exclusively on TCP/IP with a hierarchical backbone-regional-campus connection structure had enormous impact. At the time, academic networks in the US had strong consensus on the technology of packet switching, but deployed infrastructure included substantial pockets of TCP/IP, DECnet Phase IV, BITnet, UUCP, dial-up CSnet, and MFEnet. With heroic work by clueful engineers, a variety of application gateways created some degree of interoperability for email. By focusing on TCP/IP, the time of technical experts and investments in circuits and routers were dramatically improved. Equally important, a consistent IP-based infrastructure created an environment that fostered the creation and rapid adoption of innovative applications. Further, interactions among disciplines (e.g., between high-energy physics folks who had been using DECnet and fusion physics folks who had been using MFEnet) were dramatically enhanced. Technical leadership by the NSFnet Backbone engineers at MERIT paid dividends through efforts such as the Policy Routing Database. By dictating (and this was not too harsh a word in the context of the times) a good networking architecture (in essence the choice of TCP/IP and a hierarchical arrangement of backbone, regional, and campus networks), NSF created a climate in which network engineers across the country could "think globally and act locally."

At a funding level, NSF chose to "fully" fund the backbone, provide seed money for limited periods of time to the regional networks, and to provide no funding to the campus LANs. The quotes around "fully" in the previous sentence relates to substantial cost sharing by the State of Michigan, IBM, and MCI, but it was useful that the rest of the community, including regional and campus networks and their users, could view it as fully funded over a period extending from summer 1988 to spring 1995. While the pre-NSFnet wide-area infrastructure had succeeded with 50- and 56-kbps circuits, the NSFnet backbone introduced T1 (1.5 Mbps) capacity in 1988-1989 and T3 (44.7 Mbps) capacity by 1992-1993. Thus, the wide-area capacity of the NSFnet enjoyed increases of a factor of 24 followed by a factor of 28, all within a seven-year period. This was motivated, in part, by the NSFnet's very challenging narrow mission – to enable the transfer of "large" data files between computational scientists and the NSF supercomputer centers they were using. In return, however, this dramatic expansion in backbone capacity enabled both the enormous increase in pre-NSFnet applications such as electronic mail, but also created the environment necessary for the emergence of the Web.

The regional networks served as a laboratory for technical approaches (e.g., in circuit topology styles, in router technology, in operations) and in nontechnical approaches (e.g., in governance and funding models). Engineers at NYSERnet created the SGMP protocol to manage their gateways (routers), and this led directly to the creation of the modern SNMP protocol and related MIBs. The technical architectural decisions enabled these regional networks to be supporting the same interfaces, but to do so in innovative and dissimilar ways. The result was a strong degree of interoperability and innovation. Further, wide-area technologies, such as the T1 and T3 circuits pioneered in the backbone, came (after some delay) into increasingly widespread use within the regional networks. This, in turn, encouraged the emergence of a competitive IP router market.

The campus networks, similarly, explored a variety of technical and non-technical approaches. The vast majority of overall NSFnet project investment was at the campus level, with no NSF financial support. During the 1988-1995 period, these campus LANs exhibited explosive growth by any measure. Over this seven-year period, the NSFnet transitioned from a niche technical service for a narrow subset of scientists and engineers to an infrastructure that tied together academics from all university departments and from universities spread worldwide. Moreover, as undergraduate students became familiar with the Internet, and as those students graduated and entered the workforce with their Internet use skills, a key foundation of the rapidly growing commercial Internet of the mid- and late-1990s was laid.

Text provided by Guy Almes, Texas A&M University.

3. University context

There are several aspects of the challenges to effective campus bridging that have their origin in the organization of campus cyberinfrastructure and IT. Centers within colleges and universities that support research cyberinfrastructure typically report to one of two entities – the Chief Information Officer (CIO) or the Vice President or Vice Provost for Research [32]. Each situation has its own set of challenges.

Regarding CIOs and campus research cyberinfrastructure, the workshop participants made the following general observations:

- CIO organizations serve many users with a broad set of needs.
- CIOs are under considerable pressure, particularly in regards to mission critical systems, learning system support, legal compliance, and security. They sometimes simply do not have the time and resources to focus on research cyberinfrastructure. CIOs have been fired over a variety of issues related to mission critical systems, security, and regulatory compliance. Workshop participants were unable to identify a single case of a CIO being fired for failing to provide a sufficiently good campus research cyberinfrastructure.
- As a result of all of the above factors, CIOs sometimes lack the resources, ability, or motivation to focus on the needs of researchers on the campuses they support.
- There are other external forces as well. Politically, there may be challenges within some states to university / CIO involvement in creating high-end university-based networks.

Campus CI delivery and support groups that report to a Vice President or Vice Provost for Research (VPR) face challenges as well. Among those challenges workshop participants identified the following:

- Research CI organizations reporting to a VPR operate sometimes in isolation, or even in some cases in apparent competition, with the campus IT organization that reports to the CIO.
- Some of the most successful CI centers, at the national level, report to a VPR, and get their funding primarily from and primarily serve a national audience. Because of this, they can be perceived by researchers on their local campus as insufficiently interested in the research needs of local researchers.

As much as many of the workshop participants would like their concerns to be at the top of every CIO's priority list, it was recognized at this workshop that this is not the case. There was a strong feeling at the workshop that in general CIOs should be better informed about and more concerned with research cyberinfrastructure. In particular workshop participants felt that CIOs should recognize that research CI differs from the rest of their general campus information technology infrastructure and not try to shoehorn research CI into the traditional IT infrastructure model.

The issue of awareness of needs and opportunities is not limited to CIOs. The gap between what campus, regional, and national resources are available and what is understood generally by researchers is quite large within US higher education. Given this, a first step in effective campus bridging is to focus on education about availability of national resources. For researchers at institutions of higher education to bridge from where they are to regional and national facilities as effectively as possible, they must first know what resources are available and appropriate. The TeraGrid Campus Champions program [33] was widely regarded as being an effective way to educate researchers about the services offered via the TeraGrid. Discussion of this led to the following recommendation:

> ***Recommendation 2. The National Science Foundation should strengthen funding for the Campus Champions and similar campus-oriented outreach and education programs.***

At a more comprehensive level, the NSF has the opportunity to lead the national research community in ways that, while not precisely analogous to the way NSF led development of the modern Internet by specifying TCP/IP as a standard, provide the opportunity to incent and align a very large amount of investment across the nation. The workshop participants enthusiastically endorsed the idea of the NSF aligning with a national cyberinfrastructure architecture:

> ***Recommendation 3. The National Science Foundation must design its cyberinfrastructure programs, including in computational resources, software, networking, storage, and visualization, to incent campus cyberinfrastructure investment. The desired outcome is a balanced and at-least-partially coordinated pattern of investments in campus and national cyberinfrastructure.***

One possible example of this could be that if a researcher gets a TeraGrid allocation, then the campus could receive some funding to upgrade its network (to facilitate access to the TeraGrid), with this funding conditional on the campus demonstrating appropriate end-to-end network performance. Another possible example of incentives that the NSF could put into place would be to include interoperability with NSF-funded national cyberinfrastructure as a review criterion in evaluating Major Research Instrumentation proposals. In order to achieve the vision for effective bridging that was presented at the beginning of this report, coordinated and aligned investment at all levels of the US academic and open research communities will be essential. For this to happen, it will be essential for entities at all levels of higher education and the open research community to make a case for this sort of coordination. This is clearly needed to enable US competitiveness globally, but NSF incentives will aid individual researchers, departments, campuses, and multi-institution consortia to make this case locally as well as nationally.

One way to incent alignment of campus and national cyberinfrastructure would be to provide a software suite that would make standardization and effective campus bridging straightforward. This leads to the following recommendation:

Recommendation 4. The National Science Foundation should fund the architecting, implementation, and ongoing maintenance and improvement of a Campus Bridging Software Stack. This should permit use with Unix-based and other operating systems. It should be standards-based rather than implementation-based. It must be simple to use, secure, and enable effective performance of local cyberinfrastructure.

In addition to the campus and the national layers, there is some promise in exploring "regional" layers of CI resources. Existing regional organizations (including the Regional Optical Networks (RONs)) have a wide variety of styles and of missions.

4. Identity management and authentication

Effective, efficient federated identity management and authentication are among the most basic requirements for effective use of distributed cyberinfrastructure. For CI providers, this is key to identifying researchers trying to use their services. For CI users and collaborations of CI users, this is key to accessing a variety of campus and remote CI resources and also to managing access to information among collaborations of researchers from different institutions.

For all these reasons, identity management, namespace management, and authentication remain critical and ongoing challenges in coordination of cyberinfrastructure from the campus level to the national levels. Identity management is one of the critical obstacles to effective campus bridging and more effective use of the nation's human resources and CI assets. There have been demonstrated successes in delivering CI resources by making use of authentication via the InCommon Federation [34] through the SAML protocol. These successes include most notably the TeraGrid [16] and, at smaller scales, the National Institutes of Health (NIH)-funded Indiana Clinical and Translational Studies Institute [35], and the Committee on Institutional Cooperation (CIC) [36] Chief Information Officer (CIO) group.

The workshop participants reaffirm and expand upon a recommendation made in the EDUCAUSE / Coalition for Academic Scientific Computation joint report "Developing a Coherent Cyberinfrastructure from Local Campus to National Facilities: Challenges and Strategies" [3], specifically:

> ***Recommendation 5. As part of a strategy of coherence between the National Science Foundation and campus cyberinfrastructure and reducing reimplementation of multiple authentication systems, the NSF should encourage the use of the InCommon Federation global federated system by using it in the services it deploys and supports, unless there are specific technical or risk management barriers.***

Use of the InCommon Federation system requires a specific set of guidelines and quality assurance processes for every campus that becomes a member of InCommon. InCommon-based authentication is used in delivery of a service across domain boundaries, where a person with an identity in one name management / authentication domain accesses a service beyond that domain (generally in inter-institution or inter-campus situations). In discussion of the workshop recommendations on authentication systems, it became clear that a concise technical guide to implementation of authentication with InCommon-based authentication was needed. To aid campuses and projects in the effective deployment of a common identity management system through the use of InCommon, the NSF funded the creation of an "InCommon Roadmap for NSF Cyberinfrastructure" [37, 38]. The Roadmap document offers guidance for campuses and CI projects to implement a minimal level of participation in InCommon in order to support NSF researchers. The improving ability of Microsoft Active Directory Service (ADS) to support InCommon credentials makes use of such credentials much more accessible for small institutions. For institutions with very low numbers of potential users of NSF facilities, and limited ability to implement namespace management systems, it is now possible to purchase InCommon credentials for individual researchers or students through private companies (such as ProtectNetwork [39]), making this approach feasible for all institutions of higher education.

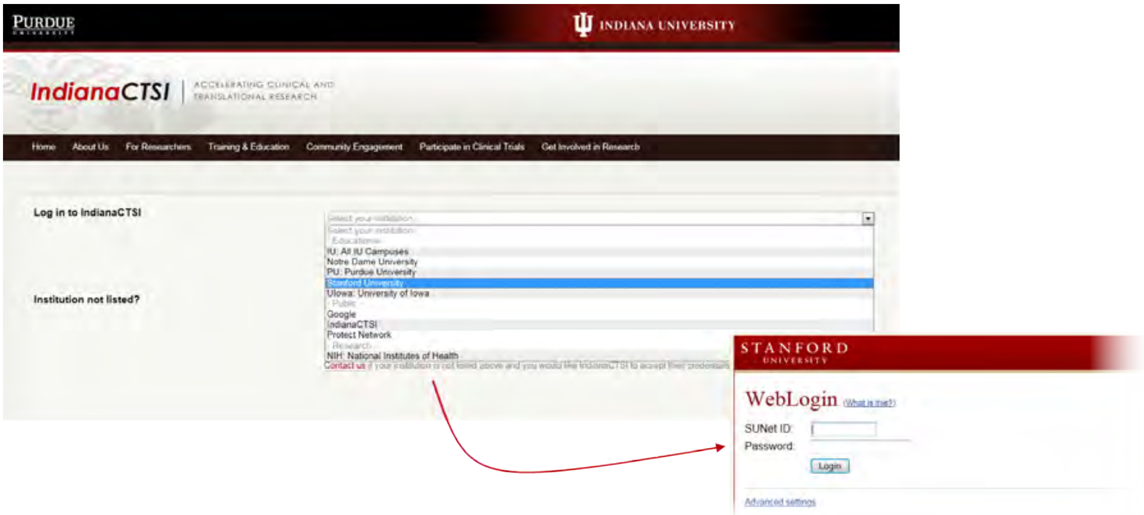


Figure 6. Use of software tools such as GridShib, the SAML protocol, and the InCommon Federation allows a researcher to access Web-based tools, as shown here, and authenticate via the authentication system at their own home institution (in the example shown here, Stanford University). NIH has developed a roadmap for use of InCommon as the basis for authentication for many of its institutional applications [40]. (Image provided by Alan Walsh, licensed under the Creative Commons 3.0 unported attribution license [12].)

The NSF has also funded the CILogon Service [41], which provides a bridge and translation service between InCommon and the International Grid Trust Federation-based (IGTF) [42] public key infrastructure certificates [43] (a.k.a. "grid certificates") that are now common for NSF cyberinfrastructure facilities.

The NIH has already announced, and partially implemented, plans to deploy a series of applications accessible only through use of InCommon-based authentication [40].

An NSF requirement to employ the InCommon Federation global federated system for identity management for all systems and services it funds – including its use for access to MRI-funded facilities when used by individuals accessing such facilities from the namespace in which said instrument resides – combined with NIH adoption of InCommon should lead the nation to consistent use of a single, interoperable, federated identity system. To the extent that this will lead to many more institutions joining the InCommon Federation, and thus improving and documenting their own identity management processes, implementation of this strategic recommendation on the part of the NSF will lead to improved cybersecurity among NSF CI facilities and services and US higher education generally.

The success of adopting use of technologies related to InCommon, such as Shibboleth and GridShib, demonstrates how the Shibboleth-based federated identity/authentication system can be used in conjunction with a system (the TeraGrid) designed for X.509 certificates.

In addition to simply requiring use of the InCommon Federation global federated system for identity management, NSF should provide continued funding for its strengthening, development, and use. In particular:

> ***Recommendation 6. The National Science Foundation should fund the strengthening of the emerging federated identity, authentication, and authorization infrastructure, with particular regard to improved scalability (including through inter-federation), improved adequacy of authorization in the face of cyberinfrastructure-related applications, and improved security.***

Future needs include support for strong (two-factor) authentication, and development of abilities to do inter-federation. For example, InCommon membership is by definition limited to organizations operating within the United States. In order to support international collaborations, it will be necessary to interoperate with European, Asian, and North and South American security systems.

5. Data production

We note, at the beginning of this section, that data present (at least) two different kinds of difficulties:

- Data are growing in size, and this very size presents difficult challenges, e.g., in storage and networking. (Indeed, the world's rate of production of digital data exceeds the rate of growth of digital storage media [44].)
- Data are growing in complexity, and this complexity presents difficult challenges, e.g., in metadata, provenance, and security.

While these are certainly related sets of issues, and while the difficulties of one set make solutions to issues in the other set more difficult, we also see value in focusing on them separately. In discussion of the recent "Blueprint for the Digital University," the report of the University of California – San Diego (UCSD) Research Cyberinfrastructure Design Team, the workshop noted the distinction between "digital curation" and "centralized disk storage" as two separate major elements of the six major elements of UCSD research cyberinfrastructure [45]. In the context of these two distinct sets of difficulties (data are big vs. data are complex), we chose to focus primarily on the difficulties related to size, and to leave leadership in complexity-related issues to the Task Force on Data and Visualization.

The size of data sets is growing more rapidly than either the growth of CPU speed or the growth in end-to-end network performance. This has been known for some time and documented particularly well by the late Jim Gray [46]. This is true both of individual data sets and, particularly, of collections of data sets. This stems from (at least) two separate causes:

- Sources of data, both modern instruments (e.g., gene sequencers) and modern computational systems (e.g., HPC clusters producing multi-terabyte data sets), generate larger data sets at a faster rate.
- The capacity of disk storage per dollar is increasing at an exponential rate exceeding that of Moore's Law.

One might imagine that, as storage capacity per dollar increases, the money being spent on disk storage will decrease. Economic theory suggests that this might not be the case. The Jevons Paradox suggests that as technology becomes more efficient, demand may actually increase. This idea is due to William Stanley Jevons, who noted that mid-1800s improvements in the efficiency of coal-based technologies did not lead to decreases in the consumption of coal [47]. This idea has more recently been applied to energy efficiency [48]. The experience of workshop participants suggests that decreased cost of disk storage has increased demand for storage.

The digital data generation capabilities of new instruments have tremendous implications. They contribute heavily to science (with greater resolution, samples/sec, etc.), but they also may overwhelm both our networks and our current structures for managing massive data storage. Further, while the disciplines using these new instruments include some usual "standard suspects" fields such as high-energy physics (e.g., the CMS and Atlas detectors on the Large Hadron Collider), chemistry (e.g., mass spectrometers), and astronomy (e.g., any of the recent optical or radio telescopes), they also include groups from biology, agriculture, veterinary, and medical research teams who are experiencing particularly sudden growth. The data production rates of major large facilities are presented in Table 1.

Data source	Data production / year
Large Synoptic Survey Telescope [49]	4.5 PB
Large Hadron Collider [50]	2 PB
One Degree Imager [51]	500 TB – 1.5 PB
COLA (Center for Ocean-Land-Atmosphere Studies) [52]	1.8 PB/year
Ocean Observatories Initiative [53]	1 PB
CRESIS [54]	20 TB/expedition
LEAD [55]	4.7 TB/prediction season
EVIA - 200 hours video in a 2-month trip [56]	1.2 TB/trip
Earthscope [57]	920 GB

Table 1. Summary of a variety of large-scale data sources and output of data. Adapted from Stewart 2010 [58].

Next generation biological instruments are in particular radically changing the topology of data production on campuses. It's quite possible for a single researcher to order a single instrument that produces more data per day than the capacity of the network connection to the entire building in which that instrument is installed. In addition, it is quite possible for such an instrument to be ordered and physically installed without anyone in the campus IT organization knowing about it until after the instrument is installed.

Type of instrument	Model	Raw image data	Data products
Light Microscopy	BD Pathway 855 Bioimager [59]	N/A	7 GB/day
Genome sequencing	Roche 454 Life Sciences genome analyzer system [60]	39 GB/day	9 GB/day
	Illumina-Solexa genome analyzer system [61]	367 GB/day	100 GB/day
	ABI SOLID 3 [62]	238 GB/day	150 GB/day
Microarray Gene Expression Chip Reader	Molecular Devices GenePix Professional 4200A Scanner [63]	N/A	8 MB/day
	NimbleGen Hybridization System 4 (110V) [64]	N/A	300 MB/day

Table 2. Data production rates of several current genomic instruments. Adapted from Stewart 2010 [58].

Data mining, understood broadly, is a key emerging paradigm. For example, consider how astronomers have learned to make use of the image data organized in the Sloan Digital Sky Survey. Life scientists need to be able to apply similar techniques to genome data. Genomics researchers will now need to take similar data mining approach to their field, building databases of genomes from many organisms, and then being able to search the database for genomes with specific properties. This leads to several challenges:

- Data storage of large numbers of such genomes, with solid reliability and performance in moving the data sets around the campus and around the country.
- Managing metadata, including ownership of the data. In some cases, successful sharing of the data will require a carefully limited sharing of the data, governed, for example, by intellectual property constraints. Even in veterinary research contexts, where no privacy issues would seem to be at play, intellectual property issues are often significant.
- Preserving the data over long periods of time. For example, the genomes of key plant and animal samples may be scientifically relevant for decades.
- Understanding the data and making discoveries based on very large amounts of data through data mining.

This general discussion leads to a finding reached by consensus of the workshop participants:

> **Finding 1. New instrumentation (including that installed at the campus lab level) is producing volumes of data that cannot be supported by most current campus networking facilities. There is a critical need to restructure and upgrade local campus networks to meet these demands.**

One example of an excellent campus network installation is project Quartzite, which has built a campus network at UCSD specifically for research data and research activities. This project, described in the NSF award abstract [65], is a key enabler of local campus data connectivity and collaboration. Indiana University has taken a more limited approach to this issue, installing dedicated 10 Gbps links from the central research systems in IU's Data Center to specific labs and buildings on campus that include high-output data production instruments.

OSG ATLAS-TIER3 supercomputer at Bellarmine University

Bellarmine University has a state-of-the-art, grid-enabled, 51-node supercomputing cluster equipped with 408 cores, 1300GB of RAM and 375TB of hard disk space (see Figure A). The cluster has been operational under the Open Science Grid (OSG) cyberinfrastructure since September 20, 2010 and is currently being used for the CERN Large Hadron Collider (LHC) ATLAS high energy physics experiment as a dedicated ATLAS Tier3 site.

We have begun the initial phases of ATLAS Monte Carlo production and are optimizing the data transfer throughput from our partnering ATLAS Tier2 OSG site at the University of Oklahoma. We have implemented various Web portals for monitoring the Tier3 cluster's performance (which is running Scientific Linux), using PCM (Platform Cluster Manager), Cacti (a Web-based graphing tool) and Nagios (network monitoring software). We have also set up an OSG monitoring dashboard for the Tier3 cluster using Netvibes (see Figure B).

Using CERN's Atlantis software package, PI Akhtar Mahmood and his students are currently analyzing both Monte Carlo and ATLAS data to study the particle tracks and decay patterns of the Higgs boson, including the identification of the W/Z bosons and the top quark (see Figure C). The undergraduate students are focusing on filtering methodologies and pattern recognition techniques.

The OSG grid site at Bellarmine University is the state of Kentucky's only OSG site. It is also OSG's first site located at a predominantly non-Ph.D. granting undergraduate institution in the US. This grid site is also part of the LHC Computing Grid (LCG). At Bellarmine University, all the grid-enabled ATLAS Monte Carlo and data analysis research tasks are being conducted collaboratively by the PI and his five undergraduate physics students, along with researchers from the University of Oklahoma and SUNY-Albany.

Figure A. OSG (Open Science Grid) ATLAS Tier3 Supercomputer at Bellarmine University.

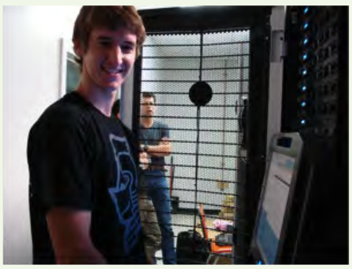


Figure B. Physics student Jovan Andjelich in front of the cluster's console, performing the various systems administration tasks.

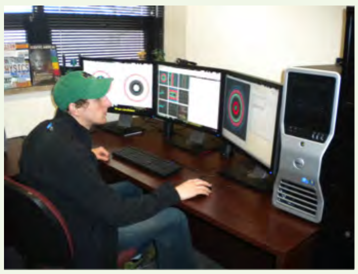


Figure C. Physics student Benjamin Draper analyzing ATLAS data and studying particle tracks using CERN's Atlantis software.

Text and images provided by Akhtar Mahmood, Bellarmine University.

6. Networks and data movement

In terms of moving the data, the problem begins within campuses and extends at every level of networking to national and international networks. The network that researchers need to keep up with data generation has different properties from the regular Internet, including: supporting multiple Gigabits per second (Gbps) both within the campus and across the nation, supporting the flow of large data objects, and supporting visualization. In the 1990s' High Performance Connections program, there was initial emphasis on "meritorious applications," but this was later dropped. This allowed the whole campus to benefit from the program, but had the downside that campus network design became disconnected from the needs of the driving meritorious applications. Furthermore, the data production capability of new instruments, such as next-generation gene sequencers, means that campus networking needs related to research may no longer be met by over-provisioning the entire campus network (as was possible until about ten years ago). Instead, CIOs and campus leadership should adopt new, targeted strategies for meeting intra-campus CI needs that focus on targeted solutions for networking from labs and buildings with high data I/O and networking needs and the main campus connecting points to high-speed research networks.

There are two aspects of data movement and networking that are particularly important to the charge of the Task Force on Campus Bridging:

- For effective bridging in the sense of an individual researcher to a state, regional, or national CI facility it is essential that there be adequate end-to-end network performance.
- New data creation capabilities drive tremendous new demand for high bandwidth network connections between campuses to support campus bridging among like CI facilities over multiple administrative points of control or among different and varied CI facilities and multiple points of control.

End-to-end performance has been understood to be primarily a last-mile problem at least since 2000. perfSONAR [66] is a mature, well supported software tool that enables analysis and correction of network tuning problems that hinder the effective use of networking facilities. In many cases network tuning contributes to a substantial limiting of network performance far below the levels expected based on installed network connections. This leads to the following recommendation:

> ***Recommendation 7. Campuses should deploy and operate perfSONAR and related tools to systematically measure, debug, record, and display the measured performance.***

Even if optimal tuning of all networks were a given, however, many researchers could not effectively move data on and off campuses in order to manage and understand them because the campuses have inadequate connections to high-speed research networks. Part of this problem resides with university and college campus leadership, CIOs, and state funding levels. Even a small campus of a higher education research institution should have a connection to research networks to support research and research education that is at least 1 Gbps now, and 10 Gbps by 2020 as a component of its basic infrastructure. This baseline will not, however, meet current, high priority data movement needs.

It is extremely important overall that the NSF focus attention on and, where appropriate, fund the scaling up of end-to-end data movement capabilities that match the growth in important data of long term value. New capabilities in dynamic allocation of bandwidth offered by Internet2 and National Lambda Rail offer the possibility of highly cost effective transfer of data within and among those networks via 10 Gbps dynamically allocated lambdas. However, it can be extremely expensive to fund a 10 Gbps connection from campus to one of these national backbone providers. The Task Force on Campus Bridging thus makes the following strategic recommendation:

> ***Recommendation 8. The National Science Foundation should create a new program funding high-speed (currently 10 Gbps) connections from campuses to the nearest landing point for a national network backbone. The design of these connections must include support for dynamic network provisioning services and must be engineered to support rapid movement of large scientific data sets.***

The overall objective of networks should be to connect the researcher to the resources the researcher needs to perform their research. This may sometimes require working outside or beyond existing networks. For example, TeraGrid Resource Partners have their own dedicated high performance network (the TeraGrid "backplane" network) and data transfer tools (e.g., gridftp) that can provide the highest performance. There is no campus coordination for high performance data transfer, however, and many high performance data transfer tools are not available at campus researcher desktops nor in the default toolsets they would normally deploy.

The concept of "Data Intensive Networks," as described by Marti and Almes' position paper on page 101, gets at the peculiarities of the data-centric networks needed for research cyberinfrastructure and campus bridging, and without specifically defining this concept suggest it can fruitfully be developed and implemented at many campuses throughout the US. The network that researchers need – a Data Intensive Network – has different properties from the regular Internet. The capabilities needed include:

- Multi-Gb/s within campus and across the nation.
- Flows of large data objects (our campuses are connected with an network designed to move megabyte (MB) objects, but instruments and other CI drivers require moving terabyte (TB) objects. We need an end-to-end network for these TB objects – otherwise we cannot share them effectively).
- Ability to support effective, high quality, remote visualization.

Over the last several decades, the bits per second rate for a given circuit/lambda has grown, but this growth is now much slower than during the 1980s. We enjoy, however, a healthy growth (via dense wavelength division multiplexing (DWDM)) in the number of lambdas per physical circuit and thus the aggregate bandwidth. DWDM networking infrastructure permits (and perhaps demands) multiple mission-specific networks. These observations suggest that future sustained growth in end-to-end network performance will require increased use of parallelism, and this will require new network architecture.

The key is that campus CI resources must be connected to the national high-speed research networks unimpeded by compromises due to the conventional campus LAN mission. Further, the Data Intensive Network of one campus must be interoperable, in the performance sense, with those at other campuses.

Campuses and the US research community generally should lead and challenge the national networking community to build networks connecting the nation's universities so that moving of large datasets is experienced as easy and fast, so that location is not a barrier to data sharing and to access to national CI resources, and so that campus CI and NSF-sponsored CI can work together flexibly. Internet2, National LambdaRail, and other networking organizations are of great value, since CIOs work with them and they can help organize the community. In challenging the network and network research community to address some of these pressing research cyberinfrastructure challenges, the workshop participants suggested the following particular areas of attention:

- Focus attention beyond issues that are purely networking problems, and thus, e.g., lead in building distributed systems such as wide-area file systems.
- Focus on networks holistically; i.e., do not focus on the backbone part of the network to the exclusion of backbone, regional, and campus combinations.
- Focus more on supporting research and in particular focus more attentively on the flow of data among campuses, the Tier 2 centers and data collections. Understanding the current flow of large data and the pent up demand for increased data flow would be useful.

If we fail to achieve dramatic growth in bandwidth, we will be telling researchers "to access the resource, you have to go there." The bandwidth of a physical storage drive shipped overnight has long been well understood. The shipping of data on physical devices will become the network of choice for a growing number of users unless there is systematic change in networking technology implementation at the campus, regional, and national levels.

SOUTHEASTERN UNIVERSITIES RESEARCH ASSOCIATION SURA

SURA regional campus bridging initiatives

SURA (http://www.sura.org/) is a consortium of over 60 leading research institutions. Established in 1980 as a non-stock, non-profit corporation, SURA serves as an entity through which colleges, universities, and other organizations cooperate with one another and with government and industry in acquiring, developing and using laboratories and other research facilities, and in furthering knowledge and the application of that knowledge in the physical, biological and other natural sciences and engineering. SURA operates the Thomas Jefferson National Accelerator Facility (http://www.jlab.org/) for the US Department of Energy through Jefferson Science Associates (http://www.jsallc.org/index.html), a SURA/Computer Sciences Corporation joint venture.

SuperRegional Coastal Modeling Testbed

A $4 million grant from the National Oceanic and Atmospheric Administration is helping SURA evaluate the readiness of marine forecasts, such as flooding from storm surge or seasonal dead zones. Focused along the Atlantic and Gulf of Mexico coasts, the effort is working to improve the ability of computational models to provide forecasts for use by emergency managers, scientific researchers and the general public. SURA has successfully obtained and is managing over 6 million Service Units on TeraGrid and LONI high performance computing systems in support of the computational needs of the coastal modeling community.

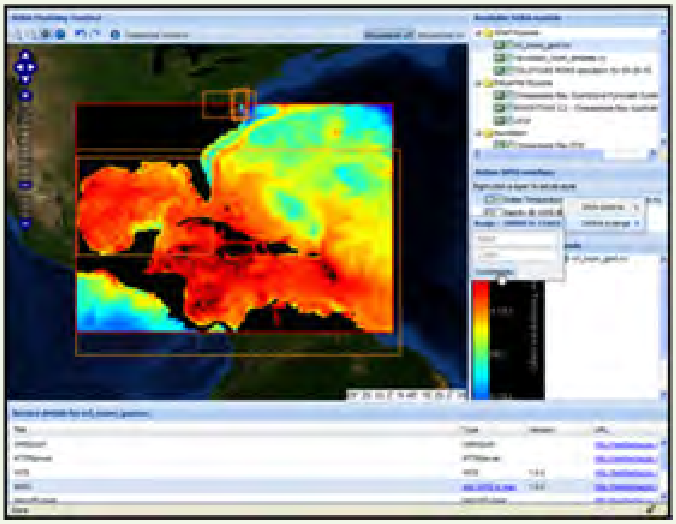


Figure D. A web-based catalog of testbed models and model data facilitates model evalucation and serves to integrate models and the critical observations needed to verify the model outputs.

Figure E. SURAgrid members.

SURAgrid (http://www.suragrid.org/) Evolving from participation in the NSF Middleware Initiative (NMI) Integration Testbed, SURAgrid is a consortium of organizations collaborating and combining resources to help bring grid technology to the level of seamless, shared infrastructure. This includes providing resources to researchers through the SURAgrid Portal and as an Open Science Grid Virtual Organization.

Education and Outreach. SURA has developed and delivered timely advanced application and information technology workshops for nearly a decade. This has included topics such as computational chemistry, digital video, biogrid development, middleware deployment and a cyberinfrastructure workshop series with an initial focus on grid applications and deployment. In support of the Texas Advanced Computing Center-led TeraGrid XD Visualization and Data Analysis (VDA) Services, SURA is leading an effort to identify established and emerging computational science programs at Minority Serving Institutions (MSIs); promote the use of VDA services to researchers from under-represented groups; and coordinate VDA training for researchers from MSIs and under-represented groups.

Innovative Corporate Partnerships. Leveraging the well-developed corporate relationships of its membership, SURA has engaged interested commercial entities in formal partnerships that enhance existing and planned SURA IT programs. A historical partnership with AT&T included a donation to SURA of the use of 8,000 miles of dark fiber optic cable on AT&T's national infrastructure. A partnership with IBM has provided SURAgrid participants with access to aggressive discount programs on IBM and Dell high performance computing hardware. This has also included the development of mutually beneficial R&D relationships and access by our members to IBM's expertise in high performance and grid computing.

SURAnet & SURA Crossroads. In the mid-1980's, SURA developed and managed SURAnet, a multi-state, regional network that connected much of the southeastern US to the former NSFNet prior to commercialization of the Internet. SURA transferred this service to private industry through the sale of SURANet in 1995 but maintained a long-term understanding of the value of a connected South. SURA built upon this understanding to bring members together through the SURA Crossroads Initiative, which supported a number of activities that leveraged the collective monetary, physical, and intellectual assets of SURA member institutions to pursue new models for advancing connectivity within and beyond the region.

Text and images provided by Gary Crane, SURA.

7. Wide area file systems

There is a significant need for a wide-area federated distributed file system to support data-intensive and data-driven open scientific research in the United States. Such a system would contribute to research effectiveness and new research breakthroughs in several ways:

- Help collaborators share data
- Provide a single global namespace (as demonstrated by the Andrew File System)
- Promote standard interfaces, while allowing some diversity of mechanism
- Leverage and drive further strengthening of emerging federated identity, authentication, and authorization infrastructure. (This system will specifically stress our understanding of authorization.)
- Drive the further strengthening of approaches to file system security
- Drive the further strengthening of algorithms for managing distributed file systems with multiple replicas of a given file, particularly in areas such as performance and replica coherence
- Leverage our current understanding of replication transparency and migration transparency
- Provide a proper measure of autonomy to the people who own given datasets
- Create computer network design possibilities
- Facilitate NSF investment
- Create a platform on which deployment of SRB and/or iRODS could benefit a broad community.
- Make available to a broader community sophisticated file system designs that optimize performance for large files or for large numbers of files or for specific scientific database needs.

Such a system might not be strictly POSIX compliant [67], but should build on the ability that many researchers have in using conventional directory/file systems. REDDNET [68] demonstrates one approach to a wide-area file system. Among other interesting aspects, it avoids slavishly following POSIX in order to realistically deal with wide-area latency.

The iRODS system [69] enables a wide-area data system that supports replication, heterogeneity of local file systems, and automated interactions between storage and applications, including with Software Agents. The Ocean Observatories Initiative mentioned earlier is a recent example of a driving application of these techniques.

The Data Capacitor [28] and GPFS [70] have both been used as file systems accessible via remote file system mount over wide area networks. To date, the Data Capacitor has had the largest adoption (in terms of number of TeraGrid Resource Partners accessing it), in part as a result of past licensing issues with GPFS.

These and other systems have demonstrated key wide-area file system technologies.

Recommendation 9. The National Science Foundation should fund the architecting, implementation, and operations of a wide-area federated distributed file system for use by the US open research community. The resulting system should support federated identity, very high-speed transfer of data (files or blocks) among major repository components of the system, and replication of files to further robustness and performance.

Such an activity must be done in collaboration with national CI partners (including the TeraGrid, the Open Science Grid, Internet2, and National LambdaRail) in order to be successful. As mentioned earlier, there would be tremendous value to a simple, standards-based, highly performing, and secure Campus Bridging Software Stack. Such a software stack would be particularly useful if it included software for accessing such a global wide area file system. To be most useful and most easily usable by individual researchers on campuses, such a software stack might leverage "Filesystems in Userspace" (FUSE) [71] to make it easy to map remote resources to local resources. Regardless of technical details, an objective for a wide area file system and the software stack that supports it would be for remote files to appear as if they were mounted on the desktop.

Project	Approach	Software used	References
Amazon Web Services	Cloud-based storage	Amazon Elastic Block Store (proprietary)	[72]
Data Capacitor	Open source, wide area file system	Lustre (requires kernel patch)	[73, 74]
Data Oasis	Parallel file system	Lustre	[45]
DropBox	Commercial cloud-based dropbox	Google sites (proprietary)	[75, 76]
Genesis II	Standards – based data grid	Genesis II	[77, 78]
Globus Online	Cloud-hosted data transfer - Point to point movement of large files	Globus, gredFTP	[79]
Google sites	Cloud-based storage		[76]
GPFS	Proprietary wide area file systems	General Parallel File System (proprietary)	[80]
iRODS (integrated Rule-Oriented Data-management System)	Separates logical names from physical location for distributed (and replicated) management of files	iRODS	[69]
OpenAFS	Distributed tree of file systems; derivative of Andrew File System	OpenAFS	[81]
Open Science Grid	Small units of data are moved to a single node for analysis		[17]
Penguin Data Caddy	Forget the network, just ship it on a physical storage device	None	[82]
REDDnet	Best effort "working storage" to manage the logistics of moving and staging large amounts of data	Suite of open source software	[68]

Table 3. A sampling of current approaches to wide area file movement.

8. Data and commercial cloud services

Commercial cloud service providers offer an interesting variety of storage options. In the long run, such cloud systems may offer interesting and valuable options for data-centric computing in the future. One aspect of use of data clouds discussed at length at the workshop was the obstacle to use presented by the licensing terms of some commercial cloud providers. Since the time of those discussions, the cloud provider discussed most particularly has changed its license terms – perhaps in part thanks to this discussion. In any event, the workshop participants are grateful for this change.

There still remain obstacles to use of commercial cloud data storage systems. Key concerns remaining today cost of data transport and regulatory compliance (export control matters, since there are presently no guarantees that data stored in a commercial cloud will be kept within the legal boundaries of the US which may created export control, HIPAA, or FERMA compliance challenges).

9. Data archives

For the past many decades, one could donate a book to a university library and it could be there forever. It is now problematic to donate a data collection or dataset to a (campus) data repository. The importance of being able to retain data in digital formats so as to enable replicability of scientific endeavors was described profoundly by Clifford Lynch during a workshop on cyberinfrastructure software sustainability held in 2009, and summarized in Stewart et al. [5].

The issue of long term storage of important research and one-of-a-kind data collections and the capabilities to serve this data to the research community must be carefully considered, both in the context of data generated by instruments (e.g., Digital Sky Surveys) and data generated by computational means.

Humans are now collecting data that is highly accurate, of potential perpetual interest, and impossible to replace if lost. Examples include weather and climate data, or genome sequence data of species in the wild. It is without doubt that the global environment is changing. To tell how and how much it will be essential to keep weather data indefinitely. To tell what impact such changes have had on the natural environment, it will be essential to keep genome sequence data collected from the wild. Species genetics may change over time and species are becoming extinct at a rapid rate (in terms of comparison with evolutionary history). Preserving genome data will allow understanding of the changes in the living ecosystems around us. At the same time, it can't be the case that every bit of data collected should be preserved forever. Digital data are produced by all sorts of instruments at a rate that exceeds our ability to store them [44], and some of that data must be noise.

However, it takes more than preserving data to make data usable and useful in the future. In order to be useful over long periods of time, data must be preserved with sufficiently rich metadata that they can be understood by people who were not involved in collecting the data and are learning about the data from the metadata. The late Dr. Dick Repasky (Indiana University) was both a computer technologist and evolutionary ecologist. When conducting field research, he would carefully create data dictionaries for data sets, which were analyzed with scripts and the open source statistical package R running on an x86-based computer using the Linux operating system. At the end of a research project, he would make a tar file of the Linux distribution, version of R, and build scripts he used to create the OS environment in which the analysis was run; the data sets; the data dictionaries; the scripts that performed the data analysis; and the output of the data analysis that he had done. To replicate his analysis then, all one really needs is the .tar file, an x86 system or emulator, and any software package that can unpack a .tar file. This is an extremely powerful approach to scientific replicability. It's also a lot of effort. For data that already exist, some sort of after-the-fact description of data takes effort, time, and money. In the long run it will be essential to develop tools that work easily (automatically or nearly so) and in real time, while data are being collected, so that the creation of metadata happens in real time. Examples of technology now being developed and used to perform this automatic metadata creation and provenance management function include Karma [83], XMC Cat [84], and Pegasus [85].

NSF guidance regarding data management plans [86] came out between the time this workshop was held and the time the report was finalized. This guidance is particularly helpful in moving toward better preservation and reusability of data collected with NSF funding. The option for researchers to include small amounts of funding for things like curation and archiving of data is a particularly helpful point. Examples of plans and guidance for development of data management plans are already in existence, including [87-89].

In addition to the matters of describing data, there is an important and ongoing discussion about what sorts of entities are responsible for maintaining data archives on an ongoing basis. The two major models are based on long-lived institutions and virtual organizations / communities. The institutional model typically focuses on libraries, harkening back to the model of the great Library of Alexandria. For example, At Princeton University, there is a recently inaugurated scheme whereby you can pay double the current cost of storage for a given dataset, and the university will commit to storing it "forever" [90]. The National Library of Medicine is an excellent model of the institutional approach as regards genomic and medical data. The other major model is based on communities and/or virtual organizations taking responsibility for long-term management and curation of data. An example of this model is the International Virtual Observatory Alliance [25]. These philosophically different approaches tend to be instantiated in technologically different ways. Institutionally-based approaches tend to focus on archival software such as Drupal [91]. Virtual organization-based and community-based approaches often focus on distributed data technology such as iRODS [69].

At the present, it is not well understood in the community whose responsibility data curation and management is, how to fund it, what architecture should be used for capturing data and metadata (including reliable provenance information), and what the best base technologies are. Experimentation with approaches will be important. NSF-funded efforts such as DataONE [22] and the Data Conservancy [23] will important parts of that. Case studies of researchers with large data movement needs would provide useful information to inform national discussions.

While experimentation is important, and there is uncertainty about what entities are responsible for maintenance, the right answer for the US research community must include NSF funding for general-purpose, long-term data archives. (This is arguably not the situation now.) This is at least in part a campus bridging problem, since for "the 4th Paradigm" [92] to become a widespread reality, researchers on campus will have to be able to access data, understand data, at times subset and selectively obtain portions of large remote data sets, and then return the results of their research to some sort of curated archive. Thus, the workshop participants make the following recommendation:

> ***Recommendation 10. The National Science Foundation should fund the architecting of costs-effective ways to archive and preserve data collections, and fund at least some facilities for archiving important data at the national level.***

10. Visualization

Visualization is a key technology in campus bridging and data issues in particular. This is particularly so as data sets get larger and larger. Many data sets are very expensive to move in toto from one location to another. Some data sets are too large to understand in any way other than visualization or statistical analysis. In these cases, the ability to create an image remotely – wherever the data live – and move or stream that image effectively back to the on-campus researcher is both a key campus bridging issue and often by far the best way to understand and extract knowledge from very large data sets.

Visualization has been stuck at or around 1 megapixel, with relatively little change in capabilities since the early 1980s. As with networking, visualization is an end-to-end problem where one of the ends is the optic nerve of the researcher. In this regard there has been considerable advance, with technologies such as VAPOR [93], developments toward high quality desktop 3D, and new systems for very high resolution such as the OptIPortal. The OptIPortal has demonstrated real-time 600 megapixel images, including stereo (described in slides included in Appendix 2 and in a special issue of Future Generation Computing Systems [94]). The NSF Advisory Committee on Cyberinfrastructure Task Force on Data and Visualization has prepared a report with a number of excellent recommendations regarding visualization and data matters [6]. From a campus bridging perspective, the workshop participants make one specific recommendation that is consistent with recommendations made in that report:

> ***Recommendation 11. The National Science Foundation should fund development of software tools and technology needed for effective remote visualization, and the NSF and institutions of higher education should fund the technology implementation and infrastructure needed for effective remote visualization.***

Visualization at Indiana University as a campus bridging case study

Indiana University's history of support for advanced visualization technologies over the past 25 years provides a useful example of the impact of campus bridging in a specific branch of cyberinfrastructure at a large, distributed institution. It illustrates the opportunities for innovation and democratization offered by the commoditization, reinforces the role of high-end "flagship" facilities, represents a sustainable blend of institutional support and grant funding, and lends greater evidence to the need for NSF to continue to invest in scalable software infrastructure for visualization. Over this time, visualization at IU has evolved through five distinct generations of technologies and support models, the first four of which can be thought to map to layers of the Branscomb Pyramid, with the fifth best representing the whole.

The first generation of visualization support centered on workstation-based graphics and video production. This generation is exemplified by the high-quality animations and interactive desktop applications developed by the Center for Innovative Computer Applications (CICA) from 1986 through 1996. These visualization delivery techniques remain important and viable today, especially with the proliferation of multi-core PCs with powerful GPUs, commodity stereo displays, and the growing standardization of Grid and Cloud technologies. However, desktop visualization and pre-rendered movies alone are not sufficient for displaying massive data sets, facilitating team collaboration, and encouraging interactive exploration.

Thus, the second generation of visualization support at IU began in 1997 with the creation of the Advanced Visualization Lab (AVL) and the installation of the University's first high-end, immersive virtual reality systems: a CAVE at IU Bloomington and an ImmersaDesk at Indiana University – Purdue University Indianapolis – facilities analogous to the "Center Supercomputers" layer in Figure 1. Those devices were funded in part by an NSF infrastructure grant and were of critical value to that initial set of research projects. Beyond the scope of that grant, these systems were highly successful in helping to generate interest in virtual reality, real-time graphics, group collaboration, and high-performance computing. Furthermore, they were invaluable in providing an accessible and visible "public face" for other advanced IT initiatives and in creating first- and second-order external funding effects. Since AVL was the University's first IT group to span campuses, the benefits of standardizing on common software and workflows for interoperability were immediately apparent.

Figure F. Renderings of IU's CAVE (left) and ImmersaDesk (center) and representative artisitic outreach event (right).

The third generation of visualization support began in 2001 with efforts to leverage advances in commodity graphics systems to supplement the high-end facilities and to address the major accessibility and usability issues of the second-generation technologies. The focus of this generation was the development and deployment of a variety of more affordable, mid-range technologies including stereoscopic displays, haptic devices, telecollaboration systems, 3D scanners, ultra-resolution displays, and small-scale graphics clusters. Notable among these systems was the IU-developed John-e-Box, which was a portable, large-format, passive stereo display system. Through licensing to a commercial producer and NSF instrumentation funding, IU was able to deploy 10 of these systems to labs, classrooms, studios, and galleries across four of its campuses. For those on campuses with high-end facilities, the John-e-Box created a technological bridge that eased access issues and spurred experimentation and innovation for faculty and students. For those on campuses without high-end facilities, it offered an introduction to the power of advanced visualization for education and research. For those reaching out to the general public and the next generation of scientists, the John-e-Box provided a way to convey the sense of excitement and engagement that scientists and artists experience.

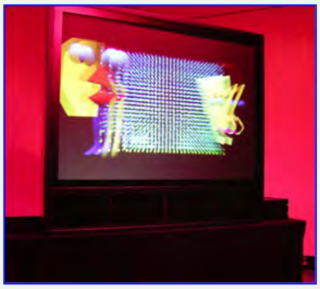


Figure G. Representative John-e-Box deployments: Chemistry department (left), art gallery installation (center), elementary school outreach (right)

The fourth generation of support coincided with IU joining the TeraGrid and the push to integrate visualization resources on all levels with the computational power represented at the top layers of Figure 1. While IU realized some moderate successes through its Condor-based rendering cluster to facilitate batch rendering and the use of its Lustre-based Data Capacitor to bring massive data sets to the visualization system, the real progress in this area was spurred by the development of scalable, open source, general-purpose visualization tools such as ParaView (by Kitware, Inc.) and VisIt (by Lawrence Livermore National Lab) that operate through a client-server paradigm and which enable a range of remote rendering options depending on the needs or constraints of the display, the data, and the network. At present, IU is partnering with Idaho National Labs (INL) and Kitware to develop an immersive version of ParaView that will properly and fully integrate the benefits of immersive virtual reality displays like the CAVE with the power, flexibility, and usability of ParaView.

The current and fifth generation of visualization support at IU is best characterized as a holistic approach where continued development, deployment, and interoperability across all levels is resulting in the blurring of the distinction between small-scale, mid-range, and high-end systems. The focus is on the delivery of a uniform, high-quality, and device-appropriate user experience across the spectrum of technologies. Casual users can utilize almost any familiar desktop application on an advanced display in order to access additional benefits of such as greater resolution, stereoscopic viewing, multi-touch interaction, or tracking, while advanced users who are accustomed to developing custom tools can still develop and use them to maximally utilize these systems. For example, the University's new Visualization and Collaboration Theater sacrifices some of the immersion of a full CAVE but gains back greater utilization for teleconferencing, research presentations, and stereoscopic investigation along with greater usability by presenting the single-system, multi-monitor metaphor with most users are familiar. Users of the University's retired CAVE system can utilize their software with full tracking and synchronized multi-screen immersion. The new IQ-wall balances the benefits of ultra-resolution offered by software tools such as SAGE and Chromium with the usability offered by a single system image and "big desktop" metaphor. The successor to the John-e-Box, the IQ-station (developed in conjunction with INL, Desert Research Institute, and University of California – Davis), leverages the rapidly advancing technologies of 3D displays, game controllers, and low-cost tracking systems to provide a system that appears like a 3D TV to some, an advanced gaming system to others, and an immersive virtual reality system to experts. Key to the success and supportability of all these systems at IU is the use of the University's common authentication/ system and the ability to mount user files from IU's OpenAFS implementation.

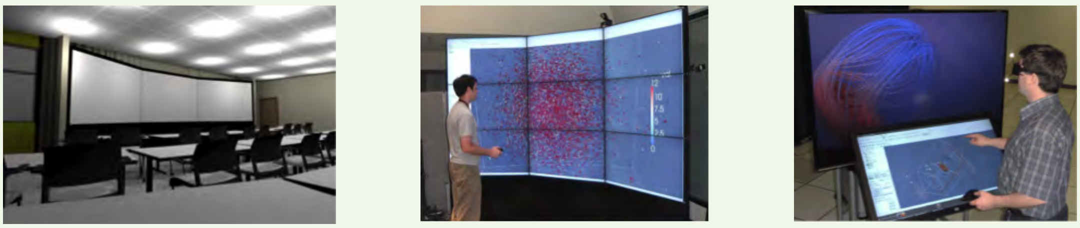


Figure H. IU's new Visualization and Collaboration Theater (left); IQ-Wall 10-megapixel display (center); IQ-station low-cost immersive visualization system (right)

Text and images provided by Eric Wernert, Indiana University.

11. Workforce and education

It is a rational (if not altogether satisfactory) expectation that a large fraction of the research faculty of tomorrow will come from the research campuses of today. However, there is a tremendous need for trained, professional staff who are or who can support computational scientists doing data-intensive computing on campuses. An earlier recommendation focused specifically on funding levels for the TeraGrid Campus Champions program [33], but this is just one example of one type of support that is needed. Much more support for computationally and data-enabled scientific research is needed for the US to maintain its pace of innovation, position in the global economic situation, and address compelling problems that face the US and humankind generally.

The workshop participants felt that the level of staff support for cyberinfrastructure and computational and data-enabled research is lower, relative to demand, than was the case 10 or 20 years ago. While there are certainly more staff (professional staff and research staff) supporting use of cyberinfrastructure, the availability of such staff has risen much more slowly than need and demand.

NSF policies regarding postdoctoral mentoring plans will likely have as a side effect a new set of challenges in terms of employing such staff. These new policies will likely result in a decrease in the use of the title 'postdoctoral fellow' and increases in ranks of research faculty, permanent research staff positions, and professional staff supporting cyberinfrastructure. (The NSF guidance about postdoctoral mentoring plans does not specifically create any limits on the time one may be a postdoctoral fellow but does call for mentoring plans to identify steps that a mentor will take to enable a postdoctoral fellow to move into a faculty position. After a person has been a postdoctoral fellow for several years, "this time for sure" becomes an increasingly less credible claim about moving a person into a faculty position.)

The gap perceived by workshop participants between the need for staff and the availability of staff supporting computational and data-intensive science engineering has a variety of sources, beginning with the digital divide as experienced socially by young children as they grow up. The workshop participants strongly endorsed recommendations made in other reports (e.g. [95-99]) to improve STEM education at the K-12 levels, and to improve attractiveness of STEM and computing as topics of study and careers.

Changing the supply of well-educated students entering the work force through getting more students who are now in primary school interested in STEM disciplines will take decades. Steps that can be enacted now to change the aspirations and career goals of students now in primary schools will take years to decades to have an impact on the supply of cyberinfrastructure professionals. Yet they are critically important for the long-term success and global competitiveness of the US.

In the shorter term there are other steps that can be taken to improve the supply of well-trained cyberinfrastructure technical experts. Small, four-year campuses constitute a prime and greatly underutilized source of talent that can be educated and cultivated to form a large and important part of a highly talented and skilled 21st century workforce, particularly the professional and academic research component of such a workforce. Education in computational and data-enabled science and engineering is a tremendous challenge at smaller institutions. While the basic principles in this area may persist for some time, the underlying cyberinfrastructure changes so rapidly that any curriculum involving practical use of CI must be updated once every two to three years in order to be up to date. Release time for faculty to create and update curriculum at most smaller schools is very scarce. MIT's release of engineering curriculum materials for use throughout the US revolutionized education in engineering [100]. There is a draft undergraduate curriculum for parallel and distributed computing [101]. However, there does not yet exist a general and widely accepted curriculum for computational and data-enabled science and engineering, nor for cyberinfrastructure software. The workshop participants strongly supported STEM efforts, and explicit inclusion of computing as a supplement to STEM education. The following text is quoted from Dreher et al. [3] and speaks to this point well:

> *A recent report from the President's Council of Advisors on Science and Technology (PCAST) makes several important recommendations regarding workforce development aimed at increasing the supply of professionals with bachelor's, master's, and doctoral degrees in networking and information technology (NIT). While the PCAST recommendations focus on actions that should increase the supply of skilled professionals in the United States in the short term, it is critically important that the academic community not only embrace these recommendations but also expand programs such as STEM (Science, Technology, Engineering, and Mathematics) in addressing long-term needs. In fact, while STEM shows signs of success, we need to continue to strengthen and expand the emphasis on STEM disciplines in elementary and secondary education so as to increase the absolute numbers and relative percentages of high school graduates who plan to enter college in an NIT-related discipline. Furthermore, we should expand the scope of STEM by recognizing that IT is a universal enabler and including computing as a core component for a C-STEM program.*

It takes more than education to make a work force, however – it takes funding for jobs, and it takes work conditions that keep people in the research workforce. The workshop participants strongly affirmed need for more funding and more reliable funding for staff. Long-term funding continuity for major and important CI projects is critical in part because of its impact on the workforce. The corrosive effects of uncertainty when staff are funded on short-term grant awards (two or three years at a time) has been noted over and over. Thus, we here note explicitly that a critical part of NSF and university support for cyberinfrastructure and campus bridging activities must be support for this "humanware" aspect of cyberinfrastructure. If we are as a nation to retain, within the fields of cyberinfrastructure and computational and data-enabled science and engineering, the best and brightest experts then the value proposition they face as individuals must be such that it is rational, and consistent with a good quality of life, to pursue and maintain a career in these fields. The workshop participants endorsed by strong consensus the following recommendations:

> ***Recommendation 12. The National Science Foundation should encourage and fund the training of more researchers of all types (especially staff) in computational and data-intensive science and engineering.***

Recommendation 13. The National Science Foundation should provide more funding for staff supporting use of cyberinfrastructure in research and in particular should provide funding that is more stable and predictable over time.

Minority Serving Institutions Cyberinfrastructure Empowerment Coalition

The Minority Serving Institutions Cyberinfrastructure Empowerment Coalition (MSI-CIEC) was established to build and enhance the social and technological mechanisms for meaningful engagement of MSIs in cyberinfrastructure (CI).

Engaging MSIs is an efficient and effective way of reaching the growing number of underrepresented minority college students – the next generation of scientists and engineers. Although only a relatively small percentage of colleges and universities in the country, MSIs serve a much greater proportion of underrepresented minority students. For example, Hispanic Serving Institutions (HSIs) are less than 10% of the higher education institutions in the country, but produce about a third of Hispanic science and engineering (S&E) baccalaureates; nine of the top ten baccalaureate alma maters of African American S&E Ph.D.'s have consistently been Historically Black Colleges and Universities (HBCUs), and Tribal Colleges have generated one of the largest pools of American Indian students that go on to complete a Ph.D. in science, technology, engineering, and mathematics (STEM) disciplines) [1-4].

MSI-CIEC is:

- Developing the CI "middleware" resources to encourage, broker, enable, and manage meaningful CI initiatives involving MSI collaborations for the use, support, deployment, development, and design of CI to enable the advancement of e-science research and education. It promotes the development of the nation's diverse STEM workforce, including the current and next generation of the STEM professoriate in an increasingly diverse society.
- Fostering a dynamic community of learning and practice, a CI-enabled distributed research and education network providing e-science education and research opportunities to MSI faculty and students. It is exploiting the synergies between CI for science and engineering and the environments supporting electronic business and communities, enabling MSIs both as national research institutions and as regional economic development hubs.
- Providing a broadly systemic approach to reaching underrepresented minority students and engaging this nation's American Indian Tribal Colleges and Universities (TCUs), HSIs, and HBCUs in the exploration, dissemination and adoption of CI tools, services and initiatives supporting research and education.

Given the pervasiveness of computing in S&E and our society more broadly, MSI-CIEC joins Nora Sabelli of SRI International, and Geoffrey Fox of Indiana University in recognizing that computing, like education, is a civil right, and that CI is the great equalizer in S&E education and research. The three organizations (American Indian Higher Education Consortium (AIHEC), Hispanic Association of Colleges and Universities (HACU), and National Association for Equal Opportunity in Higher Education (NAFEO)) comprising MSI-CIEC and leading this project represent the broadest coalition of MSIs in American higher education (more than 334 such institutions). They partner closely with Indiana University, the San Diego Supercomputer Center (SDSC) and the University of Houston-Downtown and are supported by many CI projects and leaders, e.g. Geoffrey Fox of FutureGrid and Larry Smarr of CalIT2.

MSI-CIEC works through a variety of services, such as individual campus visits* and workshops†. Attendees of these activities include faculty, students, administrators and computer technology staff from MSI's. MSI-CIEC also helps MSI's identify and respond to proposal opportunities that could strengthen their CI participation.

Perhaps the most important feature of MSI-CIEC is that it is a US Minority Serving Institutions initiative for and by MSIs. It is helping to ensure that the nation has the benefit of the best minds and talents of the next generation of Americans! Information about MSI-CIEC is online at http://www.msi-ciec.info/

Text and image provided by: Alexander Ramírez, Hispanic Association of Colleges and Universities; Karl Barnes, National Association for Equal Opportunity in Higher Education; Al Kuslikis, American Indian Higher Education Consortium; and Geoffrey C. Fox, Indiana University.

References:

[1] National Science Board (NSB). (2004). Science and Engineering Indicators - 2004. Volume 1. NSB 04-1. Arlington, VA: National Science Foundation. http://www.nsf.gov/statistics/seind04/c2/c2h.htm

[2] Institute for Higher Education Policy. (2006). Championing Success: A Report on the Progress of Tribal College and University Alumni. Washington, D.C. http://www.ihep.org/Publications/publications-detail.cfm?id=5

[3] National Science Foundation (NSF), National Center for Science and Engineering Statistics (NCSES). (2011). Women, Minorities and Persons with Disabilities in Science and Engineering. http://www.nsf.gov/statistics/wmpd/

[4] National Science Foundation (NSF), Advisory Committee for Cyberinfrastructure (ACCI). (2011). Task Force on Cyberlearning and Workforce Development, Final Report. http://www.nsf.gov/od/oci/taskforces/TaskForceReport_Learning.pdf

---

* "Cyberinfrastructure Days," e.g., http://www.ci-train.org/training/cidays/uapb.html

† E.g., workshop on use of GreenLight visualization technology, http://www.sagecommons.org/index.php?option=com_content&view=article&id=97

12. Environment and economics

The cost of the electrical power needed to power (and cool) CI resources will become an increasingly significant issue. For example, when Blue Waters [102] is turned on, it will consume a significant portion of all the electrical power at University of Illinois at Urbana-Champaign. The common practice of making electrical energy appear to be without cost to individual labs or departments distorts decision making regarding most effective use of energy resources. It has been carefully demonstrated that centralization of cyberinfrastructure facilities increased overall efficiencies in energy [103]. As members of the community of scientists, cyberinfrastructure experts should be concerned about their impact on the global environment. "Greenness" can be thought of as including several components, including: To what extent does an activity contribute to the production of greenhouse gases? And to what extent is the net result of an activity beneficial or harmful to the global environment?

Extracting the greatest possible utility out of computing hardware over its life span should minimize the unnecessary production of greenhouse gases, since such plans would minimize energy used by systems that are idle. The 'most green' approach to use of computational resources may be to maximize its energy consumption through keeping it operating at maximum possible usage levels – at least as long as the research and education activities being carried out are meritorious. This gets to the second aspect of greenness: to what extent is the net result of an activity beneficial or harmful to the global environment? There are two ways then that recommendations made in this report should maximize the net benefit of campus cyberinfrastructure in terms of global benefits. Improvements suggested in this report should decrease barriers to full use of campus cyberinfrastructure, thus enabling maximal utility of cyberinfrastructure hardware over the course of its useful life, and supporting breakthrough and practical research enabling the development of human societies in ways that are in harmony with a healthy global environment.

Rather than make a specific recommendation regarding the environment and research cyberinfrastructure, the workshop participants noted that technology that promotes effective campus bridging – as defined at the beginning of this report – could also lead to use of cyberinfrastructure in ways that minimizes environmental impact. For example, computer networks can address both energy cost issues and carbon issues (related, but not the same) by allowing power-hungry CI resources to be placed where power is cheap and/or low in its carbon impact. This is codified as a finding, as follows:

> **Finding 2. Technological development and implementation of cyberinfrastructure in ways that promote effective campus bridging will also have the natural side effect of enabling cyberinfrastructure use to have the minimal possible impact on the global environment while promoting US research capabilities.**

13. References

[5] National Science Foundation Cyberinfrastructure Council. Cyberinfrastructure Vision for 21st Century Discovery. 2007. Available from: http://www.nsf.gov/pubs/2007/nsf0728/index.jsp [cited 31 Jan 2011]

[6] NSF Advisory Committee for Cyberinfrastructure Task Force on Campus Bridging. Final Report. March 2011. [PDF] Available from: http://www.nsf.gov/od/oci/taskforces/TaskForceReport_CampusBridging.pdf or http://pti.iu.edu/campusbridging/ [cited 31 Mar 2011]

[7] Dreher, P., V. Agarwala, S.C. Ahalt, G. Almes, S. Fratkin, T. Hauser, J. Odegard, J. Pepin and C.A. Stewart. "Developing a Coherent Cyberinfrastructure from Local Campuses to National Facilities: Challenges and Strategies". EDUCAUSE. 2009. Available from: http://www.educause.edu/Resources/DevelopingaCoherentCyberinfras/169441 or http://hdl.handle.net/2022/5122 [cited 8 Mar 2011]

[8] American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences. Our Cultural Commonwealth. 2006. [PDF] Available from: http://www.acls.org/cyberinfrastructure/ourculturalcommonwealth.pdf [cited 8 Apr 2011]

[9] Cyberinfrastructure Software Sustainability and Reusability Workshop Final Report. C.A. Stewart, G.T. Almes, D.S. McCaulay and B.C. Wheeler, eds., 2010. Available from: http://hdl.handle.net/2022/6701 [cited 31 Jan 2011]

[10] NSF Advisory Committee for Cyberinfrastructure Task Force on Data and Visualization. Final Report. March 2011. [PDF] Available from: http://www.nsf.gov/od/oci/taskforces/TaskForceReport_Data.pdf [cited 31 Mar 2011]

[11] Indiana University Pervasive Technology Institute. Campus Bridging. Available from: http://pti.iu.edu/CampusBridging [cited 31 Jan 2011]

[12] National Science Foundation. The Internet - The Launch of NSFNET. Available from: http://www.nsf.gov/about/history/nsf0050/internet/launch.htm [cited 6 Apr 2011]

[13] National Research Council Evolving the High Performance Computing and Communications Initiative to Support the Nation's Information Infrastructure. National Academies Press, Washington, D.C., 1995. Available from: http://www.nap.edu/catalog.php?record_id=4948 [cited 8 Apr 2011]

[14] Branscomb, L., T. Belytschko, P. Bridenbaugh, T. Chay, J. Dozier, G.S. Grest, E.F. Hayes, B. Honig, N. Lane, W.A. Lester, G.J. McRae, J.A. Sethian, B. Smith and M.K. Vernon. From Desktop to Teraflop: Exploiting the U.S. Lead in High Performance Computing - NSF Blue Ribbon Panel on High Performance Computing. 1993. Available from: http://www.nsf.gov/pubs/stis1993/nsb93205/nsb93205.txt [cited 31 Jan 2011]

[15] Berman, F. Beyond Branscomb. 2006. PPT. Available from: http://www.cs.utk.edu/~dongarra/ccgsc2006/Slides/talk06 Fran Berman.ppt [cited 31 Jan 2011]

[16] Creative Commons. Attribution-NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0). Available from: http://creativecommons.org/licenses/by-nc-nd/3.0/ [cited 30 Mar 2011]

[17] Global Lambda Integrated Facility. GLIF Maps. 2008. Available from: http://www.glif.is/publications/maps/ [cited 8 Apr 2011]

[18] Internet2. Home page. Available from: http://www.internet2.edu/ [cited 16 Aug 2010]

[19] National LambdaRail. Home page. Available from: http://www.nlr.net/ [cited 16 Aug 2010]

[20] TeraGrid. Home page. Available from: http://www.teragrid.org/ [cited 31 Jan 2011]

[21] Open Science Grid. Home page. Available from: http://www.opensciencegrid.org/ [cited 31 Jan 2011]

[22] SURAgrid. About SURAgrid. Available from: http://www.sura.org/programs/sura_grid.html [cited 8 Apr 2011]

[23] OptIPuter. Home page. Available from: http://www.optiputer.net/ [cited 8 Apr 2011]

[24] TeraGrid. Longhorn Visualization Portal. Available from: https://viz.teragrid.org/ [cited 8 Apr 2011]

[25] National Institute for Computational Sciences. Nautilus User Guide. Available from: http://www.nics.tennessee.edu/computing-resources/nautilus [cited 8 Apr 2011]

[26] DataONE. Home page. Available from: http://www.dataone.org/ [cited 8 Apr 2011]

[27] Data Conservancy. Home page. Available from: http://dataconservancy.org/ [cited 8 Apr 2011]

[28] HathiTrust. Home page. Available from: http://www.hathitrust.org/ [cited 23 Mar 2011]

[29] International Virtual Observatory Alliance. Home page. Available from: http://www.ivoa.net/ [cited 8 Apr 2011]

[30] DigCCur. Carolina Digital Curation Curriculum Project. Available from: http://www.ils.unc.edu/digccurr/ [cited 23 Mar 2011]

[31] University of Illinois at Urbana-Champaign. Center for Informatics Research in Science and Scholarship. Available from: http://cirss.lis.illinois.edu/ [cited 23 Mar 2011]

[32] Indiana University Pervasive Technology Institute. Data Capacitor. Available from: https://pti.iu.edu/dc/ [cited 16 Aug 2010]

[33] Stewart, B. NSFNET -- National Science Foundation Network. Living Internet, 2000. Available from: http://www.livinginternet.com/i/ii_nsfnet.htm [cited 31 Jan 2011]

[34] Atkins, D.E. The Ubiquitous TeraGrid: Everywhere but Nowhere. (Presentation) TeraGrid '07 (Madison, WI, June 6-8, 2007).

[35] National Science Foundation. Dear Colleague Letter: Cyberinfrastructure Framework for 21st Century Science and Engineering (CF21). 2010. Available from: http://www.nsf.gov/pubs/2010/nsf10015/nsf10015.jsp?org=OCI [cited 31 Jan 2011]

[36] Cornell University Center for Advanced Computing. Sustainable Funding and Business Models for Academic Cyberinfrastructure (CI) Facilities. Available from: http://www.cac.cornell.edu/SRCC/ [cited 31 Jan 2011]

[37] TeraGrid. Campus Champions. Available from: https://www.teragrid.org/web/eot/campus_champions [cited 21 Mar 2011]

[38] InCommon Federation. InCommon Identity and Access Management. Available from: http://www.incommonfederation.org/ [cited 31 Jan 2011]

[39] Indiana CTSI. Home page. Available from: http://www.indianactsi.org [cited 31 Jan 2011]

[40] Committee on Institutional Cooperation. CIC Home Page. Available from: http://www.cic.net/Home.aspx [cited 31 Jan 2011]

[41] Barnett, W., V. Welch, A. Walsh and C.A. Stewart. A Roadmap for Using NSF Cyberinfrastructure with InCommon. 2011. Available from: http://hdl.handle.net/2022/13024 or http://www.incommon.org/nsfroadmap.html [cited 28 Feb 2011]

[42] Barnett, W., V. Welch, A. Walsh and C.A. Stewart. A Roadmap for Using NSF Cyberinfrastructure with InCommon: Abbreviated Version. 2011. Available from: http://hdl.handle.net/2022/13025 or http://www.incommon.org/nsfroadmap.html [cited 28 Feb 2011]

[43] ProtectNetwork. Home page. Available from: http://www.protectnetwork.org/ [cited 31 Jan 2011]

[44] National Institutes of Health. NIH Federated Identity Service. Available from: http://cit.nih.gov/ServiceCatalog/FederatedIdentity.htm [cited 4 Mar 2011]

[45] CILogon Service. Home page. Available from: https://cilogon.org [cited 31 Jan 2011]

[46] International Grid Trust Federation. Home page. Available from: http://www.igtf.net/ [cited 31 Jan 2011]

[47] Kuhn, D.R., V.C. Hu, W.T. Polk and S.-J. Chang. Introduction to Public Key Technology and the Federal PKI Infrastructure. SP 800-32, National Institute of Standards and Technology, 2001. [PDF] Available from: http://csrc.nist.gov/publications/nistpubs/800-32/sp800-32.pdf [cited 31 Jan 2011]

[48] Gantz, J. and D. Reinsel. The Digital Universe Decade – Are You Ready? , 2010. [PDF] Available from: http://idcdocserv.com/925 (PDF) [cited 22 Apr 2011]

[49] UCSD Research Cyberinfrastructure Design Team. Blueprint for the Digital University. 2009. [PDF] Available from: http://research.ucsd.edu/documents/rcidt/RCIDTReportFinal2009.pdf [cited 29 Mar 2011]

[50] Gray, J. and P. Shenoy. Rules of Thumb in Data Engineering. In: Proceedings of 16th International Conference on Data Engineering (2000). IEEE Computer Society. Available from: http://research.microsoft.com/apps/pubs/default.aspx?id=68636 [cited 8 Apr 2011]

[51] Jevons, W.S. The coal question (3rd ed.). MacMillan and Co., Limited, London, 1906. Available from: http://books.google.com/books?hl=en&lr=&id=cUgPAAAAIAAJ&oi=fnd&pg=IA1&dq=jevons+coal+question [cited 22 Apr 2011]

[52] Giampietro, M. and K. Mayumi. The Jevons Paradox: The Evolution of Complex Adaptive Systems and the Challenge for Scientific Analysis. In: The Jevons Paradox and the Myth of Resource Efficiency Improvement. J. Polimeni, K. Mayumi and M. Giampietro, eds. Earthscan, 2008.

[53] LSST Corporation. Large Synoptic Survey Telescope. Available from: http://www.lsst.org/lsst [cited 22 Apr 2011]

[54] CERN. The Large Hadron Collider. Available from: http://public.web.cern.ch/public/en/lhc/lhc-en.html [cited 22 Apr 2011]

[55] WIYN Consortium. The One Degree Imager. Available from: http://www.noao.edu/wiyn/ODI/ [cited 22 Apr 2011]

[56] Institute of Global Environment and Society. Center for Ocean-Land-Atmosphere Studies (COLA). Available from: http://www.iges.org/cola.html [cited 22 Apr 2011]

[57] Consortium for Ocean Leadership. Ocean Observatories Initiative. Available from: http://www.oceanleadership.org/programs-and-partnerships/ocean-observing/ooi/ [cited 31 Jan 2011]

[58] Center for Remote Sensing of Ice Sheets (CReSIS). Home page. Available from: https://www.cresis.ku.edu/ [cited 9 Jun 2010]

[59] Linked Environments for Atmospheric Discovery. LEAD Portal. Available from: https://portal.leadproject.org/ [cited 22 Apr 2010]

[60] Indiana University. EVIA Digital Archive: Ethnographic Video for Instruction & Analysis. Available from: http://www.eviada.org/ [cited 31 Jan 2011]

[61] EarthScope. Home page. Available from: http://www.earthscope.org/ [cited 22 Apr 2011]

[62] Stewart, C.A. It's not a data deluge - it's worse than that. (Presentation) DIDC'10 (Chicago, IL, 22 Jun, 2010). Available from: http://hdl.handle.net/2022/13195 [cited 8 Apr 2011]

[63] BD Biosciences. BD Pathway bioimaging systems. Available from: http://www.bdbiosciences.com/instruments/pathway/systems/index.jsp [cited 2 May 2010]

[64] 454 Life Sciences. Home page. Available from: http://www.454.com/ [cited 23 Apr 2010]

[65] Illumina Inc. Genome Analyzer IIe. Available from: http://www.illumina.com/systems/genome_analyzer.ilmn [cited 2 May 2010]

[66] Applied Biosystems. SOLiD 3 Plus System (110V). Available from: https://products.appliedbiosystems.com/ab/en/US/adirect/ab?cmd=catProductDetail&productID=4444315&catID=604416&backButton=true [cited 2 May 2010]

[67] Molecular Devices. GenePix Professional 4200A. Available from: http://www.moleculardevices.com/pages/instruments/gn_genepix4200.html [cited 2 May 2010]

[68] Roche NimbleGen Inc. Hybridization System 4. Available from: http://www.nimblegen.com/products/instruments/hybridization/4-slide/index.html [cited 22 Apr 2011]

[69] Papadopoulos, P., L. Smarr, Y. Fainman and J. Ford. MRI:Development of Quartzite, a Campus-wide, Terabit-Class, Field-Programmable, Hybrid Switching Instrument for Comparative Studies (Award number 0421555). National Science Foundation, 2004. Available from: http://nsf.gov/awardsearch/showAward.do?AwardNumber=0421555 [cited 8 Apr 2011]

[70] perfSONAR. Home page. Available from: http://www.perfsonar.net/ [cited 8 Apr 2011]

[71] The Open Group. The Open Group Base Specifications Issue 7. 2008. Available from: http://pubs.opengroup.org/onlinepubs/9699919799/ [cited 22 Apr 2011]

[72] REDDnet. REDDnet: Enabling Data Intensive Science in the Wide Area. Available from: http://www.reddnet.org/mwiki/index.php/Main_Page [cited 22 Apr 2011]

[73] iRODS. IRODS: Data Grids, Digital Libraries, Persistent Archives, and Real-time Data Systems. Available from: https://www.irods.org/ [cited 28 Jan 2010]

[74] TeraGrid. GPFS-WAN. Available from: https://www.teragrid.org/web/user-support/gpfswan [cited 22 Apr 2011]

[75] FUSE: Filesystem in Userspace. Available from: http://fuse.sourceforge.net/ [cited 22 Apr 2011]

[76] Amazon Web Services. Amazon Elastic Block Store (EBS). Available from: http://aws.amazon.com/ebs/ [cited 22 Apr 2011]

[77] Simms, S.C., G.G. Pike and D. Balog. Wide Area Filesystem Performance Using Lustre on theTeraGrid. In: Proceedings of Proceedings of the TeraGrid 2007 Conference (Madison, WI, 2007). Available from: http://pti.iu.edu/pubs/wide-area-filesystem-performance-using-lustre-teragrid [cited 15 Dec 2009]

[78] Simms, S., M. Davy, B. Hammon, M. Link, C.A. Stewart, R. Bramley, B. Plale, D. Gannon, M.-H. Baik, S. Teige, J. Huffman, D. McMullen, D. Balog and G. Pike. All in a Day's Work: Advancing Data-Intensive Research with the Data Capacitor. In: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing (2006). ACM Press.

[79] Dropbox. Home page. Available from: http://www.dropbox.com/ [cited 22 Apr 2011]

[80] Google. Google Sites. Available from: http://sites.google.com/ [cited 22 Apr 2011]

[81] Virginia Center for Grid Research. The Genesis II Project. Available from: http://www.cs.virginia.edu/~vcgr/wiki/index.php/The_Genesis_II_Project [cited 22 Apr 2011]

[82] Morgan, M.M. and A.S. Grimshaw. Genesis II - Standards Based Grid Computing. In: Proceedings of the Seventh IEEE International Symposium on Cluster Computing and the Grid (2007). IEEE Computer Society. Available from: http://portal.acm.org/citation.cfm?id=1251971.1252272 [cited 22 Apr 2011]

[83] Globus Online. Home page. Available from: http://www.globusonline.org/ [cited 31 Jan 2011]

[84] IBM, I. General Parallel File System. Available from: http://www-03.ibm.com/systems/software/gpfs/ [cited 22 Apr 2011]

[85] OpenAFS. Home page. Available from: http://www.openafs.org/ [cited 22 Apr 2011]

[86] Penguin Computing. POD Environment. Available from: http://www.penguincomputing.com/POD/PODEnvironment [cited 22 Apr 2011]

[87] Pervasive Technology Institute at Indiana University Data to Insight Center. Karma Provenance Collection Tool. Available from: http://pti.iu.edu/d2i/provenance_karma [cited 8 Apr 2011]

[88] Pervasive Technology Institute at Indiana University Data to Insight Center. XML Metadata Concept Catalog (XMC Cat). Available from: http://pti.iu.edu/d2i/xmccat [cited 8 Apr 2011]

[89] Pegasus. Home page. Available from: http://pegasus.isi.edu/ [cited 29 Jan 2010]

[90] National Science Foundation. Grant Proposal Guide, January 2011. Available from: http://www.nsf.gov/publications/pub_summ.jsp?ods_key=gpg [cited 23 Mar 2011]

[91] National Science Foundation. Dissemination and Sharing of Research Results. 2010. Available from: http://www.nsf.gov/bfa/dias/policy/dmp.jsp [cited 31 Jan 2011]

[92] National Science Foundation. Data Management & Sharing Frequently Asked Questions (FAQs). Available from: http://www.nsf.gov/bfa/dias/policy/dmpfaqs.jsp [cited 8 Apr 2011]

[93] Indiana University Data Management Task Force. Indiana University Guidance on NSF Data Management Plans. 2011. Available from: http://hdl.handle.net/2022/13221 [cited 22 Apr 2011]

[94] Goldstein, S.J. and M. Ratliff. DataSpace: A Funding and Operational Model for Long-Term Preservation and Sharing of Research Data. 2010. Available from: http://arks.princeton.edu/ark:/88435/dsp01w6634361k [cited 8 Apr 2011]

[95] Drupal. Home page. Available from: http://drupal.org/ [cited 8 Apr 2011]

[96] The Fourth Paradigm: Data-Intensive Scientific Discovery. T. Hey, S. Tansley and K. Tolle, eds. Microsoft Research, Redmond, WA, 2009. Available from: http://research.microsoft.com/en-us/collaboration/fourthparadigm/ [cited 31 Jan 2011]

[97] VAPOR: Visualization and Analysis Platform for Ocean, Atmosphere, and Solar Researchers. Home page. Available from: http://www.vapor.ucar.edu/ [cited 25 Apr 2011]

[98] Smarr, L., M. Brown and C. de Laat. “Special section: OptIPlanet - The OptIPuter global collaboratory”. Future Generation Computer Systems, 25(2). 2009. Available from: http://www.sciencedirect.com/science/journal/0167739X [cited 8 Apr 2011]

[99] Members of the 2005 “Rising Above the Gathering Storm” Committee. Rising Above the Gathering Storm, Revisited: Rapidly Approaching Category 5. 2010. Available from: http://www.nap.edu/catalog.php?record_id=12999 [cited 31 Jan 2011]

[100] Committee on Prospering in the Global Economy of the 21st Century. Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future. 2007. Available from: http://www.nap.edu/catalog.php?record_id=11463 [cited 31 Jan 2011]

[101] National Science Board. Broadening Participation in Science and Engineering Research and Education: Workshop Proceedings 2005. Available from: http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsb0472 [cited 22 Apr 2011]

[102] Congressional Commission on the Advancement of Women and Minorities in Science, Engineering and Technology Development. Land of Plenty: Diversity as America's Competitive Edge in Science, Engineering and Technology 2000. Available from: http://www.nsf.gov/publications/pub_summ.jsp?ods_key=cawmset0409 [cited 22 Apr 2011]

[103] National Science Board. National Action Plan for Addressing the Critical Needs of the U.S. Science, Technology, Engineering, and Mathematics Education System. 2007. [PDF] Available from: http://www.nsf.gov/nsb/documents/2007/stem_action.pdf [cited 22 Apr 2011]

[104] Massachusetts Institute of Technology. MIT OpenCourseware. Available from: http://ocw.mit.edu/index.htm [cited 28 Feb 2011]

[105] NSF/IEEE-TCPP Curriculum Initiative on Parallel and Distributed Computing – Core Topics for Undergraduates. Available from: http://www.cs.gsu.edu/~tcpp/curriculum/ [cited 2 Mar 2011]

[106] NCSA. About the Blue Waters project. Available from: http://www.ncsa.illinois.edu/BlueWaters/ [cited 31 Jan 2011]

[107] Hacker, T. and B. Wheeler. "Making Research Cyberinfrastructure a Strategic Choice". Educause Quarterly, 30(I), 21-29. 2007. Available from: http://net.educause.edu/ir/library/pdf/eqm0713.pdf [cited 31 Jan 2011]

Appendix 1. Workshop position papers

April 1, 2010
Position Paper for NSF Campus Bridging Workshop

The OptIPortal, a Scalable Visualization, Storage, and Computing Termination Device for High Bandwidth Campus Bridging

(Excerpted from Future Generation Computer Systems/The International Journal of Grid Computing: Theory, Methods and Applications, Elsevier B.V., Vol 25, issue 2, February 2009, pp. 114-123)

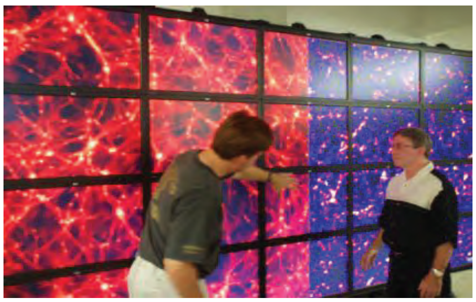
Authors: Thomas A. DeFanti[1], Jason Leigh[3], Luc Renambot[3], Byungil Jeong[5], Alan Verlo[3], Lance Long[3], Maxine Brown[3], Dan Sandin[3], Venkatram Vishwanath[6], Qian Liu[1], Mason Katz[2], Phil Papadopoulos[2], Joseph Keefe[1], Greg Hidley[1], Greg Dawe[1], Ian Kaufman[1], Bryan Glogowski[1], Kai-Uwe Doerr[1], Javier Girado[4], Jurgen P. Schulze[1], Falko Kuester[1], and Larry Smarr[1]

Affiliations:
[1] California Institute for Telecommunications and Information Technology (Calit2), University of California San Diego (UCSD)
[2] San Diego Supercomputer Center, University of California San Diego (UCSD)
[3] Electronic Visualization Laboratory (EVL), University of Illinois at Chicago (UIC)
[4] Qualcomm, Inc.
[5] Texas Advanced Computing Center (TACC), University of Texas, Austin
[6] Argonne National Laboratory

As dedicated fiber optics are deployed on campuses to bridge between data-intensive end-users and regional or national-scale optical networks, the data flow into a user's lab will go up by two-to-three orders of magnitude. This means that the shared internet termination device, the PC or server, must be scaled up as well to maintain proper impedance matching with the data flow. This scaling must be not only in storage, but also in computing power and visualization "pixel real estate" to enable scalable analysis.

Fortunately, the National Science Foundation (NSF) funded the OptIPuter project for eight years (2002-2009), and one of its research results is the OptIPortal, just such a scalable termination device. This paper describes the software systems that have been developed for the OptIPortal and gives pointers to where one can download the software and locate recipes for the hardware requirements. The main point of the OptIPuter project was to examine a "future" in which networking was not a bottleneck to local, regional, national and international computing. This is one of the key goals for NSF's campus bridging program. OptIPortals are designed to allow collaborative sharing over 1-10 Gigabit/second networks of extremely high-resolution graphic output, as well as video streams.

The OptIPortal is constructed as a tiled display wall. OptIPortals typically consist of an array of 4 to 100 LCD display panels (1- 4-megapixels each), driven by an appropriately sized PC or cluster of PCs with optimized graphics processors and



network interface cards. Rather than exist as one-of-a-kind laboratory prototypes, OptIPortals are designed to be openly and widely replicated, balancing the state of the art of PCs, graphic processing, networks, servers, software, middleware, and user interfaces, and installed in the context of a laboratory or office conference room. Some feature 3D stereo display panels (see NexCAVE and REVE below). They represent a constantly evolving technology. It is estimated that ~100 OptIPortals have been built globally and are in active use.

Some OptIPortals build on NSF's investment in the Rocks software system, which allows an end-user to easily install software across one's cluster. Since Rocks is the software environment upon which these OptIPortals are based, the hardware requirements for the OptIPortal are essentially those for Rocks, once the choice of display is made. Most of the deployments of OptIPortals have been done on commodity hardware, running Intel or AMD processors. Configurations are possible in which each computer in the cluster can drive one, two or more displays, depending on the performance and capabilities of the chosen graphics interface. OptIPortals can be optimized for specific functionality in terms of processor speed, network bandwidth, storage capacity, memory availability, and cost.

Rocks provides an easy way to configure an OptIPortal's display cluster, though OptIPortal middleware scales across different operating systems, operating system flavors and heterogeneous clusters. The middleware hides OS specific aspects and provides a cross-platform API. Locally available resources, such as the number of available graphics cards, displays and associated capabilities (resolution, swap and frame synchronization, etc.) can be probed at the device driver or the window manager level, allowing the middleware to report and adapt to hardware capabilities. Considering the number of PCs in a typical OptIPortal, mean time to failure becomes an important parameter when selecting cluster management strategies. From a system administrator's perspective, Rocks-based systems are easy to manage, largely by pruning system management overhead down to a single node.

Middleware and applications leveraging OptIPortal technology can be grouped into three major categories: stream-centric techniques, parallel distributed rendering techniques, and hybrid systems combining distributed real-time rendering and streaming within the same context. These in turn can scale from low-level visual content distribution approaches to high-performance parallel real-time rendering engines with multithread CPU support and GPU-based hardware acceleration.

OptIPortal head nodes and graphics nodes are networked together with 1 Gb/s or 10 Gb/s switches and network interface cards (NICs). Inexpensive switches allow onboard 1 Gb/s ports to be easily used, usually with a 1 Gb/s or 10 Gb/s uplink to the servers/campus networks. More expensive switches (e.g., Arista) and 10 Gb/s NICs allow much faster loading of large images and models, and, of course, facilitate streaming HD and 4K video. The latest motherboards now support up to 4 dual-ported graphics cards (GPUs), which will drive 8 two or four megapixel displays, but the load on the PC and NIC becomes quite high (just like putting a lot of disks on a PC would). One excellent feature of such systems is that GPUs can have up to 480 graphics cores and 1.792GB each, which is 1920 CUDA-programmable processors and 7GB of memory per PC. The optimal choice of OptIPortal motherboards and their NICs and GPUs is a constantly challenging design task.

1. Stream-Based Systems

SAGE (Scalable Adaptive Graphics Environment), initially funded by the OptIPuter award and now funded by NSF to harden and deploy to its growing user community, targets especially high-resolution tiled display systems, which can potentially cover all the walls and tabletop surfaces in a room, and which are interconnected to data sources and/or other OptIPortals with multi-10Gb/s optical networks. It operates on the assumption that as wall sizes increase, multiple users will naturally find a need to make full use of the available resolution to juxtapose multiple visuals and interact with them at the same time. It also assumes that it is possible for any type of application, given the appropriate middleware, to send a pixel stream to the SAGE tiled display.

SAGE middleware directs each of the incoming pixel streams from an application to the correct portion of a tiled wall allowing the system to scale to any number of streams and tiles. More importantly, it allows multiple applications on multiple distributed rendering clusters to run simultaneously and be viewed simultaneously on the tiled display, in essence, a true multi-tasking operating system for tiled displays. Anything from a parallel OpenGL application to a HD/4K video stream to a remote laptop can be displayed on the tiled display as long as the pixels from their image buffers can be extracted.

SAGE also features a capability called *Visualcasting* whereby dedicated clusters can be placed at high-speed network access points to replicate incoming pixel streams and broadcast them to multiple tiled displays at the same time, enabling users on distributed OptIPortals to look at the same visuals and therefore work collaboratively. The number of Visualcasting cluster nodes can be adjusted to suit the anticipated number of streams. This capability has been successfully demonstrated over transoceanic links. Addition of trackers or cameras for gesture input allows for richer control and interaction.

2. Parallel Distributed Rendering

Many software packages distribute visual content exclusively to multiple rendering engines in a parallel master-slave or client-server approach. A common shortcoming by most packages is scalability across multiple display tiles connected to a single machine, when the combined tile resolution exceeds the supported OpenGL display context size.

3. Hybrid Systems

CGLX explores an approach where high performance real-time parallel rendering and streaming of visual content from other applications can be combined. The middleware is based on the assumption that the rendering nodes in a cluster have sufficient CPU and GPU resources at their disposal. The framework can leverage from these resources by utilizing classical work distribution strategies in cluster systems such as culling and multi-threading for OpenGL applications and provides a freely programmable API in combination with a native container-based distributed desktop management application which accepts multiple pixel streams. To maximize the availability of network resources for data transmission related to the visualization content, CGLX implements its own lightweight network layer and message passing environment. CGLX provides users with access to parallel hardware accelerated rendering on different operating systems and aims to maximize pixel output to support high resolution tiled display systems. Natively, CGLX maps an OpenGL context to each display tile, resulting in multiple contexts when multiple displays are connected per node. This attribute makes CGLX the only fully scalable OptIPortal interface currently available.

Crucial for all distributed rendering approaches is the availability of a reliable high performance network to retrieve massive data content or to control the visualization system itself. An OptIPortal features a network solution that can provide data transfer rates up the 10Gbits/s. These maximum values can be maintained due to dedicated high performance local networks or a high-speed network grid such as OptIPuter. The access to vast amounts of distributed storage and computational resources on an OptIPortal and the additional network bandwidth enables stream-based approaches to dramatically increase their achievable performance. High performance real-time parallel visualization systems, which can also act as rendering back ends for stream-based approaches, can leverage these network resources to load and process data at remote sites and to simply stream the final results at interactive rates. This attribute of OptIPortals allows users to share, exchange and manipulate remote data sets interactively in distributed cooperative workspaces spanning the globe.

4. OptIPortal Virtual Reality/3D Systems: NexCAVE and REVE

The NexCAVE (.calit2.net/newsroom/article.php?id=1584) is a multi-panel, 3D virtual reality display that uses JVC HDTV 3D LCD screens in an array. When used with polarized stereoscopic glasses, the NexCAVE's modular, micropolarized panels and related software make it possible for a broad range of scientists — from geologists and oceanographers to archaeologists and astronomers — to visualize massive datasets in three dimensions, at a level of detail impossible to obtain on a single-screen desktop display. The NexCAVE's technology delivers a faithful and deep 3-D experience with excellent color saturation, contrast, and stereo separation. To present stereo imaging with a modified consumer HDTV, the JVC panels have a transparent surface applied that circularly polarizes alternate horizontal lines of the screen clockwise and anticlockwise. Lightweight passive polarized glasses filter out, for each eye, the corresponding clockwise or anticlockwise images. Since these HDTVs are very

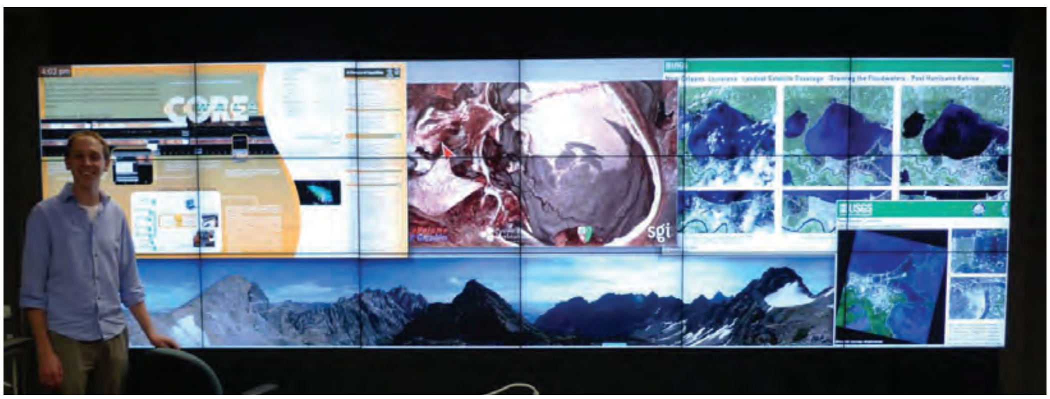
bright, 3-D data in motion can be viewed even with standard fluorescent room lights on. Thus, like other OptIPortals (and unlike projection-based VR systems), the NexCAVE will fit in any office or lab, and can be viewed in normal ambient light. The 10-panel, 3-column prototype at Calit2 has a ~6000x1500 pixel resolution, while a 21-panel, 7-column version built for KAUST has ~15,000x1500-pixel resolution in a semi-circular surround configuration.

The REVE ("Rapidly Expandable Virtual Environment") uses passive lenticular lens HDTV panel 3D technology from Alioscopy, Inc. to present very bright images to the viewer without requiring stereo glasses. The ~3:1 loss of resolution caused by autostereo spatial multiplexing is made up for by tiling the displays. Calit2 has a 6-panel REVE OptIPortal, and KAUST has an 18-panel one.

We currently support three software environments to drive the NexCAVE and the REVE: OpenCover, which is the OpenSceneGraph-based VR renderer of COVISE, CGLX, and EVL's Electro. We use ROCKS-based OS distribution and management to quickly install and recover nodes.

5. Almost Entirely Seamless OptIPortal (the AESOP)

AESOP is a nearly borderless tiled display wall built with 46” NEC ultra-narrow bezel 720p LCD monitors. These NEC displays have inter-tile borders that are 7mm thick when tiled edge-to-edge within the framing, virtually eliminating the “window pane” effect of the “classic” OptIPortal’s 35mm tiled borders. Calit2 has one configurable 16-tile (4x4) AESOP. EVL has an 18-tile (3x6) AESOP, shown above in its Cyber-Commons room. EVL built the first AESOP in the summer of 2009 using hardware funds from an existing NSF grant, shown above in EVL’s Cyber-Commons room. (*Cyber-Commons* is EVL’s term for a technology-enhanced meeting room that supports local and distance collaboration and



promotes group-oriented problem solving.) Calit2 built its AESOP shortly thereafter, also with NSF funds. These displays are scalable, support audio, and have networking and software sufficient to connect these displays to each other, to local servers, and to servers and similar displays worldwide. These OptIPortals run CGLX (UCSD) and SAGE (EVL) software, which support current and future applications.

April 1, 2010
Position Paper for NSF Campus Bridging Workshop

A High-Performance Campus-Scale Cyberinfrastructure For Effectively Bridging End-User Laboratories to Data-Intensive Sources

Philip Papadopoulos[1]
Larry Smarr[2]
University of California, San Diego

A Campus-Scale Dedicated 10Gbps Campus Data Utility
Enabled by nearly a decade of NSF investment, UCSD has been able to investigate, at campus scale, the use of dedicated optical fibers, or wavelengths on the fibers, to simplify the process of bridging data-intensive campus-resources from data generation or storage devices, internal or external to the campus, into end-users labs. To meet expected performance, much of this infrastructure must be on-campus, connecting the lab to the campus gateway or to campus data systems. This has been done in a fashion that is easily duplicated on other campuses.

Two major NSF awards made this cyberinfrastructure (CI) research possible. In 2002, our colleagues and we were recipients of an NSF Information Technology Research grant, the "OptIPuter" (NSF OCI-0225642)[3], that asked the fundamental question of "how would distributed systems be redesigned if bandwidth leaving campus laboratories was essentially unlimited?" In addition to the optical networking research, the OptIPuter project led to the design and software development needed to create tiled display wall OptIPortals[4]--scalable "termination devices" for ultra-speed data flows (typically 10Gbps dedicated per user) entering a user's laboratory.

In 2004, we began developing a prototype terabit-class, campus-scale network instrument through the Quartzite[5] Major Research Instrumentation (MRI) grant (NSF CNS-0421555). The Quartzite system simultaneously supports network-intensive applications, while using the instrument to examine different hybrid network configurations. The three level central Quartzite switch supports both packet switching as well as dense wave-division multiplexing (DWDM) switching-both wavelength conserving switching or switching flows from one wavelength to another. Our work catalyzed the design of UCSD's nascent research cyberinfrastructure overlay to the traditional shared campus Internet by utilizing inexpensive DWDM multiplexing and packet switching to provide many dedicated 10Gbps Ethernet to

---
1 Director, UC Systems, San Diego Supercomputer Center(SDSC)
2 Director, California Institute for Telecommunications and Information Technology (Calit2)
3 www.optiputer.net
4 See white paper submitted to this workshop by DeFanti, et al.
5 The OptIPuter, Quartzite, And Starlight Projects: A Campus To Global-Scale Testbed For Optical Technologies Enabling LambdaGrid Computing (invited Paper) Larry Smarr, Joe Ford, Phil Papadopoulos, Shaya Fainman, Thomas DeFanti, Maxine Brown, Jason Leigh, Optical Fiber Communication Conference & Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC) 2005, Anaheim, California, March 6-11, 2005, CD ROM. [www.optiputer.net/publications/articles/Smarr-OFC-2005.pdf]

various labs throughout campus. Quartzite today has more than sixty 10Gbps paths that interconnect computing, storage, OptIPortals, and instruments at various sites around the UCSD campus. This means a total provisioned of 600 Gbps (1.2 Tbps bidirectional).

Quartzite allows for a centralized campus compute and storage complex that is connected to end-user labs via the 10Gbps dedicated optical wavelengths. This complex at UCSD is called the Triton Resource, a new facility being built at SDSC for UCSD and UC researchers. It has ~30 Teraflops of computing capacity with a total of 15TB of memory. Triton has a small Lustre parallel file system at 110TB usable capacity, which will grow to support large temporary storage (~1 Petabyte) in a new file system called DataOasis. The interconnect is 10Gbps Myrinet (416 MX ports, 32 10GbE) with 8x10Gbps channels connected to Quartzite, TeraGrid, and the Campus Research Network. The UCSD Triton Resource[6] includes the Petascale Data Analysis Facility (PDAF) with 9TB of RAM distributed among just 28 32-way SMP systems, each with large local memory(0.25 – 0.5TB memory per) and 4GB/sec of network I/O (120GB/sec aggregate). These "fat memory" nodes will be expanded even further next year when the NSF-funded Gordon[7] data analysis supercomputer is brought online at SDSC. The Quartzite network allows us to incorporate nodes of Triton as peripherals.

For example, the Community Cyberinfrastructure for Advanced Marine Metagenomics (CAMERA) provides a targeted "computational science destination" which has a 512-processor cluster and about 200TB of project storage on the 1st floor of Calit2. This resource is used by over 3500 remote users from over 75 countries. When the local CAMERA-owned infrastructure is insufficient to meet user demands, the Quartzite campus-scale infrastructure is used to directly mount at 10Gbps (or multiples thereof) the CAMERA data resources onto the SDSC Triton Resource. CAMERA computations can overflow into Triton without any explicit movement of data, providing a relatively seamless integration of resources.

The Dawn of inexpensive 10G Ethernet

Quartzite was a multi-million dollar investment to help us investigate how the next-generation of campus networks should be designed and implemented. In 2005, incremental port cost on a Force 10 E1200 switch-router (used as the core of the network), was approximately $5000-$7000. Today, mid-scale (48-port) 10GbE switches with modest routing capability are about $500/port. Larger switches (100s of ports) with more complete routing capabilities are entering the Market in the first half of 2010 with expected pricing of about $1000/port. In other words, laboratories on campuses can be connected with significant bandwidth for the price of one or two high-end servers. This changes the economics and fundamentally allows remote campus resources to be brought "virtually" into laboratories, via the switched optical fiber infrastructure. Storage and computing (the fundamental elements of cloud computing) therefore no longer need to be located in the end-user lab, but can be elsewhere, allowing for economies of scale in these common resources.

Whither the Grid and Enter the Cloud?

The Grid as envisioned in the mid 1990s sold itself as a way to knit together distributed resources to form low-cost supercomputers. Our community has learned a great deal from the extended grid experiment, and perhaps the greatest lesson was that, in general, most lab scientists found the Grid too difficult to use and simply refused to expend the effort needed to

---

6 Triton is directed by Papadopoulos

7 www.sdsc.edu/News%20Items/PR110409_gordon.html

get “over the hump.” We argue that another issue with the Grid was that infrastructure and data did not appear as local or locally-controlled resources by the user.

It's not a surprise that cloud computing has captured the imagination of scientists, because they could control the definition of their resources while not owning actual hardware. There are probably a large number of small-data needs (email, social networking, photos, etc.) for which commercial clouds will be very useful, since the shared Internet and commercial cloud systems are well engineered for megabytes of data.

However, the cloud (especially in its current commercial forms) does not necessarily meet the needs of data-intensive scientists, for which terabyte-sized data manipulation is the norm. In CAMERA, for example, an annotation data set directly mounted on Triton is 1.6TB and is re-processed in multiple passes. Modern gene sequencers, which are appearing in increasing numbers on campuses, can easily produce a terabyte per run in less than a day. The problem is that a terabyte would take more than ten days to move from the lab to a remote cloud at the usual 10 Mbps achieved on heavily shared wide area networks. Using Amazon published transfer prices, it would also cost about $300 to copy a Terabyte data set in and out of a commercial cloud just once. Neither of these numbers makes the cloud “easy to use” or time practical for large data projects.

The Need for Data-Intensive Campus “Clouds”
On the other hand, with a well wired campus with many 10Gbps optical paths, such as UCSD, the same terabyte takes only ten minutes to transfer from lab to campus cloud. Furthermore, the campus cloud can be engineered for high I/O storage and tightly coupled high performance computing clusters, as in Triton, neither of which make financial sense for commercial clouds. As we mentioned CAMERA (and several other projects at UCSD using Triton) can have their locally-owned storage mounted directly onto such a community resource. What about reverse? That is, having centralized storage directly mounted onto lab resources.

In theory, this is no harder, but security and service scalability become important for practical implementation. This is where elements of the Grid, namely identity management and identity proxy are technically quite important. A major step forward is the Indiana University “Data Capacitor” which provides temporary file space at a campus scale. But from our view, scientists need permanent online storage that can used directly in their labs with sufficient performance. The pricing of 10Gbps makes it financially practical to provide the wiring at campus scale to make data transfer times minimal. However, there are significant technical issues to solve to implement centralized storage that meets performance, security, and data integrity requirements.

Software Integration of Community and Local Resources
Most users build local infrastructure (eg. domain scientists build their own cluster and storage), because they often need to control the software structure. They need their analysis codes (many of which are home grown) to work on their data. It's clear that having every lab “roll their own” is both highly inefficient and creates islands of data. Yet, if the CI community that understands and has a track record of building scalable infrastructure wants to impact users, special attention has to be paid to “ease-of-use,” meaning making the remote infrastructure behave and perform as close to local infrastructure as possible. This is still an open challenge.

Regional Cyberinfrastructure as a Bridge Between Campus and National CI

March 30, 2010

Greg Monaco, Great Plains Network
Rick McMullen, University of Kansas

A critical emerging problem is the ***cyberinfrastructure divide*** between the computing environments for researchers a) on their home campuses, b) at HPC centers of other universities where they may collaborate and c) at national CI centers such as the TeraGrid, the OSG and national data repositories. It is often difficult or impossible for researchers to leap this divide because they lack appropriate local support that understands both departmental/campus computing environments and those provided by national facilities. This CI divide is caused by structural mismatches between national and campus CI and related services and, in many cases, altogether missing campus CI components. Programs such as CI Days and EMERGE at RENCI are very useful for building awareness and meeting specific researchers' needs through provisioning of off-campus computing resources but do not directly address the long-term commitment needed to nurture and build campus CI so that it both meets local needs and articulates seamlessly with national CI.

Just as it makes sense to aggregate computational tools at logical levels such as the department, the campus and the nation to achieve efficiencies and economies of scale, it also makes sense to think in terms of aggregation of expertise and even resources at an intermediate, regional level. A regional focus offers advantages in terms of a) being a manageable size, b) being likely to include a broader range of expertise not available at any one campus, c) having already been successfully exploited for state and regional networking, a major CI component.

We propose that a regional focus in the form of Regional CI Centers will be effective for the coordinating the development of CI at the campus and national levels and for bridging the *CI divide*.

I. What Is Regional CI?

Regional cyberinfrastructure (CI) consists of networks, computing and data grids, shared computational and data services, infrastructure to support virtual organizations spanning regional institutions, and expertise in diverse areas such as high performance computing, grid computing, data management, and collaboration technologies *held in common across multiple institutions in a region*.

A focus on regional CI addresses several important problems including:

- How to successfully ***coordinate*** efforts across multiple institutions with shared interests,
- How to foster a ***collaborative context*** in which researchers can address problems of regional interest and importance, and
- How to bridge the ***economic component*** of the *cyberinfrastructure divide* between users at resource poor institutions and regional and national resources.

John Connolly, Director for Kentucky NSF EPSCoR, has been a strong proponent of a regional focus for cyberinfrastructure in EPSCoR states, and SURA has been effective at organizing SURAgrid across a wide area. These efforts, as well as efforts in the Great Plains region, suggest the following goals for regional CI:

- Supporting the development of CI at the campus level in a regionally coherent way at primary research institutions with the possibility of outreach to four year and two year institutions and to industry,
- Aggregating regional CI resources and making them available through a common set of policies and procedures,
- Providing a seamless path for researchers to scale their research from campus work environments to regional HPC and national capability computing centers,
- Building knowledge management structures and training programs that will allow the efficient sharing of CI and computational science expertise across the region, and
- Supporting the integration of CI into educational programs at institutions of higher education primarily through the cooperative development and delivery of computational science curricula.

II. How Can Regional CI Be Organized: The Regional CI Center

For regional CI to work, it must be organized. Historically, regional organizations in the form of the GigaPoP and Regional Optical Network (RON) were essential to physically bridge the divide between national and campus-level networks by connecting institutions in the region. These organizations have created a model for successful regional CI efforts: They span multiple campuses, have experience at developing close service provider relationships on behalf of their constituents, and have a culture that fosters cutting edge expertise. This combination puts them in a unique position to bridge the CI Divide by providing a context in which regional CI can emerge as a synthesis of local needs, capabilities, development goals and funding. In addition to providing a technically sophisticated engineering capability and a 24x7 operational support structure (e.g. operations centers) RONs typically have the social infrastructure such as social networks and administrative and technical support needed to develop and implement technical goals across multiple institutions.

The expertise of a regional CI center can support the efforts of each university HPC center in several important ways. First, CI center staff can assist in bridging campus users at one institution to resources at nearby institutions. As a corollary CI center staff can also help to establish processes and standards to make local resources available beyond one institution. Regional CI organizations can help advance state and regional economic development goals using the combined resources of the member institutions. They will also provide a context for outreach from campus HPC centers to nearby two and four year colleges that may also participate in the region's R&E network. Finally, they increase university HPC centers' reach and potential impact on state economic development priorities and responsiveness to national needs.

Regional CI centers can help to increase the effectiveness and reach of a campus HPC center by providing end-user training, research training and inter-institutional graduate research fellowship programs. They can also be springboards for the formation of new collaborations, directly supporting regional multi-institutional projects and serving as a focal point for development and specialization of multiple regional HPC centers with different but interlocking areas of expertise.

Regional CI centers may also serve as focal points for provisioning CI to be held in common by multiple institutions but co-located at a particular HPC center. Examples of this include infrastructure for federated identity management across several institutions in the region and providing a point of aggregation of services provided by several HPC centers so that all institutions in a region (or from outside the region) have a consistent view and method to access these services.

A working prototype for the Great Plains Network (GPN) consortium was developed by Gordon Springer of the University of Missouri at Columbia, and provides an example of how a regional CI organization can support campus-based research CI. Taking advantage of Federated Identity Management across a subset of Great Plains Network (GPN) member institutions, Springer demonstrates that campus-level services, including computation and storage assets, can effectively be made broadly available outside the home institution. This has important implications for sharing data, computation and storage resources across an entire region but still requires support to help researchers access shared resources and move applications to HPC or HTC centers, training and support to campuses for technologies relevant to resource sharing (e.g., Condor and Shibboleth), and perhaps, most important, ongoing development and refinement of middleware and end-user services for integration of local computing, storage and computational services into a regional pool.

III. How Will a Regional CI Center Bridge Between Campus and National HPC Centers?

A regional perspective on CI that is aligned with capabilities and practices at national facilities can accelerate development and scale-up of projects from campus resources. Following the RON model, CI centers would provide expertise, training, support and monitoring services through a service organization with deep technical expertise. This, combined with an understanding of research at their constituent campuses, provides an excellent platform for offering training and related services that connect these users to national supercomputer facilities or shared regional capabilities.

From the national facility point of view, a key issue in sustaining HPC centers is to nurture and grow a varied user base working on important problems. Regional CI centers can contribute to this by fostering collaborations between individual researchers working on large scale problems. When they outgrow local and regional CI resources these successful and productive collaborations can be connected to national resources. RONs and Internet2 provide starting points for regional CI activities like *CI Days*. Momentum gained from individual instances of such programs needs to be sustained through the development of regional *communities of practice*. Persistent regional CI organizations are a natural place from which to nurture this expertise.

Another issue in bridging campus and national HPC centers is differentiation and specialization of campus centers. Regional CI groups can support differentiation strategies and the development of nationally relevant specialized CI resources at campus HPC centers through outreach to researchers who may be interested in a center’s expertise or facilities. Regional CI centers can also serve as aggregators and filters to help researchers in their region find campus HPC centers with competencies and expertise that meet specific requirements needed to achieve their research goals.

IV. Conclusions

We believe that regional Cyberinfrastructure centers have a promising role to play in bridging the CI divide between campuses and from campus to national CI resources. Regional CI centers can support efforts on campuses within the region, support efforts among campuses, and coordinate efforts with other regional and national centers. This layered approach, from campus to regional to national, with multiple regional centers of expertise, will serve to strengthen the fabric of national CI.

The Role of a Data-Intensive Network (DIN)

Willis Marti and Guy Almes
March 2010

Over the years, universities have developed campus production access networks of increasing performance, reliability, and security. These often have modern switched Ethernet with 10-Gb/s trunks forming its backbone layer, but with a complicated and heterogeneous array of lower-performance infrastructure closer to the edge, but often constrained by firewalls, traffic shapers, and similar "packet disruption devices". They are successful as access networks, enabling good performance for a wide variety of applications for thousands of faculty, staff, and students. These access networks, however, often cannot support high-speed wide-area end-to-end performance, both because of limited capacity, but (even more significantly) because of the jitter and packet loss induced by the "packet disruption devices".

To meet the needs of the research community, it will be important to support data-intensive applications with high-speed networks optimized for that role. Current production access networks are not ideal in this role and cannot evolve to support this role in a natural, cost-effective, manner. While the current network does make some use of 10-Gb/s links and switches, it has several weaknesses, most notably:

- due to the diversity of switches and interfaces, it will not be able to support the large packets (i.e., packets larger than 1500 bytes, sometimes called jumbograms) needed for efficient bulk data transfer; and
- due to the extreme prevalence of PCs and other systems administered by (amateur) users, firewall-based approaches to security are needed, thus further limiting single-application performance.

Again, these weaknesses are natural consequences of the need to optimize the production access network for security and good support for a very wide diversity of equipment, applications, and users.

Instead, we propose a purpose-built network, which we refer to as the Data-Intensive Network (DIN), that will *complement* the production access network. The DIN will be characterized by several properties:

- The DIN will support jumbograms of up to 9000 bytes throughout.
- The DIN will support very-high-performance 10-Gb/s Ethernet switches.
- All hosts connected to the DIN will do so with 10-Gb/s 9000-byte MTU interfaces.
- All hosts connected to the DIN will be professionally sys-admined and will support best-practice practices consistent with high performance and strong security requirements.
- Some hosts connected to the DIN will also be connected to the production access network (typically with 1-Gb/s connections).
- The DIN will connect to its Regional Optical Network and national backbones at 10 Gb/s. This connection will *not* have a firewall, but will have an intrusion-

detection system. The goal of this connection is to support interconnectivity to cyberinfrastructure resources at other universities and laboratories that are similar in nature to the hosts connected to the loca DIN.

The hosts to be connected to the DIN are high-end cyberinfrastructure resources, among which there high-speed data transfers are needed. Examples include:

- High-performance and/or high-throughput systems;
- High-end visualization resources;
- High-end instruments; and
- Large parallel storage systems that support massive and/or high-speed data storage for the above systems.

The underlying argument for the DIN is that these advanced cyberinfrastructure resources, in order to achieve their potential, will need to exchange data at high speed with other campus and/or national cyberinfrastructure resources across the extended Internet2 fabric.

This approach has several attractive characteristics:

- The technical approach is simple and within the engineering abilities of many research universities;
- If done in a coordinated fashion across the nation, the benefits would grow as the square of the number of universities and cyberinfrastructure resources so attached -- that is, the benefit would grow as Metcalfe's Law;
- The value of the campus cyberinfrastructure resources so connected would be enhanced;
- Similarly, the value to university researchers of non-local cyberinfrastructure resources, particularly nationally-funded ones, would be enhanced; and
- The resulting DINs would be a natural focus for perfSONAR and other network performance measurement tools, and similarly for dynamic circuit and other advanced network architectures.

Thus, the DIN approach provides an accessible means for many research universities to enhance the value of local and remote cyberinfrastructure by cost-effectively improving high-speed wide-area flows among cyberinfrastructure resources across the country.

A Strategy for Campus Bridging for Data Logistics

Terry Moore - University of Tennessee, Knoxville

Introduction

As data intensive methods of research have escalated and spread to more and more fields over the past decade, sites at all tiers in the research community, but especially Tier 3 CI sites, have faced mounting problems of data logistics, i.e. problems in the management of the time related positioning of data. The trends producing these logistical challenges show no sign of abating. By all accounts, the size of the data flows pulsing out of new instruments, sensor networks and massive simulations are going to continue to grow, and consequently many researchers in large national and international collaborations, sitting in different administrative domains, spread across the wide area network, and using diverse local resources, are likely to find that the data sets of most current interest are not where they need to be, when they need to be there, for their work to proceed efficiently. It seems clear that an adequate cyberinfrastructure strategy for campus level bridging must address this critical problem area.

The Research and Education Data Depot Network (REDDnet) is an NSF funded initiative to build and operate a network of WAN-aware storage nodes to address just such challenges in data logistics. A basic premise of REDDnet is that, as in other types of logistics, fast transport is only one component of a successful data logistics strategy. Consider some of the most prominent factors that shape the challenges users confront in this area:

- *Data volume*: In many branches of science and engineering, the quantity of data that needs to be made accessible is already enormous, coming in regular or occasional pulses that range from multiple terabytes to petabytes.
- *Distribution*: The people who need to engage with a given data set, and the resources they need to use, are distributed across the wide area (i.e. the nation or the globe), and most are in locations where super fast networks do not go.
- *Asynchrony*: Different members of these distributed teams and communities will want to work on a given set data at different times; and sometimes, if not frequently, they need to coordinate their workflow with one another.
- *Data transformation*: The data set of interest sometimes needs to be preprocessed (e.g. reformatted, segmented, filtered, reduced) in some way, and sometimes in various ways, in order to be suitable for the different uses that members of a collaboration will want to make of it.

The REDDnet community believes that addressing problems with these elements requires the integration of high performance networking with substantial amounts of “working storage” for caching, staging, prefetching, replication and explicit buffering. To attack this problem, REDDnet builds on a unique form of storage technology that is designed for both deployment scalability and fast data transfer in the wide area. We believe that this technology, which we are currently working to package with complementary network tools and services (e.g. perfSONAR, Phoebus) in a software distribution called the Data Logistics Toolkit (DLT), can provide a powerful platform for campuses looking to create bridges for data intensive collaboration with national or regional infrastructure.

Today REDDnet provides a wide area storage facility, consisting of a substantial set of large storage depots, distributed across the nation's high performance network and available for use by a diverse set of researchers and educators with large datasets to manage and an established need for distributed collaboration. Its growing data-depot network consists of three elements: (1) the storage hardware (depots) installed at host institutions, (2) the high performance research network that connects these institutions, and (3) the software components used to manage, monitor and operate the depots and network together as a unified, sharable resource. The DLT will bundle together the key parts of (3) in a easy to install package that will enable campuses to deploy it on their own hardware, empowering users to both leverage REDDnet resources and create innovative solutions at the local level.

Use Case: CMS and "Untethered Computing"

The pattern of data distribution and access in much data-intensive research today exhibits a kind of “pulse”: there is a periodic, large-scale injection of data from some source (e.g. a detector, telescope, simulation run) that is of great interest to a widely distributed community; this data is vigorously analyzed in different ways, by different research groups, using different aggregations of local resources; then, as later injections occur, interest in it rapidly drops off . This combination of the extremely large files (which may overwhelm local storage resources), the size and distribution of the community of interest, and the scheduling of the resources that need to be applied to it, tends to create serious logistical problems in providing for the timely availability of such data.

The high energy physics experiment CMS (Compact Muon Solenoid) offers a notable and current example of this model. The CMS research community must confront a severe practical reality: CMS will generate petabytes of data every year, which in turn means that several petaflops of computing power will be required to extract the science. The large community physicists and their computational and storage resources are dispersed around the world, and the majority of researchers and students do not have local resources sufficient for the task. This is especially problematic because the current data analysis model for CMS uses a hierarchal network of distributed data and computing centers. The top level centers (Tier 0 through Tier 2) are heavily managed resources with high service guarantees to produce CMS mission-critical results. Unfortunately, this computational model is "data-tethered" -- computation is scheduled only where the data already resides, and this presents a serious problem of data logistics for Tier 3 sites.

CMS Tier 3 sites often lack either sufficient storage or computing power or both, and therefore have a different dynamic, with no minimum resource levels or service guaranties. Their storage and computational resources vary widely and may even overlap with other application communities and Grids. To “untether” Tier 3 analysis, REDDnet’s CMS collaborators will upload datasets into the geographically distributed data depots on REDDnet, which has the capacity, via its *Logistical Distribution Network (LoDN)* server, to automatically replicate and stripe the data across multiple depots according to user defined policies. Using LoDN’s unique metadata capabilities, CMS analysis jobs are then able to stream data directly from REDDnet, using multiple depots and multiple network paths, with no pre-staging needed. Dynamic selection of depots optimizes data streaming – data is always taken from the closest copy. In this way, REDDnet can act as a “global storage element” for many Tier 3 sites, acting as if all the desired data was stored locally.

The Data Logistics Toolkit (DLT) is an integrated collection of software components that have been deployed on REDDnet, are being tested for the data logistics requirements of the CMS community, and which we are packaging together for general distribution to address the logistical issues of data intensive collaboration. Its ultimate purpose is to enable services (e.g. a global “drop box” that supports automatic data postitioning) that support the automation and optimization of rapid movement and timely placement of data, across a wide range of scenarios, through policy-controlled sharing, replication, caching, control loop optimization and overlay multicast. Below we provide an overview of the early versions of the package and our plan for its deployment in conjunction with REDDnet and CMS.

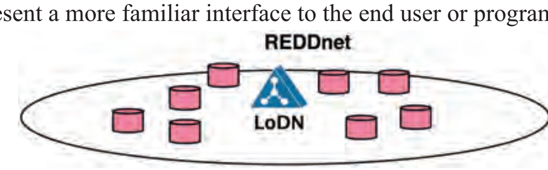
A Data Logistics Toolkit for Campus Bridging

The DLT combines software technologies for shared storage, network monitoring, enhanced control signaling and more efficient use of dynamically allocated circuits. Its main components are network storage server (“depot) technology based on the Internet Backplane Protocol (IBP), perfSONAR for network performance measurement and monitoring, and Phoebus, for optimizing the use of network resources in long haul data transfers. These components have been developed independently (with NSF funding) and their value has been demonstrated through a variety of research and infrastructure projects (e.g. REDDnet); but to achieve dramatic improvements in other production environments, like those required for campus bridging, they work together more seamlessly and come in a package (e.g. a pre-configured «LiveCD») that is simple to install and manage. That is the purpose of the DLT effort.

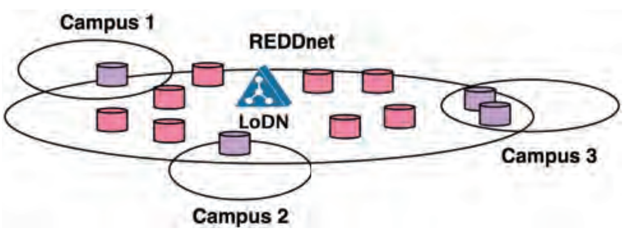
Another essential component for the DLT-based campus bridging strategy is the Logistical Distribution Network (LoDN) Directory Service. LoDN is responsible for storing and managing the unique portable file metadata (“exnodes”) that REDDnet uses. In particular, it associates exnodes with administrative metadata, such as a name in a hierarchical namespace, a concept of ownership by a specific user and/or group, and a set of goals or policies as to where replicas should be placed. LoDN’s dynamic management of exNodes includes making periodic requests for the renewal of allocation leases and making new allocations and moving data between allocations in order to take account of allocations that fail, due to depot or network problems or to non-renewal of a lease.

Many of the performance gains that come from the use of DLT-managed depot pool are the result of direct access to IBP's low level storage service from application programs. Application tools which are layered over IBP (e.g. I/0 library with core POSIX functions; a GridFTP gateway, supporting SRM; DLT-enabled version of HDF) present a more familiar interface to the end user or programmer.

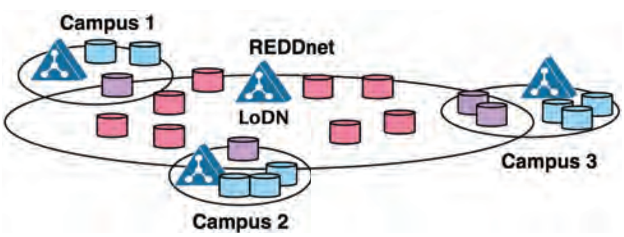
We anticipate that the deployment of DLT software for campus bridging would occur in three stages, as represented in Figures 1. Stage 1 is the current deployment, with DLT software enabling the REDDnet core. REDDnet currently has a few hundred terabytes of storage deployed on IBP depots at 11 institutions across three continents. A LoDN server at Vanderbilt University actively manages and provides policy driven replication for all the data currently available on the facility. It is important to note that these illustrations do not show either the networks involved —depots in the REDDnet core are deployed at sites on Internet2, NLR, and other advanced networks— or the critical networking components of the DLT. To achieve optimal network performance, REDDnet deploys portions of the pS Performance Node suite of performance and monitoring tools at all depot locations. We are also integrating Pheobus , which is already deployed in the Internet 2 backbone, into the DLT to optimize network paths, improve performance of transfers over congested or lossy connections, and automatically take advantage of advanced services, such as Internet2’s ION service and the ESnet SDN.



Stage 1: Representation of the current REDDnet



Stage 2: DLT distributed to Tier 3 campuses and installed on depots there (purple); all distributed data managed through REDDnet LoDN.



Stage 3: DLT distributed with LoDN; campuses use their own LoDN to manage data on their private depots (blue) and coordinate with REDDnet.

Figure 1: Staged deployment of DLT for campus bridging

In Stage 2, participating campuses (beginning collaborators in the CMS community) will use the DLT software distribution, without LoDN server software, to deploy depots on storage clusters on their own individual campuses. The central REDDnet LoDN will manage all automatic data positioning, but local campus depots could be used to improve performance and flexibility of all users, which could include any application community that requires working storage. Subsequent distributions of the DLT will include LoDN server software, which campuses can use to manage there own data logistics infrastructure and coordinate with national facilities, like REDDnet.

Extending Cyberinfrastructure Beyond its own Boundaries

Campus Champion Program

S. Kay Hunt[1], Scott Lathrop[2]

[1]Purdue University, [2]University of Chicago

March 30, 2010

TeraGrid is a national-scale high performance computing facility funded by the Office of Cyberinfrastructure at the National Science Foundation to provide resources and services in support of advancing scientific research and education. TeraGrid (TG) includes 11 Resource Providers and is a leader and major component of the emerging national and international cyberinfrastructure (CI). The TG (soon to be followed by eXtreme Digital or XD) provides free access to computing resources and services to all researchers and educators. To extend beyond the current user base to support broader national participation, the Campus Champions Program was formed just over a year ago. The Campus Champions program pro-actively engages researchers, educators, administrators, staff and students on campuses across the country to facilitate awareness of and access to TeraGrid's resources and services.

Experience has demonstrated that numerous activities, strategies and pathways are needed to engage and support broader participation in terascale and now petascale science, engineering, and scholarship among researchers and educators. The Campus Champion Program establishes strong campus partnerships with the TeraGrid that provide multiple avenues for support and advocacy among campus users and TeraGrid Resource Providers.

Further, experience has demonstrated that TeraGrid has benefitted from leveraged partnerships with other national, regional and local providers of cyberinfrastructure resources and services. In particular, the Campus Champions program works closely with EDUCAUSE, Internet2, National Lambda Rail, Open Science Grid, MSI-CIEC, SURA, Blue Waters and other organizations to provide campus partners with a wealth of resources and opportunities to best serve the campuses.

As we prepare for the TeraGrid "eXtreme Digital" (XD) future, establishing a national Campus Champion group focused on meeting the requirements of a much broader national community has tremendous potential to increase participation and provide a foundation from which TeraGrid and other Cyberinfrastructure (CI) resources and services may be broadened among traditional communities of users of TG as well as among under-represented communities. Particular emphasis must be placed on including women, minorities, and people with disabilities, as well as including under-represented institutions including Minority Serving Institutions and EPSCoR institutions. These all represent communities that have been under-represented in all areas of science, technology, engineering and mathematics (STEM), and with respect to their use of CI and TG resources, specifically.

The vision is to realize the potential of CI and High Performance Computing (HPC) to empower a larger and more diverse set of individuals and institutions to participate in science and engineering education, research and innovation.

The mission is to develop and evaluate an extensible, scalable, and comprehensive program to assist researchers, educators, staff and students at academic institutions to effectively utilize CI, TG, and XD resources and services, to advance scientific discovery in all fields.

The goal of the Campus Champion Program is to identify key people on campuses who can be pro-active advocates and brokers of information among local campus users to help them make informed choices among the range of national, regional and local CI resources available and advance scientific discovery in all fields.

The objectives of the program are to:

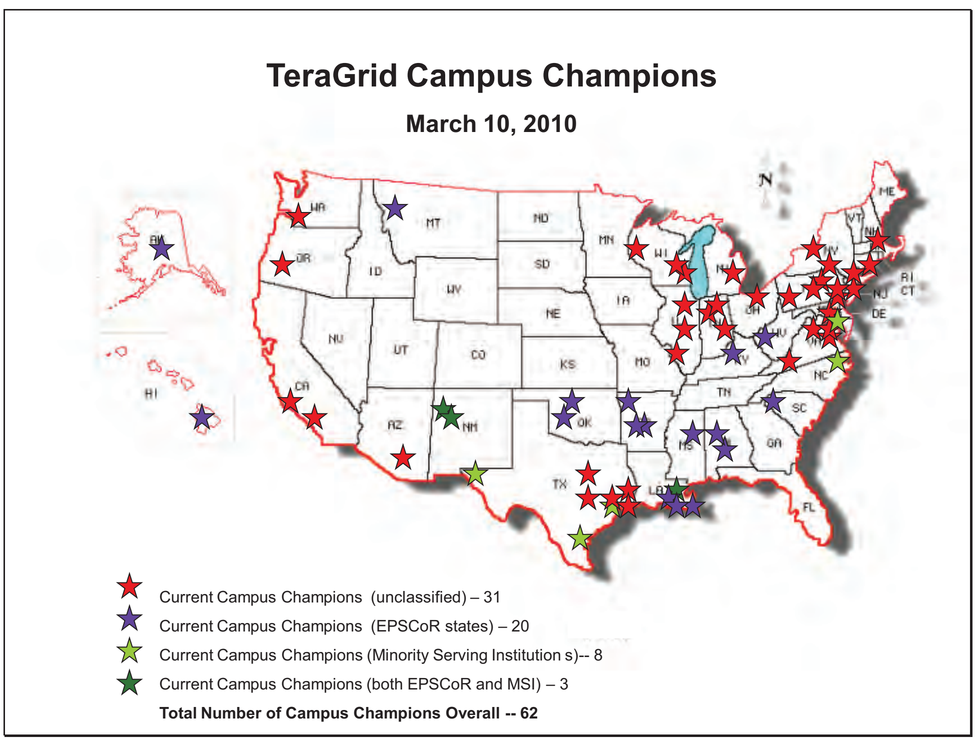
- Provide campus representatives (Campus Champions) with the requisite knowledge to train and support campus users on CI and TG resources and services
- Provide campus representatives with software tools, advice and support that they can in turn provide to local users to assist the users in connecting to and using CI and TG
- Provide campus representatives with the tools to help local uses to quickly get started using CI and to quickly overcome barriers and problems that preclude productive use of CI
- Promote campus advocates to raise awareness among campus users to foster increased an effective use of CI

The program addresses the following challenges:

- Lack of understanding of CI and TG
- Lack of knowledge among researchers, faculty and students about CI and TG resources and services
- Facilitate timely allocations and problem resolution for users of TG and CI
- Local on-site assistance with problems encountered by users
- Solving repeated problems
- Professional development for Champions

While the Campus Champions Program has been quite successful, including the recruitment of over 50 member institutions, there is tremendous potential to further improve the program to better serve diverse communities across the country. TG and all of the aforementioned collaborating organizations, continue to seek advice and recommendations from all Campus Champion institutions and potential member institutions to further improve the Program offerings and services.

In summary, the Campus Champion Program continues to offer its services while seeking advice for further improvements for achieving broader and more effective use of the national cyberinfrastructure to optimize innovation and discovery among all science and engineering communities. The lasting result will be more effective and efficient use of the current and emerging cyberinfrastructure.



Development and deployment of integrated attribute based access control for collaboration

Ken Klingenstein, Internet2

Summary:

With the emergence of federated identity and mechanisms to move attributes from sources of authority to applications and relying parties, it is time to complete the basics of a national authentication and authorization infrastructure. Given the progress on authentication, and the new capabilities it introduces, more attention can now be concentrated on authorization. In many cases, authorization and access control can be addressed by attributes. By moving forward with a coordinated development and deployment of the tools for attribute management, the foundations for a national collaboration identity infrastructure will be in place.

To be successful, there are three requirements on attribute management

- Integration across campus and virtual organization
- Integration across collaboration and domain tools
- Integration across web and command line

Description of work:

Attributes are properties and characteristics of people. (They also apply to devices and computing processes, among others, but those uses of attributes are beyond the scope of this paper.) Attributes may be short-lived and dynamic or static (e.g. date of birth.) Attributes are assigned by sources of authority, perhaps delegated, and can be assigned values by systems (e.g. a Student Information System determines who is a student) or by individuals (a VO administrator will grant permission to control a device or write to a wiki.)

Attributes are the fodder for access control decisions. Many of these kinds of attributes can be expressed simply, by group membership, for example. Other attributes, often called privileges, are finer grain and can contain limits, quotas, expiration dates and other modifiers of a basic permission.

Current approaches to access control in collaborations are problematic. They are non-scalable, non-privacy preserving, difficult to adapt to new applications, hard to use and do not integrate across service platforms or communities.

There are several elements that have to be developed or expanded.

Tools to create and store attributes – Typically on campuses there are legacy and institutional systems that generate attributes for students, faculty and staff. For virtual organizations and inter-institutional collaborations, however, there are almost no tools, and no best practices on how to create the attributes from the common roles and permissions of the collaboration.

Mechanisms for transport of attributes to relying parties and applications. Some elements, such as LDAP and SAML exist. Tools for applications and resources that consume other kinds of credentials, notably X.509, would be useful. For example, a general tool for packing attributes into X.509 certificates that could be customized into the variety of X.509 formats that applications expect.

Techniques for aggregation of attributes by a relying party. Some basics exist, for example in Shibboleth, but additional mechanisms are needed.

Mechanisms for the easy expression of policy by relying parties and the translation of policies into access control languages. Much work is needed in developing user interfaces and constructs that readily permit resource owners to set policy requirements.

Coordination of name space and attribute use by VO's. In order to integrate with campus and federated attributes, VO's will need to have a registry for their name spaces and a set of models on how to create attributes internally and consume other external attributes.

For this work to become broad cyberinfrastructure requires deployments in both campus and VO settings, nationally and abroad. The campus elements are beginning to be deployed, in a variety of systems, including Grouper and Rice, as well as the legacy applications exporting attributes into directories. To roll this out for virtual organizations will require some incentive at the national level by the research agencies. For example, developments in national collaboration services through attribute systems are being done by SURFnet in the Netherlands, and JISC in the UK. Within the US, DoE Office of Science could promote and house such authorization infrastructure. Within NSF, either the CIO's Office or individual Directorates could foster such infrastructure.

To be cyberinfrastructure also requires that relevant national and international laws be followed. This includes privacy laws, mechanisms to enfore export controls and other regulatory issues, informed consent approaches on the release of attributes, and other legal considerations. Audit controls can provide legal, diagnostic and scholarly value.

Costs and benefits:

As with much middleware, especially tightly scoped work, development costs are relatively modest. In fact many of the elements already exist or are being developed.

The costs of deployment, however, are significant, though more so in time and practices than in hardware, software or networking. Because business practices might need adaptation, and because of an embedded base of systems that use identity for authorization and depend on existing, non-scalable solutions, there is a need for change management as well as new deployment. User training for collaboration managers is also needed.

Benefits include ease of use for both the user and the collaboration manager, automatic provisioning and deprovisioning, integration of instruction and research, integration of campus and virtual organization attributes, applicability across web, web services and command line applications, reduced costs of maintenance, reduced legal exposure, compliance with international policies, and automatic management of access where possible.

<u>Non-position paper</u>:

There are critical elements of the authentication infrastructure that are not yet in place but are actively being worked on. It is assumed that these projects will have continue and be successful, else the gaps need to be addressed.

Interfederation

Integration of federated identity into non-web applications

Realizing a National Cyberinfrastructure via Federation

Andrew Grimshaw
University of Virginia

There's an old story about the person who wished his computer were as easy to use as his telephone. That wish has come true, since I no longer know how to use my telephone.
-- Bjarne Stroustrup

The Problem

As an enterprise, scientific research today is increasingly computationally-focused, data-intensive, multi-organizational, and international. This holds true not only in the "traditional" computational sciences such as physics, (bio)chemistry, astronomy, and chemical engineering, but also in emerging computational communities such as biology, clinical and translational research, economics, architecture, and English. This trend toward an ever-wider use of computational techniques for collaboration and research will likely continue for the foreseeable future.

This kind of research, though increasingly popular, poses certain difficulties. Researchers are in need of simple mechanisms to support collaboration, such as sharing data, applications, devices, cycles, and other resources across organizational boundaries. Obstacles that stand in the way of efficient collaboration include baroque, non-interoperable security policies and mechanisms, the ever-increasing quantity and types of data (and metadata), and the lack of operating system support for access to non-local resources. Unfortunately, researchers in communities new to computationally-focused research tend not to be as computationally sophisticated as the early adaptors, a circumstance that results in slower tool adoption.

Looking at just one of these issues, data sharing, we observe that the explosion of information systems in almost every organization has led to the collection of vast arrays of data stored in widely-varying formats, locations, and subject to numerous access and privacy policies. Much of this data is collected and housed in standalone data stores, with little or no integration between systems.

Unfortunately, the process of integrating data from existing sources is complex, especially when the integration involves crossing organizational boundaries. The lack of appropriate models, infrastructure, and tooling to aid the collaborative process has slowed the pace of integration, consequently slowing the pace of innovation and the quality of information available to researchers, clinicians, and decision-makers. The human energy threshold needed to overcome these problems is currently too high; what is needed is a solution on the level of infrastructure—in this case, an effective way to share data.

Achieving data integration requires tackling a wide range of issues, including: finding, naming and providing access to physical data sources; crossing firewalls; enforcing security and privacy rules; synthesizing data with multiple schema, representations, formats, accuracy, and terminology; and providing appropriate levels of performance, coherence, timeliness, reliability, and availability.

We argue that the fundamental problem preventing the integration and collaborative exploitation of data is the underdevelopment of software: what is missing is an effective system software that will not only securely and reliably manage the data but also (via an access layer) mediate interactions between users and back-end system resources.

There are those who will argue that bandwidth is the fundamental problem for research universities to "bridge" the gap across campuses and connect campuses with national resources. While bandwidth is necessary, it is not sufficient. Simply adding a 10 gbs link to a campus will not change what people do, whereas giving them easy-to-use, powerful tools that use that bandwidth will. For proof, we need only look at the fact that most tier 1 schools do not utilize the bandwidth capacity they currently possess.

Others will argue that the browser-based web access to resources is sufficient. Indeed, the web is great for read-mostly, human-driven data access or portal-based access to compute resources. But it is not good for application-driven access—whether to data, applications, or compute resources—because it requires a human in the loop to drive the process, a fact that fundamentally raises the energy required to exploit non-local resources.

Finally, many argue that new, rich, powerful APIs and tools are the answer. While these tools certainly have a place in the toolkit of high-end users, one must keep in mind that most campus researchers are not that sophisticated, and moreover are not interested in becoming so. Even in engineering and the sciences, the sophistication of users (from a command-line and scripting point of view) is decreasing.

Attributes of a Good Solution

What are the attributes of a good solution? We feel that there are at least four. First, the solution should be simple and use familiar paradigms. Second, it should have reasonable performance. Third, it must be secure. Finally, it should be standards-based. The first two attributes are focused on lowering the human energy barrier to use, while the third and final attributes regarding security and standards address institutional requirements and risk. Let's look at these requirements more closely.

Simplicity and familiarity. The first and most important thing to recognize is that most scientists do not want to become computer hackers. They look at the computer as a quotidian tool used for a wide variety of tasks, including reading email, saving attachments, opening documents, and cruising through the directory/folder structure to look for files. Therefore, rather than have scientists learn a whole new paradigm to search for and to access data, we believe the paradigm with which they are already familiar should be extended across organizational boundaries and to a wider variety of file types. Specifically, data of all types in many different organizations should be mapped into shared name spaces (directory structures) and then mapped into the scientists local operating system as a mounted file system. Scientists could then (subject to access control) access data the way they do now. Not only could they access it by clicking or *cat*ing it, their applications could also access the data as if it were local.

Similarly, when it comes to installing software, sharing their own data, or any other task, it should be as simple and straightforward as installing software on a PC. Simple installers, GUIs, and tools should be used, rather than multi-step installation processes that require the user to understand *tar* or *make* commands.

Performance. Performance requirements are hard to quantify, as different users and applications have different needs. Some applications use a few very large (> GB) files, while others use thousands of small (<32K) files. Performance goals should be focused on what the user is used to, i.e., his or her local file system and network file system performance. Researchers will not use the tool if it is too slow, i.e., if it performs noticeably slower than what they are used to.

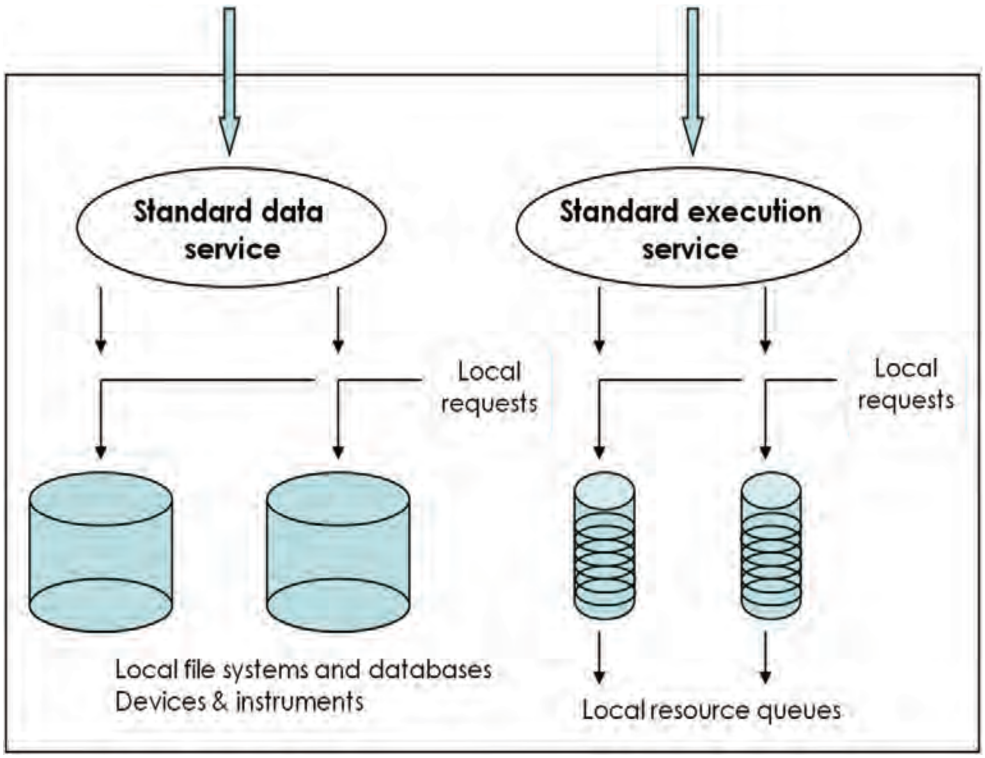
Security. The solution must be highly secure and hook into existing organizational authentication infrastructures and tools. Organizations are, for the most part, locked into their current security infrastructure. Requiring changes is costly, and thus discourages quick adoption. At the same time, the security mechanism should have flexible access control so that scientists can easily share their data with colleagues at other institutions.

Open standards. It is important for solutions to be standards-based so that different implementations by different vendors, deployed at different institutions, can interoperate. If University A adopts solution X and University B adopts solution Y, unless X and Y use standard interoperable protocols, A and B will most likely be unable to easily and securely share data.

The Solution: Federated, Standards-Based Access to Resources Using the File System Paradigm

A federated, standards-based, service-oriented architecture is the best way to connect campuses to one another, to connect campuses to the national cyberinfrastructure, and to interconnect national cyberinfrastructures.

In a federated environment, the *modus operandi* for accessing data, compute, and other resources at individual institutions remains the same. Local access mechanisms, policies, and applications remain unchanged, reducing to near zero the impact on the local operational tempo. Access via the national cyberinfrastructure to local resources represents an alternative access path, not *the* access path. Similarly, individual organizations maintain complete control of their resources; authority is not ceded to a central organization.



In a federated architecture, the local environment remains the same. External requests from the national infrastructure are executed by proxy services on behalf of external users.

A standards-based architecture allows campuses to choose any compliant implementation of the services. There are several advantages to standards-based implementations. First, each institution can choose that implementation that best meets its own quality-of-service profile (cost, reliability, availability). Additionally, the freedom of choice eliminates vendor lock-in, thus mitigating the risks associated with a particular vendor (or project that supplies the software) going out of business.

Finally, use of the file system paradigm means that users do not need to learn how to use a new tool or a new way of accessing resources. They already know how to use file systems , since they use them every day. Similarly, applications built to execute against the *stdio* libraries will work both with local resources and with resources located throughout the national cyberinfrastructure.

Why is Advanced Cyberinfrastructure Not More Widely Used?

Russ Hobby <rdhobby@hobbyfamily.org>

Cyberinfrastructure (CI) has been the said to be the way to enable modern research to work with large discipline groups where participants are from different geographic areas. However, even with all the technology that is available, use of the technology is still relatively modest. There are a few researchers that use the resources extensively and there are many researchers that use a few of the simpler resources. However, wide use of advanced CI is still not the norm today. This paper will explore some possible reasons Advanced CI has not been integrated into the everyday activities of all researchers.

For quite a number of years now there have been researchers that have embraced technology and have integrated it into their research activities. These are the "CI Champions". They have built their own computer clusters, they established high-speed network connections, they have created storage for large data sets, they have developed visualization systems, and moved the data around the globe. However none of that has been easy and it has required that the researchers become technology experts as well as doing their own research. Over time the technology has become more complex and the task of creating and maintaining the technology has gotten harder. Many of these researchers are finding that "it isn't as much fun as it used to be" and would be happy to have someone else take care of it for them.

There is also a population of researchers that in the past have not used CI to any great extent. These researchers do not want to have to become experts in technology in addition to doing their own research. These are the "CI Consumers". However because of the change in the way research is done, working in distributed discipline groups, the use of CI is now a requirement, not an option. These researchers are seeking help with their CI needs so that they can continue to focus on their research subject and not have to become CI experts.

EASE OF USE

Why don't all researcher embrace and use CI to its full extent? Two reasons are often stated. 1) CI is not easy to use and 2) the reliability of CI is less than acceptable. Let's look at ease of use first. During one Campus CI Days event, researchers, who were novices in the use of CI, were asked. "What would your perfect user interface to CI look like?" One researcher answered "It would look like an Excel spreadsheet". This met with agreement from the others in the room. It is not that Excel has the greatest user interface but rather it is one that they know. Learning a new CI tool is time consuming and viewed as time taken away from their research. If they can leverage skills and knowledge they already have, it is viewed as a better option. A common user interface that evolves incrementally is much more preferred to having to learn an entirely new interface every time a new resource is created or updated.

Another example is in the use of research computing resources. A researcher would like to develop their project on small computer systems, in their lab for example. As they move on to the testing phase they would like to use a larger system, like a campus cluster. Once they have proven the methodology they are ready to move to a large number cruncher at a supercomputer facility. However researchers find that user interface at each of these systems is different, and worse, they often have to recode the programs for each larger system. The CI Champions are willing to take this on but not the CI Consumer.

One thing that helps with the ease of use are standard software packages. All supercomputer sites have installed sets of software that are general purpose or specific to a particular discipline. Many of these packages are available at most larger computer sites. While this eases the burden in moving between computing resources for those familiar with the software, each of these packages have been developed independently and have their own unique interface to learn. Is it possible to develop a common, general user interface that can be used by many packages and ease the learning curve of the researcher? (Who remembers the graphic programming interface in AVS?)

Another tool that is used to aid researchers is the gateway (often called a science gateway). For a particular discipline or project a gateway provides coordination and an access point to resources. A gateway can automate and hide much of the complexity needed to make the various resources work together. Authentication to resources is often an issue when there are different id's and passwords for each service. Gateways can sometimes help with this particularly if they tie into federated identity management services.

Widespread use of federated identity management services in applications would help to make all applications more user friendly. It would provide a single user interface for all logins and one set of authentication credentials to remember. It would also make it easier for the applications manager by not requiring identity management for each individual application. It would also help on the reliability side.

RELIABILITY

As stated above, reliability is the other main concern of researchers. While the CI Champions are willing to dig in and figure out why systems are not performing as they should, the CI Consumer will give up and move on to methods that do not use the problematic systems, even if they lose desired capabilities. Effects of poor reliability can range from total failure to denigrated performance. With diminished performance some researcher will use the CI system and think that is "just the way it works" and not realize they could be doing better. Work has been done on end-to-end performance for several years and many of the performance and reliability problems have been identified. However more advanced applications remain fragile and often require help from technical support to get them to work, even if they worked during the last use. Why aren't things better?

The "end-to-end" part is the key. An end-to-end system is composed of many components. When experiencing reliability or performance problems diagnostics of each independent component usually show that each component is working fine. Yet the total experience is less then optimal. A big part of why is because each of these many components have been developed independently. The components are then patched together to make a system where each component is unaware of the needs of the other components. Interfaces between the components are usually quite minimal and only exchange information pertinent to the two components, not the whole end-to-end system.

Because of the lack of a system view, security and performance measurement are tacked on top of the system rather than integrated into it. Each component tries to take care of it own security and performance. However both security and performance are end-to-end issues and dealing with them at the component level without end-to-end coordination creates conflicts where neither is done well. Indeed, security often degrades performance because security measures are placed in the wrong place in the end-to-end path.

Current use of many firewalls is an example of security measures inappropriately placed. Security should reside close to the end that understand the security needs, not placed in a component, the network in the case of firewalls, to make up for security problems in other components (operating systems and applications). In a well designed CI system, the network might participate in security measures as indicated by the needs of the ends, but should not apply blanket security measures that may impede the application more than it helps.

CI ARCHITECTURE

We are in this situation because there has never been an overall coordinated CI architecture, nor could there be because we needed the experience to date to understand the problem. The closest CI architecture has been the seven layer OSI model (Figure 1). With this model each layer is treated as a "black box" and interface between layers has been reduced to the bare minimums. While this makes the design and implementation of each layer easier because it reduces constraints from outside the box, it makes the use of the overall system more difficult because it is not coordinated. Also the OSI model is focused more on the network and less on the other CI components.

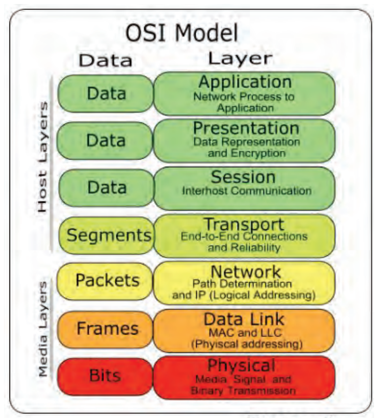


Figure 1.

The Internet itself is a perfect example of this approach. Each operational domain in the Internet only has to hand packets off to the next domain and it is done. It has no concept of what role it plays in any of the end-to-end paths that pass through it. This concept has allowed the Internet to grow to its current size with a minimum of centralized control (and a minimum of regulation that could have stunted its growth). However this structure makes end-to-end services nearly impossible. This was discovered in the early days of Internet2 when Quality-of-Service (QoS) was viewed as the solution to performance and reliability. Mechanisms for QoS (priority queuing and policing) could be implemented within an operational domain but it was found to be very difficult to coordinate QoS between two operational domains, much less across multiple operational domains in an end-to-end path. There are current efforts in the R&E network community to create new end-to-end services. We will see if they are able to overcome the obstacles of the past.

Even if end-to-end network services are created, the network is only one component in the CI system. Focus of performance has often been about the network, bandwidth in particular. With the exception of a few researchers moving large data sets, network performance is generally not the source of performance and reliability problems for researchers at universities. More often it is a configuration problem in one or more of the other components. The correct configuration of each component is often dependent on the surrounding CI environment but there are few tools integrated into the component to discover that environment. It generally requires technical expertise to get the configuration of all the components to work with the best performance and reliability. If one of the components changes then the tuned configuration is no longer valid and the system has to be manually retuned again.

What are the components to CI and how do they interrelate? CI can be viewed from many dimensions. One way to view it is from the perspective of the researcher. The diagram in Figure 2 has been used the past couple of years to depict a view of the functions and resources. It can be argued that not all the items should be included in cyberinfrastructure, but they are all important to the researcher when dealing with CI. The researcher would like them to all work together smoothly and reliably.

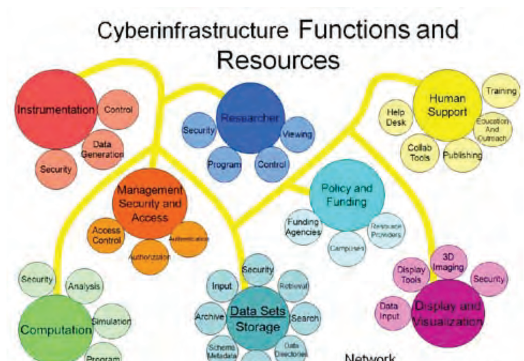


Figure 2.

Another way to view CI components would be to use a modified OSI model where the components include the various parts of the network, the end-station hardware, and the end-station software. Figure 3 depicts some of the components in this view. The Blue boxes represent the traditional computer hardware and operating system (OS), purple the peripherals, and orange the network. The green represents the application software but with today's CI environment the distributed functions that drive CI (many considered middleware) are driven by the application and are not part of the operating system. The end result is that there is a disconnect between the application and the OS. Both are trying to deal with security and performance independently, and without coordination, they are often at odds in accomplishing either. The network is another CI component that also tries to deal with security and performance but it too is not coordinated with other components and can come to similar conflicts. This problem is very hard to resolve because there is no overarching architecture to describe how these components should work together.

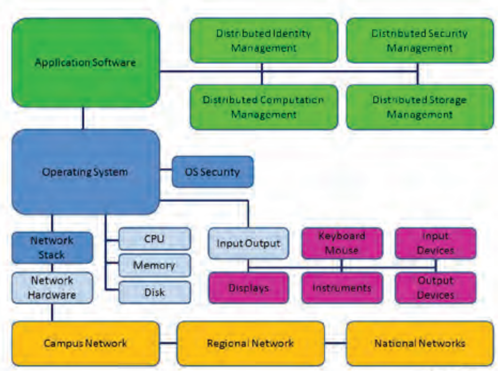


Figure 3.

There are many other ways to view an overarching CI architecture. Each view will show a different set of dependencies and areas in need of coordination. Only after exploration and discussion of these different views will we be able to understand how a CI System can be constructed to provide the features, ease of use, and reliability that most researchers require.

CONCLUSION

Impediments for researchers using advanced CI broadly are that 1) it is perceived as hard to use or too much to learn, and 2) it often breaks or performs poorly. Problems with ease of use are because every new CI tool has a new interface to learn and changing to new tools happens rapidly in today's CI environment. An evolution based on a known user interface would be better accepted by researcher to bring in new features and capabilities. If a flexible base user interface is created, it could serve for new applications with minor changes, rather than an entirely new user interface for each application.

Reliability is difficult to achieve in today's CI environment because each component interacts and depends on the other components, but there is minimal information available for one component to know the state or requirements of those other parts. It often requires a human to examine the overall state of the system and make the necessary corrections. A starting point to improve this would be an understanding of the entire CI architecture. Once dependencies and interactions are understood then means can be determined to advance the coordinated CI system.

OSG Campus Grid Working Meeting Notes

Brian Bockelman (Nebraska), Dan Bradley (Wisconsin), Keith Chadwick (Fermilab), Steve Gallo (Buffalo), Sebastien Goasguen (Clemson), Rob Gardner (Chicago), Sam Hoover (Clemson), John McGee (RENCI), Doug Olson (LBNL), Preston Smith (Purdue), Ben Cotton (Purdue), Prakashan Korambath (UCLA)

January 25, 2010

1 Executive Summary

We summarize here discussions from an OSG working meeting held at Fermilab, January 19-20, 2010, focusing on Campus Grids (CG). The goal was to create a technical summary document identifying and comparing current CG implementations, best practices and patterns, as well as address technical issues moving forward in a number of topical areas.

An important outcome of this effort is that we believe OSG may be a good forum for building a community for establishing best-practices and knowledge sharing about building and evolving such infrastructure within the campus. Conversely, in terms of connecting campuses to national infrastructure it seems obvious OSG should develop a strategy across all existing work areas to support these connections; indeed this may be uniquely the most important strategic area of focus for future generations of OSG.

The topical areas identified during phone discussions held prior to the meeting, including:

- Global file systems and storage systems in use or needed at the CG
- Creation and use of seamless user environments
- Role and use of virtual machines and cloud computing technologies in CG
- Implementation challenges resulting from provincial campus issues, local constraints and priorities

During the meeting we heard of progress or plans from the following CG efforts: GLOW (Wisconsin), Purdue, FermiGrid, University of California Grid, Nebraska, and New York State Grid (a regional CG), each project choosing to describe a number of activities, challenges, and strategic approaches taken in creating their infrastructure. There quickly emerged a number of interesting questions and points we believe are relevant to OSG, and that OSG should have well thought out answers to providing clear project and consortium-wide guidance and vision. We list some of them below, starting from the obvious:

1. Why build campus grid infrastructure and what role can or should OSG play?
2. What incentives are in place to convince faculty, departmental IT managers, CIOs, provosts (and others who may be unfamiliar with campus grids and OSG) to join their resources: a) together across the campus, and b) to national Cyberinfrastructure facilities using for example the services of the OSG?
3. What lessons can be drawn from successful efforts in pooling resources within a CG that can be useful for stimulating opportunistic sharing across OSG? For example, we've seen sharing agreements between contributed "pool" resources among campus researchers far more dynamic than what occurs in the wide-area on OSG.

4. What missing pieces could OSG provide in catalyzing opportunistic sharing or trading of compute (and ultimately storage) resources beyond and between campus grids, within the national Cyberinfrastructure context? How can potential joiners to OSG readily realize and monitor the potential benefit to their organizations?
5. Are there additional principles needed making up OSG's core mission and architecture to catalyze sharing within and across campus grids? For example, resource sharing so far within OSG has been governed within the context of the VO (virtual organization) with its associated required infrastructure and organizational frameworks. However, the diversity and size and opportunity of resources at the CG is more complex and often less VO-centric, especially those without well-established communities already within OSG but have sizable excess resources.

In this document it has been our intention to not offer any solutions or specific, prescriptive implementations or architectures to the complex issues touched by CG, but to point to key principles and issues that pose significant challenges.

2 Opportunities

We see several areas where the development of campus grids can create opportunities for our community:

- Harboring local HTC/HPC experts to help scientists get started (perhaps like TeraGrid's "Campus Champions"). An individual research group may simply not have the continuity of funds and demand for computing to support such people, but making the case for such support at the campus level has succeeded in Nebraska, Purdue, GLOW, Clemson and probably most or all of the successful campus grids. The OSG Engage group has demonstrated how effective and essential this sort of activity is. The more the better.
- Encourage more efficient and professional system administration. This could lead to less downtime, less manual configuration, better system tuning, reduced costs, and more efficient use of the scientist's time.
- Benefit from collective buying power and decision making. Examples where this has been very beneficial have been reported by FermiGrid, Purdue, Clemson and GLOW.
- Achieve more efficient usage of machines and licenses. For example, Purdue and GLOW get about 15% more out of their clusters by sharing them.
- Testing ground for new Grid technologies. If something has wild success at a campus, it might be a breakthrough for OSG as well.
- Participation in a national community which may help leverage local investments, increase revenue and diversify funding sources.

Some opportunities that OSG can create for campus grids are

- Making cross-organizational collaborations practical (sharing data and compute power, authenticating users).
- Bootstrapping campus grid activity in cases where connection to OSG happens first.
- Sharing expertise, best practices, Grid software stacks, documentation and training modules.
- Our community sometimes claims or implicitly assumes that linking resources to OSG will benefit sites by making them part of a collective pool that provides greater computing power to its members than they would get out of their own isolated resources. Therefore, it seems like an item that would

normally be added to this list. However, providing greater total computing power to the users of OSG does not necessarily mean that resource providers are the ones who benefit. If it *is* true that providing resources to OSG is strongly motivated by the expectation of increased computing power, wouldn't this lead to discussions or estimations of value exchange by those who pay the bills? Most concretely, my site expects to spend X on power and system administration to support your VO, but we expect to get Y in return, which is better than simply hibernating our idle machines. At the campus grid scale, part of this discussion is implicit in the university's typical willingness to provide free power to researchers, because the institution hopes to achieve more as a result. So the resource owners mostly just need to agree that the administrative costs of sharing are worth it (and an argument can be made that the administrative costs of a CG are actually less, not more than isolated clusters). So for a campus, sharing seems to support both the resource owners' and the institution's desires for more computing power, without considering other benefits such as increased ease of collaboration. At the inter-campus OSG level it is not obvious that any accounting or expectation of increased computing power is really taking place in the thinking of the resource providers. Resource sharing between VOs for the most part appears to be motivated more along the lines of volunteer computing. We mention this because it seems like a noteworthy difference between the dynamics of a campus grid and the dynamics of the present incarnation of OSG.

3 Difficulties on the Street

The shared cluster model (or condominium computing model) of resource sharing has been very successful on many university campuses (Wisconsin, Duke, Rice, Purdue, Clemson, UCLA among others). In this model, the benefit for a researcher to buy into the program and contribute research dollars for nodes in a community owned cluster is clear: the overall cost of the resource is shared among participating researchers and the IT organization hosting the resource; resource administration and management is typically more professional, moved out of the research lab and into an IT organization; the overall utilization of the system is greater than the alternative of researchers purchasing and managing their own individual systems; researchers maintain an important confidence level similar to individually owned resources as there is a guaranteed level of service commensurate with their level of contribution to the shared system; opportunistic access to more resources than their direct contribution.

One could argue that the OSG Campus Grids initiative, or OSG Campus Shared Resources initiative, is a scaled up version of the shared cluster model that crosses new boundaries and borders such as resources and even campuses. Given the fundamental OSG principles of local autonomy and control, OSG resource owners maintain control of and therefore a guaranteed level of service to their own resources, yet provide unused cycles out to a broader community. By joining the OSG community, researchers gain opportunistic access to significantly more resources than their local system via sharing couched in the framework of the VO model. The overall utilization of resources is greater in this shared model and the overall cost of the infrastructure is shared among the resource owners, and the OSG project.

Given these similarities between the successful shared cluster model and the OSG Campus Resource Sharing initiative, there are some important differences that affect OSG's ability to gain traction and catalyze change at the campus level on a larger scale:

- Infrastructure know-how: Cluster management has been around for a number of years, and is a reasonably well understood challenge. Enterprise wide resource sharing is significantly more complex, not as well understood, and the necessary tooling is less mature. For example configuration, identity, and data management. For a research team that owns a resource, their IT admin burden decreases (typically to zero) in a shift to a shared cluster, but increases non-trivially in a shift to CG style resource sharing.

- Usage know-how (and end-user environments): In all of the OSG connected CG deployments that were examined, it was noted that there are separate submission mechanisms for users to run on the CG vs. OSG. The infrastructure is not yet robust enough to allow for a CG to easily "spill over" jobs from the CG to OSG as the CG reaches capacity. Because of this missing piece, a researcher must make a decision before job submission whether to send it to the CG or to the OSG. Our experience in the campus engagement activity indicates that this lack of integration significantly hinders CG adoption.
- Benefits: In the shared cluster model, there is a more easily quantified *potential value* that can be gained from the opportunistic access. The researcher likely has a sense of the scientific computing landscape on their campus and the other researchers participating in the shared cluster. Thus, they can make a somewhat reasonable informed guess as to how much of the opportunistic cycles they are likely to be successful in acquiring. Given the small size of this community of sharers (relative to a national infrastructure), they can also horse trade and make informal agreements about usage time windows, or simply gain insight into likely times of high availability. Another way of thinking about this is that scope and locality in this model enables a simple marketplace where members can exchange value and benefits from the community resource. This is not possible in OSG today, and the community is too inaccessible for making arrangements for future availability in this way. This also relates to concerns that HEP VOs will overwhelm the system leaving little to no opportunistic availability.
- Incentives: In the shared cluster model, the incentive for increased overall utilization of resources is in fact a fiduciary responsibility of the campus CIO: to provide capable and cost effective research computing to faculty and staff on campus. In OSG, the incentive for increased overall resource utilization is rooted in the VO's, groups of like-minded researchers working for the betterment of a specific science community. Each of these individually works very well. However, mixing the two is like mixing oil and water. Without some form of tangible and quantifiable value exchange, CIO's simply cannot justify the sharing of campus resources beyond the borders of his/her administrative purview. In the cases where this has been successful (Purdue, Clemson, Fermilab, NY State Grid), there were trail blazing thought leaders in place who were willing to contribute campus resources and take a leadership position in the national Cyberinfrastructure. In doing so, they likely enjoyed important benefits that are difficult to quantify, such as help in winning future awards, attracting faculty, etc. However, this level of incentive quantification is not sufficient for broad adoption of CG infrastructures.

4 Outreach and Engagement

As we have been discussing, without CIO-level support a campus grid project is not likely to gain any traction. Support from the CIO is important for making resources and personnel available to the project. In some cases, for example Purdue and Clemson, it is the CIO who drives the creation of the campus grid. However, some institutions may find the CIO indifferent or even opposed to the idea of creating a campus grid (some have reported that campus resources are dedicated to students and should not be used by external users). The convincing argument will depend on the specific objections of the CIO, but it is worth discussing some of the more compelling arguments here: namely that a campus grid can be set up to take advantage of existing underutilized resources. In 2009, Purdue provided 17 million hours of compute time on it's DiaGrid, i.e. 17% of the total HPC hours that year. For a campus without a traditional cluster, the case is even stronger as scientists can spend more time analyzing data instead of waiting for computations to finish.

While having an engaged CIO is beneficial, this of course does not preclude campus grid creation within academic divisions on campus; however this usually involves senior involvement of institute heads or Deans with fiscal authority over the required campus resources (space, power, cooling, networking) and limits the size of the contribution that the CG can make globally, even though it can still prove to be a valuable local resource.

It is in fact the research scientist who is the ultimate driver of a successful campus grid. The best technical setup in the world is of little use if there is no work being done on it. As a result, reaching out to researchers is a critical part of establishing a campus grid. Other researchers who have their own success stories with a campus grid can be the best advocates, not only to their peers, but also to the CIO. Indeed, a clamor from the research faculty may be the compelling argument that gets the CIO support when all other approaches have failed. Indeed if faculty do not ask for it why would a CG be created, often other projects have higher priorities.

By approaching the researchers with real-world cases showing how a campus grid can help meet the research and instructional missions of an institution, the necessary support can be achieved. However, the outreach to and engagement of researchers does not end when the campus grid is declared operational. In order to provide ongoing success the researchers must continue to be involved. This means modifying the setup to provide the necessary environment as well as providing support for understanding how to best make use of the grid. The end result should be that researchers could focus on their area of interest, and not have to worry about maintaining their research-computing environment. While OSG can help startup the engagement and outreach efforts, long-term sustainability will be based on local support structure.

5 Strategies for Resource Aggregation

At the root of campus grids is the idea of sharing and aggregating the resources on a campus. At institutions with successful campus grid programs, several different strategies exist for aggregating resources into a grid.

5.1 Aggregation approaches in use today

GLOW

GLOW, at the University of Wisconsin, aggregates physically distributed collections of systems into a single Condor pool. For example, the engineering group acquires, houses, powers, and cools its own collection of machines in its own space, as does the Physics department. Software and operating systems are managed centrally.

A site can join GLOW with a minimum contribution of about one rack of machines. The University and the Condor Team adds value to the campus grid with additional opportunistic resources.

GLOW provides access to the AFS software repositories of its members. Other than this, there is no shared file system.

Purdue

Purdue aggregates resources both via its "community cluster" program, where faculty research dollars are pooled together to build a single large cluster, with professional system administration, support, and facilities; and the use of Condor to tie together otherwise idle machines in student labs and around the campus with idle cluster nodes. A group can join the community clusters with a single node.

Distributed Condor resources at Purdue are managed with the "CycleServer" management console - useful for maintaining Condor configuration on machines with distributed ownership. A tool for configuring, monitoring, etc is useful to administrators operating a grid of systems owned by multiple groups.

Purdue provides pre-configured packages to aid departments with adding resources to the campus grid.

Purdue clusters share centralized NFS servers.

FermiGrid

Rather than combining individual cluster node systems into a central grid or cluster, FermiGrid aggregates many distinct, previously independently owned and operated clusters together with middleware - placing all clusters behind a single point of entry. All FermiGrid clusters share site-level services (GUMS, SAZ, MyProxy, VOMS, etc)

FermiGrid clusters share centralized NFS servers (using BlueArc).

NYSGrid

The New York State Grid initially aggregated together Blue Gene systems at New York state universities for others to use. Additionally, Buffalo is aggregating campus resources with Condor, and backfilling HPC clusters with Condor, much like Purdue's model, with one or two gatekeepers functioning as an entry point.

Other Institutions

- Clemson aggregates resources in a model very similar to Purdue: condominium cluster and large condor pools centrally managed. It is important to note that Purdue former CIO is now Clemson's CIO.
- Nebraska leadership is supportive of sharing and combining resources and is taking advantage of opportunistic funding opportunities (gifts) for broad benefit to campus researchers.
- UC (California) Grid presents distributed resources (at many UC campuses) through a single portal, providing easy access for job submission and monitoring, and data transfers.

6 Connecting Researchers to Resources

During our discussions it became clear that (somewhat paradoxically) it is helpful if the focus is less on "building Campus Grid infrastructure" and more on connecting researchers to any resources available to them. We decided to dissect this approach somewhat. Thus for campus grids, there are three parts to the subject of "Connecting Researchers to Resources":

1) Resources: Some sufficiently enticing and easy to use computational or storage resource on campus that can be used for scientific computing.
2) Researchers: Finding and maintaining the interest of campus researchers who have a need for computing in order to get their science done.
3) Connecting: Selecting the right level of engagement for researcher based on the present and future need of the science. In this section, we are not going to tackle the acquisition of resources but rather finding and growing the needs of researchers and connecting them to the types of scientific computing that suit them best.

We consider two CG efforts in this section – experience garnered by doing these activities at the University of Nebraska-Lincoln and at the University of California Grid. Both of these examples provide useful insights into the thinking and considerations that are required from the resource sharing, user interface, and application / workflow porting perspectives.

6.1 Connecting Nebraska Researchers

In Nebraska, new researchers are using the following means:

1) Top-down approach: UNL enjoys strong support from the Office of Research, which encourages scientists to partner with the Holland Computing Center when additional computing is needed and does not allow new grants to purchase their own computing resources outside HCC.
2) "User Recommendation": Users often are recommended by their peers, whether they are new research groups recommended by colleagues who use HCC or by new students who are joining a research group that is already utilizing HCC.
3) Active Engagement: Some research groups have joined after having directly talked to heads of departments, deans, or research group leads. UNL feels this is one method to "break in" to a new campus or department where we have no active users.
4) Education: HCC staff members teach CSE classes almost every semester. Topics in the past have included system administration, parallel programming, cluster computing, and grid computing. Students who believe they need scientific computing (or whose advisers believe this) often take these classes and get involved with HCC through their class project. These classes are used as a recruitment tool for student workers. A local workshop is offered approximately once a year. UNL has never examined the "retention rate" for active users, or thoroughly examined the reasons why active users become inactive.

Before we going into detail of how researchers are connected to resources, a few definitions are given (one can skip this section if they are familiar with all the keywords).

Primary types of resources:

1) Commodity Linux clusters: Clusters composed of low-to-mid-range server hardware; commodity Ethernet network; small number (<=16) of cores per node; 1-3 GB RAM per core.
2) Tightly coupled clusters: Commodity Linux clusters with a low-latency network.
3) Specialty resources: Machines serving a specific niche purpose not well suited for general usage or non-dedicated applications. Examples include GPU-equipped machines, SGI Altix / large memory single-system-image machines, and possibly machines with non-x86 architectures.

We also divide the jobs up into general classes:

1) High throughput: A large number of single-core jobs; usually a large number of jobs (hundreds to tens of thousands) form a single workflow, which might have trivial or complex interdependencies (a large number of jobs should be able to be run simultaneously). Parallelism is achieved through running additional jobs.
2) High performance: MPI or other massively parallel jobs. These tend to take up significant amounts of computing resources - many, perhaps hundreds, of nodes. Usually, a small number (<10) batch system jobs per workflow
3) High throughput, high performance: Workflows that mix the characteristics of high performance and throughput; usually multi-core jobs running on a single machine.
4) Specialty - jobs that can only run on specialty resources.

At UNL the following pairings have made the most sense:

1) Specialty jobs can only be run on specialty resources. The users with specialty jobs often have more computing expertise; they know what they want and they can take care of themselves if they have access to the resource. No interest in distributed or grid computing. Care must be taken to take the added cost for hardware support and purchase into consideration.
2) HPC jobs: Generally, HPC jobs can only run on HPC resources. The portability of these jobs is low, as the researcher is often interested in intimately tuning the jobs to the machine (we've had experts state that it takes 1-2 months to "break in" a new machine for their code); it can take many recompiles per machine to get the desired performance. These workflows may not scale well by adding additional cores to the jobs, which is why compiler settings are so important. Amongst these jobs, there is usually low interest in distributed computing, and hence a low probability of success for Engagement with campus grids effort. The only potential successes we foresee are from converting those users whose jobs are really HTC to use HTC methods; for example, there are embarrassingly parallel workflows ideal for HTC that are implemented in MPI because that was the only tool the researcher was familiar with.
3) HTPC jobs are only run on tightly-coupled clusters at Nebraska. These have the potential to run on commodity cluster and a core stakeholder (CMS) may express interest in this. Nebraska will probably wait for guidance and leadership from OSG Satellite for running HTPC jobs on the grid. There are a small number of local users whose jobs might fit this description; they may eventually be a target for the campus grid. Currently, the cost for porting is too high and the potential for increased resources for these users is too low.
4) HTC jobs: These jobs can and are run almost anywhere - HPC, HTC, or even specialty resources. These workflows scale by adding additional jobs - there is less focus to highly tune the job to each machine they run on. These users are more interested in distributed computing and their jobs have a higher probability to be successfully ported to grids.

At Nebraska, the hope is to offload as many researchers off to a campus grid or the OSG as possible. Porting a job to the campus grid increases utilization of all local resources and increases the resources available to a single research group. UNL currently **does not** attempt to port a job to the campus grid if they meet any of the following criteria:

1) Software requires licensing or license server.
2) Software requires multiple cores per job.
3) Workflow can be done within the desired timeframe regardless of how busy the cluster is. I.e., almost any local user should be able to finish a 1,000 compute hour workflow overnight through fair-share; if the turn-around of 8 hours is acceptable for the user, there will likely never be a need to use distributed clusters.
4) Workflows which are highly data-intensive (more than 1GB of input per job)

(1) and (2) are software limitations that can be solved, but the solutions are complex enough that the costs outweigh the benefits. (3) is difficult as it is anticipated that clusters will become increasingly over-subscribed; if in doubt, UNL doesn't apply it. It is important to consider the potential science benefit versus the cost of HCC support time; there are cases where a HTC workflow meets all the criteria for running on the OSG except actually being "large enough". When a researcher has HTC jobs that don't meet any of the exclusion criteria above, UNL envisions the following steps for a successful campus grid user:

1) Run application interactively on any HCC cluster, and run the workflow from start to finish for one path (i.e., if the workflow is a sweep through 3,000 input parameter sets, verify they can run it on 1 parameter set). The HCC effort involved is usually the systems administrators installing new software dependencies or any HCC employee helping with Unix basics.

2) Port the application to the Condor cluster. Express the workflow dependencies are expressed as a Condor DAGMan. Each individual Condor job should encompass a "reasonable" amount of work (i.e., should be between 30 minutes and 8 hours long, with an average of 2 hours). Data dependencies should be well defined and expressed in the condor job; the user should not depend on a shared file system. An HCC application expert or integration expert will be able to help here. It is possible that, after this step is completed, the user is satisfied for quite awhile. It is possible they may not continue to step 3 until their computing needs increase or usage of the condor cluster increases.

3) Port the application to the campus grid. Currently, this is done by individual engagement with the integration expert. Usually, this involves:

    a. Getting the user a grid certificate and adding them to the GPN (or Engage?) VO. Training with OSG grid basics.

    b. Deploying their application code to all sites supporting GPN.

    c. Modifying their Condor submit script slightly to use OSG-MM (match-maker) instead of "normal" Condor. UNL already has implemented a mostly automated tool to do this. UNL evaluating GlideInWMS that should make this step even easier.

UNL believes (2) is a crucial step in order to allow a user to fully debug their application locally (where it is easier to separate condor errors from application errors than separate condor-g errors from application errors). It also provides the cleanest transition from a single local resource to an OSG-like resource, especially when the data dependencies are correctly expressed.

6.2 Connecting University of California Researchers

We now turn to experience from the UC Grid project that has focused principally on web-portal based designs. The architecture and the web interface for the UC Grid Portal was evolved from experience as an organization giving extensive consulting help to users who run high performance computing applications. Some observations are:

1) Majority of the CPU time-consuming users are using commercial or precompiled applications such as Gaussian, NWChem, Matlab, Mathematica, R, Q-Chem, Amber etc.

2) UC Grid found while most of the young researchers have extensive backgrounds in browsing and using the web, they lack experience in command line computing interfaces.

3) They also have to learn basic data management commands such as those that group files into a single tar file, transfer files between their local machine and the cluster(s) they are using, etc.

4) Different clusters use different job schedulers such as condor, SGE or Torque. This often confuses users who already don't know much about Linux or Unix.

5) Some researchers do collaborative research with researchers at other UC campuses and/or with other university campus researchers. So there is a need to authenticate them within their campus as well as outside their campus.

6) There are a lot of idle resources on many of the clusters but the cluster owners are hesitant to share those resources to others due to lack of secure transfer of those resources to unknown users.

7) License fee for some of the commercial applications are very expensive and often times a lot of unused licenses are seen.

As a solution to this UC Grid designed a Job submission service and a Data Manager Service with a web interface to upload input file or specify arguments. In order to use the job submission service the users only need to upload the input file and choose the number of processors and duration of the job. The Grid portal system will submit appropriate GRAM jobs calling suitable commands because the Grid Portal maintains a database that knows exactly how to run a specific application on each participating cluster, where the executables reside, how they are invoked, default arguments, etc. For generic job submission users will have to choose their executable and other parameters. Data Manager services uses GridFTP to transfer files. UC Grid chose Globus Toolkit as the underlying grid software because of its wide usage in Teragrid and other national grids such as OSG. This also provides a common user authentication mechanism through the use of X-509 based certificates.

UC Grid also wanted anybody in any of the UC campuses to have a unique certificate so that he or she can be uniquely identified from anywhere. This led to the creation of a single certifying authority for all campuses. As all of the ten UC campuses use Shibboleth, it was decided to use the campus Shibboleth authentication as the basic service to authenticate and issue the certificate. As of now a UC Grid certificate is used only for the services from the UC Grid portal. It was therefore decided to keep the user certificate with the portal and allow users lease only the short-lived credentials as and when they login at the portal from a MyProxy server.

Finally, UC Grid added some of the services such as interactive login through VNC due to user demands for real time code development and debugging.

7 High Availability Services at the Campus Grid Scale

In the evolution of a campus Grid, a large number of resources will eventually wind up being dependent on various services. The traditional set of these services includes:

- Space and power
- Environmental management (heating, cooling, humidification)
- Networking (physical network, DNS)
- Staff

There are various "traditional" methods to deal with these service dependencies:

- Generators and UPS for power
- Multiple CRAC units so that failure of a single unit does not impact the environment
- Use of switch and router capabilities to provide network fault tolerance
- Redundant DNS (primary and secondary)

The campus Grid may also introduce dependencies on additional Grid specific services, such as (using FermiGrid as a model):

- Virtual Organization Membership Service (VOMS)
- Grid User Mapping Service (GUMS)
- Site AuthoriZation (SAZ) Service
- Squid Web Cache
- MyProxy

For the initial building of the campus grid, these Grid specific services can be provisioned as non-redundant services, but as the campus Grid grows, there will come a point when the Grid specific services will need to be commissioned in a fault tolerant or highly available (HA) infrastructure (an outage of either GUMS or SAZ has the potential to impact >5K jobs/hour on a 20K job slot campus Grid). FermiGrid is addressing this vigorously.

In the case of a regional Grid, the need for redundant services is even greater, since a power or network outage in one administrative area of the regional Grid has the potential to impact the availability of resources across the entire regional Grid.

The need to deploy these Grid specific services in a redundant deployment may be viewed as requiring a large amount of hardware resources to accomplish. Fortunately, virtualization can be used to deploy these services on a minimal hardware footprint.

For the set of Grid specific services listed above, FermiGrid (the Fermilab Campus Grid) has developed configurations that can support in excess of 20K job slots on two (appropriately configured) “midrange” systems. These configurations are freely available, and the FermiGrid personnel are willing to consult on the necessary deployments for campus Grids that are considering joining the Open Science Grid.

8 Leveraging Emerging Technologies

There are a number of emerging technologies affecting the landscape of campus research computing and it is sensible that we consider that some or all of these will have an effect on the interfaces between the future OSG and campus research computing. Some of these technologies have been around for a while but are recently beginning to have a noticeable impact.

As just discussed, one technology is the Shibboleth identity management technology that is being used on an increasing number of campuses for campus-wide authentication; clearly it has an important role in the interface with users and newcomers to scientific computing. Coupled with other available technologies like GridShib and MyProxy one could potentially achieve a much simpler user experience around authentication than is used on OSG today. To be most effective there would need to be some policy work on the acceptance of credential stores at the IGTF level but one can consider if the international acceptance of credentials is really needed for all grid users. Clemson who is part of the InCommon federation has demonstrated peering with the NCSA Gridshib CA, which enabled Clemson faculty to obtain a short-lived NCSA proxy certificate which gave them access to TeraGrid resources.

Another technology that has been developing for years is virtual machines. These are used at a number of campuses in the “traditional” areas of server consolidation and resource management and are more recently being explored for more widespread scientific computing. From the end user perspective the key driver is being able to run the same application environment everywhere. From the resource providers perspective it is a way to provide application-specific resources on a temporary or shared basis without having to dedicate resources to individual applications. This is similar to the server consolidation concept but moving into broader application areas.

The third emerging technology, related to virtual machines, is cloud computing. The deployment of the Globus-Nimbus, Eucalyptus and Opennebula software as a means to provide a common interface across numerous distributed resources is showing to be amenable to many application deployment scenarios. A key driver is the relatively clean separation of resource management from application environment so the same resource can be quickly re-configured for different uses. In the scientific computing domain there are open questions and issues about the performance that can be achieved, primarily in the storage and network communications domains. The Magellan project at NERSC and ANL is focused on understanding the cost and performance issues of the cloud paradigm with respect to scientific computing and the relation of public and private clouds.

Data Approaches for Campus Bridging

Michael Shoffner, Ilia Baldine, Leesa Brieger, Jason Coposky, Kevin Gamiel, Jefferson Heard, Howard Lander, Anirban Mandal, John McGee, Nassib Nassar, Arcot Rajasekar, Charles Schmitt, Erik Scott, Michael Stealey, Stanley Ahalt

RENCI Data Working Group
Renaissance Computing Institute (RENCI), University of North Carolina at Chapel Hill

1. Introduction

This paper presents a brief overview of RENCI's data challenges and proposed approaches to aid the efforts of the NSF Campus Bridging Task Force as well as the broader CF21 efforts in support of discovery, research, and learning. RENCI develops and deploys advanced cyberinfrastructure (CI) to enable research discoveries and practical innovations. Founded in 2004 as a major collaboration involving University of North Carolina Chapel Hill (UNC-CH), Duke University, and North Carolina State University (NCSU), RENCI is a statewide virtual organization with facilities at campuses across North Carolina and partners that include federal agencies, other research institutes, the NC community college system, and industry.

2. Data Challenges

RENCI investigates problems in areas such as weather, genomic sciences, oceanography, visualization, and digital archives and collaborates with researchers throughout the U.S. in the usage of national resources including TeraGrid and the Open Science Grid (OSG). In these projects, we encounter a number of complex data challenges, including:

- **Discovery**—finding relevant data within distributed data systems
- **Integration**—combining data from many sources with divergent metadata into integrated data sets for specific use cases
- **Policy**—specifying data policies and access, utilization, and redistribution rights across multiple levels of collaboration
- **Curation**—implementing preservation and sharing that maintains policies on data sets replicated remotely
- **Granularity**—adequately managing both very large files (e.g., WRF weather forecasts) and very many small files (e.g., gene sequence reads)
- **Management**—general data life cycle support; support for distributed archives; controlled access to shared data resources; and adequate association of metadata with data for provenance tracking, auditing, and integrity checking
- **Data Placement Services**—higher-level functional services for the movement and placement of data as required by computational and analysis workflows
- **Privacy and Security**—privacy and security assurances for healthcare, social science, and other sensitive data
- **Scalability**—the growing number and complexity of structured and unstructured data sets, streams, and performance
- **Fault Tolerance**—support for availability, redundancy, and concurrency required for distributed scientific workflows
- **Distributed/Mobile Computing**—data management when network connectivity is limited and/or unstable
- **Provenance, Validation, and Versioning**—for observational data, annotations, and derivative data

RENCI's data collection is projected to grow a minimum of 1.5 petabytes over the next three years. The types and purposes of the data are manifold, ranging from video testimonials from Holocaust survivors to a collection of one thousand human genomes. RENCI assists researchers from across the U.S. to leverage both TeraGrid and OSG via a hosted TeraGrid Science Gateway and by leading the OSG Engagement program. To date, these efforts have focused on large-scale high throughput computing; during 2009 we brokered nearly 2 million jobs consuming 8 million CPU hours run across more than twenty sites for users from fifteen unique campuses and five science domains. The current data management landscape requires that as a broker of national CI services we manage the data flows from the researcher's home institution, through RENCI and out to the national resources, providing results back to the home institution. Our experience indicates that the majority of researchers on campuses do not have the local IT experience or capacity necessary to implement the data management security and transfer mechanisms used by national infrastructures.

3. Solutions

These data challenges represent obstacles to discovery, research, and learning faced by universities across the country, and they are without easy solutions. RENCI is actively addressing these challenges through a number of initiatives and partnerships, while investing significant resources into data systems and the development of reusable technology.

iRODS

RENCI is adopting data grid technology as a step toward addressing a number of the above challenges. The integrated Rule-Oriented Data System (iRODS) data grid is based on expertise gained over a decade of support of data grids, digital libraries, persistent archives, and real-time data systems. It is developed by the UNC-CH Data Intensive Cyberinfrastructure Environments (DICE), which collaborates closely with RENCI to promote the development and deployment of the technology to user communities under the open source BSD license. iRODS provides a middleware layer that allows the integration of distributed and heterogeneous resources into a structure that implements data management policy across administrative sites, disparate systems, and diverse user groups[1]:

- National Archives and Records Administration (NARA) Transcontinental Persistent Archives Prototype (TPAP) demonstrates the capability to set up separate NARA data grids and federate among them so that National Archive data can be mirrored at several sites around the nation using iRODS microservices that implement archival policy[2]. Federation between the distributed, separate NARA data grids allows for ongoing synchronization among them.
- TUCASI Infrastructure Project (TIP) establishes a federated data-sharing environment among UNC-CH, Duke, and NCSU. iRODS is one of the foundational elements for prototyping this environment and addresses policy issues such as intellectual property management for content such as courseware. Having an infrastructure in place that allows the controlled sharing of content gives the TIP partners a significant amount of choice and flexibility.
- The RENCI/NCSA MotifNetwork is an NSF-funded system that consists of a suite of distributed workflows incorporating ensemble and parallel processing to perform large-scale informatics analyses[3]. Collaborators have access to data produced by the workflow without needing accounts at any of the participating sites because iRODS manages authentication and authorization. iRODS also replicates output into the UNC-CH mass storage resource which backs up output results data.
- NSF DataNet. The partners of the DataNet proposal consider that a data grid will be crucial for sharing data and collaborative projects. These partners are Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI), Temporal Dynamics of Learning Center (TDLC), CIBER-U, iPlant Collaborative, Ocean Observatories Initiative (OOI), and the UNC-CH Odum Institute for Research in Social Science. Some partners will depend on the data grid to support their publishing, curation, and preservation of educational materials and others will use it for sharing data with other partners. All welcome the opportunity to implement their own data policies in their iRODS systems. RENCI is responsible for facilities and operations for DataNet, which includes deployment and customization of the technology to the six science partner groups along with user support and training.
- RENCI Virtual Organization (R-VO). RENCI's interconnected visualization resources include its multi-touch technology integrated with a 360-degree Social Computing Room and fifteen-foot diameter dome at UNC-CH. The R-VO spans twenty-seven RENCI visualization and collaboration facilities located on campuses in seven geographic regions across North Carolina and is currently the largest statewide coordinated distributed visualization effort in the nation. R-VO uses iRODS to manage data across its statewide infrastructure.

Sensor Data Bus

The Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) was formed to explore the many dimensions of emerging CI and sensing technologies that need to be brought together to serve hydrologic science and related communities. The CUAHSI community has developed an ecosystem of tooling and interoperable services that suit observations that can be related geo-spatiotemporally to a point, trajectory, or grid along with associated annotations and

---

[1] Moore, R., Marciano, R., Rajasekar, A., de Torcy, A., Hou, C.Y., Brieger, L., Crabtree, J., Ward, J., Chua, M., UNC Chapel Hill; Schroeder, W., Wan, M., Chen, S.Y., UCSD (2009). NITRD iRODS Demonstration. Hosted by NARA at NSF. https://www.irods.org/pubs/iRODS_NITRD-Report-0910.pdf

[2] Ward, J., de Torcy, A., Chua, M., Crabtree, J. (2009). Extracting and Ingesting DDI Metadata and Digital Objects from a Data Archive into the iRODS extension of the NARA TPAP using the OAI-PMH. https://www.irods.org/pubs/DICE_eScience_Paper-2Oct2009-rev-last.pdf

[3] Tilson, J.L., Rendon, G., Jakobsson, E. (2009). MotifNetwork: High throughput determination of Evolutionary Domain Network. Proceedings of the 2009 International Conference on Bioinformatics and Computational Biology (BIOCOMP09), July 13-16, 2009.

metadata[4]. RENCI's NC Sensor Data Bus (SDB) project[5] builds upon and contributes to this ecosystem, enabling us to participate in an active national network of data interoperability. By joining this community and consuming/contributing technologies, all RENCI projects involving geo-spatiotemporal relational data can now interoperate with human and software systems across campus, regional, and national boundaries.

DataLab

The RENCI DataLab is a research project to explore new approaches and standards for addressing the types of data issues that can limit collaborations between separate research groups. Approaches are needed that allow distributed members of scientific communities to share and integrate data easily without the need for database specialists, semantic experts, and Web programmers. By focusing on protocols and abstractions that lower the barrier for researchers to publish to and retrieve data from collaborators we believe the CI community can build tools for inclusive and productive community-based science[6].

One DataLab approach is lightweight data retrieval protocols that allow researchers to access, filter, and query distributed data via URIs. We have been working with the California Digital Library to implement, extend, and evaluate the use of the THUMP protocol for scientific data sharing. We are exploring the effectiveness of this approach in collaborations that involve the UNC Coastal Studies Institute, the UNC-CH Medical Genetics Department, the UNC-CH Kenan-Flagler Business School, and the UNC-CH Department of Public Policy. We are also developing a simplified data abstraction called Schema-N to allow researchers to easily generate normalized relational models of scientific data.

Interoperability

While iRODS microservices can be developed to deliver sophisticated distributed data services, the data grid technology can also be integrated with other special-purpose data presentation tools to provide a preservation environment that underpins and supports overall data services and management. The integration of Fedora (DuraSpace)[7] is on track to bring the rich metadata and searching capabilities of Fedora together with the distributed preservation services of iRODS. Similarly, the integration of DataVerse with iRODS allows social scientists to continue using the access and statistical methods of their community package on data sets that are curated by an iRODS data grid. As RENCI grows its collections for the SDB and DataLab services, the integration of these tools with iRODS will allow us to implement data curation policy while also delivering targeted services for these structured and geographical data sets. We are also pursing further integration of our data technologies with the national computing CI and investigating integration against open cloud computing platforms.

4. Data Approaches for Campus Bridging

It is our belief that to maximally enable data exchange between communities and campuses, NSF should facilitate and support these approaches:

A. Documenting and making available patterns, explanations, policies, processes, workflows, architectures, case studies, and other elements of a body of practice to address complex data interoperability challenges in the STEM community.
B. Making available open source technologies and tools that support policy-driven, shared data collections at various levels such as campus-to-campus and campus-to-national CI.
C. Encouraging standards development and implementations such as CUAHSI's Open Geospatial Consortium work.
D. Providing a mechanism to make available the body of practice, standards, and technology tools to the STEM research and campus communities. Options include standard information channels (papers, workshops, etc.) as well as funding initiatives to support collaborative partnerships for knowledge and technology transfer.
E. Leveraging organizations with existing collaborative capabilities and relationships in order to achieve effectiveness over the next twelve to eighteen months.

---

[4] Zaslavsky, I. (2010). National Science Foundation TeraGrid Workshop on Cyber-GIS, Washington, DC, Feb. 2-3, 2010. http://www.cigi.uiuc.edu/cybergis/docs/Zaslavsky_Position_Paper.pdf

[5] Sensor Data Bus. Web site at http://www.sensordatabus.org

[6] Nassar, N., Kunze, J.A., Newby, G.B., and Gamiel, K. (2009). Sarcomere: A System for Data Interoperability (poster). Presented at the 5th International Digital Curation Conference, London, England, December 2009.

[7] Zhu, B., Marciano, R., Moore, R. Enabling Inter-repository Access Management between iRODS and Fedora. 4th International Conference on Open Repositories, Atlanta, Georgia, U.S.A. May 18-21, 2009. http://smartech.gatech.edu/dspace/handle/1853/28494

Enabling and Sustaining Campus-to-Campus Cyberinfrastructure

Gary Crane, Southeastern Universities Research Association
John-Paul Robinson, University of Alabama Birmingham
Phil Smith, Texas Tech University

1. Introduction

Cyberinfrastructure resources are deployed at many levels in the nation's academic research community, but the use of those resources is critically hindered by a lack of accessibility in the software and the social models employed to administer those resources. This lack of organization and accessibility has created inefficiencies and gaps in our ability to successfully utilize available CI resources in support of the academic enterprise. Identity management and trust are particularly thorny issues, but many other technical issues loom as well. There have been concerted and well-funded Federal efforts to bridge some of these gaps between campus and national CI services. What we have not seen is a similar effort to lower the barriers to the deployment of campus-to-campus CI services, services deployed for use within and between campuses. This is a fundamental gap that campus IT organizations are seeking to bridge.

SURAgrid (http://www.sura.org/suragrid), a community of campuses engaged in the adoption of a coordinated campus-to-campus CI, is working to bridge the CI gaps between its member institutions. SURAgrid participants recognize the advantage of working with a community of peers to address the many existing barriers to the effective deployment of CI services. The SURAgrid model allows expertise to be shared and not duplicated across member campuses. The campus-to-campus barrier exhibits most if not all the problems encountered in linking up disparate CI. One could view inter-campus CI deployment efforts as addressing the "horizontal" problems of CI (intra- and inter-campus CI services) whereas efforts like the TeraGrid address the "vertical" problem (researcher to national CI services).

2. The Growing Need for and Challenges of CI Tools and Services

A desire to grow our collective body of knowledge has connected people throughout history. Each generation has leveraged the tools of their era to share insights and build common understanding. The Internet has most recently and strikingly accelerated our ability to collaborate across significant geographic and organizational boundaries. This ease and density of interconnection has led us to a point where a new infrastructure – Cyberinfrastructure[1] – is evolving to advance peer interaction and shared discovery.

The promise of a Cyberinfrastructure (CI) that provides seamless collaboration across all boundaries is widely discussed today, within and beyond the research and education (R&E) community. Within R&E in particular we have shared this vision for nearly a decade[2] but still find ourselves facing numerous hurdles in implementation. Impressive CI resources have been established through nationally focused funding initiatives. Significant obstacles must be eliminated, however, before these resources can function as a pervasive CI from which the vast majority of researchers can benefit. Many of the obstructions are well known; others we are just beginning to see and understand.

3. A Horizontal CI Perspective

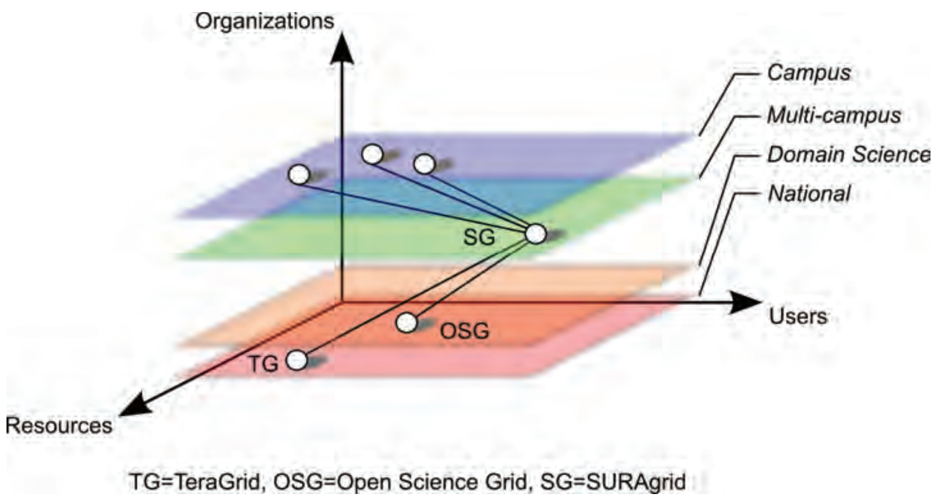
The barriers to successful adoption of CI are easier to understand by considering the three major dimensions of the CI landscape; **Resources**, **Users** and **Organization**s. The figure below is an illustration of the national CI environment that considers both horizontal and vertical communities of practice and includes campus, multi-campus and national CI resources and helps to identify unmet needs that confront the development of a coherent national CI. This diagram emphasizes the "horizontal" perspective of a campus-to-campus coordinated CI that has not yet been effectively addressed in both funding and effort. While most recent dialog and major funding initiatives have focused on the vertical issue of creating and accessing national CI resource centers (the NSF Tier 1 and Tier 2 centers), the potential to maximize and leverage the large aggregate number of users, resources and organizations at the nation's academic campuses has largely been overlooked. It is campuses, however, that are the fundamental

---

1 Cyberinfrastructure consists of computational systems, data and information management, advanced instruments, visualization environments, and people, all linked together by software and advanced networks to improve scholarly productivity and enable knowledge breakthroughs and discoveries not otherwise possible. From: "Developing a Coherent Cyberinfrastructure from Local Campus to National Facilities: Challenges and Strategies", pg 4. http://www.educause.edu/Resources/DevelopingaCoherentCyberinfras/169441

2 Atkins, Daniel E. (Panel Chair), et al. *Revolutionizing Science and Engineering Through Cyberinfrastructure:* Report of the National Science Foundation Blue Ribbon Advisory Panel on Cyberinfrastructure, January 2003, http://www.nsf.gov/publications/pub_summ.jsp?ods_key=cise051203

organization of academics. It is at this horizontal level that many barriers are encountered, many of which can be most effectively and sustainably addressed by leveraging resources, perspectives, and collaborations across campus boundaries.



4. Barriers to CI Adoption and Use

As discussed above, the landscape of CI adoption includes both vertical and horizontal perspectives. Barriers to the effective use and adoption are also varied and may be technological, sociocultural, or policy-related with interdependencies that add to the complexity. Over the past 10 years, technologies like the grid have made it significantly easier to meet some of these challenges. For example, in the areas of high performance and high throughput computing, large numbers of similarly configured resources can now be harnessed as a coordinated pool with significantly reduced effort. Still, incompatibilities between systems are real, and differences in the size of memory, disk space, inter-process communication tools and schedulers seriously impact the transparency of the infrastructure. The responsibility for resolving these incompatibilities can be shifted to dedicated personnel, however, the effort required to build traditional organizations with sufficient talent to address these challenges is considerable and effectively prohibits all but the most determined and well funded campuses or research teams from engaging in the effort to harness large resource pools.

Many research and development groups are only beginning to require resources beyond their local reach and do not have funding to build the full spectrum of expertise dedicated to integrating distributed or advanced technologies into their workflow. Most of the support required to harness CI resources would need to come from expertise they do not posses. This is especially true for individual researchers and small research groups at campuses with limited institutional support for CI services. These researchers know that more is possible; they just don't know how to get started short of becoming technology experts themselves. What is needed is an effective way for individuals and organizations to engage a targeted fraction of their efforts in a larger, open community dedicated to co-development of shared CI.

Expanding our definition of CI to include horizontal (multi-campus) initiatives can significantly complement other efforts for faster progress. These initiatives in general are driven by local (campus) needs and provide foundations for tool development based on the desire to collaborate. Broad and de-centralized initiatives such as SURAgrid are enablers of this evolution, with the ability to draw from national initiatives and further support tools and services that become integral components of the campus IT infrastructure, creating notable efficiencies of scale in their use. Effectively organizing multiple campuses to address a given problem is a significant challenge. Multi-campus initiatives, formed from the desire to solve common problems with shared solutions, are the most naturally suited to building the trust and shared work environment needed to address these challenges.

Great economies of scale can be realized through aggregation and coordination at human, system and enterprise levels. This type of sharing freely mixes viewpoints and talents from various stakeholders – researchers, faculty, students, CIOs, and IT support staff – for successful adoption of CI at the campus and beyond – a necessary foundation for large-scale CI success in R&E. Results can be as quantifiable as group-buying power, or as subtle as influencing industry direction; as simple as easing access for an individual researcher, or as far-reaching as promoting and expanding the value of federated identity.

National funding priorities have made significant progress toward establishing large national CI resource centers that are currently part of the TeraGrid, and generic interfacing technologies, like Globus, that enable coordination of infrastructure across administrative domains. Organizations like the Open Science Grid, caBIG and the science gateways of TeraGrid have adapted and evolved to meet the needs of specific science communities. However, the job of implementing and sustaining a truly broad-scale and integrated CI remains a major unsolved distributed systems challenge. It in itself poses a "big-science" problem – how to address social, cultural, technological, and economic challenges inherent in building a CI that can serve the competing needs of the many user communities found even within a single campus.

5. Towards Effective Investment in Horizontal CI Initiatives

Considering the elements described above, we arrive at a final but critical consideration in the realization of a coherent CI: the role of those charged with directing or influencing CI implementation through funding allocations. Major CI investments by federal agencies to-date have focused on the deployment of national services and (to a lesser extent) on individual researcher and campus CI systems. The potential to maximize and leverage the role and resources associated with multi-institutional collaborative communities like SURAgrid have largely been overlooked and rarely funded. A global view of CI that considers **resources**, **users** and **organizations** provides crucial guidance to effective investments in CI – Federal, State and Campus. It requires that the campus, together with its extended community, be regarded as an integral component of a national CI. Regardless of how unique each campus experience is perceived to be, the truth is that many of the same processes are duplicated between, and even within, campus environments. Supporting community efforts that build an open, operational, shared infrastructure across multiple campuses is a more effective use of funds – and ultimately more sustainable than each campus "going it alone."

The same level of support that could in the past affect change for only a handful of users could now impact a much larger community who hold in common their investment in a coherent infrastructure. From the campus perspective, sharing the burden of infrastructure development across a larger pool of stakeholders has clear financial benefits. Each campus can't be expected to individually meet competing demands of their members to connect with distinct national infrastructures or resources and services available through other campuses. Establishing connections through inter-campus communities enables shared expense in both delivery and support, and is more practical than expending limited funds to connect select groups to specialized resources.

In summary, we offer the following recommendations to achieve a more pervasive and inclusive national Cyberinfrastructure:

- **Encourage Community Models:** Re-enforce the proposition that communities implementing open CI infrastructures for broad availability and shared use are crucial to a coherent CI that bridges campus, regional and national resources.
- **Extend Campus CI Initiatives:** Recognize that engagement in multi-institutional communities with shared goals will accelerate campus CI deployments while contributing to the development of a broad, integrated national CI.
- **Reward Economies of Scale:** Invest in CI programs and communities that realize cost-savings, effectively coordinate talent, create sustained collaboration, and reduce barriers to new and non-traditional users.
- **Invest in Campus-to-Campus CI Collaborations:** Allocate funds to create and extend multi-institutional CI collaborations and communities. Within the National Science Foundation, the Office of CyberInfrastructure appears particularly well suited to sponsor such programs.

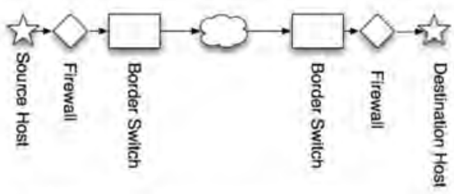
Extracted from full paper available at: http://www.sura.org/programs/docs/SGWPDraftforComment0909.pdf

Appendix 2. Workshop presentations

The following presentations are released under the Creative Commons Attribution 3.0 Unported license (http:creativecommons.org/licenses/by/3.0/) by the author(s) listed on the first slide of each presentation. This license includes the following terms: You are free to share – to copy, distribute and transmit the work and to remix – to adapt the work under the following conditions: attribution – you must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work). For any reuse or distribution, you must make clear to others the license terms of this work.

Current Realities

- Both Backbone and RON:
  - DWDM Fiber-optic plants
  - Hot and cold running 10 Gb/s
- Campus Access LANs
  - 10 GigE backbones, with 10/100baseT, gigE, and 10-GigE connections
  - Traffic Disruption Devices
  - Huge number of devices to manage



Technical: Problem is the firewall (and other PRDs)
Non-technical: Campus LAN optimized for 'safety'

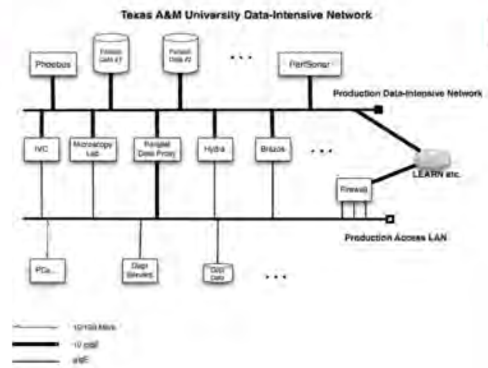
nb: This was done for Earth Day 1971, but maybe it applies also to our Internet2 "environment".

Clarity of Roles

- Campus Access LAN (a.k.a. Enterprise LAN) optimized for ordinary Internet use by many thousands of students, faculty, and staff
- We need a (possibly) separate LAN optimized for supporting the campus's Cyberinfrastructure responsibilities

Data Intensive Network

- Separate this network from the campus access LAN
- Optimize for high performance (both local and wide-area)
- Emphasize host security for hosts placed on this DIN
- Connect key campus Cyberinfrastructure resources to both DIN and the LAN



Need for Common Action

- Internet2 has about 200 universities
- Positive Statement:
  - DIN Cost grows as N
  - DIN Benefit grows as N^2 (Metcalfe's Law)
- Negative Statement:
  - We (collectively) investment millions in multi-10-Gb/s wide-area infrastructure
  - Then we kill end-to-end performance with bad engineering

Cyberinfrastructure Framework for 21st Century Science & Engineering (CF21)

NSF-wide Cyberinfrastructure Vision

People, Sustainability, Innovation, Integration

Edward Seidel/Alan Blatecky

Acting Assistant Director, Mathematical and Physical Sciences

Acting Deputy Director, Office of Cyberinfrastructure

1

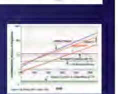
Framing the Question

Science is Radically Revolutionized by CI

- Modern science
  - Data- and compute-intensive
  - Integrative
- Multiscale Collaborations for Complexity
  - Individuals, groups, teams, communities
- Must Transition NSF CI approach to support
  - Integrative, multiscale
  - 4 centuries of constancy, 4 decades 10^{9-12} change!

2

Five Crises



- Computing Technology
  - Multicore: processor is new transistor
  - Programming model, fault tolerance, etc
  - New models: clouds, grids, GPUs,... where appropriate
- Data, provenance, and viz
  - Generating more data than in all of human history: preserve, mine, share?
  - How do we create "data scientists"?
- Software
  - Complex applications on coupled compute-data-networked environments, tools needed
  - Modern apps: 10^6+ lines, many groups contribute, take decades

3

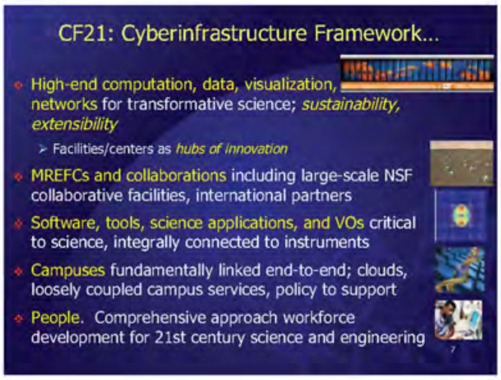
Five Crises con't

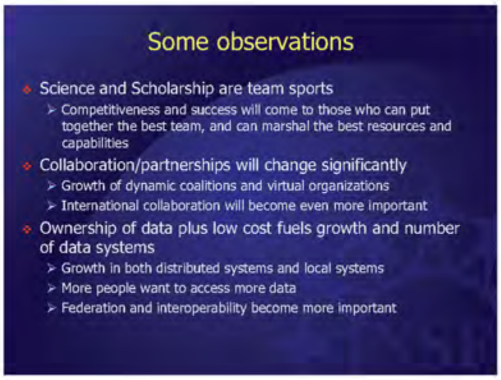
- Organization for Multidisciplinary Computational Science
  - *"Universities must significantly change organizational structures: multidisciplinary & collaborative research are needed [for US] to remain competitive in global science"*
  - *"Itself a discipline, computational science advances all science...inadequate/outmoded structures within Federal government and the academy do not effectively support this critical multidisciplinary field"*
- Education
  - The CI environment is running away from us!
  - How do we develop a workforce to work effectively in this world?
  - How do we help universities transition?

4


What is Needed?
An ecosystem, not components...
NSF-wide CI Framework for 21st Century Science & Engineering
People, Sustainability, Innovation, Integration
5


Cyberinfrastructure Ecosystem
Organizations
Expertise
Scientific Instruments
Data
Computational Resources
Networking
Software
Maintainability, sustainability, and extensibility


CF21: Cyberinfrastructure Framework...
High-end computation, data, visualization, networks for transformative science; sustainability, extensibility
Facilities/centers as hubs of innovation
MREFCs and collaborations including large-scale NSF collaborative facilities, international partners
Software, tools, science applications, and VOs critical to science, integrally connected to instruments
Campuses fundamentally linked end-to-end; clouds, loosely coupled campus services, policy to support
People. Comprehensive approach workforce development for 21st century science and engineering
7


Some observations
Science and Scholarship are team sports
Competitiveness and success will come to those who can put together the best team, and can marshal the best resources and capabilities
Collaboration/partnerships will change significantly
Growth of dynamic coalitions and virtual organizations
International collaboration will become even more important
Ownership of data plus low cost fuels growth and number of data systems
Growth in both distributed systems and local systems
More people want to access more data
Federation and interoperability become more important

More observations

- More discoveries will arise from search approaches
  - Mining vast amounts of new and disparate data
  - Collaboration and sharing of information
- Mobility and personal control will continue to drive innovation and business
- Gaming, virtual worlds, social networks will continue to transform the way we do science, research, education and business
- The Internet has collapsed six degrees of separation and is creating a world with two or three degrees.

Campus Bridging/Networking

- A goal of Virtual Proximity – as though you are one with your resources (including people)
  - Continue to collapse the barrier of distance and remove geographic location (including campus location) as an issue
  - All resources are virtually present, accessible, secure
- Leverages, informs and depends upon the whole suite of CI elements
  - HPC, Vis, Data, Software, Expertise, VOs, etc
  - Provides end-to-end connectivity
- Deployment of leading edge networking infrastructure and cybersecurity to support CF21

Campus Bridging/Networking Challenges

- Neither "campus bridging" nor "networking" accurately captures the need or concept
  - Campus bridging is vague
  - Networking is often thought of as "plumbing"

End-to-end Integrated Cyberinfrastructure

- Foundational substrate
- Involves entire protocol stack through application
- Involves user interacting with CI capabilities
  - Data, software, visualization, HPC, clouds, organizations, etc
- Throughput and usefulness is the metric

Driving Forces

- Need to support the efficient pursuit of S&E
  - Multi-domain, multi-disciplinary, multi-location
  - Leading edge CI network capabilities
  - Seamless integration
- Need to connect Researcher to Resource
  - Access to major scientific resources and instruments
  - CI resource availability – at speed and in real-time
    - (HPC, MREFC, Data Center, Vis center, Clouds, etc)
  - Campus environment including intra-campus
  - State, regional, national and international network and infrastructure transparency

12

Networking Infrastructure Issues

- Major Scientific Facility Interconnects
  - Networking infrastructure focus
- High Performance End-User Access
  - Address at-speed connection at desktop
  - Usefulness and User throughput
  - Pilot and prototype approach
- Experimental Research Networks
  - Multi layer, hybrid networks including cybersecurity
  - Apps with end-to-end focus
- Digital Divide issues
  - Geographically remote, rural areas, community colleges, etc
  - On campus, off campus

The Shift Towards Data
Implications

- All science is becoming data-dominated
  - Experiment, computation, theory
- Totally new methodologies
  - Algorithms, mathematics
  - All disciplines from science and engineering to arts and humanities
- End-to-end networking becomes critical part of CI ecosystem
  - Campuses, please note!
- How do we train "data-intensive" scientists?
- Data policy becomes critical!

Critical Factors

- Science and society profoundly changing
- Comprehensive approach to CI needed to address complex problems of 21st century
  - All elements must be addressed, not just a few
  - Many exponentials: data, compute, collaborate
- Data-intensive science increasingly dominant
  - Modern data-driven CI presents numerous crises, opportunities
- Academia and Agencies must addressed
  - New organizational structures, rebalanced investments, educational programs, policy
- End-to-end; researcher to resources

15

CF21 Plan

- Existing Task Forces
- CICC: need to recast this as CF21 WG
  - Establish CI lead in each Directorate
  - Creation of the CF21 document is the goal
- CF21 Colloquium (C^2)
- FY 2012
  - Need to have a budget building exercise for CF21
  - NSF-wide, OCI catalyzed
  - OSTP offers to help

16

Steve Corbató
University of Utah

NSF OCI CF21 Campus Bridging Workshop, Indianapolis
07 April 2010

Integration of Advanced Networks as key component of Campus and Regional CI

THE UNIVERSITY OF UTAH

emulab

First Law of (CI) Real Estate: Location, location, location

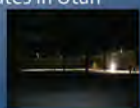
- New, greener data centers
- Facilities-based optical networks
  - Conduit & RoWs
- Utah specific factors
  - Major national fiber junction (5-way)
  - Research into network reinvention
    - GENI under NSF CISE
    - Emulab/protoGENI at Utah

2

New off-campus data center

- 74,000+ sq ft² former industrial building south of downtown SLC (~4 miles off-campus)
- Designing for enterprise, HPC/CI (1+ MW), and co-location space for research groups & external partners
- Low industrial electric power rates in Utah





3

Metro & state optical nets

But it's also virtual now…

- Converging themes
  - Data center virtualization (enterprise IT)
  - GENI emulation environments (network research)
  - Virtualized HPC & storage services (computational science)
    - Amazon S3 storage
    - NSF/Microsoft Azure program
- Cloud computing in its infancy
  - Bursty events – $\delta(t_{surge})$ in the enterprise
  - Obstacles at campuses with medical schools
    - Data privacy concerns (e.g., HIPAA, FERPA)

5

Making the net WORK for CI

- Performance montoring
  - perfSONAR, ESnet Faster Data (move 1 TB in 8 hrs)
  - Performance federation (←NSF leadership)
    - Routine, proactive testing across all paths
    - On-demand, authenticated high-capacity flow testing
  - Campus switch/router RFP input
- Dedicated, dynamic capabilities
  - An example: Internet2 ION service
  - Well suited for early trials of cloud computing applications at outset
  - Extension across RON and campus net a challenge
    - GENI OpenFlow for building virtual overlay?

6

Utah… a great place for field science

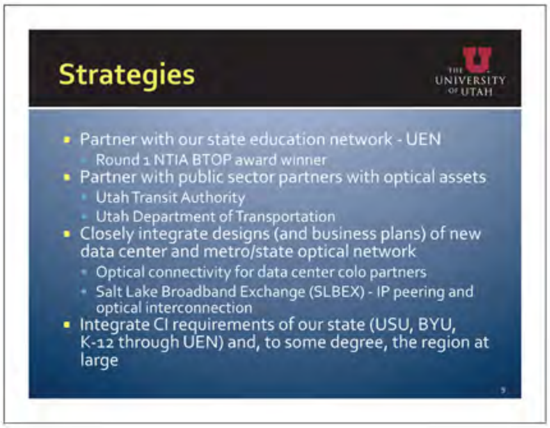
Frisco Peak Observatory
Milford, Utah

Ultra High Energy Cosmic Ray Observatory
Delta, Utah

7

Utah Field Station Network

Rio Mesa Center, Dolores River (Eastern Utah)


Strategies
THE UNIVERSITY OF UTAH
Partner with our state education network - UEN
Round 1 NTIA BTOP award winner
Partner with public sector partners with optical assets
Utah Transit Authority
Utah Department of Transportation
Closely integrate designs (and business plans) of new data center and metro/state optical network
Optical connectivity for data center colo partners
Salt Lake Broadband Exchange (SLBEX) - IP peering and optical interconnection
Integrate CI requirements of our state (USU, BYU, K-12 through UEN) and, to some degree, the region at large

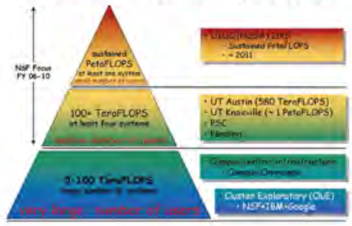

Call for NSF leadership!
THE UNIVERSITY OF UTAH
Imprimatur is extremely important
Workshops
Leveraging campus investment – O(10x)
UCAN example
Emulate success of NSF Connections (NSFNET) and HPNC (vBNS+Abilene) programs
Cultivate meritorious applications
Incent sustainable collaboration between CI-intensive researchers and IT groups
Address unintended consequences of indirect cost accounting
Will preference for equipment capitalization discourage adoption of cloud computing service models?

Enabling and Sustaining Campus-to-Campus Cyberinfrastructure

Gary Crane
Southeastern Universities Research Association
gcrane@sura.org
April 7, 2010

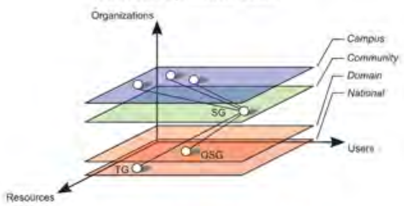
Full white paper available at: http://sura.org/programs/docs/CI_White_Paper_Final.pdf

Branscomb Pyramid



Illustrates **resource** and **user** dimensions and highlights their relative scale. Places focus on bridging gap between national resources and researchers.

Introducing **organizations** (campuses) as a third dimension leads to the depiction of a global CI environment that suggests the importance of bridging the gap between campus-to-campus CI deployments.



TG=TeraGrid OSG=Open Science Grid, SG=SURAgrid

Developing Campus CI that Scales

- Campus CI investments represent a very significant component of the nation's cyberinfrastructure.
- Open and shared campus CI deployment models are crucial to development of a national CI that bridges campus, regional, national and international boundaries.
- Lowering the barriers to deployment of campus-to-campus CI could significant improve the nation's cyberinfrastructure.

SOUTHEASTERN UNIVERSITIES RESEARCH ASSOCIATION SURA

What's Needed?

- **Encourage Community Models:** Re-enforce the proposition that communities implementing open CI infrastructures for broad availability and shared use are crucial to a coherent CI that bridges campus, regional and national resources.
- **Extend Campus CI Initiatives:** Recognize that engagement in multi-institutional communities with shared goals will accelerate campus CI deployments while contributing to the development of a broad, integrated national CI.
- **Reward Economies of Scale:** Invest in CI programs and communities that realize cost-savings, effectively coordinate talent, create sustained collaboration, and reduce barriers to new and non-traditional users.
- **Invest in Campus-to-Campus CI Collaborations:** Allocate funds to create and extend multi-institutional CI collaborations and communities. Within the National Science Foundation, the Office of CyberInfrastructure appears particularly well suited to sponsor such programs.

www.SURA.org

SOUTHEASTERN UNIVERSITIES RESEARCH ASSOCIATION SURA

Acknowledgments

SURAgrid Collaborators

- John-Paul Robinson, UAB
- Phil Smith, Texas Tech University
- Mary Fran Yafchak

SURAgrid Governance Committee

Full white paper available at: http://sura.org/programs/docs/CI_White_Paper_Final.pdf

www.SURA.org

Realizing a National Cyberinfrastructure via Federation

Andrew Grimshaw
University of Virginia

Five minutes – no agenda

- Science is increasingly collaborative and multi-organizational
- Campus-bridging is a critical component as most (NSF-relevant) researchers are physically located on campuses
- Networking is necessary, but not sufficient
- It's the software stup#d

My observations

- Most campus researchers are not very computer savvy (i.e., Unix savvy) as we think of it
  - Diverse user community beyond the usual suspects of physics, astronomy, etc.
  - Economics, English, stats, ..
  - They tend towards HTC, low degree parallel, and collaborative data sharing
  - They don't want to become guru's
  - -→ Must exploit what they know

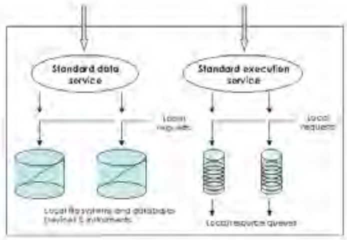
Requirements

- Simplicity
- Performance
- Security
- Standards

Solution: File based, Federated, Standards-Based Access

- Standards for risk reduction
- Map diverse resource types to files and directories – think /proc file system, Plan 9
- File system API for human uptake
- FUSE/IFS drivers to map resources into the local file system

Autonomy & Remote Access



Why is Advanced Cyberinfrastructure Not More Widely Used?

Russ Hobby <rdhobby@hobbyfamily.org>
Campus Bridging Technologies Workshop
7 April 2010

Researchers and CI

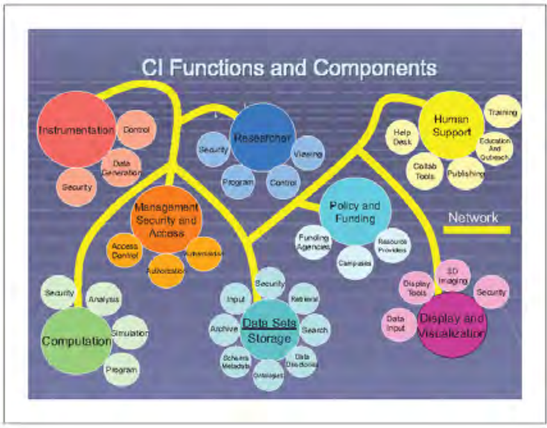
- With research being done in distributed discipline groups, CI is needed to allow the group to work together
- Some researchers, CI Champions, have integrated advanced CI into their research, but have also had to become CI experts
- Other researchers, CI Consumers, do not want to become experts and are waiting for help

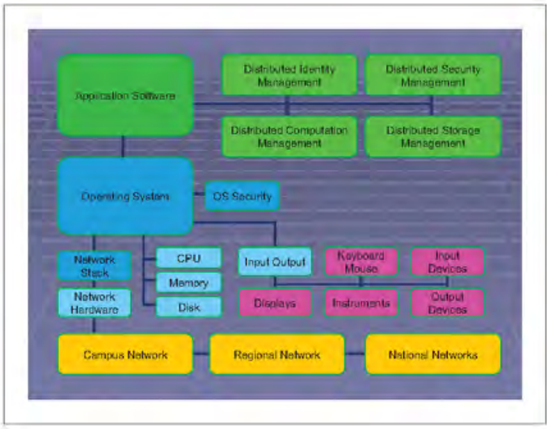
Why is CI not broadly used?

- Difficulty in use
  - A new user interface has to be learned for each new tool.
  - Combining tools together or moving to a different tool is difficult.
- Problems with reliability and performance
  - CI resources often do not operate optimally the first time. Expert assistance is required.
  - Even after optimum performance is achieved, a change in one component can require retuning.

Possible Help?

- Ease of use
  - A common, flexible interface style could serve as a base for tool development
  - Standard workflow mechanisms for combining tools
- Reliability and Performance
  - CI is a complex system with many dependencies. A better understanding is needed of the whole system
  - A CI architecture would help developers create a better coordinated CI system.


CI Functions and Components
Instrumentation
Control
Data Generation
Security
Researcher
Security
Viewing
Program
Control
Human Support
Training
Help Desk
Education And Outreach
Collab Tools
Publishing
Management Security and Access
Access Control
Authorization
Policy and Funding
Funding Agencies
Resource Providers
Campuses
Network
Security
Analysis
Computation
Simulation
Program
Security
Input
Retrieval
Archive
Data Sets Storage
Search
Schema Metadata
Data Directories
3D Imaging
Display Tools
Security
Data Input
Display and Visualization


Application Software
Distributed Identity Management
Distributed Security Management
Distributed Computation Management
Distributed Storage Management
Operating System
OS Security
Network Stack
Network Hardware
CPU
Memory
Disk
Input Output
Keyboard Mouse
Input Devices
Displays
Instruments
Output Devices
Campus Network
Regional Network
National Networks

EXTENDING CYBERINFRASTRUCTURE BEYOND ITS OWN BOUNDARIES

Campus Champion Program

Kay Hunt
kay@purdue.edu

HOW DO WE MAKE THESE RESOURCES AVAILABLE??

The Campus Champion program is intended to identify and support a campus representative who will become the local source of knowledge about high performance computing opportunities and resources. This knowledge and assistance will empower researchers and educators to advance scientific discovery.

Campus Champions Program

- Source of local, regional and national high performance computing and cyberinfrastructure information at home campus
- Source of information about resources and services that will benefit their campus
- Source of startup accounts to quickly get researchers and educators using their allocation of time on the resources
- Direct access to support staff
- http://www.teragrid.org/eot/campuschamps.html

History of the Champions

- Planning began in Fall of 2007
- Advisory Board Formed
- 1st Champion selected in May 2008
- March 2010 – 62 institutions joined

Where are these Champions????

TeraGrid Campus Champions

March 10, 2010

What is expected of each campus?

- Provide information on CI resources to researchers and educators
- Assist campus users to quickly get a start-up allocation
- Host awareness sessions
- Host training sessions
- Provide local users with contacts for quick problem resolution
- Attend Annual Conference

What does the TeraGrid provide?

- Regular correspondence on new resources, services, and offerings
- Participation in User Services Working Group
- Forum for sharing information
- Campus visits by personnel
- Training for Champions
- Allocations of resources
- Waiver of registration fee to annual conference

Why join the Champions?

Your campus will benefit by having direct access to the resources and input to the staff, resource allocations awarded for their use, and assistance in using those resources.

Who should be the Champion?

- Each institution is unique in its needs
- May need more than one Champion
- Champion likely to be from one of the following communities
  - CIO staff member
  - IT professional staff member
  - Researcher or faculty member with a passion to engage colleagues

Partnerships

Blue Waters Project

Open Science Grid

Internet2

Questions ??

CAMPUS CHAMPIONS

http://www.teragrid.org/eot/campuschampions

tgcc-help@teragrid.org

Development and deployment of integrated attribute based access control for collaboration

Collaborations and Virtual Organizations

- IdM is a critical dimension of collaboration, crossing many applications and user communities
- Virtual organizations represent critical communities of researchers sharing domain resources and applications as well as general collaboration tools. Providing a unified identity management platform for collaboration is essential in a multi-domain, multi-tool world.
- Lots of activities in domesticating applications to work in a federated world, moving from tool-based identity to collaboration-centric identity.

Collaboration Platform

- Integrated set of collaboration apps (wikis, listprocs, CVS, file share, calendaring, etc)
- Integration of at least identity and access control via group memberships
- Integration with domain science apps
- Integration of content and meta-data is harder
- Repackages successful approaches for a collaborative/ project/VO setting
  - Federated identity, group management, directories, and security token services (aka credential convertors)

Collaboration Infrastructure (COIN)

- Dutch National Collaboration Infrastructure
- Domesticated tools -Adobe Connect; Alfresco; Foodle; Filesender; Confluence; WSO2 mashup server; OpenFire; Drupal; KnowledgeTree, Sympa and Limesurvey
- Domesticated services -Google Apps; MyExperiment.org; Twitter; PubMed
- Integration across VO, institution and third-party domains
- Workflow
- Grid integration

Domestication of applications

- The work of re-factoring applications to use the emergent identity services infrastructure
- Begins with federated identity and authentication, use of directories; gains a lot from group management for access control, etc
- Needs a fine grain set of authorization tools down the road
- Domesticated apps can receive IdM attributes via LDAP, SAML, X.509, SQL, Kerberos PAC, and maybe all of the above

Typical activities in collaboration management

- Add or remove people from groups
- Create new subgroups, identify overlapping memberships, etc.
- Permit or deny access control to wiki pages, calendars, computing resources, version control systems, etc
- Add people to mailing lists, wikis, etc
- Create and delete/archive users, accounts, keys
- Identify group membership on a given date

COManage Elements

What's in a COmanage data store

Enterprise Attributes	Project/VO attributes
Federated Id	PI groups
Enrolled classes	Wiki editing permissions
Display name	Instrument permissions
Citizenship	VO certificates
Enterprise affiliation	...

Grouper

- A general purpose, extensible, open-source group management tool
- In production at many institutions in the US and overseas
- Core national infrastructure service in several countries
- Manages groups of things – people, devices, processes
- Has GUI, people picker, group math, inheritance, delegation, provisioning and deprovisioning, etc.
- Stores values in LDAP directory
- Aimed at spectrum from power user to collabmin, sysadmin and enterprise IdM

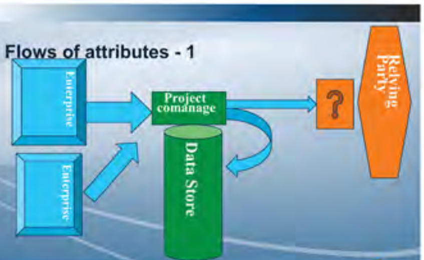
Security Token Service

- Converts the form of an existing credential or packs a set of attributes into a new credential
- Presents external security information to an application or service in the lingua of the app/service
- Conversions – SAML into X.509, SAML into Kerberos, SAML to LDAP, etc.
- Mythical in a single comprehensive package; legion in individual instances

What forms does COmanage take?

- Usually as an assembled set of services
  - A dashboard, directory product, Shibboleth IdP and SP, Grouper, and a set of applications provisioned on other servers
  - On an enterprise level to serve its collaborations and VO's, within a large VO, or at a federation level to serve a national community
- Can also be a VM, a VM in the cloud, or a service with the applications in the cloud.
- Can be embedded in a science portal or gateway

Flows of attributes - 1




Use cases it enables
• A student adds a class and is immediately enabled to use the VO wiki; a student drops the class and is immediately disabled from using the VO instruments
• A resource prohibited from use by foreign nationals is protected
• International privacy laws are adhered to
• Anonymous access is enabled but limited to those authorized to participate
• Security is commensurate with the risks

Regional Cyberinfrastructure as a Bridge Between Campus and National CI*

Greg Monaco, Great Plains Network
Rick McMullen, University of Kansas

*A brief overview based on our carefully prepared position paper

cyberinfrastructure divide

- between the computing environments for researchers
  - on their home campuses,
  - at HPC centers of other universities where they may collaborate and
  - at national CI centers such as the TeraGrid, the OSG and national data repositories

I. What Is Regional CI?

- successfully ***coordinate*** efforts across multiple institutions with shared interests,
- foster a ***collaborative context*** in which researchers can address problems of regional interest and importance, and
- bridge the ***economic component*** of the *cyberinfrastructure divide* between users at resource poor institutions and regional and national resources
- Build on existing regional orgs (e.g., RONs)

Goals

- Development of CI at the campus level in a regionally coherent way
- Outreach to four year and two year institutions and to industry,
- Aggregate regional CI resources through a common set of policies and procedures,
- seamless path for researchers to scale from campus to regional HPC and national capability computing centers,
- Build knowledge management structures and training programs that will allow the efficient sharing of CI and computational science expertise across the region, and
- Support integration of CI into educational programs through the cooperative development and delivery of computational science curricula.

How Regional CI Bridges Campus and National HPC Centers?

- develop and scale-up of projects from campus resources
- foster collaborations between individual researchers working on large scale problems
- support differentiation strategies among campuses
- Develop and share expertise across campuses in a sustainable way

Campus Grids & the OSG

April 7th 2010

Rob Gardner, University of Chicago, Integration & Sites Coordinator
Ruth Pordes, Fermilab, Executive Director

Campus Grids have been part of the OSG accessible set of resources for years, with ~single new entrants every year. They cross-faculty within an organization and are used by multiple local and remote science communities.

US ATLAS, US CMS, LIGO have/use multiple Campus single-faculty clusters and storage which form collective "grid of campuses" - Community Campus Grids spanning (widely) geographically separated sites. Currently ~20 sites for each of the LHC experiments, with ~40 more campuses coming on board in 2010 and an equivalent number over the next ~2 years.

National Laboratories and Universities act together in both spaces.

- The technologies used vary: different batch and storage systems. The interfaces are common for the submission of jobs, transfer of data, information publishing, Grid-wide services.
- Different models depend on local conditions and needs
  - GLOW,
  - FermiGrid,
  - Purdue,
  - Clemson,
  - Nebraska
  - UNC/RENCI
- Effectiveness extends out from the existing department (controlled) clusters and mode of working.
- An exercise in Change Management
  - Sociological changes must be stepwise and not rejected.
  - Planning is an ongoing and significant activity.

- Back in ~2005 for Tier-2s: Evaluation criteria for "Science Community allocations for campus resources" included points for collaboration with local IT Facility, inclusion and sharing.
- For the Tier-3s, which are separately funded, the intent is to use them as entry-points to the campus from which to osmose to other faculty "one by one"


Emerging Themes
Systems management tools & services (cross-campus)
– Configuration management and monitoring
– Policy compliance monitoring / audit
Resource sharing/trading
– Through "human communication" at least
– Incentives for sharing & trading
Usability
– "Trust management
– Available cycles and storage for users, with understood expectations
Global file systems
Common user environments to reach diverse types of capability/capacity
Interest in evaluating virtual machines and clouds

Introduction

RENCI Data Working Group

- All CI technological areas represented
- Integrative approach
- Applicable to campus bridging

renci

Data Challenges

Discovery
Integration
Policy
Curation
Granularity
Management
Data Placement
Privacy and Security
Scalability
Fault Tolerance
Mobile
Provenance

renci

Technology Solutions

Reusable, open technology systems

- iRODS
- Sensor Data Bus
- DataLab
- etc.

renci

Science of Cyberinfrastructure

- Holistic integration of data, computing, analysis
- Architecture
  - Process
  - Patterns
  - Case studies
  - Integrated CI stack for data
- Community engagement and involvement

renci

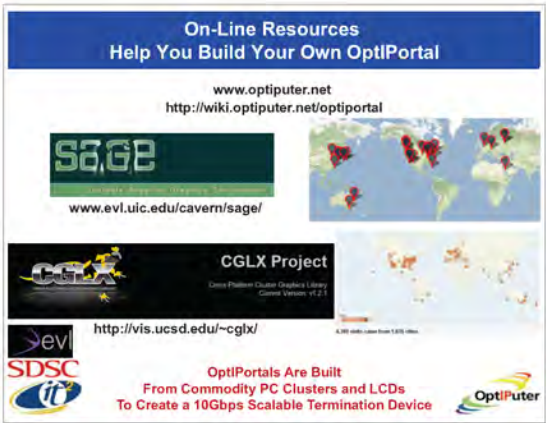
Recommendations to NSF

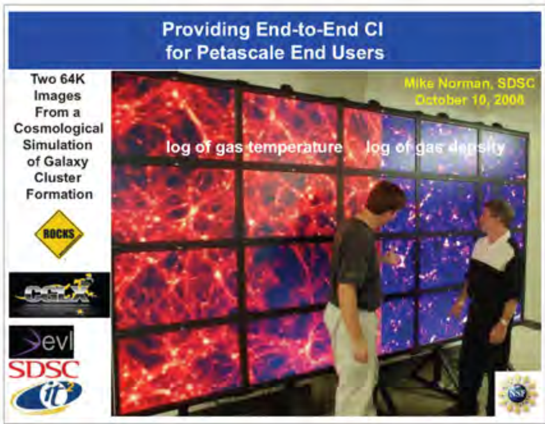
- Document and disseminate a body of practice to address complex data interoperability challenges in the STEM community
- Make available open source technologies and tools that support policy-driven, shared data collections at various levels such as campus-to-campus and campus-to-national CI
- Encourage standards development and implementations
- Provide a mechanism to make available the body of practice, standards, and technology tools to the STEM research and campus communities
- Leverage organizations with existing collaborative capabilities

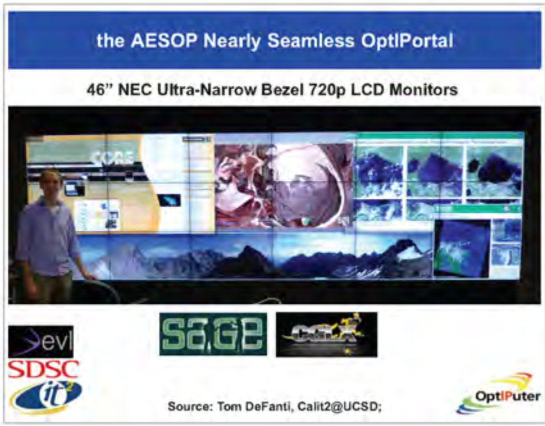
renci


The OptIPortal, a Scalable Visualization, Storage, and Computing Termination Device for High Bandwidth Campus Bridging
Presentation by Larry Smarr to the NSF Campus Bridging Workshop
April 7, 2010
University Place Conference Center
Indianapolis, IN
Calit2@UCSD: Tom DeFanti, Qian Liu, Joseph Keefe, Greg Hidley, Greg Dawe, Bryan Glogowski, Ian Kaufman, Jurgen Schulze, Kai-Uwe Doerr, Falko Kuester, Larry Smarr
SDSC@UCSD: Phil Papadopoulos, Mason Katz
EVL@UIC: Jason Leigh, Luc Renambot, Alan Verlo, Lance Long, Maxine Brown, Dan Sandin
ANL: Venkatram Vishwanath
Qualcomm, Inc.: Javier Girado
TACC@UTA: Byungil Jeong
SDSC
OptIPuter


The OptIPuter Project: Creating High Resolution Portals Over Dedicated Optical Channels to Global Science Data
Scalable Adaptive Graphics Environment (SAGE)
NCMIR
Picture Source: Mark Ellisman, David Lee, Jason Leigh
SDSC
Calit2 (UCSD, UCI), SDSC, and UIC Leads—Larry Smarr PI
Univ. Partners: NCSA, USC, SDSU, NW, TA&M, UvA, SARA, KISTI, AIST
Industry: IBM, Sun, Telcordia, Chiaro, Calient, Glimmerglass, Lucent
OptIPuter

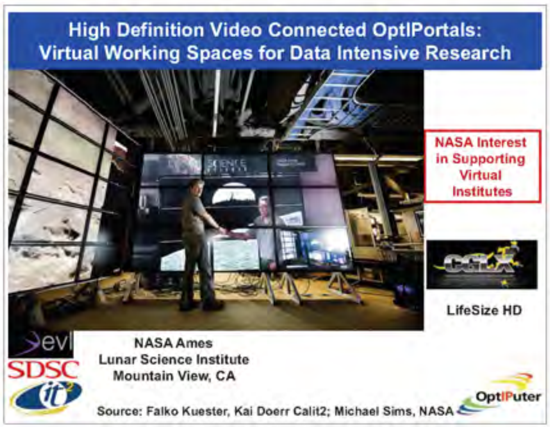

On-Line Resources Help You Build Your Own OptIPortal
www.optiputer.net
http://wiki.optiputer.net/optiportal
www.evl.uic.edu/cavern/sage/
CGLX Project
http://vis.ucsd.edu/~cglx/
SDSC
OptIPortals Are Built From Commodity PC Clusters and LCDs To Create a 10Gbps Scalable Termination Device
OptIPuter


Providing End-to-End CI for Petascale End Users
Two 64K Images From a Cosmological Simulation of Galaxy Cluster Formation
Mike Norman, SDSC
October 10, 2008
log of gas temperature
log of gas density
ROCKS
SDSC

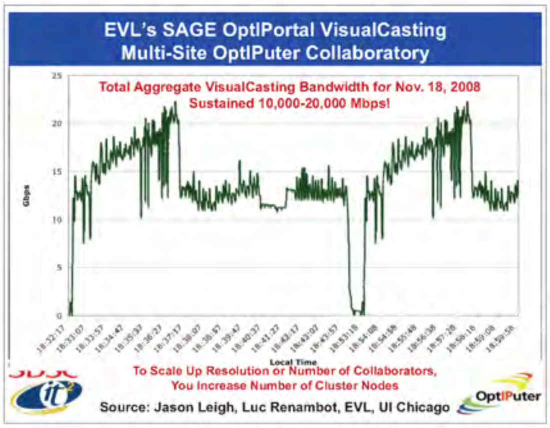


the AESOP Nearly Seamless OptIPortal
46" NEC Ultra-Narrow Bezel 720p LCD Monitors
evl
SDSC
SAGE
Source: Tom DeFanti, Calit2@UCSD;
OptIPuter

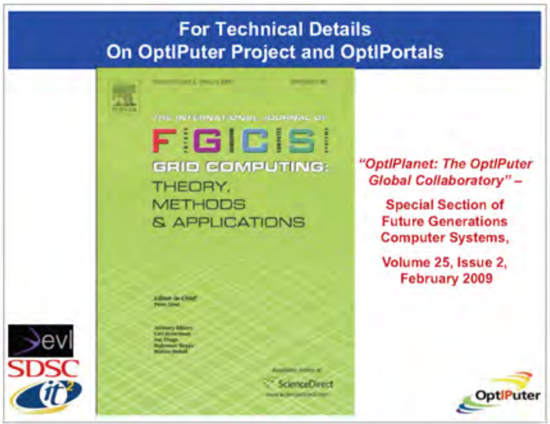
3D Stereo Head Tracked OptIPortal:
NexCAVE
evl
SDSC
Array of JVC HDTV 3D LCD Screens
KAUST NexCAVE = 22.5MPixels
www.calit2.net/newsroom/article.php?id=1584
Source: Tom DeFanti, Calit2@UCSD
OptIPuter



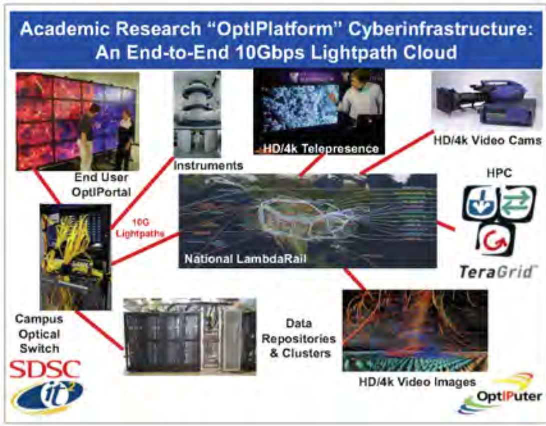
High Definition Video Connected OptIPortals:
Virtual Working Spaces for Data Intensive Research
NASA Interest in Supporting Virtual Institutes
LifeSize HD
evl
SDSC
NASA Ames
Lunar Science Institute
Mountain View, CA
Source: Falko Kuester, Kai Doerr Calit2; Michael Sims, NASA
OptIPuter

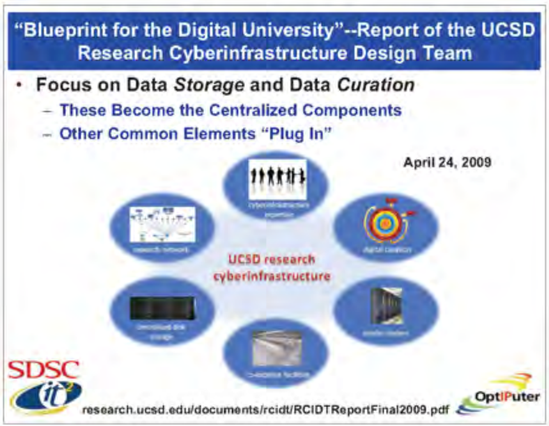


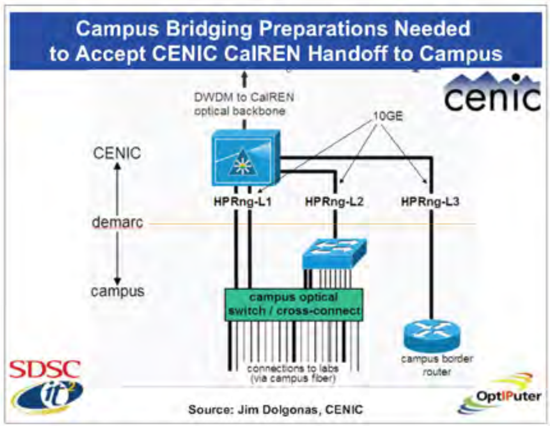
EVL's SAGE OptIPortal VisualCasting
Multi-Site OptIPuter Collaboratory
Total Aggregate VisualCasting Bandwidth for Nov. 18, 2008
Sustained 10,000-20,000 Mbps!
Gbps
Local Time
To Scale Up Resolution or Number of Collaborators,
You Increase Number of Cluster Nodes
Source: Jason Leigh, Luc Renambot, EVL, UI Chicago
OptIPuter

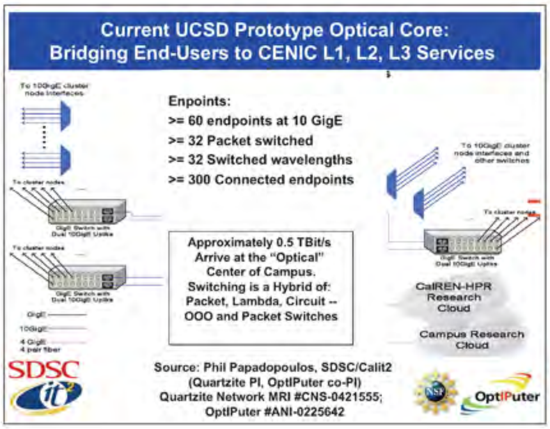

For Technical Details
On OptIPuter Project and OptIPortals
FGCS
GRID COMPUTING:
THEORY,
METHODS
& APPLICATIONS
ScienceDirect
"OptIPlanet: The OptIPuter Global Collaboratory" –
Special Section of Future Generations Computer Systems,
Volume 25, Issue 2, February 2009
evl
SDSC
OptIPuter

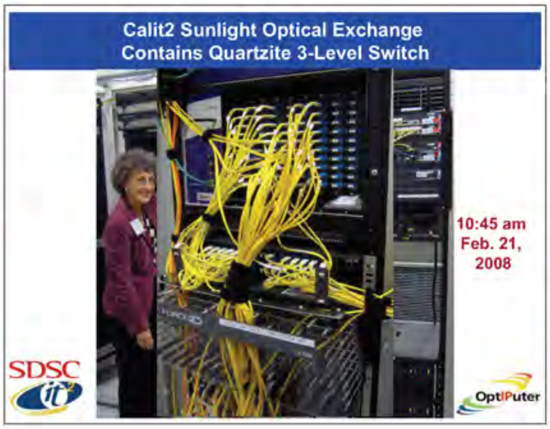

A High-Performance Campus-Scale
Cyberinfrastructure
For Effectively Bridging End-User Laboratories
to Data-Intensive Sources
Presentation by Larry Smarr to the NSF Campus Bridging Workshop
April 7, 2010
University Place Conference Center
Indianapolis, IN
Philip Papadopoulos, SDSC
Larry Smarr, Calit2
University of California, San Diego
SDSC
OptIPuter

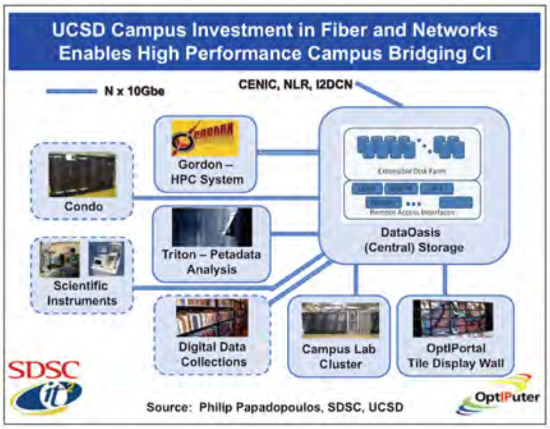

Academic Research "OptIPlatform" Cyberinfrastructure:
An End-to-End 10Gbps Lightpath Cloud
End User
OptIPortal
Instruments
HD/4k Telepresence
HD/4k Video Cams
HPC
10G
Lightpaths
National LambdaRail
TeraGrid
Campus
Optical
Switch
Data
Repositories
& Clusters
HD/4k Video Images
SDSC
OptIPuter

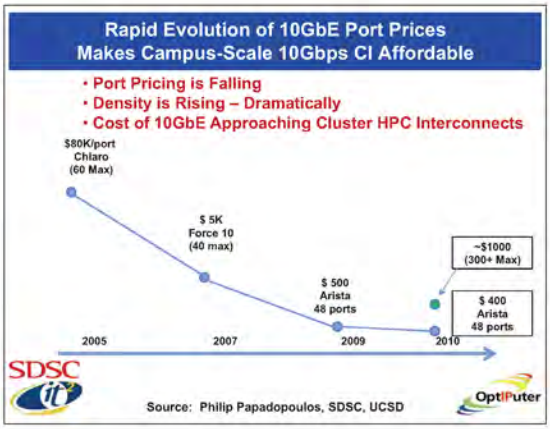

"Blueprint for the Digital University"--Report of the UCSD
Research Cyberinfrastructure Design Team
• Focus on Data Storage and Data Curation
– These Become the Centralized Components
– Other Common Elements "Plug In"
April 24, 2009
UCSD research
cyberinfrastructure
SDSC
research.ucsd.edu/documents/rcidt/RCIDTReportFinal2009.pdf
OptIPuter


Campus Bridging Preparations Needed
to Accept CENIC CalREN Handoff to Campus
cenic
DWDM to CalREN
optical backbone
10GE
CENIC
HPRng-L1
HPRng-L2
HPRng-L3
demarc
campus
campus optical
switch / cross-connect
connections to labs
(via campus fiber)
campus border
router
SDSC
Source: Jim Dolgonas, CENIC
OptIPuter


Current UCSD Prototype Optical Core:
Bridging End-Users to CENIC L1, L2, L3 Services
Enpoints:
>= 60 endpoints at 10 GigE
>= 32 Packet switched
>= 32 Switched wavelengths
>= 300 Connected endpoints
Approximately 0.5 TBit/s Arrive at the "Optical" Center of Campus.
Switching is a Hybrid of: Packet, Lambda, Circuit -- OOO and Packet Switches
CalREN-HPR Research Cloud
Campus Research Cloud
Source: Phil Papadopoulos, SDSC/Calit2 (Quartzite PI, OptIPuter co-PI)
Quartzite Network MRI #CNS-0421555; OptIPuter #ANI-0225642
SDSC
OptIPuter


Calit2 Sunlight Optical Exchange
Contains Quartzite 3-Level Switch
10:45 am
Feb. 21,
2008
SDSC
OptIPuter


UCSD Campus Investment in Fiber and Networks
Enables High Performance Campus Bridging CI
N x 10Gbe
CENIC, NLR, I2DCN
Condo
Gordon – HPC System
DataOasis (Central) Storage
Scientific Instruments
Triton – Petadata Analysis
Digital Data Collections
Campus Lab Cluster
OptIPortal Tile Display Wall
Source: Philip Papadopoulos, SDSC, UCSD
SDSC
OptIPuter


Rapid Evolution of 10GbE Port Prices
Makes Campus-Scale 10Gbps CI Affordable
• Port Pricing is Falling
• Density is Rising – Dramatically
• Cost of 10GbE Approaching Cluster HPC Interconnects
$80K/port Chiaro (60 Max)
$ 5K Force 10 (40 max)
$ 500 Arista 48 ports
~$1000 (300+ Max)
$ 400 Arista 48 ports
2005
2007
2009
2010
Source: Philip Papadopoulos, SDSC, UCSD
SDSC
OptIPuter

Improving Network Data Logistics with REDDnet

Network Logistics and Buffering

- Logistics is *time-related positioning of resources*
- Buffers are *temporary storage areas used when transferring data*
  - Storage has a continuum of persistence and durability
- Network devices have some buffer capacity
- High-performance I/O depends on buffering at various levels

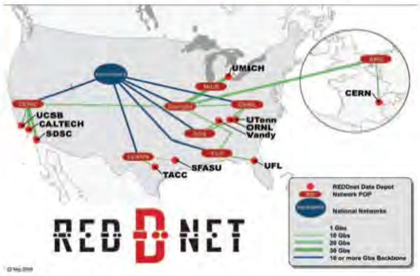
REDDnet: Infrastructure for Distributed, Data Intensive Collaboration

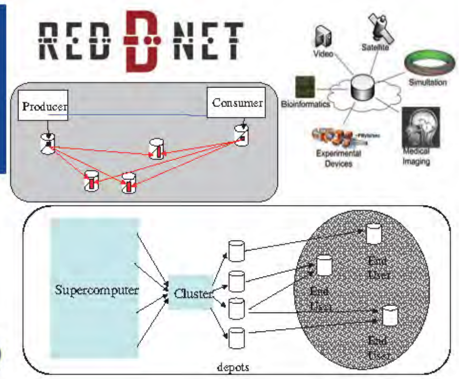
- REDDnet: Research and Education Data Depot Network
- NSF funded, additional support from Library of Congress
- "Working storage" to help manage the logistics of sharing, moving and staging large datasets across wide areas and distributed collaborations.
- Institutions: Vanderbilt, Tennessee, Stephen F. Austin, NC State, ORNL, Nevoa Networks, U. Delaware
- Host Sites: Caltech, Florida, Michigan, SDSC, TACC, UC Santa Barbara (Stephen F. Austin, Tennessee, ORNL, Vanderbilt)
- Based on the Internet Backplane Protocol (IBP) funded by the DOE and NSF

REDDnet: Infrastructure for Distributed, Data Intensive Collaboration

- REDDnet: Research and Education Data Depot Network
- NSF funded, additional support from Library of Congress
- "Working storage" to help manage the logistics of sharing, moving and staging large datasets across wide areas and distributed collaborations.
- Institutions: Vanderbilt, Tennessee, Stephen F. Austin, NC State, ORNL, Nevoa Networks, U. Delaware
- Host Sites: Caltech, Florida, Michigan, SDSC, TACC, UC Santa Barbara (Stephen F. Austin, Tennessee, ORNL, Vanderbilt)
- Based on the Internet Backplane Protocol (IBP) funded by the DOE and NSF

REDDnet Depot Map





Network Cyberinfrastructure

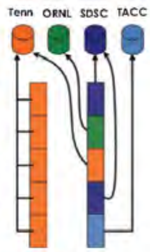
- A distributed data logistics system has broad applicability across scientific domains
  - The early adopters are those that understand their CI needs
- "We can't use the network – we just ship DVDs"
- We must bridge the gap to bring CI to the "rest of the world"
- Generic working storage can help campuses make better use of network links (campus admins should like it)
- It can also improve performance (so users like it too!)

End

- Thank you for your attention
- Questions?

Technology: The IBP Depot

- A Depot can be one storage device or a collection located in one physical location
- When you upload a file, it is first split into "slices" of a fixed size that you can specify.
- Slices can all be put on one depot or can be spread out across several/all (can be user specified or policy driven)
- Slices are stored with an expiration date set by you or the depot (depot sets max)
- An eXnode (think inode) stores file metadata



The Data Logistics Toolkit

- Our current integration effort builds on REDDnet/IBP
- We will integrate perfSONAR and Phoebus for network information and optimization
  - perfSONAR Information Service - UNIS
- perfSONAR will provide network performance information, which is vital for debugging
- More importantly, network topology can be used to inform data distribution

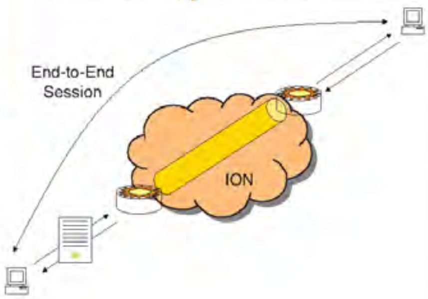
Logistical Networking

- *Logistical Networking improves data movement performance and functionality via buffering across a range of parameters and layers*
- Internet Backplane Protocol / REDDnet
  - More persistent, significant capacity
  - Buffer duration from minutes to months
  - Implements "Working Storage" used in general distributed applications
- Logistical Session Layer / Phoebus
  - Highly transitory, moderate capacity
  - Buffer duration from milliseconds to minutes
  - Implements "Store and Forward" along the end-to-end path for improved throughput

Phoebus

- Phoebus is a network optimization "inlay"
- Based on the eXtensible Session Protocol – XSP
- Protocol tuning and translation
- Transparent dynamic network resource allocation
  - Phoebus can serve as an on-ramp to ION
- Recent work with the Linux implementation performs well over 10GE

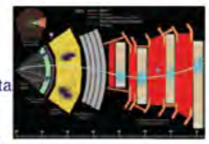
Phoebus



Working Storage and The Data Pulse Model

- REDDnet is designed to offer "working storage," i.e., a place where data can persist while you're working on it, but not indefinitely
- Often collaborations are strongly interested in a new data set – or data "pulse" - for a relatively brief period, shifting to a new set afterward
- Data from the LHC fits this model, among others
- Want data "pulse" to be widely available for period of interest
- Working storage with a "temporal" element – both space and time are important
- REDDnet is designed for data logistics

Data Pulse for CMS Analysis



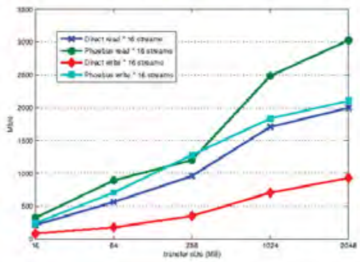
Physicists will:

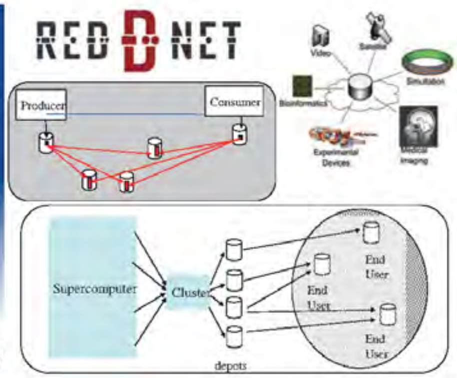
- need access to 10-100 TB Data **Sets** for short term periods.
- run over this data many times, refining, improving their analysis.
- use local computing resources (where they may not have much storage available.)
- make "opportunistic use" of compute resources at Tier 3 sites and Grid sites.
- perform "production runs" at Tier 2 sites.

CMS Tier 3 Analysis

- CMS currently has a "data-tethered" analysis model. A copy of the data must be local to the computing resources
- Using REDDnet to break the tether, transparent to users
- Upload data using standard CMS data movement tools (PhEDEx and gridFTP)
- Special gridFTP backend that uploads into REDDnet
- Data then replicated to depots to make sure copies are near users and CPU resources they can use
  - local, OSG, ...
- Plugin for CMS Software – reads directly from REDDnet

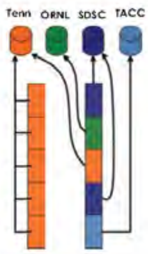
REDDnet transfers from CERN to Vandy





Technology: The IBP Depot

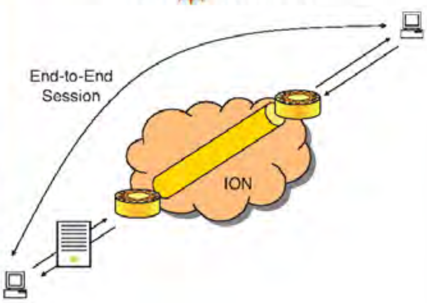
- A Depot can be one storage device or a collection located in one physical location
- When you upload a file, it is first split into "slices" of a fixed size that you can specify.
- Slices can all be put on one depot or can be spread out across several/all (can be user specified or policy driven)
- Slices are stored with an expiration date set by you or the depot (depot sets max)
- An eXnode (think inode) stores file metadata



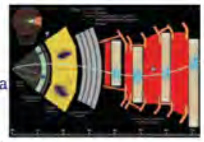
Phoebus

- Phoebus is a network optimization "inlay"
- Based on the eXtensible Session Protocol – XSP
- Protocol tuning and translation
- Transparent dynamic network resource allocation
  - Phoebus can serve as an on-ramp to ION
- Recent work with the Linux implementation performs well over 10GE

Phoebus



Data Pulse for CMS Analysis



Physicists will:

- need access to 10-100 TB Data **Sets** for short term periods.
- run over this data many times, refining, improving their analysis.
- use local computing resources (where they may not have much storage available.)
- make "opportunistic use" of compute resources at Tier 3 sites and Grid sites.
- perform "production runs" at Tier 2 sites.

The cover image is based on Joachim Bering's etching of the city of Königsberg, Prussia as of 1613 (now Kaliningrad, Russia). Seven bridges connect two islands in the Pregal River and the portions of the city on the bank. The mathematical problem of the Seven Bridges of Königsberg is to find a path through the city that crosses each bridge once and only once. Euler proved in 1736 that no solution to this problem exists or could exist. This image appears on the cover of each of the Campus Bridging Workshop reports.

The goal of campus bridging is to enable the seamlessly integrated use among: a scientist or engineer's personal cyberinfrastructure; cyberinfrastructure on the scientist's campus; cyberinfrastructure at other campuses; and cyberinfrastructure at the regional, national, and international levels; so that they all function as if they were proximate to the scientist. When working within the context of a Virtual Organization (VO), the goal of campus bridging is to make the 'virtual' aspect of the organization irrelevant (or helpful) to the work of the VO. The challenges of effective bridging of campus cyberinfrastructure are real and challenging – but not insolvable if the US open science and engineering research community works together with focus on the greater good of the US and the global community. Other materials related to campus bridging may be found at: https://pti.iu.edu/campusbridging/