UNH 2015-2020 Cyberinfrastructure Plan

Executive Summary

Inspired in part by the NSF EPSCoR CI Workshop (2007), which revealed northern New England as the "black hole of connectivity"; UNH implemented its first CI Plan in 2008. Informed by and integrated with the vision and plans of the Governor's Telecom Advisory Board, the Division of Economic Development's Director of Broadband, and the UNH Technology plan for research, education, and business; the UNH CI plan embodies the goals and strategies needed to advance technological development at UNH and across New Hampshire.

Since 2008, UNH has made significant advances in its cyberinfrastructure, including:
- ✓ 2008: USNH and the University of Maine System completed a fiber network construction project from Portland, ME to Cambridge, MA, upgrading a shared backhaul link to 10 Gbps.
- ✓ 2009: Deployed UNH Space Science Center "Fishercat" IBM/CELL supercomputer.
- ✓ 2010: UNH led the effort to construct 10Gbps fiber optic networks in all 10 NH counties, linking USNH and the Community College System of NH (CCSNH) on one network.
- ✓ 2010: Renovation of the 2,000 square foot research computing data center.
- ✓ 2011: With an Indefensible Right to Use (IRU) lease, UNH implemented a 12 fiber network from Manchester to Keene, NH, connecting Keene State College to UNH at 10 Gbps.
- ✓ 2011: UNH Director of Strategic Technology Scott Valcourt (*Proposal PI)* named an Extreme Science and Engineering Discovery Environment (XSEDE) Campus Champion for UNH.
- ✓ 2012: UNH Information Technology (IT) and Research Computing Center (RCC) staffs were trained to develop next-generation CI for database and network expansion.
- ✓ 2012: UNH IT established a parallel test IPv6 network to begin migrating the campus LAN.
- ✓ 2012: UNH IT students developed the first Shibboleth implementation to test federated authentication across UNH domains and those beyond the UNH boundary.
- ✓ 2013: Cray Supercomputer funded from NSF MRI award.
- ✓ 2013: IRU lease of 12 fibers from Manchester to Plymouth to Hanover completed, linking Plymouth State University and Dartmouth College to UNH.
- ✓ 2013: Dark fiber paths with highly scalable networks for research and education completed to connect resources and people across the Northeast.
- ✓ 2013: NSF CC-NIE award to build a Science DMZ targeting five science drivers.
- ✓ 2014: Launch of Box, a cloud computing business that provides secure, encrypted file-storage, collaboration, and other tools for working with files that are uploaded to its servers to support education, outreach and research.

As a result of rapidly achieving the 2008-2018 CI Plan goals, and in concert with the release of the UNH 2015-2020 strategic plan, the UNH CI Plan has been updated to incorporate new and evolving technologies, and to align its goals with those of the overall UNH strategic blueprint. Therefore, the UNH 2015-2020 CI Plan overarching goals for the next five years are to:

1. Offer high speed connectivity for education, outreach and research to other NH institutions, including all University System of New Hampshire (USNH) campuses (UNH, UNH at Manchester, Keene State College, Plymouth State University, and Granite State College headquartered in Concord with educational facilities throughout NH).
2. Align strategic investments in networking, compute, and storage with the collective needs of faculty and administration guided by their specific drivers.
3. Develop the personnel needed to design, build, implement and maintain the latest technologies and tools to help researchers and educators translate their needs efficiently and effectively.
4. Incorporate compute, storage and analytic services into the UNH community fabric.

5. Maintain a suite of "Data Management Services" to support each stage of the data lifecycle.
6. Design and implement a governance structure to advocate for and direct the CI Plan.

The UNH 2015-2020 CI Plan highlights the current and future states of specific technologies, followed by an execution strategy to prioritize the order in which each is advanced.

Highlighted CI Technologies

Network Capacity and Regional Connectivity
Current State: USNH, through UNH, is connected to the Internet2 ION network via two 10Gbps dense wave-division multiplexing (DWDM) waves. UNH access to ION is available via the Northern Crossroads (NoX), a peering point for all New England higher education institutions. NoX, through the US UCANN BTOP award, recently upgraded its two Internet2 connections to 100Gbps routes, providing resiliency through New York and Chicago and offering Internet2 connectivity that is shared by the nearly 40 NoX-linked institutions. With a recently awarded CC-NIE grant (#ACI-1340972) allowing a third 10Gbps wave in 2014, UNH has even greater opportunities for data movement to and from UNH devices using our DYNES instrument.

Future State (supporting CI Plan goals #1 and #2): A roadmap will chart the networking vision to keep pace with future demands. This demand may be met either by implementing a fourth 10Gbps DWDM science-dedicated wavelength from UNH to Cambridge, MA, or replacing all of the 10Gbps waves for two 100Gbps waves. Trained IT staff will engage with research faculty to assist in the use of the Science DMZ including use of the DTN, Identity and Access Management services and targeted productivity tools such as Globus.

Compute Capacity
Current State: Existing compute resources include a Cray Supercomputer, several multi-node Linux-based clusters and dozens of decentralized transaction-based web, database and application servers. HPC resources serve primarily sponsored research projects and are in constant use. A significant HPC demand exists among early career investigators and for their projects that have yet to receive funding.

Future State (supporting CI Plan goals #2 and #4): UNH will provision and sustain a shared HPC Cluster available as a free resource for all campus researchers. Removing the cost of entry into HPC will reduce the proliferation of single-purpose servers that are less powerful, efficient and secure. Transactional systems will migrate to a more efficient and effective virtual server environment, offering virtual systems for both the researchers and faculty seeking to replace their legacy stand-alone systems with virtual systems that are able to flex their compute power as required. Cloud-based HPC will be leveraged for applicable use cases augmenting campus capacity, allowing HPC solutions that address small, medium and large-sized projects in the most effective and efficient manner.

Storage Capacity
Current State: UNH IT provides centralized storage for faculty, staff, and students (50GB/user quota) through "Box," a commercial cloud-based storage provider. Research data storage is the responsibility of each investigator and is generally achieved through the purchase of direct-attached storage that is increased as needed. A data transfer node provides a 30TB staging area for large data transfers across the UNH Science DMZ.

Future State (supporting CI Plan goals #2, #4, and #5): A balance of managed and shared storage solutions will address the varying needs of the research enterprise. Central storage solutions will offer resilient tiered capacity to accommodate data management at each stage of

the data life cycle from acquisition through archival. Cloud services will be periodically assessed and expanded as deemed effective with regard to both cost and functionality. In all future storage solution configurations, having a clear alignment to data management plans, security of data and data accessibility and movement are primary factors.

IPv6 and InCommon Federation
Current State: The infrastructure to support IPv6 protocols at UNH has been in place since the IPng Working Group was proposed in the Internet Engineering Task Force. The UNH-IOL (InterOperability Laboratory) established a collaborative project in 2003, called Moonv6, in conjunction with the North American IPv6 Task Force and the Joint Interoperability Test Command. UNH-IOL's initial involvement resulted in their leading the U.S. Government's deployment of IPv6 natively across the country in government installations. Involvement by UNH students has enabled the UNH IT community to garner extensive, real world, IPv6 deployment experience. In addition to the UNH-IOL USGv6 IPv6 native network, the UNH campus network has a working native IPv6 network for production systems. Critical services (e.g., routing and name services), are resident on our IPv6 test network. A current IT project maintains electronic mailing services, web services and domain name services on a native IPv6 stack. The core campus network is running IPv6, and UNH IT has select UNH research and operations units operating either in native or bridged IPv6 on this network. UNH maintains a custom-built identity management system with LDAP backend, serving UNH and Granite State College students. This system provides a single sign-on solution for many campus-based systems for UNH students, faculty, staff, alumni and various "associated" users.

The North East Cyberinfrastructure Consortium (NECC) research work in the NH NSF EPSCoR RII Track 2 grant made evident the need to authenticate across institution domains. (Many researchers needing to access devices outside the UNH campus network needed to maintain independent login and password combinations to accomplish their research.) In 2012, UNH IT joined InCommon and deployed federation using Shibboleth between research devices on the UNH campus.

Future State (supporting CI Plan goals #1, #2, and #5): UNH's entire network will operate as a native IPv6 network. Staff members leading the migration efforts will present to private and other public institutions in NH and in the region on the process and merits of transitioning to native IPv6. Through InCommon Federation, UNH will achieve its goal of advancing certification to make remote resource access seamless for researchers, as well as making all UNH resources accessible to off-campus researchers and educators.

Cyberinfrastructure Personnel
Current State: UNH's CI investment is focused on building the tools, infrastructure, personnel and experience necessary to remain competitive in pursuing scientific research. To address the demands of providing and supporting cyberinfrastructure, UNH created a CI Strategic Working Group with representation from the IT Strategic Technology office, the Research Computing Center, the Research Office and the IT Broadband Services group. All IT personnel currently engaged in deploying CI at UNH are doing so in addition to their existing responsibilities – most notable is Scott Valcourt's role of acting CI director.

Future State (supporting CI Plan goals #3, #5, and #6): To effectively integrate and maximize CI resources, a team of dedicated CI personnel will be assembled to lead the strategic build out of CI consistent with the overarching goals listed in the above Executive Summary. A comprehensive training program will be implemented to promote and raise awareness of the availability of UNH's cyberinfrastructure. Working with the faculty from various disciplines,

students interested in CI will participate in those disciplines that have a need for advanced support including HPC and GPU programming, data analytics, and visualization. Working together, the community of CI-aware technologists will grow organically.

Data Management Services
Current State: UNH has in place existing data management policies, procedures, and services. Resources include guidance and templates on developing data management plans, proposal development collaboration tools, and a digital repository.

Future State (supporting CI Plan goals #3 and #5): UNH will enhance its data management services, aligning the efforts from IT, the Research Computing Center, the Research Office and the Library. Enhancements will include incorporating any new Federal Uniform Guidance on data management, developing more comprehensive services to PIs developing data management plans for research proposals, and incorporating the use of an institutional repository. A Research Data Services Librarian will be established to bring the provisions for archival and preservation further into the realm of digital data management. Data Management Services will address the specific needs of each stage of the data lifecycle including: data creation (conceptualization), data collection and description, data storage, archiving and preservation, data access, discovery and analysis, and data re-use and transformation.

Institutional Repository
Current State: The UNH Library maintains an institutional repository to properly store, discover, and retrieve faculty, student and institutional scholarly work including datasets, research materials, reports and published findings -- ensuring continued access for the benefit of UNH and the public at large. Associated services include statistical tracking and citation reporting, access to faculty research profiles, and assistance with author and intellectual property rights.

Future State (supporting CI Plan goals #2, #3, and #5): Expanded data curation services will include direct support to faculty in defining and achieving long-term maintenance of <u>all</u> research data such that it is available for reuse and preservation as stipulated in associated data management plans. The repository will connect to the CI to serve as the primary source for UNH-acquired and maintained data to support educational, research and outreach project goals and encourage data meshing to advance new discoveries when datasets achieve a common x-axis.

Data Analytics and Visualization
Current State: Data Analytics and Visualization is pervasive throughout UNH's academic, research, and administrative applications. The UNH Master of Science in Analytics program produces graduates equipped with the technical and professional skills to help businesses, other organizations, and individuals to use data effectively and efficiently to produce real-time results. Analytics are used extensively on UNH research projects, aided by both custom and commercial software. Administrative offices (such as admissions, finance, research, and advancement) are either actively using or assessing the use of new commercial tools to provide accurate views of current and historical data.

Future State (supporting CI Plan goals #2, #3, #4 and #5): Data Analytics and Visualization will enable UNH to leverage data to improve institutional decision-making and strengthen our ability to achieve our teaching, research, and engagement mission. The services will impact every area of UNH – from a department manager performing budget forecasting with modeling tools, to the researcher analyzing large complex datasets using new and emerging technologies, to

the student taking coursework in the field. The analytics and visualization will evolve to include real time data processing and predictive modeling.

Bridging the UNH CI Current State/Future State Gap

At UNH, numerous CI demands compete for the same resources. In order to maximize the return on investment, campus leadership makes informed decisions about expenditures by involving administrators, technical personnel and research faculty in the decision-making process. For example, UNH is unfolding a campaign to engage faculty in an ongoing effort to prioritize their research CI needs and communicate those needs to UNH's CI leadership. A similar effort is underway to identify the academic needs of the faculty. In both cases, faculty share their current and anticipated assessments of the CI as it impacts their specific UNH needs. This input is then included in a prioritization process. Those CI needs that are at the top of the list are first in line for deliberate investments including both the pursuit of targeted extramural funding as well as strategic institutional allocations.

UNH CI Plan: 2015 Technology Prioritizations

The CI Strategic Working Group, through discussions with administrators, faculty, and technical personnel, has identified the following technologies, over-arching goals impacted, and their associated rankings for 2015, highlighted in Table 1 below.

UNH CI Technology	UNH CI Plan Overarching Goal(s)	Ranking 2015
Network Capacity & Regional Connectivity	#1, #2,	6
Compute Capacity	#2, #4	2
Storage Capacity	#2, #4, #5	3
IPv6/InCommon Federation	#1, #2, #5	5
Cyberinfrastructure Personnel	#3, #5, #6	1
Data Management Services	#3, #5	4
Institutional Repository	#2, #3, #5	8
Data Analytics and Visualization	#2, #3, #4, #5	7

Table 1 – UNH CI Plan 2015 Priorities