EPSCoR

OneTenn

A 21st Century Cyberinfrastructure for Tennessee

A Plan for Creating and Sustaining a 21st Century Infrastructure for Research and Education in Tennessee

A report by the Cyberinfrastructure Commission of the Tennessee EPSCoR Project

March 2005

State University & Community College System of Tennessee

UT

Oak Ridge National Laboratory

Managed by UT-Battelle for the Department of Energy

Vanderbilt University

Acknowledgement

The CIC Co-Chairs acknowledge the effort expended by the Commission members and particularly the significant support and contribution of the subcommittee co-chairs.

Brice Bible, University of Tennessee
Paula Short, Tennessee Board of Regents

Eric Cromwell
Economic & Community Development

Optical Network Team

John Lankford (Co-Chair), UT
Tom Danford (Co-Chair), TBR
Mike Abney, UTM
Nasir Ghani, TN Tech
Jeff Kell, UTC
Lucinda Lea, MTSU
Dave Mathews, VU
Ana Preston, UT
Predrag Radulovic, UT
Mark Reavis, UM
Gregory Schaffer, MTSU
Taylor Strickland, UTHSC
Mike Turpin, ORNL
Monty Wilson, UTC

Research Computing Team

Doug Hurley (Co-Chair), UM
Terry Moore (Co-Chair), UT
Jack Buchanan, UTHSC
Arlene Garrison, UT
Edwin Koshland, UM
Joel Muehlhauser, UTSI
Jeanne Hermann-Petrin, UTHSC
Francis Otuonye, TN Tech
Abdul Rao, MTSU
Paul Sheldon, VU
Tim Swafford, UT SimCenter
Linda Warner, TBR
Dave Whitfield, UTC
Michael Woodruff, ETSU

TABLE OF CONTENTS

OneTenn: A 21st Century Cyberinfrastructure for Tennessee

A Plan for Creating and Sustaining a 21st Century Infrastructure for Research and Education in Tennessee

A report by the Cyberinfrastructure Commission
of the Tennessee EPSCoR Project

I. EXECUTIVE SUMMARY

A. Advanced Cyberinfrastructure: An Historic Opportunity for Tennessee

Across America, states are making strategic investments in essential technology infrastructure as the information technology revolution continues to redefine the conditions of competition in research, education, and business in a global economy. Today, the most critical investments are in *cyberinfrastructure*. The term was coined at the National Science Foundation (NSF) to denote the technology fabric underlying new research environments that combine computation, massive data storage, information management, networking and intelligent instruments, and sensing systems. The powerful new techniques and forms of collaboration enabled by cyberinfrastructure far outperform previous capabilities and empower scientists to attack and solve scientific, medical, and engineering problems.

Advancements in science, medicine, and engineering are dependent upon massive flows of data that are generated, moved, stored, and analyzed by research teams from different fields who are often scattered geographically and organizationally. By overcoming the traditional barriers of time and place, cyberinfrastructure enables access to the latest scientific tools and collaboration among world class researchers, top caliber students, and highly-trained and experienced technical personnel, who together put these remarkable technologies to their most creative uses

This report proposes the establishment of a cyberinfrastructure for Tennessee (**the OneTenn cyberinfrastructure**). Cyberinfrastructure, according to the National Science Foundation, combines computation, massive data storage, high-speed networking, intelligent instruments, and sensing systems to provide the foundation for solving new classes of scientific, medical, and engineering problems. OneTenn replaces the current ineffective and costly infrastructure, which has hindered collaboration, with an innovative state-of-the-art infrastructure. OneTenn will enable Tennessee to participate in the national cyberinfrastructure and to effectively compete in the 21st Century information economy. OneTenn will provide dramatic advantages for education, research and economic development statewide. And, OneTenn positions Tennessee to successfully compete with any state in the 'new economy.'

B. OneTenn Goals

1. **Accelerate the growth of statewide research opportunities:** The recent statewide assessment of research strengths by the Tennessee EPSCoR committee clearly showed a strong, widespread desire for intra-state collaboration within the Tennessee research community. OneTenn will foster the development of "co-laboratory" environments, enabling researchers to work with colleagues at distant institutions in a truly collaborative manner. It will enable new partnerships across institutions that have not been defined traditionally as research universities and permit peer interactions with junior scientists and

students despite differences in experience, community or school demographics, research facilities, or local infrastructure.

2. **Connect state researchers to cutting-edge national infrastructure:** Integrating Tennessee's research community into emerging cyberinfrastructure initiatives at the global, national, and regional level is crucial for continued scientific and technological advancement within Tennessee. OneTenn will enable Tennessee and the Tennessee Valley to emerge as the next "Silicon Valley" - a focal point for attracting leading-edge technologies, researchers, investors, and the brightest students to a beautiful geographic region with an outstanding quality of life.

3. **Support education across the state through the distribution of digital content (video and multimedia)**: OneTenn's unique combination of high-performance networking, computational resources, and distributed storage will give Tennessee's higher education community a foundation for multimedia (digital) production and distribution. Colleges and universities can collaboratively develop media-intensive high-quality course content, and then easily distribute that content virtually anywhere, including the Tech centers and P-12 schools.

4. **Expand access to essential research computing resources**: Scientists and engineers at many Tennessee colleges and universities lack access to adequate computational resources necessary to participate in computationally-intensive research. OneTenn will dramatically improve their ability to collaborate and utilize computing capacity across the state. With OneTenn's network and storage fabric to handle the computation logistics for today's large data sets, researchers at smaller schools will be able to access computing power at larger universities, research centers, and at Oak Ridge to accelerate their work.

5. **Expand research opportunities with Oak Ridge National Laboratory (ORNL)**: With its world-class scientists, its one-of-a-kind scientific instruments (e.g. the Spallation Neutron Source), its leading-edge computational resources, and its involvement in national/global research programs, ORNL is one of the world's leading research organizations. ORNL offers incomparable value to Tennessee's research community. Tennessee has made, and continues to make, substantial investments in ORNL through the Joint Institutes. OneTenn will dramatically improve the ability of all colleges and universities in the state, as well as industrial and corporate research partners, to engage with ORNL and leverage the state's investment in this world-class research facility.

6. **Raise the level of math, science, and engineering education**: Science, math, and engineering education (P-12 through post graduate) can benefit tremendously from cyberinfrastructure-based services. Expert mentoring, especially in the many emerging 'new science' interdisciplinary areas, can support students, faculty, and teachers anywhere there is access to OneTenn. OneTenn will encourage/prepare the next generation of scientists, engineers, and technical personnel through broad training in scientific simulation and modeling for fields that are both computationally intensive and interdisciplinary in nature. This will become the basis for the 21st Century global economy.

7. **Enable innovative approaches to public health care and education:** Tennessee ranks 48th among the states with regard to the health of its citizens. OneTenn could provide the mechanism to share expertise and instructional material as well as medical services over wide distances to support a statewide collaborative school of public health. OneTenn could enable collaboration between researchers and faculty at multiple physical locations, and

support community based health initiatives. Such a statewide distributed Public Health instructional, patient care, and research effort could serve as a model for the nation and enable initiatives that could not otherwise effectively be undertaken.

8. **Enhance collaboration among Tennessee institutions:** Each Tennessee higher education institution provides a variety of business solutions for effective delivery of education, research, and public service. Many of these solutions are technology-oriented and include services ranging from communications (e-mail, calendaring, video conferencing, synchronous and asynchronous course management systems), to business processes (financial and student information systems), to advanced computational applications (academic computing systems, storage/back-up, disaster recovery). The OneTenn network infrastructure provides a modern solution for efficient and effective sharing of these and may other solutions – a unified network architecture for the entire state.

9. **Serve as Catalyst for Economic Development in the 21st Century:** Economies change, and technology accelerates the pace of change. In order for Tennessee to compete and win in a global economy increasingly powered by knowledge, services, and technology, the state must have the necessary infrastructure that enables current and future economic growth. A robust cyberinfrastructure is a necessary component to a thriving research, education, and business community and its absence severely limits many economic development opportunities. The amount of data being created and transported around the world is staggering, and having the right information at the right place, and at the right time, is a competitive necessity. This enabling technology could generate new opportunities for importing knowledge workers, capital, and businesses to Tennessee. Furthermore, the network architecture is designed to traverse the state, allowing for the formation of partnerships with rural communities to explore new strategies to better connect all of Tennessee to the information superhighway.

C. The OneTenn Vision

The OneTenn plan is groundbreaking in comparison to other similar state initiatives. OneTenn incorporates a statewide cyberinfrastructure that unifies high performance networking, computational resources, and data storage into one system, creating a pool of resources that will enable collaboration across the state in ways that would otherwise be impossible. Its foundation is a state-of-the-art, high-speed communications backbone, based on dark fiber and Dense Wave Division Multiplexing (DWDM) technology to provide a variety of flexible optical network infrastructure services The OneTenn network infrastructure will connect twenty (20) major Tennessee cities and all state universities and colleges with multiple high-speed connections (see Figure 1). Every institution will be connected via a minimum of 1.0 Gigabit bandwidth with the state research universities connected at 10 Gigabits or greater speeds. This highly-scalable solution will be capable of supporting multiple, fully independent, network configurations on a single infrastructure. In addition, the OneTenn network flexibility is crucial to research collaboration and growth for the state.

Key Features of OneTenn

- *it provides* access to large scale essential technology resources for shared use
- *it enables* collaboration across organizational and geographic boundaries and barriers
- *it stimulates and encourages* innovation in scientific, engineering, and medical research

A second key OneTenn innovation will be a new form of network storage technology based on the Internet Backplane Protocol (IBP). IBP extends the Internet design for interoperability to storage resources, making it possible to aggregate and use the resources of widely scattered storage nodes, called "depots," as if it were one giant storage pool. This innovative optical network and storage infrastructure will also provide every Tennessee university campus with very high-speed connectivity to other Tennessee and DOE (ORNL) computing resources, as well as insure optical access to other national computational resources. OneTenn cyberinfrastructure will place Tennessee at the forefront of advanced network initiatives across the country.

OneTenn will create an advanced cyberinfrastructure to serve the research, education, and public service mission of Tennessee's entire higher education community in the 21st Century. Its impact, however, will be even broader. OneTenn can also provide the necessary framework for cost-effective services delivered via broadband to P-12 education and to economic development and workforce development initiatives. The OneTenn infrastructure is designed for exceptional scalability and growth, so that future investments, (public and private, local and statewide) can extend the OneTenn infrastructure to all parts of the state to provide a variety of services.

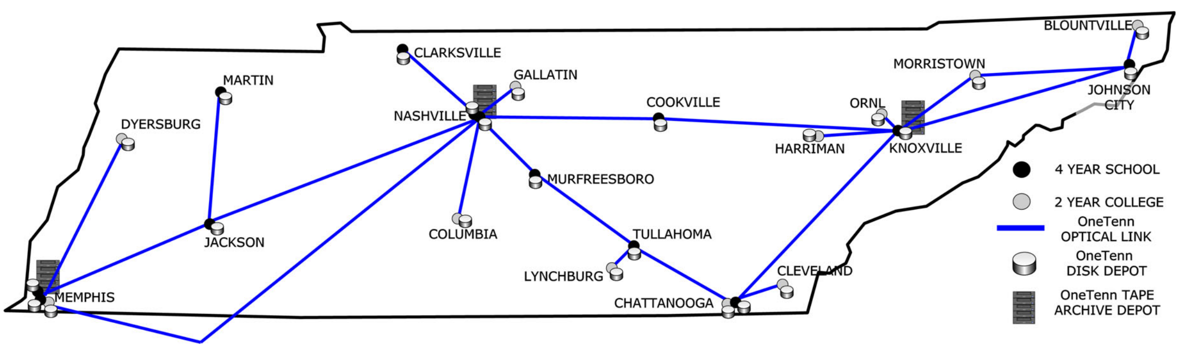

Figure 1: The OneTenn Infrastructure upon Completion

D. Three Phase Plan for Deployment

A comprehensive cyberinfrastructure such as OneTenn is complex and will require a multi-year implementation plan. The proposed plan is in three (3) phases extending over five years (see Figure 2). The OneTenn plan capitalizes on the historical role that research universities across the country have played in leading advanced cyberinfrastructure development. UT Knoxville, Vanderbilt, UT Health Sciences Center, the University of Memphis, and ORNL**,** as well as others, are already well positioned to take advantage of OneTenn. These institutions have substantial research communities involved in data intensive research, well-developed campus networks, and high-speed access to national research networks. Building on the foundation of these principal research universities, the OneTenn plan will progressively strengthen the capacity of all Tennessee universities and colleges to participate in, utilize, and benefit from the OneTenn cyberinfrastructure.

Figure 2. Three Phase Plan for OneTenn Deployment

	High Performance Networking	Distributed Storage and Compute Systems	Values and Services
Phase 1	+Connect the research sites at UTK, Vanderbilt U, the UoM, & UTHSC +Connect to ORNL +Connect to St. Jude's Children's Hospital and other state-based organizations	+Provide large scale data storage/compute services to enable research, educational, and e-Learning collaborations across the state	+Enable Tennessee to compete nationally for the most advanced research initiatives which can be linked to economic development opportunities +Shared use of computational services among the Tennessee research communities
Phase 2	+Connect remaining public universities +Enhance OneTenn connections to national research networks	+Install large scale storage/compute services at remaining public universities +Deployment of large storage/compute to selected 2 year colleges +Early storage deployment in selected P-12 sites	+Digital video distribution services for e-Learning across the network for all Tennessee higher education and P-12 +Collaborate with Tennessee communities for enhanced workforce and economic development
Phase 3	+Connect all 2 year colleges, technology centers, and selected P-12 sites +Enhance OneTenn connection to nation research networks	+Complete deployment of large scale storage/compute services at 2 year colleges +Continue storage deployment in P-12 sites	+Expand opportunities for economic development +Advanced computational and storage services to enable digital video distribution and e-Learning, as well as access to other high tech resources across Tennessee

E. Managing and Leading OneTenn

Historically the breakthrough, most advanced information technologies (e.g. the Internet and the World Wide Web) have emerged from the academic and research environments. The Web was invented as a tool for information sharing among high-energy physicists long before it became a pervasive tool for mass communication, business transactions, and entertainment. The impetus to sustain and accelerate the development of advanced technology tools for information access and dissemination comes from the research and education community.

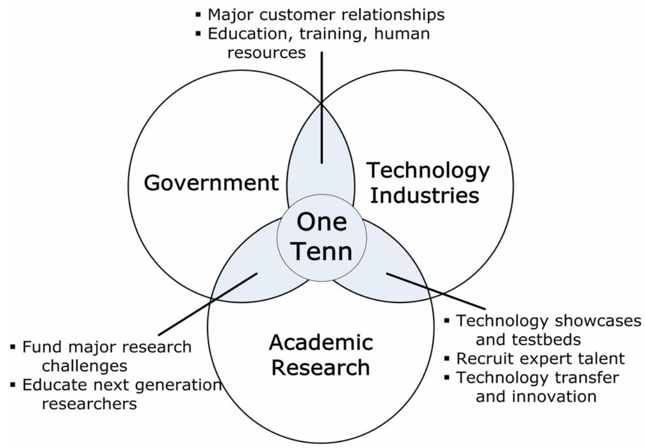

Figure 3: The shared need for cyberinfrastructure among government, industry, and the higher education research

Government, especially at the federal level, has long recognized the necessity for nurturing the environments that create these scientific advancements. At the same time, government and industry also recognize their need to exploit the potential of advanced information technologies for economic and social development. Thus OneTenn should be managed as a public infrastructure service (see Figure 3). To ensure that OneTenn retains its leading-edge scientific orientation, the Tennessee research community will be incorporated into a leadership team composed of government, industry, and higher education. The strength of OneTenn is that it can serve the needs of all three communities.

Due to its innovative technology underpinnings, OneTenn can simultaneously provision the fastest networks for leading edge research, deliver creative learning content for P-12 education, extend healthcare to underserved populations, and sustain workforce development programs to support emerging industries in rural Tennessee.

F. The Opportunity is Now

Tennessee is at a cyberinfrastructure crossroads. The window of opportunity to purchase dark fiber assets will close soon and bargain-basement pricing will disappear. As the Technical Plan will show, Tennessee can purchase, deploy, and operate the OneTenn cyberinfrastructure for less than \$30M over five years with recurring operating costs of approximately \$3.2M annually once fully deployed. There is actually no direct comparison to the current leased solution as the state's commercial infrastructure is limited to OC-48 (2.5Gbps) speeds. However, the cost of a lease option for the OneTenn capacity has proven to be much more costly in other states when compared to the proposed CIC strategic approach. In fact, the current funding level of leased line connectivity for all Tennessee public higher education institutions is more than adequate to cover operational costs associated with OneTenn acquisition. In addition, the OneTenn network will rely on flexible and reliable Ethernet switching technologies as compared to the traditional ATM technology approach still utilized by most state-based commercial carriers. It's no wonder nearly every surrounding state has initiated or completed procurement of such a versatile optical network solution.

G. Conclusion

Research powers the hi-tech engine of revolutionary new discoveries, new knowledge, and new opportunities while sustaining established knowledge and mature industries. Developing Tennessee's statewide cyberinfrastructure is essential for continued advancement of science and engineering within the state. OneTenn is based on the core strengths of its leading research universities. Critical to its success will be the development of a framework that promotes overall collaboration and sharing so that use of the OneTenn cyberinfrastructure will continue to scale over time.

In the 21st Century, collaborative research teams are essential to advances in science and engineering. Cyberinfrastructure must provide mechanisms for creating, building and maintaining communities of expertise. New researchers and research teams will emerge and join the cyberinfrastructure, contributing new technologies and creating new science, knowledge, communities, expertise, and opportunities for Tennessee. New scientific communities emerge around 'big science' large-scale research projects that become the foundation for new economic investments.

The OneTenn cyberinfrastructure plan represents a strategic opportunity for Tennessee to address a diverse set of goals, spanning research, education and economic development. By connecting the state to cutting-edge national infrastructure and expanding access to essential computing resources, OneTenn will dramatically enhance research opportunities statewide. By supporting distribution of high end video and multimedia content and providing online access to databases and instruments to all of the state's research and education institutions, OneTenn will help to raise the level of math, science and engineering education and enable new modes of health care delivery. The OneTenn plan is highly innovative but draws on experienced technology creation, deployment and user groups throughout the state and across the country. Putting cyberinfrastructure at the foundation of Tennessee's plan for competing in the 21st Century will enable the state invest strategically in knowledge and communication, the two most important drivers of innovation and economic success over the past 50 years.

II. Tennessee Analysis and OneTenn Strategy

The CIC surveyed the principle computational assets available today to support research in Tennessee and the results are summarized here. Computational resources now in place at the major research universities in the state are listed along with on going efforts to make them sharable. This section also includes a synopsis of some of the leading research groups in Tennessee and their involvement in national and international cyberinfrastructure efforts. The CIC does not consider this as an exhaustive listing of research expertise and resources within Tennessee universities, and any inadvertent omissions are unintentional.

After a review of computational and network resources across the state, the conclusion is clear: Tennessee lacks the essential components of cyberinfrastructure i.e. no robust, high speed network, no unified, accessible storage, and minimal ability to share scarce computing capabilities. UT Knoxville, Vanderbilt, and the University of Memphis, as well as other campuses may have robust network infrastructures, large data stores, and compute clusters that are shared within their campuses. But in most cases these storage and computational resources were established to serve local research communities and cannot be freely repurposed, and, with rare exceptions have not been designed to be easily sharable across the state. Moreover, even if they could be shared, Tennessee's statewide network infrastructure is inadequate.

There are key differences in how networking, storage, and computational resources fit into the cyberinfrastructure paradigm. These differences are briefly summarized as follows:

1. **Networks are shared across administrative domains**: Network resources can be shared much more easily than storage and computing services. IP networks are designed to be used across administrative domains without requiring any authorization; storage and computing are designed to be shared only among authorized users within well defined domains, not across them.
2. **Networks have standards for interoperability**: The Internet protocol suite provides a de facto standard for interoperability that makes it possible to use networks deployed statewide (and worldwide) as a single infrastructure. Nothing equivalent exists for storage or high performance computing resources. (OneTenn's storage component is unique because it is designed for interoperability.)
3. **Computing is too complex to standardize**: Various factors combine to make computer processing inherently more difficult to integrate as a shared resource into a unified scheme that spans organizations than either networking or storage.

A. Tennessee's Network Infrastructure Today

The Tennessee EPSCOR Committee has identified the lack of an adequate network infrastructure as one of the most significant barriers to collaboration and to sharing research resources within the state. Several independent production networks are in use across the state at the present time, some of which are shown in Figure 4. At best, these networks are marginally adequate for current needs of each community they serve. None effectively interoperate, nor can any of them scale to meet future needs. All of these networks are both inadequate in capabilities and excessively costly to operate. Tennessee Board of Regents institutions are connected through the TNII[1] network, K-12 institutions are connected via the ENA[2] network and the University of Tennessee system operates its own

[1] Tennessee Information Infrastructure (TNII): http://www.tnii.net
[2] http://www.ena.com

network. A few schools are connected to a separate video network via low-speed (mostly single or multiple T-1's, 1.54 Mbps each) commercial links, which cannot effectively support H.323 videoconferencing together with normal data flows. The four research universities (The University of Tennessee, Knoxville, Vanderbilt University, the University of Memphis and the UT Health Science Center) connect via OC-3 (155 Mbs) or higher speeds to Abilene, the international high-bandwidth research and education network provided by the Internet2 consortium[3]. A few other universities connect to Abilene at substantially slower rates (DS-3 level, 45 Mbps) via the principal universities listed above.

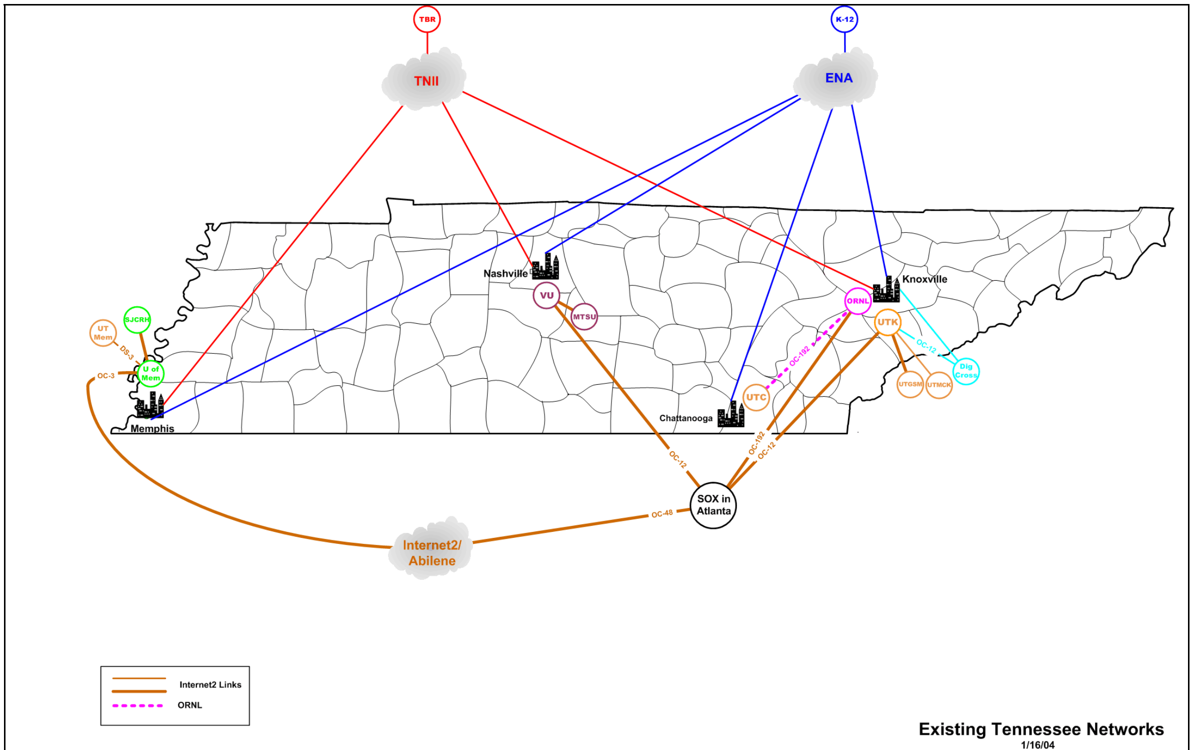

Figure 4: Existing Tennessee Networks

The following examples illustrate both the inefficiencies and the costliness of the present networking environment for research and education in Tennessee. When using the Abilene network, traffic destined from either Vanderbilt University or UT-Knoxville must go through Atlanta, Indianapolis and Kansas City before reaching the University of Memphis or the UT Health Sciences Center. Similarly, communication between UT-Knoxville campus and UT Space Institute in Tullahoma must go through Atlanta. Even more bizarre, for communications between the UT-Knoxville and Pellissippi State Technical Community College (approximately 15 miles distance) the traffic /flow from UT-Knoxville must go all the way to Dallas, TX before coming back to Pellissippi State. These kinds of communications inefficiencies cause unacceptable delays while increasing transport costs as data traverses multiple carriers over large distances and consumes out-of-state resources. To compete in the 21st Century, Tennessee must have a state-of-the-art network infrastructure; an infrastructure that can scale to meet growing needs at a cost effective price.

[3] http://www.internet2.edu

B. Tennessee's High Performance Computing Resources Today

Tennessee is limited in the computational and storage resources that are *available to be shared by a broad, statewide user community*, but the state is not unique. Although significant computational resources exist throughout the state, the lack of a network infrastructure makes them not easily accessible. This is a problem because modeling and simulation in most science and engineering fields, along with the need to analyze the huge volumes of data generated by advanced instruments, necessitates an increasing amount of computational power. While the major computing assets available to scientific researchers on some Tennessee campuses are significant, leading edge problems in fields such as Materials Science, Environmental Modeling, and Genomic Analysis are so compute and data intensive that they frequently outstrip the capabilities of even the largest computation facilities on these Tennessee campuses. In such circumstances researchers either scale back the problem they attempting to solve or go off site to national labs and NSF centers to compete for time on the nation's largest supercomputers. Researchers at smaller schools with fewer resources confront precisely the same problem on a more limited scale, but have an even harder time getting remote access to the additional processing power they need. This chronic scarcity of large scale computation assets is a fundamental problem for science, yet it is a problem that can easily be solved through the OneTenn cyberinfrastructure.

The following section describes some of the more significant high performance computing assets in Tennessee and their potential as part of the cyberinfrastructure under OneTenn. It should be noted, however, that ORNL's remarkable collection of computing resources represent a *national asset* that happens to be located in Tennessee.

1. University of Tennessee, Knoxville

One of the world's strongest research groups in the area of grid computing is the Innovative Computing Laboratory (ICL), lead by Jack Dongarra of UT Knoxville's Computer Science Department and the UT Center for Information Technology Research. Over the past five years, with the support of a major NSF infrastructure grant, Dr. Dongarra and his colleagues have deployed the Scalable Intracampus Research Grid (SInRG) on the Knoxville campus. The purpose of the SInRG infrastructure is to provide the controlled testbed required for the research necessary in order to make grid-enabled computing a reality for a broad scientific, medical, and engineering community. The current SInRG deployment is presented in Figure 5. It includes both clusters acquired under the NSF award and donations from other project partners, including Microsoft, Hewlett Packard, AMD, Myricom, Sun, and Dolphin. The UT's central IT department (OIT) has also contributed to SInRG in the form of a large cluster available to the campus community via SInRG system software.

Working with teams of CS and domain scientists, SInRG research has focused on software that makes grids easy to use and on a grid fabric that can scale globally. In fact, software developed on SInRG will play an important role in the OneTenn plan. The distributed storage software based on the Internet Backplane Protocol (IBP) will provide the foundation for the fabric of data storage that OneTenn will deploy, and NetSolve grid software, developed by Dr. Dongarra's group, will enable researchers across the state to use SInRG via OneTenn. As described below, IBP storage depots are already deployed at Vanderbilt and the University of Memphis, and a deployment of NetSolve is also being tested for deployment on the Vanderbilt campus.

This deployment of Grid hardware and software has enabled SInRG's CS researchers to be major participants in national projects and initiatives, such as the NSF Teragrid, that are exploring the scientific and technical problems that must be solved to make grid application development and performance tuning for real applications an everyday practice. More importantly, the grid system software developed at UT has enabled a variety of interdisciplinary projects that have tested and used the SInRG infrastructure. Collaborative projects, some of which included participants from Vanderbilt, ORNL, UT Health Science Center, and other out-of-state institutions, were carried out in a variety of different fields, including remote visualization of medical images, distributed data mining, community digital video recording, active digital libraries, environmental modeling, computational analysis of gene regulatory networks, and distributed gene sequencing.

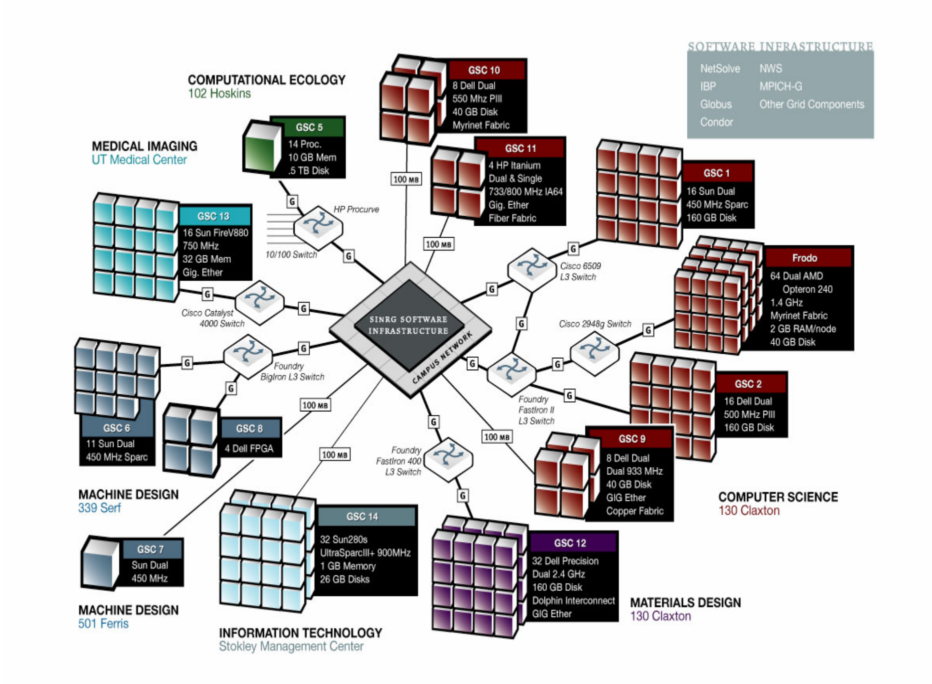

Figure 5: Hardware and software resources of UTK's Scalable Intracampus Research Grid. (SInRG)

Many other UT Knoxville colleges/departments maintain one or more compute clusters. The UT Office of Information Technology (OIT) has joined forces with the Colleges of Engineering and Arts and Sciences to centrally support a series of mid-range clusters (32-256 node clusters), designed to meet the various computations needs of the research community.

2. Vanderbilt University

To support its large community of computational scientists, Vanderbilt's Advanced Computing Center for Research and Education (ACCRE) has put in place an impressive installation of computational resources. ACCRE's compute hardware is comprised of 1408 processors of several different hardware types and a combined compute power of 6 TFLOPS. All systems involved use the same scheduler and performance monitor, and they all see the same disk for each user's home directory (9TB per user) All machines have at least 1GB of memory, a 40GB disk drive, and gigabit Ethernet connections.

To support even more data intensive efforts, Vanderbilt researchers are involved in several national cyberinfrastructure projects, including the international Virtual Data Grid Laboratory (iVDGL), the Particle Physics Data Grid (PPDG), and the Open Science Grid.

3. University of Memphis

The University of Memphis currently has most of its computational resources located in several departments on campus: Chemistry, Physics, Mechanical Engineering and Computer Science. Several computational scientists from these departments are also members of the Computational Research on Materials Institute (CROMIUM) and whose research relies heavily on modeling and simulation. Each department has one or more compute clusters and a project is currently under way to join these clusters into a campus-wide computing grid. Phase II of this project will join the computers in many of the computer labs on campus to this grid resulting in a grid of close to 1000 processors. In spring 2005 the University will acquire a centrally managed, high performance research computing facility. Configuration, performance and vendor have yet to be determined.

4. Oak Ridge National Laboratory

ORNL's Center for Computational Science has the nation's most modern computer facility for unclassified research computing. The Eugene P. Wigner Center for Computational Sciences Building, located on ORNL's secure campus, has 40,000 square feet of raised-floor computer space designed specifically for leadership-class computer systems. The facility can provide up to 12 megawatts of power and 4,800 tons of cooling with redundant capacity to allow concurrent operation and maintenance. CCS Staff has extensive experience in supporting advanced computing systems for national user communities, such as the SciDAC programs, and evaluation of computing architectures.

CCS has a professional, experienced operational and engineering staff comprised of groups in Operations, Computing Systems, Storage Systems, Networking, Future Technologies, Scientific Applications Support, and Scientific Visualization. The CCS computer facility is staffed 24 hours per day, 365 days per year to provide for continuous operation of the center and for immediate problem resolution. On evenings and weekends, the operators provide first-line problem resolution for users with additional user support and system administrators on-call for more difficult problems. Primary CCS systems include the following:

a. **Cheetah:** a 27-node IBM Power-4 system. Each Power-4 node of Cheetah has 32 1.3-GHz Power4 processors. Twenty of the nodes have 32 GB of memory, five nodes have 64 GB of memory and two nodes have 128 GB of memory. The peak performance of Cheetah is 4.5 teraflops.

b. **Phoenix:** a Cray X1, with 512 multistreaming vector processors (MSPs) and two TB of globally addressable memory. Each MSP has two MB of cache, and four MSPs form a node with 16 GB of shared memory. Memory bandwidth is very high, roughly half the cache bandwidth. The interconnect functions as an extension of the memory system, offering each node direct access to memory on other nodes at high bandwidth and low latency. The peak performance of Phoenix is 6.4 teraflops.

c. **Eagle:** a 184-node IBM RS/6000 SP with 176 "thin" nodes (four 375 MHz Power3-II processors and 2 GB of memory) and eight "wide" nodes (two 375 MHz Power3-II processors and two GB of memory). Eagle has 2.5 TB of SSA disk attached. The peak performance of Eagle is 1.0 teraflops

d. **Ram:** a 256-processor SGI Altix with two TB of shared memory. Each processor is the Intel Itanium2 1.5 GHz processor. The full system runs a single Linux image and the large shared memory facilitates analysis of very large data sets. The peak performance of Ram is 1.5 teraflops

In addition, the CCS recently won the competition to build the fastest computer in the world open to all users which would be housed in ORNL's National Leadership Computing Facility. The proposed hardware roadmap contained three purpose-built architectures optimized for scientific applications; Cray's X and Red Storm systems and the IBM Blue Gene system with primary focus on the Cray X series vector solution. Within about 6 months the NLCF will have 20 TFLOPS of Cray X1e, 20 TFLOPS of Cray Red Storm and 5 TFLOPS of IBM Blue Gene. By mid-2006 100 TFLOPS of the follow-on Cray X2 system will be made available to users. The Cray X series is a proven architecture for performance and reliability, has the most-powerful processors and interconnect available, has scalable, globally addressable memory and bandwidth, leverages commodity where possible, and offers extremely high efficiency capability computing for key applications.

The inadequacy of the state's current cyberinfrastructure is the main impediment to Tennessee researchers from having better access to these resources. Enabling the entire Tennessee research community to engage more fully with ORNL by leveraging the state's investment in the Joint Institute for Computational Science (JICS) is a major goal of OneTenn.

C. Tennessee's Storage Resources

Measured by today's standards, where scientific instruments and computer simulations are producing increasingly vast quantities of data, there is very little mass data storage for research in Tennessee. And there is none available statewide as cyberinfrastructure for the whole community. Various research groups have multi-terabyte (TB) storage servers for private file space, but only ORNL and Vanderbilt have data stores that range in size from hundreds of terabytes to multiple petabytes.

On the other hand, a research team in Tennessee has developed the most innovative storage technology for wide area, data intensive collaboration; which will form the basis for OneTenn. This section describes both the conventional mass data stores that exist as private resources in the state, as well as the public Logistical Networking infrastructure. It is important to note that the tape archive that Vanderbilt is constructing will use IBP and Logistical Networking tools to manage the petabytes of storage that it will contain, and therefore it will be fully interoperable with the OneTenn's storage infrastructure. The tape archives at Knoxville, Nashville, and Memphis that will anchor OneTenn will use the same technologies.

1. Oak Ridge National Laboratory

High-performance, scalable storage resources are vital to data-intensive applications. The High Performance Storage system (HPSS) provides archival storage. HPSS has a capacity of up to five petabytes of data and regularly supports data transfers of more than 10 TB per day. Both the bandwidth and capacity of HPSS can be increased as needed. CCS uses a combination of file system technologies for high-speed, local storage. The CCS also

provides a shared secondary file storage system to enable data sharing among computer systems, data analysis systems, visualization systems, and archival storage. A project is currently underway with Cray and other strategic partners to implement a single high-speed shared file system linking all of the computing systems within the CCS. The underlying technology of this file system will be based on the LUSTRE file system developed by Cluster File Systems Inc. ORNL also has a multi-terabyte installation of IBP depots to support wide area collaborations in Astrophysics and Energy Sciences under DOE's SciDAC program.

2. **Vanderbilt**

CCRE is actively developing a large-scale (multi-petabyte) hierarchical storage management (HSM) solution. They are collaborating with UTK's LoCI Laboratory and Fermi National Accelerator Laboratory in the effort. This storage installation will become a major archive of particle physics data produced by experiments with Fermi's particle accelerator. This solution will use IBP software to transfer data to and from a disk-based cache running the dCache software developed at Deutsches Elektronen-Synchrotron (DESY). The dCache software supports a variety of persistence models. Based on the implemented persistence model, dCache will migrate data to and from Fermi's Enstore tape library management system as necessary. All data migration from disk to tape, and vice versa; will be managed by tools co-developed with LoCI Lab. The current configuration consists of five dual tape library/drive configurations (three different models) with a combined capacity of 338 TB.

3. **University of Tennessee, Knoxville**

OneTenn's storage component involves technology transfer of research on *Logistical Networking* with funding from the NSF and DOE at UTK's Logistical Computing and Internetworking (LoCI) Laboratory, under the direction of Micah Beck. Logistical Networking (LN) technologies offer a highly scalable means to manage distributed content using shared network storage. LoCI has developed software tools allowing users to deploy local IBP storage "depot(s)" or to utilize shared IBP storage deployed worldwide to easily accomplish long haul data transfers, temporary storage of large data sets (on the order of terabytes), pre-positioning of data for fast on-demand delivery, and high performance content distribution such as streaming video.

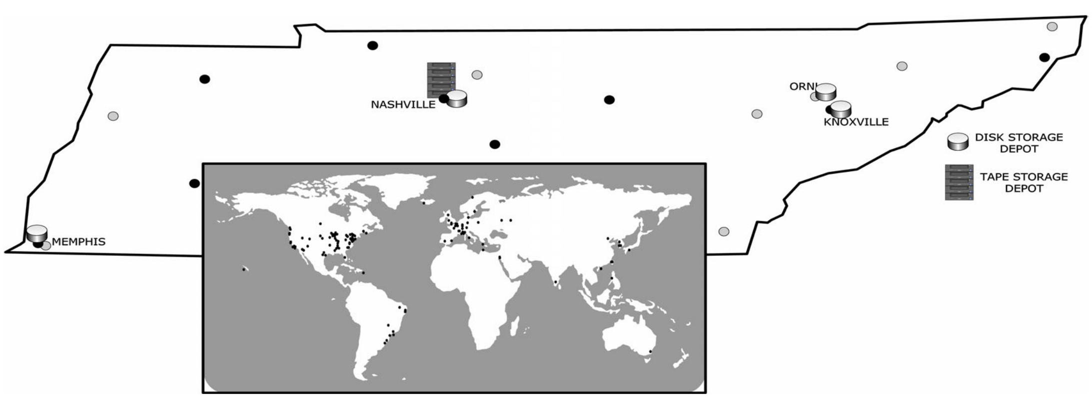

Figure 6: Current deployment of IBP storage technology in Tennessee (background) and worldwide (inset).

The NSF funded National Logistical Networking Testbed (NLNT) is now part of a global infrastructure that, together with the PlanetLab collaboration, includes over 300 depots in the US and over 500 depots worldwide, serving over 30 TB of shared storage (Figure 6).

LoCI software enables users to access and utilize any of these storage depots in an identical way. This infrastructure brings Logistical Networking directly into the hands of researchers, educators, and students nationwide. Large distributed research groups, such as DOE's Terascale Supernova Initiative based at ORNL, routinely use NLNT depots to deliver data to widely dispersed collaborators. Students from high school through university levels are using NLNT resources themselves as part of class projects and student experimentation.

In Tennessee, the disk based depot clusters in Knoxville, Oak Ridge, Nashville, and Memphis collectively have 10TB of storage available. The tape-based depot is the archival unit described above, which is under development at Vanderbilt and currently offers 338 TB of storage. As described below, ongoing research collaborations between UTK, Vanderbilt, and ORNL are exploring new uses for this technology in the areas of optical networking and archival storage.

III. General Overview of the OneTenn Plan

Compared to the cyberinfrastructure plans of other states, OneTenn breaks new ground. What sets it apart is the way that it incorporates, in a single, unified system, not only high performance and advanced optical networking, which other forward-looking states include in their plans, but also a widely distributed and fully interoperable data storage infrastructure. This unique combination creates a statewide pool of resources that will enable OneTenn to provide types and levels of collaboration, both within the state and around the world, beyond the reach of networking-only approaches. Each component of the plan plays a different critical role:

1. **Optical Networking** — Building on a statewide system of fiber optic lines that it will acquire, OneTenn will deploy leading edge optical networking technology that can create multiple, distinct networks across the same segment of fiber optic cable by using separate wavelengths of light. This approach will enable OneTenn to accommodate, on any given segment in the OneTenn "footprint," both production networks, for ongoing and mission critical projects and activities, and experimental networks, for projects that may have extreme requirements or involve research on networking itself. Different networks can be simultaneously deployed to provide precisely the levels of performance, reliability, and security that their users need. For example, wavelengths on particular segments can be dedicated to support special activities, such as remote instrument control, which must be low latency and free of jitter.
2. **Distributed Storage and Computation** —One unique innovation to OneTenn is its use of *Logistical Networking*, a revolutionary new synthesis of networking and storage technology developed at the University of Tennessee's *Center for Information Technology Research* (CITR) and now in use by research and education groups worldwide. Just as conventional logistics coordinates transportation lines and warehouses for the distribution of physical goods, Logistical Networking integrates networking and storage for the distribution, staging, and delivery of data. Its key innovation is the *Internet Backplane Protocol (IBP).* IBP extends the Internet design for interoperability to storage resources, making it possible to aggregate and use the resources of widely scattered storage nodes, called "depots," as if they were one giant storage pool. These storage and statewide computational resources described later in this report, will be supplemented with strategically placed computer clusters for distributed use by Tennessee research community.

By integrating state-of-the-art optical networking infrastructure and computational systems with high capacity, fully interoperable data depots deployed to every public four year university, as well as Vanderbilt and other private schools, and every public two year college, OneTenn will enable resource sharing, data intensive collaboration and rapid innovation across Tennessee's research and higher education community. Researchers and educators will be able to easily share huge files of scientific data or multimedia and video content without the need to create and manage accounts, to pre-position data for fast, on-demand delivery, and to create flexible and affordable content distribution networks capable of delivering media rich content services statewide, even to areas that remain bandwidth challenged.

Below we describe these components of the OneTenn plan in more detail. We also discuss the way in which this infrastructure will be used to support both improved remote access to the high performance computing resources that are currently available to Tennessee's research community, and to develop an approach to grid computing that can become the next generation in Tennessee's cyberinfrastructure.

A. Goals of successful cyberinfrastructure

Though the report of NSF's blue-ribbon advisory panel titled "Revolutionizing Science and Engineering through Cyberinfrastructure[4]" is national in scope, it highlights three key goals of cyberinfrastructure that have guided the OneTenn plan. Cyberinfrastructure should:

1. **Provide essential resources for shared use:** Although the raw capacity of basic computational resources — processing power, transmission bandwidth, and data storage — continue to increase at exponential rates, the requirements of leading edge research (e.g. the size of the simulations to be run, the amount of data to be analyzed) continue to outstrip the computational resources that research teams, and even universities, can acquire for their private use. At all levels, provisioning these basic resources so that they can be shared within larger communities is the only practical way to make sufficient quantities affordably available to researchers in any given field. The shared use of these common elements of cyberinfrastructure has become an indispensable part of large collaborative projects that now dominate advanced research initiatives.
2. **Facilitate collaboration across boundaries:** In the digital era, science and engineering has become more collaborative. Even medium-scale projects typically engage the efforts of researchers from a range of different fields. However, the required collection of experts for a given project is rarely in one place or one organization. The cyberinfrastructure necessary to move, store, and process the large flows of data within such collaborations enable research teams to work across the boundaries that separate disciplines, locations, or organizations.
3. **Incorporate innovation:** New research methods and techniques are primary accelerators of scientific progress. In today's compute and data intensive environment, such innovations must be rapidly assimilated into the shared cyberinfrastructure on which the research process is built. If the systems already installed are resistant to absorbing new innovations quickly, they become obstacles to progress. It is essential, therefore, that cyberinfrastructure be designed to be as open as possible to future innovation.

The Internet is unrivaled as a model of successful cyberinfrastructure. Its success was premised upon the adoption and use of a common communications service (IP) which is so generic that many different kinds of network technologies can implement it and many different kinds of applications can use it. The Internet created a communication facility with an unprecedented ability to enable

[4] Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure: Http://www.cise.nsf.gov/sci/reports/toc.cfm

technologies and applications to interoperate. This remarkable foundation of interoperability has made it possible for research and education communities to deploy massive amounts of network capacity to enable resources to be shared across geographic and administrative borders. New opportunities for new collaboration have emerged, and new technologies and applications have been adopted without damaging the usability of established applications.

B. OneTenn's High Performance Network

As in any other cyberinfrastructure plan, networking infrastructure always plays a major role. Over two-dozen US states are building their next-generation optical networking infrastructure and Tennessee with its OneTenn project is no exception.

In addition to the human factor, the two key building elements of any current bleeding-edge research and education network are the long-term leased dark fiber and the optical gear. Dark fiber is simply fiber not being used and is a result of the over capacity in laying fiber in the 1990s. With the burst of the Internet bubble, a large excess of this fiber capacity drove prices to a historic low. With the economy picking up, these prices will increase again. Seizing dark fiber at these low prices is a major goal of the OneTenn project.

The optical gear (to light the fiber) is based on DWDM (Dense Wave Division Multiplexing) technology. In essence, in DWDM, multiple optical carriers transmit and receive data across the optical fiber allowing clear and secure ways to move data point-to-point anywhere on the fiber footprint. This is similar to multiple radio or TV stations operating in the dedicated band, each transmitting its own programming securely and without interference from the neighboring channels.

OneTenn optical networking fabric will support multiple networks, each capable with at least 10 Gbps (Gigabits per second) rate. This speed allows transfer of the content of a single DVD between any two points on the network within four seconds. This speed is also at least 70 times higher that any currently existing Tennessee state network operated link. OneTenn will open new horizons for general connectivity and add enhanced research capability. This will allow researchers to support these dedicated links with their own grants, something that required multi-million dollar investment under the old networking model. In order to increase capacity of the network by another 10 Gbps anywhere on the OneTenn fiber footprint, only a marginal investment of about $100k is needed.

Initially, between 1 and 4 lambdas will be lit up on any segment of the OneTenn fiber footprint. The current technology (i.e. optical equipment) allows up to 64 channels to be simultaneously in use, while the theoretical limit of the fiber is known to be more than 200 of these channels or lambdas. As we speak, experimental networks already operate at 40 Gbps rates and it is widely expected that by the time that the OneTenn project is completed, rates of up to 100 Gbps on single carrier will be available.

OneTenn optical network will be scalable on the range of a couple of orders of magnitude ensuring that a large initial investment will support all future needs of research and education as well as other connectivity needs.

While OneTenn primarily addresses needs of higher education in Tennessee, the presence of a high-speed optically based network across the state can significantly stimulate the development and spread of broadband technologies in Tennessee.

C. OneTenn's Distributed Storage Services

The OneTenn plan includes a storage component not found in other state cyberinfrastructure initiatives, which typically focus on networking only. The Internet revolution showed that available data storage is fundamental to many powerful network applications. The two most revolutionary applications of the Internet era – e-mail and the Web – are both based on the shared use of remote data storage. And today, by far the vast majority of traffic on Internet2 comes from applications in which data transfer to and from disk resources is the essential mechanism. The plan for OneTenn builds on this fact. The deployment of the OneTenn network will be complemented at each stage by a corresponding deployment of both new data storage and compute services that enable the state's research and education community to take full advantage of the dramatically increased network capacity. This combination will enable rapid development of applications and real expansion of the research and educational efforts across the state to exploit the OneTenn investment.

OneTenn's storage infrastructure will be based on protocols and tools that can serve as the basis for standardization in wide area applications without placing undue constraints on local IT policymakers and administrators. By providing a uniform standard that everyone can adopt, it will enable the research and education community to break new ground in the area of statewide cooperation by helping to eliminate IT incompatibilities that make it hard for people in different organizations to pool their resources and combine their efforts. Its key innovation is a new kind of network storage service, based on the *Internet Backplane Protocol (IBP)*, which has been developed at the University of Tennessee's *Center for Information Technology Research* (*CITR*) and is now in use by research and education groups worldwide. The goal of deploying a network storage services based on a common standard is to obtain a level of success beyond what is expected from cyberinfrastructure initiatives that focus on networking alone and which provide neither standards to guide nor resources to assist in the deployment of storage.

The storage component of the OneTenn infrastructure will include a variety of storage resources. The main two types of resources include the following:

- **Large tape archives** placed at major research and education sites throughout the network and providing the entire state with services that cannot be supported at most other locations
- **Major disk-based working storage** at all higher education campuses and smaller caches placed in some schools, libraries or other field locations.
- **Major Linux-based compute systems** at many campus locations for shared access by all institutions.

The rollout of storage and compute resources will proceed from the larger sites where research applications today require an investment in new storage resources and require innovative data-intensive services.

The large tape archive systems, accessible via IBP and directly connected to the OneTenn backbone, provide a uniquely powerful foundation for the OneTenn cyberinfrastructure. Since the entire storage infrastructure uses IBP, users at any site with a disk-based "depot" can use OneTenn's high performance optical networks to transfer data to and from its highly reliable data archives. In this way, widely distributed, disk-based storage resources on each campus can be used at any time to localize, and otherwise optimize, access to data according to the specific requirements for performance throughout the network.

By separating storage, as a generic service, from computation and information processing, the OneTenn approach can provide all OneTenn sites with storage services that were previously available only at large data centers or through complex distributed file systems. The use of optical networking to couple distributed, but less reliable disk resources, to highly stable, but more centralized tape resources, is an architectural advance developed by Tennessee's cyberinfrastructure research community that can now put the state in the forefront of innovation in this critical area.

D. Strategy for Computational Resources under OneTenn

Since computer modeling, simulation, and analysis have become fundamental in all areas of science, engineering, and medical research, increasing the processing power available to researchers is a high priority. There are, however, significant technical challenges to creating interoperability standards for computer processing. As well there are challenges in sharing computing resources across administrative domains. Recognizing that access to computation resources is so critical, the OneTenn plan takes a two-pronged approach to addressing this situation.

First, OneTenn's fabric of high performance networking and distributed storage will enable and facilitate the use of public and non-public computing resources that may be available to some researchers around the state. Examples of such resources (all described in more detail previously) include ORNL's supercomputing facilities, UTK's research infrastructure for grid computing, UoM's planned research computing facility, and the national TeraGrid facility. In addition, distributed cluster compute systems will be installed strategically across the state on the OneTenn network. To ensure that researchers and students can take advantage of these opportunities, OneTenn's small, applied research and development staff will direct some of the efforts to this purpose.

Second, Tennessee already has a cadre of world-class researchers in high performance and grid computing. The UT-Knoxville *Innovative Computing Laboratory* (*ICL*) is an internationally recognized leader in the development of high performance and distributed computing. ICL is actively involved in this area within the DOE laboratory community, the national TeraGrid effort, and DoD defense programs. The University of Tennessee Chattanooga Computational Simulation and Design SIMCenter and the UT Space Institute are world-renowned in computational simulation and design research. Additional research expertise at UT, TBR, and Vanderbilt are detailed later in this report.

E. Campus Impact of OneTenn Infrastructure

Relative to individual campuses, OneTenn will promote cyberinfrastructure improvements in three main areas:

1. **Increased enhancement and utilization of external communications capabilities** — unlike some cyberinfrastructure projects, OneTenn is not based on a "Build it and they will come" philosophy. Significant prototype applications that can leverage OneTenn's high performance networking and data storage technology for multimedia learning content distribution and scientific data management already exist. Consequently, as new eLearning and research applications emerge through the use of OneTenn, all campuses will be able to justify upgrading local network connectivity to take advantage of these new applications. Increased demand for emerging education, research, and workforce development programs will more than justify the investment in higher performance network links for institutions in phases II and III of the plan.

2. **Upgrade local campus cyberinfrastructure** OneTenn storage services on individual campuses can leverage other investments in local technology infrastructure to dramatically enhance benefits to local users. For example, the phased approach for deploying OneTenn services will allow applications to emerge and give campuses both the time and the justification to improve the quality of their campus network and other local cyberinfrastructure. Enhanced campus networks will benefit virtually every aspect of campus academic and business services.
3. **Accelerate growth of collaborative campus application communities —** the tools and capabilities available through OneTenn make applications easy to develop and easy to deploy. Applications for data intensive collaboration and multimedia content distribution already exist and will be deployed over OneTenn. For example, software to enable digital distribution of video over OneTenn is available now and will benefit eLearning programs as soon as OneTenn networking and storage services are deployed. Research communities that already use the same services that OneTenn will deploy have found that new applications, tailored to fit specific needs, are relatively easy to create. Similar results can be expected through OneTenn to support eLearning, public service, health care, and workforce development in Tennessee.

IV. OneTenn Research Leadership

The reason for focusing the OneTenn plan on Tennessee's research community is that, historically, new and advanced information technologies, such as the Internet and the World Wide Web, developed first within academic and research environments. The Web was invented as a tool for information sharing among high energy physicists before it became a pervasive tool for mass communication, business transactions and entertainment. Today we expect information technologies that have been developed through research to rapidly transfer to industry, education and to the general public. Many "experts" in the early 1990's said they "knew" that the emerging "Information Superhighway" was *not* going to be the Internet. The impetus to sustain and accelerate the development of Web software, and the focus on information sharing rather than information protection, was led by the research and education community.

Today it is widely recognized that investment in advanced cyberinfrastructure is necessary to maintain competitiveness. The pace of technological change means that complacency is not an option. Ongoing *exponential* rates of improvement of the basic technologies that underlie information technology (computing, networking and storage) continue. One of the key reasons that research communities take the lead in advanced cyberinfrastructure projects is that, far more than any other segment of society, they confront challenges in their work today which are at a scale that demand the capabilities of the highest capacity, most cutting edge information technologies available. The size of the simulations and the data sets that groundbreaking science use *are also increasing exponentially*. As the applications of Tennessee's research community (described in detail below) make clear, today's research problems **require** the ability to move data thousands of times faster, and store files thousands of times bigger than almost any non-scientific applications. A strong connection among emerging tools, applications and these leading edge user communities is indispensable to ensure that the potential of cyberinfrastructure investments is realized.

Successful leaders for visionary projects like OneTenn typically emerge from the scientific and research community and continue to work with these initiatives to ensure they benefit both the scientific community as well as the wider society. Basing the leadership of OneTenn within the research capabilities of its leading institutions of higher education is essential to achieve the following critical objectives:

- **Harness creative energies for OneTenn development**: The extraordinary scale of big scientific problems forces computational scientists to work at the leading edge of information technology. But what *enables* them to do so is the talent, creativity and expertise of the people who are attracted by the excitement of scientific discovery and highly competitive intellectual environment it generates. This combination of need and talent, supported by the public and private funding it attracts, enables the research community to put new cyberinfrastructure to its most innovative uses.
- **Attract and retain world class research talent**: Perhaps the most direct and critical measure of the success of OneTenn is *the quality of the people it attracts to and retains within Tennessee*. Research leadership for OneTenn is essential in order to *generate* the kind of activity and excitement that makes the state magnetic for the most talented people — scientists, engineers, expert research staff, and students. And it also puts the state's research leaders in a position to *communicate* that excitement as broadly as possible within national and international circles. The OneTenn catalogue of research applications and contributing scientists presented below represents precisely the kind of leadership that OneTenn needs in order to succeed.
- **Engage industry participation leading edge projects**: Major cyberinfrastructure efforts such as SDSC in California, NCSA in Illinois, and MCNC in North Carolina, have proved to be irresistible draws for industry technology partners looking for opportunities to create the next technologies and products. . These kinds of facilities combine ground breaking science and highly motivated research talent in ways that enable companies to develop and demonstrate the newest technologies. OneTenn will offer a similar kind of revolutionary technology environment for Tennessee.
- **Connect Tennessee to emerging cyberinfrastructure initiatives**: The fields of network and high performance computing and data management are undergoing great changes. The Internet architecture and computation center models that have served for decades are no longer seen as adequate to support the variety and scale of demands of academic and commercial research communities. In addition to state and regional infrastructure initiatives such as OneTenn, important research projects that look to new paradigms in cyberinfrastructure are being developed through national and international collaborations to leverage new capabilities. Tennesseans must assume leadership roles in these new national cyberinfrastructure research partnerships and collaborations.

A. The OneTenn Research Community and Its Applications

Combining high performance networking with interoperable storage and deploying it across the state will dramatically enhance the ability of researchers and educators to collaborate across geographic, disciplinary and institutional boundaries. Below are brief descriptions, drawn from accounts provided by the researchers involved, of the research activities and collaborations that are poised to begin using OneTenn as soon as it is deployed. Impressive as this list is, it represents only a fraction of the work at the University of Tennessee, Vanderbilt, and the Tennessee Board of Regents universities that could be called up to support the OneTenn effort. As an example, the *TBR Vision of Excellence Report*[5] calls for a major emphasis on research among TBR institutions. This emphasis recognizes and supports the necessity for the OneTenn infrastructure. In some cases, the researchers are already using OneTenn technologies in some form. In others, the data and compute intensive nature of their work demands the scale of OneTenn resources in order to accelerate the work, elevate them in national and international standing, and make them more competitive for public and private research funding.

[5] http://www.tbr.state.tn.us/vision/research/research.htm

- **Computational Analysis of Gene Expression Data (Rob Williams UTHSC; Michael Langston, UT; Jay Snoddy, ORNL/UTK; Yan Cui at UTHSC)**
  A collaboration in Computational Neuroscience between Dr. Robert W. Williams of UTHSC and Dr. Michael Langston of UTK illustrates not only the importance and power of biologists and computer scientists combining forces today, but also the necessity of advanced cyberinfrastructure when the collaborators are hundreds of miles apart. Working with gene expression data derived from microarrays, these researchers use computational techniques to explore interactions among genes and gene products that would have been heretofore impossible to discover. Dr. Williams and his group collect, filter and prepare the data, which involves roughly five gigabytes per experiment, which is then transferred to the Knoxville campus, where Dr. Michael A. Langston's research team uses high performance algorithms, running on the clusters and supercomputers of ORNL and UTK, to identify groups of genes (cliques) that are highly likely to work together. With the state's current network, *these groups are reduced to hand-carrying compact disks between campuses*. Moreover, future work, which will use higher resolution microarrays and develop much larger and more complex genetic reference populations that will be far better models for human populations, will dramatically increase the amount of data that needs to be moved and stored. Without a statewide infrastructure like OneTenn the progress of this important collaboration, including their ability to compete for new research funding, will be effectively hobbled.
- **Distributed Visualization of 4D Medical Imaging (Zhaohua Ding and Adam Anderson, The Institute of Imaging Science, Vanderbilt; Jian Huang Computer Science UT,)**
  Medical imaging for research, especially in the study of the brain, is not only producing groundbreaking research, it has become one of the primary drivers for improved cyberinfrastructure. For the past three years, the work that Dr. Jian Huang of UTK and Drs. Zhaohua Ding and Adam Anderson of Vanderbilt Institute of Imaging Science have been doing in the area of remote visualization of diffusion tensor imaging datasets of human brain have been making this point. They have been using compute intensive imaging of the brain for studying the parts of the brain that process language signals in order to determine the factors underlying different language disabilities. But the size of these images is already large, and every increase in the resolution increases the size of the image by a factor of eight. With a dozen subjects at high resolution, the data sets will run to tens of gigabytes. Moreover, while the images are collected at Vanderbilt, all the computations are done on clusters at UT under the direction of Dr. Huang and his CS research group. Both the progress and the impact of the tools and techniques that this collaboration is developing will be directly proportional to the quality of the cyberinfrastructure that supports the work.
- **3D quantum-docking molecular modeling tools for Nanodesign (Preston MacDougall, Chemistry, MTSU, and Christopher Henze, NASA Ames Research Center)**
  The groundbreaking molecular visualization tools designed by Dr. Preston MacDougall, chemist at MTSU and Christopher Henze, a NASA visualization specialist, have proven to be both effective as a molecular design research tool, and as a teaching tool for all levels of chemical education. This molecular visualization software is based on a new quantum mechanical understanding of chemical reactivity and allows users to be highly interactive. In essence, the user scans over wide ranges of functions that uniquely highlight topological features that correlate with chemical reactivity, while simultaneously navigating in and around the molecule of interest. Because molecules are 3D objects, and reactive features are often spread over different "shells" of atoms, this visualization technology utilizes "volume rendering." For optimum visual exploration of a medium-sized molecule, 16 Mb files would need to be transferred 10 times per second, with the potential for a 1000-fold increase in

transfer rate if object density is increased.. With these tools and the OneTenn infrastructure, medicinal chemists at universities or research hospitals in Tennessee, and beyond, will be able to engage with Dr. MacDougall in real-time collaborations that explore potential drug molecules. The proposed OneTenn cyberinfrastructure will help keep Tennessee researchers at the forefront of molecular visualization technology, which is critical for nanoscience across a wide range of fields.

- **Integrating High Performance Networking and Mass Archives of Scientific Data (Micah Beck, Computer Science, UTK; Paul Sheldon, Physics, Vanderbilt, Fermi National Laboratory)**

  UTK's Logistical Computing and Internetworking (LoCI) Laboratory, directed by Dr. Micah Beck, is collaborating with Drs. Paul Sheldon and Alan Tackett of Vanderbilt University to use Logistical Networking storage depots to provide access to massive tape archives. This research will make massive persistent storage resources available at an important endpoint of both conventional high performance and optical networks. This allows an interoperable storage service to be created that has the persistence and reliability properties of tape while still making use of the locality characteristics of working disk storage resources distributed throughout the network. By implementing storage facilities at network endpoints which interoperate with these major tape archives, OneTenn will enable storage-based applications to work seamlessly between major research institutions and colleges and schools that are connected via broadband IP and facilitate the use of diverse networks that make up OneTenn as a single statewide infrastructure.

  Vanderbilt's Advanced Computing Center for Research and Education (ACCRE) will establish a multi-petabyte storage facility and deploy hundreds of terabytes of disk for caching and staging of data at locations throughout the campus. This configuration will allow for experimentation, led by Dr. Tackett, in the combined use of storage and disk resources distributed both within a campus and throughout the state under the control of policies determined by application-specific scheduling algorithms. It will provide a level of reliability and determinism not available even using highly redundant disk arrays while still providing high performance access to currently accessed data. This experimental configuration, which will be used in Dr. Sheldon's BTeV effort described below, is a model for the connection of OneTenn's massive, centralized tape and more lightweight, distributed disk resources, as well as future integration with storage depots connected to commodity IP networks and even networks that have only intermittent wide area connectivity.

- **Nanoscale Materials Science (Peter Cummings, Chemical Engineering, Vanderbilt, Director of the Nanomaterials Theory Institute, ORNL)**

  With the addition of Vanderbilt to the UT-Battelle team that manages ORNL, many joint opportunities are now arising built on the strengths of the institutions, particularly related to the computational biosciences and nanosciences. OneTenn will play a crucial role in enabling these collaborations and future growth joint research efforts between Vanderbilt and ORNL. Peter Cummings at Vanderbilt directs the Nanomaterials Theory Institute within the Center for Nanophase Materials Sciences at ORNL, and is part of the team that will build the computational endstation for nanoscience on the National Leadership Computing Facility (NLCF), which is expected to be the world's fastest supercomputer when completed. Likewise, computational biosciences faculty at Vanderbilt are expected to play a key role in the development of the endstation for biology on the NLCF. The OneTenn net offers the possibility to expand these joint efforts to include other institutions in the state of Tennessee. In addition to carrying and storing data, the OneTenn will permit frequent high-quality, high-bandwidth video conferencing between ORNL, Vanderbilt and other Tennessee institutions critical to the development of joint proposals and for the execution of state-wide multi-institutional projects.

There are two major areas (computational biosciences and computational nanosciences) in which Vanderbilt and ORNL are collaborating, for which the OneTenn network would greatly facilitate research between the two institutions:

- *Nanoscale Materials Science*. Peter Cummings leads a 14-investigator, six institutition Department of Energy $4.5m computational nanoscience project seeking to discover computationally new routes to molecular electronics devices and to optimize their performance. Of the 14 investigators, three (including Cummings) are at Vanderbilt, one is at UTK and four are at ORNL. The deployment of leadership-class computing at ORNL, and its availability to researchers at via OneTenn, would dramatically enhance our possibilities for nanoscale materials discovery, both in existing and future collaborative projects with ORNL researchers.

- *Complex Biological Sciences.* Vanderbilt is unique among major biomedical research universities in its ability to pursue individualized medicine. The combination of world-class clinical informatics with world-class research programs in human genetics, proteomics, and structural biology will position Vanderbilt to develop disease treatment and prevention strategies that are tailored to the individual biological features of each patient. Clearly, advancing such efforts will be facilitated dramatically by OneTenn through Vanderbilt's and ORNL's collaborative utilization of leadership-class computing facilities. Specific centers of excellence at Vanderbilt are the Center for Structural Biology, the Center for Mass Spectrometry Research, the Center for Human Genetics Research and the Informatics Center, all of which are world-class research institutes. Already there is a joint Vanderbilt/ORNL nanomedicine center under development, as well as a joint comparative biosystems center. Vanderbilt and ORNL are also jointly hiring a director for the Computational Biology Institute within the Center for Computational Sciences at ORNL.

- **Grid Computing Environments for Research and Education (Jack Dongarra, Computer Science, UTK)**

The emergence of "grid computing," i.e. using the aggregate processing power of computers spread across a network as if it were a unified system, is closely coupled with the creation of the idea of cyberinfrastructure. Put simply, *cyberinfrastructure enables grid computing*. UT's *Innovative Computing Laboratory* (*ICL*), which is internationally recognized as a leader in grid computing, is actively involved in projects within the DOE laboratory community, the national TeraGrid effort, and DOD defense programs. One of its primary technological innovations is *NetSolve*, which is a lightweight environment for research in grid computing that is well established in the grid community. NetSolve takes a straightforward approach to the problem of utilizing grid resources without sacrificing ease of use. It uses a modified "remote procedure call (RPC)" paradigm because this approach preserves the viability of desktop scientific computing environments, such as Matlab and Mathematica, and conserves investment that millions of researchers have made in them. With NetSolve, the technical and time consuming work of installing and configuring advanced or complex mathematical libraries is done once for an entire community of grid users, it frees individual researchers from some of the tedium and frustration that detracts from their proper work. NetSolve provides all the power of grid computing while allowing users to work in standard environments, like Matlab, which have become the *lingua franca* of Computational Science.

- **Multilevel Modeling of Electrical Impulse Propagation in the Heart (Jack Buchanan, Biomedical Engineering, UTHSC; Vasilios Alexiades, Mathematics, UT)**

  Based on advances in biological knowledge and computational techniques, Dr. Jack Buchanan in the Department of Biomedical Engineering at UTHSC and Dr. Vasilios Alexiades, a mathematician at UTK, are developing multi-scale computational models for intercellular and intracellular metabolic processes of the heart. Their approach incorporates realistic physiology and preserves the anatomical architecture of the heart when considered at scales ranging from sub cellular to whole heart.

  But these simulations are becoming increasingly data intensive. They generate data sets up to about 100 MB, which will have to be passed back and forth between Memphis and Knoxville when advanced visualization is required, and proposed extensions to this work involve anatomical studies with imaging at ~ 10 micron resolution, which could increase the size of the data by an order of magnitude or more. But to do this work, the collaborating team must link modest sized computing clusters at each of the performance sites with very high performance networks and storage resources sufficient to handle the logistical issues of data management. With the OneTenn infrastructure in place, this collaboration would be able to further exploit these newer techniques in a way which is efficient and scalable, using grid computing and network storage technology developed by the Innovative Computing Laboratory and LoCI Laboratory at the UT's Center for Information Technology Research.

- **Applications of Optical Networking for Remote Instrument Control -- e.g. remote control of SNS instruments (Carl Halford, Electrical & Computer Engineering, UoM, Jimmy Davidson Weng Kang, A.B. Bonds, Electrical Engineering, Vanderbilt)**

  The Center for Advanced Sensors (University of Memphis, University of Alabama-Huntsville, and Vanderbilt University) addresses the military need for imaging sensors. Military operations in non-traditional modes require imaging devices suitable for small unit operations. The problems of military operations in an urban terrain are well known. Similar new constraints exist in the maritime settings as the Navy operates in port and littoral environments. Furthermore, special operations present unique requirements for imaging sensors. The Center for Advanced Sensors combines expertise in knowledge engineering, nanotechnology, bio-electronics, wavefront control and imaging systems to provide the foundation for the next generation of military imaging devices. OneTenn's high speed network connections across the state and its distributed storage capabilities will enable image and data sharing associated with the research of the Center, e.g. between UoM and Vanderbilt.

- **Bridging Optical Networks and the Internet (Micah Beck, Computer Science, UTK; Nageswara Rao, Computer Science and Mathematics Division, ORNL)**

  Computer Scientists at UTK and ORNL are collaborating to deploy Logistical Networking storage depots as a bridge to switched optical networks that, due to their specialized nature, do not provide traditional IP routing services or connectivity. Using the storage technology that OneTenn will deploy, this effort, which will be prototyped on DOE's UltraScience Net, places massive storage buffers at network hubs where switched and IP networks meet, allowing the unprecedented bandwidth and control channels of switched networks to serve communities that are not yet directly served by circuits. This innovative approach will enable OneTenn's storage-based applications to work seamlessly between major research institutions connected to the switched optical backbone, and colleges and schools that are connected to broadband IP networks. The collaboration will also make the OneTenn infrastructure an important hub for DOE advanced networking activity generally, and the community leading efforts of ORNL in particular.

- **Collaborative Research on Fluid Flow and Heat Transfer Processes in Reduced Gravity (John Hochstein, Jeffrey Marchetta, Mechanical Engineering, UoM; M. Parang, Mechanical Engineering, UTK; B. Antar, Mechanical Engineering, UTSI)**

  The Flow Research Center (FRC) at the University of Memphis develops and uses computational models to study the flow of fluids and heat in reduced gravity environments. This research area is especially compute intensive, and currently the requirements of the UoM team far exceed local compute resources. New research by the Flow Research Center is constrained by the available compute and data-storage resources. As most of the research is transitioning from 2D to 3D modeling, the demands on the computing environment are going up geometrically. Efficient access to massively-parallel computers at remote sites, provided through OneTenn, would remove barriers and significantly accelerate the research. Further, OneTenn would be ideal for moving the large datasets associated with 3D modeling of complex physical phenomena. In addition, the proposed Cyberinfrastructure would significantly enhance collaboration with other researchers across Tennessee and enhance opportunities for collaboration on new research. For example, Drs. M. Parang (UTK) and B. Antar (UTSI), both conduct research on the flow of fluids and heat in a reduced gravity environment. Although there have been repeated attempts to collaborate among these research teams, the five hour (UTSI) and six hour (UTK) transit time between them and Memphis has hindered collaboration. OneTenn would facilitate the use of each other's models and the sharing of large datasets.

- **Experimental High Energy Physics (Paul Sheldon, Physics, Vanderbilt; Tom Handler, Physics, UTK)**

  Vanderbilt plays a leadership role in distributed data storage networks for high energy experimental data analysis, as in the case of the BTeV project at the Fermi National Accelerator Laboratory (Fermilab, near Chicago, IL). OneTenn would enable high speed interconnectivity to the source of this data, namely FermiLab in Chicago. BTeV is one of the first examples of a real-time petascale project: it will require petaflops of computing power to analyze tens of Petabytes of data, and it requires an evolution of the computational data grid into the (soft) real-time domain. At the time that it becomes operational late in this decade, it will be the flagship accelerator based high-energy physics experiment in the United States, with roughly 500 collaborators from all across the globe. To maximize the quality and rate of scientific discovery by BTeV physicists, all must have equal ability to access and analyze the experiment's data. The emerging computational data Grid offers a potential solution to these problems, and particle physicists are contributing significantly to Grid research and development through such projects as the *Grid Physics Network* (GriPhyN), the *International Virtual Data Grid Laboratory* (iVDGL), and the *Particle Physics Data Grid* (PPDG). Vanderbilt faculty are collaborating on the iVDGL project and are leading the Grid development efforts of the BTeV collaboration. Vanderbilt researchers are already working with researchers from Fermilab and from the Duetches Elektonen-Synchrotron Laboratory in Hamburg, Germany to develop storage solutions for the grid that will be beneficial not just for BTeV but for a broad range of application domains. All of the above activities would benefit greatly from OneTenn in terms of data transfer, as well as the planned data depositories. By using OneTenn to combine the strengths of Vanderbilt, ORNL and UT, there is an opportunity to establish Tennessee as the leader in this international research activity.

- **Distributed Data Mining for Exploring Gene Relationships (Michael W. Berry, Computer Science, UT; Ramin Homayouni, Neurology, Anatomy and Neurobiology, UTHSC)**

  DNA microarray experiments often generate lists of hundreds of genes whose expressions are coordinately affected by experimental treatments. To interpret these results, genes are further classified into functional groups using information in GO and biochemical pathway databases. Drs. Berry and Homayuouni have developed an automated method, which we call Semantic Gene Organizer, to functionally cluster genes based on conceptual modeling of the biological information in the titles and abstracts of MEDLINE citations for individual genes. Using techniques that go well beyond classical information retrieval methods, SGO can retrieve and rank the five primary genes in a specific protein signaling pathway with high average precision from a document collection containing 50 genes. A unique feature of SGO is the ability to identify both gene-gene as well as gene-keyword relationships.

  But 50 genes represent a small collection. SGO is scalable enough and fast enough for *genome-wide literature mining*, parsing text collections with over 150,000 documents and 125,000 terms with models based on nearly 300 factors. OneTenn would provide the network and storage capacity necessary to upgrade this cross state collaboration to this much more data intensive level. Moreover, Dr. Berry has already been working with UT's LoCI Lab to explore the potential of using OneTenn technology to upload data to remote IBP depots, which is then used for subsequent query processing. Moving both the data and the query processing into the network in this way would open up a new era in distributed data mining.

- **Neural Networks Analysis of Microarray Data (Jeff Knisley, Karl Joplin, Hugh Miller, The Institute for Quantitative Biology at ETSU)**

  Microarray data is used to investigate gene activation patterns by comparing relative levels of genetic activity for each gene. Even for simple organisms, the result is large data sets which do not lend themselves to standard statistical analysis. The results, using a Drosophila gene set from the Canadian Drosophila Microarray Centre (CDMC) of Toronto, suggest that for the entire genome there may be as many as 1000 genes that are differentially regulated for diapause, a period of suspended growth or physiological activity. Neural networks can be used as a Bayesian framework for the analysis of these large data sets, in particular as a data-mining tool to predict differentially-regulated genes and ultimately genetic pathways. Knisley et al, have had a great deal of success with a method adapted from other research that combines neural networks with a simple genetic algorithm. Unfortunately, the ETSU research has been restricted to a simple single-layer neural network due to the computational requirements of the algorithm (even the single layer network can take up to an entire day to analyze a single data set). Also due to limited resources, they have been forced to consider only subsets of the entire micro-array data structure. The neural network analysis is producing excellent results within these resource limitations, but much more could be done if there was an ability to tap into computational resources statewide (such as at Oak Ridge or UTSI) with these large data sets. OneTenn would allow ETSU to tap into the much needed computational resources that they lack locally and would encourage additional collaborations perhaps currently prevented by that lack of resources.

- **Percolation Model of Phase Transitions in the Central Nervous System during Sensory Information Processing (Robert Kozma, Paul Balister, Bela Bollobas, U of Memphis, Walter Freeman, UC Berkeley, Jim Houk, NorthWestern University, Lee Giles, Penn State University)**

  Kozma et al's research has achieved a breakthrough in the field of description and characterization of phase transitions in brains and brain models. In their brain models with complex interactions and topology, prediction of long-term behavior is not analytically solvable, except for simple cases. Their approach is to run extensive computer simulations of large-scale complex systems, using models arranged in 2-dimensional lattices. Depending of the size and complexity of the mapping function, an experiment may require several days, up to a couple months on an average PC with 2GHz processor. There are thousands of experiments to be run, each of which describes various configurations of brain components and their interactions. In the past 3-4 years, they have conducted hundreds of such experiments using a16-node cluster of parallel computers. As a result, their findings can not be scaled beyond a couple of lattices and interactions between mulitple lattices are what describe inter-cortical effects in brains. It is presently out of reach to conduct hemisphere-wide simulation for brains with about 10-12 lattices, each of which having 10^5 nodes. OneTenn would allow hemisphere-wide simulation, providing the ability to map and visualize a greater volume of lattices and subsequently facilitate collaboration with potential partners at Vanderbilt U/School of Engineering, ISIS (J.Sztipanovits, G. Karsai).

V. Relevant Regional, National and International Initiatives

The OneTenn initiative is also important to the research connectivity needs across the south. In fact, Tennessee is well positioned to serve as the connectivity crossroads of the south. Below are brief descriptions of some national and state initiative that are relevant to the Tennessee OneTenn plan.

A. Internet2

Internet2 is a consortium led by more than 200 US universities, working with industry and government. Internet2 develops and deploys advanced network applications and technologies for research and higher education, accelerating the creation of tomorrow's Internet. Since its inception in 1998, Internet2 has operated a high-speed, low latency research and education network for its members. This network backbone, called Abilene (see Figure7), has been a centerpiece for network and application research for over 6 years. During this time period, Abilene has been upgraded a number of times and in its present form operates at OC-192 (~10 Gbps) rate. Abilene has a large number of national and international peering agreements, spreading this high-quality backbone access beyond US borders.

Figure 7: Abilene Backbone (as of January 2005)

Three Tennessee higher education institutions (University of Memphis, Vanderbilt University, and UT- Knoxville) have participated on the Abilene network from very early on. Over the years, Internet2 has allowed for member universities to sponsor a number of other universities to participate on Abilene. There are currently over 100 individually sponsored institutions in the US. Currently, the following institutions are sponsored individual participants in Tennessee: UT Health Sciences Center, Middle Tennessee State University, Meharry Medical College and St. Jude's Children Research Hospital, and University of Tennessee at Chattanooga.

To broaden even more access to Abilene, Internet2 started the Sponsored Education Group Participant (SEGP) program. Currently, over 35 states (see Figure 8), are working with their Internet2 members in their own states, to provide expanded access to the Abilene network to aggregations of educational organizations brought together in a state-wide or other wide-area network within a state.

As of 2004, more than 10,000 schools (K-12), community colleges, universities, public libraries, museums, zoos, aquariums and science centers across the United States have expanded Internet2 access. As of today, Tennessee has not been able to come together and establish a SEGP. Further, Internet2 has approved the use of the SEGPs for use by medical care providers, researchers and educators. This has significant implications for Tennessee. For example, medical triage and assessment from schools to local/regional medical facilities provide a mechanism to alleviate the shortage of nurses available to care of sick school children. Science and biology education is one of the most expensive curriculums. Electronic learning technologies provide cost effective education to K-20 participants.

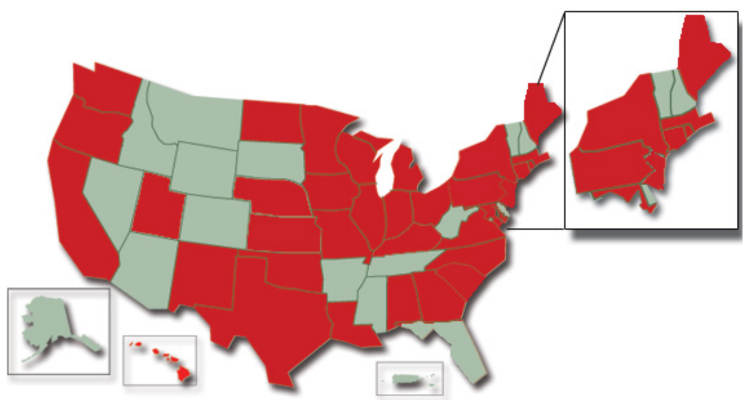

Figure 8: SEGP status (as of 12/ 2004)

The Abilene backbone operates over leased circuits from Qwest. More than ever, the Intenet2 community is moving away from a model of purchasing leased circuits towards a facility-based owned model that includes leasing dark fiber and 'owning' network infrastructure. It is expected that that the next-generation of the Abilene network will go through another transformation in the early 2006; this next-generation network will likely take advantage of the emerging new models and capitalize on the infrastructure being acquired by the National Lambda Rail project. (Internet2 invested $10 mil. into this project).

B. FiberCO tm

Internet2 has established the National Research and Education Fiber Company (FiberCo[6]) to support regional fiber optical networking initiatives dedicated to research and higher education.

The concept behind FiberCo is to help provide inter-city dark fiber to regional optical networks with the benefit of a national-scale contract and aggregate price levels. The responsibility for lighting this fiber rests with the regional networks. A secondary objective is to insure that the U.S. research university collective maintains access to a strategic fiber acquisition capability on the national scale for future initiatives.

FiberCo's extensible allocation includes over 2,600 route miles of dark fiber acquired from Level3 Communications, Inc[7]. Significant portion of this fiber has been already been allocated to Internet2 member universities and the National Lambda Rail.

The assets allocated by FiberCo are expected to enable testing of a wide variety of highly advanced network applications, including uncompressed high-definition television quality video, remote control of scientific instruments such as mountaintop telescopes and electron microscopes, collaboration using immersing virtual reality, and grid computing.

C. SURA – AT&T Fiber Agreement

On December 16, 2003, Southeastern Universities Research Association (SURA[8]) and AT&T[9] announced a collaboration agreement that will allow the nation's research and education community to use AT&T's newest national network infrastructure for experimental work on the next generation of networking technology and applications, called Grid networking[10]. Through this agreement AT&T made available, at no cost to SURA and the nations R&E community, 8,000 miles of dark fiber. Significant part of this fiber has been already allocated for the southern route of the National Lambda Rail project, as well as several regional optical networks.

The University of Tennessee and Vanderbilt University are SURA members, but as there is no next-generation AT&T fiber in Tennessee, Tennessee cannot take advantage of this agreement.

D. Oak Ridge National Laboratory's FutureNET

With the yearly budget of over $1 billion, Oak Ridge National Laboratory[11] is the premier research institution in the southeastern US.

In 1999, ORNL began building the $1.4 billion Spallation Neutron Source (SNS[12]). The SNS is an accelerator-based neutron source being built in Oak Ridge, Tennessee. This one-of-a-kind facility will be completed in 2006 and will provide the most intense pulsed neutron beams in the

[6] http://www.fiberco.org/
[7] http://www.level3.com/
[8] http://www.sura.org/
[9] http://www.att.com/
[10] SURA and AT&T Press Release, December 16, 2003: http://www.sura1.org/2000/news03/ATTPrRel.pdf
[11] http://www.ornl.gov/
[12] http://www.sns.gov/

world for scientific research and industrial development. The SNS is being built by a partnership of six DOE labs.

In addition, ORNL has recently been awarded a $40 million grant to establish the largest super-computing facility in the world.

In the joint effort with the State of Tennessee, ORNL campus hosts three science institutes: Joint Institute for Neutron Sciences (JINS), Joint Institute of Computational Sciences (JICS) and Joint Institute for Biological Sciences (JIBS).

ORNL has several projects pertinent to network infrastructure and computing: UltraNet (UltraScienceNet), Extensible TeraGrid and CHEETAH. As a part of those projects, ORNL is building a large optical infrastructure named FutureNet, which will connect Tennessee with Atlanta, Chicago and the US West coast (see Figure 9). Long-haul fiber for this project has been provided via a partnership between UT-Battelle with the Tennessee Valley Authority (TVA), Qwest Communications.

In the next couple of years, many ORNL instruments will be accessible online to the researchers all over the world. Getting good network connectivity to ORNL is an imperative to take advantage of the great resources being hosted there.

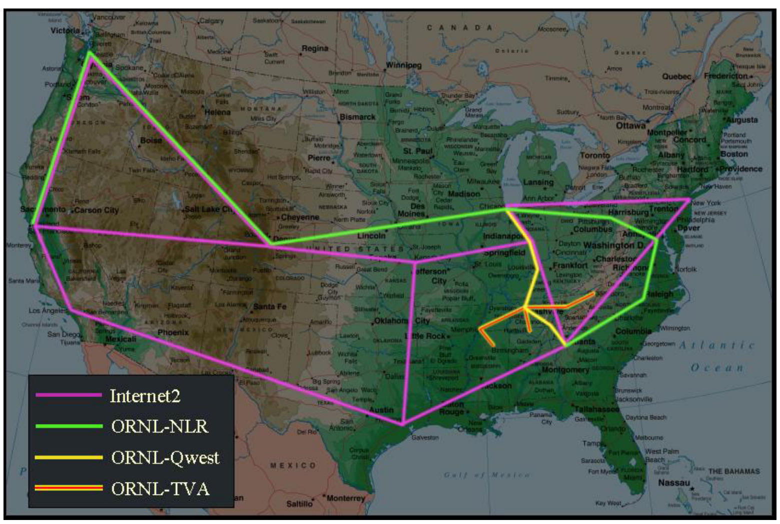

Figure 9: ORNL's FutureNet and other national initiatives

E. National Lambda Rail

National Lambda Rail (NLR) is a major initiative of U.S. research universities and private sector technology companies to provide a national scale infrastructure for research and experimentation in networking technologies and applications. NLR aims to catalyze innovative research and development into next generation network technologies, protocols, services and applications, and also put the control, the power and the promise of experimental network infrastructure in the hands of our nation's scientists and researchers.

The foundation of the NLR infrastructure (see Figure 10) is a dense wave division multiplexing (DWDM) optical footprint with a maximum of 32 and 40 wavelengths (lambdas) each capable of transmitting at least 10 billion bits per second (10 Gbps).

To acquire enabling technologies, NLR has partnered with Cisco. Cisco has provided optical DWDM multiplexers, Ethernet switches and IP routers for deployment of NLR infrastructure. About 10,000 miles of fiber assets have been provided by the Level3 Communications, WilTel Communications[13] and AT&T (via SURA).

The defining characteristic of the NLR infrastructure is its ability to support many distinct networks for the US research community using the same infrastructure. Experimental and production networks exist side-by-side, but are physically and operationally separate. While production networks support cutting-edge applications by providing guarantees levels of reliability, availability and performance, the experimental networks enable the deployment and testing of new networking technologies.

State and regional research and education optical networks have adopted the same concept and are designed in the same fashion as the NLR infrastructure, just at a smaller scale.

Figure 10: National Lambda Rail infrastructure

[13] http://www.wiltel.com

F. Gloriad

GLORIAD project aims to better integrate, with an advanced network infrastructure, the science and education (S&E) communities of the US, Russia and China, together with partners in Korea, Netherlands, Canada, etc.

GLORIAD will encompass the northern hemisphere with a ring structure (see Figure 11.) connecting national science and education networks via connection points in Chicago (US), Amsterdam (Netherlands), Moscow (Russia), Novosibirsk (Russia), Khabarovsk (Russia), Beijing (China), Pusan (Korea), Hong Kong (China) and Seattle (US).

GLORIAD provides much more than a network; it provides a stable, persistent, non-threatening means of facilitating dialog and increased cooperation between three nations who, throughout the 20th Century, have often been at odds. Consequently, they have not benefited from the level of S&E cooperation that each separately has enjoyed with other science communities.

The key scientific applications running over the GLORIAD network include those from communities in astronomy, earth sciences, high energy physics, fusion energy/ITER, network experimentation: Wavelength Disk Drives, atmospheric sciences, grids and computational resources. GLORIAD will also facilitate educational projects such as: Educational Collaboration Infrastructure, the GLORIAD classroom, The "EduCultural" channel, etc.

With the help of FutureNet project by ORNL and OneTenn project, GLORIAD network will reach Tennessee. Tennessee will greatly benefit from the GLORIAD project: not only that researchers in Tennessee will gain access to their colleagues and projects in China, Russia and Korea, but also high-school and college level students will be able to better communicate and work on joint projects with their peers abroad.

Figure 11: *GLORIAD ring around the northern hemisphere*

G. State Initiatives

As it was mentioned earlier, over 20 states in the US have started developing their respective State and Regional Optical Networks mostly lead by the higher education. Here we show examples of a few such infrastructures: Louisiana, Texas and Ohio.

1. Louisiana

The Louisiana Optical Network Initiative (LONI[14]) is a fiber network that provides access to six universities: LSU, LA Tech, Univ. of LA at Lafayette, Southern University, University of New Orleans, and Tulane University.

In March 2004, the Louisiana Board of Regents secured the $5 million National Lambda Rail membership fee by allocating $700,000 per year from its own budget and with $150,000 per year from Tulane University and LSU. In addition, Louisiana took advantage of the SURA-AT&T fiber agreement by acquiring/accepting donated AT&T next-generation fiber. LONI fiber map is shown in Figure 12.

The Louisiana Board of Regents believes that LONI will help Louisiana universities to get access to over one billion of federal research dollars from the National Science Foundation, the U.S. Department of Education and NASA.

With the advocacy of the administration and especially Governor Kathleen Babineaux Blanco, the Senate finance appropriations committee allocated $3.2 million for the LONI project. Finally, during an open forum on LONI in September 2004, Louisiana Governor surprised everyone by announcing her full support for the initiative and allocating $40 million to crate and maintain the network.

Among cited potential applications are grid computing, coastal erosion modeling and complex computer simulations to aid collaborative development of nanotechnology. Several LA universities are already very involved in nanotechnology. It is believed that LONI's potential economic impact for Louisiana is considerable, especially by attracting high-tech industries.

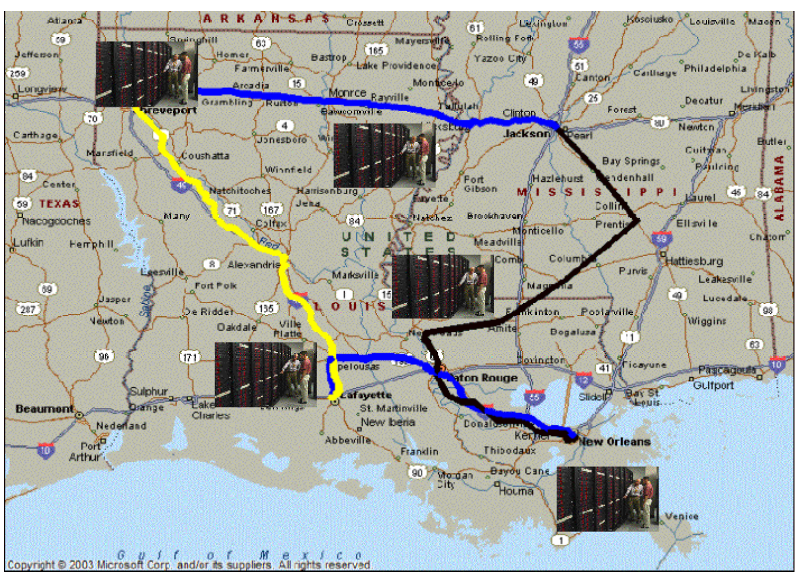

Figure 12: LONI fiber route map

[14] LONI: From Under Dog to Top Dog, http://www.cct.lsu.edu/projects/loni/underdog.php

2. Texas

Texas has a longer history in establishing networking initiatives. Examples are:

- The Grater Austin Area Telecommunications Network (GAATN), a metropolitan-wide information super-highway in Austin, Texas serving Austin Independent School District, Austin Community College, City of Austin, The University of Texas, Austin, etc.
- The North Texas GigaPOP, serves The University of Northern Texas, University of Texas at Dallas, NOAA, Southern Methodist University, Texas Christian University and The University of Texas at Arlington. It also provides SEGP access in northern Texas.
- The Texas GigaPOP provides high-speed network access to The University of Houston, Rice University, Baylor College of Medicine and Texas A&M University, etc.

Seeds of the Texas optical initiative start in late 2002 with a solicitation to establish a National Lambda Rail node in Dallas. Over the next year a new non-profit 501(c)(3) has been formed under the name Lonestar Education and Research Network or LEARN[15].

The mission of the LEARN is to link all institutions of higher education in Texas with state-of-the-art, high-speed data communications that serve the research, teaching, health care, and public service missions of the state's higher education community.

On September 28, 2004, at the opening session of the Internet2 Fall Member Meeting in Austin, Texas Governor Rick Perry and Lieutenant Governor David Dewhurst announced that $9.8 million would be made available from the Texas Enterprise Fund to support higher education and high performance computing efforts in Texas.

LEARN is allocated $7.3 million to build an optical fiber network that will interconnect 14 cities and 25 universities (see Figure 13). The remaining $2.5 million is being allocated to the members of the High Performance Computing Across Texas (HiPCAT) university consortium for the development and deployment of the Texas Internet Grid for Research and Education (TIGRE). The TIGRE initiative will be among the first beneficiaries of the advanced networking capabilities of LEARN.

LEARN currently has 32 members including colleges, universities and medical centers.

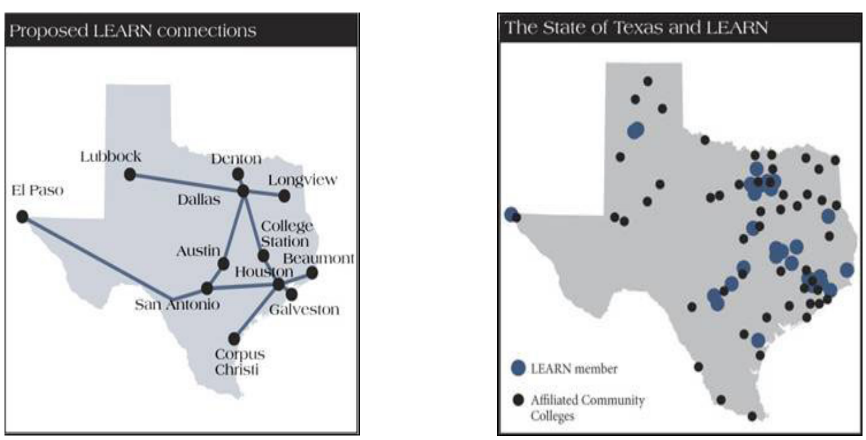

Figure 13: Lonestar Education and Research Network (LEARN) routes and members

[15] http://www.tx-learn.org/

3. Ohio

The Third Frontier Network (TFN[16]) is a dedicated high-speed fiber-optic network linking Ohio colleges and universities with research facilities to promote research and economic development (see Figure 14). Over 1,600 miles of fiber have been purchased to create the network backbone to connect colleges and universities, K-12 schools, and communities together. Nearly one hundred institutions of higher education and thousands of Ohio's primary and secondary schools are being connected in the later part of 2004.

TFN is a technology initiative of the Ohio Board of Regents, and is operated by OARnet[17], the networking division of the Ohio Supercomputer Center (OSC[18]). The Ohio SchoolNet Commission and Ohio Department of Education are also TFN founding members. TFN backbone infrastructure and equipment were purchased with state and federal funds.

The new network's massive increase in capacity will provide revolutionary ways for conducting research in fuel-cell technology, cancer treatment, bioinformatics, DNA mapping, and a host of other applications. TFN will also allow children to take virtual field trips and share classrooms around the state and around the world without ever leaving their schools. It will let doctors see patients across town or across the state without leaving their offices. And it will provide access to OSC's supercomputers in the blink of an eye.

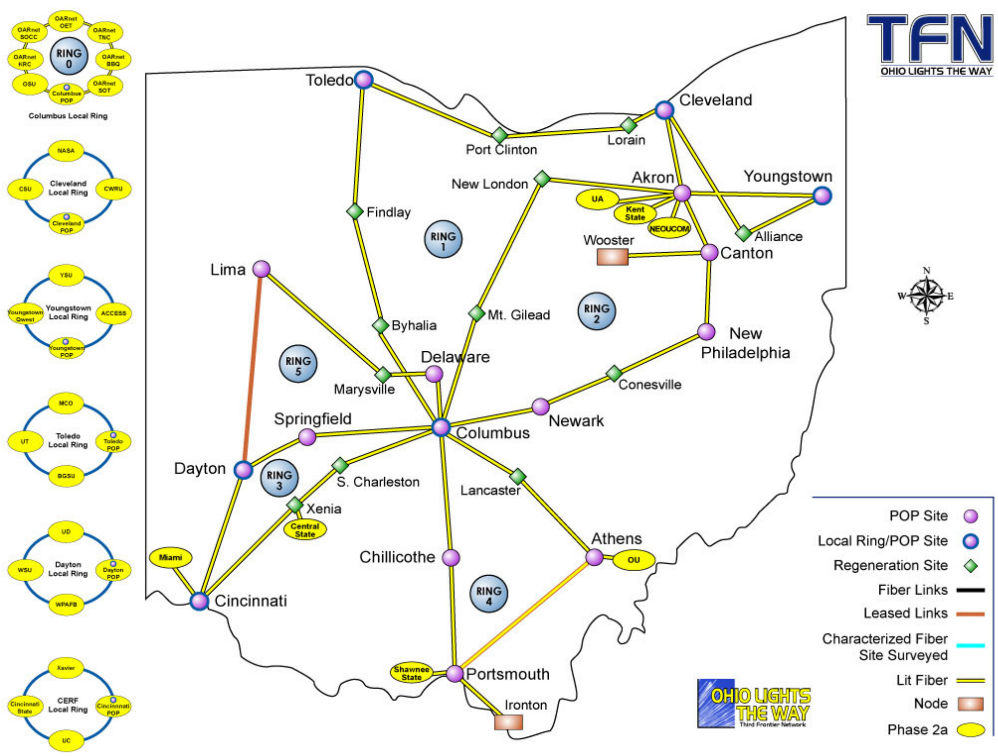

Figure 14: Ohio's Third Frontier Network layout

16 http://www.osc.edu/oarnet/tfn/
17 http://www.osc.edu/oarnet
18 http://www.osc.edu/

The Ohio Board of Regents has committed $19 million for the construction of the TFN and will support its ongoing operation through $3 million in annual funding for OARnet, the networking arm of the Ohio Supercomputer Center. Earlier this year, Congress granted Ohio $5.1 million to support the use of the TFN to improve medical and science education programs and to increase the online sharing of state-of-the-art research instruments. Also, the National Science Foundation made a $400,000 grant to OARnet to support additional academic connections to the network.

VI: Technical Plan

A. Overview of OneTenn Architecture

Scope: The proposed optical backbone design would allow connecting all sites included in the CIC request (see Table 1, page 41. and Figure 15.). While most sites are within a few miles from the backbone, a number of schools will require significant lateral builds to connect to the OneTenn backbone. Even with this increased cost, all sites will be connected within 5 years from project start and no school will be left behind. While some states opted only to connect senior research institutions, the OneTenn plan will include all public colleges and universities in Tennessee. The physical layout of the OneTenn backbone is shown in Figure 16.

OneTenn will be an innovative state-of –the-art all optical backbone network. This deign will enable and support collaboration and provide significant advantages over traditional network architectures. By using optical technology to switch light waves (lambdas), OneTenn can provide a wide range of differentiated services at very low incremental costs. For example, the network can offer production quality services to support distance learning and medical applications, while simultaneously providing very high speed experimental networks to support leading edge scientific, medical, and engineering research.

OneTenn begins with the premise that to provide cost effective yet flexible services and to prepare for future growth needs, the research community in Tennessee needs to own and operate its own fiber network. The effectiveness of this model has been proven over and over again in other states. OneTenn is based on the operation and ownership of approximately 1432 route miles of fiber across the state. This network has the capability to provide data rates up to 70x faster than any other commercial or government network in Tennessee today. Such a network provides the essential infrastructure to support the many leading edge scientific, medical, and engineering applications listed in this report. It also provides the essential link into regional, national and global research networks to enable Tennessee research to keep pace with global innovations.

Site		Network	Storage
1	UT - Knoxville, UT Health Science Center, University of Memphis, Vanderbilt University, ORNL	OC-192 OC-192 OC-192 OC-192 OC-192	
2	UT - Chattanooga, UT Space Institute, Middle Tennessee State University, Tennessee State University, Tennessee Technological University, East Tennessee State University, ----------------------------- UT - Martin Austin Peay State University Pellissippi State Technical Community College, Southwest Tennessee Community College,	OC-192 OC-192 OC-192 OC-192 OC-192 OC-192 ------ OC-192 OC-192 OC-48 OC-48	
3	Chattanooga State Technical Community College, Motlow State Community College, Northeast State Technical Community College, Volunteer State Community College, -------------------------------- Nashville State Technical Community College, Jackson State Community College, Roane State Community College, Walters State Community College, Cleveland State Community College, Columbia State Community College Dyersburg State Community College	OC-48 OC-48 OC-48 OC-48 ------ OC-48 OC-48 OC-48 OC-48 OC-48 OC-48 OC-48	

Table 1: Institutions to be connected to OneTenn backbone

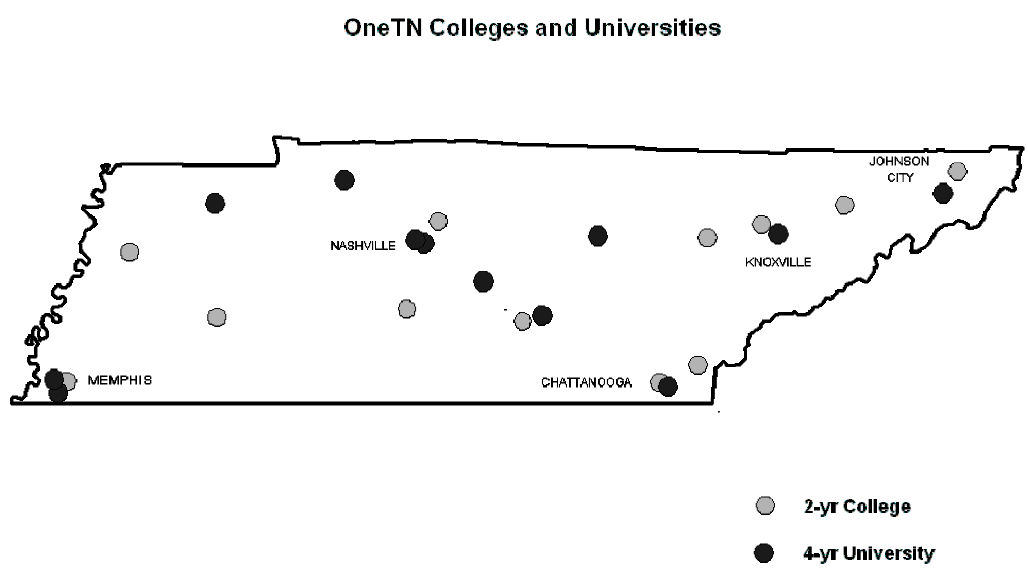

Figure 15: Sites to be connected to the OneTenn backbone

To provide for the switching and routing for the networks operating on the OneTenn backbone, Layer 2 and Layer 3 equipment will be placed in the three major nodes: Knoxville, Nashville and Memphis.

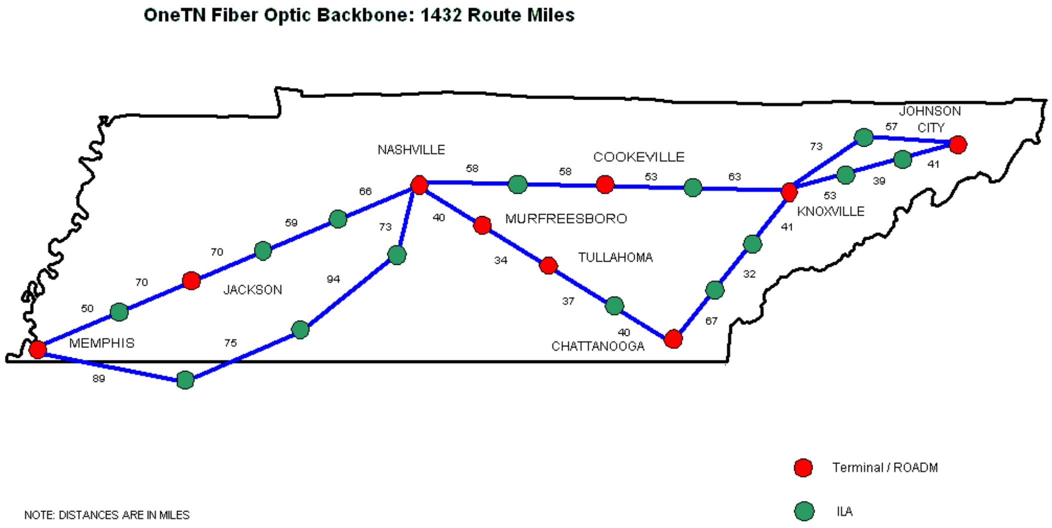

Figure 16: OneTenn backbone

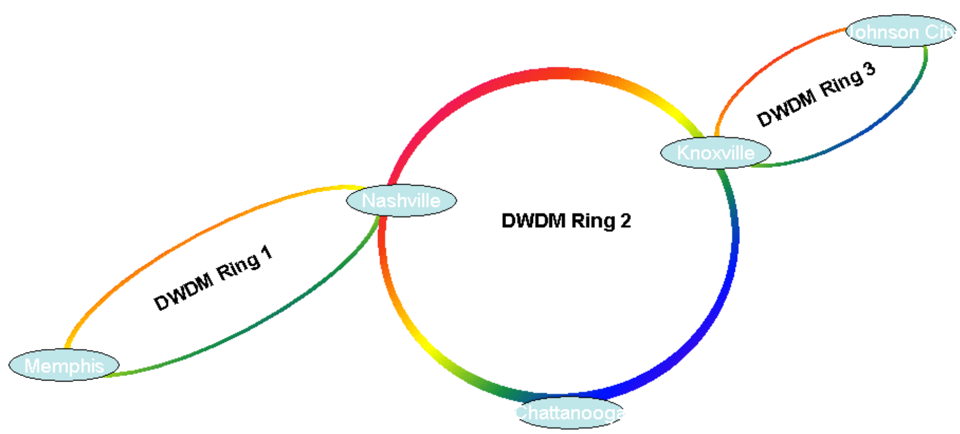

Figure 17: The three ring design of the *OneTenn backbone*

The ring design (see Figure 17) of the OneTenn backbone provides redundant protection against (single) fiber cuts on any route between the three major nodes. Layer2/Layer3 protection mechanisms will be used rather than more expensive SONET protection mechanisms.

Sites will be connected either via OC-48 (2.4 Gbps) or OC-192 (10 Gbps) to the backbone. Initially, all universities will be connected via single OC-192 connection to the backbone, while 2-year colleges will be connected via single OC-48 connection. This exceeds requirements set by the initial CIC charge to connect 25 sites with 1 GigE (1 Gbps) connection. This design was selected because it is more cost effective to establish OC-48 connectivity rather than a 1 GigE connection; an OC-48 connection can be mapped to two GigE channels by a technique called Time Division Multiplexing (TDM), effectively doubling capacity at a lower price.

This network design also offers very inexpensive opportunities for future growth. As research and other needs increase, new connections can be established with minimal hardware investment. This infrastructure will be able to provide simultaneously connections (waves/lambdas) for production services, as well as connections for research. OneTenn will connect to the ORNL's Future Net via the Knoxville and/or Nashville backbone nodes (POPs).

Private schools in Tennessee will be able to connect to OneTenn backbone. These institutions will be responsible for connecting to the backbone, and for a pro-rated share of the capital and operational costs of OneTenn.

1. Optical Network Implementation Phases

OneTenn will be implemented in 3 phases, spread over 5 years, and eventually connecting Vanderbilt University, ORNL and at least 25 public colleges and universities in Tennessee. Other colleges and universities, and research entities such as SJCRH will also be able to connect. During the build phases of OneTenn, sites will be connected in the following order:

- R-1 Research universities and ORNL
- Other universities
- Institutions located in metropolitan areas served by the OneTenn backbone
- Most 2 year colleges
- Hard-to-reach (in terms of fiber availability) institutions

a. PHASE 1

Acquire the network backbone and establish operational service between the 3 major research sites. All long-haul fiber will be acquired to establish the OneTenn network, including the main backbone between Knoxville, Nashville, and Memphis (University of Memphis and UT Health Science Center). OneTenn between the three main network POP sites (Nashville, Memphis and Knoxville) will be fully equipped and operational.

Institution	City/Location	Gateway
Vanderbilt University	Nashville	Nashville
University of Tennessee, Knoxville	Knoxville	Knoxville
University of Memphis	Memphis	Memphis
UT Health Science Center	Memphis	Memphis

Table 2: Total number of sites connected by the end of Phase 1: 4

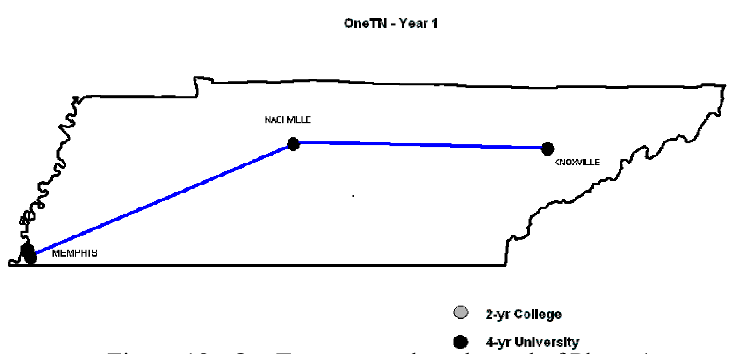

Figure 18: OneTenn network at the end of Phase 1

b. PHASE 2

Establish operational services to the remaining universities and to four community colleges. This includes the link to Johnson city, the second pair of fiber between Memphis and Nashville and the link between Chattanooga and Knoxville. Two major backbone links (to Martin and to Clarksville) will be established. Four 2-year colleges in metropolitan areas will also be connected. Connections to Atlanta and Huntsville will be established to link OneTenn to regional and national high speed research networks. All universities will be connected by the end of Phase 2.

Institution	City/Location	Gateway
University of Tennessee – Chattanooga	Chattanooga	Knoxville
UT Space Institute	Tullahoma	Nashville
Middle Tennessee State University	Murfreesboro	Nashville
Tennessee State University	Nashville	Nashville
East Tennessee State University	Johnson City	Knoxville
Tennessee Technological University	Cookeville	Nashville
University of Tennessee – Martin	Martin	Memphis
Austin Peay State University	Clarksville	Nashville
Pellissippi State TCC	Knoxville	Knoxville
Southwest TCC - Macon Cove	Memphis	Memphis
Jackson State Community College	Jackson	Memphis
Nashville State Tech. Comm. College	Nashville	Nashville

Table 3: Total number of sites connected by the end of Phase 2: 10

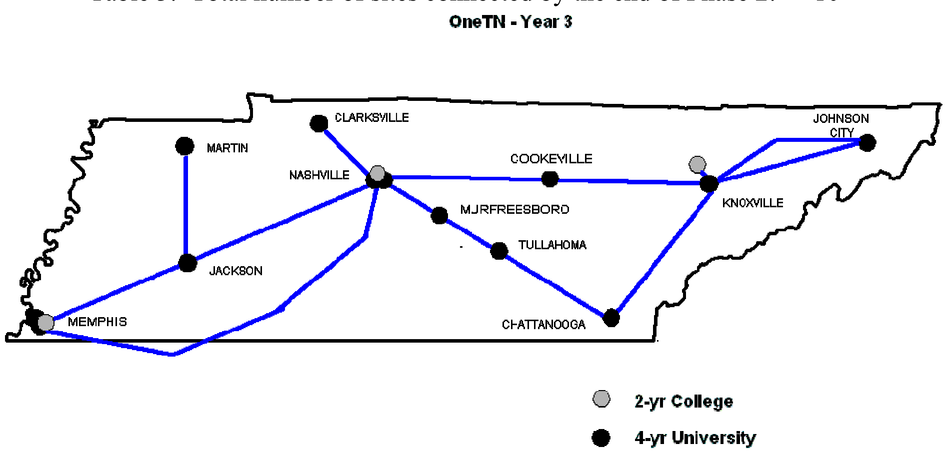

Figure 19: OneTenn network layout by the end of Phase 2

C. PHASE 3

Establish operational services to community colleges. Four additional 2-year community colleges will be connected to the OneTenn backbone.

Institution	City/Location	Gateway
Chattanooga State TCC	Chattanooga	Knoxville
Motlow State Community College	Tullahoma	Nashville
Northeast State TCC	Blountville	Knoxville
Volunteer State Community College	Gallatin	Nashville
Roane State Community College	Harriman	Knoxville
Walters State Community College	Morristown	Knoxville
Cleveland State Community College	Cleveland	Knoxville
Columbia State Community College	Columbia	Nashville
Dyersburg State Community College	Dyersburg	Memphis

Table 4: Total number of sites connected by the end of Phase 3: 25

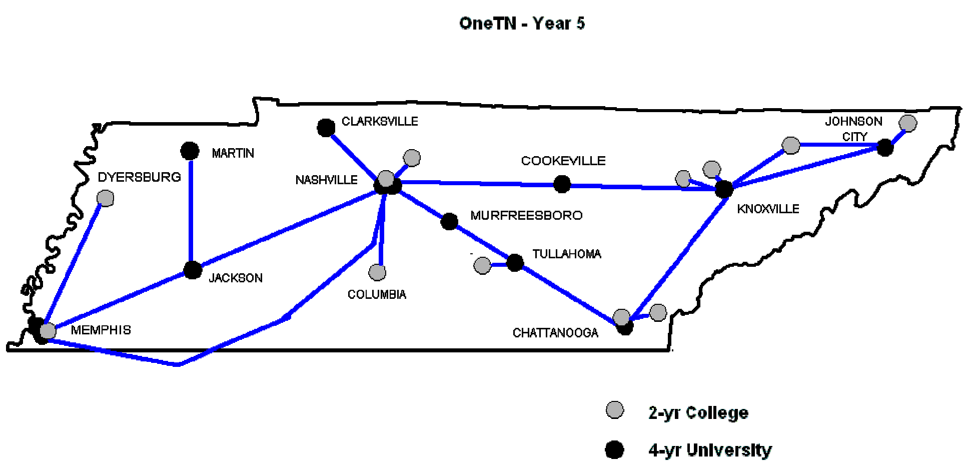

Figure 20. OneTenn network layout by the end of Phase 3

2. Storage and Computation

In addition to the networking infrastructure described above, OneTenn will provision storage and computing. The computing nodes, which will be densely packed clusters of processing power, networking and storage, are not intended to offer general purpose computing services to OneTenn users, but rather to provide the processing power necessary to provide high performance buffer services for traffic moving between optical circuits and traditional high performance IP networks. They can be described as optical network buffer processing units. Taken together, these resources will help enable users to take advantage of the high-capacity network as well as provide new services within the local institutions.

The R-1 Research universities, which correspond with the network POPs in Knoxville, Nashville and Memphis, will form the major storage and computing backbone. These sites will have storage systems, archival (long-term) storage systems, and computing clusters for high performance buffering services. All universities will have storage and compute clusters. All two-year colleges will have storage systems. OneTenn will also provision some test storage systems at select P-12 and tech centers.

Institution	Storage Systems	Computer Clusters	Archive Systems
R-1	12 (4 ea)	13 (4 ea*)	15 (5 ea)
Four Year (non R-1)	36 (4 ea)	9 (1 ea)	0
Two Year	52 (4 ea)	0	0
Select P-12/Tech Center	25 (1 ea)	0	0

*UTK will have five compute clusters.

Table 5: Total Storage and Computation Infrastructure

a. Infrastructure Goals

This plan meets several goals:

- Every higher education institution receives storage systems;
- Every university receives a computation cluster;
- The R-1 universities, which are the network POPs, also house the archival storage systems;
- Systems are provisioned for select P-12 and tech centers testing even if they are not directly connected to the OneTenn backbone; and
- All universities have access to all 400+ servers working as a state-wide computing platform.

b. Rollout Plan

- *Year 1*

  Each R-1 site gets a storage system and a cluster farm. In addition, Knoxville will get an archival storage system. All remaining universities get a storage system.

- *Year 2*

  Nashville (Vanderbilt) and Memphis add an archival storage system. Each two-year college receives a storage system and TTU and UT-Martin receive a compute cluster.

- *Year 3*

  Each R-1 site adds a second archival storage system. The remaining universities receive compute clusters. Each two year and four year institution receives a second storage system. Twenty-five select P-12 and tech centers will deploy storage systems on non-backbone connected sites.

- *Year 4*

  Each R-1 site adds two clusters and a third archival storage system. Each university adds two more storage systems.

- *Year 5*

  Each R-1 site adds another cluster (two at UTK) and two archival storage systems. Each two-year institution adds two more storage systems.

VII. Budget Summary

A budget estimate is provided below that describes both the capital and operating costs for the OneTenn infrastructure. This five–year estimate includes all costs associated with procurement, operation, and management of OneTenn as delivered to all 23 university and college locations. Annual operating costs after year 5 are not detailed in this report but are estimated to no greater than the current annual expenditures for UT and TBR wide-area connectivity.

OneTenn Projected Budget						
		Year 1	Year 2	Year 3	Year 4	Year 5
Network - capital						
	equipment	4,655,000	3,465,000	1,970,000	200,000	250,000
	fiber	4,388,000	0	0	0	0
	construction, installation, testing	205,000	312,000	245,000	100,000	125,000
	Subtotal	$9,248,000	$3,777,000	$2,215,000	$300,000	$375,000
Storage - capital						
	Depot equipment	450,000	397,500	1,479,000	900,000	795,000
	Archive equipment	188,000	376,000	564,000	564,000	1,128,000
	Subtotal	$638,000	$773,500	$2,043,000	$1,464,000	$1,923,000
TOTAL Capital		$9,886,000	$4,550,500	$4,258,000	$1,764,000	$2,298,000
Network - recurring						
	Metro fiber loops	125,000	241,000	490,000	634,000	1,133,000
	Operation and maintenance	479,000	1,085,000	1,360,000	1,360,000	1,360,000
	NLR	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
	Personnel	360,000	360,000	360,000	360,000	360,000
	Subtotal	$1,964,000	$2,686,000	$3,210,000	$3,354,000	$3,853,000
Storage - recurring						
	Personnel	325,000	325,000	350,000	350,000	375,000
	Operation and maintenance	25,000	25,000	50,000	50,000	50,000
	Subtotal	350,000	350,000	400,000	400,000	425,000
TOTAL Recurring		$2,314,000	$3,036,000	$3,610,000	$3,754,000	$4,278,000
TOTAL Costs/year		**$12,200,000**	**$7,586,500**	**$7,868,000**	**$5,518,000**	**$6,576,000**

Table 6: One Tenn Projected Budget