

“Innovation, Transformation, Community”

Remarks of Michael A. McRobbie
President, Indiana University
Spring 2012 Internet2 Member Meeting
Plenary Session
Arlington, VA
Arlington Crystal Gateway Marriott
Salon I/II/III
Monday, April 23, 2012
6:00 p.m.

1. INTRODUCTION

Let me first, on behalf of the Internet2 board and staff, welcome you to the Spring 2012 Internet2 Member Meeting. We're grateful that you've made the journey to join us and that you are part of the Internet2 Community.

Hundreds of IT leaders from universities, research groups, government, and industry from around the United States and around the world have come together under the umbrella of Internet2 to share, learn, and collaborate.

While many of you are not in academia, all of us hold a stake in higher education. Especially as we see industry increasingly retreating from pursuing long-term fundamental research, our universities become even more vital to innovation in science, engineering, health and medicine, business, the social sciences, the arts, public policy and national defense.

And we must always remember that our universities also created the Internet, an innovation of enormous power that has transformed world economies, business models, and information transfer in a way unprecedented since the Industrial Revolution.

In 1996, representatives from 34 universities met in Chicago to create a new organization that would meet the national networking needs and provide an innovation platform for higher education and related research organizations. This led to the creation in 1997 of Internet2, an organization with which I am proud to have been associated from its inception. Indiana University was a founding member of Internet2 and, ever since, has implemented, managed and operated the successive Internet2 networks through its Global Network Operations Center.

Now, once again, we as a community have been provided with an opportunity to transform the future through an innovative technology platform. Last year, Internet2 received a \$62.5 million federal stimulus grant from the National Telecommunications and Information Administration's Broadband Technology Opportunities Program, known as BTOP. The total project value is approximately \$97 million.

The grant is funding a dramatic expansion of the Internet2 network infrastructure and capacity. The upgraded network will have an unprecedented 8.8 Terabits of capacity and use brand new 100 Gigabit Ethernet technology. The Internet2 Network is the first national network to deploy 100 GigE waves on its entire footprint, and will become the most sophisticated research and education platform in the world.

In addition, the Obama Administration recently announced the "Big Data Research and Development Initiative" in which several federal agencies have made commitments to develop new technologies to manipulate and manage large quantities of data and to use those technologies in science, national security, and education.

The application of these new technologies in education has the potential to accelerate the pace of discovery in research and to transform teaching and learning. Areas with great potential benefits for global health—such as systems biology—have enormous network capacity requirements.

These kinds of projects and other new applications will dwarf what are currently state-of-the-art networks, thus new capacity will be required for innovation.

Hence, we have the opportunity to use these investments to help transform the way we deliver research and education—and perhaps change the world again.

But how might this be done?

As a university president, I try to view everything through the prism of what have been the fundamental missions of higher education from the earliest days of the most ancient universities.

These missions are:

- The creation of knowledge (research & innovation)
- The dissemination of knowledge (education & learning)
- The preservation of knowledge (information repositories)

In what follows, then, I am going to talk about three examples of where there is enormous potential for the Internet2 community, and its partner communities internationally, to both contribute to and enable these missions in visionary and transformative ways.

2. SOFTWARE DEFINED NETWORKING – THE CREATION OF KNOWLEDGE

Research in all areas of academia from anthropology to zoology has been fueled from the earliest days of the Internet by applications that have made innovative use of ready and ubiquitous access to bandwidth in higher education—relative, of course, to the most advanced technology of the day. One of the key things we have done, and must continue to do, is to lead a change in thinking from a concern with bandwidth scarcity to embracing abundant bandwidth availability. New applications do not happen without the

capacity for innovation. We must continue to provide as much end-to-end bandwidth in all areas of our networks as is economically and technically feasible.

The leading and game changing applications continue to originate on university campuses and in research labs. The Web browser, Google, Facebook, peer-to-peer networking and thousands of other applications that have changed society were born on our campuses. These innovations happened because of the ubiquitous end-to-end capabilities we provided for innovation, and we must continue to commit to bandwidth abundance.

We are also seeing the beginnings of potentially another major paradigm shift in networking with open networking. The research and innovation leading to this has once again arisen in universities, and Internet2 is one of the leaders in the deployment and use of this technology. The Internet was created, in part, by banishing proprietary approaches and moving to open standards like Ethernet and TCP/IP. By building open networks and applications, we are potentially creating a whole new market for advanced communications. Our investment on a national scale will provide major new opportunities to create scalable applications. We continue to move the world from proprietary to open networking, but now is the time to accelerate our commitment.

To catalyze these opportunities, our researchers, scientists, and educators need us to once again create a new network playing field that is at the limits of the available technology. The new 100 Gigabit Network is an example of the dramatic increase in capacity that is needed. But capacity in this next era must also be coupled with the latest ideas in open networking.

Adding Open Software-defined Network capabilities to networks with bandwidth abundance will create an Innovation Platform that will help launch the next set of global innovations.

Led originally by Stanford University, and now by many other academic and commercial leaders, new thinking to open the network layer to innovation through SDN is vigorously underway. We are beginning to virtualize hardware to allow innovators to share physical platforms with alternative software.

SDN is a way of thinking about network hardware as just one more programmable component of IT infrastructure. It holds the promise of reinventing transport protocols, economics, application awareness, security, and other aspects of networking.

As evidenced by last week's announcement by Google that it has quietly moved its entire internal global infrastructure onto an SDN-based Openflow network, this is an innovation opportunity that is moving quickly.

Even with industry quickly investing in Openflow and SDN, our community has a pivotal role to play in deploying SDN in our unique environment. The early innovations in SDN are in single enterprises. Our community has the opportunity to create the first multi-domain SDN-enabled network open for all in higher education and elsewhere to use for innovation. Internet2 and the Global Network Operations Center at Indiana University have deployed a test bed of Openflow enabled hardware and are now working with the community on details for a full network wide deployment.

True end-to-end deployment is essential—at the campus, regional, national and ultimately global level. The National Science Foundation and GENI are now investing in the campus/regional layer. The NSF has long been a key collaborator in growing the Internet from its infancy. Today, their GENI program continues to enable innovations like Openflow and concepts like SDN that are at the cutting edge in terms of network research. Internet2, our regional partners, and our university campuses must expand and deepen our collaboration with the NSF & GENI so that early innovations can be deployed and nurtured on our campuses just as TCP/IP and Ethernet were nurtured in an earlier era.

Together with other investments from campuses and the government in the regionals and Internet2, this positions the research and education community once again as the place where next generation networking will flourish. And ultimately, in collaboration with our global partners, we expect these innovations will span the globe.

The new research and education-based Software Defined Network platform—open, at scale, and with abundant bandwidth—is, then, the next innovation platform. It provides the best environment, as it has in the past, from which new network models and applications will emerge.

This is a remarkable opportunity which this community must embrace while also bringing objectivity, a critical eye, and the years of experience at the frontiers of networking to bear in this area.

3. BUILDING A HIGH-SPEED PRODUCTION GLOBAL R&E NETWORK – THE DISSEMINATION OF KNOWLEDGE

I started the previous section by noting that there is no area of academia that has not been affected by the rise of the Internet. At the same time, and not unrelated, there is no area of education that has not been affected by internationalization. It is true of research because the Internet has dissolved the boundaries of space and allowed it to become truly international, but it is also true of education where the international dimension of education in almost any field has become essential.

As education becomes truly international with degrees requiring some international component, with the rise of 2+2 and similar degrees, with global collaborative courseware platforms, with instruction becoming multilateral and virtual—and with all of this fuelled, in part, with ubiquitous very high quality video conferencing and telepresence technologies, a high-speed production global research and education network is becoming absolutely essential.

The research and education interests of American research universities are truly global, as are the needs of many other research universities elsewhere in the world, and the need to support end-to-end applications is no longer a requirement that stops at the ocean's edge. For example:

- Large-scale scientific teams are now organized at the global level, using massively expensive instruments that are funded and operated multi-nationally.
- American research universities are building entire campuses and other facilities across the globe and expect to be able to routinely use high bandwidth telepresence and video conferencing applications to support their operation.
- Globally distributed virtual research organizations require a coherent and integrated set of technology tools and applications to work effectively.

There have been a number of such high-speed production research and education networks proposed in the past—Steve Wallace and I proposed one called the Global Terabit Research Network about 10 years ago, an idea that was probably ahead of its time. And there has been much activity in this area, gradually growing in scale, complexity, and reliability. Some of the components of such a network exist at the moment allowing reasonable connectivity, for example, between the US, Europe, and parts of Asia, due to the vision and enormous hard work of many National Research and Education Network leaders world-wide. But much of it is dependent on the vagaries of agency funding and it lacks the characteristics of a true production network.

I believe with the dawn of this era of true network abundance such as Internet2's new 100 Gigabit network, with Terabit networking just over the horizon, with the promise of the paradigm change brought about by SDN, and with the increasingly vital importance of internationalization in education and research, it is time to renew our efforts to build a true high-speed stable long-term production research and education network.

I am delighted, by the way, to note the large number of international participants at this meeting. I understand there are over 100 of our international colleagues who have joined us from across the world.

As we have migrated our national and continental networks to dense wavelength division multiplexing-based services built on dark fiber IRUs capable of supporting scalable bandwidth and multiple layer 2 & 3 services, utilizing open exchange points, we will need to extend that model to our inter-continental connection fabric.

Achieving this will require innovative approaches to the way we organize and fund these efforts. While the dedicated efforts of scores of NREN leaders have created a global fabric that has met our “first generation” needs; these new “second generation” demands will require a significantly more systematic and intentional approach to the architecture of global infrastructure—an infrastructure that will need to provide a consistent and seamless advanced set of services, born from a fully integrated set of components, and operating within a common policy.

This is, in essence, a transformational challenge to the global NREN community and will require all our best efforts and dedication.

4. THE DIGITAL PRESERVATION NETWORK – THE PRESERVATION OF KNOWLEDGE

The third example I want to discuss is of less immediacy than the first two, but in some ways even more important. And this is the preservation of digital data. Digital technology has enabled unprecedented growth of knowledge in essentially all areas of scholarly activity. This knowledge, however, is inherently vulnerable and the academy has been slow to recognize or deal with the problem. In fact, it is no exaggeration to say that there is a looming crisis in this area. Though there are some efforts to systematically preserve digital data, they tend to focus on one area, be reliant again on

the vagaries of foundation founding, and be vulnerable to political and social change. There is no systematic strategy in place anywhere aimed at the long-term preservation of digital data not just for tens of years, but for hundreds of years.

Let me give just two examples of where digital data is in peril.

First: the Sloan Digital Sky Survey. Begun in 2000, this took 8 years to complete, covered 25% of the sky, mapped 930,000 galaxies, released more than 100 Terabytes of data to the scientific community, and has resulted in more than 2,000 articles and 70,000 citations to date. This is extraordinarily valuable data by any measure, as is other data from similar projects. In fact, the Space Telescope Science Institute now reports that more papers are published with archived data sets than with newly acquired data.

But the odds that they will remain available to future generations are tenuous enough to make everyone uncomfortable. In 2008, the University of Chicago Library entered into a formal agreement with Astrophysical Research Consortium to assume responsibility for archiving the Sloan Digital Sky Survey data. While that was clearly a positive step, the library funding for these preservation efforts expires in 2013. Data that took 8 years to collect and that has a scientific value measured in decades has a preservation horizon that expires just next year. While the Sloan Foundation may well see fit to continue funding the initiative beyond 2013, there is every possibility that it might not.

As a second and completely different example, consider the new Alexandria Library opened in 2002 and funded through various Egyptian government and international sources. This is a wonderful and audacious project to re-establish the great Alexandria library of antiquity but with some digital holdings being central. It is based in a magnificent new building in Alexandria, Egypt. It was, for a time, the only place that held the back-up copy of the Internet Archives and it has many other rare and magnificent collections.

But it has been controversial with some fundamentalist groups seeing it as a symbol of modernity and secularism, and the library and its director have been subject to threats. During the Arab Spring in early 2011, concern about the physical safety of the library was such that it was ringed by students to protect it from the sort of looting and destruction seen at some other Egyptian museums. Its future will have to be of concern. So this is an example of digital data being in peril due to political and religious forces.

These two examples highlight the fact that while digital collections proliferate at network speed, they are typically not durable and remain susceptible to multiple single points of failure. Moreover, the emphasis in building these collections tends to be more on providing access to current users rather than on preserving them for the future. Absent focused and coordinated effort, much of today's scholarship will be lost to future generations.

Let me digress, then, to consider lessons from the past that apply to the preservation of today's digital information. Consider any of the great works of literature, history or philosophy from the ancient Greeks or Romans. No original manuscripts from the period of their composition have survived, but a surprisingly large amount of the most significant works have survived in spite of all the vicissitudes and calamities both natural and man-made that have befallen those parts of the world since.

I submit there are two reasons for this. First, such works were constantly copied—by hand—for hundreds of years so that tens of thousands of the major works were distributed around the Mediterranean. Though the originals were gradually destroyed or wore out they were, in turn, recopied, though in decreasing numbers during the Dark Ages. Second, these precious copies were held in institutions that were part of not only the most powerful and prestigious institution in that part of the world for centuries but one that is the longest surviving human institution—the Catholic Church.

So what lessons do we draw from this for the very long-term preservation of digital data? There are two.

First, ensure that there are multiple copies of major digital data repositories geographically and politically distributed, ultimately globally.

Second, associate these copies with powerful and prestigious institutions that have the greatest chance of surviving into future centuries. And I would contend that these are universities. Universities are the longest-lived human institutions apart from the Catholic Church. The great medieval universities of Oxford and Bologna date from the 11th Century. Al Azhar University in Egypt dates from the 10th Century. The legendary Nalanda University in India survived 17 centuries before its destruction by the Mughals in the 12th Century. And Nanjing University in China claims to have had unbroken existence since the 3rd century BC! This would make it older than the Catholic Church, though this claim is controversial.

In spite of the present frenzy of enthusiasm for “clouds,” I would claim universities are a better bet for the long-term preservation of digital data than IT companies. I remember hearing John Chambers, the Chairman and CEO of Cisco (and an IU alum, I am proud to say), remarking at a speech a few years ago, that of the 10 networking companies that had existed 5 years before, only 3 were still in business. There is a famous quote by Clark Kerr, the legendary president of the University of California, who noted that of the 85 human institutions founded by the 16th Century still in existence today, 70 were universities.

The Digital Preservation Network (DPN) is an initiative that is aimed at a systematic approach to the long-term preservation of digital data. Fundamental to it is replication and the ownership by a consortia of universities. It seeks to build upon the higher education community’s current efforts to build a federated preservation network, owned by and for the academy, which will provide secure digital preservation of the scholarly and cultural record for centuries.

Last week, together with James Hilton, the CIO at the University of Virginia and an Internet2 Board member, and Ann Wolper, the librarian at MIT, I gave a presentation on digital preservation and DPN to the Spring meeting of the AAU presidents. It was received with great interest. Due to the efforts of James and others, over 50 universities have now committed to the initial exploratory phase of DPN.

At the heart of DPN is a commitment to replicate the data and metadata of research and scholarship across diverse software architectures, organizational structures, geographic regions, and political environments. Replication diversity, combined with succession rights management, will ensure that future generations have access to today's discoveries and insights.

An initial implementation of DPN would have 3 major storage nodes geographically distributed around the nation with as much diversity of software and hardware as possible for resilience acting as front doors to different types of digital data—for example, text, rich media, and large scientific data sets. But then, each would replicate the data of each of the other nodes so that each node would have a full copy of all the data. Crucial for what would be the routine replication of petabytes of data across these three initial nodes would, of course, be Internet2. And clearly, if DPN is eventually to be expanded overseas, which is consistent with its philosophy of maximum diversity, then international connectivity on a par with the NREN backbone speeds is essential.

While the Digital Preservation Network is not an Internet2 initiative—its success will require independent governance that can survive over time—it is an initiative then to which Internet2 will be an essential contributor.

5. CONCLUSION

We are still in the early stages of exploring the full potential of Internet technology and its importance to the research and education community, commerce and business,

healthcare and science, arts and humanities, and beyond. Our members believe that networking capabilities should be shared as freely as possible across disciplinary limits, political boundaries and economic divisions.

We are working with our industry partners to explore new ways to utilize this new capacity to build new environments that are suited to our needs. Our NET+ partners recognize that our organizations can help them develop new models that will address our needs for big data, research, education, and administration.

By working together, we can leverage many services and use these tools in innovative ways to transform our campuses.

So let us all take on this important work that will form the foundation for future economies, science, health care, education, and governance, not only in the U.S., but around the globe.

Thank you.