

TIP2013

APAN 35 Conference and Winter 2013 ESCC/Internet2 Joint Techs Meeting

Remarks of Michael A. McRobbie

President, Indiana University

Hawaii Imin International Conference Center

Keoni Auditorium

Honolulu, Hawaii

Monday, January 14, 2013

9:15 a.m.

1. INTRODUCTION

Thank you very much, Dave [Lambert], for that kind introduction. I really am pleased to be back among this community with so many old friends and with so many people whose accomplishments I have long admired.

I was, of course, a member of this community for many years and gave many networking and IT talks to meetings all over the world that, of course, reflected the perspective of a CIO, as I then was. But my perspective on IT and networking as president of a university with a \$3 billion budget, 110,000 students and 20,000 faculty and staff, as well as a \$4 billion hospital system, has, frankly, evolved significantly. Everything we do in IT and networking must always be seen through the prism of the university's fundamental missions of education and research. Some of the issues that once seemed of enormous importance, that filled me and, I know, many people in this room, with passionate intensity, seem of far less import when one is struggling with huge budget issues caused by government disinvestment in higher education, with an increasing mood of public skepticism about the value of education, with catastrophic

potential risk from athletics to infrastructure, and with existential threats to our very existence as institutions.

So, in this address, I want to discuss what I think, from a presidential perspective, are two major issues of fundamental importance to the broad enterprise of higher education and research on which I believe this community, as an international community, should focus. These are the issues of stable global networking—a matter of immediate importance—and very long term digital preservation—a matter of maybe less obvious importance but of vital importance to the academic enterprise.

Since I was invited to be the plenary speaker for the APAN portion of this meeting, I will focus some of these comments on APAN—the Asia Pacific Advanced Network—of which I am proud to have been a founder, and suggest the direction in which it needs to consider moving, given these issues.

In doing so, I want to try to take what is maybe a unique perspective that I hope may be of some value, of someone who has been a member of this community but who now has responsibility for all aspects of a large higher education institution. This perspective also reflects, I believe, the broad sentiment of the Internet2 Board, which I chair, and which provides presidential engagement and oversight to the organization.

2. APAN PRE-HISTORY AND HISTORY

I will start by reviewing a little of the pre-history and early history of APAN at the time I was involved in it. In so doing, I will mention a debate that took place in these early days which seems to me—now just an observer, to have remained an unresolved issue right up until today. This section will be necessarily a little personal and verge dangerously on being reminiscences.

My own interest in high performance networking was very much influenced by the work done on the US Gigabit Testbed Initiative (GTI), which commenced in about 1990 and ran for about five years. There were five testbeds, but the one I thought most suggestive and visionary was Casa—a gigabit testbed that connected Caltech, JPL¹, the San Diego Supercomputer Center and Los Alamos National Laboratory. The primary focus of Casa, under the extraordinary leadership of Paul Messina, was on distributed heterogeneous supercomputing applications involving very large computational problems. As such, it was a very influential forerunner to grid computing and much more. The GTI, in part, led to the establishment of the very high-speed Backbone Network Service (vBNS) and, indirectly, to Internet2.

The GTI and Casa, in particular, influenced the establishment of a new research center in Australia of whose board I was a member—the Research Data Network Cooperative Research Center (RDN-CRC), which was established in 1993 to carry out research on Telstra Australia’s—Australia’s major carrier—Experimental Broadband Network (EBN). Two of the research projects were in heterogeneous supercomputing.

Around the same time, discussions were being carried out between the Australian and Japanese governments—they had even risen on two occasions to the prime ministerial level—on cooperation in HPCC² that could broadly build on the success of the research partnership between the Australian National University and Fujitsu in which I was then involved (and that continues to this day). These discussions led ultimately, in 1995, to an agreement to establish a direct network connection between Australia and Japan basically to connect high-performance network testbeds in these two countries (though this does simplify a more complicated situation).

Through these activities, I was invited to give a paper at the *Asia-Pacific Economic Cooperation Symposium for Realizing the Information Society*, which was held in Tsukuba, Japan on March 27-28, 1996. The paper was called “Towards an Asia-Pacific

¹ Jet Propulsion Laboratory

² High-Performance Computing Cluster

Gigabit Testbed.” In it, I advocated generalizing the model of intercontinental connectivity that we were trying to bring about between Australia and Japan, to the broader Asia Pacific regions based on the success of GTI in the United States.

Professor Kilnam Chon, then in the Department of Computer Science at KAIST³ in Korea, also gave a paper at the same meeting called “Asian Highway – Why Not Gigabit Network?” Though we had never met, nor did we know each other, our papers were remarkably similar in what we advocated—namely a high performance Asia Pacific network.

During the meeting a small group of us—including myself, Professor Chon; Professor Goto from Wasada University; Mr. Konishi, then at KDD⁴; Dr. Tin Tan Wee from Singapore; and one or two others met in the lobby of the Dai-Ichi Hotel in Tsukuba on the evening of March 27 and agreed to try to establish such a network.

This was initially called the AP Testbed and I agreed to draft notes of our meeting summarizing what we had agreed to try to do.

Steve Goldstein from the National Science Foundation also attended the meeting to observe. He did, though, inform us that the NSF was considering releasing a solicitation called the High Performance International Internet Services (HPIIS) calling for proposals to establish connections between the vBNS and comparable networks in other regions of the world. He suggested that this program might be a way of providing a high-speed connection from the Asia Pacific to the U.S.

On June 18-20, 1996, the *APEC/APII Testbed Forum* meeting was held in Seoul, Korea. By this time, Professor Chon and I had agreed to combine and revise our papers into a joint paper called “Towards an Asia Pacific Advanced Network” which we

³ Korea Advanced Institute of Science and Technology

⁴ Kokusai Denshin Denwa Co., Japan’s international telecommunications carrier

presented at this meeting. During the *Forum* we had the first major meeting of representatives of most of the countries, institutions, and networks that would eventually become the founding partners in APAN.

During the meeting, there was a detailed discussion about a more precise mission for APAN—would it be, in my words, more of a coordinating body for advanced networking activities in the region, or would it be an organization more like, for example, Internet2, that owned and operated its own infrastructure funded by its members that connected all of them and selected others?

Some advocated for the latter. In fact, here is a diagram I drew describing this approach—this was still from the era in which you produced a hand-written transparency to get your ideas down quickly.

But finding a way to fund such an ambitious undertaking was always going to be very hard to accomplish, as it was an idea well beyond what most funding agencies could comprehend at that point. Recognizing this reality, APAN's mission basically became the former option—as it is described today, “to coordinate and promote network technology developments and advances in network-based applications and services across the Asia Pacific region...”

Soon after this, however, I accepted the position of vice president for information technology at IU and moved to the U.S. at the end of 1996. I expected my involvement with APAN had come to an end. But in the spring of 1997, the NSF finally released its HPIIS solicitation. APAN was about to become formally established. I met with Professor Chon, Professor Goto, Mr. Konishi and others in Tokyo in June 1997, and we agreed to partner on a joint proposal to the NSF HPIIS program entitled “TransPAC: A High Performance Network Connection for Research and Education between the vBNS and the Asia Pacific Advanced Network (APAN)” on which I was principal investigator.

This was funded by the NSF, initially, for five years for approximately \$10 million, and TransPAC commenced on July 30, 1998. Initially, it was a 35Mb connection. Today, it comprises two 10Gb connections. TransPAC-2 commenced in August 2005, and, as I had taken on additional administrative responsibilities at IU as vice president for research, Jim Williams replaced me as PI, and he has done a superb job in this role ever since, as have many others, including Steve Wallace, who was involved from the outset. I think it is fair to say that the credibility that this NSF HPIIS award gave to the whole idea of international connectivity helped significantly with the growth of APAN, and the NSF deserves great credit for this program.

Later in 1998, IU was also awarded the contract to manage and operate the Internet2 network—then called Abilene, and, in 2000, we formally established the Global Network Operations Center—the Global NOC, to manage TransPAC, a number of other international connections funded under the HPIIS program, and Abilene. Today, it is the premiere NOC in the world for advanced research and education networks. It manages 18 major national and international networks, employs over 100 network engineers and associated technical staff, and has over 15 graduate students working on various projects associated with the NOC.

APAN, of course, has experienced a great deal of growth and success since its establishment in 1998. It has grown to include 16 primary members, 12 affiliate members, and an additional 10 associate, liaison, and industry members.

It has been an important catalyst and coordinator for the development of regional NREN network interconnections in the Asian Pacific region, and for the interconnection of the region to other parts of the world as TransPAC for example has demonstrated.

It has also supported initiatives in fields that include bioinformatics, medical informatics, distributed computing, telemanufacturing, remote robotic control, digital libraries, wide area parallel computing, astronomy, and high-energy physics.

APAN's regular meetings have become focal points for researchers, network engineers and others involved in the uses and applications of higher performance networks in the Asia Pacific region. APAN and its leaders over 17 years are to be congratulated, then, on all they have accomplished.

3. WHAT UNIVERSITIES EXPECT FROM HIGH-PERFORMANCE NETWORKS

I have described how, when APAN was born, there was discussion as to what its mission should be: one proposal being for it to establish its own infrastructure. I also described how it did not proceed in this direction in any major way as it was then seen as being unrealistic. However, I believe it is now time for APAN to again seriously consider engaging in developing its own infrastructure as part of an effort to build a true global advanced research and education network. And I would claim the need for such a network is becoming essential for education and research around the globe.

As a university president, I try to view everything we do 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),

and

- the preservation of knowledge (information repositories).

And as all of you know, high-speed global computer networks are or are becoming fundamental in every one of these areas. They are the means by which international digitally-enabled education and research is taking place. It is no exaggeration to say that research and science has become almost totally digital. Data is being generated, collected, processed, analyzed, visualized and stored in digital form. Simulations and

modeling are being carried out completely digitally. And the historical and contemporary archives of science, certainly the main material, have been converted fully into digital form.

At the same time, science has become completely international in character. In the academy, scholarship and research in just about every discipline from anthropology to zoology is truly international—a process hugely accelerated by the Internet. There is, in general, no such thing as American anthropology or Chinese zoology—just anthropology and zoology (though there may be contending schools of theory and analysis within these disciplines).

This scholarship and research takes place within a global research or scholarly environment where it is, in general, facts and reason that determine progress, not national origin. Hence, the quality of the programs and research at universities is determined by the quality of the faculty and students who contribute to them, and they can come from anywhere in the world. And fundamental to research, especially in the sciences, is collaboration, whether it be two co-authors on opposite sides of the world, or a group of thousands from dozens of countries working with some major experimental facility, most famously for example, the CERN particle accelerator in Geneva, but it is only the most visible.

I said a moment ago that research and science have become completely international in character. The same is certainly true of education. 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 universities establishing campuses and other facilities all over the world, with many degrees now 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.

From the perspective of a university president, one of my primary concerns is that faculty, students, and staff have available to them the infrastructure that will allow them to work at the most sophisticated level and to collaborate effectively with colleagues around the world.

It is impossible to imagine a college or university of any size operating effectively today without such infrastructure. Conversely, it should by now be unacceptable for this infrastructure to be run at anything but the highest production standards as we have come to expect of our NRENs, and in the best engineered and best operated way. And they also need to be provisioned and procured in a fiscally responsible way. We are seeing research funding being reduced and coming under serious pressures in many parts of the world. Funding agencies and universities need to be providing as much funding as possible for education and research—not funding infrastructure that is in reality superfluous, or managed or operated sub-optimally.

Global networking, that is, the global interconnection of NRENs, has, to be blunt, been accomplished overall in an ad hoc way in spite of the best efforts of organizations like APAN and their local success. There is, on the one hand, massive over-provisioning in some places with all the economic inefficiencies this implies—for example, there are over 10 separately acquired and operated 10Gb connections across the Atlantic, and, on the other hand, connectivity to some parts of the world is at the mercy of the whims of funding agencies, even individual PIs, or even warring network potentates. The global interconnecting fabric is, overall, not architected for redundancy or reliability. This was acceptable once, when international infrastructure was an experimental domain with still emerging standards and high financial and technical cost of entry, and when the services that it provided were less fundamental to the world higher education enterprise.

But this is no longer true in an age of standards, of fiber and bandwidth abundance, and the increasing commodification of much of the infrastructure.

4. BUILDING A HIGH-SPEED PRODUCTION GLOBAL R&E NETWORK

There have been a number of such high-speed production research and education networks proposed in the past—my Indiana University colleague 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.

There has been much activity in this area, gradually growing in scale, complexity, and reliability. Some of the components that could comprise such a network exist at the moment allowing reasonable connectivity, for example, between the U.S., Europe, and parts of Asia, due to the vision and enormous hard work of many National Research and Education Network leaders worldwide, including those involved in APAN. But much of it is dependent on the vagaries of agency funding or individual institutional 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 developments such as software defined networking, and with the now central 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.

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 and 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 its concerted efforts. Here, I would strongly encourage APAN and the NRENs in the region to engage very proactively, as happened 17 years ago, and to seek to play a direct role in efforts to establish a production global R&E network.

Establishing such a network will also require the involvement, in some form, of the key higher education organizations in the relevant countries such as the Association of American Universities in the U.S., the key funding agencies in these countries such as the NSF and NIH in the U.S., and other key government agencies and, where possible, multilateral organizations.

In this regard, there have recently begun to be grounds for some optimism. I want to enthusiastically commend the various NRENs who met last September in Geneva to begin to seriously address these problems and I am, of course, very pleased Internet2 is among these. They deserve great credit for these efforts. Some of you may have seen the communiqué they released in which they indicated they would begin to tackle four different areas fundamental to a global research and education network. These are:

1. Global Network Architecture—developing a well-defined, inclusive, global architecture for, and a roadmap towards interconnecting the R&E networks on a global scale;

2. Global Federated Identity Management—developing a global, interworking architecture for, and a roadmap towards, the delivery of federated identity management for the global R&E community;
3. Global Realtime Communications Exchange—developing an interworking system for multi-domain video/audio conferencing systems;

and

4. Global Service Delivery—developing a model for global above-the-net service delivery to the NREN's constituencies, leveraging aggregation of supply and demand through scale (such as Net+ services in the case of Internet2).

These efforts, of course, will involve sophisticated engineering and technical efforts and will no doubt be motivated by many people as being essential to the most advanced and sophisticated research. I, of course, agree with all this. But the point I have tried repeatedly to make is that it goes even beyond this—it is now needed for nearly all that we do in higher education.

5. THE DIGITAL PRESERVATION NETWORK AND THE PRESERVATION OF KNOWLEDGE

I talked in the previous sections about the importance of high performance R&E networks to two of the fundamental missions of universities:

- the creation of knowledge (which is research & innovation) and
- the dissemination of knowledge (which is education & learning).

Now, let me discuss the role of such networks for the third of these missions—the mission of the preservation of knowledge.

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 funding, 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 percent 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 eight years to collect and that has a scientific value measured in decades has a preservation horizon that expires this year. While the Sloan Foundation may well see fit to continue funding the initiative beyond this year, 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 Ancient Library of Alexandria, 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 wave 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 Indiana University alumnus, I am proud to say), remarking at a speech a few years ago, that of the ten networking companies that had existed five years before, only three 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 here in the U.S. 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 year, together with James Hilton, the CIO at the University of Virginia and an Internet2 Board member, and Ann Wolpert, director of the MIT Libraries, I gave a presentation on digital preservation and DPN to the Spring meeting of the presidents of the Association of American Universities. It was received with great interest. Due to the efforts of James and others, over 50 universities have now committed to the establishment 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 three 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.

The success of the Digital Preservation Network will require independent governance that can survive over time—but it is an initiative to which all of our organizations need to be essential contributors.

6. CONCLUSION

Let me say just a few words in conclusion. I have discussed the early days of APAN and encouraged APAN direct involvement in efforts in the Asia Pacific to contribute to the establishment of a true production global R&E network. And I have also noted that such a network is needed not just for the most sophisticated and advanced research, but now to support the research and education missions of universities in their totality. And I have tried to argue that the long-term preservation of the digital knowledge itself will ultimately rest on a global R&E network.

Many of you will now move to attend all the specialist and technical workshops and talks that comprise a conference like TIP-2013 and which make it so valuable. However, as you do, I ask that you might remember these themes since, if progress on these problems is to be made, much of it will be made by people in this room.