The Evolution of NYU-NET in the Context of Regional and National Networking in the U.S.A.

George Sadowsky

george.sadowsky@nyu.edu

New York University
and
The Internet Society

Introduction

The purpose of this talk is to relate to you the development of the Internet in one United States research university, New York University. I will observe that its development was intermeshed with that of the regional network, NYSERNet (New York State Education and Research Network) which the University helped to being into existence and in which it continues to participate actively. A strong cooperative environment has existed among the universities in New York State and NYSERNet that has substantially helped all participants. This experience may be relevant for the CEENET region.

New York University is the largest private university in the United States. It is composed of 13 separate schools, and has approximately 15,000 undergraduate students, 16,000 graduate students, and about 15,000 people at any one time who are taking continuing education courses. It was established in 1833, and for much of its history was oriented purposely to students in the New York City area. However, since World War II it had become increasingly national and international, and now boasts among the highest percentage of foreign students among U.S. universities.

One school, the Courant Institute of Mathematical Sciences, was instrumental in the development of computing and networking at NYU. The Institute was founded with Professor Richard Courant, a famous mathematician, who emigrated from Germany to the U.S. before World War II. After the war ended, The Institute attracted substantially more talent, some of it from the Los Alamos project, which had used applied mathematical methods extensively in the design of the first nuclear weapons.

As a result of this background, the Institute quickly entered into a major research relationship with the U.S. Atomic Energy Commission (now the Department of Energy), in which mathematical modeling and simulation were important techniques. As a result, when digital computers were first commercialized, the Courant Institute received in 1957 an IBM 704 computer for use for many different kinds of scientific calculations. It was one of a small group of companies and industries for which automatic digital computation was essential for progress. The Courant Mathematics and Computing Laboratory (CMCL) was quickly formed, and it did seminal work in the development of computational methods and their applications to problems in applied mathematics. Courant has remained in the forefront of this type of computing activity to this day. In

1965, a Control Data 6600 was installed (serial no. 4), followed by a succession of high powered computational engines from various suppliers.

Early Experimental Networking (1975-1983)

In the early days of scientific computing, there were different types of computer, all having different operating systems, programming languages, and data formats. Interoperability was the exception, not the rule. The post World War II climate was quite favorable for the growth of scientific research, and applied mathematics research groups sprung up at other research centers and universities. Much of the research required collaboration among this community, and often required use of distant computing facilities to implement such collaboration.

This requirement was at the heart of the initial ARPANet. Using file transfer and remote login techniques, it would be possible for someone at one location to use a computer at another location as if they were physically present. Because of the very large research productivity gains that this mode of operation promised, mathematicians and computer specialists became very excited about achieving this goal, and the initial ARPANet results spurred them on. Courant was one of the leaders in this field, and the CMCL staff contributed substantially to parts of the ARPANet protocols and applications in the 1970s.

In 1981, with the spread of university computing opportunities outside of the sciences, Courant was asked to take over the academic computing support responsibility for all of New York University, which it did. At that time, networking was still experimental. The mode of development seemed to be towards disciplinary networks such as CSNET for computer sciences, HEPNET for high energy physicists, and BITNET for mail and file transfer. This was possibly encouraged by ARPA which, although it had the closest thing to a general purpose network at that time, restricted its use to those working on government and military activities. However, one should note that the then management of ARPANet had no concept of how large their network would become, since the pre-TCP/IP protocol set, NCP, had address space for addressing at most 256 networks of computers.

In 1983, the ARPANet made a very large and significant transition from its earlier protocol set, NCP, to TCP/IP. It was thought that the theoretical capacity of TCP/IP to address 4 billion hosts would suffice for the indefinite future, exemplifying the sense of the time that the Internet would likely remain a relatively specialized network.

There were other competing networking efforts at the time. UUCP (Unix-to-Unix copy Program) was used across dial-up links to implement a store and forward network for delivering mail files. Similarly, FidoNet was a network created by a grass roots effort both in the K-12 community and by NGOs operating in developing countries to develop a cooperative volunteer store and forward network using dial-up telephone circuits.

Technological Backdrop

Those of us who work in information technology are the beneficiaries of technological progress unmatched in hardly any other industry. Since 1955, the performance-to-price index for information technology hardware in general has been in the range of 25-30% per year. The same measure for the telecommunications industry has historically been considerably lower but in recent years, however as this technology has depended to an increasing extent upon the semiconductor industry and digital communication using optical fiber media, technological progress has accelerated sharply. Nor is there an end in sight, at least for the next 10 years which is generally the limit of product development vision.

The cumulative effect of this rate of progress can be measured quantitatively, but it is felt qualitatively, as entire major areas of applications move from being infeasible to being feasible. Markets spring up without much warning, and the demographics of capitalism accelerate, with firms being born and dying with great rapidity. We have appropriately chosen to call this phenomenon *Internet time*, and it determines the speed of our clock.

Awakening of the Academic and Research Community

In the early 1980s, the National Science Foundation (NSF), an executive agency within the U.S. government, was increasingly the target of reports that U.S. scientists were falling behind other countries because of the lower levels of investment in supercomputing in the U.S. In highly publicized testimony, the astrophysicist Larry Smarr reported that he had to travel to Germany to do his research effectively. The NSF has a mandate to encourage and support research in the sciences, and it works through branches corresponding to subject matter areas that fund programs and projects.

NSF responded to these complaints in the mid 1980's by funding the establishment of 5 supercomputer centers, and added a component for funding regional networks that would connect to an NSF national backbone for the purpose of allowing researchers to access the supercomputer centers through a network rather than having to travel to the center itself. In retrospect, it is ironic that since that investment decision the relative importance of supercomputers has diminished while the relative importance of data networks has grown enormously.

This three-tier model of network expansion — campus networks, regional networks, and the national backbone, formed a solid base for the expansion of academic networking. Starting with a backbone of 64 Kbps on the backbone, capacity increased in 1988 using fractional T1 circuits, in 1991 using T3 circuits, and finally became commercial in 1995 at ever increasing bandwidths. We no longer speak of a backbone because the topology has become much more complex in the five years since that time.

The mid 1980's were a time of significant excitement in the academic and research community. The rapid development of mini and microcomputing technology yielded the concept of the 3M machine, a computer that had 1 megabyte of primary memory, could execute 1 million instructions per second, and had a 1 megapixel display unit. This conceptualization of the scholar's workstation was thought to be an ideal computing environment for much of the work of that time. In addition, the deployment of broadband Ethernet and the expansion of the ARPANet presaged network connections ot powerful external services.

NYSERNet (New York State Research and Education Network) was the first regional network to establish connectivity in its region. NYSERNet was established in 1986 to *"advance science and education in New York State by means of high speed telecommunications techniques, and to assist … in gaining access to … resources outside of New York State by such means."* In addition, it connected the one supercomputer center within New York State, at Cornell University, to the net. NYSERNet is a not-for-profit organization, and its Board of Directors consists of representatives of the major institutions which it connects." New York state is one of 50 states, containing about 20-25 million people, most of whom live in an L-shaped region connecting the New York city metropolitan area, Albany to the north, and Buffalo to the west.

NYSERNet was founded by the leading research universities of New York State. On the one hand, an organization like NYSERNet had to happen; in no way could individual universities have afforded the cost of individually connecting to the national network. On the other hand, it was a happy coincidence, since it laid the groundwork for the formation of personal and institutional relationships as well as cooperative programs that have served the organization and its members very well during its existence.

First Generation Regional and Campus Networks

First generation regional networks were characterized by low bandwidth, few subscribers, and experimental learning. In the case of NYSERNet, an initial experiment to contract the network to the regional telephone service provider yielded unsatisfactory results, and NYSERNet decided to build its own network, using circuits leased from the company. The network infrastructure consisted of T1 circuits.

In 1990, the network design and operations team wanted to enter the then emerging commercial IP industry. An amicable settlement was reached whereby the team, led by Bill Schrader, would take ownership of the physical networking assets and would leave with the staff to form a new commercial company, Performance Systems International, Inc., and would offer a commercial service using PSINET. NYSERNet would contract with PSI for IP service for a five year period, and in return, NYSERNet would be given a portion of the common stock of PSI.

At about the same time, other companies were making similar decisions. Alternet, under the leadership of Rick Adams, founded UUNet, and established companies such as MCI, BBN and Sprint were beginning to enter the field. Advanced Networks and Services (ANS) was formed in 1990 to manage the second generation NSFNet backbone, using IBM RS/6000 minicomputers as routers for T3 circuits. During the next few years, the Internet prospered, especially with the introduction and spread of the World Wide Web, and infrastructures such as PSI benefited from that activity. ANS did especially well; after the next evolutionary step of NSFNet — Network Access Points for peering and the commercialization of the backbone — they sold their network to America OnLine (AOL) and became a charitable foundation.

Networking was evolving at NYU during this time. Since NYU has been an early entrant into networking, it had deployed early technology. In 1985, as a result of the Greene-AT&T decision, NYU installed its own internal telephone system, and at the same time, laid a broadband coaxial cable connecting all main buildings for both television and data communication. Several thousand nodes were connected into the 1990's through buffered repeaters and low cost terminal control units. (This is very similar to the same technology that is used now in home cable Internet connections.)

The shortcomings of coaxial cable transport soon became apparent. Reliable connectivity depended upon high frequency analog carriers, with active amplifiers throughout. Tuning the entire network was a major problem. Second, the entire network was bridged. Occasional broadcast storms were frustrating and their sources were sometimes difficult to locate. Occasionally someone would plug a television source into an Ethernet jack and the network would be rendered unusable. The system used two unidirectional 5 megahertz channels, and there would be occasional interference when adjacent analog channels were used. For these reasons as well as increasing utilization, in 1993 a decision was made to migrate to NYU-NET-2, a fiber based, routed network centered initially around a DEC Gigaswitch, using FDDI as the major backbone transport with Ethernet tail circuits.

The NYU-NET-2 undertaking started in 1993-94 was planned for 2-3 years, but in fact has lasted much longer. A good part of this result was that previously we perceived network implementation and upgrading as a series of discrete steps, separated widely in time. Such a concept was born in the earlier period of experimental networking, which was project based and related to the funding cycles of government agencies. With NYU-NET-2, we learned that network evolution was moving from a step function to a more continuous process as large parts of campuses wanted to become connected to the net.

Second Generation Networking: After PSI and NYU-NET-1

Near the end of the five year contract with PSI, Inc., NYSERNet began to outgrow the capacity provided by PSI. In addition, NYSERNet management as well as university representatives had become familiar with some of the disadvantages of not having control of their network which they had enjoyed prior to 1990. A decision was made to invest in its own network again.

The decision was made possible by its investment in PSI. Before that time PSI had launched an initial public offering (IPO) and its stock was commanding respectable value. NYSERNet therefore liquidated its position at a significant profit, and used a part of the funds to invest in new network infrastructure which, in cooperation with Sprint and

Verizon (then NYNEX) provided a statewide T3 network, managed by external contractors but controlled by NYSERNet.

At about the same time, it was the judgment of NYSERNet management that in an era of rapid commercialization of Internet services, regional academic and research networks had a limited future. A decision had to be made between static continuation, which was feared would lead to acquisition or death, or commercialization. A decision was made to go the route of commercialization again, and from that decision in 1995-96 was born AppliedTheory Corporation. The scenario repeated itself; much of AppliedTheory was owned by NYSERNet, the physical assets were transferred to AppliedTheory, and a five year service contract was concluded. AppliedTheory inherited the existing relationships with Sprint, which provided the WAN connectivity and overall network management within New York State, and with NYNEX, which provided the local loops.

The Internet Explosion (1995-1999)

Although the growth rate of the Internet has been consistently high from its beginning, the impact of this growth, coupled with public awareness, came in the mid-1990's. The applications explosion provided much of the fuel. Applications such as Gopher, which provided an entirely new and powerful way to organize and access information hierarchically and associatively, were smothered a year later by the emergence of the World Wide Web and Mosaic based upon the initial HTML version 1.0. Since then the Web has expanded even more, with Netscape, Internet explorer, new versions of HTML, XML, Java, back end servers, Perl scripts, applets and servlets.

During this period of time, there were substantial challenges to campus networking. At NYU, infrastructure expense levels were beginning to be quite visible in computing budgets. Bandwidths that had seemed very excessive at the beginning of the period were increasingly inadequate as applications evolved, new users joined the net, and applications became more bandwidth intensive. User expectations were shifting; instead of regarding the network as an experimental entity as it was regarded in 1990, greater reliability and availability were expected. The network began to be exploited for business applications, which heightened expectations.

At the same time, the net began to be an attractive target for hackers of the worse kind, so that security issues concerning the net not only were increasingly visible to its users, but had to be dealt with by network security staffs which were established and grew. Universities, having a large population of students who were going through an experimental period in their lives, suffered more than most. Misbehavior, and how to deal with it in an environment that championed free expression, became a major issue.

Internet-2

NYSERNet was quite right in forecasting their future as commercial growth and stagnant demise. During these years every regional network in the United States was either acquired by a commercial ISP or went out of business in some other way.

The academic community realized that it had lost control of the Internet and were now at the mercy of commodity Internet providers. With increasing congestion and no remedy in sight, the Internet-2 project was started. Internet-2 was to be a separate network, technically stable, offering high bandwidth and different qualities of service for experimental applications. The quality of service dimension (QoS) was especially important, since it held the key to rational pricing, investment, and rationing of the resource.

Internet-2 has developed substantially, and has added a substantial amount of capacity for academic use. Experimental applications are being developed, although at a disappointing rate. More disappointing, the use of the network is not being restricted to experimental applications, but is being used for all traffic between participating institutions. In theory there is a more restrictive acceptable use policy but it is not enforced. In addition, work on QoS has been more difficult than what was foreseen, and progress has been slow. Finally, the ability to access Internet-2 on an end to end basis is causing many campuses to have to make relatively expensive additions and changes to their campus networks.

Internet-2 topology is similar to that supported by NSF 15 years ago. There are regional networks, now called regional aggregation points, or *gigapops*, that are connected in turn to each other. Coalitions of geographically proximate members are charged with the responsibility of creating and operating the gigapop and connecting themselves to it.

NYSERNet participates in Internet-2 through an unusual gigapop that is 500 miles long and several optical fibers thick, which uses the right of way of the New York State Thruway, a limited access highway that connects New York City to Albany to Buffalo. The gigapop operates at OC-12 speed, and has connections at both ends to Internet-2 backbone provider points of presence. This architecture would not have been possible without using the capital gains realized on PSI stock in 1995, which provided a moderate size endowment for NYSERNet. This research network, called NYSERNet 2000 and sponsored by NYSERNet, is in addition to NYSERNet's connections to the commodity network.

Within NYSERNet, we have seen a slow move toward meaningful applications on NYSERNet 2000. Perhaps the majority of them so far have had to do with geographically distributed artistic synchronized performances. In my opinion, many of the applications on a national basis have demanded low latency to be successful, with high bandwidth being a remote second attribute of choice. To the extent that this continues, quality of service research and implementation offers a great deal of hope, even over networks that may be somewhat congested at the present time.

NYU's reaction to Internet-2 is, I think moderately typical. We have benefited from the increased bandwidth, and scientists who share large data files with their colleagues have done well. On the other had, new applications have been slow to emerge, which leads to the question of whether there is a "killer application" in Internet-2 space. One would think that some form of enhanced desktop videoconferencing would be forthcoming and would claim such a title, but this has not happened, at least not yet.

Third Generation NYSERNet Networking

NYSERNet continues to provide commodity Internet service to its members through AppliedTheory. From single T3 lines spanning the State of New York in 1996, AppliedTheory has expanded to an OC-12 network spanning much of the eastern part of the United States, and also including California and Washington State. NYSERNet's member institutions are upgrading their connections from T3 to OC-3 to meet local demand. NYSERNet, committed to pushing the networking frontier in New York State, is assisting the early adopters to make such moves by offering operating subsidies from its endowment.

The Future

NYU-NET's future is easy to see in the large, but not in the small. We anticipate that demand will continue to grow at the same rate. With all student residence hall rooms now networked and network participation rates at 60% and climbing, growth will continue. A serious challenge that student computing presents is how to provide available bandwidth for academic use in the face of recreational networking activities such as Napster and its relatives that consume enormous amounts of bandwidth.

Increased reliability and availability will be critical. Telephony engineers speak of five-9's reliability, i.e. 99.999% uptime. We are not there yet. Spread spectrum wireless technology is invading Internet space just as most of our buildings have become wired. What are the benefits of providing mobile computing, and what are the costs and the opportunity costs of doing so? What are the residual security issues and risks in implementing the current state of IEEE 802.11 based wireless technology? How quickly will we have to, or want to, make the transition to Ipv6, in light of using up the Ipv4 address space? How should regard voice over IP (VoIP) technology at this point in time? To what extent does the connecting of other non-IP hand held devices to the local network contribute to the academic mission or the business operations of the institution? There are many more questions than answers at this time.

Policy Issues are Important

Universities and similar institutions differ from Internet service providers (ISPs). Whereas an account with an ISP generally enables the subscriber to use the services provided for anything legal, academic and research institutions — and the networks that they manage — furnish network services to their various constituencies in furtherance of their mission. Earlier in academic networking the U.S. National Science Foundation formulated an acceptable use policy (AUP) for all users of its network that stated essentially that the network was not to be used for commercial or other non-mission related activities. Such AUPs are a useful rationing device in a field characterized by high and rising costs.

Similarly, institutions need to ensure that all connected institutions are paying their own way. If downstreaming of connections is permitted i.e. a participant attached to the network through one or more participants, then it is not unreasonable to establish rules that ensure that downstreamed participants pay some share of the cost of maintaining the network. Alternatively, network connections could be engineered and priced assuming that all bandwidth would be in constant use, and then downstreaming policy could be left to the institutions connected to the network. This could lead to substantially higher costs overall.

Policies regarding privacy, security, and appropriate behavior can be a significant issue in universities. By the nature of the institution, networking has some experimental component in higher education. The difficulty comes in defining what behavior is appropriate and acceptable, and what behavior is inappropriate and unacceptable. One strategy is to lay down a long list of rules; this generally generates a competitive reaction in students who then work to see what unacceptable modes of behavior still fall within the rules. Another approach is to provide general principles and illustrative cases. Regardless of the approach chosen, enforcement will be needed, and education may be one of the best tools for minimizing the need for enforcement.

Issues of content are always present. Censorship of content goes against most democracies, yet allowing all content regardless of network effect may cause inadvertent denial of service for others. The spread of Napster earlier this year illustrated this tradeoff in a dramatic manner. The tensions between freedom of expression, content choices, privacy, and institutional mission are sometimes not easy to reconcile.

Economic and Financial Issues

Early wide area networking has been substantially subsidized in many countries, at multiple levels. Early use was experimental, limited in scope, ad not a part of the essential operations of an institution. Costs to the institution were limited and often covered by research grant and contract funds.

Most networking is no longer experimental, and the costs of providing what is becoming a new and essential infrastructure to an entire institution is very high, the more so because the rapid technical advances underlying the infrastructure imply a short replacement cycle or substantial opportunity cost. These costs have grown to the point where central administrations often lean toward the side of recovering them by direct user charges. This is easy to do on a very approximate basis, but still difficult to do in a more exact manner.

At NYU for example, the cost of just our T1 Internet connection from 1990 to 1995 was approximately $25,000 per year, which was very modestly subsidized. From 1995 to 1999, the cost of just our T3 Internet connection was about $105,000 per year. This cost was more heavily subsidized, since the equivalent commercial price of such service was between $250,000 and $300,000 per year. Finally, starting at the end of year 2000, the cost of our OC-3 connection to the commodity Internet will be about $250,000 per year, less than half of the commercial cost of such a connection. Many other universities are paying for their service at or close to commercial rates.

Within NYU, historically we have set a zero marginal price for the use of the existing network, but have charged offices schools and departments the cost of the equipment to build, extend, modify, and upgrade the network. The labor for this work has been paid for by central funds. This policy was appropriate for an experimental network.

Usage pricing is appealing, but difficult to accomplish. With appropriate records and accounting and data base programs it is possible to charge for an IP address, or for use of a network jack. However, it is extremely costly to charge by volume of transmission, i.e. by the packet. One can approach this result by sampling, with the resulting inequity being determined by the sampling error of the process implemented. One can price also by specific services rendered, such as installations, deinstalls, and trouble shooting calls

From an economist's point of view, such a system is in equilibrium if the aggregate charges collected from all sources (including possibly government grants and other sources from outside the institution) are sufficient to pay for the cost of operating the network and providing the network services, and in addition, to pay for the cost of the renewals, upgrades, and network expansions needed to keep the services current when measured against peer institutions. While there are not a large number of choices, the choice is not trivial, and may well change as the technology evolves. Different institutions will come to different conclusions based upon specific circumstances and practices within their institution and their country.

Smart Policy Decisions or Luck?

Both NYU-NET and NYSERNet are examples of successful activities. They deliver services with reasonable reliability, and they have evolved in an affordable manner that has generally met almost everyone's needs. To obtain a better performance in either case would have taken considerably greater resources that were not available.

I believe that there were significant decisions taken that helped both of these projects to be successful. For NYU, it was smart policy to get involved with networking early and to continue to be involved in the forefront of networking as much as possible. This policy, although probably never enunciated formally as such, allowed the development of a core technical staff experienced in networking that allowed us to make investment decisions — both with regard to timing and technology — that were in the university's best interests. It allowed us to be more than proportionately influential in steering the NYSERNet direction.

Perhaps the most beneficial policy that emerged from our relationship with NYSERNet was the cooperative group and activities that emerged. While collaboration on the first generation regional network was necessary because of financial considerations, what emerged was a group that cooperated on a much broader basis, sharing resources, experience, software, and support. Such cooperation worked toward the interests of almost all members of the group, and allowed us as a group to get more and do more with less resources, an important consideration in a period of academic recession.

On NYSERNet's part, it was smart to recognize that it could do more for its members as a buying consortium than its members could do separately. Under the strong and able leadership of Richard Mandelbaum, NYSERNet acted as a cohesive body with respect to every significant investment decision that was made. It forged close partnerships with private sector firms to push the frontier of networking for the benefit of all of its members. Common interests and the common cultural aspects of the academic and research community were easily identified and assisted in common cooperative work both internally and with external organizations.

NYSERNet's relatively unique achievement was to be able to leverage the value that it created by spinning it off into the private sector. Twice NYSERNet reinvented itself by splitting off the major part of the organization and letting it develop and compete in the private sector. The returns to ownership in these companies have been very good so far, and have allowed NYSERNet to emerge as the only remaining regional network O(of the original set commissioned and supported by NSF) and one that has sufficient resources to assist its members, individually and collectively, to stay on the networking frontier.

Nevertheless, there were also elements of luck in the paths that were taken. These are best identified in hindsight. First, NYSERNet did not really understand how important the equity positions in PSI and in AppliedTheory were. It was not clear what the two stock issues would do in terms of valuation, but it was not anticipated that they would become so financially important as they have become. Second, there were no negative competitive relationships among NYSERNet members that disrupted good cooperation between them. Third, we had the benefit of strong and knowledgeable leadership from Richard Mandelbaum, one of the most prominent networking pioneers in the United states. And finally, our timing with regard to the strength of information technology stock issues in the equity markets has been fortuitous.

Conclusion

One might ask whether the NYU-NET and NYSERNet experience is really too U.S.-centric to serve as a model for developing countries. I believe that our experience applies, although not exactly. In support of this, I note that CEENet countries and regions have plenty of market opportunities, although perhaps not as large, as well as opportunities for entrepreneurs. These countries have both local and international partnering opportunities to build networks together and to capitalize upon the value added that they have generated by being early adopters. In most countries there is a strong academic tradition and a recognized professional class. On the other hand, access to the capital markets is clearly not as strong, and there are different legal and regulatory environments that may not be as supportive as needed.

It is clear now that the growth, use and exploitation of internetworking is of crucial importance to learning and research. Growth and evolution of today's networks is inevitable, and the financial implications of this evolution may be severe for universities, as well as other levels of education. University administrations often do not understand this issue. It is therefore important to capitalize the intellectual value added and obtain returns from it. In this regard, the academic sector can be its own enemy, since it often has difficulty extracting financial value from intellectual achievement. Partnerships and cooperative activities have worked within New York State, and are more likely to lead to identification and exploitation of value that can be capitalized for economic return. With some luck, and with proper partnering, academic and research groups can play their role in this world by experimentation and innovation, activities that historically they have done well at, thereby benefiting multiple sectors of civil society.