A Vision for the Network

A Working Paper Contribution for the Broadband Pricing Group
EDUCAUSE

Timothy Lance

In cooperation with
Gary Augustson, Erv Blythe, Jaqueline Brown, Hud Croasdale,
Joanne Hugi, Ron Hutchins, Anne Lance

Yet felt I like some watcher of the skies
when a new planet swims into his ken
Or like stout Cortez who, with wond’ring eyes
stared at the Pacific, and all his men
Looked at each other with wild surmise
silent upon a peak in Darien

John Keats
On First Looking into Chapman’s Homer

I. Prologue

Beyond the ever-elusive horizon awaits a network pervasive, intelligent, capable, one that will effect transforming social benefit and catalyze fundamental scientific breakthroughs possible only in its sustaining environment. This "nervous system for knowledge" will extend and enrich distance learning, creating wholly unanticipated means of information immersion, while its ubiquity, applications, and responsiveness create feedback loops strengthening each, and driving social value altogether new. Technology for its own sake will recede in motivational importance as political, social, economic, educational, and scientific influences assert primacy – for good and ill. And even as these fundamental shifts realize their intended results, greater, unlooked-for consequences may well occur.

It now falls to the network community, already enabling benefits deliberate and unintended, to engage in "wild surmise" regarding the possible: to design and deploy evolving networks flexibly, identify and embrace extrinsic factors that accelerate network development, and foster unexpected uses that transform networking's environment.

Science, industry and policy for the public good

History offers precedent for just such vision; indeed, the extraordinary development of telecommunications demonstrates the profound impact exerted by the confluence of science, policy, and public need. Forty years ago, "telecommunications" meant the monolithic, centralized, public switched AT&T network, which had rapidly achieved near-universal deployment in the United States. Its venerable Bell Labs supported research (also carried out independently and collaboratively at universities) that led to discoveries such as the transistor and laser. These latter tools, our developing academic/industrial research base and National Science Foundation, and the cold war, together forged the crucible from which, fired by Sputnik and the space race, a new concept of communication emerged.

Politically motivated, fueled by national pride, and culminating a post-Sputnik decade when geeks were suddenly cool, the 1969 moon landing also carried far-reaching derivative results. Just three months after Neil Armstrong's "one small step," mankind indeed took a giant leap forward, with the first Internet packet exchange. DARPANet - made possible by combining computing capabilities the space race had accelerated with the realization that those computational powers could be used to communicate – had been born.[1]

[1] http://www.lk.cs.ucla.edu/LK/Inet/1stmesg.html

Quickly embracing the conceptual framework provided by the new discipline of computational complexity,[2][3] which identified creation of computational power, rather than optimization of programming code, as key to bringing new capabilities within reach, over the next decade and a half technology exploded. The microprocessor continued its development, to a remarkable degree following "Moore's Law,"[4] i.e., the number of transistors per integrated circuit would at least double every 18 months.[5] The personal computer was introduced, and proliferates to this day because of extraordinary market forces Moore's Law supports, and because new ways to interface with this technology, notably the point and click graphical user interface pioneered by Doug Engelbart and colleagues at the Stanford Research Institute[6] and extended by Xerox PARC,[7][8] put computing within reach of a broad population.

Emergence of the public Internet

All the necessary components having been assembled for computing's most potent use, communication, in 1985 the National Science Foundation awarded supercomputer contracts to the Cornell Theory Center, Princeton's John Von Neumann Center, the National Center for Supercomputing Applications at the University of Illinois, the Pittsburgh Supercomputing Center, and the San Diego Supercomputer Center at the University of California, San Diego. NSF also supported development of a communications backbone (NSFnet) that provided high-speed (then 56 Kbps) access to these centers, and independent deployment of thirteen regional networks that extended it. Computer communication protocol candidates for these networks included public domain and commercial and proprietary options; NSFnet and the thirteen regionals all decided on TCP/IP, the Internet protocol. When the deployed network exceeded all expectations, the Internet protocol's long maturation process had been amply justified.

From the outset, NSF Program Director for Advanced Networking Steve Wolff strove to foster an environment in which regional creativity could flourish, with this latter force spawning in its turn the now nearly axiomatic concept of the Internet as a network of networks:

---

[2] Juris Hartmanis and Richard E. Stearns, *On the Computational Complexity of Algorithms,* Transactions of the American Mathematical Society, vol. 117, No. 5, May, 1965, pp. 285-306.

[3] Richard Karp, *Combinatorics, Complexity, and Randomness*, Communications of the ACM 29 (1986) 98-109. See also http://www.cs.ucr.edu/~vahid/courses/220_f00/karp_cacm86.pdf to see the paper on line.

[4] Gordon Moore's original paper can be found at http://www.intel.com/research/silicon/moorespaper.pdf.

[5] See also http://www.intel.com/research/silicon/mooreslaw.htm for a brief illustration of this remarkable phenomenon.

[6] http://sloan.stanford.edu/MouseSite/MouseSitePg1.html This work brought Dr. Engelbart the ACM Software Systems award, and later the Turing Award and the National Medal of Technology.

[7] On January 4, 2002, Xerox PARC was spun off as the independent company Palo Alto Research Center Incorporated.

[8] http://www.parc.com/company/history/ The Alto, an early personal computer from PARC, made the first commercial use of a WYSIWYG (what you see is what you get) editor, mouse, and graphical user interface.

"NASA and DOE wanted to limit the network to a defined constituency, such as researchers at the federal labs and grant recipients, and run it for them. But I thought the money wouldn't be there forever, and even if it was, we should do more things with it, build bigger networks" (SRI interview with Steve Wolff, April 23, 1996). Hans-Werner Braun, co-PI of the Merit proposal, said of Wolff, "In fact, shortly after Steve Wolff started at NSF he made comments to me, and I assume to others as well, saying 'I do not want to see this network only as a supercomputing center network'" (Merit, 1995: 17).

"In response to the Connections solicitation, NSF received innovative responses from what would become two of the major regional networks: SURANET and NYSERNET. They proposed a regional, distributed network design rather than one with all universities independently connected to the regional supercomputing center (a "star" design).

> "The NYSERNET and SURANET examples caused a major paradigm shift at NSF. Instead of funding institutional connections to supercomputer centers, the NSF shifted to funding connections of 'cohesive' regional networks. ... NSFNET is not a network. It is an internetwork - i.e., a network of networks, which are organizationally and technically autonomous but which interoperate with one another." Steven Wolff, NSF[9]

Titans like Jonathan Postel {http://www.domainhandbook.com/postel.html), who oversaw this protocol's adaptation to the increasingly complex challenges it faced, supplied a second basis for its enduring utility: its suppleness. Swashbuckling Internet engineers (very un-phone-company-like!) have enabled this protocol to dodge countless predictions, and some near-actualizations, of its demise.[10]

By assorted measures (number of nodes and users, packets shipped, etc.) Internet usage has grown, as with Moore's law, with predictable doubling times characteristic of the exponential function but faster (about twelve months[11]). Doubling times for web traffic have been somewhat shorter, as might be expected from the tool driving its overall growth in recent years. Just as the point and click interface drove broad acceptance of personal computing, its Internet recreation has done so for the network, propelling the web into every aspect of daily life.

The regional networks, and networking personnel at their member institutions, cooperated to an astonishing degree, given traditional inter-institutional rivalries. As a result, the R&E community was able to bring its full talent to bear on this burgeoning resource. The NSF and regional backbones soon migrated from 56kbs to T1, then T3.

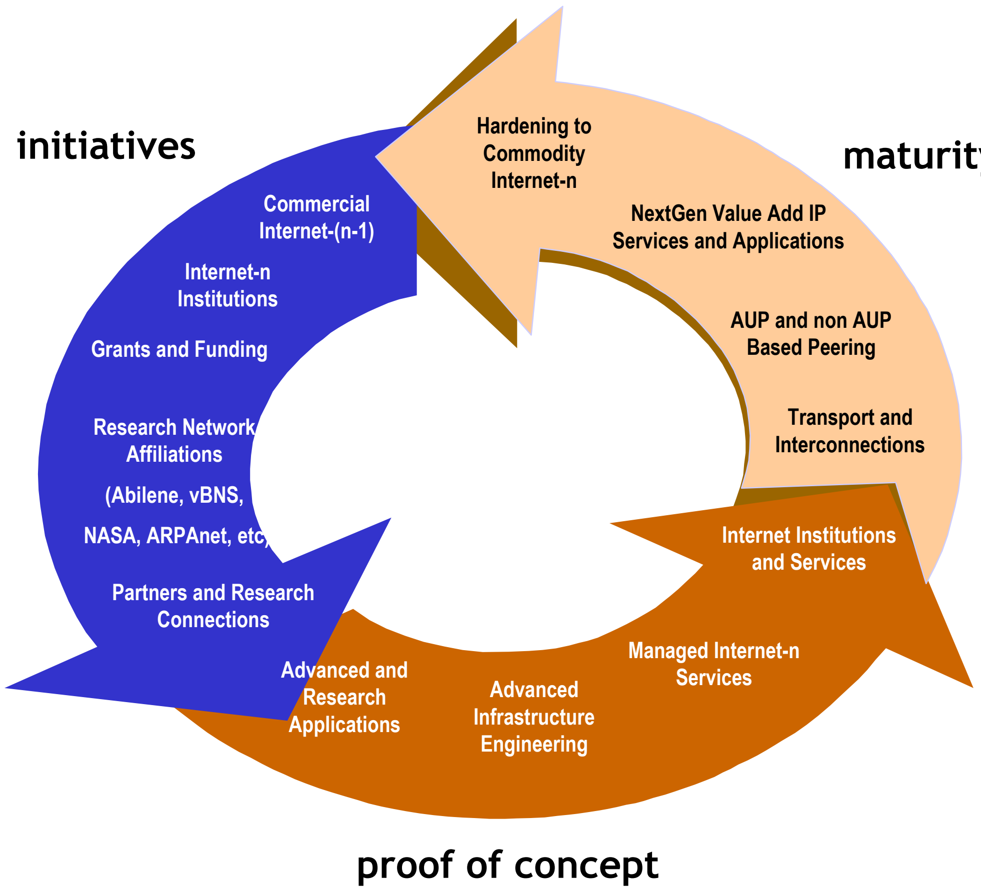
---

[9] (http://www.sri.com/policy/stp/techin/inter3.html)

[10] See *On the Internet*, The Standards Issue, Internet Society Press, Spring/Summer 2001, for a beautiful, user friendly insider's view of the IETF, RFCs, and maintaining the Internet protocol.

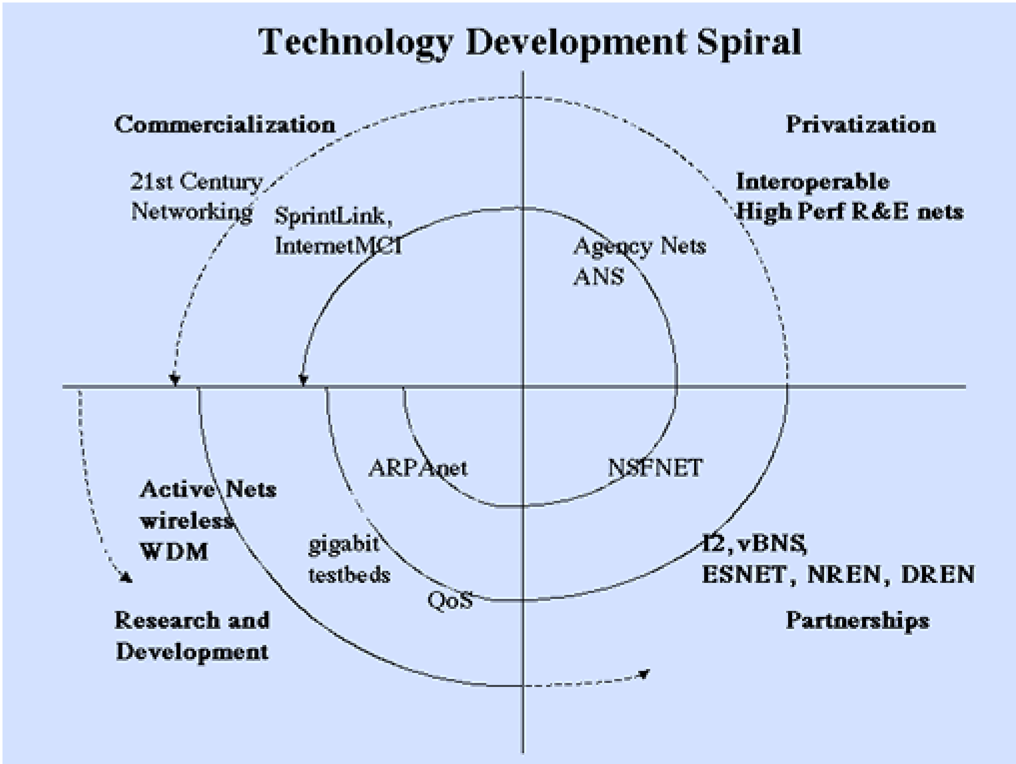
[11] Coffman and Odlyzko, *Internet growth: Is there a "Moore's Law" for data traffic*? AT&T Labs - Research, 2001 http://www.research.att.com/~amo/doc/internet.moore.pdf

Regional networks spun off commercial Internet service providers such as UUNet and PSINet. The networking community had begun the first turn around what we now take as axiomatic, the Internet development cycle:



This development spiral is in fact multithreaded, its independent developments influencing, accelerating, and occasionally merging with others. EDUCAUSE Vice President Dr. Mark Luker depicted this granularity presciently in a 1998 presentation: http://www.arl.org/arl/proceedings/132/programI/luker_ppt/index.htm[12]:

[12] The "revolutionary" Internet development spiral seems so natural, almost axiomatic, that we all feel we must have invented it. But first credit should properly be given to Ivan Moura Campos, a Brazilian computer scientist who reintroduced the spiral concept to the U.S. higher education networking community at the Cheyenne Mountain workshop in Colorado Springs in August, 1996. Dr. Campos today serves on the ICANN Board (http://www.icann.org/biog/campos.htm).



This astonishing development holds a final irony: TCP/IP, developed during the cold war's space race to ensure wartime information dissemination, proved itself an unstoppable vehicle for its peacetime distribution. Soon after the public Internet's birth the Berlin wall fell, quickly followed by the Soviet Union. Rapid improvement in affordable computing and Internet service then proceeded to reshape our lives.

The Monterey Papers

The Internet's robustly developing capabilities and explosive expansion – especially after creation of the World Wide Web, Mosaic, and its offspring browsers – together with the network's increasing commercial use, signaled the privatization of NSFNet and the end of direct Internet support as we knew it. Alarmed by this termination's possible implications for research and education networking, in Fall 1995 representatives from the networking community met in Monterey, California, formed an informal working group known as the Monterey Futures Group, or MFuG, and produced three documents:

A background document for the joint NTTF/FARNET meeting

White Paper: Telecommunications Requirements for a Virtual University

A National Higher Education Networking Organization (NHENO)

The blueprint these documents together provided helped guide advanced networking for the next seven years, influencing formation of the University Corporation of Advanced Internet Development (UCAID), reengagement of the National Science Foundation in direct support of advanced networking, and creation of research networks vBNS and Abilene for the research and education community.

II. Articulating the Role of Research and Education Networks

Rapidly developing technology, shifts in networking's political and commercial environment, and the Internet's penetration to the core of research, education, and our whole society, compel us to revisit the issues addressed by that first Monterey conference and its papers. In doing so, however, we must incorporate three fundamental new challenges:

- Identify parameters for measuring changes resulting from previous trips around the development cycle and, with both now fixed, determine what new network and other technology attributes we must foster and track to optimize development in *network* cycles to come.
- Understand, perhaps even anticipate, network users' requirements, and let their needs drive its evolution.
- In parallel, sustain these early adopters' efforts until the network evolution that their ideas drive matures and new network resources become scalable and affordable.

Lessons derived from past network deployments, and from the computer's development, suggest the importance and difficulty of these three tasks. For example, perhaps the most brilliant components of the web's development were the realizations that, first, "point and click" could make computers broadly accessible, and, second, overlaying such an environment on the Internet would do the same for the network. Yet the web was not developed, as advocated above, as an application-specific response. Rather, it came from the networking community, like many broadly used network tools not driven by users. With a powerful, broadly deployed network now in place, articulation between those who use and those who develop offers a potential wellspring of such ideas.

History suggests that technology will keep pace with insights thus derived. Complexity theory, for example, combined with predictable rapid development in hardware capacity, freed computer design: advances in computer applicability are realized not by optimizing for a given computer class, but by measuring a problem's complexity precisely. Code can then be designed for a sufficiently capacious device that, predictably, will appear.

This same process (and, indeed, many of complexity theory's mathematical tools as laid out in Hartmanis's and Stearns's seminal work[13]) applies here – presenting an extraordinary opportunity, once we understand how to gauge current and future network applications' complexities. Optical networking will massively increase capacity. By listening to its users we can understand potential applications this rich environment enables, as disciplines cross-pollinate and interact in ways not previously possible.

We advocate a five-year networking vision for *our* community, because this is where major Internet innovation occurs. Every day we support research on the network itself and applications (such as the proliferation of "bio-grid" initiatives, Project NEON[14], or the teragrid[15]) that potentially require network resources too advanced, too experimental, to form part of any reasonable corporate business plan. Research and education, however, are nearly synonymous with the institutions we represent; our institutions' very identities mandate pushing, occasionally breaking, the network, in the process learning from both its unanticipated capabilities and their limitations.

Only the R&E community tolerates this modus operandi, even requires that it pervade its research and thus creates an environment optimal for network development. The advanced networking community's comparative freedom from commercial pressures has repeatedly reaped unexpected, far-reaching benefits. Ten years ago neither the web, nor a universally available (including dialup) network capable of using it, existed. No one imagined back then what profound results might emerge from our efforts to deploy, then constantly improve, the network; yet today, how many companies can survive without a web site? How many major companies' inventories does the network support, at the outset fundamentally determine? Home Depot's dominance derives from its IT strength[16], and so does Walmart's. From the consumer's perspective, every major provider, from department stores to airlines to hotels, now has an on-line "Sears Catalog," and easy comparison-shopping has reshaped market competitiveness.

These ideas could not have exerted such far-reaching or profound effects without close interaction with the commercial sector. Xerox's Palo Alto Research Center originated a breathtaking suite of ideas (many commercialized by others) in part because of its close interaction with the Stanford Research Institute and other academic institutions. And yet, many PARC products were successfully commercialized by other companies. Though we have all (except, possibly, Xerox) benefited greatly from such generosity, secrecy, proprietary products, and aggressive attempts to make one product prevail regardless of its relative merit, generally characterize the corporate sector.

Such unsurprising characteristics – especially since the .com bubble burst, concurrent with unsustainable market pressure – require that we assume an additional critical task:

---

[13] Ibid, 2, 3.
[14] http://www.archbold-station.org/abs/NEON/history.html
[15] http://www.teragrid.org/
[16] A Google search on Home Depot plus information technology yields a wealth of information supporting this assertion. Academic courses (see e.g. http://www.cimt.luc.edu/newpages/Research/index.htm have considered the Home Depot phenomenon, and many in the IT vendor community happily featured their use by Home Depot (e.g., http://www.ipservices.att.com/realstories/databriefs/profiles/31_homedepot.cfm)

engage the vendor community in our efforts, balancing free exchange of ideas with our mutual need for successful commercialization of those ideas and their derivatives.

The end to end argument

The very commercial success of a network initially intended to assist national defense, then expanded into a tool for research and education, presents the research community with daunting challenges, as it again attempts to influence succeeding network generations. Specifically, even as we argue for networks of ever-greater capacity, sophistication and "intelligence," as discussed below, "application-level functions cannot, and preferably should not, be built into the lower levels of the system – the core of the network."[17]

Marjory Blumenthal and David Clark describe this quandary as the "end to end argument." Network research and development occur optimally when the end-to-end purpose remains paramount, which can easily be achieved in a laboratory's rarified atmosphere. Once a new network application moves to the research, then to the commodity realms, however, competing forces make the end-to-end argument difficult to sustain. As applications become more mission critical (e.g., linked massive databases in distributed research, remote instrumentation, or telemedicine), and as we ponder September 11's consequences, we look increasingly to redundant paths, rapid recovery, network tools to preserve integrity of data, and adherence to the hierarchical control principles of the end to end argument to maintain end-to-end performance.

Today, measurement and end-to-end performance assessment together comprise an important element in research network oversight, with assessment of stability at the transport layer the most difficult to control in the current R&E networking paradigm. Carriers note that the physical path components of such "data", under much scrutiny since September 11, may still prove unreliable, as carriers on whom they in turn rely combine provisioned multiple paths for cost efficiency.

The public Internet's early years offered the simplest, purest incubating environment for development - free, as we only later realized, from commercial pressure. Indeed, some of the first regionals initially turned to RBOCs to deploy and manage their IP infrastructure. While many did provide the leased circuits, they refrained from learning and exploring this relatively new protocol, freeing the "new" network designers to incorporate (or reject) ideas from the reliable, ubiquitous PSTN. Today we face a more subtle challenge: incorporating or rejecting networking ideas born in the Internet community with the same dispassionate objectivity we applied to the PSTN, even as we necessarily move toward greater control of the transport layer.

---

[17] Marjory S. Blumenthal and David D. Clark, Rethinking the design of the Internet: The end to end argument vs. the brave new world, *ACM Transactions on Internet Technology*, Vol 1 (2001), 70-109. The preprint is available on line at http://lawschool.stanford.edu/e2e/papers/TPRC-Clark-Blumenthal.pdf

Likewise, we must diligently avoid shifting our attention to off task concerns, however worthy (e.g., rural access, remote health care, digital divide, K-12). The networking community functions best when it advances networking; therein also resides its greatest likelihood for influencing ancillary issues. With nearly universal access at a level of the first few network generations achieved, today's "digital divide" solely concerns access to powerful networking. The principle governing our choices, therefore, must be this: we comprise the digital continuum's growth edge; as we push on, our earlier labors become affordable and available to society as a whole[18], the societal component of our end to end argument.

The right network on the wrong infrastructure

The Internet protocol has shown itself durable indeed. When regional networks and the NSFNet backbone were first deployed, only the telephone companies could provide long haul and local loop transport, though within their domains institutions became quite sophisticated. The Internet's diversified, logical overlay relied on the PSTN, an infrastructure designed and implemented in a centralized manner to support voice. Although much Internet transport still occurs this way, the network continues to thrive at ever-higher capacity.

The protocol has also demonstrated its survivability under duress. Although NYSERNet's research network ran through ground zero with transport provided by Verizon, on lines severed when Building 7 collapsed September 11, the network never wavered: Verizon's SONET rings self-healed, and the Internet protocol helped restore commodity service on Long Island and in Westchester County by remapping onto the NYSERNet network. Fiber facilities, and light based services dependent on them, were restored rapidly and for an order of magnitude less cost than Verizon's still-incomplete restoration of the copper-based PSTN infrastructure.

The PSTN in fact amply fulfills the end-to-end argument for voice – in a non-violent world. Its centralized nature dictates that equipment be astonishingly reliable, and it is. Its centralized framework supplies a caller with all the network control necessary for successful voice delivery. Before a phone number's last few digits are "dialed" (a linguistic legacy from the analogue world in which the PSTN developed), a best physical path to the CO near the intended destination has been reserved. Neither that level of control nor the PSTN's centrality suffices for survivability, or for handling digital data's richness, however. The Internet protocol separated the application from its underlying transport logically and brilliantly, yet still relies on a transport infrastructure designed for an analogue age. A decade and a half ago, Davis Clark's classic *The Design Philosophy*

[18] We argue below that though society must concern itself with the "digital divide," its critical component extends well beyond the network, to differential education and skills. As citizens, we must be concerned and contribute to resolving society's overall concerns, but the networking community's unique mission requires it to focus exclusively on developing the high end network.

of the DARPA Internet Protocols[19] compellingly suggested that, as the protocol was being developed de novo for a far richer suite of services, control of the transport medium would permit greater achievement.

Today's challenge and opportunity, then, require us to design from scratch, overlaying an even more advanced protocol onto the right infrastructure by controlling transport. The R&E networking community – hospitals, museums, and libraries as well as universities – must seize opportunities for affordable fiber they need when it is available, and build when it is not. Though the R&E community's focus remains on advanced networks and applications for its use, we can promote and enable mechanisms that make fiber accessible to K-12, local health care and government, business, farms, and individual households. Pervasive societal benefits, often unexpected, invariably derive from advanced network deployments.

R&E regional and metro fiber projects such as the San Diego to Seattle Pacific Light Rail, NYSERNet's Manhattan Project or SURA's RONCO, and the wavelength based Abilene II, with tier one and local providers, herald broader deployments to come. Similarly, the Quilt[20] has proved itself able to leverage size and solidarity with tier one providers, to the R&E community's tremendous benefit.

That same community can serve as an anchor tenant for commercial deployments, with these in turn driving down price. Succeeding network generations can greatly expand the R&E community's traditional role in education, and in employing science, the arts, and medicine for the public good. Properly leveraged, our deployments thus can expedite commercial providers' efforts to supply fiber to a broader community. Control of the transport medium finally provides the right base for support of a next generation architecture, flexible, layered, secure and securely overlaid.

The core of the network – revisiting the end to end argument

Nor does the challenge end with broad control of the transport medium. Today's conventional wisdom puts the next generation network's intelligence increasingly at the edge, controlling essentially limitless fiber-based transport capacity in between. The control is logical, hence easily transferable, and not dependent on a particular opto/electronic device or optical fiber segment.

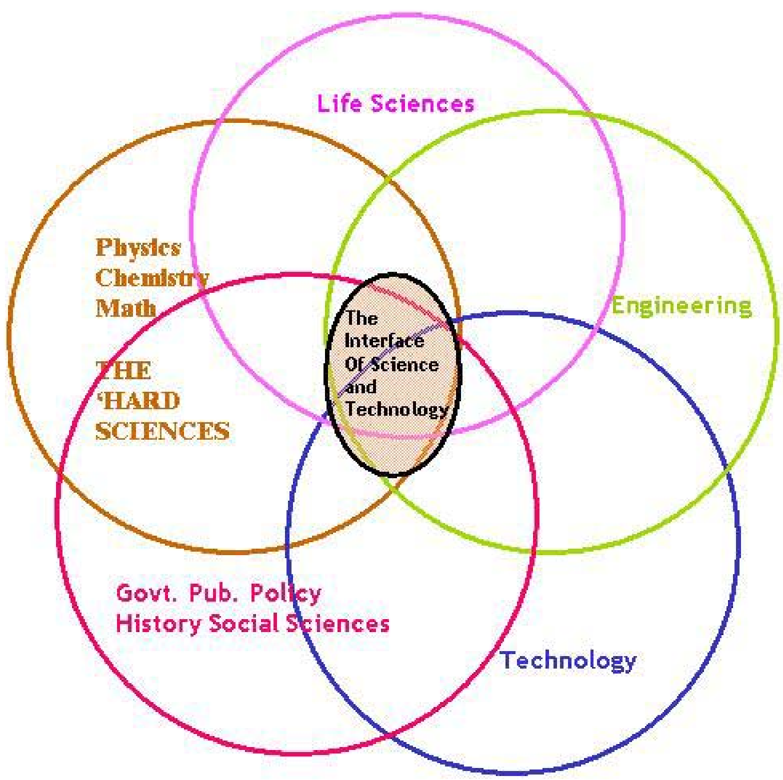
In coming Internet generations, that intelligence will continue to grow at the edge, but will also increase in between. As applications become more sophisticated and portable (that is, able to move to follow users or events to be tracked), the network itself will acquire increasing choice-making discretion, i.e., it will learn from network events as they progress. Future networks' most valuable results will transpire only if the network

[19] Available from ACM at http://www.acm.org/sigcomm/ccr/archive/1995/jan95/ccr-9501-clark.html
[20] http://www.thequilt.net/

itself can respond in a direct, end-to-end way to demands made of it – that is, we must enable the network to actively deal with an increasingly complex array of applications without placing specific application-level functions in its operational core[21].

How can the network support this? Over the past half-century, geopolitics and defense largely drove interactions and collaborations among industry, university research, and government. The fertile environments thus created pushed science and technology's frontiers, transferred enormous technology to the private sector, and profoundly affected all our lives. The transistor, laser, computing, communications, the Internet itself, the point and click environment and its network implementation as the web, all resulted from this marriage. So did new disciplines, like genomics and computational science.

Future network generations must thus be designed not only to provide greater capacity and control, but to satisfy broader functional need. While sustaining the network tools on which research and education already depend, those networks must also support, and will depend on, deeper, discipline-generating interactions among science, mathematics, technology, engineering, government and public policy. From their seminal interface dramatically new ideas can emerge.[22]



[21] Ibid 17, 19
[22] Anne Dunford of NYSERNet provided this graphical representation of the functional core.

Protein folding, global climatic modeling, forensic epidemiology (as in response to biological attack), and understanding of our universe's first moments, e.g., are problems that we are only beginning to be able to tackle, whose progress depends seminally on the network.

During the network's development, crucial ideas emerged from a crucible formed by the right people working "together" in an environment responsive to creative instincts. Personal notes by those making the discoveries in some of this paper's references furnish compelling evidence of this fertile fusion. Today the network, one of this period's most important technical and social consequences, offers new ways for people to be "together," and also expands our definition of "environment." Disciplines overlap increasingly as tools, and our understanding of how to use them in new settings, mature.

But today's network is capable of carrying just a small fraction of the communication engaged in during the most productive human interactions. Fully supporting such communication, and adding to it capabilities such as access to remote instrumentation as if right there, must be the goal of network generations to come. Mimicking human interactions so as not to lose important communication, network linkage must respond almost instinctively. Moreover, if the network itself can learn, it will carry lessons from one interaction to the next, even if that interaction's people or machines are different. Such a capability clearly enables, and would accelerate, scientific discovery, as well as open new vistas of possibilities in areas like health care and learning. It also raises sobering concerns about security, trust and privacy.

Attributes of the Network

It is axiomatic that network generations support increasingly demanding applications, with greater bandwidth and intelligence pushing their leading edge. We must ensure that each new network can grow with its applications through a securely overlaid architecture, and that end-to-end design in all the ramifications sketched above remains possible. Leased circuits no longer furnish adequate transport components for advanced networks; today's minimum deployment or increment unit is a wavelength, with the R&E community determining what happens within each λ. Partnerships with vendors should accelerate deployment of advanced optronics, benefiting both communities, and promote access to dark fiber based research networks.

In short, future network generations will be light based, with the wavelength as the minimal unit of measure, secure on many levels, portable, and organic (discussion follows). Wireless will play an intermediate role, partially resolving the last mile challenge. As fiber deployment increases, with far greater distribution of land-based network access points, wireless devices will in the end mediate multifaceted access to the research network. But our continuing focus must be development of deployment models whose risk/benefit analyses promote and accelerate network implementations measured by wavelengths. To preserve scalability and the capacity for end-to-end design, the R&E

community ultimately must deploy networks where they will control the transport medium.[23]

III. Extensibility and Scalability – Drivers for Deployment

With this paper well underway, the telecommunications industry commenced its cataclysmic implosion, raising a legitimate question: can the communications infrastructure described above be built, let alone in such a way as to support the research and education community's purposes? As it turns out, the current communication revolution, born in the space race and nurtured for its potential military value during nuclear war, was driven once it became a "public good" as much by greed and corruption as by rarified abstractions. Creative accounting and cooked books partly funded this astonishing enterprise, enriching a few and, in their rapacious wake, impoverishing many. Can we build a qualitatively new infrastructure free of similarly destabilizing influences?

Happily, no! "Human" motives always drive progress, a point Stephen Ambrose[24] argues compellingly regarding the nation's first major communications infrastructure deployment, the transcontinental railroad. Construction corners were cut, shoddy work required redoing, and speculators made or lost fortunes while laborers lost lives. Yet despite – partly because of – the motley, unsavory activities it inspired, the transcontinental railroad was ultimately completed, with all its ensuing economic and quality of life benefits.[25]

That this infrastructure seeded communication change and shrank the nation can scarcely be challenged. Ambrose writes, "A man whose birthday was in 1829 or earlier had been born into a world in which President Andrew Jackson traveled no faster than Julius Caesar, a world in which no thought or information could be transmitted any faster than in Alexander the Great's time. In 1869, with the railroad and the telegraph that was beside it, a man could move at sixty miles per hour and transmit an idea or a statistic from coast to coast almost instantly."[26] In part because it foresaw this public good, the government played a critical role in the transcontinental railroad, as a project initiator and financial backer, just as it has with the network, most notably through DARPA and NSF. As we continue to develop, the government's role as an enabler and champion of the public good will not recede in importance.

---

[23] For a cogent recent discussion of the meanings of "broadband" and their implications, see *Broadband, Bringing Home the Bits*, Report of the Computer Science and Telecommunications Board, National Research Council, National Academy Press, 2001.
[24] Stephen E. Ambrose, *Nothing Like It In The World*, Simon and Schuster, New York, 2000.
[25] Ironically, writing *Nothing Like It In The World* seemingly entailed questionable practices, with some sections of this hugely successful book apparently borrowed. http://www.forbes.com/2002/01/17/0117ambrose.html
[26] Ibid, 24, pp 357.

How did the PSTN or Internet expand and enrich continental communication? To be sure, each enabled transmission and expanded deployment of a far richer information suite. Perhaps more fundamentally, however, each pushed information control closer to the end user, and protected the message in such a way that its dissemination became nearly unstoppable. The Internet made both accomplishments particularly true: we take as axiomatic that the network's coming uses will be so data intensive that fiber must touch the user whenever possible. This goal challenges us to both nurture network growth in such a way as to lower its participation threshold, and quickly leverage its economic, scientific, and educational benefits to maintain a vibrant development cycle. Avoiding prescription, we need to explore network development drivers from multiple perspectives, knowing that their interrelated, unexpected confluence will bring about a scalable, extensible, end-user controlled fiber network.

The λ-Based Network

Our challenge and opportunity, then, are to deploy a ubiquitous broadband network for research, education and the community, defining "broadband" at a level sufficiently high that it will scale and meet future needs. To achieve the greatest educational, medical, commercial, and economic development while driving down user cost, deployment must be widespread. In *Building a Positive, Competitive Broadband Agenda,*[27] the Information Technology Association of America presents a compelling case for broadband, and for content as the core value driving such infrastructure development. As with earlier network generations, the R&E networking community must press for this optical infrastructure's design and implementation, until market forces weigh in.

Going forward, broadband should be measured in nothing less than λs, with networks designed around this basic measure[28]. Today we envision wavelengths used to establish as-needed high-speed point-to-point connections (á la PSTN) within which IP supports the communication with a λ-based QOS. Such a model will scale in several ways: deployment of increasing amounts of fiber, management of more wavelengths over a single fiber and, as these two elements progress, commercial applications spurring economic development and reinvigorating both.

Such a network could presently be structured within a financial framework of carriers providing λ-networks over some footprint, with preexisting network and financial peering agreements between carriers. Within this paradigm, the Quilt is already capable of supporting a powerful precursor to a λ-based network, with Abilene one constituent. Ultimately, though, we should seek some standardized framework within which not just a

[27] Available online as a pfd file at http://www.positivelybroadband.com.

[28] Using λs as the minimal measure of network capacity is almost within reach, and this section in no way implies that this will comprise the basic unit of measure in networks to come. Advances in DWDM, or radical new methods of communication (e.g., using solitons) could reset basic expectations of network capacity. However, a scalable, self-incentivising model in any potential minimal unit can achieve broad deployment. The λ-pool is one such model, independent of the symbol λ's actual value.

λ-path, but its associated accounting, can be negotiated on an instantaneous, as-needed basis.

Pushing this notion to its logical conclusion, we arrive at a national (or global) λ- pool whose participants adhere to universally agreed upon specifications that include touching the rest of the pool at more than one point, and with some minimal radius. Participating entities (public, private, for-profit, prisons, hospitals, etc.) deploy and light fiber, and obtain the absolute right to establish λ-paths for communication within the global λ-pool. Path payment by entity X, like the path itself, is negotiated on the fly, and can come in the form of actual usage payment and path credit for use of X's fiber. Alternate route algorithms, perhaps partly stochastic, could ensure that deployed fiber would then be used, generating λ-credits for the deployer, increasing the λ-pool's route diversity, and encouraging the users' new deployment and development of applications.[29]

To initiate a contribution to the λ-pool may require a leap of faith. NYSERNet began its Manhattan project, a dark fiber deployment in New York City, without resources, understanding of financial and regulatory obstacles, or commitment beyond a tiny handful of theoretical participants. September 11 at once slowed the project and made it more pressing. Since that date it has grown rapidly, with Phase One now likely to connect most of the R&E community, hospitals, and many libraries, museums and cultural institutions in Manhattan and nearby. Though its opportunities and obstacles appeared similarly great beforehand, this project has now acquired life and energy of its own, providing an intense microcosm of a λ-pool, shared risk taking by the R&E community and industry[30], and the R&E community's anchor tenancy.

Similarly, west coast gigaPoPs used their own advanced network needs and expanding, diverse user base to design and begin deploying Pacific Light Rail, a λ-based regional network. (See http://staff.washington.edu/gray/talks/2001/wired-wireless.ppt and http://www.ucop.edu/irc/jog/cenic.ppt to see this project as a natural regional extension of local needs, and a motivating example for the λ-pool.) Pacific Light Rail's enormous programmatic reach – education, medicine, environmental agencies, public television, and more – has grown to become the National Light Rail project.

Qualitative Changes, Tipping Points, and Dr. Fleming

The λ-pool substantially lowers the threshold for full participation in a light-based network since, having made a specific bandwidth or fiber contribution, each participant enjoys access to the λ-pool's full wealth. Since that threshold isn't zero, however, many key participants will be able to make a cost benefit argument for membership only by

[29] Though technologies currently in labs make such a vision promising, the λ-pool must be standards-based. We're a long way from that since, for example, there are no universal standards for dense wave division multiplexing, a fundamental tool for this architecture.

[30] In this case, Lexent Communications, using the R&E community as an anchor tenant to justify building in a depressed market.

factoring in the quantitative and qualitative changes this network will facilitate. Its rich virtual presence could fundamentally alter health care delivery, remote instrumentation, weather forecasting and remote learning, and along the way shatter disability boundaries.

That coming network generations will allow such qualitative life improvements constitutes a potent motivator, one the research and education community should continually remember and leverage. Helping a broad community understand a regional fiber build's potential, for example, offers many advantages: valuable ideas this wider community germinates, new participants and, with that, lower cost per participant. Every individual and institution has a "tipping point"[31] - a complex of financial, social, and scientific motivators that shifts thinking from seeing barriers to eliminating them. Working as a community we will make participation possible for more, and each such success will expand both the potential for qualitative change and network participation.

Capabilities built into the network for scientific and educational purposes will be borrowed, modified, used out of spec for purposes unimagined – working real magic. As email and the web have both shown, technical capability changes qualitatively once it functions broadly and productively, driving the network's use and speeding its successor's emergence. To expedite this organic evolution, the networking community must articulate its capacity to many communities – to administrators who control internal support, the vendor community, current or prospective advanced network participants, and funding agencies.

Perhaps our most daunting task is to spiral network development symbiotically with its dependent applications, all the while maintaining objectivity sufficient to grasp when qualitative change occurs. Dr. Fleming saw that the "spoiling" of a bacteria culture by Penicillium mold was actually a long-sought solution (from which he ultimately extracted Penicillin), then held to that realization when others failed to understand.[32] We have been privileged to participate in the public Internet's birth, and sufficiently objective and fortunate to recognize fundamental changes it has enabled. With the end user enjoying far more power, tools, and control in networks to come, we must appreciate the qualitative change to be derived, and employ this realization to drive network development.

Application Ownership and Roaming

We sometimes rely on technical tools to ensure (or attempt to ensure) quality of service (QoS) – guaranteed bandwidth to support a critical application that cannot tolerate interruption or congestion delays. Some technologies have proved difficult and expensive to develop[33], while others, available and scalable to any bandwidth, play an

[31] Thanks to Hugh O'Kane, Chairman of Lexent Communications, for this notion and his keen business perspective regarding the network and fiber's role.
[32] http://www.pbs.org/wgbh/aso/databank/entries/dm28pe.html
[33] Attempts to bring native routing to ATM provide a good case in point.

ever-evolving role. As they mature and applications' bandwidth needs grow, over-provisioning the next network generation often provides QoS, only to have the network QoS tools reemerge in the higher bandwidth. Controlling the transport medium and pressing the capabilities of dense wave division multiplexing and other tools will permit the network to migrate to an era where we can "play" concurrently with many bandwidth intensive applications.

As noted, rapid networking and computing advances have resulted in insights regarding a host of core scientific problems, intimating the seminal progress that vastly greater computational and networking resources would make possible. Protein folding, precise prediction of severe weather, advances in genomics, and understanding our universe's first moments are but a few. Meanwhile, "everyday" applications, from distance learning to remote medicine or remote instrumentation, could expand so quantitatively as to render these differences qualitative. Projects such as the teragrid[34] are already pushing out this dual capacity's boundaries. Harder to predict, but as certainly and fundamentally influential, are new network uses enabled by the rich "presence" that a more capacious network will support. For precedent, one need think only of the web's impact.

The network's natural progression, and our necessary target, is its support for applications wherever they require such high capacity. Within certain broad parameters, an application must "own" the network assets it requires: we don't know when or where a tornado or the weather that spawns it will occur, and we do know that it will move. Researchers already cluster over shared data and interactive links, and will do so more powerfully when the network grows, but at constantly shifting times and places; similar in their need for intense, relatively brief network usage, distance learning or remote medical consultations are somewhat more predictable.

For expected applications, and for many we cannot imagine, the network must support a kind of high bandwidth roaming, with network assets available as needed, automatically or requiring minimal notification. Thus we advocate a network environment so capacious and extensive that it can always support such spontaneous uses. Such capability's wide availability will directly benefit society, and drive continued network growth. We discuss some aspects of funding in section IV below.

Content, the First Mile, and the Digital Divide

The very high bandwidth roaming described above is likely achievable only across a grid linking R&E institutions, hospitals, and instrumentation and perhaps expanding to include schools and other foci for concentrated network activity, in the process producing mass sufficient to lower per-user costs. But the network must reach every home, every person – the so-called first mile problem. Every possible solution of this problem's technical aspects - innovative uses of wireless connections, optical Ethernet, DSL over

[34] http://www.ncsa.uiuc.edu/About/TeraGrid/ .

existing POTS resources, cable modems, anything that genuinely and affordably extends the network's reach – is presently receiving intense attention.

With this paper's focus on the long view, we need not review or advocate a particular first mile technology. Some presently employed will likely shift to fundamentally different uses. As discussed below, PDAs' power will increase significantly, well beyond the cell phone/network access devices available today, and the 802.11 suite will likely continuously negotiate a user's presence/access to a land based network. In the long run, content, not technology, will drive first mile deployment.

Except in specialized areas, the networking community's function is not to generate content, but to ensure that the network meets its most vigorous users' demands for content delivery. We push the network's boundaries to enable those who do own content to experiment with network use, ideally driving its further development. For the R&E networking community, content ranges from "traditional" academic material to telemetry streams, dynamically shared databases, video, distributed computing processes and, of course, "recreational traffic."[35] The key to network development is not content per se, but its inexorable capacity for pushing the network and its tools, for the more robust these become, the broader and more creative the applications.

Consider, for example, the web's first mile effect. Developed in an academic setting to replicate for the network the point and click environment, it spawned an explosion of uses, mostly commercial. These having made broad network access desirable, AOL, other subscriber services, and recent fast versions like cable modems and DSL brought affordable networking to a wide audience. A student can reach web based course material from home thanks to high consumer demand for LL Bean, stock quotes, or sports news. A tool developed to enable research and education communication spawned far broader application, which in turn made an intermediate first mile solution generally available.

Though DSL or cable modem access hardly comprises a household norm, the above example shows that the R&E networking community's salient contribution to the first mile challenge results from supporting networking's highest level and its natural constituency's work. Research and education institutions, living at technology's bleeding edge and designed to absorb the experimental failures that precede success, will create and harden the very high speed networking tools that will see wider application and drive the fiber network's broad deployment.

Separation of the nation's population by Internet access, the "digital divide"[36], is in reality intertwined with other educational issues such as literacy and math skills. Our

---

[35] The graph on page 9 of Cornell's annual information technology report shows the top 25 ResNet users consuming essentially all the bandwidth (http://www.cit.cornell.edu/oit/Reports/) on a typical March day.

[36]A growing body of literature exists on this subject, much of it on-line. See for example the Department of Commerce site http://www.ntia.doc.gov/ntiahome/dn/index.html, and in particular the NTIA study on the Internet's use, with very interesting comparisons with other technology diffusion in their recent study *A Nation Online: How Americans are expanding their use of the Internet*, available in pdf format at http://www.ntia.doc.gov/ntiahome/dn/anationonline2.pdf.

role supporting the latter is the same as with the former; the networking community must stay focused on its high-performance networking mission. Though the government may decide to intervene on the first mile problem, networking can act most effectively by facilitating natural drivers through development of advanced network tools. The broad social and economic benefits that will derive from this new network generation hold potential to narrow the digital divide and also, with sustained commitment, what Rensselaer's President Jackson calls the "continental divide": destructive differences in access to education, to jobs, to opportunity.[37]

Life inside the black box – making the λ-pool work

This paper has deliberately been technically non-prescriptive because, within its intended five-year window, unforeseen technology developments will profoundly shape – and ideally be influenced by – network advances. Indeed, if past acceleration offers a valid comparison, the development perceivable at this five-year event horizon is slower than will actually occur. Even the λ-pool was proposed not as a technical solution, but as a socio-economically scalable framework that would foster a broad, light-based network environment's emergence within the research and education community.

In fact, this model for deployment of optical networks has materialized almost as a default. The Canadian CA*net4 network[38], designs considered in the evolving National Light Rail discussion[39], or alternative hierarchical designs with a national backbone and regional optical networks, each extended locally with metropolitan builds from New York City to Wellington, Palo Alto to Pocahontas, Jacksonville to Vancouver[40], all point to a λ-pool as a viable model for scalable deployment. If this broad pool of shared wavelengths proves itself the correct socio-economic driver, speculation regarding its near and longer term uses, and some of the hard scientific, mathematical, and technical issues inherent in continued scalable use, will be worthwhile.

Bill St. Arnaud describes a vision, achieved in Canada[41], of customer owned fiber networks, and an associated optical network architecture[42] with user control of the light path at the edge, ultimately to be employed across the CA*net4 network. In principle such architecture could be employed today (or at least soon), using tools such as GMPLS, were the same technologies deployed at every node. Going beyond that, prescriptive technology would be unnecessary, once universal standards for dense wave division

---

[37] http://www.rpi.edu/web/President/speeches/NACME.html.
[38] See the CANARIE web site http://www.canarie.ca/canet4/.
[39] See e.g. http://www.internet2.edu/presentations/fall02/20021027-HENP-Reese.htm
[40] The Digital Rivers Report gives a very nice presentation of the status of many metropolitan and long haul efforts, set against the broader context of telecommunications industry developments. http://www.digitalrivers.info/digital_rivers/pdf/Digital%20Rivers%20Report.pdf
[41] http://www.canarie.ca/canet4/library/customer.html
[42] http://www.canarie.ca/canet4/library/c4design/canet4_design_document.pdf

multiplexing had been accepted, and tools deployed to make use of the carrying capacity of wavelengths.

Ultimately, though, to preserve the network's ability to support new, unexpected needs, the light path's control must return to the network core, with end-to-end application requirements causing the core to respond, without directing it *how*. The same end-to-end principle that enabled the Internet's structural design to support the breathtaking suite of applications with which it presently copes ultimately must be invoked for future architectures.

Depending as it does on technology as yet undeveloped and hard structural problems still unsolved, any detailed architectural description would be premature. Yet a little speculation, coupled with intermediate architectural builds like that proposed for CA*net4, might elucidate the problems. The goal is to embed the "best effort, store and forward" Internet protocol's durability and flexibility in a λ-based network capable of sustaining huge, reliable data streams for many concurrent users.

One possibility may be to simply mimic but scale up the current network, breaking a full message into shorter λ-streams. Achieving for many concurrent users approximately the speeds possible in a continuous light path then requires storage capacity and speed well beyond current technology. Alternatively (or, perhaps, additionally), before sending a data stream the network could, rather like lightning, determine an optimal end-to-end light λ-path - thus presenting a queuing problem growing in complexity with the number of users and nodes faster than any known algorithms could handle. Or the queuing problem might be partly tamed by confining the sort to a pre-specified grid. Very hard mathematical and scientific problems present themselves every way we look, challenging and potentially rewarding the scientific and networking community. As one of General Patton's tank group commanders put it, "They've got us surrounded again, the poor bastards."

IV. The Role of Government and the Commercial Sector in Network Development

This paper's opening section traced the complex interaction among government, the commercial sector, and the research and education community that produced technological breakthroughs such as the transistor, laser, microchip, ever more powerful computing generations, and protocols that allow computers to communicate. When the Internet protocol that emerged dominant extended communication from computers to people, the network transformed itself into a powerful social, political, and economic force. Its very success now drives development of vastly more powerful networks. We now examine government's and the commercial sector's respective ongoing roles in that development.

Initiating and Sustaining Support for the Public Good

The Federal Government subsidized the Internet's early development heavily through DARPA/ARPA.[43] As the public Internet, with its obvious scientific applications, developed, the Government continued its funding through the National Science Foundation[44] and, as its broader uses came to be understood, through still more federal agencies. As with much of our scientific enterprise, these efforts do not constitute Federal projects per se. The government provides money for basic research, and public policy sometimes heavily influences these funding agencies' budgets and priorities.[45] Peer review determines the actual awarding of grants, however, brilliantly allowing policy to suggest but not dictate where to place scientific effort, and precluding it from evaluating that science's quality.

Federal agencies point proudly to the extraordinary leveraging for science and the public good of the fund they provide, with most development support coming from institutions or from subsequent commercialization. But without such initiating government subsidies much of this work would be slowed considerably, perhaps not be done at all. We propose in this paper creation of a network of vastly superior capability, with a correspondingly advanced infrastructure and strengthening of the hardware and software tools required for its support. To achieve success, this enterprise will need government assistance at the federal, state, and local levels. Clearly, support for basic research on the network itself, and for network research related to the policy issues our next section discusses, is critical. But full integration of the qualitatively different resources these future networks will incorporate also requires significant intellectual and technical retooling by researchers and educators who wish to effectively utilize their capacity, intelligence, and end user control.

Probably only grant assistance can catalyze that integration, and the new network's subsequent use across the scientific spectrum, which will drive and support its deployment. Such backing from NSF, NIH, Defense, Energy, and other government agencies could exert an even greater impact on network deployment than direct network support. Moreover, seminal advances will occur through this interplay among the network and the computing, research and education dependent on it.

The R&E networking community focuses primarily on scientific and educational applications but, appropriately, the network's profound economic and social impact will drive government networking policy at the federal, state, and local levels. Though this paper's focus, charting the course of the network itself, precludes its advocacy of a

---

[43] See http://www.acm.org/sigcomm/ccr/archive/1995/jan95/ccr-9501-clark.pdf for an interesting history, or http://www.darpa.mil/body/pdf/transition.pdf for a broader presentation of the DARPA mission.

[44] See http://moat.nlanr.net/INFRA/NSFNET.html for the early history.

[45] In her speech to the Engineering Deans Council at the National Academy of Sciences, NSF Director Rita Colwell outlines some of the impact of September 11's attacks on funding for basic research across the scientific spectrum, including networking. http://www.nsf.gov/od/lpa/forum/colwell/rc020212naecolloq.htm.

government stance, we must help make lawmakers at every level cognizant of its profound advantages and enable them to reach informed decisions.

It is similarly important to work with lawmakers on a state and local level to facilitate the network's deployment. In advancing the research and education network, our community can speak authoritatively about the technology's capabilities. But political and education leaders who represent the broader clientele we hope to serve as the network expands can inform us as well. They provide a critical source for contacts with prospective participants, facilitate issues like rights of way, and function as powerful advocates, once aware of the benefits advanced networking will bring.

Right of way provisions for a fiber build vary greatly by locality. Existing conduit systems, from New York City's venerable Empire City Subway to Boston's new, empty conduits provide well-established rules and fees that, if met, guarantee right of way. Other localities require that permission be obtained from property owners bordering the path. (CENIC faced some right of way ribbons so wide that one build had to be scrapped.) Intra- and interstate rights of way vary still more widely among neighboring states.

As with phone companies, local or regional officials ultimately will play key roles in any network deployment. This reality presents an opportunity/necessity for the networking community to establish long-term partnerships with state and local government, learning as we teach. These informed advocates, motivated by many of the R&E network's loftiest goals as well as by economic benefit, can amplify the arguments for the new network given above and will help determine its extensibility, one locality at a time.

The R&E Network as an Anchor Tenant

The commercial sector has played a critical role in Internet development, from providing basic transport for the R&E networks in local and long line circuits, to developing commercial markets that in turn drive down the cost of point to point circuits. As a network generation matures, it is subsumed into the commercial sector, with the R&E community building a next generation network on that improved infrastructure.

But we advocate building a network de novo, starting with the transport medium itself, in part because applications to come will require affordable scalability that only comes from control of the fiber optic cable. This brings the research and education community into new alliances, forming partnerships with the commercial sector for fiber co-builds, and acting as an anchor tenant on deployments. For the R&E community the resulting economies of scale will mean greater transport capability than they could afford alone, and for the commercial sector our anchor tenancy can provide the "tipping point" for a deployment they might not yet have done. This has already been the case in NYSERNet's Manhattan Project, components of the CENIC deployment, Indiana, Ohio, the Great Plains Network, and many other places.

As in our dealings with government, this codependency with the commercial sector to create infrastructure brings a concomitant opportunity to inform and learn, with cooperation among all three for their common ends a real possibility. Our previous section examined drivers for network scalability and extensibility from a variety of perspectives, many the natural realm of government, the R&E community, or the commercial sector. Close cooperation among these three forms the nexus from which a broad-based broadband network will inevitably emerge.

Consider these statements from a metropolitan viewpoint. Such areas will likely possess a concentration of schools, libraries, hospitals, and cultural institutions sufficient to induce them, together, to consider a metropolitan fiber deployment. Future network uses will require fiber to the user. Commercial sector visionaries who understand this need a motivation, a tipping point, to bring fiber into many buildings. Local government quickly grasps the commercial benefit to the community, the prestige and cachet of such a deployment, and the obvious benefit to the R&E community. They help with rights of way, perhaps funding, and a robust infrastructure is created.

Fiber Capacity and Maintaining the R&E Role after Commercialization

The R&E community's shift to control of their transport medium both stabilizes and accelerates the Internet development spiral. Till now, research networks have been overlaid on essentially the wrong infrastructure, and successive network generations brought an escalation in circuit costs that slowed, and for some institutions prevented, participation. Fiber's enormous capacity means that the research and education community controlling this transport medium can make quantum and incremental changes needed by any size group of institutions, allowing applications, not the time required to provision and get payment approval for a circuit, to determine development's pace. Indeed, this endows the network with a dynamism closely resembling a laboratory's or classroom's: it can now truly become a fully responsive tool.

Because they will depend on separate, perhaps co-built, infrastructures, research and commercial networks can progress independently. The R&E community can commit not only to a full spectrum of network-supported applications, but to support for seminal uses that stress the network with a richer suite of network tools and protocols. In some of the excess capacity, the community can experiment on the network itself. The teragrid[46] has offered preliminary glimmerings of this; our dream now is to accomplish an extensive, high capacity, broad based network, with a dynamic, flexible response to demand.

The commercial sector can learn from the R&E community's differentiated networking model, hardening "best of breed" developments on protected portions of their own infrastructure before broader usage. This couples the Internet spiral's components more closely, encourages experimentation and progress on multiple fronts, and, potentially, dramatically accelerates concurrent developments.

---

[46] http://www.ncsa.uiuc.edu/About/TeraGrid/

Building the Next Generation's Network

These last two sections focus on drivers for creation of a new breed of fiber-based network, and on the roles of government, the commercial sector, and the research and education networking community in furthering that goal. But now let's look at their constituents. Typically, a business can handle all of its data needs with a T1 line, reflecting both rudimentary tools and limited data traffic.[47] Since the government's constituency is all the people, it necessarily focuses on broader social concerns, economic opportunity, and quality of education, in all of which the network and the digital divide form one small component. Universities, by contrast, push the boundaries of available network resources in anticipation of their research and education needs, then quickly fill this expanded network capability as intended, and also in many clever, if less desirable, ways.

Yet a curious symbiosis exists among these divergent constituencies. Born in the research and education community, the public Internet could not have accelerated without NSF support for connections, and for research that led to network tools that in turn produced the explosion in Internet usage. Those tools proved themselves so flexible and adaptable that their commercial applications encouraged deployment of cable modems, DSL, and widespread free access tools like hotmail. Students entering universities have never known the network not to exist, and many of them enjoy high-speed access at home. So now, "the wheel is come full circle."[48] Universities upgrade dorm access to match what students have at home, receive support for their network dependent research, government cooperation with issues like rights of way, and observe the magic that bubbles from these creative minds. Moreover, these students graduate, move up that business ladder, and wonder how they can possibly get by with just a T1 connection.

Perhaps that is the correct way to think about the Internet development spiral. Our mission is not to build the next generation Internet, but the next generation's Internet - what an honor.

V. Policy issues regarding the network

As the network and the tools it supports become more sophisticated, and our dependence on it grows, the policies governing its development, and our thinking about its political, social, and economic impact, must keep pace. We outline some of those here, in broad strokes.

---

[47] Reflection of Hugh O'Kane, Chairman of Lexent Communications

[48] Edmund speaking in Act IV, Scene 3, of *King Lear* by William Shakespeare.

Security

The true urgency of this long-standing problem was revealed September 11, when the Internet, designed to withstand physical attack, endured a severe test of its capacity to do so. Though NYSERNet's network ran through ground zero, it survived, thanks to its staff's ingenuity, its redundant design, luck, cooperation by vendors and the broader networking community, and phenomenal bravery by people working near the WTC.

Today security - physical, logical, layered - actively concerns the Internet2 and vendor community. Like airplanes, the network's increasingly core importance renders it an alluring target for primary / societal attack. Thus as we grow the network, we also increase the responsibility to safeguard it.

Though physical security is a problem with easy theoretical solutions, they are all costly, with redundant paths, equipment, NAPs and NOCs key ingredients. The best approach to physical security may be the broadest, since economies of scale much like the Quilt's on commodity network pricing could thus be realized. The Quilt's extraordinary internal cohesion, sharing of insights and broad buying power make it a paradigm for deploying a network with essentially unlimited bandwidth potential, very broad footprint, and high redundancy.

Supporting this collective stance regarding physical security, the network behaved remarkably well last September 11 despite collateral damage suffered in the World Trade Center attack. The protocol itself, the agility of people managing networks in or touching New York City and its vicinity, some fortunate pre-existing planned physical redundancy, SONET, and the richness of telecommunications deployed in Manhattan - all these made for a rapid, effective response, with little or no cascading failures on the global network.[49] September 11's assault was not aimed at the network itself, however, and the threat posed by direct attack is ominous.[50] In many respects, the network behaved normally on September 11 even in New York City, and it recovered rapidly, particularly when weighed against the terrible physical damage it sustained. The Nimda virus a week later had broader network effects, and even "friendly" network events like Kazaa can significantly impact performance and expense.

Above the transport layer security becomes a more complex problem, and must remain an area of active research – particularly as we become more network-dependent. We have become increasingly adept at "routine" problems like denial of service attacks, though a

---

[49] An upcoming report of the Computer Science and Telecommunications Board of the National Research Council will address the impact of September 11's attack on the network.

[50] A considerable body of literature exists on network security. See for example, *Trust in Cyberspace*, National Academy Press, 1999, a high level report by the Computer Science and Telecommunications Board Report for the National Research Council.

recent DOS attack on DNS root servers was far more effective than we'd like[51]. With more people actively trying to disrupt the network, or break in and gain access to data, and with tools that facilitate such efforts routinely available on the web, network operators face ever increasing challenges to information security.

After September 11 discussion concerning impact should the network itself be attacked and crippled by disrupting transport or by logical attacks on the network core, intensified. But the true analogy, and perhaps our greatest security test, is not that the network might be destroyed, but that it might somehow be commandeered for use as a weapon. Given our broad network dependency, the impact of even subtle alterations to it on vital functions that it supports could be profound.

Future networks, many relying on optronics not yet developed and facing a host of security challenges, should be built and managed with security designed in at the outset and, additionally, able to deploy new security measures as rapidly as they become available. The λ-pool potentially can generate a grid with sufficient richness that multiple redundancies effectively protect the physical path. The network core's security depends first on prevention of inappropriate access by obvious (e.g., never use NCC1701, the Enterprise's call sign, or other easy-to-guess strings, as root password) and subtle means, and by constant monitoring. One aspect of greater embedded network intelligence could be agents within its core capable of identifying and eliminating, or at least sounding an alarm about, untoward shifts in network operations instigated by outside influence - with corresponding issues of trust in these very security agent. With the network now fundamental to nearly every facet of everyday life, its safety has become correspondingly vital.

Future research networks should incorporate the capacity for ultra-security, to provide command and control and vital facilities in the event of attack aimed at or by the network itself, and as a locus for research on network security. Ideally, security researchers should coordinate closely with government and industry, jointly deepening awareness regarding security issues, and the tradeoff between fixed hardware changes to improve security versus remote network control (vital on September 11).

Portability

Perhaps the new network's innovative, unexpected uses will emerge only once an individual's network environment becomes truly portable – i.e., resides on the network, and so is fully available wherever the person enjoys access. Email and AOL already offer one rudimentary version of this entrée and, in another sense, so does the web. The interface and tools are sufficiently intuitive that all one requires are a keyboard and

---

[51] http://www.icannwatch.org/article.pl?sid=02/10/21/193354

screen attached to the network; this very portability has been responsible for the network's decisive social and behavioral effects.

Using 802.11a and its possible successors, the network's wireless extension will profoundly shape the nature of portability. Personal devices will continuously mediate an individual's network contact and access to local hardware, and transfer processes from one access point to another.

Supporting personal, portable network environments presents daunting technical challenges and, until standards are set and portability tools broadly adopted, equally complex ones of cost. From the outset, this degree of portability (and the network's necessary awareness of individuals) also raises difficult security and privacy concerns. Though it is not our task to resolve, or even address, these control issues, building appropriate security tools and providing full, precise information to those charged to do so, are.

In a second sense, "portability" relates to grid computing or massive, time sensitive network applications. For distributed computing processes, the network must possess sufficient intelligence to keep track of each, and also time stamp data going from one component precisely enough to transfer an entire process synchronously, when so directed. In this sense the global computation resides on the network itself, with computers or other site-based devices subservient to this network based process.

To test these concepts, the National Science Foundation has sponsored an initial test bed linking the San Diego Supercomputer Center, California Institute of Technology, the National Center for Supercomputer Applications in Urbana Champagne, and Argonne National Laboratory. More extensive near-term applications will involve CERN, Brookhaven National Laboratory, and the physics research community. Ultimately, the network must support pervasive next generation grid applications that, as clearly, will require multiple λ-pools.

Remote instrumentation provides an example of portability almost within reach. Remote access of a telescope or functional MRI will rival on-site control only when the network can provide the massive, real-time return provided by direct sensory input. Use of such control for medical intervention raises the stakes enormously, with total network security and reliability required. Medicine based on network interactivity also demands tremendous scalability if, in response to the patient's evolving condition, additional instrumentation and related feedback must be brought on-line.

Specific important applications not yet within our scientific grasp go well beyond grid computing. Atmospheric scientists have long worked toward accurately predicting severe weather events, using the latest generation radar and other newly available telemetry devices. Today's radar can already generate more data than atmospheric scientists know how to fully process on the fly, more than current networks can carry.

Even assuming no further improvements in radar, using it to predict a tornado and its path accurately is enormously difficult. Although the atmospheric physics is well understood abstractly, detecting *well in advance* from a single Nexrad site remains in its infancy, partly because the current network doesn't reach those sites at any reasonable speed. In principle the information would be better from two sites, still better from three, but the complexity of the mathematics, physics, and computer science in coordinating data from n sites, even if we knew how to interpret it, grows more than exponentially in n. And storms move. Were monitoring to be based on the three nearest sites, for example, network support for such an effort would require rapid reallocation of enormous network resources – it would need a λ-pool.

As a final example, consider the NSF-sponsored National Ecological Observatory Network, or NEON (http://www.sdsc.edu/NEON/mar2000/index.html). Initially, NEON consists of about a dozen ecological observatory sites capable of "rapidly" sharing data via the network. Following the natural progression for such a project, however, the network would become the observatory, with much raw processing of enormous amounts of telemetry done on the network itself.

Privacy

As with network security, a growing body of scholarly literature, legislation (not entirely consistent) and case law relates to privacy. HIPAA (covering medical records), FERPA (the Buckley Act governing educational records), and associated completed litigation broadly impact research and education. While the networking community has experts exploring privacy-related issues, and maintains regularly updated references on the EDUCAUSE web site,[52] legislation and the courts must sort through these issues. Though this document's concern remains the network itself, not articulation of a position on privacy, proposed solutions regarding security and privacy necessarily possess both policy and technical ramifications. The need to make the network technically responsive to policy to the extent possible, furnishes yet another argument for preserving end-to-end capability for design and implementation of network applications.

Our community must also track security and privacy issues policy makers consider, offer expert opinion and testimony, and carefully explain to a broad audience the technical implications inherent in each, so that lawmakers and the courts reach informed decisions

Configuring the Network/Network User

One probable component of portability will be the ability of network users to largely prescribe their network environment. As people interact with a new level of "presence"

[52] http://www.educause.edu/asp/doclib/subject_docs.asp?Term_ID=188

supported by the network, this in itself will lead to unexpected applications. In some sense, this represents simply a rich extension of the current Internet development cycle – the "configuring the network" half of the above title.

Perhaps the most interesting new network capabilities will derive from user/network interactions once the network becomes able to gain knowledge.[53] Educating ourselves from network "insights," while the network is learning from us, should pose quite a challenge. Young network users are already shaping their information acquisition habits rapidly, not always in ways we like, as network capabilities and tools evolve.

And how are security and privacy to be handled? Although we regard the privacy of an individual's lawful interactions with the network as axiomatic, should the network be able to apply its own insights in other private interactions? Despite occasional breakdowns in such relationships, trust with transferable insight is core to interactions with a physician, lawyer, or clergyman,.

So perhaps our biggest challenge as we look over the horizon is to be like children, play with the network we create, interact with it openly, and be open to realize what happens.

Organic Network

We need to stay focused on the network itself. Success in that endeavor will spin off tools for the user community and unexpected opportunities for K-12, narrowing the digital divide, providing remote health care or unimagined new applications. The next Internet iteration will organically benefit these: a λ-pool with its built in incentives could improve rural access, as a natural byproduct of an essential network attribute.

The fact is, we can't identify with certainty underlying, long-term issues in, for example, K-12 education. Over the next five years, while universities produce only a tiny fraction of their needed replacements, a majority of our secondary science and math teachers will retire. Although the network will play an important role in addressing this crisis, we don't yet understand precisely how. Computer- and network-dependent reading programs for high school students at risk may well permit the substitution of thoughtless computer manipulation for disciplined reading; it is uncertain what that will mean long term, and it would be risky to try to anticipate the network's role too specifically.

The new network needs to be free to grow where needed, to support experimental, robust interactions set up on an ad hoc basis, thus providing a living medium for all of us to experiment with communicating. In partnership with its managers, the network needs to be able to learn. If atmospheric scientists make progress predicting severe weather, the network itself needs to learn from observation, to anticipate patterns that become

[53] The description of the λ-pool above, with some stochastic assignment of paths (there proposed to support the cost/benefit analysis for a prospective λ-pool contributor), combined with the network's ability to learn the efficacy of certain pathways, gives real potential to joint user/network learning experiences and unanticipated network intelligence.

familiar, to grow and respond intelligently, to interact with those that manage it. We in turn must understand how that network learning occurs, to sustain network trust, security, and privacy.

Involvement with Government

With the researchers and educators they support, networking community members must actively involve themselves with government. Though technology, broadband to the home, and their economic consequences are all actively discussed at federal, state, and local government levels, most officials concerned gain their understanding of these ever-evolving issues from others.

We must help advise them, in the face of fierce competition in that effort. Technology companies, and organizations and agencies benefiting from government money that might more appropriately further the goals in this report, may view government's support for technology through tax policy and direct funding very differently.

Tax policy can exert profound effects, from promoting rapid deployment of new technologies with informed decision-making about efficacious spending, through stifling devotion to the status quo, to erecting obstacles to new development unless, and until, it can be taxed. This report promotes partnerships with our user community and with industry to accelerate deployment of advanced technology; we must likewise foster efforts by our industrial and network partners to inform government as it formulates policy.

As the network grows in power, sophistication, and intelligence, likely far exceeding the ideas in this paper, the networking community must assume an even more fundamental responsibility to government and society. It must honestly assess the network's current and anticipated impact, not merely from technical, scientific, and economic perspectives, but from social, political, and moral ones as well. Will development of a new network capability carry greater potential for harm than good, and if so should such work be terminated? Should we allow networks to become intelligent, how will we monitor the impact, and control it if it proves harmful? The community must be just as forthright about the positive effects.

Involvement with government is critical at the federal level, but even more so at the state and local level, where dramatic variations in understanding and commitment exist. Just as government and the communities they serve should enjoy access to the networking community's expertise, we in turn need to hear and be influenced by their ideas, needs, and concerns.

The Network and Government

We have repeatedly visited the theme of the networking community's interactions with government, helping inform the decision making process, advocating for network support, and cooperating with local government on specific issues, e.g., rights of way in fiber deployments and exploitation of the network's full economic potential. But the most profound impact, which calls on every element in this paper, may accrue from the network's increasingly pivotal role in the process of government itself.[54]

If one views government as a business, then it has reacted just like the business community to this new tool. A decade ago there were two email addresses on Capital Hill. Today every Congressman and Senator, every state and governor has a web site and, come election time, a campaign web site. The web has proved a most effective delivery mechanism for delivery of forms for, say, the Internal Revenue Service, or Motor Vehicles[55], and, increasingly, for direct, on-line interactions.

Voting, by contrast, often remains primitive. The last presidential race, which turned on "chads" in a few Florida counties, has accelerated the process of modernizing voting equipment, but these changes remain incremental and network shy. As the New York Task Force on Election Modernization reports[56]:

> *Internet voting has been used only on a limited experimental basis.[57] As Jacob Myer's mechanical lever machine did in the 19th century, some believe that the Internet will eventually revolutionize the way we vote in the 21st century. Others do not. Internet voting is merely another technology. The same concerns that apply to other technologies also apply to the use of the Internet for voting:*
>
> - *Adequate safeguards must guarantee a secure ballot, that votes will be counted as cast and that the person voting is the one entitled to vote.*
> - *The system must be secure against hackers who may wish to disrupt the system or intercept and change votes.*
> - *Internet voting must be at least as secure as absentee voting.*
> - *The public must have confidence that each vote will be counted accurately and fairly.[58]*

[54] Thanks are due to Jay Blaire for reminding me of this potential network function.
[55] Links to DMV sites worldwide can be found at http://cache.cow.net/~friedman/dmv.html.
[56] http://www.state.ny.us/governor/
[57] In what was reported as the first "binding election" using the Internet nearly 40,000 votes were cast on line in the Arizona Democratic Party's Presidential Preference Primary in March 2000. In the November 2000 general election in a Department of Defense pilot program, 84 overseas military voters voted by Internet. (Footnote quotes verbatim from the Task Force Report).
[58] Internet Voting Overview, A Message from R. Doug Lewis, Executive Director of the Election Center (http://www.electioncenter.org/voting/InetVotingOverview.html).

These concerns, a subset of those outlined above, will be solved as a consequence of network evolution. But we disagree with the possibly inferred view that the network is "merely another technology" in government. The network is a fundamentally democratizing tool, particularly in a society sufficiently committed to education that all can use it meaningfully. Employing the Internet as a voting tool is an incremental step, as are congressional web pages. But letting the network's uses evolve so that they pervade government constitute true revolution, potentially permitting ideas, vetted and tested in a light based public forum, to drive choices about our leaders and our laws, freeing us from the crushing influence of position and money.

VI. Conclusion

In his valedictory annual report (http://www.ibm.com/annualreport/2001/home/index.html), retiring IBM CEO Lou Gerstner describes the next wave of technology development as driven in a fundamentally different way, this time with the customer's assertion of ownership. Mr. Gerstner's warning is apt. As we envision networks to come, we must remember that their magic, breathtaking new tools, social benefits and, ultimately, the financial model supporting our efforts, all derive from the interactions and interplay they enable. Exciting though they may be, maturing of IPv6, grid computing, the network serving as backplane, high-resolution multicast, remote control of telemetry devices with real-time feedback of the output – ultimately, these are just tools. Those tools support interactions representing broad experience and expertise among people and their resources; from this crucible alone emerge qualitative new ideas that benefit society. Our task is to listen, understand, measure, and perhaps anticipate that seminal community's network needs, letting their work drive the next generations of networks.[59]

[59] Special thanks are due to Charles House (Intel) and Anne Dunford (NYSERNet) for important contributions, to Mark Luker (EDUCAUSE) for the enhanced development spiral and other ideas in that presentation, and to Bill Owens (NYSERNet) for many instructive discussions and for helping me understand the end-to-end argument. Thanks also to many who worked through earlier versions, and especially to my wife Anne for her close copy editing of several drafts.