Predicting the Unpredictable – Future Directions in Internetworking and their Implications[1]

David J. Farber
Distinguished Career Professor of Computer Science and Public Policy
Carnegie Mellon University
Institute for Software Research
5000 Forbes Ave, 5305 Wean Hall
Pittsburgh, PA, 15213
Tele: +1 412 268 5711
Cell: +1 412 726 9889
email: dave@farber.net

I have avoided discussions of network neutrality, network management etc as I have widely published on these subjects as well as widely discussed them in my IP mailing list. Rather I have tried to outline why it is so important not to freeze the Internet in a political/regulatory straitjacket that will preclude the rapid evolution, and even dramatic change, of the net – its hardware and protocols. The use of TCP/IP is not etched in stone. It was a brilliant response to a particular blend of technology and capabilities at the time it was created. As those technology shift slowly, it has been surprisingly adaptable. But dramatic changes, such as those outlined in this paper, may stretch it to the breaking point. At this time, we just don't know, we can just caution it is most likely the case.

It has always been dangerous to say "that is enough bandwidth". When we did the Gigabit test bed we were constantly faced with comments like – who needs gigabits? Now it is who needs 100 gigabits? It is like the film – build it and they will come (and ask for more). The future explosion of video, not only for TV like uses but for Virtual Reality environments – like Snow Crash will generate demand for bandwidth that will exceed what we have installed especially in the first mile.

And finally, the Internet has stretched our understanding and interpretation of our laws. In the USA the interpretation of the Bill of Rights in the electronic ago has created stresses. We must relieve the stress in our society in such a manner as not to freeze the ability of the network to evolve, improve and maybe someday be replaced by it's child.

The Future

I intend to explore the likely networking technologies that will mature over the next decade and the implications on other fields of information systems and society. I have chosen to restrict myself to the ten-year time scale since predicting outside that that time frame is likely to be writing science fiction. This paper addresses two major themes. The first explores at 10000-foot level the major technology changes that one expects to bear fruit over the next decade and the second explores the implications of these changes on our society.

---

[1] This document is a updating of an invited article for the IEEE Communications magazine in 2003.

Since the time the talk on which this paper was based was given, the world has changed. The need for security and robustness of our electronic infrastructure has dramatically increased and our understanding of how to do this unfortunately has not kept up with the need. In many countries, there is an increasing strain between the privacy desires of its citizens and the desire for safety. In many industries, such as music and motion picture, the rapid expansion of digital content and the ease of distribution of that content had created demands for increased protection much of which must be implemented in computer hardware and networking protocols. How this will impact the perceived right of the public to fair use of that material is a battle being fought in the legislature and the courts.

First let me explore the technological directions of the next decade.

.

Ultra high speed all photonic networks

With the steady progress in the technology that underlies photonic networks, I use the term photonic rather than optical to emphasize the end to end nature of these networks and to distinguish them from the electro-optical networks currently in use, it now feasible to predict with some certainty the future directions of this technology. We can expect that photonic networks with speeds in the 100-gigabit per wave (color) and number of waves in the fiber in the 100 ranges are not beyond the rational dreams of technologists. The inclusion of optical amplifiers makes it feasible even now to remain in the photonic domain for continental distances. Current work in all photonic switches strongly suggests that we may, within the next ten years achieve end to end national communications systems in which the signals remain entirely in the photonic space. This raises a set of very interesting questions. Some of them are technical; some economic and some societal.

In the latter part of the 20^{th} century, even low gigabit speed networks raised a whole new set of very difficult technical issues. Designing and building switching devices and interface devices, which can operate at these speeds, was not simple. It pushed both hardware design and VLSI technology to their limits. As a result, it was necessary to take innovative architectural approaches to even hope to achieve adequate end to end speeds. Perhaps most interesting, though, is the conclusion that many of the ideas developed over the past twenty years in computer architecture, operating system design, and networking protocols seem to be ineffectual when applied to such speeds. It is worth observing that these communication speeds are of the same order of magnitude as the main memory bus speeds of modern workstations. Thus it is not surprising that we have run into problems. In addition to issues raise by the bandwidth/latency product of such systems, when streams of data arrive at memory speeds, it becomes difficult, given the protocol systems currently in use, to get the data into memory, to allow the processor enough processing bandwidth to examine the data and move it, and still to have processing power left over for other tasks.

I will not elaborate on the possible solutions in this paper but one of the possible solutions the author has suggested, revolves around the creation of a geographically dispersed distributed machine, the components of which would be interconnected by high-speed photonic networks.

What is more important than a particular solution is the challenge of facing a future in which 100 gigabit speed networking will be considered slow, in which our communication infrastructure will consist of multi-terabit, low error, high-latency photonic networks, in which our processing units, while growing faster, will not keep up with increasing communication speeds. It is too easy to just suggest that we remove a few instructions, hack a few cures, and show that one can operate not too badly at current speeds of communication. Perhaps this is equivalent to saying, "let the next generation solve the problem."

I believe that there is a challenge facing the computer communication field of at least the same magnitude as the challenge the field faced in the very early days of networking. Attacking this problem will require the talents of people from every area of both the computer and communications fields -- people willing to experiment and willing to face the same set of challenges that those in the fifties faced with the then-new computers.

The major impact of all this new thinking will be on industry. The changing of technologies forced by these ultra high speed photonic systems will expose existing industrial leaders to the same pressures and dangers that main frame manufacturers of hardware faced in the 70s. Many of them failed to respond to those challenges and several of the then leaders are no longer forces in the computer business. Attempting to continue to market buggy whips in the era of fast cars is a formula for becoming obsolete. The same is true of software. As the focus shifts from large slow, by future standards, operating systems and individual software applications to the lean and mean systems required by the future high-speed communications, many companies will be come obsolete. Applications will tend to become much more distributed as communications capabilities expand and costs, and maybe even price, decrease dramatically. Such changes will upset the current national industrial strengths and will push the need for increased numbers of IT trained people to compete in this New World. Countries which are fast to adopt, have a pool of investment capital and have a pool of flexible trained people will most likely emerge as the leaders.

When one adds to the photonic backbone infrastructure the slower speed but still multi-megabit data rate technology capable of delivery to the last kilometer— to the home and business, dramatic changes in the communications world also becomes possible.

Bits to the Home/SOHO (small office/home office).

Over the past decade, a major change has started in the first kilometer connection area – the so-called local loop/last mile… Prior to that time, the notion was that the last mile was provided by either the telephone or cable TV companies. In both cases, they were monopoly-regulated services that offered little or no data services beyond what was available via analog modems. While ISDN (integrated services data network) was in theory deployed, for a number of reasons it failed to be a major player. The reasons were both technical – ISDN still used the telephone local switch and tended to overload it and economic -- prices were held high and deployment was sparse and strategic – namely a fear of obsolescing the legacy profitable service.

Relatively recently two alternative technologies have been used to reach the users. These have been cable based data services and the fiber based FIOS systems being deployed mainly by the former telephone companies. It should be noted that there has been a great deal of justified criticism of the telephone industry for deployment mainly in rich suburban areas leaving the city centers not served. They both offered the potential of always connected multi-megabit two way data services to the home at prices that were affordable for certainly the SOHO area and many home customers. While the capability to provide these services are there, business strategy and some technical issues have prevented the materialization of the high speed, 10s of megabit per second, use of these facilities. The growth, however, of broadband access to homes and small business has dramatically increased with adoption rates matching those of black and white television. This has been helped by the increasing capability of the personal computer, Intel and Apple based, to supply the computing cycles to process the broadband input as well as a steadily increasing amount of video and audio material on the commercial internet.

The above technologies, as well as the wireless technologies soon to be discussed, create the possibility of providing significant last mile access to the future photonic wide area networks. A note: there is the possibility of optical connectivity directly to the home However there are potential legal/regulatory roadblocks to such deployments as legacy industries attempt to slow their introduction especially if municipally supplied. However, if it is feasible to widely deploy them, then my following argument just gets even more interesting.

What will be the impact on the public and industry of this enhanced last mile capability? Again it threatens to attack the business base of several important industries. Those most susceptible to economic damage include the video aspects of the cable companies and the voice aspects of the telephone industry. Reacting to these threats (or opportunities) will require large capital investments in a time of increasing price pressures.

While some industry is threatened by new technology for the last mile, another technology is arising via wireless systems in the world/city and house.

Wireless and mobility

Wireless technology is one of the fastest growing areas of communications technology. For many years we have been offered the vision of being connected via a digital path at all times. Unfortunately for a long time the technology, regulators and the market place did not combine to deliver the promise. There are several new technologies coming into the marketplace and laboratories that may radically change the wireless world. First, and this is US centric, there are market forces at work that are changing the way wireless is used. The bulk tariffs that are now offered by almost every cellular carrier have changed the balance between the wire world and wireless cellular service. In many cases, especially in urban areas, the need for a wire line telephone is decreasing. The costs of wireless with a large allocation of talk time, about 600 to 1000 minutes per month with no toll charges and no roaming bring often the cost of cellular below that of wire service. That will have two effects. One is that again the “cream” market will be skimmed off the wire telephone service and what is left is the low margin “lifeline” service and second it will force the costs of conventional toll down to a level to match that of the flat rate cellular – that is happening already. All this, at least at a nation level, will bring reduced costs and increased capability to the public and the advantages of wireless systems in emergencies, storms etc.

The role of data in the current generation of cellular wireless systems was at best secondary. However the next generation cell systems, 3 and 4G, talk about data rates that could hit multi-megabit per second. This will greatly enhance the use of data services, as the mobile devices become extensions of the home and office computer environments. This is happening already in many countries that got an early start on enhanced mobility.

.

Home and Personal networks

The wireless world is often focused at mobility in geographic distances such as cities, states etc. The advent of short-range very inexpensive radio links such as WiFi and very short-range systems like Bluetooth allow us to talk about mobility in the home and on the person as well as in the neighborhood.

WiFi, also known as 802.11[2], wireless technology utilizes unlicensed, at least in the USA, spectrum that has been allocated worldwide to this use. It has limited range, roughly covers a house, and is easily adapted to a pseudo cellular environment with the possibility of roaming from "cell" to cell utilizing the mobile IP protocols. It can be used to distribute data within a house from the data entrance to the printers, computers etc. It is becoming widespread in airports, hotels, neighborhoods etc and has spawned several "isps" that offer access to a variety of commercial and free services.

On a more limited range basis, the very short range wireless technology represented by Bluetooth[3] opens up a wide and exciting variety of applications. Essentially everything in a house can now be equipped with network access. Personal devices such as watches and calendars can communicate to other systems in the home and office. An ID card can become an active badge that identifies the user to the environment they are in. Locks can open, lights can come on and temperature can be set to an individual taste when they enter a room. (As a side comment the badge can require a PIN prior to becoming active so as to insure the rightful user is using it.). TV sets can restrict the program material depending on who is viewing. Thermostats can sense occupants, outside weather etc and properly adjust the house saving significant fuel.

More important is the ability for devices to organize themselves into interesting new systems. For example, a PC equipped with a Bluetooth can detect that the PDA (Pilot like personal device) belonging to me (via use of my personal ID card) and update its data base with new appointments etc. The possibilities are huge and will start a new round of innovation at a much more personal level that that of the PC.

As with each of the technologies I have mentioned, there is a societal impact. The privacy of the individual will be constantly under attack. If I have an active ID card which constantly announced to all that will listen, or all that will ask, who and where I am, there is an excellent opportunity for massive violations of my privacy. Each of these threats to the individual has a corresponding technical solution and unfortunately the potential for these solutions to be blocked

[2] http://grouper.ieee.org/groups/802/11/
[3] http://www.bluetooth.com/

by national laws. For example, most of the applications envisioned by such wireless links would greatly benefit from or require strong encryption. Yet in many nations strong encryption is banned or controlled. Also in many nations they ban the export of such technology and thus damage the marketplace for devices with that capability.

Robust and secure networks

In spite of the fundamental design of the packet network and its ability to route around failures etc, there have been addition demands placed on it as the data networks become critical infrastructures in the world. There have been numerous attacks on the Internet and all communication systems by “hackers” and national players. There are attempts to penetrate these networks to gather information or to disrupt the activities on them. The Internet has grown out of an experiment in networking and many of the protocols were not designed with cyber-attack in mind. Just as is true with computer software, the reaction to attack is patch, patch, and patch. It is very difficult to fix the security of any system after it is deployed if security is not designed in at the beginning. It is critical that we make sure that in future systems; we pay attention to these issues in the beginning. Our cyber-infrastructure is just too critical to our economy, defense and peoples to not do so.

Computers

While I have focused my attention in this paper at communications technology since that is the arena where the most change will take place in the next decade, it would be inappropriate to ignore looking at computers and software.

The end of the exponential expansion in VLSI chips has often been predicted based on limits of line width etc on semiconductor devices. Each decade we predict the end of increased capability due to these effects and each decade it keeps growing. At least for the next 5 years we can expect increased speed and capability out of our microprocessors. Architectures that used to require rooms of equipment are now in a small chip. Further we see yet again the blending of software methodology with hardware architecture as computers that require complex and capable compilers are entering the marketplace. We can expect to see this sophistication increase and begin to see optical interconnects between components to overcome the limits of wire connections. Further we will see increased use of multi-processor architectures to gain increased performance from commodity components as well as increased capability to provide security on the chip. The limitations on computing, at least at the non-supercomputer end, will come from software.

A major limit to innovation has always been our ability to produce reliable working software. The scale of software systems enabled by the modern computer and communications systems has stressed our abilities to create such systems. This problem started in the 60s and has gotten worse. Further such systems have embedded themselves in critical tasks often managing life-supporting systems. In spite of the major efforts by some of the best research laboratories, this lack of a viable software engineering methodology, applicable to the scale of software experienced in modern systems is currently the major limit to the greater infusion of computers and communications into our society. The USA in its Information Technology research initiative has devoted significant resources in an attempt to solve this problem. However the size and complexity of systems may out race our abilities, under the best of circumstances, to get software under control.

Repeated failures of key software systems will have societal impact on people and nations. This impact ranges from the loss of credibility on the part of the public for computer based devices with increased pressure not to continue our technologically enabled society up to and including massive damage done by nuclear plants and missile control systems which fail.

Some social comments

While I have been asked to address the technical issues and their impact on society, I can not finish this paper without making some social comments. These comments originally were made at a meeting of the AAAS in 1996 and are modified for this paper.

Cyber-rights

John Perry Barlow is credited with having observed that *"Our Bill of Rights is but a local ordinance in Cyberspace"*[4]. He was referring to the fact that the basic rights which "we hold self evident" in the USA are only self evident to our society and are not accepted worldwide. Similarly our notions of morality, law, right and wrong are European centric and are not accepted uniformly worldwide. Our society is individually oriented. The rights of the individual often take priority over the rights of society as a whole. This view is certainly not a world wide view. Asia, especially Singapore, is fond of pointing out that the Asian view puts the group first and the individual is viewed in the light of what is good for the group. What "side" wills Cyberspace citizens take in this very profound cultural argument? Can both views live compatibly in a closely coupled cyber-world?

In Cyberspace individual national laws and customs, which are often different and contradictory, may conspire to limit the ability of individuals and corporations to freely interchange information, ideas, images and spoken works even when those items are legal and appropriate in the nation of one of the participants. Many societies currently, for example, limit the availability of satellite dishes. Several governments have equated internet access, along with the fax machine, as the prime vehicles for external disturbances to their control of their society and have stated that in the event of any future internal disturbances they will severe the internet connections rapidly. What will be the impact of such attitudes on international commerce and learning?

The privacy laws, that many governments have reasonably instituted to protect their citizens from having their personal information flow outside the control of the laws of their nation, raises many difficulties when one is engaged in an Internet environment. The establishment of directory structures which involve some nation's citizens may be in violation of the laws of that nation. Libel laws are traditionally national yet in Cyberspace, libel is instantaneous and globally damaging. Is there a notion of global liability? How do I sue a person in another nation? If I can,

---

The First Amendment in Cyberspace, *Marjorie W. Hodges and Steven L. Worona CAUSE/EFFECT* Volume 20, Number 3, Fall 1997, pp. 4-8 do we achieve the lowest common denominator? Is there a global “right to privacy”? How is it enforced? What happens to global commerce if there is not a common understanding?

Many nations and cultures have dramatically different perceptions of what is “proper” and not proper for its citizens to possess or to view. Consider an extreme case -- child pornography. We in the United States have strong laws which forbid the distribution and possession etc. of such material -- other cultures may not agree with us or have different notions of the control of such material. Suppose citizens of two such countries send each other such material and the material transits the United States; is storage on a US computer (without the knowledge of the owner of the computer) against the law? Can or should the US intercept such material and “delete” it, should they arrest the people when they next enter the US, should they close down the computer used to store the material. Is there an international agreement on the transport of cryptographic material across national boundaries? Is there a right of “innocent” passage -- that is, it is bound for another nation and just stops for a short stay -- mail relays for example? What is the right of a nation to monitor the contents or addresses of electronic communications that is accidentally transiting their nation?

All these and other questions are currently being hotly contested in the courts, by the legislatures and by advocacy groups.

The Cyber-economy

As the Internet becomes more a part of the everyday business of the nations, it will become more and more necessary for commerce to take place among the users of the infrastructure. We can expect the continued growth of an international electronic marketplace where goods of all types -- merchandise, information, software etc. are being bought and sold.

Historically there has always been a need to create a way of paying for such goods in order to motivate the supply side of the marketplace. Currently our primitive electronic marketplaces have no very effective mechanisms for paying for goods.

This creates an interesting and exciting opportunity to examine just what is needed to supply a mechanism for the exchange of electronic currency and how such a mechanism can exist in a national and international arena.

The issues raised by the potential existence of an international electronic marketplace are not limited to just how to pay for things. There is the need to have the equivalent of credit cards, checks and paper money with it various shades of traceability and privacy. There is the need for escrow mechanisms and international exchange etc.

Conclusion

I have tried to present a panorama of the major technical imperatives that will for the basis for the evolution of the networking and cyber-space in the first decade of the 21st century. I have suggested:

1. The advent of end-to-end photonic networks will cause major changes in the underlying structure of network protocols, methodology, computers and software systems as well as the industries that produce them. Such changes will be jarring to these industries and but will be necessary in order to profit fully from the new technology

2. The last mile problem will be a limiting facility especially for the home and work at home. Some municipalities will attempt to create cures to this in the same manner that they provide other infrastructures such as sewers and road

3. The advent of very high speed connectivity will have a major impact on legacy industry such as the telephone industry, the cable industry and the media companies

4. license free wireless will expand to form a very widespread facility challenging the new high speed data cellular systems

5. We need to pay much more attention to the creation of robust, trusted and secure networks and computer systems. Adding patches to legacy systems will not provide the capabilities we need.

6. And finally the impact of the global data networks on the life of the world citizens will be ever more disturbing. Protection against inappropriate intrusion on the privacy of the public will become an ever more contentious issue.

We should remember Ben Franklin -- "They that can give up essential liberty to obtain a little temporary safety deserve neither liberty nor safety.", ~1784