

“Anything you can do, I can do better; I can do anything better than you...”*

by Richard Jay Solomon,
Associate Director,
MIT Research Program on Communications Policy

We are about to witness one of those revolutionary shifts in technology that pass through similar stages of pathologic stress often chronicled for human behavior: first an imperceptible perturbation deemed a singularity or interesting curiosity; then the change grows into a faddish amusement or distraction; soon the amusement starts generating enough money or activity to become either, or both, a real business or a threat to other established enterprises or even to the social order; and finally the technological change takes hold, completely revamps its corner of the universe, economy, or society, and now enters the realm of permanent “problem” or “opportunity,” depending on your own particular history and demeanor.

The much abused shift of “paradigms” hardly suffices to describe such changes, for paradigms deal with models of behavior, business, and sociopolitical *g*estalt. The Internet is more than a new and different model for communicating electrons. The Internet doesn’t even exist, yet it has passed through the several stages of stress, and is midway between becoming a real business *and* a threat to the established order of things, and likely to completely revamp telecom, if not some segment of the universe.

These are big claims for an amorphous, somewhat virtual non-entity, so let us review briefly how we got this far and some possibilities for the near future:

The Net’s imperceptible stage lasted some 20 years, from the origins of the Arpanet and packet switching in 1969 to the building of the Internet backbone by the National Science Foundation in 1988-9, primarily to handle an anticipated moderate growth of traffic among the computer science and engineering research community. That worldwide message traffic on the Arpanet and associated, interconnected invisible networks surpassed the world total of Telex and telegraph traffic about a decade earlier seems to have escaped notice as a significant event by the conventional telecom industry — a wonderful example of denial, quite normal in stage one. We find similar historical denial parallels in other infrastructure shifts, notably in transport — trolleys and automobiles, railroads and trucks, etc.

The popular, somewhat faddish fascination with the Net has taken a mere two or three years, ranging from the amusement of downloading images of the Dead Sea Scrolls from some far distant server (indecipherable by more than a handful of Aramaic scholars), to the thrill of accessing perhaps 10 million machines chock full of pretty useless stuff. Nevertheless, the Net’s utility is not to be denied; that early base of researchers has now multiplied by millions who run their daily

operations almost entirely on electrons — email and web pointers — and from every corner of the globe. Commerce is now transacted on the Net. The ordinary consumer has begun to use it for purchasing, and for communicating with relatives, friends, and making new friends. How many?, how much?, to whom?, and where? are all poorly documented, but enough is known so that Stage Three has become interesting to the economy.

We are moving rapidly into that dangerous stage. The Internet has passed from fad, amusement, and utility, to a real business (overnight, billionaires have been created with the right-sounding Internet product at the right time). To some, the Net is also a real threat to the established order, most notably the social order as the hazard of content has reared its ugly head (not necessarily a mixed metaphor).

The last stage — some measure of permanence — may evolve via a number of paths. We will make some guesses later in the paper.

What is the Internet?

The Internet, *per se*, does not exist as an entity. It is no more than its name implies: a set of protocols or rules connecting computer, or more precisely, data communications networks to one another. Indeed, it is sometimes described as “a network of networks.” The Internet, being a virtual, some say “hostile” overlay on other physical networks, basically consists of two sets of protocols optimized to handle almost any type of digital transmission and interconnection over disparate and noisy channels. Together, these protocols are known as TCP/IP (for “Transmission Control Protocol/Internet Protocol”) IP, the Internet Protocol sets up a *virtual* connection between machines, determines the path for the connection, and handles the address details, including address translations—the so-called IP address. TCP, the Transmission Control Protocol regulates the actual traffic flow, providing both the human user and client machines an apparently robust and error-free path for streaming bits, even over paths that contain inherently poor circuits.

Up to now, the maintenance of the protocols, addressing tables, etc., have been handled by volunteer organizations consisting primarily of the user community; this is changing with the commercialization of interconnection, backbones, Internet Service Providers, and other operational features. The next stage in management evolution is an open question; it will not necessarily be an extension of conventional standards entities, nor does it appear that open, non-contentious, cooperative behavior will rule either as interfaces get more complex and diffuse. We offer as an example Figure 1, which lists and groups the bodies and firms contending for the next iteration of multimedia standards.

Figure 1

The Multimedia Standards Universe

[source: S. Neil, R. Solomon, L. McKnight, ©1995, 1996 MIT Research Program on Communications Policy]

Packet vs. Circuit Switching

The Internet follows a store-and-forward, packet data communications architecture, quite different than that of the public switched telephone network (PSTN). High-speed packet switches can minimize packet delays so that they are imperceptible for most applications. Alternatively, this delay can be exploited to guarantee delivery whether or not the recipient device is connected to the Internet at the time the data is generated. While this system has less first-order efficiency compared to networks optimized for one application (e.g., duplex, real-time voice telephony, or broadcast media), this relative lack of efficiency must be balanced against the high level of robustness built into the two basic protocols and an incredible gain in trunk-route (or backbone) efficiency because of maximized circuit sharing. Hence, it has been noted that the real incremental cost of an additional bit on the net is zero (and may be even be negative if excess traffic flows from one backbone to another). Costs rise in step functions, and the steps can be pretty far apart.

The Internet protocols accomplish this magic economic trick via sufficient intelligence in the control process, making it flexible enough to handle a range of uses from electronic mail to telephony and different types of video. Much of the control is peripheral to the Internet backbones and access links, residing principally in client devices or user's software. This use of distributed processing for network control is key to the Net's efficiency, extensibility, and ready scalability. The economics of distributed processing goes beyond volunteer standards-making to a shift in cost structure for network maintenance and upgrading — something quite alien to conventional telecommunications network Operations, Administration and Maintenance (OA&M). We will return to this point when we discuss future Internet evolutions.

Net protocols are constantly being modified; the next generation is expected to have stronger priority controls, making interactive services feasible such as full-scale, full duplex voice telephony. Furthermore, efforts are underway to create direct interfaces between the Internet's control system and protocols and those of the public switched telephone network, eventually migrating — not necessarily smoothly — to a transparent or seamless set of data and telephone systems.

Internet mini-FAQ

To understand possible future network scenarios, the important points to remember about the Internet as a concept are:

- The Internet is not a physical or unitary *thing* (not yet, at least). It is what its name implies: inter networked computer appliances. There are many “internets”: many interconnected networks use the underlying internet protocols (TCP/IP) for connectivity, but may not be directly connected with the Internet.
- The Internet is delineated by its interface protocols. It has not been delineated by its corporate structure (it has none, so far), its tariffs (currently determined by cost of access to network access points), physical routes (it uses whatever is available, public or private, selected by a mix of customers, carriers, and other vendors), or boundaries (whatever device or application connects, using the Net’s accepted protocols, becomes immediately part of the Internet—for that slice of connect time).
- Internetting is primarily software-driven. The software for connecting, routing, and controlling data flow is embedded in the operating systems of the connected computers.

Why Does the Internet Work?

The Internet approach has been successful both because of the technical strength of its protocols, which are flexible enough to accommodate both existing and emerging systems, and because of key institutional strategies:

- *Media independence.* The TCP/IP protocols have been designed to facilitate error-free data connections over virtually any kind of physical communications network—dial-up telephone, private digital and analog broadband circuits, satellites, radio, etc.
- *Interoperability with other protocols.* TCP functions within other protocols, such as broadband ATM (Asynchronous Transfer Mode) fibers and wire pairs, Ethernet and LocalTalk/AppleTalk LANs, and also over the public switched telephone network with the use of optimized dial-up synchronous line protocols (SLIP) including PPP, AppleTalk Remote Access, and on basic and primary rate ISDN.

Operating entities are now offering IP interfaces which permit the Internet to use internal telephone network signaling protocols (e.g., SS7) to control access to the PSTN and vice versa, so in principle an IP address can translate directly into a telephone number (or vice versa) opening up all sorts of possible new services making the telephone network transparent to the Internet user. This potential high level of interoperability portends dire consequences for telco traffic diversion and bypass, and is virtually unstoppable because of the inherent robust nature of IP.

Proposals to ban the sale of “Internet telephony” products would be as effective as previous attempts to forbid direct connection of modems by telephone administrations around the world. Nevertheless it is a conundrum: software to interconnect computers directly with telephone switching systems is being promoted by the telcos for direct marketing and other business purposes. They cannot have it both ways — permitting interconnection to computers for one type of call and simultaneously denying interconnection if the bits transmitted produce negative revenue.

- *Technical robustness.* IP by itself is a very unstable, unreliable connection architecture, but with TCP riding above it, it can interconnect almost anything. With TCP, the user appliance perceives IP as super-robust, and therefore one can easily build in ultra-reliability with a suitable congestion mechanism (pricing, for example). TCP provides this robustness via its built-in redundant features: automatic alternate routing, multiple simultaneous connections, scalable bitstreaming, content independence, and flexible use of IP addresses. No test data is available (i.e., no one has yet compared TCP/IP to PSTN and published verifiable results), but in principle, TCP/IP should be more robust than the PSTN, though this is a complex question with complex answers.

One slogan the Internet goes by is that IP sees all blocks as errors, and will make extreme attempts to route around blockages. There is simply no easy way to block access on the Net because IP was deliberately designed as a Cold War mechanism for communications survival during a nuclear attack. With a little help from its friends, IP tends to overcome any censorship, for example, and once connected to the Net, a computer operating system has to take extra precautions if firewalls are not to be breached by error or by deliberate attack.

- *No cost to the user.* One primary reason TCP/IP became so popular, as opposed to ITU recommendations X.400/X.500 which deemed to perform similar packet switching feats, is that the Arpanet-cum-Internet interface protocols have been bundled free of charge into Unix since the mid 1980s, the operating system used on most servers. DOS, Windows and the MacOS also have TCP/IP built-in, via free or public domain products, and now bundled into almost every PC sold.
- *User-driven.* The Internet protocols were developed in response to user needs and are flexible enough to develop in concert with emerging user requirements.

IP *domains* are increasing faster than IP *name* addresses. Since domains are completely portable from one ISP to another (like toll-free 800 and 888 numbers are now in the U.S.), it is worthwhile (and even for a very small business, not terribly expensive) to have your own domain, for then your address may never have to change wherever you set up business. Whenever you sent mail to xyz@domain.abc, the local server looks up the correct IP *number*, which tells it how to route the packets. All servers worldwide are updated frequently with the correct IP routing information. So if you change your

Internet Service Provider, you merely inform the Internic registration center of this fact, and your mail and hypertext pointers (and in the future, possibly your telephone calls) follow you to the new location.

Interfacing address schemes for interworking with global PSTN numbering plans are a potentially great opportunity to stop constant expansion of telephone numbers at great confusion and inconvenience to the public. IPng addresses will be sufficient to cover the planet with some 64,000 addresses per square inch; enough for any scheme that includes subaddresses for each and every RAM that may ever exist. Using IP techniques for lookup tables, we can all carry permanent addresses which translate into names and domains, and which will work on any network anywhere,. But such transparent flexibility also carries with it the ease of bypassing telephone switching for faxes, data transmissions, full duplex, stereo audio, and future videotelephony products.

All this technology, supported by millions of customer terminals and 100s of thousands of technical wizards, has created a totally different kind of telecom system. It is one where a large segment of the user community has more knowledge and insight of how the network works than the network provider. This is not the stuff that monopoly is made of.

The reason as to why the Internet has grown so rapidly, and why it costs so little, is that more than 90% of its costs are borne by the user upfront, and not accounted for as such (very much like the transformation of the transport economy from public transit to the automobile, whereby most auto owners discount the real costs of operating a vehicle and only look at out-of-pocket costs — tolls and parking). Users buy PCs, and then LANs to connect them in their offices; most likely this part of the investment would have been made without the Internet. Even modems, which are now a small part of the total cost of user equipment, are useful to connect with other machines (or to send faxes) and the Net was not necessary to justify the investment in connectivity. Internet costs are only marginal for most commercial, government and academic users. Marginal enough to postulate that the Net was inevitable, since the R&D community operated much more efficiently with it than without it. What is a surprise is that the general public is gradually taking to networking.

The growth of the net parallels some similar shifts in transport. For a few decades while the motorcar and then aviation drew traffic from railroads and transit, the ground public transport system remained viable enough to handle excess traffic, especially during World War II when railroads doubled their load, or when airlines broke down during snowstorms and other catastrophes. Eventually that all ended, and there is no transport redundancy left for most everyone: today you drive, fly, or don't go. With rare exceptions, there is no workable ground public transportation alternative left in the U.S. Similarly, at some point in the future, Net traffic will surpass the conventional PSTN; it will be cheaper for some segment that has direct connections; the Net will have all the features, and more, that the PSTNs dream of offering with Intelligent Networks (trivial things like call forwarding, answering, and conferencing are available now on cheap customer premise equipment; on PCs, it will be just some software upgrade).

The telcos, currently in denial mode for the catastrophe that beckons, say you still need their wires to connect to the Internet. Perhaps, though other wires, cables, radio and satellite channels will be there, too. However, the Internet does not need the expensive telco switches. And once on the Internet Service Provider (some small computer on a pole will suffice), there are many choices for Net backbones — remember, the software defines the Net, not the wires, cables, fibers, etc.

Before this sounds like the millennia has arrived, there are a few problems to resolve. While Net software can do anything the PSTN does, and probably better, there is the matter of who is going to pay for infrastructure in the first place, and who will pay for the investment mistakes, if all the Net does is ride on existing, embedded plant? How will the Internet access points and backbones be financed if the rest of the system goes slowly bankrupt?

Several basic questions must be faced in the transition:

- Charging for use based on transmission capacity will not work where complex services such as video, voice and data are mixed. The ratio of bits needed for video, for example, is several orders of magnitude higher than that needed for even stereo, full-frequency voice channels. But it is extremely difficult, if not impossible, to discriminate among services — the network only knows about bits, about packet delays, and with IPng, priorities (red and green bits). Otherwise, bits are bits.

The details of why service segmentation will not work are too complex to cover in this paper, but, in general the only costs that can be reflected in the price are the access costs, not the transmission (assuming price reflects costs at all). Pricing by size of access “pipe” is the closest an operator can come to recovering costs. Telco dreams of increasing revenues by bits transported or value-added will come to naught when IP bypass protocols are put into effect and local access competition pushes margins razor thin, as the superstores have done to retail merchandise. The new entries do not have to amortize \$30 million central office circuit switches.

Are we to provide a Universal Service Fund to finance obsolescent embedded plant? When does this stop? Who pays for it? (Where PTTs are still government-owned, the answer is clear. But what excuse will be offered for the taxpayers to pick up privatized mistakes?) Do the new Internet police expand their domain from porn to bypass by tapping into all packets? What if it is encrypted? Does the telephone company get a court order if it suspects someone is “talking” instead of typing? Do we mess up the network because we invented a technology that is too efficient.

- Today, most Internet “maintenance” is borne by the customer. This consists of tweaking the PCs and LANs,

purchasing and maintaining software, and connecting to the ISP. Proposals to extend “Universal Internet Service” seem to assume that the less “computer literate” customer will be willing or able to out-source (pay the ISP or telecom operator) the true costs of high-quality system management today found with corporate or academic system gurus, hobbyists, and friends you can call at 2 am when your disk crashes.

Universal service designs are reflected in the either naive or cynical offerings of \$500 “hollow” PCs for connection to TV sets so the masses can surf the net. Never mind that research has shown that humans cannot tolerate fine text on conventional TV screens (which use an inferior display standard as compared to PC monitor), it is not clear that consumers are willing to pay for cheap boxes with expensive community server/timesharing machines behind them.

- As the Internet enters its “mature” phase of permanent problems (or opportunities), it appears that several evolutionary scenarios may be followed, none of which were anticipated by its techie designers. The Internet is now going to be regulated in the U.S., the country that pioneered telecom deregulation, because it scared people afraid of speech, it scares firms afraid of competition, and it scares politicians just plain afraid. Other countries may do even worse; China wants to register, and regulate *users*, not just the providers. Will you need a driver’s license to drive the information highway?

Regulation from fear, rather than regulation towards a goal (such as fair rates or non-discriminatory access), often leads to chaotic results — the Law of Unexpected Consequences. The Net will not go away; if it did not exist, it would be invented today. All the Net is is communicating computers. As long as you have any kind of a network, and any kind of computer, there will be computer communications. If the Internet were to be dismantled, researchers would re-connect via private lines, corporate circuits, the PSTN (Fidonet was, and in some countries still is, an Internet wannabe using PSTN and thousands of PCs dialing each other). We would then have many new Internets with some odd configurations, but this new set would eventually coalesce into the Net once again. The Law of Unexpected Consequences meets the Sorcerer’s Apprentice. And the compression technologies, the robust routing algorithms, the powerful encryption techniques, and the latent demand for computer communications would not disappear in a dismantling.

So, what is to be done? The network that emulates all other systems cannot be simply controlled just because it violates some sense of neatness; it cannot be bought, because there is almost nothing to buy unless some newly empowered monopoly were given the resources to buy everyone’s equipment, connections, software and brainpower and then find more resources to manage it all; it cannot be harnessed to carry only proper bits and to filter or lose improper bits; and it cannot be turned off because it is being used for productive work by millions of people who would just find another mechanism to rebuild the Internet under a new name if they have to.

Perhaps instead of worrying about Internetting computers causing harm, we should be re-evaluating how we provision infrastructure from first principles. And, unlike the concrete and asphalt highways, information highways are not always visible, so simplistic solutions are not going to be easy to come by.