Metcalf's Law and Legacy

George Gilder

Forbes ASAP
September 1, 1993

The world of networks breaks into two polar paradigms. Most familiar is the Public Switched Telephone Network. From the tiniest transistor flip-flop on a modem chip through labyrinthine layers of rising complexity on up to a 4ESS supercomputer switch linking 107,520 telephone trunk lines (itself consisting of millions of interconnected transistors), the public network is a vast, deterministic web of wires and switches. Once you are connected in the public network, your message is guaranteed to get through.

In the public network, bandwidth constantly expands as you rise in the hierarchy. At the bottom are the twisted-pair copper wires of your telephone that function at four kilohertz (thousands of cycles per second). At the top are fiber-optic trunk lines that function at rates close to the 2.9-gigahertz speeds of the electronic transistors that feed the glass wires. In The Geodesic Network, writer Peter Huber has described the five tiers of the telephone switching system as a structure with "the solidity, permanence and inflexibility of the Great Pyramid of Cheops, which on paper it resembled." Although the pyramid has suffered erosion and change in recent years, it remains mostly in place today: the public network pyramid.

That is one network paradigm. The other paradigm is Robert Metcalfe's. It germinated in his mind in 1970 as he read a paper by Norman Abramson of the University of Hawaii given at a computer conference that year. Abramson told of another paradigm. He called it Aloha. With Aloha, there were no guarantees.

AlohaNet was a packet radio system used for data communications among the Hawaiian Islands. Packets are collections of bits led by a header, which is a smaller collection of bits, bearing an address; they proceed through a communications system rather like envelopes through a postal system. The key feature of AlohaNet was that anyone could send packets to anyone else at any time. You just began transmitting. If you didn't get an acknowledgment back, you knew the message had failed to get through. Presumably your packets had collided with others. In Metcalfe's words, "They were lost in the ether." At that point, you would simply wait a random period (to avoid a repeat collision as both parties returned to the channel at once). Then you would retransmit your message.

To Metcalfe, AlohaNet seemed a beautifully simple network. But Abramson showed that, because of collisions and other problems, it could exploit only 17 percent of its potential capacity. A student of computer science searching for thesis ideas, Metcalfe believed that by using a form of advanced mathematics called queuing theory he could drastically improve the performance of AlohaNet without damaging its essential elegance and simplicity. What Metcalfe, then a graduate student at Harvard, eventually discovered would bring such networks up toward 90 percent of capacity and make the Aloha concept a serious threat to the entire structure of the public network pyramid.

Metcalfe's discovery is known as Ethernet. Twenty years later, Ethernet is the world's dominant local area network and, at 47, Metcalfe is known and celebrated as its inventor. He was also founder in 1981 of 3Com Corp. of Santa Clara, Calif., the leading producer of Ethernet adapter cards and a major communications products company. In this era of networking, he is the author of what I will call Metcalfe's law of the telecosm, showing the magic of interconnections: connect any number, "n," of machines - whether computers, phones or even cars - and you get "n" squared potential value. Think of phones without networks or cars without roads. Conversely, imagine the benefits of linking up tens of millions of computers and sense the exponential power of the telecosm.

Indeed, the power of the telecosm reproduces on a larger scale - by interconnecting computers - the exponential yield of the microcosm, a law describing the near magical effect of interconnecting transistors on chips of silicon: As increasing numbers of transistors are packed ever closer together, the transistors run faster, cooler, cheaper and better. Metcalfe's law suggests that a similar spiral of gains is available in the telecosm of computer communications.

Already the world economy is beginning to reap these gains. Ethernet now links more than half of the world's 40 million networked computers, extending Metcalfe's paradigm and his law. Indeed, the law would suggest that in addition to his some $20 million of personal net worth from 3Com, Metcalfe's concept has fostered scores of billions of dollars in global wealth. Led by Novell Inc., with an equity capitalization of more than $8 billion, the top 15 publicly traded computer networking companies have a total market value of some $22 billion. Add to that sum the productivity value derived from the world's 100 million computers as they are increasingly linked in networks, and you may sense the power of the Metcalfe paradigm.

Today, 20 years after Metcalfe conceived it at Xerox's Palo Alto Research Center, Ethernet is still gathering momentum, gaining market share and generating innovations. Between 1989 and 1993, the percentage of America's computers on LANs rose from less than 10 to more than 60, and most of these gains were in Ethernets.

Ether Moves to Cable

The telecosm's powers could end up saving the American economy from itself. In an era when the new payroll taxes and regulations of Clintonomics could end up driving millions of mind workers back into their homes, Digital Equipment Corp. is now extending Ethernet's range from its current two-mile limit to some 70 miles. Called Channelworks, the DEC system can run Ethernet on the some 50 million miles of cable television coax. This will enable potential scores of millions of telecommuters to access their familiar office LAN, tap their company E-mail and their corporate databases, and generally make themselves feel at work while at home. Deployed at a profit and extended to customers at a flat monthly rate, Ethernet in the neighborhood could become a massive growth business for the cable industry over the next decade.

As Ethernet spreads and faces the challenge of remote work teams using digital images, simulations, maps, computer-assisted design schematics, visualizations, high-fidelity sounds and other exotic forms of data, the system is constantly adapting. From 3Com spin-offs Grand Junction Networks and LAN Media Corp. to smart hubmaker David Systems, from Kalpana to Synernetics, from National Semiconductor to Hewlett Packard, from Cabletron to SynOptics, from AT&T even to Token Ring leader IBM, scores of companies are pushing Ethernet into new functions and performance levels. It is emerging in full-duplex, multimedia, fast, fiber- optic, shielded, unshielded, twisted, thin, thick, hubbed, collapsed, vertebrate, invertebrate, baseband, broadband, pair, quartet, coaxial and wireless versions. It now can run at 2.9, 10, 20 and 100 megabits per second. It has moved from 2.9 megabits per second to 100 megabits per second and from a few hundred to several million users in some 10 years. At its present pace of progress, Ethernet will someday run isochronous (real-time) gigabits per second on linguine.

Aloha ATM, Gushing Cash

So why is its boyish-looking inventor - over Metcalfe's anguished protests, think of Ted Kennedy some 10 years ago - giving up on his baby just as it enters its roaring 20s? Why is he ready to abandon his basic paradigm in favor of a return to the public network vision of massive, intelligent switching systems? Why is he now talking of Ethernet as a "legacy LAN"?

Discoursing this summer from a deck chair on his yacht (a converted lobster boat) as he breezed down from his Maine retreat to a dock on the Charles River for his 25th MIT reunion, Metcalfe has the air of an elder statesman. Though humbly grateful for the benisons of Ethernet, he has seen the future in a poll of experts prophesying the universal triumph of a powerful new switching system called asynchronous transfer mode (ATM). "I have found," Metcalfe solemnly intones, "an amazing consensus among both telephone industry and computer networking experts that ATM is the future of LANs." Aloha, ATM.

Metcalfe is not alone among Ethernet pioneers flocking back to Ma Bell's pyramid of switches. Also leaving Ethernet behind is his onetime nemesis, Leonard Kleinrock of UCLA, a leading guru of gigabit networks who helped define the mathematical limits of Ethernet, and is given credit (or is it blame?) for naming its Carrier Sense Multiple Access/Collision Detection protocol (CSMA/CD). Preparing to defect to ATM is Ronald Schmidt, the brilliantly ebullient technical director of SynOptics, who created the latest Ethernet rage - sending the signals over telephone wire under the 10baseT standard (10 megabits of baseband data over twisted pair).

There has not been such a stampede to a new standard since the global rush to ISDN (Integrated Services Digital Network) in the early 1980s. Offering digital phone lines at 144 kilobits per second, ISDN is just now coming on-line in time to be aced by the megabits per second of Ethernet over cable.

In a prophetic memo launching the concept in 1973, Metcalfe foreshadowed the secret of Ethernet's success. He wrote: "While we may end up using coaxial cable trees to carry our broadcast transmissions, it seems wise to talk in terms of an ether, rather than `the cable'.... Who knows what other media will prove better than cable for a broadcast network: maybe radio or telephone circuits, or power wiring, or frequency-multiplexed cable TV or microwave environments, or even combinations thereof. The essential feature of our medium - the ether - is that it carries transmissions, propagates bits to all stations." In other words, it is the stations, rather than the network, that have to sort out and "switch" the messages.

The word Ethernet may be capitalized to signify the official standard of CSMA/CD. Or it may be lowercased to suggest a medium without switches, routers and other intelligence. In either case, the word "ether" conveys the essence of the ethernet. An ether is a passive, omnipresent, homogeneous medium. Long believed essential for the propagation of electromagnetic waves, the literal existence of an ether was disproven in the late 19th century by the famous experiments of Albert Michelson and Edward Morley. But the concept of a figurative ether - a dumb medium of propagation - survives in modern communications.

The enduring magic of ethernets stems from the law of the microcosm, favoring distributed terminals over centralized hierarchies, peer networks of PCs over mainframe pyramids. The microcosm's relentless price/performance gains on chips have endowed Metcalfe's peer-to-peer scheme with ever more powerful peers, at ever lower prices. Medium-independent from the outset, the Metcalfe systems do not require central switching. In an ethernet system the intelligence is entirely in the terminals, not in the network itself, and most of the bandwidth is local (where some 80 percent of traffic resides).

Although this ATM is expected to gush jackpots of cash for gaggles of network companies and investors, it is unrelated to its acronymic twin, automatic teller machines. Think of ATM rather as an automated postal center that takes messages (of any size or addressing scheme), chops them up, puts them into standardized little envelopes and figures the best routes to their destinations in billionths of a second. The magic of ATM comes from restricting its services to those uniform envelopes (called cells) of 53 bytes apiece (including a five-byte address) and creating for each envelope what is called a virtual circuit through the network. These features make it unnecessary for intermediate switches in the network to check the address; the cell flashes through the system on a precomputed course.

A compromise defined by phone companies as the longest packet size that can handle voice in real time, 53-byte cells are also short enough to be entirely routed and switched in cheap hardware; i.e., microchips. This means that the ATM postal center can function at speeds of up to 155 megabits per second or even higher. Perhaps most attractive of all, ATM can handle multimedia data, such as digital movies or teleconferences, with voice, text and video that must arrive together at the same time in perfect sync. As the world moves toward multimedia, the industry is flocking toward ATM, the innovation that can make it possible.

Ethernet: A Legacy LAN?

By contrast, Ethernet seems old and slow: the vacuum tube of computer communications. Think of it, crudely, as a system where all the messages are cast into the ocean and picked up by terminals on the beach which scan the tides for letters addressed to them. Obviously, this system would work only if the beach terminals could suck up and filter tremendous quantities of sea water. The magic of ethernet comes from the ever growing power of computer terminals. The microcosm supplies sufficiently powerful filtering chips - chiefly digital signal processors improving their powers some tenfold every two years - to sort mail and messages in the vasty deep. This is quite a trick. To the experts, it seems unlikely to prevail for long against the fabulously swift switching of ATM.

True, there is some confusion about just how, where and when this miracle cure will arrive. The industry's leading intellectual, Robert Lucky of Bellcore - a paragon of long-distance networks - predicts that ATM will come first in local area networks, while Metcalfe, of local area network fame, thinks it will come first in wide area networks. James Chiddix of Time- Warner Cable is probably right in predicting digital cable pay- per-view as the first big ATM customer, using it for broadcasting films in his 500-channel digital cable TV project in Orlando. But most experts agree that one way or another ATM will blow away Ethernet during the next decade or so.

Nonetheless, as usual, conventional wisdom is wrong. Ethernet is quietly preparing for a new era of hegemony in the marketplace for computer connections.

The reason Ethernet prevailed in the first place is that, in the words of Ronald Schmidt, "it was incredibly simple and elegant and robust." In other words, it is cheap and simple for the user. Customers can preserve their installed base of equipment while the network companies innovate with new transmission media. When the network moves to new kinds of copper wires or from one mode of fiber optics to another, Ethernet still looks essentially the same to the computers attached to it. Most of the processing - connecting the user to the network, sensing a carrier frequency on the wire and detecting collisions - can be done on one Ethernet controller chip that costs a few dollars.

As Metcalfe described the conception of this technology in 1981, "I explored the advantages of moving the transceiver down out of the ceiling onto the adapter board in the host computer. I had seen many actual Ethernet installations in which our brick transceivers were not up in the ceiling tapping into the ether cable, as they were supposed to be...but instead were on floors behind computers, dropped in the centers of neatly coiled transceiver cables.... We were discovering that the people buying personal computers and workstations in those days were not generally the same kind of people who were allowed to remove ceiling tiles and string cables through conduits.... The personal computer revolution was taking place in organizations from the bottom up.... It was time for Ethernet to be re- invented for bottom-up proliferation among the personal computer work group revolutionaries."

Using "silicon compiler" design tools to radically reduce the time to market, Seeq Technology created an Ethernet chip for PCs in time for a single-board version of the interface unit. Putting the transceiver on the adapter board eliminated a special transceiver cable and drastically simplified the system. There is no bulky connection between the coding device preparing information for the network and the transceiver sending or receiving the signals on the net. All this processing is done in the computer, on one printed circuit board, now reduced to the size of a credit card. While its rival from IBM - Token Ring - requires a mostly proprietary array of token-passing managers, clocking assignments and other complexities, Ethernet is an open system. Relative to the alternatives, it offers the possibility of something near plug-and-play. So advantaged, Ethernet has overcome IBM's Token Ring, 20 million nodes to 8 million in installed base.

But this does not persuade Ethernet pioneers Bob Metcalfe, Leonard Kleinrock and Ronald Schmidt. Because ATM can handle all kinds of data fast, Metcalfe sees it as the "grand unifier" bringing together WANs and LANs and effecting a convergence of television, telephony and computing in turbulent multimedia bit streams bursting into our lives early next century. "And of all the variations of multimedia," he writes in Infoworld - Metcalfe is now its publisher - "the one that will drive ATM is personal computer video conferencing - interactive, two-way, real-time, integrated digital voice, video and data." Although Ethernet will persist as a "legacy LAN," he says, it cannot compete with ATM in these crucial new roles. Schmidt makes the same essential case, stressing the need for switch-based architectures in a world of exotic new media.

Kleinrock's Formula

Why the pessimism on Ethernet? Bringing mathematics to bear on the argument, Kleinrock declares that the collision-detecting functions of Ethernet bog down with large bandwidths, short packets and long distances. Thus, the system must fail with the onset of fiber highways across the land. The oceans of Ethernet will simply grow too large to allow efficient detection of collisions in its depths. With large bandwidths, more packets can be pumped into the wire or glass before a collision is detected; by that time, most of the transmission is finished. When the distances get too long, collisions can occur far from the transmitting computer and take longer to be detected. The shorter the packets, the worse these problems become.

As Kleinrock computes these factors, the efficiency of Ethernet is roughly a function (a), computed as five times the length of the line in kilometers times the capacity of the system in megabits per second, divided by the packet size in bits. When a exceeds a certain level (Kleinrock sets it at 0.05), Ethernet's efficiency plummets.

With ATM packet sizes needed for voice traffic - or even at the minimum Ethernet packet size of 72 bytes - any Ethernet with a capacity much higher than 10 megabits per second exceeds this tipping point. Therefore, high-speed Ethernets must either use packets too long for voice or shrink in extent to far less than three kilometers. This is what Howard Charney's Grand Junction and its rival LAN Media propose with Fast Ethernet. Noticing that 10baseT hubs have reduced the length of Ethernet connections by a factor of 10, Ron Crane, founder of LAN Media, suggests that this change allows acceleration of the system by an equal amount: to 100 megabits per second.

But this seems a one-time fix that fails to address the multigigabit world of fiber optics. At some point, Kleinrock, Schmidt and Metcalfe agree, ad hoc fixes will begin to fail and ATM (or possibly some other system) will begin to prevail. Using Kleinrock's formula, that point is here today, with 100-megabit- per-second Ethernet lines.

As an increasing share of network traffic takes the form of pictures, sounds, simulations, three-dimensional visualizations, collaborative work sessions, video teleconferences and high- resolution medical images, the Ethernet model already seems to be foundering, according to many expert projections. The triumph of ATM, so it would seem, is just a matter of time.

Time, however, is precisely what is absent from all these projections. Ethernet is a system based on the intelligence of terminals; ATM is a system based on the intelligence of switches and networks. All the arguments for ATM miss the law of the microcosm: the near annual doubling of chip densities, the spiraling increase of computer power surging on the fringes of all networks as transistor sizes plummet over the next decade.

The Power of Exponents

Amazingly, most technology prophets fail to come to terms with the power of exponents. You double anything annually for long - whether deforestation in ecological nightmares or transistors on silicon in the awesome routine of microchip progress - and you soon can ignite a sudden moment of metamorphosis: a denuded world or a silicon brain.

Shortly after the year 2000, semiconductor companies will begin manufacturing microchips with more than a billion transistors on them - first as memories, and soon after as processors. A billion transistors could accommodate the central processing units of 1,000 Sun workstations or 16 Cray supercomputers. This means roughly a millionfold rise in the cost-effectiveness of computing hardware over the next decade or so.

Intelligence in terminals is a substitute for intelligence in networks; switching and routing functions migrate from the center of the web to the increasingly powerful computers on its fringe. Looming intelligence on the edge of the network will relieve all the current problems attributed to ethernets and will render the neatly calculated optimizations of ATM irrelevant.

Meanwhile, the law of the telecosm is launching a similar spiral of performance in transmission media, ultimately increasing their bandwidth, also by a factor of millions. Bandwidth is a replacement for switches. If you can put enough detailed addressing, routing, prioritization and other information on the packets, you don't have to worry about channeling the data through ATM switches. The emergence of dumb, passive all-optical networks with bandwidths some ten- thousandfold larger than existing fiber optics will obviate much of the pressure on switches. Combining microcosm and telecosm in explosive convergence makes it nothing short of ridiculous to expect a system optimized for 1995 chip densities and fiber capacities to remain optimal in 2013, when Metcalfe foresees the final triumph of ATM, or even in 2001.

Of course, ATM will be useful in various applications before then. Sun and SynOptics envisage putting ATM ports in future workstations where ISDN ports mostly languish today. AT&T, MCI, Sprint and Wiltel will incorporate ATM switches in their long- distance networks. Time-Warner may indeed use them for distributing movies. In general, however, companies that rely on an apparent trend toward centralized switches will be disappointed.

Cable firms will do better by sticking to the ethernet paradigm of dumb bandwidth that has made them the envy of all in the emerging era of digital video. IBM and other computer firms with powerful ethernet and fiber technologies should not rush to adopt the public network paradigm. Telephone companies in particular should maintain an acute interest in their ongoing experiments with all-optical networks and other passive optical technologies. Any near-term successes of ATM, afflicted with the many glitches and growing pains of any new technology, are likely to come too slowly to deflect the continuing onrush of ethernets.

Ethernet prevails because it is dumb. In the old world of dumb terminals - whether phones, IBM displays or boob tubes - a network had to be smart. There was time even to put human operators into the loop, and a need to concentrate programming at one central location. But in the emerging world of supercomputers in your pocket or living room, networks will have to be dumb bandwidth pipes. What the coming array of desktop supercomputers and cheap massively parallel servers will need is passive dark fiber, mostly unlit by switching intelligence. Dark fiber can allow for the huge variety of data forms and functions, protocols and modulation schemes that is emerging in the new era of convergence between phones and computers.

Ethernet is the protocol for a dumb pipe, a passive ether. That is why it fits so well on a cable TV line and why it will fit even into the multigigabit world of a multimedia future.

The Return of Aloha

The dumb networks of the fibersphere will be ethernets. These all-optical links that have been made possible by the creation of erbium-doped amplifiers and other passive devices give access to the full 25,000-gigahertz bandwidth of fiber optics (see "Into the Fibersphere," December 7, 1992). In these networks, fiber changes from a substitute for copper to a substitute for air. Just as the microcosm put entire computer systems on single slivers of silicon, the telecosm will put entire communications systems on seamless webs of silica. Terminals will tune into the infrared colors of the fibersphere like radios tuning into the frequencies of AM or FM.

As chips and fiber are hugely expanding their performance and bandwidth, information traffic is rapidly migrating from the wires to the air. Although many experts contend that the radio frequencies in the air - the electromagnetic spectrum - are running out, communications systems now use only a tiny sliver of spectrum, well under one percent of the usable span. As shown by Cellular Vision's success in sending cable TV signals over the air at 28 gigahertz, it is now possible to move up the spectrum into the vast domains of microwaves; other experiments show that network traffic in these portions of the spectrum can be accommodated with error rates of less than one in a billion, enough to avoid extensive error correcting.

At the same time, the replacement of today's 30-mile cells with tomorrow's closely packed microcells means an exponential rise in available spectrum and an exponential reduction in power usage. The replacement of analog systems with digital systems using code division multiple access (CDMA) will allow the reuse of all frequencies in every cell, thus further expanding available spectrum (see "New Rules of Wireless," March 29, 1993). A company called ArrayCom in Santa Clara, Calif., is developing a new system, called spatial division multiple access (SDMA), based on smart antennas that can follow an individual communicator as it moves through a cell. This technology would allow the use of all the available spectrum by each "phone."

Back to the Real "Ether" Net

Inspired by a radio network, ethernet is well adapted for this new world of wireless. The increasing movement of data communications into the air - the real ether - will give new life to Metcalfe's media-independent system. Cellular systems already operate with protocols similar to CSMA/CD. As microcells fill up with digital wireless traffic, all networks will increasingly resemble the most popular computer networks. In the ether, links will resemble ethernets far more than ATMs.

The coming age of bandwidth abundance in glass and in air converges with an era of supercomputer powers in the sand of microchips. We should build our systems of the future - the cathedrals of the Information Age - on this foundation of sand. It will not disappoint us.

Whether in glass or in air, the basic protection of Ethernet is not smarts but statistics. Ethernet is a probabilistic system. This fact has caused endless confusion. Because a probabilistic system cannot guarantee delivery of data on a specific schedule, or at all, many experts have concluded that Ethernet is unsuited for critical functions, or for isochronous data inherent in multimedia - with voice and video that must arrive in real time. When and whether anything arrives is a stochastic matter.

Nonetheless, if there is enough bandwidth for the application, ethernets work just as reliably and well as their deterministic rivals, even for advanced video traffic. As Kleinrock observes, for many image applications, very long packets can be as effective as very short ones. The long packets become a virtual circuit connection, somewhat like a phone call. It is likely that perhaps 80 percent of all multimedia will be sent in burst mode, with a store-and-forward protocol, rather than isochronously in real time. Broadband ethernets will be better for burst mode than ATM's short packets.

In any case, the combination of intelligence at the terminals and statistics in the network is more robust than the mechanistic reliability of Token Rings or ATM switches. As Metcalfe points out in explaining the triumph of his vision over Token Ring, Ethernet is a simple system that is stabilized by its own failures. The CSMA/CD algorithm uses collision detection in a negative feedback loop that delays retransmission in exponential proportion to the number of collisions, which is a reliable index of the level of traffic. Thus thriving on a worst- case assumption of frequent failure, Ethernet has outpaced all rivals that guarantee perfect performance and depend on it.

Metcalfe's Law: Transcending His Own Doubts

Now, in ATM, Ethernet is faced with a new paragon of determinism offering high speeds and rigorous guarantees, a new version of the public network paradigm, a new pyramid of switching power. But Metcalfe's law and legacy may well win again, in spite of his own defection.

As Metcalfe explains, "Ethernet works in practice but not in theory." The same could be said of all the devices of the microcosm and telecosm. Both of the supreme sciences that sustain computer and communications technology - quantum theory and information theory - are based on probabilistic rather than deterministic models. They offer the underpinnings for an age of individual freedom and entrepreneurial creativity.

Humankind's constant search for deterministic assurance defies the ascendant science of the era, which finds nature itself as probabilistic. To Einstein's disappointment, God apparently does throw dice. But chance is the measure of human ignorance and the mark of divine knowledge. Chance thus is the paradoxical root of both fate and freedom.

Nations and networks can win by shunning determinism and finding stability in a constant shuffle of collisions and contentions in ever expanding arenas of liberty.

Because of an acceptance of setbacks, capitalist markets are more robust than socialist systems that plan for perfection. In the same way, successful people and companies have more failures than failures do. The successes use their faults and collisions as sources of new knowledge. Companies that try to banish chance by relying on market research and focus groups do less well than companies that freely make mistakes and learn from them.

Because of an ability to absorb shocks, stochastic systems in general are more stable than deterministic ones. Listening to the technology, we find that ethernets resonate to the deepest hymns and harmonies of our age.