Communications Challenges of the Digital Information Utility

Joel S. Birnbaum

The Internet and World Wide Web are forerunners of a digital information utility that in time will provide computing as well as information to society, just as other utilities provide water and electric power.

This is a turbulent, uncertain time in communications and computing because major changes are taking place at the same time in related measurement, computing, and communications technologies. This is a time when suppliers of technology, operators, and users alike must attempt to predict the future, yet that's never been harder. As Yogi Berra, the former manager of the New York Mets, once said, "Prophecy is hard, particularly about the future."

In this article, I will explore the idea that the Internet and the World Wide Web built upon it are the earliest form of what will become a digital information utility that eventually will bring information and computing to most homes, schools, and businesses. Because important breakthroughs are needed in high-speed communications, network management, security, and many other issues that result from the enormous increases in the scale of these emerging global nets, we might not see this utility for a decade or more. I believe, though, that market forces will ultimately cause these barriers to be overcome and that the resultant digital information utility will create a new style of computing that will both demand and enable a new type of high-speed communications infrastructure.

Pervasive Information Systems

For about 20 years, I have pursued a dream that one day computing would become a pervasive technology— that is, more noticeable by its absence than its presence, just as automobiles, television sets, telephones, and many other technologies have become part of everyday life for most people.

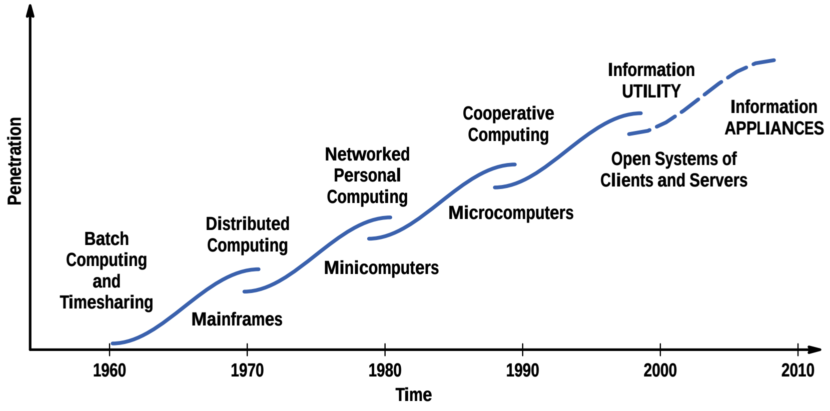
Joel S. Birnbaum

Joel S. Birnbaum is Hewlett-Packard's Senior Vice President of Research & Development and Director of HP Laboratories (the company's central research and development organization), a position he assumed in 1991. Joel joined HP in November 1980 as founding director of the Computer Research Center of HP Laboratories in Palo Alto, California. Before that, he spent 15 years at IBM Corporation's Thomas J. Watson Research Laboratory in Yorktown Heights, New York, where he last served as Director of Computer Sciences. His personal contributions are in the areas of distributed computer system architecture, real-time data acquisition, analysis and control, and RISC processor architecture. He holds a Bachelor's degree in Engineering Physics from Cornell University (1960), and MS (1961) and PhD (1965) degrees in Nuclear Physics from Yale University. He is a member of several professional science, engineering, and computer organizations. He also serves on engineering advisory councils of Carnegie Mellon University, Stanford University, and the University of California at Berkeley.

Figure 1 shows the progression of computing technology over several decades. It shows computing elements below the S-curves and shows how they are interconnected above the curves.

Figure 1

Progression of computing technology over several decades.



A key transition occurred at the beginning of the 1990s. At that point, our ability to make devices small enough let us begin specializing them according to their function and producing them in sufficient volume that they could be priced for mass markets. The architected interface between the requesters and providers of services—the so-called clients and servers that produced open systems—is the enabling transition for computing technology to become pervasive within the next decade. Because only people born after a technology has been invented think of it as part of the environment, and not as a technology, tomorrow's children will think of computers the way we think of telephones and TV today.

For a technology to become truly pervasive, it must transcend being merely manufacturable and commonplace. It must become intuitively accessible to ordinary people and deliver enough value to justify the large investment required to build the supporting infrastructure. Think for a moment how easily most of us can use a telephone or drive a car with an automatic transmission; think too of the investment needed to build the infrastructure behind the telephone and highway systems.

In the next century, computers will often be embedded in devices dedicated to performing a particular task, and they will let people think in terms of the task being performed instead of the underlying technology. A comparison with electric motors and consumer appliances helps convey this idea. We think of the function of a washing machine, electric toothbrush, or VCR, not about the motors within. In fact, the motors are very different and have required sophisticated engineering development. However, we do not think of this when we use these appliances (note that they are almost always named by what they do). We do not know, nor do we care, how many motors we own; the same will be true for computers when they too are pervasive.

I believe that we will see the emergence of information appliances, which I view as devices that will offer a means to a particular end and whose use will seem quite natural to most people. They will differ greatly from portable general-purpose computers because of what these appliances will do and because they will be easier to use. We expect appliances to evolve to hide their own complexity, just as today's one-button automatically tuned television set has replaced its less-complex but harder-to-adjust ancestor. Because of continued advances in semiconductors and software, tomorrow's information appliances will be able to do the same and likely will conform to a simple model of use, like the "Neutral-Drive-Reverse" of an automatic transmission. The Internet and the World Wide Web could not have achieved such dramatic growth without the concurrent invention of the intuitive point-and-click-style browsers, which let users think about what they want to access without concern for where or how it is stored.

The Digital Information Utility

I mentioned earlier that the Internet and World Wide Web are the forerunners of a digital information utility that in time will provide computing as well as information to society, just as other utilities provide water and electric power. User expectations certainly will change once we begin thinking of information flowing over a digital utility. Like all successful infrastructures, the digital information infrastructure will have to be so dependable that we notice it only when it fails. It will also have to be secure, endure over generations, be found everywhere, and serve a purpose important to almost all of the population.

Like other pervasive infrastructures, it will foster new industries. Think for a moment of how the public highway infrastructure, which made the personal automobile a pervasive technology, spawned many new industries such as auto insurance, car rentals, driver training, car washes, muffler shops, and gasoline stations, to name just a few.

In a world in which clients—whether general-purpose computers or information appliances—connect to this utility, most people will pay for their computing by usage. This will change what is now a capital investment into a competitive service such as that provided by the electric power and water utilities. This will not be the same as timesharing, which was proprietary, closed, and often location-dependent. In client-utility computing, open resources, located arbitrarily, will be combined as needed. In fact, one day, when the communications bandwidth is great enough and the costs low enough, it will no longer matter where the computers are located or which manufacturer makes them. Today's suicidal obsolescence schedule will be replaced by the capacity requirements of the service provider. Just as users don't notice when their electric utility replaces a generator at the power plant, so information utility users shouldn't detect upgrades.

Quality of service will become a crucial competitive differentiator at the systems level because users will expect the information utility to be available, ready and waiting, just as today they pick up a phone and expect a dial tone.

A key aspect of this information utility is that it will be digital. This means that devices that send and receive information can be independent of each other instead of having to be designed in matched pairs, as analog fax machines or TV and radio transmitters and receivers must be today. Given the appropriate interchange standards, your digital fax machine, for example, could begin communicating with your television, which could also receive your newspaper as well as diagnostic information from your car.

This change also means that many of the decisions now made at the transmission end instead can be done at the receiver. Once information interchange standards are in place, appliance and peripheral families will emerge, and many of these will be able to communicate directly without the intervention or invocation of a general-purpose computer and the attendant, cumbersome general-purpose operating system. In fact, many, if not most, of the computers of the next generation will be enormously powerful nonreprogrammable invisible processors. I'm thinking here, for example, of the several dozen embedded processors hidden in a modern car that control the car's ignition, suspension, braking, steering, engine-management, and climate-control systems and that provide diagnostic information to the driver and the mechanic. No one has ever asked a car dealer whether these computers run UNIX® or Windows® NT—the interface is the steering wheel or brake pedal, and function is augmented without the need to introduce an unfamiliar interface or force users to pay attention to the inner workings of the software. The embedded processors of the future will enable users to invoke powerful functions at a much higher level of abstraction than is common today.

Everything Could Have a Web Page

It is already practical to embed a web server, which can be a small amount of code in a very inexpensive microprocessor, in individual devices. This means that everyday appliances can have a web page at negligible cost. With a conventional browser, you could easily check things like home security and heating or cooling systems and control individual appliances like a hot tub or toaster from anywhere that you can click on a web page. We may even make it possible for ordinary adults to program a VCR! In fact, any instrument's front panel, or a subset or simpler replacement of it, could be viewed from anywhere by people with access rights. This would let engineers and scientists inexpensively collaborate across vast distances and make practical remote maintenance for low-cost devices. Virtual instruments, which are networked combinations of real ones, could be created. Notification capability can be incorporated; imagine, for example, a printer that automatically signals a supplier to ship a replacement toner or ink cartridge when some usage threshold has been crossed.

If the communications pipe is fast enough and cheap enough, this also means that you can do things far away that appear to be done locally. Distributed computing power, remote distributed measurements, very rich user interfaces in small, inexpensive devices, and remote printing and scanning all become practical and will enable applications at a scale and cost unthinkable today.

I believe that client-utility computing will provoke as great a change in the computing industry as open systems did in this decade. The new paradigm will do for computation what the Web did for data and will produce such dramatic decreases in cost of ownership, with concomitant increases in uptime and accessibility, that those companies that do not react to the opportunity will not be able to satisfy customer expectations. History tells us that this can have dire consequences.

Much technology—consisting mostly of middleware to address issues of scale, interoperability, security, robustness, and end-to-end systems management—is needed. Developing this technology won't be easy, but HP and many of the world's computer and software companies and research laboratories are today busy developing practical engineering solutions to these problems. It is just a matter of time before the greatly decreased cost of ownership of client-utility systems, coupled with their functional advantages, makes this the way that most people will access multimedia information and solve problems requiring computation.

Needed: Communications Breakthroughs

A number of telecommunications technology breakthroughs are needed for this dream of pervasive information systems to come true. We are entering a period that some at HP Laboratories have begun calling the "Tera Era" because the demanding technical requirements needed to support inexpensive high-bandwidth networks are measured in quantities of trillions—that is, trillion-bit-per-second transmission, trillion-byte memories, and trillion-instructions-per-second computers. The viability of the digital infrastructure for multimedia documents, appliance-utility computing, and distributed remote measurements depends on a number of key technologies coming together. Particularly critical is having high bandwidth at low cost, and many people are working to bring this about.

Other articles in this issue discuss some of these technologies, but I would like to focus for a minute on important developments taking place in optics technology that could have a huge effect on the telecommunications system. The theoretical capacity of the fiber is vast—something on the order of 25,000 GHz. In fact, though, we typically don't take advantage of anything but a tiny fraction of that capability. If we could find a way to send and switch more signals through a fiber-optic cable, we could increase the system's capacity by at least two or three orders of magnitude. It would mark a radical change in such systems because for the first time the electronic switching components would become the bottleneck instead of the transmission lines connected to them. This suggests that an all-optical system is needed, and the most promising approach is called wavelength-division multiplexing (WDM). This is not a new idea, but it is now becoming realistic to think of such systems being widely deployed in the not-too-distant future.

Fiber technology has advanced over the last 20 years to the point that the distance across which a usable light pulse can be sent has grown from a fraction of a kilometer to hundreds of kilometers, and cost has plummeted concurrently. If we could transmit and then select all the theoretically possible frequencies, a wavelength-division multiplexed system could work the way that a radio does. That is, at one end, a particular station chooses its frequency; at the other end, the user has what amounts to a big dial. Depending on whom you want to be connected to, you turn the dial and change the frequency of the receiver. Sometimes a movie comes over the pipe, sometimes a newspaper, sometimes the result of an economic model from a distant supercomputer, sometimes your child's voice.

This is, of course, a lot more easily said than done. Early WDM systems used mechanical electrooptical devices for frequency selection. Essentially, movable gratings would preferentially select a single frequency. However, this is not very practical because such interferometers are slow, expensive, and limited in frequency range. Laboratories today are investigating promising low-cost high-performance optoelectronic transceivers and other electronic and photonic devices that could replace or augment existing terminal equipment.

We also need other new or dramatically improved technology, including multigigabit-per-second semiconductor lasers and photodetectors, multigigabit-per-second integrated electronics for laser drivers and clock and data recovery, large input/output cross-connect switches, and optical circuits such as add/drop filters and wavelength converters.

If these technology challenges can be met, the results will be fabulous. Just one strand of fiber, in principle, could carry a billion separate audio telephone calls—all the telephone calls for a large city. Many users of the digital information utility will want enough bandwidth , say, a 100-MHz channel, for high-resolution real-time multimedia. That's ten times an entire Ethernet for each user. With WDM, a single optical fiber could carry a quarter million of these superbandwidth channels.

Network Measurement and Management

The digital utility will also require much more sophisticated network management because users will expect the quality of service they now get from their telephone system. The Internet as we know it today is nowhere close to this.

To manage applications end-to-end, there can be no alternative to making distributed, continuous measurements of performance, capacity, and usage statistics and relating them to a model of the system, or of part of it. By measuring what's actually happening across the entire system, you can, among other things, adjust its capacity, detect many types of fraud, predict where performance bottlenecks are likely to occur, locate outages, and identify unused parts of the system for reserve capacity. Once again, Internet technology will reduce the complexity and lower the cost. In fact, HP has built a successful prototype of such a system that is operating today in a London cellular telephony trial.

Network management and measurement are the Achilles heel of the robust, flexible infrastructure that operators and users want. I think these will be the pacing core technologies of the Information Age, and they demand the immediate attention of telecommunications and computer manufacturers alike. It will be important to develop international standards; and while the technology will be similar in some ways to that now used by distributed computer systems and the telephone networks, it will have to solve problems of scale and speed that people have not had to deal with before. Systems with millions, or even tens of millions, of nodes will be commonplace, and the heterogeneity of the hardware and software will be unprecedented.

Once developed, this core capability of measuring and managing the evolved Internet itself will be extended to enable a vast range of distributed measurement applications that today would require specialized, expensive systems. The utility provides the infrastructure to link distributed sensors at low cost in a ubiquitous way. Many industries, such as healthcare, agriculture, and transportation, will be transformed by this ability to operate on a continental basis.

Security

An issue much in the news is security, which stands in the way of many commercial applications. Security is a hard problem because we're on a difficult tightrope. On one hand, we want maximum interoperability for authorized, authenticated users among all computers, all appliances, and all nations. Essentially, this means that we impose no barriers to interconnection. But this aim of unfettered interoperability conflicts with access control, privacy, and system integrity. Citizens' rights to privacy are often in direct conflict with the needs of national security and criminal justice agencies. Building a truly open, global system is technically antithetical to a secure system with good performance and attractive cost, creating a difficult set of technical, social, and political trade-offs to be resolved, and compromises are not easily achieved.

I believe that the security issue will be solved more easily than the bandwidth and network-management problems. Most of the world's computing and communication companies are working feverishly on security-related issues because these issues stand in the way of the profits that can be reaped from electronic commerce. Although no perfect solution is likely to emerge, I think that an acceptable de facto standard, based on a sensible range of tradeoffs and compromises, will emerge because of the overwhelming financial and market forces demanding it.

Conclusion

In closing, the telephone system that we know today and the Internet and Web technology built upon it are precursors of a global digital information utility. The Internet is delivering today at low bandwidth and relatively high cost entirely new classes of information and services. As the bandwidth and usage grow, costs inevitably will decline. The resulting information utility will dramatically change computing as well as telephony and the delivery of multimedia information. In time, we will think of today's systems as quaint.

▶ Go to Next Article
▶ Go to Journal Home Page