Chapter 1

# Burden's Wheel

In 1851, in a field beside an ironworks in upstate New York, Henry Burden built a magnificent machine. Resembling a giant bicycle wheel, with dozens of thick iron spokes radiating from a massive central hub, Burden's machine was the largest industrial waterwheel in the country and the most powerful anywhere. Fed by a fast-moving spout of water diverted from the nearby Wynantskill River, the 60-foot-tall, 250-ton behemoth could produce as much as 500 horsepower when turning at its top speed of two and a half times a minute. The power was harnessed and delivered to the drill presses, grinding wheels, forge hammers, and lathes used by Burden's workers through an intricate system of gears, belts, and pulleys.

Henry Burden had a genius for industrial invention. A Scottish engineer, he emigrated to the United States in 1819, at the age of twenty-eight, to take a position with a farming-tool manufacturer in Albany. Within months he had invented the country's first cultivator for preparing seed beds and designed an improved version of a plow. Three years later, he moved to the nearby town of Troy to manage the Troy Iron and Nail Factory, which he would eventually buy and rename the Burden Iron Works. It didn't take him long to

see the boon in the plant's location near the junction of the Hudson River and the freshly dug Erie Canal. If he could boost the factory's output, he would be able to ship its products to new markets throughout the Northeast and Midwest. He set to work mechanizing what had for centuries been a local industry built on the handiwork of blacksmiths and other craftsmen. Within a dozen years, he had created machines that automated the production of nails and railroad spikes, and in 1835 he invented the Burden Horseshoe Machine, an ingenious contraption that turned bars of iron into finished shoes at the rate of one a second. In his spare time, Burden also managed to design a large ocean-going steamboat that became the model for many subsequent ferries and cruise ships.

But Burden's greatest creation, the one that made him rich as well as famous, was his wheel. Dubbed "the Niagara of Waterwheels" by a local poet, its unprecedented size and power provided the Burden Iron Works with a decisive advantage over other manufacturers. The company was able to expand the yield and efficiency of its factory, producing more shoes, spikes, and other goods with fewer workers and in less time than rivals could manage. It won a contract to supply nearly all the horseshoes used by the Union army during the Civil War, and it became one of the major suppliers of spikes to American railroads as they extended their lines across the country.

For Burden, the efficient production of mechanical power turned out to be every bit as important to his company's success as the skill of his workers and even the quality of his products. Like other factory owners of the time, he was as much in the business of manufacturing energy as manufacturing goods.

But a visitor to the Burden Iron Works in the early years of the twentieth century would have come upon a surprising sight. The great waterwheel stood idle in the field, overgrown with

weeds and quietly rusting away. After turning nonstop for fifty years, it had been abandoned. Manufacturers didn't have to be in the power-generation business anymore. They could run their machines with electric current generated in distant power plants by big utilities and fed to their factories over a network of wires. With remarkable speed, the new utilities took over the supply of industrial power. Burden's wheel and thousands of other private waterwheels, steam engines, and electric generators were rendered obsolete.

What made large-scale electric utilities possible was a series of scientific and engineering breakthroughs—in electricity generation and transmission as well as in the design of electric motors—but what ensured their triumph was not technology but economics. By supplying electricity to many buyers from central generating stations, the utilities achieved economies of scale in power production that no individual factory could match. It became a competitive necessity for manufacturers to hook their plants up to the new electric grid in order to tap into the cheaper source of power. The success of the utilities fed on itself. As soon as they began to supply current to factories, they were able to expand their generating capacity and scale economies even further, achieving another great leap in efficiency. The price of electricity fell so quickly that it soon became possible for nearly every business and household in the country to afford electric power.

The commercial and social ramifications of the democratization of electricity would be hard to overstate. Electric light altered the rhythms of life, electric assembly lines redefined industry and work, and electric appliances brought the Industrial Revolution into the home. Cheap and plentiful electricity shaped the world we live in today. It's a world that didn't exist a mere hundred years ago, and yet the transformation that has played out over just a few generations

has been so great, so complete, that it has become almost impossible for us to imagine what life was like before electricity began to flow through the sockets in our walls.

TODAY, WE'RE IN the midst of another epochal transformation, and it's following a similar course. What happened to the generation of power a century ago is now happening to the processing of information. Private computer systems, built and operated by individual companies, are being supplanted by services provided over a common grid—the Internet—by centralized data-processing plants. Computing is turning into a utility, and once again the economic equations that determine the way we work and live are being rewritten.

For the past half century, since the first mainframe computer was installed in a corporate data center, businesses have invested many trillions of dollars in information technology. They've assembled hardware and software into ever more complex systems to automate nearly every facet of their operations, from the way they buy materials and supplies to the way they manage their employees to the way they deliver their products to customers. They've housed those systems on-site, in their own plants and offices, and they've maintained them with their own staffs of technicians. Just as Henry Burden and other manufacturers competed in part on the sophistication of their power systems, so modern companies have competed on the sophistication of their computer systems. Whatever their main business might be, they've had no choice but to also be in the business of data processing.

Until now.

Capitalizing on advances in the power of microprocessors and the capacity of data storage systems, fledgling utilities are beginning to build massive and massively efficient information-processing plants,

and they're using the broadband Internet, with its millions of miles of fiber-optic cable, as the global grid for delivering their services to customers. Like the electric utilities before them, the new computing utilities are achieving economies of scale far beyond what most companies can achieve with their own systems.

Seeing the economic advantages of the utility model, corporations are rethinking the way they buy and use information technology. Rather than devoting a lot of cash to purchasing computers and software programs, they're beginning to plug into the new grid. That shift promises not only to change the nature of corporate IT departments but to shake up the entire computer industry. Big tech companies—Microsoft, Dell, Oracle, IBM, and all the rest—have made tons of money selling the same systems to thousands of companies. As computing becomes more centralized, many of those sales will dry up. Considering that businesses spend well over a trillion dollars a year on hardware and software, the ripple effects will be felt throughout the world economy.

But this is hardly just a business phenomenon. Many of the most advanced examples of utility computing are aimed not at companies but at people like you and me. The best example of all is probably Google's search engine. Think about it: what is Google but a giant information utility? When you need to search the Internet, you use your Web browser to connect to the vast data centers that Google has constructed in secret locations around the world. You type in a keyword, and Google's network of hundreds of thousands of interlinked computers sorts through a database of billions of Web pages, draws out the few thousand that best match up with your keyword, arranges them in order of relevance, and shoots the results back through the Internet to your screen—usually in less than a second. That amazing computing feat, which Google repeats hundreds of millions of times a day, doesn't happen inside your PC. It *couldn't*

happen inside your PC. Rather, it happens miles away, maybe on the other side of the country, maybe even on the other side of the globe. Where's the computer chip that processed your last Google search? You don't know, and you don't care—any more than you know or care which generating station produced the kilowatts that light the lamp on your desk.

All historical models and analogies have their limits, of course, and information technology differs from electricity in many important ways. But beneath the technical differences, electricity and computing share deep similarities—similarities that are easy for us to overlook today. We see electricity as a "simple" utility, a standardized and unremarkable current that comes safely and predictably through outlets in our walls. The innumerable applications of electric power, from televisions and washing machines to machine tools and assembly lines, have become so commonplace that we no longer consider them to be elements of the underlying technology—they've taken on separate, familiar lives of their own.

It wasn't always so. When electrification began, it was an untamed and unpredictable force that changed everything it touched. Its applications were as much a part of the technology as the dynamos, the power lines, and the current itself. As with today's computer systems, all companies had to figure out how to apply electricity to their own businesses, often making sweeping changes to their organizations and processes. As the technology advanced, they had to struggle with old and often incompatible equipment—the "legacy systems," to use a modern computer term, that can lock businesses into the past and impede progress—and they had to adapt to customers' changing needs and expectations. Electrification, just like computerization, led to complex, far-reaching, and often bewildering changes for individual companies and entire industries—and, as households began to connect to the grid, for all of society.

At a purely economic level, the similarities between electricity and information technology are even more striking. Both are what economists call general purpose technologies. Used by all sorts of people to do all sorts of things, they perform many functions rather than just one or a few. General purpose technologies, or GPTs, are best thought of not as discrete tools but as platforms on which many different tools, or applications, can be constructed. Compare the electric system to the rail system. Once railroad tracks are laid, you can pretty much do only one thing with them: run trains back and forth carrying cargo or passengers. But once you set up an electric grid, it can be used to power everything from robots in factories to toasters on kitchen counters to lights in classrooms. Because they're applied so broadly, GPTs offer the potential for huge economies of scale—if their supply can be consolidated.

That's not always possible. Steam engines and waterwheels were general purpose technologies that didn't lend themselves to centralization. They had to be located close to the point where their power was used. That's why Henry Burden had to build his wheel right next to his factory. If he had built it even a few hundred yards away, all the energy produced by the wheel would have been consumed in turning the long shafts and belts required to convey the energy to the factory. There wouldn't have been any left to power workers' machines.

But electricity and computing share a special trait that makes them unique even among the relatively small set of general purpose technologies: they can both be delivered efficiently from a great distance over a network. Because they don't have to be produced locally, they can achieve the scale economies of central supply. Those economies, though, can take a long time to be fully appreciated and even longer to be fully exploited. In the early stages of a GPT's development, when there are few technical standards and no

broad distribution network, the technology is impossible to furnish centrally. Its supply is by necessity fragmented. If a company wants to tap into the power of the technology, it has to purchase the various components required to supply it, install those components at its own site, cobble them together into a working system, and hire a staff of specialists to keep the system running. In the early days of electrification, factories had to build their own generators if they wanted to use the power of electricity—just as today's companies have had to set up their own information systems to use the power of computing.

Such fragmentation is wasteful. It imposes large capital investments and heavy fixed costs on firms, and it leads to redundant expenditures and high levels of overcapacity, both in the technology itself and in the labor force operating it. The situation is ideal for the suppliers of the components of the technology—they reap the benefits of overinvestment—but it's not sustainable. Once it becomes possible to provide the technology centrally, large-scale utility suppliers arise to displace the private providers. It may take decades for companies to abandon their proprietary supply operations and all the investments they represent. But in the end the savings offered by utilities become too compelling to resist, even for the largest enterprises. The grid wins.

At a conference in Paris during the summer of 2004, Apple introduced an updated version of its popular iMac computer. Since its debut in 1998, the iMac had always been distinguished by its unusual design, but the new model was particularly striking. It appeared to be nothing more than a flat-panel television, a rectangular screen encased in a thin block of white plastic and mounted on an aluminum pedestal. All the components of the computer itself—the chips, the drives, the cables, the connectors—were hidden behind

the screen. The advertising tagline wittily anticipated the response of prospective buyers: "Where did the computer go?"

But the question was more than just a cute promotional pitch. It was, as well, a subtle acknowledgment that our longstanding idea of a computer is obsolete. While most of us continue to depend on personal computers both at home and at work, we're using them in a very different way than we used to. Instead of relying on data and software that reside inside our computers, inscribed on our private hard drives, we increasingly tap into data and software that stream through the public Internet. Our PCs are turning into terminals that draw most of their power and usefulness not from what's inside them but from the network they're hooked up to—and, in particular, from the other computers that are hooked up to that network.

The change in the way we use computers didn't happen overnight. Primitive forms of centralized computing have been around for a long time. In the mid-1980s, many early PC owners bought modems to connect their computers over phone lines to central databases like CompuServe, Prodigy, and the Well—commonly known as "bulletin boards"—where they exchanged messages with other subscribers. America Online popularized this kind of online community, greatly expanding its appeal by adding colorful graphics as well as chat rooms, games, weather reports, magazine and newspaper articles, and many other services. Other, more specialized databases were also available to scholars, engineers, librarians, military planners, and business analysts. When, in 1990, Tim Berners-Lee invented the World Wide Web, he set the stage for the replacement of all those private online data stores with one vast public one. The Web popularized the Internet, turning it into a global bazaar for sharing digital information. And once easy-to-use browsers like Netscape Navigator and Internet Explorer became freely available in the mid-1990s, we all went online in droves.

Through the first decade of its existence, however, the World Wide Web was a fairly prosaic place for most of us. We used it mainly as a giant catalog, a collection of "pages" bound together with hyperlinks. We "read" the Web, browsing through its contents in a way that wasn't so different from the way we'd thumb through a pile of magazines. When we wanted to do real work, or play real games, we'd close our Web browser and launch one of the many programs installed on our own hard drive: Microsoft Word, maybe, or Aldus Pagemaker, or Encarta, or Myst.

But beneath the Web's familiar, page-like surface lay a set of powerful technologies, including sophisticated protocols for describing and transferring data, that promised not only to greatly magnify the usefulness of the Internet but to transform computing itself. These technologies would allow all the computers hooked up to the Net to act, in effect, as a single information-processing machine, easily sharing bits of data and strings of software code. Once the technologies were fully harnessed, you'd be able to use the Internet not just to look at pages on individual sites but to run sophisticated software programs that might draw information from many sites and databases simultaneously. You'd be able not only to "read" from the Internet but to "write" to it as well—just as you've always been able to read from and write to your PC's hard drive. The World Wide Web would turn into the World Wide Computer.

This other dimension of the Internet was visible from the start, but only dimly so. When you ran a Web search on an early search engine like AltaVista, you were running a software program through your browser. The code for the software resided mainly on the computer that hosted AltaVista's site. When you did online banking, shifting money between a checking and a savings account, you were also using a utility service, one that was running on your bank's computer rather than your own. When you used your browser to

check your Yahoo or Hotmail email account, or track a FedEx shipment, you were using a complicated application running on a distant server computer. Even when you used Amazon.com's shopping-cart system to order a book—or when you subsequently posted a review of that book on the Amazon site—you were tapping into the Internet's latent potential.

For the most part, the early utility services were rudimentary, involving the exchange of a small amount of data. The reason was simple: more complex services, the kind that might replace the software on your hard drive, required the rapid transfer of very large quantities of data, and that just wasn't practical with traditional, low-speed dial-up connections. Running such services would quickly overload the capacity of telephone lines or overwhelm your modem. Your PC would grind to a halt. Before sophisticated services could proliferate, a critical mass of people had to have high-speed broadband connections. That only began to happen late in the 1990s during the great dotcom investment boom, when phone and cable companies rushed to replace their copper wires with optical fibers—hair-thin strands of glass that carry information as pulses of light rather than electric currents—and retool their networks to handle virtually unlimited quantities of data.

The first clear harbinger of the second coming of the Internet—what would eventually be dubbed Web 2.0—appeared out of nowhere in the summer of 1999. It came in the form of a small, free software program called Napster. Written over a few months by an eighteen-year-old college dropout named Shawn Fanning, Napster allowed people to share music over the Internet in a whole new way. It scanned the hard drive of anyone who installed the program, and then it created, on a central server computer operated by Fanning, a directory of information on all the song files it found, cataloging their titles, the bands that performed them, the albums

they came from, and their audio quality. Napster users searched this directory to find songs they wanted, which they then downloaded directly from other users' computers. It was easy and, if you had a broadband connection, it was fast. In a matter of hours, you could download hundreds of tunes. It's no exaggeration to say that, at Napster's peak, almost every work of popular music that had ever been digitally encoded onto a compact disk—and many that had never appeared on a disk—could be found and downloaded for free through the service.

Napster, not surprisingly, became wildly popular, particularly on college campuses where high-speed Net connections were common. By early 2001, according to an estimate by market researcher Media Metrix, more than 26 million people were using the service, and they were spending more than 100 million hours a month exchanging music files. Shawn Fanning's invention showed the world, for the first time, how the Internet could allow many computers to act as a single shared computer, with thousands or even millions of people having access to the combined contents of previously private databases. Although every user had to install a little software program on his own PC, the real power of Napster lay in the network itself—in the way it created a central file-management system and the way it allowed data to be transferred easily between computers, even ones running on opposite sides of the planet.

There was just one problem. It wasn't legal. The vast majority of the songs downloaded through Napster were owned by the artists and record companies that had produced them. Sharing them without permission or payment was against the law. The arrival of Napster had turned millions of otherwise law-abiding citizens into digital shoplifters, setting off the greatest, or at least the broadest, orgy of looting in history. The musicians and record companies fought back, filing lawsuits charging Fanning's company with

copyright infringement. Their legal counterattack culminated in the closing of the service in the summer of 2001, just two years after it had launched.

Napster died, but the business of supplying computing services over the Internet exploded in its wake. Many of us now spend more time using the new Web services than we do running traditional software applications from our hard drives. We rely on the new utility grid to connect with our friends at social networks like MySpace and Facebook, to manage our photo collections at sites like Flickr and Photobucket, to create imaginary selves in virtual worlds like World of Warcraft and Disney's Club Penguin, to watch videos through services like YouTube and Joost, to write blogs with WordPress or memos with Google Docs, to follow breaking news through feed readers like Rojo and Bloglines, and to store our files on "virtual hard drives" like Omnidrive and Box.

All these services hint at the revolutionary potential of the information utility. In the years ahead, more and more of the information-processing tasks that we rely on, at home and at work, will be handled by big data centers located out on the Internet. The nature and economics of computing will change as dramatically as the nature and economics of mechanical power changed in the early years of the last century. The consequences for society—for the way we live, work, learn, communicate, entertain ourselves, and even think—promise to be equally profound. If the electric dynamo was the machine that fashioned twentieth-century society—that made us who we are—the information dynamo is the machine that will fashion the new society of the twenty-first century.

Lewis Mumford, in his 1970 book *The Pentagon of Power,* the second volume of his great critique of technology *The Myth of the Machine,* made an eloquent case against the idea that technologi-

cal progress determines the course of history. "Western society," he wrote, "has accepted as unquestionable a technological imperative that is quite as arbitrary as the most primitive taboo: not merely the duty to foster invention and constantly to create technological novelties, but equally the duty to surrender to these novelties unconditionally, just because they are offered, without respect to their human consequences." Rather than allowing technology to control us, Mumford implied, we can control technology—if only we can muster the courage to exert the full power of our free will over the machines we make.

It's a seductive sentiment, one that most of us would like to share, but it's mistaken. Mumford's error lay not in asserting that as a society we pursue and embrace technological advances with little reservation. That's hard to dispute. His error lay in suggesting that we might do otherwise. The technological imperative that has shaped the Western world is not arbitrary, nor is our surrender to it discretionary. The fostering of invention and the embrace of the new technologies that result are not "duties" that we have somehow chosen to accept. They're the consequences of economic forces that lie largely beyond our control. By looking at technology in isolation, Mumford failed to see that the path of technological progress and its human consequences are determined not simply by advances in science and engineering but also, and more decisively, by the influence of technology on the costs of producing and consuming goods and services. A competitive marketplace guarantees that more efficient modes of production and consumption will win out over less efficient ones. That's why Henry Burden built his wheel, and it's why that wheel was left to rust a few decades later. Technology shapes economics, and economics shapes society. It's a messy process—when you combine technology, economics, and human nature, you get a lot of variables—but it has an inexorable logic, even if we can trace

it only in retrospect. As individuals, we may question the technological imperative and even withstand it, but such acts will always be lonely and in the end futile. In a society governed by economic trade-offs, the technological imperative is precisely that: an imperative. Personal choice has little to do with it.

We see the interplay of technology and economics most clearly at those rare moments when a change takes place in the way a resource vital to society is supplied, when an essential product or service that had been supplied locally begins to be supplied centrally, or vice versa. Civilization itself emerged when food production, decentralized in primitive hunter-gatherer societies, began to be centralized with the introduction of the technologies of agriculture. Changes in the supply of other important resources—resources as diverse as water, transportation, the written word, and government—also altered the economic trade-offs that shape society. A hundred years ago, we arrived at such a moment with the technologies that extend man's physical powers. We are at another such moment today with the technologies that extend our intellectual powers.

The transformation in the supply of computing promises to have especially sweeping consequences. Software programs already control or mediate not only industry and commerce but entertainment, journalism, education, even politics and national defense. The shock waves produced by a shift in computing technology will thus be intense and far-reaching. We can already see the early effects all around us—in the shift of control over media from institutions to individuals, in people's growing sense of affiliation with "virtual communities" rather than physical ones, in debates over the security of personal information and the value of privacy, in the export of the jobs of knowledge workers, even in the growing concentration of wealth in a small slice of the population. All these trends either spring from or are propelled by the rise of Internet-based

computing. As information utilities grow in size and sophistication, the changes to business and society—and to ourselves—will only broaden. And their pace will only accelerate.

Many of the characteristics that define American society came into being only in the aftermath of electrification. The rise of the middle class, the expansion of public education, the flowering of mass culture, the movement of the population to the suburbs, the shift from an industrial to a service economy—none of these would have happened without the cheap current generated by utilities. Today, we think of these developments as permanent features of our society. But that's an illusion. They're the by-products of a particular set of economic trade-offs that reflected, in large measure, the technologies of the time. We may soon come to discover that what we assume to be the enduring foundations of our society are in fact only temporary structures, as easily abandoned as Henry Burden's wheel.