Network Engineering

24

A Highway for Ideas

David Clark

Courtesy of David Clark

Key Concepts
from Previous Chapters

1 Process
1 Researching the Problem
1 Building the Prototype
1 Testing the Prototype

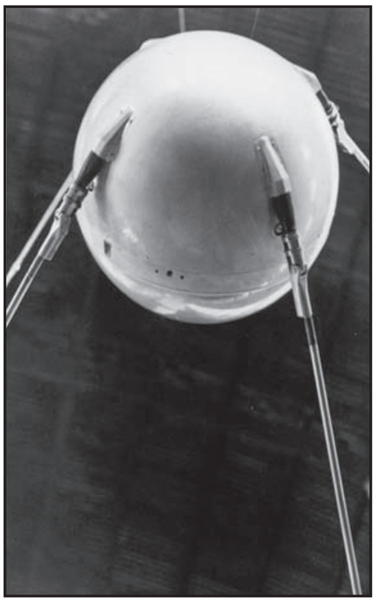
Scientists attempt to understand the world as it is. Engineers, on the other hand, take the principles from science and ask, "What is possible? Am I limited by what I know or can I reach further? Can I look at today's scientific understandings with new eyes and create something that has never existed before?" When an engineer invents something that profoundly changes the world, such as the telephone or the computer, you could say that the engineer has turned science into a constructive art, an art of the imagination. Many believe that in the virtual space of the Internet, the only real limit is our imagination. That's exactly what the creators of the Internet envisioned.

My name is David Clark, and I am a computer scientist. I teach network engineering and design at the Massachusetts Institute of Technology (MIT). Some people refer to me as one of the "fathers" of the Internet because I was on the team that designed it. I like to refer to myself as the Internet's "first cousin."

That's because so many people had a hand in developing a way for people in different locations to exchange information by linking their computers together.

Those of us on that design team never imagined that the Internet would become the incredible resource it is today. It makes information about any topic available to any user 24 hours a day, and allows people from different cultures, languages, nations, and religions to gather on common ground. Who could have foreseen that the Internet would become so widespread that people would have trouble thinking of what the world may have been like before? Many young people today have a hard time imagining life without downloadable music, Instant Messenger, e-mail, or the World Wide Web.

The roots of the Internet can be traced all the way back to World War II and its aftermath, when the United States started investing more in science and engineering to ensure a better future for our country and the world. These were complex times for the United States. In the 1940s and 1950s, society witnessed how scientists and engineers could alter the course of history. A group of researchers and engineers involved in the Manhattan Project created the first nuclear bomb by 1945. The bombs dropped on two cities in Japan, ending World War II, led to the loss of hundreds of thousands of lives. In the early 1950s, President Eisenhower oversaw the creation of the national interstate highway system. Very quickly, Americans achieved a mobility enjoyed by no other nation, but we also have suffered the unintended consequences of the automobile's widespread use.



Courtesy of NASA

Sputnik was the first artificial satellite put into orbit.

Perhaps most importantly, in 1957 the former Soviet Union launched Sputnik, the first man-made satellite to orbit the Earth. At the time, the United States was involved in the "cold war" with the Soviet Union, a nation that has since dissolved into Russia and several countries in Eastern Europe and Central Asia. The U.S. competed with the Soviet Union for military, economic, and technological superiority. At the time, Sputnik symbolized a major victory for the Soviets by beating the U.S. into space. In response, the U.S. Defense Department established the Defense Advanced Research Projects Agency (DARPA) to keep the U.S. ahead in using technology. The U.S. launched its own satellite within four months.

DARPA quickly turned its attention to what was then an emerging technology: computers. In 1960 there were no desktop computers, only large mainframe computers that filled whole rooms. Personal computers—or PCs—would not arrive on the scene for another 20 years. Most people who used computers were researchers at universities or those working in the government. Only the biggest of companies could afford their own computer. While these computers could process large amounts of data and store files, they weren't used for much else.

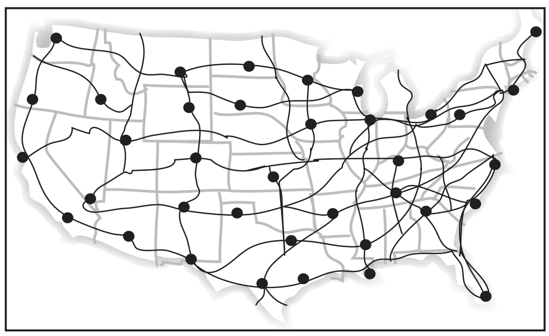
In the 1960s, engineers and computer scientists developed many of the important computer technologies we now take for granted. This includes software for sharing computers among many users, graphical displays (without which we would have no computer games today), and the computer disk (without which we would not be able to store all our music, mail, and files on our computers).

Around that time, DARPA started focusing on a new problem: networking the computers of research groups at universities and government labs. They thought that such a network might allow researchers to work together to advance the state of science and technology in America. So DARPA assembled a group of computer scientists to figure out just how to do that. First they built something called ARPAnet, the Advanced Research Project Agency Network. In the 1970s, they started on the Internet. That is when I joined the project.

Designing the Net

What would the Internet be like if you had designed it? Would it be any different than it is now? It's a question few people stop to ask themselves. The Internet, after all, seems like a huge, chaotic marketplace for ideas that is not controlled by anyone. It's hard to imagine that someone might have designed it to be that way. But we did—quite intentionally.

You see, as we designed the Internet, we drew inspiration from another new technology that was similar to what we wanted to create: the interstate highway system.



A map of the interstate highway system

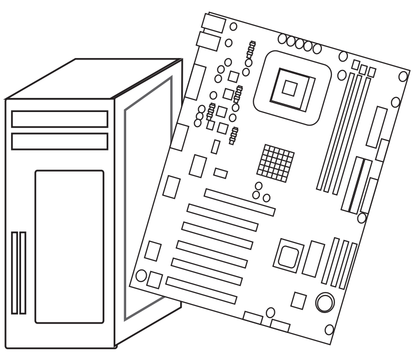
Why a highway system? Highways give users a lot of freedom. The engineers who designed the highway placed few limits on how vehicles were to be designed. They only made three main rules: a vehicle must have some type of power source to propel it; it must roll along the highway without destroying the surface of the road; and it must fit onto a highway lane.

It was up to vehicle designers to figure out the best vehicle design given those constraints. Yes, a car might use a combustion engine, but the highway itself does not require that. The engine could run on solar power, electricity, or natural gas—or even some power source not yet created. Sure, other forces such as the availability of fuel, national and state laws, or economics might affect an automobile design. But the highway itself imposed only a few restrictions. It simply provided a path for moving the machines.

In the early seventies, when we were coming up with the Internet's design, we knew that new communications technologies would emerge in the next several decades. But we had no idea what they would be. Satellites had just been launched, television was becoming mainstream, and the telephone system in the country was well established. Telephone systems were so well established that we decided to use telephone wires to connect the computers in our system. Fiber-optic cables, which allow for rapid transit of information across distances, did not yet exist, nor did CDs or DVDs. Those technologies all came later. We wanted to create a system that would be flexible enough to accommodate new technologies, like the highway system. Our system would not limit what traveled along its paths; it would only provide a way for the information to get from point A to point B.

Computer Science 101

Before I describe the rules we created for the Internet, you need to understand a few basic concepts about how computers function.


Motherboard

Computers are electronic devices that process electrical signals. The "brains" of a computer, called the ***motherboard,*** is a very complex, interconnected set of electronic devices, processor, memory, and control chips. The chips contain transistors, which are basically on-off switches that control electrical signals as they flow through the circuits. These transistors are made out of semiconductors. ***Semiconductors*** conduct currents, but not as well as a true conductor, such as copper. This makes semiconductors excellent materials for controlling the flow of electricity through a circuit.

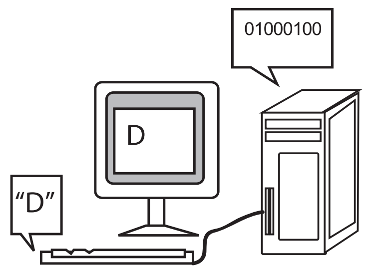
Computers recognize two states: the presence of an electrical signal or the absence of the electrical signal. In other words, they recognize when the signal is on or off. People have given the “on” and “off” signals numerical values so that they are easier for us to work with. The “on” signal is referred to as one, while the “off” signal is referred to as zero. You’ve probably heard of bits and bytes. A bit is a 1 or a 0. A byte is a string of eight bits.

So how does a computer process very complex information when it only recognizes two digits? Well, if you think about it, we only use ten digits, 0 through 9, which is called a base 10, or ***decimal,*** system. A computer describes every number, no matter how large, with 1s and 0s. So the computer uses a base 2, or binary, system, which we call binary code.

Every letter of the alphabet has been assigned a corresponding number in binary code. For instance, when a person taps the letter “D” on a computer keyboard, it will turn transistors on and off in the following pattern: 01000100. This code is called ASCII (American Standard Code for Information Interchange) and is pronounced “ask-kee.”

The strings of bits and bytes, made up of the ones and zeros that a computer reads, can be represented as ***digital signals.*** Digital signals are not continuous, but are made up of a discrete, limited set of elements. Many other communications signals are not digital. Electrical signals used in older telephone system were all ***analog signals.*** Analog signals are continuous and vary within some range, rather than using an on and off pattern. The graph of an analog signal is a continuous wave, with peaks and valleys.

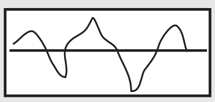
	Binary
0	0
1	1
2	10
3	11
4	100
5	101
6	110
7	111
8	1000
9	1001
10	1010



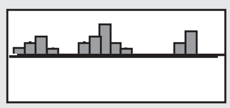
ASCII

Letter	Binary
A	01000001
B	01000010
C	01000011
D	01000100
E	01000101
F	01000110
G	01000111

An ***analog signal***
ranges over a continuous set of values.



A ***digital signal***
has a discrete limited set of numerical values, such as bits and bytes.

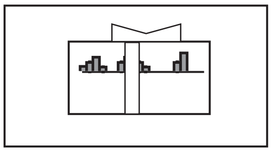


The "Rules" of the Internet

Now that you know something about the digital nature of computers, I can explain the limits Internet users face. Again, those of us on the original design team didn't want to limit users, but we did have to establish some ground rules.

1. Information Must Be Digital

The first Internet connected distant computers using analog signals over a telephone line. We wanted to make sure that the Internet could carry all sorts of data. Some of the data was created inside the computers, the programs, and data files. This data needs to be carried in digital form, which is how it was created. In addition, there are other analog signals besides voice, including music, television, and scanned images. The Internet needed to carry all these as well. So we decided that the Internet would transport digital bits, knowing that any analog signal could be converted to a digital signal (a sequence of zeros and ones).

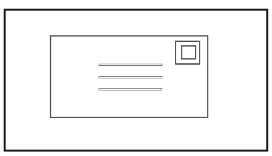


We settled on the idea of ***data packets,*** which are discrete packages of information. The data packets could contain any type of data as long as it was encoded by the computer into digital form. Then a modem, a communication device, would convert the digital signal from the computer into an analog signal before sending it over the telephone network. You can hear these signals today when a modem first connects. The tones, whistles, and buzzes are in the same frequencies that we use in speech, so they are carried by the telephone system quite naturally.

We also decided that it would be the user's responsibility to create data that could be encoded as a digital signal. The Internet's job would be to move these data packets along as efficiently as possible without any concern for what information the packets contained.

2. Every Computer on the Network Must Have an Address

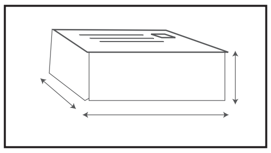
We needed to ensure that the system could deliver data to the correct recipient. We decided that every computer would have an address, just like how every home or business has a postal address. Today, every computer on the Internet has an IP, or Internet Protocol, address. Internet addresses are thirty-two bits long, expressed as a binary number, but we often write them down in a more convenient form.



You may have seen an Internet address. It looks like "71.232.4.16," four sets of numbers separated by dots. Each data packet is labeled with the address of the recipient. Routers located along the Internet pathway make sure that the data packets follow the shortest path possible to the addressee.

3. Packets Can Only Be So Large

As you may know, a data file can be very large. It may be a huge text file or contain graphic images or even short movies. We knew that hundreds and maybe thousands of computers would be sending files of all sizes across the Internet. (Little did we know that millions of users would one day communicate through the Internet!) If each computer had to wait for another computer on the network to send its data, the computers sending the largest messages would clog the network, making it difficult for shorter messages to get through in a reasonable period of time.

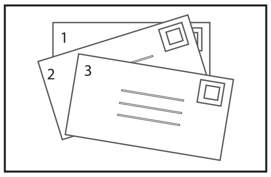


Waiting for a big message to get through would be like standing in line at the grocery store behind someone buying several carts full of groceries. The lines move fastest when customers limit the number of items in a shopping cart to ten or fifteen items. So imagine if store management did not allow anyone to pass through the checkout with more than fifteen items. A person with sixty items would have to pay for fifteen items, go to the back of the line, submit another fifteen items through the checkout, then return to the back of the line two more times to complete the sale. While this person had to go through the line four times, many others would be able to finish shopping quickly.

So our design team decided to limit the size of each packet. The actual limit has changed over time, but today each data packet sent over the Internet is normally no more than 1,500 bytes. If someone needs to send 3,000 bytes of information, the message is split into two packets. If the message is 30,000 bytes large, it is split into twenty packets.

4. Packets Must Be Sequenced

Now, in addition to carrying the address of the receiving computer, each data packet also has to have a sequence number so that the receiving computer can reassemble the message in the correct order. If there were no sequence numbers, it would be like sending several postcards to a friend without numbering the postcards. Your friend would not understand the message until all the postcards arrived, and then only if she could figure out the right sequence.



Routers along the system have only one set of tasks: read the address, locate the receiving machine, and send the packet over the fastest route available. It does not matter to the routers what is contained in the message. It could be a video transmission from CNN or someone in the Alaskan countryside sending an e-mail to a friend in New Jersey. The routers are intentionally designed to ignore the content being delivered.

Those are the four key constraints that our design imposed: Senders are responsible for encoding a message as bits and bytes on the sending side, and receivers must decode it for their own purposes on the receiving side; digital information must be broken up into small packets; each packet has to have a clear address and each has to have a sequence number. That's it. As long as a user has a connection to the Internet and followed these rules, he can send anything he wants over the Internet.

Designing for Innovation

You will notice that these principles do not tell you anything about how to build the actual network technology. They do not define how fast the links go; whether the network uses wires, fiber optics, or radio; or whether the network hooks up 10 computers or 10,000. Any kind of communication technology that can carry bits may be put to use as part of the Internet.

The first network that carried Internet packets was the original ARPAnet. Then a wireless network and a satellite network were added. When the PC was invented, and local area networks (or LANs) were deployed, those got hooked in as well. Then the interconnected set of networks called the Internet started to grow across the country and across the globe.

The next step in the growth of the Internet occurred in the eighties, when the National Science Foundation (NSF), a government organization that funds research, developed a program to start linking all of the different networks together with one "backbone." This started out being built out of phone lines, and ended up as a massive high-speed fiber-optic cable "super highway."

The NSF encouraged all kinds of users—not just researchers—to join these networks. In time, private companies such as America Online got into the game, offering users access to the Internet for a fee. Since the early days of ARPAnet, the Internet has grown into a giant web of users some 800 million strong!

Internet users are not restricted by the limits of older technologies and older versions. New technologies like Ethernet, wireless, and fiber optics have been incorporated. New applications such as the Web, voiceover IP (voIP), and multiplayer games have been devised and deployed. The same is true for the development of streaming video and Internet telephone.

That's what we had in mind. We were never sure what would happen, but we wanted to help it along, whatever "it" might become. So we designed the Internet to promote innovation. When Tim Berner-Lees, a computer programmer, figured out how to use the Internet to host Web pages, and created the World Wide Web, he didn't come ask for our permission. We didn't have to change anything about the Internet. He just followed the conventions for data packets and addressing and then went about developing a method for presenting information on interactive "pages." His pioneering work totally transformed the way we find information.

No one has to ask to use the Internet. If you want to teach classes to students 200 miles away, or consult with a physician in another country, you just do it. So long as the messages you send are broken into digital packets, properly addressed, and sequenced, you can do just about anything on the Internet.

The Age of the Internet is far from over. In some respects, that age will never end. It will simply evolve into a new form as new innovations replace the old. As computers evolve, from PCs to laptops to PDAs and even smaller, smarter devices, the Internet will evolve to hook them all together. From the initial modest goal of transferring files from one computer to another, a new chapter in the history of our society has begun. Now an entire world of users can collaborate to build the system called the Internet, giving it a life of its own. We are all engineers when it comes to building this virtual space.

What's the Story?

1. What major historical events led to the formation of the Internet?

2. What problems were the developers of the Internet trying to solve?

3. What are the design constraints placed on Internet users?

4. Why must data packets be small and sequenced?

Designing with Math and Science

5. What's the difference between an analog signal and a digital signal? Give examples of each.

What Do You Think?

6. In your opinion, what are some of the key ways the Internet has affected our society?

7. Do you believe its a good thing that anyone can put anything (or find anything) online? What are some of the unintended consequences of that freedom?