NIGEL HOLMES

The smallest ever GUIDE TO THE INTERNET
(for busy people)

The Internet
is a bunch of networks linked together.

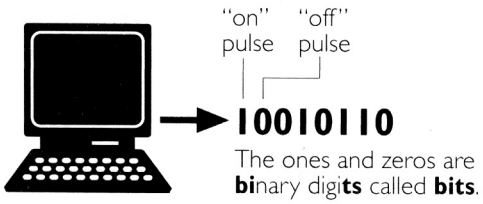
A network
is a bunch of computers linked togther.

A user
is someone who accesses data
from these computers.

The smallest ever guide to the Internet
(for busy people)

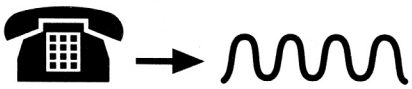
COMPUTERS, PHONES & MODEMS

Computers are **digital** machines. They produce a stream of pulses, on or off.

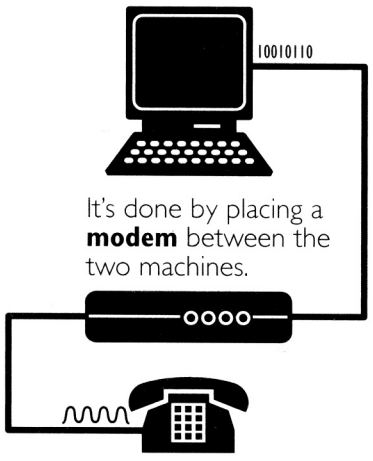


The ones and zeros are **bi**nary digi**ts** called **bits**.

To send or receive messages, computers are usually hooked up to phones. But unlike computers, phones are **analog** machines—they work by using sound waves.



For messages to be transmitted from computer to computer over phone lines, the computer's digital language must be converted to the phone's analog language.



It's done by placing a **modem** between the two machines.

The translation is called modulation (the word **modem** comes from **mo**dulation and **dem**odulation). At the other end of a message's journey, the process is reversed.

A modem's speed—the **baud rate**—is measured by the number of times a second it converts **100101** to ∿∿∿.

9.6 & 14.4 baud rates are hardly used anymore; most people connect to the web with 28.8, 33.6 or 56k modems.

Baud rate is often confused with **bits per second (bps)** which is a measure of transmission speed—the number of bits that pass a specific point per second.

The computer uses eight bits to create a single character or letter. Those eight bits = one **byte**.

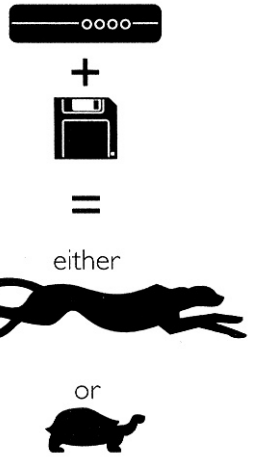
Right: 1,000 characters, about 166 words. Since each character, or letter, is one byte, this block of text is one **kilo**byte (1K*). A floppy disc holds one **mega**byte (1MB). A megabyte is a million characters, or about 166,600 words—roughly 530 pages of text.

This block of text is 1,000 characters long. That includes every letter of every word, and all the spaces and punctuation marks between the words. When calculating the length of a block of type, it's common practice to say that, on average, words and their attendant spaces and punctuation are six characters long (in the English language). So that is why a .kilobyte—1,000 characters—is calculated to be equal to roughly 166 words. (Remember, it takes one byte to create just one character). This block of text is 1,000 characters long. That includes every letter of every word, and all the spaces and punctuation marks between the words. When calculating the length of a block of type it's common practice to say that, on average, words and their attendant spaces and punctuation are six characters long (in the English language). So that is why a kilobyte—1,000 characters—is calculated to be equal to roughly 166 words. (Remember, it takes one byte to create just one character). So choose your words carefully. You are eating up valuab space with each keystroke!

*__kilo__byte=a thousand bytes / __mega__byte=a million bytes / __giga__byte=a billion bytes

Purists correct my simplification here. A kilobyte is actually 1,024 bytes, because computers use base-2 numbers, instead of the familiar base-10 system. Hey, this is the kind of stuff I'm trying to shield you from!

The time it takes to transmit messages depends firstly on the speed of the modem, and secondly on the size of the message.



But bandwidth makes a difference to transmission time too...

TRANSMISSION SPEEDS

Internet messages travel down **pipelines.** The bigger the pipeline (more **bandwidth**), the faster a message travels.

Here's **how long** it takes to transmit the contents of a **floppy disc** (one MB, or a million characters of text) using different pipelines:

Using a 28.8 phone modem: **7 minutes**

an ISDN* line (64,000-128,000 bps): **3 minutes**

DSL (1–1.5 million bps): **8 seconds**

a T-1 line (1.5 million bps): **5.2 seconds**

a cable "modem" (3–10 million bps): **1 second**

*Integrated Services Digital Network (uses standard phone lines)

At the moment, phone companies and TV cable companies are racing to upgrade their services to match increased demand for speed (and for audio and video, which require a lot more bytes).

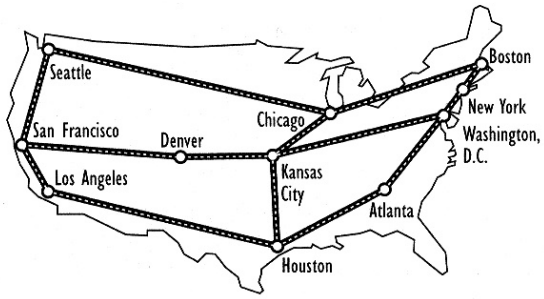
Phone companies are developing **DSL** (Digital Subscriber Line) modems. If you live more than three miles from a local switching station however, you won't be able to get this service.

The service is called **ADSL** (Asymmetrical DSL) when the speed you send (upload) data to someone is slower than the speed that you receive (download) it.

Cable companies are converting the one-way analog cables that bring you TV into high-speed two-way digital pipelines. While cable "modems" are available in parts of the US now, it will be some years before everyone can get the service.

And when we all can, the speed will depend on how many of us are logged on—at peak times cable service may be no faster than a 56k modem.

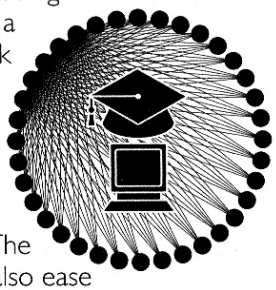
In the US, the backbones of the Internet itself (not individual connections to it) are built with fiber optic OC-3 lines that carry 155 million bits per second. Companies such as MCI Worldcom are now upgrading to lines that can handle 2.4 gigabits (2.4 billion bits) per second.



Businesses use T-1 lines and T-3 lines (45 million bps) to connect to these backbones, or build custom lines for routes they regularly use to cut down waiting time.

Next-generation Internet

The Abilene Project (also known as Internet2, or simply I2) is a cooperative effort of approximately 150 universities and companies including Cisco and IBM. They are designing a broadband network that will carry text, voice, video, and graphics fast enough to enable users to have real-time conferences. The new network will also ease the rapid increase in general net traffic.



This collaboration is a return to the origin of the Internet. In 1969, the Defense Department, in partnership with universities and other research communities, constructed the computer network ARPANET to share military and science data. In 1986, the National Science Foundation (NSF) started its own network. NSFNet took over ARPANET's functions, and in 1988 the whole system was referred to as the Internet. This is the same year that Tim Berners-Lee wrote the code for HTTP and HTML (◂ page 13), and created the World Wide Web.

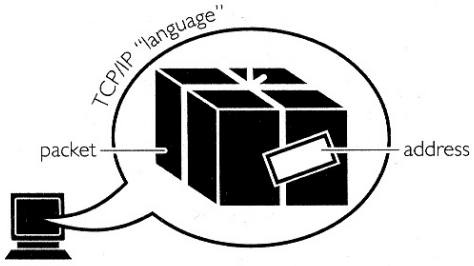
WE ALL SPEAK THE SAME LANGUAGE

A common **protocol**, or language, is used on the Internet so that people with different types of computers all over the world can correspond with each other.

Internet protocol is called **TCP/IP**, which stands for **T**ransmission **C**ontrol **P**rotocol / **I**nternet **P**rotocol.

The function of the **TCP** part of this mouthful is to break down internet messages into small **packets** of data.

The function of the **IP** part of TCP/IP is to attach an address to the packets so that the internet knows where to send them.



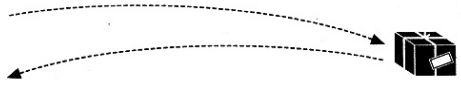
Each packet contains a mere 1,500 bytes of data (about 250 words).

Why so little data in each packet? Here's the reason:

There is **interference** on phones (it's the crackle you often hear on the line). If a computer message were sent as one continuous stream of binary digits and it encountered interference during transmission, the whole message would be destroyed and would have to be resent from the start.

But with small packets, only the interrupted ones must be resent.

Sending and resending packets between many parties, in both directions simultaneously is called **packet-switching.**



FYI, **circuit-switching** is a conversation between just two people, on one phone line.

Within the machinery of the Internet, **IP addresses** are all numerical.
The use of numbers rather than letters is less complicated for machines.

e-mail addresses substitute real names and words for numbers so we humans can remember them.

Reading from left to right, the address identifies:

1. the **user**
2. "at" (@) a **host** computer
3. . ("dot")
4. the general **domain**

The most used domains:

.com ... commercial
.edu ... educational
.gov ... government
.mil ... military
.net ... computer network entities
.org ... organizations

URL stands for Uniform Resource Locator. It is another way of saying "internet address."

A URL (address) for a site on the Wide Web looks like this:

hyper**t**ext **t**ransport **p**rotocol: the computer language used by the web

world **w**ide **w**eb

the internet domain

the name of the file

World

http://www.thelist.com/index.html

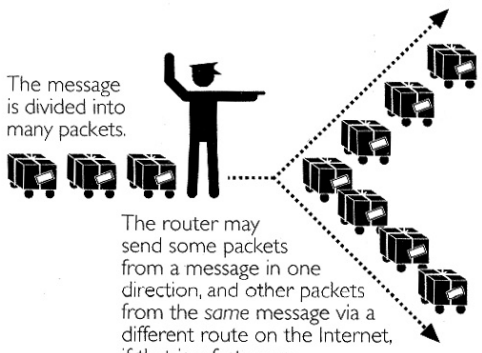
all the "dots," punctuation and slashes are needed. Following the double slashes (across to the single slash) is the name of the machine on the Internet that you will access

the name of the web site. This one has the phone numbers of all **ISP**s (Internet Service Providers) in the world

everything following a single slash is a specific file on the site. Everything to the right of this slash is case sensitive. So type carefully!

denotes the file type. **H**yper**t**ext **m**arkup **l**anguage is a format that enables the file to be read by a web browser

Routers, acting like traffic cops, control the vast numbers of packets criss-crossing cyberspace. They make sure that messages get to their destination by the fastest route.

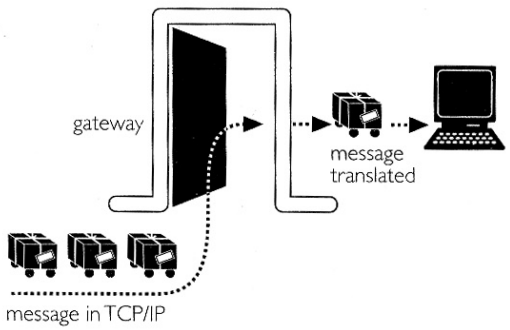


It's as though you mailed a postcard to India from the US, but the Post Office cut your card in half, then routed the two halves in different directions around the world. They also make sure that both halves arrive in India at the same time and in the right order.

The common language of the Internet, TCP/IP is used only for traffic *between* computers. It is not the native language of any of the computers themselves.

So, at the end of the message's electronic journey, it must be translated back into the language of the receiving computer.

This translation is done by a **gateway**—software residing on the Internet provider's computer.



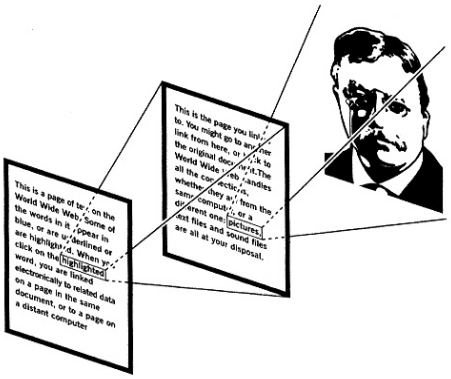
INSIDE THE WEB

To find your way around the most popular part of the Internet, the world wide web, your computer needs something like a pair of reading glasses. This is called a **browser.**

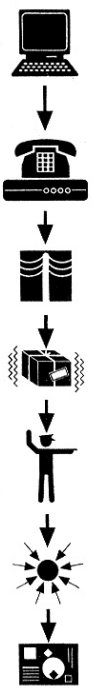
Browsers such as Netscape or Microsoft Explorer help you to move to different sites on the web, and between pages within a site, by pointing and clicking on **links** within the text or on an image.

This linkage is made possible by **hypertext.**

In the text of a file, **hypertext** links are highlighted in various ways (as blue words, for example). These are keywords that have the address of a new page embedded invisibly under them. When you click on a highlight, you are linked directly to the new page.



Hypertext links may take you to other text, to pictures and graphics, to animations, to film clips, or to audio clips.



Why is it sometimes so slow?
These factors can affect the time it takes to retrieve data from the web (old joke—the World Wide Wait):

The **speed** of your modem

The type of **phone line** between you and the host computer/**I**nternet **S**ervice **P**rovider. Also, **congestion** at your ISP

Interference on the line, which means that packets have to be resent

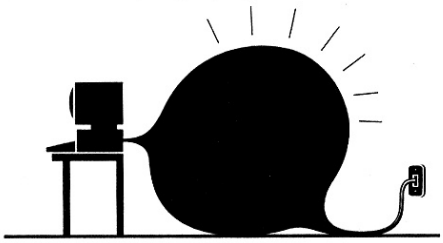
Routers, which are slower than the cables that join them

The popularity of the site you are visiting, meaning that you must wait until the **congestion** clears

Complex graphics at the site

The Internet is only as fast as the slowest link between you and the data you request.
—Stephen Manes, New York Times

Too much design
Users will quickly go elsewhere on the net if it takes them too long to **download** (see) a page.



Web designers should take into account that many people are not using the latest, fastest equipment to download pages. They should design pages that are graphically **simple** and appropriate to the restrictions of the medium. There are estimated to be 320 million pages on the Internet. You may use one of the top **portals** (such as Yahoo, or Excite) to find information, and you'll thank the designers whose work gets onto your screen the fastest.

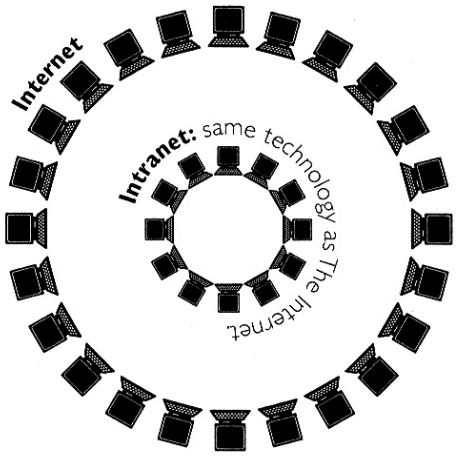
INTERNET & INTRANET

While **the Internet** is many linked **networks...**

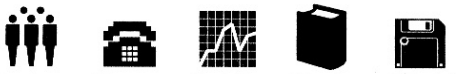
Within the networks there are **hosts.** A host is a computer that provides internet services to other computers on the network.

How many users are there? According to the US Dept. of Commerce, 80 million Americans and approximately 200 million people worldwide are connected to the Net. You'll find this and lots of other interesting numbers at www.internetindicators.com.

...an intranet is a private network within an organization—an internet inside a company.



All employees are linked together, streamlining communication. Here are some uses for intranets:



Employee training; benefits info

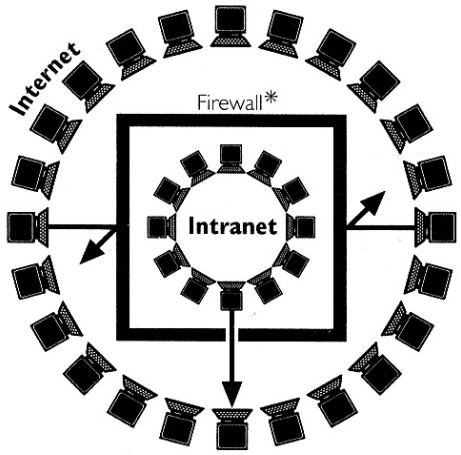
Company phone directory

Marketing materials

Libraries; research data

Technical databases; software library

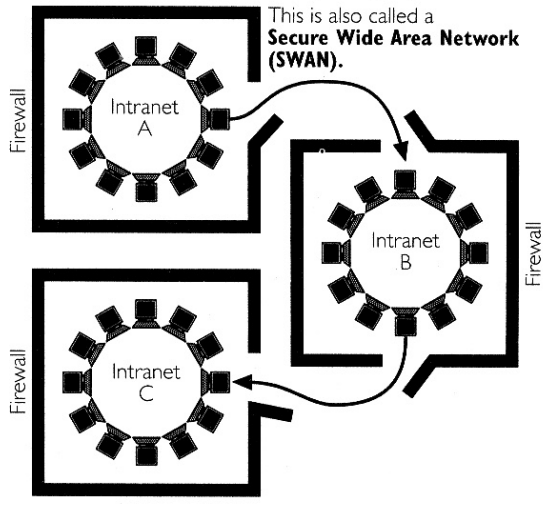
Security for intranets is achieved with a **firewall** which prevents "outsiders" from gaining access to the intranet. "Insiders," however, may cross through the firewall to retrieve data from the net proper.



If an intranet is well-engineered, users will be unaware of whether they are inside or outside the firewall.

*The firewall works by disabling part of the packet-switching activity of the net.

A **Virtual Private Network (VPN)** is an arrangement between companies to open their firewalls to each other, allowing access to their respective intranets for a specific time—for example, on a joint project.



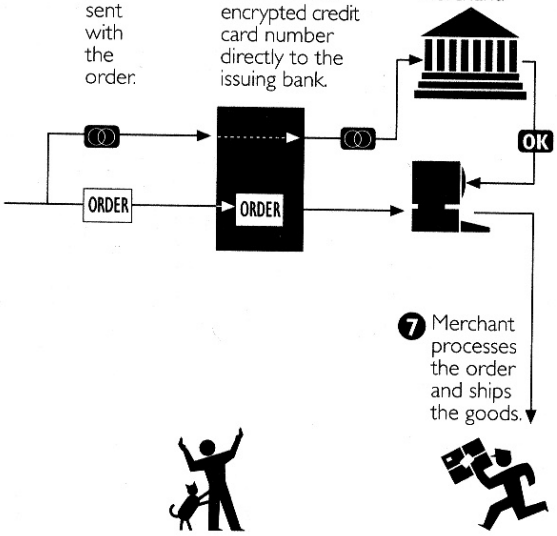
LAN = Local Area Network
(any internal network of computers in one company)
WAN = Wide Area Network
(the connection between different LANs)

COMMERCIAL TRANSACTIONS

It is estimated that worldwide Internet commerce could reach $3.2 trillion by 2003. That's 5% of total global sales. Security is a pressing concern for both buyers and sellers. Credit card companies and banks are jointly developing **Electronic Data Interchange (EDI)** technologies. Here's one way that financial transactions are made secure:

1. Buyer obtains a **digital certificate (DC)** from a bank. Digital certificates are issued to both consumers and merchants. DCs are permits that facilitate commercial transactions on the net.
2. Before an order for goods or services can go through, the buyer's software checks that the merchant has a DC, and is therefore authorized to sell over the Internet.
3. Buyer places order electronically

4. Buyer's credit card is encrypted and sent with the order.
5. Merchant's server takes the order information, but sends the encrypted credit card number directly to the issuing bank.
6. Bank decrypts the card number, authorizes payment and notifies the merchant.



7. Merchant processes the order and ships the goods.

In this kind of electronic transaction, the only party to see the buyer's credit card information is the bank which originally issued it.

COPYRIGHT PROTECTION

When the Internet was originally set up, there was no question of "ownership" of the information. Today, the web is as much a commercial medium as it is a place to exchange free information. Sites that charge users to download images, for example, naturally want to protect their property from pirates.

One way to safeguard unauthorized use of images is with a **digital watermark.** This technology encodes ownership information into the pixels of photographs. The brightness of certain pixels are changed enough to be electronically detected, but not enough to be seen.

A digital **worm** can then be sent to burrow through the Internet in search of stolen images that have this changed pattern of pixels. The rest is up to the lawyers.

But, you know, **don t steal stuff.**
Besides, lawyers have enough business.

JUST A FEW BUZZWORDS

Cookies are the coding left on your computer by websites you have visited.

Java is a programming language that can run on most computers. An **applet** (a small Java program sent to your computer over the Internet) can contain simple animation or complex math functions.

Mouse potato Someone who spends too much time at the computer.

MP3 is a format that compresses sound into files small enough for internet transmission with minimal loss of quality.

Optical computers are the next generation. They process information by sending pulses of light along optical fibers rather than bursts of electricity along metal wires.

V-commerce is e-commerce that is speech-activated.

VoiceXML is a computer language that substitutes the human voice for a mouse, and the spoken word for a click. (BUY!)

Vortal = vertical + portal. A vertical portal is one that targets a *specific* group of users. [Yahoo (p. 19) caters to *all* users, and is thus a horiz tal portal.]

XML (Extensible Markup Language) is a language like HTML (p.13) that uses **tags** (a code) to tell the computer what to display. Unlike HTML it also tells the computer what the information is. For example, in HTML, $9.99 is merely text, but in XML it can be 'tagged' as a price. This allows it to be searchable by information type.