Understanding Internet Networking and the Structure of the Web

To make the Internet work, basic networking most be used to connect computers to one another. All machines on the Internet are either servers or clients.

1. Connecting to the Internet

Every computer that is connected to the Internet is part of a network, even the one in your home.

For example, you may use a modem and dial a local number to connect to an **Internet Service Provider** (ISP). *Internet Service Provide* is a business that has permanent Internet connections and provides temporary connections to individuals and companies for free or a fee.

At work, you may be part of a **local area network** (LAN), but you most likely still connect to the Internet using an ISP that your company has contracted with.

When you connect to your ISP, you become part of their network. The ISP may then connect to a larger network and become part of their network.
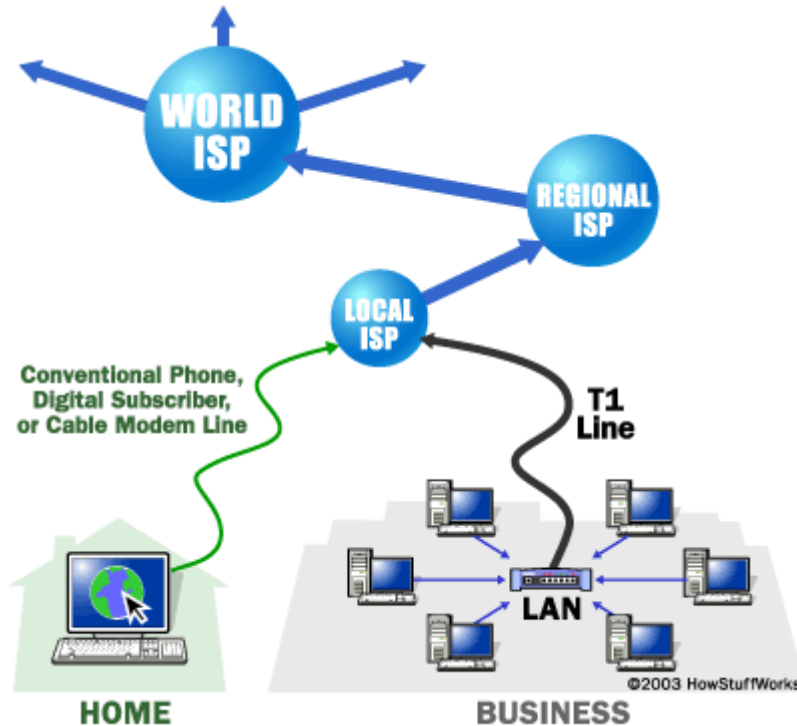
The Internet is simply a network of networks.

2. Router

We're all used to seeing the various parts of the Internet that come into our homes and offices -- the Web pages, e-mail messages and downloaded files that make the Internet a dynamic and valuable medium. But none of these parts would ever make it to your computer without a piece of the Internet that you've probably never seen the technology most responsible for allowing the Internet to exist at all: the **router**.
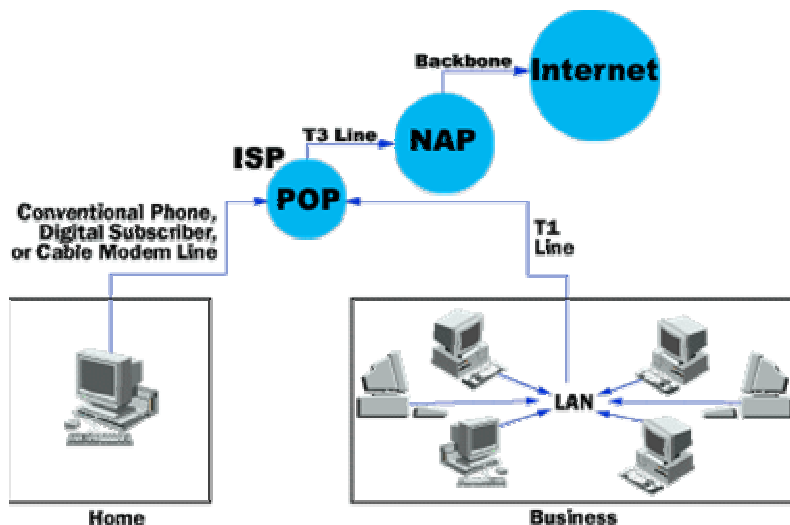
Bridging The Divide

All of these networks rely on NAPs, backbones and **routers** to talk to each other. What is incredible about this process is that a message can leave one computer and travel halfway across the world through several different networks and arrive at another computer in a fraction of a second!

POP (Points of Presence) is a place for local users to access the company's network, often through local phone number or dedicated line.

NAP(Network Access Points) are high-level networks connecting to each other.

Backbone: high –speed communication– capable of moving large amounts of data

The routers determine where to send information from one computer to another. *Routers are specialized computers that send your messages and those of every other Internet user speeding to their destinations along thousands of pathways.*

A router has two separate, but related, jobs:

1. *It ensures that information doesn't go where it's not needed. This is crucial for keeping large volumes of data from clogging the connections of "innocent bystanders."*
2. *It makes sure that information does make it to the intended destination.*

In performing these two jobs, a router is extremely useful in dealing with two separate computer networks.

1. It joins the two networks, passing information from one to the other.

2. It also protects the networks from one another, preventing the traffic on one from unnecessarily spilling over to the other.

Regardless of how many networks are attached, the basic operation and function of the router remains the same. Since the Internet is one huge network made up of tens of thousands of smaller networks, its use of routers is an absolute necessity.

3. Backbones

The **National Science Foundation** (NSF) created the first high-speed backbone in 1987. Called **NSFNET**, it was a T1 line that connected 170 smaller networks together and operated at 1.544 Mbps

(million <u>bits</u> per second). IBM, MCI and Merit worked with NSF to create the backbone and developed a T3 (45 Mbps) backbone the following year.

Backbones are typically fiber optic trunk lines. The trunk line has multiple fiber optic cables combined together to increase the capacity. Fiber optic cables are designated OC for optical carrier, such as OC-3, OC-12 or OC-48.

Speeds of Transmission:

         1. OC-3 line is capable of transmitting 155 Mbps

         2. OC-48 can transmit 2,488 Mbps (2.488 Gbps).

         3. 56K modem transmitting 56,000 bps and you see just how fast a modern backbone is.

Today there are many companies that operate their own high-capacity backbones, and all of them interconnect at various NAPs around the world. In this way, everyone on the Internet, no matter where they are and what company they use, is able to talk to everyone else on the planet. The entire Internet is a gigantic, sprawling agreement between companies to intercommunicate freely.

4. Internet Protocol: IP Addresses- Created by Vincent Cerf

Reminder: Protocols
In networking, the term **protocol** refers to a set of rules that govern communications. Protocols are to computers what language is to humans. Since this article is in English, to understand it you must be able to read English. Similarly, for two devices on a network to successfully communicate, they must both understand the same protocols.

Every machine on the Internet also has a unique identifying number, called an **IP Address**.

The IP stands for **Internet Protocol**, which is the language that computers use to communicate over the Internet. A protocol is the pre-defined way that someone who wants to use a service talks with that service. The "someone" could be a person, but more often it is a computer program like a Web browser.

         A typical IP address looks like this:

216.27.61.137

To make it easier for us humans to remember, IP addresses are normally expressed in decimal format as a *dotted decimal number* like the one above. But computers communicate in binary form. Look at the same IP address in binary:

11011000.00011011.00111101.10001001

A **<u>server has a static IP</u>** address that does not change very often.

A **<u>home machine</u>** that is dialing up through a modem, on the other hand, typically has an IP address assigned by the ISP every time you dial in. **<u>That IP address is unique for your session</u>** -- it may be different the next time you dial in. This way, an ISP only needs one IP address for each modem it supports, rather than one for each customer.

5. Domain Name System Internet Protocol: Domain Name System

When the Internet was in its infancy, it consisted of a small number of computers hooked together with modems and telephone lines. You could only make connections by providing the IP address

of the computer you wanted to establish a link with. This was fine when there were only a few hosts out there, but it became unwieldy as more and more systems came online.

The first solution to the problem was a simple text file maintained by the Network Information Center that mapped names to IP addresses. Soon this text file became so large it was too cumbersome to manage. In 1983, the University of Wisconsin created the **Domain Name System** (DNS), which maps text names to IP addresses automatically. This way you only need to remember www.howstuffworks.com, for example, instead of HowStuffWorks.com's IP address.

6. Uniform Resource Locators

When you use the Web or send an e-mail message, you use a domain name to do it. For example, the **Uniform Resource Locator** (URL) "http://www.howstuffworks.com" contains the domain name howstuffworks.com. So does this e-mail address: example@howstuffworks.com.

Every time you use a domain name, you use the Internet's DNS servers to translate the human-readable domain name into the machine-readable IP address.

Top-level domain names, also called first-level domain names, include .COM, .ORG, .NET, .EDU and .GOV.

Within every top-level domain there is a huge list of second-level domains. For example, in the .COM first-level domain there is:

- HowStuffWorks
- Yahoo
- Microsoft

Every name in the .COM top-level domain must be unique. The left-most word, like www, is the host name. It specifies the name of a specific machine (with a specific IP address) in a domain. A given domain can, potentially, contain millions of host names as long as they are all unique within that domain.

Even though it is totally invisible, DNS servers handle billions of requests every day and they are essential to the Internet's smooth functioning. The fact that this distributed database works so well and so invisibly day in and day out is a testimony to the design.

7. Web Browser

A **web browser** is a software application.

1. Eables a user to display and interact with text, images, and other information typically located on a web page at a website on the World Wide Web or a local area network.
2. Text and images on a web page can contain hyperlinks to other web pages at the same or different website.
3. Web browsers allow a user to quickly and easily access information provided on many web pages at many websites by traversing these links.
4. Web browsers format HTML information for display, so the appearance of a web page may differ between browsers.

Web browsers available for personal computers include Internet Explorer, Mozilla Firefox, Safari, Netscape, and Opera in order of descending popularity (as of August 2006).

Web browsers are the most commonly used type of HTTP user agent. Although browsers are typically used to access the World Wide Web, they can also be used to access information provided by web servers in private networks or content in file systems.

Protocols and standards

Web browsers communicate with web servers primarily using HTTP (hypertext transfer protocol) to fetch webpages.

HTTP allows web browsers to submit information to web servers as well as fetch web pages from them.

Pages are located by means of a URL (uniform resource locator), which is treated as an address, beginning with *http:* for HTTP access.

Many browsers also support a variety of other URL types and their corresponding protocols, such as *ftp:* for FTP (file transfer protocol), *rtsp:* for RTSP (real-time streaming protocol), and *https:* for HTTPS (an SSL encrypted version of HTTP).

The file format for a web page is usually HTML (hyper-text markup language) JPEG, PNG and GIF image formats, and can be extended to support more through the use of plugins. T

The combination of HTTP *content type* and URL protocol specification allows web page designers to embed images, animations, video, sound, and streaming media into a web page, or to make them accessible through the web page.

Early web browsers supported only a very simple version of HTML. The rapid development of proprietary web browsers led to the development of non-standard dialects of HTML, leading to problems with Web interoperability. Modern web browsers support standards-based HTML and XHTML, which should display in the same way across all browsers. Internet Explorer does not fully support HTML 4.01 and XHTML 1.x yet. Currently many sites are designed using WYSIWYG HTML generation programs such as Macromedia Dreamweaver or Microsoft Frontpage. These often generate non-standard HTML by default, hindering the work of the W3C in developing standards, specifically with XHTML and CSS (cascading style sheets, used for page layout).

Some of the more popular browsers include additional components to support Usenet news, IRC (Internet relay chat), and e-mail. Protocols supported may include NNTP (network news transfer protocol), SMTP (simple mail transfer protocol), IMAP (Internet message access protocol), and POP (post office protocol). These browsers are often referred to as *Internet suites* or *application suites* rather than merely web browsers.

Brief history

Tim Berners-Lee (who pioneered the use of hypertext for sharing information) as the world's first web server, and also to write the first web browser, WorldWideWeb in 1990. Berners-Lee introduced it to colleagues at CERN in March 1991. Since then the development of web browsers has been inseparably intertwined with the development of the web itself.

Marc Andreessen, who was the leader of the Mosaic team at NCSA, quit to form a company that would later be known as Netscape Communications Corporation.

Netscape released its flagship Navigator product in October 1994, and it took off the next year. Microsoft, which had thus far not marketed a browser, now entered the fray with its Internet Explorer product, purchased from Spyglass Inc. This began what is known as the **browser wars**, the fight for the web browser market between Microsoft and Netscape.

The wars put the web in the hands of millions of ordinary PC users, but showed how commercialization of the web could stymie standards efforts. Both Microsoft and Netscape liberally incorporated proprietary extensions to HTML in their products, and tried to gain an edge

by product differentiation. Starting with the acceptance of the Microsoft proposed Cascading Style Sheets over Netscape's JavaScript Style Sheets (JSSS) by W3C, the Netscape browser started being generally considered inferior to Microsoft's browser version after version, from feature considerations to application robustness to standard compliance. The wars effectively ended in 1998 when it became clear that Netscape's declining market share trend was irreversible. This trend may have been due in part to Microsoft's integrating its browser with its operating system and bundling deals with OEMs; Microsoft faced antitrust litigation on these charges.

Netscape responded by open sourcing its product, creating Mozilla. This did nothing to slow Netscape's declining market share. The company was purchased by America Online in late 1998. At first, the Mozilla project struggled to attract developers, but by 2002 it had evolved into a relatively stable and powerful internet suite. Mozilla 1.0 was released to mark this milestone. Also in 2002, a spin off project that would eventually become the popular Mozilla Firefox was released. In 2004, Firefox 1.0 was released; Firefox 1.5 was released in November 2005. Firefox 2, a major update, was released in October 2006 and work has already begun on Firefox 3 which is scheduled for release in 2007. As of 2006, Mozilla and its derivatives account for approximately 12% of web traffic.

Opera, an innovative, speedy browser popular in handheld devices, particularly mobile phones, as well as on PCs in some countries was released in 1996 and remains a niche player in the PC web browser market. It is available on Nintendo's DS, DS Lite and Wii consoles[2]. The Opera Mini browser uses the Presto (layout engine) like all versions of Opera, but runs on most phones supporting Java Midlets.

The Lynx browser remains popular for Unix shell users and with vision impaired users due to its entirely text-based nature. There are also several text-mode browsers with advanced features, such as w3m, Links (which can operate both in text and graphical mode), and the Links forks such as ELinks.

The Macintosh scene too has traditionally been dominated by Internet Explorer and Netscape. However, Apple's Safari raised in popularity. Safari is the default browser on Mac OS X.

In 2003, Microsoft announced that Internet Explorer would no longer be made available as a separate product but would be part of the evolution of its Windows platform, and that no more releases for the Macintosh would be made. However, more recently in early 2005, Microsoft changed its plans and released version 7 of Internet Explorer for Windows in October 2006.

Features

Different browsers can be distinguished from each other by the features they support. Modern browsers and web pages tend to utilize many features and techniques that did not exist in the early days of the web. As noted earlier, with the browser wars there was a rapid and chaotic expansion of browser and World Wide Web feature sets.

The following is a list of some of the most notable features:

Standards support

- HTTP and HTTPS
- HTML, XML and XHTML
- Graphics file formats including GIF, PNG, JPEG, and SVG
- Cascading Style Sheets (CSS)
- JavaScript (Dynamic HTML) and XMLHttpRequest
- Cookie
- Digital certificates
- Favicons

- RSS, Atom

Fundamental features

- Bookmark manager
- Caching of web contents
- Support of media types via plugins such as Macromedia Flash and QuickTime

Usability and accessibility features

- Autocompletion of URLs and form data
- Tabbed browsing

Annoyance removers

- Pop-up advertisement blocker
- Advert filtering
- Phishing defenses

Retrieved from "http://en.wikipedia.org/wiki/Web_browser"