Internet Protocol (IP)

The internet protocol, or *IP* for short, is the underlying system of communication for all data sent across the internet. It is a simple protocol for one computer to send packets of data to another.

IP address
A 32-bit number identifying a computer on a network using the IP protocol.

A major aspect of any protocol is the issue of addressing. That is, how do I specify to whom I want to send my message? IP deals with addressing by giving every computer on the internet a unique 32-bit ID number called an *IP address*. To a computer, an IP address is just a large integer, but when written out for people to read, an IP address is often broken into four smaller integers between 0 and 255, such as 145.10.34.3. The 32 bits of the address are divided into 4 groups of 8 bits; 8 bits of binary data can represent $2^8 = 256$ different unique integers, so each of the four parts of an IP address is a number between 0 and 255.

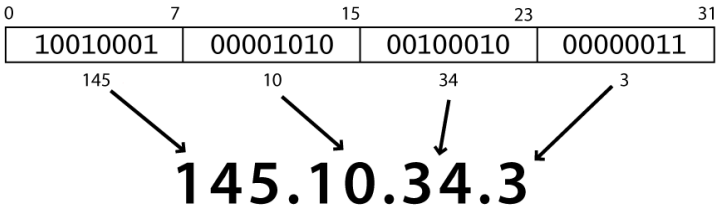

Figure 1.1 IP address

Every web server you visit has an IP address, and your own computer has one, too. You can find out your IP address by visiting a web site such as `whatismyip.com`, or by running a program on your computer such as `ipconfig` (Windows) or `ifconfig` (Mac/Linux). Most computers can connect to themselves using the special IP address 127.0.0.1 or by using the name `localhost`.

Two computers communicating with each other via the internet are almost never directly physically connected to each other. Instead, the computers are linked by a long chain of intermediate machines called routers. A *router* is a device that receives IP packets and hands them off to another computer on the network. Routers have several input and output connections and use tables of information to know where to send each packet of IP data by looking at the packet's destination IP address.

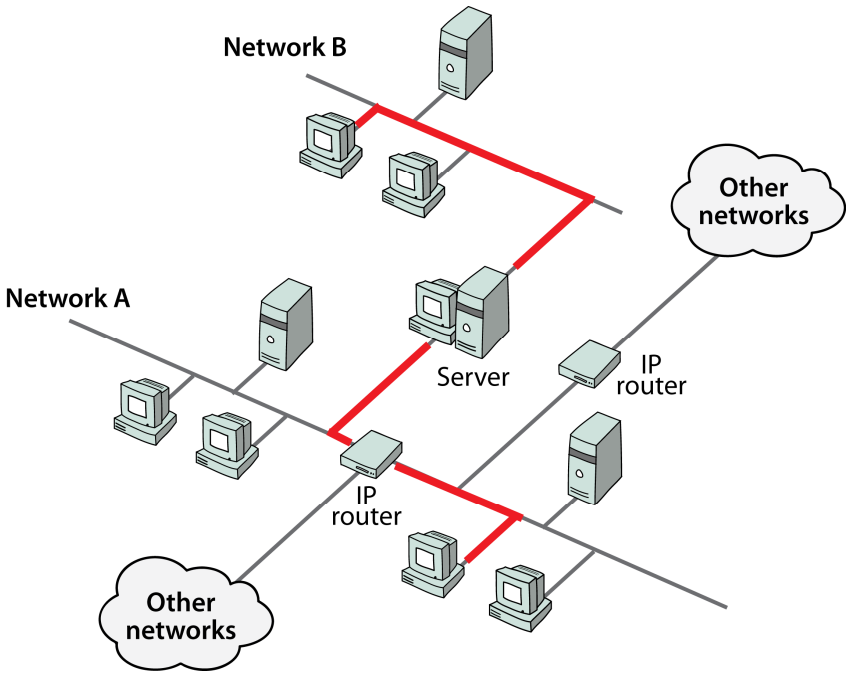

Figure 1.2 Routers

IP is a very minimal protocol; addressing and end-to-end routing are essentially all it provides. IP does not provide any functionality to prevent against problems such as:

- **Data loss**: Information might be lost in transit.
- **Corruption**: Information might arrive in a corrupted or altered state.
- **Duplication**: A single message might arrive twice.
- **Sequencing**: Information could arrive out of order. (Message A might be sent before B, but B might arrive at the destination before A.)

These issues are handled by protocols that work on top of IP.

Transmission Control Protocol (TCP)

TCP

A protocol that provides reliable, in-order delivery of information between computers.

The Transmission Control Protocol (or *TCP* for short) is another communication standard implemented on top of IP. When one computer sends a packet to another using IP, once the packet arrives, the set of software rules that interpret and handle its data are called TCP.

TCP exists to solve the limitations of IP listed previously. TCP guarantees reliable, non-corrupt, in-order delivery from one computer to another. It does this by adding extra information to IP data such as validity-checking information and unique numbers for each packet of data, so that they can be placed in order when they arrive at their destination. Computers communicating with TCP send acknowledgements when data has arrived and re-send data when no acknowledgement comes back, to make sure every packet arrives at the destination.

TCP also adds the ability for multiple programs and services to share the same physical computer and internet connection. It does this by associating each program/service with a unique integer called a *port*. You may have seen TCP ports while using the web or other internet-enabled programs. Certain common internet services have been given standard ports, such as:

TCP port	service
21	file transfer (FTP)
22	secure shell (SSH)
25, 110	email (SMTP, POP3)
80	web (HTTP)
443	secure web (HTTPS)
993, 995	secure email

Table 1.1 Common ports and services

TCP is an important layer of the internet because it frees the author of an internet application from worrying about the reliability of the underlying network. When information is sent with TCP, you have reasonable assurance that it will arrive at its destination successfully. Since TCP is always run on top of the IP protocol, the combination of the two is frequently referred to as *TCP/IP*.

TCP is not the only protocol that runs on top of IP; others such as the UDP protocol are better suited to other types of applications such as streaming media and online games.

Self-Check

1. What is the difference between the internet and the World Wide Web (WWW)?
2. What are some of the most important aspects or features of the internet?
3. What organization decides on standards for web pages and languages?
4. What features are provided by the Internet Protocol (IP), and what additional features are added by TCP?
5. Which of the following are legal IP addresses?
    a) www.google.com
    b) 150.135.1.150
    c) 123.456.789.10
    d) 241.259.17.127
    e) 10.0.0.1
6. What is your computer's public IP address? What are the IP addresses of the following sites?
    a) www.google.com
    b) www.facebook.com
    c) www.microsoft.com
    d) www.mozilla.org