Beam Me Up, Scotty: How the Internet Transports Data

If you have e-mailed, chatted, searched for information or transmitted anything across the interent, you have experienced first hand the powerful connection between people that results from using the Internet.

Just as its predecessor ARPANET was originally conceived as a network of people using computers (rather than as a computer network), the Internet promotes the development of human communities sharing information of interest. And that now extends into the economic life of the country – individuals and businesses communicating with each other to transact business.

What is the Internet?

In simple terms, the Internet is a collection of thousands of computer networks, each established, maintained and supported by individual private organizations, universities and government agencies, and all cooperating to share information by using standard procedures and protocols for communicating. Even though its use has exploded, it has remained true to its original intent – to enable communication between people located anywhere in the world.

Transporting data over the Internet

Before the Internet existed, scientists thought of computers mainly as machines to perform arithmetic functions. ARPA (Advanced Research Projects Agency), by contrast, developed a new theme — that the promise offered by computers is the potential to allow people to communicate. In fact, ARPA asserted that the communication potential "dwarfs into relative insignificance the historical beginnings of the computer as an arithmetic engine." (ARPA Completion Report).

Prior to development of the Internet, the telephone system was a familiar and widely used way for people to communicate. The telephone system is an *analog* network that uses *circuit switching* to transport information. *Analog information* is represented as an electronic signal or tone whose frequency is continuously varied – *i.e.,* modulated – in frequency, amplitude or phase. To maintain data integrity, analog information must be kept intact. Thus, to complete a phone call, it is necessary to establish a pathway for the modulated frequency (your voice) and to dedicate the pathway until you have completed your call and broken the connection (by hanging up).

In contrast, the Internet is a *digital* network that uses *packet switching* to transport information. *Digital information* is represented by a *binary code* – a 0 or 1 – that can be broken down into small blocks of data (called *packets*). Each packet can be sent independently of one another; then at the destination, recombined into their original, unified form.

To picture how this works, think of Captain James T. Kirk of the Starship Enterprise on the surface of an alien planet. He issues the command, "Beam me up, Scotty." Kirk dissolves – becomes a series of independent packets of digital data – and immediately reappears and recombines aboard the Enterprise. He has been transported using a digital network and packet switching.

Handling information packets: TCP/IP

Without some rules (called *protocols*) for breaking down, transporting and reassembling information packets, Captain Kirk might end up looking very strange indeed. On the Internet, the two most important communication protocols are the *transmission control protocol (TCP)* and *Internet protocol (IP)*. TCP breaks down and reassembles the information packets; the IP handles the routing to makes sure that all packets are sent to the right destination.

For a number of reasons (including hardware limitations), data sent across the Internet is broken into packets of less than about 1500 characters each. The TCP gives each packet a *header* that contains information such as the packet's relationship to other packets (so that the packets can be accurately assembled), and a *checksum* based on the precise amount of data in the packet. (The checksum is used on the receiving end to determine whether any errors have been introduced into the packet during transmission.) Finally, the TCP puts each packet into a separate IP envelope that contains addressing information (sender and destination address, how long to keep the packet, etc.)

As the packets make their way across the net, *routers* examine the IP envelopes, find the destination address, and determine the most efficient path for sending each packet to the next router closest to the final destination. Packets typically travel through a series of routers – the exact number and routes are a function of the Internet traffic at any given moment. Since the traffic load changes constantly, packets may arrive at the final destination out of order.

As the packets arrive at their destination, TCP calculates a checksum for each packet and compares it to the one sent in the packet. If the checksums don't match, TCP knows the data has been corrupted. It then discards the packets and requests a retransmission. When all packets have been received in perfect form, TCP assembles them into their original form.

Internet addresses and domains

Every IP address on the Internet is a series of four numbers separated by periods (called *dots*). Because remembering these numeric addresses is so difficult, the *Domain Name System (DNS)* was developed. A *domain* is a group of computers on the Internet; each computer in the domain is given a separate *domain name* (also known as an *internet address*) consisting of easily recognizable letters and words. The domains at the top of the hierarchy maintain lists and addresses of the domains just beneath them; each subsequent domain does the same, allowing every computer on the Internet to have a domain name.

Internet file types

In general the Internet contains three kinds of files: ASCII (American Standard Code for Information Exchange); EBCDIC (Extended Binary Coded Decimal Interchange Code); and binary. Binary files contain special coded data that can only be read or run by specific computers and software. EBCDIC files organize data into something understandable by people and are primarily used by mainframe computers.

ASCII files are made up of computer codes that translate into characters on your computer screen (such as an upper case *L* or an ampersand (&)). ASCII text files contain characters only, with no formatting (such as bold face, italics or paragraph indentations).

Binary files (such as *Portable Document Format*) and formatted ASCII files (such as *Postscript*) contain complex information about documents – placement of pictures, size and type of fonts, formatting information – and need special software to open and view. For PDF you need an Adobe Acrobat reader; for Postscript, a Postscript reader. With this special software you will be able to view the fully formatted screen on your screen.

Using the Internet to transfer files

There are two principal ways to transfer files across the Internet: *File Transfer Protocol (FTP)* and as an attachment to e-mail.

FTP, which was articulated in the early days of Internet development, describes the ways in which the client and server computers interact to successfully effect a file transfer. Since this protocol was specifically created to enable accurate data transfer, it is the first choice for sending files.

Files can also be transferred as attachments to e-mail. Since the Internet isn't able to handle binary files in e-mail, the file must first be encoded. (Two popular coding schemes are *Multipurpose Internet Mail Extensions* or *MIME* and *Uuencode*.) When the encoded attachment reaches its destination, it must be decoded with the same scheme used for the encoding.

Unlike FTP, which is used uniformly throughout the Internet, different networks on the Internet may use different e-mail formats. *Gateways* translate e-mail formats from one network to another as e-mail makes its way to its destination. E-mail may also have to pass through a *firewall* (a computer that shields the network from the broader internet).

Sending us files

File transfer from your office to ours is a great benefit of the Internet. If you have yet to use this convenient method of file transport, we encourage you to experiment with a test file. Call us and we'll answer any questions you may have.

For your convenience, our web site www.expressprintingusa.com incorporates FTP. To send us a file, simply go to the site and follow the instructions for uploading a file. We think you'll agree it is easy and convenient.