



**Computer  
Science**

**COMPSCI 111 SC - Lecture 26**

October 2003

## New connections - New Media

*(An introduction to data communications, networks and the Internet)*

### Introduction

Data Communications began in 1940, when Dr George Stibbitz sent data through the telephone system from Dartmouth College to New York. The value of transmitting data from one machine to another was quickly recognised. Data communications became a popular area of research among the military, business and academic communities. The importance of data communications continues to attract funding and research to this day.

### Data Transmission mediums

The path through which data is sent is known as a channel. There are many different mediums which form data channels. The most common is plain wire cables (twisted pair or coaxial) which use electrical signals to carry data. Fibre optic cables consist of very thin glass filaments which carry digital signals generated by switching lasers on and off to form pulses of light. These fibre optic cables have many advantages over wire, the most important of which is speed of transmission. A single fibre optic cable can carry up to 7000 times as much data as regular wire cables. They are also less prone to interference (up to a million times lower rate of errors), safer, last longer and are much more difficult to wiretap (giving them a higher degree of security). Microwaves are also used to transmit data via relay stations. These stations must be in line-of-sight of each other, and can transmit very large quantities of data very rapidly. Both cellular phones and communication satellites use this technology which offers a cable free medium of communication.

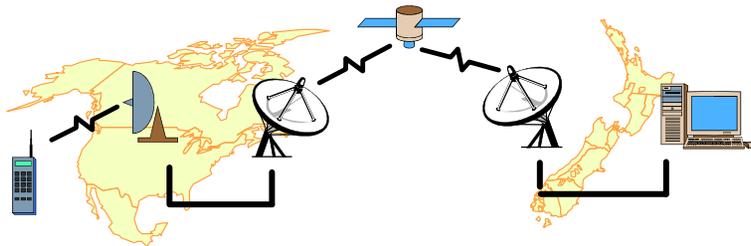


Figure 1: Data Transmission usually occurs through a variety of different communication mediums.

### Transfer rates.

The greatest concern facing specialists in data communications is the speed by which the data is transferred. As technology increases, researchers are discovering new techniques which can be used to transmit data faster than ever before. However, our expectations continue to rise, requiring an ever faster rate of data transfer (e.g. people were once happy to use Morse code to transfer messages via telegraph, now people want video conferencing). The transmission rate of data is often referred to as the baud rate (which is incorrect, but the baud rate and transmission rate are often approximately equal), and is measured in bits per second (bps). The transmission rate is dependant upon the range of frequencies which can be transmitted by a given medium (i.e. the bandwidth of the medium), and the speed at which the signal travels through the medium.

### Technical details

The transmission of data can be realised in a variety of ways. The best method of transmission is highly dependant upon the nature of the data to be transferred. Simplex channels allow data to be transmitted in one direction only (e.g. TV, Radio). Half-duplex channels allow 2-way communication, but only one direction at a time (e.g. walkie-talkies - one end must stop transmitting before the other one can begin). Full duplex channels allow data to be transmitted in both directions at the same time (e.g. telephone line - both people can speak and listen at the same time).

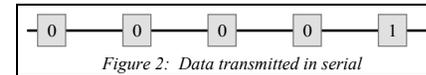


Figure 2: Data transmitted in serial

Data can be transmitted in a parallel or serial fashion depending upon hardware constraints. Inside the computer, information is usually transmitted in parallel one word at a time (a word being the number of bytes which the computer treats as a single unit). However, most data channels require data to be transmitted in serial, resulting in a much slower rate of transfer.

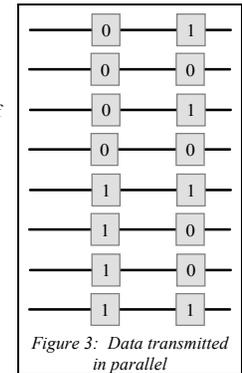


Figure 3: Data transmitted in parallel

Computers need to use accurate methods of data transfer since even a single error could render the data useless. Two common methods are called asynchronous and synchronous. Asynchronous transmission involves sending one byte of data at a time. In order to separate the bytes of data, start and stop bits are used to signal that a byte of data is being sent. Usually, 2 bits are sent as a signal that a byte is following, then the byte is sent, then 2 bits are sent to indicate that the byte has been transmitted.

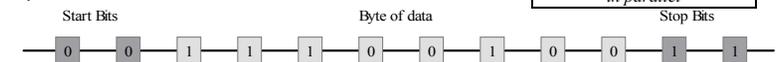


Figure 4: Transmitting the byte 11100100 asynchronously requires sending 2 start bits (00), followed by the byte (11100100), followed by the stop bits (11).

Synchronous transmission involves sending many bytes gathered together in a group known as a packet. These bytes are sent without start or stop bits. Instead, the transmission is carefully timed, so that the receiver can distinguish each individual byte.

### Protocols

Data communications is a complex issue, involving many possible ways of sending and receiving data. It is important that both ends of a transmission use the same methods. Computer professionals have realised the importance of standardising on particular communications techniques. These standard methods of communicating are known as protocols and they define a set of rules and procedures indicating how to initiate and maintain communication. Low level protocols define the electrical and physical standards to be observed, the bit-ordering and the transmission, error detection and correction of the bit stream. High level protocols deal with the data formatting, including the syntax of messages, character sets, sequencing of messages etc.

### Communications Hardware

A typical data communications system will involve many terminals connected to a central server. A network which exists over a large area would typically use modems to create the connections between the terminal and the server. These common elements of a network warrant further discussion.

Any input/output device which is at the end (or terminal point) of a data communications system is known as a terminal. However, terminal is usually used to refer to a computer which is connected to the communications system and includes a screen and keyboard for input and output. Today, many computers have "terminal" software which allow them to connect to other host computers.

A modem is a device which enables a computer to receive and transmit data over a telephone line. Traditionally, telephone lines have been analogue lines (i.e. transmits signals which have a continuous range of values), whereas computers use digital signals (discrete values - either 0 or 1 for computers). A modem converts the computer's digital signal into analogue form which can be sent through a regular telephone. A modem at the other end of the telephone converts the signal from analogue back into digital so that the remote computer can understand the signals.

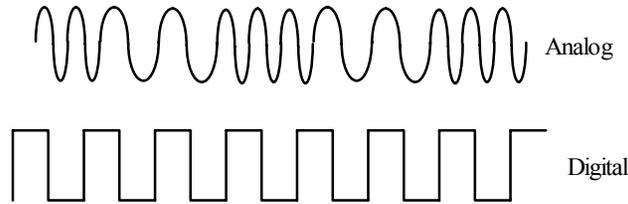


Figure 5: Analog signals have a continuous range of values, while digital signals have a limited range of discrete values.

**Server**

A server is a computer which understands a particular protocol. It receives and transmits data according to the protocol which it understands. A typical data communications system may have many different servers each of which is dedicated to a particular sort of task (and can understand the corresponding protocol).

**Traditional Networks**

Traditionally, early networks used a single computer to co-ordinate the transfer of information from one machine to another. This central computer became known as the host computer. The host computer was usually left switched on all the time, allowing other computers to send data at any time.

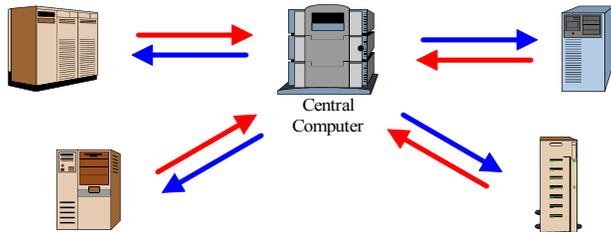


Figure 6: A central computer was used to co-ordinate communications between different machines

**Local Area Networks (LAN)**

Most small businesses today have at least one computer. Those who have more than a single computer are likely to have them connected together in order to share files and other resources such as printers and modems. A local area network (LAN) consists of two or more computers physically connected together (usually with twisted pair or coaxial cable) which are located in a limited geographic area (less than 1 km across).

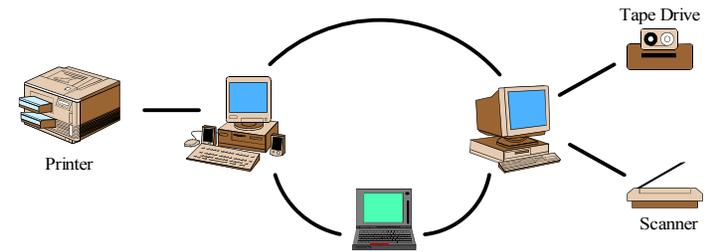


Figure 7: A Local Area Network allows users to transfer files between one another and to share expensive resources (such as printer, scanner or tape drive).

**Wide Area Networks (WAN)**

A network which spans a geographic area of greater than 1 km is usually known as a wide area network (WAN). Networks of this type often include satellite dishes, or operate through fast digital lines leased from a telecommunications company such as Telecom. The US National Science Foundation (NSF) backbone is one such example of a WAN.



Figure 8: A Wide Area Network spans a distance of greater than 1 kilometre.

**An internet**

When two or more networks are connected together, the resulting network is known as an internet. This term should not be confused with the global network of networks commonly known as the Internet (note the difference in capitalisation).



Figure 9: The connection of two or more networks forms an internet.

**The Beginning of the Internet**

During the 1950's, America was in the grip of "Cold War" paranoia. Fear of communism, and the perceived threat of a Nuclear Strike from the Soviet Union caused a intense period of industrial research. The successful launch of Sputnik by the Soviet Union in 1957 provoked the formation of the Advanced Research Projects Agency (ARPA) within the Department of Defence (DoD).

**A New Type of Network:**

Dr. J. Licklider, who was head of ARPA in 1962 began to involve universities in the ongoing research, thus tapping the resources of the academic community. After the Cuban missile crisis in 1962, a nuclear attack seemed unavoidable. Investigation into the aftermath of a Nuclear Attack suggested that neither the long-distance telephone plant nor the military command and control network would survive such an attack. Even though most of the links would be operable, the centralised switching facilities would be destroyed, rendering the system useless. During this time, Paul Baran of RAND (a Research and Development organisation) came up with a promising solution.

Baran conceived of a network where each node would have equal status; be autonomous; and be capable of receiving, routing, and transmitting information. This system would take each message and break it down into pieces (called packets). Each of these packets would be individually addressed and allowed to travel through the network by whatever route was available. If parts of the network were destroyed, the address of the destination was sufficient for a node to send the packet via a different route.

In autumn of 1969, ARPA connected four sites together using methods developed by RAND. These sites were:

1. Stanford Research Institute (SRI)
2. University of California at Los Angeles (UCLA)
3. University of California at Santa Barbara (UCSB)
4. University of Utah

## The Developing Network

The network was initially to be used primarily for long-distance computing. This would allow a user in a remote location to log into and use a distant machine as if they were actually there. Researchers could make use of the few supercomputers in the country without having to travel miles to try out their programs. Within 2 years however, the main use of the network was no longer long-distance computing, but rather sending personal messages. Each user had their own account, and could be sent an individual message, which proved to be extremely popular among the growing number of users.

The decentralised design of ARPAnet allowed computers to be connected far more easily than in traditional networks. It was not necessary to register with a central authority, but merely to connect to the closest machine already connected to the network. This design allowed the network to expand quickly and easily.

## TCP/IP

Once the network began to expand, it started to be connected to other existing networks. The original protocol for ARPAnet was designed for only a small number of machines. Robert Kahn saw a need for a protocol which would allow for unlimited expansion. In 1972 he began to work on TCP/IP (Transmission Control Protocol/ Internet Protocol). He assumed:

- No internal changes would have to be made to a network in order to connect it to the Internet.
- Communication would be on a best effort basis, with any lost packets being retransmitted
- There would be no global control of the network.

At a similar time in PARC the Ethernet protocol was developed by Bob Metcalfe for use with local area networks. After the public release of Ethernet in 1973 it quickly became the standard protocol for LAN communications and small independent networks flourished.

## From ARPAnet to Internet

The term "Internet" was first used in a planning document in 1974, but ARPAnet was still commonly referred to until the early 1980's. Around this time, TCP/IP was integrated into the UNIX system, so that standard mainframes would come with networking software already installed. Networks of varying sizes continued to become more commonplace.

On January 1st, 1983, ARPAnet changed protocol from its original limited version to the TCP/IP protocols we use today. The transition went surprisingly smoothly. This change to TCP/IP allowed the DoD to separate from ARPAnet, and form Milnet (Military Network).

## Internet today

When ARPAnet was a small network it was possible to maintain a table of all the computers and what their addresses were. With the open-ended architecture of TCP/IP, this was no longer feasible, and a system of maintaining addresses called the Domain Name Server was created. This system introduced a hierarchical scheme for addresses which operated in a similar way to regular post office addresses.

The National Science Foundation (NSF) wanted a high performance network to link supercomputers together with the aim of providing researchers with available resources. During the 80's, they created a high speed network which ran across the US, and became known as the Backbone of the Internet. Universities, Government departments and Businesses were all encouraged by the high speed link, and rapid growth of the Internet began.

ARPAnet ceased in 1990, by which time the Internet was a well established system. The creation of the WWW in 1991 allowed the general public to experience the Internet for the first time. Entertainment, Advertising, and uninformed opinion quickly became commonplace. Experienced users of the Internet longed for the "good old days" when the Internet was a small closed community. The general public was able to enjoy another form of mass media, but one in which they had control. In 1995, the NSF created a new Very High Speed backbone which was reserved solely for academic research.

## Internet Service Providers (ISP)

An Internet Service Provider (ISP) provides a way for an individual to access the Internet. The ISP usually has a host computer which maintains a permanent connection to other host computers and forms part of the Internet. The ISP will usually a client with an account on one of the host machines owned by the ISP. This account allows ISP to keep a record of how much information each user transfers, and provides a location where email and other files may be stored.

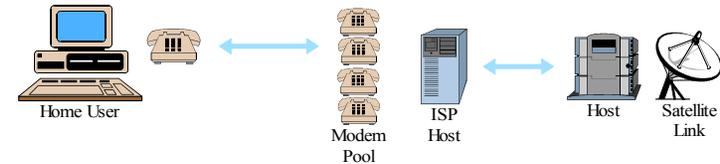


Figure 10: A home user normally accesses the Internet through an Internet Service Provider.

## Connecting to the Internet

In order for a user to access the Internet, they require a computer, a modem, software that allows the computer to connect to a host, and an account with an ISP. The communication software (or dial-up software in Windows 95) will tell the modem to dial the appropriate number and connect to the modem owned by the ISP. Once the modem has connected, the user must log into the account provided by the ISP (this is often achieved by scripts that run in the background). If the user has email waiting on the host machine, they may transfer it to their local machine to read it. Anytime the user looks at any information on the Internet (WWW pages, USENET news, email, IRC, etc.) the data must be transferred through the host computer, down the telephone line, through the modem and onto the local machine before it can be read.

## Communication

The Internet provides a world-wide communication medium. It is possible for a user to communicate to a single individual, small groups, or broadcast to a large audience. The most common methods of communication are through electronic mail, internet relay chat, or Usenet newsgroups.

## Electronic mail

Email is a mechanism that allows a user to type a message, and then send it through the Internet to a specified address. This is usually available with the lowest level of Internet access, and is considered by many to be the most essential aspect of Internet communication. The protocol by which email is transferred restricts the message to plain ASCII text. The development of MIME (Multipurpose Internet Mail Extensions) has allowed documents of any sort to be sent through email, greatly increasing its usefulness.

More recently, email facilities have been incorporated into WWW browsers such as Netscape. These browsers support the use of HTML code within the mail message, allowing different fonts, sizes, and styles to be used within the document. Some of these browsers even support the inclusion of pictures within an email message (although such messages cannot be easily read by those who are not using an email program which supports HTML).

## Internet Culture

Email usually reaches its destination within minutes. This fast delivery time has directly influenced the culture of Internet users in a variety of ways. Replies are usually received quickly (compared with snail mail), giving email a conversational flavour. This has resulted in an informal style of writing in which even spelling mistakes are taken for granted. Chat groups have almost instant delivery of text to the group, requiring people to develop methods of carrying on written conversations in real time. To reduce the time taken to get the message across, shorthand notation developed for common phrases (eg; FYI = for your information, TTYL = talk to you later, TNSTAAFL = there's no such thing as a free lunch). Emotion is difficult to express through pure text, yet it is important for conversations, and so emoticons

(smileys) were developed to fulfil this role. These are usually used to represent sarcasm, humour, and basic emotions such as happy or sad.

:-) ..... happy    ;-) ..... wink (sarcasm/ humour)  
 :-( ..... sad

## Mailing Lists

All email is handled purely by computer systems. It is therefore easy to automate the sending of email to many users. All that is needed is a list of addresses, and a computer program (called a mailer) can mail out a message to every address on the list. This feature is the essential idea behind mailing lists, which are a list of addresses of people interested in the same topic, together with a program that manages the list. Any mail which is sent to a specified address (that of the mailer) is automatically forwarded to each person whose address is on the list. Generally, mailing lists are for people who have a common interest in a particular topic (eg. Kite Making). It is estimated that there are over 30,000 different mailing lists today. The largest mailing list currently is A-Word-A-Day with over 100,000 subscribers.

## Privacy

There is no such thing as private email. There are no national laws which protect email. The administrators have access to all email that is sent to or from their system. It has been estimated that 25% of US companies read their employees' email messages. There is no reason to be alarmed however, since in practice, it would take far too long to read everyone's email. Just don't send anything *really* important (i.e. documents relating to national security) via email. The issue of email privacy has become a common concern in the US which is in the process of developing legally binding ethical standards for the computer industry.

## Usenet

The UNIX User Network (Usenet) began in 1981 when the UNIX to UNIX Copy Protocol (UUCP) was created. This protocol simply copied files from one UNIX machine to another across a network. It was quickly seen as an ideal way to build a message system, where any messages would be copied to all the other machines running UUCP. A backbone of machines supporting UUCP was set up spanning the US. The messages were stored and forwarded on a purely cooperative basis, at some expense to those providing the service. At this time, the groups fell into only 2 categories; mod were for moderated groups, and net were for all the others.

## Great Renaming

In 1986, the newsgroups were beginning to become too numerous to easily manage, and so the Great (or Grand) Renaming began. This was a move to restructure the Usenet groups into the hierarchical structure we have today and took about a year to fully realise. The main groups became comp, misc, news, rec, sci, soc and talk. The renaming made it possible for the administrators to copy only the first few groups (the important ones in their eyes), and not support the other groups. This possibility angered many, and one of the biggest flame wars of all times ensued. During this time, the proposed rec.sex, and rec.drugs groups caused the backbone to break as some of the administrators refused to support those groups. This in turn caused the creation of the anarchistic alt hierarchy (in which anyone can start up a group).

## Usenet today

The amount of information posted to Usenet newsgroups is staggering. Approximately 100 MB of information (the size of Encyclopedia Britannica) is posted each day, distributed between more than 25,000 different newsgroups. Many of these articles are kept for only a few days, others may be kept for weeks or months depending upon which newsgroup the message appears in. The estimated number of readers world-wide is over 100 million.

## Other Forums

The popularity of Internet browsers such as Internet Explorer have drawn the general public into the use of the WWW, but has not attracted so many into areas such as Usenet. Web-based forums are becoming increasingly common as a way of discussing special interest topics. Using the web for forum discussions has the added advantage of keeping the topic of discussion in the same place (same web site) as other information relating to the discussion. For example: A web site about the dangers of GE food can host pages informing the public of the danger as well as a discussion forum where questions can be raised and

answered. The more frequent use of forums on web pages is beginning to give the web a greater sense of community and a more interactive feel to users.

## Netiquette

With such a large audience able to publish (ie; post messages), there is a need for conventions governing acceptable behavior, which is commonly known as netiquette. Failing to follow these rules of netiquette has no official or formal effect, but reflects the poor judgement of the person posting. You would be well advised to read an article on netiquette before communicating via Email or posting to Usenet newsgroups.

## Timeline of some Internet Related events

- 1957 USSR launches Sputnik.  
The US forms ARPA to establish US lead in military technology
- 1962 Paul Baran, RAND: "On Distributed Communications Networks"
- 1969 ARPANET commissioned by DoD for research into networking  
First node-to-node message sent between UCLA and SRI (October)
- 1970 ARPANET hosts start using Network Control Protocol (NCP).
- 1972 Ray Tomlinson (BBN) writes basic email message send and read software  
Telnet specification
- 1973 Bob Metcalfe's Harvard PhD Thesis outlines idea for Ethernet  
File Transfer specification
- 1974 Vint Cerf and Bob Kahn publish "A Protocol for Packet Network Intercommunication" which specified in detail the design of a Transmission Control Program (TCP).
- 1976 Queen Elizabeth II sends out an e-mail  
UUCP (Unix-to-Unix CoPy) developed at AT&T Bell Labs
- 1979 USENET established using UUCP
- 1982 ARPA establishes the Transmission Control Protocol (TCP) and Internet Protocol (IP), as the protocol suite, commonly known as TCP/IP, for ARPANET. This leads to one of the first definitions of an "internet" as connected set of networks, specifically those using TCP/IP, and "Internet" as connected TCP/IP internets.
- 1983 Cutover from NCP to TCP/IP (1 January)  
ARPANET split into ARPANET and MILNET
- 1984 Domain Name System (DNS) introduced.  
Number of hosts breaks 1,000
- 1986 NSFNET created (backbone speed of 56Kbps)
- 1987 Number of hosts breaks 10,000
- 1988 Internet worm burrows through the Net, affecting ~6,000 of the 60,000 hosts  
NSFNET backbone upgraded to T1 (1.544Mbps)
- 1989 Number of hosts breaks 100,000  
NZ connects to NSFNET
- 1990 ARPANET ceases to exist  
Electronic Frontier Foundation (EFF) is founded by Mitch Kapor  
"The World" becomes first commercial ISP
- 1991 World-Wide Web (WWW) released by CERN; Tim Berners-Lee developer  
NSFNET backbone upgraded to T3 (44.736Mbps)
- 1992 Number of hosts breaks 1,000,000  
The term "Surfing the Internet" is coined by Jean Armour Polly
- 1993 Mosaic takes the Internet by storm.  
WWW proliferates at a 341,634% annual growth rate of service traffic.

- 1994 Shopping malls arrive on the Internet  
First Virtual, the first cyberbank, open up for business  
Vladimir Levin of St. Petersburg, Russia, is the first publicly-known Internet bank robber, stealing millions of dollars from Citibank between June and August.  
Arizona law firm of Canter & Siegel "spams" the Internet with email advertising green card lottery services; Net citizens flame back  
Yes, it's true - you can now order pizza from the Hut online  
Radio stations start rockin' (rebroadcasting) round the clock on the Net
- 1995 NSFNET reverts back to a research network. Main US backbone traffic now routed through interconnected network providers  
Hong Kong police disconnect all but 1 of the colony's Internet providers in search of a hacker. 10,000 people are left without Net access  
Radio HK, the first commercial 24 hr., Internet-only radio station starts broadcasting
- 1996 Internet phones catch the attention of US telecommunication companies who ask the US Congress to ban the technology (which has been around for years)  
The WWW browser war, fought primarily between Netscape and Microsoft, revolutionised the software industry with new releases made quarterly.
- 1998 Internet users get to be judges in a performance by 12 world champion ice skaters on 27 March, marking the first time a television sport show's outcome is determined by its viewers.  
Electronic postal stamps become a reality, with the US Postal Service allowing stamps to be purchased and downloaded for printing from the Web.
- 1998 Dot Coms (Internet companies) become the hottest item on the stock exchange. Millions are made as many Internet stocks increase by 10 fold.
- 1998 The game Everquest released and becomes the first really popular Massive Multiplayer Online Roleplaying game. Players pay a subscription and can create a character that they play in a huge 3D persistent world environment. It takes more than 8 hours of playing time to simply walk from one end of the world to another
- 2000 Everquest commonly known in the Internet community as EverCrack., one of the most addictive games people have played. Most players (community of 300,000+) have logged over 1000 hours of play. People reported as losing jobs and marriages broken due to playing Everquest
- 2000 The downfall of the tech stocks. Having doubled every year for the past 3 years, the bubble bursts and most tech stocks plummet in value. Dot Coms are worst hit, many going bankrupt or reducing to a fraction of their original value. Huge layoffs in IT staff follow.
- 2001 Research shows that a significant portion of children that meet a stranger online (in chat rooms etc.), go on to meet the stranger face to face in an unsupervised setting.
- 2002 Everquest blamed for the suicide of a teenage boy. Sony Online (Publisher) and Verant Interactive (Developer) of Everquest sued in US courts. Warning labels on "addictive" games a likely outcome.

Author: Andrew Luxton