COMP1200 - Perspectives on Computing

These slides were based on those of C Johnson, E McCreath and W Liang.

Outline

- Networks
  - Classifications
  - Topologies
  - Protocols
- The Internet
  - Addressing
  - Applications
  - Protocols
  - Security

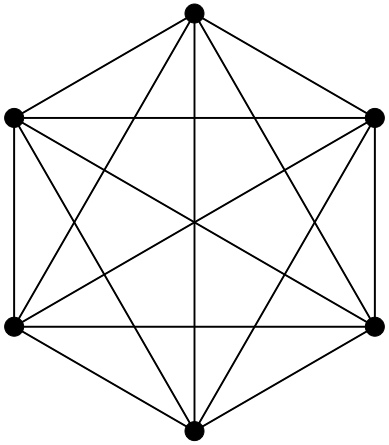
Networks

- A network is a number of computers linked together so that data can be transferred between machines
- Computers can exchange messages and share resources such as
  - printing capabilities
  - software packages
  - data storage
- The Internet (note the I) has become the most extensive network for computers to communicate across the world

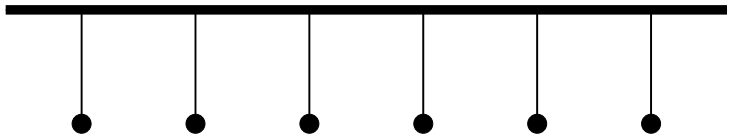
Network Classification

- Local Area Network (LAN)
  - computers in a single building/institution e.g. ANU
- Metropolitan and Wide area Networks (MAN and WAN)
  - computer networks that span larger areas
- Open Network
  - internal operations based on designs that are in the public domain e.g. the Internet
- Closed/Proprietary Network
  - privately owned e.g. developed by IBM

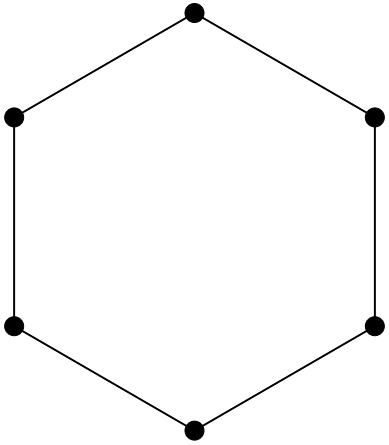
Network Topologies



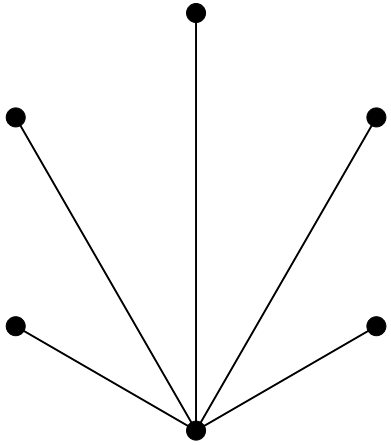
Fully Connected



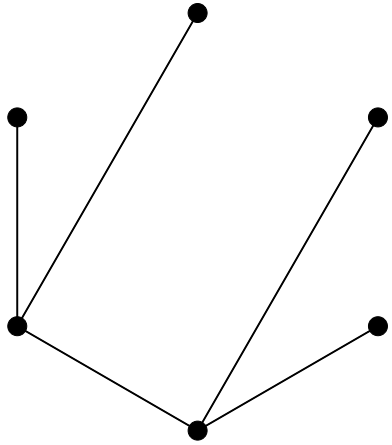
Bus



Ring



Star



Tree

Network Protocols

- Protocols are the rules by which network activities are conducted
  - e.g. transmitting messages around a network
- Token Ring Protocol
  - transmit messages in one direction around the ring
  - when the message reaches its destination, a copy is taken and the message is forwarded on
  - when the message reaches its starting point, the message is removed from the ring

Token Ring Protocol

- The token ring protocol relies on inter-machine cooperation
- Machines are not permitted to send any message they want
- A unique bit pattern (token) is passed around the ring
- Possession of the token allows a machine to transmit its own messages
- Without the token only forwarding is allowed
- When the message completes its cycle the machine forwards the token to the next machine

Carrier Sense, Multiple Access with Collision Detection (CSMA/CD)

- Used in bus topology networks
- Each message is broadcast to all machines on the bus
- Machines monitor all messages but keep only their own
- To transmit a message, a machine waits for no activity
- If two machines transmit at once, they both wait for a random period of time before trying again

Ethernet

- Ethernet is a set of standards for implementing a LAN with a bus or star topology
- Ethernet has become one of the most popular standards for transmitting data between computers on a LAN
- Ethernet will transmit data at either 10 or 100Mbps

Combining Networks

- Sometimes it is necessary to combine networks into larger ones
- This is achieved by means of different devices
- Repeater
  - connects two buses
  - passes messages back and forth without considering their meaning

Combining Networks

- Bridge
  - connects two buses
  - passes messages across the connection only if that message is destined for a machine on the other side
- Switch
  - connects more than two buses
  - when connected they look like spokes on a wheel
  - passes messages only to the spoke that contains the machine that a message is destined for

Combining Networks

- Sometimes, networks that have incompatible characteristics must be connected
  - A ring network using the token ring protocol
  - A bus network using CSMA/CD
- Build a network of networks called an internet
- The individual networks maintain their individuality
- Connection is handled by a machine called a router
- It's task is more significant than a bridge or repeater

Interprocess Communication

- This is the communication between processes on a machine within a network in order to coordinate their actions and complete their tasks
  - Client/Server Model
    - a process that makes a request is a client
    - a process that satisfies the request is a server
    - e.g. a print server
    - e.g. a file server

Interprocess Communication

- Peer-to-Peer Model (P2P)
  - involves two processes communicating as equals
  - both processes execute on a temporary basis
  - e.g. instant messaging
  - e.g. computer chess

Distributed systems

- These consist of software units that execute as processes on different computers
  - global information retrieval systems
  - company-wide accounting and inventory systems
  - computer games
- The machines on which they reside are called hosts

The Internet

- Originated from research projects in the 1960's
- The aim was to develop the ability to connect a variety of networks
- The Internet can be viewed as a collection of domains
- A domain is a network or small internet operated by a single organisation

The Internet

- Establishment of domains is overseen by the Internet Corporation for Assigned Names and Numbers (ICANN)
- To establish a domain you need to register via a registrar
- Attach a registered domain to the Internet via a router
- A domain's router is called a gateway

Internet Addressing

- An internet must be associated with an internet-wide addressing system that assigns an address
- This is the Internet Protocol (IP) address
- An IP address has 32 bit and comes in 2 parts
  - The Network Identifier
    - identifies the domain in which it resides
  - The Host Address
    - identifies the particular computer within the domain

Internet Addressing

The string

10010110110001011011000110000 1000

is the same as

150.203.99.8

which is the same as

www.anu.edu.au

This unique mnemonic address is the domain name

Internet Applications

- Electronic Mail
  - All mail goes via a mail server
  - Email address is usually user@mailservername
- The File Transfer Protocol
  - FTP is a client/server protocol
  - transfers files across the Internet
- Telnet and SSH
  - Allows computer access from a great distance
  - Telnet is not encrypted - passwords are not secure
  - Secure Shell (SSH) fixes this problem

The World Wide Web

- Multimedia information is disseminated over the Internet using hypertext
- hypertext originally meant text documents that contain links to other documents
- Sometimes use hypermedia - to include video, audio, images etc.
- Portions of a document are linked to other documents forming an intertwined web of related information
- This web is the World Wide Web, WWW, W3, The Web, The Net, . . .

Web Implementation

- Software packages that allow access to hypertext are either browsers (clients) or web servers (servers)
- Documents are transferred between browsers and web servers using the Hypertext Transfer Protocol (HTTP)
- Each document is given a unique Uniform Resource Locator (URL)

http : //cs.anu.edu.au/students/comp1200/lectures/lectures.html

HTML

- Hypertext documents are written using Hypertext Markup Language (HTML)
- Files contain text and tags
- Tags are used to say
  - how text is to be displayed
  - which multimedia resources to include
  - which other documents there are links to
- The file encoded into HTML is called the source code

HTML

```
<HTML>

<HEAD>

<TITLE>My Title</TITLE>

</HEAD>

<BODY>

<CENTER><H1>My Heading</H1></CENTER>

This is some text<BR>This text appears on a new line<BR>

<A HREF="anotherpage.html">This text links to another page</A><BR>

<CENTER><IMG SRC="mypicture.jpg"></CENTER>

</BODY>

</HTML>
```

HTML

My Title

My Heading

This is some text
This text appears on a new line
This text links to another page

eXtensible Markup Language

- A standardised style for representing notational systems for representing data as text files
- Notational systems (called markup languages) have been developed for representing
  - Mathematics, Multimedia presentations, Music, . . .
- HTML is based on the XML standard for representing web pages
- Since HTML was invented before XML we use XHTML to refer to HTML that conforms rigorously to XML standards

Internet Protocols

- A layered approach to Internet software
- Principal task of networking software is to provide an infrastructure for transferring messages
- Internet software has 4 layers of routines
  - Application layer
  - Transport layer
  - Network layer
  - Link layer

Internet Protocols

- Messages originate in the application layer
- They are passed through the transport layer and the network layer
- They are transmitted by the link layer
- The target machine receives the message via the link layer
- The message is then passed up though the network layer and the transport layer to the application layer

Internet Protocols

- Application layer
  - Software units that use Internet communication to carry out tasks
  - Responsible for providing an address for the message that is compatible with the transport layer

Internet Protocols

- Transport layer
  - Ensures messages are properly formatted for transport
  - Divides message into small segments
  - Adds sequence numbers to each segment (so that it can be re-assembled later)
  - Attaches addresses to each segment (packet)
  - Passes the packets to the network layer

Internet Protocols

- Network layer
  - Responsible for forwarding the packets it receives from one network to another
  - If the destination address is within the current network, it sends the packets to that address
  - Otherwise it sends them to a router in the network
  - Once it knows where to send the package, it appends this intermediate address to the packets and passes them to the link layer

Internet Protocols

- Link layer
  - sends the packets to the intermediate address
- The message is received by the link layer and passed to the network layer
- If the destination address is not local it is passed to the link layer with a new intermediate address
- If the packets have reached their final destination, they are passes to the transport layer
- The transport layer re-assembles the packets and passes the message to the application layer

TCP/IP

- The TCP/IP protocol suite is a collection of protocols used to implement the 4 level hierarchy
- TCP = Transmission Control Protocol
- IP = Internet Protocol
- There are many more
- TCP defines a version of the transport layer
- The User Datagram Protocol (UDP) also defines a version of the transport layer

TCP/IP

- Two basic differences
  - TCP sends its own message before sending packets
  - TCP transport layers at origin and destination communicate to ensure all packets have been sent correctly
- IP is the Internet's standard for the network layer
- A hop-count is appended to each packet
- This limits the number of times a packet can be forwarded
- 32 is usually enough

Security

- A computer connected to a network is subject to unauthorised attack
- Usually by malicious software (malware)
  - Virus
    - Software that insert itself into a program
    - Corrupts data or other programs
  - Worm
    - Transfers itself though a network
    - Forwards copies of itself to other computers
    - Can ultimately overload an entire network

Security

- Trojan Horse
  - Disguises itself as something useful and is willingly imported via email attachment
  - May lay dormant for a specific period
- Spyware
  - Collects information about a computer
  - Reports this back to the instigator
- Phishing
  - A way of obtaining information by asking for it
  - Usually via email

Security Protection

- Firewall
  - Installed at gateway
  - Blocks incoming messages with origin addresses within the domain
- Spam Filter
  - Firewall designed to block unsolicited email
- Proxy Server
  - Acts as an intermediary between client and server
  - Use for Ftp, Telnet, etc.

Security Protection

- Anti Virus software
  - Must be routinely updated
- Encryption
  - Passwords may be compromised
  - Examples of encryption include
    - FTPS
    - SSH
    - HTTPS
    - Public-key encryption

Reading and Self Assessment

- Required reading:
  - Chapter 4, sections 1, 2, 3, 4 and 5.
- Questions:
  - 4.1 : 1, 2, 3.
  - 4.2 : 1, 2, 3.
  - 4.3 : 1, 2, 3.
  - 4.4 : 1, 2, 3.
  - 4.5 : 1, 2, 3.

Algorithms

- "The study of algorithms is the cornerstone of computer science"
- An algorithm is a set of steps that define how a task is to be performed
  - e.g. consider the machine cycle
    - as long as the halt instruction has not been executed, continue to do the following steps:
      - fetch an instruction
      - decode the instruction
      - execute the instruction

Algorithms

- Formal Definition:
  - An algorithm is an ordered set of unambiguous, executable steps that defines a terminating process
- The set of steps must be ordered
- The steps must be executable
- The steps must be unambiguous
- The algorithm must define a terminating process
- The term 'algorithm' is often used to represent steps that don't necessarily define a terminating process

Algorithms

- An algorithm and its representation are distinct
- An algorithm is abstract from its representation - an algorithm may be represented in many ways
  - $F = \frac{9C}{5} + 32$
  - Multiply the temperature in Celsius by 9; divide this by 5; then add 32
- A program is a representation of an algorithm
- A process is the activity of executing an algorithm

Primitives

- The representation of an algorithm requires some form of language
- Problems are caused when
  - the language is not clearly defined
  - information is not given in adequate detail
- Computer Science establishes a well defined set of building blocks (primitives) from which algorithm representations can be constructed

Primitives

- Assigning precise definitions to primitives removes ambiguity
- Requiring algorithms to be described in terms of these primitives establishes a uniform level of detail
- A collection of primitives along with rules that state how they can be combined to construct more complex ideas constitutes a programming language

Primitives

- Primitives come in two parts
  - Syntax
    - refers to the primitive's symbolic representation
  - Semantics
    - refers to the primitive's meaning
- Algorithms are usually represented by a collection of "high level" primitives
- These are formed from the "low level" primitives in the machine's language
- The result is a formal programming language

Pseudocode

- Pseudocode is something less formal than a programming language
- Pseudocode is a notational system in which ideas can be expressed informally during the algorithmic process
- Need to develop a consistent, concise notation for representing recurring semantic structures

Pseudocode

- Saving a computed value
  - $name \leftarrow expression$
  - assign $name$ the value of $expression$
- Selecting one of two possible activities depending on the validity of of some condition (TRUE or FALSE)
  - if ($condition$) then ($activity1$) else ($activity2$)
  - If the $condition$ is TRUE then $activity1$ is performed
  - If the $condition$ is FALSE then $activity2$ is performed

Pseudocode

- The repeated execution of a statement
  - while $(condition)$ do $(activity)$
  - Repeatedly do $activity$ as long as $condition$ is TRUE
- In order that the pseudocode may be used as abstract tools in other applications, we define a procedure
  - procedure $name$
- As procedures may be used in different situations we introduce parameters to procedures
  - procedure $sort(list)$

Pseudocode

- The development of a program consists of:
  - Discovering the underlying algorithm
  - Representing that algorithm as a program
- Algorithm discovery is usually the most difficult part
- We would like to reduce the process of problem solving to an algorithm itself
- This has been proven to be impossible as there exist problems that have no algorithmic solution

Problem Solving

- Basic Principles (Polya 1945)
  - Phase 1:
    - Understand the problem
  - Phase 2:
    - Devise a plan for solving the problem
  - Phase 3:
    - Carry out the plan
  - Phase 4:
    - Evaluate the solution (for accuracy and its potential)

Program Development

- Translated into the context of program development, these become
  - Phase 1: Understand the problem
  - Phase 2: Get an idea of how an algorithmic procedure might solve the problem
  - Phase 3: Formulate the algorithm and represent it as a program
  - Phase 4: Evaluate the program (for accuracy and its potential as a tool for solving other problems)

Program Development

- These are not necessarily steps to be *followed* when trying to solve a problem
- These are phases that will be completed sometime during the solution process
- Don't *follow* - Take the initiative and **lead!**
- Phases are not necessarily completed in sequence

How Do You Make A Start?

- Try working backwards
  - If you know what the output should be, maybe you can figure out the algorithm more easily
- Look for a related/easier problem
  - This problem may have been solved already
- Apply step-wise refinement
  - Treat as several sub-problems
  - Approach the overall solution in terms of steps
  - Top-down Methodologies versus Bottom-up Methodologies

How Do You Make A Start?

- You should find algorithms for all instances of a problem (not just one)
- Suppose we are asked to sort some integers into ascending order: 5, 7, 3, 6 for example
- One approach might be
  - Swap 3 with 5 (3, 7, 5, 6)
  - Swap 7 with 5 (3, 5, 7, 6)
  - Swap 7 with 6 (3, 5, 6, 7)
- This is too specific - needs to be more general

Reading and Self Assessment

- Required reading:
  - Chapter 5, sections 1, 2 and 3.
- Questions:
  - 5.1 : 1, 2.
  - 5.2 : 3.
  - 5.3 : 3.