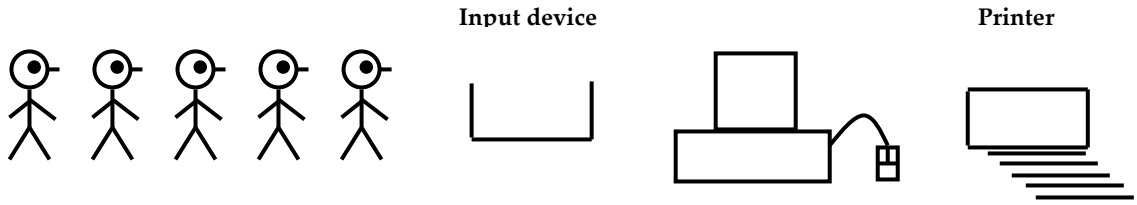
Chapter 26 - An introduction to batch processing

26.1 Introduction

Computers cost a lot of money. The equipment costs money and the infrastructure that ensures they are always working such as employing a network manager, for example, costs money. Because of the high costs involved, computer designers and software designers spend a lot of time making sure that computers can do jobs efficiently. They try to ensure that computer time is used as wisely as possible. One result of this is that there is now a wide range of types of operating system.

26.2 Batch processing - in the beginning...

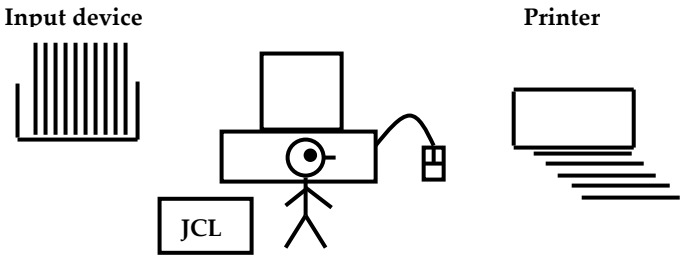
In the early days of computers, operating systems were very basic. If a person wanted to use a computer for something (a 'job') they would need to load their program into the computer and then load up the computer with the input data. The input data might be held in a coded form on punched cards. When both the program had been loaded up and the punched cards put into the input device, the program would be 'executed'. The program would read the data, process it and then send the output to a printer. When the job had been finished, the next person could use the computer. This changeover wasted a lot of the CPU's time. The CPU is a very fast device. If it has to wait while data is read into it or while data is sent out of it, then it won't be working to its full capacity. This problem is sometimes referred to as a '**speed mismatch**' problem. We have seen how buffers help overcome the problem of speed mismatch and we will meet this problem again later.



Jobs were processed very slowly, one after another.

To improve the situation, a new input device was designed that could hold lots of punched cards. Similar jobs were collected together and put into a 'batch'. It was possible to stack all of the input cards together for one batch process, and then also have some other input cards collected together for a different batch process. Of course, if you were going to put lots of jobs together into a batch, and have lots of batches together, then there needed to be a way of managing and coordinating these different processes. An operator, therefore, needed to also write some instructions to do just that. The program was written using a **Job Control Language** (JCL). A JCL program would specify such things as when to start processing the batch, which program to use with which punched cards and what printer to use for the output or what file to write the output to.

Batch processing improved matters because this type of processing ensured that the CPU was being used as efficiently as possible - it wasn't waiting around so much while an operator regularly loaded in programs and data. Although batch processing systems have their roots in systems developed many years ago, they are still much in evidence today.



Batch processing collects all the jobs together.

26.3 An example of batch processing today

Consider the example of an Internet bank. At the end of every month it might prepare everyone's bank statements and then email all of their customers, telling them that their statements were ready.

- Batch processing involves many jobs, in this example, many statements that need processing.
- The jobs are all similar in nature. Each statement is processed in broadly the same way.
- A JCL program is written, telling the computer what files to use, when to start the processing, what to do with the output and so on.
- Some instructions in the JCL program tell the computer to process the statements at a time when the network's resources are not needed for other, more important things. For example, the processing doesn't begin at 9.00am on Monday morning, but 1.00am Sunday morning.
- The processing begins. The output is not immediate - the processing might take hours. If it were really important to have the results of the processing immediately, batch processing wouldn't have been selected.
- Once the JCL has been written and loaded, it takes over the management of the batch processing. Once the job starts, there is no human intervention necessary in the process.
- There is no physical output in this particular case - the output is in the form of updated statements held in electronic form and prepared emails.
- Once the processing starts, it continues uninterrupted until it is finished.
- At a later time, when the computers are free, the file containing emails is accessed and all emails are sent.
- When a customer receives an email telling them that their statement is ready, they can log in to their account. They will then be able to access and print off their statement.

Batch processing, then, is a way of improving the efficiency of a computer or network's resources by bringing together similar work and processing them all in one go at a time when resources aren't needed. One of the main things to remember is that the output from a batch-processing run isn't immediate. Summarising the key points about batch processing:

- Jobs are collected together over a period of time.
- Jobs are processed altogether in one go, at the same time.
- It is not important that the output is available immediately.
- It is worth noting that when a group of jobs are collected together in a batch, they are often batch processed in the middle of the night, using computers and printers when nobody else needs them.
- The jobs can be set up to be batch processed because it is possible to write a little program using a Job Control Language that manages the work to be done.
- Because the JCL manages the batch processing, once it starts, no further human intervention is required.

26.4 Speed mismatch problems

The above example is interesting because data processing will be relatively fast. The data input needed is held on file and the output is also to files. There are many data processing activities, however, that still involve getting data into a system from paper-based records and this in itself can slow down the CPU. (Remember, reading data from an input device or outputting data to an output device is a very slow process compared to the speed the CPU could work at. This is the problem of **speed mismatch**). There are also systems that require physical output in the form of printouts, for example. Because printing is a very slow operation compared to the speed that a CPU works at, this is yet another source of concern for the efficiency of batch processing. Fortunately, something can be done to help!

26.5 Offline processing

To help speed up data input and output into computer systems, **offline processing** was 'invented'. This removes the need for relatively slow input devices. Instead, data is stored in files on a high-speed data storage device. This could be done in a number of ways. For example, an operator might use a different computer system to simply type in data from paper forms into a file, perhaps using a **key-to-disk system**, discussed later in this chapter. Or if the data was in a machine-readable form (such as OMR sheets), it could be read automatically and stored in a file. Then, when the computer is ready, it can get all the data it needs from these files on the high-speed storage device rather than from relatively slow input devices. This is described as 'offline processing' because the main computer doing the processing is not immediately controlling and reading the data from its input devices. The data is being prepared and stored away from that computer on a high-speed storage device and is then made available as necessary.

26.5.1 More examples of offline (and online) processing

A 'network' is a number of computers (more than one) that have been physically connected together. This enables them to communicate with each other and share files and share hardware such as printers. A computer that is part of a network can be

either ‘online’ or ‘offline’. If it is offline then it is **physically connected** to the network but not actually using the resources of the network. It is **not logged in** to the network. For example, you may have a computer at home that has a modem plugged into your telephone socket. This is known as a ‘dial-up connection’. Your computer is part of the Internet. Some people, however, write their emails ‘offline’. This means that they actually write their emails without logging on to the Internet. They are writing emails offline because they are not actually using the resources of the network. When they want to send their emails, however, they need to log on to the Internet. As soon as you connect to the Internet via your Internet Service Provider, or ISP, then you go **online**. You are using the resources of the network to send emails. You are physically connected to the network **and** you are logged into the network. This is just one example of the meaning of ‘online’ and ‘offline’.

'Online' and 'offline' are terms usually associated with networks. They describe whether a computer or workstation is or isn't using the resources available on the network. In particular, it describes whether or not they are using the processing power available on the network. If the computer or workstation is using the network's processing power then it is said to be **online**. If it isn't (but is physically connected) then it is said to be **offline**. The email example is one example where you can be either online or offline. There are other examples. Someone working on a network might be typing in data into a file on their workstation's disk. They are working 'offline'. If this data is then used to update a master file held on the server then the user will go online to do this. They contact the server by logging in, tell it where to find the data stored on disk, the server processes the data and the master file is updated.

Consider another example. Somebody who wants to ask you some questions for a questionnaire stops you in the street. As you give your answers to their questions, they record your responses directly into a database on their tablet PC. When the interviewer goes back to the office, they transfer all of the responses of all of the people that have been interviewed into the master database held on a central computer. The data is collected **offline** (because the interviewer is not using the processing power of the main computer) but the data is transferred and analysed **online**, because the processor in the main computer is required to transfer the data and to work on it.

26.6 Key-to-disk data entry

Organisations need ways of ‘storing data offline’. In other words, organisations need ways of getting data from the outside world into a file on their computer system but without tying up the resources of their computer system to do this. A ‘modern’ key-to-disk system might be employed by an organisation to do just this because it has to transfer lots of details held on paper forms into a file so that the data can then be processed later. This system might work as follows:

- A team of data entry operators key the data on the paper forms into a temporary file on the computer.
- As the data is entered, it is **validated** to ensure that only sensible data is entered into the system.
- Data is stored on disk after it is validated.
- The same paper forms are passed to a second operator who re-enters the data.
- The data is then **verified**. As the second operator types in the data, it is compared to the data held in the temporary file typed in by the first operator.
- If the data is not the same, an error is reported to the operator. They will then investigate and correct the problem.
- If the data is the same, it is accepted and can be made available to the batch processing system.

This key-to-disk system, with its validation and verification techniques, will pick up most errors in data input but it is still possible for incorrect data to be put into the system. It could be that there is a mistake on the original form that means both operators enter the incorrect data. Perhaps both operators can't read the form and both of their best guesses are incorrect!

26.6.1 In the old days …

The ‘modern’ key-to-disk system, then, is a way of offline-processing the data input. If you go back twenty years, a data input device called a ‘key-to-disk input device’ was commonly used by companies to get data on paper forms into a computer file. An operator would write data directly from a special keyboard that had a magnetic disk drive. As they typed, the data was saved directly onto a disk in the drive. When the main batch processing computer was ready, it would simply read the data from the disk. Reading data from a file on magnetic disk was much faster than reading data from a keyboard, when the CPU had to work at the comparatively slow speed of the data input operators. With the development of computer power and network sophistication, however, and the subsequent reduction in computing costs, key-to-disk input devices have largely become obsolete and have been replaced by the ‘key-to-disk’ system described earlier.

Offline processing of output can also occur. Instead of the output from a batch processing job being sent to a printer it could be sent to a file. The printing can then be done at a later date, when the computer is not being used for processing for a more important task. By not sending the output directly to a printer, the computer can choose when to manage and work at the relatively slow rate of the printer. We saw examples of this with the Internet bank example described earlier in this section.

Offline processing techniques, then, allow organisations to get lots of batch jobs ready for processing. If all of the batch jobs are prepared offline, then the CPU can get each job at will, in the order that makes the most efficient use of the CPU. Selecting what order to do jobs in is known as 'scheduling'. This is discussed fully in a later chapter.

26.7 Batch totals, control totals and check sums

We have seen examples where data is entered into a file offline so that it can be processed using batch processing. Integral to this concept is ensuring that the data can be collected into a file **accurately**. To help us ensure this, we can make use of batch totals, control totals and check sums.

26.7.1 Batch totals

One method used to ensure that the integrity of a set of data is maintained when it is entered into a computer system is to use a 'batch total'. Before an operator enters a set of data or a set of records into the computer, they calculate a number known as the 'batch total'. This is worked out by adding up the total number of data items or records in the set to be entered. This 'batch total' is then entered into the computer along with the data or records. For example, if an operator was entering a set of product codes into a computer system, they could generate an extra number by counting up manually all the actual codes that had to be entered. This number is the 'batch total' and is entered into the computer by the operator along with the product codes. Once the codes and the batch total have been entered, the batch total is re-calculated automatically by the computer. It is then compared to the original batch total entered by the operator. If they are the same, then the data's integrity has (probably) been maintained. If they are different, then the operator may have missed out a product code or may have entered a code twice accidentally. The computer system would then generate an error message to highlight the problem to the operator.

26.7.2 Control totals (also known as hash totals)

This is another check that can be easily done on sets of data and can also be used in addition to batch totals. To generate a control total, the operator might add up pieces of the data entered. In the example used to describe what batch totals were, the operator might add up the first digit of each of the product codes. This produces the 'control total'. The control total is then entered into the computer system (perhaps also with a batch total as a double-check). Once again, the computer will recalculate the control total automatically and signal to the operator if it differs from the one the operator entered. Consider an operator who wants to enter in the following product codes:

2324455 3434333 7823444 6555678 6556665

The batch total would be calculated as 5 (because there are 5 codes to be entered). A control total might be calculated by using the algorithm 'add the first numbers of all data entered'). This would give 2 + 3 + 7 + 6 + 6 = 24. The operator would then enter not only the data but also the batch total and the hash total. The computer, once all the product codes and the batch and control totals were entered, would recalculate the batch and control totals automatically. If they differed to what the operator entered then an error message would be displayed on the input screen. If the data consisted of words rather than numbers, then the characters in each word could be converted into numbers using their ASCII equivalent codes. These can then be used in a suitable algorithm to generate a control total.

26.7.3 Check Sums

When computerised data is sent from one computer system to another, the sending computer can calculate a 'check sum' automatically. This is simply a number produced by adding up all of the individual pieces of data. When the receiving computer gets the data and the check sum, it recalculates the check sum and compares it to the check sum it received. If they are different, it means the data was sent incorrectly. The receiving computer can then signal to the sending computer to resend the data. Check sums are very similar to batch totals. The main distinction is that check sums are calculated by computer systems whereas batch totals are calculated by humans!

26.7.4 Data consistency

As well as ensuring that data gets entered into the computer correctly, designers also try to ensure that the consistency of the data is maintained. This term describes the relationship between the input data, processing on the data and the output data. For example, if a person's actual age in a database is not the same as their age when calculated using their date of birth, then the data has become inconsistent. In this case, it might be because of an incorrect formula, or an incorrect date of birth or age has been entered into the system.

26.7.5 Data integrity

So far, we have considered the use of validation rules to try and maintain data consistency as well as the use of batch totals, control totals and check sums to ensure that all the data gets entered in to a system. However data already in a system can become corrupted as a result of some processing that is done on that data. The phrase 'data integrity' refers to the whether the

data already in a computer has been corrupted after some processing has taken place on it. For example, consider a computer system that had to work out the VAT on some prices and then restore the new prices (including VAT), over-writing the old ones. If, as a result of processing, some negative numbers were produced that then got restored into the system, then something went wrong with the processing of the data! As a result of the processing, the data became corrupted. It has lost its integrity.

26.7.6 Data that has passed checks can still be wrong

Data can still conform to validation rules but be entered into a database incorrectly. For example, you may have a validation rule for AGE that says you must be an integer greater than 5 and less than 100. If you enter 10 for your age, clearly this is valid data because it conforms to the validation rule. We can also say that it is consistent data because it has been entered into the database in the correct form that we want, an integer. But the data is not correct because you are not 10 (are you)? By having **verification procedures** in place such as in the key-to-disk data entry example described earlier or by visually double-checking entries, you should be able to pick up most of these kinds of problems. It is still possible, however, for incorrect data to get past all of your validation and verification checks and be entered into the system.

- If the information is wrong on a paper form in the first place it will be entered into the system incorrectly.
- If the handwriting can't be read clearly then it may be incorrectly interpreted.
- It is also possible that, as a result of some processing you have done on some data that's already in the system, the data gets corrupted and saved back into the system. This may be, for example, because of a bug in the processing software.
- It may also be that a hacker has broken into the system and has altered the data.

If an operator incorrectly calculated a batch total, and then the wrong number of records were entered, the batch total might actually match the incorrect number of records entered! The most important asset a company has is its data. Data is the one thing that takes years to amass and is very valuable. Any good system needs checks in place to try and pick up problems that might occur during data entry and processing.

26.8 Transaction files and master files - an example

Closely related to the idea of batch processing is the idea of a 'transaction file' and a 'master file'. Read through the following example carefully and see if you can follow it.

A company has many hourly-paid workers who are paid weekly. Each worker's personal details are stored in a computer record, which also holds details of how much that worker has been paid so far this tax year, how much has been paid towards their pension and insurance, how much tax has been paid and so on. All of the records about each employee are held in a nice organised order, using their Employee Identification Number. The first employee will be number 1. The second employee in the file will be number 2, and so on until the end of the file. This file (a file is a collection of records) is **not accessed frequently**. It is called the '**master file**'.

Each worker clocks on and off each day using a computerised system. Each time a worker clocks on or off, a record for that worker is accessed and updated. All the records that are **frequently** accessed and changed are known as the '**transaction file**'. At the end of the week, everyone's pay must be calculated and a payslip issued. To do this, a Job Control Language program is written. This lays out what files to use, when to use them, what program to run them with and what printer to send the output to. Both the transaction file and the master file are then accessed at the time laid down in the JCL. The details held in the transaction file are used to:

- Update the details in each employee's record in the master file.
- Calculate how much each employee should be paid and how much tax they should pay.
- Produce a payslip for each employee using the details held in the master file record for that employee and the record for that employee in the transaction file.

The details held in the master file are required so that, for example, the name can be retrieved and printed on each payslip. Each record in the master file will also need to be updated with the new details about, for example, how much a person has been paid and how much tax they have paid to date. Given that employees have been clocking in and out in a random order all week, the entries in the transaction file will all be in a different order to the nice order that the master file will be in. The first job that has to be done is to **sort** the transaction file entries so that they are in the **same order** as the master file, with all of the entries in the transaction file organised in ascending order by Employee Identification Number. If this didn't happen, it would take much longer to process a payslip because you would have to do a lot of searching for all of the entries for any particular employee throughout the transaction file once the master file entry for that employee had been retrieved. The entries for any one employee will be all over the transaction file, which is why it takes more time than if they were all together and in order! Once the transaction file is in the same order as the master file, the payslips can then be batch processed. Note that the company's

computers will be used in the middle of the night, when they are **not needed** for more important, more immediate tasks, the output from the batch processing **isn't immediate** - the payslips will be printed off and ready some hours after the processing has begun and once the payslip processing has started, there won't be any need for human intervention.

26.9 Backing up the master file and the transaction file

Once a particular week's transaction file has been used to print off payslips and update the master file, the transaction file is wiped clean, ready to be used for the next week's details. Interestingly, the only time when the master file is completely up-to-date is immediately after the master file has been updated. At any other time, the master file is out-of-date because it doesn't contain the details held in the transaction file. This means that if a company wants to ensure its data is backed-up, then it needs to ensure that both the master file and the transaction file are backed-up.

26.10 Real-time processing

When batch processing was discussed, you saw that one of the features of batch processing was that the output wasn't immediate. You may not get the output for a number of hours. In real-time processing, time is critical! The length of time a program takes is important, as is the way that the CPU deals with instructions. Real time is the most important thing. Making sure that the CPU is used as efficiently as possible as with batch processing is *not* as important. Most real-time applications can be found in booking systems or in control applications. We will discuss real-time systems in the next chapter.

26.11 Some jargon to know

We have discussed a batch processing system. We have also mentioned many times different operating systems such as Windows, Linux, DOS, and so on. We will briefly introduce some phrases here that you may come across.

26.11.1 Single-user operating systems

Some operating systems only allow **single users**. DOS, Windows and Linux are good examples. A single user has access to all of the CPU processing power available. They can be multi-tasking as well - the user can run more than one application at the same time. See below.

26.11.2 Multi-tasking operating systems

Some operating systems allow you to open **more than one application** at a time. You could have open a word processor and a chat room, for example, and have an MP3 player playing in the background. Many single-user operating systems also allow multi-tasking. Windows does, for example, but DOS doesn't. DOS is a single-user single-task operating system.

26.11.3 Multi-user operating systems

Sometimes, you need to have a database of information stored centrally but it needs to be accessed by lots of people. For example, you might have a database of products and their prices on a computer in a supermarket. Lots of terminals need to access the database. Each terminal is given a very short amount of computer time before the next terminal is serviced. This happens so quickly that each terminal thinks it is the only one with access to the database! This kind of system is known as a multi-user operating system.

Q1. Summarise the six characteristics of batch processing systems.
Q2. What does 'offline' mean?
Q3. What does 'online' mean?
Q4. State three methods of ensuring that data is entered into a system accurately.
Q5. Give a clear example that illustrates data consistency.
Q6. Describe two ways that data, which has passed all checks, may still be entered into a database incorrectly.
Q7. What is a transaction file?
Q8. Describe how a transaction file is used to update a master file.
Q9. What does multi-tasking mean?
Q10. What is a multi-user operating system?

Chapter 27 - An introduction to real-time processing

27.1 Introduction

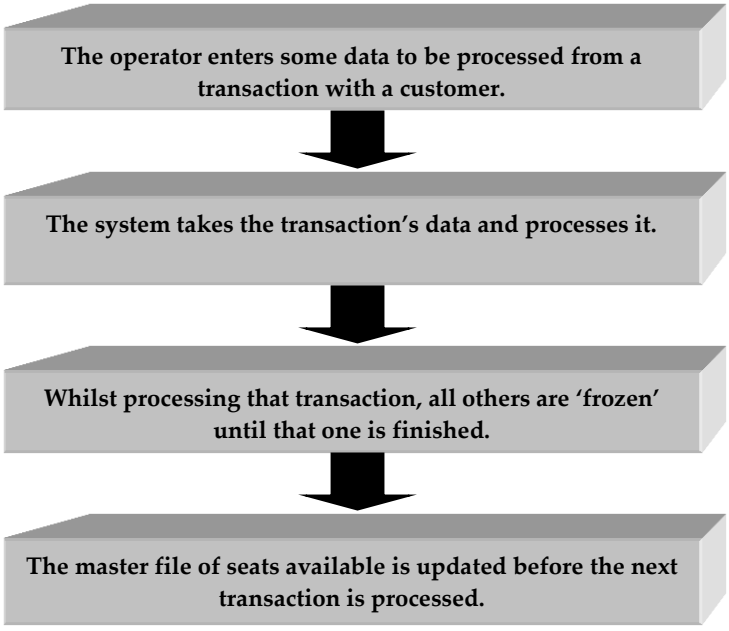
With batch processing you don't get the output straightaway. With real-time processing, an input is received and then processed. Any output generated is processed quickly enough to influence the system before the next input is received! You get the output 'immediately'.

27.1.1 Some examples of real-time systems

For example, in an air traffic control system, planes' details are displayed on a screen. When a plane has moved to a different position, the screen needs to be updated. If this didn't happen quickly enough, the screen would show a situation that didn't reflect the actual situation. The air traffic controller might then give the pilot instructions based on the inaccurate information displayed on the screen. This may result in an accident. When a missile is going towards a pre-programmed destination, it is constantly readjusting its aim because of wind conditions. It is feeding back into its processor information about its direction from input sensors, comparing that information with where it should be going and then making any necessary adjustments.

27.2 Booking systems

Real-time systems don't have to react with the speed of light, however! For example, if you buy a ticket for the theatre, the booking system will receive an input to say that you have booked a particular seat. It has to process the input and update the file that stores which seats are still available before the next ticket is sold. If the system weren't updated quickly enough then the same seat might be sold again. This is an example of a real-time system, even though the process of booking a ticket may take a few minutes. The reason it is a real-time system is that the master file (the file which holds details of which seats have been sold and which ones are free) is updated before the next seat is booked. Each booking affects the next one. Just to confuse things, the above example is also known as a '**transaction process**' as well as a **real-time process**.



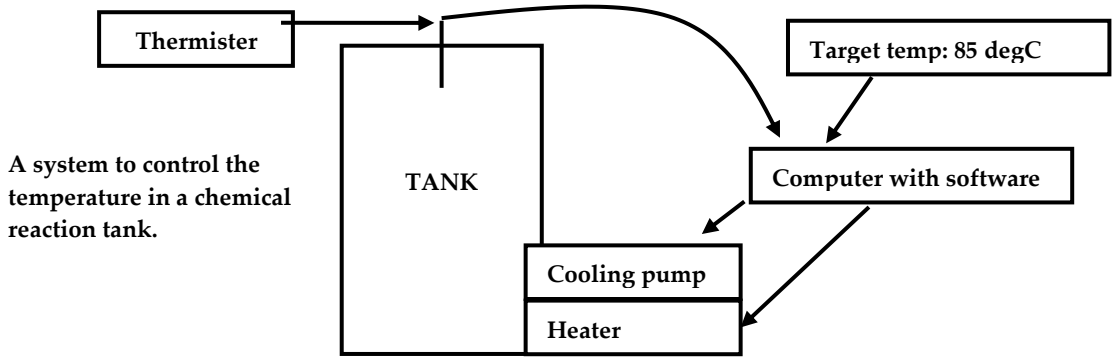
It is a little harder to argue that the air traffic control or the missile guidance systems are examples of 'transaction processing', however, because there isn't any interaction between a customer and an operator taking place! Transaction processing can also be called '**interactive processing'**, because there is an exchange of data (an interaction) between a customer and an operator.
To summarise, computer systems can be set up to process data in different ways or modes. Some systems can be designed to use batch processing. Other systems can be designed to use real-time processing, or transaction processing. The designer of any proposed system must look at the application and make a judgement about what to set up.

27.3 A process control system

A process control system is another example of a real-time system. It is a computer system that automatically monitors and reacts to changes in a process. There are many types of process that could be used as an example, from a greenhouse

temperature and humidity control system, to a chemical reaction monitoring system, to an air-conditioning system that maintains a nice cool temperature in a hot computer room.

Any processing system will have a number of common characteristics. These include inputs signals from sensors and transducers, some processing software (a program), some outputs, possibly using some **actuators** (an actuator is simply any device that causes movement and can be controlled by a computer - examples include computer-controlled motors and pumps) and some feedback, to allow a comparison of the current situation with a target situation.



1. When the system is started, the desired temperature is inputted into the system, perhaps using a keypad.
2. A signal is sent to the heater and the tank heats up.
3. The thermistor (a temperature measuring sensor) constantly sends a signal back to the computer. This is known as a 'feedback signal' because it is used to compare the current temperature with a target one.
4. The process control program in the computer checks the current temperature against the desired temperature.
5. If the desired temperature hasn't been reached yet, the heater is kept on. This process is continued until the desired temperature is reached or exceeded.
6. When it is reached or exceeded, the heater is switched off. If the tank ever becomes too hot, then a pump (an actuator) is switched on. This pumps cool water around the tank.
7. It is switched off when the temperature returns to an acceptable level.

27.4 Speed mismatch between humans and devices

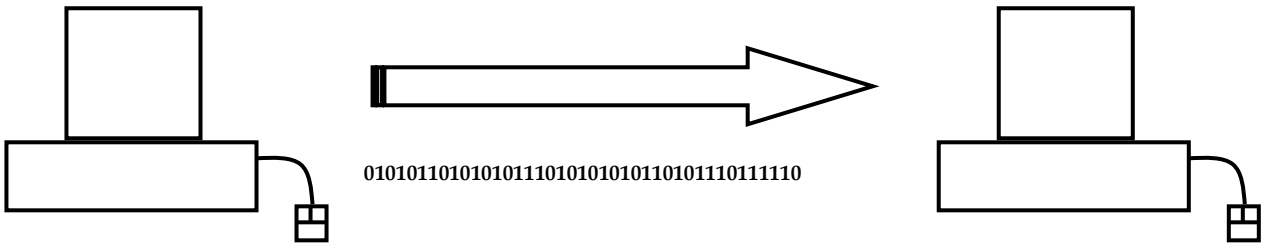
Computers work very quickly. Humans work very slowly by comparison. Consider any control system. A computer may be able to read in information from sensors, process it and then take appropriate actions in fractions of a second. A human, on the other hand, may take seconds to read information displayed on a screen, seconds to think about it (process it) and then may take more time taking appropriate action, giving the equipment instructions. For some situations, the speed mismatch between the computer and the human may be a serious problem. For example, a pilot may put the safety of passengers at risk if the on-board anti-collision systems of their plane were a manual system. They would first of all have to realise there was a problem then decide what to do about it and finally take appropriate action. It would be far safer (quicker) to rely on the computer making the necessary adjustments to the flight path immediately a problem was detected. Manufacturing lines need adjustments! Temperatures, pressures, flows of liquids and so on all need to be constantly adjusted to take into account a constantly changing set of circumstances on the production lines. If these changes were all left to operators to do, there may well be a lot of products that don't pass quality control! It would take operators time to realise a change was required to the temperature in the first place, for example. They would then spend time deciding what to do about it. Finally, they would need time to make the necessary adjustments. The time spent making adjustments would reduce the number of quality products produced. If control of all the variables were left to the computer, adjustments could be done automatically with minimal time delay and maximum production.

Q1. Define 'sensor'.
Q2. Do some research on the Internet. State the names of five sensors and the physical property they measure.
Q3. Define 'actuator'
Q4. Do some research on the Internet. State the names of two actuators describe an example of its use in each case.
Q5. What is the key characteristic of a real-time system?
Q6. Describe how the temperature in a greenhouse can be controlled using a control system.
Q7. What is meant by the term 'feedback'?
Q8. Do some research on the Internet. How are robots used in car factories? Watch a video of robots in a car factory.
Q9. What are the advantages of using robots in car factories compared to humans?
Q10. What are the disadvantages of using robots in car factories compared to humans?

Chapter 28 - File formats

28.1 Introduction

How information is stored in a computer is important. It is especially important when that information is transmitted to another place, perhaps across a LAN or over the Internet. We will look at how the speed of data transfer is measured. We then need to appreciate that a pattern of bits can be used to store different kinds of information including text, pictures, sounds and video. So far, we have used binary patterns to store different types of numbers and we know that binary patterns can be used to store machine codes. We need to also understand that there are many different ways of storing text, pictures, sound and video and we must consider the **purpose** of information each time we save it in one format or another.



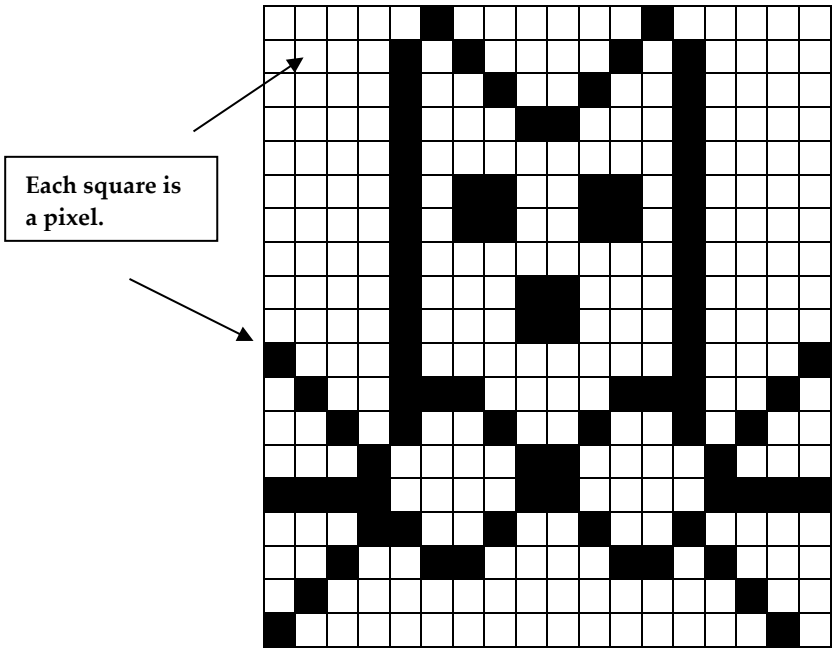
File size and data transmission speed are important across networks

28.2 Data transmission speed

The speed of data transmission is measured in 'bits per second'. Another way of saying 'bits per second' is **baud rate**. If you send one bit in a second, your transmission rate is 'one baud'. If you are sending 1200 bits per second, you are transmitting at 1200 baud. If you are sending 56K bits per second (56 thousand bits per second), you are transmitting at 56K baud. The faster the baud rate, the more data every second you are sending. The quicker you can send data, the quicker other users can gain access to a network to send their data. You may also save money if you are paying for network access by the second!

28.3 Displaying pictures using bitmaps

The higher your screen's **resolution**, the more detail about a picture you can display. Visual Display Unit (VDU) resolution is measured by the number of dots, or pixels, that make up a screen. He more pixels per unit area, the greater the detail.



It's difficult to display e.g. a picture of a cat if your screen is set at a very low resolution!!!

A picture on your VDU is made up of pixels. For example, a small square icon might be 80 pixels by 80 pixels. To completely define the picture, the computer needs to store information about every pixel, to say what colour each one is. The information about every pixel can be held in one bit, or in one byte, or in more than one byte, depending upon how much colour detail you need to store.

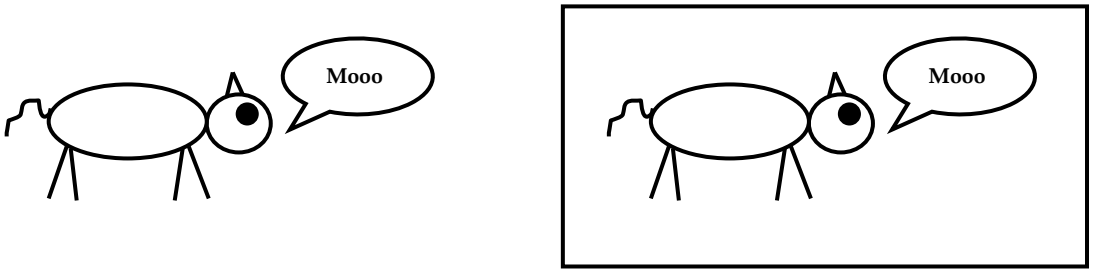
- If you want to display a black and white picture, each pixel that makes up the picture will need 1 bit to store whether that pixel is black or white.
- If you want to store a colour picture using up to 256 colours, then each pixel will need 1 byte to record the shade of colour of each pixel. (There are 256 different bit patterns in one byte.)
- If you want to store a colour picture using about 65000 colours, then each pixel will need 2 bytes to record the shade of colour of each pixel. (There are about 65000 different bit patterns in two bytes.)
- If you want to store a colour picture using about 16.7 million colours, then each pixel will need 3 bytes to record the shade of colour of each pixel. There are about 16.7 million different bit patterns in three bytes.

To recap, to store a picture, you need to know the resolution of the picture and the colour depth.

Pictures stored in the above manner are known as 'bitmaps'. The problem with bitmaps is that they take up a very large amount of storage space and therefore take a long time to load compared to e.g. vector graphics. To store one picture that is 300 by 300 pixels using any of 16.7 million shades of colour requires 300 x 300 x 3 = 270 Kbytes approximately. If you send this across the Internet, it will take a long time (in Internet terms) to be transmitted. When you click on a web page, you want it to appear as immediately as possible, pictures and all! Another problem with bitmaps is that they pixelate when you enlarge the image - you start to see each pixel individually rather than the whole picture. Because of the size of bitmaps, other ways of storing pictures have been developed. These use **compression techniques** to 'squash' the data needed to store a picture, thereby reducing the file size and making it quicker to transmit. Two of the most popular are GIFs and JPEGs.

28.4 The GIF format

Graphic Interchange Format pictures, or GIF pictures, use up to 256 colours. They are widely used for simple pictures, icons and line drawings on the Internet because they produce **very small** file sizes by using compression techniques (codecs). These are clever maths algorithms that take a file and make it smaller without reducing the information held in the file. Smaller files mean less transmission time on the Internet. GIFs can be also used to created animations known as **animated GIFs**. These are simply a number of GIFs run together in succession. Each one differs slightly from the previous one, like the frames of a film and gives the impression of an animation. GIFs can also be **transparent**. These are pictures where the subject sits on the page background rather than on a picture background. Transparent GIFs appear as part of the page.



Transparent GIFs appear as part of a page, without its own background or border.

28.5 The JPEG (JPG) format

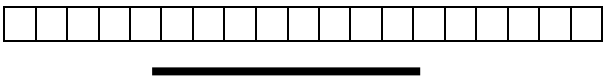
The Joint Photographic Expert Group picture format is used for photo-quality pictures that require lots of colours. GIFs can use up to 256 colours, which might sound a lot but means the true colour of high quality images will be lost (although as we have already said, GIFs are ideal for certain types of pictures such as icons and line drawings). JPEGs can use 16.7 million colours! The JPEG codec (see later in this section for information about codecs) is a compression algorithm that helps to keep files small although the files do suffer a little from 'blurring' where there are distinct colour boundaries. The price you pay for a smaller file is a very slight loss of quality.

28.6 Vector graphics

Bitmaps, GIFs and JPEGs can be used to store pictures. All three formats seek (in different ways) to store a picture bit by bit. There is a problem with this, however - these kinds of formats do not lend themselves to resizing! If you try to enlarge a picture in these formats, you end up viewing pictures that seem to be made up of jagged edges. This is true for any curve or line that is not horizontal or vertical. There are some applications where this is a major disadvantage, for example, in designing and drawing technical drawings, where designers and draftsmen need to zoom into the detail of a part of a drawing to give it more

detail. This simply cannot be done using the bitmap approach. Rather than storing a picture bit by bit, vector diagrams could be used. These store rules and equations about how lines and curves are constructed. If you zoom into a bitmap, you will eventually see the blocks of pixels. If you zoom into a vector drawing, the rules and equations are reapplied using the pixels available and the information about how much to zoom in by. In this way, a new drawing is reproduced that is clear and which does not show the blocks of pixels characterised by zooming into bitmaps. However, vector drawings are not good at showing a range of shades and tones - colouring is basic. This format would not be a good choice for colour photographs, for example.

Suppose you wanted to draw a line. To do this, you can either store information about every pixel that makes up the line (these are 'bitmap' pictures) or store a formula and some values that allow you to plot the line at any time (called 'vector' pictures).



y=mx+c

You can 'draw' a line by storing information about each pixel (bitmap graphics) or using a formula (vector graphics).

Vector graphics take up less memory than other formats but make greater demands on the processor. You know if a file is a vector drawing because it will have a certain extension associated with vector drawings. These include the Windows Metafile extension .WMF, the Encapsulated Postscript Vector graphics format (from Adobe Illustrator) .EPS and the Coral Draw Vector drawing file format .CDR.

28.7 Comma Separated Variable format (CSV)

A CSV file, or Comma Separated Variable (sometimes Comma Separated Value) file, is a file format used to **transfer data held in a table** from one application to another. The rows in the table are on separate lines in a CSV file whilst the fields in each row are separated from each other by commas. The data between each comma in a CSV file is made up of ASCII characters. They are most often used to transfer data between spreadsheets and databases.

If you **exported** a table of data about students on courses from a database and saved it as a CSV file, it would look like this:

A,Jones,1998,ICT
B,Smith,1999,Maths
G,Goodman,1998,English

- The data in each field in each record is separated by a comma.
- Each record (each row of data) is on a new line.

If you wanted to now get this data into another database or perhaps put it into a table in a spreadsheet then you would **import** the CSV file into the application. CSV files are important because they allow you to **transfer data** held in tables easily between applications. Suppose you had 10000 students in a university database whose details you needed to put into a new computerised system. You could type them in again but this would take a long time and you may make mistakes when you type in so much data. The alternative would be to export the data from the database as a CSV file and then import this data into the new system. Whenever you need data from somebody's database or spreadsheet, always ask them to give you it as a CSV file!

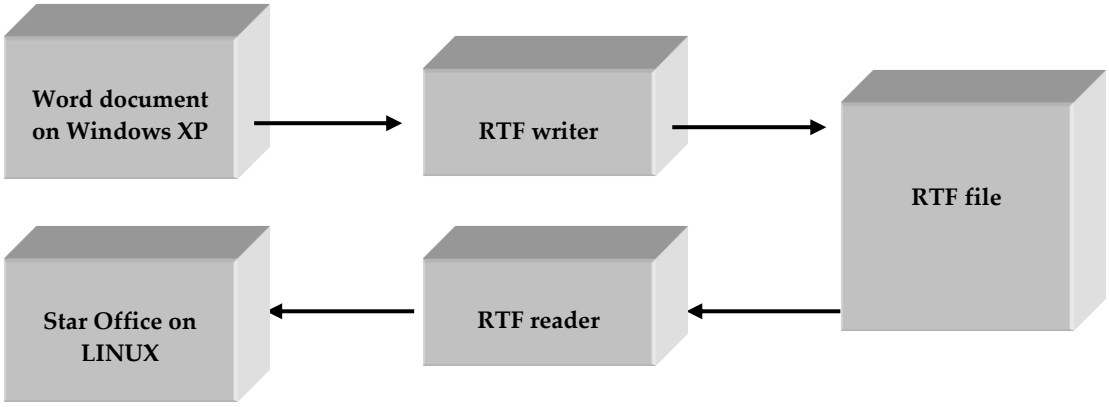
28.8 Text files

We have seen in a previous section that text is stored on PCs using ASCII codes. Each keyboard character is not stored as a bitmap as described above. The use of the ASCII character set is a very efficient way for a computer to store text. In a text file, you only get the ASCII characters that make up the data. There are no control codes with the text file so no information exists about how to format the data in the file. You won't be able to store information about which words are underlined, centre justified, in bold and so on. If you do not include control codes of any description then the text file will be very small indeed. In some circumstances, for example, when you need to send large amounts of data across a network, this may be a real benefit because it will take less time to transmit a smaller file than a bigger one.

28.9 Rich Text Format files (RTF)

RTF (Rich Text Format) files are used when you need to transfer files between computers that use different operating systems and different applications. For example, you could write a file in Word on Windows XP and save it in the RTF format. You could then email it as an attachment to somebody who uses Star Office in Linux or WordPerfect in Windows 3.1. They would be able to open the file because their application can read the RTF format.

An RTF file is made up of ASCII codes and a common set of formatting codes (common to all operating systems and applications). These are codes that tell some text to be bold, underlined, coloured red, in a new paragraph and so on. When you save a file written in Word, for example, as an RTF file, it is intercepted by a piece of software known as an **'RTF writer'**. This program processes the file, replacing Word's formatting controls with the common controls found in RTF. When you then open the file in, for example, Word Perfect, a piece of software known as the **'RTF reader'** intercepts the instruction. It converts the control codes used in the RTF file into the control codes used by Word Perfect and then displays the file. Interestingly, RTF files produced from files written in Word, for example, will produce very small files compared to the Word equivalent. This may be important if you need to transfer data across the Internet. If you saved the files in RTF format, they would be smaller and would therefore take less time to transfer. This is only true if the files do not contain pictures - the RTF format is not very efficient at shrinking files with pictures in!



RTF files can easily be transferred between computers using different operating systems and applications.

28.10 PDF files (Portable Document Format)

When you write documents or design pages in certain ways using certain types of applications, what you see on your computer screen at the time of designing or writing may sometimes not be **exactly** the same as what gets viewed on the Internet or what comes out of the printer! This is certainly true of HTML pages (pages written for the Internet using Hypertext Mark-up Language). HTML cannot reproduce the sophisticated formatting and presentation available in word processors, DTP software and graphics applications. If you designed a newsletter in a Desktop Publisher, for example, and then uploaded it to your web site, people trying to view it **may** view a document that has been distorted. The PDF format is a type of file that ensures that what you see on the screen ***is*** what gets viewed and is what gets printed out, regardless of what operating system, application or printer you happen to be using.

28.10.1 PDF files and the Internet

It is certainly very popular on the Internet because it allows the reliable viewing of documents that have been produced by an individual or organisation. When somebody produces a document that they want to display reliably, they first of all save it in, for example, Word format. They then use a piece of software (the most popular one is from the people who designed the format originally, called Adobe Acrobat, but there are many others - try searching the Internet for PDF) to convert the file from Word format into PDF format. The new PDF document can then be made available via the web site - it is uploaded to the web site. If somebody wants to view the document from the Internet, they must have Adobe Acrobat Reader (or some other PDF reader software) installed first - and it is **free** and easy to download and set up. Once they have done this then any PDF document will open automatically when it is clicked on.

28.10.2 PDF files and commercial printers

Graphics designers and writers increasingly send their finished work to their printers in PDF format. The printer can take these PDF documents and give them to their sophisticated computer-controlled printing machines. The designers and writers can be confident that what they have done will be faithfully reproduced by the printer because the PDF format has been used. This book was sent to the printers as a number of PDF files!

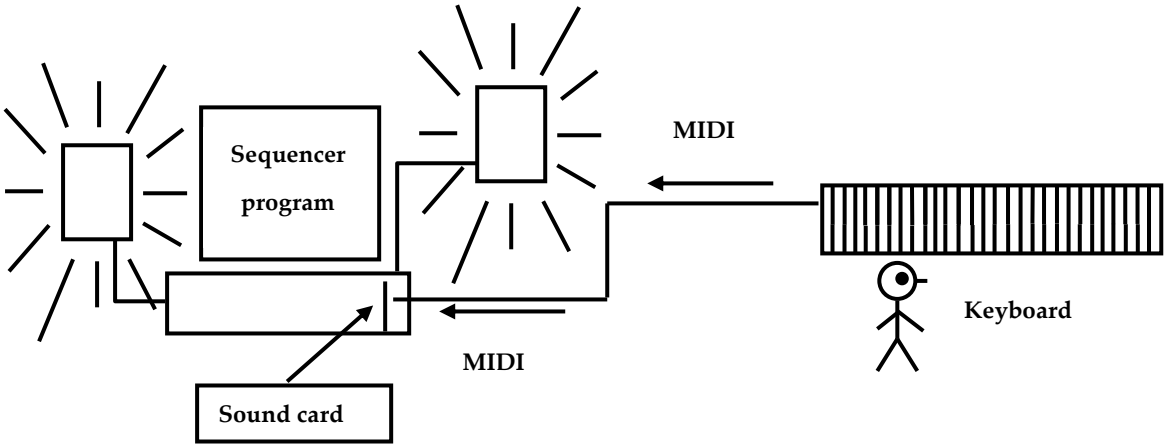
28.11 MIDI files for sound

MIDI (Musical Instrument Digital Interface) was originally designed to allow a range of electronic instruments to be connected to each other. The MIDI format specifies how connections between MIDI instruments will be done and specifies the codes that will be used and understood by those instruments. A MIDI music file will contain all the sound information (in digital form)

needed to play a song, such as the length of notes, the pitch, volume and special effects. MIDI files are sent using serial transmission.

28.11.1 Sequencers

A **sequencer** can be a piece of hardware or a software program. A sequencer software application can receive and store the notes a musician plays on a keyboard, for example. The sequencer can then be used to either tell a piece of equipment that can play MIDI files (such as a synthesiser) what to play, how to play it and when to start and finish, or can play back the song using the computer's sound system.



Using the MIDI interface to connect a keyboard to a computer.

Sequencers are powerful because they allow you to do clever things with songs once they are recorded. These include:

- speeding up or slowing down the performance of the song automatically
- changing the key the music is played in
- adding and removing special effects
- correcting any mistakes made by the musician
- combining different instruments into one file.

28.12 MP3

MP3 is an acronym for MPEG Audio Layer 3. This codec is a format that was created to create compressed music files that could then be quickly downloaded from the Internet. It has now achieved widespread use globally as millions of tracks are downloaded every day. MP3 files provide music that to all but the purists are indistinguishable from CD audio files. The main advantage of MP3 files however is that you can fit about 120 songs in MP3 format on to one CDROM!

28.13 WAV

WAV files are another format in widespread use. It is a popular format because it was developed by Windows and is therefore omnipresent on computers that use the Windows operating system. These files, however, are not highly compressed (like MP3) and can therefore lead to very large files. They are not suitable for large music files downloaded over the Internet but have many uses within Windows-based applications.

28.14 RA

RA is yet another codec that produces compressed sound files. The Real Audio format was around before the MP3 format came along. It has an excellent property that is employed by many radio stations - it can be used to 'stream' music over the Internet. This simply means it can be used to send music over the Internet, using a buffer and a Real Audio player, to play the sound as it is downloading.

28.15 Ogg Vorbis

This is another codec used to compress sound files. Compared to file compression using MP3, the files are smaller and higher quality. The other advantage of this format is that it is 'open source'. That means that anyone can write applications for and use this codec without paying any royalties. This is not so with, for example, the MP3 format, which is owned by a company.

28.16 Movie files: MPEG and AVI

There are a range of file formats that are used to store movies. One of the most common is the MPEG (Motion Pictures Experts Group) format. This is a codec (see section on codecs) that compresses both video and audio to create a movie format. The AVI format (Audio Video Interleave) is another format in widespread use. It saves a movie in much the same way as a film, frame by frame. It uses compression techniques to ensure the file is as small as possible. These files can then be played back using a movie player designed for that purpose, such as Media Player.

28.17 Pirates

On a general note, when you copied a song from a record onto a tape or copied a film on video to another video in the olden days, the quality always degenerated slightly. If you then used the copies to make further copies, the quality degenerated even further until eventually the copies were all but useless. Copying digital material is an entirely different story. There is no degradation when you copy bit patterns. For this reason, copies are always perfect. This has got the music, software and film industry very worried. It is all too easy now for songs, films and software to be copied illegally across the Internet and thereby deny companies of their royalties. The problem is becoming worse as broadband use widens.

28.18 Codecs

Any data file can be very large indeed. This can be a problem, especially on the Internet, where the larger the file the longer the download time. It would be useful if files could somehow be 'squashed' to make them smaller. We need to find a way to **compress** data files, be they text files, audio files, graphics files or video files. Fortunately, there are many standards around now that take a raw file and compress it to make it smaller. We have just been discussing many of these! For example, we discussed GIF files, JPEG, AVI and MPEG files. All of these are essentially a set of maths rules (algorithms) that can be applied to a raw data file with the result that the file is compressed. There may be some payback, however! There may be a slight loss in the quality of the original file, for example. And don't forget that if you compress a file then it needs to be decompressed before it can be used again. If you don't have the correct software to decompress a file then you cannot access it!

The algorithms used to compress text, graphics, sound and video data files are known as **codecs**. The word codec stands for **compressor / decompressor**. Each codec has its own characteristics and selecting the right codec is important. For example, some codecs, known as **lossless codecs**, will take a video or graphics file and compress it in a way that preserves the image quality. This inevitably leads to bigger files than a file that is compressed using the other type of codec known as a **lossy codec**. These types of codec concentrate on getting the file as small as possible, although this leads to a slight loss in picture quality. A JPEG file is an example of a format that uses a lossy codec; it produces a smaller file but the picture quality very slightly degrades.

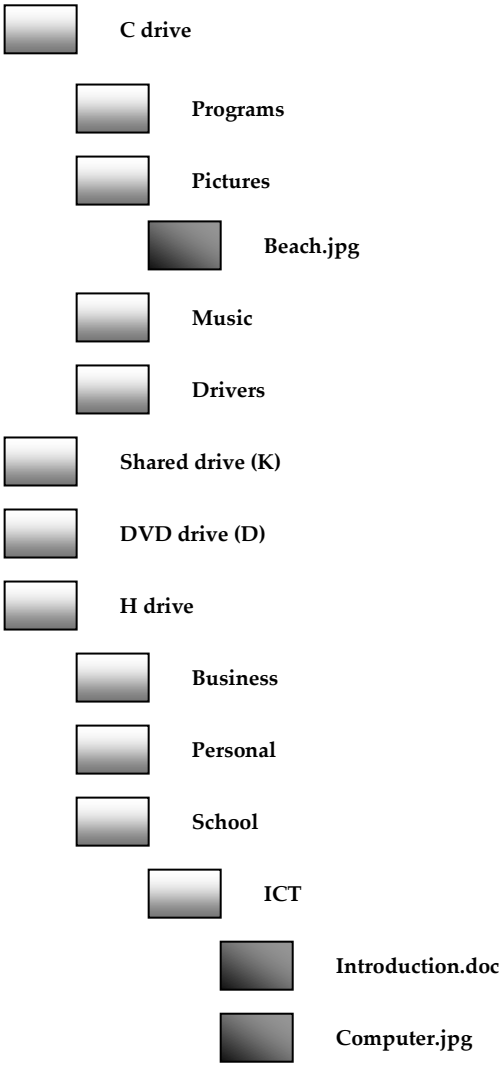
Consider a file that has been compressed on Computer A using a particular codec and then sent to somebody else over the Internet. That person is using their computer, Computer B. To view that compressed file on Computer B, the file must be correctly decompressed. This can only happen if Computer B has the **same** codec installed that was used to compress the file in the first place on Computer A. If the correct codec is not present then the file cannot be decompressed and viewed and an error message will be issued. The person on Computer B would have to identify the codec, find it on the Internet and install it.

Q1. What is meant by 'baud rate' and what is it used to measure?
Q2. What is meant by a 'high resolution picture' and a 'low resolution picture'?
Q3. What two pieces of information are always needed when a picture is to be stored?
Q4. What is meant by 'metadata'? Give examples to illustrate your understanding.
Q5. How does a vector picture differ from a bitmap?
Q6. Why are pdf files commonly used to display documents on the Internet?
Q7. What is a csv file used for?
Q8. What is a MIDI interface used for?
Q9. Do some research on the Internet. Write down some examples of lossy file formats.
Q10. Do some research on the Internet. Write down some examples of lossless file formats.

Chapter 29 – File management

29.1 Introduction

We need to be able to organise and find our files. Operating systems such as Windows have adopted a hierarchical directory structure to enable users to do this. A typical hierarchical structure is shown below.



This structure is made up of directories (often referred to as 'folders'), subdirectories and files, all stored on an actual, physical disk drive. This physical disk is usually known as the 'C drive'. The C drive shown here has four directories. One of these, Pictures, has a file in it called Beach.jpg.

This computer also has a shared drive, called the K drive. It can be seen by all users and is used to share files and folders. The other drive this computer has got is a DVD drive. This was allocated the letter D when the computer was set up.

Although this computer has just one physical drive, it is often convenient for the user to pretend that they have more. The user can take the physical drive and using some software or the operating system software, split it up (or partition it) into sections. Each section, called a logical drive, is allocated its own drive letter.

The user might now use one logical drive to store all their data files, for example. This would make backing up their personal files very easy - they just back up the whole of that one drive. They could set up and use another logical drive to store all of their music collection. A different one could be used just for movies, and so on. Partitioning a drive is another tool that helps users organise their programs and data files.

The H drive shown here is an example of a logical drive. This is also known as a virtual drive because it is not a real one - it is a pretend, or virtual, one. It is part of the physical drive known as the C drive but as far as the user is concerned, it looks like there is another drive!

The H drive has three directories in it, called Business, Personal and School. The School directory contains a subdirectory called ICT. A subdirectory is simply a directory that is inside another directory. The ICT subdirectory holds two files, called Introduction.doc and Computer.jpg.

Each drive, for example, the C drive and the H drive, is also known as a root directory. It is the directory that sits at the top of the hierarchical data structure and can be a physical drive or a virtual one.

If you wanted to describe where a particular file is on a computer, you have to specify its path. This tells you exactly what root directory to use and what directories and subdirectories you have to go through to get to any particular file. For example, suppose you wanted to know the path of the file called Introduction.doc, shown above. The path for this is:

H: \ School \ ICT \ Introduction.doc

This tells a user that the file is stored in the root directory H, in a directory called School and in a subdirectory called ICT. It is important to be able to refer to the path of any particular **file** because e.g. you may need to tell someone else where it is. You may also have two files with the same name but stored in different places. The paths would distinguish between them.

29.2 Information about files (metadata)

When files are stored (and logical drives, directories and subdirectories), it is not just the file itself that gets stored but also information about the file. If you go to the file manager of whichever operating system you are using (Windows Explorer in the

Windows operating system) you should be able to see the extra information , called 'metadata' held about the file. To do this in Windows, you right click on the file and then click on Properties. You can see that it tells you (amongst other information):

- the type of file it is
- what program is used to open this file when you double click on it
- the path of the file
- the size of the file
- when the file was created, last modified and last accessed by the user
- Whether it is read only, or whether it can be written to as well (the access rights for a file).

This extra information can be quite useful to the user. For example, if you cannot save a document onto a pen drive, perhaps it is because it is too big and you don't have enough space on the pen drive! You can firstly check the file size using the method above and then you can check how much space is on the drive by right clicking on it and going to Properties. This extra information is sometimes referred to as 'overheads'.

Q1. What is the hard drive used for on a typical home computer?
Q2. What drive letter is the hard drive usually allocated?
Q3. What is another name for a 'directory'?
Q4. What is meant by a 'physical drive'?
Q5. What is meant by a 'logical drive'?
Q6. What are you actually doing when you 'partition' a physical drive?
Q7. What is a subdirectory?
Q8. What is meant by the 'path' of a file?
Q9. Give three examples of a file's metadata.
Q10. What is meant by a file's 'access rights'?

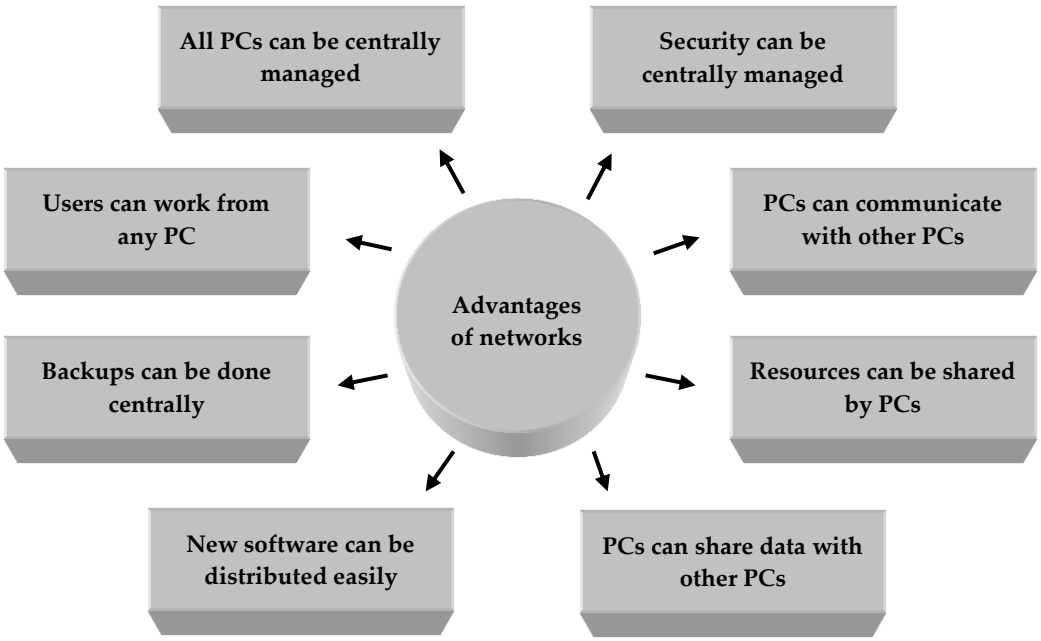
Chapter 30 - An introduction to networks

30.1 Introduction

If all the computers in a room are standalone, that means they are not connected to each other. This has a number of implications. They can't send messages to each other. They can't share a file or share a piece of equipment like a printer easily. They can't work on one particular computer one day, and then on a different one the next day and expect to be able to get back easily files that were saved on the first computer. If computers were connected together in some way then they would be able to do all of the above. Computers that are connected together are known as a 'network'.

30.2 Advantages and disadvantages of networks compared to standalone machines

The advantages of a network compared to standalone machines are:



Advantages of networks.

- Computers that are connected together can communicate with each other. Standalone machines can't.
- Users on a network can easily share resources such as printers, scanners and modems. This means that a company doesn't have to buy so much equipment and users do not waste time moving files to different machines so they can use a piece of equipment that is only available on a particular PC.
- Users on a network can share data. For example, a team might be working on a project. They can all access and work on the same files easily without having first to back them up on pen drives, for example, and then transport them to the next team member and then reload the files.
- New software need only be added once. It can then be distributed to all other PCs automatically.
- When data files are backed up they only need to be backed-up once, centrally, at the server. With standalone machines, you would have to back up the files on every machine. This would be very time-consuming!
- Users can retrieve and work on files from any machine on a network. If one machine is being used or breaks down, you simply move to a different machine!
- Networks can be managed. This means a network manager is able through the network software to control who can access the network, when they can access it, and what files and software and hardware they are allowed to use. An audit trail of each user can be built up so that projects can be more efficiently costed, for example, by keeping track of what printouts are done in connection with what project. Users, knowing that their activities are being monitored, will also be encouraged to stay on-task and not waste the organisation's time!

- Security can be centrally managed by the network manager. They can add patches centrally, ensure virus patterns are up-to-date centrally and so on.

BUT

- It costs more money to build a network than it does standalone machines. This is because you have to buy network cards, interconnections, a server and a Network Operating System.
- There is an additional support cost that is not insignificant. Networks are more complicated than sets of standalone machines. They need specialist knowledge to set them up and maintain them as well as time to maintain them. To do this, you would need to employ somebody with network management skills.
- Networks have machines in different locations, each of which potentially could be used to gain access to the server's hard disk and the data stored on it. This is a security headache and the need to protect data on a network from hackers adds an extra degree of complexity to a network that doesn't exist with standalone machines.
- If the server on a client-server network goes down, the whole network will be out of action. It is also possible for one faulty machine on a network to cause other machines on the network to stop working. In addition, if the cables on a network fail, problems on either individual machines or on the whole network may occur.

30.3 LANs and WANs

Broadly, networks can be split into two categories, LANs and WANs. **Local Area Networks** are networks that are made up of computers connected up to each other and are geographically close to each other, for example, in a room, or in a building, or in a number of close buildings. **Wide Area Networks**, on the other hand, are made up of computers connected up to each other over a wide geographical area. The Internet is a good example of a WAN. Companies who have offices scattered over the country may have a Wide Area Network.

30.4 Network hardware and software

If you had a room of standalone machines, what would you need to go out and buy to make a network?

30.4.1 Interconnections

To link the computers together in a network, you need to buy some interconnections. This might simply be some cable but you could also use a wireless technology to connect the computers together. With any interconnections, you need to think about:

- **Cost**. How much will the interconnections cost?
- **Bandwidth**. What bandwidth will users of the network need? Bandwidth is the term used to describe the amount of data transmitted across a network through a communications channel, such as a cable. Transmitting data such as music or video is a high bandwidth activity. Transmitting word-processed documents is a low bandwidth activity. Different interconnections are able to carry different bandwidths. A network designer needs to identify what users will be using the network for and then select an appropriate connection method. If users will be using video conferencing, then a coaxial cable (like a television cable) or a fibre optic cable, for example, could be selected because these can handle a high bandwidth. A wireless interconnection might be selected for a school, so long as the users just wanted to use basic applications such as word processing, databases and so on. If the users wanted good Internet access, the wireless solution might be discarded in favour of a cable solution for bandwidth reasons. That said, the bandwidth for wireless connections has improved considerably recently.
- **Cable routes**. If bandwidth isn't an issue, a wireless technology might be preferred in situations where physically laying cables might be a problem, for example, where lots of buildings needed to be connected and you don’t want to dig up roads! Distances also need to be considered. Different interconnections can send signals different distances. It may be necessary to buy some extra equipment called **repeaters** to boost signals if transmission distances are too large.
- **Electrical interference**. Cables and wireless technologies are subject to electrical interference from any electrical devices because the data itself is electrical in nature. This is not so for fibre optics, which use light pulses to send data, not electrical signals. In locations where there is a large possibility of electrical interference, fibre optic technology could be considered.
- **Security**. Wireless technology is insecure. It is relatively easy with the right equipment to 'capture' the data sent in this way. Cables are more secure but can also be tapped into although in both cases, encryption technology should be used if data is that sensitive. Fibre optic technology is difficult to tap into - the light signals degenerate if someone tries to break into the signal.

30.4.2 Network cards / network adapters

Each computer that is going to be part of a network needs to be actually connected up to the network. To do this, you will need to buy and fit a network card into an expansion slot in each computer. The network card inside a computer provides:

- a means of connecting a computer to a network
- a way of splitting up data that you want to send across a network and then getting that data actually onto the network successfully
- a way of collecting 'packets' of data addressed to that particular computer from the network. Each network card has a unique identifier called a 'MAC address' that enables this to happen.

A connection can be physically made using a cable from the network card to the main network cable. Networks could also make use of wireless network cards. These are installed in computers but do not physically connect to the network. Instead, they connect using radio signals. Another term in use is 'network adapter'. The meaning of this term has broadened to mean any card that connects a computer to a network. For example, you could talk about the thin little card that allows you to connect a laptop to a network (known as a **PCMCIA card**) as a network adapter, or a card that allows a wireless connection as a network adapter, or a basic network card as a network adapter.

30.4.3 Server

Computers on a client-server network are known as 'clients'. The clients need a very high-specification computer to organise and manage the communications around the network and to take responsibility for controlling access to the files held centrally. This high-specification computer is known as a 'server', and for a quality client-server network, you will need to go out and buy a server. Although a server is just like a normal computer it does have a huge hard disk compared to the clients. This is because it needs to store everyone's data files as well as store applications centrally. The server will also have much more RAM than the clients and a very fast processor, to enable it to work as quickly as possible. The server will usually have some kind of automated backup hardware, to ensure that users' files are backed-up regularly. There are other kinds of servers that provide extra functionality and efficiency and these are dealt with in a later chapter.

30.4.4 Network Operating System (NOS)

The server on its own is not enough. You need to also buy some software to allow the server to organise and manage the communications around the network and to allow the server to take responsibility for controlling access to the files held on its hard disk. This software is known as the 'Network Operating System', or NOS. Typically, the NOS controls access to the facilities on the network by making a user log in. They do this by providing each user with a user ID and a password, which they then have to type in when they want to use the facilities of the network.

30.4.5 An Internet connection

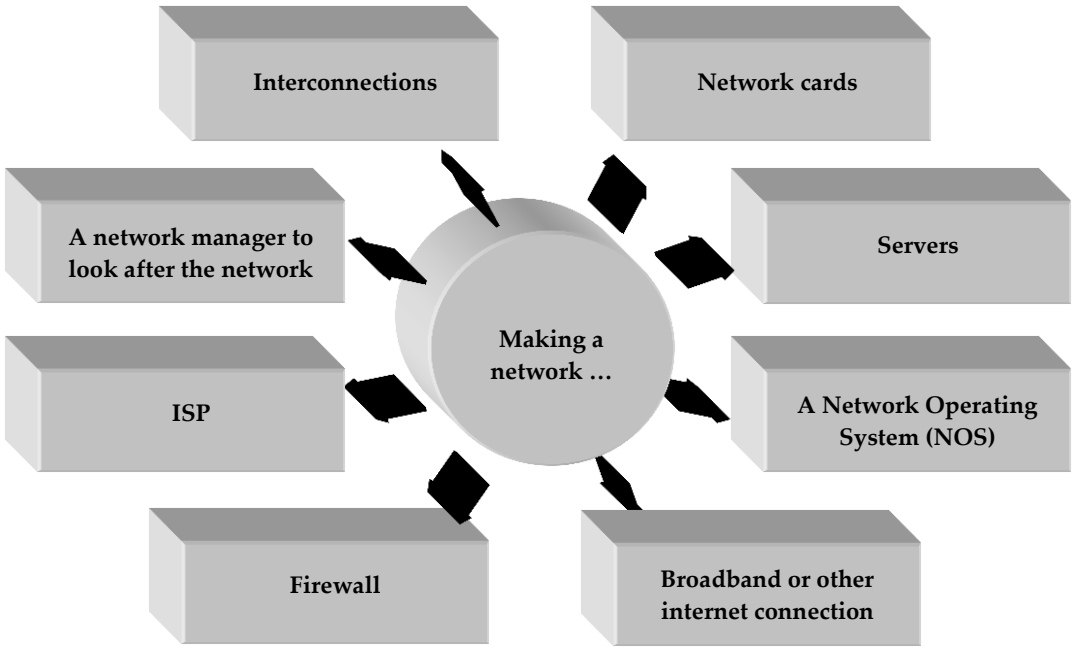
If you want your LAN to communicate with other networks via the Internet (so your LAN is part of a WAN), then you need some further equipment. Connections to the Internet can be done in a number of ways. One increasingly out-dated way is to set up a 'dial-up connection'. The network is connected to a modem and the modem is connected to a phone line. Then, when you want to send an email or use the WWW, the network dials up a phone number and makes a connection to the ISP (see below). You can then send or receive data. When you have finished, you disconnect the connection. The modem turns the digital signals of the LAN into the 'analogue' signals of the phone line, and reconverts analogue signals received back into digital signals for the network. Modems have limited bandwidth and the connection to the Internet is not permanent. In addition, if you are using an office phone line for your dial-up connection then that line cannot be used for anything else, such as making a phone call or sending a FAX. Of course, you could always go to the expense of having another phone line put in.

30.4.6 Broadband

Modems used in dial-up connections work at speeds of up to 56 Kbytes. This may be fine for light Internet use. However, a broadband connection gives much higher bandwidths, which means you can download more data every second – ideal if you download or upload lots of multimedia files or a lot of people are using the same connection, for example, in an office. A broadband connection is also always on – you don't have to 'dial-up' a connection when you want to use the Internet. Because it is always on, you should ensure you have a firewall installed. A firewall is a piece of software which stops unauthorised external users gaining access to your network. Another advantage of broadband is that you can use the broadband connection to access the Internet at the same time as sending or receiving phone calls. You can't do this with a dial-up connection. It is either one or the other!

30.4.7 ISP account

Of course, if you want to connect a network to the Internet then you need an Internet Service Provider (ISP) account. This is a company that acts as a door to the Internet. You sign up with an ISP for a fixed monthly or annual fee and they then allow you to access the Internet through them, either using your dial-up or your broadband connection.

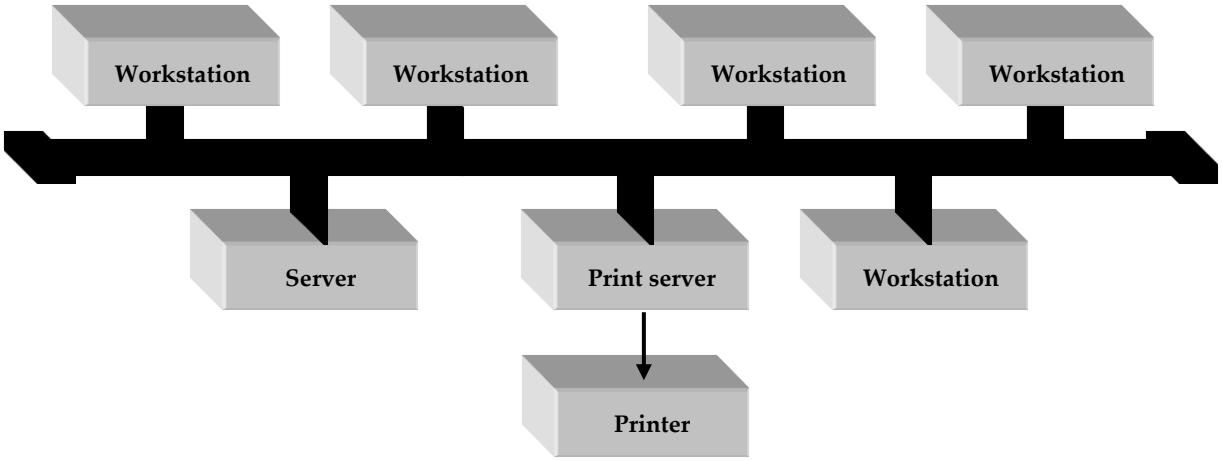


Moving from a set of standalone PCs to a network.

30.5 Network topologies

Computers that are going to be connected together can be connected in different ways. The way that the computers on a network are connected together is known as the **topology** of the network.

30.5.1 Bus networks

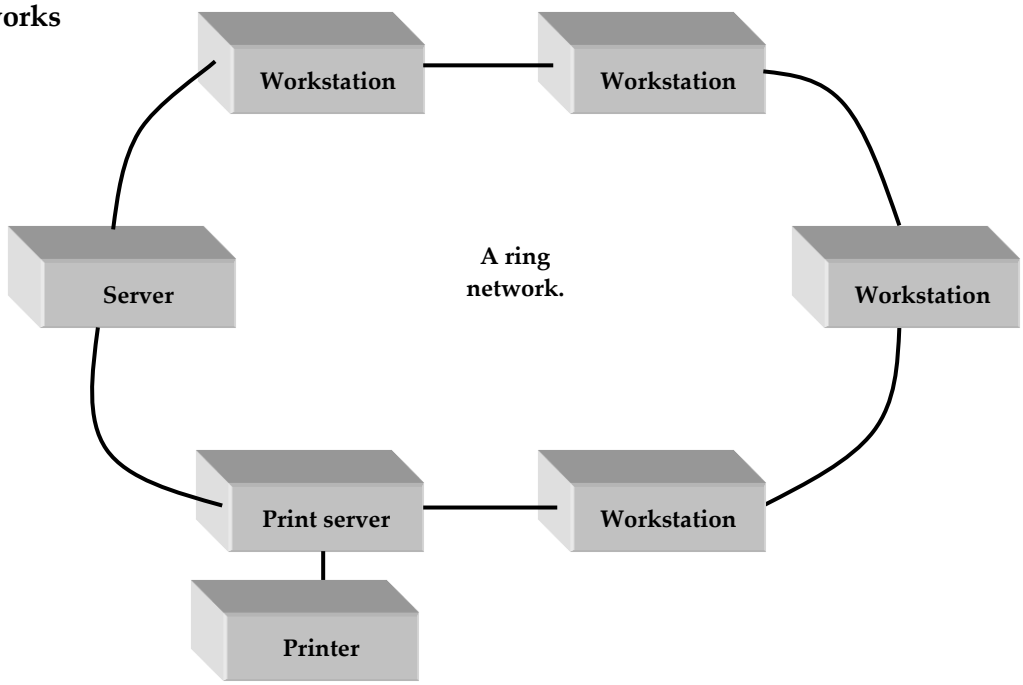


A bus network.

- A bus network is one where the stations are connected to a main communications cable, called a bus.
- The cable is often a twisted pair cable or coaxial cable.
- The bus has 'terminators' at the end of the cable.
- This set-up is very simple and very inexpensive to implement compared to alternative topologies because very little cable is needed compared to a star network, for example.
- If the main cable breaks, however, then the whole network stops working and it can be difficult to pinpoint where exactly the fault is.

Bus networks have a problem when two workstations want to transmit a message at the same time on the communications cable. The cable can only carry one message at a time. If the cable isn't busy then a station can put its message on the cable. If there is a message on the cable then it waits until it is free. This simple method **avoids** most '**collisions**'. This is when two messages are put on the network cable at exactly the same time, which results in corrupted messages. Collisions are possible because two stations can see at exactly the same time that the cable is free and at exactly that same time, can put their own messages on the cable! Fortunately, stations can also **detect** when these collisions have occurred. When they do detect a collision, the stations involved are asked to re-send their messages.

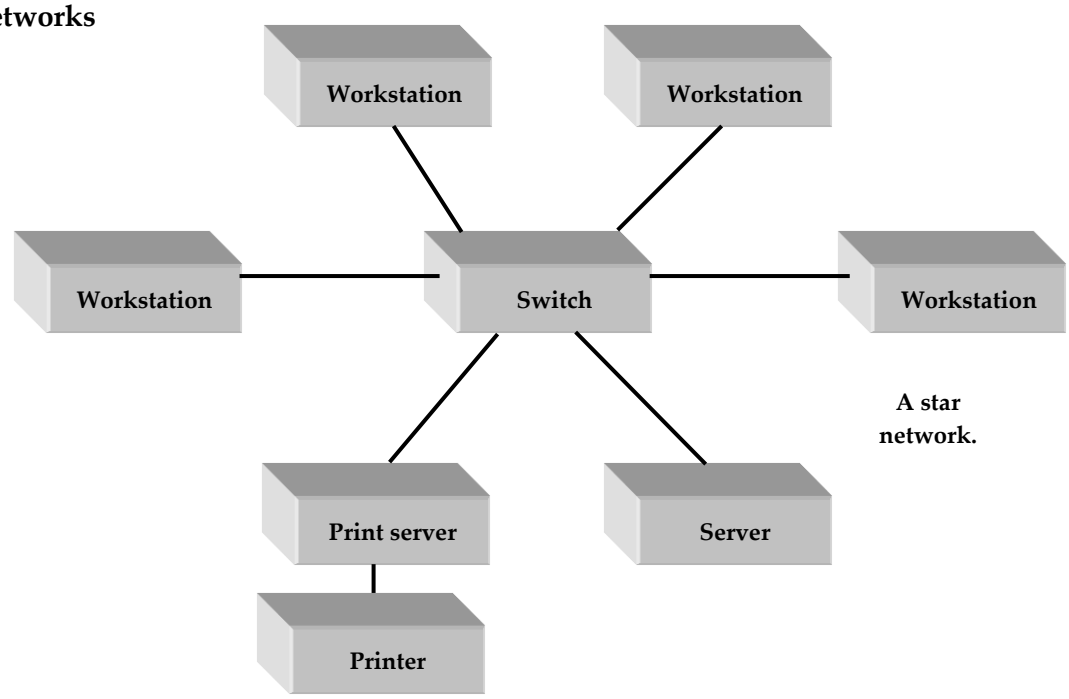
30.5.2 Ring networks



A ring network.

- A ring network is like a bus network with the ends connected together!
- This gives a better performance than a bus network, but is not as fast as a star network.
- You need a ring for the network to work properly so if the cable breaks, the network stops functioning.
- It's also difficult to find the error if the cable breaks, compared to a star network.
- Packets of data need to travel through each workstation between the sending and receiving workstations. This slows down data transmission.

30.5.3 Star networks

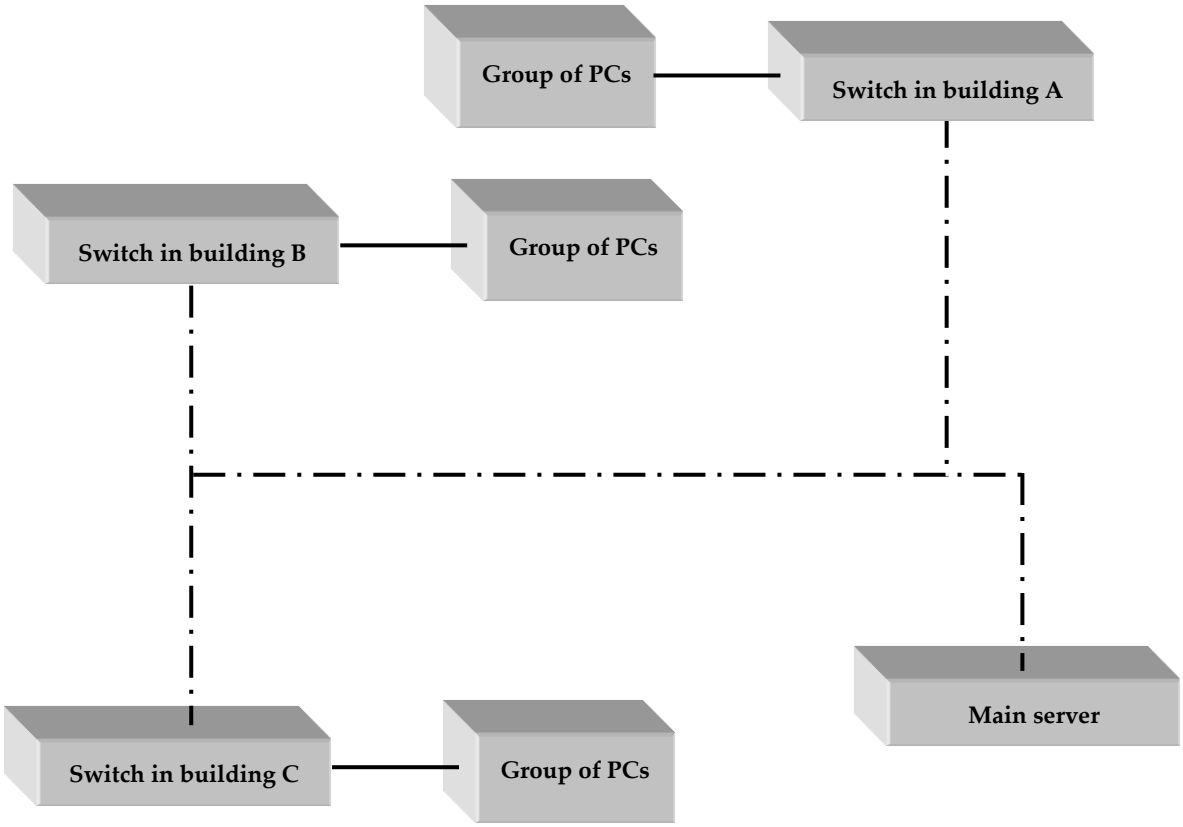


A star network.

- A star network is one in which the workstations are connected directly to a hub or switch, and this is then connected directly to a server by one cable.
- This kind of network is secure because all of the communications go from one station to the target station via the server. Other stations cannot 'grab' the communication. Security can be managed from one central point - the server.
- This kind of network takes up a lot of cable compared to a bus network and this can be expensive.
- If a fault occurs that is not central, then the rest of the network will carry on working.
- A switch provides a means for all of the cables to be physically connected to one point. The switch itself is connected to the server by one cable.
- The switch is also responsible for 'routing' packets of information to and from the server.
- The server goes from one station to the next in turn (via the switch), checking to see if each workstation needs to do something. This is known as 'polling'.
- If the server finds a station that needs to send a message, for example, then it will deal with it.

30.6 Backbone

The backbone of a network refers to the main high-speed communication channel that links smaller sub-networks of a big main network. Backbones often use fibre optics because fibre optic links provide a very high data transmission rate. Fibre optic cables would typically run the length of a building or would connect buildings close together, terminating at switches. Clusters of computers would then be connected to the switches.



The dot-dash line is the fibre optic backbone of this network.

Q1. Define 'network'.
Q2. What is the difference between a LAN and a WAN?
Q3. What is the purpose of a network card?
Q4. What is meant by bandwidth?
Q5. Do some research on the Internet. What is the purpose of a 'proxy server'?
Q6. Do some research on the Internet. What are the main jobs of the Network Operating System?
Q7. Sketch out a ring network, a bus network and a star network.
Q8. What does a router do?
Q9. Do some research on the Internet. What is a gateway used for?
Q10. What is meant by the backbone of a network?