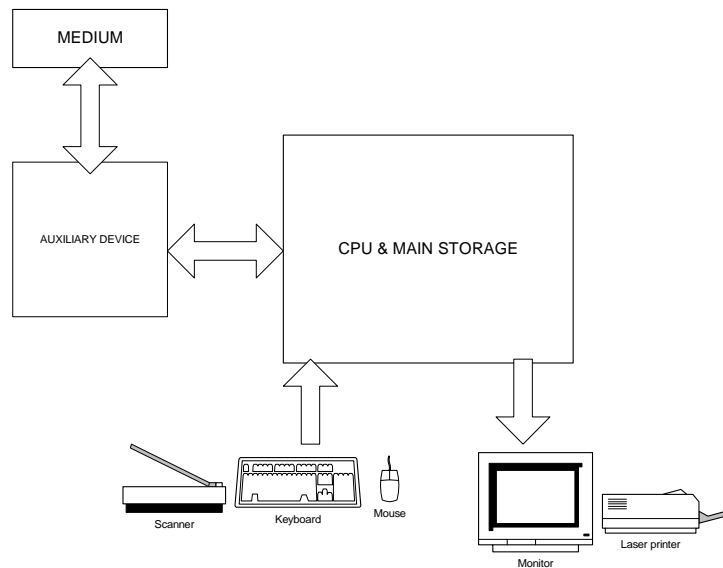


Lesson
6

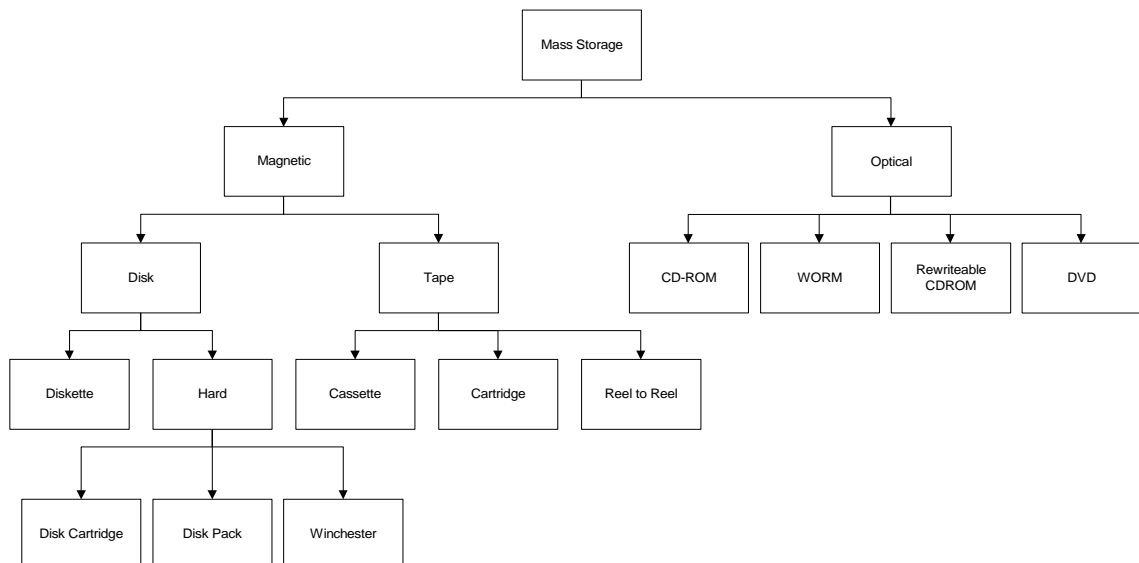
Auxiliary Storage and File Organisation

Auxiliary storage refers to storage other than the main storage (e.g. magnetic tape or direct access devices). Other names given to auxiliary storage are secondary and backing storage. Auxiliary storage devices have been covered in the previous unit, this section will concentrate on the mediums (e.g. magnetic disk, floppy disk, cd etc) and how data/information is organised.

Generally, data and information is read from the medium by the auxiliary storage device and passed to the computers main storage.



Following are the categories each storage medium falls under:-

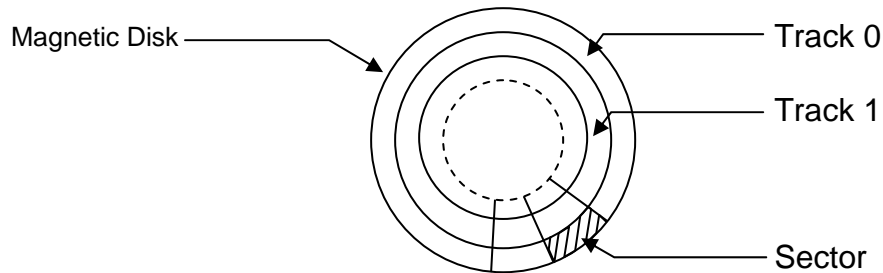


Magnetic Disk

Magnetic disks range in size from the old 3" (not used nowa days) to the 3 ½". Hard disks are usually found to be fixed within computers and are faster and have larger storage capacities (10Mb – 100Mb). Hard disks are used to store anything which is permanently required on line, for example Windows '98, MSOffice, and user documents.

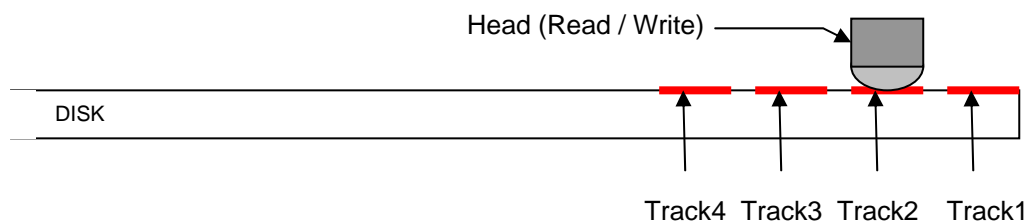
Exchangeable Hard Disks are removeable from the drive and replaced by another. They are used for storage of files not required to be always on-line (Backup).

The amount of data that can be stored on a disk depends on how many tracks there are.

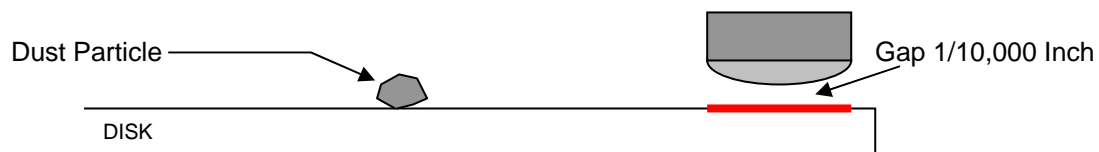


Head Positioning

To read data from a disk it is essential that the head is positioned exactly over the required track, else the output from the head may be too low to be read. Obviously the closer the tracks are to each other the more critical head positioning becomes.

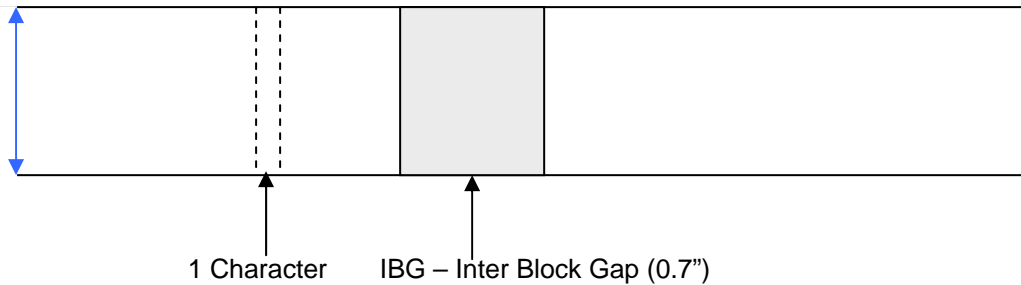


More sophisticated disk drives employ a system known as servo positioning to obtain exact head positioning. The head is moved to its appropriate position by a mechanical drive. An electronic system then measures the output from the head and moves the head slightly until its output is maximum which occurs when the head is exactly over the track.



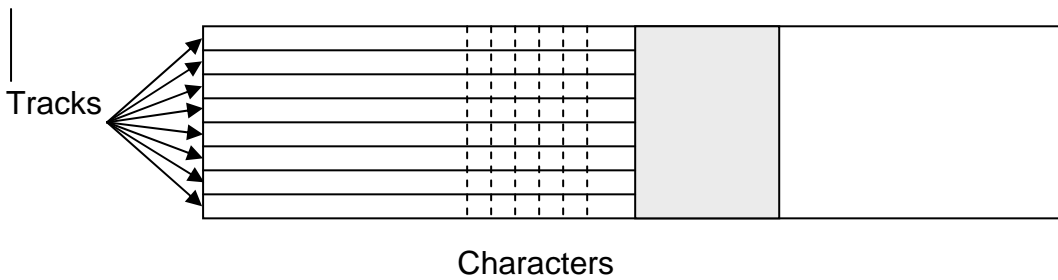
Magnetic Tape

With magnetic tapes, each character is recorded across the width of the tape.

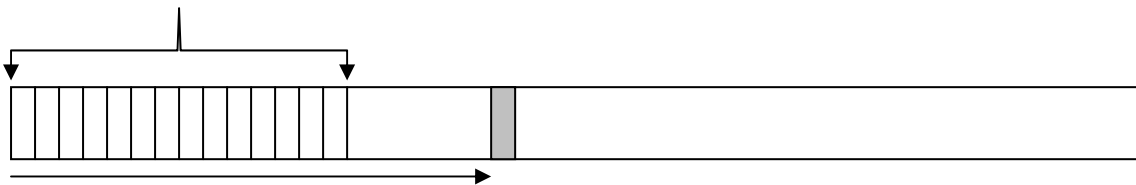


Recording density is measured in bits/inch or characters per inch. The IBG represents the end of each record or block. Hence only a fraction of the tape holds data. The purpose of the IBG is so that the tape can accelerate and decelerate to the appropriate data block and to signify the start and end of each block.

A character consists of 1's and 0's and is recorded on individual tracks running across the magnetic tape :-



Magnetic tapes can only be read serially, hence all records must be read irrespective if those records are required. For example to locate a record stored half way on the magnetic tape will entail all records stored prior to that being read!



Read all records until reaching required record

File Organization - Introduction

All organizations – large or small depend on reliable, up-to-date information. As businesses grow in size and complexity, the collection, storage and retrieval of data of information needs to be formalised.

Example

A small company employing around 5 people would not normally use a computer to do its payroll every month. However if it had 5000 employees or even 50 it would be no doubt that computerisation would be a necessity. The more employees a company has, the longer it will take to calculate and process salaries by manual methods – hence the need of computerisation.

When calculating an employee salary, it is necessary to know

- Hours Worked
- Tax
- Medical benefits
- Pension etc

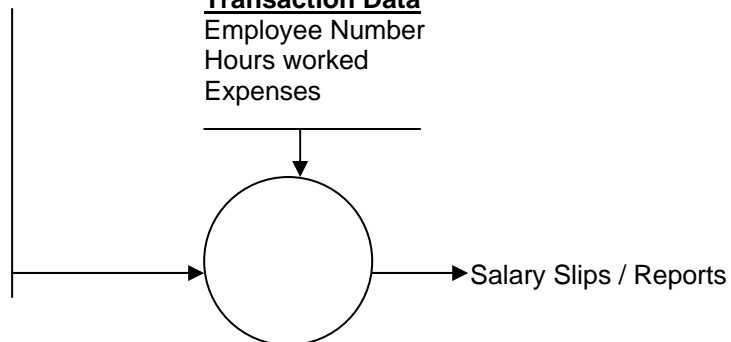
Data required to calculate the payroll for the employee :-

Master File

Total Hours worked to date
Rate of pay
Employee Number
Tax Code
Name, Surname, Address
Bank Details
Pension
Medical Insurance
Other deductions

Transaction Data

Employee Number
Hours worked
Expenses



Master File will already contained previous details of the employee. New data will need to be input every month (Employee Number, Hours Worked etc) into the transaction file. Subsequently the Master File and Transaction file are both used to generate and process the wages. The Master File will be updated with reference to Total Hours worked to date or Salaray paid to date etc.

Looking at this process – it can be seen that this is a repetitive task requiring a number of items of information and some relatively simple calculations. Nevertheless accuracy is a key necessity and computerisation enables this to be carried out effectively.

Benefits of Computerisation

- Speed of processing
- Capability of manipulating and handling large amounts of data
- Storage of large amounts of data
- Accuracy
- Improved presentation (printouts/reports/ wage slips)
- Access to additional information for management

Disadvantages of Computerisation

- Staff training
- Systems Support, failure of computer systems
- Staff redundancies
- Consistent input of accurate data
- Accurate conversion of a data file from old paper based records.

Not all applications need to be computerised, remember manual methods of data processing still exist when

The system may not be large enough to justify the use of a computerised system. Although a system may be computerised a number of procedures could still be carried out manually due to the cost benefits.

File Types

The file is the heart of any business data processing system. Information cannot be stored randomly. It needs to be stored in a logical way and files provide a convenient way of doing this. We begin by looking at the different types of files used in a computerised business system.

Transaction Files	Contains source data to access master file. Used to amend or update master files.
Master Files	Contains data of a permanent nature. The value can change in a file, they are not permanent. Also contains additional information that is regularly updated (tax paid to date).
Backup Files	Copies of Master and Transaction files are held for security purposes. Usually held on tape. Kept in secure locations.

Processing

There are two major types of processing that may be carried out on records held on files:-

UPDATE	Data values in the files records are changed to reflect a current position. Updating can also include adding or deleting records from files
ENQUIRY	Data values are retrieved from a record or a number of records to satisfy an enquiry. Example - Locate all Customers that live in the vicinity of Des Moines.

Hit Rate

This term is used to describe the rate of processing of a master file in terms of active/accessed records.

For example if 1,000 transactions are processed each day against a master file of 10,000 records, then the hit rate is said to be 10%.

The hit rate is a measure of the activity of the file.

Descriptions of high or low can be applied to the hit rate and will often influence the method of access chosen. They are subjective assessments, but in broad terms the higher the hit rate, the more likely it is that sequential processing will be preferred over direct.

File Organization and Access Methods

File organisation is a method used to store and retrieve records within a file. There are four basic forms of file organisations

- Serial
- Sequential
- Random
- Index Sequential

Serial

MEDIA Records are placed on tape or disk one after the other, or on disk into a flat file ASCII/text file.

WRITE Records within the file are stored one after the other in no particular order.

ID	Book Name	Series	Author	Type
157	A Passage to India		E M Forster	Drama
737	A Private Cosmos		Philip Jose Farmer	Science Fiction
249	A Regency Scandal		Alice Chetwynd Ley	Historical Novel
269	A Sailor of Austria		John Biggins	Historical Novel
743	A Spell for Chameleon	1	Piers Anthony	Science Fantasy
532	A Stainless Steel Rat is Born		Harry Harrison	Science Fiction
4	Aboriginals of Australia			Anthropology
347	About Time		Paul Davies	Physics

READ Records can only be accessed sequentially, in the sequence they are stored.

UPDATE Since records are in no particular order – updating is found to be extremely slow and inefficient . Many passes of this file have to be made to perform any update.

ENQUIRIES Enquiries are impractical since the whole file must be searched before the required data is found.

USE Serial organised files tend to be used for data capture where we wish to store the data in the order it arrives. Serial organisation is typically chosen for transaction files e.g. a file of unsorted transactions in a sales order processing system.

Sequential

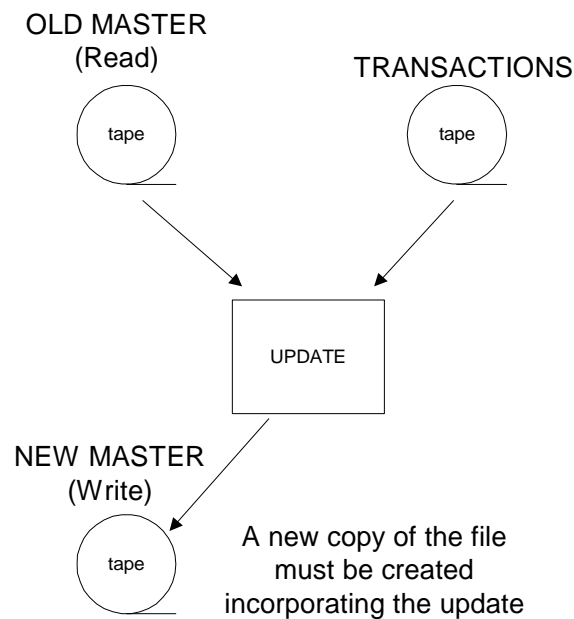
MEDIA Records are placed on tape or disk one after the other, or on disk into a flat file ASCII/text file.

WRITE Records within the file are stored one after the other in a key field (primary key).

ID	Book Name	Series	Author	Type
154	Complete DIY Manual		Readers Digest	DIY
155	The Prince		Celia Brayfield	Drama
156	The Ringed Castle		Dorothy Dunnett	Drama
157	A Passage to India		E M Forster	Drama
158	The Collected Short Stories of Saki		Hector Hugh Munro	Drama
159	The Pitman Dictionary		Isaac Pitman	English Language

READ Records can only be accessed sequentially, in the sequence they are stored.

UPDATE The file cannot be updated directly, a new copy of the file must be created incorporating the update. This is because read and write cannot be performed during the same run.



ENQUIRIES Enquiries are impractical since the whole file must be searched before the required data is found.

USE Sequential files are ideal for applications with high hit rates. A typical example is a payroll where a vast majority of records need to be updated. Sequential files are used in batch processing. Makes minimal use of index. Minimises head movement. Processes all records in block at single read – means index used one per block not once per record. Transaction file pre-sorted into sequence of master file.

Random

MEDIA Random file organisation can only be achieved using a disk.

WRITE Records are stored on disk by using a hashing algorithm. The key field is fed through hashing algorithm and a relative address is created. This address gives the position on the disk where the record is to be stored.

ame as th

READ records are retrieved by using the key field and the hashing algorithms. records are accessed directly

UPDATE Individual records and fields can be updated, since you can directly access and write to an individual record.

ENQUIRIES Data to satisfy the enquiry can be retrieved directly from the record(s), without having to search through the entire file. (Much faster).

USE Random files are ideal for applications with lower hit rates. An example of this could be a booking system where we wish to update individual records.

Sequential access to the records is difficult due to the random physical order of the records. This makes sequential access and associated processing impractical.

Direct files are used in on-line processing.
Index used once per record (read/write head moves back and forth along file).

Method used when transactions processed immediately.
Method also used when updating two files at simultaneously (for example stock file and customer file in same run).

Indexed Sequential

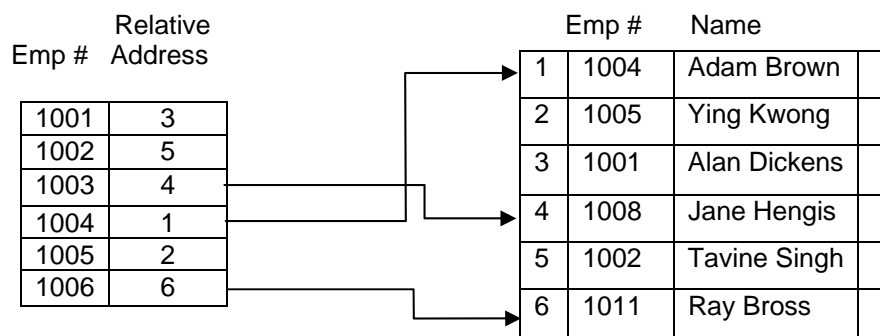
MEDIA Indexed Sequential file organisation can only be achieved using a disk.

WRITE Records are stored in blocks (buckets). No order is implied between these buckets, but within a bucket records are written to the file and an index created.

READ Records can be retrieved in two ways:-

Random Access Directly using the same method as direct file organisation.

Sequential Access Sequentially by using an index. The index is a separate file which contains the key fields and relative addresses of all records stored. It is sequentially ordered using the key field.



INDEX

DATA AREA

The problem of a random physical order of records caused by direct file organisation is overcome by the use of an index allowing sequential access.

UPDATE Individual records and can be updated sequentially and randomly.

ENQUIRIES Enquiries are carried out via Random Access.

USE Index sequential files are ideal for applications which require both random enquiries update and regular updating of the whole file.

An example of this is a stock control system which requires random enquiry updates during business hours, and stock level processing (updates) for the whole file outside business hours.

Index sequential files are used for on-line processing and batch modes.

On - Line Processing

Batch processing was the original method of data processing. Many systems in current business situations use on-line processing as it is immediate in comparison to batch processing.

It must be still emphasised that batch processing is still used in different scenarios where immediate responses are not required or where delays are acceptable. For example processing utility bills when computer systems are least being used.

On-line processing is when transactions are processed as they arise. For example when money is withdrawn from a bank account, the new balance is updated thereafter. Other situations could be room booking, to avoid double booking such room information would need to be updated immediately.

Such systems nowadays are developed as on-line, real time data capture and enquiry systems.

Characteristics of On-line Processing

- A user is linked directly to a computer for the purpose of data input or receiving output.
- There is a direct link from transmission medium to the computer processor.
- System is in constant readiness.
- Immediate processing of individual transactions.
- Online validation - Immediate notification of input error.
- Only relevant records are updated.
- Peak transaction capability

Advantages of On-line Processing

- Information is available immediately for effective decision making
- Accurate data capture
- Close control of credit and stock levels

Real Time Processing

If current information is needed, then the mode of processing required is called “real time”.

A real time system is one in which data is entered into a computer and processed, the results returned are sufficiently quickly to effect the forth coming process at a point in time

Typical examples of real time systems would be an aircraft guidance system or a flight/seat reservation system.

- The time taken to complete the processing of a transaction is known as the “Response Time” and plays a significant role in real time systems.
- To be acceptable as a real-time information processing system, the “response time” must meet the needs of the user! For example in an aircraft guidance system processing would need to be immediate (fraction of a second).
- Real time systems are developed in order to react very quickly to changes in demand, and to adjust resource allocation(s) so as to give immediate response(s) to users.
- Real time systems require more complex software and file handling techniques and faster hardware to ensure such speed is delivered.

Computer Requirements

There are several specific computer requirements for real-time processing:-

- Input/output devices which allow direct entry to the computer from remote locations using data transmission facilities.
- Immediate availability of programs that are required for processing.
- Large program storage for rapid switching of program segments to/from memory at high speed.
- Immediate access to stored data to/from direct access storage devices with very fast access and transfer rates.
- A real time operating system which can detect the occurrence of any event and quickly schedule the appropriate resources to handle it.

Examples of An On-Line System 'Drugs R Uz'



Drugs R Uz

200th Street,
Des Moines
WA 98787

Sales Order Processing

Drugs and other prescription medicines are only kept in small quantities at the drug stores supplied by 'Drugs R Us'. A batch processing system is too slow to handle these order. Drug stores placing orders will expect delivery within 3 hours of the order being placed.

Master Files used

The Stock Master File is kept as an index sequential file on disc. The product code is the primary key. This means that the index can be used to find a record if the Product Code is given (this is direct or random access).

The file can also be processed sequentially if most of the records require processing.

An in-between approach, skip sequential can be used when groups of records are processed and groups of records skipped

The Customer Account File is also an indexed sequential file. Customer A/C Number is used as a primary key.

Order Processing

Orders are received via fax/phone and immediately keyed directly in to the on-line system. The system automatically carries out validation checks:-

- Stock Availability
- Credit Worthiness
- Existing Customer etc

Subsequently the order is then processed, both the master files are updated (ONLY the records concerned are altered). The system prints the despatch note and invoice. Following this appropriate drugs / products are picked in the warehouse and despatched to the appropriate destinations.

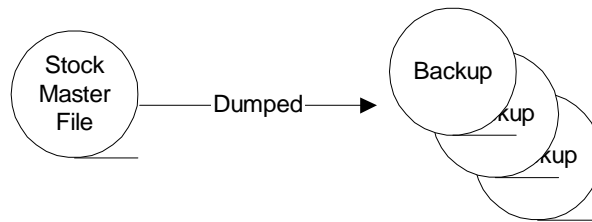
Transaction log is written to disc. This records all details of the transaction which can be used.

- To recover and back track any errors that may have occurred whilst processing.
- To provide an audit trail.

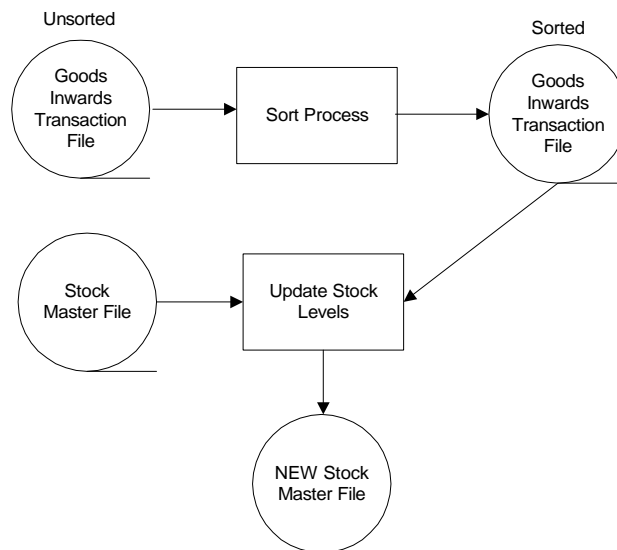
The program is on-line throughout the working day (8:00am – 8:00pm). During the evening/night shift, other processing is carried out on the Stock Master File.

Stock Master File is:-

- Back up copies are dumped to tape so that the file could be recreated if data was lost in the unforeseen future.



- Updated with details of Goods inwards received during the day (Batch processing). It is more effective to sort these transactions before updating, and then to do a sequential update (updating in place).



Subsequently the records can be accessed randomly using the index.

- When there have been considerable number of insertions and deletions, there will be records on the file not in physical sequence. Accessing the file will get slow. A utility program will then need to be executed to recognise this and sort this out by bringing together the file fragments (records).

Exercise

In groups discuss different situations and or businesses various file organisations could be used. In particular presentations of different business scenarios are expected. Typical examples could be:-

Student Enrolment System

Clothing Manufacturer

Automotive Manufacturer

Hospital Systems

Taxation Systems

Your presentation should be at least 10 minutes.