Unit 4: Transaction Processing Systems

Everyday of our lives we all come across a transaction processing system. A transaction processing system is an information system that collects, stores, modifies and retrieves transactions of an organisation. A transaction is a series of events important to an organisation. A transaction is a series of events important to an organisation that involve a request, an acknowledgement an action and outcome.

A TPS is also a unit of work composed of multiple operations that must all succeed or must all fail where each operation generates and/or modifies data. If all operations of a system are completed successfully, then a transaction will be completed successfully but if one of the operations fail, then the whole transaction is aborted - including those operations who completed their tasks successfully.

Example: withdrawing money - if the account cannot be debited then the cash will not be given and vice versa.

Therefore, a transaction and its operations are committed and if only all operations of the transaction are successful meaning that they are permanently stored within the appropriate database systems. If one or more events have failed then the transaction is “rolled back”, meaning each event is requested to abort all actions.

TPS’s are made up of properties (aspects of the system), the resources (components that contribute to the system) and behaviours (ways in which the properties and resources interact)

There are 4 basic properties and are referred to as ACID properties of a transaction

4.1 Characteristics of TPS

- Atomicity - these properties test that a transaction is carried through completely and if it fails it is undone completely.
- Consistency - the data stored is accounted for and each step of the transaction is carried out in the same way each time it occurs so to ensure the data is correct for each part of the transaction.
- Isolation - transactions must be independent of each other i.e. treat each transaction separately, keeping the data for each transaction separately.
- Durability - the effects of the completed transaction should last i.e. transactions recorded and checked at any time and backups kept of all data.
- Economy - conducting a transaction should not be expensive
- Divisibility - it must be possible to interchange many low transactions for one high-level transaction
- Scalability - system must support many users simultaneously
- Interoperability - a value should be able to move back and forth between many system i.e. the dollar value in a retail transaction is the value of the currency used to purchase a service
- Conservation - money should hold its value over time and is easy to store and retrieve. Resources are the customers or clients and equipment of an organization involved in transaction processing e.g. Hardware, software (general & specific). Most TPS’s used DBMS as their storage software.

Behaviours are the way in which a tps responds to the transactions that

The expected behaviours include:
- Reliability - a TPS should be reliable as failure could result in a collapse of a business. There should be backup and recovery procedures in place to prevent total failure and restore businesses to full working capacity a.s.a.p in large businesses mirrored systems operate i.e. two TPS’s hold identical data that work together and if one fails the other continues without affecting the system. Other hardware could also include UPS (uninterrupted power supply)
- Standardisation/inflexibility - refers to the processing of each transaction in exactly the same way. Data handling and recording methods should be the same. This prevents errors occurring in the system. A TPS should process all transactions using the same steps regardless of when, where and who is involved in the transaction. If a TPS is flexible, there would be too many opportunities for non-standard operations.
- Control - is very important as it enables roles and responsibilities in an organization to be defined and gives each person in the system a clear understanding of what they are required to do. There should be a clearly defined system
- Rapid response/performance - fast performance with a rapid response time is critical. Businesses cannot afford to have customers waiting for a TPS to respond. The turnaround time from the input of the transaction to the production of the output must be a few seconds or less. This can either relate to the pace of the transaction or the number occurring per second.

4.1.1 - Historical Significance of TPS
In 1964 IBM released its popular SystemJ360 which allowed online processing of transactions to become commonplace. In 1969 IBM released the CICS transaction processing monitor software.

IBM in the early 1970s was the first to release a relational database known as system R - based on the work of Ted Codd. Relational DBMSs became popular during the 1980s following the release of oracle then IBM’s DB2. SQL was standardised in 1986. Microsoft released SQL Sewer in the 1990s after purchasing Sybase.
From 2000 onwards the commercial Oracle, SQL server and IBM DBMS continue to be updated, however new competition in emerging from open source DBMS such as MySQL. TPSs are routinely used by large organisations whose transactions span multiple servers and systems.

Historical batch processing was the first type of transaction processing. In the early days of computers all input was via punch cards - this includes the actual program code as well as the data to be processed. Each card was manually punched by an operator for input. Completed stacks of punch cards were physically loaded into the computer and processed sequentially. In these early days online real time processing of multiple transactions was simply not possible. The hardware performed a single task at a time and the output was stored sequentially on magnetic tape. The processing resources were limited and also costly, therefore batch jobs were...

4.1.2 - Automation of a Manual TPS

The automation of manual transaction processing is at times not possible. Usually, to be able to automate a task requires the manual transaction to follow a strict sequence of events where each event must be acknowledged as complete before the next commences and if any event fails then the entire transaction is aborted and rolled back.

When assessing the work routine of a clerk to check for suitability, first check whether the manual system follows a strict sequence of events and then consider the strengths and weaknesses of each to judge whether the automation of the transaction is justified.

Manual system strengths:
- Minimal start-up costs - little or no initial capital expenditure.
- Minimal training time and costs
- Quick response to changing requirements
- Well suited to small organisations where participants have time and fulfil multiple roles.
- Responds well to human insight and intuition

Manual system weaknesses:
- Analysis of historical data is difficult and time consuming
- Transactions take considerably longer to process.
- Difficult to rigidly enforce transaction rules and sequences.
- Redundant or duplicate data is a feature of most manual system.
- System becomes more and more difficult to manage as it grows.
- Making backups of data is difficult and is rarely, if ever, performed for all data.

Automated transaction processing strengths:
- Much faster transaction processing.
- Less repetitive work for participants.
- Enforces the sequence and rules for each transaction
- Calculation errors are virtually eliminated.
- Ability to integrate transaction processing with outside organisation.
- History data available for statistical and financial analysis
- Backup easily made and restored if system fails.
- System easily grows as transaction processing needs grow.

Weaknesses of automated transaction processing:
- Significant start-up costs to purchase information technology.
- Extensive training required to operate the system.
- Changes to requirements often require specialised expertise to implement
- Rigidly enforces existing transaction rules and sequences for all data
- Less total work for humans resulting in lower employment
- Reliance on information technology - failure of one or components that can cripple the system.

4.1.3 - Components of TPS

There are two types of transaction processing systems which include:
- Realtime data transaction processing: the immediate processing of data. (e.g. uber eats, ebay, any online delivery/shopping)

And
- Batch processing: the collection and storage for data processing at a scheduled time or when there is sufficient data. (medicare, government forms, cheque clearance, payroll, etc.)

All transaction processing systems, just like any information system, are made up of components including:
- Purpose: defines who the TPS is for and what their needs are.
- data: in the majority of transaction processing systems, data is stored in databases - usually relational databases. This data is transformed into information by the system’s information processes. However in transaction processing systems the issue of integrity (correctness and accuracy) of data during transactions arises.
Components of TPS:

- **Information technology**: the hardware and software forms the information technology of the system. Hardware includes servers, large and fast access storage, communication devices and transmission media, backup devices and client workstations such as ATMs, EFTPOS machines or just personal computers connected to a network. Software includes DBMS to manage and control transaction on databases, client applications/interface and transaction processing monitors (TPM) which coordinate transaction processing of large TPSs. the TPM controls the commit and rollback of the total transaction response to requests sent to and acknowledgements received from each operation.

- **Processes**: each transaction is an information process and is therefore composed of events that are also information processes: all information processes are performed using the resources within the system. The system’s resources include participants, data and information, and information technology.

- **Participants**: anybody who interacts directly with a transaction processing system becomes a participant in that system - they are integral to the system’s operation. Therefore participants include people who work for the organisation that operates the transaction processing system and also people who enter data that initiates transactions. On the other hand indirect users are not participants, they send and/or receive data from the system but o not directly cause its entry or display.

4.1.4 - **Data Integrity**

The integrity of data is critical in all transaction processing system since inaccurate data may have dire consequences on the TPS. Data integrity is a measure of how correct and accurate data is compared to its source. In order to improve data integrity, it is necessary to:

- Apply data validation techniques
- Apply ACID properties

Data validation involves setting criteria within the software that automatically checks that the format of data is in the correct format and is reasonable as it is entered into the system. Examples of data validation checks include:

- Range checks used if restricted to a small range of values e.g. where you have a 13 digit field for an ISBN or a date field which cannot be later than today’s date.
- List checks data is compared to a set of expected data e.g. male or female
- Type checks check if the data type entered is correct, e.g. if formatted as number, text cannot be added.
- Check digits a number is calculated from the digits of a code number is then added to that number as an extra digit e.g. ISBN08585992219
- Entering data twice and comparing them to verify that they are entered correctly, e.g. passwords.
- Error detection methods to check if any errors occurred during transmission and retransmitting if true.
- Provide means that enables clients to update data to ensure details are correct because for example, they may move houses or change telephone numbers. Some companies provide forms on the back of envelopes to remind and allow customers to update their details.

**4.1.4b - Data Integrity - Data Verification**
- This difficult task is rarely 100% successful, as people and businesses change locations, phone numbers, names and credit card numbers. Ensuring these changes are reflected in the data is the aim of verification processes.
- In Australia there is no single identifier that can legally be used to identify individuals, due to privacy concerns. While the Tax File Number and Australian Business Number can be used by the government to link information, it is illegal for a private organisation to do the same.

**4.1.4c - Data Integrity - Referential Integrity**
- Referential integrity ensures all foreign keys in linked tables match a primary key in the related table, meaning a record in the primary table must exist before data is added to the linked table. If referential integrity is not enforced, orphaned records will exist.
- When there are many databases involved, errors are inevitable. Validation and verification can affect integrity also. For example, and incorrect BPay number means that primary records cannot be linked to various organisations for customer payment.

**4.1.4d - Data Integrity - ACID Properties**
- ACID properties combine to ensure the integrity of all data is maintained before, during and after each transaction. They ensure that transactions are never incomplete (atomicity), the data is never inconsistent (consistency), transactions do not intrude or affect each other (isolation) and that the results of a completed transaction are permanent (durability).

**4.2 - Real-time (Online) TPS**
- Real time transaction processing involves the immediate processing of data. They complete transactions immediately after they have been initiated and feedback is also given.
- Real-time processing occurs when a participant or device makes a request for processing to occur. The request is performed immediately and gives an immediate result.
- Usually, participants require access to an online database (does not necessarily mean through the internet) which allows them to process transactions in real time
participants of a real time transaction system often include employees of an organisation and customers who are entering details to the system via the internet or other networks.

- Real-time processing involves a large number of people who are simultaneously performing transactions to change data. Even though each individual person is processing a relatively small number of records, their requested are being made at the same time, using a lot of processing power.

Therefore, to update data in real-time processing the information technology required includes:

- Direct access storage devices which provide much faster access than sequential devices. Storage also needs to be large to store large amounts of data.
- Fast processors that reduces the time needed to access and complete transactions
- User friendly software interfaces as participants, trained or not trained, will use it. The software will also need to allow for online and simultaneous access to data as many people can perform transactions at once. Also an online database is needed to allow for real-time processing and committing of transaction data.
- Depending on the size of the organisation, critical online data can be stored on RAID storage system which uses multiple hard disk drives. Transactions are saved on each of the disk drives. This has the overall effect of speeding up data access and improving the reliability of the storage system.

The steps involved in updating in real-time transaction processing include:

1. Enter data
2. Retrieve the appropriate record from the master file.
3. Update the record
4. Send the appropriate response/feedback back to the participant

Advantages:
- it provides an instant confirmation of a transaction- response time is faster.
- User is given rapid feedback so is able to check data without delay
- It does not require an online database.
- Data is verified and validated as it is entered.
- Data is processed when needed
- Errors may be corrected immediately.

Disadvantages:
- System failure is critical and has widespread effects.
- Locating problems may be difficult as nodes are usually widespread.
- Maintaining the network may also be difficult
- Hardware and software requirements are more expensive.
Rigid procedures are needed to ensure all transactions of the same type are handled the same way. Generally, there are two types of transaction processing systems, namely online real-time and the other is batch. In an online real-time TPS, the term online does not refer to internet access, but it really refers to having access to a database where the transaction data is stored and that the participant is able to access and update transactions immediately. Online real-time systems can be two types:
Non web-based: does not require internet access
Web-based: requires internet access

4.2.2 - Point of Sale Systems
Point of sale systems process transactions within retail outlets. Retail outlets include
- Small local stores
- Chains of stores
- Hotels and clubs and
- Also large department stores and supermarkets
Although the amount of data increases significantly for larger retailers the general nature of the essential transactions remain similar. Retail stores sell directly to customers and they purchase inventory or stock from suppliers as described in the context diagram below. These two processes, purchasing and selling, form the basic transactions performed by all retailers and hence must be present within all POS systems.
- The retailer sends suppliers a purchase order, the supplier sends the products together with a delivery docket followed shortly by an invoice.
- Finally the retailer pays the invoice and sends the payment details in the form of a remittance advice.
- When selling retailers accept payment from customers and provide the customer with a receipt - often in the form of a tax invoice.

4.2.2a - POS Systems - Processes
Particular companies produce and market proprietary POS systems for specific industries. Some companies produce and market complete POS systems for jewellery stores, other specialise in
hardware stores, whilst others specialise in fruit and vegetable stores. Commonly these systems include all information technology, together with the training required to operate the system.

Let us expand on our initial POS system context diagram by considering some typical transactions performed by most POS systems.

4.2.2b - Point of Sale Systems - Data/Information
The data within POS systems is almost always stored in a relational database. For the system described above tables for products, suppliers and purchase orders would be required. These models also show the information output by POS systems - that is, receipt details, purchase orders details and remittance advice.

4.2.2c - Point of Sale Systems - Information Technology
All hardware, apart from the POS terminals, is common to many other systems. Therefore we restrict our discussion to the detail of POS terminals. Firstly, the use of the word terminal is somewhat misleading - in fact most current POS terminals are in fact personal computers that include integrated collection and display devices. In the past POS terminals were indeed terminals where processing was performed centrally. Today POS systems are largely client-server systems and hence much of the processing is performed by the client.

Some of the data used by this system:
- Product details such as: name, expiry date, country of manufacturing, delivery/order date, supplier, retail details, stock levels etc.
- Supplier details: SupplierID, Company etc.
- Purchase order details: PONumber, UPC etc.

Data and information
Information that is generated or can be extracted by the system:
- Most popular products
- Daily takings
- Products that need markdown
- Current stock levels.
The information technology for POS systems include a database server that runs DBMS software and includes sufficient storage to secure and maintain the database. For smaller retailers backups are made to CDRs, whilst larger systems include tape drives. One or more POS terminals are installed which run the POS client application, that processes sales transaction. Further personal computers are often present to perform other transaction. Commonly an ethernet LAN is used to connect to the database server. Currently most POS terminals include a standard PC motherboard including:
• Intel processor
• RAM and hard disk.

Attached or integrated devices include:

• touch screens
• magnetic stripe readers
• barcode scanners
• cash drawers
• Receipt

4.2.2d - Point of Sale Systems - Social/Ethical Issues

Ergonomic issues for participants using POS terminals are different compared to the issues present for those seated at most traditional computer workstations. POS terminals are commonly used whilst standing for extended periods of time and the collection devices are different. The task performed by POS terminal users often include a much broader range of movements as they scan products, use touch screens and interact with customers. Barcode scanners, touch screens and magnetic stripe readers reduce the likelihood of RSI and other health issues associated with keyboard data entry.
Activity Summary:

<table>
<thead>
<tr>
<th>Reservation Systems</th>
<th>Library Loans</th>
<th>Information Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Processes:</strong></td>
<td><strong>Processes:</strong></td>
<td><strong>Processes:</strong></td>
</tr>
<tr>
<td>- Collect and store</td>
<td>- Catalogue searches</td>
<td>- Database system with</td>
</tr>
<tr>
<td>service details</td>
<td>- Library finance</td>
<td>DBMS software with</td>
</tr>
<tr>
<td>- Confirm availability of</td>
<td>- management</td>
<td>sufficient storage</td>
</tr>
<tr>
<td>service</td>
<td>- Charges for</td>
<td></td>
</tr>
<tr>
<td>- Collect and store</td>
<td>- Digital data</td>
<td>- Smaller retailers have</td>
</tr>
<tr>
<td>customer details</td>
<td>- Check-in system</td>
<td>back ups made into</td>
</tr>
<tr>
<td>- Collect payment</td>
<td></td>
<td>- Ethernet LAN</td>
</tr>
<tr>
<td>details as required</td>
<td></td>
<td>connection used to</td>
</tr>
<tr>
<td>- Create and display</td>
<td></td>
<td>connect to a database</td>
</tr>
<tr>
<td>confirmation to</td>
<td></td>
<td>server</td>
</tr>
<tr>
<td>customer</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Data information:**

- Catalogue searches
- Finances
- Financial charges for printing

---

**4.3 - Batch TPS**

Batch processing involves the collection and storage of data that represents the transactions in a transaction file. The data is processed without any user interaction in a group or a batch at a different time of the actual transaction, perhaps at a time when the system is quite such as late at night or until sufficient data exists. Waiting for a large volume of data generally results in lower processing costs per transaction. The time delay before processing or completing batch may take several minutes, hours or days.

Common examples of batch processing include clearance of cheques, bulk generation of bills and payments, and payroll processing. Batch processing is common when the required data is collected on paper forms.

When batch processing commences, all data collected in a transaction file are prepared for processing where each transaction executes in much the same way as a single real time transaction. Successful transactions are committed (permanently written) to the master file or database. Transaction problems and errors result in rollback are written to an error file (or log). Writing detail to a file allows further transactions within the batch processing will halt when an error occurs.

The processing resources of all computer systems are limited, therefore batch processing are scheduled to occur during evenings or weekends when real time processing requirements are lowest. Such scheduling not only ensures CPU processing resources are available, it also reduces the wait time for transactions as it is less likely that other transactions will be
simultaneously requiring access to the same data. E.g., banks process internet transactions at night.
Batch processing can sometimes utilise an offline database because it may not need real-time data. For example, in bill generation, the bill will include information on last month’s transactions and does not need details of any other more recent/current transactions.

User interaction with batch processes is restricted to input prior to the commencement of processing and to deal with problems after batch processing completes. Furthermore employees rather than customers commonly initiate batch processing. As a result the design of user interfaces for batch processing is different – they are designed for rapid entry rather than user friendliness.

**Advantages**
- Reduced costs - set-up costs lower and processing not as demanding in batch systems. Updating data only when all data is ready keeps costs minimal whilst personnel need is greatly reduced. Carried at times resources are least used such as overnight.
- Standardisation - all data is processed the same way
- Control - organisation able to decide exactly when data processed.

**Disadvantages**
- Time delay - batch processing delayed regardless of delay whether delay of a few minutes, a few hours or a few days. Results in processing not available immediately after transaction is completed.
- E.g. - if a book is returned today, until books return processing is completed, book can’t be borrowed by the next person.
- Identical data - must be processed in any one batch. Large batches of different data collected must be sorted before processing so each individual batch contains one type of data.
- Errors - errors in a batch can’t be corrected during processing and only corrected later after another batch has passed on to be reprocessed.
- Costs - some forms of batch processing are expensive and if requiring keyed input then personnel must be hired for inconvenient times such as overnight.

Batch processing mostly involves automated tasks and does not involve any users. Participants are involved mostly in the collection and compilation of the transaction file but sometimes no participants are needed during the processing.
The information technology required in updating in batch processing includes:

- An offline database which may be a snapshot of the real-time database to avoid interference while processing and the real-time processes won’t be degraded by the batch processes. This is required when the online database may be updating in real-time while some batch processing is being performed e.g. bill generation analysis of data can be done online after the updates are completed. The user interface is not as important in batch processing as in real-time processing because most batch processing is automated.

- Distributed/parallel processing to speed up the batch processing which allows several transactions to be processed at the same time allows several transactions to be processed at the same time since each transaction is independent. This is useful in large organisations which process thousands of transactions at a time.

- Sequential access of the transaction file during the batch update where data is stored in the order of processing. This becomes quicker because all records in a transaction file will need to be processed and time is saved from using direct access. This is also useful where large amount of processing is needed.

The steps involved in updating in batch transaction processing include:

1. Compile all transaction data is a transaction file.
2. Organise transaction in order of processing.
3. Process each transaction and update the master file when transaction is complete. Repeat until transactions are finished
4. Record any errors encountered in an error log.
4.3a - Batch TPS - Systems to store Paper records

Even though we are living in the digital age, paper records are still common and important part of transaction processing. However, paper records in their original hard-copy format has several disadvantages:
- The storage of paper is bulky and requires a lot of physical space
- A separate index or catalogue system must be maintained to allow documents to be located
- It is difficult to manually find all the occurrences of text in multiple documents.
- Sorting large quantities of paper documents is a slow and unreliable process.

The alternative to paper records is to convert them into digital data in a process called digitising. Digitising can be achieved using two methods:
- Document imaging: this is where documents scanned and saved as images, even text become part of the image. This means the text cannot be searched or edited. It is basically a photocopy of the original paper but in digital form. However, separate indexing systems are required to access the scanned documents as searches cannot be done on the text contained in them. Its main advantage is that it saves storage space.
- Optical character recognition (OCR): like document imaging, pages containing text are scanned and saved as image files. Character recognition software then processes the images to identify the different shapes of the characters. When a character is recognised, it is held in a document file which can be opened and edited by a word processing application. This document can also include any images that were on the page. Scans of multiple pages can be combined in a single document file. OCR is not perfect in recognising letters as it may confuse a ‘1’ with an ‘l’ but such mistakes can be easily recognised with a spell checker.

4.4.1a - Full & Partial Backups - Full

- A full backup is a complete copy of all data within the system. It could include a complete copy of the hard disk including software applications as well as the data but most businesses perform a complete backup of only their data files on a regular basis.
- Full backups are the easiest to restore should a failure occur as it will simply be copied back into the operational system. Unfortunately copying all files takes a long time and requires a large amount of storage; therefore it is often impractical to perform full backups on a daily basis. Common backup procedures specify that full backups be made on a weekly basis, usually commencing on friday afternoon and for large systems continuing over the weekend.
-
A partial backup is done to save time and resources where only files changed since the last full backup are copied. In order to perform partial backups, there needs to be a full backup performed first. There are two types of partial backups; incremental and differential.

Incremental backup: incremental backups include making partial backups that copy all files that have changed or been created since the last backup. The last backup may have been a full or partial backups. As a consequence, performing an incremental backup is significantly faster and requires significantly less storage compared to full backup. Before incremental backups can be made it is necessary to first perform a full backup commonly faster incremental backups are then made on at least a daily basis. The significant savings in backup time is counteracted by the extra time required to recover the data. During a recover the latest full backup is first restored, then each partial backup is restored in the order in which they were made.

Differential backup: a differential backup uses partial backups to make copies of all files that have been altered or created since the last full backup. If such partial backups are made each day then each will contain copies of files within all previous partial backups since the last full backup was made. To restore to the most recent backup requires first restoring the full backup and then restoring the most recent partial backup. The size of the differential backup continues to grow until the next full backup is completed.

A transaction log is maintained that contains a record of original data and the changes that are to be made as part of a transaction. If all events within a transaction can perform their actions successfully then the transaction can be committed to the master file. Those exam changes are permanently made within the actual database. If someone part of the transaction fails then all actions are reversed using transaction log records.

The most compelling reason for maintaining before and after versions of all data changes is to provide a backup of all recent changes since the last backup. Transaction logs contain details of all the transactions made since the log and data was last backed up. Should a failure of the main database occur then it can be first be recovered from the latest backup. The transaction log is then used to roll forward all transactions that occurred since the backup was made. If at the time of failure some transaction were incomplete..

Mirroring writes all data to two or more secondary storage devices. If one of these storage devices fails then the system can continue operating using the mirrored copies.
Mirroring improves the fault tolerance of the system. However mirroring only maintains the current data it does not keep historical versions of the data, therefore it should not be used as a replacement to normal backup procedures. In addition most mirrored drives are typically within the same unit in the same location, hence there is no protection against total destruction. An example of mirroring is RAID storage systems.

4.4.3 - Backup Media
- There are different types of storage media that could be used for storing backup file. The decision which backup medium to use will depend on the systems operations and the amount of data that needs to be backed up.

4.4.3a - Backup Media - Magnetic Tape
- Magnetic tape has been the backup medium of choice for a long time. The advantage of tape is cost; its less expensive than other storage options. However, the tradeoff is performance. As the amount of data that organisations have and need to backup has grown, the amount of time it takes to backup all that data to tape has become increasingly inconvenient. Likewise, finding data on tape is a time-consuming process since it is a sequential access media.

4.4.3b - Backup Media - Hard Drives
- As hard drives prices continue to drop, many businesses and individuals are turning to external hard drives as a backup solution. The cost per gigabyte can be very cheap, and copying the data is incredibly easy. Hard drives are prone to physical malfunction and corruption by a number of forces, however, making their durability questionable.
- CD-R and CD-RW technology offers a number of advantages as a backup solution. Both types of media are incredibly cheap, write times are fast, and the data is easily accessible by virtually all computers. Capacity is 650-700MB. Recent studies have indicated that durability is less than was initially presumed, with retention dropping after five years. For cd backup for most people. Dvd technologies can either hold 4.7gb or, in the case of dual-layer dvds, 8.5gb. Media is currently comparable in pricing with cds, with discs available in bulk for $0.50 each. Dvds have similar durability to cd technologies
- Dvd-r, dvd+r, dvd+rw,dvd-ram, and dual layer dvd technologies are essentially they replacement
- Businesses are beginning to emerge on the internet that specialise in providing online backup and recovery for individuals and small businesses. The online systems totally automated the backup process for users. All data is transferred via the internet to a secure remote site. The remote site then manages the secure storage of the data on behalf of the individuals or business. Clearly the remote site must use some form of secure and permanent storage. When first using an online backup system a full backup must be made, which is a time consuming process.

- After the initial backup, incremental backups are made at regular intervals - in some cases every time a file is saved. Such systems enable recovery of different historical versions of individual files as well as recovery of computer systems.

- Large organisations that manage large volumes of critical data maintain complete operational copies of their entire system at remote locations. This will enable instant recovery from the backup online site. This is considered to be the ultimate backup procedure in the case of a fire or terrorist attack.

- Backups need to be readily available for quick recovery should a failure occur, however what if the building is destroyed by some catastrophic event? In this case backups are stored offsite are CRITICAL. But it is common for users to request old versions of files - perhaps they made a change or have deleted a file and now require the old data. Such events require quick access to backup copies. The solution is to store some backups onsite and others offsite. The organisation should assess the risks and possible losses that would occur should they lose data.

- For some organisations losing even an hour of data changes is significant, whilst for others loss of a week's work may be an acceptable risk in all cases it is wise to regularly store some backups in a secure area onsite - such as within a fireproof safe or a locked storeroom. A less regular intervals a complete backup should be stored offsite. In smaller organisations a trusted employee..

- The grandfather, father, son backup strategy is a method of maintaining backup on a daily, weekly, and monthly schedule. GFS backup schemes are based on a five-day or seven day weekly schedule.
4.5.2 - Collection from forms

Form design is an essential aspect to transaction processing as most of data collection occurs through forms - including online, web-based and paper forms. The processing of paper forms are usually done in a batch and online forms are processed in real-time. However web forms processing depends on the needs of the organisation where some forms are processed in real-time and data is collected in others and stored temporarily in a database and then processed at a later time by an employee. With form design, one of the important considerations is the friendliness of the user interface. It is also important for forms to provide and enforce data validation. This is possible with online and web-based forms, however, with paper forms instructions and indicators of the required data are provided for example instructions, examples and input areas that restrict length of data.

4.6 - Analysing data output from TPS

Once the transaction data is processed and stored, it may be analysed in a number of ways for a range of different purposes. The output for transaction processing is the input for other information systems such as data warehouses, decision support systems and management information systems.
4.6.1 - Data Warehouse

A data warehouse is a database that includes copies of data from each of an organisation’s operational/transaction databases. It is a large separate combined copy of all the operational databases which is read only and is data from each of the databases in uploaded to the warehouse at regular intervals. It contains historical data which is used to analyse the activities of the organisation and provide evidence that will help in decision making and thus improve the organisation’s performance.

The data warehouse is not in itself an analysis tool rather it is a data resource that analysis tools access to analyse the historical activities of the organisation. Since it is independent of the operational database, analysis can take place without concern over simultaneous access or updating of transaction records. Furthermore the data warehouse can act as an archive for the organisation’s historical data.

Advantages of a data warehouse include:

- Old transaction data can be purged from the operational system
- Analysis processes performed on the data warehouse do not degrade performance of the operational system.
- A data warehouse includes historical transaction data, often over 10 or more years. Systems change completely and are regularly upgraded, however data warehouses are designed such that all data is stored using a similar format. This common format greatly simplifies analysis processes.
- Data warehouses are snapshot copies of the real data. This data does not and should not change. Therefore analysis processes can proceed more efficiently. There is no need to be concerned with record locks, ACID properties and data integrity issues.
- Data warehouses centralise data from within the entire organisation
- As a data warehouse is completely separate to the operational data it can be organised differently to the operational data.

Management information systems (MIS) use existing information to assist management in organising and controlling an organisation more effectively. MIS reports give information on the current status of an organisation using summaries or statistical analysis of existing data within the system. For instance, MIS use input transactions such as the monthly sales account reports to provide a graph that summarises the annual sales performance of a company. Large management information systems link to transaction data and perhaps to a data warehouse. For instance reports that compare current productivity with historical productivity require access to current transaction data and also to historical data within the organisation’s data warehouse.

MIS are part of the regular planned reporting process and are concerned with the events that have taken place within an organisation. The reports generated are used by managers to plan and direct the operation of the organisation.
4.6.2 - Management Information Systems

- Management information systems (MIS) use existing information to assist management in organising and controlling an organisation more effectively. MIS reports give information on the current status of an organisation using summaries or statistical analysis of existing data within the system.
- For instance, MIS use input transactions such as the monthly sales account reports to provide a graph that summarises the annual sales performance of a company. Large management information systems link to transaction data and perhaps to a data warehouse. For instance reports that compare current productivity with historical productivity require access to current transaction data and also to historical data within the organisation’s data warehouse.
- MIS are part of the regular planned reporting process and are concerned with the events that have taken place within an organisation. The reports generated are used by managers to plan and direct the operation of the organisation.

4.6.3 - Decision Support Systems

Decision Support Systems (DSS) are similar to MIS in that they are used to assist managers in making decisions, however, they are more forward looking and involve a lot more processing than summarising transaction data. The analysis performed by decision support systems presents possible solutions and is able to assess the likely consequences of making particular decisions. Many decision support systems look to the future, they are able to generate forecasts and predictions based on historical or incomplete data. For example predicting future interest rates or forecasting the weather are problems that do not have a definite single correct solution. Decision support systems analyse the available data to produce or suggest the most likely outcomes.

- For example, a DSS uses input transactions such as:
  - staff records, to create a model of the management structure for planning a major restructuring of a company
  - past accounting trends, forecasted inflation and interest rates, to create a what if scenario to predict the effects of a new factory on cash flow and overall profit margins

Decision support systems that analyse transaction data commonly use a data warehouse as their data source. Data mining is one decision support technique that examines the raw data in an attempt to discover hidden patterns and relationships. It presents new information that was not originally intended to be present within the data.
4.6.4 - Enterprise System
An enterprise is simply a large organisation, for example government departments, large corporations and universities. An enterprise system is any system that performs processes central to the overall operation of an enterprise. This includes critical hardware, critical software applications and in particular critical data. For instance, a typical university would have a variety of enterprise systems in operation, including a student records system, a finance system, a payroll system, a human resources system and also a content management system. Each of these enterprise systems is central to the running of the university and operates throughout the university.

4.7.1 - The Changing Nature Of Work
Until recently, data collection and processing has been a manual job. The automation of various manual jobs has largely turned clerks into mere data entry operators. It has replaced most dull and repetitive manual tasks. The revolution of information technology has eliminated some jobs and created new ones. This brought about the need for workers to retrain whilst the issue of deskilling become a problem for other workers. People involved in retailing require fewer skills as it is no longer necessary to remember prices of goods, whether or not products are available because this information is all available from a machine.
As the technology has become more sophisticated and more widely available, the roles and identities of participants and users have changed and merged. There are many old jobs that were once performed by employees but are now performed by members of the public. The bank’s teller is the most common example of the issue as members of the public can directly withdraw money from ATMs or pay for goods using EFTPOS. An emerging trend is also self checkout which has been adopted by some department stores. Online shopping is another example that does not require a customer’s interaction with a sales assistant.
4.7.2 - The Need for Alternative Non-Computer Procedures

Inevitably, all transaction processing systems will come across problems like power failure, theft, communication line breakdown or a natural disaster like fire or floods. In case of a transaction system failure, organisations need to adopt procedures that can be used in emergencies. The most common form of alternative procedure to systems usually involves non-computer manual procedures. For example, if the EFTPOS system goes down in a shop, the sales assistant will have to resort to processing EFTPOS transactions manually using paper form. Some large organisations might have backup power generators or may even have redundant communication lines to a mirrored system in case of a breakdown. It is important however to regularly test backup plans to ensure that they operate correctly and also to train the participants on how to adopt/use alternative procedures so if something does happen, they know what to do.

4.7.3 - Bias in Data Collection

Bias is an inclination or preference that influences most aspects of the collection process; the result of bias during collection is inaccurate data leading to inaccurate outputs from the system. Those involved in collecting data must aim to minimize the amount of bias present.

Often in the cheapest or most available source of data is used rather than the best source of data that represents the entire population. Consider surveys; the data source for all surveys should aim to be a representative sample of the entire population. However for ease of collection many organisations collect survey data from users over the internet. Internet users, in most cases, are not a representative sample of the population; in general internet users are younger, have higher incomes and possess higher technology skills than the general population. Consequently results derived from such surveys would not accurately reflect the entire population.

The collecting process itself should take into account the likely perceptions held by those on whom the data is collected. People answer questions and fill out forms differently based on their perception of how the survey data will be used. For example, a survey conducted by the Australian Taxation Office (ATO) is likely to yield different results similar survey conducted by the Australian Bureau of statistics. Individuals would likely perceive the tax office as being interested in their individual responses whereas a survey conducted by the Australian Bureau of Statistics is more likely to be viewed as truly anonymous.
4.7.4 - Data Security Issues

For most organisations employees, customers and data are their most important assets. Employees are the primary system participants and hence they legitimately require access to data; however whenever access is granted potential security issues emerge. Employees can maliciously, or even inadvertently, alter data or they may access and use information inappropriately. Customers can have access to systems so that they can initiate transactions and display transaction results. This presents further security issues, particularly when access is provided via the public internet.

A summary of some common strategies to combat data security issues include:

- **Passwords** - passwords are used to confirm that a user is who they say they are.
- Once verified the user name is then used by the system to assign particular access rights to the user.
- **Backup copies** - a copy of important files is made on a regular basis. Should the original file fail or be lost then the backup copy can be used. It is important to keep backup copies in a secure location.
- **Physical barriers** - machines storing important data and information.
- **Anti-virus software** - all files are scanned to look for possible viruses.
- **Firewalls** - a firewall provides protection from outside penetration by hackers
- **Data encryption** - data is encrypted in such a way that it is unreadable by those who do not possess the decryption code.
- **Audit trails** - the information system maintains records of who and when transactions were performed.

4.7.5 - Data Integrity Issues

Data integrity is about the correctness or accuracy of the data. It is a measure of how accurately the data matches and continually match its source. Techniques to improve data integrity include:

- **Data validation** - checks, at the time of data collection, to ensure the data is reasonable and meets certain criteria it does not check the actual accuracy of data.
- **Data verification** - regular checks to ensure the data collected and stored matches the source of the data.
- **Referential integrity**
- **ACID properties (data integrity)**
- **Minimising data redundancy**
Data quality extends the meaning of data integrity to include the reliability and effectiveness of the data to the organisation. These are often higher level issues that deal with the purpose of the data. Quality issues include the design of data entry forms so users respond as intended and enter responses in a consistent manner and standardising attributes used to store the same data. Quality simplifies the analysis of data, particularly when the data is sourced from multiple databases or systems. The process of improving the quality of existing data is known as ‘scrubbing’ or ‘cleansing’.

Control in transaction processing can be considered in two ways:

Control of data in transaction processing systems: most people in the world have held about them somewhere with or without their knowledge. This is valuable information that some companies pay so they could be able to use it. Management were the only people with access and control of data while participants only had limited access. However, today participants have more control and access as this is seen as necessary to allow workers to do their job more efficiently.

Control of participants by management in transaction processing systems: the power management has over its employees is necessary but sometimes is excessive. Some managers often imose authoritative control such as excessive monitoring in combination with negative consequences in an attempt to enforce control. Such methods reduce motivation and therefore performance is reduced. More effective style of management is achieved when participants are positively motivated and this will produce higher level of productivity from participants. Participants are motivated when they are encouraged to take more responsibility for work they complete.
**4.8 - Current & Emerging trends in TPS Data Warehousing**

A data warehouse is a database that includes copies of data from each of an organisation’s operational databases. It is a large separate combined copy of all the operational databases which is read only and is data from each of the databases is uploaded to the warehouse at the regular intervals. It contains historical data which is used to analyse the activities of the organisation and provided evidence that will help in decision making and thus improve the organisation’s performance.

**Data Mining**

Data mining is the process that looks for non-obvious relationships and patterns in the data stored in a database. It sorts through the data and turns up interesting and useful connections. For example, data mining could be used to analyse the transactions at a supermarket. It might determine that there was a relationship between tomato sauce sales and meat pie sales. This information might be useful for marketing promotions. One problem with data mining is that many of the patterns occur by chance and have no value in making decisions. It also raises issues of privacy and ownership of data.

**Online Analytical Processing (OLAP) and Online Transaction Processing (OLTP)**

OLAP is an example of a real-time decision support system that analyses existing data to produce graphs and statistics. The user chooses the graphs and statistics they require and the OLAP system produces the graphs immediately so that the user can flexibly explore relationships and trends.

OLTP is the real-time processing of transactions. OLTP produces new and modified data as the transaction is completed. The transaction data resulting from OLTP is often exported and reorganised to form the data used by OLAP software.