Transaction Processing Systems

Introduction

- **Transaction** = an event that determines either a manual or computer-based activity.
- **Transaction processing systems (TPS)** = used for the processing of everyday transactions that produce large amounts of data, for example, sales outlets or retail stores.

**Characteristics of transaction processing systems**

TPS collect, store, modify and retrieve the transactions of an organisation. Main processes = collecting & storage.

The 4 characteristics of TPS include:

→ **Rapid Response** = Fast performance with a rapid response time is critical.
→ **Reliability** = Failure rate must be very low
→ **Inflexibility** = a TPS wants every transaction to be processed in the same way regardless of user. If TPS were flexible there would be too many opportunities for non-standard operations (less consistency).
→ **Controlled processing** = the processing in a TPS must support an organisation’s operations.

A TPS minimises the organisation’s costs by reducing the data that must be handled. There are 2 types of transaction processing: Batch (time delay) and Real-time (immediate).

**Batch Processing**

- **Batch Processing** = collects data as a group, batch and processes it later all at once.
It has a time delay. E.g. clearance of cheques.
It is carried out by large organisation’s using a mainframe\(^1\) or mid-range\(^2\) computer.
It involves a large batch of an identical data type, such as payroll or stock information.
Batch programs are often run at night when there is less demand for the information system.

There are 3 disadvantages:
- All processing must wait until a set time. The processing schedule is predetermined.
- Errors cannot be corrected during processing
- Sorting the transaction data is expensive and time consuming.

Less commonly used today. Below is a diagrammatic representation of batch processing:

---

1 A mainframe computer is a central computer for a large number of users. It is more powerful than a midrange computer and often has thousands of terminals connected to it. Mainframe applications include payroll computations, accounting and airline seat reservations.

2 A midrange is a central computer that performs the processing for a number of users working at terminals. A terminal is an input/output device (usually a keyboard or screen). Midrange computers are typically used for accounting, database management and specific industry applications.
The Advantages:
- Can be run as a regularly scheduled event or when enough data has been collected.
- Can be run as a fully automated process, without the need for a human operator.

**Real Time Processing**

- **Real Time Processing** = the immediate processing of data. It provides instant confirmation of a transaction but does not require access to an online database.

It involves immediate processing and updating of an online database.
It involves using a terminal or workstation to enter data and display the results of the TPS.
It involves a large number of users who are simultaneously performing transactions to change data.
Two common examples: airline reservation systems and banking systems.

Two main concerns:
- **Concurrency** – ensures that two users cannot change the same data at the same time.
  One user cannot change a piece of data before another user has finished with it.³

- **Atomicity** – ensures that all of the steps involved in a transaction is completed successfully as a group. If any step fails, no other step should be completed.⁴

Main disadvantages:
- Expense
- More hardware and software to install.

Below is a diagrammatic representation of real time processing:

³ Example: if an airline ticket agent starts to reserve the last seat on a flight, then another agent cannot tell another passenger that a seat is available.

⁴ Example: a banking transaction may involve 2 steps: withdrawing money from a cheque account & transferring it into a savings account. If the first step (withdrawal) succeeds, then the second step (transfer) must succeed otherwise, the entire transaction is abandoned. This ensures the transaction isn’t recorded twice & the transfer only once.
- More computer operators are required in real time as operations are not centralised. Real time is more difficult to maintain.

**Data Validation**

- **Data Validation** = used to check the entry of transaction data.
- Involves ensuring transaction correct & have been accurately stored.
- Involves:
  - **Transaction Initiation**
    - Used in real time to eliminate a number of possible errors.
    - Used to acknowledge that a TP monitor is ready to receive the transaction data.
    - Some TPSs add an entry time to the transaction data to trace data if lost.
  - **Field Checking**
    - Occurs when the transaction data is entered into a database
    - Data is organised into files, records, fields and characters.
    - Data validation is carried out by checking the fields, using a range check, list check, type check or check digit.

**Historical Significance of transaction processing systems**

- First type of information system
- First commercially available electronic computer = UNIVAC used by US Bureau of Census
- Could process both numerical & alphabetical calculations with ease.
- UNIVAC was used by organisations to batch process business transactions, such as paying employees and recording customer purchases & payments.
- Early computers used batch with punch cards or tapes.

**Examples of transaction processing systems**

**Components of a transaction processing system**

There is a human element to TPSs.

- **Users** of a TPS often take data provided by TPS and use it in another information system.
- **Participants** people who conduct the information processing. They need to know what to do, how, and when.
- **People** from the environment are becoming participants in real time processing systems as they directly enter transactions and perform validation.

**Examples of real-time transaction processing**

3 examples of real time processing

- **Reservation systems**
  - Used extensively in any type of business involved in setting aside a service or product for a customer to use at a future time, like travelling or booking a seat on a train.

---

5 Universal Automatic Computer

6 Example: a POS system could provide stock inventory that is used by an automated manufacturing system. The users of the other information systems belong to the same organisation that owns the TPS.

7 Example, if you withdraw from ATM, you are the participant of the system.
- This system requires acceptable response time because the transactions are made in the presence of the customers.
- Some are decision making
- Hundreds of terminals must be connected at the same time.
- The Steps for a typical motel reservation system are as follows:
  i. Answer customer enquiries about room availability and pricing for a certain period
  ii. Place a reservation on one or more rooms & receive confirmation of that reservation
  iii. Perform a cancellation of a reservation
  iv. Take up the reservation on arrival
  v. Close off a customer’s account at the end of the stay and provide an invoice
  vi. Perform consolidation activities, such as a ‘night audit’ in which transactions for the day are posted to the general ledger
  vii. Provide a report and information management functions.

- The motel system consists of participants, data/information and IT.
- Night audit is generally done by batch processing.

>> **Point of Sale terminals (POS)**
- Used by retail stores to sell G+S
- POS terminals send inventory data to a central computer\(^8\) when the sale is made
- Immediate processing minimises the costs of batch handling.
- To accomplish centralised processing, data converted to form that can easily be transmitted
- Advantage = correct price of product is received once the product number is entered. (usually by barcodes).

*Context diagram of POS.*

---

\(^8\) Central computer usually a mainframe or mid-range computer that does the processing for the entire chain.
System flow chart for point of sale system:

Library Loans
- Used to keep track of items borrowed from the library.
- Barcode is scanned, this info and data are immediately recorded on the library database.
- Library loan system completes each transaction in real time.
- Has many similarities with reservation system (operational activities like- info on products, availabilities, usage and maintenance).
- Items stored in a special database called a data warehouse.

Examples of batch processing
3 examples of batch processing:

Cheque Clearance
- Cheque is a written order asking the bank to pay a certain amount of money to a particular person.
- When a cheque is issued, he or she deposits it into a bank account.
- Money cannot be withdrawn until cheque has cleared.
- The cheques are cleared as a group during the bank’s quite period (usually 1-3 pm)

Bill Generation
- Organisations create a bill, or invoice, for G+S that have been supplied to a customer.
- Usually generate a group of bills at a scheduled time – enables user to manage their time & results in less disruption to main database.
- Not done immediately, but as a group.
- A bill cannot be performed until billing period is completed, and then bill is posted.
- Below is a data flow diagram for collection of bills.

>> **Credit Card sales transactions**
- Credit card transactions are batch, even though appearing to be real time. Only the online checks if stolen, cancelled or over the card limit are real time.
- The remainder of the transaction, that is updating stage (money transfer and calculation of new balance), is performed batch anything from hours to days.
- Added to this, older-style machines produce carboned impressions of the credit card on paper, and are sent in every day or two as a batch.

**Storage and Retrieval**

TPS requires efficient method for storage & retrieval of data.

Data is stored in a database or data warehouse.

Storage system requires well-designed backup & recovery procedures.

**Databases and Files**

Storage & Retrieval depend on databases & files.

- **Database** = an organised collection of data.
  - An organisation stores all accounting and operational records in database, often called an *operational database*.
  - Database = model of organisation’s operational aspects.
  - Data in operational database is defined in a schema.
  - TPS restricted view.
  - Designed using:

  - *Hierarchical database*: organises data in a series of levels. It uses a top-down structure consisting of nodes and branches.
b) *Network database:* organises data as a series of nodes linked by branches. Each node can have many branches & each lower level node may be linked to more than one higher node.

c) *Relational database:* organises data as a series of tables. Relationships are built between the tables to provide a flexible way of manipulating & combining data.

- When designing a database the following features are important:
  - *Good data placement.*
    Large numbers of users are simultaneously performing transactions to change data.
    Database should be designed to access patterns of data use & to place frequently accessed data together.
  - *Short Transactions.*
    Keeping transactions short enables the entire transaction to be processed quickly, which improves concurrency.
    User interaction during transaction processing avoided, slows down system
  - *Real-Time backup*
    Characterised by continuous operations with downtime kept minimum.
    Backing scheduled during times of low activity.
  - *High Normalisation*
    Redundant information kept minimum – increasing speed of updates & improve concurrency.
    Reducing speeds back-ups as there is less data.
  - *Archiving of historical data*
    Data that is rarely referenced should be archived into separate databases or moved out of the heavily updated tables.
    This keeps tables small, improving backup times & query performance.
  - *Good Hardware Configuration*
    The hardware needs to be able to handle a large number of concurrent users and to provide quick response times.
File = block of data.
- In a database, a file is divided into a set of related records.
- In a record it contains specific information such as customer or product.
- Every TPS uses files to store and organise its transaction data.
- Batch and real require different methods of storage and retrieval.
- In TPS there are 5 basic types of files:
  - **Master File**
    Contains info about business situation. Master file stores the operational database.
    Transaction data stored in master file.
  - **Transaction File**
    Collection of transaction records.
    Data in transaction file used to update master file.
    Transaction file also audits trails and history for the organisation
  - **Report File**
    Contains data that has been formatted for presentation to a user.
  - **Work File**
    Temporary file in the system used during processing
  - **Program File**
    Contains instructions for the processing of data.
    Created from something like Visual Basic and C++

**Data Warehousing**

- **Data warehouse** = a database that collects information from different data sources.
  Data gathered in real time can be used for analysis if stored in a data warehouse.
  Volume of data is a problem in analysing a database that is continually being updated. Solution is to
  have periodic downloads of data into a separate database set especially for analysis – data warehouse.
  The process of using analysis tools to find patterns and trends in that data is *data mining*.

This means that a ‘snapshot’ of the transaction database can be collected at any time, and is on-hand to
analyse database performance itself, as well as trends.

Data warehouse provides data that is:
- **Consolidated (merged)**
  Data is organised using consistent naming conventions, measurements, attributes and semantics.
  Organisations can use similar data in different formats.  
  Data is data warehouse is stored in a single, acceptable format.
  Data warehousing allows data from all over the organisation to be effectively used in a consistent manner.

---

9 Example, True or false data can be represented as one/zero, on/off, pos/neg.
- **Subject-orientated**
  A large amount of data is stored across an organisation. Some is irrelevant for executive reporting & makes querying the data difficult. A data warehouse organises only the key business info from operational sources so that it is available for analysis.

- **Historical**
  Real time systems represent current value. Do not show inventory at some time in past. Querying stock inventory a moment later may return a different response. However, data stored in a data warehouse is accurate for a specific moment in time, as it represents its historical information and cannot change. Data warehouse stores a series of snapshots of an organisation’s operational data generated over a long period of time.

- **Read-only**
  After data has moved to data warehouse – it does not change unless data incorrect. As it represents a particular point in time – it must never be updated. Only operations that occur are loading and querying data.

**Backup procedures**
- Organisations now dependent on TPS
- Well designed backup and recovery procedures minimises disruptions when TPS goes down.
- A backup is another copy of data, could be used to rebuild the system, if system down, the recovery process rebuilds the system.
- Success of backup and recovery depends on implementing appropriate procedures.
- In general, the more valuable the data the:
  i. More often the files should be backed up
  ii. More copies should be made
  iii. Greater number of locations where the backups should be stored.

  The range of different methods include:
* **Full Backups** *(Grandfather-father-son)*
  Longest to perform but contain all that is needed. 3 generations of files may be kept for successive backups and then cycle repeats. Allows 2 previous versions of files to be stored in case of error from one generation to next. Commonly used with magnetic tape. Most recent is son. Son destroyed, refer to previous generation.

* **Partial backups**
  Can be:
  i. Differential – which have full copy of files from time to time, as well as regular copies that have changed since last backup
  ii. Incremental – which are copies of all files that have changed since last full backup.

* **Off site Backup**
  Used in case of vandalism or disaster (like fire)

---

10 Deletes, inserts and updates not applicable.
• **On site**
  Secured backup, requires several levels of security including fire proof safe.

• **Backup media**
  CD's etc..
  Reliability, cost, file size will determine method used.

• **Magnetic Tape**
  Can store large quantities of data inexpensively.
  Magnetic tape is long thin plastic, coated with thin layer of magnetic material.
  Tape wound on two reels inside.
  Disadvantage – sequential.

• **Backup Software**
  Designed to manage copying of selected files to backup media, such as tapes, generally offers automated scheduling as well as options for full, differential or incremental backups.

>> **Recovery Process**

To cope with failures, TPS must detect & correct errors.

The recovery of a database involves:

• **Backup**
  Typically done twice a day. Copy protected.
• **Journal**
  Maintain an audit trail of transactions and database changes.
  Transaction logs record essential data for each transaction such as data values, time of transaction and terminal number.
  Journals is the database change log. Contains before & after copies of records that have been modified by transactions.

• **Checkpoint**
  DBMS periodically suspends all processing to synchronise its files and journals.
  Transactions in progress are completed, and journals updated. System then in ‘quiet state’ and the database with transaction logs is synchronised.
  DBMS then writes a special record to the transaction file – ‘checkpoint record’ – which contains info necessary to restart a system. This is taken frequently, when failures occur it is possible to resume processing from most recent checkpoint.

• **Recovery manager**
  Restores database to a correct condition & restarts transaction processing.

Two types of recovery procedures:

*Backward recovery* = used to undo unwanted changes to the database, due to aborted transactions.
*Forward recovery* = recovery begins with the backup copy and processes all the transactions from the partial back-ups in order, through to the most recent. This method is faster & more accurate.

**Updating in a batch**

Updating a batch used when transactions are recorded on paper\(^{11}\) or stored on magnetic tape.
Transactions collected & updated in a batch when it is convenient or economical to process them.

Historically, updating a batch was the only feasible method when transaction details were stored on punch cards or magnetic tape. IT did not exist for immediate processing.

Two stages in batch processing:

i. Collecting and storage of transaction data in a transaction file. Involves sorting that data in sequential order.

ii. Processing of the data by updating master file, may involve additions, updates and deletions that need to happen in a certain order. One error occurs in the batch, then entire batch rejected.

IT in batch requires secondary storage medium that can store large amounts of data inexpensively – usually use Magnetic tape.

Steps involved in batch update are:

• *All changes are made to the transaction file* additions, alterations, deletions and so on are stored.

• *Transaction file is sorted* by a key field to be in the same sequential order as the master file

• *Files are matched* by key field and the relevant records in the master file updated

**Hardware** – large capacity tape drive or CD-RW, DVD-RAM etc…

---

\(^{11}\) Such as credit card slips
Software – does not have to allow online work or simple user interface as only well-trained and experienced participants use.

>>> Updating in real-time systems
Online connections have allowed for real time or interactive processing systems.
Random access to data = instant responses possible once a primary key is recognised & request is sent to the online database in a master file.

Steps involved in real-time update:
- Enter
- Retrieve
- Update
- Send

Hardware – includes large capacity secondary storage (such as hard drives), with direct access files so that response time is quick (several seconds or less).
Software – must enable online work for multiple simultaneous use, as well as a simple user-friendly interface, since many participants will use it & often have only brief on-the-job training.

► Real Time processing = immediate processing of data. Provides instant confirmation of a transaction. Involves a large number of users simultaneously.

Other information Processes
Collecting
Collecting involves generating transaction data. 12

>>> Hardware
Includes MICR readers, ATMs and barcode readers.

- MICR (magnetic ink character recognition)
  Systems used by banks to read account numbers on cheques.
  Characters printed using magnetic ink that contains magnetised particles.
  MICR processes cheques at speeds of up to 2000 per minute.
  This is an example of batch processing.

- ATM (automatic teller machine)
  Banking terminal that performs common transactions, such as deposits & withdrawals.
  ATM accesses a communication system.
  Data & information transfers between ATM terminal & bank’s central computer in real time.

- Barcode readers
  Used to collect product information at point of sale.
  Often use laser to read barcodes.
  Product information (description, price and code) is held on central computer linked to POS terminal.

---

12 Example, people using an ATM generate transaction data by entering their debit card numbers & typing their requests on a keyboard.
**Forms**

- **Form** = document used to collect data from a person.
  Form completed, it is processed in batch or real-time.
  Designed to limit responses & minimise data entry errors by providing good format for user.

3 types of forms:
- **Paper** - written on paper & batch processed later.
- **Online/on-screen** – completed for data entry to populate databases, and can be processed in real time if needed. E.g. Commonwealth Bank website. User can view, enter & change data any time. Well-designed form provides information explaining the required data.
  Forms can minimise data entry errors by automatically filling in previously stored data, such as a customer’s address, once the user has entered the customer’s name.
- **Web forms** – for users purchasing on net. User’s details become data in database.

**Analysing Data**

Output from a TPS is input to other types of information systems, such as decision support systems and management information systems.

**Decision support system**

- **Decision support systems (DSS)** = assist people to make decisions by providing information, models and analysis tools.  
  Problem – amount of data processed by transaction processing systems growing ∴ more complex and harder to make decisions, whereas before there were no precedents.

- **Data Mining** = used in DSSs to find relationships and patterns in the data stored in a database.
  Sorts through data & connects data – allows for better decision making.

Data models using ‘what-if’ scenarios are used where trends are not clear or change unpredictably, so that a decision can be made.
Models follow mathematical principles based on independent variables which are governed by inputs.

**Management information systems (MIS)**

- **Management information system** = provide information for the managers. Presents basic facts about the performance of the organisation.
  Common examples: reports on sales, stock inventory, payroll, orders and budgets.
  MIS generally report.
  MIS shares data from TPS which is then manipulated to create reports.
  These reports assist managers in their decision-making.

Several types of reports:
- **Period reports** – produced on a regular schedule, for use by all levels of management.
- **Detail reports** – organised list relating to a product or an activity to be performed.
- **On-demand reports** – produced in response to an unscheduled request for information

---

13 Example, a business uses TPS to process its sales transactions. It uses DSS to periodically summarise its sales data by date, region and product. This summary information is stored in a separate database to be analysed by senior management. To make decision, management need to see trends quickly by querying data.
• **Summary reports** – generally combine data, showing totals over different areas or times for tactical & strategic planning.

• **Exception reports** - contain information which is outside normal range and is used to alert management to some situation which is either unfavourable (like a critical shortage of some component) or may require special handling (like one area suddenly having large numbers of late payments).

## Issues related to transaction processing systems

### Changing nature of work

Automation of jobs and people from the environment completing the transaction have affected the nature of work in many organisations.

>>> **Automation of jobs**

- **Automation of jobs** = use of IT to perform tasks once performed by people.  

Workers are required to learn new skills & complete ongoing training.

Automation = loss of jobs in one industry, but replaced by growth in another.

>>> **People as participants**

People = participants in TPSs as they directly enter transactions. E.g. ATM.

Internet buying will likely have similar effects on shop assistant numbers if this form of transaction keeps growing – but opens opportunities for more workers.

### Non computer procedures

Computer TPS not available, need manual backup, like a library loaning books – they must record on paper if system not working.

### Bias

- **Bias** = systematic inaccuracy due to methods used in collecting, processing & presenting data. Bias means data is unfairly skewed or gives too much weight to a particular result.

Data gathered from TPS can be presented in tables in a biased way.

Generally an issue during the processing stage and presentation stages.

Deliberate bias = ethical issue.

### Importance of data

Organisations need procedures to ensure data is secure, accurate and valid.

>>> **Data Security**

- **Data Security** = involves a series of safeguards to protect data.

Data is under threat by:
  - Being stolen
  - Destroyed
  - Maliciously modified

There is a greater risk when the data is accessible to multiple users.

14 Example, the POS terminal once performed by people in a manual transaction system, such as memorising price of products.
At the lowest security level passwords, biometric methods such as fingerprints and retinal scans work in that they keep unskilled people from accessing data.

Next level of security is based on:
- **Data Encryption** = this is the coding (changing of characters according to a pattern) of the data for transmission and storage, so that a complex decryption process, based on a key, is required before the data can be read.
- **Firewalls** = these used on networks & operate by checking passwords for each user, as well as verifying & authenticating all incoming data to the main computer. More than one the firewall may be installed, providing a barrier at several levels.

**Data accuracy**
- **Accuracy** = extent to which it is free from errors.
Data entered into TPS not always accurate. This could be costly.

Errors may occur at many stages, as shown below:
- **Data Collection errors**
Where errors are made on or inputted from collection forms because of poor writing misreading etc..
- **Data entry errors**
Where the in putter makes a typo, takes info from the wrong form etc…
- **Out of date errors**
Which occur when a person moves, changes their name on marrying etc…
- **Mismatching errors**
Where data relates to a different person with similar name.

Steps for improving data accuracy include:
- **Careful design/wording of data collection forms to limit possible errors**
- **Checking data accuracy at data entry time (also known as data validation)**
- **Missing or out-of-range data should be detected as errors during processing by TPS.**

**Data Integrity**
- **Data Integrity** = describes the reliability of data. Involves accuracy, currency and relevance of data.
Data integrity in real-time transaction processing is provided when the transaction passes the ACID test.

- **Atomicity** occurs when all of the steps involved in a transaction are completed successfully as a group. If any step fails, no other step should be completed and is abandoned. If some operations succeed & others fail, there is no atomicity.
- **Consistency** occurs when a transaction successfully transforms the system and the database from one valid state to another. Consistency in a TPS stems from the correct application programming, such as always debiting & crediting the same amount.
- **Isolation** occurs if a transaction is processed concurrently with other transactions and still behaves as if it were only transaction executing the system. Transactions must not interfere with each other’s database updates.

---

15 Example, if the price of a product has been entered incorrectly into the database, then customers buying that product will be charged at the wrong price.
16 If it is still correct after collection some time ago
- *Durability* occurs if all the changes that a transaction makes to the database become permanent when the transaction is committed.

These ACID properties guarantee that a transaction is never incomplete, the data is never inconsistent, concurrent transactions are independent, and the effects of a transaction are permanent.

**Control in transaction processing**

Control is critical because volume of data is immense & people only see parts of the operation. ∴ Errors and problems easily occur, though not be noticed, if control measures are no adhered to.

Management should not depend solely on the output from the TPS, but try to make decisions & plans incorporating outputs from both management information and decision support systems.

People in management have created false transaction data to promote their careers, just shows results of TPS aren't always correct. ∴ There should be less dependence.