Lecture 01

Overview of Databases & Transaction Processing

Lecturer: Sael Lee

Slide adapted from the author's slides and Dr. Ilchul Yoon's slides.
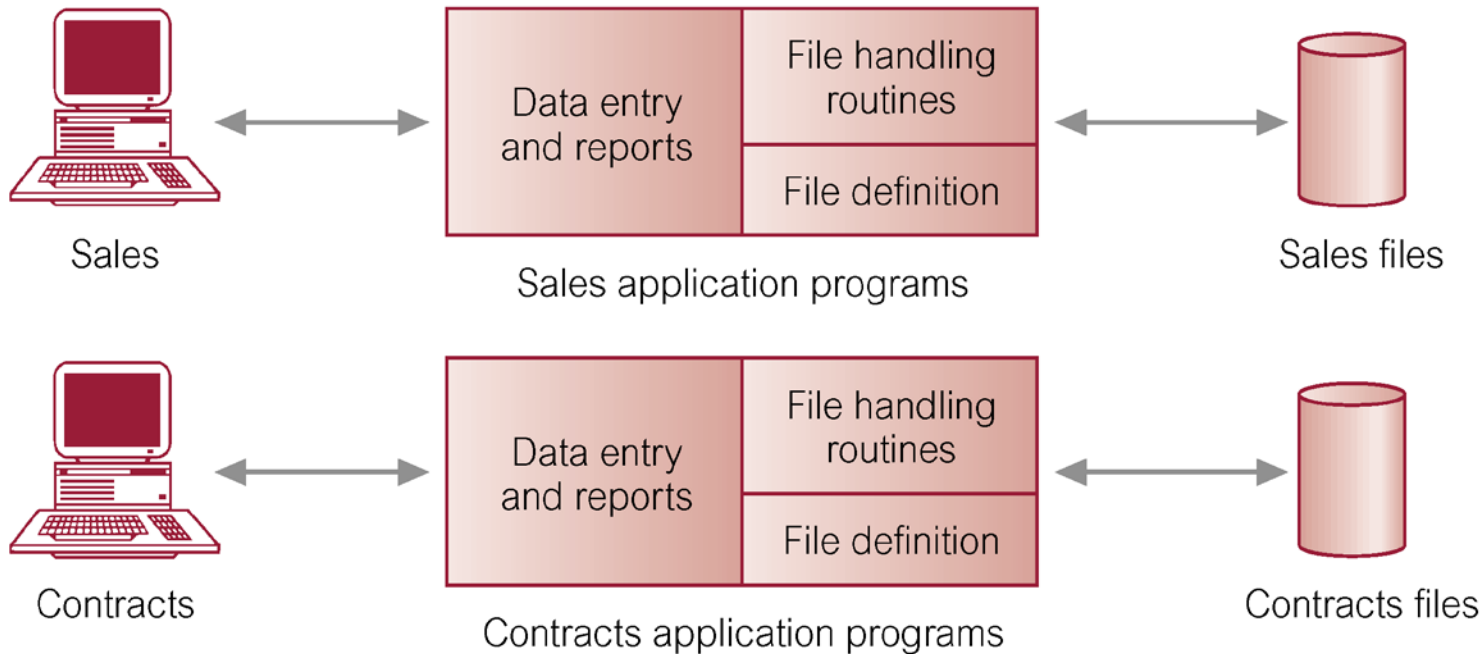
What is a Database?

- Collection of data central to some enterprise
- Essential to operation of enterprise
  - Contains the only record of enterprise activity
- An asset in its own right
  - Historical data can guide enterprise strategy
  - Of interest to other enterprises
- State of database mirrors state of enterprise
  - Database is persistent

File-based Processing

Sales
Sales application programs
Sales files

Contracts
Contracts application programs
Contracts files

Data entry and reports | File handling routines | File definition

Sales Files

PropertyForRent (propertyNo, street, city, postcode, type, rooms, rent, ownerNo)
PrivateOwner (ownerNo, fName, lName, address, telNo)
Client (clientNo, fName, lName, address, telNo, prefType, maxRent)
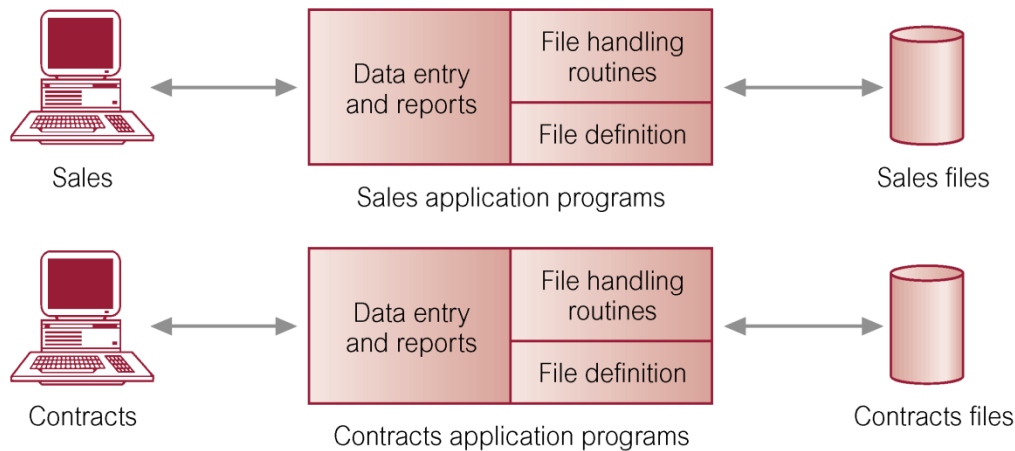
Contracts Files

Lease (leaseNo, propertyNo, clientNo, rent, paymentMethod, deposit, paid, rentStart, rentFinish, duration)
PropertyForRent (propertyNo, street, city, postcode, rent)
Client (clientNo, fName, lName, address, telNo)

File-based Processing

- What data is being used by the 2 applications?
- Do they share source code?
- Do they share data?

Sales Files

PropertyForRent (propertyNo, street, city, postcode, type, rooms, rent, ownerNo)
PrivateOwner (ownerNo, fName, lName, address, telNo)
Client (clientNo, fName, lName, address, telNo, prefType, maxRent)

Contracts Files

Lease (leaseNo, propertyNo, clientNo, rent, paymentMethod, deposit, paid, rentStart, rentFinish, duration)
PropertyForRent (propertyNo, street, city, postcode, rent)
Client (clientNo, fName, lName, address, telNo)

4
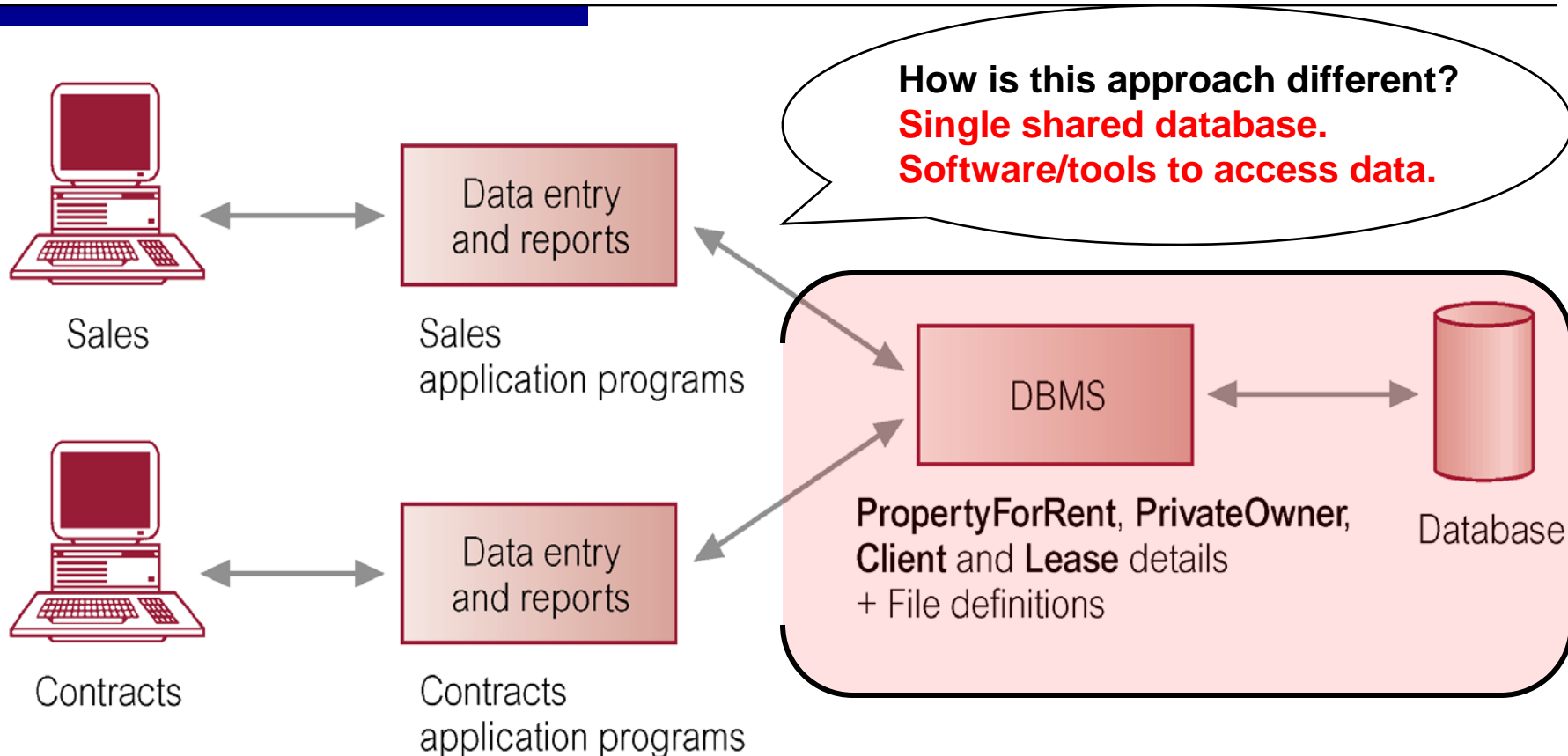
Limitations of File-based Approach

-

Think of limitations from the perspective of data management and consistency

-

Limitations of File-based Approach

-

-

Think of limitations from the perspective of application development and maintenance

-

Database Management System (DBMS)

How is this approach different?
Single shared database.
Software/tools to access data.

PropertyForRent (propertyNo, street, city, postcode, type, rooms, rent, ownerNo)

PrivateOwner (ownerNo, fName, lName, address, telNo)

Client (clientNo, fName, lName, address, telNo, prefType, maxRent)

Lease (leaseNo, propertyNo, clientNo, paymentMethod, deposit, paid, rentStart, rentFinish)

Database Approach

- Objective:
  - Define data independent of application programs
  - Provide independent access to the data to all applications and software
- Result
  - Database + Database Management System (DBMS)

- Database Management System (DBMS):
  - A program that manages a database
  - Supports a high-level access language (e.g. SQL)
    - Application describes database accesses using that language.
    - DBMS interprets statements of language to perform requested database access.

Transaction and TPS

- Real world event → corporate database update
  - Transaction is used for such updates
  - Typically... real-time operation

- A transaction is:
  - an <u>application program with special properties</u> to guarantee database correctness after execution

- Transaction Processing System (TPS)
  - TP monitor + databases + DBMS + transactions
  - TP monitor and DBMS together guarantee the special properties of transactions
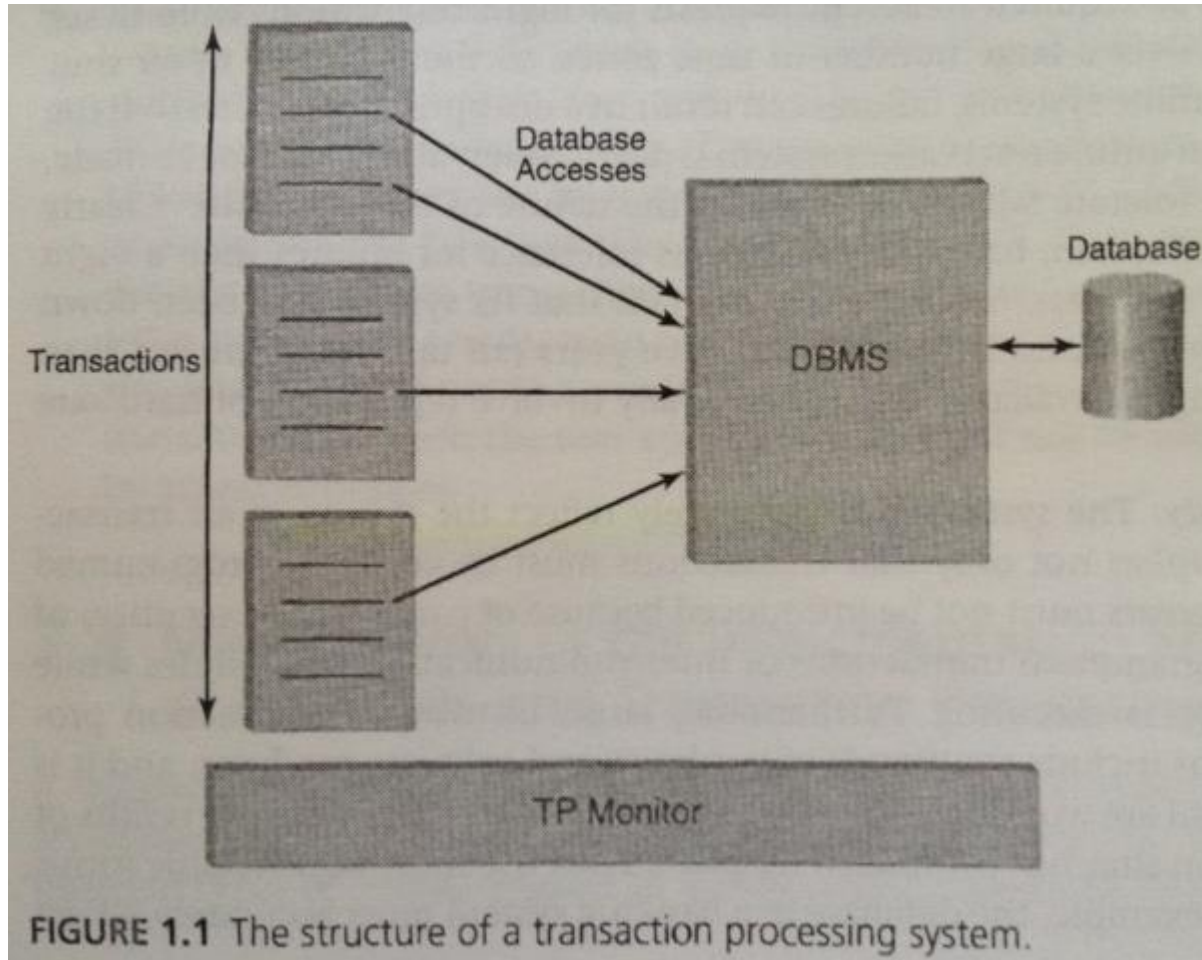
TPS – Figure 1.1

FIGURE 1.1 The structure of a transaction processing system.

TPS Requirements

- High Availability
- High Reliability
- High Throughput
- Low Response Time
- Long Lifetime
- Security

Roles in the Database Environment

- System Analyst – business description
- Database Designer – data structure in database
- Application Programmer
- Database Administrator (DBA)
- System Administrator

SUNY Korea

History of Database Systems

- First-generation
  - Hierarchical and Network

- Second generation
  - Relational

- Third generation
  - Object Relational
  - Object-Oriented

Advantages of DBMS

- Control of data redundancy
- Data consistency
- More value (higher ROI) from the same data amount
- Sharing of data
- Improved data integrity
- Improved security
- Enforcement of standards
- Economy of scale

SUNY Korea

Advantages of DBMS

- Balanced conflicting requirements
- Improved data accessibility and responsiveness
- Increased productivity
- Improved maintenance through data independence
- Increased concurrency
- Improved backup and recovery services

Disadvantages of DBMS – Challenges

- Complexity
- Size
- Cost of DBMS
- Additional hardware costs
- Cost of conversion – training, hiring specialist, …
- Performance
- Greater impact of a failure

SUNY Korea

OLTP vs. OLAP

- **On-line Transaction Processing (OLTP)**
  - Day-to-day handling of transactions that result from enterprise operation
  - Maintains correspondence between database state and enterprise state

- **On-line Analytic Processing (OLAP)**
  - Analysis of information in a database for the purpose of making management decisions

OLAP

- Queries on historical data

- Large data volume

- Often use a data warehouse
  - Data Warehouse - <u>(offline) repository</u> of historical data generated from OLTP or other sources
  - Data Mining - use of warehouse data to *discover* relationships that might influence enterprise strategy

OLTP, OLAP, and Mining
-- an example: Supermarket

- OLTP
  - Event is 3 cans of soup and 1 box of crackers bought; update database to reflect that event
- OLAP
  - Last winter in all stores in northeast, how many customers bought soup and crackers together?
- Data Mining
  - Are there any interesting combinations of foods that customers frequently bought together?