Database Management System (DBMS)

Dr. Rahul Kumar Verma
Department of Computer Science

Number of “Credits” in this course

- Lecture – 3 Credits (150 Marks)
- Tutorial – 0 Credits
- Lab – 1 Credit (100 Marks)

Total – 4 Credits

Evaluation Scheme

Components	Weightage	
Quizzes	40 Marks	150 Marks for Theory
Class Participation	10 Marks	
Mid-Term Exam	30 Marks	
End-Term Exam	70 Marks	
Continuous Lab Evaluation	100 Marks	

Books

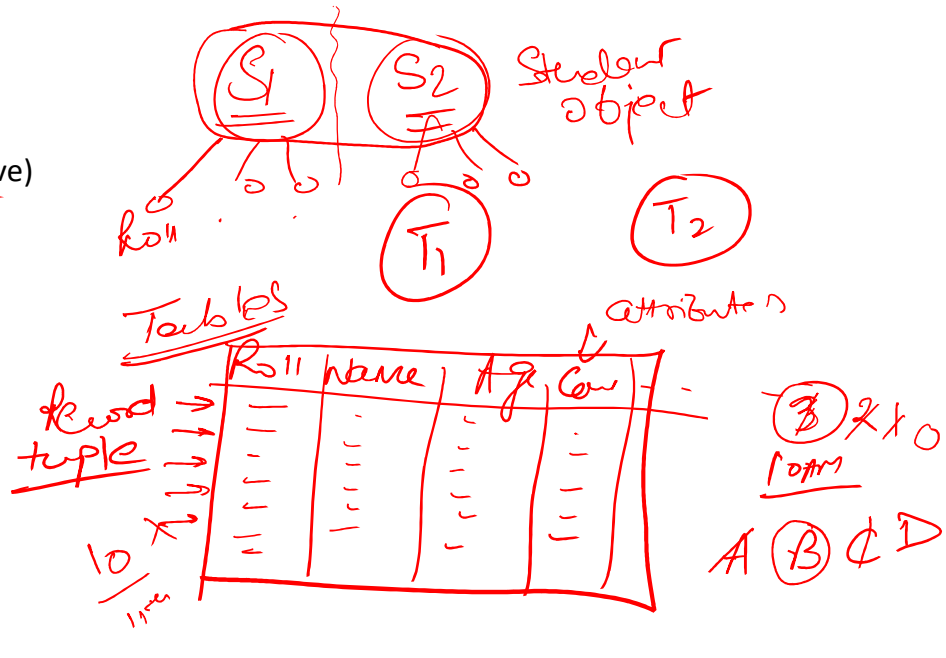
- Fundamentals of database systems (Ramez Elmsari · Shamkant B. Navathe)
- Database System Concepts (Avi Silberschatz · Henry F. Korth · S. Sudarshan)

Outlines

- Introduction of database
- File processing system
- Disadvantages of file processing system
- Transition to Database management system
- Characteristics of the database management system
- MCQs
- Summary

What is Database Management System (DBMS)?

- Data – **Facts** about an entity that can be recorded or stored
  - e.g. Person Name, Age, Gender and Weight etc.
- Information – Processed, meaningful and usable data (depends on perspective)
- Database - Collection of **logically related data (i.e., record)**
  - Traditional Database (TDB): Text and Number
  - Multimedia Database (MDP): Video, Speech, song, movie, etc.
  - Geographic Information System (GIS): Images of earth
  - Real-Time Database (RTD): Production, Supermarket – varying products data
  - Data Warehouse (DW): Huge and Historical Data.
- Management - Manipulation, Searching and Security of data
  - e.g. Viewing result on website, etc.
- System - **Programs** or **tools** used to manage database
  - e.g. SQL Server Studio Express, Oracle etc.
- Database Management System (DBMS) – Set of programs or **software used to define, manipulate, retrieve and manage data in a database**.
  - e.g. MS SQL Server, Oracle, My SQL, SQLite, MongoDB etc.
- DB+DBMS = DBS (Database System)

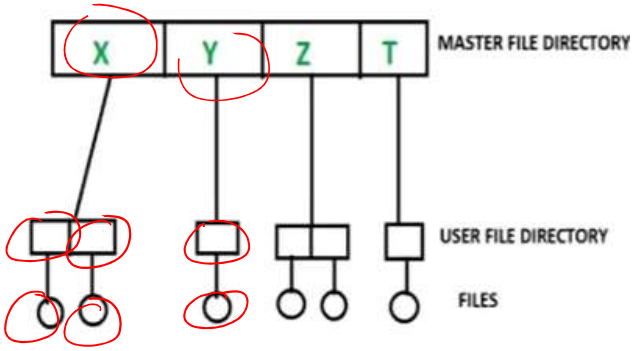


Introduction

- A number of records created or exists at various places, like:
  - Universities: students registration, grades, etc., form students records
  - Banking: transactions records
  - Airlines: reservations detail records
  - Sales: customers, products, purchase
  - Online Retailers: Order tracking
  - Many more
- All these records are to be stored and managed properly for their efficient utilization.
- Data Base Management System (DBMS) helps in all those areas where handling of the data is required.
- First of all, we will start with traditional data handling approach : File processing system.

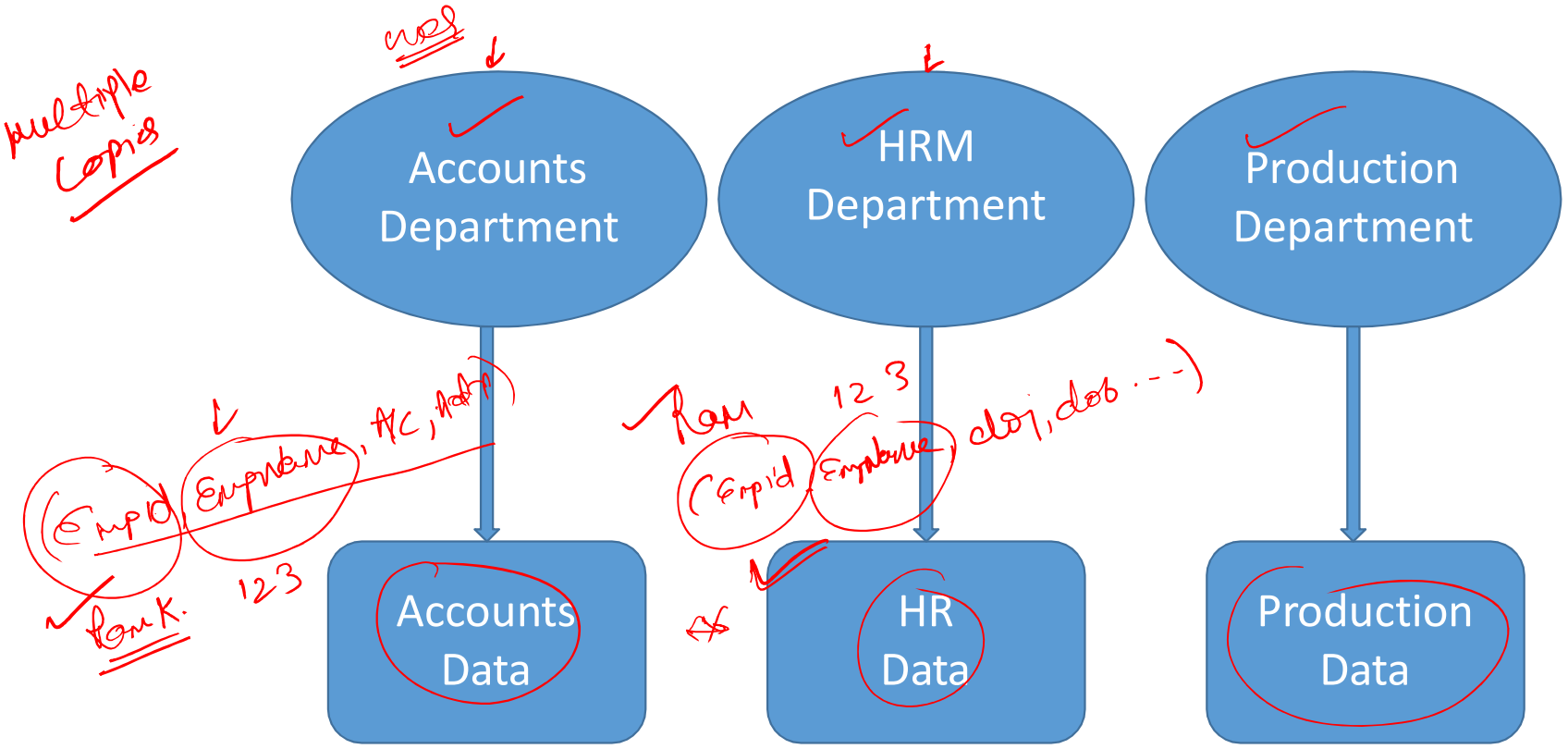
Database System vs. File-processing system

- A file processing system is a software that stores and manage files in computer hard disk.
  - It allows access to single files or tables at a time.
  - Data is directly stored in set of files.
  - It contains flat files that have no relation to other files.
  - File system consists of different files which are grouped into directories. The directories further contain other folders and files.
  - E.g. NTFS (New Technology File System), EXT (Extended File System)



- DBMS
  - A Database Management System (DBMS) is a software that allows user to efficiently define, create, maintain and share databases.
  - E.g. Oracle, MySQL, MS SQL server, IBM DB2, MS Access, dBASE, SQLite, etc.

How File Processing System Works?



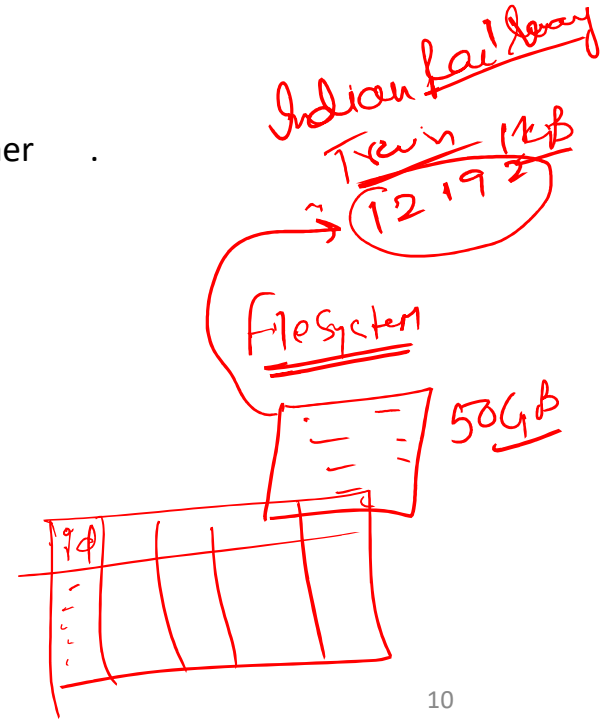
Each department maintain their own set of data. There is no link between those data pools.

Example for file-processing system

- University records,
  - The grade reporting office,
    - may keep files of students and their grades.
    - Programs to print a student's transcript and to enter new grades.
  - The accounting office,
    - may keep track of students' fees and their payments.
- Although both users are interested in data about students,
- Each user maintains separate files—and programs to manipulate these files.
- This make redundancy in defining and storing data

Disadvantages file-processing system

- Data redundancy and inconsistency
  - Redundancy occurs when same piece of the data is held in two or more separate places.
  - Inconsistency occurs when same data is kept in different format or have different values at two or more places.
  - Data redundancy leads to data inconsistency.
  - All copies may not be updated properly.
- Difficulty in accessing data
  - It does not allow needed data to be retrieved in a convenient and efficient manner .
  - More responsive data-retrieval systems are required .
- Data isolation
  - Data is scattered in various files, and files may be in different formats.
  - It is very difficult to handle.
- Integrity problems
  - The data values stored in the database must satisfy certain types of constraints.
  - It is difficult to change the programs , when new constraints are added.



Disadvantages file-processing system

- Atomicity problems
  - It is essential in database that either complete query to be executed, or none, which means either all the operations in a transaction executes or none.
  - It is difficult to ensure atomicity in a conventional file-processing system.
  - Failures may leave database in an inconsistent state with partial updates carried out.
- Concurrent-access anomalies
  - For faster response, many systems allow multiple users to update the data simultaneously.
  - In such an environment, interaction of concurrent updates is possible and may result in inconsistent data.
- Security problems
  - Not every user of the database system should be able to access all the data.
  - As application programs are added to the file-processing system in an ad hoc manner, enforcing such security constraints is difficult.

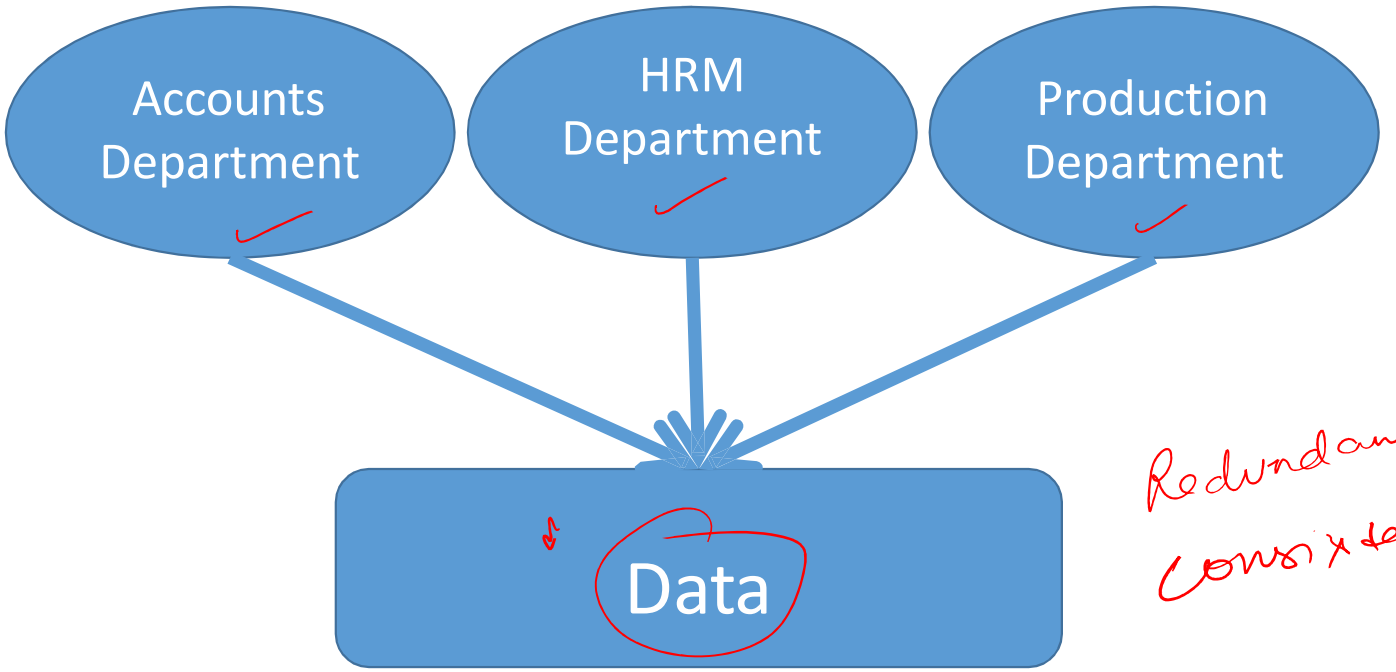
DBA

Database management system offers solutions to all the above problems

Transition phase

- These difficulties, among others, prompted both the initial development of database systems.
- The transition of file-based applications to database systems, back in the 1960s and 1970s.
- We shall see the concepts and algorithms that enable database systems to solve the problems with file-processing systems.

How DBMS works?



Redundancy ✗
Consistency ✓

Characteristics of the Database Approach

- In the database approach, a single repository maintains data.
- It should be accessed by various users repeatedly through queries, transactions, and application programs.
- The main characteristics of the database approach are:
  - Self-describing nature of a database system
  - Insulation between programs and data, and data abstraction
  - Support of multiple views of the data
  - Sharing of data and multiuser transaction processing

Self-Describing Nature of a Database System

- A fundamental characteristic of the database approach is
  - It should contain not only the database itself but also contains its complete definition.
- The definition contains
  - The structure of each file, the type and storage format of each data item, and various constraints on the data.
  - This information is called meta-data.

Example of a Database:

STUDENT

Name	Student_number	Class	Major
Smith	17	1	CS
Brown	8	2	CS

COURSE

Course_name	Course_number	Credit_hours	Department
Intro to Computer Science	CS1310	4	CS
Data Structures	CS3320	4	CS
Discrete Mathematics	MATH2410	3	MATH
Database	CS3380	3	CS

SECTION

Section_identifier	Course_number	Semester	Year	Instructor
85	MATH2410	Fall	07	King
92	CS1310	Fall	07	Anderson
102	CS3320	Spring	08	Knuth
112	MATH2410	Fall	08	Chang
119	CS1310	Fall	08	Anderson
135	CS3380	Fall	08	Stone

GRADE_REPORT

Student_number	Section_identifier	Grade
17	112	B
17	119	C
8	85	A
8	92	A
8	102	B
8	135	A

PREREQUISITE

Course_number	Prerequisite_number
CS3380	CS3320
CS3380	MATH2410
CS3320	CS1310

Figure 1.2
A database that stores student and course information.

Ramez Elmasri and Shamkant B. Navathe. 1989. Fundamentals of database systems. Benjamin-Cummings Publishing Co., Inc., USA.

Self-Describing Nature

Meta-data (data about data)

RELATIONS

Relation_name	No_of_columns
STUDENT	4
COURSE	4
SECTION	5
GRADE_REPORT	3
PREREQUISITE	2

COLUMNS

Column_name	Data_type	Belongs_to_relation
Name	Character (30)	STUDENT
Student_number	Character (4)	STUDENT
Class	Integer (1)	STUDENT
Major	Major_type	STUDENT
Course_name	Character (10)	COURSE
Course_number	XXXXNNNN	COURSE
….	….	…..
….	….	…..
….	….	…..
Prerequisite_number	XXXXNNNN	PREREQUISITE

Note: Major_type is defined as an enumerated type with all known majors. XXXXNNNN is used to define a type with four alpha characters followed by four digits

Figure 1.3
An example of a database catalog for the database in Figure 1.2.

STUDENT

Name	Student_number	Class	Major
Smith	17	1	CS
Brown	8	2	CS

COURSE

Course_name	Course_number	Credit_hours	Department
Intro to Computer Science	CS1310	4	CS
Data Structures	CS3320	4	CS
Discrete Mathematics	MATH2410	3	MATH
Database	CS3380	3	CS

SECTION

Section_identifier	Course_number	Semester	Year	Instructor
85	MATH2410	Fall	07	King
92	CS1310	Fall	07	Anderson
102	CS3320	Spring	08	Knuth
112	MATH2410	Fall	08	Chang
119	CS1310	Fall	08	Anderson
135	CS3380	Fall	08	Stone

GRADE_REPORT

Student_number	Section_identifier	Grade
17	112	B
17	119	C
8	85	A
8	92	A
8	102	B
8	135	A

PREREQUISITE

Course_number	Prerequisite_number
CS3380	CS3320
CS3380	MATH2410
CS3320	CS1310

Figure 1.2
A database that stores student and course information.

Ramez Elmasri and Shamkant B. Navathe. 1989. Fundamentals of database systems. Benjamin-Cummings Publishing Co., Inc., USA.