

Unit 1: Data And Information

SUSHMA VANKHEDE

Contents

Data and Information:

Limitations of Manual Data Processing,

Advantages of DBMS over traditional computer file-based processing approach,

Functions of DBA

Elements of DBMS: DDL, DML, Entities, Sets and attributes.

Data Base Tables : Keys - Primary, Secondary, Composite

Definitions

Data

Information

Database

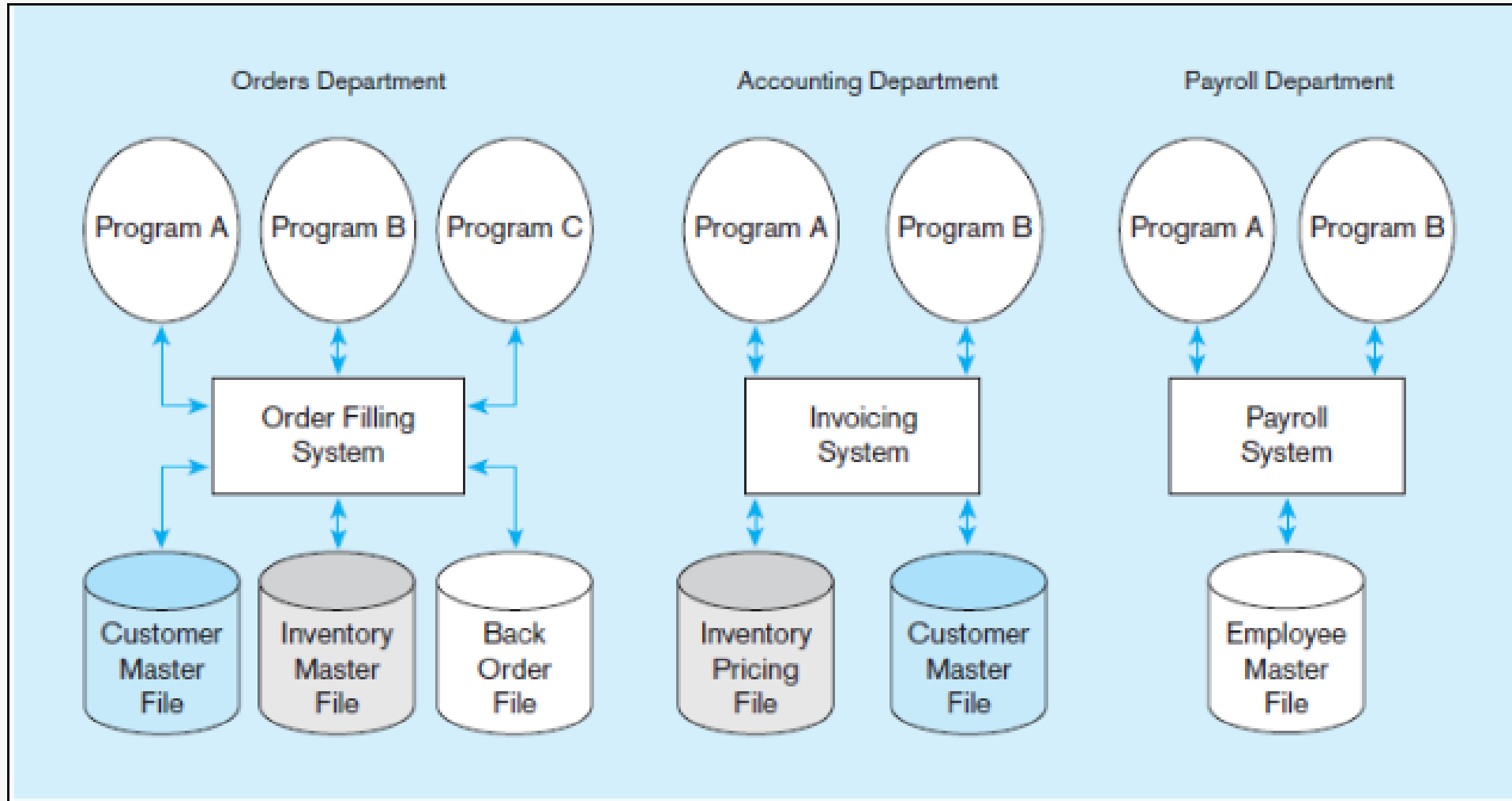
Database management system

Database System

Limitation of Manual Data Processing

- Takes up lot of space
- Prone to Damage and Being misplaced
- Hard to make changes
- Lack of security
- Access time
- Higher Cost

File Processing System



Disadvantages of File Processing System

- Data Redundancy
- Data Inconsistency
- Difficulty in accessing data
- Data Isolation
- Concurrent Access Anomalies
- Integrity
- Security

Data Redundancy

Same information May be duplicated. Leads to higher storage and access cost. Change must be reflected at all the places.

Data Inconsistency

- When data redundancy occurs data inconsistency occurs.
- Data is not updated in all the files simultaneously. When address of person changes, it should be reflected in all the files or else the data will be inconsistent as it doesn't match with each other.

Difficulty in accessing data

If we need some information urgently which is not required at time of designing, to get it we need to design another application.

Data Isolation

- Because data are scattered in various files, and files may be in different formats, writing new application programs to retrieve the appropriate data is difficult. Dependency on application programs: Changing files would lead to change in application programs.

Data Integrity problem

Integrity problem arises when the database fails to satisfy certain integrity conditions. For example, the phone number cannot be longer than 10 digits, bank balance should not go below 1000 etc. The actual problem arises when we would like to include new such conditions with the existing database.

Data Security problem

- **File processing system does not provide adequate security on data.** In some situations, it is required to provide different types of access to data for different users. For example, a data entry operator should only be allowed to enter data.

Advantages of DBMS over file processing System

- Redundancy is reduced
- Data Consistency
- Data can be shared
- Data Security
- Integrity
- Easy recovery

Functions of DBA

- Planning
- Installation
- Configuration
- Database design
- Migration
- Performance monitoring
- Security
- troubleshooting

Elements of DBMS: DDL, DML, Entities, Sets and attributes.

DDL

DML

Entities

Attributes

Data Base Tables : Keys - Primary, Secondary, Composite

A DBMS key is an attribute or set of an attribute which helps you to identify a row(tuple) in a relation(table). They allow you to find the relation between two tables. Keys help you uniquely identify a row in a table by a combination of one or more columns in that table.

1. Candidate Key: are individual columns in a table that qualifies for uniqueness of all the rows. Here in Employee table EmployeeID & SSN are Candidate keys. A table can have multiple candidate keys but only a single primary key.

2. Primary Key: is the columns you choose to maintain uniqueness in a table. Here in Employee table you can choose either EmployeeID or SSN columns, EmployeeID is preferable choice, as SSN is a secure value.

3. Alternate Key: Candidate column other the Primary column, like if EmployeeID is PK then SSN would be the Alternate key.

Data Base Tables : Keys

4. Super Key: If you add any other column/attribute to a Primary Key then it become a super key, like EmployeeID + FullName is a Super Key.

5. Composite Key: If a table don't have any individual columns that qualifies for a Candidate key, then you have to select 2 or more columns to make a row unique. Like if there is no EmployeeID or SSN columns, then you can make FullName + DateOfBirth as Composite primary Key. But still there can be a narrow chance of duplicate row.

6. Foreign Key : It is a column that creates a relationship between two tables. It acts as a cross-reference between two tables as it references the primary key of another table.

Database schema

Design of database is called as Schema.

Schema is of three types:

- **Physical/Internal Schema**
 - how data stored in block of storage
- **Logical(Conceptual) Schema**
 - What type of data stored in data structures
- **View (External) Schema**
 - End user interaction with database systems

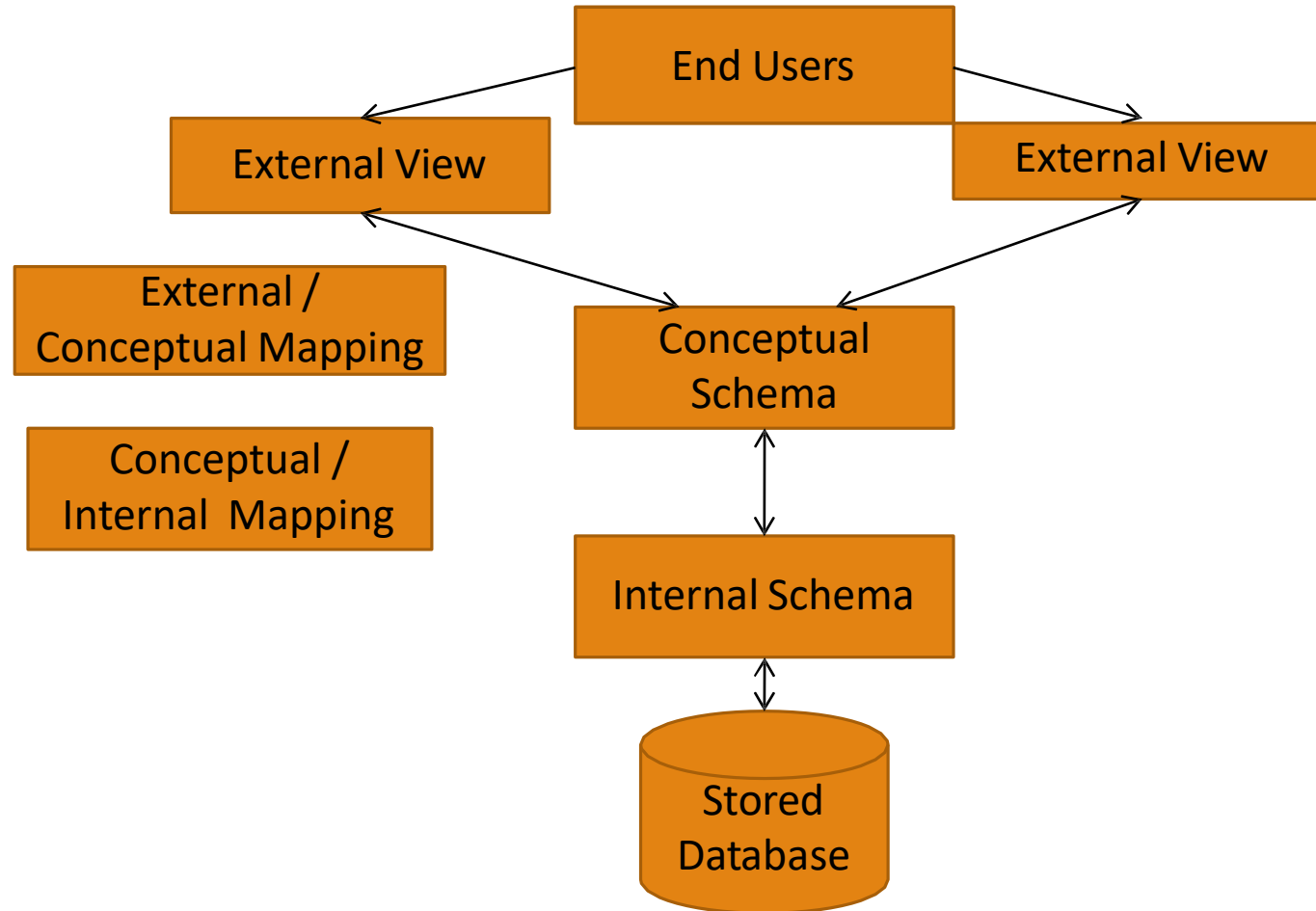
Three –Schema Architecture of database

The goal is to separate user application and Physical database.

It is proposed to support the DBMS Characteristics of:

- Program data independence
- Support of Multiple views of data

Three –Schema Architecture of database



Three –Schema Architecture of database

Defines DBMS Schema at Three levels:

External Schema

- It includes a number of external schemas or user views.
- Describes the part of the database that a particular user group is interested in and hides the rest of the database from that user group.

Conceptual Schema

- Describes the structure and constraints for the whole database

Internal Schema

- Describes physical storage and access paths (on the disk)

The Bank ATM Example

1. Insert Card into machine
2. Provide card details pin etc.
3. Specify amount to be withdrawn.
4. Machine does processing
5. Transaction completed, database updated.

Three –Schema Architecture of database

Mapping

- Mapping among schema levels are needed to transform request and data. Program refers to an external schema, and are mapped by the DBMS to internal schema for execution.

Data Models

Data model is collection of Conceptual tools for describing data, relationships, Constraints etc.

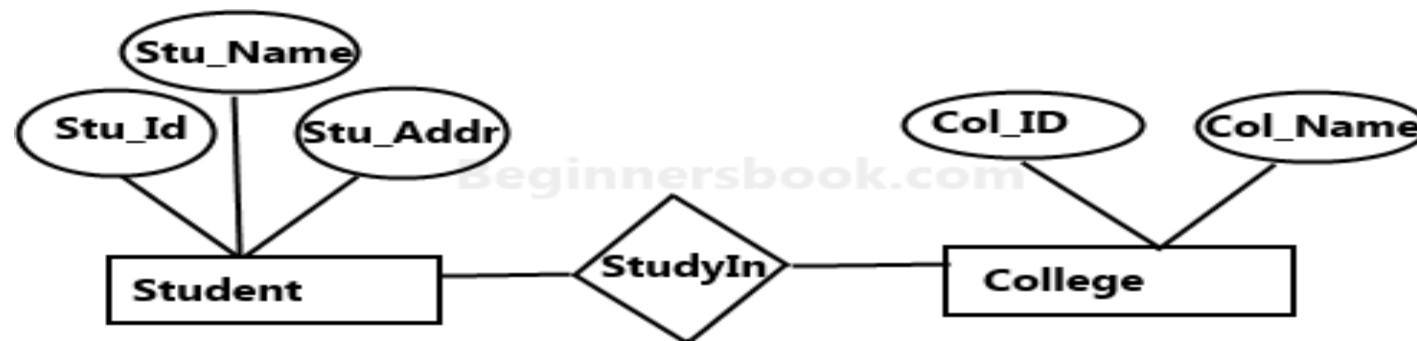
Various data models are categorized in three different groups as:

- Object based data models
- Record based data models
- Physical data models

Data Models

Object based data model

- These are used to describe data at Conceptual or view level.
- Some of the most widely used data models are:
 - 1) Entity relationship model:
 - ER data model is one of the important data model which forms the basis for the all the designs in the database world. It defines the mapping between the entities in the database.
 - It consist of collection of basic object called as Entity and relation-ship among them.



Data Models

2) Object-oriented model:

- It is based on collection of objects and object contains Variables and methods.
- Objects which contains same type of values and Methods those are grouped together into class.

CLASS: Person

PERSON
NAME
ADDRESS
AGE
PHONE
Sp_getAddress ()
Sp_getPhone ()

Objects: John

PERSON
John
Troy
25
2453545
Sp_getAddress (John): Troy
Sp_getPhone (John): 2453545

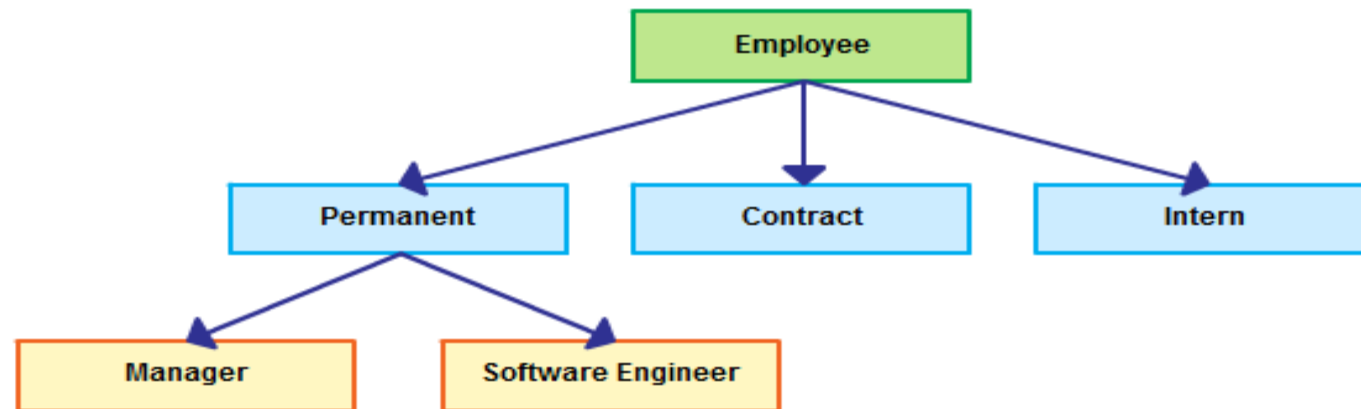
Objects: Mathew

PERSON
Mathew
Fraser Town
28
5645677
Sp_getAddress (Mathew): Fraser Town
Sp_getPhone (Mathew): 5645677

Data Models

Record-Based data models

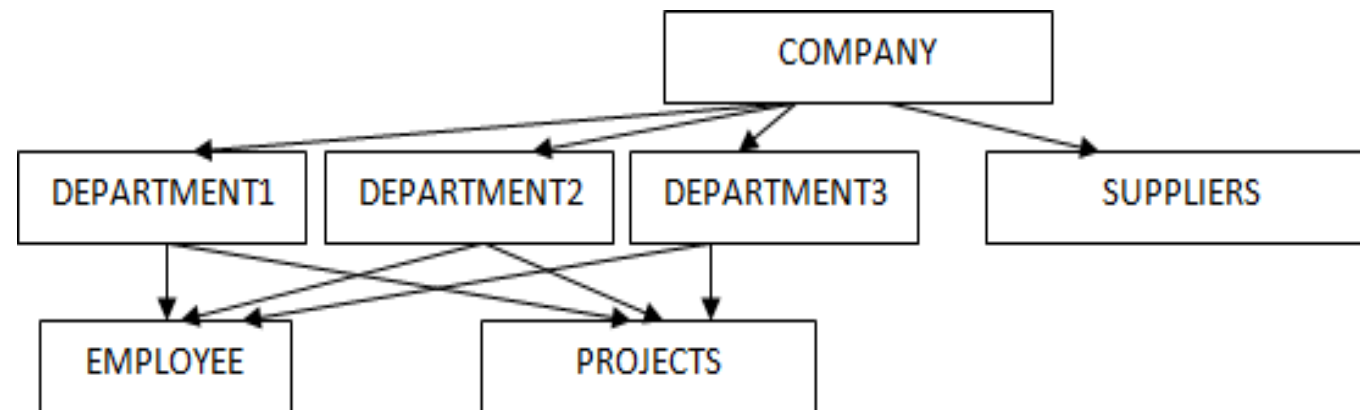
- Describe the data at conceptual or view level.
- There are three record-based models:
 - 1) Hierarchical Data model:
 - In this data and relationships are represented by records and links respectively.
 - This structure organized as collection of tree.
 - In this no node can have ,ore than one parent node.



Data Models

2) Network data model:

- Records in database are organized as collection of nodes and edges.
- Unlike hierarchical model a one node can have more than one parent node.



Data Models

3) Relational data model

- It represents data and relationships among data by collection of tables.
- Each table has number of rows and columns.
- Columns are Attributes
- Rows are Records or tuple.

EMPLOYEE			
EMP_ID	EMP_NAME	ADDRESS	DEPT_ID
100	Joseph	Clinton Town	10
101	Rose	Fraser Town	20
102	Mathew	Lakeside Village	10
103	Stewart	Troy	30
104	William	Holland	30

END

