

DBMS

(Data Base Management System)

Page 1.

Def:

(A DBMS is a software system that allows access to the data contained in a database. The objective of a DBMS is to provide a convenient and effective method of defining, storing and retrieving the information stored in the database). The DBMS interfaces with application program so that data contained in database can be used in multiple application & users.

### Components of a DBMS

A database is composed (made up) of the following four basic components.

- 1) Data
- 2) Hardware
- 3) Software
- 4) Users

208

(1) Data :- A database is a storehouse for data. Data contained in a database should possess the following properties  
Integrated a) consistent + valid } To elaborate  
b) shared.

(2) Hardware :- Physical parts of a computer system, which help in efficient storage, processing & retrieval of information, are termed as the Hardware.

(3) Software - It is group of programs which help to do all the operations related to a database. Examples, are

- a) Dbase (I, II, III, III plus)
- b) Foxbase
- c) Foxpro (for DOS & Windows)
- d) Oracle
- e) Sybase [ Unix ]
- f) Ingress

users :-  
utilizing  
or  
be  
users  
a  
to



(4) Users :- A large variety of users is utilizing a database either concurrently or in a batch environment. These can be categorized as follows:-

a) Naive (Inexperienced) users :- These users are not aware of the presence of a database system. A user of an Automatic teller machine falls in this category.

b) Online users :- These users are aware of the presence of a database system. They have gained a certain level of expertise as far as a DBMS is concerned.

c) Application Programmers :- These are professional programmers who are responsible for developing application programs or user interfaces, which are utilized by Naive & Online users.

d) Data Entry operators :- These users are proficient typists and are responsible for adding large volumes.

product.

of data in the database of organisations (rap)  
(c) Database Administrator (DBA):- DBA is  
a person or a group of persons, who  
are entrusted with the task of  
defining, constructing & maintaining  
a database. They are the users who  
are most familiar with the database  
and are responsible for its  
upkeep & security.

### Architecture for a DBMS

The generalized architecture of a  
database system is called the  
ANSI/SPARC model. A large number  
of commercial systems and research  
database models fit into this framework.

The architecture is divided into  
three levels.

- a) External level
- b) Conceptual "
- c) Internal "



The following diagram illustrates the three level architecture --

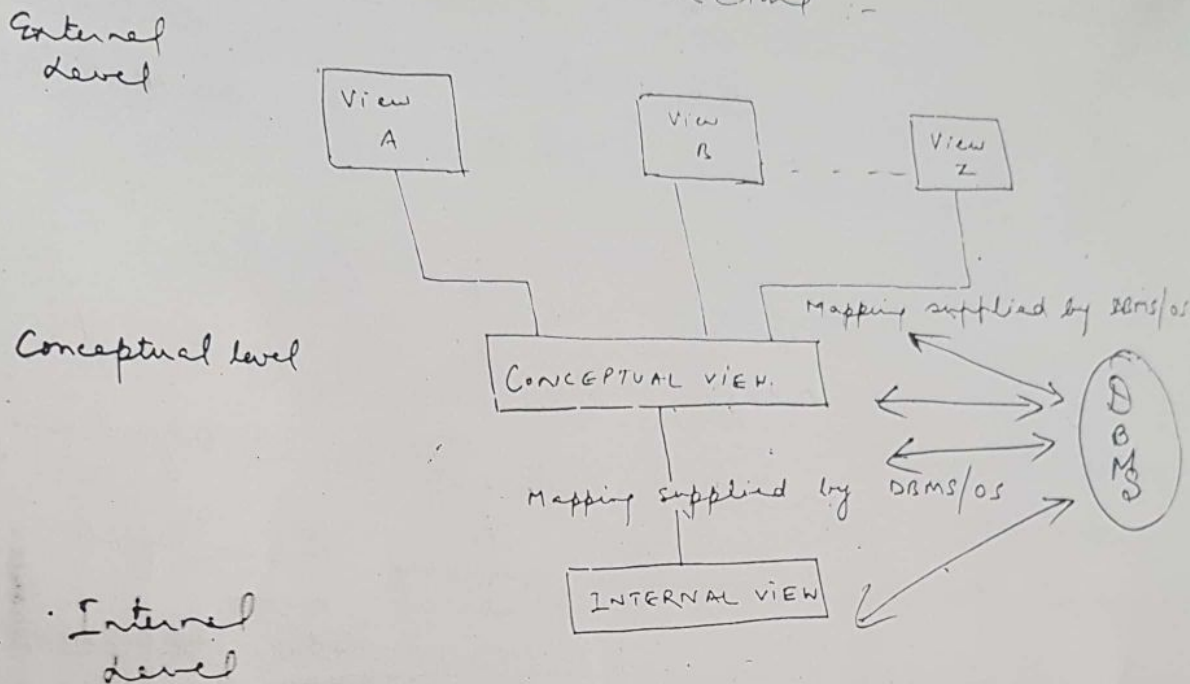


Fig I :- THREE LEVEL ARCHITECTURE OF A DBMS

External level :- The external or user view is at the highest level of database abstraction where only those portion of the database of concern for a user are included. Any number of user views may exist for a given global view. Each external view is described by means of a scheme called an external scheme. [plural for scheme]. Let

consider the definition of logical and the relationship in the external view. It also contains the method of deriving the objects in the external view from the objects in the conceptual view.

Conceptual or Global View:- At this level of database abstraction all the database entities (anything of interest to an organisation) and the relationships among them are included. One conceptual view represents the entire database. There is only one conceptual schema per database. It also contains the method of deriving the objects in the conceptual view from the objects in the internal view.

Internal View:- This view is at the lowest level of abstraction. It is closest to the physical storage method. It indicates how the data will be stored and describes the data structure access methods to be used by



database. It is expressed by the internal schema, which contains the definition of the stored record, the method of representing data fields, and the access aids used.

### Mapping between Views :-

(Two mappings are required in a database system with three different views.)

A mapping between the external and conceptual views gives the correspondence among the records and the relationships of the external & conceptual views. The external view is an abstraction of the conceptual view, which in turn is an abstraction of the internal view.

Mapping between the conceptual & the internal levels specifies ~~the~~ the method of deriving the conceptual record from the physical database.

## Data Independence

Three levels of abstraction, along with mappings from external to conceptual and from conceptual to internal, provide two distinct levels of data independence :-

- a) Logical Data Independence
- b) Physical " "

Logical Data Independence :- It indicates

that the conceptual schema can be changed without affecting the existing external schemas. The change would be absorbed by the mapping between the external and conceptual levels. It also insulates application programs from operations such as combining two records into one, or splitting an existing record into two or more records.



Physical Data Independence :- indicates that the physical storage structures or devices used for storing the data could be changed without necessitating a change in the conceptual view or any of the external views. The change would be absorbed by the mapping between conceptual and internal levels.

### Responsibilities of a DBA.

1. DBA is the custodian of the data and controls the database structure.
2. He is responsible for definition and implementation of the internal level.
3. Different mappings are also supplied by the DBA.
4. Different permissions regarding the use of database are granted by the DBA.
5. DBA is responsible for defining recovery procedures in case of a failure.

(11)

DBA is responsible for maintaining the integrity of a database.

He continuously interacts with the users to see if there is any problem.

DBA monitors the database and responds to any change in requirements.

### Advantages of using a DBMS

(1.) One of the main advantages of using a database is that the organization can exert centralized management and control over the data.

(2.) Reduction of Redundancies:- Centralized control of data by DBA avoids unnecessary duplication of data and effectively reduces the amount of data storage required.

(3.) Shared Data:- A database allows the sharing of data under its



control any any number of applications program on users.

(4) Integrity of data :- Centralized control ensures that adequate checks are there in the DBMS to provide data integrity. Data integrity means that data contained in a database is both accurate and consistent.

(5) Security :- Data is a valuable resource for any organization and can be of confidential nature. Data in a Database

can be safeguarded against unauthorized usage and access. Proper security checks can be used to ensure that only right people get right access to right data. Usage of Magnetic Ridges, username - password schemes, retina scanner, fingerprint scanner can be employed for proper security of a database.

Conflict Resolution:- Since a database is under the direct control of a DBA, the conflicting requirements of various users and applications can be resolved amicably.

Data Independence:- One of the features of a DBMS is Data Independence. Both logical & physical data independence can be achieved in a DBMS.

### Disadvantages of a DBMS

- (1) A significant disadvantage of a DBMS is cost.
- (2) Hardware has to be upgraded which increases the overheads.
- (3) Purchasing and developing the software also requires a large amount of money.
- (4) Efficiency gets reduced as a lot of time is spent in implementing security, integrity and sharing of data.
- (5) Lack of duplication requires that data should be backed up frequently. This is a time consuming & expensive operation.



## DBMS facilities

Two main types of facilities are provided by a DBMS:-

- (i) The data definition facility or DDL
- (ii) The data manipulation facility or DML

Data Definition Language (DDL) - This facility

can be used to define the conceptual scheme and also give some details about how to implement this scheme in the physical devices used to store the data.

This definition includes all the entity sets and their associated attributes as well as relationships among them.

It also includes any constraints that have to be maintained. The compiled

form of these definitions are known as Data Dictionary, directory or

System Catalog. The data dictionary contain information on the data stored in the database and is often consulted by the DBMS.

## Data Manipulation Language (DML)

The language used to manipulate data in the database is called Data Manipulation Language (DML). It involves retrieval of data from the database, insertion of new data into the database, and deletion or modification of existing data. A query is used to retrieve the data. It is a statement in the DML that requests the retrieval of data from the database.

The DML provides commands to select and retrieve data from the database. Commands are also provided to insert, update and delete records.

The DML can be procedural which means that user indicates not only what to retrieve but how to go about retrieving it. If the DML is non procedural, the user has to indicate only what is to be retrieved.



## DATA MODELS

A model is a way to represent data. There are different models for representation of data. Primary among them are

- 1) Relational Model.
- 2) Hierarchical Model.
- 3) Network Model.

To understand these models we consider data about certain suppliers, about the parts they supply and about the quantity of part supplied by them.

### Information about Supplier :-

- (i) Name of the supplier
- (ii) A Supplier Number
- (iii) A status code
- (iv) City in which supplier lives

### Information about parts

- (i) Part Number
- (ii) Part Name
- (iii) Color of part
- (iv) Weight of part
- (v) City in which part is located

Information about shipments

- (i) Part number which is supplied.
- (ii) Supplier number who supplied the part.
- (iii) quantity in which part is supplied.

The above information can be represented using the Relational, Hierarchical & Network models.

### Relational model

Information is represented in the relational model in shape of tables. There are columns in a table which represent the attributes of an entity about which the table is constructed. The rows of a table are referred to as tuples.

So to represent the information about suppliers, in relational model we need to create three tables. We can call them S-Table, P-Table & SP-Table.



# Tables in Relational Model

(Page 17)

## S-Table

S#	SNAME	STATUS	CITY
S1	Smith	10	London
S2	Clark	20	Rome
S3	Blake	30	Paris
S4	Adem	40	Amsterdam

## P-Table

P#	PNAME	WT	COLOR	CITY
P1	Nut	100	Red	London
P2	Bolt	200	Black	London
P3	Screw	300	Black	Paris

## SP-Table

S#	P#	QTY
S1	P1	100
S1	P2	250
S1	P1	250
S2	P1	300
S2	P2	200
S3	P2	250
S1	P3	

To analyse the relational model, we consider the anomalies of Insert, delete and update.

(i) Insert :- Suppose a new supplier comes from Area W. We can insert a new tuple into table S.

(ii) Delete :- Suppose supplier S3 stops supplying part P2. Remove the tuple ~~concerning~~ <sup>concerning</sup>.

Solution :-  
S3 & P2 from S1 table  
Update :- Suppose supplier S1 moves from London to Paris.

Solution :- Necessary update can be performed in table S only at one place, so no consistency is maintained.

Thus we observe that the relational model is free from anomalies of insert, delete and update and thus is very good for represent action of data.



Let us consider the anomalies of insert, delete & update for the hierarchical model.

(i) insert :- Suppose a new supplier comes from Area W

Solution :- Since we don't know the part supplied by this supplier, So information cannot be inserted. This is because part is a superior & supplier is a subordinate.

(ii) delete :- Suppose supplier S3 stops supplying part P2

Solution :- We can remove the supplier S3 information from part P2. But after that we loose any further information about supplier S3. This deletion results in loss of information

(iii) updatation :- Suppose supplier S1 moves from London to Paris

Solution :- This update can be performed very easily but at multiple places. This may result in inconsistency of database.

Page no. \_\_\_\_\_

Let us consider the anomalies of insert, delete & update for the hierarchical model.

(i) insert :- Suppose a new supplier comes from Area W

Solution :- Since we don't know the part supplied by this supplier, so information cannot be inserted. This is because part is a superior & supplier is a subordinate.

(ii) delete :- Suppose supplier S3 stops supplying part P2

Solution :- We can remove the supplier S3 information from part P2. But after that we lose any further information about supplier S3. This deletion results in loss of information

(iii) update :- Suppose supplier S1 moves from London to Paris

Solution :- This update can be performed very easily but at multiple places. It may result in inconsistency of database

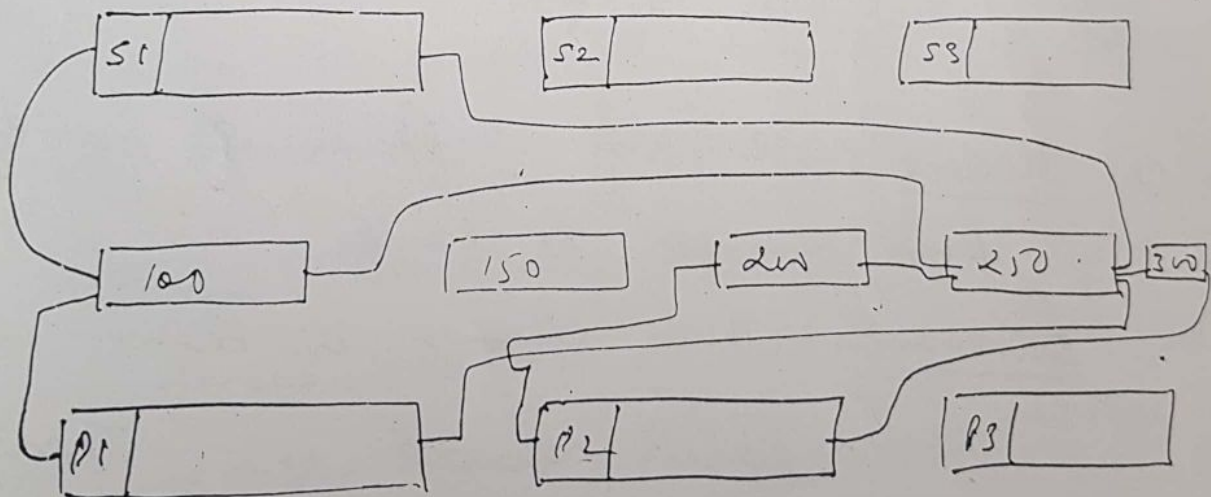


[Page 21]

So we observe that the hierarchical model suffers from anomalies of delete, insert & update and thus it is not suitable for data representation.

### Network Data model

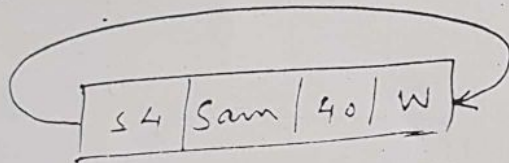
In network data model information is represented by means of pointer or chain. We may construct three records i.e. supplier record, parts record & quantity record. The arrangement is as shown below



Let us consider the three anomalies of insert, delete & update for Network model.

(1) Insert :- Suppose a new supplier comes from area W.

Solution:- A new supplier record can be constructed & when it is known that what part is supplied by him the appropriate chain can be constructed. Initially its chain can be attached to itself



(2) Update :- Suppose supplier S1 moves from London to Paris

Solution:- This change is done in the supplier record S1. Therefore there is no chance of inconsistency as the update is performed only at one place.



131 Deletion :- Suppose supplying part P2. [Page 11]  
Solution :- The chain containing S3 stops  
Record S3, P2 & qty can be removed.

Thus we see that the network Data model is free from the anomalies if Insert / update & delete.

But it has two big drawbacks

1. & 2. it is ~~more~~ expensive to implement point ②. Complexity is increased because of large number of Records.

~~These~~ Thus we observe that out of three models, Relational model is the best for representation of Data.

---