DBMS/SQL

Getting Started with Database

Capgemini

Lesson Objectives

To understand the following topics:

- Introduction to Database
- Characteristics of DBMS
- Data models
- Relational DBMS
- Database Administrator

What is Data?

Data (plural of the word datum) is a factual information used as a basis for reasoning, discussion, or calculation

Data may be numerical data which may be integers or floating point numbers, and non-numerical data such as characters, date etc.

Data by itself normally doesn't have a meaning associated with it.

e.g:- Jack
01-jan-71
15-jun-05
50000

What is Information?

Related data is called as information

Information will always have a meaning and context attached to the data element

When we add meaning and context to the data it becomes information.

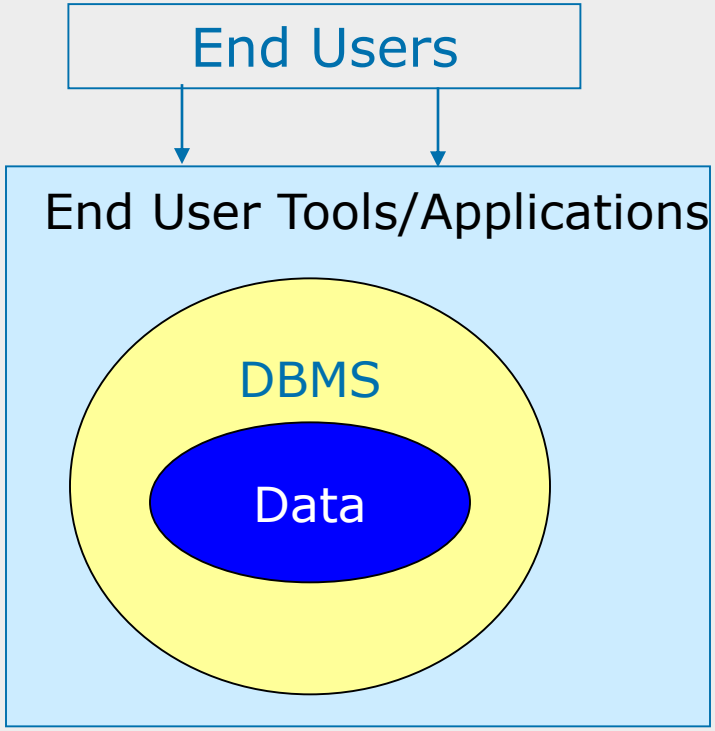
- Employee name: Jack
- Date of birth: 01-jan-71
- Data of joining: 15-jun-05
- Salary: 50000
- Department number: 10

Defining Database, DBMS & Schema

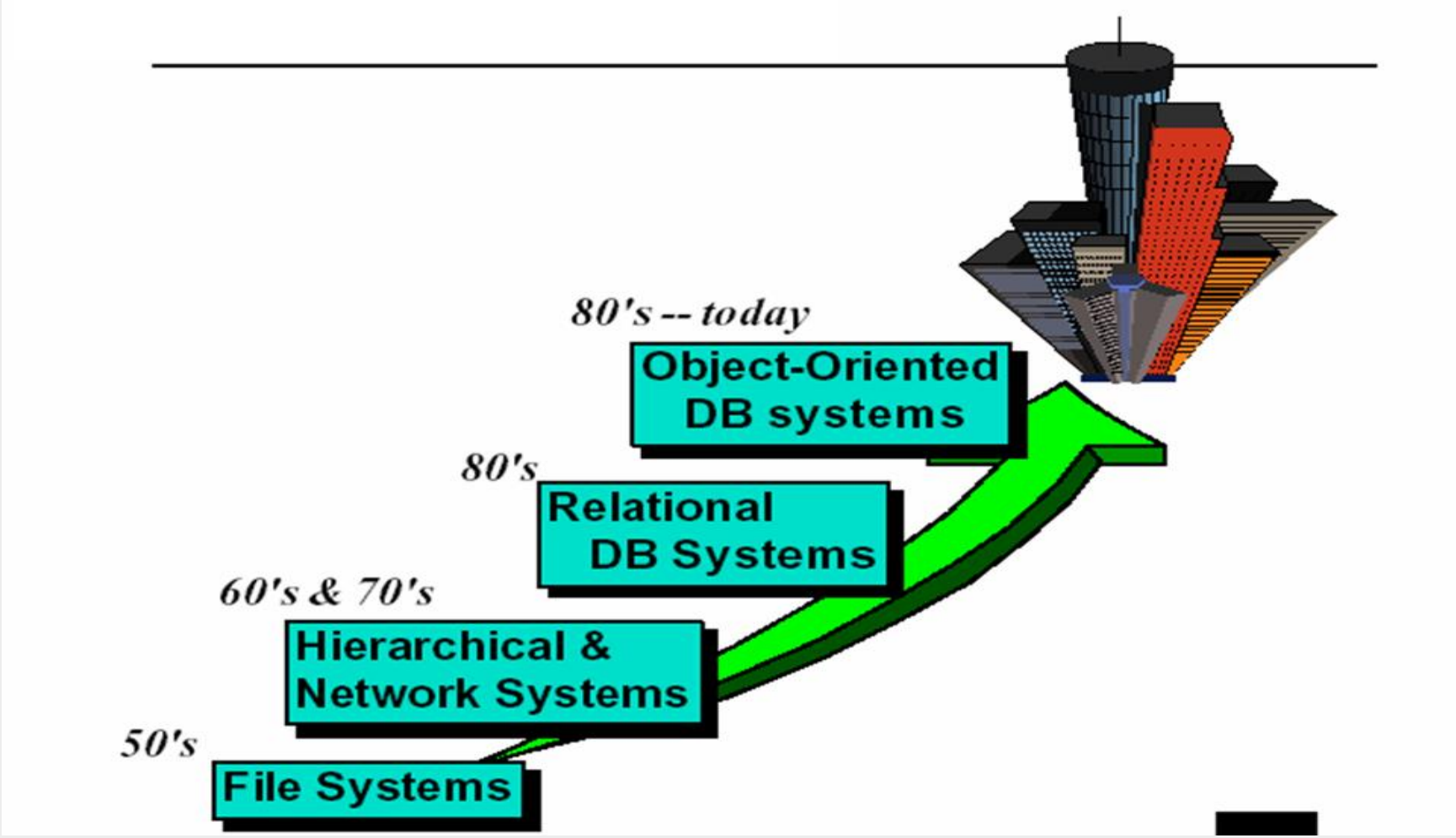
Database: It is a set of inter-related data

DBMS: It is a software that manages the data

Schema: It is a set of structures and relationships, which meet a specific need



Evolution of Databases



Characteristics of DBMS

Given below are the characteristics of DBMS:

- Control of Data Redundancy
  - Traditionally, same data is stored in a number of places
  - Gives rise to data redundancy and its disadvantages
  - DBMS helps in removing data redundancies by providing means of data- integration.
- Sharing of Data
  - DBMS allows many applications to share the data.
- Maintenance of Integrity
  - DBMS maintains the correctness, consistency, and interrelationship of data with respect to the application, which uses the data.

Characteristics of DBMS

- Support for Transaction Control and Recovery
  - DBMS ensures that updates physically take place after a logical Transaction is complete.
- Data Independence
  - In DBMS, the application programs are transparent to the physical organization and access techniques.
- Availability of Productivity Tools
  - Tools like query language, screen and report painter, and other 4GL tools are available.

Characteristics of DBMS

- Control over Security
  - DBMS provides tools with which the DBA can ensure security of the database.
- Hardware Independence
  - Most DBMS are available across hardware platforms and operating systems.

Levels of Abstraction

There are three levels of database abstraction:

- Conceptual Level:
  - The overall integrated structural organization of the database.
- Physical Level:
  - The information about how the database is actually stored in the disk.
- View / External Level:
  - The user view of the database. It is different for different users based on application requirement.

What is a Data Model?

The “Data model” defines the range of data structures supported and the availability of data handling languages.

- It is a collection of conceptual tools to describe:
  - Data
  - Data relationships
  - Constraints
- There are different data models:
  - Hierarchical Model
  - Network Model
  - Relational Model

Why is Data Modeling Important?

Why is Data Modeling important?

- The goal of the "data model" is to ensure that all the data objects required by the database are completely and accurately represented.
- The "data model" uses easily understood notations and natural language. Hence, it can be reviewed and verified as correct by the end-users.

Why is Data Modeling Important?

Why is Data Modeling important?

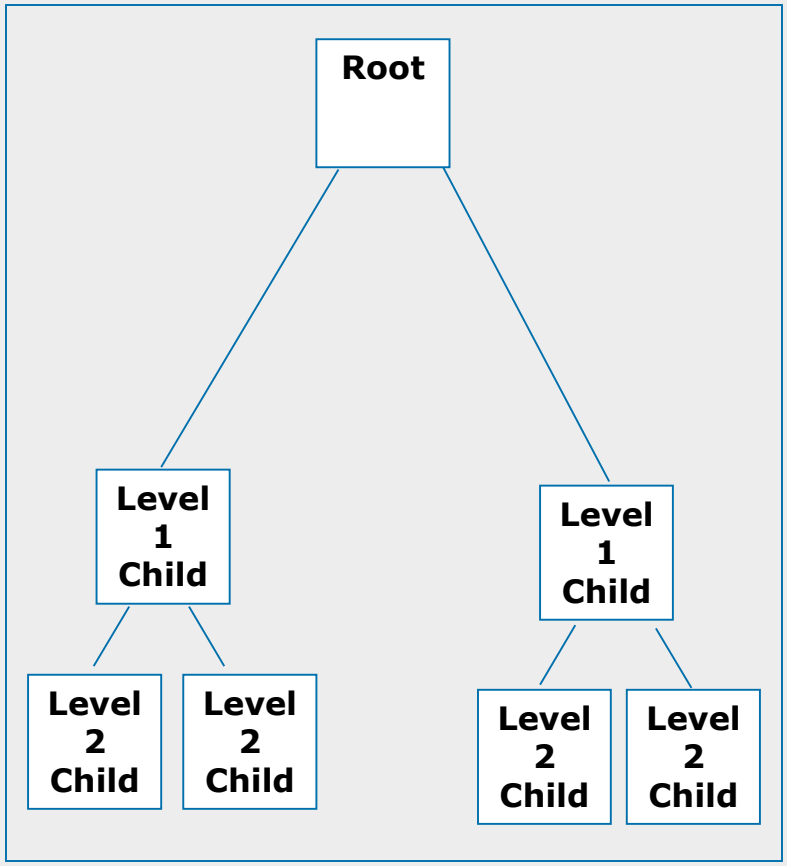
- The goal of the "data model" is to ensure that all the data objects required by the database are completely and accurately represented.

- The "data model" uses easily understood notations and natural language. Hence, it can be reviewed and verified as correct by the end-users.

Hierarchical Model

The Hierarchical model:

- In this model, data is represented by a simple tree-structure.
- Relationships between entities are represented as parent-child.
- Many-to-many relationships are not allowed.
- Parents and children are tied together by links called “pointers”.



Hierarchical Model - Example

Example:

- Consider a student course - marks database.
- In the Hierarchical model a student can register for many courses and gets marks for each course.

- A parent can have many children
- A child cannot have more than one parent
- No child can exist without its parent

Scode	Sname
S1	A

Scode	Sname
S2	B

Ccode	Cname	Marks
C1	Physics	65
C2	Chemistry	78
C3	Maths	83
C4	Biology	85

Ccode	Cname	Marks
C3	Maths	83
C4	Biology	85

Hierarchical Model - Possibilities

Possibilities in a Hierarchical model:

- INSERT
  - Insertion of Dummy student is required to introduce a new course.
- UPDATE
  - To change the course name of one course, the whole database has to be searched. This may result in data inconsistency.
- DELETE
  - Deleting a student - the only one to take the course deletes course information.

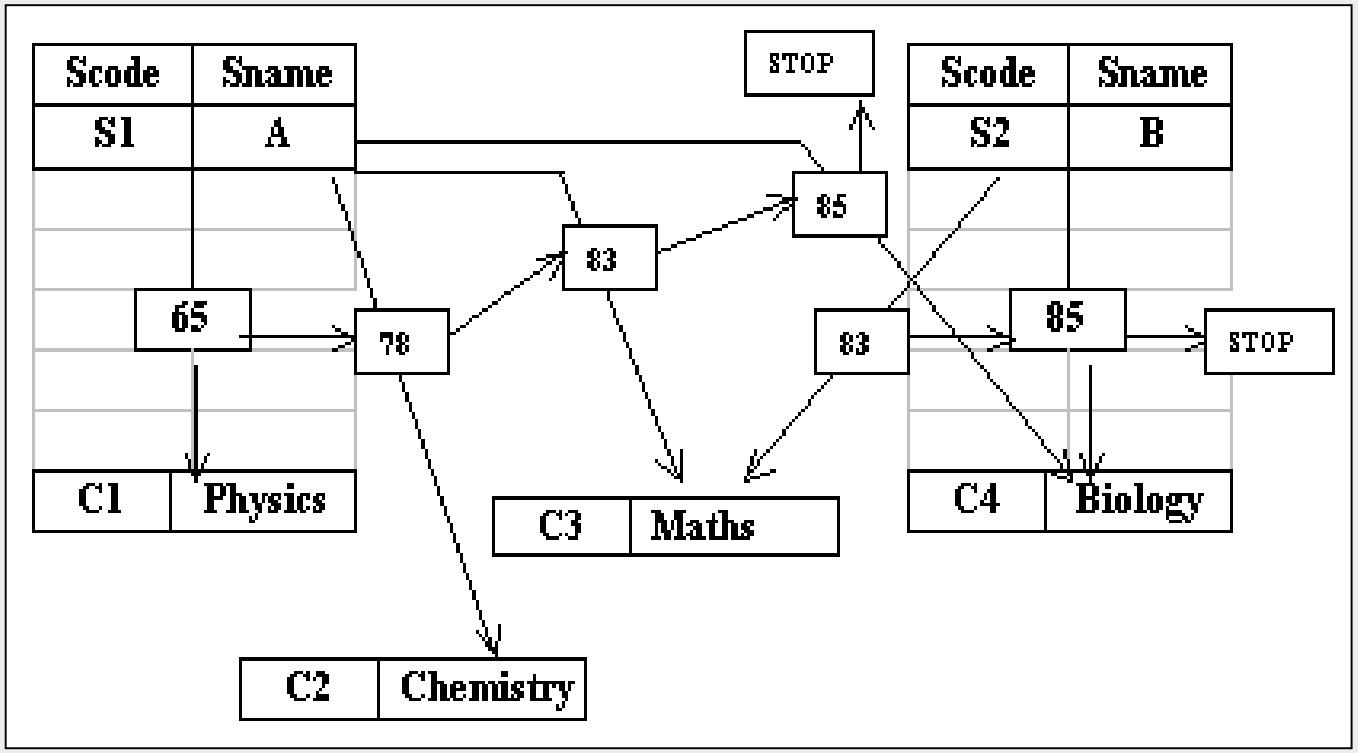
Network Model

The Network model:

- The Network model solves the problem of data redundancy by representing relationships in terms of “sets” rather than “hierarchy”.
- A record occurrence may have any number of immediate superiors.
- The Network model supports many-to-many relationships.
- There is no restriction on number of parents.
- A record type can have a number of parent and child record types.
- It is more complex than the Hierarchical model because of links.
- It is a superset of the Hierarchical model.

Network Model - Example

- In the example of student course – marks, “student record” and “course record” is linked together through “marks record”.



Network Model - Possibilities

Possibilities in a Network model:

- INSERT
  - Inserting a “course record” or “student record” poses no problems. They can exist without any connectors till a student takes the course.
- UPDATE
  - Update can be done only to a particular child record.
- DELETE
  - Deleting any record automatically adjusts the chain.

Relational Model

The Relational model:

- The Relational model developed out of the work done by Dr. E. F. Codd at IBM in the late 1960s. He was looking for ways to solve the problems with the existing models.
- At the core of the Relational model is the concept of a “table” (also called a “relation”), which stores all data.
- Each “table” is made up of:
  - “records” (i.e. horizontal rows that are also known as “tuples”), and
  - “fields” (i.e. vertical columns that are also known as “attributes”)

Relational Model

The Relational model:

- Examples of RDBMS:
  - Oracle
  - Informix
  - Sybase
- Because of lack of linkages, the Relational model is easier to understand and implement.

Student Table	
Scode	Sname
S1	A
S2	B

Course Table	
Ccode	Cname
C1	Physics
C2	Chemistry
C3	Maths
C4	Biology

Marks Table		
Ccode	Scode	Marks
C1	S1	65
C2	S1	78
C3	S1	83
C4	S1	85
C3	S2	83
C4	S2	85

Relational Model - Possibilities

Possibilities in a Relational model:

- INSERT
  - Inserting a “course record” or “student record” poses no problems because tables are separate.
- UPDATE
  - Update can be done only to a particular table.
- DELETE
  - Deleting any record affects only a particular table.

Relational Tables

Examples of Relational tables:

"column" or "attribute"

"row" or "tuple"

Dept table			
Deptno	**Dname**		**Loc**
10	Accounting		New York
20	Research		Dallas
30	Sales		Chicago
40	Operations		Boston

Emp table				
Empno	**Empname**	**Job**	**Mgr**	**Deptno**
7369	Smith	Clerk	7902	20
7499	Allen	Salesman	7839	30
7566	Jones	Manager	7839	20
7839	King	President		10
7902	Ford	Analyst	7566	20

Relational Tables - Properties

Properties of Relational Data Entities:

- Tables must satisfy the following properties to be classified as relational:
  - Entries of attributes should be single-valued.
  - Entries of attributes should be of the same kind.
  - No two rows should be identical.
  - The order of attributes is unimportant.
  - The order of rows is unimportant.
  - Every column can be uniquely identified.

Data Integrity

"Data Integrity" is the assurance that data is consistent, correct, and accessible throughout the database.

Some of the important types of integrities are:

- Entity Integrity:
  - It ensures that no "records" are duplicated, and that no "attributes" that make up the primary key are NULL.
  - It is one of the properties that is necessary to ensure the consistency of the database.

Data Integrity

Foreign Key and Referential Integrity

- The Referential Integrity rule: If a Foreign key in table A refers to the Primary key in table B, then every value of the Foreign key in table A must be null or must be available in table B.

Unique Constraint:

- It is a single field or combination of fields that uniquely defines a tuple or row.
- It ensures that every value in the specified key is unique.
- A table can have any number of unique constraints, with at most one unique constraint defined as a Primary key.
- A unique constraint can contain NULL value.

Data Integrity

Column Constraint:

- It specifies restrictions on the values that can be taken by a column.

DEPT table		
Deptno	Dname	Loc
10	Accounting	New York
20	Research	Dallas

EMP table				
Empno	Empname	Job	Mgr	Deptno
7369	Smith	Clerk	7902	20
7499	Allen	Salesman	7839	30

Database Administrator

A Database Administrator (DBA) is the database architect.

- DBA is responsible for the design and implementation of new databases, and:
  - centrally manages the database.
  - decides on the type of data, internal structures, and their relationships
  - ensures the security of the database
  - controls access to the data through user codes and passwords
  - can restrict the views or operations that the users can perform on the database

Summary

In this lesson, you have learnt:

- What is a Database?
- Characteristics of DBMS
- The Data models, including:
  - the Hierarchical model
  - the Network model
  - the Relational model
- Relational DBMS (RDBMS)
- What is Data Integrity

Review Question

Question 1: A DBA ____.

- Option 1: ensures the security of the application server.
- Option 2: controls access to the data through the user codes and passwords.
- Option 3: manages users.

Question 2: The Physical Level is ____.

- Option 1: the overall structural organization of the d/b.
- Option 2: the information about how the database is actually stored in the disk.
- Option 3: the user view of the database.

Review Question

Question 3: There are different data models such as ____.

Question 4: In Network model, each table is made up "tuples" and "fields".

- True / False

Question 5: A table can have any number of "Unique constraints".

- True / False

Review Question: Match the Following

1. Hierarchical model	a) Inconsistencies may creep in.
2. Network model	b) Many-to-many relationships are not allowed.
3. Data redundancy	c) It is a superset of the Hierarchical model.
4. In DBMS	d) Application programs are transparent to the physical organization and access techniques.