International Journal of Computers & Technology
Volume 4 No. 1, Jan-Feb, 2013 ISSN 2277-3061

REVIEW OF DATABASE AND PROMINENT PROGRAMMES

Vinod Kumar
Assistant Professor
Department of Computer ScienceDev Samaj College for Women
Ferozepur City

Gagandeep Raheja
Assistant Professor
Department of Computer Science
Dev Samaj College for Women
Ferozepur City

Sukhpreet Singh
Assistant Professor
Dept of Electrical Engg
Shaheed Bhagat Singh State Technical Campus
Ferozepur Cantt

ABSTRACT

Databases and database systems have become an essential component of everyday life in modern society. In the course of a day, most of us encounter several activities that involve some interaction with a database. For example, if we go to the bank to deposit or withdraw funds, if we make a hotel or airline reservation, if we access a computerized library catalog to search for a bibliographic item, or if we buy some item such as a book, toy, or computer- form on internet vendor through it Web page, there might be chances that our activities will involve someone or some computer program accessing a database, even purchasing items from a supermarket nowadays in many cases involves an automatic update of the database that keeps the inventory of supermarket items. So in this paper we will talk about how to manage the different type of data involved in any form in the database.

General Terms

Database, Database Management, Security, Integrity, Users

Keywords

DBA, DBMS, DATA, DDL, DML, DCL, TCL.

1.1 DATABASE- AN INTRODUCTION

A database can be defined as a collection of related data from which users can efficiently retrieve the desired information. A database can be anything from a simple collection of roll numbers, names, addresses, and phone numbers of a students to a complex collection of sound, images, and even video or film clippings. Though databases are generally computerized, instances of non-computerized databases from everyday life can be cited in abundance. A dictionary, a phone book, a collection of recipes, and a TV guide are all common examples of non-computerized databases. The examples of computerized databases include customer files, employee rosters, books catalogue, equipment inventories, and sales transactions. A single repository of data is maintained that is defined once and then accessed by many users. The fundamental characteristic of database approach is that the database system not only contains data's but it contains complete definition or description of the database structure and constraints.

1.2 DEFINITION

1. "A database is a collection of interrelated data stored in a database server; these data's will be stored in the form of tables."
2. "A database is a structured collection of records that is stored in a computer system, related together, which can be accessed by different users but the data stores only once."

1.3 DATABASE MANAGEMENT

The database is a shared collection of logically related data, designed to meet the requirement of an organization. The database is large repository of data, which can be accessed by different users at the same time. The DBMS must ensure the security of the database from unauthorized access and recovery of the data during system failures. It must also provide techniques for data sharing among several users. The database and the DBMS software are collectively known as database system. A database management system (DBMS) is computer software that manages databases. The main objective of a DBMS is to provide a convenient and effective method of defining, storing, retrieving, and manipulating the data contained in the database. A DBMS is a set of software programs that controls the organization, storage, management, and retrieval of data in a database.

1.4 DEFINITION

"A database is a computer based record keeping system whose purpose is to record and maintains the information."

1.5 APPROACHES TO DATABASE MANAGEMENT

There are different methods to manage the data contained in the database. We can use computerize or non-computerize method to manage the data.

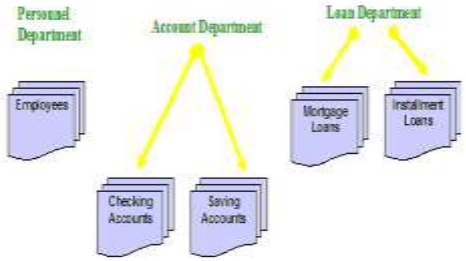
1.5.1 MANUAL APPROACH

Manual database is the record keeping system which is used to keep the records in which human beings manages the database without the support of computers.

1.5.2 FILE SYSTEM APPROACH

In the early systems data was handled manually by the different users. With the large amount of advancement in the computer field, there was a need to store the data in the database in computerized form. File processing system was an early attempt to computerize the manual system. This system stores permanent records in the various files and programs are to be written to add the records in specific files. File processing system is a collection of application programs that perform services for the end users such as the production of Reports. "File processing system was an early attempt to make the manual file system to computerized system." In the

file processing system approach Decentralized approach was used to store and control the data with more ease. When we want to find some data, we have to search through the system from starting to end. Consider a traditional banking system which using the file-based system in managing the organization's data in the figure given below.



1.5.3 DATABASE APPROACH

A database is a collection of information that is organized so that it can easily be accessed, managed, and updated. With the increasing demand of data it the need of the day to store the data in an organized way so that data can be used again and again. The primary aim of database is to provide a way to store and retrieve database information in fast and efficient manner. "A database is a collection of interrelated data's stored in a database server; these data's will be stored in the form of tables." A single repository of data is maintained that is defined once and then accessed by many users. The fundamental characteristic of database approach is that the database system not only contains data's but it contains complete definition or description of the database structure and constraints.

1.6 IMPORTANCE OF DATABASE APPROACH

Database plays an important role in the real world as we were used database in the manual system of keeping the records where large number of files was used to store the data in the database. To solve the problems found in following examples we need database.

- Consider the example of majority of people whose phone number changed from five digits to six digits then it would be difficult to change all the phone numbers from the telephone directory as we need to change all the phone numbers. To overcome this situation we can create a new directory. With this option the job will become more complex.
- Consider an example of a telephone directory: suppose we want to add a new phone number of a person whose last name start with 'J' and page allotted to that name is finished. Then to add the name we have options: 1) start with the new page at the end of directory. 2) Transfer all the names to a new directory having more pages than before. With the option 1) searching the records will become a difficult job and with option 2) the job will become time consuming.
- Consider a case of account number, suppose a customer who has forgot his account number of a particular bank. So he wants to enquire about his account according to his name. Searching would be easy if there are less number of accounts in that bank, but usually banks have more than 10000 different accounts. In that case searching would be difficult for a particular account and moreover it will be a time consuming process.

It is quite clear from the examples that traditional approach works at great accuracy if there is less number of data to be processed. The manual approach does not work well if we have to process larger amount of data. To overcome the fatal flaws in the manual system the database approach is designed.

1.7 ADVANTAGES OF DATABASE MANAGEMENT

The DBMS has a number of advantages as compared to traditional computer file-based processing approach. The DBA must keep in mind these benefits or capabilities during designing databases, coordinating and monitoring the DBMS.

The main advantages of DBMS are described below.

1. **Controlling Redundancy:** By Redundancy means storing same type of data on the different files. Centralized control of data by the DBA avoids unnecessary duplication of data and reduces the storage space. In traditional file processing, every user group maintains its own files. Each group independently keeps files on their database e.g., students. Therefore, much of the data is stored twice or more.

   Redundancy leads to several problems:

   - Duplication of effort
   - Storage space wasted when the same data is stored repeatedly
   - Files that represent the same data may become inconsistent (since the updates are applied independently by each users group).

2. **Minimized data inconsistency:** Data inconsistency exists when different versions of the same data appear in different places. By controlling the data redundancy, the data consistency is obtained. If a data item appears only once, any update to its value has to be performed only once and the updated value (new value of item) is immediately available to all users. For **example**, data inconsistency exists when a college's account branch stores a student name as **Vinod Kumar** but hostel warden stores it as **Vinod Kamboj**. If the dbms has controlled redundancy, the database system enforces consistency. It means that when data item that appears more than once, the dbms automatically updates each occurrence of a data item in the database.

3. **Restricting Unauthorized Access:** A dbms should provide a security and authorization subsystem. DBMS provided security as any database system requires you to login and then process the data depending on the rights given by the DBA to the user who has logged in. The advantage of such a system is securing the data and providing the user and the DBA the secured platform. Any user who logs in cannot do whatever he wants but his role can be defined very easily.

4. **Improved data access:** The dbms makes it possible to produce quick answers to queries. From the

database perspective, a query is a specific request for data manipulation (for example: to read or update the data) issued to the dbms. Simply put, a query is a question. The dbms sends back an answer to the applications.

5. **Data administration:** When several users share the data, centralizing the administration of data can offer significant improvements. Experienced professionals, who understand the nature of the data being managed, and how different groups of users use it, can be responsible for organizing the data representation to minimize redundancy and for retuning the storage of the data to make retrieval efficient.

6. **Enforcing Integrity Constraints:** If data is always accessed through the dbms, the dbms can enforce integrity constraints on the data. For example, before inserting mobile number of an employee, the dbms can check that the digit for mobile number i.e. it should be of ten digits. Also the dbms can enforce access controls that govern what data is visible to different classes of users.

7. **Providing Backup and Recovery:** DBMS Must provides facilities for recovering from hardware or software failures, such as disk crash, power failure, software errors. Saving the data to some other machine periodically is known as backup of data.

8. **Concurrent access and crash recovery:** A dbms schedules concurrent accesses to the data in such a manner that users can think of the data as being accessed by only one user at a time. Further, the dbms protects users from the effects of system failures. The database must include concurrency control software to ensure that several users trying to update the same data at the same time, it should maintain in a controlled manner.

1.8 DATABASE ADMINISTRATOR

In any organization where many persons use the same resources, there is a need for a chief administrator to oversee and manage these resources. In a database environment, the primary resource is the database itself, and the secondary resource is the DBMS and related software. Administering these resources is the responsibility of the database administrator (DBA). A database administrator (DBA) is a person who is responsible for the management of the database. DBA is a person or group of persons who implements the policies of an organization. He is responsible for authorizing access, monitoring database use, providing satisfactory response time, backup and recovery from the system failure. DBA has all the powers of the database.

The main Duties of the DBA are as under:

- **Installation of new software**: It is primarily the job of the DBA to install new versions of DBMS software, application software, and other software related to DBMS administration. It is important that the DBA or other staff members test this new software before it is moved into a production environment.

- **Data Dictionary Management**: Management and control of Data Dictionary is an important role of the DBA. He writes the storage structure definition which is translated by the Data Definition Language compiler. DBA creates the data definition, data validation rules; define user access rules, documentation of data dictionary and etc.

- **Scheme definition:** It is the duty of the DBA to create the original database scheme. This involves writing a set of definitions in a DDL (data storage and definition language), compiled by the DDL compiler into a set of tables stored in the data dictionary.

- **Storage structure and access method definition**: DBA will be responsible for defining the storage structure for the data in the database. He is also responsible for defining the access rights for various users according to the access given to them.

- **Define Integrity rules**: DBA is responsible for defining integrity rules for the database. These help to maintain database security. For example the mobile number of employees should be of ten digits only; such constraint must be imposed by DBA.

- **Configuration of hardware and software with the system administrator**: In many cases the system software can only be accessed by the system administrator. In this case, the DBA must work closely with the system administrator to perform software installations, and to configure hardware and software so that it functions optimally with the DBMS.

- **Security administration**: One of the main duties of the DBA is to monitor and administer DBMS security. This involves adding and removing users, administering quotas, auditing, and checking for security problems. DBA have all the rights to grant or revoke any privilege given to any particular user. He also grants such privileges according to the user's requirements.

- **Data analysis**: The DBA will frequently be called on to analyze the data stored in the database and to make recommendations relating to performance and efficiency of that data storage. This might relate to the more effective use of indexes, enabling "Parallel Query" execution, or other DBMS specific features.

1.9 DATABASE SYSTEM ENVIRONMENT

A database system is a computer based record keeping System whose purpose is to record and maintain the information according to the organization necessary for making decisions. The term database system refers to an organization of components that define and regulate the collection, storage, management and use of data within a database environment. With the continuous development and growth of the database, these systems are used in various applications that are used in real world for example

- Banking System and ATM's machines.
- Flight Reservation Systems.
- Computerized Library Systems.
- Super Market Product Inventory System.

1.10 COMPONENTS OF THE DBMS

The primary aim of database is to provide a way to store and retrieve database information in fast and efficient manner. In database approach, a single repository of data is maintained that is defined once and then accessed by many users. A database system is composed of four components which coordinate with each other to form an effective database system.

- Data
- Hardware
- Software
- Users

1.11 DATABASE LANGUAGES

A DBMS is a software package that carries out different task such as giving the facilities to user to access and modify the information in the database.

There are four types of database languages:

1) Data Definition Language(DDL)
2) Data Manipulation Language(DML)
3) Data Control Language(DCL)
4) Transaction Control Language(TCL)

1.11.1 Data Definition Language (DDL)

A Data Definition Language (DDL) is a computer language for defining the type of data structures used in the database. DDL statements are used to create, modify, and remove database objects such as tables, indexes, and users.

The following are the various DDL commands:

- **CREATE -** To create objects in the database
- **ALTER -** Alters the structure of the database
- **DROP -** Delete objects from the database
- **TRUNCATE -** Remove all records from a table, including all spaces allocated for the records are removed
- **RENAME -** Rename an object

A Data Definition Language has a pre-defined syntax for describing data. For example, to build a new table using SQL syntax, the CREATE command is used, followed by parameters for the table name and column definitions. The DDL can also define the name of each column and the associated data type. Once a table is created, it can be modified using the ALTER command. If the table is no longer needed, the DROP command will delete the table.

Example:

```
SQL>Create table Demo
    (Name varchar2 (20),
    Roll    number (60,
    Class   varchar2 (5)
    );
```

1.11.2 Data Manipulation Language (DML)

Data Manipulation Language (DML) is a family of computer languages used by computer programs database users to retrieve, insert, delete and update data in a database. Currently the most popular data manipulation language is that of SQL, which is used to retrieve and manipulate data in a Relational database. The goal is to provide efficient human interaction with the system. Once the structure of the table is created, it is necessary to insert the data into table and modify the data if necessary. Data Manipulation Languages perform the following operations:

- **SELECT** - Retrieve data from the a database
- **INSERT** - Insert data into a table
- **UPDATE** - Updates existing data within a table
- **DELETE** - Deletes all records from a table, the space for the records remain

There are two types of DML:

- **Procedural:** The user specifies what data is needed and how to get it.
- **Non-Procedural:** The user only specifies what data is needed.

1.11.3 Data Control Language Statements (DCL)

A Data Control Language (DCL) is a computer language and a subset of SQL, used to control access to data in a database. That is a user can access any data based on the privileges given to him. DCL statements are used to provide a kind of security to the database. Users can be granted privileges or roles to allow restricted access to the database. Grant Command is used to give the permission to the user.

Examples of DCL commands include:

- **GRANT:** To allow specified users to perform specified tasks.
- **REVOKE:** To cancel previously granted or denied permissions.

1.11.4 Transaction Control Language Statements (TCL)

Transaction Control (TCL) statements are used to manage the changes made by DML statements. It allows statements to be grouped together into logical transactions. For revoking the transactions and to make the data commit to the database we use TCL.

- **COMMIT** - Save work done
- **SAVEPOINT** - Identify a point in a transaction to which you can later roll back
- **ROLLBACK** - Restore database to original since the last COMMIT

1.12 DATA MODELS

The basic purpose of the data model is to make sure that all the data objects required by the database are understood fully. This is possible because data models used simple notations and natural language to represent information.

International Journal of Computers & Technology
Volume 4 No. 1, Jan-Feb, 2013 ISSN 2277-3061

Data model selected for database should be detailed enough to represent all information. The information contained in the data model is used to define all the integrity constraints and like all other information. The main properties of data models are as under:

- It should be able to represent all information diagrammatically.
- It should be simple and expressible to design the data in database.
- There should be no duplicacy in the data.
- It should be independent from application.

1.12.1 OBJECT BASED DATA MODELS

Object based data models are known as conceptual model used to represent the concepts like entities, attributes and their relationship. The various object based data models are:

1.12.1.1 The E-R Model

The entity-relationship data model is based on a collection of basic objects, called entities, and of relationships among these objects. An entity is a "thing" or "object" in the real world that is distinguishable from other objects. For example, each person is an entity, and bank accounts can be considered entities. Entities are described in a database by a set of attributes. For example, the attributes account-number and balance describe one particular account in a bank. A relationship is an association among several entities.

1.12.1.2 The Object-Oriented Model

The object-oriented model is based on a collection of objects. An object contains values stored in instance variables within the object. An object also contains bodies of code that operate on the object. These bodies of code are called methods. The only way in which one object can access the data of another object is by invoking a method of that other object. This action is called sending a message to the object.

1.12.1.3 Semantic data model

The logical data structure of a database management system (DBMS), whether hierarchical, network, or relational, cannot totally satisfy the requirements for a conceptual definition of data because it is limited in scope and biased toward the implementation strategy employed by the DBMS. Therefore, the need to define data from a conceptual view has led to the development of semantic data modeling techniques. That is, techniques to define the meaning of data within the context of its interrelationships with other data. A semantic data model is an abstraction which defines how the stored symbols relate to the real world. Thus, the model must be a true representation of the real world.

1.12.2 Physical Data Models

Physical data models are used to describe data at the lowest level. It describes the detail of data how the data is stored in the computer system with the defined structure and access paths.

1.12.3 Record-based Logical Models

In record-based models, the database is structured in fixed-format records of several types. Each record has a fixed set of fields. The three most widely accepted record-based data models are the relational, network, and hierarchical models. The latter two were widely used once, but are of declining importance. The relational model is very widely used. Databases based on the relational model are called relational databases.

1.12.3.1 The Hierarchical Data Model

A hierarchical data model is a data model in which the data is organized into a tree-like structure**.** The structure allows repeating information using parent/child relationships: each parent can have many children but each child only has one parent. All attributes of a specific record are listed under an entity type.

In a database, an entity type is the equivalent of a table; each individual record is represented as a row and an attribute as a column. Entity types are related to each other using 1: N mapping, also known as one-to-many relationships. A data model based on one–many relationships between aggregations of fixed numbers of data items, such an aggregation being termed a segment.

1.12.3.2 Network Model

The popularity of the network data model came into existence with the popularity of the hierarchical data model. Some data were more naturally modeled with more than one parent per child. So, the network model permitted the modeling of many-to-many relationships in data. In 1971, the Conference on Data Systems Languages (CODASYL) formally defined the network model. The Network model was standardized by the CODOSYL DBTG (Conference of Data System Language, Database Task Group) model. In this model a child record can have more than one parent. In this model directed graphs are used instead of tree structure. The relationships between specific record types is one to one (1:1), one to many (1: N), many to many (M: N). the relationship can be explicitly defined in the database definition of this model.

1.12.3.3 Relational Model

The Relational Model was the first theoretically founded and well thought out Data Model, proposed by E.F. Codd in 1970. It has been the foundation of most database software and theoretical database research ever since. The Relational Model is a clean and simple model that uses the concept of a relation using a table rather than a graph or shapes. The information is stored in the form of table's .i.e. in rows and columns.

The Relational database possess following properties:

- Data is represented in the form of tables only.
- The only concern of relational database is with data only not with physical structure.
- It provides the information regarding metadata.
- At the intersection of row and column there will be only one value for the tuple.
- It provides a way to handle the queries with ease.

1.13 DATABASE MANAGEMENT--- BASIC TERMINOLOGIES

1. RELATION: A relation is usually described as a table, which is organized into rows and columns. All the data and relationship is defined in the form table. An RDBMS used the table form to represent information. A table or relation consists of number of records.

2. ATTRIBUTE: An attribute is a named given to a column in a table. In the relational model relations are used to

hold the values corresponding to the attributes represented in a relation.

3. TUPLE: The rows of a relation that contains the values corresponding to the attributes are called as Tuple.

4. DOMAIN: The possible set of values that an attribute can accept is known as Domain. For each column of a table there is a set of possible values called as Domain. The domain contains set of values appear under a particular attribute.

5. DEGREE: The number of Attributes/ Columns in a relation is called as the Degree of that relation. In above table there are five attributes so we can say that the degree for the relation is 5. The relation with one attribute is called as **Unary** relation. The relation with two attribute is called as **binary** relation and with three attributes it is called as **ternary** relation.

6. CARDINALITY: The number of rows/ tuple at any one time in a relation is called as the Cardinality of that relation. Because the number of rows increased or decreased with inserts or delete operation respectively so that is why we are using the word any one time

1.14 SECURITY OF DATA

Database security is the system, processes, and procedures that protect a database from unintended activity. Unintended activity can be categorized as authenticated misuse, malicious attacks or inadvertent mistakes made by authorized individuals or processes. Traditionally databases have been protected from external connections by firewalls or routers on the network perimeter with the database environment existing on the internal network opposed to being located within a demilitarized zone.

Security is one of the important functions of database management system. It protects the data against accidental or intentional tampering**.** Database security is the protection of database against any accidental or destruction or misuse of data. As we know that the data of particular organization is very important, so it's the need of the day to protected the data from unauthorized access or misuse by the intruders.

1.14.1 SECURITY AND INTEGRITY THREATS

By threat we mean any situation that will adversely effect the database security of any organization. Security can be removed under intentional or accidental. There are basically two dimensions for the protection of data in the database**. First,** a certain class of data is available only to those people who are authorized to access it. This ensures that the confidentiality of the data is maintained. For example: the medical records of patients in a hospital are accessible only to health care officer. **Second**, the data must be protected from accidental or intentional destruction.

The security threats are broadly divided into two types:

1) Accidental security threats
2) Malicious or intentional Security threats

1.15 PRIVILEGES

Privileges define the access rights provided to a user on a database object**.** A privilege is a right to execute an SQL statement or to access another user's object. The privileges are granted by DBA. A user who has been granted some form of authorization may be allowed to pass this authorization to other users. However this authorization to different users must be given according to the needs of all the users. So the privileges are granted to users in a limit in order to maintain the security.

There are two types of privileges:

- System Privilege
- Object Privilege

1.16 CHECK CONSTRAINTS

Check constraints enforce domain integrity by limiting the values that are accepted by a column. They are similar to Foreign Key constraints in that they control the values that are put in a column. The difference is in how they determine which values are valid: Foreign Key constraints obtain the list of valid values from another table, and Check constraints determine the valid values from a logical expression that is not based on data in another column. For example, the range of values for a salary column can be limited by creating a Check constraint that allows for only data that ranges from $15,000 through $100,000. This prevents salaries from being entered beyond the regular salary range. You can create a Check constraint with any logical (Boolean) expression that returns True or False based on the logical operators. For the previous example, the logical expression is: salary >= 15000 AND salary <= 100000.

You can apply multiple Check constraints to a single column. You can also apply a single Check constraint to multiple columns by creating it at the table level. For example, a multiple-column Check constraint can be used to confirm that any row with a country/region column value of USA also has a two-character value in the state column. This allows for multiple conditions to be checked in one location.

1.17 CONCLUSION

The study of databases is a battleground of ideas. The database community is one of the oldest in the computer world, and it is almost as famous as the application programming community for the diversity of its ideas and the sharpness of the debates between its gurus. Lately events have conspired to expose these concerns to a wider audience. For instance, the seemingly inexhaustible march of the web revolution has exposed more and more developers to database issues because of the desire for ever more dynamic web sites. And the crown prince of web technologies, XML, has had the effect of increasing awareness of data design in general. This means that more and more developers find themselves choosing between database management systems (DBMSes). This can be a daunting choice considering the many available DBMSes, both open and closed source, and the broad spectrum of differences between them. This article provides some guidance through the maze of available DBMS features and methodologies, to help the developer quickly narrow the choices to the best candidate.

1.18 REFERENCES

en.wikipedia.org/wiki/Database_management_system

pages.cs.wisc.edu/~dbbook/openAccess/thirdEdition/.../slides3ed.htm...

ABS PUBLICATION'S "INTRODUCTION TO DBMS"