CHAPTER 1: INTRODUCTION TO DATABASE MANAGEMENT SYSTEMS

Introduction to Database management system (DBMS)

Database management system (DBMS) are computer software applications that interact with the user, other applications, and the database itself to capture and analyze data. A general-purpose DBMS is designed to allow the definition, creation, querying, update, and administration of databases.

It's a set of software programs that controls the organization, storage and retrieval of data (fields, records and files) in a database. It also controls the security and integrity of the database.

The following are main examples of database applications:

• Computerized library systems

• Automated teller machines

• Flight reservation systems

• Computerized parts inventory systems

Advantages of DBMS

The database management system 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 databases and monitoring the DBMS.
The Main advantages of DBMS are described below.

- Controlling Data Redundancy - In non-database systems each application program has its own private files. In this case, the duplicated copies of the same data is created in many places. In DBMS, all data of an organization is integrated into a single database file. The data is recorded in only one place in the database and it is not duplicated.

- Sharing of Data - In DBMS, data can be shared by authorized users of the organization. The database administrator manages the data and gives rights to users to access the data. Many users can be authorized to access the same piece of information simultaneously. The remote users can also share same data. Similarly, the data of same database can be shared between different application programs.

- Data Consistency - 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 is immediately available to all users. If the DBMS has controlled redundancy, the database system enforces consistency.

- Integration of Data - In Database management system, data in database is stored in tables. A single database contains multiple tables and relationships can be created between tables (or associated data entities). This makes easy to retrieve and update data.

- Integration Constraints - Integrity constraints or consistency rules can be applied to database so that the correct data can be entered into database. The constraints may be applied to data item within a single record or the may be applied to relationships between records.

- Data Security - Form is very important object of DBMS. You can create forms very easily and quickly in DBMS. Once a form is created, it can be used many times and it can be modified very easily. The created forms are also saved along with database and behave like a software component. A form provides very easy way (user-friendly) to enter data into database, edit data and display data from database. The non-technical users can also perform various operations on database through forms without going into technical details of a fatabase.

- Report Writers - Most of the DBMSs provide the report writer tools used to create reports. The users can create very easily and quickly. Once a report is created, it can be used may times and it can be modified very easily. The created reports are also saved along with database and behave like a software component.

- Control Over Concurrency - In a computer file-based system, if two users are allowed to access data simultaneously, it is possible that they will interfere with each other. For example, if both users attempt to perform update operation on the same record, then one may overwrite the values recorded by the other. Most database management systems have sub-systems to control the concurrency so that transactions are always recorded with accuracy.

- Backup and Recovery Procedures - In a computer file-based system, the user creates the backup of data regularly to protect the valuable data from damage due to failures to the computer system or application program. It is very time consuming method, if amount of data is large. Most of the DBMSs provide the 'backup and recovery' sub-systems that automatically create the backup of data and restore data if required.

- Data Independence - The separation of data structure of database from the application program that uses the data is called data independence. In DBMS, you can easily change the structure of database without modifying the application program.

Meaning of "database system

A database system is a high-level definition of the structure and relationship between stored data, a database or databases, users and the hardware or operating system used for the storage

Components of DBMS

A database management system (DBMS) consists of several components. Each component plays very important role in the database management system environment. The major components of database management system are:

- Software
- Hardware
- Data
- Procedures
- Database Access Language

Software

The main component of a DBMS is the software. It is the set of programs used to handle the database and to control and manage the overall computerized database

1. DBMS software itself, is the most important software component in the overall system
2. Operating system including network software being used in network, to share the data of database among multiple users.
3. Application programs developed in programming languages such as C++, Visual Basic that are used to to access database in database management system. Each program contains statements that request the DBMS to perform operation on database. The operations may include retrieving, updating, deleting data etc . The application program may be conventional or online workstations or terminals.

Hardware

Hardware consists of a set of physical electronic devices such as computers (together with associated I/O devices like disk drives), storage devices, I/O channels, electromechanical devices that make interface between computers and the real world systems etc, and so on. It is impossible to implement the DBMS without the hardware devices, In a network, a powerful computer with high data processing speed and a storage device with large storage capacity is required as database server.

Data

Data is the most important component of the DBMS. The main purpose of DBMS is to process the data. In DBMS, databases are defined, constructed and then data is stored, updated and retrieved to and from the databases. The database contains both the actual (or operational) data and the metadata (data about data or description about data).

Procedures

Procedures refer to the instructions and rules that help to design the database and to use the DBMS. The users that operate and manage the DBMS require documented procedures on hot use or run the database management system. These may include.

1. Procedure to install the new DBMS.
2. To log on to the DBMS.
3. To use the DBMS or application program.
4. To make backup copies of database.
5. To change the structure of database.

6. To generate the reports of data retrieved from database.

Database Access Language

The database access language is used to access the data to and from the database. The users use the database access language to enter new data, change the existing data in database and to retrieve required data from databases. The user write a set of appropriate commands in a database access language and submits these to the DBMS. The DBMS translates the user commands and sends it to a specific part of the DBMS called the Database Jet Engine. The database engine generates a set of results according to the commands submitted by user, converts these into a user readable form called an Inquiry Report and then displays them on the screen. The administrators may also use the database access language to create and maintain the databases.

The most popular database access language is SQL (Structured Query Language). Relational databases are required to have a database query language.

Users (role of key players in database design and development)

The users are the people who manage the databases and perform different operations on the databases in the database system. There are three kinds of people who play different roles in database system

1. Application Programmers
2. Database Administrators
3. End-Users

Application Programmers

The people who write application programs in programming languages (such as Visual Basic, Java, or C++) to interact with databases are called Application Programmer.

Database Administrators

A person who is responsible for managing the overall database management system is called database administrator or simply DBA.

A database developer is an IT professional responsible for working on database technologies. Where database administrators are more focused on routine maintenance and support for an existing database setup, database developers tend to focus more on improving databases, expanding their range or functionality, or otherwise developing submissions for a company's IT architecture.

End-Users

The end-users are the people who interact with database management system to perform different operations on database such as retrieving, updating, inserting, deleting data etc.

Evolution of DBMS

1. The sixties and seventies: centralized

DBMS to the sixties and seventies (IBM IMS, IDS Bull, Univac DMS, etc.) Were totally centralized systems, as befits those years operating systems, and *hardware* for which they were made: a large enterprise-wide computer and a network of dumb terminals and memory.

The first DBMS in the sixties, yet they were called and were aimed at facilitating the use of large data sets in which the interrelationships are complex.The archetype of implementation was the *Bill of materials or explosion Parts*, typical in the automotive, construction of spacecraft and related fields. These systems worked only in batches *(batch).*

Appearing keypad terminals, connected to the central computer via a telephone line, they start to build great applications *on-line* transaction processing (OLTP). The DBMS *software* were closely linked to communications and transaction management.

Although to write application programs using high level languages such as Cobol or PL / I, were also available instructions and subroutines specialized to treat BD requiring that the programmer knew many details of physical design, and that made the was very complex programming.

Since the programs were related to the physical level, it should change continuously when changes were made in the design and organization of the database.The basic concern was to maximize performance: response time and transactions per second.

2. The Eighties: relational DBMS

Computers *minis*, first, and then *micro computers*, computer spread to virtually all businesses and institutions.

This required the development of applications would be easier.The DBMS of the seventies were too complex and inflexible, and could only use highly qualified personnel.

The emergence of relational DBMS * marks a significant step to facilitate the programming of applications with BD and to ensure that programs are independent of the physical aspects of the database.

* Oracle appears in 1980.

All these factors make greater use of the DBMS.Standardization, in 1986, the SQL language was a veritable explosion of relational DBMS.

Personal computers

During the eighties appear and spread very quickly on personal computers. It also appears these teams single-user *software* (eg, dBase and its derivatives, Access), with which it is very easy to create and use data sets, and that *personal data* are called *bases.* Notice that the fact referred to these early systems DBMS PC is a bit forced, as it does not accept complex structures or

relationships, or could be used in a network that simultaneously serve many different users.But some, over time, have been turning into real DBMS.

3. The nineties: distribution, C / S and 4GL

At the end of the eighties, and relational DBMS is used in virtually all businesses. Nevertheless, until the mid-nineties, when needed a high performance have continued to use the DBMS prerelational.

In the late eighties and early nineties, companies have found that their departments have been buying departmental and personal computers, and applications have been making BD. The result has been that within the company there are numerous DBMS BD and several different types or suppliers.This phenomenon of multiplication of the BD and the DBMS has been increased by the fever of mergers.

The need to have an overview of the company and of linking different applications using different BD, together with the ease that give the networks for communication between computers, has led to the current DBMS, which allow a program to work with different BD as if it were a single. This is what is known as distributed database.

This ideal distribution is achieved when the various BD are supported by one brand of DBMS, ie when there is homogeneity.However, this is not so simple if the DBMS are heterogeneous. Today, thanks largely to the standardization of SQL language, DBMS of different brands can be serviced at each other and work together to provide service to an application program. However, in general, in cases of heterogeneity can not be reached to give the program that uses the appearance that it is a single BD.

In addition to this distribution "imposed", wanting to separate the integrated treatment of pre-existing BD, you can also make a distribution "desired" BD designing a physically distributed and replicated with parts in different systems.The basic reasons for which are interested in this distribution are:

1) Availability. The availability of a distributed system with a BD can be higher, because if it goes down one of the systems, others still work. If the data residing in the system is not available are replicated on another system, continue to be available. Otherwise, only available data from other systems.

2) Cost. A BD can reduce the cost distributed. In the case of a centralized system, all users computers that can be distributed to different and distant geographical areas are connected to the central system via communication lines.The total cost of communications can be reduced by a user to close the data used most often, eg on a computer in your office or even on your personal computer.

The technology is commonly used to distribute data is known as environment (or architecture) Client / Server (C / S). All of the relational DBMS market have been adapted to this environment.

The idea of C / S is simple. Two different processes running on one system or separate systems, they act so that one has the role of client or a service requester, and the other server or service provider.

For example, a program that a user application running on your PC (which is connected to a network) requests some data from a DB that resides on a UNIX computer which, in turn, runs the relational DBMS that manages it.The application program is the client and the DBMS is the server. A client process can request services to multiple servers. A server can receive requests from many customers. In general, a process that makes customer requesting a service to another process B can also do a service server which prompted another process C (or B, that this request would be the client).Even the client and server can reside on one system.

The ease of distribution of data available is not the only reason, not even the basic, the great success of the environments C / S in the nineties. Perhaps the main reason was the flexibility to build and grow the global computer configuration of the company, as well as making modifications to it, using very standard *hardware* and *software* and cheap.

The success of BD, including personal computers, has led to the emergence of the *Fourth Generation Languages* (4GL), very easy and powerful languages, specializing in application development based on BD.They provide many facilities at the time to define, usually visually, talks to enter, modify, and query data to the C / S.

Current Trends

Today, relational DBMS are undergoing transformation to accommodate three recent successful technologies, closely related: multimedia, object-oriented (OO) and Internet and web.

The types of data that can be defined in relational DBMS of the eighties and nineties are very limited. The incorporation of multimedia technologies, image and sound-in makes it necessary to accept relational DBMS attributes of these types.

However, some applications do not have enough with the addition of specialized media types.Need complex types that the developer can define as the application. In short, we need to abstract data types: TAD. The latest DBMS already incorporated this possibility, and a wide open market or TAD predefined class libraries.

This brings us to the object-oriented (OO). The success of the OO at the end of the eighties, the development of basic software applications in industrial engineering and construction of graphical interfaces with users, has made during the nineties is widespread in virtually all the fields of computing.

In the SI is also initiated the adoption, shy of the moment, the OO.The use of languages like C + + or Java requires relational DBMS fit them with appropriate interfaces.

The rapid adoption of the SI web makes the DBMS server resources to be incorporated into websites, such as SQL scripts including HTML, Java embedded SQL, etc.. Notice that in the world of the web are common OO data and multimedia.

In recent years it has begun rolling out an application type of the BD called Data Warehouse, or data warehouse, which also produces some changes in the relational DBMS market.

Over the years I have worked with BD in different applications, companies have accumulated large amounts of data of all kinds.If these data are analyzed appropriately can provide valuable information *.

Therefore, it is a great BD keeping with information from all kinds of enterprise applications (and even outside). The data in this big warehouse, Data Warehouse, you get a more or less elaborate replication of which is in the BD used in the daily work of the company. These data warehouses are used exclusivelymind to make inquiries, most especially to carry out studies * financial analysts, market analysts, etc..

Currently, the DBMS is adapted to this type of application, including, for example, tools such as:

a) The creation and maintenance of aftershocks, with some data processing.

b) Consolidation of data from different sources.

c) The creation of physical structures that efficiently support multidimensional analysis.

* For example, market developments in relation to pricing policy.
* These are often multi-dimensional statistics.

Note: BD refer to Big Data

Definition of the database approach

The database approach is a way in which data is stored within a computer. It is organized into various charts that are accessed by a variety of computer applications from different locations. Databases are composed of a variety of information that is pertinent and relevant to the organization that is using the database

Database Approach vs. Traditional File Processing

- Self contained nature of database systems (database contains both data and meta-data).
- Data Independence: application programs and queries are independent of how data is actually stored.
- Data sharing.
- Controlling redundancies and inconsistencies.