

COMPUTER ORGANIZATION AND ARCHITECTURE

Lecture 1

Introduction to Database Technologies

Dr Victoria Mukami

INTRODUCTION

This topic is an introduction to databases and the database management system. We will review the evolution of the database through the file-based system. Through this, we will review the drawbacks of the file-based system that led to the development of the database approach. We will also review the characteristics and functions of the database approach. Finally, we will review the various database users and roles.

Learning objectives

By the end of this topic, you should be able to:

1. Understand and describe the file-based system
2. Highlight the benefits of the database approach
3. Show the functions of the database management system (DBMS)
4. Describe the roles involved in the use and development of the DBMS

OVERVIEW

Why databases? We use databases anytime we are interacting with a technological device. We create and collect so much data in our daily lives. Let us look at a typical day in the working life of an adult. When we wake up, we may use our phones as the first device. Whether to check email, check our social media accounts or even to meditate. All these activities will be picking information from various databases. As we move along the day, we may go by a bank and withdraw funds, use these funds at the supermarket to buy groceries and supplies. We may stop at a restaurant to have a meal and make a payment using a credit card. All this data, from the bank withdrawals to the receipt with a breakdown of items bought at the supermarket and the transaction details of the credit card all need to be stored somewhere. A database would be the most obvious place to store all this data. As you go about your day to day lives, ask yourselves some vital questions. First, am I creating any data? Second, where is this data being stored and third, am I retrieving any data from a central location?

INTRODUCTION TO DATABASES

We have thrown the word data and information around. What exactly are these terms? **Data** is referred to as raw facts [1]. These are small details that have not been processed and may not make any sense in their “raw” form. For instance, let us look at the supermarket example from the overview. When we walked into the supermarket, we picked a trolley and picked various items. Then we walked to the checkout counter

and the cashier proceeded to scan and charge the items. These items are stored in a table that may pick the number of items, the costs, the name of the item, the weight and perhaps the category. This is a good example of data. Single facts that are collected. When these are compiled then we get a term known as information.

Information is defined as processed data. When data is organized and becomes meaningful then it is considered as information. The organization will of course vary based on the use. Let us go back to our supermarket example. So, the cashier has entered all these items into the systems and compiled them to show how much you owe. You, therefore, go ahead and give her some money which she enters the amount into the system, and you get back to change and a receipt. This receipt has compiled all your items, has arranged all similar items and even shows the money given and the change returned. This is no longer data but information. Information is meaningful to the user. The receipt can easily tell you where and when you bought the items, how much you spent, the specific items you bought even who served you.

Why are we going on and on about data and information? Well, the data must be stored somewhere for us to be able to mine the data and get information. Before we can fully understand databases, we need to look back at what was used before databases came into the picture. The next section is a review of the file-based system.

FILE-BASED SYSTEM

The file-based approach is one of the obsolete technologies, however, is the predecessor to the database [2]. Before databases people used the file-based system. To understand the file-based system an example is needed. Think of your local hospital during an outpatient visit. If the hospital still operates with physical files, then one would first start at the reception where they would provide their details. The receptionist would then look up your details on a large book or on a computer to locate your file number. Once they had your file number, then they would retrieve the file from a large and albeit dusty room (of course this depends on when you were in hospital last). Once armed with your file, then it would be passed on to the triage nurse who will record your vitals and perhaps give an indication of what the problem is. You would then be transferred to the consultation room, where the doctor will diagnose you and will even have a chance to go through your file. The file will then be passed on to the pharmacy, lab or cashier depending on the diagnosis by the doctor. If you are one of

the lucky ones, your file was retrieved relatively fast, and the entire process took about 30 minutes. Now look at a scenario where they cannot find your file, they might write your details on a piece of paper which will be filed upon recovery of your file. Worse still, they might open a whole new file for you as they cannot retrieve your old file. Consider how many files the hospital may have. What happens when the hospital administrator needs some information like how many patients have a particular disease?

This is one way to look at the file-based system. In some cases, various files existed for the same client with several departments. The files may have had incorrect or incomplete data. Additionally, it is impossible to do data mining. To answer the question by the administrator, then the hospital would have to provide a record book (additional records) that could be used by the doctors to record the specifications of the administrator. While this case is dealing with manual files, it could also point to a situation where each department has its system and storage. Retrieving the records might be easy but what happens when triage needs to share information with the doctor or with the pharmacy? The file-based system had quite a several drawbacks.

Drawbacks to the file-based system

Separation and isolation of data [2]

As in our example, we see that there existed either one common file or different files for different departments. This means that it is truly a task and a hard one for one to conclude how many patients were treated for a specific disease and were given a specific drug. It may even be impossible to get that data in the event the file was moved to a different department.

Lengthy development times [2]

Assuming that our files were stored on a computer system by each department, it is still a daunting task for any programmer to come up with the various system. Pharmacy has its requirements of how their digital file should look like, so does the casualty department, the laboratories, and others. This even gives way to lengthy development where if the hospital needed to change one thing across the departments, might mean the programmer having to make changes for each department.

Duplication of data [2]

Based on the hospital case, if a patient has more than one file, then their data may be duplicated. It also means that the hospital cannot control the duplication of that data nor the additional files that may be required to hold that data. This is also true for digital copies.

Complex System Administration [1]

With even as few as five departments within the hospital, administration of the file-based system becomes a complex task. Now imagine when there are 20 departments. The number of tasks, files and personnel may have to increase exponentially.

Incompatible file formats [2]

In a scenario where the departments use different file formats then it is impossible to share the data. For simple sharing, a programmer would be needed to write some code that will be able to read from the different file formats.

Based on these limitations, it was necessary to look up a better way to be able to store data and share it across various users. This paved way for the database approach.

DATABASE APPROACH

The database approach was the emergence of the database and the database management system. A **database** is an assemblage of data [3]. A database is used to store various data that is useful to a business. Of course, the data collected should be related in one way or another. For instance, a university might collect data on students, the courses they are attending and grades. The size and complexity of a database will be dependent on its use and the needs of the organization. The database allows for simultaneous access by various users across many departments [2].

Characteristics of the database approach

Several characteristics can be used to describe the database approach. These include

Self-describing nature of a database [3]

A database system contains the database and a complete definition of the database structure and constraints [3]. These definitions are stored within the DBMS. This information may come in handy to any users who may require access to information about the structure of the database.

Data and program insulation [3]

The structure of the DBMS is normally stored separately from the access programs [3]. This is known as program-data independence. This means to changing the structure of the database is easy and future views will reflect the changes made.

Support of multiple views [3]

A **view** is defined as how a user sees a database. The fact that users may not be interested to see the backend of a database, nor will they be interested to see the structure of the database, does not mean other users will not be interested in the same. A database ensures that each user has their perspective.

Data sharing [3]

A database aims to allow for sharing of data across various departments and users. It ensures that through sharing users are more productive and can make better decisions.

Database Management System (DBMS)

A **database management system (DBMS)** is a system that allows users to be able to not only create but maintain and control a database [3]. A DBMS could be considered as general-purpose software. A DBMS can provide controlled security to a database as it prevents unauthorized access which could lead to a loss or destruction of data.

Components of the DBMS

Various components make up the DBMS environment.

- **Hardware**
These are all the physical or tangible components. It may include computers, printers, storage devices, network devices like hubs and modems [1]. The hardware devices allow for the users to interact with the database.
- **Software**
These are the software programs that make it necessary for the users to use and interact with the database. The software could include the operating system, the DBMS software and application programs.
- **People**

These are considered as all the users of a database [1]. Within this lecture, we will discuss the various users and roles within a database.

- Procedures

A procedure could be defined as an instruction or rule that say how something can be used. In this case, procedures are rules that indicate how a database can be designed and used [2].

- Data

Data as has been defined within this lecture is one of the most important parts of any database. For a database to exist, data has to be collected about something.

Advantages of DBMS

There are several advantages of using DBMS.

Improved data sharing [1]

Since users access the same data from one database, the users end up having better sharing capability

Improved data integrity [2]

Introducing constraints (consistency rules that a database cannot violate) within a database ensures that various rules are met. For instance, a patient cannot be older than 150 years.

Improved security [1]

Unlike the file-based system, the DBMS can improve security as it can authorize people and further restrict access based on their level.

Minimized data inconsistency [1]

Remember the example where our patient got a new file opened within the file-based system? Well, if the person who was taking down their name spelt it incorrectly, then you have two files with two different names. DBMS ensures that since there is only one record, the inconsistencies are minimized.

Better decision making [2]

Remember the administrator's query on patients treated with a specific disease? Well, this is now possible as the data is related and all one must do is pull up queries and make reports of the results.

Improved maintenance [2]

When the structure of the file changes, then the changes are seen across the board. It makes it easier to fix issues when one is using the DBMS approach. Additionally, updates are easier to handle.

Disadvantages of DBMS

While the DBMS offers quite a range of advantages, it has some key disadvantages.

Complexity [1]

Several technologies are used to manage databases as they hold data that is core to the running of the business. The personnel required to manage databases is exponential and increases the complexity as is seen by the different roles.

Cost [2]

The initial cost plus the cost of upgrading and maintaining a database can be massive. This is also in addition to the cost of the extra personnel and other direct costs.

Vendor dependence [1]

Based on the cost disadvantage, a company may not be willing to change vendors, and this may end up causing what is known as vendor dependence.

ROLES IN THE DATABASE ENVIRONMENT

Remember that one of the components of a DBMS is the people. There are various roles associated with a database.

Database Administrators

The database administrator (DBA) is the person in charge of managing the database. This involves maintaining the design of the database and ensuring that the database is secure. The DBA is also in charge of security and providing access to users.

Database Designers

Before a database can be used, it needs to be developed. Part of the development includes a logical and physical design. The designer is responsible for working with prospective users to ensure that the business and user needs are met. The role of a

designer is like an architect. They plan the structure of the house before the house can be built.

System Analyst and Programmers

A system analyst and programmer are individuals who also design but also implement the application programs that are used to access the database by the end-user

End-Users

These are the individuals who use the database either directly or indirectly. The end-users could be individuals like the triage nurse, the cashier or the doctor in the hospital setup.

SUMMARY

We have just concluded an introduction to database terminologies and technologies. A review of data and information has been done showing the major differences between data and information. File-based systems were explored, and the major limitations were highlighted. We have also reviewed the database approach, the DBMS and its components and the advantages and disadvantages of the DBMS. Finally, we looked at the different users of a database. The next lecture will focus on the database environment.

DISCUSSION TOPIC

For a very long time, users used the file-based approach. Some offices specifically government offices still use manual files. One such office is the lands office where each client's detail is stored on a physical file. Assuming the role of a database designer or expert, what advice would you give to the lands office to start the process of digitizing the files and using a shared database?

REFERENCES

[1] Database systems: design, implementation, and management, Coronel, C., & Morris, S, Cengage Learning, 2019.

[2] Database Systems: A Practical Approach to Design, Implementation, and Management, Connolly, T., & Begg, C., Pearson, 2015.

[3] Fundamentals of database systems, Elmasri, R., & Navathe, S. B., Pearson Education Limited, 2016.