EVOLUTION AND DEVELOPMENT OF DBMS IN SOFTWARE DEVELOPMENT INDUSTRY

[1]Karan Chauhan, [2]Ayush P Kumar, [3]Aleesha Kohli
[1]Systems Engineer, [2]Designation of 2nd Author, [3] Systems Engineer
[1]Telecom Domain, Database management,
[1]Infosys, Bangalore, India
[2]Telecom Domain, Database management,
[2]Lucknow, India
[3]Sales Domain, Testing database management,
[3]Infosys, Pune, India

Abstract— This issue informs about the history of database management systems through a series of pioneer recollections, principally from people who founded the major DBMS companies or were heavily involved in the growth and development of these products and companies. These eight recollections cover the principal DBMS software products for IBM mainframe computers. IBM itself was a significant player in this marketplace with its IMS product, but all the other products were produced and marketed by independent software companies. Many historians and industry analysts believe that these products and these companies formed the foundation on which the mainframe software products industry was built.

Index Terms—DBMS, Hierarchical Databases, Network Database, Relational Databases, Object-Oriented Model, ACID. (key words)

I. INTRODUCTION

Data has always played an important role in every phase of life, whether it be fields like health, sports, education, transport, etc. Data has provided us the ability to analyze and formulate a solution or understand a tread or simply develop a plan when it is in a organized collection and this is known as a Data Base.

Since the begin, the observations and the records registered by experiences or experiments have resulted in the output of formatted data which defined the process for the person handling the process. With the evolution of the processes and the method, the methods of storing the data, analyzing the data, and maintain the data has also evolved rapidly.

Due to which the data base has become a very important aspect in all processes and the proper understanding of data and its maintenance has become essential for business and product development. Managers need to understand high volumes of data before they can make the necessary decisions. Effective business intelligence (BI) tools assist managers with understand and formulating the patterns of the data which in turn helps them in decision making activities. [1]

II. EVOLUTION OF DBMS

The existence of database started with the use of computers from the age of Babbage. The first proper database came into existence in the 1960's by the brilliant people at IBM. This database was hierarchical in nature and called IMS, developed for the Apollo program. This database was effective and beneficial by making the data less redundant, independent and secure. Yet it had drawbacks in handling complex data.

To overcome the limitations of IMS, the creators of COBOL up with the concept of network DBMS at Honeywell called as Integrated Data Store(IDS) in1970's. In this model, each record can have multiple parents in comparison with one in the hierarchical DBMS. It created sets represented by one to many relationships between the owner and the member. The main limitation of this system was in its complex design and execution. [3][4]

In the year 1990, a new system was designed for entering data and working with big databases, where the idea was to use a table of records with one to one, one to many and many to many relations with other tables by E. Codd. This system of database resulted in the creation of the Relational DBMS. DBMS with this model took the object oriented approach which used the data in the form of text, multimedia, etc. [2][5]

III. INTRODUCTION TO DBMS

A DBMS (Database management system) is used to create and maintain the structure of a database, and then to enter, manipulate and retrieve the data it stores. Producing an efficient database design is the key to effectively using a database to support an organization's business operations [2]. A database management system (DBMS) is a computer software that manages databases, it may use any of a variety of database models, such as the hierarchical DBMS, network DBMS and relational DBMS. In large systems, a DBMS allows users and other software to store and retrieve data in a structured way. [3]

The Database Management can be classified in the following heading:

- Hierarchical databases.
- Network databases.
- Relational databases.
- Object-oriented databases

a. Hierarchical Databases (DBMS)

This model of the data base can be defined as the Tree structure which serves the same purpose as an entity–relationship (E-R), in which the records are stored in the groups of parent and child and maintain mainly a one to many relations between them. This model collects the data in the record form, and these records can be stated to be equivalent of tables with the individual records being the equivalent of rows. The connection between the records is made by links. [6]

This major benefit of this model was the rapid operation. But this model limited the relationships between the records, as the relationship between the children boxes or records was not allowed. These created issues in case complex and heavy data had to be used.

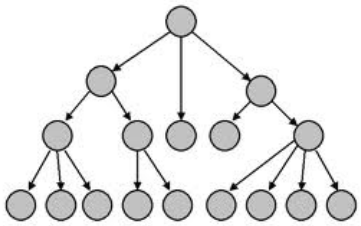


Figure 1 : Hierarchical Databases model

b. Network Database:

The network Database was similar to the Hierarchical Database, as it also supported a Hierarchical structure in this model. The only difference was that rather than having a straight flow relationship, in this method it was interlinked, like a web. The children were termed as members and the parent were termed as occupier. So as the structure was interconnected, now in this model the member could be linked between each other and could have more than one occupier. This empowered this system more flexibility. [8]

Due to this attribute of the web structure, this system supported many to many relationships in this model.

This model of Database enabled the developers and analysts to easily connect and correlate the data. But due to its complex structure pattern when handling large amounts of data, made it hard to work with and too complex to understand. [7] [9]

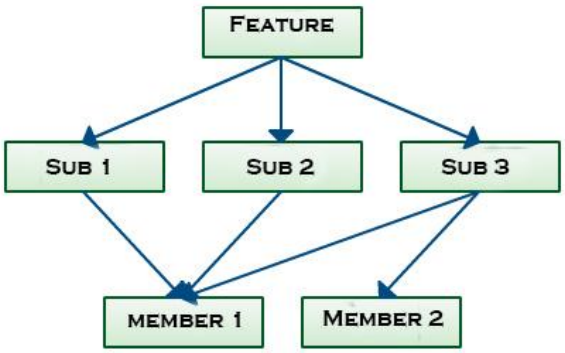


Figure 2 : Network Database model

c. Relational Databases:

The model of relation data base establishes a relational relationship between the data files. Unlike previous data bases, this data base does not support any hierarchy but connect the data directly in different files using a common data or field name.

The data is stored in different files or as it is termed in this model, it is stored in different control tables. The structure of these table is regular, with columns having field names and each row depicting a particular entry in the table. The values or entries in the table are mainly identifies based on the key field value. These values are unique to all the entries in the table.

This model is more reliable that the previous models of data base which follow hierarchical structure. The tables filled in this database are termed as Tuples, which refers to the designated row and the columns are referred to the attributes.

This databases' main advantage was that it was able to be dynamic in nature and also along with that it could easily handle large amounts of data without complicating the process or the structure of the storage, which make the understand and analysis of the data easy for the people operating on it. [9]

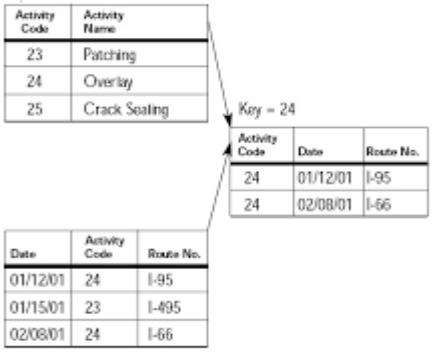


Figure 3 : Relational Databases Model

d.Object-OrientedModel:

The Object-Oriented Model has a wider scope of operations as compared to its counterpart models. This model can felicitate more functions rather than just storing data in a readable format. It provides features like increase the semantics of the languages like Java and adding functionality to the Object-oriented programming languages.

This approach is the analogical of the application and database development into a constant data model and language environment. The application can operate with lesser code, utilization of more natural data modelling and creating code bases which are easier to maintain.

The object-oriented database derivation is the integrity of object-oriented programming language systems and consistent systems. The power of the object-oriented databases comes from the cyclical treatment of both consistent data, as found in databases, and transient data, as found in executing programs. [9] [10]

This model is very vast in the scope of handling the data and the code but the cost of developing and maintaining this database was very high as compare to the RDBMS, this limited the utilization of this method, as the cost of any project is a high concern and big point in the final success of the project.

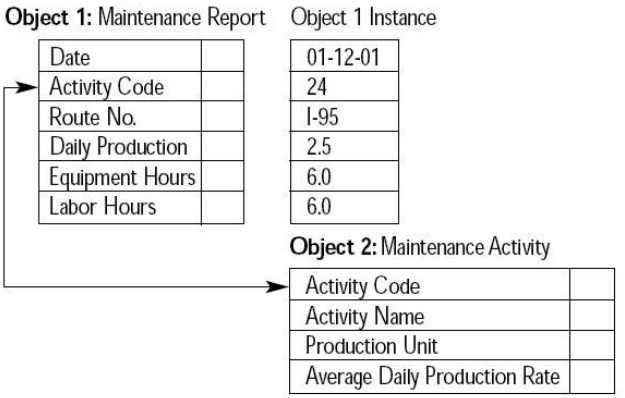


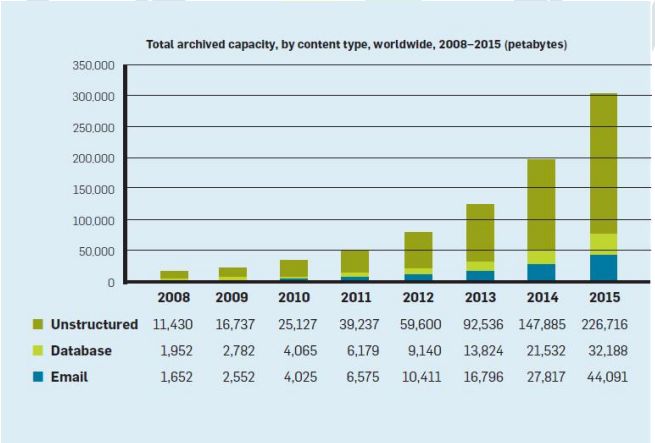
Figure 4 : Object-Oriented Model

IV. EVOLUTION OF THE ROLE OF DBMS IN THE SOFTWARE INDUSTRY:

With the dawn of programming languages and the need for new products to make life easy, the main focus was always on the development aspect of the process. The proper and efficient code writing was on high priority. But as the scenario increase and the scope of the projects became more complex along with the needs being more specific and targeted. The need to understand the environment came and the need to store the observations developed. This change in the scenario of software development industries brought the database management system at the centre of everything.

Initially the need of a database was simply, that is to enter, store and maintain the data produced. But later, the operations on the data became more complex and of higher importance. This resulted the database to be adaptive, more secure, and more dynamic in nature.

The trend of increasing data and because of which its increasing importance in the current scenario can be observed from the following graph: [11]



Graph 1 : Data growth Trend of the World

The evolution of the DBMS because of the needs resulted in the following value being inculcated in the real time industrial environment:

a) Data integrity:

The most important characteristics of database management system is the ability to maintain the integrity of the Data in objective. It ensures the quality, reliability and consistence of database system is not corrupted and compromised in any situation. It protects the unauthorized access of database and makes it more secure.

It is only possible by providing the ability to define and enforce certain constraints to ensure that users enter valid information and maintain data integrity. The constraints applied on the data are the restrictions or the rules that dictate the condition for the data that must be inserted in the database.

The movement of data and the constant updates by manual or automated means result in sometimes the compromise of the data quality. The quality of the data, results in the quality of the analysis or the work which must be carried out on that data. Then in such scenario the proper utilization of the active checks, the boundary conditions and the format constraints help in keeping the integrity of the data intact.

Hence, we can observe that the integrity of the data is an important attribute of the database.

b) Security:

As data has grown, the value of the data contained in the database has also grown rapidly. The data which is contained by a firm on the bases of their observations can be very sensitive and highly private. The data in the hands of wrong people can be harmful and destroy a lot of assets. Hence the property of data's security had to increase in the span of time.

Features like multiple access and viewing of data by different resources spread over a large geographical area require the need for the database to have security to some extent where users are unable to access data of other users and departments. The Database System offers methods to imply constraints and rules while entering data into the database and same rules and constraint in retrieving the same at a later stage. DBMS offers many different levels of security features, which enables multiple users to have different views with different features. Since a DBMS is not saved on the disk as traditional file systems, it is very hard for miscreants to break the code.

c) Independence of the Data:

In the file-based system which was stored on the physical memory and was previously widely utilised. The structure of the data files was defined in the application code itself. So, in case the developer that to change the structure of a file, under any given feature or situation then they had to access the original code as a whole and make the changes in the whole code. This aspect of the primitive system was a tedious job and was not very supportive of the changes later in the stages of development. As changing the values in the whole source code reduced the quality and increased the rate of error.

Whereas in the DBMS, the data structure is kept separate from the program code. That is, it is stored in the system catalogue and not in the programs file. This aspect of the code allowed the developer the freedom to make the changes to the structure of the database with one change. This insulation between the programs and data provided higher accuracy and make maintaining the database easier. This feature was called program-data independence.

Hence it can be summarized that system data descriptions or data describing data (metadata) are separated from the application programs or scripts. This is possible only due to the data structure being handled by the database management system and are not being embedded in the program code itself.

d) ACID Properties:

The quality of the data and the accuracy of the data in general terms was important when it came to manage and maintain large amount of data. Hence, the DBMS supported the feature of having high Accuracy, Consistency, Isolation and Durability, this property of the DBMS was known as ACID Property.

This feature of the DBMS makes sure that in every database the main purpose of the data or the quality of the data should not be lost while performing even easy and simple transactions like delete, insert an update.

It helps the data in retaining the quality a multi-transactional environment and in case of any error or failure.

e) Particular Language of DBMS:

The DBMS is equipped with a special and very useful communication feature with the developer or analysis, known as the query language. It provides the user the full and easy access to the database. The time taking tasks of retrieving and manipulate data in the database can be performed more efficiently and effectively.

The data user can apply different filters and condition using this language as required due to the features or scenario to retrieve the data. This language reduces the time, and effort of the user and increase its accuracy and quality.

The independence of the database can also be seen because of this feature on the DBMS.

Reducing the Data redundancy

The DBMS follows the rules of normalization. It can be defined as a process of organizing the data in database to avoid data redundancy, insertion anomaly, update anomaly & deletion anomaly.

It splits a relation when any of its attributes is having redundancy in values. Normalization has been proven mathematically to be rich and scientific process that reduces the redundancy in the data.

Hence making the database more reliable and durable in nature.

f) Co-relate Different Data to Meet New Requirements:

As the data from different sources and locations combine to make a database, so also the analysis and the data derived for future processes also need to be integration of these different kinds of data.

Hence the data should co-relate, and the user should be able to establish relationship between different kinds of data so that in the scenario of new requirements, the user is able to derive the required output faster from the database.

V. FUTURE DEVELOPMENT OF THE DBMS:

The Database Management System has come a long way since the start of the its implementation and proper usage. But still it has a long way to go. The data in the database is constantly increasing and is being more dynamic in nature due to the ever-changing environment. The database systems need to keep updating and evolving, to accommodate such conditions evolving.

The aspects of Security, efficient relation establishing between data on independent level, dynamic system for different platforms being used simultaneous on a single project, open and low-cost systems, etc. are the need of evolution for the upcoming age of database. These aspects of the data can increase the speed and accuracy of the operations on the data.

VI. CONCLUSION:

In this document we cover the history and the evolution of the database over the ages and due to the requirement of high quality of data, and it's easier access. It also equips us to understand where we stand with the database and also, provides us the insight of the future, where database is headed and what features and qualities would be required to provide a high functioning and high accuracy database for the ever-changing scenarios of the world.

The importance of a reliable, and high efficient database is a constant need. Seeing the trend some aspects of the database make to evolve more, and others need moving at the pace they are moving.

VII. REFERENCES:

[1] Wikipedia : Importance of Data
[2] ORACLE 9I SQL (LANNES L. MORRIS-MURPHY
[3] WIKIPEDIA : DBMS

[4] http://www.comphist.org/computing_history/new_page_9.htm
[5] History of Database Management Systems by Burton Grad and Thomas J. Bergin Published in IEEE Annals of the History of Computing (Volume: 31, Issue: 4, Oct.-Dec. 2009)
[6] Hierarchical Model - Yale codex.cs.yale.edu/avi/db-book/db6/appendices-dir/e.pdf
[7] Network Model - Yale http://codex.cs.yale.edu/avi/db-book/db6/appendices-dir/d.pdf
[8] Database Systems: Design, Implementation & Management (from the section 2.4.2) By Peter Rob, Carlos Coronel, Keeley Crockett
[9] http://www.c-sharpcorner.com
[10] Introduction to Object Oriented Modeling (Unit 1, Parts 1.2 and 1.5) www.fmuddce.org/pdf/mcs-032.pdf
[11] Instructed Data and the Enterprise by Paul Williams, part of Dataversity.
http://www.peerevaluation.org/data/8065d07da4a77621450aa84fee5656d9/PE_doc_30009.pdf