**Topics in Database Systems**

Topics that will be covered through the course of this paper are the several different types of databases such as hierarchical, network, etc. We will also cover the difference between a row based database and column based database.

**Database Types**

Depending on where you look, you’ll see a different numbers for the amount of database types that there are. We will be discussing four database types. The six database types are hierarchical, network, object oriented, relational, centralized, and decentralized.

“The schema for hierarchical databases is defined by its tree-like organization, in which there is typically a root “parent” directory of data stored as records that links to various other subdirectory branches, and each subdirectory branch, or child record, may link to various other subdirectory branches.” (1) This means that there is a parent directory that has links to other directories below it. It is there that these lesser directories are linked to other branches and receives its tree-like organization. This database structure dictates that a parent directory can have several child directories, while to the contrary a child directory can only have one parent directory. Hierarchical databases are very useful when you need to represent data in a tree like manner.

The advantage of a hierarchical database is that the model created from this database can be understood quickly and is very simple due to its one-to-many relationship format. This database model allows users to add or delete information with ease as well as allowing the top of the hierarchy to be accessed quickly. These databases also support system that work through one-to-many relationships. The main disadvantage of hierarchical database is its lack of flexibility. Due to the nature of this database type having a one-to-many structure, it means this structure is not ideal for complex structures. “Also the tree-like organization of data requires top-to-bottom sequential searching, which is time consuming, and requires repetitive storage of data in multiple different entities, which can be redundant.” (1)

This database type is still widely used in many fields where high performance and easy concept are the parameters. The hierarchical database is efficient for one-to-many relationships and is widely used in recording file system data. “The most popular hierarchical databases are IBM Information System (IMS) and RDM Mobile. Windows Registry is another example of a real-world use cases of a hierarchical database system.” (2)

With the network database system, it was a progression from the hierarchical database system and was meant to be a more flexible. The network database system achieves this flexibility by allowing multiple records to be linked to the same owner file Because of the ability to have multiple links, the network database system was very flexible and could represent many-to0many relationships as well as one-to-many relationships unlike its counterpart where the hierarchical database system could only be used with one-to-many relationships. The network database system retains a similar tree-like appearance similar to the hierarchical database.

The advantages of the network database system are that it allows many-to-many relationships. “The hierarchical database does not account for human organizational events that require an entity to have many upward links to other entities.” (3) The disadvantages of the network database are the added complexity over hierarchical databases. This newer network database is difficult to handle and maintain. “Although the Network model is more flexible than the Hierarchical model, it still has flexibility problems. Not all relations can handle by assigning them in the form of owners and members.” (4)

An object-oriented database or object database management system is a database that is based on project-oriented programming. Object-oriented databases can also be called object databases or object-oriented database management systems. “The database combines object-oriented programming concepts with relational database principles.” (5) A main difference between this database system and the previous two is that the data is represented and stored in the form of objects. These objects are then stored in memory. These objects have members such as fields, properties, and methods.

The objects are basic building blocks and instances of a class where the type is either built-in or user defined. This database type also features classes which provide a schema for objects which in-turn defined the behavior of these objects. The methods determine the behavior of the class, and the pointers help access elements of an object database and establishes relationships between objects. Since this database system deals with the data as objects, rather than outputting multiple tables, the object-oriented database instead outputs one available object package.

The advantages of using an object-oriented database system are that it allows the use of complex data and a wider variety of data types in comparison to MySQL data types. Object-oriented databases make it easy to integrate with object-oriented languages due to sharing concepts. This database type is also easy to model advanced real-world problems and allows for easy save and retrieval of data quickly.

The disadvantages of object-oriented database systems is that they are not as widely adopted as relational databases. Since there is no universal data model, it lacks theoretical foundations and standards. This database type is also very complex and can cause some performance issues for its users. It also does not feature an adequate security mechanism and access rights to the objects.

“A relational database is a type of database that stores and provides access to data points that are related to one another.” (6) In relational databases, each row in the table is a record with a unique ID called the primary key. As for the columns, these hold the attributes of the data, and each record generally has a value for each attribute, which then makes it easy to establish relationships among the data. The relational database is based on the relational model, and that means that the logical data structures such as the data tables, views, and indexes are separate from the physical storage structures. “This separation means that database administrators can manage data storage without affecting access to that data as a logical structure. For example, renaming a database file does not rename the tables stored within it.” (6)

The advantages of relational databases are their simplicity because they do not require any complex structures or querying processes. It also reduces redundancy due to the use of primary and foreign keys when relating multiple tables together. This leads to a higher accuracy then other database. The relational database is also scalable and provides a flexible structure to meet changing requirements and increasing amounts of data. “The relational model permits changes to a database structure to be implemented easily without impacting the data or the rest of the database.” (7)

With anything, there are always some disadvantages with the advantages. Relational databases have varying performance depending on the number of tables present in the database. With a higher table count, the response time to queries will be slower as opposed to a relational database with far fewer. Also, due to the column and row nature of the database type, it also requires a large amount of physical memory, and each operation depends on separate physical storage. With enough careful and proper optimization, the target application can have maximum physical memory.

“A centralized database is one that operates entirely within a single location.” (10) What this means is that the database is hosted on a central computer, and that computer runs and maintains the database and grants users the ability to be able to access the database through a computer network. These types of databases are typically used in bigger organizations because an organization may have several departments with different business processes that run simultaneously. “The basic function of a centralized database is to have a complete view of the data of the organization. Having a complete view of data reduces the time consumption in implementing the actions as the organizations do not have to analyze the data separately for various departments.” (11)

The advantage of a centralized database is that it guarantees a minimization of redundancy in data. Since the data is located solely in one location rather than several locations, this leads to the reduction of redundant data. This database type also provides a higher level of security due to protocols that organizations put in place for the individuals that have any involvement with the data. Centralized databases also increase the integrity of data due to the data being present in a single location. This means that whenever data is altered, the data is only changed the singular time due to their only being one instance of that data.

The disadvantage of centralized databases is that it may reduce the efficiency of the system when multiple users attempt to access data simultaneously. There also comes the risk of data loss or inaccessibility due to the database being located solely on one computer. Should an organization suffer any outages, due to the database being located on one computer, should that computer become offline, users would not be able to access the database.

“In decentralized systems, every node makes its own decision. The final behavior of the system is the aggregate of the decisions of the individual nodes. Note that there is no single entity that receives and responds to the request.” (12) Some advantages of a decentralized system is that more than one central unit exists in the system which can listen for connections from other nodes. Due to having several central nodes, if a central node was to fail, only part of the system would fail rather than the whole system would if it were a centralized system. Decentralized database systems has actually gained popularity, primarily due to bitcoin.

Some disadvantages of decentralized system is that there is no regulatory oversight. It is also difficult to tell which node failed if one were to as well as knowing which node responded. In a decentralized system, the request is actually served by one of the nodes in the system but it is actually difficult to find out which node served the request.

Knowing the different types of databases accessible makes choosing which one to use no easy feat. Each organization must evaluate and decide which database type they would see as a best fit for the organization. While a relational database is the most popular, if their data featured only one-to-many relationships, then hierarchical databases might be the best course of action rather than attempting to implement and force a database into a system it is not compatible with.

**Row Vs. Column**

When organizing relational databases, there are two methods to orient your database. These methods are row or column oriented. There are different databases that are either row or column oriented. For example Postgres and MySQL are row oriented and Redshift, BigQuery, and snowflake are column oriented.

“Row oriented databases are databases that organize data by record, keeping all of the data associated with a record next to eachother in memory.” (8) This method of orienting the database is the traditional way of organizing data. This method is great for storing data quickly and is optimized for reading and writing rows of data. Unfortunately, this orientation is not efficient in performing operations applicable to the entire dataset which makes them more expensive jobs.

“Column oriented databases are databases that organize data by field, keeping all of the data associated with a field next to each other in memory.” (8) This method of orienting the database has grown in popularity and provides some performance advantages in querying data. With this orientation, read and write operations are slower as compared to row oriented. “These are efficient in performing operations applicable to the entire dataset and hence aggregation over many rows and columns.” (9)

**References**

1. “What Is a Hierarchical Database?” *OmniSci*, <https://www.omnisci.com/technical-glossary/hierarchical-database>.
2. Chand, Mahesh. “What Is a Hierarchical Database.” *C# Corner*, <https://www.c-sharpcorner.com/article/what-is-a-hierarchical-database/>.
3. “Advantages & Disadvantages of Network Databases.” *Techwalla*, <https://www.techwalla.com/articles/advantages-disadvantages-of-network-databases>.
4. *Network Data Model*, <https://www.tutorialspoint.com/Network-Data-Model>.
5. “What Is an Object-Oriented Database {Concepts, Examples, Pros and Cons}.” *Knowledge Base by PhoenixNAP*, 2 Sept. 2021, <https://phoenixnap.com/kb/object-oriented-database>.
6. “What Is a Relational Database?” *Oracle*, <https://www.oracle.com/database/what-is-a-relational-database/>.
7. “What Are the Advantages of a Relational Database Model?” *Techwalla*, <https://www.techwalla.com/articles/what-are-the-advantages-of-a-relational-database-model>.
8. “Row vs Column Oriented Databases.” *The Data School*, <https://dataschool.com/data-modeling-101/row-vs-column-oriented-databases/>.
9. “Difference between Row Oriented and Column Oriented Data Stores in DBMS.” *GeeksforGeeks*, 30 Aug. 2019, <https://www.geeksforgeeks.org/difference-between-row-oriented-and-column-oriented-data-stores-in-dbms/>.
10. “What Are the Different Types of Databases?” *Indeed Career Guide*, <https://www.indeed.com/career-advice/career-development/types-of-databases>.
11. “Centralized Database: Why Do We Need a Centralized Database?” *EDUCBA*, 4 Mar. 2021, <https://www.educba.com/centralized-database/> .
12. “Comparison - Centralized, Decentralized and Distributed Systems.” *GeeksforGeeks*, 13 Sept. 2021, <https://www.geeksforgeeks.org/comparison-centralized-decentralized-and-distributed-systems/> .