Name: Abenes, Enrico O.
Autor, Flix Nino B.
Subject & Section: CC 5-2I

Date: September 11, 2024
Lab Activity No. 2

HISTORY OF DATABASE

I. History

The concept of data management dates back to the early 20th century with the development of data processing systems. Before databases, data was stored in physical files and records, which were managed manually. The need for more efficient data handling became apparent as computing technology advanced.

A. Hierarchical Database Model

The hierarchical database model emerged in the 1960s. It organizes data in a tree-like structure, where each record has a single parent and potentially many children. IBM's Information Management System (IMS), introduced in 1966, was one of the earliest and most prominent examples of this model. IMS was used extensively by industries like aerospace and banking for its ability to manage complex relationships and large volumes of data.

B. Relational Databases

In the 1970s, a significant advancement came with the development of the relational database model. Proposed by Edgar F. Codd in 1970, this model organizes data into tables (relations) that can be linked by common data attributes. Unlike hierarchical databases, the relational model allows for more flexibility in querying and data management. Codd's work led to the development of SQL (Structured Query Language), a standard language for managing relational databases.

C. Emergence of SQL

SQL was developed in the early 1970s by IBM as part of their System R project. It became a standard language for interacting with relational databases, enabling users to query, update, and manage data efficiently. SQL's introduction greatly simplified database interactions and contributed to the widespread adoption of relational databases.

D. NoSQL Databases

The late 2000s saw the rise of NoSQL databases, which emerged to address limitations in relational databases, particularly for handling unstructured data and scaling horizontally across distributed systems. NoSQL databases, such as MongoDB and Cassandra, provide flexible schemas and are optimized for specific types of data and queries. They have become crucial in handling big data and real-time applications, like social media and e-commerce platforms.

II. Findings

A. Timeline and Contributions

1. **Hierarchical Databases (1960s)**
    - The hierarchical model introduced structured data organization and was an early step towards efficient data management.
    - **Example:** IMS was used by companies like Boeing and NASA for mission-critical applications.
2. **Relational Databases (1970s):**
    - The relational model revolutionized data management with its flexibility and ease of querying. SQL allowed for complex data manipulation and analysis.
    - **Example:** Oracle, which was founded in 1977, became a major player in relational databases, used in industries ranging from finance to healthcare.

3. **SQL (1970s onward):**
    - SQL provided a standardized language for database management, making it easier for developers and analysts to work with relational databases.
    - **Example:** MySQL, a popular open-source relational database, became widely used in web development due to its ease of use and support for SQL.
4. **NoSQL Databases (2000s):**
    - NoSQL databases addressed the need for scalability and flexibility in handling diverse data types and large volumes.
    - **Example:** MongoDB, a NoSQL database, is used by companies like eBay and LinkedIn for managing large-scale, real-time data applications.

B. Impact on Data Storage, Retrieval, and Management

1. **Hierarchical Databases**
    - **Data Storage and Retrieval**

      Hierarchical databases, like IBM's IMS, structured data in a tree-like format with a single parent-child relationship. This structure facilitated efficient data organization but limited flexibility. Navigating complex hierarchies for data retrieval could be cumbersome, and any structural changes required significant reorganization.
    - **Management**

      Managing hierarchical databases involved maintaining a fixed hierarchy, which could be complex and inflexible. Adding or altering data types often required extensive modifications to the hierarchy, limiting adaptability.
2. **Relational Databases:**
    - **Data Storage and Retrieval**

      The relational model, introduced by Edgar F. Codd, organized data into tables that could be related through attributes. SQL queries enabled efficient data retrieval across multiple tables, improving the flexibility and efficiency of data access. This model simplified complex queries and data manipulation.
    - **Management**

      Relational databases introduced normalization to reduce redundancy and improve data integrity. SQL provided a standardized and powerful interface for managing data, making it easier to perform various operations and supporting robust database management tools.
3. **SQL**
    - **Data Storage and Retrieval**

      SQL, developed by IBM in the 1970s, allowed users to specify what data they wanted without detailing the retrieval process. This made complex data queries and aggregations easier and more intuitive.
    - **Management**

      SQL's standardization facilitated the widespread adoption of relational databases, providing a consistent way to interact with data and supporting transaction management for data integrity in multi-user environments.
4. **NoSQL Databases:**
    - **Data Storage and Retrieval**

      NoSQL databases, such as MongoDB and Cassandra, handle diverse data models like documents, key-value pairs, and graphs. They offer flexible schemas and horizontal scaling, making them suitable for large-scale and real-time applications.
    - **Management**

      NoSQL databases focus on scalability and performance, often using eventual consistency models to handle large volumes of data efficiently. They address limitations of traditional relational databases, particularly for unstructured data and high-velocity transactions.

C. Significant Figures and Technologies

1. **Edgar F. Codd**
   - Edgar F. Codd developed the relational database model in 1970, introducing the concept of organizing data into tables and defining relationships using keys. His work led to the creation of SQL.
   - **Impact:** Codd's relational model and normalization principles revolutionized data management, influencing the design of modern relational database systems and their widespread adoption.
2. **IBM**
   - IBM developed the hierarchical database IMS in 1966 and created SQL through the System R project in the 1970s. These innovations laid the foundation for relational databases and standardized database interactions.
   - **Impact:** IBM's work established key database technologies and standards, shaping the development of relational databases and influencing subsequent database systems.
3. **MySQL**
   - Launched in 1995, MySQL is a prominent open-source relational database known for its ease of use and cost-effectiveness. It supports SQL and has become a popular choice for web applications.
   - **Impact:** MySQL's wide adoption in web development (e.g., WordPress, Facebook) demonstrates the practical application of relational databases and highlights its role in modern software development.
4. **MongoDB:**
   - Introduced in 2009, MongoDB is a leading NoSQL database using a document-oriented model. It supports flexible schemas and horizontal scaling, catering to diverse data needs.
   - **Impact:** MongoDB's flexibility and scalability make it ideal for real-time and big data applications, addressing limitations of relational databases and supporting modern, high-performance data systems.

REFERENCES:

Abiteboul, S., Hull, R., & Vianu, V. (1995). Foundations of Databases. *Sigmod*. Retrieved from https://www.sigmod.org/publications/dblp/db/books/dbtext/abiteboul95.html

Berrington, J. (2007, November 28). Databases. *Science Direct*. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S147202990700241X

Gupta, A., Tyagi, S., Panwar, N., Sachdeva, S., & Saxena, U. (2017, October). NoSQL databases: Critical analysis and comparison. *IEEExplore*. Retrieved from https://ieeexplore.ieee.org/abstract/document/8284494/authors#authors

Jaoua, A., Belkhiter, N., Ounalli, H., & Moukam, T. (1997). Databases. *Springer Link*, 197-210. Retrieved from https://link.springer.com/chapter/10.1007/978-3-7091-6510-2_13

Revesz, P. (2010). Introduction to Databases. *Springer Link*. Retrieved from https://link.springer.com/book/10.1007/978-1-84996-095-3

Silberschatz , A. (1990). Database systems: Achievements and opportunities. *ACM Sigmod Record*, 1 - 11. Retrieved from https://dl.acm.org/doi/pdf/10.1145/125223.125272