

JOURNEY FROM SQL TO NoSQL

Jinisha Kanjiya¹, Jay Lumbhani², Prakash Gupta³, Dr. Priya Swaminarayan⁴

¹Student of MCA Department, Parul University, India

²Student of MCA Department, Parul University, India

³Student of MCA Department, Parul University, India

⁴Dean of MCA Department, Parul University, India

Abstract— A Database Management System (DBMS) is system software for creating and managing databases. There are mainly two types of Databases. Relational Database and Non-Relational Database. SQL is an example of Relational Database while NoSQL is an example of Nonrelational Database. Mainly SQL supports ACID properties while NOSQL supports CAP theorem. Hierarchical Database, Network Database, Relational Database, Object Oriented Database are types of SQL and Key –Value Databases, Document Store, Column Oriented Databases, and Graph Databases are examples of NOSQL. SQL only works with structure data but NOSQL allows to working with semi – structure data and unstructured data as well. We cannot say which one is the best database because all have their own advantages and some disadvantages too. Which database should use it is totally depending on situation? In this paper, we have covered introduction of SQL / RDBMS / Relational databases and Non – Relational database. Types of Relational Data Base and Non - Relational databases, Pros & cons of SQL and NOSQL and conclusion.

Keywords— Relational Database, Non-Relational Database, SQL, NOSQL, ACID, CAP, Semi Structure,

1. Introduction to databases

This A Database Management System (DBMS) is system software for creating and managing databases with a systematic way. It provides users and programmers to create, retrieve, update and manage data.[1] DBMS applications store data as files. for well form and related data the RDBMS are comes into the picture. RDBMS stands for Relational Database Management System. Its applications store data in a tabular form, in a RDBMS, the tables have an identifier called primary key. Relational Database Management System

(RDBMS) is an advanced version of a DBMS system. It came into existence during 1970's. It is also 2 allowing the organization to use data more efficiently than DBMS. [2] Oracle, IBM, Microsoft, there are several open source version of SQL databases available such as MySQL, Maria DB. It is easy to write SQL queries and easy to learn too, and SQL databases use an established standard, which are adopted by ANSI & ISO. But it is only allowing a structure data

2. Types of Relational Database

2.0.1 Hierarchical Database

This model organizes data into a tree-like structure, where each record has a single parent or root node. Sibling records are sorted in a particular order. That order is used as the physical order for storing the database. [4]

2.0.2 Network Database

This model is allowing multiple records to be linked to the same main file. The model can be seen as an upside-down tree where the branches are the member information linked to the main, which is the bottom of the tree. The relationship that the information has in the network database model is defined as many-to-many relationship. [12]

2.0.3 Relational Database

In relational model each table is a group of column and rows, where column represents attribute of an entity and rows represents records. [11]

2.0.4 Object Oriented Database OOD model defines a database as collection of objects, or reusable software elements, with associated features and methods. There are several kinds of object-oriented databases: [4]

1.Multimedia database:

It incorporates media, such as images, that could not be stored in a relational database

2. Hypertext database:

It allows any object to link with other object. It's useful for organizing lots of unclear data, but it's not ideal for numerical analysis. This model is known as a post-relational database model.

3.0 Introduction to NoSQL / Non - Relational databases NoSQL is a Non-Relational Database management system which differs from Relational Database management system. It does not use table form for storage of data and SQL queries for retrieval of data as relational database management system does traditionally. It does not follow any schema that is why it is also called schema less or not fixed schema database system. It can handle semi-structured and unstructured data very efficiently but relational databases have problem while handling it [6]. many companies using this NoSQL like Facebook, Netflix, Twitter and LinkedIn.

Invention of NOSQL: NoSQL is getting a lot of popularity these days but it is not new in market. The name 'NoSQL' was first used by Carlo Strozzi. In 1998 they described his open source relational database, which was a lighter version. It is a relational database but without a SQL interface in it. This term was again used in 2009 by Eric Evans for naming Non-Relational databases [6]. Why NOSQL: There is a relational database for the purpose of managing data then why we need NoSQL databases. Answer of same is given in following points: 1. Horizontal Scalability: 2. Workload Distribution 3. Big data Applications 4. Performance 5. Continuous Availability 6. Large Database 7. Higher Throughput To overcome the problem of consistency NOSQL adopt CAP theorem. It was proposed by Eric Brewer in 2000 and CAP stands for [6]:

3.1 Types of Non-Relational Database

3.1.1. Key –Value Databases:

It is a combination of two main attributes one is key and other is value that is key corresponds to a value or group of values. As an example: It is based on hash table where key is a unique and points to value. Value can be duplicate. It is a dictionary data structure which is used for insertion, deletion and search by unique key. Example. hash table and binary tree. A key can be point to different number of attributes. This database support scalability Example: It can handle a large amount of data.

among all of them this is a one of the simple databases. It supports unstructured data. And It is useful for quickly finding data from database. It uses less memory for store data. Better performance can be achieved by using these databases. Example of key-value databases: Redis, Riak, Memcached, Berkeley DB, Tokyo cabinet. If values are missing, then this will be fine in case of key-value databases. [6]

3.1.2 Document Store

Document database is consisting of two main attributes:

1)key 2) document Example: key corresponds to one document. Key used is unique key Example: different key is used for different document. Here in this database data is stored in the form of document. It is support: Structured data, Unstructured data(text) or semi structured data (XML). hashing's technique is used in document store databases. Data can be inserted, deleted and updated in document using unique key pointing to document. Example of document database: Couch DB, Mongo DB, JSON format is used in document store database with dynamic schema. Dynamic schema means not pre define schema, different documents can have different number of fields [6]. JSON stands for JavaScript Object Notation. JSON supports all basic data types like numbers, strings, Boolean, arrays In document oriented databases dynamically fields can be added by using unique key. 5 One of the document database examples is Mongo DB, there are many cloud providers who allows to create a mongo DB database on their cloud. Below figure shows same:



Figure 1: Create Database (MONGODB)



Figure2: Insert Document (MONGODB)

3.1.3 Column Oriented Databases:

Column oriented databases are also known as Extensible Record Databases. [6], These databases also contain rows and columns like relational databases. Column is a key value pair in these databases. Each column is stored in a separate file and for every row column is at same offset that is called as auto indexing [6]. Columns stored in separate files: User Table. While querying, only required column are scanned and rest are not scanned that is why they are faster in read. These databases are faster while querying than relational databases. Example of column-oriented databases: Cassandra, HBase, Bigtable. Figure Demonstration of how data stored in column-oriented databases.

In these databases, we are able to add new columns at any time as we just need to create a new file but in case of relational databases, we need to rebuild table

3.1.4 Graph Databases:

Graph databases are based on the concept of graph theory. In these databases, data is represented in form of graphs on the behalf of interconnection between data It is a collection of nodes and edges where 1. Nodes represent entities. 2. Edges represent relationship among nodes. These databases are used for shortest path calculation. These databases are scalable in nature but drawback is that the complexity also gets increased. E.g. of graph databases: Neo4J.

4.0 Pros and Cons of RDBMS and Non-RDBMS

Pros & cons

Relational Databases: Pros

- Relational databases always work with structured data.

- This DB support ACID transactional consistency and support “joins.”
- It come with built-in data integrity and a large eco-system.
- limitless indexing is there.

Relational Databases: Cons

- Relational Databases do not scale out horizontally very well, only vertically.
- Data is normalized, meaning lots of joins, that affects speed.
- They have problems it working with semi-structured data.

Non-relational / NOSQL: Pros

- Schema-free.
- High availability.
- While many NoSQL Data is denormalized.

Non-relational / NOSQL: Cons

- Weaker or eventual consistency instead of ACID.
- Limited support for joins.
- Requiring mass updates (i.e. product name change).
- Does not have built-in data integrity.
- Limited indexing.

5.0 Conclusion

Now a day instead of SQL Database NoSQL database are used. NoSQL databases have many advantages over relational databases that is why it is currently adopted by many companies. Reason of their large rate of adoption they are becoming very popular among rest of databases. But this is also true that choosing right database also depends upon the situation, we cannot say that which one is better. According to situation, we can also combine these databases in one application for better performance. All the technologies are best in there on way. It is up to the Developer to make a better use of them depending on the situations and needs.

6.0 References

Web

- [1] DBMS, <https://www.alooma.com/blog/types-of-modern-databases>
- [2] RDBMS,

<https://www.guru99.com/difference-dbms-vs-rdbms.html>

[3] DBMS Transaction,
https://www.tutorialspoint.com/dbms/dbms_transaction.htm

[4] Types Of RDBMS,
<https://www.lucidchart.com/pages/database-diagram/database-odels?a=0>

[5] Hierarchical Database,
<https://www.studytonight.com/dbms/database-model.php>

[7] NOSQL,
<https://www.dezyre.com/article/nosql-vs-sql-4-reasons-why-nosql-is-better-for-big-data-applications/86>

[8] SQL model,
<https://mpowersoft.wordpress.com/2015/08/31/sql-nosql-everything-you-need-to-know/>

[9] Importance of SQL & NoSQL,
<https://tableplus.io/blog/2018/08/why-sql-is-the-most-important-skill-to-learn.html>

[10] NOSQL,
https://www.google.com/search?q=nosql+type&rlz=1C1SQL_enIN793IN793&source=lnms&tbm=isch&sa=X&ved=0ahUKEwjYwPT7rKngAhVWSX0KHLYLmACYQ_AUIDigB#imgrc=xfztgqOrp14tiM

[11] Relational database,
<https://beginnersbook.com/2015/04/relational-model-in-dbms/>

[12] Network Database,

[13] <http://www.computerbusinessresearch.com/Home/database/network-database-model>

Research Article

[6] Priyanka," NoSQL Databases, Types and Comparison with Relational Database", ISSN 2229- 5518, pp:-4963 to 4966

<http://ijesc.org/upload/91236a5cc5518a333b9f54fa57dd9757.A%20Review%20of%20NoSQL%20Databases%20Types%20and%20Comparison%20with%20Relational%20Database.pdf>