## A COMPARATIVE STUDY OF SQL AND NoSQL DATABASES

**Mrs. Netraja Mulay**
Assistant Professor
Modern College of Engineering, Pune

**ABSTRACT:**
The Traditional Relational Database Systems (RDBMS) suffers from various limitations from the view point of volume, variety, velocity and veracity. In order to overcome the above limitations a new concept of NoSQL databases is introduced. NoSQL databases also called as Not Only SQL which becomes popular because of scalability and performance as well as their ability to handle huge amount of unstructured data. It's a schema less presentation of the data and does not require a predefined query language to store and retrieve data however the data in these databases can be managed through various NoSQL database types these are key value store, graph, columnar and document store databases. The aim of this paper is to provide a comparison of various NoSQL databases with the traditional Relational databases based on major parameters.

**Keywords**: Volume, Variety, Velocity, Veracity, Unstructured data, Key value store, Graph, Columnar, Document store

### INTRODUCTION

Today with the rapid use of internet huge amount of data is generated. Several online transactions also generate a massive amount of data. To manage such a data we require an efficient storage and fast retrieval systems .A Data Base Management System (DBMS) plays a very important role which is responsible of storage, retrieval, updating, insertion, and deletion etc.

### BACKGROUND

Two types of databases exist to manage the data SQL and NoSQL databases

### 1. SQL Database

The Relational Data Base model was invented by Edger F. Codd. It is a Relational Database Management System in which the data is arranged in the form of tables which are consists of rows and columns. Each row represents a record or tuples whereas each column represents an attribute or properties or fields. The relationships amongst tables are represented by the use of different keys such as primary key, foreign key etc. SQL databases are effective for transaction management concepts i.e. it supports strongly ACID (Atomicity, Consistency, Isolation, Durability) properties.

1. **Atomicity**:-The transaction should get completed or not completed at all
2. **Consistency**:-A transaction does not leave a database in an inconsistent state.
3 **Isolation**:-One transaction should not be involved in other.
4. **Durability**:-A complete transaction persists even though application restarts

The SQL (Structured Query Language) acts as an important tool to extract the data from the database. However Relational Data Base Systems suffers from variety of limitations such as scalability, storage, processing larger volume of the data and efficiently losing of query as volume of data goes on increasing etc. In order to overcome such limitations a new concept of NoSQL databases are introduced.

### 2. NoSQL Database

The NoSQL database is invented by Carlo Strozzi. NoSQL also called as "Not only SQL" is a non-relational database system in which there is no any fixed database schema. These databases are not

built on tables and hence they do not use SQL for data manipulation. These databases use the concept of BASE (Basic Availability, Soft state, Eventual consistency) properties.

**1. Basic availability**:-Each request is either successful or failed.

**2. Soft state**:- The state may change at times for eventual consistency.

**3. Eventual Consistency**:-The database may be temporary inconsistent but will be consistent eventually.

NoSQL databases are known to provide easier scalability, storage flexibility, and greater data manipulation and performance improvement. BASE is derived from CAP theorem which is used for distributed systems.
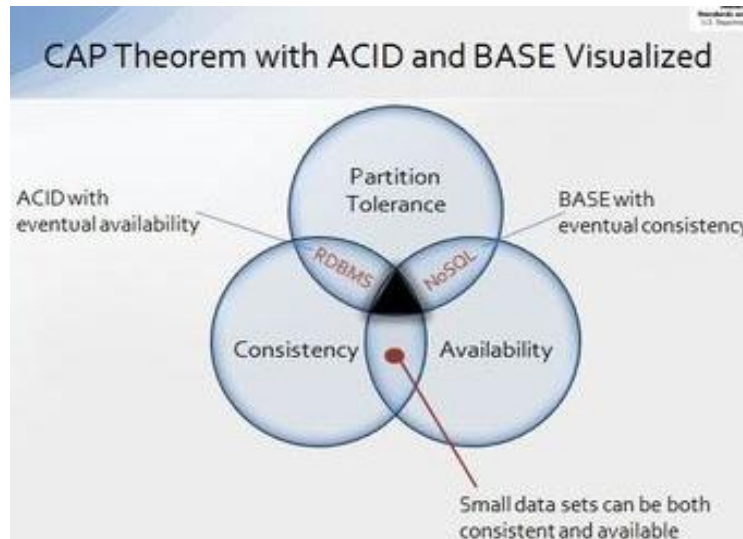


CAP stands for Consistency, Availability and Partitioning

**1. Consistency**: - The same data is present in all the machines and if it gets updated then changes should be reflected on all the machines.

**2. Availability**: - Data should be available at all the time and it should get stored permanently.

**3. Partition-tolerance**: - Even though a machine fails or any faults exist in it the database should work properly without any halts.

**TYPES OF NoSQL DATABASES**

1. **Key Value Database:**

Key value databases consists of two things one is key and another is vale. It is one of the traditional ways of data storage in NoSQL. In the key value pair a key is always unique and value consist of a single value or a group of values. It is based on the concept of hash table where each key is pointing to a value. This database can handle large amount of data. A key can point to different number of attributes

Features of Key Value Database

1. Supports unstructured data

2. Uses less memory to store the data

Example

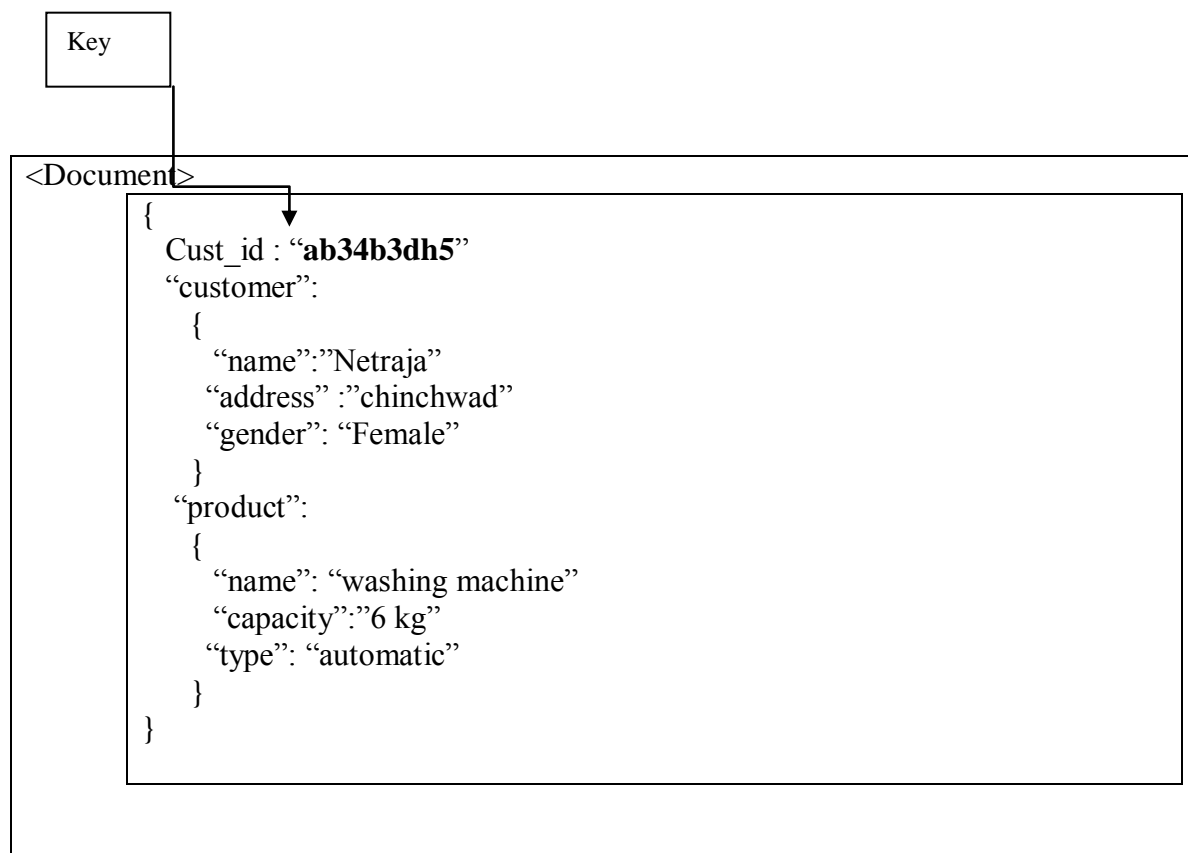| Student Database | |
|---|---|
| Key | Value |
| 1 | Rno:1<br>Name: Netraja<br>DOB:18/04/1984 |
| 2 | Rno:2<br>Name:Chetan<br>DOB:24/09/1983 |
| 3 | Rno:3<br>Name: Kimaya<br>DOB:14/04/2010 |
| 4 | Rno:4<br>Name: Nity<br>DOB:18/06/2016 |

## 2. Document Databases:

Similar to key value database as it also consists of keys which are unique. Each key is associated with a document i.e. document databases are used to store documents. The key can be a string, a URL or it can be a path.

Features of Document Databases

1. Supports structured, unstructured or semi structured data.

2. Examples or Couch DB and Mongo DB

Example

```
Key

<Document>
{
  Cust_id : "ab34b3dh5"
  "customer":
    {
       "name":"Netraja"
      "address" :"chinchwad"
      "gender": "Female"
     }
   "product":
    {
       "name": "washing machine"
      "capacity":"6 kg"
      "type": "automatic"
     }
}
```

### 3. Column Oriented Databases:

It is also called as column family store or wide column stores. A record can be a one column or in multiple columns. Columns can also be nested inside other columns. Columns can be grouped together in column families.

Examples- Couch DB,HBASE and Google's big table

Column Families
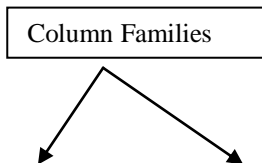
| Row key | Personal Data | | Professional Data | |
|---------|---------------|------|-------------------|--------|
| Eid | Name | City | Designation | Salary |
| 1 | Netraja | Pune | Manager | 50000 |
| 2 | Chetan | Mumbai | Engineer | 45000 |
| 3 | Kimaya | Nagpur | Software developer | 75000 |
| 4 | Nitya | Nanded | Designer | 90000 |

Features of column oriented databases
1. High performance for aggregate queries like count, sum, avg, max, min
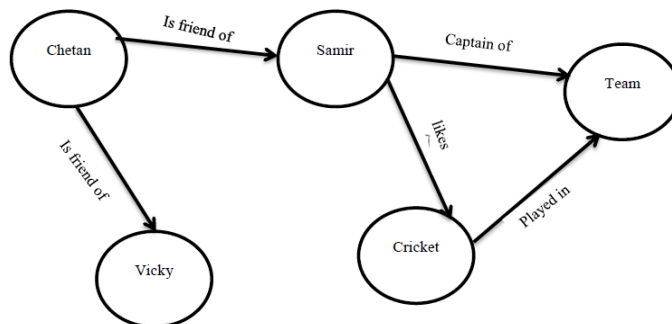2. Scalability is more.

### 4. Graph Databases:

In these type of databases data is stored in the form of graphs. Graph consists of nodes which represent objects and the edges which represents a relationship between these objects. It does not a predefined schema. It is faster than relational databases. This type of database is also scalable.

Examples-Neo4j
Features
1.Efficient storage for semi structured data
2.Graph databases are used in various applications like content management, social networking sites, bioinformatics, cloud management etc.

**SQL VS NoSQL Databases**

| 4 V's | SQL | NoSQL |
|---|---|---|
| Variety | The data is always structured. It cannot store variations of data. | The data is structured, semi structured and unstructured. It means data is present in various forms. |
| Volume | Relational databases stores limited data | These databases can able to store not only petabytes of data but also it can store zetta bytes or bronto bytes of data |
| Velocity | It refers to the speed at which data is generated and accessed. Relational database uses the concept of normalization on tables to avoid data redundancy and inconsistency but as this process involved several tables so data access speed is slow | Data is arranged in the form of collections and a single piece of data is represented in the form of entity hence reading and writing operations are much more faster and hence huge amount of data can be generated and accessed within a second |
| Veracity | SQL databases are responsible for storing precise and quality data. | NoSQL databases can store any kind of data. Data may be inconsistent incomplete or ambiguous |

**CONCLUSION**

The main aim of this research paper is to introduce the basic concepts of SQL and NoSQL databases. It also describes different types of NoSQL databases along with their features and examples. This paper also covers the basic concept of CAP theorem for distributed systems. It helps to identify the major differences of SQL and NoSQL with respect to 4 vs of Big Data. Finally I conclude that although the NoSQL databases are rich in scalability, size, speed and ability to store structured and unstructured data etc. Still it is the choice of the developer that depending upon the type of the application a developer has to choose either SQL or NoSQL database.

**REFERENCES**

1. Pramod J. Sadalage , Martin Fowler, ''*NoSQL Distilled*'',Addison-Wesley,Aug-2012
2. Supriya S. Pore,Swalaya B. Pawar. "Comparative Study of SQL and NoSQL Databases" *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, Volume 4, pp. 1747-1753
3. Yishan Li, Sathiamoorthy Manoharan. "A Performance Comparison of SQL and NoSQL databases"*Communications,computers and signal rocessing(PACRIM),*2013
4. Loginradius[Online]Available: https://www.loginradius.com/engineering/relational-database-management-system-rdbms-vs-nosql/
5. Epcc Education [Online] Available:https://static.epcc.ed.ac.uk/dissertations/hpc-msc/2012-2013/RDBMS%20vs%20NoSQL%20 %20Performance%20and%20Scaling%20Comparison.pdf
6. Mohamed A Mohamed, Obay G. Altrapi''Realtional VS NoSQL Databases: A Survey''*International Journal of Computer and Information Technology(IJCIT),*vol-3,issue-3,pp.598-601
7. The Repository at St. Cloud State[Online] Available:https://static.epcc.ed.ac.uk/dissertations/hpc-msc/2012-2013/RDBMS%20vs%20NoSQL%20-%20Performance%20and%20Scaling%20Comparison.pdf