MAT JOURNALS

Journal of Computer Science Engineering and Software Testing
Volume 4 Issue 3

Comparative Study of RDBMS, NOSQL and Graph Databases

[1]Vaibhav Sachdeva, [2]Sahil Sharma, [3]Prince Saini, [4]Dr. Mrinal Pandey
[1,2,3]UG students, [4]Assistant Professor
Department of Computer Science & Technology
Manav Rachna University
Faridabad, Haryana, India
Email: *[1]vaibhavsachdeva97@gmail.com, [2]sahilbulf.123@gmail.com, [3]princesaini2566@gmail.com, [4]mrinal@mru.edu.in*
DOI: http://doi.org/10.5281/zenodo.1468211

Abstract

The paper aims at analysis and comparison of various forms of databases particularly computer database Management System (RDBMS), Not solely SQL (NOSQL), Graph Databases. The Structured source language is employed by applications to access computer database systems containing informative during a semi declarative language whereas NOSQL databases area unit supported the key-value pairs. Graph info uses graph structures for resolution queries and to represent and store knowledge.

Keywords: *Structured Query Language, Not Only SQL, Graph Databases*

INTRODUCTION

Conventional info systems used relative kind model for the storage. , but in recent times non-relational databases have dramatically augmented in terms of quality, creating them fully completely different from the SQL databases, these databases area unit ordinarily referred to as NoSQL databases. Simple key-value pairs so the result may be keep and accessed in less time hence increasing the speed.

An info is an integrated assortment of information records, files, and different info objects [1]. Database will have a good range of users on AN on-demand basis. The world of info technologies has undergone extreme changes. The arrival of complicated knowledge has given rise to info technologies that are completely different from the standard relative databases.

Users have high demands concerning the info quantifiability, high reading and writing speed, economical access and mass knowledge storage. It's insufferable to resolve these demands with the assistance of computer database. Basic operations that area unit performed in Databases area unit CRUD (Create, Read, Update, Delete).

TYPES OF DATABASES

On a basic level, there are two core categories of databases

RELATIONAL DATABASE

These databases are majorly identified as SQL databases. These databases are a form of data structure that allows us to link data from different tables or types of data buckets. A data bucket must consist a key which helps to uniquely identify data inside the bucket. These databases have many limitations because of rapid increase in stored data. In this type of database data is stored in the form of rows and columns.

Before relational databases file based approaches were used which leads to data dependencies, duplicity and various other problems which were solved with the help of relational databases. Basically it has 3 layers of architecture. The following figure shows the layers with their corresponding

uses.

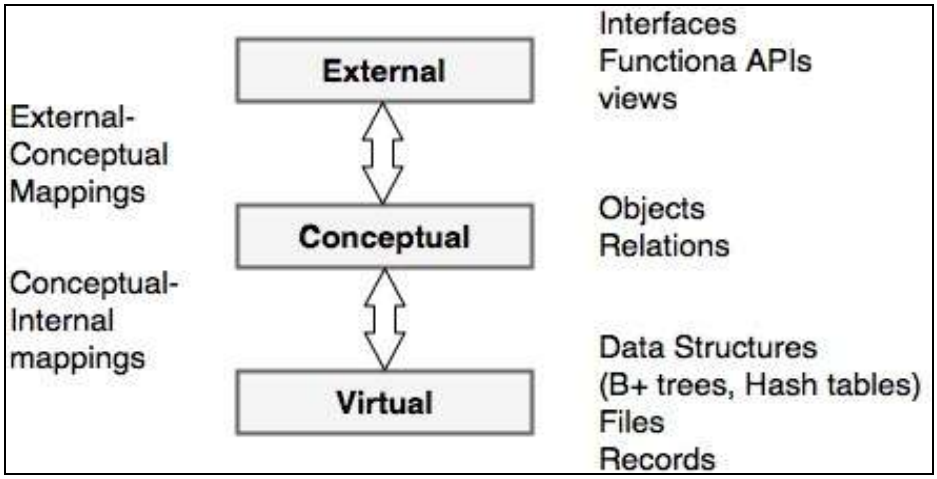

Fig: 1. *Relational Database*

Design of RDBMS

Relational databases use the notion of databases separated into tables where each row represent a record and each column represent a field. Normalization of the schema is an important aspect of relational databases. This involves following 3 steps described below:

First Normal Form (1NF): Removing teams of continuance information as in 1NF.Every attribute ought to contain atomic information. During this we have a tendency to produce a replacement table for every cluster of connected information that is acknowledged by primary key.

Second Normal Form (2NF): For multiple records having same set of values, move the records into a new table and link the tables with the foreign key.

Third Normal Form (3NF): Fields that is freelance of primary key of a table should be removed or if necessary should be placed into another table. It's conjointly necessary that 3NF satisfies 2NF and 2NF satisfies 1NF. An info schema is taken into account to be normalized if it satisfies the on top of mentioned three conditions.

Architecture of RDBMS

The internal layer consists of storage schema can be seen clearly from figure 1 which is responsible for encryption, decryption and compression of data. Conceptual layer consists of schema of the table. Conceptual layer defines the fields and columns attributes of the table. It hides the details of internal layer. External layer is the user view.

RDBMS is known for its ACID properties [2]. ACID properties comprises that the transactions are assured that the following characteristics are in place:

Atomicity: Either the complete dealing is completed or it doesn't happen in the least.

Consistency: The integrity of the info should be maintained by all transactions. Once the dealing takes place, the info should be in valid state.

Isolation: This property make sure that all the transactions ought to occur severally while not officious alternative transactions as there should be multiple transactions occurring at the same time at a similar time.

Durability: The changes created by a completed dealing are hold on in and written to a disk notwithstanding the

system failure happens or in alternative words sturdy.

NoSQL Database

NoSQL information stands for "Not solely SQL" and it's another of SQL that doesn't need any reasonably fastened tables that is needed in SQL. In spite of the actual fact that the word NoSQL databases indicated before in 1998 [6], its actual development began from year 2007.

NoSQL could be an information technology that is driven by the online, cloud computing and massive users. NoSQL area unit wide utilized in

Fig: 2. *NoSQL Database*

Qualities of NoSQL

The Main advantages of NoSQL are:

1) NoSQL reads and writes data very fast.
2) It supports storage of large quantity of data in a persisting manner
3) NoSQL is very convenient to scale.
4) It has reasonable cost.

NoSQL has some downsides also:

1. It lacks SQL support, which has high demand in industry
2. It lacks of reports, transactions and other added features
3. It is not good enough for majority of the NoSQL database items which were created in recent time.

Qualities of NoSQL databases that area unit explained higher than area unit typical ones, in real, every product befits the distinctive information model and CAP theorem. So, we'll bring NoSQL information model, and classify NoSQL in line with the CAP theorem.

CAP theorem and classification of NoSQL database

In the year 2000, academic Eric Brewer introduced the celebrated CAP theorem i.e. Consistency, Availability, Partition tolerance. CAP theorem's main plan is tells us that distributed system cannot meet the 3 district desires at the same time, however it will meet 2 of them.

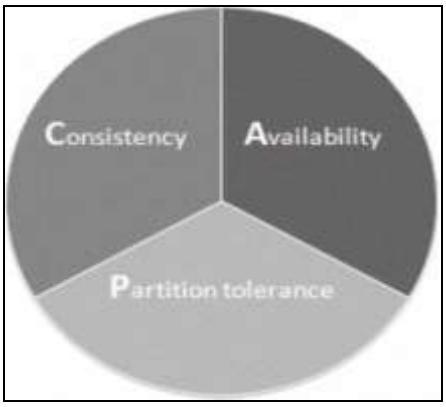

Fig: 3. *CAP theorem*

As per the CAP theorem and different issues of NoSQL databases, an initial classification of NOSQL databases are these:

Issues about availability and consistency

Some portion of database is not worried about the partition tolerance, and mostly uses Replication approach to guarantee data consistency and availability.

Systems concern the CA are: the relational database, Vertical (Column based), Aster Greenplum (Relational based), Data (Relational based), and many other.

Issues about partition tolerance (CP) and consistency

Such a database system keeps information in distributed nodes, but guarantees the consistency of data, but is not good enough for the availability. The famous CP system are: MemcacheDB (Key:value

based), BigTable (Column based), MongoDB (Document), Terrastore (Document based), Scalaris (Key: value based), Berkeley DB (Key: value), HBase (Column based).

Comparison of performance between NoSQL and SQL databases

Four crucial operations involved in comparison of databases:

```
abstract public class KeyValueStoreBase
{
abstract public bool Instantiate(string bucketName);abstract
public string Read(string key);
abstract public bool Write(string key, string value);abstract public
void Delete(string key);
}
abstract public List<string> GetAllKeys();
```

Fig: 4. *Comparison.*

1. Read: This is the specifically same because the scan of the CRUD (Create, Read, Update, Delete) model largely wont to tell information key information operations. This scans the worth admire the key provided from the key-value combine in storage.
2. Instantiate: This shows a storage bucket for key: value pairs storage
2. Write: This operation combines Create operation and Update operation of the CRUD model. If a given key: value pair is found in the storage, then this pair is put into the storage or it updates the value for the provided key in the storage.
3. Delete: This represents the Delete of CRUD model. It deletes the record key: value pair matching to a given key of the key: value pair storage.

Also, two additional operations are there which are frequently performed: iterating through all values and keys. To ensure testing these, we perform one more operation: *GetAllkeys* which grabs all the keys which are in the storage.

Major drawback of NoSQL is that it doesn't support SQL which is the current standard in most of the industries, reports and some additional features. In recent years most of the NoSQL products aren't compatible enough as compared to SQL. NOSQL databases are vey diverse which results in inadequate performance result.

Data model of NOSQL is classified in 4 groups (Key Value Stores, Document Stores, Column Family Stores, Graph Databases). NoSQL databases provide great mass storage & access and great read & write performance. NoSQL databases are more scalable and provide superior performance.

GRAPH DATABASE

In a graph database the main motive is to maintain relationship between objects which is performed by graph model which enables us to directly store objects together with relations among them, also there is a GDB which helps in accessing the query methods of both stored objects and graph structures.

(A)Challenges

1. Storing of graph properly
2. Explaining common things between graphs
3. Developing good index structure to boost pattern matching

Neo4j

Neo4j is a robust and high performance database which is backed by Neo Technology. It stores information in connected nodes [4]. Java language is the base of Neo4j and can be used on numerous systems [5].

(A) Qualities of Neo4j are

1. Supports complete ACID transaction with reliability
2. Largely expandable.
3. Easy to understand.
4. Quick and efficient retrieval of queries
5. Concentrates on what has to be retrieved not on the method of retrieval

Graph model

Graph A has a topology which can be expressed as A=(B,C),where B denotes the set of nodes and C denotes relations. Two vertices which need to exist are connected via edges. Moreover a graph can be undirected or directed, which represents whether it has direction or not. Graph contains a lot of information other than the topology

CONCLUSIONS

SQL is presently tough to exchange within the business. It had been terminated that NOSQL performs higher once measured in terms of potency, measurability, strength and reliableness. However NOSQL lacks correct relevancy in fields like transactions, reports and information product that were created in recent years. Also, it's been found that graph databases are economical however are tough to manage.

REFERENCES

1. E. F. Codd. Data models in database management. SIGMOD Rec., 11(2):112-114, June 1980.
2. Marko A. Rodriguez, July 2007, ―Problem-Solving using Graph Traversals‖, AT&Ti Technical Talk - Glendale, California.
3. Alexandra Martinez1, Rodrigo Mora2, Daniel Alvarado3, Gustavo L_opez3, and Steve Quir_os2
4. Azhi Faraj, Bilal Rashid "Comparative Study Of Relational And Non-Relations Database Performances Using Oracle And Mongodb"
5. Ali Salehnia "Comparisons of Relational Databases With Big Data: A Teaching Approach" South Dakota State University.
6. ZHANG Dong, "The Analysis of te et Operators Special Cluster- No SQL," Network World July. 2011, pp. 1-2.

Cite this article as: Vaibhav Sachdeva, Sahil Sharma, Prince Saini, & Dr. Mrinal Pandey. (2018). Comparative Study of RDBMS, NOSQL and Graph Databases. Journal of Computer Science Engineering and Software Testing, 4(3), 1–5. http://doi.org/10.5281/zenodo.1468211