

RDBMS Current Challenges and Oppurtunities with NoSQL to NewSQL

Jeyanthi Sivakumaran^a, Syed Zakir Ali^a,

Relational Database Management Systems (RDBMS) are designed to manage well-structured data, requiring users to design a schema before storing and querying data with the help of SQL (Structured Query Language). The evolution of database management system undergoes many developments for more than forty years to meet the challenges by modern applications. In this fast Growing science and technological period, where advancements in web technology, mobile devices and sensors (IoT) need to accommodate huge amount of structured and unstructured data, called Big Data. The demand for technologies with large processing and storage requirements are rising enormously. Customized RDBMS has to confront challenges against Big Data's scalability and performances. To meet these requirements such as volume, variety, velocity and variability, enterprises are adopting alternative technologies like NoSQL and NewSQL. Currently, NewSQL scalability is classified as modern RDBMS which carries NoSQL for OLTP (Online Transaction Processing) to read as well write data, whereas satisfying the Atomicity, Consistency, Isolation, and Durability (ACID) properties of a traditional DB system. Further my research paper introduces NoSQL and its key attributes and NewSQL (NuoDB, VoltDB).

Keywords: Databases; SQL; NoSQL;NewSQL

Introduction

NewSQL is a new class of database management system (DBMSs). The year 2000 onwards, Internet applications brought more challenges to resources than applications. By scaling their DBMS vertically, many tried to move their database to a machine with better hardware. But, moving the database across machines is a complicated process and often requires enough downtime. NoSQL systems are not used by many application because of its not supporting required transactional and consistency requirements. Therefore, the other option is to go for either purchase a more sophisticated machine which can vertically scale the DBMS, or to build their own sharding middleware to support transactions. But both of these methods cost more, so the NewSQL systems emerged. Protocols are assigned for large clusters running NoSQL/NewSQL databases for performances. I view NewSQL is good for data storage, security, big data support etc. Many unsolved issues in SQL and NoSQL are solved in NewSQL.

NOSQL

NoSQL stands for "not only SQL" [1]. Broadly speaking, it includes all non-relational DBMS. RDBMS transactions conforming to ACID principle, NoSQL DBMS follow the CAP theorem and thus its transactions conform to the BASE properties (Basically, Available, Soft State, and Eventually Consistent) [1]. NoSQL systems are accommodating, non-relational databases designed for storing huge volume of data (Big Data), taking into consideration its variety, data velocity and parallel data processing across a large number of NoSQL database servers. NoSQL DBMSs are not completely supporting for applications already written for RDBMS. [2]. By and large, scalability requirements NoSQL is schema-less, avoid join operations and horizontally scalable.

CAP Theorem

Brewer's CAP theorem consist Consistency, *Availability* and *Partition Tolerance*, which are the basic requirements for NoSQL

i) Consistency

A service that could manipulates successfully or not at all.

ii) Availability

The service is available at any time.

iii) Partition Tolerance

As a result of your database and application operates in the same system irrespective of scaling, your server is fully responsible for the system to run or crash. When there is no-partition and a well-distributed system a CAP maintain Consistency and Availability.

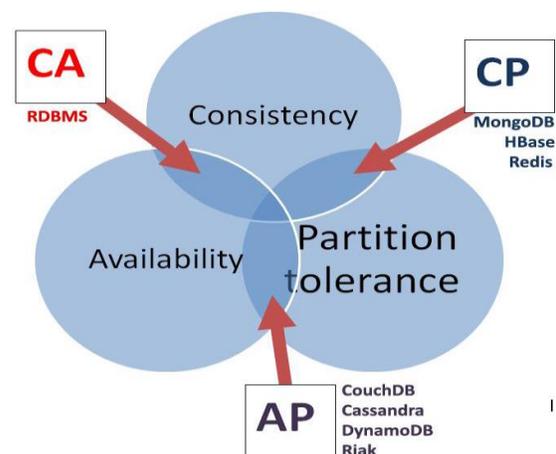


Figure 1: CAP Theorem supporting NoSQL Databases [7]

a. Department of Electronics and communication Engineering, Middle East college, Muscat, OMAN
Correspondence: zakir@mec.edu.om

Auto-Sharding

NoSQL automatically spreads data without asking for any applications across servers. Add or remove servers from data layers is feasible. Data replication and cluster data storage are supported by NoSQL databases to ensure high availability. Also it can support disaster recovery. So, there will be no work disruption even if any breakdowns or failures occurred in any one of the machine.

Distributed Query Support

RDBMS can eliminate or control the complex data queries. Even though distributed across many servers, NoSQL database systems keeping their full query power.

What's better in NoSQL

- NoSQL provides horizontal scalability better than vertical
- NoSQL support hardware getting cheaper and processing power increasing
- NoSQL support less operational complexity as against RDBMS solutions
- NoSQL provides, in most of the solutions you get automatic sharding etc. as default

New SQL

NewSQL is upcoming current scalable RDBMS system with powerful applications. Keeping the ACID properties of customized DB system, OLTPs gives high scalability for NoSQL systems workload [3], [4], [5],[2]. NewSQL databases have the similar performance as NoSQL systems; additionally provide administrators with ACID performance guarantees. NewSQL was designed to maintain SQL while addressing existing issues with traditional OLTP systems, mainly their scalability and performance to some extent.

NewSQL differs from traditional RDBMS performances, by accommodating NoSQL features namely, Column-Oriented Data Storage and Distributed Architectures, In-Memory Processing, Massively Parallel Processing (MPP), Symmetric Multiprocessing (SMP). NewSQL accepts the Big Data challenges and handle the volume, variety, velocity and variability [2].

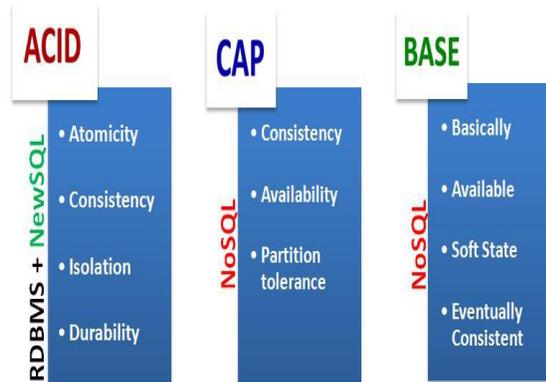


Figure 2: properties of ACID, CAP, and BASE [7]

RDBMS is stored the SQL data. Graphs and trees are stored in NoSQL and NewSQL stored all data. SQL data flows in one machine whereas data is distributed in NoSQL and in NewSQL gives the advantages of both.

Architecture for NewSQL

Newly designed NewSQL systems are to satisfy performance and scalability. Better performance is obtained if we use in-memory and flash disks as the primary data store. Solutions can be software only (Google spanner, VoltDB, NuoDB). These new solutions can support scalability requirements. NewSQL style supports VoltDB for easy execution of transactions which are much faster than traditional DBMS system. VoltDB has the capability of scale across many servers, as well as to handle millions of transactions per second. NuoDB is defined as a SQL as well as ACID based distributed DBMS. NuoDB is a peer-to-peer, asynchronous, distributed system which differs from traditional shared-nothing or shared-disk architectures.

NewSQL support to ACID properties

Atomicity is guranteed in each and every situation, including power failures, errors, and crashes. The property of consistency ensures that any transaction will change the database from one state to another. The property of isolation is to ensure that in a system transactions are executed concurrently. Concurrency control provides isolation. Once a transaction is validated durability is maintained. It will remain in the system, even in the event of any power failures, system crashes, or errors. In a relational database, for instance, once a group of SQL statements execute, the results need to be stored permanently. To defend against power loss, transactions must be recorded in a non-volatile memory. NewSQL actually satisfies the CAP theorem, but the problem of latency occurs. If the data range is small Traditional SQL can handle and the performance is quite satisfied but for big data system confronting problem. In NoSQL too volume of data increasing, for Penta bytes of data the consistency decreases the systems performance. But in NewSQL Consistency of Data is far better than both these SQL. When data volume increases NewSQL manages data perfectly. ACID properties are not supported by NoSQL therefore, security constraint prevails. The other issues are Authentication, Authorization and Confidentiality, but NewSQL cover all these bottlenecks. It's not necessary for organizations to maintain the DBMS in their own hardware, as cloud computing providers offer NewSQL Database-as-a-Service (DBaaS) product. In 2016 we find Aurora and ClearDB products are available in this NewSQL category. Splitting a database into partitions or shards is performed by NewSQL DBMSs scale-out. So, the DBMS can move data between physical resources to increase or decrease the capacity of DBMS without any disruption to the service. But, in this process NewSQL has to maintain the ACID property which is not that easy [6]. Concurrency control scheme plays the major role in implementing transaction details while processing the DBMS as it brings many changes to the system. NewSQL systems are using the protocol of decentralized multi-version concurrency control (MVCC). Replication of database is the best way for an organization to ensure high availability and data durability for their OLTP application. NewSQL systems also support some kind of replication mechanism. Customers not aware of their data are being captured by DBaaS and have been replicated. This is an added advantage to the DBaaS. NewSQL DBMS has the capability for providing fault tolerance by its crash recovery

mechanism. Newer SQL DBMSs try to maintain fault tolerance, unlike traditional DBMS to ensure that no updates are lost and also minimize the downtime [6].

Conclusion

As the data are becoming larger and larger RDBMS is finding difficult to process and organize them in arrays with a good speed. Besides, the data flow will be structured, semi-structured or unstructured but RDBMS has the ability to process only the structured data. Storage will be another hindrance for big data. SQL is not good enough and not support to modern database query. In-built optimizations of RDBMS, and its database design, enhanced performances of RDBMSs to perform faster in many applications and data sets. Improvised technology, increasing processor speeds and fixing suitable memory and storage facilities allow system administrators to build better systems that can improve database performance. In future built-in maintenance utilities help database administrators and developers to easily maintain, the database in the system, test, repair and backup. But as an alternative NoSQL and NewSQL play a big role in schemas and data storage. This enables the developers and users to choose appropriate storages and finding solutions for the current challenges and opportunities. NewSQL engines developing and employing a variety of architectures namely, Storage type, Data storage, Query method, Replication, Concurrency control etc. to assist users in choosing the best storage solution for their needs.

We can conclude that all of these new systems will support some form of relational model and SQL. Consider from the error-free performance point of view, security lapse prevails in NoSQL and NewSQL DB's than in conventional SQL databases. Issues in database security need to be focused and further research needed to safeguard the situation. NewSQL database needs to be evaluating scalability and load testing. Now we can assume that there will be a relationship with RDBMS, NoSQL and NewSQL. We can conclude that this relationship will enhance further developments and can emerge with a new form of relational database system in future.

References

- Cattell, R., (2011). Scalable SQL and NoSQL data stores. *ACM SIGMOD Record*, 39(4),12-14.
- Moniruzzaman.A,(2013). NoSQL Database: New Era of Databases for Big data Analytics - Classification, Characteristics and Comparison, *ACM SIGMOD Record*, 6(4), 608-616.
- Aslett.M, (2011). How will the database incumbents respond to NoSQL and NewSQL,. *149* (4), 923-935.
- Stonebraker, Michael (2011). NewSQL: An Alternative to NoSQL and Old SQL for New OLTP Apps. *Communications of the ACM Blog*. 12 (7), 701-708.
- Hoff, Todd (2012). "Google Spanner's Most Surprising Revelation: NoSQL is Out and NewSQL is In". *Nature*, 437 (7055), 69-87.
- Pavlo. A and Aslett. M, (2016) What's Really New with NewSQL?," *ACM SIGMOD Record*, 45(2), 45-55.

Moniruzzaman.A, (2014). NewSQL: Towards Next-Generation Scalable RDBMS for Online Transaction Processing (OLTP) for Big Data Management, *International Journal of Database Theory and Application*, 7(6), 121-130.