NoSQL Technologies

Kandalepas Konstantinos

k.kandalepas1@unimail.derby.ac.uk
University of Derby

Abstract. . As the use of Web applications and the amount of big data collected increasing, the need of NoSQL database systems becomes increasingly vital. The NoSQL DBMS, Not Only Sql, are distributed, non-relational databases that support flexible data schemas. These systems are more flexible and simpler than the traditional RDBMS, relational databases. The characteristics of NoSQL technologies are the availability, performance, and partition-tolerance. (4) On the other hand, these technologies sacrifice the consistency, atomicity, isolation, and durability in order, to achieve better soft-scale and basically-available requirements. There are two types of NoSQL systems. The aggregate-oriented and the non-aggregate oriented. The first type includes the document, the wide-column, and the key-value databases. The second one includes the graph, which have an unstructured schema and pay attention more on the connection of data, rather than the tables. (1) As the market grows, each organization must select the best blend of technologies, to stand by its specific needs. There are opportunities to improve future research in this topic such as the security management capabilities, comparative analysis of SQL and NoSQL systems and adoption of real world case studies. The research target of this paper is to identify the NoSQL systems characteristics, advantages and disadvantages comparing with the relational databases and the future research opportunities.

Keywords: NoSQL databases, NoSQL technologies, Big Data.

Introduction

NoSQL technologies are widely known as Not Only SQL. This means that NoSQL systems use relational databases, but are more efficient concerning the performance, the scalability, and the schema flexibility. In today's real world applications, the databases that a system use are often more than one or two and they are distributed. In this case, relational database management systems cannot cope with scalability and flexibility challenges and with the needs of the modern interactive systems. Furthermore, are schema-less than the traditional relational databases. NoSQL technologies are now very popular for large and unstructured data storage. The use of NoSQL databases is a vital alternative, when the applications do not work efficiently with a large quantity of data. This paper provides a literature review in NoSQL technologies, NoSQL characteristics, different types of NoSQL systems, the advantages, and disadvantages of their use. In the end, future research opportunities are discussed.

Literature review

NoSQL Categories

NoSQL systems are divided into two categories. The aggregate-oriented and the non agrreagate oriented. (1) The different types of NoSQL technologies are discussed following.

1. *Key-vlaue*. These types of databases store the data as keys in simple standalone tables. (4) Key-value DBMS are schema-less and they are very efficient when the application has as the first priority the write performance.(7)
2. *Document databases*. In these DBMS data is stored as documents and they are efficient in handling with XML documents, text documents and email messages. (4) These databases are schema-less, which is extremely vital for the performance and the flexibolity of the database.
3. *Columnar of wide-column databases*. These databases have some attributes of both relational and NoSQL databases. They have the row and comumn structure of relational databases, but without the rigorous rules that relational databases have. (1) These databases are very efficient with large ammount of data and they have applications for predictive analytics. (3)
4. *Graph databases*. These DBMS are very userfull when the data are highly connected. They can handle efficient the relationships between the data and the are very important when the data structure is like a graph. The nodes represents the entities of the database and the edges the relationships between them. (7)

NoSQL DBMS vs. Relational DBMS

The handling of unstructured data is one of the most significant advantages of NoSQL technologies. Such data can be for examples videos, emails, audio, and documents. But one of the most important issues with NoSQL systems is the encryption of the data and the authentication between the client and the server. On the other hand, relational databases have a long time of productivity and management. Furthermore, it easy for a company to find programmers with good knowledge and experience in relational database management systems. Apart from that, many of the NoSQL solutions are open source which means that they do not provide support for customers or administration and management tools, something that is extremely vital for all organizations. NoSQL systems were found that they have better flexibility than the relational databases, but tests found that relational databases have better performance than NoSQL technologies. In the end, NoSQL and relational database management systems are meant to solve different problems. So, the question for each organization is what problem wants to solve and to decide which of these two technologies fit better in its requirements. (6)

NoSQL Characteristics

Schema less. Relational databases have strict schema with a lot of constraints such the NOT NULL and UNIQUE constraints in their attributes. NoSQL systems on the other hand have schema less, which is a great attribute, concerning the new needs of big data collection.

Shared nothing architecture. In NoSQL technologies, the allowing storage is accessed at local disk speeds rather than network speeds. This is due to that every server can use only own local storage and the outcome of this attribute is the increase of capacity when more nodes added.

Elasticity. Elasticity is the dynamic expanded ability of a database. This attribute gives the advantage of data replication in new nodes, at the time that a new node is added. (1)

Sharding. Only a single server can manage the record, because they are partitioned into finders, which make them small enough. (1)

Asynchronous replication. In NoSQL technologies the writes are independent of the network traffic. This attrbute give the system the advantage to complete the writes faster. On the other hand, these data are not replicated right away, which means that they can be lost in certain windows. (1)

BASE.

Basically available. A database management system shoud ensure that when the system fails, the database will be accessible in some ways. The Basically Available property of NoSQL technology ensures that requirement, because the data stored and are replicated in many servers. This means that if a server fails, another available server will ensure the availability of data and the accessibility in the database. (3)

Soft state. When a transaction take place the updates to all data stores can be a time consuming process. This property of NoSQL technologies allows the transactions to procced even if a failure or troubles in data stores take place for a specific time. (1)

Eventually consistent. The consistency of the data is a vital issue for database management systems. Relational DBMS ensures the consistency of the data. On the other hand, NoSQL systems ensures that in the end the data will be consistent and not in the end of every operation. (1)

Future research opportunities

NoSQL technologies are new in the area of database management systems. There are a lot of reserach opportunities in this field. A very significant future research is to investigate the new opportunities that come arise of these database management systems and some security matters, because at this time NoSQL systems have security gaps and issues in contrast with relational database management systems. Apart from that, a great opportunity for future research is to establish some best practices concerning NoSQL data and models. There are some comparative analyses between NoSQL systems and relational databases in the literature, but it is vital for more information about the differences between these two technologies. The most important issue for future research is the adoption of these systems in real world case studies and the use of them in applications with different requirements, needs and data handling.

Conclusion

The motivation of NoSQL systems development was the intensive development of the Web, which created applications with huge workloads and enormous amounts of data. In the literature review, what was identified were the types and the characteristics of NoSQL technologies. The main reason was to better understand the differences between these technologies and the relational databases management systems. In some cases, the NoSQL systems are more efficient, but on the other hand relational databases fit better in some problems with an extremely need of security issues. So, every organization needs at first to understand what problem wants to solve. If they identify the problem, then it can choose which database management system fits better in their needs. For example, NoSQL systems do not fit well for applications that require transactional semantics. In the end the choice of a specific technology depends on the features of the task, the expected workload, read/write ratio, structure of records to be stored, the types of the queries to be executed and the desired level of data consistency. (NoSQL database management systems)

References

1. Deka Ganesh Chandra, BASE analysis of NoSQL database, Future Generation Computer Systems 52, p.13-21, (2015)
2. S. D. Kuznetsov and A. V. Poskonin NoSQL Data Management Systems, Progamming and Computer Software, Vol. 40, No. 6, pp. 323–332, doi: 10.1134/S0361768814060152, , (2014)
3. Madison, Michael; Barnhill, Mark; Napier, Cassie; and Godin, Joy, NoSQL Database Technologies, Journal of International Technology and Information Management: Vol. 24: Iss. 1, Article 1. (2015)
4. A B M Moniruzzaman and Syed Akhter Hossain, NoSQL Database: New Era of Databases for Big data Analytics, Classification, Characteristics and Comparison, International Journal of Database Theory and Application, Vol. 6, No.4, (2013)
5. Suman Kashyap, Shruti Zamwar, Tanvi Bhavsa3, Snigdha Singh, Benchmarking and Analysis of NoSQL Technologies, International Journal of Emerging Technology and Advanced Engineering, Volume 3, Issue 9, (2013)
6. Nance, Cory; Losser, Travis; Iype, Reenu; and Harmon, Gary, "NOSQL VS RDBMS - WHY THERE IS ROOM FOR BOTH", SAIS 2013 Proceedings. Paper 27, (2013)
7. Innocent Mapanga, Prudence Kadebu, Database Management Systems: A NoSQL Analysis, International Journal of Modern Communication Technologies & Research (IJMCTR), ISSN: 2321-0850, Volume-1, Issue-7, (2013)