

Collaborative Discussion 1: Alternatives to SQL

Initial Post

Many companies now apply NoSQL to process their real-time data or requests. Because NoSQL can resolve some issues that the traditional relational database(RDBMS) is not good at. Compared with RDBMS, NoSQL can bring some edges as below:

Flexible data structures. NoSQL is schemaless, and each data can be completely different. The feature is very suitable for Big Data(BD). When people talk about BD, the data type, size, or content could vary. However, in RDBMS, you have to predefined all tables' struct, and the data type should be the same for a column. If you want to change them in the future, it could not be an easy job, and even a slight mistake could cause data to be lost.

Easy scale horizontal. You can easily add a NoSQL database without impacting existing data because there has no relationship between data. For example, data in a key-value database has a more plain design than in a relational database, and every item stands on its own. In contrast, if you want to achieve the same effect in the RDBMS, the cost could be too high to accept (Strauch, C., Sites, U.L.S. and Kriha, W., 2011). Scalability can be considered most crucial in a business (Strauch, C., Sites, U.L.S. and Kriha, W., 2011).

High performance. NoSQL gains better performance from two parts. One is no need to limit by the ACID (atomicity, consistency, isolation, durability) principle of RDBMS, but not all the data in the business have to follow this rule. Whereas if you use RDBMS, there has no choice. Another is NoSQL can benefit from their simpler data module to high reading/write with low latency (Leavitt, N., 2010). NoSQL usually does not support some intensive operations in the RDBMS, such as joining multiple tables.

reference:

1. Strauch, C., Sites, U.L.S. and Kriha, W., 2011. NoSQL databases. Lecture Notes, Stuttgart Media University, 20, p.24.
2. Leavitt, N., 2010. Will NoSQL databases live up to their promise?. Computer, 43(2), pp.12-14.

summary post

Twenty years ago, the traditional relational database management system (RDBMS) was a mainstream way to store and process data. When people design a database with RDBMS, they often use a static approach to develop a database, such as confirming data type and evaluating possible data size in the future. On the other hand, it could be challenging to change the data struct when the system was going online.

However, as the rapid growth of data volume and variety of sources were involved in information systems, people can't relay previously ways to handle data booming. 2010 it had two zettabytes of data to be used, and 2025 could reach 181 zettabytes (Statista, 2021). People are very hard to follow RDBSM to design a database to meet internet business. In this context, NoSQL was brought up.

NoSQL includes a variety of styles, like key-value, document, graph or object-oriented; each format tries to resolve some issues that RDBMS was hard to fix. The issues resolved by NoSQL as scalability, low cost, support volume and variety of data, availability, and lower complexity (Kunda, D. and Phiri, H., 2017). However, RDBMS still has some advantages over NoSQL regarding data consistency, security and standard industry level query language (Nance, C., Losser, T., Iype, R. and Harmon, G., 2013). Especially, many RDBMS have implemented AIDC (atomicity, consistency, isolation, durability) properties and got restricted scrutiny for a long time. This characteristic is a high demand in some sections, like the banking system. Whereas NoSQL lacks confidence in AIDC and immigrating an online system from one data model to a new one could be a high-risk task.

We can't conclude that NoSQL will replace relational databases soon or that RDBMS is an outdated model. People should consider their needs or how they plan to use a database to design and select an appropriate model. Many companies could use relational databases and NoSQL simultaneously but place them in different areas in one system. "relational database and NoSQL were meant to solve different problem." (Nance, C., Losser, T., Iype, R. and Harmon, G., 2013).

Reference:

Statista (2021). Data Created Worldwide 2010-2025 | Statista. [online] Statista. Available at: <https://www.statista.com/statistics/871513/worldwide-data-created/>.

Kunda, D. and Phiri, H., 2017. A comparative study of NoSQL and relational database. *Zambia ICT Journal*, 1(1), pp.1-4.

Han, J., Haihong, E., Le, G. and Du, J., 2011, October. Survey on NoSQL database. In 2011 6th international conference on pervasive computing and applications (pp. 363-366). IEEE.

Nance, C., Losser, T., Iype, R. and Harmon, G., 2013. NoSQL vs rdbms-why there is room for both.

Discuss With Mohammad Atieh

Hi, Mohammad. Thanks for your sharing.

Many software systems were designed in object-oriented programming (OOP) thinking way. However, when the system relays a traditional relational database management system (RDBMS), the whole system can't be in a unified model. In addition, the system's code could be easily impacted by data structure changes in RDBMS.

In traditional RDBMS, it is tough to represent a complex data type, and when a new data type is coming, there could be so much trivial working need to be done. For example, insert a new column into the existing table and ensure that the current data can be compatible with the new column. But applying the OOP concept to creating a new data type could be an easy job; we can use inherited or composition ways to construct new data and no side effects on the old datasets.

If the object-oriented database was following fundamental principles of object-oriented programming, it could be a good idea that can be more naturally fix issues that RDBMS can't.

Discuss With Rob Mennell

Thanks for your sharing. Key-value, document and graph are the three majority formats in NoSQL, and graph database takes a relationship first approach to storing and querying data. When people put data to a relation database, if it has many tables and massive foreign keys scattered in different places, you can imagine how complicated a query SQL with join could be. But if data can be represented as nodes, such as LinkedIn, to show networking between people. Apply graph database users could find more hidden patterns in the data. For performance, there is research to show graph databases are better than relational databases (Vicknair, C., Macias, M., Zhao, Z., Nan, X., Chen, Y. and Wilkins, D., 2010).

However, so far graph database lacks a common standard. If you want to transfer data to another graph database, it could not be as easy as a traditional relational database.

Reference:

Vicknair, C., Macias, M., Zhao, Z., Nan, X., Chen, Y. and Wilkins, D., 2010, April. A comparison of a graph database and a relational database: a data provenance perspective. In *Proceedings of the 48th annual Southeast regional conference* (pp. 1-6).