Unit 8–11
Collaborative discussion on alternatives to NoSQL

DISCUSSION INSTRUCTIONS **2**

Discussion topic 2
Instructions 2

MY INITIAL POST ON NON-SCALABILITY OF SQL **3**

Response 1 to my post 4
Second response to my post 5

INITIAL POST BY ANOTHER STUDENT ON MANGODB **6**

My response to another student 7

INITIAL POST BY ANOTHER STUDENT ON SCHEMELESS OF NOSQL **8**

My response to another student 9

SUMMARY POST ON ALTERNATIVES TO NOSQL **10**

Unit 8–11
Collaborative discussion on alternatives to NoSQL

Discussion instructions

Discussion topic

Thanks to the advent of Big Data and our increasing demand for more data, there is a recent trend towards newer NoSQL technologies. As part of this week's activity, identify a NoSQL technology to research. Then post in the discussion forum detailing the key approaches of your chosen technology and how its relative merits when compared with a traditional relational database.

Instructions

Some types of NoSQL you might want to consider include:

- Graph databases
- Key/value stores
- Document stores
- Object databases

You should demonstrate that you understand the topic covered and ensure you use references to academic literature (journals, books, reports, etc.).

Unit 8–11
Collaborative discussion on alternatives to NoSQL

My initial post on non-scalability of SQL

Relational databases, such as SQL, have become non-scalable since the growth of data and somewhat obsolete (Mukhtar, Ercan 2020). The exponential collection of data, aka 'Big Data' increased demand for high performance has led to the implementation of NoSQL databases which permits higher scaling – horizontally, thus faster availability when used with for example cloud computing (Mukhtar, Ercan 2020).

NoSQL databases are however not as flexible as a relational one, as each NoSQL database has generally been developed to function on a specific type of data in a limited way on a large scale (Mukhtar, Ercan 2020). A lot of data collected are unorganised, unmanaged, and unstructured collected from sensors, transactions, and more, which is required to improve the output and catalyse technological advancements further, such as AI, IoT, and cloud computing (Kousiouris et al, 2016). NoSQL undouble satisfies the arisen requirements; however, NoSQL databases were initially not designed with security or authentication mechanisms (Sicari et al, 2022). The main security threats intrinsic to NoSQL are non-encryption of data, unauthorised exposure of data and its backup, and insecure communication over networks (Tewari, Gupta 2018).

The key-value data model uses a hash table, with keys stored as indexes toward data, which is suitable to use to efficiently find values in large datasets, whilst offering fast, secure, and inexpensive access to information with high availability. (Sicari et al, 2022). As the key-value data model is schemeless, its weakness may be the inability to create customised views of data (Sicari et al, 2022).

References

Tewari, A. and Gupta, B.B. (2018). Security, privacy and trust of different layers in Internet-of-Things (IoTs) framework. *Future Generation Computer Systems* 108, 909–920 DOI: https://doi.org/10.1016/j.future.2018.04.027
Sicari, S., Rizzardi, A. and Coen-Porisini, A. (2022). Security & privacy issues and challenges in NoSQL databases. *Computer Networks* 206,108828. DOI: https://doi.org/10.1016/j.comnet.2022.108828
Seriatos, G., Kousiouris, G., Menychtas, A., Kyriazis, D., Varvarigou, T. (2016). Comparison of Database and Workload Types Performance in Cloud Environments. In: Karydis, I., Sioutas, S., Triantafillou, P., Tsoumakos, D. (eds) Algorithmic Aspects of Cloud Computing. ALGOCLOUD 2015. Lecture Notes in Computer Science (), vol 9511. Springer, Cham. https://doi.org/10.1007/978-3-319-29919-8_11
Evaluating Riak Key Value Cluster for Big Data. (2020). *Tehnicki vjesnik - Technical Gazette*, 27(1), 157–165. DOI: https://doi.org/10.17559/TV-201809161205584

Unit 8–11
Collaborative discussion on alternatives to NoSQL

Response 1 to my post

Maja, thank you for discussing NoSQL databases. This family of database technologies has been fascinating to me for years!

The exponential growth of data certainly necessitated a more efficient solution to manage such massive volumes of data, and in that regard, I certainly agree with your statement that relational databases have become obsolete (Tagt, 2022). Through my research, I have found that relational databases do retain some advantages over NoSQL in certain situations (Bitnine.net, 2016; Pandora FMS Team, 2016). SQL databases are largely standardized, using a single uniform language for development and operations (Bitnine.net, 2016). Additionally, data must adhere to integrity standards pre-defined by SQL's schema (Pandora FMS Team, 2016). Thus, when working with datasets with known data points and a defined number of records, SQL databases may still be the better choice in terms of cost, compatibility, and ease of use (Bitnine.net, 2016; Pandora FMS Team, 2016).

Due to the explosion of data volume the world has experienced in the past several years, NoSQL technologies have continued to become more prominent to manage an increasingly unmanageable amount of data (Krishnan, 2016). Despite the weaknesses associated with some of these databases technologies, I believe that continued implementation and innovation will continue to improve the standardization, security, and ease of use in the near future.

Thank you for your thought-provoking post!
Rob

References

Bitnine.net. (2016). SQL vs. NoSQL: Comparative Advantages and Disadvantages. [online] Available at: https://bitnine.net/blog-computing/sql-vs-nosql-comparative-advantages-and-disadvantages/

Krishnan H., Elaydom M., & Santhanakrishnan, T. (2016). MongoDB - a comparison with NoSQL databases. International Journal of Scientific & Engineering Research, [online] 7(5). Available at: https://www.ijser.org/researchpaper/MongoDB-a-comparison-with-NoSQL-databases.pdf [Accessed 4 May 2022].

Pandora FMS Team (2016). NOSQL vs SQL. Key differences and when to choose each. [online] Pandora FMS - The Monitoring Blog. Available at: https://pandorafms.com/blog/nosql-vs-sql-key-differences/.

Tagt, M. (2022). Collaborative Discussion 2: Alternatives to SQL, Initial Post. University of Essex Online. [online]. Available at: https://www.my-course.co.uk/mod/hsuforum/discuss.php?d=306144 [Accessed 4 May 2022].

Second response to my post

Hi Maja and Rob – I was very impressed by your posts and the detailed analysis you did to concluded that relational databases are obsolete. I certainly agree that considering the vast data sets and the increasingly complex queries on those data that we currently implementing in many scenarios, this seems to be the case.

However, I work for a university and decided to discuss this topic with our CTO. He told me that our new campus management system (interestingly an American system and the first to be implemented in Europe, so we can assume the latest technology standards) which handles all our student and study program data is built on SQL. He explained to me that it is often the case that people assume that SQL is becoming obsolete, but that we are still firmly dependent on SQL. Non-relational databases make queries and data much more lightweight, but they are far from being obsolete in his opinion. I was surprised as I had imagined that the new system is much more ‘advanced’ in terms of the database technology it runs on, especially considering how the university intends to use student data in future.

To support this view, I consulted an article which raises limitations such as lack of possibility to connect databases to certain applications and third-party ecosystems causing companies to have to develop their own visualization tools (Kulkarni, 2019). Certainly, the benefit of SQL seems to be that it can function on and within many systems, questioning how lightweight NoSQL systems really are in terms of implementation across the many functions of a large organisations.

I plan to do some more research on this view to be implemented in my e-portfolio, so if you’re interested, do follow! I will be interested to hear what both of you find further on.

References

Kulkarni, A. (2019) Why SQL is beating NoSQL and what this means for the future of data. Timescale. [available at: https://www.timescale.com/blog/why-sql-beating-nosql-what-this-means-for-future-of-data-time-series-database-348b777b847a/] Accessed: 12.05.2022

Unit 8–11
Collaborative discussion on alternatives to NoSQL

Initial post by another student on MangoDB

Dealing with increasing amounts of data NoSQL databases are rising up to challenge the dominate relational databases. Most rational database systems fail to reach the scalability and flexibility required for big data. One such alternative are document-oriented databases such as MangoDB which works on the principle of storing data in 'documents' rather than in strictly defined tables. Data from the databases can be aggregated to get them into a searchable, organised form. In MangoDB documents are value/key pairs that serve as the basic data unit, each database contains collections, themselves holding documents – each document is different and can have a variable number of fields. An advantage of using this kind of database design is that data does not need to be normalised like it would in a rational database. This is an advantage because the constrains caused by normalisation can degrade the performance of a database overtime. (Margot, 2021)

Modelling the data in terms of database structure is focused on the needs of the user. However, one issue with NoSQL data bases is that designing schema for NoSQL databases can be challenging because NoSQL has no standard modelling method. (Hamouda, 2017).

References
S. Hamouda and Z. Zainol. 2017) Document-Oriented Data Schema for Relational Database Migration to NoSQL. 2017 International Conference on Big Data Innovations and Applications (Innovate-Data), pp. 43-50.
Margot P. (2021) MongoDB: all about the document-oriented NoSQL database (datascientest.com). Available from: https://datascientest.com/en/mongodb-all-about-the-document-oriented-nosql-database [accessed May 8th, 2022]

Unit 8–11
Collaborative discussion on alternatives to NoSQL

My response to another student

Thank you, Richard, for your findings.

Designing data brings clarity, as it exposes knowns and unknowns about the data and planning brings to light any discrepancies. A document outlining the structure, design, and aims of the data enables streamlined teamwork and clarity of the product.

It might even be more important in the case of NoSQL *because* it is so flexible. For example, MongoDB would allow zipCode, zipcode, ZipCode, and postalCode to be saved in the same collection, as each field would be treated as a new field without any errors or warnings. To avoid such data frustrations and quality issues, a map of the information landscape is a necessity.
One suggestion is UMLtoNoSQL, as UML is a customary modelling language. UMLtoNoSQL automatically transforms a conceptual UML model mapping out data into a NoSQL physical model (Trimane et al, 2017). The benefits are twofold; I) the data is described according to common features of NoSQL (columns, document, graph) which enables mapping on several platforms and ii) it abstracts technical details of NoSQL systems. The logical level remains stable whilst the system evolves continuously (Trimane et al, 2017). The process can be automated using Model Driven Architecture (MDA), which main aim is to separate the functional specification of a system from the details of its implementation (Trimane et al, 2017).

References

Abdelhedi, F., Brahim, A.A., Atigui, F. and Zurfluh, G. (2017). *UMLtoNoSQL: Automatic Transformation of Conceptual Schema to NoSQL Databases*. [online] IEEE Xplore. doi:10.1109/AICCSA.2017.76.

Initial post by another student on schemeless of NoSQL

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 schemeless, 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 predefine 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 plainer 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 simper 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.

References
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.

Unit 8–11
Collaborative discussion on alternatives to NoSQL

My response to another student

Hi Wang, I agree with your findings regarding NoSQL, and I believe NoSQL is better equipped to meet the demands of big data.

A relevant aspect of NoSQL is its ability to ensure data is safe, as there is increasing demand from users. NoSQL's have undoubtedly improved workflows and programs; however, the speed, performance, and scalability are often at the expense of privacy-related issues (Okman et al, 2011).

General weakness with NoSQL is the lack of encryption support and poor authentication between servers and clients (Vonitsanos et al, 2020). Like SQL, NoSQL databases are susceptible to query injection attacks, in particular those which rely on JavaScript and PHP to enhance performance (Shahriar et al, 2017). For example, a vulnerability in MongoDB is the operator "$where" – similar to "where" in SQL – where a malicious actor can pass arbitrary data using the statement. (Shahriar & Haddad, 2017).
Authentication is disabled by default in MongoDB, however, basic Mongo provides the support for it on a per-database level, as users (Shahriar & Haddad, 2017). Weak authentication provides an actor with malice intent access to the database and can conduct intentional attacks, for example, masquerade attacks or the malice actor can use the database for other uses with malice intent (Turjman, 2020).
Studies on privacy and NoSQL are fragmented and do not adequately satisfy the current challenges of data protection.

References

G. Vonitsanos, E. Dritsas, A. Kanavos, P. Mylonas and S. Sioutas, "Security and Privacy Solutions associated with NoSQL Data Stores," *2020 15th International Workshop on Semantic and Social Media Adaptation and Personalization (SMA*, 2020, pp. 1-5, doi: 10.1109/SMAP49528.2020.9248442.
Shahriar, H. and Haddad, H.M. (2017). Security Vulnerabilities of NoSQL and SQL Databases for MOOC Applications. *International Journal for Digital Society*, 8(1), pp.1244–1250. doi:10.20533/ijds.2040.2570.2017.0153.
Hossain Shahriar, Hisham M. Haddad, International Journal of Digital Society (IJDS), Volume 8, Issue 1, March 2017
L. Okman, N. Gal-Oz, Y. Gonen, E. Gudes, and J. Abramov, ``Security issues in NoSQL databases," in Proc. IEEE 10th Int. Conf. Trust, Secur. Privacy Comput. Commun., Nov. 2011, pp. 541 547
Al-turjman, F. (2020). Intelligence and security in big 5G-oriented IoNT: An overview. *undefined*. [online] Available at: https://www.semanticscholar.org/paper/Intelligence-and-security-in-big-5G-oriented-IoNT%3A-Al-turjman/5e37557d7cd34df5a31b1f8ed60b2d874211244d [Accessed 12 May 2022].

Unit 8–11
Collaborative discussion on alternatives to NoSQL

Summary post on alternatives to NoSQL

NoSQL databases are a better alternative to relational databases in regard to Big data, however relational databases still has its advantages in relation to projects that are not Big Data related. Which one who suits a particular project depends on the circumstances.