

SQL vs NoSQL

Relational vs Non-Relational Databases

Software Engineering

CSCI 771/405

Hunter College

Spring 2021

Agenda

- Definition of SQL and NoSQL databases
- Comparison
- Pros and Cons
- Performance
- Applications
- Demo

SQL vs NoSQL

- An essential ingredient of big data and cloud-computing is a database that can accommodate very large number of users
- Distributed data storage is essential to process large amount of data (in many web apps used by Google, Facebook, Amazon,...)
- The web has changed the requirements of storage database systems for the next generation of applications
- SQL databases are relational databases with a standard query language
- NoSQL (stands for Not Only SQL) databases are non-relational or distributed databases
- NoSQL is an alternative to SQL with fast access time and no tolerance to any down time during failures
- NoSQL is used by large enterprises like Facebook, Google, Amazon...

SQL vs NoSQL

SQL

- Tables with rows and columns
- Relational database (links information from different tables)
- Structured language
- Static schema
- Supports ACID transactions

NoSQL

- Key-value pairs (JSON documents)
- Non-relational database (no structured mechanism to link data)
- No structured language
- Dynamic schema
- No support for ACID transactions

USERS		
<i>ID</i>	<i>First</i>	<i>Last</i>
1	Shane	Johnson

USER SKILLS	
<i>User ID</i>	<i>Skill Name</i>
1	Big Data
1	Java
1	NoSQL

USER EXPERIENCE		
<i>User ID</i>	<i>Role</i>	<i>Company</i>
1	Technical Mktg	Red Hat
1	Product Mktg	Couchbase

RELATIONAL DATABASE

```

{
  "firstName": "Shane",
  "lastName": "Johnson",
  "skills": ["Big Data", "Java", "NoSQL"],
  "experience": [
    {
      "role": "Technical Marketing",
      "company": "Red Hat"
    },
    {
      "role": "Product Marketing",
      "company": "Couchbase"
    }
  ]
}

```

JSON DOCUMENT

SQL vs NoSQL

- Data Model
 - NoSQL databases take many modelling techniques like key-value pairs, graph, and document data model
 - NoSQL database system may use two or more of the data models to represent the data
 - NoSQL data model does not use the table as storage structure of the data
 - Schema-less and very efficient in handling the unstructured data

SQL vs NoSQL

- Scalability

- Scalability in relational databases is a challenge due to vertical scalability (by adding more hardware resources like RAM, CPU, etc...)
- Very costly and impractical given hardware limitations
- NoSQL databases depend on the horizontal scalability which is not supported by SQL databases.

- Cloud

- The relational databases are not well suited for cloud environments because they do not support full content data search and are hard to scale beyond a limit
- NoSQL databases are the best solution for cloud databases because all the characteristics that define the NoSQL databases are very desirable for cloud databases

SQL vs NoSQL

- Big data handling
 - Big data handling is an issue in relational databases and the solution is the scalability and data distribution which take two forms: vertical or horizontal
 - In horizontal scalability data must be portioned into multiple servers which add complexity and performance issues related to these operations
 - NoSQL databases are designed to handle big data so they implement methods to improve the performance of storing and retrieving data
- Complexity
 - Complexity in relational databases arises because users must convert data into tables and when the data does not fit into those tables the structure of the database could be quite complex, difficult, and slow to process, unlike the NoSQL databases which have the capabilities to store unstructured, semistructured or structured data.

NoSQL & Agile Development

- NoSQL is the best fit for Agile development
- The dynamic implementation and the scalability of NoSQL databases go hand in hand with the Agile approach
- The incremental feature-driven Agile methodology requires the flexibility of NoSQL databases in terms of choosing or changing a model during the development process
- NoSQL is the solution for unstructured data; this fits the changing specifications of the agile approach that lack strict detailed planning of the product development
- Expands the functionality of the product to reach new users

Performance

- Performance of relational databases degrades with large datasets
- NoSQL was developed to overcome performance issues with SQL databases
- NoSQL provides high scalability but lacks a standard query language – a reason to lag behind SQL in terms of the number of users

Applications

- NoSQL databases are suited for business applications – it fits the continuous growth of unstructured data
- Applications that are accessed by a large number of users (e-commerce websites)
- Websites with huge and growing data such as social media networks (millions of tweets daily on Twitter)

Summary

- NoSQL does not use the relational data model thus does not use SQL language
- NoSQL stores large volume of data
- In distributed environment (spread data to different machines), we use NoSQL without any inconsistency
- If any faults or failures exist in any machine, then there will be no discontinuation of any work
- NoSQL is open source database, i.e. its source code is available to everyone and is free to use without any overheads
- NoSQL allows data to store in any record that is not having any fixed schema
- NoSQL does not use concept of ACID properties
- NoSQL is horizontally scalable leading to high performance in a linear way.

References

- [1] Dave, Meenu. (2012). SQL and NoSQL Databases. International Journal of Advanced Research in Computer Science and Software Engineering.
- [2] Mohamed, Mohamed & Altrafi, Obay & Ismail, Mohammed (2014). Relational Vs. NoSQL databases: A survey. International Journal of Computer and Information Technology (IJCIT). 03. 598.
- [3] Nayak, A. & Poriya, A. & Poojary, Dikshay. (2013). Article: Type of nosql databases and its comparison with relational databases. International Journal of Applied Information Systems. 5. 16-19.
- [4] Padhy, R.P., Ranjan, M., Suresh, P., Satapathy, C., & Pvt, O.I (2011). RDBMS to NoSQL: Reviewing Some Next-Generation Non-Relational Database's.
- [5] MongoDB Website <https://www.mongodb.com/>
- [6] Why NoSQL Is the Perfect Fit for Agile Development <https://www.dragonspears.com/blog/why-nosql-is-the-perfect-fit-for-agile-development>