

NoSQL Database

Bhavneet Soni

NoSQL Database

Traditional databases have some common characteristics such as data being presented in tabular form, very rigid type-scripting for the columns and having a unique identifying keys. These databases can be easily queried based on the relation between various columns of the table. Languages used to query such relational databases are known as structured query language (SQL). However with the advent of big data, especially multi-structure data types (Top 5 Considerations When Evaluating NoSQL Databases, 2016) relational databases are hard to maintain and proving to be less efficient when it comes to Big Data as they were not designed to take advantage of DFS and parallel processing. NoSQL databases store data unlike traditional related tabular forms, these are very flexible and implements schema-less data models which can be easily scaled and used over distributed architectures. Although these have been around since late 1960s, (NoSQL, 2017) but have started gaining notoriety with the advent of big data. NoSQL may have data model types of Documents, Graphs, Key-value pairs and wide columns (MongoDB, Nov 2016).

There are a number of NoSQL databases available but these can be broadly classified based on how these store data in to following categories

1. Key-Values based data-model implemented in databases such as Dynamo offer fast look up in associative arrays. These are very simple to implement since it consists of indexed key and value data is stored in a schema-less way. (What is NoSQL?, 2016)
2. Document based data-model implemented in databases such as MongoDB, these are sort of key value model, where a unique key is assigned to a document and is used to retrieve it from the database, it is mostly used in document management, indexing and blogs.

3. Columnar based data-model implemented in databases such as Hbase, store data in columns instead of the rows as in traditional databases. These provide highly scalable and high performance architecture, hence used where Analytics over large datasets at high speed ETL is required.
4. Graph based data-model implemented in databases such as Allegro are used on datasets which can be represented as a graphs with vertexes and edges, these are mostly used Semantic and Link analysis.
5. Multi-Model implement more than one or more types of data models.

Since NoSQL database can work with more than one storage mechanisms and can be implemented in different ways based on the end use it is sometimes also referred to as Not Only SQL (Rouse & Beal, n.d.) .

References

MongoDB. (Nov 2016). Retrieved from <https://www.mongodb.com/nosql-explained>

NoSQL. (2017, Jan 26). Retrieved from www.wikipedia.org:

<https://en.wikipedia.org/wiki/NoSQL>

Rouse, M., & Beal, B. (n.d.). *NoSQL (Not Only SQL database)*. Retrieved from

<http://searchdatamanagement.techtarget.com/definition/NoSQL-Not-Only-SQL>

Top 5 Considerations When Evaluating NoSQL Databases. (2016, Nov). *A MongoDB White Paper*. Retrieved from <https://www.mongodb.com>.

What is NoSQL? (2016). Retrieved from <https://academy.datastax.com/planet-cassandra/what-is-nosql>