

Problem Statement

NoSQL data stores are very popular these days. This article will show why they are so popular. Then we will also discuss on the problems that nosql data stores can solve while the relational databases cannot.

References

- [1] Wikipedia. Nosql - wikipedia. <https://en.wikipedia.org/wiki/NoSQL>. Accessed April 19, 2018.
- [2] David C. Hay. A comparison of data modeling techniques. <http://www.essentialstrategies.com/publications/modeling/compare.htm>. Accessed April 5, 2018.

Overview of NoSQL Data Stores

NoSQL data stores are data stores that is modeled not in tabular relations which are used in relational databases.[1] If they are not modeled in tabular relations, then what are the structure of data stored in a NoSQL data store?

There are several types of data models used by nosql data stores. For example, key-value model is one of the simplest model used in NoSQL data stores.[1] And NoSQL data stores that use key-value models are called key-value stores. Similarly, there are document stores, graph databases, object databases and so on.

Critical Thinking

Let's compare NoSQL data stores with relational databases on the problems that NoSQL data stores can solve while relational databases cannot.

First, NoSQL data stores can free people from design predefined schema. While users of relational databases must define the schema first before add any information to the database.[2]

Second, NoSQL data stores can deal with the unstructured data such as videos, audios and email which makes up much of data on the Internet today's. While, relational databases can only handle structured data.[2]

Finally, NoSQL data stores is much easier to "horizontal" scaling to clusters of servers than relational databases due to the simpler design of the NoSQL data stores.[1]

Question

When I looked up for the different types of NoSQL data stores, I found there is a type of data store that is named tabular store. Unlike the other NoSQL data stores such as key-value stores and document stores that are very different from traditional relational databases, a tabular store like Apache HBase manages data as a table just like relational databases do.

So why are tabular stores not relational databases but NoSQL data stores? And what's the advantages of a tabular store?

Method

Describe how you are going to answer your own question stated above.

Analysis and Discussion

HBase is a database that is built on HDFS¹. It has very different concepts of processing, retrieving and storing the data or information than a relational database. First, it is column-oriented while most relational database management systems are row-oriented. And tables in HBase can be treated as inputs and outputs for MapReduce² jobs on the Hadoop framework, which is very different from using SQL on relational databases.

Let's talk about the advantages of HBase. Most importantly, HBase can be easily horizontally scaled. That means HBase based on HDFS can be used on distributed environment like a cluster of machines. And it deals with sparse tables well. Also, HBase has flexible schema which, of course, is the common advantage that a NoSQL data store has.

¹Hadoop Distributed File System, which is a file system that is well suited for distributed environment that is commonly used in clusters of servers nowadays. Many famous framework like Hadoop, Spark and HBase are built on it.

²A programming model that can be used when processing big data using distributed environment.