

New ways of managing large datasets

István Péntek

University of Debrecen
istvan.pentek@gmail.com

Abstract

A Web-based Information System should handle huge databases. A community site or a money management system can provide several hundred gigabytes data per day and need to serve web requests under a small time slot or almost real time. This database size growth cause slow Webbased systems and long processing time. The Standard Query Language (SQL) is a popular and powerful language with a well-defined and strict standard. The relational database management system (RDBMS) implementations used SQL as query tool. SQL and RDBMS are not equal but close to each other. The main problem with the RDBMS and SQL implementations are not with the SQL standard. The manufacturer implementations are different and differ from the latest SQL standards. The problem is not the query language standard or the implementations, the problem is the database management system core implementations and old architecture. The industry try to find proper solutions. There are two new ways to store and query big data. The first is NoSQL which is based on non-structured or hierarchical data. This solution is instead of SQL and RDBMS. The NoSQL implementations are very different and based on different principle. The second approach is the NewSQL which uses SQL as query tool but the core implementation is more scalable than OldSQL implementations. NewSQL extends the SQL standard and improves the readability of SQL queries. Both approach based on new ideas and give more than the OldSQL implementations. These new databases more cloud compatible than the single-noded OldSQL and can work parallel. This article discuss the possibilities of these new technologies and the architectural differencies.

Keywords: OldSQL, NoSQL, NewSQL