Abstract

There are two obvious methods to map a two-dimension relational database table onto a one-dimensional storage interface: store the table row-by-row, or store the table column-by-column. Traditionally, database system implementations and research have focused on the row-by-row data layout, since it performs best on the most common application for database systems: business transactional data processing. However, there are a set of emerging applications for database systems for which the row-by-row layout performs poorly. These applications are more analytical in nature, whose goal is to read through the data to gain new insight and use it to drive decision making and planning. In this paper, we study the poor performance of row-by-row data layout for these emerging applications, and evaluate the column-by-column data layout opportunity as a solution to this problem. The solution will be analyzed and represented by graph. At the end of the paper we will see the comparative performance of Oracle 10g and MSSQLServer database.
- Daniel J. Abadi, In CIDR, Asilomar, CA, USA, 2007, Column stores for wide and sparse data.
- P. A. Boncz and M. L. Kersten. MIL primitives for querying a fragmented world.

Index Terms

Computer Science
Information Sciences

Keywords
Databases Database Systems Row Store Column Store Performance Tuning