Abstract

In this paper, we are paying attention to the problem of poor performance of row-by-row data layout for the emerging applications, and evaluate the column-by-column data layout opportunity as a solution to this problem. There have been a variety of proposals for how to build a database system on top of column-by-column layout. These proposals have different
levels of implementation effort, and have different performance characteristics. If one wanted to build a new database system that utilizes the column-by-column data layout, it is unclear which proposal to follow. This paper provides (to the best of our knowledge) the detailed study of multiple implementation approaches of such systems, categorizing the different approaches into three broad categories, and evaluating the tradeoffs between approaches.

Reference


Index Terms

Computer Science

Communications
Key words

Column oriented database

c-store

Query execution plans

row stores