We investigate the usability and performance of the UB-Tree (universal B-Tree) for multidimensional data, as they arise in all relational databases and in particular in data-warehousing and data-mining applications. The UB-Tree is balanced and has all the guaranteed performance characteristics of B-Trees, i.e., it requires linear space for storage and logarithmic time for the basic operations of insertion, retrieval and deletion. Therefore it can efficiently support OLTP. In addition the UB-Tree preserves clustering of objects with respect to Cartesian distance. Therefore, it shows its main strengths for multidimensional data. It has very high potential for parallel processing. A single UB-Tree can replace a large number of secondary indexes and join indexes including foreign column join indexes (FCJ). For updates this means that only one UB-Tree must be managed instead of several secondary indexes. This reduces runtime and storage requirements substantially. For retrieval the UB-Tree has multiplicative complexity with respect to the relative size of the ranges for range queries, resulting in a dramatic performance improvement over multiple secondary indexes which have additive range query complexity. Furthermore, using the Tetris-Algorithm the UB-Tree enables reading data in any arbitrary sort order without the necessity of external sorting. Thus data need to be read only once to perform most of the operations of the relational algebra, such as ordering, grouping, aggregation, projection and joining. Therefore, the UB-Tree can support OLAP very efficiently. It is useful for geometric databases,
data-warehousing and data-mining applications, but even more for databases in general, where multiple secondary indexes on one relation or FCJ-indexes to join several relations are widespread, which can all be replaced by a single UB-Tree index. Therefore, the difficult index selection problem [GHRU97] largely disappears and the UB-Tree offers the potential to integrate OLAP with OLTP in the same processing environment.

Stichworte: UB-Tree; Multidimensional Range Queries; Query Optimization; Relational Algebra; Query Processing; Caching; Searching; Sorting

Jahr: 1998

Jahr / Monat: 1998-09-01 00:00:00

Seiten: 16

Occurences: · Einrichtungen > Fakultäten > Fakultät für Informatik > Technische Berichte > 1998

entries: