Database cracking

Stratos Idreos *
CWI, Amsterdam
The Netherlands

Martin Kersten
CWI, Amsterdam
The Netherlands

Stefan Manegold
CWI, Amsterdam
The Netherlands

Abstract

Database indices provide a non-discriminative navigational infrastructure to localize tuples of interest. Their maintenance cost is taken during database updates. In this work we study the complementary approach, addressing index maintenance as part of query processing using continuous physical reorganization, i.e., *cracking* the database into manageable pieces. Each query is interpreted not only as a request for a particular result set, but also as an advice to crack the physical database store into smaller pieces. Each piece is described by a query, all of which are assembled in a *cracker index* to speedup future search. The cracker index replaces the non-discriminative indices (e.g., B-trees and hash tables) with a discriminative index. Only database portions of past interest are easily localized. The remainder is unexplored territory and remains non-indexed until a query becomes interested. The cracker index is fully self-organized and adapts to changing query workloads.

With cracking, the way data is physically stored self-organizes according to query workload. Even with a huge data set, only tuples of interest are touched, leading to significant gains in query performance. In case the focus shifts to a different part of the data, the cracker index will automatically adjust to that.

We report on our design and implementation of cracking in the context of a full fledged relational system. It led to a limited enhancement to its relational algebra kernel, such that cracking could be piggy-backed without incurring too much processing overhead. Furthermore, we illustrate the ripple effect of dynamic reorganization on the query plans derived by the SQL optimizer. The experiences and results obtained are indicative of a significant reduction in system complexity with clear performance benefits.

*Stratos Idreos is the contact author (Stratos.Idreos@cwi.nl) and a Ph.D student at CWI