

Architectural optimizations in databases

Data Morphing: An Adaptive, Cache-Conscious Storage Technique

Rick van der Zwet J.K. Vis Others

LIACS - Leiden University

Advanced Compilers and Architecturers, 2010

Outline

- 1 Data Morphing
 - The Basic Problem That We Studied
 - The Approach

- 2 Results
 - Main Results

Database Storage Methods.

The Order Matters

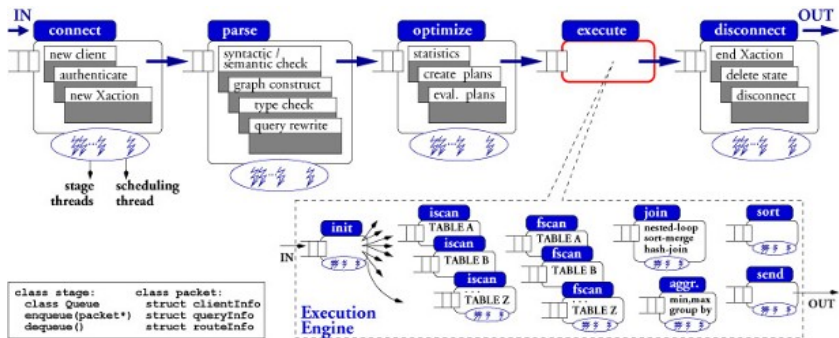
- Traditional Row Based, Space Oriented Storage
- Column based, New Trend. Search Oriented Storage.
- Both has advantages and disadvantages

Row Based Storage.

- Focused on keeping data of a related entry close to each other
- Common implementation N-ary storage model

- | | | |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 5 | 6 |

Row Based Storage Engine.



Column Based Storage

- Focused on allowing fast search on attributes
- PAX storage model used.

- | | | |
|---|---|---|
| 1 | 3 | 5 |
| 2 | 4 | 6 |

Definitions

- *group*: Set of attributes that are written to consecutive memory addresses
- *partition*: Set of groups that uniquely defines the position of every attribute in a relation
- *zone*: Area of a page where all instances of a particular group are written
- *zone-record*: defines an instance of the attributes in a particular group

The Algorithm

- 1 Calculating cache-efficient storage template.
- 2 Re-organizing the data into a cache-efficient organization.
 - 1 Static evaluation, based on heuristics
 - 2 Dynamic restructuring based a request recognition.

Conclusions and Future Work

- Speed-up made, but only on specific well-defined datasets.
- Finding a general optimisation algorithm is hard.
- Experiments did only cover memory based databases.

Summary

- Data Morphing is a Promising Method.
- Combining both Row Based Storage and Column Based Storage could lead to nice speeds-up.
- Depending which method to choose how-ever is really hard.

- Outlook
 - Make it scale so it also take the variable 'disk access' into count.
 - Build a framework to allow easy and proper testing for various combinations.

For Further Reading I



Richard A. Hankins

Data Morphing: An Adaptive, Cache-Conscious Storage Technique

Proceedings of the 29th VLDB Conference, Berlin, Germany, 2003.