17 - An Experimental Comparison

of Thirteen Relational Equi-Joins in Main Memory

1. Overview of Idea

There are many join algorithms like non-partition algorithm, partition-based algorithm, and sort-merge algorithm, but comparison of these algorithms' performance is not done yet. In order to make good query plan, we should know which algorithm is better on given situation or which algorithm is best.

2. Main Finding

Partition based join algorithm with huge page is good choice with almost situation. But, input size and NUMA-awareness matters on performance too.

3. Systems used and its Specifications

13 different join algorithms are used, classified by NOP*(no partition), PR*(partition based), CPR*(chunked partition based). NOP* and PR* are implemented according to related papers, CPR* are original algorithms by the authors.

4. Workloads evaluated

TPC-H workload #19 is used, varied by dataset size, only join operation. After that, evaluated same query to measure full performance.