14 – Efficiently Compiling Efficient Query Plans for Modern Hardware

1. Overview of Idea
   With growing of main memory capacity, query performance is determined by CPU performance. To use CPU efficiently, we should use machine code and high level language, and make query data-centric.

2. Main Finding

   Data-centric query operation model is efficient

3. Systems used and its Specifications
   HyPer (in memory DBMS) and disk-oriented DBMS are used. Also, for test, DB X, Ingres Vectorwise, MonetDB are used for comparison.

   A. Register Friendly Compiled Query Code
      simple operations are implemented in LLVM, which can use CPU register directly. Some complex operations like sorting and hashing are implemented in C++.
      vectorwise instructions are used.

   B. Pipeline Friendly Architecture
      not use volcano model or vectorwise model
      decompose query by 'pipeline breaker' so that query plan can be pipeline friendly. Each pipeline breakers are not single operator on data, these are combination of operation on data
      simply to say, reform the query to data-centric operations

4. Workloads evaluated
   TPC-C and TPC-H are used.