Contribution ID: **31**   Type: **not specified**

[DEMO] PostgresRaw

Wednesday 29 May 2013 14:20 (5 minutes)

As data collections become larger and larger, users are faced with growing bottlenecks in their data analysis. One such bottleneck is the time to prepare and load data into a database system, which is required before any queries can be executed. For many applications, this data-to-query time, i.e. the time between first getting the data and retrieving its first meaningful results, is a crucial barrier, and a major reason why many applications already avoid using traditional database systems altogether. As data collections grow, however, the data-to-query time will only grow.

In this demonstration, we will showcase a new philosophy for designing database systems called NoDB. NoDB aims at minimizing the data-to-query time, most prominently by removing the need to load data before launching queries. We will present our prototype implementation, PostgresRaw, built on top of PostgreSQL, which allows for efficient query execution over raw data files with zero initialization overhead. We will visually demonstrate how PostgresRaw incrementally and adaptively touches, parses, caches and indexes raw data files autonomously and exclusively as a side-effect of user queries. Moreover, we will demonstrate with "live races" how PostgresRaw outperforms traditional database systems across a variety of workloads.

Author: Dr BRANCO, Miguel (EPFL)

Presenter: Dr BRANCO, Miguel (EPFL)

Session Classification: Lightning Talks on Demand