Lecture 7:

Trends in the evolution of database systems

Rasmus Pagh

Goals of lecture

- Present some recent views and ideas about where DBMSs are going.
  - Opinions can be more interesting than facts!
  - Do not expect to hear the last word today.
- Complement textbook which talks mainly about developments that are well-established.
- Hope to stimulate questions and discussion.

Contents

- First I talk about a recent paper by Michael Stonebraker: *The End of an Architectural Era*.
- Then we watch a video presentation by Jeff Dean, Google, who talks about Google's internally developed DBMS, BigTable.
  - We will pause for questions and discussion.
- Along the way, some point from the paper *Beyond Relational Databases* by Margo Seltzer will be made.

Stonebraker talk

- See separate slide set.

Stonebraker summary

- Many data management scenarios have special needs, and special features that can be exploited to gain performance.
- For OLTP it is:
  - feasible to switch to main memory DBs
  - worth considering single-threaded execution
  - often possible to simplify logging
- Some H-store design goals
  - Use grid computing (cheap, extensible)
  - Achieve high availability
  - ”No knobs”

Bigtable talk

- See separate slide set (handed out) + video.

Bigtable summary

- Subset of usual DBMS functionality (”meet 7 out of 8 demands”).
- Primitive interface (not SQL!)
  - User makes some crucial choices, e.g. memory versus disk.
- Built on top of other components:
  - GFS file system
  - Chubby distributed locking system
  - MapReduce (queries + data manipulation)
- Highly scalable on clusters of cheap machines - add machines to scale up.
- Highly fault tolerant.

Guest lecture, Tuesday 1PM

- Mogens Nørgaard, Miracle A/S, talks: *"You Probably Don't Tune Right (and you probably never have)"*
- The presentation will explain the four Worst Practices to date:
  - Best Practices (vs Best Evidence)
  - Counters (don't) count,
  - Guess & Grimacing,
  - Ground-hog Day Tuning.
- Then the Grand Unifying Solution (GUS) will be presented. Finally the problems with GUS will be discussed briefly.