

My Top 10 Disappointments

by

Michael Stonebraker

Number 1

- Optimizers are still not robust 45 years later
 - Tamr guys complain bitterly about multiple commercial optimizers
 - And exclusively use hints
- No shortage of interesting ideas (plan interpreters, sampling, ...)

Number 2

- Our systems are too hard to use and have a horrible “out of the box” experience
 - Lots of ways to deal with this
 - Few implementations
 - And vendors love consulting/tuning gigs

Number 3

- New kinds of systems often built by people with no DBMS experience
 - E.g. Map-Reduce, NoSQL, Spark
- Creates a lot of “churn”
 - NoSQL really means “No yet SQL”

Number 4

- SQL is a horrible legacy
 - See Date '86 for a blistering critique
 - Thank IBM for this one

Number 5

- JDBC/ODBC is a horrible legacy
 - Thank IBM again
 - At least this is going away (Linq, Ruby-on-Rails, ..)

Number 6

- Vertica was not open source
 - Too early
- Reliable open-source parallel data warehouse systems have been slow to appear
-

Number 7

- We miss the boat on semi-structured data
 - XML vs JSON

Number 8

- We are nowhere on data curation
 - 800 pound gorilla in the corner

Number 9

- Nothing higher level than SQL has stuck
 - Certainly not Datalog

Number 10

- There is no semantic transaction theory
 - Serializable or no guarantee
 - A lot of operational systems (snapshot isolation, weak isolation, eventual consistency, ...) with no guaranteed semantics

And Yes

- The paper diarrhea of LPUs is pitiful