Lazy users and agile databases: toward timely, predictable and cheap data analytics

Renata Borovica-Gajic and Anastasia Ailamaki

The Need for Agile Approach in building DBMS

What every business analyst wants

Why timely and interactive?

Why predictable?

Why cost-efficient?

Challenges for DBMS

• Why wait for loading and tuning?
  • Issues for data exploration:
  • Long data-to-insight time due to high startup cost

Time to first insight

Data loading

Timer

Data to-insight time due to high startup cost

State of affairs in database systems

Quest for robust & predictable execution

Region Snooping with Smooth Scan

Data delivered is not necessarily data sought

Smallest granularity

Statistical - unreliable advisor

Reoptimization - risky

 Executable code

Rounding-off

Ineffective Adaptive

Robust = Near-optimal execution for all inputs

Data driven adaptation = near-optimal performance

Data analytics for a penny

Database storage tiering

Cold storage in the tiering hierarchy

Skipper: adaptive CSD-driven execution

Cold in capacity tier

Cold storage: subsume archived and capacity tiers

Support for out-of-order execution

Skipper in Action

Cold in capacity tier

Cold storage: subsume archived and capacity tiers

Support for out-of-order execution

Skipper in Action