Magdalena balazinska

Nov 29 2012

Large parallel systems

Map reduce type systems

Parallel data management

Big data

Make it easier, faster, flexible

Easy:

Autocomplete SQL

Explanations for mistakes

Fast:

Partition data

Ex: examine solar system

    Problem: dense regions of the universe are skewing the data

Solution: partition smaller, but not too small, non uniform partitions are optimal

Checkpoint write to disk SQL queries for large scale parallel databases

Flexible:

SciDB:

Partition data but have data between partitions saved in its own partition for analyzing overlapping partitioned data.