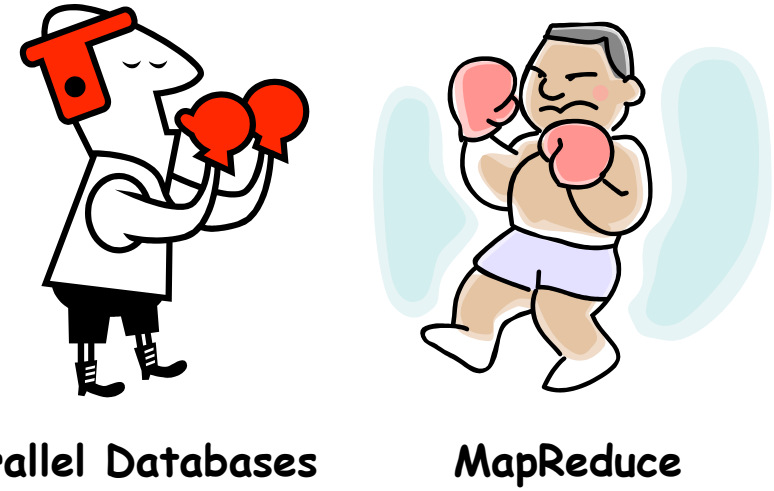
MapReduce: A Major Step Backwards (or maybe something else?)

Eugene Shekita, IBM Almaden Research



Parallel Databases MapReduce

Outline

- **What parallel databases got right**
  - parallel databases == shared nothing RDBMS
- **Interesting ideas in MapReduce (the architecture)**
  - MapReduce: Simplified Data Processing on Large Clusters, Dean and Ghemawat, SOSP 2004
- **Things people like about Hadoop (the open source MapReduce)**
  - No IBM product but lots of use at our labs and customer feedback
  - Also IBM/Google/NSF university program – hosted Hadoop clusters
- **A couple of questions for the panel**

What Parallel Databases Got Right

- **30 years of research: data partitioning, indexing, parallel sorts/joins/aggregation, column stores, etc**
  - Multiprocessor Hash-Based Join Algorithms, DeWitt and Gerber, VLDB 1985
  - The Case for Shared Nothing, Stonebraker, IEEE 1986
  - Parallel Sorting on a Shared-Nothing Architecture, DeWitt et al., PDIS 1991
  - Adaptive Parallel Aggregation Algorithms, Shatdal and Naughton, SIGMOD 1995
  - C-Store: A Column-Oriented DBMS, Stonebraker et al., VLDB 2005
- **Lots of commercial products**
  - Teradata, IBM DB2, Vertrica, Greenplum, Netezza, Paraccel, Netezza, Aster Data, Microsoft DATAllegro...
- **A multi $B industry**

Interesting Ideas in MapReduce (the architecture)

- **A very flexible UDF framework**
  - Able to handle almost any data type – records, arrays, images, inverted indexes, etc
- **A runtime job scheduler and load balancer**
  - Doesn't need to understand what's going on in the UDFs or rely on flakey complier cost models
- **Intra-job fault tolerance and straggler handling**
  - Critical for long running tasks
- **Use of a fault tolerant DFS**
  - Over direct-attached storage, 3x replication
  - Any node can get to any block but expose block locations to enable function shipping

Things People Like about Hadoop

- **$0 to get started**
  - No viable open source parallel database
- **Scalability and fault tolerance on commodity hardware**
  - Easy to set up
  - Works ok without any tuning
- **Freedom from the "Warehouse Priesthood"**
  - Analysts like to load data in HDFS and experiment with it
  - A rigid warehouse schema is often not what they want

Things People Like about Hadoop

- **Open source brings innovation and choices**
  - Query languages (Pig, Hive, Jaql, Cascading)
  - HDFS InputFormats (SequenceFiles, Thrift, Avro, ProtocolBuffers, …)
  - ML libs (Mahout)
  - R libs (R-Forge)
  - …
- **Extensibility & Programmability of the platform** 
  - Able to do stuff they probably couldn't do in a parallel database
    - Text analysis, image mining, statistical machine translation, credit card fraud analysis, drug interaction analysis, entity resolution, monte carlo simulation…

Panel Questions

- **Is it silly for MapReduce to try and be like a parallel database?**
  - SQL on Hadoop: Facebook's Hive

Hive on an 11-node cluster

TPC-H Power@Size	TPC-H Throughput@Size	QphH@Size	Price	Price/Performance
436.2926511	350.916137	391.282675	33000$	84.33800449

top 10 are under 5.0

- **Is it silly for a parallel database to try and be like MapReduce?**
  - MapReduce in SQL: Greenplum, Aster Data, …

- **Is a collision coming?**
  - Does anybody get run over?