

MapReduce Extensions and Contemporaries

- Pig (Yahoo, available open source)
 - Relational Algebra over Hadoop
- HIVE (Facebook, available open source)
 - SQL over Hadoop
- Impala *Cloudera*
 - SQL over HDFS; uses some HIVE code
- Cascading
 - Relational Algebra
- Dryad (Microsoft, sadly not available)
 - Relational Algebra
- Clustera (U of Wisconsin, not available)
 - Relational Algebra

MapReduce vs RDBMS

- RDBMS

- Declarative query languages
- Schemas
- Logical Data Independence
- Indexing
- Algebraic Optimization
- Caching/Materialized Views
- ACID/Transactions →

Pig, HIVE

HIVE, Pig, DryadLINQ

Hbase

Pig, (Dryad, HIVE)

Hadoop

- MapReduce

- High Scalability 1000x
- Fault-tolerance
- “One-person deployment”