MapReduce and Parallel DBMSs: Together at Last

Andrew Pavlo
New England Database Summit – MIT CSAIL
January 28, 2010
Co-Authors

- Daniel Abadi (Yale)
- David DeWitt (Microsoft)
- Samuel Madden (MIT)
- Erik Paulson (Wisconsin)
- Alexander Rasin (Brown)
- Michael Stonebraker (MIT)
Today’s Talk

- SIGMOD ‘09
  - A Comparison of Approaches to Large-Scale Data Analysis

- CACM ‘10
  - MapReduce and Parallel DBMSs: Friends or Foes?
  - Compare/Contrast with Jeffrey Dean & Sanjay Ghemawat (Google)
Outline

- Introduction
- Benchmark Study & Results
- Sweet Spots
- Together At Last
- Concluding Remarks
Benchmark Environment

- **Tested Systems:**
  - *Hadoop (MapReduce)*
  - *Vertica (Column-store DBMS)*
  - *DBMS-X (Row-store DBMS)*

- 100-node cluster at Wisconsin

- Additional configuration information is available on our website.
Benchmark Tasks

- Original MR Grep Task:
  - *Find 3-byte pattern in 100-byte record*
  - *Dean et al. (OSDI ‘04)*

- Analytical Tasks:
  - *Web Log Aggregation*
  - *Table Join with Aggregation*
  - *User-defined Function*
## Results Summary

<table>
<thead>
<tr>
<th>Task</th>
<th>Hadoop</th>
<th>DBMS-X</th>
<th>Vertica</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grep Task</td>
<td>284 sec</td>
<td>194 sec</td>
<td>108 sec</td>
</tr>
<tr>
<td>Web Log</td>
<td>1146 sec</td>
<td>740 sec</td>
<td>268 sec</td>
</tr>
<tr>
<td>Join</td>
<td>1158 sec</td>
<td>32 sec</td>
<td>55 sec</td>
</tr>
</tbody>
</table>

- Full results are available in our SIGMOD & CACM papers.
Outline

- Introduction
- Benchmark Study & Results
- Sweet Spots
- Together At Last
- Concluding Remarks
Extract-Transform-Load

- “Read Once” data sets:
  - Read data from several different sources.
  - Parse and clean.
  - Perform complex transformations.
  - Decide what attribute data to store.
  - Load the information into a DBMS.

- Allows for quick-and-dirty data analysis.
Semi-Structured Data

- MapReduce systems can easily store semi-structured data since no schema is needed:
  - Typically key/value records with a varying number of attributes.

- Awkward to store in relational DBMS:
  - Wide-tables with many nullable attributes.
  - Column store works better.
Limited Budget Operations

- MapReduce frameworks:
  - Community supported and driven.
  - Attractive for projects with modest budgets and requirements.

- Parallel DBMSs are expensive:
  - No open-source option.
Together At Last?

- What can *MapReduce* learn from *Databases*?
  - *Fast query times.*
  - *Schemas.*
  - *Supporting tools.*

- What can *Databases* learn from *MapReduce*?
  - *Ease of use, “out of box” experience.*
  - *Attractive fault tolerance properties.*
  - *Fast load times.*
MR+DBMS Integration

- Vertica now integrates directly with Hadoop:
  - *Hadoop jobs can use Vertica as input source.*
  - *Push Map/Reduce tasks down directly into DBMS nodes.*

- Other notable commercial MR integrations:
  - *Greenplum*
  - *AsterData*
  - *Sybase IQ*
MR+DBMS Integration

- HadoopDB (Yale+Brown):
  - Replace Hadoop’s distributed file system with multiple database instances.
  - Rewrite Hive query plans into localized SQL for each execution node.

- Position available for HadoopDB @ Yale
Other Work

- **MRi (Wisconsin):**
  - Improving Hadoop by adding DBMS technologies that are transparent to users.
  - Ported GiST Search Trees to Hadoop.

- **SQL Server 2008 R2 (Microsoft):**
  - Including “MapReduce-like” functionality into parallel data warehouse version of MSSQL (Project Madison)
Conclusion

- Complete benchmark information and source code is available at our website:
  - http://database.cs.brown.edu/sigmod09/

- Questions/Comments?