A Comparison of Approaches to Large-Scale Data Analysis

Magnus Kirø

Norwegian University of Science and Technology

October 30, 2012

Presentation Goals

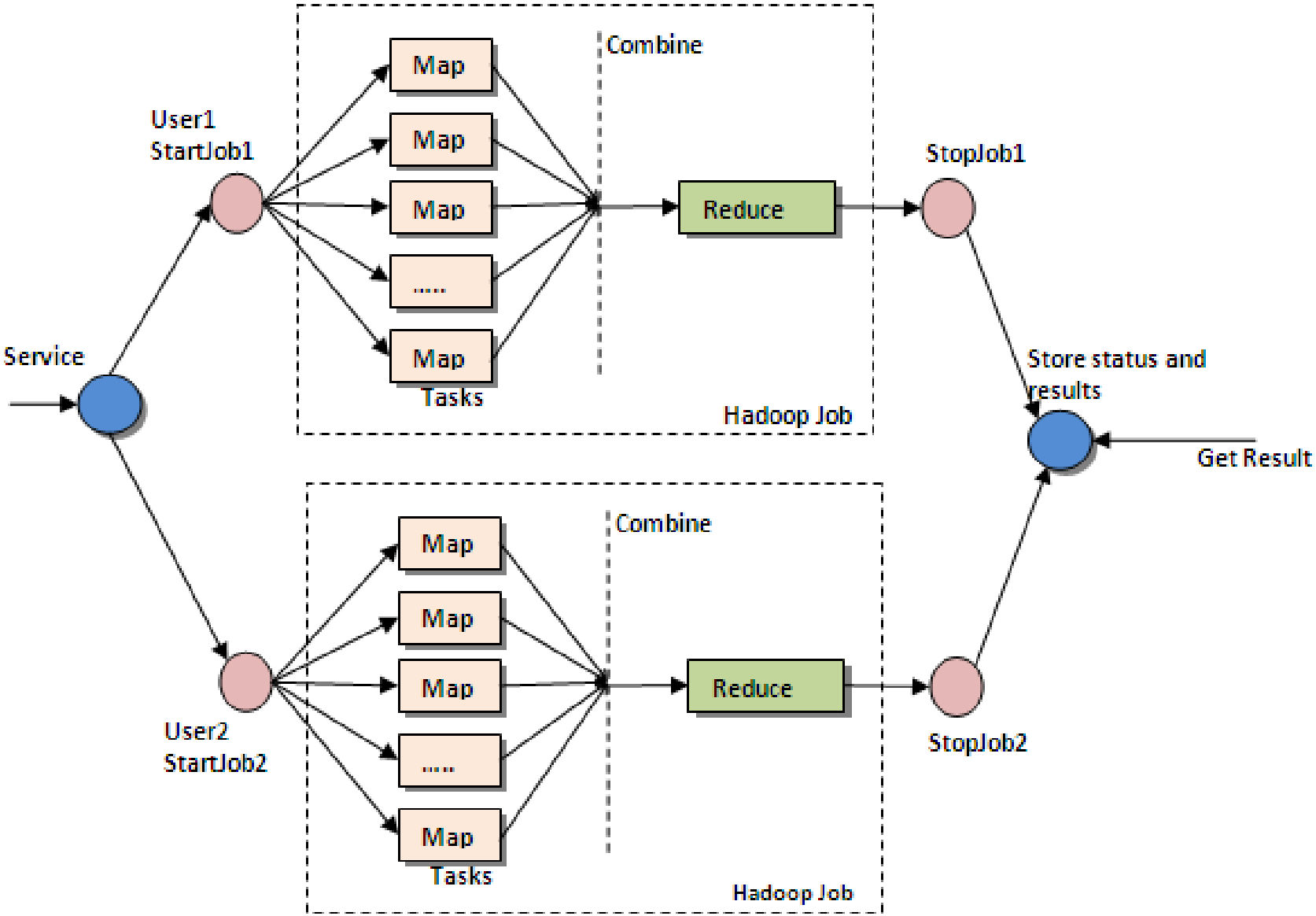
The purpose of paper is to consider MapReduce and parallel Database Management Systems for large-scale data analysis.

1. **Parallel DBMS** and **MR**, two approaches to large-scale data analysis.
2. The **Architectural Elements** of MR and DBMSs.
3. **Benchmarks**, tests and results.
4. **Discussion**, which conclusions can we draw from the tests and their results?

1. Two Approaches
   - Architectural differrences of the two approaches
2. Benchmark
   - The different tasks of the benchmark and it's execution
3. Discussion
   - Pros and cons of the results and setup
4. Conclusion
   - Summary of solutions and drawbacks

Map Reduce (MR)

1. It's simplicity makes MR attractive.
2. Basically: Throw data in a bucket and read it when needed.
3. Only two Functions:
   1. **Map**: Map data into files that are stored in the underlying distributed file system.
   2. **Reduce**: Compiles the output data from a mapping function to creata a combined result to the query.
4. The map and reduce functions have to be implemented.


Combine
Map
Map
Map
.....
Map
Tasks
User1
StartJob1
Reduce
StopJob1
Service
Store status and
results
Get Result
Hadoop Job
Combine
Map
Map
Map
.....
Map
Tasks
User2
StartJob2
Reduce
StopJob2
Hadoop Job

Parallel Database Management System (DBMS)

1. Tables are partitioned across nodes
2. Query optimizer, that translates SQL to a query plan. Execution of the query plan is divided among multiple nodes.
3. Underlying storage details can be disregarded by the programmers.

Schema Support

1. MR does not have Schema support. Manual data integrity enforcement is required.
2. DBMS has Schema support. Data integrity is automatically enforced by the schema.

Indexing

1. MR does not have inbuildt indexing. Again the programmer has to implement it, if the functionality is wanted.
2. DBMS provides indexing.

Programming Model

1. MR, Codasyl style, provide an algorithm to get the data you want.
2. DBMS, Relational style, state what you want.

Data Distribution

1. MR: get all documents, then compute the result.
2. DBMS: distributes code to all nodes, the nodes compute partial answers, answers are combined into the result.

Execution Strategy

1. MR: Pull data. Nodes*Maps files - potentially a severe performance problem.
2. DBMS: Push data.

Flexibility

1. MR has the most flexibility. You can do nearly whatevery you want. But you have to enforce your own rules.
2. DBMS is strict and limited, but comes with great support after a long development time and lots of use.

Fault Tolerance

1. MR: Node crash - task is recheduled to another node. Only that subtask is lost in computing time.
2. DBMS: Node crash - the whole transaction has to be restarted. Might be very expensive.

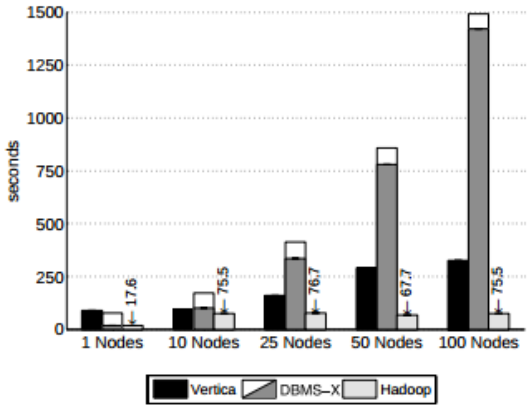
Environment

1. Hadoop, DBMS-X and Vertica.
2. Hadoop whitout compression. The rest with.
3. Task execution: Each task was executed three times.
4. All systems was optimized for the tasks given.

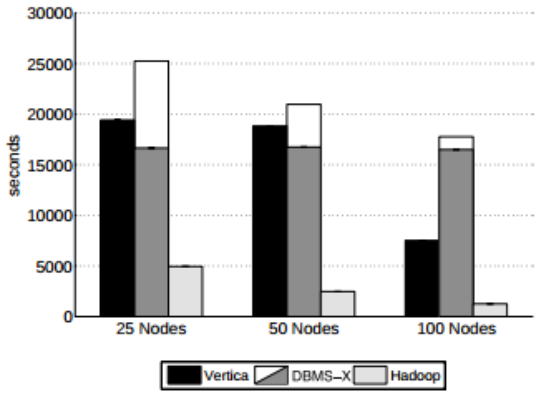
Grep Task

1. Scan all files for a string pattern.
2. 100byte records, 10byte key, 90 byte random data. once in every 10.000 records.
3. Hadoop: Command line to copy data to FS. Significant startup cost.
4. DBMS: Hash aware load data.
5. Vertica: Provides a copy cmd.

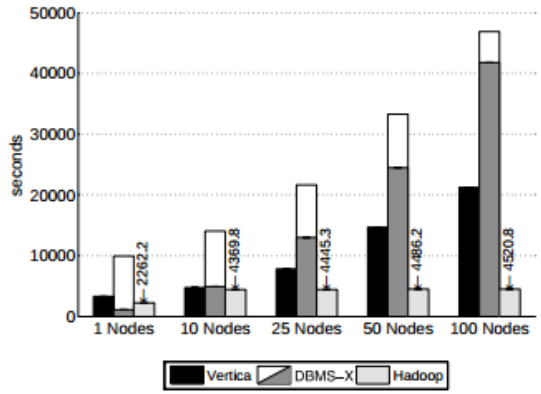
Grep Load Times



1: Load Times – Grep Task Data Set (535MB/node)

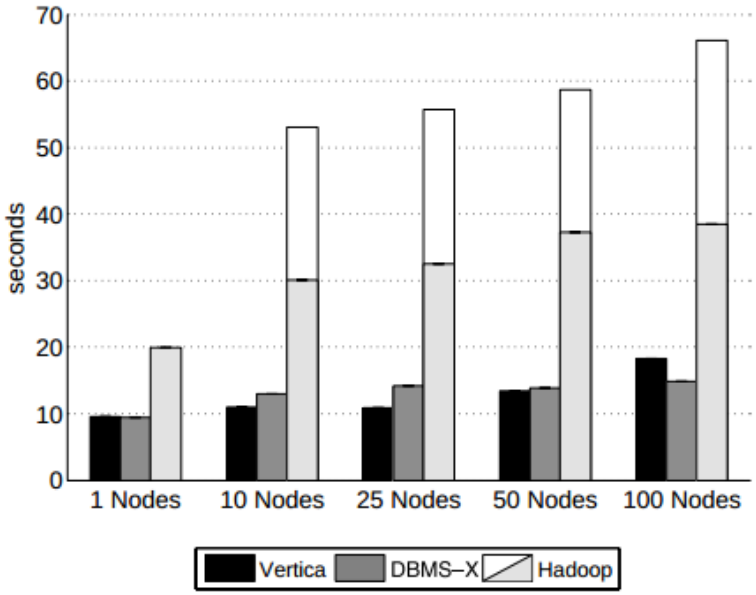


2: Load Times – Grep Task Data Set (1TB/cluster)

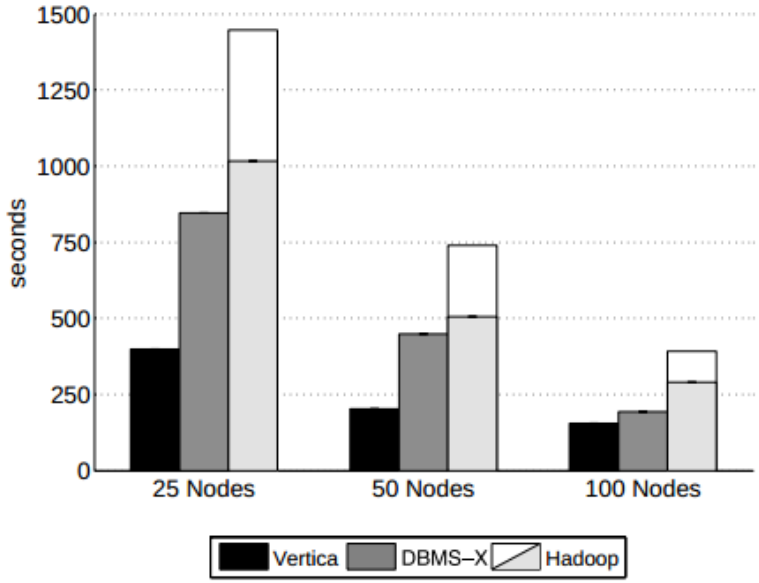


3: Load Times – UserVisits Data Set (20GB/node)

Grep Task Results



4: Grep Task Results – 535MB/node Data Set

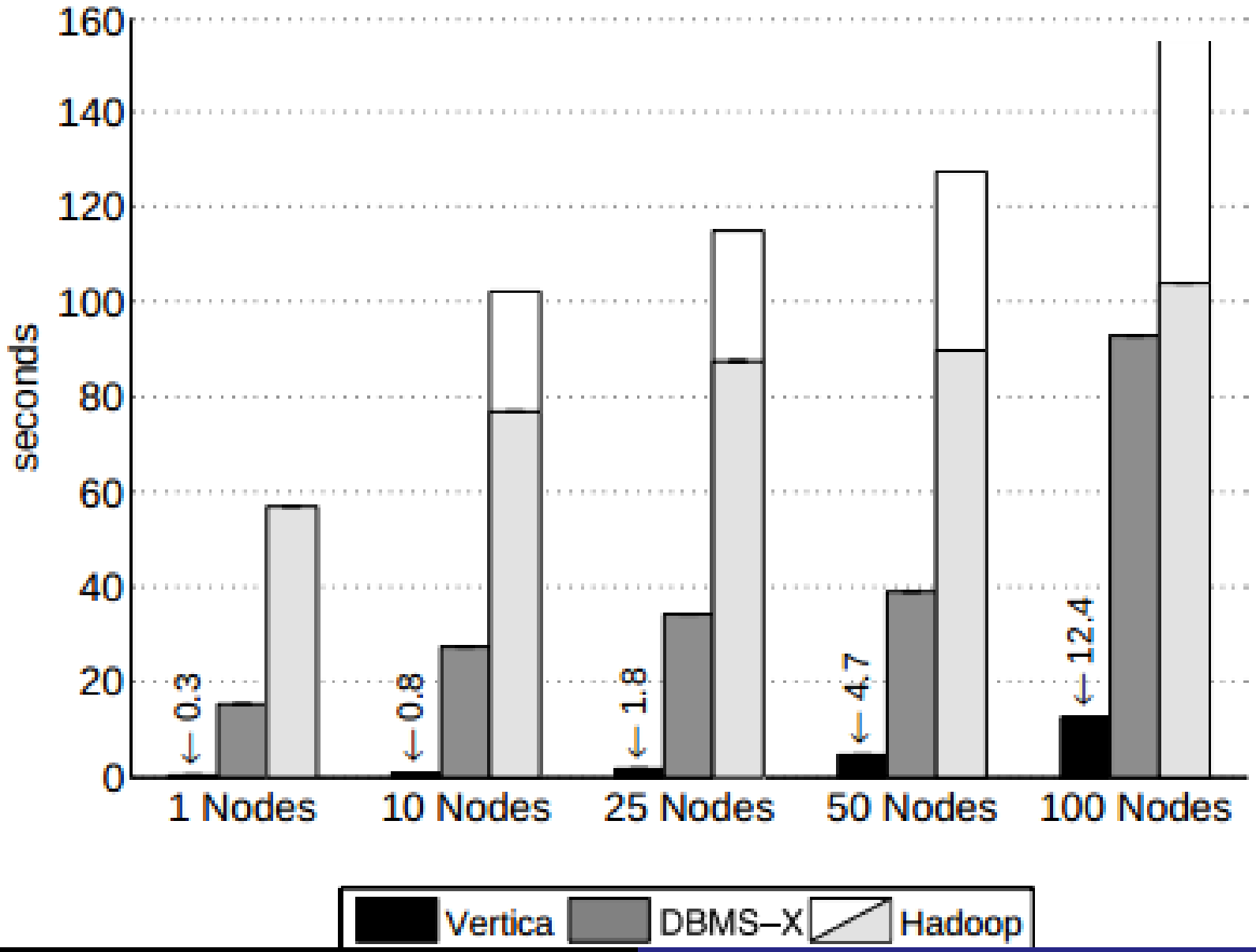


5: Grep Task Results – 1TB/cluster Data Set

Selection Task

1. 36.000 data records per file on each node.
2. Hadoop: Fisnishes so quickly that a torrent of controll messages increases the total execution time.
3. Again Hadoop is outperformed by the other two.

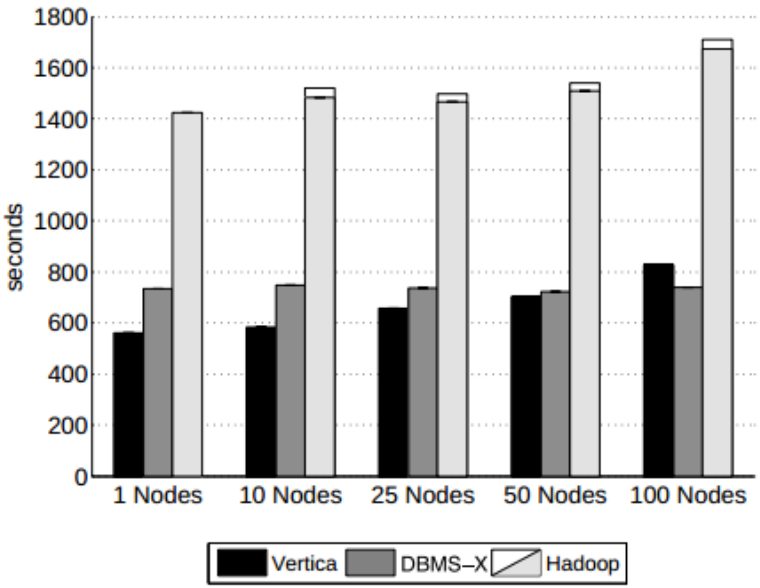
Selection Task Results



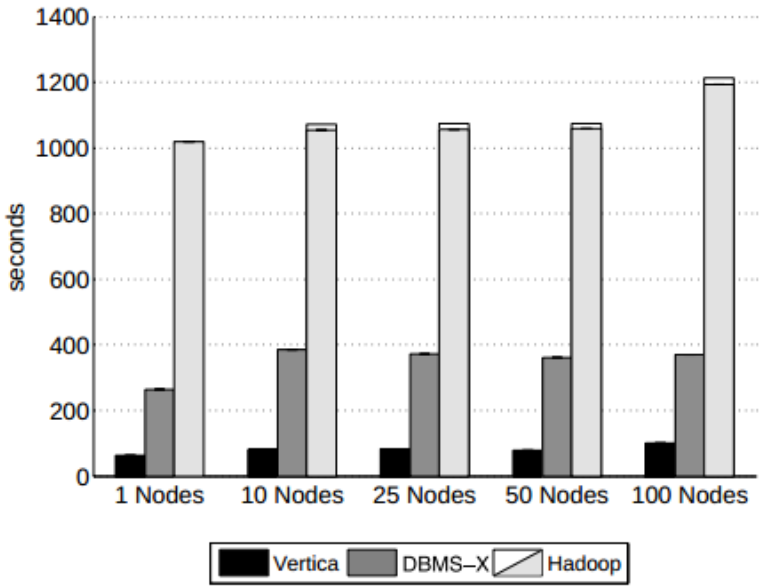
Aggregation Task

1. Task: calculate total revenue by IP.
2. Produces 2.5 million records(53MB) and 2.000 records(24KB).
3. Vertica slows down. But does not read unnecessary data columns.
4. Hadoop: finds all elements of correct type, then sums up the results.

Aggregation Task Results



7: Aggregation Task Results (2.5 million Groups)

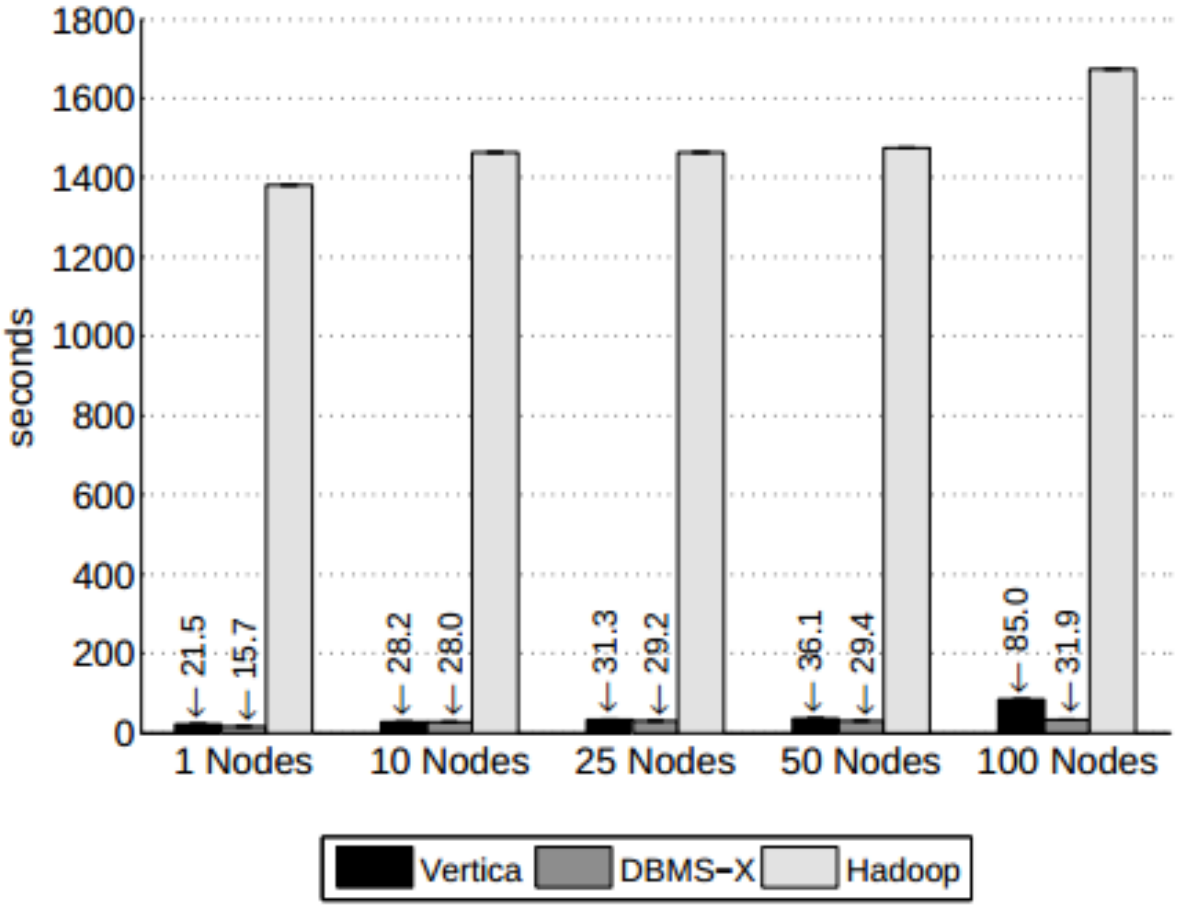


8: Aggregation Task Results (2,000 Groups)

Join Task

1. Task: Page rankings in a time period.
2. Complex MR program with three phases.
3. Reading and processing data is the most time consuming.

Join Task Results

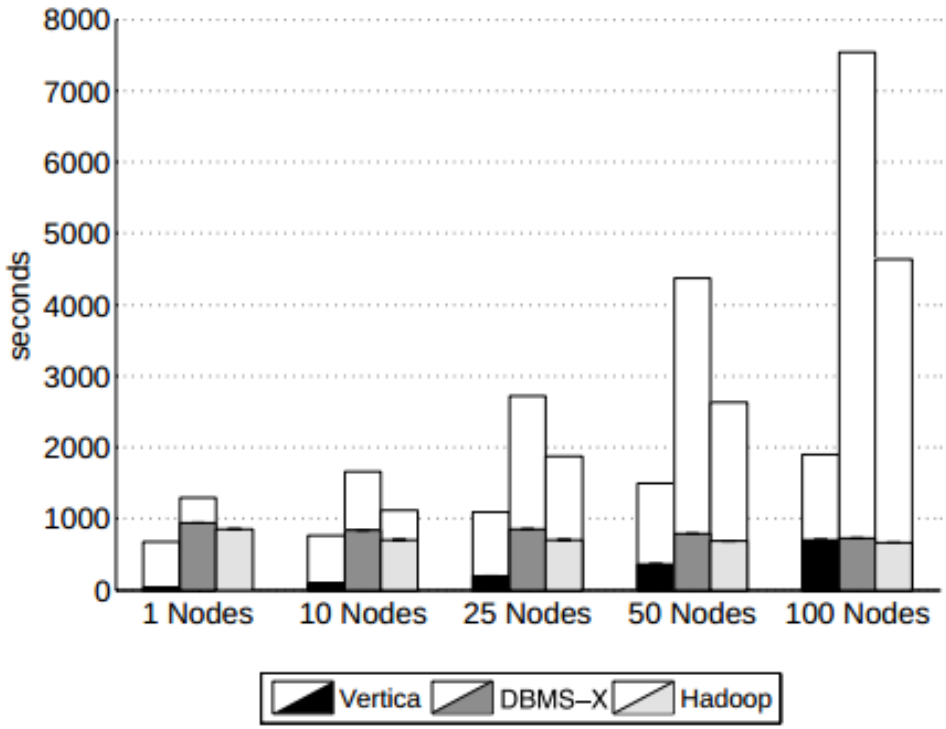


9: Join Task Results

UDF Aggregation task

1. Task: Counting links in documents.
2. DBMS-X and Hadoop has close to constant execution time.
3. Result writing gets slower with increased number of nodes.

UDF Aggregation Task Results



10: UDF Aggregation Task Results

Install

1. Hadoop: Easy install, trial and error optimization. Task tuning.
2. DBMS-X: Straight forward install. But the configuration proved difficult.
3. Vertica: Quite easy install. But too automated tuinig capabilities.

Task Startup

1. MR: 10 sec until the task is distributed. 25 sec for all nodes to start executing.
2. Hadoop reuse JVM reduced startup time by 10-15%
3. DBMS: startup time was one of the first things that was improved.
4. Resent improvements (article from 2009).

Compression

1. Both DBMS-X and Vertica worked better with data compression.
2. Hadoop worked better without compression.

Data Loading

1. Hadoop was the best system to load and read data.
2. Hadoop was more CPU intensive.
3. DBMSs can reorganize data on load.

Exectuion Strategies

1. Hadoops overhead messaging slowed it down.
2. DBMS data push strategy
3. DBMS query plan.

Failure Models

1. More HW = more failures.
2. MR is more tolerant to failure.
3. Sophisticated error recovery could improve performance.

Ease of Use

1. MR(Hadoop) was easier to get up and running. Simple structure. But algorithms have to be implemented.
2. DBMS: might be easier to maintain later. Less data enforecement to do.

Additional Tools

1. DBMS have a long history of development and have a lot of extrernal tools to use.
2. MR is still young so there is not to many tools available yet.

Conclusion and Thoughts

1. Small scale data analysis will work better with DBMSs.
2. Large scale data analysis today is way bigger then it was in 2009.
3. Hadoop and MR systems has room for improvement and will probably be improved over time.
4. Both architectures will probably remain, due to their different strenghts and areas of use.

Last slide,

Questions?