HadoopDB:
An Architectural Hybrid of MapReduce and DBMS Technologies for Analytical Workloads

Yang Zheng

- Background
- Motivation
- Details of the paper

Background

- Data Analysis
  - identify user preference
  - predict market trends
  - support decision making
- Example
- Wal-mart

Background

- Current Trend
  - High-end proprietary machines -> lower-end commodity hardware
  - Data need to be analyzed explodes

Background

- Two best-suited technologies
  - Parallel database
  - MapReduce-based system

Background

Parallel Database

- Advantages
  - Performance and efficiency
- Disadvantages
  - Poor scalability relative to MapReduce
    - assume failures are rare
    - assume homogeneity of machines
    - not tested at large scale

Background

MapReduce-based System

- Advantages
  - Fault tolerance
  - Scalability
- Disadvantages
  - Bad performance
    - not originally designed to perform structured data

Motivation

Four Desired Properties Of A Petabyte Scale System

- Performance
- Fault Tolerance
- Ability to run in a heterogeneous environment
- Flexible query interface

Motivation

Performance

- faster analysis
- cost saving

Fault Tolerance

- a query won't be restarted in the face of node failure
- low-end commodity machine
- increasing number of nodes

Motivation

Ability to run in a heterogeneous environment

- prevent the straggler to effect the query time

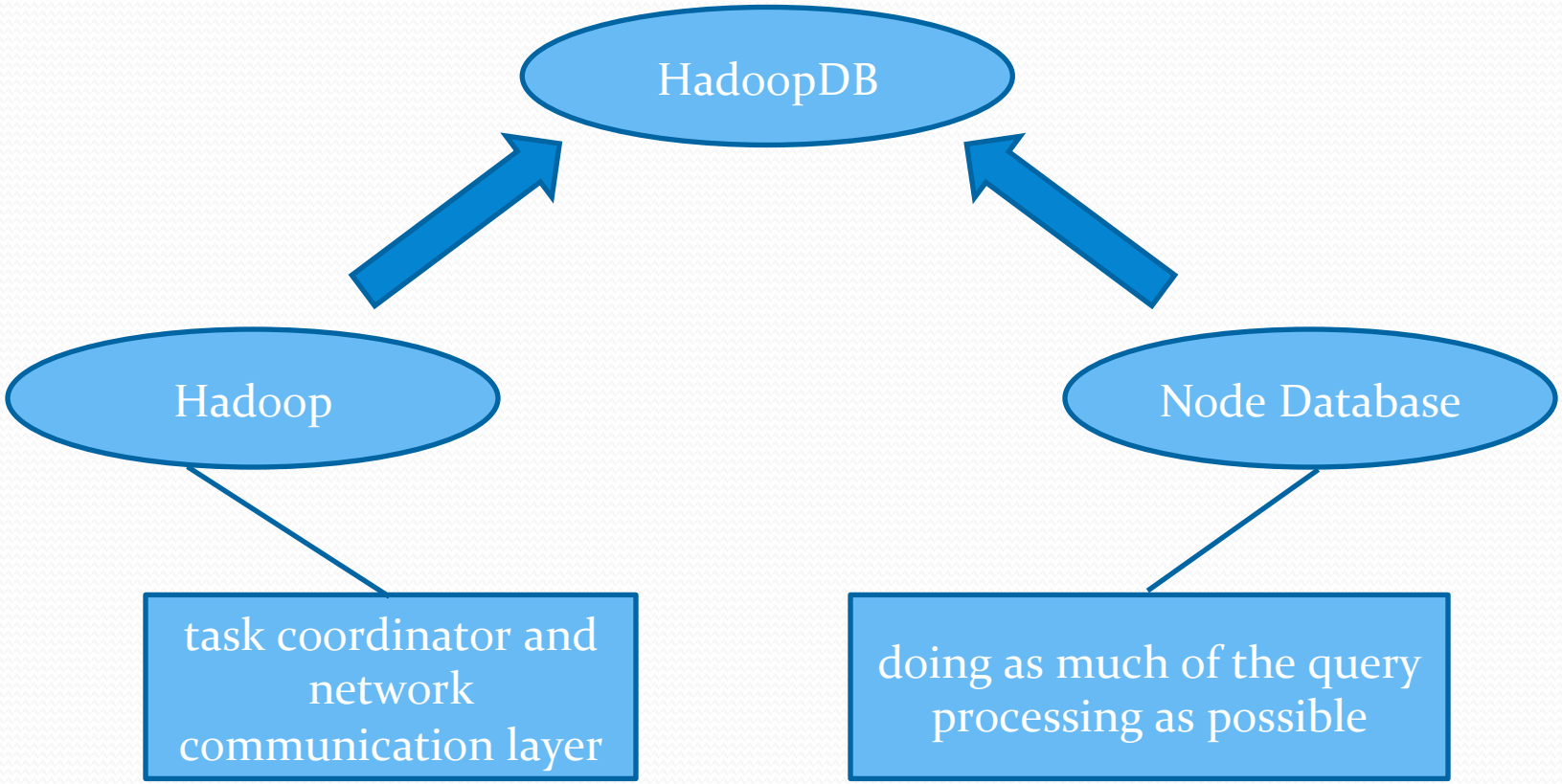
Flexible query interface

- business analyst better express their idea to the system

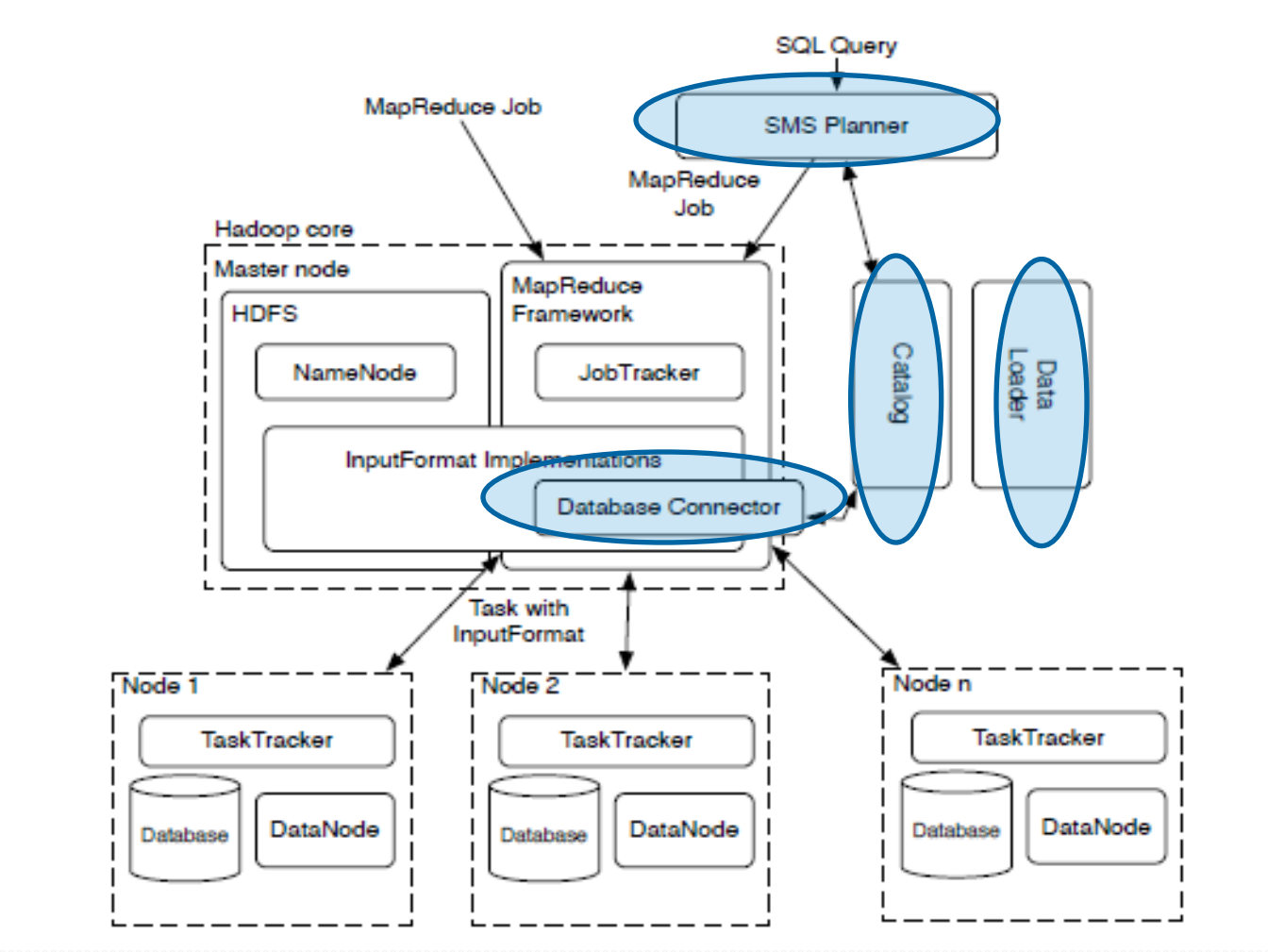
Motivation

	Performance	Fault Tolerance	Ability to run in a heterogeneous environment	Flexible query interface
Parallel Database	√			√
MapReduce-based System		√	√	√

HadoopDB


HadoopDB
Hadoop
Node Database
task coordinator and network communication layer
doing as much of the query processing as possible

HadoopDB


SQL Query
MapReduce Job
SMS Planner
MapReduce Job
Hadoop core
Master node
HDFS
NameNode
MapReduce Framework
JobTracker
InputFormat Implementations
Database Connector
Catalog
Data Loader
Task with InputFormat
Node 1
TaskTracker
Database
DataNode
Node 2
TaskTracker
Database
DataNode
Node n
TaskTracker
Database
DataNode

SQL to MapReduce to SQL (SMS) Planner

- transforms SQL queries into MapReduce jobs
- extends Hive

Hive - a high-level language that tramsforms HiveQL, a variant of SQL, into MapReduce jobs.

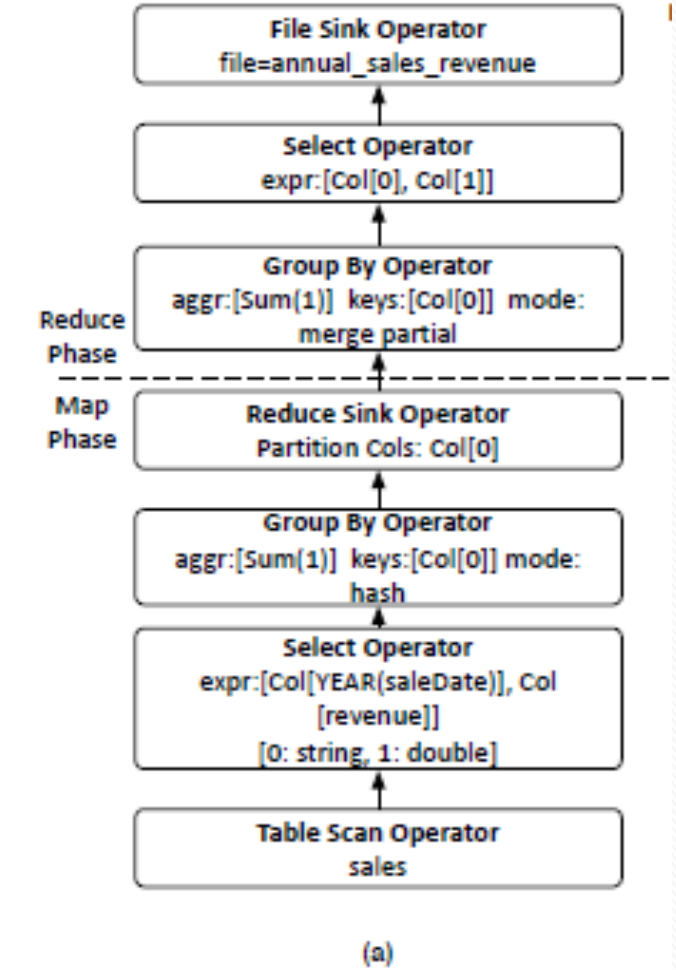
A simple example

```
SELECT YEAR(saleDate), SUM(revenue)
FROM sales GROUP BY YEAR(saleDate);
```

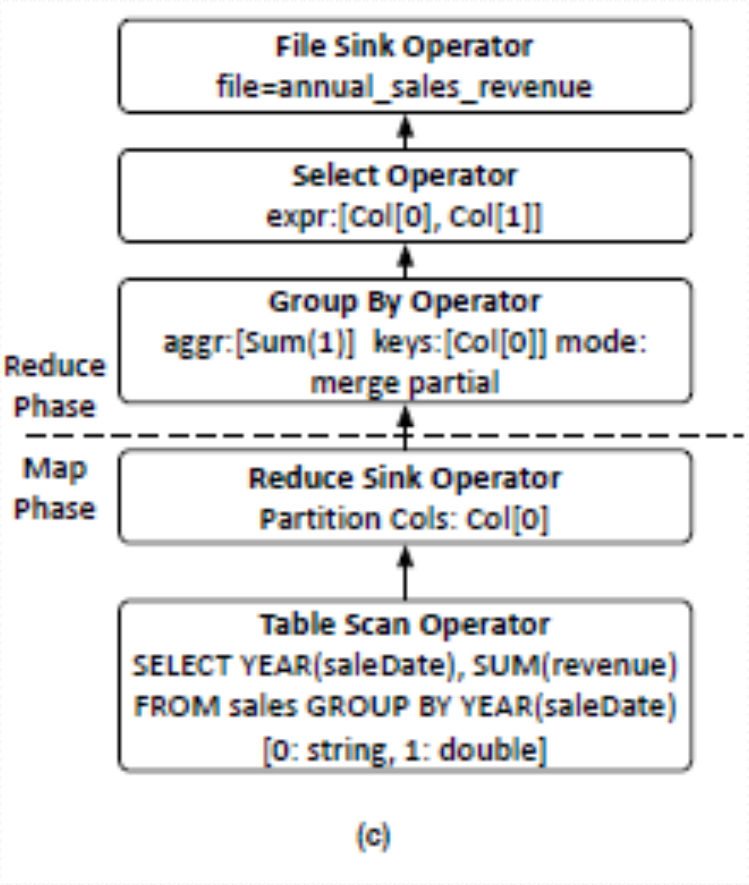
- parser transforms the query into an Abstract Syntax Tree
- Semantic Analyzer connects to Hive's internal catalog, to retrieve the schema of the sales table
- The logical plan generator then creates the query plan
- The optimizer restructures the query plan to a more optimized plan
- the physical plan generator converts the logical query plan into a physical plan
- The physical plan is then serialized into an XML plan, which is read by a Hadoop job

Physical Query Plans

Hive



SMS



Benchmark System (Components)

- Hadoop
- HadoopDB
  - use PostgreSQL as node databases
- Vertica
  - column-store parallel database
- DBMS X
  - parallel row-oriented database

Benchmark System (Tasks)

- Grep Task
- Selection
- Aggregation Task
- Join Task
- UDF Aggregation Task

Grep Task

- looking for a three characters pattern
- process unstructured data

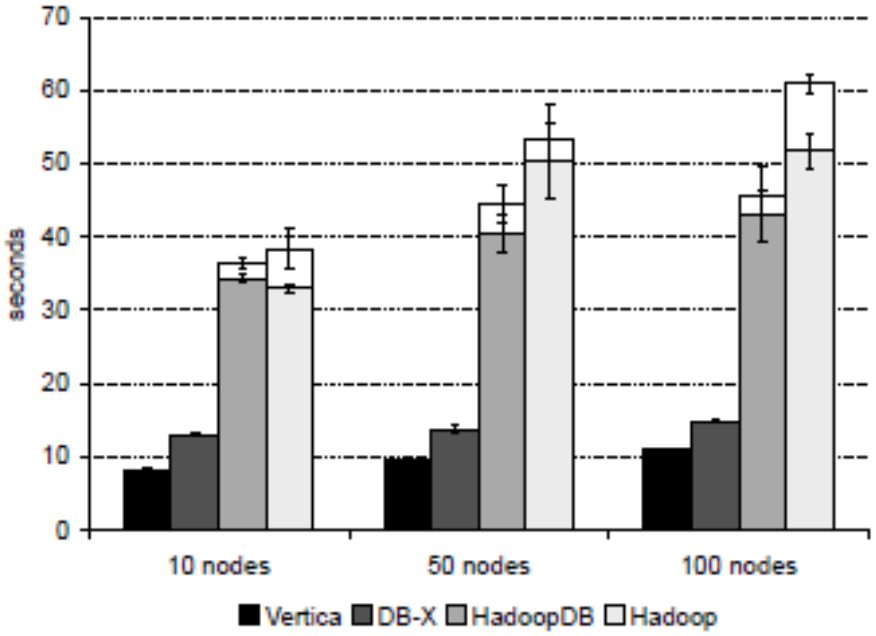


Figure 5: Grep Task

Aggregation Task

- intermediate results to be exchanged between different nodes

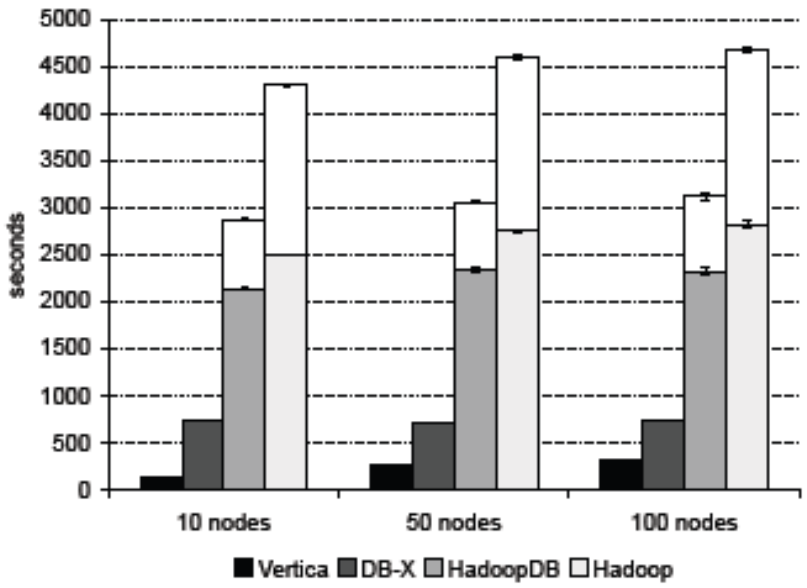


Figure 7: Large Aggregation Task

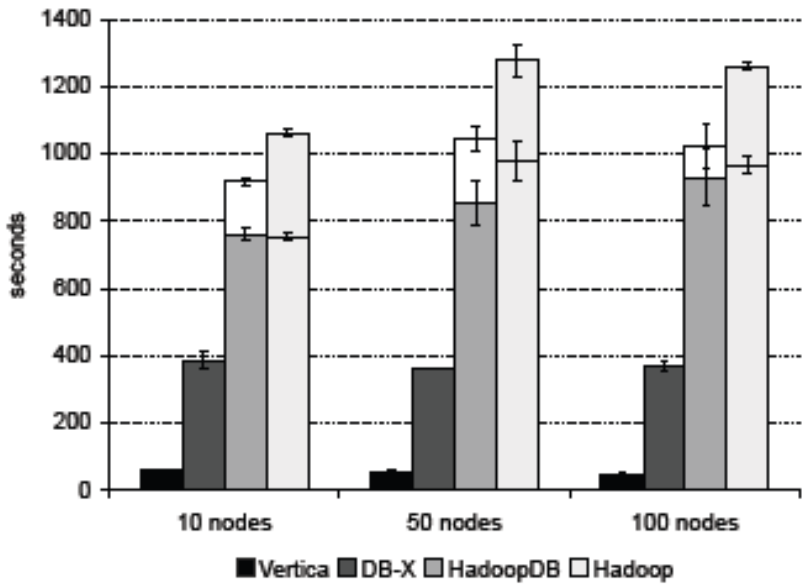


Figure 8: Small Aggregation Task

Join Task

- read in two different data sets and join them

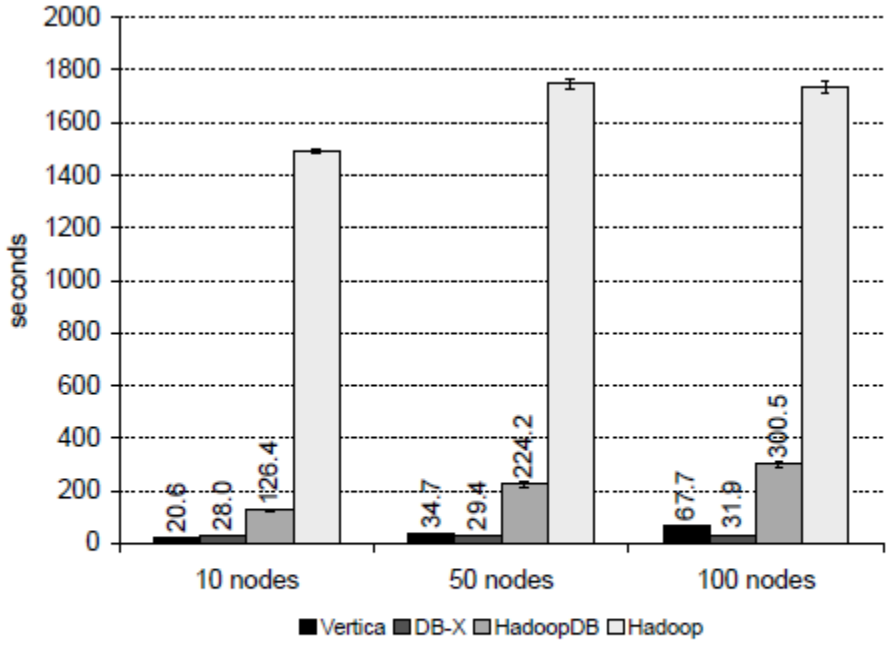
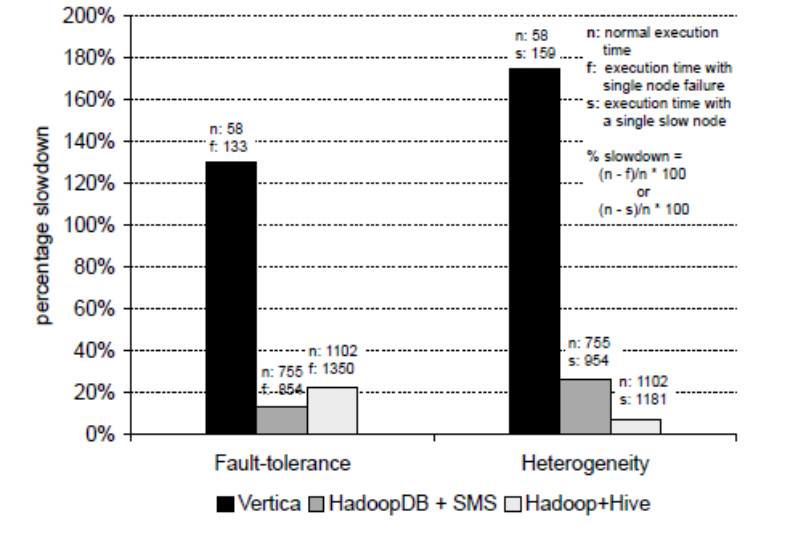


Figure 9: Join Task

FAULT TOLERANCE AND HETEROGENEOUS ENVIRONMENT

- Simulate node failure
  - terminated a node at 50% query completion
- to simulate heterogeneous environment
  - running an I/O-intensive background job



Summary

HadoopDB

- able to approach the performance of parallel database systems
- achieve similar scores on fault tolerance
- an ability to operate in heterogeneous environments