Борчук Леонид

Хранилища для аналитики – конвергенция Hadoop, СУБД In-memory и SQL

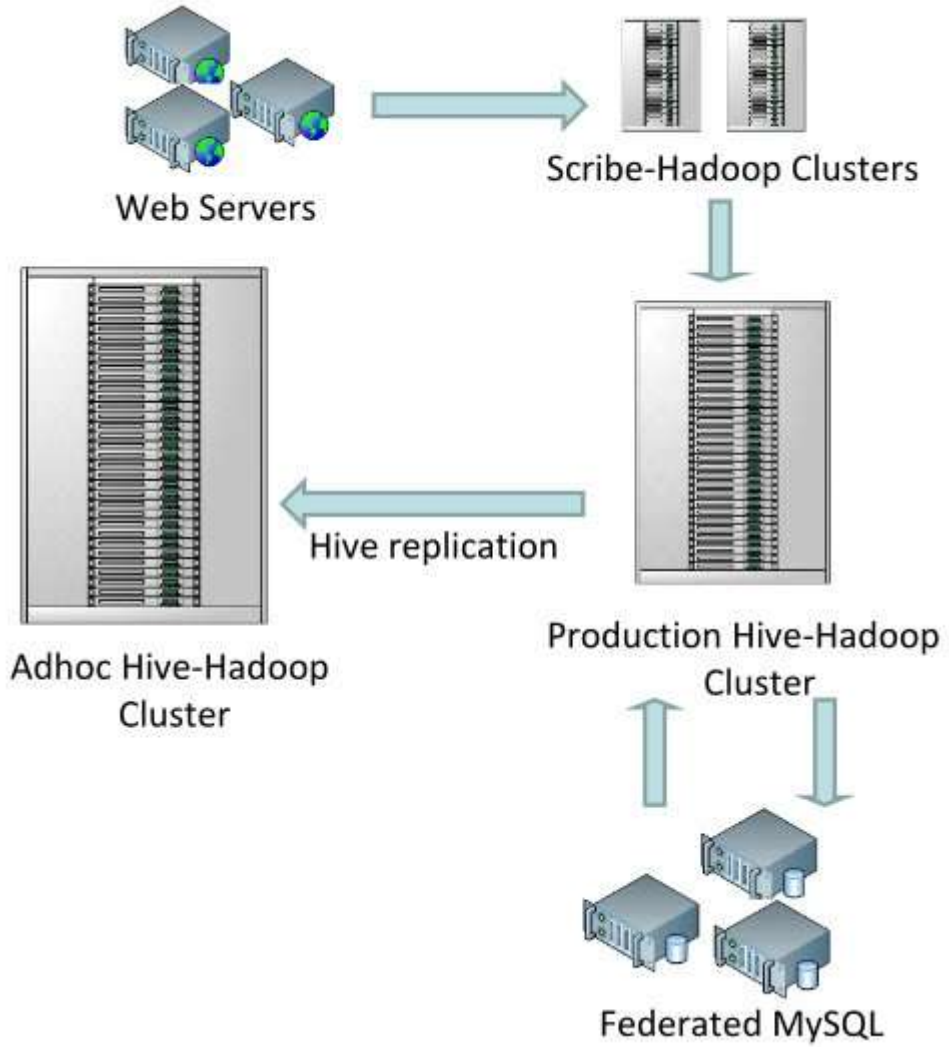
Проблема:

Предположим, у нас есть 2 Пб данных. Пусть это будет ClickStream (просто пример). Как получить по ним аналитику?

Проблема состоит в оперативности получения данных.

Варианты решения

2011, [1] Data warehousing and analytics infrastructure at facebook



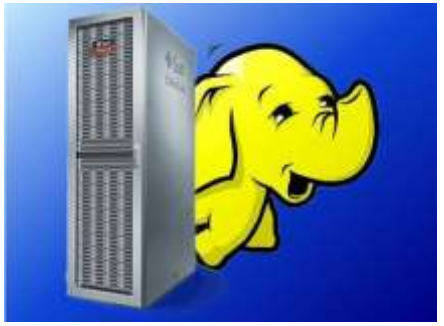
Варианты решения

Welcoming Facebook to the growing family of HP Vertica customers!

Facebook selected the HP Vertica Analytics Platform as one component of its big data infrastructure. Vertica's value to Facebook can be found in its ability to provide business insights with incredible speed and flexibility.

Коммерческие решения

cloudera hadoop



Hortonworks

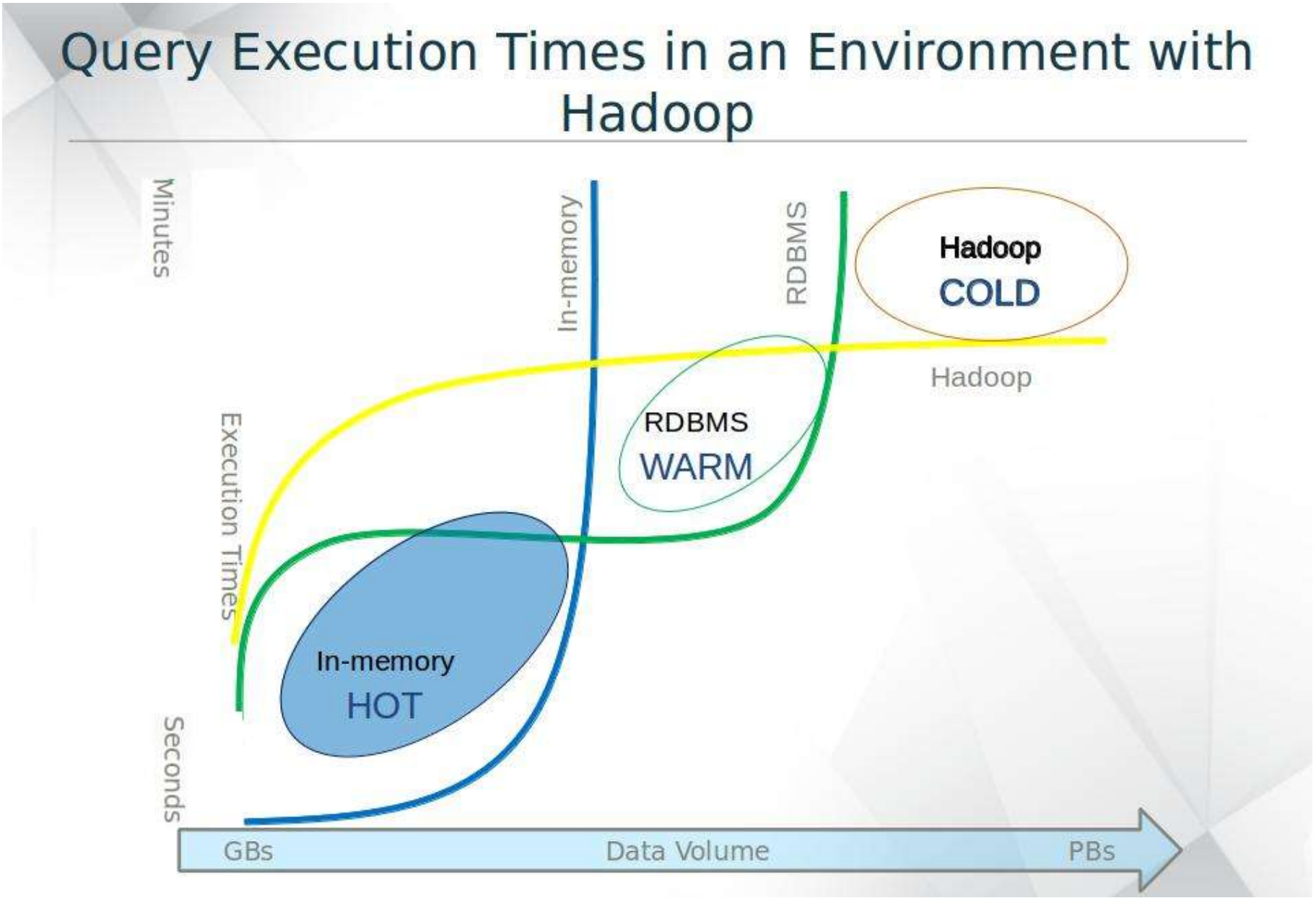
VERTICA

MAPR

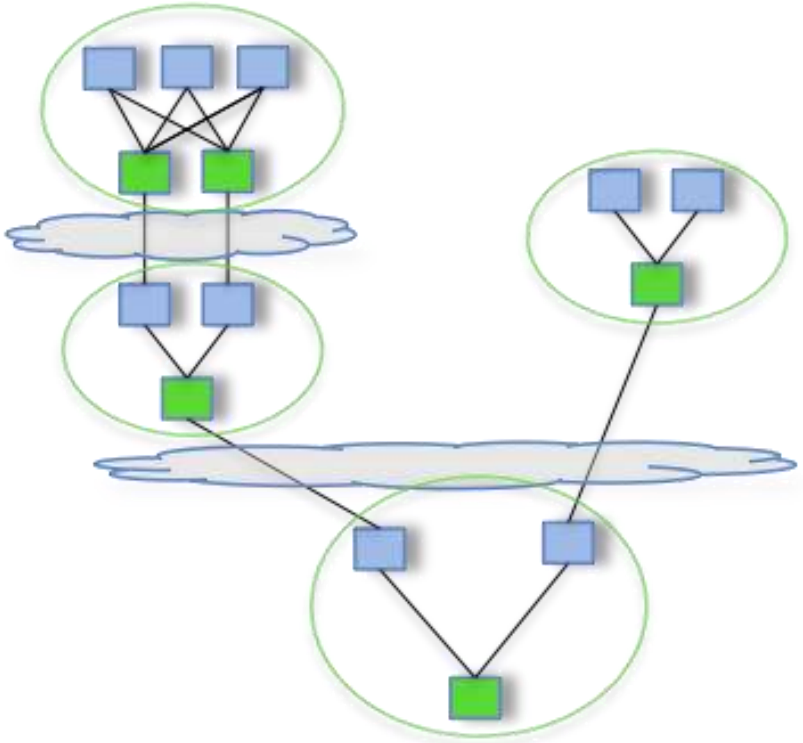
Greenplum

TERADATA

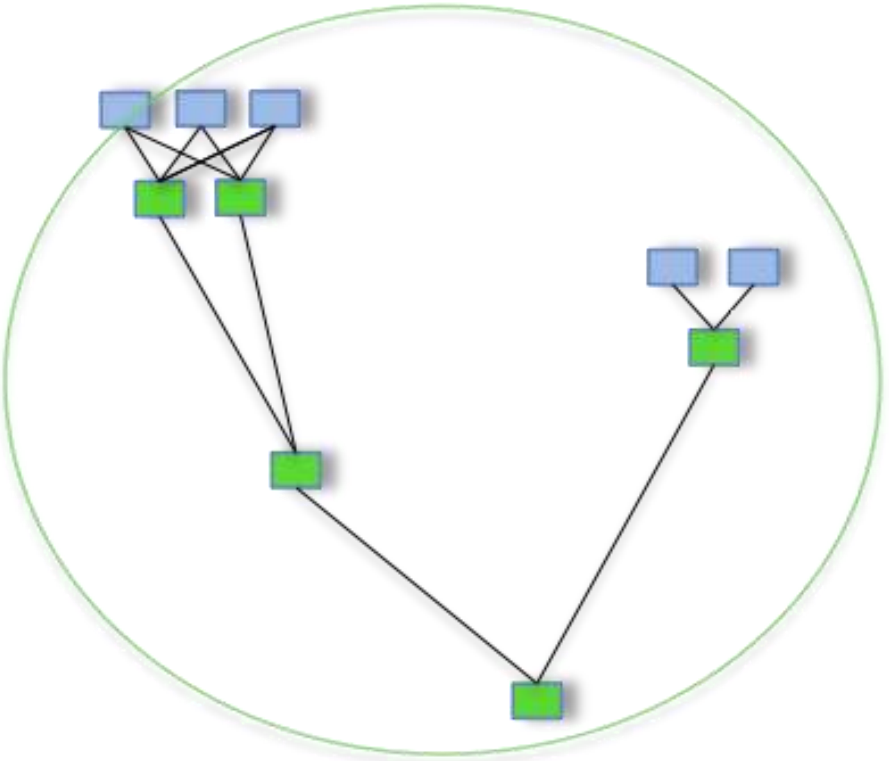
Hadoop?


Query Execution Times in an Environment with Hadoop
Minutes
Execution Times
Seconds
In-memory
RDBMS
Hadoop
COLD
Hadoop
RDBMS
WARM
In-memory
HOT
GBs
Data Volume
PBs

Map/Reduce → in-memory graph



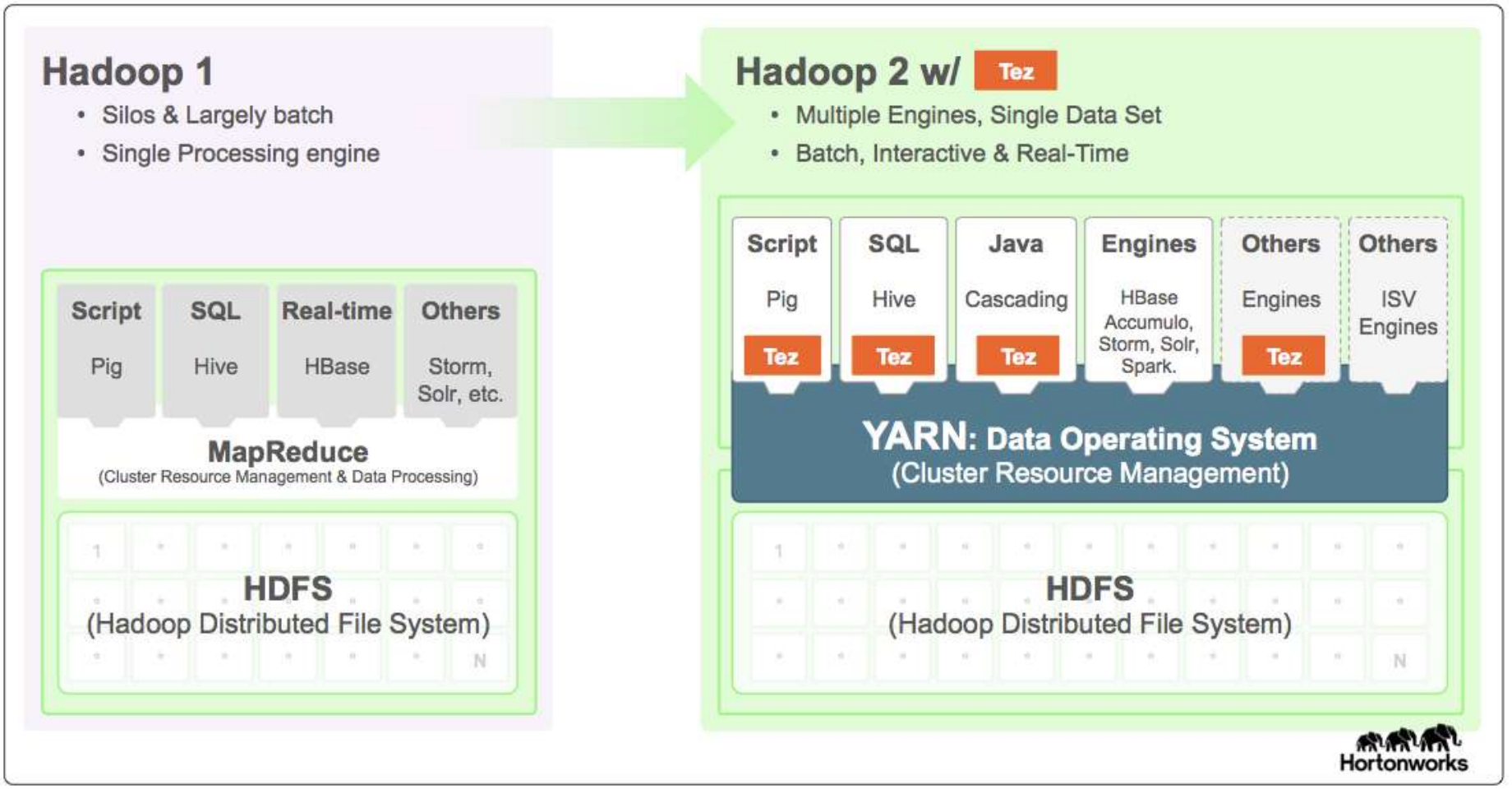
Pig/Hive - MR



Pig/Hive - Tez

Hadoop engines

Spark, Tez, Impala



Сравнение движков

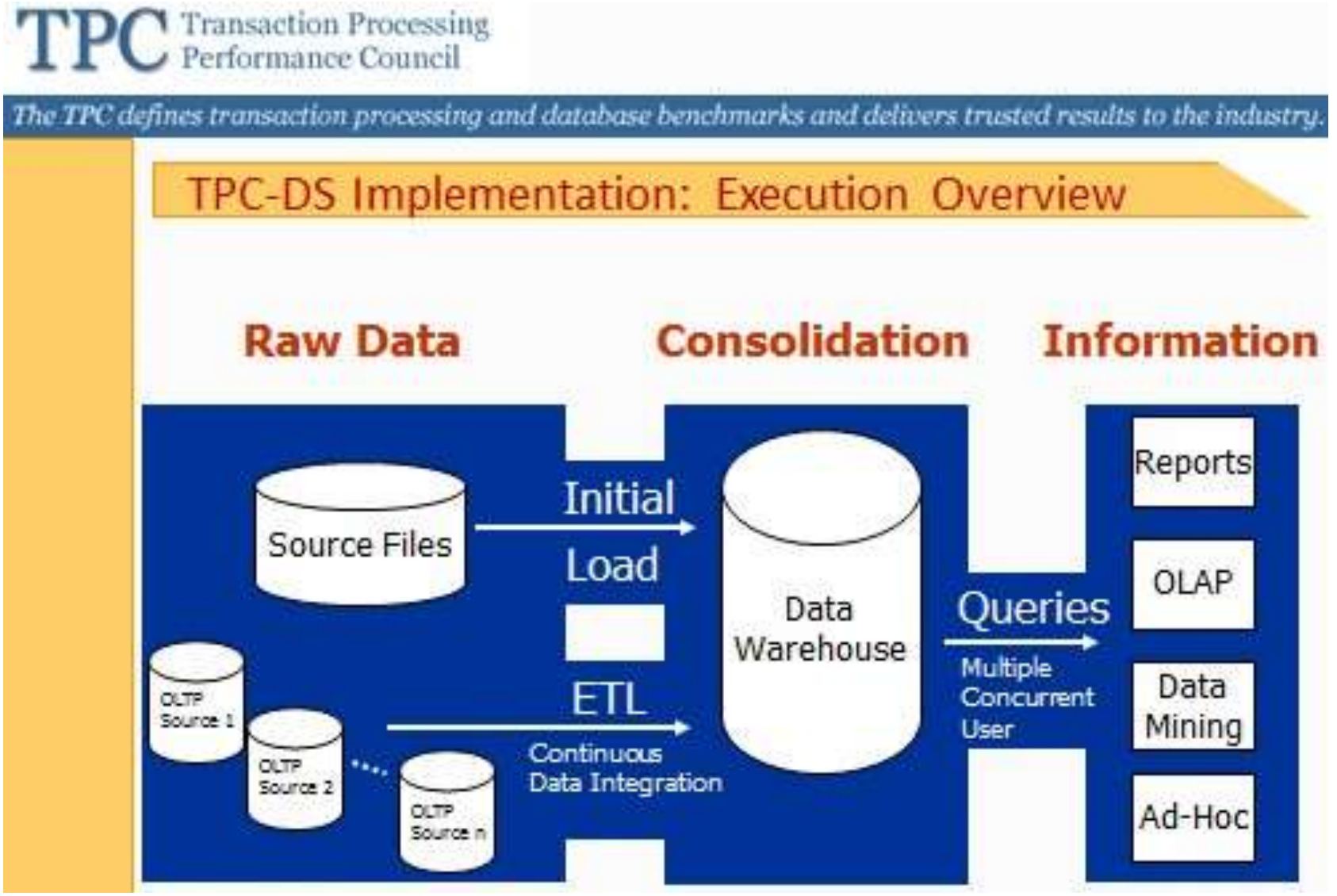
New SQL Benchmarks: Apache Impala (incubating) Uniquely Delivers Analytic Database Performance

February 11, 2016 | By Devadutta Ghat, David Rorke, and Dileep Kumar | 42 Comments

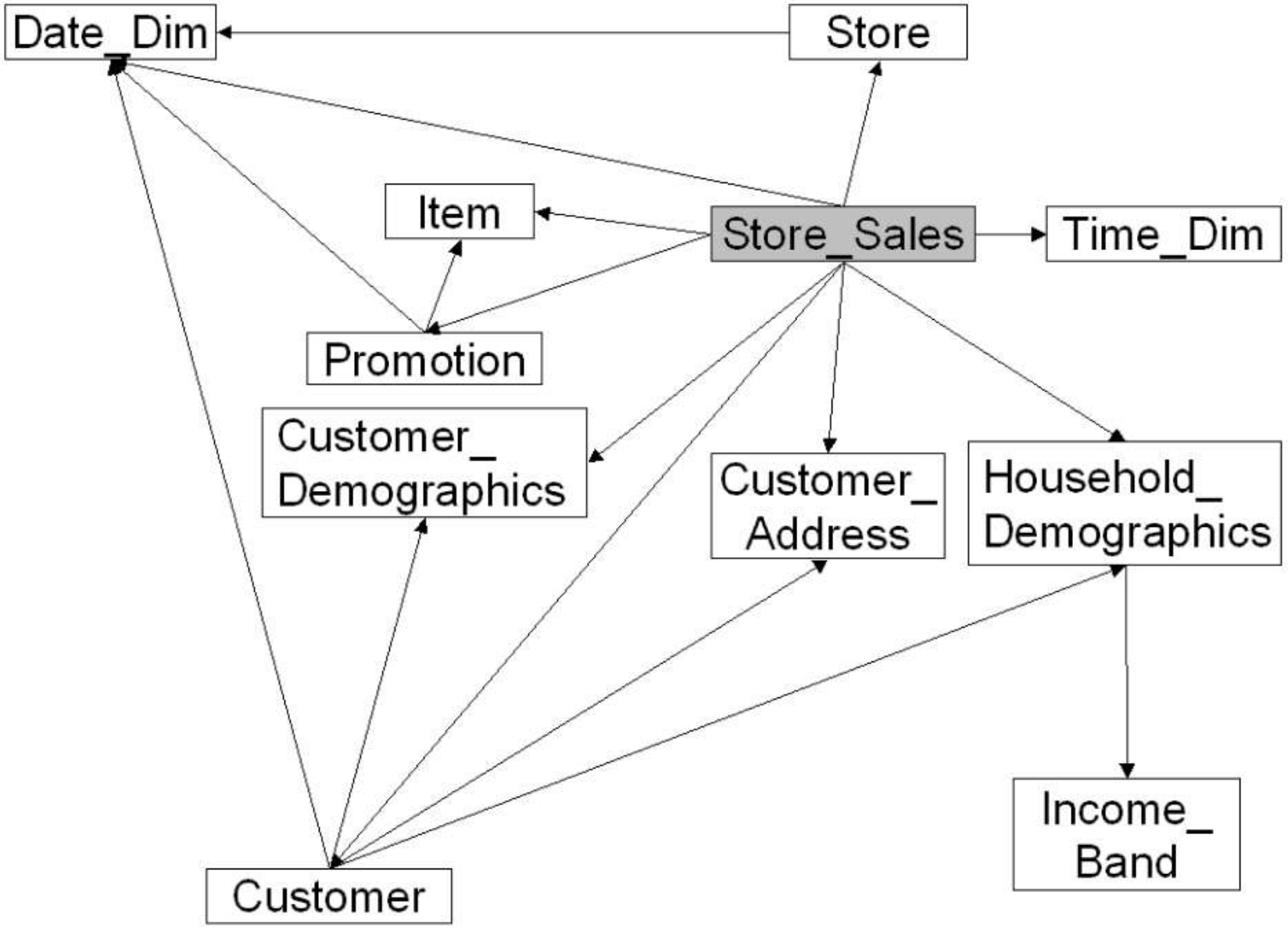
Categories: Hive | Impala | Performance | Spark

New testing results show a significant difference between the analytic database performance of Impala compared to batch and procedural development engines, as well as Impala running all 99 TPC-DS-derived queries in the benchmark workload.

TPC-DS


TPC Transaction Processing Performance Council
The TPC defines transaction processing and database benchmarks and delivers trusted results to the industry.
TPC-DS Implementation: Execution Overview
Raw Data
Consolidation
Information
Source Files
Initial
Load
ETL
Continuous
Data Integration
OLTP
Source 1
OLTP
Source 2
OLTP
Source n
Data
Warehouse
Queries
Multiple
Concurrent
User
Reports
OLAP
Data
Mining
Ad-Hoc

TPC-DS схема данных



Интерактивные запросы

B.55

For a given year, month and store manager calculate the total store sales of any combination all brands.

Qualification Substitution Parameters:
MANAGER.01 = 28
MONTH.01 = 11
YEAR.01 = 1999

Отчетные запросы

B.7

Compute the average quantity, list price, discount, and sales price for promotional items sold in stores where the promotion is not offered by mail or a special event. Restrict the results to a specific gender, marital and educational status.

Qualification Substitution Parameters:
YEAR.01=2000
ES.01=College
MS.01=S
GEN.01=M

Аналитические запросы

B.79

Compute the per customer coupon amount and net profit of Monday shoppers. Only purchases of three consecutive years made on Mondays in large stores by customers with a certain dependent count and with a large vehicle count are considered.

Qualification Substitution Parameters:
VEHCNT.01 = 2
YEAR.01 = 1999
DEPCNT.01 = 6

Прочие запросы

B.82
Find customers who tend to spend more money (net-paid) on-line than in stores.

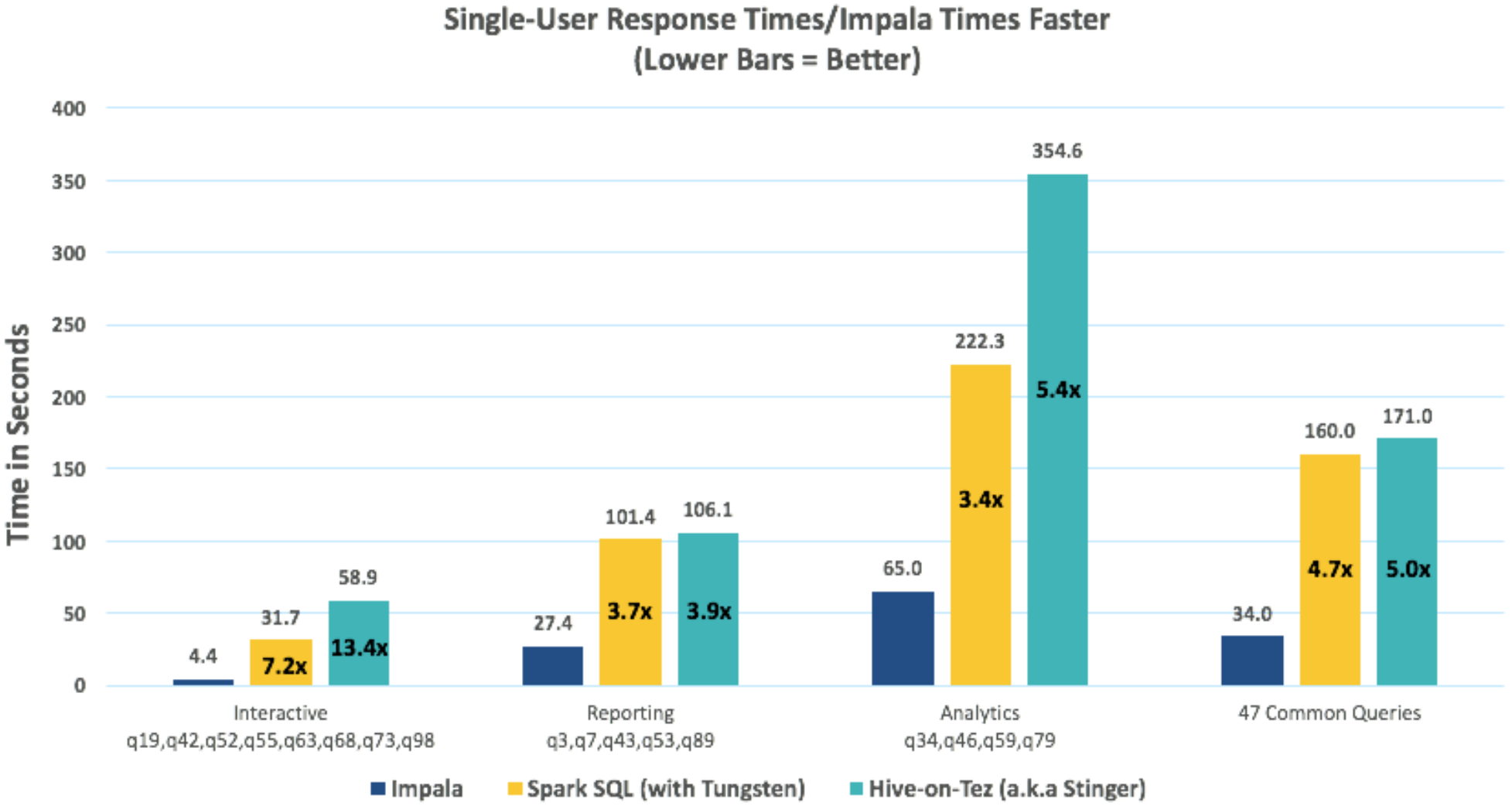
Qualification Substitution Parameters

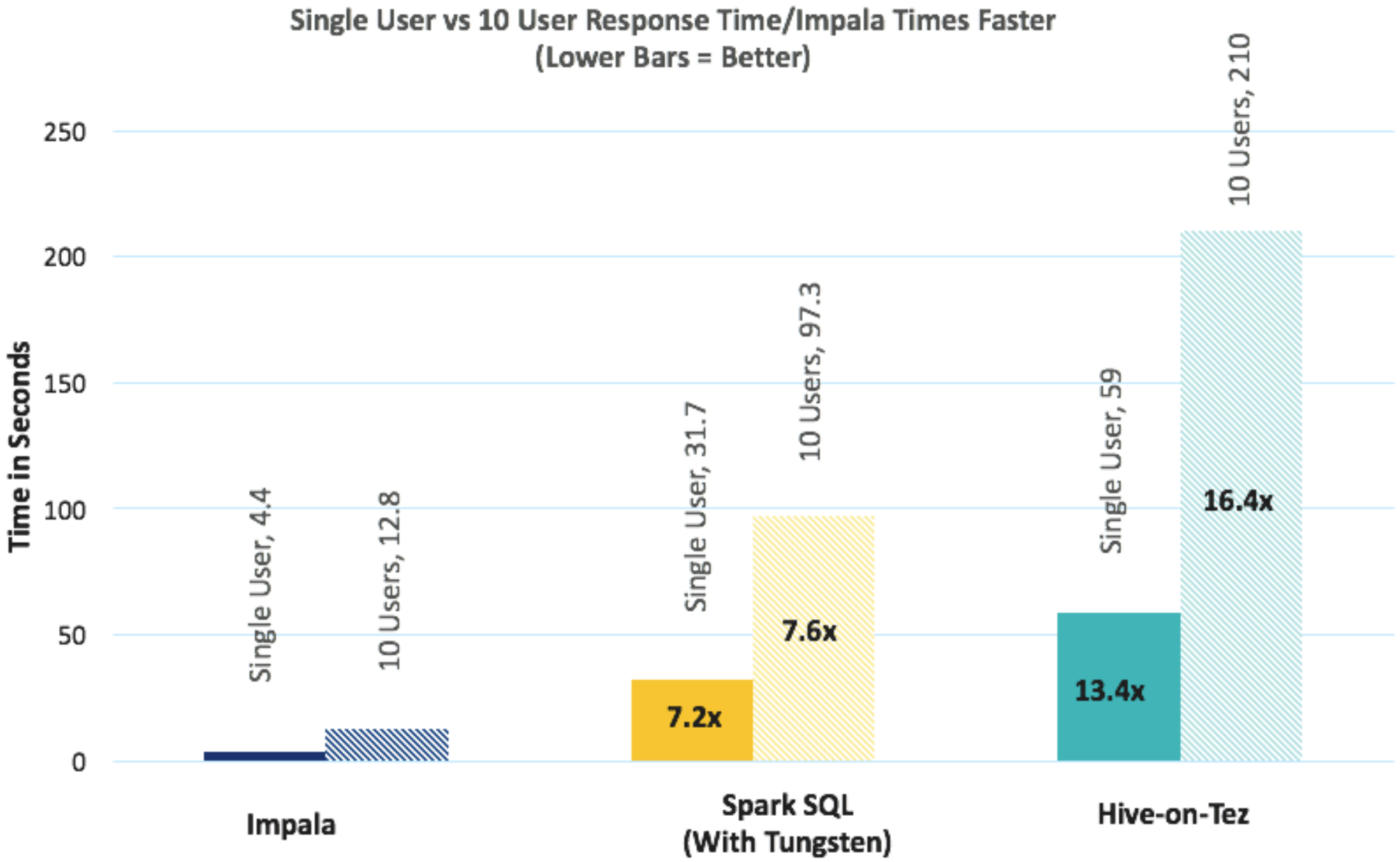
MANUFACT_ID.01 = 129
MANUFACT_ID.02 = 270
MANUFACT_ID.03 = 821
MANUFACT_ID.04 = 423
INVDATE.01 = 2000-05-25
PRICE.01 = 62

Тестируемая конфигурация

БД размером 15 Тб.
Кластер из 21 ноды:
CPU: 2 sockets, 12 total cores, Xeon E5-2630
12 disk drives at 1 Tb
384 GB memory

Сравнивались:
1. Impala 2.3
2. Hive-on-Tez: Hive 2.0 on Tez 0.5.2
3. Spark SQL 1.5 with Tungsten


Single-User Response Times/Impala Times Faster
(Lower Bars = Better)
Time in Seconds
400
350
300
250
200
150
100
50
0
4.4
31.7
7.2x
58.9
13.4x
27.4
101.4
3.7x
106.1
3.9x
65.0
222.3
3.4x
354.6
5.4x
34.0
160.0
4.7x
171.0
5.0x
Interactive
q19,q42,q52,q55,q63,q68,q73,q98
Reporting
q3,q7,q43,q53,q89
Analytics
q34,q46,q59,q79
47 Common Queries
Impala
Spark SQL (with Tungsten)
Hive-on-Tez (a.k.a Stinger)


Single User vs 10 User Response Time/Impala Times Faster
(Lower Bars = Better)
Time in Seconds
250
200
150
100
50
0
Single User, 4.4
10 Users, 12.8
Single User, 31.7
10 Users, 97.3
Single User, 59
10 Users, 210
7.6x
16.4x
7.2x
13.4x
Impala
Spark SQL
(With Tungsten)
Hive-on-Tez

Spark Tungsteen

Impala - «магия» в LLVM

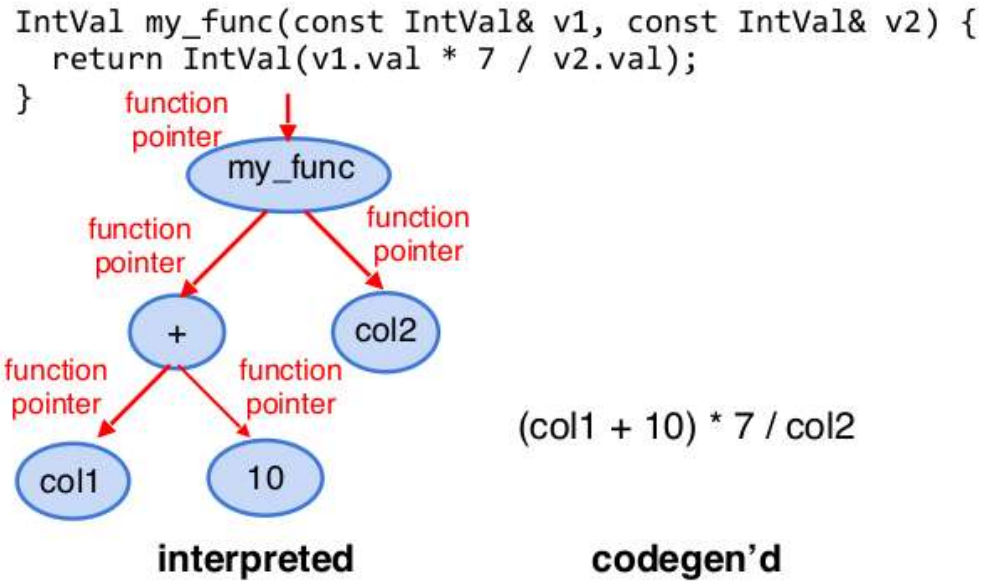
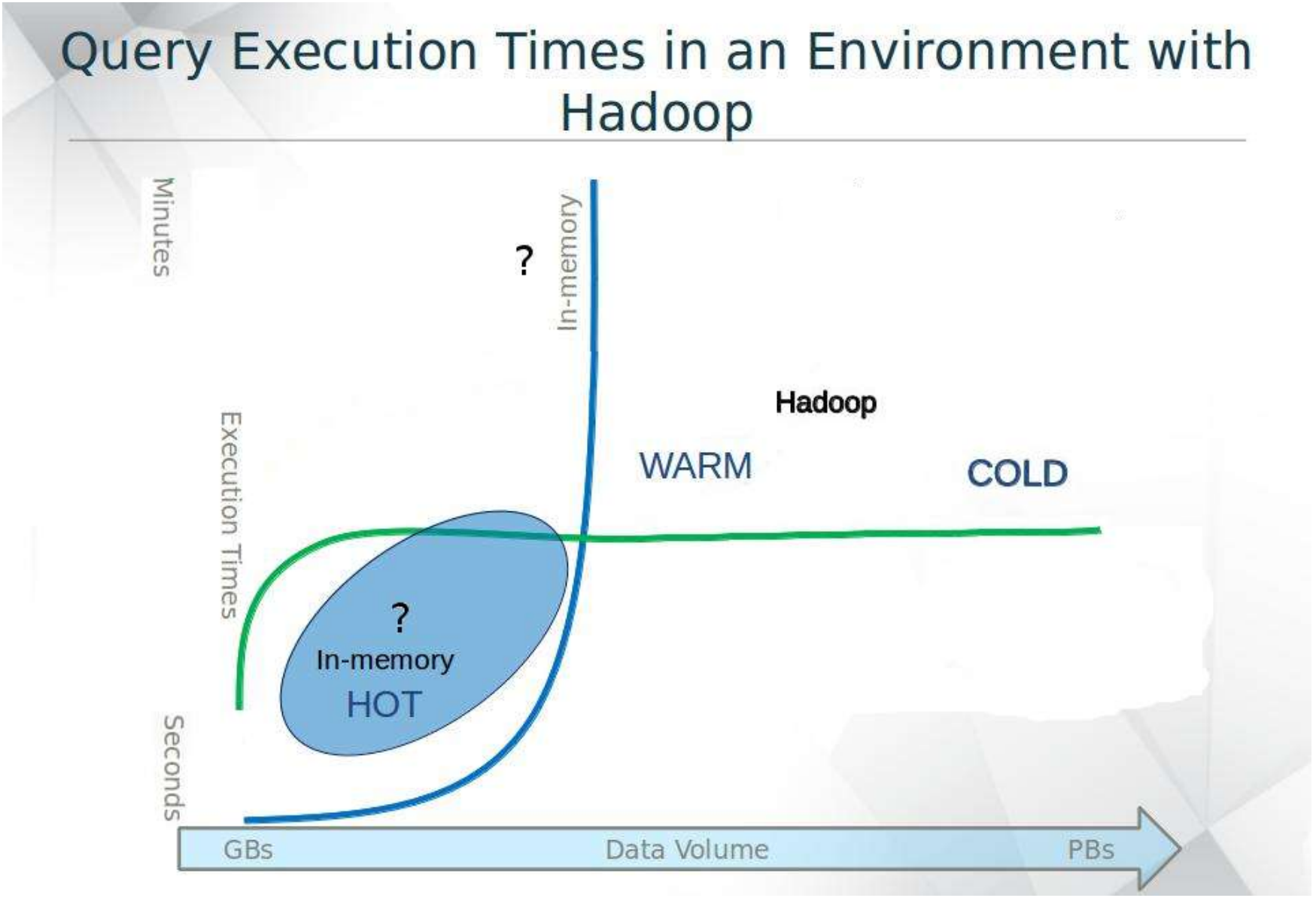


Figure 3: Interpreted vs codegen'ed code in Impala.

Project Tungsten: Bringing Apache Spark Closer to Bare Metal:

1. *Memory Management and Binary Processing*
2. *Cache-aware computation*
3. *Code generation*

Hadoop!


Query Execution Times in an Environment with Hadoop
Minutes
Execution Times
Seconds
?
In-memory
Hadoop
WARM
COLD
?
In-memory
HOT
GBs
Data Volume
PBs

Вопросы?

Борчук Леонид

Ссылки

[1] Ashish Thusoo, Zheng Shao, Suresh Anthony, Dhruba Borthakur, Namit Jain, Joydeep Sen Sarma, Raghotham Murthy, Hao Liu: Data warehousing and analytics infrastructure at facebook, SIGMOD '10 Proceedings of the 2010 ACM SIGMOD International Conference on Management of data
[2] Welcoming Facebook to the growing family of HP Vertica customers! https://community.hpe.com/t5/Big-Data/Welcoming-Facebook-to-the-growing-family-of-HP-Vertica-customers/ba-p/6304063#.WC2y8bXqVZ4
[3] Facebook's CIO Tim Campos at MicroStrategy World 2014 http://www.youtube.com/watch?v=2FWKfO14cGU
[4] . Pavlo et al. A comparison of approaches to large-scale data analysis. SIGMOD, 2009 http://database.cs.brown.edu/papers/benchmarks-sigmod09.pdf
[4] New SQL Benchmarks: Apache Impala (incubating) Uniquely Delivers Analytic Database Performance http://blog.cloudera.com/blog/2016/02/new-sql-benchmarks-apache-impala-incubating-2-3-uniquely-delivers-analytic-database-performance/
[5] Viktor Leis et all How Good Are Query Optimizers, Really? http://www.vldb.org/pvldb/vol9/p204-leis.pdf
[6] Marcel Kornacker et all Impala: A Modern, Open-Source SQL Engine for Hadoop http://cidrdb.org/cidr2015/Papers/CIDR15_Paper28.pdf