facebook

for the (data) masses

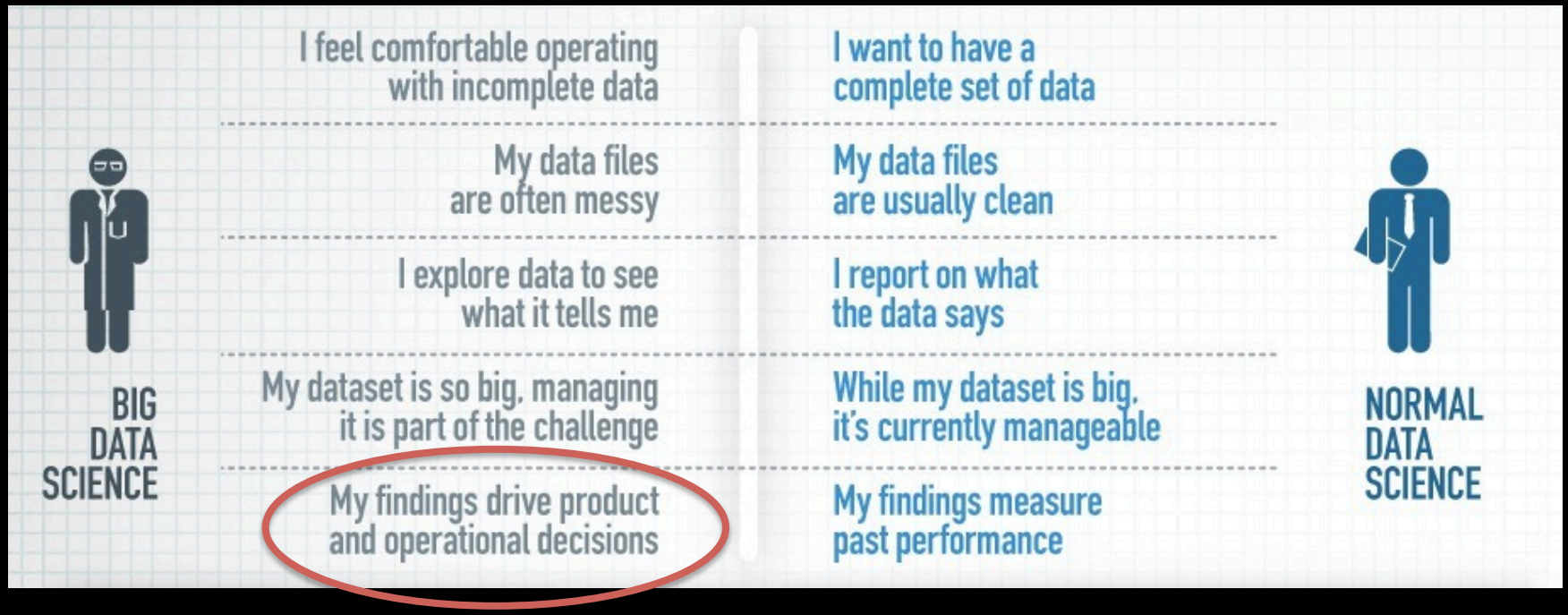
```
            -> order by 2 desc
            -> limit 5;
 n_name   | customers
----------+-----------
 VIETNAM  |   6003717
 MOROCCO  |   6003115
 ROMANIA  |   6002183
 CHINA    |   6001991
 IRAN     |   6001889
(5 rows)

Query 20131105_005539_00082_ee7y3, FINISHED, 13 nodes
Splits: 187 total, 187 done (100.00%)
0:07 [150M rows, 22.4GB] [22.2M rows/s, 3.32GB/s]

presto:default> exit
```


Lukas Gotter, M.Sc. | lukas.gotter@codecentric.de

Presto

„an interactive SQL-on-hadoop engine“

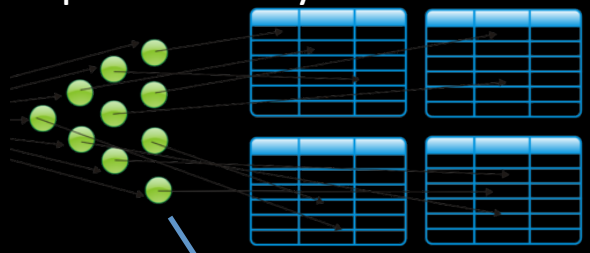
Data scientist‘s mindset



EMC Data Scientist Study, 2011

How we manage and analyze data today

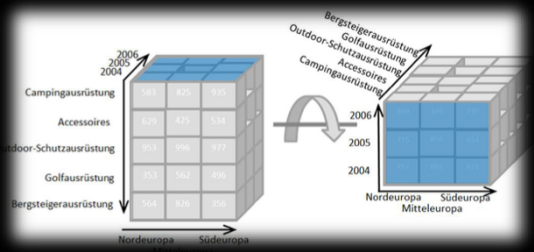
Operational Systems



Milliseconds

By David Rodma (Own work) [Public domain], via Wikimedia Commons

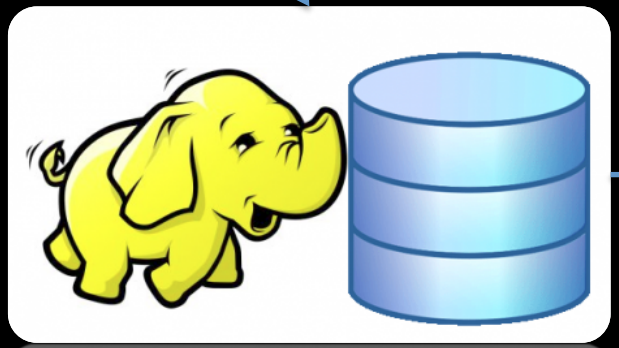
Data warehouse



Seconds

By Infopedian (Own work) [CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons

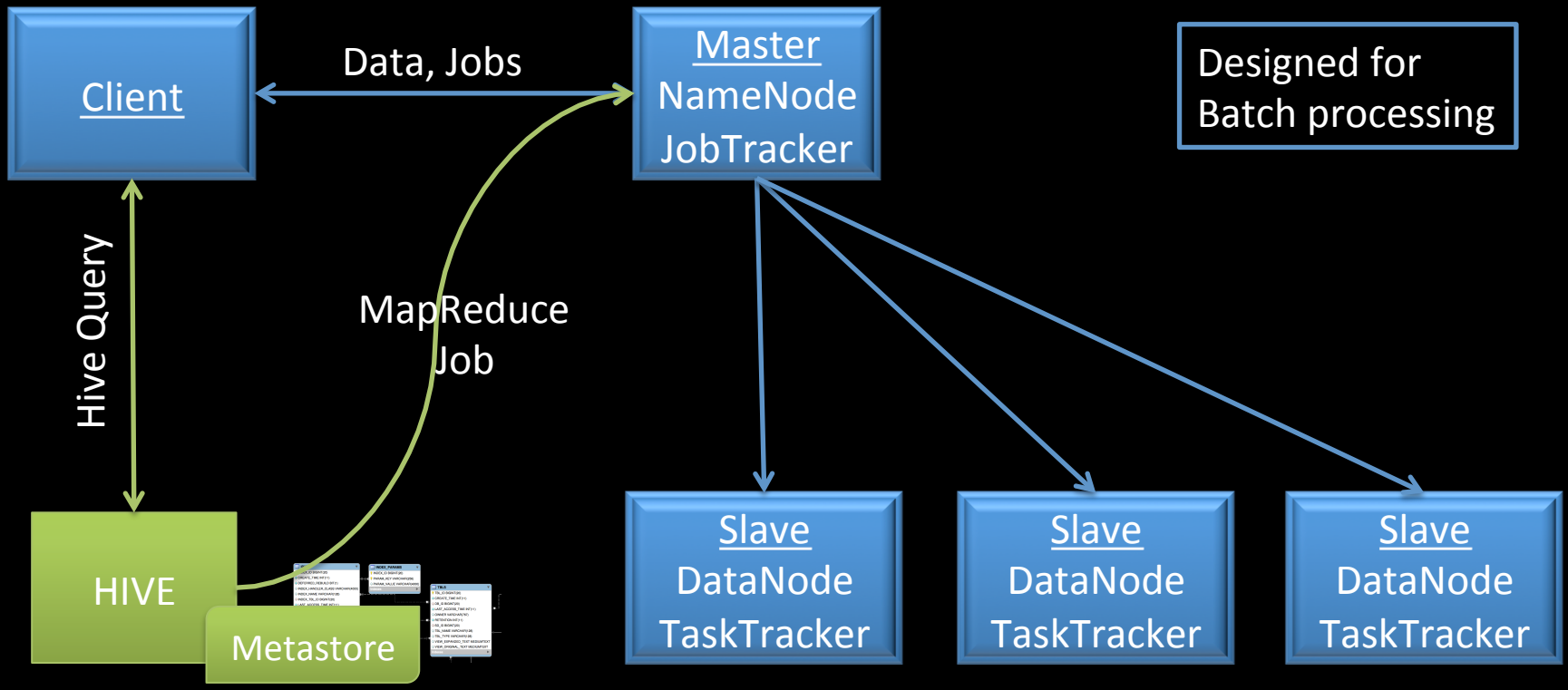
Minutes - hours



Big Data

Presto – Interactive sql-on-hadoop Engine

Simple Hadoop Cluster



„We are x times faster than Hive“- Engines

„10x faster!“

Hortonworks Stinger

„35x-45x faster. Goal: 100x!“

cloudera IMPALA

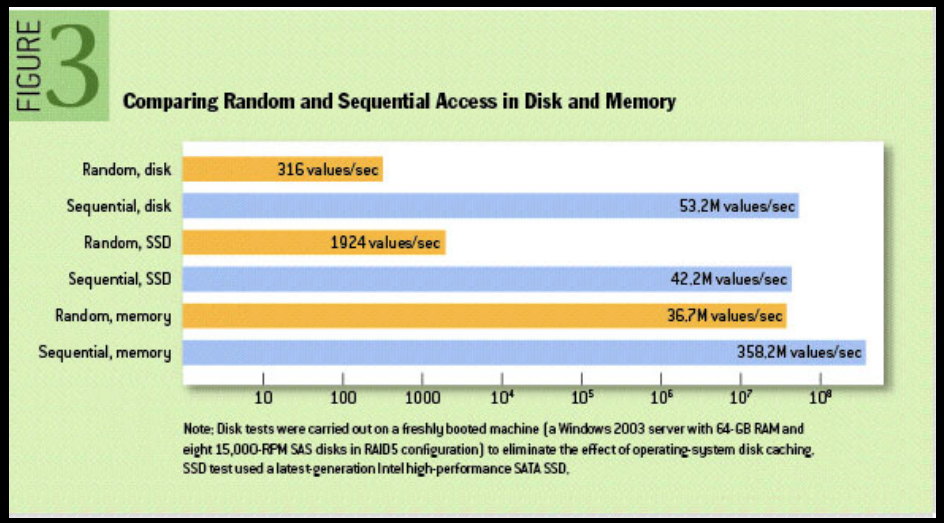
„10x-100x faster!“

„10x - 100x faster!“

...

Where to improve performance?

http://blog.scoutapp.com/articles/2011/02/08/how-much-slower-is-disk-vs-ram-latency



http://queue.acm.org/detail.cfm?id=1563874

- Disk Vs. Memory
- In-Memory Data
- In-Memory Processing
- Optimize Code

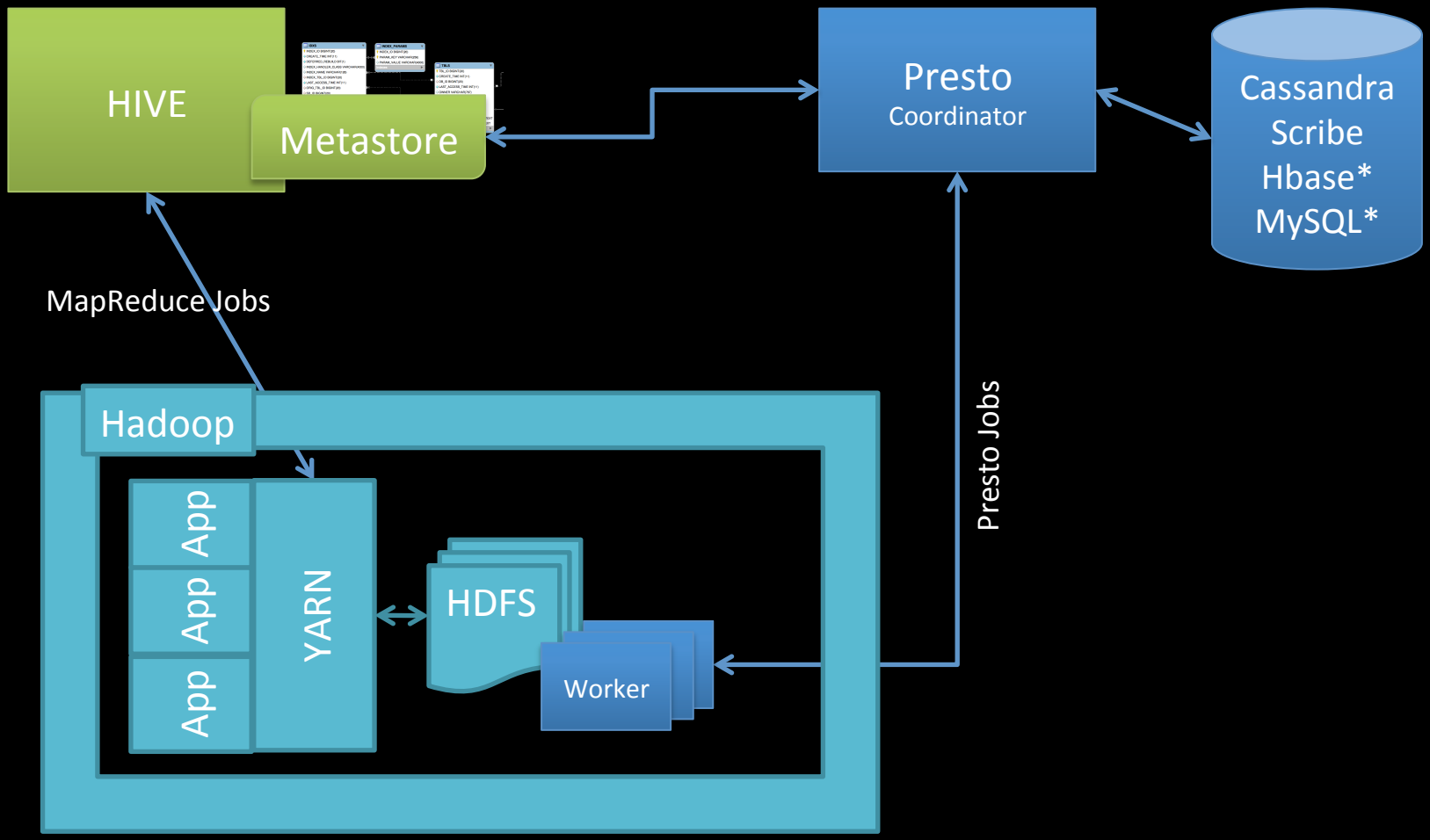
„Interactive SQL-on-Hadoop Engine“

- Target HDFS (on-disk data)
- Not based on hadoop‘s mapreduce

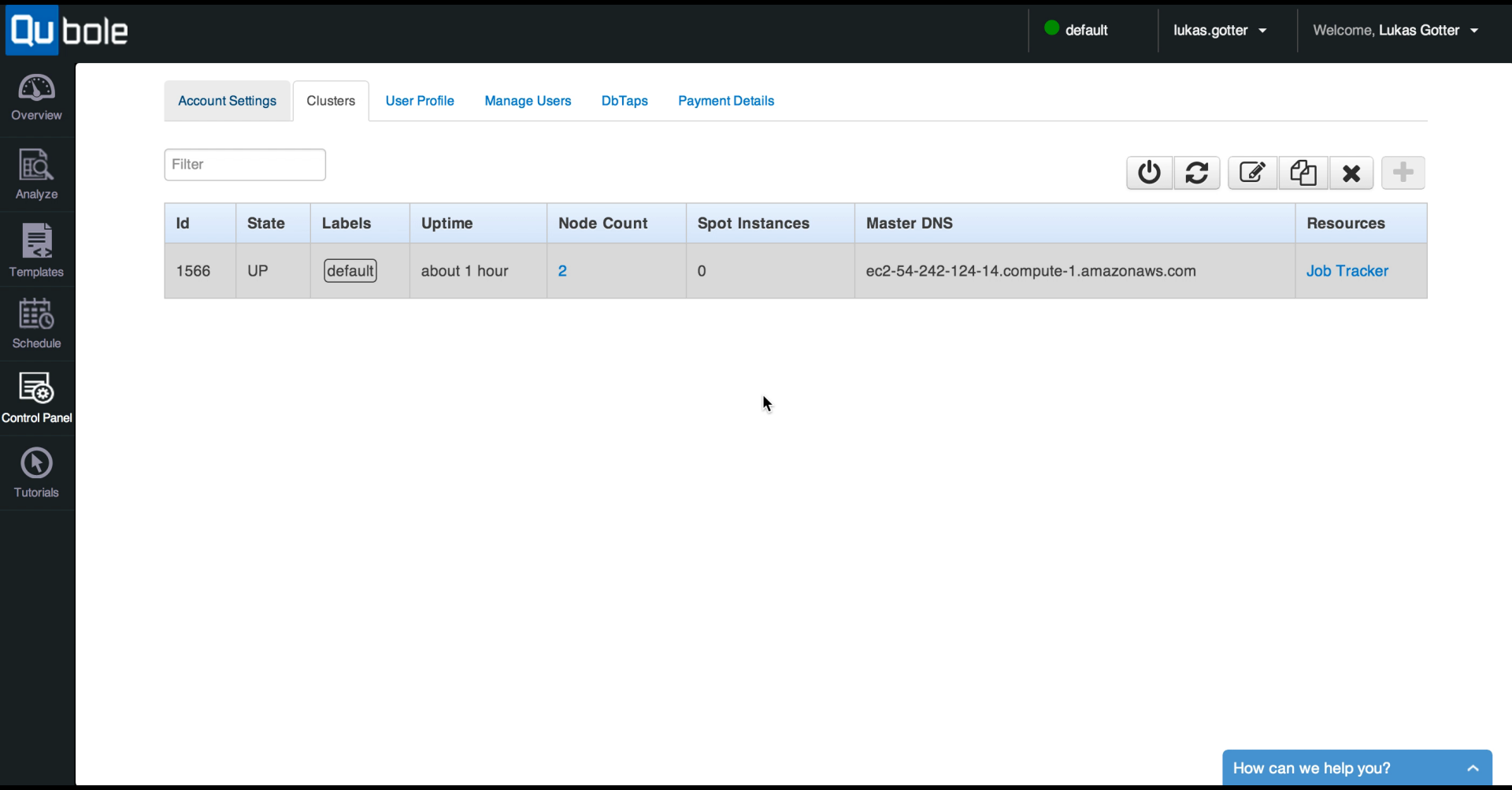
- Intermediate Results in-memory vs on-hdfs
- Pipelined execution

- Optimizations:
  - Queries are compiled
  - flat-memory data structures (low GC activity)

Presto Architecture

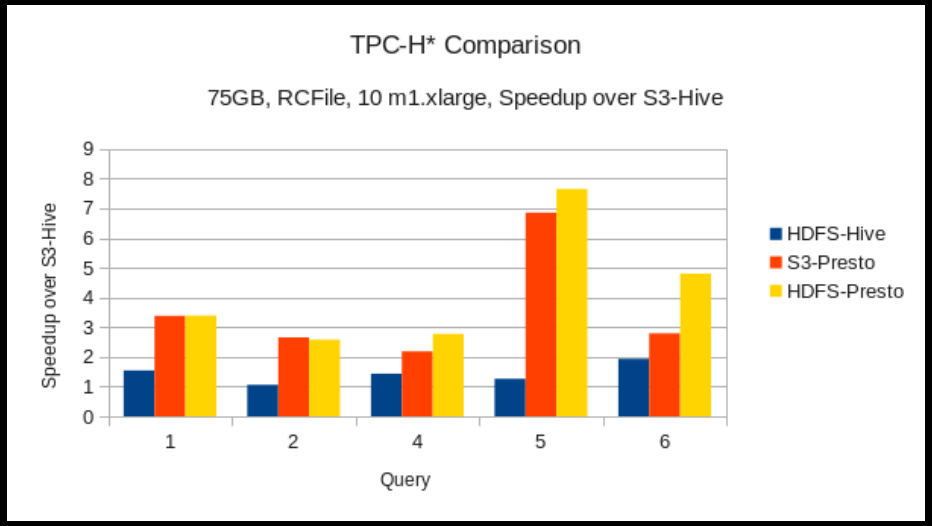


DEMO



Presto – interactive sql-on-hadoop engine

DEMO



2x-8x improvement

By Sivaramakrishnan Narayanan, Quobole

One last thing ;)

SELECT AVG(hunger)
FROM Table
WHERE event=„bedcon2014“
WITHIN 2 SECONDS

> Queries with Bounded Errors and Bounded Response Times on Very Large Data

```
Query 20131105_005539_00082_ee7y3, FINISHED, 13 nodes
Splits: 187 total, 187 done (100.00%)
0:07 [150M rows, 22.4GB] [22.2M rows/s, 3.32GB/s]

presto:default> exit

$
```

Thank You!

Lukas.Gotter@codecentric.de
@lukesolar

blog.codecentric.de

Sources

SQL is what's next for Hadoop: Here's who's doing it
http://gigaom.com/2013/02/21/sql-is-whats-next-for-hadoop-heres-whos-doing-it/

Introducing Presto - Analytics @ Scale 2013
https://www.youtube.com/watch?v=q7zfpK9bafY

The patologies of big data
http://queue.acm.org/detail.cfm?id=1563874

Facebook's New Realtime Analytics System: HBase To Process 20 Billion Events Per Day
http://highscalability.com/blog/2011/3/22/facebooks-new-realtime-analytics-system-hbase-to-process-20.html

Big data benchmark
https://amplab.cs.berkeley.edu/benchmark/

Understanding Hadoop Clusters and the Network
http://bradhedlund.com/2011/09/10/understanding-hadoop-clusters-and-the-network/#comment-11853

Shark: SQL and Rich Analytics at Scale
https://amplab.cs.berkeley.edu/wp-content/uploads/2013/02/shark_sigmod2013.pdf

Quora: What are the main differences between Facebook Presto and Amplab Shark?
http://www.quora.com/Facebook-Engineering/What-are-the-main-differences-between-Facebook-Presto-and-Amplab-Shark

Qubole – Big data as a service (company website)
http://www.qubole.com/

Presto official website
http://prestodb.io/