

# Evaluation of distributed open source solutions in CERN database use cases

HEPiX, spring 2015

Kacper Surdy IT-DB-DBF

M. Grzybek, D. L. Garcia, Z. Baranowski, L. Canali, E. Grancher

# Motivations

- propose better suited solution for big volumes of **read-only** data
- open new possibilities for data analysis
- offload Oracle in data warehouse workloads
- our users ask for this!

# Hadoop evaluation

- One interface (SQL)
- Different use cases from CERN
  - LHC log, Controls data, Atlas jobs manager (PANDA)
- Different query engines
- Different approaches for storing data

# Query engines

## Apache Hive

- SQL -> MapReduce jobs

## Cloudera Impala

- Custom SQL planner and executor
- Uses Hive metastore

## Apache Spark SQL

- Spark lightweight threading execution
- Can use Hive metastore

# Data formats

Data storage format and compression type make a difference

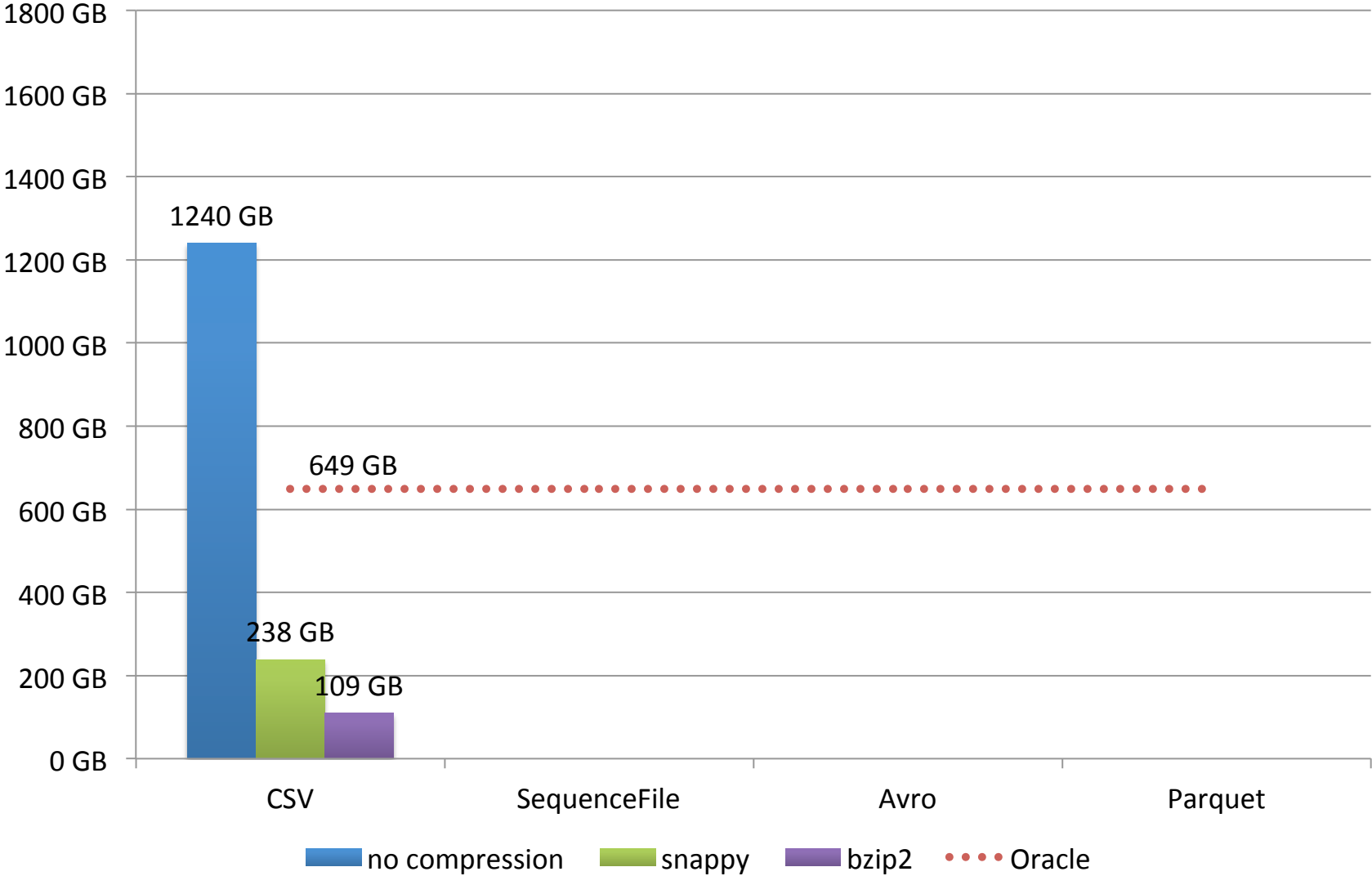
- data volume
- IO vs. CPU bound throughput
- benefits of columnar storage
- partitioning

# Data formats

## CSV

- human readable

# ACCLOG 8 days - file size comparison





# Data formats

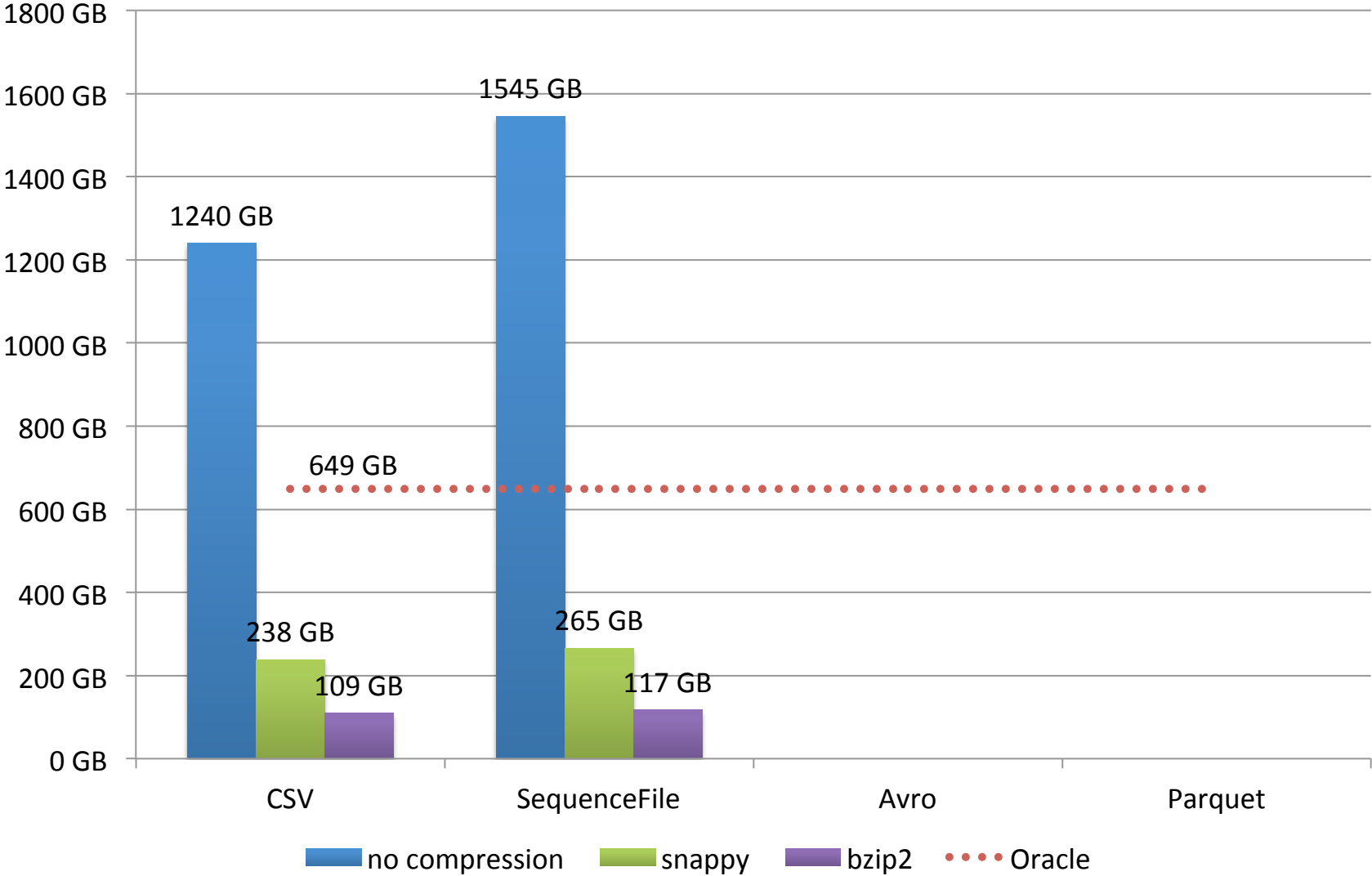
## CSV

- human readable

## SequenceFile

- flat set of binary stored (key, value) pairs

# ACCLOG 8 days - file size comparison



# Data formats

## CSV

- human readable

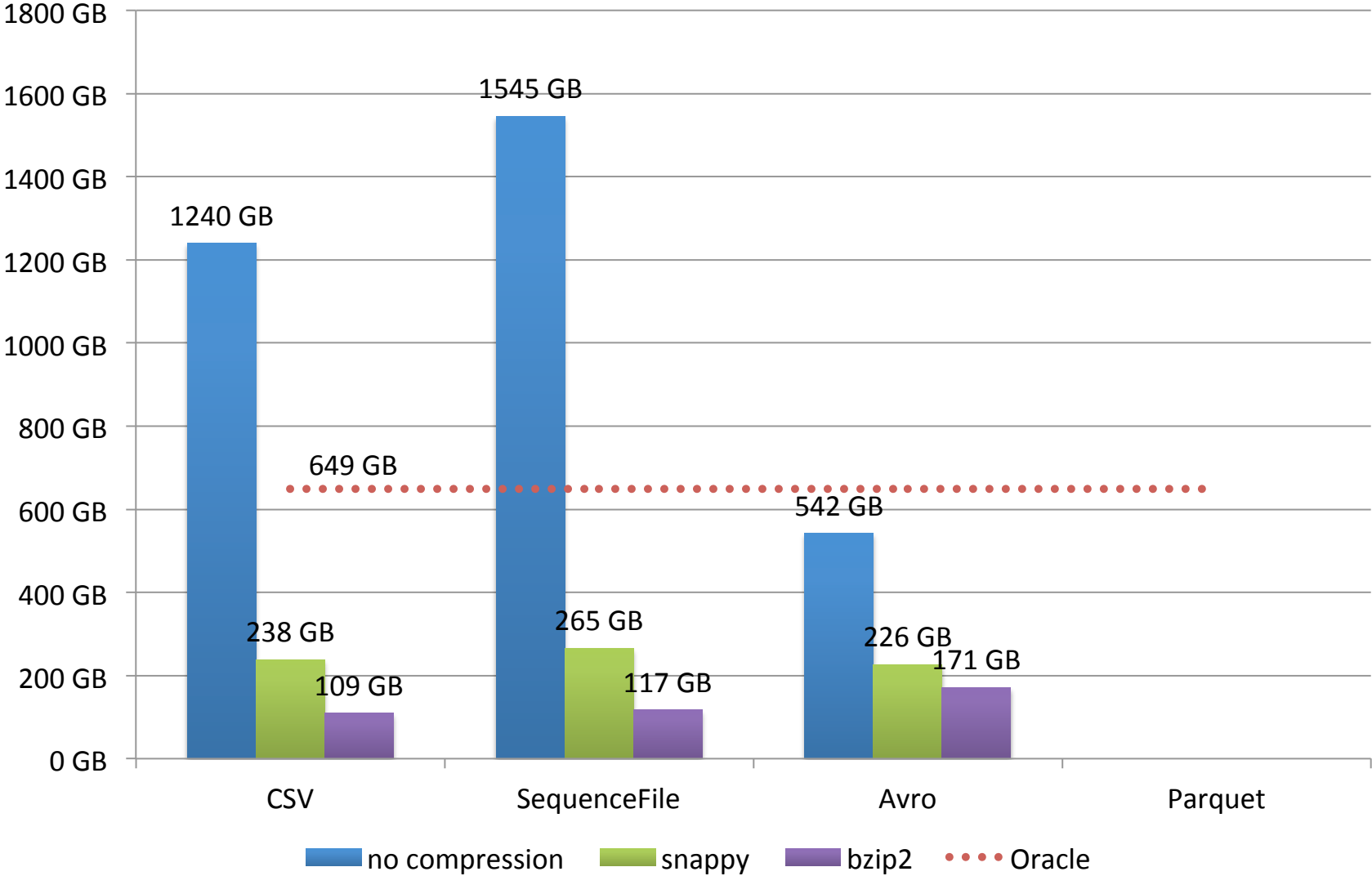
## SequenceFile

- flat set of binary stored (key, value) pairs

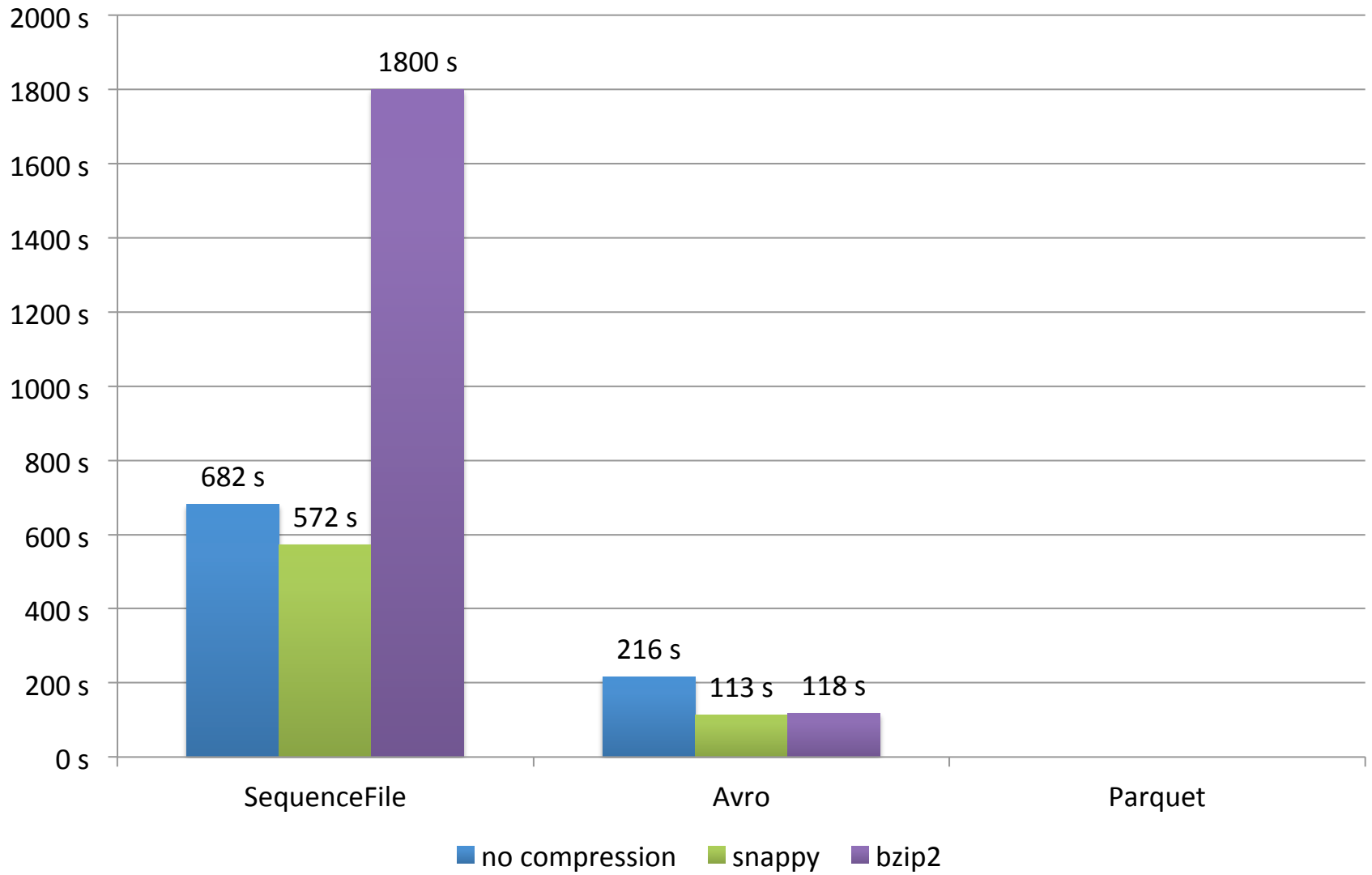
## Apache Avro

- binary serializing standard

# ACCLOG 8 days - file size comparison



## ACCLOG 8 days - union query execution time



# Data formats

## CSV

- human readable

## SequenceFile

- flat set of binary stored (key, value) pairs

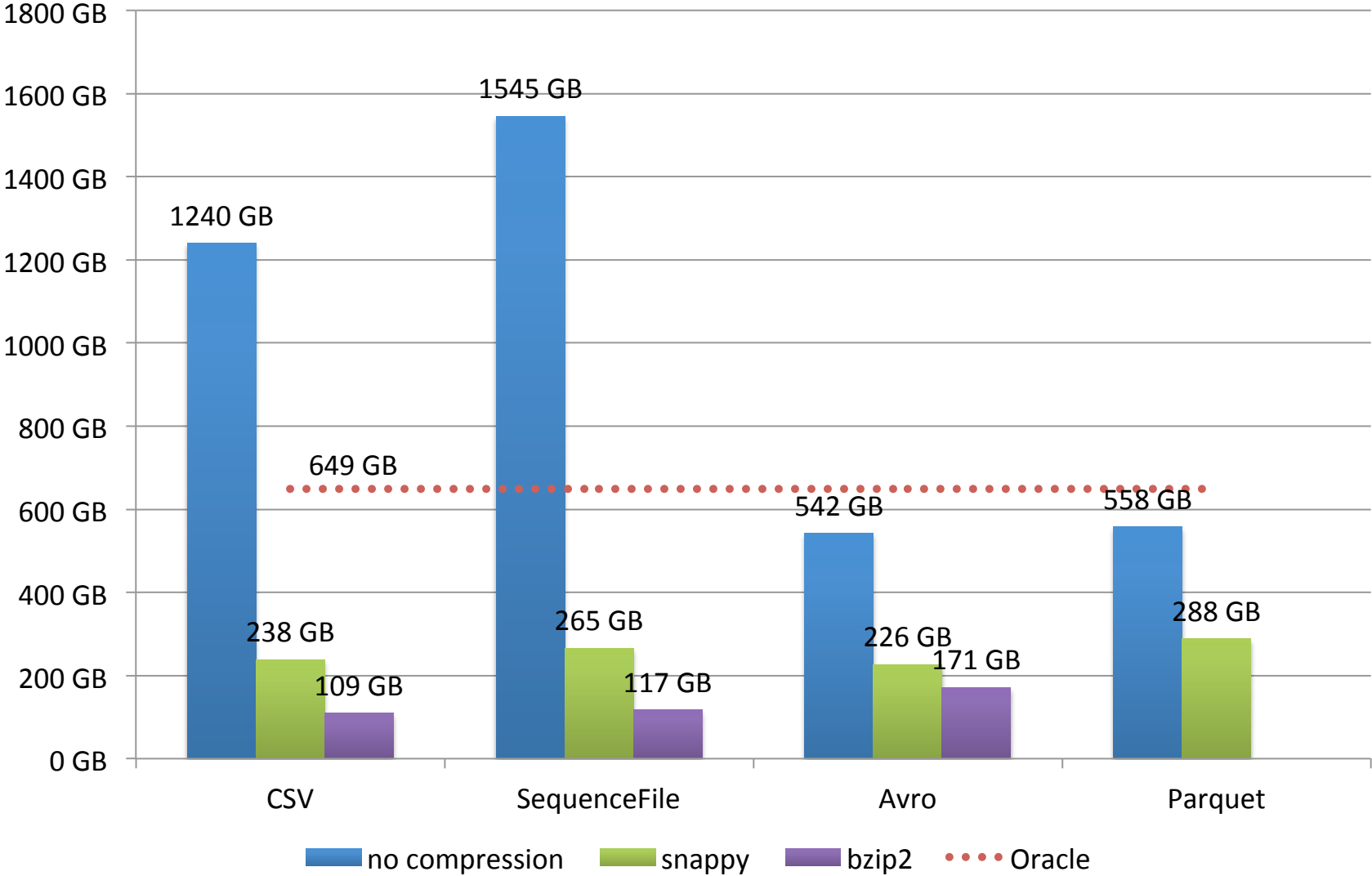
## Apache Avro

- binary serializing standard

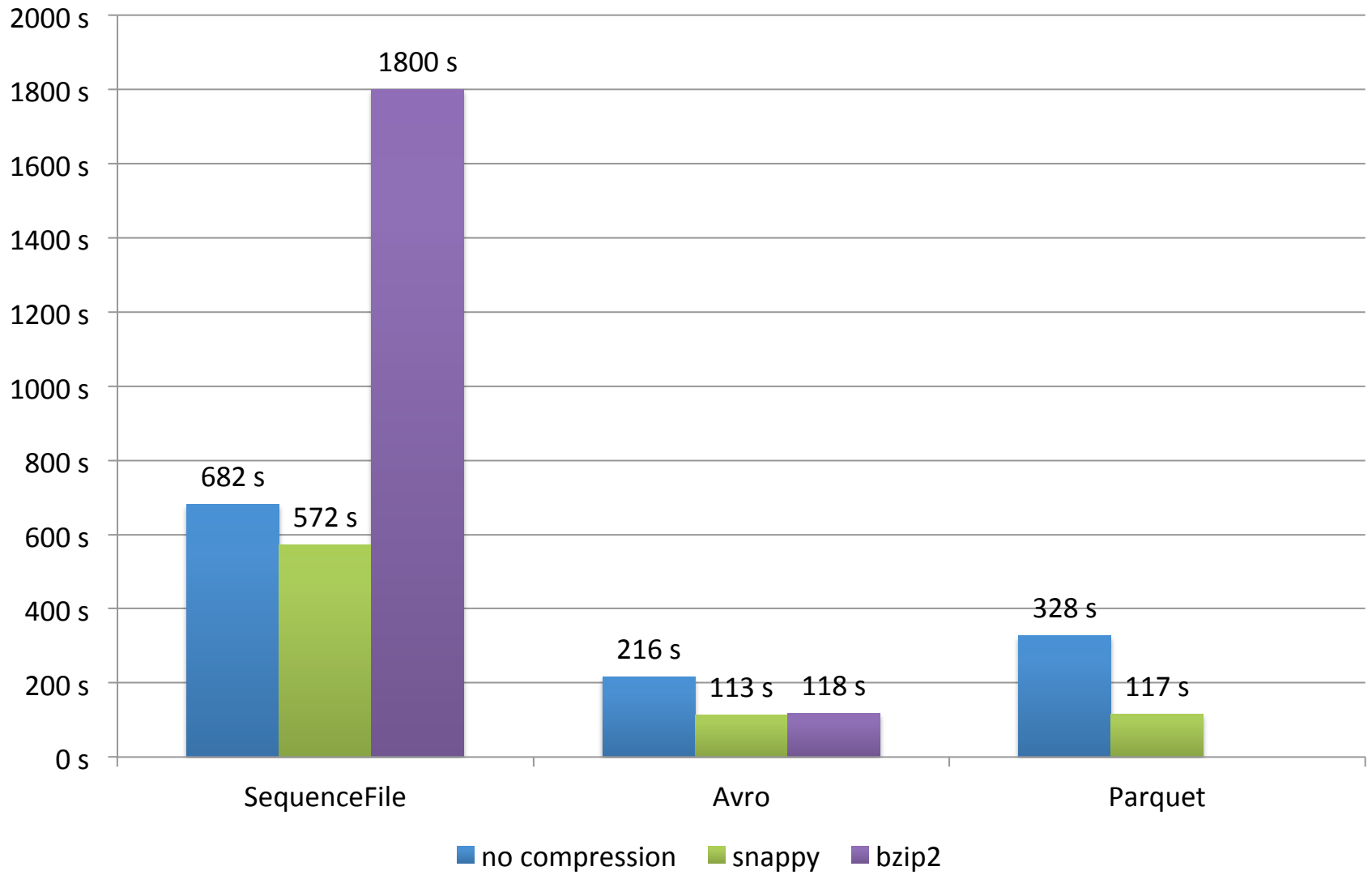
## Parquet

- binary columnar-storage

# ACCLOG 8 days - file size comparison



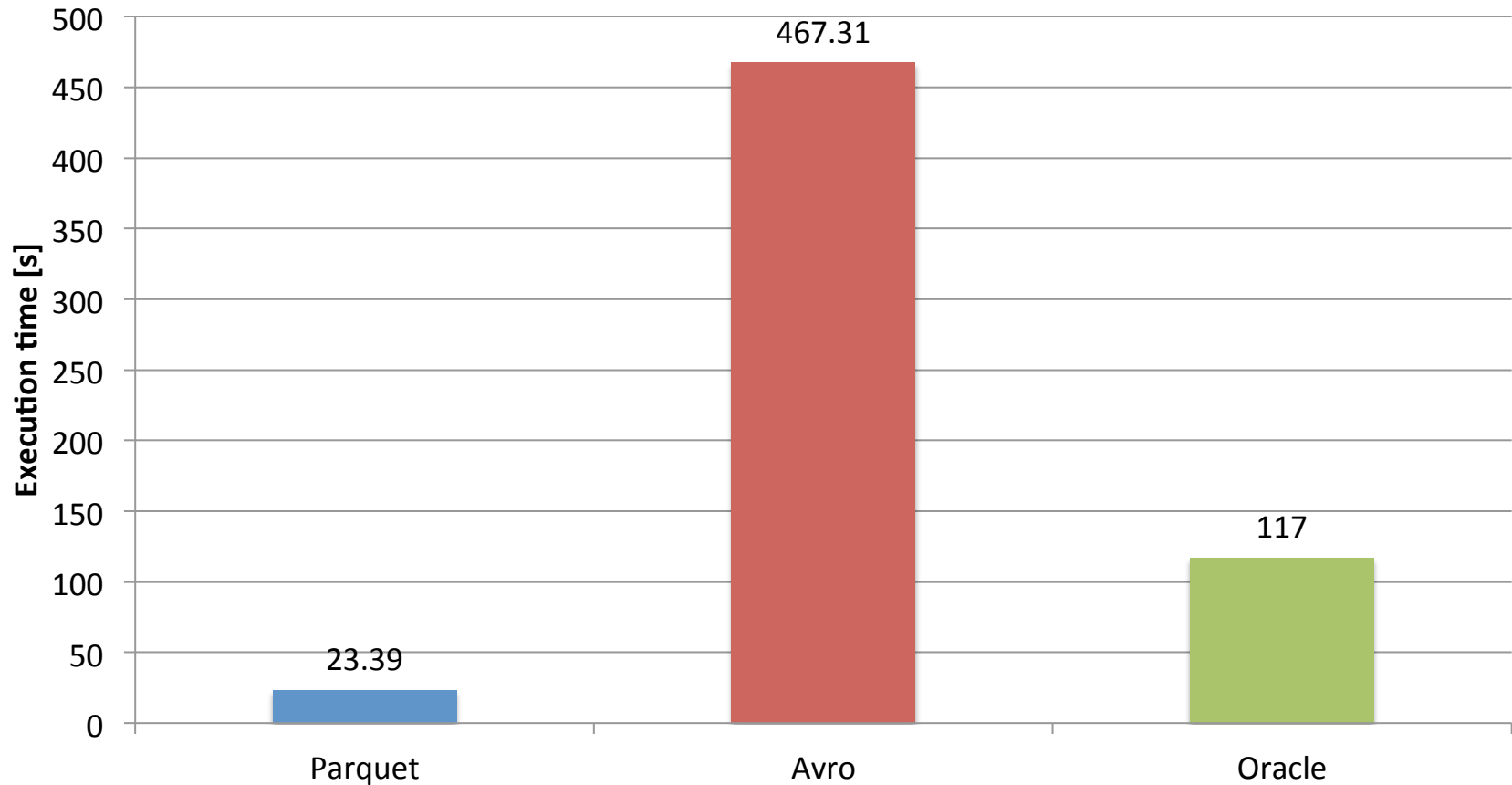
## ACCLOG 8 days - union query execution time





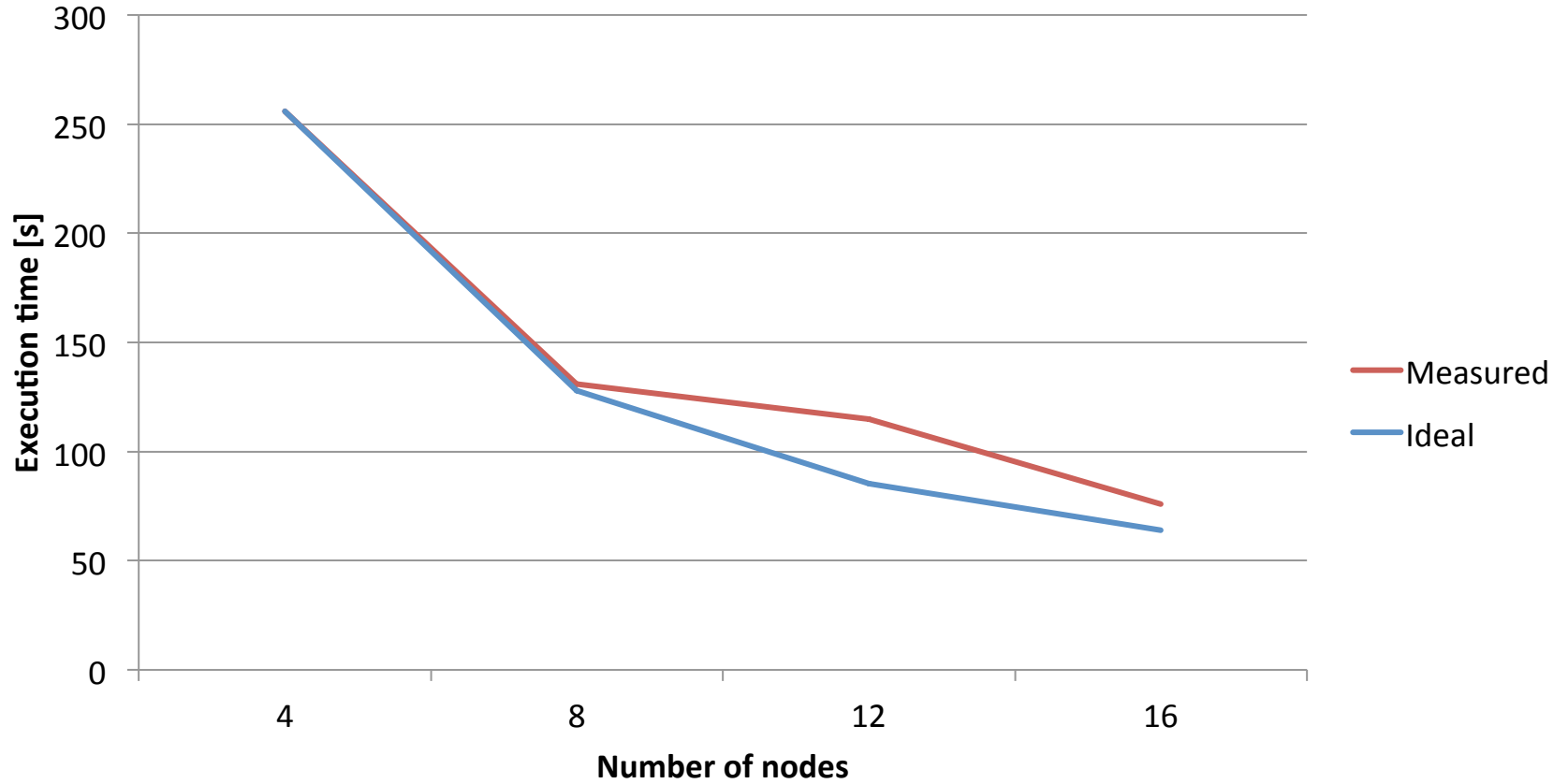
# Potential columnar store benefit

Ntuples data exeuction comparison



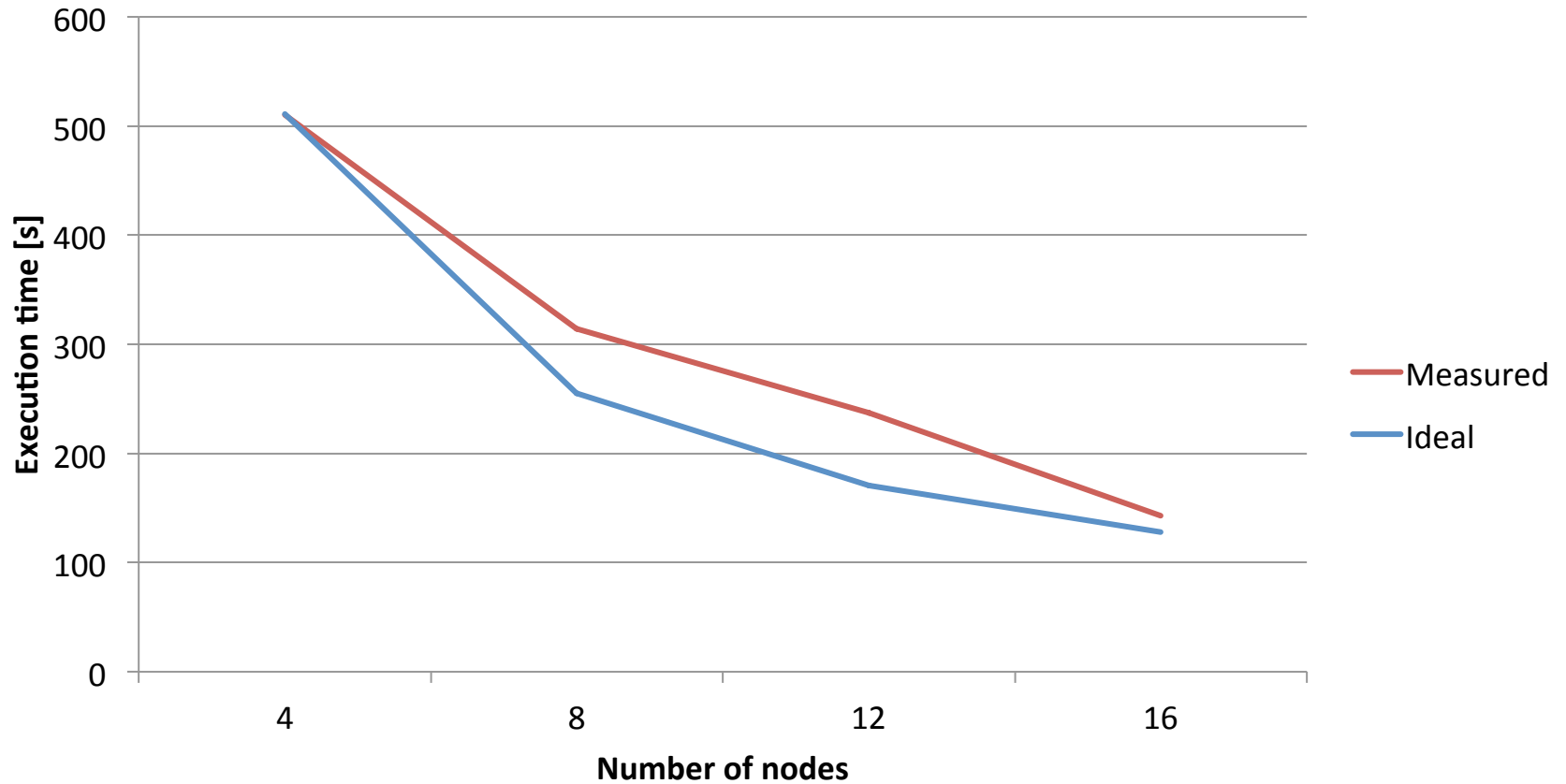
# Scalability - Impala

ACCLOG data

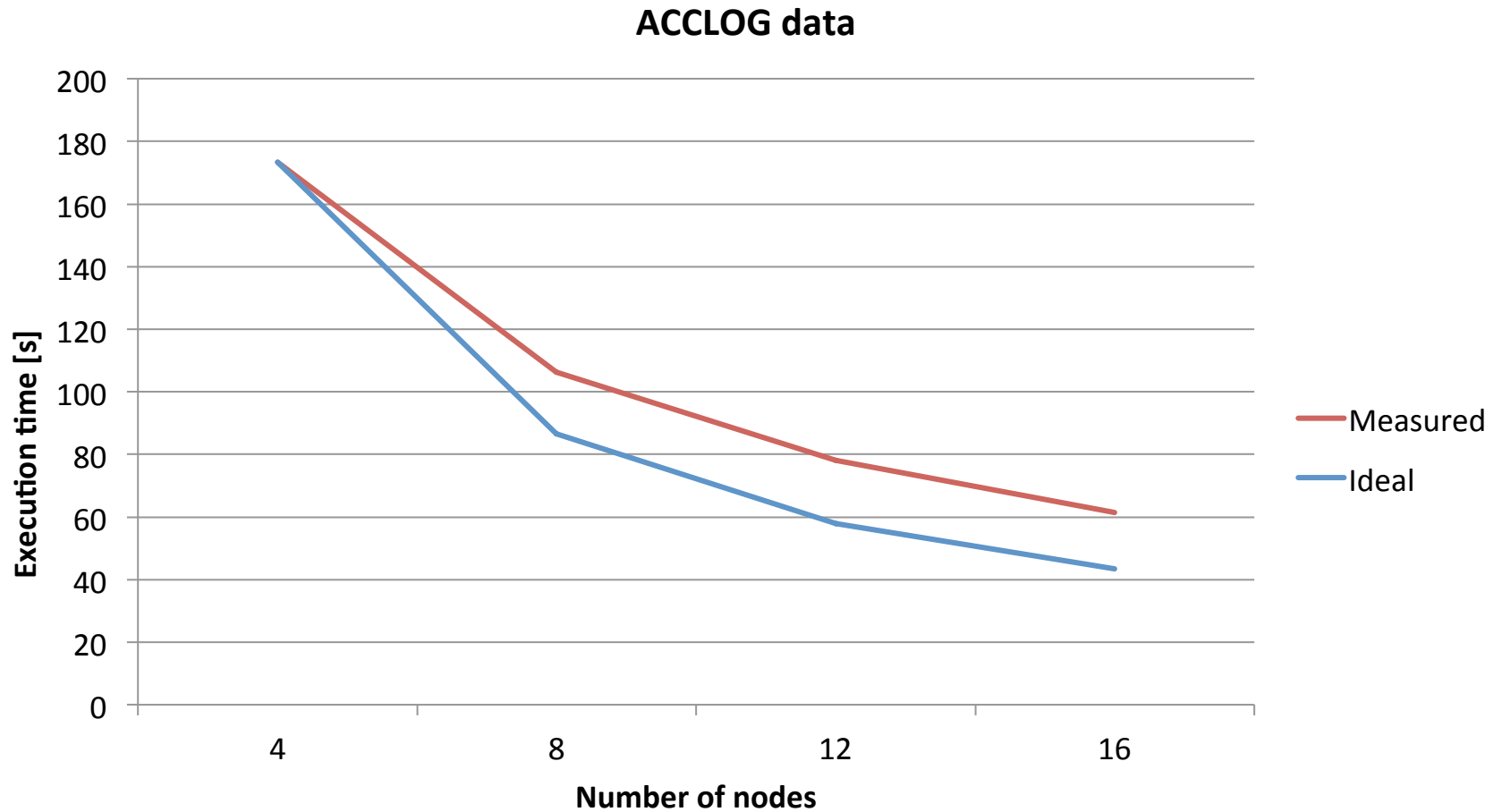


# Scalability - Hive

ACCLOG data – Hive scalability



# Scalability – Spark SQL



# Summary

- Data format choice is a key
- Heterogeneous data in the same place
- No matter what interface -> it scales



[www.cern.ch](http://www.cern.ch)