



UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG



eXascale Infolab

Big Data: an Introduction

Philippe Cudré-Mauroux

eXascale Infolab, University of Fribourg
Switzerland

October 19, 2017
IT Valley Big Data Workshop, Fribourg

Menu for Today

- Introduction to Big Data infrastructures (~1h)
- Login to DAPLAB, HDFS (~45')
- Hadoop MapReduce (~60')
- Apache Hive (~45')

(Tomorrow: Spark, Zeppelin, Latent Semantic Analysis)

Instant Quiz

- Logistic Regression?
- 3 Vs of Big Data?
- Hadoop?
- Spark?

Exascale Data Deluge

- Web companies
 - Google
 - Ebay
 - Yahoo
- Science
 - Biology
 - Astronomy
 - Remote Sensing
- Financial services, retail companies, manufacturing, governments, etc.



- New data formats, new machines
- Peta & exa-scale datasets
- Obsolescence of traditional information infrastructures
- New opportunities



Big Data Central Theorem

Data+Technology → Actionable Insight → \$\$

Big Data as a New Class of Asset

- The Age of Big Data (NYTimes Feb. 11, 2012)
<http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html>

“Welcome to the Age of Big Data. The new megarich of Silicon Valley, first at Google and now Facebook, are masters at harnessing the data of the Web — online searches, posts and messages — with Internet advertising. At the World Economic Forum last month in Davos, Switzerland, Big Data was a marquee topic. A report by the forum, “Big Data, Big Impact,” declared **data a new class of economic asset, like currency or gold.**”

Big data can generate significant financial value across sectors



US health care

- \$300 billion value per year
- ~0.7 percent annual productivity growth



Europe public sector administration

- €250 billion value per year
- ~0.5 percent annual productivity growth



Global personal location data

- \$100 billion+ revenue for service providers
- Up to \$700 billion value to end users



US retail

- 60+% increase in net margin possible
- 0.5–1.0 percent annual productivity growth



Manufacturing

- Up to 50 percent decrease in product development, assembly costs
- Up to 7 percent reduction in working capital

SOURCE: McKinsey Global Institute analysis

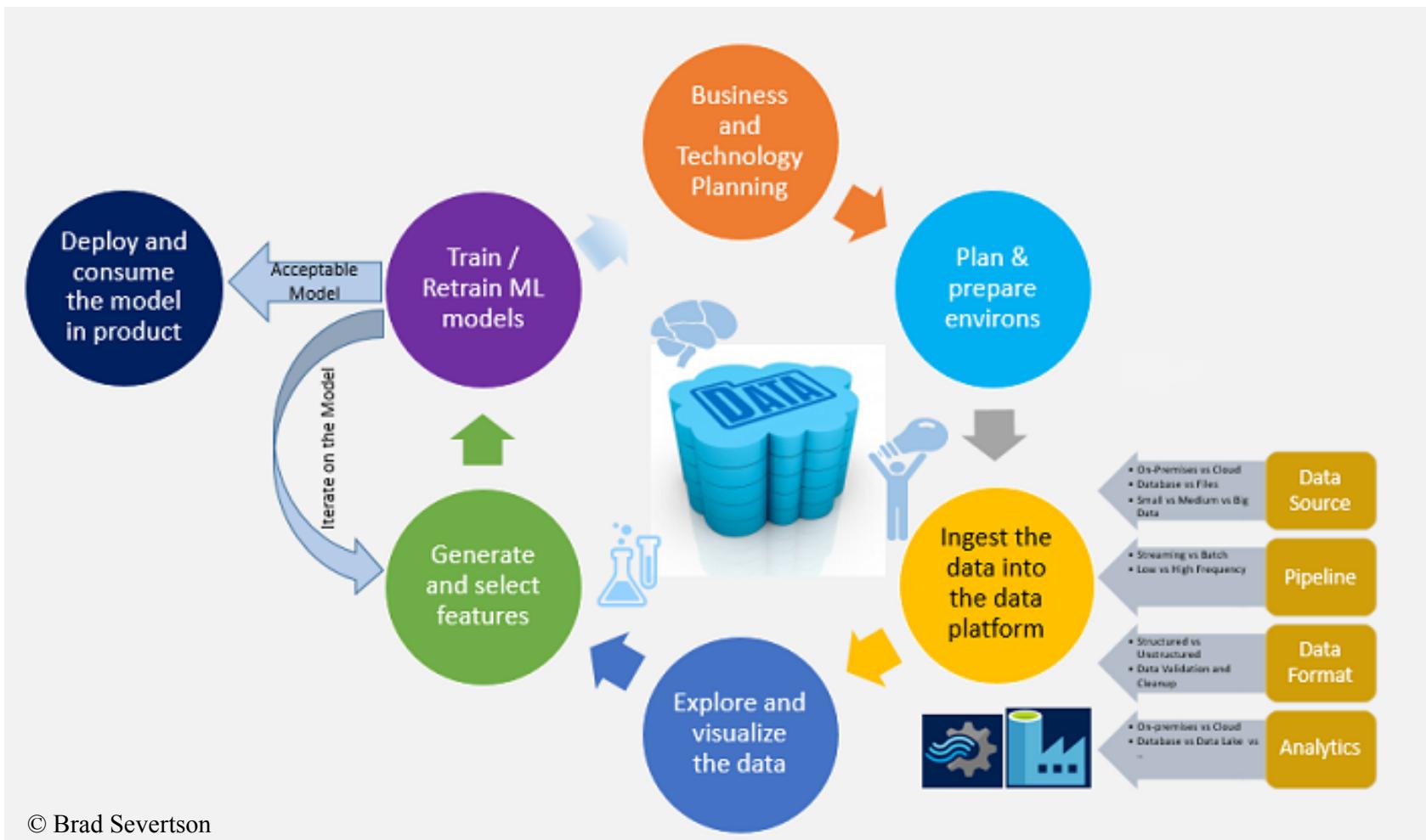
The 3-**V**s of Big Data

- **V**olume
 - Amount of data
- **V**elocity
 - speed of data in and out
- **V**ariety
 - range of data types and sources
- [Gartner 2012] "*Big Data are high-volume, high-velocity, and/or high-variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization*"

What can you do with the data

- Reporting
 - Post Hoc
 - Real time
- Monitoring (fine-grained)
- Exploration
- Finding Patterns
- Root Cause Analysis
- Closed-loop Control
- Model construction
- **Prediction**
- ...

Data Science Lifecycle

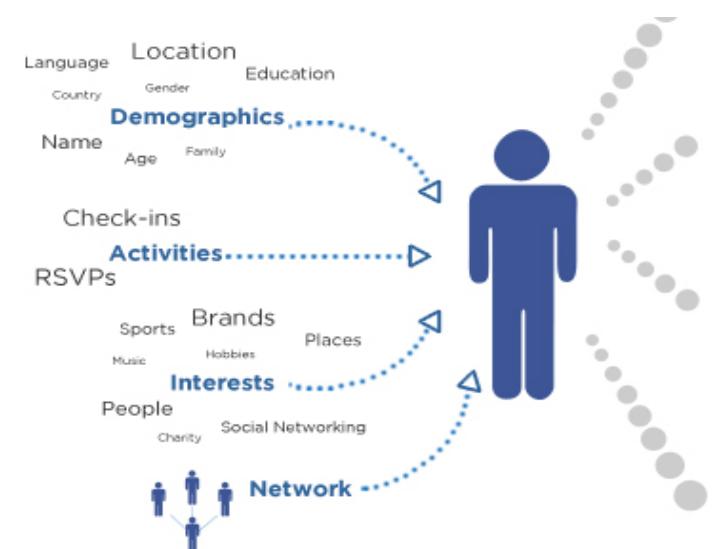


Different Models

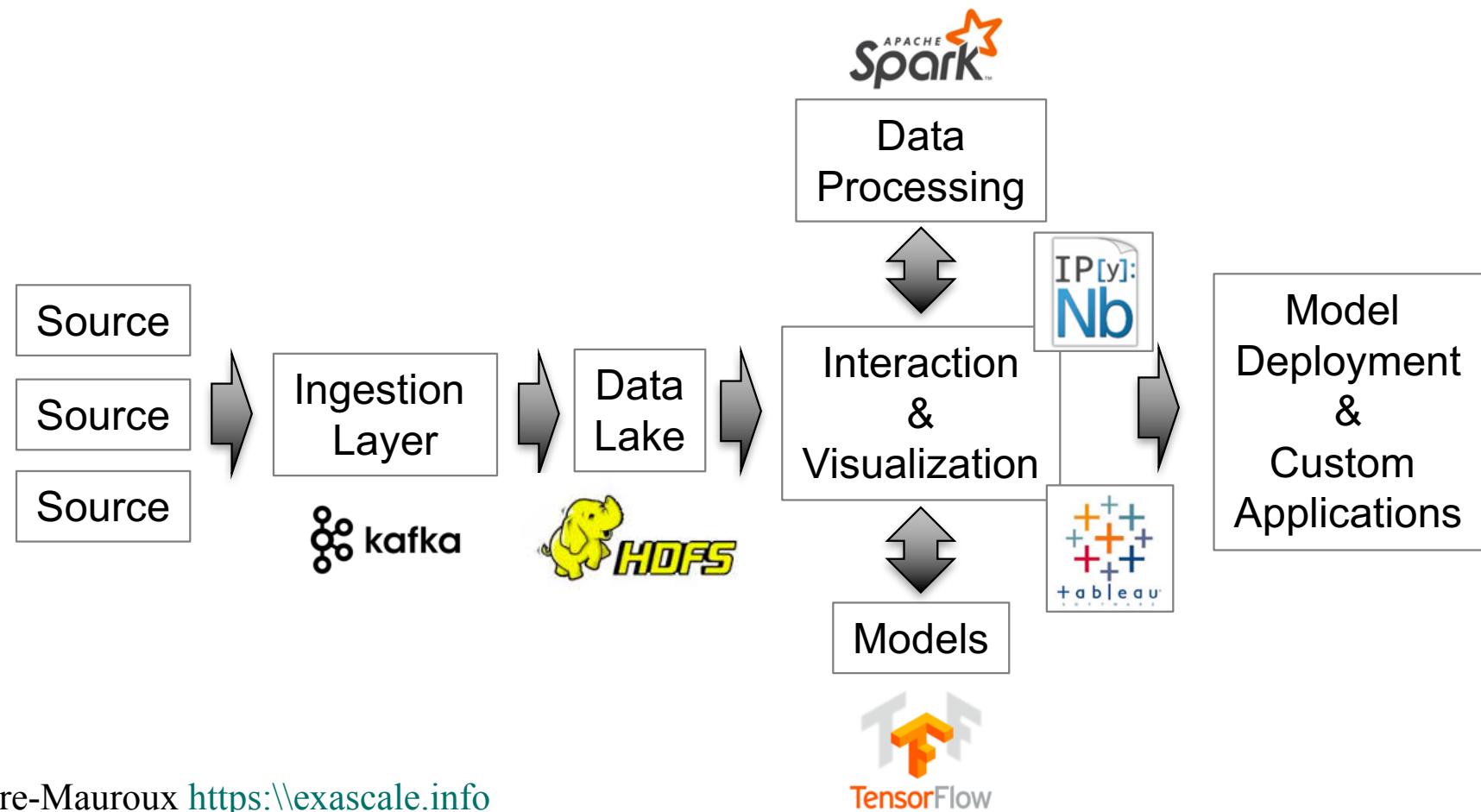
- **Supervised** learning
(classification, regression, anomaly detection)
 - Build a model from labeled examples
 - Linear / logistic regression
 - Trees, forests, jungles
 - SVMs
 - Bayesian
 - Deep Learning
- **Unsupervised** learning (clustering)
 - Data points have no labels associated with them
 - Often, goal is to organize the data in some way or to describe its structure

Typical Big Data Success Story

- **Modeling users** through Big Data
 - Online ads sale / placement [e.g., Facebook]
 - Personalized Coupons [e.g., Target]
 - Product Placement [Walmart]
 - Content Generation [e.g., NetFlix]
 - Personalized learning [e.g., Duolingo]
 - HR Recruiting [e.g., Gild]



Data Science Infrastructure (circa 2017)



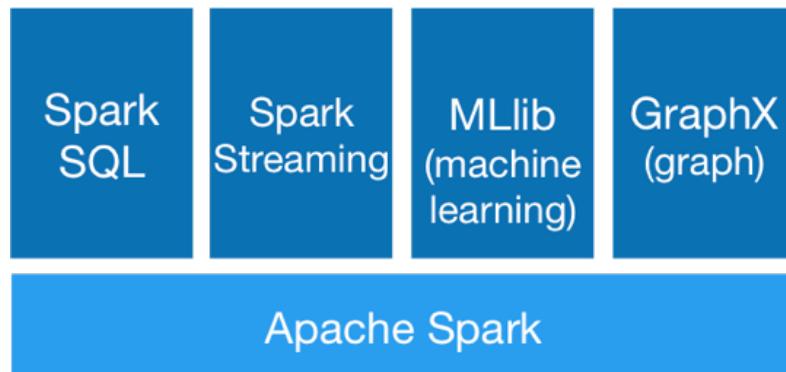
Data Lake

- Often based on **Hadoop's** Distributed File System (HDFS)
- Elastic (**scale-out**)
- Social impact
- Governance?

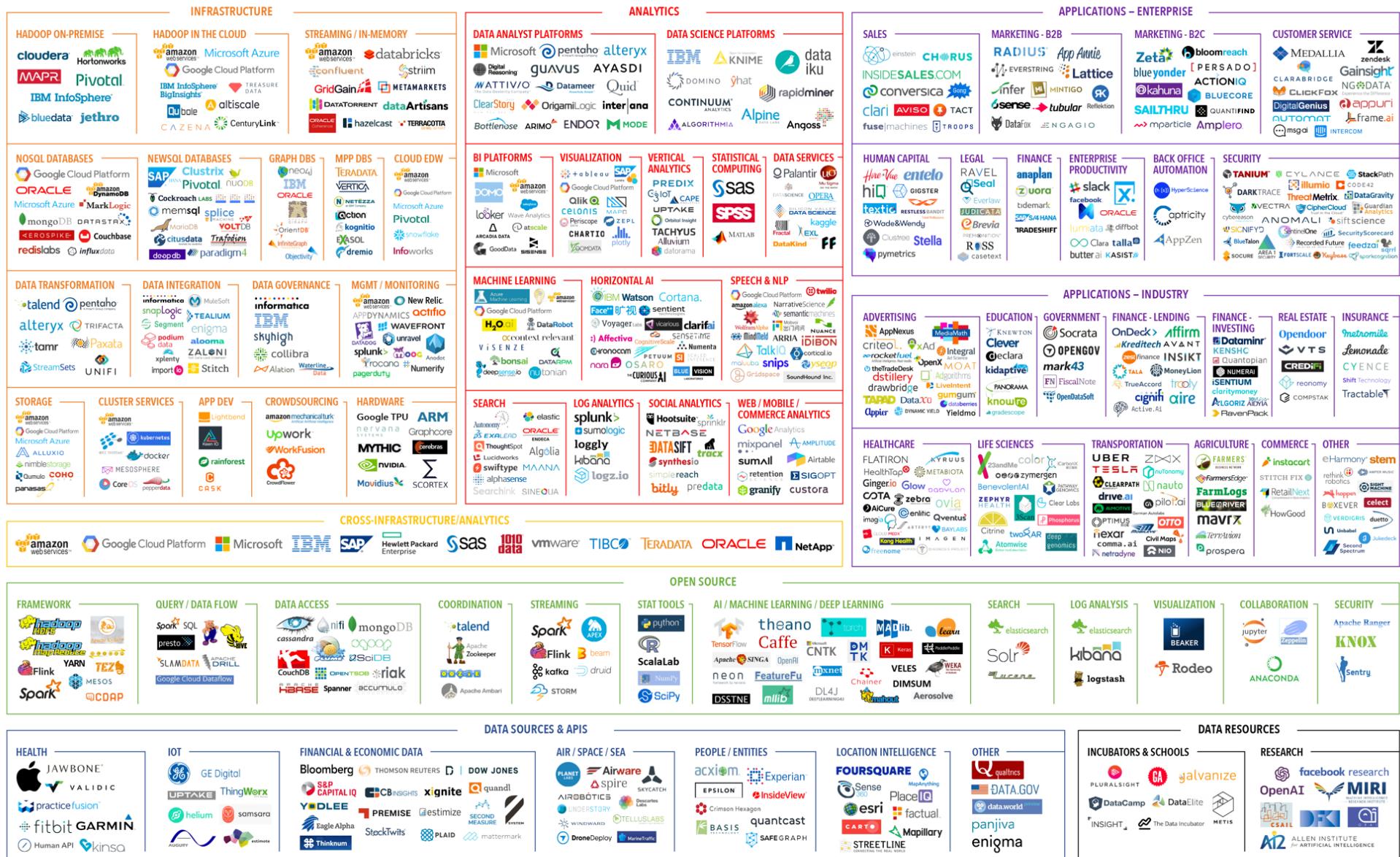


Spark

- In-memory cluster computation framework
- Can run on top of Hadoop's filesystem (HDFS)
- Versatile data processing solution



BIG DATA LANDSCAPE 2017



Hadoop



Distributed Computation
MapReduce

Distributed Storage
HDFS



Why Hadoop?

- Your job is to *process* 1TB of data
 - You: *Sure!*
 - But first, you need to read it.

```
with open(bigFile) as f:  
    line = f.readlines()  
    ...
```
- ☹ It's taking forever.
- HDD (reads : 150 MB/s; 10ms per seek) ... => hours
 - Similar issues w/ SSD, even RAM
- Multithreading will not help (especially not for HDD)

Scaling out...

- That was an exercise:
 - 10TB, 100TB, 1PB, 1EB ?
- The data storage medium is often a bottleneck
 - other cases can be bound by CPU or RAM
- We need a way to increase the throughput
 - Limited if only one machine (scale up)
 - Scale **out!**
- N Machines with K disks
 - Cost effective when using cheap machines.
- What are the consequences ?

Apache Hadoop

- A platform (**ecosystem**) for Big Data processing
 - Started by Yahoo! and was inspired by the following papers from google Google
 - MapReduce

Dean, Jeffrey, and Sanjay Ghemawat. "MapReduce: simplified data processing on large clusters." Communications of the ACM 51.1 (2008): 107-113.
 - GFS

Ghemawat, Sanjay, Howard Gobioff, and Shun-Tak Leung. "The Google file system." ACM SIGOPS operating systems review. Vol. 37. No. 5. ACM, 2003.
- Written in Java mostly
- Main core components
 - A Map Reduce implementation
 - Hadoop Distributed File System (HDFS)
 - YARN (resource manager)

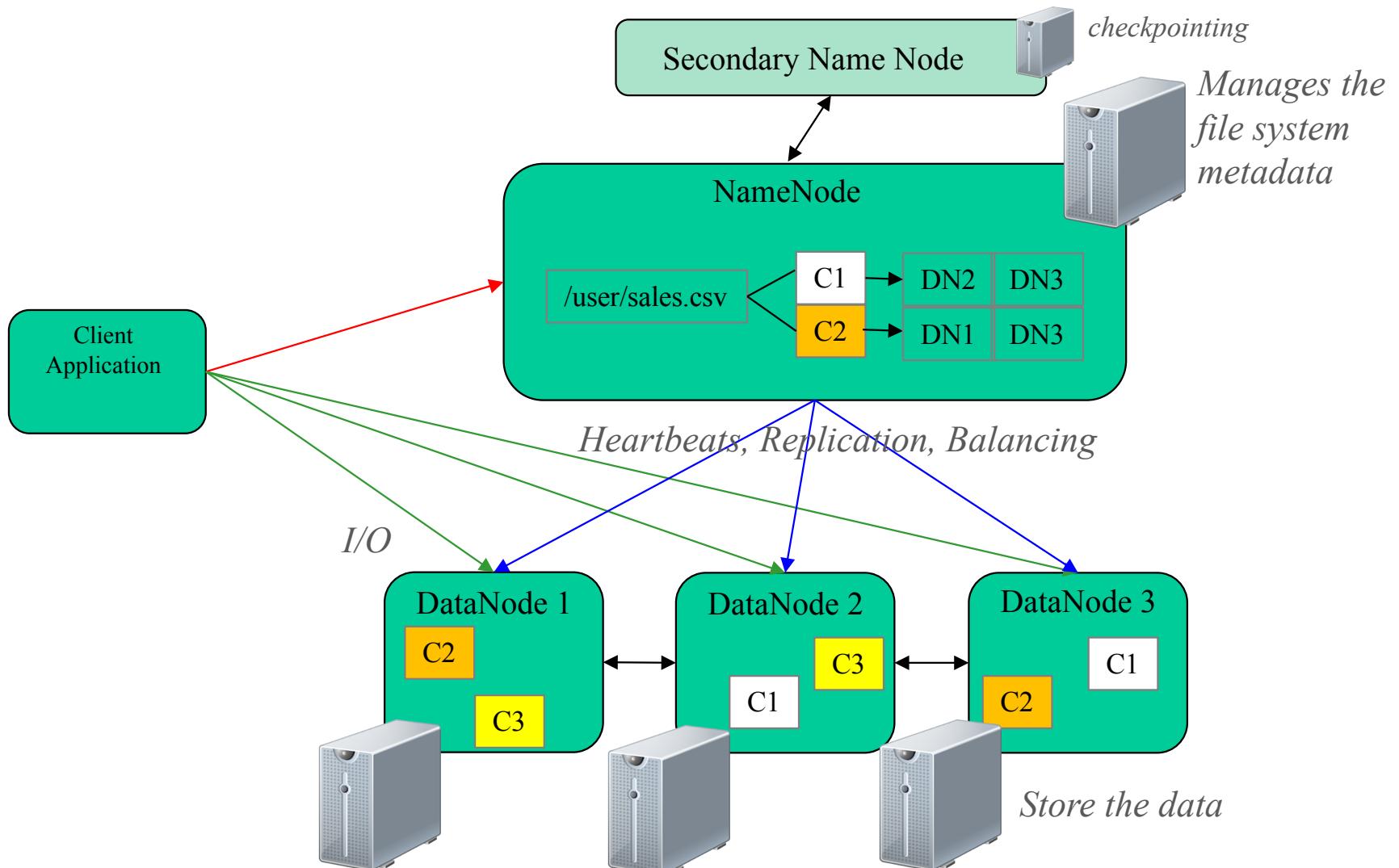
HDFS: Motivation

- Lots of commodity hardware
 - Cheap PCs
- High failure rates
 - Commodity components failure is much higher than server grade.
- Write-once read many times data
 - Logs, archives etc.
 - Processing is mostly read oriented
- Large streaming reads over random access
 - High sustained throughput over low latency

HDFS Basic Operations

- When an input file is added to HDFS
 - File is split into smaller **blocks** of **fixed size** (typically 64 or 128 MB)
 - Each block is replicated to multiple hosts (machines)
 - Each replicated block is stored on a different host
 - Replication level is configurable
 - Default is 3
 - If a host crashes or is decommissioned
 - All blocks are replicated to a new host
 - In case a new host is added
 - Blocks will be rebalanced (avoid data skew)

HDFS Architecture



MapReduce

- MapReduce is a programming model and an associated implementation for processing Big Data in parallel on clusters
- The programmer needs to write two functions:
 - **Map(Key, Value) => <KEY1, VALUE1>**
 - Performs filtering and sorting
 - All values with the same key are sent the same reducer
 - **Reduce(KEY1, VALUE1) => <KEY2, VALUE2>**
 - Performs a summary operation

Word Length Count

Abridged Declaration of Independence

A Declaration By the Representatives of the United States of America, in General Congress Assembled.

When in the course of human events it becomes necessary for a people to advance from that subordination in which they have hitherto remained, and to assume among powers of the earth the equal and independent station to which the laws of nature and of nature's god entitle them, a decent respect to the opinions of mankind requires that they should declare the causes which impel them to the change.

We hold these truths to be self-evident; that all men are created equal and independent; that from that equal creation they derive rights inherent and inalienable, among which are the preservation of life, and liberty, and the pursuit of happiness; that to secure these ends, governments are instituted among men, deriving their just power from the consent of the governed; that whenever any form of government shall become destructive of these ends, it is the right of the people to alter or to abolish it, and to institute new government, laying its foundation on such principles and organizing its power in such form, as to them shall seem most likely to effect their safety and happiness. Prudence indeed will dictate that governments long established should not be changed for light and transient causes: and accordingly all experience hath shewn that mankind are more disposed to suffer while evils are sufferable, than to right themselves by abolishing the forms to which they are accustomed. But when a long train of abuses and usurpations, begun at a distinguished period, and pursuing invariably the same object, evinces a design to reduce them to arbitrary power, it is their right, it is their duty, to throw off such government and to provide new guards for future security. Such has been the patient sufferings of the colonies; and such is now the necessity which constrains them to expunge their former systems of government. the history of his present majesty is a history of unremitting injuries and usurpations, among which no one fact stands single or solitary to contradict the uniform tenor of the rest, all of which have in direct object the establishment of an absolute tyranny over these states. To prove this, let facts be submitted to a candid world, for the truth of which we pledge a faith yet unsullied by falsehood.

Word Length Count

Map Task 1
(204 words)

Yellow: 10+

Red: 5.9

Blue: 2.4

Pink: = 1

Map Task 2
(190 words)

Abridged Declaration of Independence

A Declaration By the Representatives of the United States of America, in General Congress Assembled.
When in the course of human events it becomes necessary for a people to advance from that subordination in which they have hitherto remained, and to assume among powers of the earth the equal and independent station to which the laws of nature and of nature's god entitle them, a decent respect to the opinions of mankind requires that they should declare the causes which impel them to the change.
We hold these truths to be self-evident; that all men are created equal and independent; that from that equal creation they derive rights inherent and inalienable, among which are the preservation of life, and liberty, and the pursuit of happiness; that to secure these ends, governments are instituted among men, deriving their just power from the consent of the governed; that whenever any form of government shall become destructive of these ends, it is the right of the people to alter or to abolish it, and to institute new government, laying its foundation on such principles and organizing its power in such form, as to them shall seem most likely to effect their safety and happiness. Prudence indeed will

(key, value)

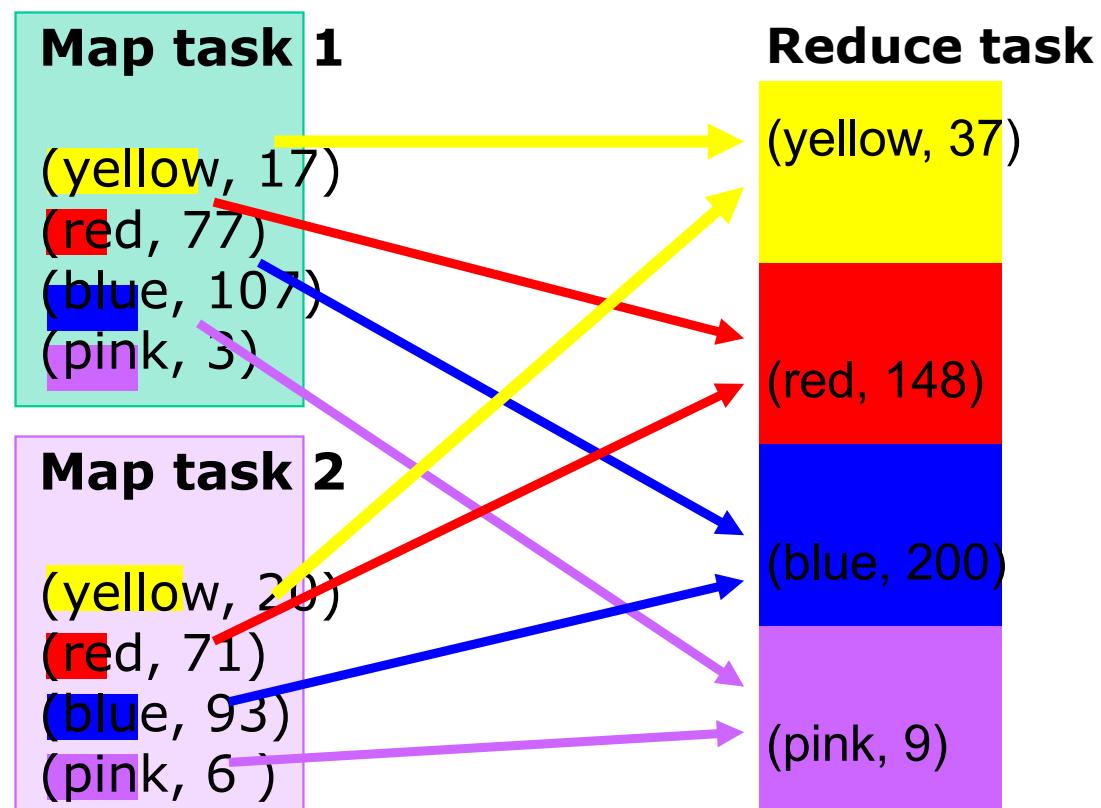
(yellow, 17)
(red, 77)
(blue, 107)
(pink, 3)

dictate that governments long established should not be changed for light and transient causes; and accordingly all experience hath shewn that mankind are more disposed to suffer while evils are sufferable, than to right themselves by abolishing the forms to which they are accustomed. But when a long train of abuses and usurpations, begun at a distinguished period, and pursuing invariably the same object, evinces a design to reduce them to arbitrary power, it is their right, it is their duty, to throw off such government and to provide new guards for future security. Such has been the patient sufferings of the colonies; and such is now the necessity which constrains them to expunge their former systems of government. the history of his present majesty is a history of unremitting injuries and usurpations, among which no one fact stands single or solitary to contradict the uniform tenor of the rest, all of which have in direct object the establishment of an absolute tyranny over these states. To prove this, let facts be submitted to a candid world, for the truth of which we pledge a faith yet unsullied by falsehood.

(yellow, 20)
(red, 71)
(blue, 93)
(pink, 6)

Example 2: Word Length Count

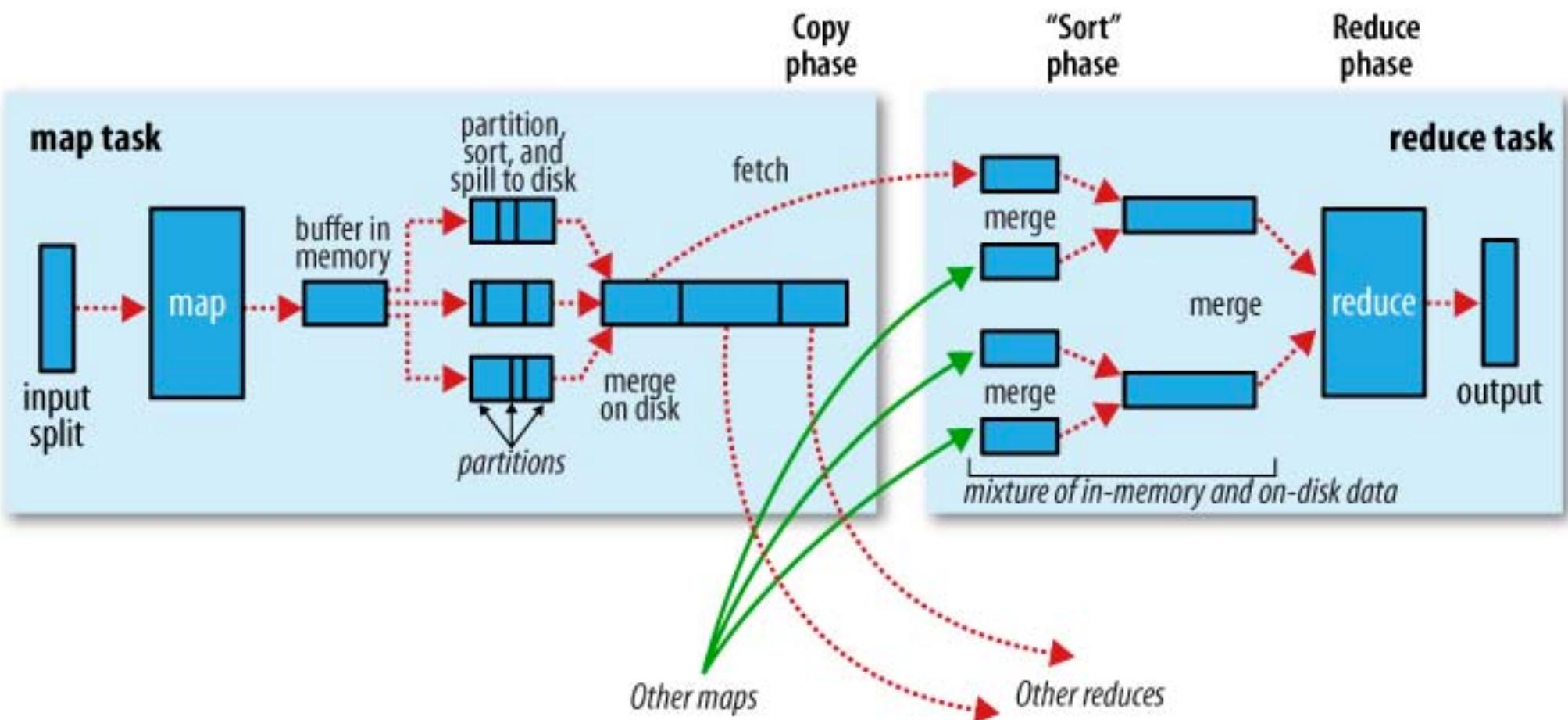
Map is a **GROUP BY** operation
Reduce is an **AGGREGATE** operation



Map Reduce Implementation

- Input reader
 - Divide input into splits (HDFS default block size is default split size), assign each split to a Map task
- Map task
 - Apply the Map function to each record in the split
 - Each Map function returns a list of (key, value) pairs
- Shuffle/Partition and Sort
 - Shuffle distributes sorting & aggregation to many reducers
 - All records for key k are directed to the same reduce processor
 - Sort groups the same keys together, and prepares for aggregation
- Reduce task
 - Apply the Reduce function to each key
 - The result of the Reduce function is a list of (key, value) pairs

The Anatomy of a M/R Execution

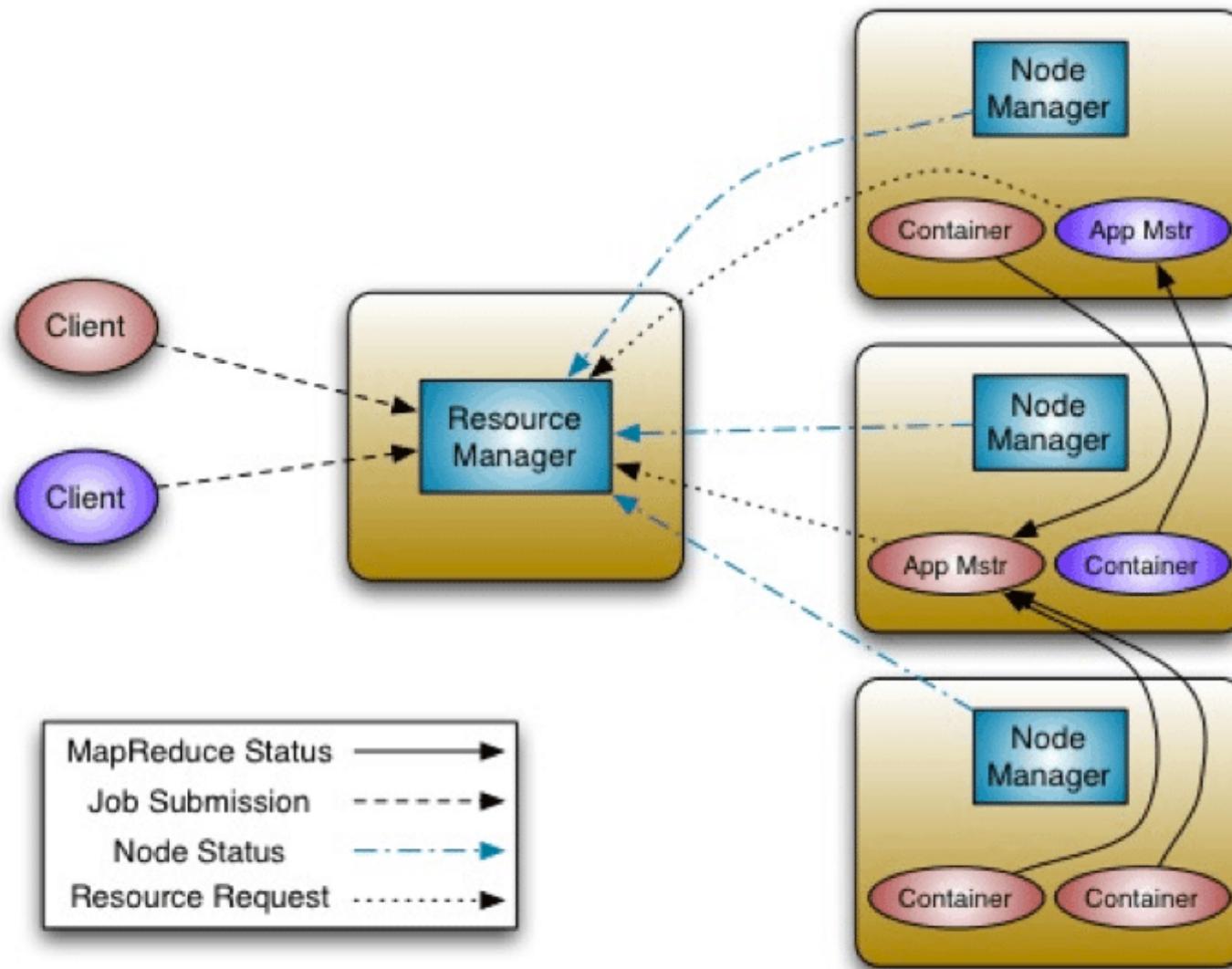


From (Tom White 2013)

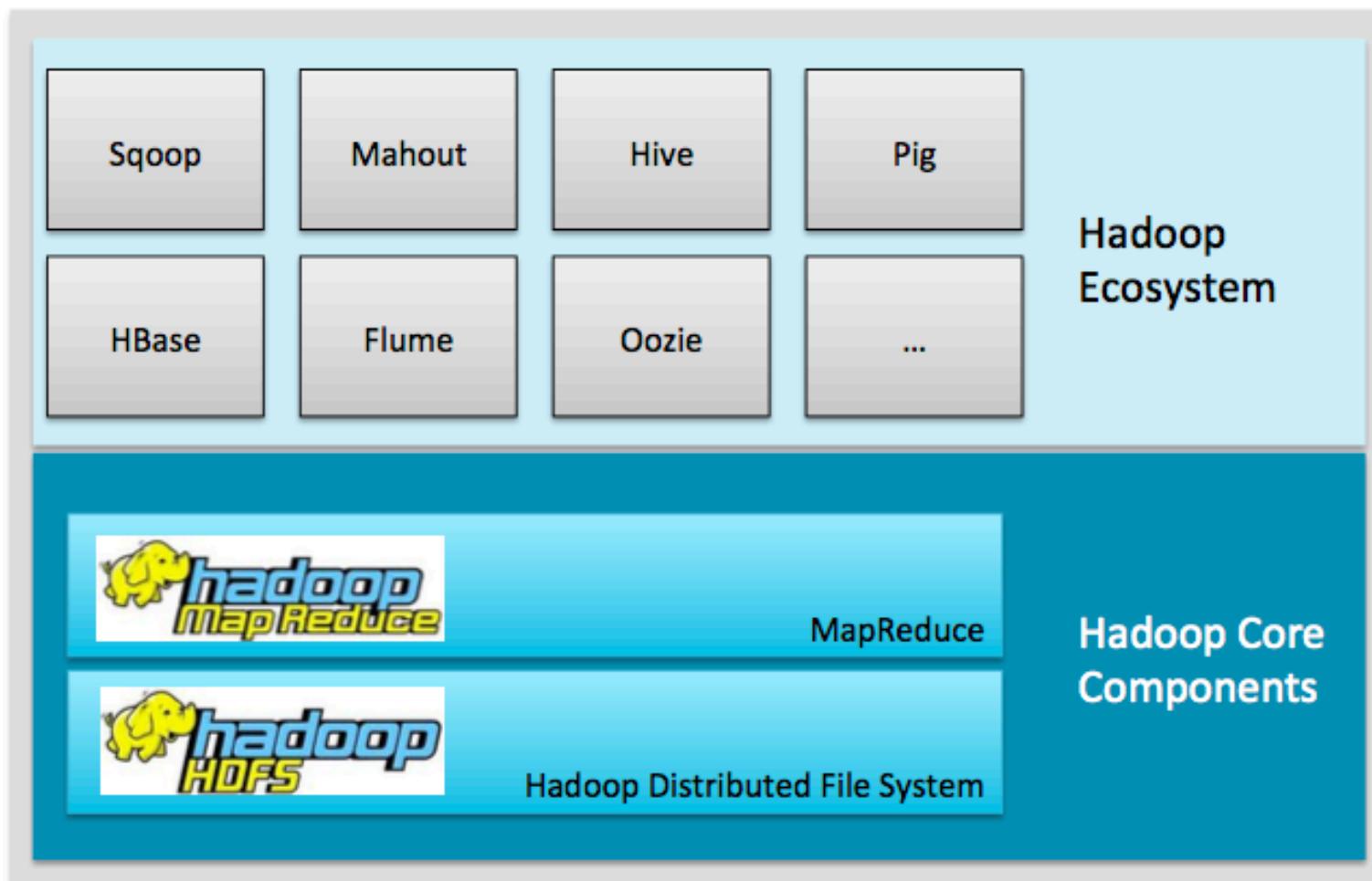
Mapper ? # Reducers ?

- The number of mappers usually depends on the input data
 - Typically 1 map per split
 - E.g., For a 1TB input file we can have 7'813 Mappers
 - when block size is 128MB
- Reducers will depend on the expected results
 - 1 Reducer
 - means 1 machine
 - All the k/v will go there, bottleneck
 - Too many reducers will create extensive network traffic
 - Strategy: set depending on the expected number of keys

MapReduce on YARN (Hadoop 2.0)

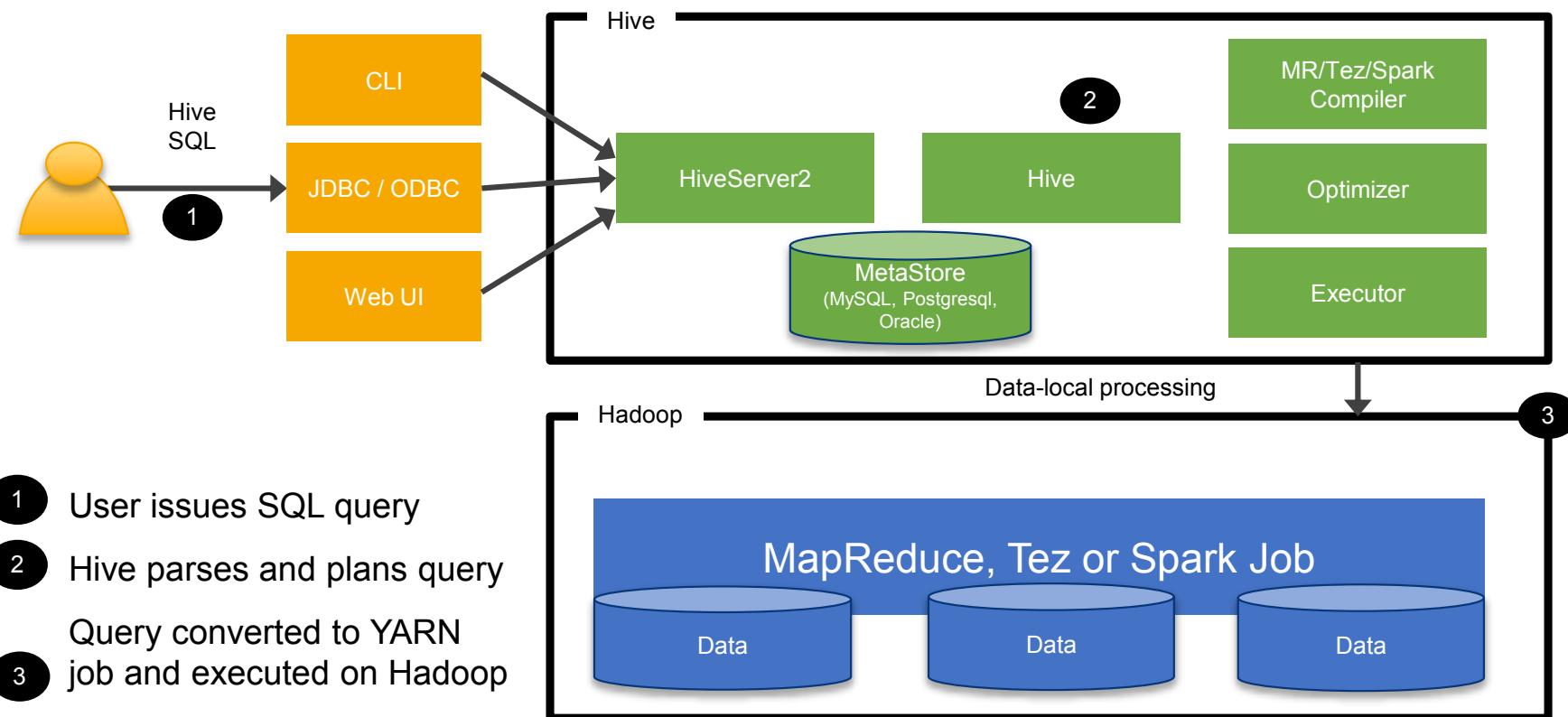


Hadoop Ecosystem



Apache Hive

- SQL (-like) queries on Hadoop



Hive-SQL Example

```
DROP TABLE IF EXISTS docs;  
CREATE TABLE docs (line STRING);  
LOAD DATA INPATH 'input_file' OVERWRITE INTO TABLE docs;  
CREATE TABLE word_counts AS  
SELECT word, count(1) AS count FROM  
    (SELECT explode(split(line, '\s')) AS word FROM docs) temp  
GROUP BY word  
ORDER BY word;
```

Thanks for your Attention!



<https://exascale.info>