Inspire…Educate…Transform.

# Applying ML to Big Data using Hadoop and Spark Ecosystem

## Day 1: Big Data Overview, Hadoop Ecosystem, HDFS

Dr. Manoj Duse
Aug 4, 2018

# Applying Machine Learning to Big Data Using Hadoop and Spark

Foundations & Distributed Storage

Resource Management & Parallel Processing

More Spark and Spark ML

Streaming

….

# Agenda

- Big Data Overview
- Use Cases
- What is Hadoop, History/Evolution
- Hadoop Ecosystem
- HDFS

# What is Big Data ?

- Big does not refer to size or volume alone !

- So what else comes into play ?

# The V's of Big Data

## Volume

Sources:

Click stream

Logs

Social Media

IoT Sensors

Text Corpus [Blogs, Proposals, emails]

# The V's of Big Data

## Velocity

- Speed at which data is generated
- Speed at which it has to be analysed for actionable insights

# The V's of Big Data

## Variety

- Structured
- Semi-structured
- Unstructured

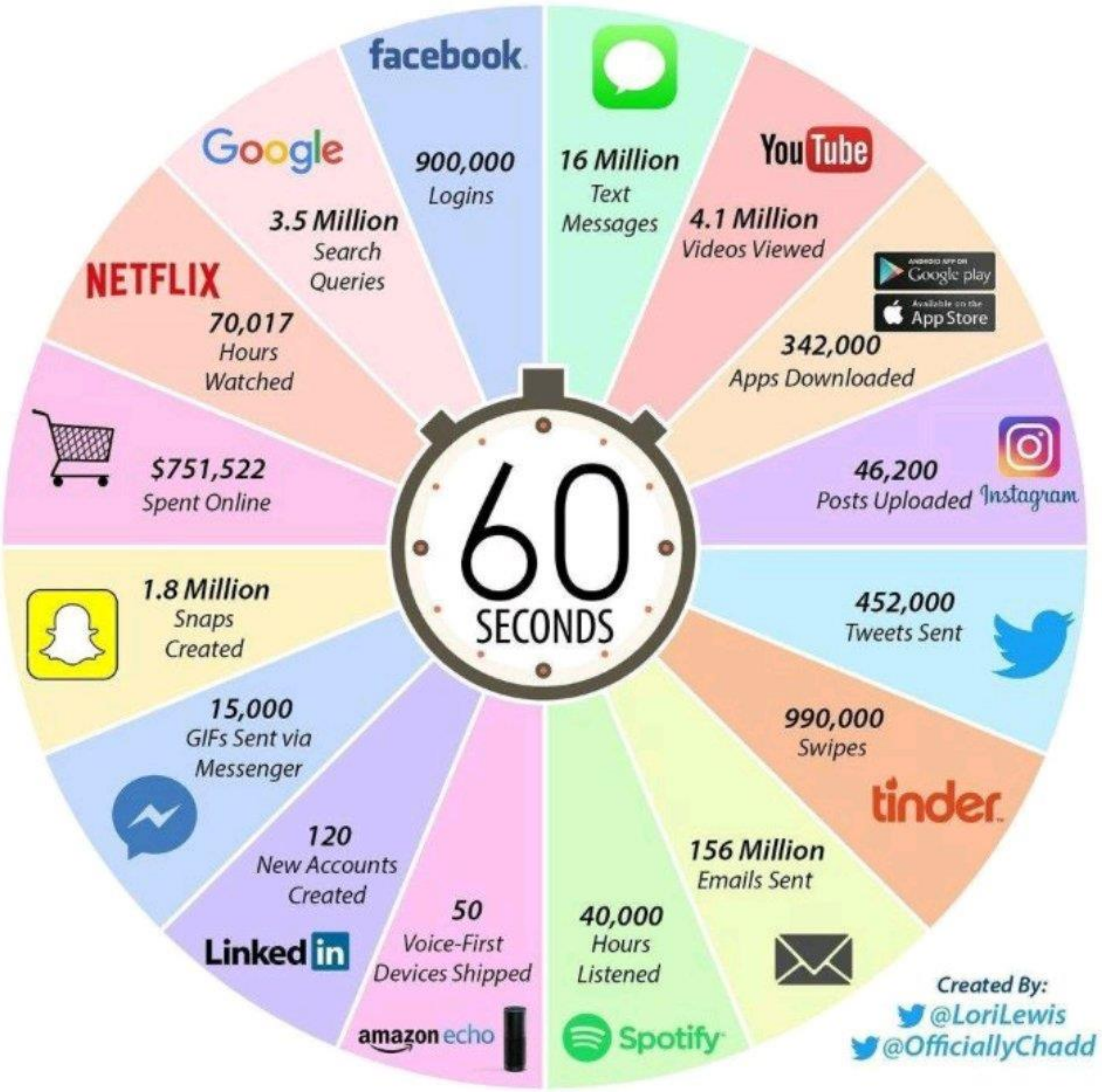
## Veracity ?

- Untrusted
- Uncleansed

# Gartner, Big Data Definition

- "Big data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation."

# 2017 This Is What Happens In An Internet Minute



# Orders of Magnitude

| Name (Symbol) | Value | Binary usage |
|---|---|---|
| kilobyte (kB) | $10^{3}$ | $2^{10}$ |
| megabyte (MB) | $10^{6}$ | $2^{20}$ |
| gigabyte (GB) | $10^{9}$ | $2^{30}$ |
| terabyte (TB) | $10^{12}$ | $2^{40}$ |
| petabyte (PB) | $10^{15}$ | $2^{50}$ |
| exabyte (EB) | $10^{18}$ | $2^{60}$ |
| zettabyte (ZB) | $10^{21}$ | $2^{70}$ |
| yottabyte (YB) | $10^{24}$ | $2^{80}$ |

# Let us understand the “scale” of data in today’s world

- Google is the largest ‘big data’ company in the world, processing 3.5 billion requests per day, storing 10 Exabytes of data.

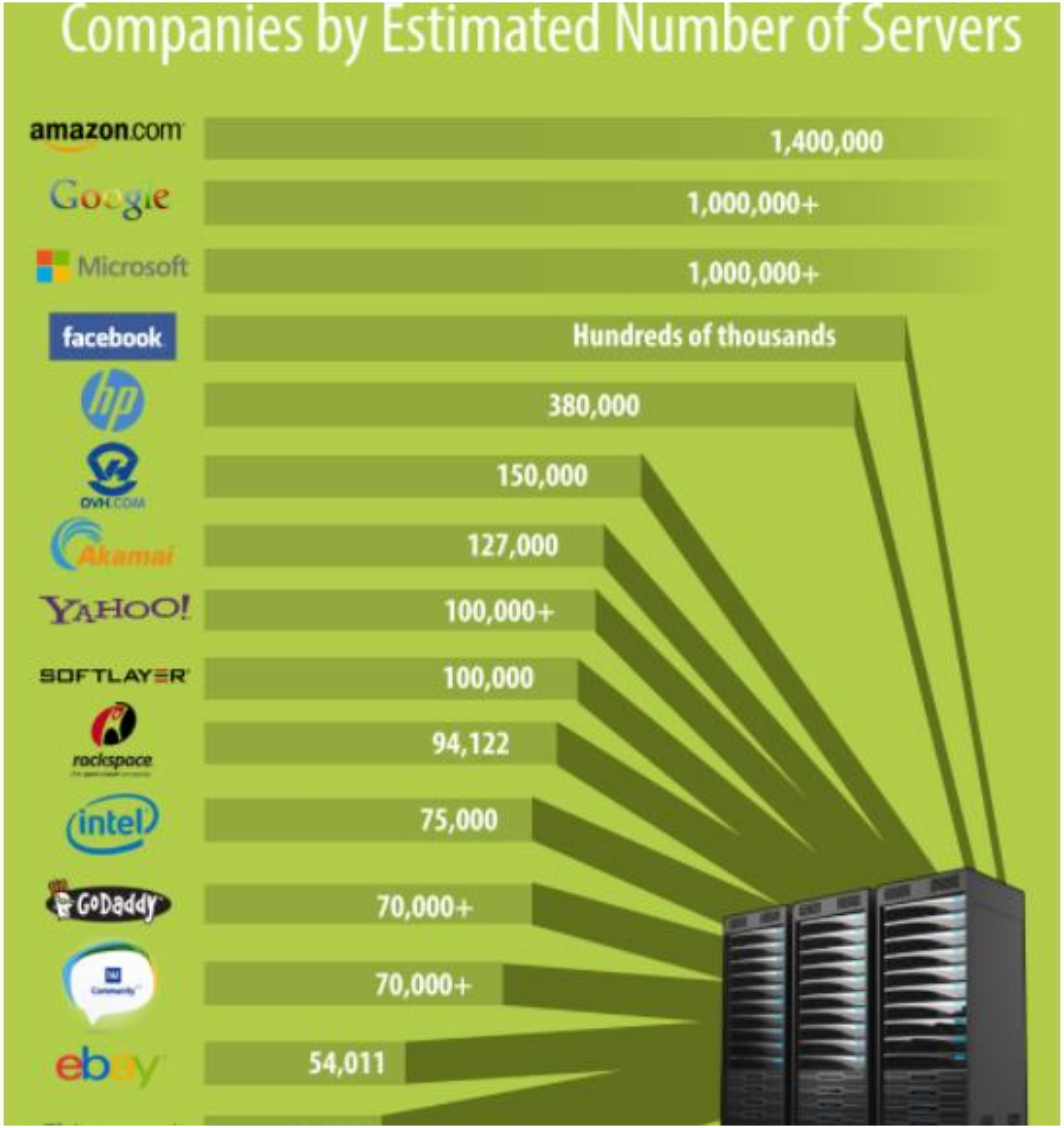
- Amazon hosts the most servers of any company, estimated at 1,400,000 servers with Google and Microsoft close behind.

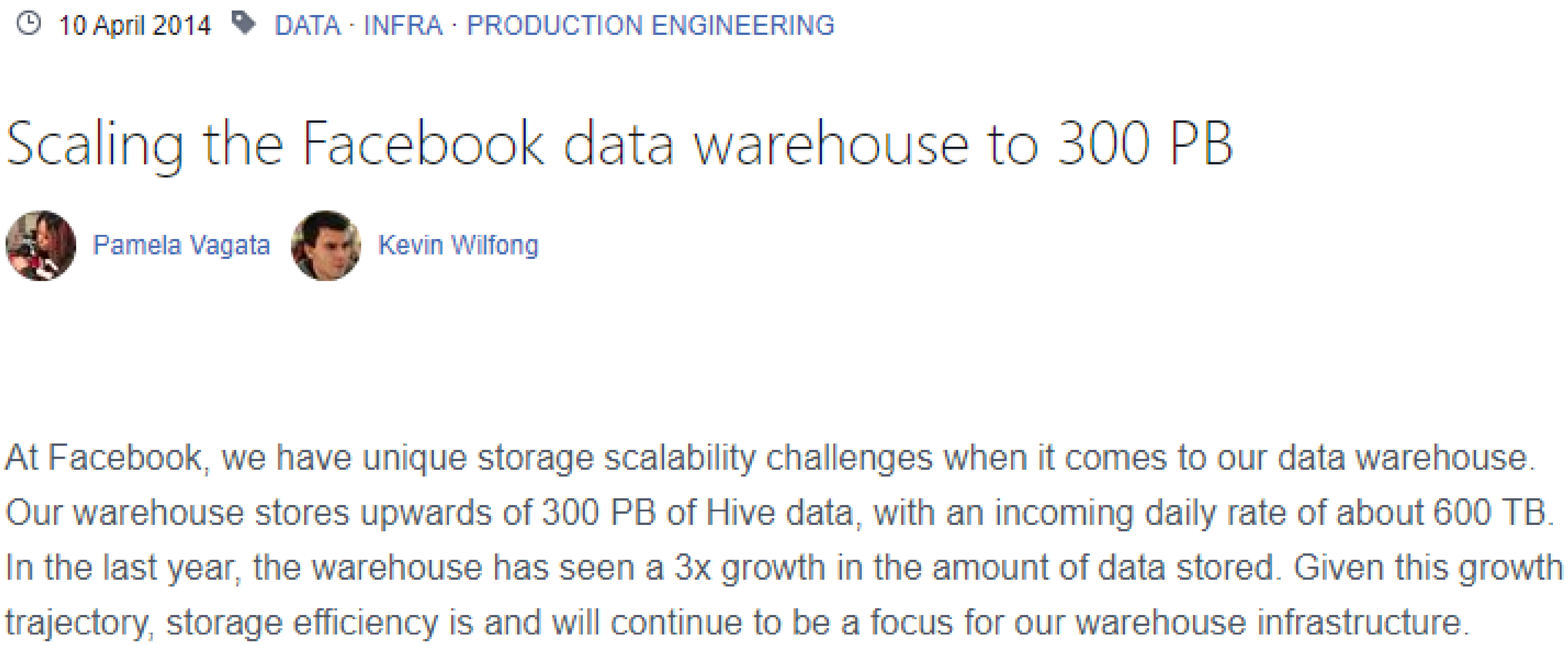
- Amazon Web Services (AWS) are used by 60,000 companies and field more than 650,000 requests every second.

- Facebook collects 500 terabytes of data daily, including 2.5 billion pieces of content, 2.7 billion likes and 300 million photos.

- 90% of all the data in the world was produced in the last 2 years.
- It is estimated that 40 zettabytes (40,000 Exabytes) of data will be created by 2020. 3.2 zettabytes [in 2014]
- The total amount of data being captured and stored by industry doubles every 1.2 years

# How big is Big?





10 April 2014 DATA · INFRA · PRODUCTION ENGINEERING

# Scaling the Facebook data warehouse to 300 PB

Pamela Vagata Kevin Wilfong

At Facebook, we have unique storage scalability challenges when it comes to our data warehouse. Our warehouse stores upwards of 300 PB of Hive data, with an incoming daily rate of about 600 TB. In the last year, the warehouse has seen a 3x growth in the amount of data stored. Given this growth trajectory, storage efficiency is and will continue to be a focus for our warehouse infrastructure.

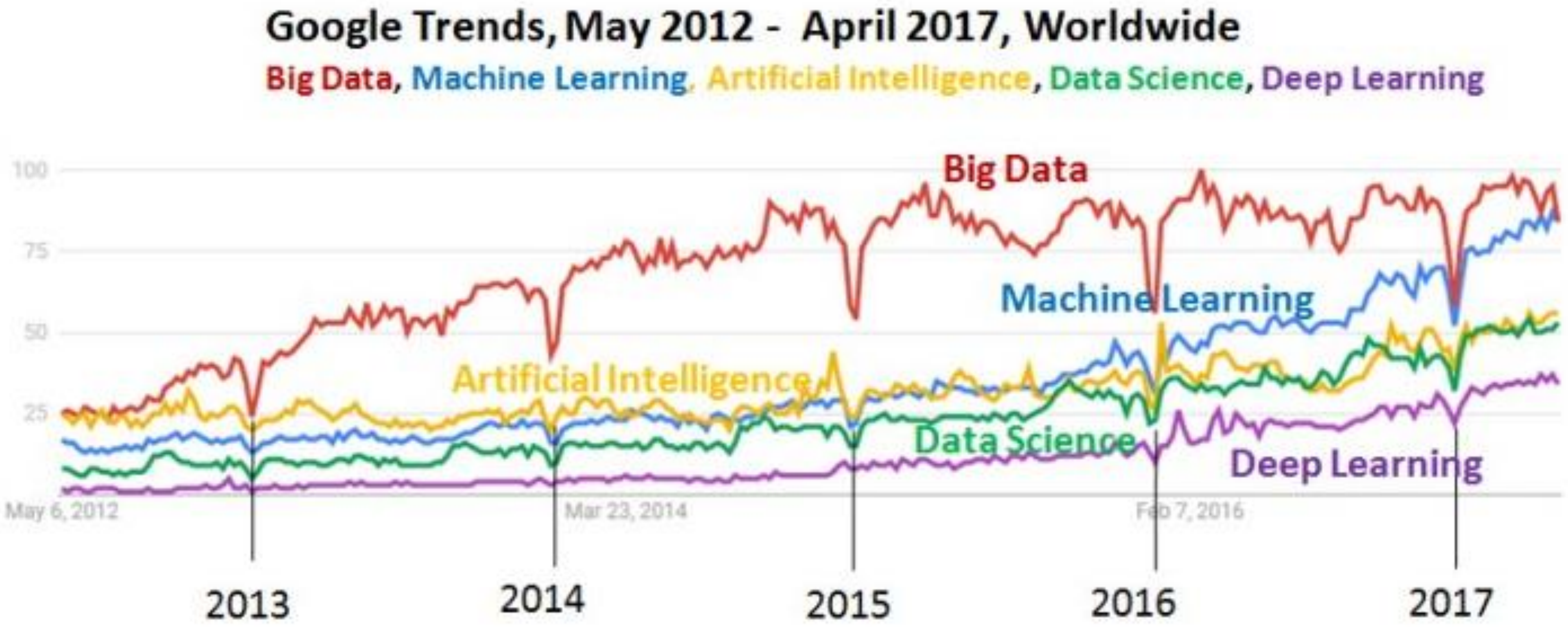


Fig. 3: Google Trends, May 2012 - April 2017, Worldwide
"Big Data" vs "Machine Learning" vs "Artificial Intelligence" vs "Data Science" vs "Deep Learning" search terms.

# Premise

- The more data you do data science with, the finer (and better) your insights will be.

# Big data - model building issues

- Data ought to be in primary storage, or even better, RAM
- Programmers ought to see the storage as monolithic.
    - Resource Management: YARN, Mesos
- “Serially written” programs ought to run in parallel.
    - Map Reduce, MR2, Spark, BSP, Flink, …
- There ought to be faster, more reliable ways of bringing in much more data
    - Data ingestion methods – Sqoop, Flume, Kafka…

# What are big data use cases?

# Big Data at CERN

- https://www.youtube.com/watch?v=j-0cUmUyb-Y

- Physicists at CERN have been pondering how to store and share their ever more massive data for decades - stimulating globalization of the internet along the way, whilst 'solving' their big data problem.

- Tim Smith plots CERN's involvement with big data from fifty years ago to today.

- Lesson by Tim Smith, animation by TED-Ed.

# One of the Google data centres powering the cloud

- https://www.youtube.com/watch?v=zDAYZU4A3w0

# Enter the world of “Hadoop”

- Hadoop is Apache Project

- What is Apache …ASF ?

## WHAT IS THE APACHE SOFTWARE FOUNDATION?

The Apache Software Foundation (ASF) is a non-profit corporation [incorporated in1999 ]

The ASF is a natural outgrowth of The Apache Group, a group of individuals that was initially formed in 1995 to develop the Apache HTTP Server.

The all-volunteer ASF develops, stewards, and incubates more than 350 Open Source projects and initiatives that cover a wide range of technologies.

## WHY WAS THE APACHE SOFTWARE FOUNDATION CREATED?

1.provide a foundation for open, collaborative software development projects

2.create an independent legal entity to which companies and individuals can donate resources and be assured that those resources will be used for the public benefit;

3.provide a means for individual volunteers to be sheltered from legal suits directed at the ASF projects;

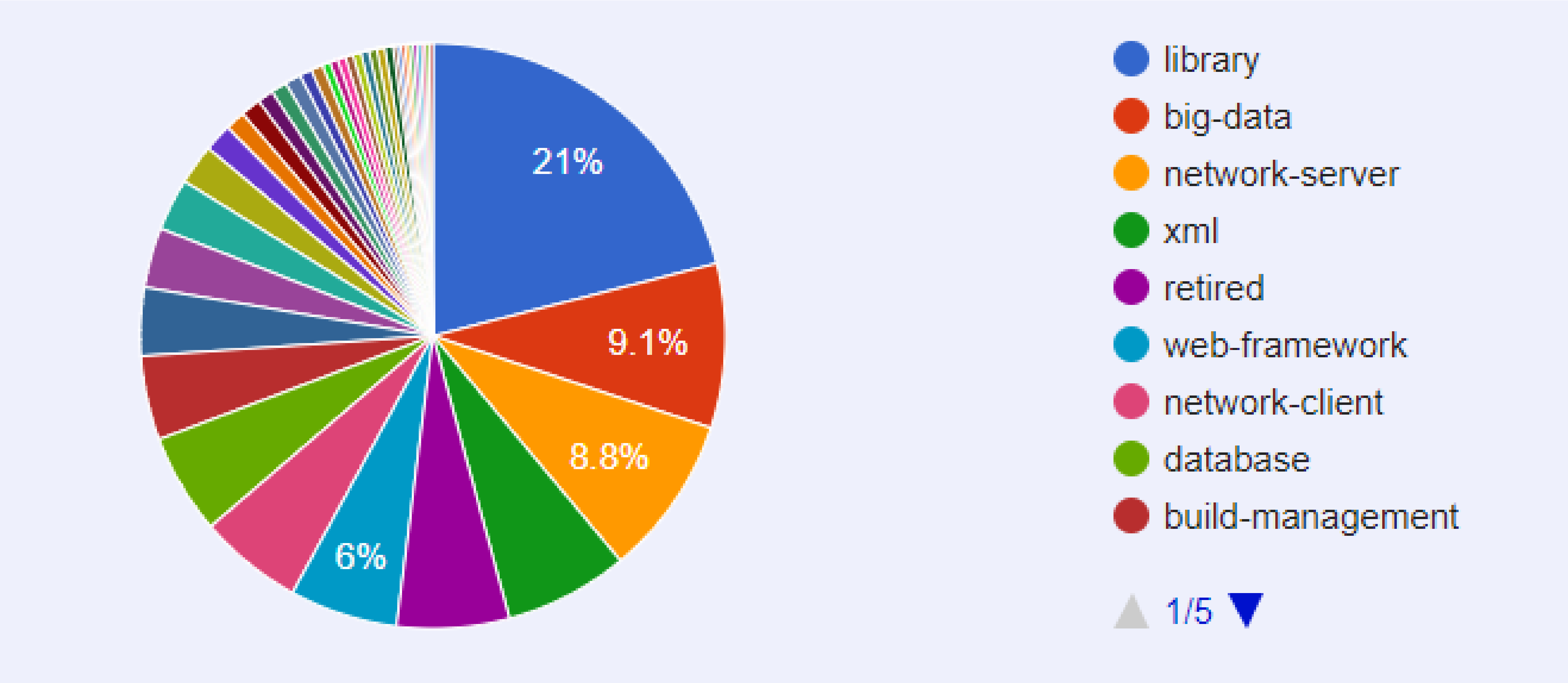
4.protect the 'Apache' brand, as applied to its software products, from being abused

## WHY WAS THE NAME 'APACHE' CHOSEN?

The name 'Apache' was chosen from respect for the various Native American nations collectively referred to as Apache, well-known for their superior skills in warfare strategy and their inexhaustible endurance.

It also makes a cute pun on "a patchy web server" -- a server made from a series of patches

# Big Data Share in Apache Projects



# What is Hadoop?

- Hadoop is an open source framework, from the Apache foundation, capable of processing large amounts of heterogeneous data sets in a distributed fashion across clusters of commodity computers and hardware using a simplified programming model.

- The Hadoop framework is based closely on the following principle:

- *In pioneer days they used oxen for heavy pulling, and when one ox couldn't budge a log, they didn't try to grow a larger ox. We shouldn't be trying for bigger computers, but for more systems of computers.*

***~Grace Hopper***

# Back to Hadoop

HDFS - Reliable Shared Storage

Distributed Storage

+

MapReduce - Distributed Computation

Parallel processing

=

# Some basic terms:

- Rack, Switch
- Daemon
- Latency
- Throughput
- Scale Up, Scale Out
- Scale vertically, Scale horizontally

# Rack

The rack contains multiple mounting slots called bays,
each designed to hold a hardware unit secured in place with screws.

A single rack can contain multiple servers stacked one above the other, consolidating network resources and minimizing the required floor space.

The rack server configuration also simplifies cabling among network components.

Cooling systems become critical aspects

# Switch

- A switch, in the context of networking is a high-speed device that receives incoming data packets and redirects them to their destination on a local area network (LAN).

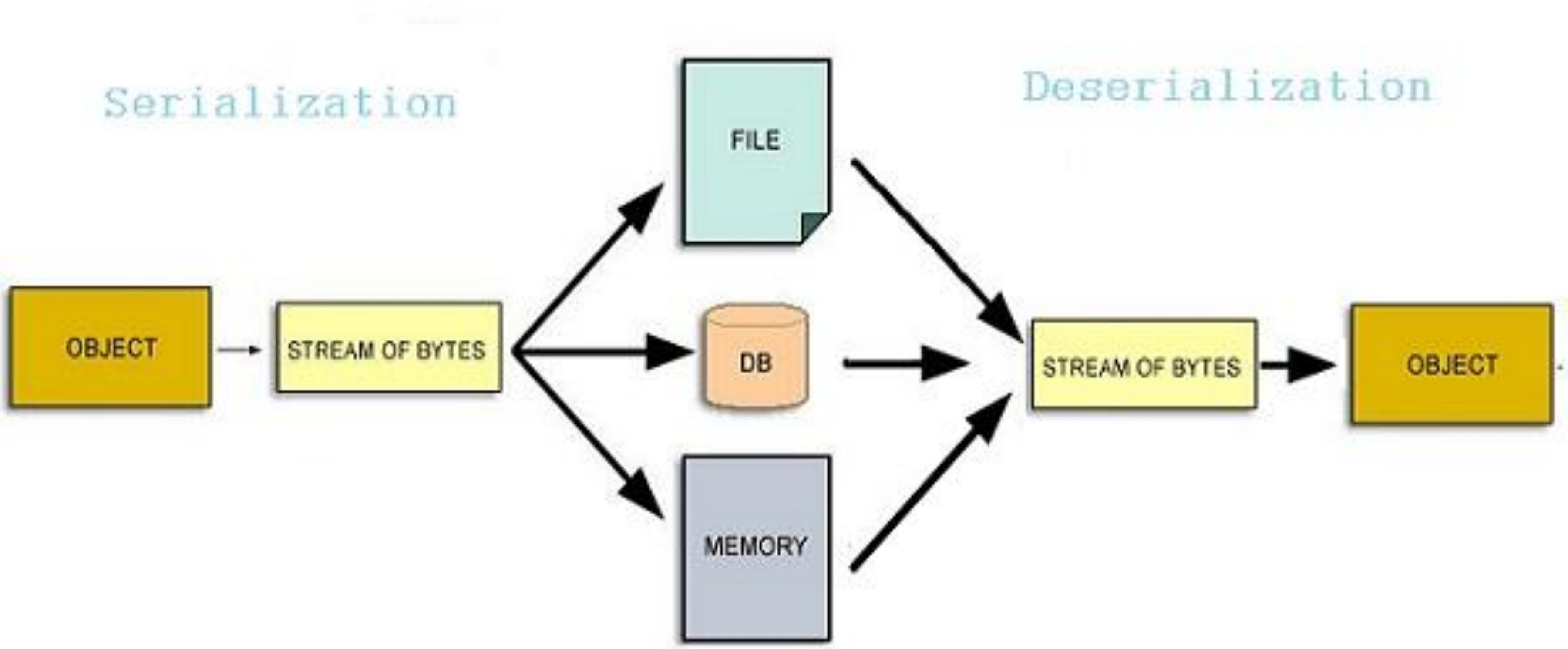
- Essentially, switches are the traffic cops of a simple local area network

- switch is limited to node-to-node communication on the same network.

# SerDe

- **Ser**ialization and **De**serialization

- What is serialization ?


Serialization
Deserialization
OBJECT
STREAM OF BYTES
FILE
DB
MEMORY
STREAM OF BYTES
OBJECT

# Serialization

- Serialization is the process of converting the state information of an object instance into a binary or textual form to persist into storage medium or transported over a network.

Worded differently…

- Serialization is the process of converting an object into a stream of bytes in order to store the object or transmit it to memory, a database, or a file. [save]

- Its main purpose is to save the state of an object in order to be able to recreate it when needed.

- The reverse process is called deserialization. [restore]

- Bandwidth :  How wide the pipe is

- Latency : How long does it take to travel from one end of pipe to the other

- Throughput : Amount flowing through per unit time

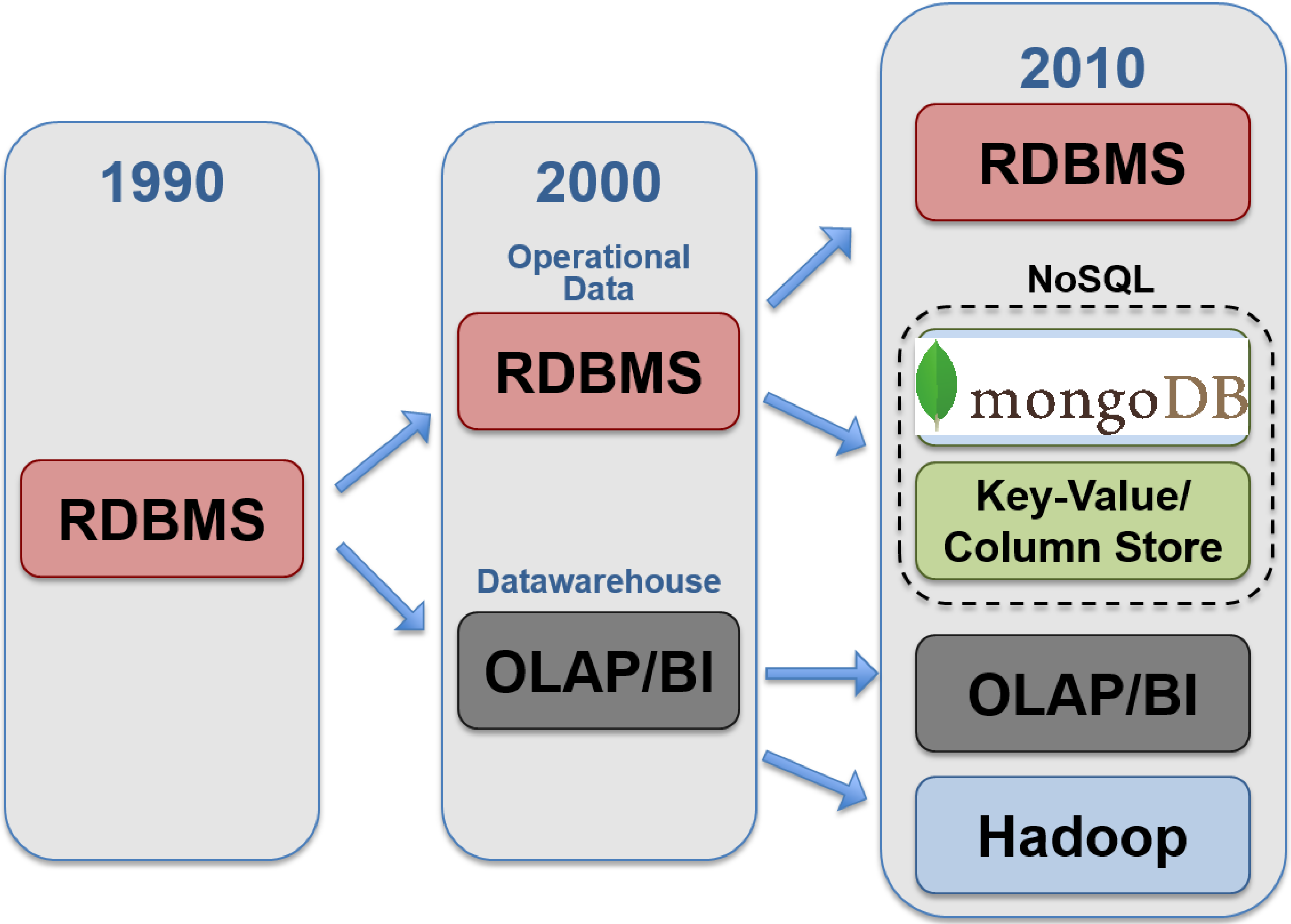
# Ways to Scale ?

- To **scale horizontally** (or **scale out**) means to add more nodes to a system, such as adding a new computer to a distributed software application.

- To **scale vertically** (or **scale up**) means to add resources to a single node in a system, typically involving the addition of CPUs or memory to a single computer.

# Let’s set the stage for next discussion:

- RDBMS
- SQL
- JDBC
- OLTP
- OLAP
- NoSQL
- Data Warehouse
- ETL
- Data Mart
- Data Lake

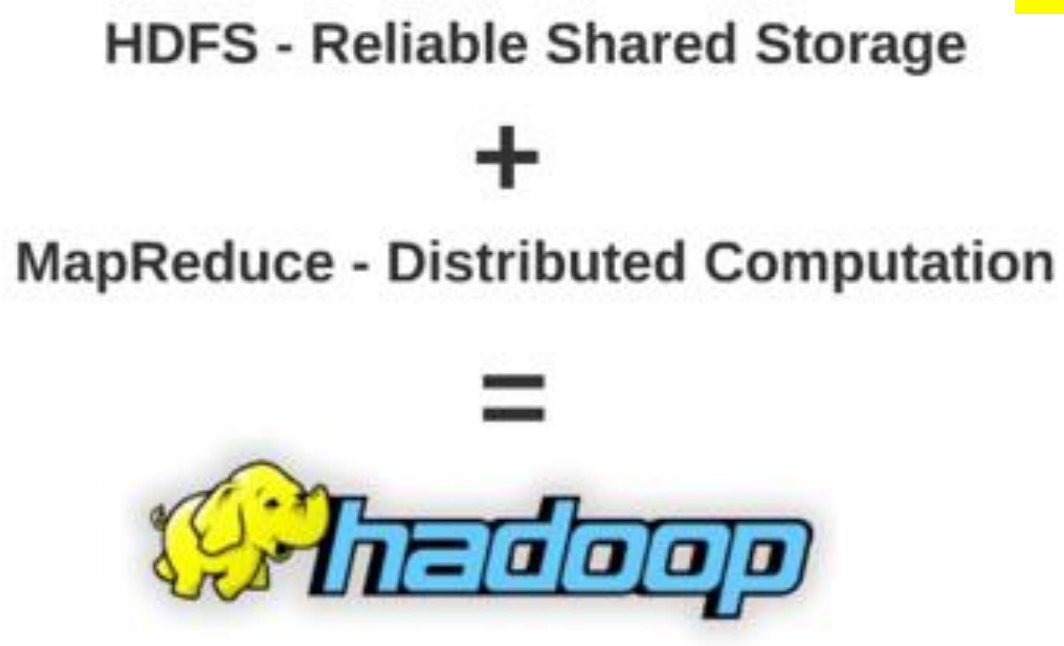


Transition from databases to data warehouses to data lakes

# Data Lakes
## *The New Big Data Thinking*

- All data has potential value
- Data hoarding
- No defined schema—stored in native format
- Schema is imposed and transformations are done at query time *(schema-on-read).*
- Apps and users interpret the data as they see fit

# Back to Hadoop



HDFS - Reliable Shared Storage

+

MapReduce - Distributed Computation

=

Distributed Storage

Parallel processing

# Hadoop Characteristics:

- Distribute data initially
  - Let processors / nodes work on local data
  - Minimize data transfer over network
  - Replicate data multiple times for increased availability
- Write applications at a high level
  - Programmers should not have to worry about network programming, low level infrastructure, etc
- Minimize talking between nodes (*share-nothing)*

# History of Hadoop

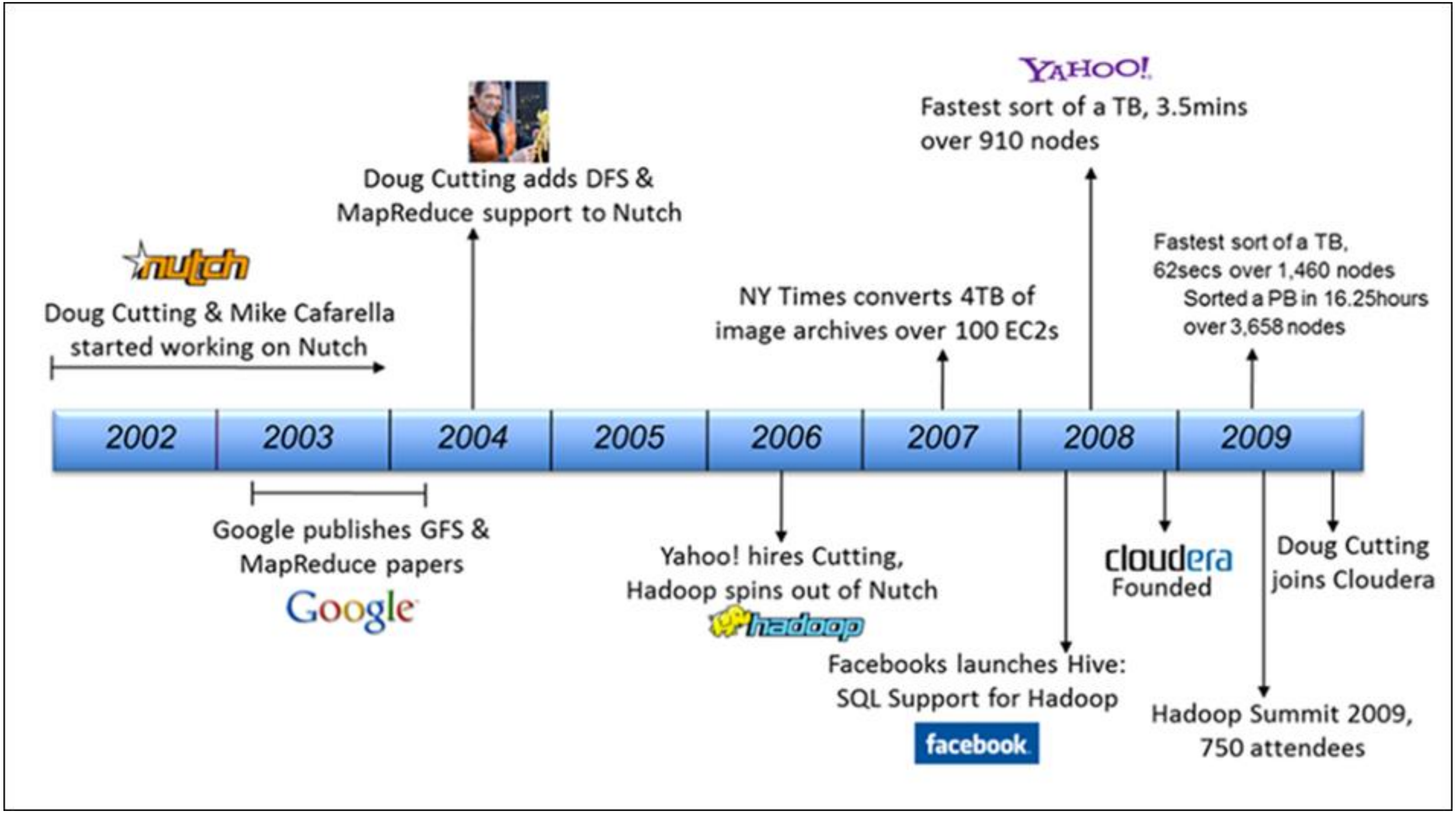
- Hadoop was created by Doug Cutting and Mike Cafarella.

- Originated from an open source web search engine called "Apache Nutch", which is part of another Apache project called "Apache Lucene“

- What does HADOOP stand for ?

- According to Hadoop's creator Doug Cutting, the name came about as follows:
- "*The name my kid gave a stuffed yellow elephant. Short, relatively easy to spell and pronounce, meaningless, and not used elsewhere: those are my naming criteria.”*

# Hadoop Timeline



2002 2003 2004 2005 2006 2007 2008 2009

Doug Cutting & Mike Cafarella started working on Nutch

Google publishes GFS & MapReduce papers

Doug Cutting adds DFS & MapReduce support to Nutch

Yahoo! hires Cutting, Hadoop spins out of Nutch

NY Times converts 4TB of image archives over 100 EC2s

Facebooks launches Hive: SQL Support for Hadoop

YAHOO! Fastest sort of a TB, 3.5mins over 910 nodes

cloudera Founded

Fastest sort of a TB, 62secs over 1,460 nodes
Sorted a PB in 16.25hours over 3,658 nodes

Hadoop Summit 2009, 750 attendees

Doug Cutting joins Cloudera

2011:
HW
And
MapR

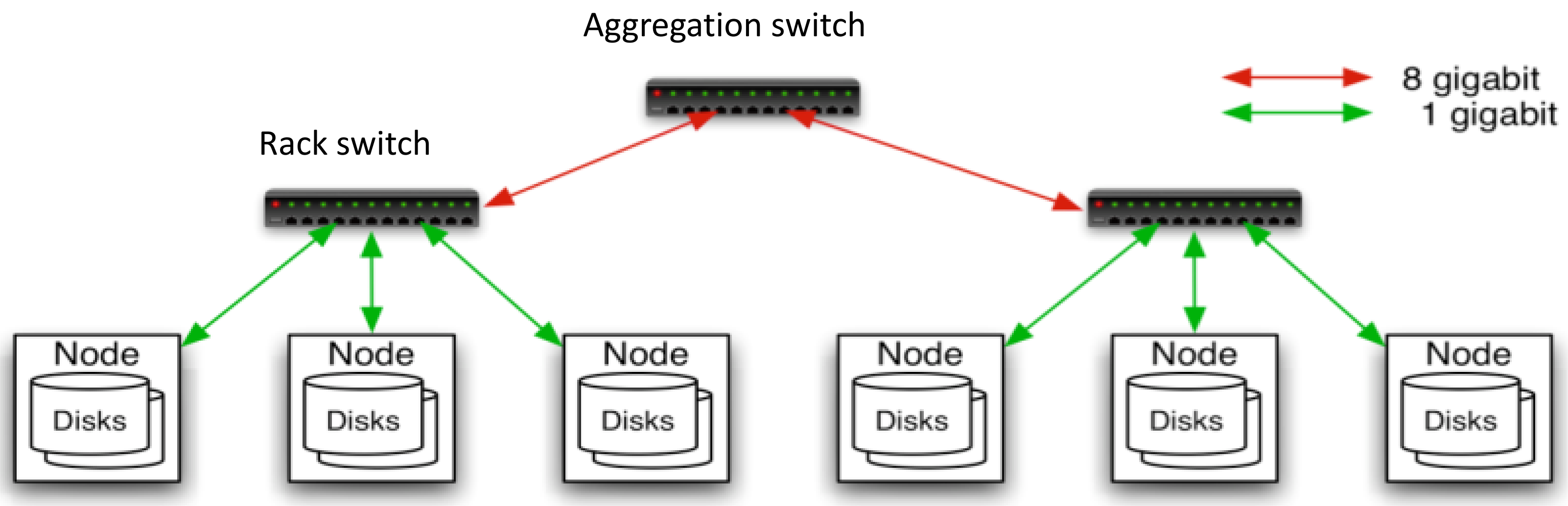
# Doug Cutting Basics of Hadoop Video

https://www.youtube.com/watch?v=0GOxDBR6VAU

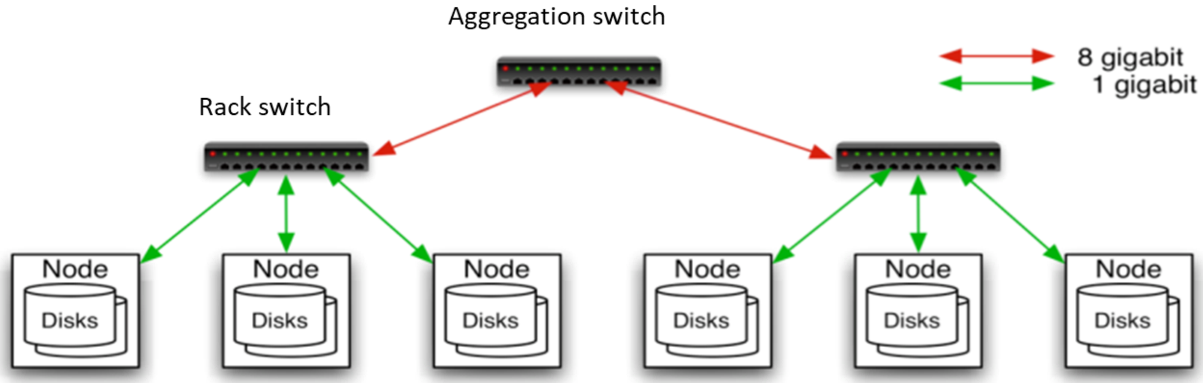
# Why Is Hadoop Important?

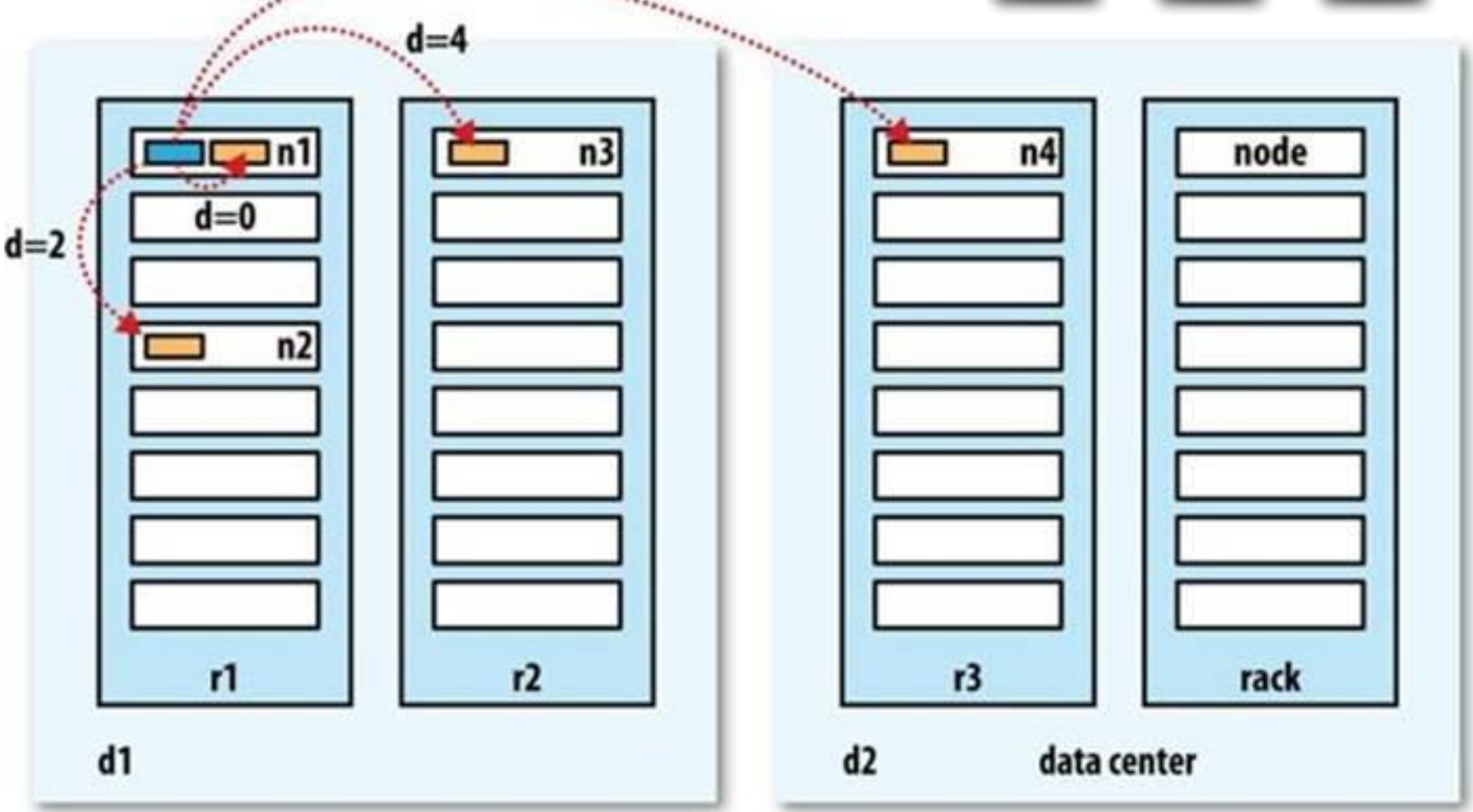
- Ability to store and process huge amounts of any kind of data, quickly.
- Computing model processes big data fast
- Fault tolerance
- Flexibility
- Low Cost
- Scalability

# Hub & Spoke Hardware



- Typically in 2 level architecture
  - Nodes are commodity PCs
  - Typically 30-40 nodes/rack





# Notion of Distance [wrt data transfer]

# The “Big Data Bazaar”

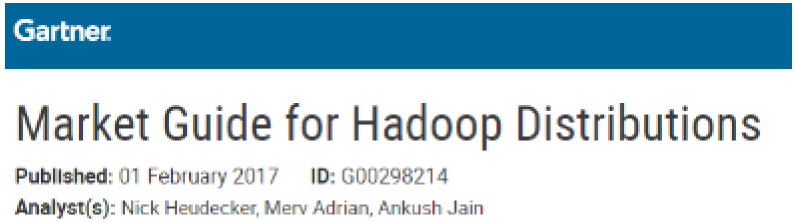
# The different options:

| On-Premise | IaaS | PaaS | SaaS |
|---|---|---|---|
| | | | 🟩 Customizations |
| 🟩 Applications | 🟩 Applications | 🟩 Applications | 🟪 Applications |
| 🟩 Data | 🟩 Data | 🟩 Data | 🟪 Data |
| 🟩 Runtime | 🟩 Runtime | 🟪 Runtime | 🟪 Runtime |
| 🟩 Middleware | 🟩 Middleware | 🟪 Middleware | 🟪 Middleware |
| 🟩 Operating System | 🟩 Operating System | 🟪 Operating System | 🟪 Operating System |
| 🟩 Virtualization | 🟪 Virtualization | 🟪 Virtualization | 🟪 Virtualization |
| 🟩 Servers | 🟪 Servers | 🟪 Servers | 🟪 Servers |
| 🟩 Storage | 🟪 Storage | 🟪 Storage | 🟪 Storage |
| 🟩 Networking | 🟪 Networking | 🟪 Networking | 🟪 Networking |

🟩 Customer Managed 🟪 Vendor Managed

# Key Hadoop Vendors [By Type]

- Pure play Hadoop vendors: Cloudera, Hortonworks, MapR, IBM OpenPlatform, Huawei FusionInsight, Seabox, Transwarp
- Cloud infrastructure as a service (IaaS): Hadoop on AWS, Hadoop on Azure
- Platform as a service (PaaS): IBM BigInsights, Microsoft HDInsight, Google Cloud Platform, Amazon EMR, Oracle Big Data Cloud Service, Qbole
- Big Data Appliances: Teradata, Oracle Big Data Appliance, Cray



Gartner

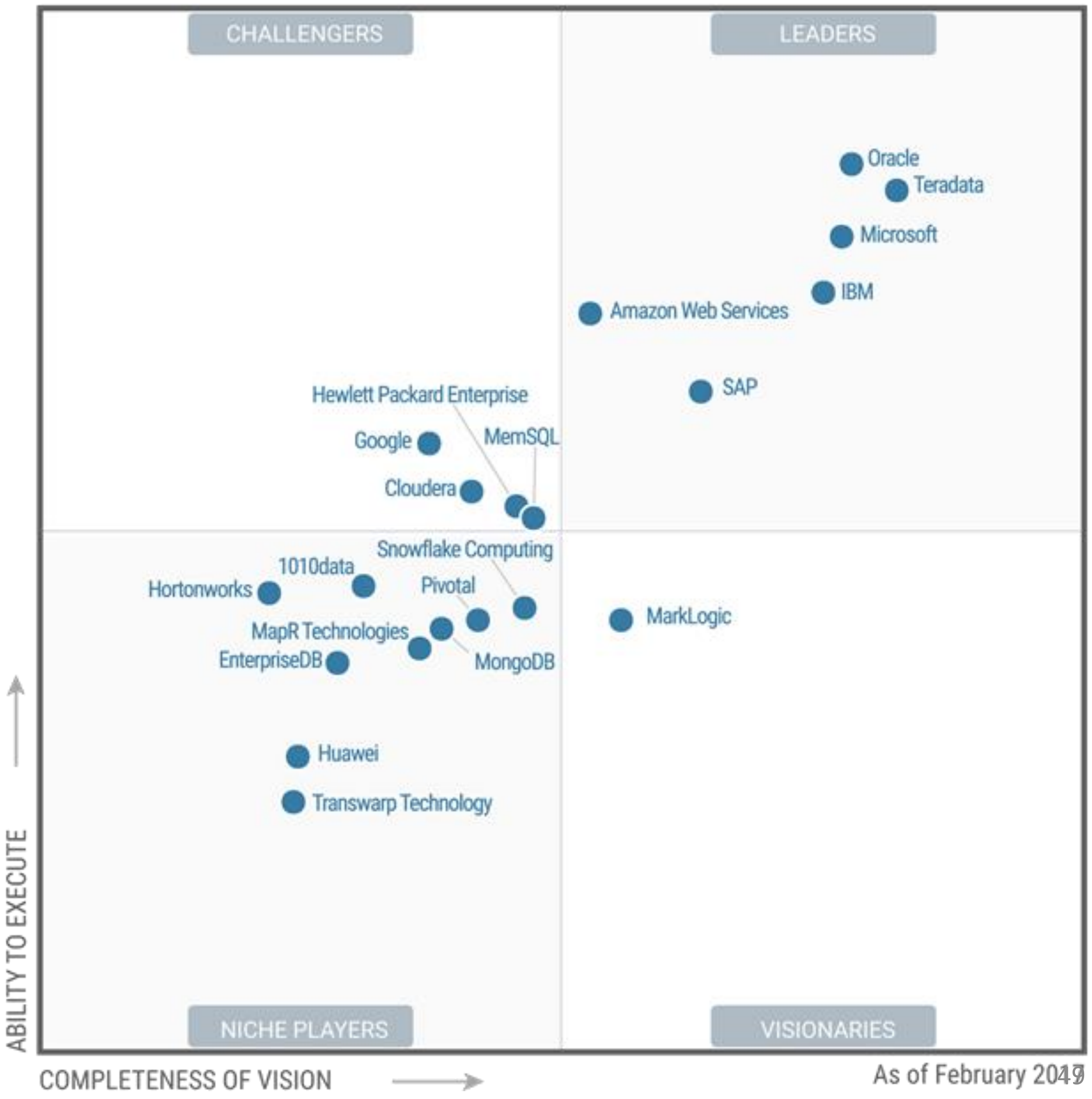
Market Guide for Hadoop Distributions

**Published:** 01 February 2017 **ID:** G00298214

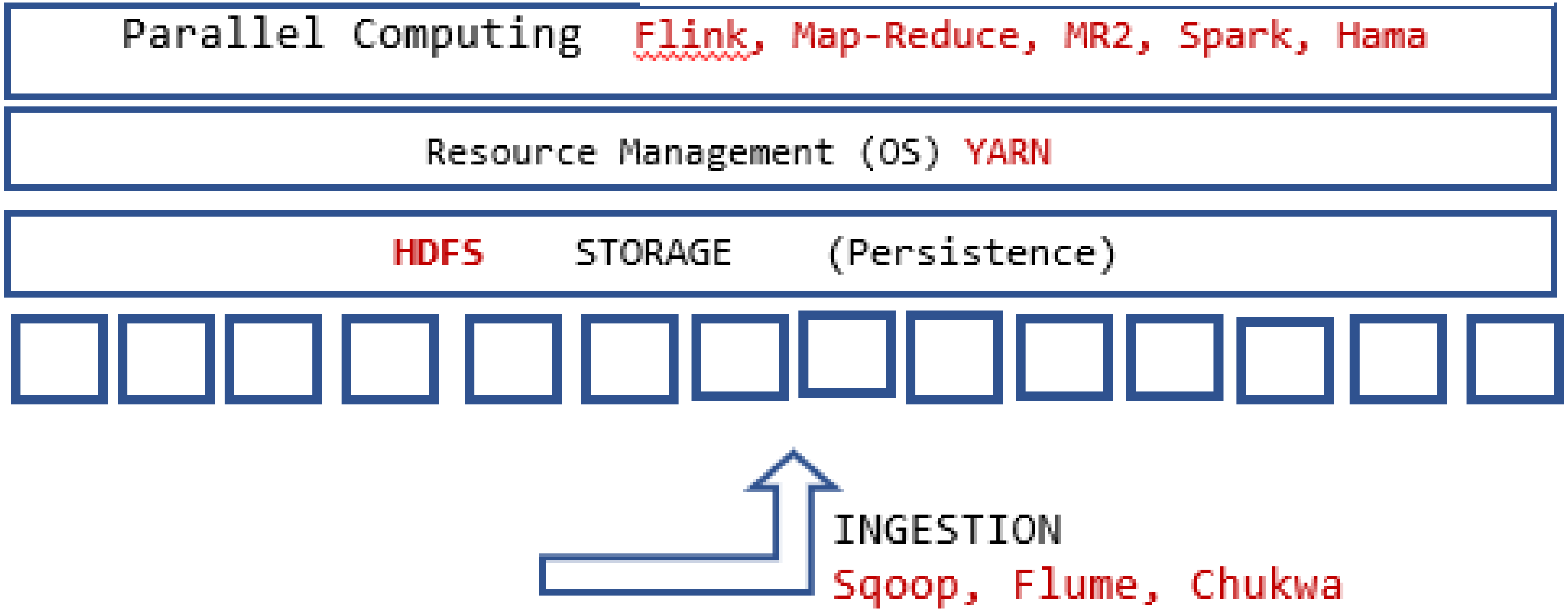
**Analyst(s):** Nick Heudecker, Merv Adrian, Ankush Jain

# Market Position of Hadoop Vendors

**Gartner Magic Quadrant for Data Management Solutions for Analytics (DMSA) – 2017**



# Everyone needs them.....



Parallel Computing Flink, Map-Reduce, MR2, Spark, Hama

Resource Management (OS) YARN

HDFS STORAGE (Persistence)

INGESTION
Sqoop, Flume, Chukwa

# Pick what you need

Machine Learning on Spark-ML, Mahout, Samsara, H2O, Hadoop

Streaming & Near Real Time Processing
KAFKA, SAMZA, STORM, TRIDENT, SPARK-STREAMING, FLINK

Application Programming
PIG, Oozie, Hadoop Streaming, Spark-R

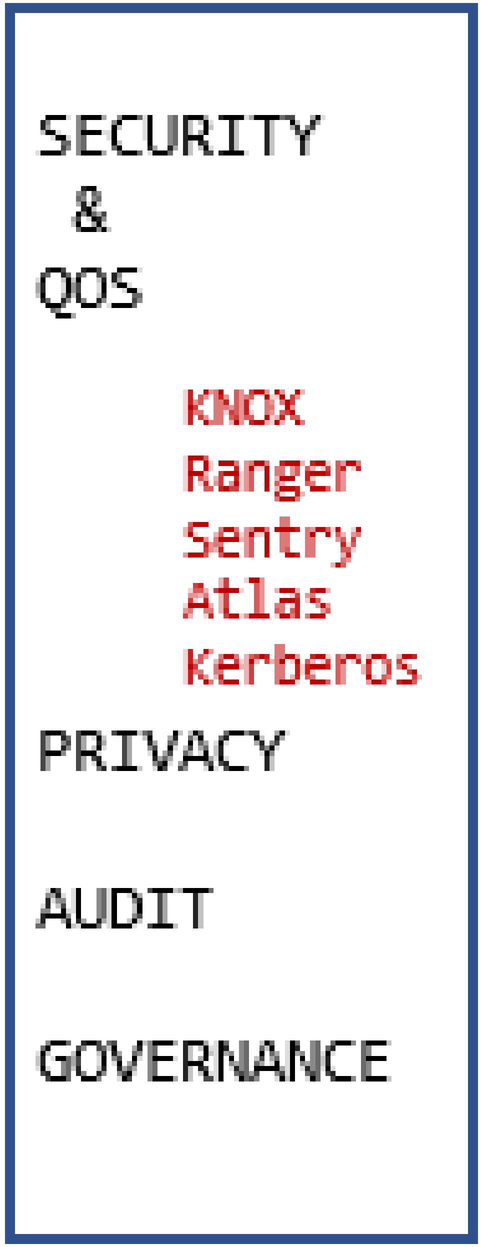
Data Organization

SQL
HIVE, IMPALA, SPARK SQL, Apache Drill

----------------------------------------

NoSQL
Hbase, Cassandra, MongoDB, Neo4J, Kudu

# Security, Audit, Governance



SECURITY
&
QOS

KNOX
Ranger
Sentry
Atlas
Kerberos

PRIVACY

AUDIT

GOVERNANCE

# Primary Content of Hadoop Distributions: Apache Projects

All major Hadoop vendors support these Apache projects:

HDFS, MapReduce, YARN, Pig, Hive, Hbase, ZooKeeper, Avro, Flume, Kafka, Oozie, Parquet, Solr, Spark and Sqoop.

Select vendors support these Apache projects:

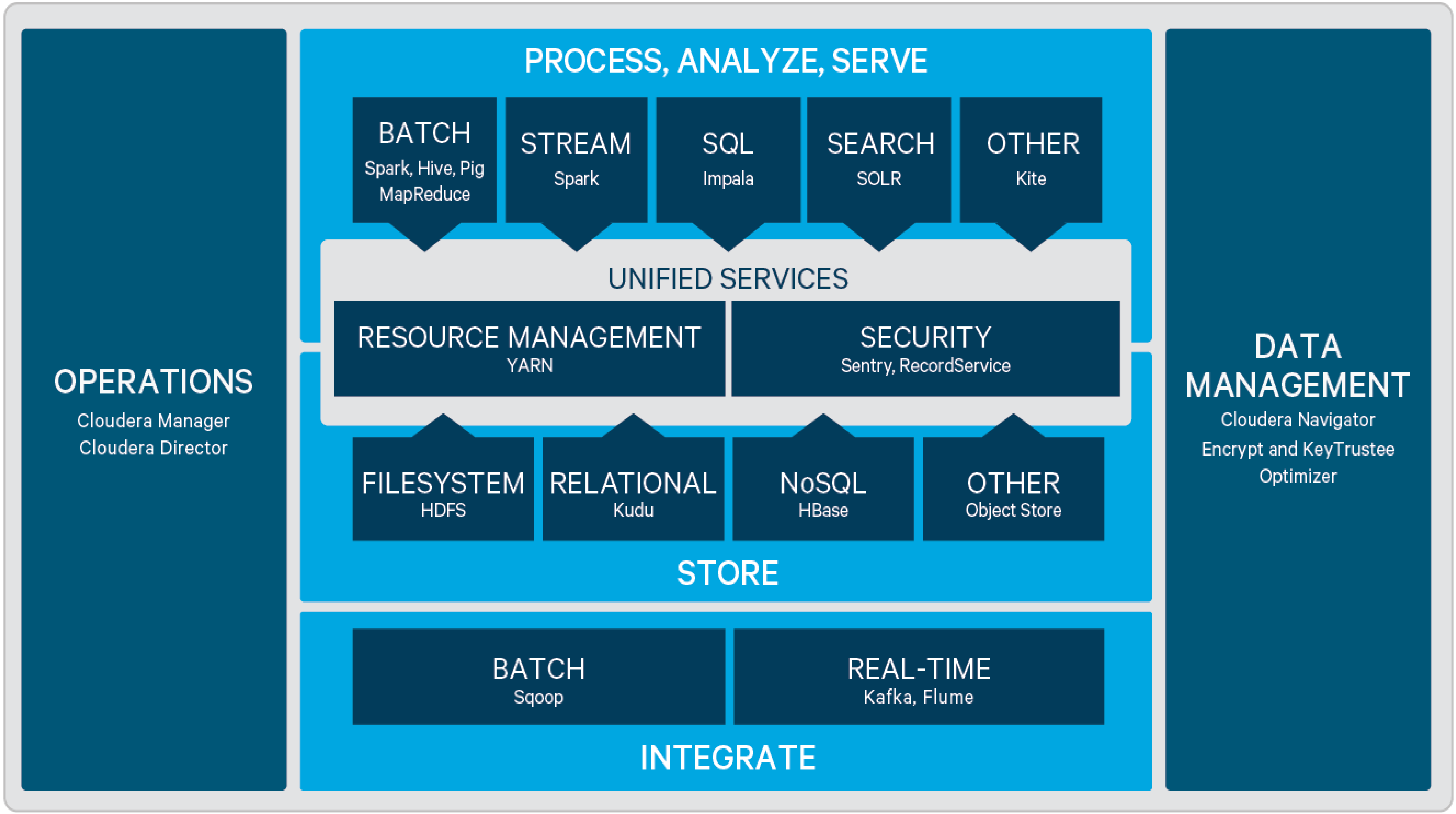
Accumulo, Ambari, Atlas, Impala, Knox, Myriad, NiFi, Phoenix, Slider, and Zeppelin

These projects show great promise:

Apache Arrow, Beam, Flink, Ignite and Kylin

Details & More: https://hadoopecosystemtable.github.io/

# Sample Hadoop Distribution: Cloudera



OPERATIONS
Cloudera Manager
Cloudera Director

PROCESS, ANALYZE, SERVE

BATCH
Spark, Hive, Pig
MapReduce

STREAM
Spark

SQL
Impala

SEARCH
SOLR

OTHER
Kite

UNIFIED SERVICES

RESOURCE MANAGEMENT
YARN

SECURITY
Sentry, RecordService

FILESYSTEM
HDFS

RELATIONAL
Kudu

NoSQL
HBase

OTHER
Object Store

STORE

BATCH
Sqoop

REAL-TIME
Kafka, Flume

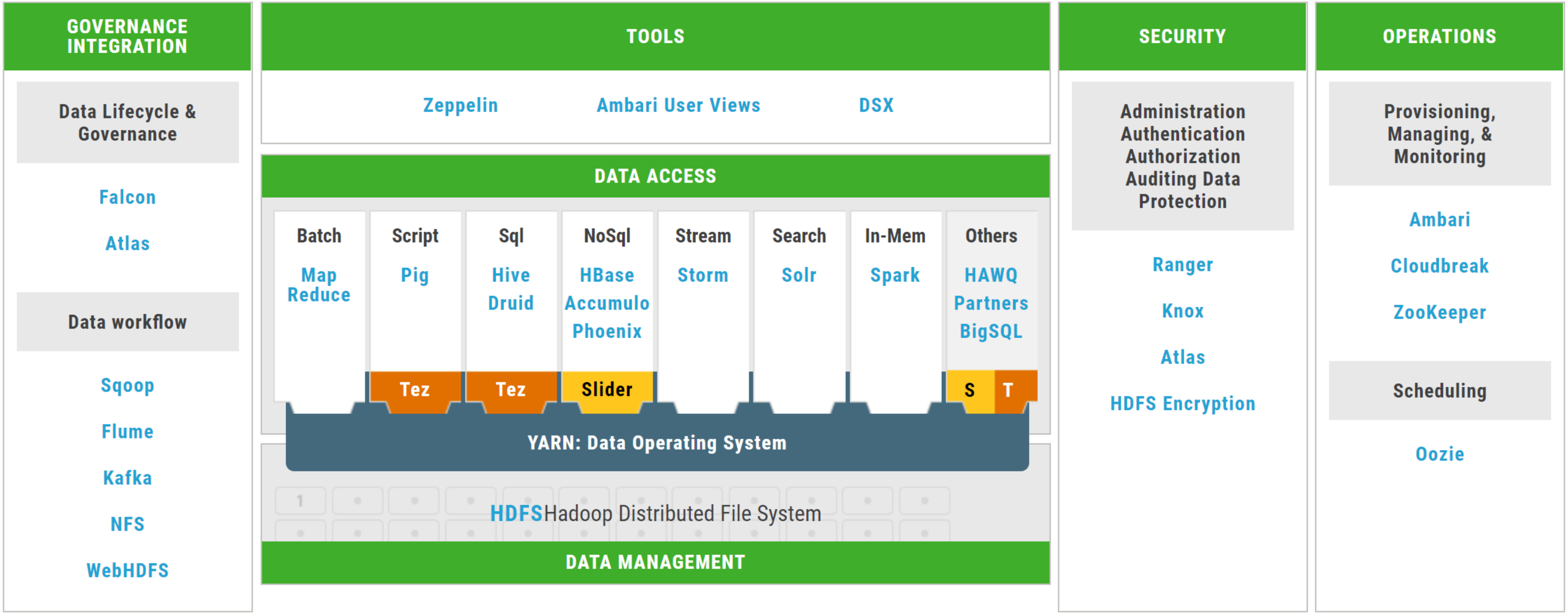
INTEGRATE

DATA MANAGEMENT
Cloudera Navigator
Encrypt and KeyTrustee
Optimizer

Key Components of Cloudera's CDH (100% open source)

- Hive, HBase
- Pig
- Kafka
- Spark
- Sqoop, Flume
- Impala
- Accumulo
- Scalability

# Hortonworks



GOVERNANCE INTEGRATION

Data Lifecycle & Governance

Falcon

Atlas

Data workflow

Sqoop

Flume

Kafka

NFS

WebHDFS

TOOLS

Zeppelin

Ambari User Views

DSX

DATA ACCESS

| Batch | Script | Sql | NoSql | Stream | Search | In-Mem | Others |
|---|---|---|---|---|---|---|---|
| Map Reduce | Pig | Hive<br>Druid | HBase<br>Accumulo<br>Phoenix | Storm | Solr | Spark | HAWQ<br>Partners<br>BigSQL |
|  | Tez | Tez | Slider |  |  |  | S T |

YARN: Data Operating System

1

HDFS Hadoop Distributed File System

DATA MANAGEMENT

SECURITY

Administration Authentication Authorization Auditing Data Protection

Ranger

Knox

Atlas

HDFS Encryption

OPERATIONS

Provisioning, Managing, & Monitoring

Ambari

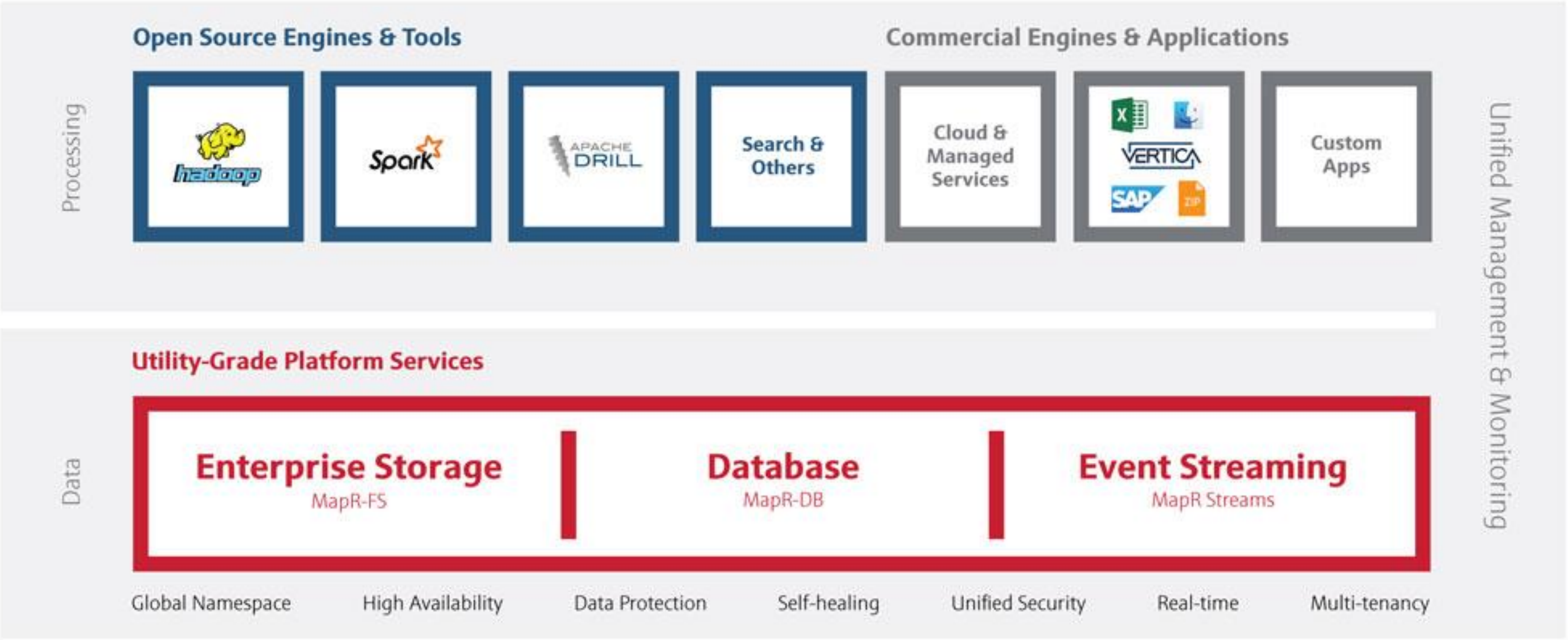
Cloudbreak

ZooKeeper

Scheduling

Oozie

# MapR Converged Data Platform



Processing

Open Source Engines & Tools

Search & Others

Commercial Engines & Applications

Cloud & Managed Services

Custom Apps

Data

Utility-Grade Platform Services

Enterprise Storage
MapR-FS

Database
MapR-DB

Event Streaming
MapR Streams

Global Namespace

High Availability

Data Protection

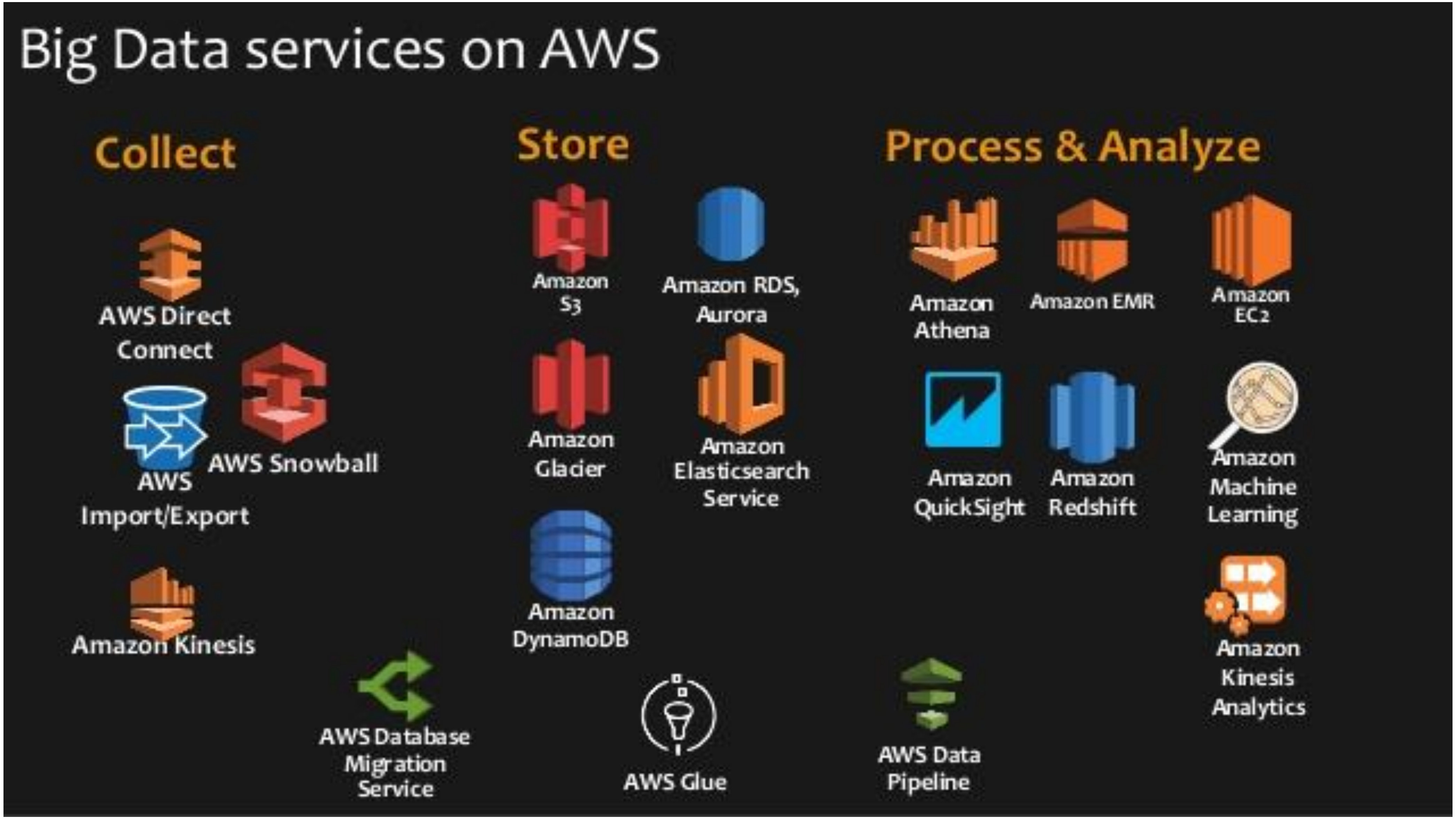
Self-healing

Unified Security

Real-time

Multi-tenancy

Unified Management & Monitoring



# Big Data services on AWS

## Collect

AWS Direct Connect

AWS Import/Export

AWS Snowball

Amazon Kinesis

## Store

Amazon S3

Amazon RDS, Aurora

Amazon Glacier

Amazon Elasticsearch Service

Amazon DynamoDB

## Process & Analyze

Amazon Athena

Amazon EMR

Amazon EC2

Amazon QuickSight

Amazon Redshift

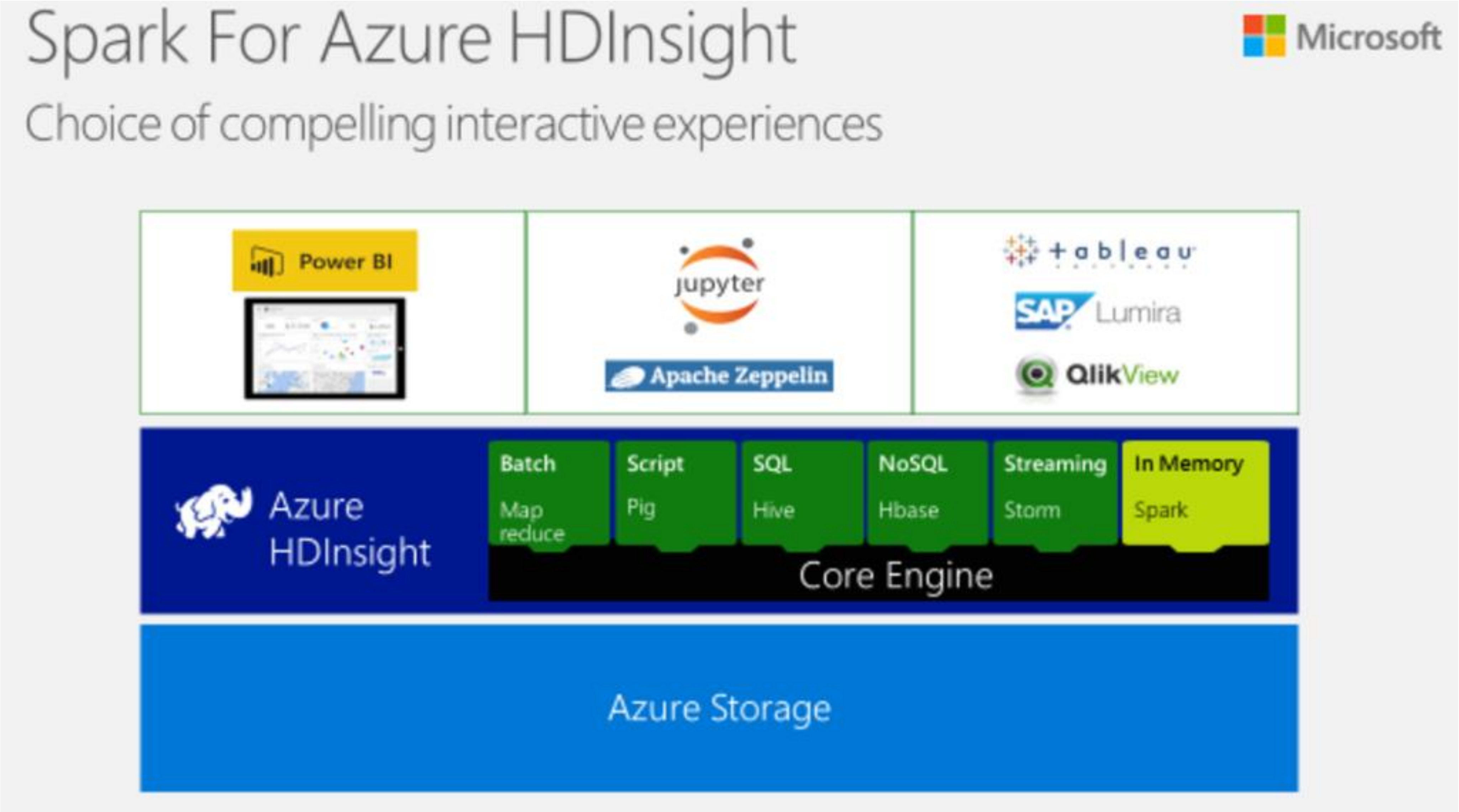
Amazon Machine Learning

Amazon Kinesis Analytics

AWS Database Migration Service

AWS Glue

AWS Data Pipeline



# Spark For Azure HDInsight

Microsoft

## Choice of compelling interactive experiences

Power BI

jupyter

Apache Zeppelin

tableau

SAP Lumira

QlikView

Azure HDInsight

Batch
Map reduce

Script
Pig

SQL
Hive

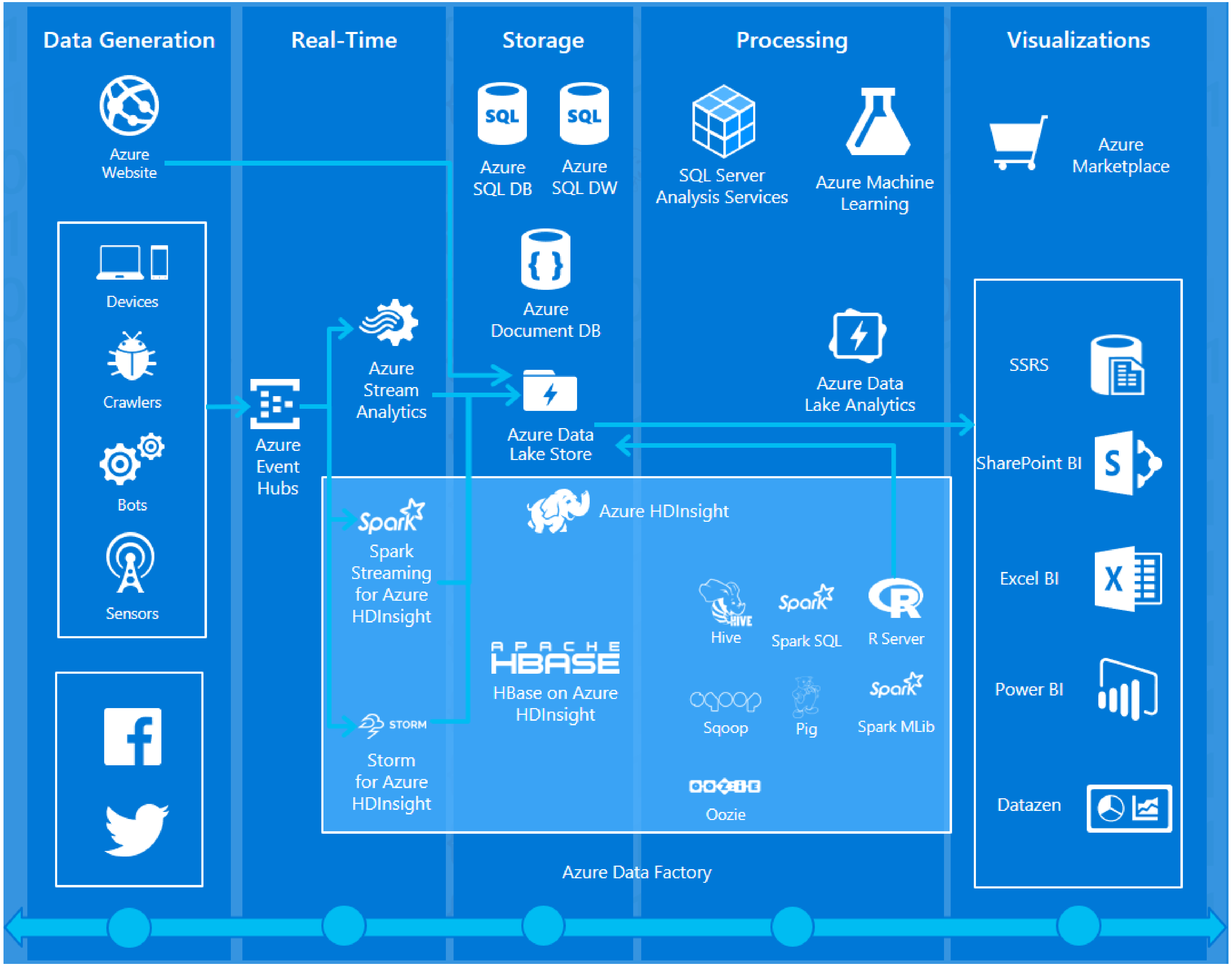
NoSQL
Hbase

Streaming
Storm

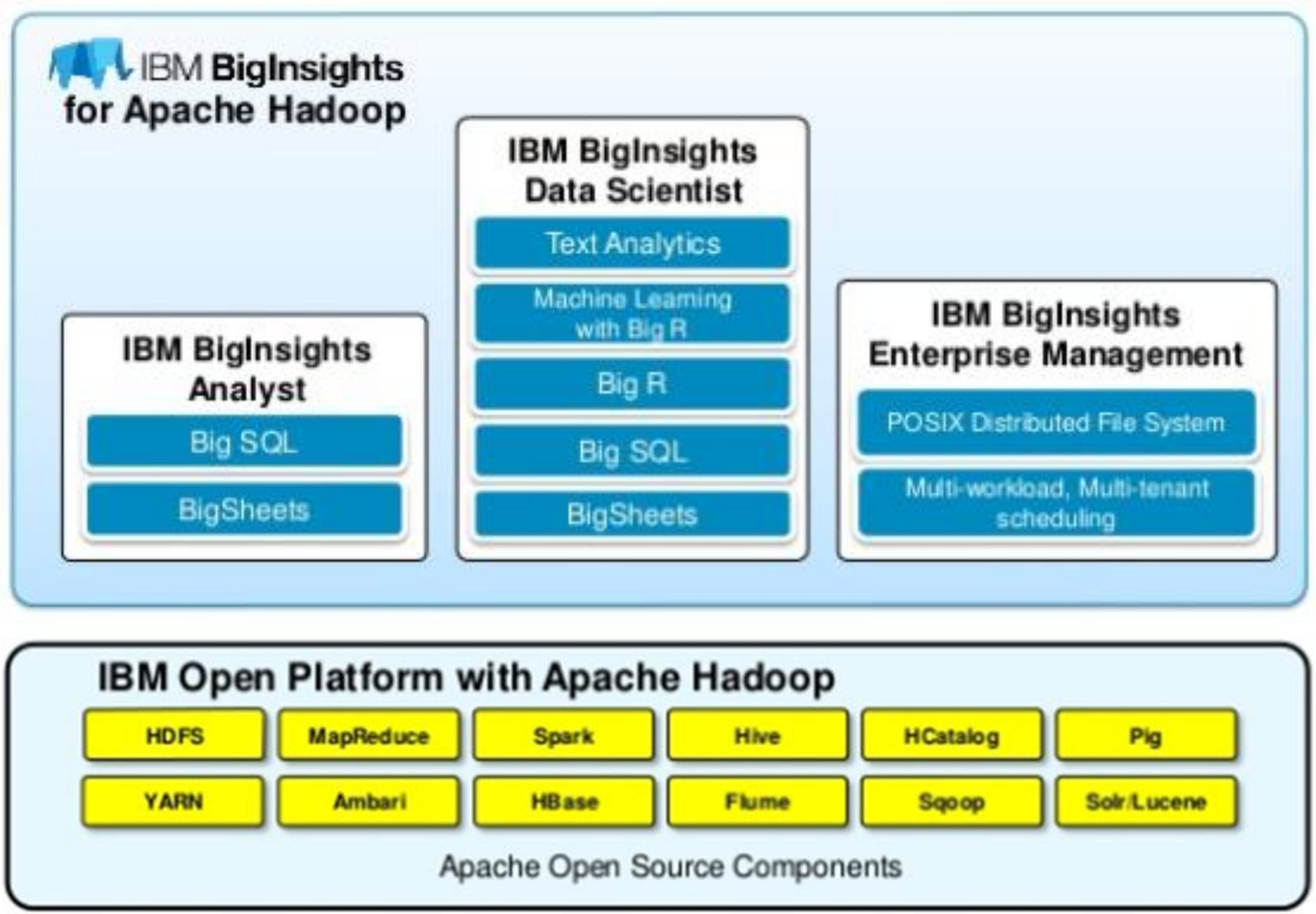
In Memory
Spark

Core Engine

Azure Storage



# Microsoft Azure HDInsight Architecture



IBM BigInsights for Apache Hadoop

IBM BigInsights Analyst
- Big SQL
- BigSheets

IBM BigInsights Data Scientist
- Text Analytics
- Machine Learning with Big R
- Big R
- Big SQL
- BigSheets

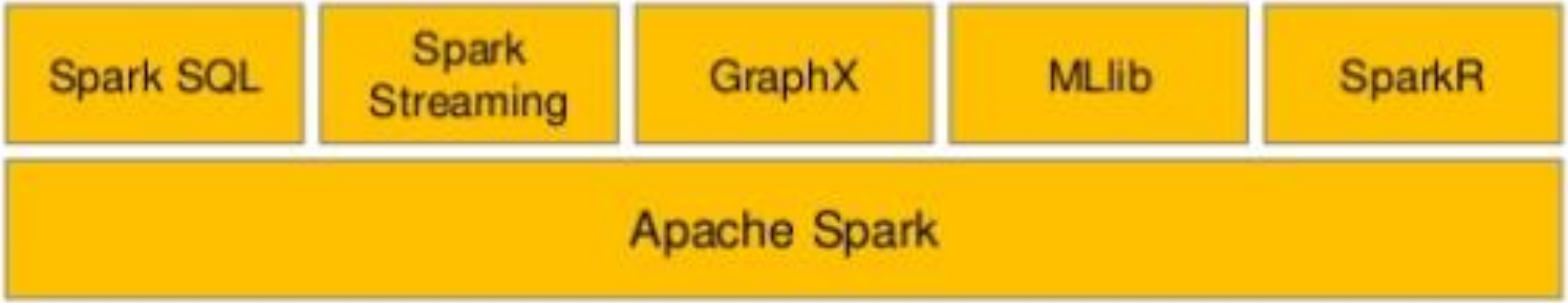
IBM BigInsights Enterprise Management
- POSIX Distributed File System
- Multi-workload, Multi-tenant scheduling

IBM Open Platform with Apache Hadoop

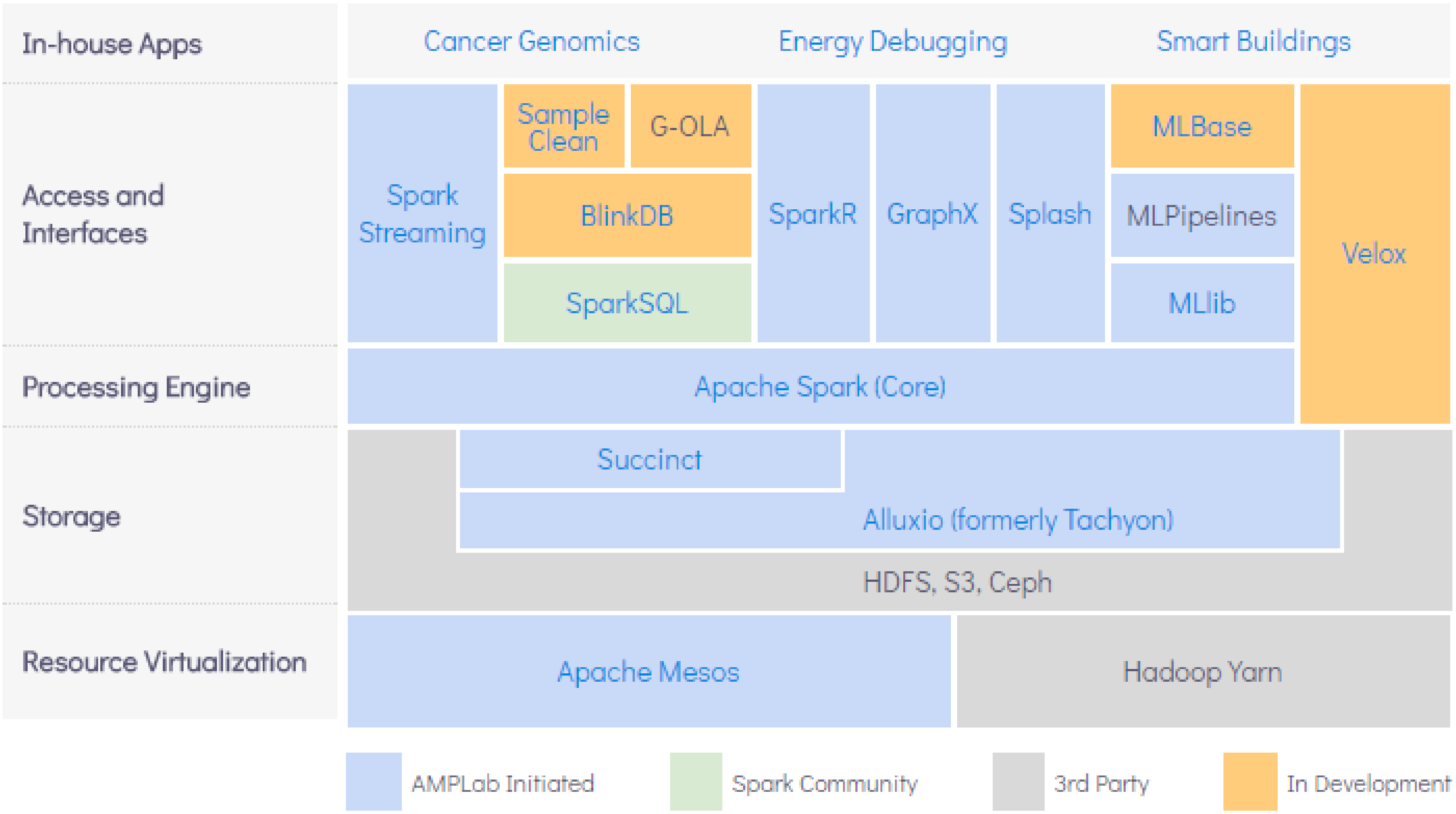
| | | | | | |
|---|---|---|---|---|---|
| HDFS | MapReduce | Spark | Hive | HCatalog | Pig |
| YARN | Ambari | HBase | Flume | Sqoop | Solr/Lucene |

Apache Open Source Components

# The Spark Ecosystem



# Berkeley Data Analytics Stack (BDAS)



In-house Apps

Cancer Genomics

Energy Debugging

Smart Buildings

Access and Interfaces

Spark Streaming

Sample Clean

G-OLA

BlinkDB

SparkSQL

SparkR

GraphX

Splash

MLBase

MLPipelines

MLlib

Velox

Processing Engine

Apache Spark (Core)

Storage

Succinct

Alluxio (formerly Tachyon)

HDFS, S3, Ceph

Resource Virtualization

Apache Mesos

Hadoop Yarn

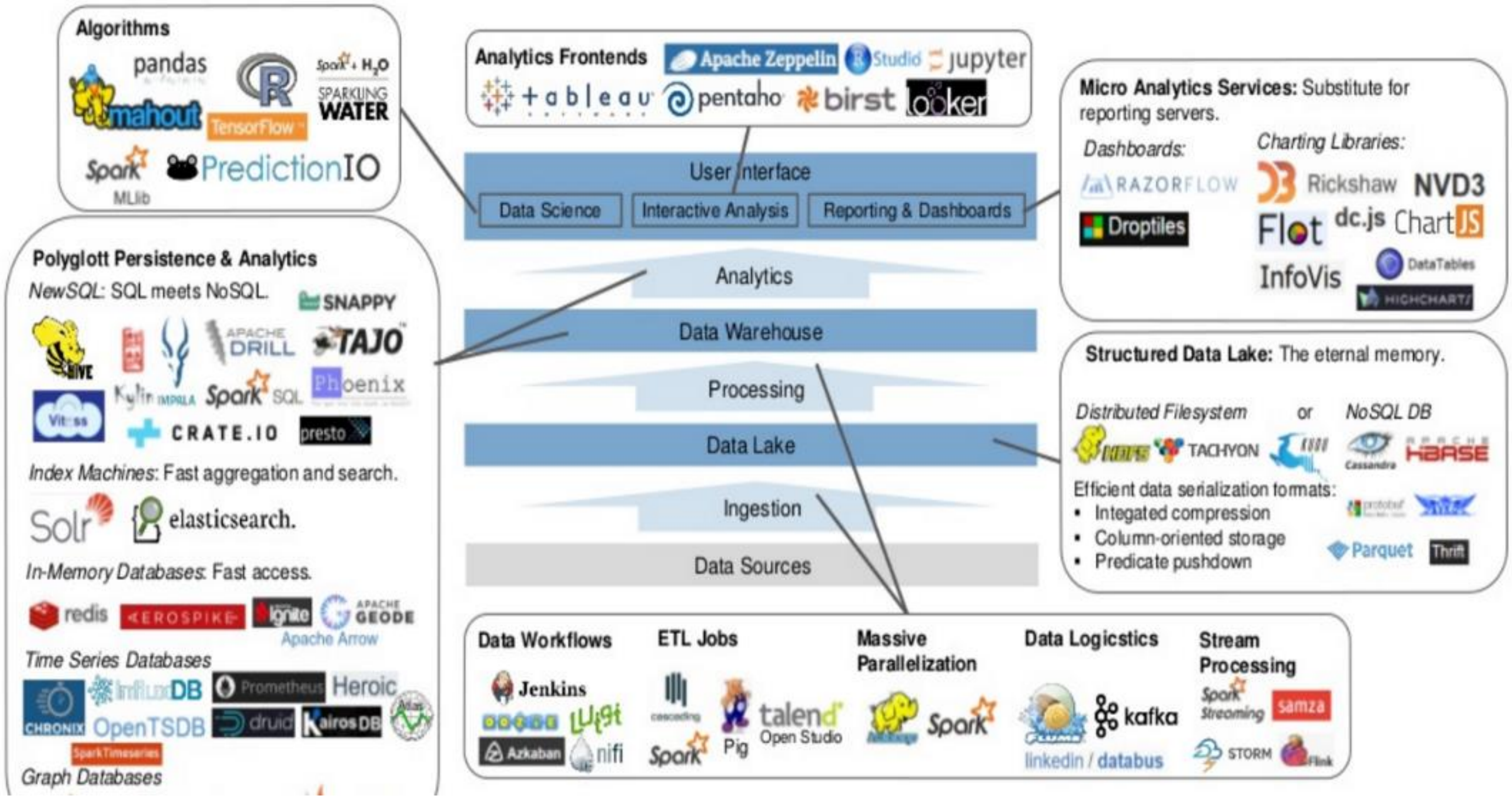
AMPLab Initiated

Spark Community

3rd Party

In Development

# Big Data Landscape 2016



Algorithms

Analytics Frontends

User Interface

Data Science

Interactive Analysis

Reporting & Dashboards

Analytics

Data Warehouse

Processing

Data Lake

Ingestion

Data Sources

**Micro Analytics Services:** Substitute for reporting servers.

*Dashboards:*

*Charting Libraries:*

**Structured Data Lake:** The eternal memory.

*Distributed Filesystem* or *NoSQL DB*

Efficient data serialization formats:

- Integated compression
- Column-oriented storage
- Predicate pushdown

**Polyglott Persistence & Analytics**

*NewSQL:* SQL meets NoSQL.

*Index Machines:* Fast aggregation and search.

*In-Memory Databases:* Fast access.

*Time Series Databases*

*Graph Databases*

Data Workflows

ETL Jobs

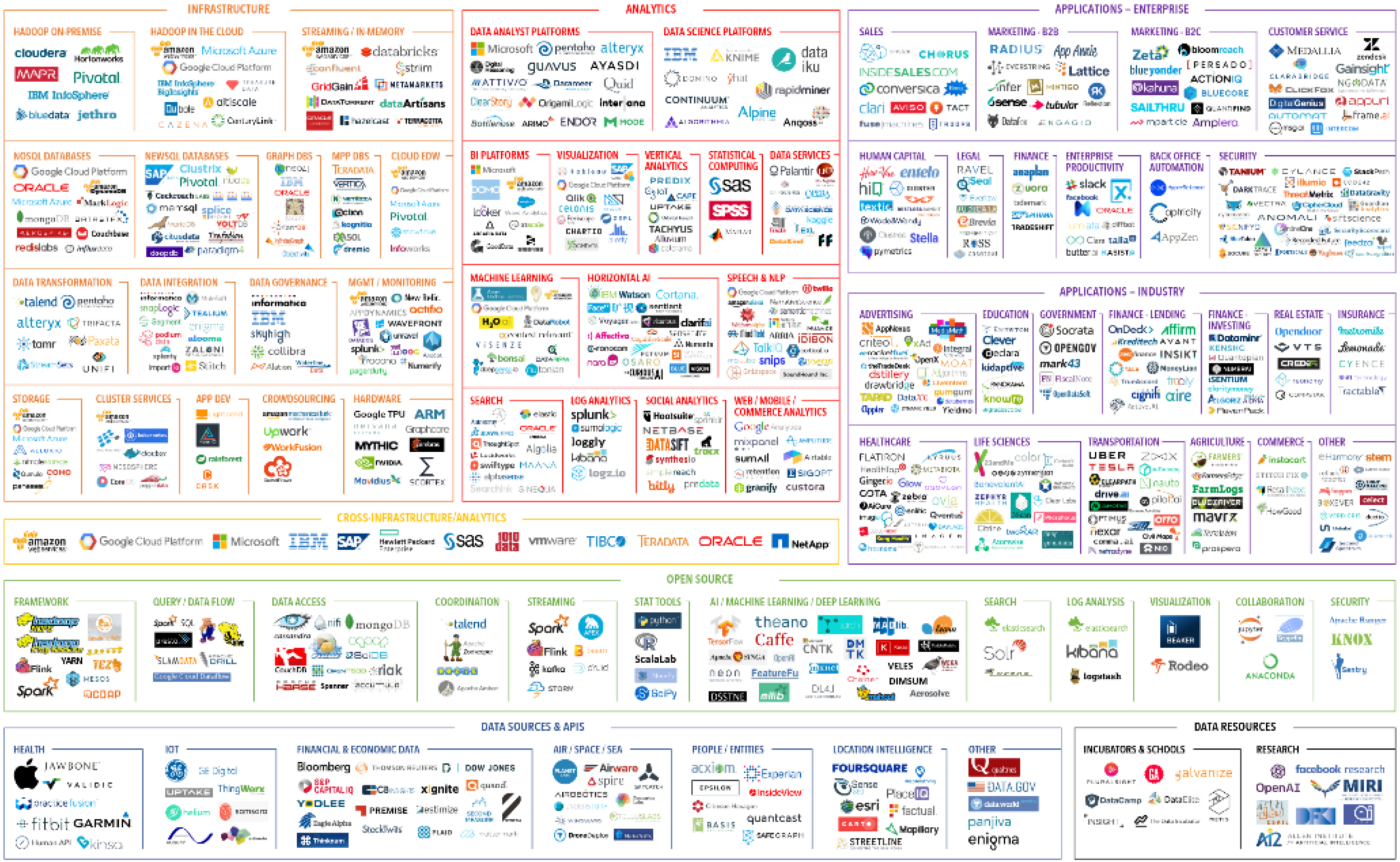
Massive Parallelization

Data Logicstics

Stream Processing

http://www.slideshare.net/adersberger/big-data-landscape-2016-58917032

# BIG DATA LANDSCAPE 2017



## INFRASTRUCTURE

HADOOP ON-PREMISE

HADOOP IN THE CLOUD

STREAMING / IN-MEMORY

NOSQL DATABASES

NEWSQL DATABASES

GRAPH DBS

MPP DBS

CLOUD EDW

DATA TRANSFORMATION

DATA INTEGRATION

DATA GOVERNANCE

MGMT / MONITORING

STORAGE

CLUSTER SERVICES

APP DEV

CROWDSOURCING

HARDWARE

## ANALYTICS

DATA ANALYST PLATFORMS

DATA SCIENCE PLATFORMS

BI PLATFORMS

VISUALIZATION

VERTICAL ANALYTICS

STATISTICAL COMPUTING

DATA SERVICES

MACHINE LEARNING

HORIZONTAL AI

SPEECH & NLP

SEARCH

LOG ANALYTICS

SOCIAL ANALYTICS

WEB / MOBILE / COMMERCE ANALYTICS

## APPLICATIONS – ENTERPRISE

SALES

MARKETING - B2B

MARKETING - B2C

CUSTOMER SERVICE

HUMAN CAPITAL

LEGAL

FINANCE

ENTERPRISE PRODUCTIVITY

BACK OFFICE AUTOMATION

SECURITY

## APPLICATIONS – INDUSTRY

ADVERTISING

EDUCATION

GOVERNMENT

FINANCE - LENDING

FINANCE - INVESTING

REAL ESTATE

INSURANCE

HEALTHCARE

LIFE SCIENCES

TRANSPORTATION

AGRICULTURE

COMMERCE

OTHER

## CROSS-INFRASTRUCTURE/ANALYTICS

## OPEN SOURCE

FRAMEWORK

QUERY / DATA FLOW

DATA ACCESS

COORDINATION

STREAMING

STAT TOOLS

AI / MACHINE LEARNING / DEEP LEARNING

SEARCH

LOG ANALYSIS

VISUALIZATION

COLLABORATION

SECURITY

## DATA SOURCES & APIS

HEALTH

IOT

FINANCIAL & ECONOMIC DATA

AIR / SPACE / SEA

PEOPLE / ENTITIES

LOCATION INTELLIGENCE

OTHER

## DATA RESOURCES

INCUBATORS & SCHOOLS

RESEARCH

V2 – Last updated 5/3/2017

© Matt Turck (@mattturck), Jim Hao (@jimrhao), & FirstMark (@firstmarkcap) mattturck.com/bigdata2017

FIRSTMARK
EARLY STAGE VENTURE CAPITAL

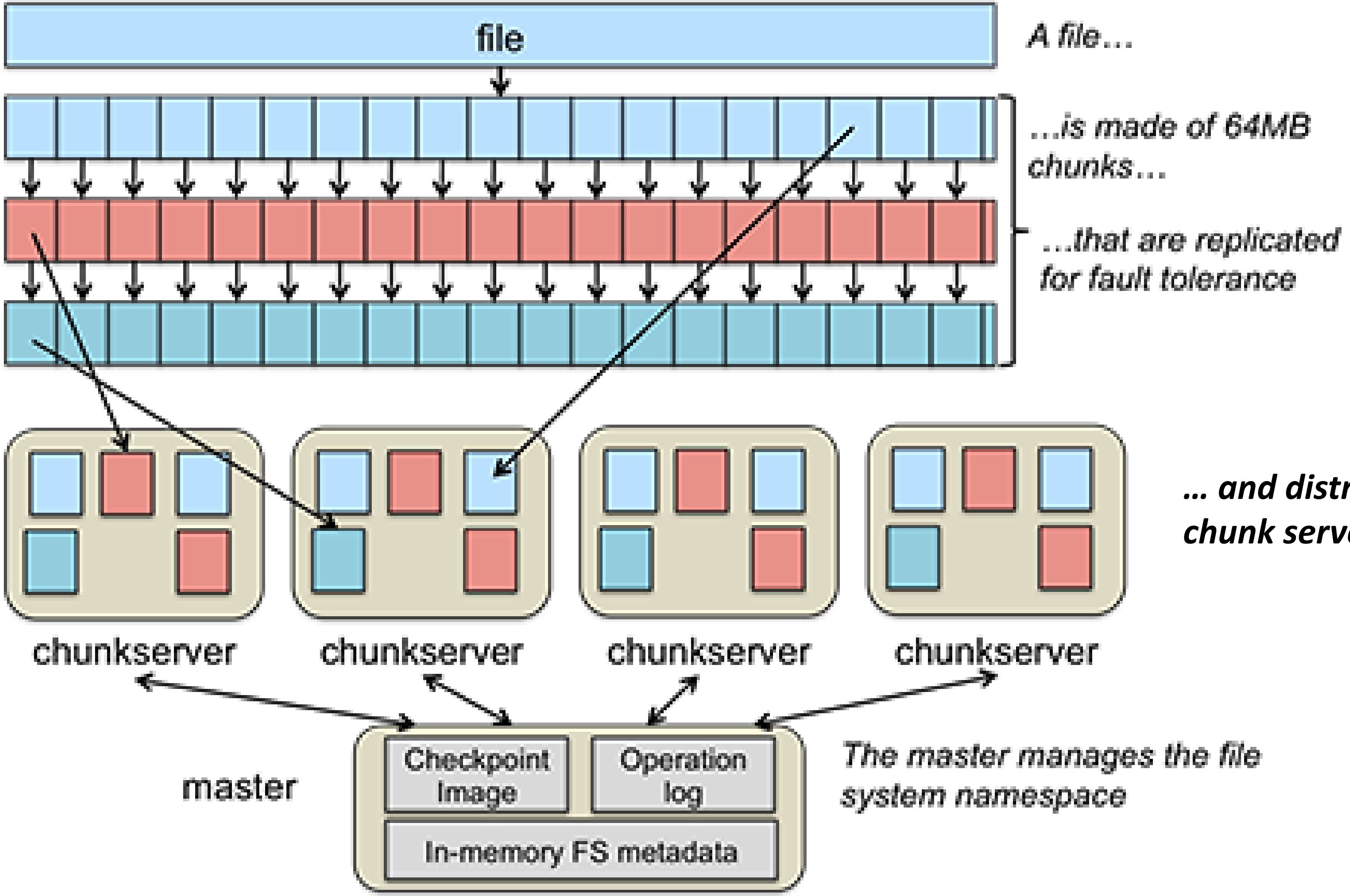
© 2014 Ted Goff www.tedgoff.com

“So you want to hire me as a Data Scientist for Intelligent Virtualized Deep Machine Learning Real-time Big Data in the Cloud for Social Networks? Ok, but if you also want Hadoop, increase my salary by 50%.”

# Hadoop : Storage

## HDFS

# GFS



*A file…*

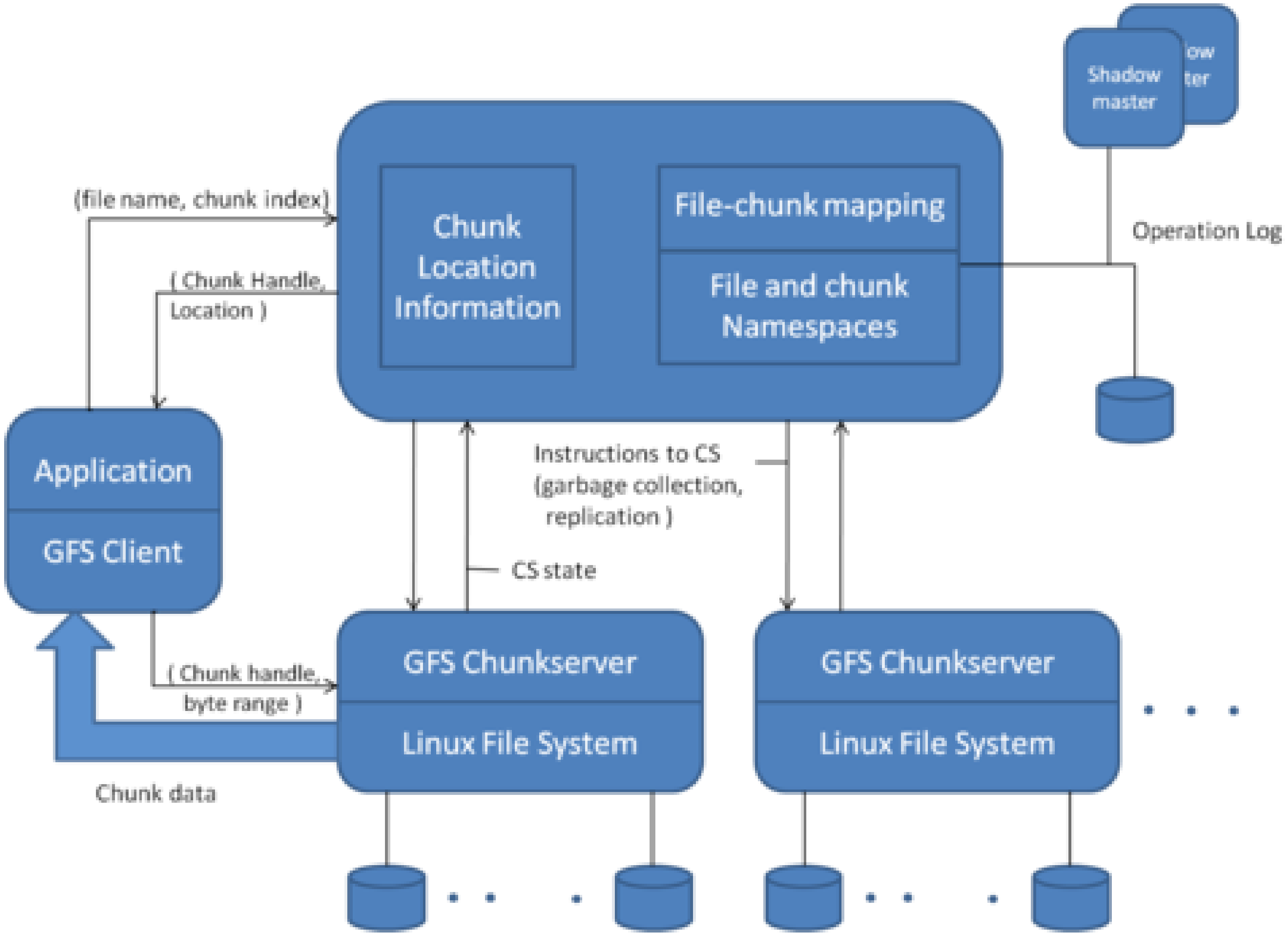
*…is made of 64MB chunks…*

*…that are replicated for fault tolerance*

***… and distributed on chunk servers***

*The master manages the file system namespace*

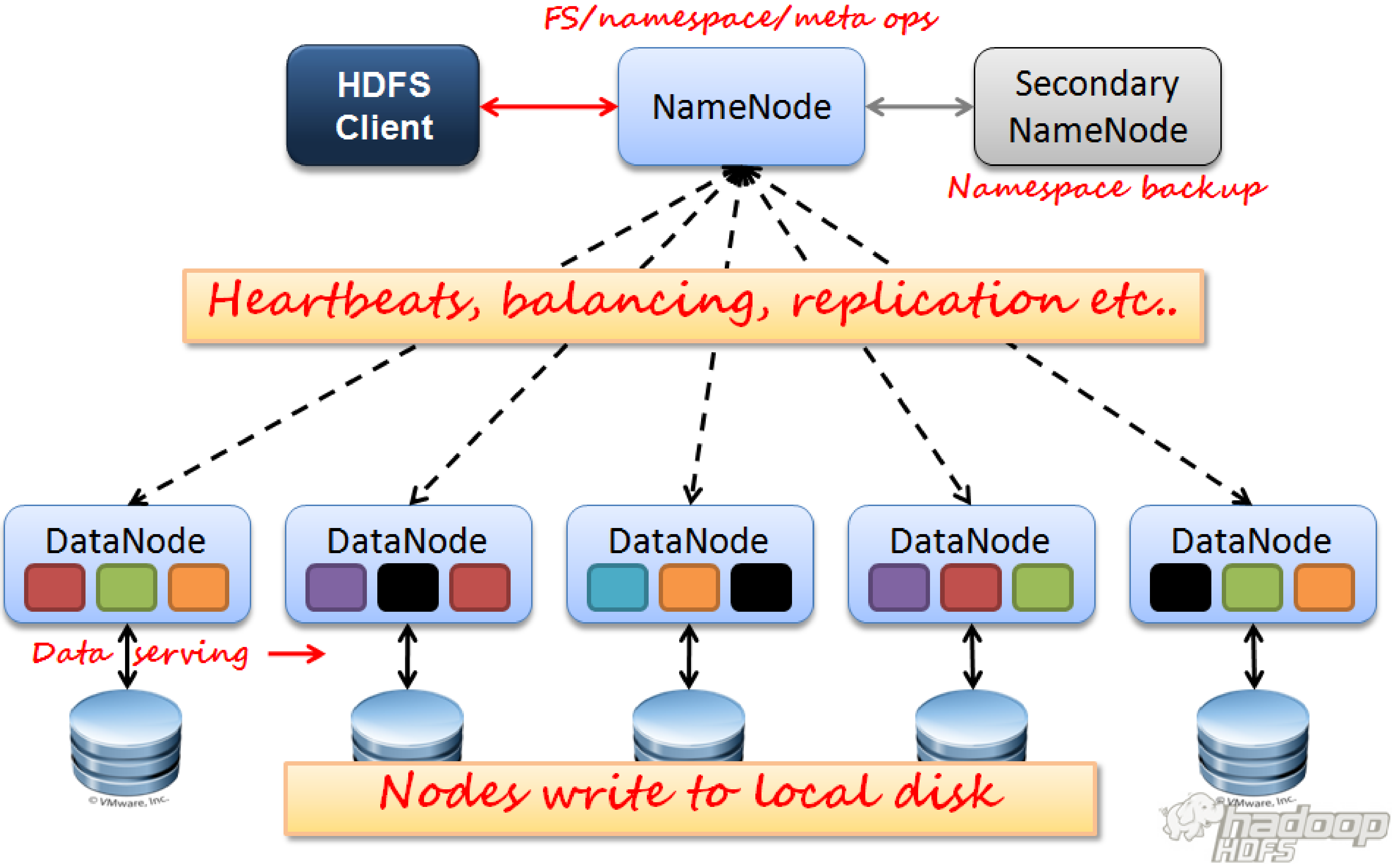
# GFS: Architecture



# GFS Master Responsibilities

- Metadata storage
- Namespace management/locking
- Periodic communication with chunkservers
  - □ give instructions, collect state, track cluster health
- Chunk creation, re-replication, rebalancing
  - □ balance space utilization and access speed
  - □ spread replicas across racks to reduce correlated failures
  - □ re-replicate data if redundancy falls below threshold
  - □ rebalance data to smooth out storage and request load
- Garbage Collection
  - □ simpler, more reliable than traditional file delete
  - □ master logs the deletion, renames the file to a hidden name
  - □ lazily garbage collects hidden files
- Stale replica deletion
  - □ detect “stale” replicas using chunk version numbers

# HDFS CDH3: Open source reimplementation of GFS



# NameNode Metadata

- Metadata in Memory
    - The entire metadata is in main memory

- Types of metadata
    - List of files
    - List of Blocks for each file
    - List of Data Nodes for each block
    - File attributes, e.g. creation time, replication factor

- A Transaction Log
    - Records file creations, file deletions etc

# Block Replica Placement

- Current Strategy
  - One replica on local node
  - Second replica on a remote rack
  - Third replica on same remote rack
  - Additional replicas are randomly placed

- Clients read from nearest replicas

- Policy is pluggable

# Replica Block Placement

- Rack-Aware strategy

# HDFS: Heartbeat and Rebalancing

- **Heart beats**
  - Data Nodes send hear beat to the Name Node
  - Once every 3 seconds
  - Name Node uses heartbeats to detect Data Node failure

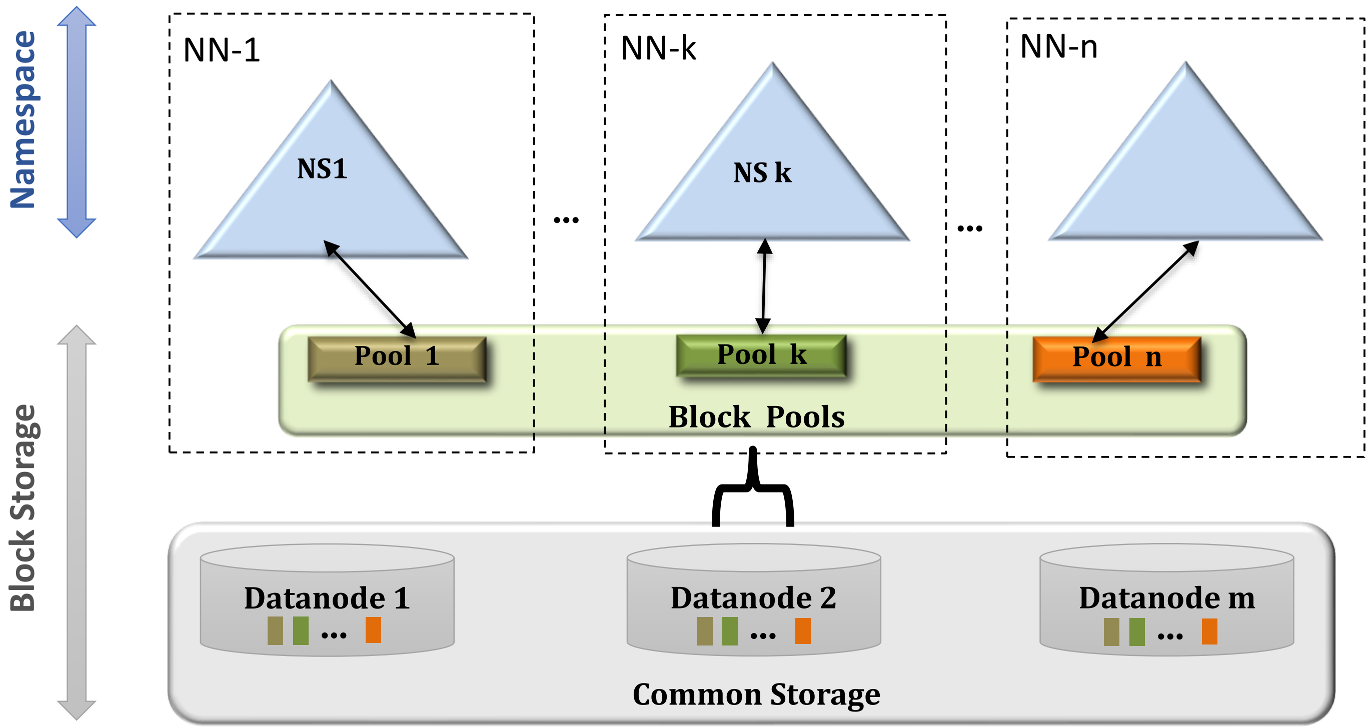
- **Rebalancing**: % disk full on Data Nodes should be similar

  - Usually run when new Data Nodes are added
  - Cluster is online when Rebalancer is active
  - Rebalancer is throttled to avoid network congestion
  - Command line tool

# What limitations you can think of ?

- Any problems you foresee wrt what we have seen so far?

# HDFS 2.0: Name Node Federation Elaborated

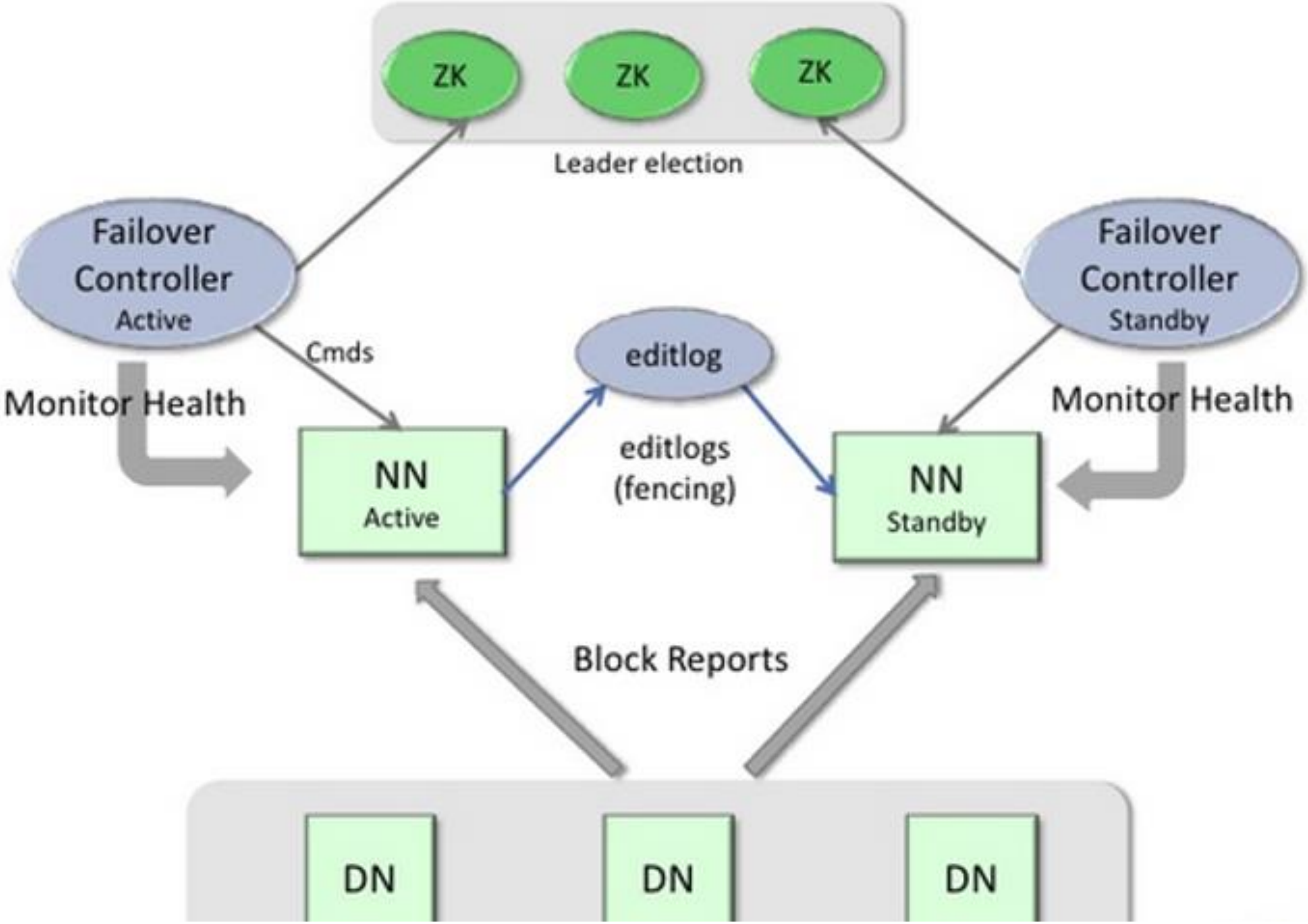


- Multiple **independent** Namenodes and Namespace Volumes in a cluster
  - Namespace Volume = Namespace + Block Pool
- Block Storage as generic storage service
  - Set of blocks for a Namespace Volume is called a ***Block Pool***
  - DNs store blocks for all the Namespace Volumes – no partitioning

# What benefits come from “Federation”?

- Bottlenecks
- Chargeback
- Isolation

# HDFS 2.0: High Availability Elaborated



Reference: Hadoop Summit 2012 | HDFS High Availability talk

In order to provide a fast failover, it is also necessary that the Standby node have up-to-date information regarding the location of blocks in the cluster.

In order to achieve this, the DataNodes are configured with the location of both NameNodes, and send block location information and heartbeats to both.

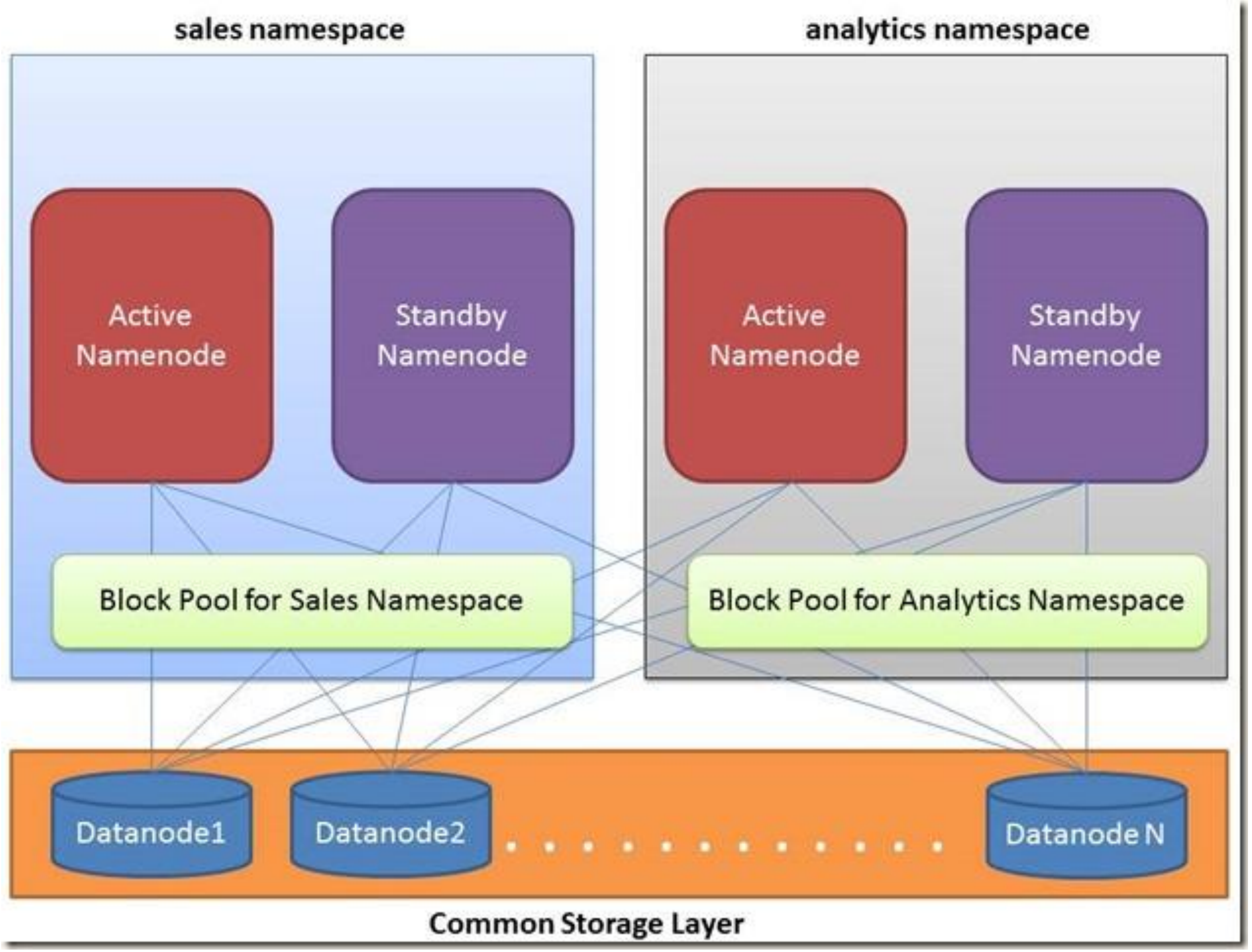
The ZKFailoverController (ZKFC) is a ZooKeeper client that also monitors and manages the state of the NameNode.

Each of the hosts that run a NameNode also run a ZKFC.

The ZKFC is responsible for **Health monitoring** of Namenode

- the ZKFC contacts its local NameNode on a periodic basis with a health-check command. So long as the NameNode responds promptly with a healthy status, the ZKFC considers the NameNode healthy.
- If the NameNode has crashed, frozen, or otherwise entered an unhealthy state, the health monitor marks it as unhealthy.

# HDFS 2.0: High Availability, Federated



http://pe-kay.blogspot.in/2016/02/configuring-federated-hdfs-cluster-with.html

# Some administrative work

- Block Checker

- Balancer

- Commissioning and decommissioning

- HDFS still needs to be backed up.
- Select and Prioritize what to backup

# So what all we have learnt ?

- V’s of Big Data
- Big Data Use Cases
- Hadoop and its ecosystem
- HDFS, HDFS 2.0

| | |
|---|---|
| Web: | http://www.insofe.edu.in |
| Facebook: | https://www.facebook.com/insofe |
| Twitter: | https://twitter.com/Insofeedu |
| YouTube: | http://www.youtube.com/InsofeVideos |
| SlideShare: | http://www.slideshare.net/INSOFE |
| LinkedIn: | http://www.linkedin.com/company/international-school-of-engineering |