CERTIFICATE IN BIG DATA AND DATA ANALYTICS

Dr. Anthony Franklin

MATRIX TRC

Training, Research & Consulting
MATRIX

Meirc Training & Consulting

PLUS SPECIALTY TRAINING

Objectives

Open Dialogue

Design your own Big Data Analytics architecture diagram.

MSFT Azure Services:
Subscription
Azure Databricks
Azure BLOB

Hands-on Examples w/ Big Data

Notes:

Modules do NOT correspond to days
Read a case study each day
Hands-on-labs with technology
End of course assignment and Presentation (Last Day)
Big Data for ANALYTICS focus

My Bio

- Father of 3
- Former college football player
- Born in New York City, NY USA
- PhD from NC State
- Former Instructor for NCSU Institute for Advanced Analytics
- 5 years at SAS
- Currently MSFT Big Data and AI Architect
- Co-founder of Fanalytical

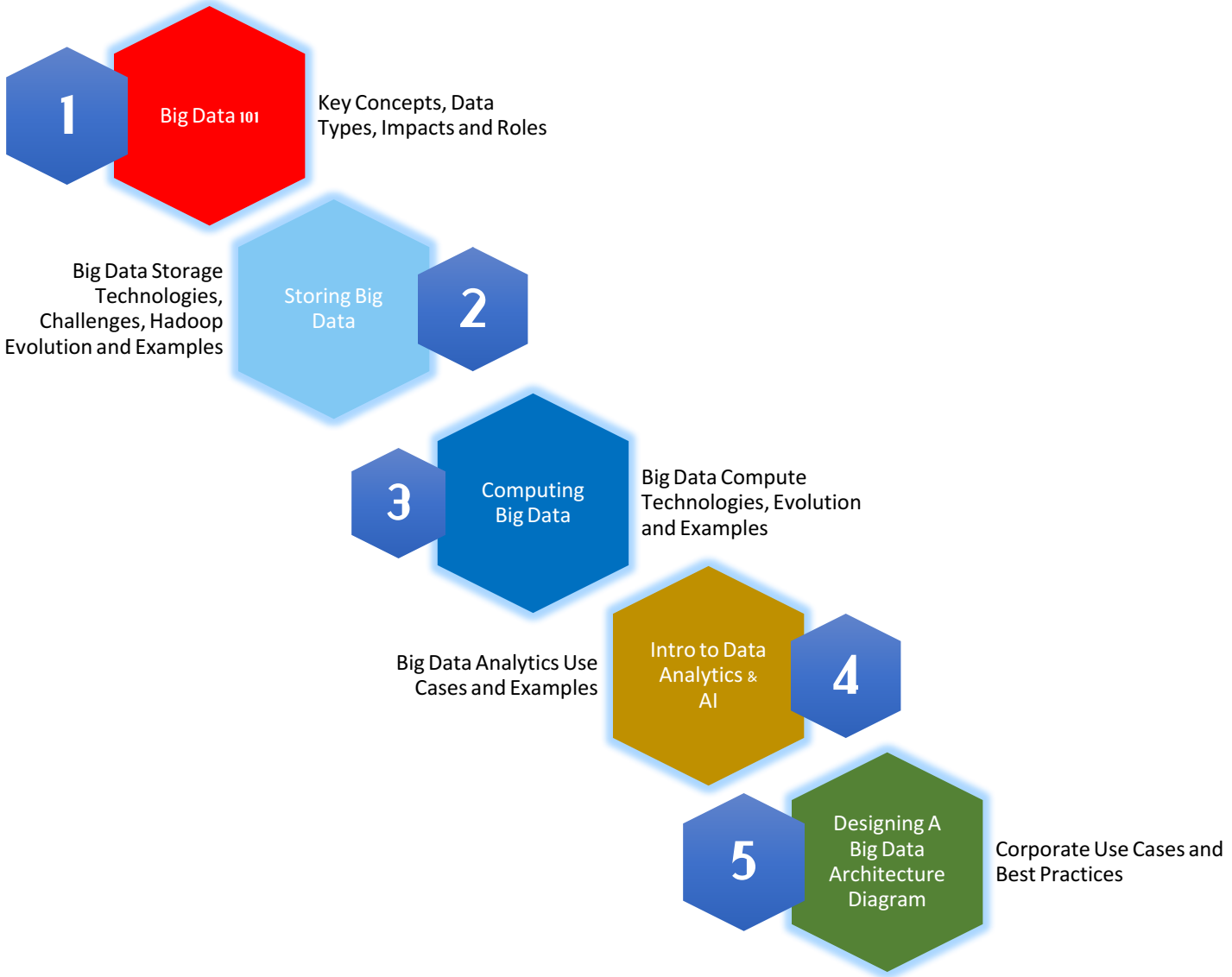
What is Github?

Git is a revision control system, a tool to manage your source code history. **GitHub** is a hosting service for **Git** repositories. So they are not the same thing: **Git** is the tool, **GitHub** is the service for projects that use **Git**.

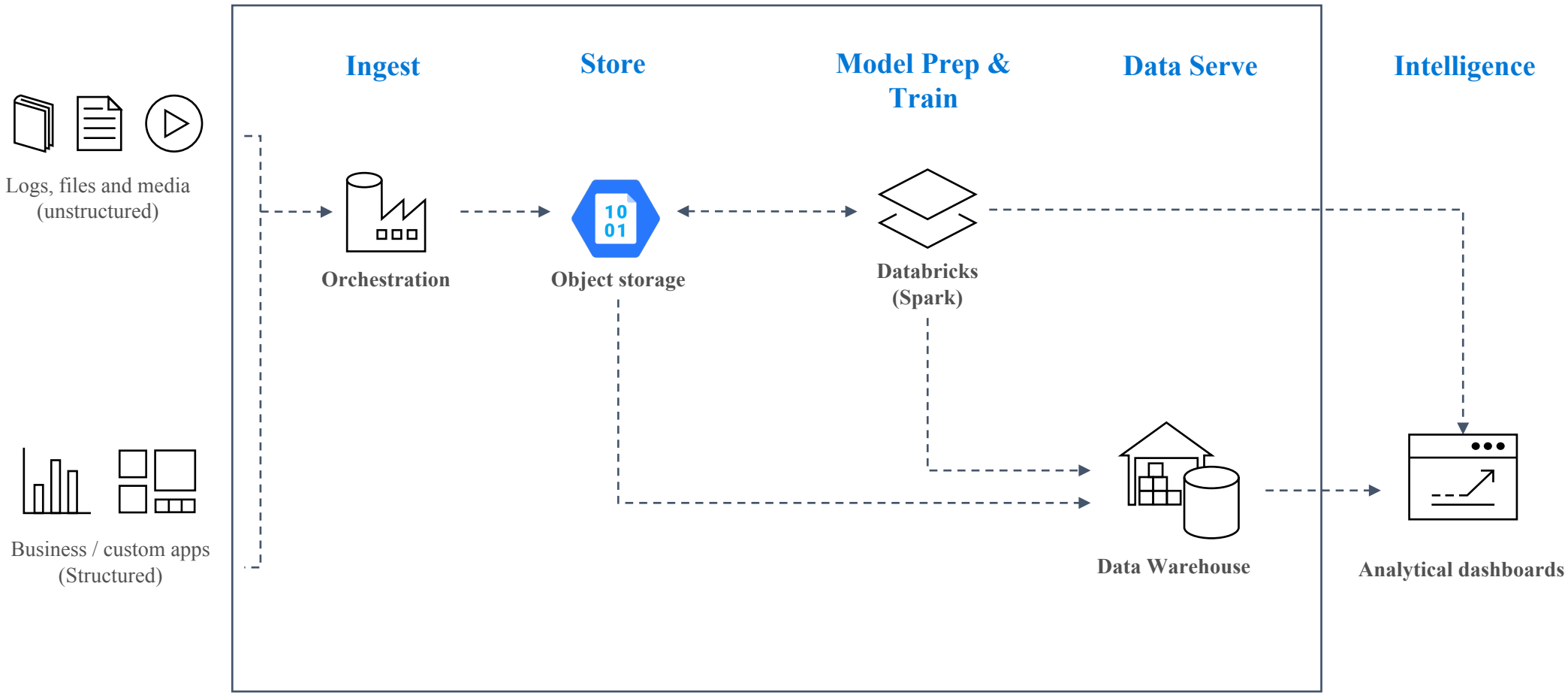
Course Hands-On-Labs:

https://github.com/franklin8705/Big-Data-and-Data-Analytics-Course

Course Flow


1
Big Data 101
Key Concepts, Data Types, Impacts and Roles
Big Data Storage Technologies, Challenges, Hadoop Evolution and Examples
Storing Big Data
2
3
Computing Big Data
Big Data Compute Technologies, Evolution and Examples
Big Data Analytics Use Cases and Examples
Intro to Data Analytics & AI
4
5
Designing A Big Data Architecture Diagram
Corporate Use Cases and Best Practices

Architecture Diagram Example



Characteristics

No directory hierarchy

Store files in flat plane

Key-value pair access

Persistent storage

Notes

Great for unstructured, static and archival data.
Requires less metadata
Competes with Hadoop.
Low latency and intended for limited writes.
Easily scalable. Compatible with cloud storage platforms.
Handles petabytes of storage.

Icons

Cost

Types of Data

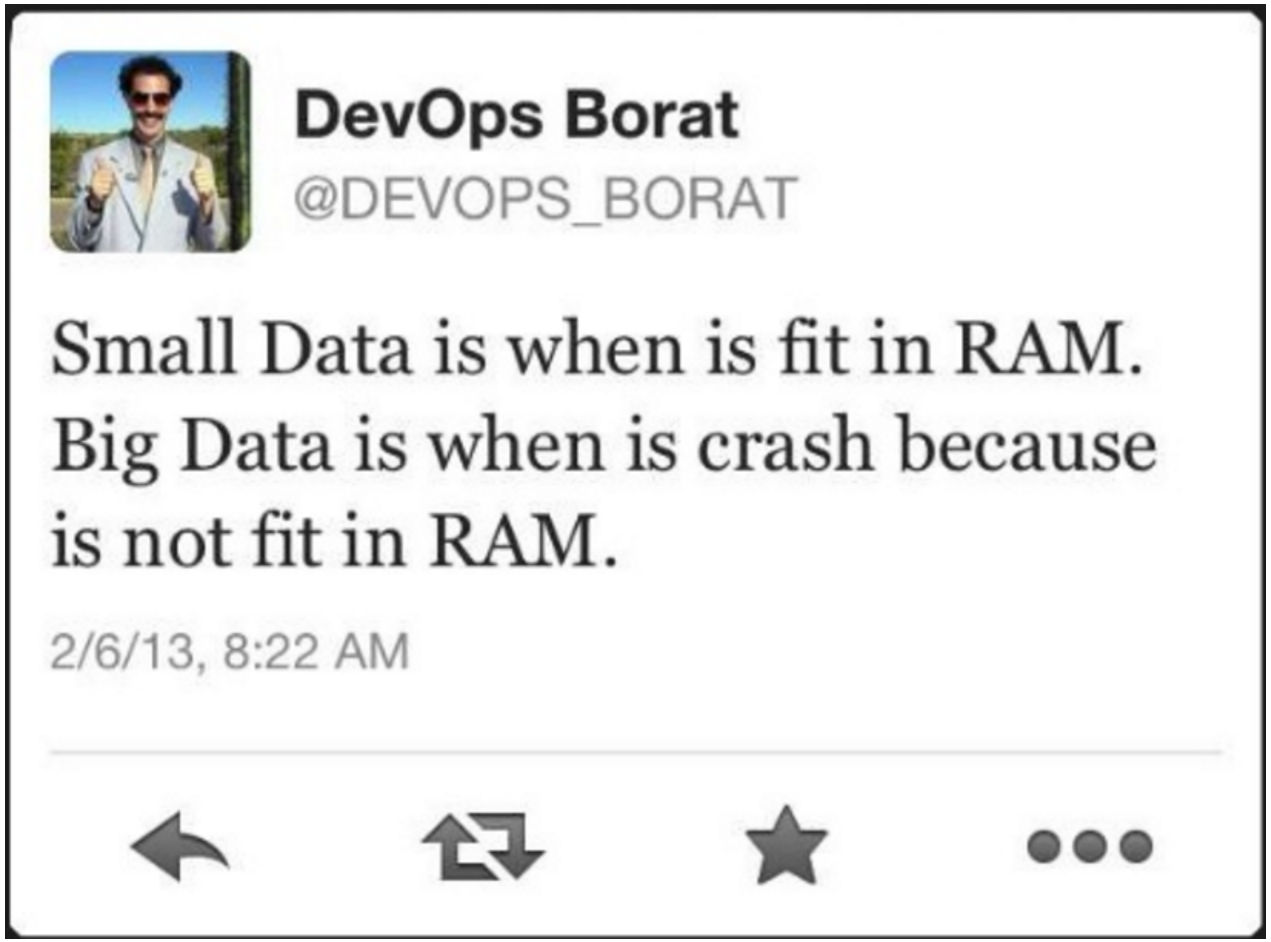
Big Data 101

- Big Data Defined
- Big Data Example
- Big Data Impact
  - Analytics
  - Technologies
- Open Source Revolution
- 360 View
- Professional Roles
- Architectures
  - SMP vs MPP
  - Distributed In-Memory

What is Big Data?

Meirc Training & Consulting

PLUS SPECIALTY TRAINING

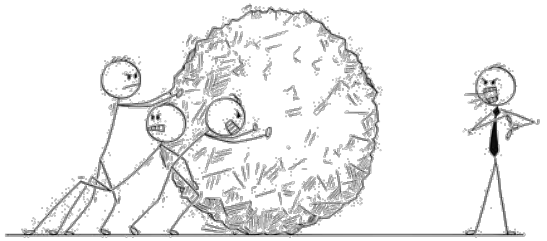

DevOps Borat
@DEVOPS_BORAT
Small Data is when is fit in RAM.
Big Data is when is crash because
is not fit in RAM.
2/6/13, 8:22 AM

Check Laptop Memory Capacity

Big Data Lab

Azure Databricks

When **traditional** data systems can *no longer* handle the tasks.



So **IT IS NOT** what is Big Data, but what are the

IMPACTS

of Big Data.

The 5 Vs of Big Data

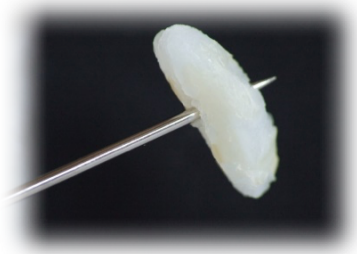
Meirc Training & Consulting

PLUS SPECIALTY TRAINING

5 Vs of Big Data

Data Volume

Byte
One grain of rice



Kilobyte
= 1,024 Bytes
Cup of rice

Megabyte
= 1,024 Kilobytes
8 bags of rice

Gigabyte
= 1,024 Megabytes
2 semi trucks

Terabyte
= 1,024 Gigabytes
3 containers

Data Volume

Petabyte
= 1,024 Terabytes
Blankets Manhattan

Exabyte
= 1,024 Petabytes
Half of India

Zettabyte
= 1,024 Exabyte
Fills Pacific Ocean

Yottabyte
= 1,024 Zettabytes
Planet Earth

Brontobyte
= 1,024 Yottabytes
A Galaxy

Geobyte
= 1,024 Brontobytes
The Universe

Data Volume

Size	JPEG	1080P Video	MP3 Audio	Relational Table
1GB	195 files	39 min	17.5 hrs	~300,000 rows
64GB	12,500 files	41.6 hrs	1,120 hrs	~20 MM rows
256GB	50,000 files	166.4 hrs	4,480 hrs	~100 MM rows

The total amount of data available today can overwhelm traditional storage systems.

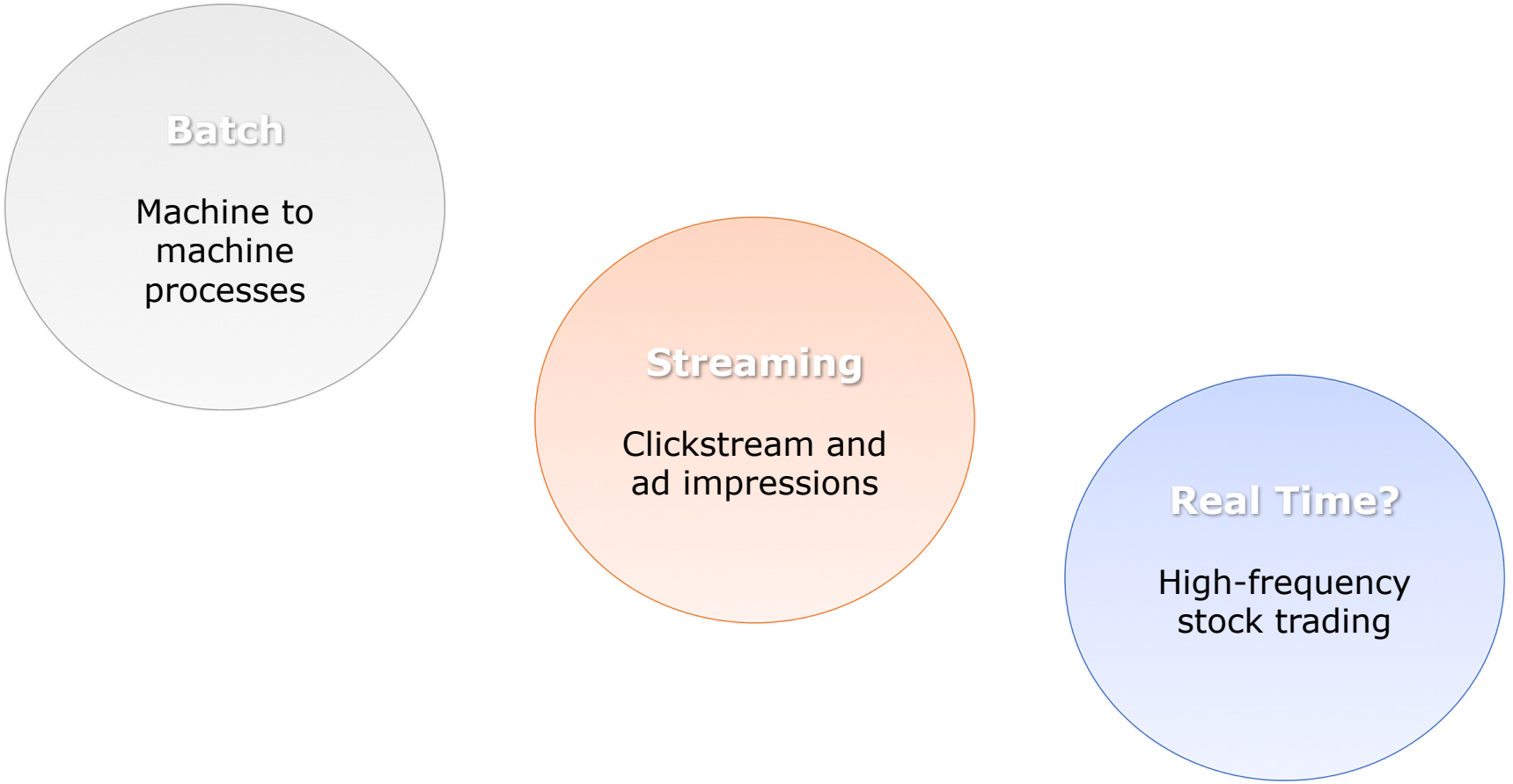
Walmart … handles more than 1MM transactions per hour.

… handles 40 Billion photos from its user base

Decoding the human genome originally took 10 years to process… now takes one week.

Velocity

Data can be created and/or received at **ranges of speed** from ***batch*** to ***real time*** processing.



Variety

With most **important sources** of big data being relatively ***new***, structured databases that stored most corporate data until recently are often **not appropriate** for *storing* and *processing* **new varieties** of data.

Videos

Log Files

Raw Survey Data

Financial Statements

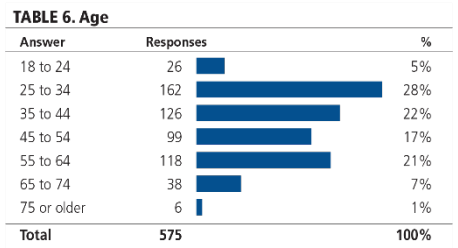


TABLE 6. Age		
Answer	Responses	%
18 to 24	26	5%
25 to 34	162	28%
35 to 44	126	22%
45 to 54	99	17%
55 to 64	118	21%
65 to 74	38	7%
75 or older	6	1%
Total	575	100%

Tables

Images

Audio Files

Email Correspondences

Notes or Presentations

Flat Files

Sensor Data

Trustworthiness of the data and *Reliability* of data source.

"**27%** of IT respondents said they were *unsure* whether their data was **accurate.**"

(IBM Study)

Worth derived from the *Information*.

Value

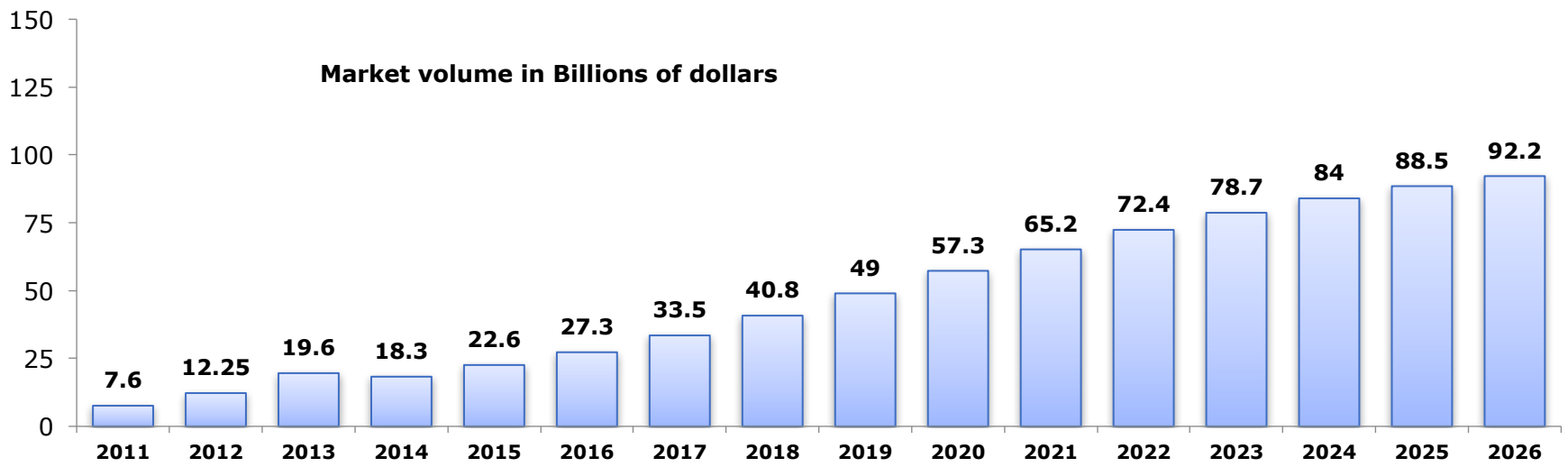
- ✓ Will the insights you gleam help you achieve your objective?
- ✓ Goal: make better decisions, take meaningful actions.
- ✓ Competitive advantage

US Market for Big Data

2016 = $27 Billion

2026 = $92 Billion

Future value



Big Data Risks

Meirc Training & Consulting

PLUS SPECIALTY TRAINING

OVERWHELMING

- Right technology
- Tech evolving
- Right people to solve right problems
- More IT
- Dark Data ,

PRIVACY

- Self-regulation
- Government regulation
- Legal regulation

COSTS

- Escalating too fast
- Must it be 100% captured?

SECURITY

- Infrastructure security
- Integrity security
- Reactive security
- Data Privacy

Principle of Finality

The principle of finality sets out that data must be collected for …

Specified *Explicit* *Legitimate purposes*

… and not further processed in a way incompatible with those purposes.

Principle of Proportionality

The proportionality principle **prevents** the personal data to be **used excessively** in relation to the purposes for which they were collected.

Thus companies and **public authorities are restricted** in the use and processing of personal data in their databases, including the …

Reuse *Transfer* *Sale of data*

Big Data Impact on Analytics

Meirc Training & Consulting

PLUS SPECIALTY TRAINING

How does BD **impact Analytics**?

Accessing data for analytics

Storing data for analytics

Processing data for analytics

Visualizing data for analytics

Operationalizing analytics jobs

Big Data Impact on Technologies

Meirc Training & Consulting

PLUS SPECIALTY TRAINING

How does BD impact technologies?

More *data*, more *problems* & *new problems*

Problem specific *technologies* **developed**

Cost of old vs **new** technologies

Proprietary vs *Open Source* development

Technology **revolution**

Rapid evolution

Did You Know?

240TB of data

will be generated by a Boeing

737 in a ***single flight***!

Typical PC in 2000 had 10GB of disk

↓

1TB of disk … today!

1.2 trillion
searches per year

We perform 40,000 search queries every second (on Google alone), which makes it 3.5 billion searches per day and …

Open Source Revolution

Meirc Training & Consulting

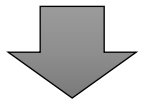
PLUS SPECIALTY TRAINING

Big Data problems are … **BIG**.

Catalyst for **openness**

Complex

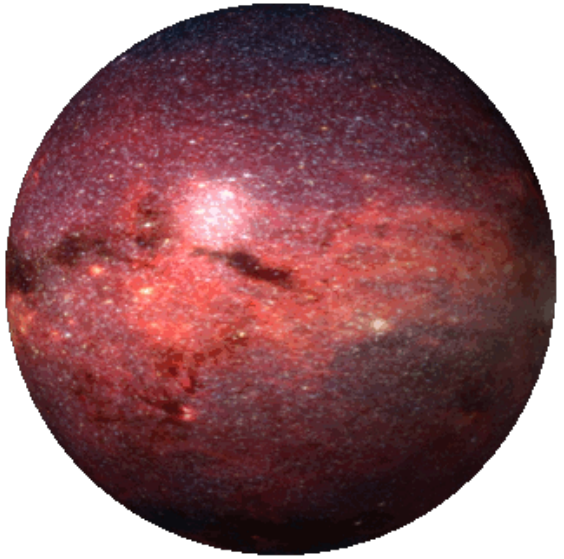
Robust



The **Open Movement**

Hardware availability for **Big Data** problems in the public.

Impact On Thinking



I need to know what I'm looking for.
↓
Recovery

I don't know what I'm looking for.
↓
Discovery

There are some things so big, they become **irresistible**. They become, an **opportunity** worth sharing.

There are some things so big, they have **implications** on our lives, whether we want it or not.

Open source VS **Proprietary software**

Freedom

Community

Unlimited access

Control

Lack of access

Immediate family

Open ...

... projects

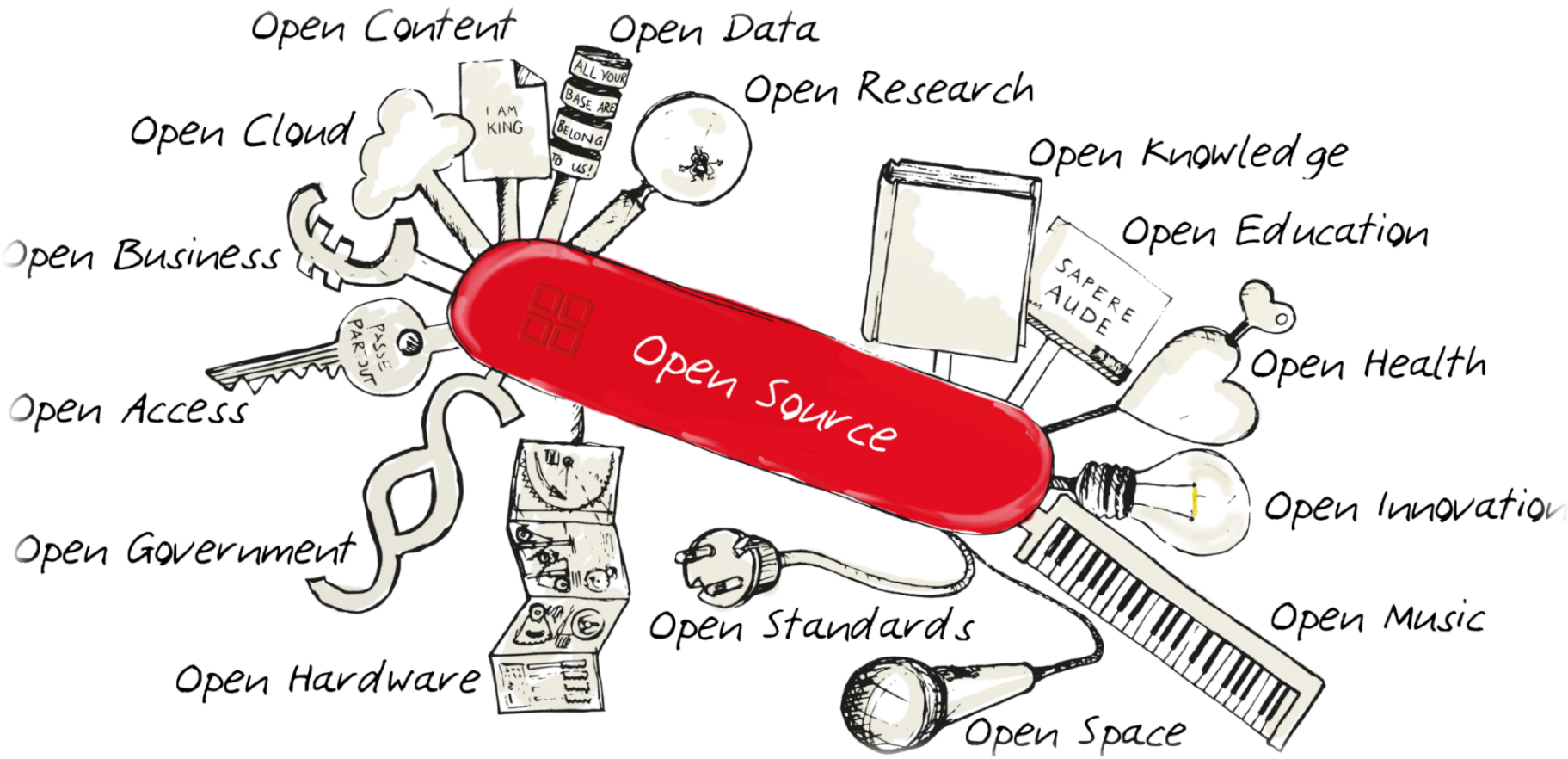
... challenges

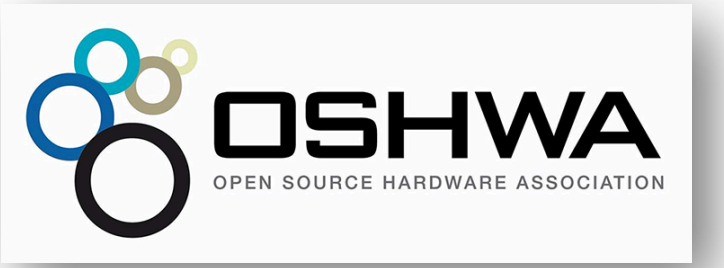
... forums

... Hardware

FOSH

The Open Movement





*"...For me, the **Open Source Hardware** is an **incredible opportunity** to show the world that this country still has a lot of **potential** and has people capable of **create**, **invent** and develop **solutions** to the problems, this empowers the feeling on the people that we are not just a consumerist country with a lot of social problems but a place were we can **create** and **solve problems** by ourselves ...."*

What is Open Source?

General

"Something that can be modified because its *design* is *publicly accessible*"

Software

"Software whose *source code* is available for modification or enhancement by anyone"

Culture and Philosophy

"Embrace and celebrate …

collaborative participation

open exchange

meritocracy

rapid prototyping

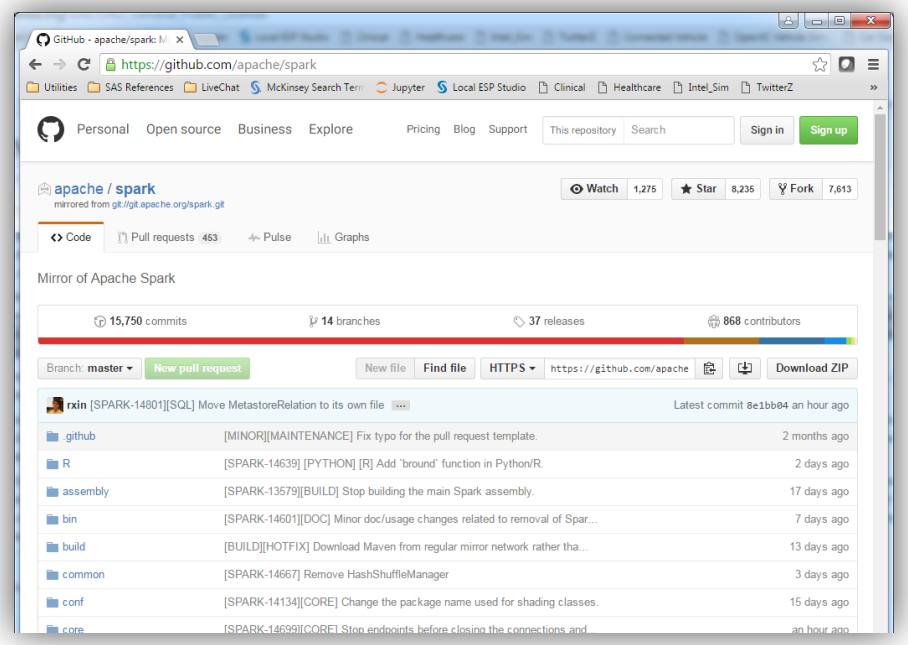
transparency

community development"

What is Open Source?

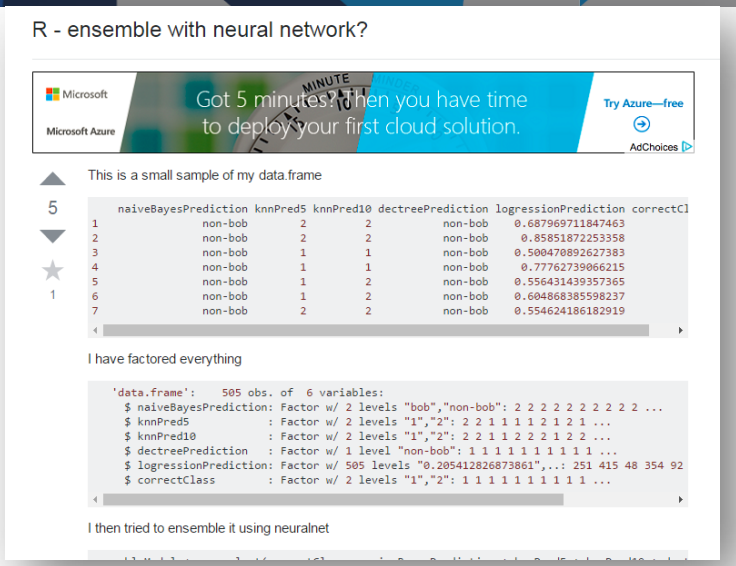
- Open source is the software for which the original source code is made freely available with an open source license for redistribution and modification by anyone.
- The pre-written programs or functions can be used as is or changed to fit the user specific need.

- Open Source technologies often have GPL (General Public License), which is a widely used *copyleft license*, which guarantees end users (individuals, organizations, companies) the freedom to run, study, share and modify the software.



Key Characteristics

- Active community of users who ***communicate***, ***share code***, and ***troubleshoot*** together
- Enhances made by open and transparent submission of code on open platforms like **GitHub**.
- Quality and testing occurs by **anyone** and **everyone** who is interested.
- More and more companies and organizations are enabling access to datasets and capabilities (Data.gov, World Bank and Open Street Map).


R - ensemble with neural network?

Microsoft Azure
Got 5 minutes? Then you have time to deploy your first cloud solution.
Try Azure—free
AdChoices

5 1

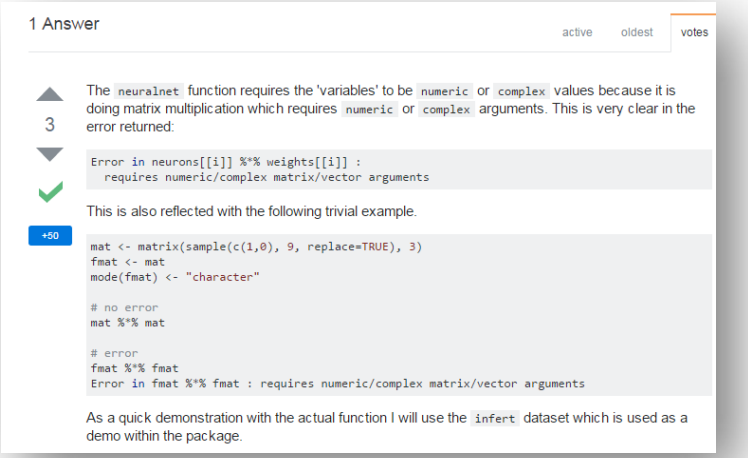
This is a small sample of my data.frame

	naiveBayesPrediction	knnPred5	knnPred10	dectreePrediction	logressionPrediction	correctCl
1	non-bob	2	2	non-bob	0.687969711847463	
2	non-bob	2	2	non-bob	0.85851872253358	
3	non-bob	1	1	non-bob	0.500470892627383	
4	non-bob	1	1	non-bob	0.77762739066215	
5	non-bob	1	2	non-bob	0.556431439357365	
6	non-bob	1	2	non-bob	0.604868385598237	
7	non-bob	2	2	non-bob	0.554624186182919	

I have factored everything

```
'data.frame':    505 obs. of  6 variables:
 $ naiveBayesPrediction: Factor w/ 2 levels "bob","non-bob": 2 2 2 2 2 2 2 2 2 2 ...
 $ knnPred5            : Factor w/ 2 levels "1","2": 2 2 1 1 1 1 2 1 2 1 ...
 $ knnPred10           : Factor w/ 2 levels "1","2": 2 2 1 1 2 2 2 1 2 2 ...
 $ dectreePrediction   : Factor w/ 1 level "non-bob": 1 1 1 1 1 1 1 1 1 1 ...
 $ logressionPrediction: Factor w/ 505 levels "0.205412826873861",..: 251 415 48 354 92
 $ correctClass        : Factor w/ 2 levels "1","2": 1 1 1 1 1 1 1 1 1 1 ...
```

I then tried to ensemble it using neuralnet


1 Answer

active oldest votes

3 +50

The `neuralnet` function requires the 'variables' to be `numeric` or `complex` values because it is doing matrix multiplication which requires `numeric` or `complex` arguments. This is very clear in the error returned:

```
Error in neurons[[i]] %*% weights[[i]] :
  requires numeric/complex matrix/vector arguments
```

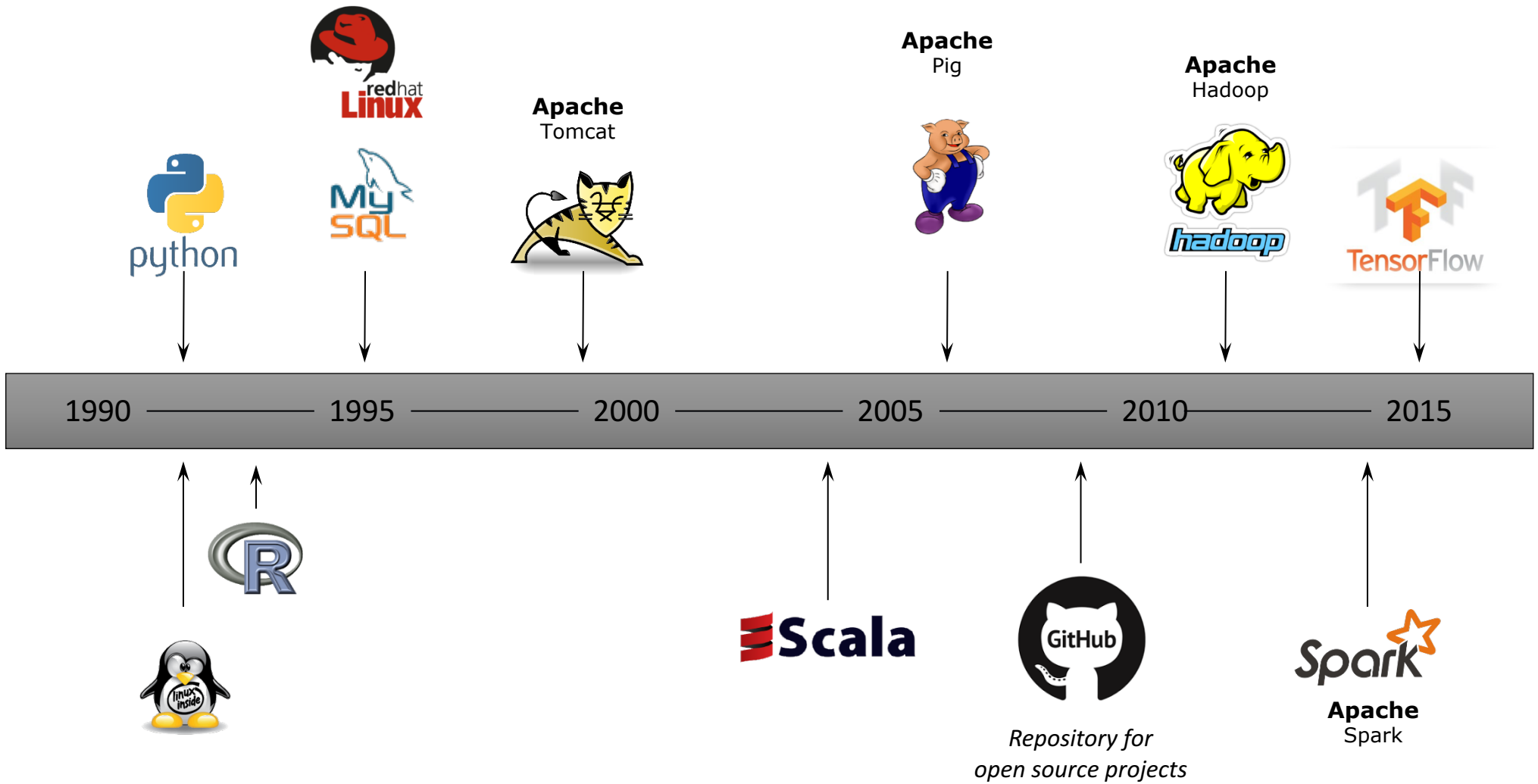
This is also reflected with the following trivial example.

```
mat <- matrix(sample(c(1,0), 9, replace=TRUE), 3)
fmat <- mat
mode(fmat) <- "character"

# no error
mat %*% mat

# error
fmat %*% fmat
Error in fmat %*% fmat : requires numeric/complex matrix/vector arguments
```

As a quick demonstration with the actual function I will use the `infert` dataset which is used as a demo within the package.

What is Open Source

"Open Source" has been around for a long time



Storage

How? Costs

Where?

Processing

How?

Who?

Where?

Programming **model**

Analytics and **Learning**

Operations

Advancements

Big Data Key Concepts & Types

Meirc Training & Consulting

PLUS SPECIALTY TRAINING

Object Based Storage vs File System Storage

Video, audio, social and images

Professional *Roles*

Hadoop

Distributed Systems

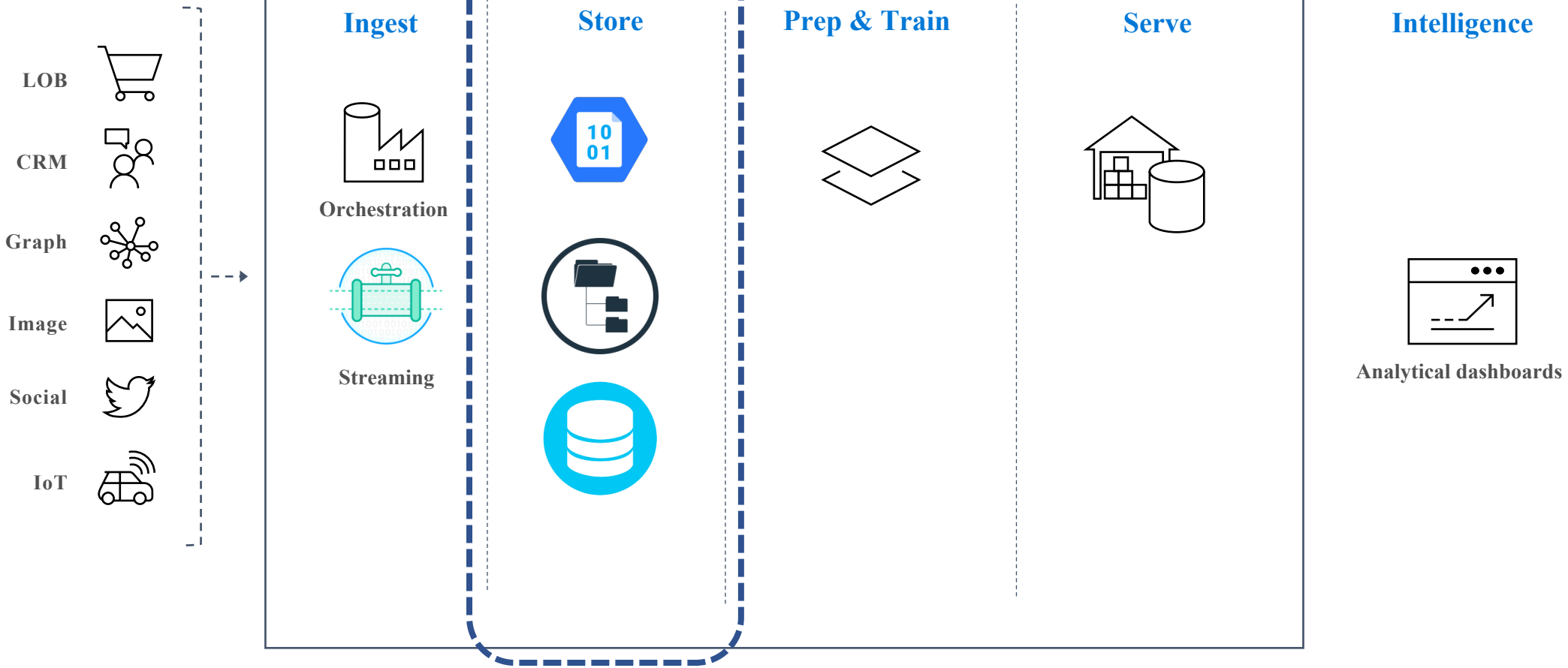
Real Time

SQL vs *NoSQL*

Data *Management*

Cloud Platforms

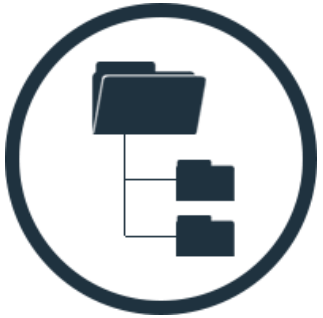
Architecture Diagram Icons



Types of Data Storage

Object Storage

Database Storage



File System Storage

Object Storage

Characteristics

No directory hierarchy

Store files in flat plane

Key-value pair access

Persistent storage

Notes

Great for unstructured, static and archival data.
Requires less metadata
Competes with Hadoop.
Low latency and intended for limited writes.
Easily scalable. Compatible with cloud storage platforms.
Handles petabytes of storage.

Icons

Cost

Types of Data

File System Storage

Characteristics

Directory hierarchy

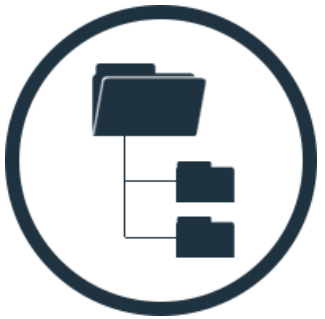
Unstructured data store

Limited data access policies

Notes

Great for unrelated data.
Most efficient ops for smaller file operations.
High latency
Intended for high read, high write

Icons



$$

Cost

Types of Data

Characteristics

Structured and managed

Relational and NoSQL

Programmatic Control

Notes

Performance monitoring is standard.
Heavy metadata.
Handles internal computations and querying.
High read, high write.
Best for connected data.
Built-in data recovery services.

Icons

$$$

Cost

Types of Data

Storage Technologies

Object

Azure Blob

Google Cloud

Hadoop

File System

Amazon

Azure Data Lake Gen2

DBMS

Columnar/NoSQL

Big Data Analytics Professional Roles

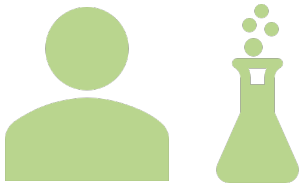
Meirc Training & Consulting

PLUS SPECIALTY TRAINING

Professional Roles

Data Engineer

- Transform data
- Create and schedule batch
- Streaming ETL jobs
- Manage platforms
- Agile development
- Knowledge of:
  - Platforms
  - Java, C++
  - Scripting
  - Python, …



Data Scientist

- Explore data
- Create and deploy analytical models
- Transform data
- Perform advanced analytics tasks
- Knowledge of:
  - Statistics
  - Machine Learning
  - Linear Algebra
  - R, …

Business Analyst

- Write SQL queries
- Generate reports
- Explore data
- Analyze and visualize data
- Knowledge of basic analysis tools such as:
  - Databricks notebooks
  - Power BI
  - Tableau

Big Data Architectures and Paradigms

Meirc Training & Consulting

PLUS SPECIALTY TRAINING

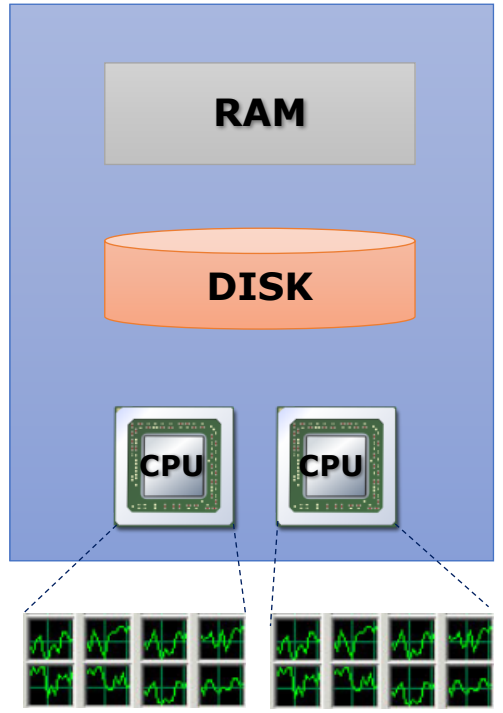
Whether it is a personal computer or an enterprise server, a computer is made up of …

Three essential components

MEMORY

STORAGE

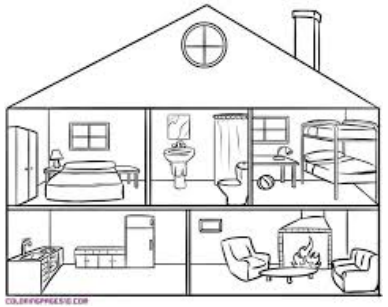
PROCESSING



RAM

(memory)

Pantry or attic



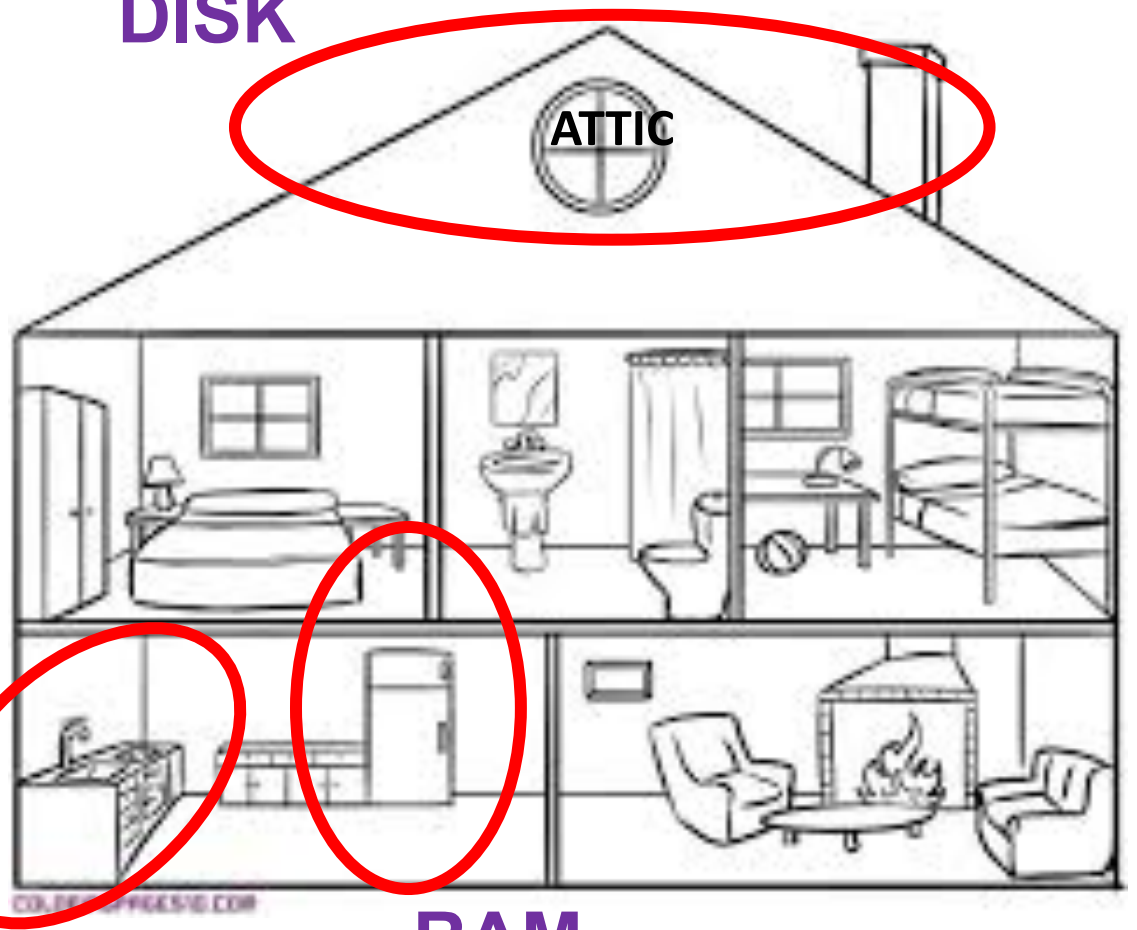
Hard Disk

(Storage)

CPU

(Processing)

DISK

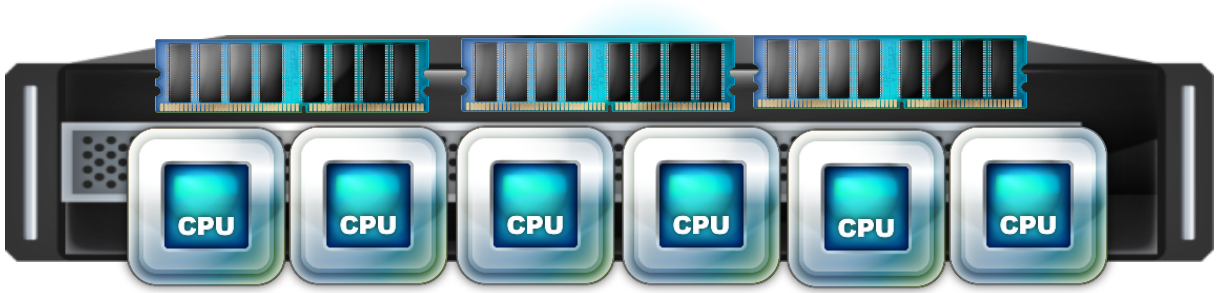


PROCESSOR

RAM

Single Machine

Workstation



Server

… how does this scale with data?

Distributed Systems

Meirc Training & Consulting

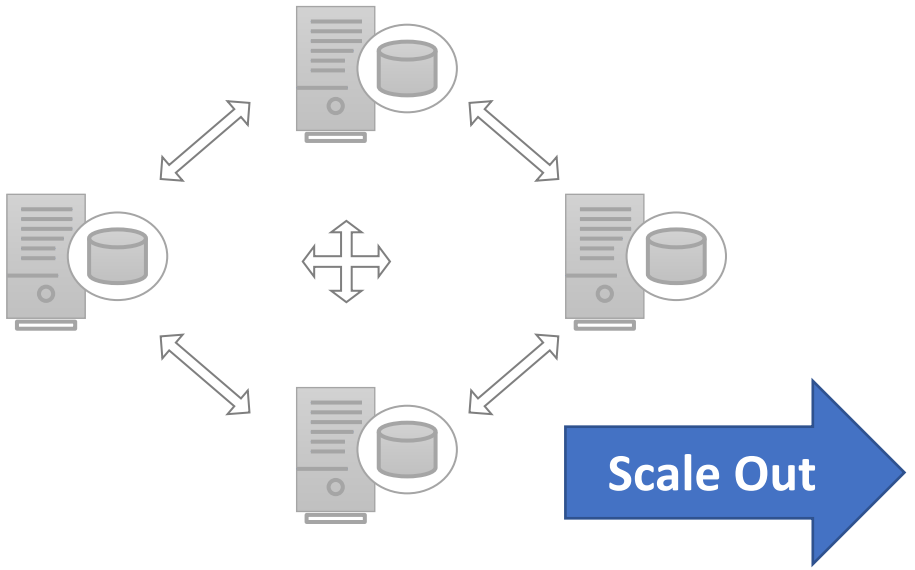
PLUS SPECIALTY TRAINING

SMP
↓
Symmetric **M**ulti-**P**rocessing

System where processors share resources

Scale Up = larger server

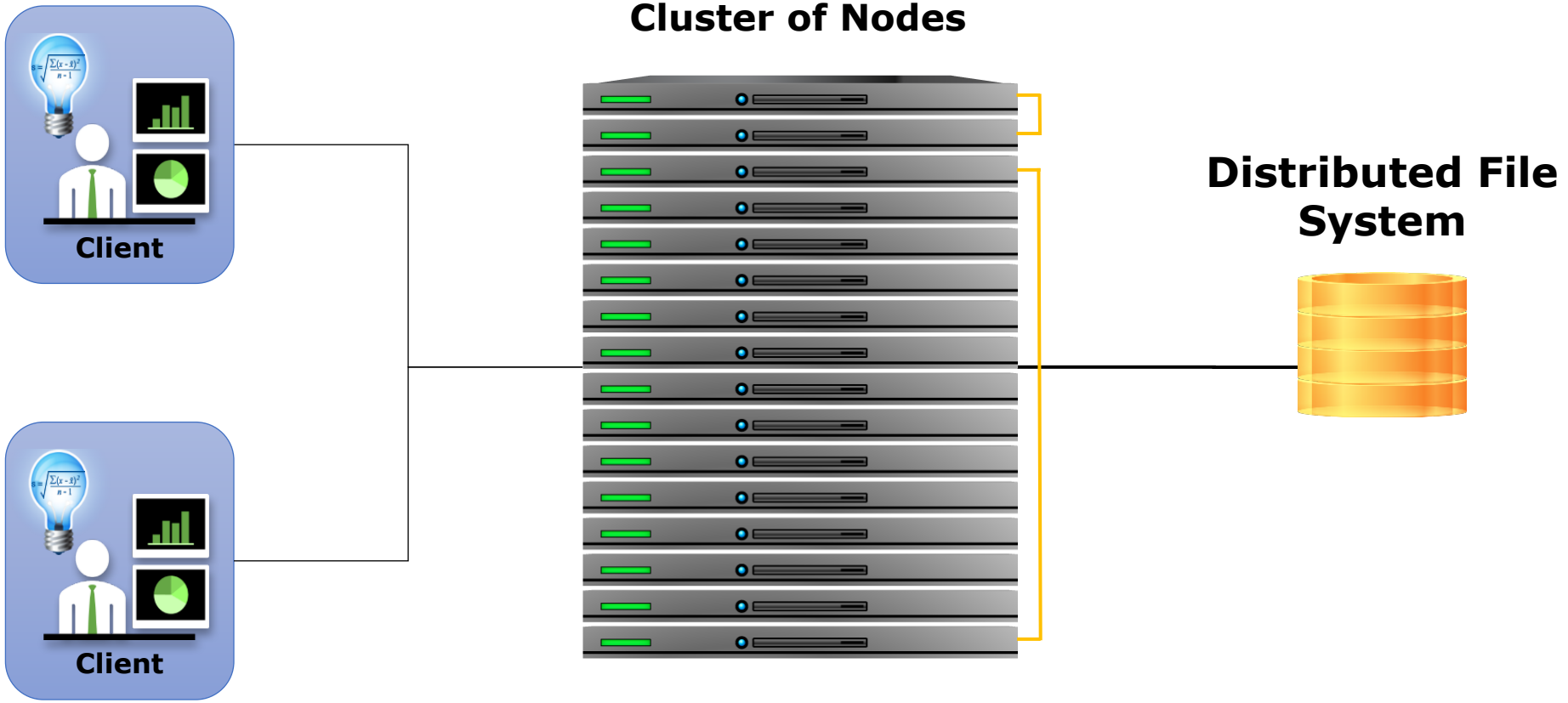
MPP
↓
Massively **P**arallel **P**rocessing



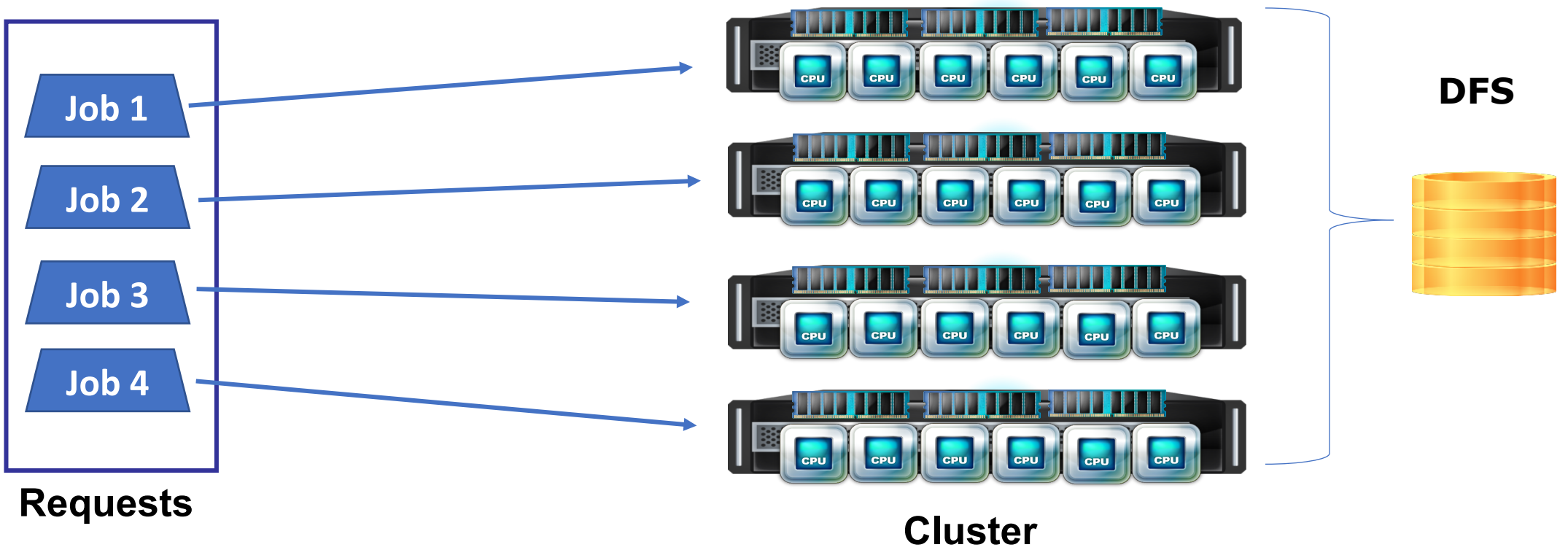
System of processors that **share nothing**

Scale Out = add more servers

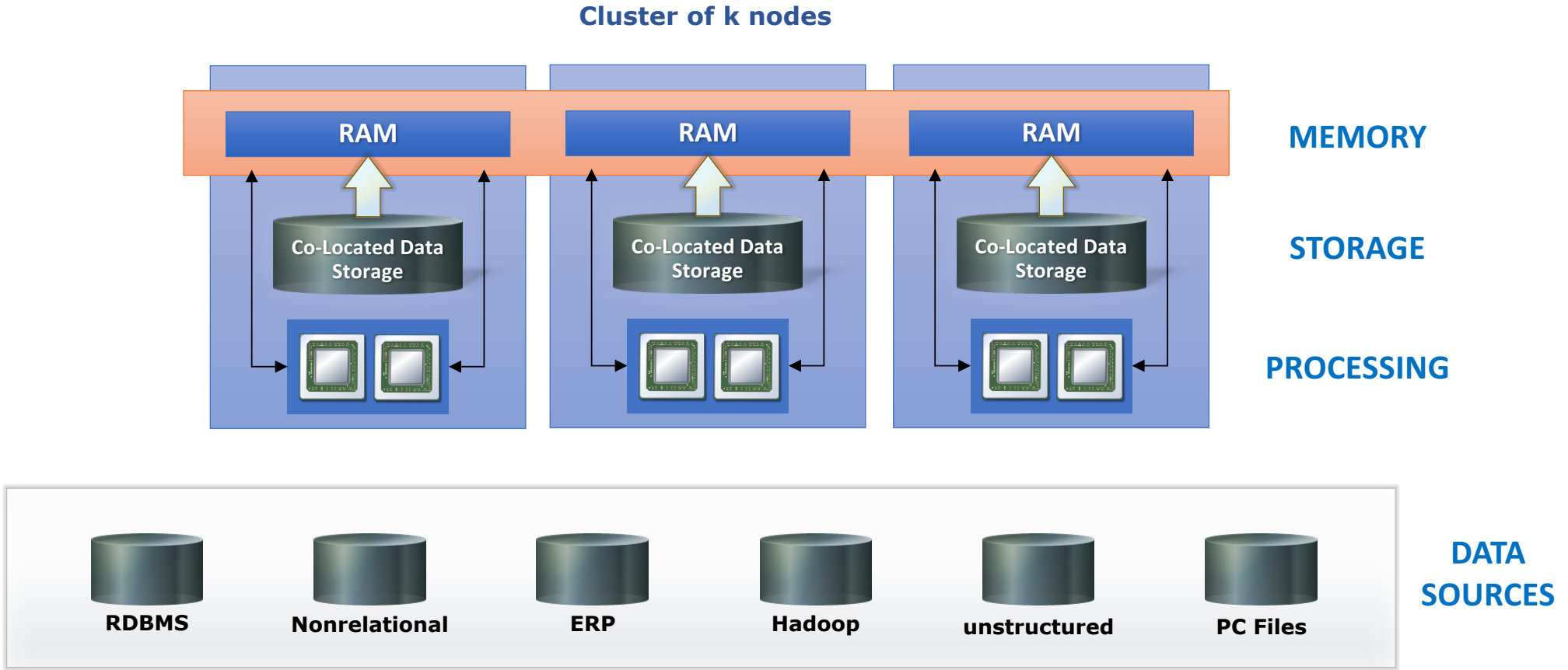
Distributed System



Parallelize Jobs



MPP vs Distributed In-Memory



What Do We Need?

Storing Big Data

Hadoop Overview

HDFS

RDBMS vs NoSQL

Hands-on-Lab

Streaming Storage

Data Warehousing vs Data Mart

Data Lake

Lambda Architecture

Case Study: Big Data Storage

What is Hadoop?

Hadoop is an open-source software for:

Scalable

Reliable

Distributed computing

A tool that can **store** and **analyze** massive amounts of data.

Allows the use of **simple programming** models for distributed large data sets across **clusters**.

Built for **distributed processing** across clusters of ***commodity hardware*** (off the shelf hardware)

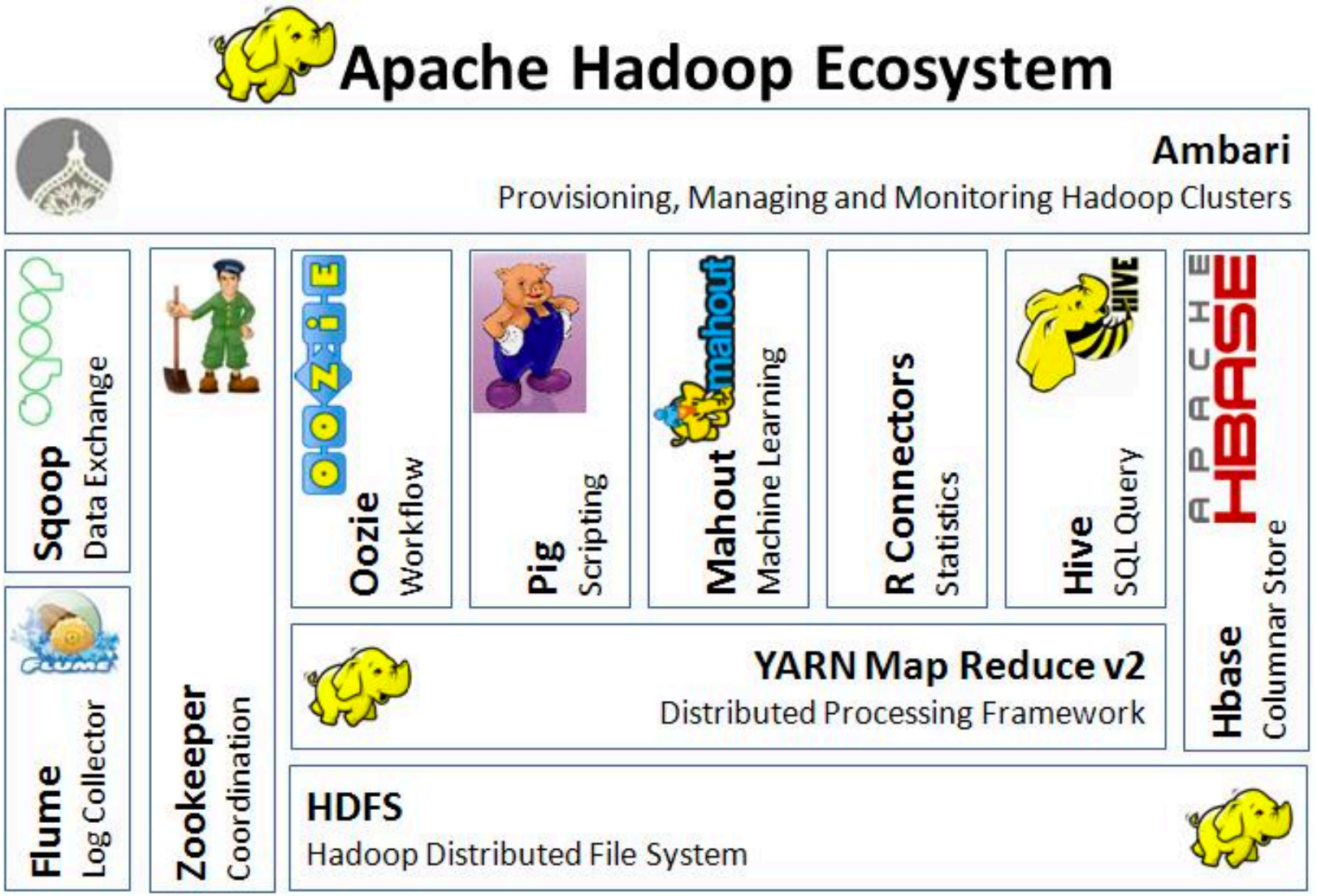
What is Hadoop?

At it's core **Hadoop** has two primary components:

- ***MapReduce*** to process data
- **HDFS** to store data.

Designed with the **assumption** that ***hardware failures*** are common and should be **automatically handled** by the framework.

- Hadoop is designed to:
  - Scale up from single servers to thousands of machines, each offering local computation and storage.
  - Detect and handle failures at the application layer, that is delivering a highly-available service on top of a cluster of computers, each of which may be prone to failures.


Apache Hadoop Ecosystem
Ambari
Provisioning, Managing and Monitoring Hadoop Clusters
Sqoop
Data Exchange
Flume
Log Collector
Zookeeper
Coordination
Oozie
Workflow
Pig
Scripting
Mahout
Machine Learning
R Connectors
Statistics
Hive
SQL Query
APACHE HBASE
Hbase
Columnar Store
YARN Map Reduce v2
Distributed Processing Framework
HDFS
Hadoop Distributed File System

Hadoop HDFS

Characteristics

Distributed file system

Open sourced technology

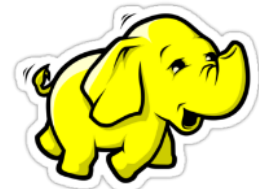
Commodity hardware

Persistent storage

Notes

Great for unstructured, static and archival data.
Great On-premise option.
HIGH latency and intended for limited writes.
Scale with design. Not common with cloud storage platforms.
Handles petabytes of storage.
Difficult to maintain.

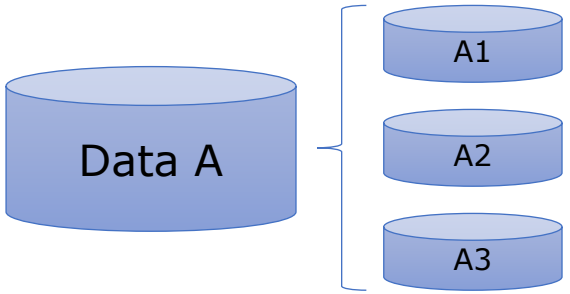
Icons

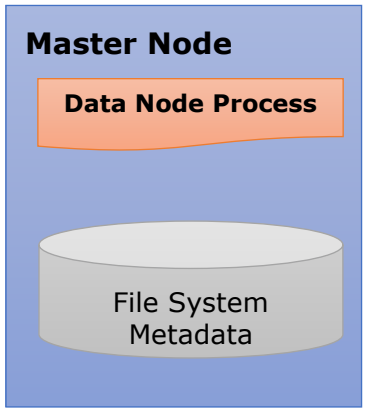


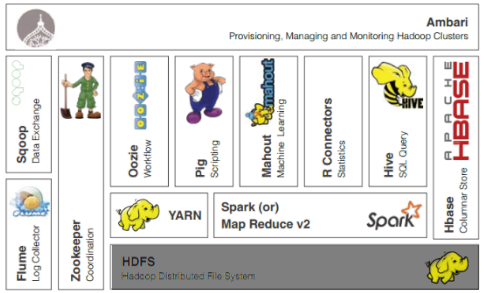
~$

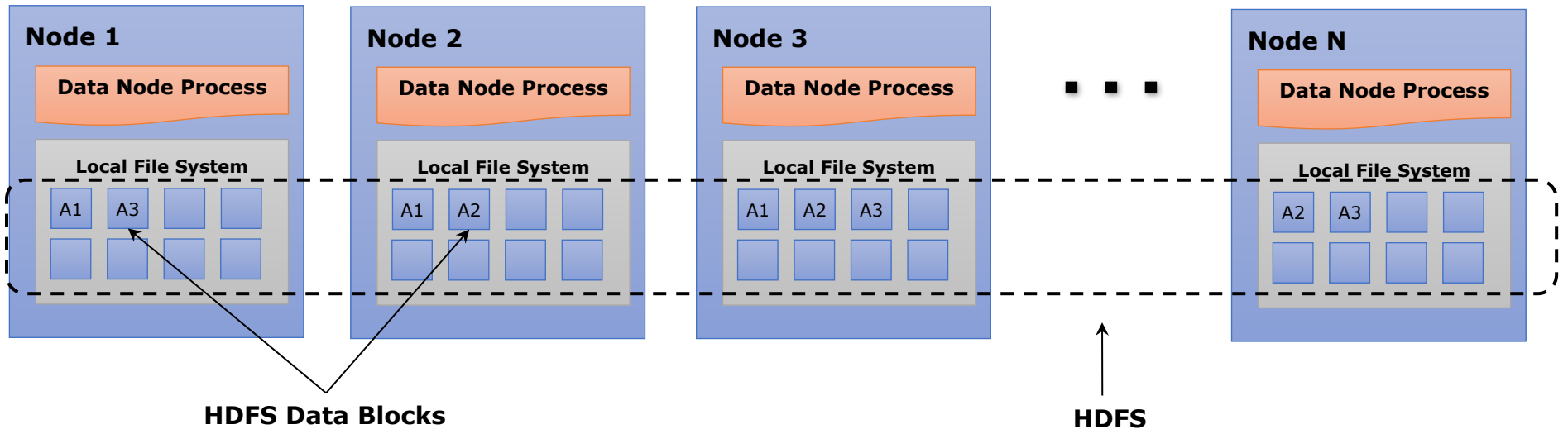
Cost

Types of Data

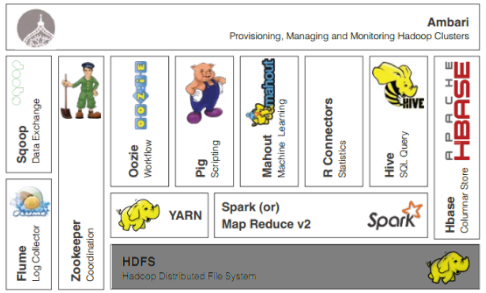

Data A
A1
A2
A3


Master Node
Data Node Process
File System
Metadata


Ambari
Provisioning, Managing and Monitoring Hadoop Clusters
Sqoop
Data Exchange
Oozie
Workflow
Pig
Scripting
Mahout
Machine Learning
R Connectors
Statistics
Hive
SQL Query
Hbase
Columnar Store
Flume
Log Collector
Zookeeper
Coordination
YARN
Spark (or)
Map Reduce v2
Spark
HDFS
Hadoop Distributed File System

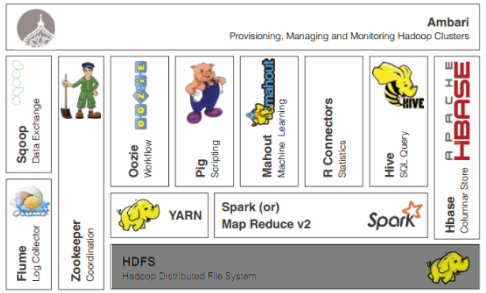

Node 1
Data Node Process
Local File System
A1
A3
Node 2
Data Node Process
Local File System
A1
A2
Node 3
Data Node Process
Local File System
A1
A2
A3
Node N
Data Node Process
Local File System
A2
A3
HDFS Data Blocks
HDFS

Hadoop HDFS


Ambari
Provisioning, Managing and Monitoring Hadoop Clusters
Sqoop
Data Exchange
Flume
Log Collector
Zookeeper
Coordination
Oozie
Workflow
Pig
Scripting
Mahout
Machine Learning
R Connectors
Statistics
Hive
SQL Query
Hbase
Columnar Store
YARN
Spark (or)
Map Reduce v2
HDFS
Hadoop Distributed File System

File
Block 1
Block 2
Block 3
Node 1
Block 1
Block 3
Node 2
Block 2
Node 3
Block 1
Block 2
Node 4
Block 3

Hadoop



cloudera
Ask Bigger Questions

+

Hortonworks

IBM

MAPR

Pivotal

RDBMS vs NoSQL

Relational Database (SQL)
Supports powerful query language.
It has a fixed schema.
Follows ACID (Atomicity, Consistency, Isolation, and Durability).
Supports transactions.

NoSql Database
Supports very simple query language.
No fixed schema.
It is only “eventually consistent”.
Does not support transactions.

To schema or not to schema…

A **NoSQL** (often interpreted as Not only SQL) **database**

- Storage and retrieval of data that is modeled in means other than the tabular relations used in relational **databases**.
- Looks at wholes instead of parts.
- Traditional storage or relational DBs are unable to accommodate increasing number of observations

Driving reasons

✓ Design simplicity

✓ Horizontal scaling

✓ Finer control over availability

"**Relax** storage" requirements **solution**

✓ Denormalize

✓ Loosen consistency

✓ Loosen schema

Characteristics

Distributed RDBMS

Open sourced technology

DBMS on top HDFS

Data Warehouse functions

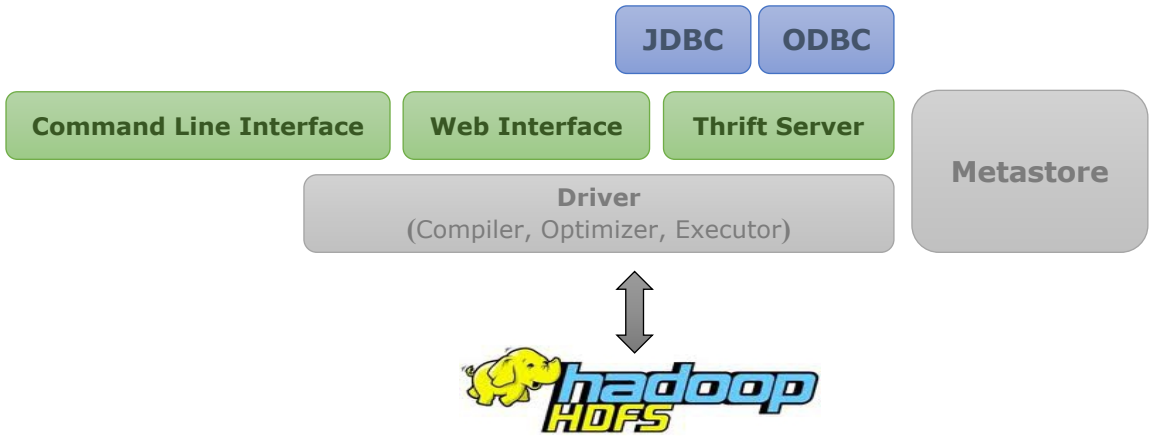
Notes

Great for data summarization, query and analysis.
Great On-premise option.
SQL-like interface.
Translates SQL into Mapreduce
Handles petabytes of storage.
Hive-metastore offers flexible schema.
Serialization/deserialization

Icons

$$

Cost



PostgreSQL or MySQL

Characteristics

Open source RDBMS

Ansi-SQL compliant

Predefined schemas

Persistent storage

Notes

Great for related information.
Great On-premise or cloud option.
Low latency and heavy read/writes.
Handles terabytes of storage.
Popular for data scientists.
Simple to maintain.
Distributed functionality.

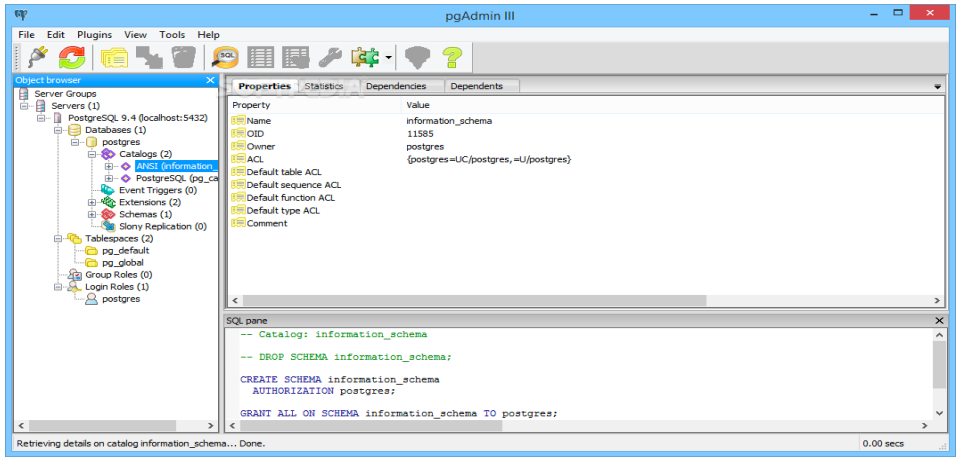
Icons

PostgreSQL

MySQL®

$$

Cost



MSFT SQL Server

Characteristics

MSFT developed RDBMS

Ansi-SQL compliant

Predefined schemas

Persistent storage

Notes

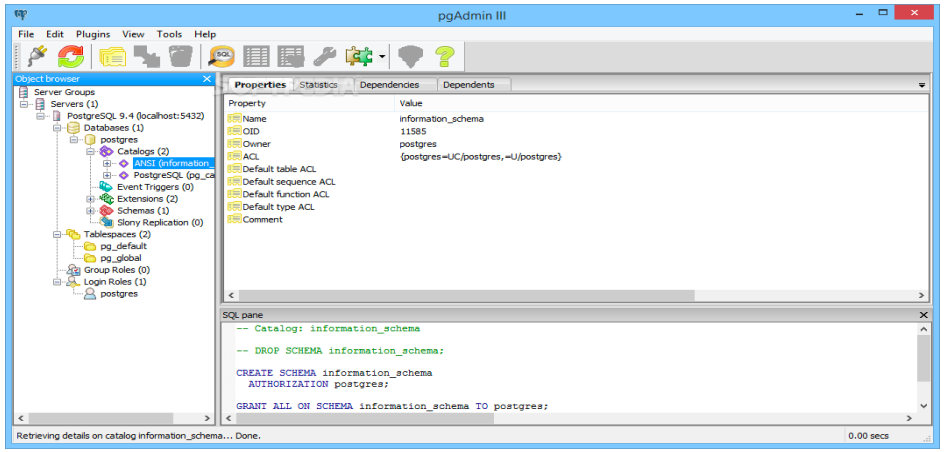
Great for related information.
Great On-premise or cloud option.
Low latency and heavy read/writes.
Handles terabytes of storage.
Popular for data scientists and applications.
Simple to maintain.
Distributed functionality.

Icons

Microsoft® SQL Server®

$$

Cost



Apache HBASE

Characteristics

Open source, distributed

Column-oriented database

Compatible with HDFS

Notes

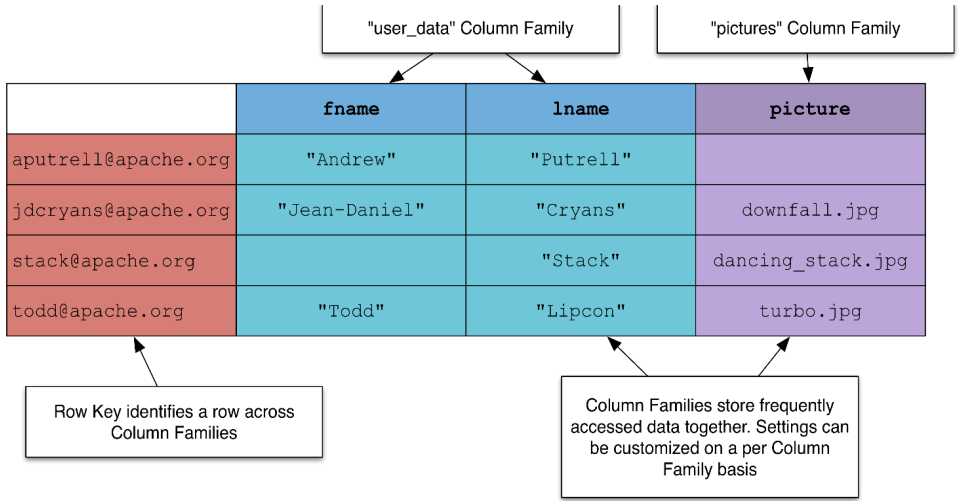
Master-worker architecture
Great On-premise option.
Best for data lake use cases.
Heavy read, heavy writes.
No SQL-like facility to interface
Handles petabytes of storage.
Challenges with availability and maintenance.

Icons

APACHE HBASE

$

Cost



Characteristics

Open source, distributed and decentralized

Column-oriented database

Persistent storage

Notes

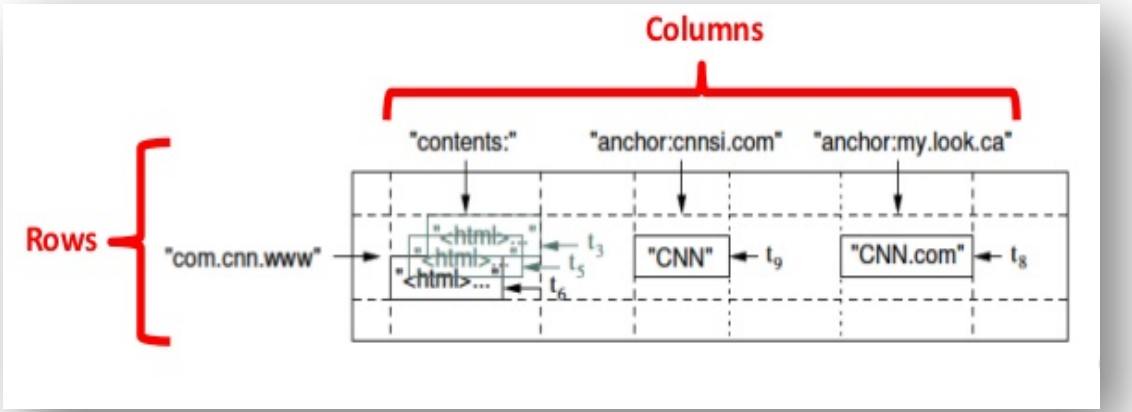
No master-worker architecture
Great On-premise option.
Low latency
Heavy read, heavy writes.
Scale on demand. Compatible with cloud storage platforms.
Handles petabytes of storage.
Global replication

Icons

cassandra

$$

Cost



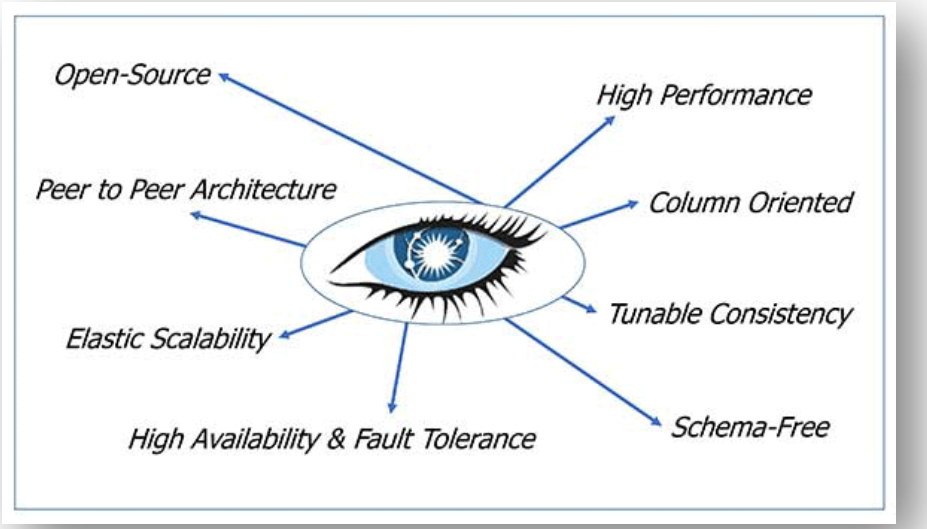
Deployment within more than 1,500 companies.

Over 75,000 nodes storing over 10 PB of data.

2,500 nodes, 420 TB, over 1 trillion requests per day.

Search engine, involving 270 nodes, 300 TB, over 800 million requests per day.

Over 100 nodes, 250 TB



Constant *The Weather Channel* *Reddit* *Intuit*

Hulu

CERN *Comcast* *GitHub* *GoDaddy* *Instagram*

Hands-on Lab

Meirc Training & Consulting

PLUS SPECIALTY TRAINING

Azure Databricks

Hands-on-lab

Streaming Data Storage

Architecture Diagram Icons

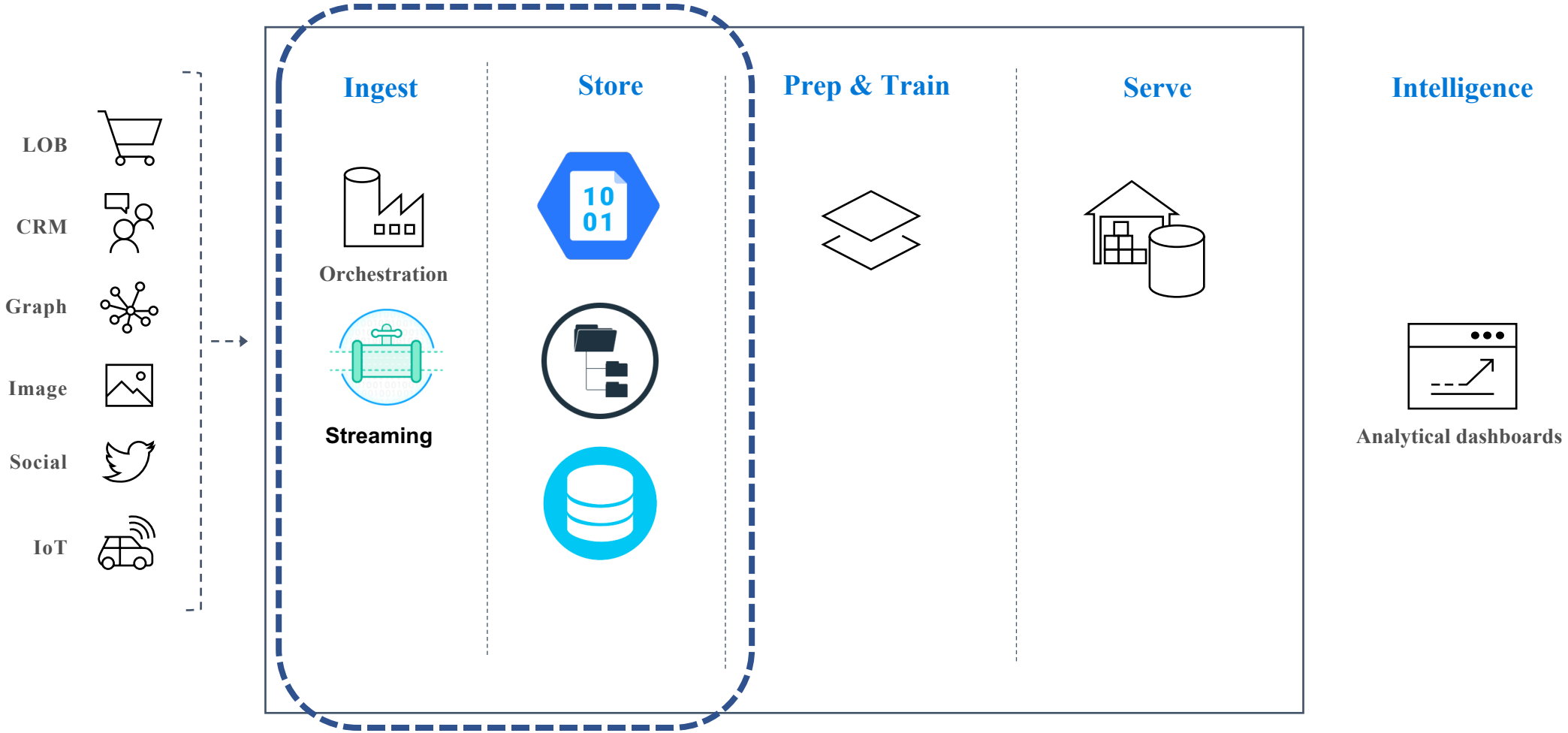
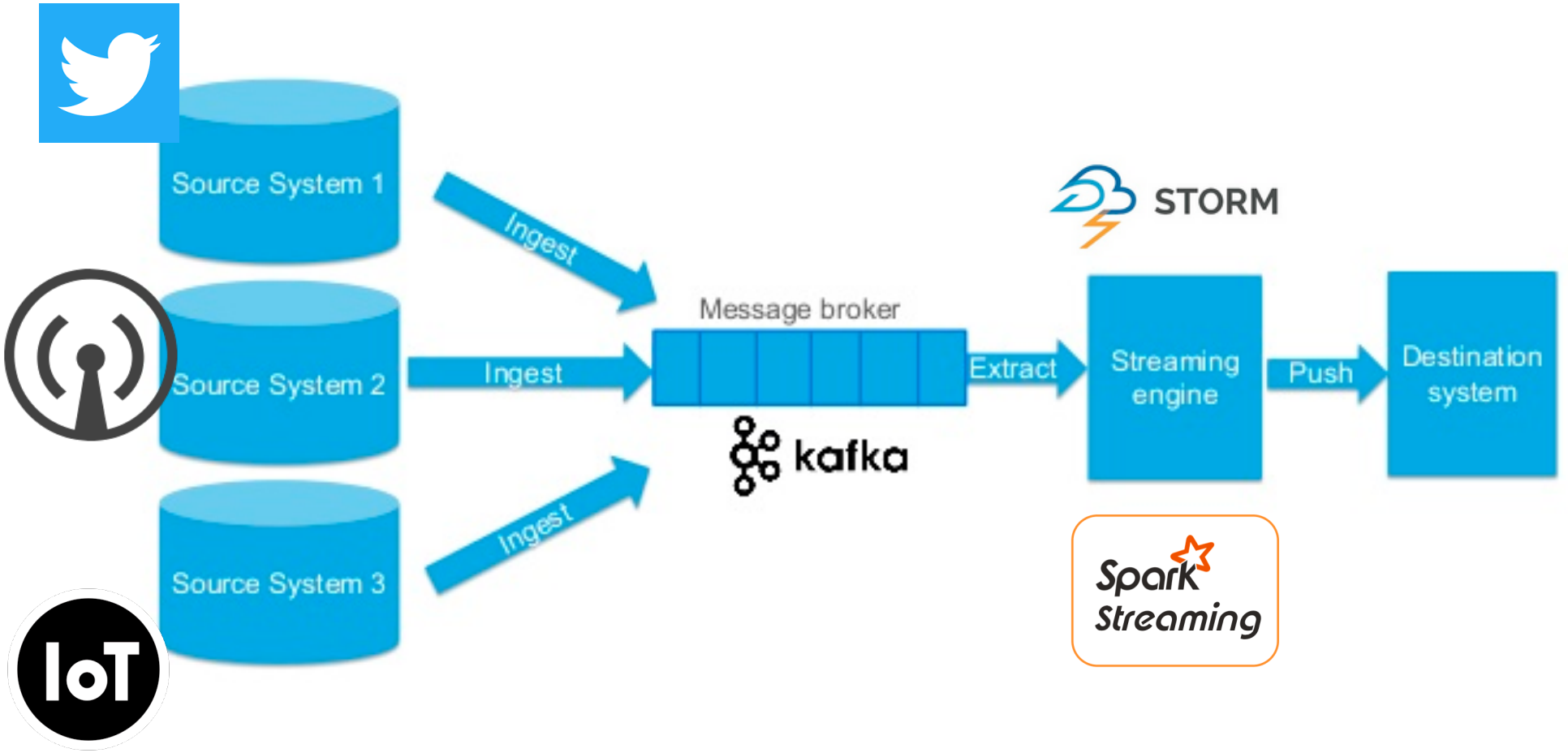


Image of Streaming Diagram



Characteristics

Open source

Distributed cluster system

Message broker

Stores records in topics

Notes

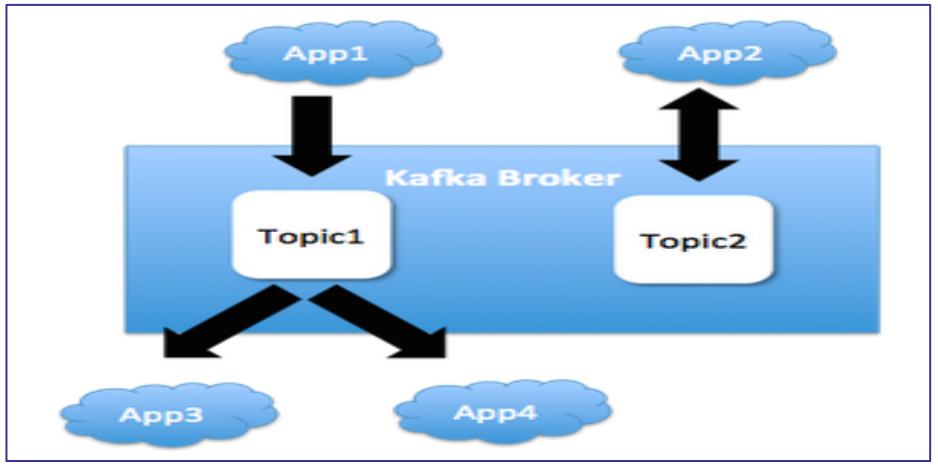
Each record consists of key, value, timestamp.
Good for real-time streams and get data between systems.
Great On-premise option.
Low latency
Heavy read, no writes.
Compatible with cloud storage platforms.
Publish and subscribe.

Icons

kafka

$$

Cost

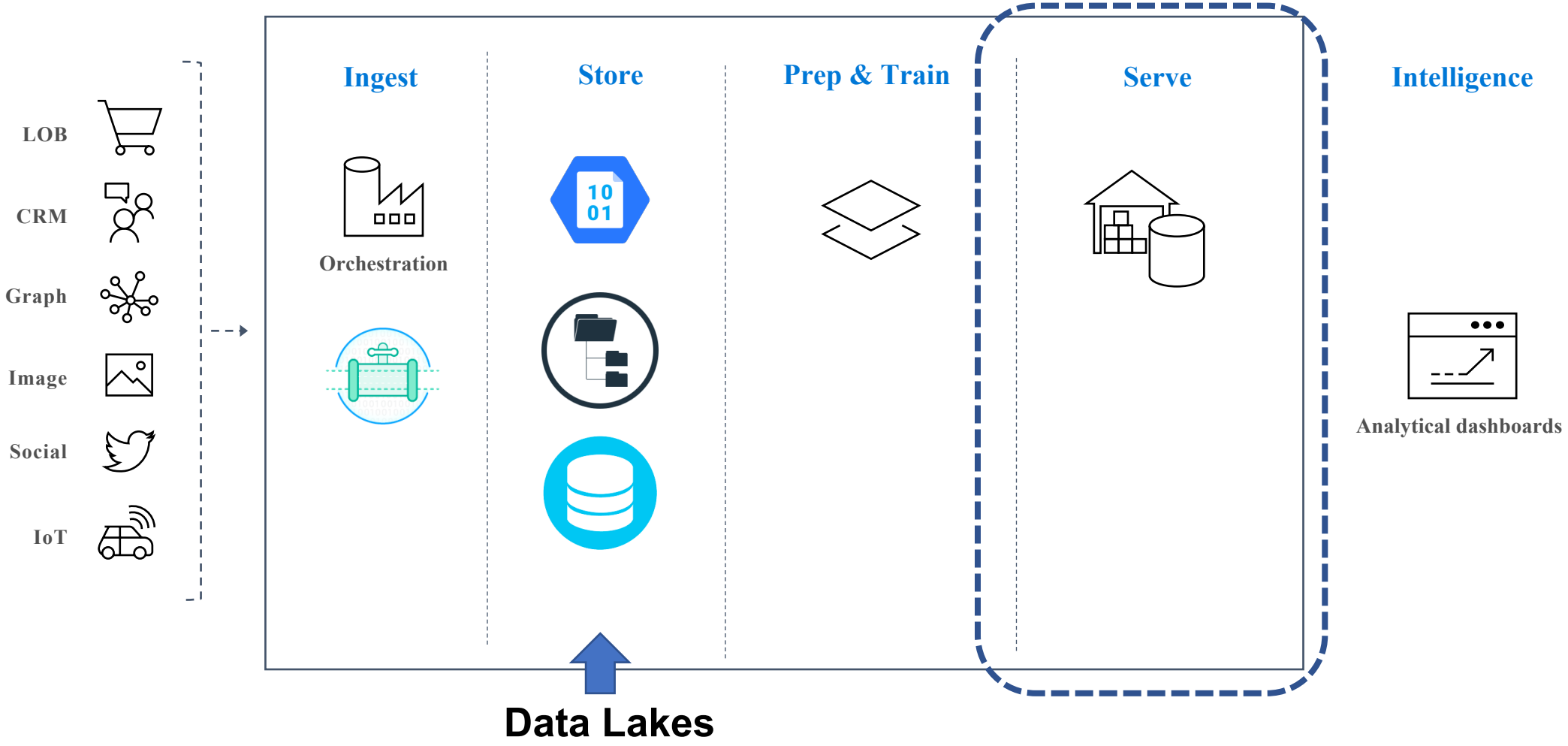


Data Warehousing
vs
Data Mart
vs
Data Lake

Meirc Training & Consulting

PLUS SPECIALTY TRAINING

Architecture Diagram Icons



Data Warehouse

Characteristics

Storing data from different sources

Distributed system

Highly organized and structured

Subject specific

Notes

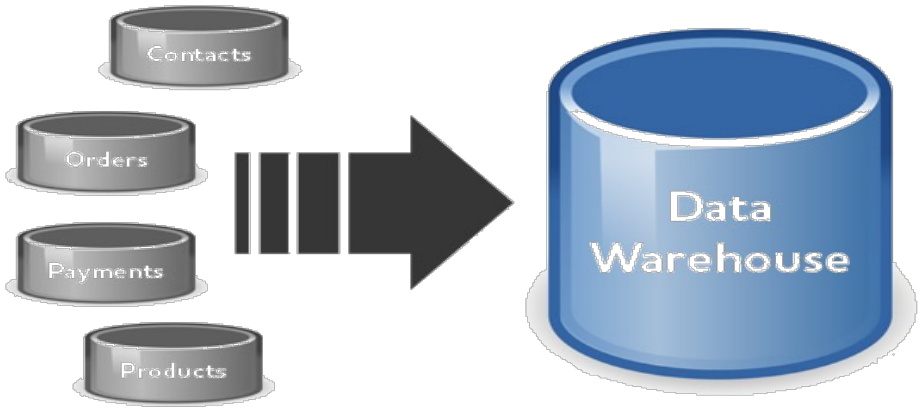
Subject-oriented historical data.
On-premise or cloud option.
Low latency
Heavy read, low writes.
Non-volatile and detailed information.
Can handle terabytes of data.
Business analysis and decision support.
Accuracy and consistency.
Enterprise-wide

Icons

snowflake

$$$

Cost



Data Mart

Characteristics

Small scale DW

Highly organized and structured

Subject/department specific

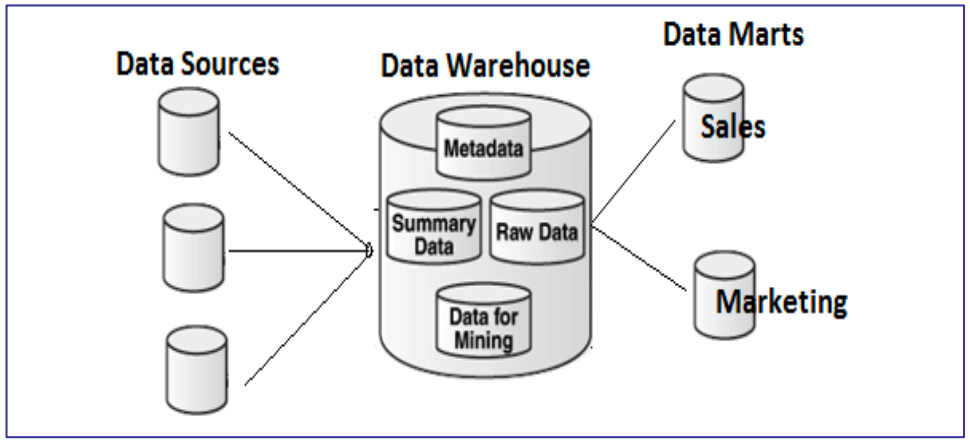
Notes

Designed to store frequently used information.
Store less than DW for efficiency.
Very Low latency
Heavy read, low writes.
Replicated or Stand-alone
Accuracy and consistency.
Department specific.

Icons

$$$

Cost



Data Lake

Characteristics

Data storage

No organization required

Raw data content

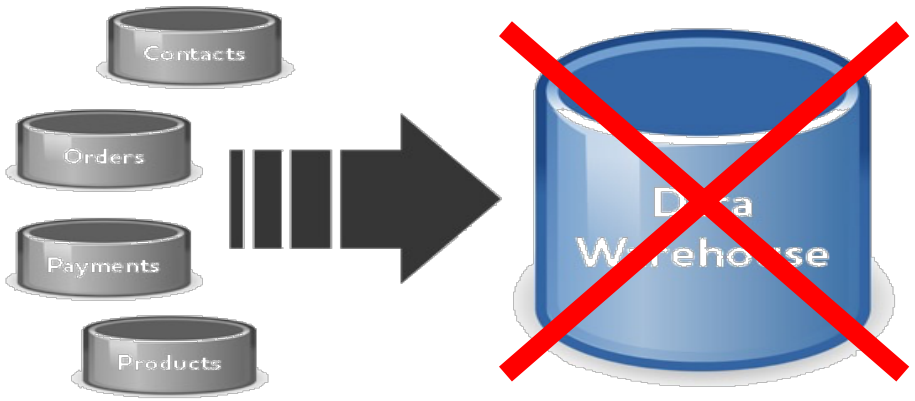
Notes

Designed to provide access for data discovery.
Raw data with no defined purpose.
Medium latency
Heavy read, heavy writes.
Not ACID compliant.
Explore the unknown.

Icons

$

Cost



Lambda Architecture

Meirc Training & Consulting

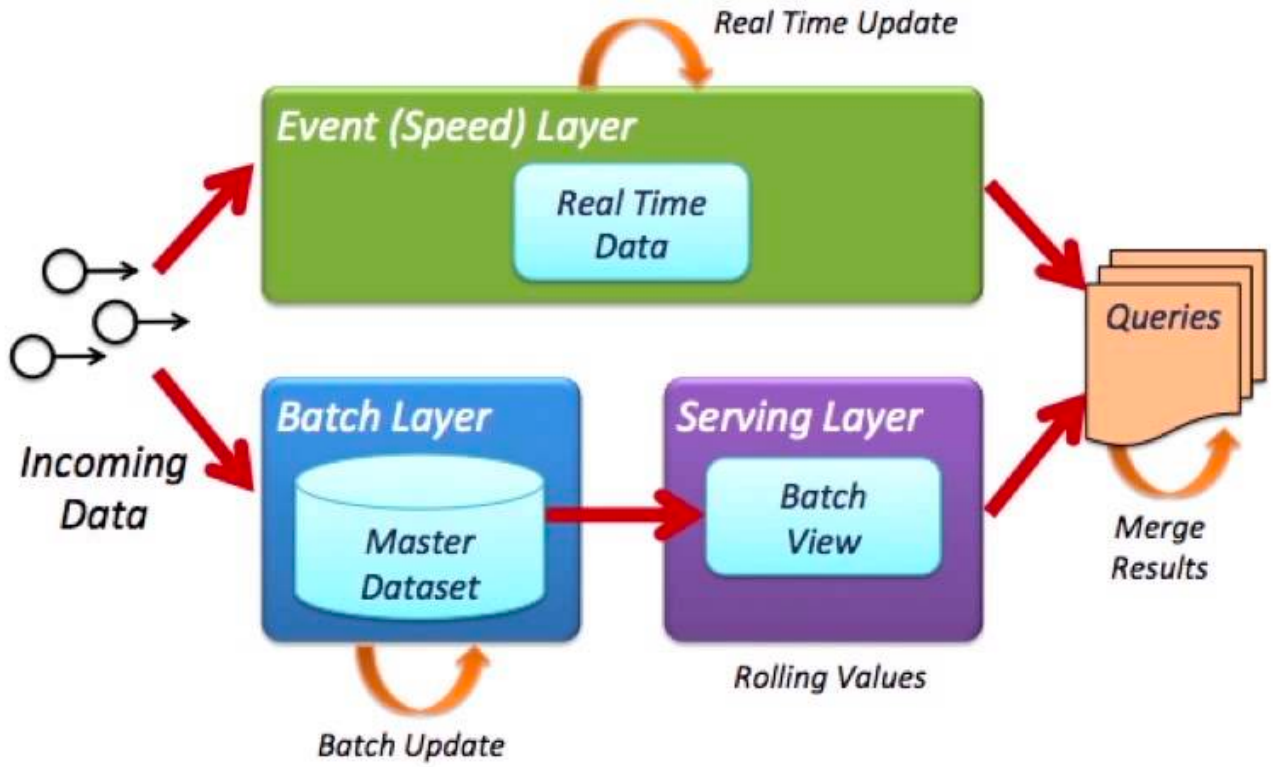
PLUS SPECIALTY TRAINING

Characteristics

Aimed to support analytics applications

Combines Batch and Real-time storage and processing

3 Layers:
Batch
Speed
Serving



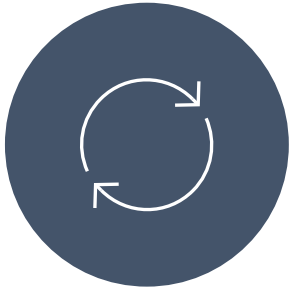
Case Study: Big Data Storage

Meirc Training & Consulting

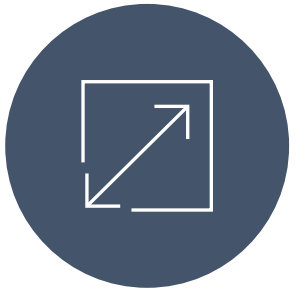
PLUS SPECIALTY TRAINING

Description

Evolv helps large global companies make better hiring and management decisions through predictive analytics. Evolv crunches more than 500 million data points on gas prices, unemployment rates, and social media usage to help companies predict when an employee is most likely to leave.



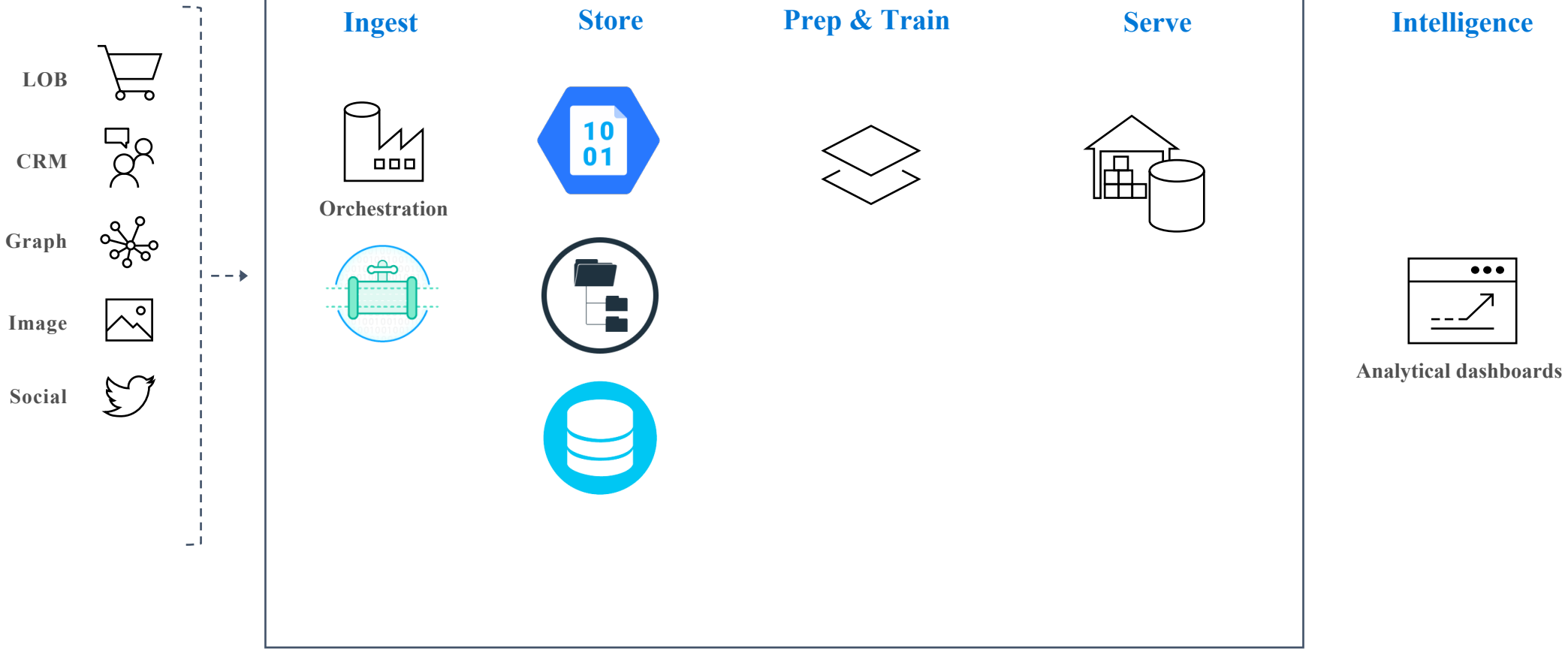
Identify the challenges



List the desired outcome

Begin designing the architecture diagram

Architecture Diagram Icons



EVOLV Diagram Example