

# Profit from all your Data

Joao Salcedo – System Engineer LATAM

February 2016

# Cloudera Fact Sheet

What we do?

The modern platform for data management and analytics.

Our Mission

Allow our customers to use *all their data* to *ask bigger questions*.

Enterprise-Grade

800+ Enterprise subscription software customers.

Partner Ecosystem

2,000+ partners.

Global Scale

Operations in 20 countries, 24x7 global support, education.

Open Source Innovation

Leading developers & contributors to Hadoop community.

Employees

1,000+

Founders

**ORACLE** **YAHOO!** **facebook.** **Google**

Headquarters

Palo Alto. Founded 2008. Private company.

# Customer success across all industries

**Financial Services** BBVA Santander BARCLAYS Experian MasterCard WESTERN UNION *moving money for better*

**Telecom** vodafone SFR T-Systems BT Comcast QUALCOMM Skybox Imaging

**Healthcare & Life Sciences** Cerner Children's Healthcare of Atlanta explorys KAISER PERMANENTE. MONSANTO RelayHealth

**Media & Technology** intel AUTO SCOUT24 box CBS Interactive CONCUR EPSILON MediaMath NetApp yp

**Retail & CP** SONAE Carrefour ebay COSTCO WHOLESALE MARKS & SPENCER otto group overstock.com

**Public Sector** HM Revenue & Customs CSC LOCKHEED MARTIN Revenue C  in agus Custaim na h  ireann Irish Tax and Customs SAN JOS   STATE UNIVERSITY SOCIAL SECURITY ADMINISTRATION Minnesota.gov U.S. ARMY

# Partner Focused. Partner Engineered.

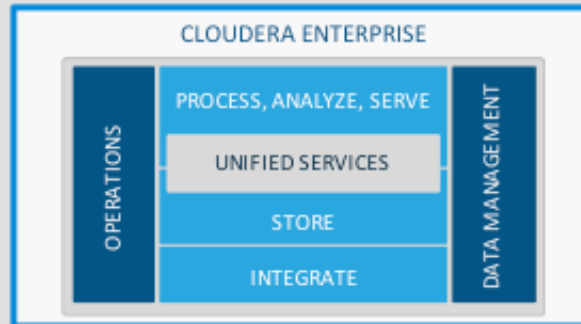
## Applications



## Operational Tools



## Data Systems



## Infrastructure



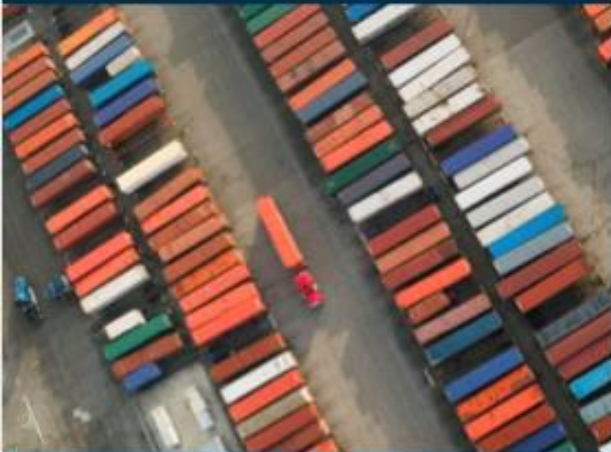
More than 2,000 partners ensure compatibility with existing investments, lower skill barriers, and help maximize value from your data.

## System Integration



# Why is Big Data Happening Now?

## Instrumentation



Everything that can be measured  
will be measured

## Personalization



Employees and customers expect  
more personal interactions, but not  
at the cost of their privacy.  
The age of *"segment of 1"*.

## Advanced Analytics



The most innovative companies  
embrace experimentation,  
predictive analytics and agility.



# Why is Our Legacy Data Architecture not Enough?



## Limited Data

**Inefficient** to manage existing data, let alone **new data**, while maintaining **performance at scale**.

Data must be **structured** first.



## Limited Insights

Limited access and poor data views cause **bad analysis** and **low adoption**.

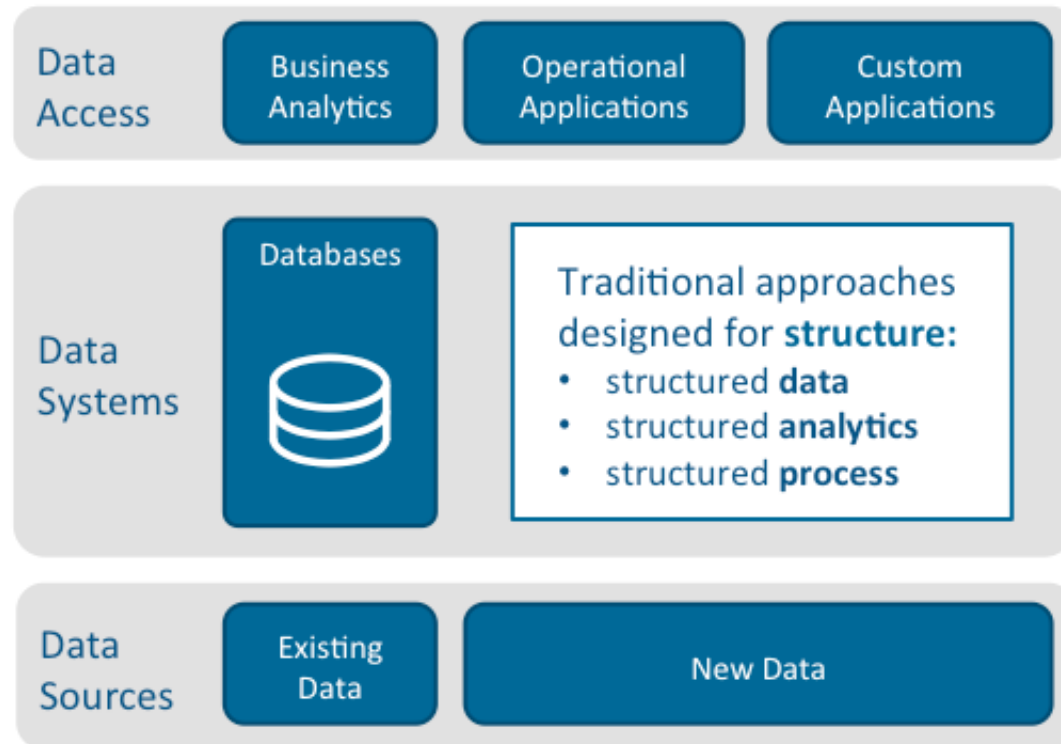
Hard to put insights into **action**.



## Complex Architecture

More data, more users, and more tools create **complexity**.

Gaps in unified security give access to **bad actors**.



# “Data Centric Apps powered by Hadoop / Cloudera”



U B E R

The world's largest taxi company owns ZERO vehicles.



The world's largest accommodation provider owns ZERO real estate.

facebook.

The world's most popular media owner creates ZERO content.



The world's most valuable retailer has ZERO inventory.

# Top Use Cases For Hadoop-based Solutions

## Operational Data Store



- Active Archive
- ETL Off-load
- EDW Optimization

## Data Discovery & Analytics



- Self-service BI
- Advanced Analytics
- Data Science

## Operational Analytics



- Recommendation Engine
- Event Detection
- Real-time Analytics

## Vertical Solutions



- FinServ: Compliance
- Telco: Churn Reduction
- Mfg: Yield Improvement

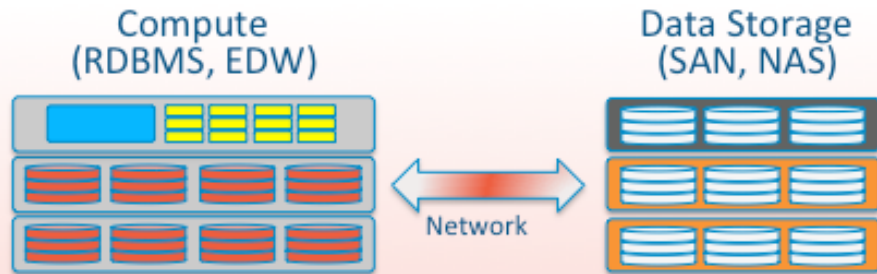
Improve Operational Efficiencies

Drive Business Value



# Why Hadoop vs Data Warehouse

## The Old Way

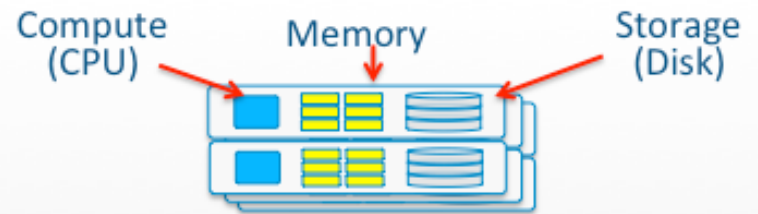


**Expensive, Special purpose, "Reliable" Servers**  
**Expensive Licensed Software**

- Hard to scale
- Network inevitably becomes a bottleneck
- Only handles structured/relational data
- Difficult to add new fields & data types

**Expensive & Unattainable**

**\$30,000+ per TB**



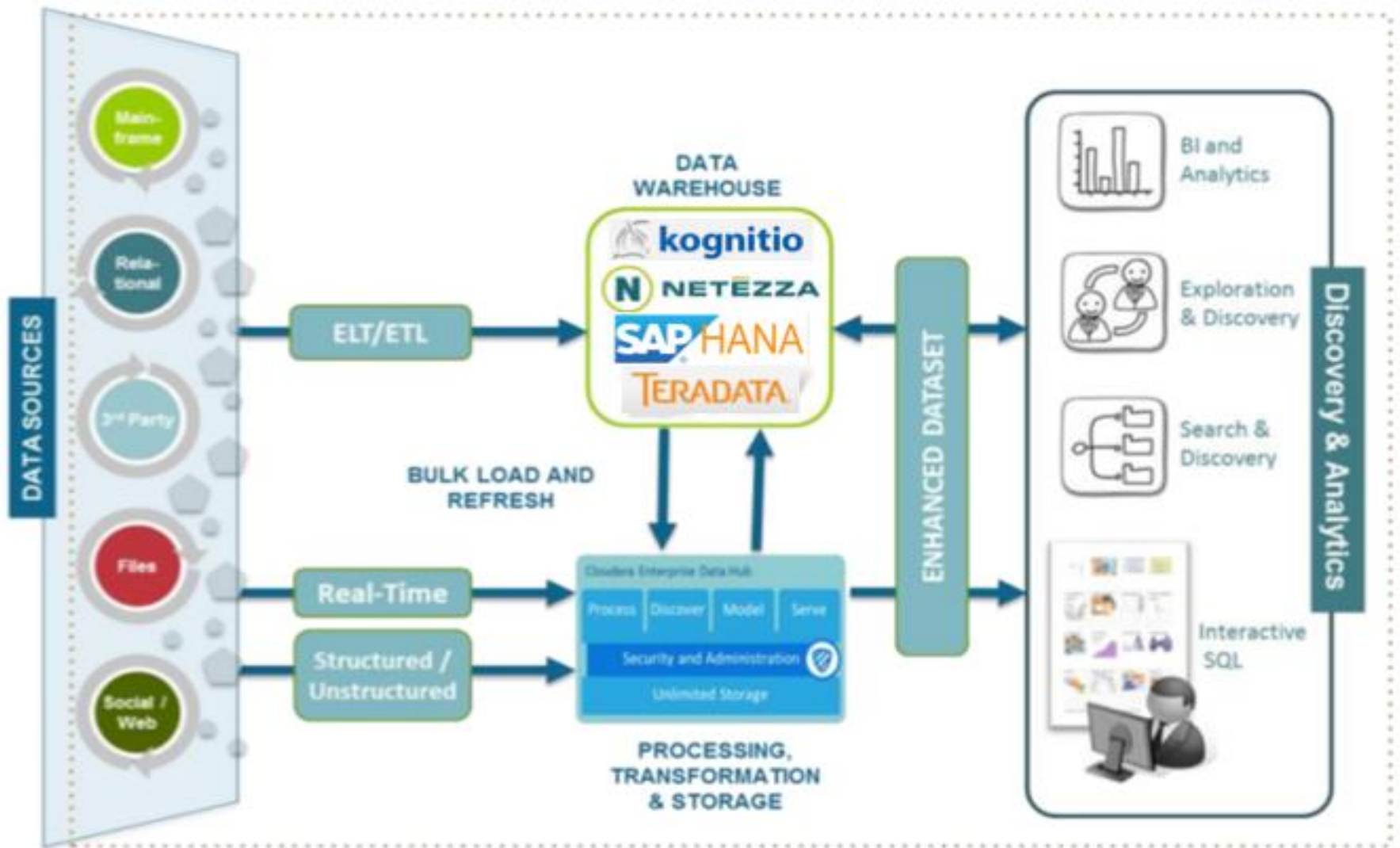
**Commodity Servers**  
**Hybrid Open Source Software**

- Scales out indefinitely and linear
- Network eliminated as a bottleneck
- Easy to ingest any type of data
- Agile "schema-on-read" data access

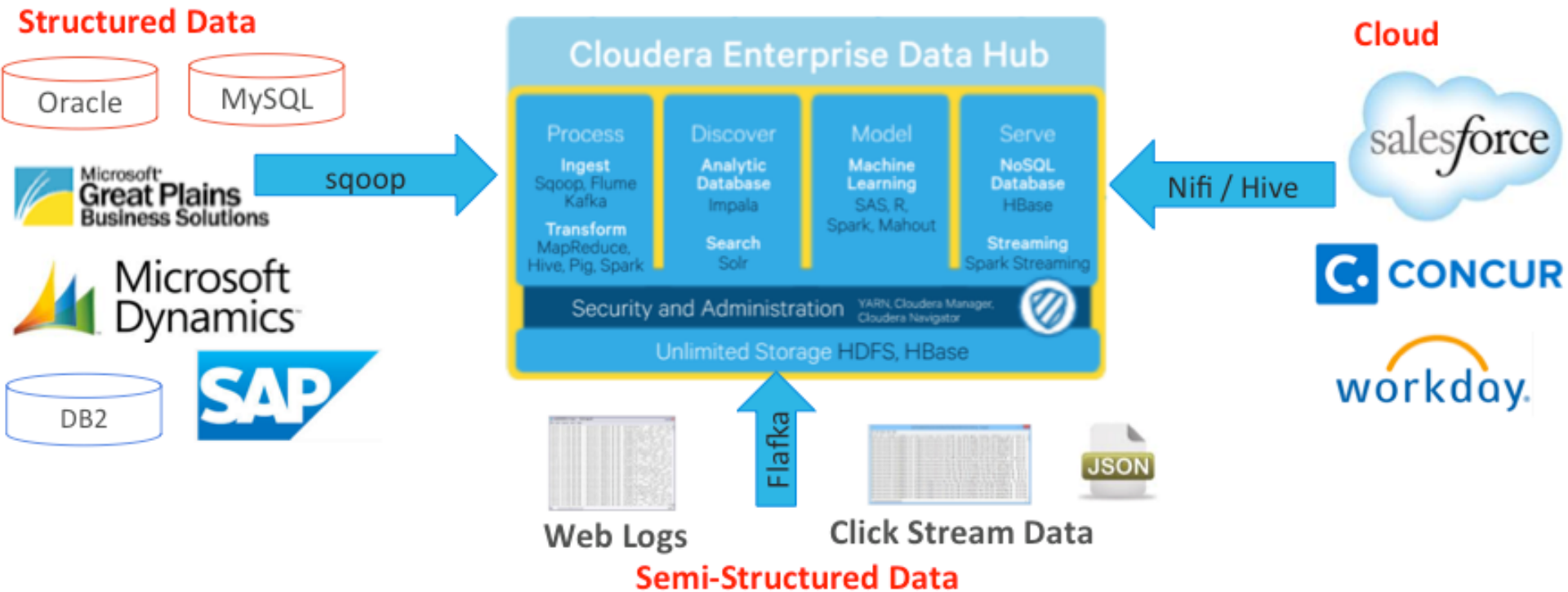
**Affordable & Attainable**

**\$300-\$1,000 per TB**

# Data Warehouse Optimization – cost reduction



# Analytics – BI and Predictive on all data



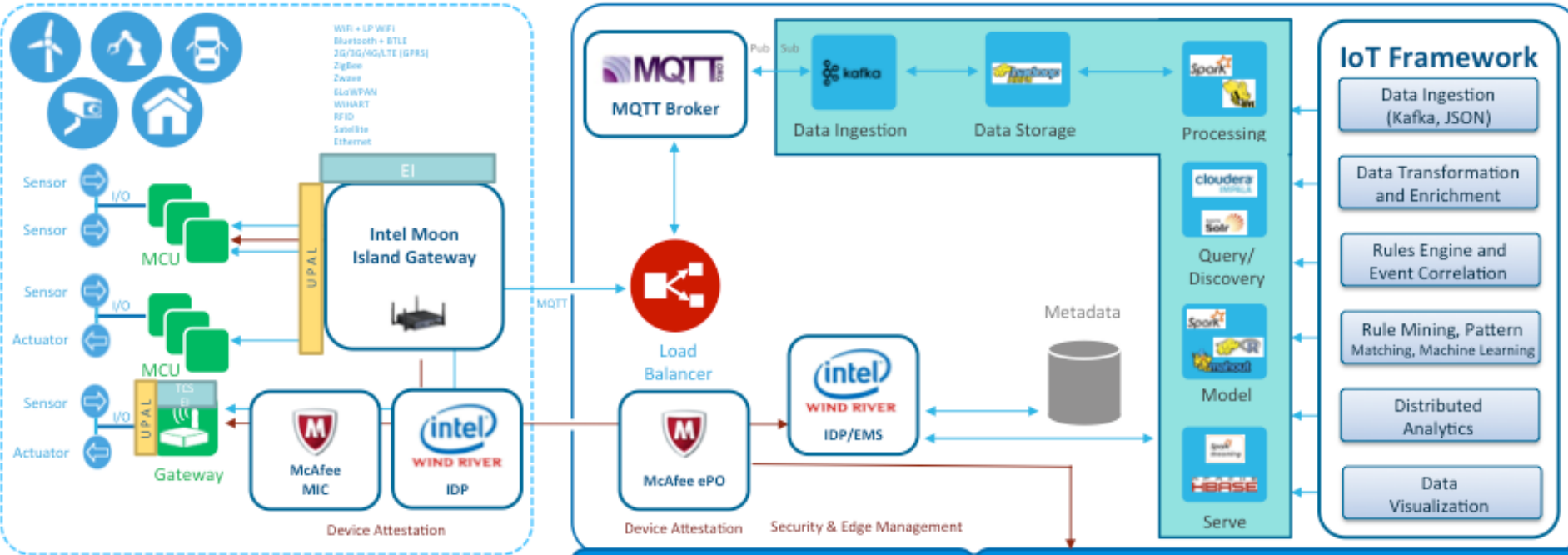
# Internet of Things / IoT – real-time use case



Cloud or On Premise Deployment



cloudera



MCU & Gateway: Identity Protection + Secure Boot  
 UPAL = Protocol Abstraction Layer  
 Rev 2.1  
 EI = Edge Intelligence  
 = Control Flow

intel Intel Discovery Peak Cloudera Manager and Navigator

Management and Governance (Monitoring, Scaling, Logging, Events, Lineage, Audit)

# Enterprise Use Cases

## Financial Services

- Customer 360
- Fraud
- Compliance
- Risk
- Operational Data Store

## Telco

- Customer 360
- Customer Insights
- Network Optimization
- Data Monetization
- EDW Augmentation

## Manufacturing

- Connected: Car, Plane, Equipment
- Agile Supply Chain
- Predictive Maintenance
- IoT Data enabled "Smart Services"

## Retail

- Ship to Store
- Agile Supply Chain
- Next Best Offer
- Connected Store
- Operational Data Store
- IoT – Stores

## Government

- Border Control
- Risk / Intelligence
- 360 Tax payer
- Tax Optimization
- Cyber Threat
- Fraud prevention

# Cloudera Enablement

Start

Train

Certify

## QuickStart VM

- [Quickstart VM >>](#)

## Partner Connect

- [Trainings >>](#)
- [Learning Paths >>](#)
- [Sales & Marketing >>](#)
- [Free Webinar Library >>](#)
- [Why Cloudera >>](#)
- [CDH 5.5 New Features >>](#)

## Documents

- [Cloudera Docs >>](#)
- [Install Guide >>](#)
- [Reference Architecture >](#)

## Courses

- [Administrator – 4 days](#)
- [Spark & Hadoop – 4 days](#)
- [Data Analyst – 4 days](#)

## Mastery / Elective

- [Big Data Apps – 4 days](#)
- [Data Science – 3 days](#)
- [Cloudera Search – 3 days](#)
- [HBase – 3 days](#)

## Cloudera University

- [Resource Library >>](#)
- [Schedule >>](#)

## Certification

- [Administrator – CCAH](#)
- [Data Scientist – CCP](#)
- [Data Engineer – CCP](#)
- [Spark and Hadoop – CCA](#)

## Services Silver Bootcamp

- [Attend Bootcamp >>](#)
- [Study Materials >>](#)

## Demonstrations

- [Cloudera Manager Demo](#)
- [Cloudera Navigator Demo](#)