



Hadoop framework: Understanding its components and how is it supporting Big Data

Leonardo Deana

Paul Kim

May 20, 2015

Imagination at work

Final Report Information

Team email

- Leonardo.deana@ge.com
- Hyeunkyu.kim@ge.com

GitHub account

- 465-leodeana
- Directory: Project_Report
- link: https://github.com/futuresystems/465-leodeana/tree/master/Project_Report

Artifacts

Survey paper, Project Presentation and Project Proposa which will be posted in the GitHub repository informed.



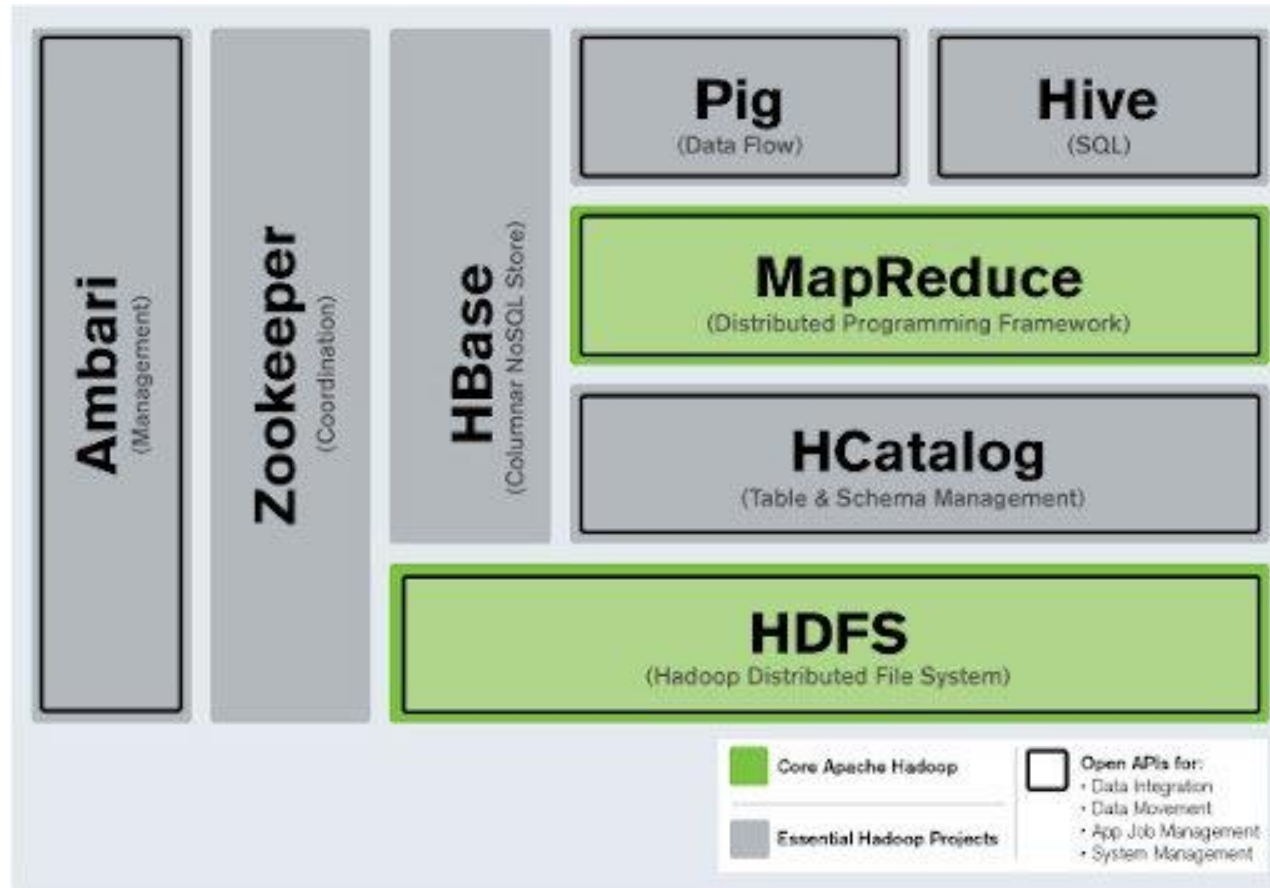
Final Report Description

We will do a survey paper on main concepts of the Apache Hadoop framework, explaining its:

- Components
- Map Reduce, HDFS, Zookeeper, etc.
- Advantages and disadvantages compared to other tools
- Best applications
- How it's supporting big data in large companies
- New trends on this technology



Apache Hadoop Subprojects



Report objectives

We aim to have an deep understanding of apache Hadoop key functionalities and its support in big data for large corporations, including GE if no proprietary issues.

Proprietary Issues:

The reports aims to use public data and other papers/articles as references which will be proper indicated at the references page.

Also, we will only use GE public information regarding this subject.



