



Hadoop Ecosystem



oozie
(Work flow)

HCatalog

Table & schema
Management



Pig
(Scripting)



Hive
(Sql Query)



(Machine
Learning)



Drill
(Interactive
Analysis)



AVRO
(JSON)

Thrift

(Cross
Language
Service)

APACHE
HBASE

HBASE
(Columnar
Store)



Sqoop
(Data Collection)



Zookeeper
(Coordination)



Ambari



Mapreduce
(Data Processing)



FLUME
Flume
(Data Collection)

Apache Ambari
(Management
& Monitoring)



HDFS
(Hadoop Distributed File system)



Yarn
(Cluster Resource Management)

Spark Framework

