

all these tools work collectively to provide such as absorption, analysis, storage, & transfer of data, etc.

Following are the components that collectively form a Hadoop ecosystem.

HDFS : Hadoop distributed file system

YARN : Yet Another Resource Navigator

MapReduce : Programming based Data processing

Spark : In-memory data processing

Pig, HIVE : Query based processing of data sources

H-Base : NoSQL database

Mahout, Spark MLlib : Machine Learning algorithm libraries

Elasticsearch, Lucene : Searching & Indexing

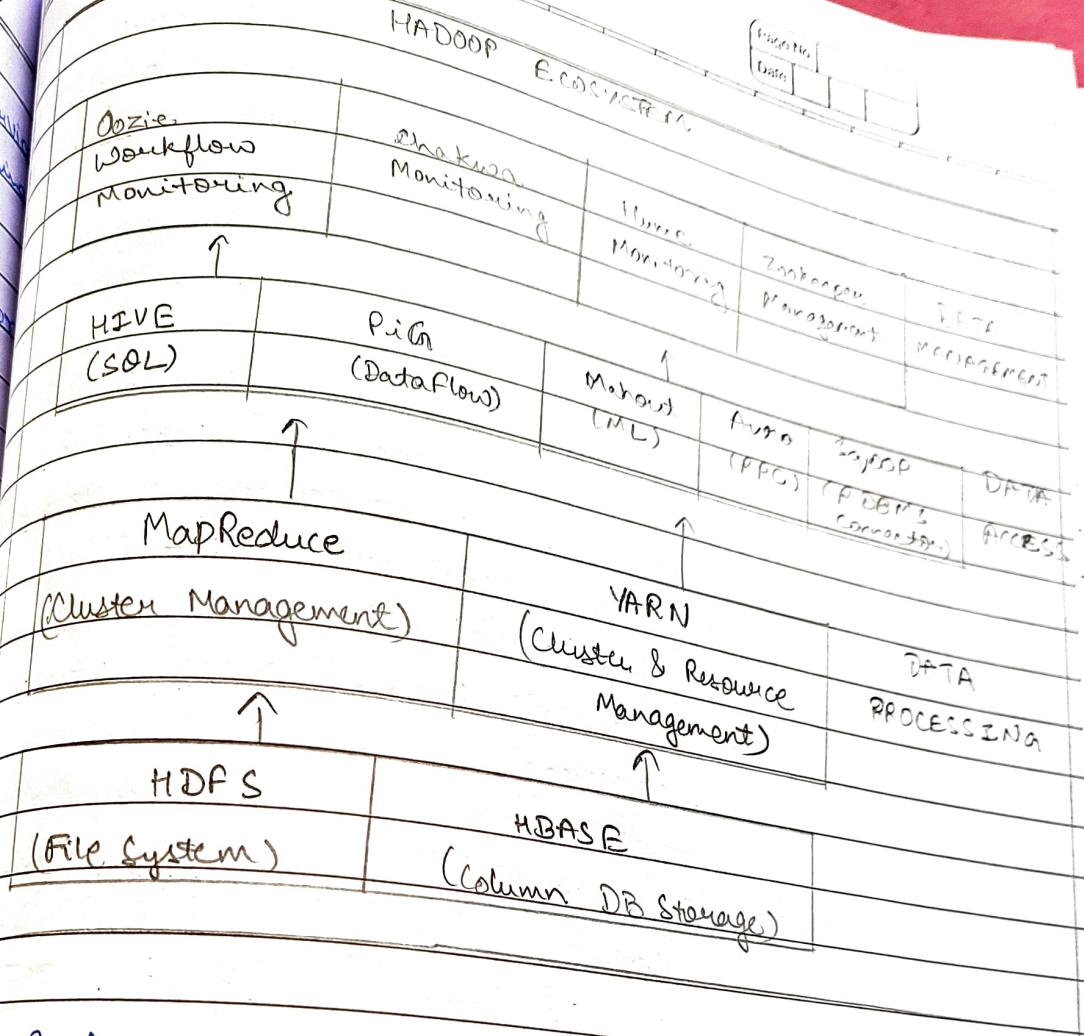
Zookeeper : Managing Cluster

Oozie : Job scheduling

Apart from the above mentioned components, there are many other components too that are part of the Hadoop ecosystem.

All these toolkits or components revolve around a term i.e., data that's the beauty & hence making its synthesis easier.

vely to provide storage & maintainance that collectively form system Navigator Data processing. ing of data serving learning algorithm dexing



components, too that are by making use of distributed & Hcl algorithms, MapReduce makes it possible to carry over the processing's logic & helps to write applications which transform big data ~~sets~~ into manageable ones.

Conclusion - Successfully studied about Hadoop Ecosystem & MapReduce.