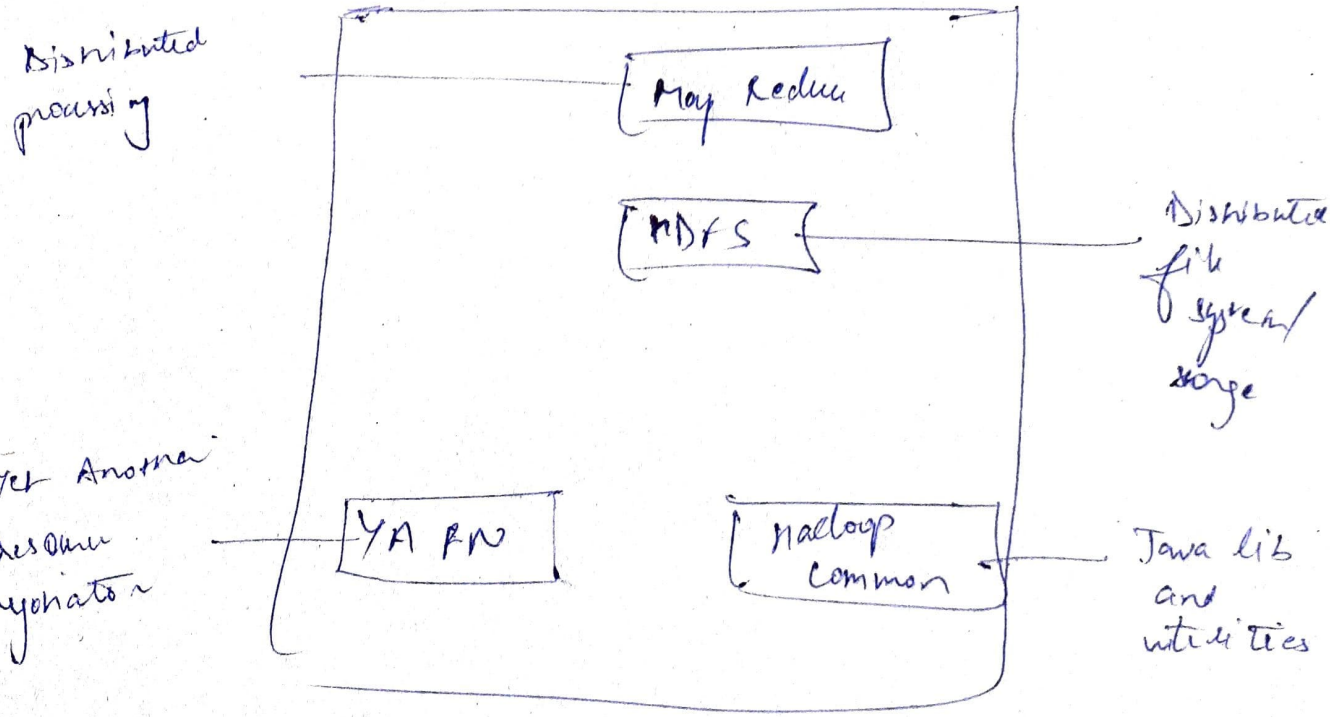
Rahul Yadav / Sec - F / 1918588

Q → Explain Hadoop Architecture

⇒ The Hadoop architecture mainly consists of 4 components.

1) Map Reduce
2) HDFS
3) YARN
4) Common Utilities or Hadoop Common



1) Map Reduce

It is like an algorithm or data structure that is based on the YARN framework. The major feature of MapReduce is to perform the distributed processing in parallel in Hadoop cluster which makes Hadoop working so fast.

2) HDFS

HDFS (Hadoop distributed file system) is a utilized for storage permission is a hadoop cluster. It mainly designed for working on commodity Hardware devices.

HDFS in Hadoop provides fault-tolerance and high availability to the storage layer and the other devices present in that Hadoop cluster. Data storage Nodes in HDFS.

- NameNode (Master)
- DataNode (Slave)

3) YARN (Yet Another Resource Negotiator)

YARN is a framework on which MapReduce works. YARN performs 2 operations that are job scheduling and Resource Management. The purpose of job schedule is to divide

a big task into small jobs so that each job can be assigned to various slaves in a Hadoop cluster and processing can be maximized.

Features of YARN -

→ Multi-Tenancy
→ Scalability
→ Cluster utilization
→ Compatibility

4) Hadoop Common or common utilities.

Hadoop common or common utilities are nothing but our 'java library and 'java files or we can say 'java scripts that we need for all the other components present in a Hadoop cluster. These utilities are used by HDFS, YARN, and MapReduce for running the cluster.