

- > Hadoop founded by GFS
- > Since 2005, it's been evolving
- > It is open source framework. It is an interference b/w DB & HDFS
- > Sqoop → open source framework used for efficient data transfer.
  - import → RDBMS → HDFS
  - export → HDFS → RDBMS
  - use: analyzing the data
- > Tools & technology: Hive, Pig, Spark, Map Reducer

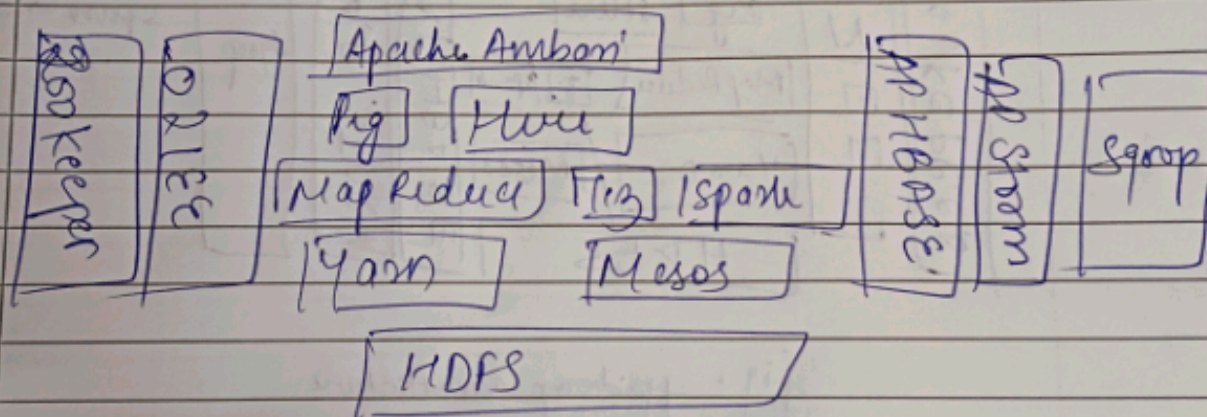


fig Hadoop architecture.

Yarn → Yet another resource Negotiator  
 func. → Tell from which resource we should get data  
 Same as mesos

Map reducers → Alternative tez  
 (can make graphical repers)

\_/\_/\_

Spark → script, lang

Apache HBase → storage work usage

Apache Storm → real time processing usage

Ozoo → scheduling purpose

Zookeepers → state of machine  
up → on  
down → off.

Apache Ambari → Access of everything