

Big Data

- > Hadoop was found by GFS (Google file system)
- > Symbol - yellow elephant
- > Since, 2005 it's been evolving.
- > Hadoop is an open source framework. It is a interface or interlinking between database and HDFS
- > Apache scoop - an open source framework used for efficient bulk data transfer.

Map Reducer → parallel processing

Sqoop

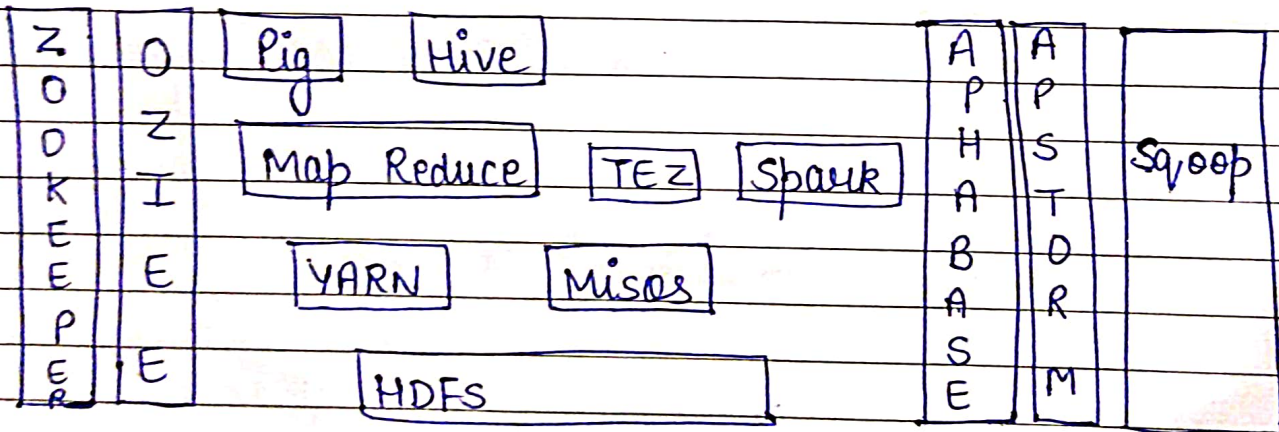
Import → RDBMS → HDFS

Export → HDFS → RDBMS

Use → Analysing the data

Tools and Technology → Hive, Pig, Spark, Map Reducer

Apache Ambari



- > Tez → Graphical representation
- > Apache → Storage
- > YARN → Yet another resource Negotiator
- > Apache storm → Real time operators
(we can update running programs)

> Zookeeper → State maintain [Up → on
Down → off]

> Apache Ambari → we can enter the field we want to access.