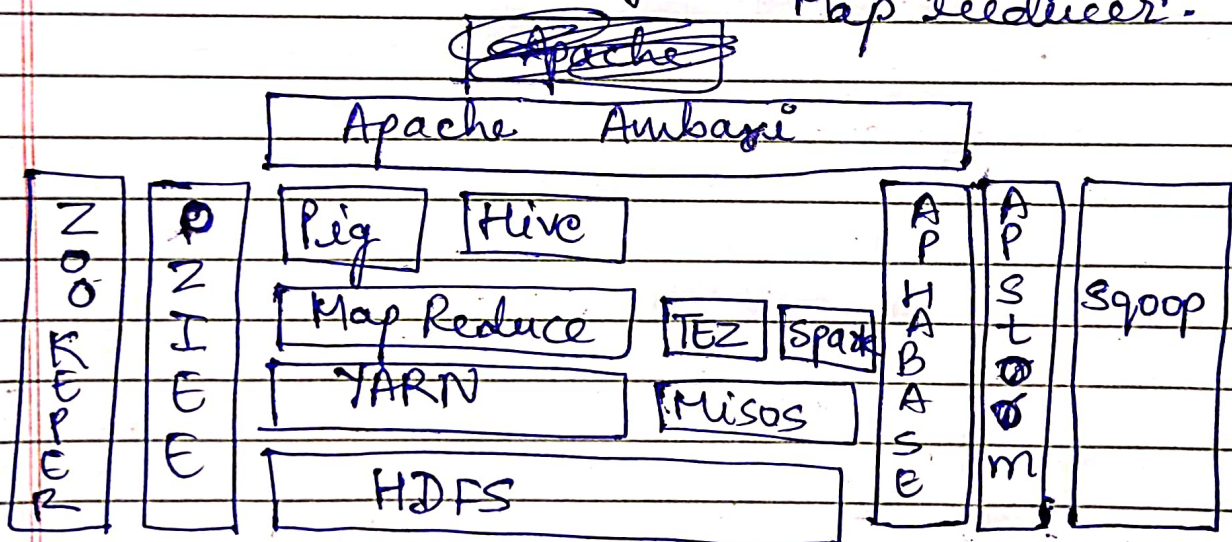


- > Hadoop was found by GFS. (Google file system)
- > Symbol — yellow elephant.
- > Since, 2005 it's been evolving.
- > Hadoop is an open source framework. It is a interface or interlinking between database and HDFS.
- > Apache Scoop → an open source framework used for efficient bulk data transfer. 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

Export → HDFS → RDBMS. 4PB-  
Use → analysing the data.

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



⇒ Ek node se jo different systems ~~also~~ connected h us system me diff. data kisse bhi system se access kiya ja sakta h → Distributed file system.

⇒ Tez → Graphical representation



Apache → storage ka kaam.

YARN → yet another resource Negotiator.

Apache storm → Real time operations.

(Chalte programme me updation  
cho to to vo real time me kaam  
karta h)

zookeeper → state maintain karne ka kaam h  
(up state me h, ya down state me h)  
up → on.  
down → off.

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