

## IAT-1

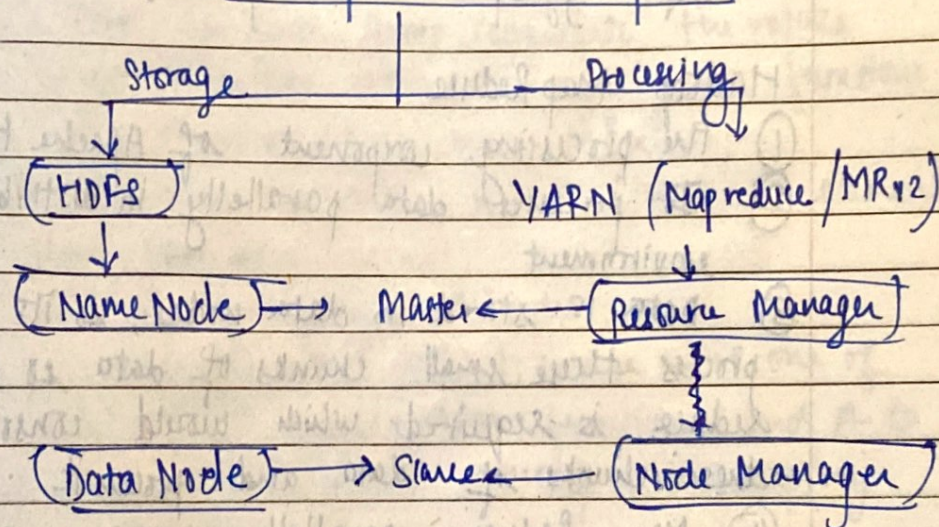
Q2]

A] Storage and processing component in Hadoop.

Hadoop consists of two main components i.e.

- i) Storage — Hadoop HDFS
- ii) Processing — Hadoop Map Reduce

B]

Hadoop 2.x Core componentsHadoop HDFS

- ① Each file is stored in HDFS in the form of Blocks
- ② The default size of block is 128 MB in Hadoop 2.x
- ③ Previous version had a default size of 64 MB
- ④ Name Node is the master daemon, it maintains and manages datanodes
- ⑤ Records metadata eg:- location of blocks, size of files, permissions, hierarchy etc
- ⑥ Receives heartbeat and block report from all the DataNodes.
- ⑦ The data node is the slave of the daemons, it stores the actual data and serves read and



write requests from the clients.

- ⑧ Hadoop has a lot of features like scalability, Reliability, storage of replicas of the block, portable file system
- ⑨ HDFS stores large files typically in ranges of gigabytes to terabytes.

### Hadoop Map Reduce

- ① The processing component of Apache Hadoop
- ② It processes data parallelly in distributed environment
- ③ Data is stored in data nodes, so its difficult to process these small chunks of data so a map reduce is required which would consolidate these chunks of data and process it.
- ④ Map Reduce parallelly processes so the processing is faster
- ⑤ Moving data to processing is very costly in map reduce, we move processing to the data.

