- I] decide 1) Job Priority 2) Job dependency
  with one Job 3) which
  - " can process the Job.
  - (ii) what is time require the1
  - Job execution these decide by
    Job Scheduling.

A resource management :-
  what resource use require to perform to
  execution on data is handle by resource
  management.
  - which resource, how many resource, which
  types resource allocate for finish task,
  it decide by resource management.

3) HDFS
  4) map reduce

3) Hadoop HDFS :-
  - HDFS is the primary or major component
    of Hadoop.
  - Is responsible for storing - it the data set
    of structured & unstructured data across
    various nodes

4) map reduce :-
  - Is a method for distributing computation
    across multiple nodes.
  - Each node processes the data that is stored
    at that node.
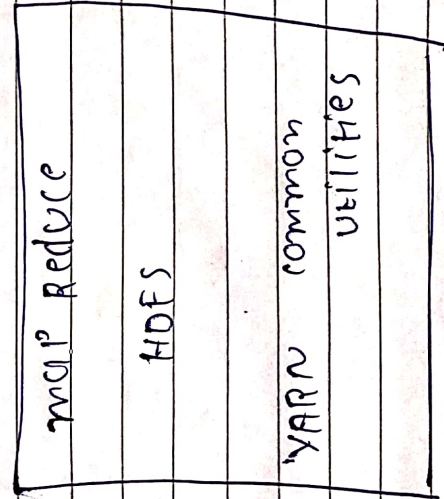 - two main phase :- map
    : Reduce.

Hadoop

17011016036

* Feature of Hadoop

1) Hadoop is open source
2) Hadoop cluster is high scalable
3) Fault tolerance
4) High Availability
5) Data Processing
6) Data locality
7) Feasibility
8) easy to use
9) Data Reliability

Hadoop Core component

→ MapReduce
→ HDFS
→ YARN
→ Common utilities

MapReduce
HDFS
YARN   Common utilities

Hadoop have main 4 components

1) HDFS 2) map Reduce 3) YARN 4)

1) Common utilities:-
- It also call Hadoop common
- It is Java files / Libraries which
  required for other module (YARN, to
  HDFS) work.
- to start and maintain performance o
  so Java files & scripts are used wh
- Provide by common utilities.

2) YARN
- YARN stend for yet another resource
- It work 2 main thing 1) Job sched
  2) Resource m
- Job scheduling:-
- large data as input, it devides i