Mangal
Date: / /201

1. Hadoop component
2. Hadoop common
3. HDFS
4. MapReduce
5. Hadoop TOOLS
6. Other open source TOOLS.

HDFS (Hadoop Distributed File system)

- primary storage system for Hadoop.
- Distributed, portable, scalable file system written in java.
- Files are broken down & stored in multiple machine.
- Designed for large scale distributed data processing.
- It follows Master & Slave architecture.

HDFS consists of

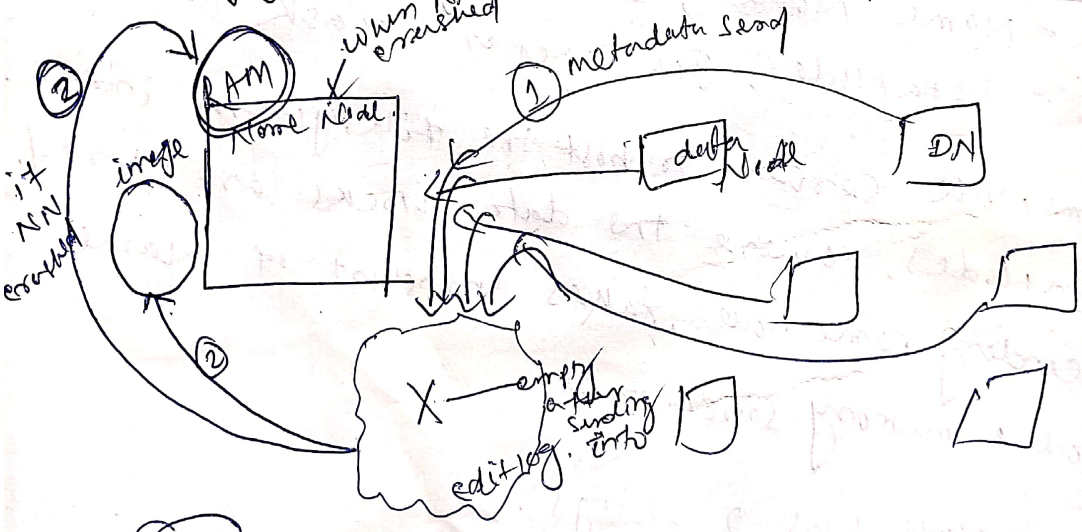
- Name Node, Secondary Name Node
- DataNode, Job Tracker, Task Tracker.

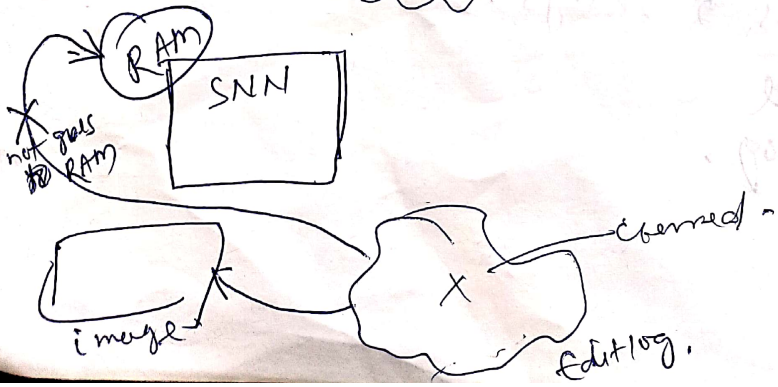
- NameNode server to host the filesystem index.
- DataNodes, where the data blocks are stored.
- secondary name node, takes snapshot of Name nodes memory structure.
- NameNode has 2 things.

1 - Image
2 - Edit log.

Mangai
Date: / /201

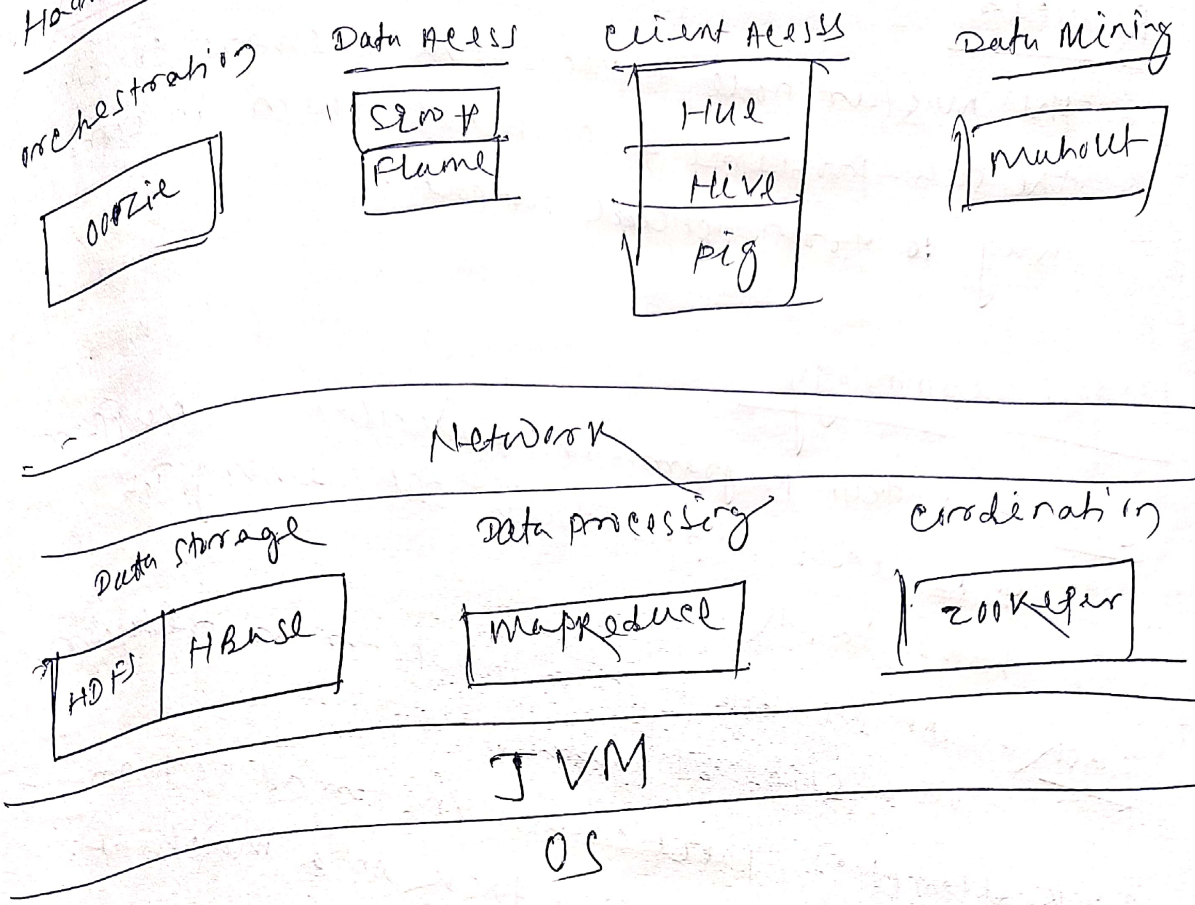
- Whenever we write file to the machine, the slave node sending their intron. Then who will receive.
- Editlog comes to picture. Editlog have all the infor?.
- Let's say Namenode crashed, when coming back the editlog send the metadata to the RAM.
- When we are Restarting, the editlog goes to RAM as well as image.
- Then the editlog has cleared.
- When all the salve node sending metadata to the Namenode, it is commit log of secondary namenode.
- But for SNN editlog can't go to RAM. it only goes to image.





Hadoop components

Mangal
Date: / /201



MapReduce

- A SW framework design for processing large volume of data.
- Divides data into set of implementation tasks to run in parallel machine.
- It will operates in 2 steps.
  - Map step
  - Reduce step.

Mapper

- Master node takes, input then partitions it into sub-problems & distributes it to worker nodes.
- Map func takes a series of key/value pairs, processes & generates 0 or more output key/value pairs.

Mangal
Date: / /201

Reducer

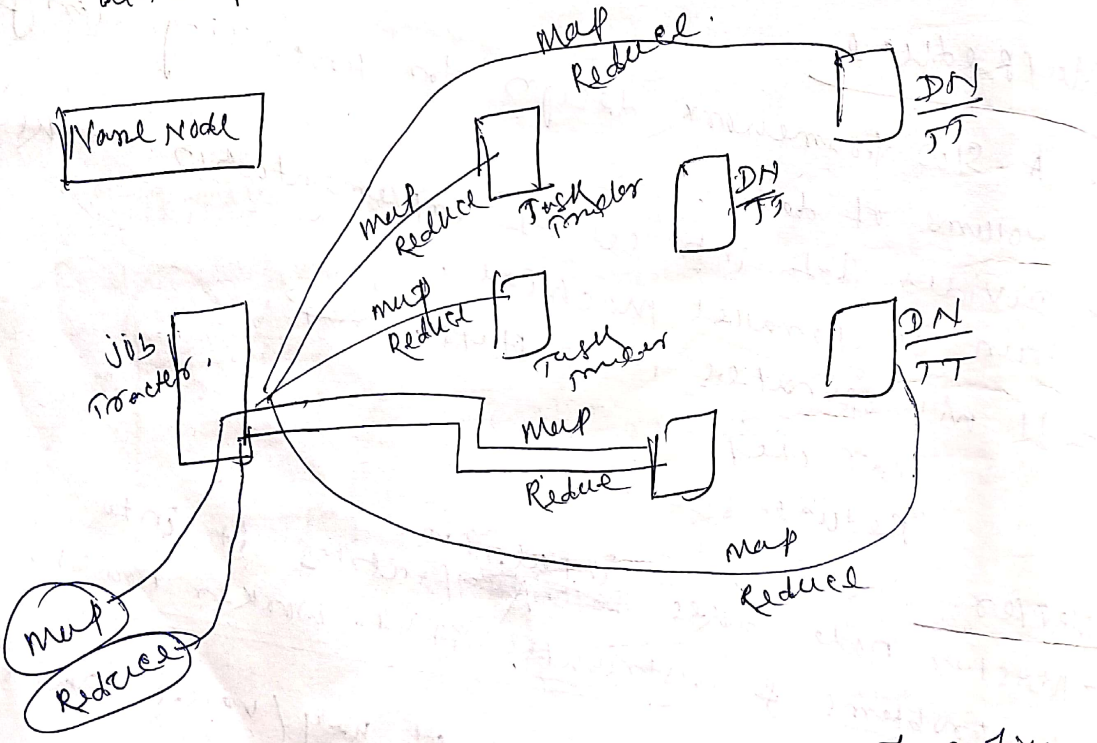
- The Master node collects the answers to all the sub-problems & combines them in some way to form actual output.

Hadoop Terminology

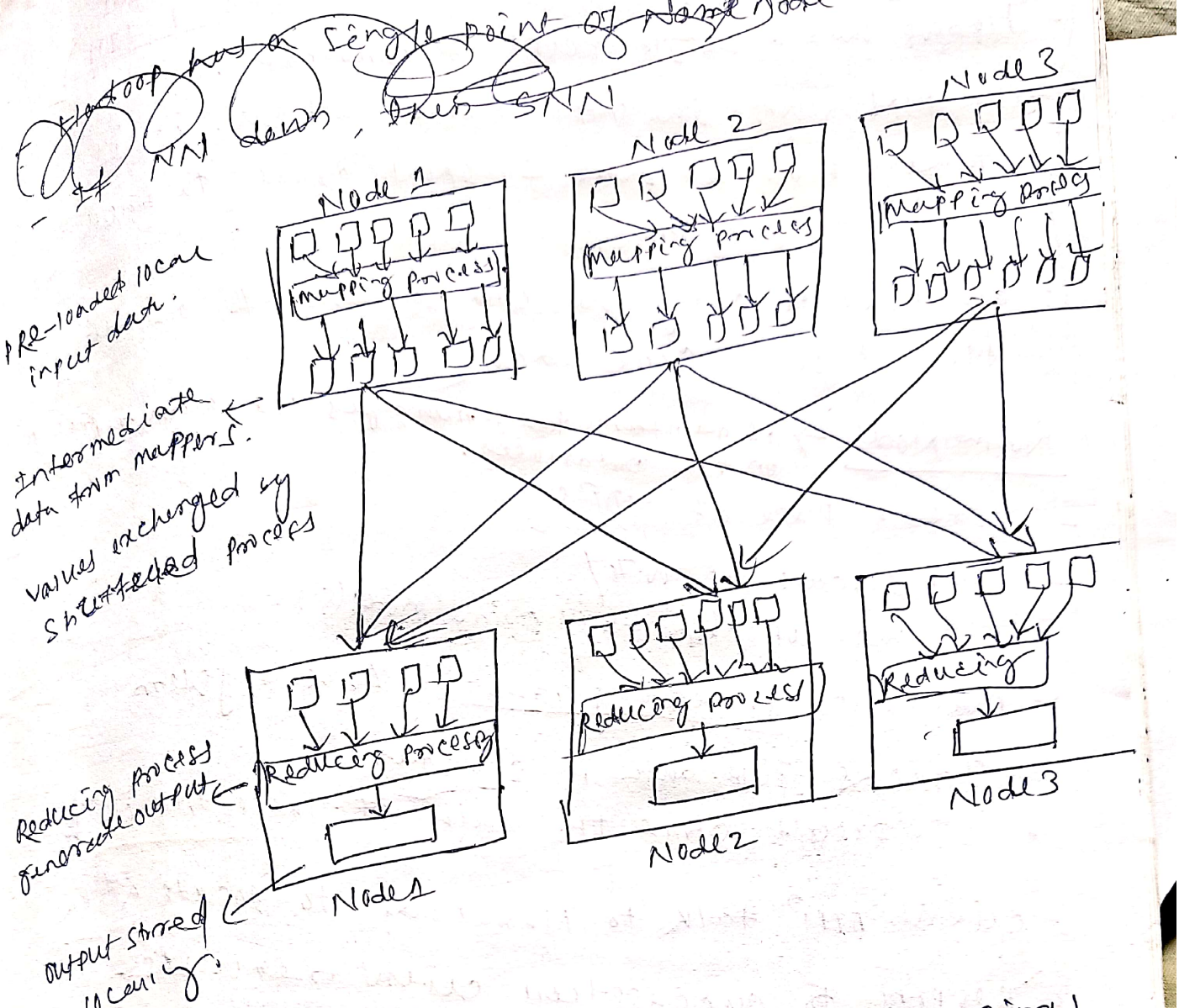
Job:- A full program - an execution of mapper & Reducer across data set. every Job have task.

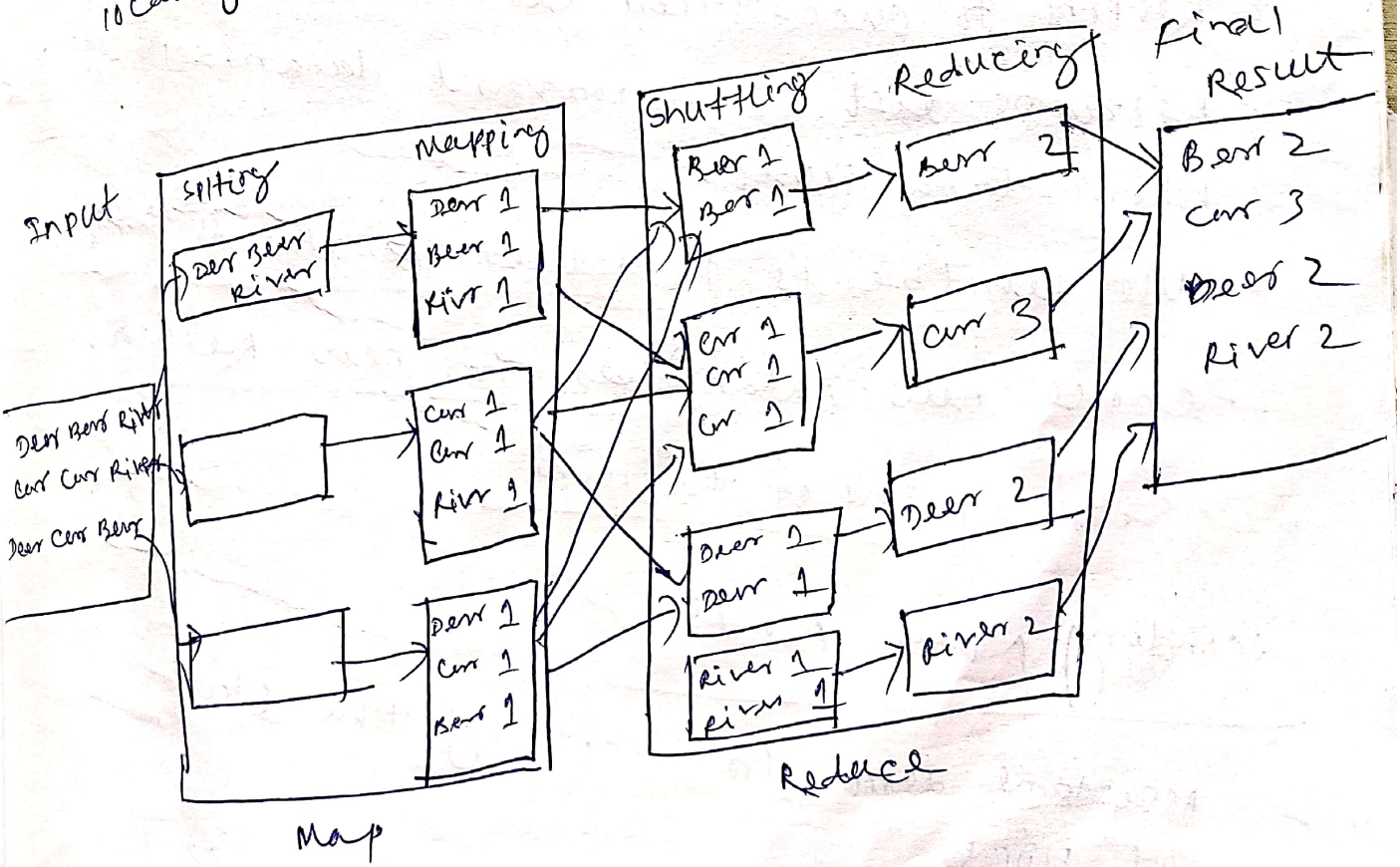
Task:- An execution of a mapper or reducer on a slice of data.

Task attempt:- particular instance of an attempt to execute a task on a machine.



- Job Tracker will take the input of a file & talk to task Tracker.
- Map & Reduce go to the every machine with the help of Job Tracker.


Node 1
Node 2
Node 3
Mapping process
Mapping Process
Mapping Process
Reducing process
Reducing Process
Reducing
Node1
Node2
Node3
Pre-loaded local input data.
Intermediate data from mappers.
Values exchanged by shuffled process
Reducing process generate output
output stored locally.


Input
Splitting
Mapping
Shuffling
Reducing
Final Result
Deer Bear River
Car Car River
Deer Car Bear
Bear 1
Bear 1
Car 1
Car 1
River 1
Deer 1
Deer 1
River 1
River 1
Bear 2
Car 3
Deer 2
River 2
Bear 2
Car 3
Deer 2
River 2
Map
Reduce

- Hadoop has a single point of name node.
- If NN. down, the SNN come back. it is taking hour of time & SNN comes to back is a manual process.
- We have to copy the image file & edit log then it will come back.

NameNode → It maintains & manages the block regions on the DataNodes.

- Center piece of HDFS
- http://localhost:50070/
  - Web UI for HDFS name node(s)
- Maintains directory tree in file system
- Tracks file data in cluster.
  - Doesn't store the data of these files itself.
- client App? talk to Namenode to locate a file.
- Respons to successful client requests.
- Returns list of relevant datanodes where data lives.
- Use heartbeats to detect Datanode failure.
- Choose new Datanodes for new replicas.
- Single point of failure.

Secondary Name Node

- Assistant demon for monitoring the state of HDFS

Mangal
Date: / /201

- Each cluster has one SSN residing on it own machine.
- No other datanode or TaskTracker runs in SNN.
- Doesn't receive any real-time changes to HDFS.
- It is a backup for NameNode.
- Take snapshot of metadata at intervals defined by the cluster configuration.
- Snapshots minimize the downtime & loss of data.

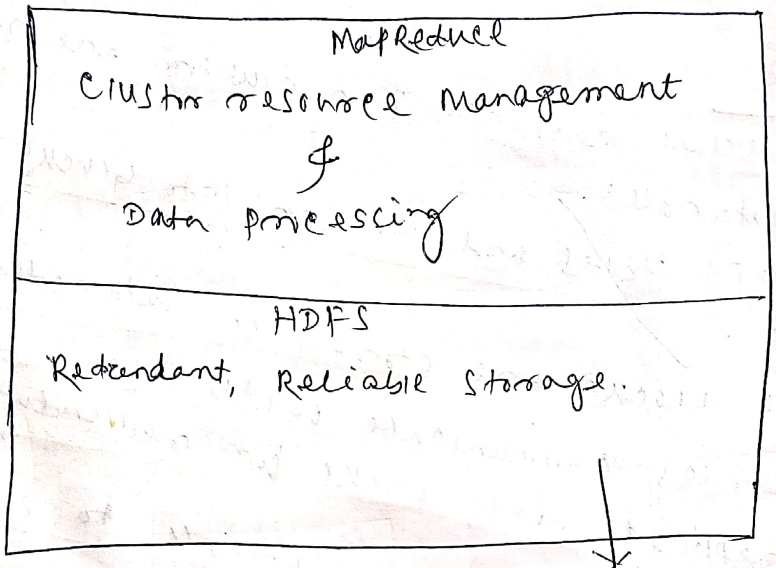
DataNode

- Individual machines in a cluster are referred to as Datanodes.
- In HDFS files are broken into blocks of fixed size.
- These blocks are stored in datanodes.
- Datanodes communicate with other Datanodes to replicate data blocks for redundancy.
- Datanodes connect to Namenode to establish service.
- Responds to requests from the Namenode.
- Datanodes send heartbeat to the Namenode.

Mangal
Date: / /201

Problem in Hadoop (1.0)

- Hadoop has a single point of Namenode.
- If Namenode down, then SNN come back. It is taking hours of time & SNN come to back is a manual process.
- We have to copy the image file & editlog then it will come back.



Which is responsible for storing the data.

- But on top of the HDFS only 1 technique is there, i.e, Mapreduce.
- Mapreduce is done using the help of JobTracker.
- When we submit a Job, then JobTracker will do the Cluster Resource Management.

Which machine have how many CPU & how many CPUs give to which Jobs.

Mangal
Date: / /201

- MapReduce doesn't have capacity of 1000 jobs, more CPU, more RAM, more VM cases to run Job Tracker can't take more than that.
- Job Tracker bottleneck
- That's why Hadoop(2.0) came to picture.
- single job tracker receive all the jobs as parallel.

~~Apache Hadoop Yarn (Hadoop 2)~~

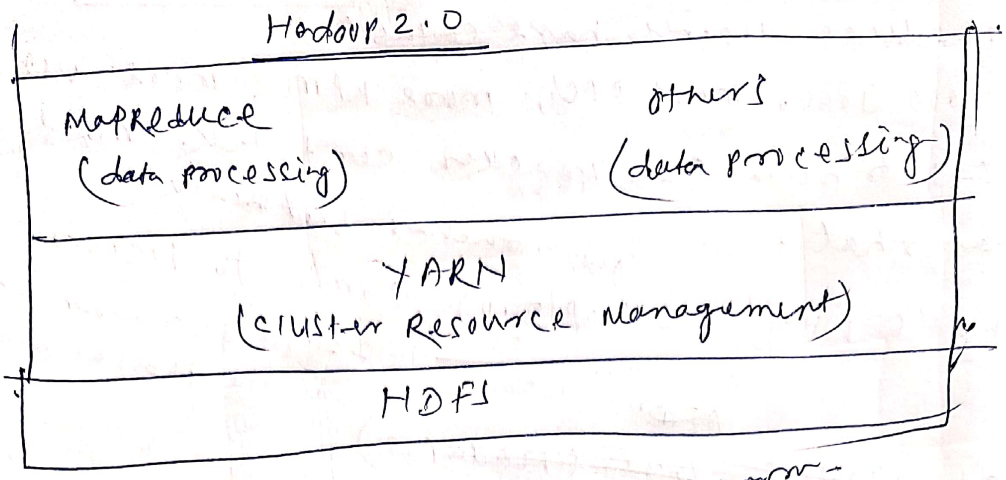
- problem with large scaling.
  - > 4000 nodes
  - > 40K concurrent tasks.
- problem with Resource utilization.
- slots only for Map or Reduce
- single NameNode, single point of failure.
- clients & cluster must be at same version.

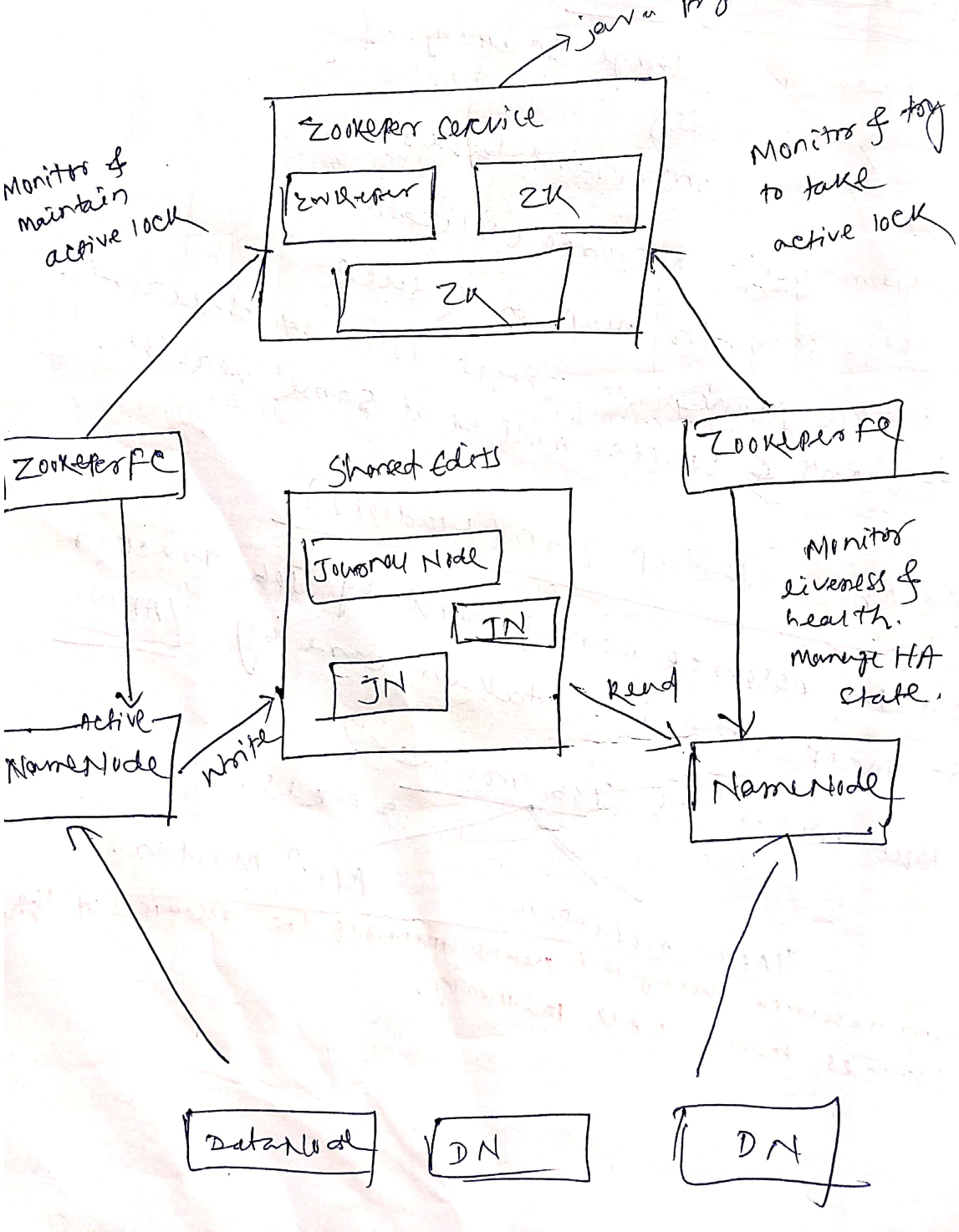
Apache Hadoop Yarn (Hadoop2)

- cluster Resource management by Job Tracker in MapReduce i.e, taken care by YARN.
- But where is Job Tracker?
  - 2 pieces.
    - YARN architecture
    - Appl' Master.

cluster resource management means managing the resources of the cluster.
- resources means CPU, memory.

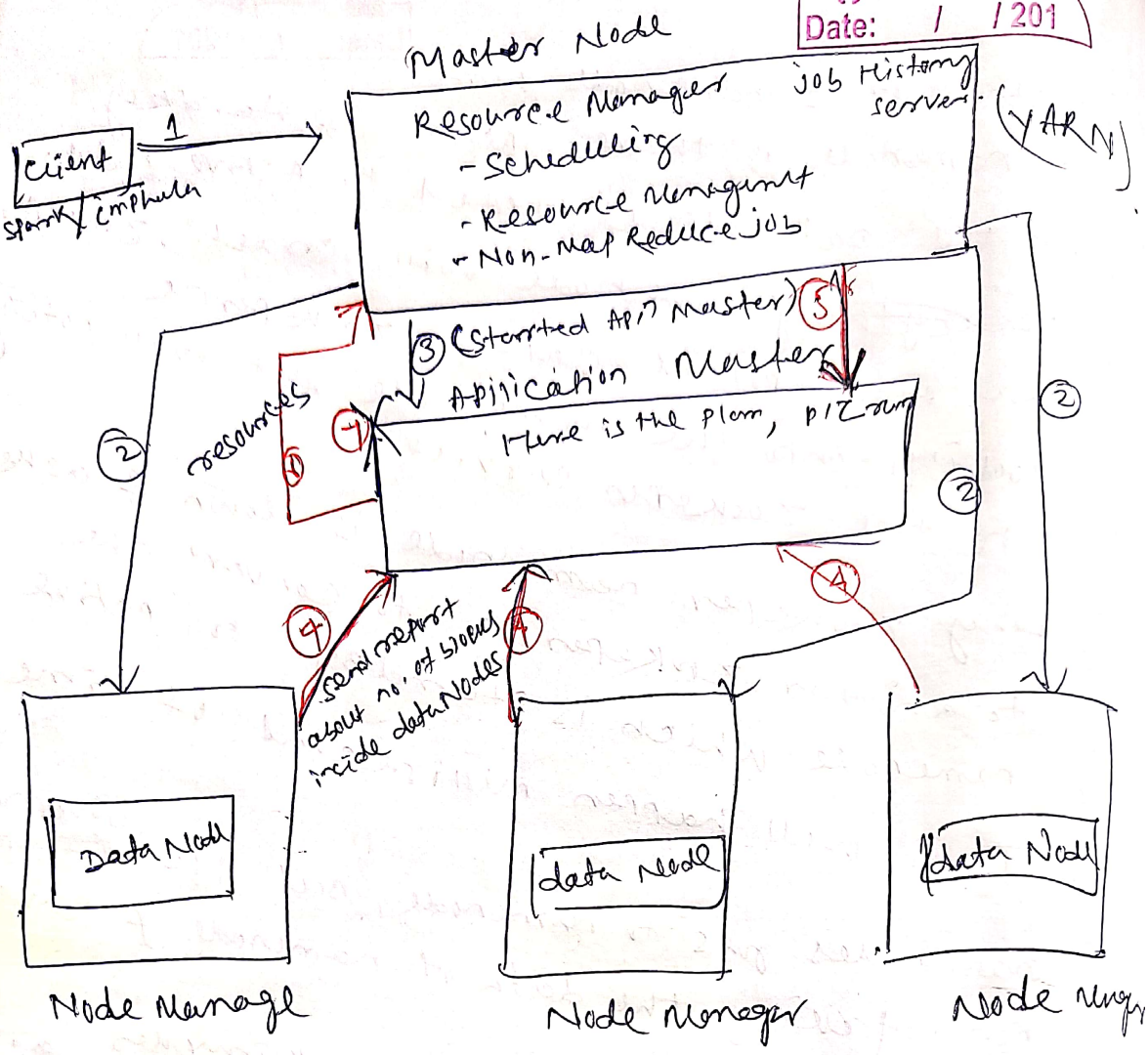
Mangal
Date: / /201





Mangal
Date: / /201

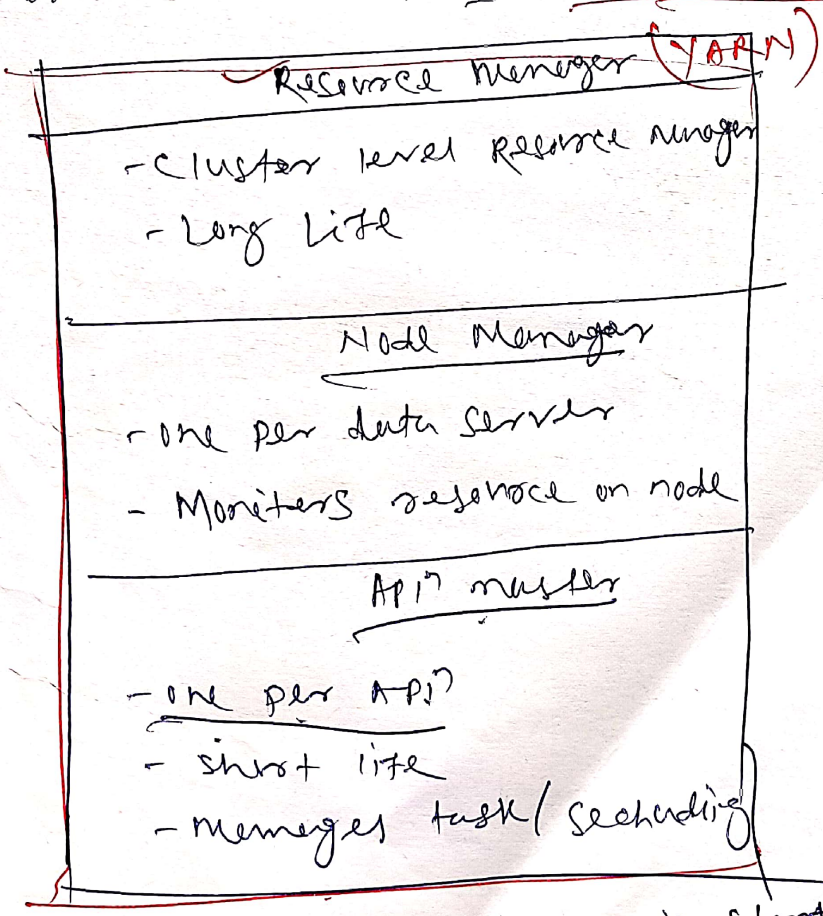
- Instead of 1 namenode here 2 namenode is there. Active & standby.
- All the datanode connected to Active node.
- If ~~anything~~ namenode will crash, Zookeper has a component called Zookeeper looking into the namenode, it will notify to the Zookeper service.
- Hey Zookeper namenode is down & asked to another Zookeeper to convert the namenode which is standby to Active.
- This will happen millisecond of time.
- Data nodes goes to namenode, but the Journal nodes taking the data of namenode & converting it to standby.
- Journal node, Zookeper, ~~Datanode~~ TaskTrackers are service only.
- Zookeper is a java program, it runs 3 machine.
- Journal node taking the commit logs from Active Namenode & writing to standby namenode & make as shink.
- Task Tracker replace by Node Manager.
- improves scaling.
= 11 resource Management



- client can send different type of apps. Impala, Spark ...
- non-mapreduce job also handle by Resource Manager.
- Resource manager take the data from different type of apps. - Q
- Resource Manager know how many resources are their.
- Resource Mnager talk to Node Managers & then started the app master.

Mangal
Date: / /201

- API master will do appl monitoring & job scheduling.
- Node manager will give the report about the no. of blocks available in the data node.
- Then Resource Manager tells to api master hey, api master here is the plan, plz run it. (process context)
- After processing the api master returns the resources to Resource Manager.
- Then api master gone.
- There are more than 1 resource Manager.



- When we submit a job the yarn will start api master
- Then api master will start the certain no. of executors.