① master-slave
② Cluster
③ Nodes.

① Scale-in.
② Scale-out

Bigdata → localSystem → 16GB / 24 GB → 32GB RAM → 64GB RAM
500 GB SSD → 1000 GB SSD → 10000 GB SSD
i5 Processor → i7 Processor → i9 Processor
GPU Processing

Nodes → System

Vertical scaling → Scale-in → Limited

(500 GB RAM)

16 GB Ram. 500 SSD i5 Processor

16 GB RAM 500 SSD i5 Processor

16 GB RAM 500 SSD i5 Processor

16 GB RAM 500 SSD i5 Processor

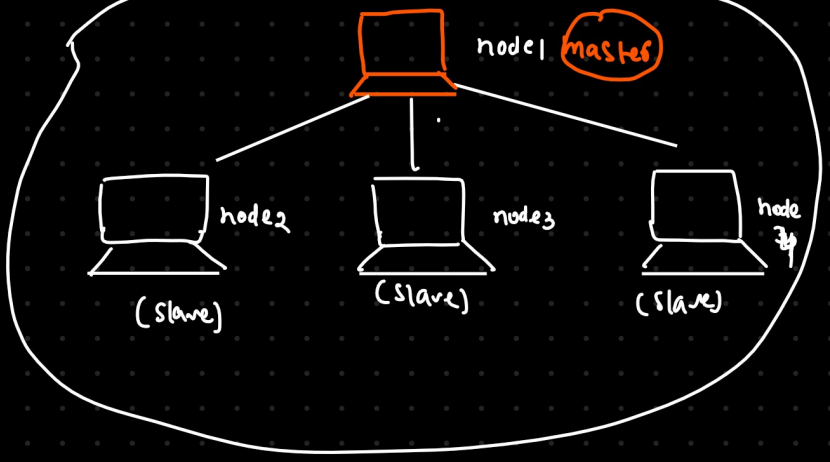
Cassandra

Ring topology

Scale-out

Horizontal Scaling

Hadoop



hadoop demons / hadoop components

Hadoop 1.x.

HDFS / Map reduce.

Name node
data node
{ task tracker
Job tracker }

HDFS → Name node
Data node

Map reduce → Job tracker
task tracker

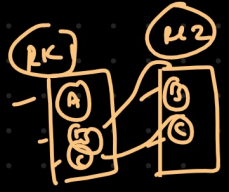
Hadoop 2.x.

HDFS / MRv2. / YARN

Name node
Secondry name node
data node
{ resource manager
node manager }

HDFS → Name node
data node

map reduce → RM
NM



Master - System
1. name node
2. task tracker

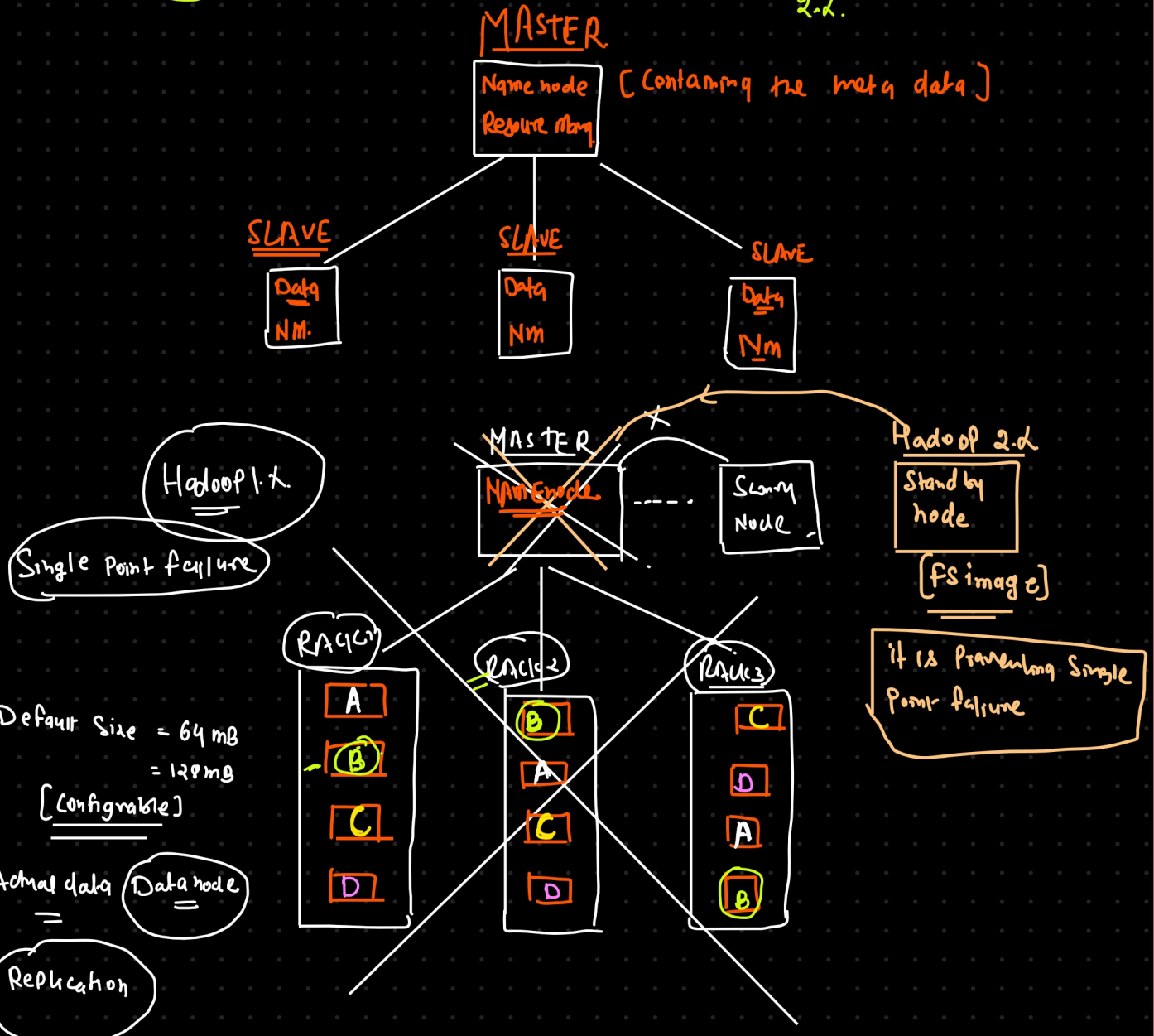
Slave System
1. Data node
2. Job tracker

1.x

1. name node
2. resource manager

1. Data node
2. node manager

2.x.



Default Size = 64 mB
= 128mg
[Configrable]

Actual data = Data node

Replication

by default replication factor → 3

Rack awarness

Hadoop 1.x.

(1) MapRed. cluster mang + batch Proc

Namenode
Metadata

(2) NO stand by node.

(3) Limited to 4000 nodes.

Hadoop 2.x.

MRV2
Processing

Cluster management
YARN

metadata / Name Space
Name node

(2) there is a stand by node.