

Week-2 FAQs

W2:1 Theoretically it is always Map -> Shuffle Sort -> Reduce. So, from the MR code where does actually shuffle and Sort taking place before giving input to Reducer? Is it by default from map output?

Ans: The output of multiple mappers is shuffled to the reduce node. Shuffling is the process of moving the intermediate mappers data to the reducer node. Once the intermediate data are shuffled to the reducer node, Sorting happens based on key before sending the data to the reduce task.



W2:2 Number of mappers = Number of data blocks. Let's say we have default replication factor 3. 500MB size file. We have totally 12 blocks (including a copy of data). How program will know it should take only 4 blocks of actual data?

Ans: As NameNode has all the metadata related to the files stored. So, this will expose only 4 blocks of data.

Ph:9108179578

W2:3 Will the number of mappers running be equal to the blocks or will they be equal to the number of nodes in a cluster? For e.g., if the data we are working on has say 100 blocks in node 1, will there be 1 mapper in node 1 or 100 mappers in node 1?

Ans: Number of mappers is always equal to number of blocks. Now the question comes: are all mappers running in parallel or not? This is determined by the resources we have. Consider you have 100 blocks on a single node, and this node is a 4-core machine. then on this node 4 mappers can run in parallel at a time. Once these 4 are done, the next 4 will start and So, on.

W2:4 Does the combiner and partitioner happen in map nodes or reducer nodes?

Ans: Combiner is always a part of the Mapper phase whereas partition is an intermediate phase that takes place after the map phase and before the Reduce phase. Refer Image in Q1

W2:5 Do Shuffle and Sort happens between every stage like mapper output -> combiner ->partitioner -> reducer?

Ans: Shuffle & Sort done only once. mapper -> combiner ->partitioner -> shuffle & Sort -> reducer. Refer Image in Q1

W2:6 When reducer is 0, then output of the mapper won't be aggregated as I understand. So, if we assume there are 4 mappers, we will get 4 outputs. So, from an end user perspective how will these 4 outputs be combined as output is distributed across 4 different files?

Ans: Using getmerge command. Example given in Week1 FAQ

W2:7 Is it possible to get a combiner only aggregation count with zero reducer for word count problems? I tried to set a custom combiner class and set the num of reduce task 0, the output of map is not aggregated.

Ans: It will not get aggregated because local aggregation will not work on word count cases. What if the word hello is present in two blocks? The output will have two records of hello. Without reducer, the global aggregation will not take place

W2:8 This is valid code. Will there be any performance gain here for using the same class as Combiner and Reducer?

```
job.setMapperClass(MyMapper.class);
job.setCombinerClass(MyReducer.class);
job.setReducerClass(MyReducer.class);
```

Or is it useless to define the same class as Combiner as well as Reducer?

Ans: Yes, there will be performance gain, as the combiner is executed in the map nodes which will provide parallelism. The combiner will perform local aggregation on the map output and hence avoid large amounts of network movement at the Shuffle and Sort Phase. Also, the amount of data to be aggregated in case of using the combiner at the reducer will be less as compared to running without a combiner.

Combiner and Reducer can be same most of the time but in few cases it has to be different.

W2:9 Number of mappers is equal to the number of blocks. What if we have a cluster of 4 machines and there is 1gb file, then the total blocks will be 8. So, can there be multiple mappers running on the same machines or how exactly it will work in this case?

Ans: 8 blocks will be divided into 4 nodes means each node will hold 2 blocks. If each node has 2 cores, then 2 mappers will run simultaneously on each node.

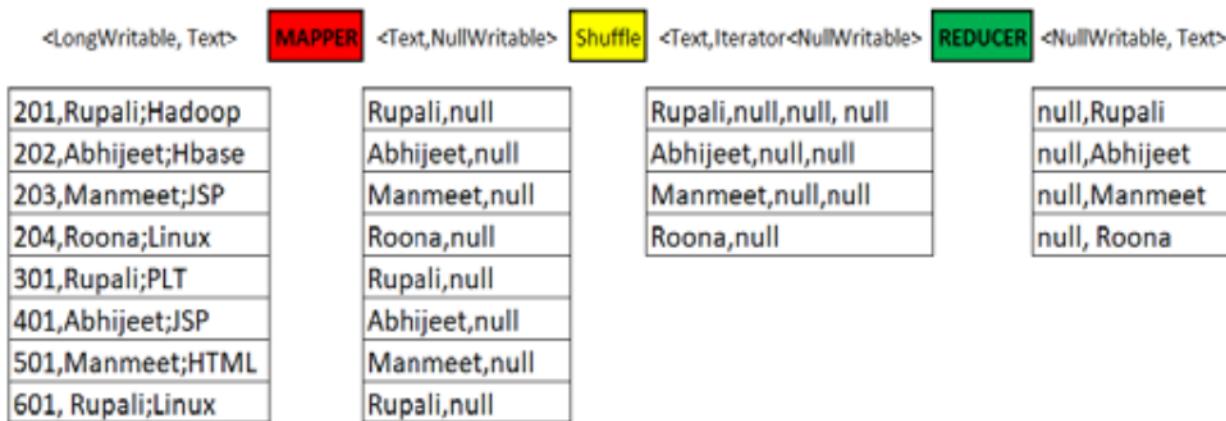
W2:10 Does the Record reader work on Mapper Machine?

Ans: Yes, the record reader works on mapper machines

W2:11 In Finding Distinct value examples we need shuffle and Sort to get the final output that is the reason to have a reducer. But what will we write in the reducer code?

Ans: In the reducer we get a key and a list of values as input right. We will just write the key and we will ignore values means we won't be looping

through values. Below pic will find distinct names



W2:12 Number of Mappers = Number of data blocks. So, if data blocks are 4 and replication factor is 3, then total mappers would be 12 or still 4?

Ans: No of mapper would be 4. As we are replicating the things then running a mapper on the same replicated data would result in the same result. Then why we will run mapper on all replicas. Processing happens only on one copy of data block. Replication is for fault tolerance.

W2:13 In case there is a reducer value set to 0 and we process two input files in the directory as a result we get 2 mapper output. My concern is if two files having size less than 128 MB that can be handled by a single mapper or rather both files are distinct So, no matter file size it will generate a separate mapper output file.

Ans: No of mappers will be equal to the no of blocks in HDFS for a file. So, even if files are small, they will be residing in two different blocks thus will be processed by two mappers and we will get 2-part m files as o/p.

W2:14 Is it possible to utilize the shuffle and Sort part of the map reduce framework without having any reducer logic or setting number of reducers to zero?

Ans: If you set no of reducer to zero then shuffle and Sort will not execute. If you write reducer without any logic then shuffle and Sort will occur but that makes no sense.

W2:15 After importing the given jar files reducer code is showing Some error



The screenshot shows a Java code editor with three tabs at the top: Main.java, Map.java, and *Reduce.java. The *Reduce.java tab is active. The code is a Reducer implementation:

```
1 import org.apache.hadoop.io.IntWritable;
2
3 public class Reduce extends Reducer<Text, IntWritable, Text, LongWritable> {
4
5     @Override
6     public void reduce(Text key,
7                        Iterable<IntWritable> values,
8                        Context context)
9                     throws IOException, InterruptedException {
10        long count = 0;
11        for (IntWritable value : values) {
12            count = count + value.get();
13        }
14
15        context.write(key, new LongWritable(count));
16    }
17}
```

Ans: You need to extract the jars folder first and then add it to project

W2:16 Where intermediate data will be stored after the mapper phase completes - on a local file system or HDFS?

Ans: - Local, as we don't want 3 replicas for intermediate data.

W2:17 If we are running a MapReduce program for filtering operations and we did all the tasks at mapper end but we kept reducer and just passed the output as it is coming from mapper then which type of output part file, we will get part-m or part-r?

Ans: - if you keep reducer then part-r if you make job.setNumReduceTasks(0) then it will be part-m

W2:18 Not able to understand the 3rd example given in MapReduce_session4 completely.

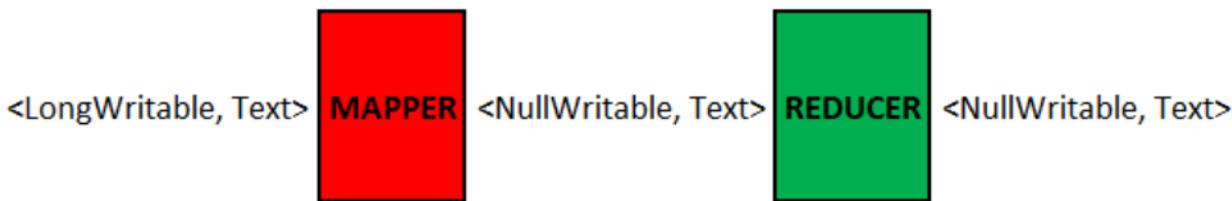
Problem Statement: We need to find out the top 3 users which are having maximum followers in twitter. There was a statement in the session like 'All mappers should output the same key for the sorting 3 records from 12 records. 'How the key value will be the same for all the 12 records from all the mappers? Can you please explain from that point?

Ans: - The key (in this case A) will be provided by the programmer. For example, in wordcount problem we provided value '1' for each key like

Ph:9108179578

(Hello, 1), similarly here programmer has to supply the key (A) to each value like (A, (1,30)).

We can also, consider key as null. Key has to be same, value can be different and values are used to find top 3.



W2:19 If in case more than one block of the same file to be processed resides on the same mapper machine then those will be run in parallel or they be processed sequentially?

Ans: - No of blocks = no of mappers. How many mappers will run concurrently that depends upon how many cores one slave node has. Suppose one data node has 4 cores then it can run 4 mappers in parallel. Coming to your question - if it is 2 blocks then 2 mappers will run in parallel.

For your information, in production environment you will generally find 16 cores/node

W2:20 Is there any order in which output from different mappers are collected together.

Ans: - No, As Soon as mapper completes, shuffle can start

W2:21 In Hadoop 2.x when do reducer tasks start? Do they start after a certain percentage (threshold) of mappers complete? If So, is this threshold fixed? What kind of threshold is typically used?

Ans: - Reduce task start after completion of Map & Shuffle and Sort

W2:22 As it is mentioned that Map is a program run on each block and output of all the mappers will be input for the reducer. How it takes care of redundant values like Block 1 is available in 2 nodes then final output will have two values for block 1, coming from node 1 and node 3. How will it be taken care of?

Ans: - Mapper runs only on 1 block of data not on all replications hence no chance of multiple outputs to reducers. In case of Speculative Execution (a

process that triggers when slower execution of a task encountered at a node), the name node starts executing another copy of the same block on the other node. The task which is finished first is accepted and the execution of others is stopped by killing the task. So, in no case multiple output of the same task is sent to the reducer.

W2:23 Why the ToolRunner interface is used to run the MapReduce application, what is the disadvantage of not using it and running the jar from the public static void main () directly?

Ans: Using ToolRunner interface makes a developer's job much easier as he no need to deal with configuration objects unnecessarily. It is easy to bind command line arguments with the conf object that you create within the 'run' method.

Developers can very well go by the 'main' method approach and deal with configuration objects on their own. However, in the Hadoop ecosystem, developers always prefer to write MapReduce programs using ToolRunner interface.

W2:24 I am getting error during adding jar from shared folder to word-count program

Ans: Try to unmount and mount again the shared folder

If still unresolved, it could be a permission issue. use terminal + sudo to check if you are able to see the contents. If So, copy it on desktop and on this copied folder, use sudo chmod 777

W2:25 Combining is like executing a reducer locally, then why can't do shuffle and Sort in combiner? Ans: Combiner will give the results locally, meaning we will have the results of one node which will only have a part of your file. But we need the output/result of entire file, that's why we need to shuffle data e.g. We want to do word count and our file is spread to 2 nodes

Let's say

node 1 has this part of file data "big data is very big big" and

node 2 has this data "big data big book"

Combiner would just give u local results

Ph:9108179578

node1 - (big, 3)

node2 - (big, 2)

If we just stop here then this is wrong results rather the word is appearing 5 times, that's why these 2 keys need to be shuffled /sent to the reducer which will give (big, 5), which is the correct count.

W2:26 If you have 10 node clusters and a 50 GB file then what would you consider doing, increase the block size?

Do we need to convert to GB-MB.. how can identify the block size if we have the file in GB. I am confused here.

Ans:

The default block size in Hadoop2 is 128MB

Here file size is 50gb = $(50*1024)/128 = 400$ blocks of 128MB.

Total no of nodes here = 10, Each node will get $400/10 = 40$ blocks.

If we will make block size = 512MB

then we will get $50*1024/512 = 100$ blocks of 512MB

Each node will get $100/10 = 10$ blocks

Here by making block size to 512MB we can process the file much quicker and better cluster utilization.

W2:27 When we create a folder over HDFS? Where is it going to create? master node or slave node if we have 3 master and 3 slaves is it going to be on all?

Ans: Master node only stores Metadata. So, obviously the folder will be created in data nodes as data nodes are meant for to store actual data. All DataNode will have the folder but few DataNode will have data which is known by NameNode. If you have a huge amount of data, the data will be divided as per the default block size(128MB) and split across your cluster

W2:28 Could there be a chance that a metadata of a large file can be split across multiple Name nodes?

Ans: No, there will be one NameNode only. In Hadoop 2 secondary NameNode has been introduced but it becomes active only when our current active NameNode grows down. Secondary Name node is not for load balancing or metadata balancing, but just for fault tolerance. No matter how big the metadata file is, it has to be persisted in the primary name node's disk, while the secondary name node does the checkpointing

W2:29 Where the output of the mapper is stored? Is it in HDFS or local?

Ans: Output of mapper is stored on local and not HDFS as its an intermediate output and storing it in HDFS won't be meaningful. Anything stored in HDFS will have by default 3 replicas and we do not want 3 replicas of intermediate data, that will be waste of HDFS space. This is when there are reducers available. If its map only job then map output is final output n stored in HDFS.

W2:30 What if intermediate data is lost in the process?

Ans: Hadoop initiates another computational resource in order to perform failed map or reduce tasks.

W2:31 If Hadoop provides high latency that means Hadoop gives more delayed response?

This is a contradictory statement?

Ans: Hadoop is for batch processing. MR is relatively slower because it's designed to work on many formats, structure and huge data. HBase sits on HDFS and is very fast because they do not follow MapReduce execution for data retrieval and processing.

The slowness occurs only because of the nature of the map/reduce based execution, where it produces lots of intermediate data, much data exchanged between nodes, thus causes huge disk IO latency. Furthermore, there is no way to persist data in disk for synchronization between phases So, that it can support Job recovery from failures. Also, there are no ways in MapReduce to cache the all/subset of the data in memory.

Conclusion: Hadoop is faster but MR is slower.

W2:32 Can anyone please let me know a top N example in MapReduce concept if the Node1 has greater number than node2 but that number is not in the top3 from node1. How can we handle this kind of example?

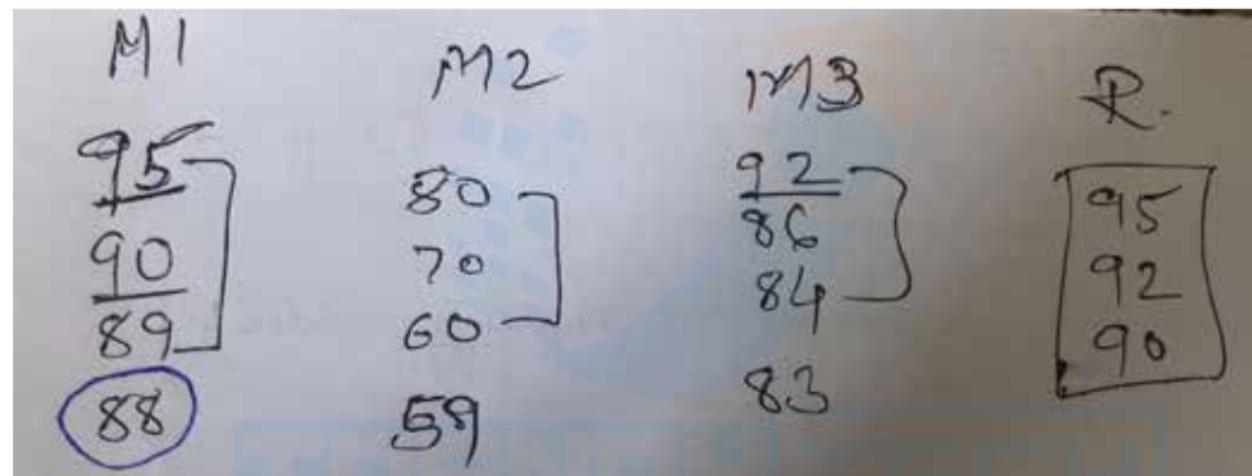
Ans: Any way those numbers would be smaller than the TOP3 from node1, So, those can't be top numbers. Hence, we can ignore them.

e.g. node 1 -> 5,6,9,8,7

node2 -> 1,2,3,4

In this case even if 5 and 6 are greater than top 3 from node2 i.e., 2,3,4 ultimately 7 8 9 are going to be greater in comparison.

Another example using picture



At the end you will always get top 3.

If you consider top 3 students from every university then this problem may arise but it's not a problem because that's your requirement.

W2:33 Does the record reader come into picture always? What if the input data is already in a key, value pair then no need for a record reader right?

Ans: Input data is not coming in the form of KEY, VALUE form, So, record reader will always work in term of giving input to the mapper in the form of key and value. It adds random key and respective line becomes a value for it

W2:34 Still I am having an issue that I moved the files to the shared folder but I cannot take these files to the package explorer nor I am not able to copy these files to the word count?

Ans:

First connect Desktop Shared folder to Cloudera shared folder

Copy MapReduce codes on Desktop shared folder then Open MapReduce codes on Cloudera by Shared folder

Open Eclipse create project in project click on source and create Class (class name Map, Reduce, Main)

Copy MapReduce codes from Cloudera shared folder

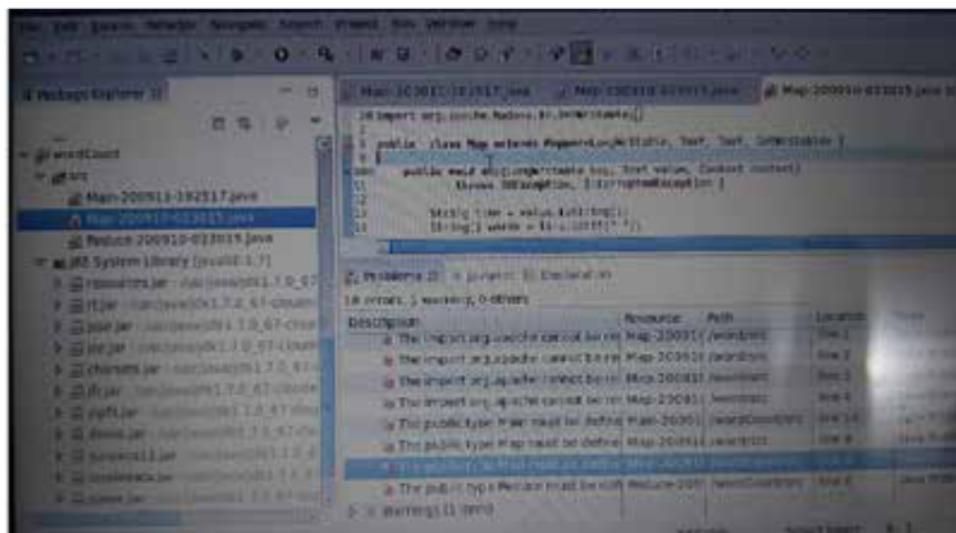
Change class name, class name and file name must be same (class WCReducer {....}, file name must be WCReducer.java). do it for all files

Run

W2:35 I've created a shared folder and mounted it on Cloudera VM, but whenever I switch it off and on, my shared folder becomes empty in VM and in order to get any files from outside VM to Cloudera VM, I need to mount again.

Ans: Always save the machine state rather than power off the machine. Refer W1 FAQ

W2:36 I have added all map reduce jar files in word count java project in eclipse. But I am getting error in map, reduce and main code.



Ph:9108179578

Ans: In Java, class name and file name must be same.

```
java - wordCount/src/Map-200910-023015.java - Eclipse
```

```
Map-200910-023015.java [Map-200910-023015.java]
```

```
18 errors, 1 warning, 0 others
```

```
Resource Path Location
```

```
The import org.apache.hadoop.io cannot be resolved line 1
```

```
The import org.apache.hadoop.mapreduce.Mapper cannot be resolved line 2
```

```
The import org.apache.hadoop.mapreduce.Mapper cannot be resolved line 3
```

```
The import org.apache.hadoop.mapreduce.Mapper cannot be resolved line 4
```

```
The public type Mapper must be defined in Map-200910-023015 line 5
```

```
The public type Mapper must be defined in Map-200910-023015 line 6
```

```
The public type Mapper must be defined in Map-200910-023015 line 7
```

```
The public type Mapper must be defined in Map-200910-023015 line 8
```

```
1 warning (1 item)
```

```
Available Smart insert 8.1.1
```

```
[ChatterList | Welcome ] [Java - wordCount.java] [File] [Search] [Help] [File] [Edit] [Source] [Navigator] [Search] [Project] [Run] [Window] [Help]
```

W2:38 John cannot find Eclipse in itversity lab, So, do he need to download it on my PC?

Ans: yes, you have to download it in your PC. Create the jar file and take it to itversity lab and execute the jar file.

W2:39 How to upload map reduce files which has been shared to word count Java project?

Ans: Way 1: Copy the files in shared location on Linux then it can be copy paste to java project.

Way 2: You can also, run MR program on eclipse available on windows.

W2:40 I understood the flow of MR operation:

Map-->Combiner-->Shuffle-->Sort-->Reducer

So, Shuffle and Sort happens once. But then in our combiner code how are we defining the value for the method as Iterable?

```
public class CustomCombiner extends Reducer<Text, IntWritable, Text, IntWritable>
{
    public void reduce(Text key, Iterable<IntWritable> values, Context context) throws
    IOException, InterruptedException
```

The output from Mapper to the Combiner would be this:

Hello,1
Hello,1
Hello,1

and not

Hello, {1,1,1} right?

Ans:

If we look at the code of `CustomCombiner` class carefully then it does not extend `Reducer` Interface and one needs to define the `reduce` unimplemented method. This phase takes the map phase output as input and the output of combiner phase is key-value collections pair. If we read the official documentation of hadoop2.0 then they have given very concise definition.

Combine When the map operation outputs its pairs they are already available in memory. For efficiency reasons, sometimes it makes sense to take advantage of this fact by supplying a combiner class to perform a reduce-type function. If a combiner is used then the map key-value pairs are not immediately written to the output. Instead, they will be collected in lists, one list per each key value. When a certain number of key-value pairs have been written, this buffer is flushed by passing all the values of each key to the combiner's `reduce` method and outputting the key-value pairs of the combine operation as if they were created by the original map operation. For example, a word count MapReduce application whose map operation outputs (word, 1) pairs as words are encountered in the input can use a combiner to speed up processing. A combine operation will start gathering the output in in-memory lists (instead of on disk), one list per word. Once a certain number of pairs is output, the combine operation will be called once per unique word with the list available as an iterator. The combiner then emits (word, count-in-this-part-of-the-input) pairs. From the viewpoint of the Reduce operation this contains the same information as the original Map output, but there should be far fewer pairs output to disk and read from disk.

Ph:9108179578

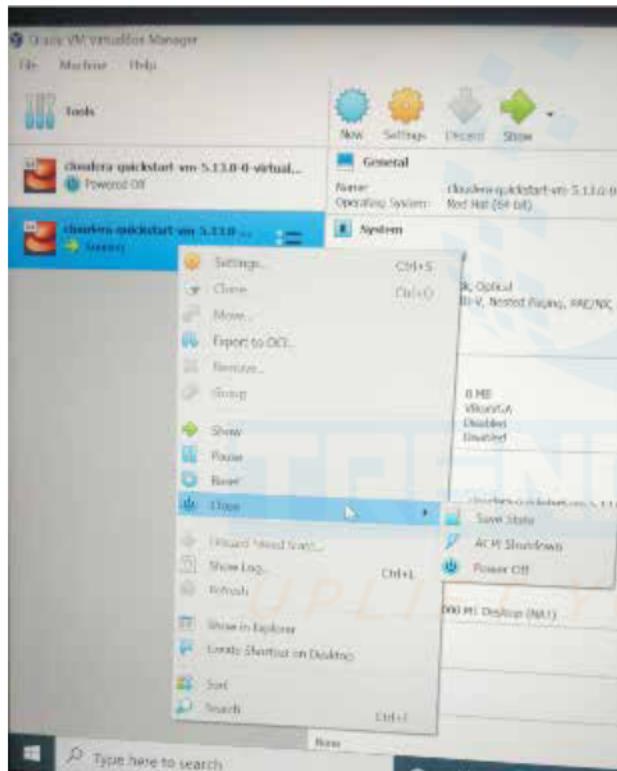
Reference Link - <https://cwiki.apache.org/confluence/display/HADOOP2/HadoopMapReduce>

Things to note - Instead they will be collected in lists, one list per each key value.

Shuffle & Sort is done only once. mapper -> combiner ->partitioner -> shuffle & Sort ->
reducer

W2: 41 How to save VM changes before shutting down the VM?

Ans: Right click on VM machine and go to close under it select saved state. Next time you start the previous things will be saved



W2:42 Unable to find files in the shared folder in VM machine. How to fix it? This worked for me previously but now I can find the shared folder

Ans: This happens because you did not save the state during previous shutdown. Now You need to configure the shared folder with the command

Open the terminal and type the commands as shown. “su” command will ask for password.

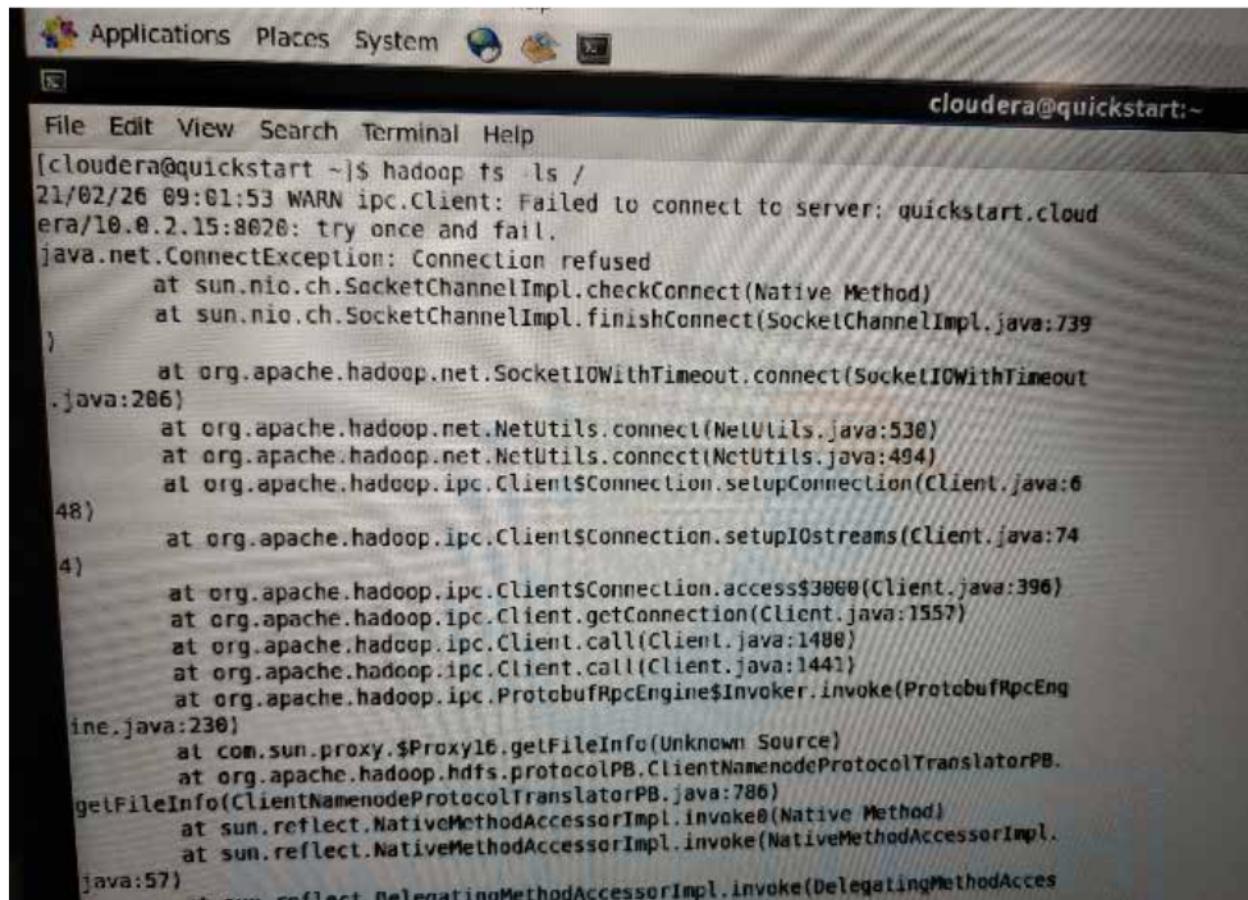
The password is Cloudera

Ph:9108179578

run command **mount -t vboxsf shared**

Steps to save state during VM shutdown: refer Q no 41

W2:43 Getting below error during executing HDFS command in terminal.



The screenshot shows a terminal window with a blue header bar containing icons for Applications, Places, System, and network status. The title bar says "cloudera@quickstart:~". The terminal window displays a stack trace from a Java application, specifically a Hadoop client, showing a connection refused error. The stack trace starts with a warning message and continues with the following stack trace:

```
[cloudera@quickstart ~]$ hadoop fs ls /  
21/02/26 09:01:53 WARN ipc.Client: Failed to connect to server: quickstart.cloud  
era/10.0.2.15:8020: try once and fail.  
java.net.ConnectException: Connection refused  
    at sun.nio.ch.SocketChannelImpl.checkConnect(Native Method)  
    at sun.nio.ch.SocketChannelImpl.finishConnect(SocketChannelImpl.java:739)  
}  
    at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout  
.java:206)  
    at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:538)  
    at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:494)  
    at org.apache.hadoop.ipc.Client$Connection.setupConnection(Client.java:6  
48)  
    at org.apache.hadoop.ipc.Client$Connection.setupIOstreams(Client.java:74)  
4}  
    at org.apache.hadoop.ipc.Client$Connection.access$3000(Client.java:396)  
    at org.apache.hadoop.ipc.Client.getConnection(Client.java:1557)  
    at org.apache.hadoop.ipc.Client.call(Client.java:1488)  
    at org.apache.hadoop.ipc.Client.call(Client.java:1441)  
    at org.apache.hadoop.ipc.ProtobufRpcEngine$Invoker.invoke(ProtobufRpcEng  
ine.java:238)  
    at com.sun.proxy.$Proxy16.getFileInfo(Unknown Source)  
    at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolTranslatorPB.  
getFileInfo(ClientNamenodeProtocolTranslatorPB.java:786)  
    at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)  
    at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.  
java:57)  
    at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAcces
```

Ans: Run this command check the status of services:

service --status-all

you may see

Ph:9108179578

```
cloudera-quickstart-vm-5.13.0-0-virtualbox [Running] Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
File Edit View Search Terminal Help
nohup python -m SimpleHTTPServer 80 &
}

# TODO: check for expired CM license and update config.js accordingly
fi
['status = start']

:cloudera-scm-agent (pid 5246) is running...
:cloudera-scm-server (pid 5280) is running...
:puspeed is stopped
:crond (pid 5365) is running...
:usage: /etc/init.d/dkms_autoinstaller {start|stop}
:insmasq is stopped
:flume NG agent is not running [FAILED]
:hadoop datanode is not running [FAILED]
:hadoop journalnode is not running [FAILED]
:hadoop namenode is not running [FAILED]
:hadoop secondarynamenode is not running [FAILED]
:hadoop https is not running [FAILED]
:hadoop historyserver is not running [FAILED]
:hadoop nodemanager is not running [FAILED]
:hadoop proxyserver is not running [FAILED]
:hadoop resourcemanager is not running [FAILED]
:yald (pid 4566) is running...
:Base master daemon is not running [FAILED]
:base-regionserver is not running [FAILED]
:Base rest daemon is not running [FAILED]
:Base Solr Indexer is not running [FAILED]
:Base thrift daemon is not running [FAILED]
```

All the services are dead, you need reinstall Cloudera VM

W2:44 Is it possible to run the Map Reduce program without starting Hadoop cluster?

Ans: We can run the map reduce program without starting the Hadoop cluster.

Hadoop work on 3 modes one is Local Mode So, we can run the program even if we didn't start the Hadoop cluster but we make sure that in this case we are not reading/writing the data in Hadoop cluster. we should read the data from local and also, output should place to local system.

Production environment works in cluster mode, So, hence it's not possible

W2:45 Where does mapper write its output? HDFS or LFS?

Ans: Outputs of mapper are written into local disk because these intermediate results don't need replications (because of which we can have less I/O operations) and only needed for the reducers.

W2:46 If we have multiple reducers then which data goes to which reducer?

Ans: Default Partitioner will take care of it using hash function.

We can provide Custom partitioner if we want particular data to go in particular partition.

Every partition has one reducer.

mapper -> partitioning -> shuffle & Sort -> reducer

W2:47: If we have 4 node cluster and each node contain 3 blocks then what would be number of mappers? is it 4 or 12?

Ans: No. of mappers = No. of Blocks (No of nodes are not considered)
So, if a Block is 128MB and the data is 500MB then it would need 4 blocks.

It should 3 or $3 \times 3 = 9$: Block replication (default 3) is done only for fault tolerance.
It doesn't mean that all the 3 same replicated blocks will be processed.

Consider a case where you have 500MB size file and we have 3 node cluster.
500MB size file will have 4 blocks (b1, b2, b3, b4).

Suppose block 1(b1) is stored on all the 3 machines as replication factor is 3,
but when map phase is launched, b1 will be processed on only 1 machine out
of 3 based on data locality.

W2:48: Hadoop provides high latency and high throughput.

Ans: HADOOP ecosystem is for dealing with huge amount of data which is distributed across multiple nodes in the cluster and mostly for batch processing and not for real time processing. That's the reason it takes time to process the input (high latency) but process millions of records in the single batch run (high throughput)

W2:49 The combiner performs same operations as the reducer. So, without reducer, we can consider the final output from combiner?

Ans: Combiner can have same code as reducer or it may differ based on the requirement.

If the expected output can be obtained from combiner then the reducer is set to 0 and combiner output will be considered as final output

Sometime combiner is used to decrease the load of reducer but it doesn't give us final output. In that case combiner output is passed on to reducer and reducer output will be considered as final output

W2:50 If we consider combiner output as final then partitioner will be invoked or not?

Ans: No, Partitioner is used when we have more than 1 reducer, it is used to distribute the data equally to all the reducer. With reducer equal to 0 partitioner is not required.

