

Simplified Steps to Execute MapReduce Program in Cloudera Quickstart VM

1. Create a shared folder in the host and mount it to the guest VM:

```
```bash

sudo mkdir -p /mnt/shared

sudo mount -t vboxsf mapreduce /mnt/shared

...`
```

2. Verify files in the shared folder:

```
```bash

ls /mnt/shared

...`
```

3. Upload the required files to HDFS:

```
```bash

hadoop fs -put /mnt/shared/access_log_short.csv /user/cloudera/

hadoop fs -put /mnt/shared/UserLogMapper.java /user/cloudera/

hadoop fs -put /mnt/shared/UserLogDriver.java /user/cloudera/

hadoop fs -put /mnt/shared/UserLogReducer.java /user/cloudera/

...`
```

4. Create a directory for compiled class files:

```
```bash

mkdir -p /home/cloudera/userlog/classes

...`
```

5. Compile Java code:

```
```bash

javac -classpath $(hadoop classpath) -d /home/cloudera/userlog/classes UserLogMapper.java
UserLogDriver.java UserLogReducer.java

...`
```

6. Create a JAR file:

```
```bash
```

```
jar -cvf /home/cloudera/userlog/userlog.jar -C /home/cloudera/userlog/classes .
```

```
```
```

#### 7. Run the Hadoop MapReduce job:

```
```bash
```

```
hadoop jar /home/cloudera/userlog/userlog.jar mrLogFile_demo.UserLogDriver
```

```
/user/cloudera/access_log_short.csv /user/cloudera/output
```

```
```
```

#### 8. Check output directory in HDFS:

```
```bash
```

```
hadoop fs -ls /user/cloudera/output
```

```
```
```

#### 9. View the job output:

```
```bash
```

```
hadoop fs -cat /user/cloudera/output/part-r-00000
```

```
```
```