CLOUD COMPUTING

Assignment 03

ALI ABBAS

BCSF17A502

```
        HDFS: Number of write operations=2
        HDFS: Number of bytes read erasure-coded=0
Job Counters
        Launched map tasks=1
        Launched reduce tasks=1
        Data-local map tasks=1
        Total time spent by all maps in occupied slots (ms)=2239
        Total time spent by all reduces in occupied slots (ms)=2504
        Total time spent by all map tasks (ms)=2239
        Total time spent by all reduce tasks (ms)=2504
        Total vcore-milliseconds taken by all map tasks=2239
        Total vcore-milliseconds taken by all reduce tasks=2504
        Total megabyte-milliseconds taken by all map tasks=2292736
        Total megabyte-milliseconds taken by all reduce tasks=2564096
Map-Reduce Framework
        Map input records=3
        Map output records=8
        Map output bytes=83
        Map output materialized bytes=80
        Input split bytes=110
        Combine input records=8
        Combine output records=6
        Reduce input groups=6
        Reduce shuffle bytes=80
        Reduce input records=6
        Reduce output records=6
        Spilled Records=12
        Shuffled Maps =1
        Failed Shuffles=0
        Merged Map outputs=1
        GC time elapsed (ms)=154
        CPU time spent (ms)=1450
        Physical memory (bytes) snapshot=524750848
        Virtual memory (bytes) snapshot=5219487744
        Total committed heap usage (bytes)=456130560
        Peak Map Physical memory (bytes)=292489344
        Peak Map Virtual memory (bytes)=2606850048
        Peak Reduce Physical memory (bytes)=232341504
        Peak Reduce Virtual memory (bytes)=2612637696
Shuffle Errors
        BAD_ID=0
        CONNECTION=0
        IO_ERROR=0
        WRONG_LENGTH=0
        WRONG_MAP=0
        WRONG_REDUCE=0
File Input Format Counters
        Bytes Read=51
```

Activities   Terminal ▾   Mon 08:01

Tabs: Cloud-Computing-Project× | Google × | Cloud Computing BS CS P× | How to Install Hadoop on× | Apache Download Mirror× | +

https://phoenixnap.com/kb/install-hadoop-ubuntu

Please only use the backup mirrors to download KEYS, PGP signatures and hashes (SHA* etc) -- or if no other mirrors are working.

HTTP

https://downloads.apache.org/hadoop/common/hadoop-3.2.1/hadoop-3.2.1.tar.gz

-- or if no other mirrors are working.

```
                hdoop@dev-HP-ProBook-450-G7: ~
File  Edit  View  Search  Terminal  Help
|     o.Booo     |
|  .  o.+o=.o..+|
|     oo.*=o.oE.+*|
+----[SHA256]-----+
hdoop@dev-HP-ProBook-450-G7:~$ cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
hdoop@dev-HP-ProBook-450-G7:~$ chmod 0600 ~/.ssh/authorized_keys
hdoop@dev-HP-ProBook-450-G7:~$ ssh localhost
The authenticity of host 'localhost (127.0.0.1)' can't be established.
ECDSA key fingerprint is SHA256:Ch8qIMzToFuV++d+05SyPRHjl9dsQtD5MxPoCit8gmA.
Are you sure you want to continue connecting (yes/no)? y
Please type 'yes' or 'no':
Host key verification failed.
hdoop@dev-HP-ProBook-450-G7:~$ wget https://downloads.apache.org/hadoop/common/hadoop-3.2.1/had
oop-3.2.1.tar.gz
--2021-06-14 07:59:23--  https://downloads.apache.org/hadoop/common/hadoop-3.2.1/hadoop-3.2.1.t
ar.gz
Resolving downloads.apache.org (downloads.apache.org)... 88.99.95.219, 135.181.209.10, 135.181.
214.104, ...
Connecting to downloads.apache.org (downloads.apache.org)|88.99.95.219|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 359196911 (343M) [application/x-gzip]
Saving to: 'hadoop-3.2.1.tar.gz'

hadoop-3.2.1.tar.gz      22%[=====>                ] 77.04M  1.55MB/s    eta 6m 17s
```

...mirror sites. The
...e official

...-3.1.4.tar.gz

...org

(Default) ▾   ...oop-3.2.1.

...e this from now on.   ...op-3.2.1/hadoo

Cancel   OK   ...2.1/hadoop-3.2

```
Resolving downloads.apache.org (downloads.apache.org)... ...:f8:10a:201a::2
Connecting to downloads.apache.org (downloads.apache.org)|88.99.95.219|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 359196911 (343M) [application/x-gzip]
Saving to: 'hadoop-3.2.1.tar.gz'

hadoop-3.2.1.tar.gz    100%[===========================>] 342.56M  10.2MB/s    in 31s

2020-04-22 09:29:23 (10.9 MB/s) - 'hadoop-3.2.1.tar.gz' saved [359196911/359196911]
```
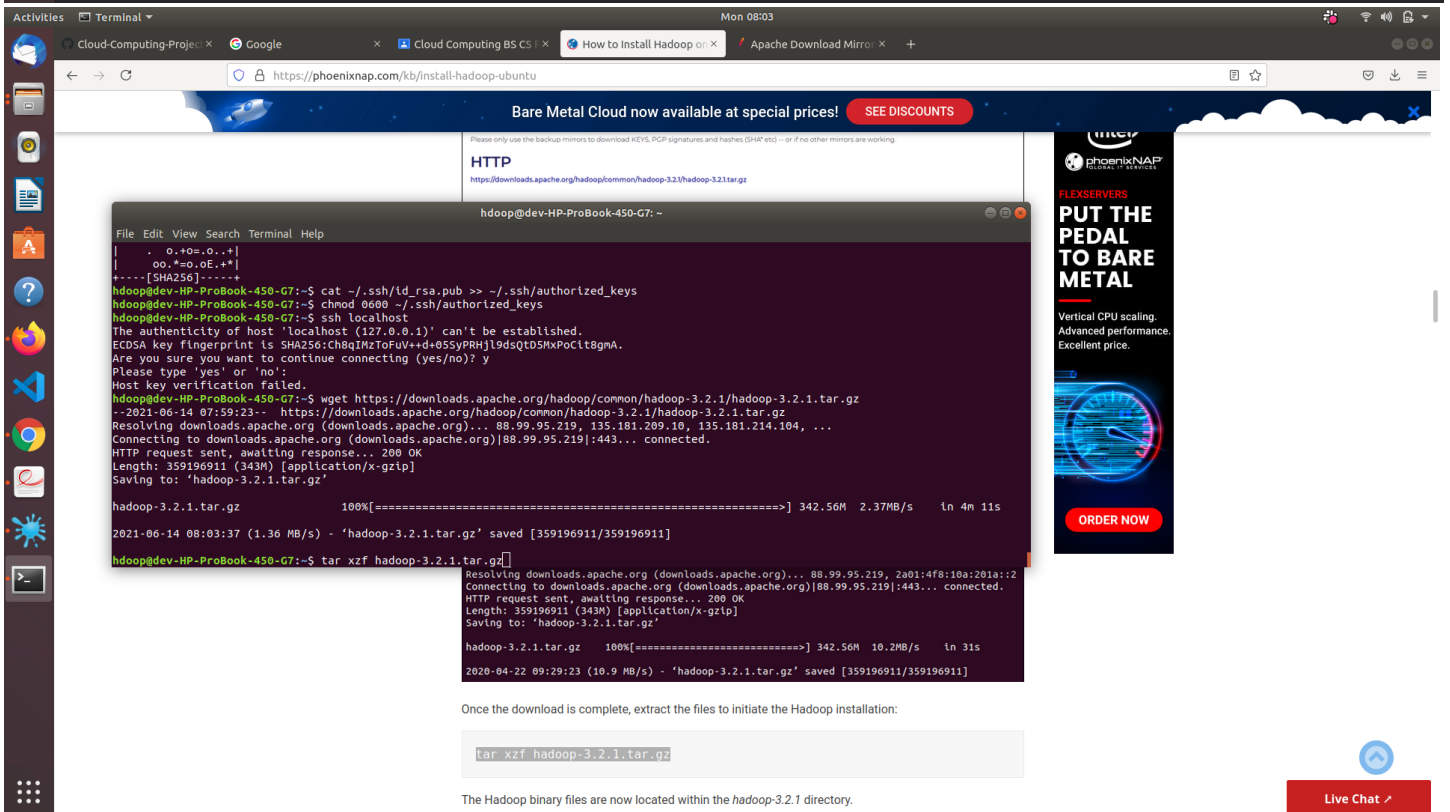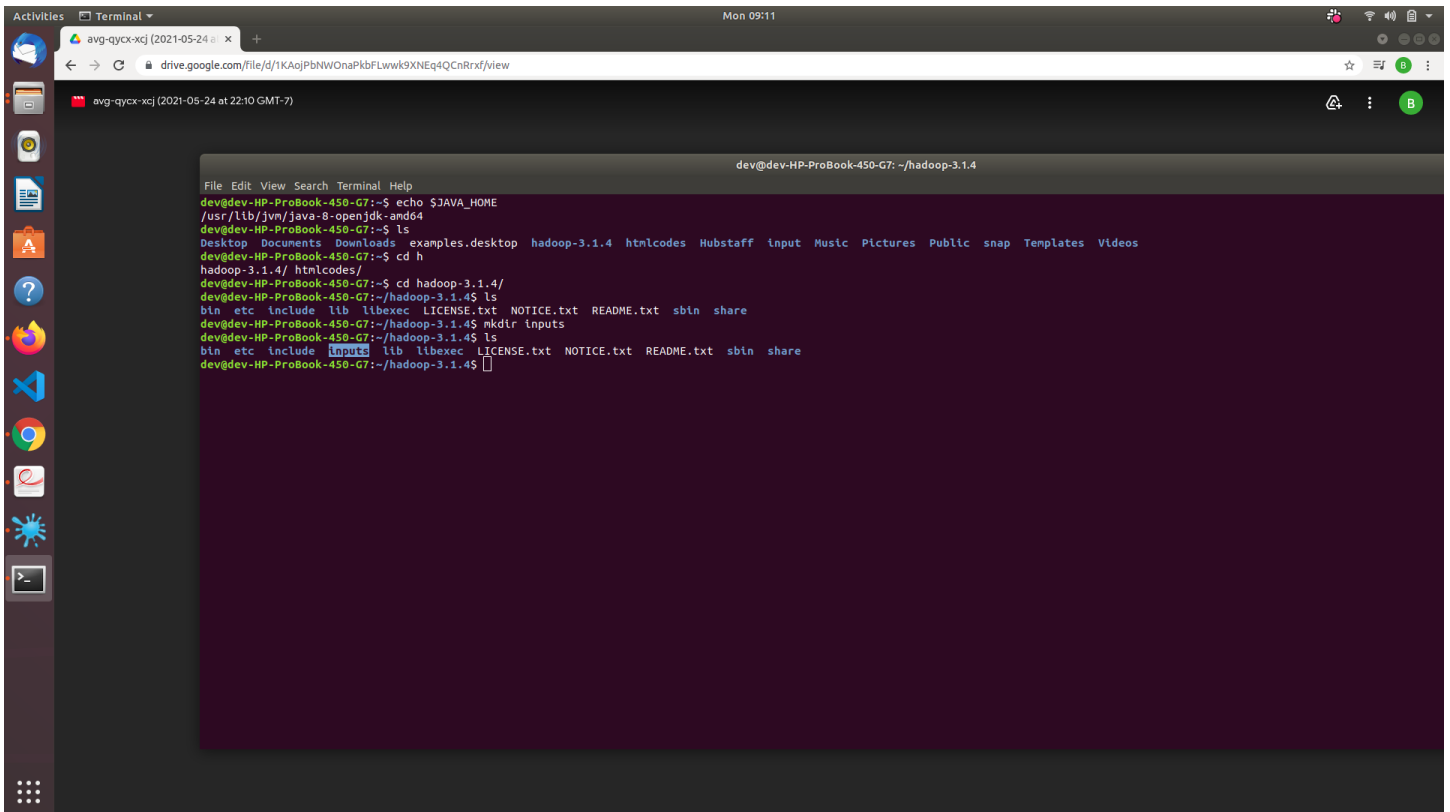
Once the download is complete, extract the files to initiate the Hadoop installation:

```
tar xzf hadoop-3.2.1.tar.gz
```

The Hadoop binary files are now located within the *hadoop-3.2.1* directory.

```
2021-06-14 10:10:33,067 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2021-06-14 10:10:33,222 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2021-06-14 10:10:33,286 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2021-06-14 10:10:33,286 INFO impl.MetricsSystemImpl: JobTracker metrics system started
2021-06-14 10:10:33,379 INFO input.FileInputFormat: Total input files to process : 0
2021-06-14 10:10:33,392 INFO mapreduce.JobSubmitter: number of splits:0
2021-06-14 10:10:33,484 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1671980215_0001
2021-06-14 10:10:33,484 INFO mapreduce.JobSubmitter: Executing with tokens: []
2021-06-14 10:10:33,585 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
2021-06-14 10:10:33,585 INFO mapreduce.Job: Running job: job_local1671980215_0001
2021-06-14 10:10:33,586 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2021-06-14 10:10:33,592 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2021-06-14 10:10:33,592 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2021-06-14 10:10:33,593 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter
2021-06-14 10:10:33,618 INFO mapred.LocalJobRunner: Waiting for map tasks
2021-06-14 10:10:33,618 INFO mapred.LocalJobRunner: map task executor complete.
2021-06-14 10:10:33,621 INFO mapred.LocalJobRunner: Waiting for reduce tasks
2021-06-14 10:10:33,622 INFO mapred.LocalJobRunner: Starting task: attempt_local1671980215_0001_r_000000_0
2021-06-14 10:10:33,640 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2021-06-14 10:10:33,640 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2021-06-14 10:10:33,652 INFO mapred.Task:  Using ResourceCalculatorProcessTree : [ ]
2021-06-14 10:10:33,654 INFO mapred.ReduceTask: Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@1c343f0d
2021-06-14 10:10:33,655 WARN impl.MetricsSystemImpl: JobTracker metrics system already initialized!
2021-06-14 10:10:33,666 INFO reduce.MergeManagerImpl: MergerManager: memoryLimit=3240257024, maxSingleShuffleLimit=810064256, mergeThreshold=2138569728, ioSortFactor=10, memToMemMergeOutputsThreshold=10
eshold=10
2021-06-14 10:10:33,668 INFO reduce.EventFetcher: attempt_local1671980215_0001_r_000000_0 Thread started: EventFetcher for fetching Map Completion Events
2021-06-14 10:10:33,673 INFO reduce.EventFetcher: EventFetcher is interrupted.. Returning
2021-06-14 10:10:33,676 INFO mapred.LocalJobRunner:
2021-06-14 10:10:33,676 INFO reduce.MergeManagerImpl: finalMerge called with 0 in-memory map-outputs and 0 on-disk map-outputs
2021-06-14 10:10:33,677 INFO reduce.MergeManagerImpl: Merging 0 files, 0 bytes from disk
2021-06-14 10:10:33,677 INFO reduce.MergeManagerImpl: Merging 0 segments, 0 bytes from memory into reduce
2021-06-14 10:10:33,680 INFO mapred.Merger: Merging 0 sorted segments
2021-06-14 10:10:33,680 INFO mapred.Merger: Down to the last merge-pass, with 0 segments left of total size: 0 bytes
2021-06-14 10:10:33,681 INFO mapred.LocalJobRunner:
2021-06-14 10:10:33,707 INFO Configuration.deprecation: mapred.skip.on is deprecated. Instead, use mapreduce.job.skiprecords
2021-06-14 10:10:33,718 INFO mapred.Task: Task:attempt_local1671980215_0001_r_000000_0 is done. And is in the process of committing
2021-06-14 10:10:33,719 INFO mapred.LocalJobRunner:
2021-06-14 10:10:33,719 INFO mapred.Task: Task attempt_local1671980215_0001_r_000000_0 is allowed to commit now
2021-06-14 10:10:33,721 INFO output.FileOutputCommitter: Saved output of task 'attempt_local1671980215_0001_r_000000_0' to file:/home/hdoop/hadoop-3.2.1/grep-temp-1466437422
2021-06-14 10:10:33,721 INFO mapred.LocalJobRunner: reduce > reduce
2021-06-14 10:10:33,721 INFO mapred.Task: Task 'attempt_local1671980215_0001_r_000000_0' done.
2021-06-14 10:10:34,812 INFO mapred.LocalJobRunner: Finishing task: attempt_local305787015_0002_r_000000_0
2021-06-14 10:10:34,812 INFO mapred.LocalJobRunner: reduce task executor complete.
2021-06-14 10:10:35,710 INFO mapreduce.Job: Job job_local305787015_0002 running in uber mode : false
2021-06-14 10:10:35,711 INFO mapreduce.Job:  map 100% reduce 100%
2021-06-14 10:10:35,711 INFO mapreduce.Job: Job job_local305787015_0002 completed successfully
2021-06-14 10:10:35,712 INFO mapreduce.Job: Counters: 30
        File System Counters
                FILE: Number of bytes read=1266784
                FILE: Number of bytes written=3356368
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
        Map-Reduce Framework
                Map input records=0
                Map output records=0
                Map output bytes=0
                Map output materialized bytes=6
                Input split bytes=128
                Combine input records=0
                Combine output records=0
                Reduce input groups=0
                Reduce shuffle bytes=6
                Reduce input records=0
                Reduce output records=0
                Spilled Records=0
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=0
                Total committed heap usage (bytes)=790626304
        Shuffle Errors
                BAD_ID=0
                CONNECTION=0
                IO_ERROR=0
                WRONG_LENGTH=0
                WRONG_MAP=0
                WRONG_REDUCE=0
        File Input Format Counters
                Bytes Read=98
        File Output Format Counters
                Bytes Written=8
```

```
                HDFS: Number of write operations=2
                HDFS: Number of bytes read erasure-coded=0
        Job Counters
                Launched map tasks=1
                Launched reduce tasks=1
                Data-local map tasks=1
                Total time spent by all maps in occupied slots (ms)=2239
                Total time spent by all reduces in occupied slots (ms)=2504
                Total time spent by all map tasks (ms)=2239
                Total time spent by all reduce tasks (ms)=2504
                Total vcore-milliseconds taken by all map tasks=2239
                Total vcore-milliseconds taken by all reduce tasks=2504
                Total megabyte-milliseconds taken by all map tasks=2292736
                Total megabyte-milliseconds taken by all reduce tasks=2564096
        Map-Reduce Framework
                Map input records=3
                Map output records=8
                Map output bytes=83
                Map output materialized bytes=80
                Input split bytes=116
                Combine input records=8
                Combine output records=6
                Reduce input groups=6
                Reduce shuffle bytes=80
                Reduce input records=6
                Reduce output records=6
                Spilled Records=12
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=154
                CPU time spent (ms)=1450
                Physical memory (bytes) snapshot=524750848
                Virtual memory (bytes) snapshot=5219487744
                Total committed heap usage (bytes)=456130560
                Peak Map Physical memory (bytes)=292409344
                Peak Map Virtual memory (bytes)=2606850048
                Peak Reduce Physical memory (bytes)=232341504
                Peak Reduce Virtual memory (bytes)=2612637696
        Shuffle Errors
                BAD_ID=0
                CONNECTION=0
                IO_ERROR=0
                WRONG_LENGTH=0
                WRONG_MAP=0
                WRONG_REDUCE=0
        File Input Format Counters
                Bytes Read=51
```