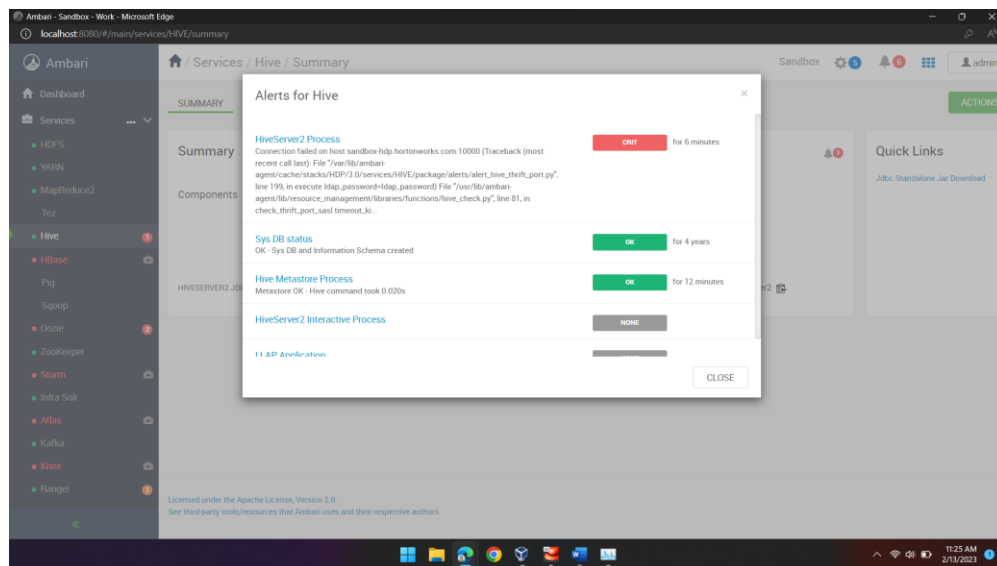


Praktikum Teknologi Perencanaan Data

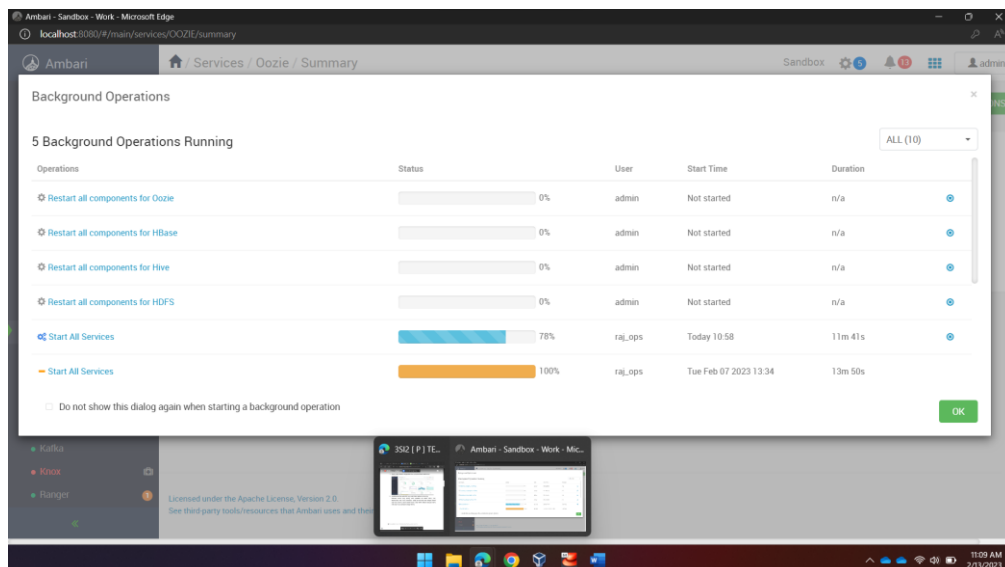
Tanggal Praktikum : Monday, February 13, 2023

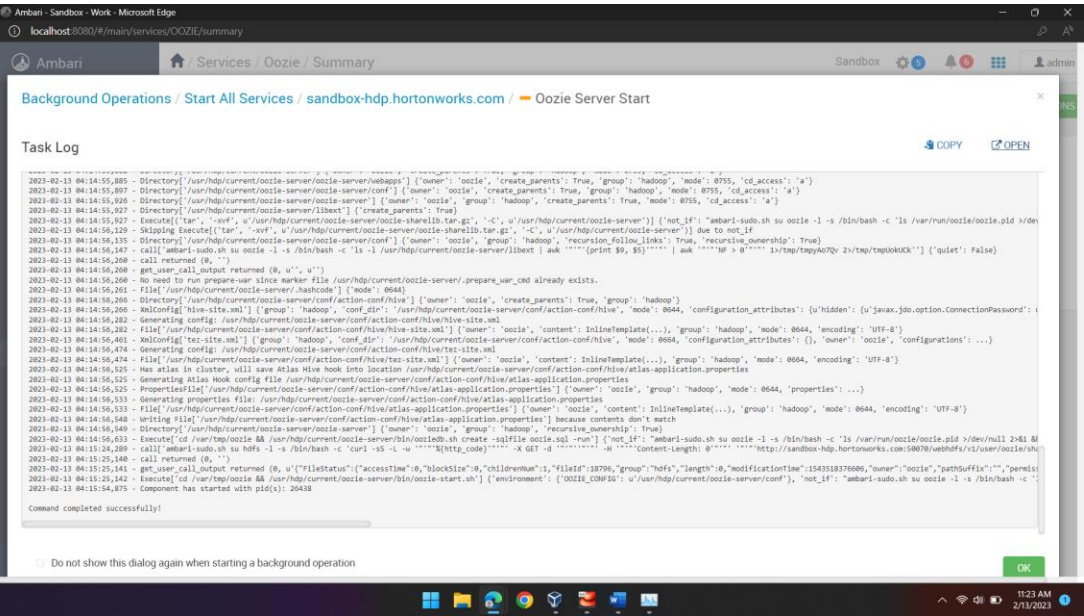
1. Troubleshooting

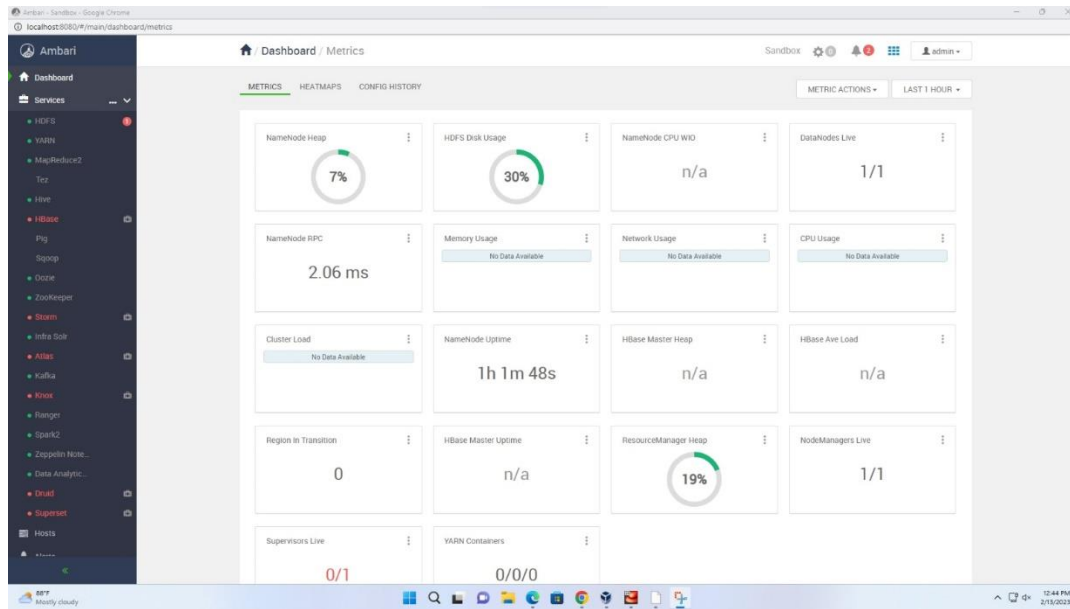
Warning yang muncul : connection failed ke laman hdp sandbox



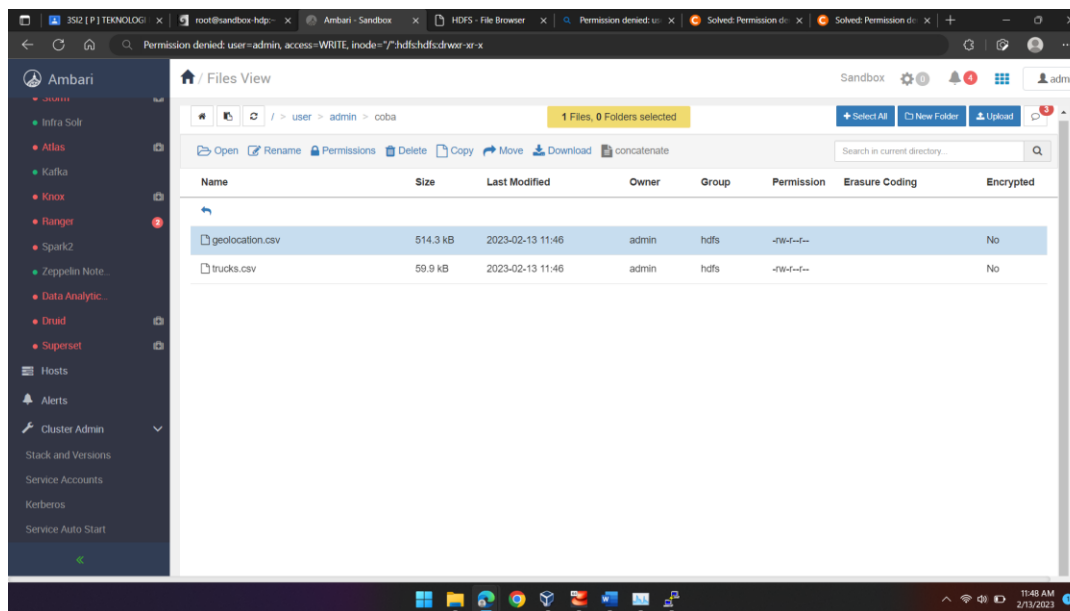
Melakukan Restart Services



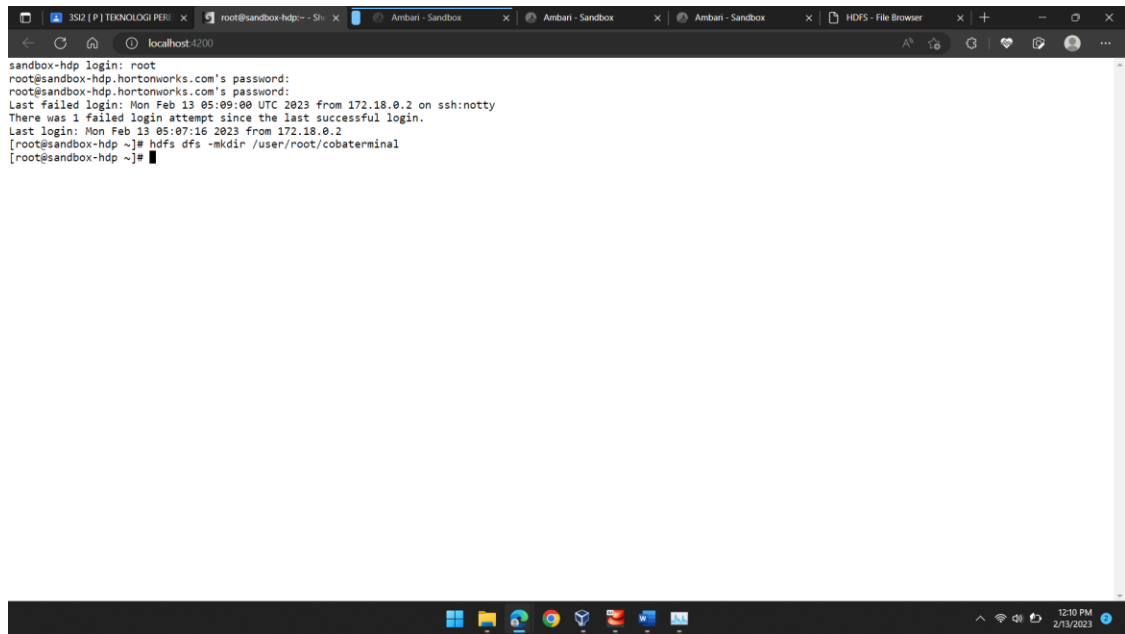




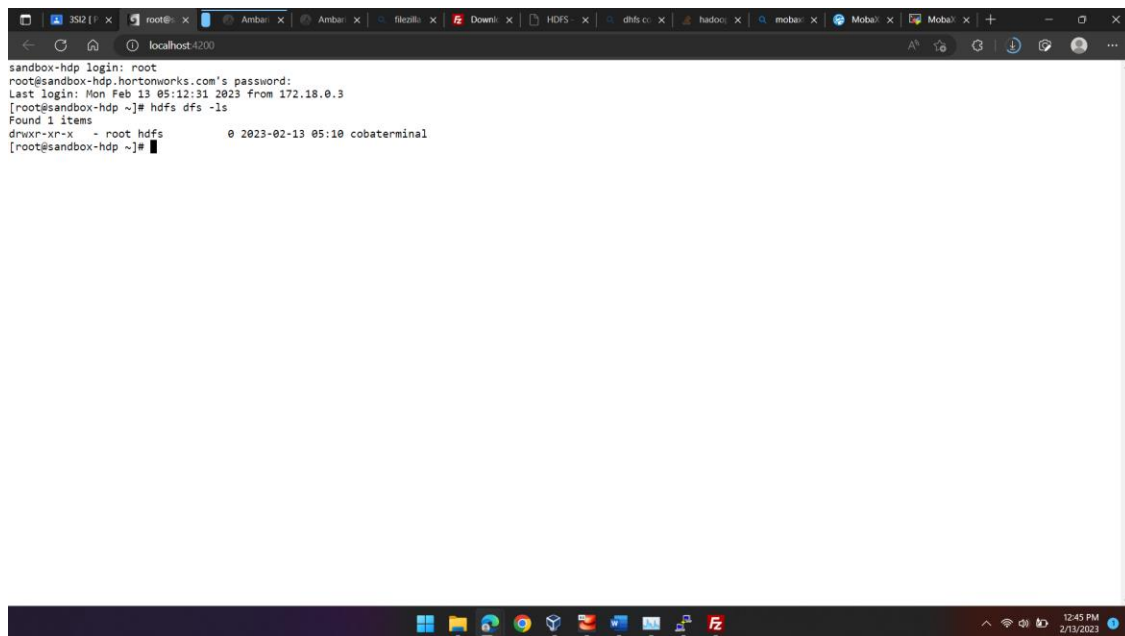
Membuat direktori dan upload file



Create Directory



View Directory With Permission



Remove Directory

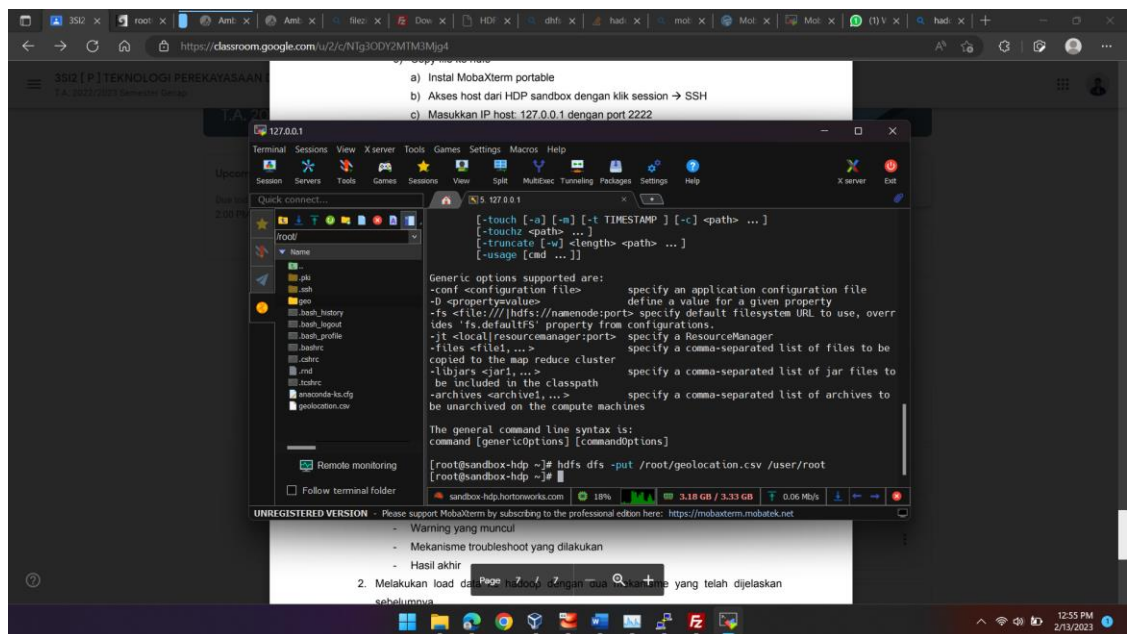
```
[-get [-f] [-p] [-ignoreCrc] [-crc] <src> ... <localdst>]
[-getfacl [-R] <path>]
[-getfattr [-R] {-n name | -d} [-e en] <path>]
[-getmerge [-nl] [-skip-empty-file] <src> <localdst>]
[-head <file>]
[-help [cmd ...]]
[-ls [-C] [-d] [-h] [-q] [-R] [-t] [-S] [-r] [-U] [-e] <path> ...]]
[-mkdir [-p] <path> ...]
[-moveFromLocal <localsrc> ... <dst>]
[-moveToLocal <src> <localdst>]
[-mv <src> ... <dst>]
[-put [-f] [-p] [-l] [-d] <localsrc> ... <dst>]
[-renameSnapshot <snapshotDir> <oldName> <newName>]
[-rm [-f] [-r] [-R] [-skipTrash] [-safely] <src> ...]
[-rmdir [--ignore-fail-on-non-empty] <dir> ...]
[-setfacl [-R] [{-b|-k} {-m|-x <acl_spec>} <path>][[--set <acl_spec> <path>]]
[-setfattr {-n name [-v value] | -x name} <path>]
[-setrep [-R] [-w] <rep> <path> ...]
[-stat [format] <path> ...]
[-tail [-f] <file>]
[-test [-defsz] <path>]
[-text [--ignoreCrc] <src> ...]
[-touch [-a] [-m] [-t TIMESTAMP] [--c] <path> ...]
[-touchz <path> ...]
[-truncate [-w] <length> <path> ...]
[-usage [cmd ...]]

Generic options supported are:
-conf <configuration file>      specify an application configuration file
-D <property=value>             define a value for a given property
-fs <file:///hdfs://namenode:port> specify default filesystem URL to use, overrides 'fs.defaultFS' property from configurations.
-specify a ResourceManager
-files <file1,...>               specify a comma-separated list of files to be copied to the map reduce cluster
-libjars <jar1,...>              specify a comma-separated list of jar files to be included in the classpath
-archives <archive1,...>        specify a comma-separated list of archives to be unarchived on the compute machines

The general command line syntax is:
command [genericOptions] [commandOptions]

[root@sandbox-hdp ~]# hdfs dfs -rm -r /user/root/cobaterminal
[root@sandbox-hdp ~]#
```

Upload File dan Memindah File menggunakan mobaXterm



1. Help

Untuk melihat bantuan command

```
localhost:4200
methods for the value. If the argument is enclosed in double quotes,
then the value is the string inside the quotes. If the argument is
prefixed with 0x or 0X, then it is taken as a hexadecimal number. If
the argument begins with 0s or 0S, then it is taken as a base64
encoding.
-x name Remove the extended attribute.
<path> The file or directory.

-setrep [-R] [-w] <rep> <path> ... :
Set the replication level of a file. If <path> is a directory then the command
recursively changes the replication factor of all files under the directory tree
rooted at <path>. The EC files will be ignored here.

-w It requests that the command waits for the replication to complete. This
can potentially take a very long time.
-R It is accepted for backwards compatibility. It has no effect.

-stat [format] <path> ... :
Print statistics about the file/directory at <path>
in the specified format. Format accepts permissions in
octal (%a) and symbolic (%A), filesize in
bytes (%b), type (%F), group name of owner (%g),
name (%n), block size (%o), replication (%r), user name
of owner (%u), access date (%x, %X),
modification date (%y, %Y),
%a and %y show UTC date as "yyyy-MM-dd HH:mm:ss" and
%X and %Y show milliseconds since January 1, 1970 UTC.
If the format is not specified, %y is used by default.

-tail [-f] <file> :
Show the last 1KB of the file.

-f Shows appended data as the file grows.

-test [-defsz] <path> :
Answer various questions about <path>, with result via exit status.
-d return 0 if <path> is a directory.
-e return 0 if <path> exists.
-f return 0 if <path> is a file.
-s return 0 if file <path> is greater than zero bytes in size.
-w return 0 if file <path> exists and write permission is granted.
-r return 0 if file <path> exists and read permission is granted.
```

2. Count

untuk melihat jumlah file yang ada di directory

```
localhost:4200
[ -d [-h] <path> ... ]
[ -du [-s] [-h] [-v] [-x] <path> ... ]
[ -expunge ]
[ -find <path> ... <expression> ... ]
[ -get [-f] [-p] [-ignorecrc] [-crc] <src> ... <localdst> ]
[ -getfacl [-R] <path> ]
[ -getfattr [-R] [-n name [-d] [-e an] <path> ]
[ -getmerge [-n] [-skip-empty-file] <src> <localdst> ]
[ -head <file> ]
[ -help [cmd ...] ]
[ -ls [-C] [-d] [-h] [-q] [-R] [-t] [-S] [-r] [-u] [-e] <path> ... ]
[ -mkdir [-p] <path> ... ]
[ -movefromlocal <localsrc> ... <dst> ]
[ -movetolocal <src> <localdst> ]
[ -mv <src> ... <dst> ]
[ -put [-f] [-p] [-l] [-d] <localsrc> ... <dst> ]
[ -renameSnapshot snapshotDir oldName newName ]
[ -rm [-f] [-r] [-R] [-skipTrash] [-safely] <src> ... ]
[ -rmDir [-ignore-fail-on-non-empty] <dir> ... ]
[ -setfacl [-R] [-h] [-x] [-m] [-x <acl_spec>] <path> ] [[--set <acl_spec>] <path>]]
[ -setfattr [-n name [-v value] | -x name] <path> ]
[ -setrep [-R] [-w] <rep> <path> ... ]
[ -stat [format] <path> ... ]
[ -tail [-f] <file> ]
[ -test [-defsz] <path> ]
[ -test [-ignorecrc] <src> ... ]
[ -touch [-a] [-m] [-t TIMESTAMP] [-c] <path> ... ]
[ -touchs <path> ... ]
[ -truncate [-u] <length> <path> ... ]
[ -usage [cmd ...] ]

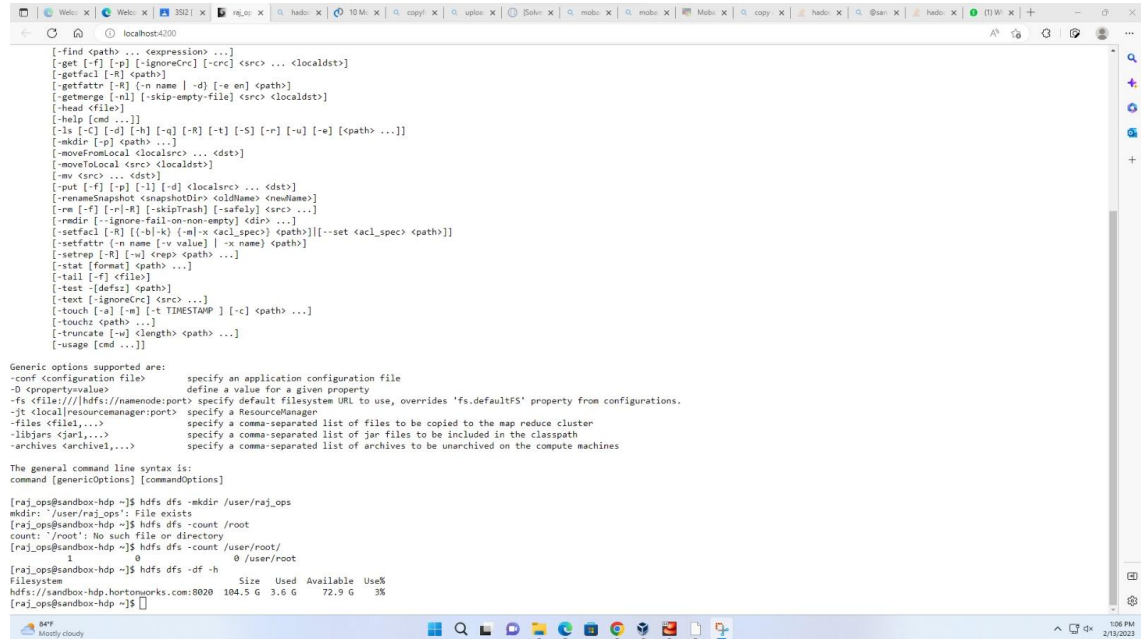
Generic options supported are:
-conf <configuration file> specify an application configuration file
-D <property=value> define a value for a given property
-fs <file>:///hdfs://namenode:port specify default filesystem URI to use, overrides 'fs.defaultFS' property from configurations.
-it <localResourceManager:port> specify a ResourceManager
-files <file>,... specify a comma-separated list of files to be copied to the map reduce cluster
-libjars <jar>,... specify a comma-separated list of jar files to be included in the classpath
-archives <archive>,... specify a comma-separated list of archives to be unarchived on the compute machines

The general command line syntax is:
command [genericOptions] [commandOptions]

[raj_ops@sandbox-hdp ~]$ hdfs dfs -mkdir /user/raj_ops
mkdir: '/user/raj_ops': File exists
[raj_ops@sandbox-hdp ~]$ hdfs dfs -count /root
count: /root: no such file or directory
[raj_ops@sandbox-hdp ~]$ hdfs dfs -count /user/root/
1 0 /user/root/
[raj_ops@sandbox-hdp ~]$
```

3. Df

Untuk melihat free space



```
[~] find <path> ... <expression> ...
[~] get [-f] [-p] [-ignorecrc] [-crc] <src> ... <localdst>
[~] getfacl [-R] <path>
[~] getfattr [-R] [-n name] [-d] [-e en] <path>
[~] getmerge [-nl] [-skip-empty-file] <src> <localdst>
[~] head <file>
[~] help [cmd ...]
[~] ls [-C] [-d] [-h] [-q] [-R] [-t] [-S] [-r] [-u] [-e] [<path> ...]
[~] mkdir [-p] <path> ...
[~] movefromlocal <localsrc> ... <dst>
[~] movefromlocal <src> <localdst>
[~] mv <src> ... <dst>
[~] put [-f] [-p] [-l] [-d] <localsrc> ... <dst>
[~] renameSnapshot <snapshotDir> <oldName> <newName>
[~] rm [-f] [-r] [-R] [-skipTrash] [-safely] <src> ...
[~] rmdir [-ignore-fail-on-non-empty] <dir> ...
[~] setfacl [-R] [[-b|-k] {-m|-x <acl_spec>} <path>][[--set <acl_spec> <path>]]
[~] setfattr [-n name [-v value] | -x name] <path>
[~] setrep [-R] [-w] <rep> <path> ...
[~] stat [format] <path> ...
[~] tail [-f] <file>
[~] test [-defsz] <path>
[~] text [-ignorecrc] <src> ...
[~] touch [-a] [-m] [-t TIMESTAMP ] [-c] <path> ...
[~] touch <path> ...
[~] truncate [-w] <length> <path> ...
[~] usage [cmd ...]

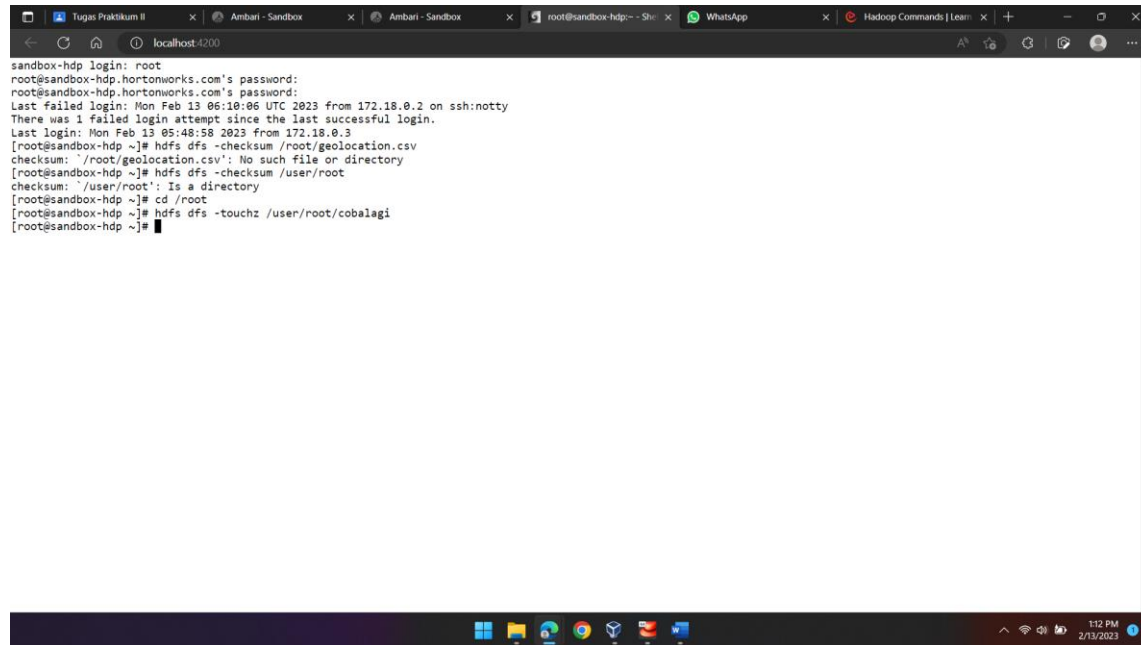
Generic options supported are:
-conf <configuration file>          specify an application configuration file
-D <property=value>                 define a value for a given property
-fs <file:///hdfs://namenode:port> specify default filesystem URL to use, overrides 'fs.defaultFS' property from configurations.
-jt <local[resource]manager:port>    specify a ResourceManager
-files <file1,...>                  specify a comma-separated list of files to be copied to the map reduce cluster
-libjars <jar1,...>                 specify a comma-separated list of jar files to be included in the classpath
-archives <archive1,...>            specify a comma-separated list of archives to be unarchived on the compute machines

The general command line syntax is:
command [genericOptions] [commandOptions]

[raj_ops@sandbox-hdp ~]$ hdfs dfs -mkdir /user/raj_ops
mkdir: '/user/raj_ops': File exists
[raj_ops@sandbox-hdp ~]$ hdfs dfs -count /root
count: '/root': No such file or directory
[raj_ops@sandbox-hdp ~]$ hdfs dfs -count /user/root/
      1      0
[raj_ops@sandbox-hdp ~]$ hdfs dfs -df -h
Filesystem              Size      Used Available Use%
hdfs://sandbox-hdp.hortonworks.com:8020 104.5 G  3.6 G    72.9 G   3%
[raj_ops@sandbox-hdp ~]$
```

4. Touchz

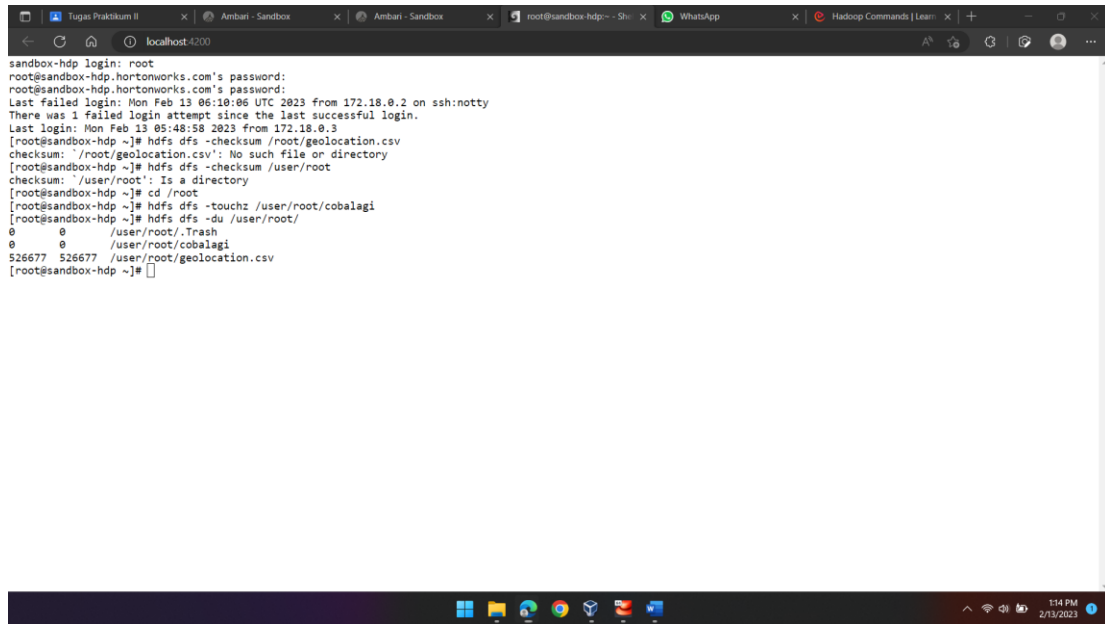
Membuat file dengan timestamp



```
sandbox-hdp login: root
root@sandbox-hdp.hortonworks.com's password:
root@sandbox-hdp.hortonworks.com's password:
Last failed login: Mon Feb 13 06:10:06 UTC 2023 from 172.18.0.2 on ssh:notty
There was 1 failed login attempt since the last successful login.
Last login: Mon Feb 13 05:48:56 2023 from 172.18.0.3
[root@sandbox-hdp ~]# hdfs dfs -checksum /root/geolocation.csv
checksum: '/root/geolocation.csv': No such file or directory
[root@sandbox-hdp ~]# hdfs dfs -checksum /user/root
checksum: '/user/root': Is a directory
[root@sandbox-hdp ~]# cd /root
[root@sandbox-hdp ~]# hdfs dfs -touchz /user/root/cobalagi
[root@sandbox-hdp ~]#
```

5. Du

Melihat usage directory



```
sandbox-hdp login: root
root@sandbox-hdp.hortonworks.com's password:
root@sandbox-hdp.hortonworks.com's password:
Last failed login: Mon Feb 13 06:10:06 UTC 2023 from 172.18.0.2 on ssh:notty
There was 1 failed login attempt since the last successful login.
Last login: Mon Feb 13 05:48:58 2023 from 172.18.0.3
[root@sandbox-hdp ~]# hdfs dfs -checksum /root/geolocation.csv
checksum: /root/geolocation.csv: No such file or directory
[root@sandbox-hdp ~]# hdfs dfs -checksum /user/root
checksum: /user/root: Is a directory
[root@sandbox-hdp ~]# cd /root
[root@sandbox-hdp ~]# hdfs dfs -touchz /user/root/cobalagi
[root@sandbox-hdp ~]# hdfs dfs -du /user/root/
0      0      /user/root/.Trash
0      0      /user/root/cobalagi
526677 526677 /user/root/geolocation.csv
[root@sandbox-hdp ~]#
```

[Hadoop Commands | Learn Top 23 Useful Hadoop Commands \(educba.com\)](https://educba.com/hadoop-commands/)