

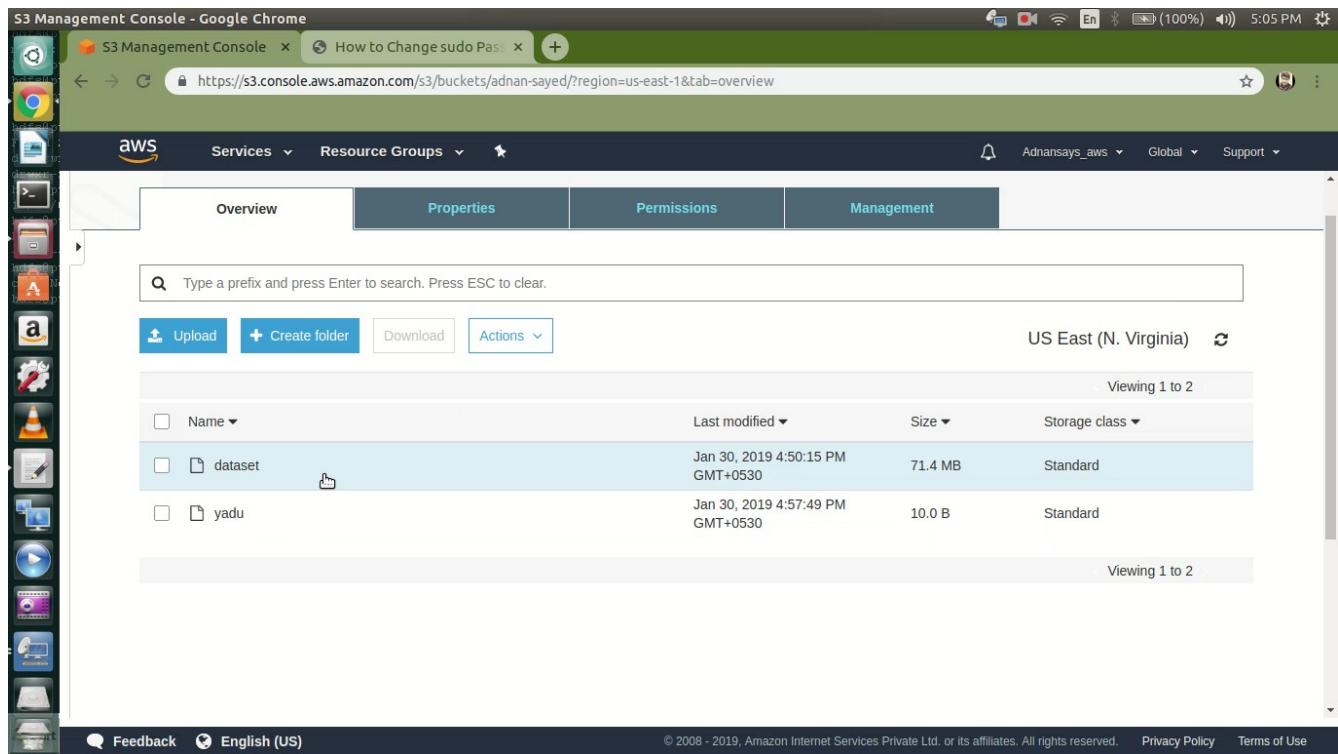
Data Pipeline: Configure core-site.xml by adding Access key and secret Access key

```
cloudage@pune: /etc/hadoop/conf
xml version="1.0" encoding="UTF-8">
<!-- autogenerated by Cloudera Manager -->
<configuration>
<property>
<name>fs.defaultFS</name>
<value>hdfs://ip-10-0-1-102.ec2.internal:8020</value>
</property>
<property>
<name>fs.s3n.awsAccessKeyId</name>
<value>AKIAJHM7742SH7GCRLGVA</value>
</property>
<property>
<name>fs.s3n.awsSecretAccessKey</name>
<value>W8iKCoasI9b5kU/akIvmkoUzDokQ52bVwevfCpC</value>
</property>
<property>
<name>fs.trash.interval</name>
<value>1</value>
</property>
<property>
<name>io.compression.codecs</name>
<value>org.apache.hadoop.io.compress.DefaultCodec,org.apache.hadoop.io.compress.GzipCodec,org.apache.hadoop.io.compress.BZip2Codec,org.apache.hadoop.io.compress.DeflateCodec,org.apache.hadoop.io.compress.SnappyCodec,org.apache.hadoop.io.compress.Lz4Codec</value>
</property>
<property>
<name>hadoop.security.authentication</name>
<value>simple</value>
</property>
<property>
<name>hadoop.security.authorization</name>
<value>false</value>
</property>
<property>
<name>hadoop.rpc.protection</name>
<value>authentication</value>
</property>
<property>
<name>hadoop.security.auth_to_local</name>
<value>DEFAULT</value>
</property>
:wq!
```

Refresh Nodes (make sure you drop to hdfs user)

```
hdfe@pune:~$ sudo vi /usr/local/hadoop/etc/hadoop/core-site.xml
sudo: unable to resolve host pune
[1]+  Stopped                  sudo vi /usr/local/hadoop/etc/hadoop/core-site.xml
cloudage@pune:~$ cd /etc/hadoop/
hadoop/      hadoop-htdfs/   hadoop-kms/    hbase/        hbase-solr/    hive/        hive-hcatalog/  hive-webhcatt/  hue/
cloudage@pune:~$ cd /etc/hadoop/conf
cloudage@pune:~/etc/hadoop/conf$ ls
__cloudera_generation__  core-site.xml  hdfs-site.xml  mapred-site.xml  topology.map  yarn-site.xml
__cloudera_metadata__   hadoop-env.sh  log4j.properties  ssi-client.xml  topology.py
cloudage@pune:~/etc/hadoop/conf$ sudo vi core-site.xml
sudo: unable to resolve host pune
cloudage@pune:~/etc/hadoop/conf$ cd
cloudage@pune:~$ hdfs dfsadmin -refreshNodes
refreshNodes: Access denied for user cloudage. Superuser privilege is required
cloudage@pune:~$ su hdfs
Password:
hdfe@pune:/home/cloudage$ cd
hdfe@pune:~$ hdfs dfsadmin -refreshNodes
Refresh nodes successful
hdfe@pune:~$
```

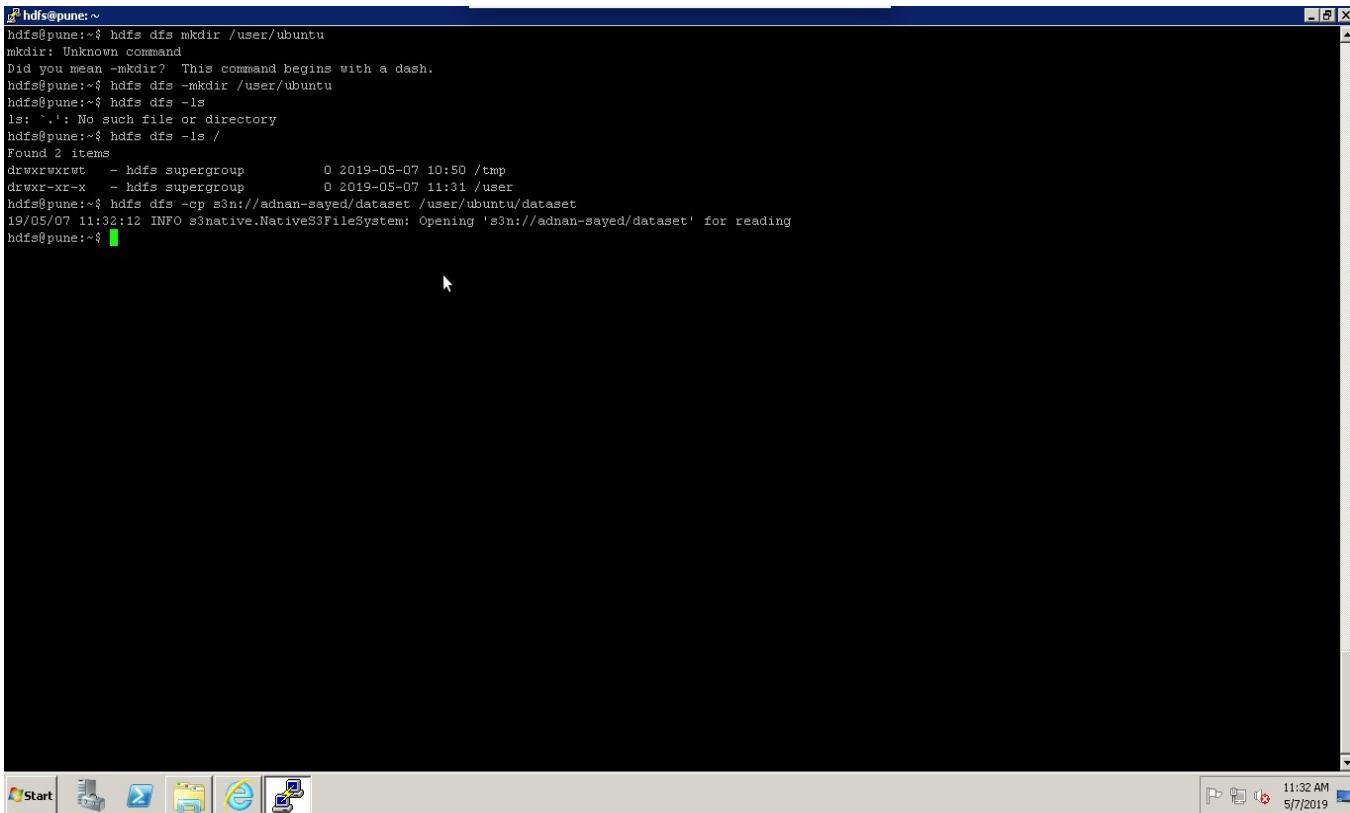
Some dataset on s3 to copy on our cluster



The screenshot shows the AWS S3 Management Console. The URL in the address bar is <https://s3.console.aws.amazon.com/s3/buckets/adnan-sayed?region=us-east-1&tab=overview>. The page displays an 'Overview' tab with two objects listed:

Name	Last modified	Size	Storage class
dataset	Jan 30, 2019 4:50:15 PM GMT+0530	71.4 MB	Standard
yadu	Jan 30, 2019 4:57:49 PM GMT+0530	10.0 B	Standard

Copying dataset from s3 to cluster



```
hdfs@pune:~$ hdfs dfs mkdir /user/ubuntu
mkdir: Unknown command
Did you mean -mkdir? This command begins with a dash.
hdfs@pune:~$ hdfs dfs -mkdir /user/ubuntu
hdfs@pune:~$ hdfs dfs -ls
ls: .': No such file or directory
hdfs@pune:~$ hdfs dfs -ls /
Found 2 items
drwxrwxrwt - hdfs supergroup          0 2019-05-07 10:50 /tmp
drwxr-xr-x - hdfs supergroup          0 2019-05-07 11:31 /user
hdfs@pune:~$ hdfs dfs -cp s3n://adnan-sayed/dataset /user/ubuntu/dataset
19/05/07 11:32:12 INFO s3native.NativeS3FileSystem: Opening 's3n://adnan-sayed/dataset' for reading
hdfs@pune:~$
```

```
hdfs@pune:~$ hdfs dfs mkdir /user/ubuntu
mkdir: Unknown command
Did you mean -mkdir? This command begins with a dash.
hdfs@pune:~$ hdfs dfs -mkdir /user/ubuntu
hdfs@pune:~$ hdfs dfs -ls
ls: .': No such file or directory
hdfs@pune:~$ hdfs dfs -ls /
Found 2 items
drwxrwxrwt  - hdfs supergroup          0 2019-05-07 10:50 /tmp
drwxr-xr-x  - hdfs supergroup          0 2019-05-07 11:31 /user
hdfs@pune:~$ hdfs dfs -cp s3n://adnan-sayed/dataset /user/ubuntu/dataset
19/05/07 11:32:12 INFO s3native.NativeS3FileSystem: Opening 's3n://adnan-sayed/dataset' for reading
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--r--  3 hdfs supergroup  74845419 2019-05-07 11:32 /user/ubuntu/dataset
hdfs@pune:~$
```

```
copying dataset from hdfs to s3 (name changed to Newdataset)
hdfs@pune:~$ hdfs dfs mkdir /user/ubuntu
mkdir: Unknown command
Did you mean -mkdir? This command begins with a dash.
hdfs@pune:~$ hdfs dfs -mkdir /user/ubuntu
hdfs@pune:~$ hdfs dfs -ls
ls: .': No such file or directory
hdfs@pune:~$ hdfs dfs -ls /
Found 2 items
drwxrwxrwt  - hdfs supergroup          0 2019-05-07 10:50 /tmp
drwxr-xr-x  - hdfs supergroup          0 2019-05-07 11:31 /user
hdfs@pune:~$ hdfs dfs -cp s3n://adnan-sayed/dataset /user/ubuntu/dataset
19/05/07 11:32:12 INFO s3native.NativeS3FileSystem: Opening 's3n://adnan-sayed/dataset' for reading
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--r--  3 hdfs supergroup  74845419 2019-05-07 11:32 /user/ubuntu/dataset
hdfs@pune:~$ hdfs dfs -cp NewData s3n://adnan-sayed/
cp: 'NewData': No such file or directory
hdfs@pune:~$ hdfs dfs -cp /user/ubuntu/dataset s3n://adnan-sayed/NewDataSet
19/05/07 11:35:40 INFO s3native.NativeS3FileSystem: OutputStream for key 'NewDataSet._COPYING_' writing to tempfile '/tmp/hadoop-hdfs/s3/output-3414653515873043598.tmp'
19/05/07 11:35:41 INFO s3native.NativeS3FileSystem: OutputStream for key 'NewDataSet._COPYING_' closed. Now beginning upload
19/05/07 11:35:44 INFO s3native.NativeS3FileSystem: OutputStream for key 'NewDataSet._COPYING_' upload complete
hdfs@pune:~$
```

Newdataset on s3

The screenshot shows the AWS S3 Management Console. The URL in the address bar is <https://s3.console.aws.amazon.com/s3/buckets/adnan-sayed/?region=us-east-1&tab=overview>. The main interface displays the 'Overview' tab for the bucket 'adnan-sayed'. The bucket contains three objects: 'NewDataSet' (51.4 MB, last modified May 7, 2019), 'dataset' (71.4 MB, last modified Jan 30, 2019), and 'yadu' (10.0 B, last modified Jan 30, 2019). The interface includes a search bar, upload/download buttons, and an 'Actions' dropdown.

Trash : By default trash is enabled in cloudera , configurations can be changed from this screen

The screenshot shows the Cloudera Manager HDFS configuration page. The URL is <http://10.0.1.80:7180/cmfservices/14/config#filterfree>. The 'Clusters' menu is open, and the 'HDFS' cluster is selected. The 'Configuration' tab is active. A search bar at the top has 'trash' typed into it. On the left, there's a 'Filters' sidebar with sections for 'SCOPE' and 'CATEGORY'. The 'SCOPE' section lists various HDFS components with their counts: HDFS (Service-Wide) 1, Balancer 0, DataNode 0, Gateway 1, HttpFS 0, JournalNode 0, NFS Gateway 0, NameNode 2, SecondaryNameNode 0, and Failover Controller 0. The 'CATEGORY' section lists Advanced 0 and Checkpointing 0. The main configuration area shows settings for 'Use Trash' (checkbox is off, note: HDFS Trash is disabled. Deleted files will be unrecoverable. [Suppress...](#)), 'Filesystem Trash Interval' (set to 1 day), 'Filesystem Trash Checkpoint Interval' (set to 1 hour), and an 'Audit Event Filter' set to 'HDFS (Service-Wide)'. A 'Save Changes' button is at the bottom right.

Deleting dataset (trash enabled)

```
hdfe@pune:~$ hdfs dfs -mkdir /user/ubuntu/
hdfe@pune:~$ ls
hdfe@pune:~$ su clouddage
Password:
clouddage@pune:/var/lib/hadoop-hdfs$ cd
clouddage@pune:~$ pwd
/home/clouddage
clouddage@pune:~$ su hdfe
Password:
hdfe@pune:/home/clouddage$ cd
hdfe@pune:~$ hdfs dfs -copyFromLocal /home/clouddage/data.txt /user/ubuntu/
hdfe@pune:~$ hdfs dfs -ls
ls: `': No such file or directory
hdfe@pune:~$ hdfs dfs -ls /
Found 2 items
drwxrwxrwt  - hdfe supergroup          0 2019-05-08 06:35 /tmp
drwxr-xr-x  - hdfe supergroup          0 2019-05-08 07:01 /user
hdfe@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--r--  3 hdfe supergroup         72 2019-05-08 07:04 /user/ubuntu/data.txt
hdfe@pune:~$ hdfs dfs -rm /user/ubuntu/data.txt
19/05/08 07:06:54 INFO fs.TrashPolicyDefault: Moved: 'hdfs://ip-10-0-1-115.ec2.internal:8020/user/ubuntu/data.txt' to trash at: hdfs://ip-10-0-1-115.ec2.internal:8020/user/hdfs/.Trash/Current/user/ubuntu/data.txt
hdfe@pune:~$
```

```
hdfe@pune:~$ hdfs dfs -mkdir /user/ubuntu/
hdfe@pune:~$ ls
hdfe@pune:~$ su clouddage
Password:
clouddage@pune:/var/lib/hadoop-hdfs$ cd
clouddage@pune:~$ pwd
/home/clouddage
clouddage@pune:~$ su hdfe
Password:
hdfe@pune:/home/clouddage$ cd
hdfe@pune:~$ hdfs dfs -copyFromLocal /home/clouddage/data.txt /user/ubuntu/
hdfe@pune:~$ hdfs dfs -ls
ls: `': No such file or directory
hdfe@pune:~$ hdfs dfs -ls /
Found 2 items
drwxrwxrwt  - hdfe supergroup          0 2019-05-08 06:35 /tmp
drwxr-xr-x  - hdfe supergroup          0 2019-05-08 07:01 /user
hdfe@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--r--  3 hdfe supergroup         72 2019-05-08 07:04 /user/ubuntu/data.txt
hdfe@pune:~$ hdfs dfs -rm /user/ubuntu/data.txt
19/05/08 07:06:54 INFO fs.TrashPolicyDefault: Moved: 'hdfs://ip-10-0-1-115.ec2.internal:8020/user/ubuntu/data.txt' to trash at: hdfs://ip-10-0-1-115.ec2.internal:8020/user/hdfs/.Trash/Current/user/ubuntu/data.txt
hdfe@pune:~$ hdfs dfs -ls /user/ubuntu/
hdfe@pune:~$
```

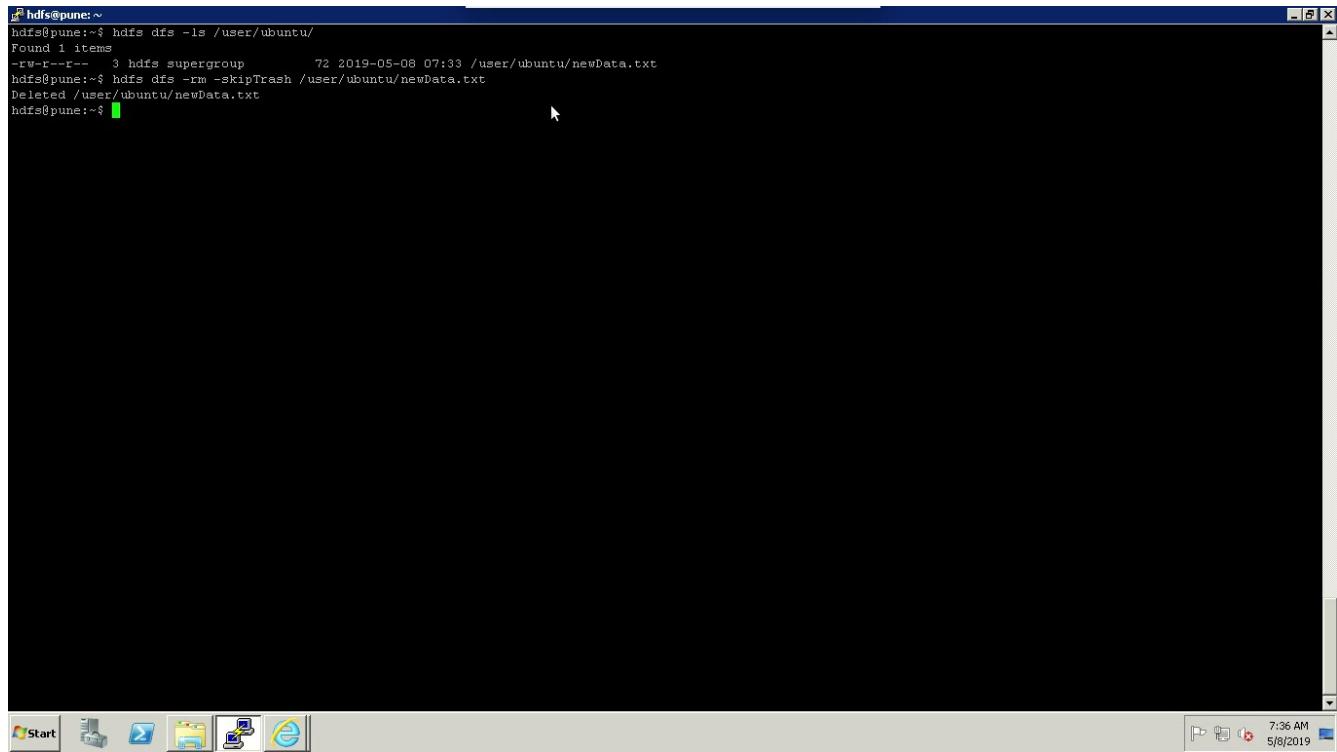
backup available even after deleting

```
hdfe@pune:~$ hdfs dfs -mkdir /user/ubuntu/
hdfe@pune:~$ ls
hdfe@pune:~$ su cloudage
Password:
cloudage@pune:~/var/lib/hadoop-hdfs$ cd
cloudage@pune:~$ pwd
/home/cloudage
cloudage@pune:~$ su hdfs
Password:
hdfe@pune:/home/cloudage$ cd
hdfe@pune:~$ hdfs dfs -copyFromLocal /home/cloudage/data.txt /user/ubuntu/
hdfe@pune:~$ hdfs dfs -ls
ls: './': No such file or directory
hdfe@pune:~$ hdfs dfs -ls /
Found 2 items
drwxrwxrwt  - hdfs supergroup          0 2019-05-08 06:35 /tmp
drwxr-xr-x  - hdfs supergroup          0 2019-05-08 07:01 /user
hdfe@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--r--  3 hdfs supergroup         72 2019-05-08 07:04 /user/ubuntu/data.txt
hdfe@pune:~$ hdfs dfs -rm /user/ubuntu/data.txt
19/05/08 07:06:54 INFO fs.TrashPolicyDefault: Moved: 'hdfs://ip-10-0-1-115.ec2.internal:8020/user/ubuntu/data.txt' to trash at: hdfs://ip-10-0-1-115.ec2.internal:8020/user/hdfs/.Trash/Current/user/ubuntu/data.txt
hdfe@pune:~$ hdfs dfs -ls /user/ubuntu/
hdfe@pune:~$ hdfs dfs -ls /user/ubuntu/.Trash/
ls: '/user/ubuntu/.Trash/': No such file or directory
hdfe@pune:~$ hdfs dfs -ls /user/hdfs/.Trash
Found 1 items
drwx-----  - hdfs supergroup         0 2019-05-08 07:06 /user/hdfs/.Trash/Current
hdfe@pune:~$ hdfs dfs -ls /user/hdfs/.Trash/Current/user/ubuntu/
Found 1 items
-rw-r--r--  3 hdfs supergroup         72 2019-05-08 07:04 /user/hdfs/.Trash/Current/user/ubuntu/data.txt
hdfe@pune:~$
```

Restoring dataset from trash and renaming it to newData

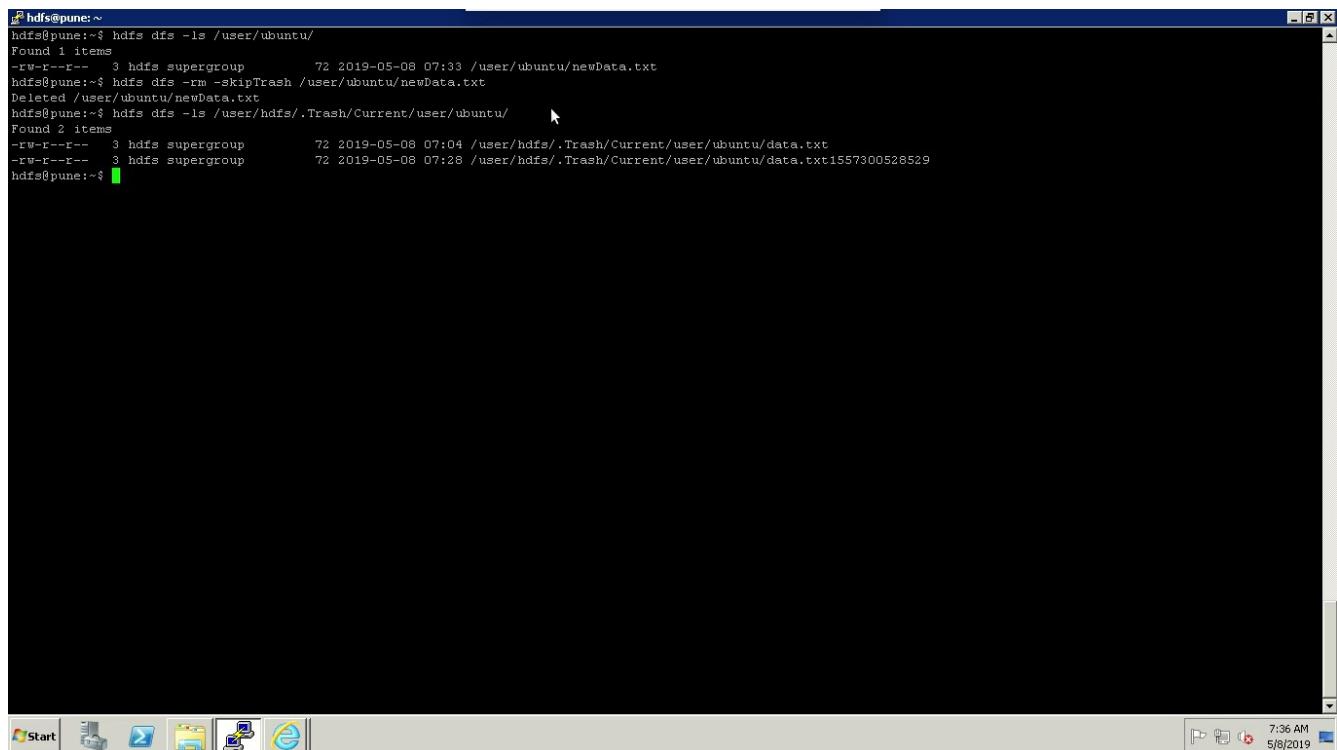
```
hdfe@pune:~$ hdfs dfs -rm /user/ubuntu/data.txt
19/05/08 07:28:48 INFO fs.TrashPolicyDefault: Moved: 'hdfs://ip-10-0-1-115.ec2.internal:8020/user/ubuntu/data.txt' to trash at: hdfs://ip-10-0-1-115.ec2.internal:8020/user/hdfs/.Trash/Current/user/ubuntu/data.txt1557300528529
hdfe@pune:~$ hdfs dfs -ls /user/hdfs/.Trash/Current/user/ubuntu/
Found 2 items
-rw-r--r--  3 hdfs supergroup         72 2019-05-08 07:04 /user/hdfs/.Trash/Current/user/ubuntu/data.txt
-rw-r--r--  3 hdfs supergroup         72 2019-05-08 07:28 /user/hdfs/.Trash/Current/user/ubuntu/data.txt1557300528529
hdfe@pune:~$ hdfs dfs -cp /user/hdfs/.Trash/Current/user/ubuntu/data.txt /user/ubuntu/
hdfe@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--r--  3 hdfs supergroup         72 2019-05-08 07:29 /user/ubuntu/data.txt
hdfe@pune:~$ hdfs dfs -rm -skipTrash /user/ubuntu/
rm: '/user/ubuntu': Is a directory
hdfe@pune:~$ hdfs dfs -rm -skipTrash /user/ubuntu/data.txt
Deleted /user/ubuntu/data.txt
hdfe@pune:~$ hdfs dfs -ls /user/hdfs/.Trash/Current/user/ubuntu/
Found 2 items
-rw-r--r--  3 hdfs supergroup         72 2019-05-08 07:04 /user/hdfs/.Trash/Current/user/ubuntu/data.txt
-rw-r--r--  3 hdfs supergroup         72 2019-05-08 07:28 /user/hdfs/.Trash/Current/user/ubuntu/data.txt1557300528529
hdfe@pune:~$ hdfs dfs -cp /user/hdfs/.Trash/Current/user/ubuntu/data.txt /user/ubuntu/newData.txt
hdfe@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--r--  3 hdfs supergroup         72 2019-05-08 07:33 /user/ubuntu/newData.txt
hdfe@pune:~$
```

-skipTrash will delete the data permanently (newData.txt)



```
hdfe@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--r-- 3 hdfs supergroup      72 2019-05-08 07:33 /user/ubuntu/newData.txt
hdfe@pune:~$ hdfs dfs -rm -skipTrash /user/ubuntu/newData.txt
Deleted /user/ubuntu/newData.txt
hdfe@pune:~$
```

newData is not present in Trash after -skipTrash



```
hdfe@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--r-- 3 hdfs supergroup      72 2019-05-08 07:33 /user/ubuntu/newData.txt
hdfe@pune:~$ hdfs dfs -rm -skipTrash /user/ubuntu/newData.txt
Deleted /user/ubuntu/newData.txt
hdfe@pune:~$ hdfs dfs -ls /user/hdfs/.Trash/Current/user/ubuntu/
Found 2 items
-rw-r--r-- 3 hdfs supergroup      72 2019-05-08 07:04 /user/hdfs/.Trash/Current/user/ubuntu/data.txt
-rw-r--r-- 3 hdfs supergroup      72 2019-05-08 07:28 /user/hdfs/.Trash/Current/user/ubuntu/data.txt1557300528529
hdfe@pune:~$
```

expung will empty the Trash

```

hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--k-- 3 hdfs supergroup          72 2019-05-08 07:33 /user/ubuntu/newData.txt
hdfs@pune:~$ hdfs dfs -rm -skipTrash /user/ubuntu/newData.txt
Deleted /user/ubuntu/newData.txt
hdfs@pune:~$ hdfs dfs -ls /user/hdfs/.Trash/Current/user/ubuntu/
Found 2 items
-rw-r--k-- 3 hdfs supergroup      72 2019-05-08 07:04 /user/hdfs/.Trash/Current/user/ubuntu/data.txt
-rw-r--k-- 3 hdfs supergroup      72 2019-05-08 07:28 /user/hdfs/.Trash/Current/user/ubuntu/data.txt1557300528529
hdfs@pune:~$ hdfs dfs expunge
expunge: Unknown command
Did you mean -expunge? This command begins with a dash.
hdfs@pune:~$ hdfs dfs -expunge
19/05/08 07:36:58 INFO fs.TrashPolicyDefault: TrashPolicyDefault#deleteCheckpoint for trashRoot: hdfs://ip-10-0-1-115.ec2.internal:8020/user/hdfs/.Trash
19/05/08 07:36:58 INFO fs.TrashPolicyDefault: TrashPolicyDefault#deleteCheckpoint for trashRoot: hdfs://ip-10-0-1-115.ec2.internal:8020/user/hdfs/.Trash
19/05/08 07:36:58 INFO fs.TrashPolicyDefault: TrashPolicyDefault#createCheckpoint for trashRoot: hdfs://ip-10-0-1-115.ec2.internal:8020/user/hdfs/.Trash
19/05/08 07:36:58 INFO fs.TrashPolicyDefault: Created trash checkpoint: /user/hdfs/.Trash/190508073658
hdfs@pune:~$ 

```

The terminal window shows a series of HDFS commands being run. It starts with listing files in the /user/ubuntu directory, then deleting a file named newData.txt using the -rm -skipTrash option. It then lists the contents of the .Trash/Current directory. A command to expunge files is entered but fails because it's not a recognized command. Finally, it runs several log entries related to the TrashPolicyDefault class, which handle the creation and deletion of trash checkpoints.

Block Size: block size changed from 128 to 256 (need to restart hdfs)

HDFS - Cloudera Manager - Internet Explorer

cloudera MANAGER Clusters Hosts Diagnostics Audits Charts Backup Administration

HDFS (Cluster 1) Actions

Status Instances Configuration Commands File Browser Charts Library Cache Statistics Audits NameNode

Configuration changes have been saved successfully.

blocksiz

Show 1 Suppressed Warning(s) Show All Descriptions

Filters

SCOPE

- HDFS (Service-Wide) 1
- Balancer 0
- DataNode 0
- Gateway 0
- HttpFS 0
- JournalNode 0
- NFS Gateway 0
- NameNode 0
- SecondaryNameNode 0
- Failover Controller 0

CATEGORY

- Advanced 0
- Checkpointing 0

HDFS Block Size HDFS (Service-Wide) 256 MiB

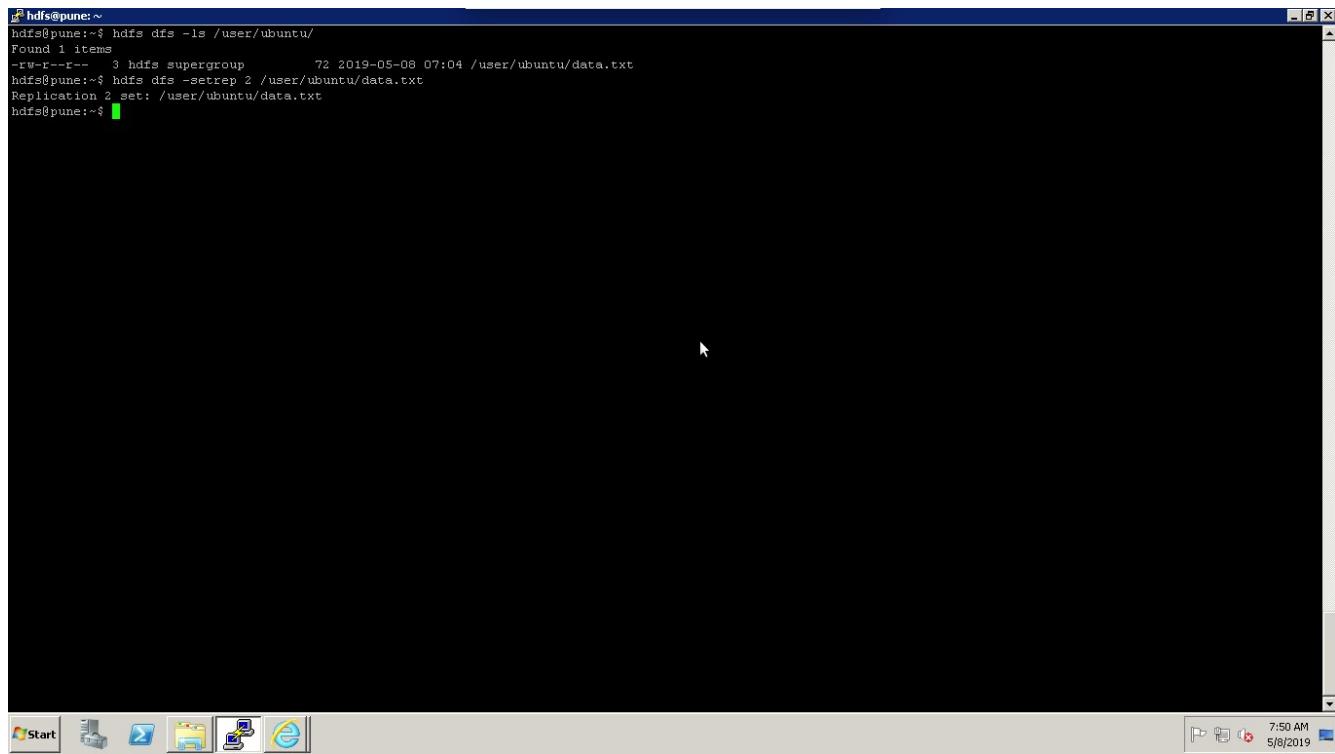
25 Per Page

Save Changes

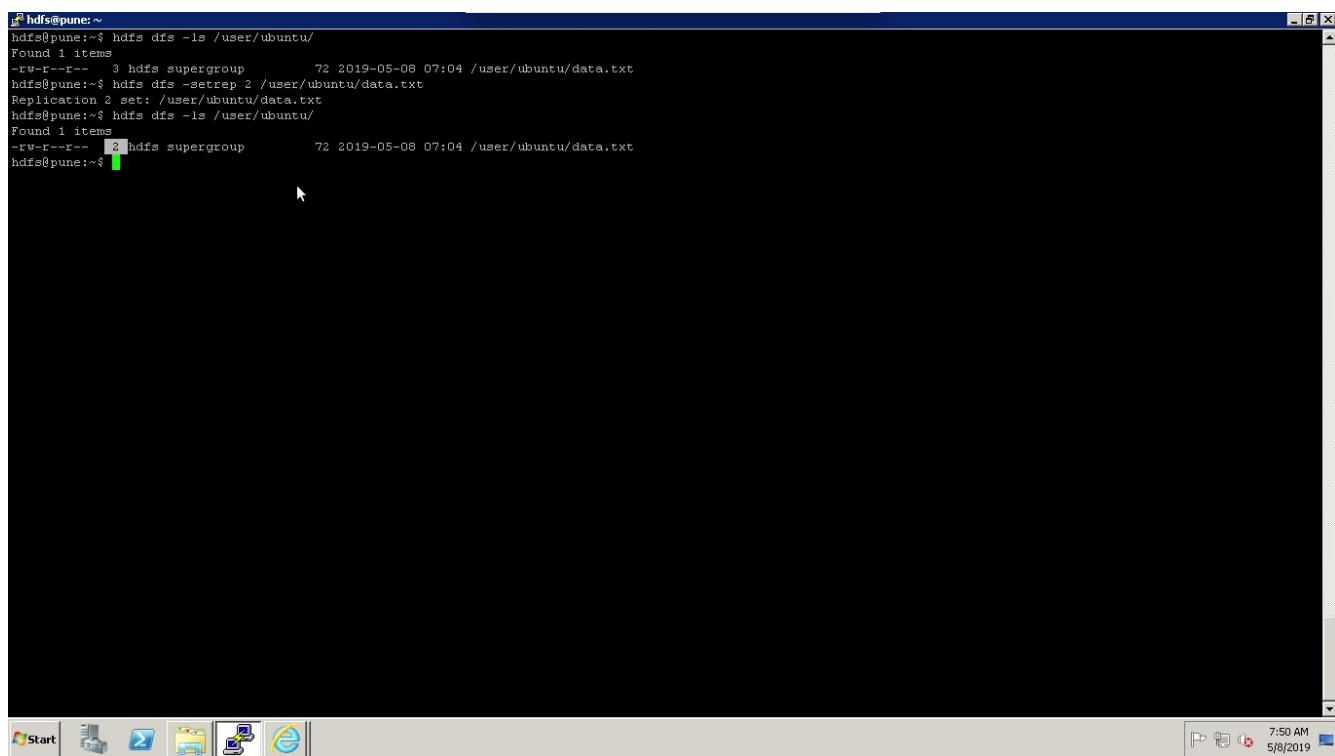
7:41 AM 5/8/2019

This screenshot shows the Cloudera Manager interface for managing the HDFS service. The 'Configuration' tab is selected. A success message indicates that configuration changes have been saved. In the main area, the 'Configuration' section is open, showing the 'HDFS Block Size' setting has been updated from its previous value of 128 MiB to 256 MiB. The 'Blocks' filter is applied, showing one suppressed warning. The bottom status bar shows the current time as 7:41 AM on 5/8/2019.

Setting Replication:



```
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--r-- 3 hdfs supergroup 72 2019-05-08 07:04 /user/ubuntu/data.txt
hdfs@pune:~$ hdfs dfs -setrep 2 /user/ubuntu/data.txt
Replication 2 set: /user/ubuntu/data.txt
hdfs@pune:~$
```



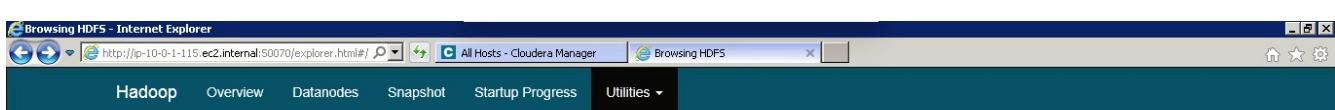
```
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--r-- 3 hdfs supergroup 72 2019-05-08 07:04 /user/ubuntu/data.txt
hdfs@pune:~$ hdfs dfs -setrep 2 /user/ubuntu/data.txt
Replication 2 set: /user/ubuntu/data.txt
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--r-- 2 hdfs supergroup 72 2019-05-08 07:04 /user/ubuntu/data.txt
hdfs@pune:~$
```

MetaSave:

```

hdfs@pune:~$ hdfs dfsadmin -metasave metasave-report.txt
Created metasave file metasave-report.txt in the log directory of namenode hdfs://ip-10-0-1-115.ec2.internal:8020
hdfs@pune:~$ 

```



Browsing HDFS - Internet Explorer

http://ip-10-0-1-115.ec2.internal:50070/explorer.html#/ OK All Hosts - Cloudera Manager Browsing HDFS

7:55 AM 5/8/2019

Hadoop Overview Datanodes Snapshot Startup Progress Utilities ▾

Utilities ▾

Browse the file system

Logs

Browse Directory

/user

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	hdfs	supergroup	0 B	Wed May 08 07:26:45 +0000 2019	0	0 B	abc
drwx-----	hdfs	supergroup	0 B	Wed May 08 07:06:54 +0000 2019	0	0 B	hdfs
drwxrwxrwx	mapred	hadoop	0 B	Wed May 08 06:34:42 +0000 2019	0	0 B	history
drwxrwxr-t	hive	hive	0 B	Wed May 08 06:35:33 +0000 2019	0	0 B	hive
drwxrwxr-x	hue	hue	0 B	Wed May 08 06:36:06 +0000 2019	0	0 B	hue
drwxrwxr-x	oozie	oozie	0 B	Wed May 08 06:36:18 +0000 2019	0	0 B	oozie
drwxr-xr-x	hdfs	supergroup	0 B	Wed May 08 07:48:25 +0000 2019	0	0 B	ubuntu

Hadoop, 2018.



Directory: /logs/ - Internet Explorer					
SecurityAuth-hdfs.audit					
audit/					0 bytes May 8, 2019 6:33:42 AM
hadoop-cmf-hdfs-NAMENODE-ip-10-0-1-115.ec2.internal.log.out					4096 bytes May 8, 2019 6:33:41 AM
hadoop-cmf-hdfs-SECONDARYNAMENODE-ip-10-0-1-115.ec2.internal.log.out					2156233 bytes May 8, 2019 8:10:04 AM
hdfs-audit.log					179167 bytes May 8, 2019 7:43:37 AM
metasave-report.txt					1162928 bytes May 8, 2019 8:10:08 AM
stacks/					123764 bytes May 8, 2019 7:55:41 AM
					4096 bytes May 8, 2019 6:33:41 AM

http://ip-10-0-1-115.ec2.internal:50070/logs/hadoop-cmf-hdfs-NAMENODE-ip-10-0-1-115.ec2.internal.log.out - Internet Explorer					
Start	File	Edit	View	Favorites	8:10 AM 5/8/2019

http://ip-10-0-1-115.ec2.internal:50070/logs/metasave-report.txt - Internet Explorer					
890 files and directories, 747 blocks = 1637 total					
Live DataNodes:	2				
Dead DataNodes:	1				
Metasave: Blocks waiting for replication: 747					
/user/oozie/share/lib/lib_20190508063616/hive/libthrift-0.9.3.jar: blk_1073741827_1003 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/jpm-1.1.jar: blk_1073741828_1004 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/asm-tree-3.1.jar: blk_1073741828_1005 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/mail-1.4.jar: blk_1073741830_1006 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/jackson-databind-2.1.3-cloudera.1.jar: blk_1073741831_1007 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/xz-1.0.jar: blk_1073741832_1008 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/jettison-1.1.jar: blk_1073741833_1009 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/hive-shims-23.jar: blk_1073741834_1010 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/jersey-server-1.9.jar: blk_1073741835_1011 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/datanucleus-api-jdo-3.2.1.jar: blk_1073741836_1012 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/hbase-protocol.jar: blk_1073741837_1013 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/metastore.jar: blk_1073741838_1014 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/asm-commons-3.1.jar: blk_1073741838_1015 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/hive-common.jar: blk_1073741840_1016 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/jaxb-api-2.2.2.jar: blk_1073741841_1017 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/hive-contrib.jar: blk_1073741842_1018 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/jackson-core-asl-1.8.10.jar: blk_1073741843_1019 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/boncpc-0.7.1.RELEASE.jar: blk_1073741844_1020 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/conoms-lang3-3.3.2.jar: blk_1073741845_1021 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/jackson-annotations-2.2.3.jar: blk_1073741846_1022 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/commons-collections-3.2.2.jar: blk_1073741847_1023 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/eigenbase-properties-3.1.4.jar: blk_1073741848_1024 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/commons-beanutils-core-1.8.0.jar: blk_1073741849_1025 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/commons-pool-1.5.4.jar: blk_1073741850_1026 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/stax-api-1.0.2.jar: blk_1073741851_1027 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/jackson-jaxrs-1.8.10.jar: blk_1073741852_1028 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/slf4j-log4j12-1.7.8.jar: blk_1073741853_1029 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/jackson-xc-1.8.10.jar: blk_1073741854_1030 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/curator-framework-2.7.1.jar: blk_1073741855_1031 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/jaxb-impl-2.2.3-1.jar: blk_1073741856_1032 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/curator-recipes-2.7.1.jar: blk_1073741857_1033 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/ST4-4.0.4.jar: blk_1073741858_1034 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/hadoop-common.jar: blk_1073741859_1035 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/commons-dhcp-1.4.jar: blk_1073741860_1036 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/libfb303-0.9.3.jar: blk_1073741861_1037 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/jersey-servlet-1.14.jar: blk_1073741862_1038 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/commons-net-3.1.jar: blk_1073741863_1039 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/httpcore-4.2.5.jar: blk_1073741864_1040 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					
/user/oozie/share/lib/lib_20190508063616/hive/activation-1.1.jar: blk_1073741865_1041 (replicas: 1: 2 d: 0 c: 0 e: 0) 10.0.1.162:50010 : 10.0.1.149:50010 :					

Decommissioning :

All Hosts - Cloudera Manager - Internet Explorer

http://10.0.1.80:7180/cmft/hardware/hosts

cloudera MANAGER

Clusters Hosts Diagnostics Audits Charts Backup Administration

Search Support admin

All Hosts

Configuration Add New Hosts to Cluster Re-run Upgrade Wizard Inspect All Hosts

Search

Actions for Selected (1) ▾

Columns: 10 Selected ▾

Status	Name	IP	Roles	Commission State	Last Heartbeat	Load Average	Disk Usage	Physical Memory	Swap
<input type="checkbox"/>	ip-10-0-1-115.ec2.internal	10.0.1.115	> 12 Role(s)	Commissioned	2.5s ago	0.00 0.04 0.14	9.4 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB	
<input type="checkbox"/>	ip-10-0-1-149.ec2.internal	10.0.1.149	> 8 Role(s)	Commissioned	2.51s ago	0.48 0.15 0.11	10 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB	
<input type="checkbox"/>	ip-10-0-1-162.ec2.internal	10.0.1.162	> 4 Role(s)	Commissioned	2.78s ago	0.00 0.02 0.05	9.9 GiB / 29.4 GiB	1.2 GiB / 7.8 GiB	
<input checked="" type="checkbox"/>	ip-10-0-1-80.ec2.internal	10.0.1.80	> 4 Role(s)	Commissioned	2.81s ago	0.00 0.06 0.15	12 GiB / 29.4 GiB	2.5 GiB / 7.8 GiB	

Feedback

Start

7:59 AM 5/8/2019

All Hosts - Cloudera Manager - Internet Explorer

http://10.0.1.80:7180/cmft/hardware/hosts

cloudera MANAGER

Clusters Hosts Diagnostics Audits Charts Backup Administration

Search Support admin

All Hosts

Configuration Add New Hosts to Cluster Re-run Upgrade Wizard Inspect All Hosts

Search

Actions for Selected (1) ▾

Commission State Last Heartbeat Load Average Disk Usage Physical Memory Swap

Commissioned	2.5s ago	0.00 0.04 0.14	9.4 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB	
Commissioned	2.51s ago	0.48 0.15 0.11	10 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB	
Commissioned	2.78s ago	0.00 0.02 0.05	9.9 GiB / 29.4 GiB	1.2 GiB / 7.8 GiB	
Commissioned	2.81s ago	0.00 0.06 0.15	12 GiB / 29.4 GiB	2.5 GiB / 7.8 GiB	

Actions for Selected (1) ▾

Assign Rack

Regenerate Keytab

Apply Host Template

Stop Roles on Hosts

Begin Maintenance (Suppress Alerts/Decommission)

End Maintenance (Enable Alerts/Recommission)

Remove From Cluster

Remove From Cloudera Manager

Feedback

Start

8:00 AM 5/8/2019

All Hosts - Cloudera Manager - Internet Explorer

cloudera MANAGER Clusters Hosts Diagnostics Audits Charts Backup Administration Search Support admin

Stop Roles on Hosts Command

Status **Finished** May 8, 8:00:09 AM 7ms
Successfully stopped all the roles on selected hosts.

Completed 4 of 4 step(s).

Show All Steps Show Only Failed Steps Show Running Steps

> Execute command Stop on service Hive	Hive	May 8, 8:00:09 AM	2ms
> Execute command Stop on service YARN (MR2 Included)	YARN (MR2 Included)	May 8, 8:00:09 AM	2ms
> Execute command Stop on service HDFS	HDFS	May 8, 8:00:09 AM	1ms
> Execute command Stop on service ZooKeeper	ZooKeeper	May 8, 8:00:09 AM	2ms

Close

Feedback

All Hosts - Cloudera Manager - Internet Explorer

cloudera MANAGER Clusters Hosts Diagnostics Audits Charts Backup Administration Search Support admin

All Hosts Configuration Add New Hosts to Cluster Re-run Upgrade Wizard Inspect All Hosts

Search

Filters

- STATUS**
 - Concerning Health 1
 - Good Health 3
- CLUSTERS**
- CORES**
- COMMISSION STATE**
- LAST HEARTBEAT**
- LOAD (1 MINUTE)**
- LOAD (5 MINUTES)**
- LOAD (15 MINUTES)**
- MAINTENANCE MODE**
- RACK**
- SERVICES**
- HEALTH TESTS**

Actions for Selected (1) ▾

- Assign Rack
- Regenerate Keytab
- Apply Host Template
- Start Roles on Hosts
- Stop Roles on Hosts
- Begin Maintenance (Suppress Alerts/Decommission)**
- End Maintenance (Enable Alerts/Recommission)
- Remove From Cluster
- Remove From Cloudera Manager

Columns: **10 Selected ▾**

Commission State	Last Heartbeat	Load Average	Disk Usage	Physical Memory
Commissioned	13.85s ago	0.00 0.03 0.13	9.4 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB
Commissioned	13.87s ago	0.05 0.10 0.10	10 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB
Commissioned	14.16s ago	0.00 0.01 0.05	9.9 GiB / 29.4 GiB	1.2 GiB / 7.8 GiB
Commissioned	14.19s ago	0.00 0.04 0.13	12 GiB / 29.4 GiB	2.5 GiB / 7.8 GiB

Feedback

All Hosts - Cloudera Manager - Internet Explorer

http://10.0.1.80:7180/cmft/hardware/hosts

cloudera MANAGER

Clusters Hosts Diagnostics Audits Charts Backup Administration

Search Support admin

All Hosts

Begin Maintenance (Suppress Alerts/Decommission)

Host	Role Instances
ip-10-0-1-80.ec2.internal	DN G NM S

SUPPRESS ALERTS

The selected host(s) will enter Maintenance Mode and all alerts will be suppressed.

Decommission Host(s)

All roles on the selected host(s) will be decommissioned (where applicable) and stopped.

For HDFS DataNodes on the selected host(s):

Decommission DataNodes

This action replicates all HDFS data to other hosts before stopping and could take significant time.

Cancel Begin Maintenance

Feedback

All Hosts - Cloudera Manager - Internet Explorer

http://10.0.1.80:7180/cmft/hardware/hosts

cloudera MANAGER

Clusters Hosts Diagnostics Audits Charts Backup Administration

Search Support admin

All Hosts

Configuration Add New Hosts to Cluster Re-run Upgrade Wizard Inspect All Hosts

Search

FILTERS

STATUS

- Concerning Health 1
- Good Health 3

CLUSTERS

CORES

COMMISSION STATE

LAST HEARTBEAT

LOAD (1 MINUTE)

LOAD (5 MINUTES)

LOAD (15 MINUTES)

MAINTENANCE MODE

RACK

SERVICES

HEALTH TESTS

ACTIONS FOR SELECTED

Columns: 10 Selected

Status	Name	IP	Roles	Commission State	Last Heartbeat	Load Average	Disk Usage	Physical Memory
<input checked="" type="checkbox"/>	ip-10-0-1-115.ec2.internal	10.0.1.115	12 Role(s)	Commissioned	3.56s ago	0.50 0.24 0.19	9.4 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB
<input checked="" type="checkbox"/>	ip-10-0-1-149.ec2.internal	10.0.1.149	8 Role(s)	Commissioned	6.72s ago	0.80 0.25 0.15	10 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB
<input checked="" type="checkbox"/>	ip-10-0-1-162.ec2.internal	10.0.1.162	4 Role(s)	Commissioned	7.09s ago	0.00 0.01 0.05	9.9 GiB / 29.4 GiB	1.2 GiB / 7.8 GiB
<input checked="" type="checkbox"/>	ip-10-0-1-80.ec2.internal	10.0.1.80	4 Role(s)	Decommissioned	7.12s ago	0.00 0.01 0.10	12 GiB / 29.4 GiB	2.5 GiB / 7.8 GiB

Feedback

All Hosts - Cloudera Manager - Internet Explorer

http://10.0.1.80:7180/cmft/hardware/hosts

cloudera MANAGER Clusters ▾ Hosts ▾ Diagnostics ▾ Audits Charts ▾ Backup ▾ Administration ▾ Search Support ▾ admin ▾

All Hosts Configuration Add New Hosts to Cluster Re-run Upgrade Wizard Inspect All Hosts

Search

Filters

STATUS

- Concerning Health 1
- Good Health 3

CLUSTERS

CORES

COMMISSION STATE

LAST HEARTBEAT

LOAD (1 MINUTE)

LOAD (5 MINUTES)

LOAD (15 MINUTES)

MAINTENANCE MODE

RACK

SERVICES

HEALTH TESTS

Actions for Selected (1) ▾

Commission State Last Heartbeat Load Average Disk Usage Physical Memory

(s)	Commissioned	3.56s ago	0.50 0.24 0.19	9.4 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB
(i)	Commissioned	6.72s ago	0.80 0.25 0.15	10 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB
(i)	Commissioned	7.09s ago	0.00 0.01 0.05	9.9 GiB / 29.4 GiB	1.2 GiB / 7.8 GiB
(i)	Decommissioned	7.12s ago	0.00 0.01 0.10	12 GiB / 29.4 GiB	2.5 GiB / 7.8 GiB

Assign Rack Regenerate Keytab Apply Host Template Start Roles on Hosts Stop Roles on Hosts Begin Maintenance (Suppress Alerts/Decommission) End Maintenance (Enable Alerts/Recommission)

Remove From Cluster Remove From Cloudera Manager

Columns: 10 Selected ▾

Feedback

Start

8:07 AM 5/8/2019

All Hosts - Cloudera Manager - Internet Explorer

http://10.0.1.80:7180/cmft/hardware/hosts

cloudera MANAGER Clusters ▾ Hosts ▾ Diagnostics ▾ Audits Charts ▾ Backup ▾ Administration ▾ Search Support ▾ admin ▾

All Hosts

Remove Hosts From Cluster

Removing these hosts will stop and delete all roles running on them and then remove them from their clusters. The hosts will still be managed by Cloudera Manager and can be utilized after being added to new or existing clusters.

Role data directories will not be deleted.

Host	Role Instances
ip-10-0-1-80.ec2.internal	DN G NM S

Decommission Roles (Warning: Removing the hosts without decommissioning the roles running on them can result in permanent data loss.)

Skip Management Roles

Cancel Confirm

Feedback

Start

8:07 AM 5/8/2019

All Hosts - Cloudera Manager - Internet Explorer

http://10.0.1.80:7180/cmft/hardware/hosts

cloudera MANAGER

Clusters Hosts Diagnostics Audits Charts Backup Administration

Search Support admin

All Hosts

Configuration Add New Hosts to Cluster Re-run Upgrade Wizard Inspect All Hosts

Search

Filters

Status

- Concerning Health 1
- Good Health 3

Clusters

Cores

Commission State

Last Heartbeat

Load (1 minute)

Load (5 minutes)

Load (15 minutes)

Maintenance Mode

Rack

Services

Health Tests

Actions for Selected (1)

Assign Rack

Regenerate Keytab

Apply Host Template

Start Roles on Hosts

Stop Roles on Hosts

Begin Maintenance (Suppress Alerts/Decommission)

End Maintenance (Enable Alerts/Recommission)

Remove From Cluster

Remove From Cloudera Manager

Columns: 10 Selected

	Commission State	Last Heartbeat	Load Average	Disk Usage	Physical Memory
(s)	Commissioned	13.18s ago	0.48 0.29 0.21	9.4 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB
(s)	Commissioned	2.83s ago	0.16 0.21 0.14	10 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB
(s)	Commissioned	3.23s ago	0.00 0.01 0.05	9.9 GiB / 29.4 GiB	1.2 GiB / 7.8 GiB
(s)	Commissioned	6.99s ago	0.15 0.05 0.10	12 GiB / 29.4 GiB	2.4 GiB / 7.8 GiB

Feedback

Start Taskbar 8:09 AM 5/8/2019

All Hosts - Cloudera Manager - Internet Explorer

http://10.0.1.80:7180/cmft/hardware/hosts

cloudera MANAGER

Clusters Hosts Diagnostics Audits Charts Backup Administration

Search Support admin

All Hosts

Remove Hosts From Cloudera Manager

1 host(s) can be removed.

Host	Role Instances
ip-10-0-1-80.ec2.internal	

Hosts are automatically added when Cloudera Manager agent communicates with Cloudera Manager. To remove these hosts, the agent needs to be stopped on these hosts first.

- Stop each Cloudera Manager Agent manually on each host.
\$ ssh username@hostname
\$ service cloudera-scm-agent stop
- Click Confirm to remove these hosts from Cloudera Manager.

Cancel Confirm

Columns: 10 Selected

	Physical Memory
29.4 GiB	2.9 GiB / 7.8 GiB
29.4 GiB	2.9 GiB / 7.8 GiB
29.4 GiB	1.2 GiB / 7.8 GiB
29.4 GiB	2.4 GiB / 7.8 GiB

Feedback

Start Taskbar 8:09 AM 5/8/2019

All Hosts - Cloudera Manager - Internet Explorer

http://10.0.1.80:7180/cmf/hardware/hosts

cloudera MANAGER

Clusters Hosts Diagnostics Audits Charts Backup Administration

Search Support admin

All Hosts

Configuration Add New Hosts to Cluster Re-run Upgrade Wizard Inspect All Hosts

Search

Actions for Selected

Columns: 10 Selected

Status	Name	IP	Roles	Commission State	Last Heartbeat	Load Average	Disk Usage	Physical Memory
<input type="checkbox"/>	ip-10-0-1-115	10.0.1.115	> 12 Role(s)	Commissioned	786ms ago	0.61 0.33 0.23	9.4 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB
<input type="checkbox"/>	ip-10-0-1-149	10.0.1.149	> 8 Role(s)	Commissioned	5.44s ago	0.20 0.21 0.14	10 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB
<input type="checkbox"/>	ip-10-0-1-162	10.0.1.162	> 4 Role(s)	Commissioned	5.85s ago	0.00 0.01 0.05	9.9 GiB / 29.4 GiB	1.2 GiB / 7.8 GiB

Start

8:09 AM 5/8/2019

Commissioning a Node:

All Hosts - Cloudera Manager - Internet Explorer

http://10.0.1.80:7180/cmf/hardware/hosts

cloudera MANAGER

Clusters Hosts Diagnostics Audits Charts Backup Administration

Search Support admin

All Hosts

Configuration Add New Hosts to Cluster Re-run Upgrade Wizard Inspect All Hosts

Search

Actions for Selected

Columns: 10 Selected

Status	Name	IP	Roles	Commission State	Last Heartbeat	Load Average	Disk Usage	Physical Memory
<input type="checkbox"/>	ip-10-0-1-115	10.0.1.115	> 12 Role(s)	Commissioned	8.37s ago	0.10 0.39 0.37	9.4 GiB / 29.4 GiB	2.9 GiB / 7.8 GiB
<input type="checkbox"/>	ip-10-0-1-149	10.0.1.149	> 8 Role(s)	Commissioned	8.35s ago	0.69 0.32 0.27	10 GiB / 29.4 GiB	3 GiB / 7.8 GiB
<input type="checkbox"/>	ip-10-0-1-162	10.0.1.162	> 4 Role(s)	Commissioned	8.46s ago	0.00 0.02 0.05	9.9 GiB / 29.4 GiB	1.2 GiB / 7.8 GiB

Start

8:32 AM 5/8/2019

final.mp4

Cloudera MANAGER

Specify hosts for your CDH cluster installation.

Hosts should be specified using the same hostname (FQDN) that they will identify themselves with. Cloudera recommends including Cloudera Manager Server's host. This also enables health monitoring for that host.

Hint: Search for hostnames and IP addresses using patterns 

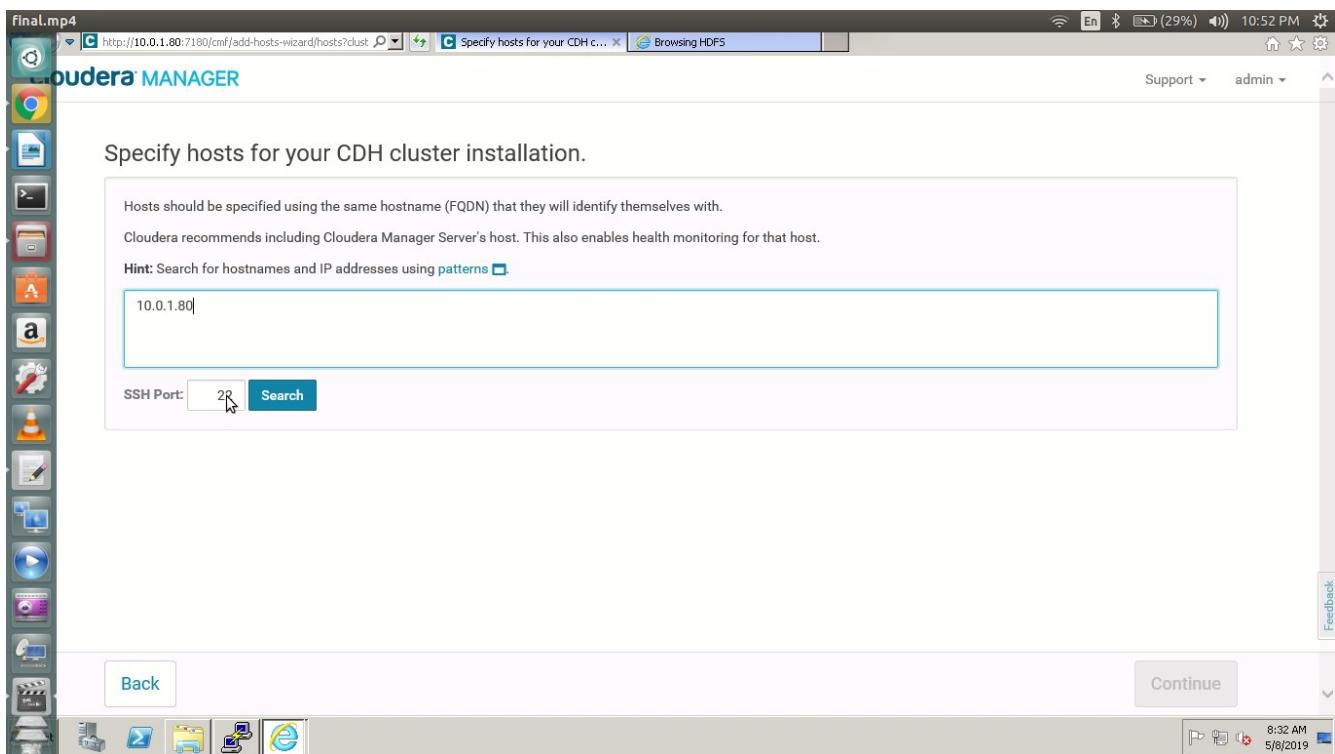
10.0.1.80

SSH Port:

Feedback

Back Continue

8:32 AM 5/8/2019



final.mp4

Browsing HDFS

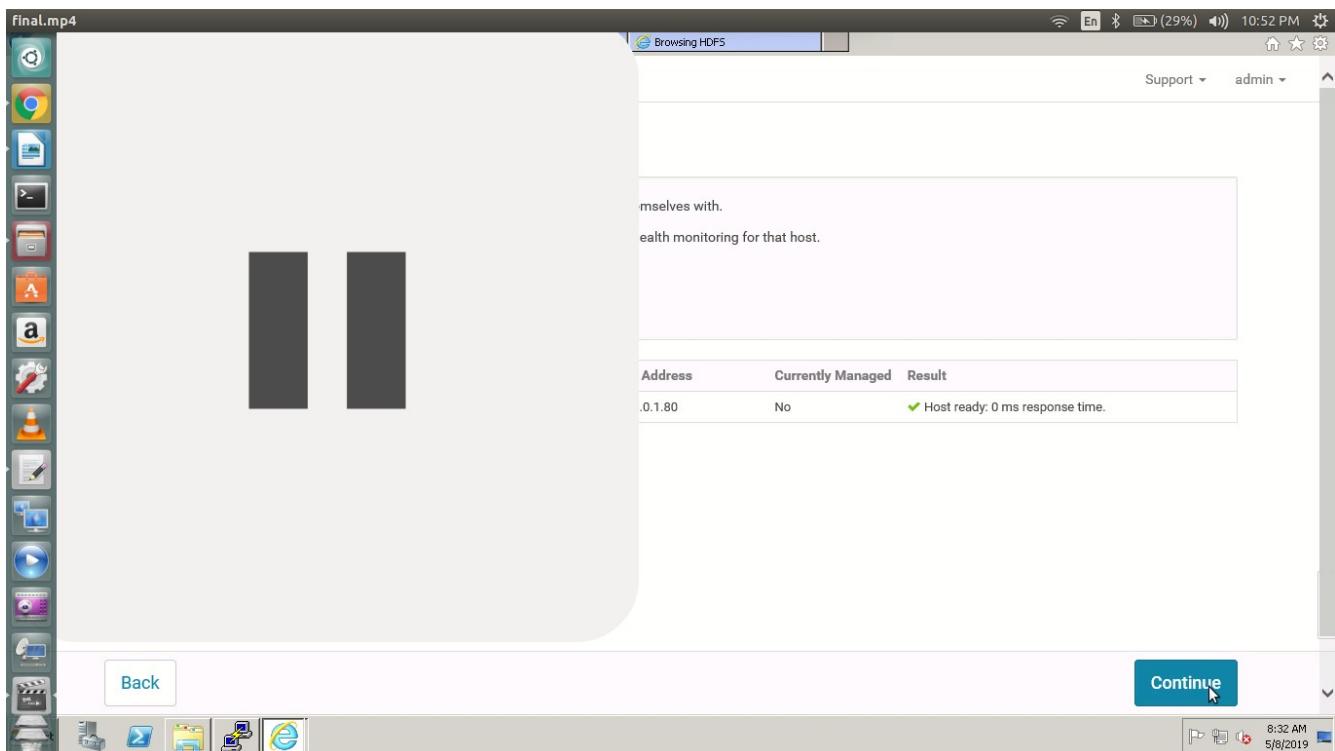
Support admin

mselves with.
ealth monitoring for that host.

Address	Currently Managed	Result
10.0.1.80	No	✓ Host ready: 0 ms response time.

Back Continue

8:32 AM 5/8/2019



final.mp4

Cloudera MANAGER

Add New Hosts to Cluster

Select Repository

Select the specific release of the Cloudera Manager Agent you want to install on your hosts.

Matched release for this Cloudera Manager Server
 Custom Repository

Back Continue

1 2 3 4 5 6 7 8

8:32 AM 5/8/2019

final.mp4

Oracle Binary Code License Agreement for the Java SE Platform Products and JavaFX

ORACLE AMERICA, INC. ("ORACLE"), FOR AND ON BEHALF OF ITSELF AND ITS SUBSIDIARIES AND AFFILIATES UNDER COMMON CONTROL, IS WILLING TO LICENSE THE SOFTWARE TO YOU ONLY UPON THE CONDITION THAT YOU ACCEPT ALL OF THE TERMS CONTAINED IN THIS BINARY CODE LICENSE AGREEMENT AND SUPPLEMENTAL LICENSE TERMS (COLLECTIVELY "AGREEMENT"). PLEASE READ THE AGREEMENT CAREFULLY. BY SELECTING THE "ACCEPT LICENSE AGREEMENT" (OR THE EQUIVALENT) BUTTON AND/OR BY USING THE SOFTWARE YOU ACKNOWLEDGE THAT YOU HAVE READ THE TERMS AND AGREE TO THEM. IF YOU ARE AGREEING TO THESE TERMS ON BEHALF OF A COMPANY OR OTHER LEGAL ENTITY, YOU REPRESENT THAT YOU HAVE THE LEGAL AUTHORITY TO BIND THE LEGAL ENTITY TO THESE TERMS. IF YOU DO NOT HAVE SUCH AUTHORITY, OR IF YOU DO NOT WISH TO BE BOUND BY THE TERMS, THEN SELECT THE "DECLINE LICENSE AGREEMENT" (OR THE EQUIVALENT) BUTTON AND YOU MUST NOT USE THE SOFTWARE ON THIS SITE OR ANY OTHER MEDIA ON WHICH THE SOFTWARE IS CONTAINED.

1. DEFINITIONS. "Software" means the software identified above in binary form that you selected for download, install or use (in the version You selected for download, install or use) from Oracle or its authorized licensees, any other machine readable materials (including, but not limited to, libraries, source files, header files, and data files), any updates or error corrections provided by Oracle, and any user manuals, programming guides and other documentation provided to you by Oracle under this Agreement. "General Purpose Desktop Computers and Servers" means computers, including desktop and laptop computers, or servers, used for general computing functions under end user control (such as but not specifically limited to email, general purpose Internet browsing, and office suite productivity tools). The use of Software in systems and solutions that provide dedicated functionality (other than as mentioned above) or designed for use in embedded or function-specific software applications, for example but not limited to: Software embedded in or bundled with industrial control systems, wireless mobile telephones, wireless handheld devices, kiosks, TV/STB, Blu-ray Disc devices, telematics and network control switching

Install Oracle Java SE Development Kit (JDK 7)

Check this box to accept the Oracle Binary Code License Agreement and install the JDK. Leave it unchecked to use a currently installed JDK.

WARNING: This Cloudera offering includes Oracle's Unlimited Strength Java(TM) Cryptography Extension (JCE) Policy Files for the Java(TM) Platform, Standard Edition (Java SE) Runtime Environment. Due to import restrictions of some countries, the version of the JCE Policy Files that are bundled in the Java Runtime Environment, or JRE(TM), allow "strong" but limited cryptography to be used. The Unlimited Strength JCE Policy Files included in this Cloudera offering, however, provides "unlimited strength" policy files which contain no restrictions on cryptographic strengths. Please note that some countries may legally prohibit the import of unlimited encryption strength policy files. You are responsible for determining whether you are subject to legal restrictions on cryptographic strength, and if so, you should not download this file. If legally permitted, you may configure the transport-layer security to reduce the cryptographic strength available with the default policy files. Additional information may be found at www.oracle.com. You are advised to consult your local laws before proceeding.

Install Java Unlimited Strength Encryption Policy Files

Check this checkbox if local laws permit you to deploy unlimited strength encryption and you are running a secure cluster.

Back Continue

1 2 3 4 5 6 7 8

8:32 AM 5/8/2019

Add New Hosts to Cluster - Cloudera Manager - Internet Explorer

Support admin

cloudera MANAGER

Add New Hosts to Cluster

Enter Login Credentials

Root access to your hosts is required to install the Cloudera packages. This installer will connect to your hosts via SSH and log in either directly as root or as another user with password-less sudo/pbrun privileges to become root.

Login To All Hosts As: root Another user ubuntu (with password-less sudo/pbrun to root)

You may connect via password or public-key authentication for the user selected above.

Authentication Method: All hosts accept same password All hosts accept same private key

Private Key File:

Enter Passphrase:

Confirm Passphrase:

SSH Port:

1 2 3 4 5 6 7 8

Back Continue

Feedback

8:33 AM 5/8/2019

final.mp4 Add New Hosts to Cluster - Cloudera Manager - Internet Explorer

8:33 AM 5/8/2019

Add New Hosts to Cluster

Install Agents

Installation in progress.

0 of 1 host(s) completed successfully. [Abort Installation](#)

Hostname	IP Address	Progress	Status
ip-10-0-1-80.ec2.internal	10.0.1.80	<div style="width: 100%;"> </div>	<input type="radio"/> Waiting for newly installed agent to heartbeat... Details

Feedback

Back Continue

Do you want AutoComplete to remember web form entries? [Learn about AutoComplete](#) Yes No

8:33 AM 5/8/2019

Snapshots: Allow snapshots

```
hdfe@pune:~  
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1924)  
at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2272)  
  
cloudage@pune:~$ su hdfe  
Password:  
hdfe@pune:~$ cd  
hdfe@pune:~$ hdfs dfsadmin -allowSnapshot /user/ubuntu/data.txt  
allowSnapshot: '/user/ubuntu/data.txt': Is not a directory  
    at org.apache.hadoop.hdfs.server.namenode.INodeDirectory.valueOf(INodeDirectory.java:61)  
    at org.apache.hadoop.hdfs.server.namenode.SnapshotManager.setSnapshottable(SnapshotManager.java:179)  
    at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.allowSnapshot(FSNamesystem.java:8486)  
    at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.allowSnapshot(NameNodeRpcServer.java:1508)  
    at org.apache.hadoop.hdfs.server.namenode.AuthorizationProviderProxyClientProtocol.allowSnapshot(AuthorizationProviderProxyClientProtocol.java:743)  
    at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.allowSnapshot(ClientNamenodeProtocolServerSideTranslatorPB.java:1105)  
    at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)  
    at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:617)  
    at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:1073)  
    at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2278)  
    at java.security.AccessController.doPrivileged(Native Method)  
    at javax.security.auth.Subject.doAs(Subject.java:415)  
    at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1924)  
    at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2272)  
' Is not a directory: '/user/ubuntu/data.txt': Is not a directory  
    at org.apache.hadoop.hdfs.server.namenode.INodeDirectory.valueOf(INodeDirectory.java:61)  
    at org.apache.hadoop.hdfs.server.namenode.SnapshotManager.setSnapshottable(SnapshotManager.java:179)  
    at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.allowSnapshot(FSNamesystem.java:8486)  
    at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.allowSnapshot(NameNodeRpcServer.java:1508)  
    at org.apache.hadoop.hdfs.server.namenode.AuthorizationProviderProxyClientProtocol.allowSnapshot(AuthorizationProviderProxyClientProtocol.java:743)  
    at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.allowSnapshot(ClientNamenodeProtocolServerSideTranslatorPB.java:1105)  
    at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)  
    at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:617)  
    at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:1073)  
    at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2278)  
    at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2274)  
    at java.security.AccessController.doPrivileged(Native Method)  
    at javax.security.auth.Subject.doAs(Subject.java:415)  
    at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1924)  
    at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2272)  
  
hdfe@pune:~$ hdfs dfsadmin -allowSnapshot /user/ubuntu/  
Allowing snapshot on /user/ubuntu/ succeeded  
hdfe@pune:~$
```

create a snapshotable directory

```
hdfe@pune:~  
hdfe@pune:~$ hdfs dfs -createSnapshot /user/ubuntu/ snapShotFolder  
Created snapshot /user/ubuntu/.snapshot/snapShotFolder  
hdfe@pune:~$
```

Browsing through Namenode web ui:

Namenode information - Internet Explorer
http://ip-10-0-1-115.ec2.internal:50070/dfshealth.html# This page can't be displayed

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

Path	Snapshot Number	Snapshot Quota	Modification Time	Permission	Owner	Group
/user/ubuntu	1	65536	Wed May 08 09:27:39 +0000 2019	rwxr-xr-x	hdfs	supergroup

Snapshot Summary

Snapshottable directories: 1

Path	Snapshot Number	Snapshot Quota	Modification Time	Permission	Owner	Group
/user/ubuntu	1	65536	Wed May 08 09:27:39 +0000 2019	rwxr-xr-x	hdfs	supergroup

Snapshotted directories: 1

Snapshot ID Snapshot Directory Modification Time

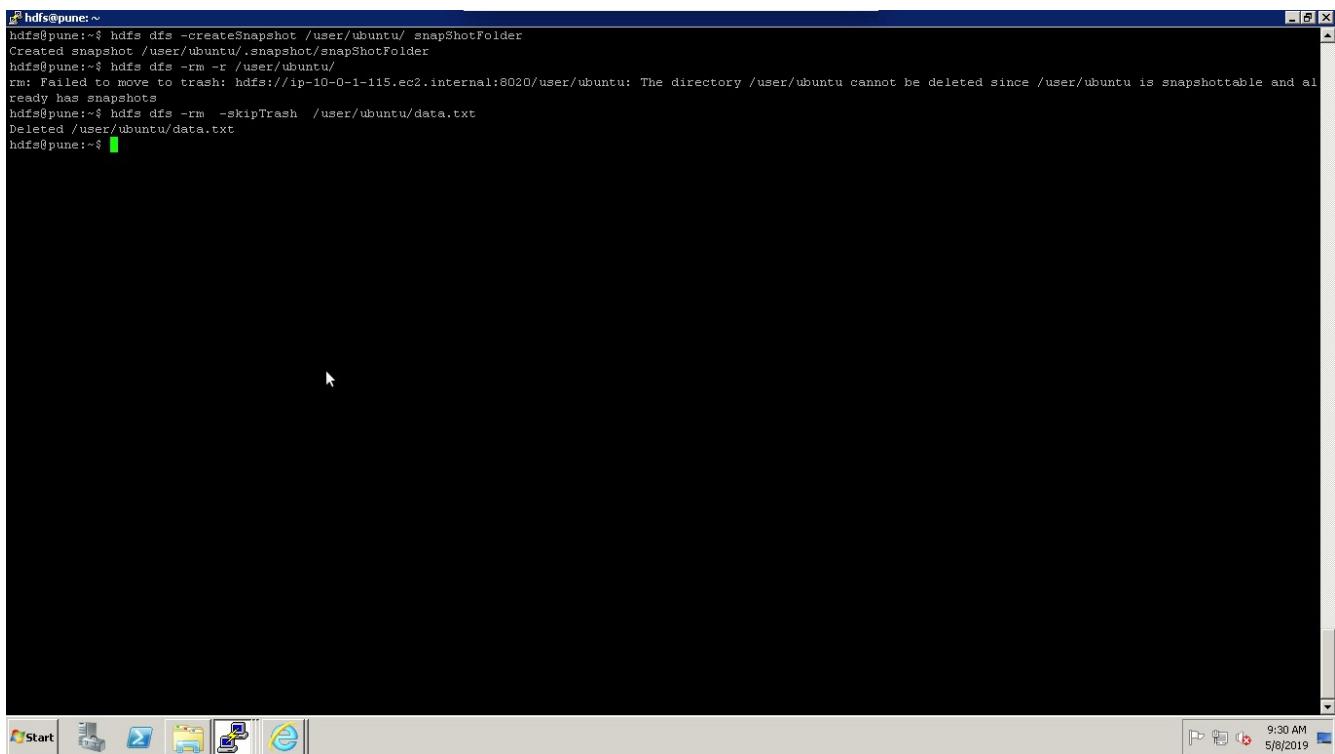
Start File Explorer Task View Internet Explorer 9:28 AM 5/8/2019

Snapshot-able directory cannot be deleted.

```
hdfs@pune:~$ hdfs dfs -createSnapshot /user/ubuntu/ snapShotFolder
Created snapshot /user/ubuntu/.snapshot/snapShotFolder
hdfs@pune:~$ hdfs dfs -rm -r /user/ubuntu/
rm: Failed to move to trash: hdfs://ip-10-0-1-115.ec2.internal:8020/user/ubuntu: The directory /user/ubuntu cannot be deleted since /user/ubuntu is snapshottable and already has snapshots
hdfs@pune:~$
```

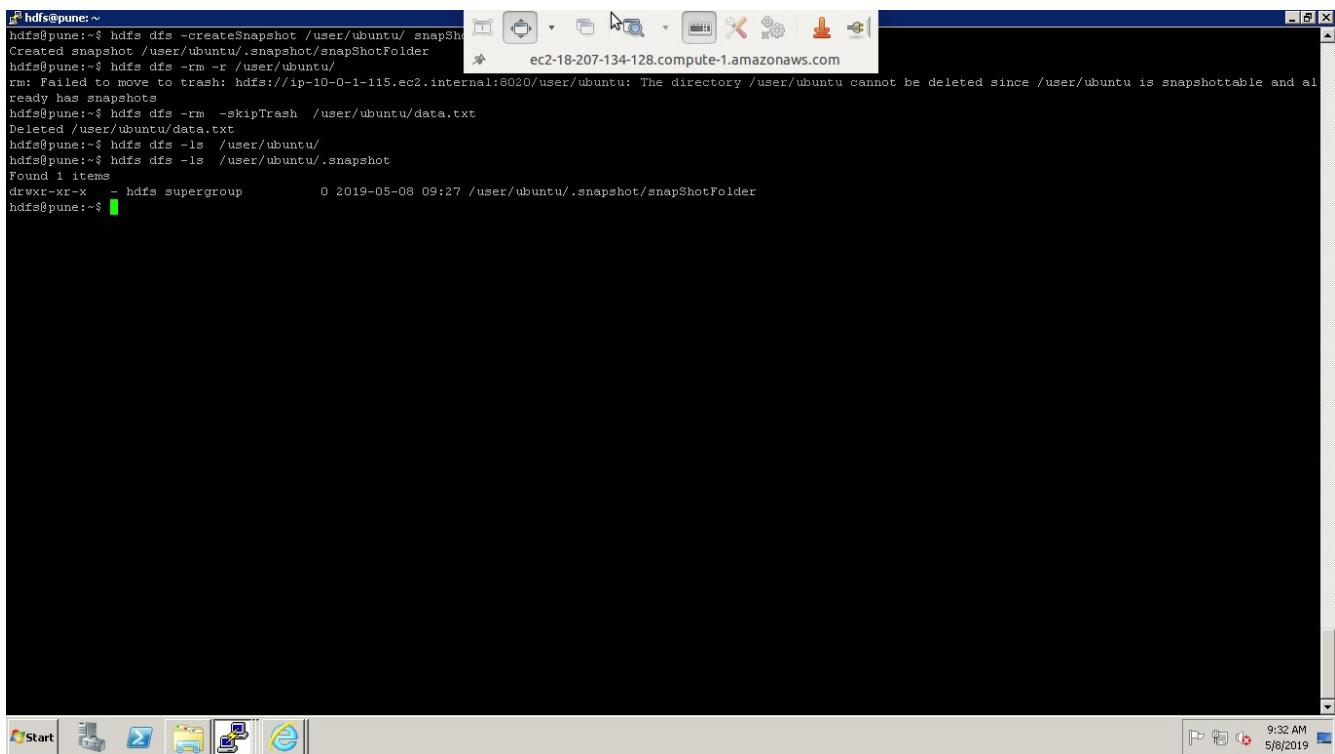
Start File Explorer Task View Internet Explorer 9:28 AM 5/8/2019

Deleting dataset.txt permanently



```
hdfs@pune:~$ hdfs dfs -createSnapshot /user/ubuntu/ snapShotFolder
Created snapshot /user/ubuntu/.snapshot/snapShotFolder
hdfs@pune:~$ hdfs dfs -rm -r /user/ubuntu/
rm: Failed to move to trash: hdfs://ip-10-0-1-115.ec2.internal:8020/user/ubuntu: The directory /user/ubuntu cannot be deleted since /user/ubuntu is snapshottable and already has snapshots
hdfs@pune:~$ hdfs dfs -rm -skipTrash /user/ubuntu/data.txt
Deleted /user/ubuntu/data.txt
hdfs@pune:~$
```

data.txt present in snapshot directory



```
hdfs@pune:~$ hdfs dfs -createSnapshot /user/ubuntu/ snapShotFolder
Created snapshot /user/ubuntu/.snapshot/snapShotFolder
hdfs@pune:~$ hdfs dfs -rm -r /user/ubuntu/
rm: Failed to move to trash: hdfs://ip-10-0-1-115.ec2.internal:8020/user/ubuntu: The directory /user/ubuntu cannot be deleted since /user/ubuntu is snapshottable and already has snapshots
hdfs@pune:~$ hdfs dfs -rm -skipTrash /user/ubuntu/data.txt
Deleted /user/ubuntu/data.txt
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/.snapshot
Found 1 items
drwxr-xr-x 1 hdfs supergroup 0 2019-05-08 09:27 /user/ubuntu/.snapshot/snapShotFolder
hdfs@pune:~$
```

```
hdfs@pune:~$ hdfs dfs -createSnapshot /user/ubuntu/ snapShotFolder
Created snapshot /user/ubuntu/.snapshot/snapShotFolder
hdfs@pune:~$ hdfs dfs -rm -r /user/ubuntu/
rm: Failed to move to trash: hdfs://ip-10-0-1-115.ec2.internal:8020/user/ubuntu: The directory /user/ubuntu cannot be deleted since /user/ubuntu is snapshottable and already has snapshots
hdfs@pune:~$ hdfs dfs -rm -skipTrash /user/ubuntu/data.txt
Deleted /user/ubuntu/data.txt
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/.snapshot
Found 1 items
drwxr-xr-x  - hdfs supergroup          0 2019-05-08 09:27 /user/ubuntu/.snapshot/snapShotFolder
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/.snapshot/snapShotFolder
Found 1 items
-rw-r--r--  4 hdfs supergroup      72 2019-05-08 07:04 /user/ubuntu/.snapshot/snapShotFolder/data.txt
hdfs@pune:~$
```

Resorting from Snapshot

```
hdfs@pune:~$ hdfs dfs -createSnapshot /user/ubuntu/ snapShotFolder
Created snapshot /user/ubuntu/.snapshot/snapShotFolder
hdfs@pune:~$ hdfs dfs -rm -r /user/ubuntu/
rm: Failed to move to trash: hdfs://ip-10-0-1-115.ec2.internal:8020/user/ubuntu: The directory /user/ubuntu cannot be deleted since /user/ubuntu is snapshottable and already has snapshots
hdfs@pune:~$ hdfs dfs -rm -skipTrash /user/ubuntu/data.txt
Deleted /user/ubuntu/data.txt
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/.snapshot
Found 1 items
drwxr-xr-x  - hdfs supergroup          0 2019-05-08 09:27 /user/ubuntu/.snapshot/snapShotFolder
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/.snapshot/snapShotFolder
Found 1 items
-rw-r--r--  4 hdfs supergroup      72 2019-05-08 07:04 /user/ubuntu/.snapshot/snapShotFolder/data.txt
hdfs@pune:~$ hdfs dfs -cp /user/ubuntu/.snapshot/snapShotFolder/data.txt /user/ubuntu/data.txt
hdfs@pune:~$
```

```
hdfs@pune:~$ hdfs dfs -createSnapshot /user/ubuntu/ snapShotFolder
Created snapshot /user/ubuntu/.snapshot/snapShotFolder
hdfs@pune:~$ hdfs dfs -rm -r /user/ubuntu/
rm: Failed to move to trash: hdfs://ip-10-0-1-115.ec2.internal:8020/user/ubuntu: The directory /user/ubuntu cannot be deleted since /user/ubuntu is snapshottable and already has snapshots
hdfs@pune:~$ hdfs dfs -rm -skipTrash /user/ubuntu/data.txt
Deleted /user/ubuntu/data.txt
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/.snapshot
Found 1 items
drwxr-xr-x - hdfs supergroup          0 2019-05-08 09:27 /user/ubuntu/.snapshot/snapShotFolder
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/.snapshot/snapShotFolder
Found 1 items
-rw-r--r--  4 hdfs supergroup      72 2019-05-08 07:04 /user/ubuntu/.snapshot/snapShotFolder/data.txt
hdfs@pune:~$ hdfs dfs -cp /user/ubuntu/.snapshot/snapShotFolder/data.txt /user/ubuntu/data.txt
hdfs@pune:~$ hdfs dfs -ls /user/ubuntu/
Found 1 items
-rw-r--r--  3 hdfs supergroup      72 2019-05-08 09:36 /user/ubuntu/data.txt
hdfs@pune:~$
```

