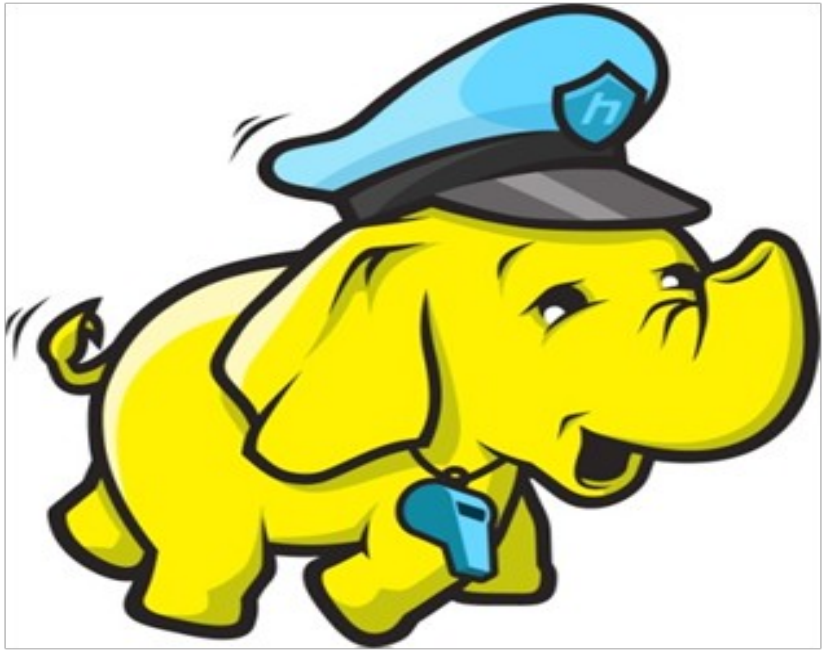
Upgrade your Hadoop cluster



Agenda

What is Hadoop ugradation
Why upgradation is required
Save old Hadoop directory structure
Stop old Hadoop
Install new version of Hadoop
Verify Namenode & data directory
Start Namenode with upgrade option
Start remaining Hadoop processes
Check upgradation status
Rollback the upgrade
Finalize the upgrade

What is Hadoop ugradation

Apache releases new versions of Hadoop.

We can use newer version of Hadoop without loosing data.

Moving from old version of Hadoop to newer version of Hadoop is known as upgradation.

Moving from new version of Hadoop to older version of Hadoop is known as downgradation.

Why upgradation is required

Newer version of Hadoop are better than older versions.

New versions has bugs fixed which were present in older version.

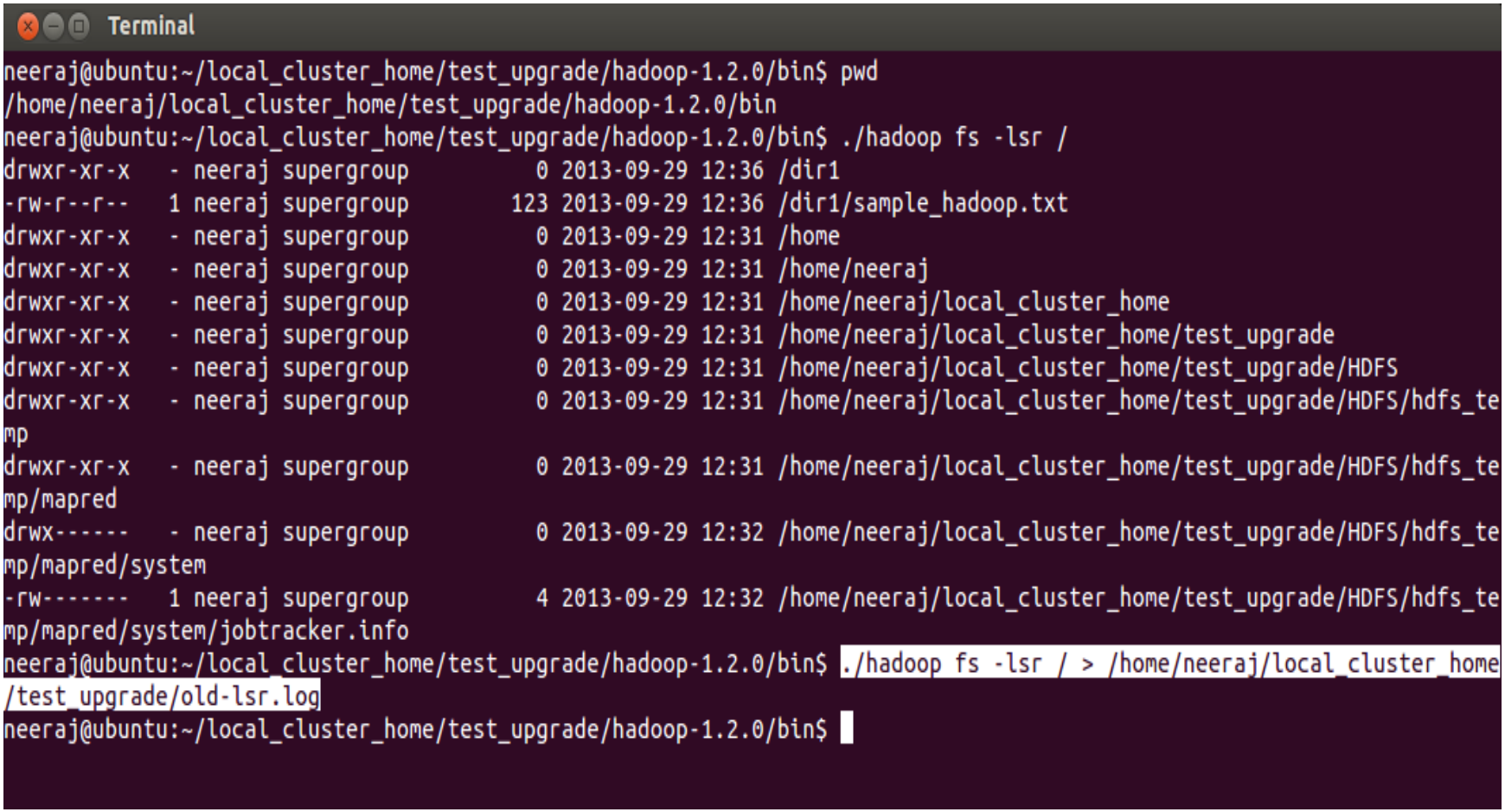
New version has new features.

New versions comes with updated functionality.

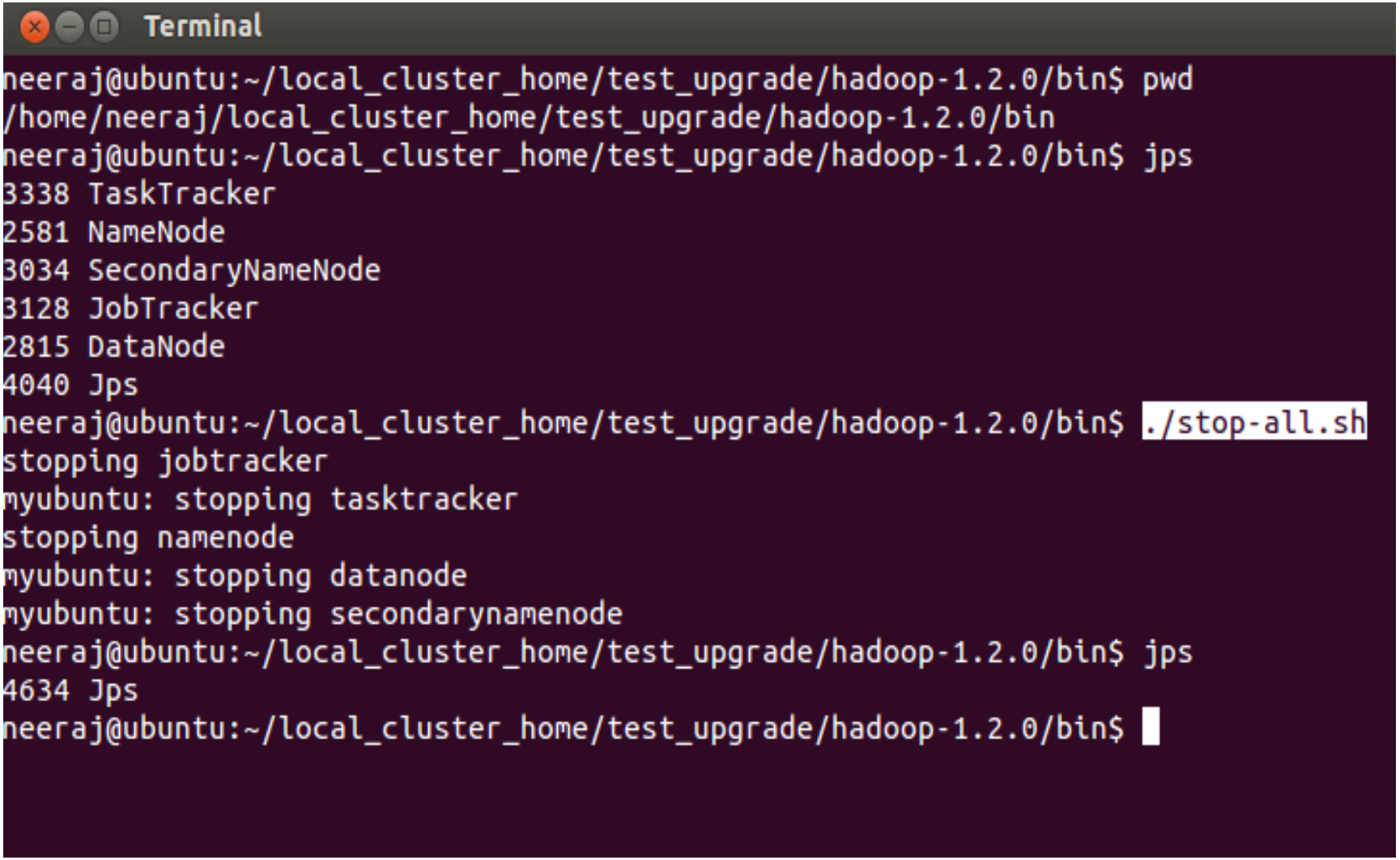
New version may be stable or unstable.

Save old directory structure

Save the complete directory structure in a file for future use.



Stop old Hadoop



Install new version of Hadoop

Download new version of Hadoop from Apache website.

Make the necessary changes in configuration files.

Repeat these steps on all machines on cluster.

Follow Hadoop Installation PDF for more details.

New version of Hadoop uses the old fsimage to create new metadata.

Verify Namenode & data directory

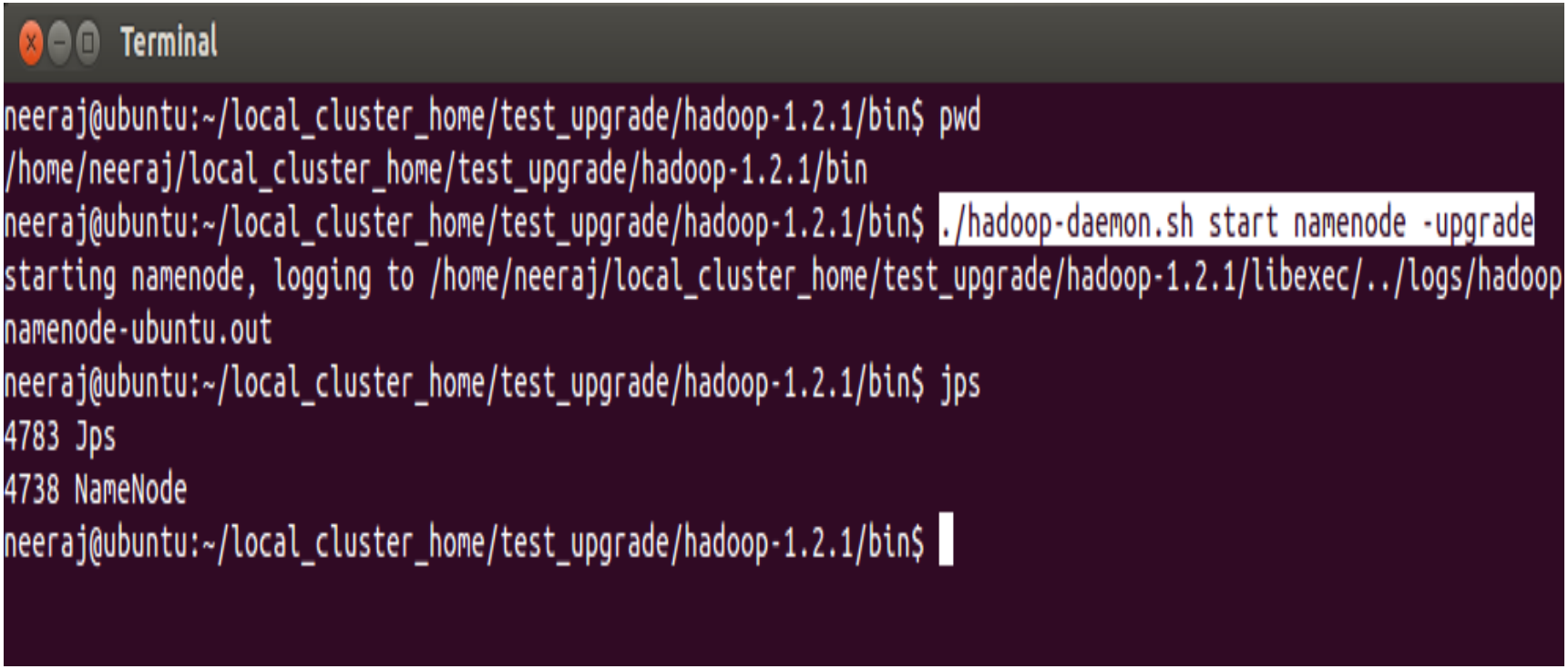
Make sure that the Namenode directory & Data directory points to old version of Hadoop

```
<property>

    <name>dfs.name.dir</name>
    <value>/home/neeraj/local_cluster_home/test_upgrade/HDFS/hdfs_metadata</value>
    <description>This directory will be used to store the Namenode metadata </description>

 </property>

<property>

    <name>dfs.data.dir</name>
    <value>/home/neeraj/local_cluster_home/test_upgrade/HDFS/hdfs_data</value>
    <description>This directory will be used to store the data </description>

 </property>
```

Start Namenode with upgrade option

Fsimage contains all the metadata.

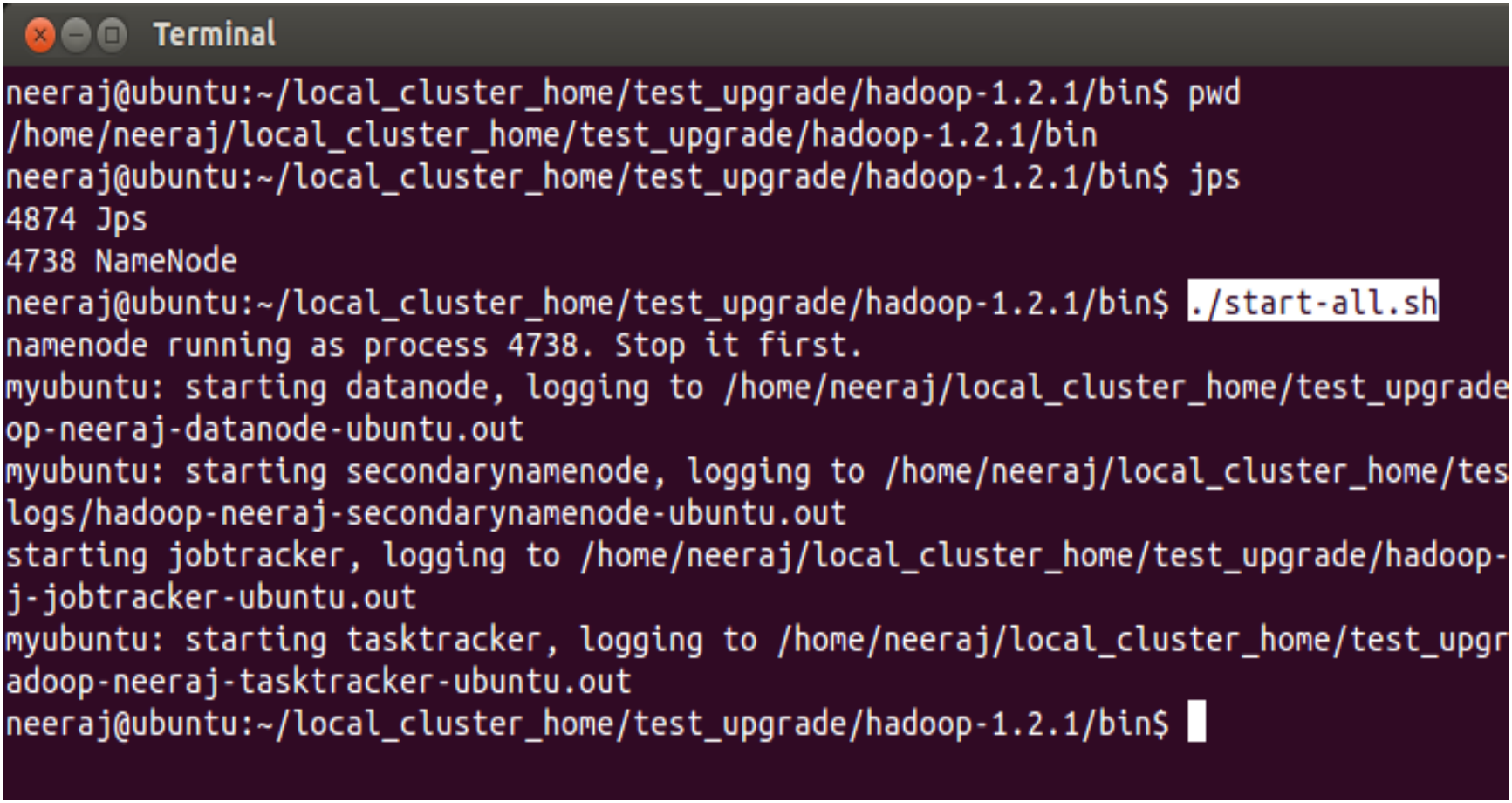
New Hadoop takes old fsimage & create a new fsimage for new version



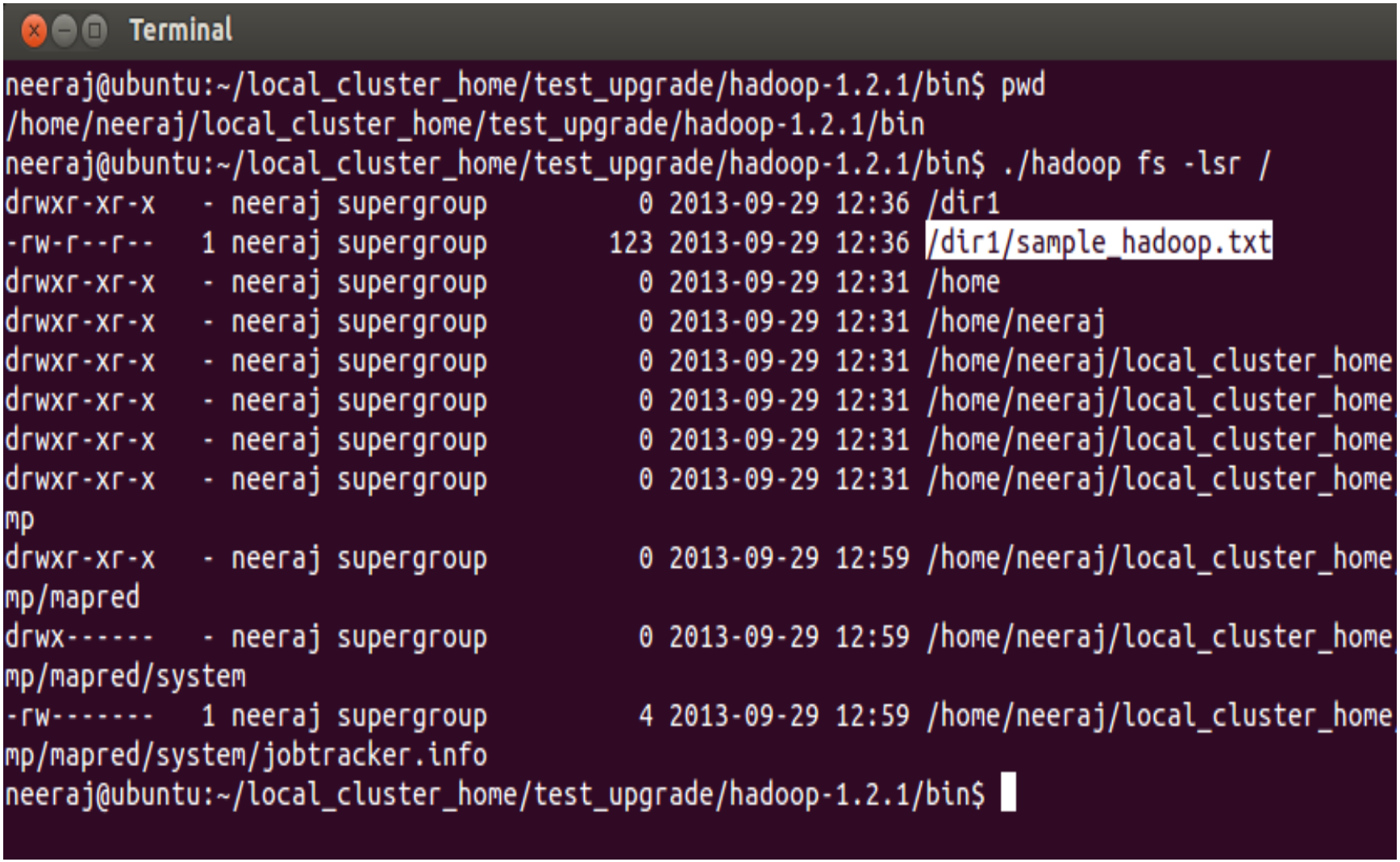
Start new Hadoop

Start new Hadoopusing the below command.

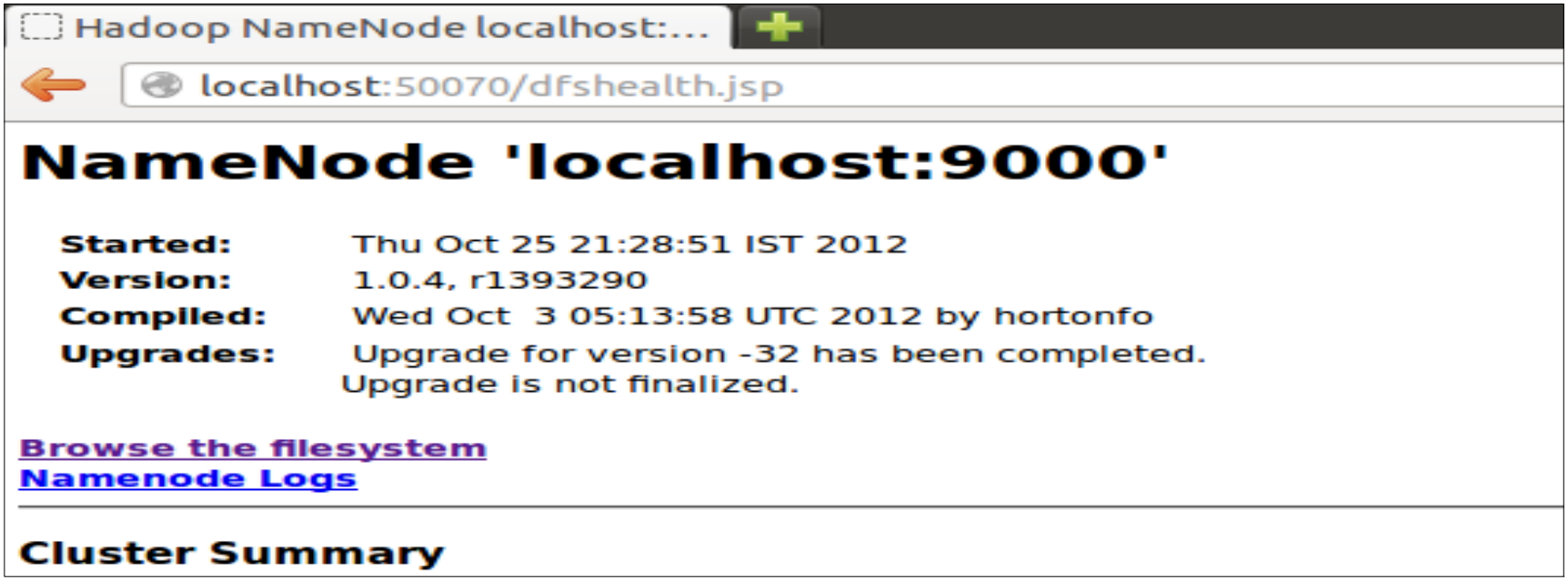
./start-all.sh

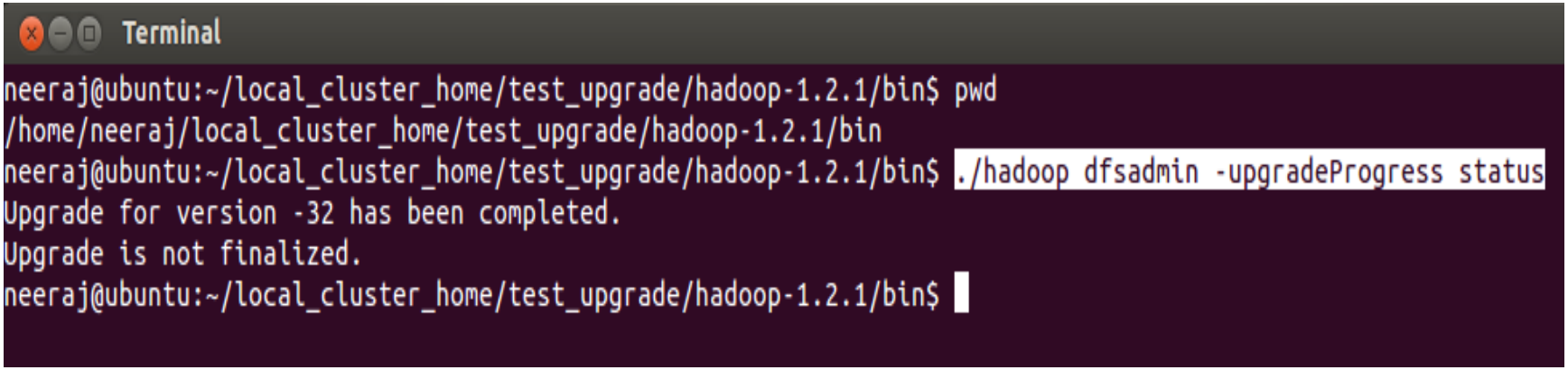


Verify HDFS data in new Hadoop



Check upgradation status

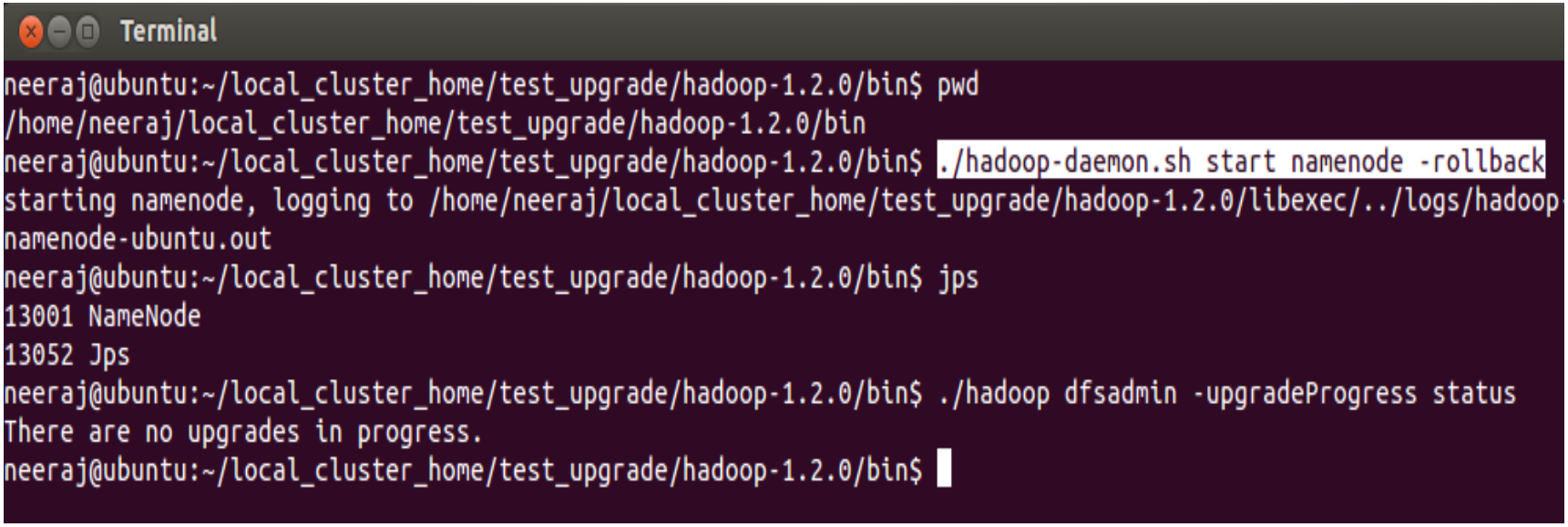




Rollback the upgrade

Rollback the upgrade by running below command from old Hadoop bin directory

./hadoop-daemmon.sh start namenode -rollback

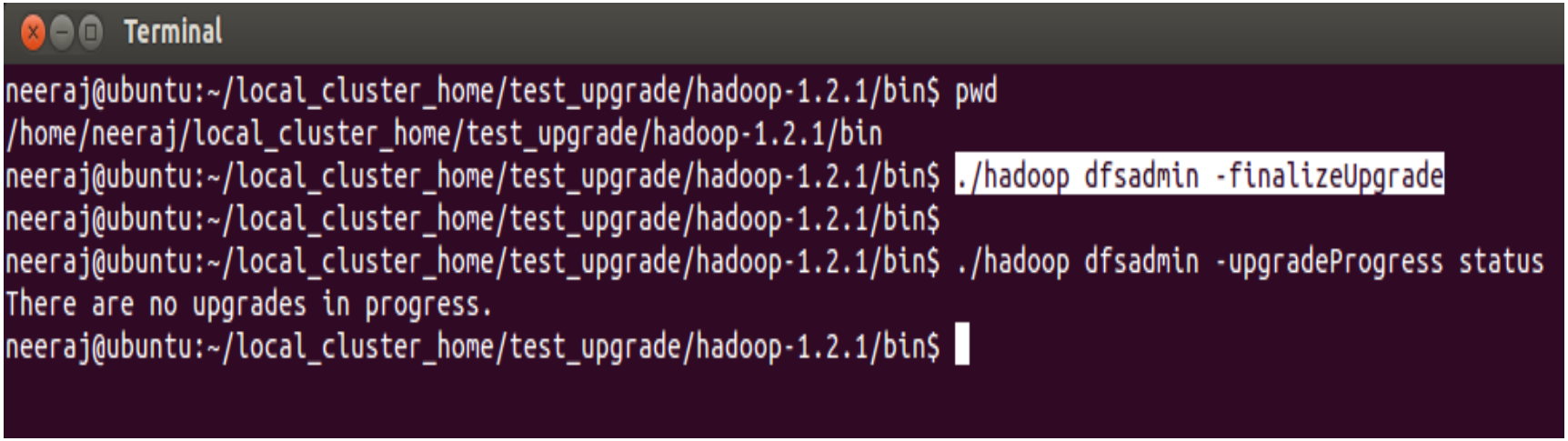


You have successfully rolled back your Hadoop

Finalize the upgrade

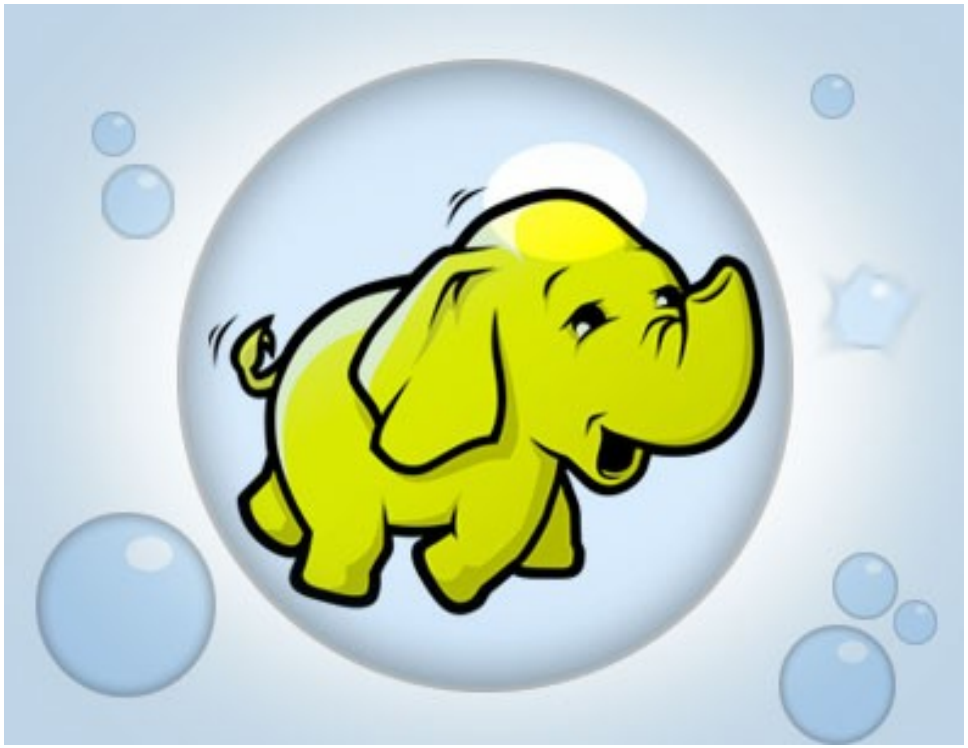
Finalize the upgrade by running below command from new Hadoop bin directory

./hadoop dfsadmin -finalizeUpgrade



You have successfully upgraded your Hadoop

...Thanks...



For online Hadoop training, send mail to neeraj.ymca.2k6@gmail.com