Lab - 2

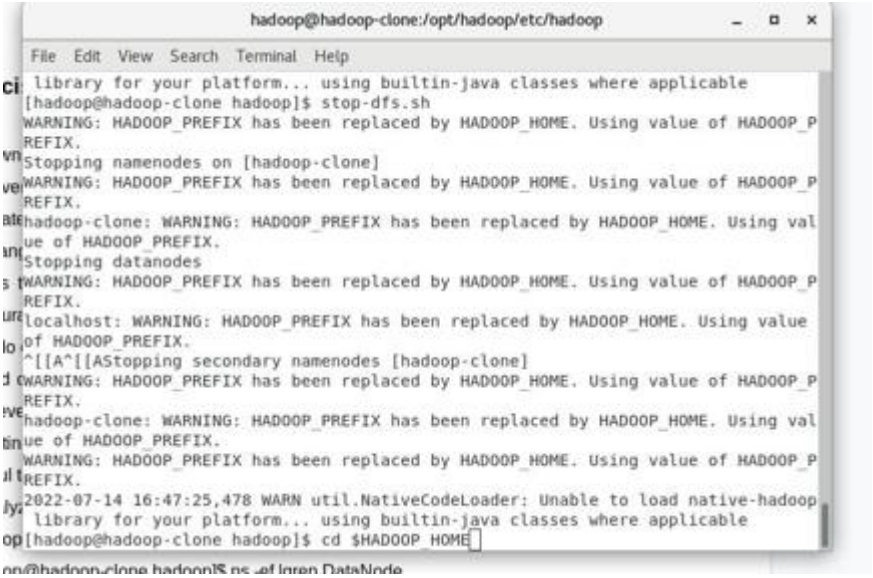
Aim: "Configuring and experiencing Hadoop Single node cluster. Start/stop script and monitoring."

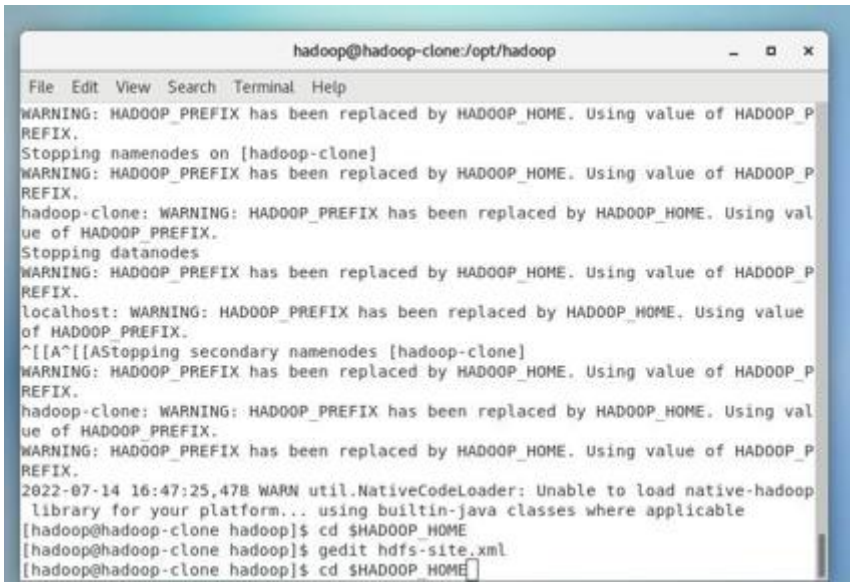
Rahil Acharya

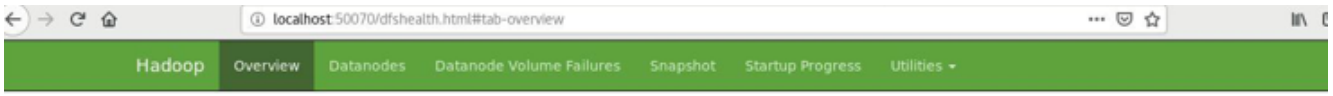
CE 004

20CEUOD004

Change the URL of HDFS Web UI to localhost:51234 instead of localhost:50070 and the access the web UI using new URL. [Hint: refer configuration files] [Know that traditionally configuration based servers need to be restarted after any server configuration change.]







Overview 'hadoop-clone:9000' (active)

Started:	Thu Jul 14 16:50:43 +0530 2022
Version:	3.2.0, re97acb3bd8f3befd27418996fa5d4b50bf2e17bf
Compiled:	Tue Jan 08 11:38:00 +0530 2019 by sunilg from branch-3.2.0
Cluster ID:	CID-a0711dac-8474-44f9-8368-d68c1ff4c838
Block Pool ID:	BP-1266388418-127.0.0.1-1560101340487

Summary

Security is off.

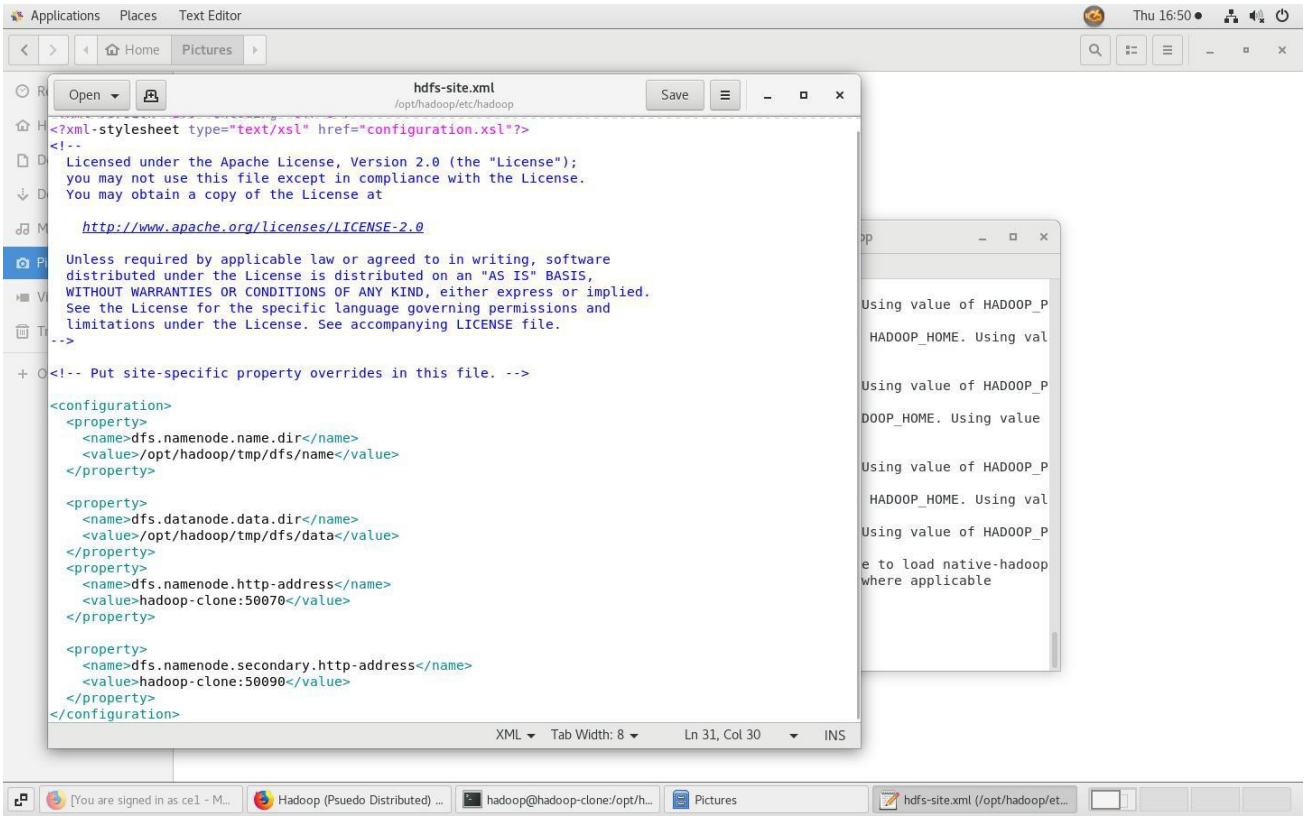
Safe mode is ON. The reported blocks 5 has reached the threshold 0.9990 of total blocks 5. The number of live datanodes 1 has reached the minimum number 0. In safe mode extension. Safe mode will be turned off automatically in 20 seconds.

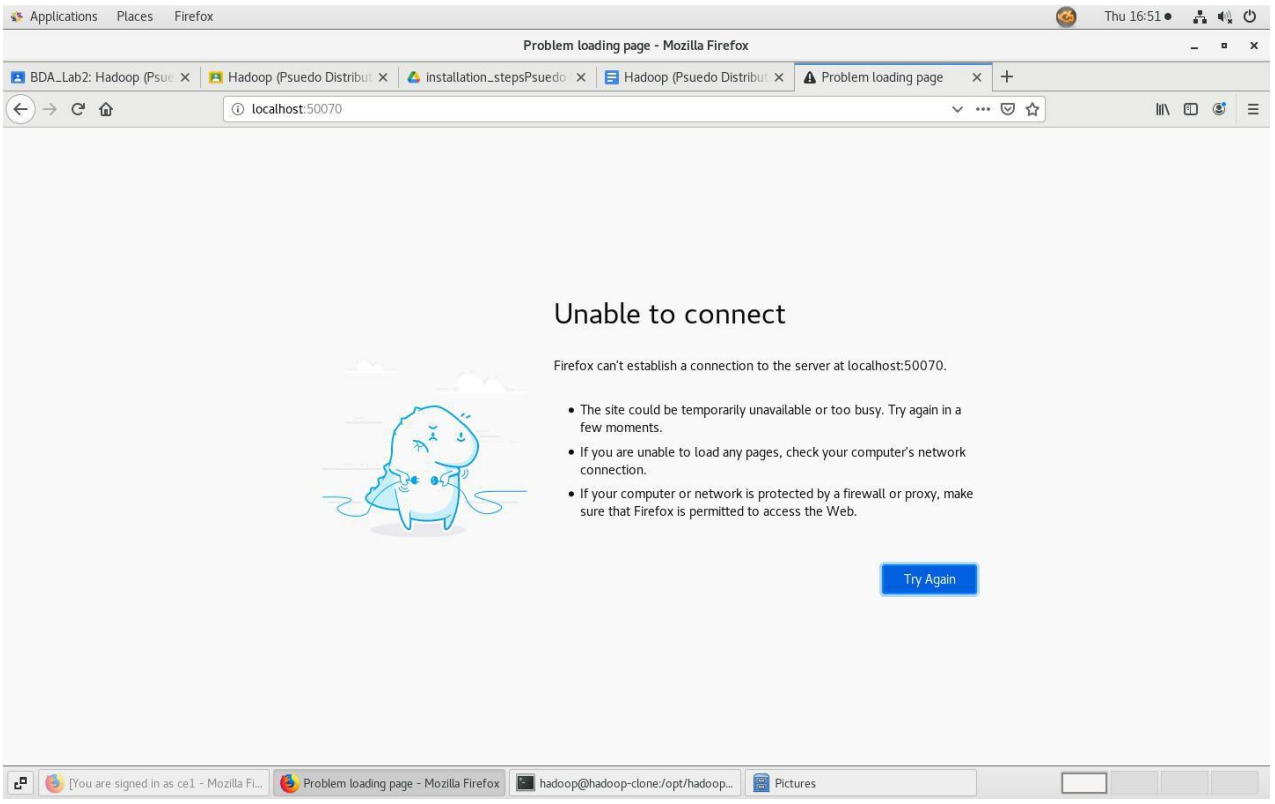
42 files and directories, 5 blocks (5 replicated blocks, 0 erasure coded block groups) = 47 total filesystem object(s).

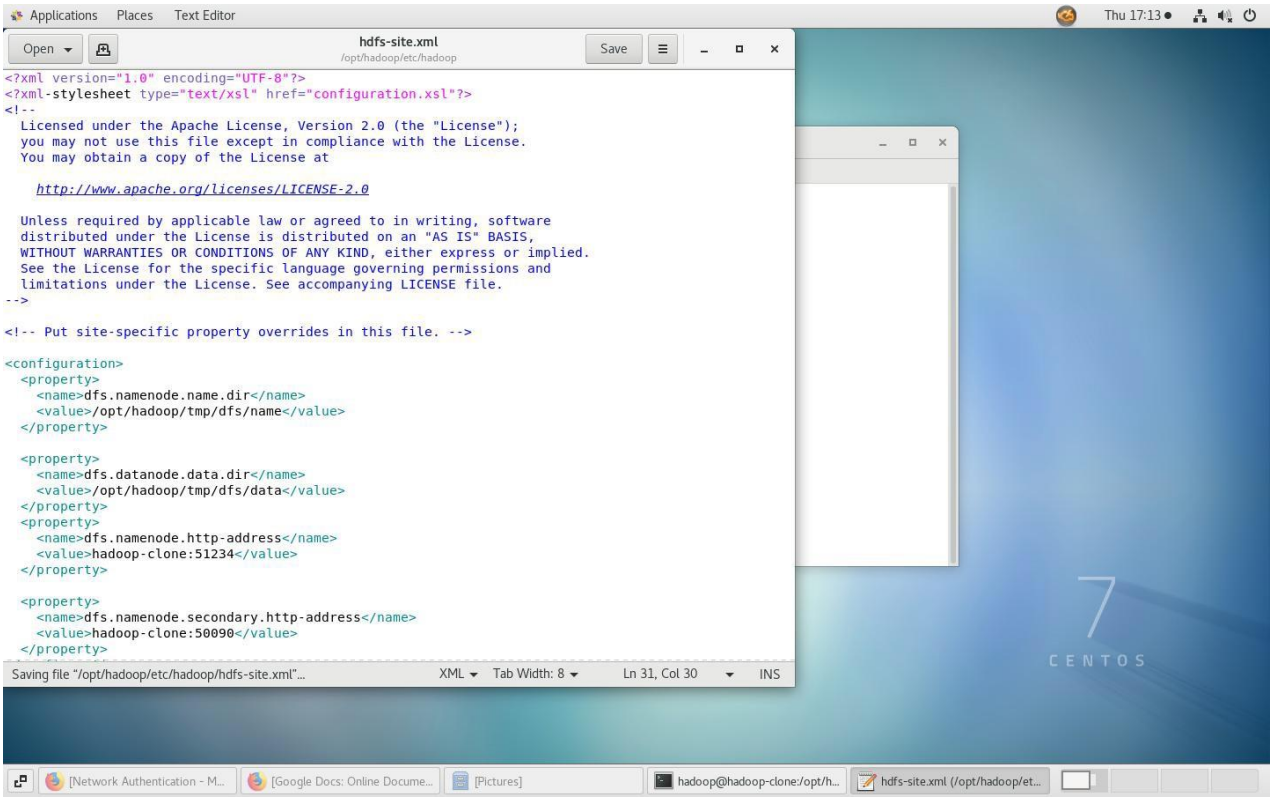
Heap Memory used 84.03 MB of 423.5 MB Heap Memory. Max Heap Memory is 3.42 GB.

Non Heap Memory used 46.69 MB of 48.03 MB Commited Non Heap Memory. Max Non Heap Memory is <unbounded>.

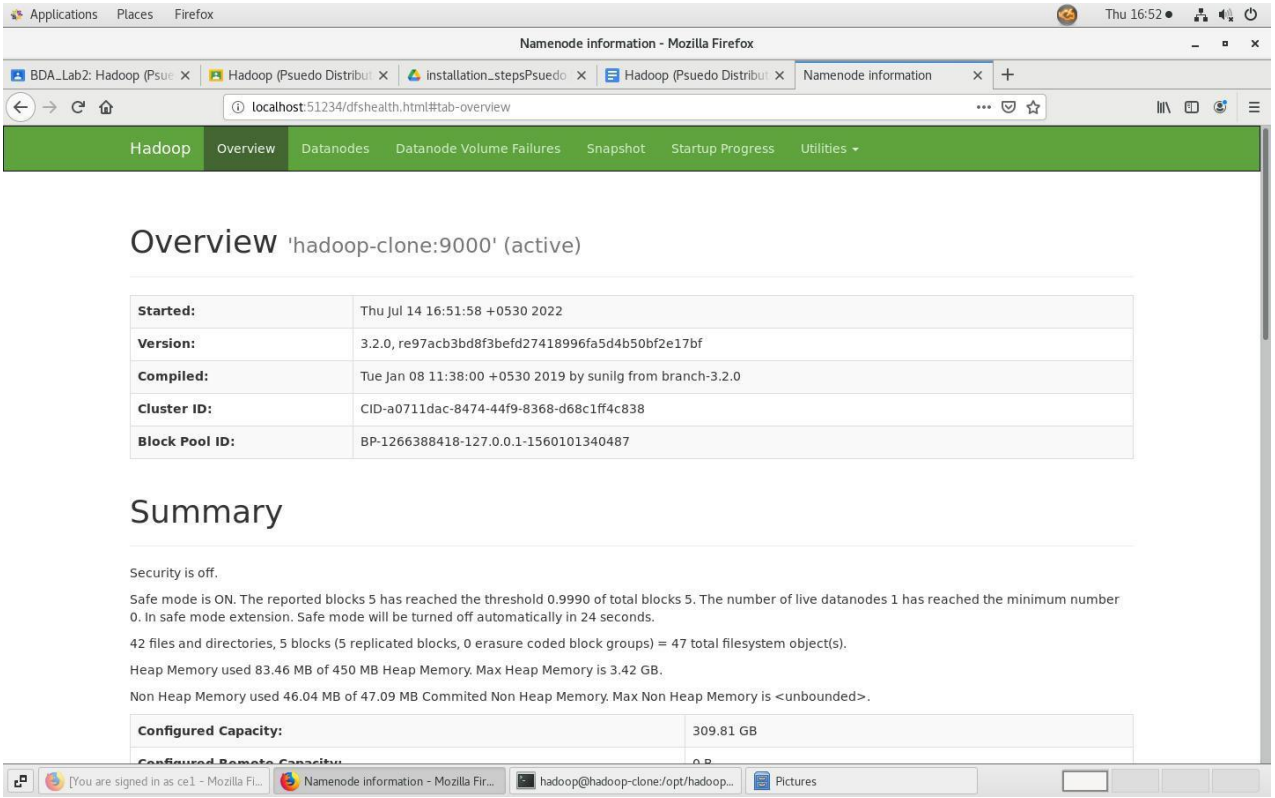
Configured Capacity:	309.81 GB


Applications
Places
Text Editor
Thu 16:50
Home
Pictures
Open
hdfs-site.xml
/opt/hadoop/etc/hadoop
Save
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
Licensed under the Apache License, Version 2.0 (the "License");
you may not use this file except in compliance with the License.
You may obtain a copy of the License at
http://www.apache.org/licenses/LICENSE-2.0
Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->
<!-- Put site-specific property overrides in this file. -->
<configuration>
<property>
<name>dfs.namenode.name.dir</name>
<value>/opt/hadoop/tmp/dfs/name</value>
</property>
<property>
<name>dfs.datanode.data.dir</name>
<value>/opt/hadoop/tmp/dfs/data</value>
</property>
<property>
<name>dfs.namenode.http-address</name>
<value>hadoop-clone:50070</value>
</property>
<property>
<name>dfs.namenode.secondary.http-address</name>
<value>hadoop-clone:50090</value>
</property>
</configuration>
XML
Tab Width: 8
Ln 31, Col 30
INS
Using value of HADOOP_P
HADOOP_HOME. Using val
Using value of HADOOP_P
DOOP_HOME. Using value
Using value of HADOOP_P
HADOOP_HOME. Using val
Using value of HADOOP_P
e to load native-hadoop
where applicable
[You are signed in as ce1 - M...
Hadoop (Psuedo Distributed) ...
hadoop@hadoop-clone:/opt/h...
Pictures
hdfs-site.xml (/opt/hadoop/et...


Applications
Places
Firefox
Thu 16:51
Problem loading page - Mozilla Firefox
BDA_Lab2: Hadoop (Psue
Hadoop (Psuedo Distribut
installation_stepsPsuedo
Hadoop (Psuedo Distribut
Problem loading page
localhost:50070
Unable to connect
Firefox can't establish a connection to the server at localhost:50070.
The site could be temporarily unavailable or too busy. Try again in a few moments.
If you are unable to load any pages, check your computer's network connection.
If your computer or network is protected by a firewall or proxy, make sure that Firefox is permitted to access the Web.
Try Again
[You are signed in as ce1 - Mozilla Fi...
Problem loading page - Mozilla Firefox
hadoop@hadoop-clone:/opt/hadoop...
Pictures

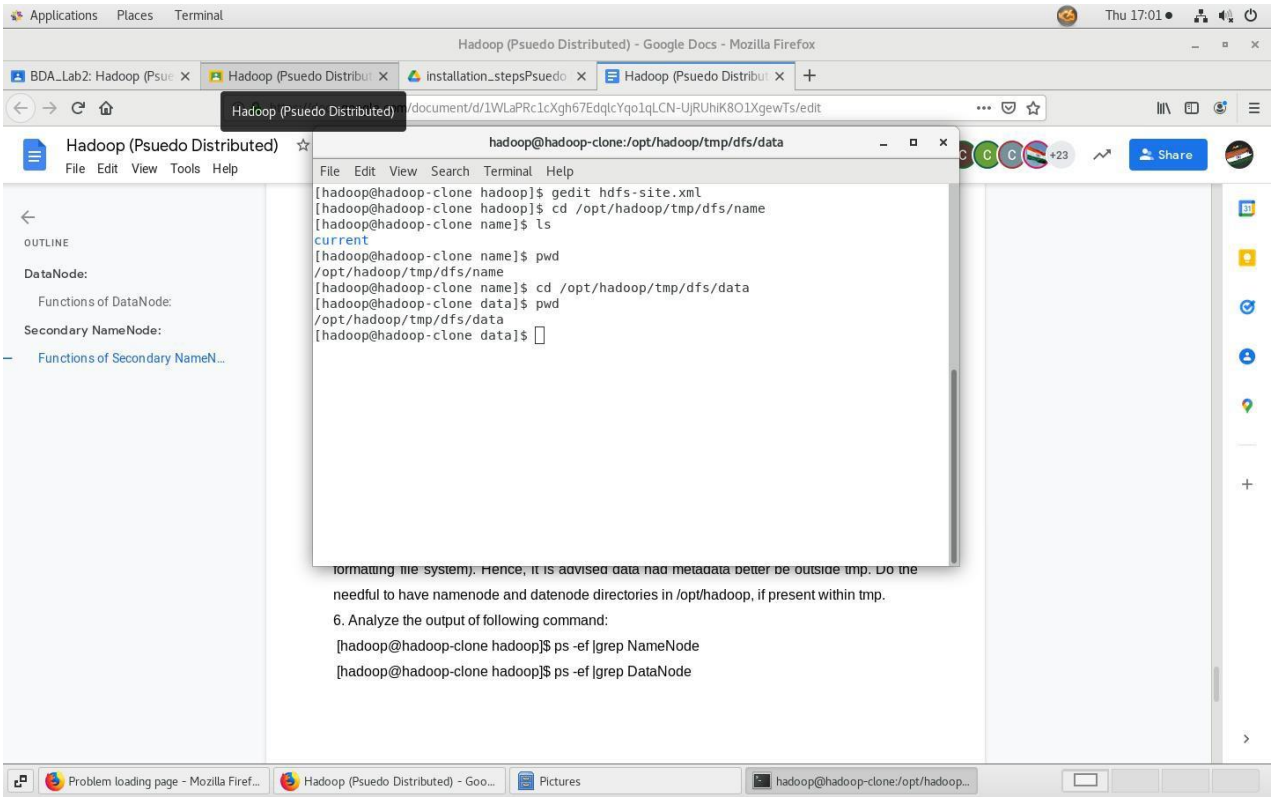


Undo changes done in task1.

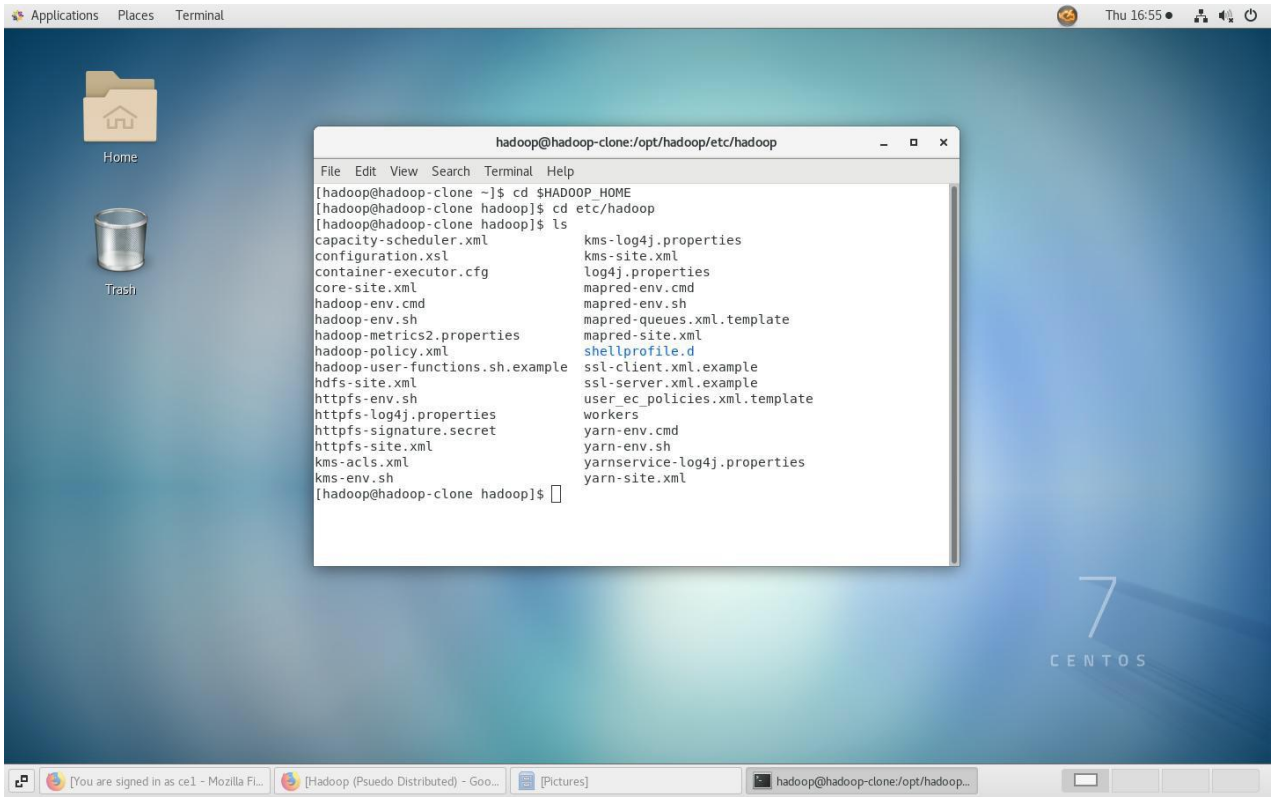


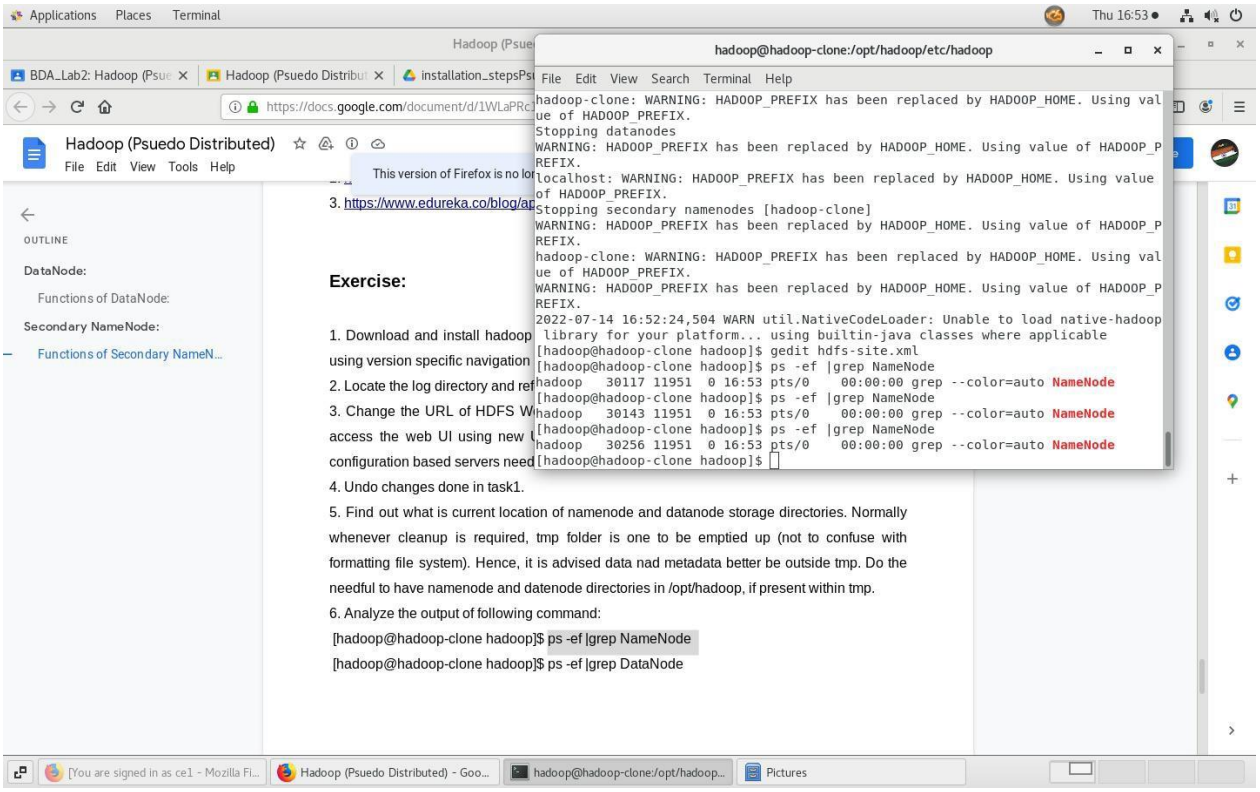
Find out what is current location of namenode and datanode storagedirectories. Normally whenever cleanup is required, tmp folder is one tobe emptied up (not to confuse with formatting file system). Hence, it is advised data and metadata better be outside tmp. Do the

needful to have namenode and datenode directories in /opt/hadoop, ifpresent within tmp.



Analyze the output of following command: [hadoop@hadoop-clone hadoop]$ ps -ef |grep NameNode[hadoop@hadoop-clone hadoop]$ ps -ef |grep DataNode




Applications Places Terminal
Thu 16:53
hadoop@hadoop-clone:/opt/hadoop/etc/hadoop
File Edit View Search Terminal Help
hadoop-clone: WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using val
ue of HADOOP_PREFIX.
Stopping datanodes
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_P
REFIX.
localhost: WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value
of HADOOP_PREFIX.
Stopping secondary namenodes [hadoop-clone]
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_P
REFIX.
hadoop-clone: WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using val
ue of HADOOP_PREFIX.
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_P
REFIX.
2022-07-14 16:52:24,504 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
[hadoop@hadoop-clone hadoop]$ gedit hdfs-site.xml
[hadoop@hadoop-clone hadoop]$ ps -ef |grep NameNode
hadoop 30117 11951 0 16:53 pts/0 00:00:00 grep --color=auto NameNode
[hadoop@hadoop-clone hadoop]$ ps -ef |grep NameNode
hadoop 30143 11951 0 16:53 pts/0 00:00:00 grep --color=auto NameNode
[hadoop@hadoop-clone hadoop]$ ps -ef |grep NameNode
hadoop 30256 11951 0 16:53 pts/0 00:00:00 grep --color=auto NameNode
[hadoop@hadoop-clone hadoop]$
Hadoop (Psuedo Distributed)
File Edit View Tools Help
OUTLINE
DataNode:
Functions of DataNode:
Secondary NameNode:
Functions of Secondary NameN...
Exercise:
1. Download and install hadoop
using version specific navigation
2. Locate the log directory and ref
3. Change the URL of HDFS W
access the web UI using new
configuration based servers need
4. Undo changes done in task1.
5. Find out what is current location of namenode and datanode storage directories. Normally whenever cleanup is required, tmp folder is one to be emptied up (not to confuse with formatting file system). Hence, it is advised data nad metadata better be outside tmp. Do the needful to have namenode and datenode directories in /opt/hadoop, if present within tmp.
6. Analyze the output of following command:
[hadoop@hadoop-clone hadoop]$ ps -ef |grep NameNode
[hadoop@hadoop-clone hadoop]$ ps -ef |grep DataNode
Pictures