Exp. No : 1

Installation of Hadoop

1. Download Hadoop Binary file
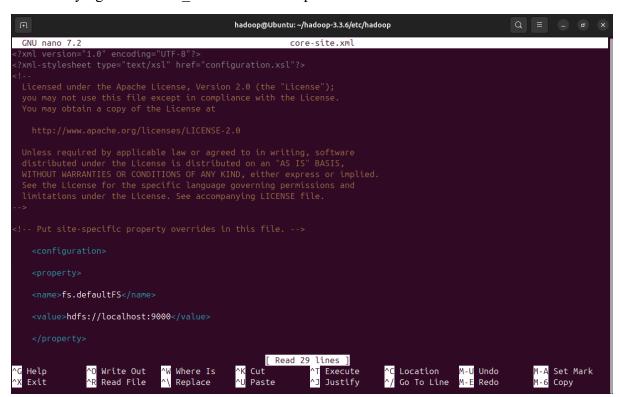

2. Modifying .bashrc file

3. Updating JAVA_HOME in $HADOOP_HOME/etc/hadoop/hadoop-env.sh

```
hadoop@Ubuntu: ~/hadoop-3.3.6/etc/hadoop
  GNU nano 7.2                              hadoop-env.sh
## {YARN_xyz|HDFS_xyz} > HADOOP_xyz > hard-coded defaults
##

# Many of the options here are built from the perspective that users
# may want to provide OVERWRITING values on the command line.
# For example:
#
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
#
# Therefore, the vast majority (BUT NOT ALL!) of these defaults
# are configured for substitution and not append.  If append
# is preferable, modify this file accordingly.

###
# Generic settings for HADOOP
###

# Technically, the only required environment variable is JAVA_HOME.
# All others are optional.  However, the defaults are probably not
# preferred.  Many sites configure these options outside of Hadoop,
# such as in /etc/profile.d

# The java implementation to use. By default, this environment
# variable is REQUIRED on ALL platforms except OS X!
# export JAVA_HOME=

# Location of Hadoop.  By default, Hadoop will attempt to determine
# this location based upon its execution path.

^G Help       ^O Write Out   ^W Where Is   ^K Cut      ^T Execute   ^C Location   M-U Undo   M-A Set Mark
^X Exit       ^R Read File   ^\ Replace    ^U Paste    ^J Justify   ^/ Go To Line M-E Redo   M-6 Copy
```

4.   Modifying $HADOOP_HOME/etc/hadoop/core-site.xml

```
hadoop@Ubuntu: ~/hadoop-3.3.6/etc/hadoop
  GNU nano 7.2                              core-site.xml
<?xml version="1.0" encoding="UTF-8"?>
<?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>fs.defaultFS</name>

    <value>hdfs://localhost:9000</value>

    </property>

                              [ Read 29 lines ]
^G Help       ^O Write Out   ^W Where Is   ^K Cut      ^T Execute   ^C Location   M-U Undo   M-A Set Mark
^X Exit       ^R Read File   ^\ Replace    ^U Paste    ^J Justify   ^/ Go To Line M-E Redo   M-6 Copy
```

5.   Modifying $HADOOP_HOME/etc/hadoop/hdfs-site.xml

```
hadoop@Ubuntu: ~/hadoop-3.3.6/etc/hadoop
  GNU nano 7.2                                    hdfs-site.xml
    </configuration>
-->
    <configuration>

    <property>

    <name>dfs.replication</name>

    <value>1</value>

    </property>

    <property> <name>dfs.namenode.name.dir</name>

    <value>file:///home/hadoop/hadoopdata/hdfs/namenode</value>

    </property>

    <property>

    <name>dfs.datanode.data.dir</name>

    <value>file:///home/hadoop/hadoopdata/hdfs/datanode</value>

    </property>

    </configuration>

^G Help        ^O Write Out   ^W Where Is    ^K Cut         ^T Execute     ^C Location    M-U Undo      M-A Set Mark
^X Exit        ^R Read File   ^\ Replace     ^U Paste       ^J Justify     ^/ Go To Line  M-E Redo      M-6 Copy
```

6. Modifying $HADOOP_HOME/etc/hadoop/mapred-site.xml

```
hadoop@Ubuntu: ~/hadoop-3.3.6/etc/hadoop
  GNU nano 7.2                                    mapred-site.xml
    <configuration>

    <property>

    <name>yarn.app.mapreduce.am.env</name>

    <value>HADOOP_MAPRED_HOME=$HADOOP_HOME/home/hadoop/hadoop/bin/hadoop</value>

    </property>

    <property>

    <name>mapreduce.map.env</name>

    <value>HADOOP_MAPRED_HOME=$HADOOP_HOME/home/hadoop/hadoop/bin/hadoop</value>

    </property>

    <property>

    <name>mapreduce.reduce.env</name>

    <value>HADOOP_MAPRED_HOME=$HADOOP_HOME/home/hadoop/hadoop/bin/hadoop</value>

    </property>

    </configuration>

^G Help        ^O Write Out   ^W Where Is    ^K Cut         ^T Execute     ^C Location    M-U Undo      M-A Set Mark
^X Exit        ^R Read File   ^\ Replace     ^U Paste       ^J Justify     ^/ Go To Line  M-E Redo      M-6 Copy
```

7. Modifying $HADOOP_HOME/etc/hadoop/yarn-site.xml

8. Execute hdfs namenode -format

```
2024-08-15 16:44:40,084 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.num.users = 10
2024-08-15 16:44:40,084 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.windows.minutes = 1,5,25
2024-08-15 16:44:40,121 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
2024-08-15 16:44:40,121 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total heap and retry cache entry expiry
 time is 600000 millis
2024-08-15 16:44:40,146 INFO util.GSet: Computing capacity for map NameNodeRetryCache
2024-08-15 16:44:40,150 INFO util.GSet: VM type       = 64-bit
2024-08-15 16:44:40,153 INFO util.GSet: 0.029999999329447746% max memory 871.5 MB = 267.7 KB
2024-08-15 16:44:40,153 INFO util.GSet: capacity      = 2^15 = 32768 entries
Re-format filesystem in Storage Directory root= /home/hadoop/hadoopdata/hdfs/namenode; location= null ? (Y or N) Y
2024-08-15 16:44:43,736 INFO namenode.FSImage: Allocated new BlockPoolId: BP-1847114295-127.0.1.1-1723720483718
2024-08-15 16:44:43,736 INFO common.Storage: Will remove files: [/home/hadoop/hadoopdata/hdfs/namenode/current/edits_inp
rogress_0000000000000000002, /home/hadoop/hadoopdata/hdfs/namenode/current/fsimage_0000000000000000001.md5, /home/hadoop
/hadoopdata/hdfs/namenode/current/fsimage_0000000000000000000.md5, /home/hadoop/hadoopdata/hdfs/namenode/current/VERSION
, /home/hadoop/hadoopdata/hdfs/namenode/current/fsimage_0000000000000000001, /home/hadoop/hadoopdata/hdfs/namenode/curre
nt/seen_txid, /home/hadoop/hadoopdata/hdfs/namenode/current/fsimage_0000000000000000000, /home/hadoop/hadoopdata/hdfs/na
menode/current/edits_0000000000000000001-0000000000000000001]
2024-08-15 16:44:43,917 INFO common.Storage: Storage directory /home/hadoop/hadoopdata/hdfs/namenode has been successful
ly formatted.
2024-08-15 16:44:44,225 INFO namenode.FSImageFormatProtobuf: Saving image file /home/hadoop/hadoopdata/hdfs/namenode/cur
rent/fsimage.ckpt_0000000000000000000 using no compression
2024-08-15 16:44:44,577 INFO namenode.FSImageFormatProtobuf: Image file /home/hadoop/hadoopdata/hdfs/namenode/current/fs
image.ckpt_0000000000000000000 of size 401 bytes saved in 0 seconds .
2024-08-15 16:44:44,630 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
2024-08-15 16:44:44,661 INFO namenode.FSNamesystem: Stopping services started for active state
2024-08-15 16:44:44,661 INFO namenode.FSNamesystem: Stopping services started for standby state
2024-08-15 16:44:44,675 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when meet shutdown.
2024-08-15 16:44:44,685 INFO namenode.NameNode: SHUTDOWN_MSG:
/************************************************************
SHUTDOWN_MSG: Shutting down NameNode at Ubuntu/127.0.1.1
************************************************************/
hadoop@Ubuntu:~/hadoop-3.3.6/sbin$
```

9. Execute start-dfs.sh

```
hadoop@Ubuntu:~/hadoop-3.3.6/sbin$ start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [Ubuntu]
2024-08-15 16:47:24,236 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using buil
tin-java classes where applicable
```
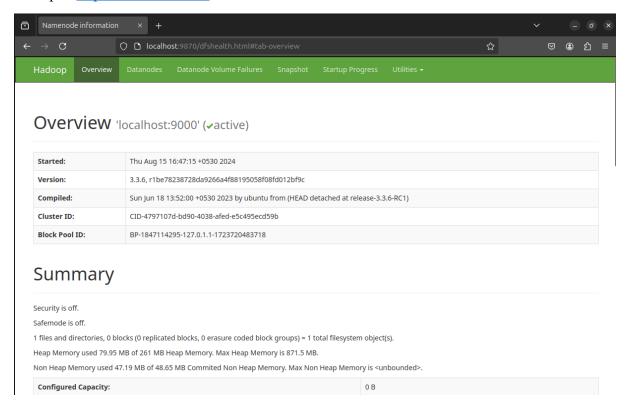
10. Execute start-yarn.sh

```
hadoop@Ubuntu:~/hadoop-3.3.6/sbin$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
hadoop@Ubuntu:~/hadoop-3.3.6/sbin$
```
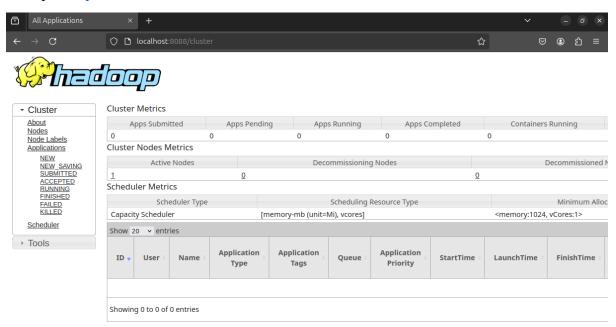
11. Execute jps command

```
hadoop@Ubuntu:~/hadoop-3.3.6/sbin$ jps
4758 NodeManager
4390 SecondaryNameNode
5111 Jps
4632 ResourceManager
4015 NameNode
```

12. Open http://localhost:9870

13. Open http://localhost:8088

14. Open http://localhost:8042