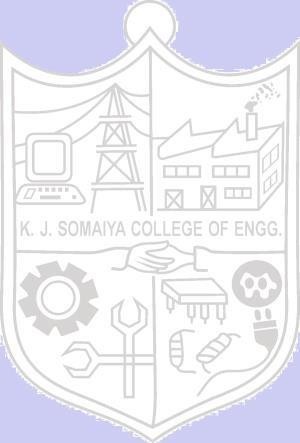
KJSCE/IT/B.Tech./SEMVII/BDA/2019-2020



**Experiment No.4**

**Title: Installation and configuration of Hadoop**

KJSCE/IT/B.Tech./SEMVII/BDA/2019-2020

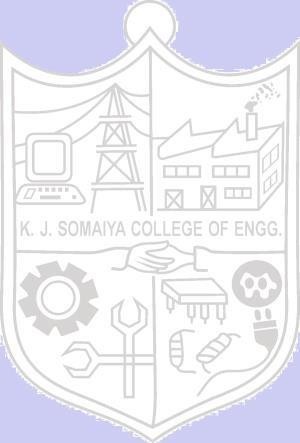
# Batch: B3 Roll No.:1724005 Experiment No.:4 Title: Installation and configuration of Hadoop

**Resources needed:** 64 bit LTS Ubuntu OS (16.04/14.04), Internet

# Theory:

Hadoop is an open-source framework that allows to store and process big data in a distributed environment across clusters of computers using simple programming models. It is designed to scale up from single servers to thousands of machines, each offering local computation and storage.

# Apache Hadoop:

The Apache Hadoop project develops open-source software for reliable, scalable, distributed computing. The Apache Hadoop software library is a framework that allows for the distributed processing of large data sets across clusters of computers using simple programming models. It is designed to scale up from single servers to thousands of machines, each offering local computation and storage. Rather than rely on hardware to deliver high- availability, the library itself is designed to detect and handle failures at the application layer, so delivering a highly-available service on top of a cluster of computers, each of which may be prone to failures.

The project includes these modules:

**Hadoop Common**: The common utilities that support the other Hadoop modules. **Hadoop Distributed File System (HDFS):** A distributed file system that provides high- throughput access to application data.

**Hadoop YARN**: A framework for job scheduling and cluster resource management. **Hadoop MapReduce**: A YARN-based system for parallel processing of large data sets. Other Hadoop-related projects at Apache include:

**Ambari**: A web-based tool for provisioning, managing, and monitoring Apache Hadoop clusters which includes support for Hadoop HDFS, Hadoop MapReduce, Hive, HCatalog, HBase, ZooKeeper, Oozie, Pig and Sqoop. Ambari also provides a dashboard for viewing cluster health such as heatmaps and ability to view MapReduce, Pig and Hive applications visually alongwith features to diagnose their performance characteristics in a user-friendly manner.

**Avro**: A data serialization system.

**Cassandra**: A scalable multi-master database with no single points of failure.

**Chukwa**: A data collection system for managing large distributed systems.

**HBase**: A scalable, distributed database that supports structured data storage for large tables. **Hive**: A data warehouse infrastructure that provides data summarization and ad hoc querying. **Mahout**: A Scalable machine learning and data mining library.

**Pig**: A high-level data-flow language and execution framework for parallel computer

**Spark**: A fast and general compute engine for Hadoop data. Spark provides a simple and expressive programming model that supports a wide range of applications, including ETL, machine learning, stream processing, and graph computation.

**Tez**: A generalized data-flow programming framework, bKuiJlSt ConE/HITa/dBo.TopecYh.A/SRENM,VwIhI/iBchDpAr/o2v0i1d9e-s2a020 powerful and flexible engine to execute an arbitrary DAG of tasks to process data for both batch and interactive use-cases. Tez is being adopted by Hive, Pig and other frameworks in the Hadoop ecosystem, and also by other commercial software (e.g. ETL tools), to replace Hadoop MapReduce as the underlying execution engine.

**ZooKeeper**: A high-performance coordination service for distributed applications.

# Procedure:

**Refer the Hadoop Installation file uploaded in writeup folder to set the Hadoop node. Results: (Document as per the format)**

**ubuntu@ubuntu-OptiPlex-3020:~$ sudo addgroup hadoop\_group**

[sudo] password for ubuntu:

Adding group `hadoop\_group' (GID 1014) ... Done.

# ubuntu@ubuntu-OptiPlex-3020:~$ sudo rm -rf /usr/local/Hadoop.

**ubuntu@ubuntu-OptiPlex-3020:~/Desktop$ cd hadoop ubuntu@ubuntu-OptiPlex-3020:~/Desktop/hadoop$ bin/hadoop** Usage: hadoop [--config confdir] [COMMAND | CLASSNAME]

CLASSNAME run the class named CLASSNAME or

where COMMAND is one of:

fs run a generic filesystem user client version print the version

jar <jar> run a jar file

note: please use "yarn jar" to launch YARN applications, not this command.

checknative [-a|-h] check native hadoop and compression libraries availability distcp <srcurl> <desturl> copy file or directories recursively

archive -archiveName NAME -p <parent path> <src>\* <dest> create a hadoop archive classpath prints the class path needed to get the

Hadoop jar and the required libraries credential interact with credential providers daemonlog get/set the log level for each daemon trace view and modify Hadoop tracing settings

Most commands print help when invoked w/o parameters.

# ubuntu@ubuntu-OptiPlex-3020:~/Desktop/hadoop$ sudo addgroup nithya

Adding group `nithya' (GID 1015) ... Done.

# ubuntu@ubuntu-OptiPlex-3020:~/Desktop/hadoop$ sudo adduser --ingroup nithya file1

Adding user `file1' ...

Adding new user `file1' (1009) with group `nithya' ... Creating home directory `/home/file1' ...

Copying files from `/etc/skel' ... Enter new UNIX password: Retype new UNIX password:

passwd: password updated successfully

Changing the user information for file1

Enter the new value, or press ENTER for the default Full Name []: Nithya Mani

Room Number []:

KJSCE/IT/B.Tech./SEMVII/BDA/2019-2020

Work Phone []: Home Phone []: Other []:

Is the information correct? [Y/n] y

KJSCE/IT/B.Tech./SEMVII/BDA/2019-2020

ubuntu@ubuntu-OptiPlex-3020:~/Desktop/hadoop$ sudo adduser file1 sudo Adding user `file1' to group `sudo' ...

Adding user file1 to group sudo Done.

# ubuntu@ubuntu-OptiPlex-3020:~/Desktop/hadoop$ su file1

Password:

file1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$

# file1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$ sudo apt-get install apache2

[sudo] password for file1: Reading package lists... Done Building dependency tree Reading state information... Done

apache2 is already the newest version.

The following packages were automatically installed and are no longer required: libboost-filesystem1.54.0 libboost-program-options1.54.0

libboost-thread1.54.0 libgoogle-perftools4 libpcrecpp0 libsnappy1 libtcmalloc-minimal4 libunwind8

Use 'apt-get autoremove' to remove them.

0 upgraded, 0 newly installed, 0 to remove and 417 not upgraded.

# file1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$ sudo service apache2 start

\* Starting web server apache2 \*

file1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$ sudo apt-get install apache2 Reading package lists... Done

Building dependency tree Reading state information... Done

apache2 is already the newest version.

The following packages were automatically installed and are no longer required: libboost-filesystem1.54.0 libboost-program-options1.54.0

libboost-thread1.54.0 libgoogle-perftools4 libpcrecpp0 libsnappy1 libtcmalloc-minimal4 libunwind8

Use 'apt-get autoremove' to remove them.

0 upgraded, 0 newly installed, 0 to remove and 417 not upgraded.

file1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$ apt-get autoremove

E: Could not open lock file /var/lib/dpkg/lock - open (13: Permission denied) E: Unable to lock the administration directory (/var/lib/dpkg/), are you root? file1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$ clear

# file1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$ sudo apt-get install apache2

Reading package lists... Done Building dependency tree Reading state information... Done

apache2 is already the newest version.

The following packages were automatically installed and are no longer required: libboost-filesystem1.54.0 libboost-program-options1.54.0

libboost-thread1.54.0 libgoogle-perftools4 libpcrecpp0 libsnappy1 libtcmalloc-minimal4 libunwind8

Use 'apt-get autoremove' to remove them.

KJSCE/IT/B.Tech./SEMVII/BDA/2019-2020

0 upgraded, 0 newly installed, 0 to remove and 417 not upgraded.

file1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$ sudo service apache2 start

\* Starting web server apache2

\*

file1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$ sudo apt-get install default-jdk Reading package lists... Done

Building dependency tree Reading state information... Done

The following packages were automatically installed and are no longer required: libboost-filesystem1.54.0 libboost-program-options1.54.0

libboost-thread1.54.0 libgoogle-perftools4 libpcrecpp0 libsnappy1 libtcmalloc-minimal4 libunwind8

Use 'apt-get autoremove' to remove them.

The following extra packages will be installed:

ca-certificates-java default-jre default-jre-headless fonts-dejavu-extra

java-common libatk-wrapper-java libatk-wrapper-java-jni libgif4 libice-dev libpthread-stubs0-dev libsctp1 libsm-dev libx11-dev libx11-doc libxau-dev libxcb1-dev libxdmcp-dev libxt-dev lksctp-tools openjdk-7-jdk openjdk-7-jre openjdk-7-jre-headless tzdata tzdata-java x11proto-core-dev

x11proto-input-dev x11proto-kb-dev xorg-sgml-doctools xtrans-dev Suggested packages:

equivs libice-doc libsm-doc libxcb-doc libxt-doc openjdk-7-demo openjdk-7-source visualvm icedtea-7-plugin icedtea-7-jre-jamvm

sun-java6-fonts fonts-ipafont-gothic fonts-ipafont-mincho fonts-wqy-microhei fonts-wqy-zenhei ttf-indic-fonts

The following NEW packages will be installed:

ca-certificates-java default-jdk default-jre default-jre-headless

fonts-dejavu-extra java-common libatk-wrapper-java libatk-wrapper-java-jni libgif4 libice-dev libpthread-stubs0-dev libsctp1 libsm-dev libx11-dev libx11-doc libxau-dev libxcb1-dev libxdmcp-dev libxt-dev lksctp-tools openjdk-7-jdk openjdk-7-jre openjdk-7-jre-headless tzdata-java

x11proto-core-dev x11proto-input-dev x11proto-kb-dev xorg-sgml-doctools xtrans-dev

The following packages will be upgraded: tzdata

1 upgraded, 29 newly installed, 0 to remove and 416 not upgraded. Need to get 61.8 MB of archives.

After this operation, 104 MB of additional disk space will be used. Do you want to continue? [Y/n] y

Ge[t:1 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main ca-certificates-java all 20130815ubuntu1 [13.4 kB]

Ge[t:2 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty-updates/main tzdata all 2017c-0ubuntu0.14.04 [166 kB]

Ge[t:3 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty-updates/main tzdata-java all 2017c- 0ubuntu0.14.04 [70.1 kB]

Ge[t:4 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main java-common all 0.51 [130 kB]

Ge[t:5 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main libsctp1 amd64 1.0.15+dfsg-1 [9,226 B]

Er[r http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty-updates/main openjdk-7-jre-headless amd64 7u181- 2.6.14-0ubuntu0.1

404 Not Found [IP: 91.189.88.162 80]

Ge[t:6 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main default-jre-headless amd64 2:1.7-51 [3,834

B] KJSCE/IT/B.Tech./SEMVII/BDA/2019-2020

Er[r http://security.ubuntu.com/ubuntu/](http://security.ubuntu.com/ubuntu/) trusty-security/main openjdk-7-jre-headless amd64 7u181- 2.6.14-0ubuntu0.1

404 Not Found [IP: 91.189.88.149 80]

Ge[t:7 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main libgif4 amd64 4.1.6-11 [28.6 kB]

Er[r http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty-updates/main openjdk-7-jre amd64 7u181-2.6.14- 0ubuntu0.1

404 Not Found [IP: 91.189.88.161 80]

Ge[t:8 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main default-jre amd64 2:1.7-51 [940 B]

Er[r http://security.ubuntu.com/ubuntu/](http://security.ubuntu.com/ubuntu/) trusty-security/main openjdk-7-jre amd64 7u181-2.6.14- 0ubuntu0.1

404 Not Found [IP: 91.189.88.149 80]

Ge[t:9 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main libatk-wrapper-java all 0.30.4-4 [30.2 kB] Get:10 <http://in.archive.ubuntu.com/ubuntu/>trusty/main libatk-wrapper-java-jni amd64 0.30.4-4 [25.2 kB]

Er[r http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty-updates/main openjdk-7-jdk amd64 7u181-2.6.14- 0ubuntu0.1

404 Not Found [IP: 91.189.88.152 80]

Er[r http://security.ubuntu.com/ubuntu/](http://security.ubuntu.com/ubuntu/) trusty-security/main openjdk-7-jdk amd64 7u181-2.6.14- 0ubuntu0.1

404 Not Found [IP: 91.189.88.149 80]

Ge[t:11 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main default-jdk amd64 2:1.7-51 [934 B] Get:12 <http://in.archive.ubuntu.com/ubuntu/>trusty/main fonts-dejavu-extra all 2.34-1ubuntu1 [1,736 kB]

Ge[t:13 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main xorg-sgml-doctools all 1:1.11-1 [12.9 kB] Get:14 <http://in.archive.ubuntu.com/ubuntu/>trusty-updates/main x11proto-core-dev all 7.0.26- 1~ubuntu2 [700 kB]

Ge[t:15 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main libice-dev amd64 2:1.0.8-2 [57.6 kB] Get:16 <http://in.archive.ubuntu.com/ubuntu/>trusty/main libpthread-stubs0-dev amd64 0.3-4 [4,068 B]

Ge[t:17 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main libsm-dev amd64 2:1.2.1-2 [19.9 kB] Get:18 <http://in.archive.ubuntu.com/ubuntu/>trusty/main libxau-dev amd64 1:1.0.8-1 [11.1 kB] Get:19 <http://in.archive.ubuntu.com/ubuntu/>trusty/main libxdmcp-dev amd64 1:1.1.1-1 [26.9 kB] Get:20 <http://in.archive.ubuntu.com/ubuntu/>trusty/main x11proto-input-dev all 2.3-1 [139 kB] Get:21 <http://in.archive.ubuntu.com/ubuntu/>trusty/main x11proto-kb-dev all 1.0.6-2 [269 kB] Get:22 <http://in.archive.ubuntu.com/ubuntu/>trusty-updates/main xtrans-dev all 1.3.5- 1~ubuntu14.04.2 [70.7 kB]

Ge[t:23 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main libxcb1-dev amd64 1.10-2ubuntu1 [76.6 kB]

Ge[t:24 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main libx11-dev amd64 2:1.6.2-1ubuntu2 [629 kB]

Ge[t:25 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main libx11-doc all 2:1.6.2-1ubuntu2 [1,448 kB] Get:26 <http://in.archive.ubuntu.com/ubuntu/>trusty/main libxt-dev amd64 1:1.1.4-1 [455 kB]

Ge[t:27 http://in.archive.ubuntu.com/ubuntu/](http://in.archive.ubuntu.com/ubuntu/) trusty/main lksctp-tools amd64 1.0.15+dfsg-1 [51.3 kB]

Fetched 6,185 kB in 9s (632 kB/s)

E: Failed to fetc[h http://security.ubuntu.com/ubuntu/pool/main/o/openjdk-7/openjdk-7-jre-](http://security.ubuntu.com/ubuntu/pool/main/o/openjdk-7/openjdk-7-jre-) headless\_7u181-2.6.14-0ubuntu0.1\_amd64.deb 404 Not Found [IP: 91.189.88.149 80]

E: Failed to fetc[h http://security.ubuntu.com/ubuntu/pool/main/o/openjdk-7/openjdk-7-jre\_7u181-](http://security.ubuntu.com/ubuntu/pool/main/o/openjdk-7/openjdk-7-jre_7u181-) 2.6.14-0ubuntu0.1\_amd64.deb 404 Not Found [IP: 91.189.88.149 80]

E: Failed to fetc[h http://security.ubuntu.com/ubuntu/poolK/mJSaCinE//oIT/o/Bpe.Tnejdchk.-/7S/EoMpeVnIjId/Bk-D7A-j/d2k0\_179u-2108210-](http://security.ubuntu.com/ubuntu/pool/main/o/openjdk-7/openjdk-7-jdk_7u181-) 2.6.14-0ubuntu0.1\_amd64.deb 404 Not Found [IP: 91.189.88.149 80]

E: Unable to fetch some archives, maybe run apt-get update or try with --fix-missing? file1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$ sudo apt-get install ssh Reading package lists... Done

Building dependency tree Reading state information... Done ssh is already the newest version.

The following packages were automatically installed and are no longer required: libboost-filesystem1.54.0 libboost-program-options1.54.0

libboost-thread1.54.0 libgoogle-perftools4 libpcrecpp0 libsnappy1 libtcmalloc-minimal4 libunwind8

Use 'apt-get autoremove' to remove them.

0 upgraded, 0 newly installed, 0 to remove and 417 not upgraded.

# file1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$ sudo nautilus

(nautilus:6114): Gtk-WARNING \*\*: Failed to register client: GDBus.Error:org.freedesktop.DBus.Error.ServiceUnknown: The name org.gnome.SessionManager was not provided by any .service files

Nautilus-Share-Message: Called "net usershare info" but it failed: 'net usershare' returned error 255: net usershare: cannot open usershare directory /var/lib/samba/usershares. Error No such file or directory

Please ask your system administrator to enable user sharing.

# file1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$ sudo mkdir -p

**/var/lib/samba/usershares/**

f**ile1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$ su file1 Password:**

**file1@ubuntu-OptiPlex-3020:/home/ubuntu/Desktop/hadoop$ cd file1@ubuntu-OptiPlex-3020:~$ sudo gedit ~/.bashrc**

file1@ubuntu-OptiPlex-3020:~$ exec bash file1@ubuntu-OptiPlex-3020:~$ source ~/.bashrc file1@ubuntu-OptiPlex-3020:~$ source gedit ~/.bashrc bash: source: /usr/bin/gedit: cannot execute binary file

file1@ubuntu-OptiPlex-3020:~$ cd file:///home/ubuntu/Desktop/hadoop/etc/hadoop bash: cd: file:///home/ubuntu/Desktop/hadoop/etc/hadoop: No such file or directory file1@ubuntu-OptiPlex-3020:~$ cd Desktop/hadoop/etc/hadoop

bash: cd: Desktop/hadoop/etc/hadoop: No such file or directory file1@ubuntu-OptiPlex-3020:~$ cd Desktop

file1@ubuntu-OptiPlex-3020:~/Desktop$ cd hadoop bash: cd: hadoop: No such file or directory file1@ubuntu-OptiPlex-3020:~/Desktop$ cd hadoop bash: cd: hadoop: No such file or directory file1@ubuntu-OptiPlex-3020:~/Desktop$ cd /hadoop bash: cd: /hadoop: No such file or directory file1@ubuntu-OptiPlex-3020:~/Desktop$ clear

file1@ubuntu-OptiPlex-3020:~/Desktop$ cd hadoop bash: cd: hadoop: No such file or directory file1@ubuntu-OptiPlex-3020:~/Desktop$ cd

KJSCE/IT/B.Tech./SEMVII/BDA/2019-2020

file1@ubuntu-OptiPlex-3020:~$ cd /usr/local/hadoop-2.5.2/etc/hadoop file1@ubuntu-OptiPlex-3020:~$ cd

file1@ubuntu-OptiPlex-3020:~$ cd /usr file1@ubuntu-OptiPlex-3020:/usr$ cd /local bash: cd: /local: No such file or directory file1@ubuntu-OptiPlex-3020:/usr$ cd local

file1@ubuntu-OptiPlex-3020:/usr/local$ cd hadoop bash: cd: hadoop: No such file or directory file1@ubuntu-OptiPlex-3020:/usr/local$ cd file1@ubuntu-OptiPlex-3020:~$ dir

Desktop examples.desktop hadoop file1@ubuntu-OptiPlex-3020:~$ cd /usr file1@ubuntu-OptiPlex-3020:/usr$ cd /local bash: cd: /local: No such file or directory file1@ubuntu-OptiPlex-3020:/usr$ dir

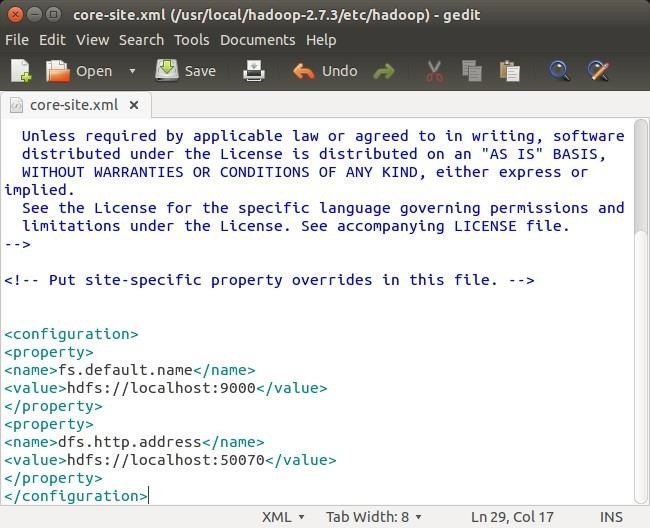
bin games hadoop-2.7.3 hadoop-2.7.3.tar.gz include lib local sbin share src file1@ubuntu-OptiPlex-3020:/usr$ cd local

file1@ubuntu-OptiPlex-3020:/usr/local$ dir

bin etc games go hadoop-2.7.3 hadoop-2.7.31 hadoop-2.7.3.tar.gz hadoop\_store include lib man sbin share src

file1@ubuntu-OptiPlex-3020:/usr/local$ cd hadoop-2.7.3 file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3$ cd etc file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3/etc$ cd hadoop

# file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3/etc/hadoop$ sudo gedit core-site.xml



KJSCE/IT/B.Tech./SEMVII/BDA/2019-2020

**file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3$ sudo mkdir -p**

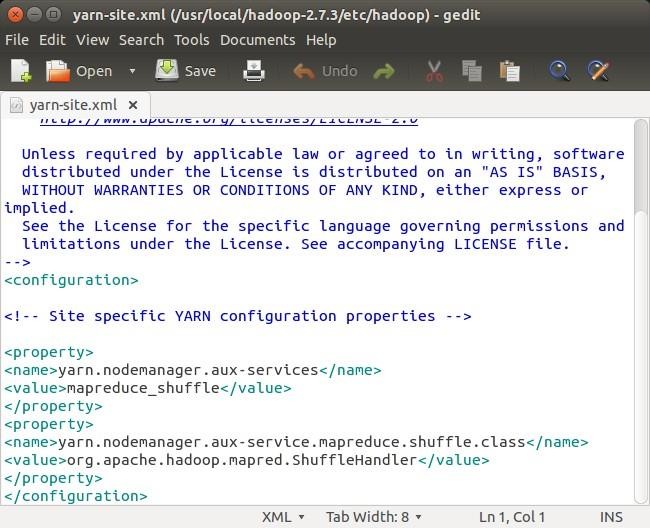
**/hadoop2\_data/hdfs/datanode**

file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3$ cd hadoop2\_data file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3/hadoop2\_data$ dir hdfs

file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3/hadoop2\_data$ cd .. file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3$ sudo chown -R file1:hadoop

/usr/local/hadoop-2.7.3/hadoop2\_data

# file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3$ $sudo gedit hdfs-site.xml

KJSCE/IT/B.Tech./SEMVII/BDA/2019-2020

# file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3$ sudo gedit mapred-site.xml

file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3$ cd /usr/local/hadoop-2.7.3/bin/ file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3/bin$ sudo ./hadoop namenode -format DEPRECATED: Use of this script to execute hdfs command is deprecated.

Instead use the hdfs command for it.

18/08/13 20:33:43 INFO namenode.NameNode: STARTUP\_MSG:

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* STARTUP\_MSG: Starting NameNode

STARTUP\_MSG: host = ubuntu-OptiPlex-3020/127.0.1.1 STARTUP\_MSG: args = [-format]

STARTUP\_MSG: version = 2.7.3

# file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3/bin$ cd /usr/local/hadoop-2.7.3/sbin/ file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3/sbin$ sudo ./hadoop-daemon.sh start datanode

**starting datanode, logging to /usr/local/hadoop-2.7.3/logs/hadoop-root-datanode-ubuntu-**

**OptiPlex-3020.out**

**file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3/sbin$ sudo ./hadoop-daemon.sh start**

**namenode**

KJSCE/IT/B.Tech./SEMVII/BDA/2019-2020

# starting namenode, logging to /usr/local/hadoop-2.7.3/logs/hadoop-root-namenode-ubuntu-

**OptiPlex-3020.out**

**file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3/sbin$ sudo ./hadoop-daemon.sh start namenode**

**file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3/sbin$ sudo ./yarn-daemon.sh start resourcemanager**

**starting resourcemanager, logging to /usr/local/hadoop-2.7.3/logs/yarn-root-resourcemanager- ubuntu-OptiPlex-3020.out**

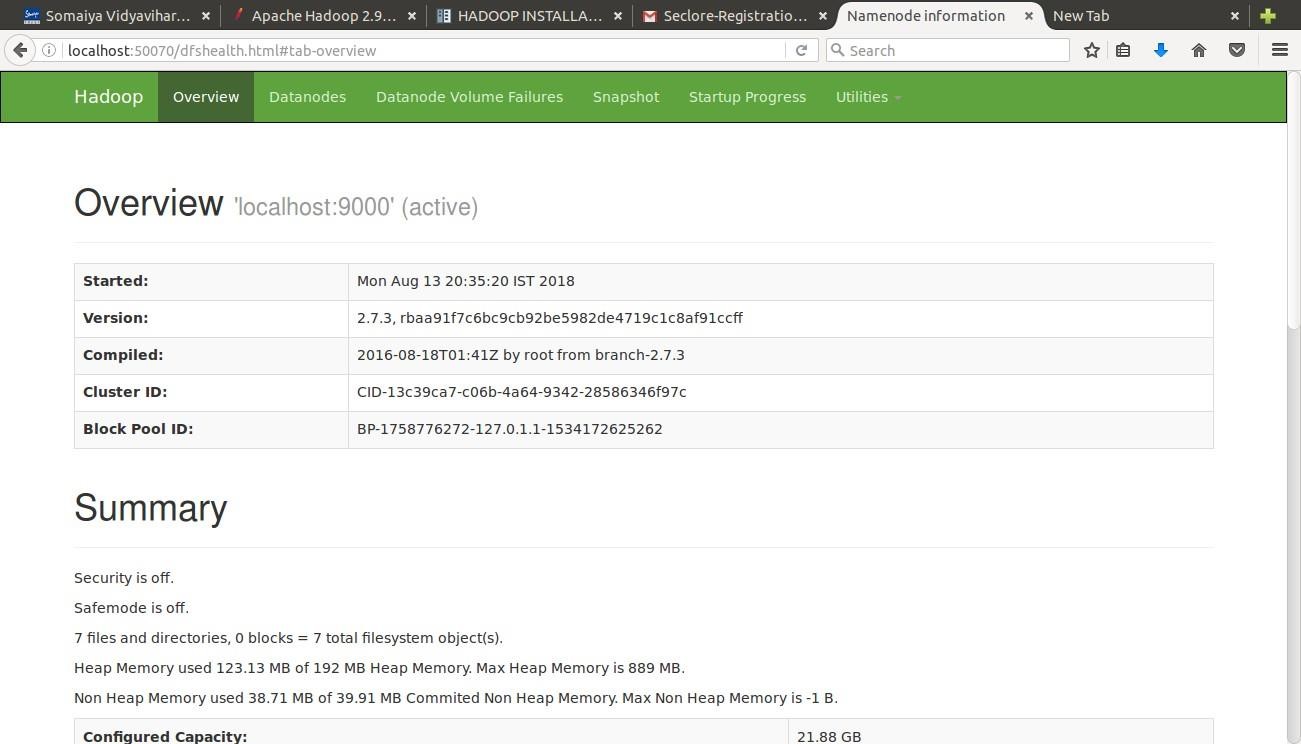
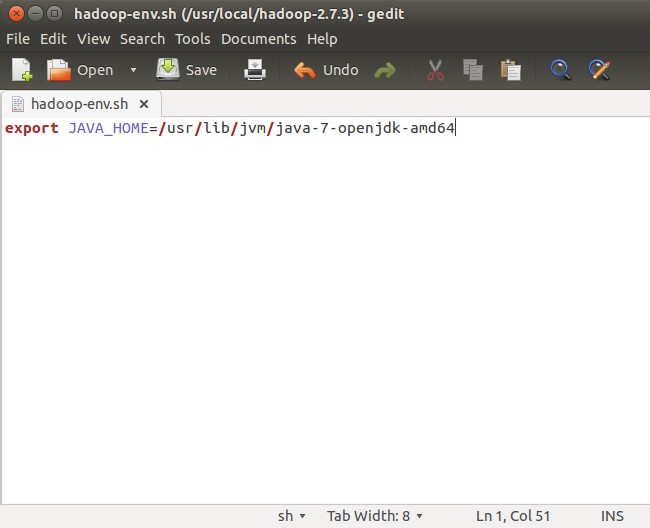
**file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3/sbin$ sudo ./yarn-daemon.sh start nodemanager**

starting nodemanager, logging to /usr/local/hadoop-2.7.3/logs/yarn-root-nodemanager-ubuntu-

OptiPlex-3020.out

file1@ubuntu-OptiPlex-3020:/usr/local/hadoop-2.7.3/sbin$ sudo ./mr-jobhistory-daemon.sh start historyserver

KJSCE/IT/B.Tech./SEMVII/BDA/2019-2020



# Outcomes:

Installation of Hadoop in Ubuntu.

# Conclusion:

We have installed Hadoop on Ubuntu successfully.

# Grade: AA / AB / BB / BC / CC / CD /DD Signature of faculty in-charge with date

**References:**

**Books/ Journals/ Websites:**

1. Judith Hurwitz, Alan Nugent, Dr. Fern Halper, Marcia Kaufman, “Big Data for Dummies”, Wiley India
2. McCreary and Ann Kelly “Making Sense of NoSQL –A guide for managers and the rest of us”, Manning Press

KJSCE/IT/B.Tech./SEMVII/BDA/2019-2020

1. https://docs.mongodb.com/getting-started/shell/import-data/