Big Data (ECE 595-004/007), Fall 2018
Hands-on 1, Part2: Hadoop Set-up (15 pts)
Due date: September 21, 2018, 11:59 PM

This assignment is straightforward. You have to set-up Hadoop in your virtual machine. For grading, submit the screenshot of the steps that you followed for set-up in one word file and upload the PDF for it.

Virtual machine

The tutorial for VMWare player has been given in the tutorial document.

Those who have Mac, they can install Virtual Box and create a virtual machine using it.
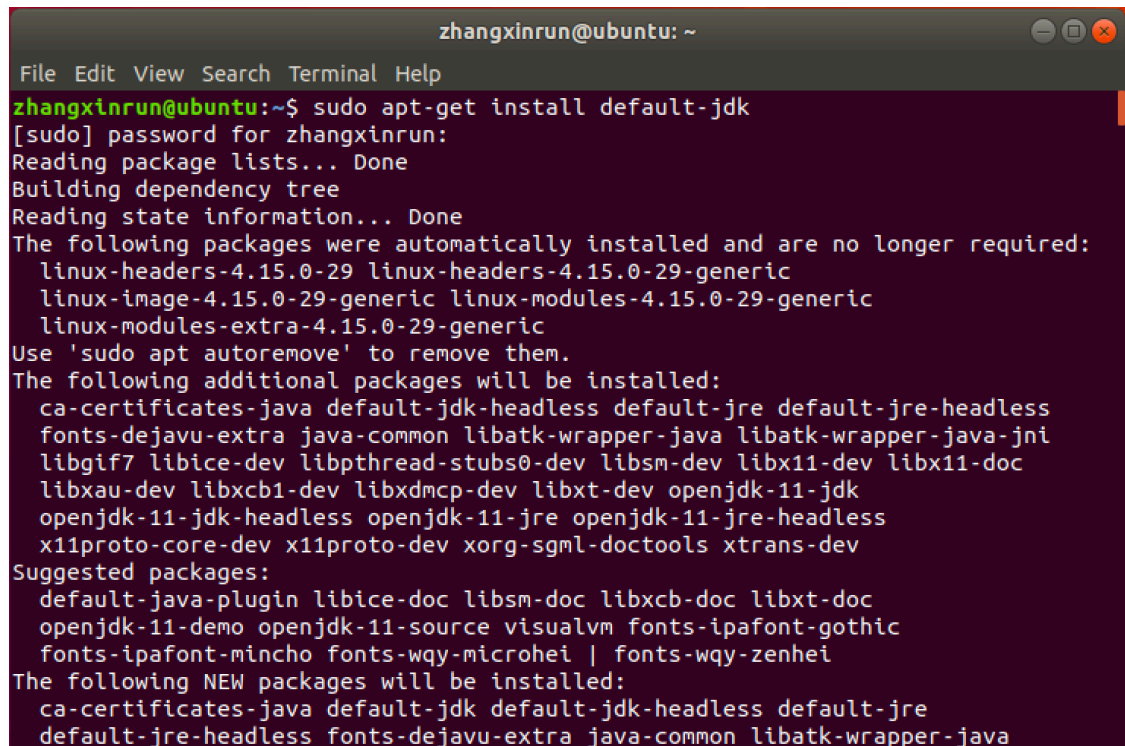
Link for Virtual Box installation: https://www.youtube.com/watch?v=lEvM-No4eQo

Download Ubuntu 16: http://releases.ubuntu.com/16.04/ubuntu-16.04.5-desktop-amd64.iso if your machine is 64-bit or http://releases.ubuntu.com/16.04/ubuntu-16.04.5-desktop-i386.iso if your machine is 32-bit.

Create Ubuntu virtual machine: https://www.youtube.com/watch?v=fh8OdDd0K30

Hadoop set-up

1. Install Java: **`sudo apt-get install default-jdk`**
    `A few steps will be executed.`

```
                        zhangxinrun@ubuntu: ~

 File Edit View Search Terminal Help
Adding debian:DigiCert_Assured_ID_Root_G3.pem
Adding debian:IdenTrust_Commercial_Root_CA_1.pem
Adding debian:AC_RAIZ_FNMT-RCM.pem
Adding debian:Certigna.pem
Adding debian:Autoridad_de_Certificacion_Firmaprofesional_CIF_A62634068.pem
Adding debian:SecureSign_RootCA11.pem
Adding debian:USERTrust_ECC_Certification_Authority.pem
done.
Setting up default-jre (2:1.10-63ubuntu1~02) ...
Setting up openjdk-11-jdk:amd64 (10.0.2+13-1ubuntu0.18.04.2) ...
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/appletviewer t
o provide /usr/bin/appletviewer (appletviewer) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jconsole to pr
ovide /usr/bin/jconsole (jconsole) in auto mode
Setting up default-jdk (2:1.10-63ubuntu1~02) ...
Processing triggers for libc-bin (2.27-3ubuntu1) ...
Processing triggers for ca-certificates (20180409) ...
Updating certificates in /etc/ssl/certs...
0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...

done.
done.
zhangxinrun@ubuntu:~$
```

Confirm the installation： `java -version`

```
zhangxinrun@ubuntu:~$ java -version
openjdk version "10.0.2" 2018-07-17
OpenJDK Runtime Environment (build 10.0.2+13-Ubuntu-1ubuntu0.18.04.2)
OpenJDK 64-Bit Server VM (build 10.0.2+13-Ubuntu-1ubuntu0.18.04.2, mixed mode)
zhangxinrun@ubuntu:~$
```

2. Install SSH server: `sudo apt-get install openssh-server`

```
zhangxinrun@ubuntu:~$ sudo apt-get install openssh-server
[sudo] password for zhangxinrun:
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  ncurses-term openssh-sftp-server ssh-import-id
Suggested packages:
  molly-guard monkeysphere rssh ssh-askpass
The following NEW packages will be installed:
  ncurses-term openssh-server openssh-sftp-server ssh-import-id
0 upgraded, 4 newly installed, 0 to remove and 17 not upgraded.
Need to get 637 kB of archives.
After this operation, 5,316 kB of additional disk space will be used.
Do you want to continue? [Y/n] y
Get:1 http://us.archive.ubuntu.com/ubuntu bionic-updates/main amd64 ncurses-term
 all 6.1-1ubuntu1.18.04 [248 kB]
Get:2 http://us.archive.ubuntu.com/ubuntu bionic/main amd64 openssh-sftp-server
amd64 1:7.6p1-4 [45.5 kB]
Get:3 http://us.archive.ubuntu.com/ubuntu bionic/main amd64 openssh-server amd64
 1:7.6p1-4 [332 kB]
Get:4 http://us.archive.ubuntu.com/ubuntu bionic-updates/main amd64 ssh-import-i
d all 5.7-0ubuntu1.1 [10.9 kB]
```

```
Setting up ncurses-term (6.1-1ubuntu1.18.04) ...
Processing triggers for ufw (0.35-5) ...
Processing triggers for ureadahead (0.100.0-20) ...
Setting up openssh-sftp-server (1:7.6p1-4) ...
Processing triggers for systemd (237-3ubuntu10.3) ...
Processing triggers for man-db (2.8.3-2) ...
Setting up ssh-import-id (5.7-0ubuntu1.1) ...
Setting up openssh-server (1:7.6p1-4) ...

Creating config file /etc/ssh/sshd_config with new version
Creating SSH2 RSA key; this may take some time ...
2048 SHA256:r/C6NeMDycqmmnc0OhWvMZ1/k7MogD943cw35fzNQGs root@ubuntu (RSA)
Creating SSH2 ECDSA key; this may take some time ...
256 SHA256:hSMC8tIKtRYW95zn39inBsWlNeJcLfJVQXNVwab5teg root@ubuntu (ECDSA)
Creating SSH2 ED25519 key; this may take some time ...
256 SHA256:aMp6rJpLiHqFXU3LYKOjH503N1/Ez8PBHVTd9enEIYk root@ubuntu (ED25519)
Created symlink /etc/systemd/system/sshd.service → /lib/systemd/system/ssh.servi
ce.
Created symlink /etc/systemd/system/multi-user.target.wants/ssh.service → /lib/s
ystemd/system/ssh.service.
Processing triggers for ureadahead (0.100.0-20) ...
Processing triggers for systemd (237-3ubuntu10.3) ...
Processing triggers for ufw (0.35-5) ...
zhangxinrun@ubuntu:~$
```

3. Download Hadoop 2.9.1 from the following link:
   https://www.apache.org/dyn/closer.cgi/hadoop/common/hadoop-2.9.1/hadoop-2.9.1.tar.gz
4. **Go to the Downloads directory: `cd Downloads`**
5. **Extract Hadoop files from the downloaded file: `tar xzvf hadoop-2.9.1.tar.gz`**

```
hadoop-2.9.1/share/doc/hadoop/images/expanded.gif
hadoop-2.9.1/share/doc/hadoop/images/apache-maven-project-2.png
hadoop-2.9.1/share/doc/hadoop/images/icon_error_sml.gif
hadoop-2.9.1/share/doc/hadoop/images/logo_apache.jpg
hadoop-2.9.1/share/doc/hadoop/images/icon_warning_sml.gif
hadoop-2.9.1/share/doc/hadoop/images/newwindow.png
hadoop-2.9.1/share/doc/hadoop/images/maven-logo-2.gif
hadoop-2.9.1/share/doc/hadoop/images/banner.jpg
hadoop-2.9.1/share/doc/hadoop/images/logo_maven.jpg
hadoop-2.9.1/share/doc/hadoop/images/collapsed.gif
hadoop-2.9.1/share/doc/hadoop/images/external.png
hadoop-2.9.1/share/doc/hadoop/images/h5.jpg
hadoop-2.9.1/share/doc/hadoop/index.html
hadoop-2.9.1/share/doc/hadoop/project-reports.html
hadoop-2.9.1/include/
hadoop-2.9.1/include/hdfs.h
hadoop-2.9.1/include/Pipes.hh
hadoop-2.9.1/include/TemplateFactory.hh
hadoop-2.9.1/include/StringUtils.hh
hadoop-2.9.1/include/SerialUtils.hh
hadoop-2.9.1/LICENSE.txt
hadoop-2.9.1/NOTICE.txt
hadoop-2.9.1/README.txt
zhangxinrun@ubuntu:~/Downloads$
```

6. Move extracted Hadoop directory to /usr/local directory:
   `Sudo mv Hadoop-2.9.1 /usr/local/hadoop`

Change ownership of Hadoop directory:

`sudo chown -R bigdata:bigdata /usr/local/hadoop`

Note: `bigdata` is a user here. It may be different in your case.

```
zhangxinrun@ubuntu:~/Downloads$ sudo chown -R zhangxinrun:zhangxinrun /usr/lo
cal/hadoop
zhangxinrun@ubuntu:~/Downloads$
```

7. Find directory for Java: `readlink -f /usr/bin/java | sed "s:/bin/java::"`

The command output will be the directory for Java.

```
zhangxinrun@ubuntu:~/Downloads$  readlink -f /usr/bin/java | sed "s:/bin/java
::"
/usr/lib/jvm/java-11-openjdk-amd64
```

8. Go to home directory: `cd`

9. Open .bashrc file: `gedit .bashrc`

Add following lines:

`export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64/jre`
`export HADOOP_HOME=/usr/local/hadoop`
`export PATH=$PATH:$HADOOP_HOME/bin:$HADOOP_HOME/sbin`

Note: Value for JAVA_HOME is the output of the previous command for finding Java directory

Close bashrc file and execute command: `. .bashrc`

```
export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64
export HADOOP_HOME=/usr/local/hadoop
export PATH=$PATH:$HADOOP_HOME/bin:$HADOOP_HOME/sbin
```

```
zhangxinrun@ubuntu:~$ . .bashrc
zhangxinrun@ubuntu:~$
```

You can verify the changes with the command: `echo $JAVA_HOME`

```
zhangxinrun@ubuntu:~$ echo $JAVA_HOME
/usr/lib/jvm/java-11-openjdk-amd64
zhangxinrun@ubuntu:~$
```

10. Open /usr/local/hadoop/etc/hadoop/hadoop-env.sh:

`gedit /usr/local/hadoop/etc/hadoop/hadoop-env.sh`

Replace `export JAVA_HOME=${JAVA_HOME}` with
`export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64`
Also add this: `export HADOOP_OPTS=-Djava.net.preferIPv4Stack=true`
Close the file

```
# Set Hadoop-specific environment variables here.

# The only required environment variable is JAVA_HOME.  All others are
# optional.  When running a distributed configuration it is best to
# set JAVA_HOME in this file, so that it is correctly defined on
# remote nodes.

# The java implementation to use.
export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64

# The maximum amount of heap to use, in MB. Default is 1000.
#export HADOOP_HEAPSIZE=
#export HADOOP_NAMENODE_INIT_HEAPSIZE=""

# Enable extra debugging of Hadoop's JAAS binding, used to set up
# Kerberos security.
# export HADOOP_JAAS_DEBUG=true

# Extra Java runtime options.  Empty by default.
# For Kerberos debugging, an extended option set logs more invormation
# export HADOOP_OPTS="-Djava.net.preferIPv4Stack=true -Dsun.security.krb5.debug=true -
Dsun.security.spnego.debug"
export HADOOP_OPTS=-Djava.net.preferIPv4Stack=true
```

11. Create a directory /app/hadoop/tmp: `sudo mkdir -p /app/hadoop/tmp`
    Change ownership: `sudo chown -R bigdata:bigdata /app/hadoop/tmp`

```
zhangxinrun@ubuntu:~$ sudo mkdir -p /app/hadoop/tmp
zhangxinrun@ubuntu:~$ sudo chown -R zhangxinrun:zhangxinrun /app/hadoop/tmp
zhangxinrun@ubuntu:~$
```

12. Open /usr/local/hadoop/etc/hadoop/core-site.xml:
    `gedit /usr/local/hadoop/etc/hadoop/core-site.xml`
    Add these lines inside <configuration> </configuration> tags
    `<property>`
    `<name>hadoop.tmp.dir</name>`
    `<value>/app/hadoop/tmp</value>`
    `</property>`
    `<property>`
    `<name>fs.default.name</name>`
    `<value>hdfs://localhost:9000</value>`
    `</property>`
    Close the file

```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>hadoop.tmp.dir</name>
<value>/app/hadoop/tmp</value>
</property>
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:9000</value>
</property>
</configuration>
```

13. Open /usr/local/hadoop/etc/hadoop/hdfs-site.xml:
    gedit /usr/local/hadoop/etc/hadoop/hdfs-site.xml
    Add these lines inside <configuration> </configuration> tags
    <property>
    <name>dfs.replication</name>
    <value>1</value>
    </property>
    Close the file.

```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>dfs.replication</name>
<value>1</value>
</property>
</configuration>
```

```
gedit /usr/local/hadoop/etc/hadoop/mapred-site.xml
```
Add these lines inside \<configuration\> \</configuration\> tags
```
<property>
<name>mapreduce.framework.name</name>
<value>yarn</value>
</property>
```
Note: mapred-site.xml is not available. Rename mapred-site.xml.template file to mapred-site.xml file: mv mapred-site.xml.template mapred-site.xml

Close the file.

```xml
<?xml version="1.0"?>
<?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>mapreduce.framework.name</name>
<value>yarn</value>
</property>
</configuration>
```

```
zhangxinrun@ubuntu:~$ cd /usr/local/hadoop/etc/hadoop
zhangxinrun@ubuntu:/usr/local/hadoop/etc/hadoop$ ls
capacity-scheduler.xml          kms-acls.xml
configuration.xsl               kms-env.sh
container-executor.cfg          kms-log4j.properties
core-site.xml                   kms-site.xml
hadoop-env.cmd                  log4j.properties
hadoop_env.sh                   mapred-env.cmd
hadoop-env.sh                   mapred-env.sh
hadoop-metrics2.properties      mapred-queues.xml.template
hadoop-metrics.properties       mapred-site.xml.template
hadoop-policy.xml               slaves
hdfs-site.xml                   ssl-client.xml.example
httpfs-env.sh                   ssl-server.xml.example
httpfs-log4j.properties         yarn-env.cmd
httpfs-signature.secret         yarn-env.sh
httpfs-site.xml                 yarn-site.xml
```

15. Open /usr/local/hadoop/etc/hadoop/yarn-site.xml:
```
gedit /usr/local/hadoop/etc/hadoop/yarn-site.xml
```
Add these lines inside \<configuration\> \</configuration\> tags
```
<property>
<name>yarn.nodemanager.aux-services</name>
<value>mapreduce_shuffle</value>
```

```xml
<?xml version="1.0"?>
<!--
  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.
-->
<configuration>

<!-- Site specific YARN configuration properties -->
<property>
<name>yarn.nodemanager.aux-services</name>
<value>mapreduce_shuffle</value>
</property>

</configuration>
```
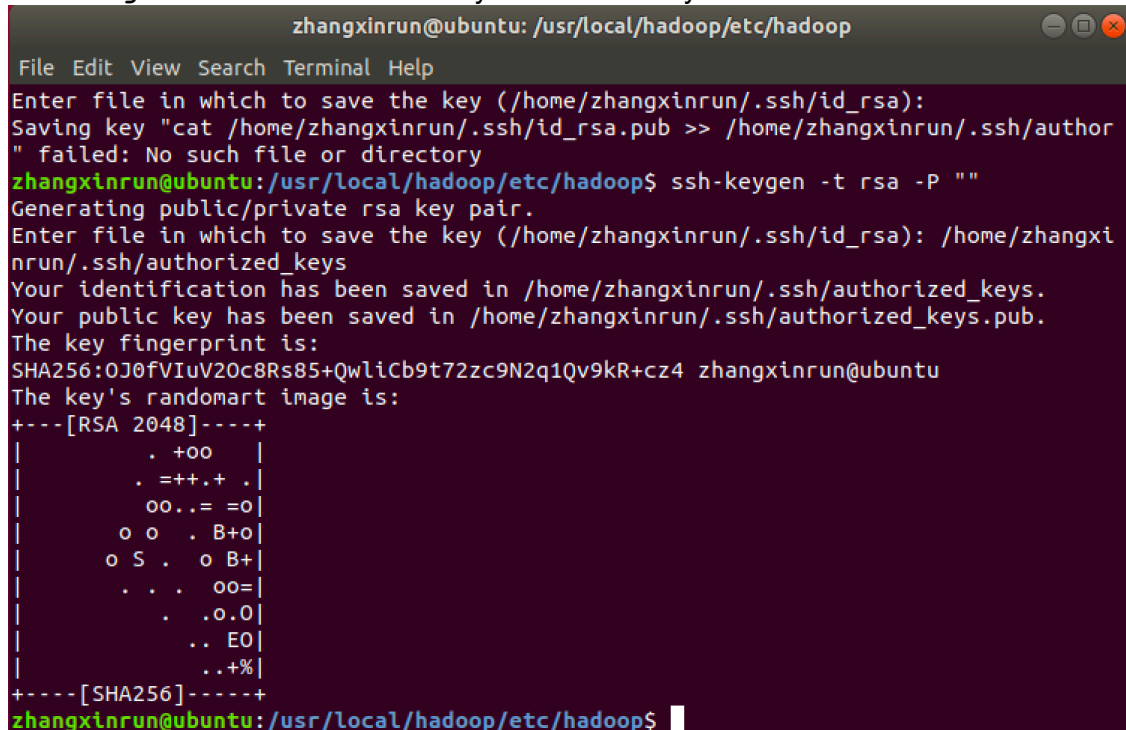
16. Perform following commands to set-up SSH public key authentication

    ```
    ssh-keygen -t rsa -P ""
    cat /home/zhangxinrun/.ssh/id_rsa.pub >>
    /home/zhangxinrun/.ssh/authorized_keys
    ```

    Note: **bigdata** is a user here. It may be different in your case.

17. Format HDFS: **hadoop namenode –format**

```
zhangxinrun@ubuntu:~$ hadoop namenode -format
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.

18/09/13 14:54:54 INFO namenode.NameNode: STARTUP_MSG:
/************************************************************
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = ubuntu/127.0.1.1
STARTUP_MSG:   args = [-format]
STARTUP_MSG:   version = 2.9.1
STARTUP_MSG:   classpath = /usr/local/hadoop/etc/hadoop:/usr/local/hadoop/share/ha
doop/common/lib/commons-codec-1.4.jar:/usr/local/hadoop/share/hadoop/common/lib/pr
otobuf-java-2.5.0.jar:/usr/local/hadoop/share/hadoop/common/lib/jetty-util-6.1.26.
jar:/usr/local/hadoop/share/hadoop/common/lib/api-util-1.0.0-M20.jar:/usr/local/ha
doop/share/hadoop/common/lib/commons-beanutils-1.7.0.jar:/usr/local/hadoop/share/h
adoop/common/lib/jackson-xc-1.9.13.jar:/usr/local/hadoop/share/hadoop/common/lib/c
ommons-beanutils-core-1.8.0.jar:/usr/local/hadoop/share/hadoop/common/lib/junit-4.
11.jar:/usr/local/hadoop/share/hadoop/common/lib/jaxb-api-2.2.2.jar:/usr/local/had
```

```
18/09/13 14:54:55 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
18/09/13 14:54:55 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total h
eap and retry cache entry expiry time is 600000 millis
18/09/13 14:54:55 INFO util.GSet: Computing capacity for map NameNodeRetryCache
18/09/13 14:54:55 INFO util.GSet: VM type       = 64-bit
18/09/13 14:54:55 INFO util.GSet: 0.029999999329447746% max memory 966.7 MB = 297.
0 KB
18/09/13 14:54:55 INFO util.GSet: capacity      = 2^15 = 32768 entries
18/09/13 14:54:55 INFO namenode.FSImage: Allocated new BlockPoolId: BP-280999354-1
27.0.1.1-1536875695729
18/09/13 14:54:55 INFO common.Storage: Storage directory /app/hadoop/tmp/dfs/name
has been successfully formatted.
18/09/13 14:54:55 INFO namenode.FSImageFormatProtobuf: Saving image file /app/hado
op/tmp/dfs/name/current/fsimage.ckpt_0000000000000000000 using no compression
18/09/13 14:54:55 INFO namenode.FSImageFormatProtobuf: Image file /app/hadoop/tmp/
dfs/name/current/fsimage.ckpt_0000000000000000000 of size 328 bytes saved in 0 sec
onds .
18/09/13 14:54:56 INFO namenode.NNStorageRetentionManager: Going to retain 1 image
s with txid >= 0
18/09/13 14:54:56 INFO namenode.NameNode: SHUTDOWN_MSG:
/************************************************************
SHUTDOWN_MSG: Shutting down NameNode at ubuntu/127.0.1.1
************************************************************/
zhangxinrun@ubuntu:~$
```

18. Start service: **start-all.sh**

```
zhangxinrun@ubuntu:/usr/local/hadoop/etc/hadoop$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.ut
il.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.1.
jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.se
curity.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflectiv
e access operations
WARNING: All illegal access operations will be denied in a future release
Starting namenodes on [localhost]
zhangxinrun@localhost's password:
localhost: starting namenode, logging to /usr/local/hadoop/logs/hadoop-zhangxinrun
-namenode-ubuntu.out
zhangxinrun@localhost's password:
localhost: starting datanode, logging to /usr/local/hadoop/logs/hadoop-zhangxinrun
-datanode-ubuntu.out
Starting secondary namenodes [0.0.0.0]
zhangxinrun@0.0.0.0's password:
0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop/logs/hadoop-zhan
gxinrun-secondarynamenode-ubuntu.out
```

19. Execute command to check services are running: **jps**

```
zhangxinrun@ubuntu:/usr/local/hadoop/etc/hadoop$ jps
22562 NodeManager
22677 Jps
21321 ResourceManager
```

20. Once your job is done, you can stop services: `stop-all.sh`
    Do not execute this command now, first do something in HDFS cluster as described below.

HDFS command

1. Assume you are at your home directory or execute: `cd`
2. Examine files in HDFS cluster: `hdfs dfs -ls`
   You will see nothing in the output as you have not created anything yet.

```
zhangxinrun@ubuntu:~$ hdfs dfs -ls
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (fil
e:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.1.jar) to method sun.security.krb5.Config.
getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authenticatio
n.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
ls: `.': No such file or directory
```

3. Create a directory in HDFS cluster: `hdfs dfs -mkdir /user`

```
zhangxinrun@ubuntu:~$ hdfs dfs -mkdir /user
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (fil
e:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.1.jar) to method sun.security.krb5.Config.
getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authenticatio
n.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
```

4. Examine files again in HDFS cluster: `hdfs dfs -ls`
   You will see the listing of /user directory in the output
5. Create a directory inside /user directory using your user name:
   `hdfs dfs -mkdir /user/bigdata`
6. Examine that directory in HDFS cluster: `hdfs dfs -ls /user/`

```
zhangxinrun@ubuntu:~$ hdfs dfs -ls /user/
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.Ke
rberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.1.jar) to me
thod sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.securit
y.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective acc
ess operations
WARNING: All illegal access operations will be denied in a future release
Found 1 items
drwxr-xr-x   - zhangxinrun supergroup          0 2018-09-18 12:02 /user/bigdata
```

7. Create a text file, say `demo.txt`, using gedit and write something in it and save it.
8. Upload the file in the cluster：
   `hdfs dfs -copyFromLocal demo.txt /user/bigdata/`

```
zhangxinrun@ubuntu:~$ vi demo.txt
zhangxinrun@ubuntu:~$ hdfs dfs -copyFromLocal demo.txt /user/bigdata/
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.Ke
rberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.1.jar) to me
thod sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.securit
y.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective acc
ess operations
WARNING: All illegal access operations will be denied in a future release
zhangxinrun@ubuntu:~$
```

9. List the file in the cluster: `hdfs dfs -ls /user/bigdata/`

```
zhangxinrun@ubuntu:~$ hdfs dfs -ls /user/bigdata/
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.Ke
rberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.1.jar) to me
thod sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.securit
y.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective acc
ess operations
WARNING: All illegal access operations will be denied in a future release
Found 1 items
-rw-r--r--   1 zhangxinrun supergroup         19 2018-09-18 12:06 /user/bigdata/demo.tx
t
zhangxinrun@ubuntu:~$
```

10. See the content of file: **hdfs dfs -cat /user/bigdata/demo.txt**

```
zhangxinrun@ubuntu:~$ hdfs dfs -cat /user/bigdata/demo.txt
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.Ke
rberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.9.1.jar) to me
thod sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.securit
y.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective acc
ess operations
WARNING: All illegal access operations will be denied in a future release
llalalallala
sadwa
zhangxinrun@ubuntu:~$
```

Every time I tried to execute hdfs commands, some warnings showed up. I asked my professor for help and according to this link: https://stackoverflow.com/questions/52155078/how-to-fix-hadoop-warning-an-illegal-reflective-access-operation-has-occurred-e/52155228.

There is nothing you can do about these warnings, this is related to jigsaw project and strong(er) encapsulation.

Basically there is some class called `sun.security.krb5.Config` that is part of some "module" called `java.security.jgss`. This module "defines" what it exports (what others can use out of it) and to whom. This also means in plain english that this is not for the public usage - don't touch it; well `hadoop` did, it's part of their effort to fix this. You can report this or try to upgrade hadoop, may be this is already fixed.

Maybe a higher version of Hadoop can fix this problem?