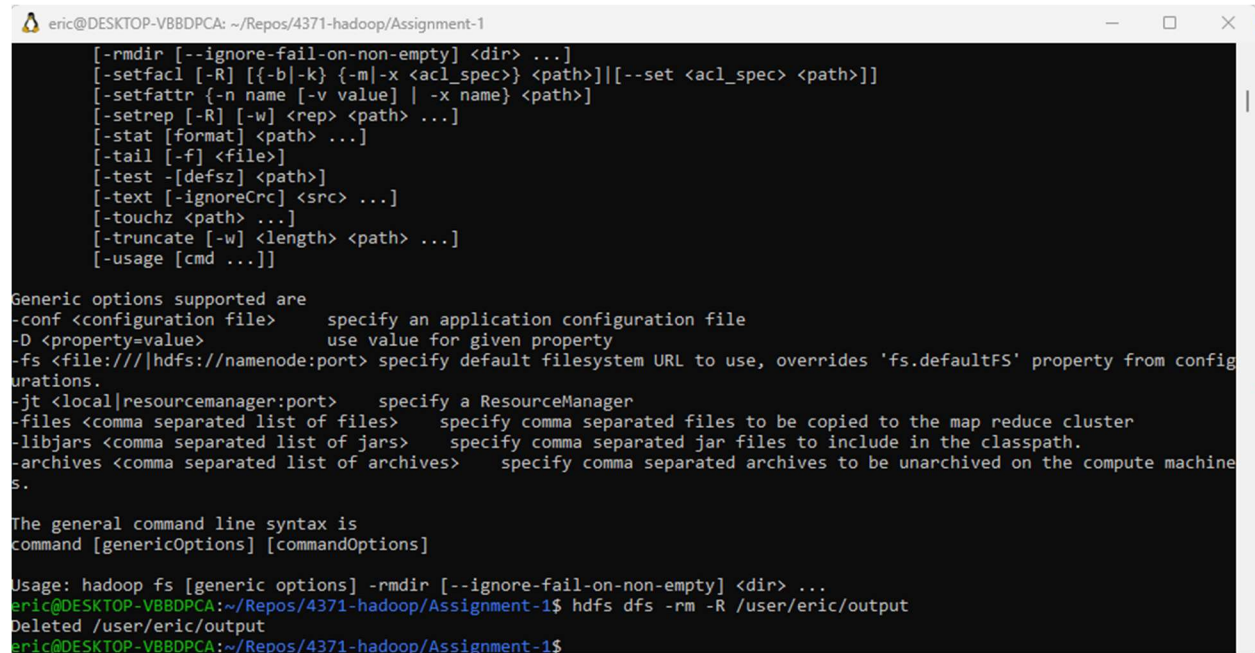


Homework 1:

Question 1: Screenshots

Part A – My OS environment is an Ubuntu distribution in wsl.



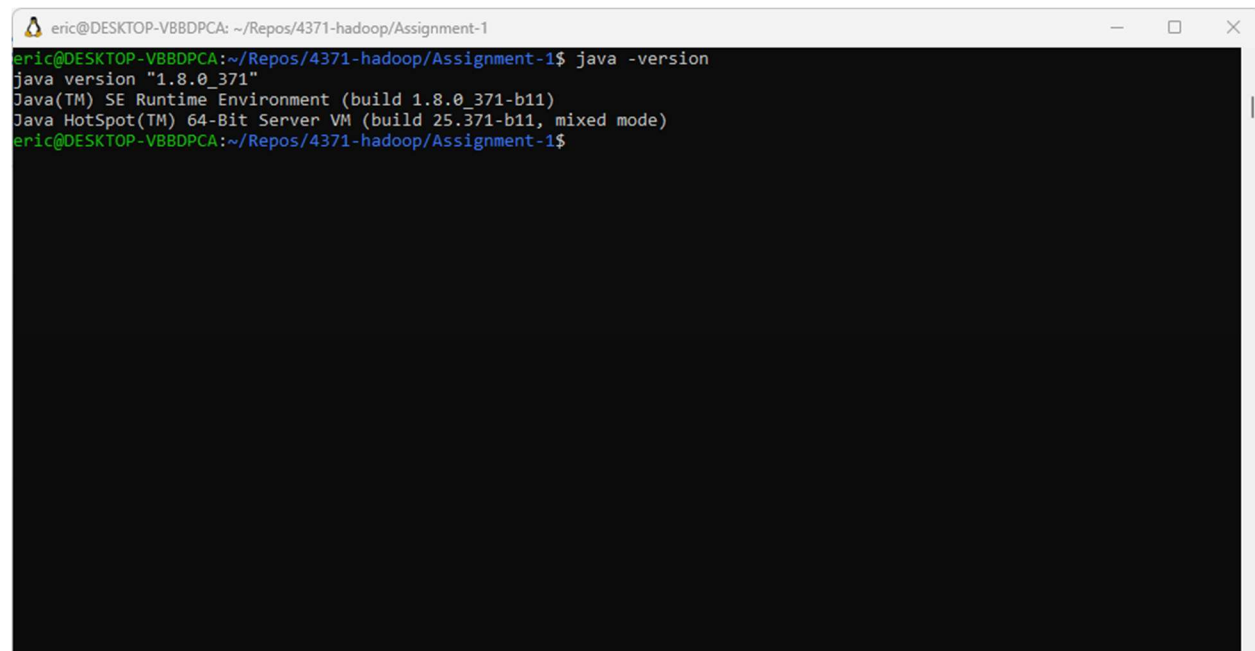
```
eric@DESKTOP-VBBDPCA: ~/Repos/4371-hadoop/Assignment-1
[-rmdir [--ignore-fail-on-non-empty] <dir> ...]
[-setfacl [-R] [{-b|-k} {-m|-x <acl_spec>} <path>][[--set <acl_spec> <path>]]
[-setfattr {-n name [-v value] | -x name} <path>]
[-setrep [-R] [-w] <rep> <path> ...]
[-stat [format] <path> ...]
[-tail [-f] <file>]
[-test [-defsz] <path>]
[-text [-ignoreCrc] <src> ...]
[-touchz <path> ...]
[-truncate [-w] <length> <path> ...]
[-usage [cmd ...]]

Generic options supported are
-conf <configuration file>      specify an application configuration file
-D <property=value>             use value for given property
-fs <file:///|hdfs://namenode:port> specify default filesystem URL to use, overrides 'fs.defaultFS' property from configurations.
-jt <local|resourcemanager:port> specify a ResourceManager
-files <comma separated list of files> specify comma separated files to be copied to the map reduce cluster
-libjars <comma separated list of jars> specify comma separated jar files to include in the classpath.
-archives <comma separated list of archives> specify comma separated archives to be unarchived on the compute machines.

The general command line syntax is
command [genericOptions] [commandOptions]

Usage: hadoop fs [generic options] -rmdir [--ignore-fail-on-non-empty] <dir> ...
eric@DESKTOP-VBBDPCA:~/Repos/4371-hadoop/Assignment-1$ hdfs dfs -rm -R /user/eric/output
Deleted /user/eric/output
eric@DESKTOP-VBBDPCA:~/Repos/4371-hadoop/Assignment-1$
```

Part B – My output of the java -version



```
eric@DESKTOP-VBBDPCA: ~/Repos/4371-hadoop/Assignment-1
eric@DESKTOP-VBBDPCA:~/Repos/4371-hadoop/Assignment-1$ java -version
java version "1.8.0_371"
Java(TM) SE Runtime Environment (build 1.8.0_371-b11)
Java HotSpot(TM) 64-Bit Server VM (build 25.371-b11, mixed mode)
eric@DESKTOP-VBBDPCA:~/Repos/4371-hadoop/Assignment-1$
```

Part C – Output of the Hadoop version command

```
eric@DESKTOP-VBBDPCA: ~/Repos/4371-hadoop/Assignment-1
eric@DESKTOP-VBBDPCA:~/Repos/4371-hadoop/Assignment-1$ hadoop version
Hadoop 2.8.1
Subversion https://git-wip-us.apache.org/repos/asf/hadoop.git -r 20fe5304904fc2f5a18053c389e43cd26f7a70fe
Compiled by vinodkv on 2017-06-02T06:14Z
Compiled with protoc 2.5.0
From source with checksum 60125541c2b3e266cbf3becc5bda666
This command was run using /home/eric/Documents/hadoop/share/hadoop/common/hadoop-common-2.8.1.jar
eric@DESKTOP-VBBDPCA:~/Repos/4371-hadoop/Assignment-1$
```

Part D – Various snips of my files

File contents of my Path Variables in .profile

```
eric@DESKTOP-VBBDPCA: ~/Repos/4371-hadoop/Assignment-1
GNU nano 6.2 /home/eric/.profile
# ~/.profile: executed by the command interpreter for login shells.
# This file is not read by bash(1), if ~/.bash_profile or ~/.bash_login
# exists.
# see /usr/share/doc/bash/examples/startup-files for examples.
# the files are located in the bash-doc package.

# the default umask is set in /etc/profile; for setting the umask
# for ssh logins, install and configure the libpam-umask package.
#umask 022

# if running bash
if [ -n "$BASH_VERSION" ]; then
    # include .bashrc if it exists
    if [ -f "$HOME/.bashrc" ]; then
        . "$HOME/.bashrc"
    fi
fi

# set PATH so it includes user's private bin if it exists
if [ -d "$HOME/bin" ] ; then
    PATH="$HOME/bin:$PATH"
fi

# set PATH so it includes user's private bin if it exists
if [ -d "$HOME/.local/bin" ] ; then
    PATH="$HOME/.local/bin:$PATH"
fi

export JAVA_HOME=/usr/lib/java
export JRE_HOME=$JAVA_HOME/jre
export CLASSPATH=$CLASSPATH:.$JAVA_HOME/lib:$JRE_HOME/lib
export PATH=$JAVA_HOME/bin:$JRE_HOME/bin:$PATH
export PATH=$PATH:~/Documents/hadoop/bin:~/Documents/hadoop/sbin

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

Snip of Core-site.xml

```

eric@DESKTOP-VBBDPCA: ~/Documents/hadoop/etc/hadoop
GNU nano 6.2 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>
</configuration>

```

Read 24 lines

Help Write Out Where Is Cut Execute Location M-U Undo M-A Set Mark
Exit Read File Replace Paste Justify Go To Line M-E Redo M-G Copy

Snip of hdfs-site.xml

```

eric@DESKTOP-VBBDPCA: ~/Documents/hadoop/etc/hadoop
GNU nano 6.2 hdfs-site.xml
<!--
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>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>/home/eric/Documents/hadoop_hdfs_data/name</value>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>/home/eric/Documents/hadoop_hdfs_data/data</value>
  </property>
  <property>
    <name>dfs.namenode.checkpoint.dir</name>
    <value>/home/eric/Documents/hadoop_hdfs_data/secondary</value>
  </property>
</configuration>

```

Help Write Out Where Is Cut Execute Location M-U Undo M-A Set Mark
Exit Read File Replace Paste Justify Go To Line M-E Redo M-G Copy

Snip of Hadoop-env

```

eric@DESKTOP-VBBDPCA: ~/Documents/hadoop/etc/hadoop
GNU nano 6.2 hadoop-env.sh
# Licensed to the Apache Software Foundation (ASF) under one
# or more contributor license agreements. See the NOTICE file
# distributed with this work for additional information
# regarding copyright ownership. The ASF licenses this file
# to you 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.
#
# 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/java
#
# The jsvc implementation to use. Jsvc is required to run secure datanodes
# that bind to privileged ports to provide authentication of data transfer
# protocol. Jsvc is not required if SASL is configured for authentication of
# data transfer protocol using non-privileged ports.
#export JSVC_HOME=${JSVC_HOME}
#
export HADOOP_CONF_DIR=${HADOOP_CONF_DIR:-"/etc/hadoop"}
#
# Extra Java CLASSPATH elements. Automatically insert capacity-scheduler.
Read 108 lines
^G Help      ^O Write Out ^W Where Is  ^K Cut       ^T Execute   ^C Location  ^U Undo      ^A Set Mark
^X Exit      ^R Read File ^N Replace   ^U Paste     ^J Justify   ^_/ Go To Line ^-E Redo     ^-G Copy

```

Part E – The output of jps command

```

eric@DESKTOP-VBBDPCA: ~/Documents/hadoop/etc/hadoop$ jps
11602 NodeManager
60642 Jps
11156 SecondaryNameNode
10982 DataNode
11308 ResourceManager
10845 NameNode
36319 org.eclipse.equinox.launcher_1.6.500.v20230717-2134.jar
eric@DESKTOP-VBBDPCA: ~/Documents/hadoop/etc/hadoop$

```

Part F – Snip of the Web Interface

Overview 'localhost:9000' (active)

Started:	Wed Sep 13 13:25:38 -0500 2023
Version:	2.8.1, r20fe5304904fc2f5a18053c389e43cd26f7a70fe
Compiled:	Fri Jun 02 01:14:00 -0500 2017 by vinodkv from branch-2.8.1-private
Cluster ID:	CID-e9955f1d-e62e-4b4a-9f26-b88cc8a9de89
Block Pool ID:	BP-717907232-127.0.1.1-1694629513343

Summary

Security is off.
Safemode is off.
6 files and directories, 2 blocks = 8 total filesystem object(s).
Heap Memory used 72.31 MB of 307 MB Heap Memory. Max Heap Memory is 889 MB.
Non Heap Memory used 50.83 MB of 51.97 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

Configured Capacity:	1006.85 GB
DFS Used:	8.48 MB (0%)
Non DFS Used:	5.38 GB
DFS Remaining:	950.25 GB (94.38%)
Block Pool Used:	8.48 MB (0%)
DataNodes usages% (Min/Median/Max/stdDev):	0.00% / 0.00% / 0.00% / 0.00%
Live Nodes	1 (Decommissioned: 0)

Part G – Screenshot of “hdfs dfs -ls -R /”

Eric Shields

```
eric@DESKTOP-V8BDPCA: ~/Documents/hadoop/etc/hadoop
eric@DESKTOP-V8BDPCA:~/Documents/hadoop/etc/hadoop$ hdfs dfs -ls -R /
drwxr-xr-x  - eric supergroup      0 2023-09-13 21:16 /user
drwxr-xr-x  - eric supergroup      0 2023-09-13 21:29 /user/eric
drwxr-xr-x  - eric supergroup      0 2023-09-13 21:28 /user/eric/input
-rw-r--r--   1 eric supergroup    6617121 2023-09-13 21:16 /user/eric/input/big.txt
-rw-r--r--   1 eric supergroup   2154522 2023-09-13 21:28 /user/eric/input/input_hw1.txt
eric@DESKTOP-V8BDPCA:~/Documents/hadoop/etc/hadoop$
```