**Name – Akanksha Rani Big Data Analytics**

**Reg. No. – 21MCA0220 ITA6008**

**Digital Assignment – 1**

Q. Consider the below two files and implement the word count Map Reduce program.

File 1:

Big data is a combination of structured, semistructured and unstructured data collected by organizations.

File 2:

Big data also encompasses a wide variety of data types, including the following: Structured data, such as transactions and financial records. Unstructured data, such as text, documents and multimedia files. Semistructured data, such as web server logs and streaming data from sensors.

**Solution:**

**After installing Hadoop following commands to be followed:**

hadoop@akanksharani-VirtualBox:~/Desktop$ **su - hadoop**

Password:

hadoop@akanksharani-VirtualBox:~$ **sudo apt update**

[sudo] password for hadoop:

Get:1 http://security.ubuntu.com/ubuntu jammy-security InRelease [110 kB]

Hit:2 http://in.archive.ubuntu.com/ubuntu jammy InRelease

Get:3 http://security.ubuntu.com/ubuntu jammy-security/main amd64 DEP-11 Metadata [11.4 kB]

Hit:4 http://in.archive.ubuntu.com/ubuntu jammy-updates InRelease

Get:5 http://security.ubuntu.com/ubuntu jammy-security/universe amd64 DEP-11 Metadata [608 B]

Hit:6 http://in.archive.ubuntu.com/ubuntu jammy-backports InRelease

Fetched 122 kB in 36s (3,387 B/s)

Reading package lists... Done

Building dependency tree... Done

Reading state information... Done

168 packages can be upgraded. Run 'apt list --upgradable' to see them.

hadoop@akanksharani-VirtualBox:~$ **ssh-keygen -t rsa -P '' -f ~/.ssh/id\_rsa**

Generating public/private rsa key pair.

/home/hadoop/.ssh/id\_rsa already exists.

Overwrite (y/n)? y

Your identification has been saved in /home/hadoop/.ssh/id\_rsa

Your public key has been saved in /home/hadoop/.ssh/id\_rsa.pub

The key fingerprint is:

SHA256:+6BEqeaDlnDUAP+wp5L7xRbHYX5M8rZiBCYaLI2VG3U hadoop@akanksharani-VirtualBox

The key's randomart image is:

+---[RSA 3072]----+

|...o. E |

|.=+ . |

|+.==o + . |

|.oo\*.= B |

|... + B S |

|...+ \* o o |

|oo.o\* + + |

| o+=.o o o |

|.o. ... . |

+----[SHA256]-----+

hadoop@akanksharani-VirtualBox:~$ **cat ~/.ssh/id\_rsa.pub >> ~/.ssh/authorized\_keys**

hadoop@akanksharani-VirtualBox:~$ **chmod 0600 ~/.ssh/authorized\_keys**

hadoop@akanksharani-VirtualBox:~$ **ssh localhost**

Welcome to Ubuntu 22.04 LTS (GNU/Linux 5.15.0-41-generic x86\_64)

\* Documentation: https://help.ubuntu.com

\* Management: https://landscape.canonical.com

\* Support: https://ubuntu.com/advantage

168 updates can be applied immediately.

87 of these updates are standard security updates.

To see these additional updates run: apt list --upgradable

Last login: Sat Jul 16 00:07:23 2022 from 127.0.0.1

hadoop@akanksharani-VirtualBox:~$ **hdfs namenode -format**

2022-07-16 10:01,949 INFO namenode.NameNode: STARTUP\_MSG:

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

STARTUP\_MSG: Starting NameNode

STARTUP\_MSG: host = akanksharani-VirtualBox/127.0.1.1

STARTUP\_MSG: args = [-format]

STARTUP\_MSG: version = 3.2.3

STARTUP\_MSG: classpath = /home/hadoop/hadoop-3.2.3/etc/hadoop:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/re2j-1.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/commons-compress-1.21.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/animal-sniffer-annotations-1.17.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/hadoop-annotations-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/curator-recipes-2.13.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/commons-math3-3.1.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jettison-1.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/json-smart-2.4.7.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/j2objc-annotations-1.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/commons-collections-3.2.2.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jackson-jaxrs-1.9.13.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/spotbugs-annotations-3.1.9.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/commons-cli-1.2.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/dnsjava-2.1.7.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jersey-json-1.19.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/failureaccess-1.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/javax.servlet-api-3.1.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerb-server-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/slf4j-api-1.7.25.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/commons-codec-1.11.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jetty-util-ajax-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/zookeeper-3.4.14.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jackson-annotations-2.10.5.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jetty-util-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/paranamer-2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jersey-server-1.19.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/commons-net-3.6.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerb-core-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jackson-core-asl-1.9.13.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/curator-client-2.13.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerb-simplekdc-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jsp-api-2.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jersey-servlet-1.19.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/htrace-core4-4.1.0-incubating.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/snappy-java-1.0.5.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/httpclient-4.5.13.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jackson-xc-1.9.13.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jsr311-api-1.1.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jackson-mapper-asl-1.9.13.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/avro-1.7.7.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jetty-security-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/commons-text-1.4.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/commons-logging-1.1.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerby-util-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/checker-qual-2.5.2.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jetty-servlet-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/asm-5.0.4.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/stax2-api-4.2.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/httpcore-4.4.13.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/commons-beanutils-1.9.4.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/woodstox-core-5.3.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jsr305-3.0.2.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/audience-annotations-0.5.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/token-provider-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/metrics-core-3.2.4.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/curator-framework-2.13.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerb-client-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jsch-0.1.55.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerb-common-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerb-util-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jetty-server-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/commons-io-2.8.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/error\_prone\_annotations-2.2.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/javax.activation-api-1.2.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jcip-annotations-1.0-1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jackson-core-2.10.5.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerby-asn1-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerb-identity-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jul-to-slf4j-1.7.25.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jetty-io-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/accessors-smart-2.4.7.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerb-admin-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/log4j-1.2.17.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jetty-webapp-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jersey-core-1.19.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/nimbus-jose-jwt-9.8.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jetty-xml-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jaxb-api-2.2.11.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerby-xdr-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/hadoop-auth-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerby-pkix-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jetty-http-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jaxb-impl-2.2.3-1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/netty-3.10.6.Final.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/listenablefuture-9999.0-empty-to-avoid-conflict-with-guava.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/gson-2.2.4.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/jackson-databind-2.10.5.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerb-crypto-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/commons-lang3-3.7.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/guava-27.0-jre.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/kerby-config-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/protobuf-java-2.5.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/lib/commons-configuration2-2.1.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/hadoop-kms-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/hadoop-nfs-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/hadoop-common-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/common/hadoop-common-3.2.3-tests.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/okio-1.6.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/re2j-1.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/commons-compress-1.21.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/animal-sniffer-annotations-1.17.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/hadoop-annotations-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/curator-recipes-2.13.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/commons-math3-3.1.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jettison-1.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/json-smart-2.4.7.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/j2objc-annotations-1.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/commons-collections-3.2.2.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jackson-jaxrs-1.9.13.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/spotbugs-annotations-3.1.9.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/commons-cli-1.2.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/dnsjava-2.1.7.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jersey-json-1.19.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/failureaccess-1.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/javax.servlet-api-3.1.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerb-server-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/commons-codec-1.11.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jetty-util-ajax-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/zookeeper-3.4.14.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jackson-annotations-2.10.5.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jetty-util-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/paranamer-2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jersey-server-1.19.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/commons-net-3.6.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerb-core-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jackson-core-asl-1.9.13.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/curator-client-2.13.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerb-simplekdc-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jersey-servlet-1.19.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/htrace-core4-4.1.0-incubating.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/snappy-java-1.0.5.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/httpclient-4.5.13.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jackson-xc-1.9.13.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jsr311-api-1.1.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jackson-mapper-asl-1.9.13.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/avro-1.7.7.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jetty-security-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/commons-text-1.4.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/commons-logging-1.1.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerby-util-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/checker-qual-2.5.2.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jetty-servlet-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/netty-all-4.1.68.Final.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/asm-5.0.4.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/stax2-api-4.2.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/okhttp-2.7.5.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/httpcore-4.4.13.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/json-simple-1.1.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/commons-beanutils-1.9.4.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/woodstox-core-5.3.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jsr305-3.0.2.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/audience-annotations-0.5.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/token-provider-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/commons-daemon-1.0.13.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/curator-framework-2.13.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerb-client-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jsch-0.1.55.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerb-common-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerb-util-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jetty-server-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/commons-io-2.8.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/error\_prone\_annotations-2.2.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/javax.activation-api-1.2.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jcip-annotations-1.0-1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jackson-core-2.10.5.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/leveldbjni-all-1.8.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerby-asn1-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerb-identity-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jetty-io-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/accessors-smart-2.4.7.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerb-admin-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/log4j-1.2.17.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jetty-webapp-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jersey-core-1.19.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/nimbus-jose-jwt-9.8.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jetty-xml-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jaxb-api-2.2.11.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerby-xdr-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/hadoop-auth-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerby-pkix-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jetty-http-9.4.40.v20210413.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jaxb-impl-2.2.3-1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/netty-3.10.6.Final.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/listenablefuture-9999.0-empty-to-avoid-conflict-with-guava.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/gson-2.2.4.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/jackson-databind-2.10.5.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerb-crypto-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/commons-lang3-3.7.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/guava-27.0-jre.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/kerby-config-1.0.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/protobuf-java-2.5.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/lib/commons-configuration2-2.1.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/hadoop-hdfs-client-3.2.3-tests.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/hadoop-hdfs-client-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/hadoop-hdfs-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/hadoop-hdfs-3.2.3-tests.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/hadoop-hdfs-httpfs-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/hadoop-hdfs-native-client-3.2.3-tests.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/hadoop-hdfs-native-client-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/hadoop-hdfs-rbf-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/hadoop-hdfs-nfs-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/hdfs/hadoop-hdfs-rbf-3.2.3-tests.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/mapreduce/lib/hamcrest-core-1.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/mapreduce/lib/junit-4.13.2.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/mapreduce/hadoop-mapreduce-client-common-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/mapreduce/hadoop-mapreduce-client-core-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/mapreduce/hadoop-mapreduce-client-uploader-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/mapreduce/hadoop-mapreduce-client-app-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/mapreduce/hadoop-mapreduce-client-jobclient-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/mapreduce/hadoop-mapreduce-client-jobclient-3.2.3-tests.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/mapreduce/hadoop-mapreduce-client-hs-plugins-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/mapreduce/hadoop-mapreduce-client-hs-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/mapreduce/hadoop-mapreduce-client-shuffle-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/mapreduce/hadoop-mapreduce-client-nativetask-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/objenesis-1.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/aopalliance-1.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/jackson-module-jaxb-annotations-2.10.5.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/jackson-jaxrs-base-2.10.5.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/guice-servlet-4.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/HikariCP-java7-2.4.12.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/jackson-jaxrs-json-provider-2.10.5.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/guice-4.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/swagger-annotations-1.5.4.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/jakarta.xml.bind-api-2.3.2.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/javax.inject-1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/java-util-1.9.0.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/metrics-core-3.2.4.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/geronimo-jcache\_1.0\_spec-1.0-alpha-1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/fst-2.50.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/jersey-client-1.19.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/bcprov-jdk15on-1.60.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/jakarta.activation-api-1.2.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/bcpkix-jdk15on-1.60.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/ehcache-3.3.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/snakeyaml-1.26.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/mssql-jdbc-6.2.1.jre7.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/json-io-2.5.1.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/lib/jersey-guice-1.19.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-server-sharedcachemanager-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-applications-unmanaged-am-launcher-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-registry-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-client-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-server-resourcemanager-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-server-timeline-pluginstorage-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-services-core-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-submarine-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-common-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-server-router-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-server-common-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-applications-distributedshell-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-server-web-proxy-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-api-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-server-applicationhistoryservice-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-server-nodemanager-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-server-tests-3.2.3.jar:/home/hadoop/hadoop-3.2.3/share/hadoop/yarn/hadoop-yarn-services-api-3.2.3.jar

STARTUP\_MSG: build = https://github.com/apache/hadoop -r abe5358143720085498613d399be3bbf01e0f131; compiled by 'ubuntu' on 2022-03-20T01:18Z

STARTUP\_MSG: java = 1.8.0\_312

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

2022-07-16 10:01,968 INFO namenode.NameNode: registered UNIX signal handlers for [TERM, HUP, INT]

2022-07-16 10:02,089 INFO namenode.NameNode: createNameNode [-format]

Formatting using clusterid: CID-c3a53bf9-b22d-40c1-a113-2e238dcedf8b

2022-07-16 10:02,640 INFO namenode.FSEditLog: Edit logging is async:true

2022-07-16 10:02,686 INFO namenode.FSNamesystem: KeyProvider: null

2022-07-16 10:02,687 INFO namenode.FSNamesystem: fsLock is fair: true

2022-07-16 10:02,689 INFO namenode.FSNamesystem: Detailed lock hold time metrics enabled: false

2022-07-16 10:02,700 INFO namenode.FSNamesystem: fsOwner = hadoop (auth:SIMPLE)

2022-07-16 10:02,701 INFO namenode.FSNamesystem: supergroup = supergroup

2022-07-16 10:02,701 INFO namenode.FSNamesystem: isPermissionEnabled = true

2022-07-16 10:02,701 INFO namenode.FSNamesystem: HA Enabled: false

2022-07-16 10:02,761 INFO common.Util: dfs.datanode.fileio.profiling.sampling.percentage set to 0. Disabling file IO profiling

2022-07-16 10:02,779 INFO blockmanagement.DatanodeManager: dfs.block.invalidate.limit: configured=1000, counted=60, effected=1000

2022-07-16 10:02,779 INFO blockmanagement.DatanodeManager: dfs.namenode.datanode.registration.ip-hostname-check=true

2022-07-16 10:02,783 INFO blockmanagement.BlockManager: dfs.namenode.startup.delay.block.deletion.sec is set to 000:00:00:00.000

2022-07-16 10:02,783 INFO blockmanagement.BlockManager: The block deletion will start around 2022 Jul 16 10:02

2022-07-16 10:02,785 INFO util.GSet: Computing capacity for map BlocksMap

2022-07-16 10:02,785 INFO util.GSet: VM type = 64-bit

2022-07-16 10:02,787 INFO util.GSet: 2.0% max memory 908.5 MB = 18.2 MB

2022-07-16 10:02,787 INFO util.GSet: capacity = 2^21 = 2097152 entries

2022-07-16 10:02,792 INFO blockmanagement.BlockManager: Storage policy satisfier is disabled

2022-07-16 10:02,793 INFO blockmanagement.BlockManager: dfs.block.access.token.enable = false

2022-07-16 10:02,799 INFO blockmanagement.BlockManagerSafeMode: dfs.namenode.safemode.threshold-pct = 0.9990000128746033

2022-07-16 10:02,799 INFO blockmanagement.BlockManagerSafeMode: dfs.namenode.safemode.min.datanodes = 0

2022-07-16 10:02,799 INFO blockmanagement.BlockManagerSafeMode: dfs.namenode.safemode.extension = 30000

2022-07-16 10:02,800 INFO blockmanagement.BlockManager: defaultReplication = 1

2022-07-16 10:02,800 INFO blockmanagement.BlockManager: maxReplication = 512

2022-07-16 10:02,800 INFO blockmanagement.BlockManager: minReplication = 1

2022-07-16 10:02,800 INFO blockmanagement.BlockManager: maxReplicationStreams = 2

2022-07-16 10:02,800 INFO blockmanagement.BlockManager: redundancyRecheckInterval = 3000ms

2022-07-16 10:02,801 INFO blockmanagement.BlockManager: encryptDataTransfer = false

2022-07-16 10:02,801 INFO blockmanagement.BlockManager: maxNumBlocksToLog = 1000

2022-07-16 10:02,830 INFO namenode.FSDirectory: GLOBAL serial map: bits=29 maxEntries=536870911

2022-07-16 10:02,830 INFO namenode.FSDirectory: USER serial map: bits=24 maxEntries=16777215

2022-07-16 10:02,830 INFO namenode.FSDirectory: GROUP serial map: bits=24 maxEntries=16777215

2022-07-16 10:02,831 INFO namenode.FSDirectory: XATTR serial map: bits=24 maxEntries=16777215

2022-07-16 10:02,850 INFO util.GSet: Computing capacity for map INodeMap

2022-07-16 10:02,850 INFO util.GSet: VM type = 64-bit

2022-07-16 10:02,851 INFO util.GSet: 1.0% max memory 908.5 MB = 9.1 MB

2022-07-16 10:02,851 INFO util.GSet: capacity = 2^20 = 1048576 entries

2022-07-16 10:02,852 INFO namenode.FSDirectory: ACLs enabled? false

2022-07-16 10:02,852 INFO namenode.FSDirectory: POSIX ACL inheritance enabled? true

2022-07-16 10:02,852 INFO namenode.FSDirectory: XAttrs enabled? true

2022-07-16 10:02,853 INFO namenode.NameNode: Caching file names occurring more than 10 times

2022-07-16 10:02,885 INFO snapshot.SnapshotManager: Loaded config captureOpenFiles: false, skipCaptureAccessTimeOnlyChange: false, snapshotDiffAllowSnapRootDescendant: true, maxSnapshotLimit: 65536

2022-07-16 10:02,888 INFO snapshot.SnapshotManager: SkipList is disabled

2022-07-16 10:02,893 INFO util.GSet: Computing capacity for map cachedBlocks

2022-07-16 10:02,893 INFO util.GSet: VM type = 64-bit

2022-07-16 10:02,893 INFO util.GSet: 0.25% max memory 908.5 MB = 2.3 MB

2022-07-16 10:02,893 INFO util.GSet: capacity = 2^18 = 262144 entries

2022-07-16 10:02,899 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.window.num.buckets = 10

2022-07-16 10:02,899 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.num.users = 10

2022-07-16 10:02,899 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.windows.minutes = 1,5,25

2022-07-16 10:02,902 INFO namenode.FSNamesystem: Retry cache on namenode is enabled

2022-07-16 10:02,902 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total heap and retry cache entry expiry time is 600000 millis

2022-07-16 10:02,910 INFO util.GSet: Computing capacity for map NameNodeRetryCache

2022-07-16 10:02,910 INFO util.GSet: VM type = 64-bit

2022-07-16 10:02,911 INFO util.GSet: 0.029999999329447746% max memory 908.5 MB = 279.1 KB

2022-07-16 10:02,911 INFO util.GSet: capacity = 2^15 = 32768 entries

Re-format filesystem in Storage Directory root= /home/hadoop/tmpdata/dfs/name; location= null ? (Y or N) Y

2022-07-16 10:10,743 INFO namenode.FSImage: Allocated new BlockPoolId: BP-7431878-127.0.1.1-1657955650731

2022-07-16 10:10,743 INFO common.Storage: Will remove files: [/home/hadoop/tmpdata/dfs/name/current/edits\_inprogress\_0000000000000000001, /home/hadoop/tmpdata/dfs/name/current/fsimage\_0000000000000000000, /home/hadoop/tmpdata/dfs/name/current/seen\_txid, /home/hadoop/tmpdata/dfs/name/current/VERSION, /home/hadoop/tmpdata/dfs/name/current/fsimage\_0000000000000000000.md5]

2022-07-16 10:10,773 INFO common.Storage: Storage directory /home/hadoop/tmpdata/dfs/name has been successfully formatted.

2022-07-16 10:10,823 INFO namenode.FSImageFormatProtobuf: Saving image file /home/hadoop/tmpdata/dfs/name/current/fsimage.ckpt\_0000000000000000000 using no compression

2022-07-16 10:10,915 INFO namenode.FSImageFormatProtobuf: Image file /home/hadoop/tmpdata/dfs/name/current/fsimage.ckpt\_0000000000000000000 of size 401 bytes saved in 0 seconds .

2022-07-16 10:10,925 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0

2022-07-16 10:10,954 INFO namenode.FSNamesystem: Stopping services started for active state

2022-07-16 10:10,954 INFO namenode.FSNamesystem: Stopping services started for standby state

2022-07-16 10:10,980 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when meet shutdown.

2022-07-16 10:10,980 INFO namenode.NameNode: SHUTDOWN\_MSG:

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

SHUTDOWN\_MSG: Shutting down NameNode at akanksharani-VirtualBox/127.0.1.1

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

**Start all (namenode, datanode, nademanager, yarn)**

hadoop@akanksharani-VirtualBox:~$ **start-all.sh**

WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 10 seconds.

WARNING: This is not a recommended production deployment configuration.

WARNING: Use CTRL-C to abort.

Starting namenodes on [localhost]

Starting datanodes

Starting secondary namenodes [akanksharani-VirtualBox]

Starting resourcemanager

Starting nodemanagers

hadoop@akanksharani-VirtualBox:~$ **jps**

6003 ResourceManager

5491 NameNode

6133 NodeManager

5829 SecondaryNameNode

5624 DataNode

6474 Jps

hadoop@akanksharani-VirtualBox:~$ **cd /home/hadoop/hadoop-3.2.3**

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$ **hdfs dfs -mkdir /input**

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$ **hdfs dfs -ls /**

Found 1 items

drwxr-xr-x - hadoop supergroup 0 2022-07-16 12:45 /input

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$ **hdfs dfs -chmod 777 /input**

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$ **hdfs dfs -ls /**

Found 1 items

drwxrwxrwx - hadoop supergroup 0 2022-07-16 12:45 **/input**

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$ **hdfs dfs -put /home/hadoop/inputfile.txt /input**

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$ **hdfs dfs -ls /input**

Found 1 items

-rw-r--r-- 1 hadoop supergroup 106 2022-07-16 12:46 /input/inputfile.txt

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$ **hdfs dfs -chmod 777 /input/inputfile.txt**

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$ **hdfs dfs -ls /input**

Found 1 items

-rwxrwxrwx 1 hadoop supergroup 106 2022-07-16 12:46 **/input/inputfile.txt**

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$ **hdfs dfs -put /home/hadoop/inputfile2.txt /input**

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$ **hdfs dfs -ls /input**

Found 2 items

-rwxrwxrwx 1 hadoop supergroup 106 2022-07-16 12:46 /**input/inputfile.txt**

-rw-r--r-- 1 hadoop supergroup 285 2022-07-16 12:48 **/input/inputfile2.txt**

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$ **hdfs dfs -chmod 777 /input/inputfile2.txt**

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$ **hdfs dfs -ls /input**

Found 2 items

-rwxrwxrwx 1 hadoop supergroup 106 2022-07-16 12:46 /input/inputfile.txt

-rwxrwxrwx 1 hadoop supergroup 285 2022-07-16 12:48 /input/inputfile2.txt

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3**$ hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.3.jar wordcount /input /output**

2022-07-16 10:49,408 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032

2022-07-16 10:49,784 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/hadoop/.staging/job\_1657955686364\_0001

2022-07-16 10:49,991 INFO input.FileInputFormat: Total input files to process : 2

2022-07-16 10:49,471 INFO mapreduce.JobSubmitter: number of splits:2

2022-07-16 10:49,663 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1657955686364\_0001

2022-07-16 10:49,665 INFO mapreduce.JobSubmitter: Executing with tokens: []

2022-07-16 10:49,834 INFO conf.Configuration: resource-types.xml not found

2022-07-16 10:49,835 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.

2022-07-16 10:49,020 INFO impl.YarnClientImpl: Submitted application application\_1657955686364\_0001

2022-07-16 10:49,081 INFO mapreduce.Job: The url to track the job: http://akanksharani-VirtualBox:8088/proxy/application\_1657955686364\_0001/

2022-07-16 10:49,082 INFO mapreduce.Job: Running job: job\_1657955686364\_0001

2022-07-16 10:49,207 INFO mapreduce.Job: Job job\_1657955686364\_0001 running in uber mode : false

2022-07-16 10:49,209 INFO mapreduce.Job: map 0% reduce 0%

2022-07-16 10:49,295 INFO mapreduce.Job: map 100% reduce 0%

2022-07-16 10:49,358 INFO mapreduce.Job: map 100% reduce 100%

2022-07-16 10:49,366 INFO mapreduce.Job: Job job\_1657955686364\_0001 completed successfully

2022-07-16 10:49,491 INFO mapreduce.Job: Counters: 54

File System Counters

FILE: Number of bytes read=616

FILE: Number of bytes written=710026

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=604

HDFS: Number of bytes written=402

HDFS: Number of read operations=11

HDFS: Number of large read operations=0

HDFS: Number of write operations=2

HDFS: Number of bytes read erasure-coded=0

Job Counters

Launched map tasks=2

Launched reduce tasks=1

Data-local map tasks=2

Total time spent by all maps in occupied slots (ms)=6129

Total time spent by all reduces in occupied slots (ms)=2828

Total time spent by all map tasks (ms)=6129

Total time spent by all reduce tasks (ms)=2828

Total vcore-milliseconds taken by all map tasks=6129

Total vcore-milliseconds taken by all reduce tasks=2828

Total megabyte-milliseconds taken by all map tasks=6276096

Total megabyte-milliseconds taken by all reduce tasks=2895872

Map-Reduce Framework

Map input records=5

Map output records=56

Map output bytes=615

Map output materialized bytes=622

Input split bytes=213

Combine input records=56

Combine output records=45

Reduce input groups=40

Reduce shuffle bytes=622

Reduce input records=45

Reduce output records=40

Spilled Records=90

Shuffled Maps =2

Failed Shuffles=0

Merged Map outputs=2

GC time elapsed (ms)=242

CPU time spent (ms)=1340

Physical memory (bytes) snapshot=745877504

Virtual memory (bytes) snapshot=7591682048

Total committed heap usage (bytes)=656408576

Peak Map Physical memory (bytes)=282697728

Peak Map Virtual memory (bytes)=2527752192

Peak Reduce Physical memory (bytes)=182513664

Peak Reduce Virtual memory (bytes)=2537734144

Shuffle Errors

BAD\_ID=0

CONNECTION=0

IO\_ERROR=0

WRONG\_LENGTH=0

WRONG\_MAP=0

WRONG\_REDUCE=0

File Input Format Counters

Bytes Read=391

File Output Format Counters

Bytes Written=402

**Output:**

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$ **hdfs dfs -ls /output**

Found 2 items

-rw-r--r-- 1 hadoop supergroup 0 2022-07-16 10 49utput/\_SUCCESS

-rw-r--r-- 1 hadoop supergroup 402 2022-07-16 10 49utput/part-r-00000

hadoop@akanksharani-VirtualBox**:~/**hadoop-3.2.3$ **hdfs dfs -cat /output/part-r-00000**

Big 2

Semistructured 1

Structured 1

Unstructured 1

a 2

also 1

and 4

as 3

by 1

collected 1

combination 1

data 5

data, 3

documents 1

encompasses 1

files. 1

financial 1

following: 1

from 1

including 1

is 1

logs 1

multimedia 1

of 2

organizations. 1

records. 1

semistructured 1

sensors. 1

server 1

streaming 1

structured, 1

such 3

text, 1

the 1

transactions 1

types, 1

unstructured 1

variety 1

web 1

wide 1

hadoop@akanksharani-VirtualBox:~/hadoop-3.2.3$

**github link:**

**https://github.com/Akanksha-rani700/wordcount\_in\_hadoop.git**