

FR. CONCEICAO RODRIGUES COLLEGE OF ENGINEERING

Department of Computer Engineering Academic Year 2025-26

Rubrics for Lab Experiments

Class : *B.E. Computer Engineering* Subject Name : *BDA Lab*

Semester : VII

Subject Code : CSL702

Practical No:	1
Title:	Study and Installation of Hadoop Ecosystem
Date of Performance:	20/07/2025
Roll No:	9913
Name of the Student:	Mark Lopes

Evaluation:

Performance Indicator	Below average	Average	Good	Excellent	Marks
On time Submission (2)	Not submitted (0)	Submitted after deadline (1)	Early or on time submission(2)	---	
Test cases and output (4)	Incorrect output (1)	The expected output is verified only a for few test cases (2)	The expected output is Verified for all test cases but is not presentable (3)	Expected output is obtained for all test cases. Presentable and easy to follow (4)	
Coding efficiency (2)	The code is not structured at all (0)	The code is structured but not efficient (1)	The code is Structured and efficient. (2)	-	
Knowledge(2)	Basic concepts not clear (0)	Understood the basic concepts (1)	Could explain the concept with suitable example (1.5)	Could relate the theory with real world application(2)	
Total					

Experiment No 1

Aim: Study and Installation of Hadoop Ecosystem

Objective:

The objective of this lab experiment is to familiarize students with the Hadoop ecosystem by guiding them through the installation and setup of core components. Students will gain hands-on experience in configuring a basic Hadoop cluster, understanding its architecture, and verifying its functionality.

Tools and Technologies:

- Hadoop: A framework that allows for the distributed processing of large data sets across clusters of computers using simple programming models.
- Hadoop Ecosystem Components: HDFS (Hadoop Distributed File System), YARN (Yet Another Resource Negotiator), and MapReduce.

Pre-requisites:

- Basic understanding of Linux/Unix commands.
- Familiarity with Java programming (helpful but not mandatory).

Equipment Required:

- Virtual or physical machines capable of running a Linux distribution (e.g., Ubuntu, CentOS).
- Sufficient memory and disk space to accommodate Hadoop's requirements (minimum of 4GB RAM recommended per node).

Experiment Steps:

1. Setting Up the Environment:

- Prepare the environment by setting up virtual machines (VMs) or physical machines with a Linux distribution (e.g., Ubuntu Server).
- Ensure that each machine has a static IP address and can communicate with each other over the network.

2. Installing Java Development Kit (JDK):

- Hadoop requires Java, so install JDK on all machines that will be part of the Hadoop cluster.
- Example command to install OpenJDK:

```
bash
Copy code
sudo apt-get update
sudo apt-get install openjdk-8-jdk
```

3. Downloading and Extracting Hadoop:

- Download the desired version of Hadoop from the Apache Hadoop website (<https://hadoop.apache.org/releases.html>).
- Extract the downloaded Hadoop tarball to a suitable directory on each machine in your cluster.

```
bash
Copy code
```

```
tar -xvzf hadoop-3.x.x.tar.gz -C /opt
```

4. **Configuring Hadoop Environment Variables:**

- Set up Hadoop environment variables in the `.bashrc` or `.bash_profile` file for each user:

```
bash
Copy code
export HADOOP_HOME=/opt/hadoop-3.x.x
export PATH=$PATH:$HADOOP_HOME/bin:$HADOOP_HOME/sbin
```

5. **Configuring Hadoop Cluster:**

- **HDFS Configuration:**
 - Edit `core-site.xml` to configure Hadoop core settings, including HDFS filesystem URI and default filesystem.
 - Edit `hdfs-site.xml` to define HDFS block size, replication factor, and namenode/datanode directories.
- **YARN Configuration:**
 - Edit `yarn-site.xml` to configure YARN ResourceManager and NodeManager settings.
 - Optionally, configure `mapred-site.xml` for MapReduce framework settings if not managed by YARN.
- **Setup SSH Authentication:**
 - Enable SSH access between nodes without requiring a password for seamless communication.
 - Generate SSH keys (`ssh-keygen`) and distribute the public key (`ssh-copy-id`) to each node.

6. **Starting Hadoop Cluster:**

- Format the HDFS filesystem on the namenode:

```
bash
Copy code
hdfs namenode -format
```

- Start Hadoop daemons using the provided scripts:

```
bash
Copy code
start-dfs.sh
start-yarn.sh
```

7. **Verifying Hadoop Installation:**

- Access the Hadoop web interfaces:
 - HDFS Namenode: `http://namenode_host:9870/`
 - YARN ResourceManager: `http://resourcemanager_host:8088/`
- Run basic Hadoop commands to ensure functionality:

```
bash
Copy code
hdfs dfs -ls / # List contents of root directory in HDFS
yarn node -list # List nodes in the YARN cluster
```

8. **Performing a Simple MapReduce Job (Optional):**

- Write a basic MapReduce program (e.g., WordCount) or use a pre-existing example.
- Compile and package the program into a JAR file.

- Submit the job to the YARN ResourceManager and monitor its progress using the web interface.

9. Observations and Conclusion:

- Document any issues encountered during setup and how they were resolved.
- Discuss the scalability and fault-tolerance features provided by Hadoop.
- Reflect on the importance of Hadoop in big data processing and its role in modern data architectures.

Expected Outcome:

By the end of this experiment, students should have successfully set up a basic Hadoop cluster comprising HDFS and YARN components. They should be able to navigate Hadoop's web interfaces, execute basic Hadoop commands, and understand the distributed nature of Hadoop processing.

Conclusion:

In this experiment, we successfully installed and configured a basic Hadoop cluster with HDFS and YARN. We learned how to set up the environment, configure core components, and verify the installation using web interfaces and basic commands. This hands-on setup provided foundational insight into Hadoop's architecture, showcasing its scalability, distributed processing, and fault-tolerant capabilities essential for big data applications.

SCREENSHOT:

```
default [tmux] /home/shadow 19:32 19:32:26 [10/557] 3m3s
Adobe_Hack ( main) --- v3.13.5
cd

docker run -it --name hadoop ubuntu bash
root@77756546f17d:/# apt update && apt upgrade -y
Get:1 http://archive.ubuntu.com/ubuntu noble InRelease [256 kB]
Get:2 http://security.ubuntu.com/ubuntu noble-security InRelease [126 kB]
Get:3 http://security.ubuntu.com/ubuntu noble-security/multiverse amd64 Packages [23.0 kB]
Get:4 http://archive.ubuntu.com/ubuntu noble-updates InRelease [126 kB]
Get:5 http://security.ubuntu.com/ubuntu noble-security/restricted amd64 Packages [1820 kB]
Get:6 http://archive.ubuntu.com/ubuntu noble-backports InRelease [126 kB]
Get:7 http://archive.ubuntu.com/ubuntu noble/restricted amd64 Packages [117 kB]
Get:8 http://archive.ubuntu.com/ubuntu noble/main amd64 Packages [1808 kB]
Get:9 http://archive.ubuntu.com/ubuntu noble/universe amd64 Packages [19.3 MB]
Get:10 http://archive.ubuntu.com/ubuntu noble/multiverse amd64 Packages [331 kB]
Get:11 http://archive.ubuntu.com/ubuntu noble-updates/universe amd64 Packages [1443 kB]
Get:12 http://archive.ubuntu.com/ubuntu noble-updates/restricted amd64 Packages [1942 kB]
Get:13 http://archive.ubuntu.com/ubuntu noble-updates/multiverse amd64 Packages [45.2 kB]
Get:14 http://archive.ubuntu.com/ubuntu noble-updates/main amd64 Packages [1608 kB]
Get:15 http://archive.ubuntu.com/ubuntu noble-backports/main amd64 Packages [48.8 kB]
Get:16 http://archive.ubuntu.com/ubuntu noble-backports/universe amd64 Packages [33.0 kB]
Get:17 http://security.ubuntu.com/ubuntu noble-security/universe amd64 Packages [1128 kB]
Get:18 http://security.ubuntu.com/ubuntu noble-security/main amd64 Packages [1270 kB]
Fetched 31.6 MB in 25s (1268 kB/s)
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
21 packages can be upgraded. Run 'apt list --upgradable' to see them.
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
Calculating upgrade... Done
The following packages will be upgraded:
  apt bsdtar gpgv gzip libapt-pkg6.0t64 libblkid1 libc-bin libc6 libgnutls30t64 libmount1 libpam-modules libpam-modules-bin libpam-runtime libpam0g
  libsmartcols1 libsystemd0 libudev1 libuuid1 mount perl-base util-linux
0 packages can be upgraded. Run 'apt list --upgradable' to see them.
```

```
default [tmux] yazi /home/shadow 19:34 19:33:59 [31/31]
root@77756546f17d:/# apt get openjdk-11-jdk openssl-server openssl-client wget curl vim net-tools python3
E: Invalid operation get
root@77756546f17d:/# apt install openjdk-11-jdk openssl-server openssl-client wget curl vim net-tools python3
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following additional packages will be installed:
  adduser alsa-topology-conf alsa-ucm-conf at-spi2-common at-spi2-core ca-certificates ca-certificates-java dbus dbus-bin dbus-daemon
  dbus-session-bus-common dbus-system-bus-common dbus-user-session dconf-gsettings-backend dconf-service dmsetup fontconfig-config fonts-dejavu-core
  fonts-dejavu-extra fonts-dejavu-mono gir1.2-girepository-2.0 gir1.2-glib-2.0 gsettings-desktop-schemas java-common krb5-locales libapparmor1
  libargon2-1 libasound2-data libasound2t64 libatk-bridge2.0-0t64 libatk-wrapper-java libatk-wrapper-java-jni libatk1.0-0t64 libatspi2.0-0t64
  libavahi-client3 libavahi-common-data libavahi-common3 libbrotli1 libbsd0 libcbor0.10 libcryptsetup12 libcups2t64 libcurl4t64 libdbus-1-3 libdconf1
  libdevmapper1.02.1 libdrm-amdgpu1 libdrm-common libdrm-intel1 libdrm-nouveau2 libdrm-radeon1 libdrm2 libedit2 libelf1t64 libexpat1 libfdisk1
  libfido2-1 libfontconfig1 libfreetype6 libgbm1 libgif7 libgirepository-1.0-1 libgl1 libgl1-amd-glx libgl1-mesa-dri libgl1-mesa-dri libglapi-mesa libglvnd0 libglvnd0
  libglvnd0 libglx-mesa0 libglx0 libgpm2 libgraphite2-3 libgssapi-krb5-2 libharfbuzz0b libice-dev libice6 libicu74 libjpeg-turbo8
  libjpeg8 libjson-c5 libk5crypto3 libkeyutils1 libkmod2 libkrb5-3 libkrb5support0 liblcms2-2 libldap-common libldap2 libllvm19 libnghttp2-14
  libnss-pr4 libnss-systemd libnss3 libpam-systemd libpciaccess0 libpcsc-lite1 libpng16-16t64 libpsl5t64 libpthread-stubs0-dev libpython3-stdlib
  libpython3.12-minimal libpython3.12-stdlib libpython3.12t64 libreadline8t64 librtmp1 libsass2-2 libsass2-modules libsass2-modules-db
  libensors-config libensors5 libsm-dev libsm6 libsodium23 libsqlite3-0 libssh-4 libsystemd-shared libvulkan1 libwayland-client0 libwayland-server0
  libwrap0 libx11-6 libx11-data libx11-dev libx11-xcb1 libxau-dev libxau6 libxaw7 libxcb-dri2-0 libxcb-dri3-0 libxcb-glx0 libxcb-present0
  libxcb-randr0 libxcb-shape0 libxcb-shm0 libxcb-sync1 libxcb-xf86vm0 libxcb1 libxcb1-dev libxcomposite1 libxdmcp-dev libxdmcp6 libxext6 libxfixes3
  libxft2 libxi6 libxinerama1 libxkbfile1 libxml2 libxmu6 libxmuu1 libxpm4 libxrandr2 libxrender1 libxshmfence1 libxt-dev libxt6t64 libxtst6 libxv1
  libxxf86dga1 libxxf86vm1 lsb-release media-types mesa-libgallium mesa-vulkan-drivers ncurses-term netbase networkd-dispatcher
  openjdk-11-jdk-headless openjdk-11-jre openjdk-11-jre-headless openssl-sft-server openssl publicsuffix python3-blinker python3-cffi-backend
  python3-cryptography python3-dbus python3-distutils python3-gi python3-httplib2 python3-jwt python3-launchpadlib python3-lazr python3-lazr.restfulclient
  python3-lazr.uri python3-minimal python3-oauthlib python3-pkg-resources python3-pyparsing python3-six python3-wadllib python3.12 python3.12-minimal
  readline-common session-migration shared-mime-info ssh-import-id systemd systemd-dev systemd-resolved systemd-sysv systemd-timesyncd tzdata ucf
  vim-common vim-runtime x11-common x11-utils x11proto-dev xauth xdg-user-dirs xorg-sgml-doctools xtrans-dev xxd
Suggested packages:
  liblocale-gettext-perl perl cron quota ecryptfs-utils default-jre alsa-utils libasound2-plugins cups-common low-memory-monitor gpm krb5-doc
  krb5-user libice-doc liblcms2-utils pciutils pcscd libsass2-modules-gssapi-mit | libsass2-modules-gssapi-heimdal libsass2-modules-ldap
  libsass2-modules-otp libsass2-modules-sql lm-sensors libsm-doc libx11-doc libxcb-doc libxt-doc iw | wireless-tools openjdk-11-demo
  openjdk-11-source visualvm libnss-mdns fonts-ipafont-gothic fonts-ipafont-mincho fonts-wqy-microhei | fonts-wqy-zenhei fonts-indic keychain
  libpam-ssh monkeysphere ssh-askpass molly-guard ufw python3-doc python3-tk python3-venv python-blinker-doc python-cryptography-doc
  python3-cryptography-vectors python-dbus-doc python3-dbus python3-crypto python3-keyring python3-testresources python3-setuptools python-pyparsing-doc
  python3.12-venv python3.12-doc binutils binfmt-support readline-doc systemd-container systemd-homed systemd-userdbd systemd-boot libip4tc2
  libqrencode4 libtss2-esys-3.0.2-0 libtss2-mu-4.0.1-0 libtss2-rc0 libtss2-tcti-device0 polkitd ctags vim-doc vim-scripts mesa-utils
```



```
default 1 docker 2 yazi /home/shadow 19:36
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/javadoc to provide /usr/bin/javadoc (javadoc) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/javap to provide /usr/bin/javap (javap) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jcmd to provide /usr/bin/jcmd (jcmd) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jdb to provide /usr/bin/jdb (jdb) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jdeprscan to provide /usr/bin/jdeprscan (jdeprscan) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jdeps to provide /usr/bin/jdeps (jdeps) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jfr to provide /usr/bin/jfr (jfr) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jimage to provide /usr/bin/jimage (jimage) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jinfo to provide /usr/bin/jinfo (jinfo) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jlink to provide /usr/bin/jlink (jlink) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jmap to provide /usr/bin/jmap (jmap) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jmod to provide /usr/bin/jmod (jmod) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jps to provide /usr/bin/jps (jps) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jrunscript to provide /usr/bin/jrunscript (jrunscript) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jshell to provide /usr/bin/jshell (jshell) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstack to provide /usr/bin/jstack (jstack) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstat to provide /usr/bin/jstat (jstat) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstatd to provide /usr/bin/jstatd (jstatd) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/rmic to provide /usr/bin/rmic (rmic) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/serialver to provide /usr/bin/serialver (serialver) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jaotc to provide /usr/bin/jaotc (jaotc) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jhsdb to provide /usr/bin/jhsdb (jhsdb) in auto mode
Setting up openjdk-11-jre:amd64 (11.0.27+6-ubuntu1~24.04) ...
Setting up openjdk-11-jdk:amd64 (11.0.27+6-ubuntu1~24.04) ...
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jconsole to provide /usr/bin/jconsole (jconsole) in auto mode
Processing triggers for ca-certificates (20240203) ...
Updating certificates in /etc/ssl/certs...
0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...
done.
Processing triggers for ca-certificates-java (20240118) ...
done.
root@77756546f17d:/# java --version
openjdk 11.0.27 2025-04-15
OpenJDK Runtime Environment (build 11.0.27+6-post-Ubuntu-0ubuntu124.04)
OpenJDK 64-Bit Server VM (build 11.0.27+6-post-Ubuntu-0ubuntu124.04, mixed mode, sharing)
root@77756546f17d:/#
```

```
default 1 docker 2 yazi /home/shadow 19:36
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jcmd to provide /usr/bin/jcmd (jcmd) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jdb to provide /usr/bin/jdb (jdb) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jdeprscan to provide /usr/bin/jdeprscan (jdeprscan) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jdeps to provide /usr/bin/jdeps (jdeps) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jfr to provide /usr/bin/jfr (jfr) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jimage to provide /usr/bin/jimage (jimage) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jinfo to provide /usr/bin/jinfo (jinfo) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jlink to provide /usr/bin/jlink (jlink) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jmap to provide /usr/bin/jmap (jmap) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jmod to provide /usr/bin/jmod (jmod) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jps to provide /usr/bin/jps (jps) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jrunscript to provide /usr/bin/jrunscript (jrunscript) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jshell to provide /usr/bin/jshell (jshell) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstack to provide /usr/bin/jstack (jstack) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstat to provide /usr/bin/jstat (jstat) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstatd to provide /usr/bin/jstatd (jstatd) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/rmic to provide /usr/bin/rmic (rmic) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/serialver to provide /usr/bin/serialver (serialver) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jaotc to provide /usr/bin/jaotc (jaotc) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jhsdb to provide /usr/bin/jhsdb (jhsdb) in auto mode
Setting up openjdk-11-jre:amd64 (11.0.27+6-ubuntu1~24.04) ...
Setting up openjdk-11-jdk:amd64 (11.0.27+6-ubuntu1~24.04) ...
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jconsole to provide /usr/bin/jconsole (jconsole) in auto mode
Processing triggers for ca-certificates (20240203) ...
Updating certificates in /etc/ssl/certs...
0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...
done.
Processing triggers for ca-certificates-java (20240118) ...
done.
root@77756546f17d:/# java --version
openjdk 11.0.27 2025-04-15
OpenJDK Runtime Environment (build 11.0.27+6-post-Ubuntu-0ubuntu124.04)
OpenJDK 64-Bit Server VM (build 11.0.27+6-post-Ubuntu-0ubuntu124.04, mixed mode, sharing)
root@77756546f17d:/# python3 --version
Python 3.12.3
root@77756546f17d:/#
```

```
default 1 docker 2 yazi /home/shadow 19:37
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jimage to provide /usr/bin/jimage (jimage) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jinfo to provide /usr/bin/jinfo (jinfo) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jlink to provide /usr/bin/jlink (jlink) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jmap to provide /usr/bin/jmap (jmap) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jmod to provide /usr/bin/jmod (jmod) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jps to provide /usr/bin/jps (jps) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jrunscript to provide /usr/bin/jrunscript (jrunscript) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jshell to provide /usr/bin/jshell (jshell) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstack to provide /usr/bin/jstack (jstack) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstat to provide /usr/bin/jstat (jstat) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstatd to provide /usr/bin/jstatd (jstatd) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/rmic to provide /usr/bin/rmic (rmic) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/serialver to provide /usr/bin/serialver (serialver) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jaotc to provide /usr/bin/jaotc (jaotc) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jhsdb to provide /usr/bin/jhsdb (jhsdb) in auto mode
Setting up openjdk-11-jre:amd64 (11.0.27+6~us1-0ubuntu1~24.04) ...
Setting up openjdk-11-jdk:amd64 (11.0.27+6~us1-0ubuntu1~24.04) ...
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jconsole to provide /usr/bin/jconsole (jconsole) in auto mode
Processing triggers for ca-certificates (20240203) ...
Updating certificates in /etc/ssl/certs...
0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...
done.
Processing triggers for ca-certificates-java (20240118) ...
done.
root@77756546f17d:/# java --version
openjdk 11.0.27 2025-04-15
OpenJDK Runtime Environment (build 11.0.27+6-post-Ubuntu-0ubuntu124.04)
OpenJDK 64-Bit Server VM (build 11.0.27+6-post-Ubuntu-0ubuntu124.04, mixed mode, sharing)
root@77756546f17d:/# python3 --version
Python 3.12.3
root@77756546f17d:/# echo 'export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64' >> ~/.bashrc
root@77756546f17d:/# cd
root@77756546f17d:/# source .bashrc
root@77756546f17d:/# echo $JAVA_HOME/
/usr/lib/jvm/java-11-openjdk-amd64/
root@77756546f17d:/#
```

```
default 1 docker 2 yazi /home/shadow 19:38
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jrunscript to provide /usr/bin/jrunscript (jrunscript) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jshell to provide /usr/bin/jshell (jshell) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstack to provide /usr/bin/jstack (jstack) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstat to provide /usr/bin/jstat (jstat) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jstatd to provide /usr/bin/jstatd (jstatd) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/rmic to provide /usr/bin/rmic (rmic) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/serialver to provide /usr/bin/serialver (serialver) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jaotc to provide /usr/bin/jaotc (jaotc) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jhsdb to provide /usr/bin/jhsdb (jhsdb) in auto mode
Setting up openjdk-11-jre:amd64 (11.0.27+6~us1-0ubuntu1~24.04) ...
Setting up openjdk-11-jdk:amd64 (11.0.27+6~us1-0ubuntu1~24.04) ...
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jconsole to provide /usr/bin/jconsole (jconsole) in auto mode
Processing triggers for ca-certificates (20240203) ...
Updating certificates in /etc/ssl/certs...
0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...
done.
Processing triggers for ca-certificates-java (20240118) ...
done.
root@77756546f17d:/# java --version
openjdk 11.0.27 2025-04-15
OpenJDK Runtime Environment (build 11.0.27+6-post-Ubuntu-0ubuntu124.04)
OpenJDK 64-Bit Server VM (build 11.0.27+6-post-Ubuntu-0ubuntu124.04, mixed mode, sharing)
root@77756546f17d:/# python3 --version
Python 3.12.3
root@77756546f17d:/# echo 'export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64' >> ~/.bashrc
root@77756546f17d:/# cd
root@77756546f17d:/# source .bashrc
root@77756546f17d:/# echo $JAVA_HOME/
/usr/lib/jvm/java-11-openjdk-amd64/
root@77756546f17d:/# useradd -m -s /bin/bash hadoop
root@77756546f17d:/# echo "hadoop:hadoop" | chpasswd
root@77756546f17d:/# usermod -aG sudo hadoop
root@77756546f17d:/# su - hadoop
hadoop@77756546f17d:~$ pwd
/home/hadoop
hadoop@77756546f17d:~$
```



```

default 1 docker 2 yazi /home/shadow 19:40
E: Could not open lock file /var/lib/dpkg/lock-frontent - open (13: Permission denied)
E: Unable to acquire the dpkg frontend lock (/var/lib/dpkg/lock-frontent), are you root?
hadoop@77756546f17d:~$ su
Password:
su: Authentication failure
hadoop@77756546f17d:~$ exit
logout
root@77756546f17d:~# apt install sudo
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following NEW packages will be installed:
  sudo
0 upgraded, 1 newly installed, 0 to remove and 0 not upgraded.
Need to get 948 kB of archives.
After this operation, 3551 kB of additional disk space will be used.
Get:1 http://archive.ubuntu.com/ubuntu noble-updates/main amd64 sudo amd64 1.9.15p5-3ubuntu5.24.04.1 [948 kB]
Fetched 948 kB in 3s (361 kB/s)
debconf: delaying package configuration, since apt-utils is not installed
Selecting previously unselected package sudo.
(Reading database ... 16741 files and directories currently installed.)
Preparing to unpack .../sudo_1.9.15p5-3ubuntu5.24.04.1_amd64.deb ...
Unpacking sudo (1.9.15p5-3ubuntu5.24.04.1) ...
Setting up sudo (1.9.15p5-3ubuntu5.24.04.1) ...
Processing triggers for libc-bin (2.39-0ubuntu8.5) ...
root@77756546f17d:~# passwd
New password:
Retype new password:
passwd: password updated successfully
root@77756546f17d:~# su - hadoop
To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

hadoop@77756546f17d:~$ sudo service ssh start
[sudo] password for hadoop:
* Starting OpenBSD Secure Shell server sshd
[ OK ]
hadoop@77756546f17d:~$

```

```

default 1 docker /home/shadow 19:47
0 upgraded, 1 newly installed, 0 to remove and 0 not upgraded.
Need to get 948 kB of archives.
After this operation, 3551 kB of additional disk space will be used.
Get:1 http://archive.ubuntu.com/ubuntu noble-updates/main amd64 sudo amd64 1.9.15p5-3ubuntu5.24.04.1 [948 kB]
Fetched 948 kB in 3s (361 kB/s)
debconf: delaying package configuration, since apt-utils is not installed
Selecting previously unselected package sudo.
(Reading database ... 16741 files and directories currently installed.)
Preparing to unpack .../sudo_1.9.15p5-3ubuntu5.24.04.1_amd64.deb ...
Unpacking sudo (1.9.15p5-3ubuntu5.24.04.1) ...
Setting up sudo (1.9.15p5-3ubuntu5.24.04.1) ...
Processing triggers for libc-bin (2.39-0ubuntu8.5) ...
root@77756546f17d:~# passwd
New password:
Retype new password:
passwd: password updated successfully
root@77756546f17d:~# su - hadoop
To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

hadoop@77756546f17d:~$ sudo service ssh start
[sudo] password for hadoop:
* Starting OpenBSD Secure Shell server sshd
[ OK ]
hadoop@77756546f17d:~$ cd
hadoop@77756546f17d:~$ wget https://downloads.apache.org/hadoop/common/hadoop-3.4.1/hadoop-3.4.1.tar.gz
--2025-07-21 19:40:38-- https://downloads.apache.org/hadoop/common/hadoop-3.4.1/hadoop-3.4.1.tar.gz
Resolving downloads.apache.org (downloads.apache.org)... 135.181.214.104, 88.99.208.237, 2a01:4f9:3a:2c57::2, ...
Connecting to downloads.apache.org (downloads.apache.org)|135.181.214.104|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 974002355 (929M) [application/x-gzip]
Saving to: 'hadoop-3.4.1.tar.gz'

hadoop-3.4.1.tar.gz          100%[=====>] 928.88M  2.91MB/s   in 6m 31s

2025-07-21 19:47:10 (2.38 MB/s) - 'hadoop-3.4.1.tar.gz' saved [974002355/974002355]

hadoop@77756546f17d:~$

```



```
default [tmux] /home/shadow 19:49
Setting up sudo (1.9.15p5-3ubuntu5.24.04.1) ...
Processing triggers for libc-bin (2.39-0ubuntu8.5) ...
root@77756546f17d:~# passwd
New password:
Retype new password:
passwd: password updated successfully
root@77756546f17d:~# su - hadoop
To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

hadoop@77756546f17d:~$ sudo service ssh start
[sudo] password for hadoop:
* Starting OpenBSD Secure Shell server sshd [ OK ]
hadoop@77756546f17d:~$ cd
hadoop@77756546f17d:~$ wget https://downloads.apache.org/hadoop/common/hadoop-3.4.1/hadoop-3.4.1.tar.gz
--2025-07-21 19:40:38-- https://downloads.apache.org/hadoop/common/hadoop-3.4.1/hadoop-3.4.1.tar.gz
Resolving downloads.apache.org (downloads.apache.org)... 135.181.214.104, 88.99.208.237, 2a01:4f9:3a:2c57::2, ...
Connecting to downloads.apache.org (downloads.apache.org)[135.181.214.104]:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 974002355 (929M) [application/x-gzip]
Saving to: 'hadoop-3.4.1.tar.gz'

hadoop-3.4.1.tar.gz 100%[=====] 928.88M 2.91MB/s in 6m 31s

2025-07-21 19:47:10 (2.38 MB/s) - 'hadoop-3.4.1.tar.gz' saved [974002355/974002355]

hadoop@77756546f17d:~$ tar -xvzf hadoop-3.4.1.tar.gz
hadoop-3.4.1/
hadoop-3.4.1/include/
hadoop-3.4.1/include/SerialUtils.hh
hadoop-3.4.1/include/TemplateFactory.hh
hadoop-3.4.1/include/hdfs.h
hadoop-3.4.1/include/StringUtils.hh
hadoop-3.4.1/include/Pipes.hh
hadoop-3.4.1/share/
hadoop-3.4.1/share/doc/
hadoop-3.4.1/share/doc/hadoop/

Screenshot copied to clipboard [32126/33805]

default [tmux] /home/shadow 19:49
Connecting to downloads.apache.org (downloads.apache.org)[135.181.214.104]:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 974002355 (929M) [application/x-gzip]
Saving to: 'hadoop-3.4.1.tar.gz'

hadoop-3.4.1.tar.gz 100%[=====] 928.88M 2.91MB/s in 6m 31s

2025-07-21 19:47:10 (2.38 MB/s) - 'hadoop-3.4.1.tar.gz' saved [974002355/974002355]

hadoop@77756546f17d:~$ tar -xvzf hadoop-3.4.1.tar.gz
hadoop-3.4.1/
hadoop-3.4.1/include/
hadoop-3.4.1/include/SerialUtils.hh
hadoop-3.4.1/include/TemplateFactory.hh
hadoop-3.4.1/include/hdfs.h
hadoop-3.4.1/include/StringUtils.hh
hadoop-3.4.1/include/Pipes.hh
hadoop-3.4.1/share/
hadoop-3.4.1/share/doc/
hadoop-3.4.1/share/doc/hadoop/
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/expanded.gif
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/maven-logo-2.gif
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/banner.jpg
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/bg.jpg
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/collapsed.gif
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/icon_info_sml.gif
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/logo_apache.jpg
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/logos/
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/logos/build-by-maven-white.png
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/logos/maven-feather.png
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/logos/build-by-maven-black.png
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/icon_success_sml.gif
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/logo_maven.jpg
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/apache-maven-project-2.png
hadoop-3.4.1/share/doc/hadoop/hadoop-kms/images/h3.jpg
```

```

default 1 docker
alias l='ls -CF'

# Add an "alert" alias for long running commands.  Use like so:
# sleep 10; alert
alias alert='notify-send --urgency=low -i "${[ $? = 0 ]} && echo terminal || echo error)" "$(history|tail -n1|sed -e '\''s/^\s*[0-9]\+\s*//;s/[:&|]\s*alert$/'\''")"'

# Alias definitions.
# You may want to put all your additions into a separate file like
# ~/.bash_aliases, instead of adding them here directly.
# See /usr/share/doc/bash-doc/examples in the bash-doc package.

if [ -f ~/.bash_aliases ]; then
    . ~/.bash_aliases
fi

# enable programmable completion features (you don't need to enable
# this, if it's already enabled in /etc/bash.bashrc and /etc/profile
# sources /etc/bash.bashrc).
if ! shopt -oq posix; then
    if [ -f /usr/share/bash-completion/bash_completion ]; then
        . /usr/share/bash-completion/bash_completion
    elif [ -f /etc/bash_completion ]; then
        . /etc/bash_completion
    fi
fi

export HADOOP_HOME=/home/hadoop/hadoop
export HADOOP_INSTALL=$HADOOP_HOME
export HADOOP_MAPRED_HOME=$HADOOP_HOME
export HADOOP_COMMON_HOME=$HADOOP_HOME
export HADOOP_HDFS_HOME=$HADOOP_HOME
export HADOOP_YARN_HOME=$HADOOP_HOME
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export PATH=$PATH:$HADOOP_HOME/sbin:$HADOOP_HOME/bin
hadoop@77756546f17d:~$

```

```

default 1 docker
alias alert='notify-send --urgency=low -i "${[ $? = 0 ]} && echo terminal || echo error)" "$(history|tail -n1|sed -e '\''s/^\s*[0-9]\+\s*//;s/[:&|]\s*alert$/'\''")"'

# Alias definitions.
# You may want to put all your additions into a separate file like
# ~/.bash_aliases, instead of adding them here directly.
# See /usr/share/doc/bash-doc/examples in the bash-doc package.

if [ -f ~/.bash_aliases ]; then
    . ~/.bash_aliases
fi

# enable programmable completion features (you don't need to enable
# this, if it's already enabled in /etc/bash.bashrc and /etc/profile
# sources /etc/bash.bashrc).
if ! shopt -oq posix; then
    if [ -f /usr/share/bash-completion/bash_completion ]; then
        . /usr/share/bash-completion/bash_completion
    elif [ -f /etc/bash_completion ]; then
        . /etc/bash_completion
    fi
fi

export HADOOP_HOME=/home/hadoop/hadoop
export HADOOP_INSTALL=$HADOOP_HOME
export HADOOP_MAPRED_HOME=$HADOOP_HOME
export HADOOP_COMMON_HOME=$HADOOP_HOME
export HADOOP_HDFS_HOME=$HADOOP_HOME
export HADOOP_YARN_HOME=$HADOOP_HOME
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export PATH=$PATH:$HADOOP_HOME/sbin:$HADOOP_HOME/bin
hadoop@77756546f17d:~$ echo "export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64" >> $HADOOP_HOME/etc/hadoop/hadoop-env.sh
-bash: /etc/hadoop/hadoop-env.sh: No such file or directory
hadoop@77756546f17d:~$ source .bashrc
hadoop@77756546f17d:~$ echo "export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64" >> $HADOOP_HOME/etc/hadoop/hadoop-env.sh
hadoop@77756546f17d:~$

```

```

default [tmux] /home/shadow 19:52 [0/128]
export HADOOP_HDFS_HOME=$HADOOP_HOME
export HADOOP_YARN_HOME=$HADOOP_HOME
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export PATH=$PATH:$HADOOP_HOME/sbin:$HADOOP_HOME/bin
hadoop@77756546f17d:~$ echo "export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64" >> $HADOOP_HOME/etc/hadoop/hadoop-env.sh
-bash: /etc/hadoop/hadoop-env.sh: No such file or directory
hadoop@77756546f17d:~$ source .bashrc
hadoop@77756546f17d:~$ echo "export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64" >> $HADOOP_HOME/etc/hadoop/hadoop-env.sh
hadoop@77756546f17d:~$ mkdir -p ~/hadoopdata/hdfs/namenode
hadoop@77756546f17d:~$ mkdir -p ~/hadoopdata/hdfs/datanode
hadoop@77756546f17d:~$ cat > $HADOOP_HOME/etc/hadoop/core-site.xml << 'EOF'
<?xml version="1.0" encoding="UTF-8"?>
<configuration>
  <property>
    <name>fs.defaultFS</name>
    <value>hdfs://localhost:9000</value>
  </property>
</configuration>
EOF
hadoop@77756546f17d:~$ cat > $HADOOP_HOME/etc/hadoop/hdfs-site.xml << 'EOF'
<?xml version="1.0" encoding="UTF-8"?>
<configuration>
  <property>
    <name>dfs.replication</name>
    <value>1</value>
  </property>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>/home/hadoop/hadoopdata/hdfs/namenode</value>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>/home/hadoop/hadoopdata/hdfs/datanode</value>
  </property>
</configuration>
EOF
hadoop@77756546f17d:~$

```

```

default [tmux] /home/shadow 19:53 [0/146]
EOF
hadoop@77756546f17d:~$ cat > $HADOOP_HOME/etc/hadoop/hdfs-site.xml << 'EOF'
<?xml version="1.0" encoding="UTF-8"?>
<configuration>
  <property>
    <name>dfs.replication</name>
    <value>1</value>
  </property>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>/home/hadoop/hadoopdata/hdfs/namenode</value>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>/home/hadoop/hadoopdata/hdfs/datanode</value>
  </property>
</configuration>
EOF
hadoop@77756546f17d:~$ cat > $HADOOP_HOME/etc/hadoop/mapred-site.xml << 'EOF'
<?xml version="1.0" encoding="UTF-8"?>
<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
</configuration>
EOF
hadoop@77756546f17d:~$ cat > $HADOOP_HOME/etc/hadoop/yarn-site.xml << 'EOF'
<?xml version="1.0" encoding="UTF-8"?>
<configuration>
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
</configuration>
EOF
hadoop@77756546f17d:~$

```



```

default [tmux]
~/configuration> EOF
hadoop@77756546f17d:~$ cat > $HADOOP_HOME/etc/hadoop/yarn-site.xml << 'EOF'
<?xml version="1.0" encoding="UTF-8"?>
<configuration>
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
</configuration>
EOF
hadoop@77756546f17d:~$ hdfs
Usage: hdfs [OPTIONS] SUBCOMMAND [SUBCOMMAND OPTIONS]

OPTIONS is none or any of:

--buildpaths          attempt to add class files from build tree
--config dir          Hadoop config directory
--daemon (start|status|stop) operate on a daemon
--debug              turn on shell script debug mode
--help              usage information
--hostnames list[,of,host,names] hosts to use in worker mode
--hosts filename      list of hosts to use in worker mode
--loglevel level      set the log4j level for this command
--workers            turn on worker mode

SUBCOMMAND is one of:

Admin Commands:

cacheadmin            configure the HDFS cache
crypto                configure HDFS encryption zones
debug                run a Debug Admin to execute HDFS debug commands
dfsadmin              run a DFS admin client
dfsrouteradmin        manage Router-based federation
ec                    run a HDFS ErasureCoding CLI

```

```

[1] default [1] [tmux]
namenode      run the DFS namenode
nfs3          run an NFS version 3 gateway
portmap       run a portmap service
secondarynamenode run the DFS secondary namenode
sps           run external storagepolicysatisfier
zkfc          run the ZK Failover Controller daemon

SUBCOMMAND may print help when invoked w/o parameters or with -h.
hadoop@77756546f17d:~$ hdfs namenode -format -force # only once
WARNING: /home/hadoop/hadoop/logs does not exist. Creating.
2025-07-21 19:54:05,921 INFO namenode.NameNode: STARTUP_MSG:
/*****
STARTUP_MSG: Starting NameNode
STARTUP_MSG: host = 77756546f17d/172.18.0.2
STARTUP_MSG: args = [-format, -force]
STARTUP_MSG: version = 3.4.1
STARTUP_MSG: classpath = /home/hadoop/hadoop/etc/hadoop:/home/hadoop/hadoop/share/hadoop/common/lib/jetty-servlet-9.4.53.v20231009.jar:/home/hadoop/h
adoop/share/hadoop/common/lib/jackson-databind-2.12.7.1.jar:/home/hadoop/hadoop/share/hadoop/common/lib/netty-all-4.1.100.Final.jar:/home/hadoop/hadoop
/share/hadoop/common/lib/netty-transport-native-unix-common-4.1.100.Final.jar:/home/hadoop/hadoop/share/hadoop/common/lib/jetty-webapp-9.4.53.v20231009
.jar:/home/hadoop/hadoop/share/hadoop/common/lib/netty-resolver-dns-classes-macos-4.1.100.Final.jar:/home/hadoop/hadoop/share/hadoop/common/lib/jsr305-
3.0.2.jar:/home/hadoop/hadoop/share/hadoop/common/lib/jakarta.activation-api-1.2.1.jar:/home/hadoop/hadoop/share/hadoop/common/lib/netty-codec-smtp-4.1
.100.Final.jar:/home/hadoop/hadoop/share/hadoop/common/lib/jettison-1.5.4.jar:/home/hadoop/hadoop/share/hadoop/common/lib/netty-buffer-4.1.100.Final.ja
r:/home/hadoop/hadoop/share/hadoop/common/lib/curator-recipes-5.2.0.jar:/home/hadoop/hadoop/share/hadoop/common/lib/netty-transport-sctp-4.1.100.Final.ja
r:/home/hadoop/hadoop/share/hadoop/common/lib/jetty-http-9.4.53.v20231009.jar:/home/hadoop/hadoop/share/hadoop/common/lib/httpcore-4.4.13.jar:/home/h
adoop/hadoop/share/hadoop/common/lib/jersey-servlet-1.19.4.jar:/home/hadoop/hadoop/share/hadoop/common/lib/jsr311-api-1.1.1.jar:/home/hadoop/hadoop/sha
re/hadoop/common/lib/netty-handler-4.1.100.Final.jar:/home/hadoop/hadoop/share/hadoop/common/lib/netty-resolver-dns-4.1.100.Final.jar:/home/hadoop/hado
op/share/hadoop/common/lib/jetty-io-9.4.53.v20231009.jar:/home/hadoop/hadoop/share/hadoop/common/lib/dnsjava-3.6.1.jar:/home/hadoop/hadoop/share/hadoop/c
ommon/lib/jersey-core-1.19.4.jar:/home/hadoop/hadoop/share/hadoop/common/lib/jetty-xml-9.4.53.v20231009.jar:/home/hadoop/hadoop/share/hadoop/common/l
ib/kerb-core-2.0.3.jar:/home/hadoop/hadoop/share/hadoop/common/lib/netty-codec-dns-4.1.100.Final.jar:/home/hadoop/hadoop/share/hadoop/common/lib/hadoop
-auth-3.4.1.jar:/home/hadoop/hadoop/share/hadoop/common/lib/commons-daemon-1.0.13.jar:/home/hadoop/hadoop/share/hadoop/common/lib/jersey-json-1.22.0.ja
r:/home/hadoop/hadoop/share/hadoop/common/lib/netty-transport-classes-kqueue-4.1.100.Final.jar:/home/hadoop/hadoop/share/hadoop/common/lib/jul-to-slf4j
-1.7.36.jar:/home/hadoop/hadoop/share/hadoop/common/lib/slf4j-reload4j-1.7.36.jar:/home/hadoop/hadoop/share/hadoop/common/lib/commons-compress-1.26.1.j
ar:/home/hadoop/hadoop/share/hadoop/common/lib/audience-annotations-0.12.0.jar:/home/hadoop/hadoop/share/hadoop/common/lib/avro-1.9.2.jar:/home/hadoop/hado
op/share/hadoop/common/lib/hadoop-shaded-protobuf-3.25-1.3.0.jar:/home/hadoop/hadoop/share/hadoop/common/lib/kerby-pkix-2.0.3.jar:/home/hadoop/hado
op/share/hadoop/common/lib/netty-resolver-dns-native-macos-4.1.100.Final-osx-aarch_64.jar:/home/hadoop/hadoop/share/hadoop/common/lib/javax.servlet-api
-3.1.0.jar:/home/hadoop/hadoop/share/hadoop/common/lib/netty-transport-native-epoll-4.1.100.Final.jar:/home/hadoop/hadoop/share/hadoop/common/lib/commo
ns-text-1.10.0.jar:/home/hadoop/hadoop/share/hadoop/common/lib/netty-resolver-4.1.100.Final.jar:/home/hadoop/hadoop/share/hadoop/common/lib/woodstox-co

```



```
default [tmux] /home/shadow 19:55 [0/480]
2025-07-21 19:54:06,712 INFO namenode.FSImage: Allocated new BlockPoolId: BP-1531782715-172.18.0.2-1753107846707
2025-07-21 19:54:06,734 INFO common.Storage: Storage directory /home/hadoop/hadoopdata/hdfs/namenode has been successfully formatted.
2025-07-21 19:54:06,763 INFO namenode.FSImageFormatProtobuf: Saving image file /home/hadoop/hadoopdata/hdfs/namenode/current/fsimage.ckpt_00000000000000000000 using no compression
2025-07-21 19:54:06,862 INFO namenode.FSImageFormatProtobuf: Image file /home/hadoop/hadoopdata/hdfs/namenode/current/fsimage.ckpt_00000000000000000000 of size 401 bytes saved in 0 seconds .
2025-07-21 19:54:06,876 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
2025-07-21 19:54:06,881 INFO blockmanagement.DatanodeManager: Slow peers collection thread shutdown
2025-07-21 19:54:06,897 INFO namenode.FSNamesystem: Stopping services started for active state
2025-07-21 19:54:06,898 INFO namenode.FSNamesystem: Stopping services started for standby state
2025-07-21 19:54:06,901 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when meet shutdown.
2025-07-21 19:54:06,902 INFO namenode.NameNode: SHUTDOWN_MSG:
/*****
SHUTDOWN_MSG: Shutting down NameNode at 77756546f17d/172.18.0.2
*****/
hadoop@77756546f17d:~$ start-
start-all.cmd      start-balancer.sh      start-dfs.sh          start-stop-daemon      start-yarn.sh
start-all.sh        start-dfs.cmd          start-secure-dns.sh   start-yarn.cmd
hadoop@77756546f17d:~$ start-dfs.sh
Starting namenodes on [localhost]
localhost: Warning: Permanently added 'localhost' (EO25519) to the list of known hosts.
Starting datanodes
Starting secondary namenodes [77756546f17d]
77756546f17d: Warning: Permanently added '77756546f17d' (EO25519) to the list of known hosts.
2025-07-21 19:54:47,032 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@77756546f17d:~$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
hadoop@77756546f17d:~$ jps
7793 Jps
7139 SecondaryNameNode
6964 DataNode
7336 ResourceManager
6825 NameNode
7451 NodeManager
hadoop@77756546f17d:~$

default [tmux] /home/shadow 19:56 [0/480]
Starting namenodes on [localhost]
localhost: Warning: Permanently added 'localhost' (EO25519) to the list of known hosts.
Starting datanodes
Starting secondary namenodes [77756546f17d]
77756546f17d: Warning: Permanently added '77756546f17d' (EO25519) to the list of known hosts.
2025-07-21 19:54:47,032 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@77756546f17d:~$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
hadoop@77756546f17d:~$ jps
7793 Jps
7139 SecondaryNameNode
6964 DataNode
7336 ResourceManager
6825 NameNode
7451 NodeManager
hadoop@77756546f17d:~$ hdfs dfs -ls
2025-07-21 19:55:40,733 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
ls: '.': No such file or directory
hadoop@77756546f17d:~$ hdfs dfs -mkdir /user
2025-07-21 19:56:00,376 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@77756546f17d:~$ hdfs dfs -mkdir /user/hadoop
2025-07-21 19:56:06,207 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@77756546f17d:~$ echo "Hello Hadoop World" > test.txt
hadoop@77756546f17d:~$ hdfs dfs -put test.txt /user/hadoop
2025-07-21 19:56:35,130 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@77756546f17d:~$ hdfs dfs -ls /user/hadoop
2025-07-21 19:56:47,048 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
-rw-r--r-- 1 hadoop supergroup      19 2025-07-21 19:56 /user/hadoop/test.txt
hadoop@77756546f17d:~$
```

```
default 1 [tmux] /home/shadow 19:57 [0/503]
77756546f17d: Warning: Permanently added '77756546f17d' (E025519) to the list of known hosts.
2025-07-21 19:54:47,032 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicab
le
hadoop@77756546f17d:~$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
hadoop@77756546f17d:~$ jps
7793 Jps
7139 SecondaryNameNode
6964 DataNode
7336 ResourceManager
6825 NameNode
7451 NodeManager
hadoop@77756546f17d:~$ hdfs dfs -ls
2025-07-21 19:55:40,733 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicab
le
ls: '.': No such file or directory
hadoop@77756546f17d:~$ hdfs dfs -mkdir /user
2025-07-21 19:56:00,376 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicab
le
hadoop@77756546f17d:~$ hdfs dfs -mkdir /user/hadoop
2025-07-21 19:56:06,207 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicab
le
hadoop@77756546f17d:~$ echo "Hello Hadoop World" > test.txt
hadoop@77756546f17d:~$ hdfs dfs -put test.txt /user/hadoop
2025-07-21 19:56:35,130 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicab
le
hadoop@77756546f17d:~$ hdfs dfs -ls /user/hadoop
2025-07-21 19:56:47,048 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicab
le
Found 1 items
-rw-r--r-- 1 hadoop supergroup 19 2025-07-21 19:56 /user/hadoop/test.txt
hadoop@77756546f17d:~$ hdfs dfs -cat /user/hadoop/test.txt
2025-07-21 19:57:11,552 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicab
le
Hello Hadoop World
hadoop@77756546f17d:~$
```

```
default 1 docker 2 yazi /home/shadow 20:09 2s
↳ docker run -it -p 9870:9870 --name hadoop local-hadoop:latest bash
root@d9288aefa69f:/# stop-yarn.sh
bash: stop-yarn.sh: command not found
root@d9288aefa69f:/# su - hadoop
hadoop@d9288aefa69f:~$ jps
14 Jps
hadoop@d9288aefa69f:~$ stop-yarn.sh
Stopping nodemanagers
localhost: ssh: connect to host localhost port 22: Connection refused
Stopping resourcemanager
hadoop@d9288aefa69f:~$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
localhost: ssh: connect to host localhost port 22: Connection refused
hadoop@d9288aefa69f:~$ start-dfs.sh
Starting namenodes on [localhost]
localhost: ssh: connect to host localhost port 22: Connection refused
Starting datanodes
localhost: ssh: connect to host localhost port 22: Connection refused
Starting secondary namenodes [d9288aefa69f]
d9288aefa69f: ssh: connect to host d9288aefa69f port 22: Connection refused
2025-07-21 20:08:20,198 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicab
le
hadoop@d9288aefa69f:~$ sudo service ssh start
[sudo] password for hadoop:
* Starting OpenBSD Secure Shell server sshd
hadoop@d9288aefa69f:~$ stop-yarn.sh
Stopping nodemanagers
Stopping resourcemanager
hadoop@d9288aefa69f:~$ stop-dfs.sh
Stopping namenodes on [localhost]
Stopping datanodes
Stopping secondary namenodes [d9288aefa69f]
d9288aefa69f: Warning: Permanently added 'd9288aefa69f' (E025519) to the list of known hosts.
2025-07-21 20:09:05,350 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicab
le
hadoop@d9288aefa69f:~$
```



```
default [1] [tmux] [2] yazi /home/shadow 20:10
hadoop@f68125dfa980:~$ stop-dfs.sh
Stopping namenodes on [localhost]
localhost: ssh: connect to host localhost port 22: Connection refused
Stopping datanodes
localhost: ssh: connect to host localhost port 22: Connection refused
Stopping secondary namenodes [f68125dfa980]
f68125dfa980: ssh: connect to host f68125dfa980 port 22: Connection refused
2025-07-21 20:07:05,695 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicab
le
hadoop@f68125dfa980:~$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
localhost: ssh: connect to host localhost port 22: Connection refused
hadoop@f68125dfa980:~$ ^C
hadoop@f68125dfa980:~$ exit
logout
root@f68125dfa980:/# exit
exit

ld 47s

docker run -it -p 9870:9870 --name hadoop local-hadoop:latest bash 2s
root@d9288aefa69f:/# stop-yarn.sh
bash: stop-yarn.sh: command not found
root@d9288aefa69f:/# su - hadoop
hadoop@d9288aefa69f:~$ jps
14 Jps
hadoop@d9288aefa69f:~$ stop-yarn.sh
Stopping nodemanagers
localhost: ssh: connect to host localhost port 22: Connection refused
Stopping resourcemanager
hadoop@d9288aefa69f:~$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
localhost: ssh: connect to host localhost port 22: Connection refused
```

Hadoop

Overview

Datanodes

Datanode Volume Failures

Snapshot

Startup Progress

Utilities

Datanode Information

✓ In service

❗ Down

🔄 Decommissioning

🚫 Decommissioned

🛑 Decommissioned & dead

👉 Entering Maintenance

🔧 In Maintenance

🛑 In Maintenance & dead

Datanode usage histogram

Disk usage of each DataNode (%)	Count
0	1

In operation

DataNode State

All

Show

25

entries

Search:

Node	Http Address	Last contact	Last Block Report	Used	Non DFS Used	Capacity	Blocks	Block pool used	Block pool usage	StdDev	Version
------	--------------	--------------	-------------------	------	--------------	----------	--------	-----------------	------------------	--------	---------

Hadoop

Overview

Datanodes

Datanode Volume Failures

Snapshot

Startup Progress

Utilities

Startup Progress

Elapsed Time: 0 sec, Percent Complete: 100%

Phase	Completion	Elapsed Time
Loading fsimage /home/hadoop/hadoopdata/hdfs/namenode/current/fsimage_000000000000000000 401 B	100%	0 sec
erasure coding policies (0/0)	100%	
inodes (1/1)	100%	
delegation tokens (0/0)	100%	
cache pools (0/0)	100%	
Loading edits	100%	0 sec
/home/hadoop/hadoopdata/hdfs/namenode/current/edits_0000000000000000001-0000000000000000002 42 B (2/2)	100%	
/home/hadoop/hadoopdata/hdfs/namenode/current/edits_0000000000000000003-0000000000000000011 1 MB (9/9)	100%	
Saving checkpoint	100%	0 sec
Safe mode	100%	0 sec
awaiting reported blocks (0/1)	100%	

Hadoop, 2024.

Postlab:-

1. What are the main components of a Hadoop application?

HDFS (Hadoop Distributed File System):

Stores large files across multiple machines with fault tolerance using replication.

YARN (Yet Another Resource Negotiator):

Manages cluster resources and job scheduling.

MapReduce:

A programming model used for distributed data processing (map = split, reduce = aggregate).

Hadoop Common:

Provides essential Java libraries and utilities used by other modules.

2. Difference between NameNode, Backup Node, and Checkpoint Node:

Component	Function	Real-Time Sync	Failure Recovery Role
NameNode	Manages file system metadata like file names, directories, and block locations.	Yes	Acts as the master; essential for HDFS operation.
Backup Node	Maintains an in-memory, up-to-date copy of metadata from the NameNode.	Yes	Can immediately take over if NameNode fails.
Checkpoint Node	Periodically downloads and merges fsimage and edits, then sends a new fsimage to NameNode.	No	Reduces NameNode startup time, not used for failover.

3. Explain the use of cat, du, du -s:

- `cat` (concatenate):
Used to view the contents of files in the terminal. Example: `cat file.txt`
- `du` (disk usage):
Shows the space used by files and directories. Example: `du myfolder/`
- `du -s` (summary):
Displays the total size of a folder, instead of listing all subdirectories. Example: `du -s myfolder/`