EXP NO.: 1 **CREATE A WEB SERVICE USING JAX-WS AND INVOKING IT FROM**
DATE: **A JAVA CLIENT**

AIM:

To create a web service using JAX-WS and invoke it from a standalone java client.

PROGRAM:

Server: Creating a web service using JAX-WS

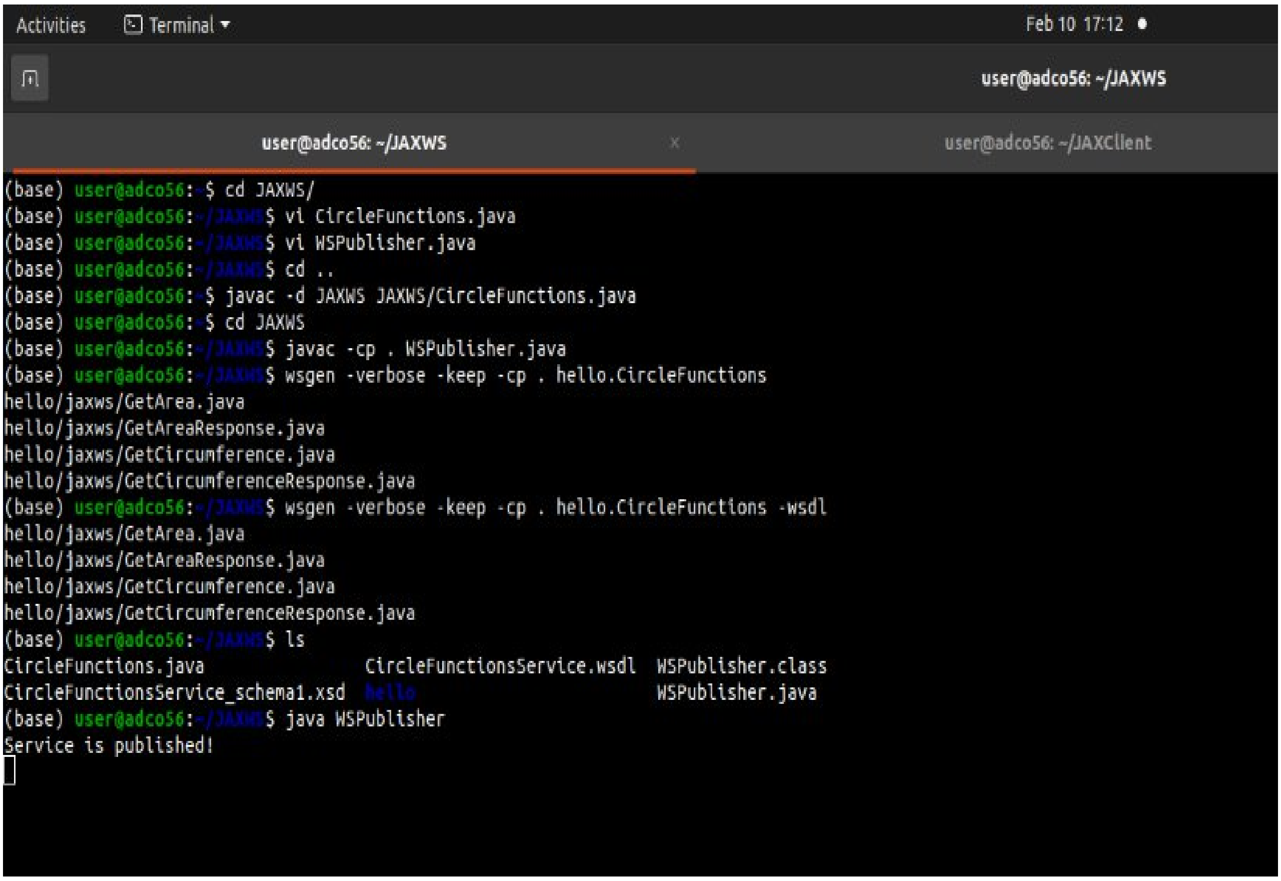
```
//CircleFunctions.java
package hello;
import javax.jws.WebService;
@WebService
public class CircleFunctions {
        public double getArea(double r) {
        return java.lang.Math.PI * (r * r);
                                        }
        public double getCircumference(double r) {
        return 2 * java.lang.Math.PI * r;
                                        }
}
```

```
//WSPublisher.java
import javax.xml.ws.Endpoint;
import hello.CircleFunctions;
public class WSPublisher{
        public static void main(String[] args) {
        Endpoint.publish(
        "http://localhost:8080/WebServiceExample/circlefunctions",
        new CircleFunctions());
        System.out.println("Service is published!");
}
}
```

Directory Structure:

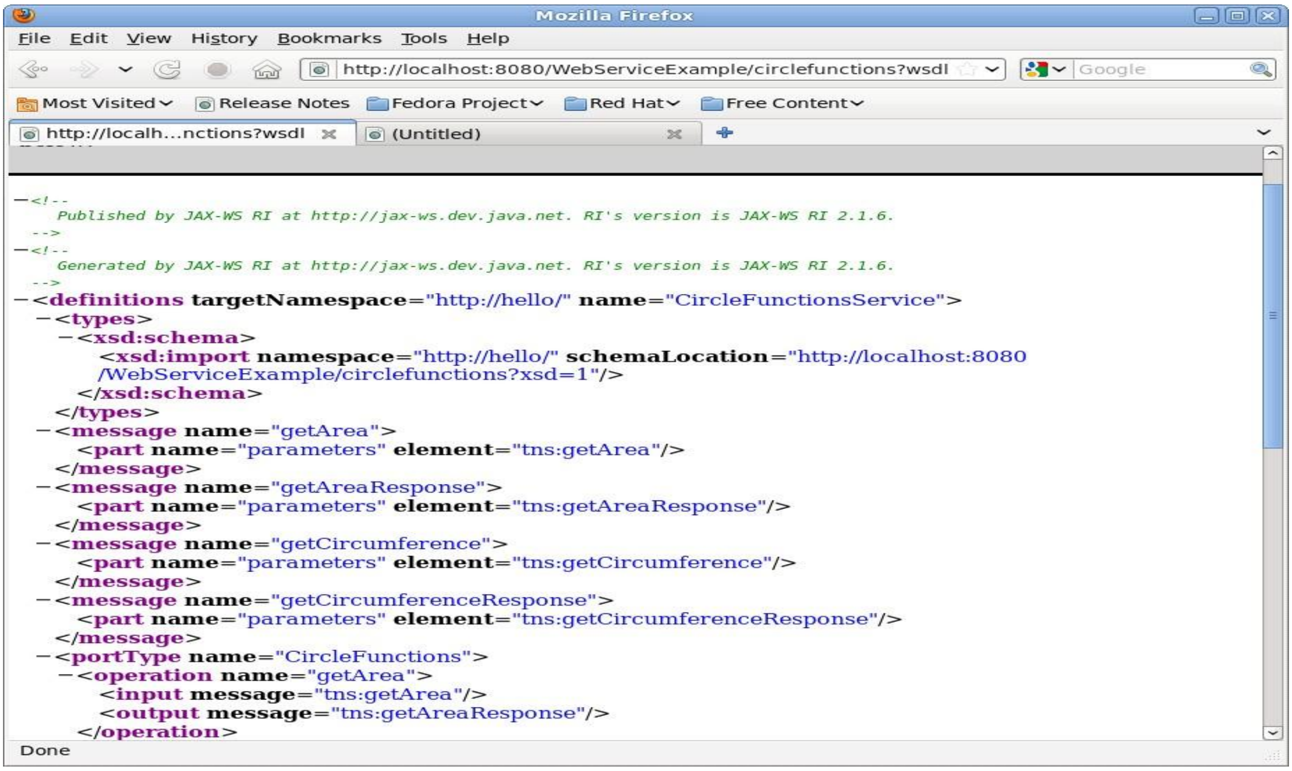
```
C:\softwares\JAXWS
→ hello/
→  CircleFunctions.java
→  WSPublisher.java
```

OUTPUT:

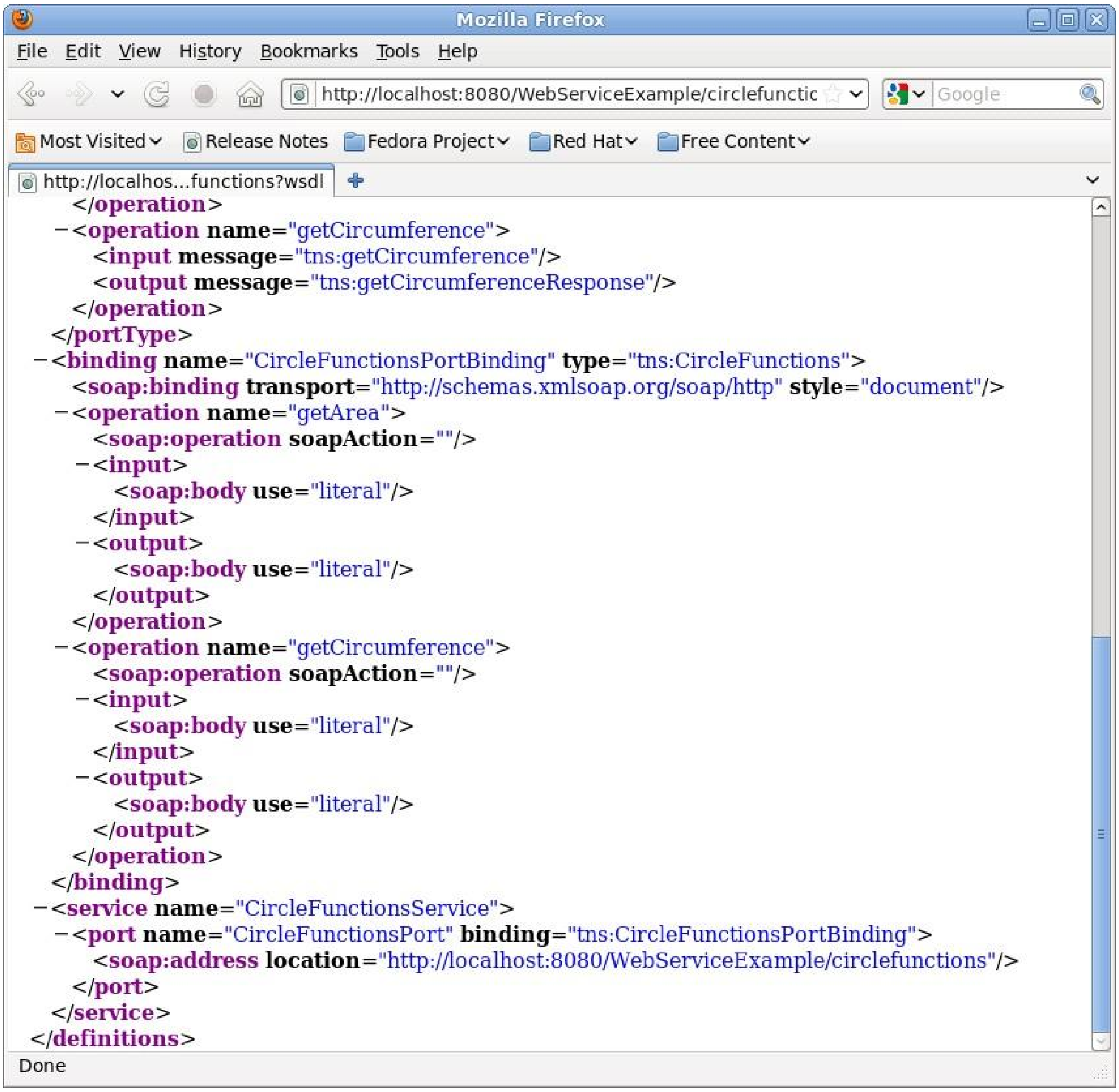


```
(base) user@adco56:~$ cd JAXWS/
(base) user@adco56:~/JAXWS$ vi CircleFunctions.java
(base) user@adco56:~/JAXWS$ vi WSPublisher.java
(base) user@adco56:~/JAXWS$ cd ..
(base) user@adco56:~$ javac -d JAXWS JAXWS/CircleFunctions.java
(base) user@adco56:~$ cd JAXWS
(base) user@adco56:~/JAXWS$ javac -cp . WSPublisher.java
(base) user@adco56:~/JAXWS$ wsgen -verbose -keep -cp . hello.CircleFunctions
hello/jaxws/GetArea.java
hello/jaxws/GetAreaResponse.java
hello/jaxws/GetCircumference.java
hello/jaxws/GetCircumferenceResponse.java
(base) user@adco56:~/JAXWS$ wsgen -verbose -keep -cp . hello.CircleFunctions -wsdl
hello/jaxws/GetArea.java
hello/jaxws/GetAreaResponse.java
hello/jaxws/GetCircumference.java
hello/jaxws/GetCircumferenceResponse.java
(base) user@adco56:~/JAXWS$ ls
CircleFunctions.java              CircleFunctionsService.wsdl  WSPublisher.class
CircleFunctionsService_schema1.xsd  hello                       WSPublisher.java
(base) user@adco56:~/JAXWS$ java WSPublisher
Service is published!
```

http://localhost:8080/WebServiceExample/circlefunctions?wsdl


Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://localhost:8080/WebServiceExample/circlefunctions?wsdl
Google
Most Visited
Release Notes
Fedora Project
Red Hat
Free Content
http://localh...nctions?wsdl
(Untitled)
Done

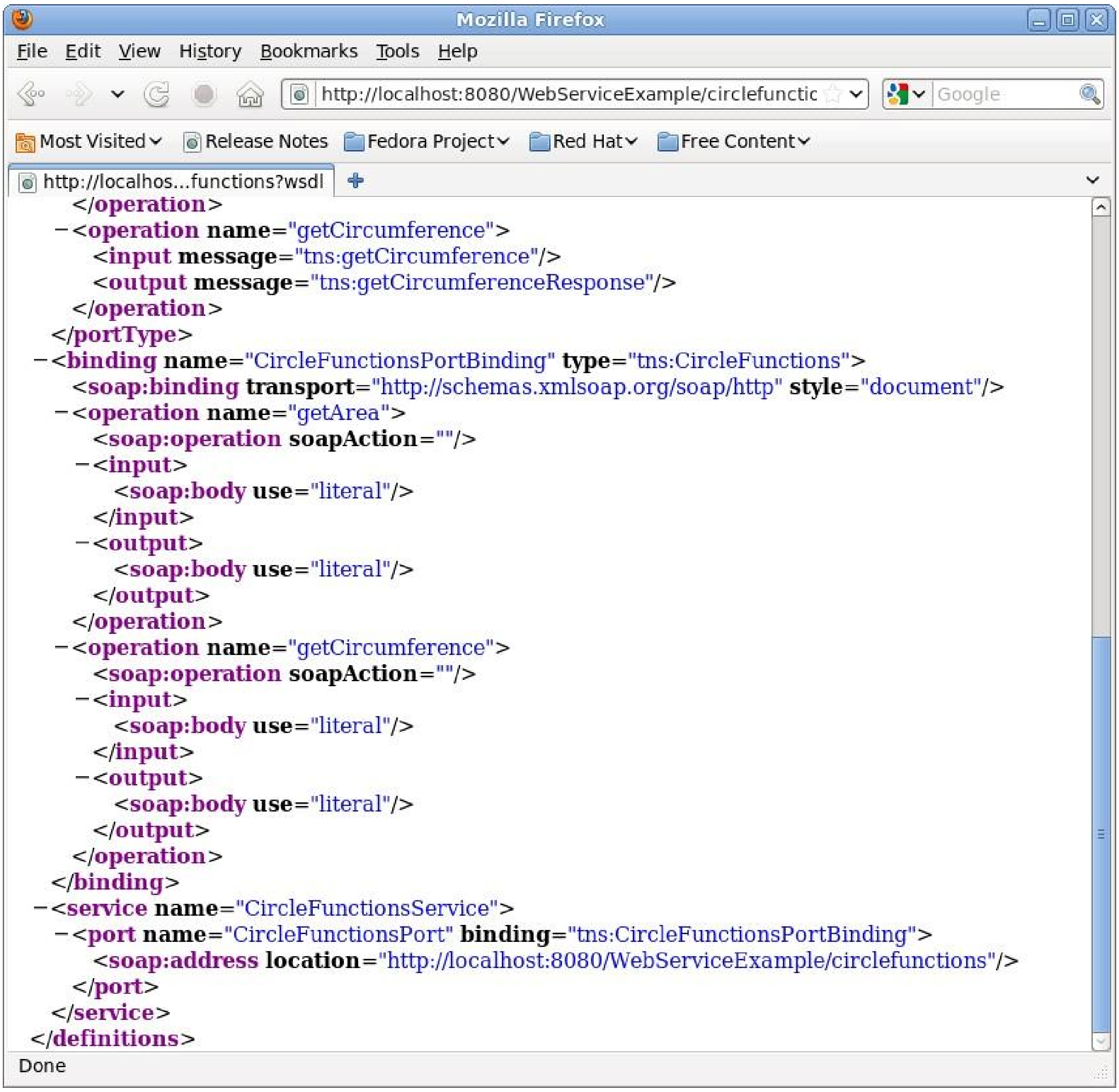
```
-<!--
   Published by JAX-WS RI at http://jax-ws.dev.java.net. RI's version is JAX-WS RI 2.1.6.
 -->
-<!--
   Generated by JAX-WS RI at http://jax-ws.dev.java.net. RI's version is JAX-WS RI 2.1.6.
 -->
-<definitions targetNamespace="http://hello/" name="CircleFunctionsService">
  -<types>
    -<xsd:schema>
       <xsd:import namespace="http://hello/" schemaLocation="http://localhost:8080
       /WebServiceExample/circlefunctions?xsd=1"/>
     </xsd:schema>
   </types>
  -<message name="getArea">
     <part name="parameters" element="tns:getArea"/>
   </message>
  -<message name="getAreaResponse">
     <part name="parameters" element="tns:getAreaResponse"/>
   </message>
  -<message name="getCircumference">
     <part name="parameters" element="tns:getCircumference"/>
   </message>
  -<message name="getCircumferenceResponse">
     <part name="parameters" element="tns:getCircumferenceResponse"/>
   </message>
  -<portType name="CircleFunctions">
    -<operation name="getArea">
       <input message="tns:getArea"/>
       <output message="tns:getAreaResponse"/>
     </operation>
```



Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://localhost:8080/WebServiceExample/circlefunctic
Google
Most Visited
Release Notes
Fedora Project
Red Hat
Free Content
http://localhos...functions?wsdl

```
      </operation>
    -<operation name="getCircumference">
       <input message="tns:getCircumference"/>
       <output message="tns:getCircumferenceResponse"/>
     </operation>
   </portType>
  -<binding name="CircleFunctionsPortBinding" type="tns:CircleFunctions">
     <soap:binding transport="http://schemas.xmlsoap.org/soap/http" style="document"/>
    -<operation name="getArea">
       <soap:operation soapAction=""/>
      -<input>
         <soap:body use="literal"/>
       </input>
      -<output>
         <soap:body use="literal"/>
       </output>
     </operation>
    -<operation name="getCircumference">
       <soap:operation soapAction=""/>
      -<input>
         <soap:body use="literal"/>
       </input>
      -<output>
         <soap:body use="literal"/>
       </output>
     </operation>
   </binding>
  -<service name="CircleFunctionsService">
    -<port name="CircleFunctionsPort" binding="tns:CircleFunctionsPortBinding">
       <soap:address location="http://localhost:8080/WebServiceExample/circlefunctions"/>
     </port>
   </service>
 </definitions>
```

Done


Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://localhost:8080/WebServiceExample/circlefunctic
Google
Most Visited
Release Notes
Fedora Project
Red Hat
Free Content
http://localhos...functions?wsdl
Done

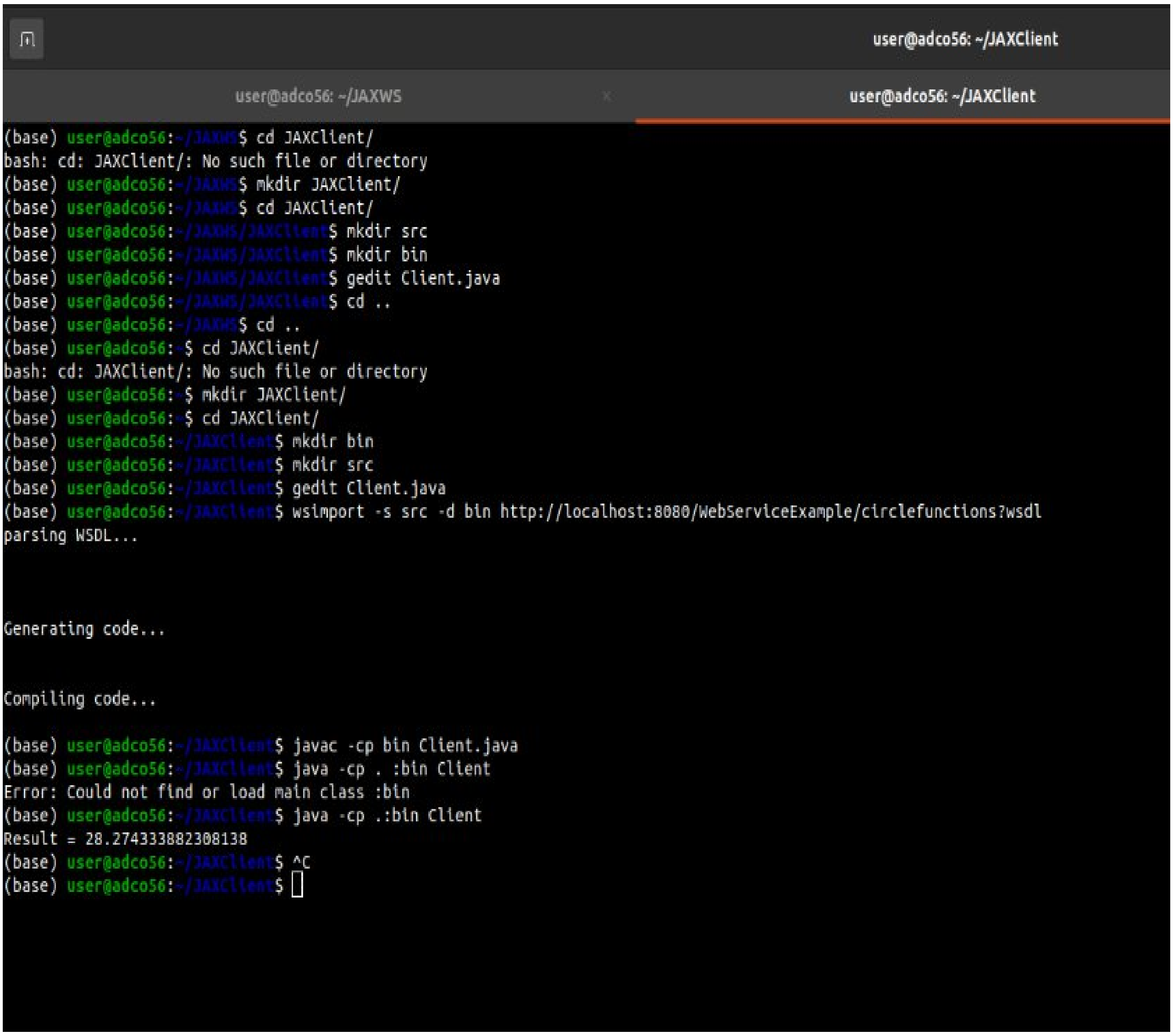
```
      </operation>
    -<operation name="getCircumference">
       <input message="tns:getCircumference"/>
       <output message="tns:getCircumferenceResponse"/>
     </operation>
   </portType>
  -<binding name="CircleFunctionsPortBinding" type="tns:CircleFunctions">
     <soap:binding transport="http://schemas.xmlsoap.org/soap/http" style="document"/>
    -<operation name="getArea">
       <soap:operation soapAction=""/>
      -<input>
         <soap:body use="literal"/>
       </input>
      -<output>
         <soap:body use="literal"/>
       </output>
     </operation>
    -<operation name="getCircumference">
       <soap:operation soapAction=""/>
      -<input>
         <soap:body use="literal"/>
       </input>
      -<output>
         <soap:body use="literal"/>
       </output>
     </operation>
   </binding>
  -<service name="CircleFunctionsService">
    -<port name="CircleFunctionsPort" binding="tns:CircleFunctionsPortBinding">
       <soap:address location="http://localhost:8080/WebServiceExample/circlefunctions"/>
     </port>
   </service>
 </definitions>
```

Client.java

```
import hello.CircleFunctions;
import hello.CircleFunctionsService;
public class Client {
public static void main(String[] args)
{
try {
CircleFunctionsService service=new CircleFunctionsService();
CircleFunctions port = service.getCircleFunctionsPort();
double arg0 = 3.0;

double result = port.getArea(arg0);
System.out.println("Result = "+result);
} catch (Exception ex) {
}
}
```

Directory Structure:
c:\Softwares\JAXClient
→ bin/
→ src/
→ Client.java

OUTPUT:



```
(base) user@adco56:~/JAXWS$ cd JAXClient/
bash: cd: JAXClient/: No such file or directory
(base) user@adco56:~/JAXWS$ mkdir JAXClient/
(base) user@adco56:~/JAXWS$ cd JAXClient/
(base) user@adco56:~/JAXWS/JAXClient$ mkdir src
(base) user@adco56:~/JAXWS/JAXClient$ mkdir bin
(base) user@adco56:~/JAXWS/JAXClient$ gedit Client.java
(base) user@adco56:~/JAXWS/JAXClient$ cd ..
(base) user@adco56:~/JAXWS$ cd ..
(base) user@adco56:~$ cd JAXClient/
bash: cd: JAXClient/: No such file or directory
(base) user@adco56:~$ mkdir JAXClient/
(base) user@adco56:~$ cd JAXClient/
(base) user@adco56:~/JAXClient$ mkdir bin
(base) user@adco56:~/JAXClient$ mkdir src
(base) user@adco56:~/JAXClient$ gedit Client.java
(base) user@adco56:~/JAXClient$ wsimport -s src -d bin http://localhost:8080/WebServiceExample/circlefunctions?wsdl
parsing WSDL...



Generating code...


Compiling code...

(base) user@adco56:~/JAXClient$ javac -cp bin Client.java
(base) user@adco56:~/JAXClient$ java -cp . :bin Client
Error: Could not find or load main class :bin
(base) user@adco56:~/JAXClient$ java -cp .:bin Client
Result = 28.274333882308138
(base) user@adco56:~/JAXClient$ ^C
(base) user@adco56:~/JAXClient$ 
```

RESULT:

Hence the given program for Creating a Web Service using JAX-WS and Invoking it from a java client is compiled successfully and corresponding output is obtained.

EXP NO.: 2
DATE:

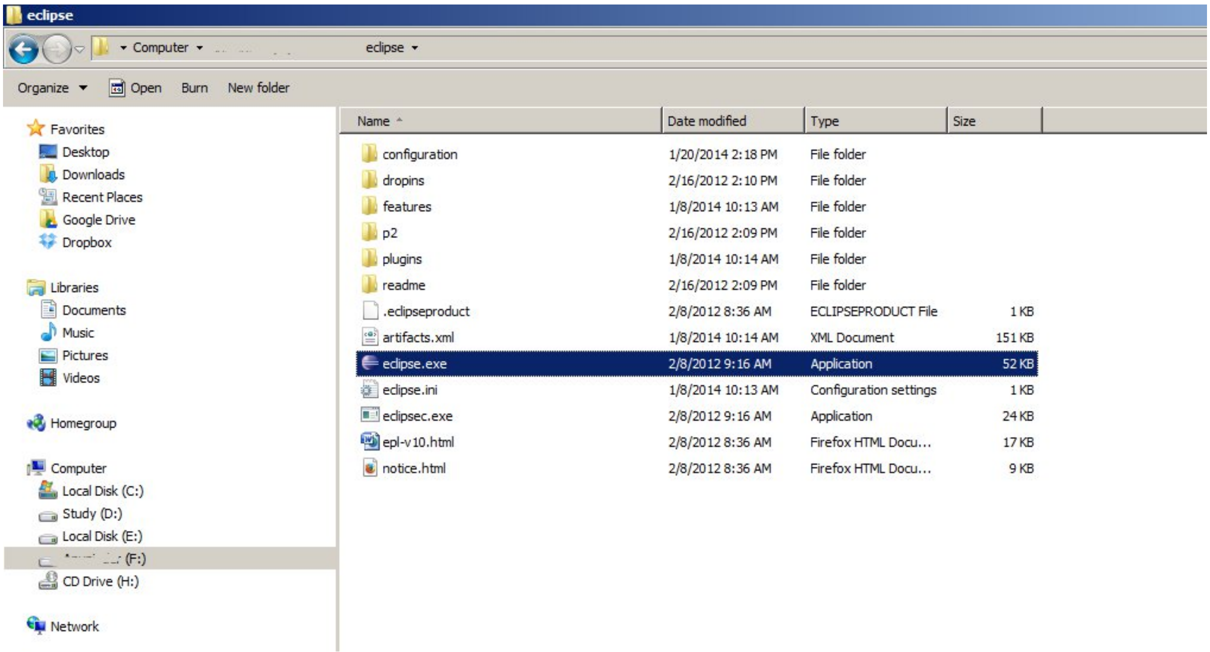
SETUP A CLOUD ENVIRONMENT AND EXECUTE A SCHEDULING ALGORITHM USING CLOUDSIM

AIM:

To set up a cloud environment and execute a scheduling algorithm using CloudSim.

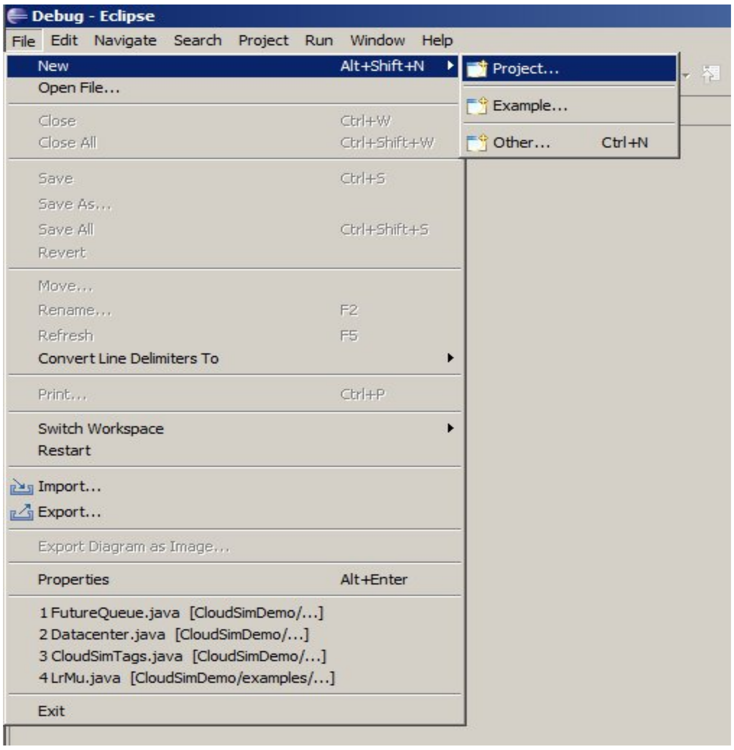
PROCEDURE:

1. Navigate to the folder where you have unzipped the Eclipse folder and open Eclipse.exe

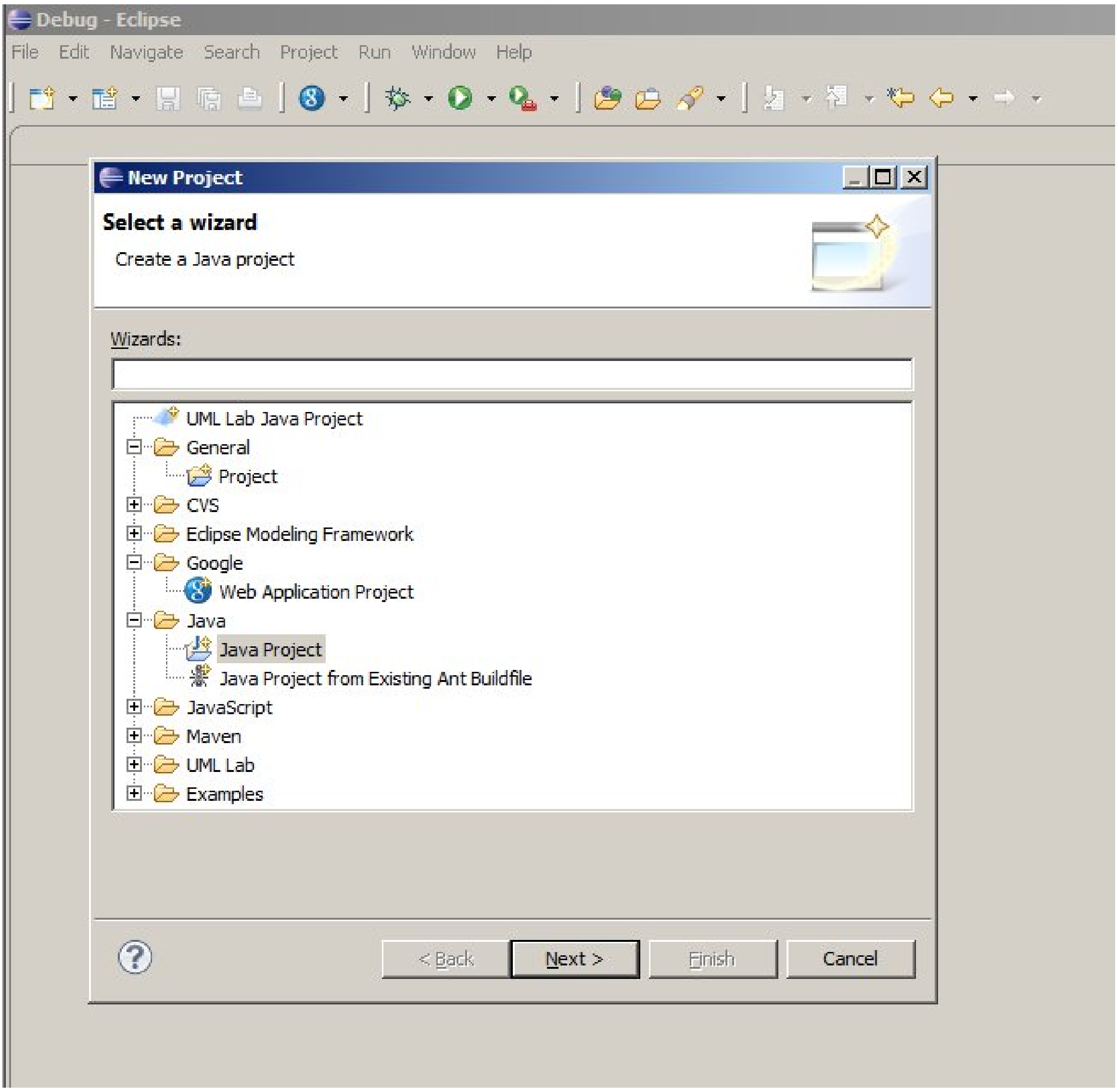


Name	Date modified	Type	Size
configuration	1/20/2014 2:18 PM	File folder	
dropins	2/16/2012 2:10 PM	File folder	
features	1/8/2014 10:13 AM	File folder	
p2	2/16/2012 2:09 PM	File folder	
plugins	1/8/2014 10:14 AM	File folder	
readme	2/16/2012 2:09 PM	File folder	
.eclipseproduct	2/8/2012 8:36 AM	ECLIPSEPRODUCT File	1 KB
artifacts.xml	1/8/2014 10:14 AM	XML Document	151 KB
eclipse.exe	2/8/2012 9:16 AM	Application	52 KB
eclipse.ini	1/8/2014 10:13 AM	Configuration settings	1 KB
eclipsec.exe	2/8/2012 9:16 AM	Application	24 KB
epl-v10.html	2/8/2012 8:36 AM	Firefox HTML Docu...	17 KB
notice.html	2/8/2012 8:36 AM	Firefox HTML Docu...	9 KB

2. Now within Eclipse window navigate the menu: ***File -> New -> Project,*** to open the new project wizard



3. A ‘***New Project***‘ wizard should open. There are a number of options displayed and you have to find & select the ‘***Java Project***‘ option, once done Click ***‘Next***



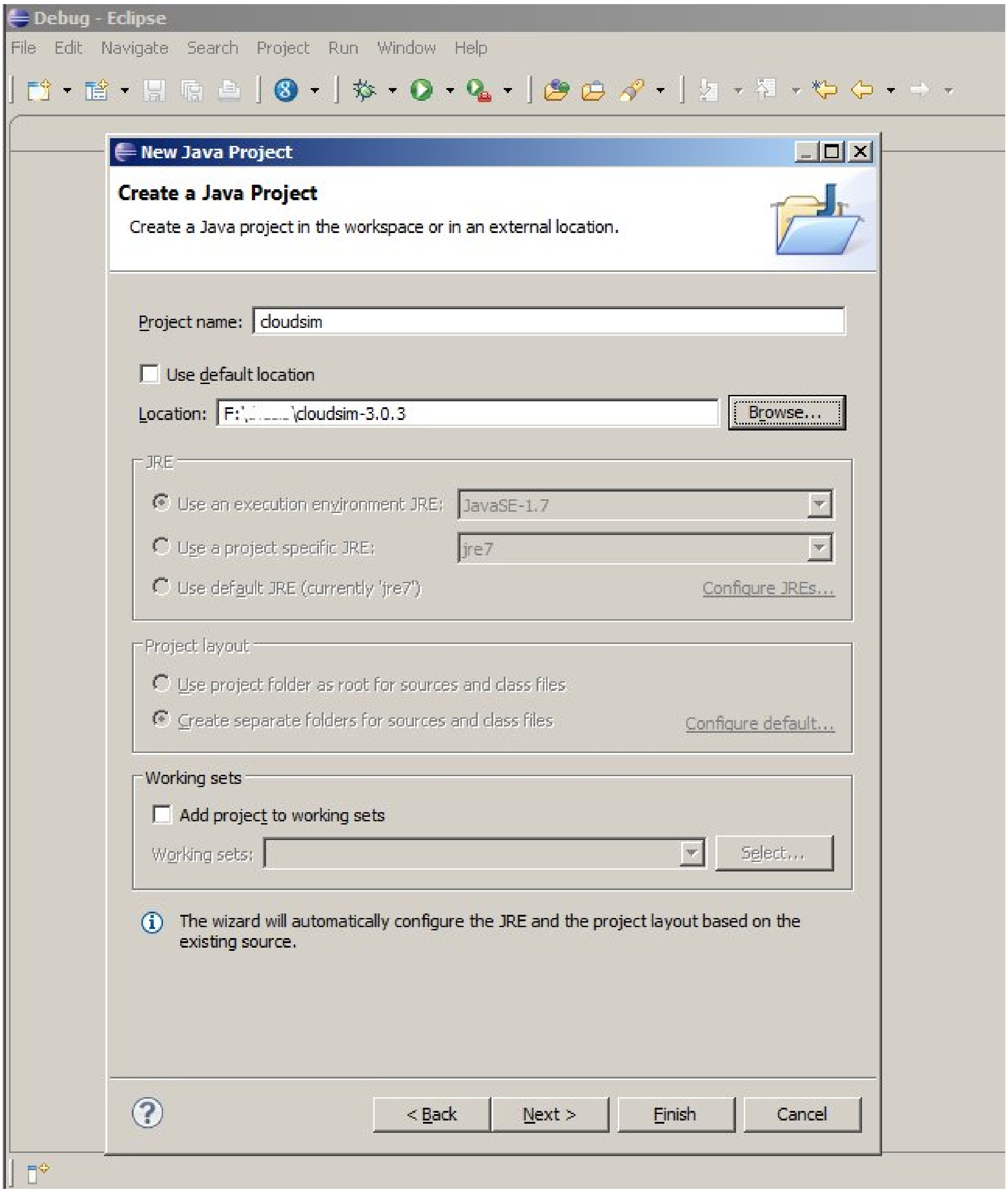
Select a wizard

Create a Java project

‘

4. Now a detailed new project window will open, here you will provide the project name and the path of the CloudSim project source code, which will be done as follows:

- Project Name: CloudSim.
- Unselect the ***‘Use default location’*** option and then click on ***‘Browse’*** to open the path where you have unzipped the Cloudsim project and finally click Next to set project settings
- Make sure you navigate the path till you can see the bin, docs, examples, etc folder in the navigation plane.
- click ‘Next’ to go to the next step i.e. setting up of project settings



Create a Java Project

Create a Java project in the workspace or in an external location.

Project name: cloudsim

☐ Use default location

Location: F:\[redacted]\cloudsim-3.0.3 Browse...

JRE

● Use an execution environment JRE: JavaSE-1.7

○ Use a project specific JRE: jre7

○ Use default JRE (currently 'jre7') Configure JREs...

Project layout

○ Use project folder as root for sources and class files

● Create separate folders for sources and class files Configure default...

Working sets

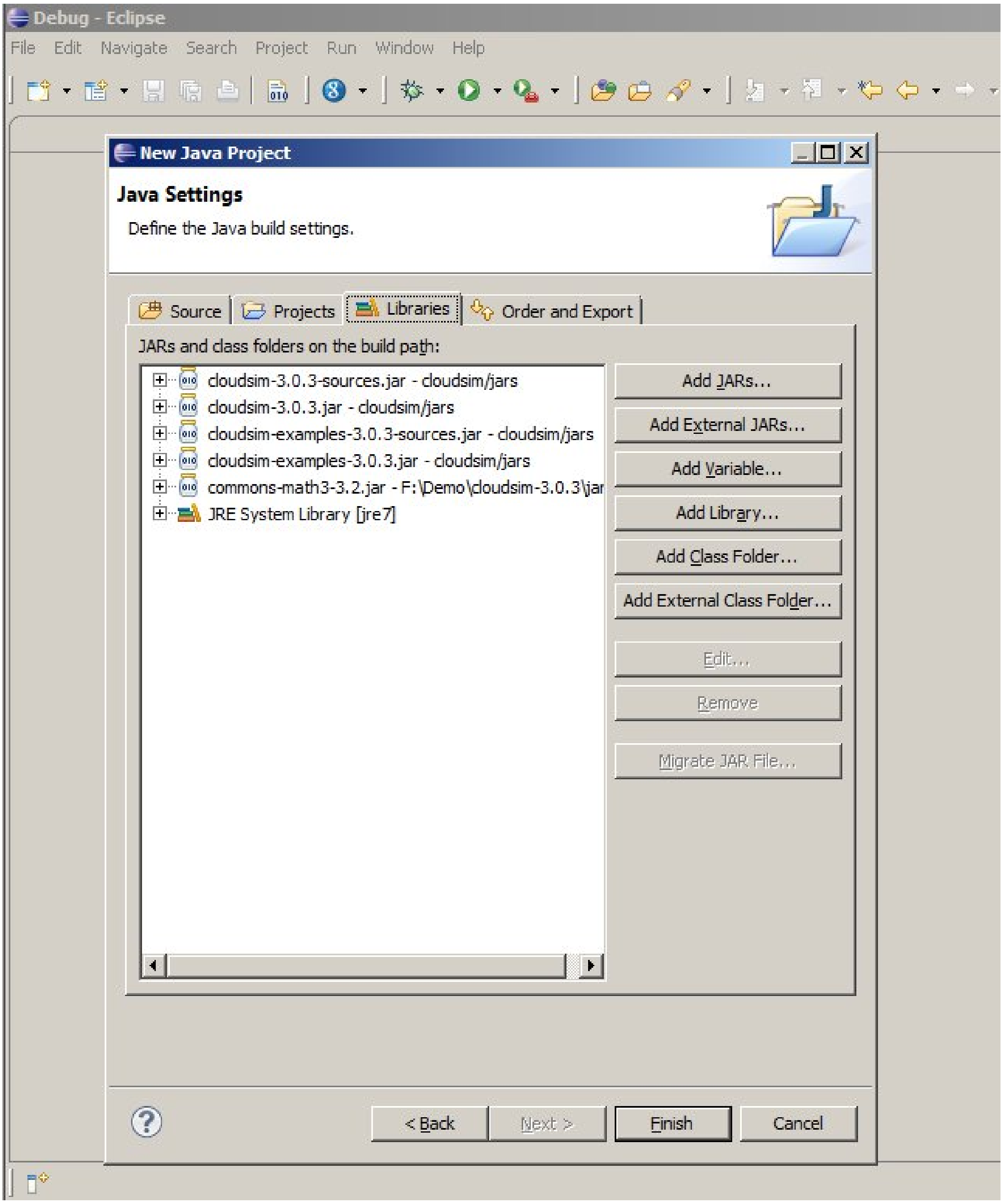
☐ Add project to working sets

Working sets: Select...

The wizard will automatically configure the JRE and the project layout based on the existing source.

5. Steps for Project Setup :

- Now open the ***'Libraries'*** tab and if you do not find commons-math3-3.x.jar in the list then simply click on ***'Add External Jar'.***
- Once you have clicked on '***Add External JAR's***' Open the path where you have unzipped the commons-math binaries and select '***Commons-math3-3.x.jar***' and click on Open.
- Ensure the external jar that you opened in the previous step is displayed in the list and then click on '***Finish***' (your system may take 2-3 minutes to configure the project).



Java Settings

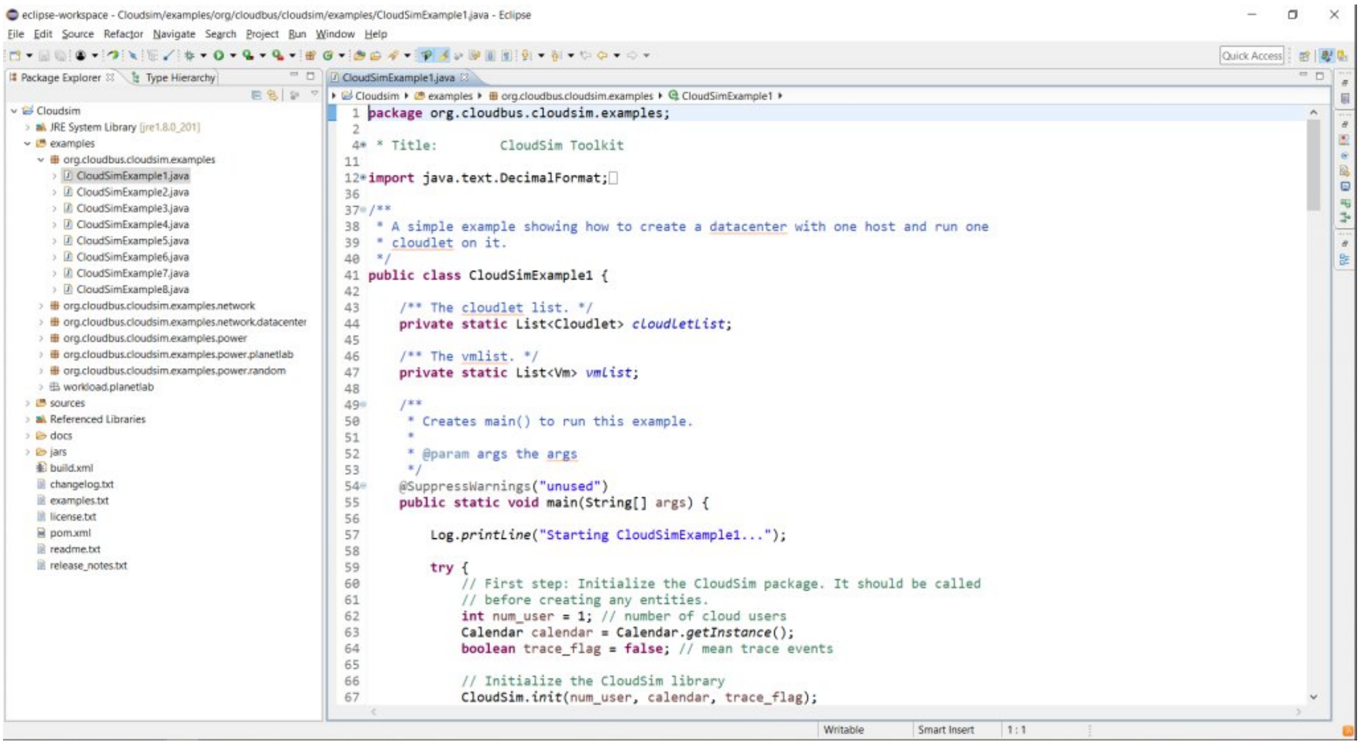
Define the Java build settings.

JARs and class folders on the build path:

- cloudsim-3.0.3-sources.jar - cloudsim/jars
- cloudsim-3.0.3.jar - cloudsim/jars
- cloudsim-examples-3.0.3-sources.jar - cloudsim/jars
- cloudsim-examples-3.0.3.jar - cloudsim/jars
- commons-math3-3.2.jar - F:\Demo\cloudsim-3.0.3\jar
- JRE System Library [jre7]

6. Once the project is configured you can open ‘Project Explorer’ and start exploring the Cloudsim project. Also for the first time eclipse automatically start building the workspace for the newly configured Cloudsim project, which may take some time depending on the configuration of the computer system.
Following is the final screen which you will see after Cloudsim is configured.

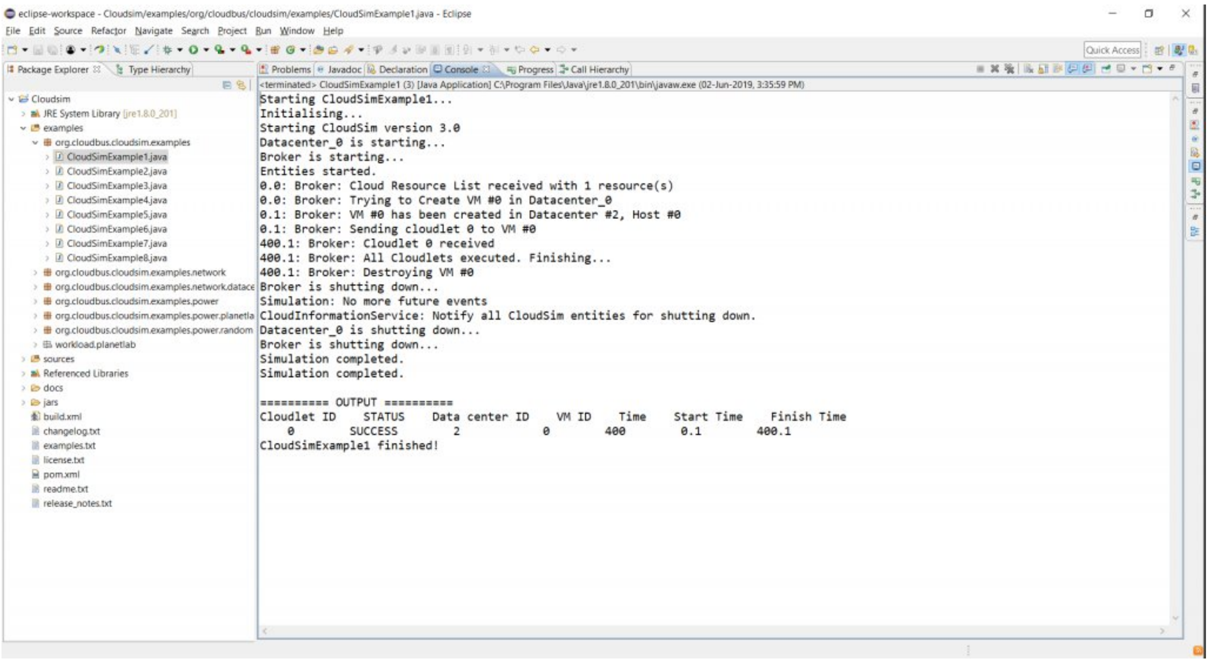
- Now just to check you within the ‘**Project Explorer**‘, you should navigate to the ‘**examples**‘ folder, then expand the package ‘***org.cloudbus.cloudsim.examples***‘ and double click to open the ‘***CloudsimExample1.java***‘.



eclipse-workspace - Cloudsim/examples/org/cloudbus/cloudsim/examples/CloudSimExample1.java - Eclipse
File Edit Source Refactor Navigate Search Project Run Window Help
Package Explorer
Type Hierarchy
CloudSimExample1.java
Cloudsim ▸ examples ▸ org.cloudbus.cloudsim.examples ▸ CloudSimExample1 ▸
Writable
Smart Insert
1 : 1

```java
package org.cloudbus.cloudsim.examples;

 * Title:        CloudSim Toolkit

import java.text.DecimalFormat;

/**
 * A simple example showing how to create a datacenter with one host and run one
 * cloudlet on it.
 */
public class CloudSimExample1 {

    /** The cloudlet list. */
    private static List<Cloudlet> cloudletList;

    /** The vmlist. */
    private static List<Vm> vmlist;

    /**
     * Creates main() to run this example.
     *
     * @param args the args
     */
    @SuppressWarnings("unused")
    public static void main(String[] args) {

        Log.printLine("Starting CloudSimExample1...");

        try {
            // First step: Initialize the CloudSim package. It should be called
            // before creating any entities.
            int num_user = 1; // number of cloud users
            Calendar calendar = Calendar.getInstance();
            boolean trace_flag = false; // mean trace events

            // Initialize the CloudSim library
            CloudSim.init(num_user, calendar, trace_flag);
```

7. Now navigate to the Eclipse menu ‘***Run -> Run***‘ or directly use a keyboard shortcut ***‘Ctrl + F11’*** to execute the ‘***CloudsimExample1.java***‘ and it is executed and output is displayed in the console window of the Eclipse IDE.



eclipse-workspace - Cloudsim/examples/org/cloudbus/cloudsim/examples/CloudSimExample1.java - Eclipse
File Edit Source Refactor Navigate Search Project Run Window Help
Package Explorer
Type Hierarchy
Problems
Javadoc
Declaration
Console
Progress
Call Hierarchy

<terminated> CloudSimExample1 (3) [Java Application] C:\Program Files\Java\jre1.8.0_201\bin\javaw.exe (02-Jun-2019, 3:35:59 PM)

```
Starting CloudSimExample1...
Initialising...
Starting CloudSim version 3.0
Datacenter_0 is starting...
Broker is starting...
Entities started.
0.0: Broker: Cloud Resource List received with 1 resource(s)
0.0: Broker: Trying to Create VM #0 in Datacenter_0
0.1: Broker: VM #0 has been created in Datacenter #2, Host #0
0.1: Broker: Sending cloudlet 0 to VM #0
400.1: Broker: Cloudlet 0 received
400.1: Broker: All Cloudlets executed. Finishing...
400.1: Broker: Destroying VM #0
Broker is shutting down...
Simulation: No more future events
CloudInformationService: Notify all CloudSim entities for shutting down.
Datacenter_0 is shutting down...
Broker is shutting down...
Simulation completed.
Simulation completed.

========== OUTPUT ==========
Cloudlet ID    STATUS    Data center ID    VM ID    Time    Start Time    Finish Time
    0        SUCCESS        2            0        400       0.1        400.1
CloudSimExample1 finished!
```

RESULT:

Hence the given program for developing a scheduling algorithm is compiled successfully complied and corresponding output is obtained.

EXP NO.: 3
DATE:

CREATE VIRTUAL MACHINE WITH DIFFERENT CONFIGURATION IN OPENSTACK CLOUD

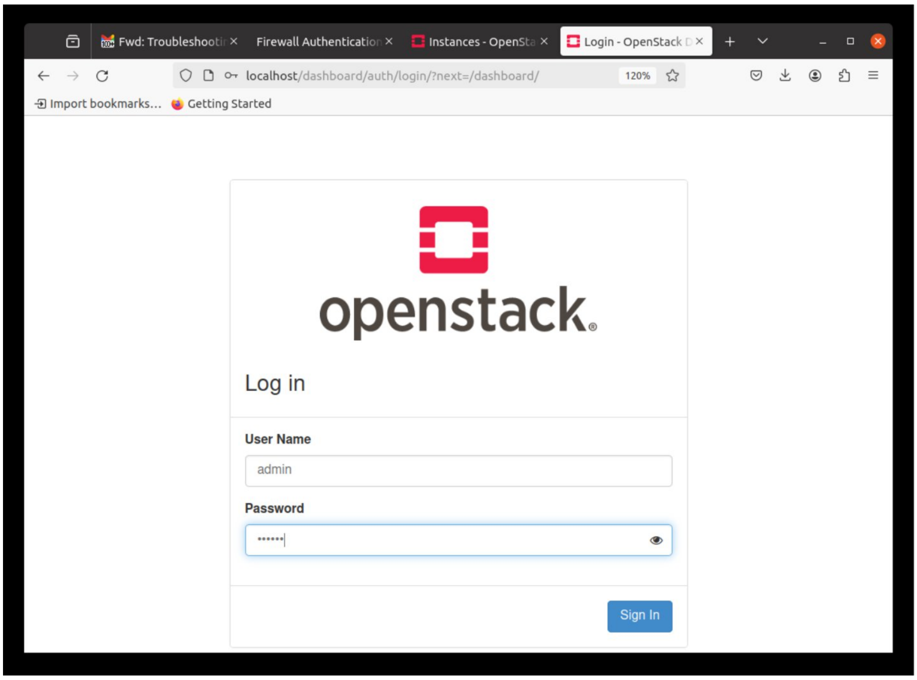
AIM:

To create a virtual machine with different configuration in openstack cloud.

PROCEDURE:

1. Login:

Access the OpenStack platform using the provided credentials: username as "admin" and password as "secret".



Log in

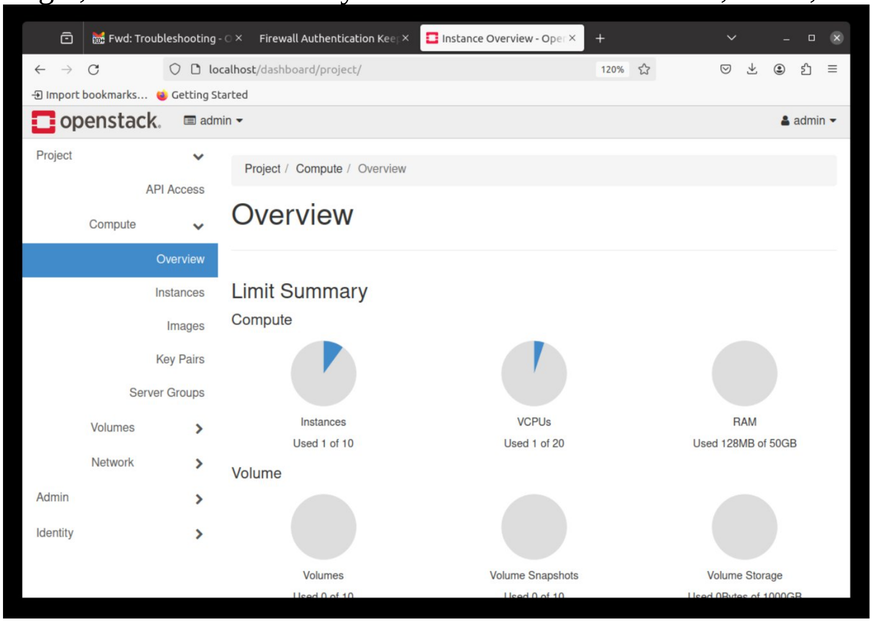
User Name: admin

Password: ••••••

Sign In

2. Resource Status:

Upon successful login, review the summary to assess the status of CPUs, RAM, volume spaces, etc.



Overview

Limit Summary

Compute

Instances: Used 1 of 10

VCPUs: Used 1 of 20

RAM: Used 128MB of 50GB

Volume

Volumes: Used 0 of 10

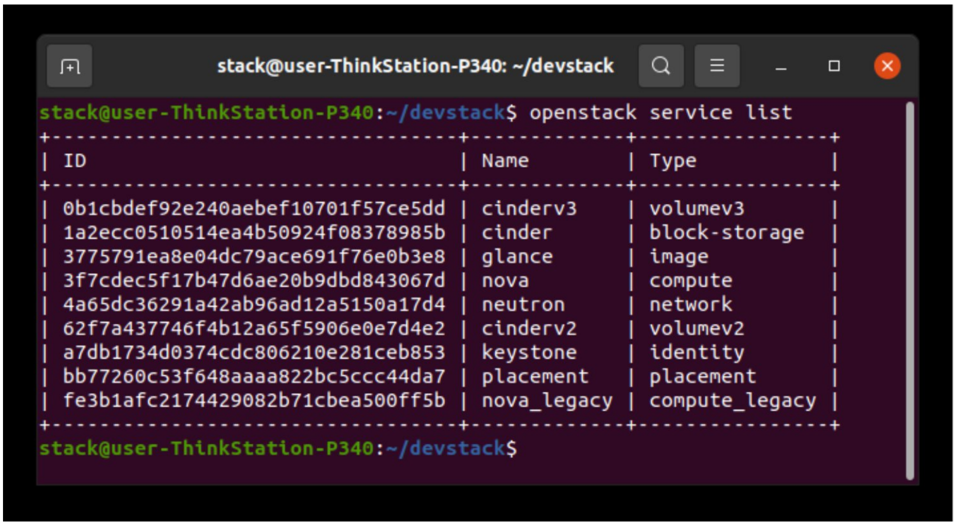
Volume Snapshots: Used 0 of 10

Volume Storage: Used 0Bytes of 1000GB

3. OpenStack Services:

Access the terminal and execute the appropriate command to display the available services.

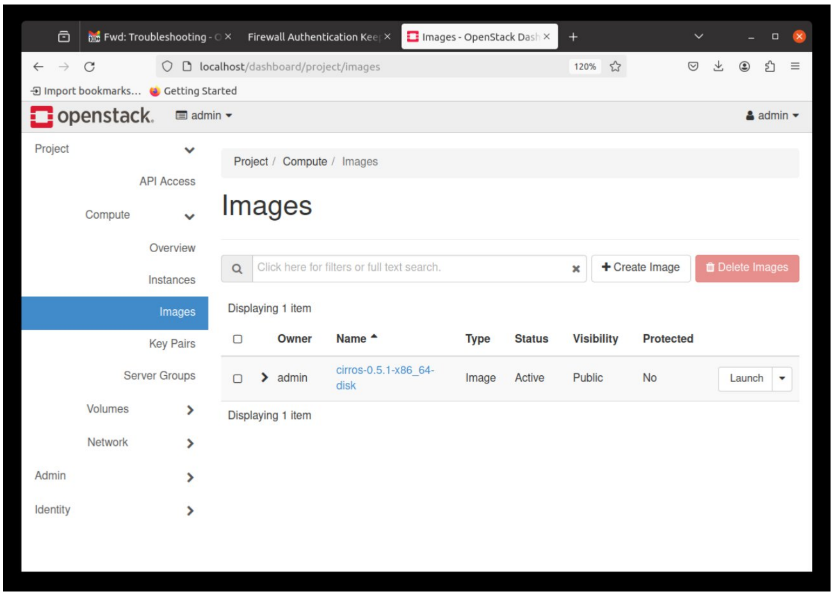
//Download open stack RC file from dashboard (Right top corner)
$sudo cp admin-openrc.sh /opt/stack/devstack
user$sudo su - stack
stack$cd devstack
stack$source admin-openrc.sh
stack$openstack service list



```
stack@user-ThinkStation-P340:~/devstack$ openstack service list
+----------------------------------+-------------+----------------+
| ID                               | Name        | Type           |
+----------------------------------+-------------+----------------+
| 0b1cbdef92e240aebef10701f57ce5dd | cinderv3    | volumev3       |
| 1a2ecc0510514ea4b50924f08378985b | cinder      | block-storage  |
| 3775791ea8e04dc79ace691f76e0b3e8 | glance      | image          |
| 3f7cdec5f17b47d6ae20b9dbd843067d | nova        | compute        |
| 4a65dc36291a42ab96ad12a5150a17d4 | neutron     | network        |
| 62f7a437746f4b12a65f5906e0e7d4e2 | cinderv2    | volumev2       |
| a7db1734d0374cdc806210e281ceb853 | keystone    | identity       |
| bb77260c53f648aaaa822bc5ccc44da7 | placement   | placement      |
| fe3b1afc2174429082b71cbea500ff5b | nova_legacy | compute_legacy |
+----------------------------------+-------------+----------------+
stack@user-ThinkStation-P340:~/devstack$
```

4. View Available Images:

Review the existing images within OpenStack, noting the option to add additional images if required.



Images

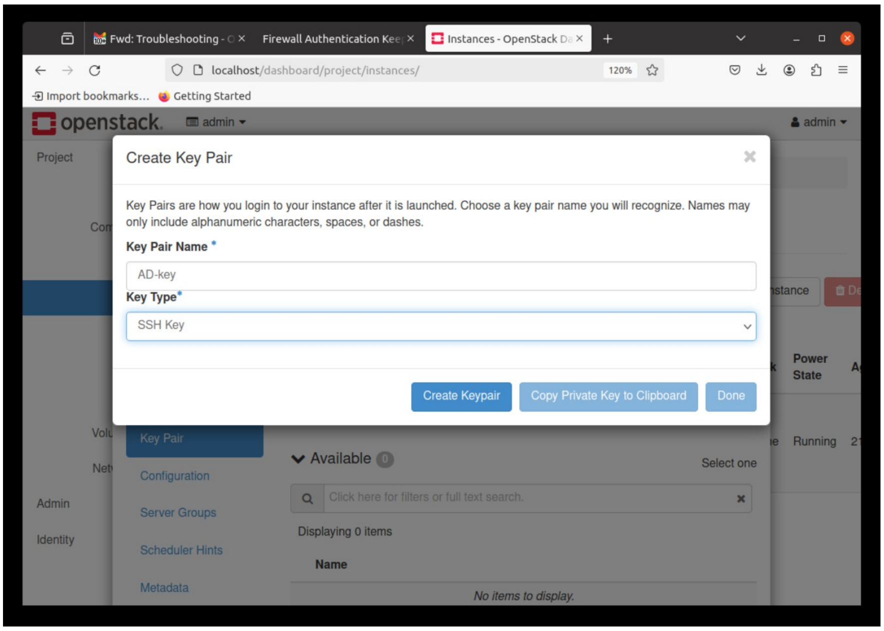
Displaying 1 item

□	Owner	Name ▲	Type	Status	Visibility	Protected
□	admin	cirros-0.5.1-x86_64-disk	Image	Active	Public	No

Displaying 1 item

5. Create a Key Pair:

Create a new key pair named "AD_Key" and select the SSH key type.



Create Key Pair

Key Pairs are how you login to your instance after it is launched. Choose a key pair name you will recognize. Names may only include alphanumeric characters, spaces, or dashes.

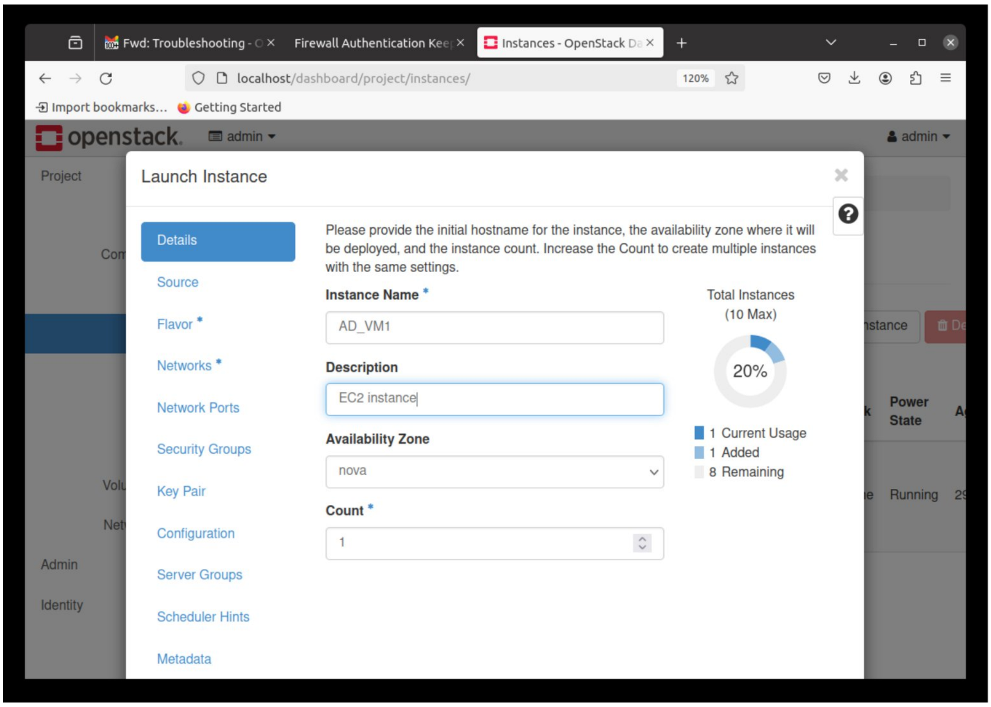
Key Pair Name *: AD-key

Key Type *: SSH Key

6. Launch Instance:

- Provide the instance name as "AD_VM1".
- Enter the description as "EC2 instance".
- Set the availability zone to "Nova".
- Define the count as "1".

then, proceed to the next step.



Launch Instance

Please provide the initial hostname for the instance, the availability zone where it will be deployed, and the instance count. Increase the Count to create multiple instances with the same settings.

Instance Name *: AD_VM1

Description: EC2 instance

Availability Zone: nova

Count *: 1

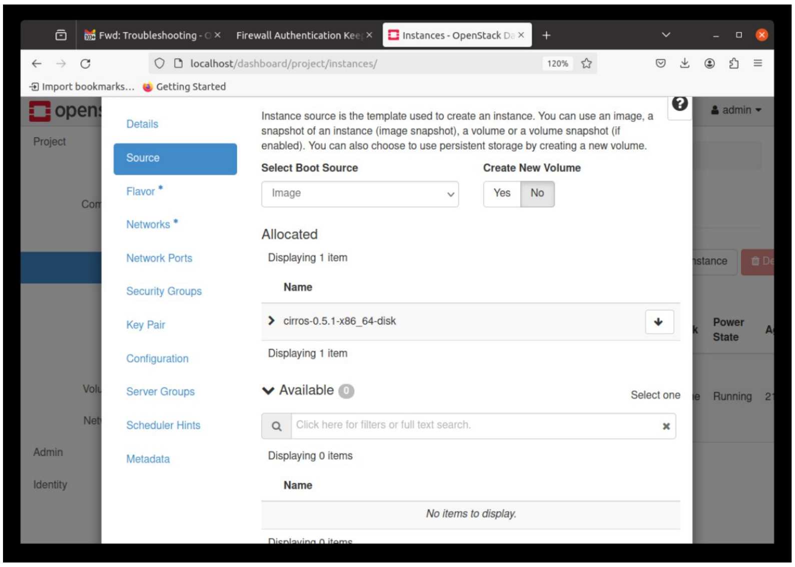
Total Instances (10 Max)

20%

1 Current Usage
1 Added
8 Remaining

7. Add an Image:

- Select the boot source as an image.
- Set "create new volume" option to "NO".
- Choose CirrOS from the available list.



Instance source is the template used to create an instance. You can use an image, a snapshot of an instance (image snapshot), a volume or a volume snapshot (if enabled). You can also choose to use persistent storage by creating a new volume.

Select Boot Source: Image

Create New Volume: Yes [✓ No]

Allocated

Displaying 1 item

Name
cirros-0.5.1-x86_64-disk

Displaying 1 item

Available 0 — Select one

Click here for filters or full text search.

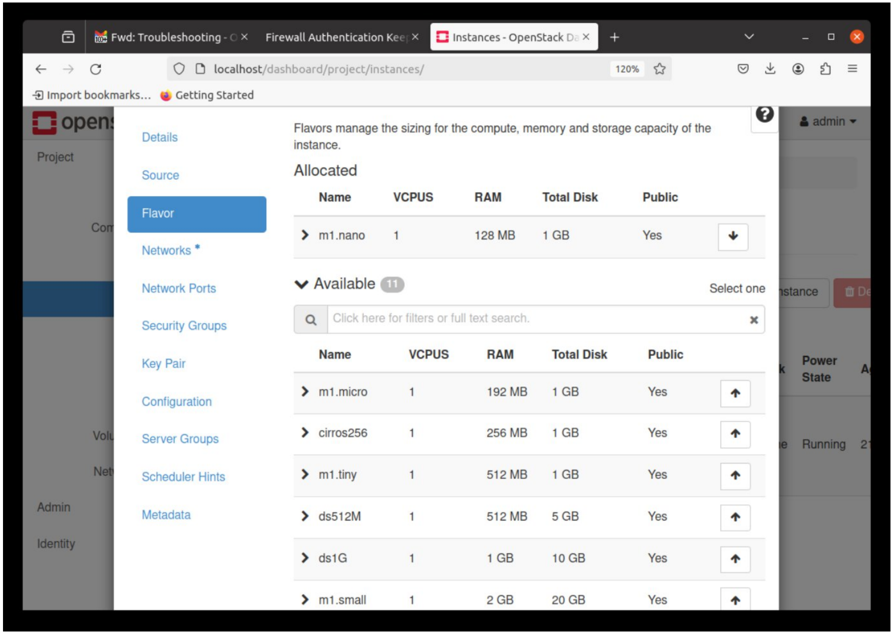
Displaying 0 items

Name
No items to display.

Displaying 0 items

8. Add Flavor:

- Select the "m1.Nano" flavor from the list of available flavors.
- Allocate resources for "m1.Nano" and proceed to the next step.



Flavors manage the sizing for the compute, memory and storage capacity of the instance.

Allocated

Name	VCPUS	RAM	Total Disk	Public
m1.nano	1	128 MB	1 GB	Yes

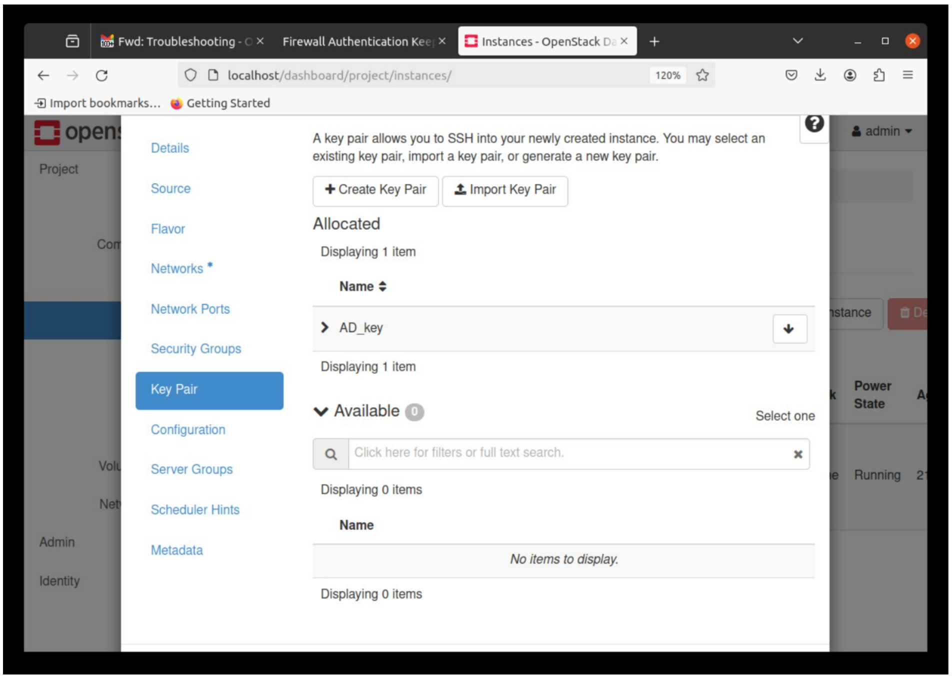
Available 11 — Select one

Click here for filters or full text search.

Name	VCPUS	RAM	Total Disk	Public
m1.micro	1	192 MB	1 GB	Yes
cirros256	1	256 MB	1 GB	Yes
m1.tiny	1	512 MB	1 GB	Yes
ds512M	1	512 MB	5 GB	Yes
ds1G	1	1 GB	10 GB	Yes
m1.small	1	2 GB	20 GB	Yes

9. Add Key Pair:

- Choose the previously created key pair from the list of available key pairs.
- Allocate the selected key pair and proceed.



A key pair allows you to SSH into your newly created instance. You may select an existing key pair, import a key pair, or generate a new key pair.

+ Create Key Pair | Import Key Pair

Allocated

Displaying 1 item

Name
AD_key

Displaying 1 item

Available 0 — Select one

Click here for filters or full text search.

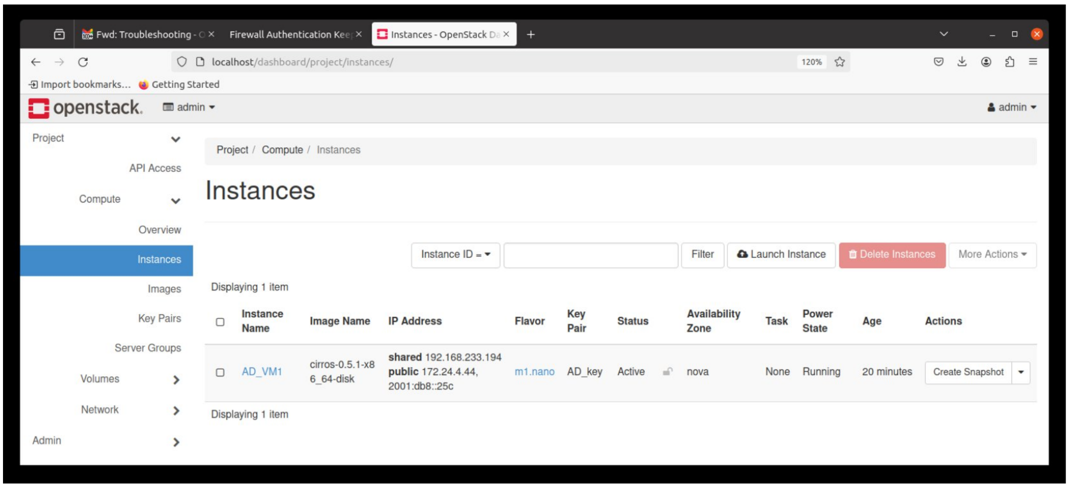
Displaying 0 items

Name
No items to display.

Displaying 0 items

10. Launch Instance:

- Follow the specified steps to complete the instance launch process.
- Now, the instance status should be "active".



Instances

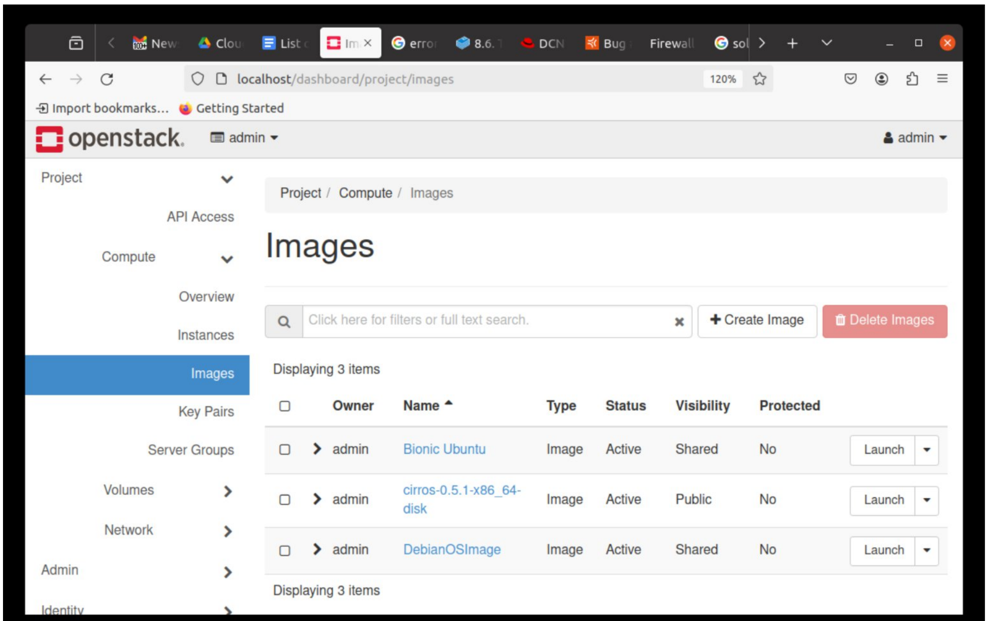
Displaying 1 item

	Instance Name	Image Name	IP Address	Flavor	Key Pair	Status		Availability Zone	Task	Power State	Age	Actions
□	AD_VM1	cirros-0.5.1-x86_64-disk	**shared** 192.168.233.194 **public** 172.24.4.44, 2001:db8::25c	m1.nano	AD_key	Active		nova	None	Running	20 minutes	Create Snapshot

Displaying 1 item

11. Adding more images:

Now repeat the same process for adding multiple images .

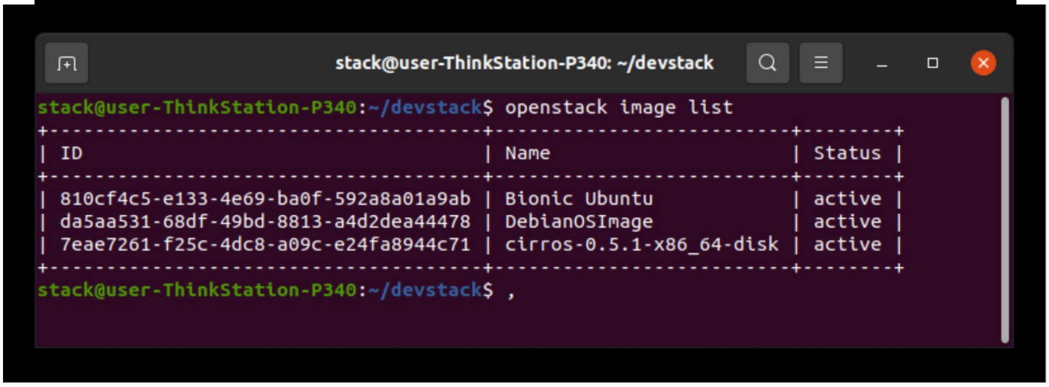


Images

Displaying 3 items

Owner	Name	Type	Status	Visibility	Protected	
admin	Bionic Ubuntu	Image	Active	Shared	No	Launch
admin	cirros-0.5.1-x86_64-disk	Image	Active	Public	No	Launch
admin	DebianOSImage	Image	Active	Shared	No	Launch

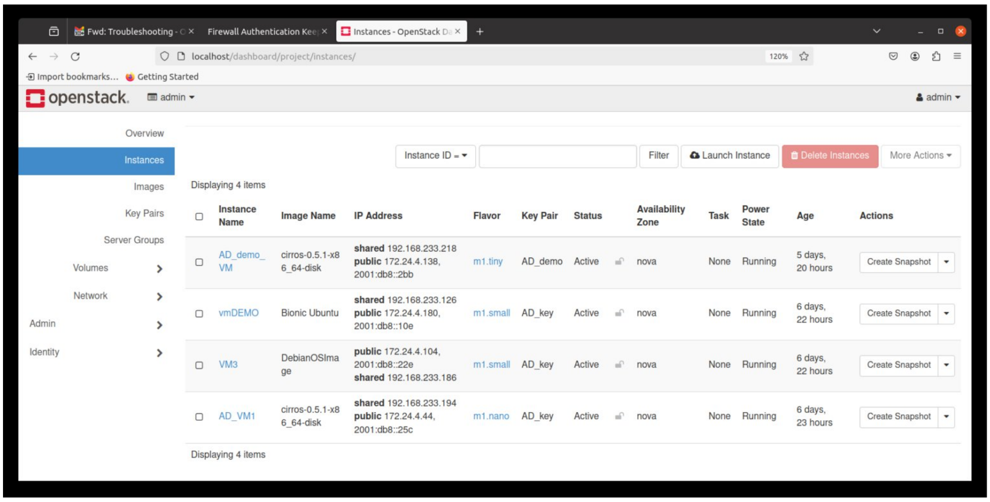
Displaying 3 items



```
stack@user-ThinkStation-P340: ~/devstack
stack@user-ThinkStation-P340:~/devstack$ openstack image list
+--------------------------------------+--------------------------+--------+
| ID                                   | Name                     | Status |
+--------------------------------------+--------------------------+--------+
| 810cf4c5-e133-4e69-ba0f-592a8a01a9ab | Bionic Ubuntu            | active |
| da5aa531-68df-49bd-8813-a4d2dea44478 | DebianOSImage            | active |
| 7eae7261-f25c-4dc8-a09c-e24fa8944c71 | cirros-0.5.1-x86_64-disk | active |
+--------------------------------------+--------------------------+--------+
stack@user-ThinkStation-P340:~/devstack$ ,
```

12. Add Instances with Different Configurations:

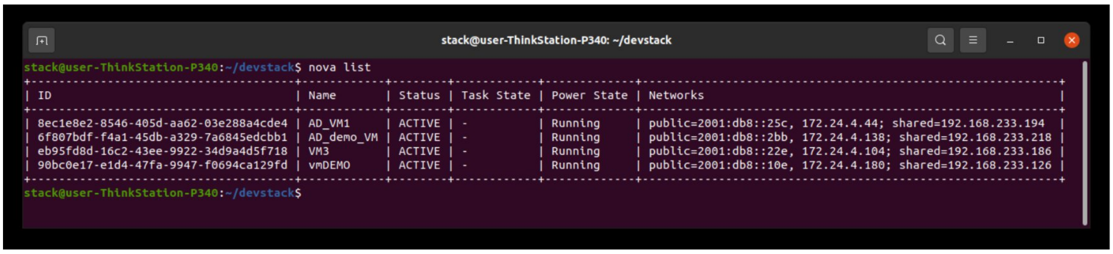
Repeat the aforementioned process to launch instances with varied configurations as needed.



Displaying 4 items

Instance Name	Image Name	IP Address	Flavor	Key Pair	Status	Availability Zone	Task	Power State	Age	Actions
AD_demo_VM	cirros-0.5.1-x86_64-disk	**shared** 192.168.233.218 **public** 172.24.4.138, 2001:db8::2bb	m1.tiny	AD_demo	Active	nova	None	Running	5 days, 20 hours	Create Snapshot
vmDEMO	Bionic Ubuntu	**shared** 192.168.233.126 **public** 172.24.4.180, 2001:db8::10e	m1.small	AD_key	Active	nova	None	Running	6 days, 22 hours	Create Snapshot
VM3	DebianOSImage	**public** 172.24.4.104, 2001:db8::22e **shared** 192.168.233.186	m1.small	AD_key	Active	nova	None	Running	6 days, 22 hours	Create Snapshot
AD_VM1	cirros-0.5.1-x86_64-disk	**shared** 192.168.233.194 **public** 172.24.4.44, 2001:db8::25c	m1.nano	AD_key	Active	nova	None	Running	6 days, 23 hours	Create Snapshot

Displaying 4 items



```
stack@user-ThinkStation-P340:~/devstack$ nova list
+--------------------------------------+------------+--------+------------+-------------+------------------------------------------------------------+
| ID                                   | Name       | Status | Task State | Power State | Networks                                                   |
+--------------------------------------+------------+--------+------------+-------------+------------------------------------------------------------+
| 8ec1e8e2-8546-405d-aa62-03e288a4cde4 | AD_VM1     | ACTIVE | -          | Running     | public=2001:db8::25c, 172.24.4.44; shared=192.168.233.194  |
| 6f807bdf-f4a1-45db-a329-7a6845edcbb1 | AD_demo_VM | ACTIVE | -          | Running     | public=2001:db8::2bb, 172.24.4.138; shared=192.168.233.218 |
| eb95fd8d-16c2-43ee-9922-34d9a4d5f718 | VM3        | ACTIVE | -          | Running     | public=2001:db8::22e, 172.24.4.104; shared=192.168.233.186 |
| 90bc0e17-e1d4-47fa-9947-f0694ca129fd | vmDEMO     | ACTIVE | -          | Running     | public=2001:db8::10e, 172.24.4.180; shared=192.168.233.126 |
+--------------------------------------+------------+--------+------------+-------------+------------------------------------------------------------+
stack@user-ThinkStation-P340:~/devstack$
```

RESULT:

Thus, the creation of virtual machine with different configuration in open stack cloud was executed successfully.

EXP NO.: 4 CREATE VIRTUAL MACHINE AND ATTACH A VOLUME IN OPENSTACK

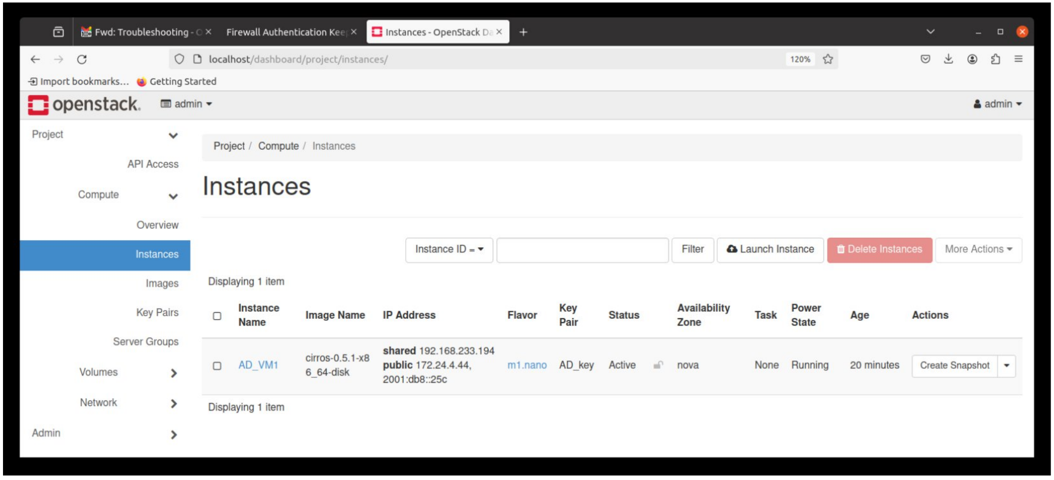
DATE:

AIM:

To create a virtual machine and attach a volume in openstack.

PROCEDURE:

- **Follow procedures [1 – 4] from the previous experiment (Experiment 3) and continue with the below mentioned steps.**



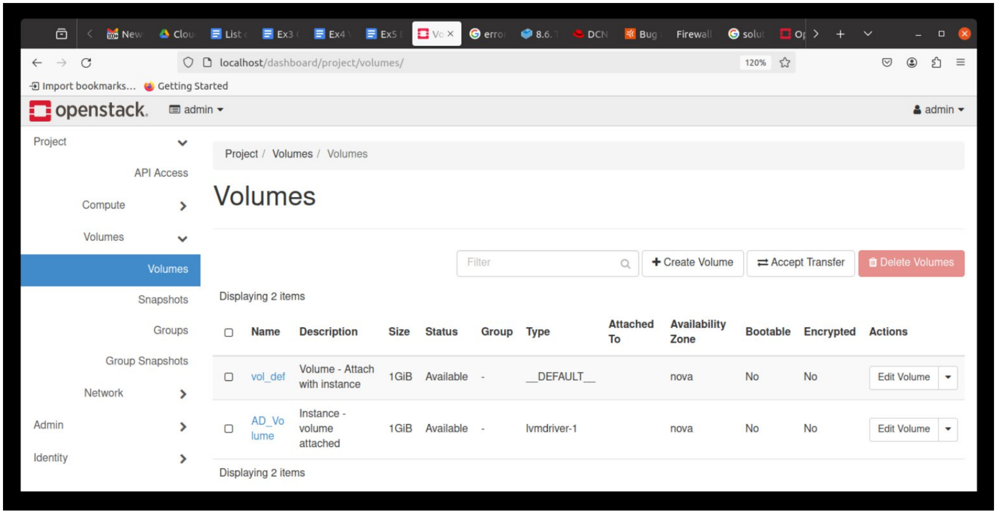
Displaying 1 item

	Instance Name	Image Name	IP Address	Flavor	Key Pair	Status		Availability Zone	Task	Power State	Age	Actions
□	AD_VM1	cirros-0.5.1-x86_64-disk	**shared** 192.168.233.194 **public** 172.24.4.44, 2001:db8::25c	m1.nano	AD_key	Active		nova	None	Running	20 minutes	Create Snapshot

Displaying 1 item

1. Create Volume:

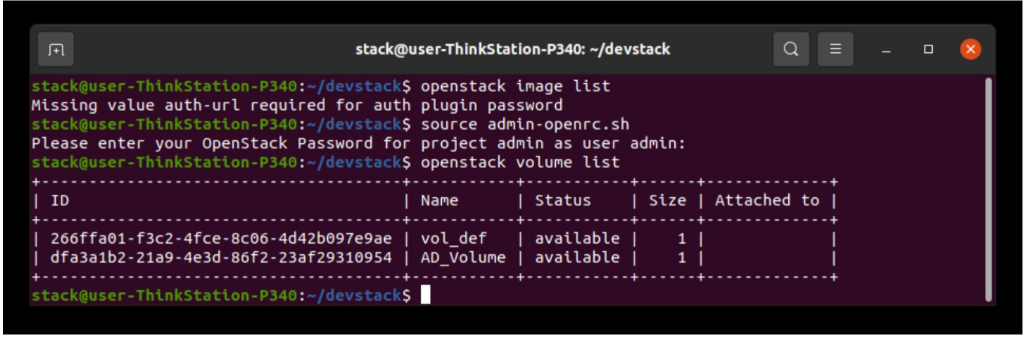
- Once instance is successfully launched, proceed to create a volume.
- Access the volume creation section in the OpenStack dashboard.



Displaying 2 items

	Name	Description	Size	Status	Group	Type	Attached To	Availability Zone	Bootable	Encrypted	Actions
□	vol_def	Volume - Attach with instance	1GiB	Available	-	__DEFAULT__		nova	No	No	Edit Volume
□	AD_Volume	Instance - volume attached	1GiB	Available	-	lvmdriver-1		nova	No	No	Edit Volume

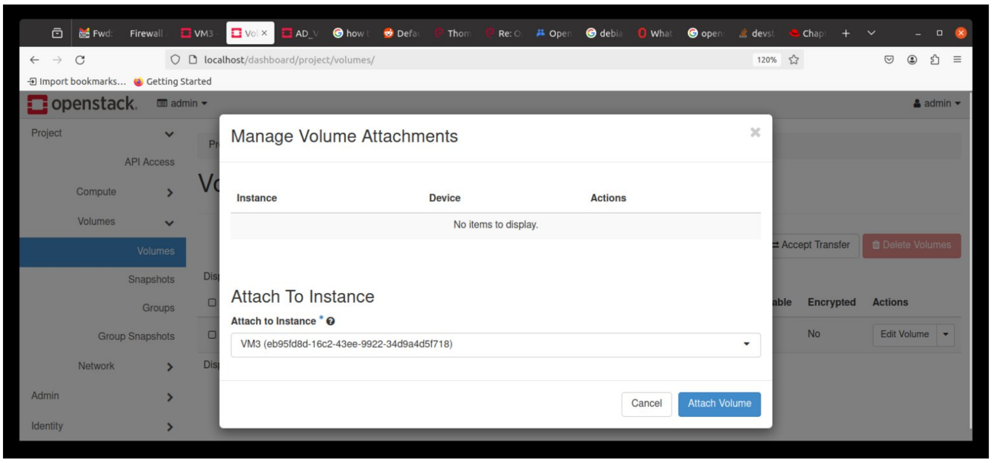
Displaying 2 items


stack@user-ThinkStation-P340: ~/devstack

```
stack@user-ThinkStation-P340:~/devstack$ openstack image list
Missing value auth-url required for auth plugin password
stack@user-ThinkStation-P340:~/devstack$ source admin-openrc.sh
Please enter your OpenStack Password for project admin as user admin:
stack@user-ThinkStation-P340:~/devstack$ openstack volume list
+--------------------------------------+-----------+-----------+------+-------------+
| ID                                   | Name      | Status    | Size | Attached to |
+--------------------------------------+-----------+-----------+------+-------------+
| 266ffa01-f3c2-4fce-8c06-4d42b097e9ae | vol_def   | available |    1 |             |
| dfa3a1b2-21a9-4e3d-86f2-23af29310954 | AD_Volume | available |    1 |             |
+--------------------------------------+-----------+-----------+------+-------------+
stack@user-ThinkStation-P340:~/devstack$ 
```

2. Attach Volume to Instance:

- Once volume is created, navigate to the list of available volumes.
- Select the required volume from the list.
- Attach the selected volume to the instance that was launched earlier.



Manage Volume Attachments

Instance	Device	Actions
No items to display.		

Attach To Instance

Attach to Instance *: VM3 (eb95fd8d-16c2-43ee-9922-34d9a4d5f718)

RESULT:

Thus, creation of a virtual machine and attachment of a volume in openstack cloud was executed successfully.

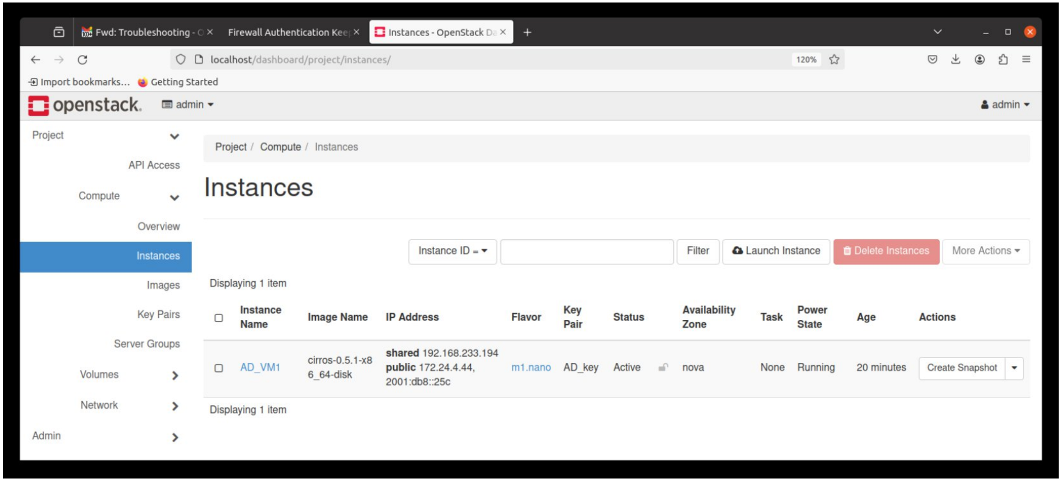
EXP NO.: 5 **EXECUTION OF COMMANDS IN VIRTUAL MACHINE**
DATE:

AIM:

To execute the commands in virtual machine.

PROCEDURE:

- **Follow procedures [1 – 4] from the previous experiment (Experiment 3) and continue with the below mentioned steps.**



localhost/dashboard/project/instances/

Project / Compute / Instances

Instances

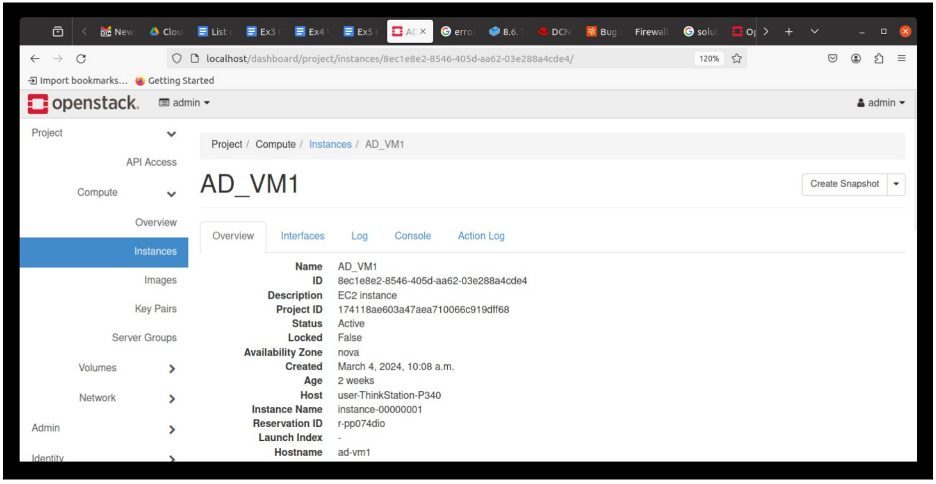
Displaying 1 item

	Instance Name	Image Name	IP Address	Flavor	Key Pair	Status		Availability Zone	Task	Power State	Age	Actions
□	AD_VM1	cirros-0.5.1-x86_64-disk	**shared** 192.168.233.194 **public** 172.24.4.44, 2001:db8::25c	m1.nano	AD_key	Active		nova	None	Running	20 minutes	Create Snapshot

Displaying 1 item

1. Instance Details:

- Access the OpenStack dashboard.
- Navigate to the section for viewing instance details.



localhost/dashboard/project/instances/8ec1e8e2-8546-405d-aa62-03e288a4cde4/

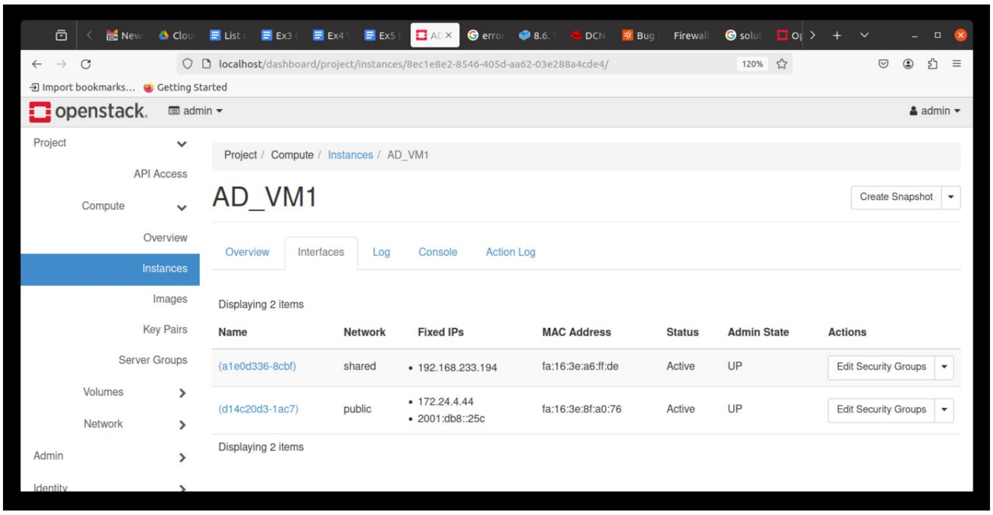
Project / Compute / Instances / AD_VM1

AD_VM1

Name	AD_VM1
ID	8ec1e8e2-8546-405d-aa62-03e288a4cde4
Description	EC2 instance
Project ID	174118ae603a47aea710066c919dff68
Status	Active
Locked	False
Availability Zone	nova
Created	March 4, 2024, 10:08 a.m.
Age	2 weeks
Host	user-ThinkStation-P340
Instance Name	instance-00000001
Reservation ID	r-pp074dio
Launch Index	-
Hostname	ad-vm1

2. Check Instance Interfaces:

- Access the OpenStack dashboard.
- Navigate to the "Interfaces" section.
- Ensure that the launched instance is active and its admin state is UP.



Project / Compute / Instances / AD_VM1

AD_VM1

Displaying 2 items

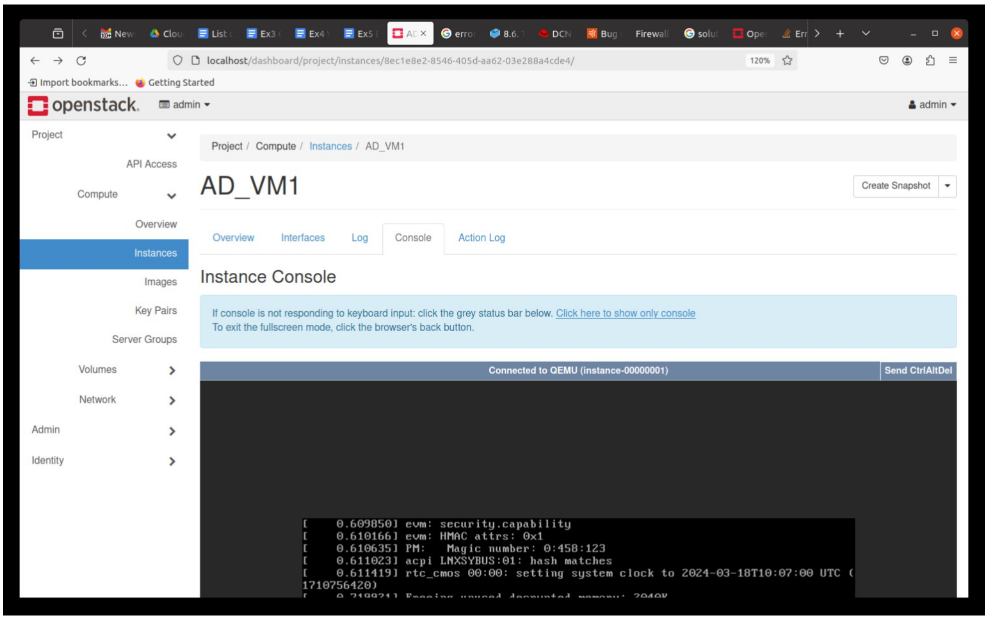
Name	Network	Fixed IPs	MAC Address	Status	Admin State	Actions
(a1e0d336-8cbf)	shared	• 192.168.233.194	fa:16:3e:a6:ff:de	Active	UP	Edit Security Groups
(d14c20d3-1ac7)	public	• 172.24.4.44 • 2001:db8::25c	fa:16:3e:8f:a0:76	Active	UP	Edit Security Groups

Displaying 2 items

3. Access Instance Console:

- Proceed to the console section in the OpenStack dashboard.
- Access the console of the instance.
- Utilize the operating system that was selected during the instance creation process.

CirrOS (Guest OS):



Project / Compute / Instances / AD_VM1

AD_VM1

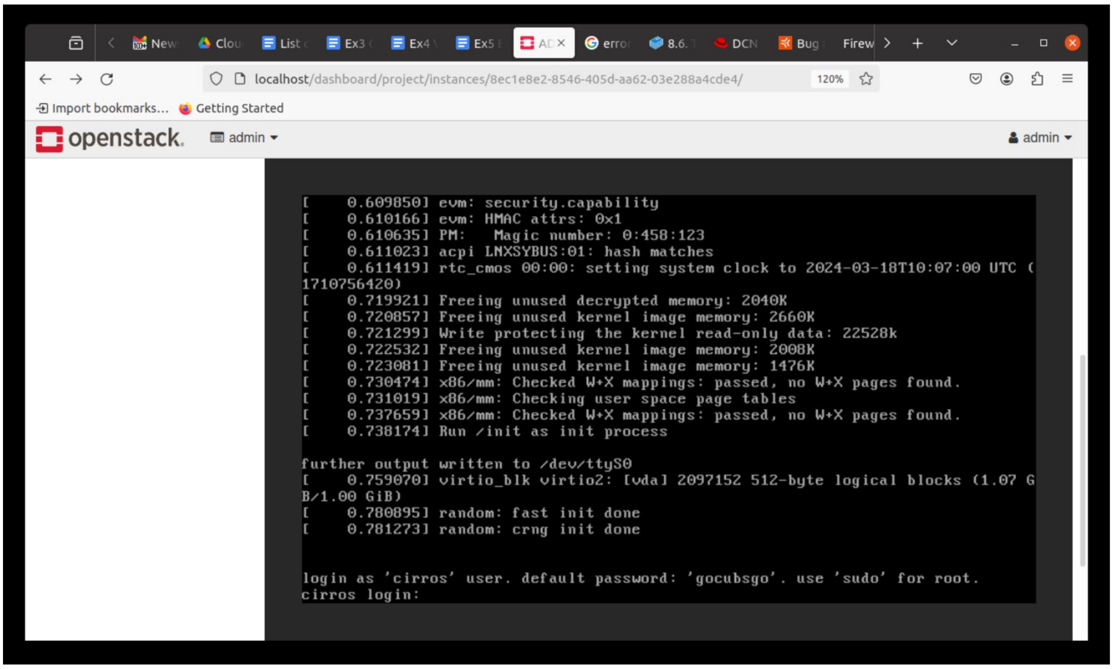
Instance Console

If console is not responding to keyboard input: click the grey status bar below. Click here to show only console
To exit the fullscreen mode, click the browser's back button.

Connected to QEMU (instance-00000001)

```
[    0.609850] evm: security.capability
[    0.610166] evm: HMAC attrs: 0x1
[    0.610635] PM:   Magic number: 0:458:123
[    0.611023] acpi LNXSYBUS:01: hash matches
[    0.611419] rtc_cmos 00:00: setting system clock to 2024-03-18T10:07:00 UTC (
1710756420)
[    0.719921] Freeing unused decrypted memory: 2040K
```

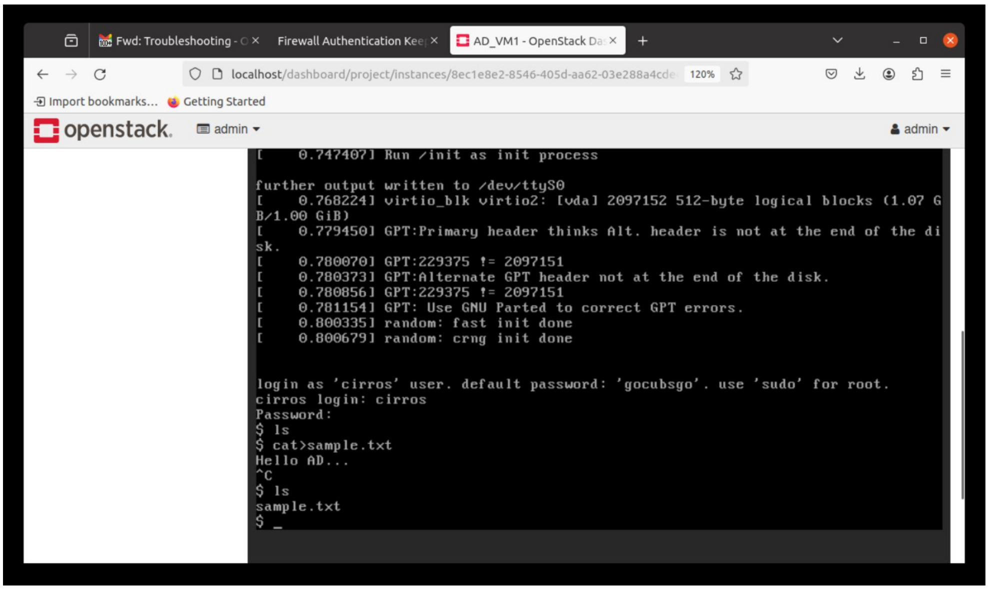
CirrOS login: Login into the CirrOS in the console. Login into the guest account.


localhost/dashboard/project/instances/8ec1e8e2-8546-405d-aa62-03e288a4cde4/
120%
Import bookmarks...
Getting Started
openstack.
admin
admin

```
[    0.609850] evm: security.capability
[    0.610166] evm: HMAC attrs: 0x1
[    0.610635] PM:   Magic number: 0:458:123
[    0.611023] acpi LNXSYBUS:01: hash matches
[    0.611419] rtc_cmos 00:00: setting system clock to 2024-03-18T10:07:00 UTC (
1710756420)
[    0.719921] Freeing unused decrypted memory: 2040K
[    0.720857] Freeing unused kernel image memory: 2660K
[    0.721299] Write protecting the kernel read-only data: 22528k
[    0.722532] Freeing unused kernel image memory: 2008K
[    0.723081] Freeing unused kernel image memory: 1476K
[    0.730474] x86/mm: Checked W+X mappings: passed, no W+X pages found.
[    0.731019] x86/mm: Checking user space page tables
[    0.737659] x86/mm: Checked W+X mappings: passed, no W+X pages found.
[    0.738174] Run /init as init process

further output written to /dev/ttyS0
[    0.759070] virtio_blk virtio2: [vda] 2097152 512-byte logical blocks (1.07 G
B/1.00 GiB)
[    0.780895] random: fast init done
[    0.781273] random: crng init done


login as 'cirros' user. default password: 'gocubsgo'. use 'sudo' for root.
cirros login:
```

4. Execute Commands:

- Once logged in, execute the desired commands in the console.
- Utilize the console interface to interact with the CirrOS guest OS and execute necessary tasks or commands.


Fwd: Troubleshooting
Firewall Authentication Kee
AD_VM1 - OpenStack Da
localhost/dashboard/project/instances/8ec1e8e2-8546-405d-aa62-03e288a4cde
120%
Import bookmarks...
Getting Started
openstack.
admin
admin

```
[    0.747407] Run /init as init process

further output written to /dev/ttyS0
[    0.768224] virtio_blk virtio2: [vda] 2097152 512-byte logical blocks (1.07 G
B/1.00 GiB)
[    0.779450] GPT:Primary header thinks Alt. header is not at the end of the di
sk.
[    0.780070] GPT:229375 != 2097151
[    0.780373] GPT:Alternate GPT header not at the end of the disk.
[    0.780856] GPT:229375 != 2097151
[    0.781154] GPT: Use GNU Parted to correct GPT errors.
[    0.800335] random: fast init done
[    0.800679] random: crng init done


login as 'cirros' user. default password: 'gocubsgo'. use 'sudo' for root.
cirros login: cirros
Password:
$ ls
$ cat>sample.txt
Hello AD...
^C
$ ls
sample.txt
$ _
```

RESULT:

Thus, execution of commands in the virtual machine was successfully completed.

EXP NO.: 6(a)
DATE:

DEVELOPING “HELLO WORLD” WEB APPLICATION AND DEPLOYING IN GOOGLE APP ENGINE

AIM:

To create a Hello World web application and deploying in google app engine.

SOURCE CODE:

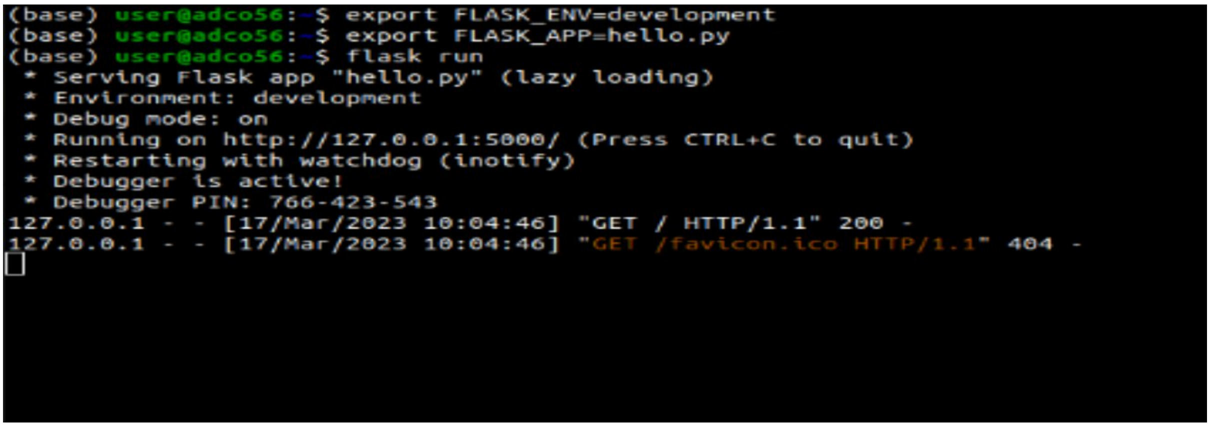
hello.py

```
from flask import Flask
app = Flask(__name__)
@app.route("/")
def hello_world():
    return "<p>AD...Hello, World!</p>"
```

COMMANDS:

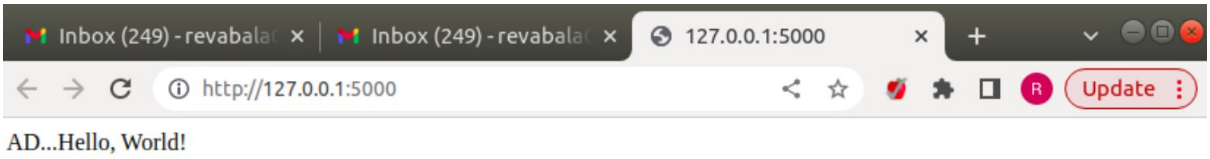
```
export FLASK_ENV=development
export FLASK_App=hello.py
flask run
```

EXECUTION:



```
(base) user@adco56:~$ export FLASK_ENV=development
(base) user@adco56:~$ export FLASK_APP=hello.py
(base) user@adco56:~$ flask run
 * Serving Flask app "hello.py" (lazy loading)
 * Environment: development
 * Debug mode: on
 * Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)
 * Restarting with watchdog (inotify)
 * Debugger is active!
 * Debugger PIN: 766-423-543
127.0.0.1 - - [17/Mar/2023 10:04:46] "GET / HTTP/1.1" 200 -
127.0.0.1 - - [17/Mar/2023 10:04:46] "GET /favicon.ico HTTP/1.1" 404 -
```

OUTPUT:



AD...Hello, World!

RESULT:

Hence the Hello World Web Application is successfully deployed in Google App Engine.

EXP NO.: 6(b) **DESIGN A REGISTRATION FORM AND PROCESS THE FORM DATA**
DATE:

AIM:

To design a registration form and process the form data.

SOURCE CODE:

hello.py

```
from flask import *
app = Flask(__name__)
@app.route("/", methods=["GET", "POST"])
def home():
    if request.method == "POST":
       name=request.form.get("name")
       email=request.form.get("email")
       phno=request.form.get("ph.no")
       address=request.form.get("addr")

       return"<p><strong>NAME:</strong>"+name+"</p><p><strong>EMAIL:</strong>"+email+"
</p><p><strong>PHONE:</strong>"+phno+"</p><p><strong>ADDRESS:</strong>"+address+"<
/p>"
    return render_template("index.html")

if __name__ == "__main__":
    app.run()
```

index.html:

```
<h1>
    This is home.html
</h1>
<form method="post">
    Name: <input name="name"/>
<br/><br/>
Email: <input name="email"/>
 <br/><br/>
 Ph.no:<input name="ph.no"/>
<br/><br/>
```

Address:<input name="addr"/>

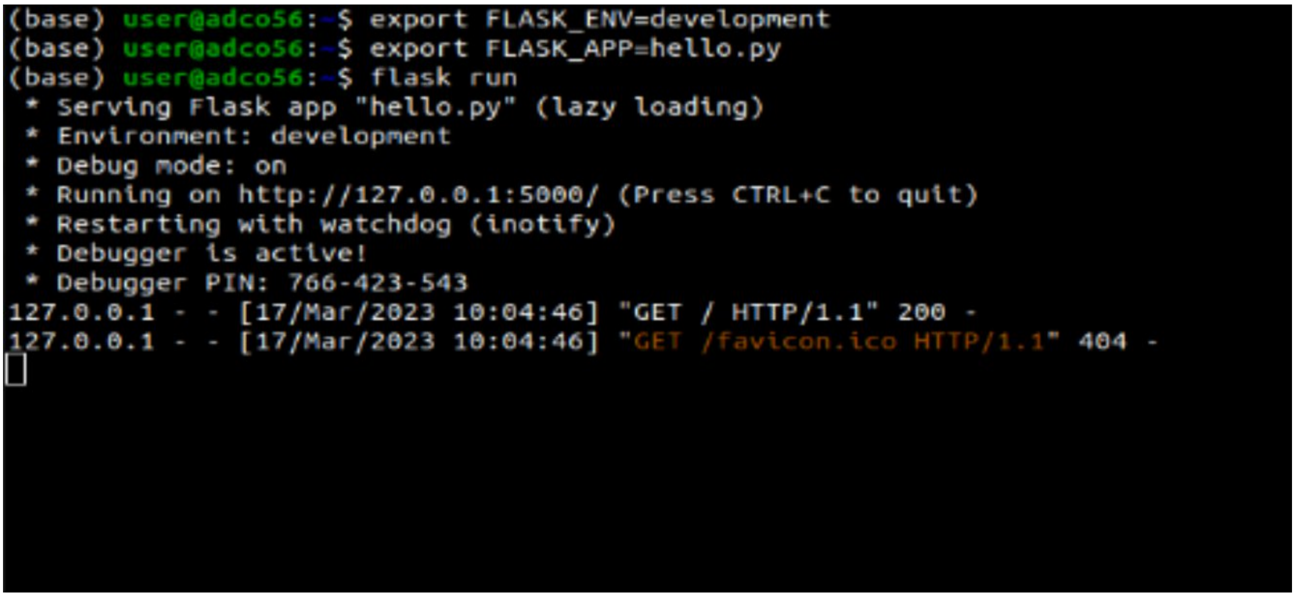
<button type="submit">Submit</button>

</form>

COMMANDS:

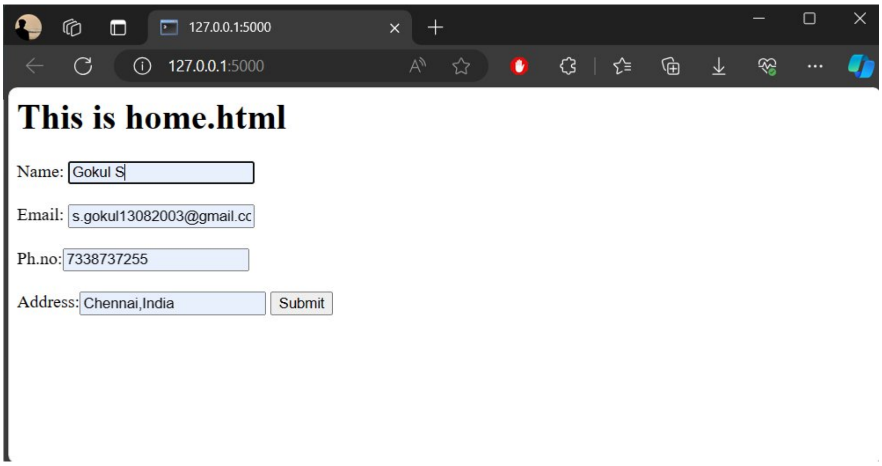
```
export FLASK_ENV=development
export FLASK_App=hello.py
flask run
```

EXECUTION:



```
(base) user@adco56:~$ export FLASK_ENV=development
(base) user@adco56:~$ export FLASK_APP=hello.py
(base) user@adco56:~$ flask run
 * Serving Flask app "hello.py" (lazy loading)
 * Environment: development
 * Debug mode: on
 * Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)
 * Restarting with watchdog (inotify)
 * Debugger is active!
 * Debugger PIN: 766-423-543
127.0.0.1 - - [17/Mar/2023 10:04:46] "GET / HTTP/1.1" 200 -
127.0.0.1 - - [17/Mar/2023 10:04:46] "GET /favicon.ico HTTP/1.1" 404 -
```

OUTPUT:



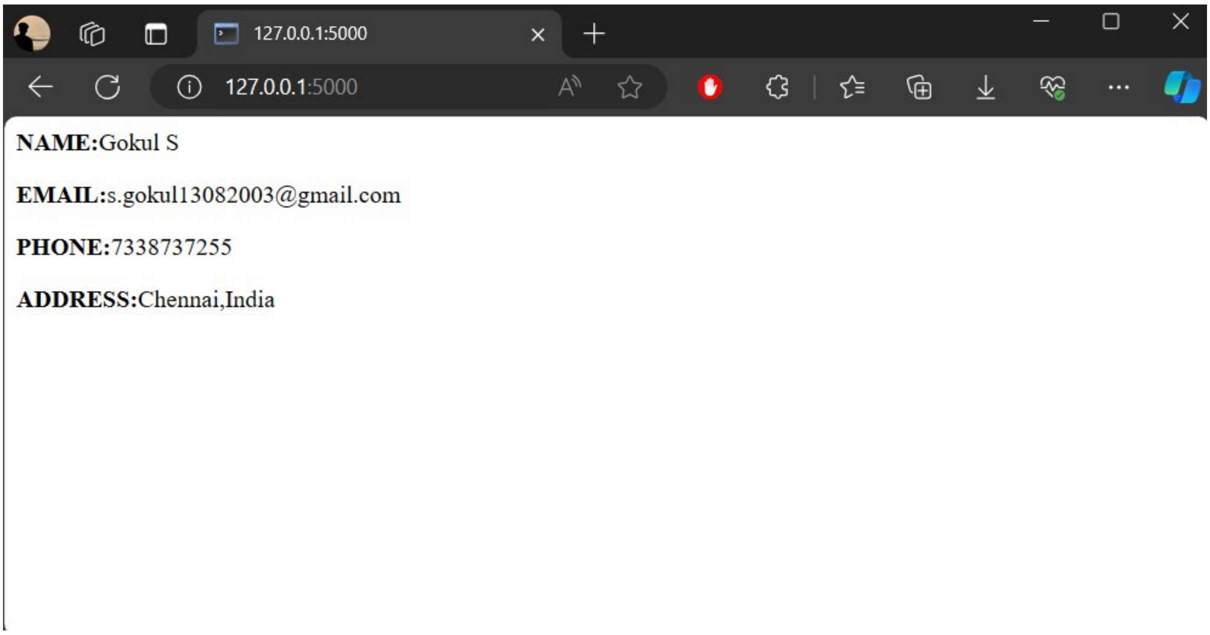
This is home.html

Name: Gokul S

Email: s.gokul13082003@gmail.cc

Ph.no: 7338737255

Address: Chennai,India Submit



NAME: Gokul S

EMAIL:s.gokul13082003@gmail.com

PHONE:7338737255

ADDRESS: Chennai,India

RESULT:
Hence the designing of registration form and processing the form data is successfully deployed in Google App Engine.

EXP NO.: 6(c) DESIGN A REGISTRATION FORM WITH VARIOUS CONTROLS AND VALIDATE
DATE:

AIM:

To design a registration form with various control and validation using javascript.

SOURCE CODE:

hello.py

```
from flask import *
app = Flask(__name__)
@app.route("/", methods=["GET", "POST"])
def home():
    if request.method == "POST":
        na=request.form.get("name")
        return render_template("response.html", na=na)
    return render_template("temp.html")

if __name__ == "__main__":
    app.run()
```

temp.html:

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible"
        content="IE=edge">
    <meta name="viewport"
        content="width=device-width, initial-scale=1.0">
    <title>
        Build a Survey Form using HTML and CSS
    </title>
    <style>
        /* Styling the Body element
        i.e. Color, Font, Alignment */
        body {
            background-color: black;
    font-family: Verdana;
    text-align: center;
}

/* Styling the Form (Color, Padding, Shadow) */
form {
    background-color: red;
    max-width: 500px;
    margin: 50px auto;
    padding: 30px 20px;
    box-shadow: 2px 5px 10px rgba(0, 0.5, 0, 0.5);
}

/* Styling form-control Class */
.form-control {
    text-align: left;
    margin-bottom: 25px;
}

/* Styling form-control Label */
.form-control label {
    display: block;
    margin-bottom: 10px;
}

/* Styling form-control input,
select, textarea */
.form-control input,
.form-control select,
.form-control textarea {
    border: 1px solid #777;
    border-radius: 2px;
    font-family: inherit;
    padding: 10px;
    display: block;
```

```
            width: 95%;
        }

        /* Styling form-control Radio
        button and Checkbox */
        .form-control input[type="radio"],
        .form-control input[type="checkbox"] {
            display: inline-block;
            width: auto;
        }

        /* Styling Button */
        button {
            background-color: green;
            border: 1px solid #777;
            border-radius: 2px;
            font-family: inherit;
            font-size: 21px;
            display: block;
            width: 100%;
            margin-top: 50px;
            margin-bottom: 20px;
        }

        /*Error validation */
        .error {
        color: white;
    }
    </style>
</head>

<body>
    <h1>Course Survey Form</h1>

    <!-- Create Form -->
```

```
<form id="form" method="post">

    <!-- Details -->
    <div class="form-control">
        <label for="name" id="label-name">
            Name
        </label>

        <!-- Input Type Text -->
        <input type="text" id="name" name="name"
            placeholder="Enter your name" />
        <span class="error" id="nameError"></span>
    </div>

    <div class="form-control">
        <label for="email" id="label-email">
            Email
        </label>
        <!-- Input Type Email-->
        <input type="email" id="email" name="email"
            placeholder="Enter your email" />
        <span class="error" id="emailError"></span>
    </div>

    <div class="form-control">
        <label for="role" id="label-role">
            Which option best describes you?
        </label>

        <!-- Dropdown options -->
        <select name="role" id="role">
            <option value="student">Student</option>
            <option value="intern">Intern</option>
            <option value="professional">Professional</option>
            <option value="other">Other</option>
        </select>
    </div>




    <div class="form-control">
        <label for="comment">
            Any comments or suggestions
        </label>

        <!-- multi-line text input control -->
        <textarea name="comment" id="comment"
                placeholder="Enter your comment here">
        </textarea>
    </div>

    <!-- Multi-line Text Input Control -->
    <button type="submit" value="submit">
        Submit
    </button>
</form>
<script>
    const form = document.getElementById('form');
    const nameInput = document.getElementById('name');
    const emailInput = document.getElementById('email');
    const nameError = document.getElementById('nameError');
    const emailError = document.getElementById('emailError');

    form.addEventListener('submit', function(event) {
        let valid = true;
        nameError.textContent = '';
        emailError.textContent = '';

        if (nameInput.value === '') {
            nameError.textContent = 'Name is required';
```

```
                valid = false;
            }

            if (emailInput.value === '') {
                emailError.textContent = 'Email is required';
                valid = false;
            } else if (!isValidEmail(emailInput.value)) {
                emailError.textContent = 'Invalid email format';
                valid = false;
            }
            if (!valid) {
                event.preventDefault();
            }
        });

        function isValidEmail(email) {
            const emailRegex = /^[^\s@]+@[^\s@]+\.[^\s@]+$/;
            return emailRegex.test(email);
        }
    </script>
</body>
</html>
```

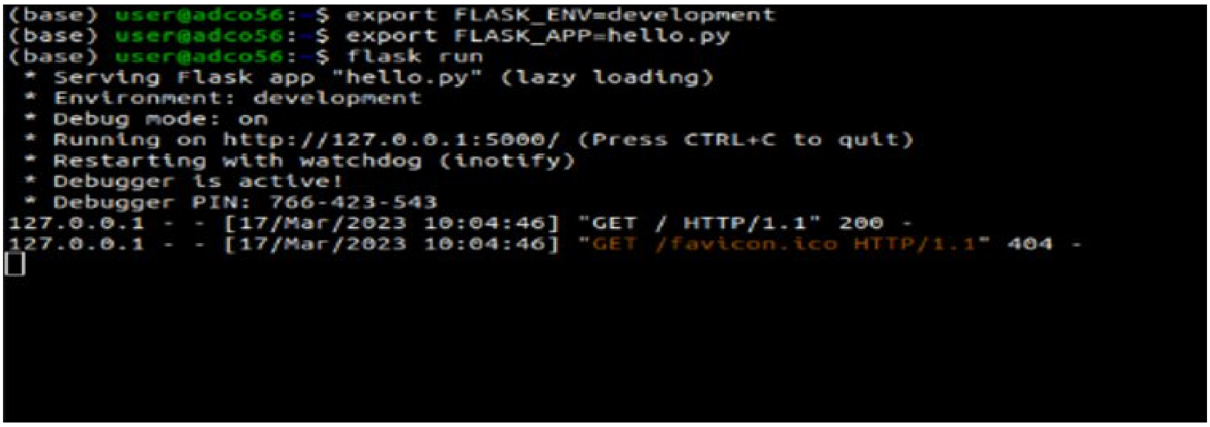
response.html:

```
<html lang="en">
<head>
    <meta charset="utf-8" />
    <meta http-equiv="X-UA-Compatible" content="IE=edge,chrome=1" />
    <meta name="viewport" content="width=device-width, initial-scale=1">
    <title></title>
    <link href='https://fonts.googleapis.com/css?family=Lato:300,400|Montserrat:700'
rel='stylesheet' type='text/css'>
    <style>
        @import url(//cdnjs.cloudflare.com/ajax/libs/normalize/3.0.1/normalize.min.css);
        @import url(//maxcdn.bootstrapcdn.com/font-awesome/4.2.0/css/font-awesome.min.css);
    </style>
    <link rel="stylesheet" href="https://2-22-4-dot-lead-pages.appspot.com/static/lp918/min/default_thank_you.css">
    <script src="https://2-22-4-dot-lead-pages.appspot.com/static/lp918/min/jquery-1.9.1.min.js"></script>
    <script src="https://2-22-4-dot-lead-pages.appspot.com/static/lp918/min/html5shiv.js"></script>
</head>
<body>
    <header class="site-header" id="header">
        <h1 class="site-header__title" data-lead-id="site-header-title">THANK YOU!</h1>
    </header>

    <div class="main-content">
        <i class="fa fa-check main-content__checkmark" id="checkmark"></i>
        <p class="main-content__body" data-lead-id="main-content-body">Thank you {{ na }} for filling the survey out.</p>
    </div>
</body>
</html>
```

COMMANDS:

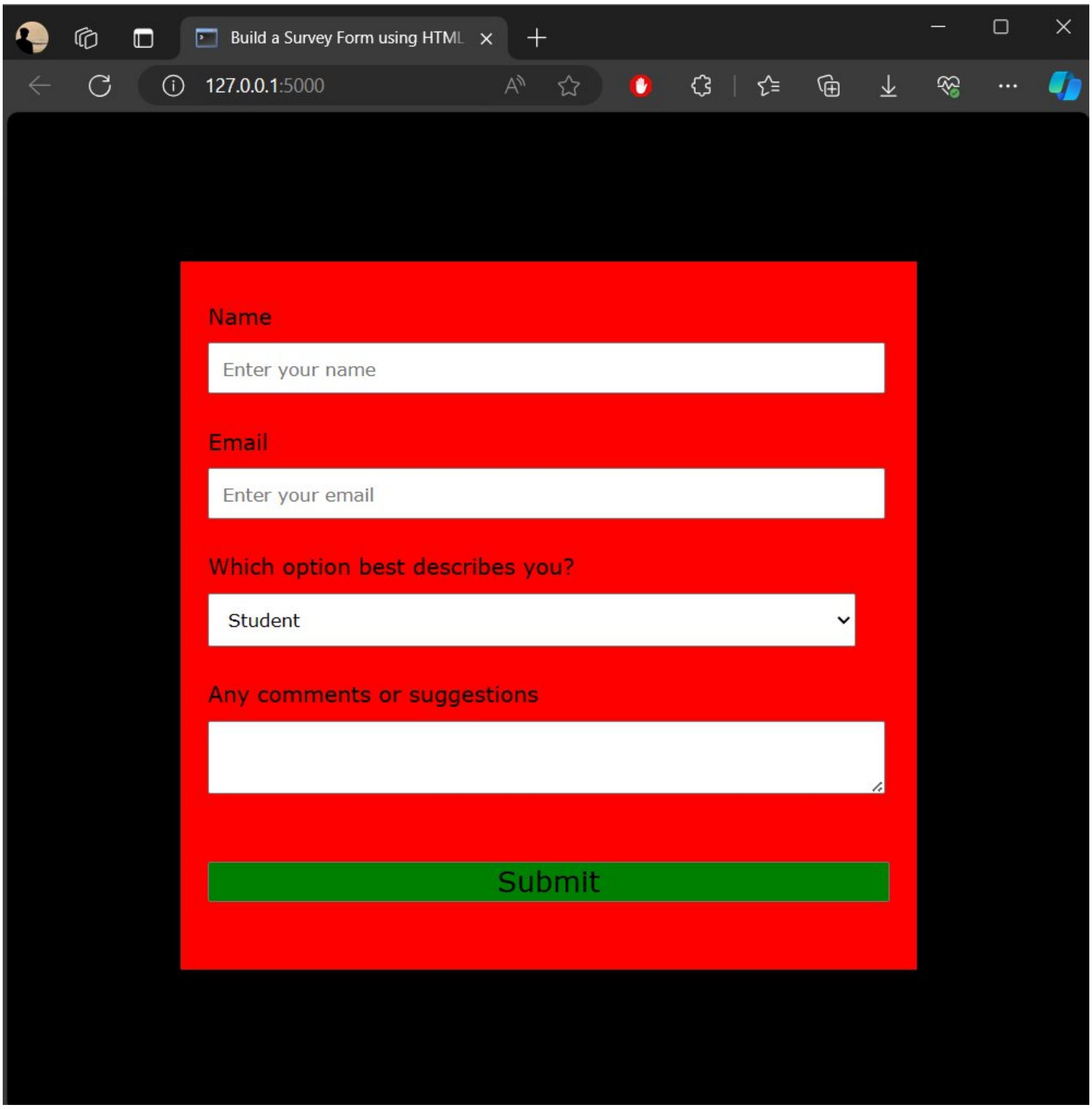
```
export FLASK_ENV=development
export FLASK_App=hello.py
flask run
```

EXECUTION:



```
(base) user@adco56:~$ export FLASK_ENV=development
(base) user@adco56:~$ export FLASK_APP=hello.py
(base) user@adco56:~$ flask run
 * Serving Flask app "hello.py" (lazy loading)
 * Environment: development
 * Debug mode: on
 * Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)
 * Restarting with watchdog (inotify)
 * Debugger is active!
 * Debugger PIN: 766-423-543
127.0.0.1 - - [17/Mar/2023 10:04:46] "GET / HTTP/1.1" 200 -
127.0.0.1 - - [17/Mar/2023 10:04:46] "GET /favicon.ico HTTP/1.1" 404 -
```

OUTPUT:



Name

Enter your name

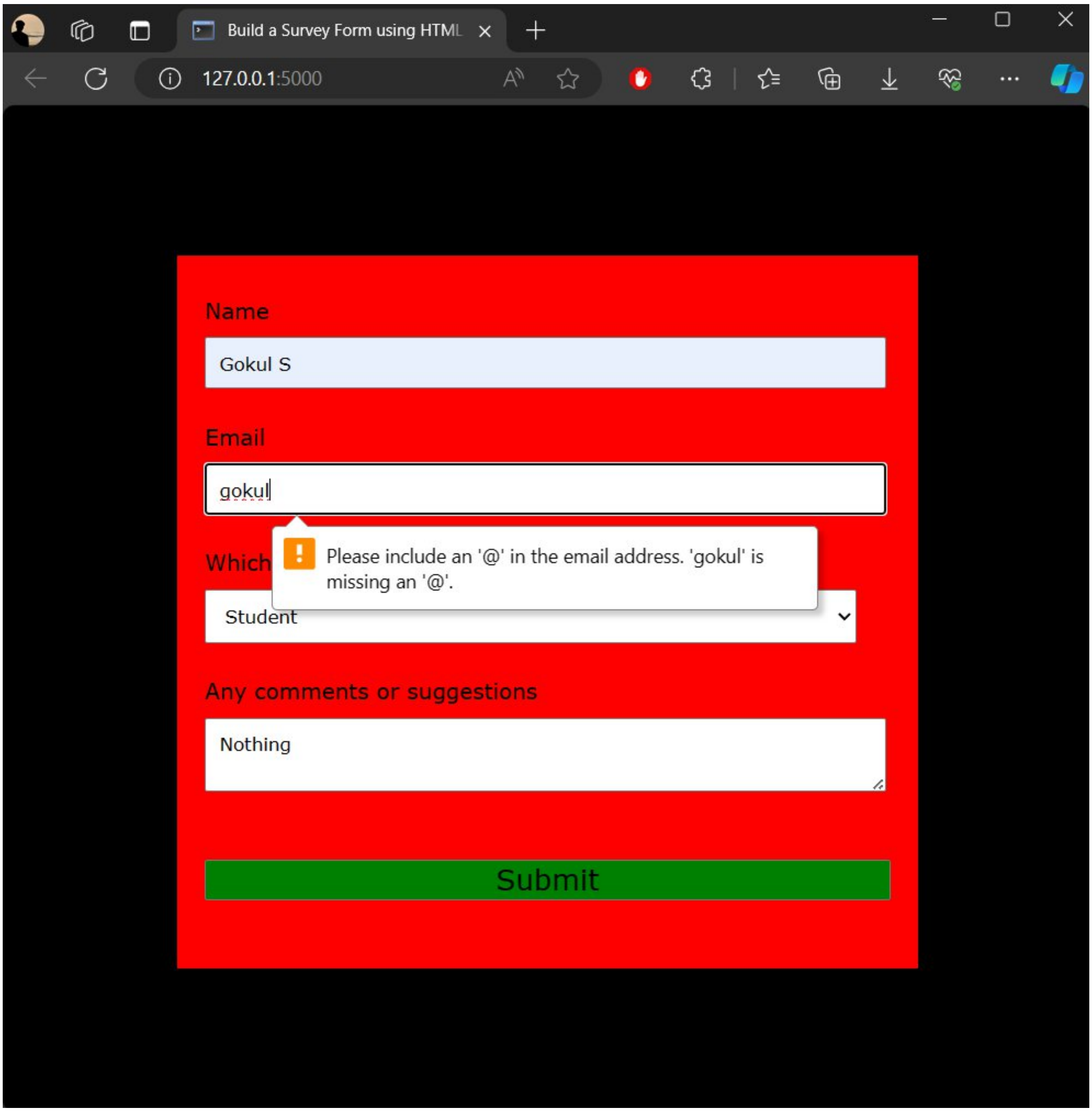
Email

Enter your email

Which option best describes you?

Student

Any comments or suggestions



Name

Gokul S

Email

gokul

Please include an '@' in the email address. 'gokul' is missing an '@'.

Which

Student

Any comments or suggestions

Nothing

RESULT:

Hence the designing of a registration form with various control and validation using javascript is successfully deployed in Google App Engine.

EX NO: 7

SET UP THE ONE NODE HADOOP CLUSTER

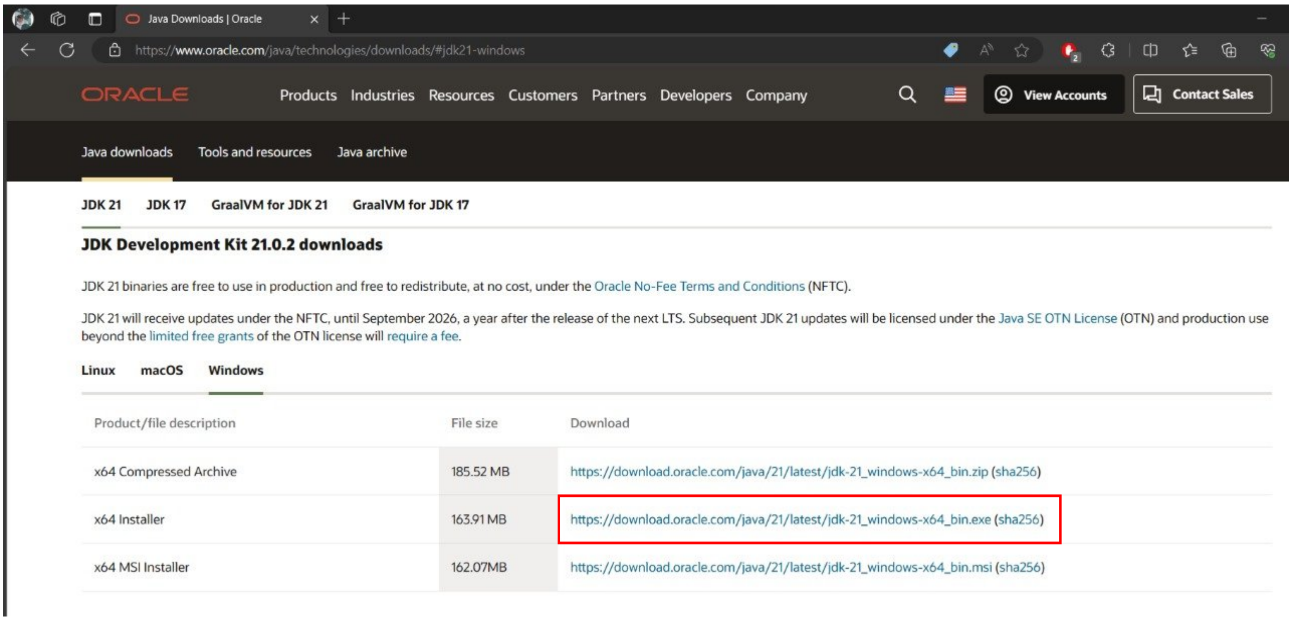
DATE:

AIM: To set up the one node hadoop cluster

Steps for Hadoop Installation

STEP 1- Install Java

https://www.oracle.com/in/java/technologies/downloads/


Java Downloads | Oracle
https://www.oracle.com/java/technologies/downloads/#jdk21-windows
ORACLE
Products Industries Resources Customers Partners Developers Company
View Accounts
Contact Sales
Java downloads Tools and resources Java archive
JDK 21 JDK 17 GraalVM for JDK 21 GraalVM for JDK 17
Linux macOS Windows

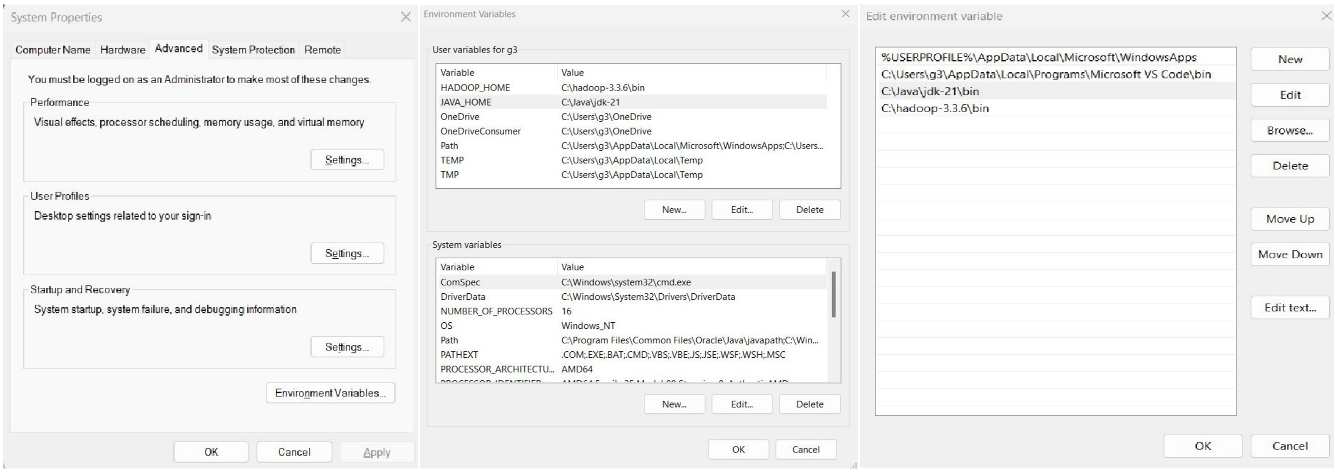
JDK Development Kit 21.0.2 downloads

JDK 21 binaries are free to use in production and free to redistribute, at no cost, under the Oracle No-Fee Terms and Conditions (NFTC).

JDK 21 will receive updates under the NFTC, until September 2026, a year after the release of the next LTS. Subsequent JDK 21 updates will be licensed under the Java SE OTN License (OTN) and production use beyond the limited free grants of the OTN license will require a fee.

Product/file description	File size	Download
x64 Compressed Archive	185.52 MB	https://download.oracle.com/java/21/latest/jdk-21_windows-x64_bin.zip (sha256)
x64 Installer	163.91 MB	https://download.oracle.com/java/21/latest/jdk-21_windows-x64_bin.exe (sha256)
x64 MSI Installer	162.07MB	https://download.oracle.com/java/21/latest/jdk-21_windows-x64_bin.msi (sha256)

STEP 2- Place the path of the bin folder into the Environment variable in your system



You must be logged on as an Administrator to make most of these changes.

Visual effects, processor scheduling, memory usage, and virtual memory

Desktop settings related to your sign-in

System startup, system failure, and debugging information

User variables for g3

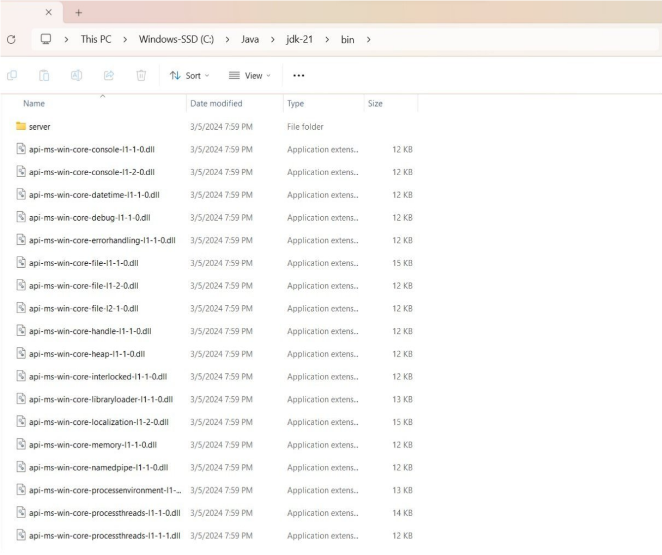
Variable	Value
HADOOP_HOME	C:\hadoop-3.3.6\bin
JAVA_HOME	C:\Java\jdk-21
OneDrive	C:\Users\g3\OneDrive
OneDriveConsumer	C:\Users\g3\OneDrive
Path	C:\Users\g3\AppData\Local\Microsoft\WindowsApps;C:\Users...
TEMP	C:\Users\g3\AppData\Local\Temp
TMP	C:\Users\g3\AppData\Local\Temp

System variables

Variable	Value
ComSpec	C:\Windows\system32\cmd.exe
DriverData	C:\Windows\System32\Drivers\DriverData
NUMBER_OF_PROCESSORS	16
OS	Windows_NT
Path	C:\Program Files\Common Files\Oracle\Java\javapath;C:\Win...
PATHEXT	.COM;.EXE;.BAT;.CMD;.VBS;.VBE;.JS;.JSE;.WSF;.WSH;.MSC
PROCESSOR_ARCHITECTU...	AMD64

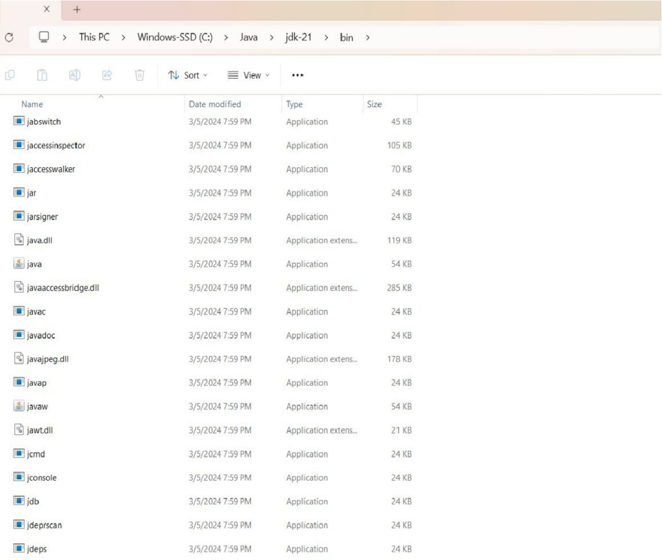
```
%USERPROFILE%\AppData\Local\Microsoft\WindowsApps
C:\Users\g3\AppData\Local\Programs\Microsoft VS Code\bin
C:\Java\jdk-21\bin
C:\hadoop-3.3.6\bin
```

STEP 3- Add JDK file to the path



This PC > Windows-SSD (C:) > Java > jdk-21 > bin >

Name	Date modified	Type	Size
server	3/5/2024 7:59 PM	File folder	
api-ms-win-core-console-l1-1-0.dll	3/5/2024 7:59 PM	Application extens...	12 KB
api-ms-win-core-console-l1-2-0.dll	3/5/2024 7:59 PM	Application extens...	12 KB
api-ms-win-core-datetime-l1-1-0.dll	3/5/2024 7:59 PM	Application extens...	12 KB
api-ms-win-core-debug-l1-1-0.dll	3/5/2024 7:59 PM	Application extens...	12 KB
api-ms-win-core-errorhandling-l1-1-0.dll	3/5/2024 7:59 PM	Application extens...	12 KB
api-ms-win-core-file-l1-1-0.dll	3/5/2024 7:59 PM	Application extens...	15 KB
api-ms-win-core-file-l1-2-0.dll	3/5/2024 7:59 PM	Application extens...	12 KB
api-ms-win-core-file-l2-1-0.dll	3/5/2024 7:59 PM	Application extens...	12 KB
api-ms-win-core-handle-l1-1-0.dll	3/5/2024 7:59 PM	Application extens...	12 KB
api-ms-win-core-heap-l1-1-0.dll	3/5/2024 7:59 PM	Application extens...	12 KB
api-ms-win-core-interlocked-l1-1-0.dll	3/5/2024 7:59 PM	Application extens...	12 KB
api-ms-win-core-libraryloader-l1-1-0.dll	3/5/2024 7:59 PM	Application extens...	13 KB
api-ms-win-core-localization-l1-2-0.dll	3/5/2024 7:59 PM	Application extens...	15 KB
api-ms-win-core-memory-l1-1-0.dll	3/5/2024 7:59 PM	Application extens...	12 KB
api-ms-win-core-namedpipe-l1-1-0.dll	3/5/2024 7:59 PM	Application extens...	12 KB
api-ms-win-core-processenvironment-l1-...	3/5/2024 7:59 PM	Application extens...	13 KB
api-ms-win-core-processthreads-l1-1-0.dll	3/5/2024 7:59 PM	Application extens...	14 KB
api-ms-win-core-processthreads-l1-1-1.dll	3/5/2024 7:59 PM	Application extens...	12 KB

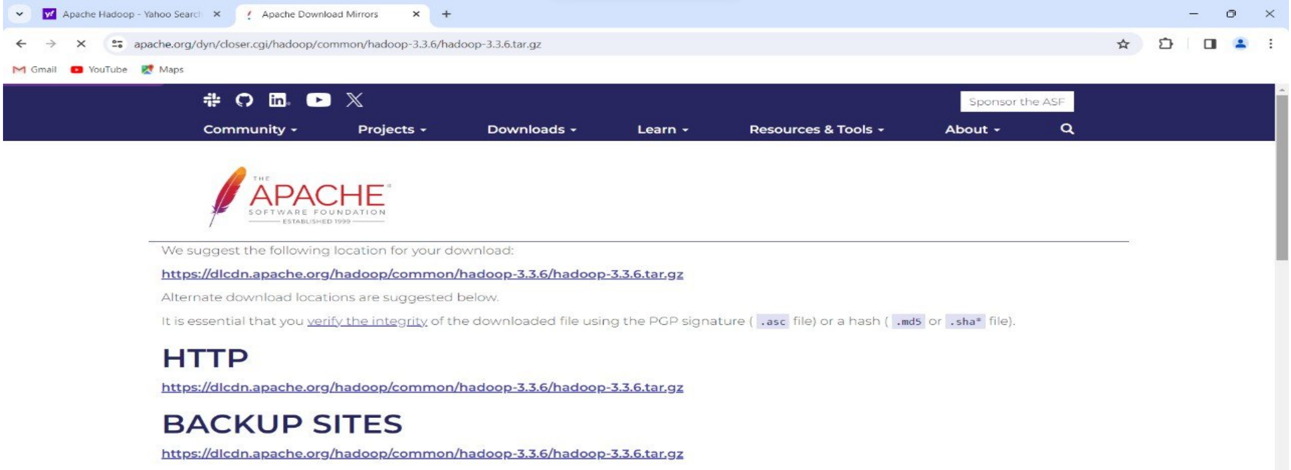


This PC > Windows-SSD (C:) > Java > jdk-21 > bin >

Name	Date modified	Type	Size
jabswitch	3/5/2024 7:59 PM	Application	45 KB
jaccessinspector	3/5/2024 7:59 PM	Application	105 KB
jaccesswalker	3/5/2024 7:59 PM	Application	70 KB
jar	3/5/2024 7:59 PM	Application	24 KB
jarsigner	3/5/2024 7:59 PM	Application	24 KB
java.dll	3/5/2024 7:59 PM	Application extens...	119 KB
java	3/5/2024 7:59 PM	Application	54 KB
javaaccessbridge.dll	3/5/2024 7:59 PM	Application extens...	285 KB
javac	3/5/2024 7:59 PM	Application	24 KB
javadoc	3/5/2024 7:59 PM	Application	24 KB
javajpeg.dll	3/5/2024 7:59 PM	Application extens...	178 KB
javap	3/5/2024 7:59 PM	Application	24 KB
javaw	3/5/2024 7:59 PM	Application	54 KB
jawt.dll	3/5/2024 7:59 PM	Application extens...	21 KB
jcmd	3/5/2024 7:59 PM	Application	24 KB
jconsole	3/5/2024 7:59 PM	Application	24 KB
jdb	3/5/2024 7:59 PM	Application	24 KB
jdeprscan	3/5/2024 7:59 PM	Application	24 KB
jdeps	3/5/2024 7:59 PM	Application	24 KB

STEP 4 – Install Hadoop

https://dlcdn.apache.org/hadoop/common/hadoop-3.3.6/hadoop-3.3.6.tar.gz



Apache Hadoop - Yahoo Search
Apache Download Mirrors
apache.org/dyn/closer.cgi/hadoop/common/hadoop-3.3.6/hadoop-3.3.6.tar.gz
Gmail
YouTube
Maps
Sponsor the ASF
Community
Projects
Downloads
Learn
Resources & Tools
About
THE APACHE SOFTWARE FOUNDATION ESTABLISHED 1999

We suggest the following location for your download:

https://dlcdn.apache.org/hadoop/common/hadoop-3.3.6/hadoop-3.3.6.tar.gz

Alternate download locations are suggested below.

It is essential that you verify the integrity of the downloaded file using the PGP signature (.asc file) or a hash (.md5 or .sha* file).

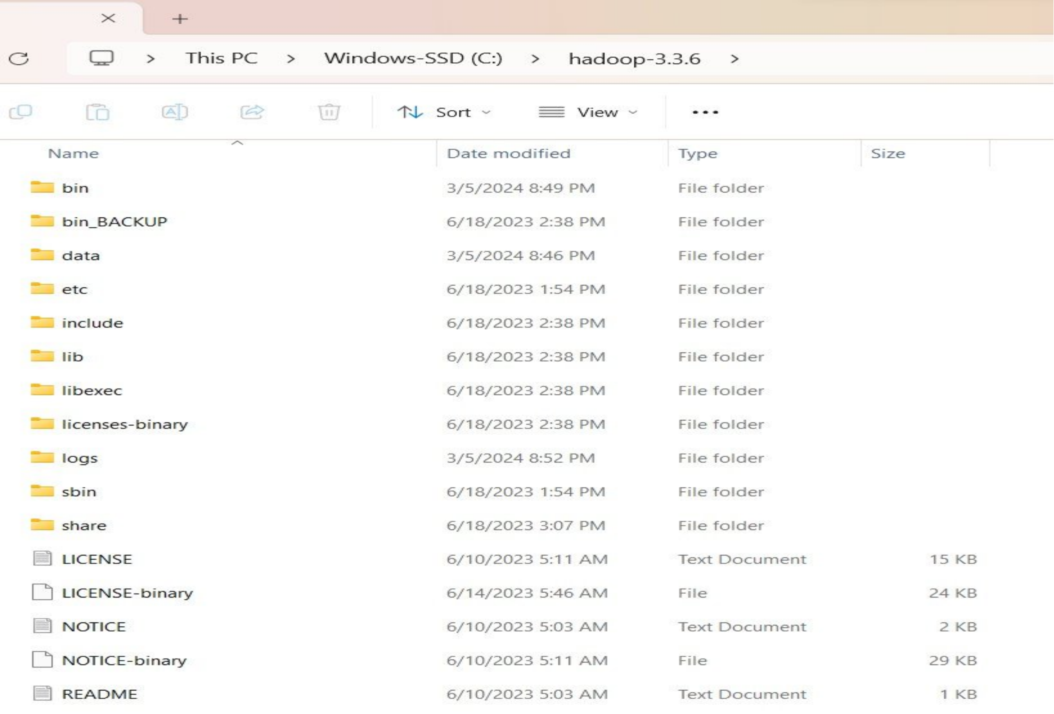
HTTP

https://dlcdn.apache.org/hadoop/common/hadoop-3.3.6/hadoop-3.3.6.tar.gz

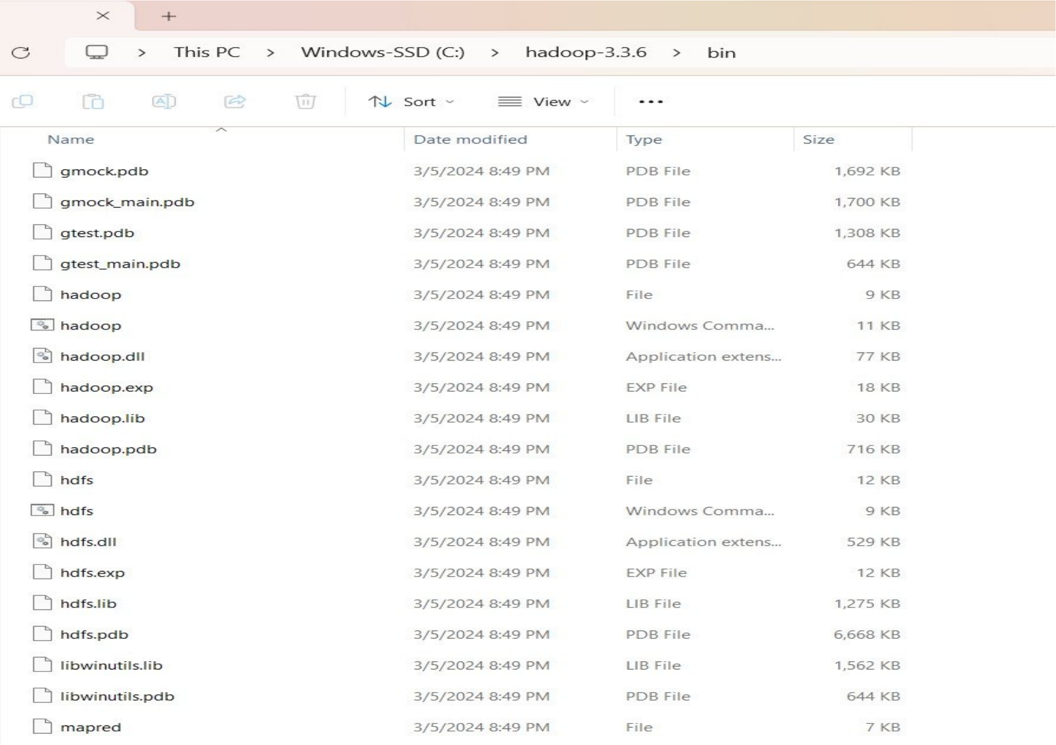
BACKUP SITES

https://dlcdn.apache.org/hadoop/common/hadoop-3.3.6/hadoop-3.3.6.tar.gz

STEP 5- Place the path of the bin folder into the Environment variable



Name	Date modified	Type	Size
bin	3/5/2024 8:49 PM	File folder	
bin_BACKUP	6/18/2023 2:38 PM	File folder	
data	3/5/2024 8:46 PM	File folder	
etc	6/18/2023 1:54 PM	File folder	
include	6/18/2023 2:38 PM	File folder	
lib	6/18/2023 2:38 PM	File folder	
libexec	6/18/2023 2:38 PM	File folder	
licenses-binary	6/18/2023 2:38 PM	File folder	
logs	3/5/2024 8:52 PM	File folder	
sbin	6/18/2023 1:54 PM	File folder	
share	6/18/2023 3:07 PM	File folder	
LICENSE	6/10/2023 5:11 AM	Text Document	15 KB
LICENSE-binary	6/14/2023 5:46 AM	File	24 KB
NOTICE	6/10/2023 5:03 AM	Text Document	2 KB
NOTICE-binary	6/10/2023 5:11 AM	File	29 KB
README	6/10/2023 5:03 AM	Text Document	1 KB



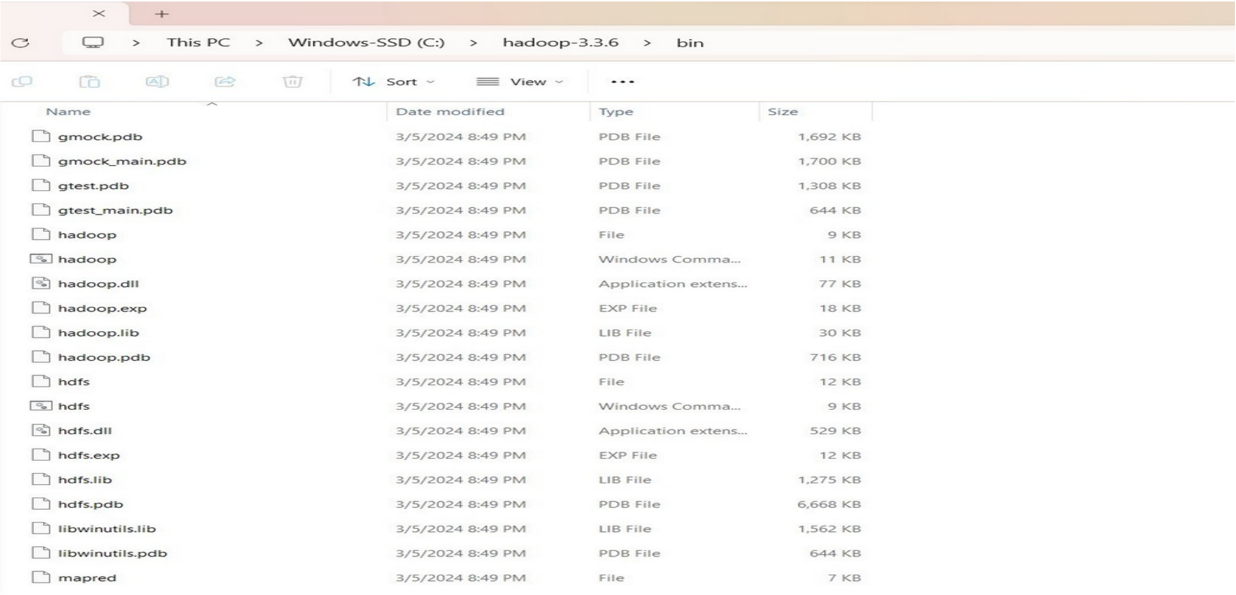
Name	Date modified	Type	Size
gmock.pdb	3/5/2024 8:49 PM	PDB File	1,692 KB
gmock_main.pdb	3/5/2024 8:49 PM	PDB File	1,700 KB
gtest.pdb	3/5/2024 8:49 PM	PDB File	1,308 KB
gtest_main.pdb	3/5/2024 8:49 PM	PDB File	644 KB
hadoop	3/5/2024 8:49 PM	File	9 KB
hadoop	3/5/2024 8:49 PM	Windows Comma...	11 KB
hadoop.dll	3/5/2024 8:49 PM	Application extens...	77 KB
hadoop.exp	3/5/2024 8:49 PM	EXP File	18 KB
hadoop.lib	3/5/2024 8:49 PM	LIB File	30 KB
hadoop.pdb	3/5/2024 8:49 PM	PDB File	716 KB
hdfs	3/5/2024 8:49 PM	File	12 KB
hdfs	3/5/2024 8:49 PM	Windows Comma...	9 KB
hdfs.dll	3/5/2024 8:49 PM	Application extens...	529 KB
hdfs.exp	3/5/2024 8:49 PM	EXP File	12 KB
hdfs.lib	3/5/2024 8:49 PM	LIB File	1,275 KB
hdfs.pdb	3/5/2024 8:49 PM	PDB File	6,668 KB
libwinutils.lib	3/5/2024 8:49 PM	LIB File	1,562 KB
libwinutils.pdb	3/5/2024 8:49 PM	PDB File	644 KB
mapred	3/5/2024 8:49 PM	File	7 KB

STEP 6- Perform configuration file setup step by step
There are 4 configuration files:

1. core-site.xml
2. mapred-site.xml
3. hdfs-site.xml
4. yarn-site.xml

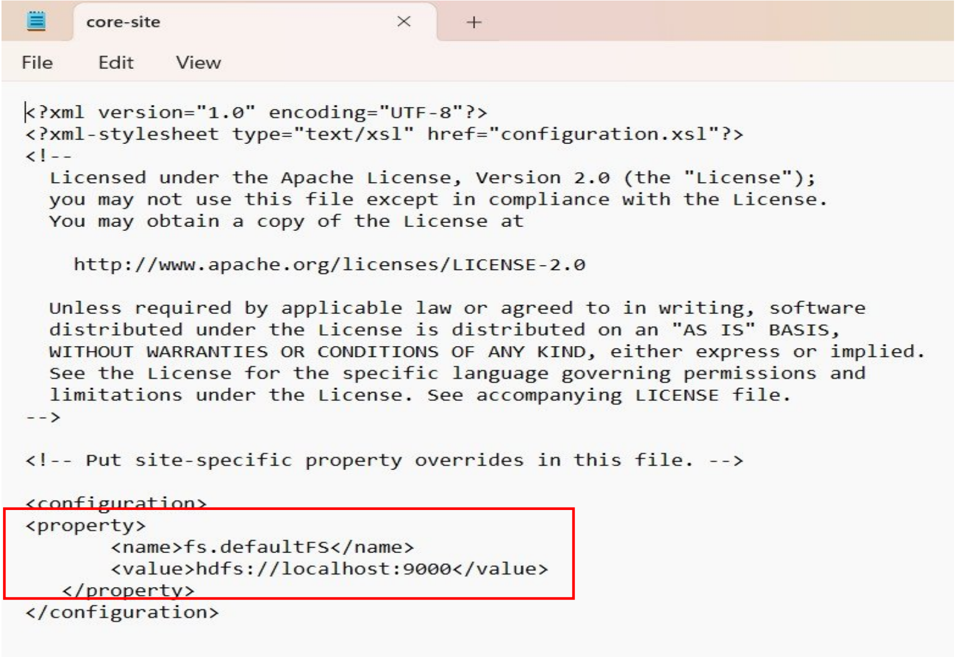
These files will be present inside the ETC folder:

For Hadoop Configuration, modify the above listed files:



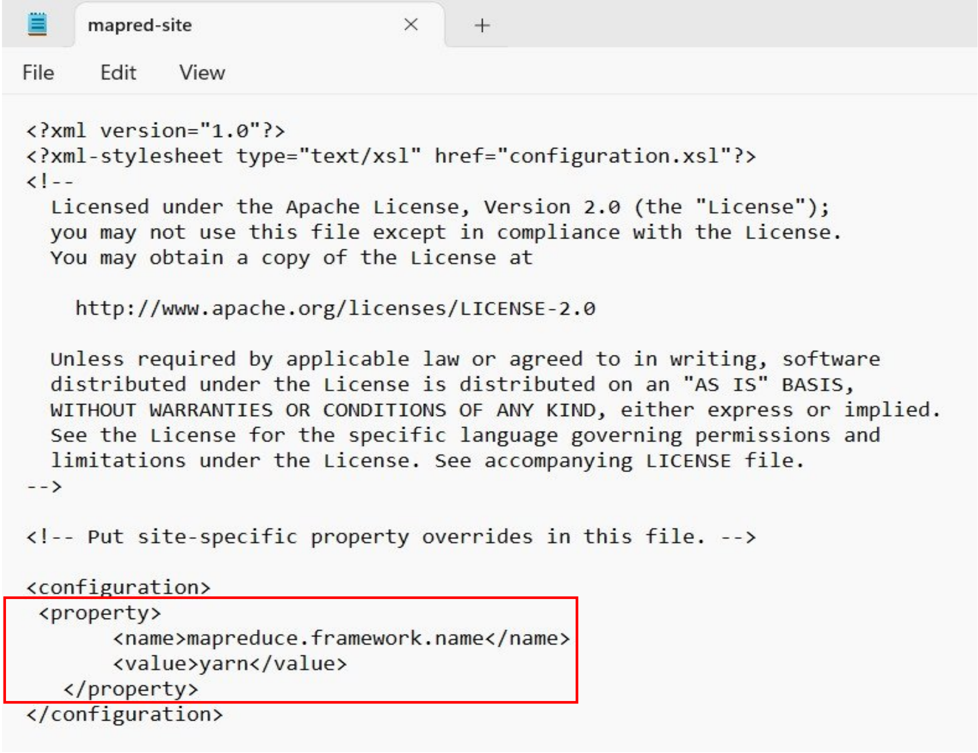
Name	Date modified	Type	Size
gmock.pdb	3/5/2024 8:49 PM	PDB File	1,692 KB
gmock_main.pdb	3/5/2024 8:49 PM	PDB File	1,700 KB
gtest.pdb	3/5/2024 8:49 PM	PDB File	1,308 KB
gtest_main.pdb	3/5/2024 8:49 PM	PDB File	644 KB
hadoop	3/5/2024 8:49 PM	File	9 KB
hadoop	3/5/2024 8:49 PM	Windows Comma...	11 KB
hadoop.dll	3/5/2024 8:49 PM	Application extens...	77 KB
hadoop.exp	3/5/2024 8:49 PM	EXP File	18 KB
hadoop.lib	3/5/2024 8:49 PM	LIB File	30 KB
hadoop.pdb	3/5/2024 8:49 PM	PDB File	716 KB
hdfs	3/5/2024 8:49 PM	File	12 KB
hdfs	3/5/2024 8:49 PM	Windows Comma...	9 KB
hdfs.dll	3/5/2024 8:49 PM	Application extens...	529 KB
hdfs.exp	3/5/2024 8:49 PM	EXP File	12 KB
hdfs.lib	3/5/2024 8:49 PM	LIB File	1,275 KB
hdfs.pdb	3/5/2024 8:49 PM	PDB File	6,668 KB
libwinutils.lib	3/5/2024 8:49 PM	LIB File	1,562 KB
libwinutils.pdb	3/5/2024 8:49 PM	PDB File	644 KB
mapred	3/5/2024 8:49 PM	File	7 KB

STEP 7- Edit the file core-site.xml



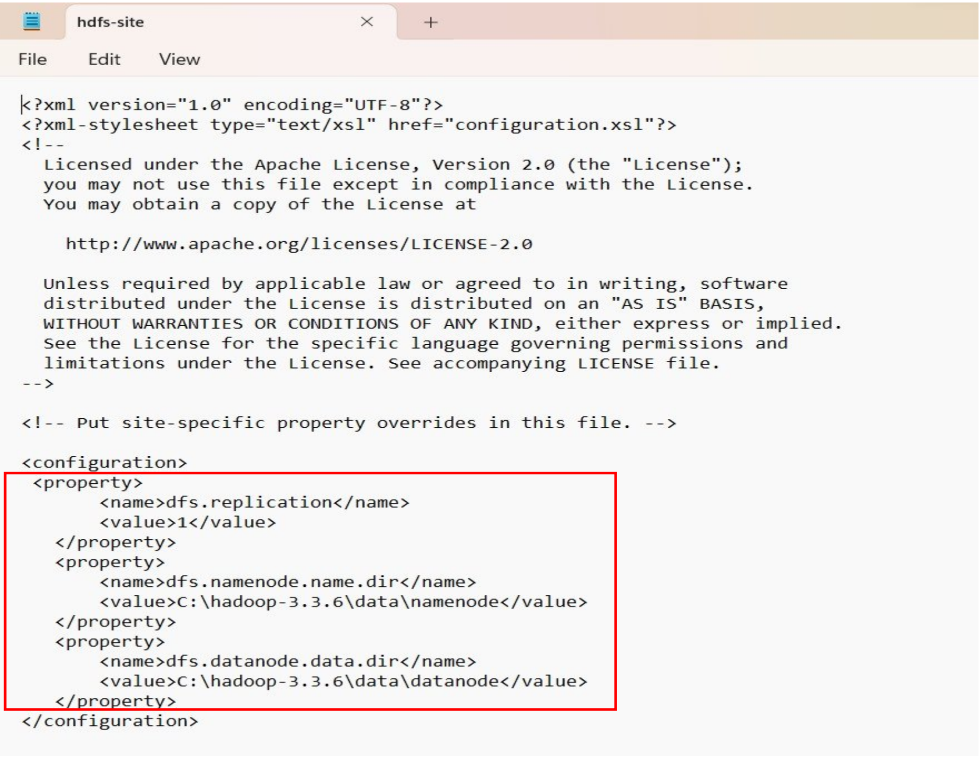
```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
  Licensed under the Apache License, Version 2.0 (the "License");
  you may not use this file except in compliance with the License.
  You may obtain a copy of the License at

    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>
```

STEP 8- Edit the file mapred-site.xml



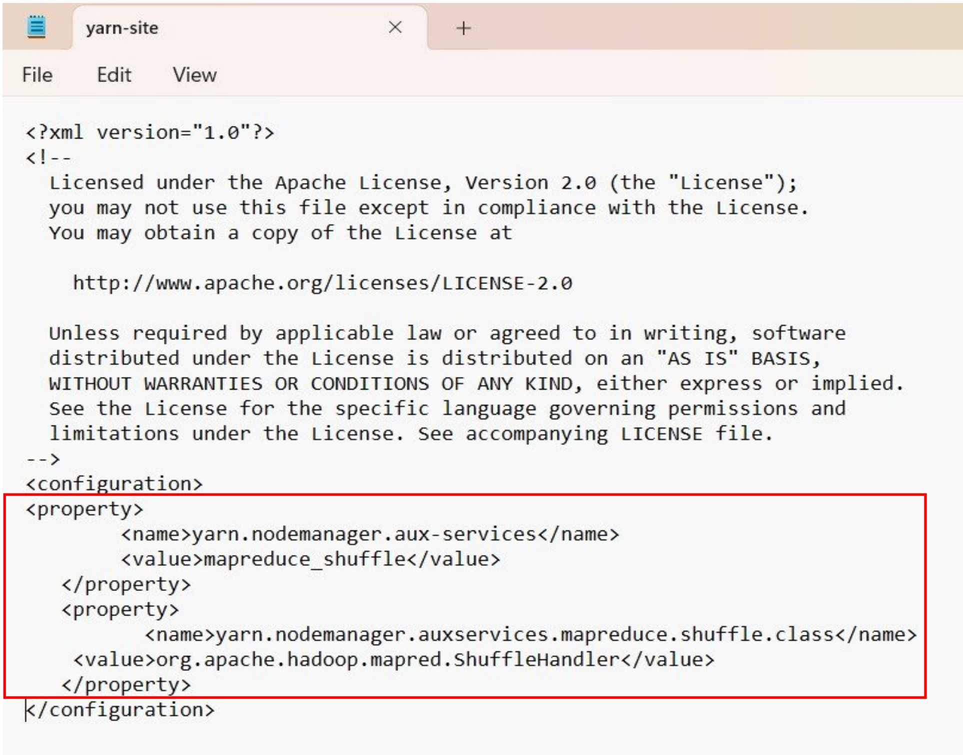
```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
  Licensed under the Apache License, Version 2.0 (the "License");
  you may not use this file except in compliance with the License.
  You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the License for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this file. -->

<configuration>
 <property>
       <name>mapreduce.framework.name</name>
       <value>yarn</value>
   </property>
</configuration>
```

STEP 9- Edit the file hdfs-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>dfs.replication</name>
       <value>1</value>
   </property>
   <property>
       <name>dfs.namenode.name.dir</name>
       <value>C:\hadoop-3.3.6\data\namenode</value>
   </property>
   <property>
       <name>dfs.datanode.data.dir</name>
       <value>C:\hadoop-3.3.6\data\datanode</value>
   </property>
</configuration>
```

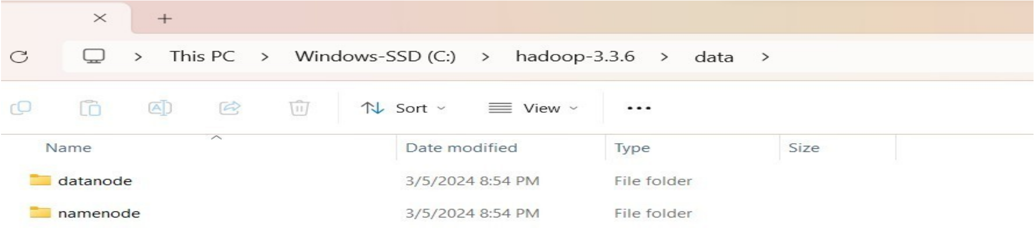
STEP 10- Edit the file yarn-site.xml:



```
<?xml version="1.0"?>
<!--
  Licensed under the Apache License, Version 2.0 (the "License");
  you may not use this file except in compliance with the License.
  You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the License for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->
<configuration>
<property>
        <name>yarn.nodemanager.aux-services</name>
        <value>mapreduce_shuffle</value>
   </property>
   <property>
          <name>yarn.nodemanager.auxservices.mapreduce.shuffle.class</name>
    <value>org.apache.hadoop.mapred.ShuffleHandler</value>
   </property>
</configuration>
```

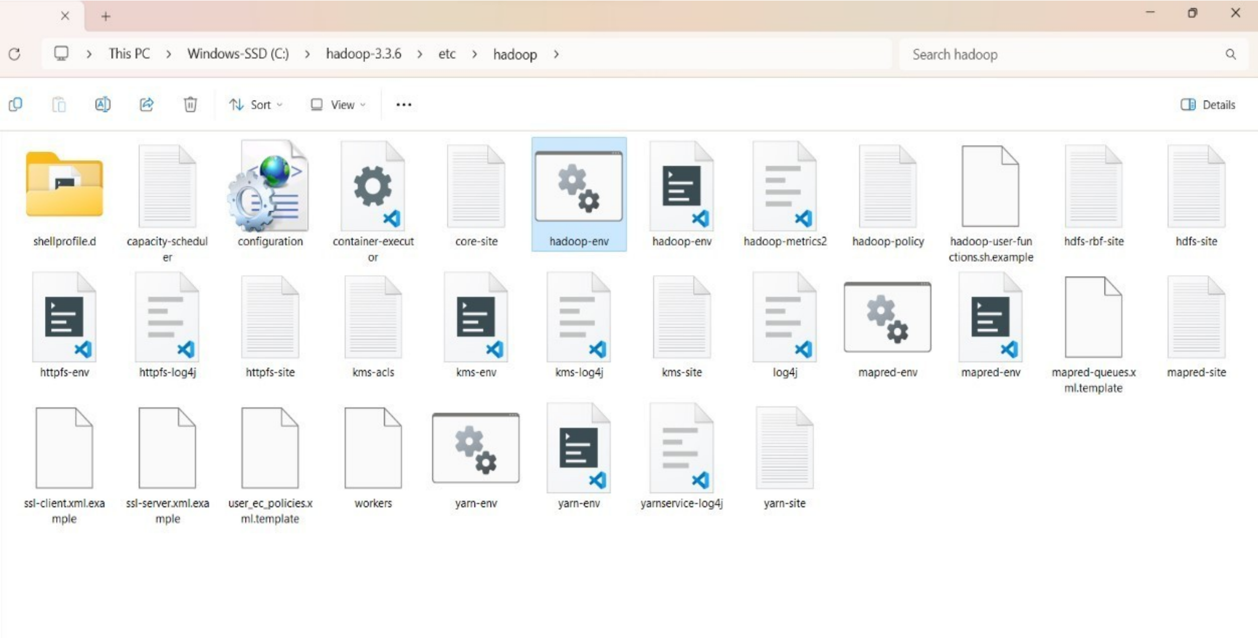
STEP 11- Create two folders:

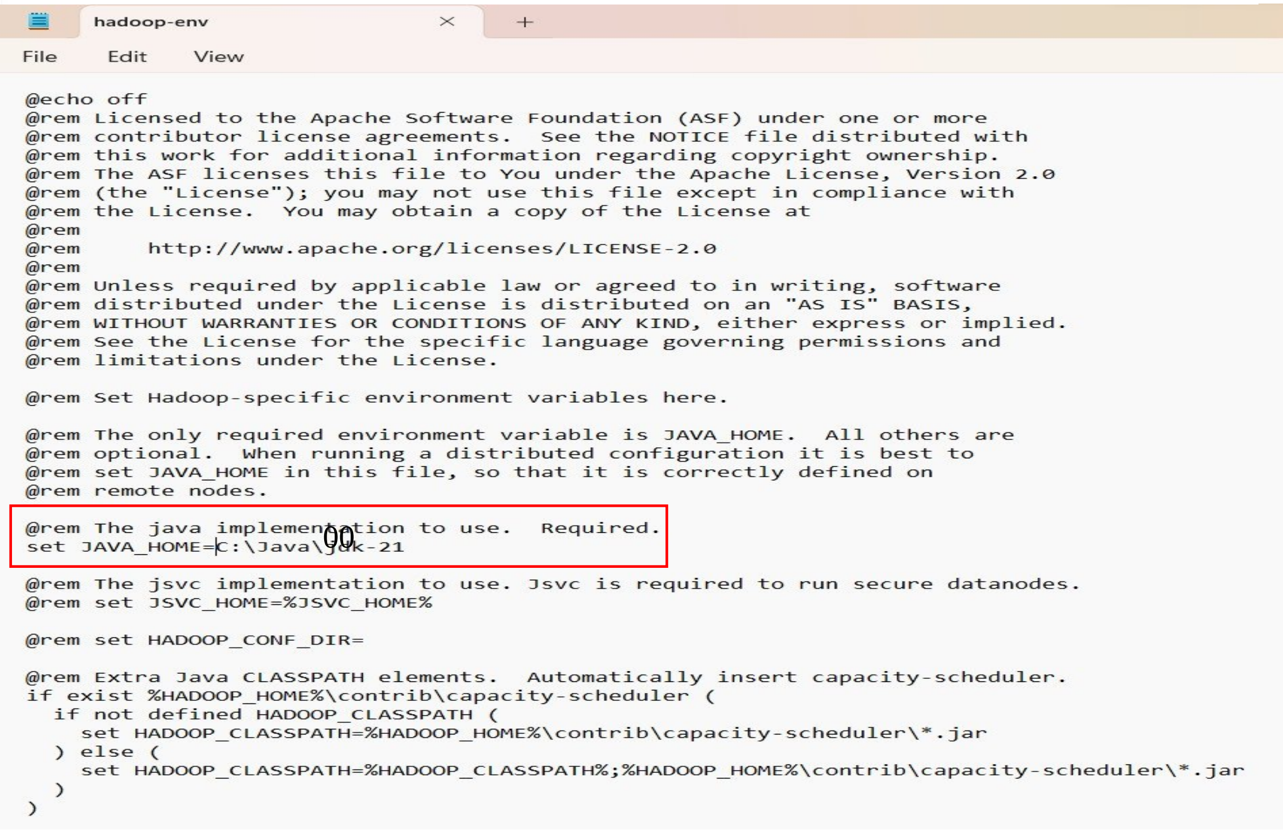
1. Name node
2. Data node



Name	Date modified	Type	Size
datanode	3/5/2024 8:54 PM	File folder	
namenode	3/5/2024 8:54 PM	File folder	

STEP 12- Open hadoop-env present in hadoop folder

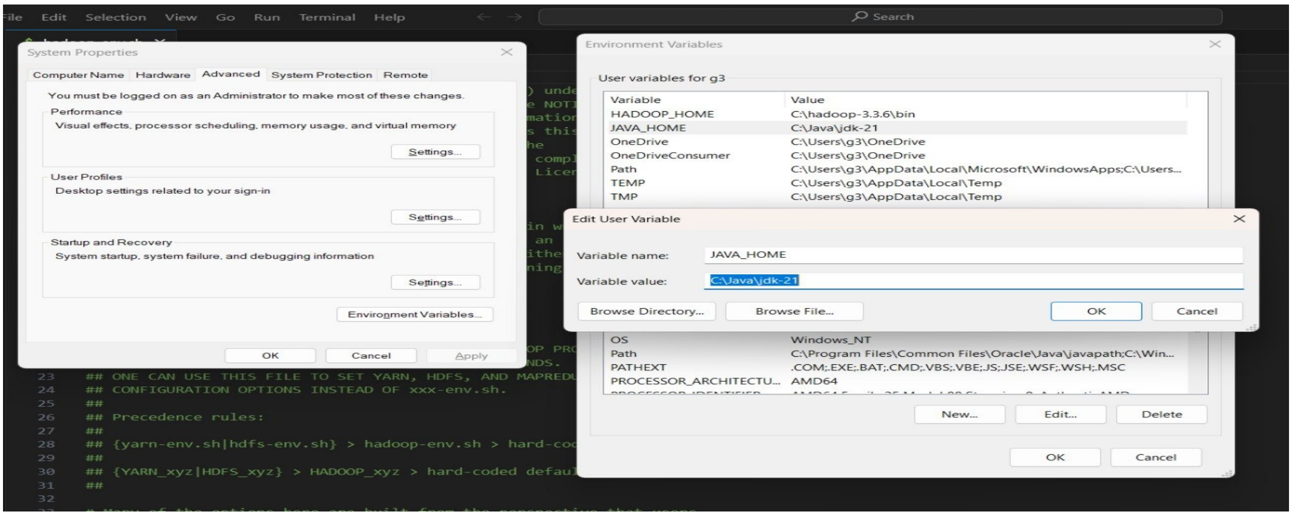




```
@echo off
@rem Licensed to the Apache Software Foundation (ASF) under one or more
@rem contributor license agreements.  See the NOTICE file distributed with
@rem this work for additional information regarding copyright ownership.
@rem The ASF licenses this file to You under the Apache License, Version 2.0
@rem (the "License"); you may not use this file except in compliance with
@rem the License.  You may obtain a copy of the License at
@rem
@rem     http://www.apache.org/licenses/LICENSE-2.0
@rem
@rem Unless required by applicable law or agreed to in writing, software
@rem distributed under the License is distributed on an "AS IS" BASIS,
@rem WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
@rem See the License for the specific language governing permissions and
@rem limitations under the License.

@rem Set Hadoop-specific environment variables here.

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

@rem The java implementation to use.  Required.
set JAVA_HOME=C:\Java\jdk-21

@rem The jsvc implementation to use. Jsvc is required to run secure datanodes.
@rem set JSVC_HOME=%JSVC_HOME%

@rem set HADOOP_CONF_DIR=

@rem Extra Java CLASSPATH elements.  Automatically insert capacity-scheduler.
if exist %HADOOP_HOME%\contrib\capacity-scheduler (
  if not defined HADOOP_CLASSPATH (
    set HADOOP_CLASSPATH=%HADOOP_HOME%\contrib\capacity-scheduler\*.jar
  ) else (
    set HADOOP_CLASSPATH=%HADOOP_CLASSPATH%;%HADOOP_HOME%\contrib\capacity-scheduler\*.jar
  )
)
```

STEP 13- Add the system variables

JAVA_HOME:


File Edit Selection View Go Run Terminal Help
Search
System Properties
Computer Name Hardware Advanced System Protection Remote
Settings...
Environment Variables...
OK Cancel Apply
Environment Variables
Edit User Variable
Browse Directory... Browse File... OK Cancel
New... Edit... Delete
OK Cancel

You must be logged on as an Administrator to make most of these changes.

Performance: Visual effects, processor scheduling, memory usage, and virtual memory

User Profiles: Desktop settings related to your sign-in

Startup and Recovery: System startup, system failure, and debugging information

User variables for g3

Variable	Value
HADOOP_HOME	C:\hadoop-3.3.6\bin
JAVA_HOME	C:\Java\jdk-21
OneDrive	C:\Users\g3\OneDrive
OneDriveConsumer	C:\Users\g3\OneDrive
Path	C:\Users\g3\AppData\Local\Microsoft\WindowsApps;C:\Users...
TEMP	C:\Users\g3\AppData\Local\Temp
TMP	C:\Users\g3\AppData\Local\Temp

Edit User Variable

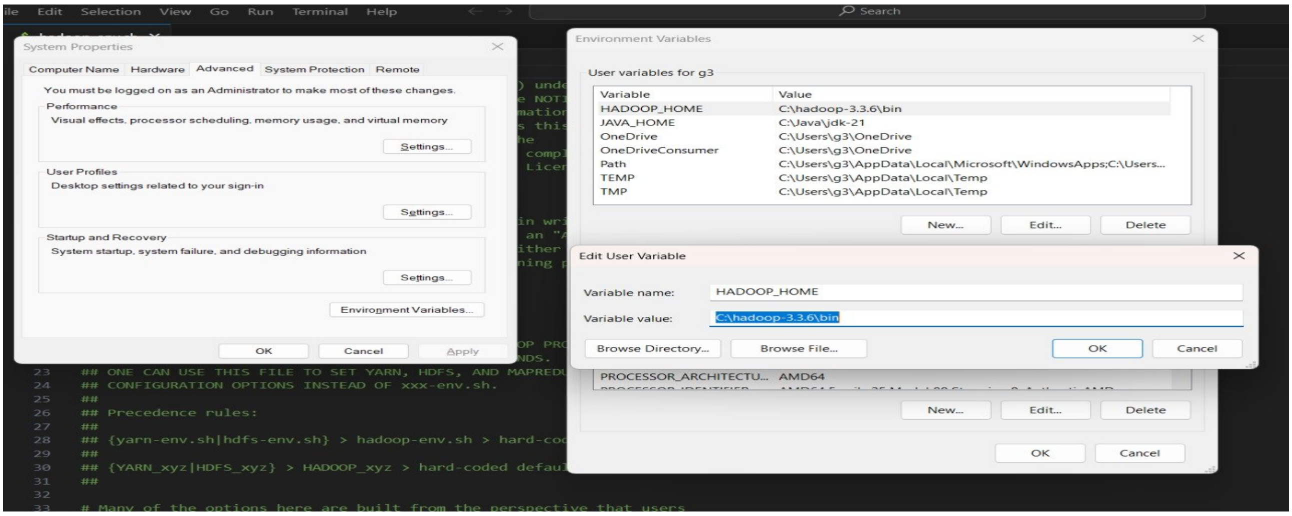
Variable name: JAVA_HOME

Variable value: C:\Java\jdk-21

Variable	Value
OS	Windows_NT
Path	C:\Program Files\Common Files\Oracle\Java\javapath;C:\Win...
PATHEXT	.COM;.EXE;.BAT;.CMD;.VBS;.VBE;.JS;.JSE;.WSF;.WSH;.MSC
PROCESSOR_ARCHITECTU...	AMD64

```
## ONE CAN USE THIS FILE TO SET YARN, HDFS, AND MAPRED
## CONFIGURATION OPTIONS INSTEAD OF xxx-env.sh.
##
## Precedence rules:
##
## {yarn-env.sh|hdfs-env.sh} > hadoop-env.sh > hard-co
##
## {YARN_xyz|HDFS_xyz} > HADOOP_xyz > hard-coded defau
##
```

HADOOP_HOME:


File Edit Selection View Go Run Terminal Help
Search
System Properties
Computer Name Hardware Advanced System Protection Remote
Settings...
Environment Variables...
OK Cancel Apply
Environment Variables
New... Edit... Delete
Edit User Variable
Browse Directory... Browse File... OK Cancel
OK Cancel

You must be logged on as an Administrator to make most of these changes.

Performance: Visual effects, processor scheduling, memory usage, and virtual memory

User Profiles: Desktop settings related to your sign-in

Startup and Recovery: System startup, system failure, and debugging information

User variables for g3

Variable	Value
HADOOP_HOME	C:\hadoop-3.3.6\bin
JAVA_HOME	C:\Java\jdk-21
OneDrive	C:\Users\g3\OneDrive
OneDriveConsumer	C:\Users\g3\OneDrive
Path	C:\Users\g3\AppData\Local\Microsoft\WindowsApps;C:\Users...
TEMP	C:\Users\g3\AppData\Local\Temp
TMP	C:\Users\g3\AppData\Local\Temp

Edit User Variable

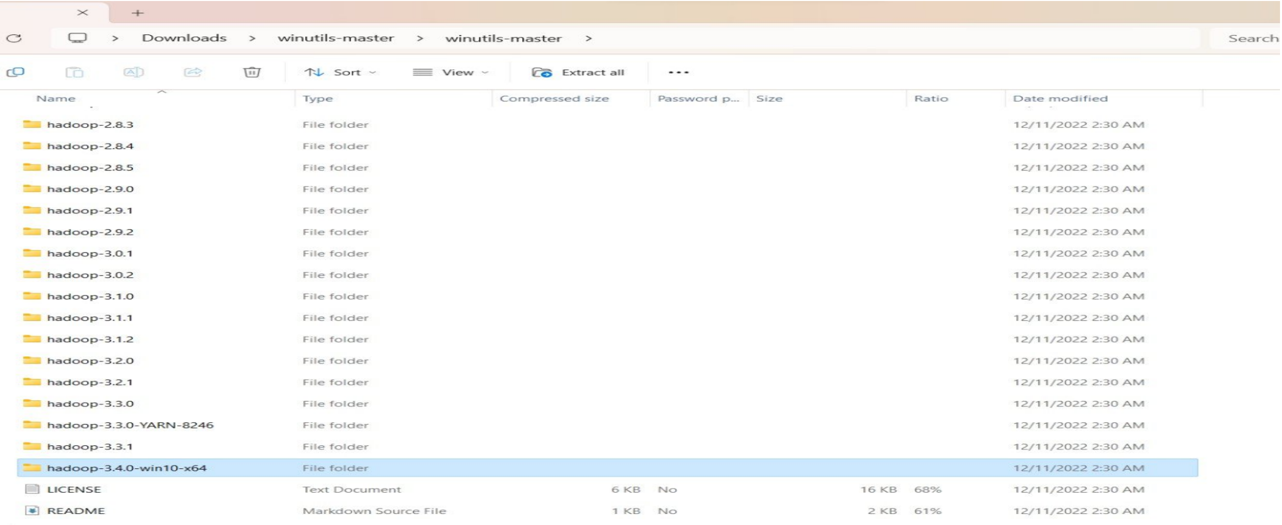
Variable name: HADOOP_HOME

Variable value: C:\hadoop-3.3.6\bin

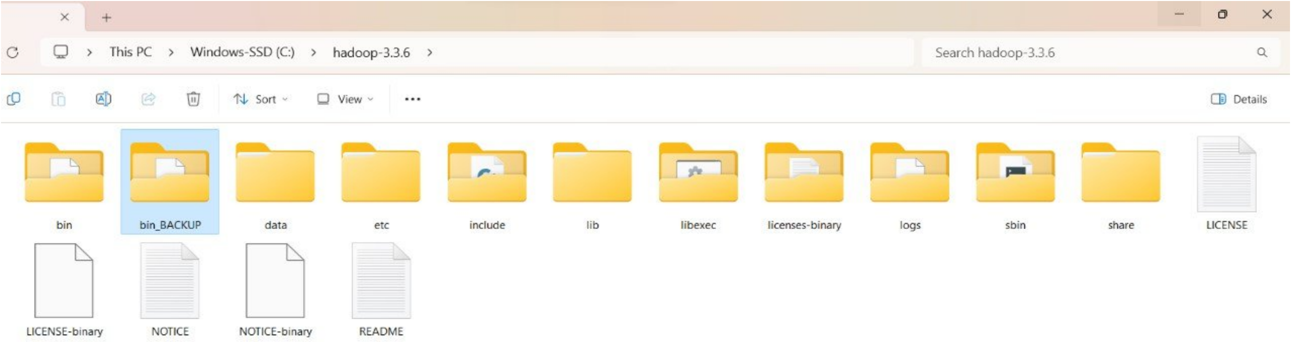
PROCESSOR_ARCHITECTU... AMD64

```
## ONE CAN USE THIS FILE TO SET YARN, HDFS, AND MAPRED
## CONFIGURATION OPTIONS INSTEAD OF xxx-env.sh.
##
## Precedence rules:
##
## {yarn-env.sh|hdfs-env.sh} > hadoop-env.sh > hard-co
##
## {YARN_xyz|HDFS_xyz} > HADOOP_xyz > hard-coded defau
##
# Many of the options here are built from the perspective that users
```

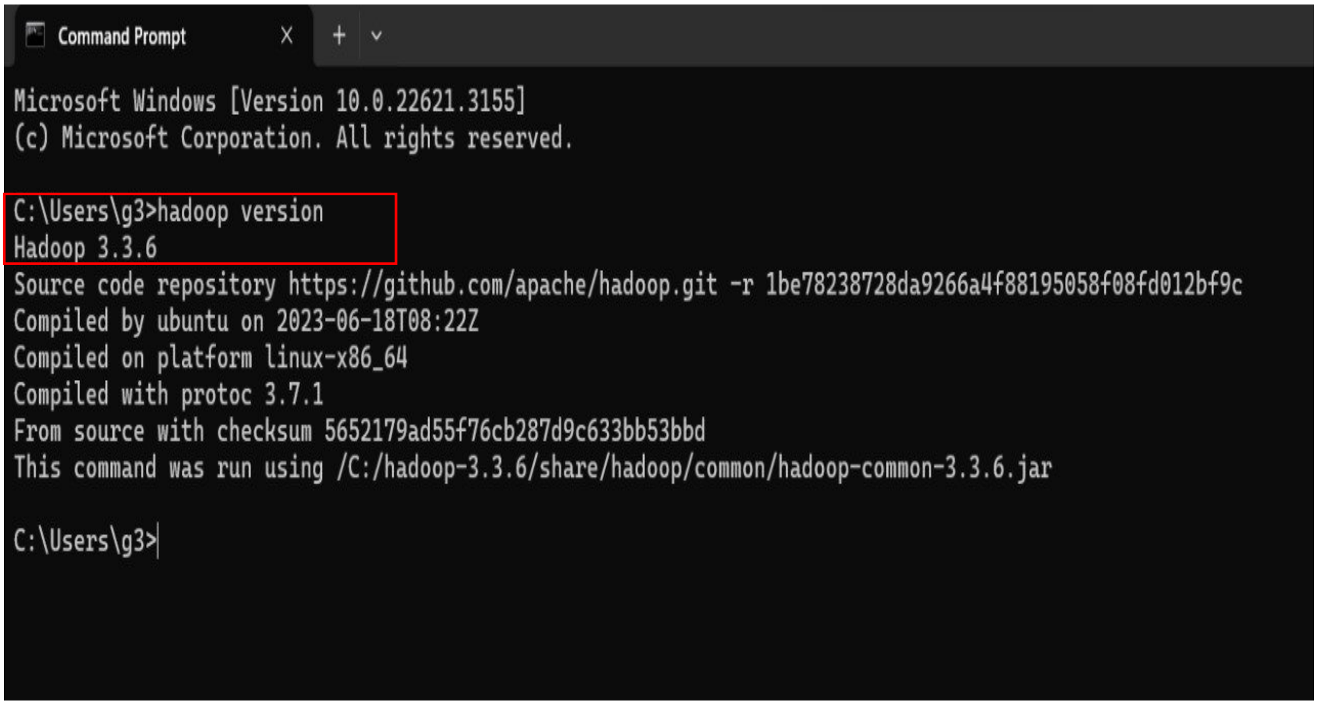
STEP 14- Download new WINUTILS and place it in the bin folder


Downloads > winutils-master > winutils-master >
Search
Sort View Extract all ···

Name	Type	Compressed size	Password p...	Size	Ratio	Date modified
hadoop-2.8.3	File folder					12/11/2022 2:30 AM
hadoop-2.8.4	File folder					12/11/2022 2:30 AM
hadoop-2.8.5	File folder					12/11/2022 2:30 AM
hadoop-2.9.0	File folder					12/11/2022 2:30 AM
hadoop-2.9.1	File folder					12/11/2022 2:30 AM
hadoop-2.9.2	File folder					12/11/2022 2:30 AM
hadoop-3.0.1	File folder					12/11/2022 2:30 AM
hadoop-3.0.2	File folder					12/11/2022 2:30 AM
hadoop-3.1.0	File folder					12/11/2022 2:30 AM
hadoop-3.1.1	File folder					12/11/2022 2:30 AM
hadoop-3.1.2	File folder					12/11/2022 2:30 AM
hadoop-3.2.0	File folder					12/11/2022 2:30 AM
hadoop-3.2.1	File folder					12/11/2022 2:30 AM
hadoop-3.3.0	File folder					12/11/2022 2:30 AM
hadoop-3.3.0-YARN-8246	File folder					12/11/2022 2:30 AM
hadoop-3.3.1	File folder					12/11/2022 2:30 AM
hadoop-3.4.0-win10-x64	File folder					12/11/2022 2:30 AM
LICENSE	Text Document	6 KB	No	16 KB	68%	12/11/2022 2:30 AM
README	Markdown Source File	1 KB	No	2 KB	61%	12/11/2022 2:30 AM

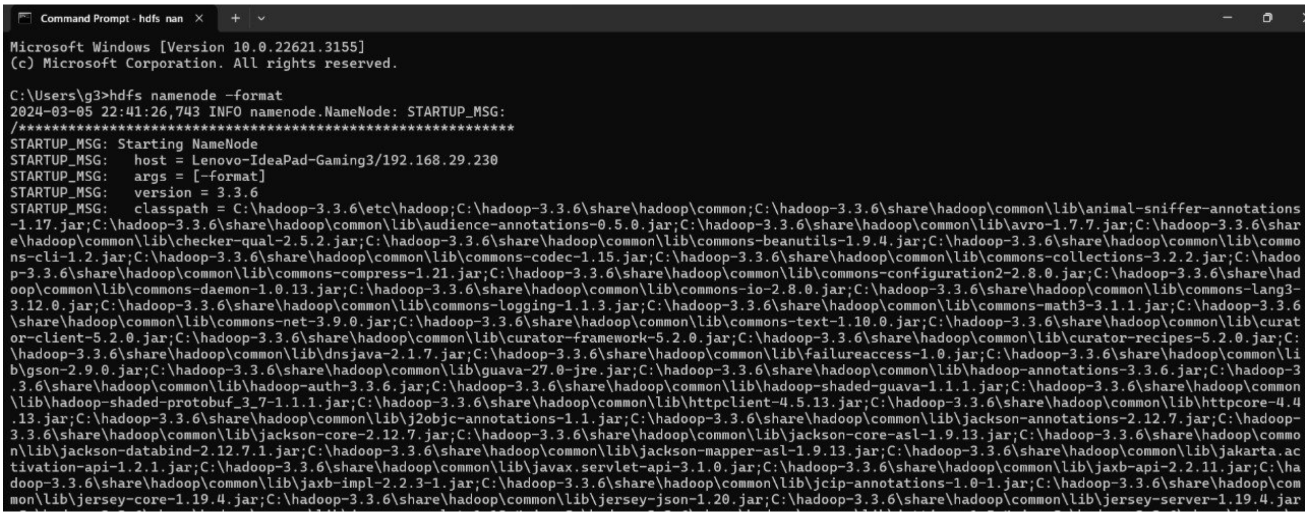


STEP 15- Test the hadoop installation



```
Microsoft Windows [Version 10.0.22621.3155]
(c) Microsoft Corporation. All rights reserved.

C:\Users\g3>hadoop version
Hadoop 3.3.6
Source code repository https://github.com/apache/hadoop.git -r 1be78238728da9266a4f88195058f08fd012bf9c
Compiled by ubuntu on 2023-06-18T08:22Z
Compiled on platform linux-x86_64
Compiled with protoc 3.7.1
From source with checksum 5652179ad55f76cb287d9c633bb53bbd
This command was run using /C:/hadoop-3.3.6/share/hadoop/common/hadoop-common-3.3.6.jar

C:\Users\g3>
```

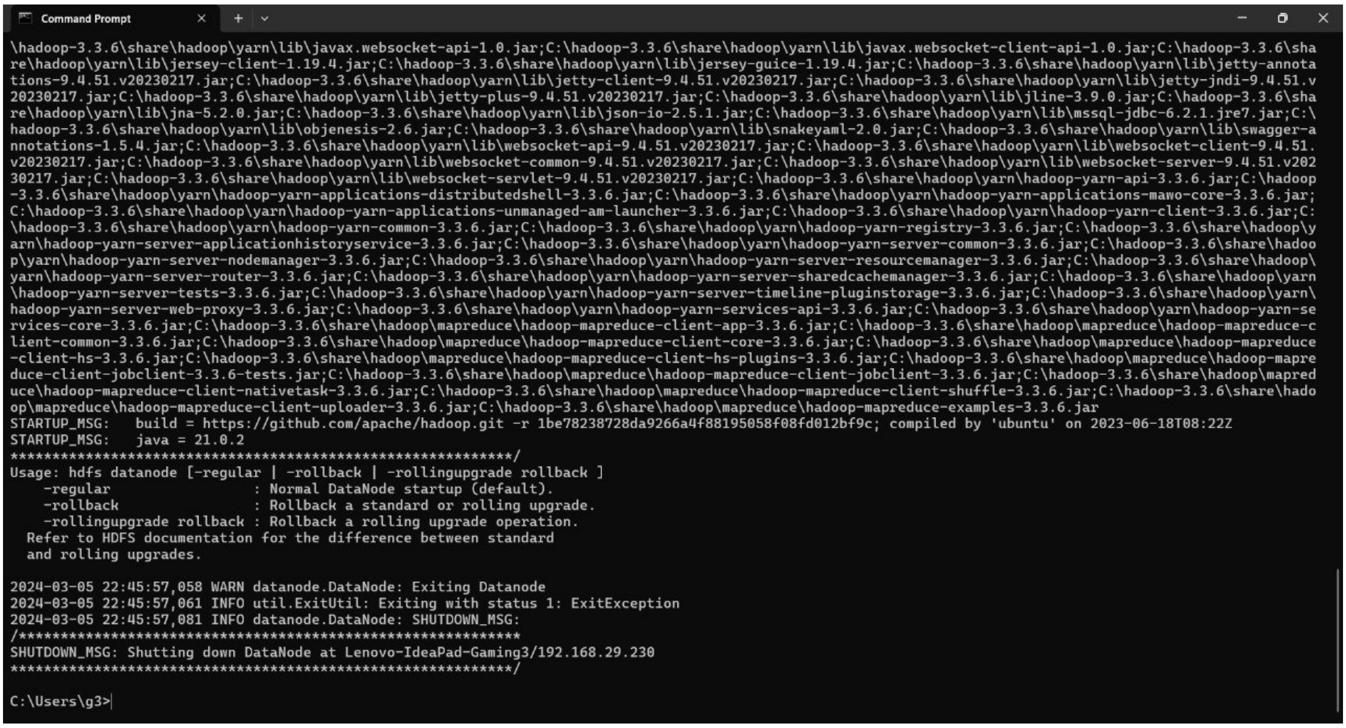
STEP 16- Format the namenode
Command: namenode -format



```
Microsoft Windows [Version 10.0.22621.3155]
(c) Microsoft Corporation. All rights reserved.

C:\Users\g3>hdfs namenode -format
2024-03-05 22:41:26,743 INFO namenode.NameNode: STARTUP_MSG:
/************************************************************
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = Lenovo-IdeaPad-Gaming3/192.168.29.230
STARTUP_MSG:   args = [-format]
STARTUP_MSG:   version = 3.3.6
STARTUP_MSG:   classpath = C:\hadoop-3.3.6\etc\hadoop;C:\hadoop-3.3.6\share\hadoop\common;C:\hadoop-3.3.6\share\hadoop\common\lib\animal-sniffer-annotations
-1.17.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\audience-annotations-0.5.0.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\avro-1.7.7.jar;C:\hadoop-3.3.6\shar
e\hadoop\common\lib\checker-qual-2.5.2.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\commons-beanutils-1.9.4.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\commo
ns-cli-1.2.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\commons-codec-1.15.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\commons-collections-3.2.2.jar;C:\hadoo
p-3.3.6\share\hadoop\common\lib\commons-compress-1.21.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\commons-configuration2-2.8.0.jar;C:\hadoop-3.3.6\share\had
oop\common\lib\commons-daemon-1.0.13.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\commons-io-2.8.0.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\commons-lang3-
3.12.0.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\commons-logging-1.1.3.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\commons-math3-3.1.1.jar;C:\hadoop-3.3.6
\share\hadoop\common\lib\commons-net-3.9.0.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\commons-text-1.10.0.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\curat
or-client-5.2.0.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\curator-framework-5.2.0.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\curator-recipes-5.2.0.jar;C:
\hadoop-3.3.6\share\hadoop\common\lib\dnsjava-2.1.7.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\failureaccess-1.0.jar;C:\hadoop-3.3.6\share\hadoop\common\li
b\gson-2.9.0.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\guava-27.0-jre.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\hadoop-annotations-3.3.6.jar;C:\hadoop-3
.3.6\share\hadoop\common\lib\hadoop-auth-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\hadoop-shaded-guava-1.1.1.jar;C:\hadoop-3.3.6\share\hadoop\common
\lib\hadoop-shaded-protobuf_3_7-1.1.1.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\httpclient-4.5.13.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\httpcore-4.4
.13.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\j2objc-annotations-1.1.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\jackson-annotations-2.12.7.jar;C:\hadoop-
3.3.6\share\hadoop\common\lib\jackson-core-2.12.7.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\jackson-core-asl-1.9.13.jar;C:\hadoop-3.3.6\share\hadoop\commo
n\lib\jackson-databind-2.12.7.1.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\jackson-mapper-asl-1.9.13.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\jakarta.ac
tivation-api-1.2.1.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\javax.servlet-api-3.1.0.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\jaxb-api-2.2.11.jar;C:\ha
doop-3.3.6\share\hadoop\common\lib\jaxb-impl-2.2.3-1.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\jcip-annotations-1.0-1.jar;C:\hadoop-3.3.6\share\hadoop\com
mon\lib\jersey-core-1.19.4.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\jersey-json-1.20.jar;C:\hadoop-3.3.6\share\hadoop\common\lib\jersey-server-1.19.4.jar
```

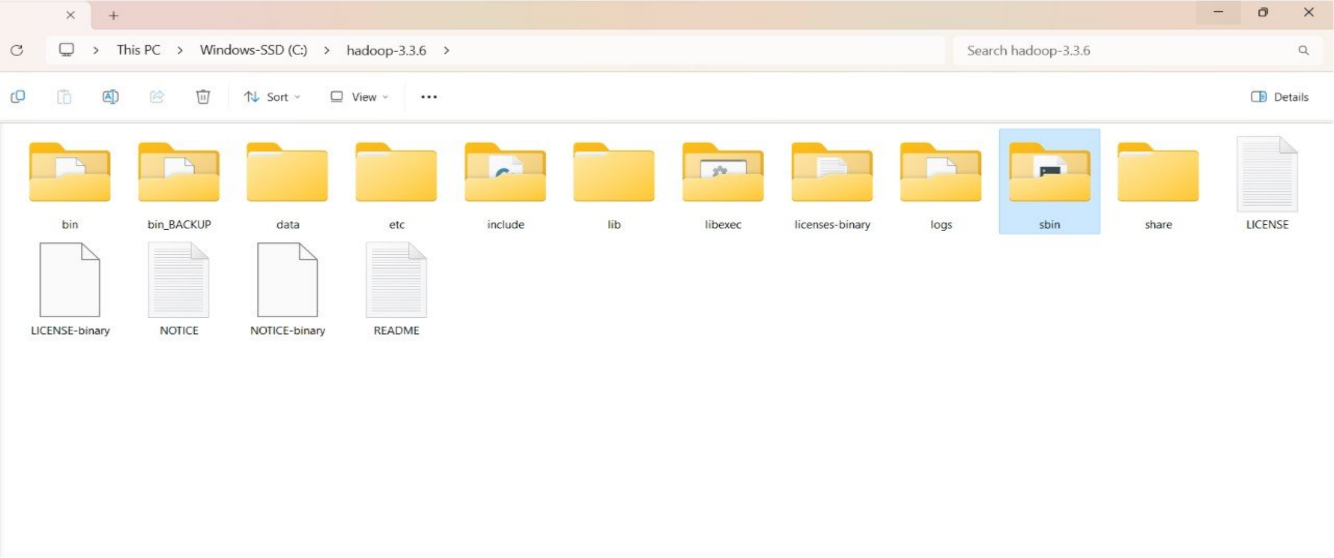
STEP 17- Format the datanode

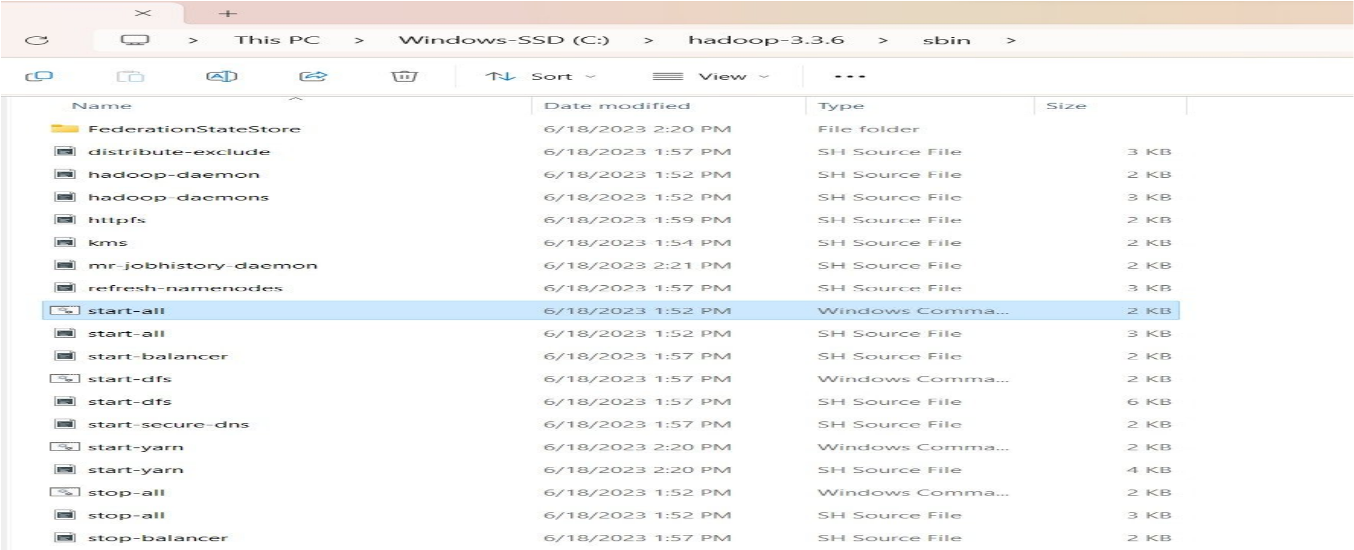
Command: datanode -format



```
\hadoop-3.3.6\share\hadoop\yarn\lib\javax.websocket-api-1.0.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\javax.websocket-client-api-1.0.jar;C:\hadoop-3.3.6\sha
re\hadoop\yarn\lib\jersey-client-1.19.4.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\jersey-guice-1.19.4.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\jetty-annota
tions-9.4.51.v20230217.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\jetty-client-9.4.51.v20230217.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\jetty-jndi-9.4.51.v
20230217.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\jetty-plus-9.4.51.v20230217.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\jline-3.9.0.jar;C:\hadoop-3.3.6\sha
re\hadoop\yarn\lib\jna-5.2.0.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\json-io-2.5.1.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\mssql-jdbc-6.2.1.jre7.jar;C:\
hadoop-3.3.6\share\hadoop\yarn\lib\objenesis-2.6.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\snakeyaml-2.0.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\swagger-a
nnotations-1.5.4.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\websocket-api-9.4.51.v20230217.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\websocket-client-9.4.51.
v20230217.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\websocket-common-9.4.51.v20230217.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\websocket-server-9.4.51.v202
30217.jar;C:\hadoop-3.3.6\share\hadoop\yarn\lib\websocket-servlet-9.4.51.v20230217.jar;C:\hadoop-3.3.6\share\hadoop\yarn\hadoop-yarn-api-3.3.6.jar;C:\hadoop
-3.3.6\share\hadoop\yarn\hadoop-yarn-applications-distributedshell-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\yarn\hadoop-yarn-applications-mawo-core-3.3.6.jar;
C:\hadoop-3.3.6\share\hadoop\yarn\hadoop-yarn-applications-unmanaged-am-launcher-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\yarn\hadoop-yarn-client-3.3.6.jar;C:
\hadoop-3.3.6\share\hadoop\yarn\hadoop-yarn-common-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\yarn\hadoop-yarn-registry-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\y
arn\hadoop-yarn-server-applicationhistoryservice-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\yarn\hadoop-yarn-server-common-3.3.6.jar;C:\hadoop-3.3.6\share\hadoo
p\yarn\hadoop-yarn-server-nodemanager-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\yarn\hadoop-yarn-server-resourcemanager-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\
yarn\hadoop-yarn-server-router-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\yarn\hadoop-yarn-server-sharedcachemanager-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\yarn
\hadoop-yarn-server-tests-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\yarn\hadoop-yarn-server-timeline-pluginstorage-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\yarn\
hadoop-yarn-server-web-proxy-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\yarn\hadoop-yarn-services-api-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\yarn\hadoop-yarn-se
rvices-core-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\mapreduce\hadoop-mapreduce-client-app-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\mapreduce\hadoop-mapreduce-c
lient-common-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\mapreduce\hadoop-mapreduce-client-core-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\mapreduce\hadoop-mapreduce
-client-hs-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\mapreduce\hadoop-mapreduce-client-hs-plugins-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\mapreduce\hadoop-mapre
duce-client-jobclient-3.3.6-tests.jar;C:\hadoop-3.3.6\share\hadoop\mapreduce\hadoop-mapreduce-client-jobclient-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\mapred
uce\hadoop-mapreduce-client-nativetask-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\mapreduce\hadoop-mapreduce-client-shuffle-3.3.6.jar;C:\hadoop-3.3.6\share\hado
op\mapreduce\hadoop-mapreduce-client-uploader-3.3.6.jar;C:\hadoop-3.3.6\share\hadoop\mapreduce\hadoop-mapreduce-examples-3.3.6.jar
STARTUP_MSG:   build = https://github.com/apache/hadoop.git -r 1be78238728da9266a4f88195058f08fd012bf9c; compiled by 'ubuntu' on 2023-06-18T08:22Z
STARTUP_MSG:   java = 21.0.2
************************************************************/
Usage: hdfs datanode [-regular | -rollback | -rollingupgrade rollback ]
    -regular                 : Normal DataNode startup (default).
    -rollback                : Rollback a standard or rolling upgrade.
    -rollingupgrade rollback : Rollback a rolling upgrade operation.
  Refer to HDFS documentation for the difference between standard
  and rolling upgrades.

2024-03-05 22:45:57,058 WARN datanode.DataNode: Exiting Datanode
2024-03-05 22:45:57,061 INFO util.ExitUtil: Exiting with status 1: ExitException
2024-03-05 22:45:57,081 INFO datanode.DataNode: SHUTDOWN_MSG:
/************************************************************
SHUTDOWN_MSG: Shutting down DataNode at Lenovo-IdeaPad-Gaming3/192.168.29.230
************************************************************/

C:\Users\g3>
```

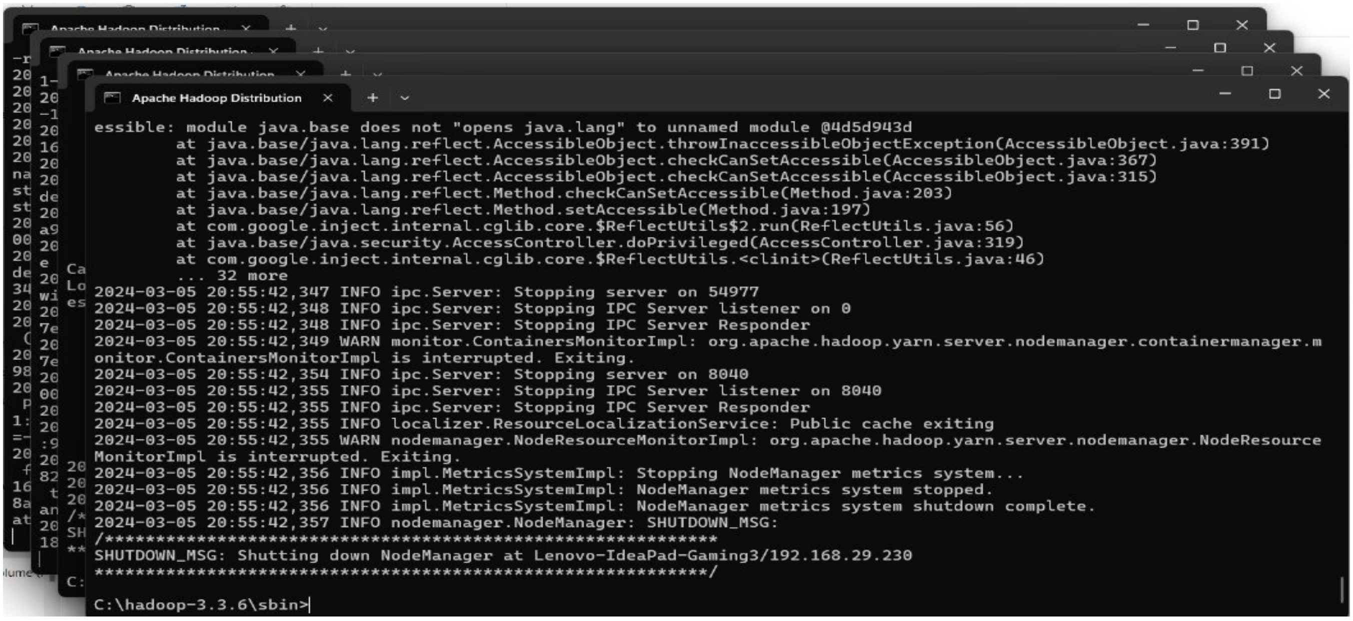
STEP 18- Start the Hadoop daemons. Open sbin dir from hadoop dir. Execute start-all





Name	Date modified	Type	Size
FederationStateStore	6/18/2023 2:20 PM	File folder	
distribute-exclude	6/18/2023 1:57 PM	SH Source File	3 KB
hadoop-daemon	6/18/2023 1:52 PM	SH Source File	2 KB
hadoop-daemons	6/18/2023 1:52 PM	SH Source File	3 KB
httpfs	6/18/2023 1:59 PM	SH Source File	2 KB
kms	6/18/2023 1:54 PM	SH Source File	2 KB
mr-jobhistory-daemon	6/18/2023 2:21 PM	SH Source File	2 KB
refresh-namenodes	6/18/2023 1:57 PM	SH Source File	3 KB
start-all	6/18/2023 1:52 PM	Windows Comma...	2 KB
start-all	6/18/2023 1:52 PM	SH Source File	3 KB
start-balancer	6/18/2023 1:57 PM	SH Source File	2 KB
start-dfs	6/18/2023 1:57 PM	Windows Comma...	2 KB
start-dfs	6/18/2023 1:57 PM	SH Source File	6 KB
start-secure-dns	6/18/2023 1:57 PM	SH Source File	2 KB
start-yarn	6/18/2023 2:20 PM	Windows Comma...	2 KB
start-yarn	6/18/2023 2:20 PM	SH Source File	4 KB
stop-all	6/18/2023 1:52 PM	Windows Comma...	2 KB
stop-all	6/18/2023 1:52 PM	SH Source File	3 KB
stop-balancer	6/18/2023 1:57 PM	SH Source File	2 KB

Start-all: (4 windows will be opened)



```
essible: module java.base does not "opens java.lang" to unnamed module @4d5d943d
        at java.base/java.lang.reflect.AccessibleObject.throwInaccessibleObjectException(AccessibleObject.java:391)
        at java.base/java.lang.reflect.AccessibleObject.checkCanSetAccessible(AccessibleObject.java:367)
        at java.base/java.lang.reflect.AccessibleObject.checkCanSetAccessible(AccessibleObject.java:315)
        at java.base/java.lang.reflect.Method.checkCanSetAccessible(Method.java:203)
        at java.base/java.lang.reflect.Method.setAccessible(Method.java:197)
        at com.google.inject.internal.cglib.core.$ReflectUtils$2.run(ReflectUtils.java:56)
        at java.base/java.security.AccessController.doPrivileged(AccessController.java:319)
        at com.google.inject.internal.cglib.core.$ReflectUtils.<clinit>(ReflectUtils.java:46)
        ... 32 more
2024-03-05 20:55:42,347 INFO ipc.Server: Stopping server on 54977
2024-03-05 20:55:42,348 INFO ipc.Server: Stopping IPC Server listener on 0
2024-03-05 20:55:42,348 INFO ipc.Server: Stopping IPC Server Responder
2024-03-05 20:55:42,349 WARN monitor.ContainersMonitorImpl: org.apache.hadoop.yarn.server.nodemanager.containermanager.m
onitor.ContainersMonitorImpl is interrupted. Exiting.
2024-03-05 20:55:42,354 INFO ipc.Server: Stopping server on 8040
2024-03-05 20:55:42,355 INFO ipc.Server: Stopping IPC Server listener on 8040
2024-03-05 20:55:42,355 INFO ipc.Server: Stopping IPC Server Responder
2024-03-05 20:55:42,355 INFO localizer.ResourceLocalizationService: Public cache exiting
2024-03-05 20:55:42,355 WARN nodemanager.NodeResourceMonitorImpl: org.apache.hadoop.yarn.server.nodemanager.NodeResource
MonitorImpl is interrupted. Exiting.
2024-03-05 20:55:42,356 INFO impl.MetricsSystemImpl: Stopping NodeManager metrics system...
2024-03-05 20:55:42,356 INFO impl.MetricsSystemImpl: NodeManager metrics system stopped.
2024-03-05 20:55:42,356 INFO impl.MetricsSystemImpl: NodeManager metrics system shutdown complete.
2024-03-05 20:55:42,357 INFO nodemanager.NodeManager: SHUTDOWN_MSG:
/************************************************************
SHUTDOWN_MSG: Shutting down NodeManager at Lenovo-IdeaPad-Gaming3/192.168.29.230
************************************************************/

C:\hadoop-3.3.6\sbin>
```

RESULT:

Thus the one node hadoop cluster has been set up successfully.

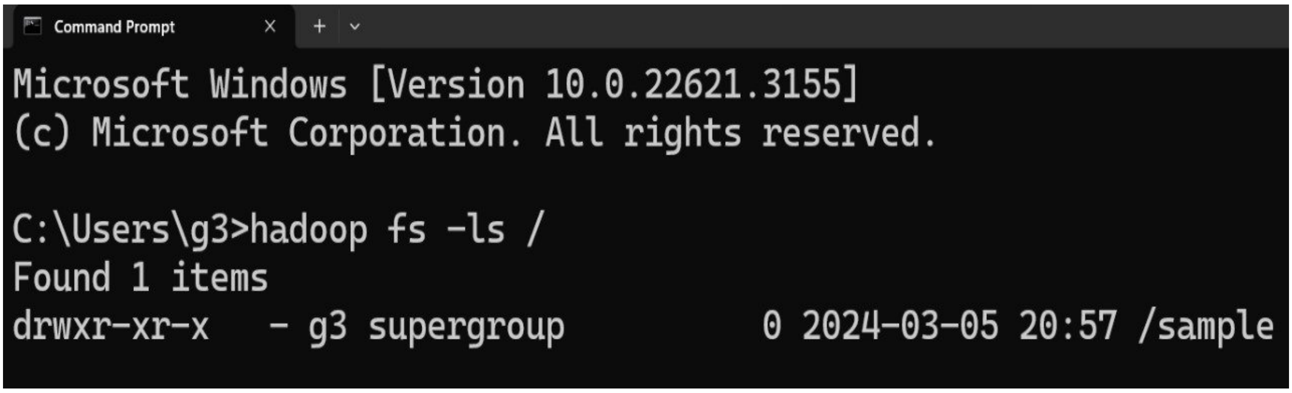
EX NO: 8 **FILE MANAGEMENT IN HADOOP**

DATE:

AIM: To execute file management commands in hadoop.

1. Get the list of directories and files at the root of HDFS

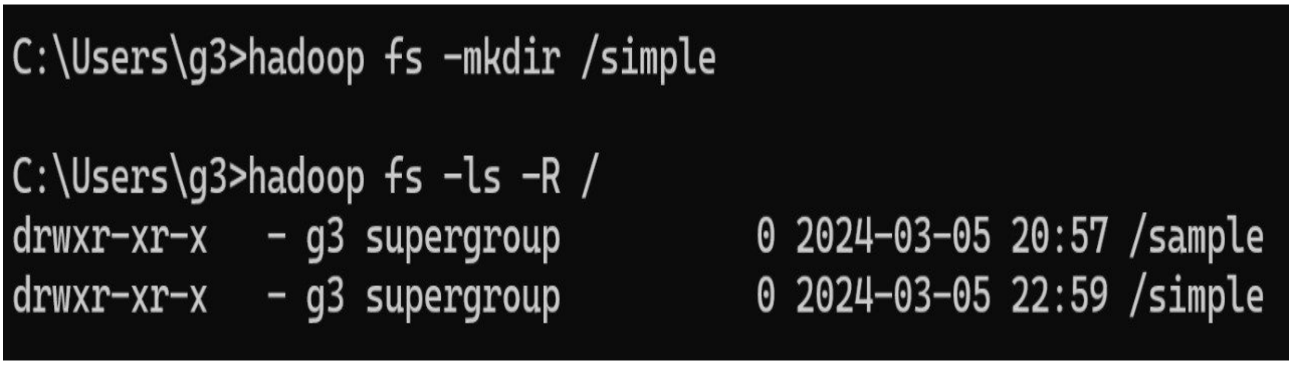
Command: hadoop fs -ls/



```
Microsoft Windows [Version 10.0.22621.3155]
(c) Microsoft Corporation. All rights reserved.

C:\Users\g3>hadoop fs -ls /
Found 1 items
drwxr-xr-x   - g3 supergroup          0 2024-03-05 20:57 /sample
```

2. Create a directory in HDFS

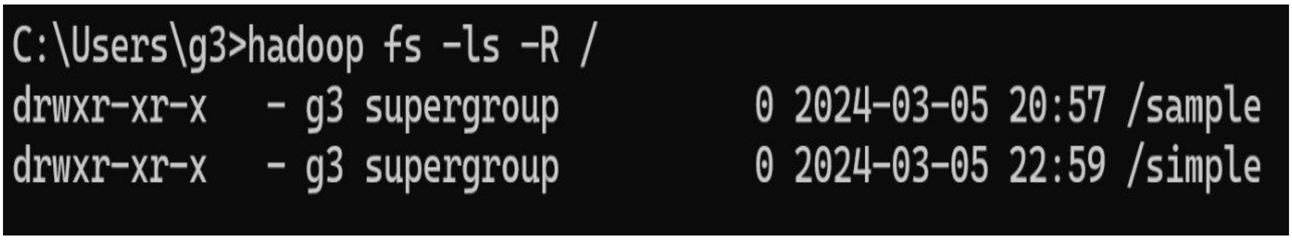
Command: hadoop fs -mkdir /sample



```
C:\Users\g3>hadoop fs -mkdir /simple

C:\Users\g3>hadoop fs -ls -R /
drwxr-xr-x   - g3 supergroup          0 2024-03-05 20:57 /sample
drwxr-xr-x   - g3 supergroup          0 2024-03-05 22:59 /simple
```

3. Get the list of complete directories and files of HDFS

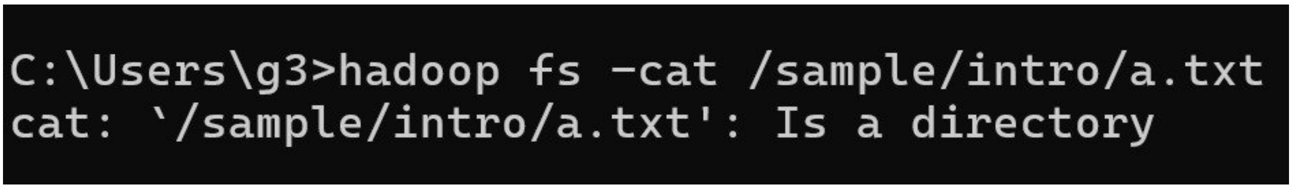
Command: hadoop fs -ls -R/



```
C:\Users\g3>hadoop fs -ls -R /
drwxr-xr-x   - g3 supergroup          0 2024-03-05 20:57 /sample
drwxr-xr-x   - g3 supergroup          0 2024-03-05 22:59 /simple
```

4. Display the contents of an HDFS file on console

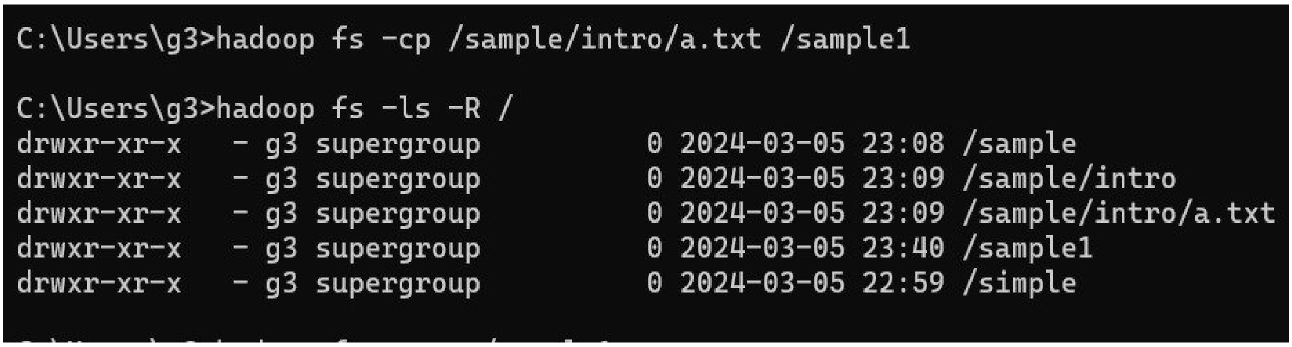
Command: hadoop fs -cat /sample/test.txt



```
C:\Users\g3>hadoop fs -cat /sample/intro/a.txt
cat: `/sample/intro/a.txt': Is a directory
```

5. Copy a file from one directory to another on HDFS

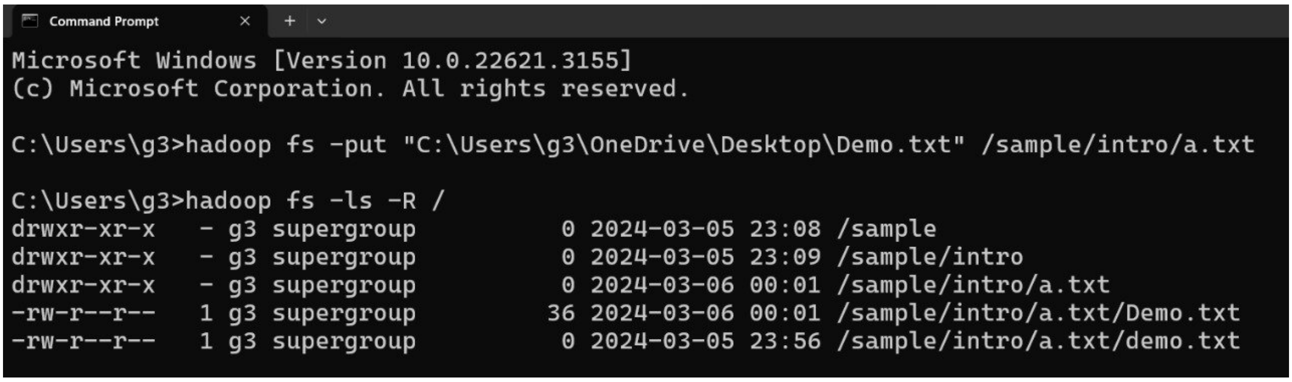
Command: hadoop fs -cp /sample/test.txt /sample1



```
C:\Users\g3>hadoop fs -cp /sample/intro/a.txt /sample1

C:\Users\g3>hadoop fs -ls -R /
drwxr-xr-x   - g3 supergroup          0 2024-03-05 23:08 /sample
drwxr-xr-x   - g3 supergroup          0 2024-03-05 23:09 /sample/intro
drwxr-xr-x   - g3 supergroup          0 2024-03-05 23:09 /sample/intro/a.txt
drwxr-xr-x   - g3 supergroup          0 2024-03-05 23:40 /sample1
drwxr-xr-x   - g3 supergroup          0 2024-03-05 22:59 /simple
```

6. Copy a file from local file system to HDFS

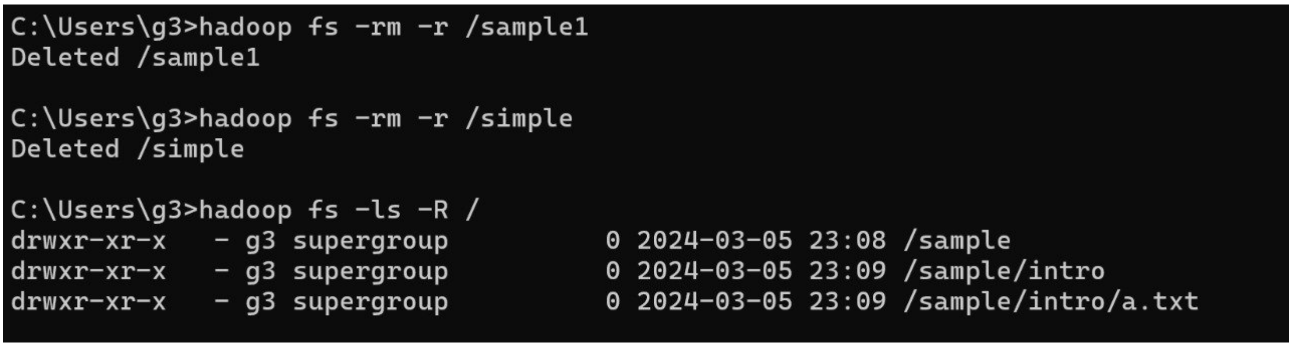
Command: hadoop fs -put/root/sample/test.txt/sample/test.txt



```
Microsoft Windows [Version 10.0.22621.3155]
(c) Microsoft Corporation. All rights reserved.

C:\Users\g3>hadoop fs -put "C:\Users\g3\OneDrive\Desktop\Demo.txt" /sample/intro/a.txt

C:\Users\g3>hadoop fs -ls -R /
drwxr-xr-x   - g3 supergroup          0 2024-03-05 23:08 /sample
drwxr-xr-x   - g3 supergroup          0 2024-03-05 23:09 /sample/intro
drwxr-xr-x   - g3 supergroup          0 2024-03-06 00:01 /sample/intro/a.txt
-rw-r--r--   1 g3 supergroup         36 2024-03-06 00:01 /sample/intro/a.txt/Demo.txt
-rw-r--r--   1 g3 supergroup          0 2024-03-05 23:56 /sample/intro/a.txt/demo.txt
```

7. Remove a directory from HDFS

Command: hadoop fs -rm-r /sample1



```
C:\Users\g3>hadoop fs -rm -r /sample1
Deleted /sample1

C:\Users\g3>hadoop fs -rm -r /simple
Deleted /simple

C:\Users\g3>hadoop fs -ls -R /
drwxr-xr-x   - g3 supergroup          0 2024-03-05 23:08 /sample
drwxr-xr-x   - g3 supergroup          0 2024-03-05 23:09 /sample/intro
drwxr-xr-x   - g3 supergroup          0 2024-03-05 23:09 /sample/intro/a.txt
```

RESULT:

Thus the file management commands in hadoop has been executed successfully.

EX NO:9 USING HADOOP APIs

DATE:

AIM:

To Study Hadoop API - PIG Commands

DESCRIPTION:

- Apache Pig is an alternative to MapReduce Programming
- It provides a high-level scripting language, known as Pig Latin which is used to develop the data analysis codes.
- Pig uses both structured and unstructured data as input to perform analytics and uses HDFS to store the results.
- There are two major components of the Pig:

  1. Pig Latin script language
  2. A runtime engine

COMMANDS:

1. Create a directory

Command: fs -mkdir /piglatindemos;

2. To select tuples from a relation based on specified conditions

Objective: Find the tuples of those student where the GPA is greater than 4.0

Command:

```
A=load'/pigdemo/student.tsv' as (rollno:int,name:chararray,gpa:float);

B=filter by gpa > 4.0

DUMP
```

3. Data Transformation based on columns of data

Objective:.Display the name of all students in uppercase

Command:

```
A=load’/pigdemo/student.tsv’ as (rollno:int,name:chararray,gpa:float);
B=foreach A generate UPPER (name);
DUMP B;
```

4. To group Data

Objective: Group tuples of students based on their GPA

Command:

```
A=load’/pigdemo/student.tsv’ as (rollno:int,name:chararray,gpa:float);
B=GROUP A BY gpa;
DUMP B;
```

5. DISTINCT -To remove duplicate tuples

Objective: To remove duplicate tuples of students

Command:

```
A=load’/pigdemo/student.tsv’ as (rollno:int,name:chararray,gpa:float);
B=DISTINCT A;
DUMP B;
```

6. LIMIT

Objective: To limit the number of output tuples

Command:

```
A=load’/pigdemo/student.tsv’ as (rollno:int,name:chararray,gpa:float);
B=LIMIT A 3;
DUMP B;
```

7. ORDER BY

Objective: To sort a relation based on specific value

Command:

```
A=load’/pigdemo/student.tsv’ as (rollno:int,name:chararray,gpa:float);
B=ORDER A BY name;
DUMP B;
```

8. JOIN

Objective: To join two relations or more based on values in the common field

Command:

```
A=load’/pigdemo/student.tsv’ as (rollno:int,name:chararray,gpa:float);
B=load’/pigdemo/department.tsv’ as (rollno:int,deptno:int,deptname:chararray);
C=JOIN A BY rollno,B BY rollno;
DUMP C;
DUMP B;
```

9. UNION

Objective: To merge the contents of two relations

Command:

```
A=load’/pigdemo/student.tsv’ as (rollno,name,gp);
B=load’/pigdemo/department.tsv’ as (rollno,deptno,deptname);
C=UNION A,B;
STORE C INTO ‘/pigdemo/uniondemo’;
DUMP B;
```

10. SPLIT

Objective:To partition a relation into two or more relations

Command:

```
A=load'/pigdemo/student.tsv' as (rollno:int,name:chararray,gpa:float);

SPLIT A INTO X IF gpa==4.0, Y IF gpa<=4.0;

DUMP X;
```

RESULT:

The study of Hadoop API – PIG was carried out successfully.

Ex-10 IMPLEMENT WORD COUNT USING MAPREDUCE

Date:

Aim

To implement word count using mapreduce.

Algorithm

Mapper Algorithm

1) START

2) Import sys module.

3) Read each of lines in STDIN one-by-one.

4) Split the line with space as separator, to obtain words.

5) Print the obtained words along with count as 1 to STDOUT for reducer.py to read.

6) REPEAT step 4 and step 5 till all lines in STDIN have been read.

7) STOP

Reducer Algorithm

1) START

2) Import sys module.

3) Create dictionary mapper to store count of each words

4) Read each of lines in STDIN one-by-one.

5) Read the word and count from the line.

6) Increase the count of the word in the Counter by int(count).

7) REPEAT step 5 and step 6 till all lines are read.

8) Print the word and its associated count to the STDOUT.

9) STOP

Source Code

```
#mapper.py

import sys
for line in sys.stdin:
 line = line.strip()
 words = line.split()
 for word in words:
  print(word,"\t",1)

#reducer.py

import sys
dictionaryWord={}
for line in sys.stdin:
 line = line.strip()
 word, count = line.split('\t', 1)
 dictionaryWord[word]=dictionaryWord.get(word,0)+int(count)
for x,y in dictionaryWord.items():
 print(x,"\t",y)
```

Execution Steps

1. Create a folder “WordCount”.
2. Place the required files(mapper.py, reducer.py, input.txt) in “WordCount” folder.
3. Execute the command on the terminal:

```
  cat input.txt | python mapper.py | sort -k1,1 | python reducer.py
```

Sample Input and Output

Input:

#input.txt

Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged. It was popularised in the 1960s with the release of Letraset sheets containing Lorem Ipsum passages, and more recently with desktop publishing software like Aldus PageMaker including versions of Lorem Ipsum.

Output:

```
1500s,	1

1960s	1

a	2

Aldus	1

also	1

an	1

and	3

been	1

book.	1

but	1

centuries,	1

containing	1

desktop	1

dummy	2

electronic	1

essentially	1

ever	1

five	1

galley	1
```

```
has	2
in	1
including	1
industry.	1
industry's	1
into	1
Ipsum	3
Ipsum.	1
is	1
it	1
It	2
leap	1
Letraset	1
like	1
Lorem	4
make	1
more	1
not	1
of	4
only	1
PageMaker	1
passages,	1
popularised	1
printer	1
printing	1
publishing	1
recently	1
release	1
```

```
remaining    1
scrambled    1
sheets    1
simply    1
since    1
software    1
specimen    1
standard    1
survived    1
text    2
the    6
to    1
took    1
type    2
typesetting    1
typesetting,    1
unchanged.    1
unknown    1
versions    1
was    1
when    1
with    2
```

Result:

Thus implementation of word count using mapreduce has been executed successfully.

Ex-11 IMPLEMENT LINE COUNT USING MAPREDUCE

Date:

Aim

To implement line count using mapreduce.

Algorithm

Mapper Algorithm

1) START
2) Import sys module.
3) Read each of lines in STDIN one-by-one.
4) Remove the leading and trailing whitespace from the line.
5) Print the line along with count as 1 to STDOUT for reducer.py to read.
6) REPEAT step 4 and step 5 till all lines in STDIN have been read.
7) STOP

Reducer Algorithm

1) START
2) Import sys module.
3) Create dictionary mapper to store count of each lines
4) Read each of lines in STDIN one-by-one.
5) Read the line and count the line.
6) Increase the count of the line in the Counter by int(count).
7) REPEAT step 5 and step 6 till all lines are read.
8) Print the line and its associated count to the STDOUT.
9) STOP

Source Code

#mapper.py

```
import sys
for line in sys.stdin:
    line = line.strip()
    print('%s\t%s' % (line, 1))
```

#reducer.py

```
from operator import itemgetter
import sys

current_line = None
current_count = 0
line = None

for line in sys.stdin:
    line = line.strip()

    line, count = line.split('\t', 1)

    try:
        count = int(count)
    except ValueError:
        continue

    if current_line == line:
        current_count += count
    else:
        if current_line:
            print('%s\t%s' % (current_line, current_count))
        current_count = count
        current_line = line

if current_line == line:
    print('%s\t%s' % (current_line, current_count))
```

Execution Steps

1. Create a folder “Line Count”.

2. Place the required files(mapper.py, reducer.py, input.txt) in “Line Count” folder.

3. Execute the command on the terminal:

```
cat input.txt | python mapper.py | sort -k1,1 | python reducer.py
```

Sample Input and Output

Input:

#input.txt

1.Big data primarily refers to data sets that are too large or complex to be dealt with by traditional data-processing application software.

2.Data with many entries (rows) offer greater statistical power, while data with higher complexity (more attributes or columns) may lead to a higher false discovery rate.

3.Big data analysis challenges include capturing data, data storage, data analysis, search, sharing, transfer, visualization, querying, updating, information privacy, and data source.

4.Big data was originally associated with three key concepts: volume, variety, and velocity.

5.The analysis of big data presents challenges in sampling, and thus previously allowing for only observations and sampling.

6.Big data was originally associated with three key concepts: volume, variety, and velocity.

Output:

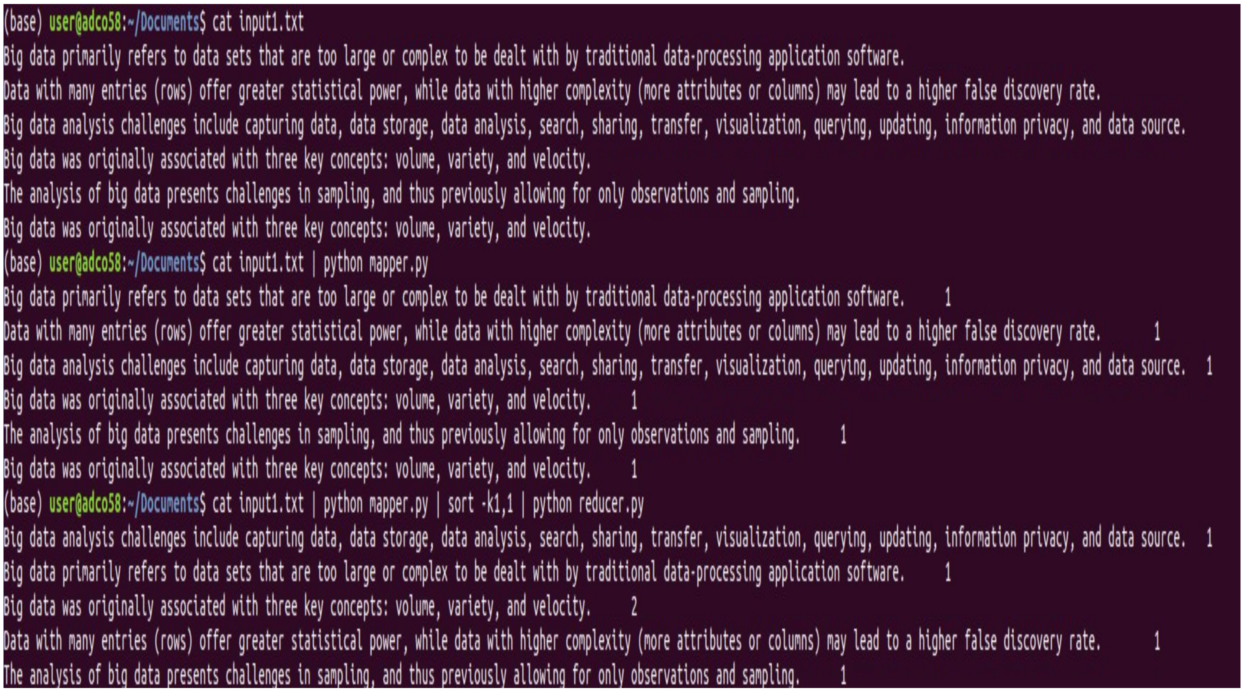
Big data analysis challenges include capturing data, data storage, data analysis, search, sharing, transfer, visualization, querying, updating, information privacy, and data source. 1

Big data primarily refers to data sets that are too large or complex to be dealt with by traditional data-processing application software. 1

Big data was originally associated with three key concepts: volume, variety, and velocity. 2

Data with many entries (rows) offer greater statistical power, while data with higher complexity (more attributes or columns) may lead to a higher false discovery rate. 1

The analysis of big data presents challenges in sampling, and thus previously allowing for only observations and sampling. 1



```
(base) user@adco58:~/Documents$ cat input1.txt
Big data primarily refers to data sets that are too large or complex to be dealt with by traditional data-processing application software.
Data with many entries (rows) offer greater statistical power, while data with higher complexity (more attributes or columns) may lead to a higher false discovery rate.
Big data analysis challenges include capturing data, data storage, data analysis, search, sharing, transfer, visualization, querying, updating, information privacy, and data source.
Big data was originally associated with three key concepts: volume, variety, and velocity.
The analysis of big data presents challenges in sampling, and thus previously allowing for only observations and sampling.
Big data was originally associated with three key concepts: volume, variety, and velocity.
(base) user@adco58:~/Documents$ cat input1.txt | python mapper.py
Big data primarily refers to data sets that are too large or complex to be dealt with by traditional data-processing application software.	1
Data with many entries (rows) offer greater statistical power, while data with higher complexity (more attributes or columns) may lead to a higher false discovery rate.	1
Big data analysis challenges include capturing data, data storage, data analysis, search, sharing, transfer, visualization, querying, updating, information privacy, and data source.	1
Big data was originally associated with three key concepts: volume, variety, and velocity.	1
The analysis of big data presents challenges in sampling, and thus previously allowing for only observations and sampling.	1
Big data was originally associated with three key concepts: volume, variety, and velocity.	1
(base) user@adco58:~/Documents$ cat input1.txt | python mapper.py | sort -k1,1 | python reducer.py
Big data analysis challenges include capturing data, data storage, data analysis, search, sharing, transfer, visualization, querying, updating, information privacy, and data source.	1
Big data primarily refers to data sets that are too large or complex to be dealt with by traditional data-processing application software.	1
Big data was originally associated with three key concepts: volume, variety, and velocity.	2
Data with many entries (rows) offer greater statistical power, while data with higher complexity (more attributes or columns) may lead to a higher false discovery rate.	1
The analysis of big data presents challenges in sampling, and thus previously allowing for only observations and sampling.	1
```

Result

Thus implementation of line count using mapreduce has been executed successfully.

Ex-12 **IMPLEMENT WORD SEARCH USING MAPREDUCE**

Date:

Aim

To implement word search using mapreduce.

Algorithm

Mapper Algorithm

1) START

2) Import re module.

3) Read each of lines one-by-one.

4) Split the input_text into smaller chunks.

5) Initialize an empty list to store the results of the mapper function for each chunk.

6) STOP

Reducer Algorithm

1) START

2) Create dictionary mapper to store count of each lines

3) Append the result of the mapper function to the list of results.

4) Combine the results from all mapper calls into a single list.

5) Initialize an empty dictionary to store the final word counts.

6) For each word and its count, add the count to the corresponding entry in the word_counts dictionary.

7) Return the word_counts dictionary containing the counts of each search word found in the input_text.

8) STOP

Source Code

#mapper.py

```
import re

def mapper(chunk, search_words):
    word_count = {}
    for word in search_words:
        word_count[word] = len(re.findall(r'\b' + word + r'\b', chunk))
    return word_count
```

#reducer.py

```
def reducer(results):
    word_count = {}
    for result in results:
        for word, count in result.items():
            word_count[word] = word_count.get(word, 0) + count
    return word_count
```

#word_search.py

```
from multiprocessing import Pool
import mapper
import reducer

def mapreduce(data, mapper_func, reducer_func, search_words, chunk_size=100):
    chunks = [data[i:i+chunk_size] for i in range(0, len(data), chunk_size)]
    with Pool() as pool:
        mapped = pool.starmap(mapper_func, [(chunk, search_words) for chunk in chunks])
    reduced = reducer_func(mapped)
    return reduced

if __name__ == "__main__":

    with open('input.txt', 'r') as file:
        input_data = file.read()



    search_words = ['Sun', 'position', 'light', 'marvel', 'temperature', 'due', 'soar', 'moon', 'star']


    result = mapreduce(input_data, mapper.mapper, reducer.reducer, search_words)
    for word, count in result.items():
        print(f"{word}: {count}")
```

Execution Steps

1. Create a folder “Word Search”.

2. Place the required files(mapper.py, reducer.py, word_search.py, input.txt) in “Word Search” folder.

3. Execute the command on the terminal:

```
python word_search.py
```

Sample Input and Output

Input:

#input.txt

As the closest planet to the Sun, Mercury is a celestial marvel with a bounty of unique attributes waiting to be explored. The planet experiences extreme temperature fluctuations due to its proximity to the

Sun. Daytime temperatures can soar up to a blistering 800 degree Fahrenheit (427 degrees Celsius), while night-time temperatures plummet to a frigid -290 degrees Fahrenheit (-180 degrees Celsius). Orbiting the Sun at a breakneck pace, Mercury completes a full orbit in just 88 Earth days, making it the shortest year of any planet in our solar system. But studying the planet is very difficult due to its position and the impact of the Sun's gravity.

Output:

Sun: 4

position: 1

light: 0

marvel: 1

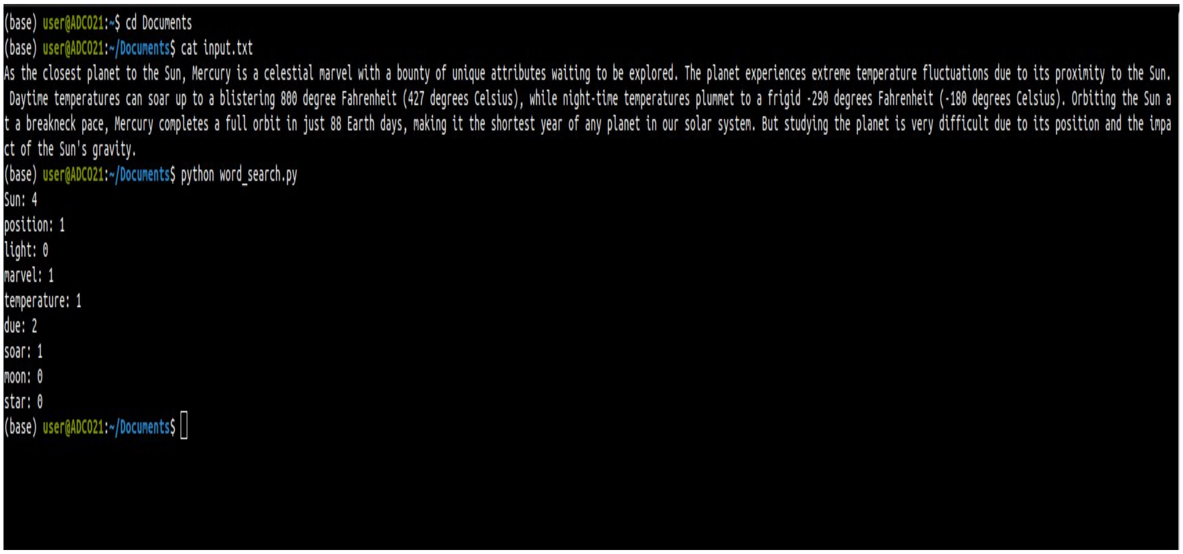
temperature: 1

due: 2

soar: 1

moon: 0

star: 0



```
(base) user@ADC021:~$ cd Documents
(base) user@ADC021:~/Documents$ cat input.txt
As the closest planet to the Sun, Mercury is a celestial marvel with a bounty of unique attributes waiting to be explored. The planet experiences extreme temperature fluctuations due to its proximity to the Sun.
 Daytime temperatures can soar up to a blistering 800 degree Fahrenheit (427 degrees Celsius), while night-time temperatures plummet to a frigid -290 degrees Fahrenheit (-180 degrees Celsius). Orbiting the Sun a
t a breakneck pace, Mercury completes a full orbit in just 88 Earth days, making it the shortest year of any planet in our solar system. But studying the planet is very difficult due to its position and the impa
ct of the Sun's gravity.
(base) user@ADC021:~/Documents$ python word_search.py
Sun: 4
position: 1
light: 0
marvel: 1
temperature: 1
due: 2
soar: 1
moon: 0
star: 0
(base) user@ADC021:~/Documents$ 
```

Result

Thus implementation of word search using mapreduce has been executed successfully.