National College of Ireland

Configuration Manual

MSc Research Project
MSc Cloud Computing

Vaishnavi Waghmare
Student ID: 21172846

School of Computing
National College of Ireland

Supervisor: Prof. Shivani Jaswal

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Vaishnavi Sarjerao Waghmare
Student ID:	21172846
Programme:	MSc Cloud Computing **Year:** 2024
Module:	Research Project
Lecturer:	Prof. Shivani Jaswal
Submission Due Date:	14th January 2024
Project Title:	A Comparative Analysis of Metaheuristic Algorithms for Optimizing Tasks in Serverless Frameworks for MapReduce Applications
Word Count:	500 **Page Count:** 10

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Vaishnavi Sarjerao Waghmare

Date: 12th January 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Vaishnavi Sarjerao Waghmare
21172846

1 Introduction

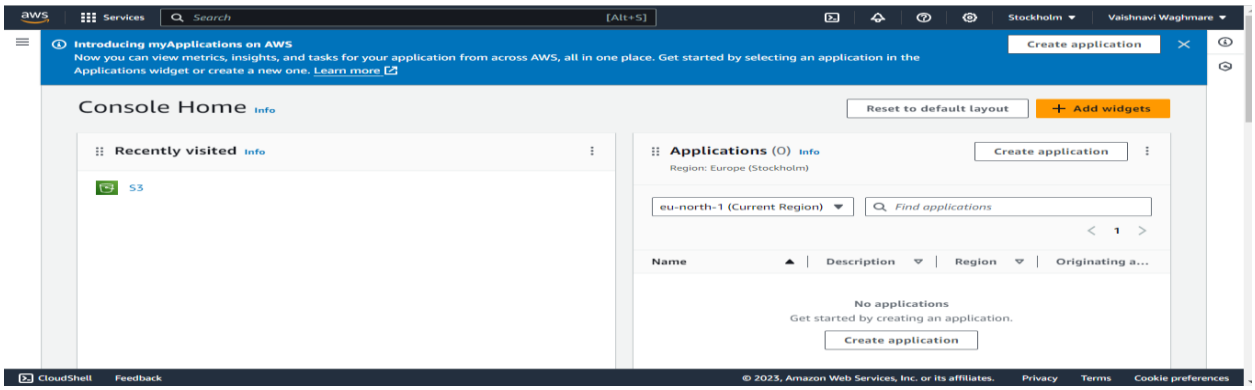
The configuration manual demonstrates running of the integrated code.

2 Requirements

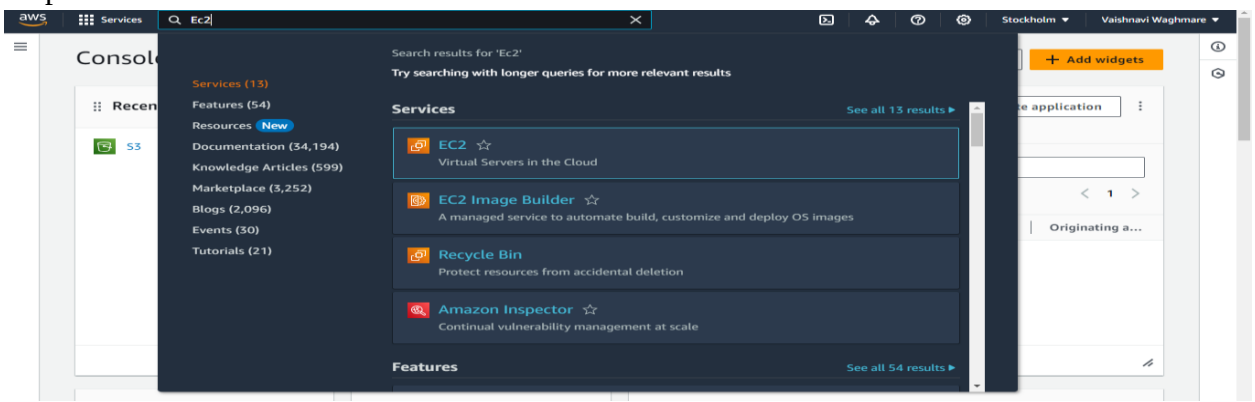
- AWS Account
- Windows/Mac Laptop
- Hadoop Installation
- AWS CLI Installation
- OpenSSH installation

3 Code Implementation

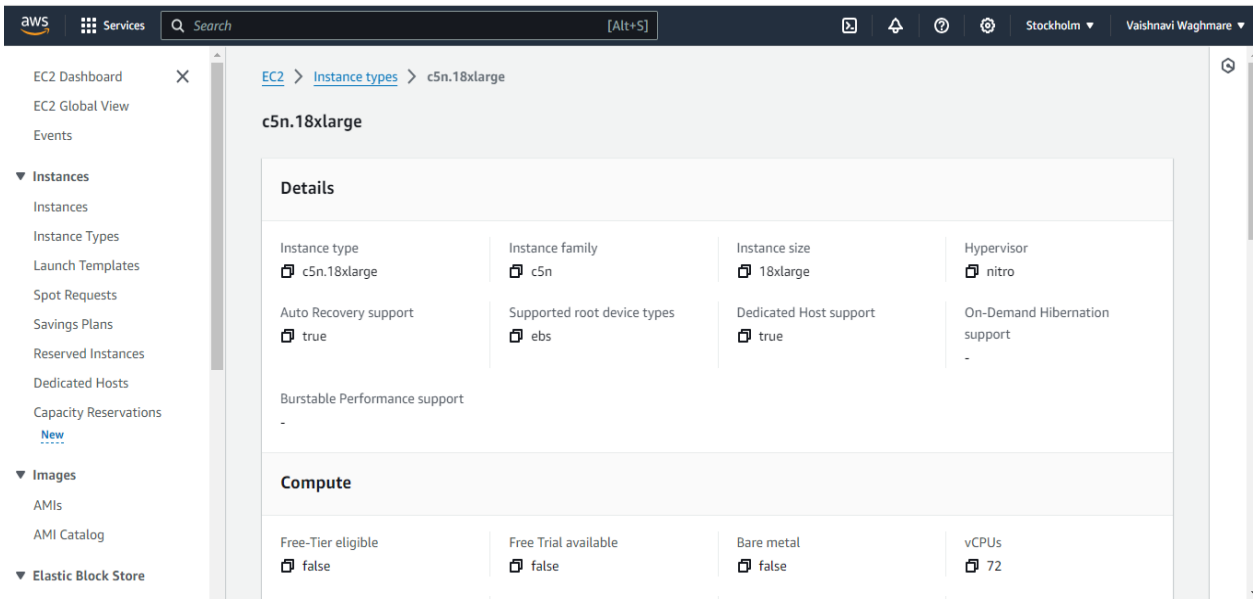
Step 1. Login to your AWS Account



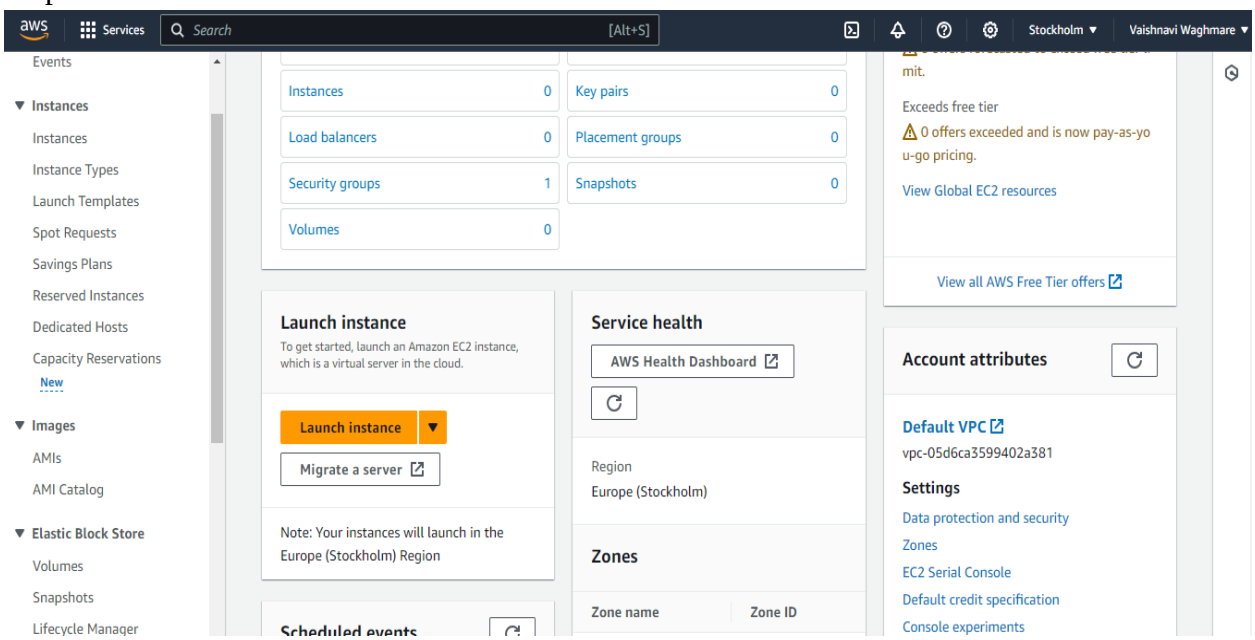
Step 2. Choose EC2 Instance



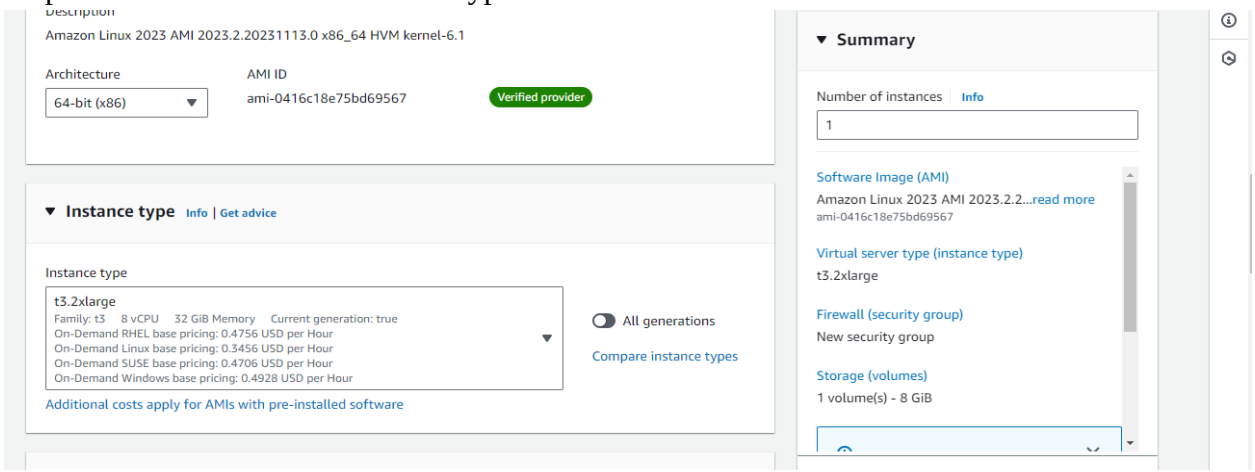
Step 3. Choose an Instance



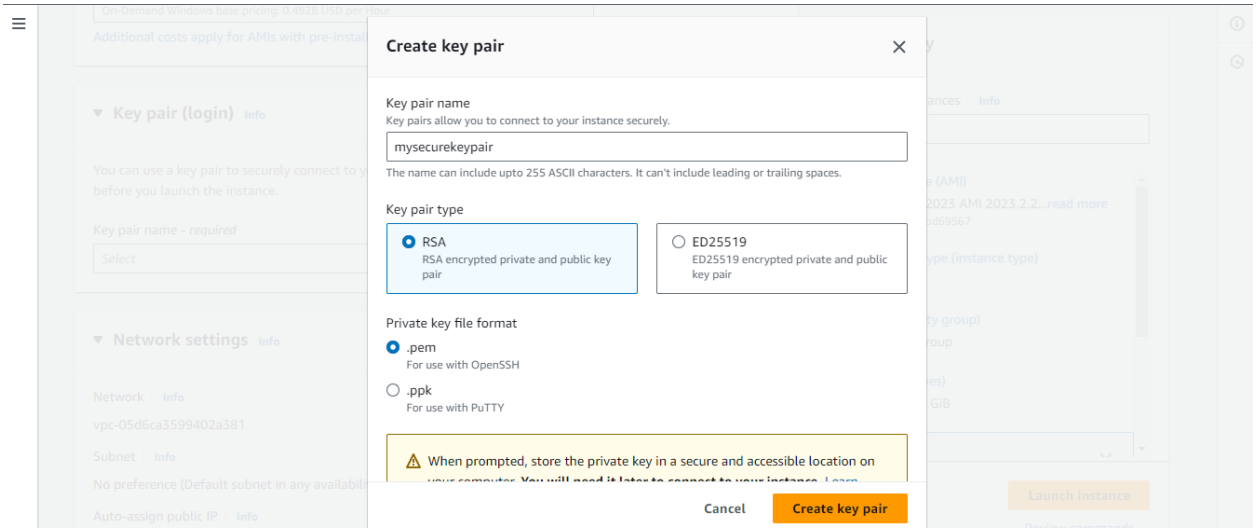
Step 4. Launch instance



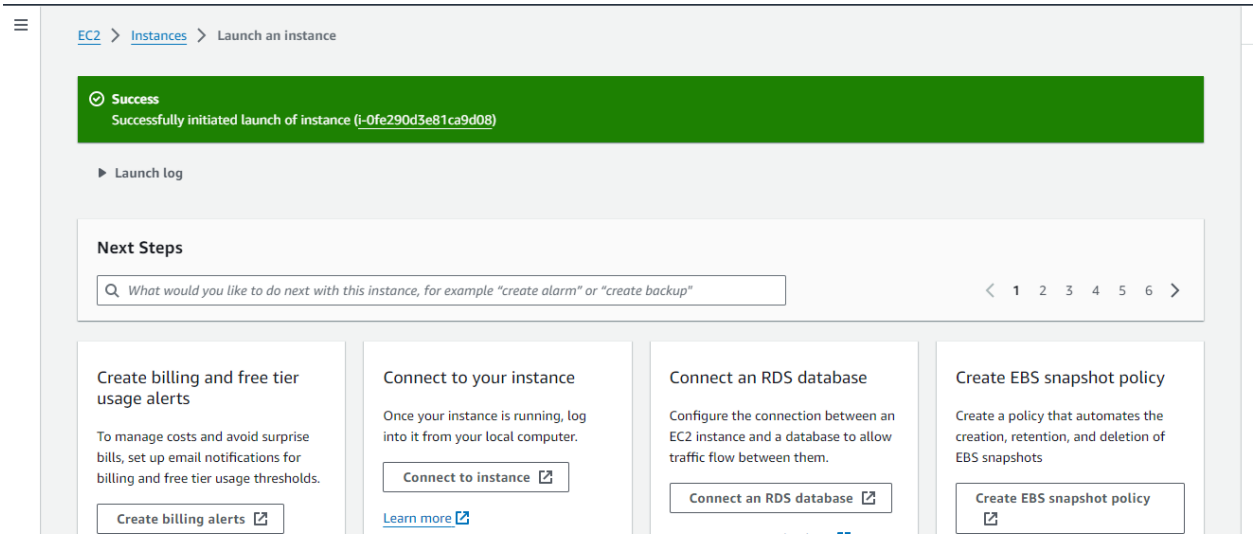
Step 5. Choose kernel and instance type



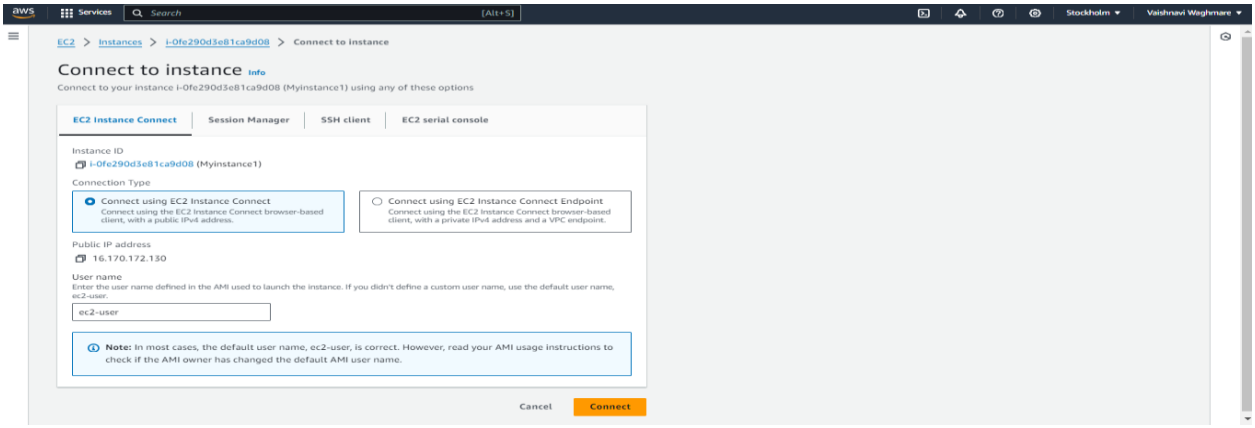
Step 6. Create a key – pair



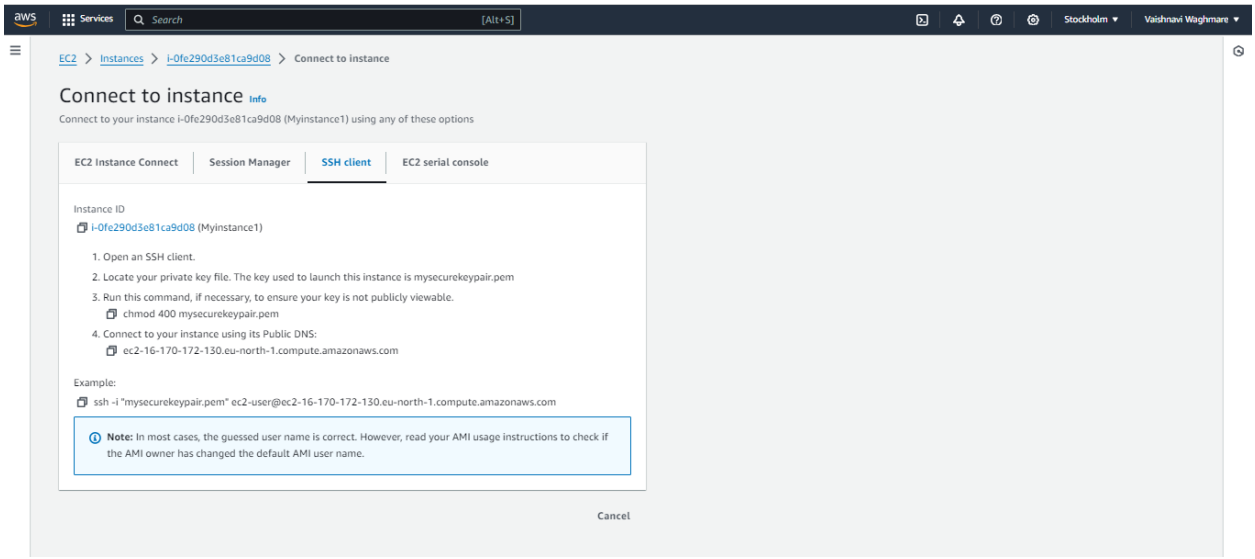
Step 7. Launch Instance



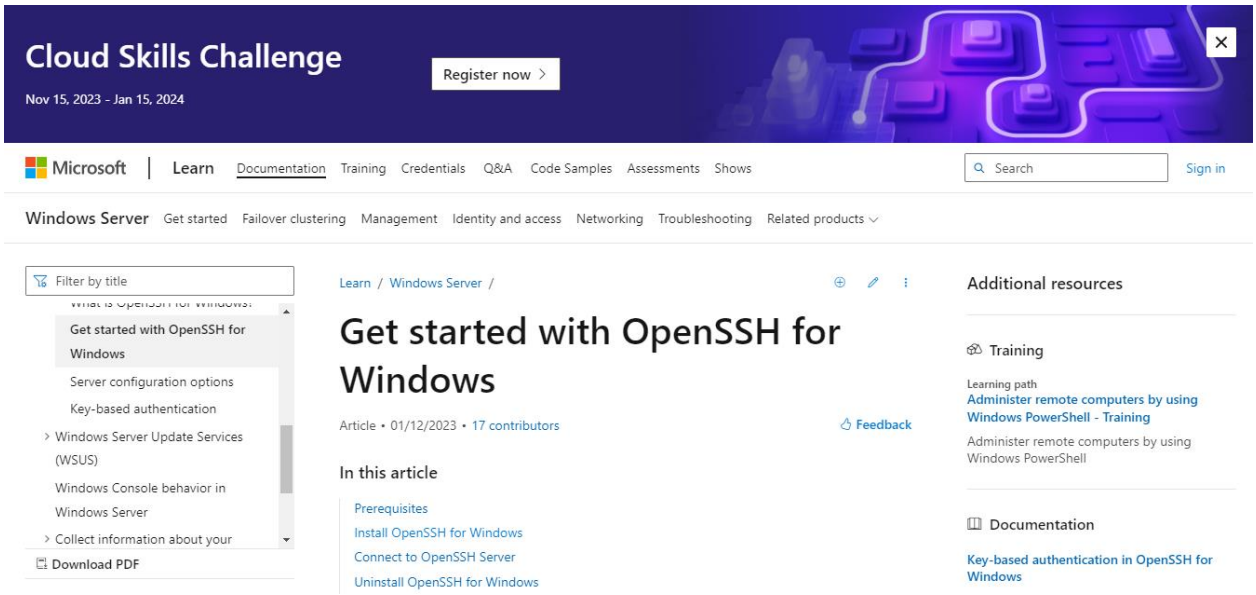
Step 8. Connecting your AWS instance



Step 9. configure sshclient

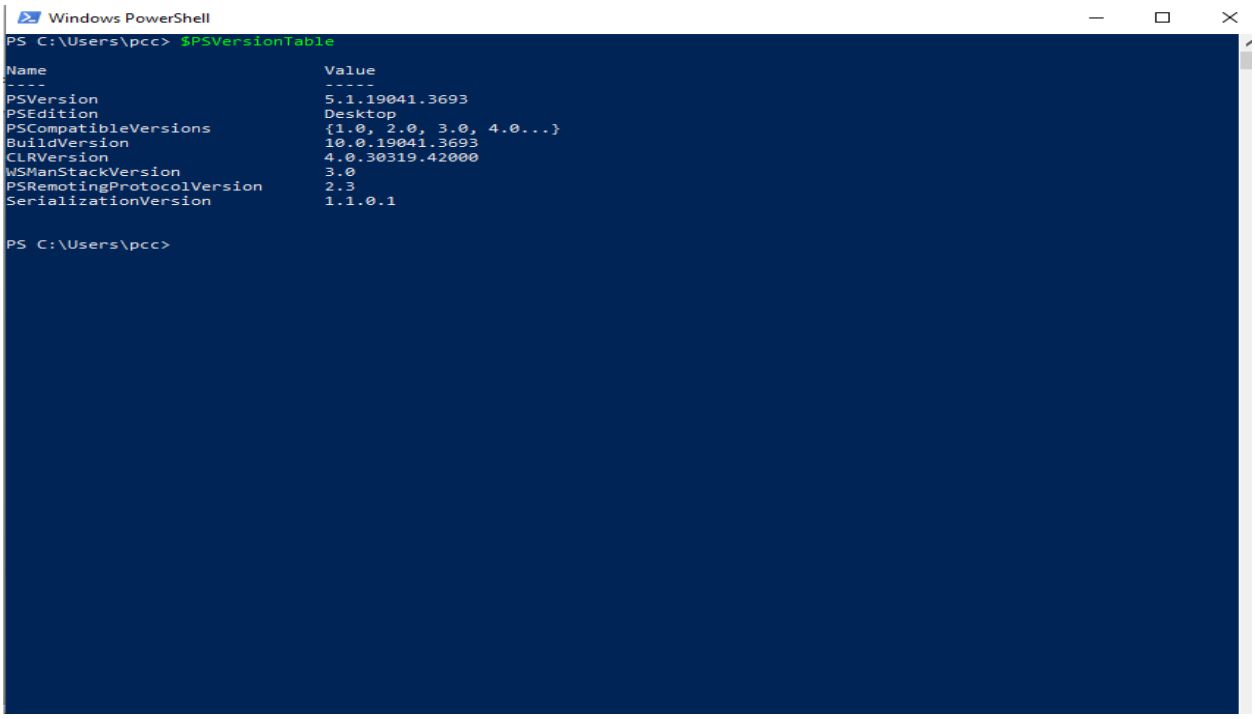


Step 10. Install OpenSSH client

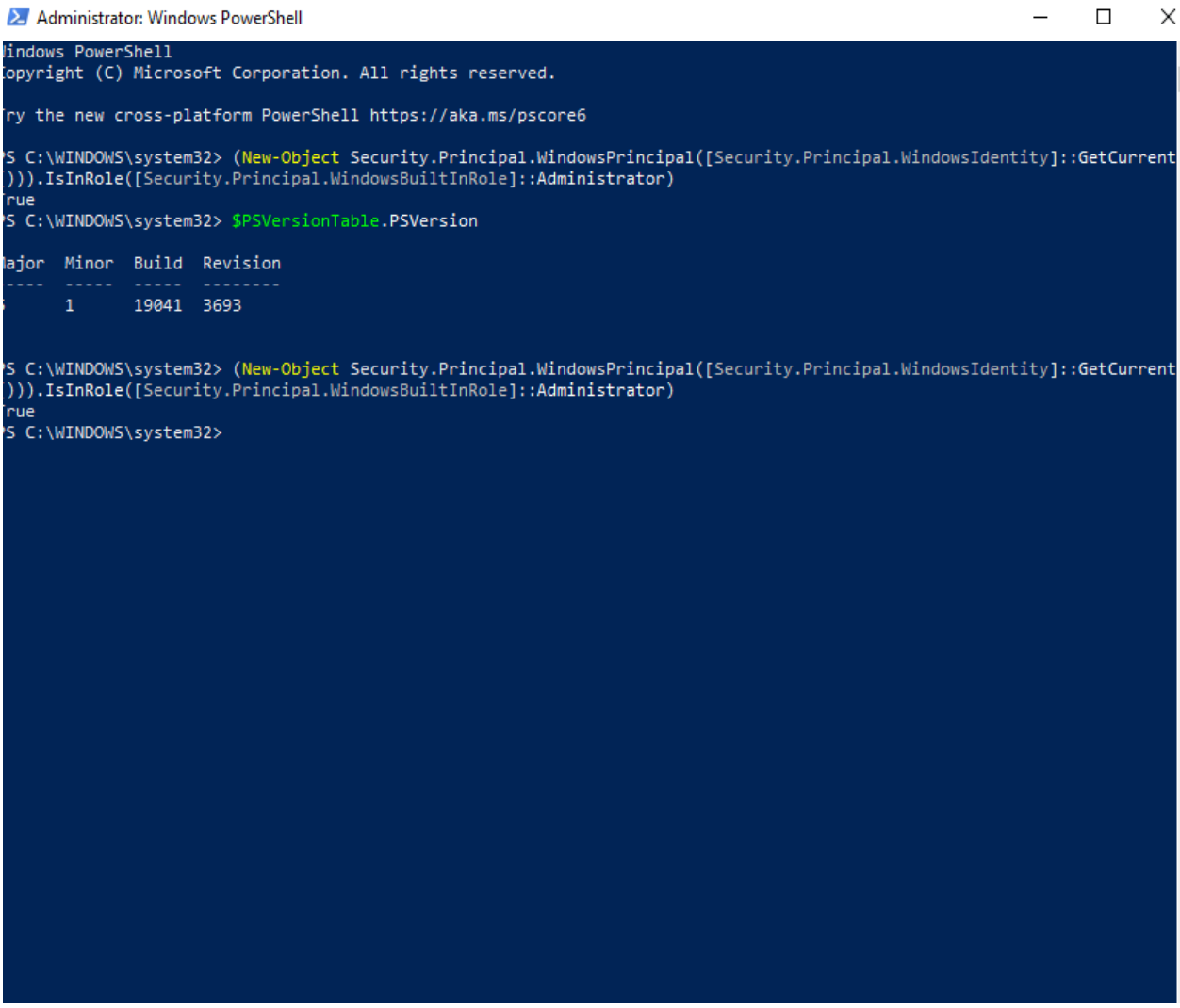


Step 11. Check for prerequisites

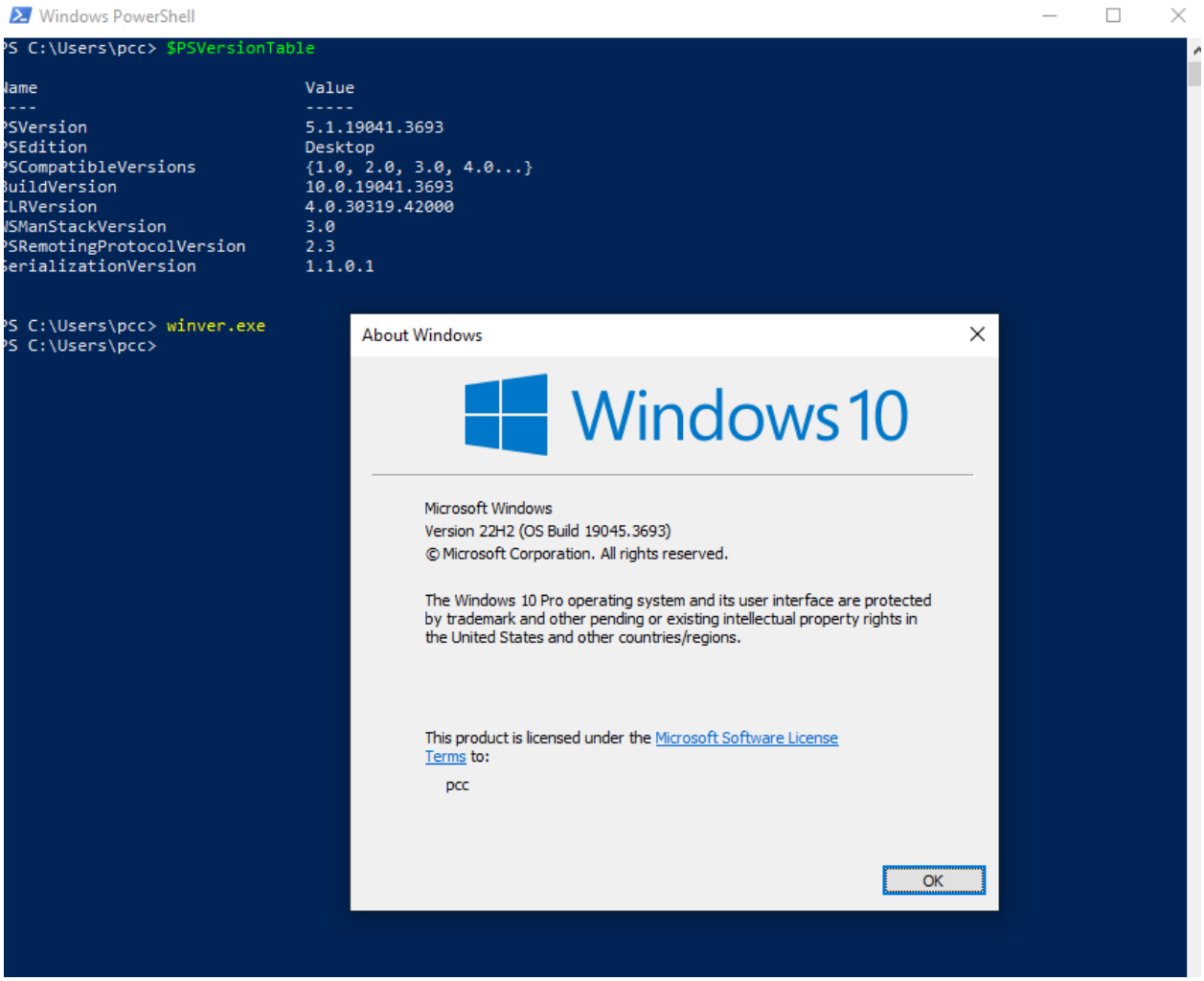
Step 12. Open Powershell as administrator and Check powershell version



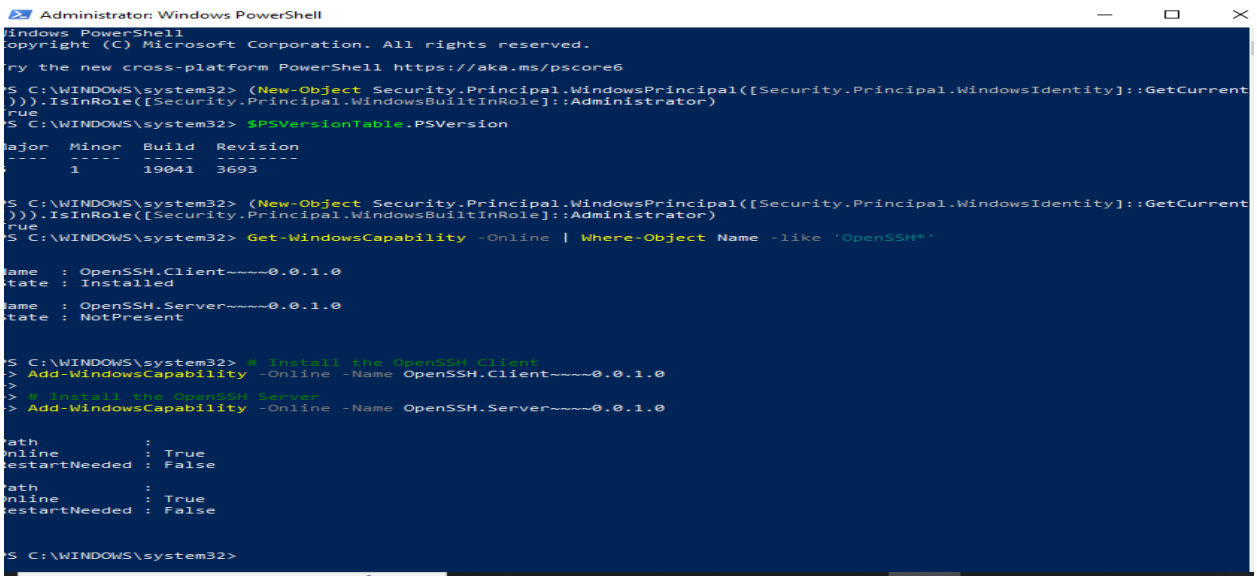
Step 13. Check for other prerequisites



Step 14. Check windows version

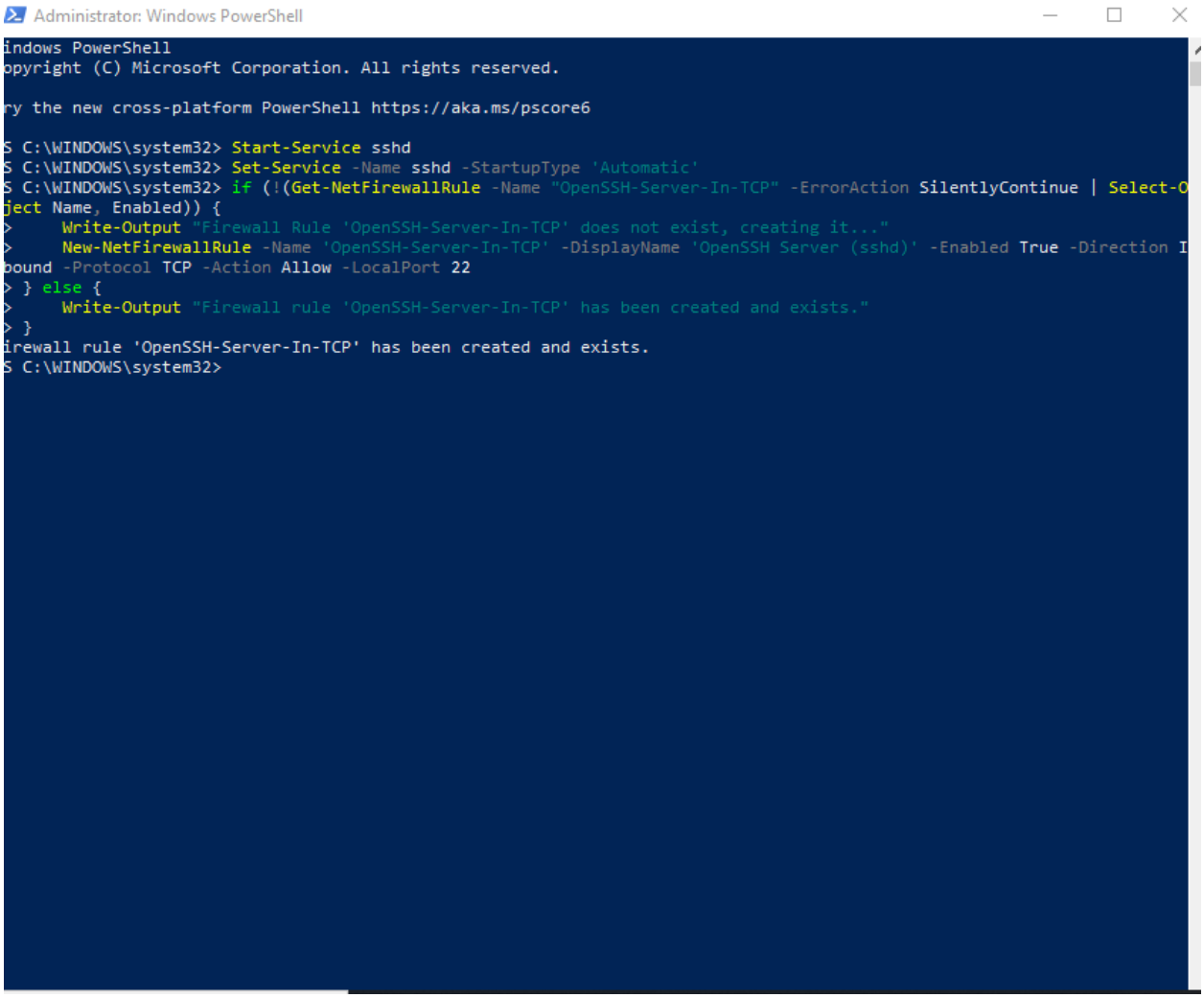


Step 15. Install OpenSSH

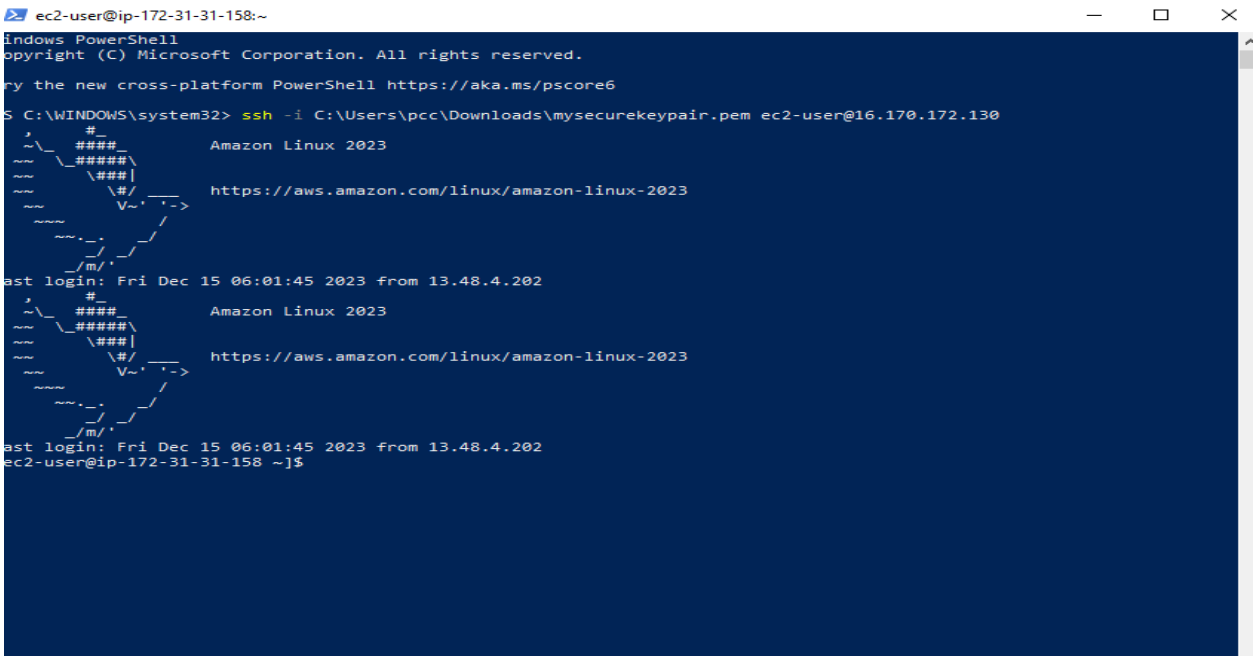


Step 16. Restart the powershell as administrator

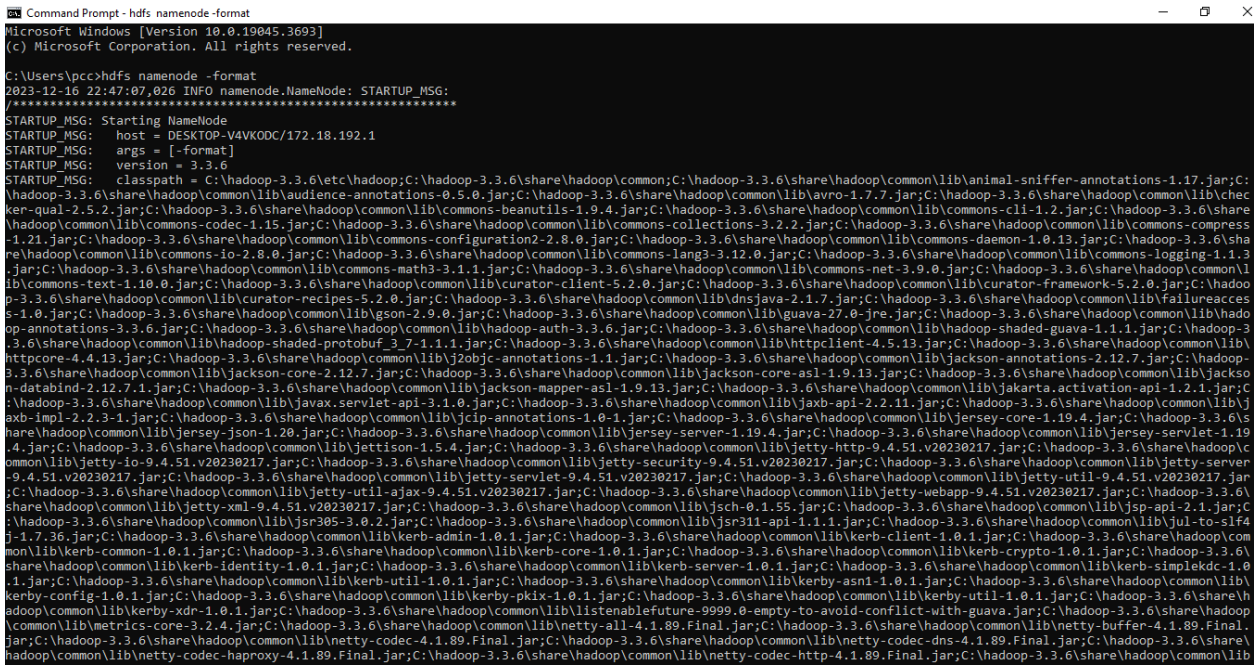
Step 17. Configure SSH



Step 18. Access SSH



Step 19. Verify response if Hadoop is installed in your instance



Step 20. Configure your Hadoop on our Ec2 instance and run the .sh functions

```
1   #!/bin/bash
2
3   # Run Teragen job
4   yarn jar $HADOOP_HOME/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.3.6.jar teragen 500000 ./I
5
6   # Check if the Teragen job was successful
7   if [ $? -eq 0 ]; then
8     echo "Teragen Hadoop job completed successfully."
9   else
10    echo "Error: Teragen Hadoop job failed."
11    exit 1
12  fi
13
14
15  # Run Terasort job
16  yarn jar $HADOOP_HOME/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.3.6.jar terasort ./INPUT_D
17
18  # Check if the Terasort job was successful
19  if [ $? -eq 0 ]; then
```

Step 21. Execute the script successfully

```
2023-12-18 10:43:08,815 INFO mapreduce.Job: Job job_local460332852_0001 running in uber mode : false
2023-12-18 10:43:08,816 INFO mapreduce.Job:  map 100% reduce 100%
2023-12-18 10:43:08,817 INFO mapreduce.Job: Job job_local460332852_0001 completed successfully
2023-12-18 10:43:08,824 INFO mapreduce.Job: Counters: 30
        File System Counters
                FILE: Number of bytes read=101344528
                FILE: Number of bytes written=1846222
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
        Map-Reduce Framework
                Map input records=500000
                Map output records=3
                Map output bytes=81
                Map output materialized bytes=93
                Input split bytes=113
                IO_ERROR=0
                WRONG_LENGTH=0
                WRONG_MAP=0
                WRONG_REDUCE=0
        File Input Format Counters
                Bytes Read=50390636
        File Output Format Counters
                Bytes Written=35
Teravalidate Hadoop job completed successfully.
All Hadoop jobs completed successfully.
[ec2-user@ip-172-31-31-158 ~]$ nano run_benchmarking.sh
[ec2-user@ip-172-31-31-158 ~]$ client_loop: send disconnect: Connection reset
PS C:\Users\pcc\Desktop\office\Sana\Vaishanavi\Code>
```