K.R. MANGALAM UNIVERSITY
THE COMPLETE WORLD OF EDUCATION

Cloud Computing
Lab File (ETCA362A)

Submitted To:

Mrs. Neetu GUPTA
(Assistant Professor)

SOET Department

Submitted By:

Harsh
Shokeen
1912730002

BTechTIIIE

Cloud Computing Lab Index

Q.1 Case Study on Cloud Computing and different Cloud vendors like AWS, AZURE, IBM, GCP

Abstract:

This case study aims to provide an in-depth analysis and comparison of cloud computing services offered by major cloud vendors, namely Amazon Web Services (AWS), Microsoft Azure, IBM Cloud, and Google Cloud Platform (GCP). Cloud computing has revolutionized the IT landscape, enabling organizations to leverage scalable and cost-effective solutions. By examining the offerings of these prominent vendors, this case study provides insights into the features, capabilities, pricing models, and customer experiences associated with each cloud platform.

Introduction

Cloud computing refers to the delivery of computing services, including storage, processing power, software applications, and databases, over the internet. It eliminates the need for organizations to invest in and manage their own physical infrastructure, such as servers and data centers, by providing access to shared and scalable resources hosted on remote servers.

The significance of cloud computing in modern IT infrastructure can be understood through the following key points:

- Scalability and Flexibility: Cloud computing offers unprecedented scalability, allowing organizations to quickly scale their resources up or down based on demand. This flexibility enables businesses to adapt to changing needs without incurring significant costs or experiencing downtime.

- Cost Efficiency: Cloud computing follows a pay-as-you-go model, where organizations pay only for the resources they consume. This eliminates the need for large upfront investments in hardware and infrastructure. Additionally, cloud providers handle maintenance and updates, reducing operational costs.

- Global Accessibility: Cloud computing enables access to data, applications, and services from anywhere with an internet connection. This fosters collaboration and remote work capabilities, facilitating seamless communication and productivity across geographically dispersed teams.

- Disaster Recovery and Data Protection: Cloud providers often offer robust disaster recovery and data backup solutions. By storing data in multiple locations, organizations can minimize the risk of data loss due to hardware failures, natural disasters, or other unforeseen events.

- Innovation and Time-to-Market: Cloud computing allows businesses to experiment and innovate rapidly. It provides easy access to a wide range of tools, frameworks, and development platforms, enabling faster development and deployment of applications and services.

- Security and Compliance: Cloud vendors invest heavily in security measures and compliance certifications to protect customer data. They often have dedicated teams of security experts and offer features such as encryption, access controls, and identity management to ensure data confidentiality and integrity.

- Big Data and Analytics: Cloud computing provides the infrastructure and services required for processing and analyzing large volumes of data. It offers scalable storage and computing capabilities, enabling organizations to leverage data-driven insights for decision-making and gaining a competitive edge.

Cloud computing has revolutionized the IT industry, empowering organizations of all sizes to access advanced technologies and capabilities that were previously limited to large enterprises. It has become an essential component of modern IT infrastructure, enabling agility, cost savings, scalability, and innovation for businesses across various sectors.

Amazon Web Services (AWS)

Amazon Web Services (AWS) is a comprehensive cloud computing platform provided by Amazon.com. It offers a vast array of services that enable organizations to build, deploy, and manage applications and infrastructure in the cloud. Here is an overview of AWS cloud services and their key features:

Compute Services:

Amazon Elastic Compute Cloud (EC2): Provides scalable virtual servers on-demand, allowing users to choose the desired computer capacity.

- AWS Lambda: Enables serverless computing, allowing developers to run code without provisioning or managing servers' storage and Content Delivery Services:
- Amazon Simple Storage Service (S3): Offers highly scalable and durable object storage for storing and retrieving data.
- Amazon Elastic Block Store (EBS): Provides persistent block-level storage volumes for EC2 instances.
- Amazon Glacier: Secure and durable storage for data archiving and long-term backup.
- Amazon Relational Database Service (RDS): Managed database service for popular relational databases like MySQL, PostgreSQL, Oracle, and SQL Server.
- Amazon DynamoDB: Fully managed NoSQL database for fast and flexible document and key-value data storage.

- Amazon Redshift: Fully managed data warehousing service for analyzing large datasets.
- Amazon Athena: Interactive query service for analyzing data stored in S3 using standard SQL.
- Amazon Kinesis: Real-time streaming data platform for collecting, processing, and analyzing data.

Case studies highlighting successful implementations of AWS in various industries:

1. **Netflix**: Netflix migrated its entire video streaming platform to AWS, taking advantage of AWS's scalability and global infrastructure to deliver seamless streaming to millions of users worldwide.
2. **Airbnb**: Airbnb relies on AWS for its infrastructure needs, enabling the company to handle peak traffic and scale rapidly during periods of high demand, ensuring a smooth booking experience for users.
3. **Capital One**: Capital One, a financial services company, migrated its infrastructure to AWS, leveraging AWS's security features and compliance certifications to ensure the protection of sensitive customer data.
4. **NASA Jet Propulsion Laboratory (JPL)**: JPL uses AWS for its mission-critical workloads, including the Mars Rover project. AWS's scalability and high-performance computing capabilities enable JPL to process and analyze vast amounts of data from space missions.

Microsoft Azure

Azure, provided by Microsoft, is a comprehensive cloud computing platform that offers a wide range of services for building, deploying, and managing applications and services. Azure provides a robust and scalable infrastructure to support organizations' digital transformation initiatives. Here is an introduction to Azure cloud services and its distinguishing features:

Compute Services:

- Azure Virtual Machines: Offers scalable virtual machines that run Windows or Linux, allowing users to deploy a wide range of applications.
- Azure Functions: Provides serverless computing, allowing developers to run code in response to events without managing infrastructure.
- Azure Blob Storage: Scalable object storage for storing and retrieving large amounts of unstructured data.
- Azure Files: Fully managed file shares accessible via the Server Message Block (SMB) protocol.
- Azure Content Delivery Network (CDN): Accelerates the delivery of web content, videos, and applications globally.

- Azure Virtual Network (VNet): Provides isolated and customizable virtual networks to securely connect resources.
- Azure Load Balancer: Distributes inbound and outbound network traffic to improve availability and scalability.
- Azure Traffic Manager: Routes incoming traffic across multiple endpoints based on defined policies.
- Azure Monitor: Provides comprehensive monitoring and diagnostics for applications, infrastructure, and network resources.
- Azure Automation: Automates manual, time-consuming tasks with workflow and process automation.
- Azure Resource Manager: Simplifies resource management and allows for consistent deployment and management across Azure.
- Azure Machine Learning: Offers a cloud-based environment for building, training, and deploying machine learning models.
- Azure Cognitive Services: Provides pre-built AI models for vision, speech, language, and decision-making capabilities.
- Azure Bot Service: Enables the creation of intelligent, conversational bots using natural language understanding.

Case studies illustrating the successful adoption of Azure in different organizations:

1. **Maersk:** Maersk, a global logistics company, adopted Azure to transform its operations and improve efficiency. By leveraging Azure's scalability and data analytics capabilities, Maersk gained real-time visibility into its supply chain, optimized route planning, and enhanced customer experience.
2. **Adobe:** Adobe, a leading software company, migrated its Creative Cloud services to Azure. This migration enabled Adobe to provide a seamless experience to its users, scale resources based on demand, and achieve global availability and performance.
3. **Rolls-Royce:** Rolls-Royce, an aerospace and defense company, chose Azure as its cloud platform to power its digital transformation. Azure's AI and IoT capabilities helped Rolls-Royce collect and analyze vast amounts of data from its aircraft engines, leading to improved maintenance processes and fuel efficiency.
4. **NBC Sports:** NBC Sports, a major sports broadcasting network, relies on Azure for streaming live sports events to millions of viewers. Azure's scalability and media services enable NBC Sports to deliver high-quality video streams, handle peak demand, and provide a seamless viewing experience.

Azure offers various pricing structures and cost optimization mechanisms to help customers manage their cloud spending effectively:

Pay-as-you-go: Customers pay for Azure services based on actual usage, with no upfront costs or long-term commitments. This model provides flexibility and allows for cost control.

IBM Cloud

IBM Cloud is a comprehensive cloud computing platform offered by IBM. It provides a wide range of infrastructure, platform, and software services to support organizations in building, deploying, and managing applications and services. Here is an overview of IBM Cloud services:

- IBM Virtual Servers: Offers scalable virtual server instances with a variety of configurations, allowing users to deploy and manage their applications.
- IBM Bare Metal Servers: Provides dedicated, single-tenant servers for workloads that require high performance, security, and control.
- IBM Cloud Foundry: An open-source platform that simplifies the deployment and management of applications across multiple cloud environments.
- IBM Kubernetes Service: Managed Kubernetes service for deploying and managing containerized applications.
- IBM Watson Assistant: Enables the creation of AI-powered virtual assistants to interact with users and provide personalized experiences.
- IBM Watson Studio: A platform for data scientists and developers to build and deploy AI models using a variety of tools and frameworks.
- IBM Watson IoT Platform: Helps organizations securely connect and manage IoT devices, collect and analyze data, and gain insights.

Case studies showcasing successful utilization of IBM Cloud in diverse industries:

1. **Medtronic**: Medtronic, a global leader in medical technology, used IBM Cloud to build and deploy a secure and scalable platform for analyzing and storing medical device data. By leveraging IBM Cloud's capabilities, Medtronic improved data management, accelerated insights, and enhanced patient care.
2. **The Weather Company**: The Weather Company, an IBM subsidiary, relies on IBM Cloud to process and analyze vast amounts of weather data in real-time. By leveraging IBM Cloud's scalability and analytics capabilities, The Weather Company delivers accurate weather forecasts and insights to millions of users worldwide.
3. **Bausch + Lomb: Bausch + Lomb,** a leading eye health company, adopted IBM Cloud to enable global collaboration, data analysis, and research. IBM Cloud provided Bausch + Lomb with a secure and scalable platform for storing and processing critical healthcare data.
4. **Bank of Tokyo-Mitsubishi UFJ**: Bank of Tokyo-Mitsubishi UFJ (BTMU) utilized IBM Cloud to develop a blockchain-based payment system. By leveraging IBM Cloud's blockchain platform, BTMU enhanced transaction security, reduced costs, and improved operational efficiency in cross-border payments.

Google Cloud Platform (GCP)

Google Cloud Platform (GCP) is a suite of cloud computing services provided by Google. It offers a wide range of services that cater to various aspects of cloud computing, including infrastructure, data analytics, and machine learning. Here is an introduction to GCP services:

Infrastructure Services:

Compute Engine: Provides virtual machines (VMs) that allow users to run applications on Google's infrastructure.

App Engine: Offers a platform for building and hosting web applications, providing automatic scaling and managed infrastructure.

Kubernetes Engine: Managed Kubernetes service that simplifies the deployment, management, and scaling of containerized applications.

Cloud Storage: Provides scalable and durable object storage for storing and accessing data from anywhere in the world.

Cloud SQL: Fully managed relational database service that supports MySQL and PostgreSQL.

Cloud Bigtable: A NoSQL wide-column database suitable for large-scale, low-latency workloads.

TensorFlow: An open-source machine learning framework that is widely used for developing and deploying ML models.

Virtual Private Cloud (VPC): Provides networking functionality for creating and managing private virtual networks.

Case studies highlight successful implementation of GCP in different use cases.

1. **Spotify:** Spotify, the popular music streaming platform, migrated its infrastructure to GCP to improve scalability and reliability. By leveraging GCP's robust data storage and processing capabilities, Spotify was able to handle massive amounts of data, enhance its recommendation algorithms, and provide a seamless user experience.
2. **Home Depot**: Home Depot, a leading home improvement retailer, utilized GCP's data analytics and machine learning services to improve its inventory management and supply chain processes. By analyzing customer data and using machine learning algorithms, Home Depot optimized inventory levels, improved product availability, and enhanced customer satisfaction.
3. **Snap Inc.**: Snap Inc., the parent company of Snapchat, turned to GCP to handle its massive data requirements and support its real-time image and video processing needs. GCP's storage, analytics, and machine learning services enabled Snap Inc. to efficiently store, process, and analyze user-generated content, delivering an engaging and personalized experience to millions of users.
4. **Airbus**: Airbus, a leading aerospace manufacturer, partnered with Google Cloud to leverage its machine learning and data analytics capabilities. By utilizing GCP's tools

and technologies, Airbus enhanced aircraft maintenance and optimized fuel consumption. The collaboration also involved analyzing vast amounts of sensor data to improve flight operations and reduce costs.

5. **Niantic**: Niantic, the company behind popular augmented reality games like Pokémon GO, chose GCP to handle the massive scale and real-time demands of its gaming platforms. GCP's infrastructure and global reach provided Niantic with the necessary resources to deliver a seamless and immersive gaming experience to millions of players worldwide.

Q2. Case study of Public, Private, Hybrid Cloud.

Volkswagen Group, one of the world's largest automobile manufacturers, embarked on a digital transformation journey that involved implementing a hybrid cloud infrastructure to support its diverse IT needs. This case study showcases the successful adoption of a hybrid cloud model by Volkswagen Group.

Challenge:

Volkswagen Group faced the challenge of managing a complex IT landscape consisting of various applications, systems, and data across multiple brands and regions. They needed a flexible and scalable cloud solution that could accommodate their diverse requirements, including data privacy and compliance regulations, while ensuring cost efficiency and operational agility.

Solution:

Volkswagen Group opted for a hybrid cloud approach, combining both public and private cloud environments to address their specific needs. They partnered with several cloud providers, including AWS (public cloud) and OpenStack (private cloud), to establish a hybrid infrastructure.

Public Cloud: Volkswagen Group utilized the public cloud for certain non-sensitive workloads and applications that required scalability and cost optimization. They leveraged AWS's services to deploy applications, run simulations, and process large datasets. The public cloud provided on-demand scalability and access to a wide range of tools and services.

Private Cloud: Volkswagen Group maintained a private cloud environment built on the OpenStack platform. This private cloud facilitated the management of critical and sensitive applications, ensuring data privacy, compliance, and enhanced security controls. The private cloud offered greater control and customization options required for specific business requirements.

Hybrid Cloud Integration: Volkswagen Group implemented a robust hybrid cloud integration strategy to seamlessly connect and manage workloads across their public and private cloud environments. They utilized various tools and technologies, such as hybrid cloud management platforms and application programming interfaces (APIs), to enable data exchange, workload migration, and workload orchestration between the two cloud environments.

Benefits:

By adopting a hybrid cloud model, Volkswagen Group achieved several significant benefits:

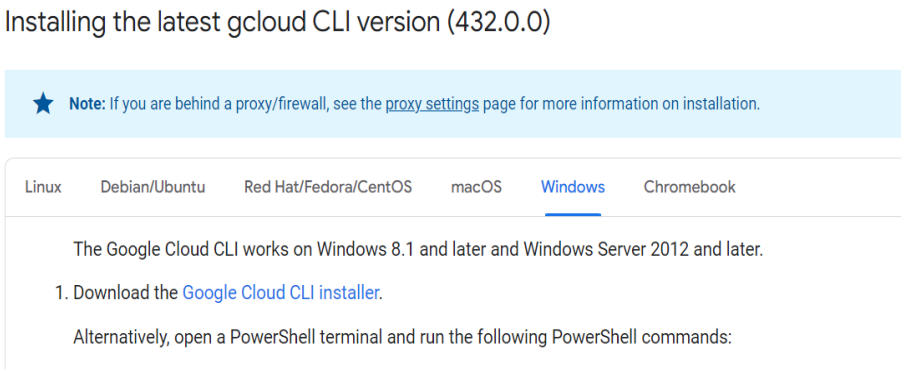
- **Flexibility and Scalability**: The hybrid cloud infrastructure allowed Volkswagen Group to scale resources up or down based on demand, ensuring optimal performance and cost efficiency for their applications and workloads.
- **Data Security and Compliance**: The private cloud component ensured that critical data and applications adhered to stringent security and compliance regulations, providing Volkswagen Group with the necessary control and governance over sensitive information.
- **Cost Optimization**: By utilizing the public cloud for non-sensitive workloads, Volkswagen Group optimized costs by paying only for the resources consumed. They could leverage the elasticity and cost-effective pricing models of the public cloud, resulting in potential cost savings.
- **Operational Efficiency**: The hybrid cloud approach enhanced operational efficiency by providing a unified management and deployment framework. Volkswagen Group could deploy workloads seamlessly across public and private clouds, enabling agility and reducing time-to-market for new applications and services.

This case study exemplifies how Volkswagen Group successfully implemented a hybrid cloud strategy, leveraging the benefits of both public and private cloud environments. The hybrid cloud infrastructure provided the necessary agility, scalability, security, and compliance capabilities required by the automotive giant, supporting their digital transformation initiatives, and ensuring efficient management of their IT ecosystem.

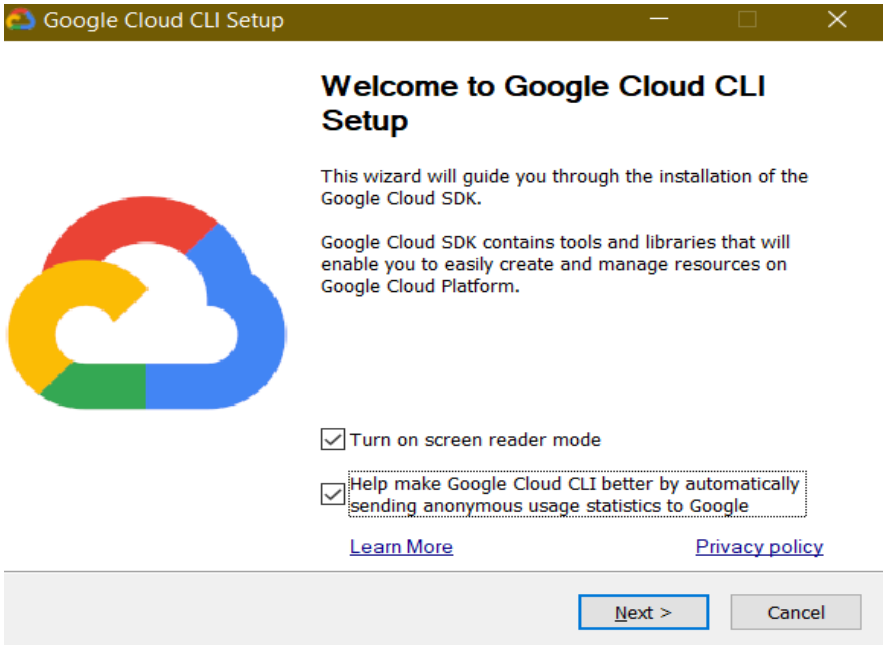
Q3. Install Google App Engine. Create a Hello world app and other simple web applications using python/java.

Procedure:

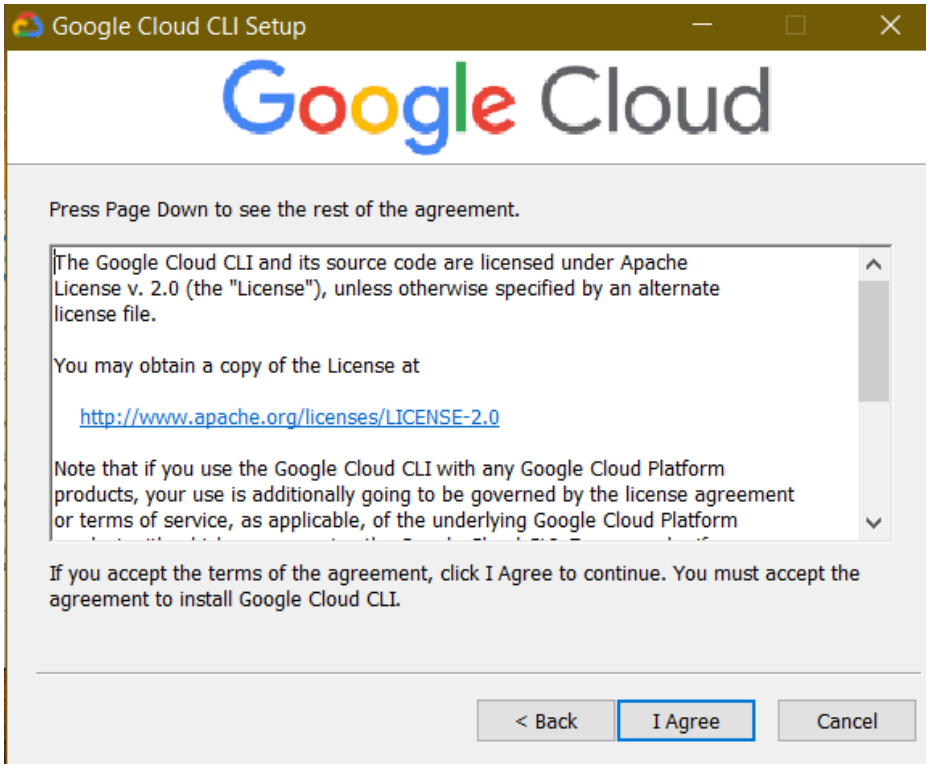
Step 1: Download Google Cloud SDK from the official website of Google.



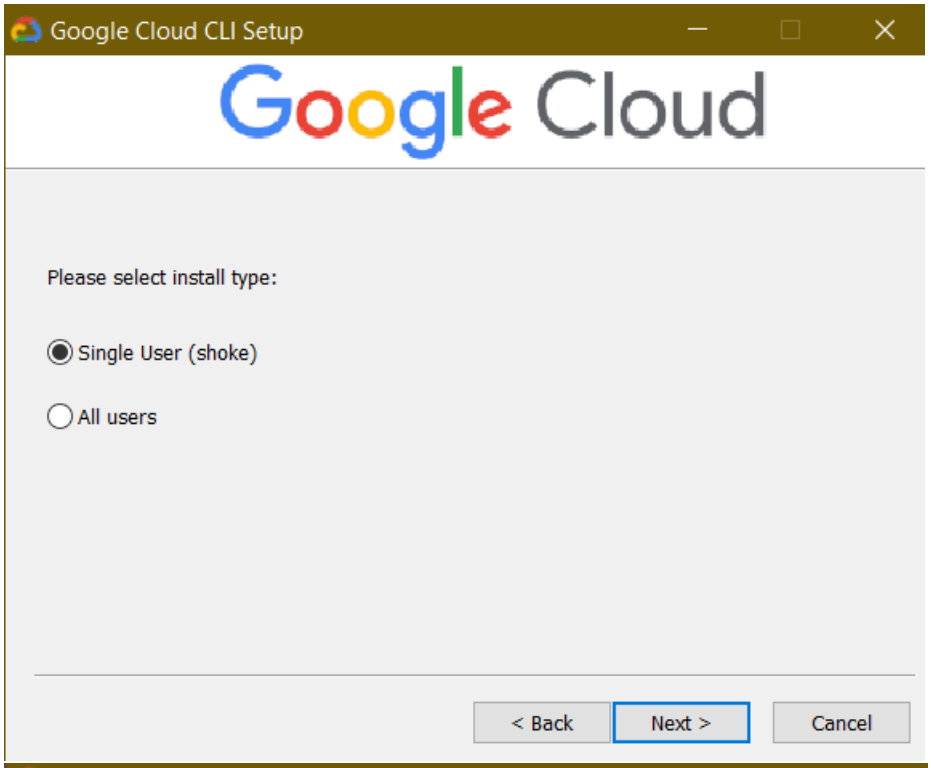
Step 2: Open the install wizard of the Google SDK/CLI(Command Line Interface)

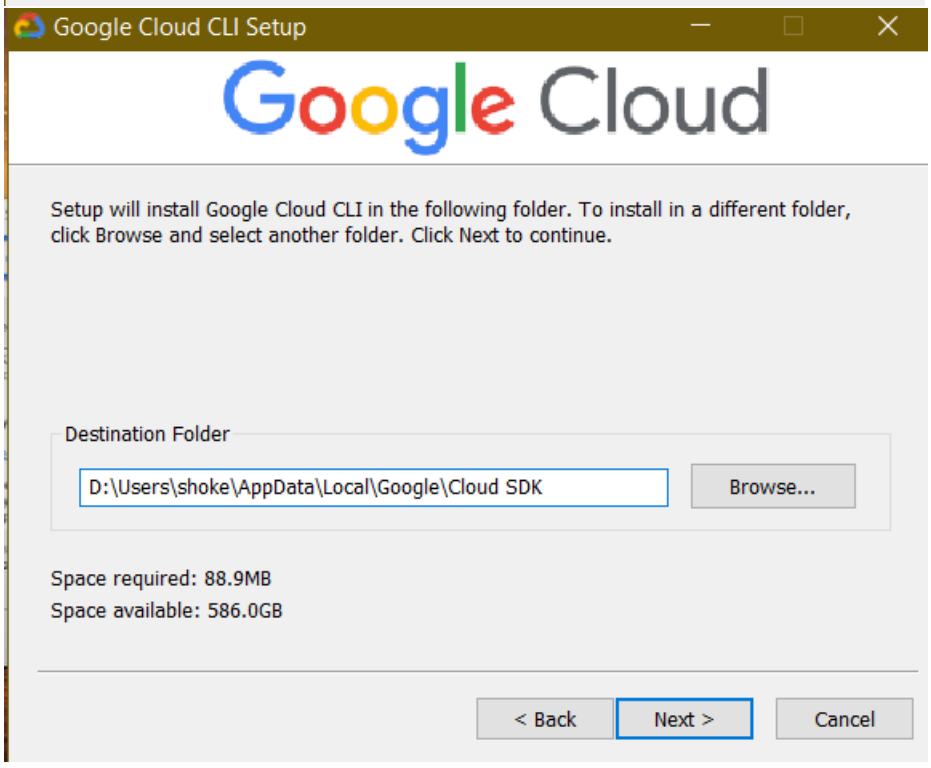


Step 3: Agree to the steps.

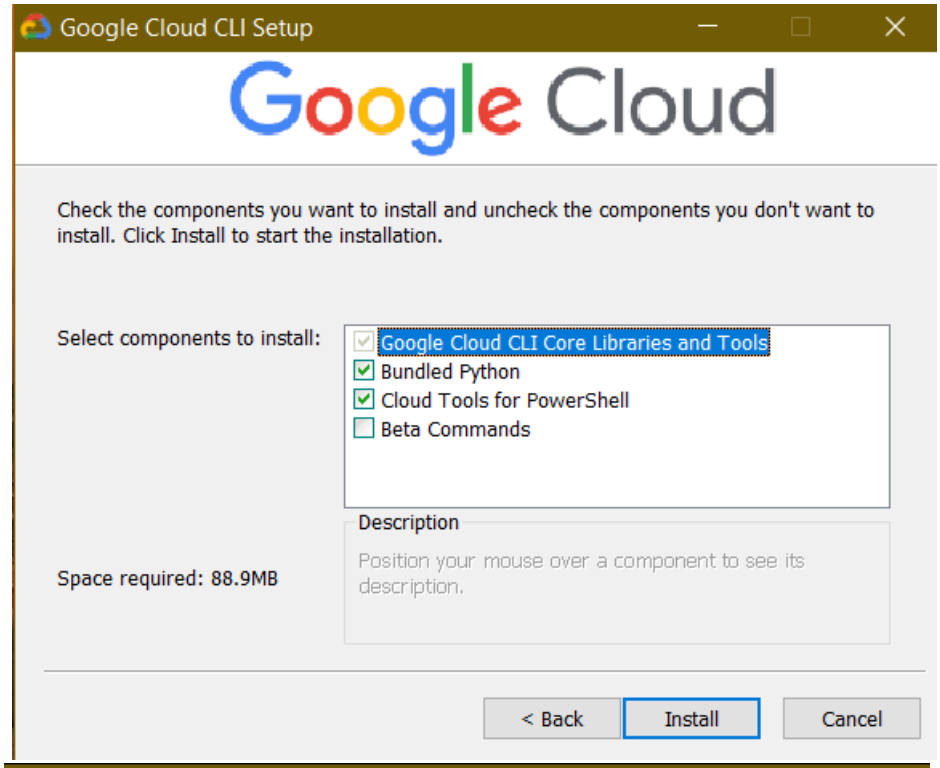


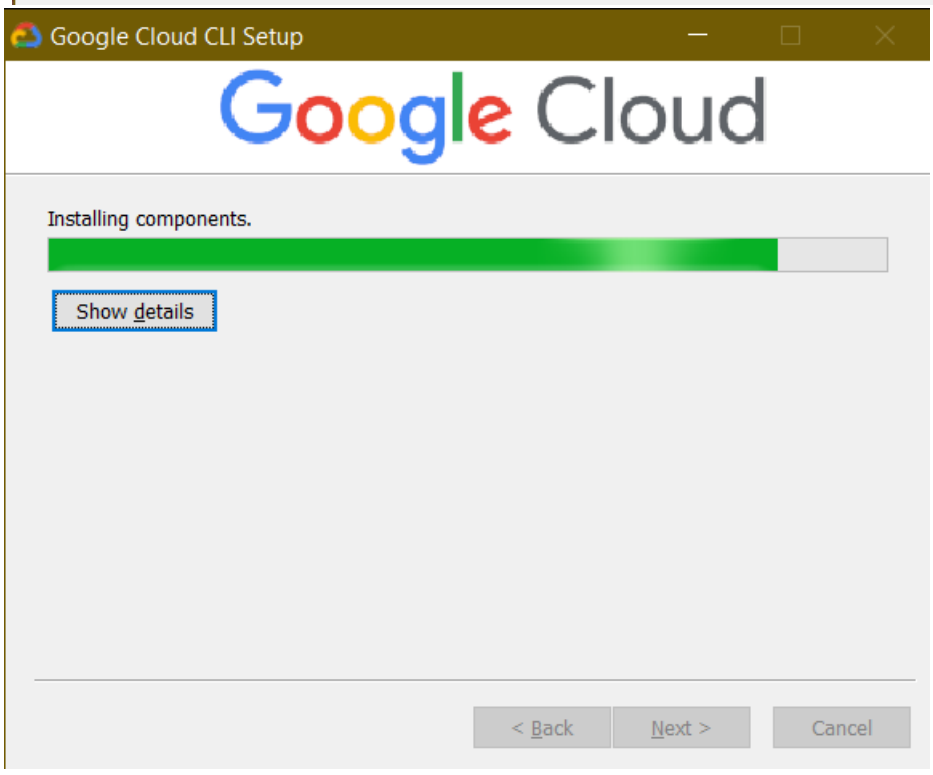
Step 4:


Google Cloud CLI Setup
Google Cloud
Please select install type:
Single User (shoke)
All users
< Back
Next >
Cancel

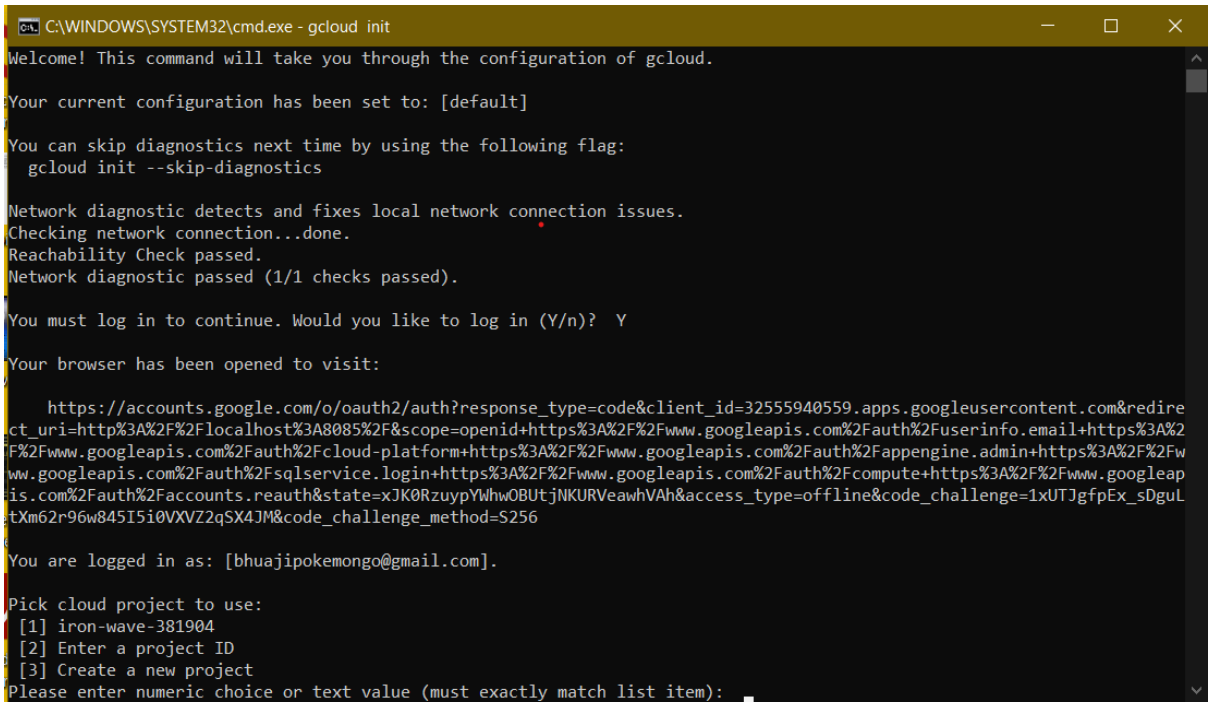

Google Cloud CLI Setup
Google Cloud
Setup will install Google Cloud CLI in the following folder. To install in a different folder, click Browse and select another folder. Click Next to continue.
Destination Folder
D:\Users\shoke\AppData\Local\Google\Cloud SDK
Browse...
Space required: 88.9MB
Space available: 586.0GB
< Back
Next >
Cancel

Step 5:

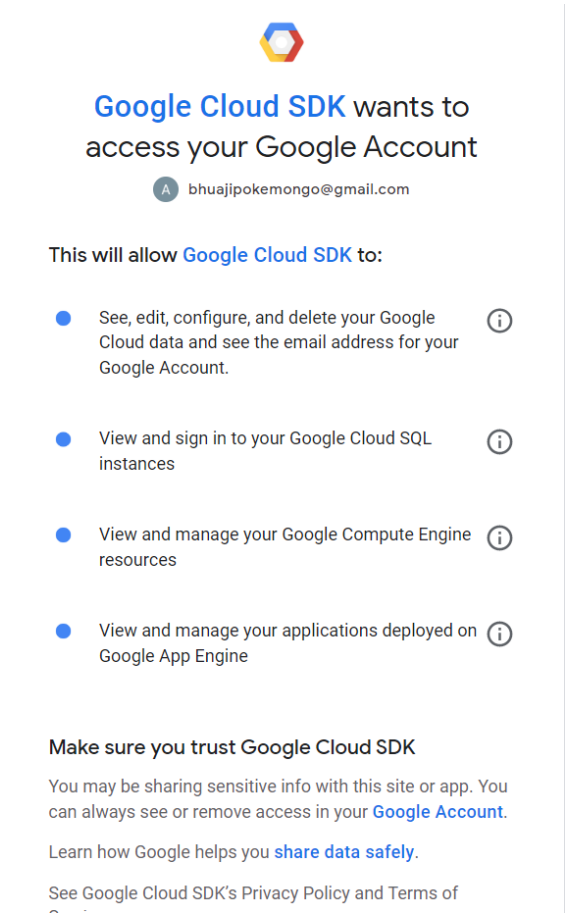

Google Cloud CLI Setup
Google Cloud
Check the components you want to install and uncheck the components you don't want to install. Click Install to start the installation.
Select components to install:
Google Cloud CLI Core Libraries and Tools
Bundled Python
Cloud Tools for PowerShell
Beta Commands
Description
Position your mouse over a component to see its description.
Space required: 88.9MB
< Back
Install
Cancel


Google Cloud CLI Setup
Google Cloud
Installing components.
Show details
< Back
Next >
Cancel

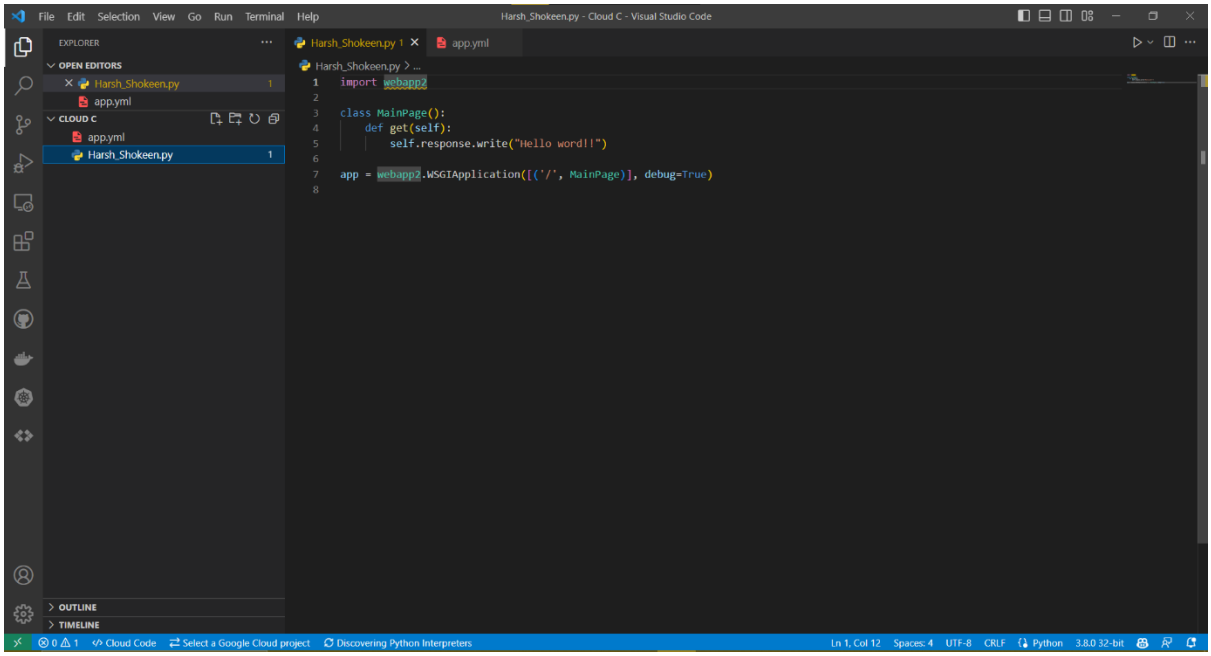
Step 6: After the installation of the Google Cloud SDK; Open the GCSDK and Accept the Login from your default account.



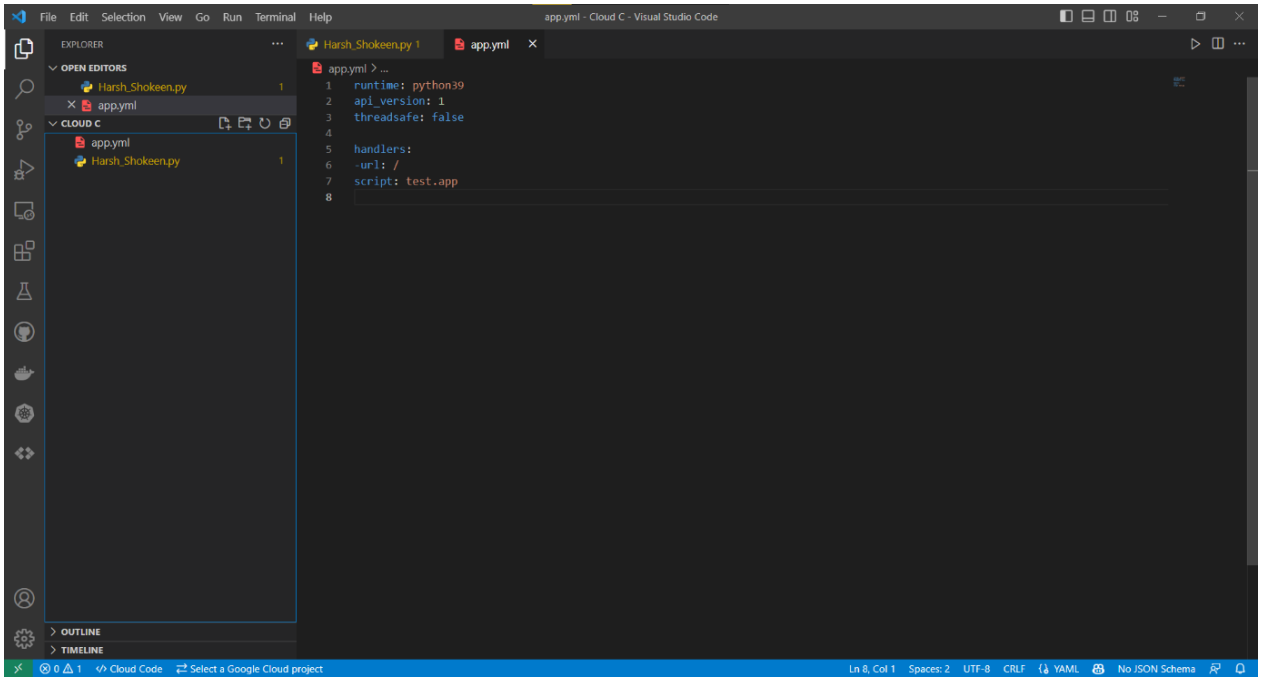
Step 7: Accept the login from your default account.



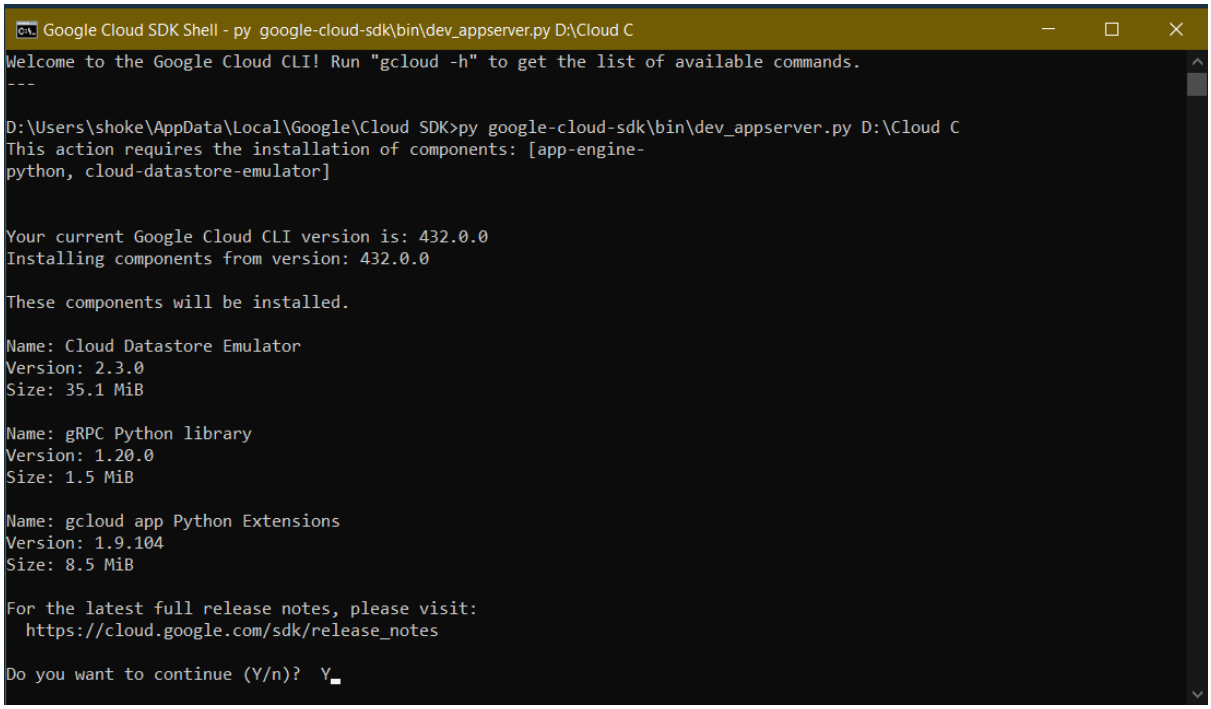
Step 8: Write the code for the Webapp in VScode



Step 9: Write the code for the YAML file.



Step 10: Open the code/folder in GCSDK and run on Google App Engine

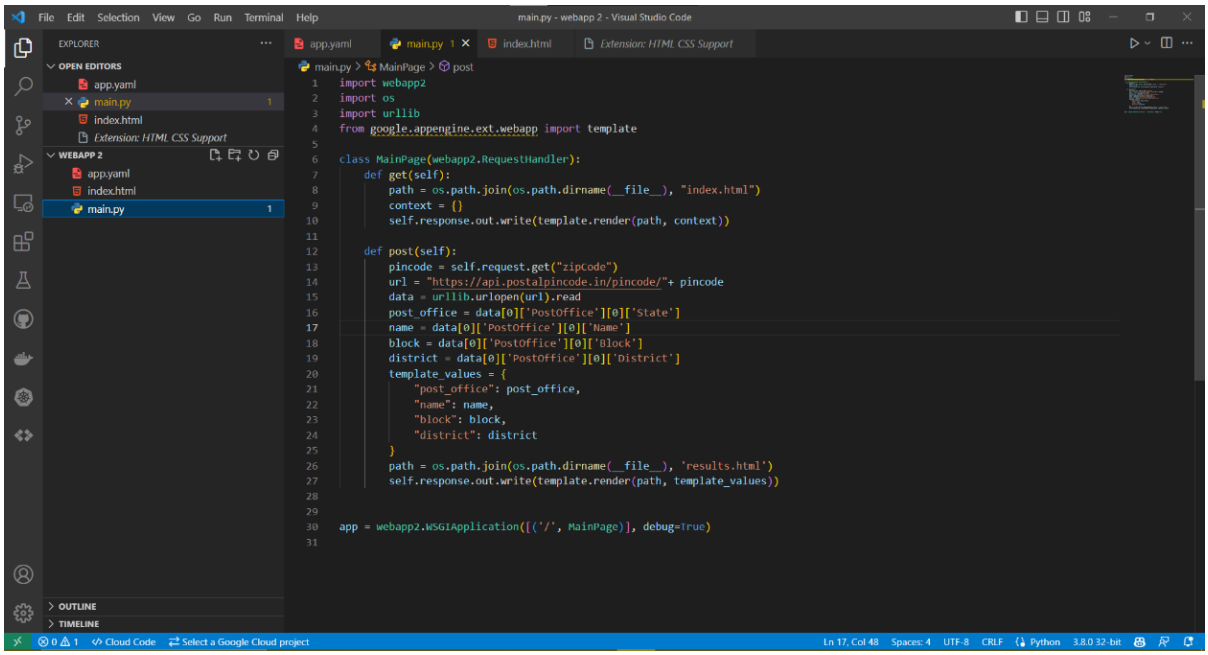


OUTPUT:

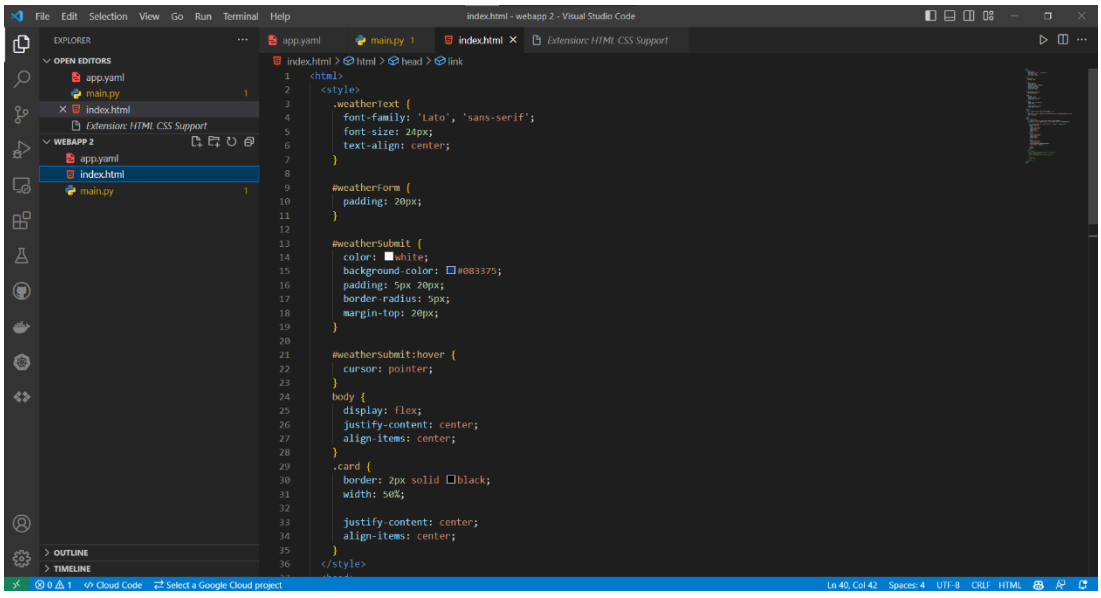
Q4. Use GAE launcher to launch the web applications.

PROCEDURE:

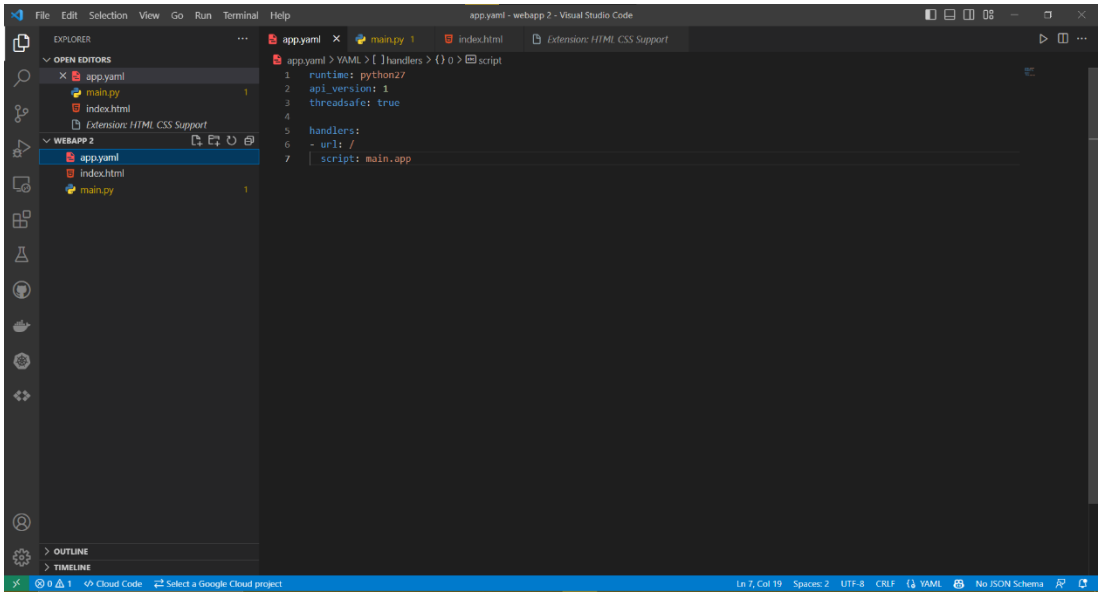
Step 1: Write the Python code for your web application; for me it is about the fetching data of the corresponding PINCODE.



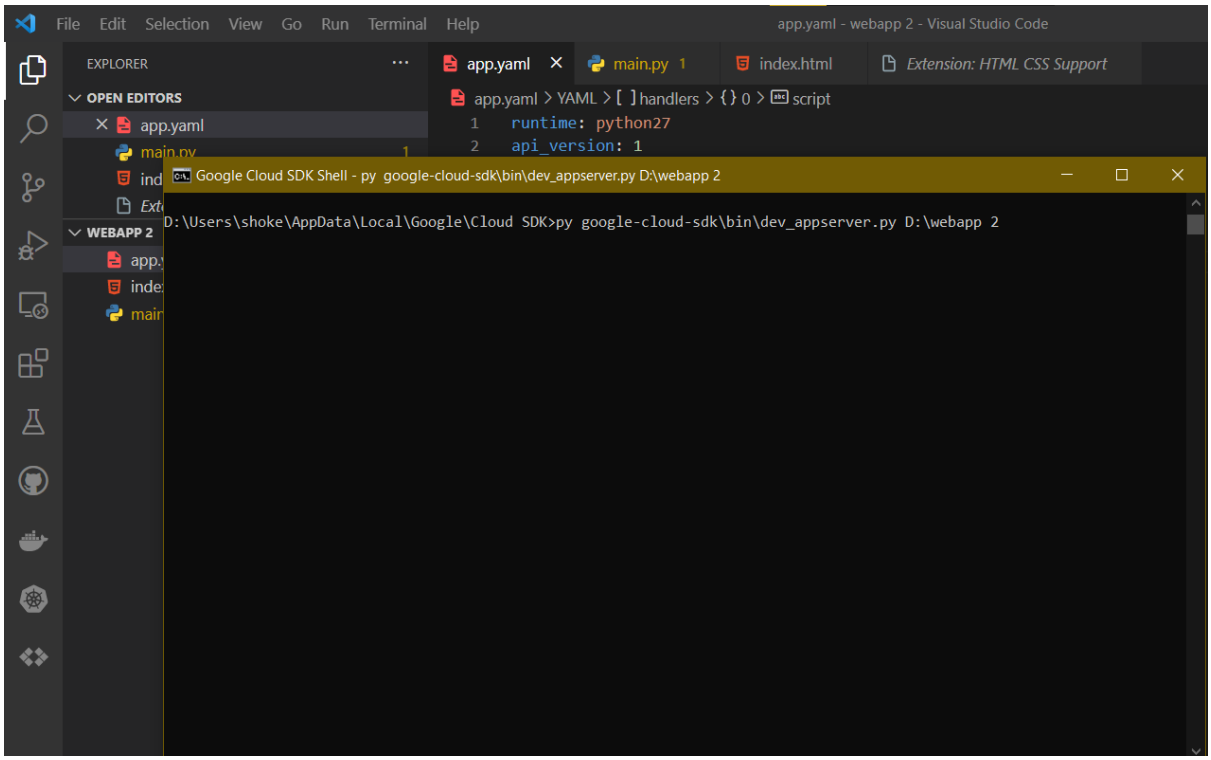
Step 2: Write the subsequent HHTML CODE for it specifying all the details you want to add.



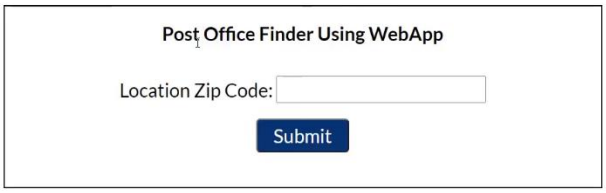
Step 3: Write the corresponding YAML for it specifying threads, APIs, and script.



Step 4: Locate the file where your appserver.py is and run with the location where your project is in Google SDK Shell and run the command.



OUTPUT:

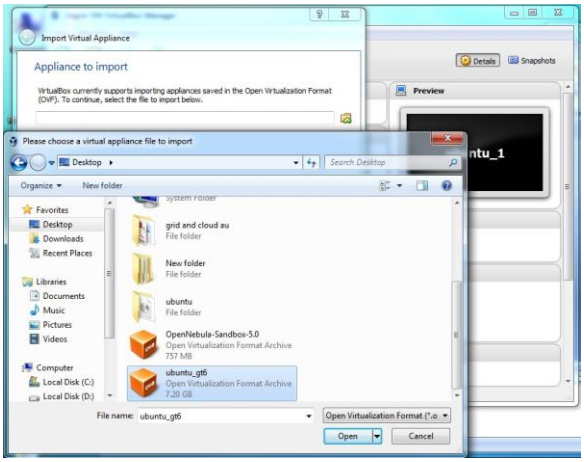


Q5. Install a C compiler in the virtual machine created using virtual box and execute Simple Programs.

PROCEDURE:

Steps to import .ova file:

1. Open Virtual box
2. File import Appliance
3. Browse ubuntu_gt6.ova file
4. Then go to setting, select Usb and choose USB 1.1
5. Then start the ubuntu_gt6
6. Login using username: Harsh, password:1912730002.

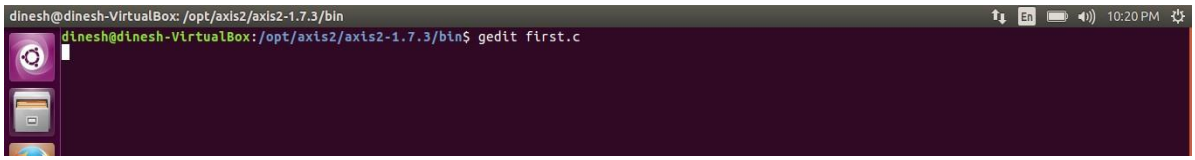


Steps to run c program:

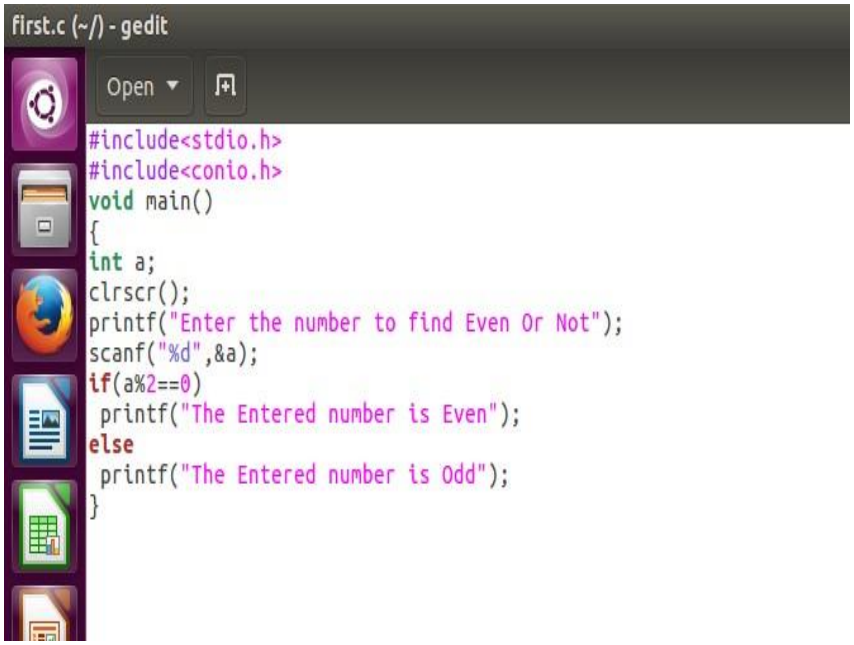
1. Open the terminal.
2. Type cd /opt/axis2/axis2-1.7.3/bin then press enter.
3. gedit hello.c
4. gcc hello.c

Step 1.Type cd /opt/axis2/axis2-1.7.3/bin then press enter.

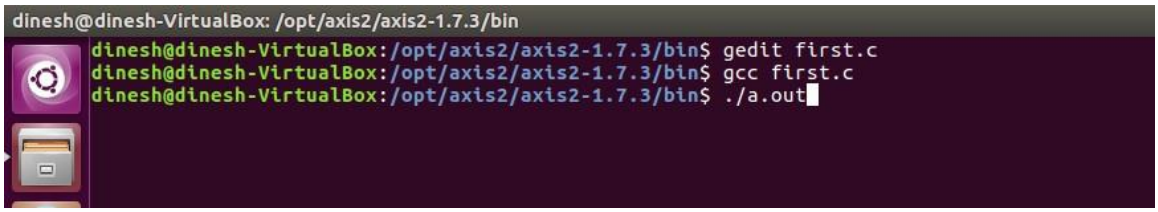
Step 2: Type gedit first.c



Step 3: Type the C Program

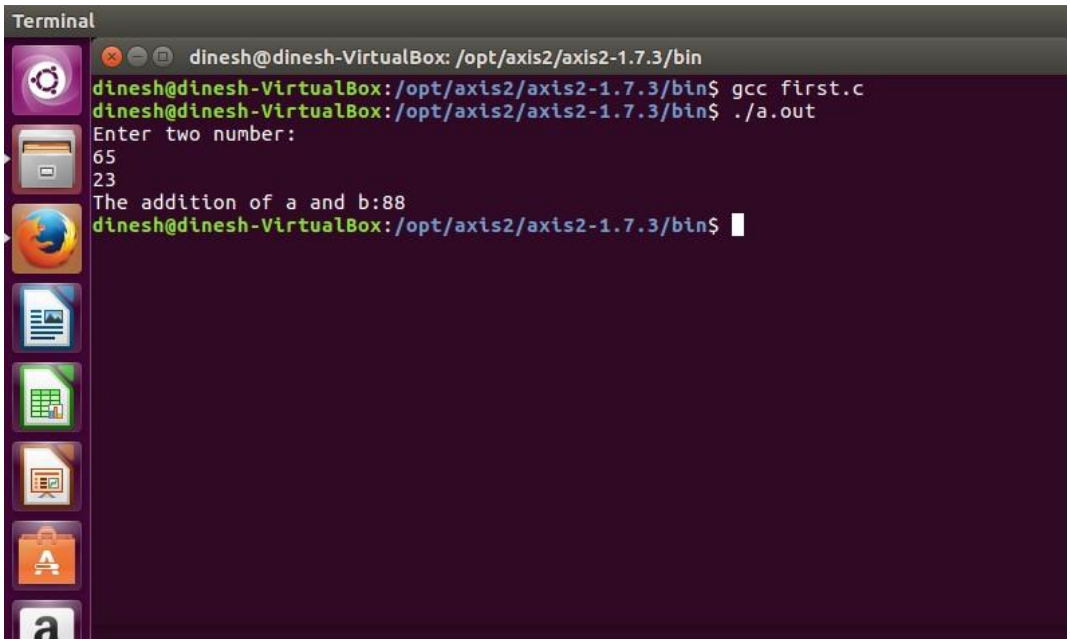


Step 4: Running the C Program



Step 5: Display The output.

RESULT:



Q6. Find a procedure to transfer the files from one virtual machine to another virtual machine.

Steps:

1. For this you need to share clipboard between host OS and guest OS, installing Guest Addition on both the virtual machines (probably setting bidirectional and restarting them). You copy from guest OS in the clipboard that is shared with the host OS.
2. Then you paste from the host OS to the second guest OS.
3. You can enable drag and drop too with the same method (Click on the machine, settings, general, advanced, drag and drop set to bidirectional)
4. You can have common Shared Folders on both virtual machines and use one of the directories shared as buffer to copy.
5. Installing Guest Additions, you have the possibility to set Shared Folders too. As you put a file in a shared folder from host OS or from guest OS, is immediately visible to the other. (Keep in mind that can arise some problems for date/time of the files when there are different clock settings on the different virtual machines).
6. If you use the same folder shared on more machines, you can exchange files directly copying them in this folder.
5. You can use the usual method to copy files between 2 different computers with a client-server application. (e.g., scp with sshd active for linux, winscp... you can get some info about SSH servers e.g., here)
7. You need an active server (sshd) on the receiving machine and a client on the sending machine. Of course you need to have the authorization setted (via password or, better, via an automatic authentication method).
8. Note: many Linux/Ubuntu distribution install sshd by default: you can see if it is running with pgrep sshd from a shell. You can install with sudo apt-get install openssh-server.
6. You can mount part of the file system of a virtual machine via NFS or SSHFS on the other, or you can share file and directory with Samba. You may find interesting the article Sharing files between guest and host without VirtualBox shared folders with detailed step by step instructions.

You should remember that you are dealing with a little network of machines with different operative systems, and in particular:

- Each virtual machine has its own operating system running on and acts as a physical machine.
- Each virtual machine is an instance of a program owned by a user in the hosting operating system and should undergo the restrictions of the user in the hosting OS.

  E.g., Let we say that Hastur and Meow are users of the hosting machine, but they did not allow each other to see their directories (no read/write/execute authorization). When each of them runs a virtual machine, for the hosting OS those virtual machines

are two normal programs owned by Hastur and Meow and cannot see the private directory of the other user. This is a restriction due to the hosting OS. It's easy to overcome it: it's enough to give authorization to read/write/execute to a directory or to choose a different directory in which both users can read/write/execute.

Windows likes mouse and Linux fingers. :-)

I mean I suggest you enable Drag & drop to be cozy with the Windows machines and the Shared folders or to be cozy with Linux.

When you will need to be fast with Linux you will feel the need of ssh-keygen and

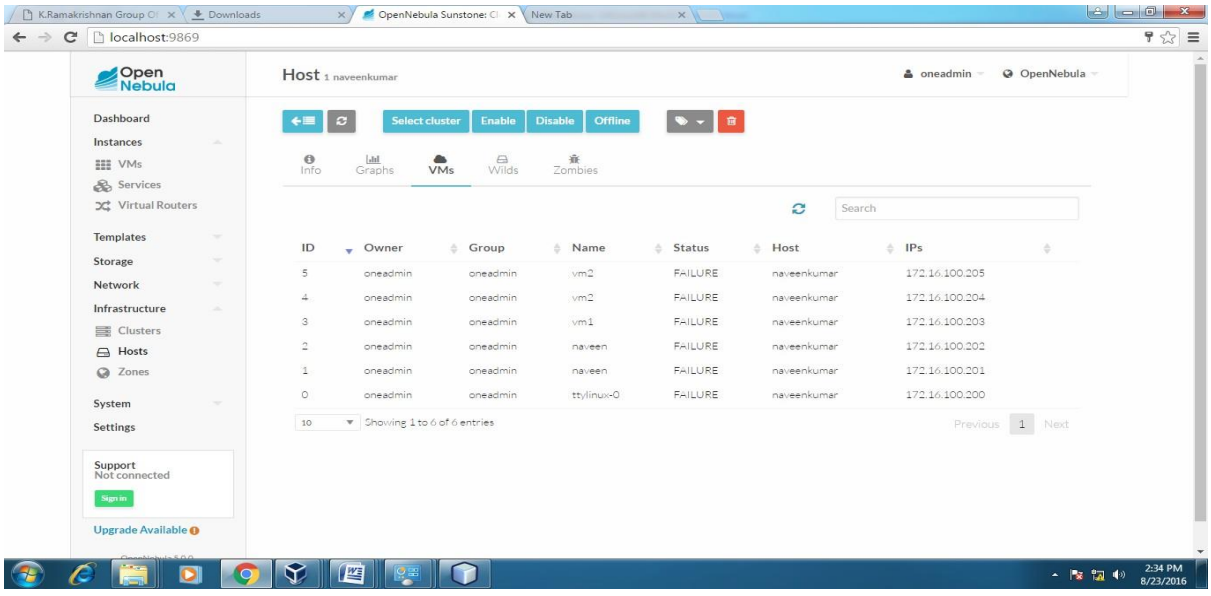
to Generate once SSH Keys to copy files on/from a remote machine without writing password anymore. In this way it functions bash auto-completion remotely too!

PROCEDURE:

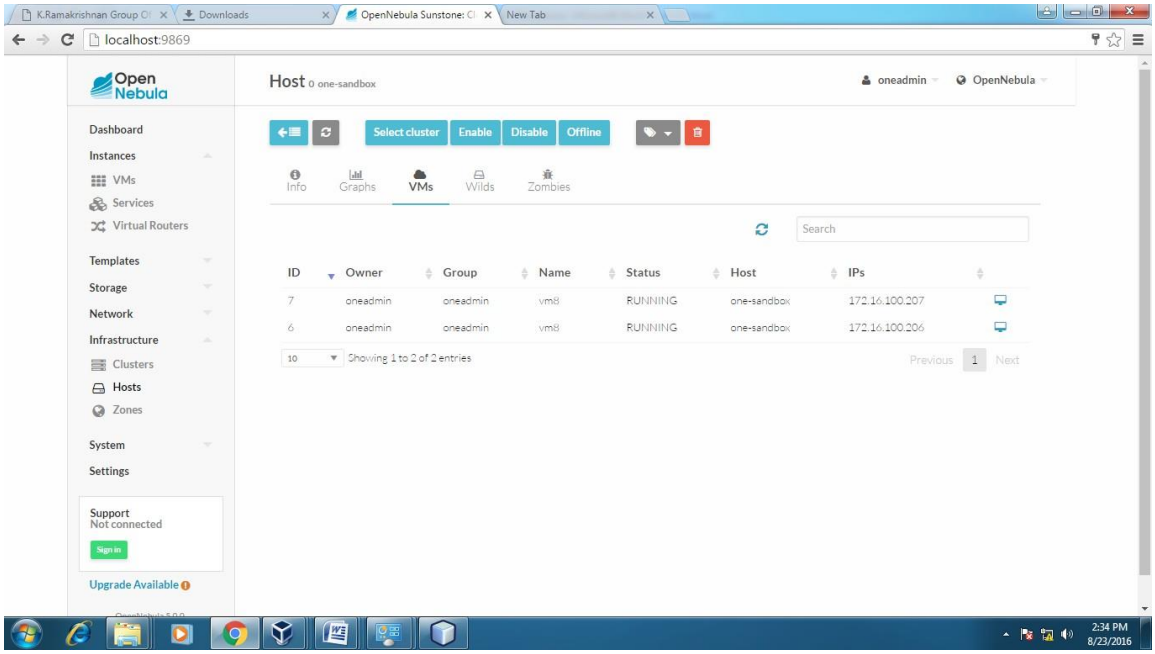
Steps:

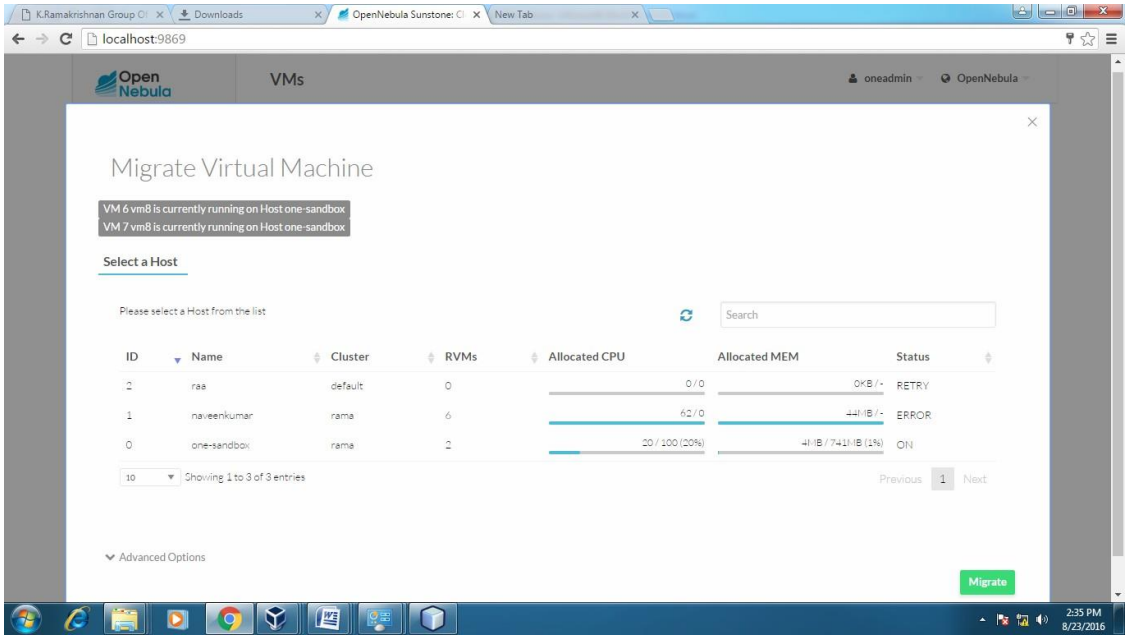
1. Open Browser, type localhost:9869
2. Login using username: oneadmin, password: opennebula
3. Then follow the steps to migrate VMs.
   a. Click on infrastructure.
   b. Select clusters and enter the cluster name.
   c. Then select host tab and select all host.
   d. Then select Vnets tab, and select all vnet
   e. Then select the datastores tab and select all datastores.
   f. And then choose host under infrastructure tab.
   g. Click on + symbol to add new host, name the host then click on create.
4. on instances, select VMs to migrate then follow the steps
   a. Click on the 8th icon ,the drop-down list display.
   b. Select migrate on that ,the popup window display.
   c. On that select the target host to migrate then click on migrate.

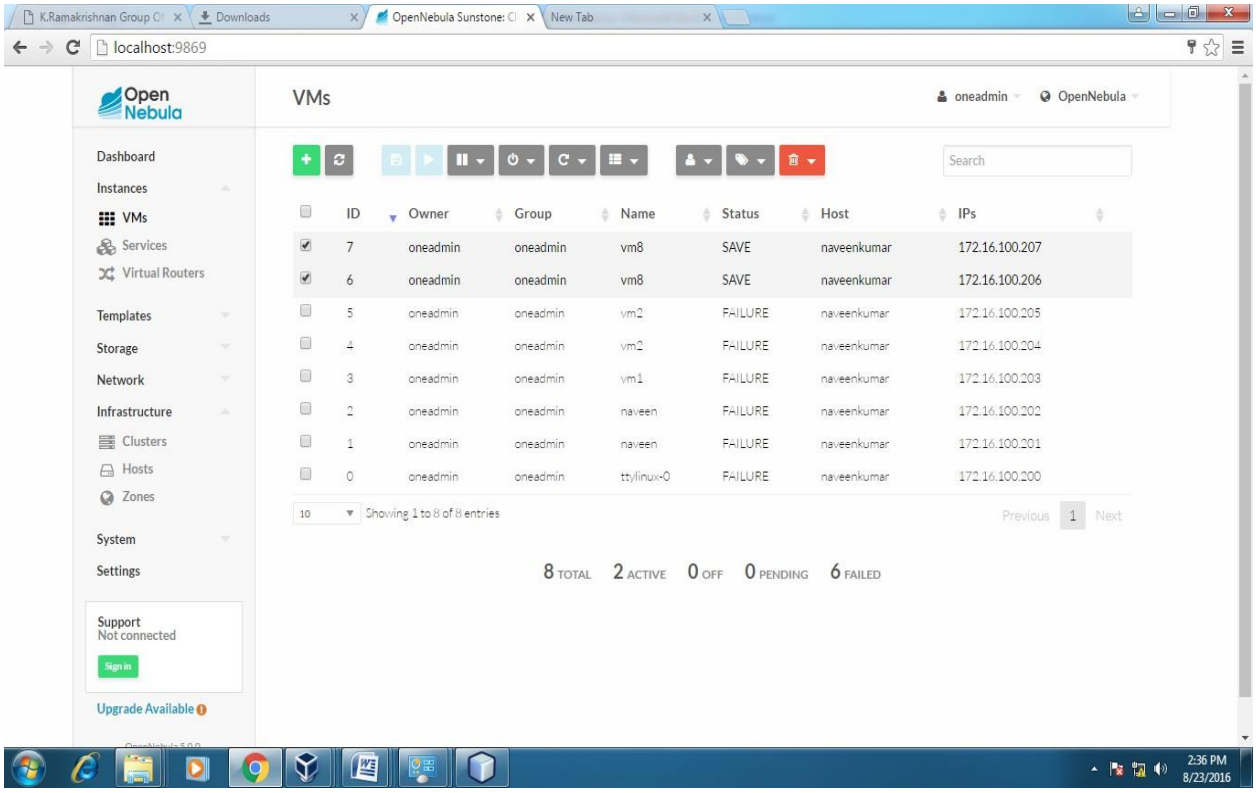
Before Migration



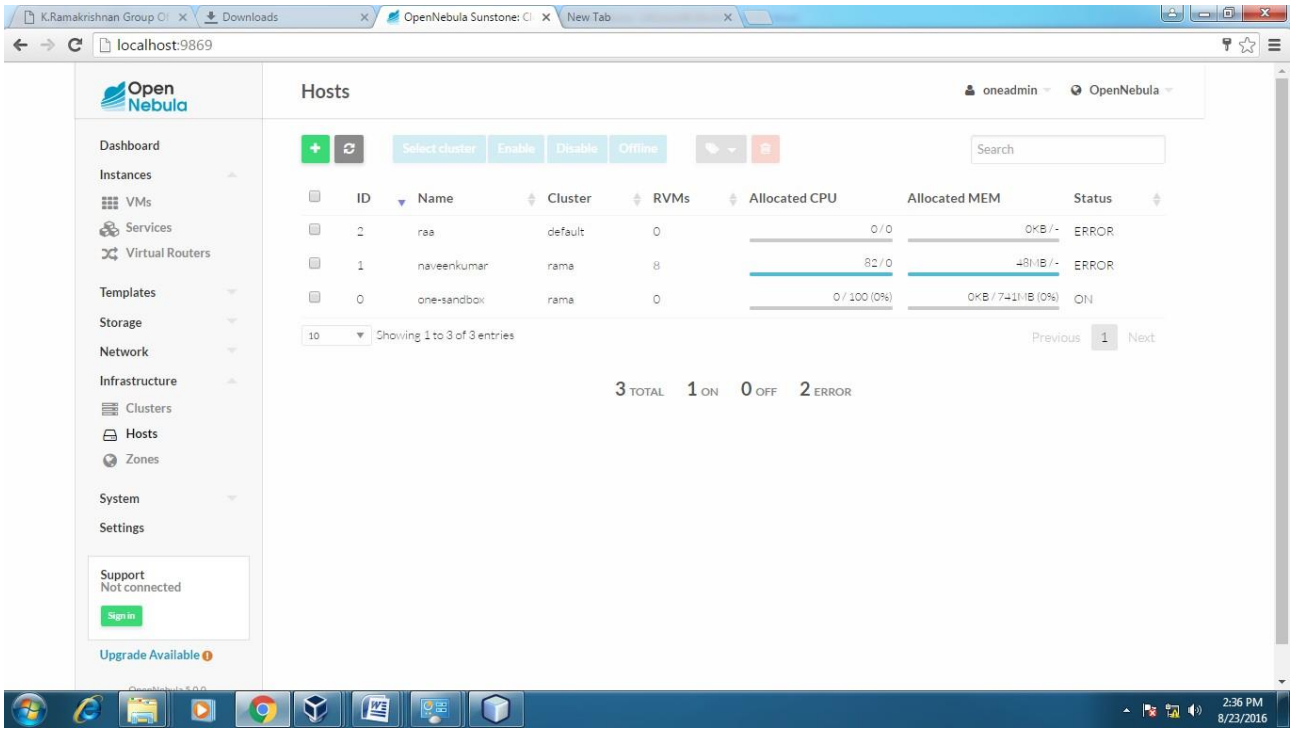
Host: one_sandbox



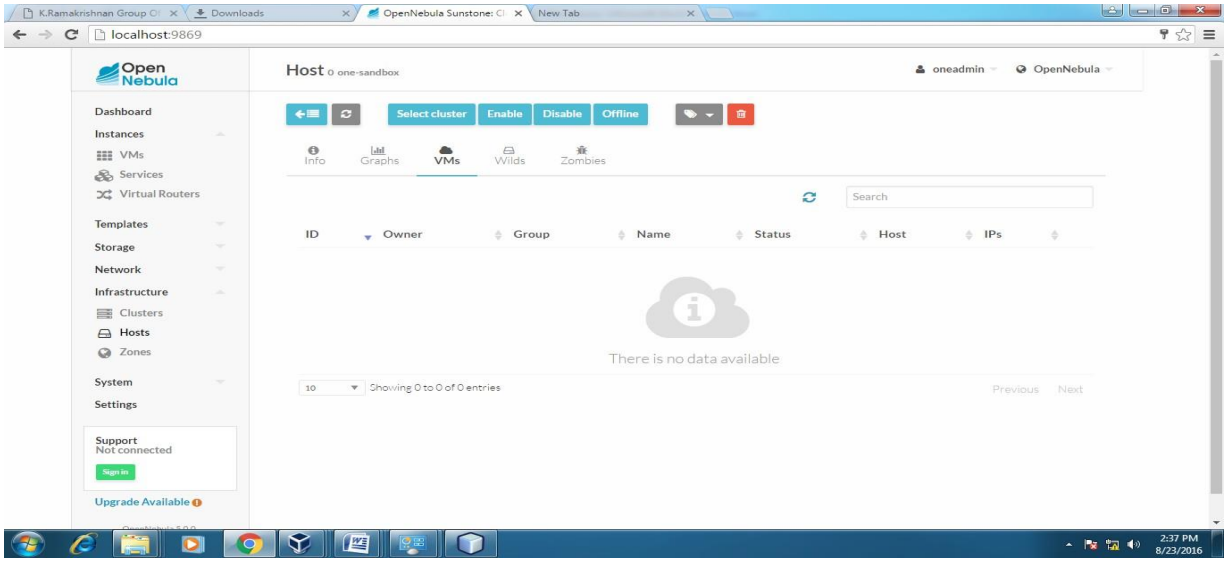




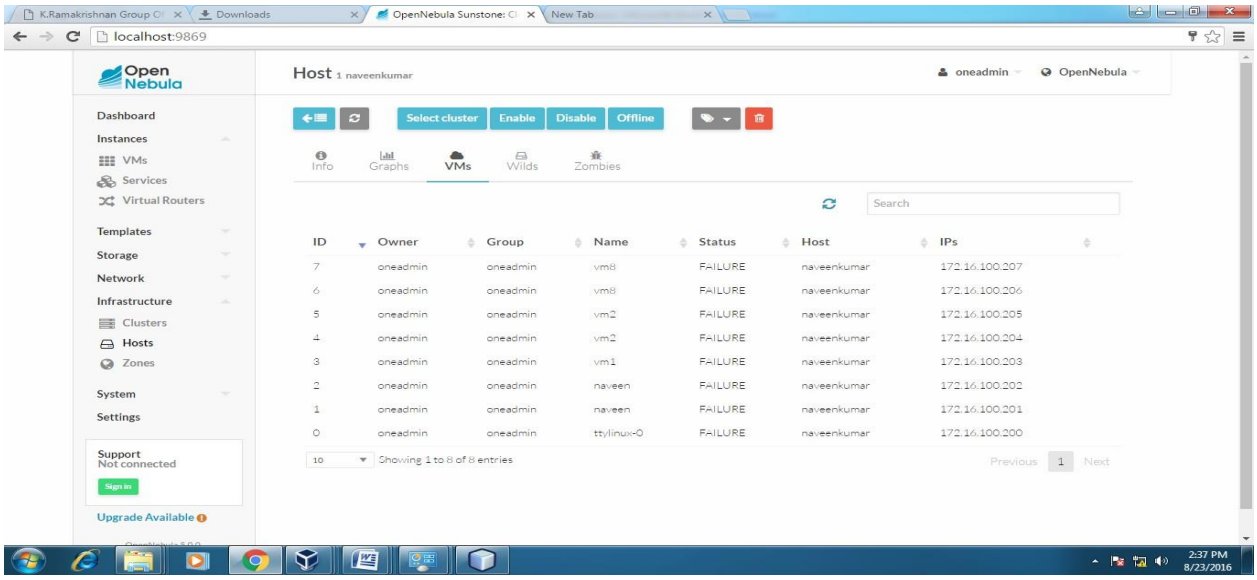
After Migration



Host: one_sandbox



Host:SACET



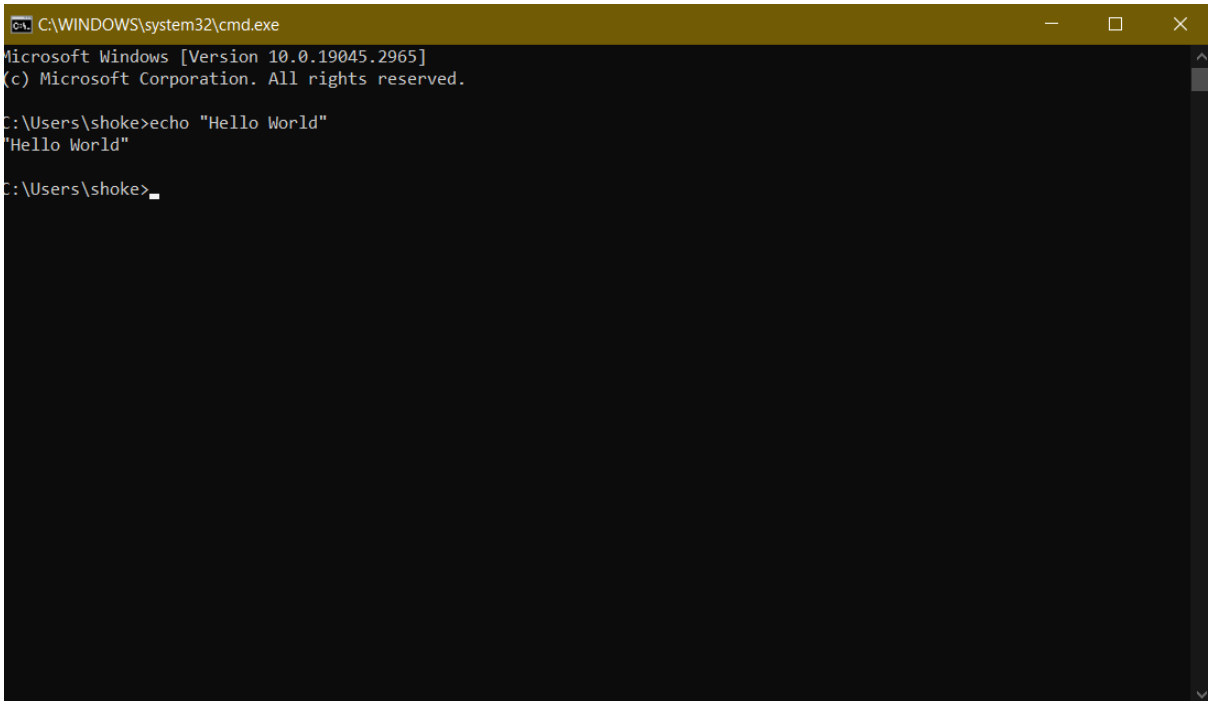
Result:

Thus, the file transfer between VM was successfully completed

Q7. Bash Command Execution

Here is a list of some of the most used bash commands:

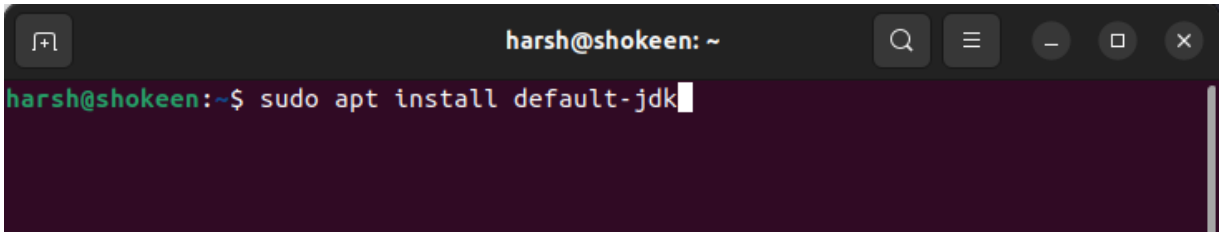
cd: Change the directory to a different location.
ls: List the contents of the current directory.
mkdir: Create a new directory.
touch: Create a new file.
rm: Remove a file or directory.
cp: Copy a file or directory.
mv: Move or rename a file or directory.
echo: Print text to the terminal.
cat: Concatenate and print the contents of a file.
grep: Search for a pattern in a file.
chmod: Change the permissions of a file or directory.
sudo: Run a command with administrative privileges.
df: Display the amount of disk space available.
history: Show a list of previously executed commands.
ps: Display information about running processes.



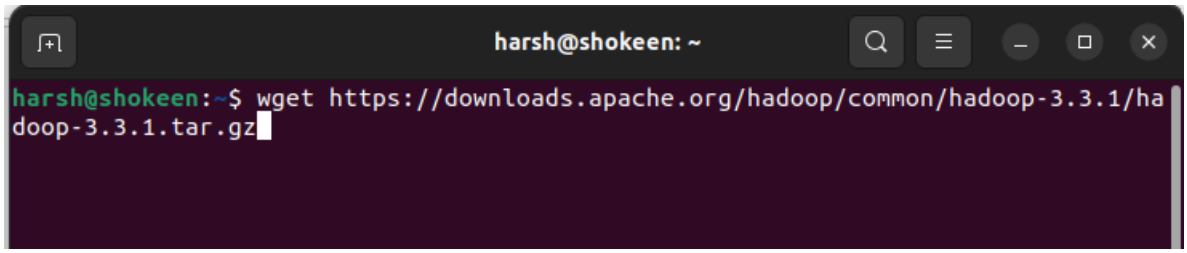
Q8. Install Hadoop single node cluster and run simple applications like wordcount.

Procedure:

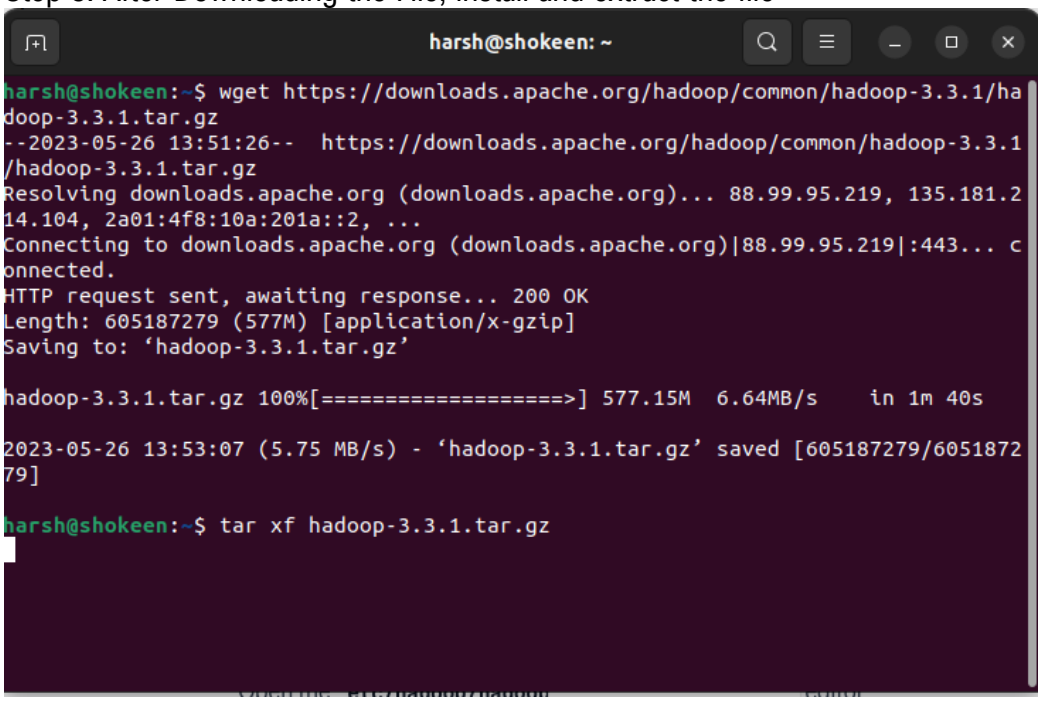
Step 1: Install JDK first in terminal by writing the following command: **sudo apt install openjdk-8-jdk**



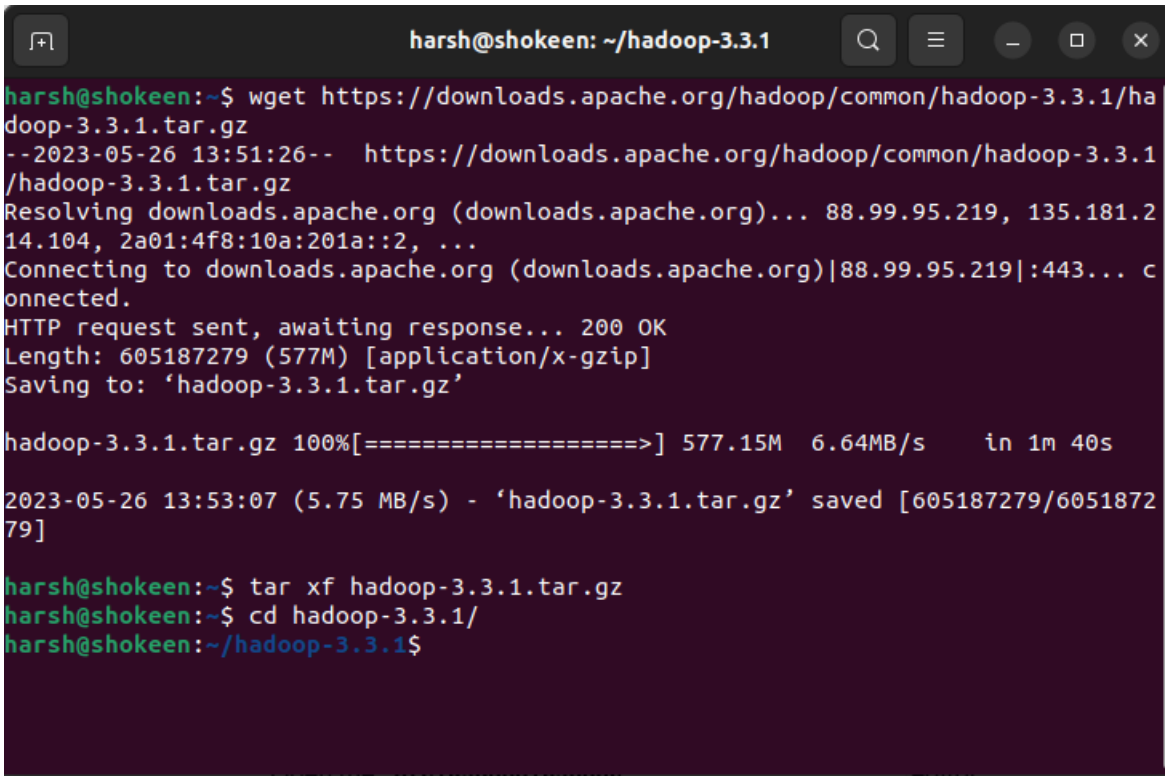
Step 2: Download Hadoop to your system using this command.



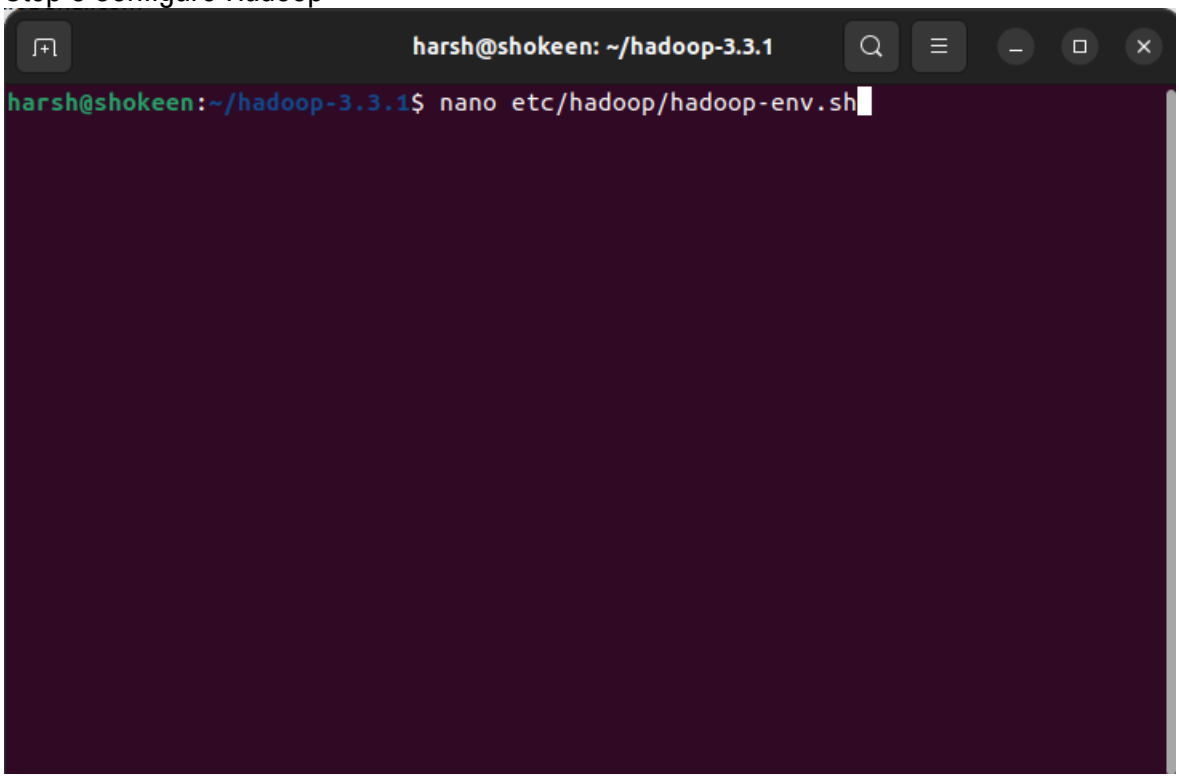
Step 3: After Downloading the File, install and extract the file



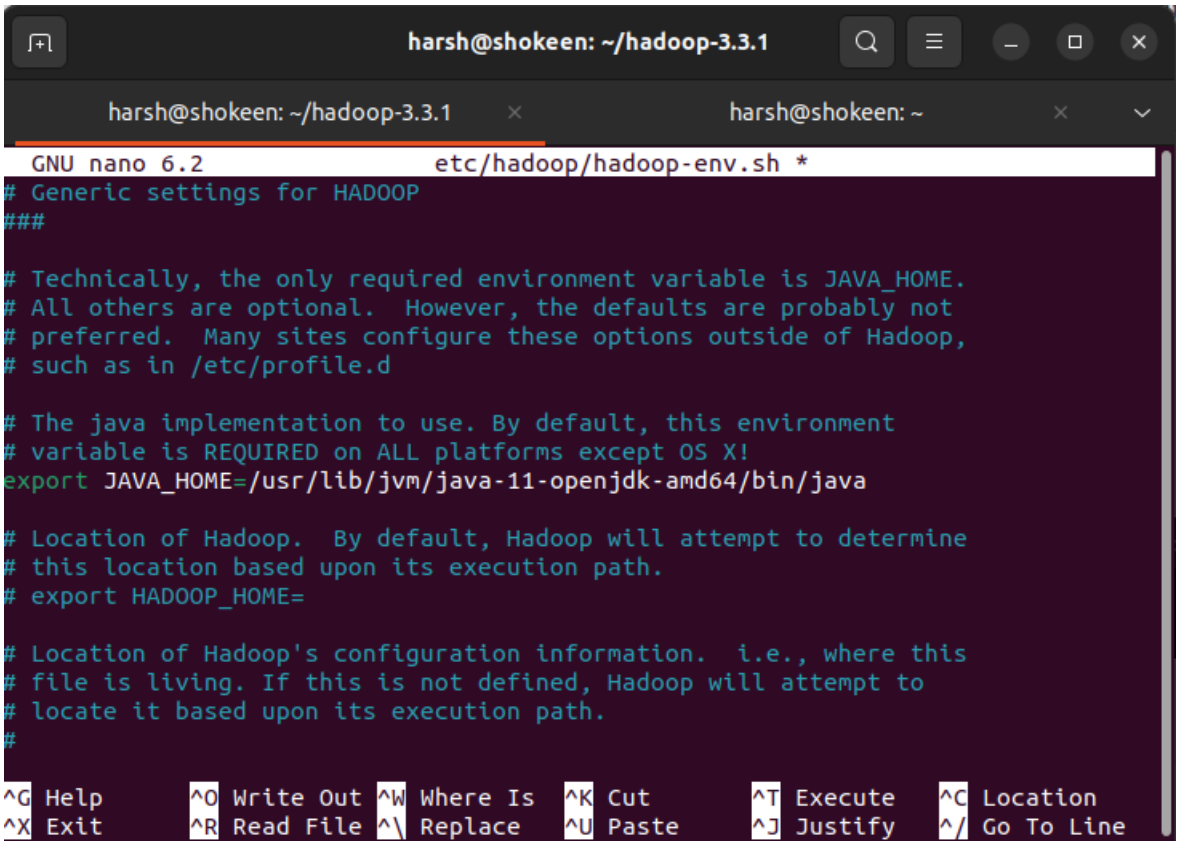
Step 4: Open the file location using cd to navigate to that folder.



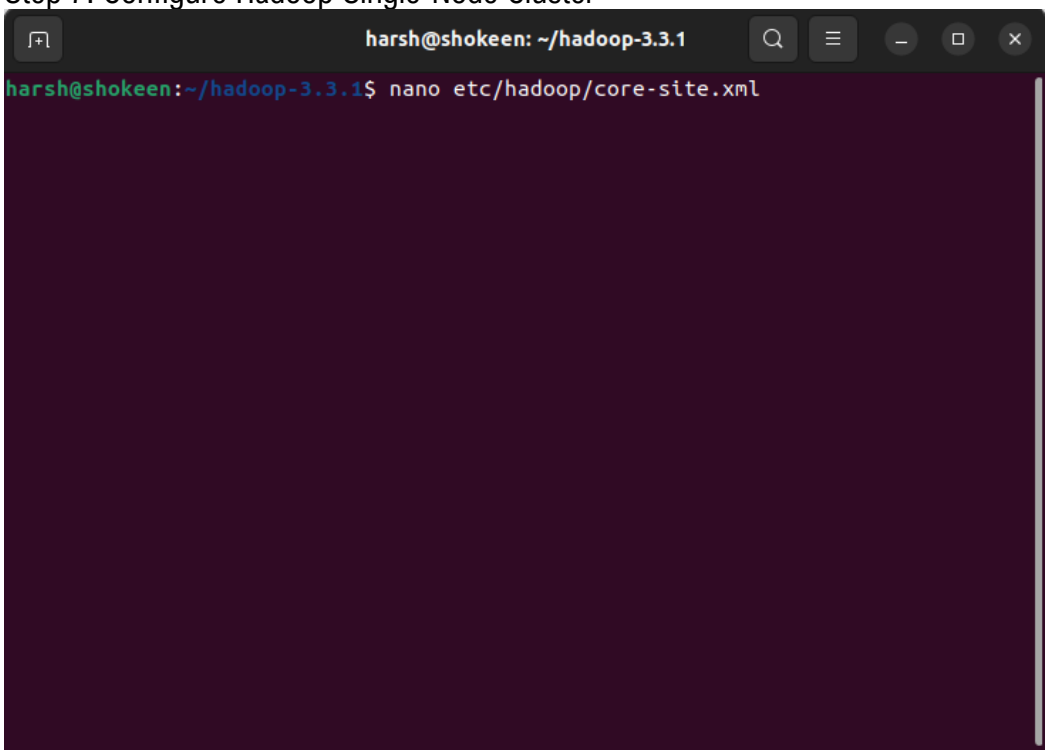
Step 5 Configure Hadoop



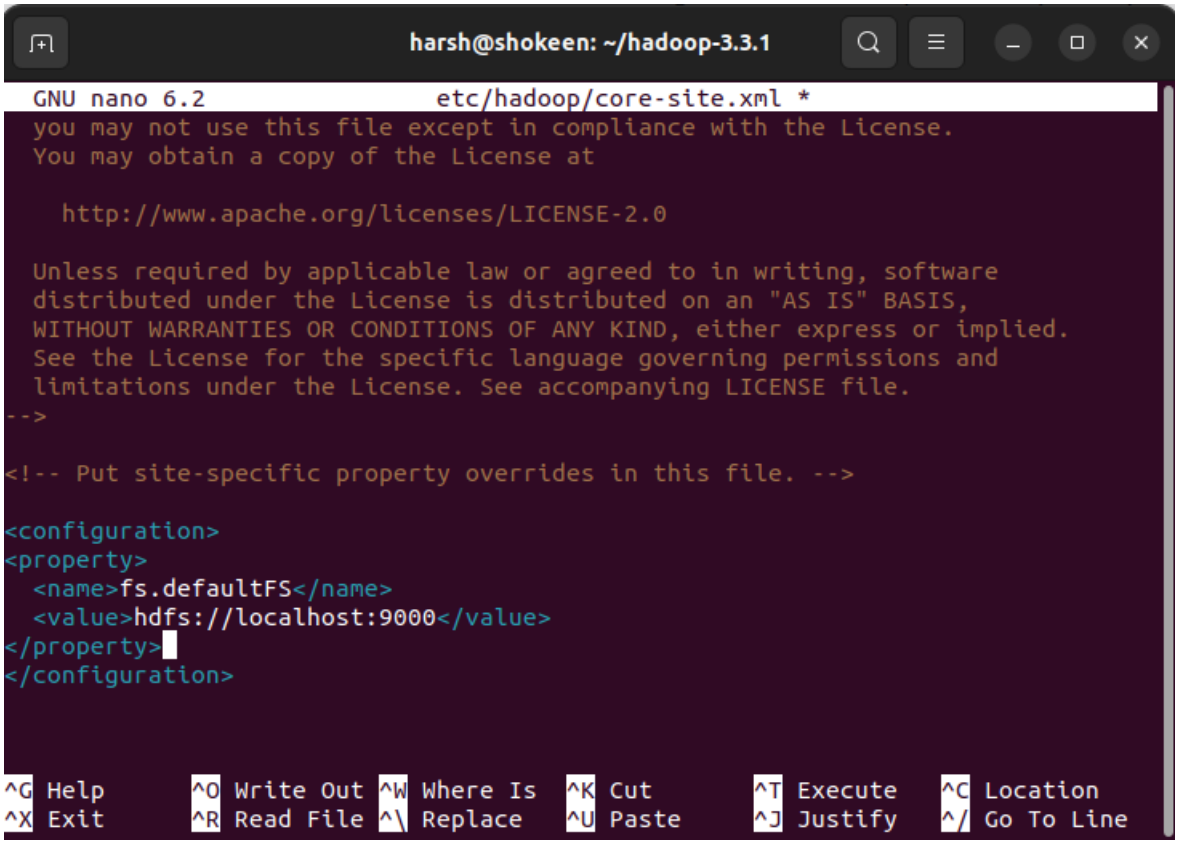
Step 6: Uncomment the line that sets the JAVA_HOME variable and provide the path to your JDK installation. Save the file and exit the text editor. This is to enable the path for JDK to function in Hadoop.



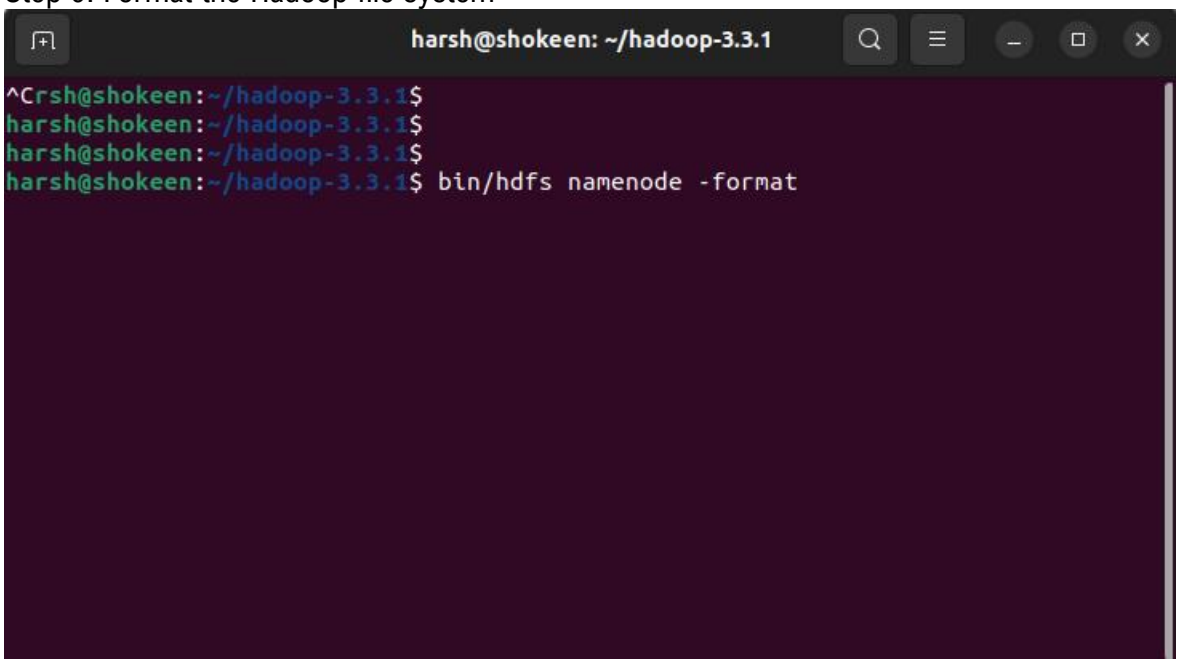
Step 7: Configure Hadoop Single-Node Cluster



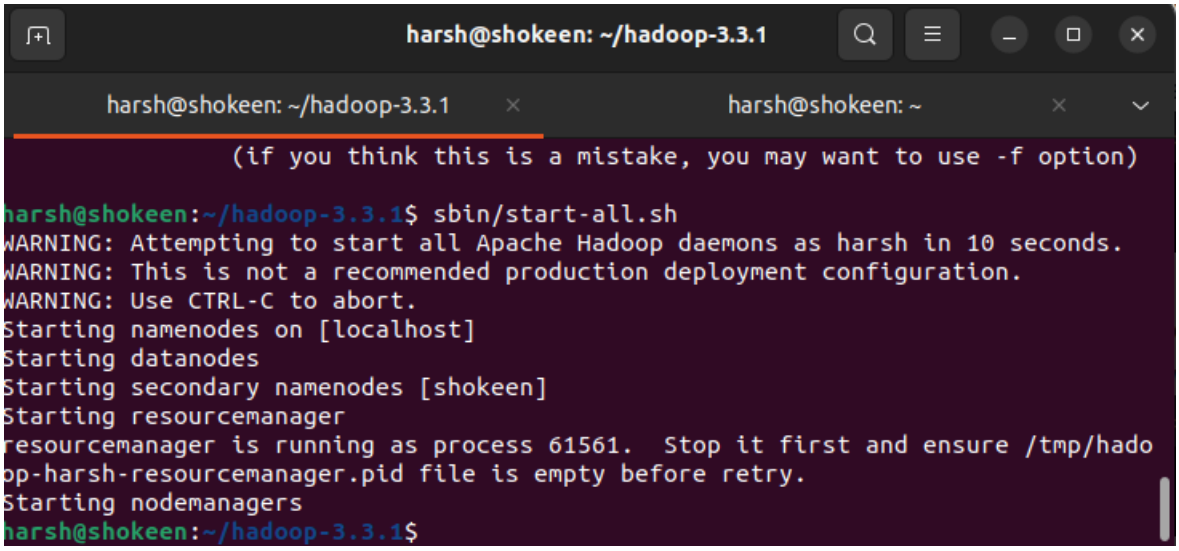
Step 8: Add the following configuration inside the <configuration> section.



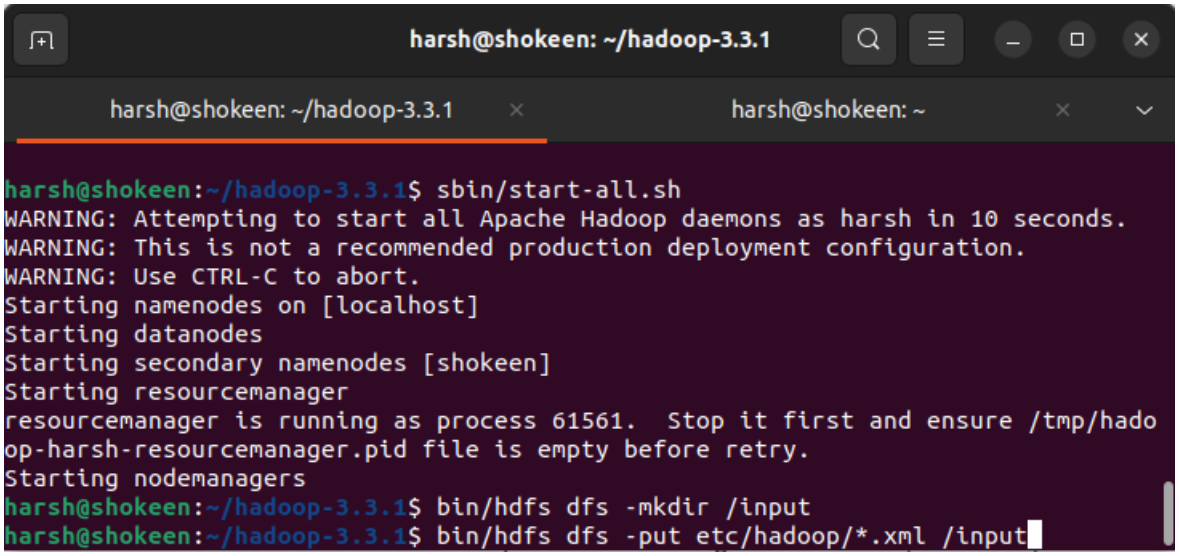
Step 9: Format the Hadoop file system



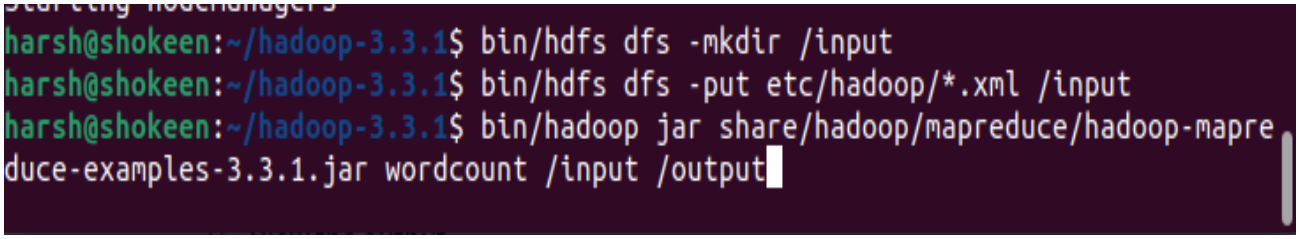
Step 10: Start the Hadoop cluster. This will start the Hadoop services including the NameNode, DataNode, ResourceManager, and NodeManager.



Step 11: Create input directory and copy input files to HDFS:

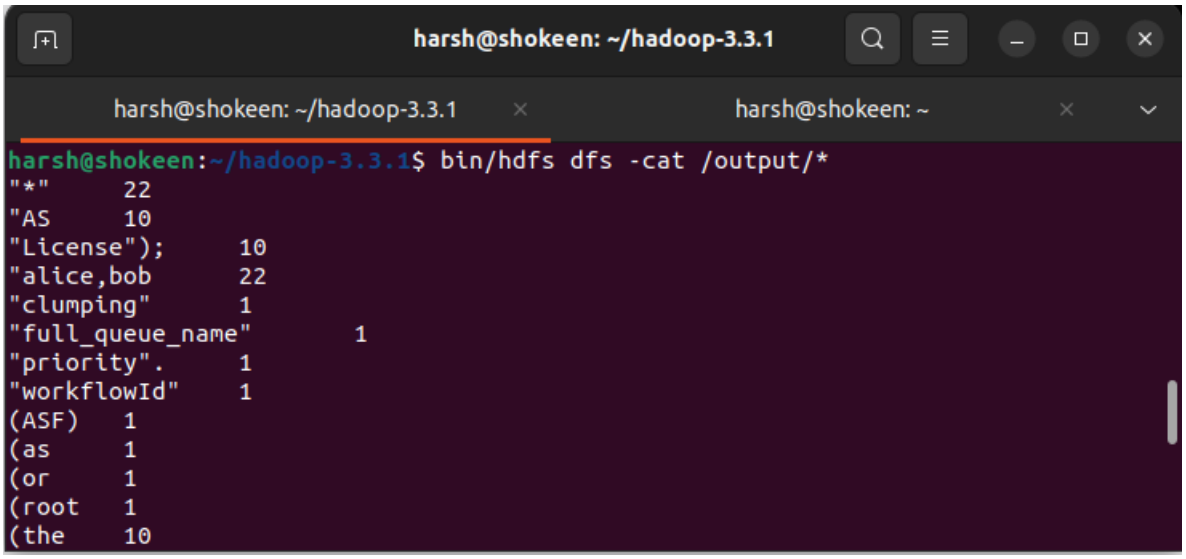


Step 12: Run the Word Count example using this command**: bin/hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-3.3.1.jar wordcount /input /output**

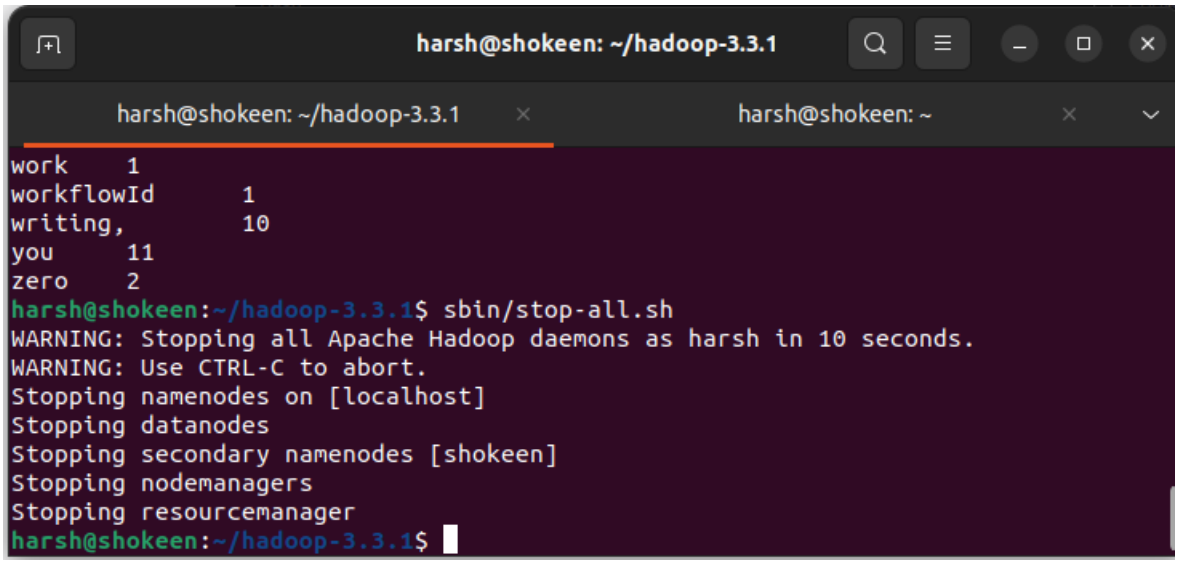


View the output:

bin/hdfs dfs -cat /output/* - This will display the word count output on the console.



Step 14: Stop the Hadoop cluster



These steps should guide you through setting up a Hadoop single-node cluster and running the Word Count

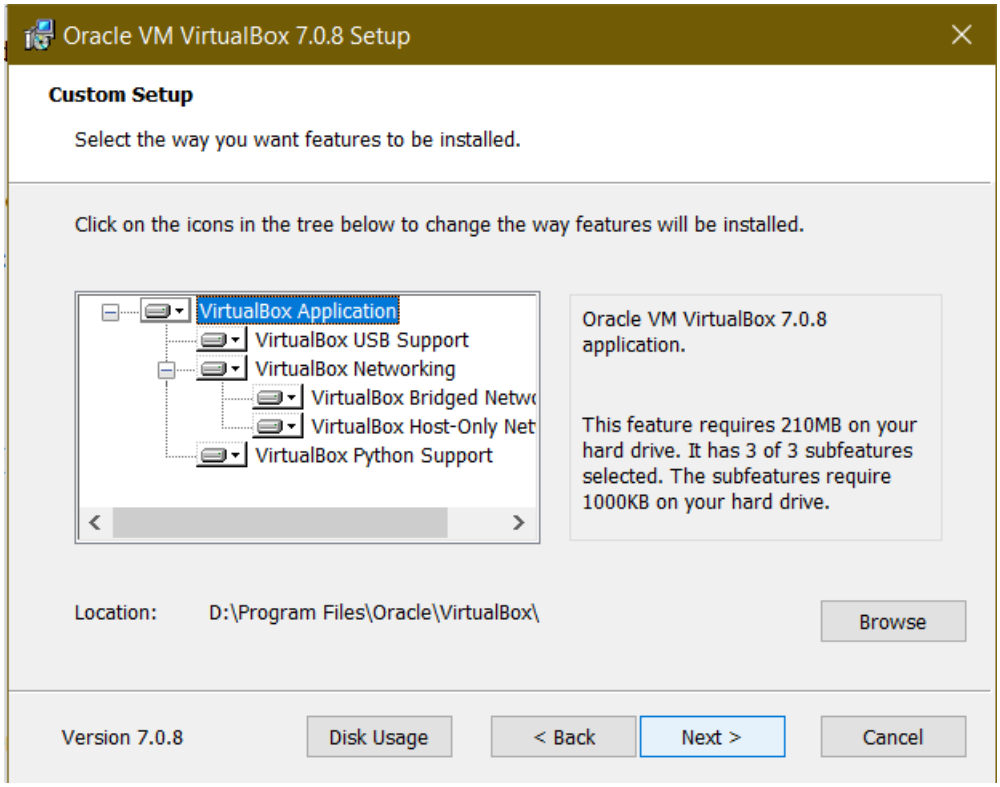
Q9. Install VirtualBox/VMware Workstation with different flavors of Linux or windows OS on top of windows7 or 8.

Procedure:

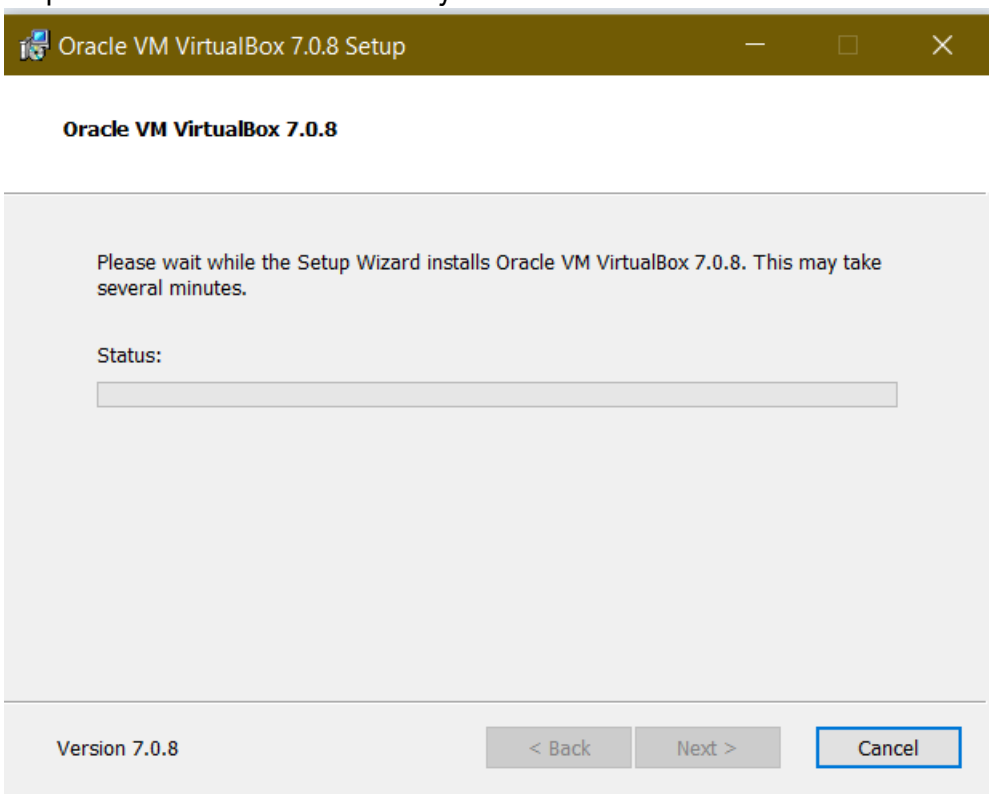
Step 1: Download the Virtual box exe and click the exe file.

Step 2: Click the Next Button of the install Wizard

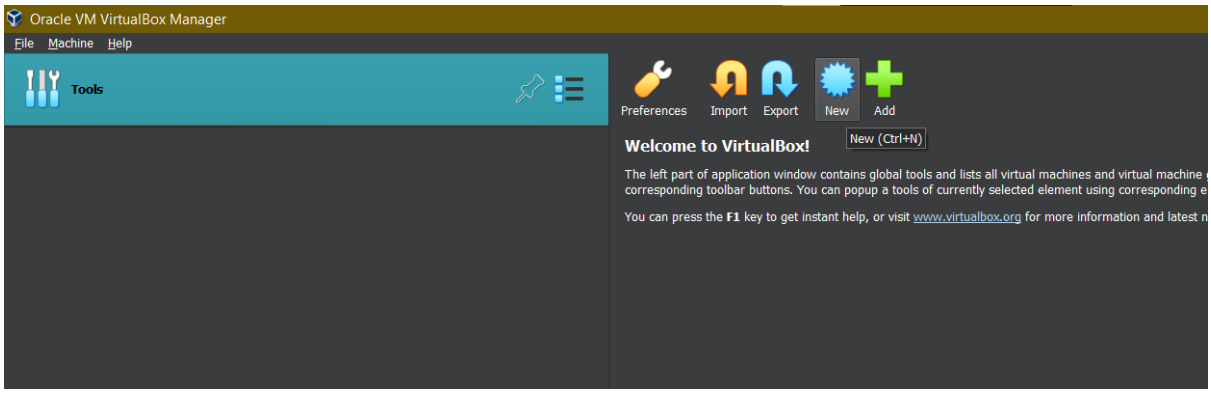
Step 3: Click the next button.



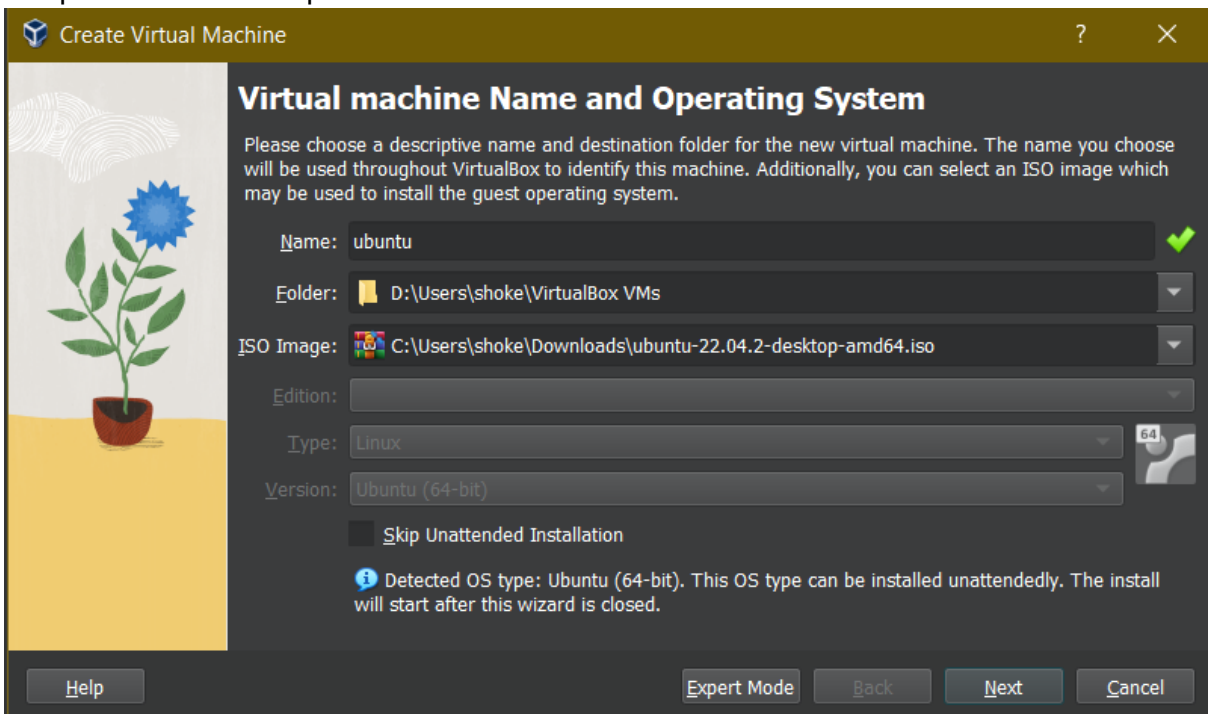
Step 4: Wait for the installation to your desired location



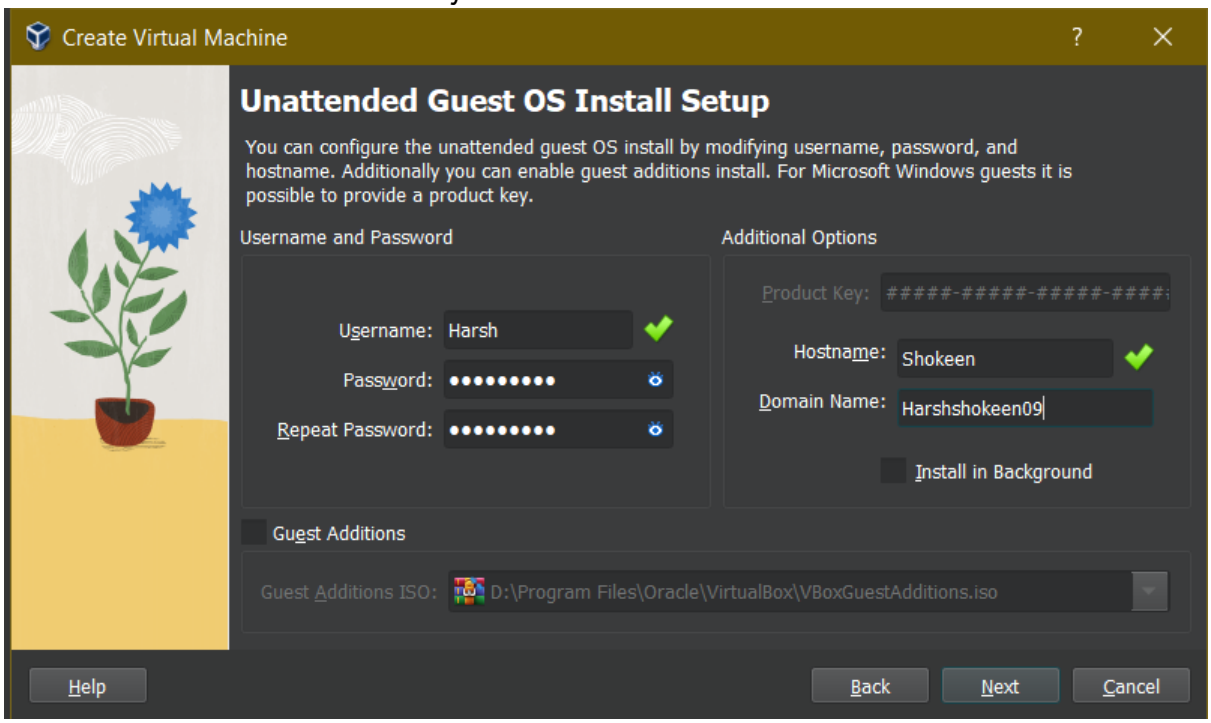
Step 5: After the installation, the Virtualbox will open; click New button in blue colour to make a new workstation



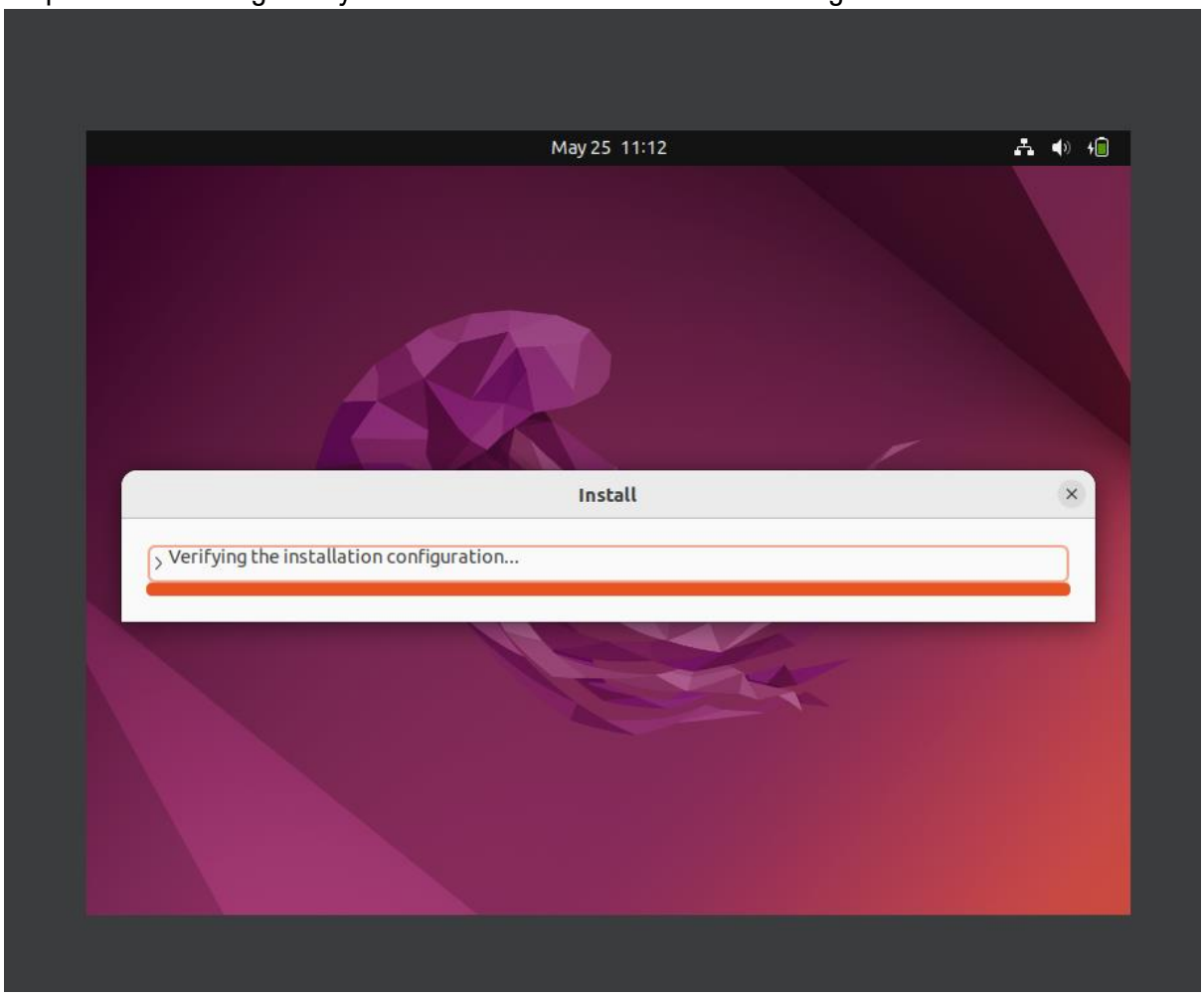
Step 6: After opening the installation Wizard selected the .iso file of your ubuntu downloaded and proceed to next steps.



Step 7: Give your name or the desired name to your system/workstation going to be created. Remember it is the same name as your localhost.



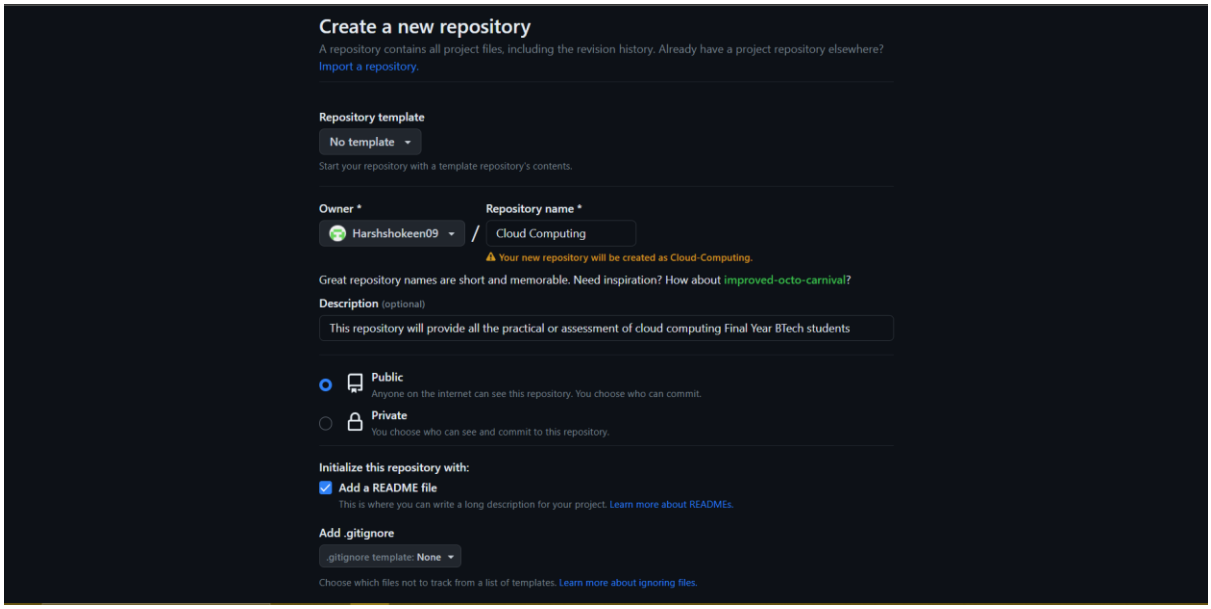
Step 8: After clicking next your ubuntu will load and start installing in virtualbox.



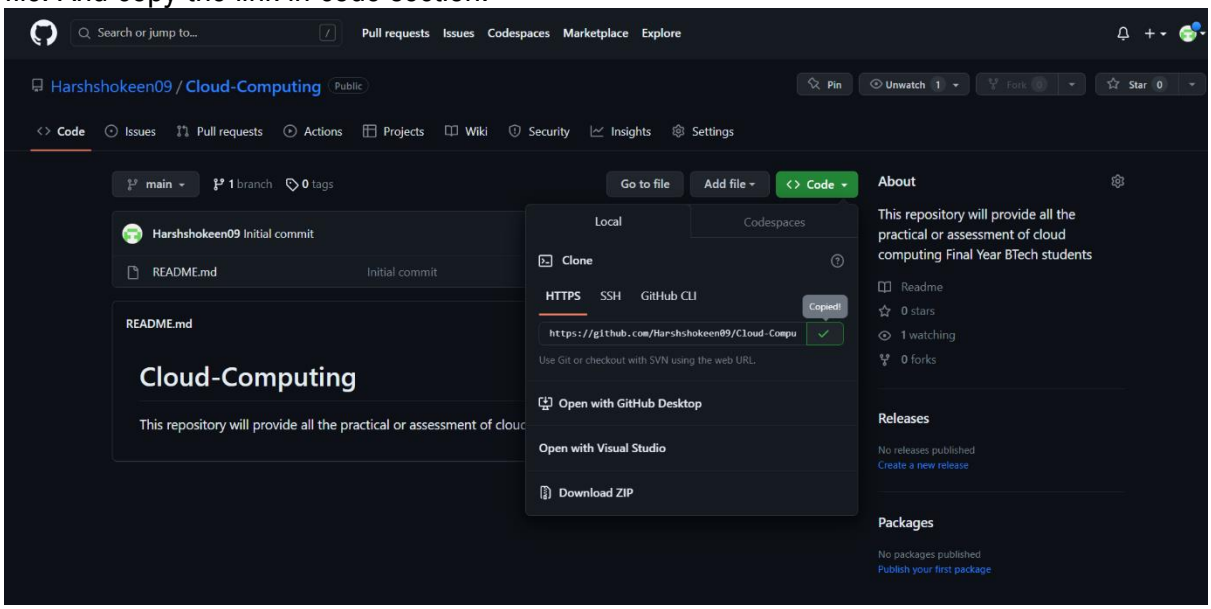
Q10. GIT Commands Execution and GUI & CLI understanding.

Procedure

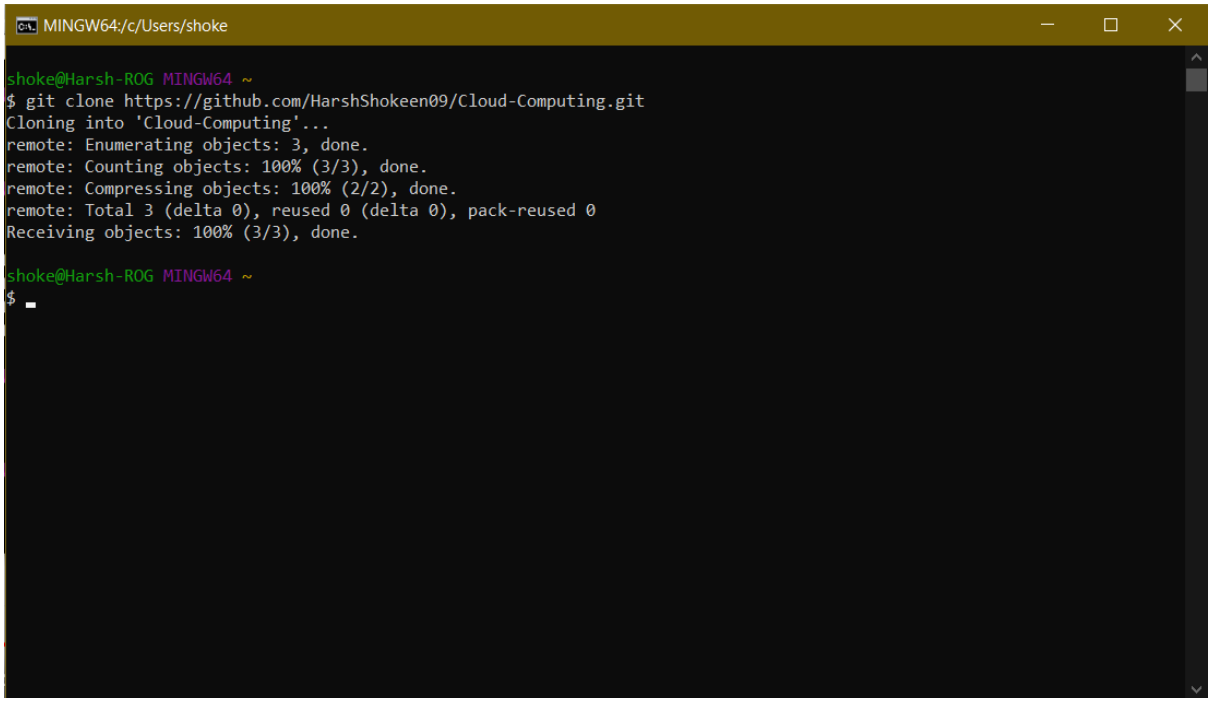
Step 1: Login to your GitHyb account and click on create repository.



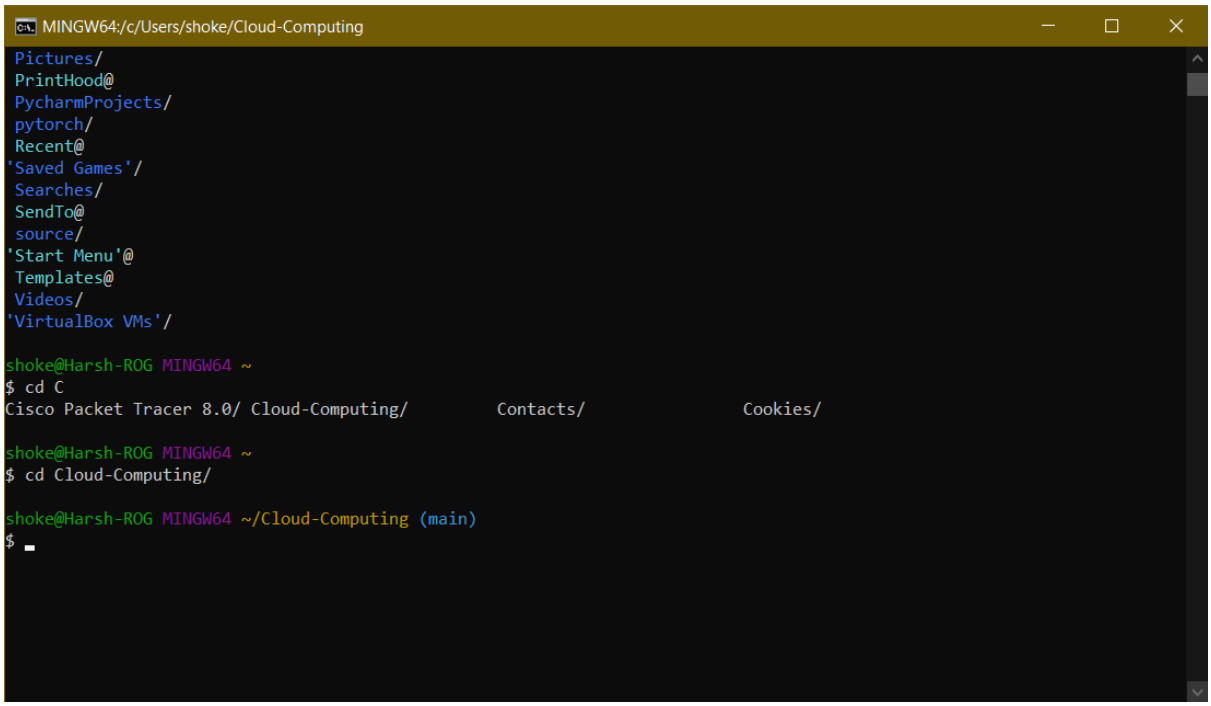
Step 2: After creating the repository manually transfer a file or just add information in readme file. And copy the link in code section.



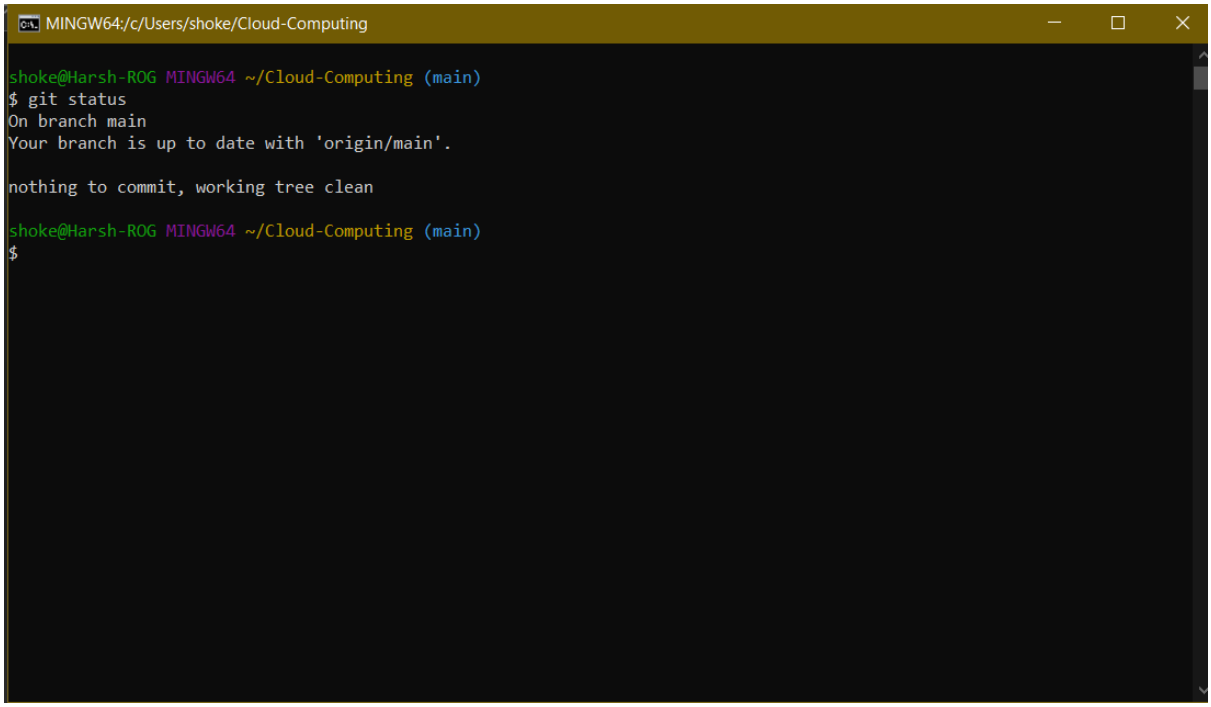
Step 3: Open git-bash and clone the repository by pasting the link and writing the following command.



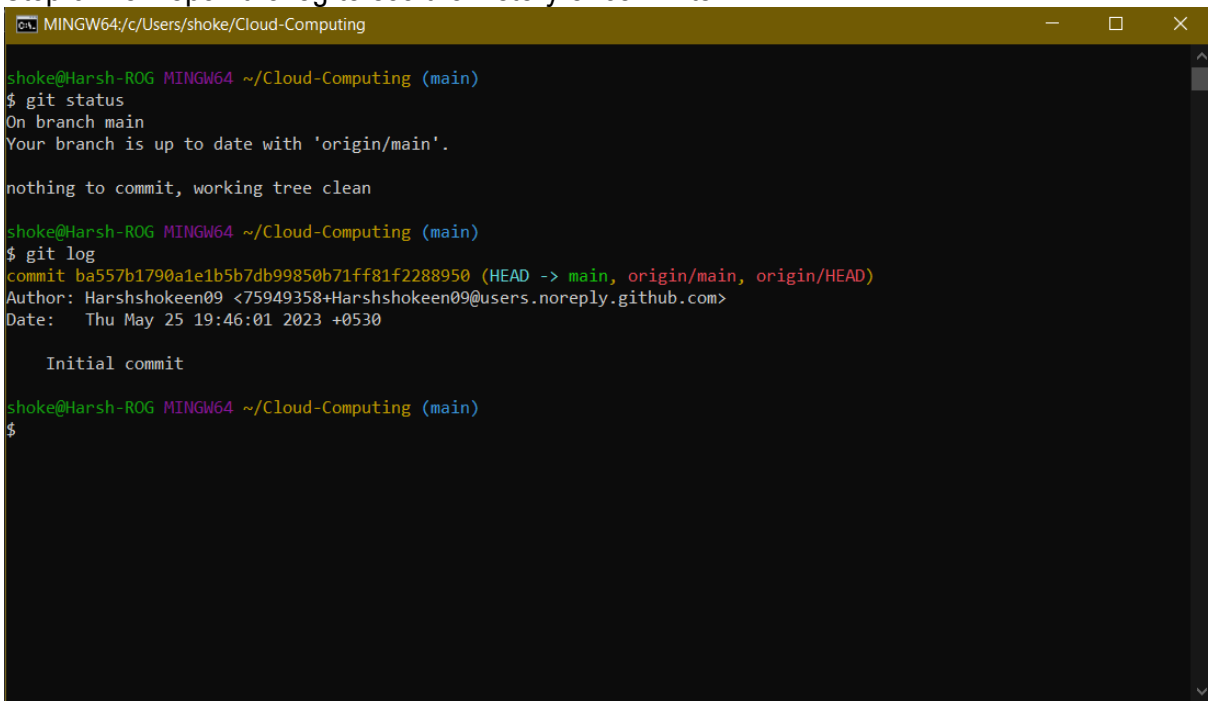
Step 4: Navigate to the desired folder. Therefore, making it the main/master branch where all the actions would be reflected.



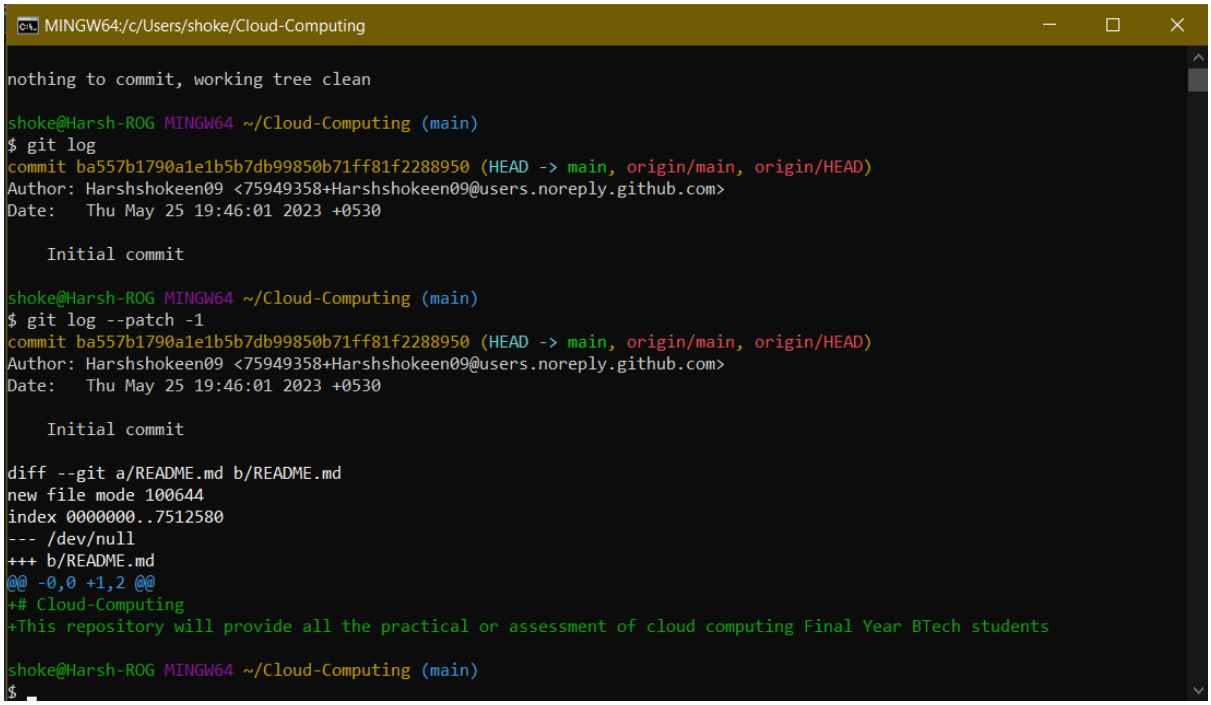
Step 5: Now check the status of your repository.



Step 6: Now open the log to see the history of commits.

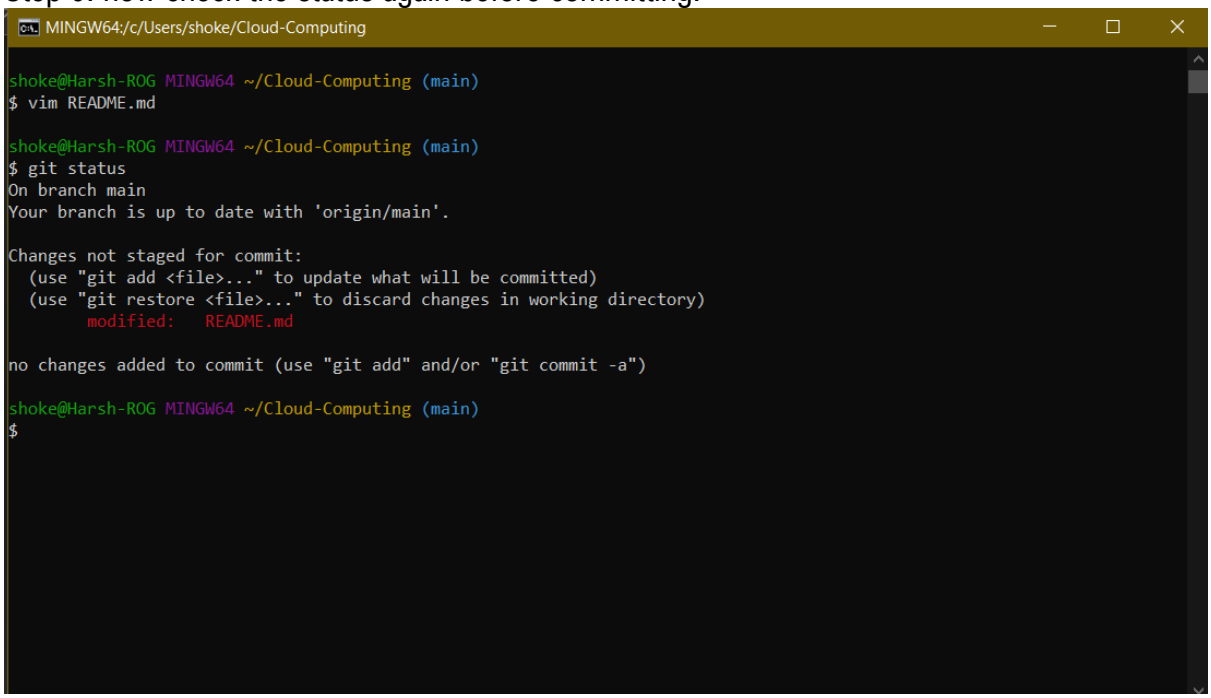


Step 7: Now to see the history of commits content open its patch version

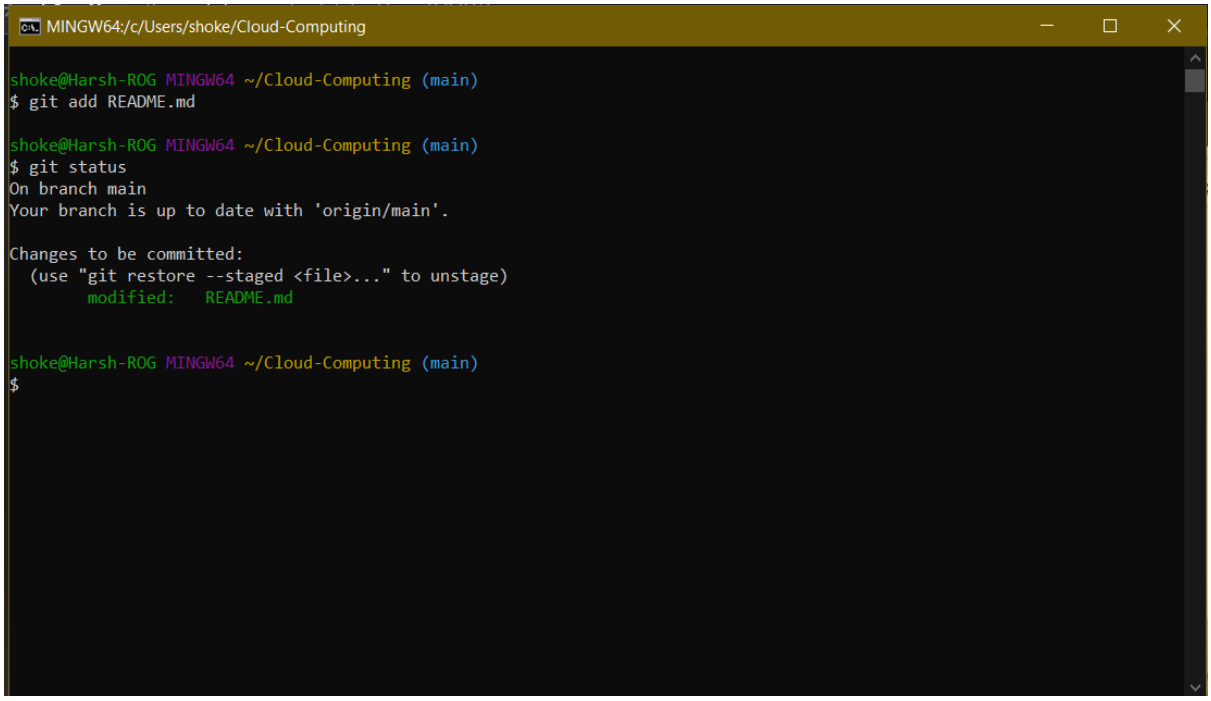


Step 8: Now open the readme file to make the first commit.

Step 9: now check the status again before committing.



Step 10: Now add changes in the readme.md file and save them using Crlt+O and exit with Crlt+X. Again, check the status to see the committed changes.



Step 11: Now finally Commit the changes.