CLOUD COMPUTING METHODOLOGIES LAB ASSESSMENT-5

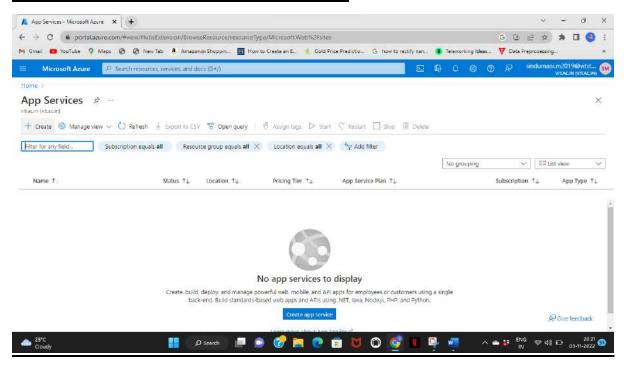
NAME:SINDUMANI.M

REGNO:19MIC0002

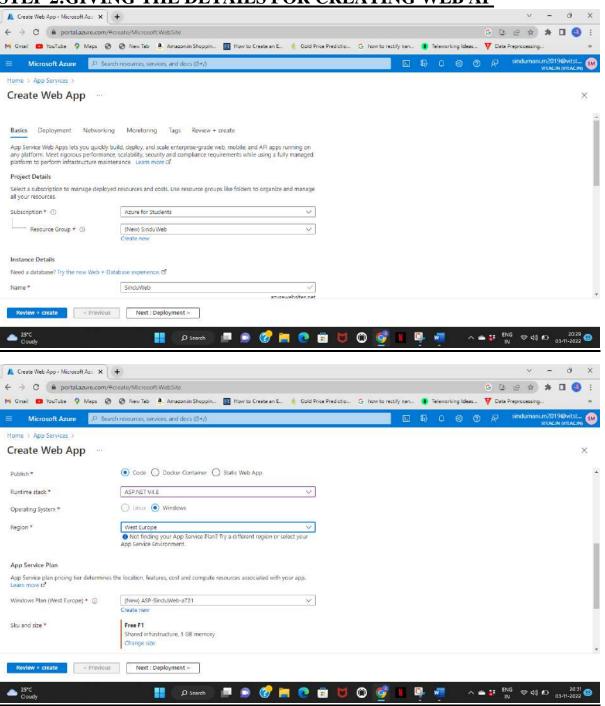
SLOT:L9+L10

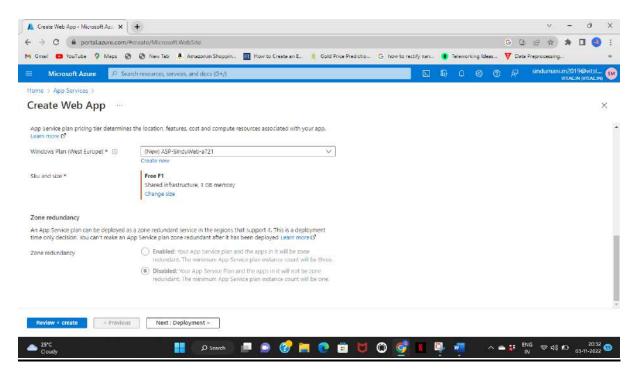
1.DEPLOYMENT OF A BASIC WEBAPP AND ADD ADDITIONAL FUNCTIONALITY

STEP 1:TO CREATE AN APP SERVICE

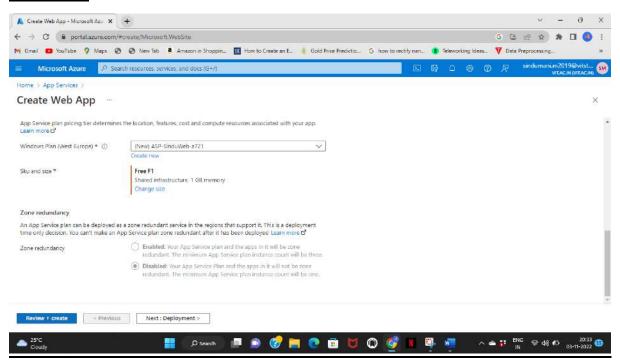


STEP 2:GIVING THE DETAILS FOR CREATING WEB AP

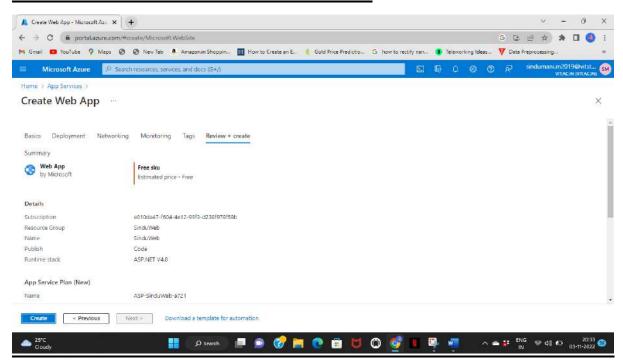




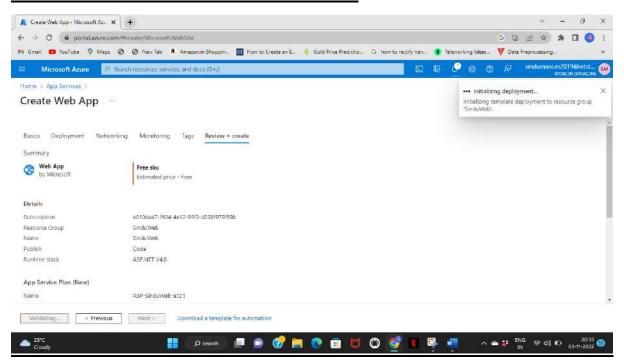
STEP 3:CLICKING REVIEW+CREATE FOR CREATING THE WEBAPP



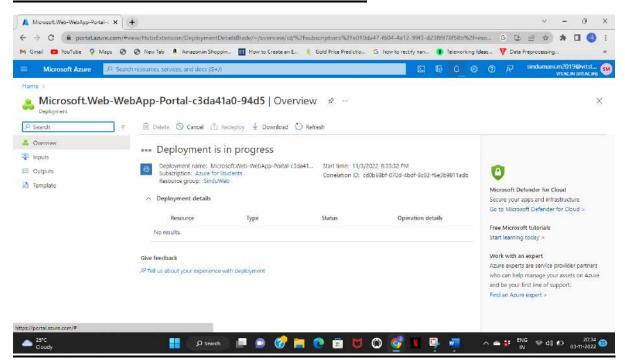
STEP 4:CLICKING CREATE TO DEPLOY



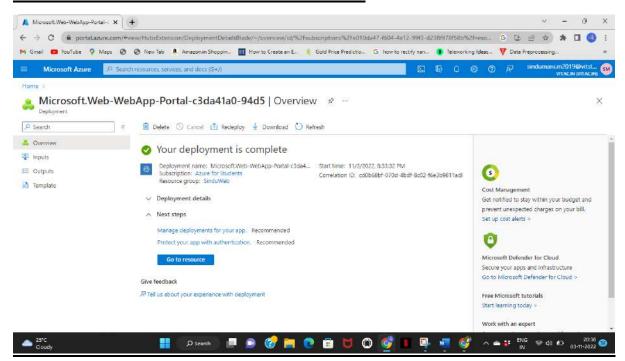
STEP 5:DEPLOYMENT IS INITIALIZED



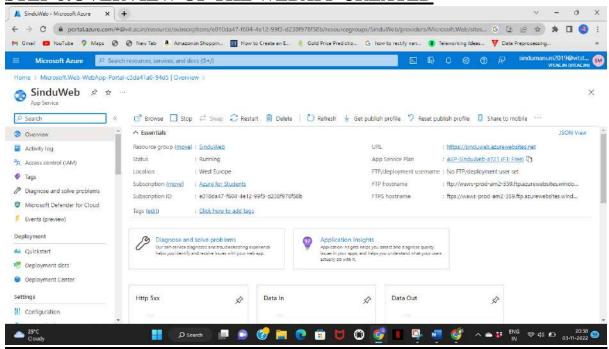
STEP 6:DEPLOYMENT IS IN PROGRESS



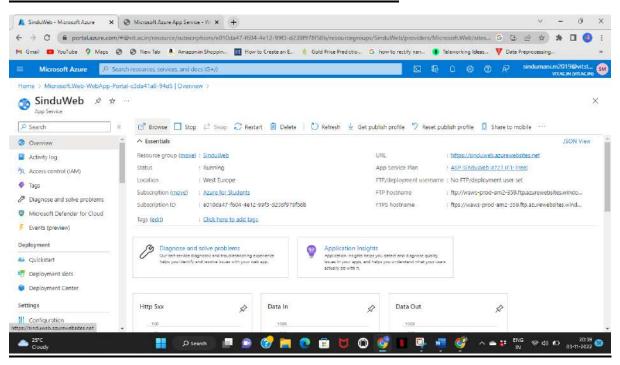
STEP 7:DEPLOYMENT IS COMPLETED



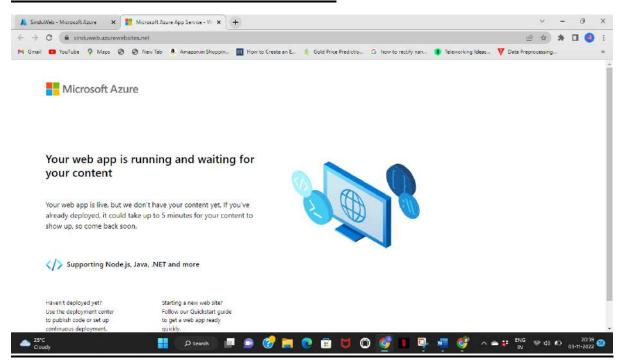
STEP 8:OVERVIEW OF THE WEBAPP CREATED



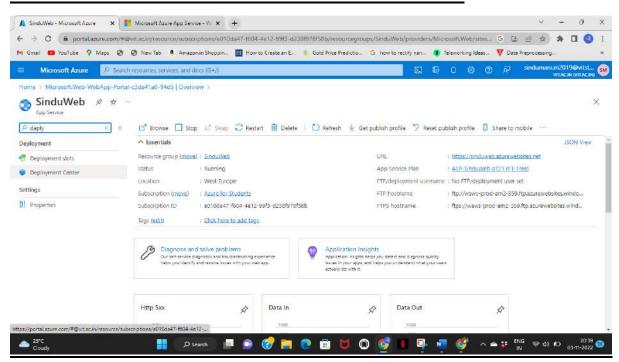
STEP 9: COPYING THE URL FROM WEBAPP



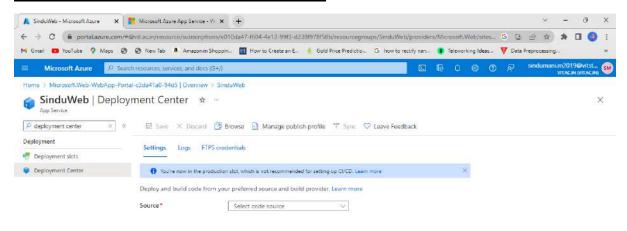
STEP 10:PASTING THE LINK IN URL



STEP 11:SEARCHING FOR DEPLOYMENT CENTER

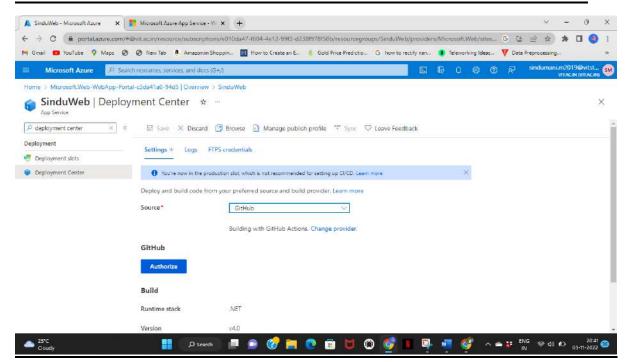


STEP 12:SELECTING THE SOURCE

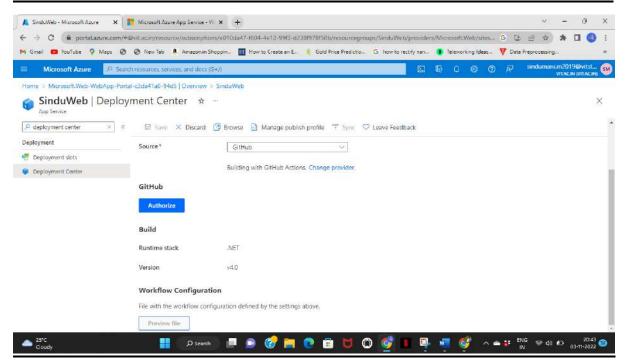




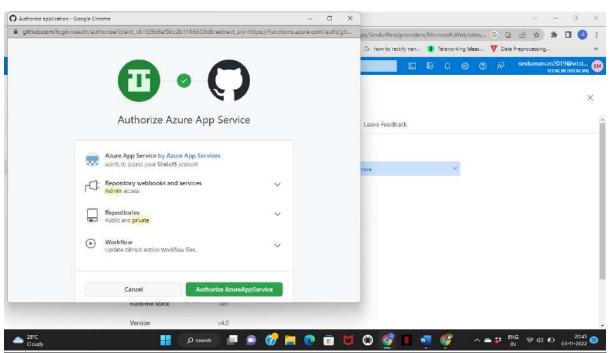
STEP 13:CHOOSING GITHUB FROM THE AVAILABLE SOURCES



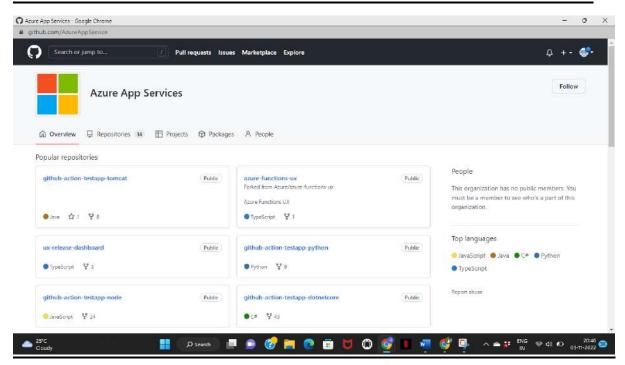
STEP 14:CLCIKING AUTHORIZE TO LINK OUR GITHUB ACCOUNT



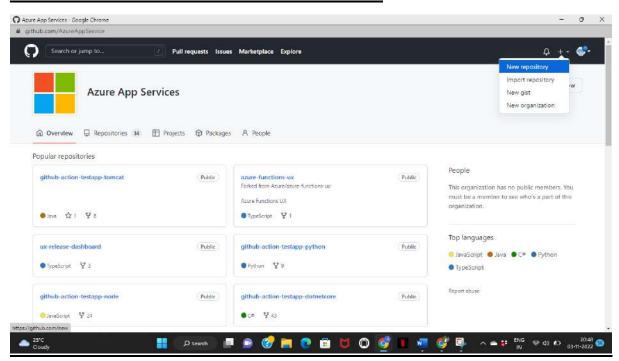
STEP 15:CHOOSING AZURE APP SERVICE AT THE FIRST DIALOGUE BOX



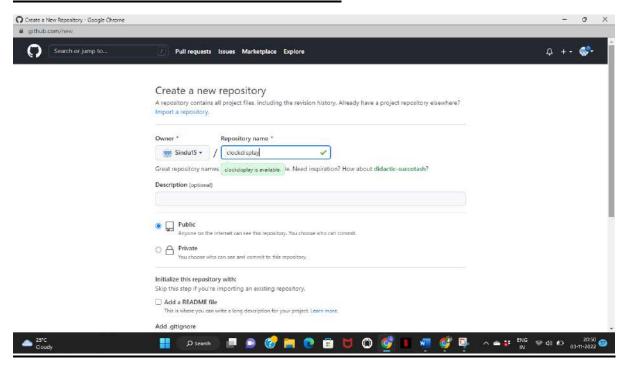
STEP 16:THE GITHUB IS OPENED UNDER AZURE APP SERVICES



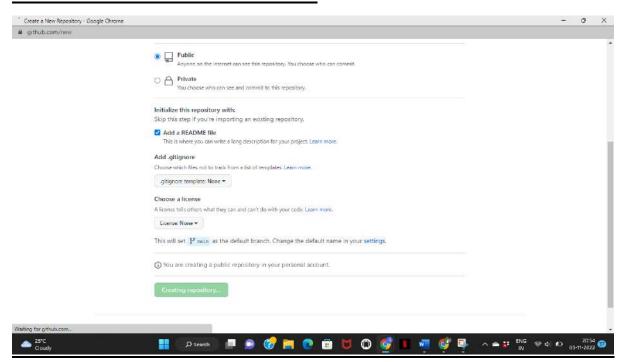
STEP 17: CREATING A NEW REPOSITORY



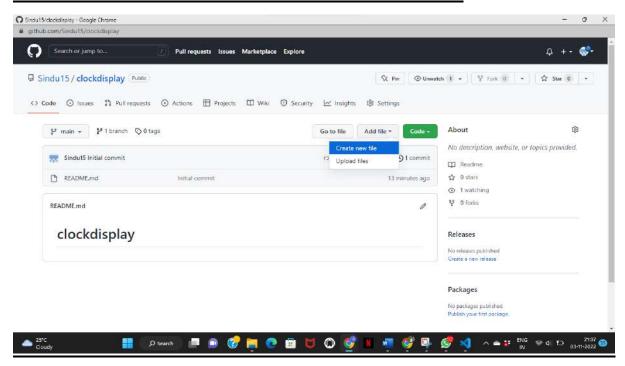
STEP 18:NAMING THE REPOSITORY



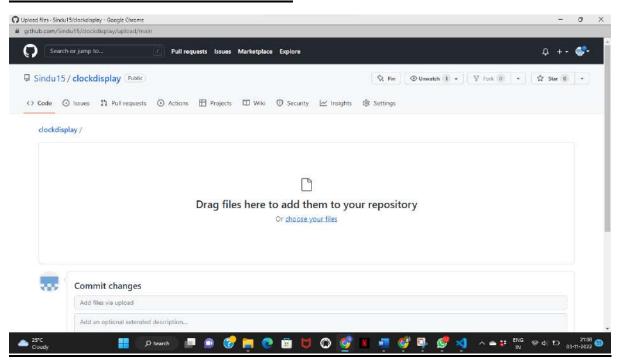
STEP 19: CREATING REPOSITORY



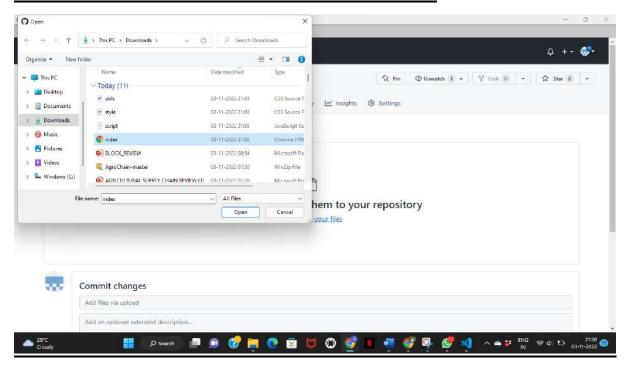
STEP 20:CLICKING ADD FILES TO UPLOAD FILES



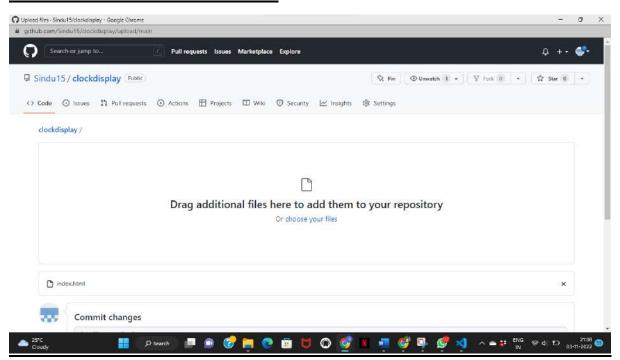
STEP 21:CHOOSING OUR FILES



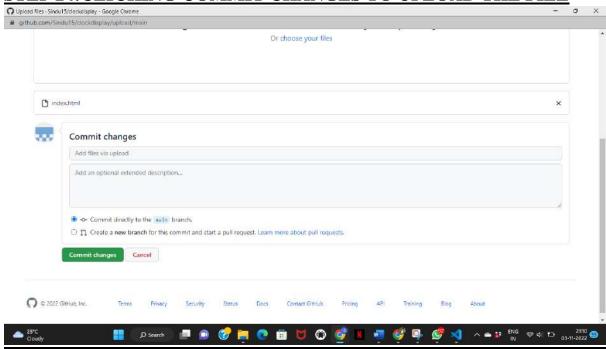
STEP 22:OPENING index.html which is the mainfile



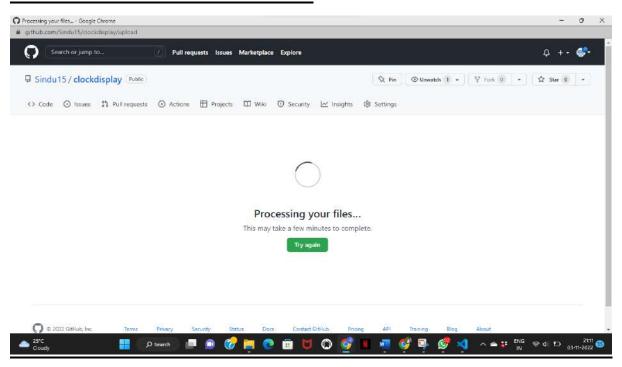
STEP 23:CHOOSING index.html



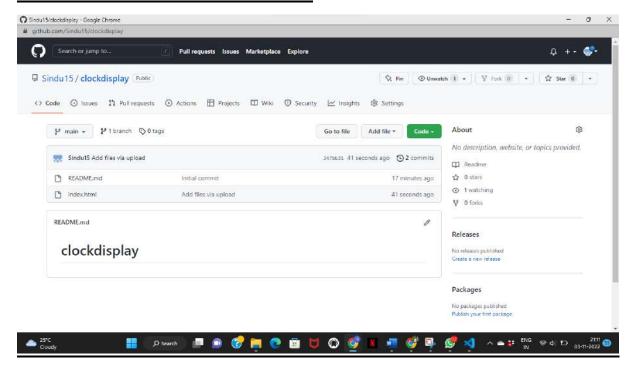
STEP 24:CLICKING COMMIT CHANGES TO UPLOAD THE FILE



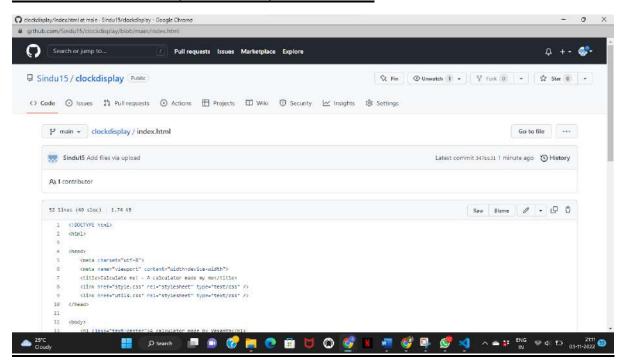
STEP 25:THE FILE IS PROCESSED



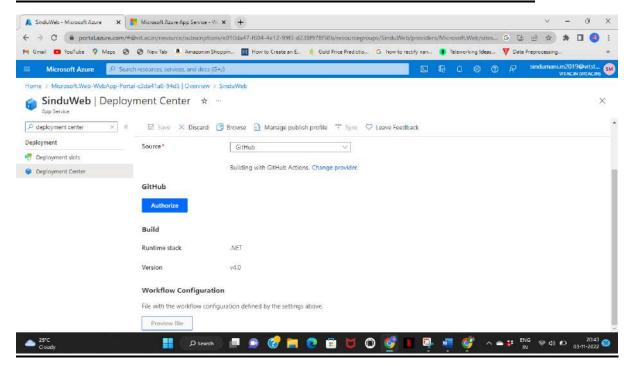
STEP 26:THE FILE IS UPLOADED



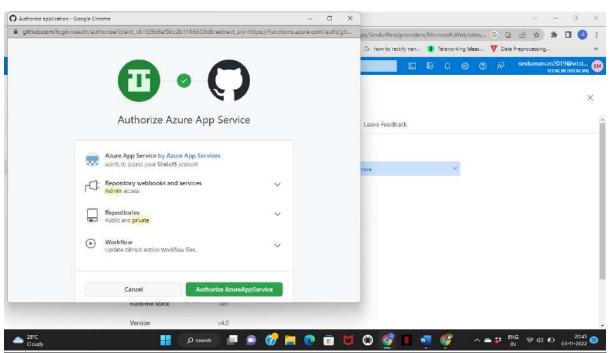
STEP 27:THE FILE(index.html)IS VIEWED



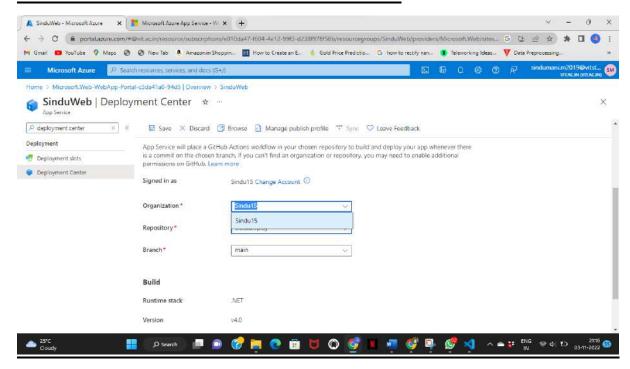
STEP 28:CLICKING AUTHORIZE TO LINK THE ACCOUNT



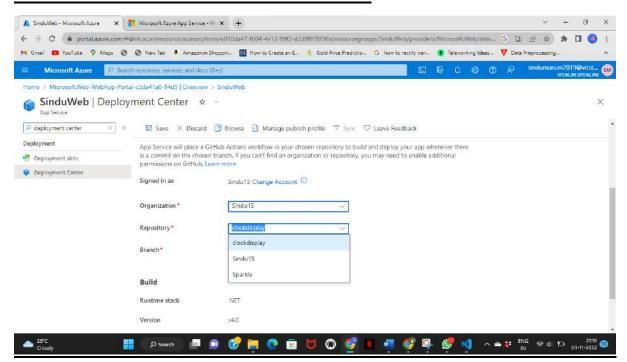
STEP 29:CLICKING AUTHORIZE AZUREAPPSERVICE TO LINK OUR ACCOUNT FINALLY



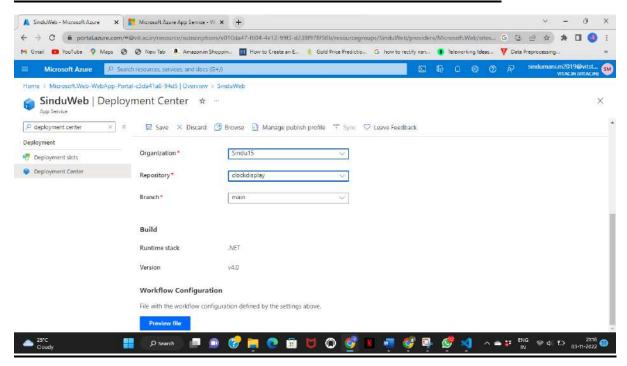
STEP 30:SELECTING THE ORGANIZATION



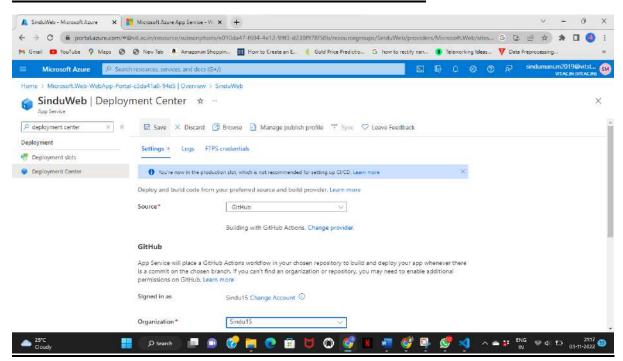
STEP 31:SELECTING THE REPOSITORY



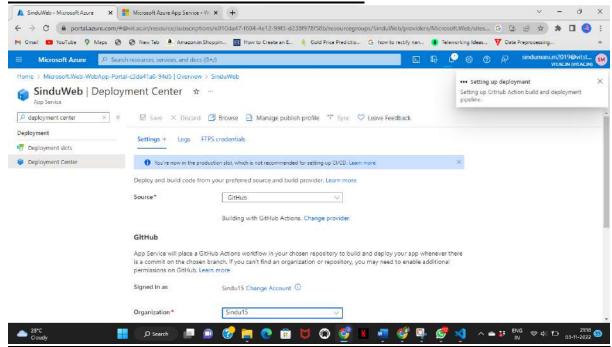
STEP 32:CHOOSING BRANCH AND PREVIEWING FILE



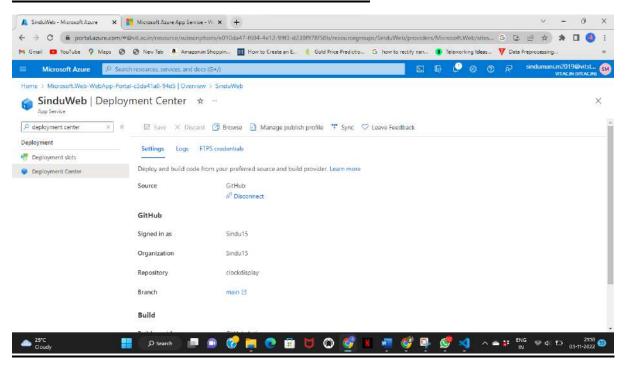
STEP 33:SAVING THE DEPLOYMENT CENTER

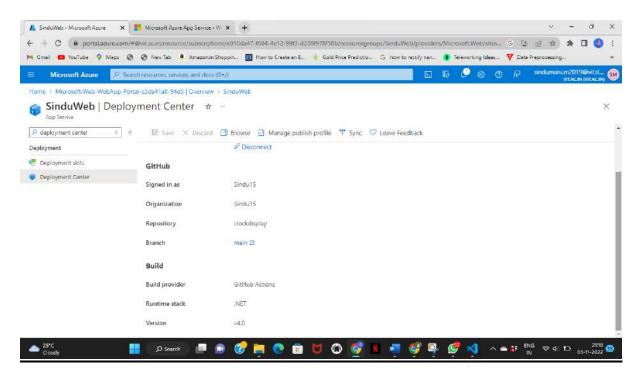


STEP 34:DEPLOYMENT IS INITIALIZED

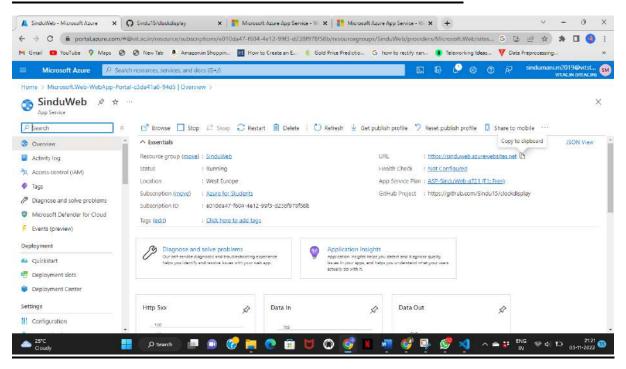


STEP 35:DEPLOYMENT IS COMPLETED

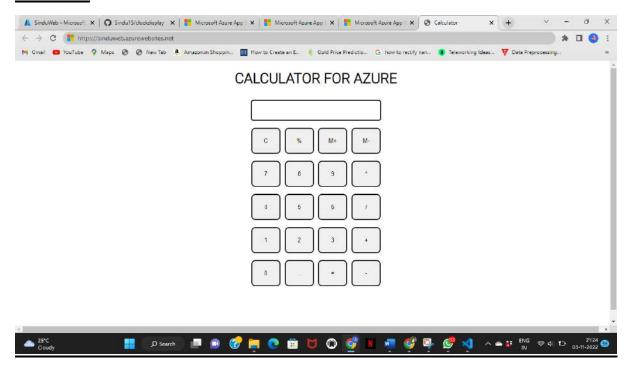




STEP 36:COPYING THE LINK FROM APP SERVICE



STEP 37:PASTING THE URL AND DEPLOYING THE FILE FROM GITHUB



2.CASE STUDY ON HADOOP AS A SERVICE

ABSTRACT

Using straightforward programming concepts, the Apache Hadoop software library provides a framework for the distributed processing of massive data volumes across computer clusters. From a single server to thousands of devices, each providing local computing and storage, it is intended to scale up. The library itself is designed to identify and handle problems at the application layer rather than relying on hardware to provide high availability. As a result, a highly-available service is delivered on top of a cluster of computers, each of which may be prone to failures. In this post, we will examine each HDFS component in detail and learn how it operates internally. HDFS has a master/slave architecture. HDFS and MapReduce are the two main components of the Hadoop framework. A single NameNode, a master server that oversees the file system namespace and controls client access to files, makes up an HDFS cluster. Additionally, there are a number of DataNodes that manage storage related to the nodes that they run on, often one per node in the cluster. User data may be saved in files thanks to HDFS, which offers a namespace for the file system. A file is internally divided into one or more blocks, which are then stored in a collection of **DataNodes.** File system namespace activities like opening, shutting, and renaming files and directories are carried out by the NameNode. Additionally, it chooses how blocks are mapped to DataNodes. Serving read and write requests from the file system's customers is the responsibility of the DataNodes. Upon receiving a command from the NameNode, the DataNodes also carry out block creation, deletion, and replication.

HDFS ANALYSIS

Here is the dependency graph of the hdfs project following the analysis of <u>Hadoop with JArchitect</u>. Hdfs leverages a variety of third-party libraries, including <u>guava</u>, <u>jetty</u>, <u>jackson</u>, <u>and others</u>, to complete its task. More information regarding the burden of using each library is provided by the DSM (Design Structure Matrix). Most of the libraries used by HDFS are rt, hadoop-common, and protobuf. When using external libraries, it's best to see if we can quickly replace a third party library with another one without having an adverse effect on the entire programme. There are numerous factors that can lead us to do so. The other library might:

Have more features

- More performent
- More Secure

We have taken <u>jetty library</u> as an example and see which hdfs methods use it directly. m.IsUsing ("jetty-6.1.26") && m.ParentProject.Name=="hadoop-hdfs-0.23.6" from m in Methods choose new'm, m.NbBCInstructions'Only a small number of methods directly use the Jetty library, and switching to another one is relatively simple. In general, it's highly interesting to limit the use of an external library to just a few classes when you can, since this makes it easier to maintain and develop the project. The main **HDFS components** are as follows:

I-DATA NODE

STARTUP

Let's perform a search before all hdfs jar entry points from m in Methods where m.Name to find out how to start a data node. There are several entries in HDFS, including **DFSAdmin, DfSsc, Balancer, and HDFSConcat.** Contains("main(String[])") && m.IsStatic select new m, m.NbBCInstructions Here is what happens when the DataNode class's main method is called, which is the entry point for the data node. When the node is launched in a non-secure cluster, this parameter is null; however, when it is started in a secure environment, the param is assigned with the secure resources. The main method then calls securemain after passing it the securityresources parameter. Two attributes may be found in the SecureResources class:

- 1. **streamingSocket**: secure port for data streaming to datanode.
- 2. **listner**: a secure listener for the web server.

And here are the methods invoked from DataNode.StartDataNode.This method initialize IPCServer,DataXceiver which is the thread for processing incoming/outgoing data stream, create data node metrics instance.

HOW DATA IS MANAGED?

The DataNode class has a data attribute with the **FSDatasetinterface** data data type. The interface for the underlying storage that houses the blocks for a data node is called FSDatasetinterface. Let's look up the Hadoop implementations that are available. where t is taken from Types.

<u>Hadoop offers FSDataset, which handles a collection of data blocks and stores them on dirs.</u>

Implement("org.apache.hadoop.hdfs.server.datanode.FSDatasetInterface") choose new t, t.NbBCInstructions. Interfaces enforce low coupling and make the design very flexible; however, if the implementation is used instead of the

interface, we lose this advantage. To see if interfaceDataSet is used anywhere to represent the data, let's search for all methods using FSDataSet. Only FSDataSet inner classes use it directly, and for all other places, we need to use the implementation. from m in Methods where m.IsUsing

("org.apache.hadoop.hdfs.server.datanode.FSDataset") select new {m, m.NbBCInstructions}Only FSDataSet inner classes use it directly, and for all the other places the <u>interfaceDataSet</u> is used instead, what makes the possibility to change the dataset kind very easy. from m in Methods let depth0 = m.DepthOfCreateA("org.apache.hadoop.hdfs.server.datanode.FSDataset") where depth0 == 1 select new {m, depth0} The factory pattern is used to create the instance; the problem is if this factory create the implementation directly inside getFactory method, we have to change the Hadoop code to give it our custom DataSet manager. Methods are used by the getFactory method. Where m.IsUsedBy("org.apache.hadoop.hdfs.server.datanode.FSDatasetInterface\$Fact ory.getFactory(Configuration)") is taken from m in Methods. choose new'm, m.NbBCInstructions'.The good news is that since the factory uses configuration to obtain class implementation, we may search for any classes that can be provided by configuration in addition to receiving our custom DataSet through configuration. from m in Methods where m.IsUsing ("org.apache.hadoop.conf.Configuration.getClass(String,Class,Class)") select new {m, m.NbBCInstructions} Many classes could be injected inside the Hadoop framework without changing its source code, what makes it very flexible.

NAMENODE

All HDFS metadata is arbitrated and stored on the **NameNode**. The NameNode never receives any user data because of the way the system is set up. When the name node is launched, the following methods are called. The following is an quick overview of the fsnamesystem and the RPC Server after they have been constructed and loaded:

NAMENODERPSERVER

All <u>RPC requests</u> sent to the <u>NameNode</u> are handled by <u>NameNodeRpcServer</u>. As an illustration, when a data node is launched, it must register itself with the NameNode. The rpc server receives this request and forwards it to the fsnamesystem, which then directs it to the dataNodeManager. from m where m.IsUsed By is true

("org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.blockReceiv

<u>d(DatanodeRegistration,String,Block[],String[])</u>") Each rectangle in the graph corresponds to the number of bytes of code instructions, and we can see that BlockManager.addBlock performs the majority of the work. select new m, m.NbBCInstructions Ha ddop is intriguing in that each class has a specific role to play and that all requests are routed to the appropriate manager.

FS NAME SYSTEM

Traditional hirarchical file organisation is **supported** by

<u>HDFS</u>. Directories can be made by a user or an application, and files can be stored there. Similar to the majority of other current file systems, the file system namespace structure allows for the creation and removal of files as well as the movement of files between directories and file renaming. Here is a dependency graph for creating a symbolic connection as an illustration.

HDFS CLIENT

Basic file operations can be carried out using a Hadoop

<u>Filesystem connection and DFSClient</u>. While interacting with a NameNode daemon, it does so using the ClientProtocol, and when reading/writing block data, it connects directly to DataNodes.

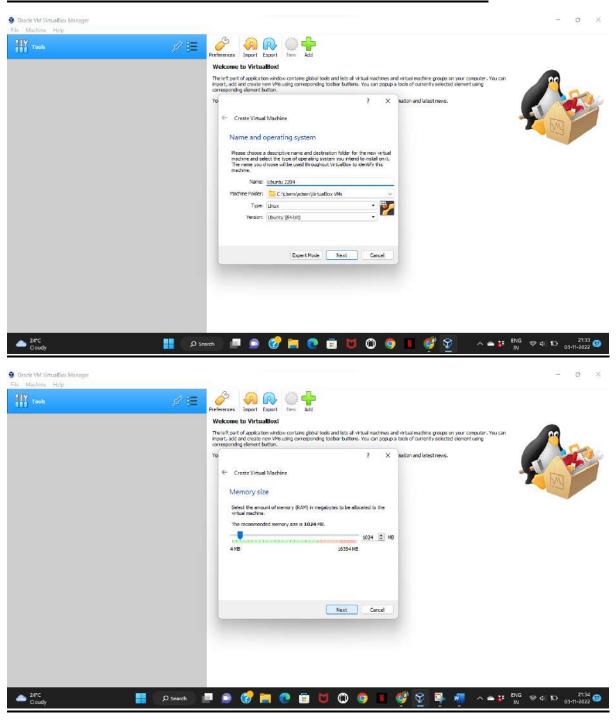
<u>Users of Hadoop DFS</u> need <u>acquire</u> a copy of <u>DistributedFileSystem</u>, which employs DFSClient to manage filesystem operations. The dependency diagram for creating a directory request shows how DistributedFileSystem serves as a facade and directs requests to the DFSClient class.

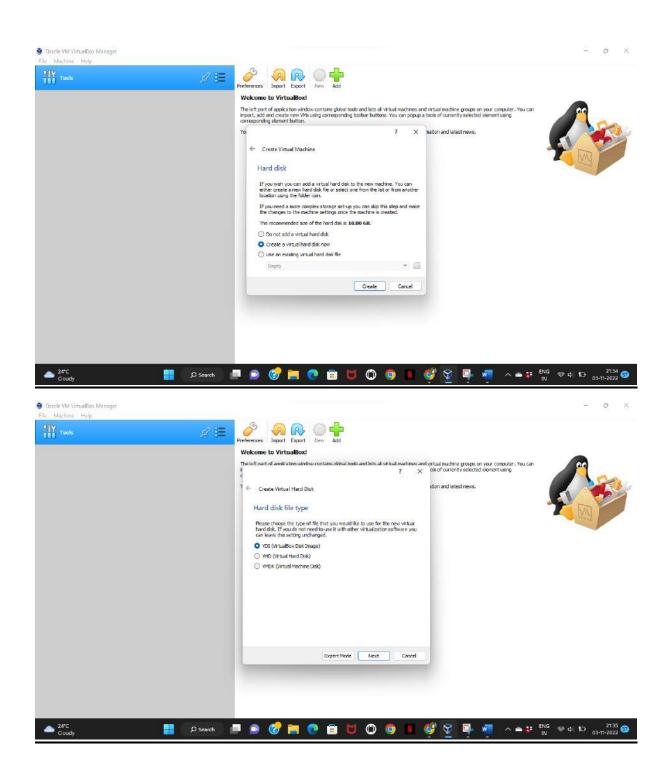
CONCLUSION

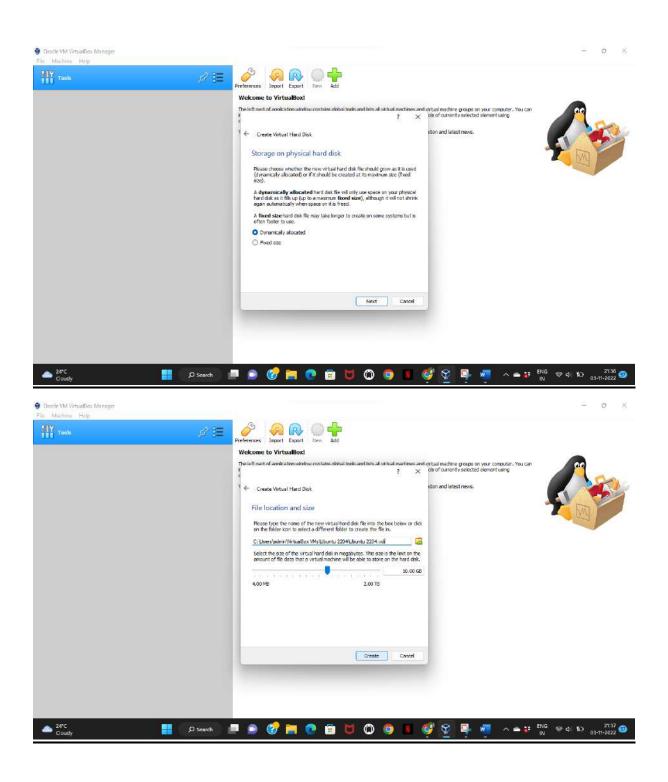
Although using frameworks as a user is quite intriguing, digging deeper into the framework could provide us with more information that would help us better grasp it and more easily adapt it to our needs. <u>Many</u> <u>businesses use the powerful Hadoop framework</u>, and the majority of them need to customise it. Fortunately, <u>Hadoop is quite adaptable</u>, allowing us to change the behaviour without changing the source code.

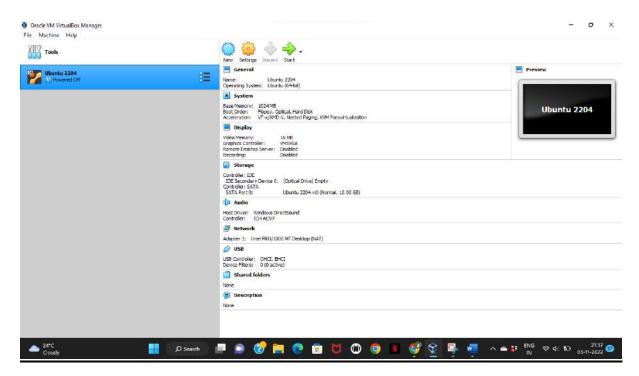
3.OPENSSTACK INSTALLATION USING DEVSTACK IN VIRTUAL BOX AND UBUNTU

STEP 1:INSTALLING UBUNTU IN VIRTUAL MACHINE

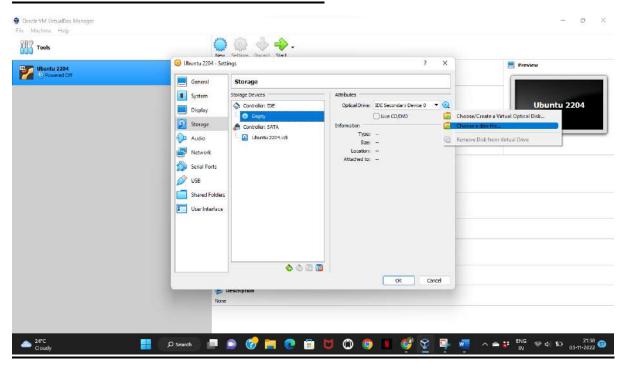


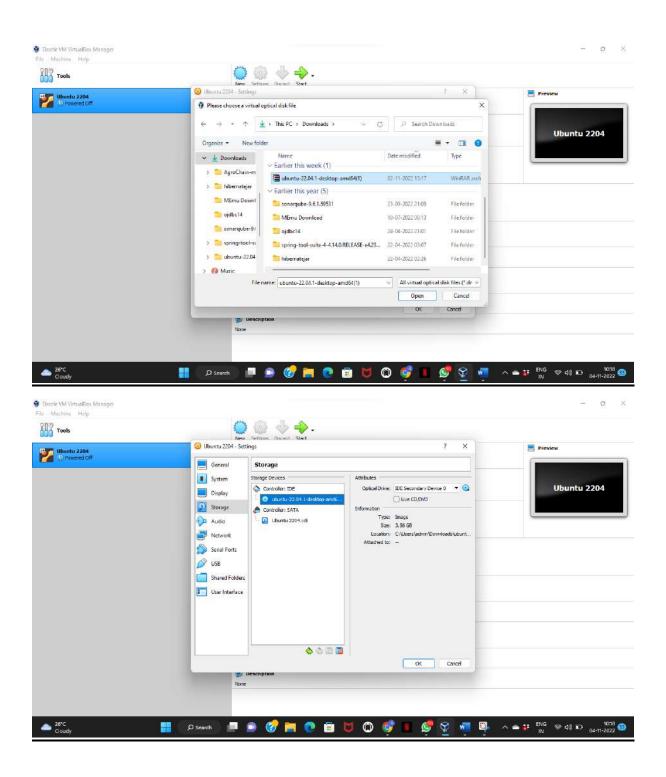




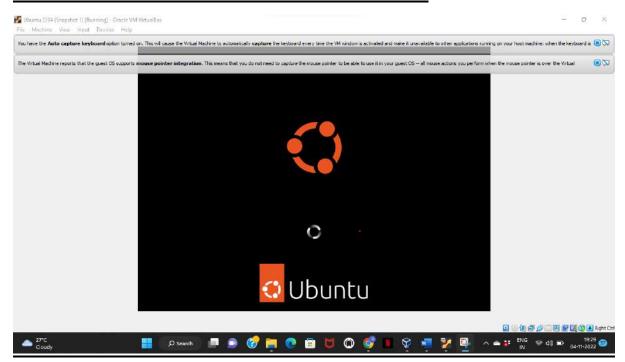


STEP 2:CHOOSING THE DISK FILE

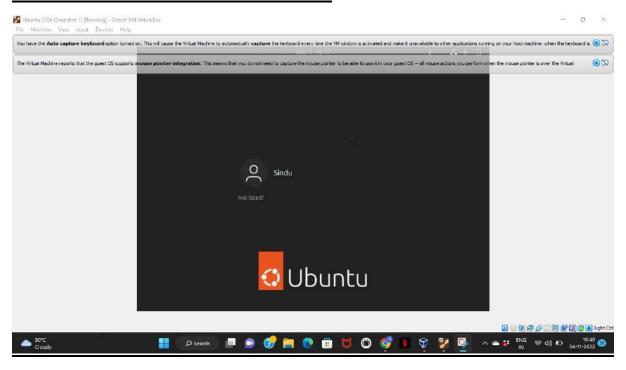


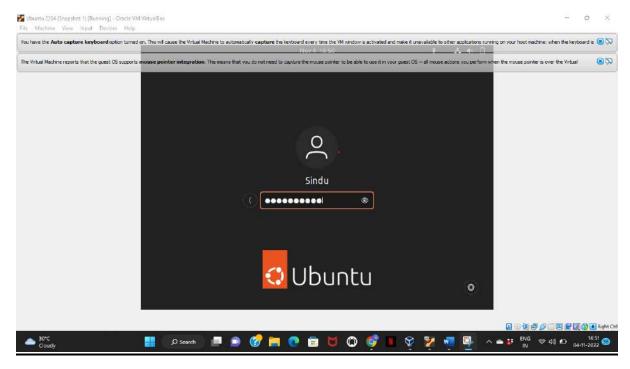


STEP 4:INSTALLING UBUNTU AND STARTING

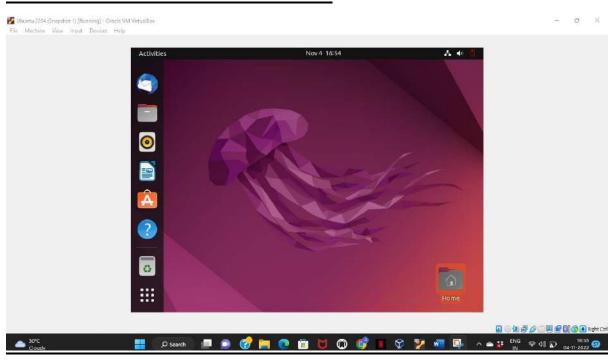


STEP 5:LOGGING IN INTO UBUNTU

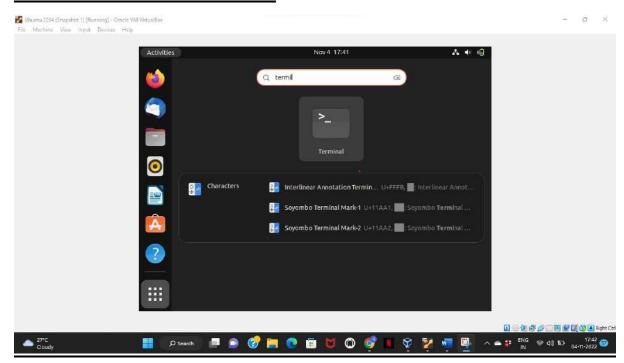




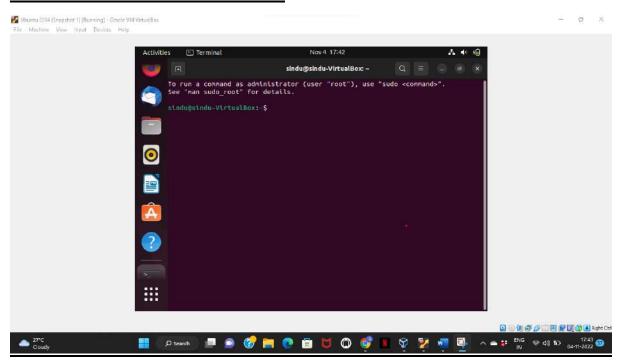
STEP 6:SUCCESSFULLY LOGGED IN



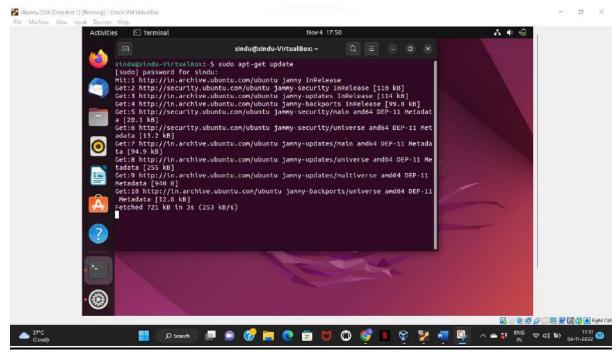
STEP 7:OPENING TERMINAL



STEP 8:TERMINAL IS OPENED



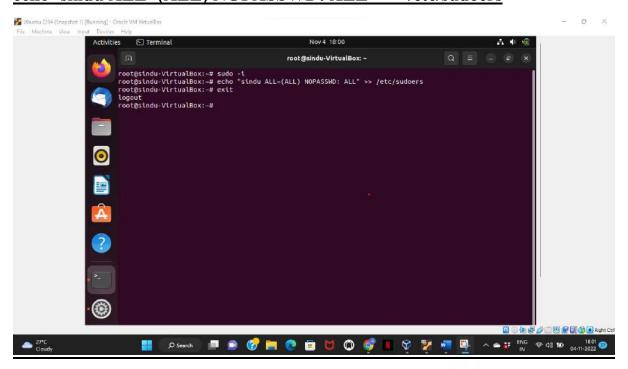
STEP 9:UPDATING THE SYSTEM USING COMMAND:sudo apt-get update



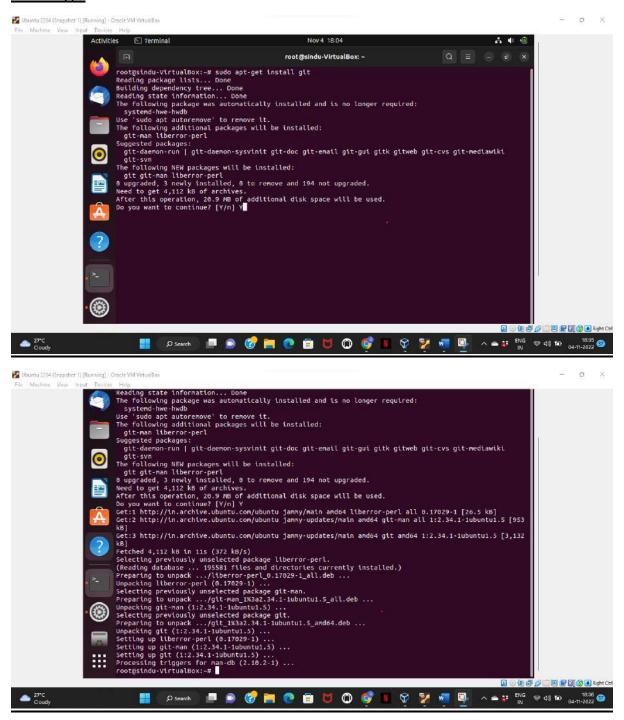
STEP 10:ASSIGNED SUDO PRIVILEDGED AND INSTALLED GIT USING THE COMMANDS:

sudo -l

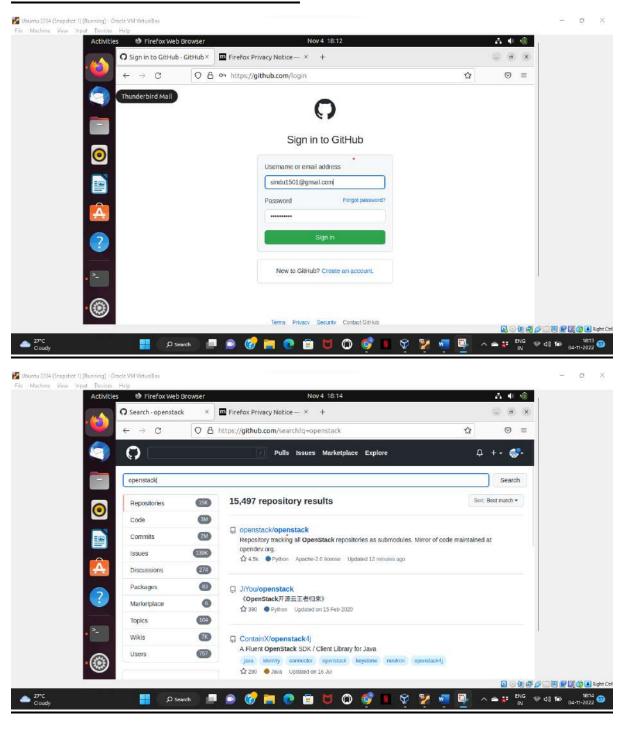
echo 'sindu ALL=(ALL) NOPASSWD: ALL" >> /etc/sudoers



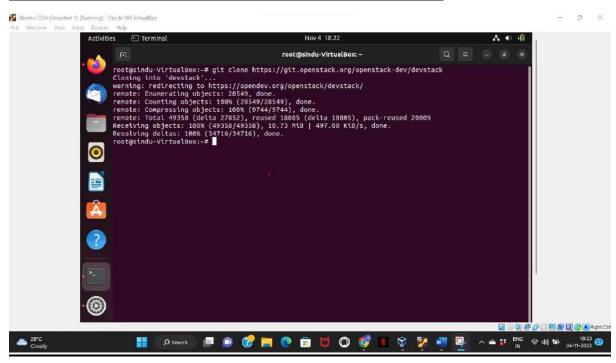
STEP 11:INSTALLING GIT USING THE COMMAND:sudo apt-get install git



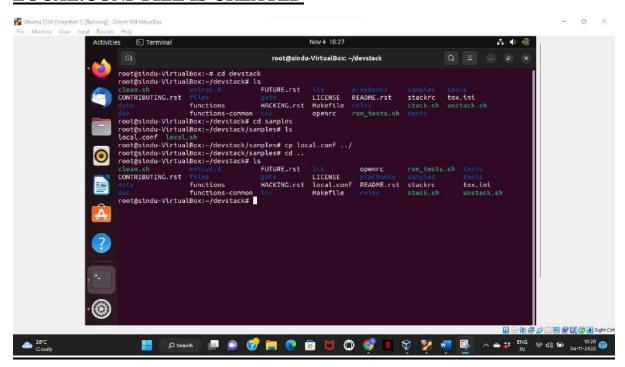
STEP 12:SIGNING INTO GITHUB



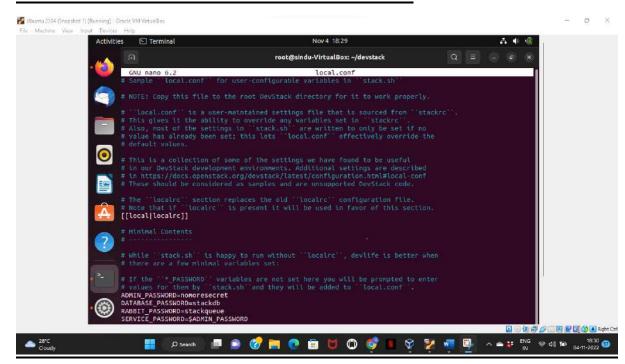
STEP 13:CLONING THE OPENSTACK USING THE COMMAND:git clone https://git.openstack.org/openstack-dev/devstack



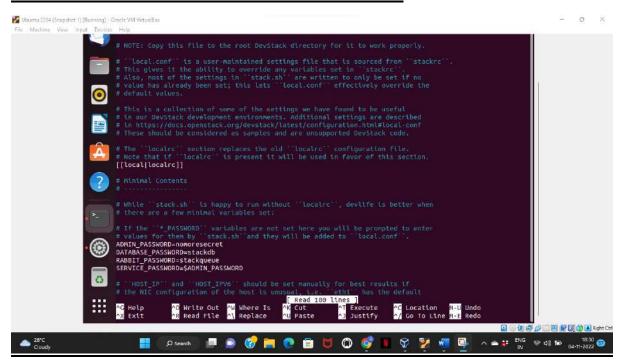
STEP 14:DEVSTACK CONFIGURATION FILE IS CREATED, LISTED AND DIRECTORY IS CHANGED TO SAMPLES AND LISTED LOCAL.CONF FILE IS CREATED



STEP 15:THE LOCAL.CONF FILE IS OPENED

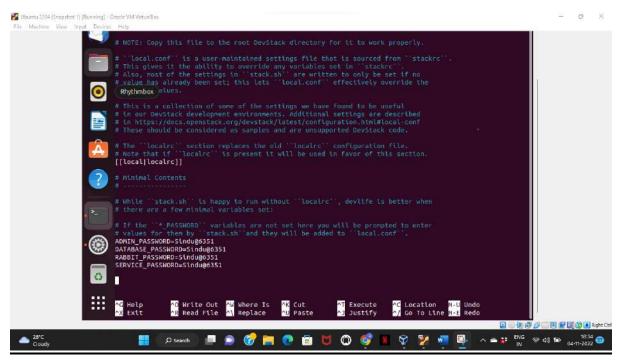


STEP 16:LOOKING FOR ADMIN PASSWORD



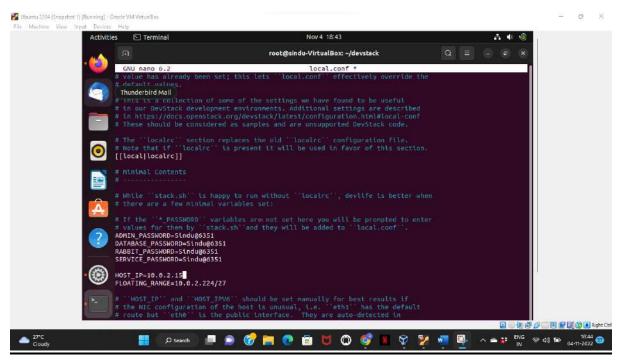
STEP 17: CHANGING THE ADMIN PASSWORD

THIS PASSWORD WILL BE USED TO LOG IN TO THE OPEN STACK LOGIN PAGE

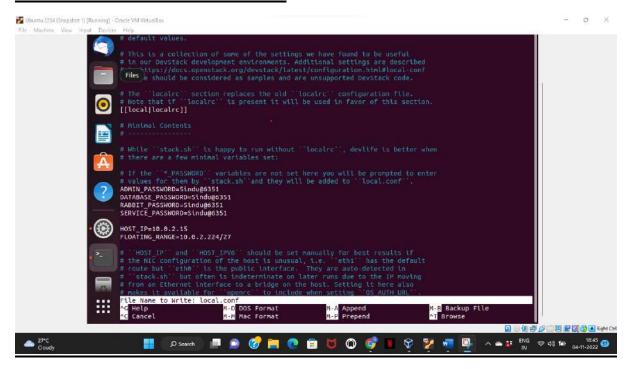


STEP 18:GIVING THE HOST IP(OUR SYSTEM IP) AND FLOATING_RANGE

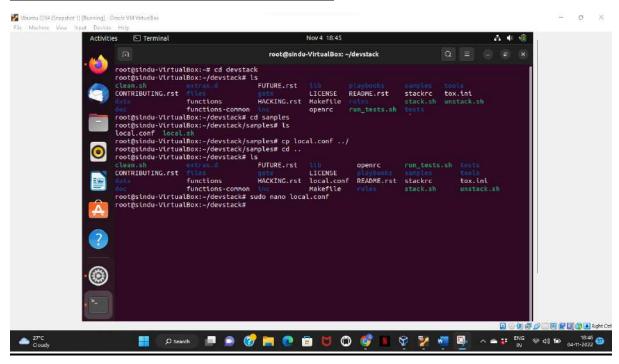
THE HOST-IP IS MY SYSTEM'S IP ADDRESS THAT IS OBTAINED BY RUNNING ifconfig or ipaddress COMMANDS



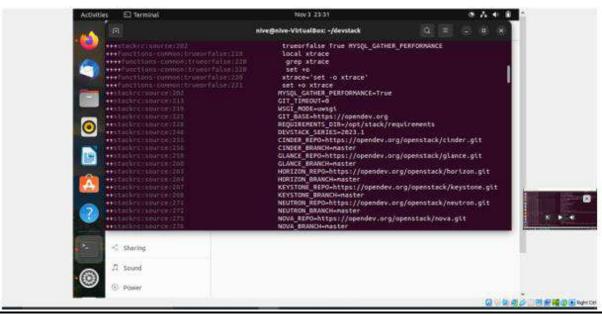
STEP 19:SAVE THE CHANGES TO THE LOCAL.CONF FILEUSING THE COMMAND CONTROL+X



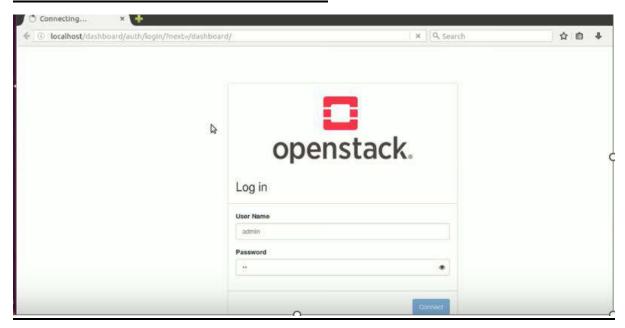
STEP 20:VIEWING THE SAVED FILE



STEP 21:THE OPENSTACK IS INSTALLED WITH DEVSTACK TO INSTALLATION OF THE OPENSTACK CAN BE COMMENCED BY USING THE COMMAND /stack.sh



STEP 22:DEPLOYING IN BROWSER



4.CASE STUDY ON CLOUD TM

AN APPROACH OF FIDELITY TOWARD CLOUD COMPUTING FOR ACHIEVING CLOUD PROABILITY

ABSTRACT

Fidelity's journey toward <u>cloud native</u> computing started in 2018 when the cloud platform team started providing Kubernetes as a platform to selected application teams, supported by a variety of managed cloud Kubernetes services. Rajarajan Pudupatti, a cloud platform architect, explains that "we looked at the challenges the applications teams have in general when implementing a brand-new technology." The team worked on developing a platform that also took into account Fidelity-specific criteria, particularly information security and data protection, and relied on an active feedback loop with developer focus groups.

CHALLENGES FACED

Fidelity Investments, one of the biggest financial services firms in the world, provides services to over 35 million investors through more than 76 million accounts. According to Amr Abdelhalem, SVP, Head of Cloud Platforms, the organisation started a digital transformation a few years ago with the goal of "leveraging next generation platforms and technologies to bo ost business value, increase speed to market, and harness the power of innovation."

Fidelity had Kubernetes distributions on-prem and on other cloud providers, which quickly became a problem.

SOLUTION

This initiative includes adopting a <u>multi-cloud approach</u>, which required the multi-year migration of thousands of crucial, stringently regulated, low-latency apps. A multi-level platform built on top of Fidelity <u>Cloud Fabric</u>, which is based on <u>CNCF technologies</u> like Kubernetes, would be the solution.

PROCESS

According to Niraj Amin, Cloud Platform Architect, "It's important to concentrate on developing uniformity across all of the platforms that a business unit may consume with an aim of a consistent developer experience." "Kubernetes is Kubernetes if it is. We therefore try to remove

some of the <u>challenges or distinctions between operating Kubernetes locally</u> versus using a certain cloud provider.

Pudupatti continues, "Now everything flows through the Kubernetes construct. A specific add-on can be easily implemented in a certain release, and it will affect the hundreds of Kubernetes clusters uniformly. A simple update causes all of the thousands of microservices that are active in a cluster to begin adhering to a specific security process. When they first started, this was one of the key points.

By developing its own operators and automating procedures particular to the financial services sector, Fidelity was able to fulfil some of its most critical regulatory and security requirements. They can actually plug in our own logic because of Kubernetes' extensibility feature, according to Pudupatti. When a problem arises, we can always refer to the Kubernetes design since there is always something we can do to fix it.

For instance, the group created and released KConnect, a CLI that enables users to sign up for the platform, find the <u>Kubernetes</u> <u>clusters</u> they have access to across various operating environments, and establish secure connections with them. Another operator was developed to restrict who can create namespaces and to ensure that, once they are formed, they adhere to the predetermined constraints. Amin claims that the entire procedure is automated. "Used to potentially be able to do other things like governance, linking constructs that we've established internally and building on top of that"

As they develop the platform, it is noted that "the very first thing they do is look at the <u>CNCF tool set</u> and the projects that are going in. They always considered the path of the community, so even if there is a simple solution, they tookthe stance of not pursuing anything short term because they believe the community is truly moving in a different route for specific reasons.

A multi-level platform built on top of what the team refers to as <u>Fidelity Cloud Fabric</u> is the final consequence of all this work. According to Pudupatti, "Fidelity Investments is leveraging a number of CNCF technologies to power our next generation of cloud native platforms." "We rely on <u>CNI for Networking API, CoreDNS for service discovery</u>, Fluentd for logging, Helm as a package manager, Kubernetes for container orchestration, and Open Policy Agent for policy administration. etcd serves as the KV store.

Additionally, they are using CNCF sandbox projects like Flux for GitOps and cert-manager for certificate administration.

According to Abdelhalem, "the Fabric itself is a method to build and innovate inside a multi-cloud, hybrid-cloud approach" for the thousands of application teams within Fidelity. "This multi-cloud provider manages our ecosystem itself, all of our application lifecycle management tools, our observability layers, our caching layers, our security and governance levels, and our AI and machine learning layers.

The Fidelity team admits that there is still much to be done in order for the corporation to make the transition to the cloud. According to Amin, some business units <u>are still in communication with on-premises services, accounts, and other cloud-based services via SaaS solutions.</u>
"Moving everything to the cloud will be difficult, but it is hoped that this platform will make it easier for developers to travel there. <u>With Kubernetes, they had a platform within their platform that is simple to frequently expand on.</u>

According to Pudupatti, the advantages have so far been obvious: "The CNCF technological breadth has had a big impact for Fidelity, as we have effortlessly sped application migrations to the cloud." They have reached about 3,000 Kubernetes services on the cloud, nearly 200 Kubernetes clusters, over 1,000 namespaces, and 10,000 containers in a short amount of time.

Given Fidelity's <u>data centre migrations and</u> <u>eventual shift of many apps to the cloud, portability has also been a huge asset.</u> "Migrating to a different cloud provider is now possible in a matter of hours, as opposed to the past, when it required months or was occasionally impossible. In the Kubernetes universe, anything we are running locally may theoretically operate on any cloud provider, according to Amin.

Cloud native will have an effect in a broader sense for years to come. The key impetus, according to Pudupatti, is Kubernetes. "One of the directives that they are getting from our CTO is a very aggressive cloud journey in a few years, and at the same time, prevent vendor lock-in, technology lock-in, and cloud lock-in. Right now, it is how cloud portability is being accomplished.

CONCLUSION

This initiative includes adopting a <u>multi-cloud</u> <u>approach</u>, which required the multi-year migration of thousands of crucial, stringently regulated, low-latency apps. A multi-level platform built on top of Fidelity <u>Cloud Fabric</u>, which is based on <u>CNCF technologies</u> like Kubernetes, would be the solution.