**CLOUD COMPUTING**

**ASSIGNMENT 3**

**Avinash Ganguri**

**16293133**

**Step1: Downloaded the Virtual box hosted hypervisor for Mac**

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

**Step2: Download Hortonworks HDP Image from the site,**

A screenshot of a video game

Description automatically generated

A screenshot of a cell phone

Description automatically generated

And my system configurations are,

A screenshot of a cell phone

Description automatically generated

*To run the Image Hortonworks HDP in Virtual Box, we need* ***RAM*** *more than 8GB. On running the virtual box my system crashes for every 5 mins, so I have to choose another alternative for running Hortonworks HDP Image. Therefore, I used* ***Microsoft Azure services*** *to create a Virtual Machine and run the* ***Hortonworks HDP 2.5 Version*** *in it***.**

A screenshot of a computer screen

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

And then connecting the VM to my localhost through SSH,

A close up of a screen

Description automatically generated

**Step 4:**

Using the public IP from Azure VM, open it in the browser,

Enter username: maria\_dev, password: maria\_dev

A screenshot of a social media post

Description automatically generated

The Ambari login page shows up, after successful login you can see the dashboard,

A screenshot of a computer

Description automatically generated

Changing the username and password and also, creating admin credentials by,

A screenshot of a cell phone

Description automatically generated

A screenshot of a social media post

Description automatically generated

After logging with the admin details, the dashboard page

A screenshot of a computer

Description automatically generated

**Step 5: Explore Ambari using web browser. Explore its capabilities and show at least three things with screenshots of its function (eg. installing other modules, setting system alert parameter and method, etc.)**

1. A Log Search Service is added, and only an Admin can add it,so we have to login only using admin details.

A screenshot of a computer screen

Description automatically generated

A screenshot of a computer

Description automatically generated

After successfully deploying,

A screenshot of a computer screen

Description automatically generated

A screenshot of a computer screen

Description automatically generated

A screenshot of a computer

Description automatically generated

1. Adding NameNode Host CPU Utilization Alert by changing the interval from 5 minsto 10 mins and connection timeout 5 secs to 10 secs.

A screenshot of a computer

Description automatically generated

1. Adding Ambari Metrics Service,

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Creating admin password,

A screenshot of a computer screen

Description automatically generated

After succesfully deploying,

A screenshot of a computer screen

Description automatically generated

A screenshot of a cell phone

Description automatically generated

**Step 6: To explore Ambari -> Pig View and find top 5 highly rated movies.**

This is the file view, where we upload the csv files such as ratings.csv here,

A screenshot of a computer

Description automatically generated

In the Pig view, a new script is created and we write our code here,

A screenshot of a cell phone

Description automatically generated

A screenshot of a social media post

Description automatically generated

Our script file is executed here and we get the Top 5 results,  
A screenshot of a cell phone

Description automatically generated

**Step 7: Do the above Step 6 with Tez enabled,**

A screenshot of a social media post

Description automatically generated

A screenshot of a cell phone

Description automatically generated

A screenshot of a computer

Description automatically generated

With the Tez enabled, the script was executed in 1 min 12 secs and with the Mapreduce it was executed in 1 min 55 sec. Hence, *Tez runs faster than MapReduce.*