SNJB's Late Sau. K. B. Jain College of Engineering, Chandwad

Department of Computer Engineering

Course Name: Laboratory Practice II(310258):Cloud Computing

Class: Third Year (TE) Div A/ Div B

Batch:T1/T2/T3/T4

Name:

Roll No:

Assignment No: 11

Answers (A) – 5M	Coding Efficiency (C) – 5M	Viva (V) – 5M	Timely Completion (T) – 5M	Total(20M)	Sign

Date of Performance:..... **Date of Completion:**.....

1. Title of Assignment: Mini-Project

Setup your own cloud for Software as a Service (SaaS) over the existing LAN in your laboratory. In this assignment you have to write your own code for cloud controller using open-source technologies to implement with HDFS. Implement the basic operations may be like to divide the file in segments/blocks and upload/ download file on/from cloud in encrypted form.

2. Objective:

1. To Learn how to setup your own LAN
2. To design application using HDFS

3. Outcome: Design application using HDFS

4. Software and Hardware Requirement:

Software Requirement:

Hardware Requirement: Internet Connection, PC with Min. 2GB RAM, Core i5 Processor

5.Relevant Theory :

Steps for Mini Project

- Set up a local area network (LAN) in your laboratory.
- This can be done by connecting all the devices in your laboratory to a switch or a router.
- Make sure that all the devices are on the same network and can communicate with each other.
- Next, you need to choose an open-source cloud controller technology that can be installed on your LAN.
- There are several options available such as OpenStack, CloudStack, and Eucalyptus. You can choose any one of these or any other that suits your needs.
- After selecting a cloud controller technology, you need to install it on a dedicated server or a virtual machine.
- Once the cloud controller is installed, you can start configuring it by creating users, assigning roles, and setting up storage.
- To implement HDFS (Hadoop Distributed File System), you need to install the Hadoop framework.
- Hadoop is an open-source software framework that provides support for distributed storage and processing of large data sets.
- Once Hadoop is installed, you can configure it to use HDFS as the underlying file system.
- To implement the basic operations of dividing the file into segments/blocks and uploading/downloading the file from the cloud in encrypted form, you can use HDFS commands.
- HDFS provides commands for uploading and downloading files, and also for dividing the files into segments/blocks.
- Hadoop Distributed File System (HDFS) is a distributed file system that allows for the storage and processing of large datasets across clusters of commodity hardware. While HDFS does provide commands for managing files, it is not designed to handle encryption natively.
- To upload and download files in encrypted form, you would need to use a third-party tool or library that supports encryption, such as OpenSSL or the Java Cryptography Extension (JCE). You would then need to integrate these tools into your HDFS workflow to ensure that files are properly encrypted and decrypted during transfer.
- To divide a file into segments or blocks, HDFS provides a command called hdfs dfs -Ddfs.blocksize=<blocksize> -put <src> <dest>. This command allows you to specify the block size in bytes and the source and destination paths for the file. By default, HDFS uses a block size of 128 MB.

- Once the file is uploaded to HDFS, you can use commands such as hdfs dfs -get <src> <dest> to download the file in its encrypted form. If you need to decrypt the file, you would need to use a separate decryption tool or library after the file has been downloaded.
- It is worth noting that there are also other tools and technologies, such as Amazon S3 and Azure Blob Storage, that offer native support for encryption and may be more suitable for certain use cases.
- Overall, setting up your own cloud for SaaS over the existing LAN in your laboratory requires several steps, including choosing an open-source cloud controller technology, installing and configuring it, setting up HDFS, and implementing the basic operations of dividing the file into segments/blocks and uploading/downloading the file from the cloud in encrypted form. It can be a complex task, but with careful planning and implementation, it can be achieved successfully.

How to Install HDFS:installation process for HDFS:

1. Download the Hadoop distribution: Go to the Apache Hadoop website and download the latest stable release of Hadoop. You can choose to download either the binary distribution or the source code.
2. Install Java: Hadoop requires Java to run, so make sure that you have a compatible version of Java installed on your system. You can check the Hadoop documentation to find the specific version of Java that is required for your version of Hadoop.
3. Configure environment variables: Set the HADOOP_HOME environment variable to the directory where you extracted the Hadoop distribution. Add the bin directory to your PATH environment variable.
4. Configure Hadoop: Edit the core-site.xml, hdfs-site.xml, and mapred-site.xml files in the conf directory of the Hadoop distribution. These files contain configuration settings for Hadoop, such as the location of the HDFS name node and data nodes.
5. Create the necessary directories: Create the directories that Hadoop requires for storing data and logs. The default directories are /tmp/hadoop-${user.name} and /var/log/hadoop-${user.name}.
6. Set up permissions: Make sure that the directories created in the previous step are writable by the user that will be running Hadoop. You may also need to create a Hadoop user and group and assign ownership of the Hadoop directories to this user and group.
7. Start HDFS: Use the start-dfs.sh script in the sbin directory of the Hadoop distribution to start the HDFS daemons. Use the jps command to verify that the daemons are running.

Frequently Asked Questions:

i) What is the use of HDFS?

ii) How to encrypt and decrypt files?

iii) Explain OpenStack, CloudStack?

iv) Explain OpenSSL and Java Cryptography Extension (JCE).?

v) how to divide files into segments and upload into the cloud?

Conclusion:

Hence, successfully understand LAN setup and HDFS.