A.I.S.S.M.S
COLLEGE OF ENGINEERING
NO - 1, Kennedy Road, Near R.T.O Pune - 411001

Practical - 05.

- AIM: Setup your own cloud for Software as a Service (SaaS) over the existing LAN. 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 / block and upload / download on/from cloud in encrypted form.

- Objective: To make own cloud SaaS.

- THEORY :

1) Cloud computing and Saas.

- Cloud computing - refers to the delivery of computing resources like servers, storage, databases, networking, software and analytics over the internet (the 'cloud')

- Software as a service (Saas) - is a cloud based service where applications are hosted on the cloud and made available to users over the internet. Instead of installing and maintaining software, users access it via a web browser.

A.I.S.S.M.S
COLLEGE OF ENGINEERING
NO - 1, Kennedy Road, Near R.T.O Pune - 411001

2] HDFS (Hadoop distributed file system) -

- HDFS is a distributed file system designed to store large volumes of data across multiple machines in a cluster. It's a part of the Apache Hadoop ecosystem and ensures high availability, fault tolerance and scalability.
- HDFS stores data in large blocks (typically 128 MB or 256 MB) across nodes in a cluster. These blocks are replicated across nodes to ensure redundancy in case of node failure.

3] Encryption -

- Process of encoding data to prevent unauthorized access. This ensures confidentiality of data being stored or transmitted.

- How to configure 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 specific version of java needed for hadoop.

A.I.S.S.M.S
COLLEGE OF ENGINEERING
NO - 1, Kennedy Road, Near R.T.O Pune - 411001

(3). Configure environment variables : SET the Hadoop_home environment variable to the directory where you extracted hadoop distribution. Add bin directory to your PATH environment variable.

(4). Configure Hadoop - Edit core site.xml file, hdfs-site.xml. and map-red-site.xml files in conf directory of Hadoop distribution. These files contain configuration settings HDFS name node, data nodes.

(5). Create necessary directories: that Hadoop requires for storing data and logs. The default directories are /tmp/hadoop-${user.name} and /var/log/hadoop.${}

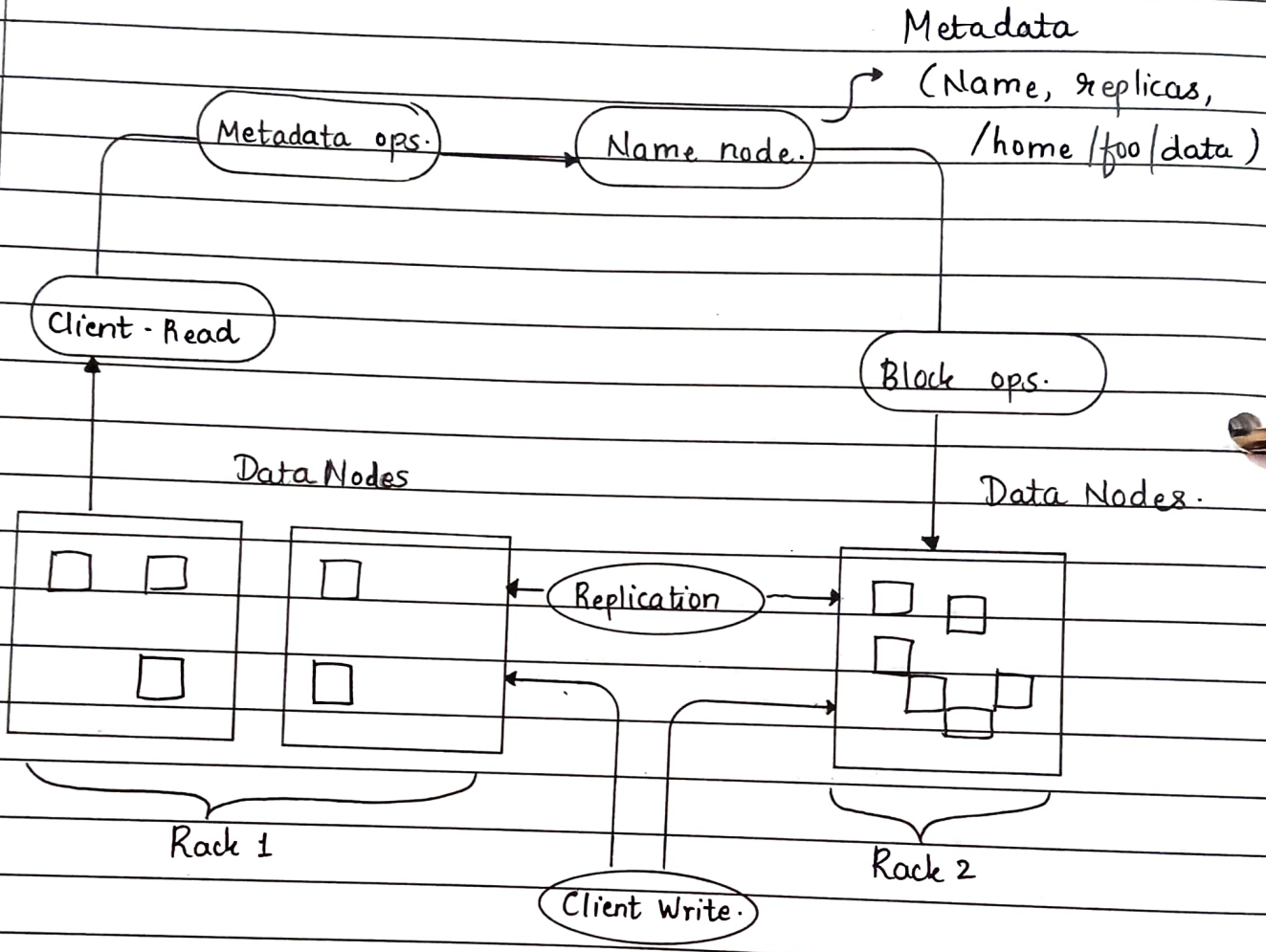
(6). Setup permissions - Make sure that the directories created in previous step are writable by user that will be running hadoop.

(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 daemons are running.

- The HDFS architecture is as given below -

→

A.I.S.S.M.S
COLLEGE OF ENGINEERING
NO - 1, Kennedy Road, Near R.T.O Pune - 411001



HDFS Architecture.

Conclusion : Hence we successfully configured the HDFS as a software as a service.