

EXERCISE NO.: 01

INSTALLATION OF HADOOP

AIM:

To install and configure Hadoop to enable distributed storage and processing of large datasets using the Hadoop Distributed File System (HDFS) and MapReduce framework.

PROCEDURE:

1. Install Java Development Kit (JDK) as Hadoop requires Java.
2. Set the JAVA_HOME environment variable.
3. Download the latest stable Hadoop tarball from the Apache Hadoop official website.
4. Extract the tarball to the desired installation directory.
5. Update .bashrc or .profile.
6. Source the file to apply changes.
7. Edit Hadoop Configuration Files
 - a. core-site.xml: Set the default filesystem.
 - b. hdfs-site.xml: Configure NameNode and DataNode directories.
 - c. mapred-site.xml: Configure the MapReduce framework.
 - d. yarn-site.xml: Configure ResourceManager and NodeManager.
8. Format the Hadoop Filesystem
9. Start Hadoop Daemons
10. Verify running daemons using jps.
11. Create directories in HDFS.
12. Copy a sample file to HDFS and list files.

RESULT:

Thus, Hadoop was successfully installed and configured on the system.