Matt Ranillo M. Orbita IT31

Introduction

Hadoop is a distributed file system that enables processing and fault tolerance simultaneously. For faster storage and recovery of data from its nodes, the Hadoop MapReduce programming model is used. Hadoop is an open-source software framework for storing data and running applications on clusters of commodity hardware. It provides massive storage for any kind of data, enormous processing power and the ability to handle virtually limitless concurrent tasks or jobs. For storing and analyzing big data, Hadoop is used. Data is stored in Hadoop on inexpensive commodity servers operating as clusters.

How to Install Hadoop

Step 1: https://hadoop.apache.org/releases.html
Step 2: choose 3.3.0 and click the binary
Step 3: After clicking The binary click the https://downloads.apache.org/hadoop/common/hadoop-3.3.0/hadoop-3.3.0.tar.gz
Step 4: Before you install the hadoop you must install the JDK. Link: https://www.oracle.com/java/technologies/javase/javase-jdk8-downloads.html
Step 5: Install the JDK First Then Move The Jdk Folder to local C.
Step 6: Now go to Environment variable for the setup. Click the Environment Variable then click the New button then Input the Variable Name:(it depends on you) Variable Value:(the location of file/ Path of the folder).
Step 7: In System Variable Click the Path Column then paste the (Path of the file) also.
Step 8: Do this in CMD: cd Program Files\Java\jdk-13.0.2\bin then type javac.
Step 9: After You installed the JDK u must install now the hadoop file. First you need to extract the file then after you extracted the Zip file. now move the file to the local C.
Step 10: After Extracted the Zip file of hadoop. Now go to ETC folder then find the. -- core-site-- xml mode -- hadoop-env-- cmd mode -- mapred-site -- xml mode -- yarn site -- xml mode -- hdfs site - xml mode
Step 11: Create folder inside the hadoop folder (data).
Step 12: Open the folder that you created then create folder again for (datanode and namenode).

Importance of Hadoop

Hadoop is an open source platform that provides excellent data management provision. It is a framework that supports the processing of large data sets in a distributed computing environment. It is designed to expand from single servers to thousands of machines, each providing computation and storage. Its distributed file system facilitates rapid data transfer rates among nodes and allows the system to continue operating uninterrupted in case of a node failure, which minimizes the risk of catastrophic system failure, even if a significant number of nodes become out of action. Hadoop is very valuable for large scale businesses basing on its proven benefits for enterprises given below:

Advantages for Enterprises:

- Hadoop is now widely used across industries, including finance, media and entertainment, government, healthcare, information services, retail, and other industries.
- It facilitates businesses to easily access new data sources and tap into different types of data to produce value from that data.
- Unique storage method of Hadoop is based on a distributed file system that basically 'maps' data wherever it is located on a cluster. The tools for data processing are often on the same servers where the data is located, resulting in much faster data processing.
- It is a highly scalable storage platform.
- Hadoop provides a cost effective storage solution for business.

Stored Data to Data Node

If you successfully connected to Hadoop then do this command:

Step 1: C:\Users\artam>hadoop fs –mkdir /sample_dir (ex jem_dir)
Step 2: C:\Users\artam>hadoop fs –put [File_location_that_you_want_to_stored] /sample_dir
Step 3: C:\Users\artam>hadoop fs –cat /sample_dir/[File_location_that_you_want_to_stored]

Explanation:

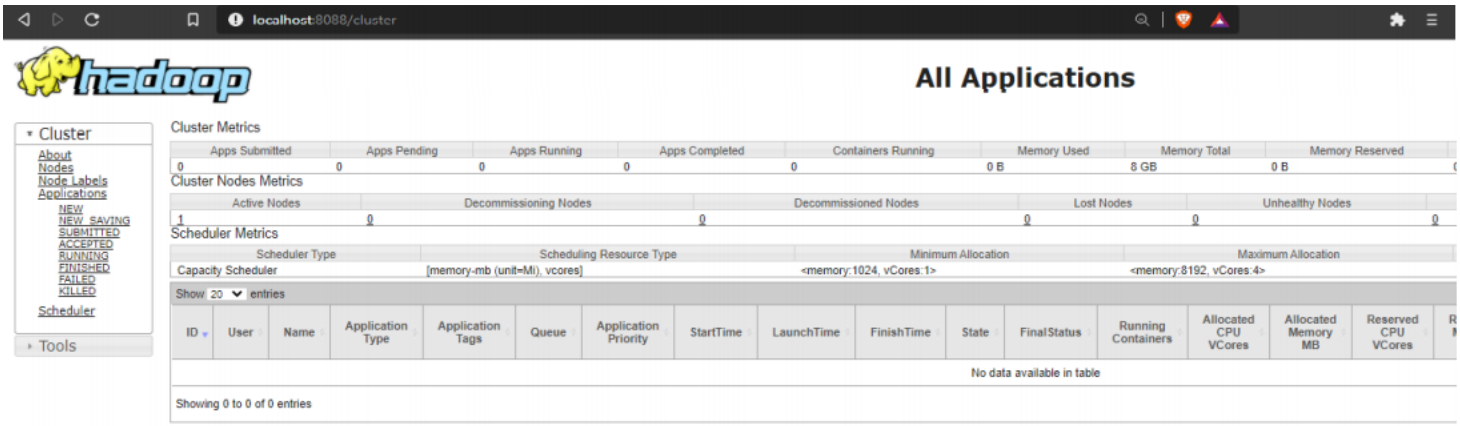
Step 1: You are making a file directory in DATANODE.
Step 2: Inserting File to the sample_dir where sample_dir is a folder in DATANODE.
Step 3: Display Data from DATANODE

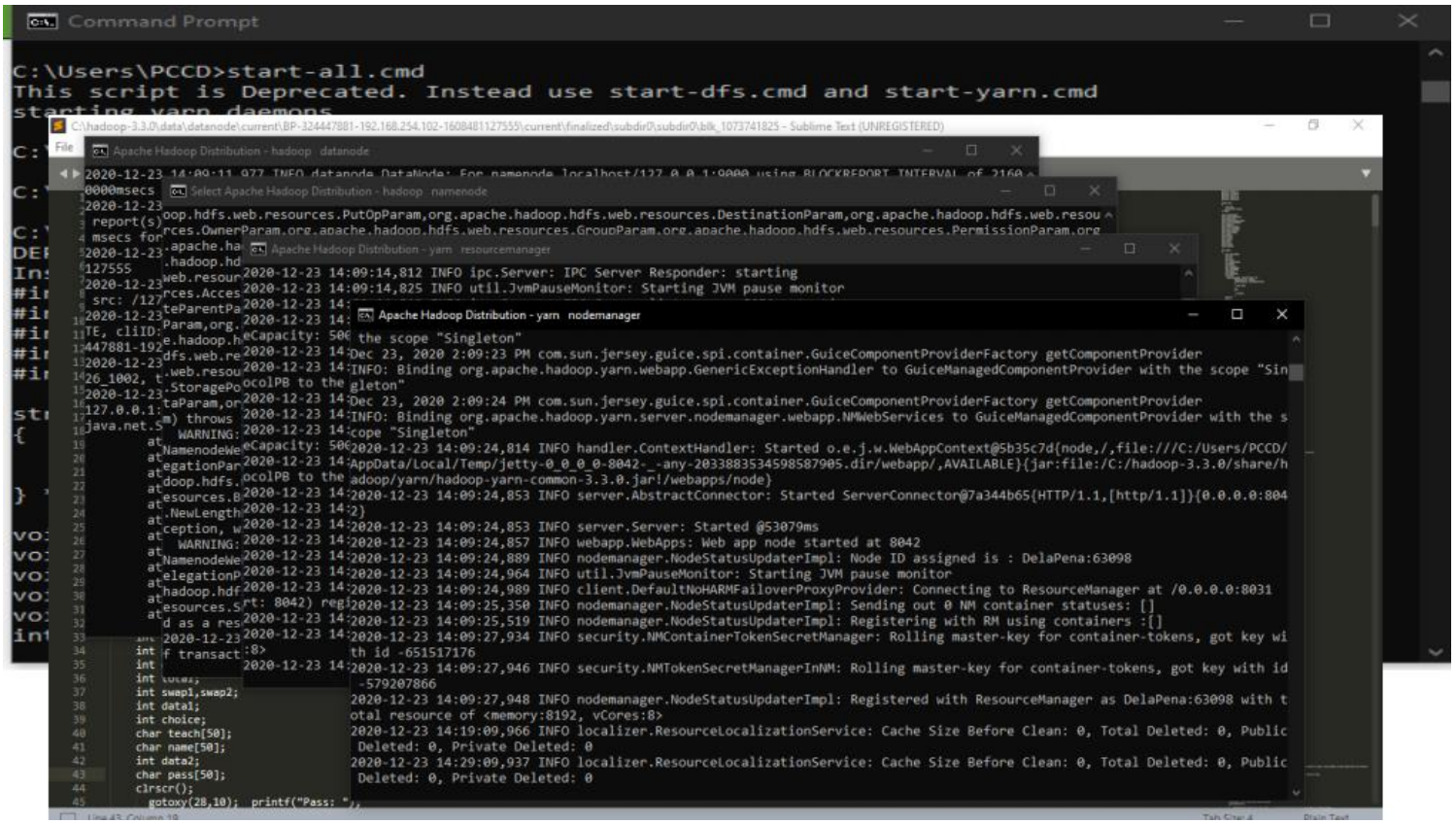
Other Command:

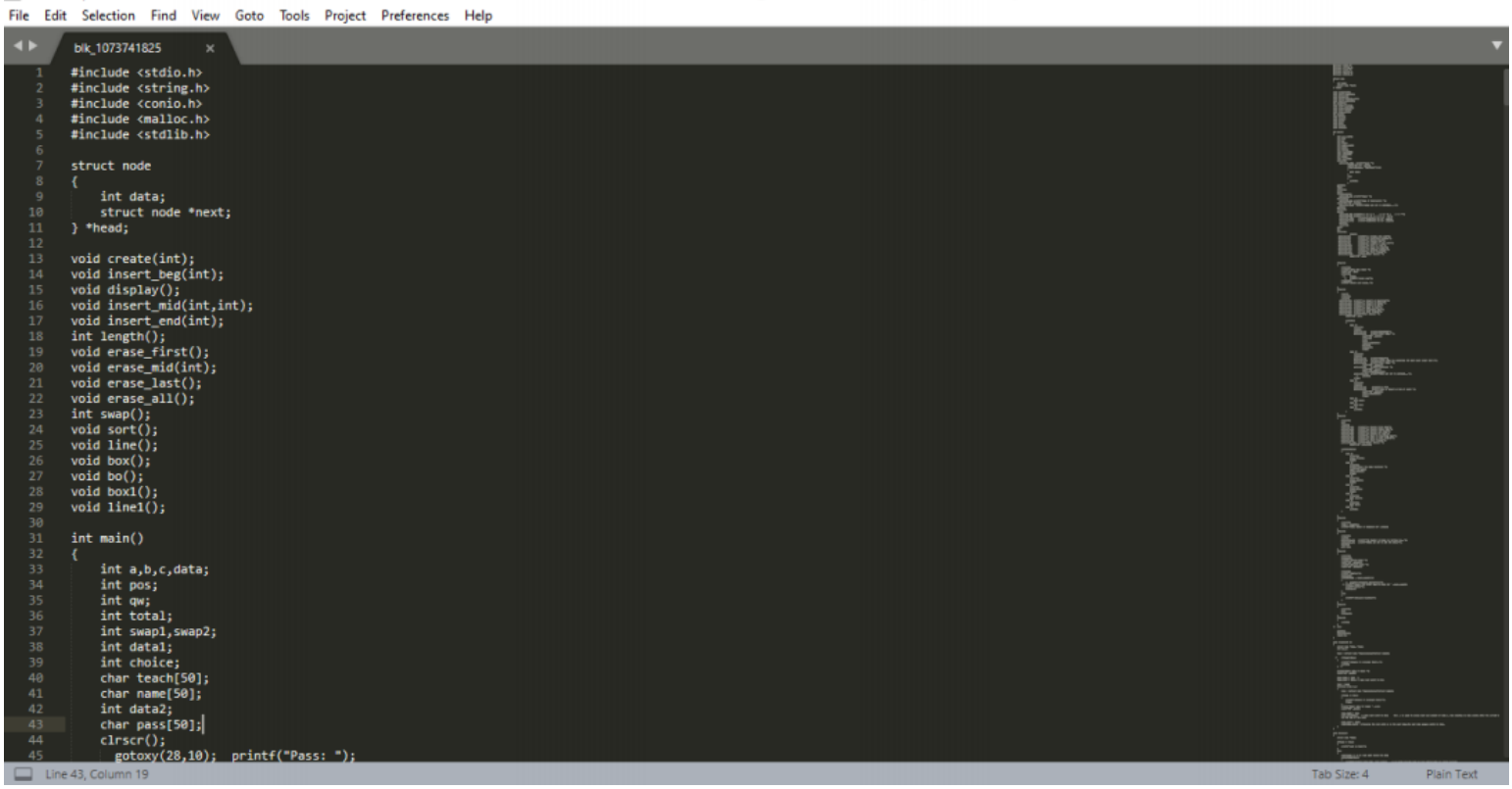
- C:\Users\artam> hadoop dfsadmin -safemode leave
- C:\Users\artam> hadoop dfsadmin –safemode enter
- C:\Users\artam> hadoop fs -rm -r /input_dir/sample.txt
- C:\Users\artam> hadoop fs -rm -r /input_dir

Explanation:

- You are leaving in a Safe Mode.
- You are Entering in a Safe mode.
- Deleting file/data inside the DATANODE where you pointing to inside the input_dir.
- Deleting input_dir where input_dir is a folder of DATANODE.




Command Prompt
C:\Users\PCCD>start-all.cmd
This script is Deprecated. Instead use start-dfs.cmd and start-yarn.cmd
starting yarn daemons
Apache Hadoop Distribution - hadoop datanode
Select Apache Hadoop Distribution - hadoop namenode
Apache Hadoop Distribution - yarn resourcemanager
Apache Hadoop Distribution - yarn nodemanager
2020-12-23 14:09:14,812 INFO ipc.Server: IPC Server Responder: starting
2020-12-23 14:09:14,825 INFO util.JvmPauseMonitor: Starting JVM pause monitor
2020-12-23 14:09:24,814 INFO handler.ContextHandler: Started o.e.j.w.WebAppContext@5b35c7d{node,/,file:///C:/Users/PCCD/AppData/Local/Temp/jetty-0_0_0_0-8042-_-any-2033883534598587905.dir/webapp/,AVAILABLE}{jar:file:/C:/hadoop-3.3.0/share/hadoop/yarn/hadoop-yarn-common-3.3.0.jar!/webapps/node}
2020-12-23 14:09:24,853 INFO server.AbstractConnector: Started ServerConnector@7a344b65{HTTP/1.1,[http/1.1]}{0.0.0.0:8042}
2020-12-23 14:09:24,853 INFO server.Server: Started @53079ms
2020-12-23 14:09:24,857 INFO webapp.WebApps: Web app node started at 8042
2020-12-23 14:09:24,889 INFO nodemanager.NodeStatusUpdaterImpl: Node ID assigned is : DelaPena:63098
2020-12-23 14:09:24,964 INFO util.JvmPauseMonitor: Starting JVM pause monitor
2020-12-23 14:09:24,989 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8031
2020-12-23 14:09:25,350 INFO nodemanager.NodeStatusUpdaterImpl: Sending out 0 NM container statuses: []
2020-12-23 14:09:25,519 INFO nodemanager.NodeStatusUpdaterImpl: Registering with RM using containers :[]
2020-12-23 14:09:27,934 INFO security.NMContainerTokenSecretManager: Rolling master-key for container-tokens, got key with id -651517176
2020-12-23 14:09:27,946 INFO security.NMTokenSecretManagerInNM: Rolling master-key for container-tokens, got key with id -579207866
2020-12-23 14:09:27,948 INFO nodemanager.NodeStatusUpdaterImpl: Registered with ResourceManager as DelaPena:63098 with total resource of <memory:8192, vCores:8>
2020-12-23 14:19:09,966 INFO localizer.ResourceLocalizationService: Cache Size Before Clean: 0, Total Deleted: 0, Public Deleted: 0, Private Deleted: 0
2020-12-23 14:29:09,937 INFO localizer.ResourceLocalizationService: Cache Size Before Clean: 0, Total Deleted: 0, Public Deleted: 0, Private Deleted: 0


File Edit Selection Find View Goto Tools Project Preferences Help
blk_1073741825
#include <stdio.h>
#include <string.h>
#include <conio.h>
#include <malloc.h>
#include <stdlib.h>
struct node
{
int data;
struct node *next;
} *head;
void create(int);
void insert_beg(int);
void display();
void insert_mid(int,int);
void insert_end(int);
int length();
void erase_first();
void erase_mid(int);
void erase_last();
void erase_all();
int swap();
void sort();
void line();
void box();
void bo();
void box1();
void line1();
int main()
{
int a,b,c,data;
int pos;
int qw;
int total;
int swap1,swap2;
int data1;
int choice;
char teach[50];
char name[50];
int data2;
char pass[50];
clrscr();
gotoxy(28,10); printf("Pass: ");
Line 43, Column 19
Tab Size: 4
Plain Text