

Lab 5

-Các bước cài đặt

1. tải các file
 - Tải jdk 8u_321 :<https://www.oracle.com/java/technologies/downloads/#java8-windows>
 - Tải hadoop 3.2.2 : <https://hadoop.apache.org/release/3.2.2.html>
 - bin 3.2.2 : <https://github.com/hmtdev/winutils>
 - Cài Path :
 - JAVA_HOME = C:\java\jdk1.8.0
 - HADOOP_HOME = C:\hadoop-3.2.2
 - %JAVA_HOME%\bin
 - %HADOOP_HOME%\bin
 - %HADOOP_HOME%\sbin
 - Kiểm tra java với hadoop
 - java -version
 - hadoop -version
 - Cấu hình các file :
 - hadoop-env.cmd
 - core-site.xml

```
<configuration>
    <property>
        <name>fs.defaultFS</name>
        <value>hdfs://localhost:9000</value>
    </property>
</configuration>
```

- hdfs-site.xml

```
<configuration>
  <property>
    <name>dfs.replication</name>
    <value>1</value>
  </property>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>C:\hadoop-3.2.2\data\namenode</value>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>C:\hadoop-3.2.2\data\datanode</value>
  </property>
</configuration>
```

- **hdfs-site.xml**

```
<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
</configuration>
```

- **yarn-site.xml**

```
<configuration>
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
  <property>
    <name>yarn.nodemanager.auxservices.mapreduce.shuffle.class</name>
    <value>org.apache.hadoop.mapred.ShuffleHandler</value>
  </property>
</configuration>
```

- link github cua trung: <https://scarce-freedom-e76.notion.site/Lab-5-d0fb7f0a2cc248a494aed3e1f445d92a>
- run cmd admin
- bin\hdfs namenode -format

- sbin\start-all.cmd
- hdfsadmin -safemode leave
-