**BÁO CÁO TUẦN 10**

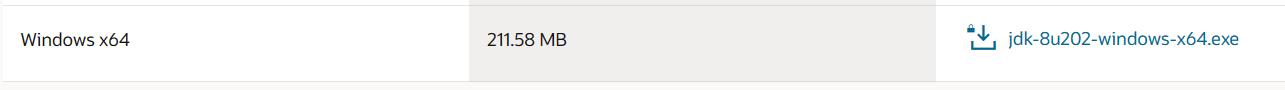
**CÀI ĐẶT HADOOP TRÊN WINDOW**

# 1. CÀI ĐẶT JDK

## 1.1 Tải JDK 1.8

Hadoop sử dụng JDK 1.8

Ta vào link sau: [Java Archive Downloads - Java SE 8 (oracle.com)](https://www.oracle.com/java/technologies/javase/javase8-archive-downloads.html)

****

Sau khi tải về ta có file



Sau đó chạy để cài đặt các gói về

A screenshot of a computer

Description automatically generated

## 1.2 Cài đặt biến môi trường JDK 1.8

Có thể tìm ở thanh công cụ “Edit the system environment variables”

**A black text on a white background

Description automatically generated**

Chọn “Environment variables”

**A screenshot of a computer program

Description automatically generated**

Dẫn đường dẫn phù hợp cho JAVA\_HOME

**A screenshot of a computer

Description automatically generated**

Và Path

**A white rectangular object with black lines

Description automatically generated**

Kiểm tra xem hệ thống đã nhận jdk chưa

**A black background with white text

Description automatically generated**

# 2. CÀI ĐẶT HADOOP

## 2.1 Tải Hadoop

Hadoop sử dụng Hadoop-3.3.6

Ta vào link sau: https://hadoop.apache.org/release/3.3.6.html

Sau khi tải về và giải nén

A screenshot of a computer

Description automatically generated

## 2.2 Cài đặt biến môi trường Hadoop

Có thể tìm ở thanh công cụ “Edit the system environment variables”

**A black text on a white background

Description automatically generated**

Chọn “Environment variables”

**A screenshot of a computer program

Description automatically generated**

Dẫn đường dẫn phù hợp cho JAVA\_HADOOP

**A screenshot of a computer

Description automatically generated**

Và Path

****

# 3. CẤU HÌNH CÁC FILE CỦA HADOOP

## 3.1 Tạo folder data

Tạo folder “data”

Tạo folder “datanode” và “namenode” trong folder “data”

A screenshot of a computer

Description automatically generated

## 3.2 core-site.xml

Vào folder “hadoop “ trong “etc”

“S:\Applications\Hadoop\hadoop-3-3-6\etc\hadoop\core-site.xml”

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

A screenshot of a computer

Description automatically generated

## 3.3 mapred-site.xml

Vào folder “hadoop “ trong “etc”

“S:\Applications\Hadoop\hadoop-3-3-6\etc\hadoop\mapred-site.xml”

|  |
| --- |
| <configuration>  <property>  <name>mapreduce.framework.name</name>  <value>yarn</value>  </property>  </configuration> |

A screenshot of a computer

Description automatically generated

## 3.4 hdfs-site.xml

Vào folder “hadoop “ trong “etc”

“S:\Applications\Hadoop\hadoop-3-3-6\etc\hadoop\hdfs-site.xml”

|  |
| --- |
| <configuration>  <property>  <name>dfs.replication</name>  <value>1</value>  </property>  <property>  <name>dfs.namenode.name.dir</name>  <value>/hadoop-3.3.6/data/namenode</value>  </property>  <property>  <name>dfs.datanode.data.dir</name>  <value>/hadoop-3.3.6/data/datanode</value>  </property>  </configuration> |

**A screenshot of a computer

Description automatically generated**

## 3.5 yarn-site.xml

Vào folder “hadoop “ trong “etc”

“S:\Applications\Hadoop\hadoop-3-3-6\etc\hadoop\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> |

A screenshot of a computer

Description automatically generated

## 3.6 hadoop-env.cmd

Vào folder “hadoop “ trong folder “etc”

“S:\Applications\Hadoop\hadoop-3-3-6\etc\hadoop\hadoop-env.cmd”

Setup biến môi trường

|  |
| --- |
| @rem The java implementation to use.  Required.  set JAVA\_HOME=S:\Applications\Java\jdk-1.8 |

A screenshot of a computer

Description automatically generated

# 4. Cập nhật file bin CẬP NHẬT FILE BIN

Tải từ link này: <https://github.com/s911415/apache-hadoop-3.1.0-winutils>

Tải về và extract ra

A screenshot of a computer

Description automatically generated

Tìm folder gần với phiên bản của phiên bản mình dùng nhất

Em dùng hadoop-3.3.6 nên em sẽ dùng folder hadoop-3.3.5

A screenshot of a computer

Description automatically generated

Copy file bin và đè lên file bin của “S:\Applications\Hadoop\hadoop-3-3-6”

A screenshot of a computer

Description automatically generated

# 5. FORMAT HỆ THỐNG

## 5.1 Format

Hdfs namenode -format



Hdfs datanode -format



## 5.2 Chép file

“S:\Applications\Hadoop\hadoop-3-3-6\share\hadoop\yarn\timelineservice”

Copy file hadoop-yarn-server-timelineservice-3.3.6

Và dán ra “S:\Applications\Hadoop\hadoop-3-3-6\share\hadoop\yarn”

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

# 6. Chạy thử

Start-all

A computer screen with white text

Description automatically generated

Namenode

A screenshot of a computer

Description automatically generated

Datanode

A screenshot of a computer

Description automatically generated

Nodemanager

A screenshot of a computer

Description automatically generated

Resourcemanager

A screenshot of a computer

Description automatically generated

**MỤC LỤC**

[1. CÀI ĐẶT JDK 1](#_Toc163423932)

[1.1 Tải JDK 1.8 1](#_Toc163423933)

[1.2 Cài đặt biến môi trường JDK 1.8 1](#_Toc163423934)

[2. CÀI ĐẶT HADOOP 3](#_Toc163423935)

[2.1 Tải Hadoop 3](#_Toc163423936)

[2.2 Cài đặt biến môi trường JDK 1.8 3](#_Toc163423937)

[3. CẤU HÌNH CÁC FILE CỦA HADOOP 4](#_Toc163423938)

[3.1 Tạo folder data 4](#_Toc163423939)

[3.2 core-site.xml 5](#_Toc163423940)

[3.3 mapred-site.xml 6](#_Toc163423941)

[3.4 hdfs-site.xml 7](#_Toc163423942)

[3.5 yarn-site.xml 8](#_Toc163423943)

[3.6 hadoop-env.cmd 9](#_Toc163423944)

[4. Cập nhật file bin CẬP NHẬT FILE BIN 9](#_Toc163423945)

[5. FORMAT HỆ THỐNG 10](#_Toc163423946)

[5.1 Format 10](#_Toc163423947)

[5.2 Chép file 10](#_Toc163423948)

[6. Chạy thử 11](#_Toc163423949)