

## Downloading and installing Hadoop, Understanding different Hadoop modes, Startup scripts, Configuration files.

### AIM:


To Download and install Hadoop, Understanding different Hadoop modes, Startup scripts, Configuration files.

### PROCEDURE:

**1. Install Java 8:** Download Java 8 from the link:

<http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html>

- a. Set environmental variables:
  - i. User variable:
    - Variable: JAVA\_HOME
    - Value: C:/java
  - ii. System variable:
    - Variable: PATH
    - Value: C:/java/bin
- b. Check on cmd, see below:



```
Command Prompt
Microsoft Windows [Version 10.0.19045.4780]
(c) Microsoft Corporation. All rights reserved.

C:\Users\user>cd \

C:\>java -version
java version "1.8.0_202"
Java(TM) SE Runtime Environment (build 1.8.0_202-b08)
Java HotSpot(TM) 64-Bit Server VM (build 25.202-b08, mixed mode)

C:\>
```

**2. Download Hadoop-3.3.6:** download Hadoop 3.3.6 from the link:

<http://www.apache.org/dyn/closer.cgi/hadoop/common/hadoop-3.3.6/hadoop-3.3.6.tar.gz>

- a. Put extracted Hadoop-3.3.6 files into C drive. Note that do not put these extracted files into Cdrive, where you installed your Windows.
- b. Download “*hadoop-common-3.3.6-bin-master*” from the link: <https://github.com/amihalik/hadoop-common-3.3.6-bin/tree/master/bin>. You will see 11 files there. Paste all these files into the “bin” folder of Hadoop-3.3.6.
- c. Create a “data” folder inside Hadoop-3.3.6, and also create two more folders in the “data” folder as “data” and “name.”
- d. Create a folder to store temporary data during execution of a

- project, such as “D:/hadoop/temp.”
- e. Create a log folder, such as “D:/hadoop/userlog”
  - f. Go to Hadoop-3.3.6 □ etc □ Hadoop and edit four files:
    - i. core-site.xml
    - ii. hdfs-site.xml
    - iii. mapred-site.xml
    - iv. yarn-site.xml

**core-site.xml:**

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<configuration>
<property>
  <name>fs.defaultFS</name>
  <value>hdfs://localhost:9000</value>
</property>
</configuration>
```

**hdfs-site.xml:**

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<configuration>
  <property>
    <name>dfs.replication</name>
    <value>1</value>
  </property>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>file:///C:/hadoop-3.3.6/data/namenode</value>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>file:///C:/hadoop-3.3.6/data/datanode</value>
  </property>
</configuration>
```

**mapred-site.xml:**

```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<configuration>
<property>
  <name>mapreduce.framework.name</name>
  <value>yarn</value>
</property>
</configuration>
```

**yarn-site.xml:**

```
<?xml version="1.0"?>
<configuration>
<property>
  <name>yarn.nodemanager.aux-services</name>
  <value>mapreduce_shuffle</value>
</property>
</configuration>
```

- g. Go to the location: "Hadoop-3.3.6\etc\hadoop,"  
and edit "hadoop-env.cmd" by  
writingset  
JAVA\_HOME=C:/java/jdk1.8.0\_2  
02
- h. Set environmental variables: Do: My computer > Properties >  
Advance system settings >  
Advanced > Environmental variables
  - i. User variables:
    - Variable: HADOOP\_HOME
    - Value: C:/hadoop-3.3.6
  - ii. System variable
    - Variable: Path
    - Value: C:/hadoop-3.3.6/bin  
C:/hadoop-3.3.6/sbin  
C:/hadoop-3.3.6/share/hadoop/common/\*  
C:/hadoop-3.3.6/share/hadoop/hdfs  
C:/hadoop-3.3.6/share/hadoop/hdfs/lib/\*  
C:/hadoop-3.3.6/share/hadoop/hdfs/\*  
C:/hadoop-3.3.6/share/hadoop/yarn/lib/\*  
C:/hadoop-3.3.6/share/hadoop/yarn/\*

C:/hadoop-3.3.6/share/hadoop/common/lib/\*

- j. **Format name-node:** On cmd go to the location “Hadoop-2.6.0\bin” by writing on cmd “cdhadoop-2.6.0\bin” and then “hdfs namenode –format”
- k. Start Hadoop. Go to the location: “D:\hadoop-2.6.0\sbin.” Run the following files as administrator “start-all.cmd”.
- l. Go to your web browser and search “localhost:9870” to access Hadoop NameNode. For Resource Manager, search “localhost:8088”.

[illegible]



