# INSTALL CONFIGURE AND RUN HADOOP AND HDFS

**Aim:**

To install configure and run hadoop and hdfs.

# Procedure:

## To install Java

* 1. Check if java is available in the system
  2. If not install java jdk 1.8 and set the environment variables



* 1. Set the path variable



## Hadoop Installation

* 1. Install Hadoop 3.3.6 from <https://hadoop.apache.org/releases.html>



Download the binary(checksum signature)

* 1. Extract the jar files to C://Hadoop



* 1. Add environment variables for Hadoop

Add path variable



* 1. Check if Hadoop is installed successfully using the command prompt



* 1. Thus Hadoop is installed successfully

## Hadoop Configuration

* 1. Configure core-site.xml in C:\hadoop\etc\hadoop by adding

*<configuration>*

*<property>*

*<name>fs.defaultFS</name>*

*<value>hdfs://localhost:9000</value>*

*</property>*

*</configuration>*

* 1. Configure the httpfs-site.xml file by adding the following xml code

*<configuration>*

*<property>*

*<name>dfs.replication</name>*

*<value>1</value>*

*</property>*

*<property>*

*<name>dfs.namenode.name.dir</name>*

*<value>C:\hadoop\data\namenode</value>*

*</property>*

*<property>*

*<name>dfs.datanode.data.dir</name>*

*<value>C:\hadoop\data\datanode</value>*

*</property>*

*</configuration>*

1. Configure mapred-site.xml file by adding the following xml code

*<configuration>*

*<property>*

*<name>mapreduce.framework.name</name>*

*<value>yarn</value>*

*</property>*

*</configuration>*

1. Configure yarn-site.xml file by adding the following xml code

*<configuration>*

*<property>*

*<name>yarn.nodemanager.aux-services</name>*

*<value>mapreduce\_shuffle</value>*

*</property>*

*<property>*

*<name>yarn.nodemanager.aux-services.mapreduce.shuffle.class</name>*

*<value>org.apache.hadoop.mapred.ShuffleHandler</value> </property>*

1. Change the bin shell command files.
2. Thus hadoop is configured.

## Hadoop execution

* 1. To check whether hadoop is running we must start the hadoop. To start hadoop we must use the command **start-all.cmd**





* 1. Check if hadoop runs in localhost.

To check this go to browser and type localhost:9870



# Result:

Thus hadoop has been installed, configured and run successfully.