# Hadoop Install Notes for Windows Server 2012

## Prerequisites

The latest stable release as of writing this document is 1.1.x is stable and 1.2.x is in beta and 2.x.x is in alpha release. For this trial, I’ve used 1.2.0

* [Download Hadoop](http://www.apache.org/dyn/closer.cgi/hadoop/common/) from one of the Apache Download Mirrors
* Required software for Hadoop installation is Java, ssh, sshd running, and Cygwin.

## Install and Configure Hadoop

1. In Cygwin, execute the following command to unpack Hadoop

|  |
| --- |
| tar -xzf hadoop-1.2.0-bin.tar.gz |

1. When the command prompt is seen again, execute the following command to check the directory

|  |
| --- |
| cd hadoop-1.2.0  ls -l |

1. To configure Hadoop, execute the following commands

|  |
| --- |
| cd conf  explorer . |

This opens the explorer window of **conf** directory.

1. Open the following files - **core-site.xml**, **hdfs-site.xml**, **mapred-site.xml**, and **hadoop-env.sh** to make additions to the respective files.

***core-site.xml***

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

***hdfs-site.xml***

|  |
| --- |
| <configuration>  <property>  <name>dfs.replication</name>  <value>1</value>  </property>  </configuration> |

***mapred-site.xml***

|  |
| --- |
| <configuration>  <property>  <name>mapred.job.tracker</name>  <value>localhost:9001</value>  </property>  </configuration> |

***Hadoop-env.xml***

|  |
| --- |
| Uncomment JAVA\_HOME and set it to C:/Program Files/Java/{java-home} |

1. It is time to format the namenode to create HDFS. Get to the root of Hadoop and make a directory for logs before formatting namenode

|  |
| --- |
| cd hadoop-1.2.0  mkdir logs  bin/hadoop namenode -format |

1. To start all of the Hadoop daemons that will simulate a distributed system, execute the following command

|  |
| --- |
| bin/start-all.sh |

Assuming no errors, check the NameNode and JobTracker in browser

* NameNode - <http://localhost:50070/>
* JobTracker - <http://localhost:50030/>

1. To copy the input files into HDFS, execute the following from Hadoop root directory

|  |
| --- |
| bin/hadoop dfs -mkdir input  bin/hadoop dfs -put conf input  bin/hadoop jar hadoop-\*-examples.jar grep input output 'dfs[a-z.]+'  bin/hadoop dfs -cat output/\* |

1. Assuming no errors, configure Eclipse for Hadoop

## Configure Eclipse for Hadoop

1. Navigate to Eclipse installation folder and open the plugins folder. Copy the **hadoop-1.2.0-eclipse-plugin.jar** to **eclipse\plugins** folder
2. Start Eclipse and click on open perspective icon located on the upper-right corner of IDE
3. Select **Map/Reduce** and click **OK**
4. Select the **Map/Reduce Locations** tab located at the bottom of IDE and right click on the blank space in that tab to select **New Hadoop location....**
5. Update the following and click **Finish**

* **Location Name** 🡪 localhost
* Map/Reduce Master
  + **Host** 🡪 localhost
  + **Port** 🡪 9101
* DFS Master
  + Check "Use M/R Master Host"
  + **Port** 🡪 9100
* **User name** 🡪 User

1. This gets **Map/Reduce Locations** tab at the bottom and in the **Project Explorer** on the left hand side of the IDE, find **DFS Locations**. Expand to find **localhost**. Now Eclipse is configured to run Map/Reduce project.