

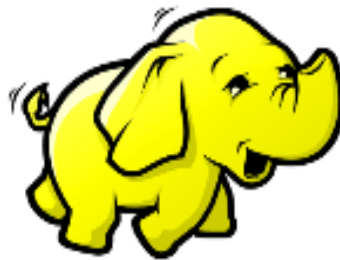


Wisdom Sprouts

We foster knowledge here

BIG DATA HADOOP

Lab Manual



Configuring Hadoop in Distributed Mode

Step 1: Stop all services from all the Machines

```
$> for service in /etc/init.d/hadoop*
> do
> sudo $service stop
> done
```

*Step 2: Create a copy of the configuration template files
from all the Machines*

```
$> sudo cp -r /etc/hadoop/conf.empty /etc/hadoop/conf.wisdom
```

*Step 3: Install the newly created configurations from all the
Machines*

```
$>sudo /usr/sbin/alternatives --install /etc/hadoop/conf hadoop-conf  
/etc/hadoop/conf.wisdom 99
```

*Step 4: Configure Hadoop to point to the newly installed
configuration from all the Machines*

```
$> /usr/sbin/update-alternatives --display hadoop-conf
```

Step 5: Check the ip_address of all your machines

```
$> /sbin/ifconfig
```

Step 6: Configure the hosts file from all the Machines

Using the Vi editor, open the /etc/hosts file. Delete all the existing content and add all the members of your group (including yourself) in the following way

ip_address_of_member host_name_of_member

For example

192.168.1.235 member1

192.168.1.236 member2

Step 7: Configure the Network file from all the Machines

Using the Vi editor, open the /etc/sysconfig/network file. Update the Hostname field with the name that you have decided for your machine. For example

Hostname = member1

Step 8: Configure the new set of configuration files from all the Machines (here host name of Name Node machine is considered as master)

Name	Value
Filename: /etc/hadoop/conf.wisdom/core-site.xml	
fs.default.name	hdfs://master:8020
Filename: /etc/hadoop/conf.wisdom/hdfs-site.xml	
dfs.name.dir	/home/disk1/dfs/name_space,/home/disk2/dfs/name_space
dfs.data.dir	/home/disk1/dfs/data_space,/home/disk2/dfs/data_space
dfs.http.address	master:50070
Filename: /etc/hadoop/conf.wisdom/mapred-site.xml	
mapred.local.dir	/home/disk1/mapred/local,/home/disk2/mapred/local
mapred.job.tracker	master:8021
mapred.jobtracker.staging.root.dir	/user

Step 9: Create the directories mentioned in the configuration files from all the Machines

```
$>sudo mkdir -p /home/disk1/dfs/name_space
$> sudo mkdir -p /home/disk2/dfs/name_space
$> sudo mkdir -p /home/disk1/dfs/data_space
$> sudo mkdir -p /home/disk2/dfs/data_space
$> sudo mkdir -p /home/disk1/mapred/local
$> sudo mkdir -p /home/disk2/mapred/local
```

Step 10: Configure permissions to the newly created directories from all the Machines

```
$> sudo chown -R hdfs:hadoop /home/disk1/dfs/name_space
$> sudo chown -R hdfs:hadoop /home/disk2/dfs/name_space
$> sudo chown -R hdfs:hadoop /home/disk1/dfs/data_space
$> sudo chown -R hdfs:hadoop /home/disk2/dfs/data_space
$> sudo chown -R mapred:hadoop /home/disk1/mapred/local
$> sudo chown -R mapred:hadoop /home/disk2/mapred/local
```

Step 11: Format the Name Node from the Master Machine

```
$> sudo -u hdfs hadoop namenode -format
```

Step 12: Start the HDFS master services from the Master Machine

```
$> sudo /etc/init.d/hadoop-hdfs-namenode start
$> sudo /etc/init.d/hadoop-hdfs-secondarynamenode start
```

Step 13: Start the HDFS slave service from the Slave Machines

```
$> sudo /etc/init.d/hadoop-hdfs-datanode start
```

Step 14: Create HDFS directory for the client from any 1 Machine

```
$> sudo -u hdfs hadoop fs -mkdir /user/wisdom
```

```
$> sudo -u hdfs hadoop fs -chown wisdom /user/wisdom
```

Step 15: Create MapReduce Specific directories from any 1 Machine

```
$> sudo -u hdfs hadoop fs -mkdir /tmp
```

```
$> sudo -u hdfs hadoop fs -mkdir /mapred/system
```

```
$> sudo -u hdfs hadoop fs -chmod -R 1777 /tmp
```

```
$> sudo -u hdfs hadoop fs -chmod -R 1777 /mapred/system
```

```
$> sudo -u hdfs hadoop fs -chown -R mapred /mapred/system
```

Step 16: Start the MapReduce Master service from the Master machine

```
$> sudo /etc/init.d/hadoop-0.20-mapreduce-jobtracker start
```

Step 17: Start the MapReduce Slave service from the Slave Machines

```
$> sudo /etc/init.d/hadoop-0.20-mapreduce-tasktracker start
```

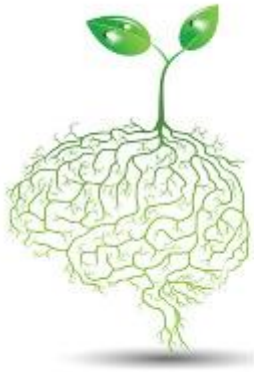
Step 18: Verify your HDFS Cluster

Go to url: master:50070

Step 19: Verify your MapReduce Cluster

Go to url: master:50030

www.wisdomsprouts.com



WISDOM SPROUTS

We Foster Knowledge Here

Contact Us

You can call us on +91-976-220-6123

Or write to us at helpdesk@wisdomsprouts.com

Or Visit www.wisdomsprouts.com