LINUX HADOOP KURULUMU

Geçen hafta Hadoop konusunun arge kısmını tamamlamış bulundum.Artık bu hafta Hadoop uygulamalarına geçiş yapıyoruz.Öncelikle Linux işletim sistemine Hadoop kurmam gerekiyordu.Benim için oldukça karmaşık birşeydi , çünkü daha önce Linux ile çalışma fırsatım çok az olmuştu.O yüzden tekrar hatırlamak için öncelikle Linux çalışmaya karar verdim.

Hadoop için kuruluma başlamadan önce ö*ncelikle bilgisayarımızda java kurulumunun yapıldığından emin olmalıyız . Aşağıdaki kod satırı terminal*e yazılarak bu bilgiyi öğrenebiliriz

|  |
| --- |
| java -version |

Benim kullandığım bilgisayarın java versiyonu eski çıktığından ötürü güncellemek zorunda kaldım.

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10 | root@laptop:~$ cd ~    root@laptop:~$ sudo apt-get update    root@laptop:~$ sudo apt-get install default-jdk    root@laptop:~$ java -version  java version "1.8.0\_141"  OpenJDK Runtime Environment (IcedTea 2.5.3) (8u144-2.5.3-0ubuntu0.14.04.1)  OpenJDK 64-Bit Server VM (build 24.65-b04, mixed mode) |
|  |  |

*openssh-server ve rsync paketlerini kuruyoruz.*

|  |  |
| --- | --- |
| 1 | root@laptop:~$ sudo apt-get install ssh |

**SSH Sertifikası oluşturma ve kurulum.**

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27 | root@laptop:~$ ssh-keygen -t rsa -P ""    Generating public/private rsa key pair.  Enter file in which to save the key (/home//.ssh/id\_rsa):  Created directory '/home//.ssh'.  Your identification has been saved in /home//.ssh/id\_rsa.  Your public key has been saved in /home//.ssh/id\_rsa.pub.  The key fingerprint is:  50:6b:f3:fc:0f:32:bf:30:79:c2:41:71:26:cc:7d:e3 root@laptop  The key's randomart image is:  +--[ RSA 2048]----+  |        .oo.o    |  |       . .o=. o  |  |      . + .  o . |  |       o =    E  |  |        S +      |  |         . +     |  |          O +    |  |           O o   |  |            o..  |  +-----------------+      root@laptop:/home/k$ cat $HOME/.ssh/id\_rsa.pub >> $HOME/.ssh/authorized\_keys |

Tüm bu ön hazırlıktan sonra sıra geldi Hadoop kurulumuna.Bunun için Hadoop’un kendi sitesine girip son versiyonunun linkini buluyoruz.Bulduğumuz linki indirdikten sonra Windowstaki zip dosyasına benzer olan dosyamızı tar.gz içinden çıkarmamız gerekiyor.Bunlar için gerekli komutlar :

|  |  |
| --- | --- |
| 1  2  3 | root@laptop:~$ wget http://mirrors.sonic.net/apache/hadoop/common/hadoop-2.7.1/hadoop-2.7.1.tar.gz    root@laptop:~$ tar xvzf hadoop-2.7.1.tar.gz |

Kısa bir beklemeden sonra Hadoop bilgisayara kuruldu.Sıra yapılandırma ayarlarına geldi.

**Configuration Dosyalarının Kurulumu**

**Aşağıdaki dosyalar kurulumu tamamlamak için modifiye edilecektir.**

1. ~/.bashrc
2. /usr/local/hadoop/etc/hadoop/hadoop-env.sh
3. /usr/local/hadoop/etc/hadoop/core-site.xml
4. /usr/local/hadoop/etc/hadoop/mapred-site.xml.template
5. /usr/local/hadoop/etc/hadoop/hdfs-site.xml

**1. ~/.bashrc**:

Home dizininde .bashrc dosyasını düzenlemeden önce, aşağıdaki komutu kullanarak JAVA\_HOME ortam değişkeni ayarlamak için javanın kurulu olduğu yeri bulamamız gerek.

|  |  |
| --- | --- |
| 1  2  3  4 | root@laptop update-alternatives --config java    There is only one alternative in link group java (providing /usr/bin/java): <span style="color: blue;">/usr/lib/jvm/java-7-openjdk-amd64/jre/bin/java</span> |

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18 | root@laptop:~$ vi ~/.bashrc    #HADOOP VARIABLES START  export JAVA\_HOME=/usr/lib/jvm/java-7-openjdk-amd64  export HADOOP\_INSTALL=/usr/local/hadoop  export PATH=$PATH:$HADOOP\_INSTALL/bin  export PATH=$PATH:$HADOOP\_INSTALL/sbin  export HADOOP\_MAPRED\_HOME=$HADOOP\_INSTALL  export HADOOP\_COMMON\_HOME=$HADOOP\_INSTALL  export HADOOP\_HDFS\_HOME=$HADOOP\_INSTALL  export YARN\_HOME=$HADOOP\_INSTALL  export HADOOP\_COMMON\_LIB\_NATIVE\_DIR=$HADOOP\_INSTALL/lib/native  export HADOOP\_OPTS="-Djava.library.path=$HADOOP\_INSTALL/lib"  #HADOOP VARIABLES END    root@laptop:~$ source ~/.bashrc |

**2. /usr/local/hadoop/etc/hadoop/hadoop-env.sh**

|  |  |
| --- | --- |
| 1  2  3 | root@laptop:~$ vi /usr/local/hadoop/etc/hadoop/hadoop-env.sh    export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64 |

|  |  |
| --- | --- |
| 1  2 | root@laptop:~$ sudo mkdir -p /app/hadoop/tmp  root@laptop:~$ sudo chown hduser:hadoop /app/hadoop/tmp |

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19 | hduser@laptop:~$ vi /usr/local/hadoop/etc/hadoop/core-site.xml    <configuration>  <property>    <name>hadoop.tmp.dir</name>    <value>/app/hadoop/tmp</value>    <description>A base for other temporary directories.</description>  </property>    <property>    <name>fs.default.name</name>    <value>hdfs://localhost:54310</value>    <description>The name of the default file system.  A URI whose    scheme and authority determine the FileSystem implementation.  The    uri's scheme determines the config property (fs.SCHEME.impl) naming    the FileSystem implementation class.  The uri's authority is used to    determine the host, port, etc. for a filesystem.</description>  </property>  </configuration> |

**4. /usr/local/hadoop/etc/hadoop/mapred-site.xml**

|  |  |
| --- | --- |
| 1 | root@laptop:~$ cp /usr/local/hadoop/etc/hadoop/mapred-site.xml.template /usr/local/hadoop/etc/hadoop/mapred-site.xml |

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10 | <configuration>  <property>    <name>mapred.job.tracker</name>    <value>localhost:54311</value>    <description>The host and port that the MapReduce job tracker runs    at.  If "local", then jobs are run in-process as a single map    and reduce task.    </description>  </property>  </configuration> |

**5. /usr/local/hadoop/etc/hadoop/hdfs-site.xml**

|  |  |
| --- | --- |
| 1  2  3 | root@laptop:~$ sudo mkdir -p /usr/local/hadoop\_store/hdfs/namenode  root@laptop:~$ sudo mkdir -p /usr/local/hadoop\_store/hdfs/datanode  root@laptop:~$ sudo chown -R hduser:hadoop /usr/local/hadoop\_store |

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19  20 | root@laptop:~$ vi /usr/local/hadoop/etc/hadoop/hdfs-site.xml    <configuration>  <property>    <name>dfs.replication</name>    <value>1</value>    <description>Default block replication.    The actual number of replications can be specified when the file is created.    The default is used if replication is not specified in create time.    </description>  </property>  <property>     <name>dfs.namenode.name.dir</name>     <value>file:/usr/local/hadoop\_store/hdfs/namenode</value>  </property>  <property>     <name>dfs.datanode.data.dir</name>     <value>file:/usr/local/hadoop\_store/hdfs/datanode</value>  </property>  </configuration> |

* Son olarak Hadoop namenode formatlanması kaldı.

|  |
| --- |
| bin/hadoop namenode -format |

Tüm bu işlemlerden sonra Hadoop yüklenmiş ve çalışır duruma gelmiştir. *Hadoop dizini altında jps komutunda aşağıdaki çıktıları görüyorsak kurulumun başarılı oldugunu söyleyebiliriz.*

Hadoop Başlatma :

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