

金融大数据-作业6

1、下载安装HBase

HBase安装

解压安装包hbase-1.1.2-bin.tar.gz至路径 /usr/local:

```
$ sudo tar -zxf ~/下载/hbase-1.1.2-bin.tar.gz -C /usr/local
```

将解压的文件名hbase-1.1.2改为hbase，以方便使用:

```
$ sudo mv /usr/local/hbase-1.1.2 /usr/local/hbase
```

配置环境变量

编辑~/.bashrc文件,文件尾行添加如下内容:

```
$ export PATH=$PATH:/usr/local/hbase/bin
```

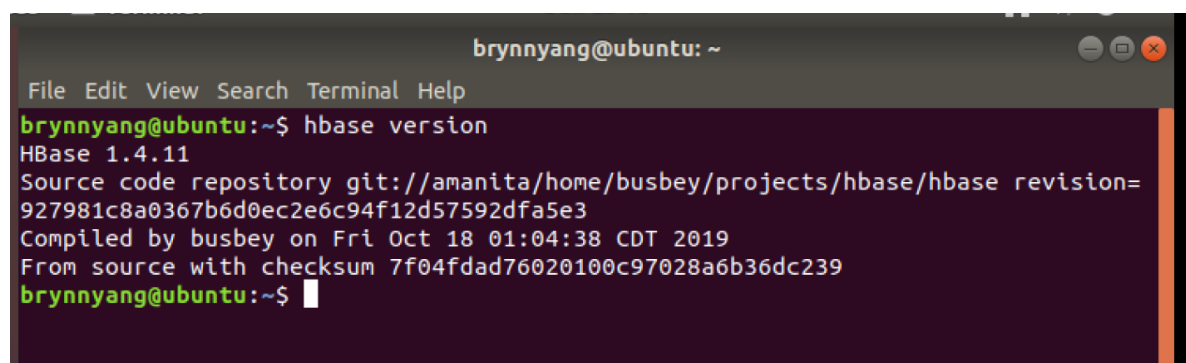
如果已经引入过PATH请在export PATH这行追加/usr/local/hbase/bin，这里的“:”是分隔符。

编辑完成后，再执行source命令使上述配置在当前终端立即生效。

添加HBase权限

```
$ cd /usr/local
$ sudo chown -R hadoop ./hbase      #将hbase下的所有文件的所有者改为hadoop，hadoop是
当前用户的用户名
```

查看HBase版本，确定hbase安装成功

A terminal window titled 'brynnnyang@ubuntu: ~' with a menu bar (File, Edit, View, Search, Terminal, Help). The command 'hbase version' is entered and executed. The output shows 'HBase 1.4.11' and details about the source code repository, compilation date (Fri Oct 18 01:04:38 CDT 2019), and checksum. The prompt returns to 'brynnnyang@ubuntu:~\$'.

HBase配置

1、单机模式

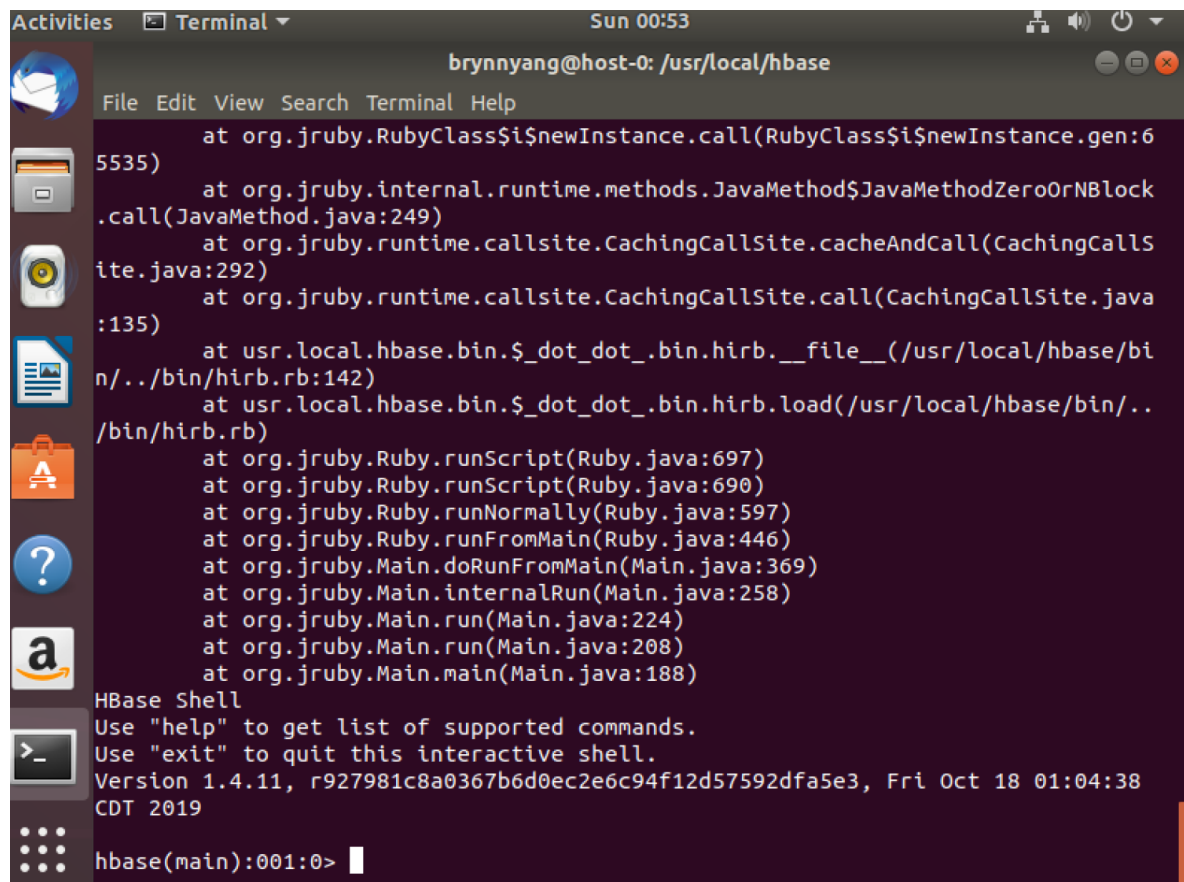
```
vi /usr/local/hbase/conf/hbase-env.sh
export JAVA_HOME=/usr/java/jdk1.8.0_221
export HBASE_MANAGES_ZK=true
```

配置/usr/local/hbase/conf/hbase-site.xml

```
<configuration>
  <property>
    <name>hbase.rootdir</name>
    <value>file:///usr/local/hbase/hbase-tmp</value>
  </property>
</configuration>
```

测试运行

```
cd /usr/local/hbase
bin/start-hbase.sh
bin/hbase shell
```



```
at org.jruby.RubyClass$Instance.newInstance.call(RubyClass$Instance.gen:6
5535)
  at org.jruby.internal.runtime.methods.JavaMethod$JavaMethodZeroOrNBlock
.call(JavaMethod.java:249)
  at org.jruby.runtime.callsite.CachingCallSite.cacheAndCall(CachingCallS
ite.java:292)
  at org.jruby.runtime.callsite.CachingCallSite.call(CachingCallSite.java
:135)
  at usr.local.hbase.bin.$_dot_dot_.bin.hirb.__file__(/usr/local/hbase/bi
n/./bin/hirb.rb:142)
  at usr.local.hbase.bin.$_dot_dot_.bin.hirb.load(/usr/local/hbase/bin/./
bin/hirb.rb)
  at org.jruby.Ruby.runScript(Ruby.java:697)
  at org.jruby.Ruby.runScript(Ruby.java:690)
  at org.jruby.Ruby.runNormally(Ruby.java:597)
  at org.jruby.Ruby.runFromMain(Ruby.java:446)
  at org.jruby.Main.doRunFromMain(Main.java:369)
  at org.jruby.Main.internalRun(Main.java:258)
  at org.jruby.Main.run(Main.java:224)
  at org.jruby.Main.run(Main.java:208)
  at org.jruby.Main.main(Main.java:188)
HBase Shell
Use "help" to get list of supported commands.
Use "exit" to quit this interactive shell.
Version 1.4.11, r927981c8a0367b6d0ec2e6c94f12d57592dfa5e3, Fri Oct 18 01:04:38
CDT 2019
hbase(main):001:0>
```

2、伪分布式

```
sudo vi /usr/local/hbase/conf/hbase-env.sh
export JAVA_HOME=/usr/java/jdk1.8.0_221
export HBASE_CLASSPATH=/usr/local/hadoop/conf
export HBASE_MANAGES_ZK=true
sudo vi /usr/local/hbase/conf/hbase-site.xml
<configuration>
  <property>
    <name>hbase.rootdir</name>
    <value>hdfs://localhost:9000/hbase</value>
  </property>
  <property>
    <name>hbase.cluster.distributed</name>
    <value>true</value>
  </property>
```

```
</configuration>
```

测试运行HBase

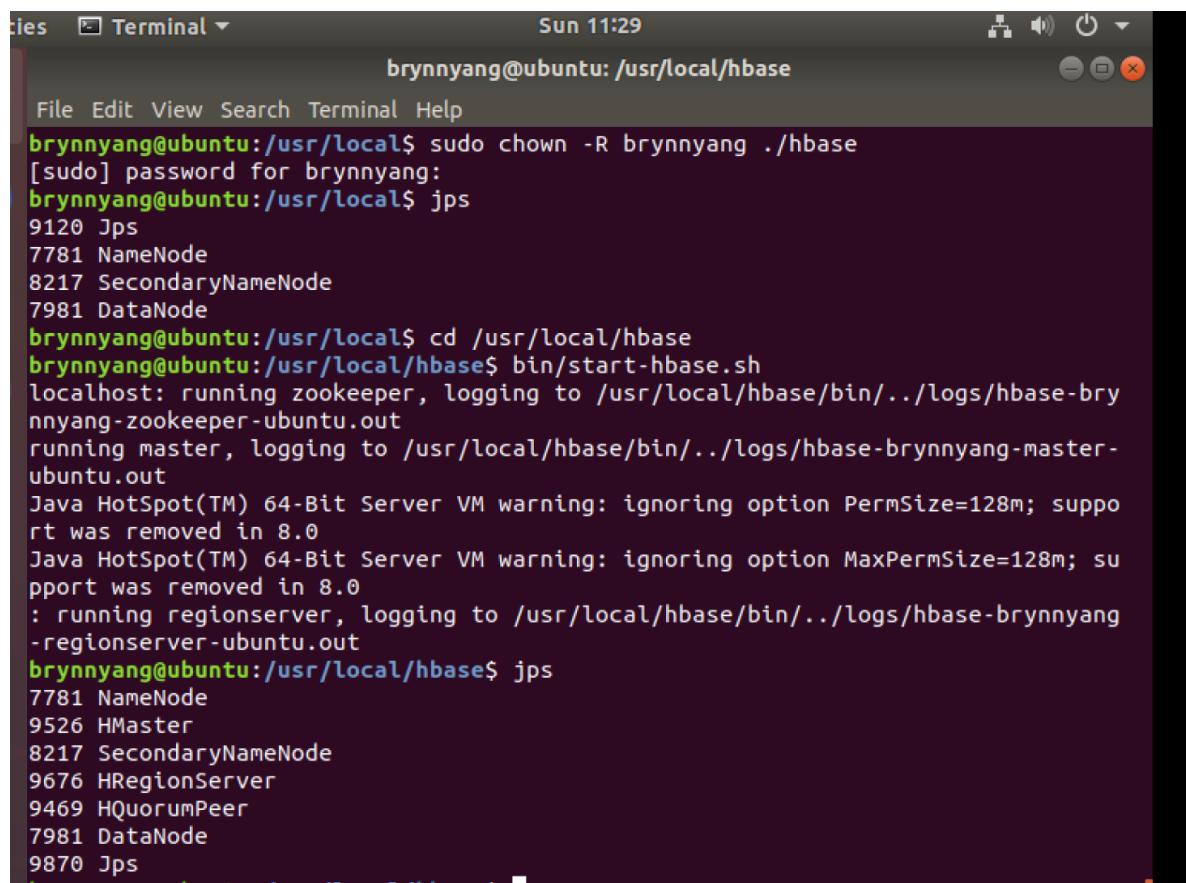
```
ssh localhost
cd /usr/local/hadoop
./sbin/start-dfs.sh

cd /usr/local/hbase
bin/start-hbase.sh

bin/hbase shell
```

如果在操作HBase的过程中发生错误，可以通过{HBASE_HOME}目录 (/usr/local/hbase) 下的 logs子目录中的日志文件查看错误原因。

这里启动关闭Hadoop和HBase的顺序一定是：**启动Hadoop—>启动HBase—>关闭HBase—>关闭Hadoop。**

A terminal window titled 'Terminal' with a dark background. The prompt is 'brynnnyang@ubuntu: /usr/local/hbase'. The user enters 'sudo chown -R brynnnyang ./hbase', followed by 'jps'. The output shows NameNode, SecondaryNameNode, and DataNode. Then the user enters 'cd /usr/local/hbase' and 'bin/start-hbase.sh'. The output shows the Zookeeper and HBase Master starting, with some Java warnings. Finally, the user enters 'jps' again, showing the full HBase cluster including HMaster, HRegionServer, HQuorumPeer, and DataNode.

```
File Edit View Search Terminal Help
brynnnyang@ubuntu: /usr/local/hbase
brynnnyang@ubuntu:/usr/local$ sudo chown -R brynnnyang ./hbase
[sudo] password for brynnnyang:
brynnnyang@ubuntu:/usr/local$ jps
9120 Jps
7781 NameNode
8217 SecondaryNameNode
7981 DataNode
brynnnyang@ubuntu:/usr/local$ cd /usr/local/hbase
brynnnyang@ubuntu:/usr/local/hbase$ bin/start-hbase.sh
localhost: running zookeeper, logging to /usr/local/hbase/bin/../logs/hbase-brynnnyang-zookeeper-ubuntu.out
running master, logging to /usr/local/hbase/bin/../logs/hbase-brynnnyang-master-ubuntu.out
Java HotSpot(TM) 64-Bit Server VM warning: ignoring option PermSize=128m; support was removed in 8.0
Java HotSpot(TM) 64-Bit Server VM warning: ignoring option MaxPermSize=128m; support was removed in 8.0
: running regionserver, logging to /usr/local/hbase/bin/../logs/hbase-brynnnyang-regionserver-ubuntu.out
brynnnyang@ubuntu:/usr/local/hbase$ jps
7781 NameNode
9526 HMaster
8217 SecondaryNameNode
9676 HRegionServer
9469 HQuorumPeer
7981 DataNode
9870 Jps
```

学习资料：Ubuntu下HBase安装与使用(shell+Java API)

<https://blog.csdn.net/flyfish111222/article/details/51893096>

2、熟悉shell命令

hbase shell命令	描述
alter	修改列族（column family）模式
count	统计表中行的数量
create	创建表
describe	显示表相关的详细信息
delete	删除指定对象的值（可以为表，行，列对应的值，另外也可以指定时间戳的值）
deleteall	删除指定行的所有元素值
disable	使表无效
drop	删除表
enable	使表有效
exists	测试表是否存在
exit	退出hbase shell
get	获取行或单元（cell）的值
incr	增加指定表，行或列的值
list	列出hbase中存在的所有表
put	向指向的表单元添加值
tools	列出hbase所支持的工具
scan	通过对表的扫描来获取对用的值
status	返回hbase集群的状态信息
shutdown	关闭hbase集群（与exit不同）
truncate	重新创建指定表
version	返回hbase版本信息

学习资料：HBase 常用Shell命令

<https://www.cnblogs.com/cxzdy/p/5583239.html>

3、创建讲义中的students表

先保证所有进程都已经开启

```
brynnYang@ubuntu:/usr/local/hbase/conf$ jps
3362 DataNode
6483 Launcher
5699 HRegionServer
5588 HMaster
5525 HQuorumPeer
3175 NameNode
2731 Main
3181 RemoteMavenServer36
6494 Jps
3583 SecondaryNameNode
brynnYang@ubuntu:/usr/local/hbase/conf$ cd /usr/local/hbase/bin
brynnYang@ubuntu:/usr/local/hbase/bin$ hbase shell
2019-11-25 10:50:41,787 WARN [main] util.NativeCodeLoader: Unable to load nati
```

pom.xml 文件（导入需要的依赖包）

```
<?xml version="1.0" encoding="UTF-8"?>
<project xmlns="http://maven.apache.org/POM/4.0.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://maven.apache.org/POM/4.0.0
http://maven.apache.org/xsd/maven-4.0.0.xsd">
  <modelVersion>4.0.0</modelVersion>

  <groupId>yf</groupId>
  <artifactId>students</artifactId>
  <version>1.0-SNAPSHOT</version>
  <dependencies>
    <dependency>
      <groupId>org.apache.hbase</groupId>
      <artifactId>hbase-common</artifactId>
      <version>1.4.11</version>
    </dependency>
    <dependency>
      <groupId>org.apache.hbase</groupId>
      <artifactId>hbase-client</artifactId>
      <version>1.4.11</version>
    </dependency>
    <dependency>
      <groupId>org.apache.hbase</groupId>
      <artifactId>hbase-server</artifactId>
      <version>1.4.11</version>
    </dependency>
    <dependency>
      <groupId>org.apache.hbase</groupId>
      <artifactId>hbase-protocol</artifactId>
      <version>1.4.11</version>
    </dependency>
  </dependencies>

</project>
```

hbase-site.xml文件

一定要与core-site中的主机地址一致！

```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
```

```

<!--
/**
 *
 * Licensed to the Apache Software Foundation (ASF) under one
 * or more contributor license agreements. See the NOTICE file
 * distributed with this work for additional information
 * regarding copyright ownership. The ASF licenses this file
 * to you under the Apache License, Version 2.0 (the
 * "License"); you may not use this file except in compliance
 * with the License. You may obtain a copy of the License at
 *
 *      http://www.apache.org/licenses/LICENSE-2.0
 *
 * Unless required by applicable law or agreed to in writing, software
 * distributed under the License is distributed on an "AS IS" BASIS,
 * WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
 * See the License for the specific language governing permissions and
 * limitations under the License.
 */
-->
<configuration>
  <!--      <property>
                <name>hbase.rootdir</name>
                <value>hdfs://master:9000</value>
            </property>-->
    <property>
      <name>hbase.rootdir</name>
      <value>hdfs://localhost:9000/hbase</value>
    </property>
    <property>
      <name>dfs.replication</name>
      <value>1</value>
    </property>
    <property>
      <name>hbase.cluster.distributed</name>
      <value>true</value>
    </property>
  </configuration>

```

随后编写代码运行即可，关键是hbase与java的连接问题。