Hive 安装手册

* 解压缩hive安装包

|  |
| --- |
| tar zxvf apache-hive-2.1.1-bin.tar.gz |

* 安装mysql (仅支持mysql 5.7以下版本，不支持5.7或更高版本，原因：若采用高版本mysql，部分文档内容需要改写)

|  |
| --- |
| sudo yum install mysql-server |

* 安装 mysql connector

|  |
| --- |
| sudo yum install mysql-connector-java |

该命令会在/usr/share/java/下产生mysql-connector-java.jar

* 建立链接

|  |
| --- |
| ln -s /usr/share/java/mysql-connector-java.jar /home/bigdata/software/apache-hive-2.1.1-bin/lib/mysql-connector-java.jar |

该命令在hive安装目录的lib目录下建立软链接，指向/usr/share/java/mysql-connector-java.jar

* 启动mysql

|  |
| --- |
| sudo service mysqld start |

可通过以下命令验证mysql启动成功(显示mysql进程)：

|  |
| --- |
| ps aux | grep mysql |

* 修改mysql密码

|  |
| --- |
| 修改mysql密码(改为newpass，请根据需要自行替换newpass)  mysql -u root  　　mysql> use mysql;  　　mysql> UPDATE user SET Password = PASSWORD('newpass') WHERE user = 'root';  　　mysql> FLUSH PRIVILEGES; |

允许用户root通过任意机器访问mysql：

|  |
| --- |
| mysql> GRANT ALL PRIVILEGES ON \*.\* TO 'root'@'%' IDENTIFIED BY 'newpass' WITH GRANT OPTION;  　　mysql> FLUSH PRIVILEGES; |

* 配置环境变量

/home/bigdata/.bash\_profile

|  |
| --- |
| export HIVE\_HOME=/home/bigdata/software/apache-hive-2.1.1-bin  export PATH=$HIVE\_HOME/bin:$PATH |

运行source生效：

|  |
| --- |
| source ~/.bash\_profile |

* 配置hive

/home/bigdata/software/apache-hive-2.1.1-bin/conf/**hive-site.xml**

|  |
| --- |
| <?xml version="1.0" encoding="UTF-8"?>  <?xml-stylesheet type="text/xsl" href="configuration.xsl"?>  <!--  Licensed 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. See accompanying LICENSE file.  -->  <!-- Put site-specific property overrides in this file. -->  <configuration>  <property>          <name>hive.metastore.uris</name>          <value>thrift://bigdata:9083</value>  </property>  <property>      <name>hive.server2.thrift.port</name>      <value>10000</value>  </property>  <property>  <name>javax.jdo.option.ConnectionURL</name>  <value>jdbc:mysql://bigdata/metastore?createDatabaseIfNotExist=true</value>  </property>  <property>  <name>javax.jdo.option.ConnectionDriverName</name>  <value>com.mysql.jdbc.Driver</value>  </property>  <property>  <name>javax.jdo.option.ConnectionUserName</name>  <value>root</value>  </property>  <property>  <name>javax.jdo.option.ConnectionPassword</name>  <value>root</value>  </property>  <property>  <name>hive.metastore.schema.verification</name>  <value>false</value>  </property>  <property>  <name>hive.metastore.warehouse.dir</name>  <value>/warehouse</value>  </property>  <property>  <name>fs.defaultFS</name>  <value>hdfs://bigdata:9000</value>  </property>  <property>  <name>datanucleus.autoCreateSchema</name>  <value>true</value>  </property>  <property>  <name>datanucleus.autoStartMechanism</name>  <value>SchemaTable</value>  </property>  <property>  <name>datanucleus.schema.autoCreateTables</name>  <value>true</value>  </property>  <property>  <name>beeline.hs2.connection.user</name>  <value>bigdata</value>  </property>  <property>  <name>beeline.hs2.connection.password</name>  <value>bigdata</value>  </property>  </configuration> |

注意 ：

* hive.metastore.uris中的“bigdata“含义为metastore server所在的机器（启动metastore的方法见下一节）
* javax.jdo.option.ConnectionURL中的“bigdata”为mysql安装机器的hostname
* javax.jdo.option.ConnectionUserName和javax.jdo.option.ConnectionPassword分别为mysql的访问用户和密码，可通过以下命令验证是否有效（期中bigdata为javax.jdo.option.ConnectionURL中配置的地址，xxx为mysql用户名）：
  + mysql –h bigdata-u xxx –p
* fs.defaultFS为HDFS的namenode启动的机器地址
* beeline.hs2.connection.user和beeline.hs2.connection.password是beeline方式访问的用户名和密码，可任意指定，但在beeline访问时要写入你指定的这个（具体参考最后一部分）
* 如果是第一次启动hive，需要执行初始化命令：

|  |
| --- |
| schematool -dbType mysql -initSchema |

注意：仅在第一次启动hive时，运行该命令，以后则只需直接启动metastore和hiveserver。

* 启动metastore

|  |
| --- |
| nohup hive --service metastore >> /home/bigdata/software/apache-hive-2.1.1-bin/metastore.log 2>&1 & |

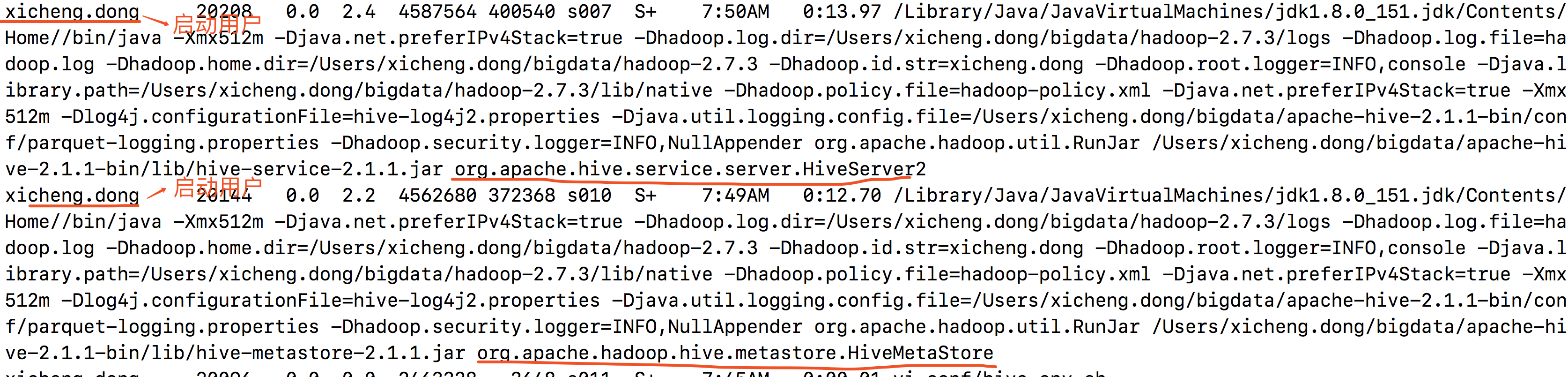
* 启动hive server

|  |
| --- |
| nohup hive --service hiveserver2 >> /home/bigdata/software/apache-hive-2.1.1-bin/hiveserver.log 2>&1 & |

* 查看hive metastore和hiveserver2是否启动成功

|  |
| --- |
| ps aux | grep hive |

能输出两个进程，分别对应metastore和hiveserver2。



* Hive常见两种访问方式

（1）**（不建议使用，已经被淘汰）**

|  |
| --- |
| hive |

（2）(**建议使用 ）**

|  |
| --- |
| beeline |

之后输入以下命令发起一个连接：

|  |
| --- |
| !connect jdbc:hive2://bigdata:10000/default bigdata bigdata |

其中bigdata和bigdata分别是在hive-site.xml配置文件中由beeline.hs2.connection.user和beeline.hs2.connection.password设置的。

注：如果要使用beeline或JDBC连接hive时，遇到报错：“User: **xxx** is not allowed to impersonate yyy”，需在hadoop的配置文件core-site.xml中加入以下配置（其中红色标志的“xxx”是你启动hive server2和hive metastore所采用的用户，**用户名中不要包含“.”，比如“cheng.dong”是不支持的**），并重启**hiveserver2， hive metastore，HDFS和YARN**：

<property>

      <name>hadoop.proxyuser.**xxx**.groups</name>

      <value>\*</value>

 </property>

 <property>

      <name>hadoop.proxyuser.**xxx**.hosts</name>

      <value>\*</value>

  </property>

比如，如果你使用linux用户bigdata启动的hive server2和hive metastore，需要增加以下两个配置：

<property>

      <name>hadoop.proxyuser.**bigdata**.groups</name>

      <value>\*</value>

 </property>

 <property>

      <name>hadoop.proxyuser.**bigdata**.hosts</name>

      <value>\*</value>

  </property>