

hive安装与配置

一、安装hive

将hive安装包放到共享文件夹，再复制到linux系统中，并赋予操作权限

```
sudo cp /mnt/hgfs/VMShare/apache-hive-2.3.4-bin.tar.gz /home/hadoop/VMShare/
```

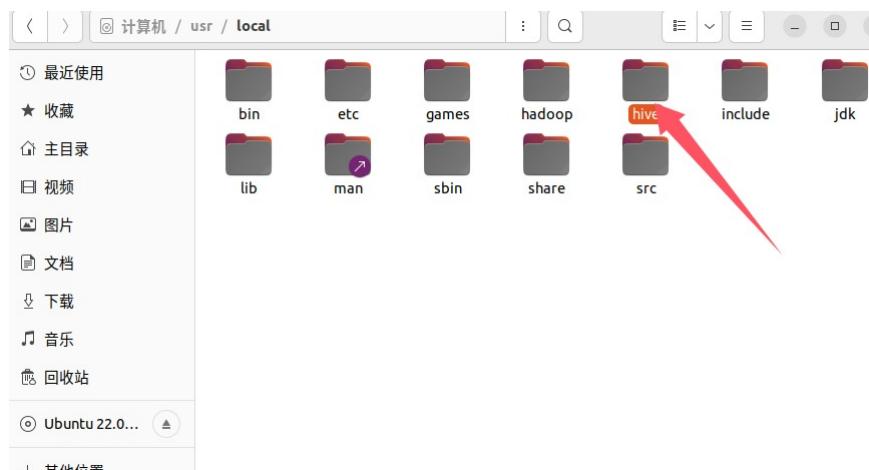
```
sudo chmod -R 777 /home/hadoop/VMShare
```

如果复制失败，重新挂载共享文件夹，再复制

```
sudo vmhgfs-fuse .host:/ /mnt/hgfs
```

进入到VMShare目录，下载和解压apache-hive-2.3.4-bin.tar.gz，并将解压后的文件移动到usr/local

```
cd /home/hadoop/VMShare/  
tar -zvxf apache-hive-2.3.4-bin.tar.gz  
sudo mv apache-hive-2.3.4-bin/ /usr/local/hive
```



二、配置环境变量

打开配置文件

```
vi ~/.bashrc
```

在末尾添加以下代码（路径与实际位置一致）

```
export HIVE_HOME=/usr/local/hive  
export PATH=$PATH:$HIVE_HOME/bin:$HIVE_HOME/sbin
```

```
if [ -f ~/bash_aliases ]; then
    . ~/bash_aliases
fi

# enable programmable completion features (you don't need to enable
# this, if it's already enabled in /etc/bash.bashrc and /etc/profile
# sources /etc/bash.bashrc).
if ! shopt -oq posix; then
    if [ -f /usr/share/bash-completion/bash_completion ]; then
        . /usr/share/bash-completion/bash_completion
    elif [ -f /etc/bash_completion ]; then
        . /etc/bash_completion
    fi
fi

export JAVA_HOME=/usr/local/jdk
export PATH=$PATH:$JAVA_HOME/bin

export HADOOP_HOME=/usr/local/hadoop
export PATH=$PATH:$HADOOP_HOME/bin:$HADOOP_HOME/sbin

export HIVE_HOME=/usr/local/hive
export PATH=$PATH:$HIVE_HOME/bin:$HIVE_HOME/sbin
-- 插入 --
```

保存退出后，输入命令使配置生效

```
source ~/.bashrc
```

三、修改hive的配置文件

在hive/conf目录下，首先将hive-env.sh.template修改为hive-env.sh，底部有三个路径，去掉注释，填上实际路径

```
42 # Larger heap size may be required when running queries over large number of files or partitions.
43 # By default hive shell scripts use a heap size of 256 (MB). Larger heap size would also be
44 # appropriate for hive server.
45
46
47 # Set HADOOP_HOME to point to a specific hadoop install directory
48 HADOOP_HOME=/usr/local/hadoop
49
50 # Hive Configuration Directory can be controlled by:
51 export HIVE_CONF_DIR=/usr/local/hive/conf
52
53 # Folder containing extra libraries required for hive compilation/execution can be controlled by:
54 export HIVE_AUX_JARS_PATH=/usr/local/hive/lib
```

在该目录下，创建一个空白文件hive-site.xml

```
vi hive-site.xml
```

打开文件，写入以下配置内容

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<configuration>
    <property>
        <name>javax.jdo.option.ConnectionURL</name>
        <value>jdbc:derby:;databaseName=/usr/local/hive/metastore_db;create=true</value>
    <description>创建 metastore_db 目录存放Derby数据库文件</description>
```

```

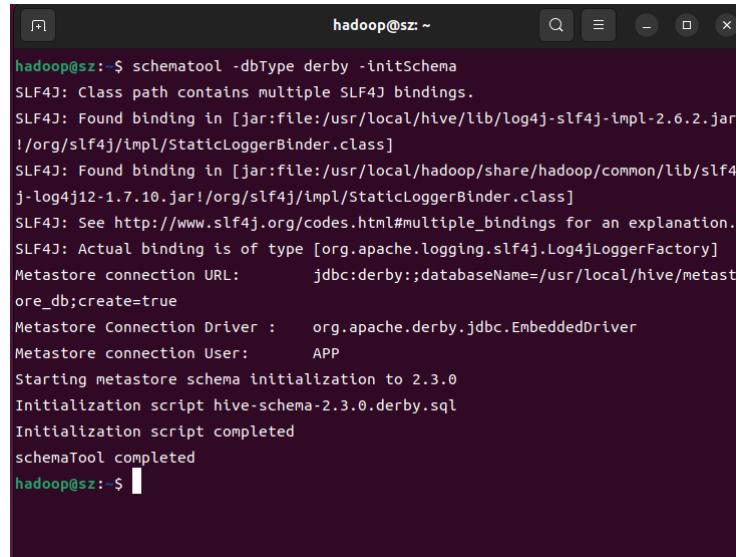
</property>
<property>
    <name>javax.jdo.option.ConnectionDriverName</name>
    <value>org.apache.derby.jdbc.EmbeddedDriver</value>
    <description>Driver class name for a JDBC metastore.</description>
</property>
<property>
    <name>javax.jdo.option.ConnectionUserName</name>
    <value>APP</value>
    <description>Username to use against metastore database.</description>
</property>
<property>
    <name>javax.jdo.option.ConnectionPassword</name>
    <value>123456</value>
    <description>password to use against metastore database.</description>
</property>
<property>
    <name>hive.metastore.warehouse.dir</name>
    <value>/user/hive/warehouse</value>
    <description>HDFS路径, Hive表数据默认存储于此。</description>
</property>
<property>
    <name>hive.exec.local.scratchdir</name>
    <value>/tmp/hive</value>
    <description>Hive本地临时目录。</description>
</property>
<property>
    <name>hive.querylog.location</name>
    <value>/tmp/hive/logs</value>
    <description>Hive查询日志目录。</description>
</property>
</configuration>

```

四、启动测试

首次使用，初始化元数据库

```
schematool -dbType derby -initSchema
```



```

hadoop@sz:~$ schematool -dbType derby -initSchema
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/:org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Metastore connection URL:      jdbc:derby:;databaseName=/usr/local/hive/metastore_db;create=true
Metastore Connection Driver :   org.apache.derby.jdbc.EmbeddedDriver
Metastore connection User:     APP
Starting metastore schema initialization to 2.3.0
Initialization script hive-schema-2.3.0.derby.sql
Initialization script completed
schemaTool completed
hadoop@sz:~$ 

```

启动hadoop集群

```
start-all.sh
```

```
hadoop@sz: ~$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
Starting namenodes on [hadoop01]
hadoop01: starting namenode, logging to /usr/local/hadoop/logs/hadoop-hadoop-namenode-sz.out
hadoop01: starting datanode, logging to /usr/local/hadoop/logs/hadoop-hadoop-datanode-sz.out
hadoop02: starting datanode, logging to /usr/local/hadoop/logs/hadoop-hadoop-datanode-sz.out
hadoop03: starting datanode, logging to /usr/local/hadoop/logs/hadoop-hadoop-datanode-sz.out
Starting secondary namenodes [hadoop02]
hadoop02: starting secondarynamenode, logging to /usr/local/hadoop/logs/hadoop-hadoop-secondarynamenode-sz.out
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-hadoop-resourcemanager-sz.out
hadoop02: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-hadoop-nodemanager-sz.out
hadoop03: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-hadoop-nodemanager-sz.out
hadoop01: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-hadoop-nodemanager-sz.out
hadoop@sz: ~$ jps
4000 NameNode
4824 Jps
4394 ResourceManager
4587 NodeManager
4173 DataNode
hadoop@sz: ~$
```

启动hive

```
hive
```

```
hadoop@sz: ~$ hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]

Logging initialized using configuration in jar:file:/usr/local/hive/lib/hive-common-2.3.4.jar!/hive-log4j2.properties Async: true
Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
hive> 
```

执行一条语句，没有报错，看到默认数据库则配置成功

```
show databases;
```

```
hive> show databases;
OK
default
Time taken: 2.735 seconds, Fetched: 1 row(s)
hive> 
```

可测试建表、导入数据等操作

实验总结

本实验在hadoop上搭建了hive环境。

