

micmiu — 软件开发+生活点滴

Hive教程之metastore的三种模式

作者: Michael 日期: 2014 年 2 月 20 日

查看评论 发表评论 (0)

Hive中metastore（元数据存储）的三种方式：

- 内嵌Derby方式
- Local方式
- Remote方式

【一】、内嵌Derby方式

这个是Hive默认的启动模式，一般用于单元测试，这种存储方式有一个缺点：在同一时间只能有一个进程连
接使用数据库。

hive-site.xml 中jdbc URL、驱动、用户名、密码等的配置信息如下：

```
1 <property>
2   <name>javax.jdo.option.ConnectionURL</name>
3   <value>jdbc:derby;;databaseName=metastore_db;create=true</value>
4   <description>JDBC connect string for a JDBC metastore</description>
5 </property>
6
7 <property>
8   <name>javax.jdo.option.ConnectionDriverName</name>
9   <value>org.apache.derby.jdbc.EmbeddedDriver</value>
10  <description>Driver class name for a JDBC metastore</description>
11 </property>
12
13 <property>
14   <name>javax.jdo.option.ConnectionUserName</name>
15   <value>APP</value>
16   <description>username to use against metastore database</description>
17 </property>
18
19 <property>
20   <name>javax.jdo.option.ConnectionPassword</name>
21   <value>mine</value>
22   <description>password to use against metastore database</description>
23 </property>
24
25 <property>
26   <name>hive.metastore.warehouse.dir</name>
27   <value>file:///Users/micmiu/tmp/hive/warehouse</value>
28   <description>unit test data goes in here on your local
29   filesystem</description>
30 </property>
31 <!-- micmiu.com -->
```

执行初始化命令：`schematool -dbType derby -initSchema`

查看初始化后的信息：`schematool -dbType derby -info`

配置完成后就可在shell中以CLI的方式访问hive 进行操作验证。

【二】、Local方式

以本地Mysql数据库为例：`创建好用户：hive；database：hive。`

微博

孙吴空的爸爸 上海 浦东新区

+ 加关注

哎。。。。//@南方都市报: 警方认为大学生将保安打成轻伤，行为存在“过度”的嫌疑，按相关法律要被追究刑事责任。

南都深圳读本 V：【“色狼”猥亵女网友被拘五天 大学生救人伤“狼”被刑拘】“色狼”保安猥亵女网友，路过的大学生救人，打伤该保安。保安因猥亵被行政拘留五天，还大开口索赔偿，而大学生被警方以涉嫌故意伤害刑事拘留，令被救的女网友想不通。警方称，保安轻伤，大学生行为

近期文章

Spring AOP拦截Struts1记录操作日志

Baby

Hadoop2.x在Ubuntu系统中编译源码

Ubuntu设置默认jdk

Ubuntu12.04中Eclipse安装svn插件问题

eclipse中开发Hadoop2.x的Map/Reduce项目

Hadoop2.x eclipse plugin插件编译安装配置

maven编译Spark源码

java.net.ConnectException: to 0.0.0.0:10020 failed on connection exception

ERROR tool.ImportTool: Imported Failed: Attempted to generate class with no columns!

浏览

图文介绍openLDAP在windows上的安装配置 - 12,282 views

Thrift入门及Java实例演示 - 11,745 views

SSO之CAS单点登录实例演示 - 11,720 views

java keytool证书工具使用小结 - 9,575 views

java自带的MD5、SHA1算法演示 - 9,268 views

VirtualBox安装Ubuntu图文介绍 - 8,801 views

Openfire服务端源代码开发配置指南 - 8,755 views

Eclipse+Maven构建web项目及部署时Maven lib依赖问题的解决 - 8,203 views

Mac OSX 10.8安装SVN的方法 - 7,061 views

配置文件 hive-site.xml 中jdbc URL、驱动、用户名、密码等属性值配置如下：

```
1 <property>
2   <name>javax.jdo.option.ConnectionURL</name>
3   <value>jdbc:mysql://localhost/hive?createDatabaseIfNotExist=true</value>
4   <description>JDBC connect string for a JDBC metastore</description>
5 </property>
6
7 <property>
8   <name>javax.jdo.option.ConnectionDriverName</name>
9   <value>com.mysql.jdbc.Driver</value>
10  <description>Driver class name for a JDBC metastore</description>
11 </property>
12
13 <property>
14   <name>javax.jdo.option.ConnectionUserName</name>
15   <value>hive</value>
16   <description>username to use against metastore database</description>
17 </property>
18
19 <property>
20   <name>javax.jdo.option.ConnectionPassword</name>
21   <value>micmiu</value>
22   <description>password to use against metastore database</description>
23 </property>
24
25 <property>
26   <name>hive.metastore.warehouse.dir</name>
27   <!-- base hdfs path -->
28   <value>/user/hive/warehouse</value>
29   <description>location of default database for the warehouse</description>
30 </property>
31 <!-- micmiu.com -->
```

ps：需要把mysql的驱动包copy到目录 <HIVE_HOME>/lib 中

如果是第一次需要执行初始化命令：`schematool -dbType mysql -initSchema`

```
1 micmiu-mbp:mysql micmiu$ schematool -dbType mysql -initSchema
2 14/02/17 14:46:08 INFO Configuration.deprecation: mapred.input.dir.recursive
3 is deprecated. Instead, use
4 mapreduce.input.fileinputformat.input.dir.recursive
5 14/02/17 14:46:08 INFO Configuration.deprecation: mapred.max.split.size is
6 deprecated. Instead, use mapreduce.input.fileinputformat.split.maxsize
7 14/02/17 14:46:08 INFO Configuration.deprecation: mapred.min.split.size is
8 deprecated. Instead, use mapreduce.input.fileinputformat.split.minsize
9 14/02/17 14:46:08 INFO Configuration.deprecation:
10 mapred.min.split.size.per.rack is deprecated. Instead, use
11 mapreduce.input.fileinputformat.split.minsize.per.rack
12 14/02/17 14:46:08 INFO Configuration.deprecation:
13 mapred.min.split.size.per.node is deprecated. Instead, use
14 mapreduce.input.fileinputformat.split.minsize.per.node
15 14/02/17 14:46:08 INFO Configuration.deprecation: mapred.reduce.tasks is
16 deprecated. Instead, use mapreduce.job.reduces
17 14/02/17 14:46:08 INFO Configuration.deprecation:
18 mapred.reduce.tasks.speculative.execution is deprecated. Instead, use
19 mapreduce.reduce.speculative
20 Metastore connection URL:      jdbc:mysql://localhost/hive?
21 createDatabaseIfNotExist=true
22 Metastore Connection Driver :   com.mysql.jdbc.Driver
23 Metastore connection User:     hive
24 Starting metastore schema initialization to 0.12.0
25 Initialization script hive-schema-0.12.0.mysql.sql
26 Initialization script completed
27 schemaTool completeted
```

查看初始化后信息 `schematool -dbType mysql -info`

初始化后查看mysql中表情况：`show tables;`

```
1 mysql> show tables;
2 +-----+
3 | Tables_in_hive |
4 +-----+
5 | BUCKETING_COLS |
6 | CDS              |
```

java.lang.OutOfMemoryError: GC overhead
limit exceeded - 6,519 views

分类目录

- J2EE (41)
 - Hibernate (12)
 - Struts (5)
 - Spring (11)
 - JDBC (6)
 - JTA (3)
 - EJB (4)
- 海量数据 (28)
 - Hadoop (12)
 - Hive (8)
 - HBase (4)
 - Spark (2)
 - Sqoop (2)
- SOA (14)
 - webservice (12)
 - RPC (2)
- OpenSource (27)
 - openflashchart (2)
 - 生成文档 (5)
 - Openfire (4)
 - Security (1)
 - CORBA (4)
- 架构设计 (7)
 - 集群 (6)
 - 缓存 (1)
- 企业应用 (19)
 - SNMP (8)
 - SSO (4)
 - 服务器软件 (4)
- 编程语言 (34)
 - Java (24)
 - javascript (5)
 - Groovy (4)
 - nodejs (1)
- NoSQL (12)
 - Berkeley (6)
- 数据库 (10)
 - Oracle (7)
 - mysql (2)
- Android (2)
- web UI (1)
- 软件工具 (32)
 - 常用软件 (1)
 - 项目构建 (15)
 - 测试工具 (2)
- 操作系统 (21)
 - linux (15)
 - windows (1)
 - Mac (5)
- 小小 (12)
 - 心底软绵绵 (4)
 - 时间静悄悄 (5)
 - 肚皮圆滚滚 (3)

7	COLUMNS_V2	
8	DATABASE_PARAMS	
9	DBS	
10	DB_PRIVS	
11	DELEGATION_TOKENS	
12	GLOBAL_PRIVS	
13	IDXS	
14	INDEX_PARAMS	
15	MASTER_KEYS	
16	NUCLEUS_TABLES	
17	PARTITIONS	
18	PARTITION_EVENTS	
19	PARTITION_KEYS	
20	PARTITION_KEY_VALS	
21	PARTITION_PARAMS	
22	PART_COL_PRIVS	
23	PART_COL_STATS	
24	PART_PRIVS	
25	ROLES	
26	ROLE_MAP	
27	SDS	
28	SD_PARAMS	
29	SEQUENCE_TABLE	
30	SERDES	
31	SERDE_PARAMS	
32	SKEWED_COL_NAMES	
33	SKEWED_COL_VALUE_LOC_MAP	
34	SKEWED_STRING_LIST	
35	SKEWED_STRING_LIST_VALUES	
36	SKEWED_VALUES	
37	SORT_COLS	
38	TABLE_PARAMS	
39	TAB_COL_STATS	
40	TBLS	
41	TBL_COL_PRIVS	
42	TBL_PRIVS	
43	TYPES	
44	TYPE_FIELDS	
45	VERSION	
46	+-----+	
47	41 rows in set (0.00 sec)	

配置完成后就可在shell中以CLI的方式访问hive 进行操作验证。

[三]、Remote方式

以Mysql数据库(192.168.6.77)为例：创建好用户：[hive](#)；database：[hive_meta](#)。Remote方式需要分别配置[服务端](#)和[客户端](#)的配置文件：

[服务端](#)的 hive-site.xml 中jdbc URL、驱动、用户名、密码等属性值配置如下：

1	<property>
2	<name>javax.jdo.option.ConnectionURL</name>
3	<value>jdbc:mysql://192.168.6.77/hive_meta?
4	createDatabaseIfNotExist=true</value>
5	<description>JDBC connect string for a JDBC metastore</description>
6	</property>
7	<property>
8	<name>javax.jdo.option.ConnectionDriverName</name>
9	<value>com.mysql.jdbc.Driver</value>
10	<description>Driver class name for a JDBC metastore</description>
11	</property>
12	
13	<property>
14	<name>javax.jdo.option.ConnectionUserName</name>
15	<value>hive</value>
16	<description>username to use against metastore database</description>
17	</property>
18	
19	<property>
20	<name>javax.jdo.option.ConnectionPassword</name>
21	<value>micmiu</value>
22	<description>password to use against metastore database</description>
23	</property>
24	
25	<property>
26	<name>hive.metastore.warehouse.dir</name>
27	<!-- base hdfs path -->

异常处理 (19)

杂谈 (2)

文章归档

- 2014年七月 (1)
- 2014年五月 (1)
- 2014年四月 (12)
- 2014年三月 (4)
- 2014年二月 (6)
- 2014年一月 (8)
- 2013年十二月 (1)
- 2013年十月 (9)
- 2013年九月 (7)
- 2013年八月 (8)
- 2013年六月 (4)
- 2013年五月 (4)
- 2013年四月 (3)
- 2013年三月 (10)
- 2013年二月 (1)
- 2013年一月 (2)
- 2012年十二月 (1)
- 2012年十一月 (7)
- 2012年十月 (2)
- 2012年九月 (5)
- 2012年八月 (13)
- 2012年七月 (14)
- 2012年六月 (12)
- 2012年五月 (16)
- 2012年四月 (20)
- 2012年三月 (22)
- 2012年二月 (88)

标签云 (3D)

技术链接

- 10→me@iteye
- 21→nosqlfan
- 22→mysqlops
- 23→OpenCloudDB

```
28 <value>/user/hive/warehouse</value>
29 <description>base hdfs path :location of default database for the
warehouse</description>
30 </property>
31 <!-- micmiu.com -->
```

ps: 需要把mysql的驱动包copy到目录 <HIVE_HOME>/lib 中

如果是第一次需要执行初始化命令: `schematool -dbType mysql -initSchema`

客户端中配置内容修改如下:

```
1 <!-- thrift://<host_name>:<port> 默认端口是9083 -->
2 <property>
3 <name>hive.metastore.uris</name>
4 <value>thrift://192.168.6.77:9083</value>
5 <description>Thrift uri for the remote metastore. Used by metastore client
to connect to remote metastore.</description>
6 </property>
7
8 <!-- hive表的默认存储路径 -->
9 <property>
10 <name>hive.metastore.warehouse.dir</name>
11 <value>/user/hive/warehouse</value>
12 <description>location of default database for the warehouse</description>
13 </property>
```

hive metastore 服务端启动命令:

```
hive --service metastore -p <port num>
```

如果不加端口默认启动: `hive --service metastore`, 则默认监听端口是: 9083, 注意客户端中的端口配置需要和启动监听的端口一致。服务端启动正常后, 客户端就可以执行hive操作了。

参考:

<https://wiki.apache.org/confluence/display/Hive/AdminManual+MetastoreAdmin>


----- EOF @micmiu.com -----

原创文章, 转载请注明: 转载自micmiu – 软件开发 + 生活点滴[<http://www.micmiu.com/>]

本文链接地址: <http://www.micmiu.com/bigdata/hive/hive-metastore-config/>

 Hive  Hive

[← Hive教程之DML数据导入导出](#) [HBase基于Hadoop2的源码编译 →](#)

 发表评论? 0 条评论。


发表评论

昵称

邮箱

网址



-  31→CrazyJvm
-  32→封神无度
-  33→董的博客

功能

-  登录
-  文章RSS
-  评论RSS
-  WordPress.org

注意 - 你可以用以下 HTML tags and attributes:
 <abbr title=""> <acronym title=""> <blockquote
cite=""> <cite> <code> <del datetime=""> <i> <q cite=""> <strike>
