

(超详细, 带你踩坑) linux centos7 基于CDH6.0.1配置hive on tez 执行引擎(亲测有效)

参考:

<https://cloud.tencent.com/developer/article/1349482>

<https://blog.csdn.net/cmzvd2222/article/details/102461358>

<https://blog.csdn.net/Shea1992/article/details/101041244>

下载已编译好的tez部署到HDFS

```
hdfs dfs -mkdir /user/tez
```

将tez-0.9.1.tar.gz上传至hdfs

```
hdfs dfs -put tez-0.9.1.tar.gz /user/tez/
```

文件权限 (可不修改)

```
hdfs dfs -chmod -R 777 /user/tez/
```

. 创建tez-site.xml, 配置客户端

在CDH的lib目录下创建tez:

一般CDH6的安装目录都在该地方

```
cd /opt/cloudera/parcels/CDH-6.0.1-1.cdh6.0.1.p0.590678/lib/
```

```
mkdir tez
```

```
cd tez
```

将tez-0.9.1-minimal文件夹下的jar及lib下的jar拷贝到tez中

从 /opt/cloudera/parcels/CDH-6.0.1-1.cdh6.0.1.p0.590678/jars中拷贝kryo-2.22.jar到tez文件夹下的lib文件夹下, 防止出现以下异常:

```
java.lang.ClassNotFoundException: com.esotericsoftware.kryo.Serializer
```

tez/lib中包含slf4j的jar包, 会打印较多日志, 可以在客户端中去掉slf4j-api-1.7.10.jar、slf4j-log4j12-1.7.10.jar这两个jar包, 减少日志打印,

--测试下来 这个必须操作, 不然报错。

```
cd /opt/cloudera/parcels/CDH/lib/hive/auxlib
```

```
mv hive-exec-2.1.1-cdh6.0.1-core.jar hive-exec-2.1.1-cdh6.0.1-core.jar.bak
```

```
mv hive-exec-core.jar hive-exec-core.jar.bak
```

到这里就客户端的基本配置结束。

将 tez这个文件就 copy的集群中的其他主机的cloudera目录下

```
scp /opt/cloudera/parcels/CDH-6.0.1-1.cdh6.0.1.p0.590678/lib/tez root@drgs2:/opt/cloudera/parcels/CDH-6.0.1-1.cdh6.0.1.p0.590678/lib/
```

配置hive环境变量

hive-site.xml 的 Hive 客户端高级配置代码段 (安全阀)

Gateway Default Group

名称	tez.lib.uris
值	/engine/tez-0.9.1-minimal.tar.gz
说明	/engine/tez-0.9.1.tar.gz <input type="checkbox"/> Final
名称	tez.use.cluster.hadoop-libs
值	true
说明	使用hadoop自身的lib包,设置为true的话可以使用minimal的tez包,false的话需要使用te <input type="checkbox"/> Final
名称	tez.runtime.compress
值	true
说明	说明 <input type="checkbox"/> Final
名称	tez.runtime.compress.codec
值	org.apache.hadoop.io.compress.SnappyCodec
说明	说明 <input type="checkbox"/> Final

[+](#)

```
<configuration>
  <property>
    <name>tez.lib.uris</name>
    <value>hdfs://ns1/engine/tez-0.9.1-minimal.tar.gz</value>
  </property>
  <property>
    <name>tez.use.cluster.hadoop-libs</name>
    <value>true</value>
  </property>
  <property>
    <name>tez.runtime.compress</name>
    <value>true</value>
  </property>
  <property>
    <name>tez.runtime.compress.codec</name>
    <value>org.apache.hadoop.io.compress.SnappyCodec</value>
  </property>
</configuration>
```

</configuration>

HiveServer2 环境高级配置代码段 (安全阀)

HiveServer2 Default Group 

```
HADOOP_CLASSPATH=/opt/cloudera/parcels/CDH/lib/tez/conf:/opt/cloudera/parcels/CDH/lib/tez*/opt/clo
```

```
HADOOP_CLASSPATH=/opt/cloudera/parcels/CDH/lib/tez/conf:/opt/cloudera/parcels/CDH/lib/tez*/opt/cloudera/parcels/CDH/lib/
```

测试

--每次执行必须带队列执行

```
set hive.execution.engine=tez;  
set hive.tez.container.size=2048;
```

```
set tez.queue.name=bf_yarn_pool.development ;  
select count(salary),sum(salary) from sample_07;
```