Flink入门-3(Table API & SQL)

先说，这个玩意其实我的内心是排斥的。但是还是要学呀。首先Table Api是一个独立的插件存在，如果使用它，需要单独导入。方法也很简单。在maven中添加如下代码：

<dependency>

<groupId>org.apache.flink</groupId>

<artifactId>flink-table\_2.11</artifactId>

<version>1.7.0</version>

</dependency>

<dependency>

<groupId>org.apache.flink</groupId>

<artifactId>flink-scala\_2.11</artifactId>

<version>1.7.0</version>

</dependency>

<dependency>

<groupId>org.apache.flink</groupId>

<artifactId>flink-streaming-scala\_2.11</artifactId>

<version>1.7.0</version>

</dependency>

然后我们继续修改之前的例子。添加Table API的使用。

//1.创建Table环境

//2.注册表

//3.逻辑处理

创建Table环境，因为我是批量任务所示用BatchTableEnvironment这个。

BatchTableEnvironment bTableEnv = TableEnvironment.getTableEnvironment(env);

注册表的方法很多，最简单的应该是registerDataSet，也就是将一个DataSet声明成一张表。

bTableEnv.registerDataSet("test", dataset,"name,value");

这个方法有3个参数，分别为声明的表名，对应的dataset和表的字段说明。

有了这个以后就可以写sql了。执行sql的方法可以用sqlQuery。举个例子：

Table sqlResult = bTableEnv.sqlQuery("SELECT name,count(1) FROM test group by name");

这方法很简单，就是写sql就可以，并且结果还是一张表。

至此就可以用Table Api了。但是还是很不建议用这个的。原本开发一个flink的程序也就几十K。但是用了Table。一下就70多MB了。图什么呀。而且写sql我觉得是挺不好的方式。程序员还是少用吧。

