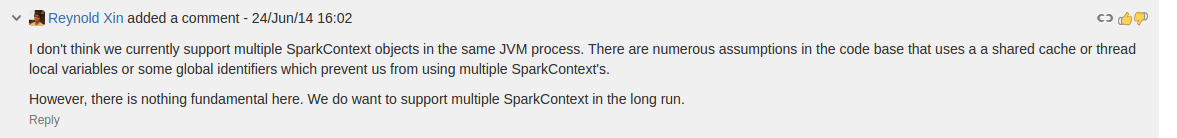
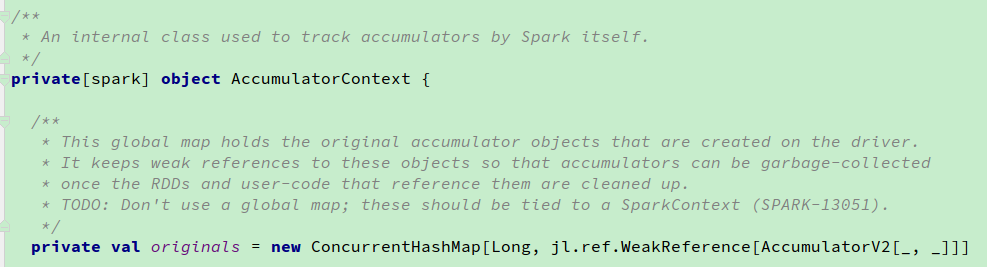
**spark core-SparkContext**

1. **Basic**

**1.SparkContext:目前spark每个JVM只能有一个SparkContext**



**2.与此对应的AccumulatorContext 也是每个JVM一个**



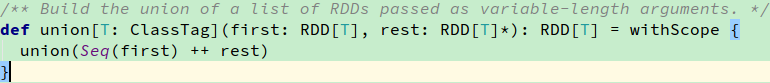
1. **下面重点研究submitJob**

**首先是：RDD的runJob,比如**

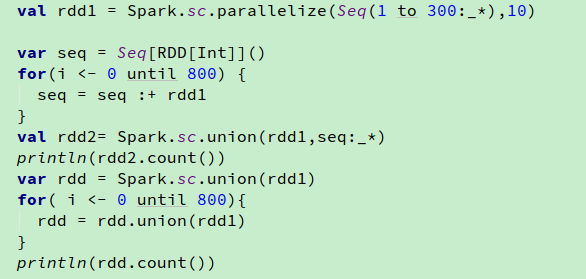
**二．Operation**

1. **Union**

Union就是把多个RDD，整合成一个RDD，注意在面的这种代码风格。查看代码发现sparkContext和RDD中都有union，但是归根结底RDD中的union调用了SparkContext中的union。



唯一一个不同的地方是，当非常多的RDD进行union时，使用RDD中的union可能存在如下问题，但是使用SparkContext中的union却没有问题。测试代码如下：



错误如下：

