说明书

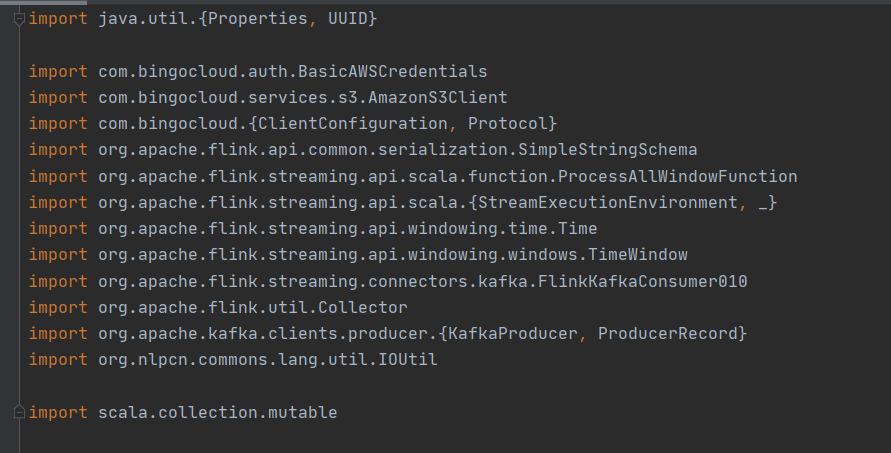
文件目录:

有两个scala文件:S3writer.scala 和Demo.scala

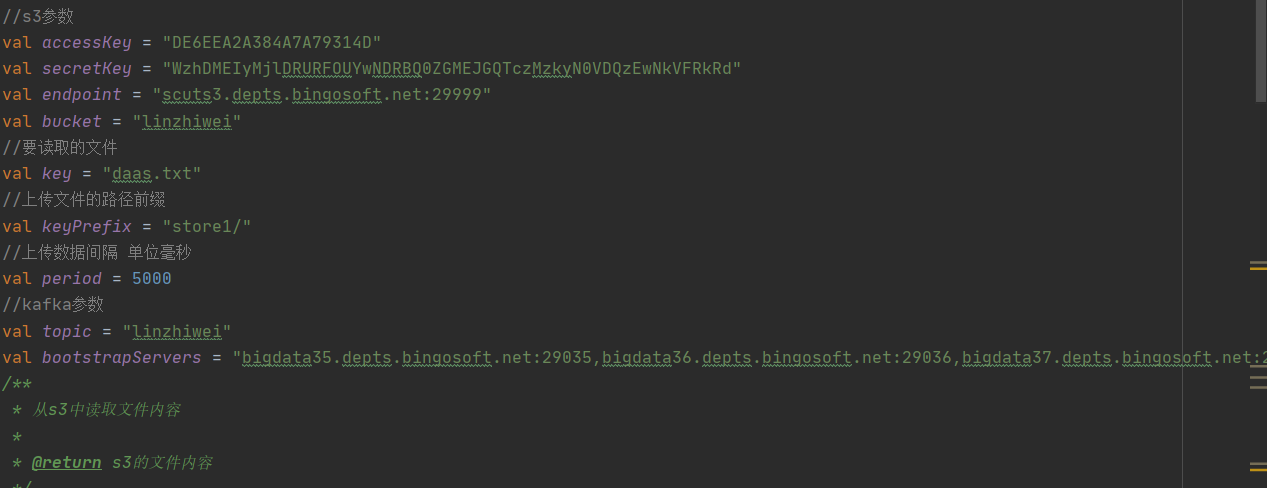
S3writer.scala是将按destination归类后的数据存放到S3存储桶中

下面着重解释Demo.scala：

导包:

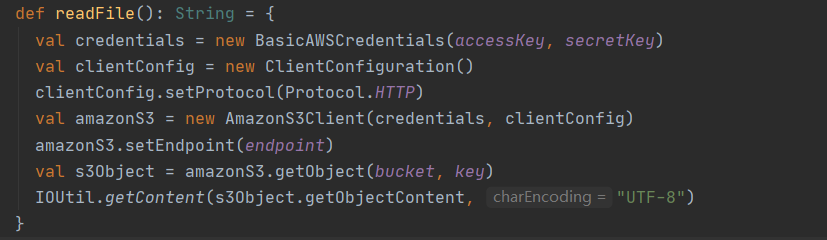


配置参数:

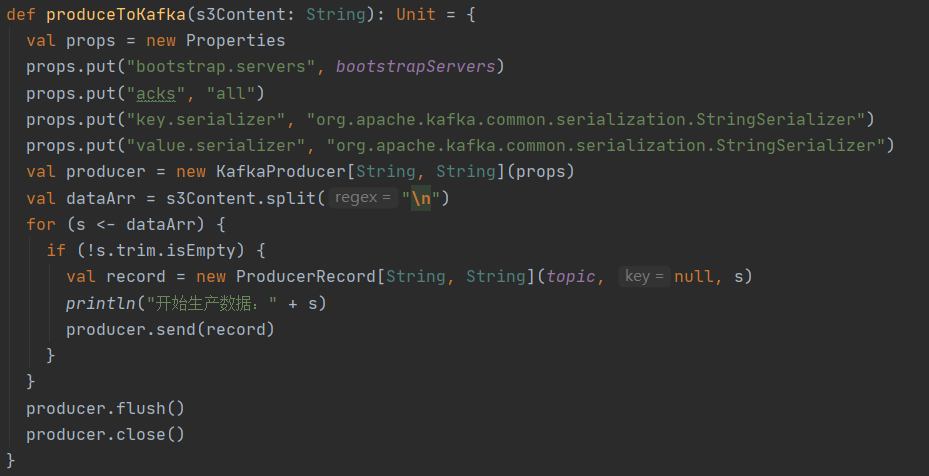


以下是逻辑实现部分:

1. readFile()实现读取S3的文件内容，返回值为字符串



1. produceTokafka() 将读取的内容上传到kfaka



(3)重点是以下这个函数，如何将从kafka读取的数据按destination进行归类





ProcessAllWindows()代码解释:

首先定义两个数组city和touristNum和一个二维数组result，分别存储城市名，游客数，和要返回的元素.

然后对每个窗口的每个元素kv进行遍历，第一次遍历：

如果city数组不存在kv的destination，将其存入，并将cityNum+1

接着进行第二遍遍历:  
对于每个元素kv：根据kv的目的地找到city数组的下标i，根据下标i找到touristNum数组的下标，并将结果存在result二维数组

经过两次遍历,result数组的每一个元素都存放着一条记录，每一行都是同个destination的记录

接着将resul的每个元素collect

最后new S3Writer将记录写入S3存储桶中。

结果截图:

