**Lab 5-分布式交易结算系统**

**Handout: May 31, 2018**

**Deadline: July 1, 2018 (No extension)**

**Assignment Overview:**

本实验要求同学们采用多种最新的分布式计算支撑技术，搭建和实现一个综合性的系统。必须要用到的技术包括Spark Streaming、Kafka、MongoDB、Zookeeper，当然你也可以在此基础上应用更多的技术使得系统的某些特性得到提升，那将会获得Bonus。

**应用场景描述：**

XXX市场会实时产生大量的外贸交易订单，有4个国家（中国、美国、日本、德国）在该市场里相互交易，它们使用的货币分别为——人民币（RMB）、美元（USD）、日元（JPY）、欧元（EUR）。该市场采用联盟币（UNI）作为结算货币，每个国家的货币对联盟币有兑换汇率，该汇率以分钟为单位动态变化。现需要同学们设计实现一个分布式交易结算系统，实时接收并处理交易订单，并计算出每分钟各个国家的交易情况。

**规定与示例：**

（1）系统初识时刻为2018-01-01 0:00，当时的汇率表如下：

|  |  |
| --- | --- |
| 源货币 | 兑换率 |
| RMB | 2.0 |
| USD | 12.0 |
| JPY | 0.5 |
| EUR | 6.0 |

意思是：在这一分钟里产生的交易，1 RMB=2.0 UNI；1 USD=12.0 UNI；1 JPY=0.5 UNI；1 EUR=6.0 UNI。

要求：**汇率表用Zookeeper进行管理**。自行编程4个并行计算的实体（多线程或多进程）修改4国的汇率，修改的规则是每分钟汇率值加0.1。

（2）某一分钟的交易情况计算规则如下：

假设2018-01-01 0:00开始的一分钟内（上述汇率表有效）产生了3个订单：

订单1：

发起国：RMB

接受国：USD

交易额：100（RMB） //交易额以发起国的货币单位计量

交易时间：2018-01-01 0:00:05

订单2：

发起国：USD

接受国：JPY

交易额：200（USD）

交易时间：2018-01-01 0:00:28

订单3：

发起国：RMB

接受国：JPY

交易额：300（RMB）

交易时间：2018-01-01 0:00:53

故2018-01-01 0:00开始的这一分钟的交易情况结果如下：

中国：收入0 RMB，支出400 RMB //交易情况以本国货币为计量单位

美国：收入16.67 USD，支出200 USD //转换金额精确到小数点后两位

日本：收入6000 JPY，支出0 JPY //(200\*12+300\*2)/0.5 =6000

德国：收入0 EUR，支出为0 EUR

（3）交易订单输入与结果输出

订单数据通过HTTP Get请求发出，为JSON形式，数据格式定义如下：

订单数据格式：

{

src\_name: “CNY”,

dst\_name: “USD”,

value: 100,

time: “2018-01-24 12:59:59”

}

订单数据将持续发送，时间将从2018-01-01 0:00:00开始不断递增，可能存在时间完全相等的订单。

结果输出的数据格式为：

{

name: “CNY”,

income: 16.67

expand: 200

time: “2018-01-24 12:59”

}

1. 其他要求

* 订单数据产生端和测试数据已经给出，需要自行编写订单数据接收端
* 接收到的数据必须使用**kafka**进行管理
* 数据计算必须使用**Spark Streaming**
* 结果输出必须保存到**MongoDB**中

**实验报告：**

实验报告分三个部分。第一，一份文档用来描述系统环境搭建与配置过程，程序设计思路，以及过程中遇到的问题和解决方案等，第二，程序设计源代码，第三，测试数据的运行结果。

Upload your lab report as a gzipped tar file and name it as {Your student ID}.tar.gz to

ftp://solostagev9051:public@public.sjtu.edu.cn/upload/lab5

**[Grading]:**

40%: the quality of lab report.

60%: laboratory examinations, scheduled to early July.