**西南石油大学实验报告**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **课 程** | 大数据处理技术 | | **实验项目** | | | PageRank设计 | | | | | **成绩** | | |  |
| **专业班级** | 网络工程2016 | | **组别** | | 无 | | **学号** | 201631063211 | | **指导教师** | | 赖俊良 | | |
| **姓 名** | 古皓 | **同组人姓名** | | 无 | | | | | **实验日期** | | | | 2019/6/24 | |

1. **实验目的**

掌握Map/Reduce相关原理和设计方法，设计PageRank应用。

1. **实验内容**

数据集page-link.txt，包含从A~Z一共26个网页的链出记录，根据此链出记录应用PageRank算法，最终输出每个网页的PageRank权值。

1. **实验步骤**

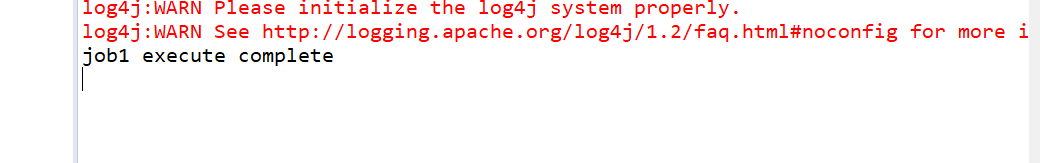
Eclipse导入工程项目，代码进行补充修改调试，

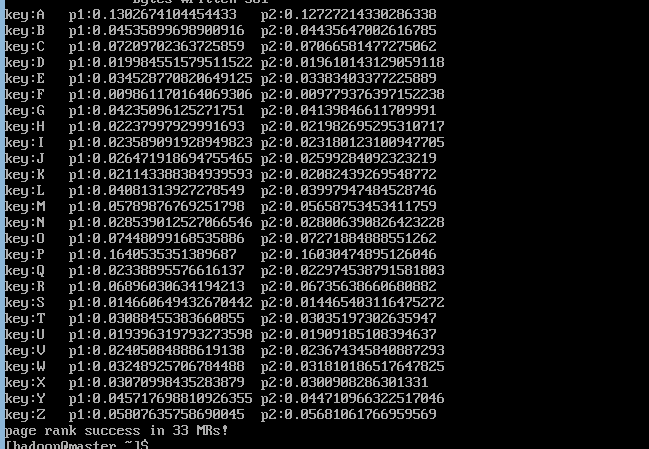
补充代码：AcountOutMapper和AcountOutReducer，求得转移矩阵

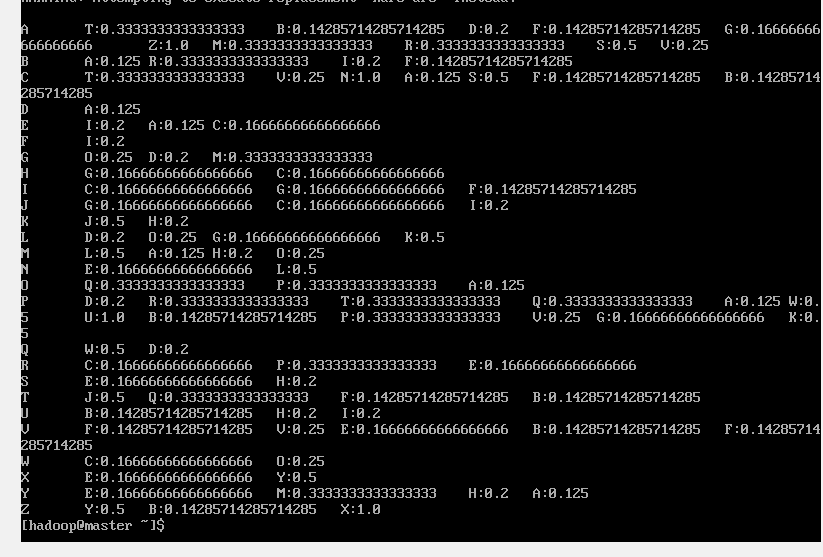
打包成可执行jar包，上传centos虚拟机进行运行调试

**四、实验结果**

试验运行过程及结果







**五、实验体会**

这次实验代码运行jar包遇到一点问题，就是进行迭代时候迭代了一次就抛出了异常，查阅资料，改了很多地方也未曾解决，最后尝试从新操作了几次，最后在同学的帮助下解决了问题，可能是粗心哪里没写对，可见细心很重要啊