**spark 第一天作业-李昱**

1. 讲集群搭建起来，运行，并在web界面观察



1. 自行选择5个org\apache\spark\examples包下面的demo 在集群里面跑，并观察效果

1.计算pi值



2.逻辑回归算法



3.kmean聚类算法



4.ALS交替最小二乘法，推荐系统的算法。



5.14. MultiBroadcastTest 测试多个Broadcast共享变量。



3.Spark安装目录下面的文件夹一个一个cd 进去观察，理解每一个文件夹的作用



drwxr-xr-x. 2 hdp-node-02 hdp-node-02 4096 Jun 30 2017 bin 可执行文件

drwxr-xr-x. 2 hdp-node-02 hdp-node-02 4096 Jul 3 04:55 conf 配置文件

drwxr-xr-x. 5 hdp-node-02 hdp-node-02 4096 Jun 30 2017 data 例子中用到的数据

drwxr-xr-x. 4 hdp-node-02 hdp-node-02 4096 Jun 30 2017 examples 自带的例子源码

drwxr-xr-x. 2 hdp-node-02 hdp-node-02 12288 Jun 30 2017 jars jar包

-rw-r--r--. 1 hdp-node-02 hdp-node-02 17881 Jun 30 2017 LICENSE 一些关于版权的声明

drwxr-xr-x. 2 hdp-node-02 hdp-node-02 4096 Jun 30 2017 licenses 关于软件使用的许可声明

drwxr-xr-x. 2 root root 4096 Jul 4 02:17 logs 日志文件

-rw-r--r--. 1 hdp-node-02 hdp-node-02 24645 Jun 30 2017 NOTICE

drwxr-xr-x. 6 hdp-node-02 hdp-node-02 4096 Jun 30 2017 python 对py支持的包

drwxr-xr-x. 3 hdp-node-02 hdp-node-02 4096 Jun 30 2017 R 对r语言支持的包

-rw-r--r--. 1 hdp-node-02 hdp-node-02 3809 Jun 30 2017 README.md 入门的spark说明

-rw-r--r--. 1 hdp-node-02 hdp-node-02 128 Jun 30 2017 RELEASE

drwxr-xr-x. 2 hdp-node-02 hdp-node-02 4096 Jun 30 2017 sbin 集群启停，因为spark自带集群环境

drwxr-xr-x. 2 root root 4096 Jul 3 07:44 spark-warehouse

drwxr-xr-x. 2 hdp-node-02 hdp-node-02 4096 Jun 30 2017 yarn 对于yarn支持的jar包