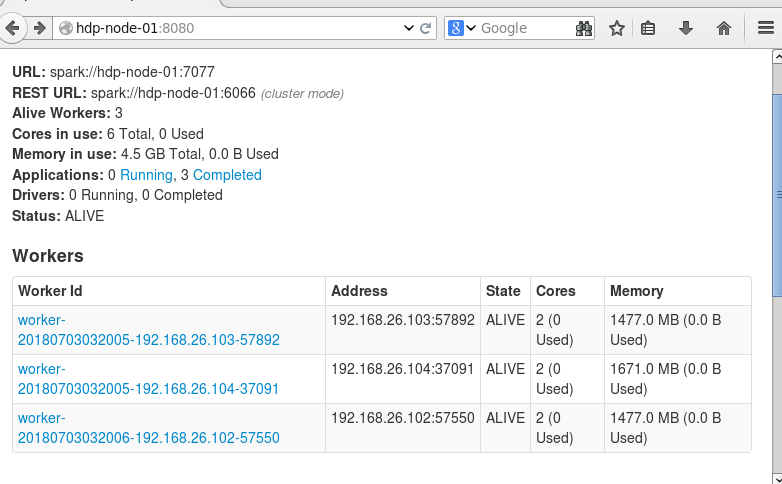
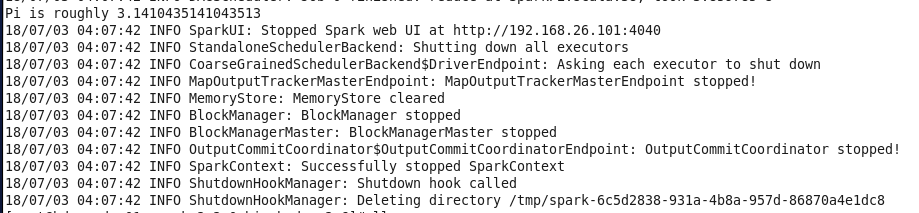
spark 第一天作业

1. 讲集群搭建起来，运行，并在web界面观察

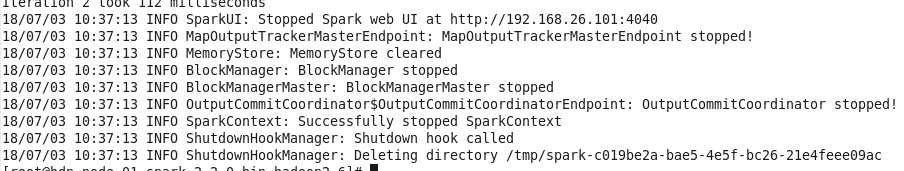


1. 自行选择5个org\apache\spark\examples包下面的demo 在集群里面跑，并观察效果

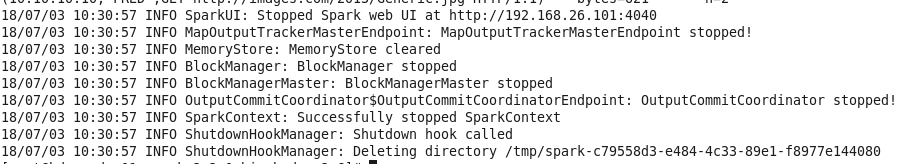
**SparkPi**



**BroadcastTest**



**LogQuery**

****

**GroupByTest**

****

**LocalPi**

****

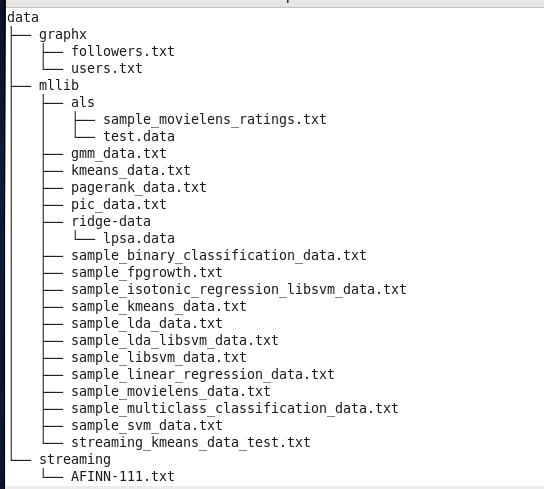
1. Spark安装目录下面的文件夹一个一个cd 进去观察，理解每一个文件夹的作用
2. **bin : 里面存放的是操作spark集群的命令**

****

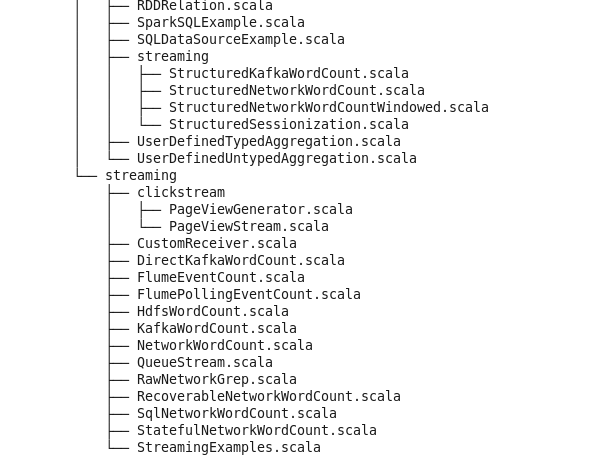
1. **conf: 存放的是配置文件，文件的配置包括slaves , spark运行环境配置等等**

****

1. **data: 存放的是机器学习demo会用到的数据**

****

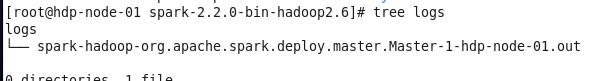
1. **examples: spark自带的小demo**



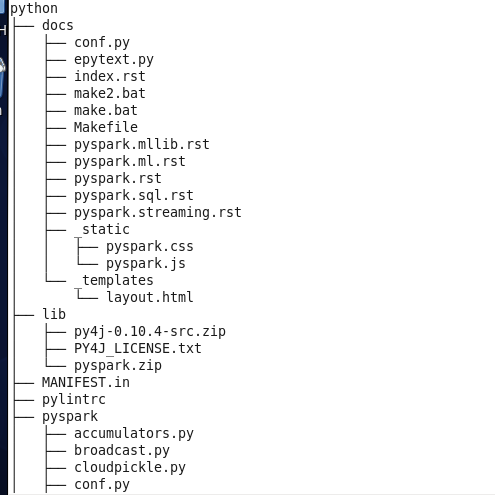
1. **Licenses ：存放的都是许可文件**

****

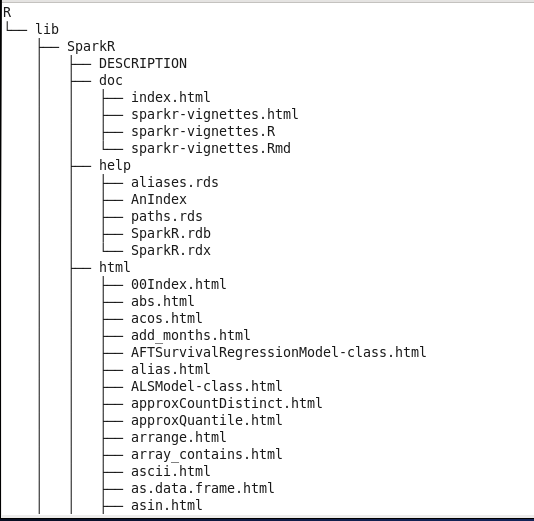
1. **logs: 运行时生成的文件夹，存放的都是运行期的日志**

****

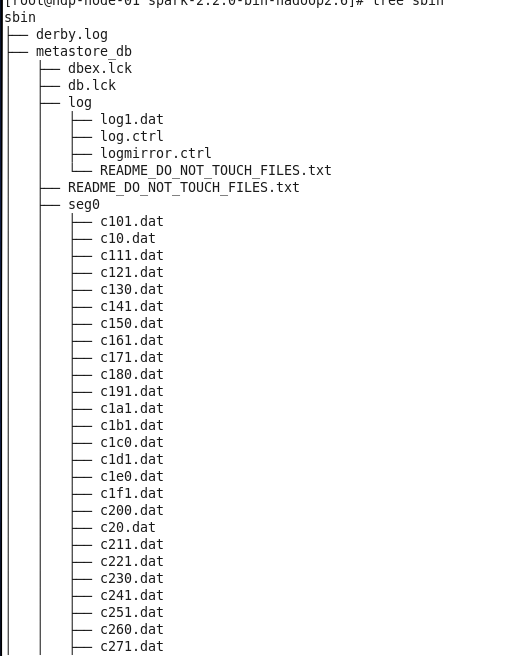
1. **python: 对python的支持**

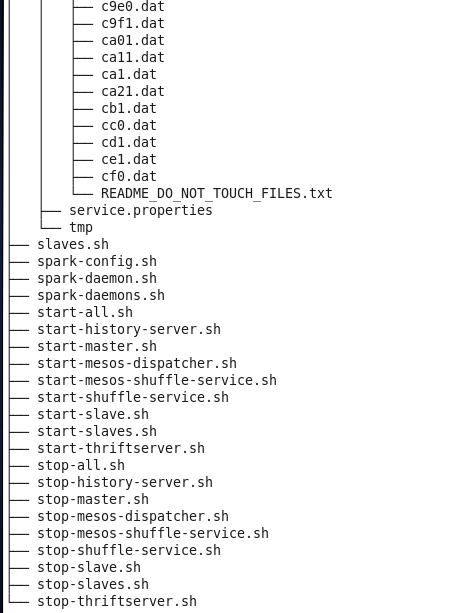
****

1. **R: 对R语言的支持**

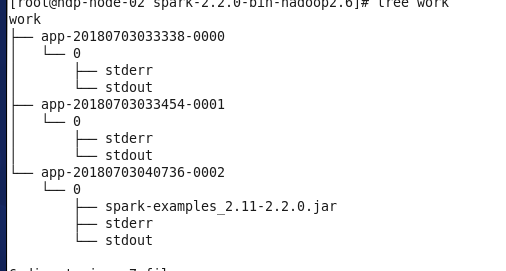
****

1. **sbin: 关闭、打开集群的shell**

****

****

**10.work: 工作目录，运行后生成。存放的就是提交的应用记录**

****