## 实训：动手搭建Spark集群

准备好了两台虚拟机。一台是master，一台是slave1。现在，基于这两台虚拟机继续搭建Spark集群。

操作步骤：

在slave1节点上安装Spark程序。

步骤01：执行完Standalone模式操作后，通过如下命令，将整个Spark目录拷贝到slave1节点。

scp -r /usr/local/spark root@slave1:/usr/local/

步骤02：修改slave1环境变量，保持和master一致。

vi /etc/profile

修改内容如下：

export JAVA\_HOME=/home/app/jdk1.8

export PATH=$PATH:$JAVA\_HOME/bin

export HADOOP\_HOME=/usr/local/hadoop

export SPARK\_HOME=/usr/local/spark

export PYTHONPATH=$SPARK\_HOME/python:$SPARK\_HOME/python/lib/py4j-0.10.7-src.zip:$PYTHONPATH

export PYSPARK\_PYTHON=python3

export PATH=$HADOOP\_HOME/bin:$SPARK\_HOME/bin:$JAVA\_HOME/bin:$PATH

步骤02：执行命令使修改生效。

source /etc/profile

步骤03：启动Spark集群。

cd $SPARK\_HOME

./sbin/start-all.sh

刷新<http://192.168.192.128:8080/页面，如图14-14>所示，出现两个工作节点。

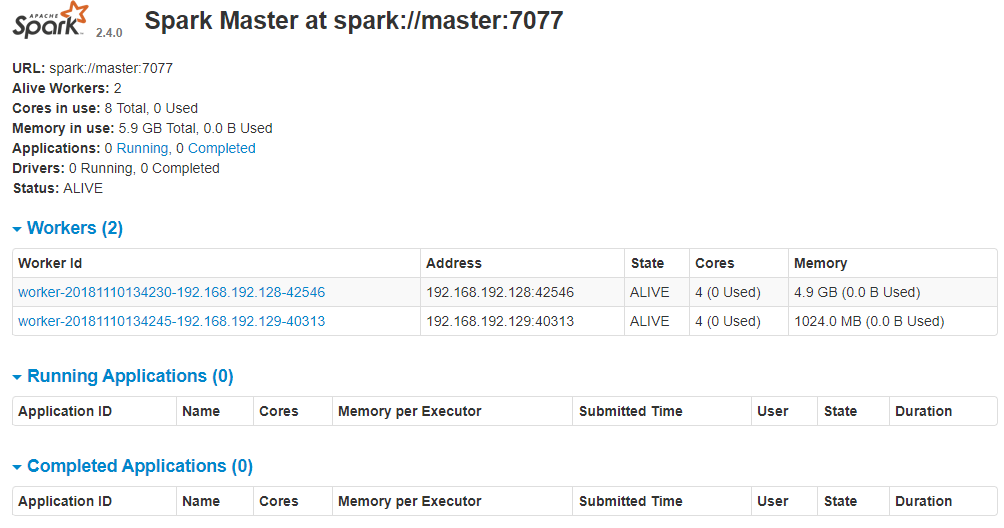
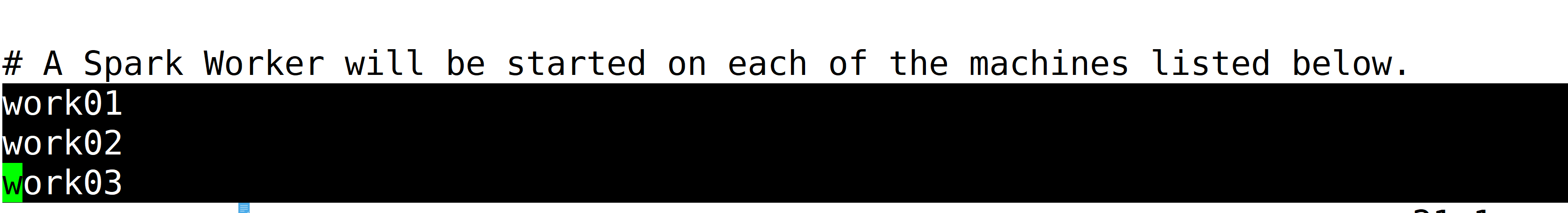


图14-14 Spark集群页面

修改配置文件：

Slaves



spark-env.sh

