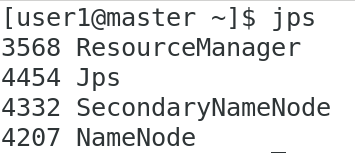
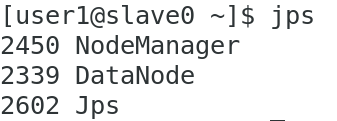
大数据第6周

1. 启动集群，检查是否正常

用命令：start-all.sh启动集群，用jps检查进程，按照我配置，master节点：



Slave节点：



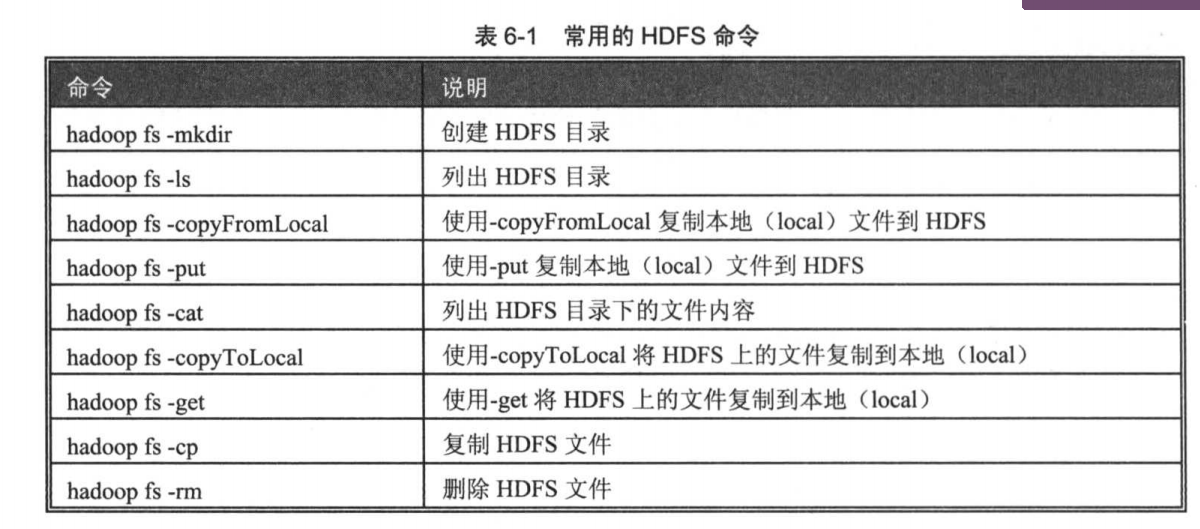
如果要全面检测，需要运行mapreduce的example程序。hadoop-2.7/share/hadoop/mapreduce下有example程序，运行方式：

hadoop jar hadoop-mapreduce-examples-2.7.7.jar pi 10 10

如果显示如下信息，说明运行正常。



1. hdfs命令



上传一个文件，用本地命令找到文件位置，理解hdfs文件存储方式。

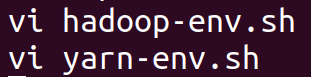
上传一个>block的文件，用本地命令找到多个文件的位置，然后重新组合，理解HDFS分块的方式。

1. 从非集群主机使用集群的方法

从非集群主机访问集群，首先需要网络畅通，还需要有命令解释器，其次要知道集群配置情况，最简单的办法就是从集群中拷贝一个hadoop软件包。

还需要一个同版本的java。

一般情况下，java路径与集群中的java路径不同，需要修改：修改



注意用户名问题。

4．Mapreduce程序开发

4.1 下载解压eclipse。

解压后启动eclipse。