

Hadoop配置历史服务器及访问历史服务器日志

Hadoop历史服务器

Hadoop自带了一个历史服务器，可以通过历史服务器查看已经运行完的Mapreduce作业记录，比如用了多少个Map、用了多少个Reduce、作业提交时间、作业启动时间、作业完成时间等信息。默认情况下，Hadoop的历史服务器是没有启动的，我们可以通过下面的指令来启动Hadoop历史服务器

配置步骤

1. 配置mapred-site.xml

```
<!-- 指定历史服务器的IP和端口 -->
<property>
  <name>mapreduce.jobhistory.address</name>
  <value>hadoop102:10020</value>
</property>

<!-- 指定历史服务器web访问的IP和端口 -->
<property>
  <name>mapreduce.jobhistory.webapp.address</name>
  <value>hadoop102:19888</value>
</property>
```

2. 启动历史服务器

```
sbin/mr-jobhistory-daemon.sh start historyserver
```

```
[root@master ~]# cd /usr/local/hadoop/
[root@master hadoop]# cd etc/hadoop/
[root@master hadoop]# vim mapred-site.xml
[root@master hadoop]# cd ..
[root@master etc]# cd ..
[root@master hadoop]# sbin/mr-jobhistory-daemon.sh start historyserver
starting historyserver, logging to /usr/local/hadoop/logs/mapred-root-historyserver-master.out
[root@master hadoop]# jps
32100 NameNode
13044 Jps
27047 JobHistoryServer
32233 DataNode
13004 JobHistoryServer
32589 NodeManager
32479 ResourceManager
```

如图所示，出现JobHistoryServer进程

4. Web端查看

前提是hosts文件配置了域名映射
<http://master:19888/jobhistory>

