# 1. Standalone模式

Standalone模式是Spark自带的资源调动引擎，构建一个由Master + Slave构成的Spark集群，Spark运行在集群中。

这个要和Hadoop中的Standalone区别开来。这里的Standalone是指只用Spark来搭建一个集群，不需要借助其他的框架。是相对于Yarn和Mesos来说的。

1.解压一份Spark安装包，并修改解压后的文件夹名称为spark-standalone

[root@bigdata1 software]# tar -zxvf spark-3.0.3-bin-hadoop3.2.tgz -C /opt/module/  
[root@bigdata1 module]# mv spark-3.0.3-bin-hadoop3.2/ spark-standalone

2.配置集群节点

[root@bigdata1 conf]# mv slaves.template slaves  
[root@bigdata1 conf]# vim slaves  
bigdata1  
bigdata2  
bigdata3

3.修改spark-env.sh文件，添加master节点

[root@bigdata1 conf]# mv spark-env.sh.template spark-env.sh  
[root@bigdata1 conf]# vim spark-env.sh  
SPARK\_MASTER\_HOST=bigdata1  
SPARK\_MASTER\_PORT=7077

4.向其他机器分发spark-standalone包

在其他机器创建spark-standalone目录。

[root@bigdata1 spark-standalone]# scp -r /opt/module/spark-standalone/ bigdata2:/opt/module/spark-standalone/  
[root@bigdata1 spark-standalone]# scp -r /opt/module/spark-standalone/ bigdata3:/opt/module/spark-standalone/

5.启动spark集群

[root@bigdata1 spark-standalone]# sbin/start-all.sh  
starting org.apache.spark.deploy.master.Master, logging to /opt/module/spark-standalone/logs/spark-root-org.apache.spark.deploy.master.Master-1-bigdata1.out  
bigdata1: starting org.apache.spark.deploy.worker.Worker, logging to /opt/module/spark-standalone/logs/spark-root-org.apache.spark.deploy.worker.Worker-1-bigdata1.out  
bigdata3: starting org.apache.spark.deploy.worker.Worker, logging to /opt/module/spark-standalone/logs/spark-root-org.apache.spark.deploy.worker.Worker-1-bigdata3.out  
bigdata2: starting org.apache.spark.deploy.worker.Worker, logging to /opt/module/spark-standalone/logs/spark-root-org.apache.spark.deploy.worker.Worker-1-bigdata2.out  
​  
#jps 检查是否有master（bigdata1）和woker（bigdata1、bigdata2、bigdata3）

6.官方求PI案例

[root@bigdata1 spark-standalone]# bin/spark-submit \  
 --class org.apache.spark.examples.SparkPi \  
 --master spark://bigdata1:7077 \  
 ./examples/jars/spark-examples\_2.12-3.0.3.jar \  
 10  
​

bin/spark-submit --class org.apache.spark.examples.SparkPi --master spark://bigdata1:7077 ./examples/jars/spark-examples\_2.12-3.0.3.jar 10  
​  
结果：  
Pi is roughly 3.1408591408591406  
​  
------------------------------ 命令 ---------------------------------------  
bin/spark-submit \  
--class org.apache.spark.examples.SparkPi \  
--master spark://bigdata1:7077 \  
--executor-memory 2G \  
--total-executor-cores 2 \  
./examples/jars/spark-examples\_2.12-3.0.3.jar \  
10

# 2. Yarn模式（仅这个就可以）

saprk客户端连接Yarn，不需要额外构建集群。

1.停止Standalone模式下的spark集群

[root@bigdata1 spark-standalone]# sbin/stop-all.sh

2.再单独解压一份spark用来做Spark on Yarn 模式

[root@bigdata1 software]# tar -zxvf spark-3.0.3-bin-hadoop3.2.tgz -C /opt/module/  
[root@bigdata1 module]# mv spark-3.0.3-bin-hadoop3.2/ spark-yarn

3.配置环境变量

#SPARK\_HOME  
export SPARK\_HOME=/opt/module/spark-yarn  
export PATH=$PATH:$SPARK\_HOME/bin

4.修改配置

#当机器内存较少时，防止执行过程进行被意外杀死，可以做如下配置：  
#修改hadoop配置文件/opt/module/hadoop-3.1.3/etc/hadoop/yarn-site.xml，添加如下内容  
<property>  
     <name>yarn.nodemanager.pmem-check-enabled</name>  
     <value>false</value>  
</property>  
​  
<property>  
     <name>yarn.nodemanager.vmem-check-enabled</name>  
     <value>false</value>  
</property>

5.分发配置文件

[root@bigdata1 hadoop]# scp -r /opt/module/hadoop-3.1.3/etc/hadoop/yarn-site.xml bigdata2:/opt/module/hadoop-3.1.3/etc/hadoop/  
​  
[root@bigdata1 hadoop]# scp -r /opt/module/hadoop-3.1.3/etc/hadoop/yarn-site.xml bigdata3:/opt/module/hadoop-3.1.3/etc/hadoop/

6.修改spark-env.sh

[root@bigdata1 conf]# mv spark-env.sh.template spark-env.sh  
[root@bigdata1 conf]# vim spark-env.sh  
YARN\_CONF\_DIR=/opt/module/hadoop-3.1.3/etc/hadoop

7.重启Hadoop

[root@bigdata1 opt]# start-dfs.sh  
[root@bigdata1 opt]# start-yarn.sh

8.求PI

[root@bigdata1 spark-yarn]# spark-submit \  
> --class org.apache.spark.examples.SparkPi \  
> --master yarn \  
> ./examples/jars/spark-examples\_2.12-3.0.3.jar \  
> 10  
​  
结果：  
Pi is roughly 3.142211142211142

bin/spark-submit --master yarn --deploy-mode cluster --class org.apache.spark.examples.SparkPi --name SparkPi --num-executors 1 --executor-memory 1g --executor-cores 1 /opt/model/spark-3.0.3/examples/jars/spark-examples\_2.12-3.0.3.jar 100

启动spark

|  |
| --- |
| bin/spark-shell --master local |