



<http://dxmall.com.cn>

## Spark集群配置

---

### NOTES

---

- 进行下一步之前要先检查上一步的目标是否达成
- 以下操作未特殊说明的均在主节点进行
- 参考文档仅作参考，根据实际情况进行配置
- 进行删改操作之前建议对配置文件进行备份
- 设计其他详细命令及操作可参考其他文档

### 修改配置文件

---

```
cd /opt/spark-2.2.0-bin-hadoop2.7/conf/  
vim slaves
```

```
root@ubuntu1:/opt/spark-2.2.0-bin-hadoop2.7/conf
# Licensed to the Apache Software Foundation (ASF) under one or more
# contributor license agreements. See the NOTICE file distributed with
# this work for additional information regarding copyright ownership.
# The ASF licenses this file to You under the Apache License, Version 2.0
# (the "License"); you may not use this file except in compliance with
# the License. You may obtain a copy of the License at
#
# http://www.apache.org/licenses/LICENSE-2.0
#
# Unless required by applicable law or agreed to in writing, software
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the License.
#
# A Spark Worker will be started on each of the machines listed bel
ubuntu1
ubuntu2
ubuntu3
~
"slaves" 22L, 878C                                1,1      All
```

```
vim spark-env.sh
```

```
root@ubuntu1:/opt/spark-2.2.0-bin-hadoop2.7/conf
# The ASF licenses this file to You under the Apache License, Version 2.0
# (the "License"); you may not use this file except in compliance with
# the License. You may obtain a copy of the License at
#
# http://www.apache.org/licenses/LICENSE-2.0
#
# Unless required by applicable law or agreed to in writing, software
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the License.
#
export SCALA_HOME=/home/sys-01/spark/usr/share/scala
export JAVA_HOME=/home/sys-01/jdk/jdk1.8.0_131/
export SPARK_MASTER_IP=ubuntu1
export HADOOP_HOME=/home/hadoop/apps/hadoop-3.1.0
export SPARK_MASTER_PORT=7077
export HADOOP_CONF_DIR=/home/hadoop/apps/hadoop-3.1.0/etc/hadoop
export LD_LIBRARY_PATH=$HADOOP_HOME/lib/native
# This file is sourced when running various Spark programs.
# Copy it as spark-env.sh and edit that to configure Spark for your site.
29,1      12%
```

发送给其他节点

```
cd /opt/spark-2.2.0-bin-hadoop2.7
scp -r conf ubuntu2:/opt/spark-2.2.0-bin-hadoop2.7
scp -r conf ubuntu3:/opt/spark-2.2.0-bin-hadoop2.7
```

## 启动spark集群

```
cd /opt/spark-2.2.0-bin-hadoop2.7/sbin
./start-all.sh
```

启动成功后在浏览器输入localhost:8080查看集群状态

# 进入spark-shell

---

```
cd /opt/spark-2.2.0-bin-hadoop2.7/bin  
./spark-shell --master spark://ubuntu1:7077
```

不使用这个命令进入spark-shell的话还是用单机模式

## 其它

---

## 问题及解决

---

- 注意在spark集群环境下读取或写入数据不要用localhost

```
source /etc/profile
```

## 参考文档:

---

spark集群配置<https://www.cnblogs.com/yanshw/p/11614988.html>

<https://blog.csdn.net/zkhong07/article/details/94993053>

---

@author MurrayZhao 感谢指正