

# 添加环境变量

1. hadoop 安装目录
2. 使配置文件生效

## hadoop-env.sh

jdk 路径

## core-site.xml

```
<!--用于设置 Hadoop 的文件系统，由 URI 指定-->
<property>
  <name>fs.defaultFS</name>
  <!--用于指定 NameNode 地址在 hadoop01 机器上-->
  <value>hdfs://hadoop01:9000</value>
</property>
<!--配置 Hadoop 的临时目录，默认/tmp/hadoop-${user.name}-->
<property>
  <name>hadoop.tmp.dir</name>
  <value>/export/servers/hadoop-2.7.7/tmp</value>
</property>
</configuration>
```

## hdfs-site.xml

```
<configuration>
  <!--指定 HDFS 副本数量-->
  <property>
    <name>dfs.replication</name>
    <value>3</value>
  </property>
  <!--secondary namenode 所在主机的 IP 和端口-->
  <property>
    <name>dfs.namenode.secondary.http-address</name>
    <value>hadoop02:50090</value>
  </property>
</configuration>
```

## mapred-site.xml

首先需要复制

```
cp mapred-site.xml.template mapred-site.xml
```

```
<configuration>
  <!--指定 MapReduce 运行时框架, 这里指定在 YARN 上, 默认 local-->
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
</configuration>
```

## yarn-site.xml

```
<configuration>
  <!--指定 YARN 集群的管理者(ResourceManager)的地址-->
  <property>
    <name>yarn.resourcemanager.hostname</name>
    <value>hadoop01</value>
  </property>
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
</configuration>
```

## slaves

```
hadoop01
hadoop02
hadoop03
```

## 将主节点的配置和软件分发到子节点

```
scp /etc/profile hadoop02:/etc/profile
scp /etc/profile hadoop03:/etc/profile

scp -r /export/ hadoop02:/
scp -r /export/ hadoop03:/
```

## 格式化文件系统

```
hdfs namenode -format
```

## 启动集群

```
start-dfs.sh
start-yarn.sh
```

## 使用UI查看

关闭防火墙 或开放端口

192.168.198.130:8088

192.168.198.130:50070