

一. 安装YARN

1. 从<http://archive.cloudera.com/cdh4/cdh/4/> 下载最新的Hadoop 2.0版本cdh的tar.gz
 - wget <http://archive.cloudera.com/cdh4/cdh/4/hadoop-2.0.0-cdh4.1.0.tar.gz>
2. 解开新的tar.gz包
 - tar xvfz hadoop-2.2.0.tar.gz
3. 移动hadoop-2.2.0到/usr/local目录下
 - sudo mv hadoop-2.2.0 /usr/local/hadoop
4. 修改hadoop目录权限
 - sudo chown -R xx:xx hadoop
5. 设置允许远程登录
 - 系统偏好设置-共享-远程登录打开
6. 设置ssh免密码登录
 - cd ~/.ssh
 - cp id_rsa.pub authorized_keys
 - ssh localhost测试是否需要密码

二. 配置YARN

1. 设置环境变量
 - /usr/local/hadoop/etc/hadoop/hadoop-env.sh 文件中添加JAVA_HOME
 - /usr/local/hadoop/etc/hadoop/yarn-env.sh 文件中添加JAVA_HOME
 - export JAVA_HOME=/Library/Java/JavaVirtualMachines/jdk1.8.0_144.jdk/Contents/Home
 - Mac查看JAVA_HOME： /usr/libexec/java_home -V
2. 修改mapred-site.xml
 - mv /usr/local/hadoop/etc/hadoop/mapred-site.xml.template /usr/local/hadoop/etc/hadoop/mapred-site.xml
 - 修改mapred-site.xml文件

```
<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
</configuration>
```
3. 修改core-site.xml
 - 修改 /usr/local/hadoop/etc/hadoop/core-site.xml文件

```
<configuration>
  <property>
    <name>hadoop.tmp.dir</name>
    <value>file:/usr/local/hadoop/tmp</value>
    <description>Abase for other temporary directories.</description>
  </property>
  <property>
    <name>fs.defaultFS</name>
    <value>hdfs://localhost:9000</value>
  </property>
</configuration>
```
4. 修改hdfs-site.xml
 - 修改 /usr/local/hadoop/etc/hadoop/hdfs-site.xml文件

```
<configuration>
  <property>
    <name>dfs.replication</name>
    <value>1</value>
  </property>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>file:/usr/local/hadoop/tmp/dfs/name</value>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>file:/usr/local/hadoop/tmp/dfs/data</value>
  </property>
</configuration>
```
5. 修改yarn-site.xml
 - 修改 /usr/local/hadoop/etc/hadoop/yarn-site.xml文件

```
<configuration>
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
</configuration>
```

三. 启动HDFS

1. namenode格式化
 - cd /usr/local/hadoop/bin
 - ./hdfs namenode -format
2. 启动namenode和datanode
 - cd /usr/local/hadoop/sbin
 - ./start-dfs.sh （可以用下面2个命令分别启动namenode和datanode进程）
 - ./sbin/hadoop-daemon.sh start namenode
 - ./sbin/hadoop-daemon.sh start datanode （如果有多个datanode，需使用hadoop-daemons.sh）
3. 通过jps命令查看是否启动成功

四. 启动YARN

1. 启动yarn
 - cd /usr/local/hadoop/sbin
 - ./start-yarn.sh （可以用下面2个命令分别启动resourcemanager和nodemanager进程）
 - ./sbin/yarn-daemon.sh start resourcemanager
 - ./sbin/yarn-daemon.sh start nodemanager
2. 通过jps命令查看是否启动成功
3. 通过浏览器访问<http://localhost:8088/cluster> YARN是否启动成功