HADOOP分布式集群环境搭建

# 准备工作

1. **修改主机名**

使用命令：

hostnamectl set-hostname master 【在master主机上执行】  
hostnamectl set-hostname slave01 【在slave01主机上执行】  
hostnamectl set-hostname slave02 【在slave02主机上执行】

**（2）修改hosts文件，三台机器依次执行**

#sudo vim /etc/hosts

192.168.137.128 master

192.168.137.129 slave01

192.168.137.130 slave02

**（3）关闭防火墙（root权限）**

# Centos7 中使用 systemctl 命令来管理服务，命令格式如下

# systemctl [start 开启]|[stop 停止]|[restart 重启]|[status 状态][enable 开机启动]| [disable 禁止开机启动] 服务名称

[root@localhost ~]# systemctl start firewalld # 开启防火墙

[root@localhost ~]# systemctl status firewalld # 查看防火墙状态

[root@localhost ~]# systemctl disable firewalld # 永久关闭防火墙（重启生效）

[root@localhost ~]# systemctl stop firewalld # 临时关闭防火墙（立即生效）

[root@localhost ~]# systemctl status firewalld # 检查修改

**（4）关闭SELinux（root权限）**

因为CentOS的所有访问权限都是有SELinux来管理的，为了避免我们安装中由于权限关系而导致的失败，需要先将其关闭，以后根据需要再进行重新管理。

在CentOS中关闭SELinux使用如下操作命令：

[root@localhost ~]# getenforce # 查看当前的 SELinux 状态Enforcing

# setenforce 1 可以设置 SELinux 为 enforcing 模式

[root@localhost ~]# setenforce 0 # 将 SELinux 的状态临时设置为 Permissive 模式（立即生效）

[root@localhost ~]# getenforce # 检查修改Permissive

[root@localhost ~]# vi /etc/selinux/config

 # 编辑 config 文件将 SELINUX=enforcing 修改为 SELINUX=disabled（重启生效）

[root@localhost ~]# cat /etc/selinux/config #检查修改

注意:使用 getenforce 命令获取当前 SELinux 的运行状态为 permissive 或者 disabled 时均表示关闭。

**(5) 创建接下来hadoop新用户**

1.#su root #以root账号登录并创建接下来hadoop新用户

2.#useradd -m hadoop -s /bin/bash #创建新用户hadoop

3.#passwd hadoop 给hadoop新用户设置登录密码

4.visudo 增加管理员权限（root权限角色）

5. 找到root ALL=（ALL）ALL这一行，一般在98行附近，然后在这一行下面 增加下面指令： hadoop ALL=（ALL）ALL （当中的间隔为tab）

6. 添加完之后，保存退出

7. #su hadoop,切换到hadoop用户上来

**(6)配置ssh免密登录**

#生成ssh免密登录密钥

ssh-keygen –t rsa (四个回车)

执行完这个命令后，会生成id\_rsa(私钥)、id\_rsa.pub(公钥)

将公钥拷贝到要免密登录的目标机器上

ssh-copy-id localhost

ssh-copy-id master

ssh-copy-id slave1

ssh-copy-id slave2

测试：

#ssh master 登录成功

#ssh slave01 登录成功

#ssh slave02 登录成功

# HADOOP分布式部署

# Zookeeper分布式部署

# Hbase分布式部署

linux （CentOS 7）报错：yum 命令报错 “ There are no enabled repos ”

<https://blog.csdn.net/qq_39720249/article/details/84104646>

yum 出现 Loaded plugins: fastestmirror解决办法

<https://blog.csdn.net/wang_magento/article/details/93205584>

**关闭或者禁用防火墙， systemctl  stop firewalld.service  关闭防火墙；systemctl disable firewalld.service  关闭防火墙**

**firewall-cmd --state  查看状态**

1. 安装JDK

查看当前系统是否已安装JDK

#yum list installed |grep java

#yum -y list java\*

安装JDK

#yum -y install java-1.8.0-openjdk\*

查看jdk安装的路径：#which java

(5) 安装jdk

配置环境变量：

#JAVA\_HOME

export JAVA\_HOME=/usr/local/java/jdk-13.0.2

export PATH=$PATH:$JAVA\_HOME/bin

#HADOOP\_HOME

export HADOOP\_HOME=/usr/local/hadoop/hadoop-3.2.1

export PATH=$PATH:$HADOOP\_HOME/bin

export PATH=$PATH:$HADOOP\_HOME/sbin

#zookeeper environment

export ZOOKEEPER\_HOME=/usr/local/zookeeper

export PATH=$ZOOKEEPER\_HOME/bin:$PATH

#hbase environment

export HBASE\_HOME=/usr/local/hbase

export PATH=$HBASE\_HOME/bin:$PATH

(6) 安装hadoop

基于centos7部署hadoop集群的准备环境部署

<https://blog.51cto.com/19940919/2174218>

wget <https://mirrors.tuna.tsinghua.edu.cn/apache/hadoop/common/stable/hadoop-2.9.1.tar.gz>

wget: unable to resolve host address解决方法

<https://blog.csdn.net/zong596568821xp/article/details/80718652>

Centos环境下yum安装更新jdk、删除自带的jdk

<https://blog.csdn.net/u013641234/article/details/76158026>

1.查看CentOS自带JDK是否已安装:yum list installed | grep java

2.假使存在自带的jdk，删除centos自带的JDK

   yum -y remove java-1.7.0-openjdk\*

   yum -y remove tzdata-java.noarch

结果显示为Complete!表示卸载完成!

3. ll(再次检查安装包是否删除)

CentOS7下Hadoop3.2.1的安装与部署（单机模式）

<https://blog.csdn.net/pengjunlee/article/details/104290537>

**hadoop提示权限不够**

解决方法：sudo chmod -R a+w /usr/local/hadoop 添加权限即可

大数据开发一 Hadoop3.2.1 集群搭建与启动

<https://blog.csdn.net/qq_41515513/article/details/101873098>

master是我们的namanode，该机器的IP是192.168.10.198，在本地电脑访问如下地址:

<http://192.168.42.134:9870/>

find / name 'zookeeper\_server.pid'