

大数据基础环境配置

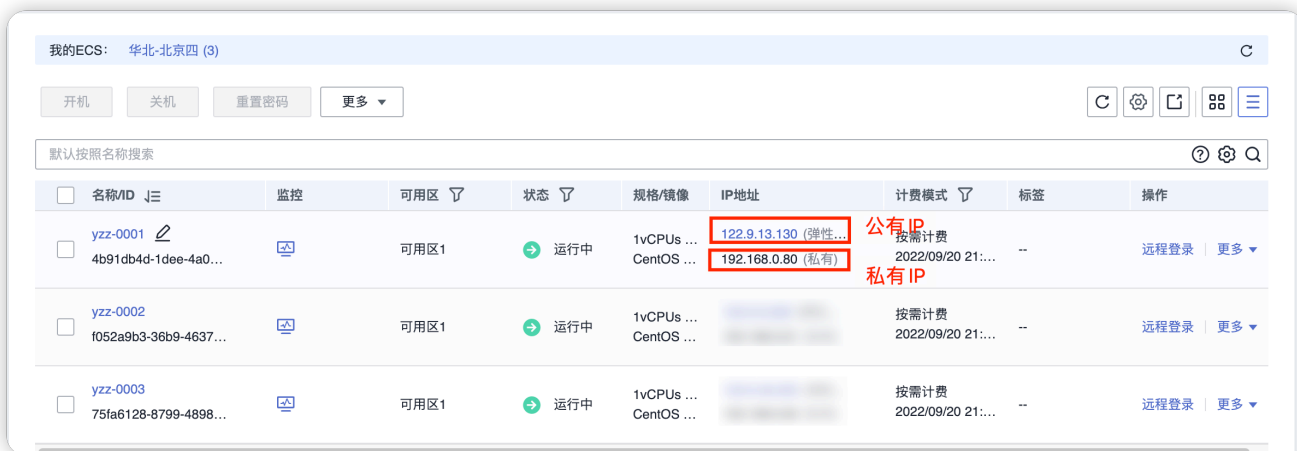
基础环境：CentOS 7.8

本教程包括：

- 连接服务器
- yum基础软件包安装
- 目录创建
- 主机名称修改
- ssh免密配置
- JDK环境配置
- Linux用户配置（可选）

0. 连接服务器

1. 在申请到华为云服务器之后，我们可以在 [服务器控制台](#) 中查看服务器的 [公网IP地址](#) 和 [私有IP地址\(局域网\)](#)。
如下图所示，通过 [远程终端工具](#) 或 `ssh` 命令，我们可以先以 `root` 用户登录三台服务器。
 1. 对于Mac用户，使用系统的默认终端，通过 `ssh root@ip` 命令即可连接服务器，命令中的ip需要替换为你的公网ip。以下图第一个服务器为例，连接命令为 `ssh root@122.9.13.130`。命令执行后，根据提示输入密码即可连接。
 2. 对于Windows用户，可以使用 [Putty](#)、[Xshell](#) 等工具，具体教程可参考 [Putty使用方法](#) 和 [XShell教程](#)



2. 在连接完服务器后，就可以在终端内执行命令

Last login: Wed Sep 21 14:10:06 2022 from 123.119.238.176

Welcome to Huawei Cloud Service

```
[root@yzz-0001 ~]# ifconfig
eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.0.80 netmask 255.255.255.0 broadcast 192.168.0.255
    inet6 fe80::f816:3eff:feec:4494 prefixlen 64 scopeid 0x20<link>
    ether fa:16:3e:ec:44:94 txqueuelen 1000 (Ethernet)
    RX packets 547 bytes 108888 (106.3 KiB)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 456 bytes 64242 (62.7 KiB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

lo: flags=73<UP,LOOPBACK,RUNNING> mtu 65536
    inet 127.0.0.1 netmask 255.0.0.0
    inet6 ::1 prefixlen 128 scopeid 0x10<host>
    loop txqueuelen 1000 (Local Loopback)
    RX packets 0 bytes 0 (0.0 B)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 0 bytes 0 (0.0 B)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

[root@yzz-0001 ~]#
```

1.

yum基础软件包安装

下列命令需要在各个节点上分别执行

执行如下命令来安装必要的软件包

```
sudo yum install -y epel-release
sudo yum install -y psmisc nc net-tools rsync vim lrzsz ntp libzstd openssl-static gcc
tree iotop git htop iperf hdparm
```

其中，**htop** 工具可以很方便地查看服务器的资源使用情况，在之后的综合实验中，我们需要观察服务器的负载情况，如果内存不足，我们可以选择对服务器进行扩容

CPU[

Mem[|||||

Swp[

0.0%]

143M/1.79G]

0K/0K]

Tasks: 29, 35 thr; 1 running

Load average: 0.59 0.22 0.12

Uptime: 00:54:14

PID	USER	PRI	NI	VIRT	RES	SHR	S	CPU%	MEM%	TIME+	Command
1	root	20	0	122M	3972	2608	S	0.0	0.2	0:01.32	/usr/lib/systemd/systemd --switched-root --system --deser
342	root	20	0	39056	2764	2452	S	0.0	0.1	0:00.09	/usr/lib/systemd/systemd-journald
367	root	20	0	46264	2728	1340	S	0.0	0.1	0:00.11	/usr/lib/systemd/systemd-udev
396	root	16	-4	55532	860	456	S	0.0	0.0	0:00.00	/sbin/auditd
395	root	16	-4	55532	860	456	S	0.0	0.0	0:00.00	/sbin/auditd
571	polkitd	20	0	597M	10076	4668	S	0.0	0.5	0:00.00	/usr/lib/polkit-1/polkitd --no-debug
572	polkitd	20	0	597M	10076	4668	S	0.0	0.5	0:00.00	/usr/lib/polkit-1/polkitd --no-debug
573	polkitd	20	0	597M	10076	4668	S	0.0	0.5	0:00.00	/usr/lib/polkit-1/polkitd --no-debug
574	polkitd	20	0	597M	10076	4668	S	0.0	0.5	0:00.00	/usr/lib/polkit-1/polkitd --no-debug
575	polkitd	20	0	597M	10076	4668	S	0.0	0.5	0:00.00	/usr/lib/polkit-1/polkitd --no-debug
576	polkitd	20	0	597M	10076	4668	S	0.0	0.5	0:00.00	/usr/lib/polkit-1/polkitd --no-debug
482	polkitd	20	0	597M	10076	4668	S	0.0	0.5	0:00.01	/usr/lib/polkit-1/polkitd --no-debug

之后为了防止JDK环境冲突，我们尝试卸载掉系统自带的OpenJDK，方便我们后续安装Oracle的JDK1.8。也有可能系统并没有自带OpenJDK

```
rpm -qa | grep -i java | xargs -n1 sudo rpm -e --nodeps
```

2. 目录创建

下列命令需要在各个节点上分别执行

- 1. 为了规范后续的安装步骤，我们创建一些目录用于安装环境或存放安装包

```
sudo mkdir /opt/module # 该目录后续用于安装环境
sudo mkdir /opt/software # 存放软件包
```

注：mkdir 命令用于新建文件夹

- 2. 如果后续不是使用 root 用户，还需要修改用户权限

```
sudo chown hadoop:hadoop /opt/module
sudo chown hadoop:hadoop /opt/software
```

注：chown 命令用于修改目录或文件权限

- 3. 创建环境变量文件

环境变量文件可以配置终端命令的扫描路径(PATH)，以及全局的环境变量，这些环境变量可以被命令或者程序调用

```
sudo touch /etc/profile.d/hadoopenv.sh
```

注：touch 命令用于创建文件

3. 主机名称修改

下列命令需要在各个节点上分别执行

1. 修改本机主机名称，格式为 `hadoop1-[学号] \ hadoop2-[学号] \ hadoop3-[学号]` 如 `hadoop1-2022110946`

```
sudo hostnamectl --static set-hostname hadoop1-2022110946
```

可以通过 `hostname` 命令查看是否修改成功

```
[root@server-0001 ~]$ hostname  
hadoop1-2022110946
```

2. 使用sudo权限修改 `/etc/hosts` 文件，添加配置好的 `主机名称` 和 `私有ip`

```
# 内容供参考，需要根据配置的静态ip修改  
192.168.0.80 hadoop1-2022110946  
192.168.0.81 hadoop2-2022110946  
192.168.0.82 hadoop3-2022110946
```

4. ssh免密配置

为了集群节点直接相互访问，我们需要配置节点之间的ssh免密登录。

1. 关闭防火墙，三节点都需要执行

```
sudo systemctl stop firewalld  
sudo systemctl disable firewalld
```

2. 生成密钥，三节点都需要执行

```
ssh-keygen -t rsa
```

执行完一直按回车即可

1. 相互拷贝密钥。三节点都需要执行,命令一条一条执行,需要输密码。

```
ssh-copy-id hadoop1-2022110946  
ssh-copy-id hadoop2-2022110946  
ssh-copy-id hadoop3-2022110946
```

5. JDK环境配置

下列安装过程需要在各个节点上执行

软件包 `jdk-8u341-linux-x64.tar.gz` 需要登录Oracle账号下载, 下载后通过 `sftp` 或 `ftp` 或 `scp` 上传文件到服务器

1. 下载jdk软件包。参考上述方法自行下载上传。
2. 解压软件包到指定文件夹

```
tar -xzvf jdk-8u341-linux-x64.tar.gz -C /opt/module/
```

3. 配置环境变量, 在 `/etc/profile.d/hadoopenv.sh` 中加入以下内容

```
# JAVA_HOME  
export JAVA_HOME=/opt/module/jdk1.8.0_341  
  
export PATH=$JAVA_HOME/bin:$PATH
```

4. 使环境变量生效

```
source /etc/profile
```

5. 测试jdk环境

```
java -version
```

```
[hadoop@hadoop1-2022110946 jdk1.8.0_341]$ java -version
java version "1.8.0_341"
Java(TM) SE Runtime Environment (build 1.8.0_341-b10)
Java HotSpot(TM) 64-Bit Server VM (build 25.341-b10, mixed mode)
```

6. Linux用户配置(可选)

下列命令需要在各个节点上分别执行

在后续的实验中，大家可以选择创建一个新的linux用户，并在该用户下进行操作。虽然使用root用户并不影响后续操作，但并不推荐直接使用root来进行项目开发和环境配置。**本节内容可选做！**

1. 添加新用户

以root用户登录服务器后，执行 `useradd` 命令添加用户

```
# 添加用户 -m选项表示自动为新用户创建home目录
```

```
useradd -m hadoop
```

```
# 配置用户密码 在实验中为了方便，可以使用一些简短的密码，虽然终端会提示密码复杂度不足，但不影响
passwd hadoop
```

2. 为新用户添加 `sudo` 权限。对于一些敏感命令，或特殊目录下的文件操作，我们需要以sudo权限去执行命令，所以我们需要为新用户赋予该权限，使得我们可以执行一些高权限命令。

修改 `/etc/sudoers` 文件

vim编辑器包含insert模式和命令模式，进入vim后，按 `i` 键即可进入insert模式，该模式下可以按正常的编辑逻辑进行修改。完成修改后，按 `esc` 退出insert模式，再输入 `:wq` 即可保存退出。如果对文件没有修改权限可能会提示无法保存，这时需要通过 `:q!` 退出后，使用sudo来执行vim编辑，或者使用 `:wq!` 退出

- **vim下尽量不要使用中文输入法！！**

```
sudo vim /etc/sudoers
```

找到如下内容，并添加 `hadoop` 用户权限。提示，可以在vim的命令模式下，输入 `/content` 查找匹配content的内容

```
## Allow root to run any commands anywhere
root    ALL=(ALL)    ALL
# 下面是添加的内容
hadoop  ALL=(ALL)    NOPASSWD:ALL
```

编辑效果：

```
## Allow root to run any commands anywhere
root    ALL=(ALL)    ALL
hadoop  ALL=(ALL)    NOPASSWD:ALL
```

保存退出后，可以再次进入vim，查看文件是否成功修改！

3. 切换为 `hadoop` 用户

执行如下命令，即可切换为新建用户，之后可以直接通过该用户连接服务器

```
su - hadoop
```

##