

Hadoop 2.x 集群系统安装准备工作

侯在钱

目 录

1. 集群系统规划 1
2. 系统环境准备 2

本文安装的版本是：hadoop-2.6.0.tar.gz。

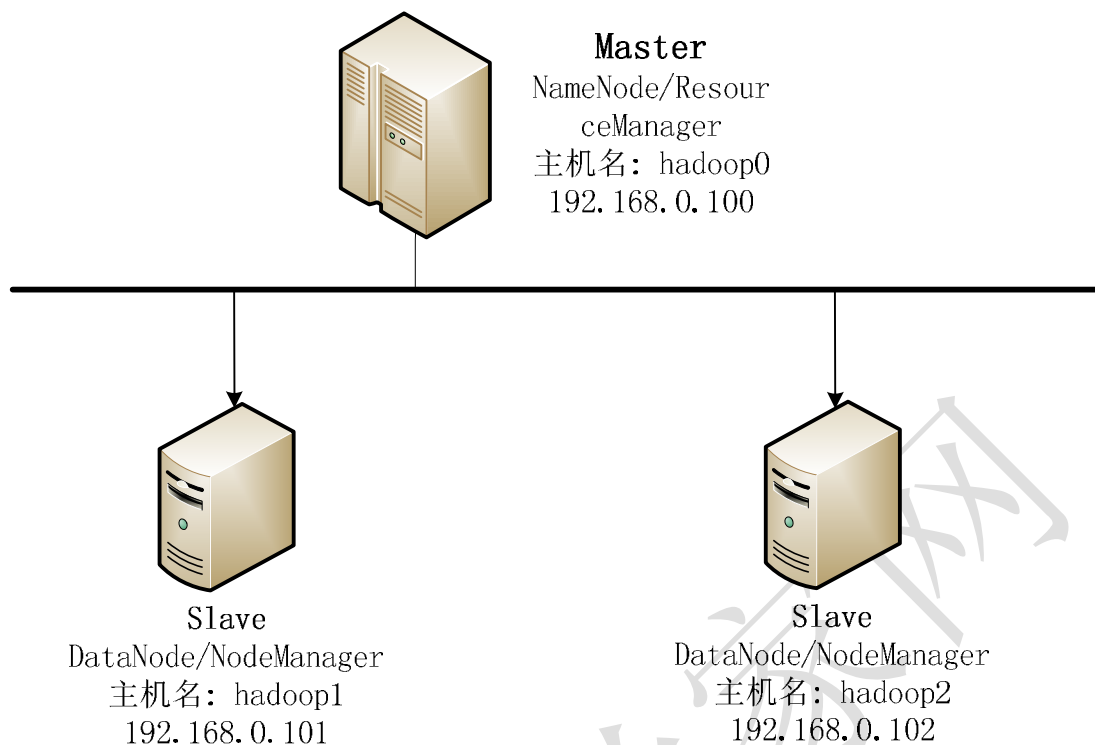
1. 集群系统规划

在安装 Hadoop 系统前，必须对整个 Hadoop 集群系统做一下规划。规划的内容主要包括：

- (1) 有几台机器作为集群中的节点。
- (2) 每个节点担任的角色分别是什么。
- (3) 每台机器的 IP 和主机名。

以下是本文准备实现的 Hadoop 集群系统：

主机名	IP	角色	说明
hadoop0	192.168.0.100	Master	NameNode/ResourceManager
hadoop1	192.168.0.101	Slave	DataNode/NodeManager
hadoop2	192.168.0.102	Slave	DataNode/NodeManager



2. 系统环境准备

安装 Hadoop 前,需要做以下工作,如果不会设置,请参看《23_附录(3)Hadoop 安装需要的 Linux 相关操作》。

- (1) 搭建三台 Linux 虚拟机。
- (2) 在每台 Linux 上安装 Java 开发环境 JDK1.7 版本及以上。
- (3) 修改主机名: 把三台机器的主机名分别设置为 hadoop0,hadoop1,hadoop2。配置文件: /etc/sysconfig/network。
- (4) 给三台机器分别做好 IP 映射。在每台 Linux 的/etc/hosts 文件末尾增加如下映射(三台 Linux 映射内容一样,但 IP 地址和主机名需要根据实际情况设置),本文的映射如下:

192.168.0.100	hadoop0
192.168.0.101	hadoop1
192.168.0.102	hadoop2

修改不用重启即生效。

(5) Master 上配置 SSH 能免密码登录到所有 Slave。本文也就是需要设置从 hadoop0 能免密码登录到 hadoop1, hadoop2。

然后分别在 hadoop0、hadoop1 和 hadoop2 上执行如下命令创建目录:

```
[root @hadoop0 ~]# mkdir /root/.ssh/
```

在 hadoop0 上执行的代码如下:

```
[root@hadoop0 ~]# ssh-keygen -t rsa -P "" -f /root/.ssh/id_rsa
[root@hadoop0 ~]# cat /root/.ssh/id_rsa.pub >> /root/.ssh/authorized_keys
[root@hadoop0 ~]# chmod 640 /root/.ssh/authorized_keys
[root@hadoop0 ~]#
[root@hadoop0 ~]# scp /root/.ssh/authorized_keys hadoop1:/root/.ssh/
[root@hadoop0 ~]# scp /root/.ssh/authorized_keys hadoop2:/root/.ssh/
```

然后分别在 hadoop1 和 hadoop2 上执行如下命令修改 authorized_keys 文件的读写权限:

```
[root @hadoop1 ~]# chmod 640 /root/.ssh/authorized_keys
```

设置完后, 在 hadoop0 上执行如下命令, 测试是否能免密码登录到 hadoop1:

```
[root @hadoop0 ~]# ssh hadoop1
```

如果不用输入密码, 则说明成功, 再键入 “exit” 退出 hadoop1。

再在 hadoop0 上执行如下命令, 测试是否能免密码登录到 hadoop2:

```
[root @hadoop0 ~]# ssh hadoop2
```

(6) 在每台机器上建立 hadoop 用户 (包括另外两台 Slave)。使用如下命令:

```
useradd hadoop
passwd hadoop
```

说明: 第 1 条语句是新增一个用户 “hadoop”。第 2 条命令是给 hadoop 这个用户设置密码。输入密码的时候, 感觉没响应, 没关系, 继续输入密码, 输入完毕后回车即可。设置的密码需要记住。本文设置的密码和用户名一样。)