Hadoop全分布式环境搭建

# 文件准备

Vmware

Centos7

Jdk8

Hadoop2.9.2

Ant1.9

Hadoop-eclipse-plugin-2.x.x

ssh

ntp

# 安装流程

1、安装vmware虚拟机

2、安装centos7, 并配置网络，使用nat模式

3、创建hadoop用户

4、配置ssh

5、在hadoop用户的home中创建文件夹bigdata

6、将jdk8、hadoop2.9.2的二进制文件解压到bigdata，并分别该文件夹名为jdk、hadoop

7、在hadoo用户的.bsahrc文件，配置java环境，配置路径

8、配置hadoop，修改bigdata下的hadoop/etc 下的Hadoop-env.sh, core-site.xml, hdfs-site.xml, mapred-sist.xml, yarn-site.xml

9、完整克隆虚拟机，克隆两个

10、修改虚拟机的网络配置，更改ip

11、在每个克隆机上重新配置ssh

12、建立各主机之间的免密登录

13、编译hadoop-eclipse-plugin-2.9.2

14、在第一台虚拟机上安装eclipse，并配置hadoop的编程环境

# 出现问题

1、ssh免密，无法免密登录

2、多次hdfs namenode –format 后，造成start-all.sh 无法开启全部的进程，这时需要删除配置core-site.xml, hdfs-site.xml, mapred-sist.xml, yarn-site.xml时所定义的文件夹内自动产生的内容

3、无法创建hadoop location 是因为hadoop-eclipse-plugin-2.9.2确实某些包，可以通过日志分析找出缺少的包，然后添加到hadoop-eclipse-plugin-2.9.2中并重启eclipse 就可以了