1. 环境

1.Windows7 内存4G以上

2.Centos7 64位系统

3.jdk8(jdk7以上) 64位

4.vmware 12

5.Xshell5

二、配置

1.选择相应的网络模式选为NAT或者桥接

2.修改主机名

3.如果选择的是NAT，设置static IP

4.修改主机名和IP的映射关系

5.关闭防火墙chkconfig iptables off

6.安装jdk并配置环境变量

7.安装hadoop并配置环境变量

export JAVA\_HOME=/usr/java/jdk8

export HADOOP\_HOME=/home/hadoop-2.4.1

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

1. 设置ntp时间服务器

yum install ntp

ntpdate -u s1a.time.edu.cn

9.修改配置文件

(1):hadoop-env.sh

export JAVA\_HOME=/usr/java/jdk8

1. :core-site.xml

<!-- 指定HADOOP所使用的文件系统schema（URI），HDFS（NameNode）的地址 -->

<property>

<name>fs.defaultFS</name>

<value>hdfs://zdy1:8020</value>

</property>

<!-- 指定hadoop运行时产生文件的存储目录 -->

<property>

<name>hadoop.tmp.dir</name>

<value>/home/hadoop/hadoop-2.4.1/tmp</value>

</property>

1. :hdfs-site.xml

<property>

<name>dfs.replication</name>

<value>1</value>

</property>

<property>

<name>dfs.secondary.http.address</name>

<value>zdy1:50090</value>

</property>

1. :mapred-site.xml

<property>

<name>mapreduce.framework.name</name>

<value>yarn</value>

</property>

1. :yarn-site.xml

<property>

<name>yarn.resourcemanager.hostname</name>

<value>zdy1</value>

</property>

<!-- reducer获取数据的方式 -->

<property>

<name>yarn.nodemanager.aux-services</name>

<value>mapreduce\_shuffle</value>

</property>

创建masters文件在文件里输入secondname的主机名

在slaves文件里输入DataNode的主机名。格式为：每个主机名写一行

1. 格式化namenode

hdfs namenode -format

1. 启动hadoop

启动namenode ：hadoop-daemon.sh start namenode

启动datanode ：hadoop-daemon.sh start datanode

注：ssh免密登录

ssh-keygen -t rsa

authorized\_keys

HA；参考官方文档配置

1.先启动zk集群(zdy1,zdy2,zdy3)

2.再在指定的journal机器上启动journalnode（zdy2,zdy3,zdy4）

hadoop-daemon.sh start journalnode

1. 格式化（zdy1）

hdfs namenode -format

启动namenode

1. 同步数据(zdy2)

*hdfs namenode -bootstrapStandby*

1. *格式化zk(在任何一台namenode上)*

*注：此命令必须在zk集群启动成功后执行*

hdfs zkfc -formatZK

1. 启动dfs

start-dfs.sh