NSD ARCHITECTURE DAY05

1. 案例1：安装Hadoop
2. 案例2：安装配置Hadoop

1 案例1：安装Hadoop

1.1 问题

本案例要求安装单机模式Hadoop：

- 单机模式安装Hadoop
- 安装JAVA环境
- 设置环境变量，启动运行

1.2 步骤

实现此案例需要按照如下步骤进行。

步骤一：环境准备

1）配置主机名为nn01，ip为192.168.1.21，配置yum源（系统源）

备注：由于在之前的案例中这些都已经做过，这里不再重复，不会的学员可以参考之前的案例

2）安装java环境

```
[root@nn01 ~]# yum -y install java-1.8.0-openjdk-devel
[root@nn01 ~]# java -version
openjdk version "1.8.0_131"
OpenJDK Runtime Environment (build 1.8.0_131-b12)
OpenJDK 64-Bit Server VM (build 25.131-b12, mixed mode)
[root@nn01 ~]# jps
1235 Jps
```

3）安装hadoop

```
[root@nn01 ~]# tar -xf hadoop-2.7.6.tar.gz
[root@nn01 ~]#  mv hadoop-2.7.6 /usr/local/hadoop
[root@nn01 ~]# cd /usr/local/hadoop/
[root@nn01 hadoop]# ls
bin  include  libexec      NOTICE.txt  sbin
etc  lib      LICENSE.txt  README.txt  share
[root@nn01 hadoop]# ./bin/hadoop  //报错，JAVA_HOME没有找到
Error: JAVA_HOME is not set and could not be found.
```

Top

```
[root@nn01 hadoop]#
```

4）解决报错问题

```
[root@nn01 hadoop]# rpm -ql java-1.8.0-openjdk
[root@nn01 hadoop]# cd ./etc/hadoop/
[root@nn01 hadoop]# vim hadoop-env.sh
25 export \
JAVA_HOME="/usr/lib/jvm/java-1.8.0-openjdk-1.8.0.131-11.b12.el7.x86_64/jre"

33 export HADOOP_CONF_DIR="/usr/local/hadoop/etc/hadoop"
[root@nn01 ~]# cd /usr/local/hadoop/
[root@nn01 hadoop]# ./bin/hadoop
Usage: hadoop [--config confdir] [COMMAND | CLASSNAME]
  CLASSNAME            run the class named CLASSNAME
 or
  where COMMAND is one of:
  fs                   run a generic filesystem user client
  version              print the version
  jar <jar>            run a jar file
                       note: please use "yarn jar" to launch
                             YARN applications, not this command.
  checknative [-a|-h]  check native hadoop and compression libraries availability
  distcp <srcurl> <desturl> copy file or directories recursively
  archive -archiveName NAME -p <parent path> <src>* <dest> create a hadoop archive
  classpath            prints the class path needed to get the
  credential           interact with credential providers
                       Hadoop jar and the required libraries
  daemonlog            get/set the log level for each daemon
  trace                view and modify Hadoop tracing settings

Most commands print help when invoked w/o parameters.

[root@nn01 hadoop]# mkdir /usr/local/hadoop/aa
[root@nn01 hadoop]# ls
bin  etc  include  lib  libexec  LICENSE.txt  NOTICE.txt  aa  README.txt  sbin  share
[root@nn01 hadoop]# cp *.txt /usr/local/hadoop/aa
[root@nn01 hadoop]# ./bin/hadoop jar  \
 share/hadoop/mapreduce/hadoop-mapreduce-examples-2.7.6.jar  wordcount aa bb
[root@nn01 hadoop]#  cat   bb/part-r-00000   //查看
```

Top

2 案例2：安装配置Hadoop

2.1 问题

本案例要求：

- 另备三台虚拟机，安装Hadoop
- 使所有节点能够ping通，配置SSH信任关系
- 节点验证

2.2 方案

准备四台虚拟机，由于之前已经准备过一台，所以只需再准备三台新的虚拟机即可，安装hadoop，使所有节点可以ping通，配置SSH信任关系，如图-1所示：

主机	角色	软件
192.168.1.21 nn01	NameNode SecondaryNameNode	HDFS
192.168.1.22 node1	DataNode	HDFS
192.168.1.23 node2	DataNode	HDFS
192.168.1.24 node3	DataNode	HDFS

图-1

2.3 步骤

实现此案例需要按照如下步骤进行。

步骤一：环境准备

1）三台机器配置主机名为node1、node2、node3，配置ip地址（ip如图-1所示），yum源（系统源）

2）编辑/etc/hosts（四台主机同样操作，以nn01为例）

```
[root@nn01 ~]# vim /etc/hosts
192.168.1.21 nn01
192.168.1.22 node1
192.168.1.23 node2
192.168.1.24 node3
```

Top

3）安装java环境，在node1，node2，node3上面操作（以node1为例）

```
[root@node1 ~]# yum -y install java-1.8.0-openjdk-devel
```

4）布置SSH信任关系

```
[root@nn01 ~]# vim /etc/ssh/ssh_config    //第一次登陆不需要输入yes
Host *
        GSSAPIAuthentication yes
        StrictHostKeyChecking no
[root@nn01 .ssh]# ssh-keygen
Generating public/private rsa key pair.
Enter file in which to save the key (/root/.ssh/id_rsa):
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /root/.ssh/id_rsa.
Your public key has been saved in /root/.ssh/id_rsa.pub.
The key fingerprint is:
SHA256:Ucl8OCezw92aArY5+zPtOrJ9ol1ojRE3EAZ1mgndYQM root@nn01
The key's randomart image is:
+---[RSA 2048]----+
|       o*E*=.    |
|        +XB+.    |
|       ..=Oo.    |
|       o.+o...   |
|       .S+.. o   |
|        + .=o    |
|         o+oo    |
|       o+=.o     |
|       o==O.     |
+----[SHA256]-----+
[root@nn01 .ssh]# for i in 21 22 23 24 ; do  ssh-copy-id  192.168.1.$i; done
//部署公钥给nn01，node1，node2，node3
```

5）测试信任关系

```
[root@nn01 .ssh]# ssh node1
Last login: Fri Sep  7 16:52:00 2018 from 192.168.1.21
[root@node1 ~]# exit
logout
```

Top

```
Connection to node1 closed.
[root@nn01 .ssh]# ssh node2
Last login: Fri Sep  7 16:52:05 2018 from 192.168.1.21
[root@node2 ~]# exit
logout
Connection to node2 closed.
[root@nn01 .ssh]# ssh node3
```

步骤二：配置hadoop

1）修改slaves文件

```
[root@nn01 ~]# cd  /usr/local/hadoop/etc/hadoop
[root@nn01 hadoop]# vim slaves
node1
node2
node3
```

2）hadoop的核心配置文件core-site

```
[root@nn01 hadoop]# vim core-site.xml
<configuration>
<property>
        <name>fs.defaultFS</name>
        <value>hdfs://nn01:9000</value>
    </property>
    <property>
        <name>hadoop.tmp.dir</name>
        <value>/var/hadoop</value>
    </property>
</configuration>

[root@nn01 hadoop]# mkdir /var/hadoop        //hadoop的数据根目录
[root@nn01 hadoop]# ssh node1 mkdir /var/hadoop
[root@nn01 hadoop]# ssh node2 mkdir /var/hadoop
[root@nn01 hadoop]# ssh node3 mkdir /var/hadoop
```

Top

3）配置hdfs-site文件

```
[root@nn01 hadoop]# vim hdfs-site.xml
<configuration>
 <property>
      <name>dfs.namenode.http-address</name>
      <value>nn01:50070</value>
   </property>
   <property>
      <name>dfs.namenode.secondary.http-address</name>
      <value>nn01:50090</value>
   </property>
   <property>
      <name>dfs.replication</name>
      <value>2</value>
   </property>
</configuration>
```

4）同步配置到node1，node2，node3

```
[root@nn01 hadoop]# yum –y install rsync  //同步的主机都要安装rsync
[root@nn01 hadoop]# for i in 22 23 24; do rsync -aSH --delete /usr/local/hadoop/
\   192.168.1.$i:/usr/local/hadoop/  -e 'ssh' &done
[1] 23260
[2] 23261
[3] 23262
```

5）查看是否同步成功

```
[root@nn01 hadoop]# ssh node1 ls /usr/local/hadoop/
bin
etc
include
lib
libexec
LICENSE.txt
NOTICE.txt
bb
README.txt
sbin
share
```

Top

```
aa
[root@nn01 hadoop]# ssh node2 ls /usr/local/hadoop/
bin
etc
include
lib
libexec
LICENSE.txt
NOTICE.txt
bb
README.txt
sbin
share
aa
[root@nn01 hadoop]# ssh node3 ls /usr/local/hadoop/
bin
etc
include
lib
libexec
LICENSE.txt
NOTICE.txt
bb
README.txt
sbin
share
aa
```

步骤三：格式化

```
[root@nn01 hadoop]# cd /usr/local/hadoop/
[root@nn01 hadoop]# ./bin/hdfs namenode -format        //格式化 namenode
[root@nn01 hadoop]# ./sbin/start-dfs.sh        //启动
[root@nn01 hadoop]# jps        //验证角色
23408 NameNode
23700 Jps
23591 SecondaryNameNode
[root@nn01 hadoop]# ./bin/hdfs dfsadmin -report        //查看集群是否组建成功
Live datanodes (3):        //有三个角色成功
```

Top

Top