

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

```
01. [root@nn01 ~]# yum -y install java-1.8.0-openjdk-devel
02. [root@nn01 ~]# java -version
03. openjdk version "1.8.0_131"
04. OpenJDK Runtime Environment (build 1.8.0_131-b12)
05. OpenJDK 64-Bit Server VM (build 25.131-b12, mixed mode)
06. [root@nn01 ~]# jps
07. 1235 Jps
```

3) 安装hadoop

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

[Top](#)

09. [root@nn01 hadoop] #

4) 解决报错问题

```

01. [ root@nn01 hadoop] # rpm -ql java-1.8.0-openjdk
02. [ root@nn01 hadoop] # cd ./etc/hadoop/
03. [ root@nn01 hadoop] # vim hadoop-env.sh
04. 25 export \
05. JAVA_HOME="/usr/lib/jvm/java-1.8.0-openjdk-1.8.0.131-11.b12.el7.x86_64/jre"
06.
07. 33 export HADOOP_CONF_DIR="/usr/local/hadoop/etc/hadoop"
08. [ root@nn01 ~] # cd /usr/local/hadoop/
09. [ root@nn01 hadoop] # ./bin/hadoop
10. Usage: hadoop [-config confdir] [COMMAND | CLASSNAME]
11. CLASSNAME      run the class named CLASSNAME
12. or
13. where COMMAND is one of:
14. fs              run a generic filesystem user client
15. version         print the version
16. jar <jar>       run a jar file
17.                note: please use "yarn jar" to launch
18.                YARN applications, not this command.
19. checknative [-a] [-h] check native hadoop and compression libraries availability
20. distcp <srcurl> <desturl> copy file or directories recursively
21. archive -archiveName NAME -p <parent path> <src>* <dest> create a hadoop archive
22. classpath       prints the class path needed to get the
23. credential      interact with credential providers
24.                Hadoop jar and the required libraries
25. daemonlog       get/set the log level for each daemon
26. trace           view and modify Hadoop tracing settings
27.
28. Most commands print help when invoked w/o parameters.
29.
30. [ root@nn01 hadoop] # mkdir /usr/local/hadoop/aa
31. [ root@nn01 hadoop] # ls
32. bin etc include lib libexec LICENSE.txt NOTICE.txt aa README.txt sbin share
33. [ root@nn01 hadoop] # cp *.txt /usr/local/hadoop/aa
34. [ root@nn01 hadoop] # ./bin/hadoop jar \
35. share/hadoop/mapreduce/hadoop-mapreduce-examples-2.7.6.jar wordcount aa bb
36. [ 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）

```
01. [root@nn01 ~] # vim /etc/hosts
02. 192.168.1.21 nn01
03. 192.168.1.22 node1
04. 192.168.1.23 node2
05. 192.168.1.24 node3
```

[Top](#)

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

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

4) 布置SSH信任关系

```
01. [root@nn01 ~]# vim /etc/ssh/ssh_config //第一次登陆不需要输入yes
02. Host *
03.     GSSAPIAuthentication yes
04.     StrictHostKeyChecking no
05. [root@nn01 .ssh]# ssh-keygen
06. Generating public/private rsa key pair.
07. Enter file in which to save the key (/root/.ssh/id_rsa):
08. Enter passphrase (empty for no passphrase):
09. Enter same passphrase again:
10. Your identification has been saved in /root/.ssh/id_rsa.
11. Your public key has been saved in /root/.ssh/id_rsa.pub.
12. The key fingerprint is:
13. SHA256:Ucl80Cezw92aArY5+zPtOrJ9ol1ojRE3EAZ1mgndYQMroot@nn01
14. The key's randomart image is:
15. +---[RSA 2048]-----+
16. |      o*E*=. |
17. |      +XB+. |
18. |      ..=Oo. |
19. |      o.+o... |
20. |      .St.. o |
21. |      +.=o |
22. |      o+oo |
23. |      o+=.o |
24. |      o==O. |
25. +---[SHA256]-----+
26. [root@nn01.ssh]# for i in 21 22 23 24; do ssh-copy-id 192.168.1.$i; done
27. //部署公钥给nn01, node1, node2, node3
```

5) 测试信任关系

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

[Top](#)

05. Connection to node1 closed.
06. [root@nn01 .ssh] # ssh node2
07. Last login: Fri Sep 7 16: 52: 05 2018 from 192. 168. 1. 21
08. [root@node2 ~] # exit
09. logout
10. Connection to node2 closed.
11. [root@nn01 .ssh] # ssh node3

步骤二：配置hadoop

1) 修改slaves文件

01. [root@nn01 ~] # cd /usr/local/hadoop/etc/hadoop
02. [root@nn01 hadoop] # vim slaves
03. node1
04. node2
05. node3

2) hadoop的核心配置文件core-site

01. [root@nn01 hadoop] # vim core-site.xml
02. <configuration>
03. <property>
04. <name>fs.defaultFS</name>
05. <value>hdfs://nn01:9000</value>
06. </property>
07. <property>
08. <name>hadoop.tmp.dir</name>
09. <value>/var/hadoop</value>
10. </property>
11. </configuration>
- 12.
13. [root@nn01 hadoop] # mkdir /var/hadoop //hadoop的数据根目录
14. [root@nn01 hadoop] # ssh node1 mkdir /var/hadoop
15. [root@nn01 hadoop] # ssh node2 mkdir /var/hadoop
16. [root@nn01 hadoop] # ssh node3 mkdir /var/hadoop

[Top](#)

3) 配置hdfs-site文件

```

01. [ root@nn01 hadoop] # vim hdfs-site.xml
02. <configuration>
03.   <property>
04.     <name>dfs.namenode.http-address</name>
05.     <value>nn01:50070</value>
06.   </property>
07.   <property>
08.     <name>dfs.namenode.secondary.http-address</name>
09.     <value>nn01:50090</value>
10.   </property>
11.   <property>
12.     <name>dfs.replication</name>
13.     <value>2</value>
14.   </property>
15. </configuration>

```

4) 同步配置到node1, node2, node3

```

01. [ root@nn01 hadoop] # yum -y install rsync //同步的主机都要安装rsync
02. [ root@nn01 hadoop] # for i in 22 23 24; do rsync -aSH --delete /usr/local/hadoop/
03. \ 192.168.1.$i:/usr/local/hadoop/ -e 'ssh' & done
04. [ 1] 23260
05. [ 2] 23261
06. [ 3] 23262

```

5) 查看是否同步成功

```

01. [ root@nn01 hadoop] # ssh node1 ls /usr/local/hadoop/
02. bin
03. etc
04. include
05. lib
06. libexec
07. LICENSE.txt
08. NOTICE.txt
09. bb
10. README.txt
11. sbin
12. share

```

[Top](#)

```
13. aa
14. [ root@nn01 hadoop] # ssh node2 ls /usr/local/hadoop/
15. bin
16. etc
17. include
18. lib
19. libexec
20. LICENSE.txt
21. NOTICE.txt
22. bb
23. README.txt
24. sbin
25. share
26. aa
27. [ root@nn01 hadoop] # ssh node3 ls /usr/local/hadoop/
28. bin
29. etc
30. include
31. lib
32. libexec
33. LICENSE.txt
34. NOTICE.txt
35. bb
36. README.txt
37. sbin
38. share
39. aa
```

步骤三：格式化

```
01. [ root@nn01 hadoop] # cd /usr/local/hadoop/
02. [ root@nn01 hadoop] # ./bin/hdfs namenode - format //格式化 namenode
03. [ root@nn01 hadoop] # ./sbin/start-dfs.sh //启动
04. [ root@nn01 hadoop] # jps //验证角色
05. 23408 NameNode
06. 23700 Jps
07. 23591 Secondary NameNode
08. [ root@nn01 hadoop] # ./bin/hdfs dfsadmin - report //查看集群是否组建成功
09. Live datanodes ( 3): //有三个角色成功
```

[Top](#)

[Top](#)