

TTS 10.0 COOKBOOK

(NSD ARCHITECTURE DAY05)

版本编号 10.0

2018-08

达内 IT 培训集团

NSD ARCHITECTURE DAY05

1. 案例 1：安装 Hadoop

• 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

• 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.
[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.e17.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
//wordcount 为参数 统计 aa 这个文件夹，存到 bb 这个文件里面（这个文件不能存在，要是存在会
报错，是为了防止数据覆盖）
[root@nn01 hadoop]# cat bb/part-r-00000 //查看
```

2. 案例 2：安装配置 Hadoop

• 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

• 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

• 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
```

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:Uc180Cezw92aArY5+zPtOrJ9o11ojRE3EAZ1mgndYQM root@nn01
The key's randomart image is:
+---[RSA 2048]-----+
|          o*E*=.      |
```

```
+XB+.
..=0o.
o.+o...
.S+.. o
+ . =o
o+oo
o+=.o
o==0.
+----[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
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
```

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
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
```