对主机、用户、目录、软件的规划

## 1. 对主机、用户、目录、软件的规划

### 1.1. 主机规划

|  |  |  |  |
| --- | --- | --- | --- |
|  | djt11/192.168.137.21 | djt12/192.168.137.22 | djt13/192.168.137.23 |
| namenode | 是 | 是 | 否 |
| datanode | 是 | 是 | 是 |
| resourcemanager | 是 | 是 | 否 |
| journalnode | 是 | 是 | 是 |
| zookeeper | 是 | 是 | 是 |

### 1.2. 软件规划

|  |  |  |  |
| --- | --- | --- | --- |
| 软件 | 版本 | 位数 | 说明 |
| jdk | java version "1.7.0\_45" | 64-Bit |  |
| centos | centos7 | 64-Bit |  |
| zookeeper | zookeeper-3.4.8.tar.gz |  | 稳定版 |
| hadoop | hadoop-2.6.5.tar.gz |  | 稳定版 |

### 1.3. 用户规划

|  |  |  |
| --- | --- | --- |
| 节点名称 | 用户组 | 用户 |
| djt11 | hadoop | hadoop |
| djt12 | hadoop | hadoop |
| djt13 | hadoop | hadoop |

### 1.4. 目录规划

|  |  |
| --- | --- |
| 名称 | 路径 |
| 所有软件目录 | /home/hadoop/app/ |
| 所有数据和日志目录 | /home/hadoop/data/ |

## 2. 集群安装前的环境检查，比如时钟同步、hosts文件检查、禁用防火墙等

1. 时钟同步

* [root@djt11 hadoop]# date  
  Wed Sep 13 04:54:00 EDT 2017

1. hosts

* [hadoop@djt11 hadoop]$ cat /etc/hosts  
  192.168.137.21 djt11  
  192.168.137.22 djt12  
  192.168.137.23 djt13

1. 禁用防火墙

* [hadoop@djt11 hadoop]$ systemctl status firewall  
  firewall.service  
   Loaded: not-found (Reason: No such file or directory)  
   Active: inactive (dead)

## 3. Ansible 批量搭建hadoop环境

1. 在一台机器上指定软件目录

* [root@slave1 tar]# ls  
  apache-hive-1.2.1-bin-spark1.5.tar.gz jdk-7u45-linux-x64.tar.gz  
  bigdata.sh spark-1.5.3-SNAPSHOT-bin-hadoop2-without-hive.tgz  
  hadoop-2.6.5.tar.gz spark-1.6.1-bin-hadoop2.6.tar.gz  
  hadoop-conf zoo.cfg  
  jdk1.7.0\_71.tar zookeeper-3.4.8.tar.gz

1. ansible bigdata 配置文件目录

* [root@slave1 ansible]# tree bigdata/  
  bigdata/  
  ├── bigdata.yml  
  ├── group\_vars  
  │   └── bigdata  
  ├── hadoop-conf  
  │   ├── core-site.xml  
  │   ├── hdfs-site.xml  
  │   ├── mapred-site.xml  
  │   ├── slaves  
  │   └── yarn-site.xml  
  ├── hosts  
  └── roles  
   ├── bigdata  
   │   ├── handlers  
   │   │   └── main.yml  
   │   └── tasks  
   │   └── main.yml  
   ├── bigdata-conf  
   │   └── tasks  
   │   └── main.yml  
   └── bigdata-delete  
   └── tasks  
   └── main.yml

1. 执行命令安装

* ansible-playbook -i hosts bigdata.yml

## 4. 运行 Hadoop 环境

1. djt11

* [hadoop@djt11 hadoop]$ jps  
  14831 DataNode  
  3082 QuorumPeerMain  
  15951 ResourceManager  
  14732 NameNode  
  15018 JournalNode  
  15483 NodeManager  
  15286 DFSZKFailoverController

1. djt12

* [hadoop@djt12 ~]$ jps  
  12882 NodeManager  
  13034 ResourceManager  
  12766 DFSZKFailoverController  
  14432 Jps  
  12461 DataNode  
  3071 QuorumPeerMain  
  12392 NameNode  
  12556 JournalNode

1. djt13

* [hadoop@djt13 ~]$ jps  
  3064 QuorumPeerMain  
  10726 Jps  
  9490 NodeManager  
  9281 DataNode  
  9373 JournalNode

## 运行 wordcount

