

# 第2章 Hadoop集群的搭建及配置实训

---

(VMware上安装Hadoop3.x)

(源自:<https://biglab.site>)

(版本:Ver2.0-20231129)

## 第2章 Hadoop集群的搭建及配置实训

CH02实训1搭建Hadoop完成分布式集群

实训要点

需求说明

实现思路及步骤

作业要求

## CH02实训1搭建Hadoop完成分布式集群

---

### 实训要点

1. 掌握虚拟机的安装及配置方法
2. 掌握在Linux下JDK的安装方法
3. 掌握Hadoop完全分布式集群环境的搭建过程
4. 掌握Hadoop集群的监控方法

### 需求说明

为了贴近真实的生产环境，需搭建完成分布式模式的Hadoop集群环境，本实验要求在个人计算机安装及配置虚拟机，并在虚拟机中搭建Hadoop完全分布式集群环境。

### 实现思路及步骤

1. 安装VMware Workstation(简称VMware)
2. 在VMware上安装Linux虚拟机
3. 进入Linux虚拟修改IP地址为固定IP
4. 使用crt或xshell远程连接虚拟机，在上面安装Hadoop、安装JDK1.8
5. 使用crt或xshell远程连接虚拟机，修改Hadoop配置文件，如core-site.xml、hadoop-env.sh、yarn-env.sh、mapred-site.xml、yarn-site.xml、hdfs-site.xml、workers、start-dfs.sh、stop-dfs.sh、start-yarn.sh、stop-yarn.sh
6. 使用crt或xshell远程连接虚拟机，修改系统的/etc/hosts、/etc/profile文件
7. 完成上述mater虚拟机后，从master虚拟机克隆出slave1和slave2虚拟机，并修改slave1和slave2的IP地址和hostname
8. 配置虚拟机时间同步,master虚拟机向北京时间服务器同步，slave1和slave2向master同步
9. 启动hadoop集群，使用start-all.sh启动集群
10. 使用crt或xshell远程连接虚拟机，使用jps命令观察master和slave1,slave2进程，并截图到作业中
11. 修改windows的host文件：c:\windows\system32\drivers\etc\hosts，添加master,slave1,slave2以及对应的IP
12. 在浏览器上观察<http://master:9870>站点和<http://master:8088>站点信息并截图到作业中

## 作业要求

1. 请截取Hadoop集群的拓扑结构图,拓扑图上需标识虚拟机名、IP地址、jps进程名。
2. 请截取本集群master、slave1、slave2中jps进程截图
3. 请截取本集群中<http://master:9870>站点的datanode列表信息截图
4. 请截图本集群中<http://master:8088>站点的nodes列表信息截图