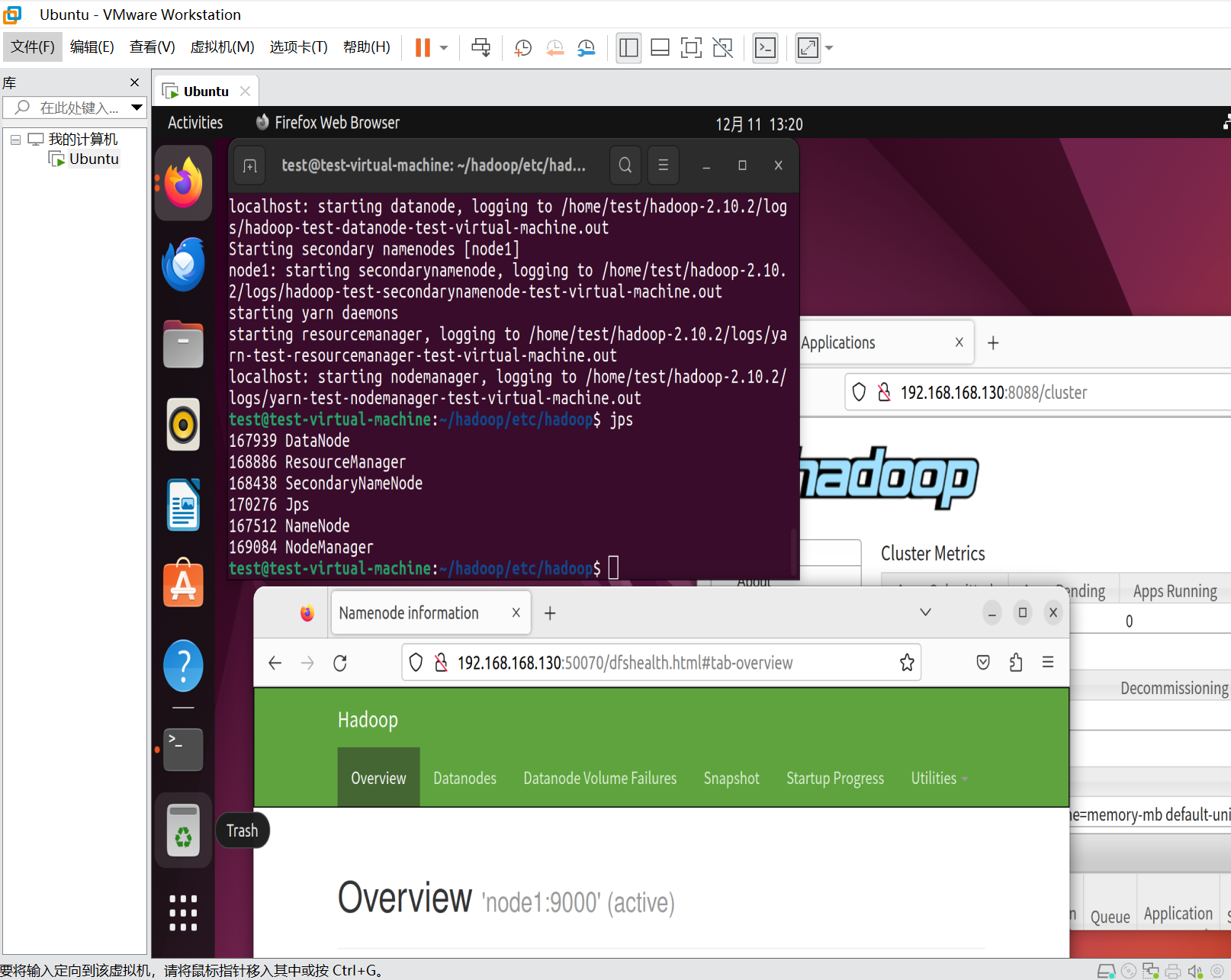
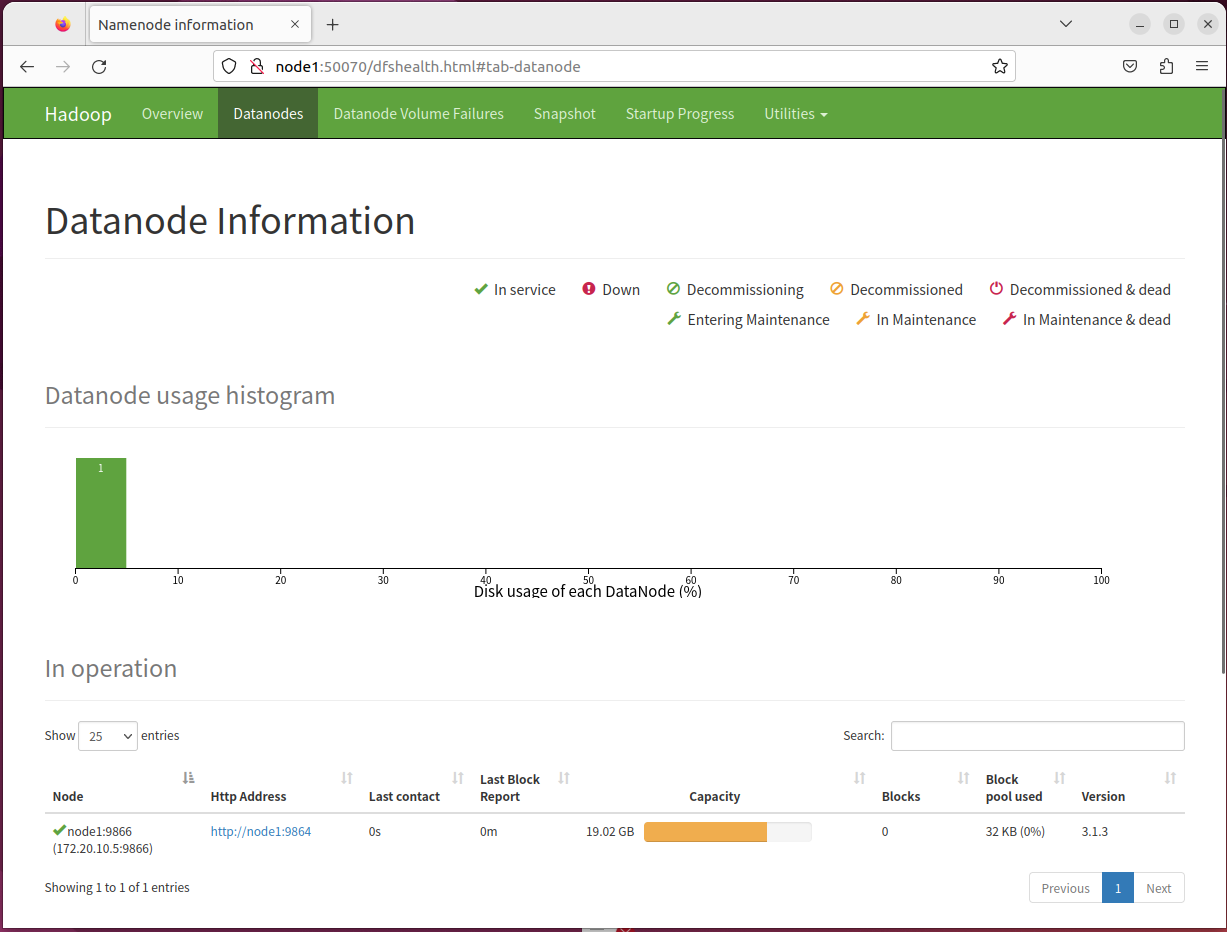
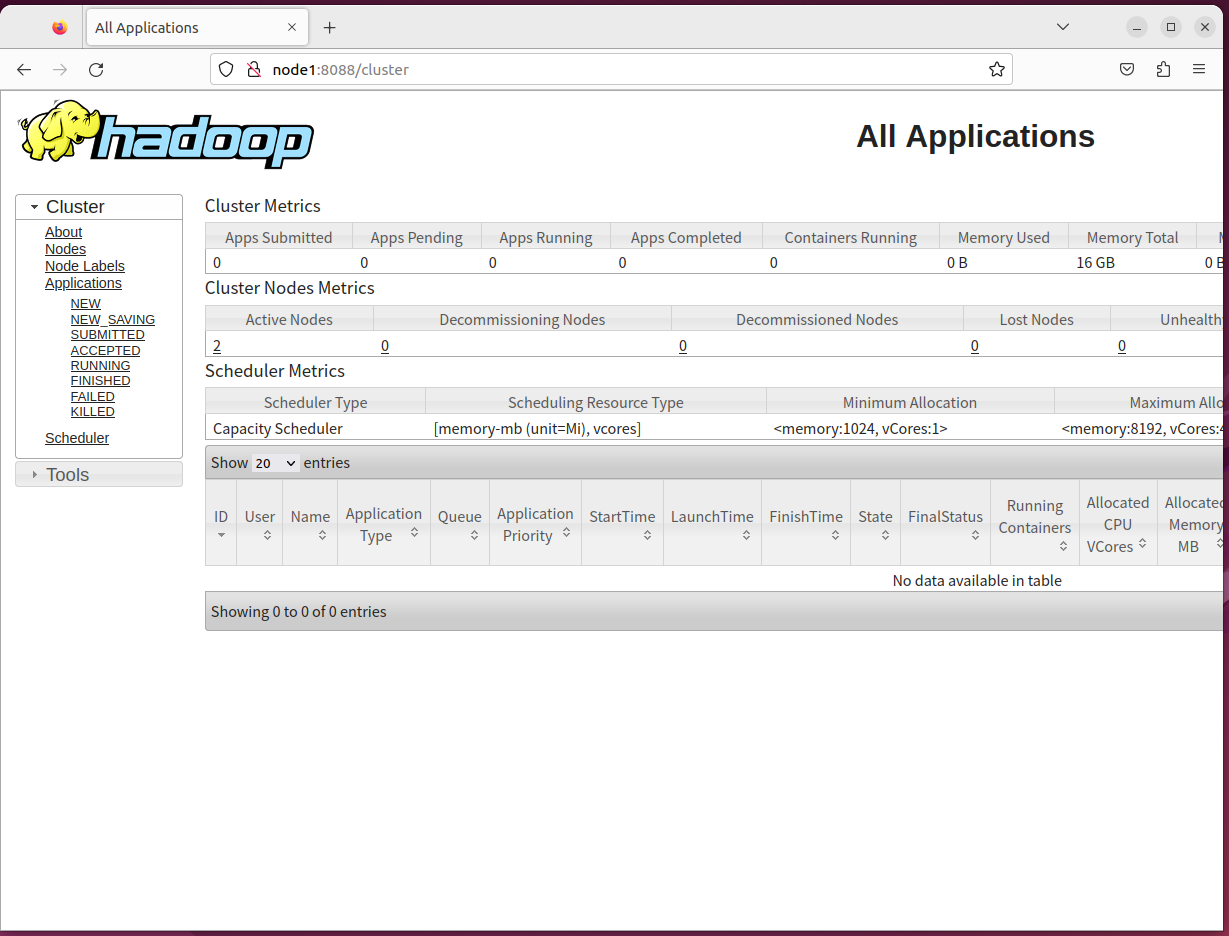
单机开发环境部署

实验结果图：

（自己电脑配置单机成功）







一、实验目标

本实验旨在单机环境下部署Hadoop伪分布式环境，通过实践操作深入了解Hadoop的架构、安装与配置过程，并验证其运行效果，熟练掌握linux的命令。

二、实验原理

Hadoop是一个分布式计算框架，包括HDFS（Hadoop Distributed File System）和MapReduce。HDFS为海量数据提供了稳定的存储环境，而MapReduce则为数据处理提供了计算能力。伪分布式环境是单机下的模拟分布式环境，用于开发和测试。

三、实验步骤

1.安装linux

2.关闭防火墙

3.安装jdk

4.安装hadoop

5.安装ssh-server

6.配置环境变量，修改文件配置

7.配置主机名，免登录

8.格式化，启动

9.测试

四、遇到的问题

1.在配置文件时，配置内容有误导致启动不成功

2.未找到指定的路径导致找不到相应的配置文件和启动文件

3.漏输sudo的服务命令Sudo service ssh restart，导致一些服务未能启用

4.hostname未被更改

5.JAVA\_HOME的路径有无找不到jdk的文件

6.Datanode的tmp需删除才能重新启动

五、实验结果与分析

在实验过程中，我们成功地在单机环境下部署了Hadoop伪分布式环境，并验证了其运行效果。通过浏览器访问Web界面，我们能够查看HDFS和MapReduce的状态和运行信息。这为我们后续进行数据分析和处理提供了稳定的基础。

Hadoop集群部署

1.实验目的：掌握多机Hadoop分布式运行环境

2.实验原理：Hadoop集群的搭建，分布式文件系统HDFS提供了基础存储支持，需要存储的数据被HDFS切割成块分布到集群环境当中进行存储，具有高容错、高可靠性、高可扩展性、高获得性、高吞吐率。

3.实验遇到的问题：

需要版本一致

需要文件路径相应的做出改动

将需要修改ip地址和ip的名

须在同一网段下面

设置主次机

修改hdfs里的机子的个数