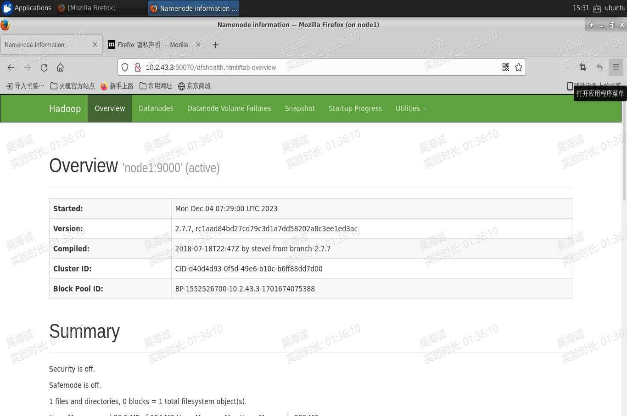
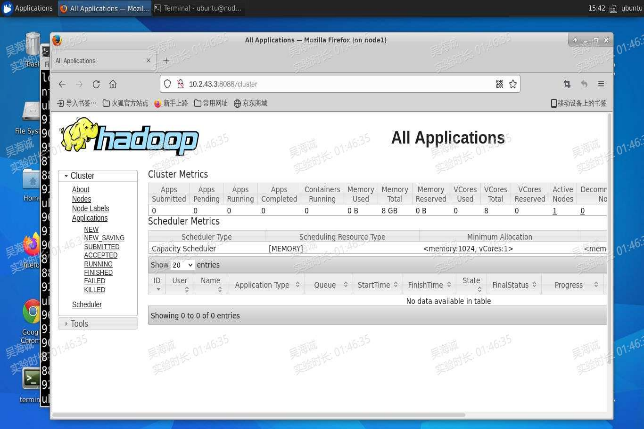
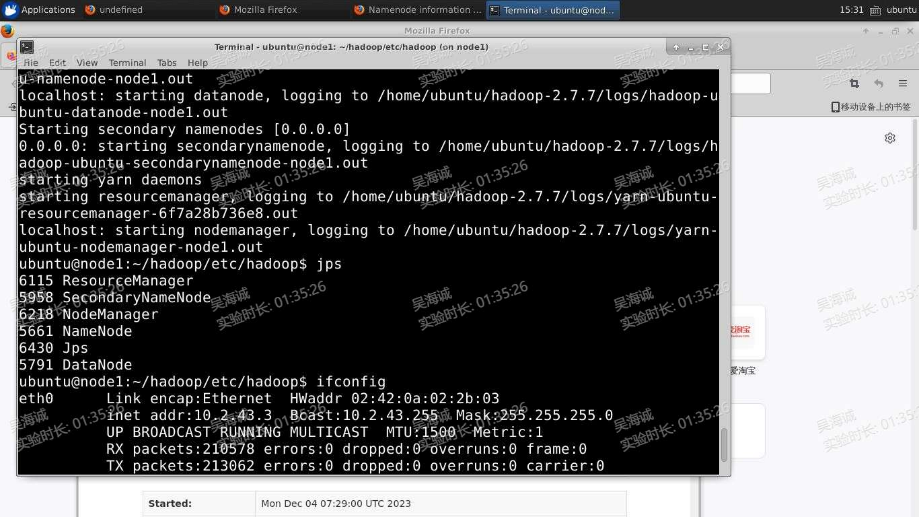
Hadoop部署

首先部署结果

有两个部署结果，一是曙光平台的保姆式教学部署结果截图如下：



关于过程中遇到的问题：在修改配置的步骤时，

文本

描述已自动生成

保姆式教学的mapreduce-shuffle需改为mapreduce\_shuffle，否则会使得nodemanager进程在start-all.sh后无法持续开启，同时在设置主机名时，

sudo hostname 主机名 后 不仅要修改/etc/hosts中的IP地址及主机名，同时还要修改/etc/hostname文件中的文件名。

Ssh设置登录时 若ssh免密登录的设置，执行sudo service ssh restart命令后方可成功登录。

其次，不仅在曙光的平台上成功部署单机模式，我在自己的机器上也成功部署，以下是证明截图：

图形用户界面

描述已自动生成

使用的平台是Ubuntu 22.4.3TLS 64位,在部署自己电脑上的环境时遇到了较多的问题，但是主要过程并无区别，下载JDK，Hadoop，解压，配置环境变量，此时遇到第一个问题，默认教程中的环境变量代码是/home/ubuntu/jdk\hadoop,需要结合自己的机器环境修改则我的环境配置代码为：/home/root1/jdk （root1是我的主机名），由于主机名不同在，配置环境时 比如 hadoop-env.sh,core-site.xml此类文件中的相关路径也需要修改，这些都是小地方，遇到的主要问题有两个。

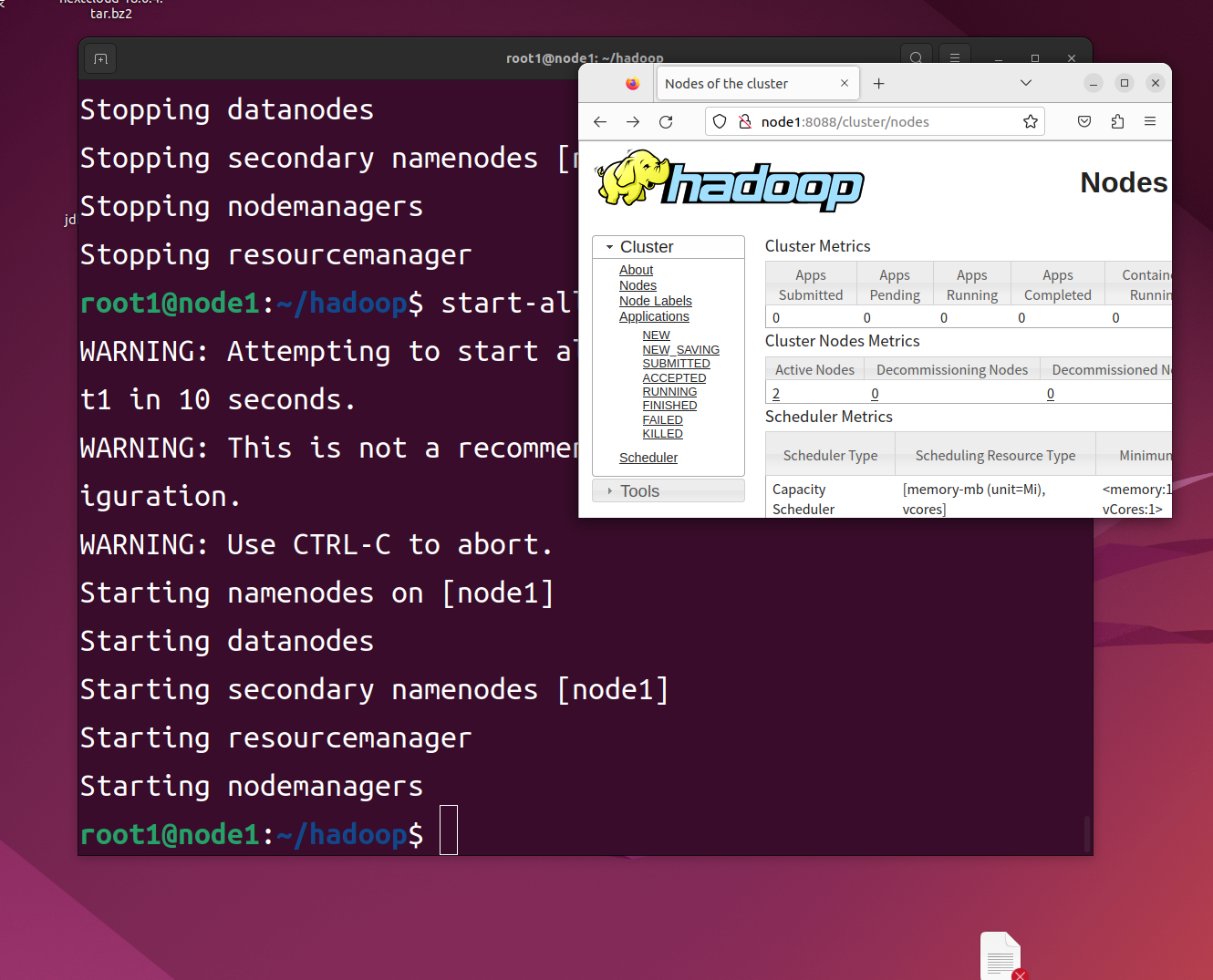
其一：在多次的 hdfs namenode -format 格式化后，我从可以jps出6个变成了5个，少了一个datanode，由此，我查阅相关资料，发现他们遇到这种情况是由于多次格式化导致的/tmp/dfs/name/current/VERSION中的namespace和clusterID与/tmp/dfs/data/current/VERSION中的namespace和clusterID不一样导致，经过我的修改及重启之后并未解决此问题，之后，我把tmp全删了，然后重新格式化hdfs之后就顺利解决。

其二：在成功出现6个之后，我可以在浏览器中，网址栏中输入[ip:8088] 后可以进入网站，但是输入[ip:50070]后不能进入网站，起初我认为是ssh登录的问题，然后疯狂的restart ssh，之后我发现，csdn上说是防火墙的问题，随后我通过重复关闭防火墙和禁用防火墙来进入这个网址，但是还是不行，最后我在/hadoop/etc/hadoop目录下的hdfs-site.xml文件中添加了

文本

描述已自动生成

四行代码，至此我成功在我的机器上配置了单机的hadoop环境。



关于多机部署，以上是截图证明，为两台虚拟机直接的连接，过程中同样遇到了许多问题，其一，在ssh登录的配置的时候，开始ifconfig出的ip地址并没有用，因为当时处于VMW的NET模式下，ip地址无法用于连接，修改为桥接模式后，由于学校的网无法满足实验要求，在热点的环境下，host机的ip地址及名称是172.20.10.5，另一台机器的ip地址及名称为172.20.10.10，在/etc/hosts中同时添加对方的ip及主机名，后ssh-keygen经过成功ping成功后，互相ssh-copy-id对方后，即完成了第一步。

但同时遇到了第二个问题，并不能在两台机器上免密登录，究其原因是因为用户名不同导致，并且我发现，两台电脑中的hadoop的版本不一样，host主机的版本是3.1.3，另一台主机的版本是2.10.1，那么如何解决呢？我一开始认为重新在非host的机器上创建一个名称相同的用户之后把hadoop复制过去，但怕会有问题出现，后面选择直接克隆了一遍host的虚拟机，之后，重新配置了hosts及hostname，使其正确的在两台虚拟机上，此时，修改/hadoop/etc/hadoop/中的hdfs-site.xml文件中的：

文本

中度可信度描述已自动生成

之后在/hadoop/etc/hadoop下新建workers文件（3.0版本后的名称，之前好像是slavers）,据说是在3.0版本后改成这样的（政治正确xD ）,不知道是否是这个原因，导致环境配置之后启动dfs之后node2上确实被拉起了DataNode，同时在所有6个进程都启动后，在网站中输入node1:8088能明确看到两个ip地址访问，但是在node1:50070中只能看到1个IP地址，但是是node2就此，同时经老师的提醒认为可能是在workers中，需要把node1修改为localhost即: 文本

描述已自动生成之后并未对上述的现象产生任何变化。就此完成了两台虚拟机在不同主机上的多机部署。