

## 实验二： Hadoop 的安装和使用

### 实验基本信息：

课程：大规模分布式系统综合实践 实验时间： 2024 年 9 月 9 日

实验类型： ☐验证性 ☒设计性 ☐综合性

专业班级：计算机科学与技术专业（中外合作）3，4 班

姓名： 张乐之 学号： 222022321102120

实验指导： 邱开金

### 实验报告提交说明：

本次实验需要撰写实验报告，实验报告填写时以完成实验任务为目的，先简要回答完成实验任务需要的步骤或需要执行的命令代码，再配以结果截图予以说明，图片数量不宜过多，能说明问题即可。最后配上总结，并提交到教师指定位置。

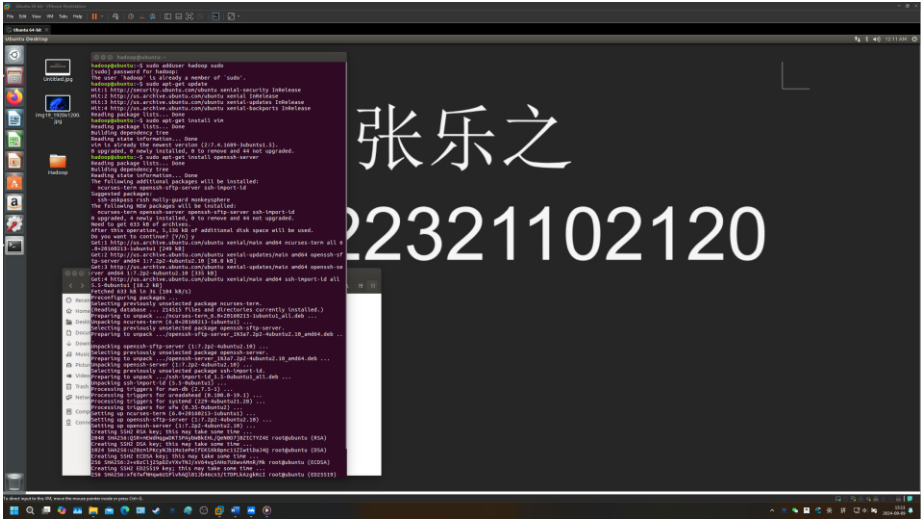
### 实验目的：

1. 掌握在 Linux 虚拟机中安装 Hadoop 的方法。
2. 熟悉 HDFS 的基本使用方法。

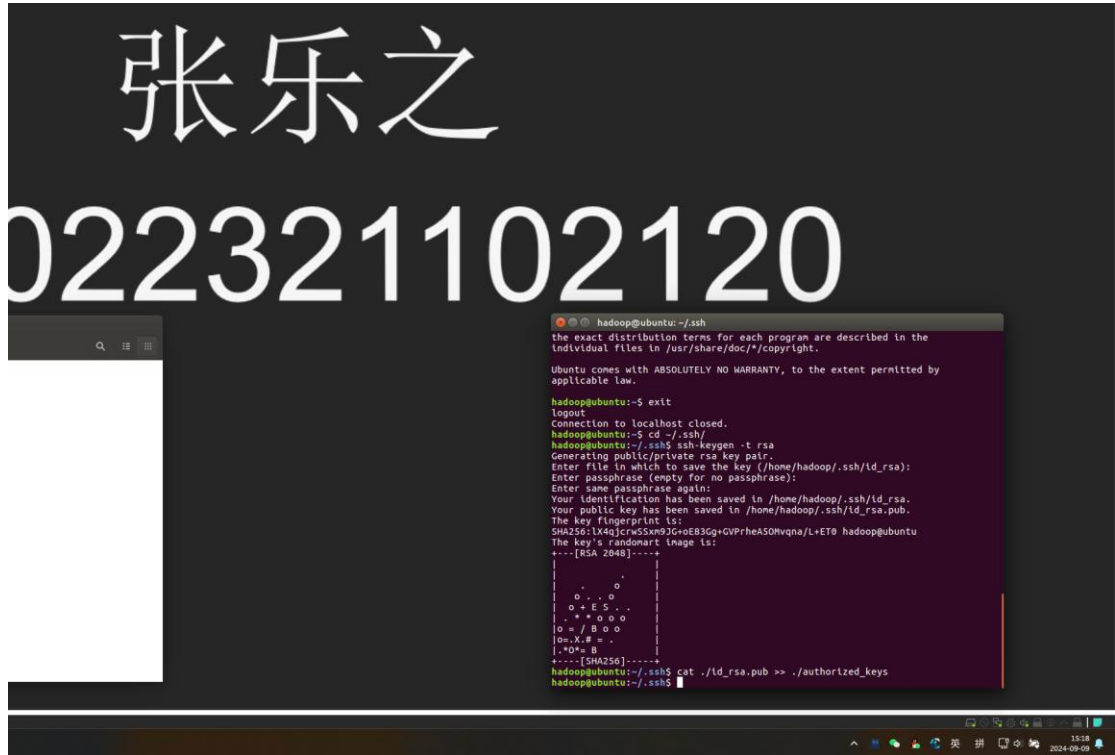
### 实验任务：

#### 1. 安装 Hadoop

准备环境，更新或安装必要软件：



远程连接，加入 SSH 授权密钥：



解压 JVM:



配置环境变量:

# 张乐之

## 222022321102120

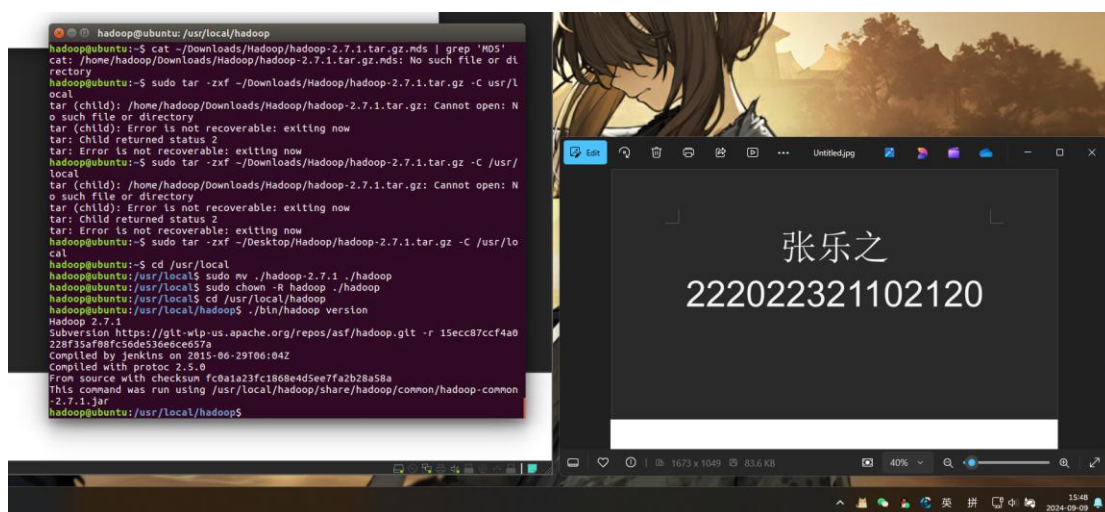
jvm 安装成功:

# 张乐之

# 2222321102120

The image displays a Linux desktop environment. The top section features a large white title "张乐之" (Zhang Lezhi) and a large white number "2222321102120". Below this, there are two screenshots of a desktop. The left screenshot shows a file manager window titled "Desktop" containing numerous software icons such as IntelliJ IDEA, Kettle, Apache Hive, Apache Tomcat, Hadoop, Hbase, Jdk-8u162-linux-x64.tar.gz, Kafka, MySQL Server, R, Redis, Sbt-launch.jar, Scala, Scala Intellij Bin, Scala SDK, Spark, Spark Streaming, and XerxCES.dll. The right screenshot shows a terminal window with a pink header bar labeled "root@hadoopgubuntu:~#". The terminal contains commands and their outputs, including setting environment variables like JAVA\_HOME, PATH, and CLASSPATH, and running tests for Java and Scala.

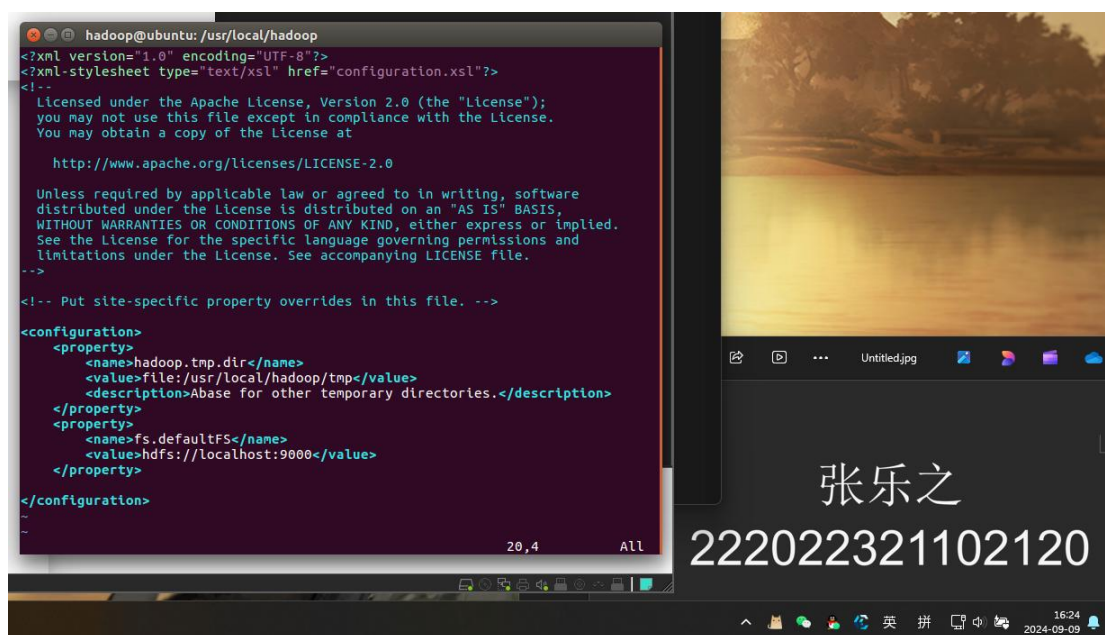
## 解压并安装 hadoop:



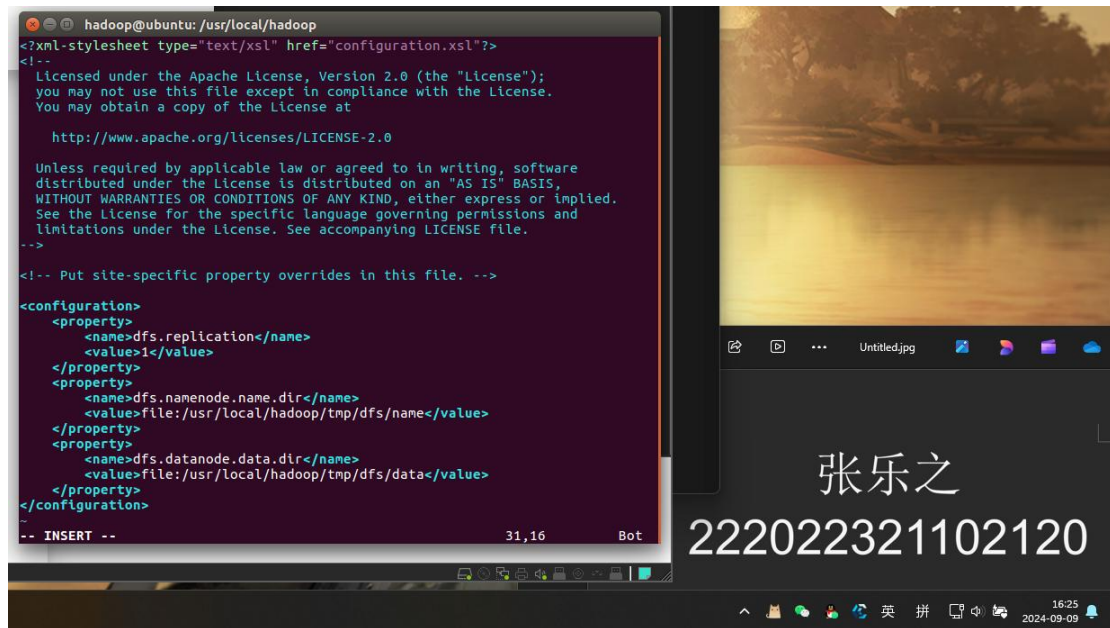
## 2. HDFS 常用操作（使用 Hadoop 提供的 Shell 命令完成如下操作）

1) 启动 Hadoop，在 HDFS 中创建用户目录“/user/hadoop”。

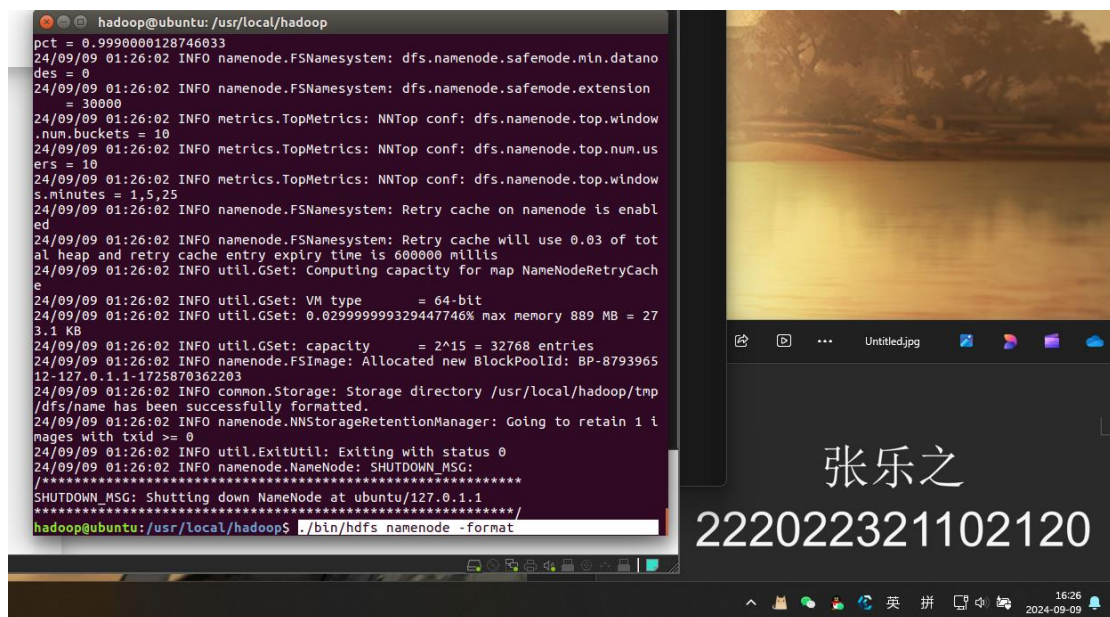
配置分布式节点：



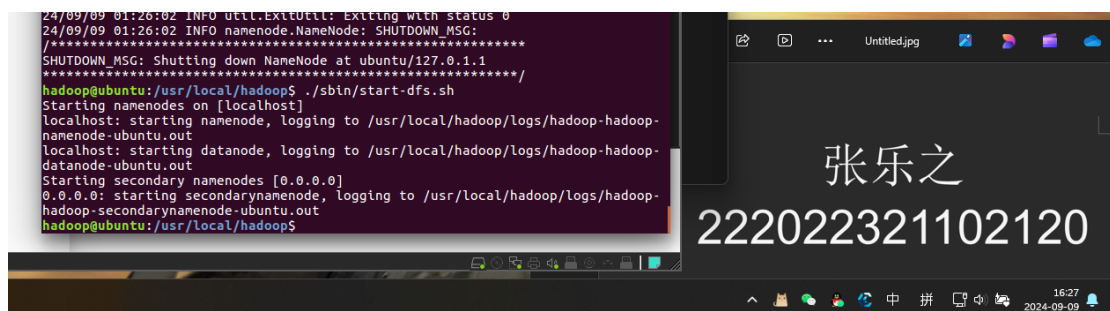


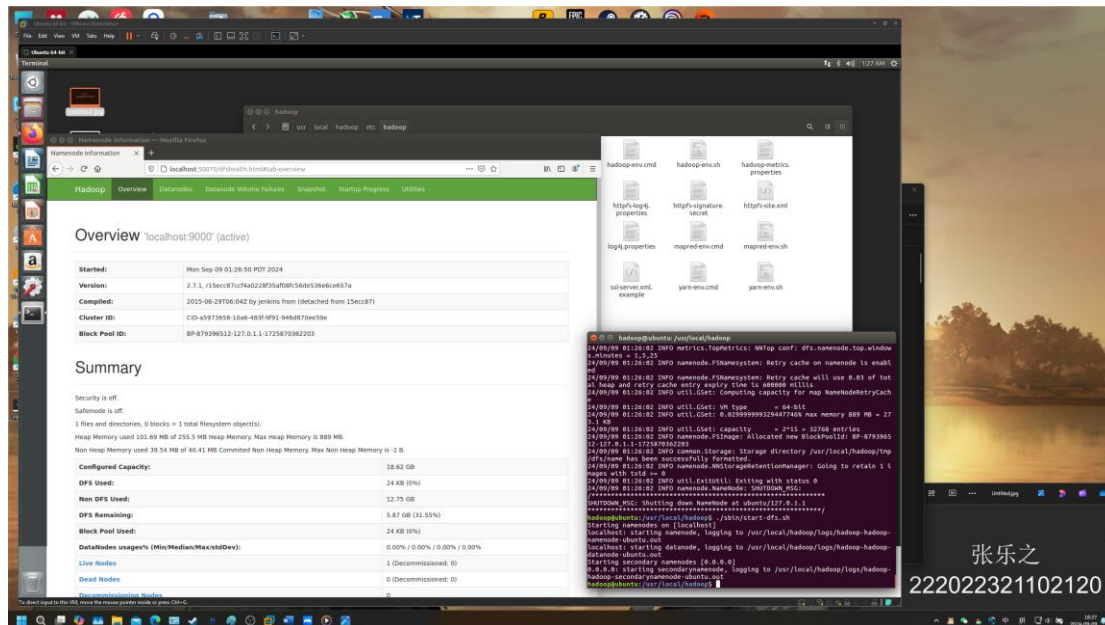


初始化: `./bin/hdfs namenode -format`

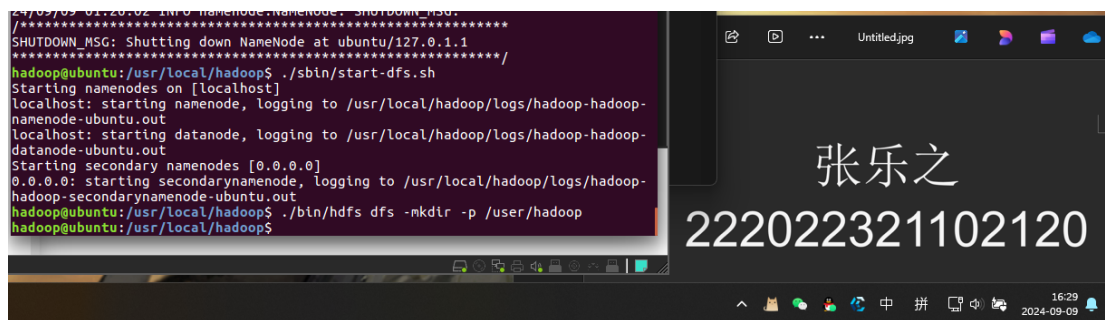


启动: `./sbin/start-dfs.sh`





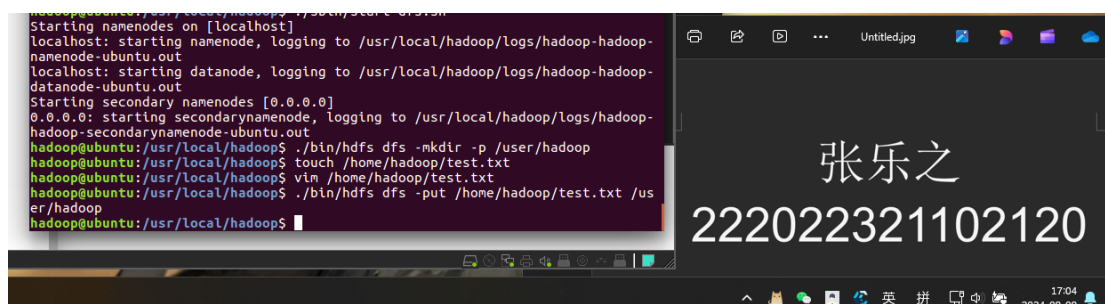
创建文件夹

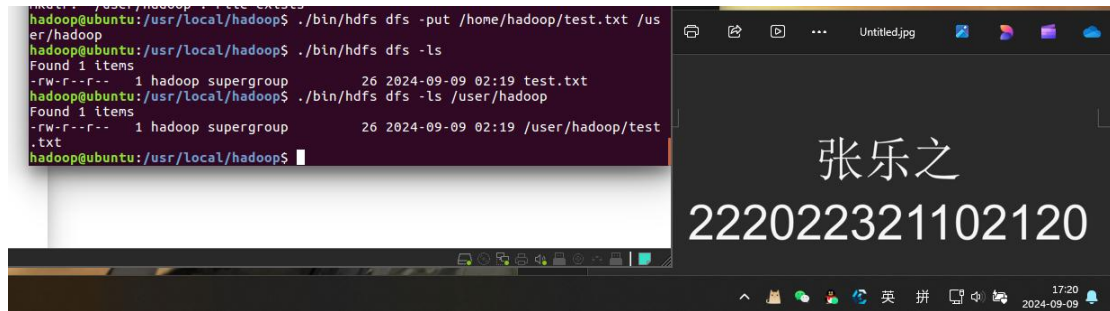


（作答要求：在每一个问题后写上所需要的命令，如果命令已经能够清楚回答问题，则不需要抓图。在所有命令执行完毕后，在终端输入 `history` 命令，将该命令执行结果抓图（抓图应能看清命令以及学生所用的电脑的当前时间或当前桌面背景等能区分出是自己做的证据信息）附在最后）

2) 在 Linux 系统的本地文件系统的“/home/hadoop”目录下新建一个文本文件 `test.txt`，并在该文件中输入自己的学号姓名，然后上传到 HDFS 的“/user/hadoop”目录下。

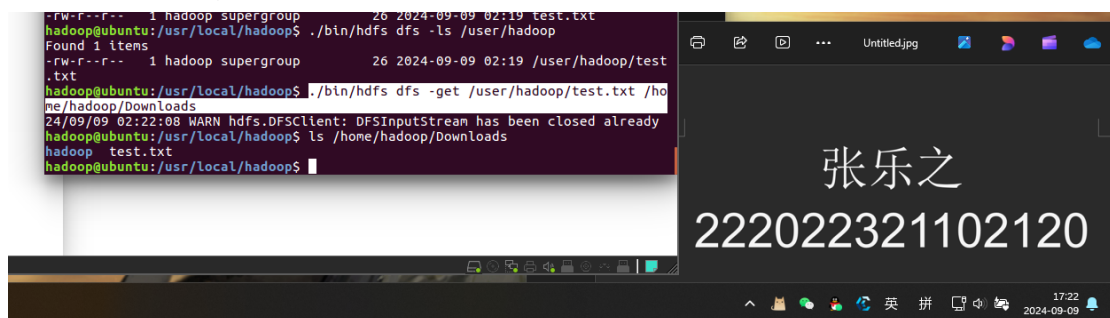
```
touch /home/hadoop/test.txt
vim /home/hadoop/test.txt
./bin/hdfs dfs -put /home/hadoop/test.txt /user/hadoop
```





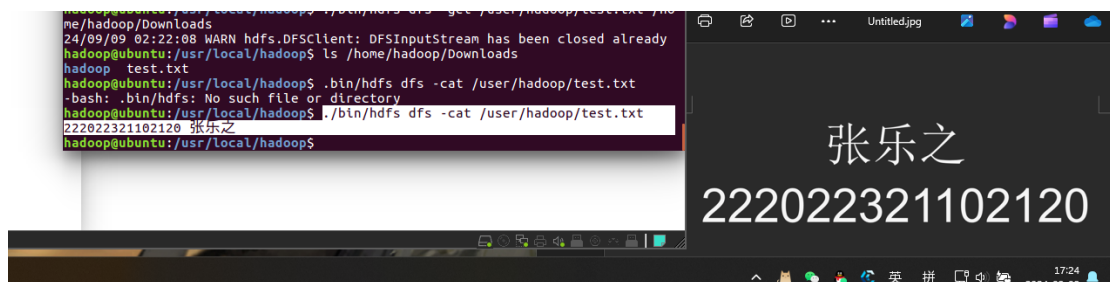
3) 把 HDFS 中“/user/hadoop”目录下的 `test.txt` 文件，下载到 Linux 系统的本地文件系统中的“/home/hadoop/下载（英文版为 Downloads）”目录下。

`./bin/hdfs dfs -get /user/hadoop/test.txt /home/hadoop/Downloads`



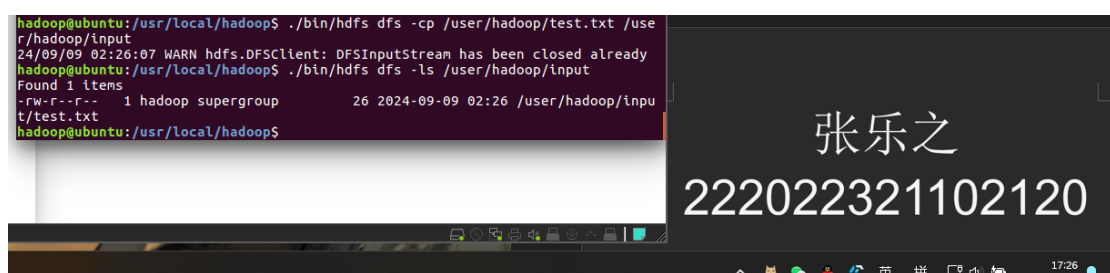
4) 将 HDFS 中“/user/hadoop”目录下的 `test.txt` 文件的内容输出到终端中进行显示。

`./bin/hdfs dfs -cat /user/hadoop/test.txt`



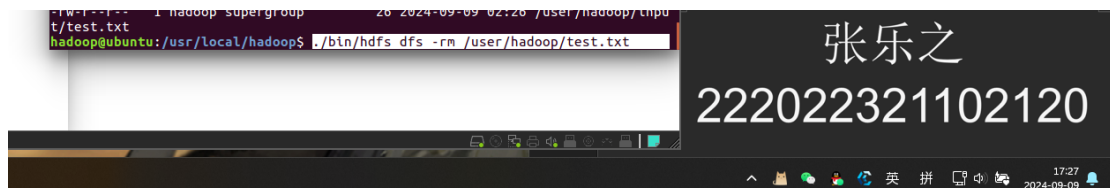
5) 在 HDFS 中的“/user/hadoop”目录下，创建子目录 `input`，把 HDFS 中“/user/hadoop”目录下的 `test.txt` 文件，复制到“/user/hadoop/input”目录下。

`./bin/hdfs dfs -cp /user/hadoop/test.txt /user/hadoop/input`



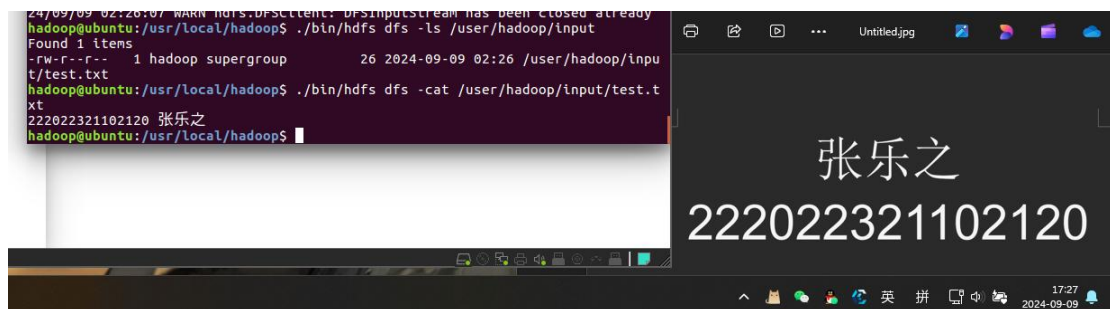
6) 模拟删除 HDFS 中“/user/hadoop”目录下的 test.txt 文件（输入这条删除命令，但并不执行）。

```
./bin/hdfs dfs -rm /user/hadoop/test.txt
```



7) 在终端输出 HDFS 中“/user/Hadoop/input/test.txt”文件中的内容（此题除写出命令外，还需要写出输出的内容是什么）。

```
./bin/hdfs dfs -cat /user/hadoop/input/test.txt
```



## 参考资料及提示：

1. 在 ubuntu16.04 的中安装 hadoop  
可参考网址：

<http://dblab.xmu.edu.cn/blog/install-hadoop/>

2. Hadoop 中 hdfs 常见 shell 命令可参考

<http://dblab.xmu.edu.cn/blog/1625-2/>