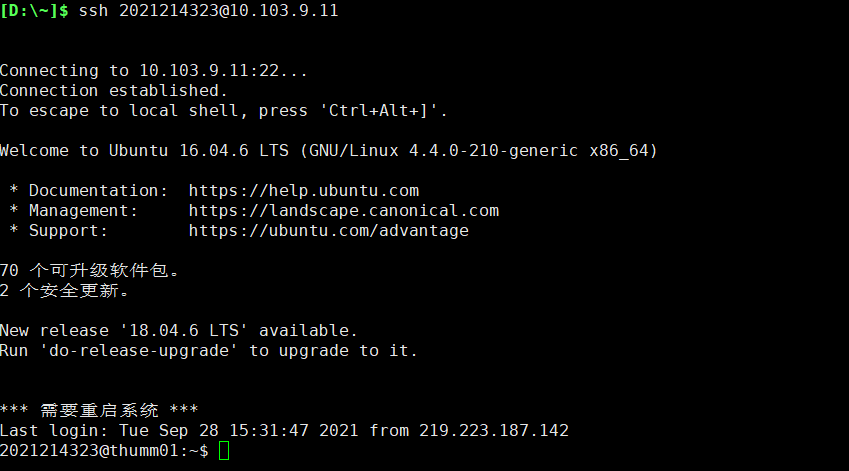
任务一：



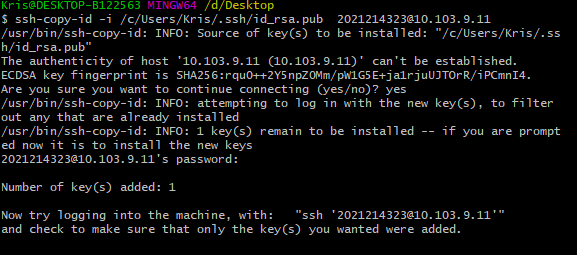
输入指令，登录成功

任务二：

在个人电脑上生成公钥和私钥

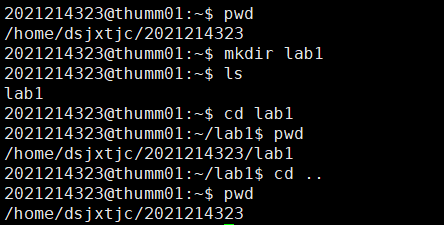


将公钥内容复制到服务器

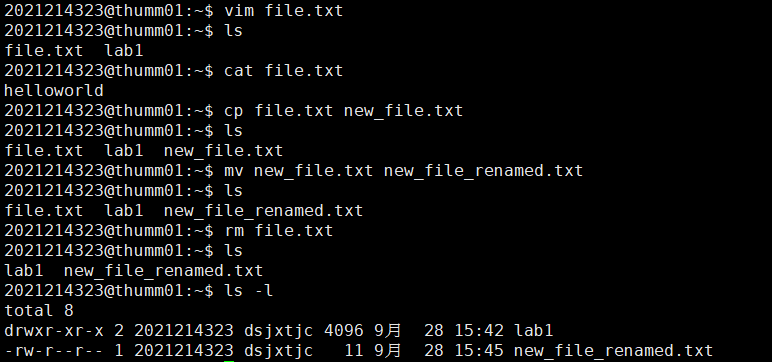


测试已可免密登录

任务三：pwd mkdir cd等命令的使用

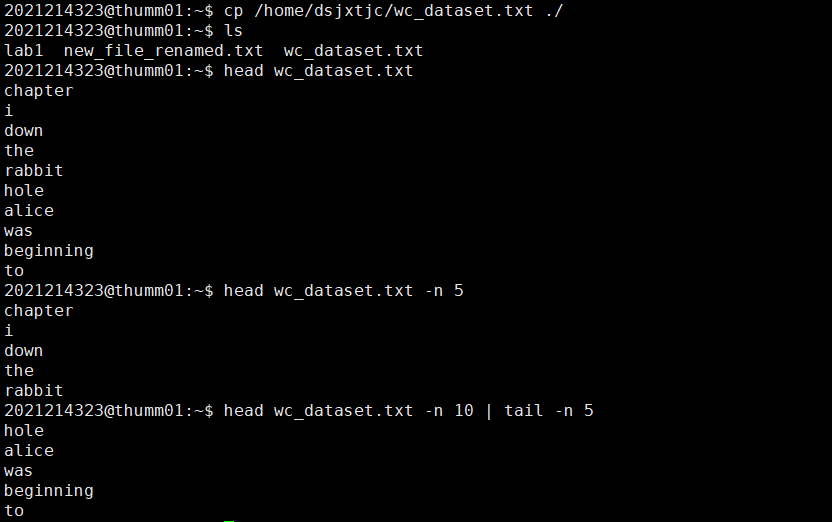


任务四 ：cp vim ls mv rm 等指令的使用



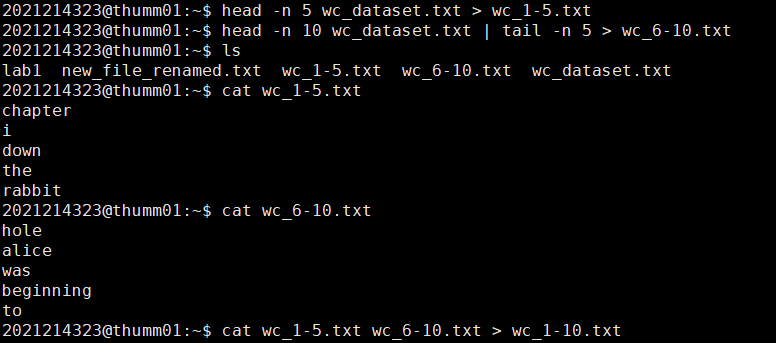
任务五

将数据集复制后用head和tail查看

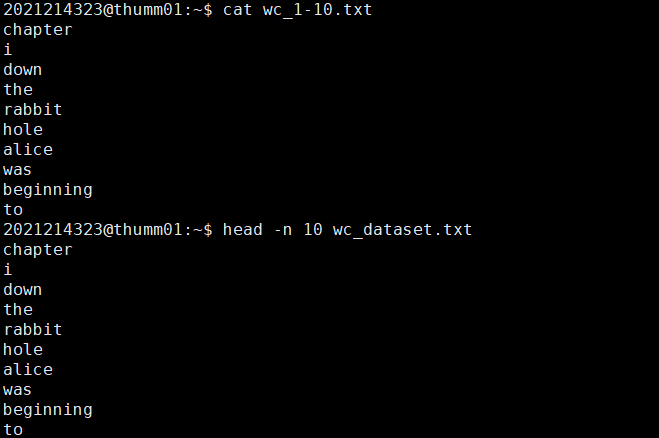


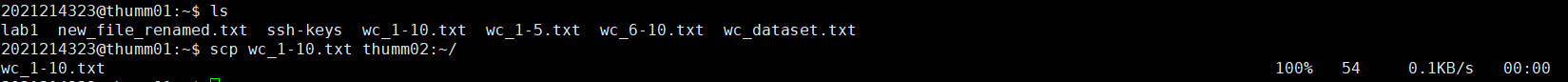
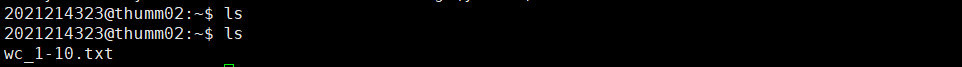
重定向符 ‘>’ 的使用 并使用cat查看结果

结合cat和重定向符 合并文件

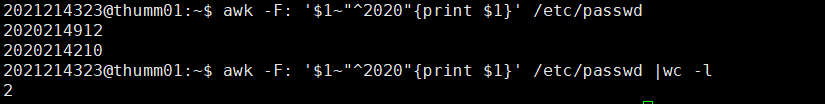


查看合并结果符合预期



使用scp在两个主机间传文件

使用awk查看2020级的学生 以及人数



使用awk查看2021级的学生 以及人数





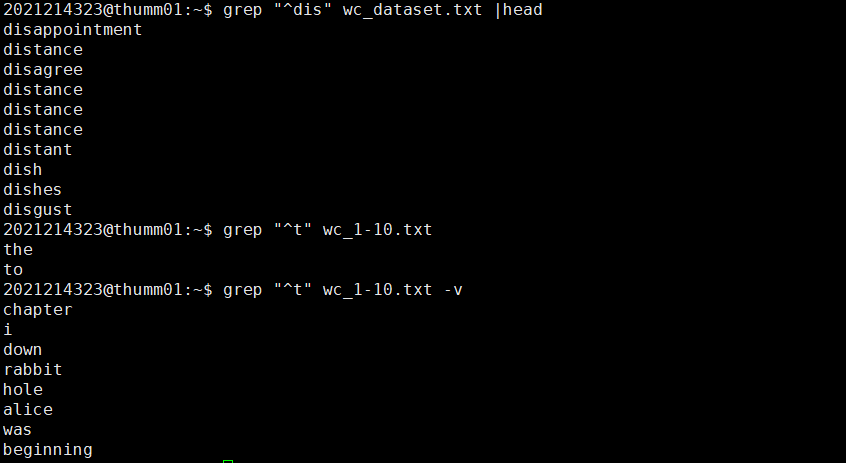
显示1000~2000行中所有以"dis"开头的单词（显示前10条）



显示所有以"dis"开头的单词（显示前10条）

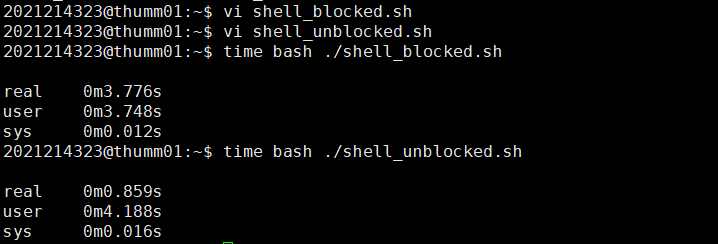
查找wc\_1-10.txt中以t字母开头的单词

接着添加参数-v，过滤以t字母开头的单词



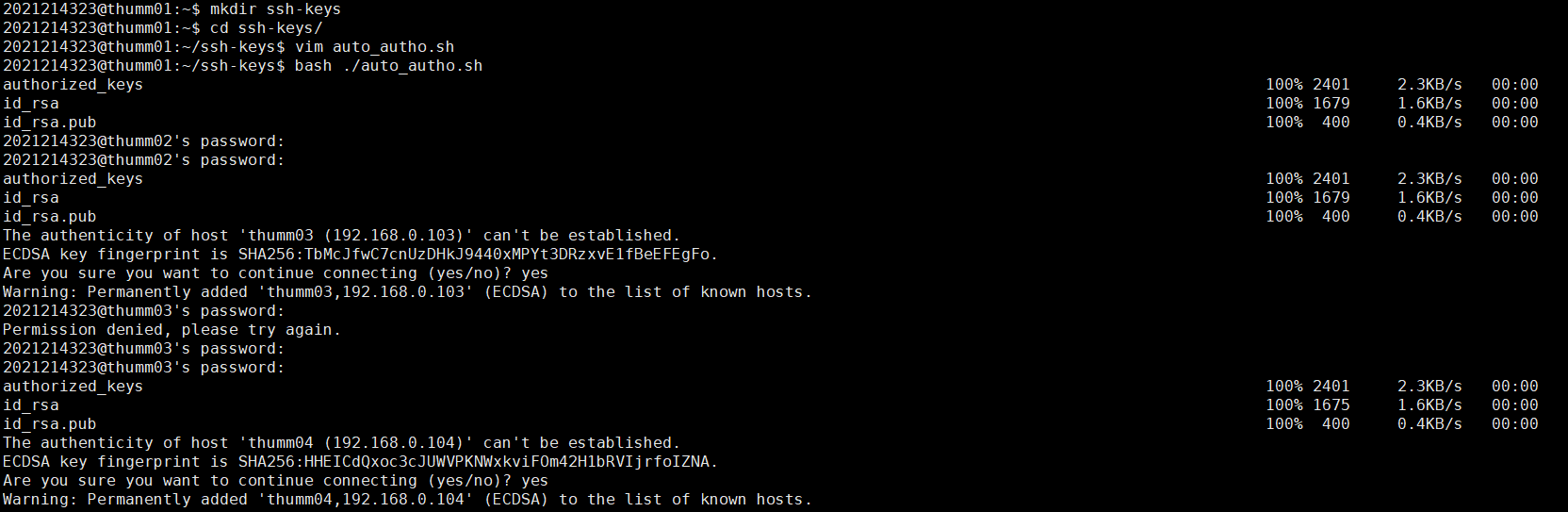
任务六 阻塞与非阻塞对比

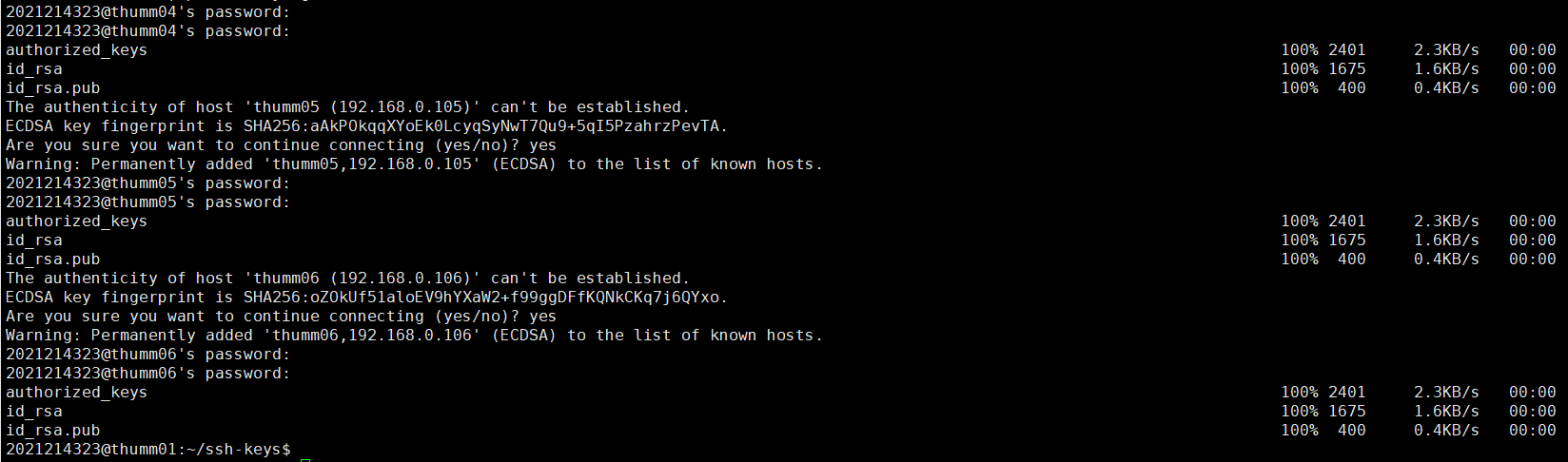
用户时间相差较少，非阻塞真实时间大幅减少



任务七

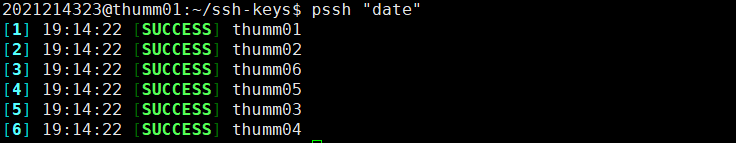
7.1 集群之间免密登录配置



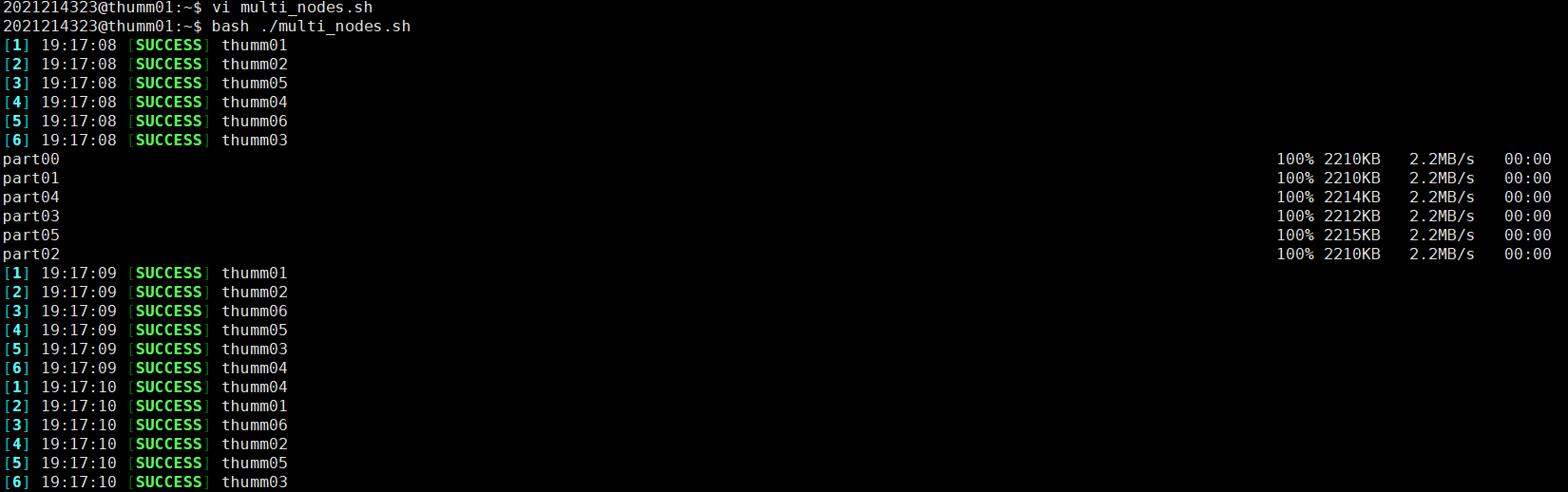


7.2集群间批管理

使用pssh指令



7.3在多主机上并行执行任务



任务8

改变split方法，将-l参数改为-C，将文件按字节划分

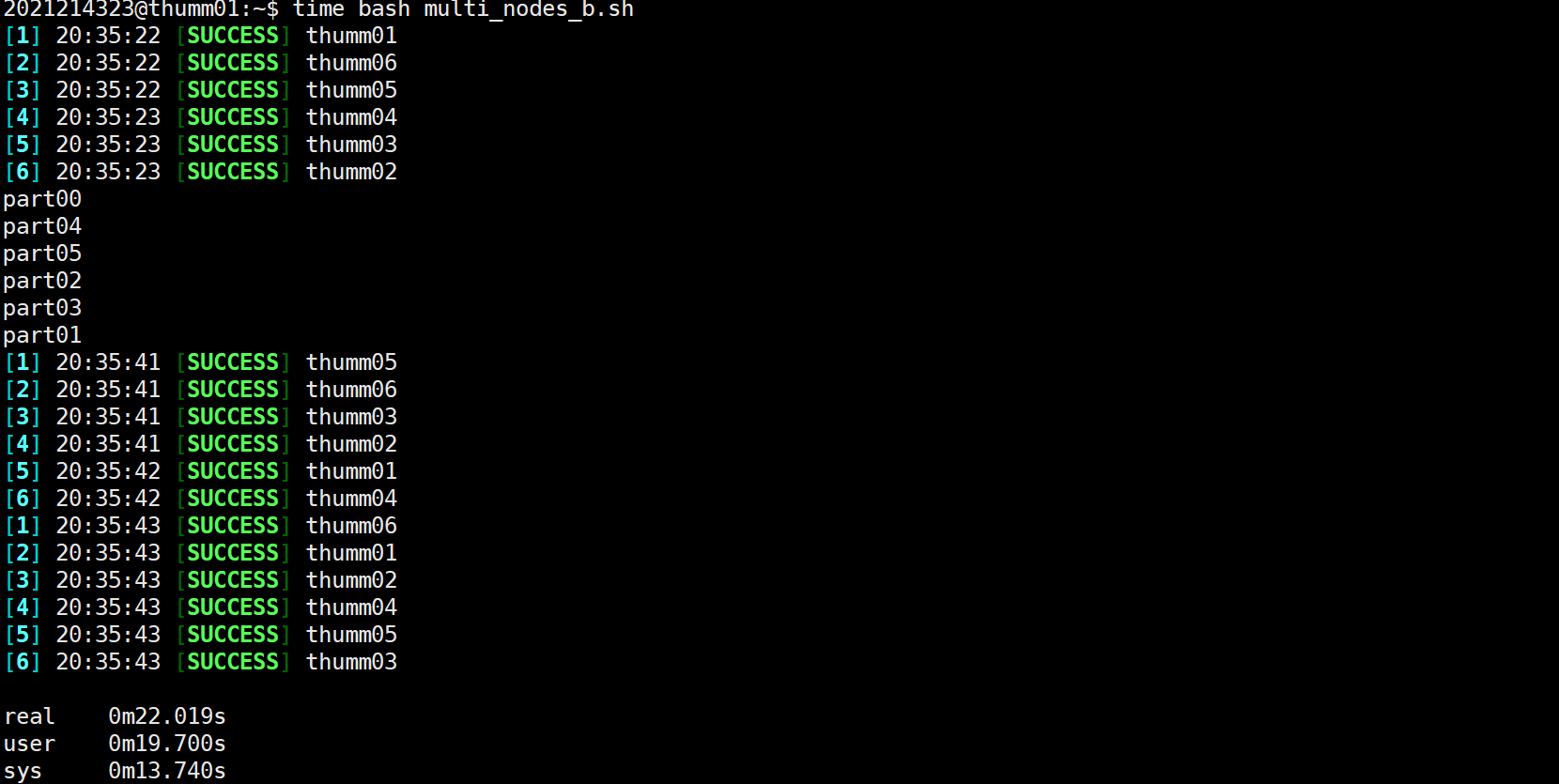
验证：在自制wc\_bigdata.txt上执行类似任务7过程

用time进行时间对比

代码1为-l参数以行数进行划分，代码2为-C参数以字节进行划分，结果分别如下

文本

描述已自动生成



# 代码1



# 代码2



任务9

Thumm01作为master，thumm02-06作为agent

Master生成子任务，并负责分发数据，最后实验结果如下

文本

描述已自动生成

文本

描述已自动生成

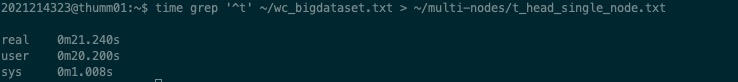
多机部分见代码3

多级执行时间

文本

描述已自动生成

单机执行时间



多机时间竟然比单机久，分析可能原因

多机数据分割方法问题，可能可以按任务7方式更好的分割，提高效率

ssh建立连接耗时，后续可改为其他方式

# 代码3

