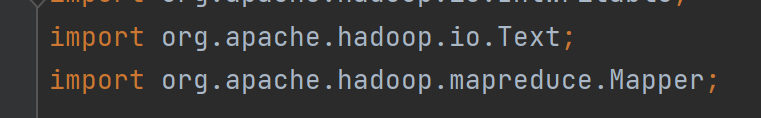
## bug1

WordCount类中添加代码中，会报红线



原因是Text类在很多包里面都有，项目没办法自己选择所需要的包，需要我们手动添加。

导入以下两个包后红色报错消失

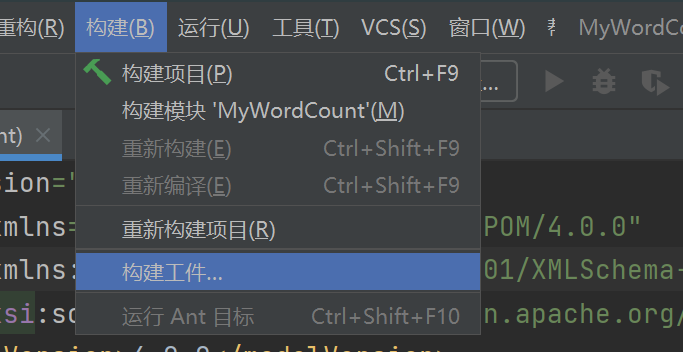


## bug2

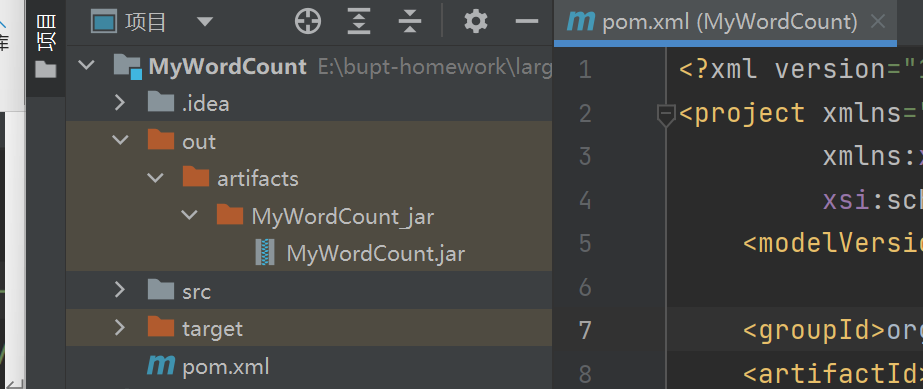
重新写完代码后，选择重新构建项目之后没有出现out包

原因：没有对WordCout进行工件构建，我们选择

构建 -> 构建工件 -> MyWordCount.jar -> 构建

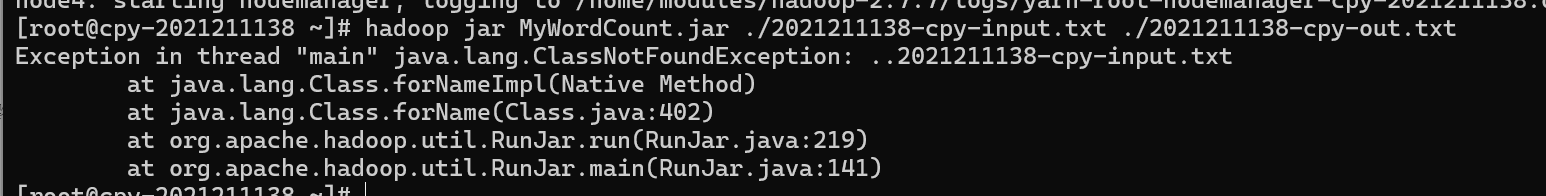


几秒过后，out包就会出现

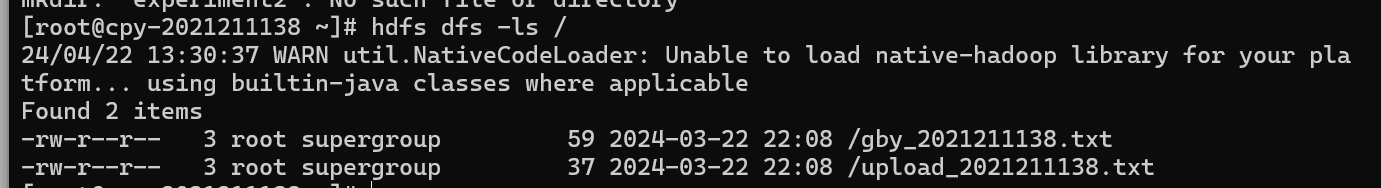


## bug3

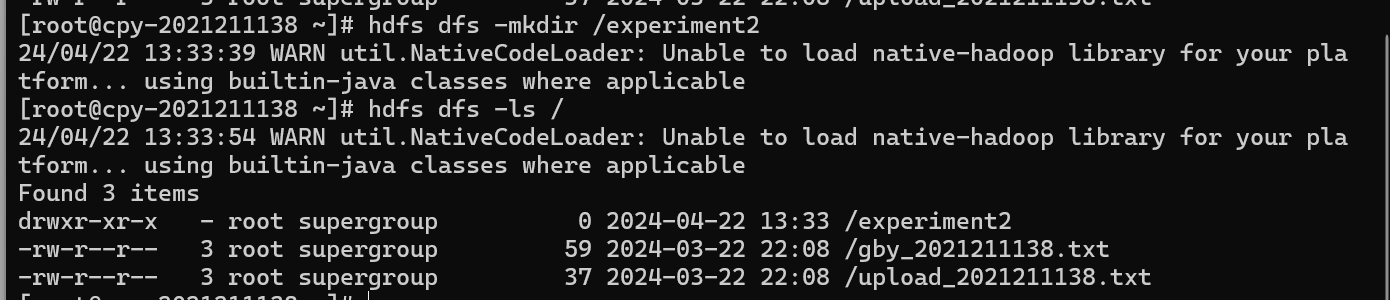
尝试运行jar包，主函数抛出异常



我觉得应该是hdfs找不到文件导致的

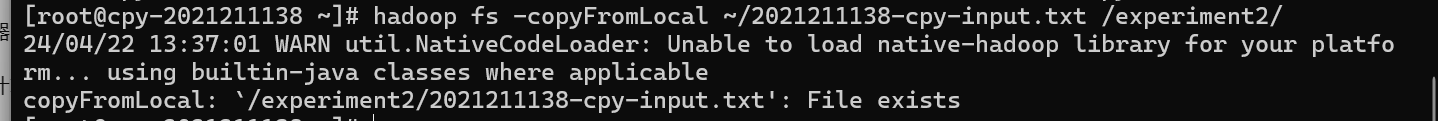


使用hdfs dfs -ls / 查看hdfs系统根目录 / 文件夹下的内容，发现并没有我们想要的输入文件。



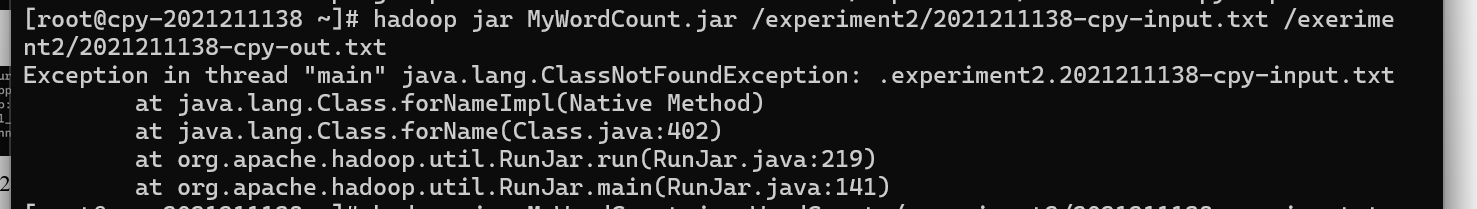
使用hdfs dfs -mkdir /experiment2，创建一个hdfs文件夹

使用hadoop fs -copyFromLocal <localSrc> <hdfsDst>将本地文件拷到hdfs文件系统的路径下



再次运行，就不报这个错误了。

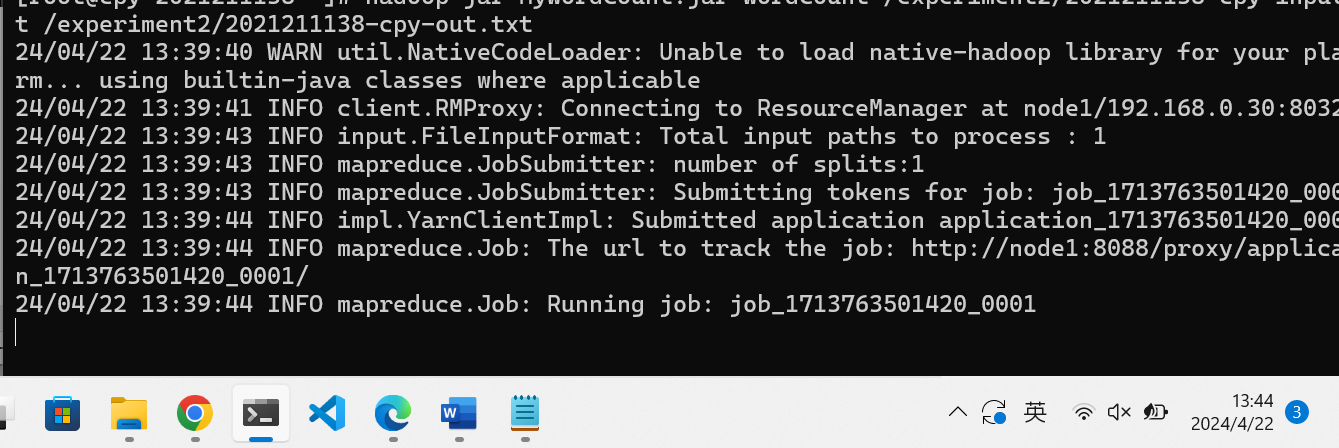
## bug4



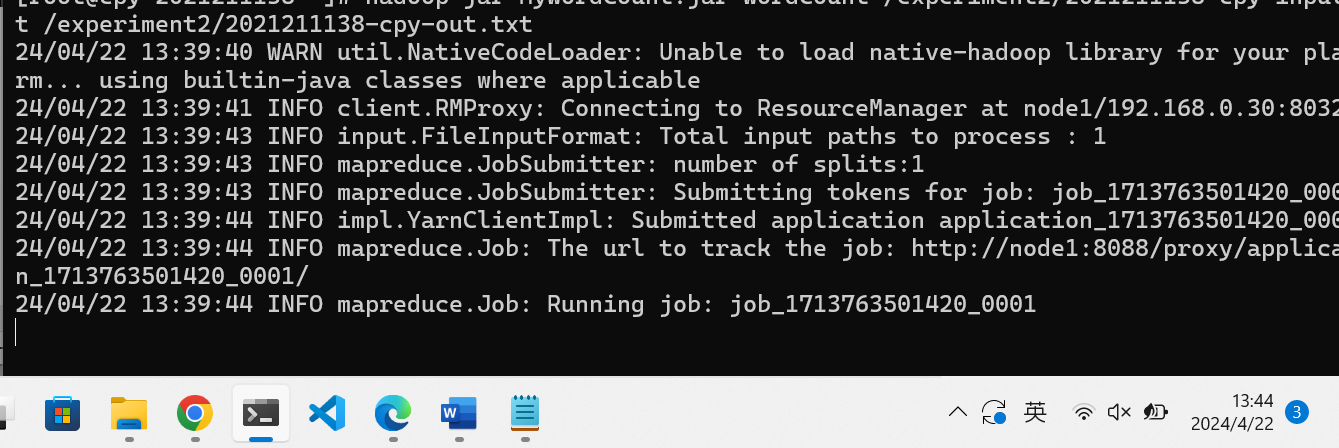
ClassNotFound错误。

原因是hadoop jar 命令的格式应该是 hadoop jar <JAR 文件路径> <主类> [参数]，而我的命令中缺少了主类的指定。所以应该将命令改为如下所示。

hadoop jar <jar包> <main函数的入口类> <input 文件或路径> <output路径>

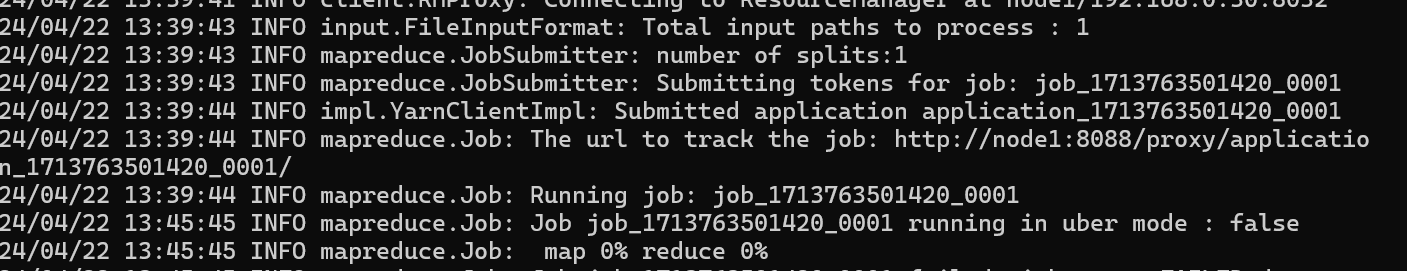


## bug5



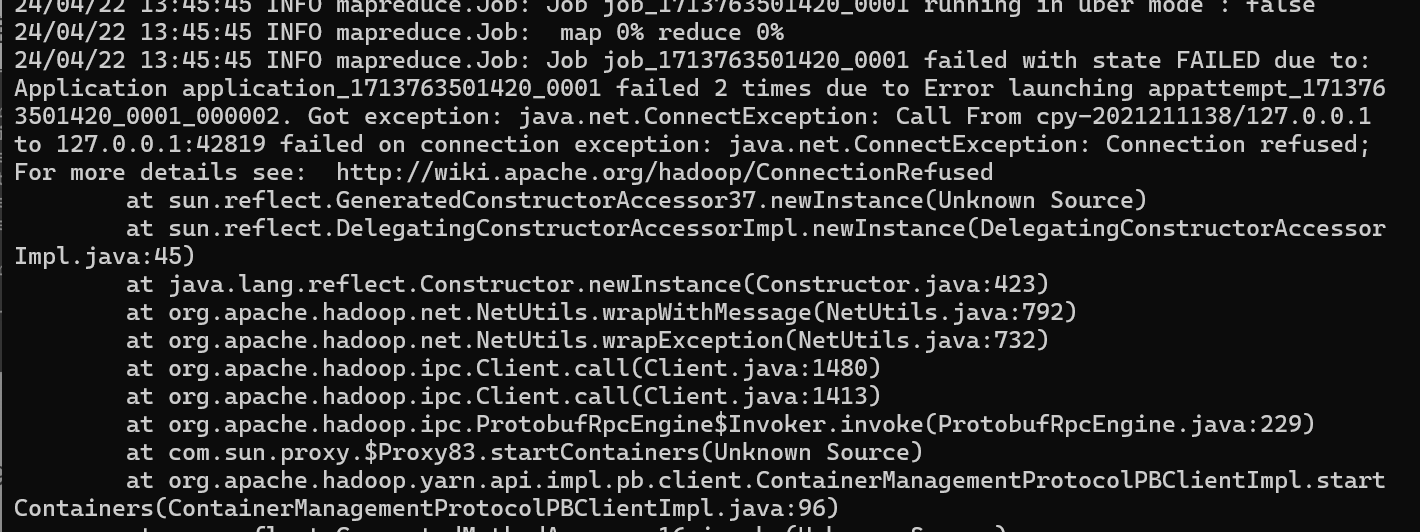
但是可以从时间戳上看到卡住了从39分到44分，很久都没有动静。

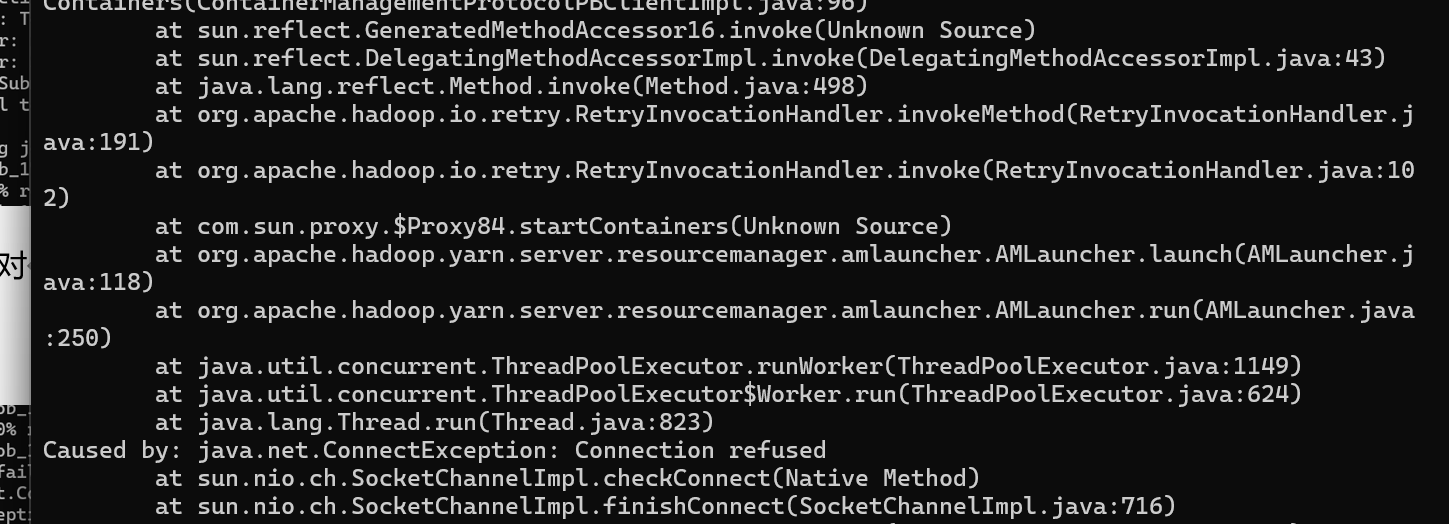
一直到45分，又有动静了

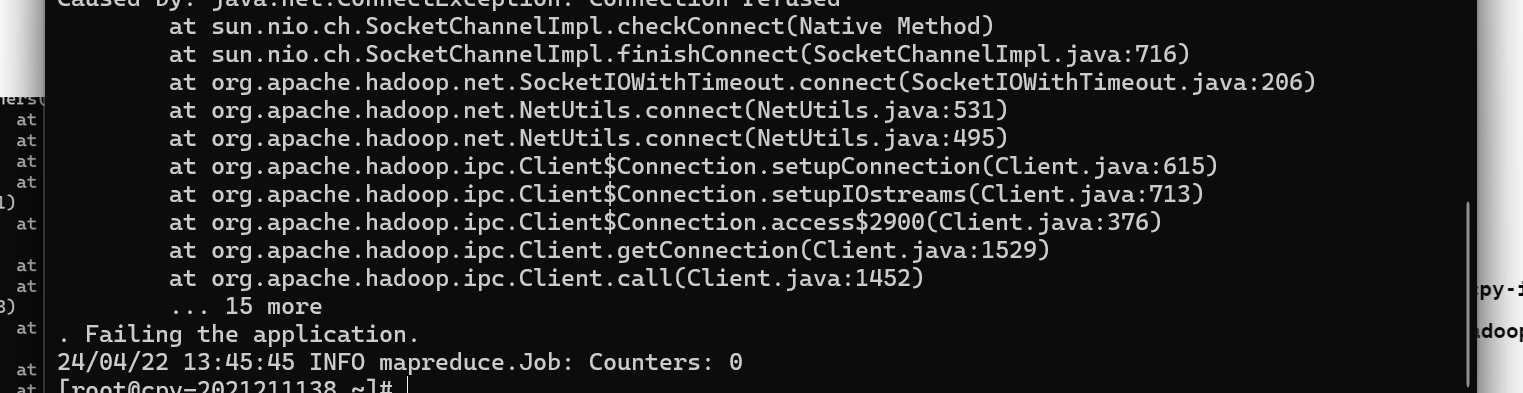


但是map 0% reduce 0%，显然不对

剩下的信息如下图所示





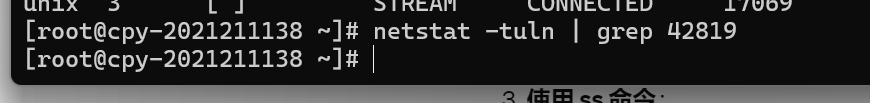


解决过程：

从下面这句话中，能得出，连接到42819端口失败了。

24/04/22 13:45:45 INFO mapreduce.Job: Job job\_1713763501420\_0001 failed with state FAILED due to: Application application\_1713763501420\_0001 failed 2 times due to Error launching appattempt\_1713763501420\_0001\_000002. Got exception: java.net.ConnectException: Call From cpy-2021211138/127.0.0.1 to 127.0.0.1:42819 failed on connection exception: java.net.ConnectException: Connection refused;

之后我使用netstat命令，并没有看到42819端口打开的信息，我觉得应该是端口没打开导致的。



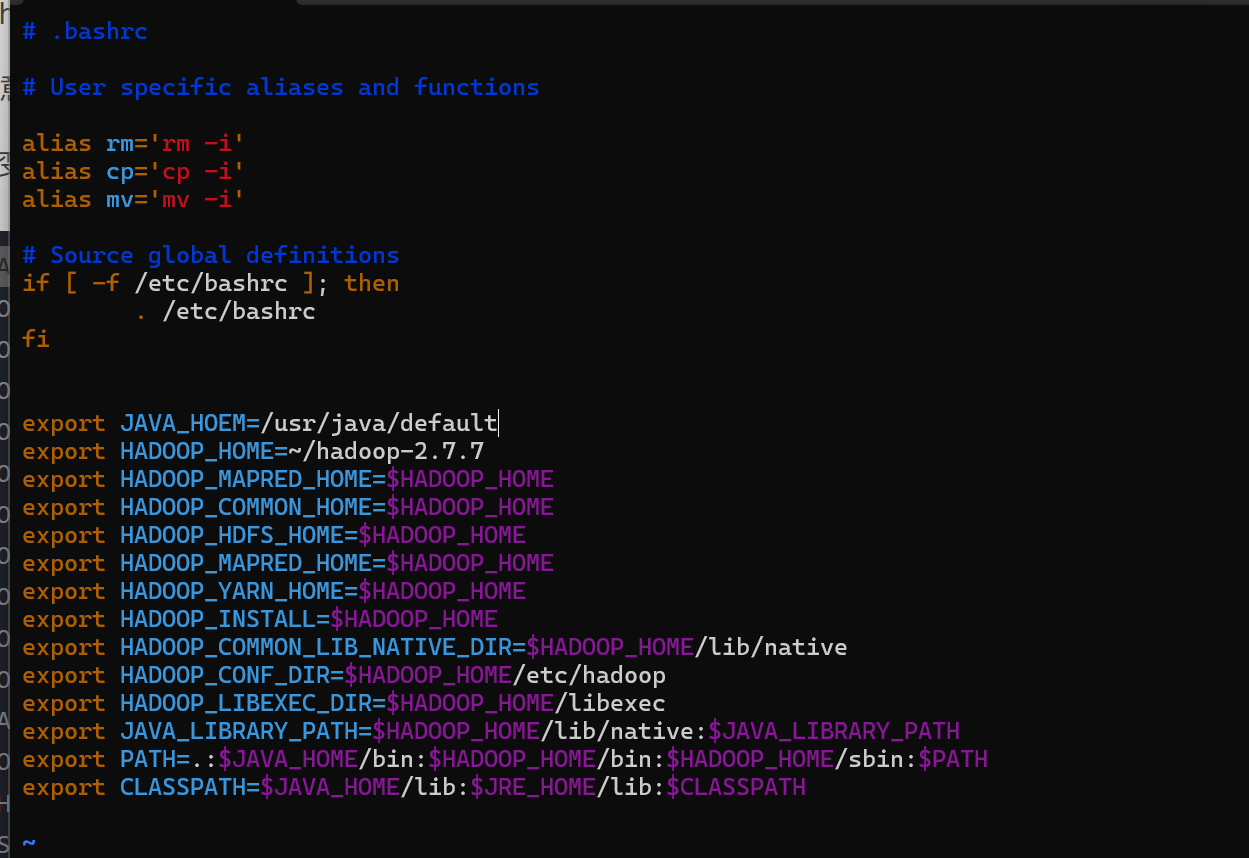
但是重新尝试了很多次每次的端口都不一样，应该不是端口的问题

我查阅了相关资料，有用的如下：

<https://blog.csdn.net/u014646662/article/details/82890443>

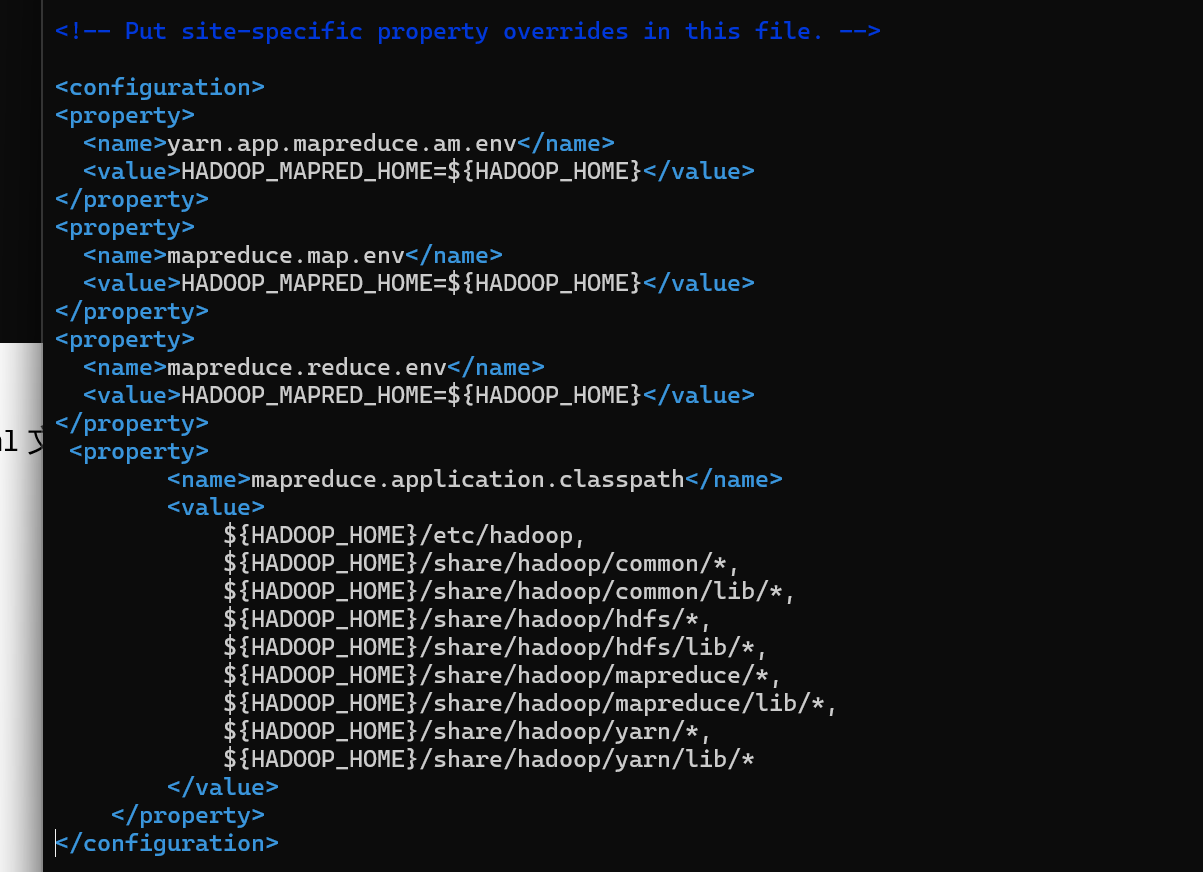
对于我这个问题，应该先配置环境变量

首先输入 vim ~/.bashrc，在最后面添加几行，如下



然后在hadoop-2.7.7/etc/hadoop 中添加一个mapred-site.xml文件

文件内容如下：



这个文件是为了配置hadoop框架的相关参数的

yarn.app.mapreduce.am.env: 这个属性用于配置 MapReduce 应用程序的 ApplicationMaster 环境变量。在这里，通过设置 HADOOP\_MAPRED\_HOME 为 ${HADOOP\_HOME}，将 MapReduce 应用程序的环境变量指向 Hadoop 安装目录。

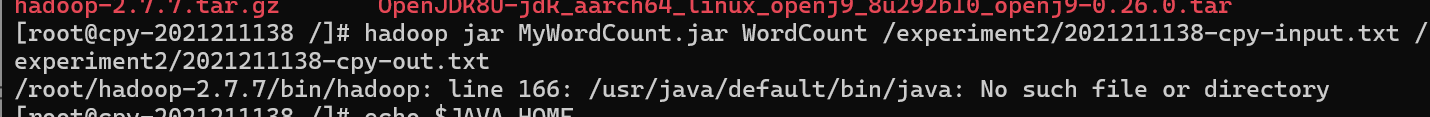
mapreduce.map.env: 这个属性用于配置 Map Task 执行器的环境变量。同样地，通过设置 HADOOP\_MAPRED\_HOME 为 ${HADOOP\_HOME}，将 Map Task 执行器的环境变量指向 Hadoop 安装目录。

mapreduce.reduce.env: 这个属性用于配置 Reduce Task 执行器的环境变量，同样地，通过设置 HADOOP\_MAPRED\_HOME 为 ${HADOOP\_HOME}，将 Reduce Task 执行器的环境变量指向 Hadoop 安装目录。

mapreduce.application.classpath: 这个属性用于配置 MapReduce 应用程序的类路径，以便能够加载所需的类和库文件。在这里，指定了各种 Hadoop 相关的目录和库文件的路径，以确保 MapReduce 任务能够正常运行。

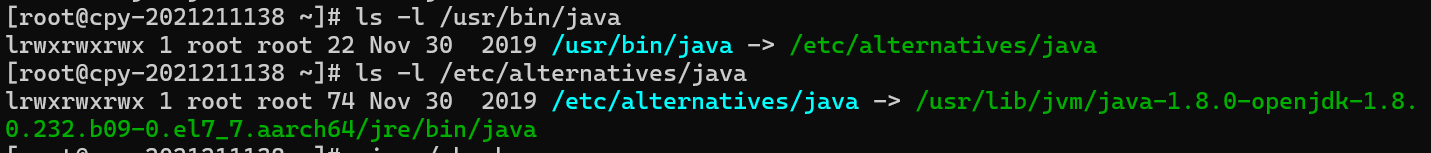
保存后，执行source ~/.bashrc 激活新的环境变量

但是这个时候又报错了，如下：

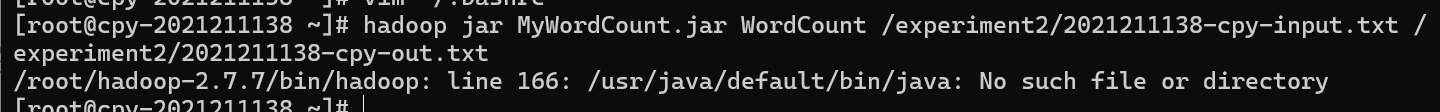


找不到/usr/java/deault/ 路径

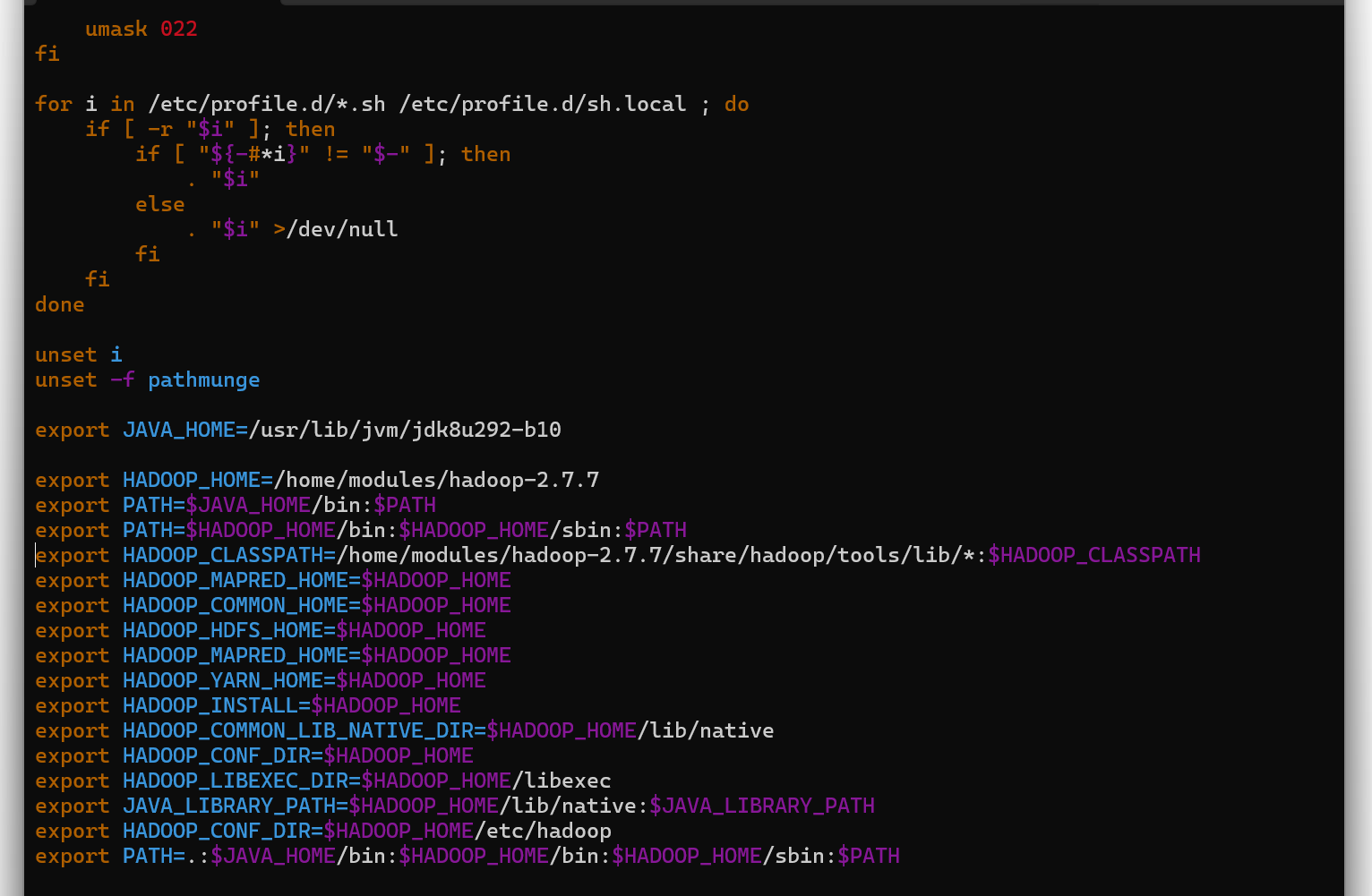
我通过如下命令找到了java的安装路径



修改了~/.bashrc的JAVA\_HOME后，还是不行



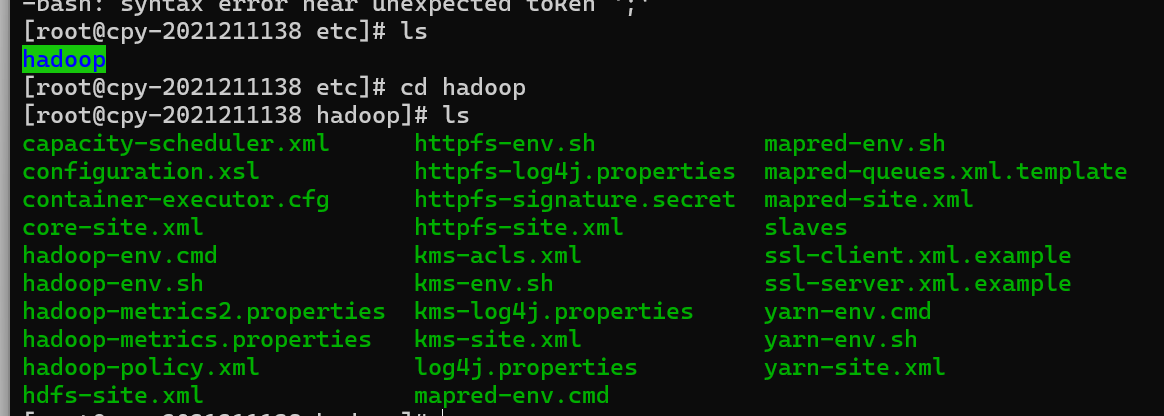
尝试修改 /etc/profile



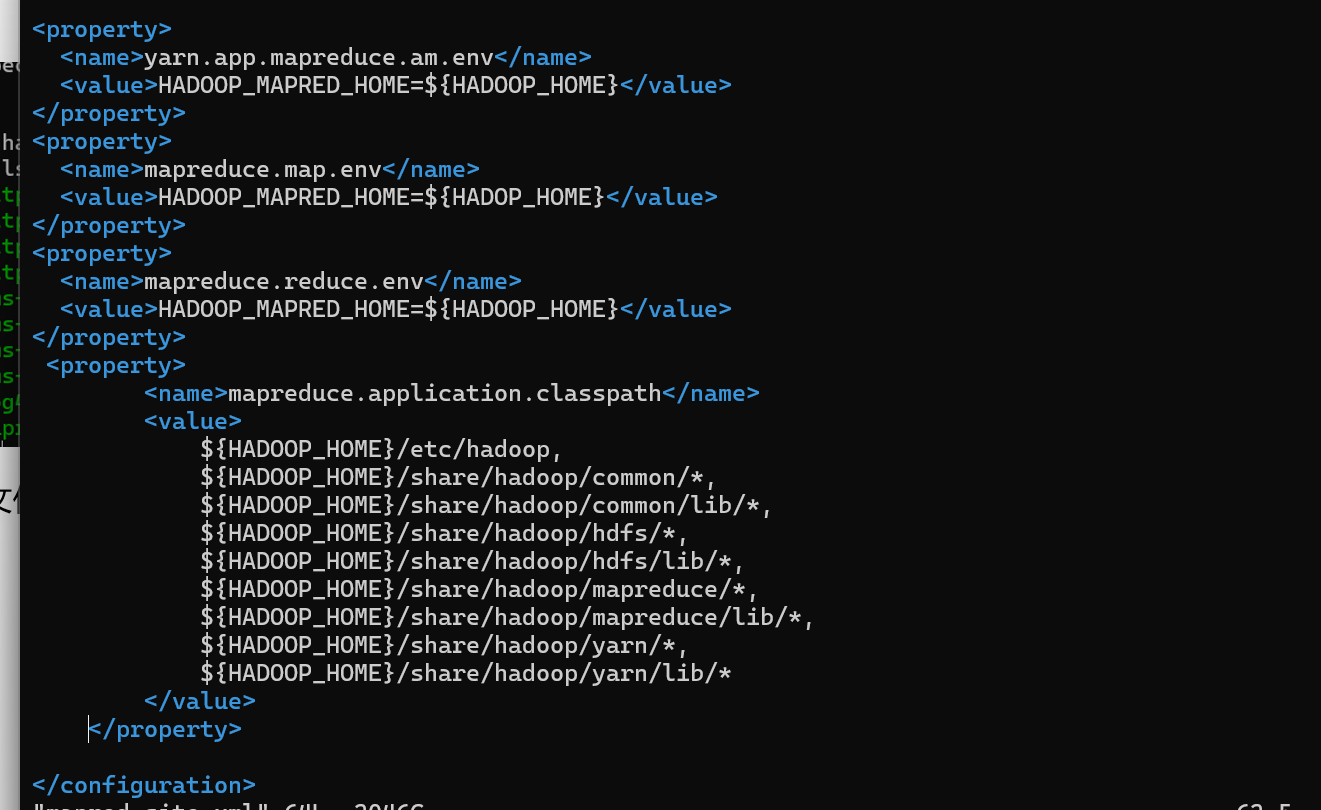
使用source /etc/profile重新激活

现在进入到HADOOP\_HOME目录中，如上图我的是/home/modules/hadoop-2.7.7

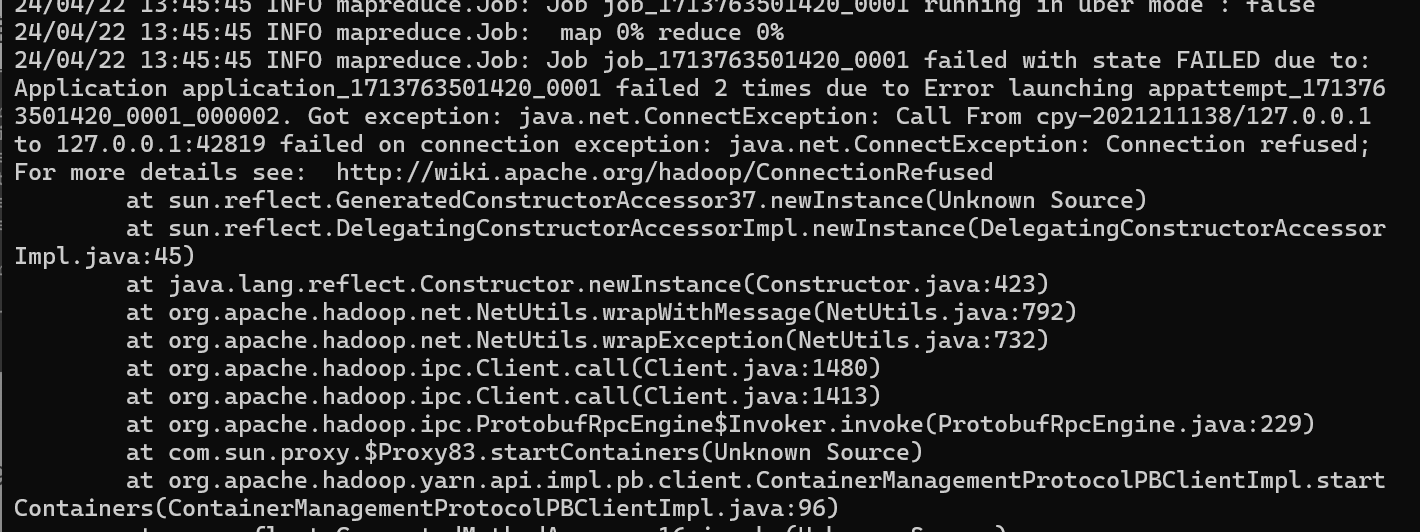
进入到etc/hadoop中



修改mapred-site.xml文件，添加如下信息

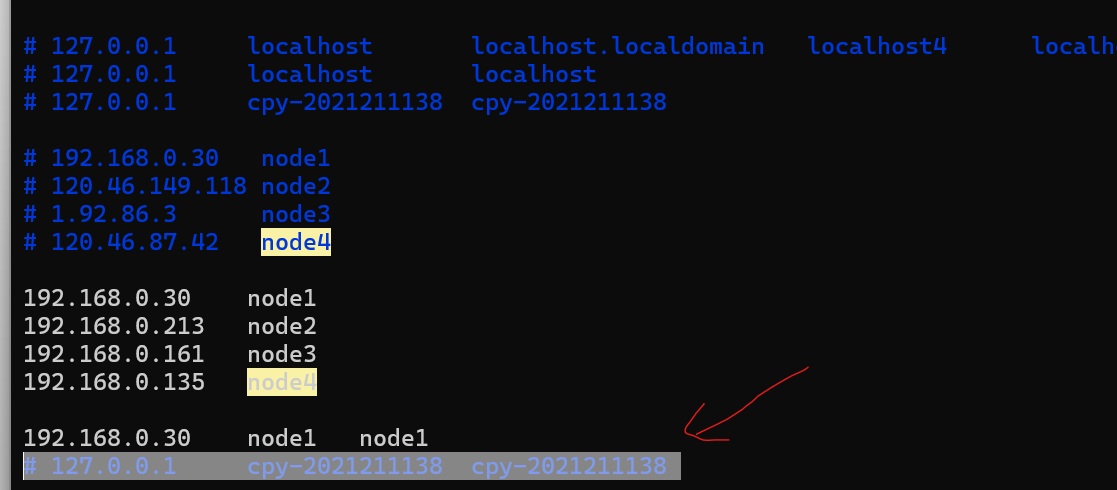


重新执行jar指令，但还是不行。这时候我认真看了一下报错信息



它其实是java连接错误，在连接到127.0.0.1的时候出错了。为什么它会连到42819端口呢？我们的hadoop服务的端口号并不是这个。我重新回顾了一下这个实验和上一个实验，定位了问题所在：hosts文件有问题。

打开/etc/hosts文件



因为服务器每次启动都会自己添加一个127.0.0.1的配置，所以它这里自动多了一行127.0.0.1 的配置信息。这也是为什么我始终运行不成功的原因了。把这一行删掉，或者注释掉之后，重新执行jar指令。

成功

