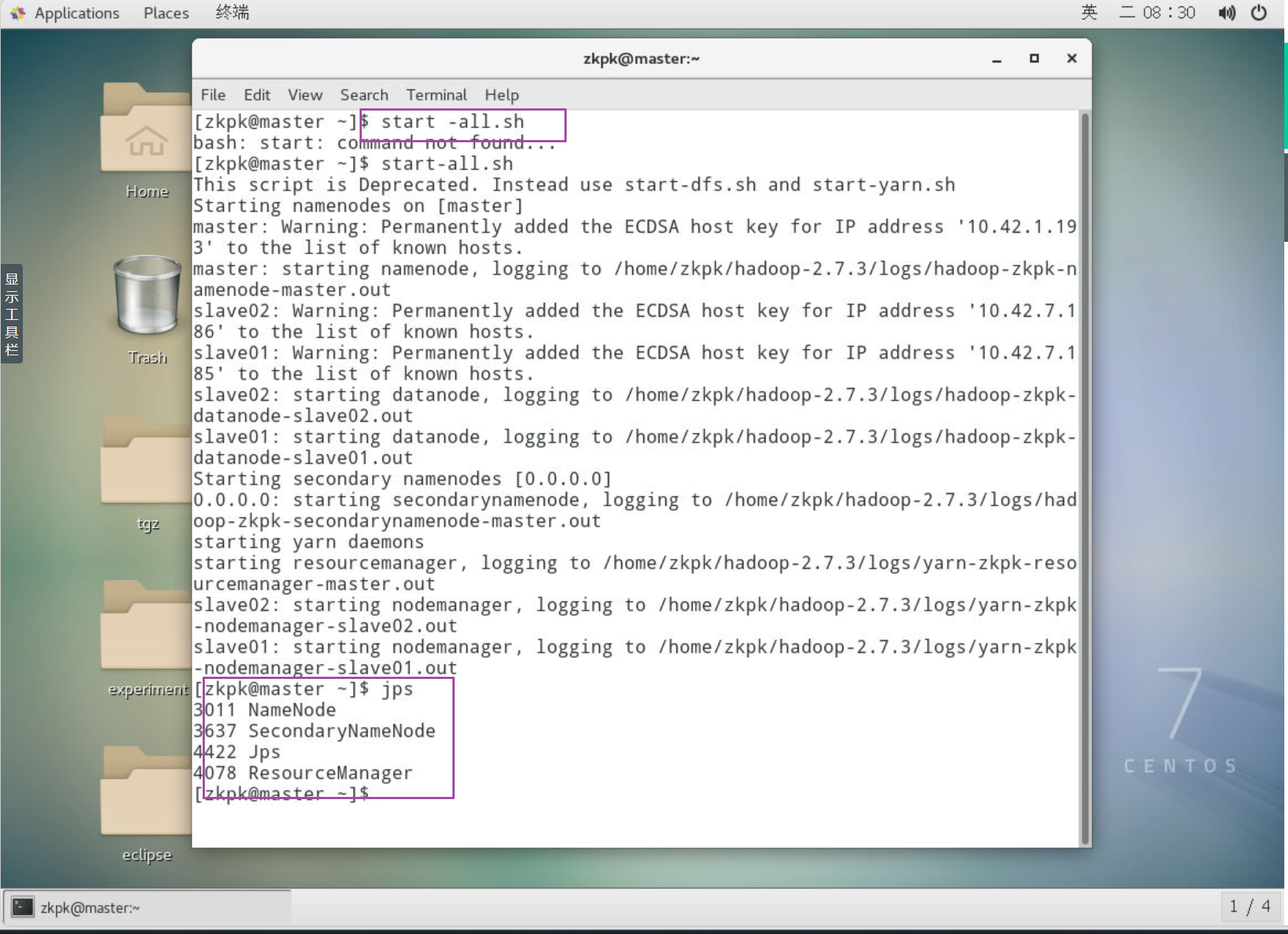
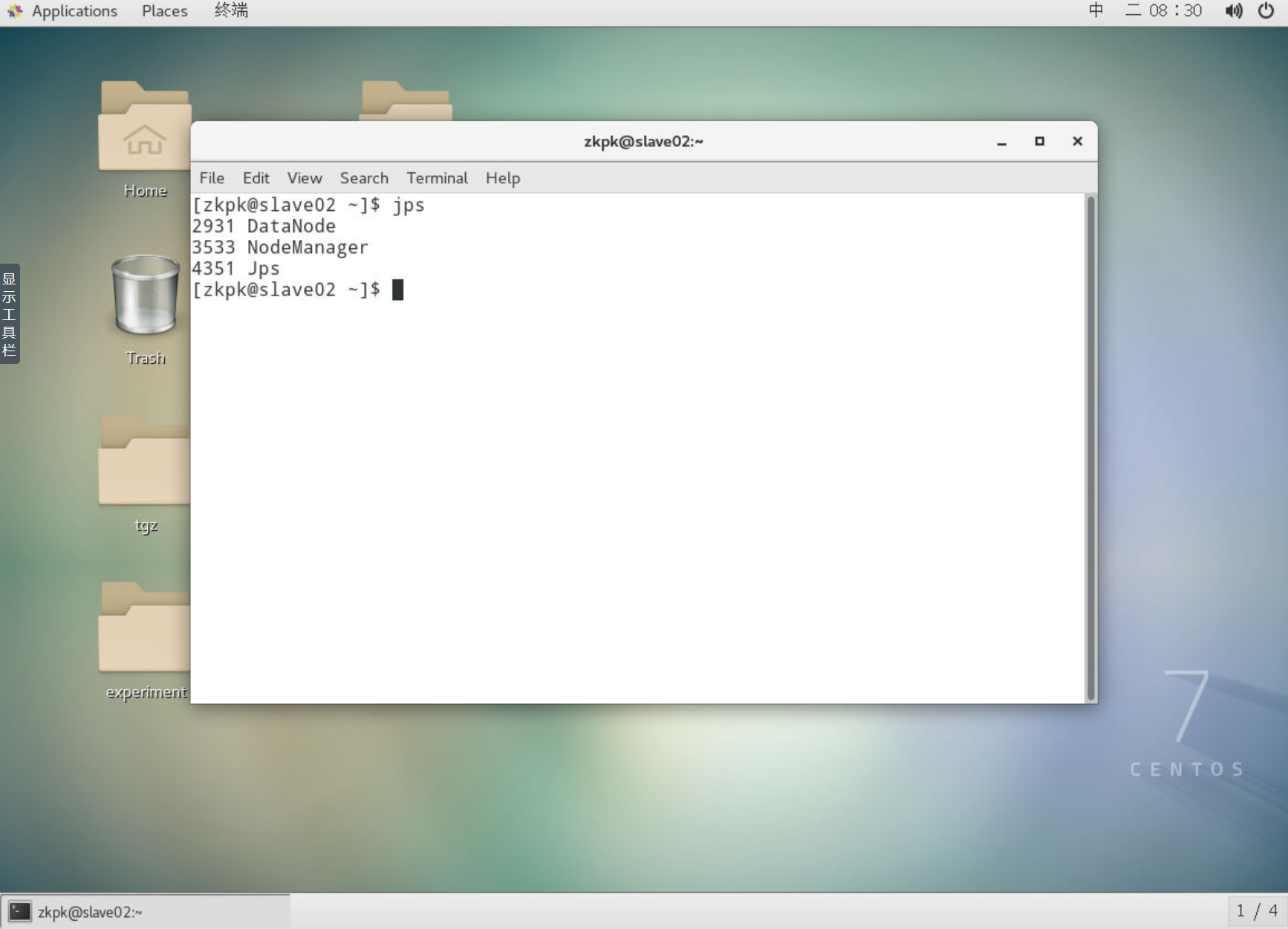
实验三、HDFS：IOUtils方式读取文件

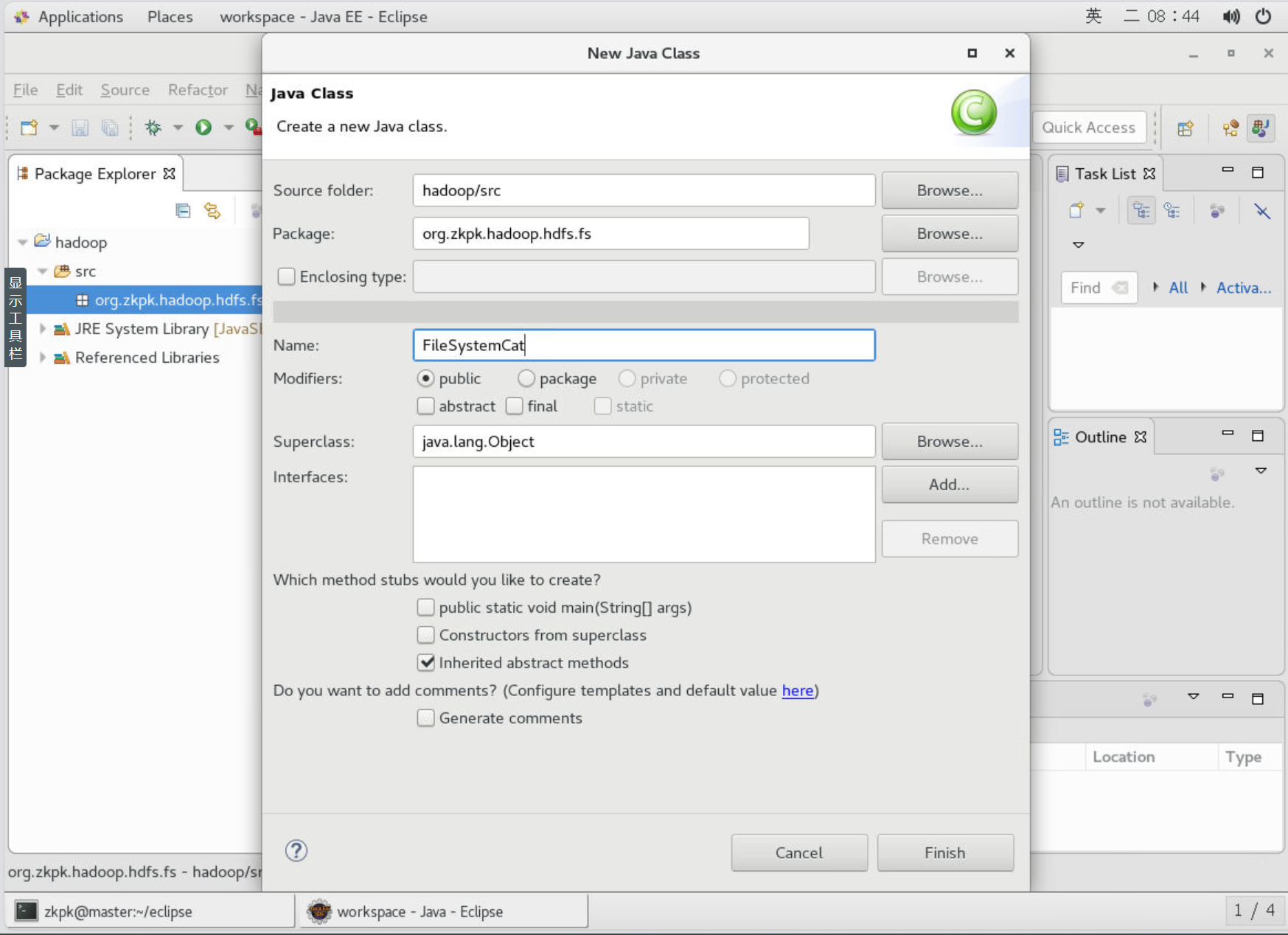
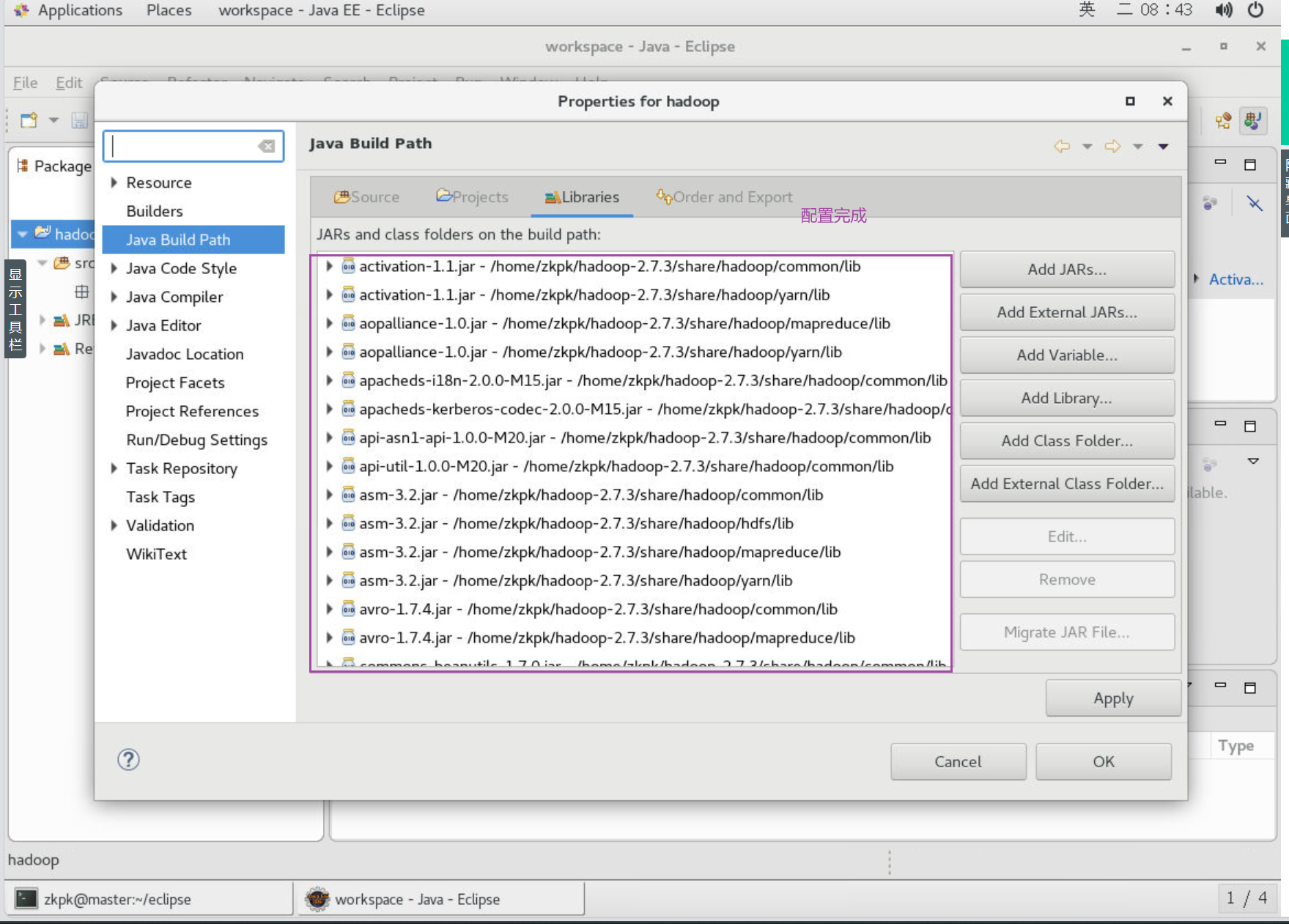
6.1

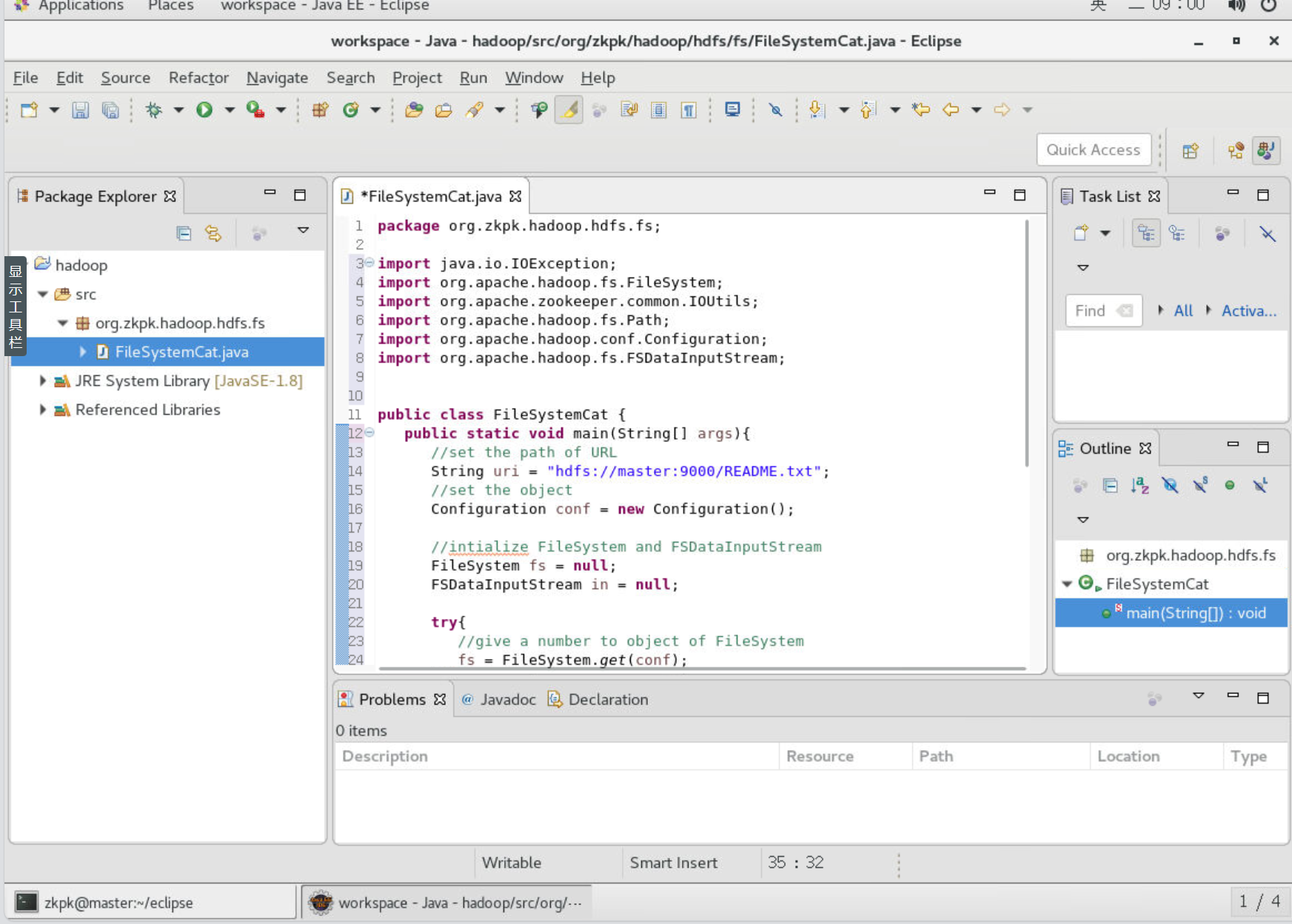


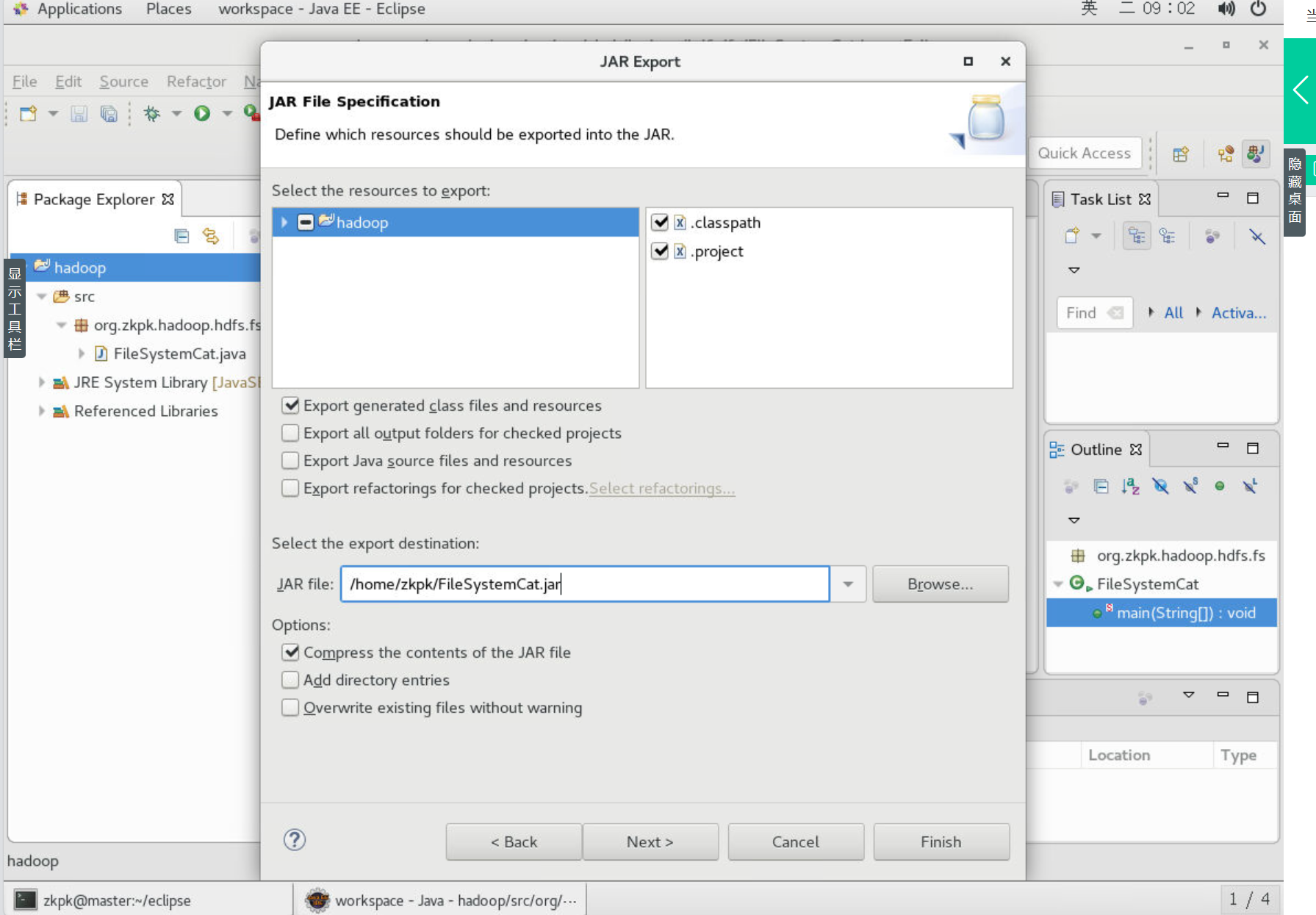


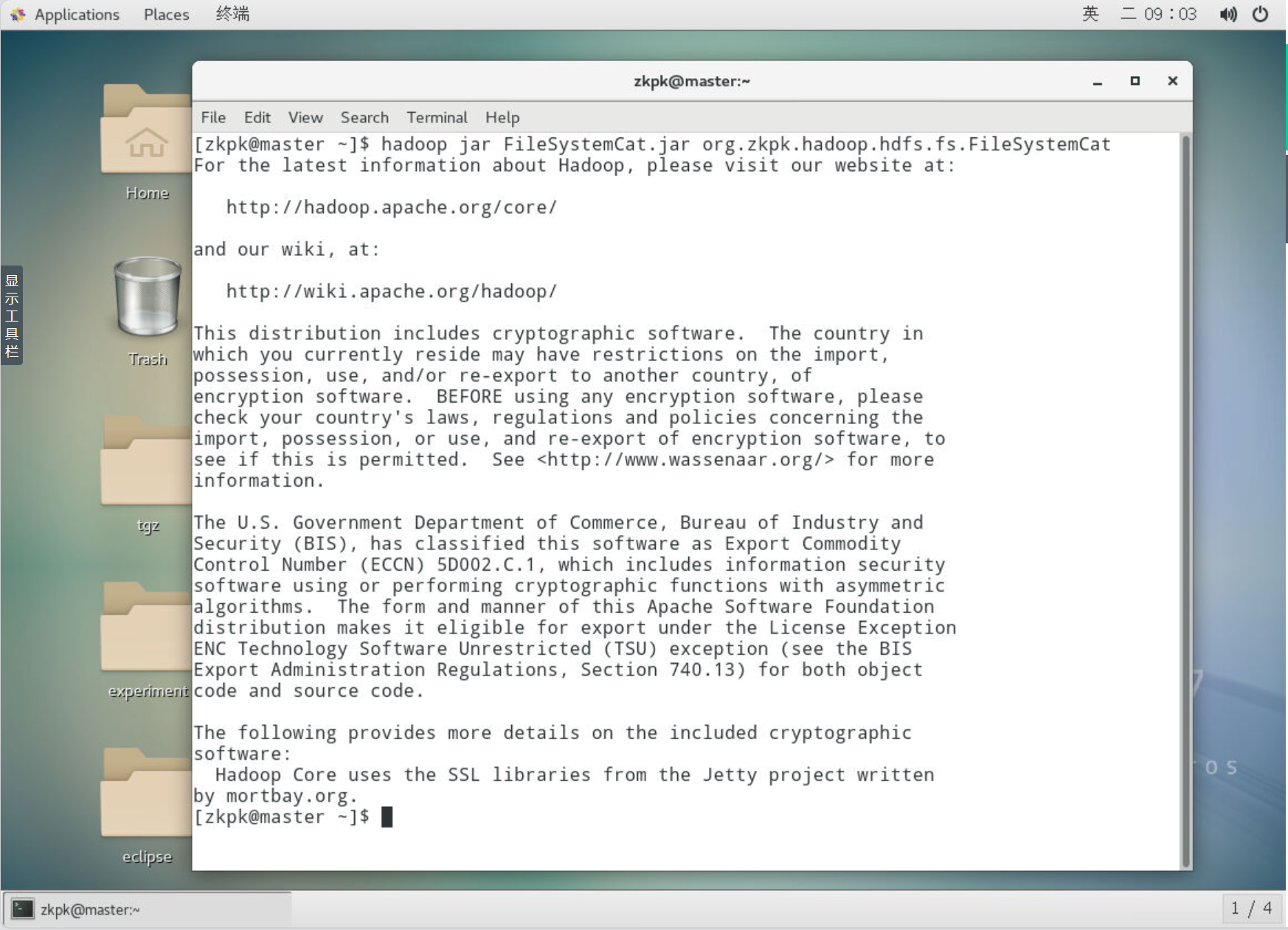












实验步骤：

通过URL类读取了HDFS上的文件内容 jps查看进程

1.在master启动Hadoop集群[zkpk@master ~]$ start-all.sh

2.上传README.txt文件到HDFS上，文件内容可以自定义

3.从Hadoop的公共目录下拷贝数据文件README.txt到/home/zkpk

[zkpk@master ~]$ cp ~/experiment/README.txt ~/

将上一步拷贝的数据文件README.txt上传到HDFS的根目录

[zkpk@master ~]$ hadoop fs –put README.txt /

验证上传成功

[zkpk@master ~]$ hadoop fs –ls /

4.启动Eclipse客户端

[zkpk@master ~]$ cd eclipse/

[zkpk@master eclipse]$ ./eclipse

创建Java工程 --->src右键创建package，名为org.zkpk.hadoop.hdfs.fs --->在工程名上单击鼠标右键，点击Build Path，点击Configure Build Path，弹出窗口，选择Add External Jars--->导入配置

share/hadoop/common

share/hadoop/common/lib

share/hadoop/hdfs

share/hadoop/hdfs/lib

share/hadoop/mapreduce

share/hadoop/mapreduce/lib

share/hadoop/yarn

share/hadoop/yarn/lib

---->所有的jar包添加完成后，点击Cofigure Build Path页面中的OK按钮完成环境配置

---->开始基于HDFS的API进行开发,在包内创建类，类名为FileSystemCat

----->代码编写完成，利用打包的方式生成java jar包并验证代码正确性,在工程名上右击鼠标，选择Export

---->在弹出窗口中，选择java的 JAR file

---->指定导出的jar包的位置及名字

5.---->代码打包完毕，关闭eclipse，回到linux命令行，提交jar包至HDFS上运行（其中org.zkpk.hadoop.hdfs.FileSystemCat是自定义FileSystemCat类的全限定名），查看实验结果