Hadoop Java编程

在linux服务器上安装了Hadoop3.2, 可用的端口号是9000/9870, 和8088.

其中9000端口在core-site.xml中配置的，注意配置成10.0.0.37:9000, 而不是localhost:9000。

配置成localhost，在windows链接hdfs的时候，出现了connect Exception连不上的问题。

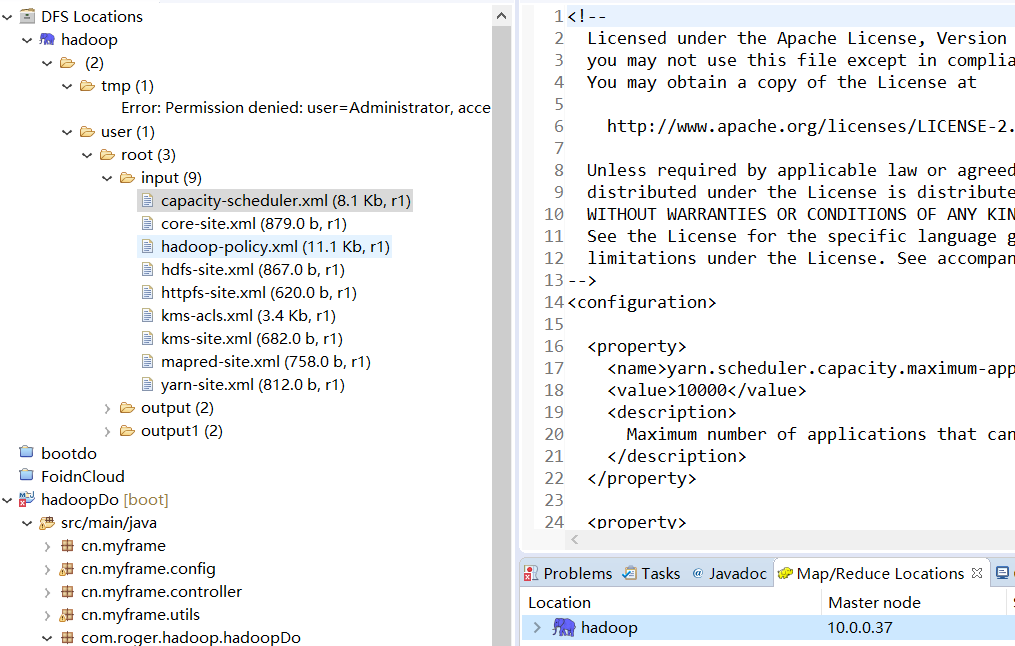
参考这个进行测试Java Hadoop编程：https://blog.csdn.net/weixin\_40096730/article/details/83818795

https://blog.csdn.net/cowbin2012/article/details/86680202

1）拦路虎 1

如果下载的过程中出现了 HADOOP\_HOME and hadoop.home.dir are unset,那么就说明你没有配置windows本地的hadoop环境变量.你可能会想我是远程调用linux下的hadoop,与我本地的hadoop有什么关系?如果你的操作只对远程的hadoop生效,如上传,创建目录,文件改名(写)等那么你是不需要在windows本地配置hadoop的,可一旦涉及到下载(读),hadoop内部的缓存机制要求本地也必须有hadoop,于是也会出现HADOOP\_HOME and hadoop.home.dir are unset,解决办法配置HADOOP\_HOME并加入%HADOOP\_HOME%\bin到PATH中,之后测试下hadoop version命令,有效的话重启你的eclipse/myeclipse,但这还不够,windows下的hadoop还需要winutils.exe,否则会报Could not locate Hadoop executable: xxxx\winutils.exe

设置好hadoop\_home后，安装eclipse hadoop插件碰到的问题也能解决了。



2）拦路虎 2

在配置测试windows下hadoop命令时，出现java错误：原文：https://blog.csdn.net/u012115730/article/details/80308797

在开发当中需要hadoop的环境搭建，当环境都搭好后运行hadoop时出现以下情况：

Error: JAVA\_HOME is incorrectly set.

Please update D:\javatemps\hadoop-2.7.2\conf\hadoop-env.cmd

提示我java环境搭建错误，找不见，这时我就慌了，刚刚还java -version了，怎么就找不见了。见鬼了吧。。。

这时我查了以下网络，才知道并不是环境问题，而是要将你的

D:\javatemps\hadoop-2.7.2\conf\hadoop-env.cmd这个东西中的java\_home 环境改下。

原来是配置文件的问题，我的JAVA\_HOME目录是C:\Program Files\Java\jdk1.8.0\_121，因为Program Files中存在空格，所以出现错误，只需要用

PROGRA~1代替Program Files即可，即改为C:\PROGRA~1\Java\jdk1.8.0\_121，当然，你也可以讲jdk装到根目录或者不存在空格等目录下。

改成上面这样子。

---------------------

System.***out***.println( "Hello World!" );

URL.*setURLStreamHandlerFactory*(**new** FsUrlStreamHandlerFactory());

//InputStream in = new URL("hdfs://10.0.0.37:9000/hello.txt").openStream();

InputStream in = **new** URL("hdfs://10.0.0.37:9000/user/root/input/core-site.xml").openStream();

IOUtils.*copyBytes*(in, System.***out***,4096, **true**);