

# 作业4

@owen

注意安装jdk1.8 (jdk11不行)

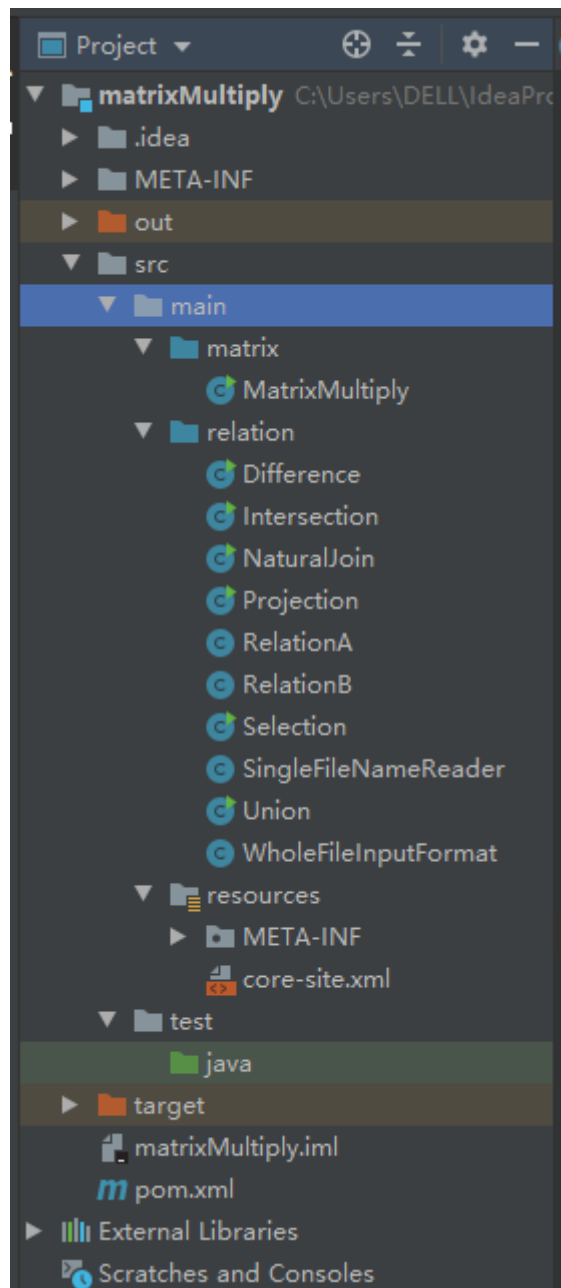
java路径 program files中间有空格，使用PROGRA~1代替

实验流程：使用IDEA本机写代码，打包成jar放到服务器Hadoop集群运行

## 1.配置环境

### 1.0 新建项目

新建maven项目（先安装maven），结构如下



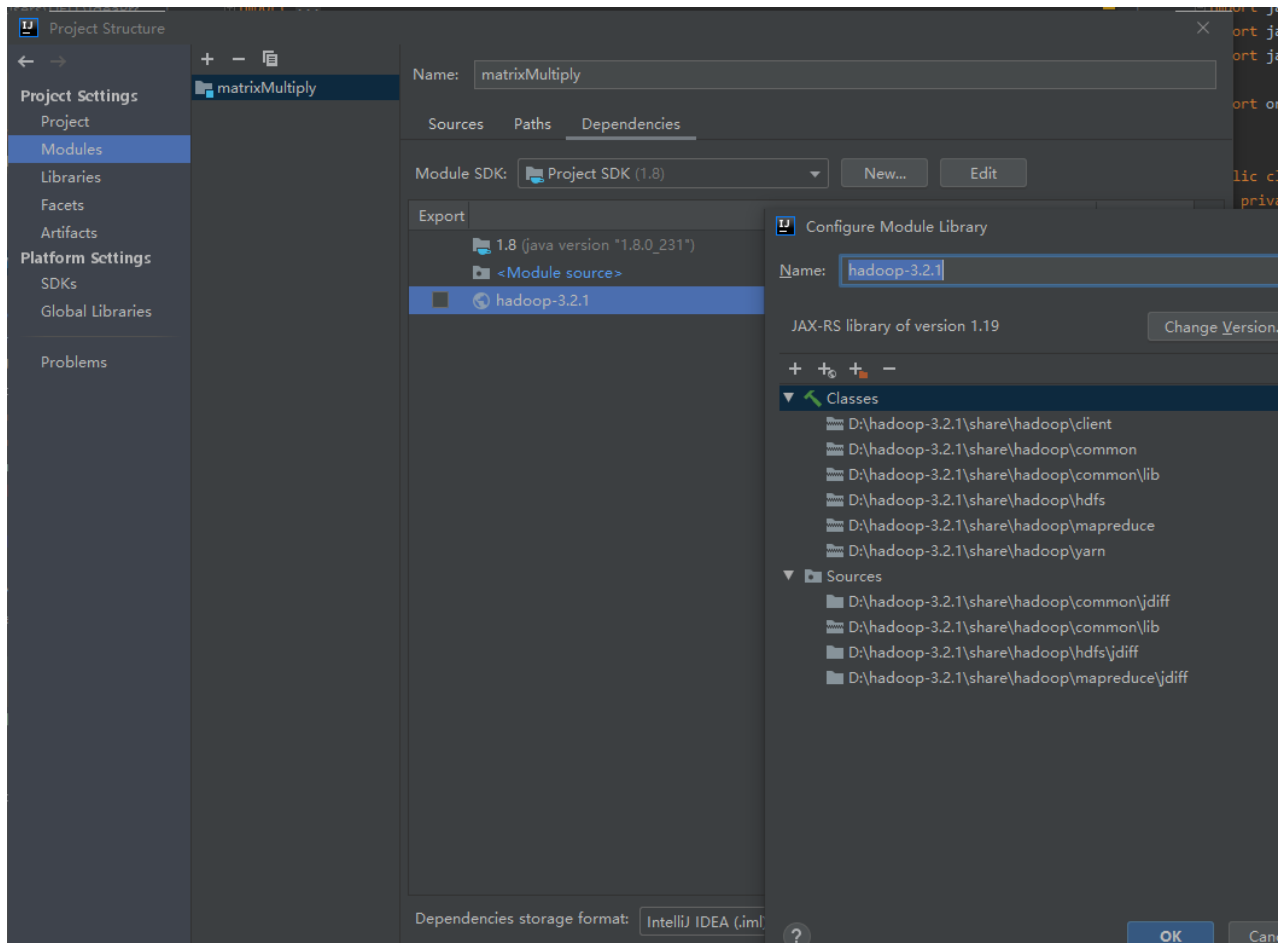
## 1.1添加jar包依赖

### 1. 方法一

通过编辑maven的pom.xml文件夹，这时会自动下载hadoop需要的jar文件

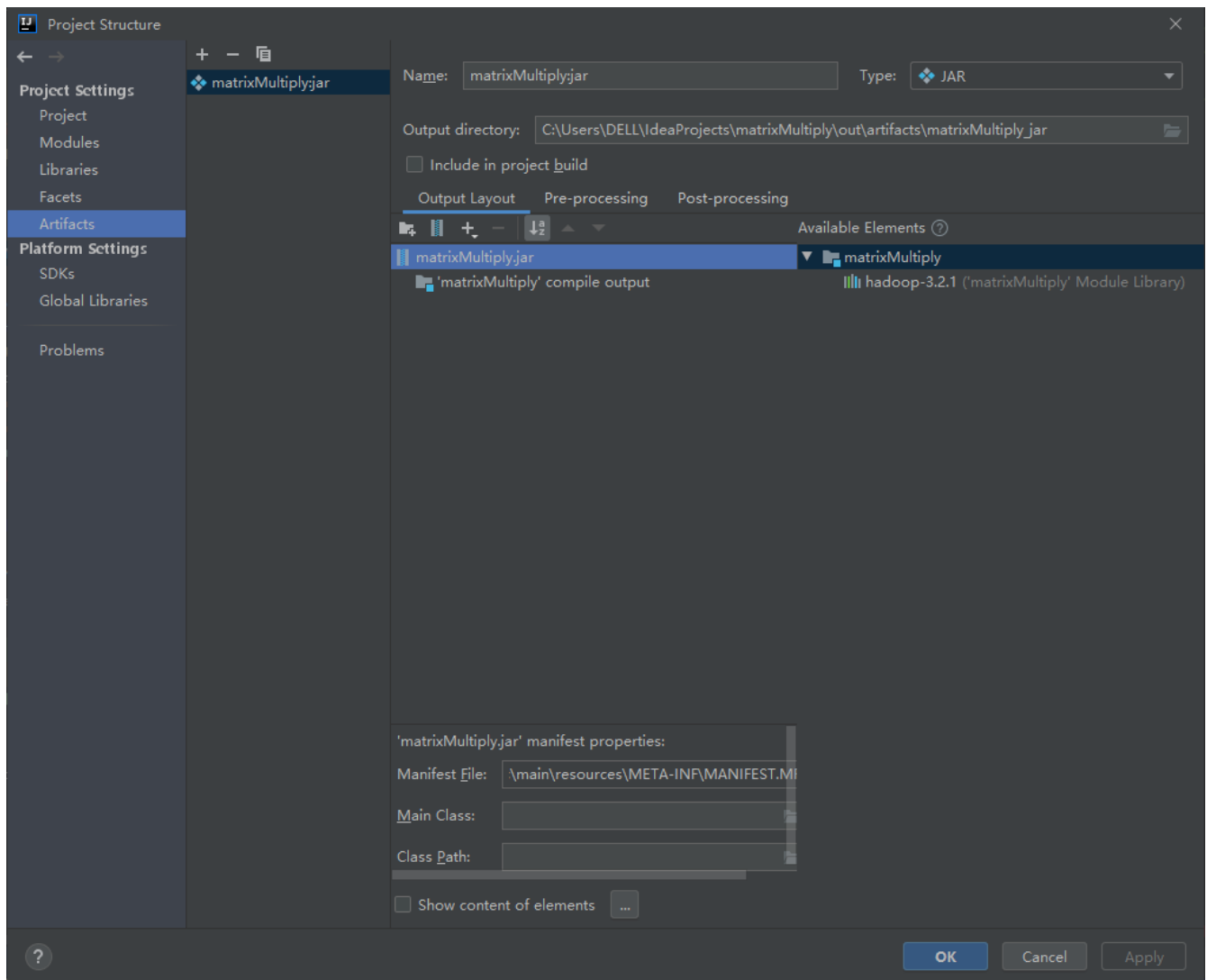
### 2. 方法二

手动从hadoop安装包中的share/hadoop文件夹中添加相关jar依赖：IDEA中设置，file-> project Structure-> Moduels-> dependencies ->+号添加



## 1.2配置Build Artifacts

同样在project Structure下，artifacts->将available Elements中的compile output加到左边（注意不添加1.1配好的hadoop依赖jar文件）-> manifest properties中，manifest File 路径选在resource下，mainClass不填（运行时候自行指定）



### 1.3 配置core-site.xml

把集群中core-site.xml和hdfs-site.xml两个文件拿下来，放到项目中resources目录下

### 1.4 配置target bytecode version

file-> settings->compiler->java compiler-> target bytecode version 选为1.8

## 2 运行结果

代码见仓库[https://github.com/VoidFly/hadoop\\_practice](https://github.com/VoidFly/hadoop_practice)

### 2.1 矩阵乘法结果

bin/hadoop jar ../share-files/matrixMultiply.jar input/M\_3\_4 input/N\_4\_2 mOutput

注意课本程序运行命令，.jar 后面不用加类名（MatrixMultiply）

```

root@f6815da6c635:/home/Hadoop/hadoop-3.2.1# bin/hdfs dfs -cat mOutput/*
2019-10-21 03:23:31,570 INFO sasl.SaslDataTransferClient: SASL encryption trust
check: localHostTrusted = false, remoteHostTrusted = false
1,1      6851
1,2      4947
2,1      13220
2,2      6935
3,1      12523
3,2      13426

```

## 2.2 关系代数

选择运算 age=18

```

627613756_0011/
2019-10-21 10:20:45,105 INFO mapreduce.Job: Running job: job_1571627613756_0011
2019-10-21 10:20:53,375 INFO mapreduce.Job: Job job_1571627613756_0011 running in uber mode : false
2019-10-21 10:20:53,379 INFO mapreduce.Job: map 0% reduce 0%
2019-10-21 10:20:59,457 INFO mapreduce.Job: map 100% reduce 0%
2019-10-21 10:20:59,475 INFO mapreduce.Job: Job job_1571627613756_0011 completed successfully
2019-10-21 10:20:59,583 INFO mapreduce.Job: Counters: 33
    File System Counters
        FILE: Number of bytes read=0
        FILE: Number of bytes written=226164
        FILE: Number of read operations=0

```

```

root@f6815da6c635:/home/Hadoop/hadoop-3.2.1# bin/hdfs dfs -cat rOutput/*
2019-10-21 09:51:56,638 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = f
alse, remoteHostTrusted = false
1,tom,18,60.0
4,tony,18,62.0
7,brown,18,65.0
12,ivy,18,58.0
13,sam,18,67.0
16,steven,18,60.0
19,coco,18,55.0

```

age< 18 (注意bash中输入小于符号要写成 "\<" 转义)

```

bash: 18: No such file or directory
root@f6815da6c635:/home/Hadoop/hadoop-3.2.1# bin/hadoop jar ../share-files/matrixMultiply.jar Selection inp
ut/Ra rOutput age \< 18

```

```

root@f6815da6c635:/home/Hadoop/hadoop-3.2.1# bin/hdfs dfs -cat rOutput/*
2019-10-21 10:23:59,871 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = f
alse, remoteHostTrusted = false
2,jack,16,59.0
3,lily,17,58.0
8,candy,17,56.0
10,grace,16,56.0
11,henry,17,61.0
17,jimmy,16,62.0
18,lucas,17,59.0
20,zoe,17,56.0

```

对name进行投影

```

2019-10-21 11:52:47,547 INFO mapreduce.Job: Running job: job_1571627613756_0014
2019-10-21 11:52:55,808 INFO mapreduce.Job: Job job_1571627613756_0014 running in uber mode : false
2019-10-21 11:52:55,812 INFO mapreduce.Job: map 0% reduce 0%
2019-10-21 11:53:01,910 INFO mapreduce.Job: map 100% reduce 0%
2019-10-21 11:53:07,958 INFO mapreduce.Job: map 100% reduce 100%
2019-10-21 11:53:07,979 INFO mapreduce.Job: Job job_1571627613756_0014 completed successfully

```

```

root@f6815da6c635:/home/Hadoop/hadoop-3.2.1# bin/hdfs dfs -cat rOutput/*
2019-10-21 11:53:13,867 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
bill
bob
brown
candy
cici
coco
grace
henry
ivy
jack
jimmy
leon
lily
lucas
owen
sam
steven
tom
tony
zoey

```

### 求交运算

```

root@f6815da6c635:/home/Hadoop/hadoop-3.2.1# bin/hdfs dfs -cat rOutput/*
2019-10-21 12:04:41,973 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
3,lily,17,58.0
8,candy,17,56.0
9,cici,19,55.0
10,grace,16,56.0
12,ivy,18,58.0
19,coco,18,55.0
20,zoey,17,56.0

```

### 求差运算

```

root@f6815da6c635:/home/Hadoop/hadoop-3.2.1# bin/hadoop jar ../share-files/matrixMultiply.jar Difference in
tersecInput rOutput Ra2
2019-10-21 15:27:21,193 INFO client.RMPProxy: Connecting to ResourceManager at h0/172.18.0.2:8032
2019-10-21 15:27:21,781 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.

root@f6815da6c635:/home/Hadoop/hadoop-3.2.1# bin/hdfs dfs -cat rOutput/*
2019-10-21 15:28:14,214 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
21,linda,19,60.0
22,tina,17,56.0
25,monica,20,61.0

```

### 求并运算

```

root@f6815da6c635:/home/Hadoop/hadoop-3.2.1# bin/hdfs dfs -cat rOutput/*
2019-10-21 15:36:57,009 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
1, tom, 18, 60.0
2, jack, 16, 59.0
3, lily, 17, 58.0
4, tony, 18, 62.0
5, bob, 20, 65.0
6, leon, 19, 58.0
7, brown, 18, 65.0
8, candy, 17, 56.0
9, cici, 19, 55.0
10, grace, 16, 56.0
11, henry, 17, 61.0
12, ivy, 18, 58.0
13, sam, 18, 67.0
14, owen, 20, 63.0
15, bill, 19, 62.0
16, steven, 18, 60.0
17, jimmy, 16, 62.0
18, lucas, 17, 59.0
19, coco, 18, 55.0
20, zoey, 17, 56.0
21, linda, 19, 60.0
22, tina, 17, 56.0
25, monica, 20, 61.0
root@f6815da6c635:/home/Hadoop/hadoop-3.2.1#

```

```

root@f6815da6c635:/home/Hadoop/hadoop-3.2.1# bin/hadoop jar ../share-files/matrixMultiply.jar Union intersectInput rOutput

```

自然连接 Ra Rb 属性为id

```

root@f6815da6c635:/home/Hadoop/hadoop-3.2.1# bin/hadoop jar ../share-files/matrixMultiply.jar NaturalJoin naturalInput rOutput 0 Ra
2019-10-22 01:22:12,854 INFO client.RMProxy: Connecting to ResourceManager at 10.178.18.0:8032

```

```

root@f6815da6c635:/home/Hadoop/hadoop-3.2.1# bin/hdfs dfs -cat rOutput/*
2019-10-22 01:23:45,355 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
1, tom, 18, 60.0, 1, 178.0
10, grace, 16, 56.0, 0, 170.0
11, henry, 17, 61.0, 1, 173.0
12, ivy, 18, 58.0, 0, 162.0
13, sam, 18, 67.0, 1, 182.0
14, owen, 20, 63.0, 1, 177.0
15, bill, 19, 62.0, 1, 177.0
16, steven, 18, 60.0, 1, 175.0
17, jimmy, 16, 62.0, 1, 178.0
18, lucas, 17, 59.0, 1, 183.0
19, coco, 18, 55.0, 0, 160.0
2, jack, 16, 59.0, 1, 175.0
20, zoey, 17, 56.0, 0, 168.0
3, lily, 17, 58.0, 0, 165.0
4, tony, 18, 62.0, 1, 173.0
5, bob, 20, 65.0, 1, 179.0
6, leon, 19, 58.0, 1, 180.0
7, brown, 18, 65.0, 1, 182.0
8, candy, 17, 56.0, 0, 166.0
9, cici, 19, 55.0, 0, 168.0

```