**实验一 HDFS操作与RDD操作**

**【实验目的】**

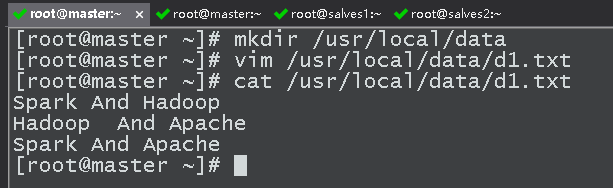
**1．熟练掌握HDFS命令；**

**2．掌握RDD的三种创建方法。**

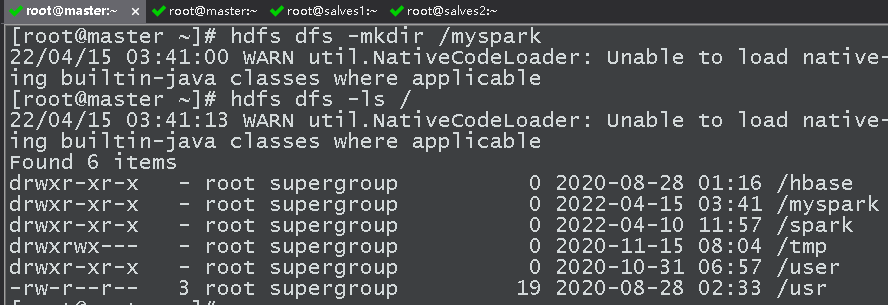
**【实验内容】**

**一、HDFS的操作：**

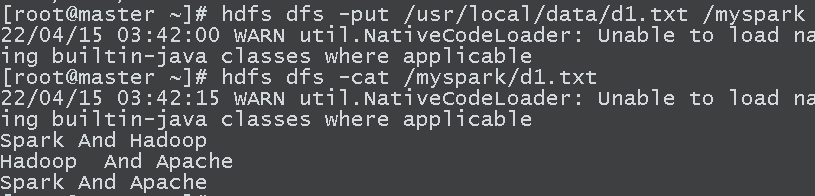
**1、在虚拟机本地盘/usr/local下创建文件夹data，在此创建一个名为才d1.txt的文件，第一行：“Spark And Hadoop”、第二行：“Hadoop  And Apache”、第三行：“Spark And Apache”;**

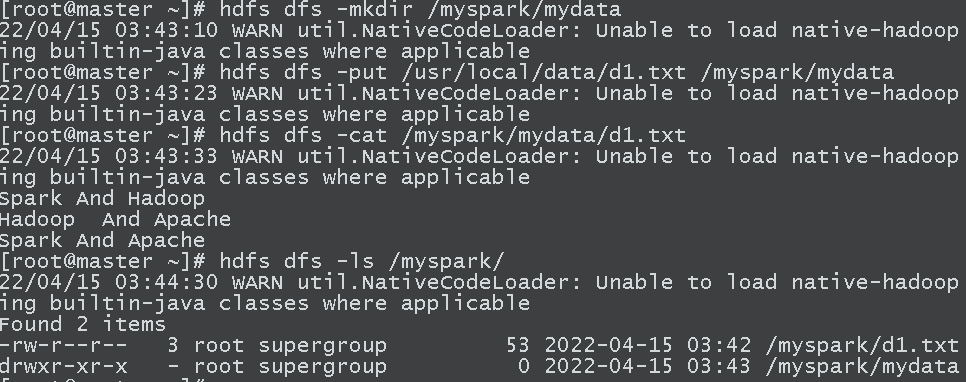


**2、在hdfs中新建目录myspark，并通过命令查看创建结果；**

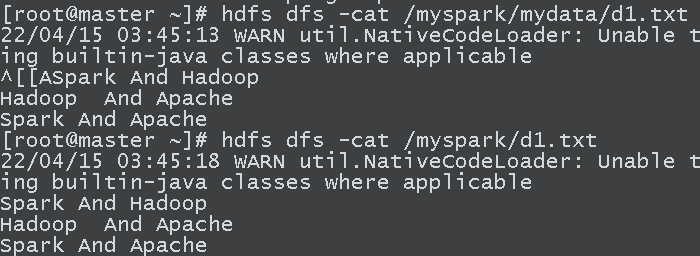


**3、将本地盘/usr/local/data下方的d1.txt上传到hdfs的myspark目录下方，并查看结果；**

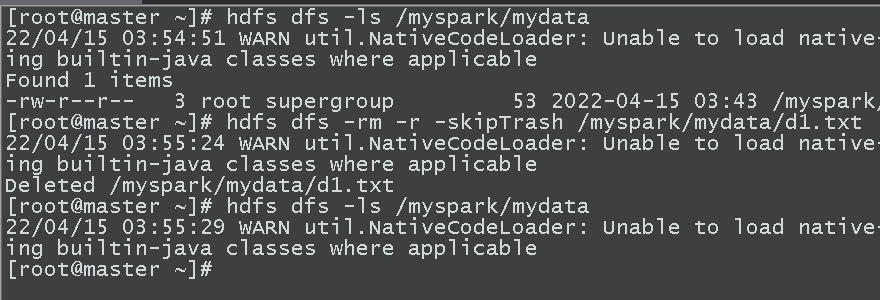


**4、在hdfs中的目录spark下方再创建二级目录mydata，并将本地磁盘的d1.txt文件上传至此，通过命令查看结果；** 

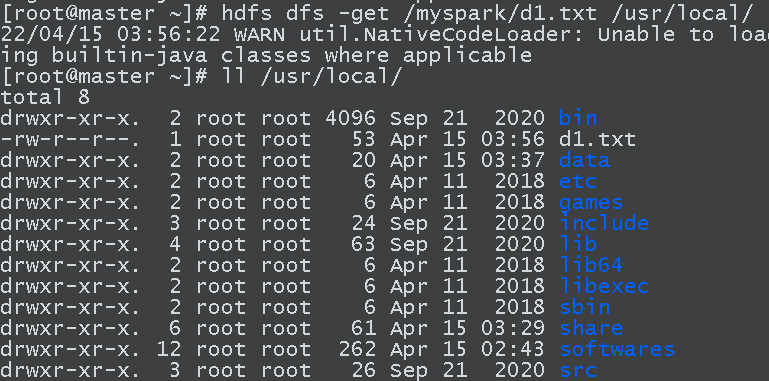
**5、在hdfs上查看d1.txt文件的内容，通过命令查看结果；**



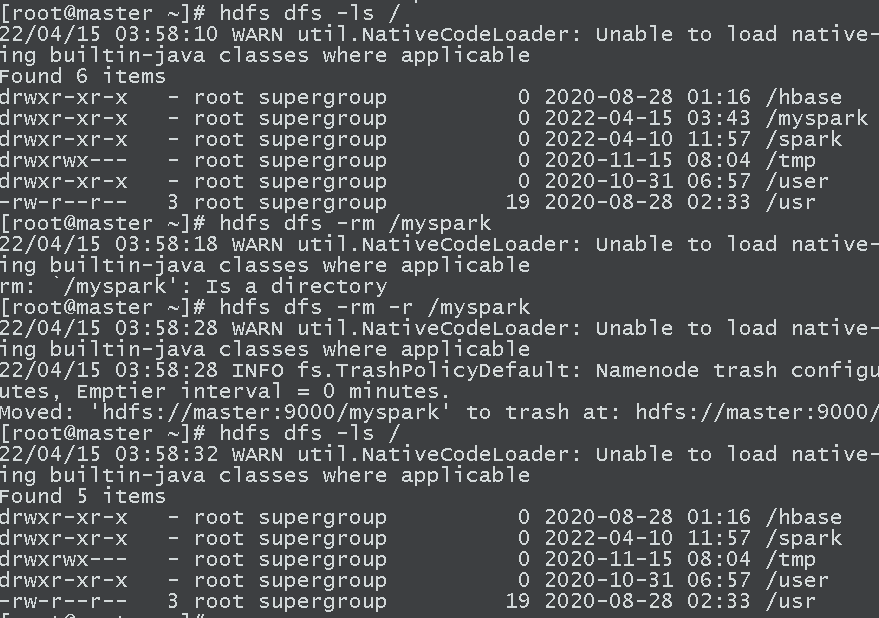
**6、将hdfs上mydata目录中的d1.txt文件删除，通过命令查看结果；**



**7、将hdfs上myspark目录中的d1.txt下载到虚拟机本地盘/usr/local/下方，并查看结果；**

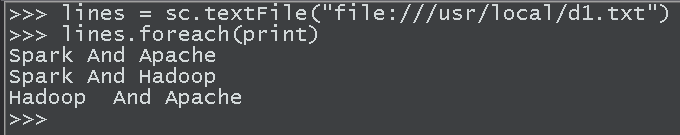


**8、删除hdfs上创建的目录myspark，通过命令查看结果。**

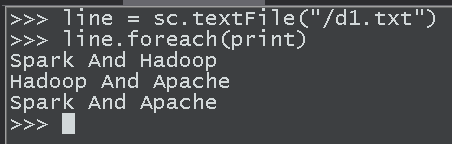


**二、spark之RDD的创建。**

**1、将本地文件d1.txt转换成RDD，并打印输出其结果；**



**2、将hdfs上文件d1.txt转换成RDD，并打印输出其结果；**



**3、使用列表形式创建RDD（列表内容为20以内的偶数）**

