

存储越困难，提取越容易

## HDFS客户端操作---开发环境准备

### 步骤一：编译对应HadoopJar包，配置Hadoop变量

hadoop > hadoop-2.7.2		搜索"hadoop-2.7.2"	
名称	修改日期	类型	大小
bin			
etc			
include			
lib			
libexec			
sbin			
share			
LICENSE.txt	2016/1/26 8:20	文本文档	16 KB
NOTICE.txt	2016/1/26 8:20	文本文档	1 KB
README.txt	2016/1/26 8:20	文本文档	2 KB

### 步骤二：创建Maven工程，导入pom依赖

```
<dependencies>
  <dependency>
    <groupId>junit</groupId>
    <artifactId>junit</artifactId>
    <version>RELEASE</version>
  </dependency>
  <dependency>
    <groupId>org.apache.logging.log4j</groupId>
    <artifactId>log4j-core</artifactId>
    <version>2.8.2</version>
  </dependency>
  <dependency>
    <groupId>org.apache.hadoop</groupId>
    <artifactId>hadoop-common</artifactId>
    <version>2.7.2</version>
  </dependency>
  <dependency>
    <groupId>org.apache.hadoop</groupId>
    <artifactId>hadoop-client</artifactId>
    <version>2.7.2</version>
  </dependency>
  <dependency>
    <groupId>org.apache.hadoop</groupId>
    <artifactId>hadoop-hdfs</artifactId>
    <version>2.7.2</version>
  </dependency>
</dependencies>
```

### 步骤三：创建日志文件

在resources目录下创建log4j.properties，输入如下信息

```
log4j.rootLogger=INFO, stdout
log4j.appender.stdout=org.apache.log4j.ConsoleAppender
log4j.appender.stdout.layout=org.apache.log4j.PatternLayout
log4j.appender.stdout.layout.ConversionPattern=%d %p [%c] - %m%n
log4j.appender.logfile=org.apache.log4j.FileAppender
log4j.appender.logfile.File=target/spring.log
log4j.appender.logfile.layout=org.apache.log4j.PatternLayout
log4j.appender.logfile.layout.ConversionPattern=%d %p [%c] - %m%n
```

## 步骤四：创建HDFSClient类

```
public class HDFSClient{

@Test
public void testMkdirs() throws IOException, InterruptedException,
URISyntaxException{

    // 1 获取文件系统
    Configuration configuration = new Configuration();

    // 2 连接集群
    FileSystem fs = FileSystem.get(new URI("hdfs://hadoop102:9000"),
configuration, "zhutiansama");

    // 3 创建目录
    fs.mkdirs(new Path("/1108/daxian/banzhang"));

    // 4 关闭资源
    fs.close();
}
}
```

客户端去操作HDFS时，是有一个用户身份的。默认情况下，HDFS客户端API会从JVM中获取一个参数来作为自己的用户身份：-DHADOOP\_USER\_NAME=zhutiansama，zhutiansama为用户名称。至此Hadoop的Win开发环境搭建完成，大家可以通过API去操作Hadoop啦【有问题都可以私聊我WX：focusbigdata，或者关注我的公众号：FocusBigData，注意大小写】