

3.3.1 中的 Shell 命令

ls

查看指定路径的当前目录结构

```
[root@node01 ~]# hadoop fs -ls /
Found 16 items
drwxr-xr-x - abc supergroup 0 2019-07-18 17:02 /abc
drwxr-xr-x - root supergroup 0 2019-07-22 11:32 /avro
drwxr-xr-x - root supergroup 0 2019-07-22 15:42 /flume01
drwxr-xr-x - root supergroup 0 2019-07-21 16:08 /hivedat
as
drwxr-xr-x - root supergroup 0 2019-07-20 15:37 /in
drwxr-xr-x - root supergroup 0 2019-07-20 10:42 /out
drwxr-xr-x - root supergroup 0 2019-07-20 15:41 /out1
drwxr-xr-x - root supergroup 0 2019-07-22 09:30 /spooldi
```

Mkdir

在指定路径下创建子目录

```
[root@node01 ~]# hadoop fs -mkdir -p /itcast/hadoop
You have new mail in /var/spool/mail/root
[root@node01 ~]#
```

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	abc	supergroup	0 B	Thu Jul 18 17:02:22 +0800 2019	0	0 B	abc
drwxr-xr-x	root	supergroup	0 B	Mon Jul 22 11:32:32 +0800 2019	0	0 B	avro
drwxr-xr-x	root	supergroup	0 B	Mon Jul 22 15:42:18 +0800 2019	0	0 B	flume01
drwxr-xr-x	root	supergroup	0 B	Sun Jul 21 16:08:18 +0800 2019	0	0 B	hivedatas
drwxr-xr-x	root	supergroup	0 B	Sat Jul 20 15:37:19 +0800 2019	0	0 B	in
drwxr-xr-x	root	supergroup	0 B	Thu Sep 26 14:53:36 +0800 2019	0	0 B	itcast

Browse Directory

/itcast							Go!
Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	root	supergroup	0 B	Thu Sep 26 14:53:36 +0800 2019	0	0 B	hadoop

Put

将本地系统的文件或文件夹复制到 HDFS 上

```
[root@node01 ~]# hadoop fs -put -f install.log /
[root@node01 ~]#
```



Browse Directory

/

Go!

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	abc	supergroup	0 B	Thu Jul 18 17:02:22 +0800 2019	0	0 B	abc
drwxr-xr-x	root	supergroup	0 B	Mon Jul 22 11:32:32 +0800 2019	0	0 B	avro
drwxr-xr-x	root	supergroup	0 B	Mon Jul 22 15:42:18 +0800 2019	0	0 B	flume01
drwxr-xr-x	root	supergroup	0 B	Sun Jul 21 16:08:18 +0800 2019	0	0 B	hivedatas
drwxr-xr-x	root	supergroup	0 B	Sat Jul 20 15:37:19 +0800 2019	0	0 B	in
-rw-r--r--	root	supergroup	27.85 KB	Thu Sep 26 14:57:26 +0800 2019	3	128 MB	install.log

3.4.2 中的代码

Pom.xml

```
<project xmlns="http://maven.apache.org/POM/4.0.0"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://maven.apache.org/POM/4.0.0
http://maven.apache.org/xsd/maven-4.0.0.xsd">

    <modelVersion>4.0.0</modelVersion>

    <groupId>com.itcast</groupId>

    <artifactId>HadoopDemo</artifactId>

    <version>0.0.1-SNAPSHOT</version>

    <dependencies>

        <dependency>

            <groupId>org.apache.hadoop</groupId>

            <artifactId>hadoop-common</artifactId>

            <version>2.7.4</version>
```

```
</dependency>
```

```
<dependency>
```

```
<groupId>org.apache.hadoop</groupId>
```

```
<artifactId>hadoop-hdfs</artifactId>
```

```
<version>2.7.4</version>
```

```
</dependency>
```

```
<dependency>
```

```
<groupId>org.apache.hadoop</groupId>
```

```
<artifactId>hadoop-client</artifactId>
```

```
<version>2.7.4</version>
```

```
</dependency>
```

```
<dependency>
```

```
<groupId>org.apache.hadoop</groupId>
```

```
<artifactId>hadoop-mapreduce-client-  
core</artifactId>
```

```
<version>2.7.4</version>
```

```
</dependency>
```

```
<dependency>
```

```
<groupId>junit</groupId>
```

```
<artifactId>junit</artifactId>
```

```
<version>4.12</version>
```

```
</dependency>
```

```
</dependencies>
```

```
</project>
```

HDFS_CRUD.java

初始化客户端对象

```
public class HDFS_CRUD {
    FileSystem fs = null;

    @Before
    public void init() throws Exception {
        // 构造一个配置参数对象, 设置一个参数: 我们要访问的hdfs的URI
        Configuration conf = new Configuration();
        // 这里指定使用的是HDFS文件系统
        conf.set("fs.defaultFS", "hdfs://192.168.52.100:8020");
        // 通过如下的方式进行客户端身份的设置
        System.setProperty("HADOOP_USER_NAME", "root");
        // 通过FileSystem的静态方法获取文件系统客户端对象
        fs = FileSystem.get(conf);
    }
}
```

上传文件到 hdfs

```
@Test
public void testAddFileToHdfs() throws IOException {
    // 要上传的文件所在本地路径
    Path src = new Path(pathString: "E:/test.txt");
    // 要上传到hdfs的目标路径
    Path dst = new Path(pathString: "/testFile");
    // 上传文件方法
    fs.copyFromLocalFile(src, dst);
    // 关闭资源
    fs.close();
}

// 从hdfs中复制文件到本地文件系统
@Test
HDFS_CRUD -> testAddFileToHdfs()

1 test passed - 2s 109ms
at org.apache.hadoop.fs.FileSystem$Cache.get(FileSystem.java:2684)
at org.apache.hadoop.fs.FileSystem.get(FileSystem.java:373)
at org.apache.hadoop.fs.FileSystem.get(FileSystem.java:172)
at com.itcast.hdfsdemo.HDFS_CRUD.init(HDFS_CRUD.java:28) <21 internal calls> <1 internal cal

Process finished with exit code 0
```

drwxrwxrwx	root	supergroup	0 B	Wed Jul 24 08:54:50 +0800 2019	0	0 B	sqoop
drwxrwxrwx	root	supergroup	0 B	Thu Jul 18 09:53:08 +0800 2019	0	0 B	test
-rw-r--r--	root	supergroup	4.12 KB	Thu Sep 26 15:07:20 +0800 2019	3	128 MB	testFile
drwxrwxrwx	root	supergroup	0 B	Sun Jul 21 09:34:50 +0800 2019	0	0 B	tmp
drwxrwxrwx	root	supergroup	0 B	Sun Jul 21 15:48:01 +0800 2019	0	0 B	user
drwxrwxrwx	root	supergroup	0 B	Mon Jul 22 11:04:57 +0800 2019	0	0 B	weblog

从 hdfs 下载文件到本地

```
// 从hdfs中复制文件到本地文件系统
@Test
public void testDownloadFileToLocal() throws IllegalArgumentException, IOException {
    // 下载文件
    fs.copyToLocalFile(new Path(pathString: "/testFile"), new Path(pathString: "E:/"));
}

HDFS_CRUD > testDownloadFileToLocal()

1 test passed - 2s 516ms
at org.apache.hadoop.util.Shell.<init>(Shell.java:390)
at org.apache.hadoop.util.StringUtils.<clinit>(StringUtils.java:80)
at org.apache.hadoop.fs.FileSystem$Cache$Key.<init>(FileSystem.java:2823)
at org.apache.hadoop.fs.FileSystem$Cache$Key.<init>(FileSystem.java:2818)
at org.apache.hadoop.fs.FileSystem$Cache.get(FileSystem.java:2684)
at org.apache.hadoop.fs.FileSystem.get(FileSystem.java:373)
at org.apache.hadoop.fs.FileSystem.get(FileSystem.java:172)
at com.itcast.hdfsdemo.HDFS_CRUD.init(HDFS_CRUD.java:28) <21 internal calls> <1 internal call>

process finished with exit code 0
```

电脑 > Workspaces (E:)				搜索"Workspaces (E:)"
名称	修改日期	类型	大小	
test.txt	2019/9/24 17:19	文本文档	5 KB	
testFile	2019/9/26 15:09	文件	5 KB	
.testFile.crc	2019/9/26 15:09	CRC 文件	1 KB	

目录操作

```
// 创建, 删除, 重命名文件
@Test
public void testMkdirAndDeleteAndRename() throws Exception {
    // 创建目录
    fs.mkdirs(new Path(pathString: "/a/b/c"));
    fs.mkdirs(new Path(pathString: "/a2/b2/c2"));
    // 重命名文件或文件夹
    fs.rename(new Path(pathString: "/a"), new Path(pathString: "/a3"));
    // 删除文件夹, 如果是非空文件夹, 参数2必须给值true
    fs.delete(new Path(pathString: "/a2"), b: true);
}

HDFS_CRUD > testDownloadFileToLocal()

1 test passed - 1s 773ms
at org.apache.hadoop.fs.FileSystem$Cache$Key.<init>(FileSystem.java:2818)
at org.apache.hadoop.fs.FileSystem$Cache.get(FileSystem.java:2684)
at org.apache.hadoop.fs.FileSystem.get(FileSystem.java:373)
at org.apache.hadoop.fs.FileSystem.get(FileSystem.java:172)
at com.itcast.hdfsdemo.HDFS_CRUD.init(HDFS_CRUD.java:28) <21 internal calls> <1 intern
ocess finished with exit code 0
```

/								Go!
Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
drwxr-xr-x	root	supergroup	0 B	Thu Sep 26 15:12:51 +0800 2019	0	0 B	a3	
drwxrwxrwx	abc	supergroup	0 B	Thu Jul 18 17:02:22 +0800 2019	0	0 B	abc	

查看目录信息

```
@test
public void testListFiles() throws FileNotFoundException, IllegalAccessException {
    // 获取迭代器对象
    RemoteIterator<LocatedFileStatus> listFiles = fs.listFiles(
    while (listFiles.hasNext()) {
        LocatedFileStatus fileStatus = listFiles.next();
        // 打印当前文件名
        System.out.println(fileStatus.getPath().getName());
        // 打印当前文件块大小
        System.out.println(fileStatus.getBlockSize());
        // 打印当前文件权限
    }
}
```

HDFS_CRUD > testListFiles()

1 test passed - 2s 303ms

```
access_log.1563764757395
134217728
rwxrwxrwx
3654
block-length:3654--block-offset:0
node01.hadoop.com
-----
```