

31 压缩文件，解压无压力

更新时间：2019-07-31 09:22:30



“

勤学如春起之苗，不见其增，日有所长。

——陶潜

”

1. 前言
2. tar 命令：将多个文件归档
3. gzip 和 bzip2 命令：压缩归档
4. zip / unzip 和 rar / unrar 命令：压缩 / 解压 zip 和 rar 文件
5. 总结
6. 第四部分第二课预告

1. 前言

上一课我们完成了第三部分的测验：[带你玩转Linux和Shell编程](#) | [第三部分测试题](#)。恭喜大家完成前三部分的学习，进入新的旅程了！

第四部分我们会讲 Linux 网络方面的一些知识点。首先我们会介绍一下如何压缩和解压文件，因为我们在网络上传输文件的时候，通常会先将其压缩。

说到压缩文件，如果是 Windows 用户，那应该对 rar 和 zip 两种格式不陌生。一般在 Windows 下，我们会使用 WinRAR，7-Zip 等压缩解压的软件。这些软件中有些是收费的，比如 WinRAR。

在 Linux 下，我们有很多免费的压缩解压软件可供选择，比如 gzip 和 bzip2。不过，gzip 和 bzip2，不能同时压缩多个文件，一次只能压缩一个文件。

但是，Linux 的开源社群可厉害了，他们还有一个软件 tar，算是 Linux 下用得挺多的程序了吧。tar 可以同时压缩多个文件。

当然了，Linux 下也有解压 zip 和 rar 格式的软件。

这一课没什么技术难点，不过有些概念还是比较重要的。就当"茶余饭后，休闲娱乐"吧。

2. tar 命令：将多个文件归档

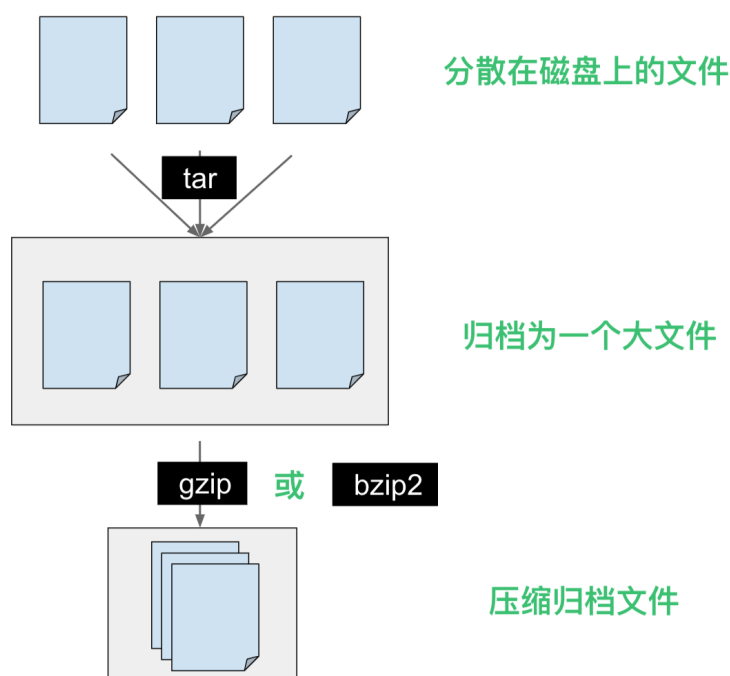
首先，我们需要弄清楚两个概念：打包和压缩。

- 打包：是将多个文件变成一个总的文件，这个总的文件我们通常称为 **archive**，是英语“存档，归档”的意思。
- 压缩：是将一个大文件通过某些压缩算法变成一个小文件。

所以，其实引言中所提到的 tar 命令，gzip 和 bzip2 命令是这样使用的：

1. 用 tar 将多个文件归档为一个总的文件，称为 archive。
2. 用 gzip 或 bzip2 命令将 archive 压缩为更小的文件。

可以用下图来描述此过程：



所以，我们会首先介绍 tar 命令，然后再学习 gzip 和 bzip2 命令。当然了，我们也会学习如何在 Linux 下解压 zip 和 rar 格式的文件。

zip 和 rar 格式是可以同时归档和压缩的。

将多个文件合并为一个文件

我们来实际操作一下。

为了不和家目录其他文件搞混，我们先创建一个目录，叫做 **compression**，是英语“压缩”的意思。然后在此文件夹中再创建一个目录，叫做 **sorting**，是英语“排序”的意思。

在 **sorting** 目录中，我放入三个 Java 语言的源文件（MergeSort.java, QuickSort.java, SelectionSort.java。分别是“归并排序”，“快速排序”和“选择排序”的算法实现）。如下图所示：

```
oscar@oscar-laptop: ~/compression/sorting$ ls
MergeSort.java QuickSort.java SelectionSort.java
oscar@oscar-laptop: ~/compression/sorting$
```

当然了, 你可以给目录起不同的名字, 并且随便放什么文件进去。

接下来, 我们就可以用 **tar** 命令对这三个文件进行归档。

-cvf: 创建一个 **tar** 归档

我们可以这样来创建一个 **tar** 归档:

```
tar -cvf sorting.tar sorting/
```

tar 命令的这三个参数分别表示:

- **c**: **create** 的缩写, 表示“创建”。
- **v**: **verbose** 的缩写, 表示“冗余”。会显示操作的细节。
- **f**: **file** 的缩写, 表示“文件”。指定归档文件。

执行以上的命令之后, 可以从 **sorting** 这个文件夹 (包含 **MergeSort.java**, **QuickSort.java**, **SelectionSort.java** 三个文件) 创建 **sorting.tar** 这个归档。

```
oscar@oscar-laptop: ~/compression$ tar -cvf sorting.tar sorting/
sorting/
sorting/MergeSort.java
sorting/QuickSort.java
sorting/SelectionSort.java
oscar@oscar-laptop: ~/compression$ ls
sorting  sorting.tar
oscar@oscar-laptop: ~/compression$
```

因为加了 **v** 选项, 所以列出了被归档的三个文件。

难道我们一定要先把三个文件都放到 **sorting** 这样的一个目录, 才能用 **tar** 归档吗?

不是的, 我们也可以直接对文件进行归档操作, 不需要先把所有文件放到一个目录里, 如下:

```
tar -cvf archive.tar file1.txt file2.txt file3.txt
```

就会把 **file1.txt**, **file2.txt**, **file3.txt** 归档为 **archive.tar**。

-tf: 显示归档里的内容, 并不解开归档

我们可以用 **tf** 选项参数来看一下我们刚生成的归档 **sorting.tar** 中的内容。

```
tar -tf sorting.tar
```

```
File Edit View Search Terminal Help
oscar@oscar-laptop: ~/compression
oscar@oscar-laptop:~/compression$ tar -tf sorting.tar
sorting/
sorting/MergeSort.java
sorting/QuickSort.java
sorting/SelectionSort.java
oscar@oscar-laptop:~/compression$
```

可以看到，`sorting.tar` 这个文件里，首先包含的是 `sorting` 这个目录，目录里有三个文件：`MergeSort.java`，`QuickSort.java`，`SelectionSort.java`

一般在建立归档之前，最好先把所有文件放置到一个目录里。因为假如你拿到的归档文件是不包含目录的（如上面 `tar -cvf archive.tar file1.txt file2.txt file3.txt` 这样），并且里面有成百上千个文件，你将其解压到你的家目录，那会乱成一团的。

-rvf: 追加文件到归档

例如：

```
tar -rvf archive.tar file_extra.txt
```

表示将 `file_extra.txt` 这个文件添加到 `archive.tar` 归档中。

-xvf: 解开归档

算是 `-cvf` 的相反操作。将生成的归档解开。`x` 是 `extract` 的缩写，表示“提取，取出”。

```
tar -xvf sorting.tar
```

```
File Edit View Search Terminal Help
oscar@oscar-laptop: ~/compression
oscar@oscar-laptop:~/compression$ tar -xvf sorting.tar
sorting/
sorting/MergeSort.java
sorting/QuickSort.java
sorting/SelectionSort.java
oscar@oscar-laptop:~/compression$ ls sorting
MergeSort.java QuickSort.java SelectionSort.java
```

3. gzip 和 bzip2 命令：压缩归档

现在我们已经学会用 `tar` 命令来创建一个归档（`archive`）了，接下来，我们就学习如何压缩这个归档吧。

Linux 中常用的压缩命令是 `gzip` 和 `bzip2`，它们的功能类似。

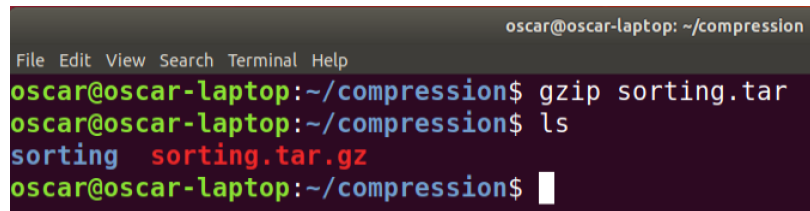
- `gzip`: 比较常用。
- `bzip2`: 不是那么常用。压缩率比 `gzip` 更大，但是更耗时。

这两个命令的用法也很简单，只需传入要压缩的归档作为参数。经过压缩，归档的后缀名也会改变。

- `.tar.gz`: 这是用 `gzip` 命令压缩后的文件后缀名。
- `.tar.bz2`: 这是用 `bzip2` 命令压缩后的文件后缀名。

两个压缩命令的用法如下：

```
gzip sorting.tar
```

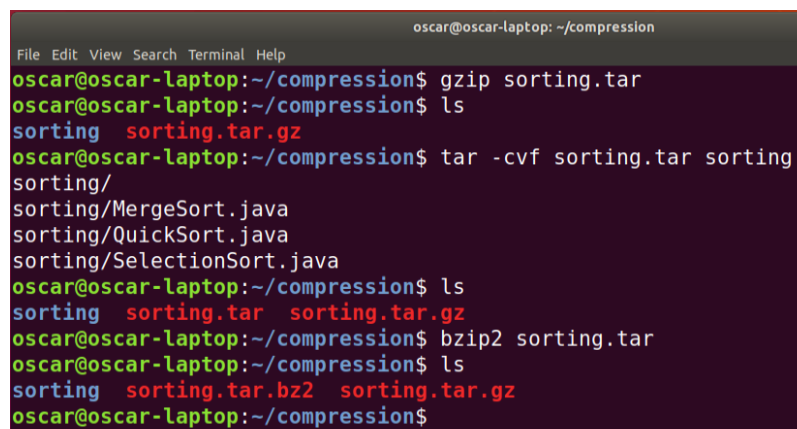


```
oscar@oscar-laptop: ~/compression
File Edit View Search Terminal Help
oscar@oscar-laptop:~/compression$ gzip sorting.tar
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.gz
oscar@oscar-laptop:~/compression$
```

可以看到，用 `gzip` 命令来将归档 `sorting.tar` 压缩后，是直接将 `sorting.tar` 文件改变为 `sorting.tar.gz` 压缩文件。所以为了演示 `bzip2` 命令，我们再生成一次 `sorting.tar`。

```
tar -cvf sorting.tar sorting
```

```
bzip2 sorting.tar
```

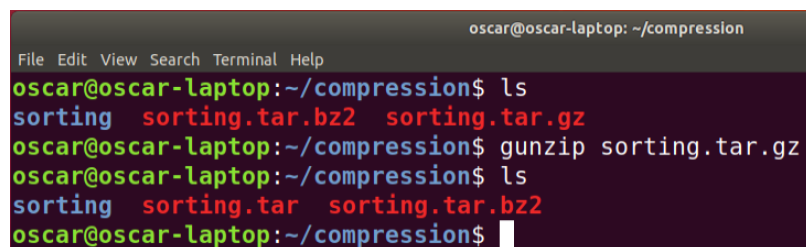


```
oscar@oscar-laptop: ~/compression
File Edit View Search Terminal Help
oscar@oscar-laptop:~/compression$ gzip sorting.tar
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.gz
oscar@oscar-laptop:~/compression$ tar -cvf sorting.tar sorting
sorting/
sorting/MergeSort.java
sorting/QuickSort.java
sorting/SelectionSort.java
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar  sorting.tar.gz
oscar@oscar-laptop:~/compression$ bzip2 sorting.tar
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.bz2  sorting.tar.gz
oscar@oscar-laptop:~/compression$
```

可以看到，`bzip2` 会将 `sorting.tar` 压缩为 `sorting.tar.bz2` 压缩文件。

如果要对这两个命令生成的压缩文件分别进行解压，可以使用 `gunzip` 和 `bunzip2` 两个命令：

```
gunzip sorting.tar.gz
```



```
oscar@oscar-laptop: ~/compression
File Edit View Search Terminal Help
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.bz2  sorting.tar.gz
oscar@oscar-laptop:~/compression$ gunzip sorting.tar.gz
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar  sorting.tar.bz2
oscar@oscar-laptop:~/compression$
```

```
bunzip2 sorting.tar.bz2
```

```
oscar@oscar-laptop: ~/compression
File Edit View Search Terminal Help
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar  sorting.tar.bz2
oscar@oscar-laptop:~/compression$ bunzip2 sorting.tar.bz2
bunzip2: Output file sorting.tar already exists.
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar  sorting.tar.bz2
oscar@oscar-laptop:~/compression$ rm sorting.tar
oscar@oscar-laptop:~/compression$ bunzip2 sorting.tar.bz2
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar
oscar@oscar-laptop:~/compression$
```

用 `bunzip2 sorting.tar.bz2` 来解压时，因为之前用 `gunzip sorting.tar.gz` 已经解压出 `sorting.tar` 这个归档了，所以显示 “Output file sorting.tar already exists”，意思是“输出文件 `sorting.tar` 已经存在”。

我们只需要把 `sorting.tar` 删除，再运行 `bunzip2 sorting.tar.bz2` 即可。

用 `tar` 命令同时归档和压缩

之前我们只介绍了 `tar` 命令的归档作用，其实我们可以用 `tar` 命令同时完成归档和压缩的操作，一气呵成。

当然了，也就是给 `tar` 命令多加一个选项参数，使之完成归档操作后，还是调用 `gzip` 或 `bzip2` 命令来完成压缩操作。

-zcvf: 归档，然后用 `gzip` 来压缩归档

为了演示清楚，我们先把 `sorting.tar` 归档删除。然后再运行：

```
tar -zcvf sorting.tar.gz sorting
```

```
oscar@oscar-laptop: ~/compression
File Edit View Search Terminal Help
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar
oscar@oscar-laptop:~/compression$ rm sorting.tar
oscar@oscar-laptop:~/compression$ ls
sorting
oscar@oscar-laptop:~/compression$ tar -zcvf sorting.tar.gz sorting
sorting/
sorting/MergeSort.java
sorting/QuickSort.java
sorting/SelectionSort.java
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.gz
oscar@oscar-laptop:~/compression$
```

一步到位，从 `sorting` 目录，归档、压缩，生成 `sorting.tar.gz` 文件。

要解压，也可以一步到位。为了演示清楚，我们先把 `sorting` 目录删除。然后再运行：

```
tar -zxvf sorting.tar.gz
```

```

oscar@oscar-laptop: ~/compression
File Edit View Search Terminal Help
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.gz
oscar@oscar-laptop:~/compression$ rm -rf sorting
oscar@oscar-laptop:~/compression$ ls
sorting.tar.gz
oscar@oscar-laptop:~/compression$ tar -zxvf sorting.tar.gz
sorting/
sorting/MergeSort.java
sorting/QuickSort.java
sorting/SelectionSort.java
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.gz
oscar@oscar-laptop:~/compression$

```

-jcvf: 归档，然后用 **bzip2** 来压缩归档

```
tar -jcvf sorting.tar.bz2 sorting/
```

```

oscar@oscar-laptop: ~/compression
File Edit View Search Terminal Help
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.gz
oscar@oscar-laptop:~/compression$ tar -jcvf sorting.tar.bz2 sorting/
sorting/
sorting/MergeSort.java
sorting/QuickSort.java
sorting/SelectionSort.java
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.bz2  sorting.tar.gz
oscar@oscar-laptop:~/compression$

```

一步到位，从 **sorting** 目录，归档、压缩，生成 **sorting.tar.bz2** 文件。

要解压，也可以一步到位。为了演示清楚，我们先把 **sorting** 目录删除。然后再运行：

```
tar -jxvf sorting.tar.bz2
```

```

oscar@oscar-laptop: ~/compression
File Edit View Search Terminal Help
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.bz2  sorting.tar.gz
oscar@oscar-laptop:~/compression$ rm -rf sorting
oscar@oscar-laptop:~/compression$ ls
sorting.tar.bz2  sorting.tar.gz
oscar@oscar-laptop:~/compression$ tar -jxvf sorting.tar.bz2
sorting/
sorting/MergeSort.java
sorting/QuickSort.java
sorting/SelectionSort.java
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.bz2  sorting.tar.gz
oscar@oscar-laptop:~/compression$

```

zcat / bzcat, zmore / bzmores, zless / bzless : 显示用 **gzip / bzip2** 压缩的文件的内容

cat, more 和 **less** 这三个命令我们已经很熟悉了，是用于显示文件内容。

如果是压缩文件，比如我们上面生成的 **sorting.tar.gz**，如果直接用 **cat, more** 或 **less** 命令来显示，会显示乱码：

```
cat sorting.tar.gz
```



```
File Edit View Search Terminal Help
oscar@oscar-laptop: ~/compression
oscar@oscar-laptop:~/compression$ cat sorting.tar.gz
g
I0I80
0{0>0-apL00>{0q000_w0[00 0000060600
S000001000x00000600 0000000
0000Q0000s000Ccp00U0000000+9'0C0KM000商Zr00??00a=?0n0T000S0002X0p0_0Y07d0J\000WB
2p0Wf0000e00
00^_00Z000-0004\DW=0[0W0$000Y0000e:0
00M0&y]100000fDj0;{00-q001Z000$;0"G
0mg0S00070$ 0_0000000000pE@000/_0000[00+000s0
H00c0\<0[0 0600S00
000000I0#0Z000VmA0GWZt~ 0cH\0e0Fu^00t&0k0NlX000000p00p0600.#c0000000,a0=00VL00R0
00B00Ak0l0100X0"\0.0:0Z0000000w-0g0000\0080:0
04'0d0M>0|0
00/B00A0D0k0
00^02iE]00002
/x0M00C%00000x000y0Q; )00000000-0000'
0S0000p0yyN0#00000N0@K00X|oC
00000Dh!500000C0|0h]0J0|00n0020srrm8#0|t700000000J00;z=00=%0p
z00000C00000020f0a0000c0=0000001R000\j0{00S000r078006?q00'0x0'TZ00%k?";0000hK0
0ad0!0000000E00k0m0300%7000h
2kq{UXUj0Be9000 0:0000
0.90>00T0'0`nk000000v0+0000nH0(000Ga20Fe70R00 00
```

但是用 `zcat`, `zmore`, `zless` 命令就可以显示用 `gzip` 命令压缩的文件的内容了：

```
zcat sorting.tar.gz
```

```
File Edit View Search Terminal Help
oscar@oscar-laptop: ~/compression
oscar@oscar-laptop:~/compression$ zcat sorting.tar.gz
sorting/000075500017500001750000000000013471277317011557 5ustar  oscaroscarsortin
g/MergeSort.java0000644000175000017500000000544413471277317014340 0ustar  oscarosc
arpackage Sorts;

import static Sorts.SortUtils.print;

/**
 * This method implements the Generic Merge Sort
 *
 * @author Varun Upadhyay (https://github.com/varunu28)
 * @author Podshivalov Nikita (https://github.com/nikitap492)
 * @see SortAlgorithm
 */

class MergeSort implements SortAlgorithm {

    /**
     * This method implements the Generic Merge Sort
     *

```

要显示 `bzip2` 命令压缩的文件内容，则用 `bzcat`, `bzmore`, `bzless` 命令。例如：

```
bzcat sorting.tar.bz2
```

4. zip / unzip 和 rar / unrar 命令：压缩 / 解压 zip 和 rar 文件

尽管 `.tar.gz` 和 `.tar.bz2` 结尾的压缩文件在 Linux 世界很常见，但是如果某个 Windows 用户给你发送压缩文件，大部分情况可能是 `.zip` 或 `.rar` 结尾的压缩文件，那么在 Linux 下如何解压这类来自 Windows 世界的“外来生物”呢？

不必担心，我们可以用 `unzip` 和 `unrar` 命令来分别解压 `.zip` 和 `.rar` 格式的压缩文件。

zip / unzip：压缩 / 解压 zip 文件

一般 Linux 发行版（Ubuntu 自带了 `unzip` 这个命令）中默认没有安装这个程序，我们可以用以下命令来安装：

```
sudo apt install unzip # 这是在 Debian 一族中的安装方式
```

安装完之后，要解压 `.zip` 格式的压缩文件，可以这样：

```
unzip archive.zip
```


很简单吧。

如果不想解开 .zip 文件，只想看其中的内容的话，可以加上 -l 参数：

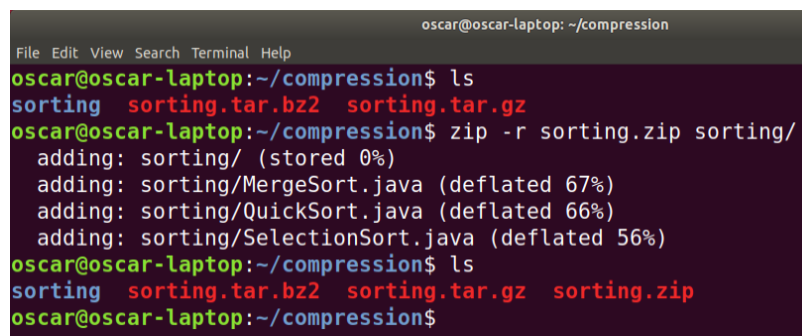
```
unzip -l archive.zip
```

在 Linux 中，有时候我们也需要创建 .zip 格式的文件，我们可以安装 zip 这个程序（Ubuntu 自带了 zip 这个命令）：

```
sudo apt install zip      # 这是在 Debian 一族中的安装方式
```

我们来压缩一下我们的 sorting 目录试试：

```
zip -r sorting.zip sorting/
```



```
oscar@oscar-laptop: ~/compression
File Edit View Search Terminal Help
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.bz2  sorting.tar.gz
oscar@oscar-laptop:~/compression$ zip -r sorting.zip sorting/
adding: sorting/ (stored 0%)
adding: sorting/MergeSort.java (deflated 67%)
adding: sorting/QuickSort.java (deflated 66%)
adding: sorting/SelectionSort.java (deflated 56%)
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.bz2  sorting.tar.gz  sorting.zip
oscar@oscar-laptop:~/compression$
```

注意：这里有一个 -r 参数（r 是 recursively 的缩写，表示“递归地”），须要加上。如果不加，则只会压缩空文件夹而已。加上 -r 参数，则会递归压缩目录和子目录中的所有文件。

rar / unrar：压缩 / 解压 rar 文件

类似于 unzip，安装 unrar 的命令如下（Ubuntu 一般没有自带 unrar 这个命令）：

```
sudo apt install unrar    # 这是在 Debian 一族中的安装方式
```

安装完之后，要解压 .rar 格式的压缩文件，可以这样：

```
unrar e archive.rar
```

可以看到，e 这个选项参数（e 是 extract 的缩写，表示“提取，取出”）前面没有 -（短横）。

如果不想解开 .rar 文件，只想看其中的内容的话，可以加上 l 参数（是的，也不加 -）：

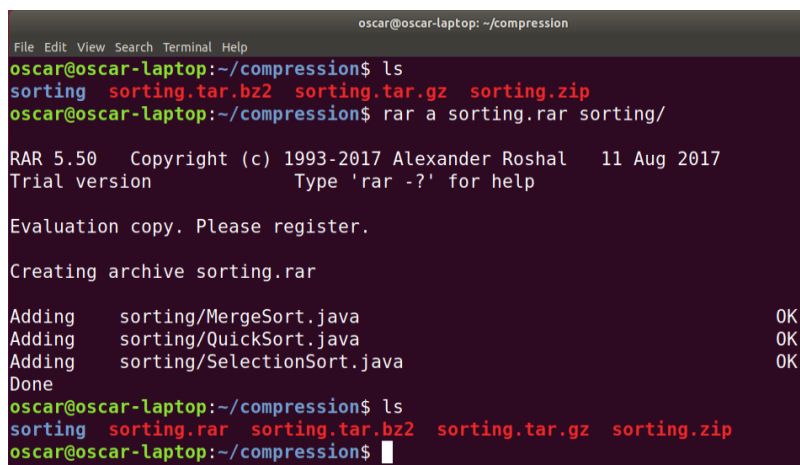
```
unrar l archive.rar
```

在 Linux 中，有时候我们也需要创建 .rar 格式的文件，我们可以安装 rar 这个程序（Ubuntu 一般没有自带 rar 这个命令）：

```
sudo apt install rar      # 这是在 Debian 一族中的安装方式
```

我们来压缩一下我们的 `sorting` 目录试试：

```
rar a sorting.rar sorting/
```



```
File Edit View Search Terminal Help
oscar@oscar-laptop: ~/compression
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.tar.bz2  sorting.tar.gz  sorting.zip
oscar@oscar-laptop:~/compression$ rar a sorting.rar sorting/

RAR 5.50 Copyright (c) 1993-2017 Alexander Roshal 11 Aug 2017
Trial version Type 'rar -?' for help

Evaluation copy. Please register.

Creating archive sorting.rar

Adding sorting/MergeSort.java OK
Adding sorting/QuickSort.java OK
Adding sorting/SelectionSort.java OK
Done
oscar@oscar-laptop:~/compression$ ls
sorting  sorting.rar  sorting.tar.bz2  sorting.tar.gz  sorting.zip
oscar@oscar-laptop:~/compression$
```

我们用了 `a` 参数，`a` 是 `add` 的缩写，表示“添加”。

可以看到，`rar` 这个软件不是免费的，我们目前用的是试用版（`Trial version`）：

```
RAR 5.50 Copyright (c) 1993-2017 Alexander Roshal 11 Aug 2017
Trial version Type 'rar -?' for help

Evaluation copy. Please register.
```

你可以购买 `rar`。当然了，我们只是演示一下，暂时用试用版也足够了。

5. 总结

1. 打包：是将多个文件变成一个总的文件，这个总的文件我们通常称为 `archive`，是英语“存档，归档”的意思。Linux 下常用的打包命令是 `tar`。
2. 压缩：是将一个大文件通过某些压缩算法变成一个小文件。Linux 下常用的压缩命令是 `gzip` 和 `bzip2`。
3. Linux 中有多种方式来压缩文件。可以用 `tar` 命令先将文件打包（`tar -cvf archive.tar archive/`），再用 `gzip` 或 `bzip2` 命令来压缩（`gzip archive.tar` 或 `bzip2 archive.tar`）。也可以将文件一次性打包压缩（`tar -zcvf archive.tar.gz archive` 或 `tar -jcvf archive.tar.bz2 archive`）。
4. `zcat`，`zmore`，`zless` 用于显示用 `gzip` 压缩的文件的内容。`bzcat`，`bzmore`，`bzless` 用于显示用 `bzip2` 压缩的文件的内容。
5. `zip` 和 `unzip` 命令分别用于压缩和解压 `zip` 文件。`rar` 和 `unrar` 命令分别用于压缩和解压 `rar` 文件。

今天的课就到这里，一起加油吧！