Help Topic: Creating and Expanding Tar Files

Spring 2020

Michael Benjamin, mikerb@mit.edu
Department of Mechanical Engineering
MIT, Cambridge MA 02139

Creating and Expanding Tar Files

At times it is useful to pass around a whole directory (folder) as a single file, for example, to email it to someone. There are a few ways to do this. Here we discuss one of the oldest and most ubiquitous forms for doing this, using the `tar` utility. This utility is natively available in MacOS (in `/usr/bin/tar`), and Linux (in `/bin/tar`).

As with many command line utilities, you can find much more information by just typing `"man tar"` on the command line. You'll find there are a lot of options. We'll discuss a couple uses to ensure you can `tar` a directory and "untar" a tar'ed file. To `tar` a directory:

```
$ ls
my_directory/
$ tar cvf my_directory.tar my_directory
$ ls
my_directory/  my_directory.tar
```

Note that (a) you should `tar` a directory when the target directory is in your current working directory (don't do it while inside the directory). (b) the tar file is made without affecting the directory itself, i.e., it does not replace the directory. (c) a common convention is to call the tar file with the same name as the directory, but just with the `.tar` suffix. Indeed you can call the tar file anything you want, even excludig the suffix. But it's very poor form to name a tar file with anything but a filename ending in `.tar`.

If you took a look at the manpage for `tar`, you would see that the above uses three separate command line option, '`-c`', '`-v`', '`-f`'. By convention, a few shortcuts are supported. The following are all equivalent invocations:

```
$ tar -c -v -f my_directory.tar my_directory
$ tar -cvf my_directory.tar my_directory
$ tar cvf my_directory.tar my_directory
```

To untar the file:

```
$ ls
my_directory.tar
$ tar xvf my_directory.tar my_directory
$ ls
my_directory/  my_directory.tar
```

Note that the only difference is the `-x` command line option replaces the `'c'` option.

Using Compression Along the Way

Often, if you're tar'ing a folder to send or post, you want to make this file as small as possible for the recipient's email box, or to facilitate the speed of the download. If the directory contains source code or other text files, the compression can be very effective. (If the directory contains images or video, compression will have little effect since images and video are already highly compressed).

To compress while tar'ing use:

```
$ ls
my_directory/
$ tar cvfz my_directory.tgz my_directory
$ ls
my_directory/  my_directory.tgz
```

Note (a) the addition of the `'z'` command line options and (b) the suffix has been changed to `.tgz`. A less-cool suffix is `.tar.gz`.

To untar the compressed file:

```
$ ls
my_directory.tgz
$ tar xvfz my_directory.tgz my_directory
$ ls
my_directory/  my_directory.tgz
```

Note that using the correct file suffix is really important as a hint to the user so they know what command line options are needed to expand the file. (Sending a compressed tar file with a `.tar` suffix and not `.tgz` suffix will be really annoying to the recipient!).

TIP: By the way, if you're wondering why, in my `ls` output above, directories always in the `'/'` character, it is due to two things. (1) I use the `-F` command line switch when invoking `ls`, and (2) I make an alias for ls to make this my default option. Using this visual cue can be super helpful and I recommend trying/adopting it.