Compression for Better Storage

- [Open Source Compression Compared](#)

Open Source Compression Compared

https://www.youtube.com/embed/KB2qJMRbmbg

File compression isn't an overly glamorous topic, but when you stop to think about it, it's a topic well worth discussing. We don't even think about it these days. You download a .zip, .tar, .tar.gz, or .7z file and sometimes your OS (MacOS in particular) will decompress the file into a folder before you can even navigate to it in the file system.

But why compress files at all these days? Drives are larger and less expensive than ever. You can get an 8 Terabyte HDD for just over $100 US.

Even with all of that storage, the truth is, applications are larger, our lives are stored digitally, and making the most out of the space we have is still extremely important.

As we've moved to a more digitized world, we have begun to realize the need for redundancy in the backup of our important photos, videos, music, and documents. It's the "just in case" portion of living in a digital world that keeps file compression so important.

Open Source, as is the case in most of the software world, is at the center of file compression. Today we look at three tools for compressing files and folders that are open source, and available on MacOS, Windows, and Linux.

The Tools:

Zip

Ah yes, good old Zip archiving. If you're a Windows user you'll be very familiar with the zip archive as it's probably what you've seen most of any of the compression formats, though 7zip (furhter down our list) has become more popular over the years, and for good reason.

Tar

If you're a Linux or Mac user, you may be more familiar with the tar file. Tar has multiple extensions, and compression types and levels, and is a tremendous archiving utility.

You've probably seen at least one of the following at some point: .tar.gz, .tar, .tar.xz, .bzip, .bzip2.

7z

And 7-zip. more of a new comer than the other two, but a definite stnad out when it comes to ease of use and compression performance, as you'll see.

Installation

All of these are available freely as open source tools. You can of course get them for Windows (as an .exe), MacOS (as a .dmg or .pkg file), and Linux from your various distro repositories.

For Ubuntu / Debian systems you can install them with

`sudo apt install zip -y`

`sudo apt install tar -y`

`sudo apt install p7zip-full -y`

Usage in the UI

All three major platforms (and BSD as well I'm sure) have the ability to open our file explorer application, select, right-click, and choose the "compress" or "archive" option, and have the selected items put into a compressed archive of some type.

Usage in the CLI (Command Line)

in order to use these tools in the command line, we only need the syntax for each.

Zip

`zip -r <archive_name_you_want>.zip file1 file2 folder1 etc`

The above command will zip up any folder and / or files, as well as the subfolders and files within any specified folders. The `-r` is the recursive flag which tells it to zip everything in a named folder.

You can list as many files and folders after the archive name as you want to include.

Tar

`tar czvf <archive_name_you_want>.tar.gz file1 file2 folder1 etc`

Again, this will compress all of the files and / or folders, as well as the files and folders within any name folders into a `.tar.gz` file archive.
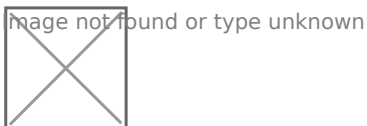
7-zip

`7z a <archive_name_you_want>.7z file1 file2 folder1 etc`

This will archive, thus the `a` flag all of the files and / or folders, as well as the files and folders within any name folders into a `.7z` file archive.

Performance

In the table belwo, you'll see that we have a set of folders, their uncompressed sizes, and then their compressed sizes using Zip, Tar, and 7-zip both in the GUI and the CLI.

Image not found or type unknown

As you can see, the 7-zip utility, from the GUI or the CLI is in general better at compressing files, but does take significantly longer to achieve that level of compression.

Additionally, the `.tar.xz` archiving we get from the GUI compression is very close to on par with the 7-zip archiving, both in size, and in terms of how long it takes to create the more compact archive.

There you have it, 3 great, open source, archiving utilities that you can use in the CLI or the GUI.