Linux - Archives

# Overview

Archives are sets of compressed files which allow for a reduction of space during storage and transport. In linux, files are collected in archive called .tar (tape archive) files, then compressed seperately, commonly using GnuZip into .tar.gz files.

# Tar

## Create

Create a snapshot of a directory using the tar command:

tar -cvf <output-name>.tar <directory>

Then compress the archive with gzip:

gzip <tar-archive>.tar

Tar can be used to combine these tasks using the -z option:

tar -cvzf <output-name>.tgz <directory>

## Extract

Extract files from the compressed archive using tar

tar -xvf <archive-path>

The extracted file will be placed in a directory with the same name as the original archived directory.

Paths are stored as they are created in tar, so if an absolute path is defined in creation, the archive will be extracted with absolute paths. The output path can be changed using the -C option, for example:

Files archived:

tar -cjf site1.bz2 /var/www/site1/

var/www/site1/style.css

var/www/site1/index.html

var/www/site1/page2.html

To remove the outer site structure but keep the inner:

tar -cjf site1.tar.bz2 -C /var/www/site1 .

## List

List the files in an archive using the -t flag:

tar -tvf <archive-path>.tar.gz

# bzip

Another common compression tool is bzip, which creates .bz2 files. bzip is used to compress single files and takes more time than gzip, but does reduce the file size futher.

## Create

use the bzip command on a single file:

bzip <filename>

use the -j switch on the tar command:

tar cvfj <filename>.tar.bz2 <directory>

## Extract

bzip -d <filename>