# Introduction

File compression is essential in Linux for reducing storage space and optimizing file transfers. Various tools like **tar, gzip, bzip2, and xz** provide powerful methods for compressing and archiving files while maintaining data integrity.

File compression saves disk space and/or network transmission time. We always increase the efficiency of compression at the expense of longer compression time: - **tar**: designed to create a single archive from multiple files or directories, the archive can be a regular file or device - **cpio**: reads and writes files in binary or ASCII format. - **gzip**: uses Lempel-Ziv (LZ77) encoding and creates .gz files - **bzip2**: uses a text compression algorithm with Burrows-Wheeler block sorting and Huffman encoding, and produces .bz2 files - **xz**: produces .xz files. and also supports the older .lzma format

## TAR

This is the default archive management program available on Linux. It is designed to create a single archive from multiple files or directories and to manipulate them. The archive can be a simple file or a device (such as a tape drive, hence the program’s name - tape archiver), which can reside on a local or remote machine, for example.

tar -cvfv archive.tar /usr/bin/.

Flags: - -c, --create: create a new archive. - -x, --extract, --get: extracting files from the archive - -f, --file=ARCHIWUM: use ARCHIVE file or device. - -v, --verbose: print out details about the files being processed - -j, --bzip2: filter the archive by bzip2 - -j, --xz: filter the archive by xz - -z, --gzip, --gunzip, --ungzip: filtering the archive by gzip - -p, --preserve-permissions: restore file permissions information - -w, --verify: attempt to verify the archive after saving

Examples of archive creation and compression:

tar zcvf archive.tar.gz /usr/bin/  
tar jcvf archive.tar.bz2 /usr/bin/  
tar Jcvf archive.tar.xz /usr/bin/

Unpacking the archive:

tar xzvf source.tar.gz  
tar xjvf source.tar.bz2  
tar xJvf source.tar.xz

## GZIP

The**gzip** command compresses files using Lempel-Ziv (LZ77) encoding. Symbolic links are ignored by default.

Flags: - -k, --keep: keep (not delete) the input file while performing compression or decompression - -d, --decompress: decompress - -v, --verbose: verbose mode - -r, --recursive: compress recursively (directory and subdirectories) - -c, --stdout: save the output, keeping the original files unchanged - -l, --list: display archive details (contents and compression data) - -q, --quiet: do not display warnings

**gzip** lets you specify a range of compression levels, from 1 to 9: - -1 or --fast: indicates fastest compression speed with minimum compression ratio - -9 or --best: means the slowest compression speed with maximum compression ratio

Compression:

gzip file  
gzip -k file  
gzip -rk /tmp/folder <- will create an archive of each file in the /tmp/folder/ directory.  
gzip -l archive.gz

Decompression:

gzip -d archive.gz  
gzip -dk archive.gz

The default compression level is -6.

Compression test:

gzip -rck --best /usr/bin/ > archive-best.gz <- redirect compressed files to a specific archive file  
gzip -rck --fast /usr/bin/ > archive-fast.gz <- compare files

# Common Practical Examples

## 1. Compressing Files with tar

### Create a tar archive without compression

tar -cvf archive.tar /path/to/files

### Extract a tar archive

tar -xvf archive.tar

### List contents of a tar archive

tar -tvf archive.tar

## 2. Compressing with Gzip

### Create a tar archive and compress with gzip

tar -cvzf archive.tar.gz /path/to/files

### Extract a tar.gz archive

tar -xvzf archive.tar.gz

### Compress a single file with gzip

gzip filename.txt

### Decompress a gzip file

gunzip filename.txt.gz

## 3. Compressing with Bzip2

### Create a tar archive and compress with bzip2

tar -cvjf archive.tar.bz2 /path/to/files

### Extract a tar.bz2 archive

tar -xvjf archive.tar.bz2

### Compress a single file with bzip2

bzip2 filename.txt

### Decompress a bzip2 file

bunzip2 filename.txt.bz2

## 4. Compressing with XZ

### Create a tar archive and compress with xz

tar -cvJf archive.tar.xz /path/to/files

### Extract a tar.xz archive

tar -xvJf archive.tar.xz

### Compress a single file with xz

xz filename.txt

### Decompress an xz file

unxz filename.txt.xz

## 5. Comparing Compression Ratios

### Check file size before and after compression

ls -lh filename.txt filename.txt.gz

### Compare compression efficiency of different algorithms

du -sh archive.tar archive.tar.gz archive.tar.bz2 archive.tar.xz

# Additional Notes

* **gzip** is faster but provides moderate compression.
* **bzip2** has better compression but is slower.
* **xz** offers the best compression but is the slowest.
* **tar** is useful for grouping multiple files into a single archive before compression.