NAME

gzip, gunzip, zcat - compress or expand files

SYNOPSIS

gzip [**-acdfhlLnNrtvV19**] [**-S suffix**] [*name ...*]
gunzip [**-acfhlLnNrtvV**] [**-S suffix**] [*name ...*]
zcat [**-fhLV**] [*name ...*]

DESCRIPTION

Gzip reduces the size of the named files using Lempel-Ziv coding (LZ77). Whenever possible, each file is replaced by one with the extension **.gz**, while keeping the same ownership modes, access and modification times. (The default extension is **-gz** for VMS, **z** for MSDOS, OS/2 FAT, Windows NT FAT and Atari.) If no files are specified, or if a file name is "-", the standard input is compressed to the standard output. *Gzip* will only attempt to compress regular files. In particular, it will ignore symbolic links.

If the compressed file name is too long for its file system, *gzip* truncates it. *Gzip* attempts to truncate only the parts of the file name longer than 3 characters. (A part is delimited by dots.) If the name consists of small parts only, the longest parts are truncated. For example, if file names are limited to 14 characters, gzip.msdos.exe is compressed to gzi.msd.exe.gz. Names are not truncated on systems which do not have a limit on file name length.

By default, *gzip* keeps the original file name and timestamp in the compressed file. These are used when decompressing the file with the **-N** option. This is useful when the compressed file name was truncated or when the time stamp was not preserved after a file transfer.

Compressed files can be restored to their original form using *gzip -d* or *gunzip* or *zcat.* If the original name saved in the compressed file is not suitable for its file system, a new name is constructed from the original one to make it legal.

gunzip takes a list of files on its command line and replaces each file whose name ends with .gz, -gz, .z, -z, _z or .Z and which begins with the correct magic number with an uncompressed file without the original extension. *gunzip* also recognizes the special extensions **.tgz** and **.taz** as shorthands for **.tar.gz** and **.tar.Z** respectively. When compressing, *gzip* uses the **.tgz** extension if necessary instead of truncating a file with a **.tar** extension.

gunzip can currently decompress files created by *gzip, zip, compress, compress -H* or *pack.* The detection of the input format is automatic. When using the first two formats, *gunzip* checks a 32 bit CRC. For *pack, gunzip* checks the uncompressed length. The standard *compress* format was

not designed to allow consistency checks. However *gunzip* is sometimes able to detect a bad .Z file. If you get an error when uncompressing a .Z file, do not assume that the .Z file is correct simply because the standard *uncompress* does not complain. This generally means that the standard *uncompress* does not check its input, and happily generates garbage output. The SCO compress -H format (lzh compression method) does not include a CRC but also allows some consistency checks.

Files created by *zip* can be uncompressed by gzip only if they have a single member compressed with the 'deflation' method. This feature is only intended to help conversion of tar.zip files to the tar.gz format. To extract zip files with several members, use *unzip* instead of *gunzip.*

zcat is identical to *gunzip* **-c**. (On some systems, *zcat* may be installed as *gzcat* to preserve the original link to *compress.)* *zcat* uncompresses either a list of files on the command line or its standard input and writes the uncompressed data on standard output. *zcat* will uncompress files that have the correct magic number whether they have a **.gz** suffix or not.

Gzip uses the Lempel-Ziv algorithm used in *zip* and PKZIP. The amount of compression obtained depends on the size of the input and the distribution of common substrings. Typically, text such as source code or English is reduced by 60-70%. Compression is generally much better than that achieved by LZW (as used in *compress*), Huffman coding (as used in *pack*), or adaptive Huffman coding (*compact*).

Compression is always performed, even if the compressed file is slightly larger than the original. The worst case expansion is a few bytes for the gzip file header, plus 5 bytes every 32K block, or an expansion ratio of 0.015% for large files. Note that the actual number of used disk blocks almost never increases. *gzip* preserves the mode, ownership and timestamps of files when compressing or decompressing.

OPTIONS

-a --ascii
Ascii text mode: convert end-of-lines using local conventions. This option is supported only on some non-Unix systems. For MSDOS, CR LF is converted to LF when compressing, and LF is converted to CR LF when decompressing.

-c --stdout --to-stdout
Write output on standard output; keep original files unchanged. If there are several input files, the output consists of a sequence of independently

compressed members. To obtain better compression, concatenate all input files before compressing them.

-d --decompress --uncompress
Decompress.

-f --force
Force compression or decompression even if the file has multiple links or the corresponding file already exists, or if the compressed data is read from or written to a terminal. If the input data is not in a format recognized by *gzip,* and if the option --stdout is also given, copy the input data without change to the standard ouput: let *zcat* behave as *cat.* If **-f** is not given, and when not running in the background, *gzip* prompts to verify whether an existing file should be overwritten.

-h --help
Display a help screen and quit.

-l --list
For each compressed file, list the following fields:

```
    compressed size: size of the compressed file
    uncompressed size:  size  of  the  uncompressed
file
    ratio: compression ratio (0.0% if unknown)
    uncompressed_name:  name  of  the  uncompressed
file
```

The uncompressed size is given as -1 for files not in gzip format, such as compressed .Z files. To get the uncompressed size for such a file, you can use:

```
    zcat file.Z | wc -c
```

In combination with the --verbose option, the following fields are also displayed:

```
    method: compression method
    crc: the 32-bit CRC of the uncompressed data
    date & time: time stamp  for  the  uncompressed
file
```

The compression methods currently supported are deflate, compress, lzh (SCO compress -H) and pack. The crc is given as ffffffff for a file not in gzip format.

With --name, the uncompressed name, date and time are those stored within the compress file if

present.

With --verbose, the size totals and compression ratio for all files is also displayed, unless some sizes are unknown. With --quiet, the title and totals lines are not displayed.

-L --license
Display the *gzip* license and quit.

-n --no-name
When compressing, do not save the original file name and time stamp by default. (The original name is always saved if the name had to be truncated.) When decompressing, do not restore the original file name if present (remove only the *gzip* suffix from the compressed file name) and do not restore the original time stamp if present (copy it from the compressed file). This option is the default when decompressing.

-N --name
When compressing, always save the original file name and time stamp; this is the default. When decompressing, restore the original file name and time stamp if present. This option is useful on systems which have a limit on file name length or when the time stamp has been lost after a file transfer.

-q --quiet
Suppress all warnings.

-r --recursive
Travel the directory structure recursively. If any of the file names specified on the command line are directories, *gzip* will descend into the directory and compress all the files it finds there (or decompress them in the case of *gunzip*).

-S .suf --suffix .suf
Use suffix .suf instead of .gz. Any suffix can be given, but suffixes other than .z and .gz should be avoided to avoid confusion when files are transferred to other systems. A null suffix forces gunzip to try decompression on all given files regardless of suffix, as in:

```
    gunzip -S "" *       (*.* for MSDOS)
```

Previous versions of gzip used the .z suffix. This was changed to avoid a conflict with *pack*(1).

-t --test
Test. Check the compressed file integrity.

-v --verbose
Verbose. Display the name and percentage reduction for each file compressed or decompressed.

-V --version
Version. Display the version number and compilation options then quit.

-# --fast --best
Regulate the speed of compression using the specified digit #, where **-1** or **--fast** indicates the fastest compression method (less compression) and **-9** or **--best** indicates the slowest compression method (best compression). The default compression level is **-6** (that is, biased towards high compression at expense of speed).

ADVANCED USAGE

Multiple compressed files can be concatenated. In this case, *gunzip* will extract all members at once. For example:

```
gzip -c file1  > foo.gz
gzip -c file2 >> foo.gz
```

Then

```
gunzip -c foo
```

is equivalent to

```
cat file1 file2
```

In case of damage to one member of a .gz file, other members can still be recovered (if the damaged member is removed). However, you can get better compression by compressing all members at once:

```
cat file1 file2 | gzip > foo.gz
```

compresses better than

```
gzip -c file1 file2 > foo.gz
```

If you want to recompress concatenated files to get better compression, do:

```
gzip -cd old.gz | gzip > new.gz
```

If a compressed file consists of several members, the uncompressed size and CRC reported by the --list option applies to the last member only. If you need the

uncompressed size for all members, you can use:

```
gzip -cd file.gz | wc -c
```

If you wish to create a single archive file with multiple members so that members can later be extracted independently, use an archiver such as tar or zip. GNU tar supports the -z option to invoke gzip transparently. gzip is designed as a complement to tar, not as a replacement.

ENVIRONMENT

The environment variable **GZIP** can hold a set of default options for *gzip.* These options are interpreted first and can be overwritten by explicit command line parameters. For example:

```
for sh:    GZIP="-8v --name"; export GZIP
for csh:   setenv GZIP "-8v --name"
for MSDOS: set GZIP=-8v --name
```

On Vax/VMS, the name of the environment variable is GZIP_OPT, to avoid a conflict with the symbol set for invocation of the program.

SEE ALSO

znew(1), zcmp(1), zmore(1), zforce(1), gzexe(1), zip(1), unzip(1), compress(1), pack(1), compact(1)

DIAGNOSTICS

Exit status is normally 0; if an error occurs, exit status is 1. If a warning occurs, exit status is 2.

Usage: gzip [-cdfhlLnNrtvV19] [-S suffix] [file ...]
Invalid options were specified on the command line.

file: not in gzip format
The file specified to *gunzip* has not been compressed.

file: Corrupt input. Use zcat to recover some data.
The compressed file has been damaged. The data up to the point of failure can be recovered using
zcat file > recover

file: compressed with *xx* bits, can only handle *yy* bits
File was compressed (using LZW) by a program that could deal with more *bits* than the decompress code on this machine. Recompress the file with gzip, which compresses better and uses less memory.

file: already has .gz suffix -- no change
The file is assumed to be already compressed. Rename the file and try again.

file already exists; do you wish to overwrite (y or n)?
Respond "y" if you want the output file to be replaced; "n" if not.

gunzip: corrupt input
A SIGSEGV violation was detected which usually

means that the input file has been corrupted.

xx.x%
Percentage of the input saved by compression. (Relevant only for **-v** and **-l**.)

-- not a regular file or directory: ignored
When the input file is not a regular file or directory, (e.g. a symbolic link, socket, FIFO, device file), it is left unaltered.

-- has *xx* other links: unchanged
The input file has links; it is left unchanged. See *ln*(1) for more information. Use the **-f** flag to force compression of multiply-linked files.

CAVEATS

When writing compressed data to a tape, it is generally necessary to pad the output with zeroes up to a block boundary. When the data is read and the whole block is passed to *gunzip* for decompression, *gunzip* detects that there is extra trailing garbage after the compressed data and emits a warning by default. You have to use the --quiet option to suppress the warning. This option can be set in the **GZIP** environment variable as in:

```
  for sh:  GZIP="-q"  tar -xfz --block-compress /dev/rst0
  for csh:  (setenv  GZIP  -q;  tar  -xfz   --block-compr
/dev/rst0
```

In the above example, gzip is invoked implicitly by the -z option of GNU tar. Make sure that the same block size (-b option of tar) is used for reading and writing compressed data on tapes. (This example assumes you are using the GNU version of tar.)

BUGS

The --list option reports incorrect sizes if they exceed 2 gigabytes. The --list option reports sizes as -1 and crc as ffffffff if the compressed file is on a non seekable media.

In some rare cases, the --best option gives worse compression than the default compression level (-6). On some highly redundant files, *compress* compresses better than *gzip.*