

NÉV

gzip, *gunzip*, *zcat* – fájlok tömörítése vagy kicsomagolása

ÁTTEKINTÉS

gzip [**-acdfhlLnNrtvV19**] [**-S suffix**] [*név ...*]

gunzip [**-acfhLlNrtvV**] [**-S suffix**] [*név ...*]

zcat [**-fhLV**] [*név ...*]

LEÍRÁS

Gzip csökkenti a megnevezett fájlok méretét a Lempel-Ziv kódolás (LZ77) felhasználásával. Ha csak lehetséges, minden fájl egy **.gz**, kiterjesztésűvel lesz helyettesítve azonos tulajdonossal, engedélyekkel, elérési- és módosítási időkkel. (Az alapértelmezett kiterjesztés **-gz** VMS alatt, **z** MSDOS, OS/2 FAT, Windows NT FAT és Atari esetén.) Amennyiben nincs fájl megadva, vagy a fájlnev a "-" jel, a standard bemenetet tömöríti a szabványos kimenetre.

Gzip csak a szabályos fájlokat kísérli meg tömöríteni. Főként a szimbolikus kötéseket hagyja figyelmen kívül.

Amennyiben a tömörített fájlnev túl hosszú a fájlrendszernek *gzip* csonkolni fogja. *Gzip* csak a fájlnev 3 karakternél hosszabb részeit kísérli meg csonkolni. (A részeket pontok határolják.) Ha a név csak kis részekből áll, a leghosszabb részeket csonkolja. Például ha a fájlnevek maximum 14 karakteresek lehetnek, *gzip.msdos.exe* tömörített neve *gzi.msdx.gz* lesz. Olyan rendszereken, ahol nincs korlát a fájlnevek hosszára, nem történik csonkolás.

Alapértelmezés szerint *gzip* megőrzi az eredeti fájl nevét és időbélyegét (timestamp) a tömörített fájlban. Ezeket akkor használja, ha a kicsomagolás (decompression) a **-N** opcióval történik. Ez akkor hasznos, ha a tömörített fájl neve csonkolva volt, vagy az időbélyeg megváltozott egy fájlátvitel miatt.

A tömörített fájlok visszaállíthatók eredeti formájukra a *gzip -d* vagy *gunzip* vagy *zcat* használatával. Ha a tömörített fájlban elmentett eredeti név nem megfelelő a fájlrendszernek, új, legális név meghatározása történik az eredeti alapján.

Gunzip egy fájllistát kap a parancssoron keresztül, és minden **.gz**, **-gz**, **.z**, **-z**, **_z** vagy **.Z** végződésű fájlt, illetve azokat, amelyek megfelelő mágikus számmal kezdődnek helyettesít a kitömörített változatukkal az eredeti kiterjesztés nélkül. A *gunzip* program felismeri a **.tgz** és **.taz** speciális kiterjesztéseket, mint a **.tar.gz** illetve **.tar.Z** rövidítéseit. Tömörítéskor *gzip* a **.tgz** kiterjesztést használja ha szükséges, ahelyett, hogy a **.tar** kiterjesztésű fájlok tömörített nevét csonkolná.

A *gunzip* program pillanatnyilag a *gzip*, *zip*, *compress*, *compress -H* vagy *pack* programok által tömörített fájlokat tudja kicsomagolni. A bemenő formátum érzékelése automatikus. Az első két esetben *gunzip* egy 32 bites CRC ellenőrzést végez. A *pack* esetében *gunzip* a tömörítetlen hosszát ellenőrzi. A standard *compress* formátum nem ad lehetőséget az önellenőrzésre. Ennek ellenére *gunzip* néha képes felismerni a rossz **.Z** fájlt. Ha egy **.Z** fájl kicsomagolása során hibaüzenet keletkezik, de a standard *uncompress* nem jelez semmit, az nem jelenti azt, hogy a tömörített fájl korrekt volt. Ez többnyire az miatt van, hogy a standard *uncompress* nem ellenőrzi a bemenetet, és vidáman generál összezagyvált kimenetet. A 'SCO *compress -H*' formátum (lzh tömörítési módszer) nem tartalmaz CRC-t, de megenged néhány önellenőrzési lehetőséget.

A *zip* által létrehozott programokat csak akkor tudja a *gunzip* kicsomagolni, ha egyetlen tagja van, mely a 'deflation' módszerrel lett tömörítve. Ez a lehetőség csak a *tar.zip* fájlok *tar.gz*-re alakítására ajánlott. Többtagú *zip* fájlok kicsomagolására az *unzip* alkalmas.

Zcat azonos *gunzip -c* -vel. (Néhány rendszeren a *zcat* program *gzcat* néven lehet telepítve, hogy megvédjék a *compress* -re mutató eredeti kötést.) *Zcat* a parancssorból érkező fájllistát és a szabványos bemenetet is képes kicsomagolni, és a kicsomagolt adatokat a szabványos kimenetre írja.

A *zcat* program a megfelelő mágikus számmal rendelkező fájlokat kicsomagolja, függetlenül a **.gz** meglététől.

A *gzip* program a *zip* és PKZIP programok által is használt Lempel-Ziv kódolást használja. Az elért tömörítés mértéke a bemenet méretétől és az azonos stringek eloszlásától függ. Például egy

tipikus szöveget, mint egy forráskód vagy angol szöveg, az eredeti méret 30–40%-ára tömörít össze. A tömörítés általában sokkal jobb, mint az LZW-vel (*compress*), a Huffman kódolással (*pack*), vagy az adaptív Huffman-kódolással (*compact*) érhető el.

A tömörítési folyamat mindig megtörténik, még akkor is, ha a tömörített fájl kicsivel nagyobb, mint az eredeti. A legrosszabb eset néhány bájtos gzip fejléccel és 32K-nként 5 bájttal való méretnövekedést jelent, illetve 0.015%-os növekedési arányt nagy fájlknál. Ez általában nem jelenti több lemezblokk felhasználását. *Gzip* megőrzi a feldolgozott fájlok módját, tulajdonosát és időbélyegeit.

OPCIÓK

—a --ascii

ASCII szöveg mód: a sorvégeleket a helyi szokásoknak megfelelően alakítja. Ez az opció csak néhány nem-Unix rendszeren támogatott. MSDOS esetén a CR+LF jelből LF lesz tömörítéskor, LF-ből pedig CR+LF kicsomagolásakor.

—c --stdout --to-stdout

A kimenetet a szabványos kimenetre írja, az eredeti fájlokat változatlanul hagyja. Több bemenő fájl esetén a kimenet függetlenül tömörített tagok sorozata lesz. Jobb tömörítést lehet elérni a bemeneti fájlok tömörítés előtti összekapcsolásával.

—d --decompress --uncompress

Kicsomagolás (decompress).

—f --force

A tömörítés vagy kicsomagolás erőltetése még akkor is, ha a fájlnek többszörös kötése van, vagy a megfelelő fájl már létezik, vagy a tömörített adatot terminálról olvassa vagy oda írja a program. Ha a bemeneti adatok nem felismerhetők a *gzip* -nek és ha a '-stdout' opció is adott, a bemenő adatokat változás nélkül a szabványos kimenetre másolja. Ilyenkor *zcat* ugyanúgy viselkedik, mint *cat*. Ha —f nincs megadva és nem a háttérben fut a program, *gzip* visszakérdez, hogy a már létező fájlok felülírhatók-e.

—h --help

Segítő információkat ír ki, majd kilép.

—l --list

Minden tömörített fájlra listát készít a következő mezőkkel:

```
compressed size: a tömörített fájl mérete
uncompressed size: a tömörítetlen fájl mérete
ratio: tömörítési arány (0.0% ha ismeretlen)
uncompressed_name: a tömörítetlen fájl neve
```

Az 'uncompressed size' mező értéke -1 a nem gzip formátumú fájlokra, mint pl. a tömörített .Z fájlok. Egy ilyen fájl tömörítetlen méretének meghatározására használhatjuk pl. a következő parancsot:

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

A -l opciót együtt használva a --verbose-zal a következő mezők is megjelennek:

```
method: tömörítési módszer
crc: a tömörítetlen adatok 32 bites CRC értéke
date & time: a tömörítetlen fájlok időbélyege
```

Pillanatnyilag a támogatott tömörítési módszerek: deflate, compress, lzh (SCO compress -H) és pack. A crc értéke ffffffff lesz nem gzip formátumú fájlokra.

A `--verbose` használata esetén a méretek és a tömörítési arányok összegzése is kiíródik, hacsaknem van ismeretlen méret. A `--quiet` opcióval a cím és az összegzés nem kerül kijelzésre.

—**L** `--license`

Kijelzi a *gzip* engedélyeit és kilép.

—**n** `--no-name`

Tömörítéskor nem menti el az eredeti fájl nevét és időbélyegét. (Az eredeti név mindig elmentődik, ha a nevet csonkolni kellett.) Kicsomagoláskor nem állítja vissza az eredeti fájlnevet, azaz csak a *gzip* fájlvégződését távlítja el, és nem állítja vissza az időbélyegeket, hanem a tömörített fájléből másolja.

Ez az opció kicsomagoláskor alapértelmezett.

—**N** `--name`

Tömörítéskor mindenképpen elmenti az eredeti fájlnevet és időbélyegét. (Alapértelmezett.) Kicsomagoláskor az eredeti fájlnevet és időbélyegét állítja vissza. Ez az opció olyan rendszereken hasznos, ahol a fájlnevek hossza korlátos, vagy az időbélyeg elveszett egy fájlátvitel során.

—**q** `--quiet`

Elhagyja a figyelmeztetéseket.

—**r** `--recursive`

Rekurzívan végigmegy a könyvtárstruktúrán. Ha a parancssorban megadott fájlnevek könyvtárakat is tartalmaznak, *gzip* bemegy a könyvtárba és tömöríti (illetve *gunzip* kicsomagolja) az ott talált fájlokat.

—**S** `.suf --suffix .suf`

A `.suf` fájlnev-végződést használja `.gz` helyett. Bármilyen végződés használható, de `.z` és `.gz`-től különböző használata nem ajánlott más rendszerekre való esetleges átvitel esetén. Üres végződés esetén *gunzip* megpróbál minden adott fájl kicsomagolni a végződéstől függetlenül. Pl. a következő parancs az aktuális könyvtár minden fájlját megpróbálja kicsomagolni:

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

A *gzip* korábbi verziói a `.z` végződést használták. Ez azért lett megváltoztatva, hogy ne legyen keveredés a *pack(1)*. programmal.

—**t** `--test`

Teszteli a tömörített fájl épségét (integrity).

—**v** `--verbose`

Bőbeszédű üzemmód. Kijelzi az összes tömörített vagy kicsomagolt fájl tömörítési arányát és nevét.

—**V** `--version`

Kijelzi a program verziószámát és a fordítási opciókat, majd kilép.

—**#** `--fast --best`

A tömörítés sebességét szabályozza a `#`, számmal, ahol `—1` vagy `—fast` a leggyorsabb, de kisebb tömörítésű tömörítési módszert jelöli, és `—9` vagy `—best` a leglassabb, de legjobban tömörítő módszert. Alapértelmezés: `—6` között.

HALADÓ HASZNÁLAT

Több tömörített fájl összekapcsolható. Ebben az esetben *gunzip* minden tagot egyszer csomagol ki. Pl.:

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

```
majd
  gunzip -c foo
```

egyenértékű ezzel:

```
cat file1 file2
```

A .gz fájl egyik tagjának sérülésekor a többi tag még visszaállítható, ha a sérült részt eltávolítjuk. Jobb tömörítés érhető el, ha az összes tagot egyszerre tömörítjük. Pl.

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

jobban tömörít, mint

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

Ha összefűzött fájlokat újra akarunk tömöríteni nagyobb tömörség elérése miatt, csinálhatjuk a következő módon:

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

Ha a tömörített fájl több tagból áll, a tömörítetlen méret és a CRC adat, amit a --list opció kijelez, csak az utolsó tagra vonatkozik. Ha mindegyik tag tömörítetlen méretére kíváncsiak vagyunk, használhatjuk a következő parancsot:

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

Amennyiben egyetlen olyan archív fájlt akarunk készíteni, mely több, egymástól függetlenül kicsomagolható tagból áll, valamilyen archiváló programot kell használnunk, mint pl. a tar, vagy a zip. A GNU tar támogatja a -z opciót, mely esetén magától meghívja a gzip-et. A gzip így kiegészíti, és nem helyettesíti a tar-t.

KÖRNYEZET

A **GZIP** környezeti változó a *gzip* alapértelmezett opcióit tartalmazhatja. Először ezek az opciók kerülnek értelmezésre, de a parancssor paramétereivel felülbírálhatók. Pl.:

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

Vax/VMS alatt a környezeti változó neve GZIP_OPT, hogy elkerüljük a konfliktust a környezeti változó és a programhívás közti konfliktust.

LÁSD MÉG

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

DIAGNOSZTIKA

A normál kilépési állapot 0, hiba esetén 1, figyelmeztetés esetén 2.

Az angol változat lehetséges hibaiüzenetei és figyelmeztetései:

```
Usage: gzip [-cdfhLnNrtvV19] [-S suffix] [file ...]
```

A parancssorban érvénytelen opciók kerültek megadásra.

file: not in gzip format

A *gunzip* -nek átadott fájl nem tömörített, vagy ismeretlen formátumú.

file: Corrupt input. Use zcat to recover some data.

A tömörített fájl megsérült. A hiba pontjáig az adatok visszaállíthatók a következő paranccsal:

```
zcat file > recover
```

file: compressed with *xx* bits, can only handle *yy* bits

File egy olyan LZW-t használó programmal lett tömörítve, amely több bitet használt, mint az aktuális gép kicsomagoló kódja. Célszerű újratömöríteni a fájlt gzip-pel, ami jobban tömörít és kevesebb memóriát használ.

file: already has .gz suffix -- no change

A fájl feltételezhetően már tömörítve van, mert kiterjesztése .gz . Ha mégsem tömörített, át kell nevezni, majd újra próbálkozni.

file already exists; do you wish to overwrite (y or n)?

Válaszolj "y"-t, ha azt akarod, hogy a kimeneti fájl felülírja az eredetit, különben "n"-t.

gunzip: corrupt input

A program "SIGSEGV violation"-t detektált, ami általában sérült bemeneti fájlt jelent.

xx.x%

A tömörítés által megtakarított terület aránya százalékokban. (`-v` és `-l` opciók esetén érvényes.)

-- not a regular file or directory: ignored

Amikor a bemeneti fájl nem szabályos fájl vagy könyvtár (hanem pl. szimbolikus kötés, socket, FIFO, eszközfájl), a gzip változatlanul hagyja. -- has *xx* other links: unchanged

A bemeneti fájlnek kötése van, a gzip változatlanul hagyja. Lásd még: *ln(1)* használatát a bővebb információért. Használj a `-f` opciót a többszörösen kötött fájlok tömörítéséhez.

FIGYELMEZTETÉSEK

Tömörített adatok szalagra írásakor általában szükséges a kimenet feltöltése nullákkal a blokkhatár eléréséig. Az adatok olvasásakor a teljes blokk átadásra kerül a *gunzip* programnak kicsomagolásra, de *gunzip* érzékeli, hogy valamilyen szemét követi a tömörített adatokat, ami alapértelmezés szerint figyelmeztetéshez vezet. A figyelmeztetés elhagyásához a `--quiet` opció használható. Ezt a **GZIP** környezeti változóban is beállíthatjuk:

```
sh esetén: GZIP="-q" tar -xfz --block-compress /dev/rst0
```

```
csh esetén: (setenv GZIP -q; tar -xfz --block-compr /dev/rst0
```

A fenti példában a gzip-et a GNU tar hívja meg a `-z` opció miatt. Legyünk biztosak, hogy azonos blokkméret (a tar `-b` opciója) lett az olvasáskor és az íráskor is használva. (Természetesen ez a példa a GNU tar használatát feltételezte.)

HIBÁK

A `--list` opció rossz méretet jelez 2 gigabájt felett. A `--list` opció `-l -et` ad meg méretnek és `ffffff` -et CRC-nek ha a tömörített fájl nem elérhető adathordozón van.

Néhány ritka esetben a `--best` opció rosszab tömörítést ad, mint az alapértelmezett tömörítési szint. Néhány nagyon redundáns fájl esetén *compress* jobban tömörít, mint *gzip*.

MAGYAR FORDÍTÁS

Horváth András <horvatha@rs1.szif.hu>