

gzip(1)

НАЗВАНИЕ

gzip, **gunzip**, **gzcat** - сжатие (упаковка) и распаковка файлов

СИНТАКСИС

```
gzip [ -acdfhlLnNrtvV19 ] [-S суффикс] [ имя ... ]
gunzip [ -acfhLlnNrtvV ] [-S суффикс] [ имя ... ]
gzcat [ -fhLV ] [ имя ... ]
```

ОПИСАНИЕ

Утилита **gzip** уменьшает размер указанных файлов за счет кодирования по *алгоритму Лемпела-Зива* (Lempel-Ziv coding - LZ77). По возможности каждый файл заменяется одноименным файлом с расширением **.gz**, с сохранением тех же прав доступа, владельца и времени изменения. (Стандартное расширение - **-gz** для ОС VMS, **z** для MSDOS, OS/2 FAT, Windows NT FAT и Atari.) Если файлы не указаны или в качестве имени файла указан дефис (-), содержимое стандартного входного потока сжимается и направляется в стандартный выходной поток. Утилита **gzip** будет пытаться сжимать только обычные файлы. В частности, она будет игнорировать символические связи.

Если имя сжатого файла - слишком длинное для соответствующей файловой системы, утилита **gzip** его усекает. Утилита **gzip** пытается усечь только те части имени файла, которые длиннее 3 символов. (Часть имени ограничивается точками.) Если имя состоит только из коротких частей, усекаются самые длинные части. Например, если имена файлов должны быть не длиннее 14 символов, файл **gzip.msdos.exe** после сжатия получает имя **gzi.msdx.gz**. Имена не усекаются в системах, не имеющих ограничений на длину имени файла.

По умолчанию, утилита **gzip** сохраняет имя и время изменения сжимаемого файла. Они используются при разжатии файла с использованием опции **-N**. Это удобно, когда имя сжатого файла было усечено или когда после переноса не сохранилась временная отметка файла.

Сжатые файлы могут быть восстановлены в исходный вид с помощью **gzip -d**, **gunzip** или **gzcat**. Если исходное имя файла до сжатия не подходит для текущей файловой системы, на основе исходного имени конструируется новое, допустимое.

Утилита **gunzip** принимает список файлов в командной строке и заменяет каждый файл, имя которого заканчивается суффиксом **.gz**, **-gz**, **.z**, **-z**, **_z** или **.Z** и который начинается с соответствующей сигнатуры распакованным файлом без расширения. Утилита **gunzip** также распознает специальные расширения **.tgz** и **.taz** в качестве сокращений для **.tar.gz** и **.tar.Z**, соответственно. При сжатии утилита **gzip**, при необходимости, использует расширение **.tgz** вместо усечения имени файла с расширением **.tar**.

Утилита **gunzip** в настоящее время может разжимать файлы, созданные утилитами **gzip**, **zip**, **compress**, **compress -H** или **pack**. Формат входных данных определяется автоматически. При использовании первых двух форматов, **gunzip** проверяет 32-битовую *контрольную сумму CRC*. Для формата **pack**, **gunzip** проверяет длину распакованного файла. Стандартный формат **compress** не предполагает возможности проверки

согласованности. Однако, утилита **gunzip** иногда способна выявить сбойный файл с расширением **.Z**. Если при разжатии файла с расширением **.Z** вы получаете сообщение об ошибке, не думайте, что этот файл корректен только потому, что стандартная утилита **uncompress** не сообщает об ошибках. Это просто означает, что стандартная утилита **uncompress** не проверяет входные данные и молча генерирует "мусорные" результаты. Формат **compress -H** в SCO (метод сжатия **lzh**) не включает контрольной суммы CRC, но позволяет выполнять определенные проверки согласованности.

Файлы, сжатые с помощью утилиты **zip** могут быть распакованы с помощью **gzip** только если они имеют один компонент, сжатый по методу **'deflation'**. Эта возможность предназначена для упрощения преобразования файлов **tar.zip** в формат **tar.gz**. Для распаковки **zip**-файлов из нескольких компонентов используйте вместо **gunzip** утилиту **unzip**.

Команда **gzcat** идентична **gunzip -c**. (В некоторых системах утилита **zcat** может устанавливаться как **gzcat**, чтобы сохранилась исходная связь с **compress**.) Утилита **gzcat** распаковывает либо список файлов, заданный в командной строке, либо стандартный входной поток, и записывает распакованные данные в стандартный выходной поток. Утилита **gzcat** будет распаковывать файлы, имеющие соответствующую сигнатуру, независимо от наличия суффикса **.gz**.

Утилита **gzip** использует алгоритм Лемпела-Зива (Lempel-Ziv), как и утилиты **zip** и **PKZIP**. Степень сжатия зависит от размера входного файла и распределения типичных подстрок. Обычно файлы типа исходных кодов программ или текстов на английском языке упаковываются на 60-70%. Степень сжатия обычно намного лучше, чем достигаемая алгоритмом LZW (который используется в **compress**), кодированием Хаффмана (Huffman coding, используемым утилитой **pack**), или адаптивным кодированием Хаффмана (как в утилите **compact**).

Сжатие выполняется всегда, даже если сжатый файл несколько больше исходного. В худшем случае файл увеличивается на несколько байтов заголовка файла **gzip** плюс 5 байтов на каждый блок размером 32 Кбайта, что означает рост порядка 0.015% для больших файлов. Учтите, что фактическое количество блоков диска, занимаемых файлом, практически никогда не увеличивается. Утилита **gzip** сохраняет права доступа, владельца и временные отметки файлов при сжатии и распаковке.

ОПЦИИ

-a --ascii

Текстовый режим **ascii**: преобразует символы конца строки на базе локальных установок. Эта опция поддерживается только на некоторых системах, отличных от Unix. Для MSDOS, CR LF преобразуется при сжатии в LF, а LF преобразуется в CR LF при распаковке.

-c --stdout --to-stdout

Выдает результат в стандартный выходной поток; исходные файлы не изменяются. Если имеется несколько входных файлов, результат состоит из последовательности независимо сжатых компонентов. Для получения большей степени сжатия, конкатенируйте все входные файлы перед сжатием.

-d --decompress --uncompress

Распаковка.

-f --force

Принудительно упаковывает или распаковывает файл даже если он имеет несколько связей или результирующий файл уже существует, или если сжатые данные читаются с терминала или записываются на терминал. Если входные данные не соответствуют формату, распознаваемому утилитой **gzip** и если указана также опция **--stdout**, копирует входные данные в стандартный выходной поток без изменений, что позволяет утилите **gzcat** работать как **cat**. Если опция **-f** не задана, и утилита не работает в фоновом режиме, **gzip** запрашивает у пользователя, надо ли перезаписывать существующий файл.

-h --help

Выдает справочную информацию и завершает работу.

-l --list

Для каждого сжатого файла выдает следующие поля:

`compressed size`: размер сжатого файла
`uncompressed size`: размер исходного файла
`ratio`: коэффициент сжатия (0.0% если неизвестен)
`uncompressed_name`: имя файла до сжатия

Размер разжатого файла имеет значение -1 для файлов, формат которых не поддерживается утилитой **gzip**, например, для сжатых файлов с расширением **.Z**. Для получения исходного размера файла, можно использовать команду:

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

В сочетании с опцией **--verbose** выдаются также следующие поля:

`method`: метод сжатия
`crc`: 32-битовая контрольная сумма (CRC) не сжатых данных
`date & time`: временная отметка исходного файла

В настоящее время поддерживаются следующие методы сжатия: **deflate**, **compress**, **lzh** (аналог **compress -H** в SCO) и **pack**. Для файлов не в формате **gzip** в качестве контрольной суммы **crc** выдается значение **ffffff**.

Если указана опция **--name**, в качестве имени, даты и времени изменения исходного несжатого файла выдаются значения, хранящиеся в сжатом файле, если они доступны.

Если указана опция **--verbose** для всех файлов выдаются также общие размеры и коэффициенты сжатия, за исключением случаев, когда размеры не известны. Если указана опция **--quiet**, заголовки и общие значения не выдаются.

-L --license

Выдает лицензию **gzip** и завершает работу.

-n --no-name

При сжатии по умолчанию не сохраняет имя и временную отметку исходного файла. (Исходное имя всегда сохраняется, если имя после сжатия придется усекавать.) При распаковке не восстанавливает исходное имя файла, если оно есть (просто удаляет суффикс **gzip** у имени сжатого файла) и не восстанавливает исходную временную отметку, если она есть (она просто копируется из сжатого файла). Эта опция является стандартной при распаковке.

-N --name

При сжатии всегда сохраняет исходное имя файла и временную отметку; так и происходит по умолчанию. При распаковке восстанавливает исходное имя и временную отметку файла, если они доступны. Эта опция полезна в системах с ограничениями на длину имени файла или когда после переноса файлов была потеряна временная отметка.

-q --quiet

Подавляет выдачу любых предупреждений.

-r --recursive

Рекурсивно проходит по структуре каталогов. Если любое из имен файлов, указанных в командной строке, представляет собой каталог, утилита **gzip** перейдет в этот каталог и сожмет все находящиеся в нем файлы (или разожмет их в случае **gunzip**).

-S .suf --suffix .suf

Использует суффикс **.suf** вместо **.gz**. Можно задать любой суффикс, но не стоит использовать другие суффиксы, кроме **.z** и **.gz**, чтобы избежать недоразумений при переносе файлов на другие системы. Пустой суффикс вынуждает программу **gunzip** попытаться распаковать все указанные файлы, независимо от суффикса, как в следующем примере:

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

В прежних версиях **gzip** использовался суффикс **.z**. Этот суффикс был изменен, чтобы избежать конфликтов с программой **pack(1)**.

-t --test

Проверка. Проверяет целостность сжатого файла.

-v --verbose

Подробна выдача. Выдает имя и процент сжатия для каждого сжимаемого или разжимаемого файла.

-V --version

Версия. Выдает номер версии и использованные опции компиляции и завершает работу.

-# --fast --best

Управляет скоростью сжатия, задавая ее указанной цифрой #, причем опция **-1** или **--fast** задает самый быстрый метод (с наименьшим коэффициентом сжатия), а **-9** или **--best** задает самый медленный метод (с наибольшим коэффициентом сжатия). Стандартный уровень сжатия - **-6** (т.е. имеется определенное смещение в сторону повышения коэффициента сжатия за счет снижения скорости).

РАСШИРЕННЫЕ ВАРИАНТЫ ИСПОЛЬЗОВАНИЯ

Можно конкатенировать несколько сжатых файлов. В этом случае утилита **gunzip** будет извлекать все файлы сразу. Например:

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

После этого команда

```
gunzip -c foo
```

будет эквивалентна команде

```
cat file1 file2
```

В случае повреждения одного из компонентов файла **.gz**, другие компоненты можно восстановить (если удалить поврежденный компонент). Однако большей степени сжатия можно достичь, сжимая все компоненты сразу:

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

сжимается лучше, чем

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

Если необходимо переупаковать конкатенированные файлы для получения большей степени сжатия, используйте команду следующего вида:

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

Если сжатый файл состоит из нескольких компонентов, размер исходного файла и контрольная сумма CRC, выдаваемые опцией **--list**, относятся только к последнему компоненту. Если необходимо узнать исходный размер всех компонентов, можно использовать команду:

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

Если необходимо создать единый архивный файл с несколькими компонентами так, чтобы можно было извлекать независимо друг от друга, используйте архиватор типа **tar** или **zip**. Утилита GNU **tar** поддерживает опцию **-z** для прозрачного вызова утилиты **gzip**. Утилита **gzip** дополняет программу **tar**, но не заменяет ее.

СРЕДА

Переменная среды **GZIP** может задавать набор стандартных опций для **gzip**. Эти опции интерпретируются первыми и могут переопределяться явными параметрами командной строки. Например:

для **sh**: GZIP="-8v --name"; export GZIP
для **csh**: setenv GZIP "-8v --name"
для **MSDOS**: set GZIP=-8v --name

В ОС Vax/VMS соответствующая переменная среды называется **GZIP_OPT**, чтобы избежать конфликта с именем программы.

ССЫЛКИ

gzip(1), **gzcmp(1)**, **gzmore(1)**, **gzforce(1)**, **gzexe(1)**, **zip(1)**, **unzip(1)**, [compress\(1\)](#), **pack(1)**

ДИАГНОСТИКА

Статус выхода обычно имеет значение 0; если возникает ошибка, статус выхода имеет значение 1. Если выдавались предупреждения, статус выхода имеет значение 2.

Usage: gzip [-cdfhlLnNrtvV19] [-S suffix] [file ...]
В командной строке заданы недопустимые опции.

file: not in gzip format
Указанный программе **gunzip** файл не был сжат утилитой **gzip**.

file: Corrupt input. Use gzcata to recover some data.
Сжатый файл был поврежден. Данные до точки сбоя могут быть восстановлены с помощью команды
gzcata file > recover

file: compressed with xx bits, can only handle yy bits
Файл был сжат (с помощью алгоритма LZW) программой, способной работать с большим количеством битов, чем код распаковки на данной машине. Сожмите файл заново с помощью утилиты **gzip**, сжимающей лучше и использующей меньше памяти.

file: already has .gz suffix -- no change
Предполагается, что файл уже сжат. Переименуйте файл и попробуйте еще раз.

file already exists; do you wish to overwrite (y or n)?
Ответьте утвердительно (**y**), если хотите заменить существующий файл, или отрицательно (**n**), если не хотите.

gunzip: corrupt input
Был получен сигнал **SIGSEGV**, что обычно вызвано повреждением входного файла.

xx.x%
Процент входного файла, сэкономленный за счет сжатия. (Только для опций **-y** и **-l**.)

-- not a regular file or directory: ignored
Если входной файл не является обычным файлом или каталогом, (т.е. это символическая связь, сокет, именованный канал FIFO или специальный файл устройства, он не изменяется.

-- has xx other links: unchanged
Входной файл имеет несколько связей; он не изменяется. Подробнее о связях см. на странице справочного руководства **ln(1)**. Для принудительного сжатия файлов с несколькими связями используйте флаг **-f**.

АВТОРСКИЕ ПРАВА

Copyright (C) 1987, 88, 89, 90, 91, 92, 1993 Free Software Foundation, Inc.

Эта программа является свободно распространяемым программным обеспечением; вы можете распространять и/или изменять ее в соответствии с условиями Генеральной общественной лицензии GNU ([GNU General Public License](#)), опубликованной Фондом свободно распространяемого программного обеспечения (Free Software Foundation); либо версии 2, либо (по вашему выбору) любой более поздней версии.

Эта программа распространяется в надежде, что она окажется полезной, но БЕЗО ВСЯКИХ ГАРАНТИЙ; даже без подразумеваемой гарантии ТОВАРНОЙ ПРИГОДНОСТИ или ПРИГОДНОСТИ ДЛЯ КОНКРЕТНОЙ ЦЕЛИ. Подробнее см. в Генеральной общественной лицензии GNU.

Вы должны были получить экземпляр Генеральной общественной лицензии GNU вместе с этой программой; если же нет, обратитесь за ней по адресу Free Software Foundation, 675 Mass Ave, Cambridge, MA 02139, USA.

ПОТЕНЦИАЛЬНЫЕ ПРОБЛЕМЫ

При записи сжатых данных на ленту обычно необходимо дополнять результат нулями до границы блока. При чтении данных и передаче их для распаковки утилите **gunzip**, она выявляет дополнительный "мусор" после сжатых данных и, по умолчанию, выдает предупреждение. Чтобы подавить выдачу этого предупреждения, используйте опцию **--quiet**. Эту опцию можно задать в переменной среды **GZIP**, как в следующем примере:

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

В примере выше утилита **gzip** вызывается неявно опцией **-z** утилиты GNU **tar**. Проверьте, что при чтении и записи сжатых данных на ленты используется один и тот же размер блока (опция **-b** программы **tar**). (В этом примере предполагается использование GNU-версии программы **tar**.)

ОШИБКИ

Опция **--list** выдает некорректные размеры, если они превышают 2 Гбайта. Опция **--list** выдает значение размера **-1** и контрольной суммы **сrc - ffffffff**, если сжатый файл находится на носителе, не допускающем произвольную адресацию (*non seekable media*).

В некоторых редких случаях опция **--best** дает худшее сжатие, чем стандартный уровень сжатия (**-6**). Некоторые особо избыточные файлы сжимаются утилитой **compress** лучше, чем утилитой **gzip**.

ПРИМЕЧАНИЯ

Исходные коды для утилиты **gzip** доступны в пакете **SUNWgzipS**.