

Linux

Работа с архивами

Команды работы с архивами:

- gzip, gunzip, zcat
- zmore
- bzip2, bunzip2, bzip2recover
- zcmp, zdiff
- gzexe
- xz
- lzma
- tar

СЖАТИЕ ФАЙЛОВ ПО АЛГОРИТМУ LZ77

```
gzip [ -acdfhlLnNrtvV19 ] [-S suffix] [ name ... ]
```

gzip сжимает (архивирует) файлы **name** с добавлением расширения «.gz» и с сохранением прав и времени. **Исходные файлы удаляются**. Если файлы **не указаны** или в качестве имени используется дефис «-», то выполняется **сжатие стандартного ввода (STDIN) в стандартный вывод (STDOUT)**. Все файлы обрабатываются **раздельно**. Параметры возможно задать **переменной окружения GZIP**

- | | |
|-----------------------------------------------|--------------------------------------------------------------------|
| -c, --stdout, --to-stdout | Использовать стандартный вывод; исходные файлы не удаляются |
| -d, --decompress | |
| --uncompress | Распаковать (аналогично gunzip) |
| -l, --list | Отобразить информацию о сжатых файлах |
| -r, --recursive | Обрабатывать директории рекурсивно, каждый файл отдельно |
| -S <.suf>, --suffix <.suf> | Задать суффикс для архива (по умолчанию «.gz») |
| -t, --test | Проверить целостность архива |
| -<num>, --fast, --best | Задать степень сжатия, где num=1..9 (1=fast, 9=best) |

```
$ gzip -c file1 > foo.gz  
$ gzip -c file2 >> foo.gz  
$ gzip -vl foo.gz  
$ gzip -vt foo.gz
```

РАСПАКОВКА ФАЙЛОВ СЖАТЫХ ПО АЛГОРИТМУ LZ77

```
gunzip [ -acfhlLnNrtvV ] [-S suffix] [ name ... ]
```

gunzip – распаковывает (разархивирует) файлы **name**, удаляя расширение «.gz». Исходные файлы удаляются. Если файлы не указаны или в качестве имени используется дефис «-», то выполняется распаковка **стандартного ввода (STDIN)** в **стандартный вывод (STDOUT)**. Архивными считаются файлы с расширениями «.gz», «-gz», «.z», «-z», «_z», «.Z», «.tgz», «.taz», «.tar.gz» или «.tar.Z» и корректными **magic number (file)**

-c, --stdout, --to-stdout	Использовать стандартный вывод; исходные файлы не удаляются
-l, --list	Отобразить информацию о сжатых файлах
-r, --recursive	Обрабатывать директории рекурсивно, каждый файл отдельно
-S <.suf>, --suffix <.suf>	Задать суффикс для архива (по умолчанию «.gz»)
-t, --test	Проверить целостность архива

```
$ gunzip file.gz  
$ gunzip -c file.gz  
$ gunzip -vl file.gz  
$ gunzip -vt foo.gz
```

ОТОБРАЖЕНИЕ ФАЙЛОВ СЖАТЫХ ПО АЛГОРИТМУ LZ77

```
zcat [ -fhLV ] [ name ... ]
```

zcat эквивалентна распаковке файлов name в стандартный вывод (gunzip -с или gzip -cd).

```
$ zcat foo.gz  
$ zcat foo1.gz foo2.gz > foo
```

- По сути, команда zcat работает идентично команде gunzip -с т.е. распаковывает файлы [name ...], либо стандартный ввод, направляя результат на стандартный вывод.
- Команда zcat работает аналогично команде cat, с тем отличием, что zcat предназначена для запакованных файлов.

ПОСТРАНИЧНЫЙ ПРОСМОТР СОДЕРЖИМОГО АРХИВА

zmore [name...]

zmore постранично отображает текстовые файлы, упакованные программами **compress**, **pack** и **gzip**, или вообще не упакованные. Если файл не существует, выполняется поиск по имени с добавлением расширения «.gz», «.z» или «.Z». Для отображения используется программа-пейджер **more**, которую можно переопределить переменной окружения **PAGER (EDITOR)**. Отображение выполняется с учетом параметров терминала **/etc/termcap (termcap(5))**

Управляющие комбинации клавиш (зависят от программы-пейджера):

<up>, **<down>**, **<left>**, **<right>**, **<pgup>**, **<pgdn>**, **<home>**, **<end>**

<space> Перемотка на полный экран вниз

g Переход в начало файла

G Переход в конец файла

q или **G** Закрытие текущего файла и переход на следующий (если задан)

= Отобразить номер текущей строки

/<expr> Искать выражение **expr** вперед

?<expr> Искать выражение **expr** (назад)

n и **N** Перейти на следующее (предыдущее для **N**) вхождение при поиске

```
$ zmore file1
```

```
$ zmore file1.gz file2.z
```

СЖАТИЕ ФАЙЛОВ ПО АЛГОРИТМУ BURROWS-WHEELER

```
bzip2 [ -cdfkqstvzVL123456789 ] [ filenames ... ]
```

bzip2 сжимает (архивирует) файлы `filenames`, заменяя оригинал и добавляя расширение «.bz2». Права и время сохраняются, каждый файл обрабатывается по отдельности. Существующие файлы не перезаписываются (нужно указать `-f`). Если файлы `filenames` не указаны или в качестве имени используется дефис, то выполняется сжатие STDIN в STDOUT, причем на экран сжатый поток не выводится (нужно использовать перенаправление)

<code>-c, --stdout</code>	Использовать стандартный вывод; исходные файлы не удаляются
<code>-d, -decompress</code>	Разархивировать (аналогично <code>bunzip2</code>)
<code>-f, --force</code>	Принудительно перезаписать существующие выходные файлы
<code>-k, --keep</code>	Не удалять исходных файлов
<code>-<num>, --fast, --best</code>	Задать степень сжатия, где <code>num=1..9</code> (1=fast, 9=best)

```
$ bzip2 -v9 file  
$ bzip2 -v9 file  
$ bzip2 -d file.bz2
```

РАСПАКОВКА ФАЙЛОВ СЖАТЫХ ПО АЛГОРИТМУ BURROWS-WHEELER

```
bunzip2 [ -fkvsVL ] [ filenames ... ]
```

gunzip2 распаковывает (разархивирует) файлы **filenames**, удаляя расширение «.bz2». Существующие файлы не перезаписываются (нужно указать **-f**). Если файлы не указаны или в качестве имени используется дефис, то выполняется распаковка STDIN в STDOUT. Архивными считаются файлы с расширениями «.bz2», «.bz», «.tbz2» или «.tbz», «bzip2»

-c, --stdout	Упаковка/распаковка в стандартный вывод
-f, --force	Принудительно перезаписать существующие выходные файлы
-k, --keep	Не удалять исходных файлов

```
$ bunzip2 file.bz2
```


ОТОБРАЖЕНИЕ ФАЛОВ СЖАТЫХ ПО АЛГОРИТМУ BURROWS-WHEELER

```
bzcat [ -s ] [ filenames ... ]
```

bzcat эквивалентна распаковке файлов filenames в стандартный вывод (bzip2 -с или bunzip -cd)

```
$ bzcat file.bz2
```

УПАКОВКА/РАСПАКОВКА ПО АЛГОРИТМУ BURROWS-WHEELER

bzip2recover filename

- Восстановить поврежденный bzip2 архив
- Используется для восстановления данных из испорченных файлов, заархивированных с помощью bzip

```
$ bzip2recover file.bz2
```

СРАВНИТЬ УПАКОВАННЫЕ ФАЙЛЫ

zcmp [options] file [file2]

zdiff [options] file [file2]

Zcmp и zdiff – фильтры, которые вызывают cmp и diff соответственно для распакованных во временную директорию. Если указан только один файл, сравнение происходит с таким же файлом, но без расширения.

TMPDIR

Временная директория для распаковки

```
$ zcmp file.gz
```

```
$ zdiff file1.gz file2.gz
```

СОЗДАНИЕ АВТОМАТИЧЕСКИ РАСПАКОВЫВАЕМЫХ ИСПОЛНЯЕМЫХ ФАЙЛОВ

`gzexe [-d] file ...`

- Распаковка происходит на лету, при запуске архива.
- Исходный файл сохраняется с суффиксом «~».
- Программа не обрабатывает не исполняемых, не регулярных файлов и файлов с установленным битом `setuid` или `setgid`.

`-d`

Распаковка

```
$ gzexe ./ls  
$ gzexe -d ./ls
```

КОМПРЕССИЯ/ДЕКОМПРЕССИЯ .xz И .lzma ФАЙЛОВ

Опции xz:

- z, --compress
- d, --decompress
- t, --test
- l, --list
- c, --stdout, --to-stdout
- k, --keep
- files[=file]
- 0 ... -3 – fast compression
- 4 ... -6 – good to very good compression
- 7 ... -9 – useful for big files (e.g. 8...16...32 MB)

СРАВНЕНИЕ ЭФФЕКТИВНОСТИ РАЗЛИЧНЫХ АЛГОРИТМОВ КОМПРЕССИИ

Input data	Algo	Time (C)	Time (D)	Ratio
ImageMagick modules	gzip	0.30	0.04	25.28%
	bzip2	0.61	0.27	20.44%
	lzma	4.97	0.10	12.64%
Lena (TIFF)	gzip	0.05	0.01	93.24%
	bzip2	0.15	0.08	74.31%
	lzma	0.33	0.06	80.69%
Lena (PNG)	gzip	0.02	0.00	100.02%
	bzip2	0.13	0.05	100.32%
	lzma	0.18	0.04	100.67%
Linux Sources	gzip	12.98	1.88	21.81%
	bzip2	54.61	14.25	17.07%
	lzma	287.58	5.46	14.47%
War and Peace	gzip	0.30	0.03	37.00%
	bzip2	0.58	0.26	26.96%
	lzma	3.30	0.09	28.34%
Overall (Average)	gzip	2.73	0.39	55.47%
	bzip2	11.22	2.98	47.82%
	lzma	59.27	1.15	47.36%

lenna.tiff

The compression testing poster girl

lenna.png

The same image as above, as a PNG, to see what the algorithms make of already-compressed data

linux-2.6.25.tar

The sources for the linux kernel, version 2.6.25

wrnp12.txt

Leo Tolstoy: War and Peace. From Project Gutenberg. ASCII Plaintext.

ImageMagick-6.4.0_modules-Q16.tar

A tar file /usr/lib64/ImageMagick-6.4.0/modules-Q16/ from amd64 system, containing many .so, .la and .a files

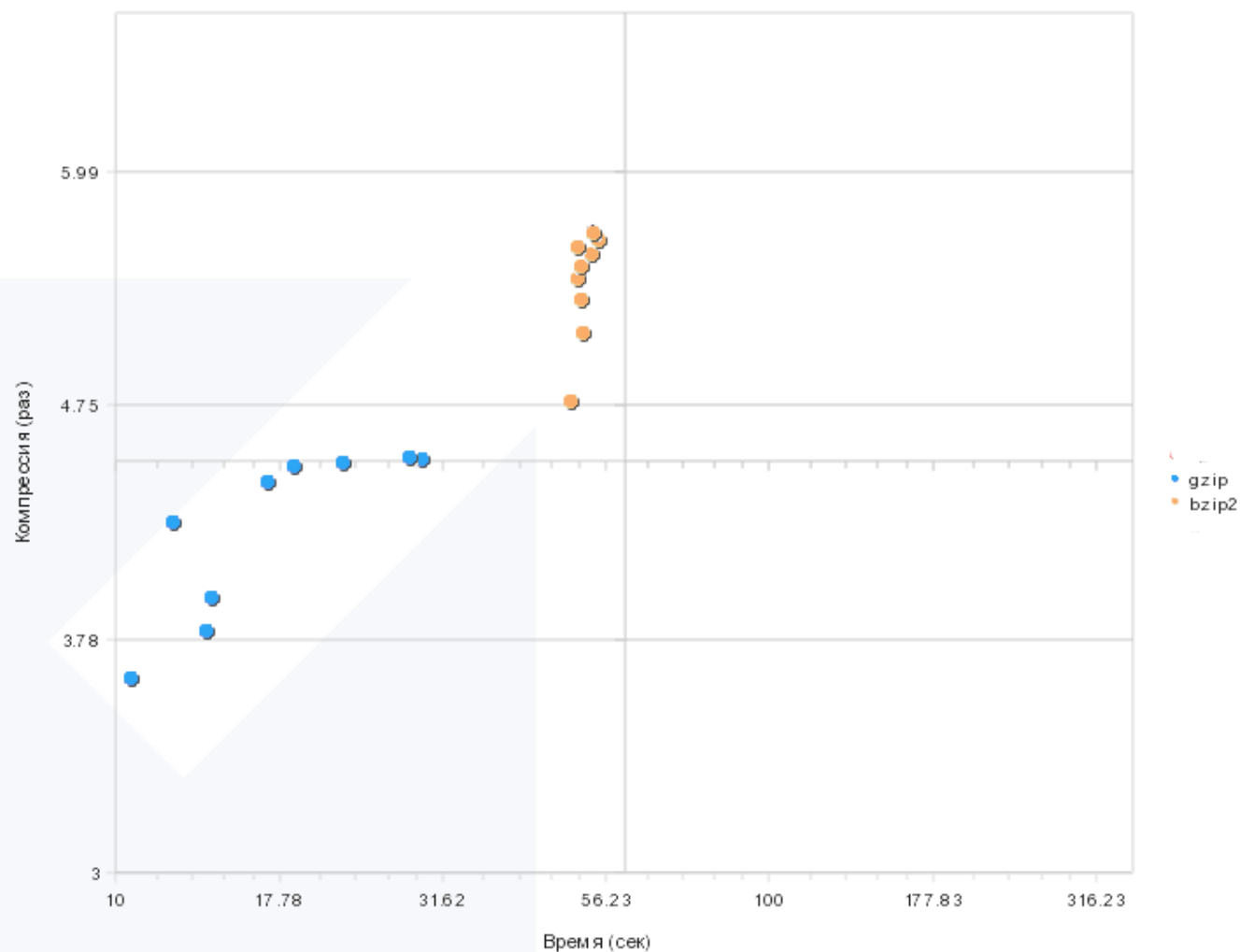
Algo – алгоритм

Time (C) – время компрессии

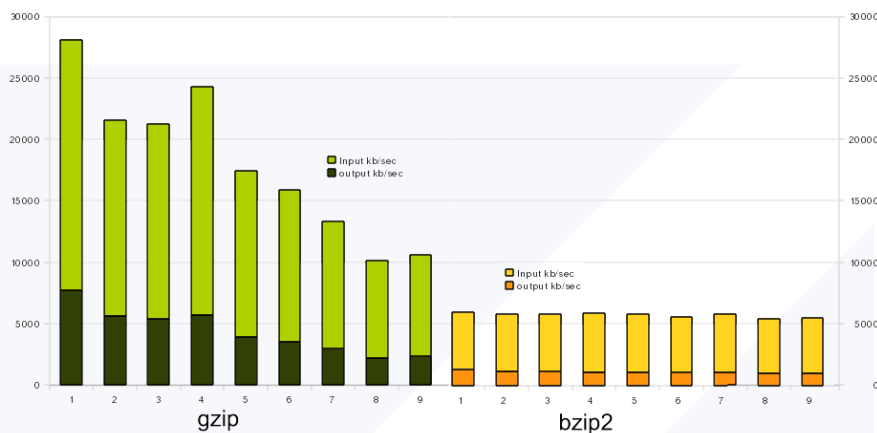
Time (D) – время декомпрессии

Ratio, % – степень сжатия

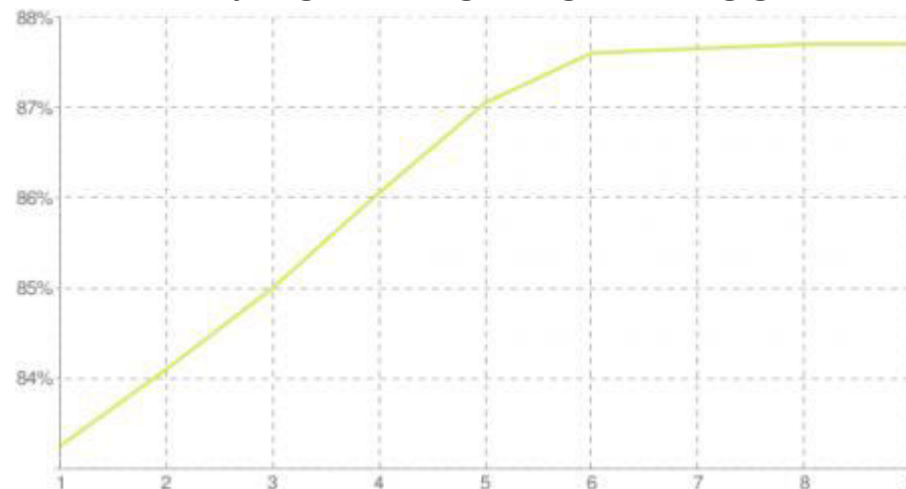
СРАВНЕНИЕ ЭФФЕКТИВНОСТИ РАЗЛИЧНЫХ АЛГОРИТМОВ КОМПРЕССИИ



СРАВНЕНИЕ ЭФФЕКТИВНОСТИ РАЗЛИЧНЫХ АЛГОРИТМОВ КОМПРЕССИИ

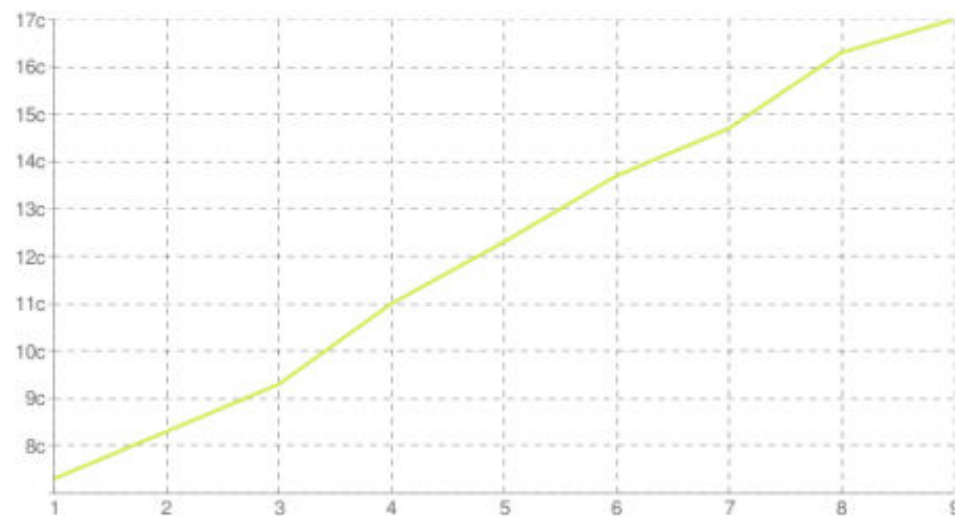


эффективность сжатия



степень сжатия

процессорное время



степень сжатия

РАБОТА С АРХИВАМИ

tar {-c} [options] [files | directories]

tar -f archive-file [options] [files | directories]

tar {-t | -x} [options] [patterns]

-c	Создать новый архив, содержащий перечисленные файлы (директории)
-t	Прочитать содержание архива
-x	Извлечь файлы из архива (с заменой)
-j	Упаковать/распаковать архив используя bzip2
-z	Упаковать/распаковать архив используя gzip
-f file	Работа с указанным файлом
-C directory	Сменить директорию до выполнения основных действий
-m	Не «распаковывать» время модификации
-p	Сохранить права файлов (только для -x)
-v	Отображать дополнительную информацию
--strip-components count	Удалить указанное число начальных компонентов пути (-x и -t)

```
$ tar -czf file.tar.gz source.c source.h  
$ tar -ztvf file.tar.gz  
$ tar -cf - -C srcdir . | tar -xpf - -C destdir  
$ cd srcdir ; tar -cf - . | (cd destdir ; tar -xpf -)
```