Bzip2 mini-KAKO

David Fetter, *david@fetter.org* verzija 2.00, 22. kolovoza 1999.

Ovaj dokument govori o korištenju `bzip2`, novog programa za sažimanje. Autorativni SGML nalazi se na `<http://fetter.org/Bzip2-HOWTO/>`.

Sadržaj

1 Uvod

bzip2 je zgodan novi algoritam za sažimanje podataka. Općenito, datoteke čini 30-40% manjim od veličine njihovih gzip-anih pandana.

Ovaj dokument vodi vas kroz nekoliko uobičajenih primjena bzip2-a.

Buduće verzije dokumenta govorit će o primjenama libbzip2, bzip2 C librarya kojeg je ljubazno napisao autor bzip2, Julian Seward, *Julian_Seward@muraroa.demon.co.uk*. Detaljne informacije o libraryu sadrži bzip2 priručnik dostupan na <http://www.bzip2.org/bzip2/docs/manual_toc.html>.

Buduće verzije dokumenta možda će također sadržavati sažetak rasprava o tome bi li (i kako bi) trebalo koristiti bzip2 u Linux kernelu.

1.1 Popis promjena

Verzija 2.00

Promijenio poglavlje 6 (Korištenje bzip2 uz less) tako da se .tar.bzip2 datoteke mogu čitati. Hvala Nicoli Fabianu *mc8644@mclink.it* za ispravak.

Osvježio alat buzzit.

Osvježio informacije o taru.

Verzija 1.92

Osvježio poglavlje 2.2 (Nabavljanje već kompajliranih izvršnih datoteka bzip2) i dodao S.u.S.E.-ove.

Verzija 1.91

Ispravio pravopisnu grešku i razjasnio neke idiome ljuske u poglavlju 5 (Korištenje bzip2 uz tar). Za to hvala Alessandru Rubiniju.

Sredio buzzit alat tako da se ne zagrcne na postojećoj bzip2 arhivi.

Dodao bgrep, zgrep-olik alat.

Verzija 1.9

Razjasnio gcc 2.7.* problem. Hvala Ulriku Dickowu što me upozorio.

Dodao elegantni način rada s tarom Leonarda Jean-Marca.

Dodao švedski prijevod Linusa Åkerlunda.

Sredio wu-ftpd poglavlje prema prijedlogu Arnauda Launaya.

Prijevode premjestio u zasebno poglavlje.

Verzija 1.8

buzzit i tar.diff u SGML-u stavio tamo gdje spadaju. Sredio interpunkcije i formatiranje. Hvala Arnaudu Launayu na pomoći u ispravljanju mog primjerka. :-)

Za sada ispustio xv projekt zbog nedostatnog zanimanja čitatelja.

Dodao neodgovorena pitanja za buduće verzije dokumenta.

Verzija 1.7

Dodao alat buzzit. Sredio zakrpu za GNU tar.

Verzija 1.6

Dodao TenThumbsov program za Netscape.

Promijenio i lesspipe.sh prema njegovom prijedlogu. Sad bi trebao raditi bolje.

Verzija 1.5

Dodao francuski prijevod Arnauda Launaya i njegovu datoteku za `wu-ftpd`.

Verzija 1.4

Dodao japanski prijevod Tetsua Isajia.

Verzija 1.3

Dodao `.emacs` Ulrika Dickowa za 19.30 i bolji.

(Ispravio i `jka-compr.el` prema njegovu prijedlogu. Ups! `bzip2` još uvijek (?) nema opciju `-append`.)

Verzija 1.2

Promijenio zakrpu za `emacs` tako da automatski prepoznaje `.bz2` datoteke.

Verzija 1.1

Dodao zakrpu za `emacs`.

Verzija 1.0

Prva runda.

2 Nabavljanje bzip2

WWW stranica `bzip2` nalazi se na `<http://www.bzip2.org/>`. Mirror u SAD je `<http://www.digistar.com/bzip2/index.html>`.

2.1 Bzip2 mini-HOWTO na vašem jeziku

Francuski čitatelji mogu pročitati francuski prijevod Arnauda Launaya. WWW verzija je na `<http://www.freenix.fr/linux/HOWTO/mini/Bzip2.html>`, a FTP-om ga možete dobiti s `<ftp://ftp.lip6.fr/pub/linux/french/docs/HOWTO/mini/Bzip2.gz>`. Arnauda možete kontaktirati elektronskom poštom na *zoro@mygale.org*.

Japanski čitatelji mogu pročitati japanski prijevod Tetsua Isajia na `<http://jf.gee.kyoto-u.ac.jp/JF/JF.html>`. Do Isajia možete doći preko njegove WWW stranice, `<http://www2s.biglobe.ne.jp/~kaien/>`.

Švedski čitatelji mogu pročitati švedski prijevod Linusa Åkerlunda na `<http://user.tninet.se/~uxm165t/linux_doc.html>`. Linusova e-mail adresa je *uxm165t@tninet.se*.

2.2 Nabavljanje već kompajliranih izvršnih datoteka bzip2

Pogledajte WWW stranice `bzip2`.

2.3 Nabavljanje izvornog koda bzip2

Može se naći na službenim stranicama (pogledajte 2 (Nabavljanje bzip2)).

3 Kompajliranje bzip2 za vaš stroj

Ako imate gcc **2.7.***, promijenite red

```
CFLAGS = -O3 -fomit-frame-pointer -funroll-loops
```

u

```
CFLAGS = -O2 -fomit-frame-pointer
```

To jest, `-O3` zamijenite s `-O2` i izostavite `-funroll-loops`. Također možete dodati `-m*` prekidače (`-m486`, na primjer) koje koristite kod kompajliranja kernela.

Najvažnije je izbjeći `-funroll-loops` jer će zbog toga mnogi `gcc`-i verzije 2.7 generirati neispravan kod, a svi `gcc`-i 2.7 generirati sporiji i veći kod. Za druge kompajlere (`lcc`, `egcs`, `gcc` 2.8.x) CFLAGS nije potrebno mijenjati.

Nakon toga samo `make` i instalirajte ga prema `README`-u.

4 Korištenje samog bzip2

RTFM (pročitajte tu dobru man stranicu :).

5 Korištenje bzip2 uz tar

Evo tri načina korištenja `bzip2` uz `tar`, to jest:

5.1 Najlakši za namještanje

Ova metoda ne zahtijeva nikakvo petljanje. Za dearhiviranje `bzip2`-ane `tar` arhive `foo.tar.bz2` u trenutnom direktoriju napišite:

```
/staza/do/bzip2 -cd foo.tar.bz2 | tar xf -
```

ili

```
tar --use-compress-prog=bzip2 xf foo.tar.bz2
```

To radi, ali toliko tipkati često može biti mučno.

5.2 Lak za namještanje, relativno lak za korištenje, bez potrebe za root privilegijama

Hvala Leonardu Jean-Marcu, *leonard@sct1.is.belgacom.be*, što mi je ovo spomenuo. Također hvala Alessandru Rubiniju, *rubini@morgana.systemy.it*, za razlike između `basha` i `csha`.

U svoj `.bashrc` možete staviti ovakav red:

```
alias btar='tar --use-compress-program /usr/local/bin/bzip2 '
```

Odgovarajući red za `.tcshrc` ili `.cshrc` izgleda ovako:

```
alias btar 'tar --use-compress-program /usr/local/bin/bzip2'
```

5.3 Isto tako lak za korištenje, ali zahtijeva root pristup

Nadogradite svoj `tar` na najnoviju GNU verziju, trenutno 1.13.10. Ona se može naći na GNU-ovom FTP poslužitelju <ftp://alpha.gnu.org/gnu/tar/> i svim mirrorima.

6 Korištenje bzip2 uz less

Za dekompresiju `bzip2`-anih datoteka u hodu, odnosno mogućnost korištenja `lessa` bez `bunzip2`-avanja, možete napraviti `lesspipe.sh` (pogledajte `man less`) poput ovog:

```
#!/bin/sh
# Ovo je preprocesor za less. Koristi se kada je postavljena varijabla
# okružja:  LESSOPEN="|lesspipe.sh %s"

  case "$1" in
  *.tar) tar tvvf $1 2>/dev/null ;; # Prikaz sadržaja raznih tar datoteka
  *.tgz) tar tzvvf $1 2>/dev/null ;;
# Ovo će raditi za nepromijenjenu verziju tara:
  *.tar.bz2) bzip2 -cd $1 $1 2>/dev/null | tar tvvf - ;;
# Ovo radi za zakrpanu verziju tara:
# *.tar.bz2) tyvvf $1 2>/dev/null ;;
  *.tar.gz) tar tzvvf $1 2>/dev/null ;;
  *.tar.Z) tar tzvvf $1 2>/dev/null ;;
  *.tar.z) tar tzvvf $1 2>/dev/null ;;
  *.bz2) bzip2 -dc $1  2>/dev/null ;; # Ispravan pregled sažetih datoteka
  *.Z) gzip -dc $1  2>/dev/null ;;
  *.z) gzip -dc $1  2>/dev/null ;;
  *.gz) gzip -dc $1  2>/dev/null ;;
  *.zip) unzip -l $1 2>/dev/null ;;
  *.1|*.2|*.3|*.4|*.5|*.6|*.7|*.8|*.9|*.n|*.man) DATOTEKA=`file -L $1` ; # u groffu
    DATOTEKA=`echo $DATOTEKA | cut -d ' ' -f 2`
    if [ "$DATOTEKA" = "troff" ]; then
      groff -s -p -t -e -Tascii -mandoc $1
    fi ;;
  *) cat $1 2>/dev/null ;;
#  *) DATOTEKA=`file -L $1` ; # Provjeravamo je li je binarna -- pregled preko strings
#    DATOTEKA1=`echo $DATOTEKA | cut -d ' ' -f 2`
#    DATOTEKA2=`echo $DATOTEKA | cut -d ' ' -f 3`
#    if [ "$DATOTEKA1" = "Linux/i386" -o "$DATOTEKA2" = "Linux/i386" \
#         -o "$DATOTEKA1" = "ELF" -o "$DATOTEKA2" = "ELF" ]; then
#      strings $1
#    fi ;;
  esac
```

7 Korištenje bzip2 uz emacs

7.1 Mijenjanje emacsa za sve

Napisao sam sljedeću zakrpu za `jka-compr.el`. Ona dodaje `bzip2` u auto-compression-mode.

Napomena: Ovo sam testirao samo na `emacsu` 20.2, ali ne vidim razloga zašto sličan pristup ne bi radio na drugim verzijama.

Primijenite ju ovako:

1. Otiđite u direktorij `emacs-20.2/lisp` s izvornim kodom (gdje god ste ga otpakirali).
2. Snimite ovu zakrpu u datoteku `jka-compr.el.diff` (u toj bi datoteci trebala biti samo ona ;).
3. Napišite:

   ```
   patch < jka-compr.el.diff
   ```

4. Pokrenite `emacs` i napišite

   ```
    M-x byte-compile-file jka-compr.el
   ```

5. Izađite iz `emacsa`.
6. Svoj izvorni `jka-compr.elc` maknite na sigurnije mjesto za slučaj greške.
7. Uživajte!

```
--- jka-compr.el        Sat Jul 26 17:02:39 1997
+++ jka-compr.el.new    Thu Feb  5 17:44:35 1998
@@ -44,7 +44,7 @@
 ;; The variable, jka-compr-compression-info-list can be used to
 ;; customize jka-compr to work with other compression programs.
 ;; The default value of this variable allows jka-compr to work with
-;; Unix compress and gzip.
+;; Unix compress and gzip.  David Fetter added bzip2 support :)
 ;;
 ;; If you are concerned about the stderr output of gzip and other
 ;; compression/decompression programs showing up in your buffers, you
@@ -121,7 +121,9 @@


 ;;; I have this defined so that .Z files are assumed to be in unix
-;;; compress format; and .gz files, in gzip format.
+;;; compress format; and .gz files, in gzip format, and .bz2 files,
+;;; in the snappy new bzip2 format from http://www.muraroa.demon.co.uk.
+;;; Keep up the good work, people!
 (defcustom jka-compr-compression-info-list
   ;;[regexp
   ;; compr-message  compr-prog  compr-args
@@ -131,6 +133,10 @@
      "compressing"    "compress"     ("-c")
      "uncompressing"  "uncompress"   ("-c")
      nil t]
+    ["\\.bz2\\'"
+     "bzip2ing"        "bzip2"         ("")
+     "bunzip2ing"      "bzip2"         ("-d")
+     nil t]
     ["\\.tgz\\'"
      "zipping"        "gzip"         ("-c" "-q")
      "unzipping"      "gzip"         ("-c" "-q" "-d")
```

7.2 Mijenjanje emacsa za jednu osobu

Za ovo hvala Ulriku Dickowu, *ukd@kampsax.dk*, sistemskom programeru iz Kampsax Technology.

Ako želite automatski koristiti `bzip2` kada niste root, samo dodajte ovo svojoj `.emacs` datoteci.

```
;; Automatska (de)kompresija kod učitavanja/snimanja datoteka (gzip(1) i
;;   slično)
;; Pokrećemo ju u ugašenom stanju tako da se može dodati podrška za
;; bzip2(1). Kod skupio Ulrik Dickow za ~/.emacs uz Emacs 19.34. Trebalo bi
;; raditi i s mnogim starijim i novijim Emacsima. Ipak, nema jamstva.
;;
(if (fboundp 'auto-compression-mode) ; Emacs 19.30+
    (auto-compression-mode 0)
  (require 'jka-compr)
  (toggle-auto-compression 0))
;; Sad dodajmo bzip2 podršku i ponovo uključimo automatsku kompresiju
(add-to-list 'jka-compr-compression-info-list
             ["\\.bz2\\(~\\|\\.~[0-9]+~\\)?\\'"
              "zipping"        "bzip2"         ()
              "unzipping"      "bzip2"         ("-d")
              nil t])
(toggle-auto-compression 1 t)
```

8 Korištenje bzip2 uz wu-ftpd

Hvala Arnaudu Launayu za ovo spašavanje bandwidtha. Sljedeće treba staviti u `/etc/ftpconversions` za de/kompresiju u hodu pomoću `bzip2`. Provjerite jesu li staze (kao što je `/bin/compress`) ispravne.

```
 :.Z:  :  :/bin/compress -d -c %s:T_REG|T_ASCII:O_UNCOMPRESS:UNCOMPRESS
 :   : :.Z:/bin/compress -c %s:T_REG:O_COMPRESS:COMPRESS
 :.gz: :  :/bin/gzip -cd %s:T_REG|T_ASCII:O_UNCOMPRESS:GUNZIP
 :   : :.gz:/bin/gzip -9 -c %s:T_REG:O_COMPRESS:GZIP
 :.bz2: :  :/bin/bzip2 -cd %s:T_REG|T_ASCII:O_UNCOMPRESS:BUNZIP2
 :   : :.bz2:/bin/bzip2 -9 -c %s:T_REG:O_COMPRESS:BZIP2
 :   : :.tar:/bin/tar -c -f - %s:T_REG|T_DIR:O_TAR:TAR
 :   : :.tar.Z:/bin/tar -c -Z -f - %s:T_REG|T_DIR:O_COMPRESS|O_TAR:TAR+COMPRESS
 :   : :.tar.gz:/bin/tar -c -z -f - %s:T_REG|T_DIR:O_COMPRESS|O_TAR:TAR+GZIP
 :   : :.tar.bz2:/bin/tar -c -I -f - %s:T_REG|T_DIR:O_COMPRESS|O_TAR:TAR+BZIP2
```

9 Korištenje bzip2 s grepom

Ovaj alat, kojeg sam nazvao `bgrep`, je malčice izmijenjen `zgrep` koji dolazi s Linuxom. Možete ga koristiti za `grep` kroz datoteke bez njihovog `bunzip2`-anja.

```
#!/bin/sh

# bgrep -- omotač oko programa grep koji po potrebi dekomprimira datoteke
PATH="/usr/bin:$PATH"; export PATH
```

```
prog=`echo $0 | sed 's|.*/||'`
case "$prog" in
        *egrep) grep=${EGREP-egrep}     ;;
        *fgrep) grep=${FGREP-fgrep}     ;;
        *)      grep=${GREP-grep}       ;;
esac
pat=""
while test $# -ne 0; do
  case "$1" in
  -e | -f) opt="$opt $1"; shift; pat="$1"
           if test "$grep" = grep; then  # grep pati od grešaka s -e na SVR4
             grep=egrep
           fi;;
  -*)      opt="$opt $1";;
   *)      if test -z "$pat"; then
             pat="$1"
           else
             break;
           fi;;
  esac
  shift
done

if test -z "$pat"; then
  echo "grep kroz bzip2 datoteke"
  echo "usage: $prog [grep_opcije] uzorak [datoteke]"
  exit 1
fi

list=0
silent=0
op=`echo "$opt" | sed -e 's/ //g' -e 's/-//g'`
case "$op" in
  *l*) list=1
esac
case "$op" in
  *h*) silent=1
esac

if test $# -eq 0; then
  bzip2 -cd | $grep $opt "$pat"
  exit $?
fi

res=0
for i do
  if test $list -eq 1; then
    bzip2 -cdfq "$i" | $grep $opt "$pat" > /dev/null && echo $i
    r=$?
```

```
  elif test $# -eq 1 -o $silent -eq 1; then
    bzip2 -cd "$i" | $grep $opt "$pat"
    r=$?
  else
    bzip2 -cd "$i" | $grep $opt "$pat" | sed "s|^|${i}:|"
    r=$?
  fi
  test "$r" -ne 0 && res="$r"
done
exit $res
```

10 Korištenje bzip2 uz Netscape pod X-om

tenthumbs@cybernex.net kaže:

> Shvatio sam i kako natjerati Netscape za Linux da koristi `bzip2` za Content-Encoding baš kao što i koristi `gzip`. Dodajte ovo u `\$HOME/.Xdefaults` ili `\$HOME/.Xresources`:
>
> Koristim opciju `-s` samo zato što više volim sporiju dekompresiju nego manje RAM-a. Ako želite, možete ju izostaviti.

```
Netscape*encodingFilters:      \
        x-compress :  : .Z     : uncompress -c  \n\
        compress   :  : .Z     : uncompress -c  \n\
        x-gzip     :  : .z,.gz : gzip -cdq      \n\
        gzip       :  : .z,.gz : gzip -cdq      \n\
        x-bzip2    :  : .bz2   : bzip2 -ds \n
```

11 Korištenje bzip2 za ponovno sažimanje drugih sažetih datoteka

Ovaj perl program uzima datoteke sažete u drugim formatima (`.tar.gz`, `.tgz`, `.tar.Z` i `.Z` za ovu iteraciju) i ponovno ih sažima radi bolje kompresije. Izvorni kod ima dosta zgodne dokumentacije o tome što radi i kako to radi. Ova, najnovija verzija ulazne datoteke prima preko naredbenog retka. Ako nema argumenata, pokušat će prepakirati svaku datoteku u radnom direktoriju.

```
#!/usr/bin/perl -w

#######################################################
#                                                     #
# Ovaj program uzima compress-ane i gzip-ane datoteke #
# u trenutnom direktoriju i pretvara ih u bzip2       #
# format. S nastavkom .tgz ponaša se razumno,         #
# stvarajući .tar.bz2 datoteku.                       #
#                                                     #
#######################################################
$brojač = 0;
$bajta_snimljeno = 0;
$ukupno_datoteka = '/tmp/bzip2_ukupno_na_stroju';
```

```
$bzip2_ukupno_na_stroju = 0;

@zadane = (defined @ARGV)?@ARGV:<*>;

foreach(@zadane) {
    next if /^bzip/;
    next unless /\.(tgz|gz|Z)$/;
    push @datoteke, $_;
}
$ukupno = scalar(@datoteke);

foreach (@datoteke) {
    if (/tgz$/) {
        ($novo=$_) =~ s/tgz$/tar.bz2/;
    } else {
        ($novo=$_) =~ s/\.g?z$/.bz2/i;
    }
    $izv_koliko = (stat $_)[7];
    ++$brojač;
    print "Ponovo sažimam $_ ($brojač/$ukupno)...\n";
    if ((system "gzip -cd $_ |bzip2 >$novo") == 0) {
        $novo_koliko = (stat $novo)[7];
        $faktor = int(100*$novo_koliko/$izv_koliko+.5);
        $bajta_snimljeno += $izv_koliko-$novo_koliko;
        print "$novo zauzima oko $faktor% veličine $_. :",($faktor<100)?')':'(',"\n";
        unlink $_;
    } else {
        print "Arrgg! Nešto se dogodilo s $_: $!\n";
    }
}
print ($bajta_snimljeno>=0)?"Dobili ":"Izgubili "
    , "ste "
    , abs($bajta_snimljeno)
    , " bajta diskovnog prostora :"
    , ($bajta_snimljeno>=0)?")":"("
    , "\n"
    ;

unless (-e '/tmp/bzip2_ukupno_na_stroju') {
    system ('echo "0" >/tmp/bzip2_ukupno_na_stroju');
    system ('chmod', '0666', '/tmp/bzip2_ukupno_na_stroju');
}


chomp($bzip2_ukupno_na_stroju = `cat $ukupno_datoteka`);
open UKUPNO, ">$ukupno_datoteka"
     or die "Ne mogu otvoriti datoteku s ukupnim rezultatem stroja: $!";
$bzip2_ukupno_na_stroju += $bajta_snimljeno;
print UKUPNO $bzip2_ukupno_na_stroju;
close UKUPNO;
```

```
print "Ukupni rezultat cijelog stroja je ",`cat $ukupno_datoteka`," bajta.\n";
```

12 Hrvatski prijevod

Najnovija verzija ovog prijevoda može se naći na `<http://dokumentacija.linux.hr/Bzip2.html>`. Održava ga Matej Vela, *mvela@public.srce.hr*. Svi su prilozi, primjedbe i prijedlozi dobrodošli.