Bzip2 Howto

David Fetter, *dfetter@best.com* <mailto:dfetter@best.com>. Översättning: Linus Åkerlund, *uxm165t@tninet.se* <mailto:uxm165t@tninet.se>. v1.8 Tue Jun 23 07:44:14 PDT 1998. Översättning: 29 juni 1998

Detta dokument går igenom hur du kan använda det nya komprimeringsprogrammet bzip2. Den ursprungliga SGML-versionen finns *här* <http://www.best.com/~dfetter/Bzip2-HOWTO/Bzip2-HOWTO.sgml>.

Innehåll

1 Inledning

Bzip2 är en häftig ny algoritm för att komprimera data. Vanligtvis gör det filer till 60-70% av storleken av sina gzippade motsvarigheter.

Detta dokument kommer leda dig igenom ett par vanliga användningsområden för bzip2.

Framtida versioner av det här dokumentet kommer diskutera det kommande bzip2-biblioteket, vilket programmeraren bakom bzip2, (`JulianSeward` beskriver på följande sätt:

```
Jag arbetar på nästa version av bzip2, vilket kommer använda samma
filformat (.bz2); det viktigaste tillägget är ett zlib-liknande
bibliotek för att läsa och skriva data i detta filformat inifrån
program.
```

Framtida versioner av detta dokument kan också komma att innehålla en summering av diskussionen om huruvida (och hur) bzip2 bör användas i Linux-kärnan.

1.1 Ändringar

1.1.1 v1.9

Redde upp gcc 2.7.*-problemet. Tack till Ulrik Dickow för att han påpekade detta.

la till Leonard Jean-Marc's eleganta sätta att arbeta med tar.

La till Linus Åkerlunds svenska översättning.

Fixade avsnittet om wu-ftpd enligt Arnaud Launays förslag.

Flytta översättningarna till en egen avdelning.

1.1.2 v1.8

La buzzit och tar.diff där de hör hemma i SGMLen. Fixade kommatering och formattering. Tack till Arnaud Launay för hans hjälp med att rätta till mitt exemplar. :-)

La ner xv-projektet för tillfället, eftersom det inte verkar intressera folk i tillräckligt hög grad.

La till några aptitretare om framtida versioner av dokumentet.

1.1.3 v1.7

La till buzzit-verktyget. Fixade patchen mot gnu tar.

1.1.4 v1.6

La till TenThumbs Netscape-tillägg.

Ändrade även lesspipe.sh enligt hans förslag. Det bör fungera bättre nu.

1.1.5 v1.5

La till Arnaud Launays franska översättning och hans wu-ftpd-fil.

1.1.6 v1.4

la till Tetsu Isajis japanska översättning.

1.1.7 v1.3

La till Ulrik Dickows .emacs för 19.30 och högre.

(Rättade också till jka-compr.el-patchen för emacs enligt hans förslag. Hoppsan! Bzip2 har ännu(?) inte en "append"-flagga.)

1.1.8 v1.2

Ändrade patchen för emacs så att det automatiskt känner igen .bz2-filer.

1.1.9 v1.1

La till patchen för emacs.

1.1.10 v1.0

Första omgången.

2 Skaffa bzip2

Bzip2s hemsida finns *här* `<http://www.muraroa.demon.co.uk/>`. En spegel i USA finns *här* `<http://www.digistar.com/bzip2/index.html>`. Du kan också hitta det på *Red Hats ftp-sajt* `<ftp://ftp.redhat.com/pub/contrib>`.

2.1 Bzip2-HOWTOn på ditt eget språk

Fransktalande kan ta en titt på Arnaud Launays franska översättning. Webb-versionen finns *här* `<http://www.freenix.fr/linux/HOWTO/mini/Bzip2.html>`, och du kan ftp-a den från *denna sajt* `<ftp://ftp.lip6.fr/pub/linux/french/docs/HOWTO/mini/Bzip2.gz>` Arnaud kan kontaktas via e-post på `dennaadress`.

De som talar japanska kan ta en titt på Tetsu Isajis japanska översättning *här* `<http://jf.gee.kyoto-u.ac.jp/JF/JF.html>`. Isaji når du via *hans hemsida* `<http://www2s.biglobe.ne.jp/~kaien/>`, eller via e-post, på `denh\"aradressen`.

Svensktalande kan hitta Linus Åkerlunds svenska dokument *här* `<http://user.tninet.se/~uxm165t/linux_doc.html>`. Linus kan nås med e-post via *denna adress* `<mailto:uxm165t@tninet.se>`. (Jo, ähum, jag valde att översätta även detta.övers.anm :-)

2.2 Skaffa förkompilerade binärfiler av bzip2

Se hemsidan. Red Hats binärpaket för Intel finns *här* `<ftp://ftp.redhat.com/pub/contrib/i386/bzip2-0.1pl2-1.i386.rpm>`. Debians finns *här* `<ftp://ftp.debian.org/debian/hamm/hamm/binary-i386/utils/bzip2_0.1pl2-4.deb>`, och Slackwares finns *här* `<http://www.cdrom.com/pub/linux/slackware-3.4/slakware/a9/bzip2.tgz>`. Du kan också hitta dem på motsvarande ställen på diverse spegel-sajter.

2.3 Skaffa källkoden till bzip2

De finns på den officiella sajten (se 2 (Skaffa Bzip2), eller hämta den från *Red Hat* `<ftp://ftp.redhat.com/pub/contrib/SRPMS/bzip2-0.1pl2-1.src.rpm>`).

2.4 Kompilera bzip2 för din maskin

Om du har gcc 2.7.*, ändra raden

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

till

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

Byt alltså ut -O3 mot -O2 och ta bort -funroll-loops. Du kan även lägga till -m*-flaggan (t.ex. -m486) som du brukar använda då du kompilerar kärnor.

Att undvika -funroll-loops är det viktigaste, eftersom det kan få många gcc 2.7-or att generera felaktig kod och får alla gcc 2.7-or att generera långsammare och större kod. För andra kompilatorer (lcc, egcs, gcc 2.8.x) går det bra med de vanliga CFLAGS.

Sedan är det bara att köra ztt/make/ och installera enligt README-filen.

3 Använda bzip2 självt

Läs den fina man-sidan :)

4 Använda bzip2 med tar

Här kommer tre sätt att använda bzip2 med tar:

4.1 Enklast att konfigurera:

Denna metod kräver ingen konfigurering alls. För att packa upp det bzip2-ade tar-arkivet foo.tar.bz i den aktuella katalogen, skriv

```
/path/to/bzip2 -cd foo.tar.bz2 | tar xf -
```

Detta fungerar, men det kan vara jobbigt att behöva göra det ofta.

4.2 Enkelt att konfigurera, ganska enkelt att använda, du behöver inte vara root:

Tack till *Leonard Jean-Marc* `<mailto:leonard@sct1.is.belgacom.be>` för tipset.

Lägg in en rad som denna i din .bashrc (eller .tcshrc eller .cshrc):

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

Also easy to use, but needs root access.

4.3 Också enkelt att använda, men kräver root-rättigheter:

Lägg till patchen nedan till gnu tar 1.12 enligt följande metod:

```
cd tar-1.12/src; patch < /sökväg/till/tar.diff
```

Kompilera och installera det och sedan är det bara att köra. Se efter så att både tar och bzip2 finns i din $PATH genom "which tar" och "which bzip2". För att använda det nya tar, skriv bara

```
tar xyf foo.tar.bz2
```

för att packa upp filen.

Att skapa ett nytt arkiv går till på ett liknande sätt:

```
tar cyf foo.tar.bz2 fil1 fil2 fil3...katalog1 katalog2...
```

Här är patchen:

```
*** tar.c       Thu Jun 11 00:09:23 1998
--- tar.c.new   Thu Jun 11 00:14:24 1998
***************
*** 196,201 ****
--- 196,203 ----
    {"block-number", no_argument, NULL, 'R'},
    {"block-size", required_argument, NULL, OBSOLETE_BLOCKING_FACTOR},
    {"blocking-factor", required_argument, NULL, 'b'},
+   {"bzip2", required_argument, NULL, 'y'},
+   {"bunzip2", required_argument, NULL, 'y'},
    {"catenate", no_argument, NULL, 'A'},
    {"checkpoint", no_argument, _option, 1},
    {"compare", no_argument, NULL, 'd'},
```

```
***************
*** 372,377 ****
--- 374,380 ----
               PATTERN                at list/extract time, a globbing PATTERN\n\
    -o, --old-archive, --portability   write a V7 format archive\n\
        --posix                        write a POSIX conformant archive\n\
+   -y, --bzip2, --bunzip2             filter the archive through bzip2\n\
    -z, --gzip, --ungzip               filter the archive through gzip\n\
    -Z, --compress, --uncompress       filter the archive through compress\n\
        --use-compress-program=PROG    filter through PROG (must accept -d)\n"),
***************
*** 448,454 ****
     Y  per-block gzip compression */

  #define OPTION_STRING \
!   "-01234567ABC:F:GK:L:MN:OPRST:UV:WX:Zb:cdf:g:hiklmoprstuvwxz"

  static void
  set_subcommand_option (enum subcommand subcommand)
--- 451,457 ----
     Y  per-block gzip compression */

  #define OPTION_STRING \
!   "-01234567ABC:F:GK:L:MN:OPRST:UV:WX:Zb:cdf:g:hiklmoprstuvwxyz"

  static void
  set_subcommand_option (enum subcommand subcommand)
***************
*** 805,810 ****
--- 808,817 ----
        case 'X':
        exclude_option = 1;
        add_exclude_file (optarg);
+       break;
+
+       case 'y':
+       set_use_compress_program_option ("bzip2");
        break;

        case 'z':
```

5 Använda bzip2 med less

För att automatiskt packa upp bzip2-komprimerade filer, det vill säga för att läsa dem med ”less” utan att första bunzip2-a dem, så kan du göra en lesspipe.sh (man less) så här:

```
#!/bin/sh
# Detta är en preprocessor för 'less'. Den används när denna miljö-
# variabel är angiven:   LESSOPEN="|lesspipe.sh %s"
```

```
  case "$1" in
  *.tar) tar tvvf $1 2>/dev/null ;; # Se innehåll i diverse tar-filer
  *.tgz) tar tzvvf $1 2>/dev/null ;;
# För den icke modifierade versionen av tar:
  *.tar.bz2) bzip2 -cd $1 $1 2>/dev/null | tar tzvvf - ;;
# För den patchade versionen av tar:
# *.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 ;; # Läs komprimerade filer korrekt
  *.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) FILE=`file -L $1` ; # groff src
    FILE=`echo $FILE | cut -d ' ' -f 2`
    if [ "$FILE" = "troff" ]; then
      groff -s -p -t -e -Tascii -mandoc $1
    fi ;;
  *) cat $1 2>/dev/null ;;
#  *) FILE=`file -L $1` ; # Binär? använd i så fall 'strings'.
#    FILE1=`echo $FILE | cut -d ' ' -f 2`
#    FILE2=`echo $FILE | cut -d ' ' -f 3`
#    if [ "$FILE1" = "Linux/i386" -o "$FILE2" = "Linux/i386" \
#         -o "$FILE1" = "ELF" -o "$FILE2" = "ELF" ]; then
#      strings $1
#    fi ;;
  esac
```

6 Använda bzip2 med emacs

6.1 Ändra emacs för alla:

Jag har skrivit följande patch till jka-compr.el, vilket lägger till bzip2 till auto-compression-mode.

Tillkännagivande: Jag har endast testat detta med emacs-20.2, men har ingen anledning att tro att en liknande metod inte skulle fungera med andra versioner.

För att använda den,

1. Gå till katalogen med emacs-20.2/lisp-källkod (var du nu packade upp det).
2. Placera den nedanstående patchen i en fil som heter jka-compr.el.diff (den ska vara ensam i den filen ;).
3. Gör

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

4. Kör igång emacs och tryck in

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

5. Lämna emacs.
6. Flytta din ursprungliga jka-compr.elc till ett säkert ställe för säkerhets skull.
7. Ersätt den med den nya jka-compr.elc.
8. Ha skoj!

```
--- 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")
```

6.2 Ändra emacs för en användare:

Tack för det här går till Ulrik Dickow (htmlurl url="mailto:ukd@kampsax.dk" name="ukd@kampsax.dk">, systemprogrammerare på Kampsax Technology:

För att få till det så att du automatiskt kan använda bzip2 när du inte är systemadministratör, lägg bara in följande i din .emacs-fil.

```
;; Automatisk (de)komprimering på laddade/sparade filer (gzip(1) eller
;; liknande). Vi startar det i av-läge, så att bzip2(1)-stöd kan
```

```
;; läggas till. Koden kastades ihop av Ulrik Dickow för ~./emacs med
;; Emacs 19.34. Borde funka med många äldre och nyare Emacsar också.
;; Ingen garanti dock.
;;
(if (fboundp 'auto-compression-mode) ; Emacs 19.30+
    (auto-compression-mode 0)
  (require 'jka-compr)
  (toggle-auto-compression 0))
;; Lägg nu till bzip2-stöd och slå på auto-komprimering igen.
(add-to-list 'jka-compr-compression-info-list
             ["\\.bz2\\(~\\|\\.~[0-9]+~\\)?\\'"
              "zipping"        "bzip2"         ()
              "unzipping"      "bzip2"         ("-d")
              nil t])
(toggle-auto-compression 1 t)
```

7 Använda bzip2 med wu-ftpd

Tack till Arnaud Launay för denna bandbreddssparare. Det följande ska läggas in i /etc/ftpconversions för att direkt packning och uppackning med bzip2 ska kunna göras. Se till att sökvägarna (som /bin/compress) är korrekta.

```
 :.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
```

8 Använda bzip2 med Netscape under X

tenthumbs@cybernex.net säger:

> Jag kom också på ett sätt att få Linux Netscape att använda bzip2 för Content-Encoding (innehållsavkodning), precis som det använder gzip. Lägg till detta till $HOME/.Xdefaults eller $HOME/.Xresources. Jag använder argumentet -s eftersom jag hellre byter bort en del uppackningshastighet mot RAM-användning. Du kan låta bli att använda den om du vill.

```
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
```

9 Använda bzip2 för att packa om andra packningsformat

Följande perl-program tar filer som komprimerats i andra format (.tar.gz, .tgz, .tar.Z och .Z för denna gång) och packar om dem med bättre komprimering. Perl-koden har alla typer av trevlig dokumentation om vad den gör och hur den gör det den gör.

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

#######################################################
#                                                     #
# Detta program tar packade och gzippade program i    #
# i den aktuella katalogen och gör om dem till bzip2- #
# format. Det klarar av .tgz-ändelsen på ett hyffsat  #
# sätt och skapar en .tar.bz2-fil.                    #
#                                                     #
#######################################################
$counter = 0;
$saved_bytes = 0;
$totals_file = '/tmp/machine_bzip2_total';
$machine_bzip2_total = 0;

@files = <*[Zz]>;
$total = scalar(@files);

foreach (@files) {
    if (/tgz$/) {
        ($new=$_) =~ s/tgz$/tar.bz2/;
    } else {
        ($new=$_) =~ s/\.g?z$/.bz2/i;
    }
    $orig_size = (stat $_)[7];
    ++$counter;
    print "Packar om $_ ($counter/$total)...\n";
    if ((system "gzip -cd $_ |bzip2 >$new") == 0) {
        $new_size = (stat $new)[7];
        $factor = int(100*$new_size/$orig_size+.5);
        $saved_bytes += $orig_size-$new_size;
        print "$new är ungefär $factor% av storleken på $_. :",($factor<100)?')':'(',"\n";
        unlink $_;
    } else {
        print "Arrgghh!  Något hände med $_: $!\n";
    }
}
print "Du har ",
      ($saved_bytes>=0)?"sparat":"förlorat",
      " $saved_bytes byte lagringsutrymme :",
       ($saved_bytes>=0)?")":"(", "\n";

unless (-e '/tmp/machine_bzip2_total') {
    system ('echo "0" >/tmp/machine_bzip2_total');
    system ('chmod', '0666', '/tmp/machine_bzip2_total');
```

```
}


chomp($machine_bzip2_total = `cat $totals_file`);
open TOTAL, ">$totals_file"
     or die "Kan inte öppna systemvid totalitet: $!";
$machine_bzip2_total += $saved_bytes;
print TOTAL $machine_bzip2_total;
close TOTAL;

print "Maskinvid totalitet av ",`cat $totals_file`," byte sparade.\n";
```