Mini Howto Bzip2

David Fetter, *david@fetter.org* ;
Version Française par Arnaud Launay, *asl@launay.org* v2.0, 22 août 1999

Ce document vous expliquera comment utiliser le nouveau programme de compression bzip2. Le document sgml original se trouve *ici* `http://fetter.org/Bzip2-HOWTO/` .

Contents

1 Introduction

Bzip2 est un nouvel algorithme délirant de compression de données. Il produit généralement des fichiers qui ne font que 60 ou 70% de la taille obtenue par compression par gzip.

Ce document vous montrera quelques applications courantes pour bzip2.

Les futures versions de ce document auront des applications de libbzip2, la bibliothèque bzip2 en C que l'auteur de bzip2, *Julian Seward* `mailto:Julian_Seward@muraroa.demon.co.uk` , a sympathiquement écrite. Le manuel de bzip2, qui dispose des informations de bas-niveau, se trouve

ici `http://www.bzip2.org/bzip2/docs/manual_toc.html` .

Les futures versions de ce document pourront également inclure un résumé de la discussion sur l'éventualité (et la méthode) d'utilisation de bzip2 dans le noyau Linux.

1.1 Historique des révisions

1.1.1 v2.00

Changement de la section 5 (Utiliser bzip2 avec less) pour que les fichiers .tar.bz2 puissent être lus. Merci à *Nicola Fabiano* `mailto:mc8644@mclink.it` pour la correction.

Mise à jour de l'utilitaire buzzit.

Mise à jour des informations sur tar.

1.1.2 v1.92

Mise à jour de la section 2.2 (Obtenir des binaires de bzip2), incluant les binaires S.u.S.E.

1.1.3 v1.91

Correction d'une type et clarification de quelques problèmes liés au shell dans la section 4 (section sur l'utilisation de bzip2 avec tar). Merci à Alessandro Rubini.

Mise à jour de l'utilitaire buzzit pour ne pas exploser l'archive bzip2 originale.

Ajout de bgrep, un outil de type zgrep.

1.1.4 v1.9

Clarification du problème avec gcc 2.7.*. Merci à Ulrik Dickow de l'avoir fait remarquer.

Ajout de la manière élégante d'utilisation de tar par Leonard Jean-Marc.

Ajout de la traduction en suédois par Linus Ãkerlund.

Correction de la section wu-ftpd sur la suggestion d'Arnaud Launay.

Déplacement des traductions vers leur propre section.

1.1.5 v1.8

Ajout de buzzit et tar.diff dans le sgml où ils doivent se trouver. Correction de la ponctuation et de la présentation. Merci à Arnaud Launay pour m'aider à corriger ma copie. :-)

Suppression du projet xv pour le moment, dû au manque d'intérêt populaire.

Ajout d'un pense-bête pour les versions futures du document.

1.1.6 v1.7

Ajout de l'utilitaire buzzit. Correction du patch pour gnu tar.

1.1.7 v1.6

Ajout du truc de TenThumbs pour Netscape.

Egalement, changement de lesspipe.sh d'après sa suggestion. Il doit mieux fonctionner maintenant.

1.1.8 v1.5

Ajout de la traduction en français d'Arnaud Launay, ainsi que son fichier pour wu-ftpd.

1.1.9 v1.4

Ajout de la traduction en japonais de Tetsu Isaji.

1.1.10 v1.3

Ajout du .emacs d'Ulrik Dickow pour les 19.30 et plus récents.

(Egalement corrigé le patch jka-compr.el pour emacs d'après sa suggestion. Oops! Bzip2 ne possède pas encore(?) de drapeau "append").

1.1.11 v1.2

Correction du patch pour emacs afin qu'il reconnaisse automagiquement les fichiers .bz2.

1.1.12 v1.1

Ajout du patch pour emacs.

1.1.13 v1.0

1ère version.

2 Obtenir bzip2

La page web de bzip2 se trouve sur le *site britannique* `http://www.bzip2.org/` . Le miroir américain est *ici* `http://www.digistar.com/bzip2/index.html` .

2.1 Le Bzip2-HOWTO dans votre langue

Les francophones peuvent se référer aux traductions d'Arnaud Launay. La version web se trouve *ici* `http://www.freenix.org/unix/linux/HOWTO/mini/Bzip2.html` , et vous pouvez utiliser le ftp *ici* `ftp://ftp.traduc.org/pub/HOWTO/FR/Bzip2.gz` . Arnaud peut être contacté par courrier électronique à *cette adresse* `mailto:asl@launay.org` .

Les nippons peuvent se référer aux traductions de Tetsu Isaji, *ici* `http://jf.gee.kyoto-u.ac.jp/JF/JF.html` . Isaji peut être joint sur *sa page principale* `http://www2s.biglobe.ne.jp/~kaien/` , ou par courrier électronique à *cette adresse* `mailto:isaji@mxu.meshnet.or.jp` .

Les suédois peuvent se référer aux traductions de Linus Ãkerlund

ici `http://user.tninet.se/~uxm165t/linux_doc.html` . Linus peut être contacté par courrier électronique à *cette adresse* `mailto:uxm165t@tninet.se` .

2.2 Obtenir des binaires précompilés de bzip2

Voir la page web.

2.3 Obtenir les sources de bzip2

Elles se trouvent sur les sites officiels (voir 2 (Obtenir Bzip2) pour les emplacements.

2.4 Compiler bzip2 pour votre machine

Si vous avez gcc 2.7.*, changez la ligne indiquant

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

pour avoir

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

C'est-à-dire, remplacez le -O3 par un -O2 et supprimez le -funroll-loops. Vous pouvez également ajouter une option quelconque du type -m* (comme -m486, par exemple) que vous utilisez lorsque vous compilez un noyau.

Eviter le -funroll-loops est le plus important, car la plupart des gcc 2.7 génèreront le mauvais code, et tous les gcc 2.7 génèreront du code plus lent et plus gros. Pour les autres compileurs (lcc, egcs, gcc 2.8.x) les CFLAGS par défaut sont bons.

Après ça, lancez simplement `make` et installez-le d'après le README.

3 Utiliser bzip2 lui-même

RTFM (Read the Fine Manual Page). En clair: lisez le manuel.

4 Utiliser bzip2 avec tar

Sont listés ci-dessous trois méthodes d'utilisation de bzip2 avec tar, soit:

4.1 Le plus facile à mettre en place

Cette méthode ne nécessite pas de configurer quoi que ce soit. Pour décompresser une archive foo.tar.bz2 en bzip2 dans le répertoire courant, tapez

```
/chemin/vers/bzip2 -cd foo.tar.bz2 | tar xf -
```

ou

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

Ceci fonctionne, mais peut être ennuyeux à taper souvent.

4.2 Facile à mettre en oeuvre, relativement simple à utiliser, pas de nécessité d'avoir les privilèges de super-utilisateur

Merci à *Leonard Jean-Marc* `mailto:leonard@sct1.is.belgacom.be` pour ce truc.

Dans votre .bashrc, vous pouvez mettre une ligne comme celle-ci:

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

Dans votre .tcshrc, ou .cshrc, la ligne ressemblera à:

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

4.3 Aussi simple à utiliser, mais nécessite les privilèges de super-utilisateur

Mettez votre tar à jour avec la nouvelle version de GNU tar, qui est aujourd'hui la 1.13.10. On peut le trouver sur *le site ftp de GNU* `ftp://alpha.gnu.org/gnu/tar/` ou tout miroir.

5 Utiliser bzip2 avec less

Pour décompresser les fichiers de bzip2 au vol, c-à-d pour pouvoir utiliser "less" sur eux sans commencer par les bunzip2-er, vous pouvez créer un lesspipe.sh (man less) comme celui-ci:

```
#!/bin/sh
# Ceci est un préprocesseur pour 'less'. Il est utilisé lorsque cette
# variable d'environnement existe:  LESSOPEN="|lesspipe.sh %s"
```

```
  case "$1" in
  *.tar) tar tvvf $1 2>/dev/null ;; # Voir le contenu de fichiers .tar et .tgz
  *.tgz) tar tzvvf $1 2>/dev/null ;;
# Celui-ci fonctionne pour la version non-modifiée de tar:
  *.tar.bz2) bzip2 -cd $1 $1 2>/dev/null | tar tvvf - ;;
# Celui-ci fonctionne avec la version patchée de 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 ;; # Voir correctement les fichiers compressés
  *.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` ; # Vérifier si c'est un binaire, alors -> voir avec '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 Utiliser bzip2 avec emacs

6.1 Changer emacs pour tous

J'ai écrit le patch suivant pour jka-compr.el qui ajoute bzip2 au mode d'auto-compression.

Avertissement: Je l'ai uniquement testé avec emacs-20.2, mais je n'ai pas de raisons de croire qu'une approche similaire ne fonctionnerait pas avec d'autres versions.

Pour l'utiliser,

1. Allez dans le répertoire des sources emacs-20.2/lisp (quel que soit l'endroit où il se trouverait décompacté)

2. Enregistrez le patch ci-dessous dans un fichier nommé jka-compr.el.diff (il doit être seul dans ce fichier ;).

3. Faites

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

4. Lancez emacs, et utilisez

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

5. Quittez emacs.

6. Déplacez votre jka-compr.elc original vers un endroit sûr en cas d'erreurs.

7. Remplacez le par le nouveau jka-compr.elc.

8. Amusez-vous !

```
--- 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 Changer emacs pour une seule personne

Merci à Ulrik Dickow, *ukd@kampsax.dk* `mailto:ukdATkampsax.dk` , ingénieur système à Kampsax Technology, pour celui-ci:

Pour faire en sorte de pouvoir utiliser bzip2 automatiquement lorsque vous n'êtes pas administrateur système, ajoutez simplement ce qui suit à votre fichier .emacs.

```
;; (Dé)compression automatique pour le chargement/sauvegarde de fichiers
;; (gzip(1) et simililaires)
;; Nous le lançons en état stoppé, ainsi le support pour bzip2(1) peut être
;; ajouté.
;; Codé par Ulrik Dickow pour ~/.emacs avec Emacs 19.34.
;; Doit fonctionner avec beaucoup d'anciens et de nouveaux emacs également.
;; Pas de garantie, cependant.
;;
(if (fboundp 'auto-compression-mode) ; Emacs 19.30+
    (auto-compression-mode 0)
  (require 'jka-compr)
  (toggle-auto-compression 0))
;; Ajouter le support bzip2 et autoriser la compression.
(add-to-list 'jka-compr-compression-info-list
             ["\\.bz2\\(~\\|\\.~[0-9]+~\\)?\\'"
              "zipping"        "bzip2"         ()
              "unzipping"      "bzip2"         ("-d")
              nil t])
(toggle-auto-compression 1 t)
```

7 Utiliser bzip2 avec wu-ftpd

Merci à Arnaud Launay pour ce gain de bande passante. Ce qui suit doit aller dans /etc/ftpconversions pour faire de la compression et décompression au vol avec bzip2. Vérifiez que les chemins (comme /bin/compress) sont bons.

```
 :.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 -y -f -
%s:T_REG|T_DIR:O_COMPRESS|O_TAR:TAR+BZIP2
```

8 Utiliser bzip2 avec Netscape sous X

tenthumbs@cybernex.net nous dit:

> J'ai également trouvé un moyen pour obtenir l'utilisation de bzip2 par Netscape/Linux pour le Content-Encoding de la même manière qu'il utilise gzip. Ajoutez ceci à $HOME/.Xdefaults ou $HOME/.Xresources. J'utilise l'option -s car je préfère perdre un peu de vitesse de décompression et conserver de la mémoire. Vous pouvez supprimer cette option si vous le désirez.

```
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 Utiliser bzip2 pour recompresser en remplacement d'autres formats de compression

Ce programme perl prend les fichiers compressés dans d'autres formats (.tar.gz, .tgz. .tar.Z, et .Z pour cette itération) et les recompresse pour un meilleur résultat. Le source perl dispose de tous les types de documentation nécessaires sur ce qu'il fait et comment il fait ce qu'il fait. Cette dernière version prend les fichiers sur la ligne de commande. Sans argument en ligne de commande, il tente de recomprimer tous les fichiers du répertoire courant.

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

#######################################################
#                                                     #
# Ce programme prend les fichiers compressés et       #
# gzip-és dans le répertoire courant et les tranforme #
# en fichiers bzip2. Il supporte proprement           #
# l'extension .tgz, produisant un fichier .tar.bz2.   #
#                                                     #
#######################################################
$counter = 0;
$saved_bytes = 0;
$totals_file = '/tmp/machine_bzip2_total';
$machine_bzip2_total = 0;

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

foreach(@raw) {
    next if /^bzip/;
    next unless /\.(tgz|gz|Z)$/;
    push @files, $_;
}
$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 "Recompacte $_ ($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 is about $factor% of the size of $_. :",($factor<100)?')':'(',"\n";
        unlink $_;
    } else {
        print "Arrgghh! Quelque chose est arrivé à $_: $!\n";
    }
}
print "Vous avez "
    , ($saved_bytes>=0)?"sauvé ":"perdu "
    , abs($saved_bytes)
    , " octets d'espace disque :"
    , ($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 "Ne peut ouvrir le total sur tout le système: $!";
$machine_bzip2_total += $saved_bytes;
print TOTAL $machine_bzip2_total;
close TOTAL;

print "Ceci nous donne un total de ",`cat $totals_file`," octets sauvés sur
tout le système.\n";
```