

UTSにおけるファイルの圧縮法について

総合情報処理センター 内本 佳彦

1. はじめに

一般的にコンピュータを使っていて、個人が所有するファイルの数は、自分で意識しているよりもずっと速いペースで増えていくようである。要らないファイルは消せばよいのだが、滅多に使わないが保存の必要があるファイルなどは残して置きたいものである。

ただし、ファイルを格納するディスク装置の容量には制限があり、UTS（汎用大型計算機で運用しているUNIXシステムのこと）では、利用者登録時に、利用者が使えるディスク容量の制限値を設定しており、むやみやたらにファイルを増やすことができないようになっている。

ここでは、利用者のカレント・ディレクトリ配下のあるディレクトリに存在する数個のファイルを一つのファイルにまとめたり、圧縮して保存したり、圧縮されたファイルからの復元など、ファイルの整理をするために必要なコマンドの使い方を紹介したいと思う。

2. ファイル整理のためのコマンド

(1) tarでアーカイブ・ファイルをつくる。

アーカイブ・ファイル？。これは関連するたくさんのファイルをまとめて1つのファイルにしたものをそう呼んでいる。アーカイブ・ファイルを作る方法の一つはtarコマンドを使うことである。tarは、個々のファイルの所有者、最終更新日付、保護モードなどを含むヘッダー情報とオリジナルのファイルのコピーからなるバイナリ・イメージのアーカイブ・ファイルを作成するものである。

一般にtarコマンドはテープやフロッピィへコピーするときによく使われるが、ここでは、ディスク装置上の通常のファイルへコピーする場合を説明する。

tarコマンドは、一般に

```
tar cvf アーカイブ・ファイル名 コピーするファイル名 . . . .
```

(例: tar cvf exall.tar ex1 ex2 ex3 ex4)

という形式で使う。

tarのfオプションは、作成するアーカイブ・ファイルのファイル名を指定するものであるが、テープやフロッピィにコピーするときは、デバイスファイル名（例

ば、/dev/rst8など)を指定し、通常のファイルにコピーするときは、単なるファイル名を指定すればよい。

また、マニュアルには、vオプションはコピーしたファイルの名前などを画面表示するとき、cオプションはツリー構造をテープにコピーするときを使うとなっているが、通常のファイルに作成するときもこのオプションを使う。アーカイブ・ファイル名にはtarで作ったことがファイル名から推測できるように、.tarという拡張子をつけるのが普通である(例えば、exall.tar)。

```
music% ls -lt (下線は入力部分を示す)
total 1
drwxr-xr-x 2 f9999  g999999 144 Jan 22 19:48 work
```

```
music% cd work・・・ディレクトリworkに移る。
music% ls -lt ・・・ファイル表示
-rw-r--r-- 1 f9999  g999999 401 Jan 22 19:48 ex4
-rw-r--r-- 1 f9999  g999999 401 Jan 22 19:48 ex3
-rw-r--r-- 1 f9999  g999999 401 Jan 22 19:47 ex2
-rw-r--r-- 1 f9999  g999999 401 Jan 22 19:47 ex1
```

```
music% tar cvf exall.tar ex1 ex2 ex3 ex4
a ex1 1 blocks
a ex2 1 blocks
a ex3 1 blocks
a ex4 1 blocks
```

表示されるメッセージの意味は、先頭の1文字がファイルを追加(a:append)している旨、次がファイル名、最後が書き込んだ大きさとなっている。

```
music% ls -lt
-rw-r--r-- 1 f9999  g999999 5632 Jan 22 20:01 exall.tar ・・・ 作成された
-rw-r--r-- 1 f9999  g999999 401 Jan 22 19:48 ex4          アーカイブ
-rw-r--r-- 1 f9999  g999999 401 Jan 22 19:48 ex3          ファイル
-rw-r--r-- 1 f9999  g999999 401 Jan 22 19:47 ex2
-rw-r--r-- 1 f9999  g999999 401 Jan 22 19:47 ex1
```

コピーするファイル名は相対パス名で指定する。絶対パス名にすると、アーカイブから個々のファイルを取り出すときに絶対パス名の位置にファイルが作られてし

まうからである。

(2) アーカイブ・ファイルの内容確認する。

ひとまとめにしたアーカイブ・ファイルの内容を確認するためには、tarコマンドを

```
tar tvf ファイル名
```

(例: tar tvf exall.tar)

という形式で使う。

tオプションはファイル名を表示するために、またvオプションはファイルの許可情報など詳しい情報を表示するために使う。

```
nusic% tar tvf exall.tar
Tar:blocksiz=11
rw-r--r-- 613/346 401 Jan 22 19:47 1992 ex1
rw-r--r-- 613/346 401 Jan 22 19:47 1992 ex2
rw-r--r-- 613/346 401 Jan 22 19:48 1992 ex3
rw-r--r-- 613/346 401 Jan 22 19:48 1992 ex4
```

(3) ファイルを圧縮する。

ファイルを圧縮するコマンドとして、System V系ではpackとunpackが、BSD系ではcompress、uncompressなどが提供されている。最も圧縮率が高く、利用の多いcompress、uncompressのコマンドはBSD系のコマンドであるが、UTS上でも利用できる。ここでは、compress、uncompressコマンドの使い方を説明する。

```
compress ファイル名
```

(例: compress exall.tar)

という形式で使う。圧縮したファイルは、"ファイル名.Z"という名前がついたファイルに書き込まれ、元のファイルは削除される。

compressコマンドでどのくらい圧縮されたかを知るためには、vオプションを使えばよい。

compressコマンドでは、圧縮した内容を書き込むファイル名がすでに存在していると、そのファイルに上書きしてよいかどうか尋ねてくるので、"y"を答えると上書きがおこなわれ、"n"を答えると上書きはされず、コマンドも実行されない。

```
nusic% compress -v exall.tar
exall.tar:Compression: 92.82% -- replace with exall.tar.Z
nusic% ls -lt
-rw-r--r-- 1 f9999  g999999  404 Jan 22 20:01 exall.tar.Z .....圧縮された
-rw-r--r-- 1 f9999  g999999  401 Jan 22 19:48 ex4          ファイル
-rw-r--r-- 1 f9999  g999999  401 Jan 22 19:48 ex3
-rw-r--r-- 1 f9999  g999999  401 Jan 22 19:47 ex2
-rw-r--r-- 1 f9999  g999999  401 Jan 22 19:47 ex1
```

(4) 圧縮したファイルの内容を見る。

圧縮したファイルはテキストファイルではないので、catコマンドで内容を見ることができない。それで圧縮したファイルのまま、ファイルの内容を見るためには、zcatコマンドを使うとよい。また圧縮されたファイルの中の特定のファイルの内容を見るときは、grepコマンドをパイプを使って組み合わせれば簡単にできる。

zcat 圧縮したファイル名 | grep 探したいファイル名

(例: zcat exall.tar.Z | grep ex1)

```
nusic% zcat exall.tar | grep ex1
ex1 644 1145 532 621 5137245663 4564 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaa<ファイルの中身>
.....
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
```

(5) 圧縮したファイルを復元する。

compressで圧縮したファイルは、uncompressコマンドで復元できる。

uncompress ファイル名

(例: uncompress exall.tar.Z)

という形式で使う。

```
nusic% uncompress exall.tar.Z
nusic% ls -lt
total 6
-rw-r--r-- 1 f9999  g999999  5632 Jan 22 20:01 exall.tar
-rw-r--r-- 1 f9999  g999999  401 Jan 22 19:48 ex4
```

```
-rw-r--r-- 1 f9999 g999999 401 Jan 22 19:48 ex3
-rw-r--r-- 1 f9999 g999999 401 Jan 22 19:47 ex2
-rw-r--r-- 1 f9999 g999999 401 Jan 22 19:47 ex1
```

圧縮したファイルを復元するとファイル名の ".Z" が取れたファイル名になっている。復元した内容を書き込むファイル名と同じファイル名が存在すると、上書きしても良いかどうかを尋ねてくるので、"n (またはy以外)" を答える必要がある。

(6) 圧縮したファイルから必要なファイルを取り出す。

tarでまとめて、圧縮したファイルが出来上がると、不要なファイルはrmコマンドで消せば、かなりファイルの整理ができる。その後、必要なときに、圧縮したファイルから、必要なファイルを取り出せばよい。

zcat 圧縮したファイル名 | tar xf - 取り出すファイル名

(例 : zcat exall.tar.Z | tar xf - ex1)

の形式で使う。

tarのオプションxでは、"- "とともに用いると標準入力を意味する。zcatでは、圧縮したファイルの内容が標準出力に出力されるので、その結果をパイプを介してtarに送り、tar xf - ではパイプを介して標準入力から入力されるデータのうち、指定されたファイルだけを取り出せる。この方法でコマンドを使用すれば、1つのファイルを取り出すために、わざわざ圧縮したtarのファイル全体を復元する必要はないので便利である。

```
nusic% ls -lt
-rw-r--r-- 1 f9999 g999999 404 Jan 22 20:17 exall.tar.Z
nusic% zcat exall.tar.Z | tar xf - ex1
nusic% ls -lt
-rw-r--r-- 1 f9999 g999999 404 Jan 22 20:17 exall.tar.Z
-rw-r--r-- 1 f9999 g999999 401 Jan 22 19:47 ex1 .....復元されたファイル
```

3. おわりに

以上、簡単にUTS上でファイルを整理をするために必要なコマンドの使い方について説明してきた。

実際の利用例で比較すると、ディレクトリwork配下にあった4個のファイル容量の合計は1604バイトであったのに対して、アーカイブ・ファイルにして圧縮したフ

ファイル (exall.tar.Z) では、404バイトになり、およそ4分の1になっているのがわかる。

計算機のディスク資源は、無尽蔵ではありません。利用者の皆さん、是非一度、ファイルの整理に取り組まれたらいかがでしょう。塵も積もれば何とかでファイル使用料金も馬鹿になりませんから。

参考文献

- [1] 坂本 文：「たのしいUNIX」，アスキー出版局，1991年2月
- [2] UNIX magazine：「フリーソフトウェアでグレードアップ」，UNIX magazine，1990年8月号