

# Kompresja, archiwizacja plików w Linuksie – podstawy

Umiejętność kompresji/dekompresji archiwów jest niezwykle przydatna. Do czego może przydać się umiejętność archiwizacji? Chociażby do tego, aby zrobić kopię zapasową folderu i zapisać go na jednej z dostępnych chmur. Archiwizacja pozwala nam zachować porządek w naszych zbiorach. Jeśli do tego dochodzi kompresja – możemy powiedzieć że jest wyśmienicie. Nie tylko wszystko sobie poukładamy, ale także oszczędzimy miejsce na dysku. A to kosztuje. Archiwizacja przydaje się także wtedy, kiedy chcesz wysłać komuś kilka plików jako załącznik do emaila. Nie ma innego wyjścia jak ich kompresja do jednego pliku.

Czy jest coś trudnego w archiwizacji plików na Linuksie? Odpowiedź jest oczywista – nie. Podobnie jak w systemie Windows, możemy korzystać z graficznych menedżerów. Wśród najpopularniejszych można wymienić Ark (domyślnie umieszczony w KDE) czy File Roller (dostępny w GNOME). Ale to nie im poświęcimy ten wpis. Zajmiemy się konsolową kompresją/dekompresją za pomocą narzędzia `tar` i innych narzędzi pomocniczych.

## Narzędzie `tar` – archiwizacja nigdy nie była tak prosta!

Do ćwiczeń przygotujemy sobie specjalny katalog. Np.: `kompikownia`. Wewnątrz niego utworzymy trzy pliki tekstowe – `wpis1.txt`, `wpis2.txt` i `wpis3.txt`. Możesz wypełnić je jakimś Lorem Ipsum – jeśli chcesz. Do utworzenia zestawu plików może posłużyć poniższy zestaw poleceń:

```
mkdir kompikownia
touch kompikownia/wpis{1..3}.txt
```

### Archiwizacja plików – zadanie 1

#### Składnia polecenia `tar`

**`tar`** – polecenie, którego będziemy używali najczęściej. Jego nazwa pochodzi od angielskiego Tape Archiver. Program był początkowo używany do archiwizacji na taśmach magnetycznych. Ale te odeszły do lamusa i twórcy `tara` przystosowali go do zapisu archiwów na dowolnych nośnikach.

To, co ujrzysz za moment może wydać się skomplikowane. W istocie takie będzie, jeśli nie zrozumiesz jakie są poszczególne części polecenia `tar` oraz za co one odpowiadają. Podstawowy schemat prezentuje się następująco:

```
tar [-przełączniki] [nazwa archiwum].tar [plik1] [plik2] ... [plikN]
```

Nawiasy kwadratowe w powyższym schemacie oznaczają, że coś w te miejsce musisz wstawić. W miejsce oznaczone jako *przełączniki* podstawiasz „literki” sterujące działaniem programu i

mówiące mu, co ma właściwie zrobić. Najbardziej podstawowe i przydatne poznasz czytając dalszą część artykułu.

Zamiast `[nazwa_archiwum]` wpisujesz nazwę pliku, do którego chcesz zapisać archiwum. Pliki zarchiwizowane za pomocą programu tar domyślnie posiadają rozszerzenie `.tar`.

Jeśli pakujemy archiwum, to ostatnią częścią polecenia będzie lista plików, które chcemy do tego archiwum dodać. Przy rozpakowywaniu nie bierzemy tej części polecenia pod uwagę.

Teoria omówiona. Pora zabrać się za praktykę, która naprawdę nauczy cię posługiwania się tym archiwizatorem

## Zadanie 1.1 – utworzenie archiwum za pomocą narzędzia tar

Wejźmy do folderu kompikownia (odbywa się to za pomocą polecenia `cd`). Na początku postawimy sobie proste zadanie – po prostu utworzymy archiwum za pomocą narzędzia tar, w którym będą znajdowały się trzy utworzone wcześniej pliki tekstowe.

```
tar -cf archiwum.tar wpis1.txt wpis2.txt wpis3.txt
```

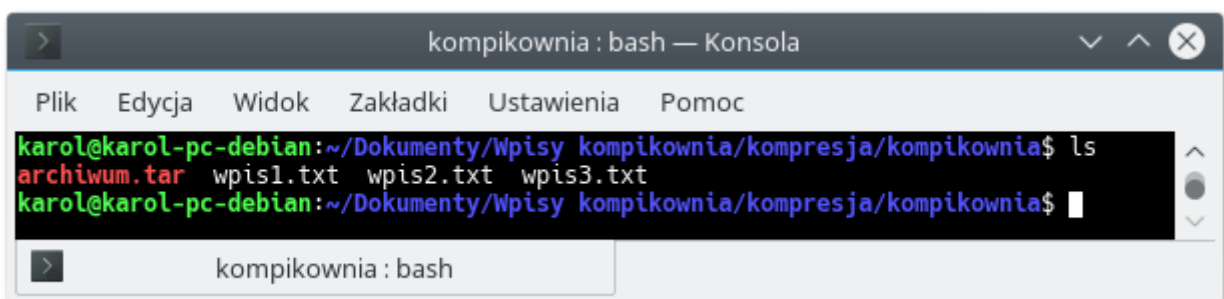
Poznajemy podstawową składnię polecenia w praktyce. Po nazwie programu (`tar`) znajdują się przełączniki. Co one oznaczają i jak łatwo je zapamiętać?

```
c - (od angielskiego słowa create) - tworzy archiwum tar  
f - (od angielskiego słowa file) - mówi, że następane słówko po liście  
przełączników oznacza nazwę archiwum, które modyfikujemy
```

Przełącznik `f` będzie występował zawsze. Umieszczamy go zawsze na końcu złożonego z liter korowodu.

Domyślamy się, co to jest `archiwum.tar` – po prostu nazwa tworzonego przez nas archiwum. W dalszej części po spacji wymieniamy wszystkie pliki, które chcemy dodać.

Archiwum zostało utworzone bez problemów. Wyświetli listę plików w katalogu za pomocą polecenia `ls`. Powinna ona prezentować się tak jak na poniższym screenie.



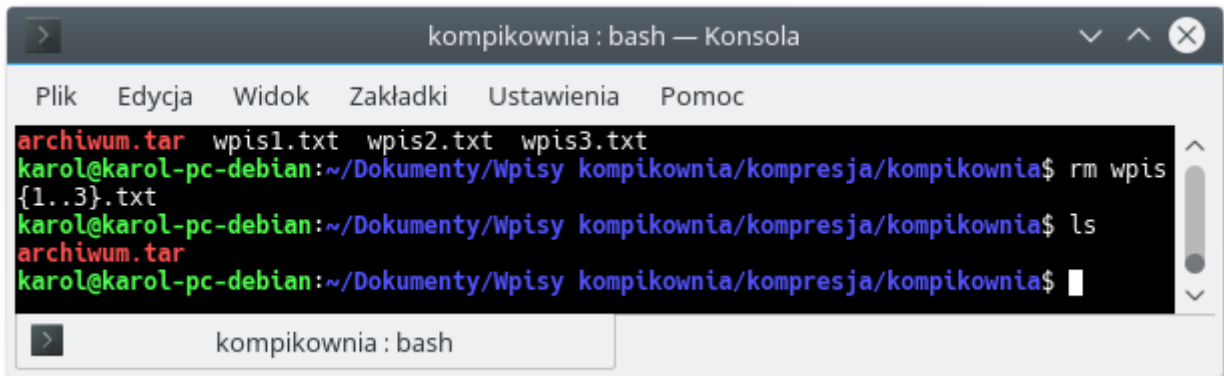
```
kompikownia : bash — Konsola  
Plik  Edycja  Widok  Zakładki  Ustawienia  Pomoc  
karol@karol-pc-debian:~/Dokumenty/Wpisy kompikownia/kompresja/kompikownia$ ls  
archiwum.tar wpis1.txt wpis2.txt wpis3.txt  
karol@karol-pc-debian:~/Dokumenty/Wpisy kompikownia/kompresja/kompikownia$
```

## Zadanie 1.2 – rozpakowanie archiwum za pomocą narzędzia tar

Udało się nam utworzyć archiwum. Świetnie, ale co dalej? Skoro stworzyliśmy paczkę, to musimy umieć także wydobyć z niej zawartość, prawda? Usuńmy wszystkie pliki tekstowe w folderze, zostawiając tylko plik tar. Zrobimy to za pomocą polecenia:

```
rm wpis{1..3}.txt
```

W tym momencie w folderze powinien znajdować się tylko plik o nazwie archiwum.tar. Sprawdź czy tak jest za pomocą polecenia ls. Powinno wyglądać to podobnie jak na screenie poniżej.



```
kompikownia : bash — Konsola
Plik  Edycja  Widok  Zakładki  Ustawienia  Pomoc
archiwum.tar wpis1.txt wpis2.txt wpis3.txt
karol@karol-pc-debian:~/Dokumenty/Wpisy kompikownia/kompresja/kompikownia$ rm wpis
{1..3}.txt
karol@karol-pc-debian:~/Dokumenty/Wpisy kompikownia/kompresja/kompikownia$ ls
archiwum.tar
karol@karol-pc-debian:~/Dokumenty/Wpisy kompikownia/kompresja/kompikownia$
```

Wszystko się zgadza? Wyśmienicie. Przechodzimy do następnego kroku.

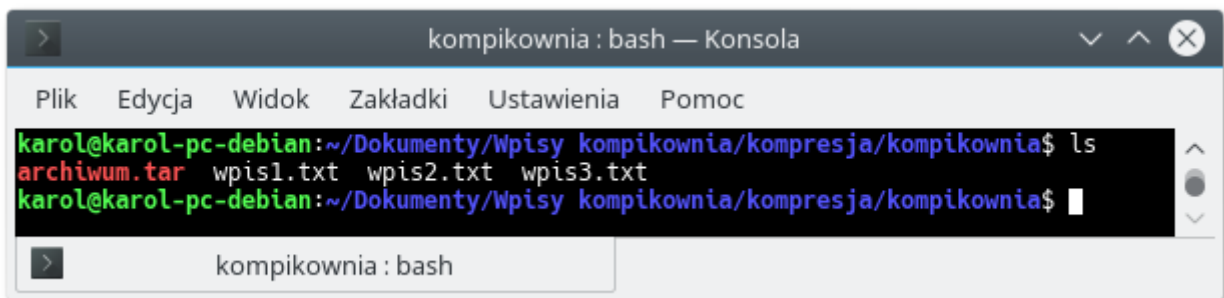
Archiwum możemy rozpakować za pomocą następującego polecenia:

```
tar -xf archiwum.tar
```

Pojawił się nowy, nieznany przełącznik – x. Cóż może oznaczać?

x - (od angielskiego eXtract) - informuje program tar o tym, że chcemy coś rozpakować

Po rozpakowaniu archiwum zawartość folderu powinna prezentować się następująco:



```
kompikownia : bash — Konsola
Plik  Edycja  Widok  Zakładki  Ustawienia  Pomoc
karol@karol-pc-debian:~/Dokumenty/Wpisy kompikownia/kompresja/kompikownia$ ls
archiwum.tar wpis1.txt wpis2.txt wpis3.txt
karol@karol-pc-debian:~/Dokumenty/Wpisy kompikownia/kompresja/kompikownia$
```

### Zadanie 1.3 – dodanie pliku do archiwum. Podgląd archiwum

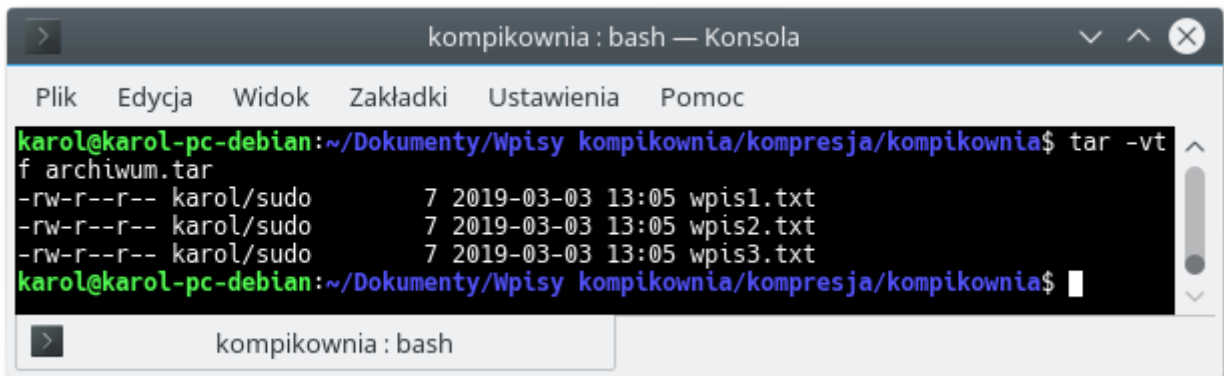
Załóżmy, że mamy ogromny plik archiwalny, ważący kilkaset megabajtów, zawierający np.: zestawienie faktur z ubiegłego miesiąca. Szef wymaga od nas dodania kolejnej faktury. Czas nas goni. Co robimy? Rozpakowanie archiwum, a następnie dodanie do niego dodatkowych plików może nas kosztować dużo cennych chwil. Nie przejmuj się, nie trzeba wykonywać aż tak

karkołomnych czynności. Zostały stworzone odpowiednie przełączniki

Najpierw wyświetlimy aktualną zawartość archiwum. Możemy zrobić to za pomocą polecenia:

```
tar -vtf archiwum.tar
```

Wynik polecenia powinien prezentować się następująco:



```
kompikownia : bash — Konsola
Plik  Edycja  Widok  Zakładki  Ustawienia  Pomoc
karol@karol-pc-debian:~/Dokumenty/Wpisy kompikownia/kompresja/kompikownia$ tar -vtf
f archiwum.tar
-rw-r--r-- karol/sudo      7 2019-03-03 13:05 wpis1.txt
-rw-r--r-- karol/sudo      7 2019-03-03 13:05 wpis2.txt
-rw-r--r-- karol/sudo      7 2019-03-03 13:05 wpis3.txt
karol@karol-pc-debian:~/Dokumenty/Wpisy kompikownia/kompresja/kompikownia$
```

Co widzimy na ekranie? Polecenie wypuściło nam listę plików, które znajdują się wewnątrz archiwum. Zauważ jak szczegółowe informacje otrzymaliśmy. Dowiedzieliśmy się nie tylko jaką nazwę ma każdy z plików, ale także kiedy został utworzony. Informacje na temat uprawnień także zostały zachowane, co jest nieocenioną zaletą archiwizatora tar.

Pojawiły się dwa nowe przełączniki. Co one oznaczają oraz jak je zapamiętać?

```
-v - (od angielskiego Verbose) - wyświetla szczegółowe informacje o plikach, które są przetwarzane
-t - (od angielskiego list) - wyświetla zawartość archiwum
```

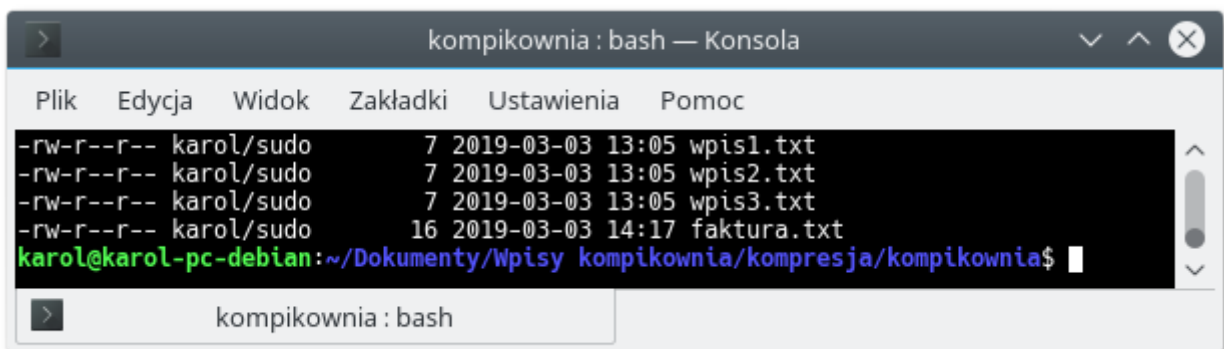
Wiemy, co znajduje się w środku. Dołóżmy teraz dodatkowy plik. Musimy go najpierw wyczarować. Do tego celu posłuży nam touch. Utwórz samodzielnie plik faktura.txt (przykład użycia polecenia touch znajduje się na samym początku artykułu).

Dodajmy plik faktura.txt do archiwum.

```
tar -rf archiwum.tar faktura.txt
```

Wyświetlmy zawartość archiwum. Wiem, że to umiesz. Zrób to sam

Pokażę ci tylko, jak powinien mniej więcej wyglądać rezultat.



```
kompikownia : bash — Konsola
Plik  Edycja  Widok  Zakładki  Ustawienia  Pomoc
-rw-r--r-- karol/sudo      7 2019-03-03 13:05 wpis1.txt
-rw-r--r-- karol/sudo      7 2019-03-03 13:05 wpis2.txt
-rw-r--r-- karol/sudo      7 2019-03-03 13:05 wpis3.txt
-rw-r--r-- karol/sudo     16 2019-03-03 14:17 faktura.txt
karol@karol-pc-debian:~/Dokumenty/Wpisy kompikownia/kompresja/kompikownia$
```

Tradycyjnie przeanalizujemy jakie nowe przełączniki poznaliśmy.

-r - dodaje wymienione na końcu polecenia pliki do archiwum

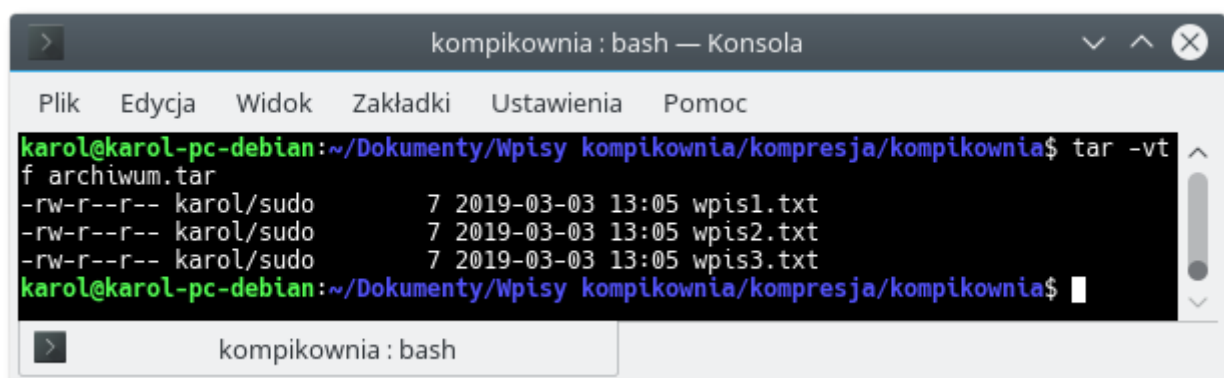
## Zadanie 1.4 – Usuwanie plików z archiwum

Okazało się, że szef dał nam zły pakiet faktur. Niewątpliwie musimy poprawić tę gafę. Jak to zrobimy? Usuniemy stary plik o nazwie faktura.txt z archiwum.

```
tar -f archiwum.tar --delete faktura.txt
```

Polecenie wygląda trochę inaczej niż wszystkie inne do tej pory. Przełącznik `--delete` nie łączy się z tymi tuż po słówku `tar`. Z tego powodu dopisaliśmy go później, po nazwie archiwum. Po przełączniku `delete` pojawia się nazwa pliku/plików, które chcemy usunąć z archiwum.

Aktualna zawartość archiwum prezentuje się następująco:



```
kompikownia : bash — Konsola
Plik  Edycja  Widok  Zakładki  Ustawienia  Pomoc
karol@karol-pc-debian:~/Dokumenty/Wpisy kompikownia/kompresja/kompikownia$ tar -vt
f archiwum.tar
-rw-r--r-- karol/sudo      7 2019-03-03 13:05 wpis1.txt
-rw-r--r-- karol/sudo      7 2019-03-03 13:05 wpis2.txt
-rw-r--r-- karol/sudo      7 2019-03-03 13:05 wpis3.txt
karol@karol-pc-debian:~/Dokumenty/Wpisy kompikownia/kompresja/kompikownia$
```

`--delete` - usuwa z archiwum pliki o nazwach podanych po przełączniku

## Kompresja i dekompresja – gz, bzip2, xz

Poznana do tej pory umiejętność tworzenia i zarządzania archiwami jest ważna, ale niewystarczająca. Marnujemy ogromną ilość miejsca przechowując pliki w nieskompresowanej postaci. Informatycy po to stworzyli i zaimplementowali różnorodne algorytmy kompresji danych abyśmy z nich – jako zwykli użytkownicy – sukcesywnie korzystali.

Najpopularniejszymi rozszerzeniami skompresowanych plików są `gz`, `bzip2` i `xz`. Każda z tych metod może być mniej lub bardziej skuteczna w kompresji poszczególnych rodzajów plików. Nauczmy się korzystać z wszystkich trzech. Jest to bardzo proste pod warunkiem, że zapamiętasz prosty schemat

Pierwszą częścią polecenia służącą do kompresji pliku jest nazwa programu, a zarazem algorytmu, którego chcemy użyć. Do wyboru mamy następujące:

- `gzip` (algorytm deflate, rozszerzenie pliku `.gz`)
- `bzip2` (rozszerzenie pliku `.bzip2`)
- `xz` (algorytm LZMA2, rozszerzenie pliku `.xz`)

Jeśli kompresujemy archiwum, po nazwie kompresora podajemy jedynie nazwę archiwum `tar` (lub innego pliku), które chcemy skompresować.

Gdybyśmy odczuwali potrzebę dekompresji utworzonego archiwum wystarczy, że po nazwie programu używanego do kompresji dodamy przełącznik -d.

## Przykłady

```
gzip archiwum.tar - kompresuje archiwum.tar za pomocą gzip, powstaje plik
archiwum.tar.gz
gzip -d archiwum.tar.gz - dekompresuje archiwum.tar.gz powstaje plik archiwum.tar

bzip2 archiwum.tar - kompresuje archiwum.tar za pomocą bzip2. Powstaje plik
archiwum.tar.bz2
bzip2 -d archiwum.tar.bz2 - dekompresuje plik bzip2. Powstaje archiwum.tar
```

## Kompresja jednym poleceniem – przełączniki tar

Używanie dwóch poleceń do wygenerowania jednego, prostego archiwum wydaje się dosyć denerwujące. Są one intuicyjne – to fakt. Ale czy nie da się tego zrobić jakoś krócej? Programiści pomyśleli o nas – szarych użytkownikach. Poznajmy kolejne przełączniki tar, które pozwalają operować na plikach gzip, bzip2 i xz.

```
-j - użyj algorytmu bzip2 do kompresji/dekompresji
-J - użyj algorytmu związanego z formatem xz do kompresji/dekompresji
-z - użyj formatu gzip do kompresji/dekompresji
```

Jak w praktyce użyć powyższych przełączników? Spójrz na przykłady

### Zadanie 2.1 – utwórz archiwum bz2 składające się z plików wpis1.txt i wpis3.txt

```
tar -cjf archiwum.tar.bz2 wpis1.txt wpis3.txt
```

### Zadanie 2.2 – rozpakuj archiwum bz2

```
tar -xjf archiwum.tar.bz2
```

### Zadanie 2.3 – utwórz archiwum gzip składające się z plików faktura.txt i wpis1.txt

```
tar -czf archiwum.tar.gz faktura.txt wpis1.txt
```

## Słowem zakończenia

Tar posiada oczywiście dużo większą ilość przełączników. Jednakże przydają się one niezwykle rzadko. Z pewnością zauważysz kiedy zajdzie taka potrzeba. Wystarczy wtedy odpalić dokumentację za pomocą polecenia `man tar`. Najważniejsze jest zrozumienie, w jaki sposób tar interpretuje listę argumentów, a także co oznaczają poszczególne literki.

Potrenuj pakowanie i rozpakowywanie różnych rodzajów archiwów. Polecenia w końcu wejdą ci w krew. Kiedy tak się stanie, spakowanie czy rozpakowanie jakiegokolwiek archiwum na Linuksie będzie dla ciebie zwykłą igraszką.