

# System Plików Dokumentacja

---

## Autor

Tomasz Jakubczyk

# Struktura Dysku Wirtualnego

## Opis Nagłówka

Na samym początku pliku występuje nagłówek dysku.

### Nagłówek:

- 4 bajty napis "TJFS" do rozpoznawania mojego pliku.
- 4 bajty (unsigned long int) rozmiar pliku (a zarazem jego koniec) w bajtach.
- 4 bajty adres pierwszego fragmentu dziennika. Jeśli jest zero, to znaczy, że nie ma utworzonego dziennika, a zatem też żadnych plików.

## Opis Dziennika

Na powyższym adresie zaczyna się pierwszy fragment dziennika. Dziennik może składać się z wielu fragmentów i jest swego rodzaju listą jednokierunkową.

### Fragment dziennika:

- 2 bajty długość fragmentu dziennika w bajtach (najprawdopodobniej będzie wyrównana do dword'a) Po informacji o fragmencie dziennika następują informacje o plikach.
- 1 bajt długość nazwy pliku w bajtach.
- n bajtów nazwy pliku (zapewne na końcu uzupełniona zerami -napisy są zwykle null terminated-, tak aby uzyskać globalne dopełnienie do dword'a).
- 4 bajty rozmiar pliku w bajtach.
- 4 bajty adres pierwszego fragmentu pliku. Pliki mogą być podzielone na fragmenty tworzące listę jednokierunkową.
- Teraz mogą wystąpić informacje o innych plikach w takim samym formacie, lub przynajmniej 4 następne bajty mają wartość 0.
- 4 bajty wskazujące na ewentualny następny fragment dziennika lub 0 znaczące, że to ostatni fragment dziennika. Ten dword najlepiej jeśli będzie wyrównany globalnie do dworda w prawo.

## Opis Plików

Pod adresem z dziennika znajduje się początek listy fragmentów zawierających plik.

### Fragmenty plików:

- 4 bajty na rozmiar danego fragmentu pliku. Z uwzględnieniem wyrównania ostatniego dword'a
- bajty pliku

- 4 bajty adres następnego fragmentu pliku, lub 0 jeśli to ostatni fragment.

## Znaczenia Rozmiarów

Bajty pełniące funkcje dworda trzymającego rozmiar lub adres dobrze, żeby miały wyrównanie naturalne do 4, bo wtedy łatwiej jest oglądać dysk w hex edytorze, lub wykonywać obliczenia oparte o system dwujkowy. Jeśli odnoszę się do rozmairu fragmentu, to mam na mysil cały rozmiar fragmentu, tak, żeby bez dodatkowych operacji wyliczyć koniec fragmentu. Dword rozmiaru pliku zawiera w sobie tylko i wyłącznie rozmiar pliku, taki, jak odczytany, przy jego kopiowaniu/otwieraniu.

## Szacowanie Rozmiarów Fragmentów Dziennika

Niech najmniejszy dysk będzie miał 100kB i będzie zapełniony plikami po 4kB. wychodzi, że nie zmieści się więcej niż 25 plików -1 na trzymanie nazw, rozmiarów i adresów. To przy założeniu, że nazwa ma średnio koło 10 znaków.  $24\text{pliki} * (19\text{B na nazwę, długości i adres}) + 18 \text{ na nagówek z jednym fragmentem dziennika} = 474\text{B}$  Czyli rozmiar bloku przyjmijmy 476B razem z wyrównaniem. Od tego momentu ze wzrostem rozmiaru dysku rozmiar fragmentu powinien wzrastać proporcjonalnie, aż do osiągnięcia maksymalenj długości pojedynczego fragmentu. Czyli średnio przy 4,6MB będzie potrzebny następny blok.  $y - \text{rozmiar fragmentu dziennika}$   $x - \text{rozmiar pliku}$   $y := \text{ceil}(0.00476 * x) + \text{ceil}(0.00476 * x) \% 4$  jeśli  $y > 65532$  to  $y := 65532$

## Ułożenie Plików

Jeśli będzie to możliwe, to pliki będą zapisywane w jednym fragmenci, jeśli nie, to będą one podzielone na największe możliwe pasujące w wolne miejsca fragmenty.

## Dlaczego tak?

Jak powszechnie wiadomo jeśli elementy nie są do siebie idealnie wpasowane, to zajmują więcej miejsca (fragmentacja), jeśli mamy wiele elementów przynajmniej dwóch różnych rozmiarów, to zajmą one przestrzeń lepiej niż taka sama obętość identycznych elementów. To trochę tak jak chaotyczne pakowanie rzeczy do plecaka. Jeśli skasujemy jakiś plik i na jego miejsce wsadzimy inny, to powstanie przerwa, którą trudno by było wypełnić jeszcze innym plikiem, ale przychodzi nam z pomocą mniejszy fragment danych dziennika, które z pewnością tam się zmieszczą. :-)

W taki rozwiązanie jest też chyba dość ciekawe.