

## Optymalizacja wydajności dysków pendrive. *Cluster alignment.*

Na wielu forach internetowych można spotkać się z pytaniami o przyczynę małej wydajności dysków przenośnych pendrive. Ludzie skarżą się, że pendrive nie osiąga prędkości podanej przez producenta i nie mówię to o dyskach noname, ale o dyskach uznanych producentów. Jedną z możliwych przyczyn takiego stanu rzeczy jest niedopasowanie logicznego podziału dysku, czyli partycji i sektorów do podziału fizycznego, który wynika z technologii pamięci błyskowych (flash EEPROM).



Niedopasowanie takie może powodować nawet dwukrotne spowolnienie dysku głównie podczas operacji zapisu. Niestety najczęstszą przyczyną spowolnienia pendrivów jest ingerencja użytkownika np. poprzez ponowne utworzenie partycji. Ale do rzeczy. Pamięci flash są specyficznym rodzajem nieulotnych pamięci EEPROM, w których operacje odczytu i zapisu przeprowadza się na bloku danych a nie na pojedynczych komórkach pamięci. Tak samo wygląda sprawa z logicznym podziałem dysku. Najmniejszą jednostką logiczną jest sektor, który też jest zawsze odczytywany w całości. W momencie, kiedy sektory i bloki pamięci flash nie są tym samym obszarem do odczytania jednego sektora potrzeba odczytać większą liczbę bloków pamięci flash. Przy odczycie nie powoduje to tragicznych opóźnień wg informacji znalezionej w Internecie może to powodować spowolnienie rzędu poniżej 15%. Większym problemem staje się zapis na dysku, w którym występuje niedopasowanie sektorów i bloków flash. Zapis na pamięciach typu flash może odbywać się tylko w przypadku czystych bloków pamięci. W momencie, kiedy blok zawiera już jakieś informacje trzeba go najpierw skasować. To powoduje, że zapis na dyskach pendrive z niedopasowaniem może być spowolniony nawet o 100%.

### Wynik optymalizacji

prędkość	przed	po	wzrost
odczytu	27,3	27,3	0 %
zapisu	15,6	19,4	25 %

Dyski uznanych producentów sprzedawane są przeważnie z partycją umieszczoną optymalnie w stosunku do bloków pamięci, a formatowanie przeprowadzone jest w ten sposób, że dopasowuje wielkość sektora do

bloku. W takim przypadku przesunięcie partycji zawsze doprowadzi do pogorszenia wydajności. Są jednak przypadki, kiedy chcemy zmienić partycjonowanie dysku pendrive. Przykładem może być mój dysk, na którym producent SanDisk umieścił dodatkową partycję z oprogramowaniem U3 do uruchamiania oprogramowania portable z którego nie korzystam. Partycja zajmowała 150MB, niby niewiele, ale chciałem odzyskać tę przestrzeń i nie doprowadzić do pogorszenia wydajności.

Zasada postępowanie w takich przypadkach jest następująca:

- odnalezienie wielkości bloków na pamięci flash;
- odnalezienie przesunięcia bloków w stosunku do sektorów;
- przesunięcie początku partycji w taki sposób, aby zniwelować przesunięcie;

- formatowanie partycji w taki sposób, aby dopasować wielkość sektorów do wielkości bloków.

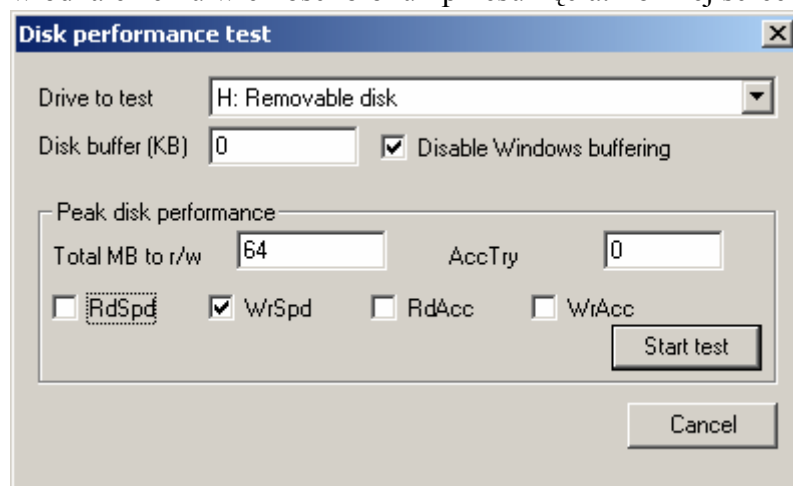
Oprogramowanie potrzebne do wykonania tych operacji:

- Dowolny program do testowania prędkości dysku flash, proponuję CrystalDisk, można go pobrać np. ze strony <http://www.programosy.pl>.

- Program dfboff64 do badania wielkości bloków i przesunięcia, do pobrania ze strony <http://vlo.nm.ru/tmp/dfboff64.rar>. Jeżeli przestanie być dostępny umieszczę go na mojej stronie.

- Dowolny program do swobodnej edycji partycji proponuję Ranish Partition Manager, można go pobrać ze [strony oficjalnej](#), lub co wygodniejsze pobrać (mało legalną) płytę CD z oprogramowaniem narzędziowym Hiren's BOOT. Płyta jest bootowalna i zawiera m.in. Ranish Partition Manager.

Sprawdzenie dysku rozpoczynamy od wykonania testu programem dfboff64. Wynik pomoże w odnalezieniu wielkości bloku i przesunięcia. Poniżej screen z ustawień programu.

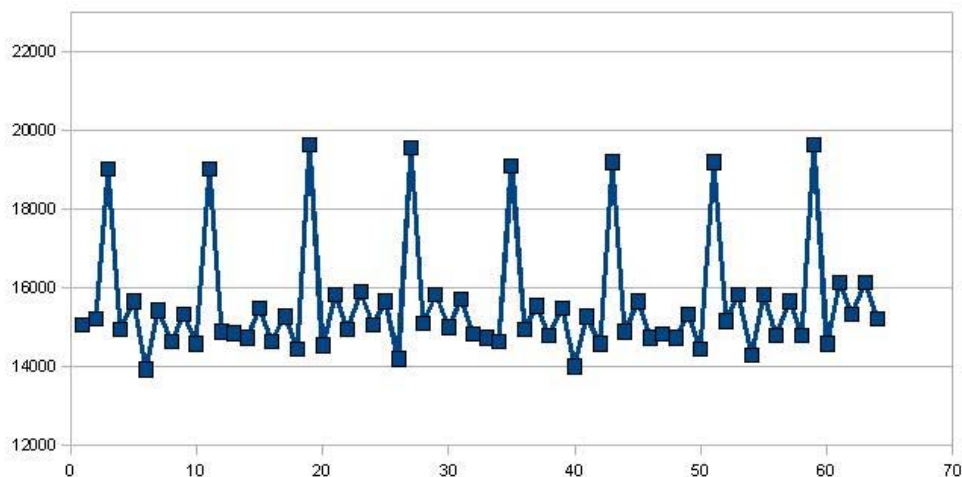


W wyniku działania programu otrzymujemy plik tekstowy benchlog.txt z prędkościami zapisu dla kolejnych obszarów partycji, podzielonych po 512B. Zawartość pliku wygląda następująco.

```
12 Sustained write : 14590 KB/s (+9/4608)
13 Sustained write : 19017 KB/s (+10/5120)
14 Sustained write : 14894 KB/s (+11/5632)
15 Sustained write : 14844 KB/s (+12/6144)
16 Sustained write : 14742 KB/s (+13/6656)
17 Sustained write : 15490 KB/s (+14/7168)
18 Sustained write : 14638 KB/s (+15/7680)
19 Sustained write : 15266 KB/s (+16/8192)
20 Sustained write : 14441 KB/s (+17/8704)
21 Sustained write : 19628 KB/s (+18/9216)
22 Sustained write : 14536 KB/s (+19/9728)
23 Sustained write : 15836 KB/s (+20/10240)
24 Sustained write : 14947 KB/s (+21/10752)
25 Sustained write : 15897 KB/s (+22/11264)
26 Sustained write : 15052 KB/s (+23/11776)
27 Sustained write : 15661 KB/s (+24/12288)
```

Aby łatwiej było analizować dane z testu proponuję zaimportować je do arkusza kalkulacyjnego np. programu Calc z pakietu Open Office i wykonać wykres z liczb

reprezentujących prędkości. Jest to kolumna pomiędzy znakiem dwukropka a napisem KB/s. Wykres prezentuje się następująco.



Szczyty na wykresie pokazują początki fizycznych bloków pamięci flash. Jak widać na osi odciętych jest to 3, 11, 19, 27 ..., czyli okres szczytów wynosi 8. Każdy zapis wykonany przez program to 512B wynika z tego, że fizyczne bloki na pamięci flash mają 4KB ( $512B \times 8 = 4096B$ ). Widać też, że pierwszy początek bloku nie pokrywa się z początkiem partycji. Takie niedopasowanie ma wpływ na spadek wydajności dysku. W tym momencie test prędkości wykonany programem Crystal Disk Mark prezentuje się następująco.

```
Sequential Read : 27.292 MB/s
Sequential Write : 15.636 MB/s
Random Read 512KB : 27.214 MB/s
Random Write 512KB : 2.591 MB/s
Random Read 4KB : 5.451 MB/s
Random Write 4KB : 0.025 MB/s
```

Prędkość odczytu nie odbiega od deklaracji producenta a nawet ją przewyższa o 2MB/s, jednak prędkość zapisu według producenta powinna wynosić 18MB/s, aby osiągnąć taką prędkość trzeba przesunąć początek partycji w celu dopasowania jej do początku pierwszego bloku pamięci flash.

Kilka słów na temat formatowania partycji. Partycja sformatowana jest w systemie FAT32 z wielkością sektora 4KB, czyli formatowanie jest pod tym względem jest optymalne, bo taka jest też wielkość fizycznego bloku pamięci flash. System FAT w odróżnieniu od np. systemu NTFS przetrzymuje meta dane w początkowym obszarze partycji. Przez co początek zapisywanych danych nie pokrywa się z początkiem partycji, jest przesunięty o blok meta danych systemu FAT. Oznacza to, że dysk po sformatowaniu innym programem, który odmiennie utworzy blok meta danych, może wykazywać inne przesunięcie pierwszego bloku. Teoretycznie możliwe jest takie skonstruowanie bloku meta danych, aby przesunięcie partycji nie było konieczne. Niestety nie znalazłem softu, który by pozwalał na taką ingerencję w strukturę FAT, a ręczna edycja danych na dysku wydaje mi się za skomplikowana.

Z powyższego wynika, że przed przystąpieniem do przesuwania partycji warto przynajmniej spróbować sformatować dysk i wykonać test powtórnie. Może okazać się, że przesunięcie partycji nie będzie konieczne. W opisywanym przypadku formatowanie pod Windowsami poleceniem `format [litera dysku]: /FS:FAT32 /A:4096` nie przyniosło rezultatu.

Przesuwanie partycji wykonać można programem Ranish Partition Manager. Korzystania z programu nie będę opisywał, wiele opisów można znaleźć z internetu. Przy operowaniu na partycjach zalecam szczególną rozważę. Poniżej screen z partycji, która była testowana.

```

Ranish Partition Manager          Version 2.44 (beta) by Muthu  June 09, 2002
Hard Disk 2  7,815 Mbytes [15,879 cylinders x 16 heads x 63 sectors ]
Using LBA

```

#	Type	Row	File System Type	Starting Cyl	Starting Head	Starting Sect	Ending Cyl	Ending Head	Ending Sect	Partition Size [KB]
0	MBR		Master Boot Record	0	0	1	0	0	1	0
1	Pri		Unused	0	0	2	0	0	63	31
2	Pri	1	Windows FAT-32 LBA	0	1	1 15,878	15	63		8,002,984
3	Pri		Unused	15,879	0	1 15,879	12	0		382
4			Unused	0	0	0	0	0	0	0
5			Unused	0	0	0	0	0	0	0
6			Unused	0	0	0	0	0	0	0
7			Unused	0	0	0	0	0	0	0

```

B - Boot flag on/off  INS - select file system  DEL - clear record
- MBR
# Partition  Size  Volume label: NO NAME  Starting:  63  20%
1 Unused    0    System id: MSWIN4.1   Drive num: 129  0M
2 Unused    0    File system: FAT32    Minimum size: 160,319  78M
3 Unused    0    Cluster Size: 4k     Partition size: 16,005,969  7,815M
4 Unused    0    FAT Size: 7,815k    Maximum size: 16,005,969  7,815M
F1 Help      F2 Save = F3 Undo = F4 Mode = F5 Disk  ESC Quit

```

Partycję należy przesunąć zgodnie z tym, co można odczytać z wykresu o 3 sektory (3 x 512B), tak, aby początek fizycznego bloku pamięci wypadł na początku regionu danych partycji FAT. Zgodnie z zasadą, że partycja powinna zaczynać się od pierwszego sektora przesuwać ją należy o cylinder lub o głowicę (nie sprawdzałem czy w przypadku pamięci flash ma to znaczenie). Partycję można przesunąć też o wielokrotność klastra 4kB (8 x sektor 512B) plus 3 sektory, czyli przesunięcia, które spełnią oczekiwaną rolę to np. 3, 11, 19, 27, 35 itd. W adresacji Cylinder / Głowica / Sektor, która widoczna jest na screenie widać, że przesunięcie o jedną głowicę daje przesunięcie o 63 sektory. Przesuwając partycję o kolejne głowice otrzymamy przesunięcia o 126, 189, 252 itd. Przesunięcie o 5 głowic daje nam przesunięcie o 315 sektorów (5 x 63 = 315), które odpowiada też planowanemu przesunięciu o wielokrotność klastra plus trzy (39 x 8 + 3 = 315). Poniżej screen z partycją przesuniętą o 5 głowic, czyli o 315 sektorów.

```

Ranish Partition Manager          Version 2.44 (beta) by Muthu  June 09, 2002
Hard Disk 2  7,815 Mbytes [15,879 cylinders x 16 heads x 63 sectors ]
Using LBA

```

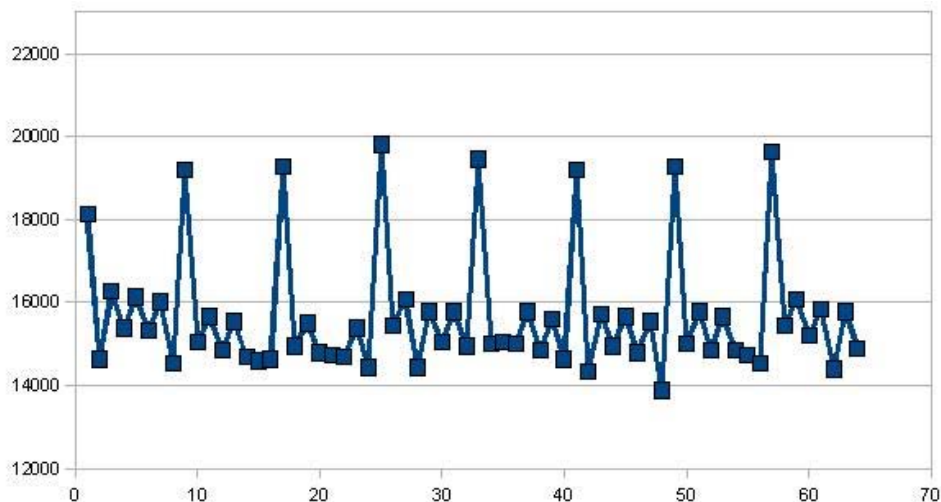
#	Type	Row	File System Type	Starting Cyl	Starting Head	Starting Sect	Ending Cyl	Ending Head	Ending Sect	Partition Size [KB]
0	MBR		Master Boot Record	0	0	1	0	0	1	0
1	Pri		Unused	0	0	2	0	4	63	157
2	Pri	1	Windows FAT-32 LBA	0	5	1 15,878	4	63		8,002,512
3	Pri		Unused	15,878	5	1 15,879	12	0		728
4			Unused	0	0	0	0	0	0	0
5			Unused	0	0	0	0	0	0	0
6			Unused	0	0	0	0	0	0	0
7			Unused	0	0	0	0	0	0	0

```

B - Boot flag on/off  INS - select file system  DEL - clear record
- MBR
# Partition  Size  Volume label: NO NAME  Starting:  315  34%
1 FAT-32    7,814  System id: MSWIN4.1   Drive num: 129  0M
2 Unused    0    File system: FAT32    Minimum size: 31,293  15M
3 Unused    0    Cluster Size: 4k     Partition size: 16,005,024  7,814M
4 Unused    0    FAT Size: 7,815k    Maximum size: 16,005,024  7,814M
F1 Help      F2 Save = F3 Undo = F4 Mode = F5 Disk  ESC Quit

```

Tak utworzoną partycję formatujemy też przy pomocy programu Ranish Partition Manager. Następnie wykonujemy test programem dffoff64 i wyniki analizujemy po przedstawieniu ich na wykresie.

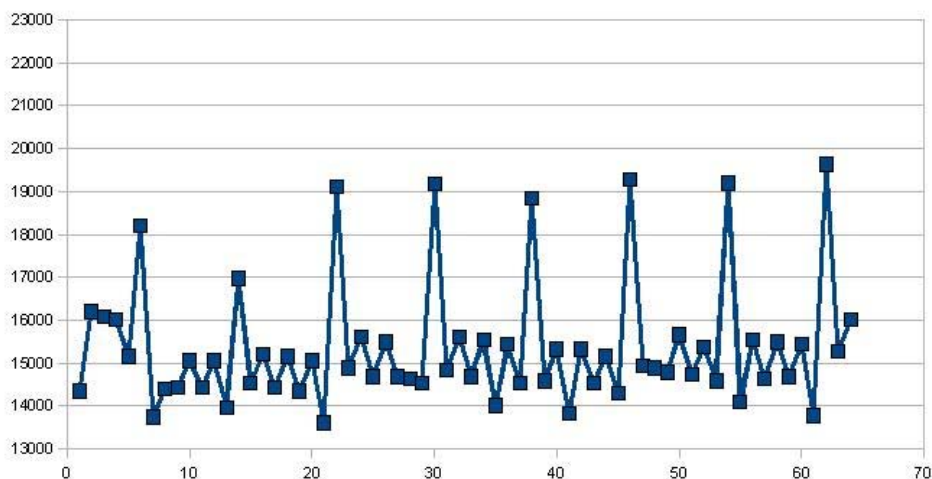


Jak widać partycja dopasowana jest idealnie i początek pierwszego bloku pamięci flash pokrywa się z początkiem regionu danych na partycji FAT. W tym momencie test prędkości wykonany programem Crystal Disk Mark prezentuje się następująco.

```
Sequential Read : 27.317 MB/s
Sequential Write : 19.452 MB/s
Random Read 512KB : 27.271 MB/s
Random Write 512KB : 2.530 MB/s
Random Read 4KB : 5.484 MB/s
Random Write 4KB : 0.026 MB/s
```

Nastąpił wzrost wydajności w przypadku zapisu z 15,6MB/s do 19,4MB/s czyli o prawie 25%. W przypadku zapisu nie widać zmiany prędkości. Po dopasowaniu partycji i formatowania dysk uzyskał a nawet nieznacznie przewyższył prędkości deklarowane przez producenta.

Dla testów po wykonaniu optymalizacji wykonałem formatowanie dysku w systemie Windows. Okazało się, że dysk przestał mieć optymalne ustawienie partycji. Oznacza to, że należałoby przeprowadzić optymalizację przesuwania partycji dla konkretnego formatowania. Poniżej wykres po formatowaniu w Windowsach.



Przesunięcie wynosi 6 sektorów i proces optymalizacji należało by przeprowadzać od początku, ale formatując dysk za każdym razem pod Windowsami a nie w programie Ranish Partition Manager. Zrezygnowałem z tego i przeprowadziłem ponowne formatowanie w Ranish Partition Manager pamiętając żeby nie formatować dysku pod Windowsami.

Opis przeprowadzonej operacji dopasowania nie wyczerpuje tematu. Podobne dopasowanie można przeprowadzić dla kart pamięci opartych na pamięciach flash stosowanych w aparatach cyfrowych, nawigacjach i innych urządzeniach przenośnych. Ciekawym zagadnieniem jest też wpływ technologii pamięci flash na optymalizację wydajności. Aktualnie stosuje się dwie technologie tych pamięci SLC i MLC. W pamięciach SLC jedna komórka pamięci może przyjmować dwa stany napięć. W technologii MLC komórka może przyjmować wiele stanów napięć, co umożliwia zapisanie więcej niż jednego bitu informacji w pojedynczej komórce. Rozważać też można wpływ systemów plików na wydajność dysków pendrive, biorąc oczywiście pod uwagę ich dopasowanie do właściwości fizycznych nośnika. Być może w wolnej chwili przeprowadzę parę testów dotyczących wydajności systemu plików na dyskach USB. Zapraszam do zagłędania na stronę [www.olek.waw.pl](http://www.olek.waw.pl)