

Informatyka 2

Politechnika Białostocka - Wydział Elektryczny
Elektrotechnika, semestr III, studia stacjonarne I stopnia
Rok akademicki 2018/2019

Wykład nr 11 (11.12.2018)

dr inż. Jarosław Forenc

Plan wykładu nr 11

- Systemy plików
 - FAT (FAT12, FAT16, FAT32, exFAT)
 - NTFS

System plików FAT (File Allocation Table)

- opracowany na przełomie lat 70. i 80. dla systemu MS-DOS
- występuje w czterech wersjach: FAT12, FAT16, FAT32 i exFAT (FAT64)
- numer występujący po słowie FAT oznacza liczbę bitów przeznaczonych do kodowania (numeracji) **jednostek alokacji pliku** (JAP), tzw. **klastrów** (ang. cluster) w tablicy alokacji plików
 - 12 bitów w systemie FAT12
 - 16 bitów w systemie FAT16
 - 32 bity w systemie FAT32 (praktycznie 28)
 - 64 bity w systemie exFAT (FAT64)
- ogólna struktura dysku logicznego / dyskietki w systemie FAT:

Rekord ładujący + sektory zarezerwowane	Tablica rozmieszczenia plików - FAT	Kopia FAT	Katalog główny (FAT12 i FAT16)	Miejsce na pliki i podkatalogi
---	-------------------------------------	-----------	--------------------------------	--------------------------------

FAT12

- system plików FAT12 przeznaczony jest dla nośników o małej pojemności
- obsługuje $2^{12} = 4096$ jednostek alokacji, max. rozmiar partycji to 16 MB
- **rekord ładujący** zajmuje pierwszy sektor dyskietki lub dysku logicznego

Rekord ładujący + sektory zarezerwowane	Tablica rozmieszczenia plików - FAT	Kopia FAT	Katalog główny (FAT12 i FAT16)	Miejsce na pliki i podkatalogi
---	-------------------------------------	-----------	--------------------------------	--------------------------------

- rekord ładujący zawiera następujące dane:
 - instrukcja skoku do początku programu ładującego (3 bajty)
 - nazwa wersji systemu operacyjnego (8 bajtów)
 - struktura BPB (ang. BIOS Parametr Block) - blok parametrów BIOS (25 bajtów)
 - rozszerzony BPB (ang. Extended BPB, 26 bajtów)
 - wykonywalny kod startowy uruchamiający system operacyjny (448 bajtów)
 - znacznik końca sektora - 55AAH (2 bajty)

FAT12

- tablica rozmieszczenia plików FAT tworzy swego rodzaju „mapę” plików zapisanych na dysku
- za tablicą FAT znajduje się jej kopia, która nie jest wykorzystywana



- za kopią tablicy FAT znajduje się katalog główny zajmujący określoną dla danego typu dysku liczbę sektorów



FAT12

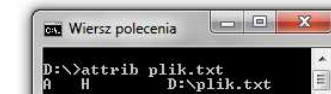
- katalog główny zawiera 32-bajtowe pola mogące opisywać pliki, podkatalogi lub etykietę dysku

Zawartość pola:

Bajty	Rozmiar	Zawartość
00H-07H	8	Nazwa pliku w kodach ASCII
08H-0AH	3	Rozszerzenie nazwy pliku
0BH	1	Atrybuty pliku
0CH-15H	10	Zarezerwowane
16H-17H	2	Czas utworzenia lub aktualizacji pliku
18H-19H	2	Data utworzenia lub aktualizacji pliku
1AH-1BH	2	Numer pierwszej JAP
1CH-1DH	2	Mniej znaczące słowo rozmiaru pliku
1EH-1FH	2	Bardziej znaczące słowo rozmiaru pliku

Atrybuty pliku:

Bit	Znaczenie
0	Plik tylko do odczytu (read only)
1	Plik ukryty (hidden)
2	Plik systemowy (system)
3	Etykieta dysku (volume label)
4	Podkatalog
5	Plik archiwalny (archive)
6,7	Nie wykorzystywane

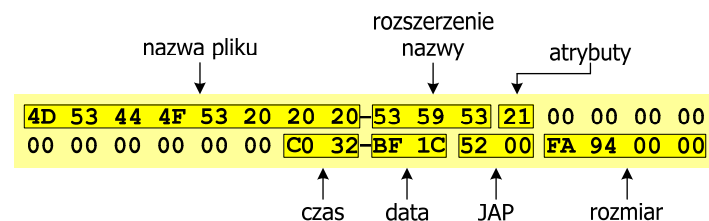


FAT12

- przykładowa zawartość katalogu głównego:

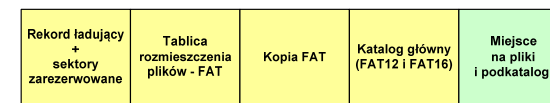
```

0000 49 4F 20 20 20 20 20-53 59 53 21 00 00 00 00 IO SYS!...
0010 00 00 00 00 00 00 C0 32-BF 1C 02 00 46 9F 00 00 .....2...F...
0020 4D 53 44 4F 53 20 20 20-53 59 53 21 00 00 00 00 MSDOS SYS!...
0030 00 00 00 00 00 00 C0 32-BF 1C 52 00 FA 94 00 00 .....2..R....
0040 43 4F 4D 4D 41 4E 44 20-43 4F 4D 20 00 00 00 00 COMMAND COM...
0050 00 00 00 00 00 00 C0 32-BF 1C 9D 00 75 D5 00 00 .....2...u...
0060 41 54 54 52 49 42 20 20-45 58 45 20 00 00 00 00 ATTRIB EXE...
0070 00 00 00 00 00 00 C0 32-BF 1C 08 01 C8 2B 00 00 .....2.....+
    
```



FAT12

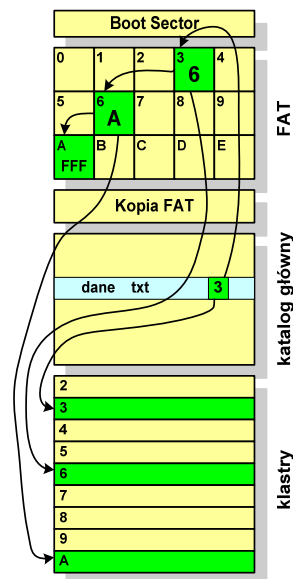
- pozostałą część dysku zajmuje miejsce na pliki i podkatalogi



- podkatalogi nie są ograniczone co do wielkości, zapisywane są na dysku w sposób identyczny jak pliki użytkowe i także zawierają 32-bajtowe pola

FAT12 - położenie pliku na dysku

- w katalogu, w 32-bajtowym polu każdego pliku wpisany jest początkowy numer JAP
- numer ten określa logiczny numer sektora, w którym znajduje się początek pliku
- ten sam numer JAP jest jednocześnie indeksem do miejsca w tablicy FAT, w którym wpisany jest numer kolejnej JAP
- numer wpisany we wskazanym miejscu tablicy rozmieszczenia plików wskazuje pierwszy sektor następnej części pliku i równocześnie położenie w tablicy FAT numeru następnej JAP
- w ten sposób tworzy się łańcuch, określający położenie całego pliku
- jeśli numer JAP składa się z samych FFF, to oznacza to koniec pliku



FAT12 - struktura dyskietki (1,44 MB)



- całkowita liczba sektorów na dyskietce: 2880
- liczba sektorów systemowych: 33

Rekord ładujący + sektory zarezerwowane	Tablica rozmieszczenia plików - FAT	Kopia FAT	Katalog główny (FAT12 i FAT16)	Miejsce na pliki i podkatalogi
1 sektor	9 sektorów	9 sektorów	14 sektorów	2847 sektorów

- maksymalna liczba plików w katalogu głównym: 224
- liczba sektorów na pliki i podkatalogi: $2880 - 33 = 2847$
- dostępne miejsce na pliki i podkatalogi: $2847 \times 512 = 1\,457\,664$ bajty

FAT16

- po raz pierwszy pojawił się w systemie MS-DOS 3.3
- ogólna struktura dyskietki / dysku logicznego w systemie FAT16 jest taka sama jak w przypadku FAT12

Rekord ładujący + sektory zarezerwowane	Tablica rozmieszczenia plików - FAT	Kopia FAT	Katalog główny (FAT12 i FAT16)	Miejsce na pliki i podkatalogi

- maksymalna liczba jednostek alokacji ograniczona jest do 2^{16} czyli 65536
- w systemach **DOS** i **Windows 95** maksymalny rozmiar JAP to 2^{15} bajtów czyli 32 kB, stąd maksymalny rozmiar dysku logicznego w tych systemach to 65536×32 kB czyli **ok. 2 GB**
- w systemie **Windows 2000** górna granica rozmiaru JAP wynosi 2^{16} bajtów (64 kB), czyli rozmiar dysku logicznego zwiększa się do **4 GB**.

FAT32

- po raz pierwszy wprowadzony w systemie Windows 95 OSR2
- ogólna struktura systemu FAT32 jest taka sama jak w FAT12/FAT16 - nie ma tylko miejsca przeznaczonego na katalog główny

Rekord ładujący + sektory zarezerwowane	Tablica rozmieszczenia plików - FAT	Kopia FAT	Miejsce na pliki i katalogi

- do adresowania JAP stosuje się, obcięty o 4 najstarsze bity, adres 32-bitowy i dlatego dysk z FAT32 może zawierać maksymalnie 2^{28} JAP
- dla JAP od 4 kB do 32 kB, teoretycznie dysk może mieć rozmiar **8 TB**, ale praktycznie ograniczenie (MBR) to liczba 2^{32} sektorów, czyli **2 TB**
- w systemie FAT32 można formatować tylko dyski, nie można natomiast zainstalować go na dyskietkach

FAT32

- w systemie FAT32 katalog główny może znajdować się w dowolnym miejscu na dysku i może zawierać maksymalnie 65 532 pliki i katalogi

Bajty	Rozmiar	Zawartość
00H-07H	8	Nazwa pliku w kodach ASCII
08H-0AH	3	Rozszerzenie nazwy pliku
0BH	1	Atrybuty pliku
0CH	1	Wielkość liter nazwy i rozszerzenia pliku
0DH	1	Czas utworzenia w milisekundach
0EH-0FH	2	Czas utworzenia
10H-11H	2	Data utworzenia
12H-13H	2	Czas ostatniego dostępu
14H-15H	2	Numer pierwszej JAP (16 starszych bitów)
16H-17H	2	Czas utworzenia lub aktualizacji pliku
18H-19H	2	Data utworzenia lub aktualizacji pliku
1AH-1BH	2	Numer pierwszej JAP (16 młodszych bitów)
1CH-1DH	2	Mniej znaczące słowo rozmiaru pliku
1EH-1FH	2	Bardziej znaczące słowo rozmiaru pliku

FAT32 - długie nazwy plików

- wprowadzone w systemie Windows 95
- informacje o nazwie pliku zapamiętywane są jako:
 - długa nazwa
 - skrócona nazwa (tzw. alias długiej nazwy)
- metoda tworzenia skróconej nazwy pliku:
 - rozszerzenie długiej nazwy staje się rozszerzeniem skróconej nazwy
 - pierwsze sześć znaków długiej nazwy staje się pierwszymi sześcioma znakami skróconej nazwy (nieodzwolone znaki zamieniane są na znak podkreślenia, małe litery zamieniane są na wielkie litery)
 - pozostałe dwa znaki nazwy skróconej to ~1 lub jeśli plik o takiej nazwie istnieje ~2, itd.

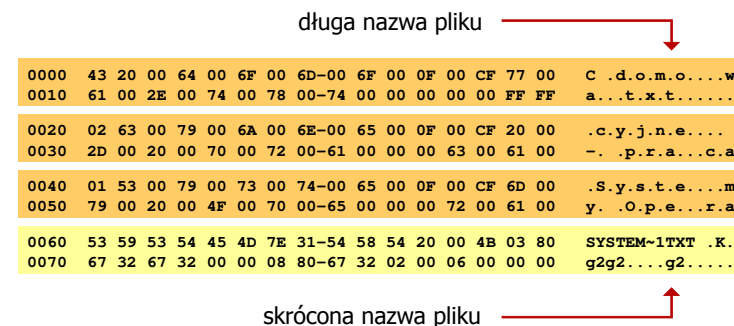
FAT32 - długie nazwy plików

- **skrócona nazwa pliku** przechowywana jest w identycznej, 32-bajtowej strukturze jak w przypadku plików w starym formacie 8+3
- **długie nazwy plików** zapisywane są także w 32-bajtowych strukturach, przy czym jedna nazwa zajmuje kilka struktur (w jednej strukturze umieszczonych jest 13 kolejnych znaków w formacie Unicode)

Bajty	Rozmiar	Zawartość
00H	1	Pierwsze 6 bitów określa numer fragmentu nazwy, bit 7 - czy jest to ostatni fragment nazwy, a bit 8 - czy plik został usunięty lub jego nazwa skrócona
01H-0AH	10	Pierwsze 5 znaków nazwy
0BH	1	Atrybut (zawsze F)
0CH	1	Zarezerwowany (zawsze 0)
0DH	1	Suma kontrolna wersji krótkiej 8+3
0EH-19H	12	Kolejne 6 znaków nazwy
1AH-1BH	2	Numer początkowego kłastera (zawsze 0)
1CH-1FH	4	Dwie ostatnie litery nazwy

FAT32 - długie nazwy plików

- Nazwa pliku: **Systemy Operacyjne - praca domowa.txt**



FAT - wady systemu plików FAT

- **fragmentacja wewnętrzna** - nawet najmniejszy plik zajmuje całą JAP - gdy rozmiar klastra jest duży, a na dysku znajduje się dużo małych plików - pewna część miejsca jest tracona
- **fragmentacja zewnętrzna** - silna fragmentacja plików pomiędzy wiele klastrów o bardzo różnym fizycznym położeniu na dysku (konieczność okresowej defragmentacji przy użyciu specjalnych narzędzi programowych)
- duże prawdopodobieństwo powstawania błędów zapisu, polegających na przypisaniu jednego klastra dwóm plikom (tzw. **crosslinks**), co kończy się utratą danych z jednego lub obu „skrzyżowanych” plików
- typowym błędem, pojawiającym się w systemie FAT, jest również pozostawianie tzw. **zagubionych klastrów (lost chains)**, tj. jednostek alokacji nie zawierających informacji, ale opisanych jako zajęte
- brak mechanizmów ochrony - **praw dostępu**

exFAT (FAT64)

- stworzony przez Microsoft na potrzeby pamięci Flash
- po raz pierwszy pojawił się w listopadzie 2006 roku w Windows Embedded CE 6.0 i Windows Vista SP1
- obsługiwany także przez Windows 7/8/10, Windows Server 2003/2008, Windows XP SP2/SP3, Linux
- może być używany wszędzie tam, gdzie NTFS nie jest najlepszym rozwiązaniem ze względu na dużą nadmiarowość struktury danych
- podstawowe cechy:
 - maksymalna wielkość pliku to 2^{64} = 16 EB
 - maksymalna wielkość klastra - do 32 MB
 - nieograniczona liczba plików w pojedynczym katalogu
 - prawa dostępu do plików i katalogów

NTFS (New Technology File System)

- **wersja 1.0** (połowa 1993 r.) - Windows NT 3.1
- **wersja 1.1** (jesień 1994 r.) - Windows NT 3.5
- **wersja 1.2** (NTFS 4) - Windows NT 3.51 (1995 r.)
- **wersja 3.0** (NTFS 5) - Windows 2000
- **wersja 3.1** (NTFS 5.1) - Windows XP/Server 2003/Vista/7/8/10
- teoretyczny **rozmiar partycji** NTFS wynosi $2^{64}-1$ klastrów, ale Windows potrafi obsłużyć tylko $2^{32}-1$ klastrów (dla klastra 64 kB - **ok. 256 TB**)
- tabela partycji w MBR dysku twardego ogranicza rozmiar partycji do **2 TB**
- teoretyczna **wielkość pliku** wynosi 2^{64} bajtów minus 1 kB, ale Windows ogranicza ten rozmiar do 2^{44} bajtów minus 64 kB (**ok. 16 TB**)

NTFS

- struktura wolumenu (dysku) NTFS:



- **Boot Sector** rozpoczyna się od zerowego sektora partycji, może zajmować 16 kolejnych sektorów, zawiera podobne dane jak w systemie FAT
- **MFT (Master File Table)** - specjalny plik, niewidoczny dla użytkownika, zawiera wszystkie dane niezbędne do odczytania pliku z dysku, składa się z rekordów o stałej długości (1 kB - 4 kB)
- pierwsze 16 (NTFS 4) lub 26 (NTFS 5) rekordów jest zarezerwowane dla tzw. metaplików, np.
 - rekord nr: 0 plik: **\$Mft** (główna tablica plików)
 - rekord nr: 1 plik: **\$MftMirr** (główna tablica plików 2)
 - rekord nr: 5 plik: **\$** (indeks katalogu głównego)

NTFS

- struktura wolumenu (dysku) NTFS:



- pozostała część pliku MFT przeznaczona jest na rekordy wszystkich plików i katalogów umieszczonych na dysku
- jeśli pierwszy rekord MFT jest uszkodzony to system automatycznie odczytuje drugi rekord, w którym zapisana jest kopia pierwszego
- położenie obu metaplików **\$Mft** i **\$MftMirr** zapisane jest w sektorze startowym partycji

NTFS

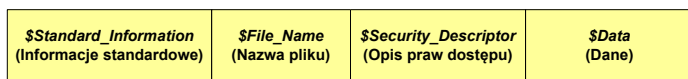
- struktura wolumenu (dysku) NTFS:



- plik w NTFS to **zbiór atrybutów**
- wszystkie atrybuty mają dwie części składowe: **nagłówek** i **blok danych**
- **nagłówek** opisuje atrybut, np. liczbę bajtów zajmowanych przez atrybut, rozmiar bloku danych, położenie bloku danych, znacznik czasu
- **bloku danych** zawiera informacje zgodne z przeznaczeniem atrybutu

NTFS - Pliki

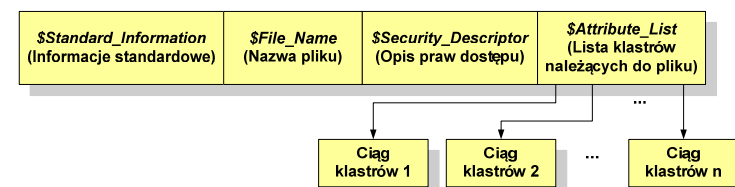
- pliki w systemie NTFS są reprezentowane w MFT przez rekord zawierający atrybuty:
 - **\$Standard_Information**
 - **\$File_Name**
 - **\$Security_Descriptor**
 - **\$Data**



- w przypadku małych plików wszystkie jego atrybuty zapisywane są bezpośrednio w MFT (atrybuty **rezydentne**)

NTFS - Pliki

- jeśli atrybuty pliku są duże (najczęściej dotyczy to atrybutu **\$Data**), to w rekordzie w MFT umieszczany jest tylko nagłówek atrybutu oraz wskaźnik do jego bloku danych, a sam blok danych przenoszony jest na dysk poza MFT (atrybuty **nierezydentne**)
- blok danych atrybutu nierezydentnego zapisywany jest w przyległych klastrach
- jeśli nie jest to możliwe, to dane zapisywane są w kilku ciągach jednostek alokacji i wtedy każdemu ciągowi odpowiada wskaźnik w rekordzie MFT



NTFS - Katalogi

- katalogi reprezentowane są przez rekordy zawierające trzy takie same atrybuty jak pliki:
 - `$Standard_Information`
 - `$File_Name`
 - `$Security_Descriptor`

<code>\$Standard_Information</code> (informacje standardowe)	<code>\$File_Name</code> (Nazwa pliku)	<code>\$Security_Descriptor</code> (Opis praw dostępu)	<code>\$Index_Root</code>	<code>\$Index_Allocation</code>	<code>\$Bitmap</code>
---	---	---	---------------------------	---------------------------------	-----------------------

- zamiast atrybutu `$Data` umieszczone są trzy atrybuty przeznaczone do tworzenia list, sortowania oraz lokalizowania plików i podkatalogów
 - `$Index_Root`
 - `$Index_Allocation`
 - `$Bitmap`

Koniec wykładu nr 11

Dziękuję za uwagę!