

Podstawy systemów operacyjnych

Wykład 3

Zarządzanie dyskiem

Formatowanie dysku HDD

- Nowy dysk magnetyczny – tarcze wykonane z materiału magnetycznego.
- Przed rozpoczęciem użytkowania dysku należy podzielić go na sektory, które sterownik dysku będzie potrafił zapisywać i odczytywać.
- Proces ten nazywa się **formatowaniem niskopoziomowym** lub **fizycznym**.
- Formatowanie polega na umieszczeniu specjalnej struktury danych w miejscach na dysku odpowiadających sektorom.
- Struktura danych sektora składa się z nagłówka, sekcji danych (najczęściej 512B) i zakończenia.

Nagłówek sektora

- Nagłówek i zakończenie sektora są używane przez sterownik dysku.
- Zawiera informacje:
 - nr sektora
 - kod korygujący (ang. ECC – error-correcting code)
- Podczas zapisu danych sterownik oblicza i uaktualnia kod ECC na podstawie obszaru danych.
- Przy czytaniu sektora, kod korygujący jest obliczany i porównywany z zapisanym. Jeśli wartości się **nie zgadzają**, to oznacza, że *sektor uległ uszkodzeniu*.

Formatowanie logiczne

- Aby użyć dysk do przechowywania plików należy podzielić go na partycje i utworzyć na nim system plików.
- Formatowanie logiczne – tworzenie systemu plików – polega na zapisaniu wstępnych struktur danych systemu plików, np. mapa wolnych i przydzielonych obszarów (np. tablica FAT) oraz początkowy, pusty katalog.

Surowy dysk

- **Linkowa tablica bloków logicznych** – bez żadnych struktur danych systemu plików.
- Surowe operacje wejścia-wyjścia – operacje przeprowadzane na surowym dysku.
- Z operacji tych korzystają niektóre *bazy danych*, gdyż pozwalają one na ścisłe kontrolowanie miejsca przechowywania na dysku poszczególnych rekordów.
- **Surowe operacje wejścia – wyjścia** wykonywane są z *pominięciem* wszelkich usług systemu plików, takich jak:
 - buforów podręcznych,
 - pobierania z wyprzedzeniem,
 - blokowanie pliku,
 - przydział miejsca,
 - nazwy plików i katalogi.

Blok rozruchowy

- W czasie uruchamiania komputera musi zostać wykonany jakiś program wstępny – **program rozruchowy** (bootstrap).
- Program rozruchowy ustawia stan początkowy wszystkich elementów systemu – od rejestru systemu po sterowniki urządzeń i zawartość pamięci głównej.
- W większości komputerów, program rozruchowy jest przechowywany w pamięci ROM (*Read Only Memory*) – wygodne rozwiązanie – nie wymaga żadnych zabiegów wstępnych, nie zmienia położenia i nie zostanie zainfekowany.
- **Kłopot?** – *przy zmianie programu rozruchowego należy wymienić układ scalony.*

Blok rozruchowy

- Zastąpienie programu **bootstrap** przez mały ładowacz programu rozruchowego (**bootstrap loader**).
- Pełny program rozruchowy można łatwo wymienić.
- Program rozruchowy umieszczany jest na dysku w strefie nazwanej blokami rozruchowymi w ustalonym miejscu.
- Dysk posiadający partycję rozruchową nazywa się dyskiem rozruchowym lub systemowym.

Zarządzanie obszarem wymiany

- Zarządzanie obszarem wymiany (*swap space management*) – jest niskopoziomowym zadaniem systemu operacyjnego.
- Pamięć wirtualna korzysta z przestrzeni dyskowej jako z *rozszerzenia* pamięci głównej.
- Podstawowym celem w projektowaniu obszaru wymiany jest umożliwienie *najlepszej przepustowości systemowej* pamięci wirtualnej.

Wykorzystanie obszaru wymiany

- Obszar wymiany jest wykorzystywany na wiele sposobów przez różne systemy operacyjne, zależnie od zaimplementowanych algorytmów zarządzania pamięcią.
- Obszar wymiany może zostać wykorzystany dp:
 - przechowywania całych procesów, łącznie z ich kodem i segmentami danych;
 - przechowywania stron (dla systemów ze stronicowaniem);
- Rozmiar obszaru wymiany uzależniony jest od rozmiaru pamięci fizycznej, rozmiaru pamięci wirtualnej i zaimplementowanych metod obsługi pamięci.

Położenie obszaru wymiany

- Obszar wymiany może rezydować w dwóch miejscach:
 - może znajdować się w obszarze zwykłego systemu plików,
 - może znajdować się na osobnej partycji
- Jeśli obszar wymiany jest jednym wielkim plikiem w obrębie systemu plików, to trzeba użyć zwykłych procedur systemu plików. Podejście niezbyt wydajne.
- Na osobnej partycji – bez żadnego systemu plików i bez struktury katalogów. Do tego obszaru wykorzystuje się osobnego zarządcę pamięci obszaru wymiany. Nie optymalizuje on zużycia pamięci lecz używa algorytmów zoptymalizowanych ze względu na szybkość.

Przykładowe systemy plików

- ext2,
- ext3
- ext4
- FAT
- NTFS
- reiserfs
- reiser 4

RAID

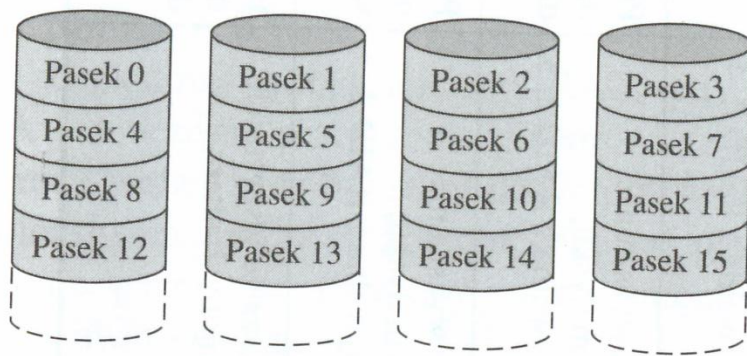
- » RAID (ang. Redundand Array of Independent Disks) to sposób połączenia dwóch lub większej ilości dysków twardych w jedną macierz, która zapewnia dodatkową funkcjonalność w porównaniu z oddzielnie podłączonymi pojedynczymi dyskami twardymi.
- » Macierze RAID są powszechnie stosowane w rozwiązaniach serwerowych, dzięki nim uzyskujemy:
 - > odporność na awarie,
 - > zwiększenie prędkości transmisji w porównaniu z pojedynczym dyskiem,
 - > zależnie od rodzaju pamięć widoczna jako jedno urządzenie.

RAID c.d

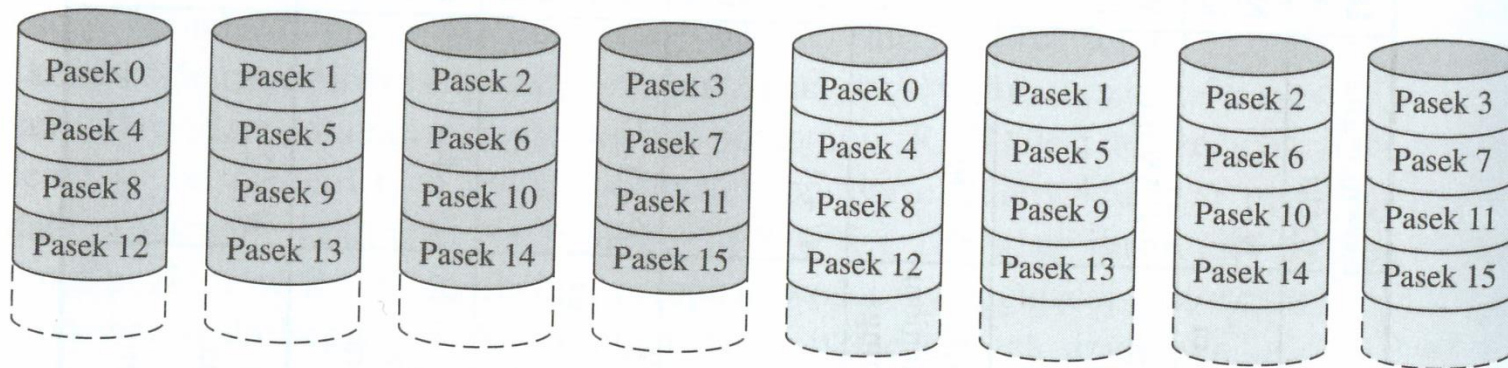
- » Macierze potrafią pracować zarówno równolegle jak też niezależnie.
- » W systemie wielodyskowym istnieje duża różnorodność organizacji danych zaś stabilność systemu można zwiększyć poprzez nadmiarowość.
- » Macierz RAID składa się z siedmiu poziomów ponumerowanych od zera do sześć. Poziomy te nie sugerują relacji hierarchicznej ale symbolizują zróżnicowane rodzaje architektury projektowej:
 1. RAID to zestaw napędów dysków fizycznych, które system operacyjny traktuje jako jeden napęd logiczny
 2. Dane są przesyłane do poszczególnych napędów macierzy.
 3. Nadmiarowa pojemność dysku jest wykorzystywana do przechowywania informacji o parzystości, co gwarantuje możliwość odzyskania danych na wypadek awarii dysku.
- » Poszczególne poziomy RAID różnią się szczegółami 2 i 3 właściwości.

Tabela 11.4. Poziomy macierzy RAID

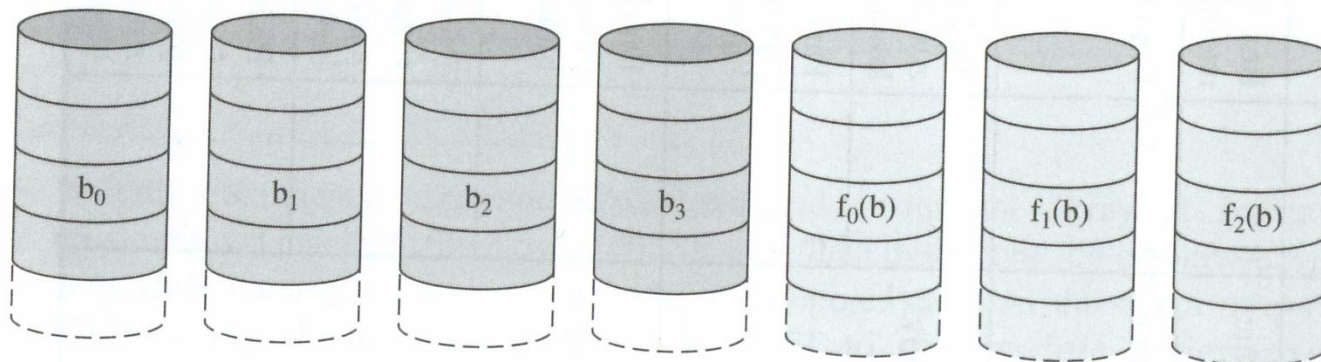
Kategoria	Poziom	Opis	Wymagane dyski	Dostępność danych	Duża przepustowość operacji we/wy	Niewielki współczynnik żądań we/wy
Paskowanie	0	Bez nadmiarowości	N	Niższa niż w przypadku jednego dysku	Bardzo duża	Bardzo wysoki zarówno dla odczytu, jak i zapisu
Kopiowanie lustrzane	1	Kopia lustrzana	$2N$, $3N$ itd.	Wyższa niż w RAID 2, 3, 4 lub 5, niższa niż w RAID 6	Wyższa niż odczyt z jednego dysku, podobna przy zapisie na jednym dysku	Do dwóch razy wyższy dla odczytu z pojedynczego dysku; podobny dla zapisu
Dostęp równoległy	2	Nadmiarowość z wykorzystaniem kodu Hamminga	$N + m$	O wiele wyższa niż w jednym dysku, wyższa niż RAID 3, 4 lub 5	Najwyższa ze wszystkich wymienionych alternatyw	W przybliżeniu dwukrotnie wyższy od jednego dysku
	3	Parzystość z przeplotem bitów	$N + 1$	O wiele wyższa niż w jednym dysku, porównywalna z RAID 2, 4 lub 5	Najwyższa ze wszystkich wymienionych alternatyw	W przybliżeniu dwukrotnie wyższy od jednego dysku
Dostęp niezależny	4	Parzystość z przeplotem bloków	$N + 1$	O wiele wyższa niż w jednym dysku, porównywalna z RAID 2, 3 lub 5	Podobna do RAID 0 dla odczytu, znacząco niższa dla zapisu na jednym dysku	Podobny do RAID 0 dla odczytu, znacząco niższy dla zapisu na jednym dysku
	5	Parzystość rozproszona z przeplotem bloków	$N + 1$	O wiele wyższa niż w jednym dysku, porównywalna z RAID 2, 3 lub 4	Podobna do RAID 0 dla odczytu, mniejsza niż w jednym dysku dla zapisu	Podobny do RAID 0 dla odczytu, ogólnie niższy niż w jednym dysku dla zapisu
	6	Podwójnie rozproszona parzystość z przeplotem bloków	$N + 2$	Najwyższa ze wszystkich wymienionych alternatyw	Podobna do RAID 0 dla odczytu, niższa niż RAID 5 dla zapisu	Podobny do RAID 0 dla odczytu, znacząco niższy niż RAID 5 dla zapisu



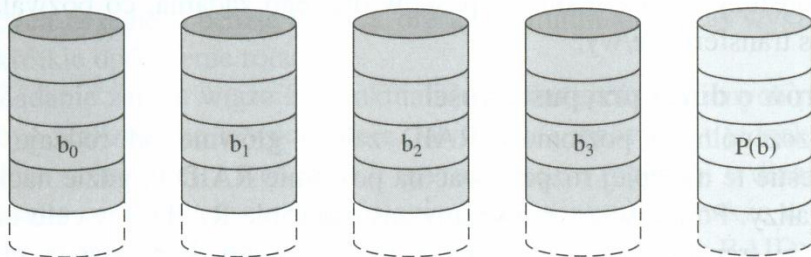
(a) RAID 0 (bez nadmiarowości)



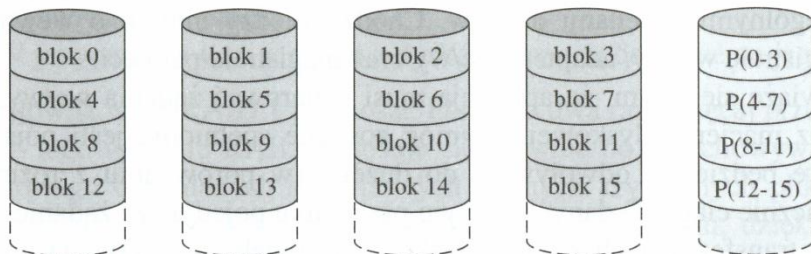
(b) RAID 1 (kopia lustrzana)



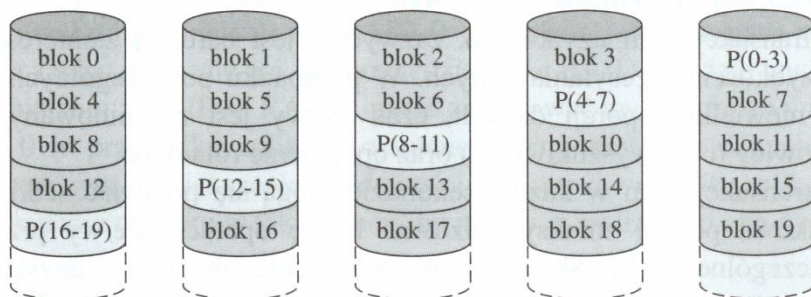
(c) RAID 2 (nadmiarowość z wykorzystaniem kodu Hamminga)



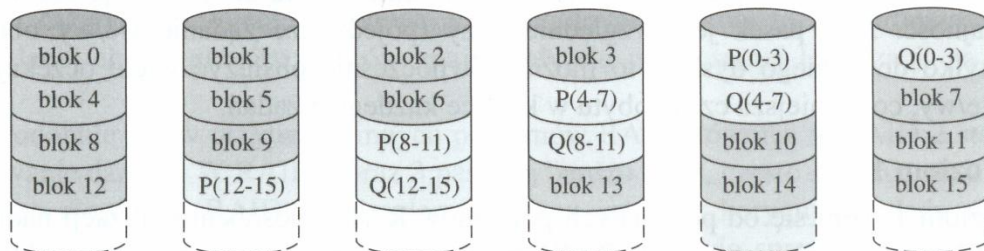
(d) RAID 3 (parzystość z przeplotem bitów)



(e) RAID 4 (parzystość z przeplotem bloków)



(f) RAID 5 (parzystość rozproszona z przeplotem bloków)



(g) RAID 6 (podwójna nadmiarowość)