

MACIERZE DYSKOWE RAID

Cel ogólny lekcji

Poznanie zasad działania macierzy dyskowych RAID, ich budowy, poziomów oraz zastosowań w systemach informatycznych, a także umiejętność doboru odpowiedniego poziomu RAID w zależności od wymagań dotyczących wydajności i bezpieczeństwa danych.

Cele szczegółowe Uczeń:

1. Wyjaśnia pojęcie **RAID (Redundant Array of Independent Disks)**.
2. Rozróżnia **rodzaje RAID**: sprzętowy, programowy, hybrydowy (Fake RAID).
3. Opisuje podstawowe mechanizmy RAID: striping, mirroring, parzystość (parity).
4. Rozpoznaje i porównuje poziomy RAID: RAID 0, 1, 5, 6, 10, RAID 50, 60 (poziomy złożone).
5. Wymienia **zalety i wady poszczególnych poziomów RAID**.
6. Wyjaśnia pojęcia: hot swap, hot spare, rebuild, degraded.
7. Analizuje wpływ RAID na: wydajność systemu, bezpieczeństwo danych, dostępność usług.
8. Wyjaśnia, dlaczego: RAID nie zastępuje kopii zapasowych, proces odbudowy (rebuild) jest krytyczny dla bezpieczeństwa danych.
9. Dobiera odpowiedni poziom RAID do zastosowania: system operacyjny, serwer plików, baza danych, archiwizacja danych.
10. Rozróżnia podstawowe interfejsy stosowane w systemach RAID: SATA, SAS, NVMe.

Cele operacyjne (praktyczne) Uczeń potrafi:

- zaprojektować prostą macierz RAID,
- wskazać optymalną konfigurację dla danego scenariusza,
- zinterpretować stan pracy macierzy RAID.

1. Cele kształcenia

Uczeń:

- definiuje RAID,
- rozróżnia poziomy RAID,
- analizuje wydajność i niezawodność,
- dobiera konfigurację,
- rozumie ograniczenia i ryzyka.

2. Definicja i idea RAID

RAID (Redundant Array of Independent Disks) – technologia łączenia wielu fizycznych dysków w jedną logiczną jednostkę.

Główne cele:

- wydajność (performance),
- odporność na awarie (fault tolerance),
- dostępność (availability).

RAID wykorzystuje:

- nadmiarowość (redundancję),
- równoległe operacje I/O.

3. Typy implementacji RAID

3.1 RAID sprzętowy

- kontroler RAID z procesorem
- cache (write-back, write-through)
- BBU (Battery Backup Unit)

Zalety:

- niskie obciążenie CPU
- wysoka wydajność
- zaawansowane funkcje

3.2 RAID programowy

- Linux: mdadm
- Windows: Storage Spaces / Dynamic Disks

Zalety:

- tani
- elastyczny

Wady:

- obciąża CPU
- zależność od systemu

3.3 Fake RAID

- BIOS/UEFI RAID
- wymaga sterowników

w praktyce: **pomiędzy sprzętowym a programowym**

4. Stany macierzy

Stan	Znaczenie
Optimal	pełna sprawność
Degraded	uszkodzony dysk
Rebuild	odbudowa
Failed	awaria macierzy

5. Mechanizmy RAID (kluczowe)

Striping

- podział danych na bloki (np. 64 KB)
- zapis równoległy

> wzrost wydajności

Mirroring

- kopia bloków danych

> pełna redundancja

Parzystość (XOR)

- operacja XOR:

$A \oplus B \oplus C = \text{Parzystość}$

- brak jednego dysku > możliwa rekonstrukcja

Hot Swap / Hot Spare

- hot swap: wymiana dysku bez shutdown
- hot spare: automatyczny zapasowy dysk

6. Poziomy RAID – pełna analiza

RAID 0

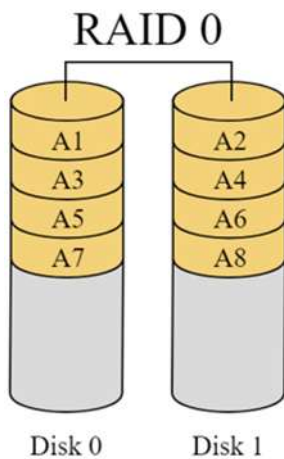
Mechanizm: striping

Min: 2 dyski

bardzo wysoka wydajność

brak odporności

MTBF spada wraz z liczbą dysków



RAID 1

Mechanizm: mirroring

odporność na awarię 1 dysku

szybki odczyt

50% pojemności

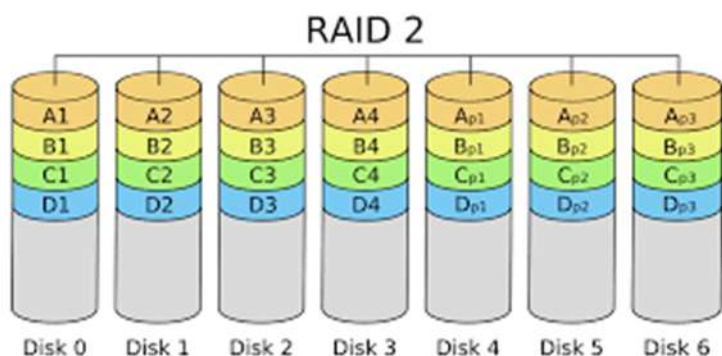


RAID 2 (historyczny)

- kod Hamminga
- korekcja błędów sprzętowa

obecnie NIEużywany

> zastąpiony przez ECC w dyskach



Zalety:

- wysoka dokładność dzięki kodowi Hamminga
- teoretycznie wysoka niezawodność

Wady:

- bardzo złożona implementacja
- wymaga wielu dysków
- obecnie nieużywany (zastąpiony przez ECC w dyskach)

RAID 3

Mechanizm:

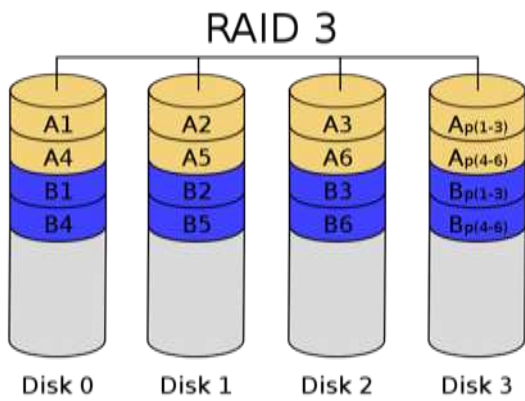
- striping na poziomie bajtów
- dedykowany dysk parzystości

Charakterystyka:

dobrze dla dużych plików

brak równoległości operacji

Jeden dysk = bottleneck występuje w RAID 3 i RAID 4, jeden dysk parzystości ogranicza wydajność w RAID 5 problem ten został wyeliminowany poprzez rozproszenie parzystości.



Zalety:

- wysoka wydajność przy dużych plikach (np. multimedia)
- efektywne wykorzystanie przepustowości

Wady:

- brak równoczesnego dostępu wielu operacji
- jeden dysk parzystości stanowi wąskie gardło (bottleneck)
- rzadko stosowany

W RAID 5 problem wąskiego gardła (bottleneck) został rozwiązany dzięki rozproszeniu parzystości na wszystkie dyski.

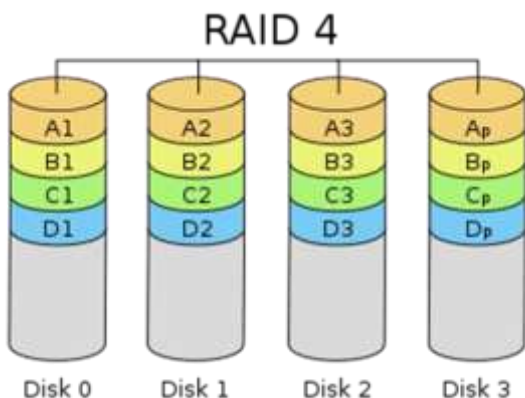
RAID 4

Mechanizm:

- striping blokowy
- dedykowany dysk parzystości

Charakterystyka:

- szybki odczyt
- wolny zapis (wąskie gardło)



Jeden dysk parzystości, każdy zapis musi trafić właśnie tam efekt: ten dysk jest stale obciążony, tworzy się wąskie gardło (bottleneck).

Zalety:

- szybki odczyt danych
- prostsza odbudowa niż RAID 3

Wady:

- wolny zapis
- jeden dysk parzystości = wąskie gardło
- praktycznie wyparty przez RAID 5

W RAID 5 problem wąskiego gardła (bottleneck) został rozwiązany dzięki rozproszeniu parzystości na wszystkie dyski.

RAID 5

Mechanizm:

- striping blokowy
- rozproszona parzystość
- kompromis wydajność / bezpieczeństwo

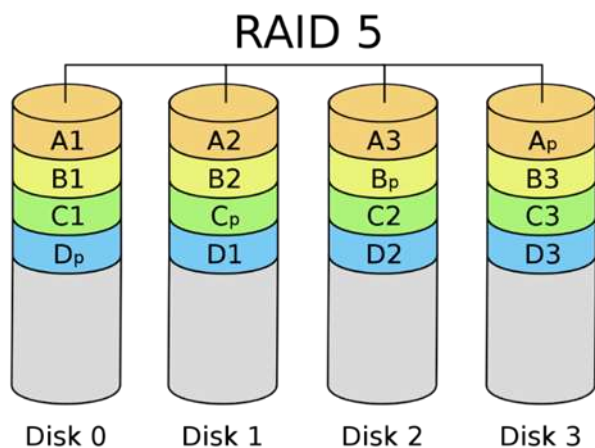
Charakterystyka:

- brak pojedynczego wąskiego gardła (bottleneck) żaden pojedynczy dysk nie jest stale przeciążony
- zapis jest wolniejszy niż odczyt
- wydajność zależy od kontrolera RAID

WRITE PENALTY RAID 5 i 6. Podczas zapisu wymagane są dodatkowe operacje:

1. odczyt danych
2. odczyt parzystości
3. zapis danych
4. zapis parzystości

powoduje to spowolnienie zapisu



Ważna różnica względem RAID 4

RAID 4 ma jeden „spowalniający dysk” (dysk parzystości), natomiast RAID 5 rozkłada obciążenie na wszystkie dyski, eliminuje pojedynczy bottleneck

Zalety:

- dobry kompromis między wydajnością a bezpieczeństwem
- brak pojedynczego wąskiego gardła
- stosunkowo efektywne wykorzystanie pojemności

Wady:

- występuje „write penalty” (wolniejszy zapis)
- wydajność spada podczas odbudowy (rebuild)
- duże ryzyko utraty danych przy awarii drugiego dysku w trakcie odbudowy

Uwaga: RAID 5 jest coraz rzadziej stosowany przy dużych dyskach ze względu na ryzyko utraty danych podczas odbudowy.

RAID 6

Mechanizm:

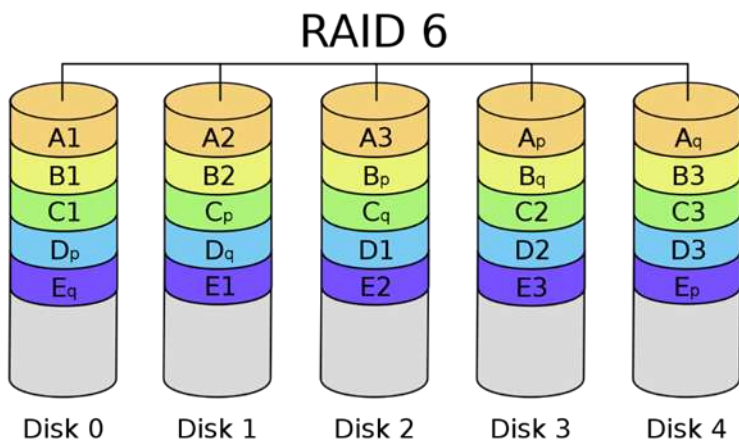
- podwójna parzystość

toleruje 2 awarie

bezpieczny dla dużych dysków

większy narzut CPU

wolniejszy zapis



RAID 6 - Striping z podwójną parzystością

Dodatkowe zalety:

- bezpieczny dla dużych dysków (np. 4 TB+)

Dodatkowe wady:

- jeszcze większy narzut obliczeniowy niż RAID 5
- wolniejszy zapis

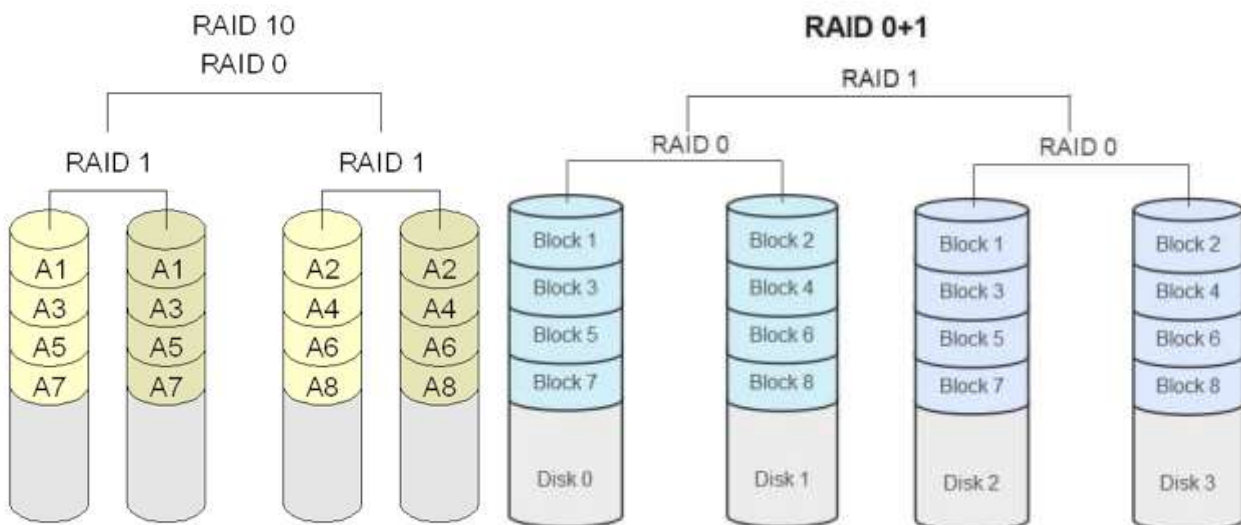
RAID 10

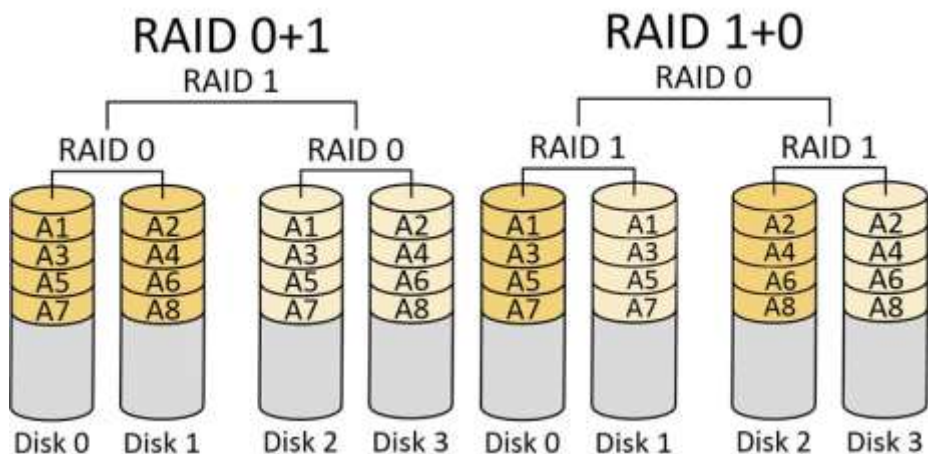
Mechanizm:

- RAID 1 + RAID 0

najwyższa wydajność, szybka odbudowa

koszt (50% pojemności)





Dodatkowe zalety:

- bardzo szybka odbudowa danych
- brak operacji parzystości (brak dużego narzutu CPU)

Dodatkowe wady:

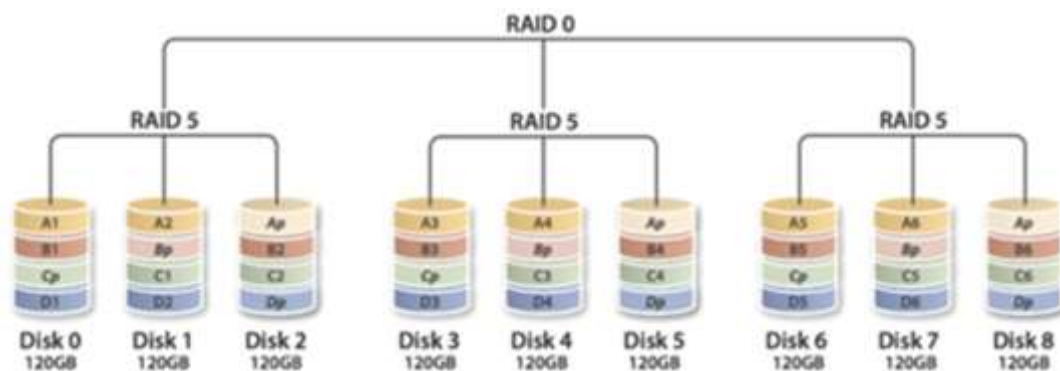
- wysoki koszt (duża liczba dysków)

7. RAID złożone

RAID 50

- RAID 5 + RAID 0
- wysoka wydajność

toleruje wiele awarii (po jednej na grupę)



Zalety:

- wysoka wydajność odczytu i zapisu
- większa odporność niż RAID 5

Wady:

- duża liczba wymaganych dysków

- większa złożoność konfiguracji

RAID 60

- RAID 6 + RAID 0

bardzo wysoka odporność

Zalety:

- bardzo wysoka odporność na awarie
- możliwość utraty kilku dysków (w różnych grupach)

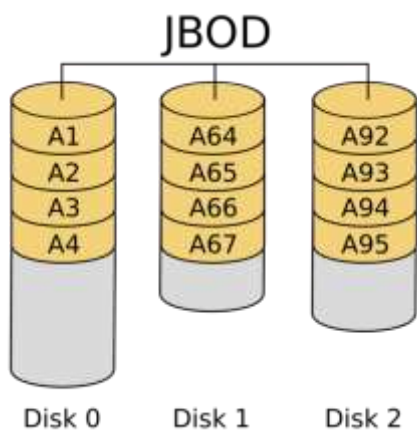
Wady:

- duże wymagania sprzętowe
- wysoki koszt
- wolniejszy zapis niż RAID 50

8. Inne konfiguracje

JBOD

- brak RAID
- dyski osobno



SPAN (NRAID)

- pojemność sumowana

brak redundancji

9. Interfejsy i architektura

SATA

- niski koszt
- niższa wydajność

SAS

- serwerowy
- wyższa niezawodność

NVMe

- PCIe
- bardzo szybkie (SSD)

Fibre Channel

- SAN
- duża odległość i przepustowość

W nowoczesnych systemach coraz częściej stosuje się macierze RAID oparte o dyski SSD i interfejs NVMe.

10. Kontrolery RAID (rozszerzenie)

PCI > SCSI

- starsze macierze enterprise
- wysoka wydajność
- cache + BBU

PCI > IDE / SATA

- tańsze
- małe serwery

SCSI > SCSI

- SAN
- redundantne kontrolery
- monitoring (SAF-TE)

NOWOCZESNE:

- RAID w kontrolerach SAS
- RAID w NAS (Synology, QNAP)
- RAID w chmurze (software-defined storage)

11. Rebuild (odbudowa) – krytyczny proces

- proces przywracania danych po awarii

Cechy:

- może trwać wiele godzin (szczególnie przy dużych dyskach)
- obciąża system i obniża wydajność
- w tym czasie macierz jest bardziej narażona na kolejną awarię

Szczególnie niebezpieczny w RAID 5 – awaria drugiego dysku oznacza utratę danych.

Czas odbudowy:

- HDD 4TB > nawet 24h+
- SSD szybciej

Ryzyko:

RAID 5: podczas rebuild = wysokie ryzyko drugiej awarii

RAID 6: bezpieczniejszy

12. Awaryjność systemu RAID

Najczęstsze przyczyny:

- awaria dysku (najczęściej)
- przegrzewanie
- błędna konfiguracja
- awaria kontrolera
- błędy firmware

13. RAID ≠ Backup

RAID NIE chroni przed:

- usunięciem pliku
- ransomware
- uszkodzeniem systemu plików
- kradzieżą
- pożarem
- błędem użytkownika
- uszkodzeniem systemu plików

obowiązkowy backup (3-2-1)

14. RAIT – RAID dla taśm

RAIT (Redundant Array of Independent Tapes)

striping i parzystość dla taśm
duża przepustowość

Zastosowanie:

- archiwizacja
- centra danych
- backup offline

15. Porównanie RAID

RAID	Min	Awaria	Wydajność	Pojemność
0	2	brak	bardzo wysoka	100%
1	2	1 dysk	wysoka	50%
5	3	1 dysk	średnia	~70–90%
6	4	2 dyski	średnia	~60–80%
10	4	wiele	bardzo wysoka	50%

16. Dobór RAID – praktyka

Scenariusz	RAID
gaming / render	0
system OS	1
NAS domowy	5
firma / backup	6
baza danych	10

Porównanie podstawowych poziomów RAID

RAID 0 – najwyższa wydajność, brak bezpieczeństwa

RAID 1 – wysoka niezawodność, niska efektywność pojemności

RAID 5 – kompromis, ale wolniejszy zapis

RAID 6 – wyższe bezpieczeństwo niż RAID 5

RAID 10 – najwyższa wydajność i bardzo dobra niezawodność

17. Wnioski końcowe

- RAID zwiększa **wydajność i dostępność danych**, a w zależności od poziomu także ich bezpieczeństwo.
- RAID **nie zastępuje kopii zapasowych (backupów)** – konieczne jest stosowanie dodatkowych mechanizmów zabezpieczających dane.

Dobór RAID w praktyce

- **RAID 5**
 - > dobry kompromis między wydajnością a bezpieczeństwem,
 - > jednak **coraz rzadziej zalecany przy dużych dyskach** (wysokie ryzyko podczas odbudowy).
- **RAID 6**
 - > **bezpieczniejszy wybór** niż RAID 5,
 - > odporny na awarię dwóch dysków jednocześnie.
- **RAID 10**
 - > **najlepsza wydajność i wysoka niezawodność**,
 - > stosowany w systemach krytycznych (np. bazy danych).

Najważniejsze zasady

- duże dyski (np. 4 TB i więcej) > preferowany **RAID 6 lub RAID 10**
- im większa macierz > tym większe ryzyko podczas procesu **rebuild**
- RAID zwiększa niezawodność systemu, ale **nie chroni przed wszystkimi zagrożeniami**