Materiały do wykładu

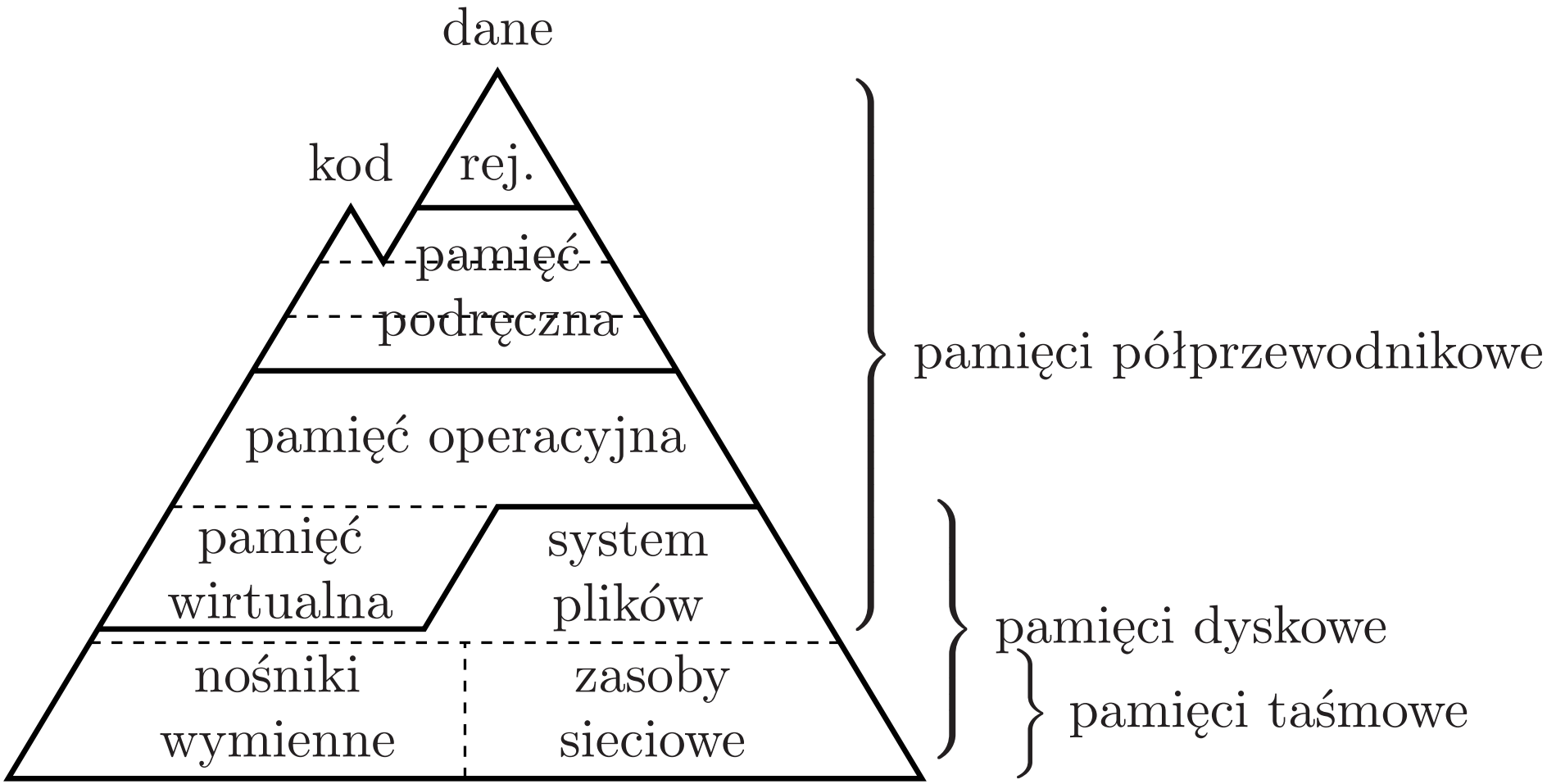
5. Pamięci

Marcin Peczarski

Instytut Informatyki
Uniwersytet Warszawski

12 kwietnia 2014

Hierachia pamięci



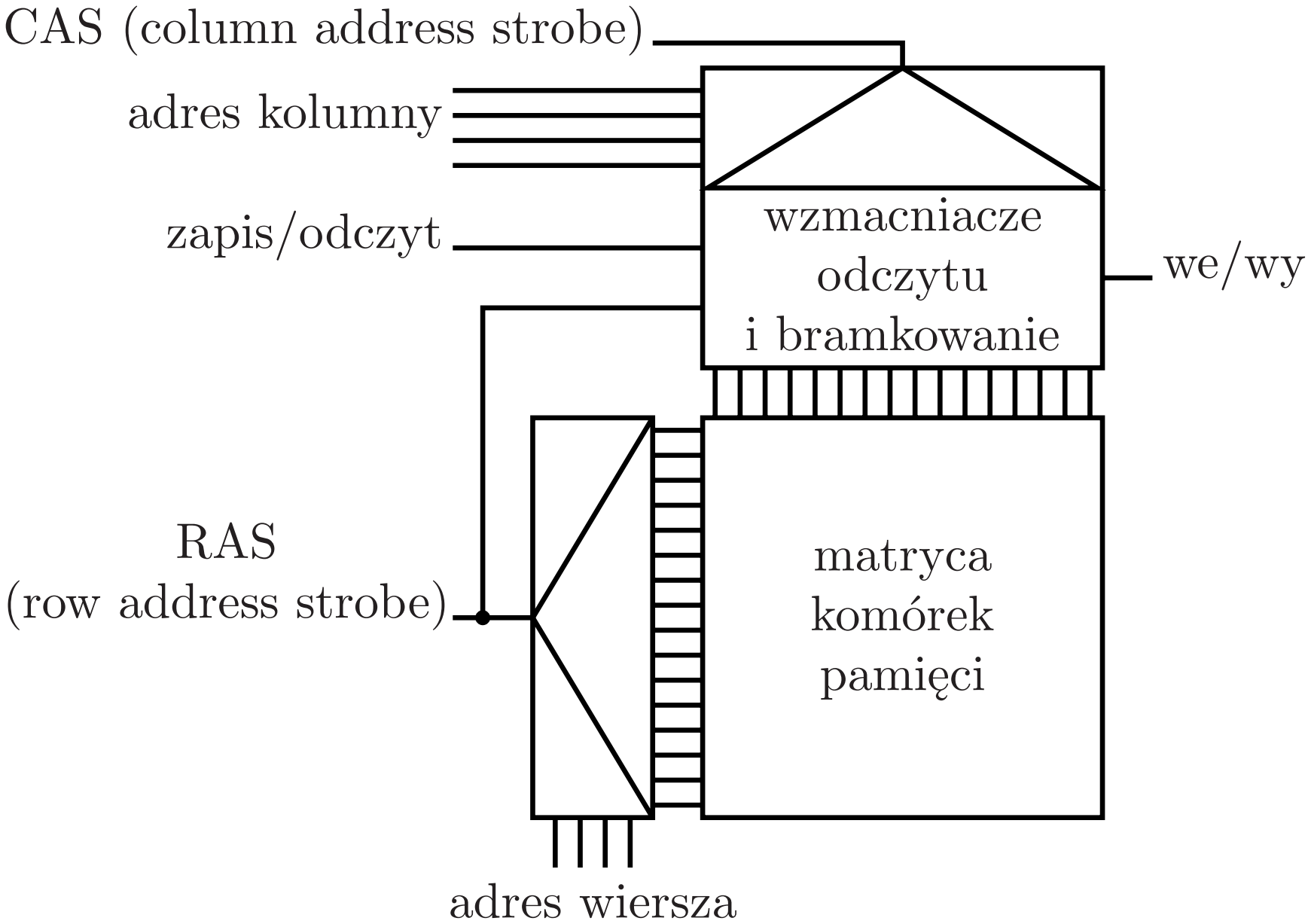
Pamięci półprzewodnikowe (1)

- Trwałość przechowywanej informacji
  - ulotne
  - nieulotne
- Sposób adresowania
  - adresowanie poszczególnych bajtów lub słów
  - bezadresowe (np. pamięci FIFO)
  - adresowanie zawartością, asocjacyjne
- Liczba magistral
  - jednobramowe, jednodostępne
  - wielobramowe, wielodostępne

Pamięci półprzewodnikowe (2)

- Sposób wprowadzania i wyprowadzania informacji
  - równoległe
  - szeregowe
  - szeregowo-równoległe
- Rodzaj synchronizacji
  - asynchroniczne
  - synchroniczne
- Organizacja wewnętrzna
  - liczba matryc
  - rozmiar matrycy
  - liczba bitów pamiętana w pojedynczej komórce

Organizacja wewnętrzna pamięci półprzewodnikowych



Parametry użytkowe pamięci półprzewodnikowych

- Pojemność
- Szybkość dostarczania danych
- Koszt na jeden bit
- Ziarnistość
- Czas dostępu
- Czas cyklu
- Organizacja zewnętrzna

Organizacja zewnętrzna pamięci półprzewodnikowych 5.6

- **układ scalony**

$$\text{liczba adresów} \times \text{liczba bitów w układzie} = \text{pojemność układu}$$

$$64\ \text{M} \times 4\ \text{b} = 256\ \text{Mb}$$

- **moduł**

$$\text{liczba układów} \times \text{liczba bitów w układzie} = \text{liczba bitów w module}$$

$$18 \times 4\ \text{b} = 72\ \text{b} = 64\ \text{b danych} + 8\ \text{b ECC}$$

$$\text{liczba adresów} \times \text{liczba bitów danych} = \text{pojemność modułu}$$

$$64\ \text{M} \times 64\ \text{b} = 512\ \text{MB}$$

- **pamięć**

$$\text{pojemność modułu} \times \text{liczba modułów} = \text{pojemność pamięci}$$

$$512\ \text{MB} \times 2 = 1\ \text{GB}$$

Pamięci nieulotne

- MROM, ROM – mask programmable read only memory
- PROM – programmable read only memory
- UV-EPROM, EPROM – erasable programmable read only memory
- OTPROM – one time programmable read only memory
- EEPROM – electrically erasable programmable read only memory
- Flash
  - NOR
  - NAND

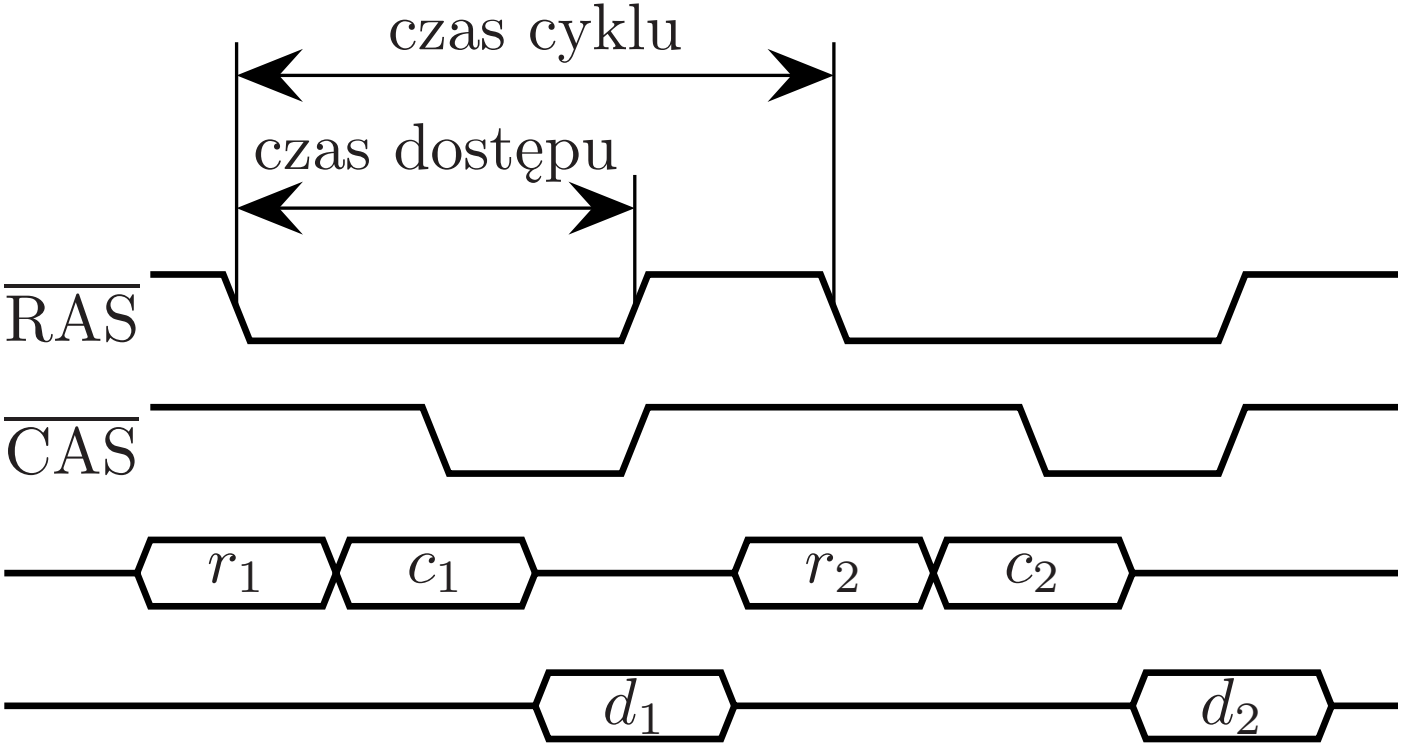
Pamięci o dostępie swobodnym

- RAM – random access memory
  - statyczne – SRAM
  - dynamiczne – DRAM
- Pamięci dynamiczne wymagają odświeżania.
- Pamięci statyczne są szybsze niż dynamiczne.
- Cena za bit w pamięciach statycznych jest większa niż w pamięciach dynamicznych.

Pamięci dynamiczne

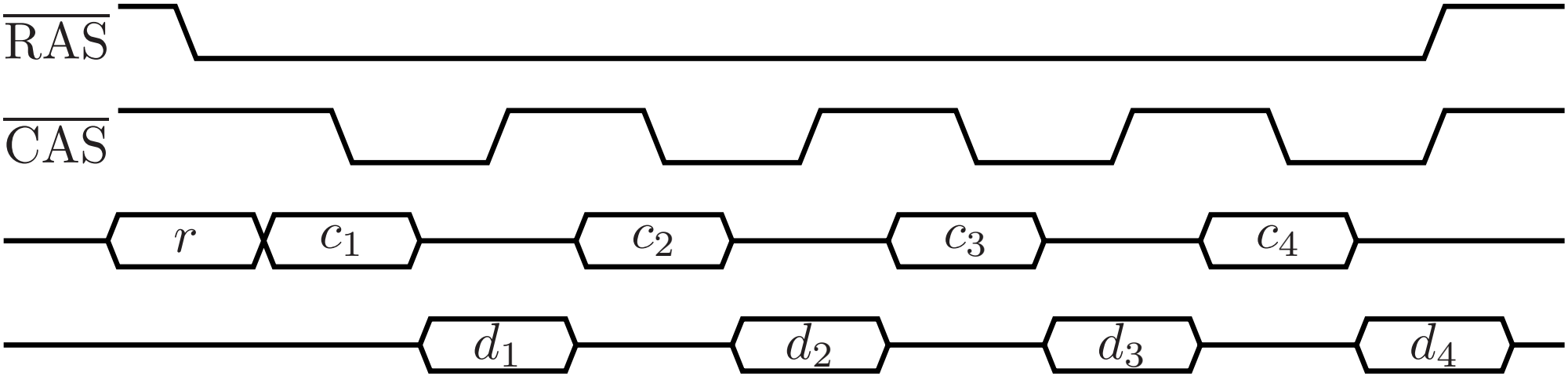
- Asynchroniczne
- Synchroniczne
  - (SDR) SDRAM
  - DDR(x) SDRAM

Cykl odczytu pamięci dynamicznej

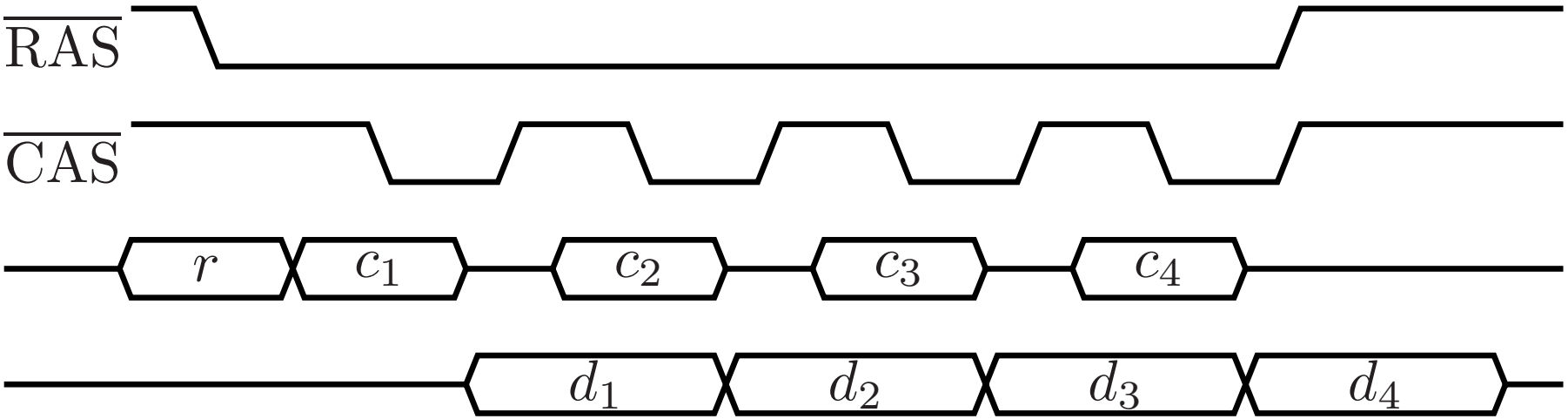


Tryb stronicowy

▷ FPM – fast page mode

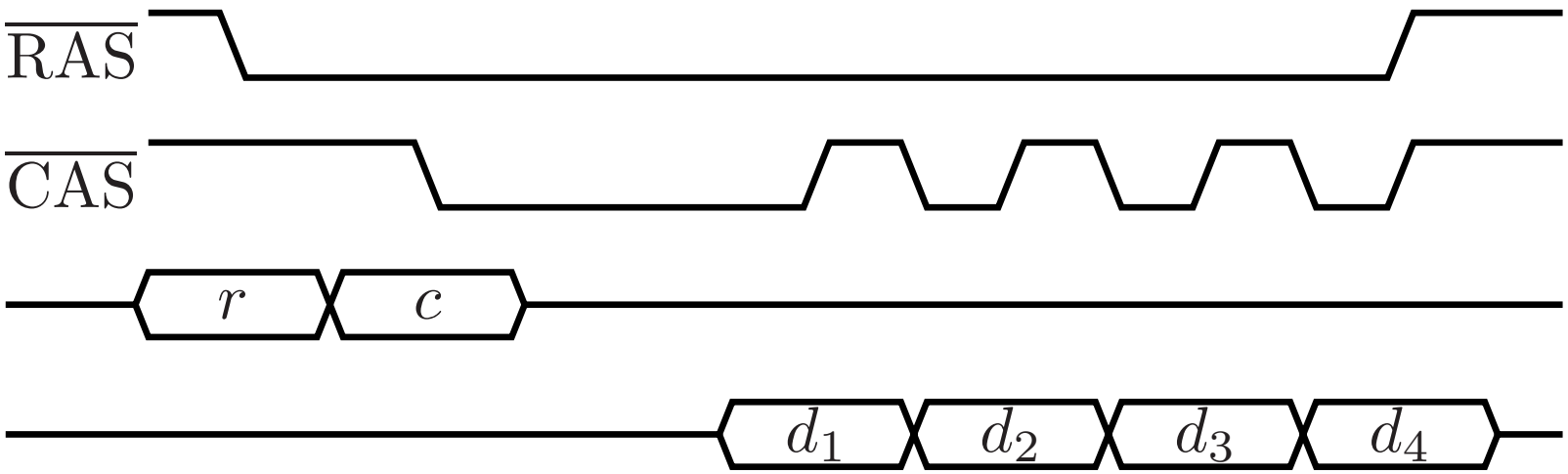


▷ EDO – extended data out

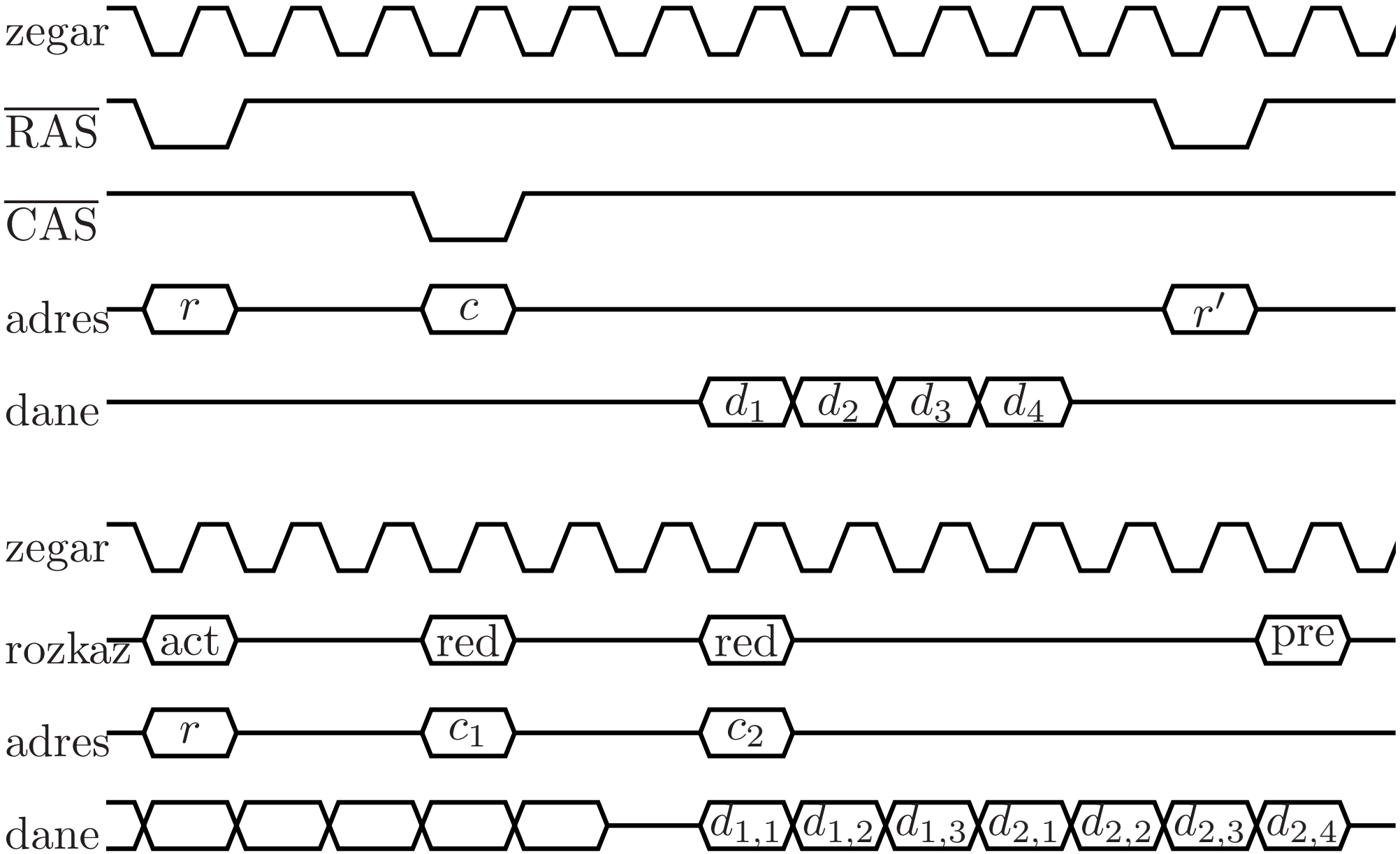


Tryb pakietowy

3-1-1-1



Pamięci synchroniczne



Pamięci podręczne

- Niedopasowanie między szybkościami działania procesora i pamięci operacyjnej
- Mniejsza pamięć może być szybsza
- Zasada lokalności odwołań do pamięci
  - czasowa
  - przestrzenna
- Do 3 poziomów
- Stopniowa integracja ze strukturą mikroprocesora
- Osobne pamięci pierwszego poziomu dla danych i kodu

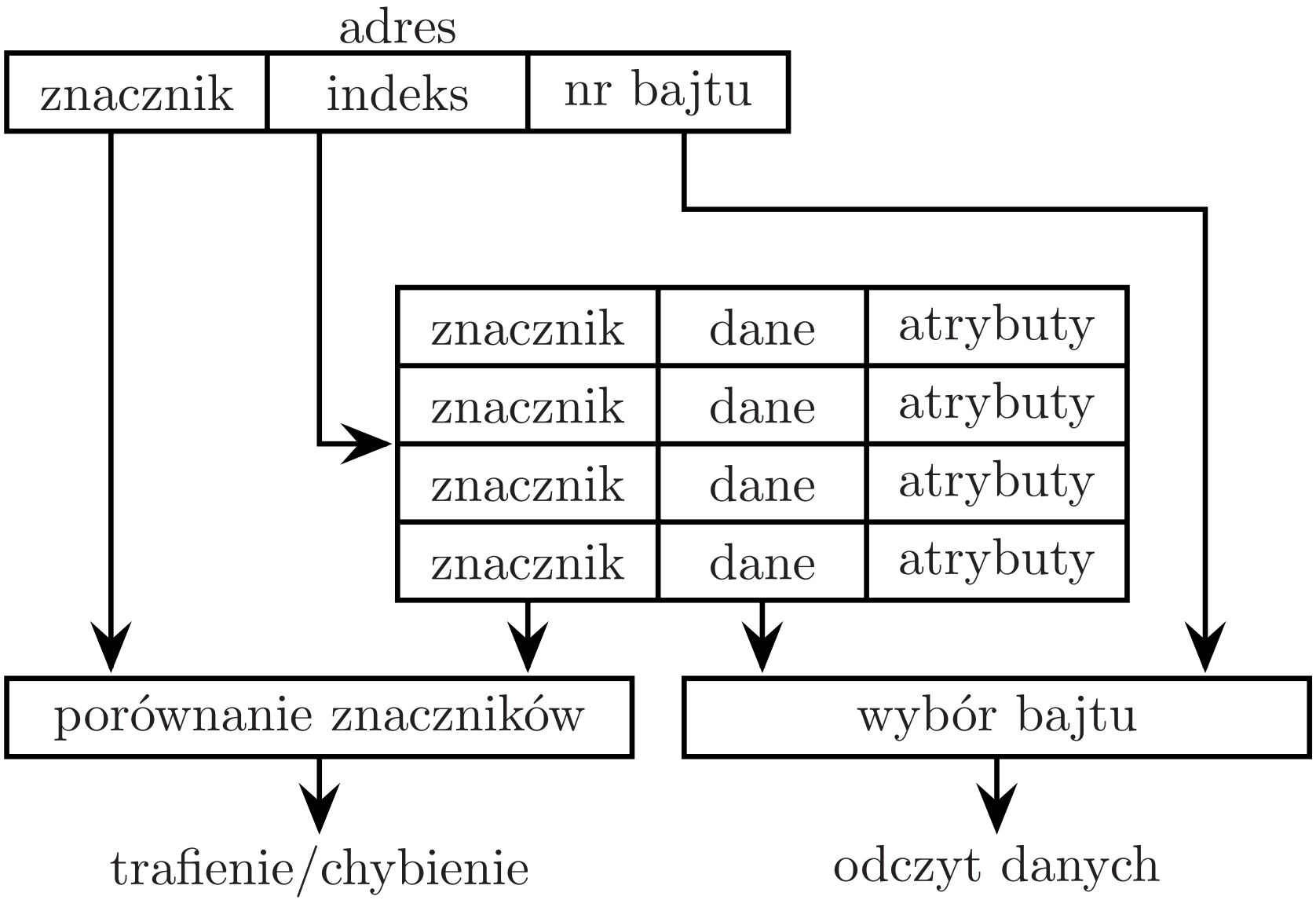
Zasada działania

$$t_e = pt_c + (1-p)(t_c + t_m)$$

$$\frac{t_e}{t_m} = \frac{t_c}{t_m} + 1 - p$$

t_e – efektywny średni czas dostępu
t_c – czas dostępu do pamięci podręcznej
t_m – czas dostępu do pamięci głównej
p – prawdopodobeństwo trafienia (ang. *hit rate*)
$1-p$ – prawdopodobeństwo chybienia (ang. *miss rate*)

Odwzorowanie bezpośrednie, 1-skojarzeniowe



Zjawisko migotania

- Naprzemienne odczytywanie danych spod adresów o tym samym indeksie, ale różnych znacznikach

- Skojarzeniowość – liczba bloków (wierszy, linii), w których mogą znajdować się dane o takim samym znaczniku

- Odwzorowanie sekcyjno-skojarzeniowe, n-skojarzeniowe, $n > 1$
  - Wymaga równoległego porównania n znaczników przy każdym odczycie.

- Odwzorowanie (w pełni) skojarzeniowe, asocjacyjne
  - Wymaga równoległego porównania wszystkich znaczników przy każdym odczycie.

Algorytmy zastępowania

- Jeśli skojarzeniowość jest większa niż jeden, potrzebny jest algorytm zastępowania.
- LRU (least recently used, najdawniej ostatnio używany) – zastępowany blok, który pozostawał w pamięci podręcznej najdłużej bez odwoływania się do niego
- Pseudo-LRU – uproszczona wersja LRU
- Non-MRU – uproszczona wersja LRU
- FIFO (first in first out, pierwszy wchodzi pierwszy wychodzi) – zastępowany blok, który pozostawał w pamięci podręcznej najdłużej
- LFU (least frequently used, najrzadziej używany) – zastępowany blok, do którego było najmniej odwołań
- Zastępowany losowo wybrany blok

Uzgadnianie zawartości między różnymi poziomami

- Zapis jednoczesny (ang. *write through*)
  - Zapis do pamięci głównej odbywa się równocześnie z zapisem do pamięci podręcznej.
  - Potrzebny jest bufor dla operacji zapisu do pamięci głównej.
- Zapis opóźniony (ang. *copy/write back*)
  - Zapis odbywa się zawsze tylko do pamięci podręcznej.
  - Zapis do pamięci głównej odbywa się, gdy zachodzi konieczność usunięcia z pamięci podręcznej bloku, który ma inną zawartość niż jego odpowiednik w pamięci głównej.
- Wcześniejszy odczyt (ang. *prefetch*)

Uzgadnianie zawartości na jednym poziomie

- Algorytmy
  - MESI
  - MSI
  - MOSI
  - MOESI

- Stany
  - Modified – mam ważną kopię bloku, który został zmodyfikowany.
  - Owned – jestem właścicielem tego bloku i ponoszą odpowiedzialność za wszystkie odwołania do niego.
  - Exclusive – mam ważną kopię bloku, który nie został zmodyfikowany.
  - Shared – wszyscy mają kopię tego bloku.
  - Invalid – blok zawiera nieważne dane.

Prawdopodobieństwo trafienia

- Zależy od:
  - stosunku wielkości pamięci podręcznej do pamięci głównej,
  - wielkości bloku pamięci podręcznej (nieliniowo),
  - skojarzeniowości.

Pamięci masowe

- Urządzenia wejścia-wyjścia
- Operowanie na blokach, sektorach danych
- Technologie „mechaniczne”:
  - dyski magnetyczne (dyski twarde, dyskietki)
  - taśmy magnetyczne
  - dyski optyczne
  - dyski magnetyczno-optyczne
- Technologie półprzewodnikowe:
  - pamięci USB (ang. *pendrive*)
  - SSD (ang. *solid-state drive*, *solid-state disk*)

Budowa dysku twardego

- Jeden do kilku dysków umieszcza się na wspólnej osi w hermetycznej obudowie.
- Dyski wykonuje się z aluminium lub szkła.
- Dyski pokrywa się twardym ferromagnetykiem:
  - pierwiastki ferromagnetyczne: żelazo, kobalt, nikiel, gadolin, dysproz;
  - stopy i związki chemiczne (np. tlenki) pierwiastków ferromagnetycznych;
  - inne związki chemiczne, np. dwutlenek chromu.

Zapisywanie i odczytywanie dysku twardego

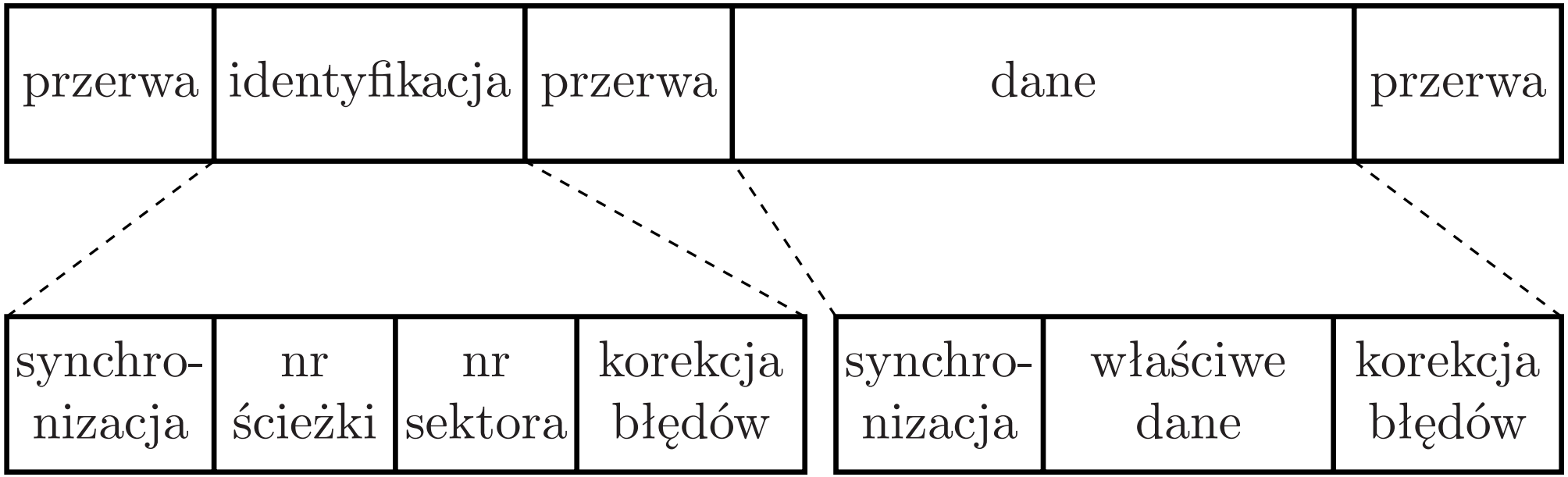
- Zespół głowic porusza się w przybliżeniu wzdłuż promienia dysku.
- Głowica zapisująca zbudowana jest z rdzenia ze szczeliną, wykonanego z miękkiego ferromagnetyka, na którym „nawinięto” cewkę z drutu miedzianego.
- Przepływ prądu w cewce powoduje powstawanie w szczelinie pola magnetycznego, które przemagnesowuje nośnik.
- Odczyt:
  - indukcja elektromagnetyczna – głowica odczytująca ma konstrukcję analogiczną jak głowica zapisująca, może to być ta sama głowica;
  - zjawisko magnetorezystancji – zmiana rezystancji półprzewodnika pod wpływem pola magnetycznego.

Organizacja dysku twardego (1)

- Liczba głowic odpowiada liczbie wykorzystanych stron dysków.
- Dane rozmieszczone są na koncentrycznych ścieżkach.
- Cylindrem nazywamy zbiór ścieżek, które mogą być odczytane przez różne głowice bez ich przesuwania.
- Ścieżka podzielona jest na sektory.
- Liczba sektorów na ścieżce może być stała lub zmienna.
- Przeplot
- Eliminowanie uszkodzonych sektorów

Organizacja dysku twardego (2)

▷ Format sektora



▷ Czas dostępu

- ◇ czas poszukiwania ścieżki (ang. *seek delay*)
- ◇ opóźnienie obrotowe (ang. *rotational delay*)
- ◇ czas przesyłania (ang. *transfer time*)

Dyski optyczne

- Odczyt za pomocą lasera
  - landy – fragmenty odbijające
  - pity – fragmenty wytłumiające

- Zapis
  - tłoczenie z matrycy – dyski tylko do odczytu (typu ROM)
  - wypalanie laserem barwnika polimerowego – dyski do jednokrotnego zapisu (typu R)
  - zmiana postaci nośnika (krystaliczna, amorficzna) przez kontrolowanie laserem procesu jego podgrzewania i chłodzenia – dyski do wielokrotnego zapisu (typu RW, RAM)

Technologie dysków optycznych

Technologia	Długość fali lasera [nm]
CD	780
DVD	650
Blu-ray	405

Organizacja dysku CD

- Jedna spiralna ścieżka fizyczna podzielona na ramki i sektory, zaczynająca się przy środku dysku
- Ścieżki logiczne (ang. *track*) odpowiadające utworom muzycznym na płycie audio
- System plików: ISO 9660, RockRidge (Unix), Joliet (Windows), HFS (MacOS), HPFS (OS/2)
- Kodowanie dźwięku: Linear PCM

Organizacja dysku DVD

- Jedna lub dwie strony
- Jedna lub dwie warstwy na stronie
- Ścieżka w warstwie zewnętrznej zaczynająca się przy środku dysku
- Ścieżka w warstwie wewnętrznej zaczynająca się przy krawędzi dysku (tryb OTP) lub przy środku dysku (tryb PTP)
- System plików: UDF
- Kodowanie dźwięku: Linear PCM, Dolby Digital, DTS Digital Surround
- Kodowanie obrazu: MPEG-2
- Konkurencyjne technologie dysków zapisywalnych: DVD−R, DVD−RW, DVD+R, DVD+RW

Organizacja dysku Blu-ray

- Jedna lub dwie warstwy
- System plików: UDF
- Kodowanie dźwięku: Linear PCM, Dolby Digital, Dolby Digital Plus, Dolby TrueHD, DTS Digital Surround, DTS-HD
- Kodowanie obrazu: MPEG-2, MPEG-4

Dyski magnetyczno-optyczne

- Odczyt za pomocą lasera wykorzystuje zjawisko Kerra:
  - skręcenie płaszczyzny polaryzacji światła przy odbiciu od materiału ferromagnetycznego zależy od kierunku namagnesowania.
- Zapis wykorzystuje przejście między ferromagnetykiem a paramagnetykiem w temperaturze Curie:
  - laser podgrzewa punktowo warstwę magnetyczną,
  - głowica magnetyczna dokonuje przemagnesowania domen magnetycznych,
  - obniżenie temperatury utrwala zapis.
- Organizacja jak na dyskach twardych

Metody zapisu, zapisywania i odczytywania

- CAV (constant angular velocity) – stała prędkość kątowa
- CLV (constant linear velocity) – stała prędkość liniowa
- ZCLV (zoned CLV) – dysk podzielony jest na strefy o stałych prędkościach liniowych.
- CAA (constant angular acceleration) – odmiana CLV, w której prędkość kątowa zmienia się krokowo ze stałym przyspieszeniem (opóźnieniem).

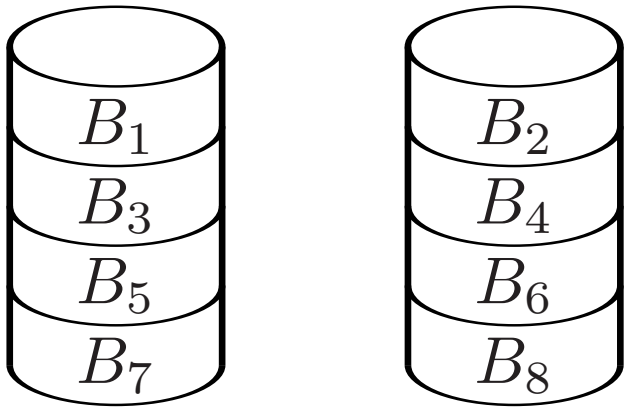
Taśmy magnetyczne

- Archiwizacja danych
- Dostęp sekwencyjny – długi średni czas dostępu
- Podłoże polimerowe pokryte warstwą magnetyczną
- Kasety jedno- lub dwuszpulowe
- Zapis
  - liniowy lub serpentynowy – wiele równoległych głowic, ścieżki równoległe do krawędzi taśmy
  - helikalny – wirujący zespół głowic, ścieżki ukośne

RAID

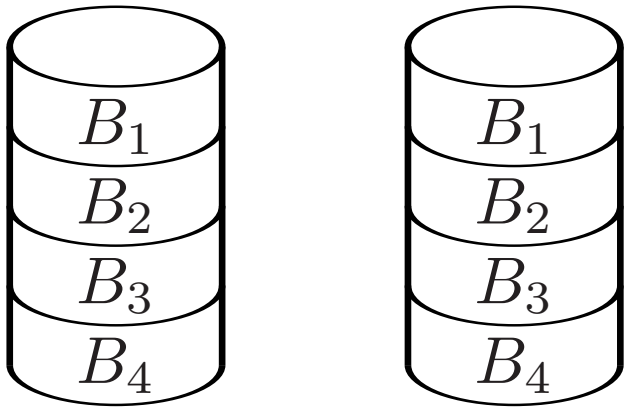
- Redundant Array of Independent Disks
- D. A. Patterson, G. A. Gibson, R. H. Katz, A case for redundant array of inexpensive disks, SIGMOD, 1988, 109–116.
  - oryginalne poziomy 1 do 5
- Implementacja
  - sprzętowa
  - programowa
  - hybrydowa

RAID 0



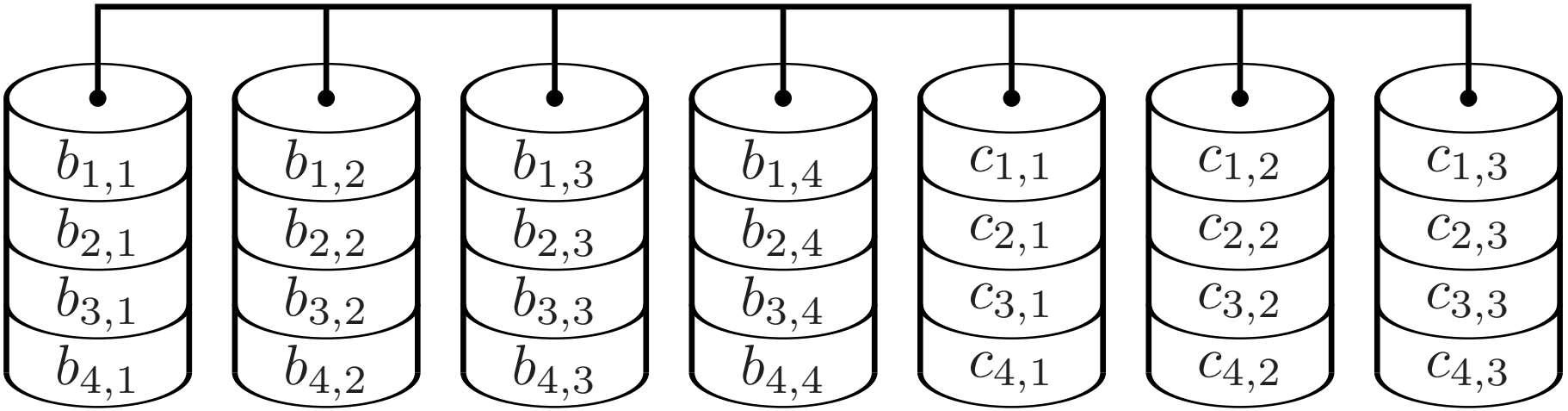
B_i – blok danych

RAID 1



B_i – blok danych

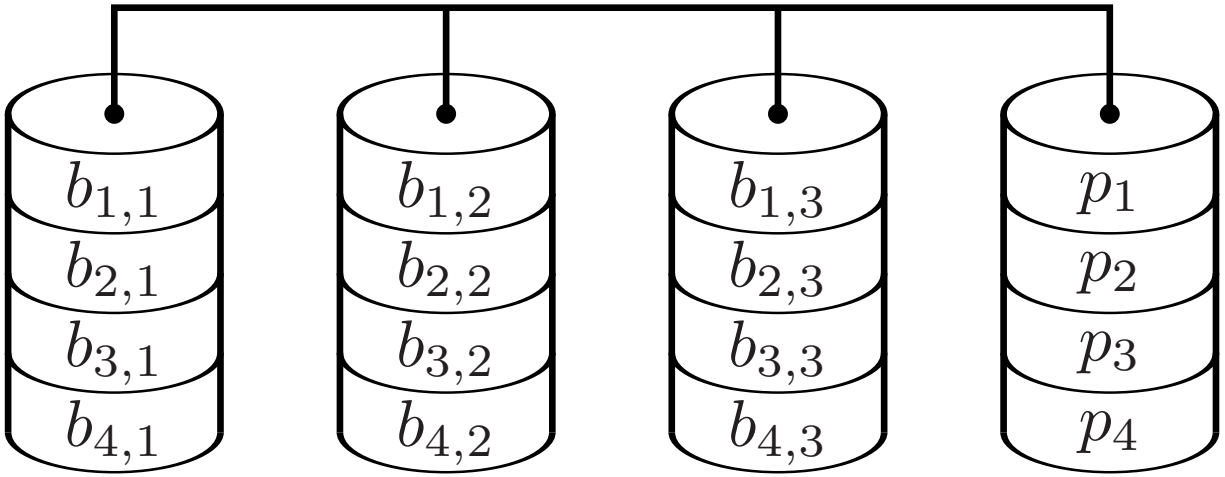
RAID 2



$b_{i,j}$ – bit danych, $c_{i,k}$ – bit kontrolny kodu Hamminga

$$c_{i,1} = b_{i,1} \oplus b_{i,2} \oplus b_{i,4}$$
$$c_{i,2} = b_{i,1} \oplus b_{i,3} \oplus b_{i,4}$$
$$c_{i,3} = b_{i,2} \oplus b_{i,3} \oplus b_{i,4}$$

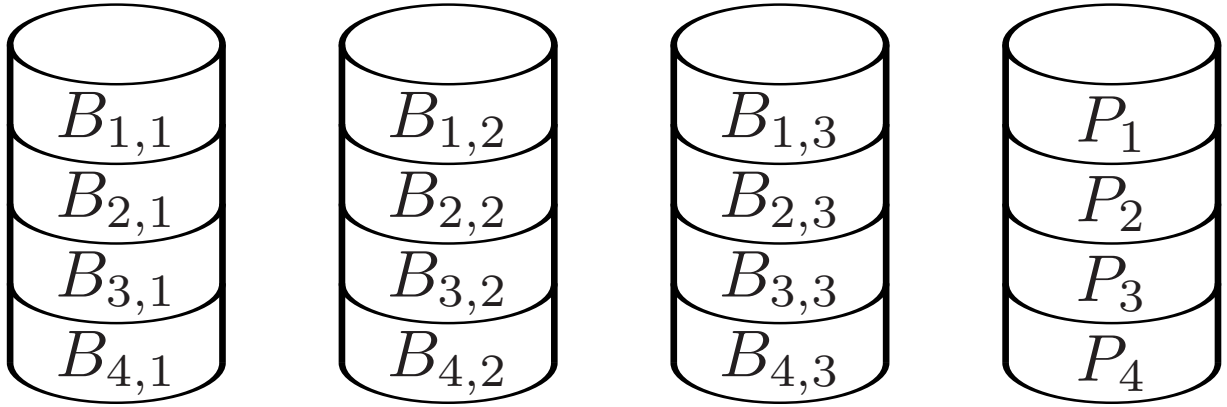
RAID 3



$b_{i,j}$ – bit danych, p_i – bit parzystości

$$p_i = b_{i,1} \oplus b_{i,2} \oplus b_{i,3}$$

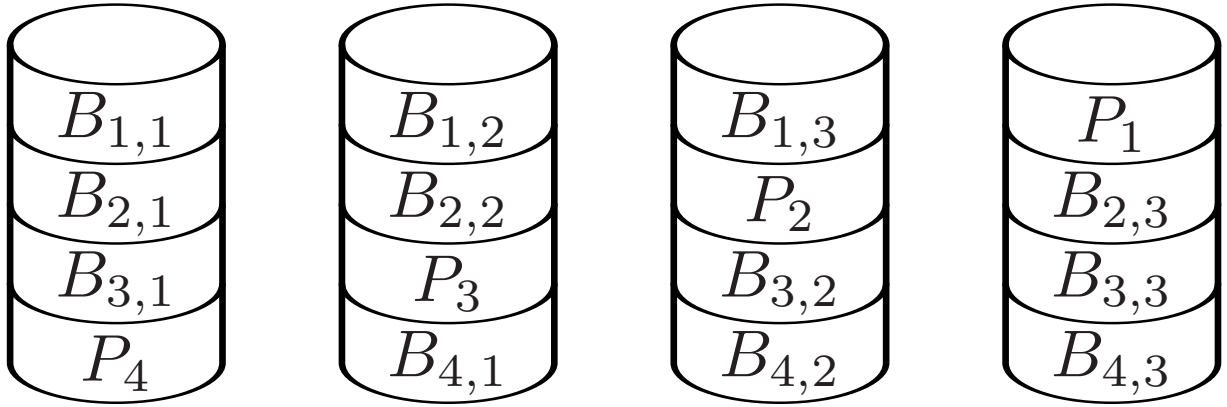
RAID 4



$B_{i,j}$ – blok danych, P_i – blok parzystości

$$P_i = B_{i,1} \oplus B_{i,2} \oplus B_{i,3}$$

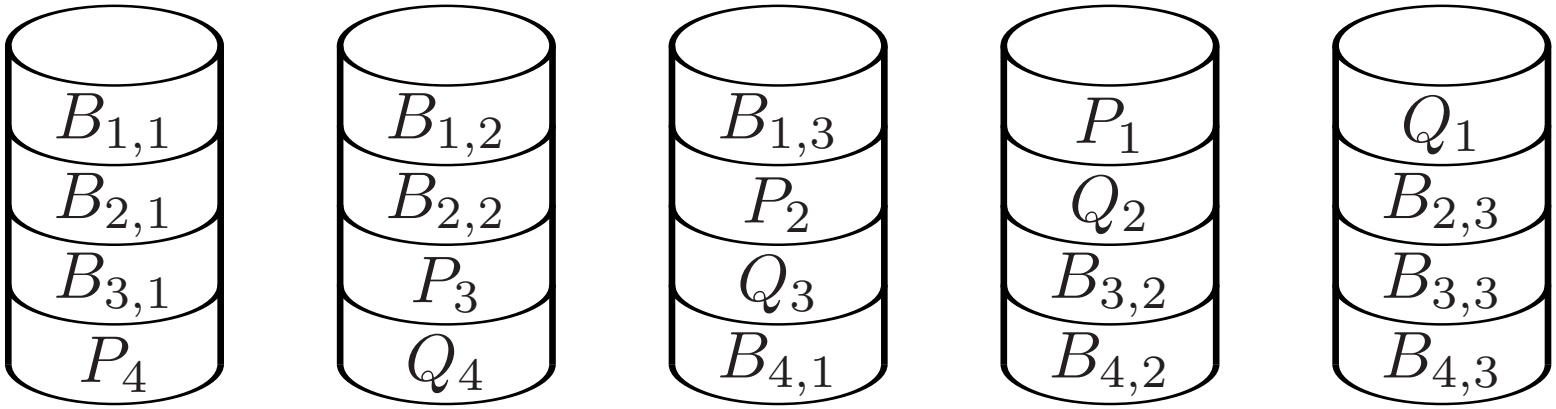
RAID 5



$B_{i,j}$ – blok danych, P_i – blok parzystości

$$P_i = B_{i,1} \oplus B_{i,2} \oplus B_{i,3}$$

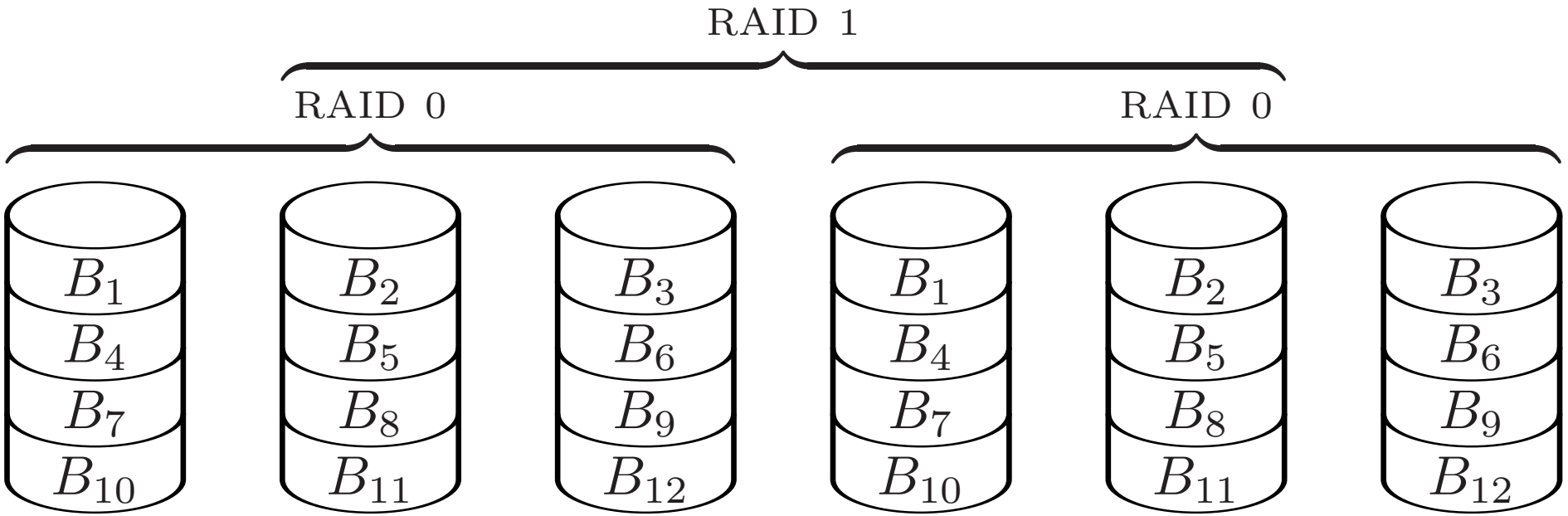
RAID 6



$B_{i,j}$ – blok danych

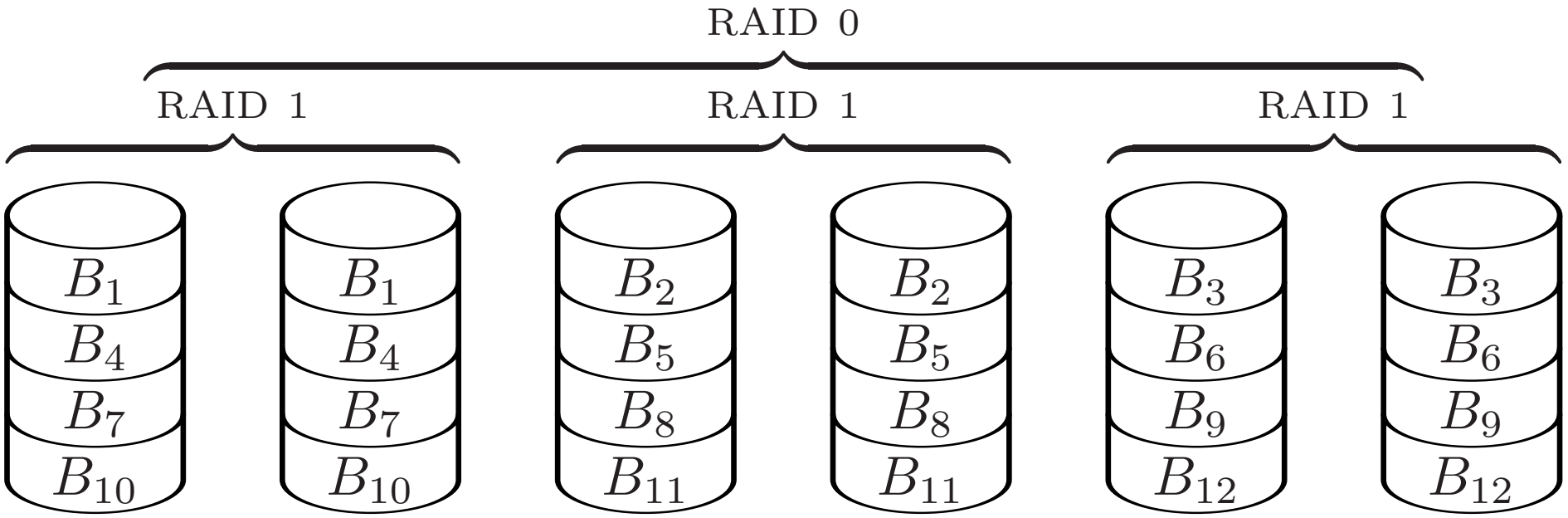
P_i, Q_i – bloki korekcyjne kodu Reeda–Solomona

RAID 0+1



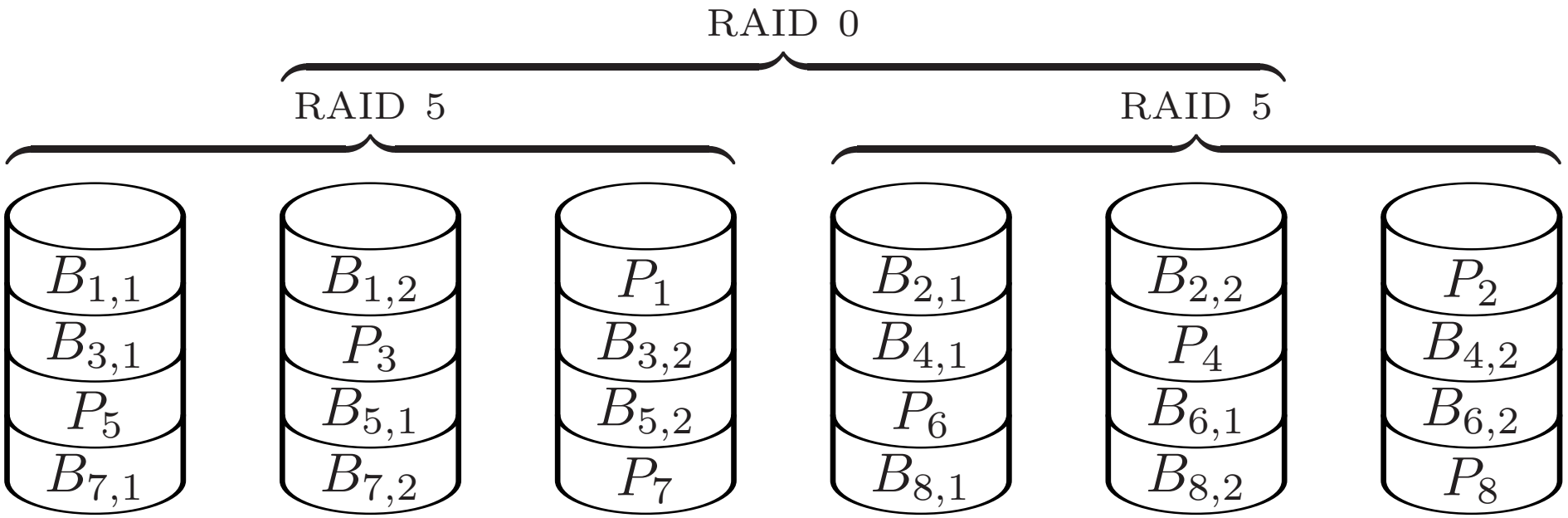
B_i – blok danych

RAID 1+0



B_i – blok danych

RAID 5+0



$B_{i,j}$ – blok danych, P_i – blok parzystości

$$P_i = B_{i,1} \oplus B_{i,2}$$

Pamięci przyszłości

- Czas dostępu jak w SRAM
- Niepotrzebne wielopoziomowe pamięci podręczne
- Gęstość jak DRAM
- Cena za bit jak w DRAM
- Nieulotne
- Praktycznie nieograniczona liczba zapisów
- Ujednolicenie pamięci operacyjnej i masowej
- MRAM?