

Pamięć

Podstawowe własności komputerowych systemów pamięciowych:

- **Położenie:** procesor, wewnętrzna (główna), zewnętrzna (pomocnicza);
- **Pojemność:** rozmiar słowa, liczba słów;
- **Jednostka transferu:** słowo, blok
- **Sposób dostępu:** sekwencyjny, bezpośredni, swobodny, skojarzeniowy
- **Wydajność:** czas dostępu, czas cyklu, szybkość transferu
- **Rodzaj fizyczny:** półprzewodnikowa, magnetyczna
- **Własności fizyczne:** ulotna/nieulotna, wymazywalna/niewymazywalna

Pamięć

Pamięć podręczna (*cache memory*) – szybka pamięć półprzewodnikowa typu SRAM (*static random access memory*), tymczasowo przechowująca dane, względnie niewielkiej pojemności, zwykle zintegrowana w strukturze jednostki centralnej.

Pamięć główna (*main memory*) – pamięć półprzewodnikowa typu DRAM (*dynamic random access memory*) najczęściej synchroniczna SDRAM lub DDR (*double data rate*) SDRAM, tymczasowo przechowująca dane, zwykle pobierane z pamięci masowej.

Pamięć masowa (*mass storage*) – pamięć trwała dużej pojemności, z wykorzystaniem różnych typów nośników, np. dysk twardy, dysk optyczny, taśma magnetyczna.

Pamięć wirtualna (*virtual memory*) – pamięć widziana przez procesor jako pamięć operacyjna, złożona ze stosunkowo niewielkiej pamięci głównej i dużej pamięci dyskowej.

Dostęp do pamięci

Dostęp sekwencyjny – pamięć jest zorganizowana za pomocą jednostek zwanych rekordami. Dostęp jest możliwy w określonej sekwencji liniowej. Przykład: **pamięć taśmowa**.

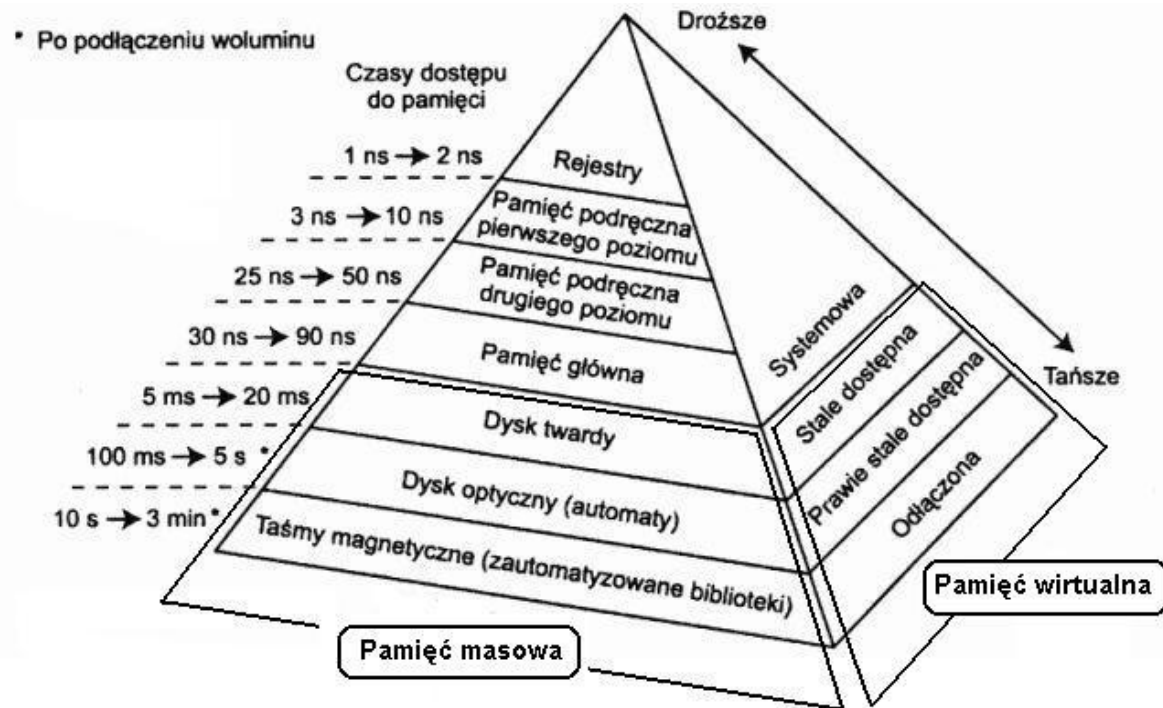
Dostęp bezpośredni – jest realizowany przez bezpośredni dostęp do najbliższego otoczenia, po którym następuje sekwencyjne poszukiwanie w celu osiągnięcia lokalizacji finalnej. Przykład: **pamięć dyskowa**.

Dostęp swobodny – każda adresowana lokalizacja w pamięci ma unikatowy, fizycznie wbudowany mechanizm adresowania.. Przykład: **pamięć główna**.

Dostęp skojarzeniowy – dostęp swobodny, który umożliwia porównywanie i badanie zgodności wybranych bitów wewnątrz słowa. Ta procedura jest prowadzona dla wszystkich słów jednocześnie. Przykład: **pamięć podręczna**.

Pamięć

Hierarchia pamięci opiera się na „oddaleniu” od procesora, gdzie odległość określa czas dostępu do danych.

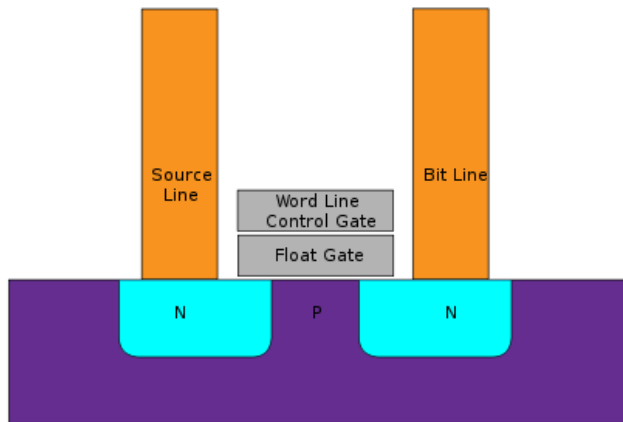


Rodzaje pamięci

Rodzaj pamięci	Kategoria	Wymazywanie	Sposób zapisu	Ulotność
Pamięć o dostępie swobodnym (RAM)	odczyt-zapis	elektryczne, na poziomie bajta	elektryczny	ulotna
Pamięć stała (ROM)	tylko odczyt	niemożliwe	maski	nieulotne
Programowalna pamięć stała (PROM)			elektryczny	
Wymazywalna PROM (EPROM)				
Pamięć błyskawiczna	głównie odczyt	światłem UV, na poziomie układu	elektryczny	nieulotne
Elektrycznie wymazywalna PROM (EEPROM)		elektryczne, na poziomie bajta		

Pamięci FLASH

- Pamięci Flash są popularną odmianą pamięci nieulotnych EEPROM
- kart pamięci lub dyski SSD (Solid State Data) o pojemnościach dochodzących do terabajta
- Do największych zalet należą - wysoka wydajność, niski pobór energii, niewielkie gabaryty i masa, wysoka odporność na uszkodzenia.



- Zasada działania opiera się na przechowywaniu informacji w tranzystorach polowych MOSFET.
- Jedna komórka pamięci Flash to tranzystor MOSFET, który pomiędzy podłożem, a bramką sterującą (Control Gate - CG) posiada odizolowaną bramkę pływającą (Floating Gate - FG).
- Raz umieszczone elektrony na FG mogą pozostać tam przez wiele lat pamiętając zaprogramowany stan.

Rodzaje pamięci RAM

RAM dynamiczna

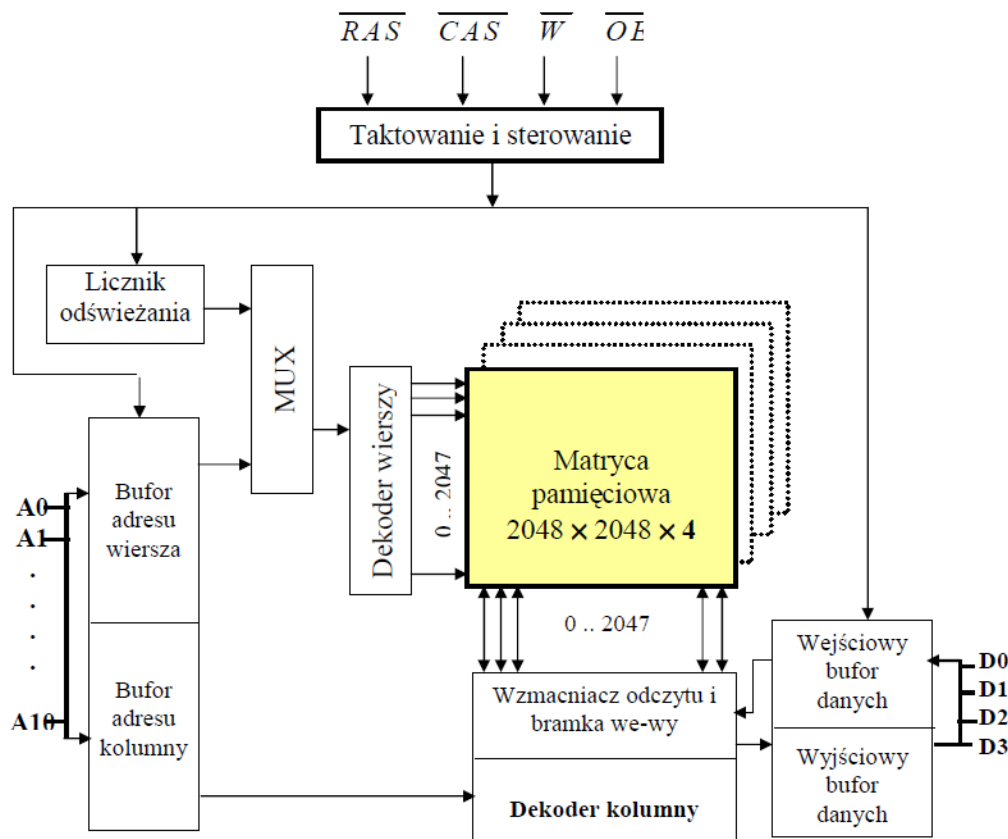
Komórki pamięci przechowują ładunek elektryczny. Obecność lub brak ładunku są interpretowane jako 0 lub 1. Okresowe odświeżanie ładunku w celu zachowania danych.

RAM statyczna

Dane są przechowywane za pomocą przerzutnikowych konfiguracji bramek logicznych. Statyczne pamięci RAM zachowują dane tak długo, jak długo są zasilane.

Blok pamięci RAM

Organizacja pamięci DRAM o pojemności 16 Mbit ($4 \text{ M} \times 4$)



Pojemność 16 Mbit

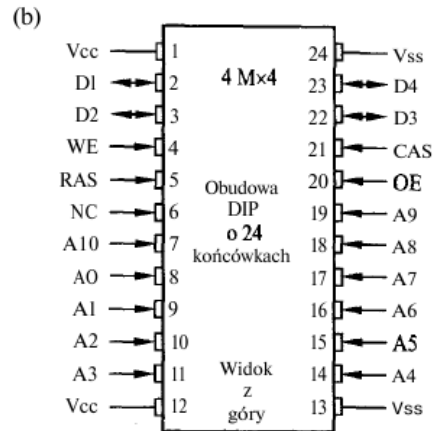
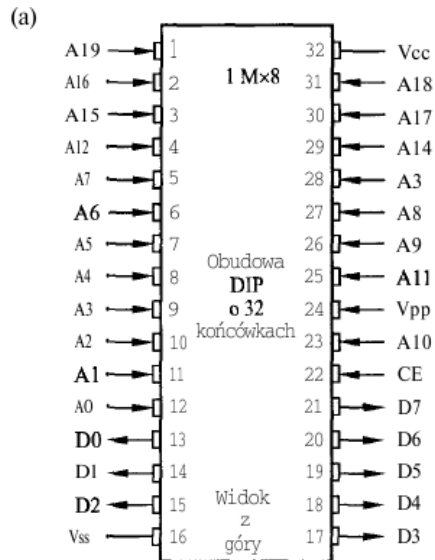
$2^{11} = 2048$ - 11 linii potrzebnych do wybrania jednego z 2048 wierszy (towarzyszy im sygnał wyboru wiersza RAS)

sygnały potrzebne do wyboru kolumny też doprowadzane są tymi samymi liniami, towarzyszy im sygnał CAS

w czasie gdy komórki są odświeżane, to nie można czytać/wpisywać danych

WE - zezwolenie zapisu
OE – zezwolenie na odczyt

Bloki pamięci



Vcc – napięcie zasilania układu

Vss – uziemienie układu

Vpp – napięcie programowania układu (operacja zapisu)

a) pamięć EPROM 8Mbit

b) Pamięć DRAM 16 Mbit

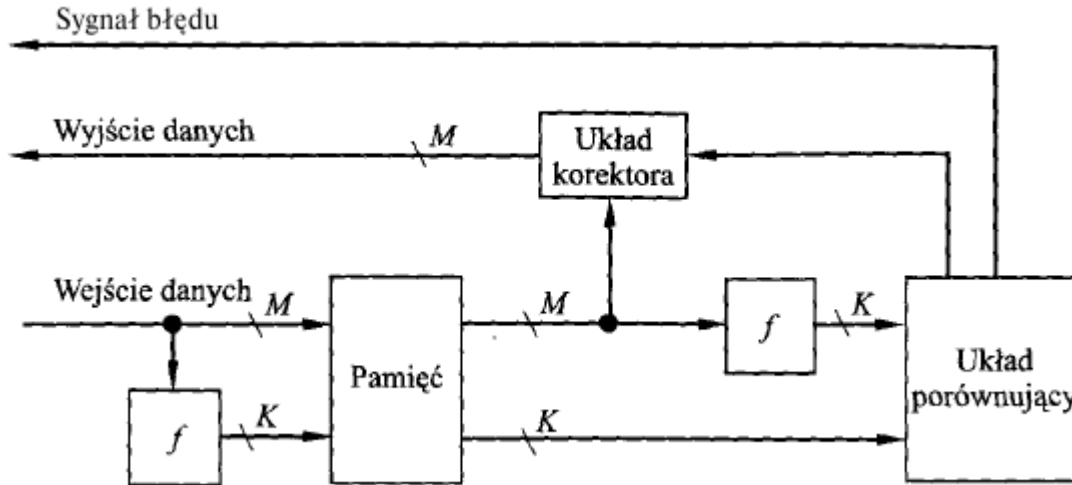
Korekcja błędów

W systemach pamięci półprzewodnikowych występują błędy:
stałe i przypadkowe

Błędy stałe – permanentny defekt fizyczny powodujący, że uszkodzona komórka lub komórki pamięci pozostają stale w stanie 0 lub 1 lub błędnie przeskakują pomiędzy 0 i 1.

Błędy przypadkowe – wywołane losowymi nieniszczącym zjawiskiem, które zmienia zawartość 1 lub wielu komórek. Np. jonizacja wywołane przez cząstkę α

Układ korekcyjny



Jeśli ma być zapisane M bitowe słowo danych a kod korekcyjny ma K bitów, to musimy zapisać $M+K$ bitów

Porównywanie prowadzi do jednego z trzech wyników:

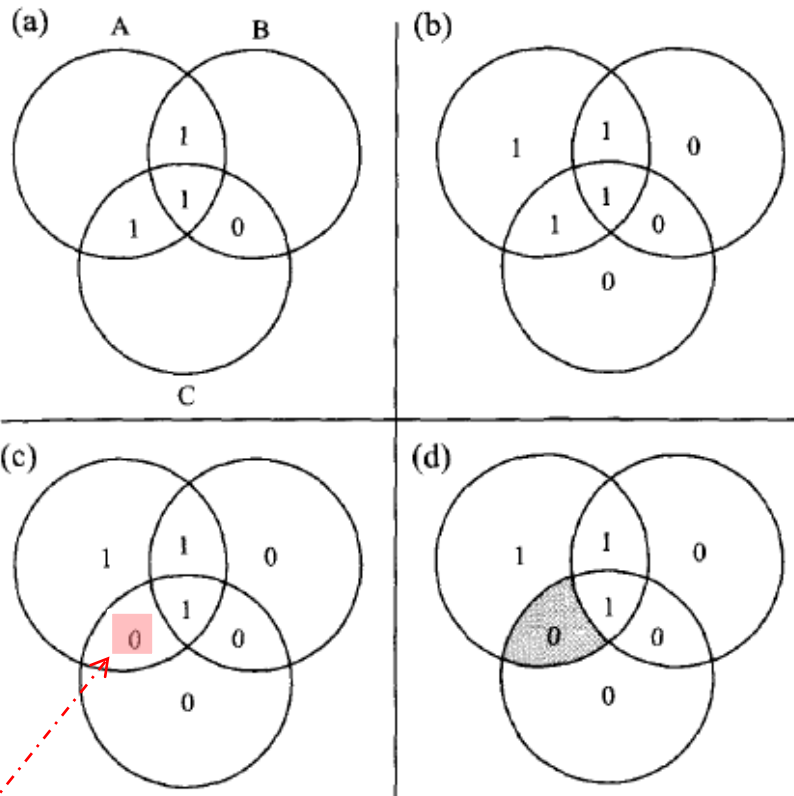
- nie wykryto żadnych błędów. Pobierane bity są wysyłane.
- wykryto błąd, którego korekta jest możliwa. Korektor tworzy poprawiony zestaw M bitów do wysłania.
- korekta jest niemożliwa. Jest generowany sygnał błędu.

Kody korekcyjne

kod Hamminga

słowo 4-bitowe

Wykres Venna



tutaj powstał błąd

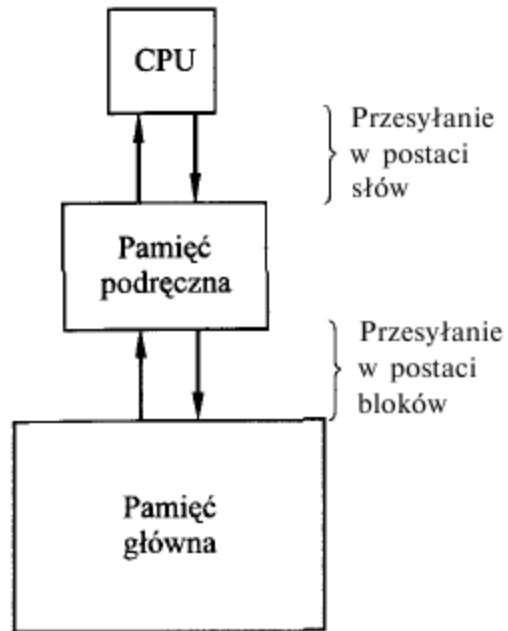
- przypisujemy 4 bity danych przedziałom wewnętrznym
- pozostałe przedziały wypełniamy bitami parzystości. Bity parzystości wybieramy tak, że całkowita liczba 1 w każdym okręgu jest parzysta.
- powstaje błąd
- sprawdzamy bity parzystości. Jest sprzeczność w okręgach A i C. Dobrze w B. Znamy pozycję złego bitu możemy go poprawić

Kody korekcyjne

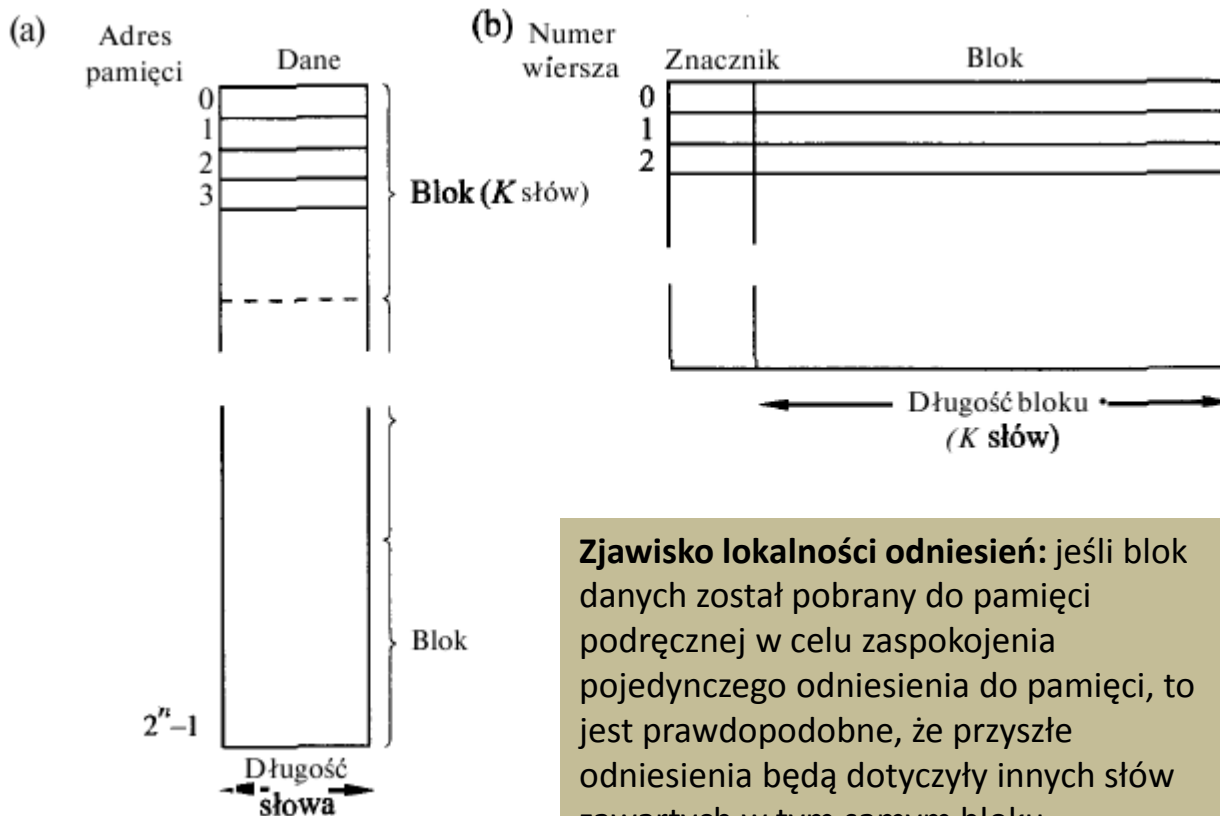
Wzrost długości słowa po uwzględnieniu bitów korekcyjnych

Liczba bitów danych	Poprawienie błędu pojedynczego		Poprawienie błędu pojedynczego , wykrycie błędu podwójnego	
	Bity kontrolne	% wzrostu	Bity kontrolne	% wzrostu
8	4	50	5	62.5
16	5	31.25	6	37.5
32	6	18.75	7	21.87
64	7	10.94	8	12.5
128	8	6.25	9	7.03
256	9	3.52	10	3.91

Pamięć podręczna



Pamięć podręczna



Zjawisko lokalności odniesień: jeśli blok danych został pobrany do pamięci podręcznej w celu zaspokojenia pojedynczego odniesienia do pamięci, to jest prawdopodobne, że przyszłe odniesienia będą dotyczyły innych słów zawartych w tym samym bloku

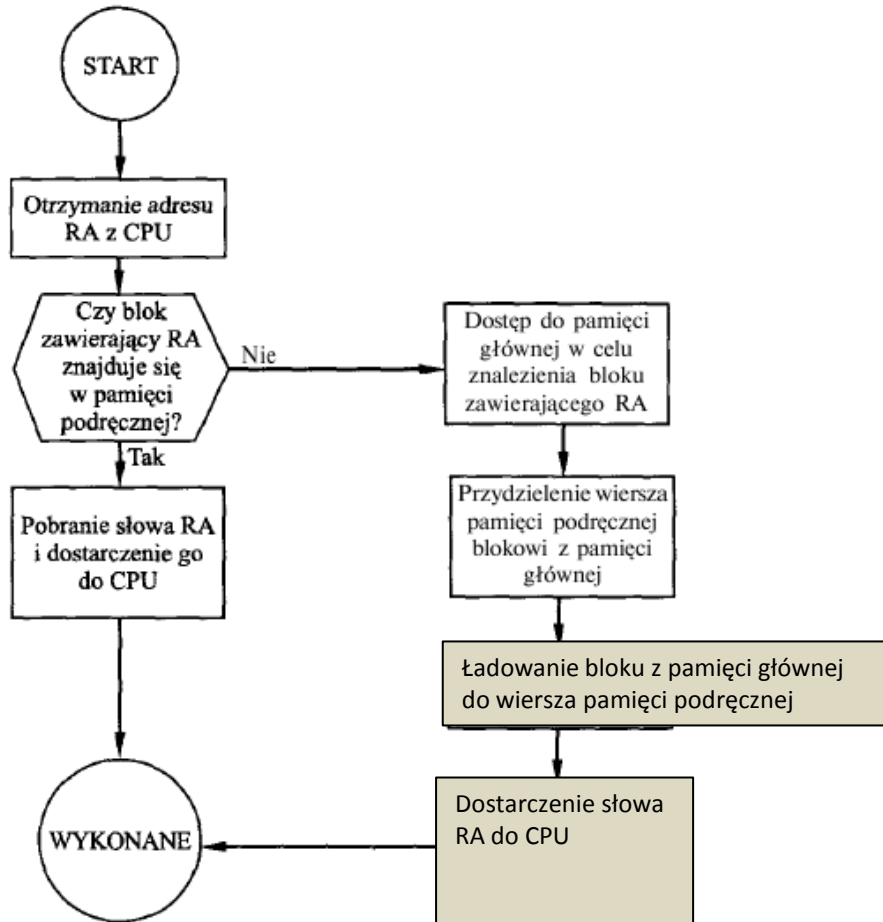
2^n - liczba adresowalnych słów

$M = 2^n / K$ liczba bloków

Pamięć podręczna zawiera C wierszy zawierających K słów: $C \ll M$

Znacznik zawiera informację o tym, który blok przechowywany jest w danym wierszu

Operacja odczytu z pamięci podręcznej



Pamięci podręczne

Funkcje odwzorowania:

Bezpośrednia, skojarzeniowa, sekwencyjno-skojarzeniowa

Rozważmy przykład pamięci podręcznej o następujących parametrach:

- 64 kB = 65536 B
- dane są przenoszone pomiędzy pamięcią główną a pamięcią podręczną w blokach po cztery bajty
- pamięć podręczna ma więc 16 k = 16 348 wierszy po 4 bajty każdy
- pamięć główna ma 16 MB adresowalnych za pomocą 24 bitowych adresów ($2^{24} = 16 \text{ M}$).
- można ją traktować jako 4 M bloków po 4 bajty każdy

Odwzorowanie bezpośrednie

Każdy adres pamięci głównej można widzieć jako składający się z trzech pól:

- najmniej znaczące bity określają położenie bajtu w bloku pamięci
- pozostałe s bitów określa jeden z 2^s bloków pamięci głównej.
- układy logiczne pamięci podręcznej interpretują te s bitów jako:

znacznik - $s-r$ bitów

pole linii - r bitów

- to ostatnie pole identyfikuje jeden z $m = 2^r$ wierszy pamięci podręcznej

znacznik	wiersz	słowo
$s-r$	r	$n-s$

Odwzorowanie bezpośrednie

wiersz pamięci podręcznej	przypisane bloki pamięci głównej
0	0, m, ..., $2^s - m$
1	1, m+1, ..., $2^s - m + 1$
....
m-1	m - 1, 2m - 1, ..., $2^s - 1$

Odwzorowanie bezpośrednie

- pamięć podręczna: 16 k = 65 384 wierszy po 4 bajty każdy - 64KB (2^{14})
- pamięć główna ma 16 MB adresowalnych za pomocą 24 bitowych adresów ($2^{24} - 1 = \text{FF'FFFF}$).
- można ją traktować jako 4 M = 4 194 304 bloków po 4 bajty każdy

znacznik	wiersz	słowo
s-r = 8	r = 14	2

wiersz pamięci podręcznej	przypisane bloki pamięci głównej
0	00'0000, 01'0000, ..., FF'0000
1	00'0001, 01'0001, ..., FF'0001
...	...
3FFF (=65383)	00'FFFC, 01'FFFC, ..., FF'FFFC

żadne dwa bloki odwzorowania w tym samym wierszu nie mają takiego samego znacznika

FF'FFFC = 16 777 212 – pierwszy bajt ostatniego bloku

FF'FFFF = 16 777 215 – ostatni bajt ostatniego bloku

Pojemność i liczba pamięci podręcznych

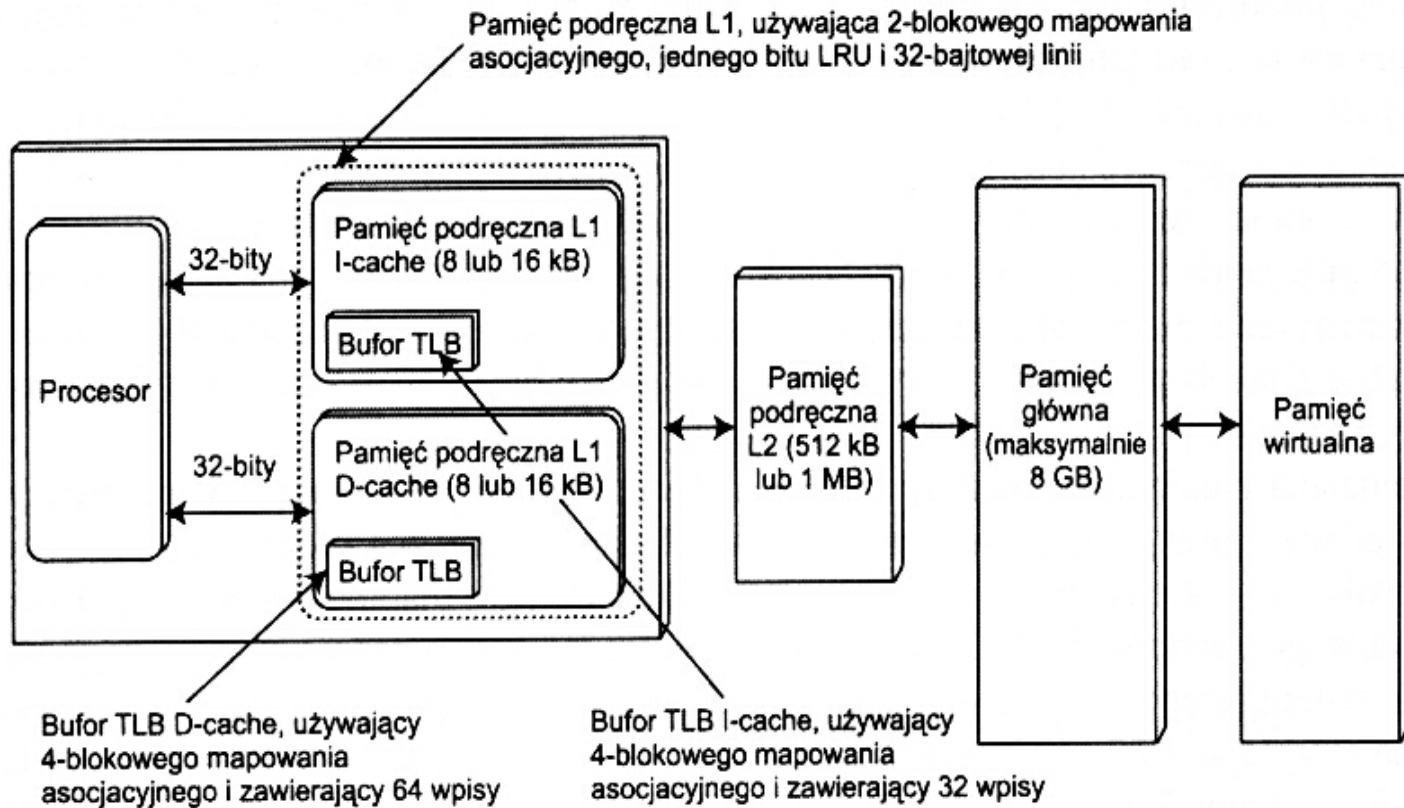
Pojemność : optymalna wielkość 1 k – 512 k słów

Liczba:

- zastosowanie pamięci podręcznej wewnątrz procesora L1 (**on-chip cache**) pozwala na ograniczenie komunikacji procesora z pamięcią główną oraz na skrócenie czasu wykonywania operacji;
- pamięć podręczna zewnętrzna L2

Jednolita a podzielona pamięć podręczna: obecnie stało się powszechne dzielenie pamięci podręcznej na dwie. Jedną przeznaczoną na rozkazy , druga zaś na dane.

Pamięć



Hierarchia pamięci procesora Pentium

Wielkości pamięci podręcznych

Processor	Type	Year of Introduction	L1 cache ^a	L2 cache	L3 cache
IBM 360/85	Mainframe	1968	16 to 32 KB	—	—
PDP-11/70	Minicomputer	1975	1 KB	—	—
VAX 11/780	Minicomputer	1978	16 KB	—	—
IBM 3033	Mainframe	1978	64 KB	—	—
IBM 3090	Mainframe	1985	128 to 256 KB	—	—
Intel 80486	PC	1989	8 KB	—	—
Pentium	PC	1993	8 KB/8 KB	256 to 512 KB	—
PowerPC 601	PC	1993	32 KB	—	—
PowerPC 620	PC	1996	32 KB/32 KB	—	—
PowerPC G4	PC/server	1999	32 KB/32 KB	256 KB to 1 MB	2 MB
IBM S/390 G4	Mainframe	1997	32 KB	256 KB	2 MB
IBM S/390 G6	Mainframe	1999	256 KB	8 MB	—
Pentium 4	PC/server	2000	8 KB/8 KB	256 KB	—
IBM SP	High-end server/ supercomputer	2000	64 KB/32 KB	8 MB	—
CRAY MTA ^b	Supercomputer	2000	8 KB	2 MB	—
Itanium	PC/server	2001	16 KB/16 KB	96 KB	4 MB
SGI Origin 2001	High-end server	2001	32 KB/32 KB	4 MB	—
Itanium 2	PC/server	2002	32 KB	256 KB	6 MB
IBM POWER5	High-end server	2003	64 KB	1.9 MB	36 MB
CRAY XD-1	Supercomputer	2004	64 KB/64 KB	1MB	— ²³