|  |  |
| --- | --- |
| ***ARCHITEKTURA SYSTEMU KOMPUTEROWEGO WG VON NEUMANA*** | |
| **POJĘCIA:** |
| **System mikroprocesorowy** |
| **Szyny (magistrale) i ich przeznaczenie** |
| **Architektura mikroprocesora (Schemat blokowy procesora)** |

### System mikroprocesorowy

**...**

**URZ.**

**ZEWN**

**PROCESOR**

**PAMIĘC**

**UKŁADY  
WE/WY**

**ZEGAR**

**SZYNA DANYCH**

**SZYNA ADRESOWA**

**SZYNA STEROWANIA**

**UŻYTKOWNIK**

**DANE  
PROGRAMY  
WYNIKI**

**MONITORY, DRUKARKI,   
CZYTNIKI,...**

**Centralna jednostka   
przetwarzająca (CPU)**

**Liczba linii wyznacza możliwości adresowania np. liczba linii 20 to można zaadresować 220 komórek**

**Liczba linii określa długość słowa procesora (8, 16, 32, .... (+bit detekcji i korekcji błędów)**

**Kilkanaście (kilkadziesiąt) linii**

**Uniwersalny układ przetwarzający informację i sterujący pracą pozostałych elementów systemu**

### Architektura komputera wg von Neumana

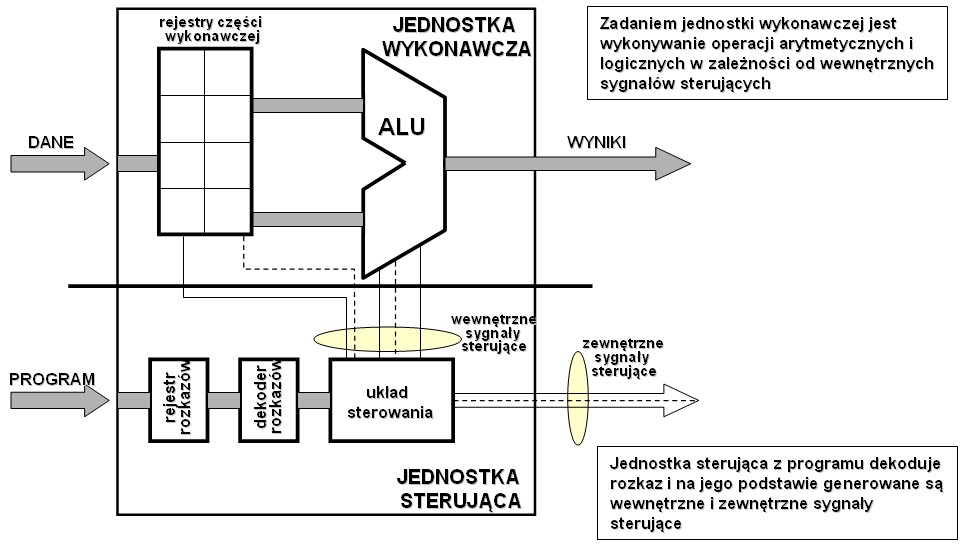
* Architektura klasycznego komputera (wg von Neumana) opiera się na założeniach:
  + program wykonywany przez procesor wraz z danymi jest umieszczony w pamięci;
  + kolejność wykonywanych rozkazów zależy od ich umieszczenia w programie (w kolejnych komórkach pamięci), a zmiana tej zasady może być wykonana tylko:
    - przez program (rozkaz skoku);
    - przez system operacyjny np. wykrycie błędu;
    - przez operatora np. *reset*
  + procesor odczytuje kolejne rozkazy z pamięci wysyłając odpowiednie adresy

W architekturze von Neumana komputer składa się z czterech głównych komponentów:

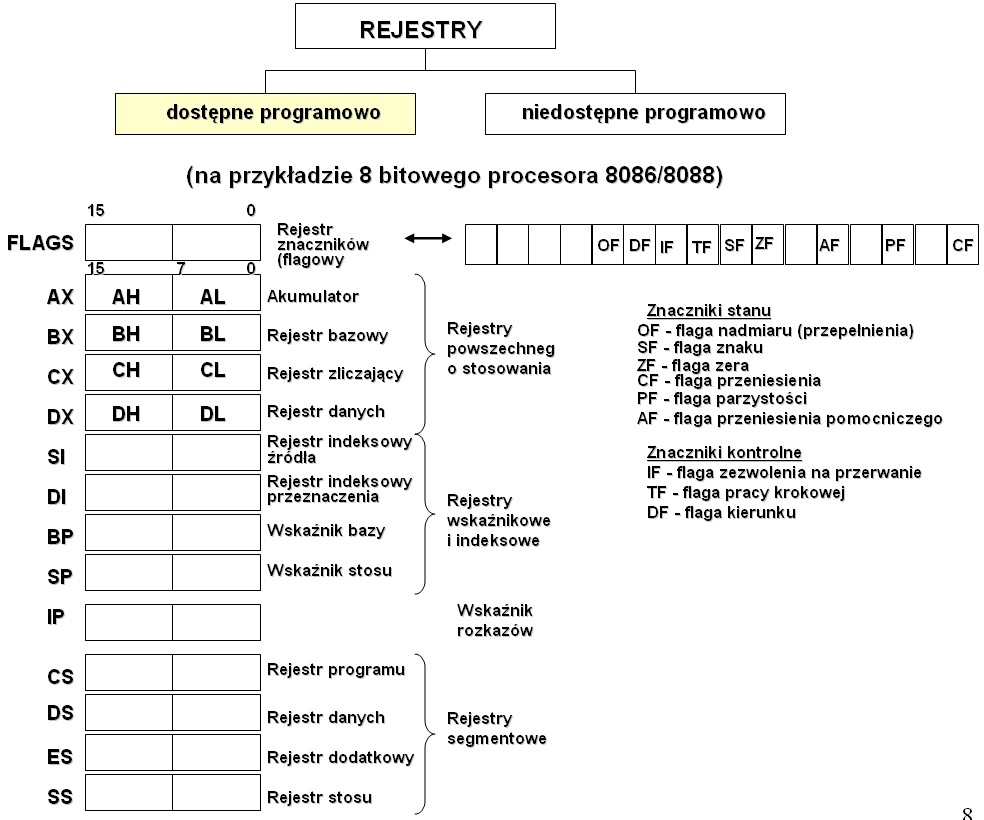
* pamięci komputerowej przechowującej dane programu oraz instrukcje programu; każda komórka pamięci ma unikalny identyfikator nazywany jej adresem
* jednostki kontrolnej odpowiedzialnej za pobieranie danych i instrukcji z pamięci oraz ich sekwencyjne przetwarzanie
* jednostki arytmetyczno-logicznej odpowiedzialnej za wykonywanie podstawowych operacji arytmetycznych.
* urządzeń wejścia/wyjścia służących do interakcji z operatorem

### ARCHITEKTURA MIKROPROCESORA

* **Schemat blokowy mikroprocesora**

****

* **Rejestry mikroprocesora**

****

**Rejestr znaczników (flagowy) - rejestr przeznaczony do przechowywania dodatkowych cech wyniku operacji (np. znak, przekroczenie zakresu, znak parzystości, ...) w postaci tzw. flagi (znacznika). Każdy znacznik jest bitem w rejestrze, który wskazuje czy określony stan wystąpił. Znaczniki mogą być wykorzystywane przez procesor lub programistę na dwa sposoby:**

* + **ustawienie znacznika dla zapamiętania określonego stanu po wykonaniu rozkazu**
  + **testowanie znacznika celem umożliwienia decyzji o sposobie dalszego postępowania (przetwarzania danych)**

**Znaczniki stanu:**

* + **CF - flaga przeniesienia lub pożyczki, ustawiana przy przekroczeniu zakresu długości słowa, w którym zapisywany jest wynik (1 wystąpiło przeniesienie)**
  + **PF - flaga parzystości (1) wskazuje na parzystą lub nieparzystą liczbę bitów o wartości jedynki**
  + **ZF - flaga zera - sygnalizująca, że wynikiem ostatnio wykonywanej operacji jest zero**
  + **PF - flaga przeniesienia pomocniczego jest wykorzystywana przy działaniach na liczbach w kodzie BCD. Ustawiana jest gdy następuje przeniesienie lub pożyczka z najstarszego bitu pierwszej tetrady wyniku**
  + **SF - flaga znaku, ustawiana gdy najstarszy bit wyniku jest równy 1 (operacje liczb ze znakiem - kod U2)**
  + **OF - flaga przepełnienia, sygnalizująca przekroczenie zakresu dla operacji arytmetycznych (operacje liczb ze znakiem - kod U2)**

**Znaczniki kontrolne:**

* + **TF- znacznik pracy krokowej - ustawienie 1 pozwala na wykonanie przez procesor po wykonaniu każdego rozkazu przerwania i przejście do specjalnych procedur obsługi**
  + **IF- znacznik zezwolenia na przerwanie - ustawienie 1 powoduje odblokowanie systemu przerwań w procesorze (0 - procesor ignoruje przerwania)**
  + **DF - znacznik kierunku - umożliwia realizację przetwarzania łańcuchów (ciągu słów) przy rosnących adresach (1) lub malejących adresach (0)**

**Rejestry ogólnego przeznaczenia - rejestry przeznaczone do przechowywania dowolnych danych i wykonywania operacji (arytmetycznych i logicznych), ale jednocześnie spełniające pewne funkcje specjalne. Szczególną funkcję pełni akumulator**

**Akumulator - rejestr przeznaczony do przechowywania jednego z operandów (argumentów) wykonywanej operacji oraz wyniku wykonywanej operacji (czasami wynik może być umieszczany w innym rejestrze)**

**Rejestry segmentowe - są rejestrami wykorzystywanymi do adresowania pamięci operacyjnej. Rejestry te zawierają adresy początkowe segmentów**

* + **CS rejestr segmentowy programu wskazuje segment programu, z którego aktualnie są pobierane kolejne rozkazy do wykonania,**
  + **DS rejestr segmentowy stosu wskazuje segment, w którym zapamiętane są zmienne używane w programie,**
  + **ES rejestr segmentowy dodatkowy wskazuje dodatkowy segment danych,**
  + **SS rejestr segmentowy stosu wskazuje segment pamięci, w którym zdefiniowany jest stos.**

**Rejestry wskaźnikowe i indeksowe - posiadają dwa rejestry wskaźnikowe i dwa rejestry indeksowe.   
Rejestry wskaźnikowe (SP, BP) są stosowane do adresowania danych w obrębie wydzielonego obszaru pamięci (stosu).   
Rejestry indeksowe (SI, DI) służą do adresowania danych w obszarze pamięci zwanym segmentem danych.   
Wszystkie rejestry mogą być wykorzystywane jako argumenty większości rozkazów arytmetycznych i logicznych.**

Pytania:

Jaka jest szerokość szyny danych w płycie GX-110?

Jaka jest szerokość szyny adresowej w płycie GX-110?

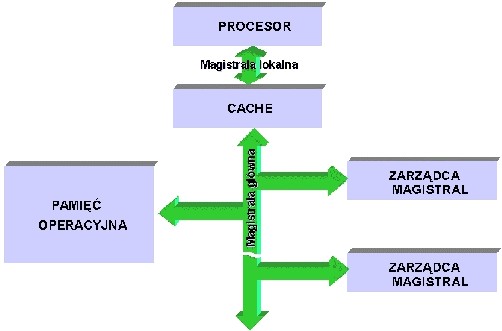
Ile komórek 8-bitowych pamięci może zaadresować magistrala o szerokości 32-bity?

|  |
| --- |
| ***PAMIĘĆ PODRĘCZNA (CACHE)*** |
| **POJĘCIA:** | |
| **Idea stosowania pamięci cache** | |
| **Rodzaje architektury pamięci cache** | |
| **Poziomy pamięci cache: L1, L2, L3, data, instruction** | |

**Pamięć podręczna** (ang. *cache*) – mechanizm, w którym część spośród danych zgromadzonych w źródłach o długim czasie dostępu i niższej przepustowości jest dodatkowo przechowywana w pamięci o lepszych parametrach. Ma to na celu poprawę szybkości dostępu do tych informacji, które przypuszczalnie będą potrzebne w najbliższej przyszłości. Współczesny procesor ma 2 albo 3 poziomy pamięci podręcznej oddzielającej go od pamięci RAM.

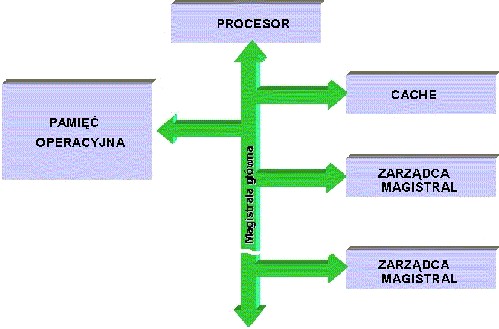
**Architektura Look-through**

Pamięć cache połączona jest magistralą lokalną z procesorem oddzielając go od reszty systemu. W przypadku trafienia główna magistrala systemu nie jest używana, co powoduje dodatkowe zwiększenie szybkości działania systemu, poprzez możliwość wykorzystania tej magistrali przez innego zarządcę (np. sterownik DMA). Żądanie dostępu do pamięci głównej tylko w przypadku chybienia. Inicjowany jest wówczas cykl magistrali głównej z koniecznymi stratami oczekiwania.



**Architektura  Look-aside**

W tej architekturze, w przeciwieństwie do architektury Look-through, procesor nie jest odizolowany od reszty układów. Dostęp do pamięci cache oraz pamięci głównej realizowany jest poprzez magistralę główną co powoduje, że nawet w przypadku trafienia magistrala jest niedostępna dla innych zarządców. Niewątpliwą jednak zaletą tej architektury jest prostsza konstrukcja, niższy koszt i nieco szybsza realizacja operacji w przypadku chybienia.



**Pamięć cache L1**

Zlokalizowana we wnętrzu procesora pamięć podręczna pierwszego poziomu przyspiesza dostęp do bloków pamięci wyższego poziomu, który stanowi zależnie od konstrukcji pamięć operacyjną lub pamięć podręczną drugiego poziomu (L-2). Z uwagi na ograniczenia rozmiarów i mocy procesora zawsze jest najmniejsza. Umieszczona jest najbliżej głównego jądra procesora i umożliwia najszybszą komunikację procesora.

**Pamięć cache L2**

Pamięć drugiego poziomu, o rozmiarze od 64 KB do 12 MB, jest wykorzystywana jako bufor pomiędzy stosunkowo wolną pamięcią RAM a jądrem procesora i pamięcią cache L1.

**Pamięć cache L3**

Pamięć podręczna procesora trzeciego poziomu jest wykorzystywana, kiedy pamięć L2 jest niewystarczająca aby pomieścić potrzebne dane. Najczęściej spotykana jest w procesorach przeznaczonych do zastosowań serwerowych. Obecność cache trzeciego poziomu pozwala na znaczącą poprawę wydajności w stosunku do procesorów o konstrukcji pamięci dwupoziomowej w wielu aplikacjach i programach. W systemach z wieloma procesorami lub rdzeniami, pamięć cache trzeciego poziomu najczęściej jest współdzielona przez wszystkie rdzenie i ma od kilku do kilkunastu megabajtów. Jej wysoka wydajność wynika z droższych i lepszych komponentów oraz bliższego fizycznego ulokowania przy procesorze.

**Podział pamięci cache na segmenty dla kodu i danych**

Pamięć podręczna L1 może być podzielona na pamięć cache dla kodu (umieszczane w niej są instrukcje) i danych (umieszczane są w niej dane): **INSTRUCTION, DATA**.

Pytania:

Dlaczego stosuje się pamięć podręczną?

Podaj główną zaletę architektury pamięci podręcznej w architekturze look-aside

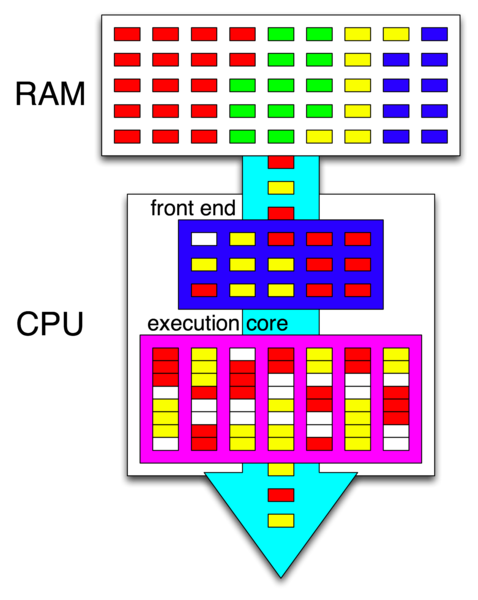
Podaj główną wadę architektury pamięci podręcznej w architekturze look-aside

Podaj główną zaletę architektury pamięci podręcznej w architekturze look- through

Podaj główną wadę architektury pamięci podręcznej w architekturze look-through

**HT - Hyper-threading**

**Hyper-threading** (nazwa oficjalna **Hyper-Threading Technology**, nazwy skrócone **HT Technology**, **HTT** lub **HT**) – jest to implementacja wielowątkowości współbieżnej (ang. *simultaneous multithreading*, SMT) opracowana przez firmę Intel i stosowana w procesorach Atom, Core i3, Core i5, Core i7, Itanium, Pentium 4 oraz Xeon.



Schemat funkcjonowania technologii Intel Hyper-Threading Technology.

**Hyper-threading** funkcjonuje poprzez duplikowanie pewnych fragmentów procesora – tych, które przechowują stany procesów (architectural state) – ale nie jest to duplikowanie głównych zasobów wykonawczych. To pozwala procesorowi wykorzystującemu Hyper-threading być widocznym dla systemu operacyjnego jako dwa „logiczne” procesory, pozwalając mu na zaplanowanie wykonania dwóch wątków lub procesów jednocześnie. Jeżeli zasoby wykonawcze nie byłyby wykorzystywane przez bieżące zadanie w procesorze bez technologii Hyper-threading, a w szczególności gdy procesor jest w stanie zatrzymania, procesor z technologią Hyper-threading potrafi użyć wolnych w tym momencie zasobów wykonawczych w celu wykonania innych zaplanowanych przez system operacyjny zadań. (Procesor może być w stanie zatrzymania ze względu na brak danych w pamięci cache **cache miss**, nietrafione przewidywanie rozgałęzień branch misprediction, lub zależności danych data dependency.)

Technologia ta jest przezroczysta dla systemu operacyjnego i uruchomionych programów. Jedynie aby skorzystać z zalet jakie oferuje Hyper-threading wymagane jest wsparcie obsługi maszyn wieloprocesorowych z SMP przez system operacyjny, ponieważ „logiczne” procesory są widziane przez system operacyjny jako dwa odrębne procesory fizyczne.

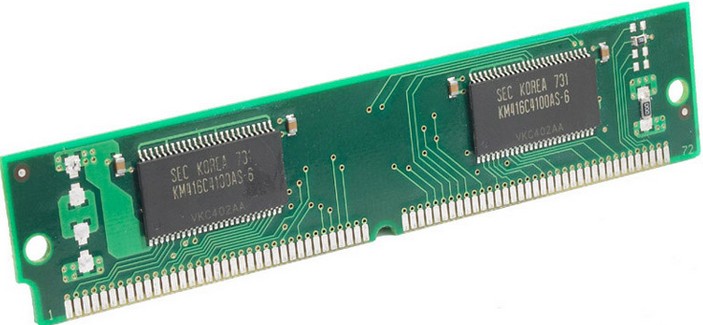
**Hyper-threading** służy zwiększeniu wydajności obliczeń prowadzonych równolegle (czyli wykonywaniu wielu zadań jednocześnie) przez mikroprocesory. Dla każdego fizycznego rdzenia procesora system operacyjny przypisuje dwa procesory wirtualne (ang. *virtual processors*), a następnie dzieli obciążenie obliczeniami między nimi jeżeli jest to możliwe. Hyper-threading wymaga nie tylko wsparcia ze strony systemu operacyjnego ale również oprogramowania specyficznie zoptymalizowanego dla obsługi tej technologii, Intel zaleca wyłączanie jej jeżeli używany jest system operacyjny bez takich optymalizacji.

|  |  |
| --- | --- |
| ***KLASYFIKACJA PAMIĘCI (MODUŁY SIMM, DIMM, RODZAJE SRAM, DRAM)*** | |
| **POJĘCIA:** |
| **Podział ze względu na budowę: SIMM, DIMM** |
| **Podział ze względu na działanie: SRAM, DRAM, SDRAM** |
| **Zasada działania i rodzaje SDRAM DDR: DDR1, DDR2, DDR3, DDR4** |

**Klasyfikacja pamięci ze względu na budowę**

**Moduły pamięci SIMM**

**SIMM** (z ang. **S**ingle **I**nline **M**emory **M**odule) – moduł pamięci RAM komputerów w postaci płytki drukowanej z polami stykowymi znajdującymi się z jednej strony płytki



**Moduły pamięci DIMM**

**DIMM** (ang. *Dual In-Line Memory Module*) – szereg standardów modułów pamięci RAM, w których styki złącza krawędziowego modułu znajdują się po obu stronach płytki drukowanej



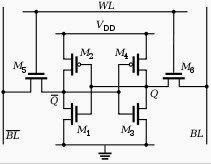
**Klasyfikacja pamięci ze względu na działanie**

**SRAM**

**SRAM** (ang. *Static Random Access Memory*), statyczna pamięć o dostępie swobodnym – typ pamięci półprzewodnikowej stosowanej w komputerach, służy jako pamięć buforująca między pamięcią operacyjną i procesorem.

Słowo "statyczna" oznacza, że pamięć SRAM przechowuje dane tak długo, jak długo włączone jest zasilanie, w odróżnieniu od pamięci typu DRAM, która wymaga okresowego odświeżania.

Każdy bit przechowywany jest w pamięci SRAM w układzie zbudowanym z czterech tranzystorów, które tworzą przerzutnik, oraz z dwóch tranzystorów sterujących. Taka struktura umożliwia znacznie szybsze odczytanie bitu niż w pamięci typu DRAM, oraz w przeciwieństwie do pamięci DRAM nie wymaga odświeżania.



**DRAM**

**Pamięć dynamiczna**, **DRAM** (ang. *Dynamic Random Access Memory*) – rodzaj ulotnej pamięci półprzewodnikowej RAM, która przechowuje każdy bit danych w oddzielnym kondensatorze wewnątrz układu scalonego. Poszczególne jej elementy zbudowane są z tranzystorów MOS, z których jeden pełni funkcję kondensatora, a drugi elementu separującego.

Pamięć wymaga okresowego odświeżania zawartości (ze względu na rozładowywanie się kondensatorów). Jednocześnie pojedyncza komórka pamięci dynamicznej składa się z mniejszej liczby elementów niż analogiczna komórka pamięci statycznej. Powyższe cechy pozwalają na większe upakowanie elementów w układach scalonych, co daje efekt w postaci niższych kosztów produkcji i pozwala na budowę tańszych układów pamięci o danych pojemnościach.

Odświeżanie musi następować w regularnych odstępach czasu i wewnętrznie polega na ponownym zapisie odczytanej wartości w tych samych komórkach pamięci. Za odświeżanie odpowiedzialne są układy pamięci, specjalizowane układy wspomagające (kontroler pamięci – obecnie najczęściej stanowi on integralną część chipsetu)

**Generacje pamięci SDRAM (Synchroniczne DRAM)**

Obecnie do rodziny pamięci SDRAM zaliczamy:

* SDR (**S**ingle **D**ata **R**ate)
  + SDR SDRAM (przepustowość od 533 MB/s do 1066 MB/s)
* DDR (**D**ouble **D**ata **R**ate)
  + DDR1 SDRAM (przepustowość od 1600 MB/s do 3200 MB/s)
  + DDR2 SDRAM (przepustowość od 3200 MB/s do 6400 MB/s)
  + DDR3 SDRAM (przepustowość od 6400 MB/s do 19200 MB/s)
  + DDR4 SDRAM (przepustowość od 12800 MB/s do 25600 MB/s)

**Zasada działania pamięci DDR**

Pamięć DDR wykorzystuje do przesyłania danych narastające i opadające zbocze sygnału zegarowego.

Stosowane są dwa rodzaje oznaczeń pamięci DDR SDRAM. Mniejszy (np. DDR-200) mówi o częstotliwości efektywnej w porównaniu do SDR-SDRAM, z jaką działają kości. Natomiast większy (np. PC1600) mówi o teoretycznej przepustowości jaką mogą osiągnąć. Szerokość magistrali pamięci wynosi 64 bity. Przepustowość obliczana jest według wzoru:

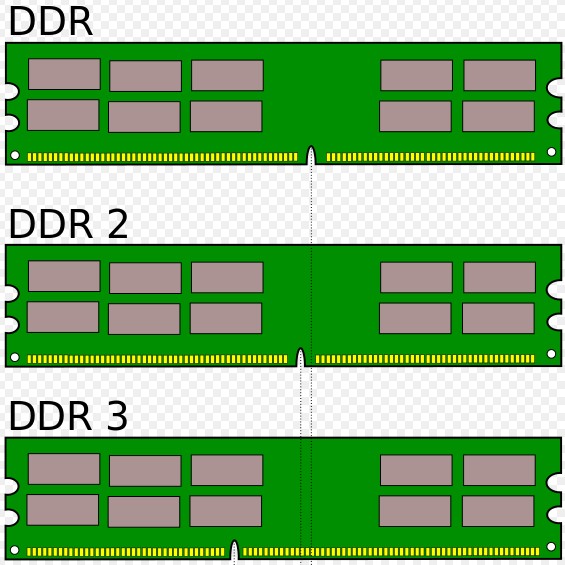
**(Szerokość magistrali) \* 2 \* (Częstotliwość magistrali) / 8 bitów**

gdzie 2 wynika z tego, że DDR to przesył o podwójnym taktowaniu (pamięć DDR2 wykorzystuje do przesyłania danych narastające i opadające zbocze sygnału zegarowego).

Przykłady:

* DDR-200 (PC-1600) – (64 bity \* 2 \* 100 MHz)/8 = 1,6 GB/s
* DDR-266 (PC-2100) – (64 bity \* 2 \* 133 MHz)/8 = 2,1 GB/s
* DDR-333 (PC-2700) – (64 bity \* 2 \* 166 MHz)/8 = 2,7 GB/s
* DDR-400 (PC-3200) – (64 bity \* 2 \* 200 MHz)/8 = 3,2 GB/s

**Rozpoznawanie wersji pamięci DDR**



Źródła: zasoby internetowe