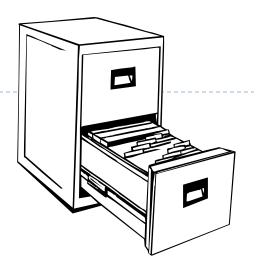
Architektura systemów komputerowych



Wykład II

Pamięci półprzewodnikowe





Pamięci półprzewodnikowe

Pamięci półprzewodnikowe

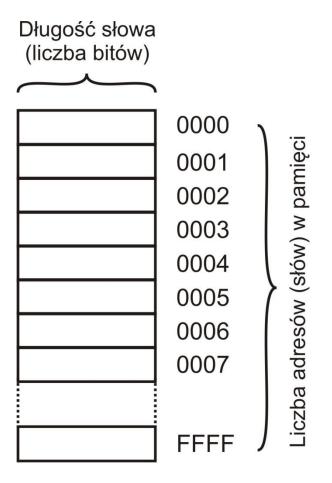
Definicja:

Pamięciami półprzewodnikowymi nazywamy cyfrowe układy scalone przeznaczone do przechowywania większych ilości informacji w postaci binarnej.

Definicja:

Pojemnością pamięci nazywamy maksymalną ilość informacji, jaką możemy przechować w danej pamięci.

Przestrzeń adresowa pamięci



Pamięć podzielona jest na słowa.

Podczas operacji wymiany danych zapisywane lub odczytywane jest zawsze całe słowo (np. podczas komunikacji z pamięcią o słowie 8 bitowym odczytywany lub zapisywany jest zawsze cały bajt).

Każdemu słowu przyporządkowany jest niepowtarzalny numer nazywany adresem.

Sposób podziału pamięci na słowa nazywamy organizacją pamięci.

Pamięci półprzewodnikowe

Definicja:

Pamięcią o dostępie swobodnym nazywamy pamięć, dla której czas dostępu praktycznie nie zależy od adresu słowa w pamięci, czyli od miejsca, w którym jest przechowywana informacja,

UWAGA:

Pamięci ROM są pamięciami o dostępie swobodnym w przeciwieństwie do pamięci o dostępie sekwencyjnym (kolejki danych, pamięci taśmowe), czy pamięci o dostępie blokowym (pamięci masowe).

Pamięci półprzewodnikowe

Definicja:

Pamięcią RAM nazywamy pamięć półprzewodnikową o dostępie swobodnym przeznaczaną do zapisu i odczytu. Pamięć RAM jest pamięcią ulotną, co oznacza, że po wyłączeniu jej zasilania informacja w niej przechowywana jest tracona.

Definicja:

Pamięcią ROM nazywamy pamięć półprzewodnikową o dostępie swobodnym, której zapis następuje na innej drodze niż odczyt (dawniej pamięć tylko do odczytu). Pamięć ROM jest pamięcią nieulotną.

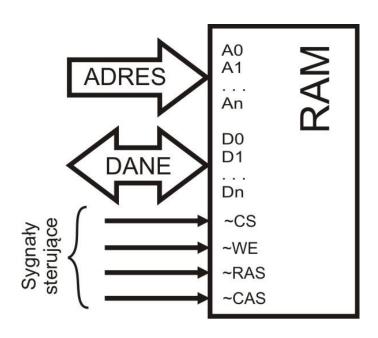
Pamięci półprzewodnikowe

Pamięci RAM są pamięciami ulotnymi. Oznacza to, że dane są tracone po odłączeniu zasilania. Wyróżniamy dwa podstawowe rodzaje pamięci RAM:

- pamięci statyczne SRAM (ang. Static Random Access Memory) są najszybszą odmianą pamięci półprzewodnikowych; Ze względu na skomplikowaną budowę są one trudne w scaleniu, co oznacza, że nie można zmieścić dużych pojemności pamięci w pojedynczych układach scalonych; Pamięci SRAM wykorzystywane są do tworzenia niewielkich lecz szybkich pamięci podręcznych np. pamięci Cache procesora;
- pamięci dynamiczne DRAM (ang. Dynamic Random Access Memory) pamięci znacznie wolniejsze od statycznych lecz dzięki prostej budowie tanie i łatwo poddające się scalaniu; Pamięci DRAM wykorzystywane są między innymi do produkcji pamięci operacyjnych komputerów i pamięci kart graficznych.

Kości pamięci





D0 - Dn - Szyna wejścia/wyjścia danych służy do wprowadzania i wyprowadzania informacji do i z pamięci.

A0 - An - Wejście adresowe służy do dokonania wyboru, na którym z wielu słów w pamięci zastanie wykonana operacja (zapisu bądź odczytu).

WE - Wejście sterujące informuje układ pamięci, jakiego rodzaju operacja będzie wykonywana: odczyt czy zapis.

CS - służy do uaktywnienia układu pamięci. Wejście to jest używane przy budowie zespołów pamięci metodą łączenia dwóch lub więcej układów scalonych pamięci.

RAS i CAS – sygnały sterujące cyklem pracy

Organizacja pamięci

Definicja:

Adresem nazywamy niepowtarzalną liczbę (numer) przypisaną danemu miejscu (słowu) w pamięci celem jego identyfikacji

Definicja:

Słowem w pamięci nazywamy zestaw pojedynczych komórek pamięci, do którego odwołujemy się pojedynczym adresem.

Definicja:

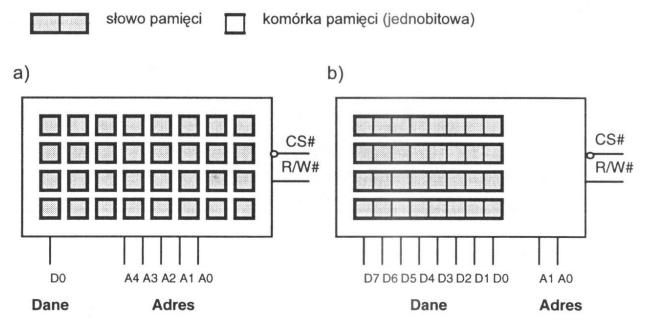
Organizacją pamięci nazywamy sposób podziału obszaru pamięci na słowa.



Organizacja pamięci

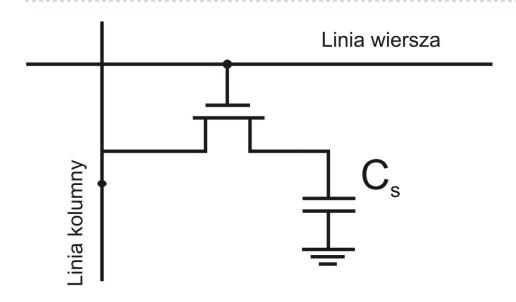
Pamięci na obu rysunkach mają tą samą pojemność (32b). Różnią się natomiast organizacją.

- ✓ Pamięć z rysunku a) ma organizację bitową. Możemy o niej powiedzieć, że jest to pamięć 32 x 1b.
- Pamięć z rysunku b) ma organizacje bajtową, czyli jest to pamięć $4 \times 8b$ (lub inaczej $4 \times 1B$).





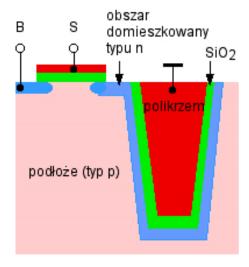
Pamięci półprzewodnikowe – budowa wewnętrzna



http://www.bryk.pl/teksty/liceum/pozosta%C5%82e/informatyka/

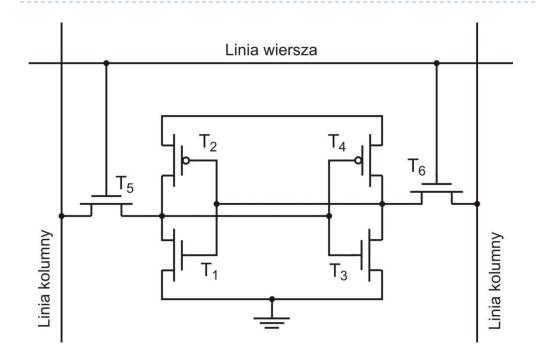
Informacje przechowywane są w pamięci DRAM pod postacią ładunku zgromadzonego przez kondensator.

- Kondensator naładowany "I"
- Kondensator rozładowany "0"



Komórka (bit) pamięci dynamicznej to zaledwie 2 jeden tranzystor i 2 kondensatory.

Pamięci półprzewodnikowe – budowa wewnętrzna



Do przechowywania każdego bitu w pamięci statycznej RAM wykorzystywane układy 4 tranzystorów CMOS oznaczonych na rysunku jako TI,T2 i T3,T4. Tworzą one prosty przerzutnik posiadający dwa stabilne stany wykorzystywane do zapamiętywania wartości logicznych 0 i I. Dwa dodatkowe tranzystory - T5 i T6 służą do sterowania dostępem do komórki podczas zapisu i odczytu danych.

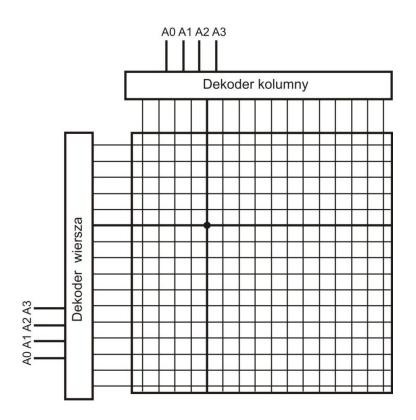
Każda komórka (bit) pamięci statycznej składa się z kilku bramek logicznych i tranzystorów.



Pamięci półprzewodnikowe

- ✓ Pamięci dynamiczne stosowane są do budowy głównej pamięci operacyjnej komputera (RAM).
- ✓ Szybkie pamięci statyczne stosowane są do budowy pamięci podręcznej (cache)

Budowa matrycy pamięci

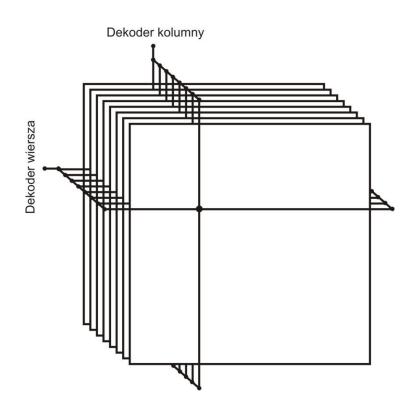


Pamięć półprzewodnikowa jest rodzajem prostokątnej matrycy. Każdy bit pamięci znajduje się na przecięciu dwóch prostopadłych linii: linii wiersza i linii kolumny.

Wyboru właściwej linii dokonują dwa dekodery - dekoder wiersza oraz dekoder kolumny.

Dekodery to układy cyfrowe które mają kilka (n) linii wejściowych i znacznie więcej linii wyjściowych (2ⁿ). Dekoder otrzymuje na wejściu liczbę binarną, dekoduje ją a następnie podaje sygnał na jedną z linii wyjściowych – linię o numerze odpowiadającym otrzymanej licznie.

Budowa matrycy pamięci



Układ pamięci o wielobitowym słowie składa się kilka matryc ułożonych warstwowo. Liczba warstw odpowiada długości słowa. Np. pamięć 8-io bitowa składa się z ośmiu matryc.

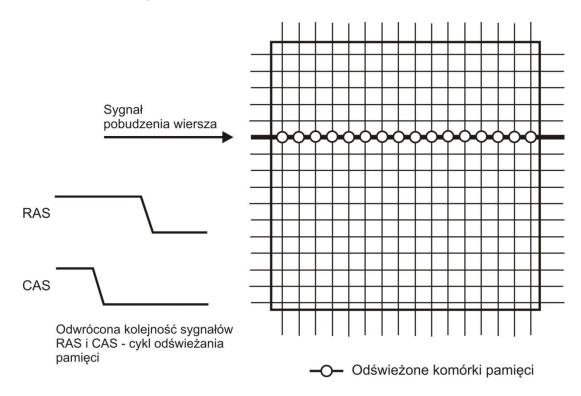
Ponieważ dekodery adresu są wspólne dla wszystkich matryc, wybranie wiersza i kolumny oznacza automatycznie zaadresowanie odpowiadających sobie komórek we wszystkich warstwach, a więc wybranie całego słowa.

16

B

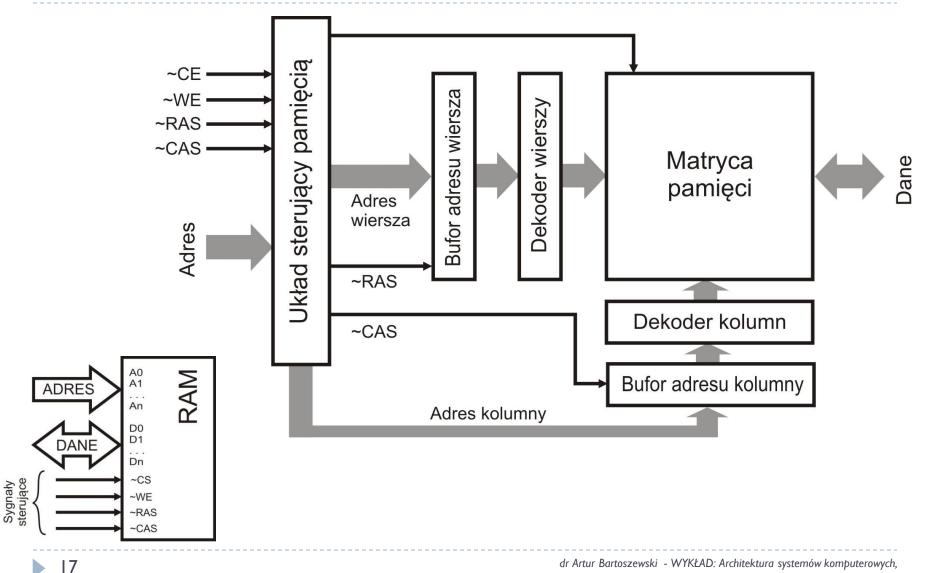
Odświeżanie pamięci

Pamięci dynamiczne, wykorzystywane powszechnie jako pamięci operacyjne komputerów, wymagają odświeżania czyli cyklicznego uzupełniania ładunku w każdej komórce pamięci.



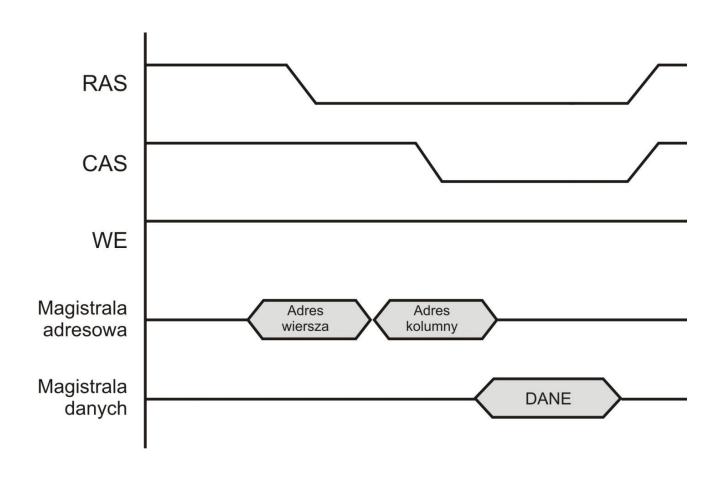
Na szczęście nie jest konieczne adresowanie i odświeżanie każdej komórki z osobna. Wysłanie sygnału pobudzenia na linię wiersza powoduje odświeżenie wszystkich leżących na niej komórek.

Budowa matrycy pamięci DRAM



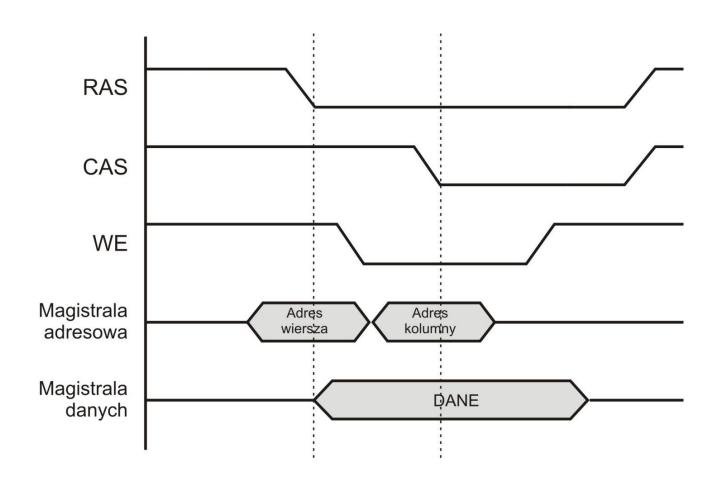


Operacja odczytu dla pamięci DRAM





Operacja zapisu dla pamięci DRAM





Odświeżanie pamięci DRAM

Definicja:

Czas, jaki upływa od momentu podania prawidłowego adresu przez zarządcę magistral do momentu pojawienia się poprawnych danych na magistrali adresowej, nazywamy czasem dostępu (t_a).

Po odczycie zawartości słowa musi upłynąć kolejny odcinek czasu potrzebny do doładowania komórek pamięci odczytywanego słowa (ang. precharge delay). Dopiero wówczas może się rozpocząć kolejny cykl dostępu do pamięci.

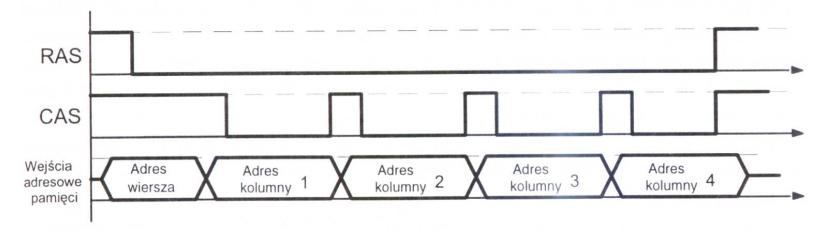
Definicja:

Odświeżanie komórek pamięci DRAM polega na cyklicznym doładowywaniu pojemności pamiętających przechowujących wartość I.

Dostęp w trybie stronicowanym

Dostęp do pamięci w trybie stronicowania jest sposobem na przyspieszenie współpracy z pamięcią DRAM. Wykorzystuje się tu dwa fakty.

- 1. Większość odczytów dokonywana jest spod kolejnych, położonych koło siebie adresów. Oznacza to, że starsza część adresu, adres wiersza, nie zmienia się, a zmienia się jedynie adres kolumny.
- 2. Czas przesłania adresu wiersza stanowi około 50% czasu dostępu. Jeżeli przy odczytach kolejnych słów nie będziemy zmieniać adresu wiersza, a jedynie adres kolumny, to czas dostępu do pamięci ulegnie skróceniu.

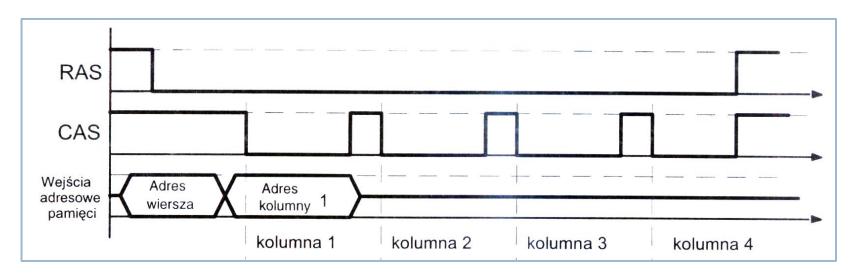




Dostęp w trybie seryjnym (bust)

Dostęp w trybie seryjnym (ang. bursf) stosowany jest przy współpracy pamięci głównej z pamięcią cache. Pamięć ta odczytuje bądź zapisuje informacje liniami, których długość zależy od rozwiązania pamięci cache.

Przykładowo dla systemów z procesorem 80486 wynosi 16 bajtów. Ponieważ procesor ten ma magistrale danych 32-bitową (4 bajty), do wypełnienia linii potrzeba 4 dostępów do pamięci. Operacje te dotyczą jednak kolejnych, leżących obok siebie słów. Oznacza to, że adres wiersza nie będzie się zmieniał, zaś adres kolumny przy każdym kolejnym dostępie będzie większy o jeden.

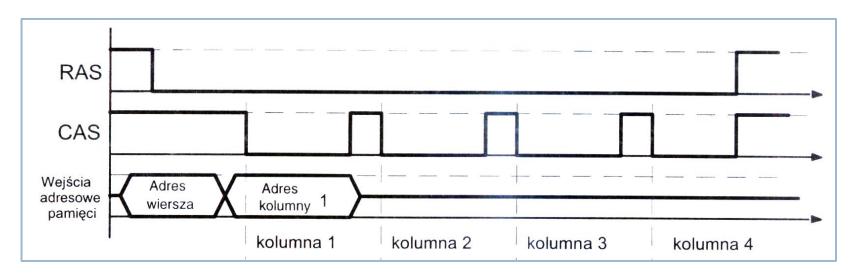




Dostęp w trybie seryjnym (bust)

Dostęp w trybie seryjnym (ang. bursf) stosowany jest przy współpracy pamięci głównej z pamięcią cache. Pamięć ta odczytuje bądź zapisuje informacje liniami, których długość zależy od rozwiązania pamięci cache.

Przykładowo dla systemów z procesorem 80486 wynosi 16 bajtów. Ponieważ procesor ten ma magistrale danych 32-bitową (4 bajty), do wypełnienia linii potrzeba 4 dostępów do pamięci. Operacje te dotyczą jednak kolejnych, leżących obok siebie słów. Oznacza to, że adres wiersza nie będzie się zmieniał, zaś adres kolumny przy każdym kolejnym dostępie będzie większy o jeden.





SDRAM (S = Synchroniczne)

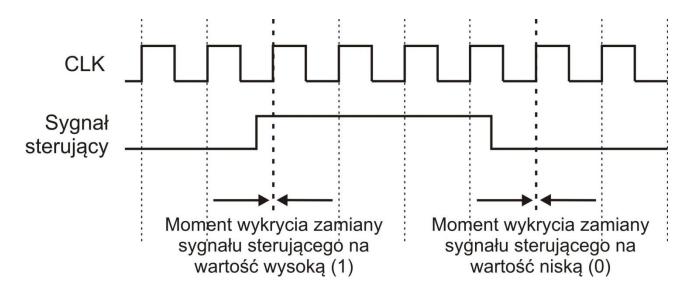
- ✓ Wszystkie sygnały sterujące pamięci SDRAM synchronizowane są przez jeden przebieg zegarowy. Ułatwia to współpracę pamięci z magistralami.
- ✓ Sygnał taktujący pozwala na dokładne zsynchronizowanie ze sobą operacji wykonywanych przez pamięć i jej kontroler – umożliwia to skrócenie czasu przeznaczonego na te operacje, a więc przyspieszenie działania pamięci.
- Obsługa odświeżania pamięci przeniesiona została do wnętrza kości. Każdy chip wyposażony jest we własny generator pobudzający w odpowiednim rytmie wszystkie komórki pamięci. Powoduje to odciążenie kontrolera pamięci i umożliwia dalsze przyspieszenie jej pracy.

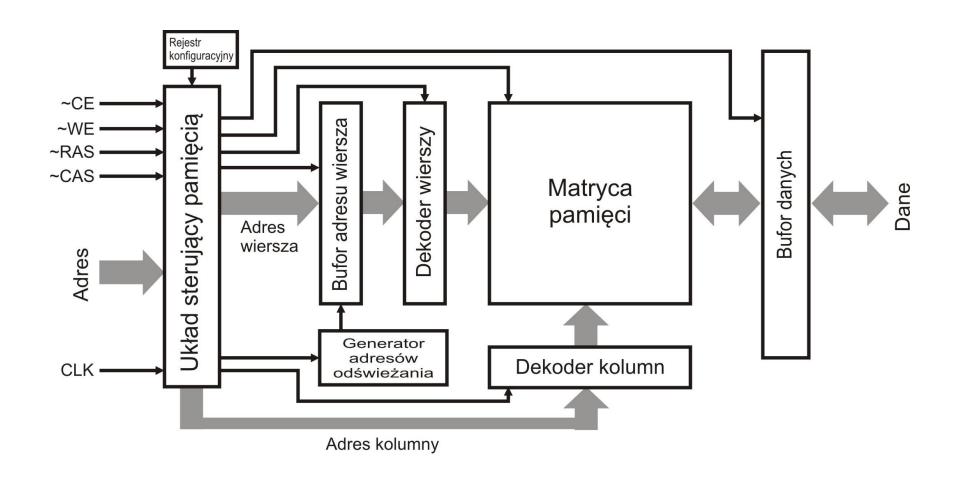


Pamieci SDRAM (ang. Synchronous Dynamic Random Access Memory).

Są one odmianą pamięci dynamicznych DRAM różniącą się od swojego poprzednika synchronicznym przesyłem danych.

Kość SDRAM posiada wejście zegarowe CLK poprzez które podawany jest sygnał taktujący. Sygnał ten wyznacza dokładnie momenty odczytania adresu i danych oraz momenty reakcji na sygnały sterujące. Reakcja układu stępuje zawsze w chwili wykrycia zbocza narastającego sygnału zegarowego CLK



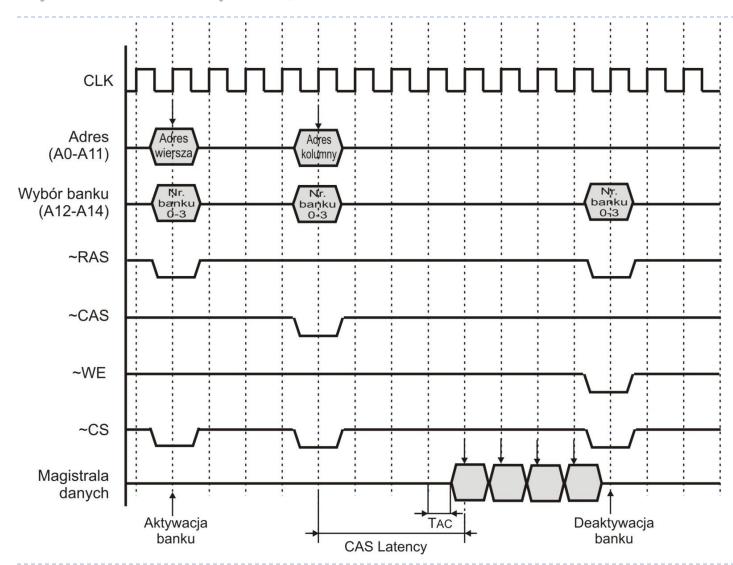




Pamięć SDRAM podzielona jest na kilka (najczęściej cztery) niezależne części nazywane bankami.

Każdy z nich może być aktywowany oddzielnie. Wybór banku dokonywany jest na podstawie ostatnich trzech bitów magistrali adresowej. Na przykład dla pamięci SDRAM o 15-to bitowej magistrali adresowej dwanaście pierwszych bitów (A0 - AII)to adres, a trzy ostatnie (A12 - A14) to numer banku.

Synchroniczna pamięć SDRAM





Cykl odczytu danych pamięci SDRAM, tak jak pokazano na poprzednim slajdzie, przebiega w pięciu etapach:

- 1. Na magistralę adresową (A0 A11) podawana jest starsza część adresu (adres wiersza); Na trzech ostatnich liniach magistrali adresowej (A12 A14) podawany jest numer banku który zostaje aktywowany; Podawane są sygnały sterujące (zbocze malejące) ~CE uaktywniający układ oraz ~RAS wyznaczający podanie adresu wiersza; Sygnały sterujące muszą być utrzymane przez przynajmniej jeden takt zegara, ich odczyt następuje w chwili wyznaczonej przez zbocze narastające CLK;
- Na magistralę adresowa podawana jest młodsza część adresu (adres kolumny) oraz ponownie numer banku, a na odpowiednie wejścia sterujące sygnały ~CE oraz ~CAS (wyznaczający moment podania adresu kolumny);
- 3. Układ odczekuje określoną liczbę taktów zegara (CAS Latency);



- 4. Na magistralę danych podawana jest seria kolejnych słów (transmisja w trybie Burst); seria rozpoczyna się od słowa o podanym adresie; W określonych odstępach pamięć podaje określoną liczbę słów o kolejnych adresach; Seria danych może mieć składać się z 1, 2, 4 lub 8 słów lub nawet całego wiersza; Długość serii ustalana jest w trakcie konfiguracji pamięci; Każde słowo utrzymywane jest przez jeden cykl zegara, a moment podania pierwszego słowa serii opóźniony jest (względem zbocza narastającego CLK) o czas TAC, dzięki czemu moment odczytu danych, wyznaczony przez zbocze narastające sygnału zegarowego, przypada w jej środku;
- 5. Po zakończeniu odczytu konieczna jest dezaktywacja banku pamięci; Dokonywana jest ona poprzez podanie numeru banku oraz kombinacji sygnałów ~RAS, ~WE i ~CE.



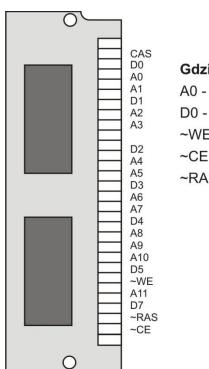
Budowa bloków (banków) pamięci polega na łączeniu układów scalonych pamięci o określonej pojemności i organizacji w ten sposób, aby uzyskać zespoły pamięci o większej pojemności i/lub o zmienionej długości słowa.

Stąd problem rozbudowy pamięci możemy podzielić na dwa podstawowe przypadki:

- zwiększanie (rozszerzanie) długości słowa przy niezmienionej ilości słów
- zwiększanie ilości słów przy niezmienionej długości słowa.

Oczywiście obydwa przypadki mogą występować (i w praktyce często występują) jednocześnie.

Zwiększanie długości słowa



Gdzie:

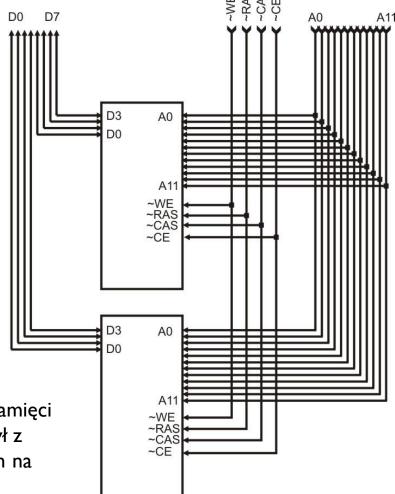
A0 - A11: Magistrala adresowa (12 bitowa)

D0 - D7: Magistrala danych (8 bitowa)

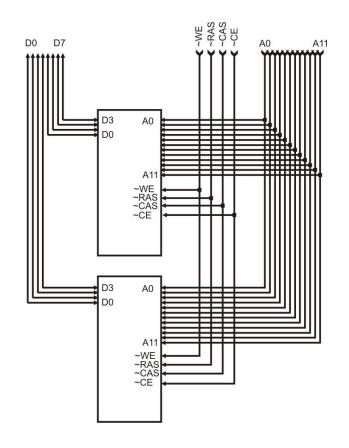
~WE: wybór trybu zapis/odczyt

~CE: uaktywnienie układu

~RAS, ~CAS: sygnaly sterujące cyklem pracy



Objaśnić można na przykładzie 8 bitowego modułu pamięci SIMM. Układ ten, w najprostszej wersji, zbudowany był z dwóch czterobitowych kości pamięci zamontowanych na płytce o 30-pinowym złączu



W celu zwiększenia długości słowa pamięci szerszą magistralę danych budujemy z bitów linii danych kolejnych układów scalonych pamięci, natomiast magistralę adresową i sygnały sterujące łączymy równolegle.

- ✓ We wszystkich połączonych układach, wybieramy słowa położone w takim samym miejscu.
- ✓ Wszystkie układy dostają te same sygnały sterujące.
 - Wszystkie układy są równocześnie uaktywnione (sygnał CS).
 - Na wszystkich słowach składowych wykonujemy tę samą operację, zapis lub odczyt, (ten sam sygnał WE)

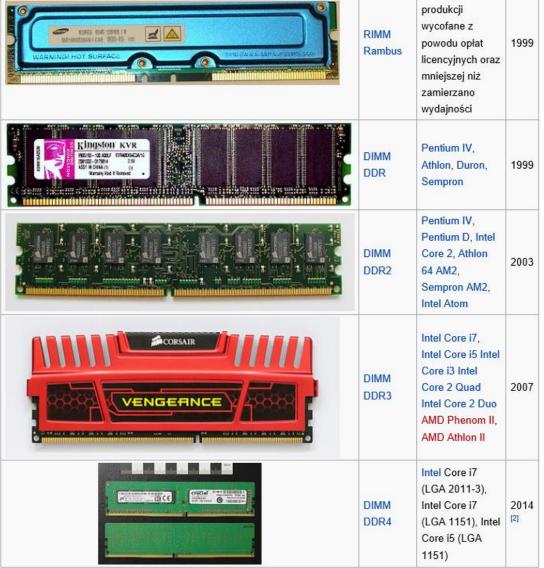
Moduły pamięci

| Wygląd wielkość naturalna na monitorze 17" (1024x768) | Obudowa Pamięć | Użycie | Rok |
|---|----------------------|--|------|
| TATALAN | DIP - | PC, XT, AT | 1981 |
| | SIPP - | 286, AT, 386 | 1983 |
| CWattherson -Wattherson - Commence of the control of the contro | SIMM (30- pinowe) | Niektóre 286, 386, 486 | 1984 |
| | SIMM (72- pinowe) | PS/2, 486, Pentium, AMD K6, AMD K5 | 1990 |
| Źródło: https://pl.wikipedia.org/wiki/RAM#Technologie_pami.C4.99ci_RAM | DIMM SDR SDRAM | Niektóre Pentium, Pentium II, Pentium IV i Celeron, a także AMD K6 | 1997 |

Moduły pamięci

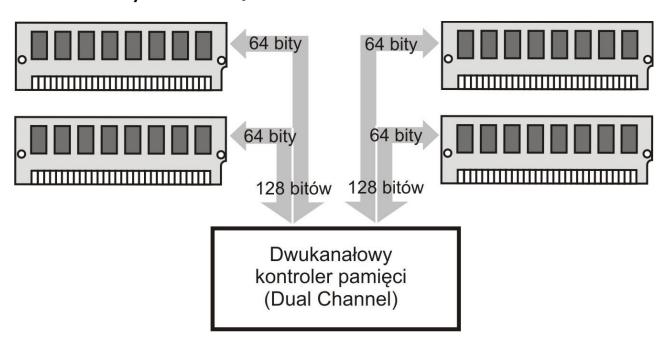




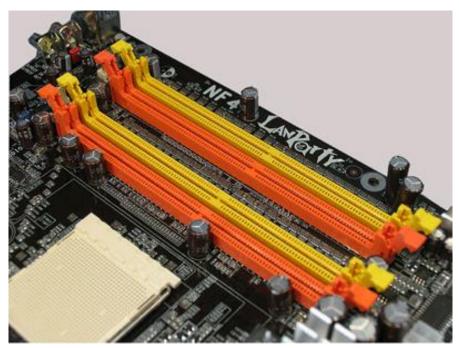


Pamięci wielokanałowe

Specyficznym przykładem łączenia układów pamięci w celu zwiększenia długości słowa są, powszechnie stosowane we współczesnych komputerach PC, wielokanałowe układy pamięci RAM. Dwukanałowy kontroler pamięci (ang. Dual Channel) łączy dwa 64 bitowe moduły pamięci w jeden 128 bitowy układ. Jednakowe moduły umieszczane są parami w skorelowanych ze sobą slotach.



Pamięci wielokanałowe





TRIPLE CHANNEL

Łączenie układów pamięci

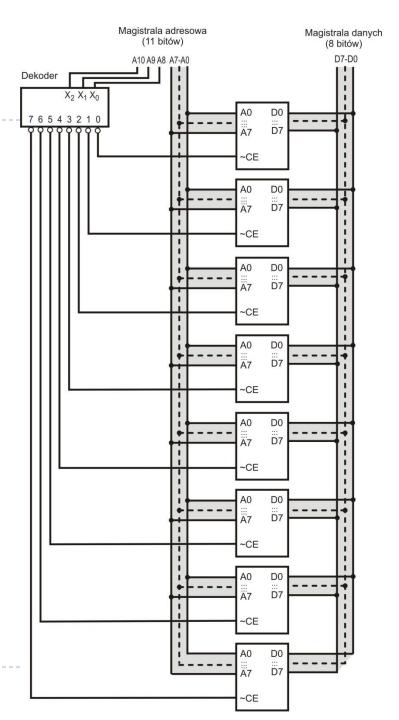
Zwiększenie liczby słów w pamięci (zwiększenie pojemności pamięci)

Każdy z układów scalonych posiada 8 wejść magistrali adresowej (A0 - A7).

Do zaadresowania 2048 słów potrzeba 11-bitowego adresu.

Trzy najstarsze bity adresu (A8 – A10) służą do wyboru jednej z ośmiu kości pamięci.

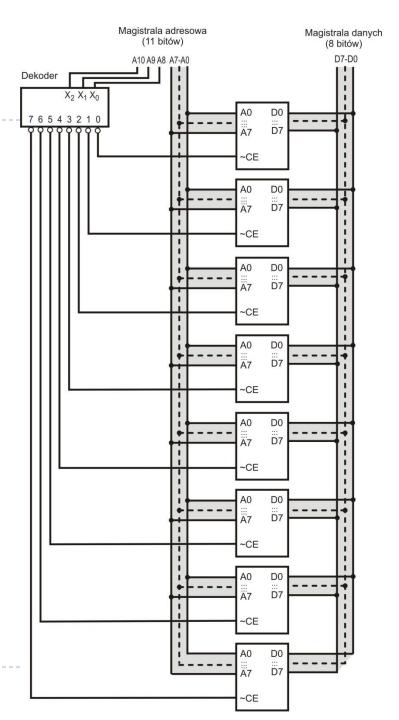
Osiem młodszych bitów służy do wyboru słowa we wskazanej kości.



Łączenie układów pamięci

Zwiększenie liczby słów w pamięci (zwiększenie pojemności pamięci)

Przykładowo, adres 000 0000 1000 oznacza szesnaste słowo w pierwszym układzie scalonym (na rys.. pierwsza kość od góry). Adres III 0000 1000 to również szesnaste słowo, lecz w ósmej kości pamięci. Wybór właściwej kości dokonywany jest,, na podstawie trzech najstarszych bitów adresu. Trafiają one na wejścia dekodera, który wysyła sygnał (w tym wypadku stan niski napięcia) na wyjście o podanym numerze (000 to wyjście pierwsze, 001 drugie itd.). Sygnał trafia na wejście ~CE jednej z kości pamięci i uaktywnia ja. Pozostałe nie są aktywne.



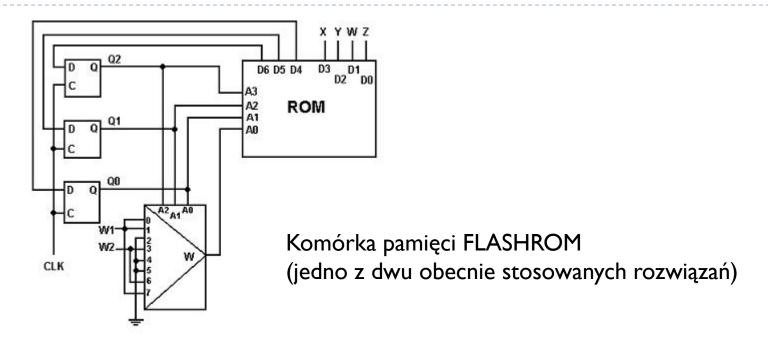


Łączenie układów pamięci

- ✓ Zwiększenie ilości słów pamięci oznacza zwiększenie ilości potrzebnych adresów, a co za tym idzie **rozbudowę szyny adresowej** o dodatkowe bity potrzebne do uzyskania tych adresów.
- ✓ Przy niezmienionej długości słowa szyna danych pozostaje bez zmian.
- ✓ Dodatkowe bity adresu służą, przy wykorzystaniu dekodera, do wyboru jednego z łączonych układów pamięci, z którego odczytamy lub do którego zapiszemy informacje.
- ✓ Wyboru dokonujemy przy użyciu wejścia CS# uaktywniającego układy scalone pamięci.
- ✓ Magistrale adresowe, danych i sygnały sterujące układów, z których budujemy nowy blok pamięci, łączymy równolegle.

B

Pamięć ROM



Definicja:

Pamięcią ROM nazywamy pamięć półprzewodnikową o dostępie swobodnym, której zapis następuje na innej drodze niż odczyt (dawniej pamięć tylko do odczytu). Pamięć ROM jest pamięcią nieulotną.

Pamięć ROM

Pamięć ROM (ang. read only memory) jest pamięcią nieulotną, gdzie zapis następuje na innej drodze niż odczyt.

- Nieulotność oznacza, że po wyłączeniu napięcia zasilania tej pamięci, informacja w niej przechowywana nie jest tracona (zapominana).
- ✓ Do pamięci tej nie możemy zapisywać danych w trakcie jej normalnej pracy w systemie (cyklu rozkaowego).
- Pamięci ROM są pamięciami o dostępie swobodnym.
- ✓ Określenie, że jest to pamięć tylko do odczytu, nie jest równoznaczne z tym, że zawartości tej pamięci w określonych warunkach nie można zmieniać. Dla niektórych typów technologicznych pamięci ROM jest to możliwe.



MROM (ang. mascable ROM) - pamięci, których zawartość jest ustalana w procesie produkcji (przez wykonanie odpowiednich masek - stąd nazwa) i nie może być zmieniana. Przy założeniu realizacji długich serii produkcyjnych jest to najtańszy rodzaj pamięci ROM.W technice komputerowej dobrym przykładem zastosowania tego typu pamięci jest BIOS obsługujący klawiaturę.



PROM (ang. programmable ROM) - pamięć jednokrotnie programowalna. Oznacza to, że użytkownik może sam wprowadzić zawartość tej pamięci, jednakże potem nie można jej już zmieniać. Cecha ta wynika z faktu, że programowanie tej pamięci polega na nieodwracalnym niszczeniu niektórych połączeń wewnątrz niej. Obecnie ten typ pamięci nie jest już używany.



EPROM - pamięć wielokrotnie programowalna, przy czym kasowanie poprzedniej zawartości tej pamięci odbywa się drogą naświetlania promieniami UV. Programowanie i kasowanie zawartości tej pamięci odbywa się poza systemem w urządzeniach zwanych odpowiednio kasownikami i programatorami pamięci. Pamięć ta wychodzi już z użycia.





Pamięci PROM – programuje się w specjalnym urządzeniu nazywanym programatorem pamięci.



B

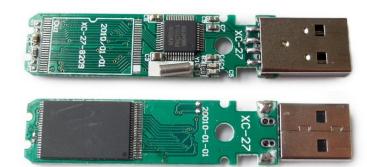
Rodzaje pamięci ROM

4. EEPROM - pamięć kasowana i programowana na drodze czysto elektrycznej. Istnieje możliwość wprowadzenia zawartości tego typu pamięci bez wymontowywania jej z systemu (jeżeli oczywiście jego projektant przewidział taką opcje) choć czas zapisu informacji jest nieporównywalnie dłuższy niż czas zapisu do pamięci RAM.

Wykonywana w różnych postaciach (np. jako FLASH), różniących się sposobem organizacji kasowania i zapisu.

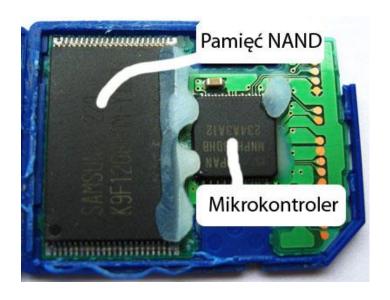






Karty pamięci





Karta pamięci ma prostą konstrukcję. Składa się z mikrokontrolera, który steruje zapisem i odczytem danych z pamięci NAND złącza, które wpinamy do czytnika.

Źródło: http://www.centrumodzyskiwaniadanych.pl/







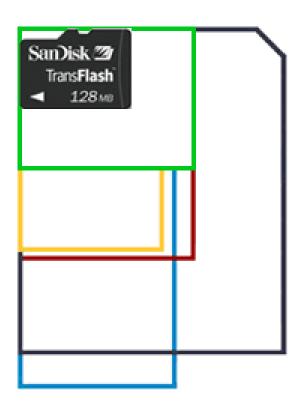






Karty pamięci





xD-Picture **Memory Stick Duo** Secure Digital MultiMedia Card SmartMedia 5 4 1 **Memory Stick**

© benchmark.pl



Literatura:

Metzger Piotr - Anatomia PC, wydanie XI, Helion 2007

Wojtuszkiewicz Krzysztof - Urządzenia techniki komputerowej, część I: Jak działa komputer, MIKOM, Warszawa 2000

Wojtuszkiewicz Krzysztof - Urządzenia techniki komputerowej, część II: Urządzenia peryferyjne i interfejsy, MIKOM, Warszawa 2000

Komorowski Witold - Krótki kurs architektury i organizacji komputerów, MIKOM Warszawa 2004

Gook Michael - Interfejsy sprzętowe komputerów PC, Helion, 2005