

Organizacja pamięci i kontrolery DRAM

Jan Mazur

Instytut Informatyki Uniwersytetu Wrocławskiego
Seminarium: Architektury Systemów Komputerowych

Trochę o pamięciach

DRAM (*dynamic random access memory*) - dynamiczna, półprzewodnikowa, ulotna pamięć swobodnego dostępu i wielokrotnego zapisu

pamięć dynamiczna - wymaga odświeżania zawartości podczas pracy z powodu rozładowywania się kondensatorów

pamięć ulotna - nie zachowuje danych bez zasilania

SRAM (*static random access memory*) - ~7x szybsza, ~4x większa, zużywa mniej prądu, droższa, używana np. w x86 do budowy cache

pamięć nieulotna, trwała - nośniki magnetyczne (kasety, dyskietki, dyski twarde), dyski optyczne (CD, DVD), pamięci półprzewodnikowe (*ROM), dyski SSD itd.

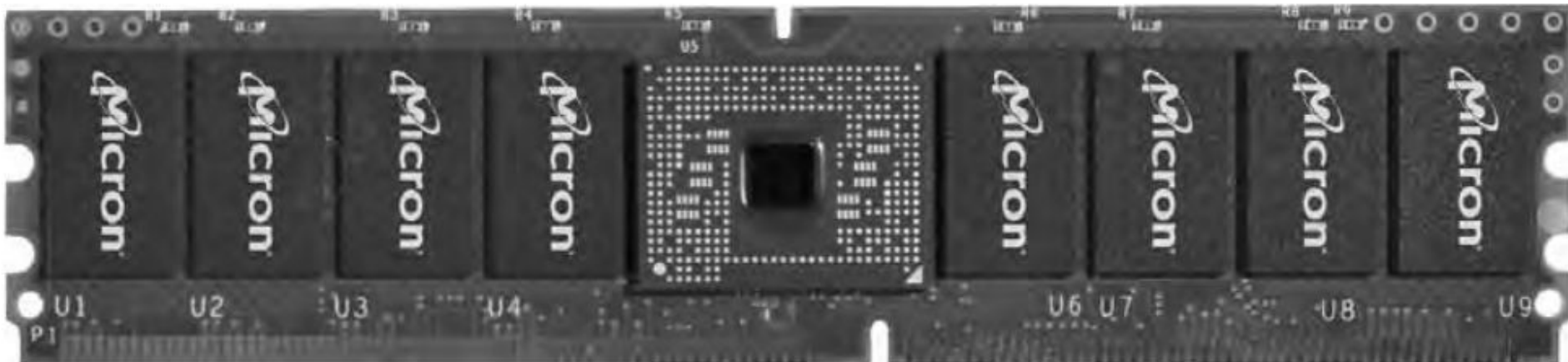
Moduł pamięci - “kość pamięci”

PCB (*printed circuit board*) - płytko drukowana

moduł pamięci (*memory module*) - PCB na której są chipy DRAM

DRAM chip - plastikowe obudowy na urządzenia DRAM

urządzenie DRAM (*DRAM die*) - odwód zrealizowany w krzemie, spełniający funkcję urządzenia DRAM



Moduł pamięci, DIMM*

DIMM - szereg standardów modułów pamięci RAM, w których styki na złączu PCB występują po 2 stronach płytki i są one elektrycznie inne; termin używany również w odniesieniu do całości modułu pamięci

SIMM - złącza po 2 stronach PCB są elektrycznie identyczne

RANK - zbiór urządzeń DRAM działający sekwencyjnie (*in lockstep*) w stosunku do innych rank przy obsłudze komend systemu pamięci; termin używany również do określenia urządzeń DRAM po jednej stronie PCB

Budowa urządzenia DRAM

Sercem urządzenia DRAM są **tablice pamięci** (*memory arrays*), które zawierają **komórki pamięci** (*storage cell*) przechowujące bit informacji w kondensatorze jako ładunek elektryczny. Urządzenie DRAM zawiera co najmniej jedną tablicę pamięci.

Każdy kondensator musi być co jakiś czas odświeżony (przeczytany i ponownie zapisany), aby permanentnie nie utracić ładunku elektrycznego (bitu informacji).

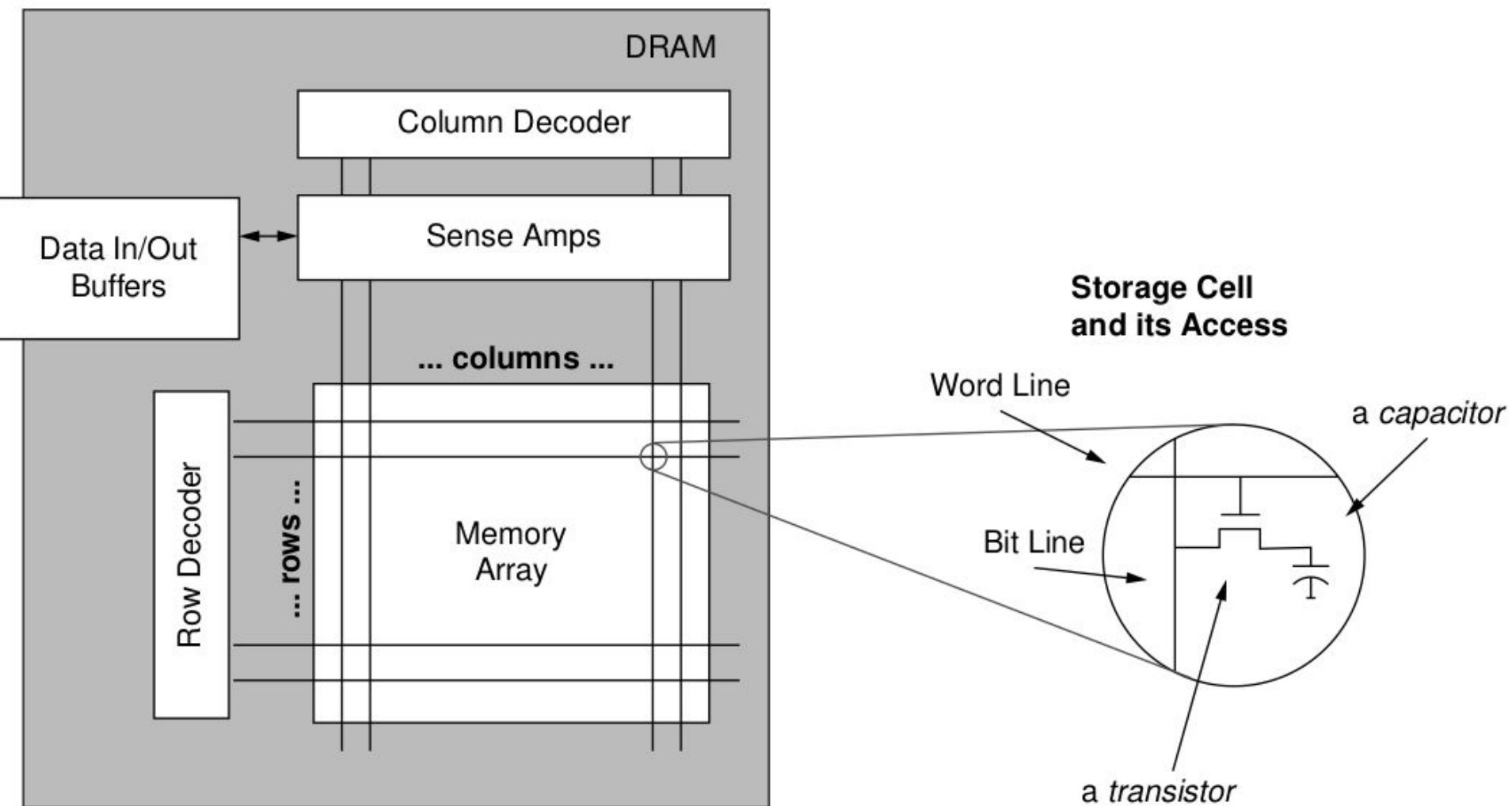
Komórki pamięci w tablicach są adresowane poprzez **adres kolumny**, **adres wiersza**.

Zapis/odczyt komórki odbywa się poprzez odpowiednie manipulowanie napięciem na przewodach **linii słowa** i **linii bitu**.

Zajmują się tym **dekoder kolumny** i **dekoder wiersza**, oraz **wzmacniacz**.

Wzmacniacz wykrywa obecność bitu, odnawia ładunek i potem ładuje nim z powrotem kondensator.

Budowa urządzenia DRAM



Banki

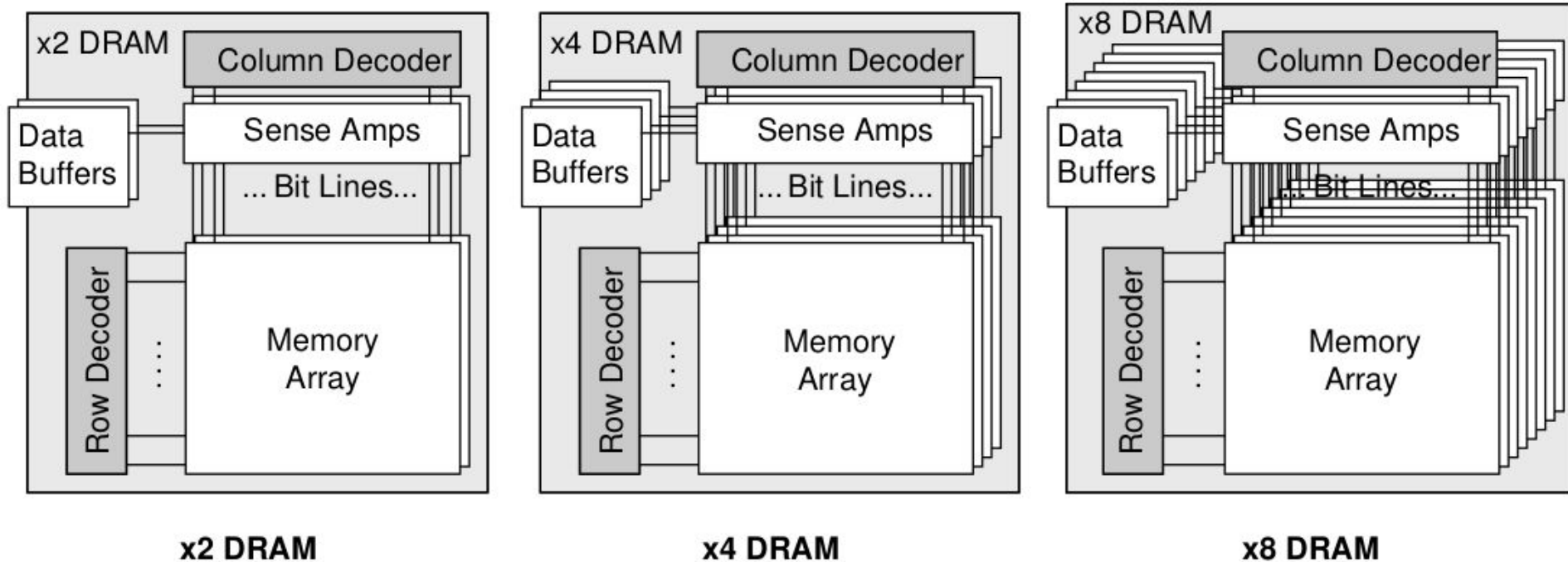
Tablice w urządzeniu DRAM mogą pracować niezależnie bądź działać zgodnie (*in unison*). W drugim przypadku zbiór takich tablic nazywamy **bankiem**.

Proste urządzenie DRAM zawierające n tablic w jednym banku czyta/zapisuje jednorazowo n bitów. Z tego samego miejsca w każdej tablicy. Oznaczamy to jako xn DRAM ($x4$ DRAM - *by four* DRAM, $x8$ DRAM - *by eight* DRAM).

Urządzenie DRAM może zawierać wiele niezależnych banków.

Banki

Skoro banki są niezależne a operacja odczytu/zapisu składa się z kilku kroków, możemy osiągnąć lepszą przepustowość, przy urządzeniach DRAM o małej przepustowości i szynie o większej przepustowości, przeplatając operację na bankach, powodując ich współbieżną pracę (IBM System/360 Model 91, 1967).



Rank

Termin **rank** został wprowadzony aby rozróżnić niezależność operacji na poziomie DIMM i w obrębie banku.

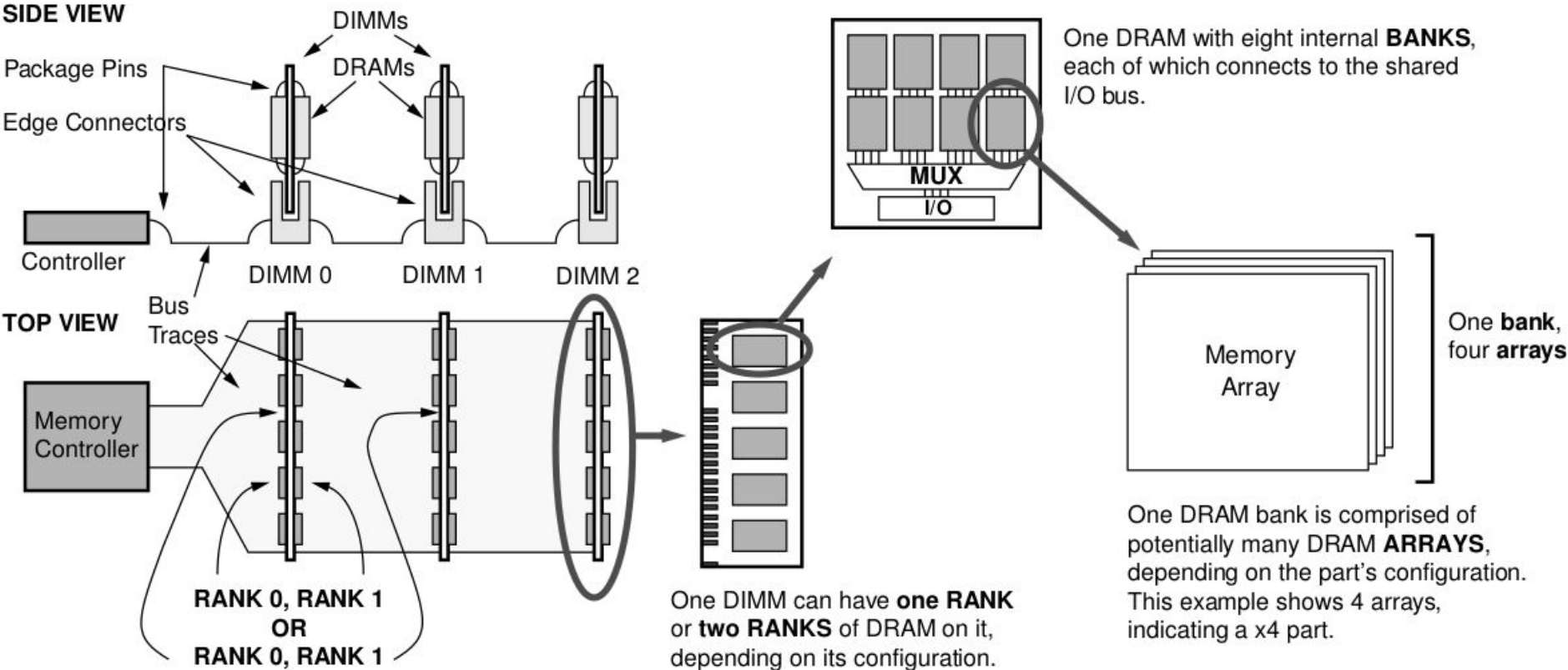
DIMM zawiera jeden lub więcej ranków.

Rank to zbiór urządzeń DRAM działających zgodnie (in unison).

Having concurrency at the rank and bank levels provides bandwidth through the ability to pipeline requests.

Having multiple DRAMs acting in unison at the rank level and multiple arrays acting in unison at the bank level provides bandwidth in the form of parallel access.

System pamięci

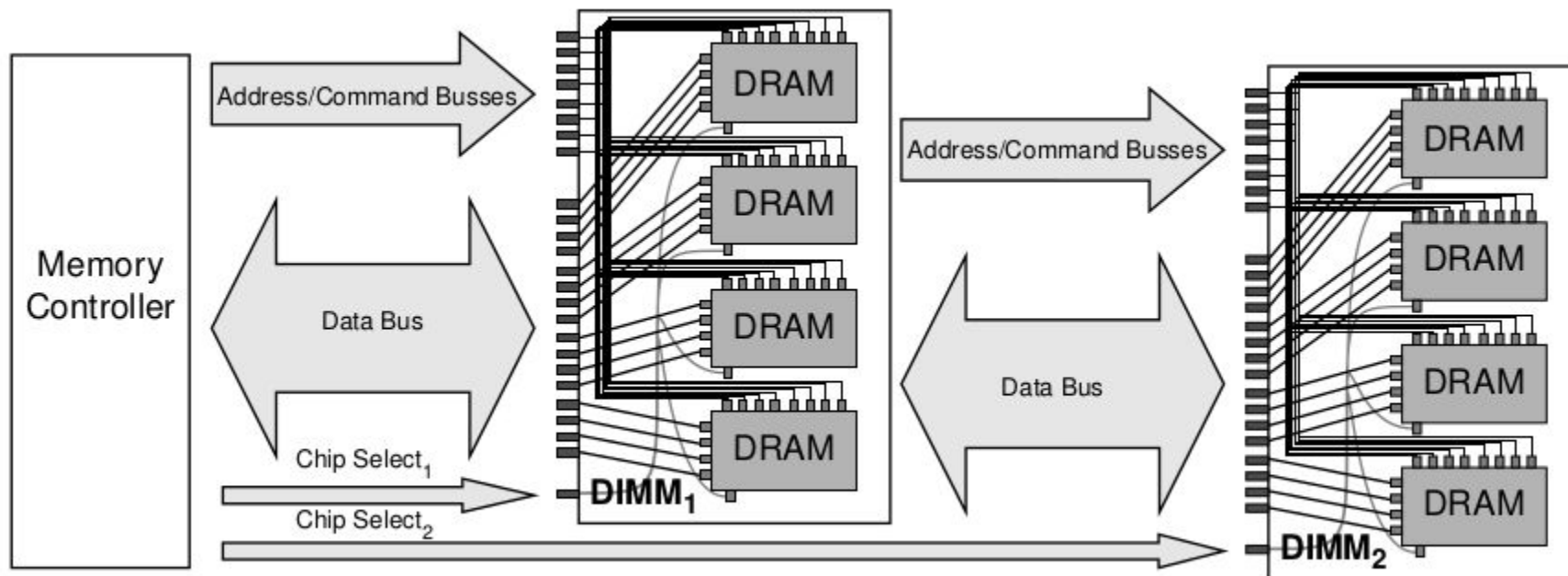


Szyna JEDEC

Organizacja szyny podzielona na 4 mniejsze szyny:

- danych - do transmisji danych, zazwyczaj szerokości 64 bitów; połączenie podzielone pomiędzy urządzenia DRAM w rank
- adresową - numery kolumny i wiersza, zazwyczaj szerokości 15 bitów; podłączona do każdego urządzenia DRAM w systemie
- kontroli - sygnały (strokes) wiersza (RAS) i kolumny (CAS), zegar i inne sygnały kontrolne; podłączona do każdego urządzenia DRAM w systemie
- wyboru chipu - osobna dla każdego rank; podłączona do każdego urządzenia DRAM w rank; wybiera urządzenia DRAM które mają odpowiedzieć na żądanie (blokuje inne)

Szyna adresowa JEDEC

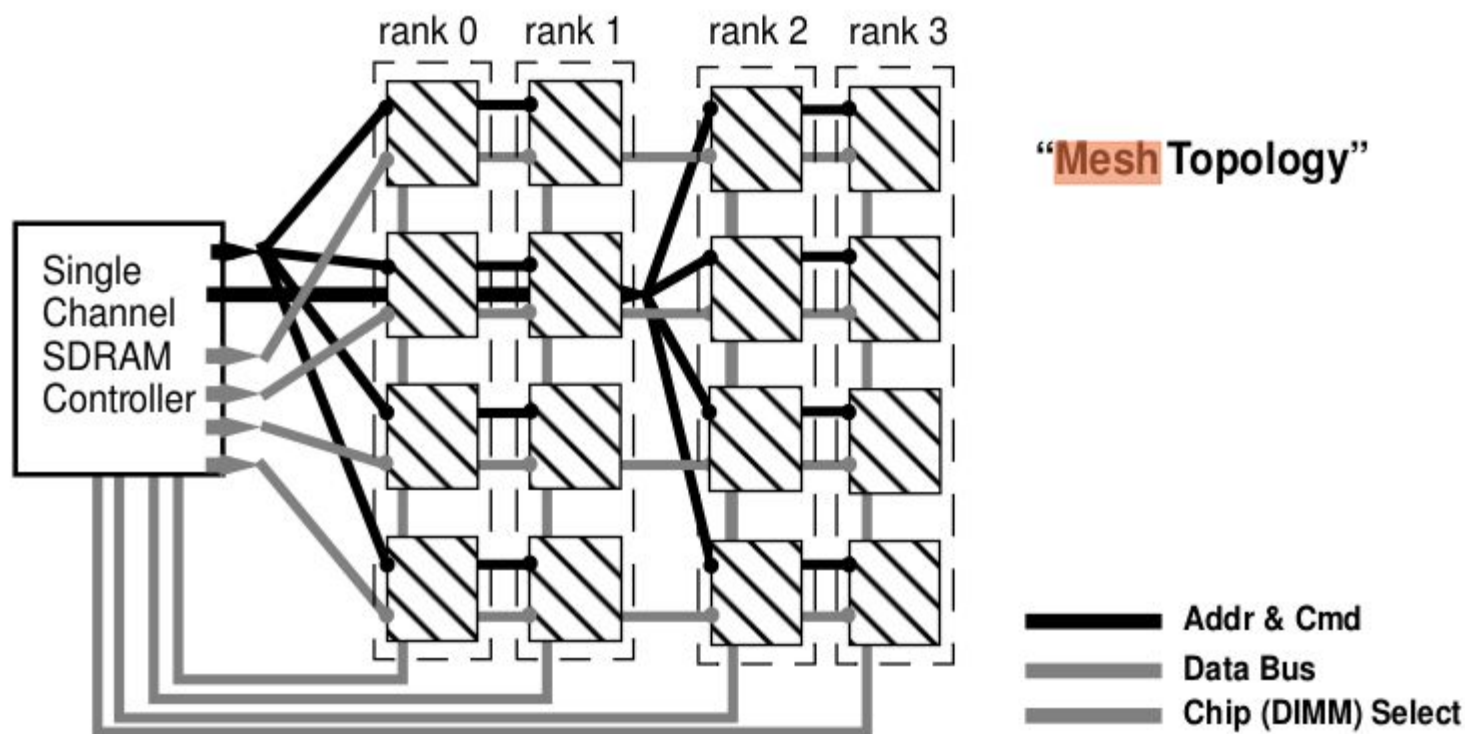


2 moduły pamięci po 1 rank każdy. W rank 4 urządzenia DRAM.

16 bitowa szyna danych, 8 bitowa szyna adresowa, 8 bitowa szyna kontrolna

Szyna wyboru chipu rozbita na 2 osobne dla każdego rank w systemie.

Wysokopoziomowy ogląd na szyny

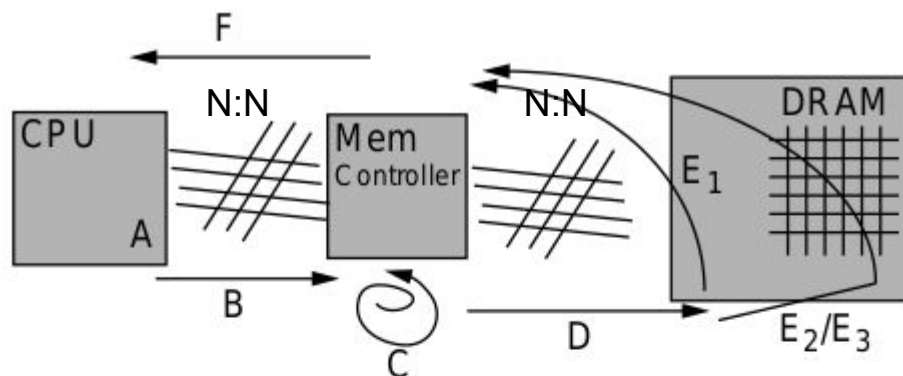


2 DIMM x 2 rank x 4 urządzenia DRAM. Zwróćmy szczególną uwagę na prowadzenie szyn i ich współdzielenie.

Jak wygląda dostęp?

CPU, kontroler(y) pamięci i system pamięci (urządzenie(a) DRAM) są połączone pewnego rodzaju siecią.

Kontroler pamięci stanowi interfejs pomiędzy CPU a systemem pamięci.



- A: Transaction request may be delayed in Queue
- B: Transaction request sent to Memory Controller
- C: Transaction converted to Command Sequences (may be queued)

D: Command/s Sent to DRAM

- E₁: Requires only a **CAS** or
- E₂: Requires **RAS + CAS** or
- E₃: Requires **PRE + RAS + CAS**

W zależności czy otwarty wiersz i jaki typ DRAM.

F: Transaction sent back to CPU

$$\text{DRAM Latency} = A + B + C + D + E + F$$

FIGURE 7.7: System organization and the steps of a DRAM read. Reading data from a DRAM is not as simple as an SRAM, and at several of the stages the request can be stalled.

Odczyt

Aby wykonać odczyt kontroler pamięci musi przekształcić adres który dostał od CPU na rank, bank, wiersz i kolumnę.

- wybór rank - sygnał przesyłany przez jedną z szyn wyboru chipu
- wybór banku - jeden lub więcej bitów szyny adresowej
- wybór kolumny i wiersza - adres kolumny i wiersza po szynie adresowej

Kiedy kontroler pamięci rozwiąże adresy to następnie odpowiedni bank musi zostać **naładowany**, czyli linie bitu muszą mieć napięcie odpowiadające wartości pomiędzy 0 a 1.

Kiedy odpowiedni bank został naładowany, drugim krokiem jest aktywowanie odpowiedniego wiersza w zidentyfikowanym rank i banku poprzez ustawienie sygnału na szynie wyboru chipu aktywującego zbiór urządzeń DRAM zawierający żądany bank. Następnie adres wiersza i banku umieszczany jest na szynie adresowej, po czym ustawiamy jest pin RAS (*Row Address/Access Strobe*).

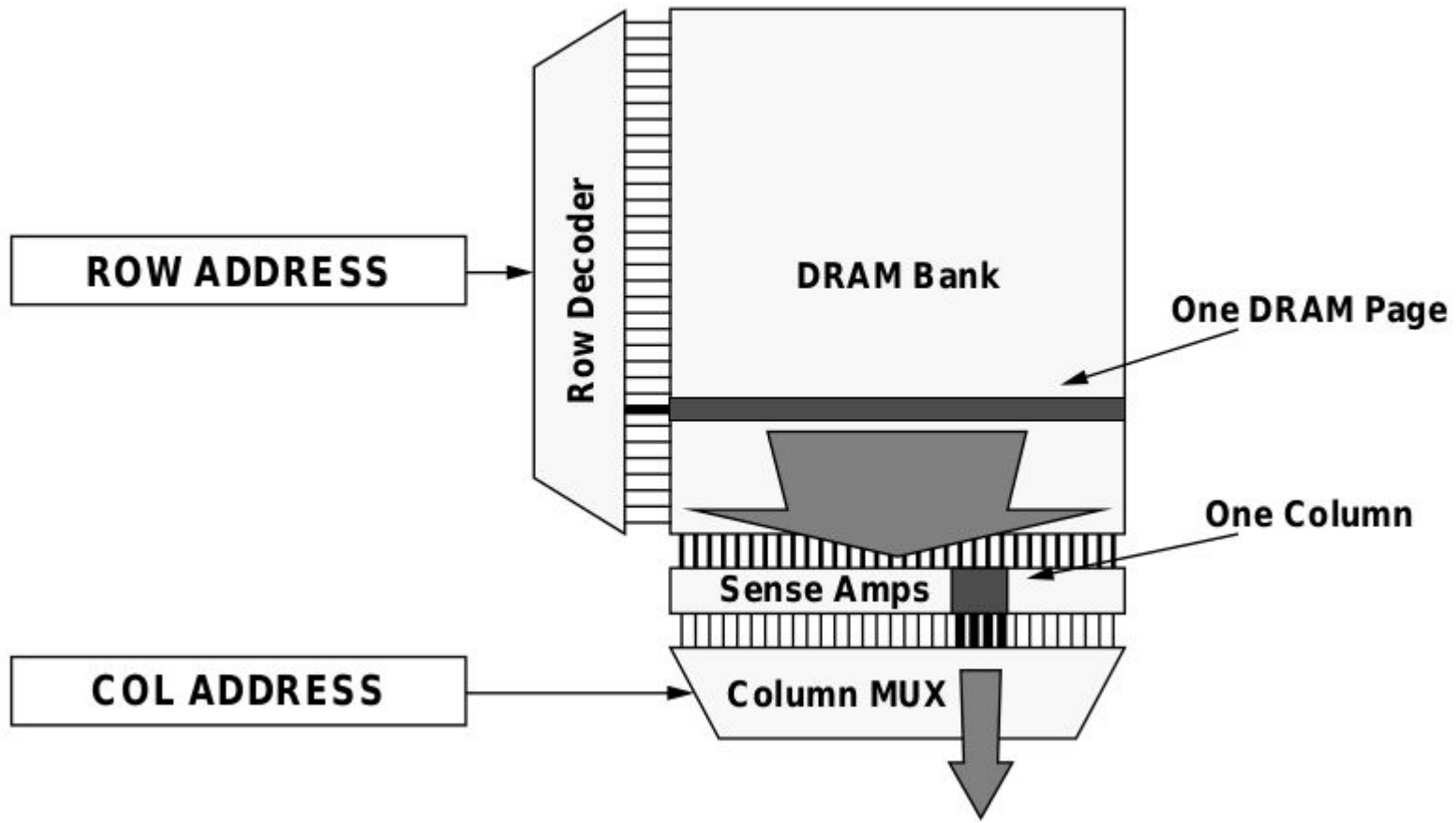
Powoduje to wysłanie całego wiersza do wzmacniacza. Wiersz jest *otwarty*.

Odczyt

Jeśli wiersz był już otwarty wcześniej, to wyżej opisane kroki pobierania wiersza nie są potrzebne i można je pominąć.

Kiedy wartości w wzmacniaczu zostały już wykryte i podniesione do napięć odpowiadających wartości 0 lub 1, pozostaje przeczytać żadaną kolumnę (podzbiór wiersza).

Szyną wyboru chipu aktywujemy zbiór urządzeń DRAM zawierających żądany bank, wysyłamy adres kolumny i identyfikator banku szyną adresową, po czym ustawiamy pin CAS (*Column Address/Access Strobe*). Powoduje to wysłanie żądanych danych szyną danych do kontrolera pamięci.



Wzmacniacz odpowiada za wykrycie ładunku (bitu informacji) jak i ich wzmocnienie.

Ewolucja DRAM

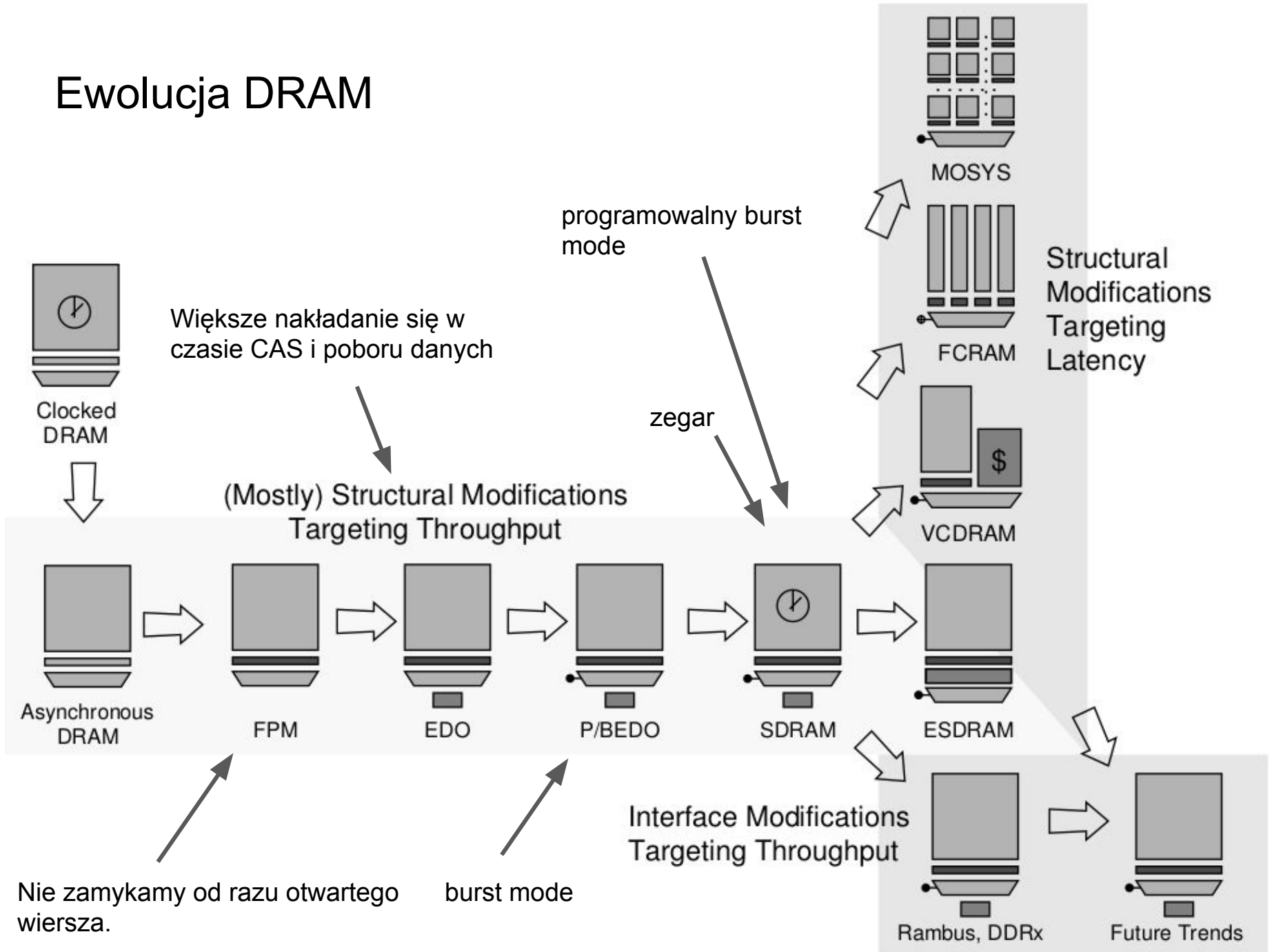
DRAM optymalizuje się głównie pod dwoma względami:

- czasu dostępu (opóźnień)
- przepustowości

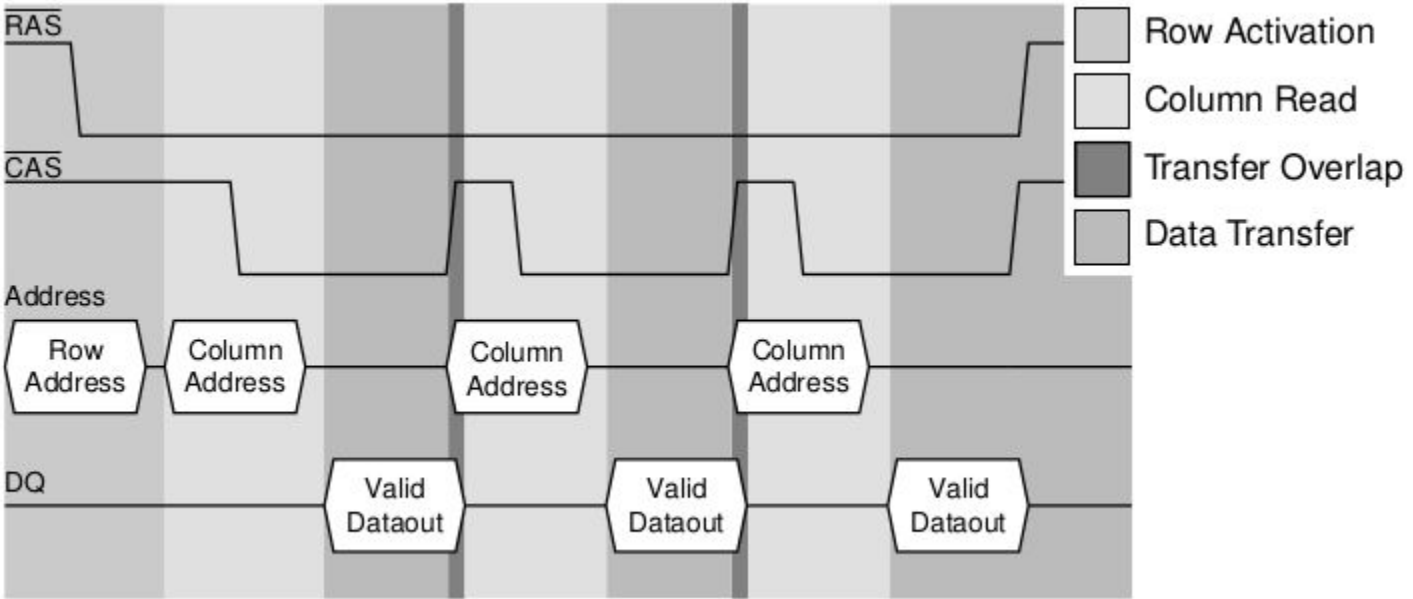
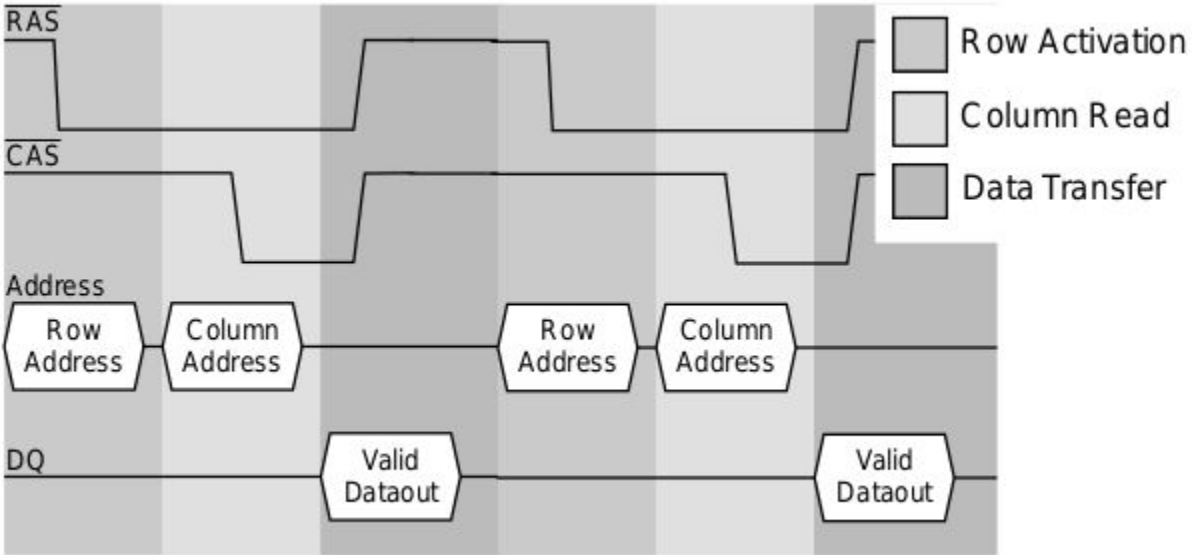
DRAM może być asynchroniczny - czas w którym zostaną wykonane pośrednie operacje dostępu zależy od asynchronicznie dostarczanych sygnałów RAS i CAS.

Synchroniczny DRAM posiada dodatkowo zegar, względem którego wykonują się wszystkie operacje, (również dostaje sygnały CAS i RAS).

Ewolucja DRAM

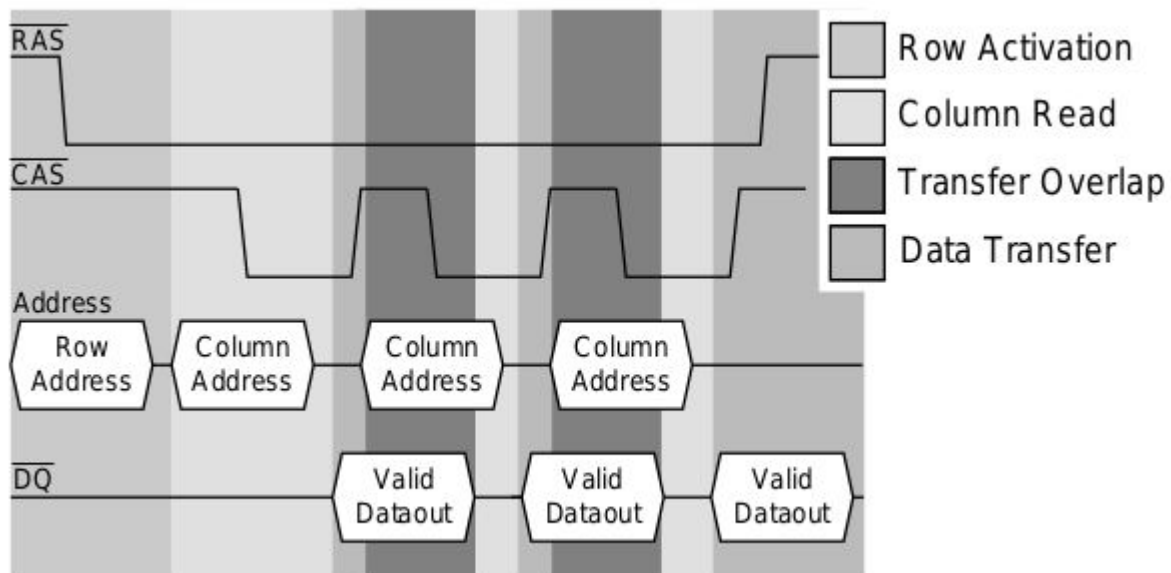


Asynchroniczny DRAM



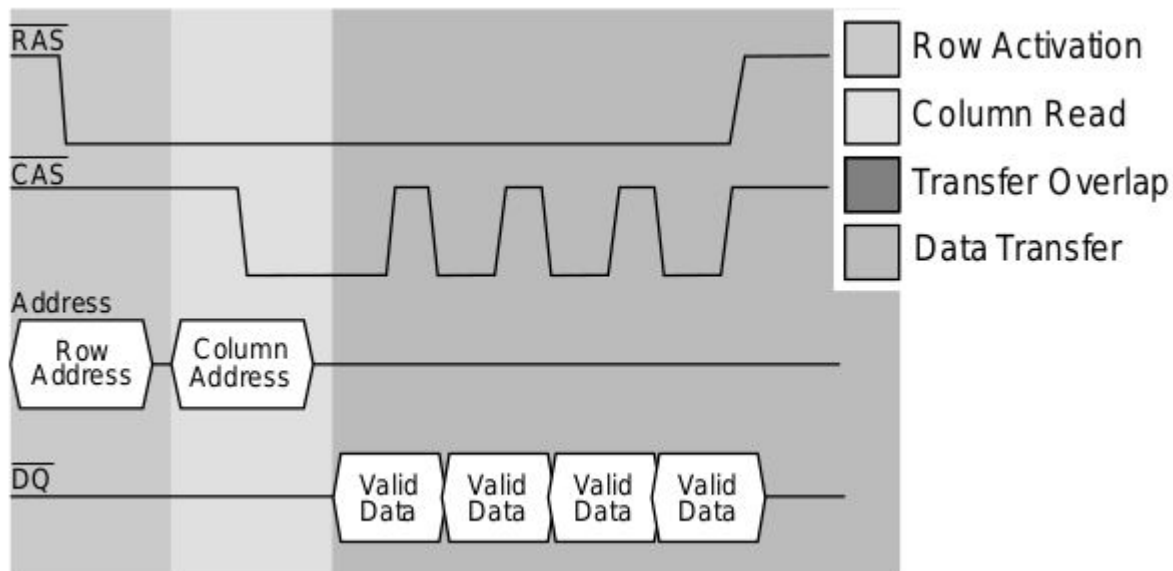
FPM DRAM - Fast Page Mode DRAM,

Zapamiętuje adres wiersza pomiędzy odczytami. Przydatne przy wielokrotnym czytaniu z jednego wiersza.



EDO DRAM - Extended Data-Out DRAM.

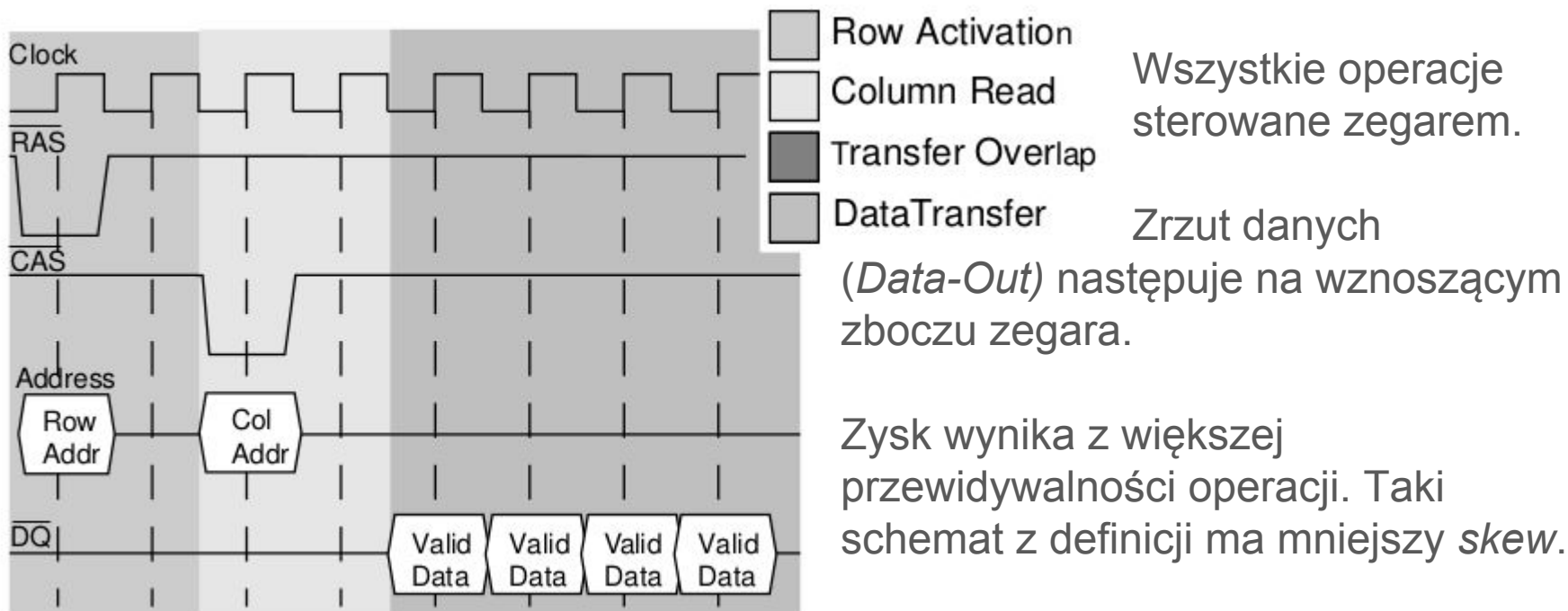
Pozwala na większe nakładanie się procesu wyboru kolumny i zrzutu danych.



BEDO DRAM - Burst-Mode EDO DRAM.

Burst mode kontrolowany przez CAS

SDRAM



Redukcja *skew* oznacza potencjalnie szybsze przetwarzanie zapytań przez system, co oznacza większą przepustowość.

Dodanie zegara pozwoliło wyeliminować także sygnały odpowiedzialne za synchronizację w czasie.

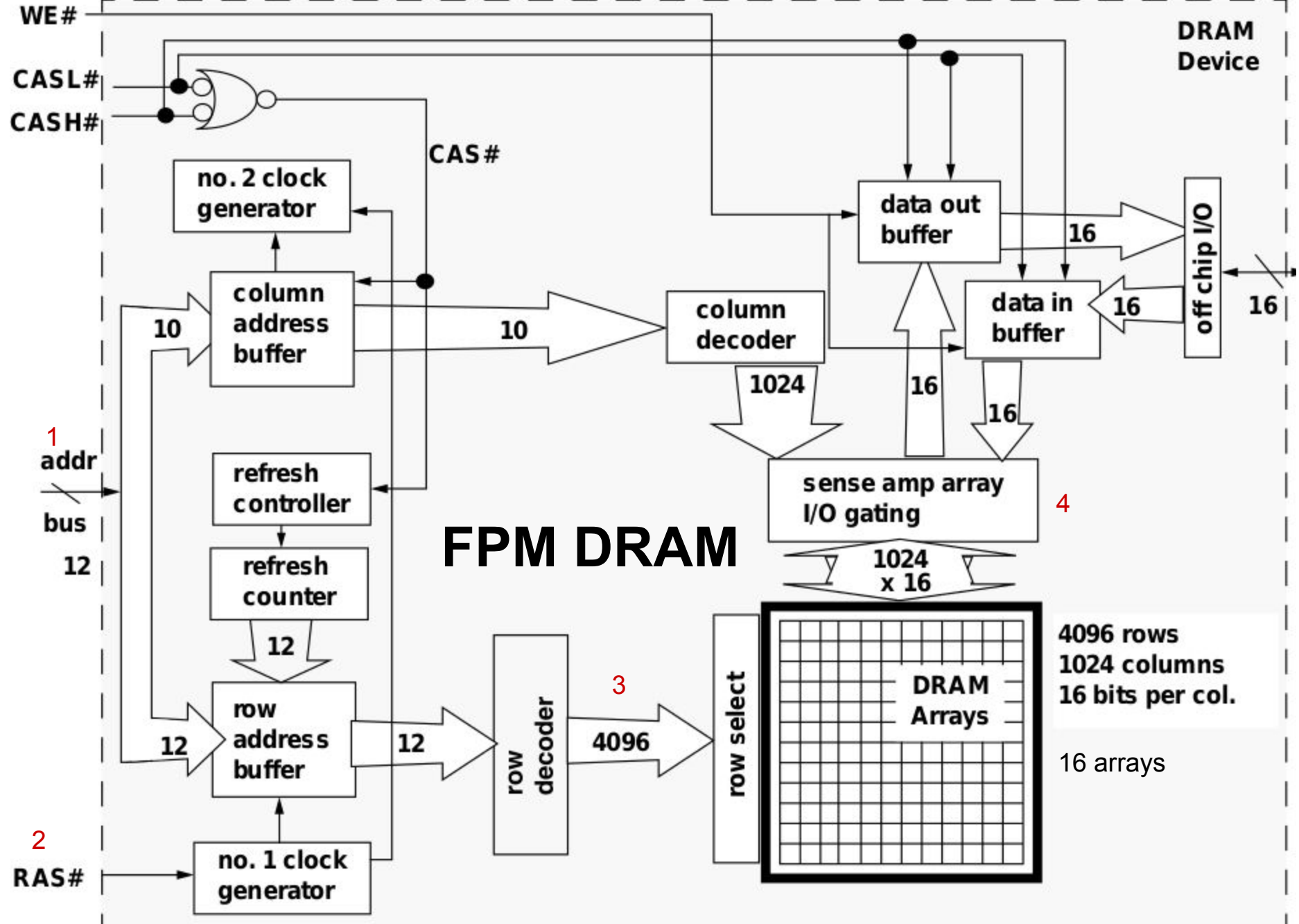
DDR - Dual Data Rate

DDR SDRAM - Data-Out na wznoszącym i opadającym zboczu zegara.



Organizacja urządzenia DRAM

WRITE ENABLE

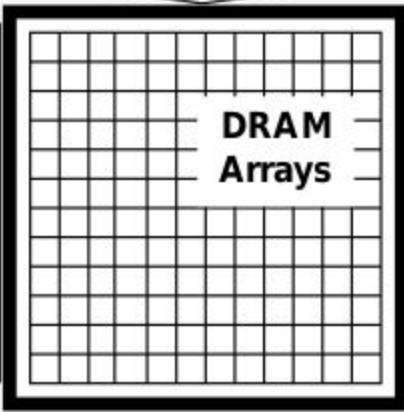


FPM DRAM

DRAM Device

4096 rows
 1024 columns
 16 bits per col.

16 arrays



1 addr bus 12

2 RAS#

CAS#

WE#
CASL#
CASH#

no. 2 clock generator

column address buffer

refresh controller

refresh counter

row address buffer

no. 1 clock generator

column decoder

data out buffer

data in buffer

sense amp array I/O gating

1024 x 16

row decoder

row select

DRAM Arrays

off chip I/O

10

10

1024

16

16

12

12

12

12

3

4096

16

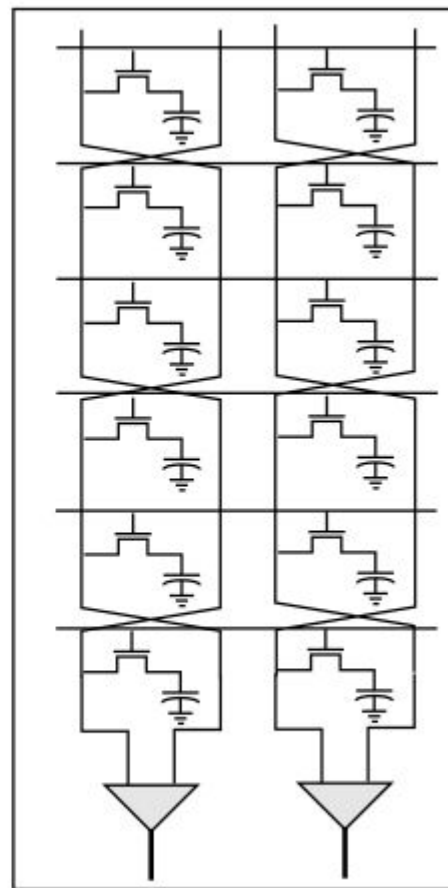
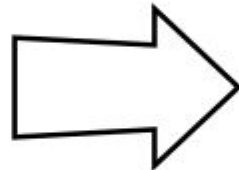
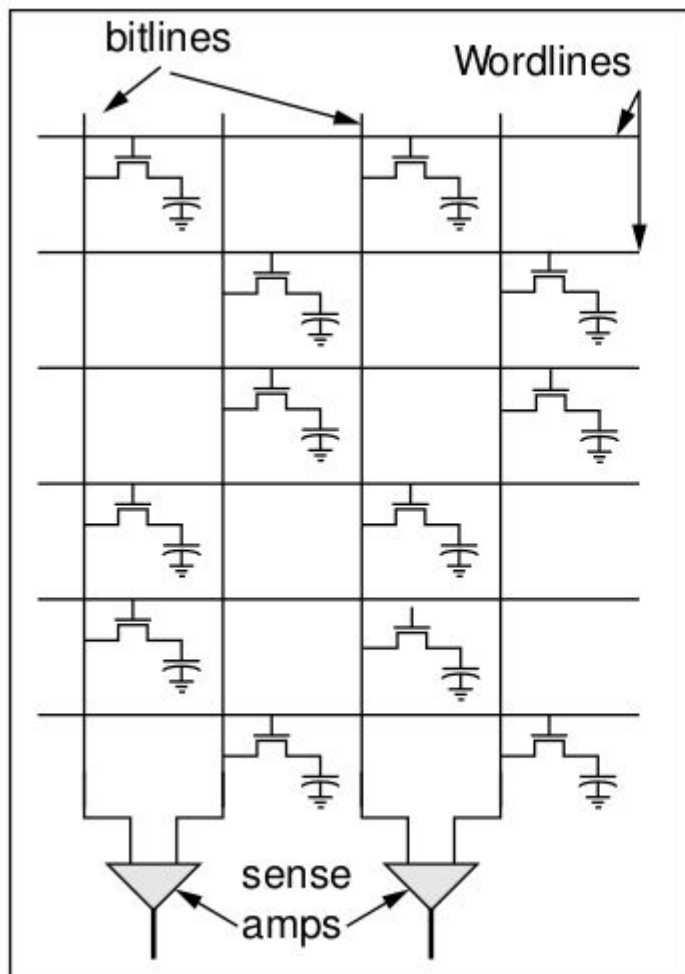
16

16

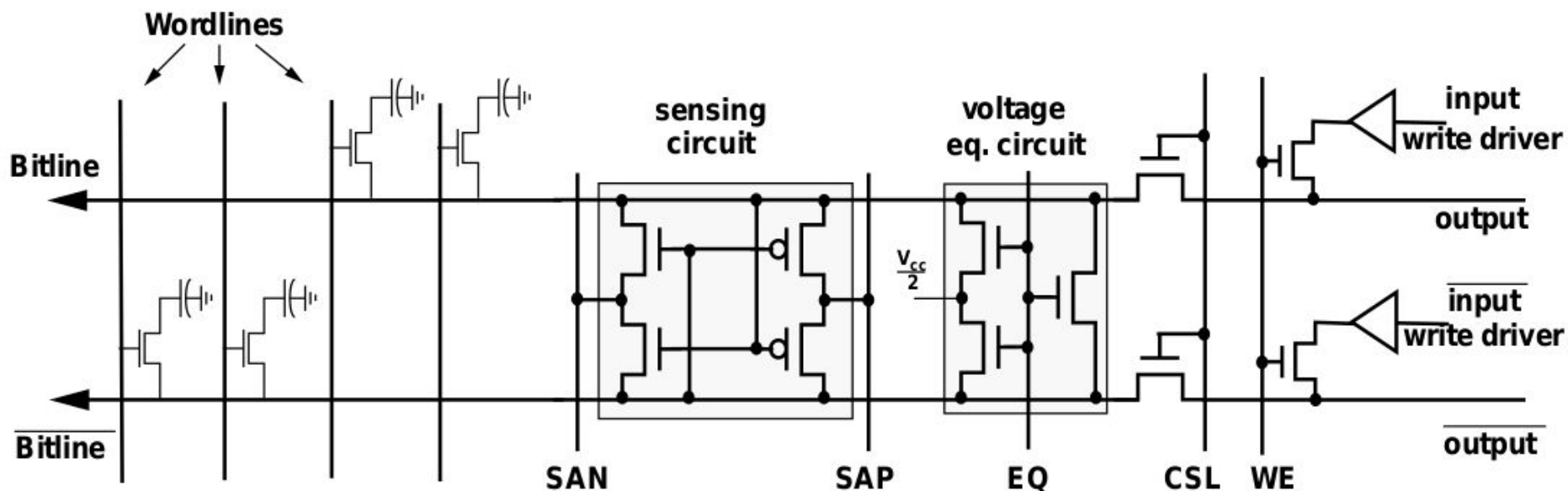
16

4

Jeden ze sposobów organizacji tablicy pamięci

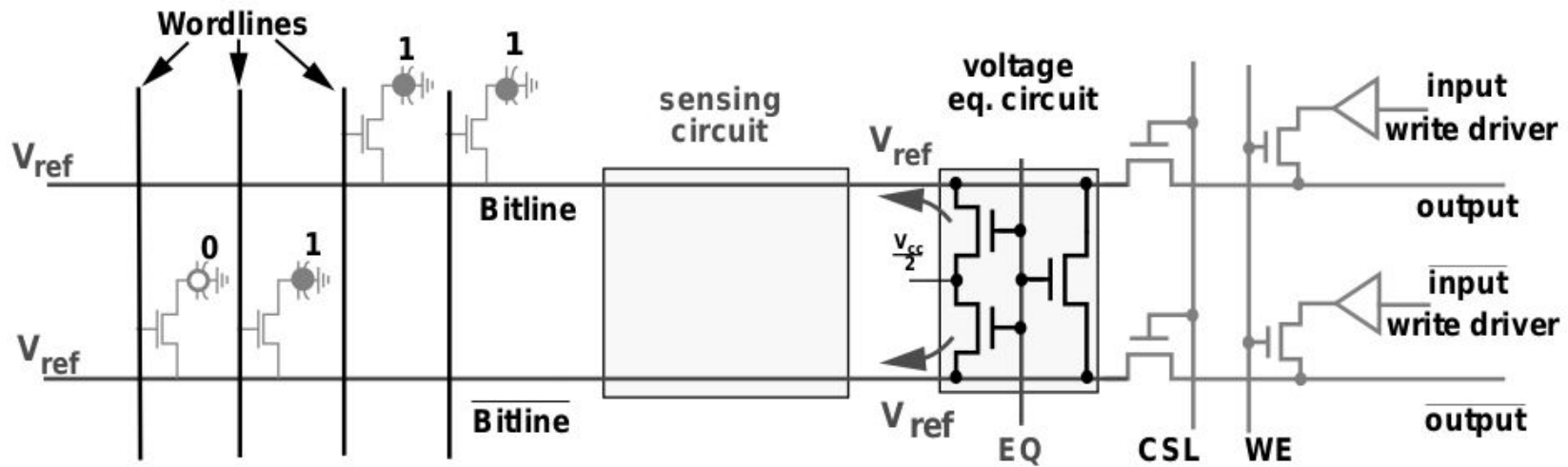


Sense amplifier

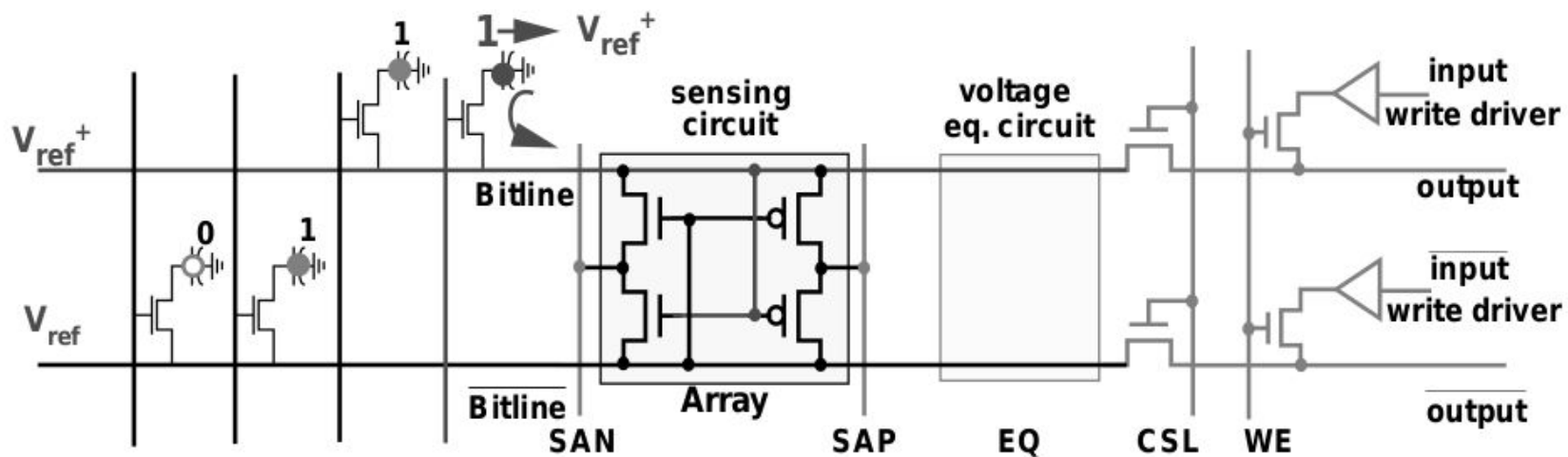


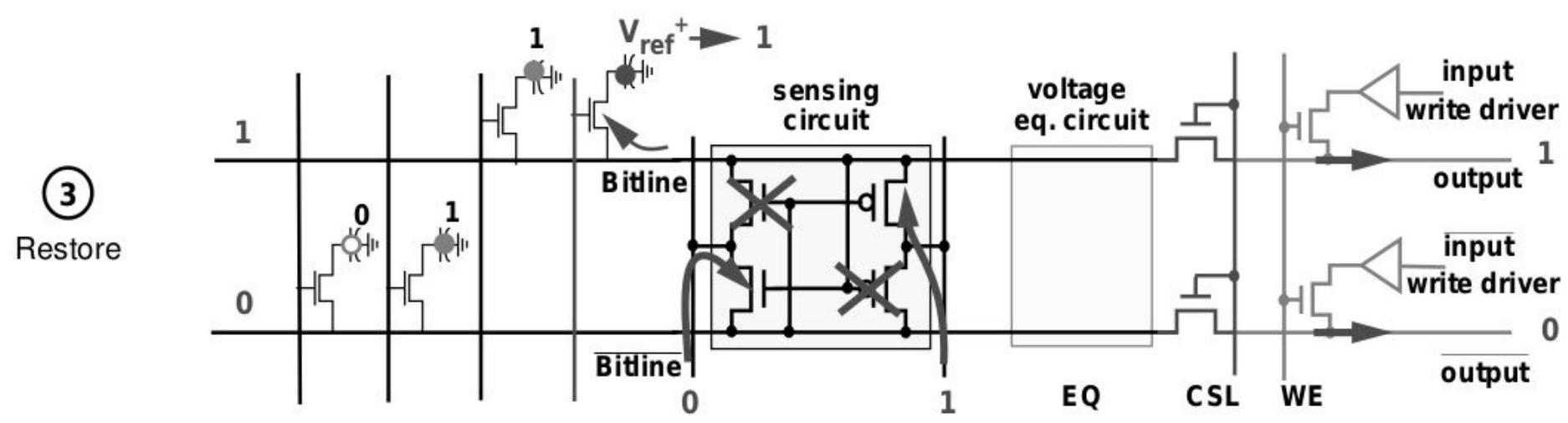
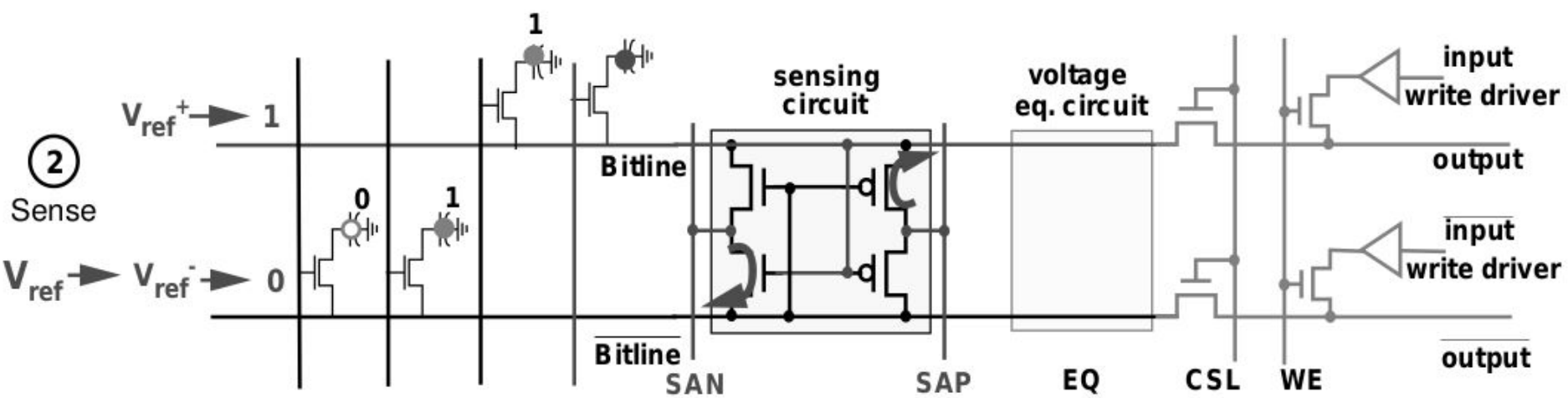
Rolą wzmacniacza jest wykrywanie ładunków w kondensatorach, oraz ich wzmacnianie.

0
Precharge

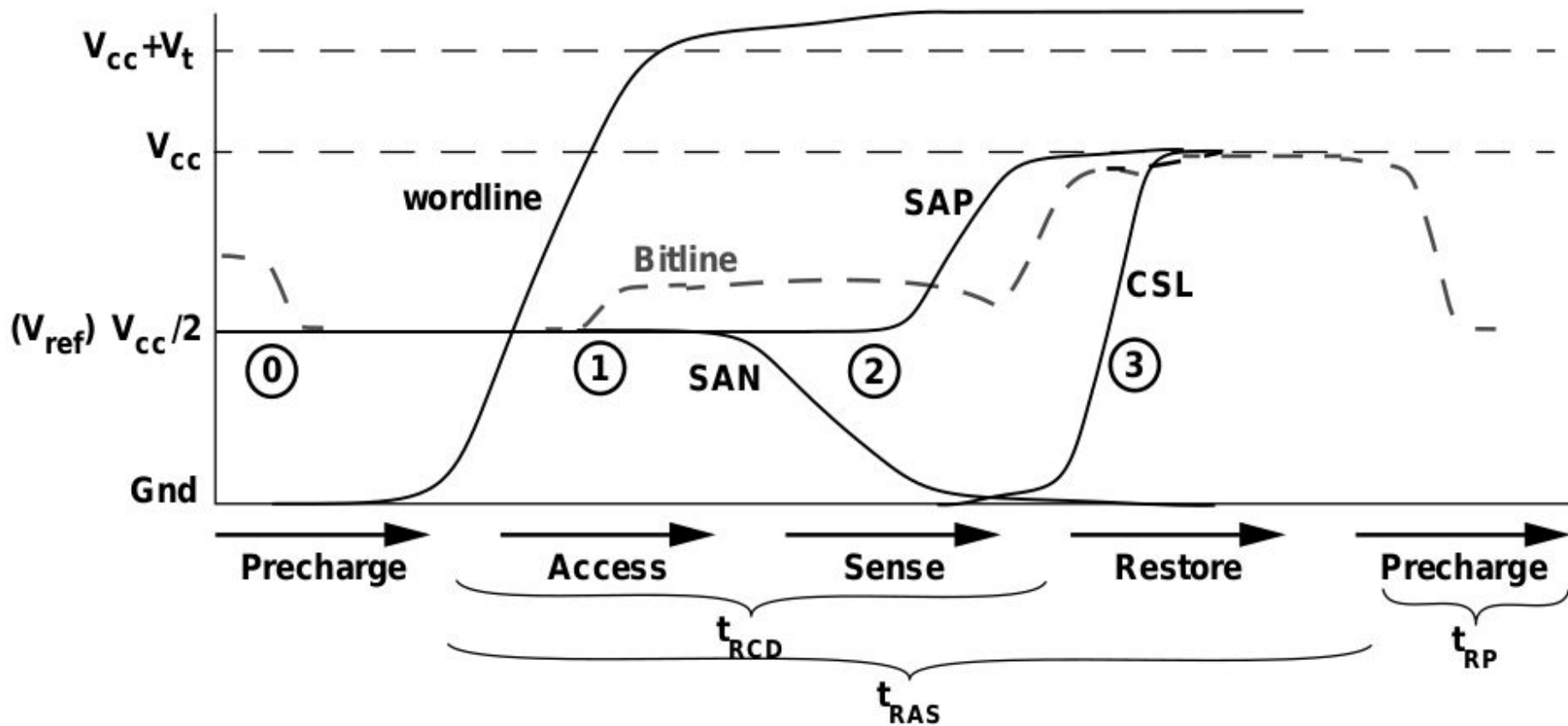


1
Access

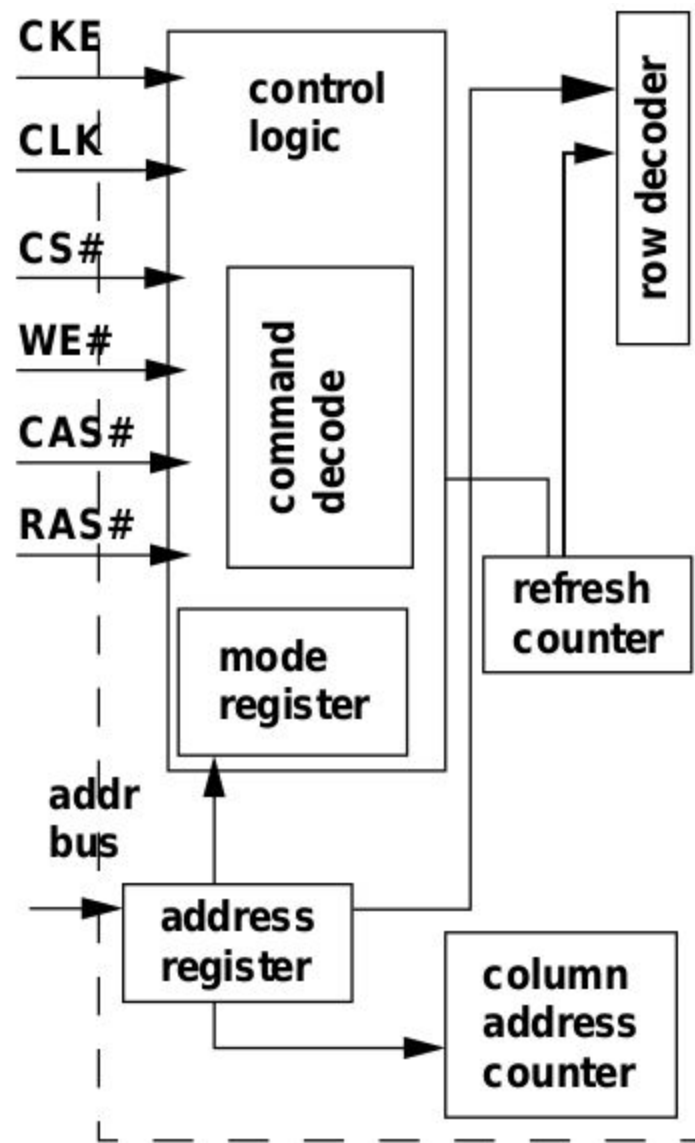




Napięcia przy odczycie we wzmacniaczu

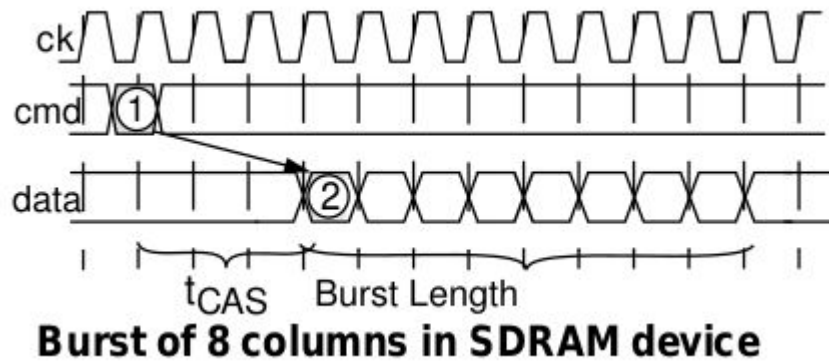


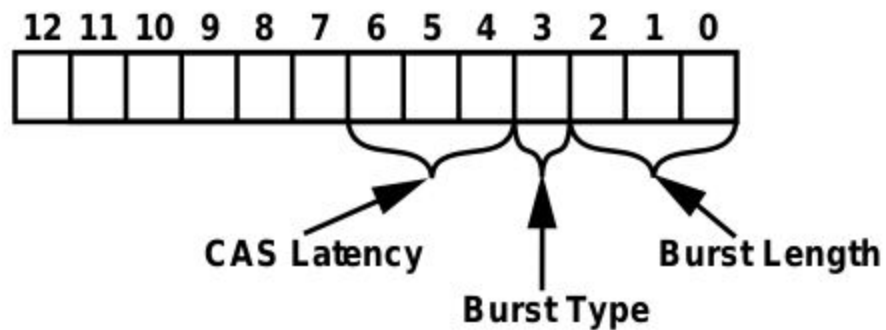
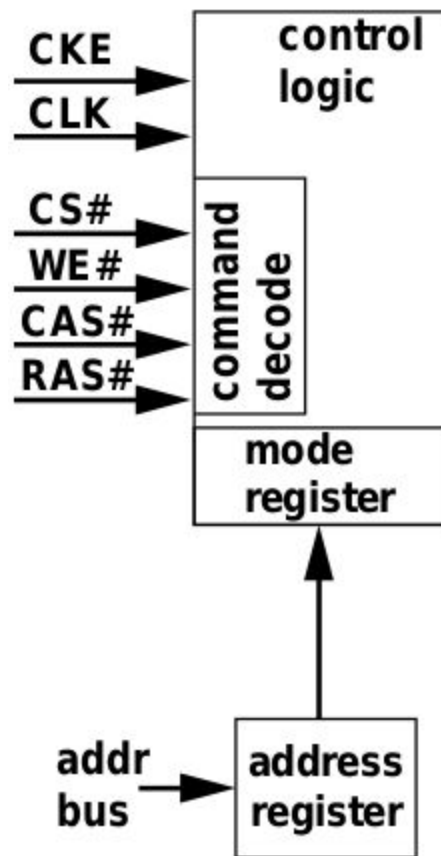
SDRAM control logic



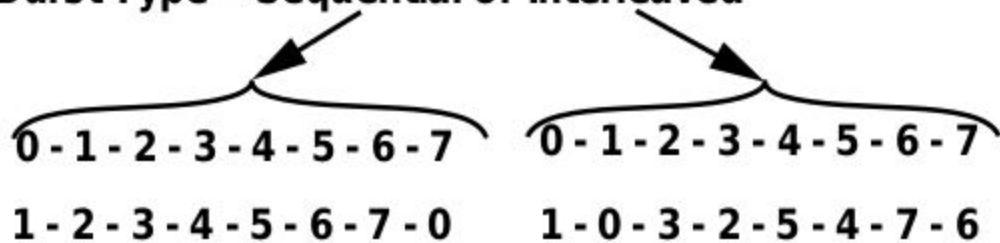
Nowoczesne urządzenia DRAM są zaimplementowane jako maszyny stanu zależne od sygnałów kontrolnych oraz od **rejestru trybu**.

W SDRAM w rejestrze trybu zapisać można opóźnienie CAS (t_{CAS} - ile cykli czekać pomiędzy komendą odczytu kolumny a zrzutem danych), typ odczytu burst (w jakiej kolejności), długość odczytu burst (ile kolumn).

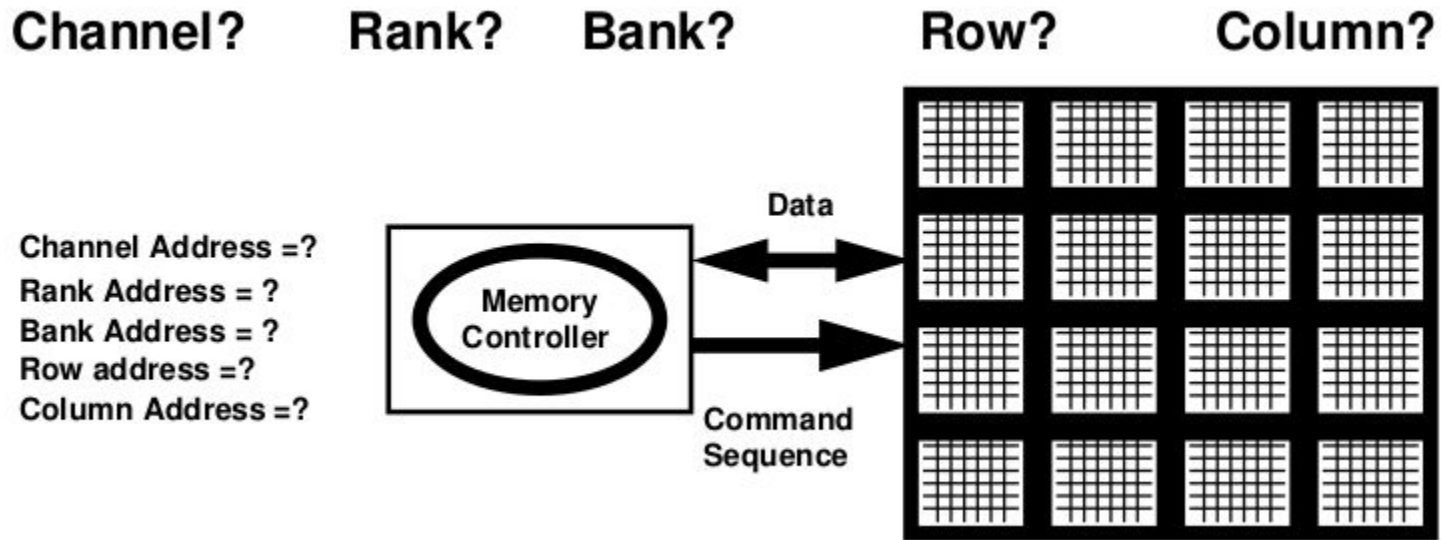




Burst Length = 1, 2, 4, 8, or Page mode
CAS Latency = 2, 3 (4, 5, etc. in special versions)
Burst Type = Sequential or Interleaved

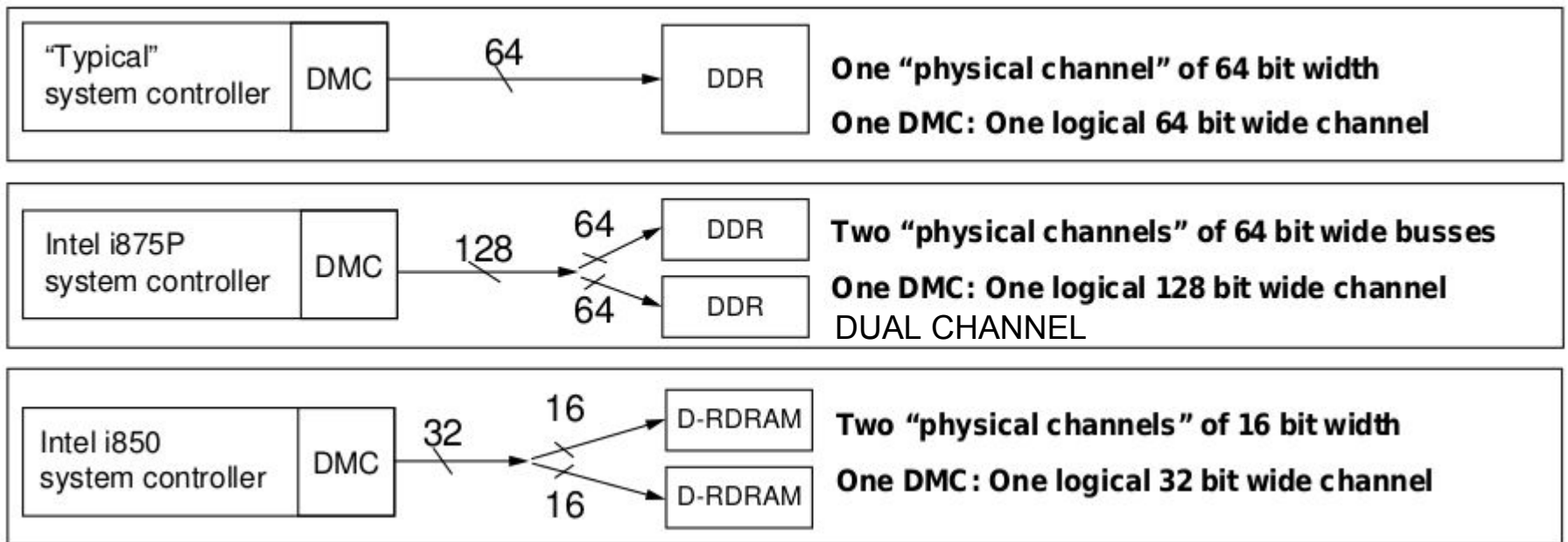


Organizacja systemu pamięci



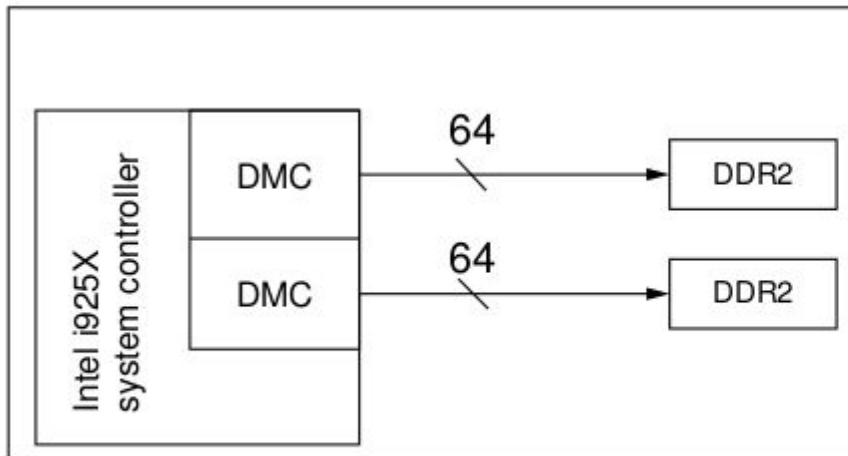
Kontroler pamięci:

- przyjmuje żądania odczytu / zapisu
- tłumaczy adres z liniowej przestrzeni na krotkę: kanał, rank, bank, wiersz, kolumna
- zajmuje się komunikacją z urządzeniami DRAM (odpowiednie wysłanie sygnałów)
- pobiera dane i przesyła z powrotem do żądającego.

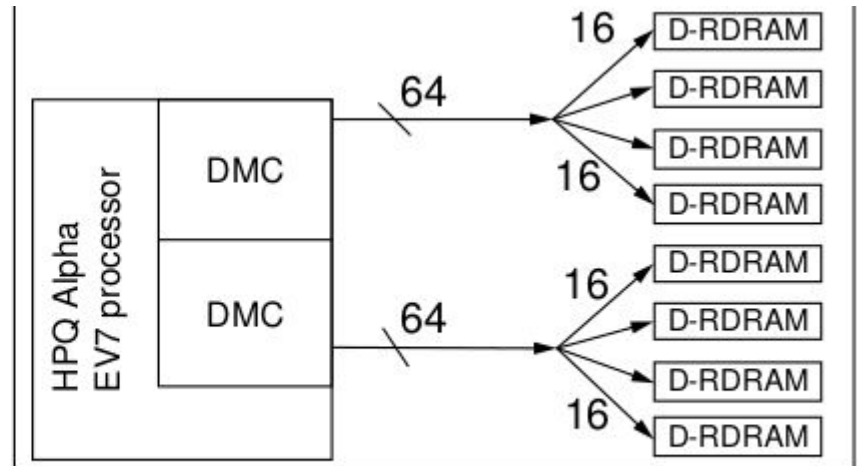


Systemowy kontroler pamięci zawiera DMC (*DRAM memory controller*). DMC kontroluje jeden kanał. Kontroler pamięci Intel i875P posiada kanał 128 bitowy, ale moduły pamięci zazwyczaj oferują 64-bitowy kanał. Wymaga to złączenia 2 kanałów 64-bitowych w jeden 128-bitowy (Dual Channel). Tylko jeden DMC więc moduły pracują współbieżnie (*in lockstep*) a nie równolegle. Logicznie jest to jeden kanał 128-bitowy, fizycznie 2 64-bitowe. Jest asymetryczny.

Więcej DMC -> większe zrównoleglenie. Kanały pracują w pełni równoległe.

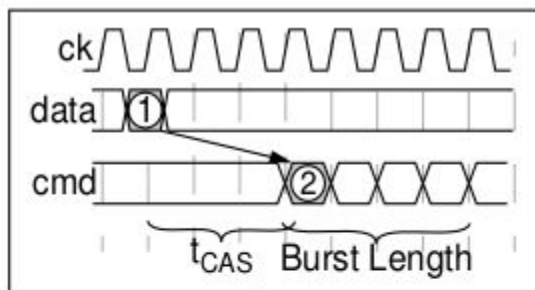


Two Channels: 64 bit wide per channel



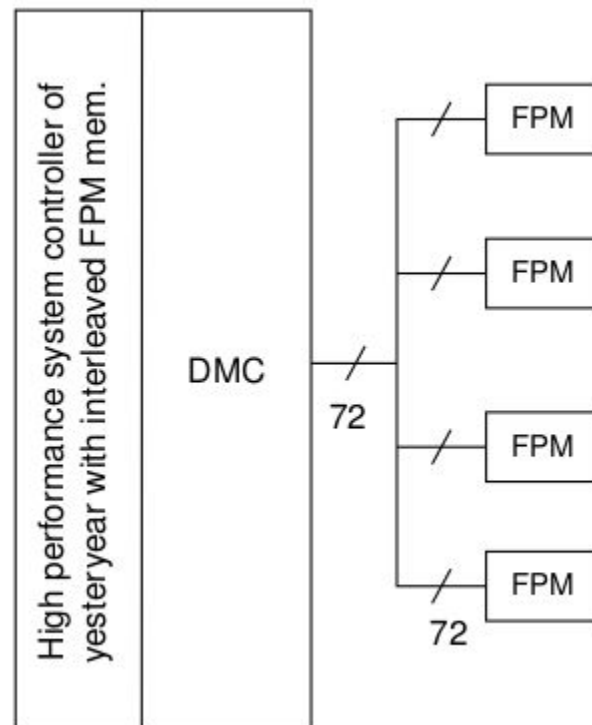
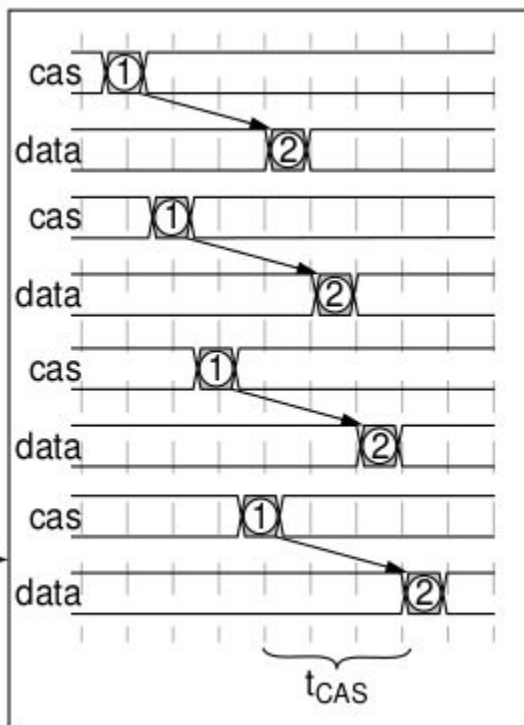
Two Channels: 64 bit wide per channel

SDRAM a FPM DRAM



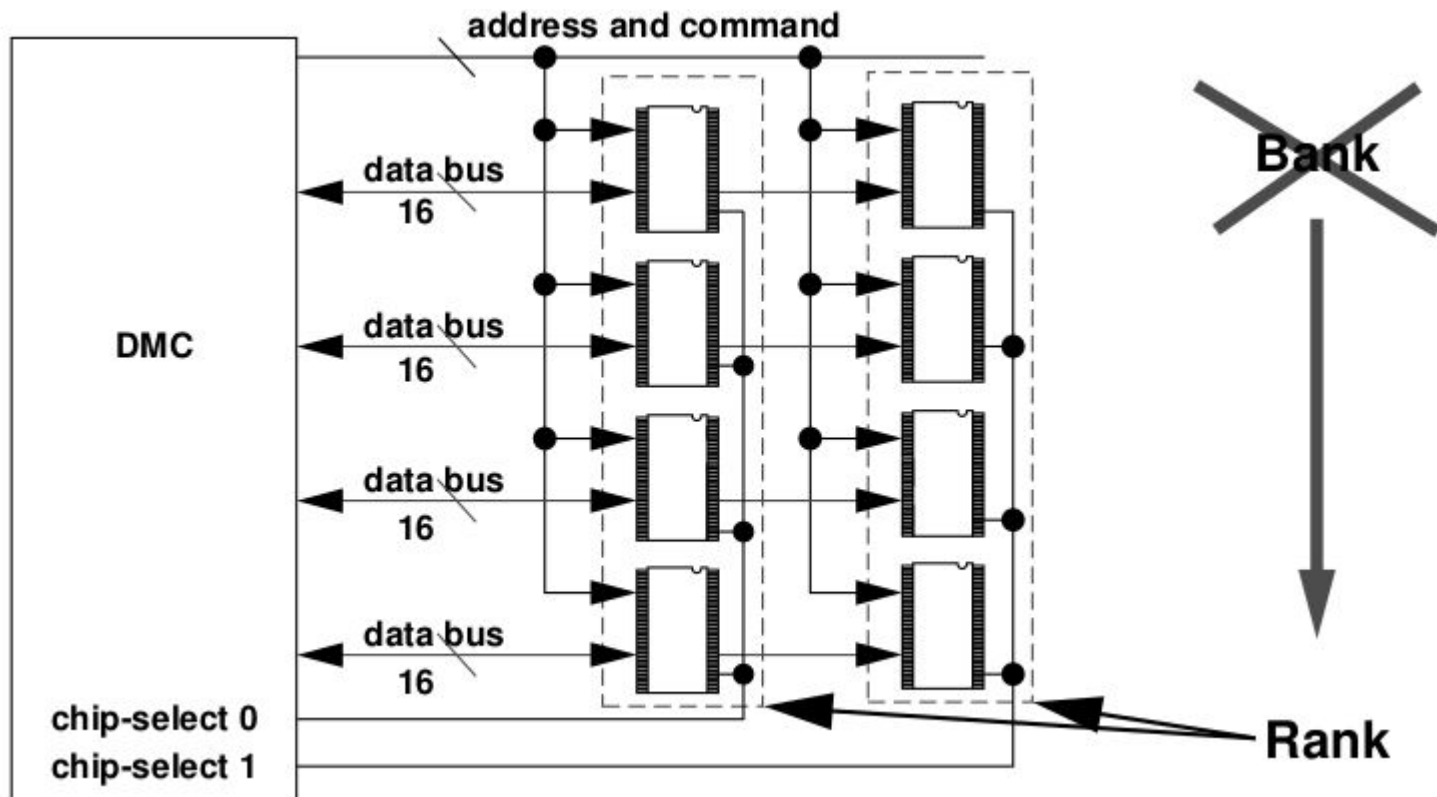
SDRAM memory bursts multiple columns of data (2) for each column access command (1).

FPM DRAM returns one column of data (2) for each column access command (1). Column accesses cannot be pipelined. Solution: stagger column accesses to different physical channels of FPM DRAM devices



Z jednym DMC możemy potokować dostępy do wielu kolumn FPM DRAM, (mimo że samo urządzenie FPM DRAM nie pozwala ani na burst ani na potokowanie) uzyskując efekt podobny do SDRAM burst.

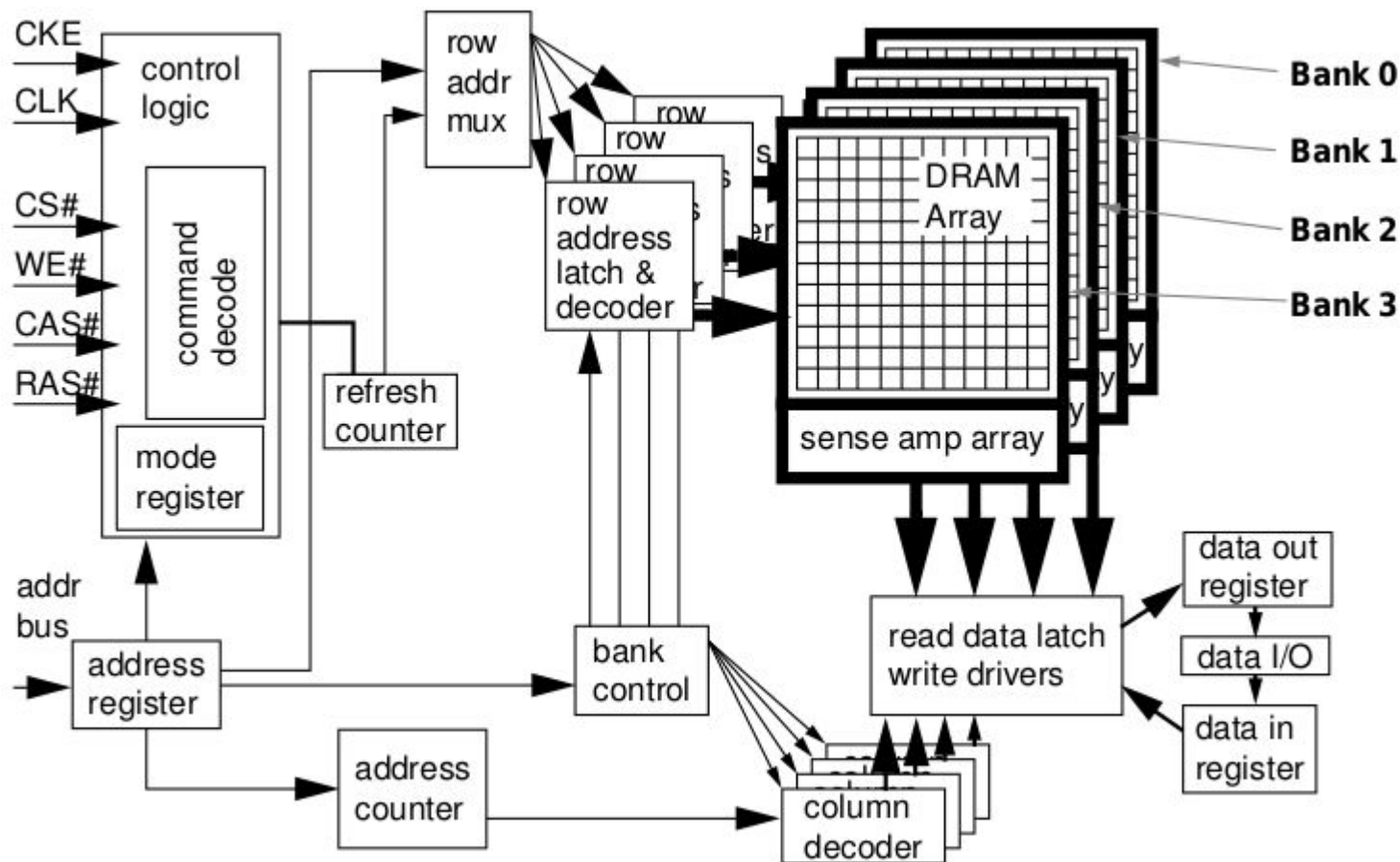
2 x RANK a DMC



Połączenie wielu urządzeń DRAM aby utworzyć szynę odpowiedniej szerokości.

Szyną wyboru chipu wybieramy odpowiedni rank który odpowiada na komendę systemu pamięci.

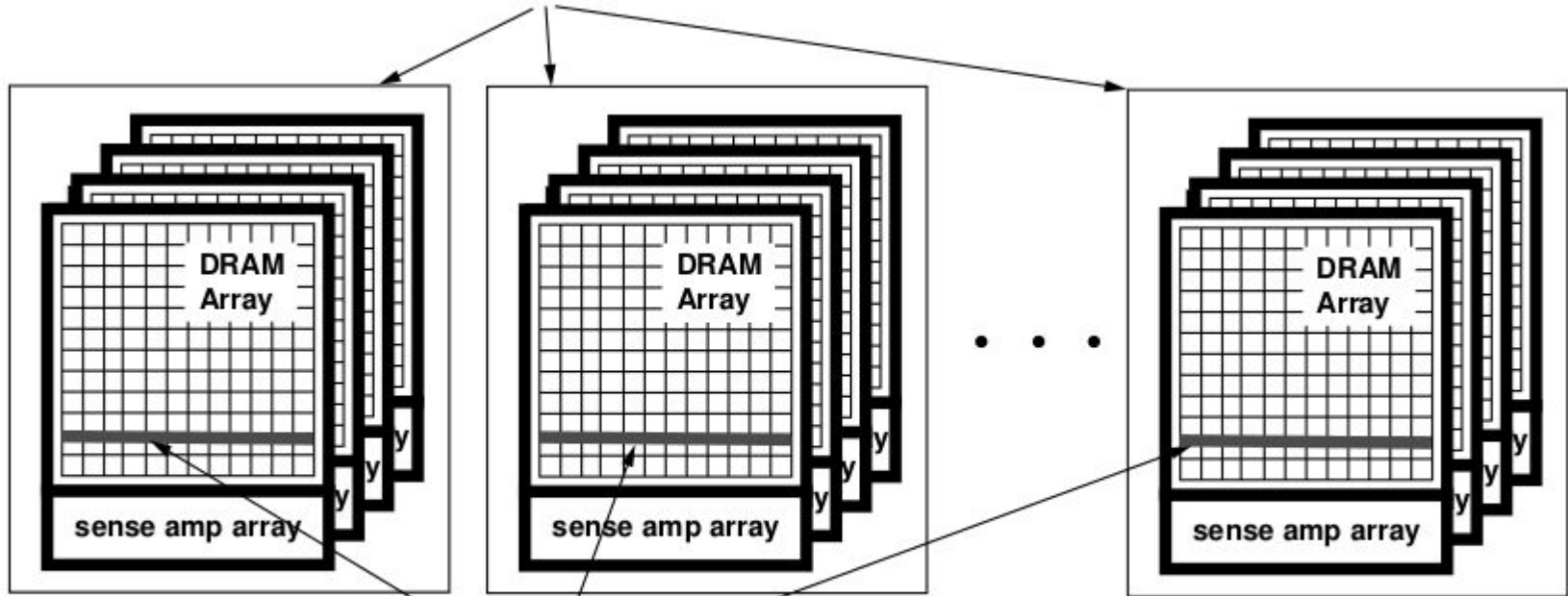
SDRAM z 4 niezależnymi bankami



Tablice pamięci mogą pracować współbieżnie bądź równolegle, w zależności od tego czy dana część układu urządzenia DRAM jest zwielokrotniona, czy też musi być współdzielona.

Równoległe działające wiersze

DRAM devices arranged in parallel in a given rank



Generic DRAM devices with 4 banks, 8196 rows, 512 columns per row, and 16 data bits per column.

one row spanning multiple DRAM devices

Wiersz to grupa komórek pamięci aktywowana równoległe na komendę aktywacji wiersza. Aktywujemy “wiersze” tablic pamięci w obrębie jednego ranku. Z punktu widzenia kontrolera pamięci wiersz (strona pamięci) to długość wiersza tablicy pamięci pomnożona razy ilość urządzeń DRAM w ranku i ilość tablic w banku. Taka strona jest “kaszowana” we wzmacniaczach do ponownego odczytu.

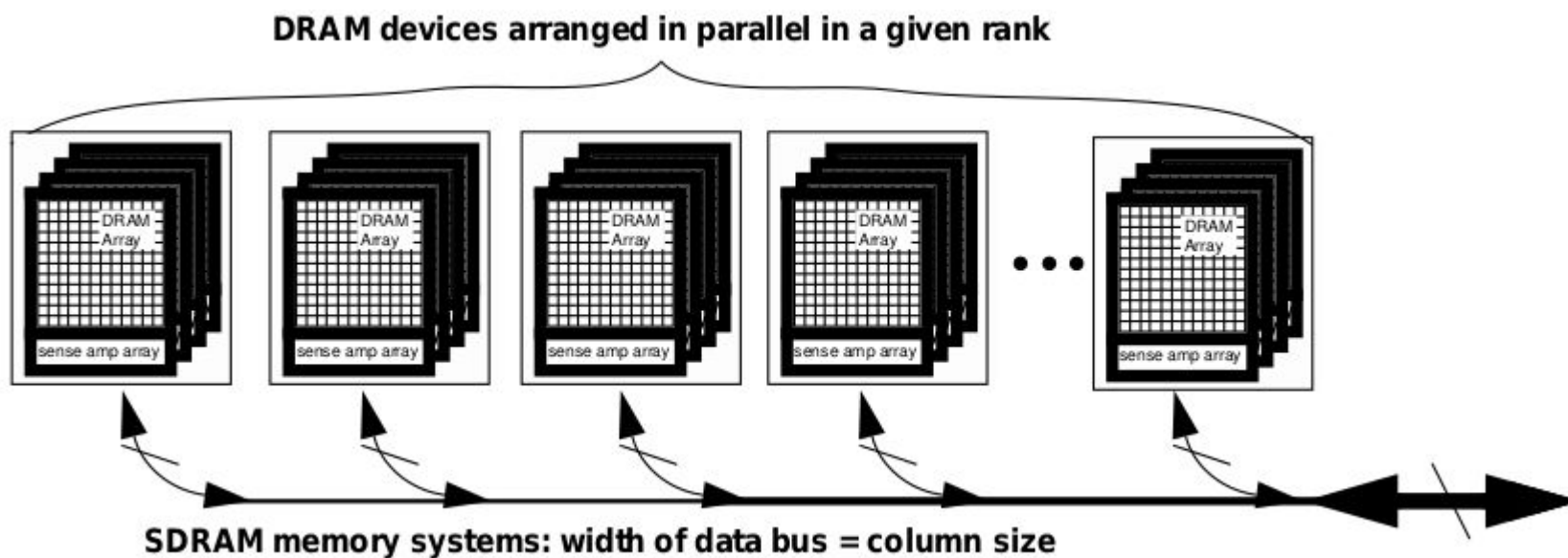
Kolumna

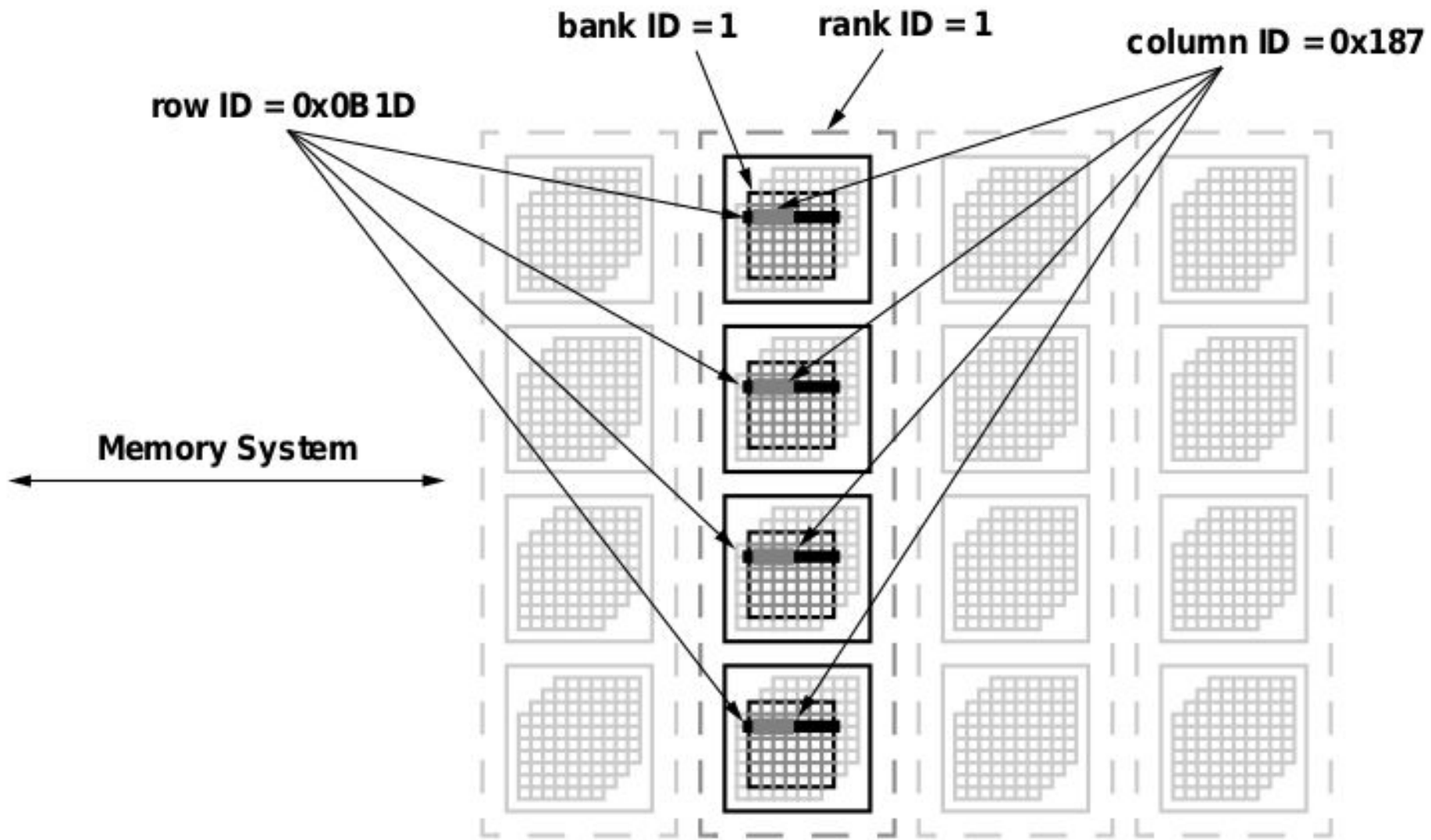
Kolumna to najmniejsza adresowalna jednostka pamięci w DRAM.

Szerokość kolumny często określana jest jako szerokość szyny danych.

Jeden odczyt kolumny może aczkolwiek sprowadzić więcej danych ze względu na burst.

Szyna danych podzielona pomiędzy urządzenia DRAM.

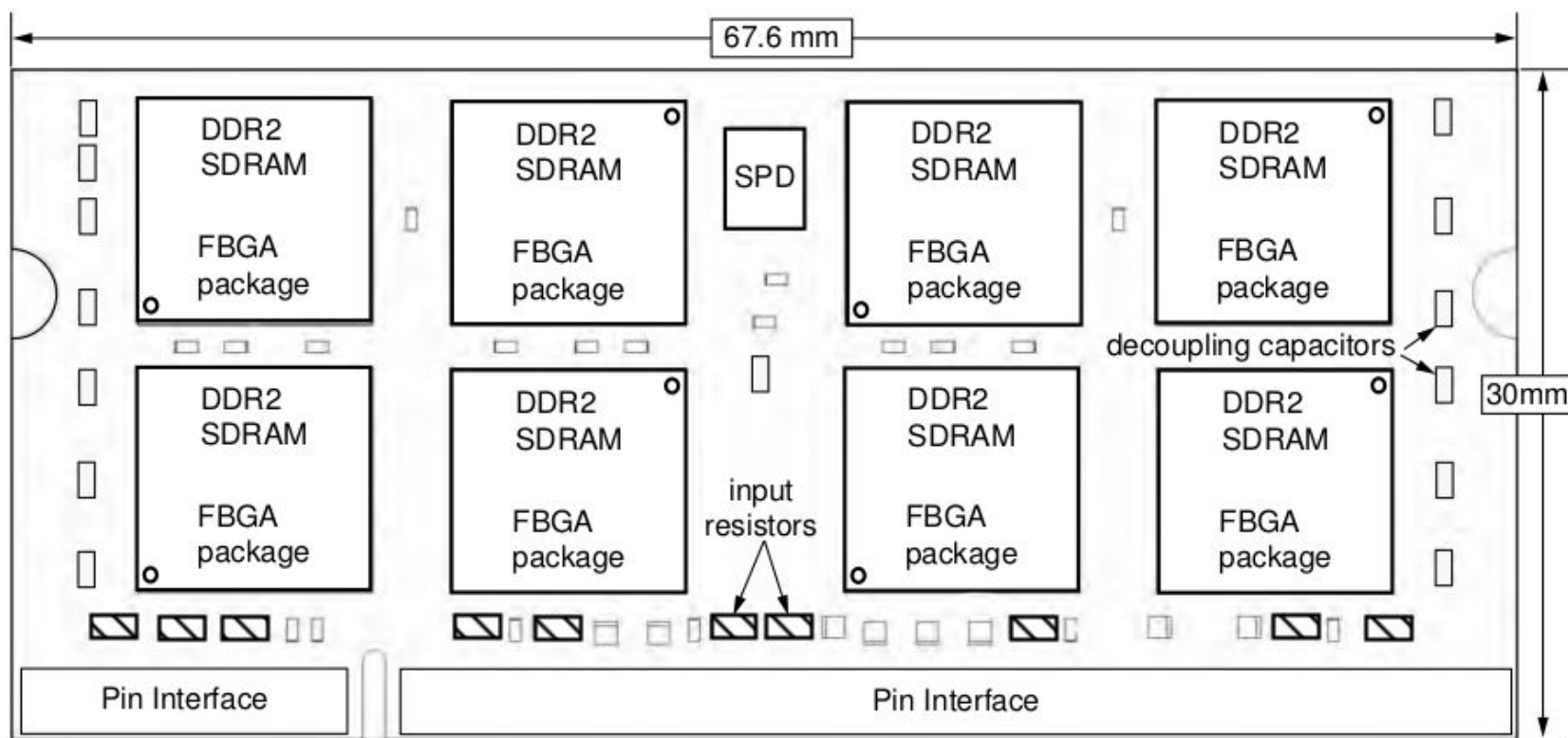




4 rank x 4 urządzenia DRAM x 4 banki x 8192 wierszy w banku x 512 kolumn w wierszu

SO-DIMM - Small Outline Dual In-line Memory Module

Specyfikacja modułu pamięci DRAM używanego w urządzeniach mobilnych (np. w laptopach).



Podstawy protokołu dostępu do pamięci

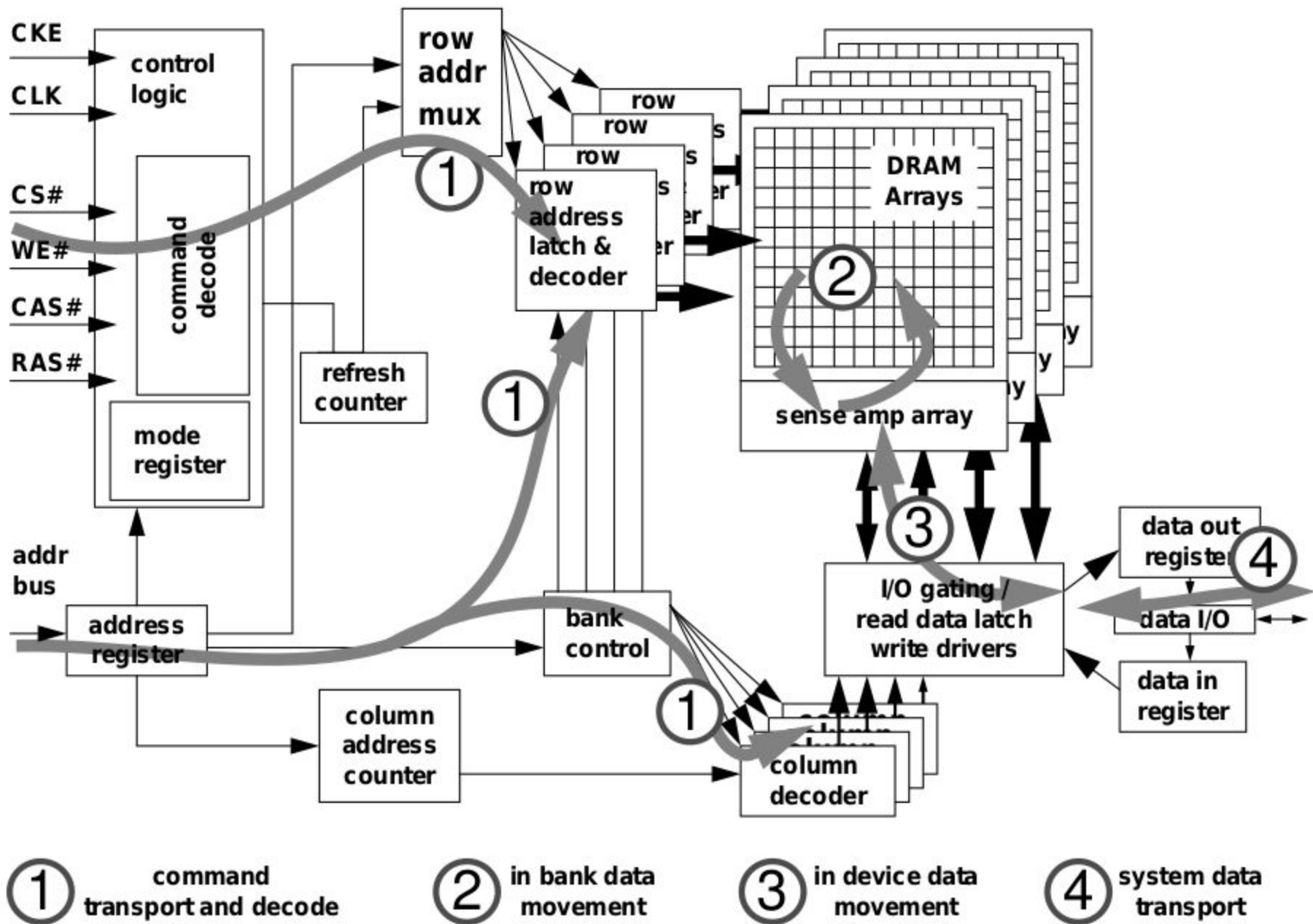


FIGURE 11.1: Command and data movement on a generic SDRAM device.

Opóźnienia pamięci

t_{RAS} Row Access Strobe. Minimalny czas między poleceniem wyboru wiersza, a przywróceniem danych w wierszu po wykonaniu operacji.

t_{RP} Row Precharge. Czas na przygotowanie innego wiersza na dostęp.

t_{RCD} Row-to-Column command Delay. Czas między wydaniem polecenia wyboru wiersza, a dostępnością danych na wyjściu z układu wzmacniającego.

t_{CAS} , t_{CL} Column Access Strobe latency. Minimalny czas między wydaniem polecenia odczytu kolumny, są początkiem transferu danych.

t_{RC} Row Cycle. Czas między dostęпами do różnych wierszy w banku.

$$t_{\text{RC}} = t_{\text{RAS}} + t_{\text{RP}}$$

Odczyt

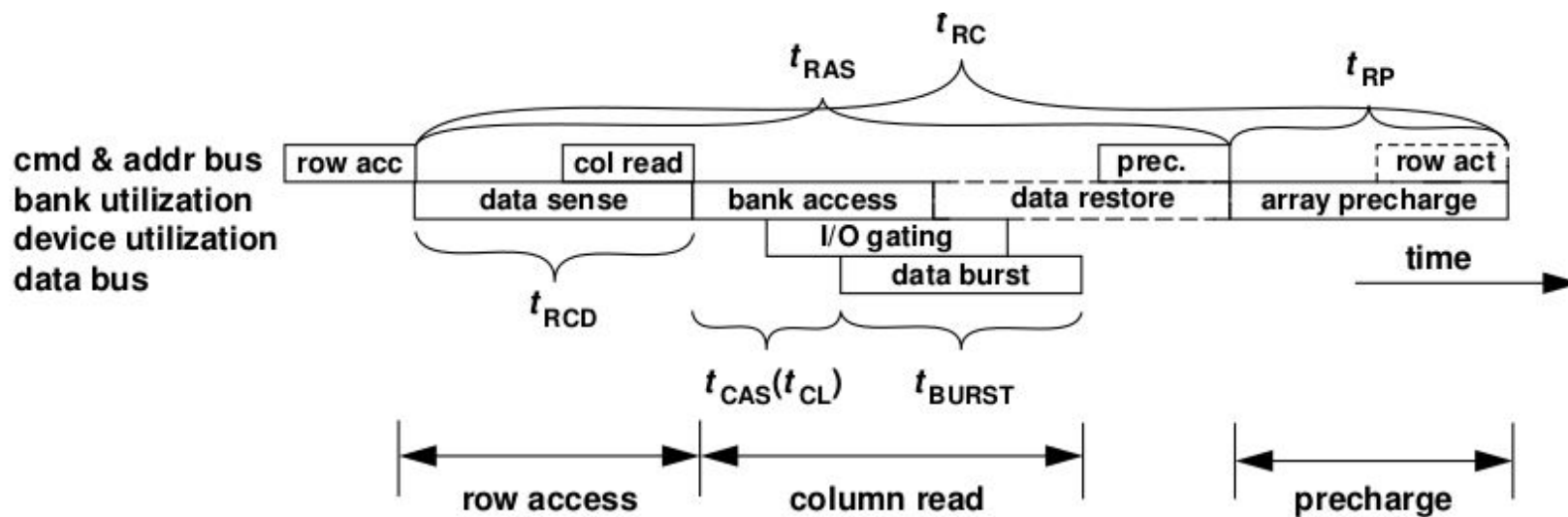


FIGURE 11.8: A read cycle.

Zapis

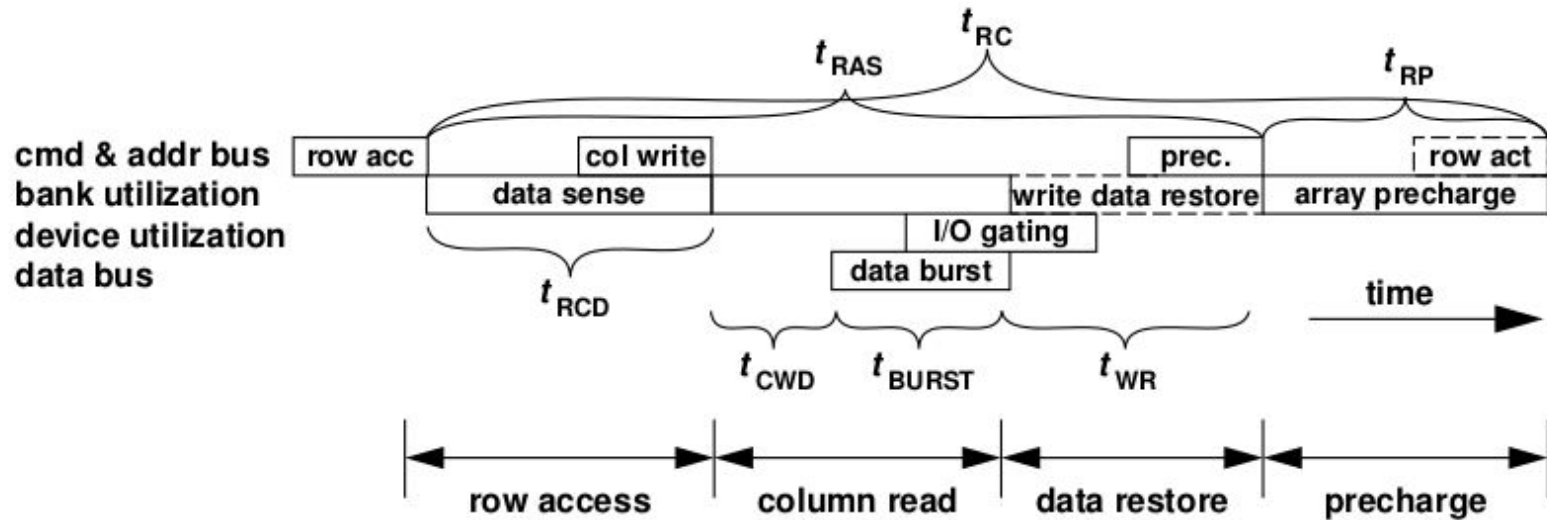


FIGURE 11.9: A write cycle.

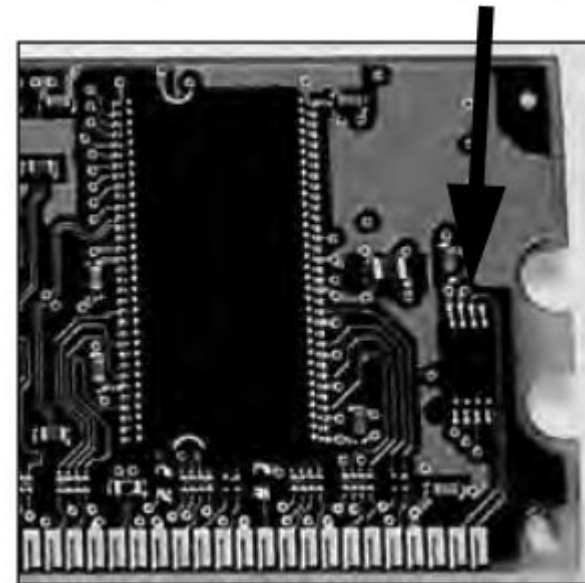
Serial Presence Detect (SPD)

W systemie pamięci może być wiele modułów pamięci o różnych parametrach (np. o różnych minimalnych opóźnieniach t_{RAS} , t_{CAS} , t_{RCD} czy t_{RP}).

Żeby usunąć tę złożoność i zapobiec konfuzji przy dokładaniu nowych modułów pamięci do systemu, przemysł komputerowy umieszcza na PCB modułu pamięci małą pamięć tylko do odczytu (SPD), w której zapisuje parametry modułu pamięci. Kontroler pamięci może odczytywać te dane i odpowiednio organizować dostępy do pamięci.

Prezentacja narzędzi: [decode-dimms](#), `lspci -v`, `lshw`, `dmidecode`

serial presence detect (SPD)



Prezentacja:

Wielokrotne otwieranie wiersza jest kosztowne - google benchmark

KONIEC

Do poczytania:

- [rowhammer](#), [Project Zero](#), [exploit1](#), [exploit2](#), [gith ub](#)

Źródła

- Memory Systems: Cache, DRAM, Disk 1st Edition, (Bruce Jakob, Spencer Ng, David Wang)
- wikipedia.org