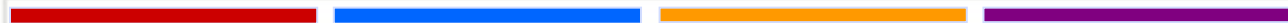


Меморија



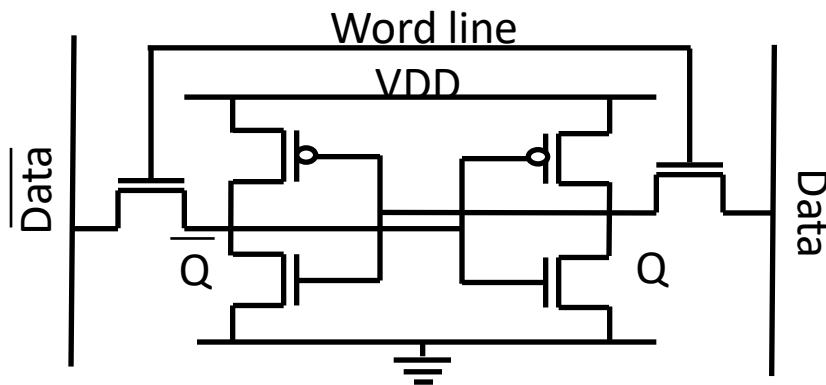
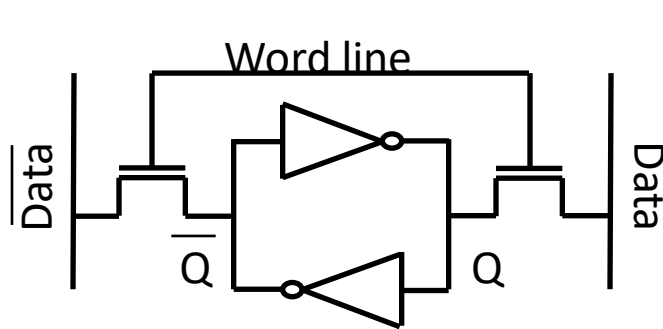
Садржај

- Технологија израде
- Интерна организација
- Циклуси на магистрали
- Потрошња енергије меморије
- Типови меморије



Технологија израде

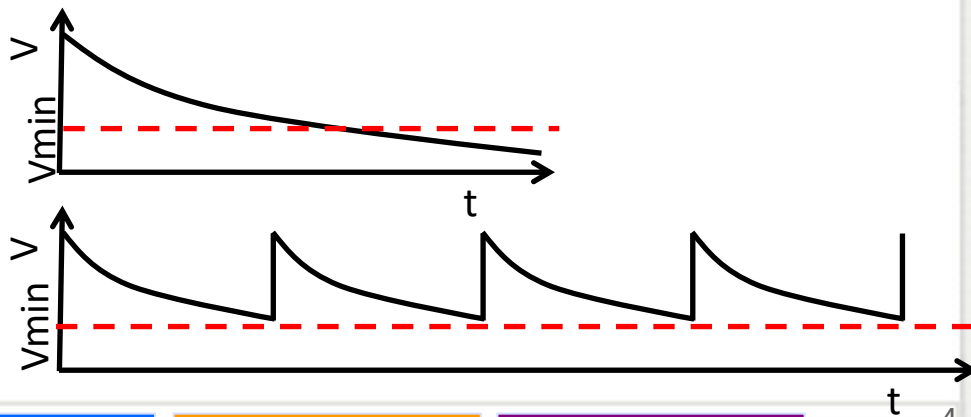
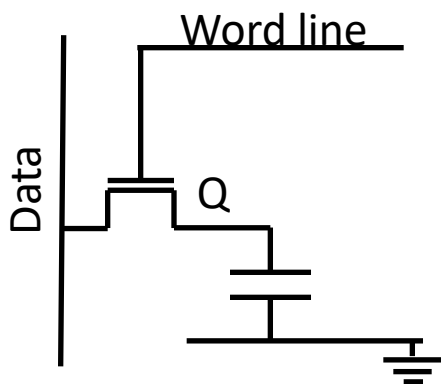
- *SRAM*
 - Потребно је мало струје за одржавање вредности
 - Потребно је 6 транзистора по биту



Технологија израде

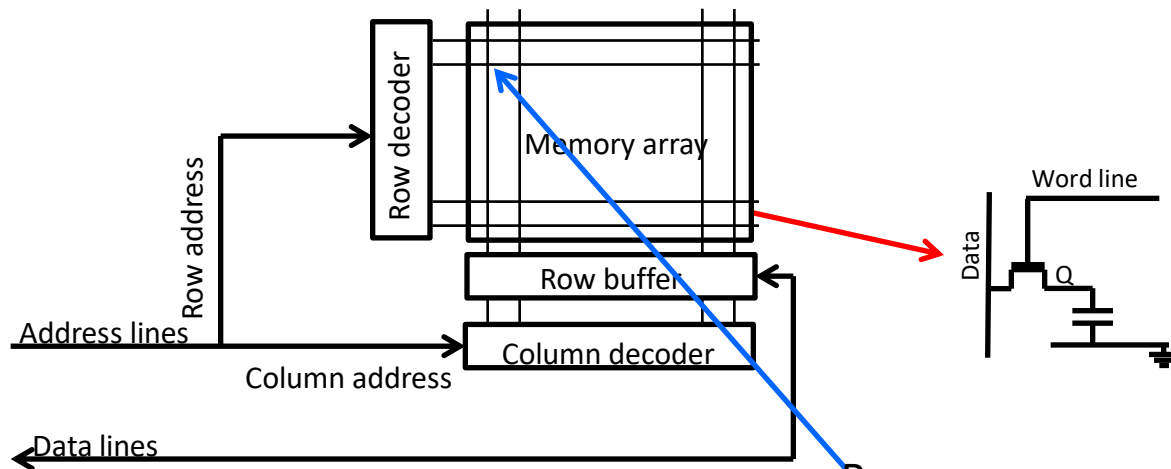
- *DRAM*

- Један транзистор по биту
- Податак се мора преписивати након читања
- Податак се мора периодично освежавати
 - На приближно $\sim 8 - 64 \text{ ms}$ (што је на 5% времена)
 - Редови се освежавају секвенцијално



Интерна организација *DRAM*

- Адресне линије се мултиплексирају:
 - Виша половина адресе – Адреса реда
 - Нижа половина адресе – Адреса колоне



Веома спора операција!

Интерна организација *DRAM*

- [Пуњење и] приступ реду

AKA: OPEN a DRAM Page/Row

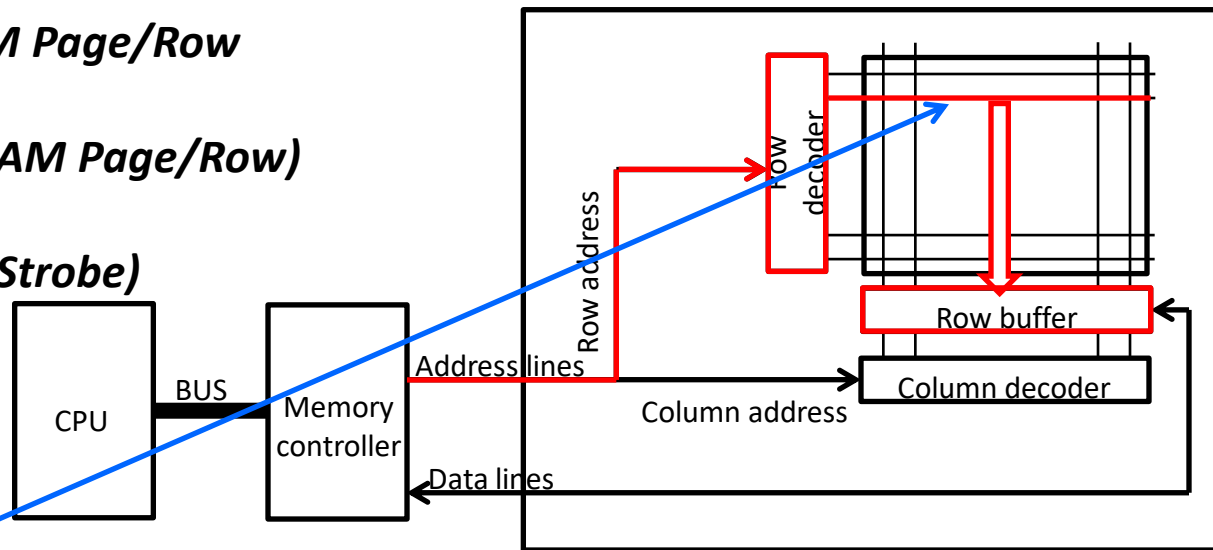
или

ACT (Activate a DRAM Page/Row)

или

RAS (Row Address Strobe)

Веома спора операција!



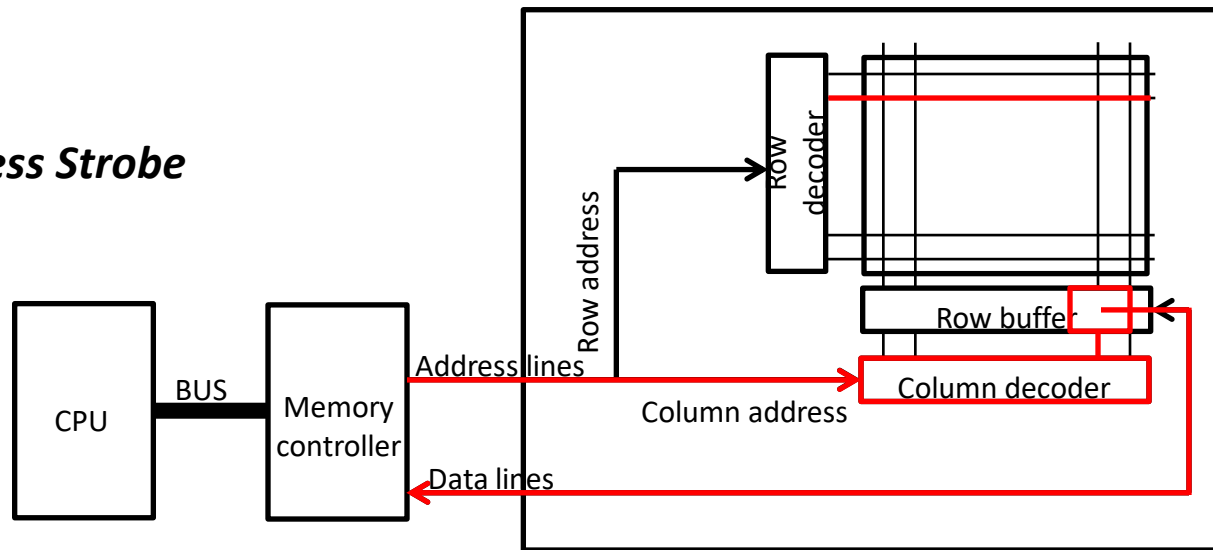
Интерна организација *DRAM*

- Приступ колони

READ Command

или

CAS: Column Address Strobe



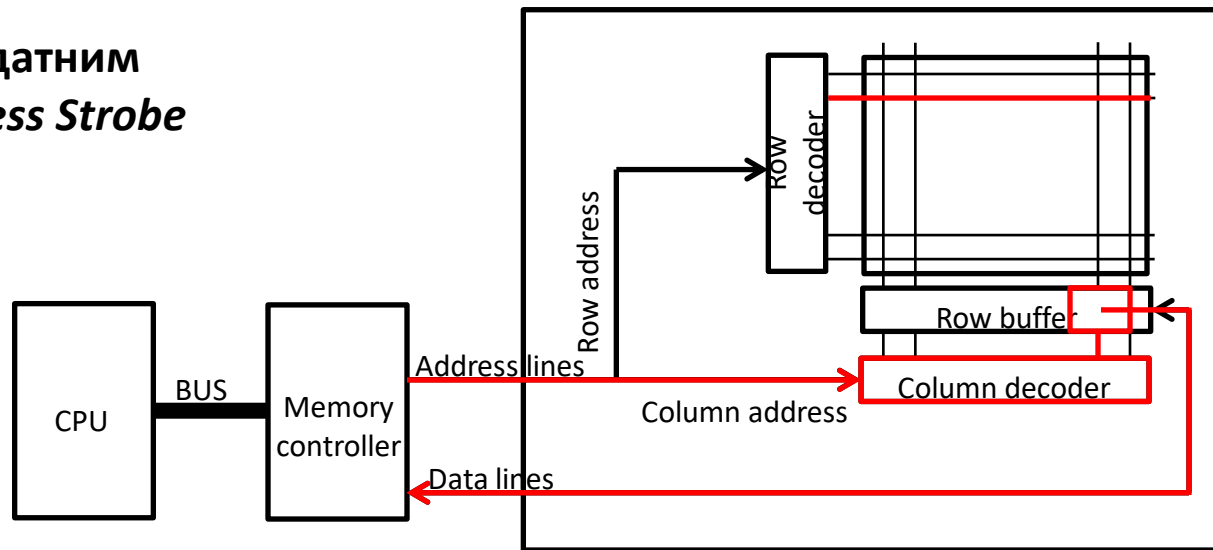
Інтерна організація *DRAM*

- Пребацивање података

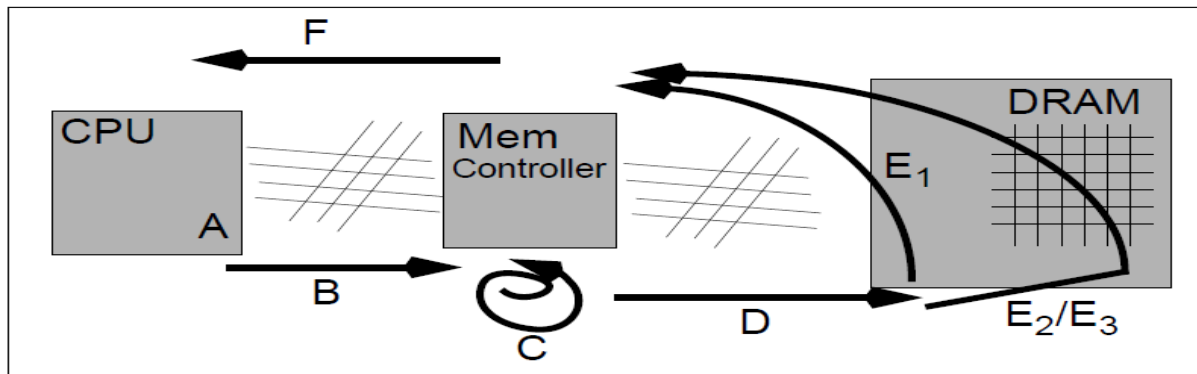
Data Out

... са опциним додатним

CAS: Column Address Strobe

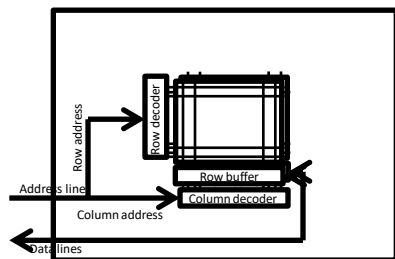


Кашњење меморије

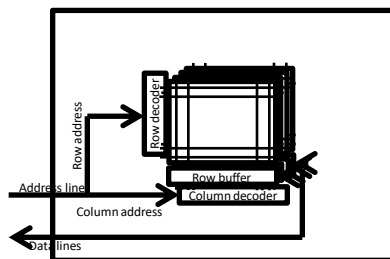


- A: Захтев за приступом (може чекати у реду)
- B: Захтев се шаље меморијском контролеру
- C: Захтев се трансформише у секвенцу наредби (може чекати у реду)
- D: Наредба послата DRAM
- E₁: Потребно је послати само **CAS** или
- E₂: Потребно је послати **RAS + CAS** или
- E₃: Потребно је послати **PRE + RAS + CAS**
- F: Враћање податка процесору
- Кашњење меморије $DRAM\ Latency = A + B + C + D + E + F$

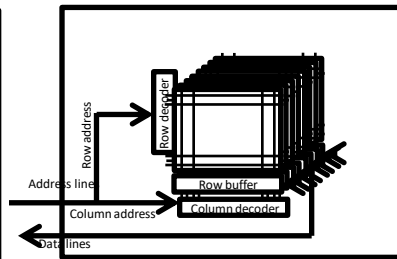
Интерна организација *DRAM*



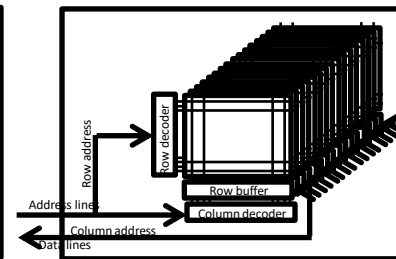
x2 DRAM



x4 DRAM

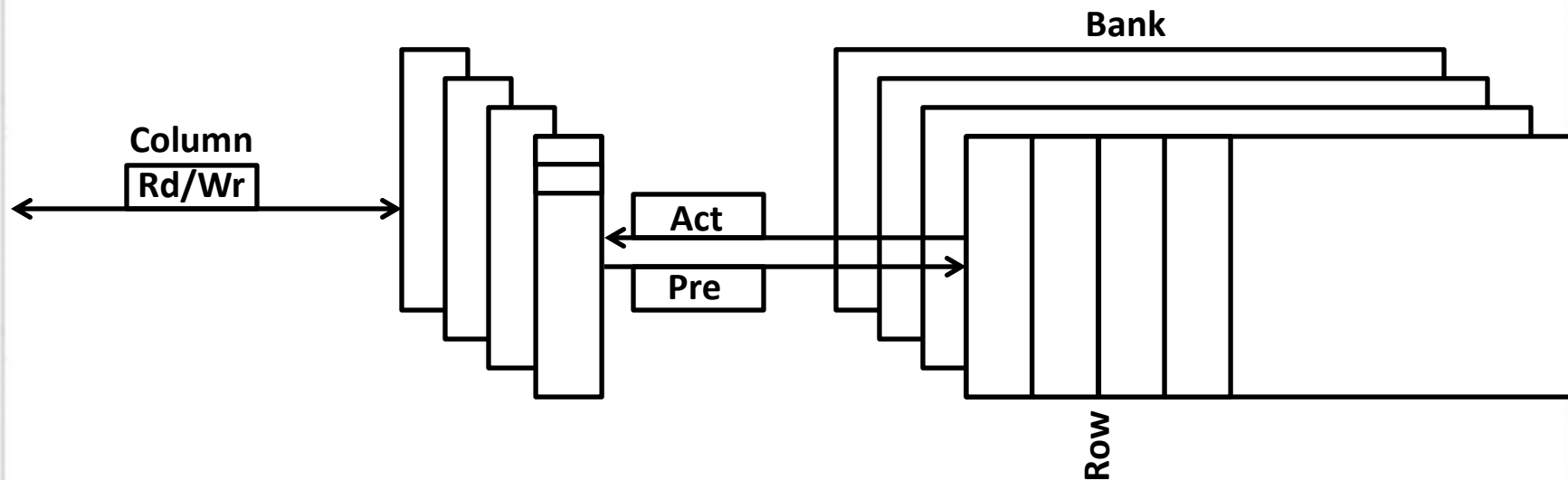


x8 DRAM

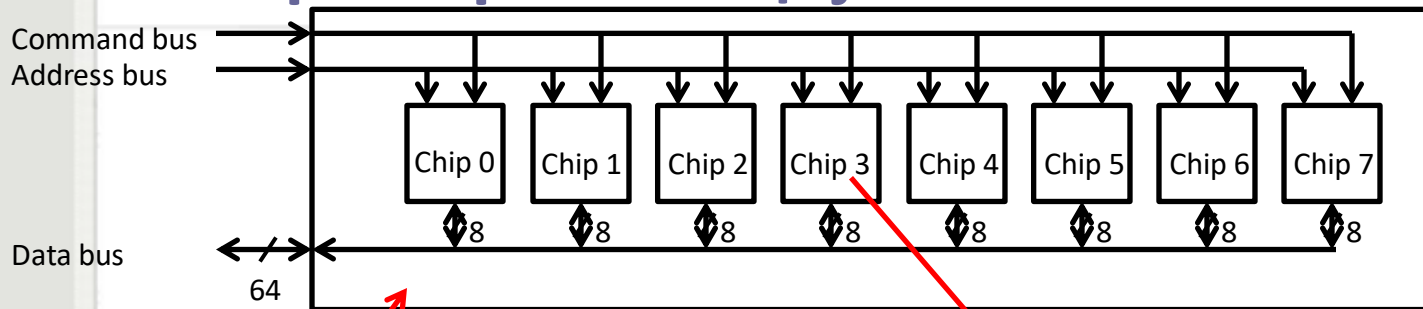


x16 DRAM

Интерна организација *DRAM*



Интерна организација *DRAM*



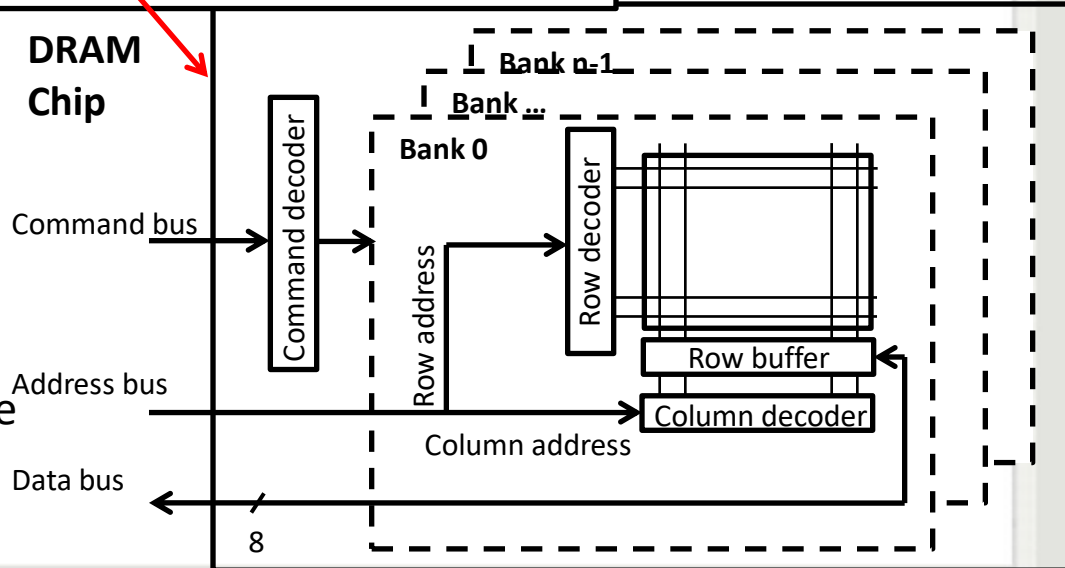
Dual inline memory modules (DIMMs)

- различни прикључци са обе стране
- 64 битни податак
- новија (184-288 прикључака)

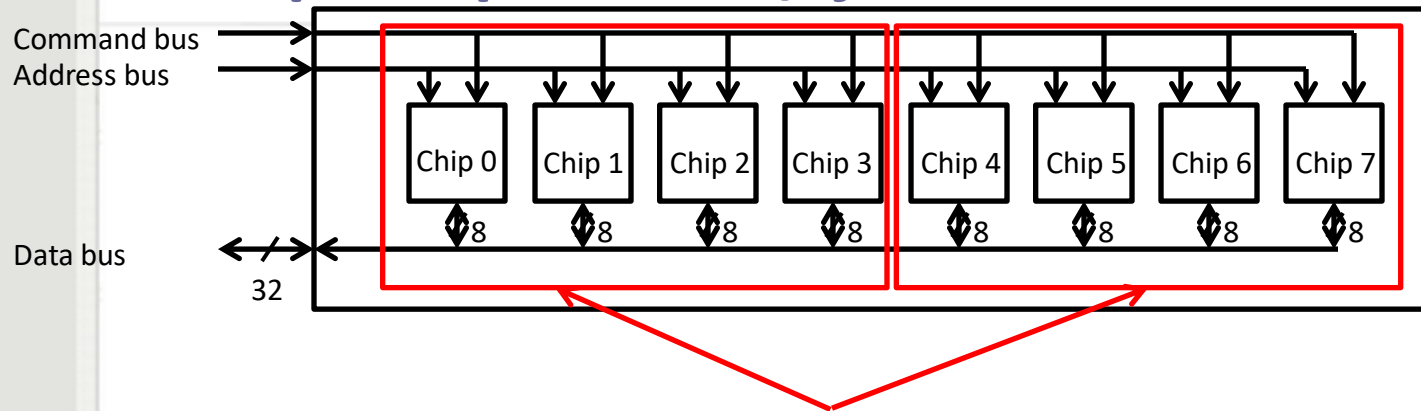
Single inline memory modules (SIMMs)

- Редундантни прикључци са обе стране
- 32 битни података
- старија, не користи се

DRAM Chip



Интерна организација *DRAM*



DIMM садржи више чипова организованих у групе/ранкове.

Група је скуп блокова (*DRAM*) чипова који раде усклађено како би одрадили појединачне наредбе коју је издао меморијски контролер

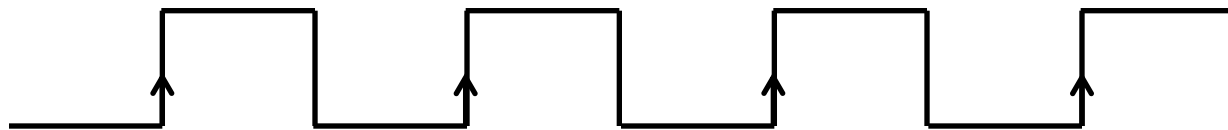
Технологије

- Амадалово правило:
 - Меморијски капацитет треба линеарно да расте са брзином рада процесора
 - Нажалост, капацитет меморије и њена брзина нису пратили карактеристике процесора
- Оптимизације:
 - Вишеструки приступ истом реду
 - Синхрони *DRAM*
 - Додат је такт *DRAM* интерфејсу
 - Пакетски мод и критична реч прво
 - Шири интерфејс
 - Двострука брзина слања - *Double data rate (DDR)*
 - Више група на сваком *DRAM* уређају
 - Мања потрошња енергије

Технологије

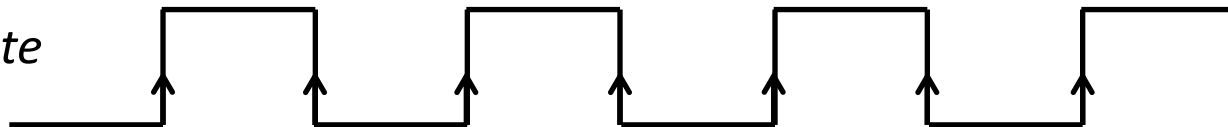
Sing data rate

SDR



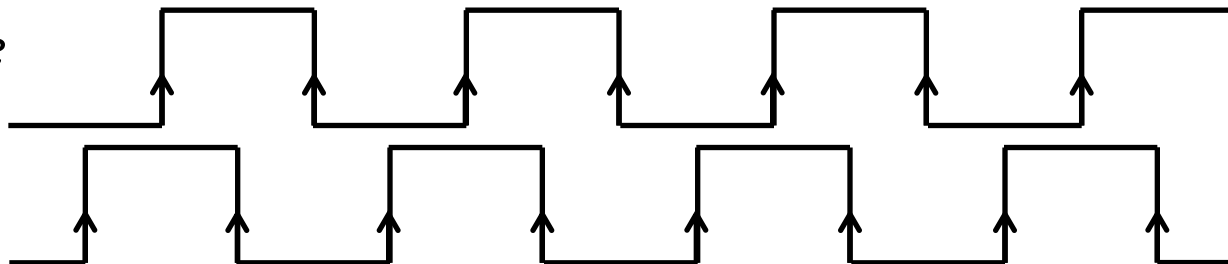
Double data rate

DDR



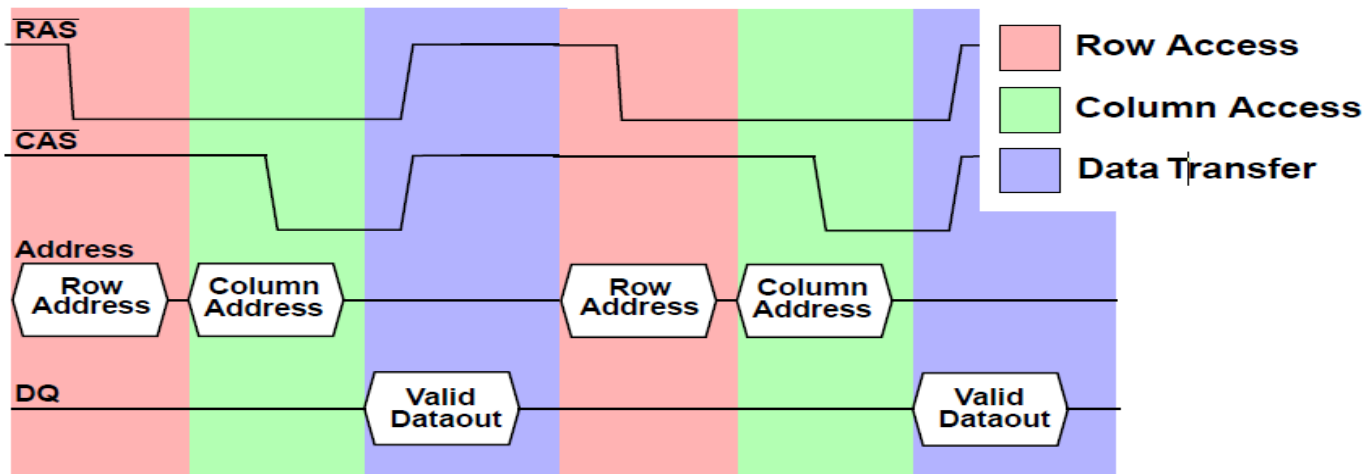
Quad data rate

QDR



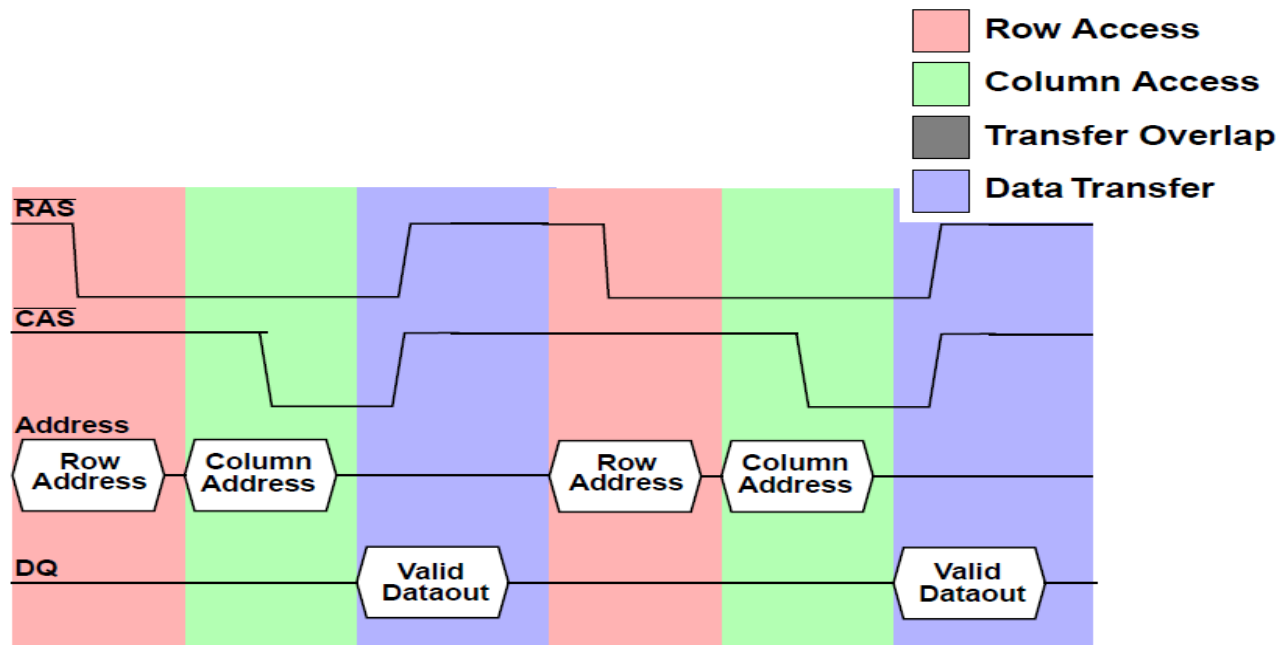
Основе

- Циклус читања стандардног *DRAM* модула



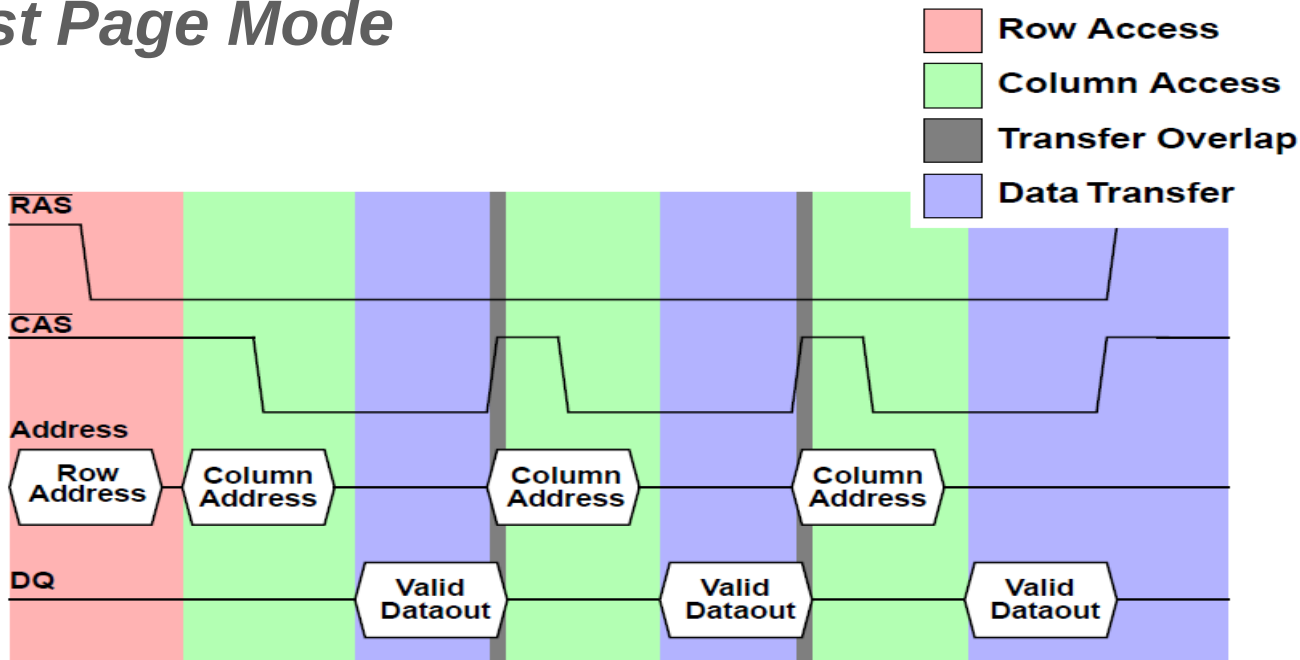
Еволуција циклуса

- Циклус читања стандардног *DRAM* модула



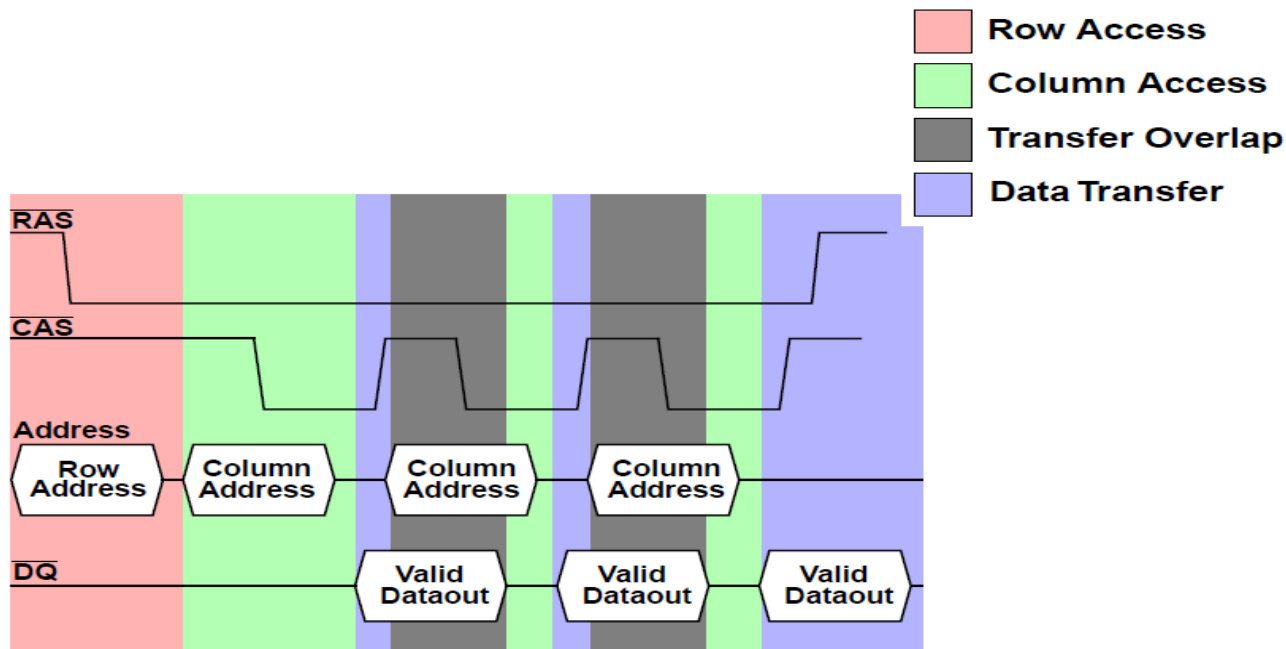
Еволуција циклуса

- Циклус читања више података из истог реда
Fast Page Mode



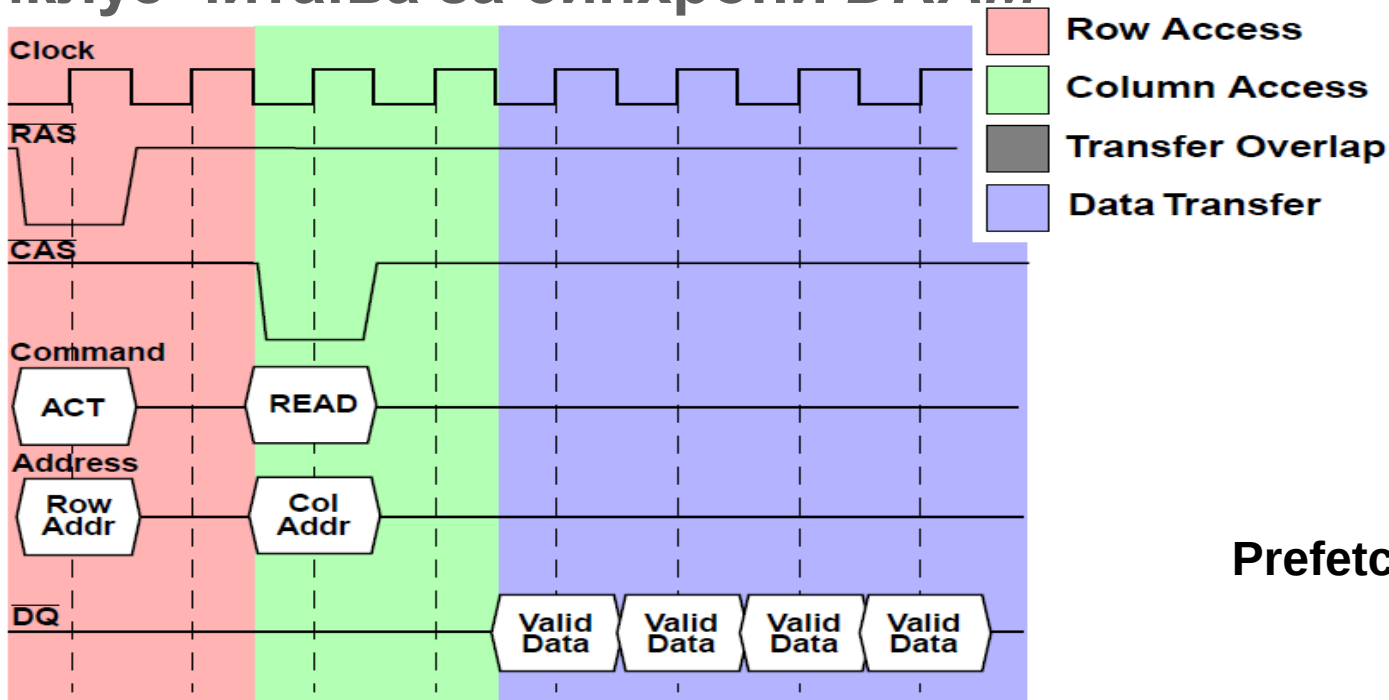
Еволуција циклуса

- Циклус читања за продужно слање података



Еволуција циклуса

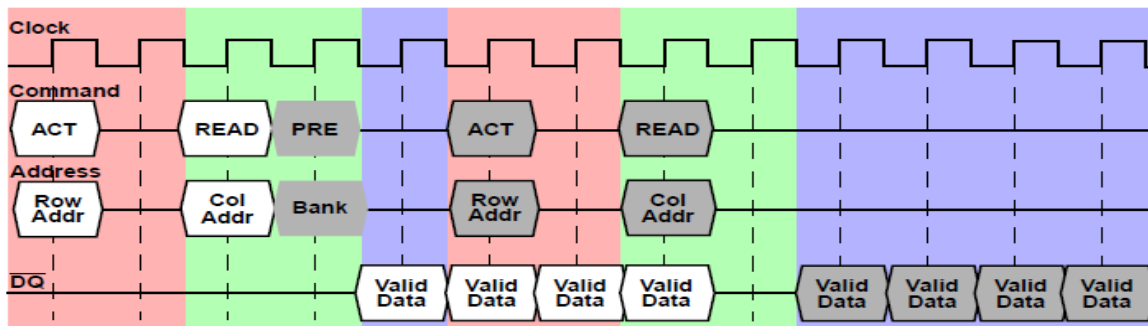
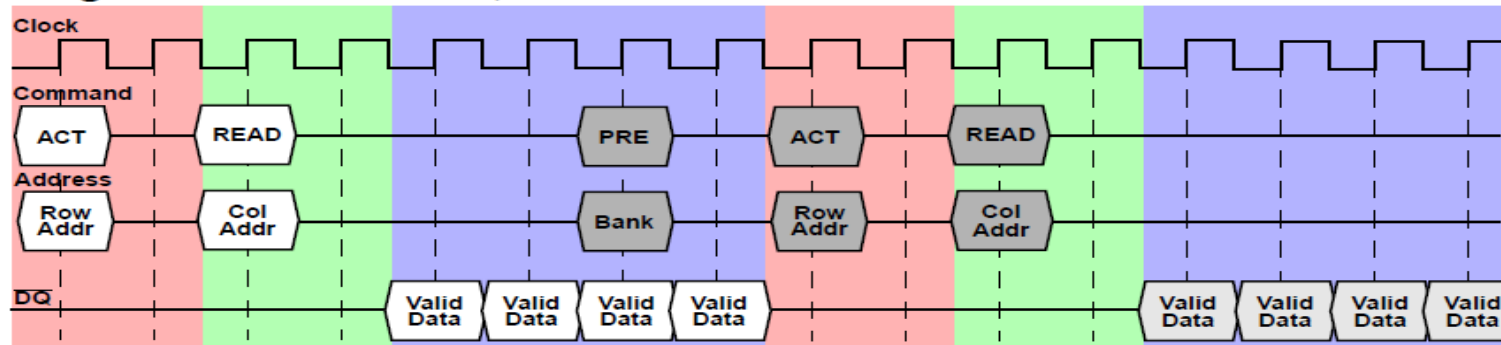
- Циклус читања за синхрони *DRAM*



Prefetch

Еволуција циклуса

- Преклапање приступа *ESDRAM*



Карактеристике меморије

			Best case access time (no precharge)			Precharge needed
Production year	Chip size	Dram type	RAS time (ns)	CAS time (ns)	Total (ns)	Total (ns)
2000	256 M bit	DDR1	21	21	42	63
2002	512 M bit	DDR1	15	15	30	45
2004	1G bit	DDR2	15	15	30	45
2006	2G bit	DDR2	10	10	20	30
2010	4G bit	DDR3	13	13	26	39
2016	8G bit	DDR4	13	13	26	39
2021	8G bit	DDR5	13	13	26	39

Карактеристике меморије

Standard	Memory Clock rate (MHz)	I/O Clock rate (MHz)	Transfer rate (MT/s)	DRAM name	Bandwidth (MB/s)	DIMM name
DDR1	133	133	266	DDR266	2128	PC2100
DDR1	150	150	300	DDR300	2400	PC2400
DDR1	200	200	400	DDR400	3200	PC3200
DDR2	133	266	533	DDR2-533	4264	PC4300
DDR2	166	333	667	DDR2-667	5336	PC5300
DDR2	200	400	800	DDR2-800	6400	PC6400
DDR3	133	533	1066	DDR3-1066	8528	PC8500
DDR3	166	666	1333	DDR3-1333	10664	PC10700
DDR3	200	800	1600	DDR3-1600	12800	PC12800
DDR4	333	1333	2666	DDR4-2666	21300	PC21300
DDR5	300	2400	4800	DDR5-4800	38400	PC38400

Карактеристике меморије

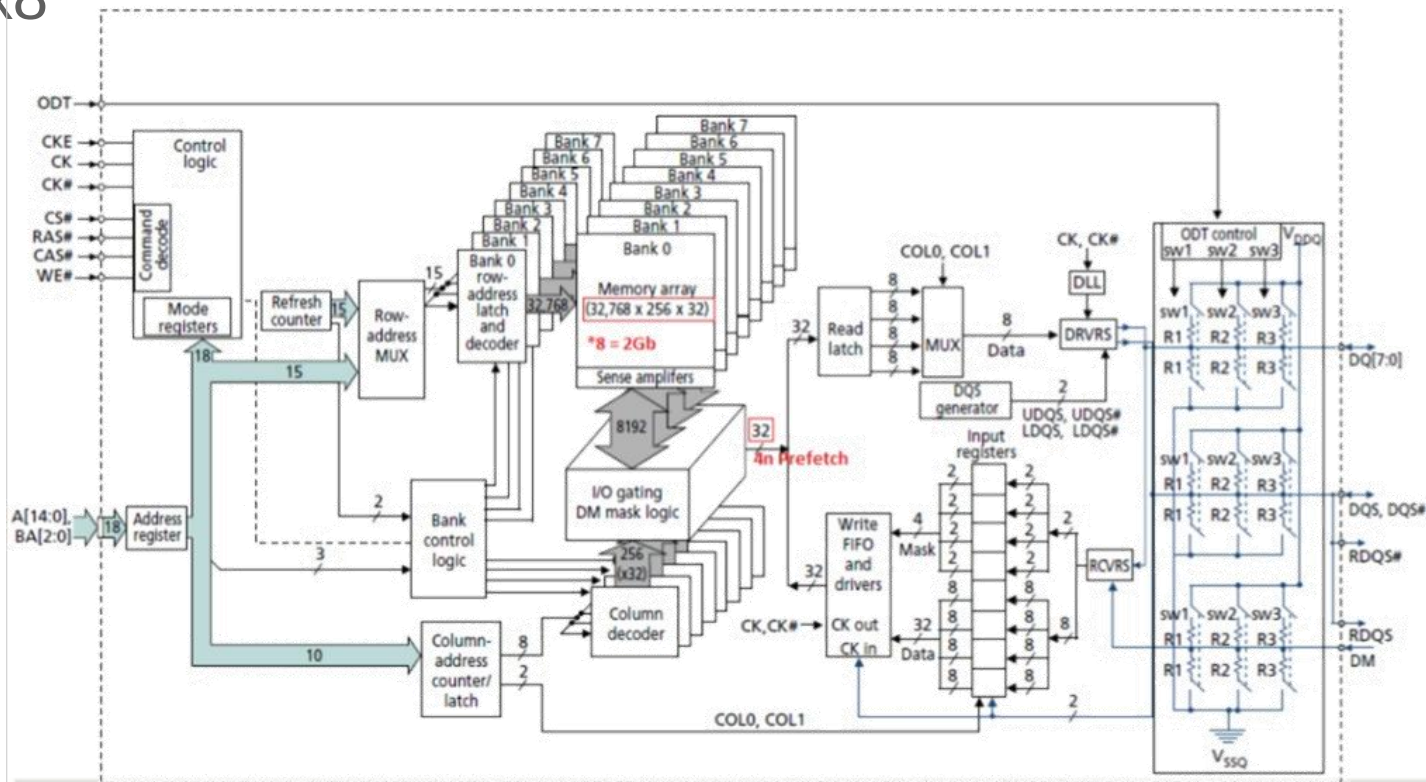
Standard	Memory Clock rate (MHz)	I/O Clock rate (MHz)	Voltage (V)	Prefetch
DDR1	100-200	100-200	2.5-2.6	2n
DDR2	100-266	200-533	1.8	4n
DDR3	100-266	400-1066	1.35-1.5	8n
DDR4	200-400	1066-1600	1.05-1.2	8n
DDR5	200-450	1600-3600	1.1	16n

- 64Mx8



Интерна структура DDR2 меморије

- 256Mx8

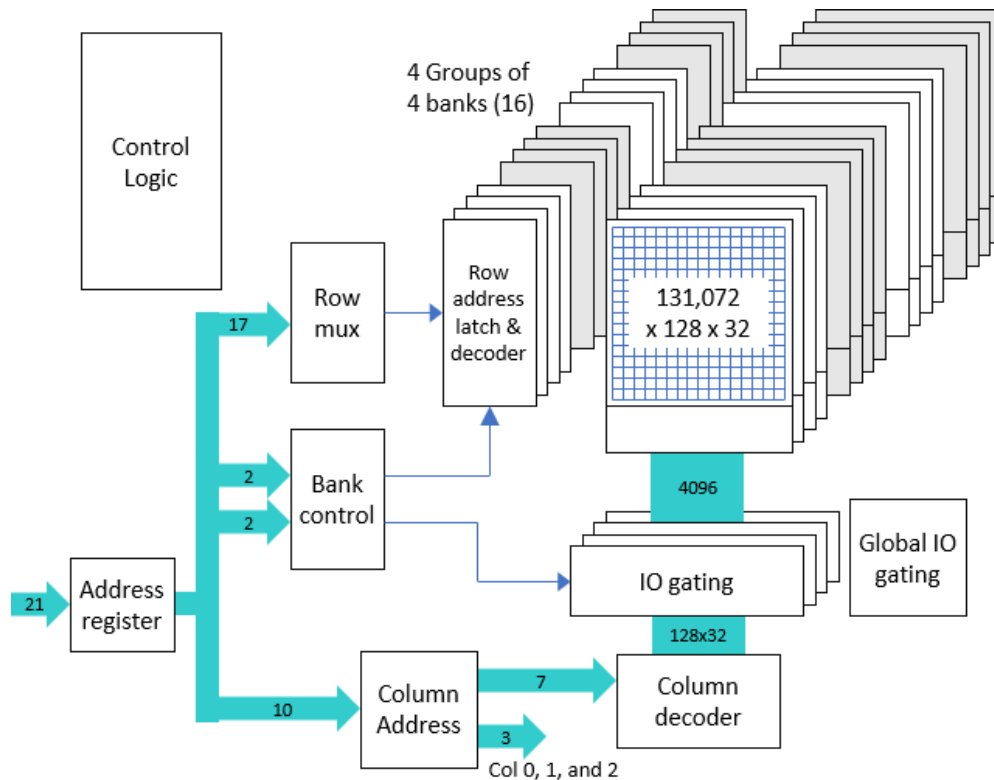


- 512Mx8

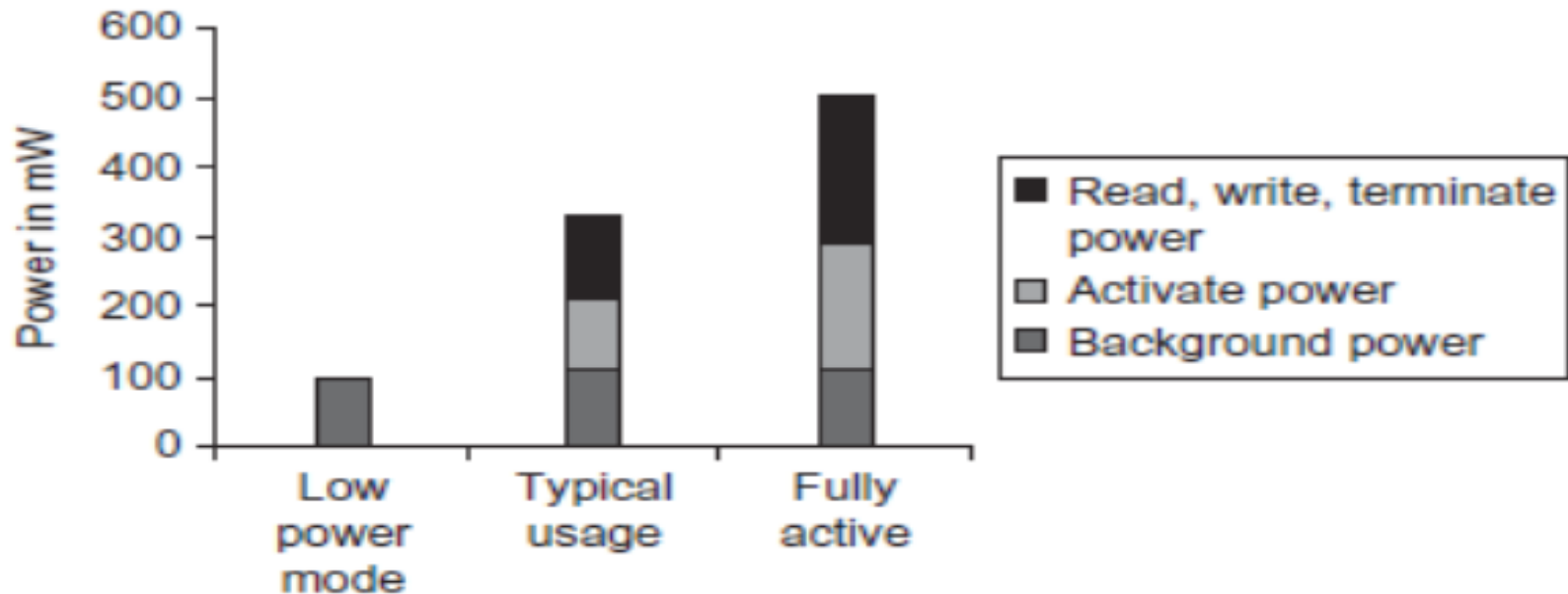


Интерна структура DDR4 меморије

- 8Gx8



Потрошња енергије меморије

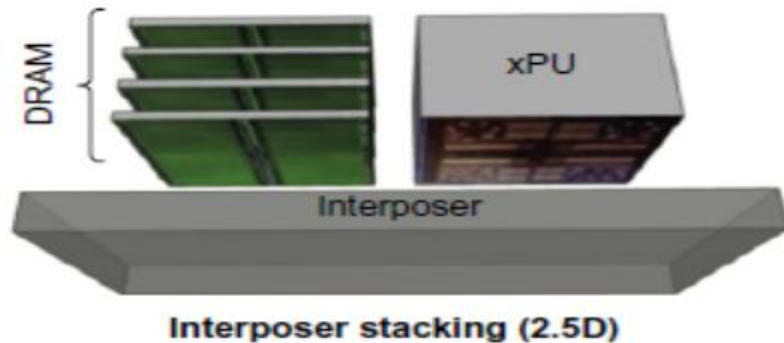
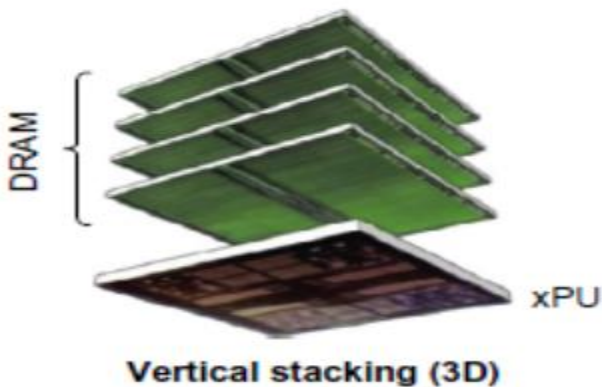


LPDDR – (Low Power DDR)

- DRAM меморије оптимизован за ниску потрошњу енергије.
- **Кључне карактеристике:**
 - Ради на nižем напону у односу на стандардни *DDR*.
 - Има механизме за уштеду енергије (динамичко искључивање делова меморије, *clock gating*).
 - Подржава *deep power-down* и *self-refresh* модове за минималну потрошњу у стању мировања.
 - Балансира брзину и енергетску ефикасност – мање пропусног опсега него *DDR* за десктоп, али довољно за мобилне апликације.
- **Примена:**
 - Смартфони, таблети, лаптопи, аутомобилска електроника, *IoT* уређаји.

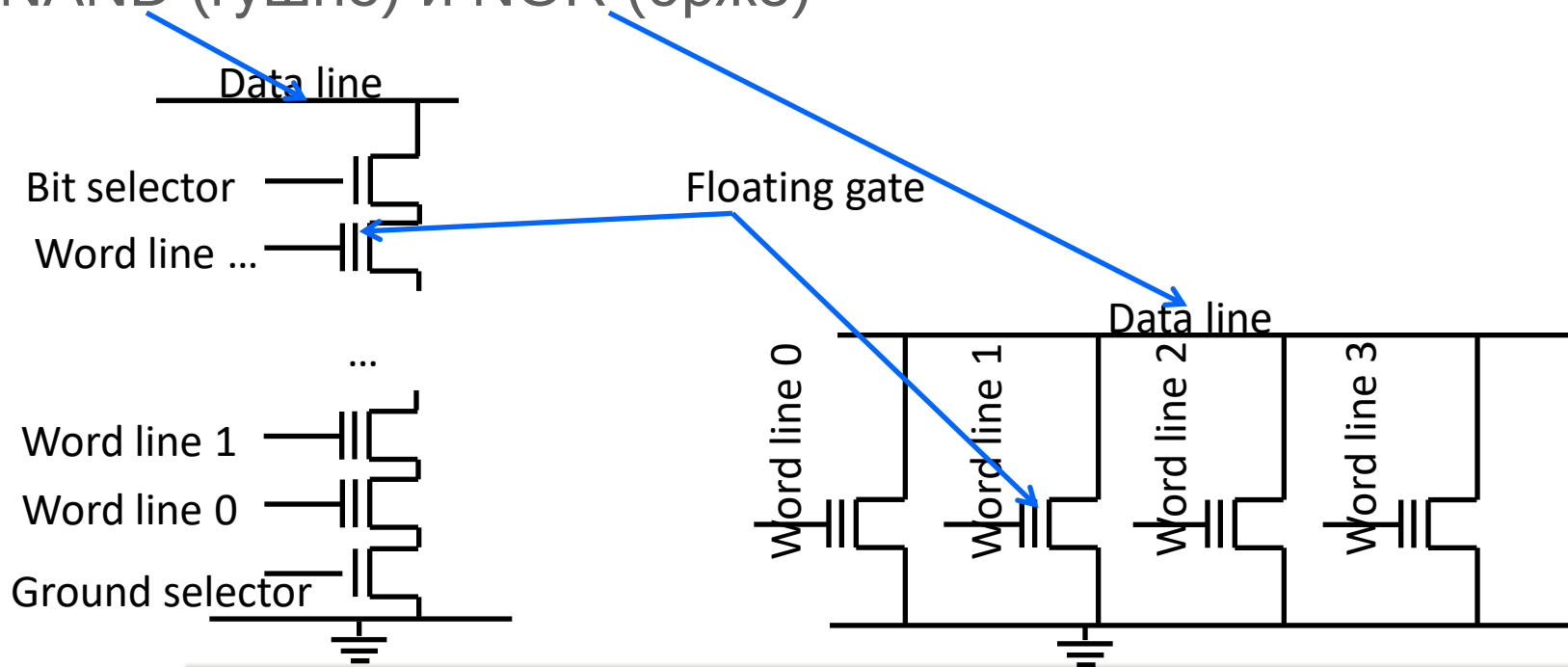
Упакован/уграђен DRAMs

- Паковање меморије и процесора на исти чип
 - *High Bandwidth Memory (HBM)*



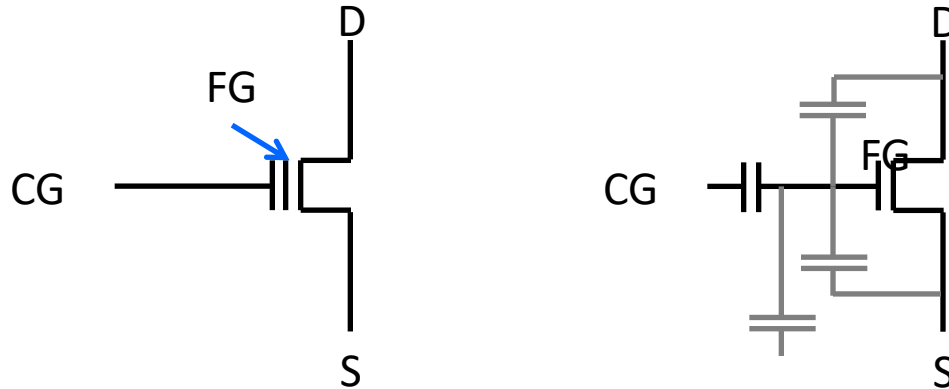
Флеш меморија - *Flash Memory*

- Типови *EEPROM*
- NAND (гушће) и NOR (брже)



Плутајући гејт - *Floating gate*

- Електрони се убацују/ваде помоћу квантног ефекта који се назива *Fowler–Nordheim* тунеловање



Флеш меморија - *NAND*

- Читање је секвенцијално
- Увек се чита читава страница (0,5 - 4 KB)
- Брзина приступа 25 μ s за први бајт и 40 MB/s за преостале бајтове
- SDRAM: 40 ns за први бајт и 4,8 GB/s за преостале бајтове
- Пример: трансфер 2 KB:
 - 75 μ s наспрам 500 ns за SDRAM, 150X спорије
- 300 – 500 пута брже него магнетни диск

Флеш меморија - *NAND*

- Мора се обрисати читав блок пре него што се упише нова вредност
- Не губи вредност, веома мало троши
- Ограничен број уписа (~100 000)
- Цена:
 - NAND – 0.12-0.2\$/GB
 - SDRAM – 7.5-12\$/GB
 - Магнетни диск - 0,02-0.05\$/GB

Флеш меморија - *Flash Memory*

Карактеристика	NAND	NOR
Густина	Велика (јефтинија)	Мала (скупља)
Читање	Умерено (<i>Block access</i>)	Веома брзо (<i>Random access</i>)
Уписивање	Брзо (по страницама)	Споро (по речима)
Брисање	Брзо (велики блокови)	Споро (мањи блокови)
Животни век	Краћи (10K - 100K циклуса)	Дужи (100K - 1М циклуса)
Примена	<i>SSD, USB, SD</i> картице	<i>Firmware (BIOS, рутери, аутомобили)</i>

Мемристор - *Phase-change memory (PCM)*

- Компоненте засноване на промени стања материјала (аморфно и кристално) под утицајем струје
- Бит одговара пресеку у дводимензионалној мрежи
- Читање се обавља одређивањем отпорности између тачака x и y
- Не користи транзисторе што би требало да омогући мању цену и већу густину паковања у поређењу са *NAND*.
- Могуће убрзање од 10 пута приликом уписа и 2 пута приликом читања у односу на *NAND*.

Алтернативна решења

MRAM (Magnetoresistive RAM)

- Чување података у магнетним доменима (*MTJ – Magnetic Tunnel Junction*). Брзина као *SRAM*, неволатилна, висока издржљивост. Висока цена, сложена производња.

ReRAM (Resistive RAM)

- Промене отпора у оксидном слоју (кисеоничне празнине). Предности: скалабилност < 10 nm, брзина, ниска потрошња. Примене: *SSD*, *IoT*, неуроморфно рачунарство. Недостаци: још увек ограничена комерцијализација.

FeRAM (Ferroelectric RAM)

- Користи фероелектрични материјал (*PZT*) уместо диелектрика. Предности: ниска потрошња, брзо уписивање, огромна издржљивост (10^{12} циклуса). Недостаци: мала густина, висока цена. Примене: паметне картице, аутомобилска електроника, индустријски уређаји.

Отклањање грешака

- Меморија је осетљива на космичке зраке
- Динамичке грешке *Soft errors*:
 - Детектују их и исптављају кодови за исправњање грешака *error correcting codes (ECC)*
- Трајне грешке *Hard errors*:
 - Користе се додатни редови за замене оштећене

ECC – Error Correcting Codes

- Вероватноће грешака:

$P(\text{без грешака}) > P(\text{једна грешка}) > P(\text{две грешке}) \gg P(>\text{две грешке})$

- Детекција грешака – обавештавање о грешци
- Исправљање – враћање исправне вредности
- Најчешће
 - SED – откривање једне грешке
 - SECDED – исправљање једне грешке, откривање две грешке

Парност

ЕСС за један бит

Тип	Исправно	Број бита	Коментар
Без	0,1	1	-
SED	00,11	2	01,10 грешке
SEC	000,111	3	001,010,100 => 0 110,101,011 => 1
SECDDED	0000,1111	4	једна јединица => 0 две јединице => грешка три јединице => 1

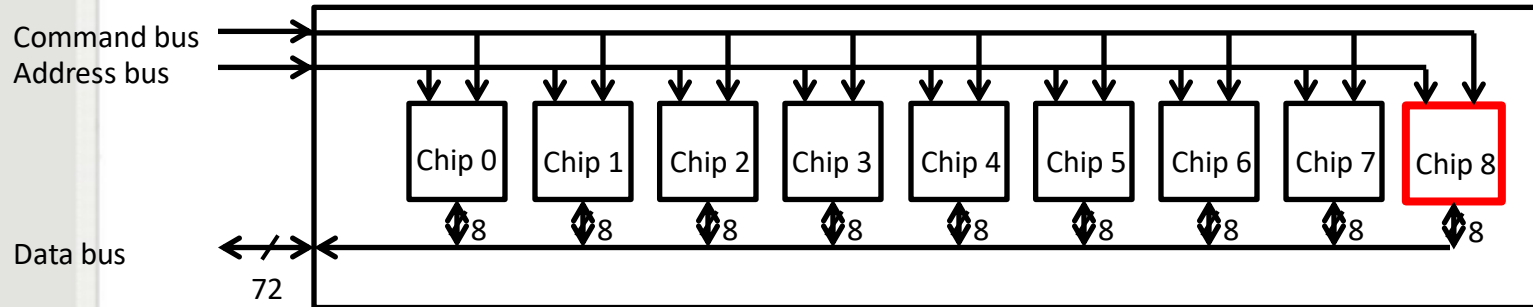
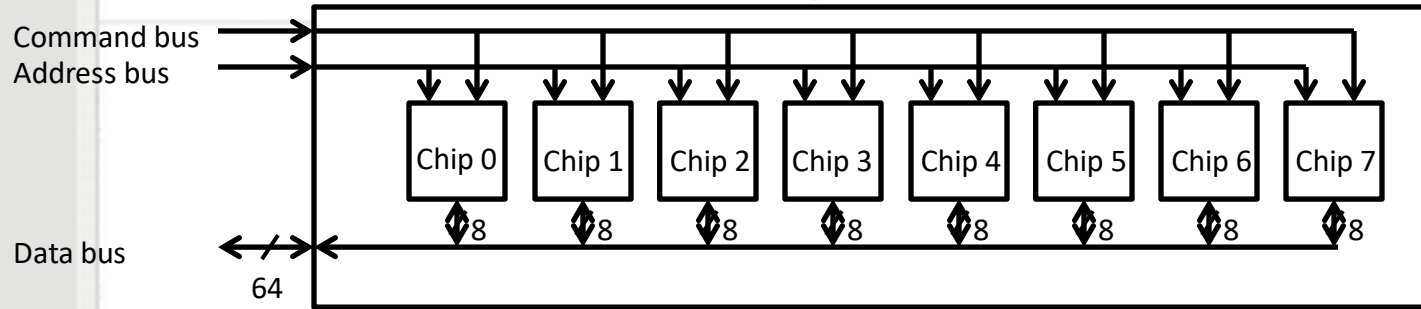
ECC – Error Correcting Codes

Смањење додатака са повећањем величине речи

Број бита	SED додатак	SECDED додатак
1	1 (100%)	3 (300%)
8	1 (12,5%)	5 (63)
16	1 (6,3%)	6 (38%)
32	1 (3%)	7 (22%)
64	1 (1,6%)	8 (13%)
n	1 (1/n)	$1 + \log_2 n + \dots$

Дужа реч => већа вероватноћа појављивања више грешака!

ECC – Error Correcting Codes



ECC – Error Correcting Codes

- Код 64 битних података користи се додатних 8 бита
dddd.....d ccccccccc
- *DIMM* се састоји из 9 осмобитних *DRAM* чипова
зашта је потребно 72 бита
- Више имплементација
 - Оригинални Хемингов алгоритам *Hamming* (72, 64)
(Старији системи, *FPGA* дизајн)
 - Оптимизовани алгоритам *Hsiao* (72, 64)
(Неки x86 сервери (Intel/AMD))
 - Оптимизовани алгоритам *Marvell* (72, 64)
(Специјализовани уређаји)

ECC – Error Correcting Codes

- Израчунавање $H = [P \mid I]$, где је I јединична матрица димензија 8×8 (која представља контролне битове), а P је матрица димензија 8×64 која дефинише како се подаци комбинују.
- *Hamming* колоне у P матрици су бинарни прикази бројева од 1 до 72, изузимајући степене двојке (који су резервисани за парност).
- *Hsiao* колоне у P матрици се бирају по броју јединица у колони. Прво колоне са тачно **3 јединице** (њих 56) па колоне са тачно **5 јединица** (допуна до 64).

Пример *Hamming (7, 4) SEC*

Позиција	1	2	3	4	5	6	7
Тип	P1	P2	D1	P3	D2	D3	D4

Приликом слања рачунају се битови парности:

$P1 = D1 \text{ XOR } D2 \text{ XOR } D4$ (проверава позиције 1, 3, 5, 7)

$P2 = D1 \text{ XOR } D3 \text{ XOR } D4$ (проверава позиције 2, 3, 6, 7)

$P3 = D2 \text{ XOR } D3 \text{ XOR } D4$ (проверава позиције 4, 5, 6, 7)

За 1010 пошаље се **1011010**

Пример *Hamming (7, 4) SEC*

Позиција	1	2	3	4	5	6	7
Тип	P1	P2	D1	P3	D2	D3	D4

Приликом пријема рачуна се синдром (S) који указује на грешку ако је различит од 0:

$$S1 = P1 \text{ XOR } D1 \text{ XOR } D2 \text{ XOR } D4$$

$$S2 = P2 \text{ XOR } D1 \text{ XOR } D3 \text{ XOR } D4$$

$$S3 = P3 \text{ XOR } D2 \text{ XOR } D3 \text{ XOR } D4$$

За пријем 1011010 добије се S=000 нема грешака

За пријем 1011**1**10 добија се S=101 што каже да је грешка на биту 5.

Да би постао SECDED потребно је додати још један бит парности.

Питања?

Електротехнички Факултет
Универзитет у Београду

