



VIETNAM NATIONAL UNIVERSITY HANOI (VNU)
VNU UNIVERSITY OF ENGINEERING AND TECHNOLOGY

Computer Architecture

Lecture 4: Internal Memories

Duy-Hieu Bui, PhD

VNU Information Technology Institute
Laboratory for Smart Integrated System (SISLAB)

Email: hieubd@vnu.edu.vn
www.uet.vnu.edu.vn/~hieubd



1. Khái niệm chung

2. Bộ nhớ trong

- RAM
- ROM
- DRAM
- FLASH
- Non-volatile Memory

Tham khảo chương 5, “Computer Organization and Architecture: Designing for Performance”, William Stallings, 10th edition



1. Khái niệm

- Bộ nhớ: thiết bị có thể *bảo quản* và *khôi phục* một thông tin
- Từ nhớ: tập bits có thể được đọc hay ghi đồng thời
- Các kiểu vật liệu nhớ:
 - Bán dẫn – semiconductor (register, cache, bộ nhớ chính, ...)
 - Từ - magnetic (đĩa mềm, đĩa cứng, ...)
 - Optic (CD-ROM, DVD-ROM)
 - ...



Đặc điểm

1. Vị trí
2. Dung lượng
3. Đơn vị truyền
4. Kiểu truy cập
5. Hiệu năng
6. Kiểu vật liệu
7. Đặc trưng vật liệu
8. Tổ chức

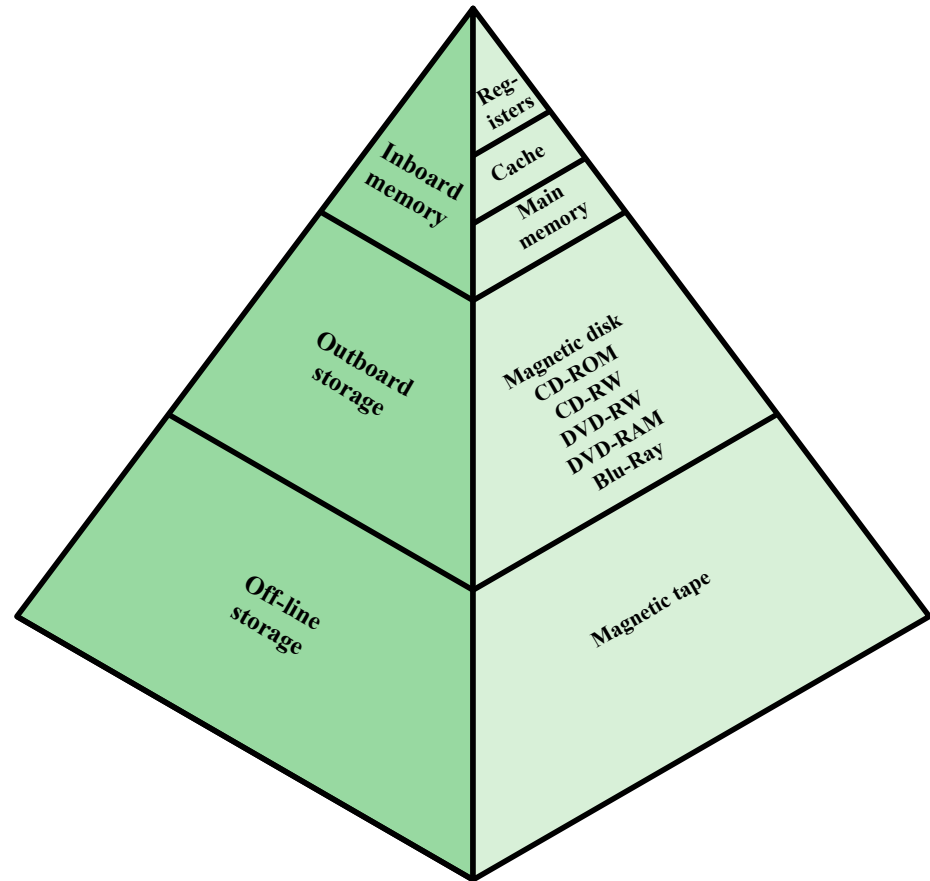


Đặc điểm...

- Vị trí
 - CPU
 - Internal
 - External
- Dung lượng
 - Phụ thuộc vào kích thước từ nhớ, và
 - Số lượng từ nhớ
- Đơn vị truyền
 - Bên trong: phụ thuộc vào độ rộng bus dữ liệu
 - Bên ngoài: block(>từ nhớ)
 - Đơn vị có thể đánh địa chỉ được

- Kiểu truy cập
 - Tuần tự: VD băng từ
 - Trực tiếp:
 - Mỗi 1 block có 1 địa chỉ duy nhất
 - Truy cập = cách nhảy đến vùng lân cận và tìm tuần tự
 - Thời gian truy cập vào vị trí hiện tại hiện tại và trước đó
 - VD: HardDisk, Floppy Disk,...
 - Ngẫu nhiên:
 - Mỗi địa chỉ xác định chính xác một vị trí
 - Thời gian truy cập không phụ thuộc vào vị trí cũng như lần truy cập trước
 - VD: RAM, ...
 - Kết hợp:
 - Một từ được định vị thông qua việc sử dụng một phần nội dung của từ đó
 - Thời gian truy cập không phụ thuộc vào vị trí cũng như lần truy cập trước
 - VD: cache, ...

- Chiến thuật phân cấp bộ nhớ: How much? How fast? How expensive?
 - Registers
 - L1 Cache
 - L2 Cache
 - Main memory
 - Disk cache
 - Disk
 - Optical
 - Tape





Đặc điểm...

- **Hiệu năng:**
 - Thời gian truy cập: khoảng thời gian từ khi gửi địa chỉ cho đến khi thu được dữ liệu trọn vẹn
 - Thời gian chu trình nhớ - Memory Cycle Time:
 - Thời gian bộ nhớ đòi hỏi để “hồi phục” trước lần truy cập kế tiếp
 - = access + recovery
 - Tốc độ chuyển dữ liệu
- **Kiểu vật liệu:**
 - Semiconductor :RAM
 - Magnetic: Disk & Tape
 - Optical: CD & DVD
 - Others: Bubble, Hologram



Đặc điểm...

- Đặc trưng vật liệu:
 - Phân rã - Decay
 - Dễ thay đổi - Volatility
 - Có thể xoá được - Erasable
 - Năng lượng tiêu thụ
- Tổ chức:
 - Cách thức sắp xếp các bits trong một từ
 - Thường không rõ ràng
 - VD: interleaved



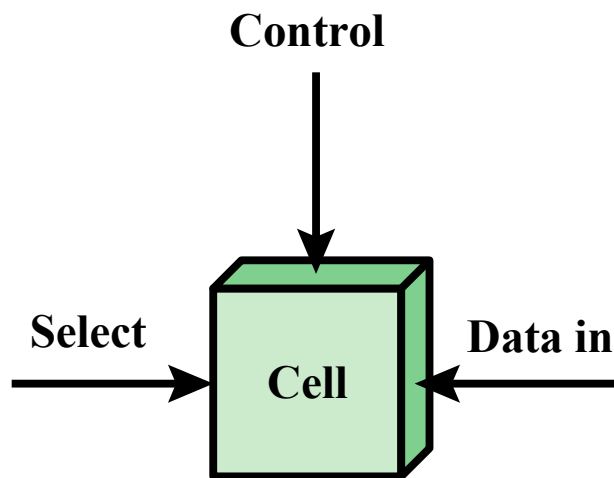
1. Khái niệm chung

2. Bộ nhớ trong

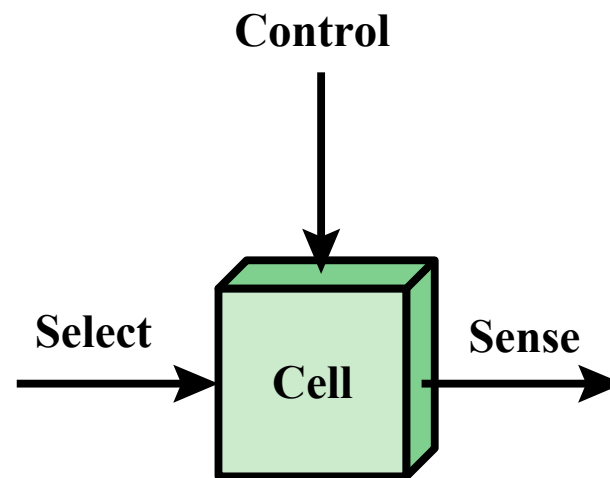
- RAM
- ROM
- DRAM
- FLASH
- Non-volatile Memory

Tham khảo chương 4 của [1]

- Bộ nhớ bên trong máy tính
 - Semiconductor
 - Truy cập ngẫu nhiên



(a) Write



(b) Read

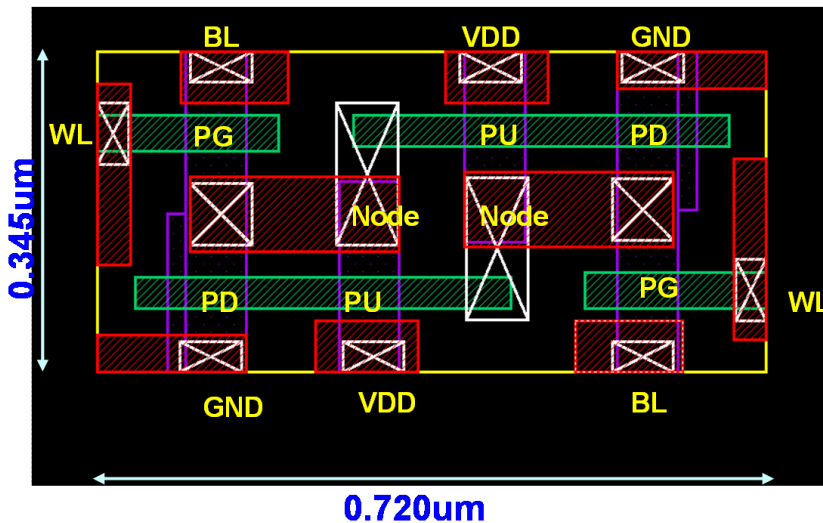
Semiconductor Memory types

- Random Access Memory (RAM): lưu giữ những dữ liệu tạm thời
- Read Only Memory (ROM): lưu giữ thông tin cố định

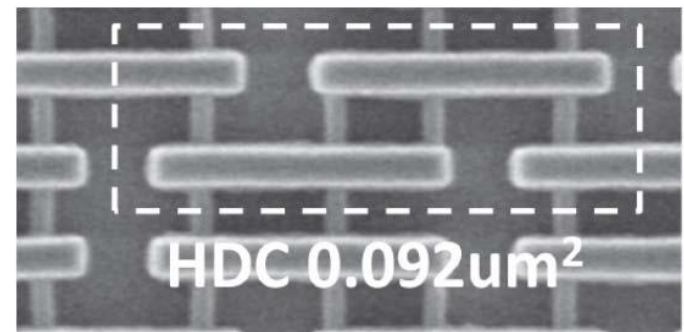
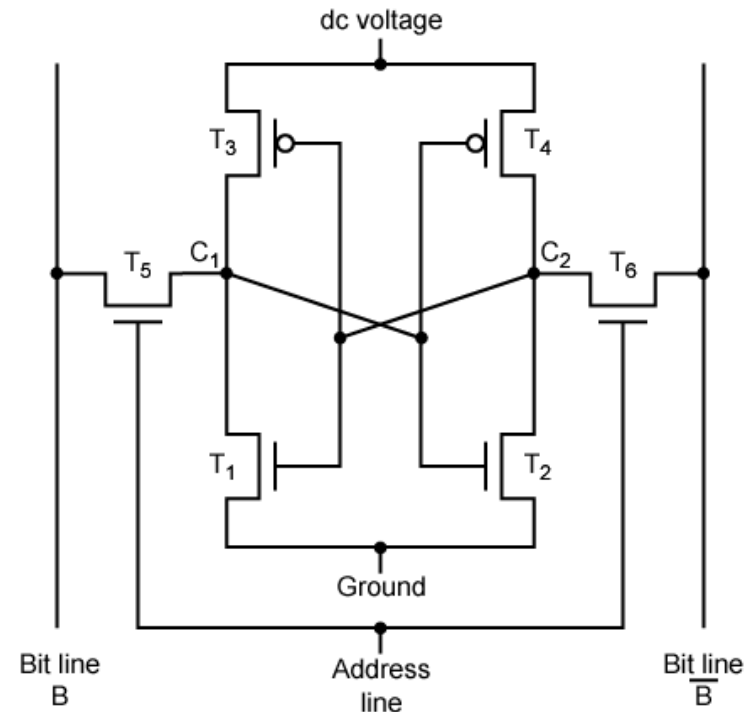
Memory Type	Category	Erasure	Write Mechanism	Volatility
Random-access memory (RAM)	Read-write memory	Electrically, byte-level	Electrically	Volatile
Read-only memory (ROM)	Read-only memory	Not possible	Masks	Nonvolatile
Programmable ROM (PROM)				
Erasable PROM (EPROM)	Read-mostly memory	UV light, chip-level	Electrically	
Electrically Erasable PROM (EEPROM)		Electrically, byte-level		
Flash memory		Electrically, block-level		

• SRAM – Static RAM

- Bits được lưu trong các flip-flops
- Không cần làm tươi, có tốc độ cao
- Phức tạp, kích thước to hơn, giá thành cao



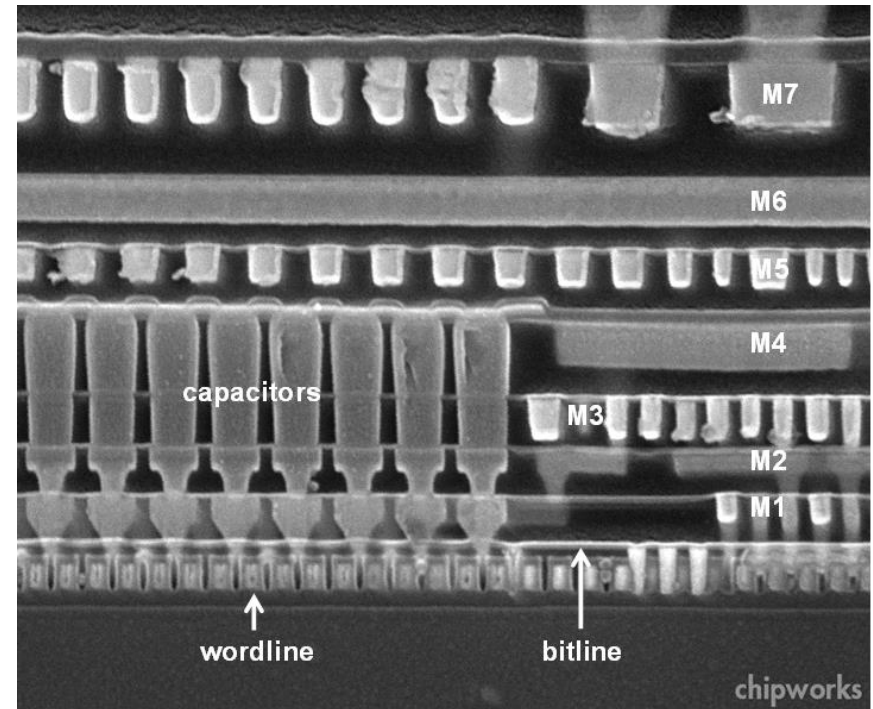
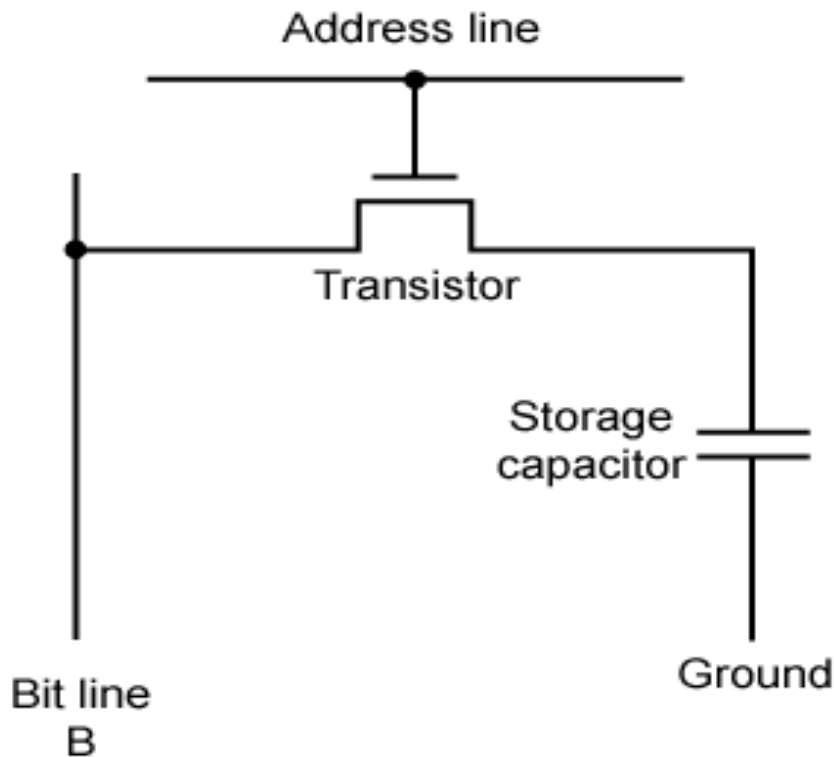
SRAM layout CMOS 45nm



Intel 22nm Tri-gate

- DRAM – Dynamic RAM

- Bits được lưu trữ trong các tụ điện
- Đơn giản, kích thước bé, giá thành rẻ
- Chậm, cần 1 chu trình làm tươi ngay cả khi đã được cung cấp nguồn



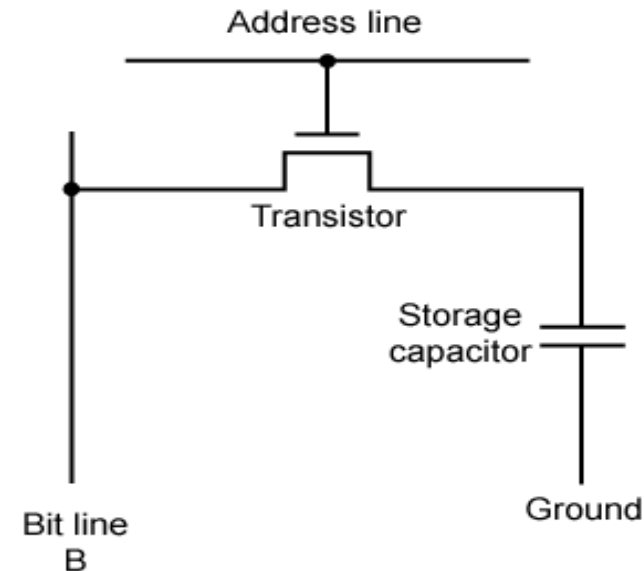
Intel embedded DRAM

<http://chipworksrealchips.blogspot.com/2014/02/intels-e-dram-shows-up-in-wild.html>

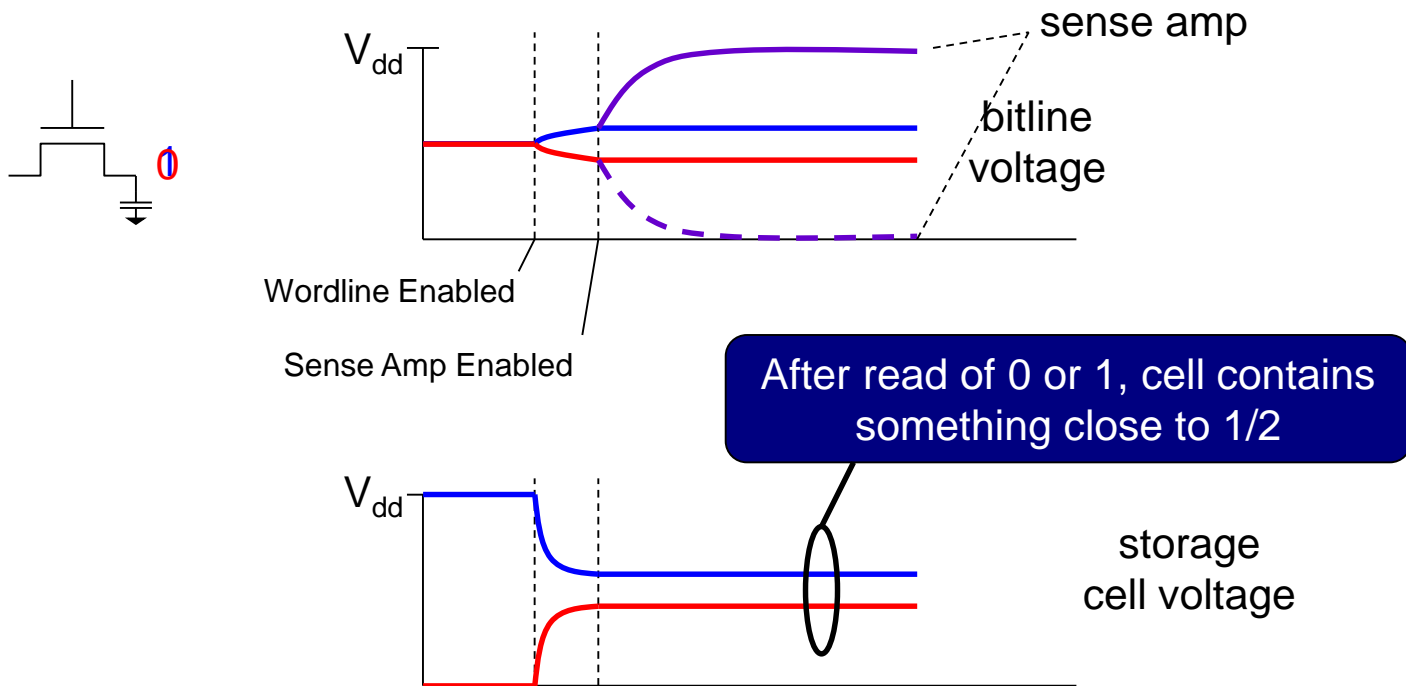


Dynamic RAM

- Đường địa chỉ được kích hoạt khi đọc/ghi bit
 - Transistor switch closed (current flows)
- Ghi
 - Voltage to bit line
 - High for 1 low for 0
 - Then signal address line
 - Transfers charge to capacitor
- Đọc
 - Address line selected
 - transistor turns on
 - Charge from capacitor fed via bit line to sense amplifier
 - Compares with reference value to determine 0 or 1
 - Capacitor charge must be restored

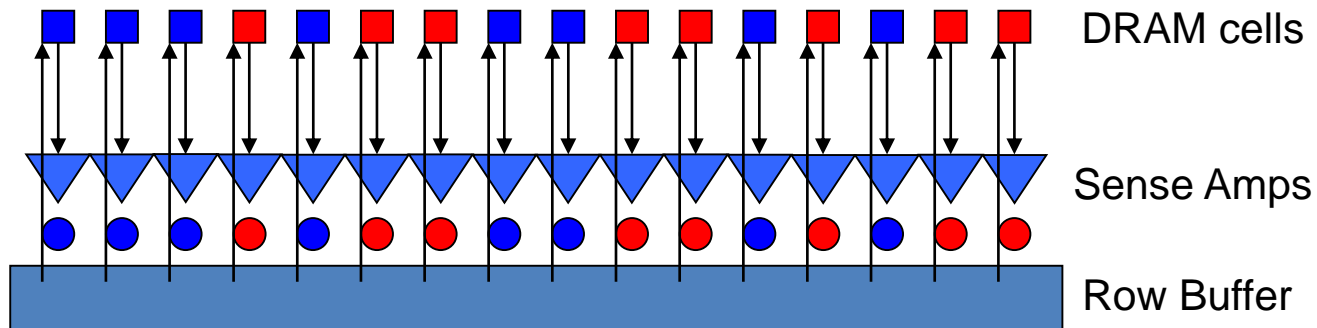


Destructive Read

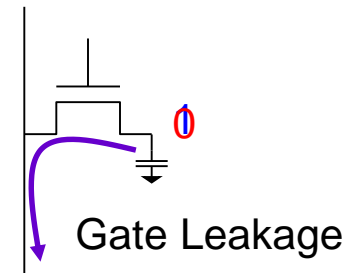


Cơ chế làm tươi - refresh

- Sau khi đọc, nội dung của DRAM cell đã bị thay đổi
- Lưu các giá trị cells trong bộ đệm hàng *row buffer*
- Ghi lại các giá trị đó cho các cells trong lần đọc kế tiếp



- Thực tế, DRAM cell sẽ mất nội dung ngay cả khi không có tác vụ đọc
→ lý do được gọi là “dynamic”
- Vì thế các cells trong DRAM cần được định kỳ đọc và ghi lại nội dung



- Transistor arrangement gives stable logic state

- State 1

- C_1 high, C_2 low
- T_1 T_4 off, T_2 T_3 on

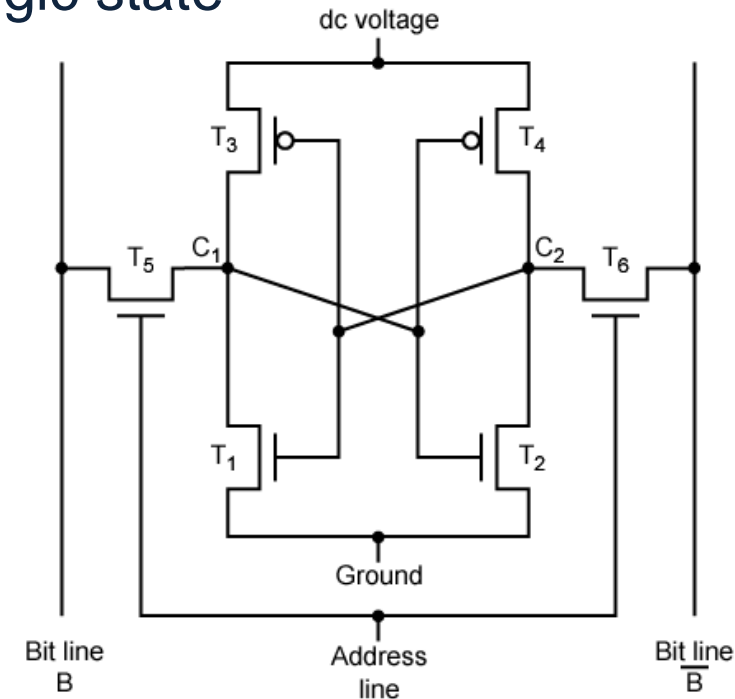
- State 0

- C_2 high, C_1 low
- T_2 T_3 off, T_1 T_4 on

- Address line transistors T_5 T_6 is switch

- Write – apply value to B & compliment to B

- Read – value is on line B





SRAM vs. DRAM

- Cả hai đều có tính chất volatile
 - Cần cung cấp năng lượng để bảo quản dữ liệu
- Dynamic cell
 - Đơn giản, kích thước nhỏ gọn
 - Mật độ cell cao
 - Chi phí thấp
 - Cần chu kỳ làm tươi
 - Cho phép kết hợp thành các đơn vi nhớ lớn
- Static
 - Nhanh hơn, cổng kênh hơn
 - Cho phép xây dựng các bộ nhớ Cache



Read Only Memory (ROM)

- Lưu giữ thông tin cố định - permanent storage, nonvolatile
 - Microprogramming
 - Library subroutines
 - Systems programs (BIOS)
 - Function tables
- Kiểu:
 - Written during manufacture
 - Very expensive for small runs
 - Programmable (once)
 - PROM
 - Needs special equipment to program
 - Read “mostly”
 - Erasable Programmable (EPROM)
 - Erased by UV
 - Electrically Erasable (EEPROM)
 - Takes much longer to write than read
 - Flash memory
 - Erase whole memory electrically



Programmable ROM (PROM)

- Less expensive alternative
- Nonvolatile and may be written into only once
- Writing process is performed electrically and may be performed by supplier or customer at a time later than the original chip fabrication
- Special equipment is required for the writing process
- Provides flexibility and convenience
- Attractive for high volume production runs



Read-Mostly Memory

EPROM

Erasable programmable read-only memory

Erase process can be performed repeatedly

More expensive than PROM but it has the advantage of the multiple update capability

EEPROM

Electrically erasable programmable read-only memory

Can be written into at any time without erasing prior contents

Combines the advantage of non-volatility with the flexibility of being updatable in place

More expensive than EPROM

Flash Memory

Intermediate between EPROM and EEPROM in both cost and functionality

Uses an electrical erasing technology, does not provide byte-level erasure

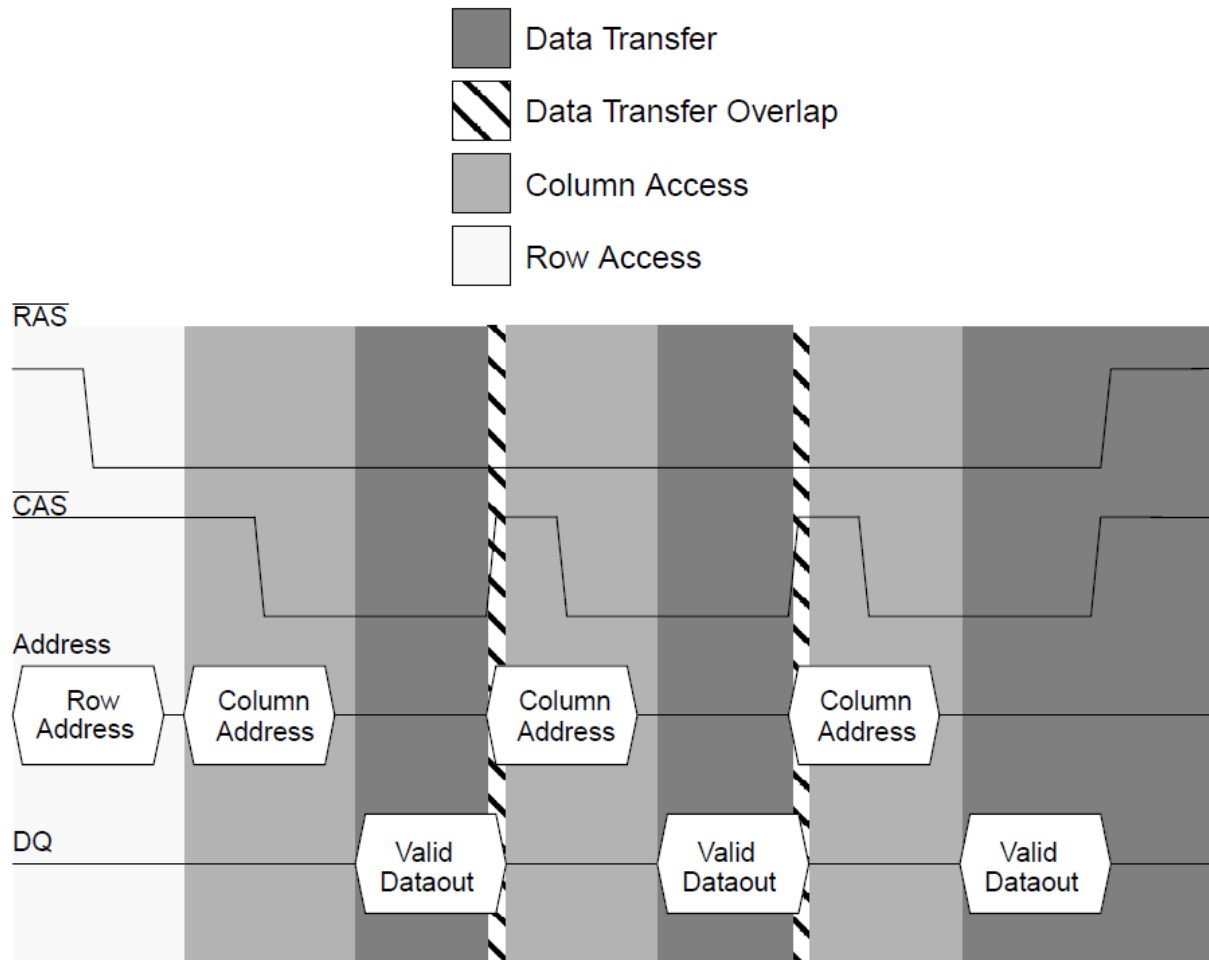
Microchip is organized so that a section of memory cells are erased in a single action or “flash”



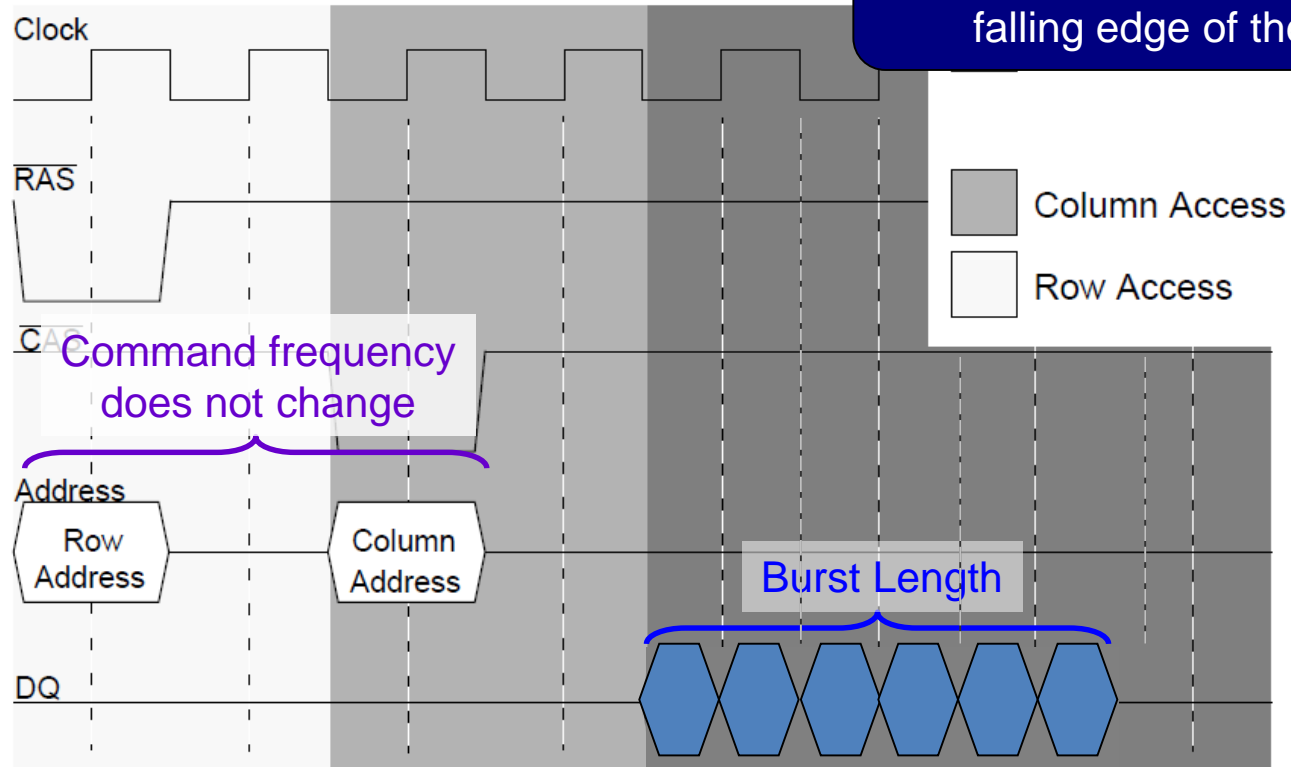
Synchronous DRAM (SDRAM)

- Truy cập được đồng bộ hoá với một đồng hồ bên ngoài
 - Địa chỉ được truyền đến RAM
 - RAM tìm dữ liệu (CPU đợi như DRAM thông thường)
 - Khi SDRAM chuyển dữ liệu theo thời gian đồng bộ với system clock, CPU biết được khi nào dữ liệu sẵn sàng
- ⇒ **CPU không cần phải chờ và có thể làm việc khác**
 - Burst mode: cho phép SDRAM thiết lập dòng dữ liệu theo từng block
 - Chỉ chuyển dữ liệu 1 lần trong 1 chu kỳ đồng hồ
- DDR-SDRAM - Double-data-rate 1 SDRAM
 - Gửi dữ liệu 2 lần trong một chu kỳ đồng hồ (leading & trailing edge)
- DDR2-SDRAM - Double-data-rate 2 SDRAM
- DDR3-SDRAM - Double-data-rate 3 SDRAM
- Cache DRAM: (misubishi)
 - Tích hợp SRAM cache (16k) vào trong DRAM chip

Việc truy cập là không đồng bộ: được kiểm soát bởi các tín hiệu RAS & CAS, các tín hiệu này có thể được sinh ra ngẫu nhiên




SDRAM Read Timing



Timing figures taken from "A Performance Comparison of Contemporary DRAM Architectures" by Cuppu, Jacob, Davis and Mudge

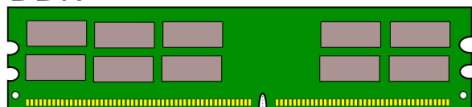
DDR SDRAM

DDR SDRAM Standard	Release year	Bus Clock (MHz)	Internal Memory Clock (MHz)	Prefetch (min burst)	Transfer Rate (MT/s)	Voltage	DIMM pins	SO-DIMM pins	MicroDIMM pins
DDR1	2000	100–200	100–200	2n	200–400	2.5/2.6	184	200	172
DDR2	2003	200–533.33	100–266.67	4n	400–1066.67	1.8	240	200	214
DDR3	2007	400–1066.67	100–266.67	8n	800–2133.33	1.5/1.35	240	204	214
DDR4	2014	1066.67–2133.33	133.33–266.67	8n	2133.33–4266.67	1.05/1.2	288	256	—

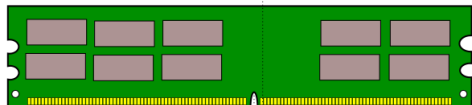
Standard name	Memory clock	Cycle time	I/O Bus clock	Data transfers per second	Module name	Peak transfer rate
DDR-200	100 MHz	10 ns ^[1]	100 MHz	200 Million	PC-1600	1600 MB/s
DDR-266	133 MHz	7.5 ns	133 MHz	266 Million	PC-2100	2100 MB/s
DDR-300	150 MHz	6.67 ns	150 MHz	300 Million	PC-2400	2400 MB/s
DDR-333	166 MHz	6 ns	166 MHz	333 Million	PC-2700	2700 MB/s
DDR-400	200 MHz	5 ns	200 MHz	400 Million	PC-3200	3200 MB/s
DDR2-400	100 MHz	10 ns	200 MHz	400 Million	PC2-3200	3200 MB/s
DDR2-533	133 MHz	7.5 ns	266 MHz	533 Million	PC2-4200 PC2-4300 ¹	4266 MB/s
DDR2-667	166 MHz	6 ns	333 MHz	667 Million	PC2-5300 PC2-5400 ¹	5333 MB/s
DDR2-800	200 MHz	5 ns	400 MHz	800 Million	PC2-6400	6400 MB/s
DDR2-1066	266 MHz	3.75 ns	533 MHz	1066 Million	PC2-8500	8533 MB/s
DDR3-800	100 MHz	10 ns	400 MHz	800 Million	PC3-6400	6400 MB/s
DDR3-1066	133 MHz	7.5 ns	533 MHz	1066 Million	PC3-8500	8533 MB/s
DDR3-1333	166 MHz	6 ns	667 MHz	1333 Million	PC3-10600	10667 MB/s[4] 
DDR3-1600	200 MHz	5 ns	800 MHz	1600 Million	PC3-12800	12800 MB/s

DDR 2,3,4

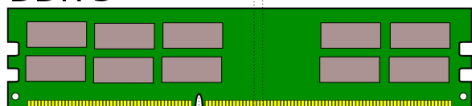
DDR



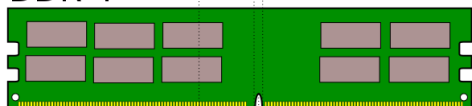
DDR 2



DDR 3

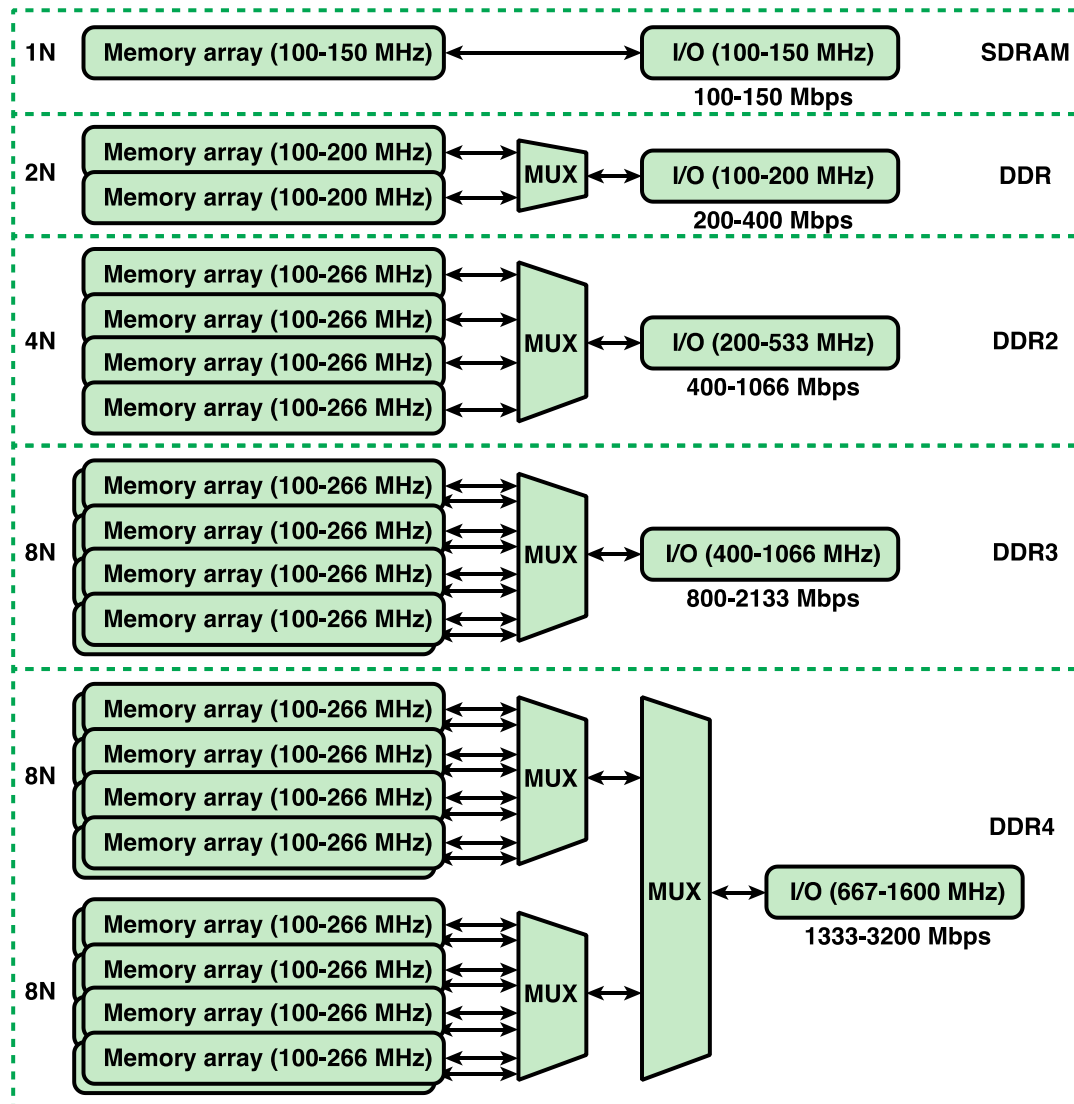


DDR 4



DDR SDRAM Standard	Internal rate (MHz)	Bus clock (MHz)	<u>Prefetch</u>	Data rate (MT/s)	Transfer rate (GB/s)	Voltage (V)
SDRAM	100-166	100-166	1n	100-166	0.8-1.3	3.3
DDR	133-200	133-200	2n	266-400	2.1-3.2	2.5/2.6
DDR2	133-200	266-400	4n	533-800	4.2-6.4	1.8
DDR3	133-200	533-800	8n	1066-1600	8.5-14.9	1.35/1.5
DDR4	133-200	1066-1600	8n	2133-3200	17-21.3	1.2

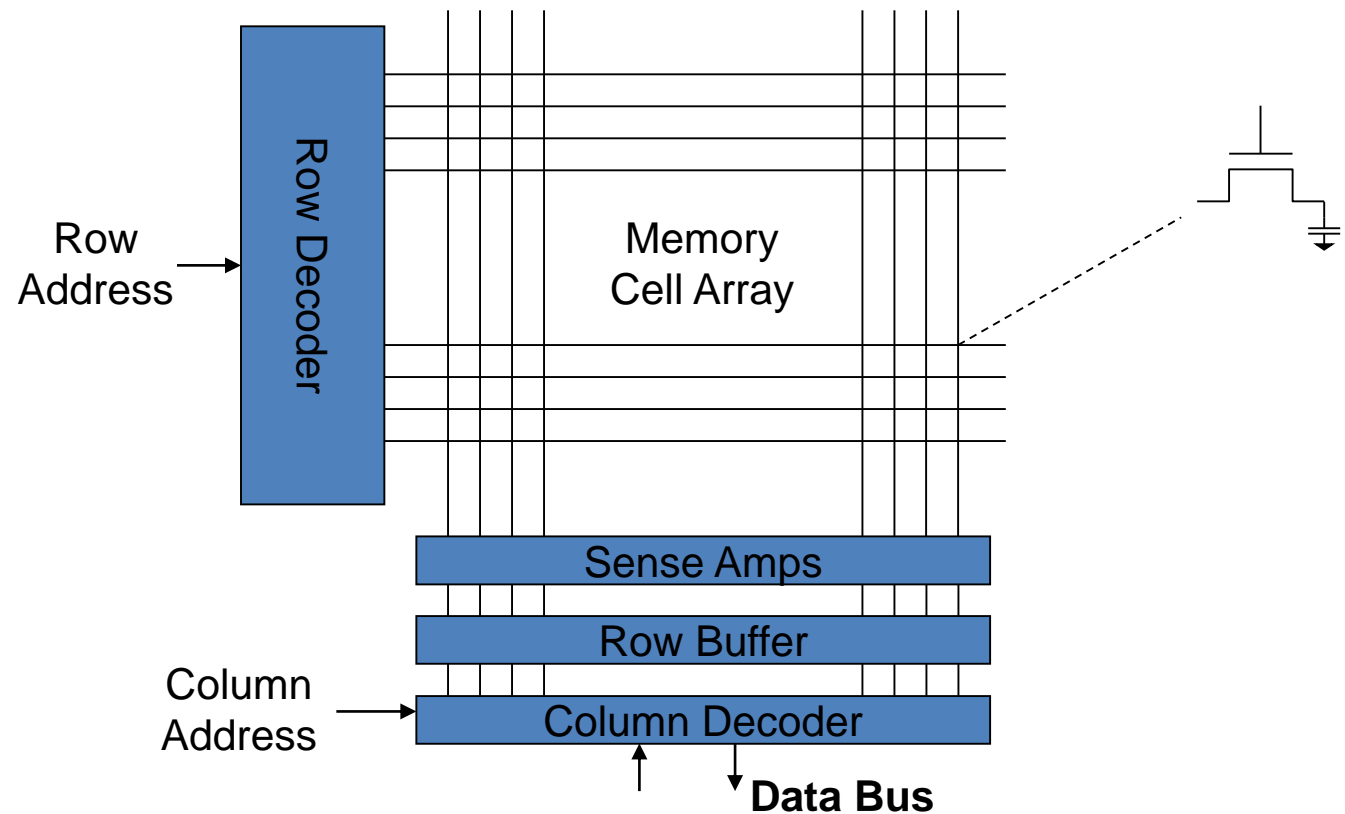
DDR Generation



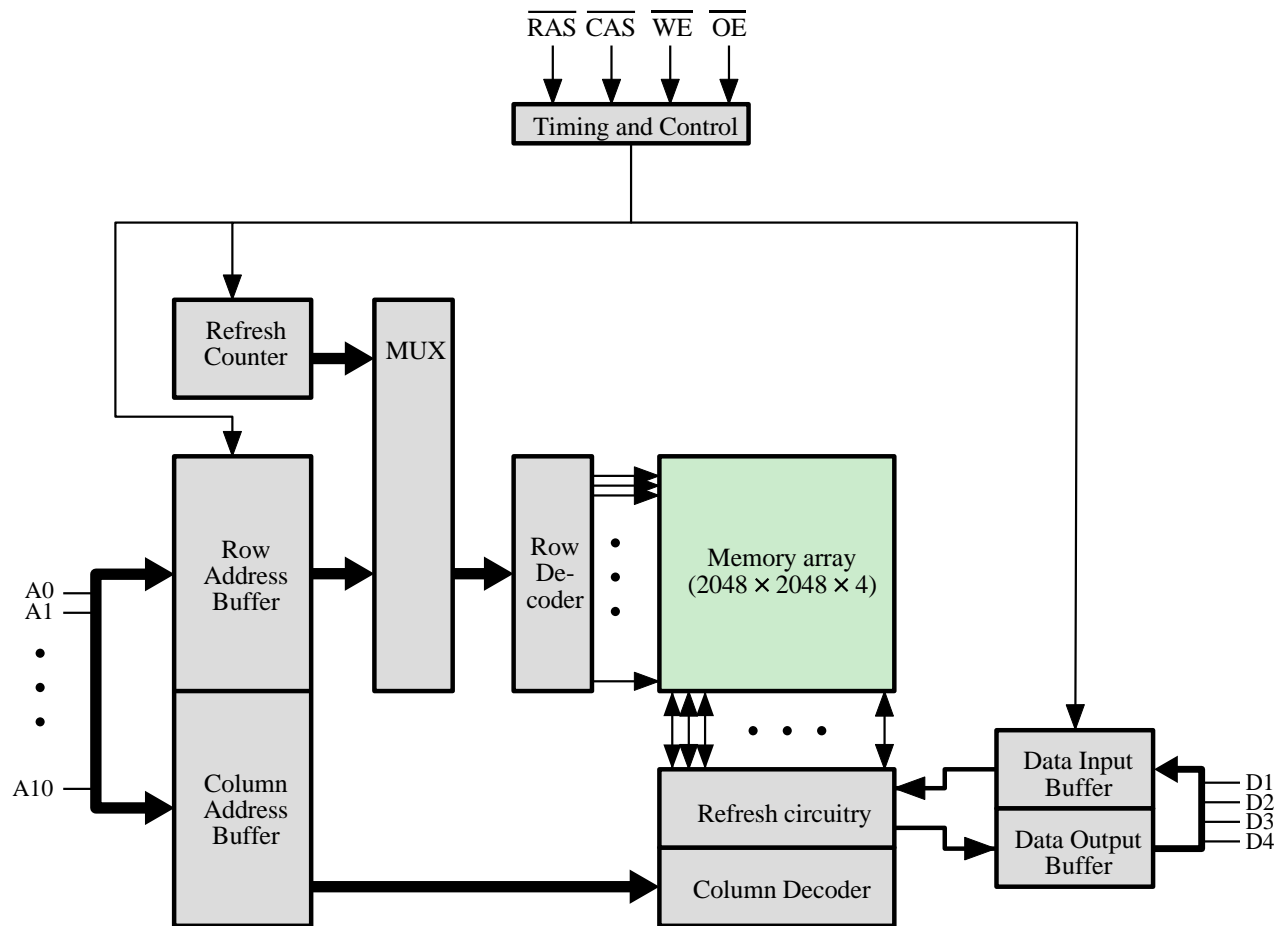


Tổ chức bộ nhớ

- Mạch nhớ W từ B bits được tổ chức dưới dạng 1 ma trận n hàng và m cột từ nhớ B bits
 - Cần $n \cdot m = W$ words, kích thước bus địa chỉ = $\lceil \log_2 W \rceil$
 - B đường dữ liệu



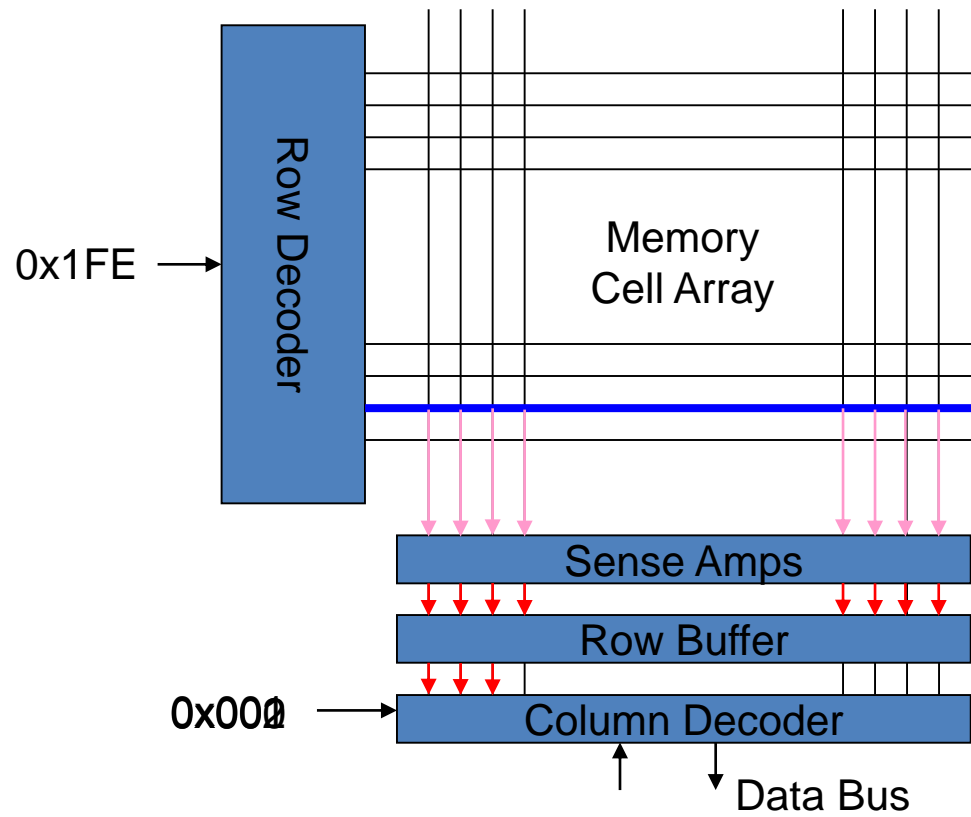
Ví dụ : 16 Mb DRAM (4M x 4)



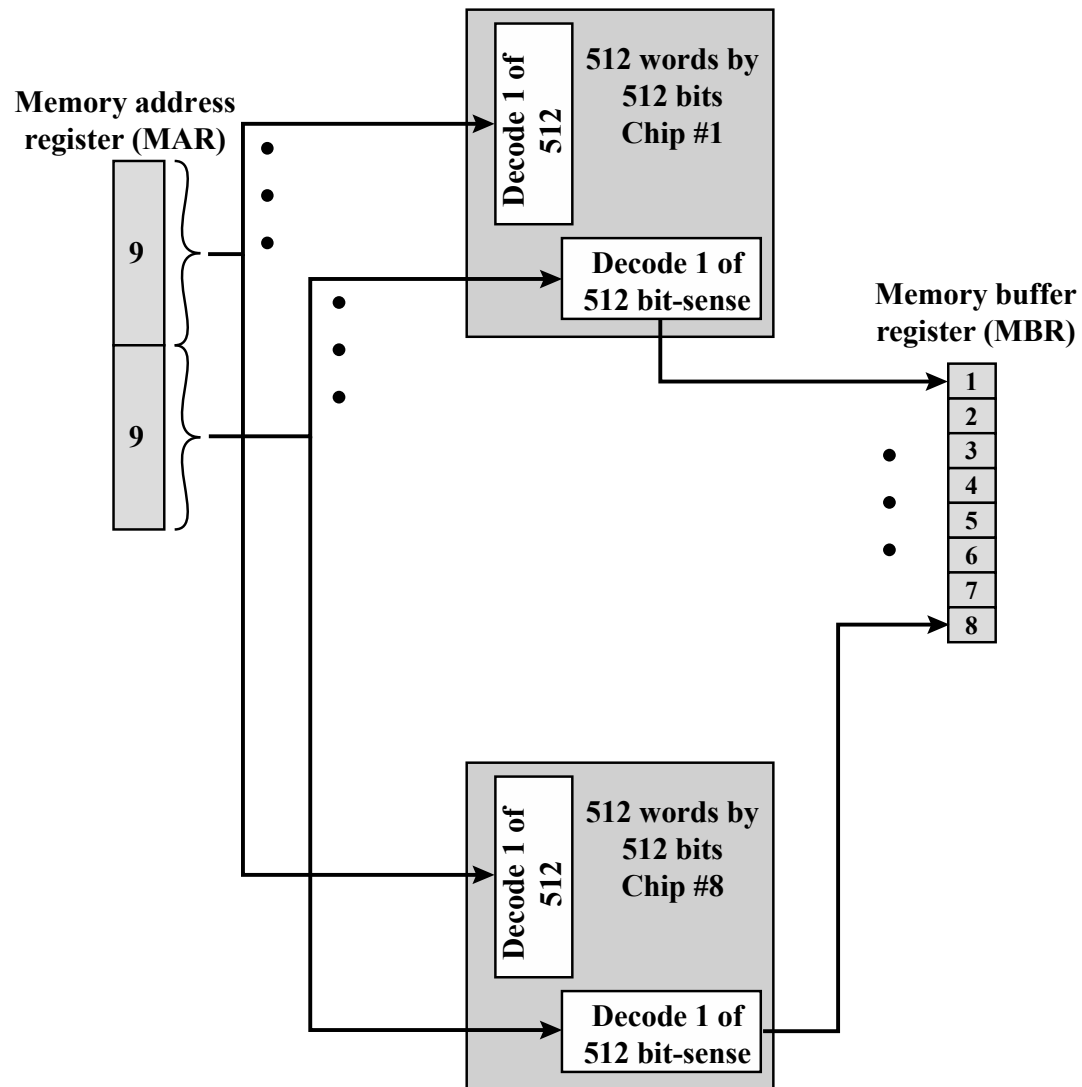
RAS: Row Address Select; CAS: Column Address Select

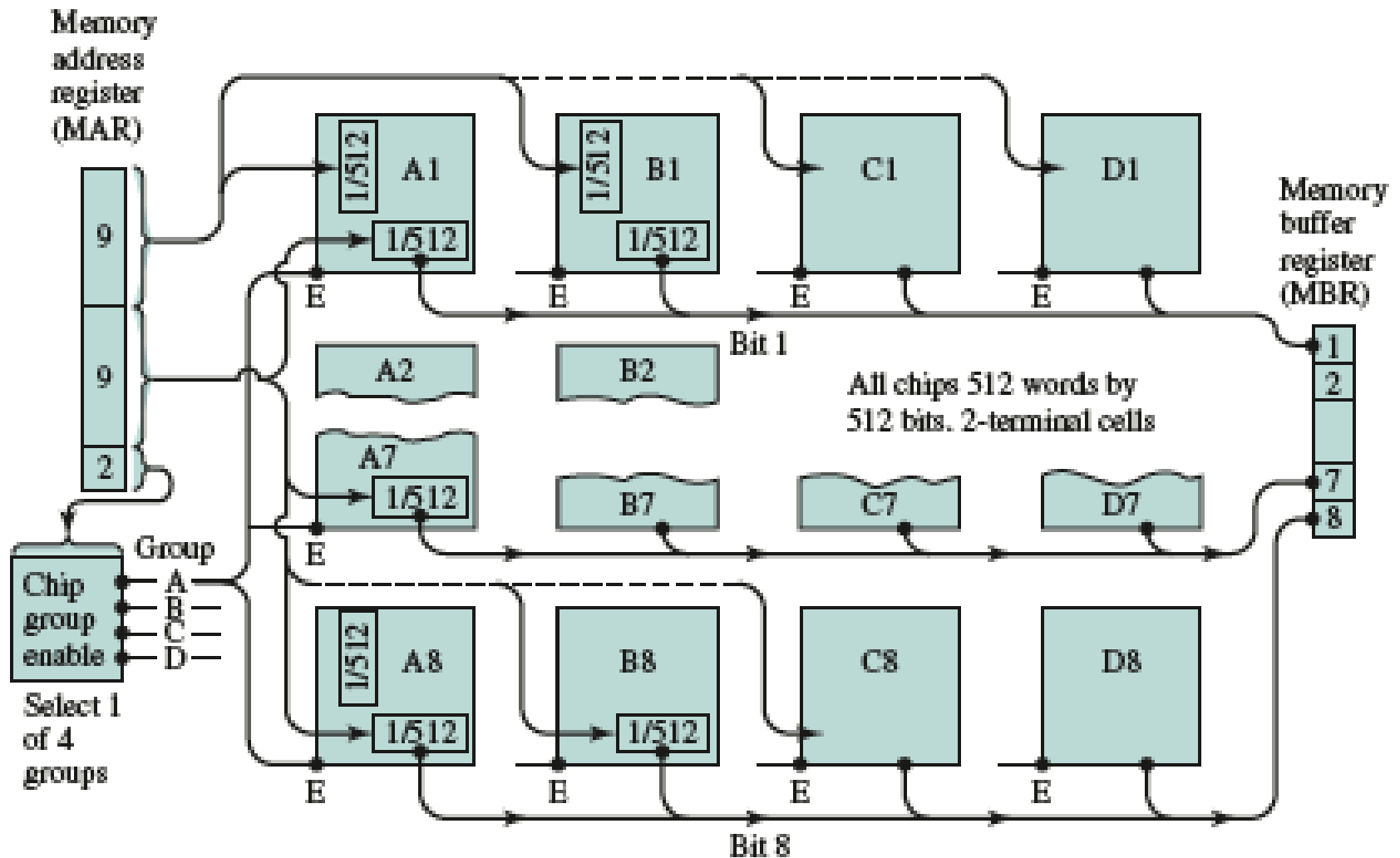
OE: Output Enable ; WE: Write Enable

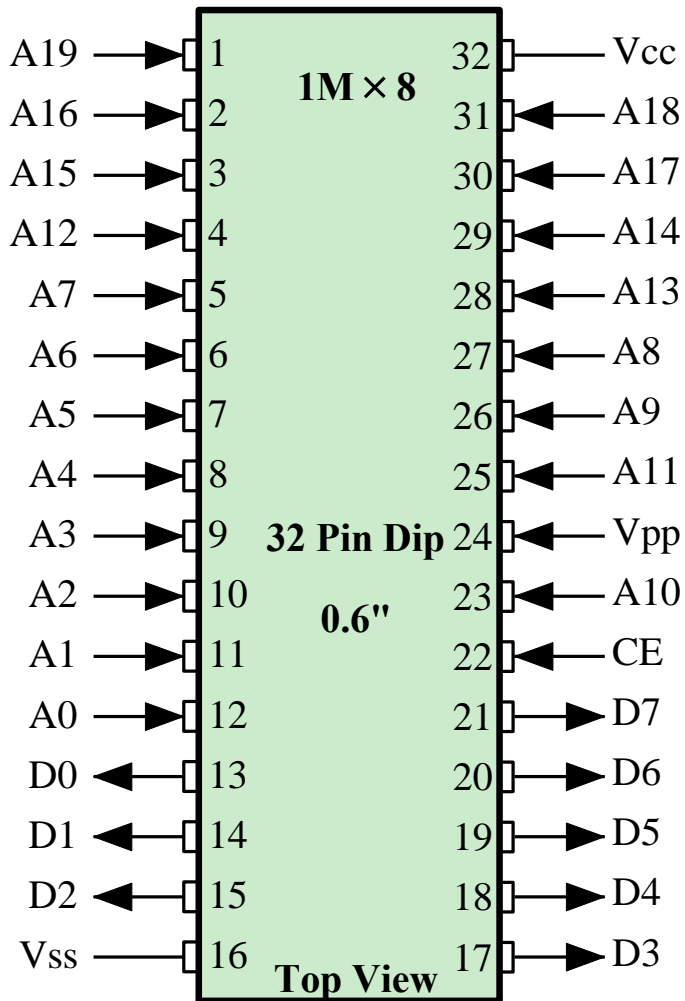
DRAM Read Operation



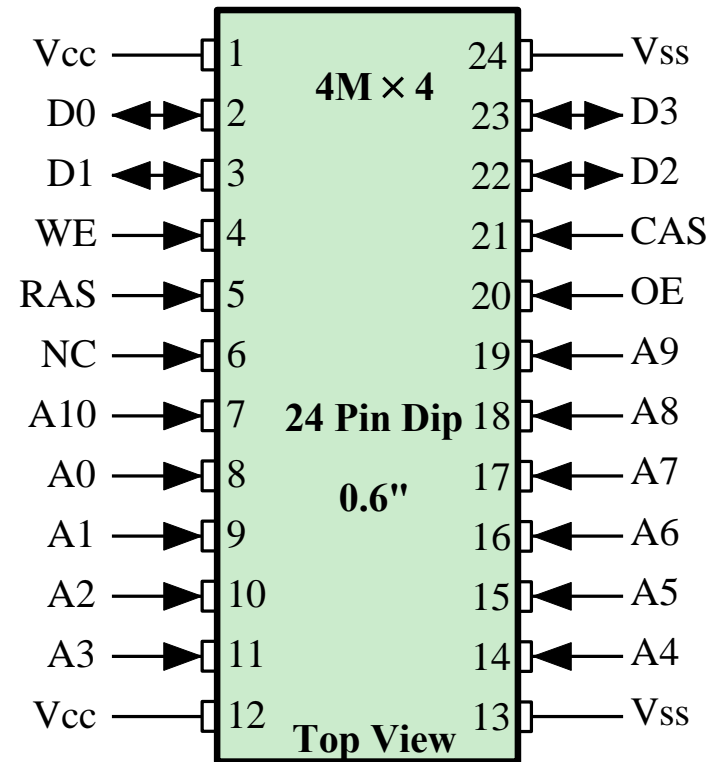
Multi-bank memory



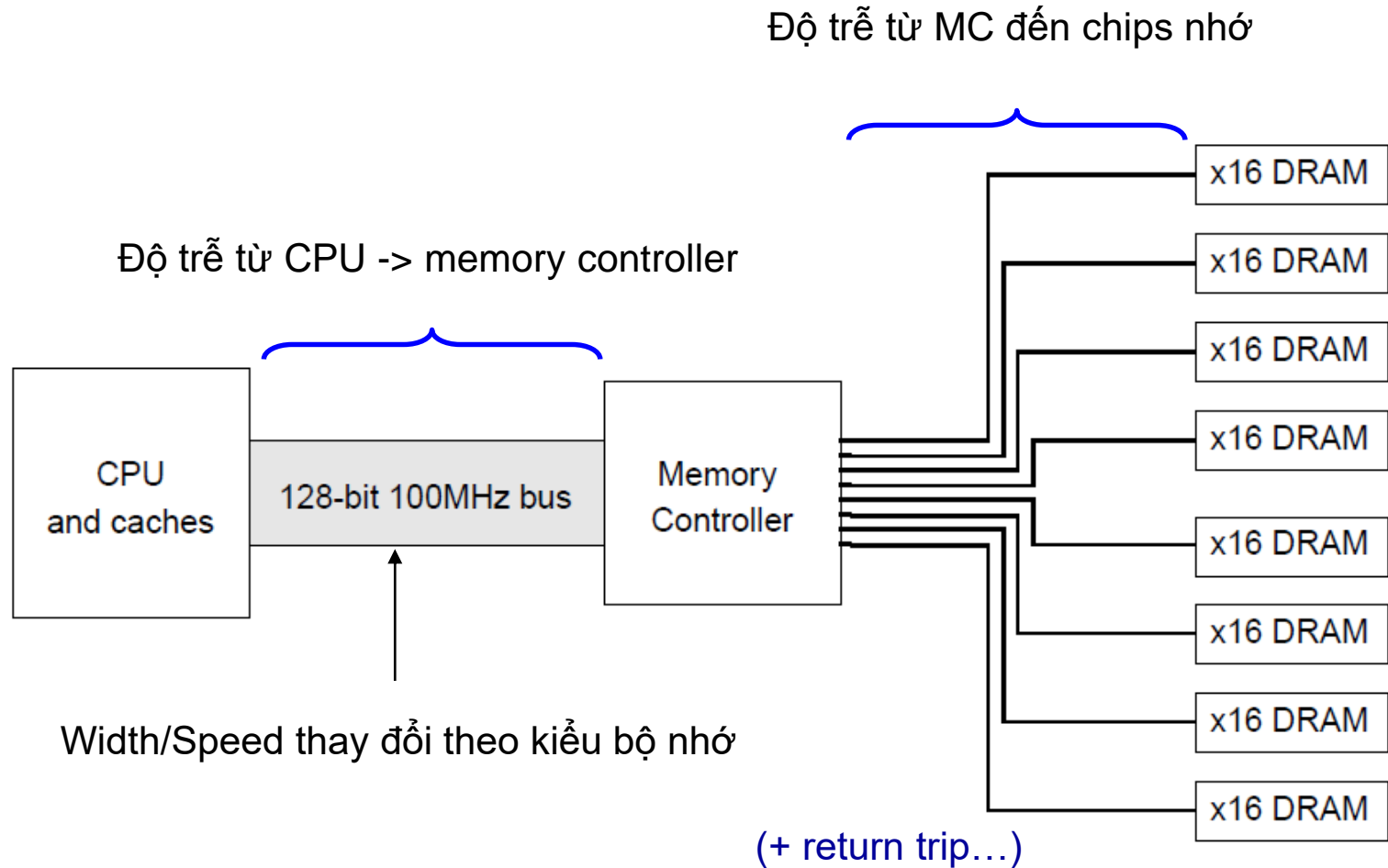


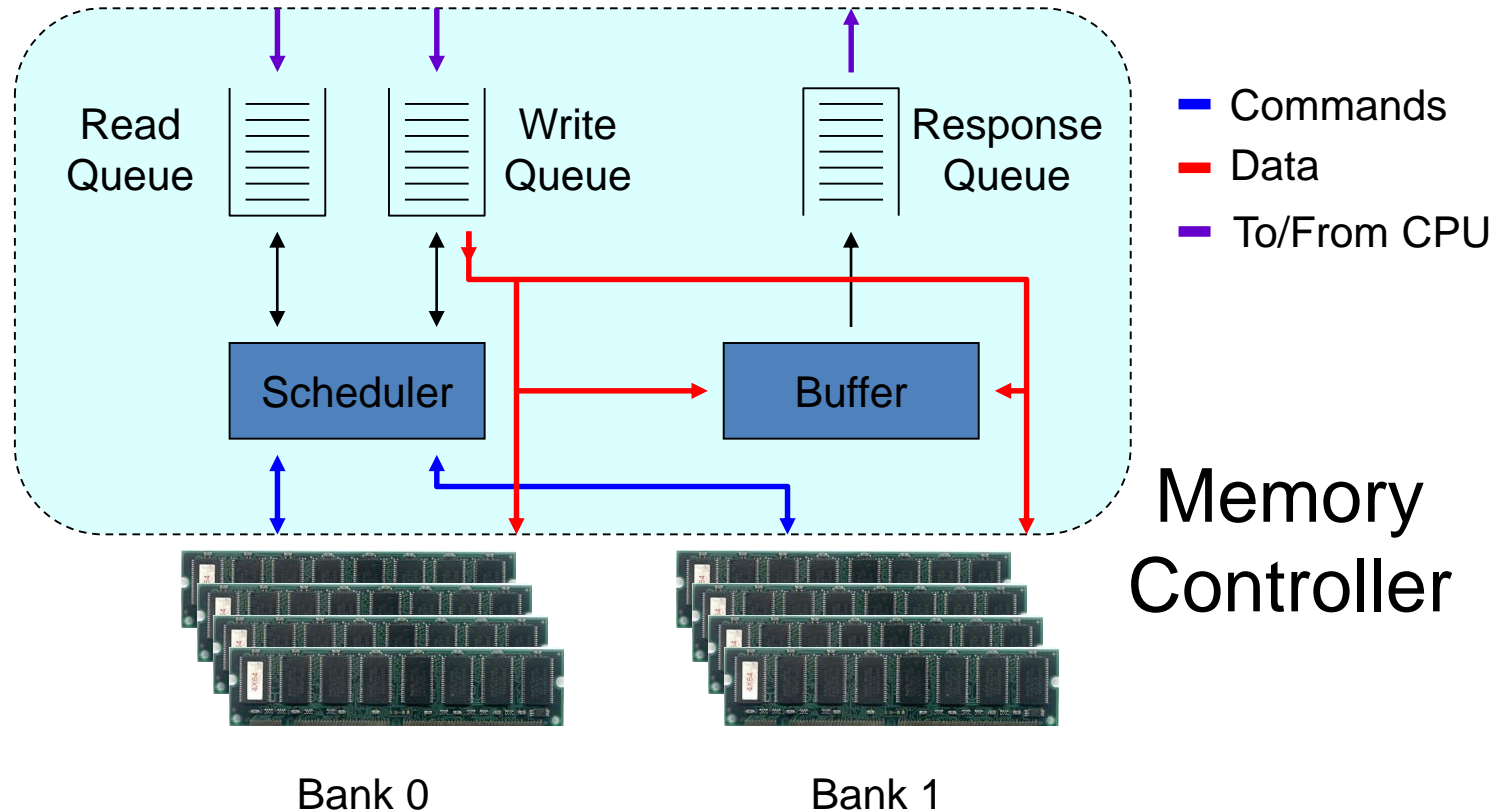


(a) 8 Mbit EPROM



(b) 16 Mbit DRAM

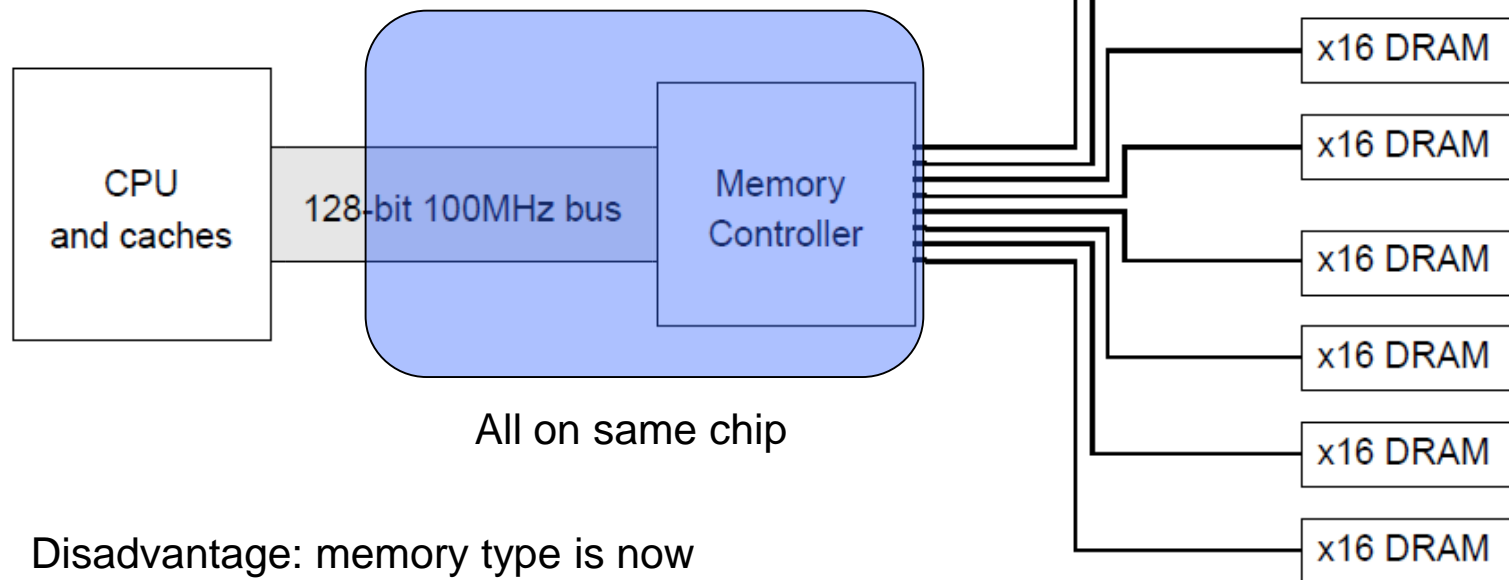




On-Chip Memory Controller

Also: cơ chế lập lịch phức tạp hơn

Memory controller thi hành cùng tốc độ
CPU thay vì tốc độ xung
FSB



Disadvantage: memory type is now
tied to the CPU implementation



Memory Latency is *Long*

- Thường từ 60-100ns
- Lãng phí thời gian của CPU:
 - 2GHz CPU
 - $\rightarrow 0.5\text{ns} / \text{cycle}$
 - 100ns memory $\rightarrow 200$ cycle CPU latency!
- Solution: sử dụng bộ nhớ đệm có tốc độ cao hơn - Caches



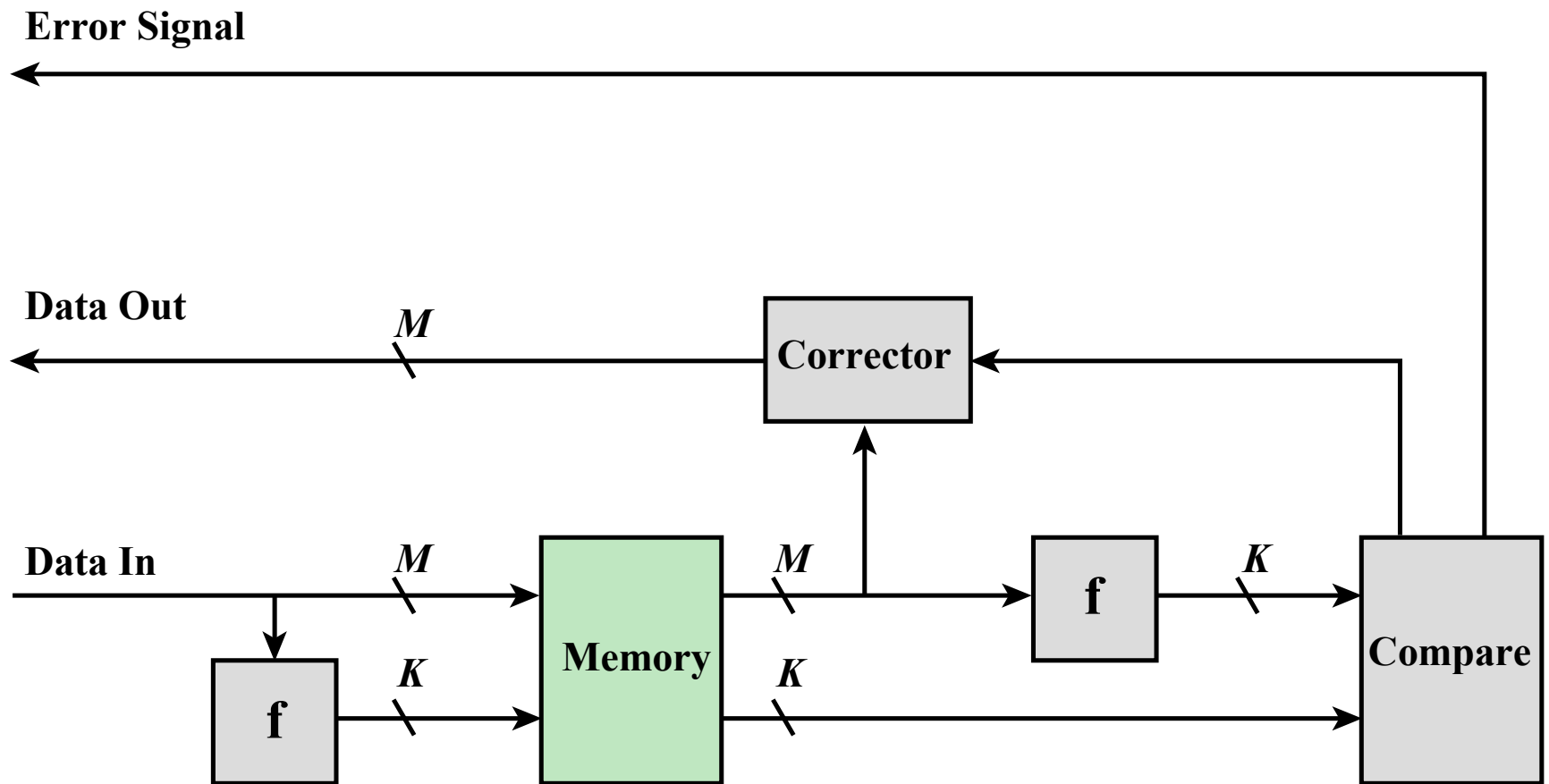
Memory Latency is *Long*

- Thường từ 60-100ns
- Lãng phí thời gian của CPU:
 - 2GHz CPU
 - $\rightarrow 0.5\text{ns} / \text{cycle}$
 - 100ns memory $\rightarrow 200$ cycle CPU latency!
- Solution: sử dụng bộ nhớ đệm có tốc độ cao hơn - Caches

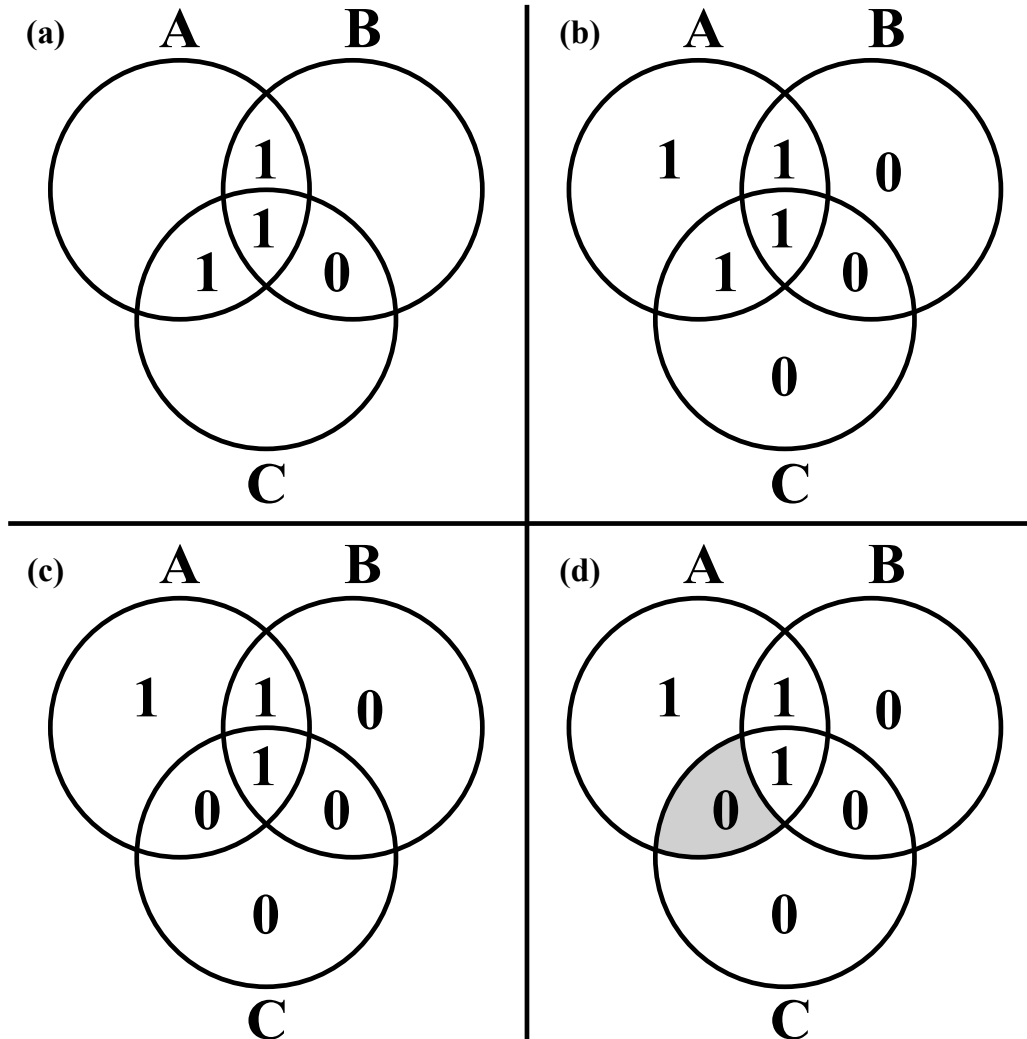


Error Correction

- **Hard Failure**
 - Permanent physical defect
 - Memory cell or cells affected cannot reliably store data but become stuck at 0 or 1 or switch erratically between 0 and 1
 - Can be caused by:
 - Harsh environmental abuse
 - Manufacturing defects
 - Wear
- **Soft Error**
 - Random, non-destructive event that alters the contents of one or more memory cells
 - No permanent damage to memory
 - Can be caused by:
 - Power supply problems
 - Alpha particles



- Parity bit



Error correcting code efficiency

Data Bits	Single-Error Correction		Single-Error Correction/ Double-Error Detection	
	Check Bits	% Increase	Check Bits	% Increase
8	4	50	5	62.5
16	5	31.25	6	37.5
32	6	18.75	7	21.875
64	7	10.94	8	12.5
128	8	6.25	9	7.03
256	9	3.52	10	3.91



Parity bit layout

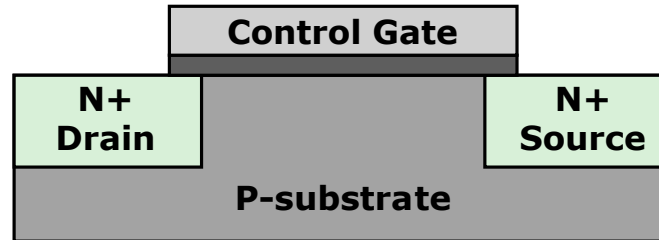
Bit Position	12	11	10	9	8	7	6	5	4	3	2	1
Position Number	1100	1011	1010	1001	1000	0111	0110	0101	0100	0011	0010	0001
Data Bit	D8	D7	D6	D5		D4	D3	D2		D1		
Check Bit					C8				C4		C2	C1

Bit position	12	11	10	9	8	7	6	5	4	3	2	1
Position number	1100	1011	1010	1001	1000	0111	0110	0101	0100	0011	0010	0001
Data bit	D8	D7	D6	D5		D4	D3	D2		D1		
Check bit					C8				C4		C2	C1
Word stored as	0	0	1	1	0	1	0	0	1	1	1	1
Word fetched as	0	0	1	1	0	1	1	0	1	1	1	1
Position Number	1100	1011	1010	1001	1000	0111	0110	0101	0100	0011	0010	0001
Check Bit					0				0		0	1

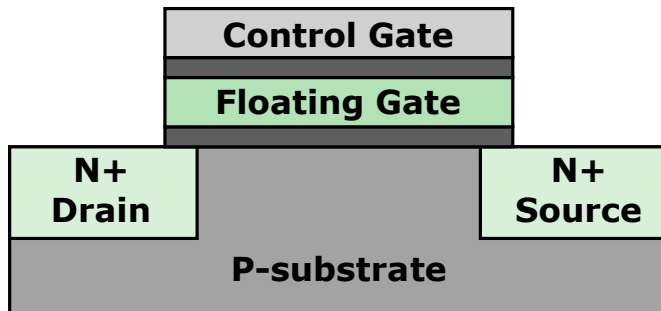


Flash Memory

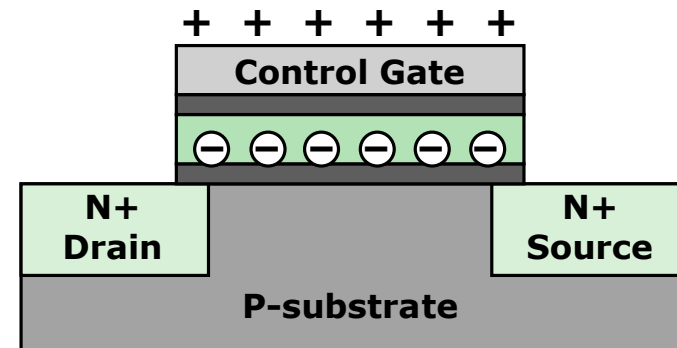
- Used both for internal memory and external memory applications
- First introduced in the mid-1980's
- Is intermediate between EPROM and EEPROM in both cost and functionality
- Uses an electrical erasing technology like EEPROM
- It is possible to erase just blocks of memory rather than an entire chip
- Gets its name because the microchip is organized so that a section of memory cells are erased in a single action
- Does not provide byte-level erasure
- Uses only one transistor per bit so it achieves the high density of EPROM



(a) Transistor structure



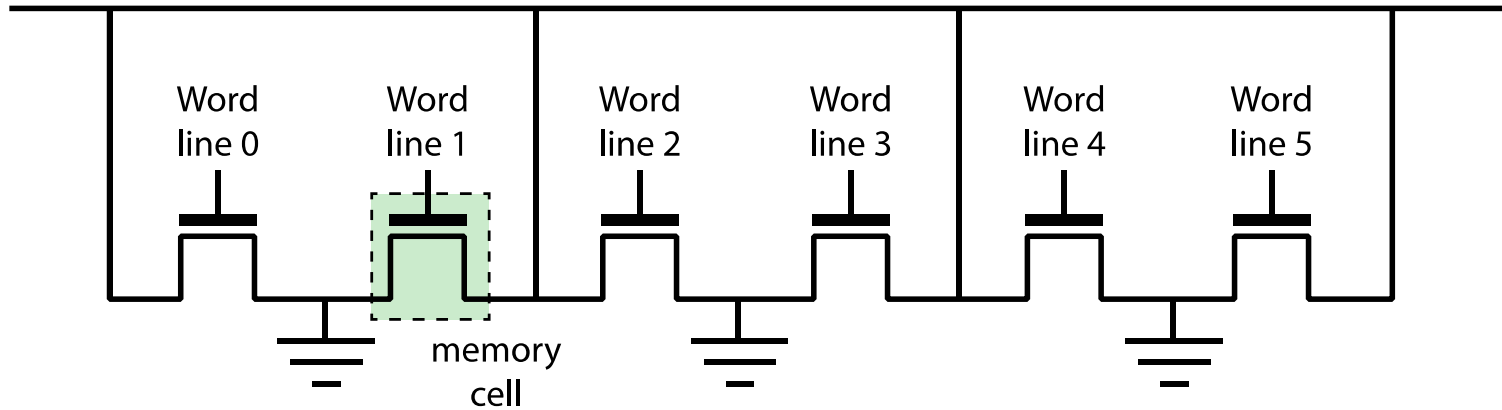
(b) Flash memory cell in one state



(c) Flash memory cell in zero state

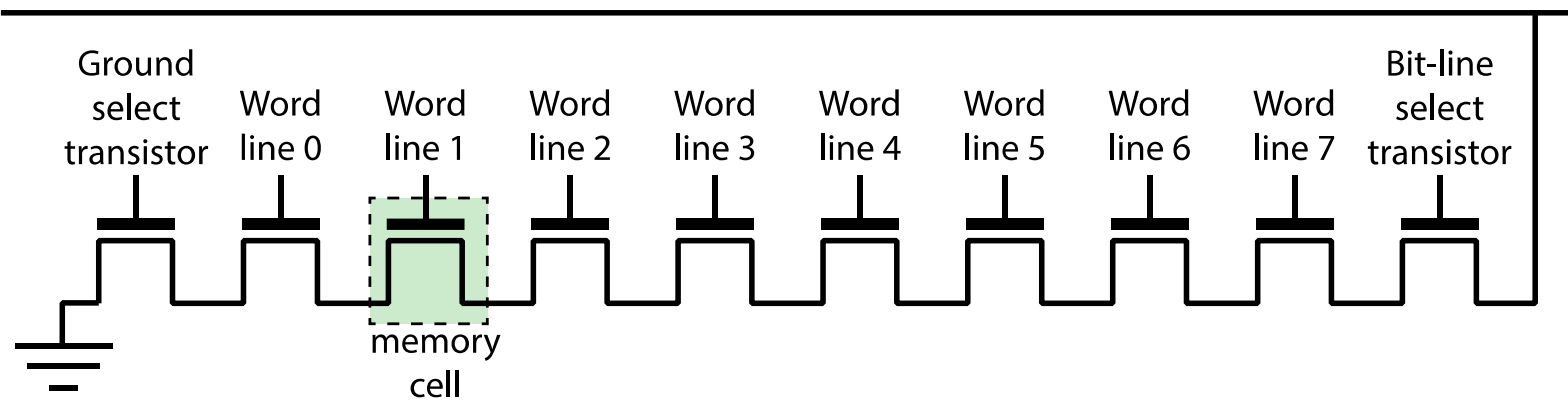
Accessing flash memory

Bit line



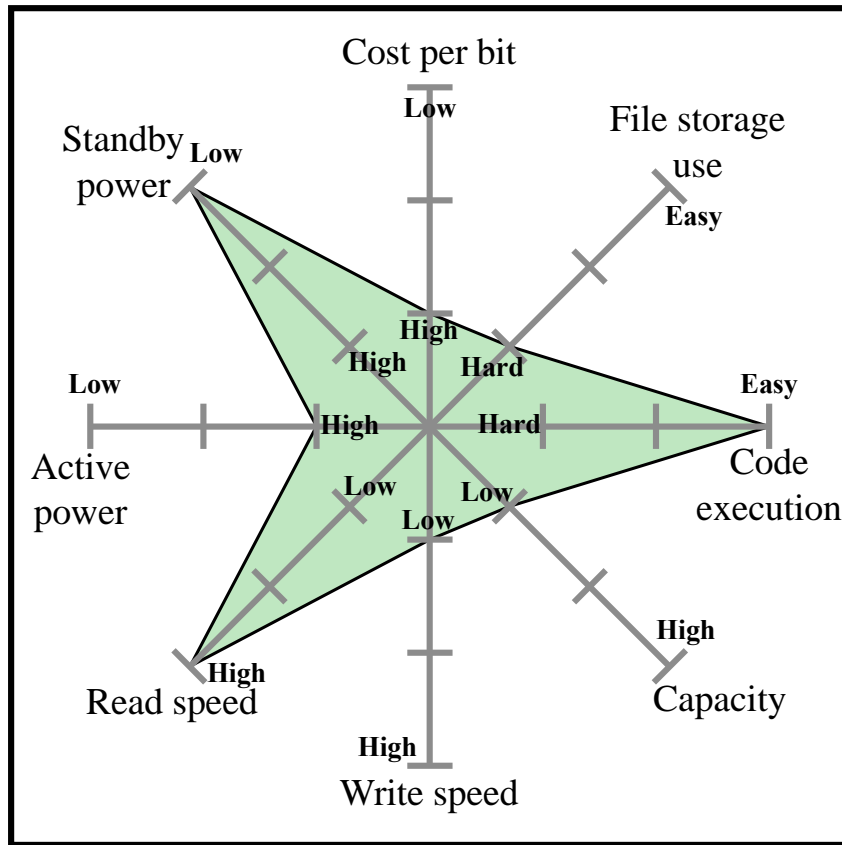
(a) NOR flash structure

Bit line

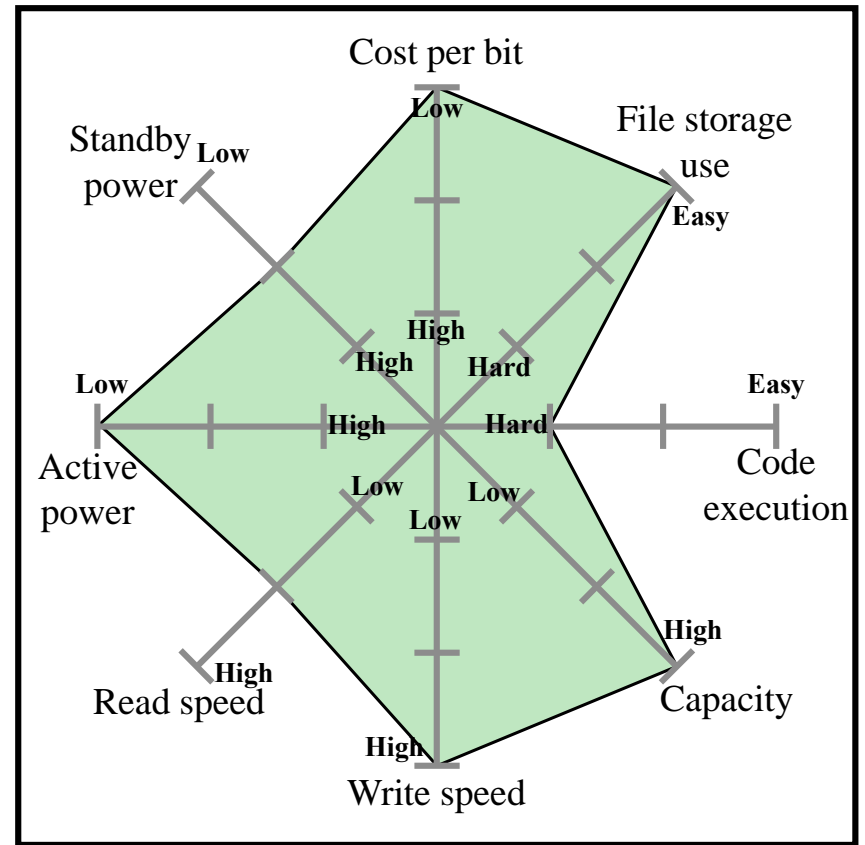


(b) NAND flash structure

NOR Flash vs. NAND Flash

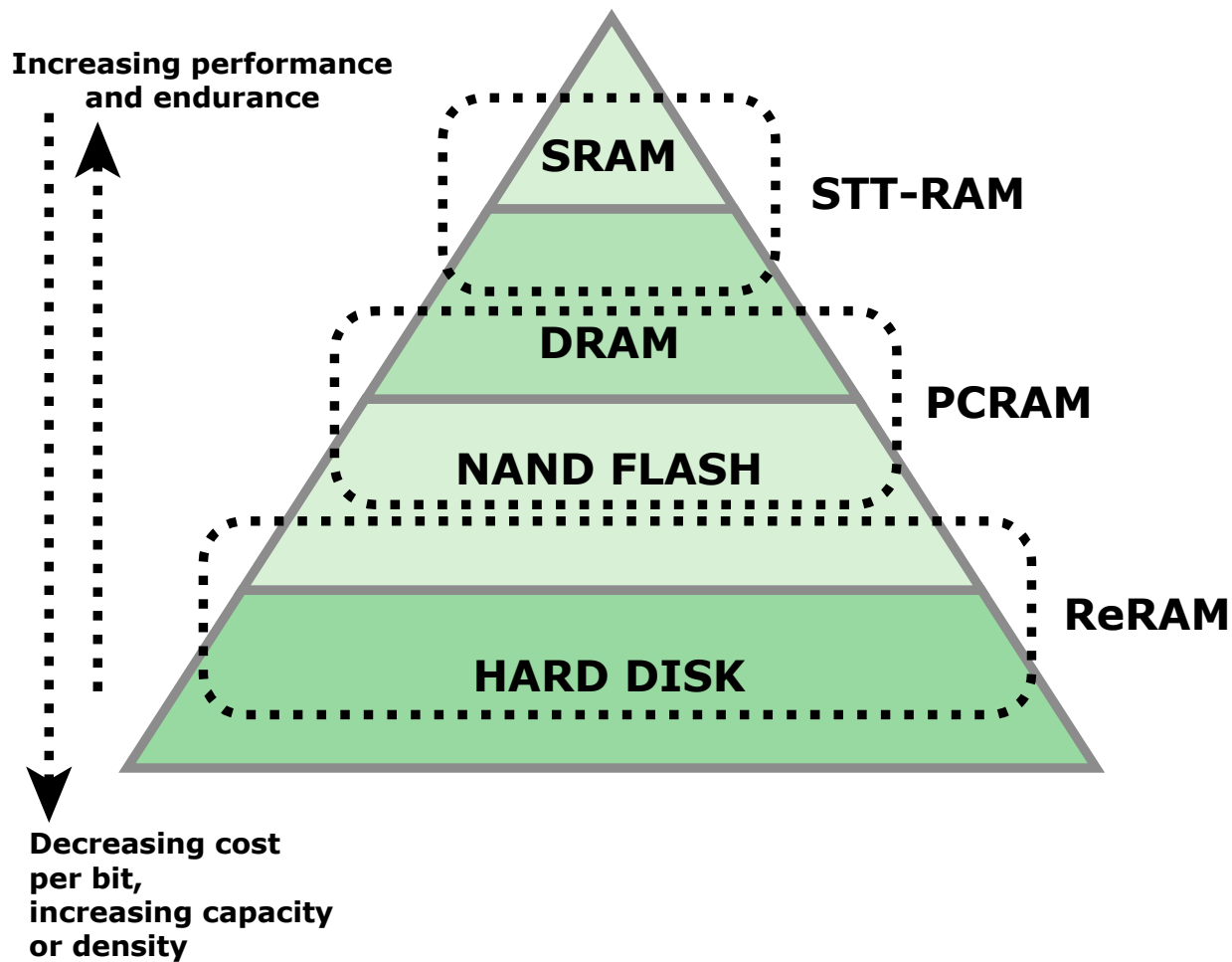


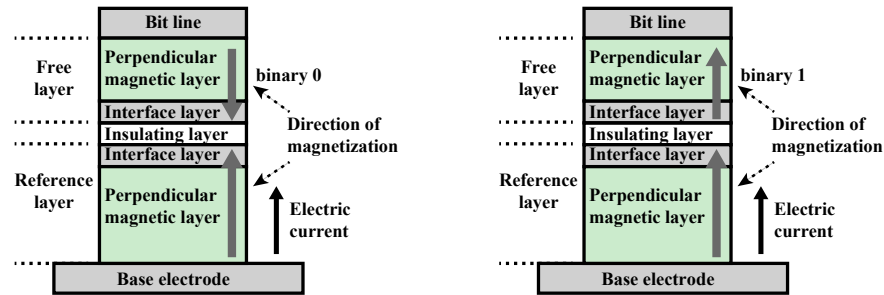
(a) NOR



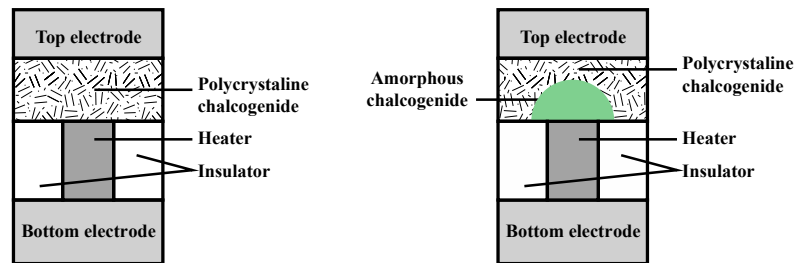
(b) NAND

Performance vs. Cost

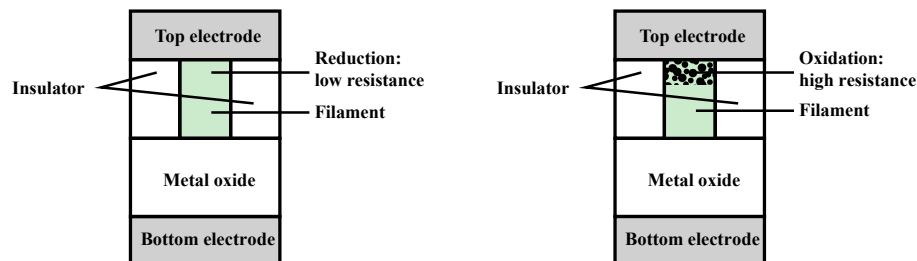




(a) STT-RAM

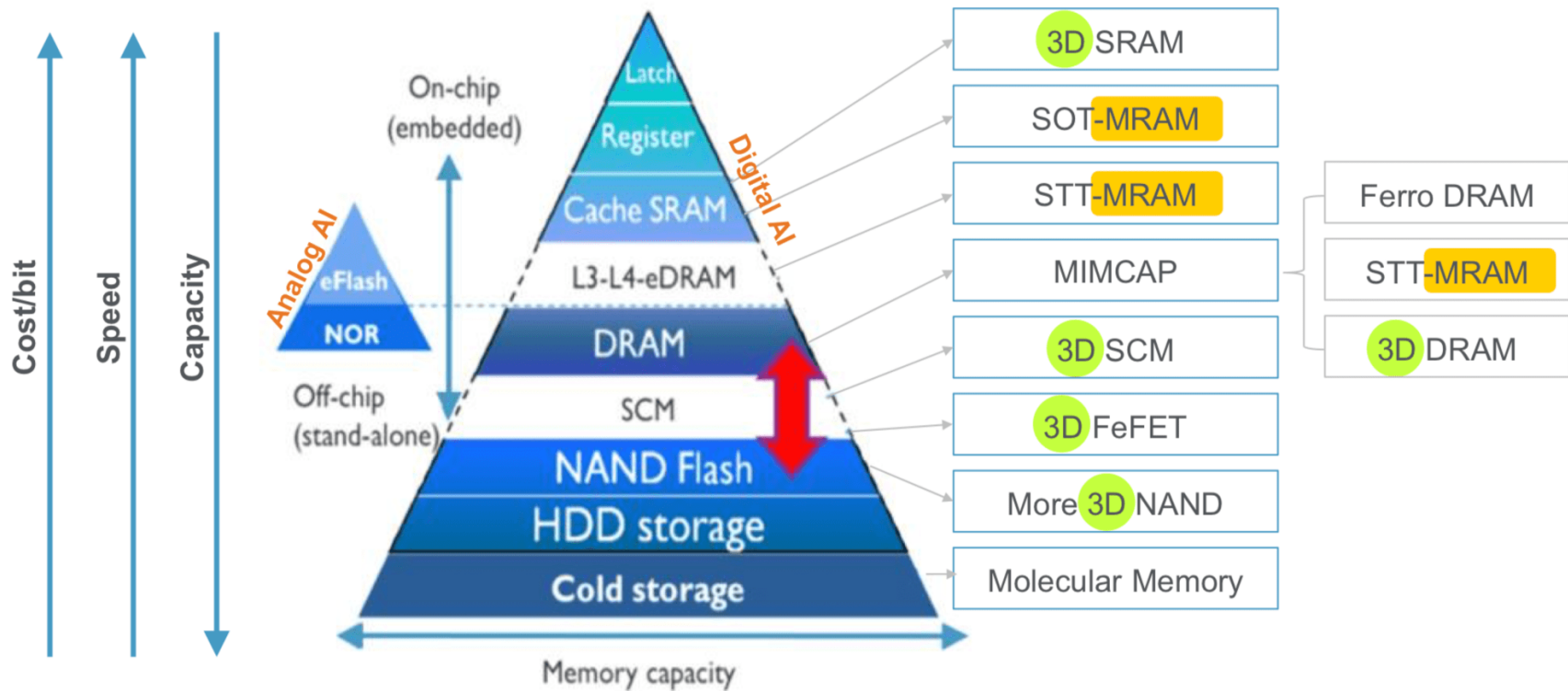


(b) PCRAM



(c) ReRAM

Emerging memories





Tổng kết

- Khái niệm về bộ nhớ trong máy tính, các đặc điểm chính, ...
- Bộ nhớ chính
 - Nguyên tắc
 - Phân loại
 - Tổ chức bộ nhớ lớn