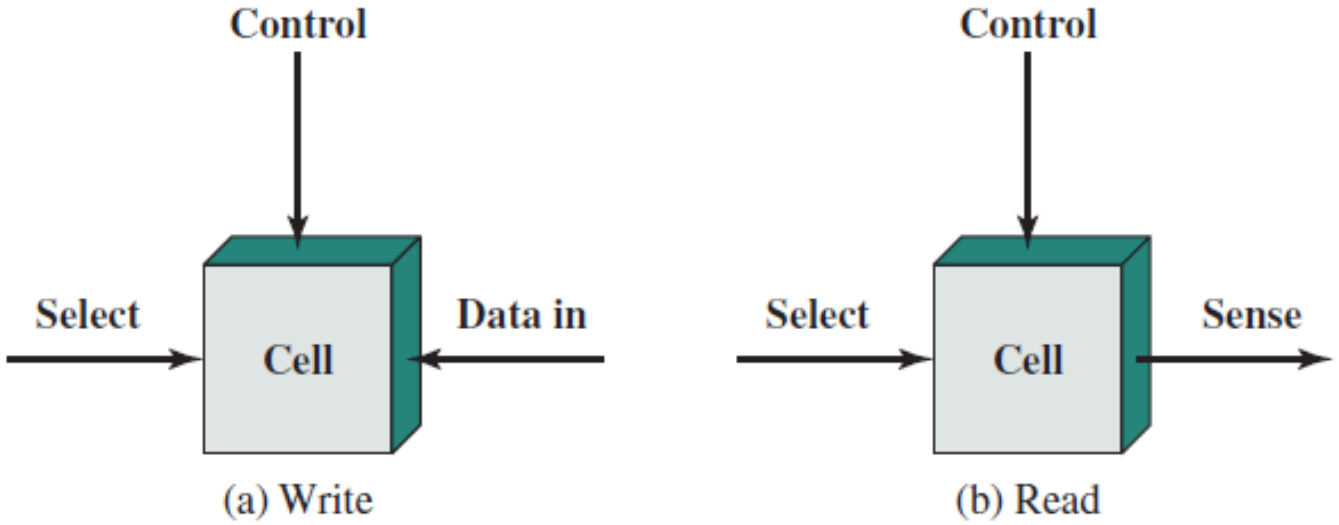
LAU
الجامعة اللبنانية الأميركية
Lebanese American University

Internal Memory

Chapter 6

Based on:
William Stallings
Computer Organization and Architecture, 11th Global Edition

Memory Cell Operation



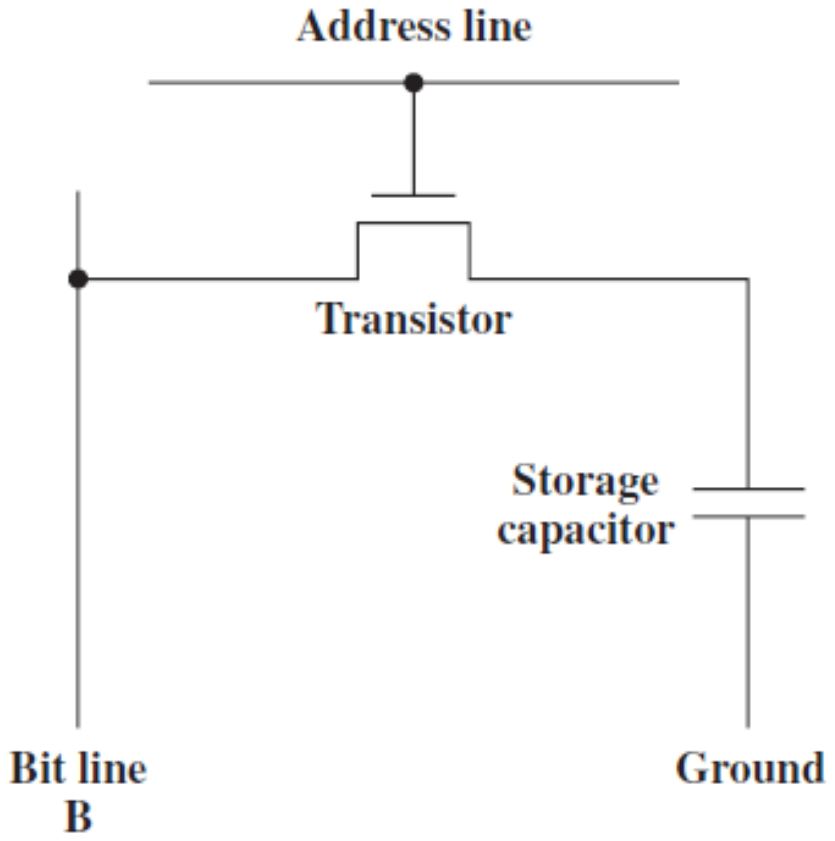
Semiconductor Memory Types

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)			Electrically	
Erasable PROM (EPROM)	Read-mostly memory	UV light, chip-level		
Electrically Erasable PROM (EEPROM)		Electrically, byte-level		
Flash memory		Electrically, block-level		

Semiconductor Memory

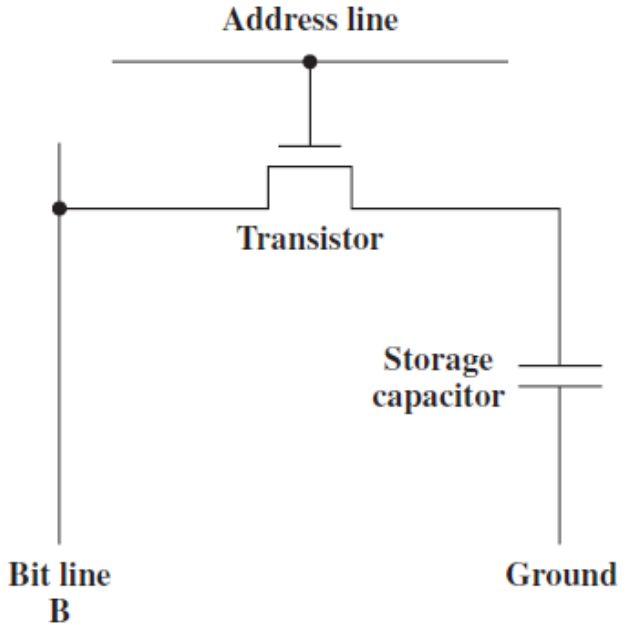
- RAM
  - Misuse of the term as all semiconductor memory is random access
  - Read/Write
  - Volatile
  - Temporary storage
  - Static or dynamic

Dynamic RAM Structure



DRAM Operation

- Address line active when bit read or written
  - Transistor switch closed (current flows)
- Write
  - Voltage to bit line
    - High for 1, low for 0
  - Then signal address line
    - Transfers charge to capacitor
- Read
  - Address line selected
    - Transistor turns on
  - Charge from capacitor fed via bit line to sense amplifier
    - Compares with reference value to determine 1 or 0
  - Capacitor charge must be restored



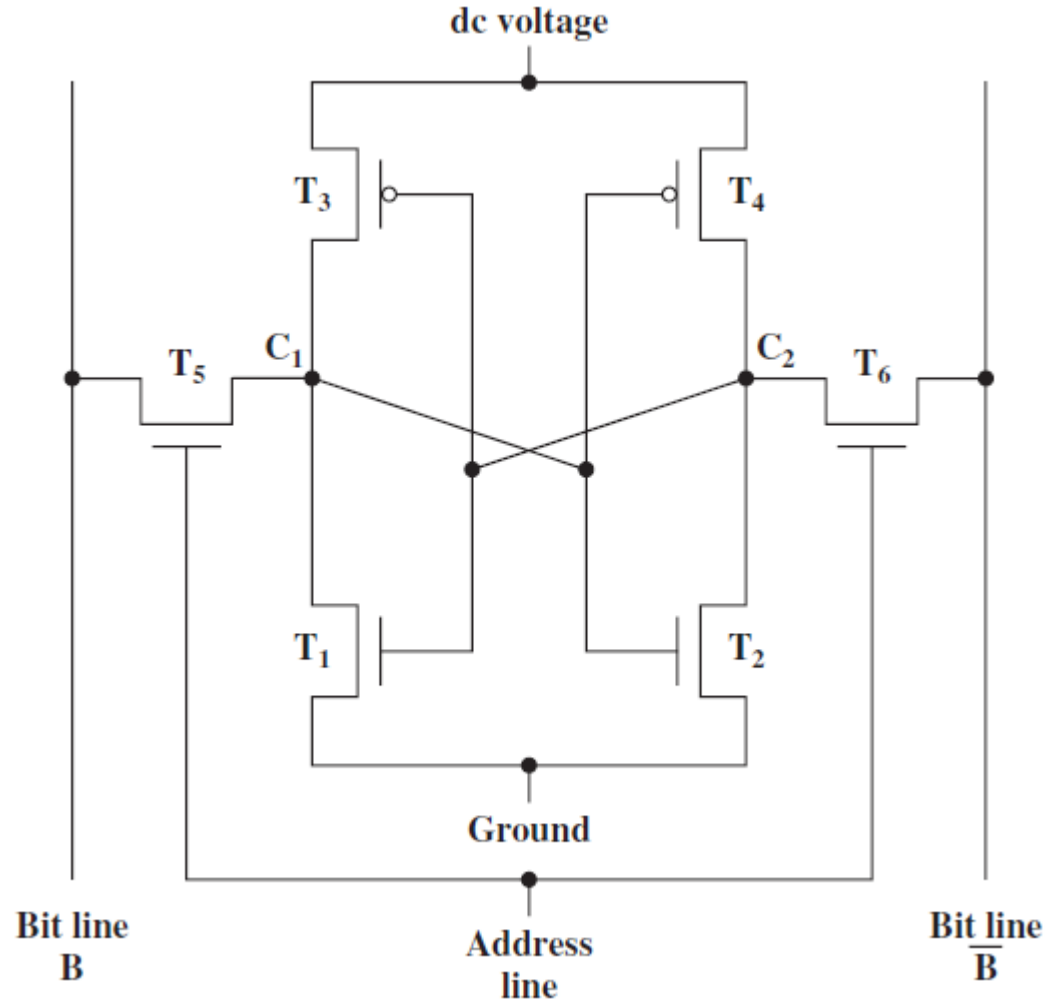
Dynamic RAM

- Bits (data) stored as charge in capacitors
  - Presence or absence of charge in a capacitor is interpreted as a binary 1 or 0
- Charges leak
- Need periodic refreshing to maintain data storage
- **Dynamic***:* tendency of the stored charge to leak away, even with power continuously applied
- Simpler construction
- Smaller per bit
- Less expensive
- Need refresh circuits
- Slower
- Main memory
- Essentially analogue: level of charge determines value

Static RAM

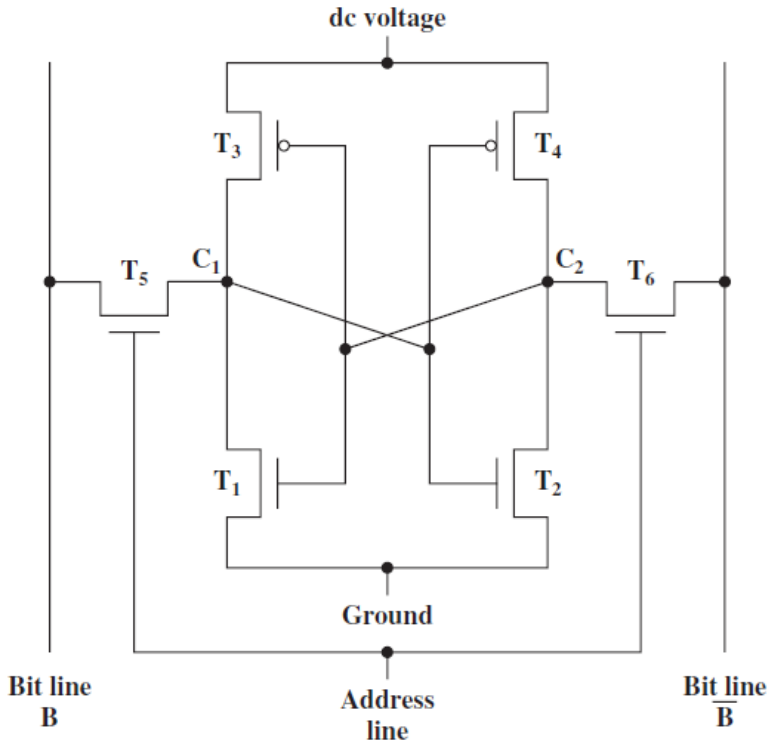
- Bits stored as on/off switches
- No charges to leak
- No refreshing needed when powered
- More complex construction
- Larger per bit
- More expensive
- Does not need refresh circuits
- Faster
- Cache

Static RAM Structure



Static RAM Operation

- 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 and complement to
- Read
  - Value is on line B



SRAM vs DRAM

- Both volatile
  - Power needed to preserve data

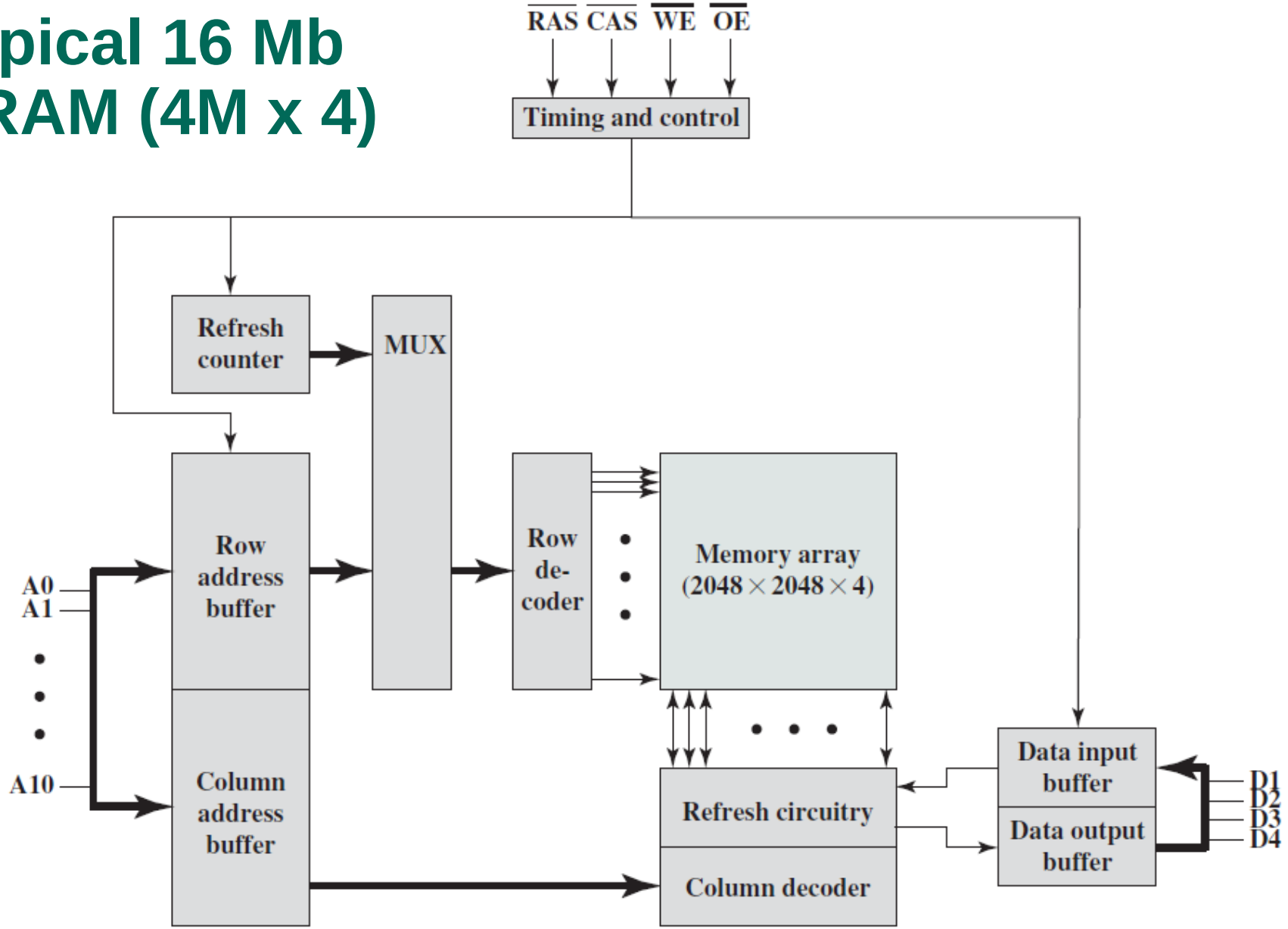
- Dynamic cell
  - Simpler to build, smaller
  - More dense (smaller cells $\rightarrow$ more cells per unit area)
  - Less expensive
  - Needs refresh
  - Tend to be favored for large memory requirements
  - Used for main memory

- Static cell
  - Faster
  - Used for cache memory (on and off chip)

Organization in detail

- A 16Mbit chip can be organized as 1M of 16 bit words
- A bit per chip system has 16 lots of 1 Mbit chip with bit 1 of each word in chip 1 and so on
- A 16 Mbit chip can be organized as a 2048 x 2048 x 4 bit array
  - Reduces number of address pins
    - Multiplex row and column address
    - 11 pins to address ($2^{11} = 2048$)
    - Adding one more pin doubles range of values so x4 capacity

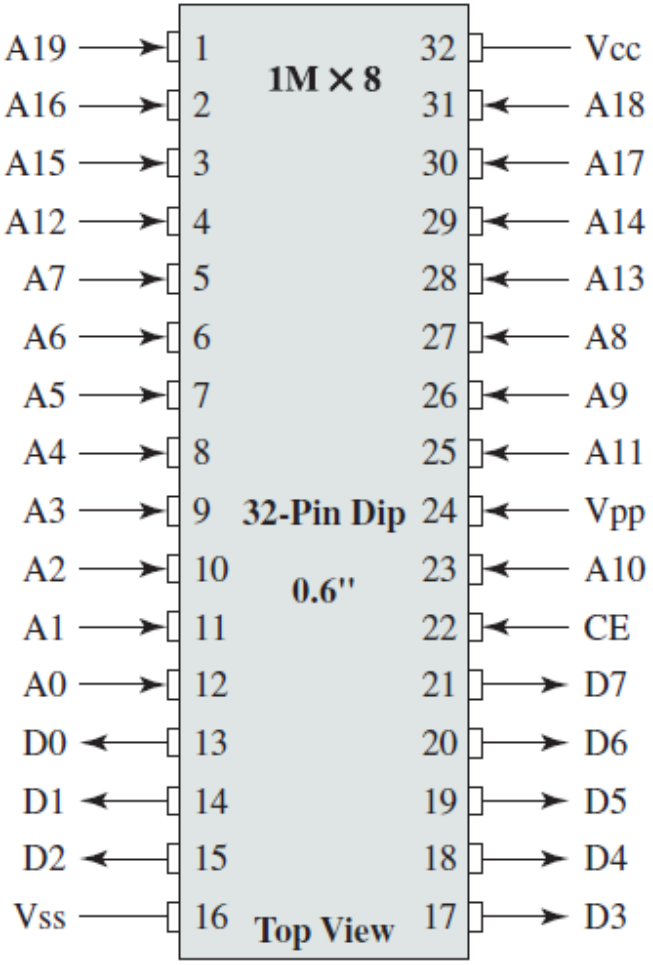
Typical 16 Mb DRAM (4M x 4)



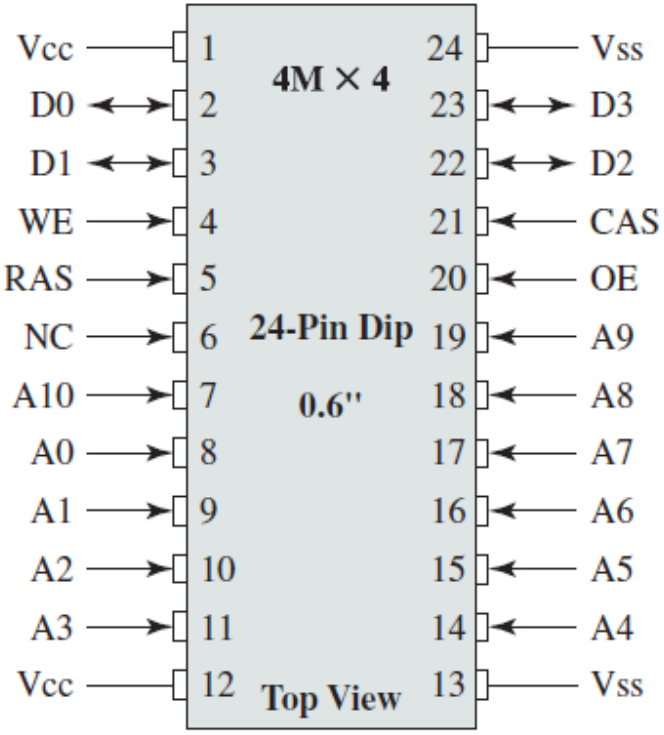
Refreshing

- Refresh circuit included on chip
- Disable chip
- Count through rows
- Read and Write back
- Takes time
- Slows down apparent performance

Packaging

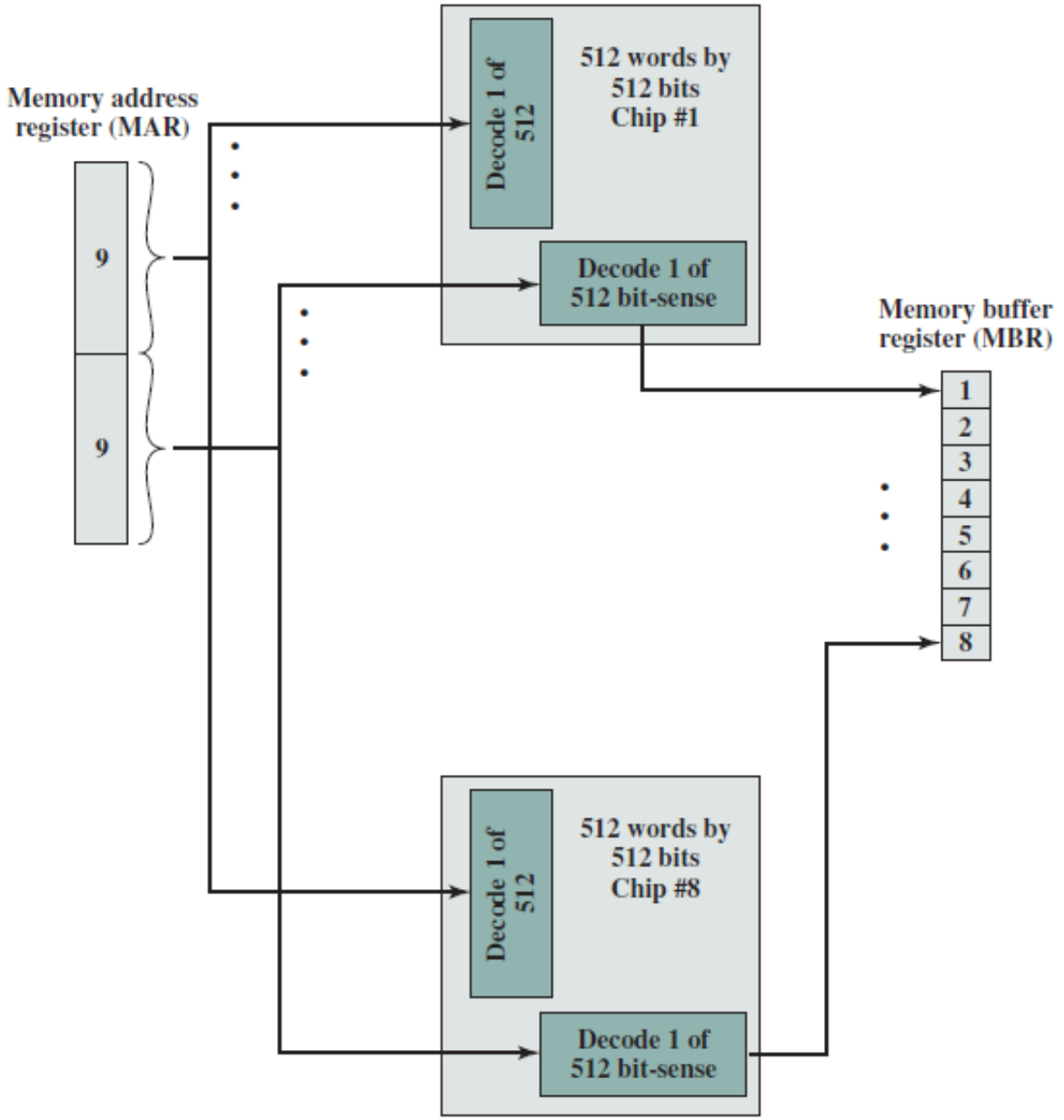


(a) 8-Mbit EPROM



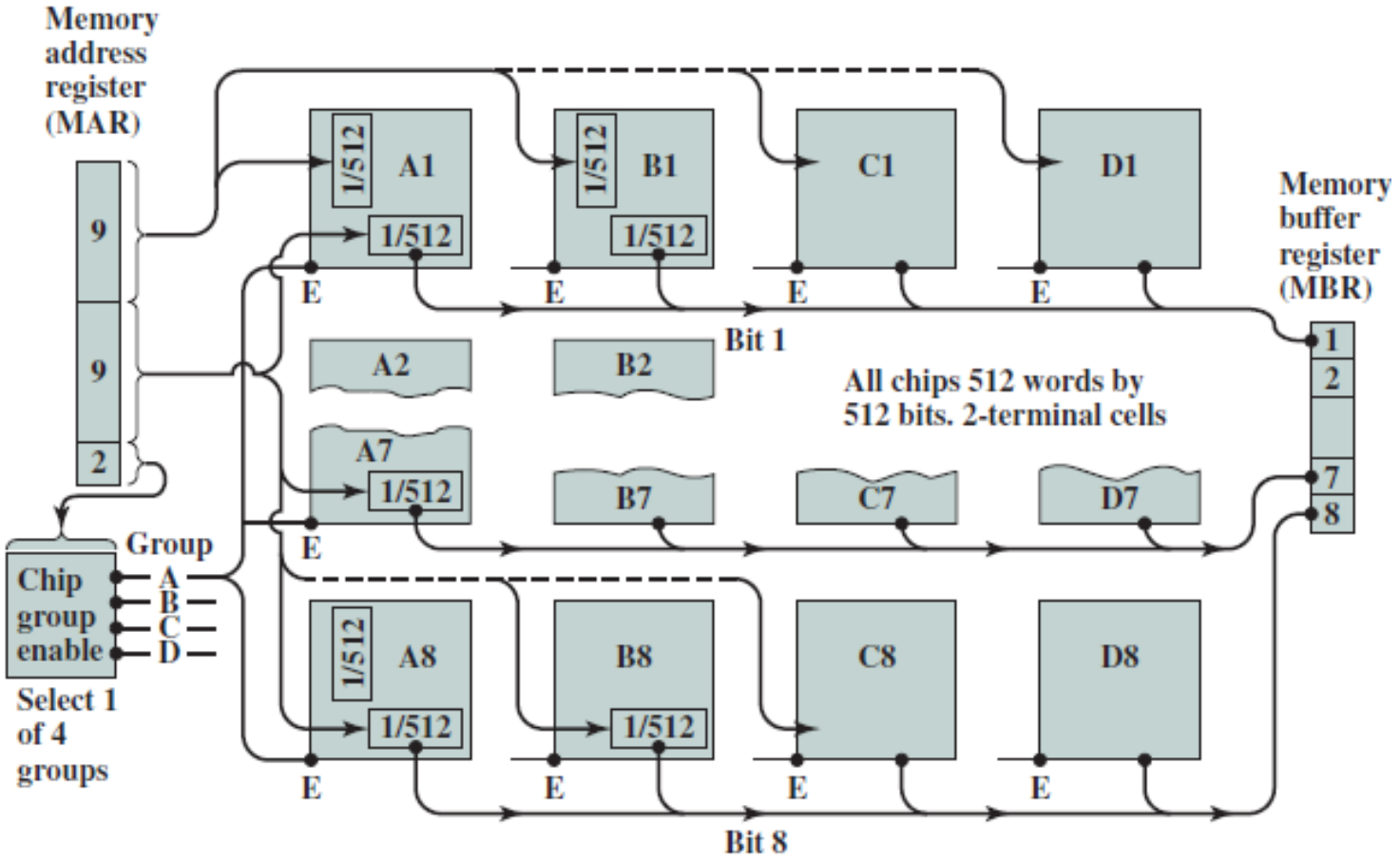
(b) 16-Mbit DRAM

Module Organization



256-KByte memory organization

Module Organization



1-MB memory organization

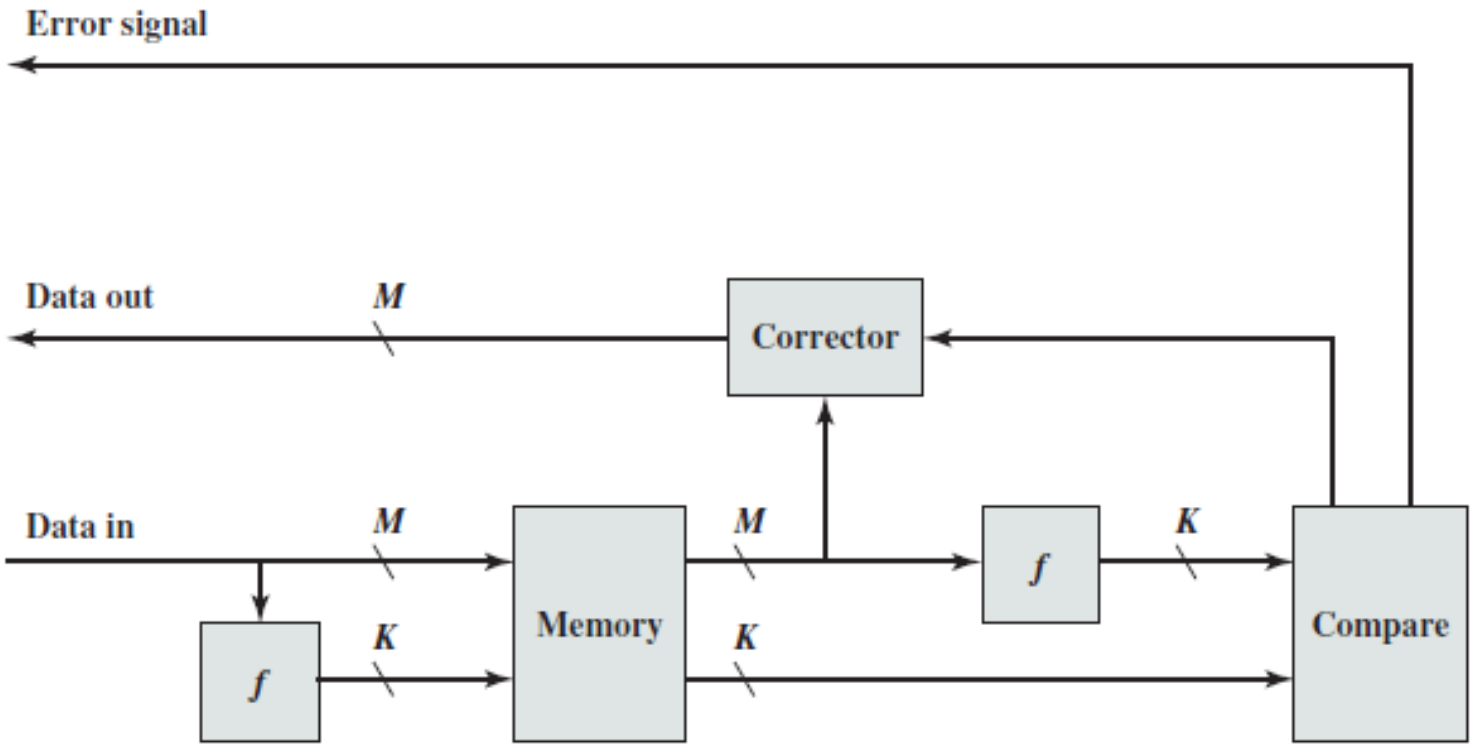
Interleaved Memory

- Main memory is composed of a collection of DRAM memory chips.
- A number of chips can be grouped together to form a **memory bank**.
- It is possible to organize the memory banks as interleaved memory.
  - Each bank is independently able to service a memory read or write request.
  - A system with *K* banks can service *K* requests simultaneously, increasing memory read or write rates by a factor of *K*.
- If consecutive words of memory are stored in different banks, then the transfer of a block of memory is speeded up.

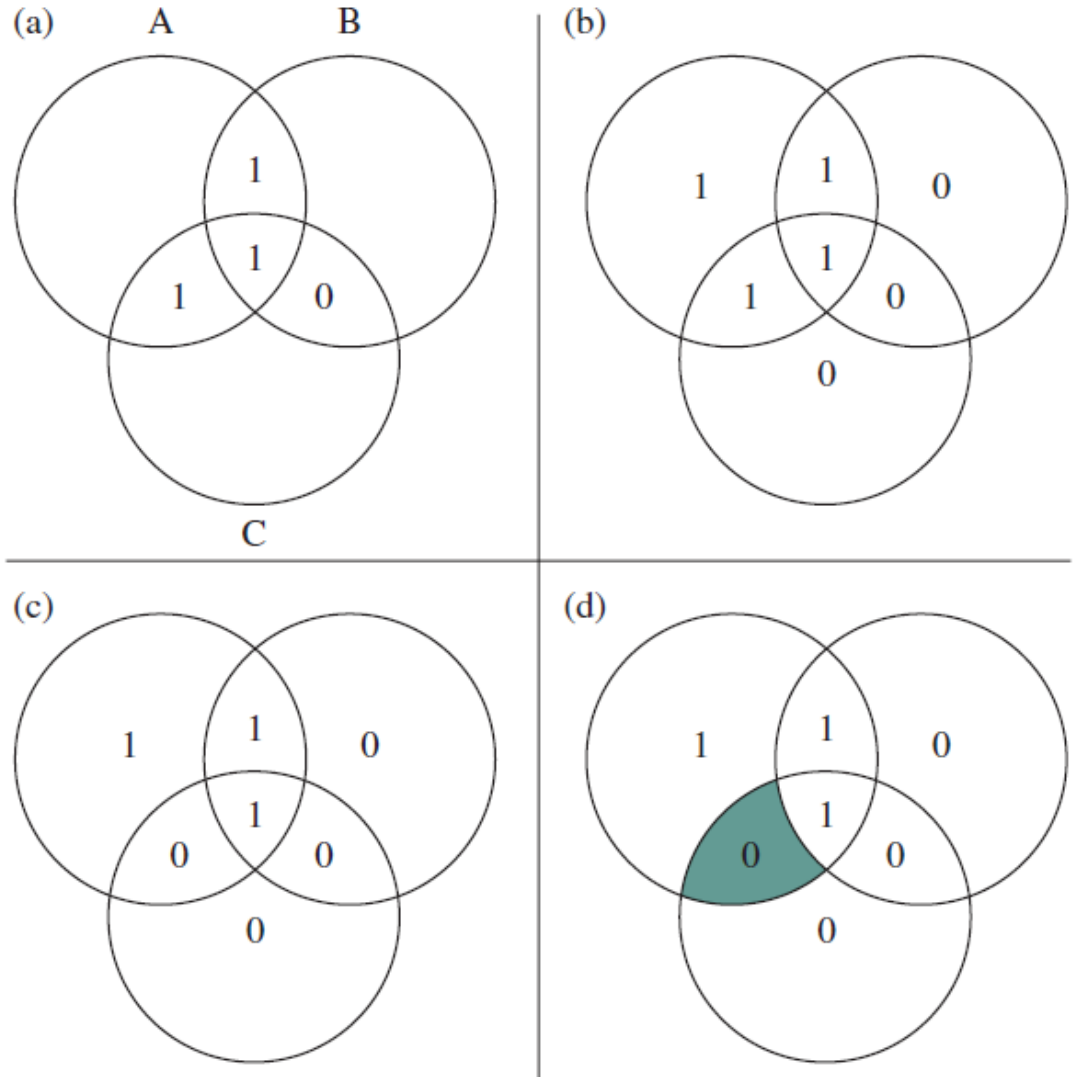
Error Correction

- Hard failure
  - Permanent defect
- Soft error
  - Random, non-destructive
  - No permanent damage to memory
- Detected using **Hamming error correcting code**

Error Correcting Code Function



Hamming Error Correcting Code



Increase in Word Length with Error Correction

	Single-Error Correction		Single-Error Correction/ Double-Error Detection	
Data Bits	Check Bits	% Increase	Check Bits	% Increase
8	4	50.0	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

Layout of Data Bits and Check Bits

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

$$C1 = D1 \oplus D2 \oplus D4 \oplus D5 \oplus D7$$

$$C2 = D1 \oplus D3 \oplus D4 \oplus D6 \oplus D7$$

$$C4 = D2 \oplus D3 \oplus D4 \oplus D8$$

$$C8 = \oplus D5 \oplus D6 \oplus D7 \oplus D8$$

Check Bit Calculation

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

	C8	C4	C2	C1
	0	1	1	1
⊕	0	0	0	1
	0	1	1	0

The result is 0110, indicating that bit position 6, which contains data bit 3, is in error.

Advanced DRAM Organization

- Interface to main internal memory is the most critical bottleneck when using high performance processors

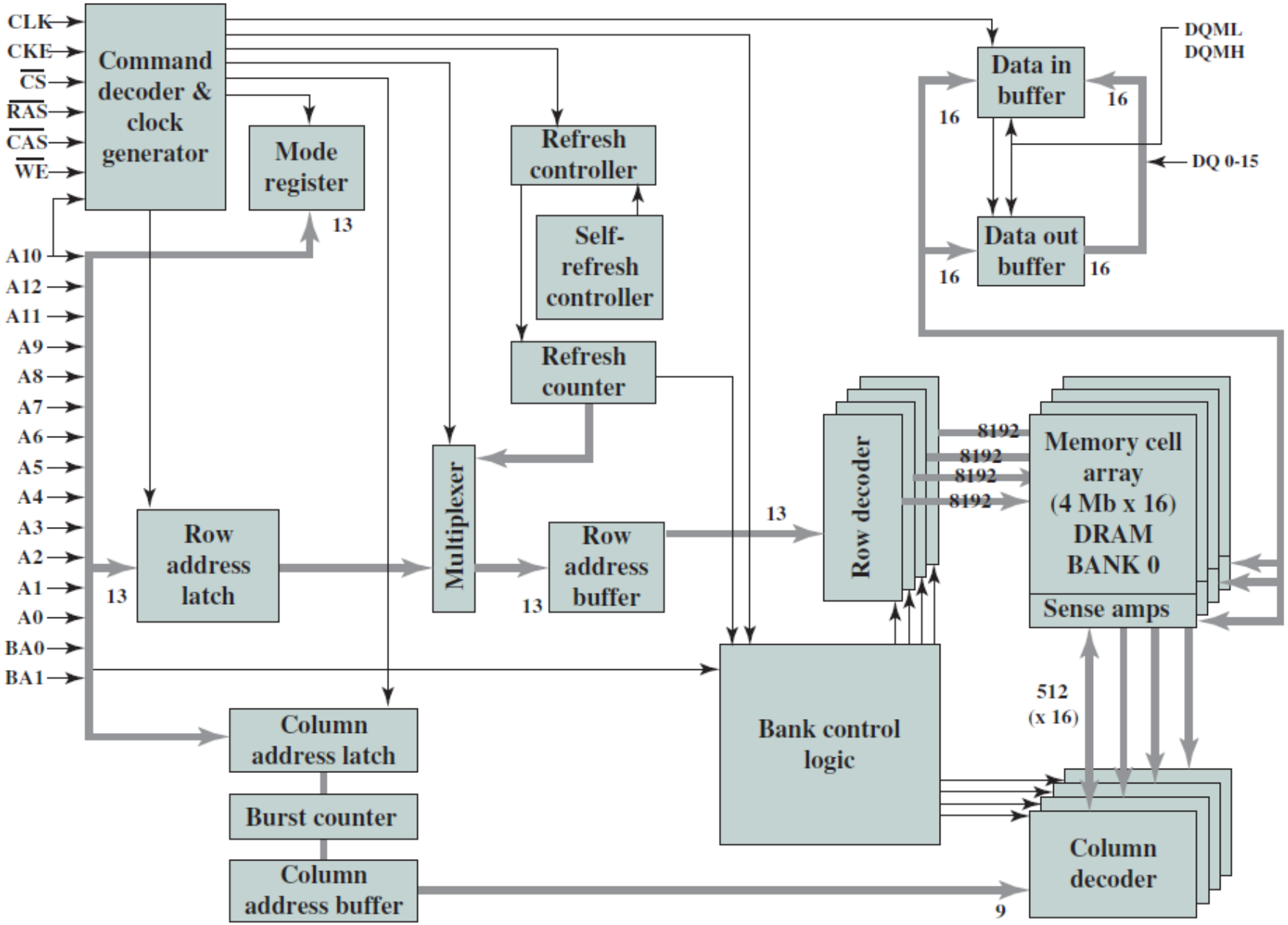
- Traditional DRAM constrained by
  - Internal architecture
  - Interface to processor's memory bus

- Enhanced DRAM
  - SDRAM
  - DDR-DRAM

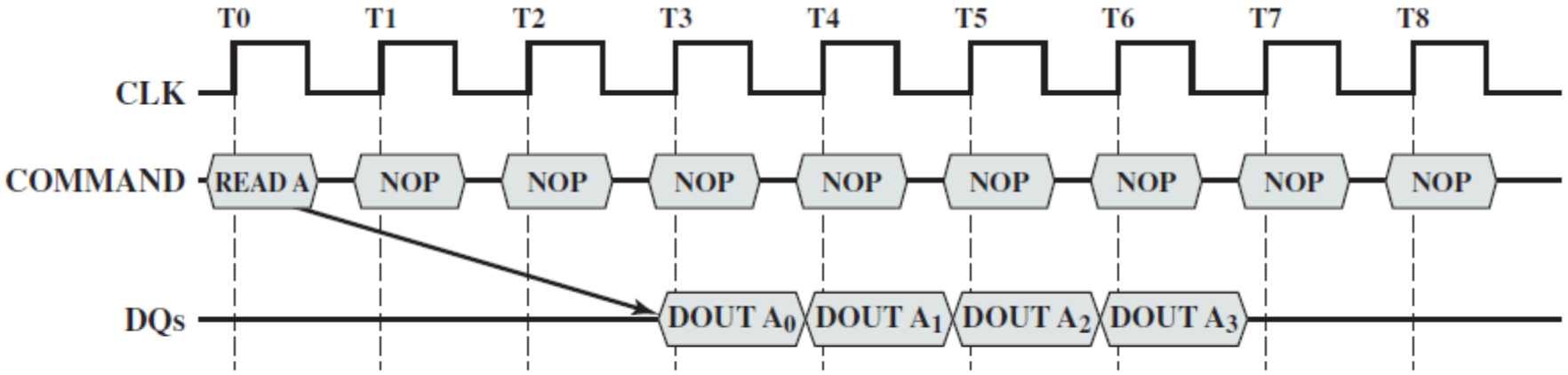
Synchronous DRAM (SDRAM)

- Access is synchronized with an external clock
- Address is presented to RAM
- RAM finds data (CPU waits in conventional DRAM)
- Since SDRAM moves data in time with system clock, CPU knows when data will be ready
- CPU does not have to wait, it can do something else
- Burst mode allows SDRAM to set up stream of data and fire it out in block

256 Mb SDRAM


CLK
CKE
CS
RAS
CAS
WE
Command decoder & clock generator
Mode register
13
A10
A12
A11
A9
A8
A7
A6
A5
A4
A3
A2
A1
A0
BA0
BA1
Refresh controller
Self-refresh controller
Refresh counter
Multiplexer
Row address latch
13
Row address buffer
13
13
Row decoder
8192
8192
8192
8192
Memory cell array (4 Mb x 16) DRAM BANK 0
Sense amps
512 (x 16)
Bank control logic
Column address latch
Burst counter
Column address buffer
9
Column decoder
Data in buffer
16
16
DQML
DQMH
DQ 0-15
Data out buffer
16
16

SDRAM Operation



SDRAM Read Timing (burst length = 4, $\overline{\text{CAS}}$ latency = 2)

Double Data Rate SDRAM – DDR SDRAM

- Sends data twice per clock cycle (leading and trailing edge)
- Achieves higher data rates in 3 ways:
  - Data transfer is synchronized to both the rising and falling edge of the clock, rather than just the rising edge
  - Uses higher clock rate on the bus to increase the transfer rate
  - Buffering scheme is used

	DDR1	DDR2	DDR3	DDR4
Prefetch buffer (bits)	2	4	8	8
Voltage level (V)	2.5	1.8	1.5	1.2
Front side bus data rates (Mbps)	200—400	400—1066	800—2133	2133—4266

Read Only Memory (ROM)

- Permanent storage
  - Nonvolatile
- Microprogramming
- Library subroutines
- Systems programs (BIOS)
- Function tables

Types of ROM

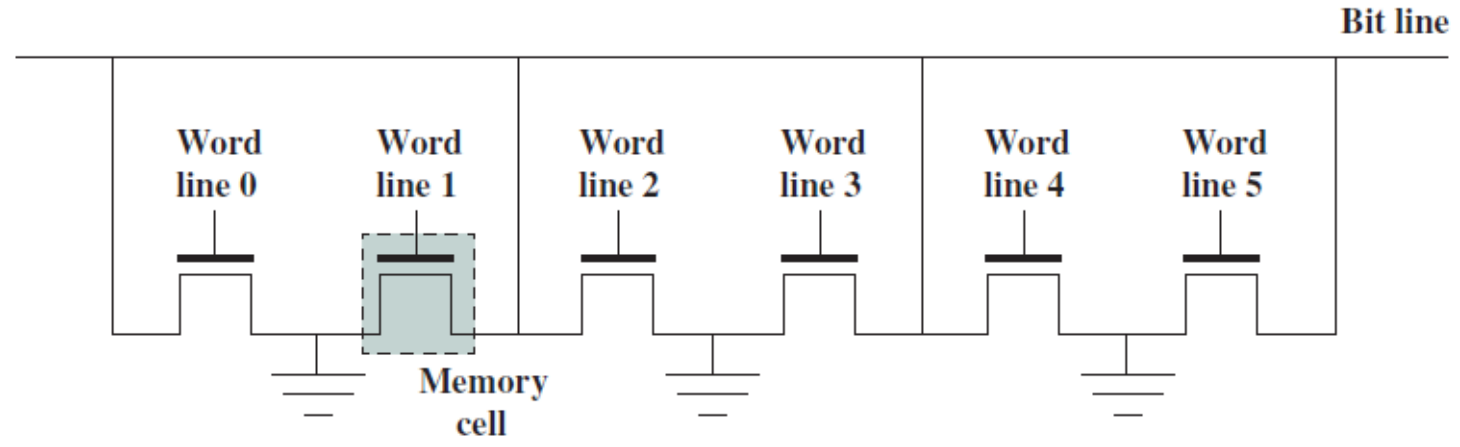
- Written during manufacture
  - Very expensive for small runs
- Programmable (once)
  - PROM
  - Needs special equipment to program
- Read “mostly”
  - Erasable Programmable (EPROM)
    - Erased by UV
  - Electrically Erasable (EEPROM)
    - Takes much longer to write than read
  - Flash memory
    - Erase whole memory electrically

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)			Electrically	
Erasable PROM (EPROM)	Read-mostly memory	UV light, chip-level		
Electrically Erasable PROM (EEPROM)		Electrically, byte-level		
Flash memory		Electrically, block-level		

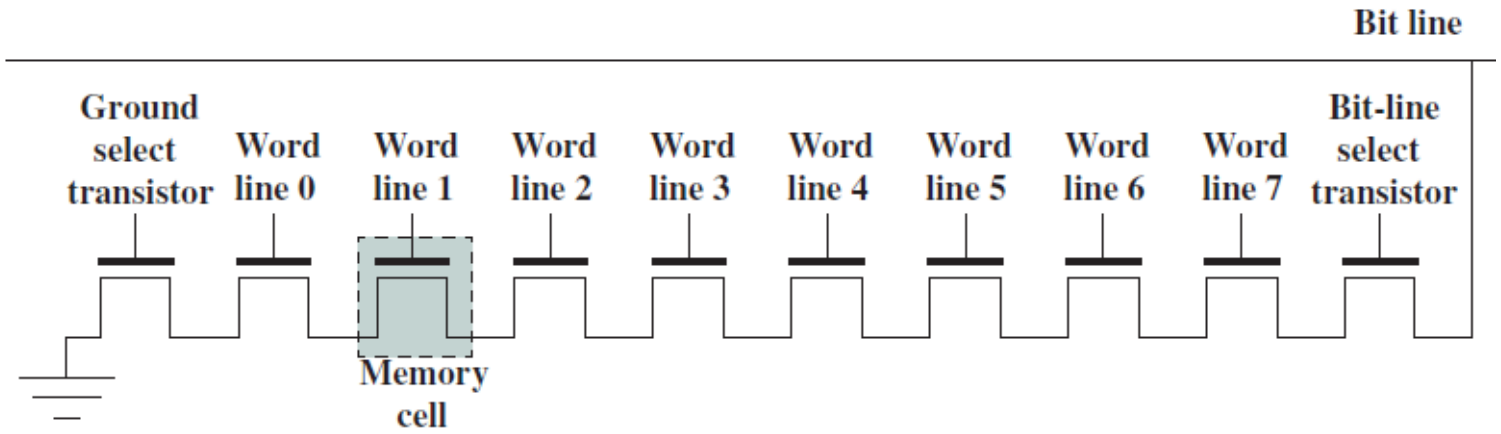
Flash Memory

- 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
- Uses only one transistor per bit so it achieves the high density of EPROM
- Persistent
  - Retains data when there is no power applied to the memory
  - Useful for secondary (external) storage, and as an alternative to random access memory

Flash Memory Structures

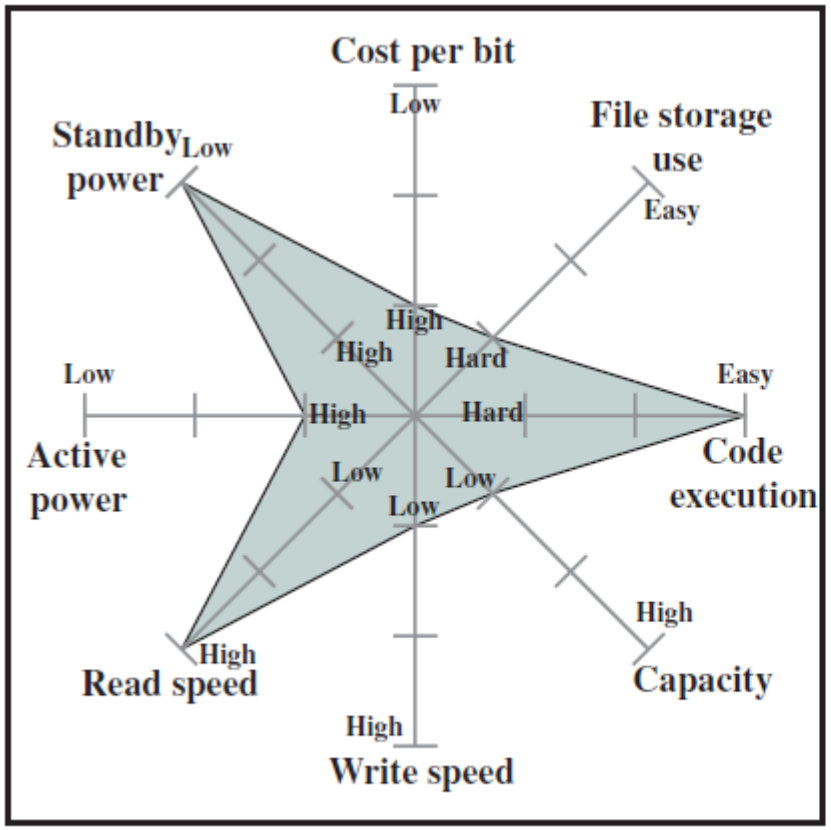


(a) NOR flash structure

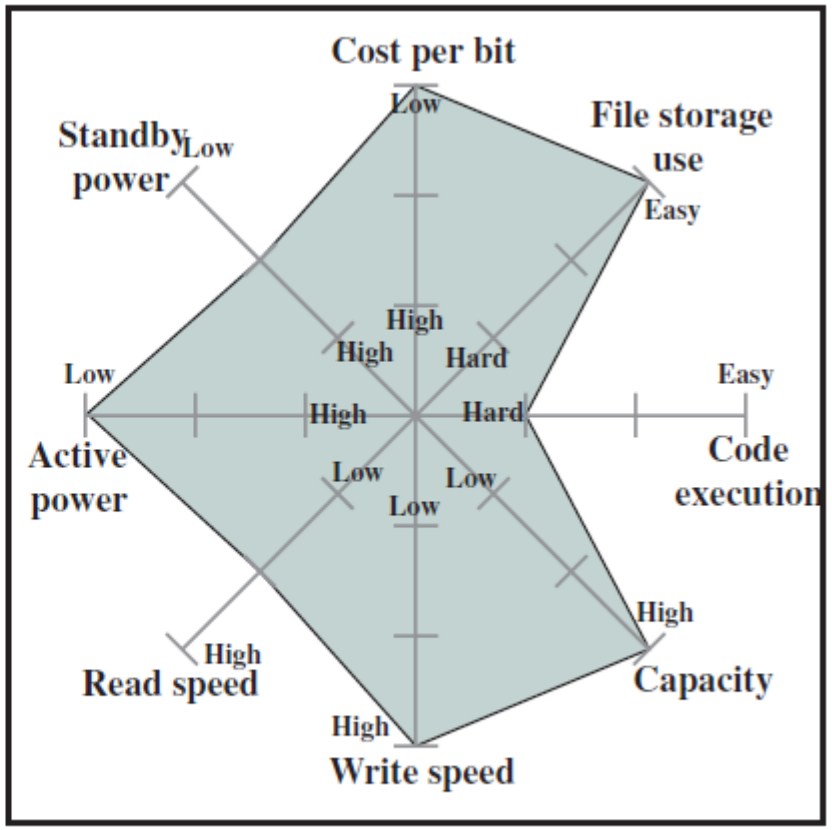


(b) NAND flash structure

Kiviat Graphs for Flash Memory

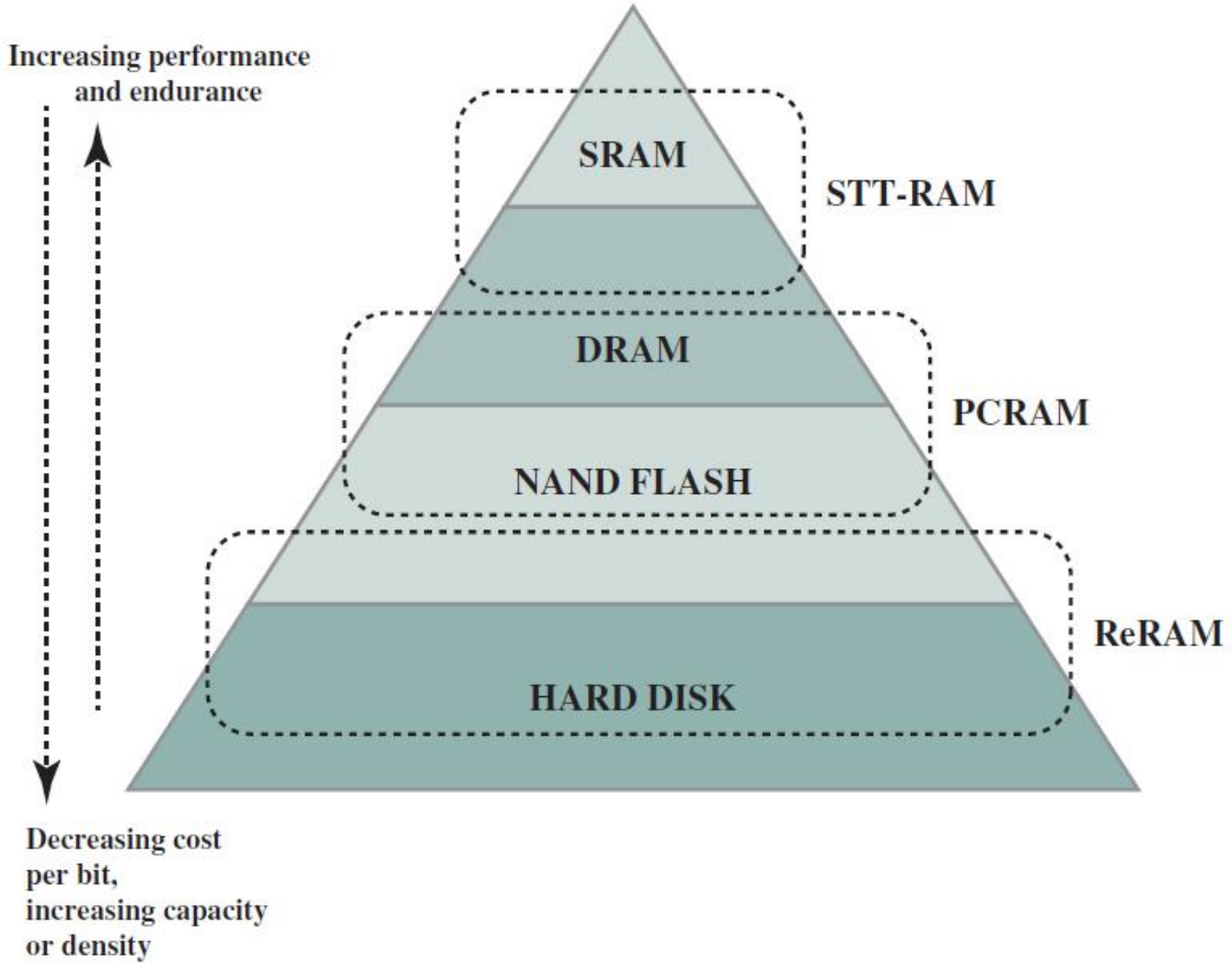


(a) NOR



(b) NAND

Nonvolatile RAM within the Memory Hierarchy



Nonvolatile RAM Technologies

- STT-RAM
  - new type of magnetic RAM (MRAM), which features non-volatility, fast writing/reading speed, and high programming endurance and zero standby power
  - storage capability or programmability of MRAM from magnetic tunneling junction (MTJ), in which a thin tunneling dielectric is sandwiched between two ferromagnetic layers
  - good candidate for either cache or main memory
- PCRAM
  - Phase-change RAM, the most mature of the new technologies
  - based on a chalcogenide alloy material, similar to those commonly used in optical storage media (compact discs and digital versatile discs)
  - good candidate to replace or supplement DRAM for main memory
- ReRAM (also known as RRAM)
  - works by creating resistance rather than directly storing charge. An electric current is applied to a material, changing the resistance of that material.
  - finding appropriate materials and measuring the resistance state of the cells
  - low voltage, endurance is far superior to flash memory, and the cells are much smaller—at least in theory
  - good candidate to replace or supplement both secondary storage and main memory