Overview

- Memory definitions
- Random Access Memory (RAM)
- Static RAM (SRAM) integrated circuits
  - Cells and slices
  - Cell arrays and coincident selection
- Arrays of SRAM integrated circuits
- Dynamic RAM (DRAM) integrated circuits
- DRAM Types

Memory Definitions

- Memory — A collection of storage cells together with the necessary circuits to transfer information to and from them.
- Memory Organization — the basic architectural structure of a memory in terms of how data is accessed.
- Random Access Memory (RAM) — a memory organized such that data can be transferred to or from any cell (or collection of cells) in a time that is not dependent upon the particular cell selected.
- Memory Address — A vector of bits that identifies a particular memory element (or collection of elements).

Memory Definitions (Continued)

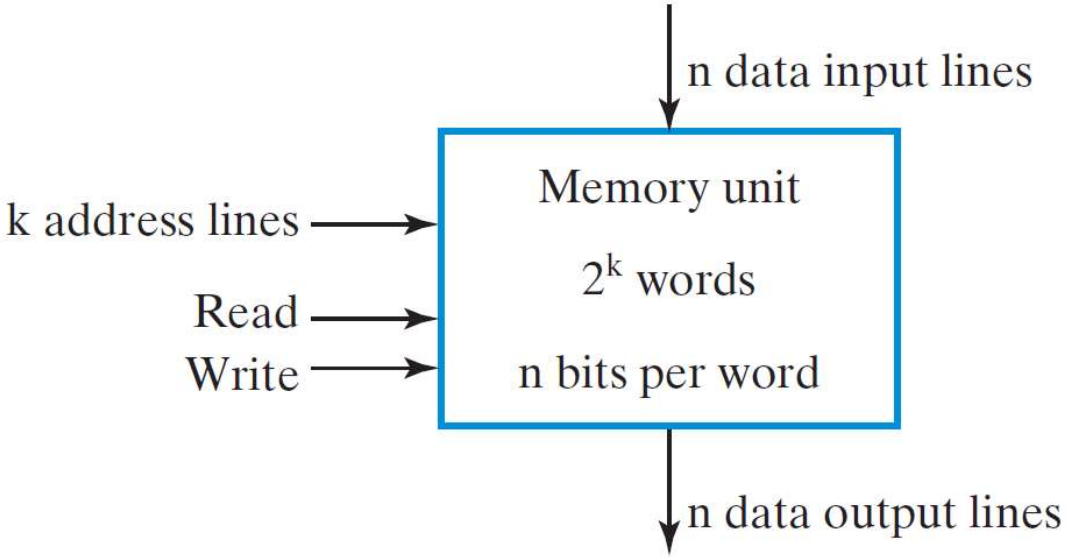
- Typical data elements are:
  - bit — a single binary digit
  - byte — a collection of eight bits accessed together
  - word — a collection of binary bits whose size is a typical unit of access for the memory. It is typically a power of two multiple of bytes (e.g., 1 byte, 2 bytes, 4 bytes, 8 bytes, etc.)
- Memory Data — a bit or a collection of bits to be stored into or accessed from memory cells.
- Memory Operations — operations on memory data supported by the memory unit. Typically, *read* and *write* operations over some data element (bit, byte, word, etc.).

Memory Organization

- Organized as an indexed array of words. Value of the index for each word is the memory address.
- Often organized to fit the needs of a particular computer architecture. Some historically significant computer architectures and their associated memory organization:
  - Digital Equipment Corporation PDP-8 – used a 12-bit address to address 4096 12-bit words.
  - IBM 360 – used a 24-bit address to address 16,777,216 8-bit bytes, or 4,194,304 32-bit words.
  - Intel 8080 – (8-bit predecessor to the 8086 and the current Intel processors) used a 16-bit address to address 65,536 8-bit bytes.

Memory Block Diagram

- A basic memory system is shown here:
- k address lines are decoded to address 2^k words of memory.
- Each word is n bits.
- Read and Write are single control lines defining the simplest of memory operations.



Memory Organization Example

- Example memory contents:
  - A memory with 3 address bits & 8 data bits has:
  - k = 3 and n = 8 so 2^3 = 8 addresses labeled 0 to 7.
  - 2^3 = 8 words of 8-bit data

Memory address		
Binary	Decimal	Memory contents
0000000000	0	10110101 01011100
0000000001	1	10101011 10001001
0000000010	2	00001101 01000110
	.	.
	.	.
	.	.
	.	.
	.	.
1111111101	1021	10011101 00010101
1111111110	1022	00001101 00011110
1111111111	1023	11011110 00100100

Basic Memory Operations

- Memory operations require the following:
  - *Data* — data written to, or read from, memory as required by the operation.
  - *Address* — specifies the memory location to operate on. The address lines carry this information into the memory. Typically: n bits specify locations of 2^n words.
  - An operation — Information sent to the memory and interpreted as control information which specifies the type of operation to be performed. Typical operations are READ and WRITE. Others are READ followed by WRITE and a variety of operations associated with delivering blocks of data. Operation signals may also specify timing info.

Control Inputs to a Memory Chip

- Read Memory — an operation that reads a data value stored in memory:
  - Place a valid address on the address lines.
  - Wait for the read data to become stable.
- Write Memory — an operation that writes a data value to memory:
  - Place a valid address on the address lines and valid data on the data lines.
  - Toggle the memory write control line
- Chip Select — to enable the particular RAM chip or chips containing the word to select.

Control Inputs to a Memory Chip

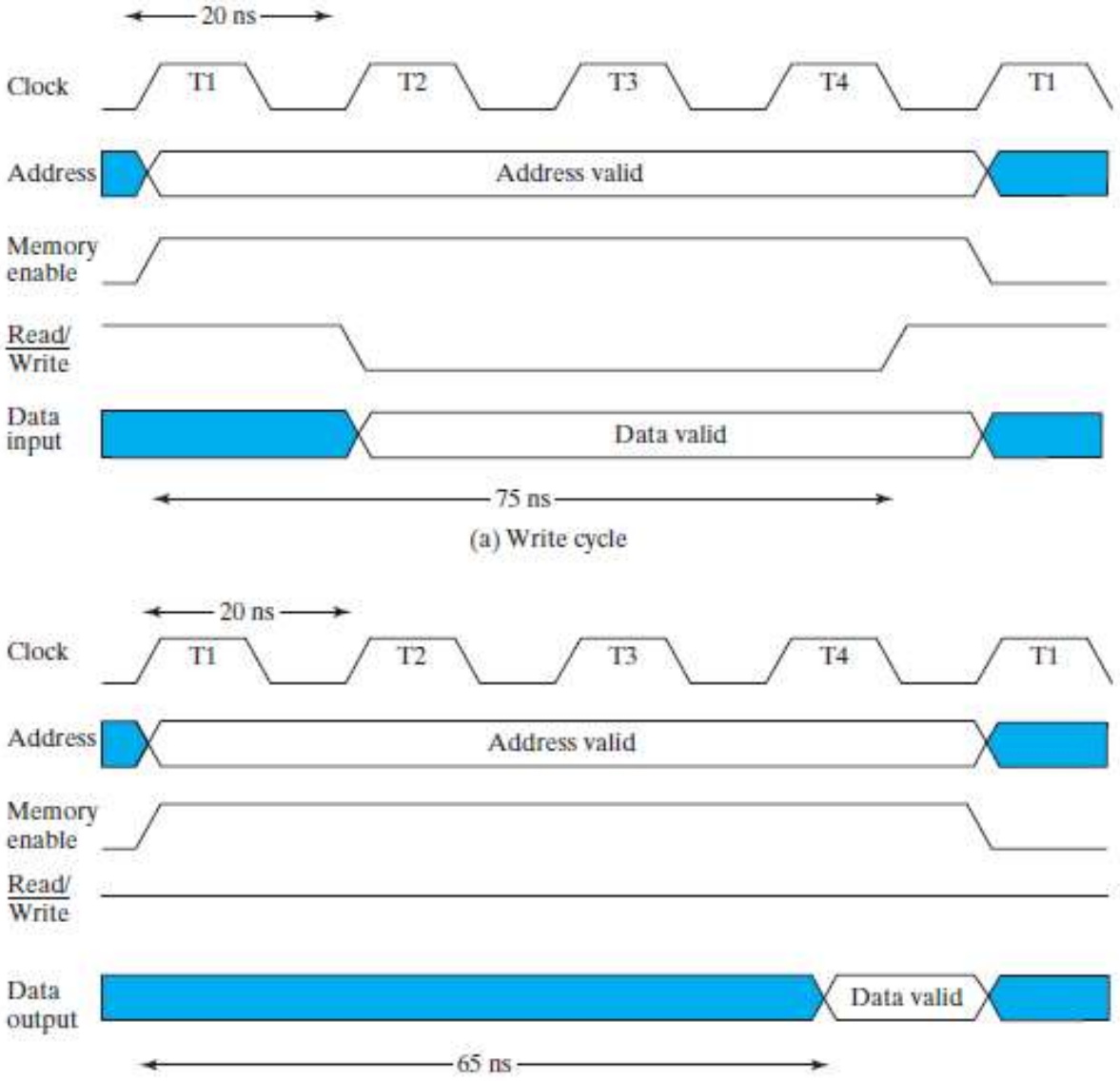
Chip select CS	Read/$\overline{\text{Write}}$ R/$\overline{\text{W}}$	Memory operation
0	X	None
1	0	Write to selected word
1	1	Read from selected word

Memory Operation

- The operation of memory unit is controlled bt an external device, such as a CPU.
- The **access time** of a memory read operation is the maximum time from the application fo the adress to the appeareance of the data at the Data Output.
- The **write cycle time** is the maximum time from the application of the address to the completion of all internal memory operationss at the intervals of the cycle time.
- As an example: a CPU operates 50 MHz, giving a period of 20 ns for one clock pulse. The CPU communicates with a memory with an access time of 65 ns and write cycle time of 75 ns.

Memory Operation

Memory Cycle Timing Waveforms. Four pulses T1, T2,T3, and T4 with a cycle of 20 ns.

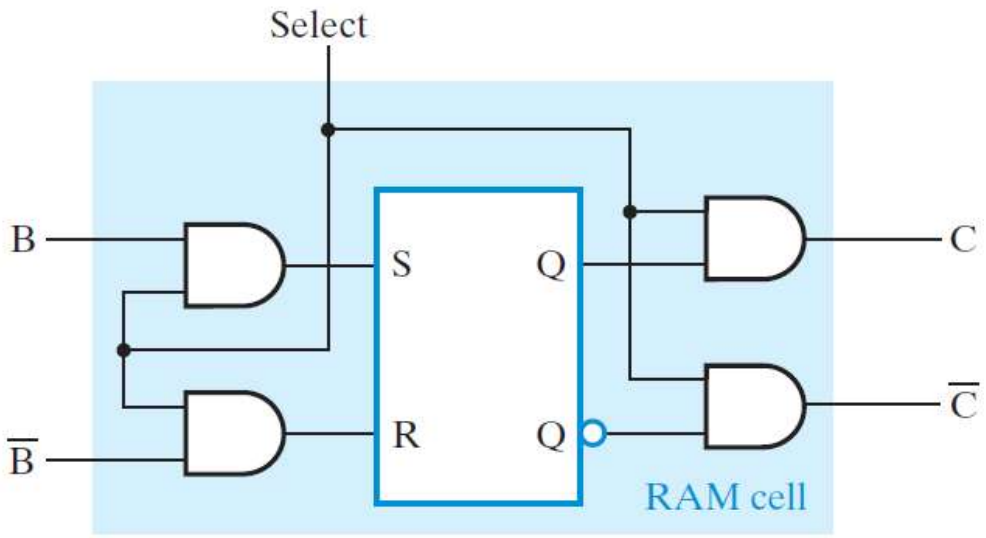


RAM Integrated Circuits

- Types of random access memory
  - *Static* – information stored in latches
  - *Dynamic* – information stored as electrical charges on capacitors
    - Charge "leaks" off
    - Periodic *refresh* of charge required
- Dependence on Power Supply
  - *Volatile* – loses stored information when power turned off
  - *Non-volatile* – retains information when power turned off

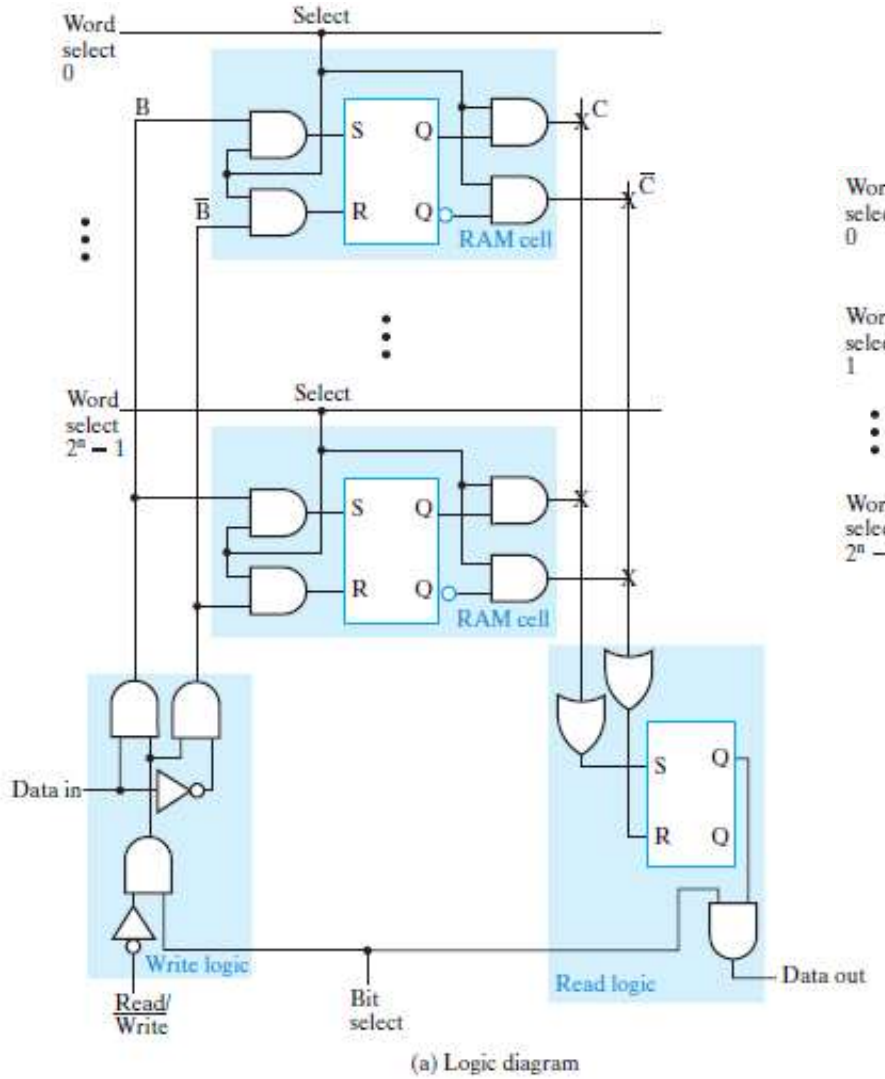
Static RAM Cell

- The internal structure of a RAM chip of ***m*** words with ***n*** bits per word cobsists og an array of ***mn*** binary storage cells and associated circuitry.
- The **RAM cell** is the basic binary storage elenmets used in the RAM chip.
- Storage Cell
  - SR Latch
  - Select input for control
  - Dual Rail Data Inputs B and B'
  - Dual Rail Data Outputs C and C'



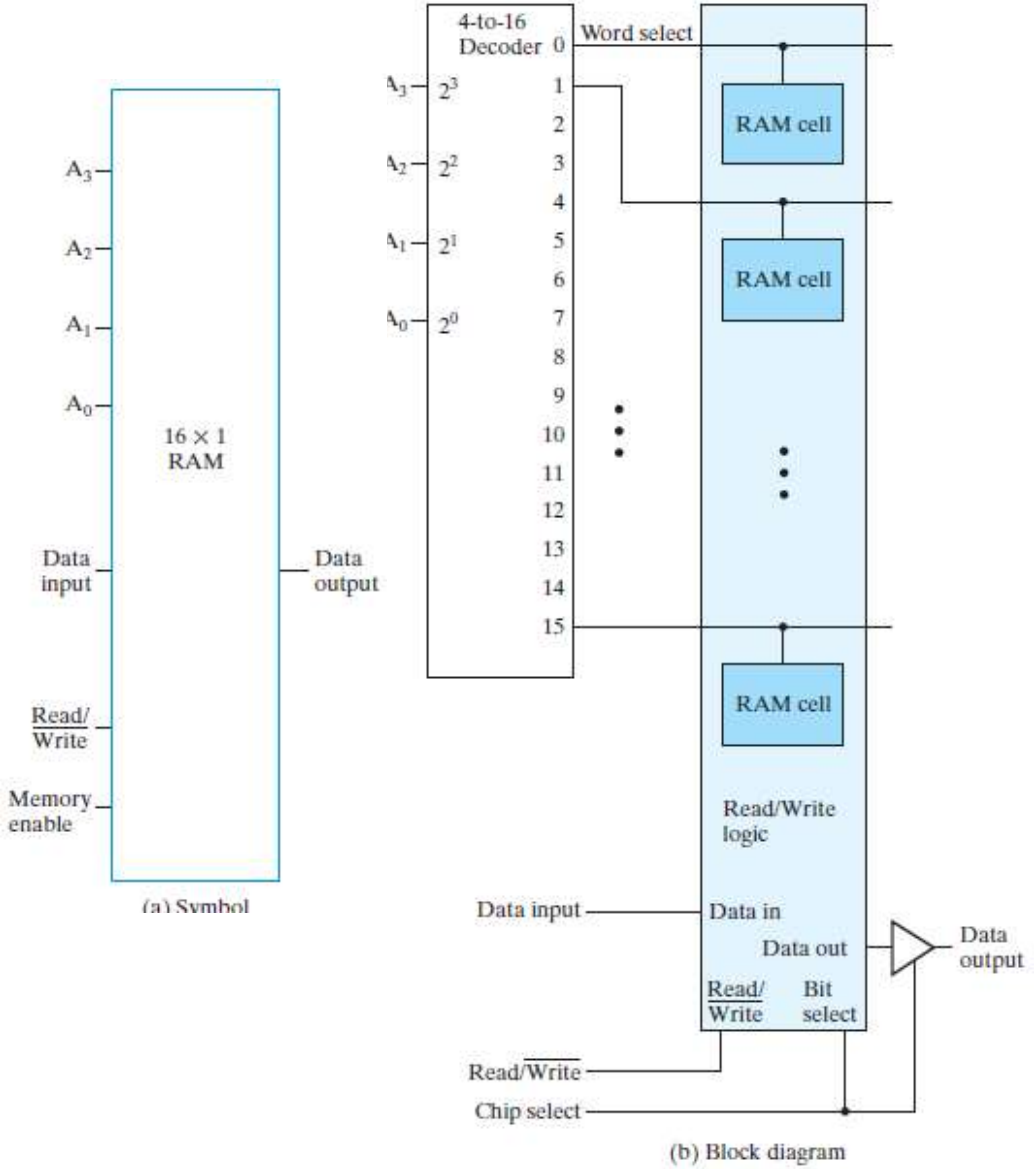
Static RAM Bit Slice

- Represents all circuitry that is required for 2^n 1-bit words
  - *Multiple RAM cells*
  - *Control Lines:*
    - Word select i – one for each word
    - Read/$\overline{\text{Write}}$.
    - Bit Select
  - *Data Lines:*
    - Data in
    - Data out



2^n-Word × 1-Bit RAM IC

- To build a RAM IC from a RAM slice, we need:
  - **Decoder** : decodes the n address lines to 2^n word select lines
  - A **3-state buffer** on the permits RAM ICs to be combined into a RAM with c × 2^n words

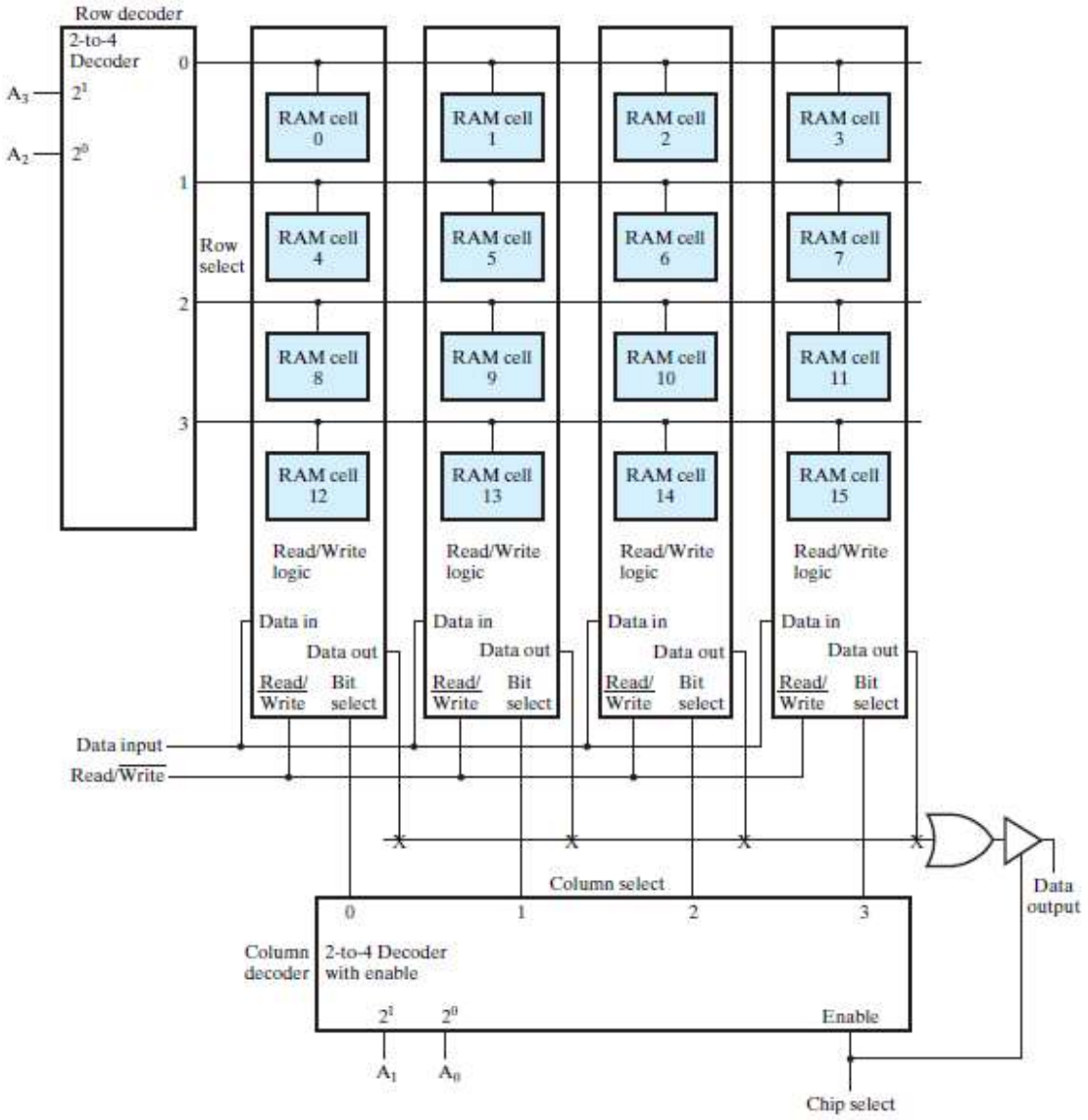


Cell Arrays and Coincident Selection

- Memory arrays can be very large =>
  - Large decoders
  - Large fanouts for the bit lines
  - The decoder size and fanouts can be reduced by approximately $\sqrt{n}$ by using a coincident selection in a 2-dimensional array
  - Uses two decoders, one for words and one for bits
  - Word select becomes Row select
  - Bit select becomes Column select
- See next slide for example
  - A_3 and A_2 used for Row select
  - A_1 and A_0 for Column select

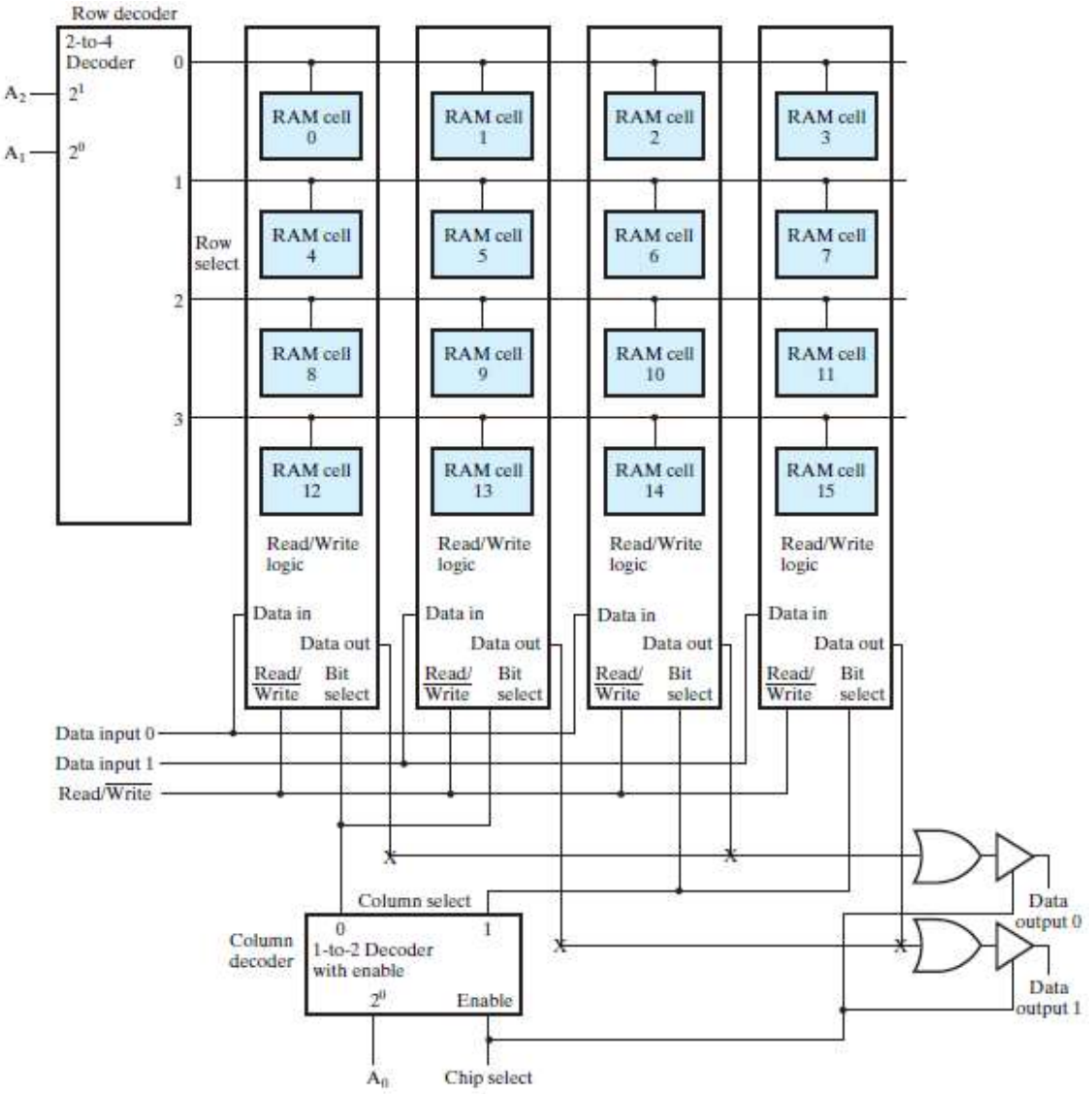
Cell Arrays and Coincident Selection (continued)

- Diagram of 16x1 RAM Using a 4 x 4 RAM Cell Array
  - For example, for the address 1001, the first two address bits are decoded to select row 10 , the sonce two address bits are decoded to selcet column 01 of the array.



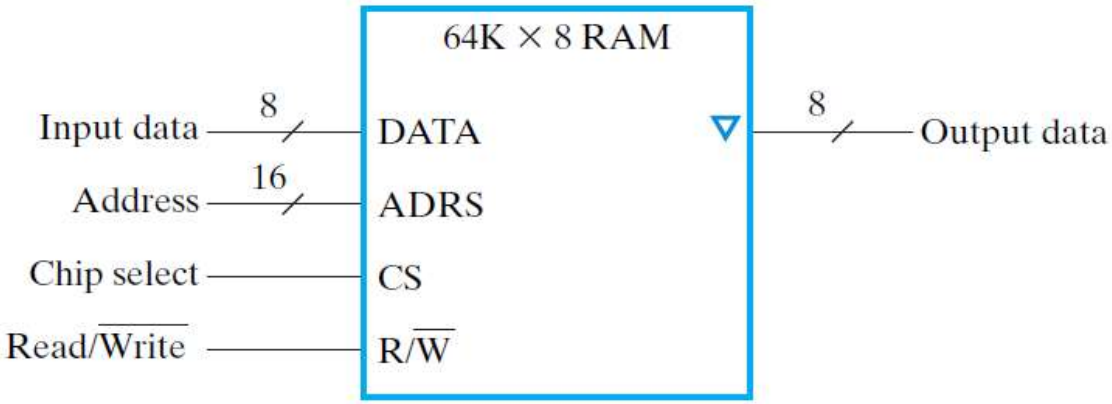
Cell Arrays and Coincident Selection (continued)

- Diagram of 8x2 RAM Using a 4 x 4 RAM Cell Array
  - For example, for the address 3 (011), the first two address bits are decoded to select row 1, the final bit,1, access column 1, which consists of bit slices 2 and 3.



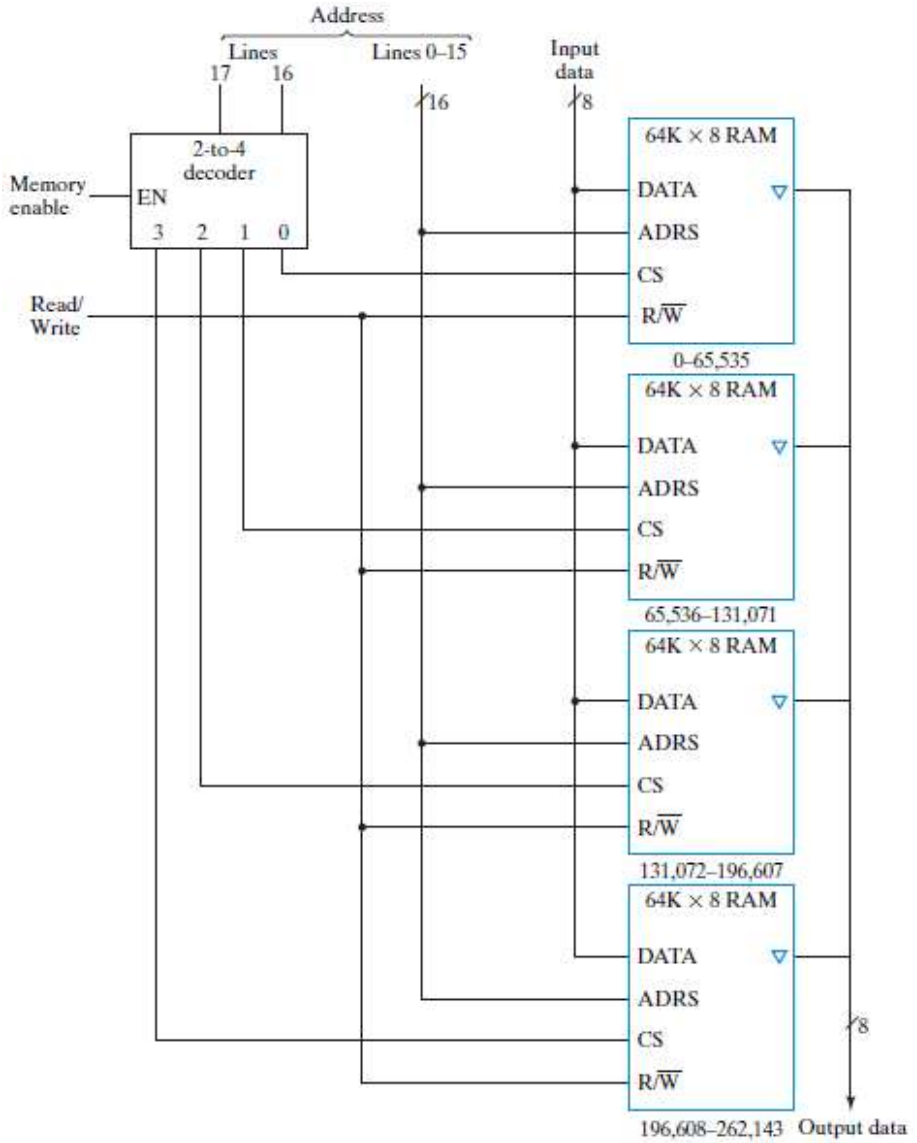
Array of SRAM ICs

- If the memory unit need for an application is larger than the capacity of one chip, it is necessary to combine a number of chips in an array to form the required size of a memory.
  - An increase in the number of words requires that we increase the address length.
  - An increase in the number of bits per word requires that we increase the number of data input and output lines.
- Using the CS lines, we can make larger memories from smaller ones by trying all address, data, and R/W lines in parallel, and using the decoded higher order address bits to control CS.



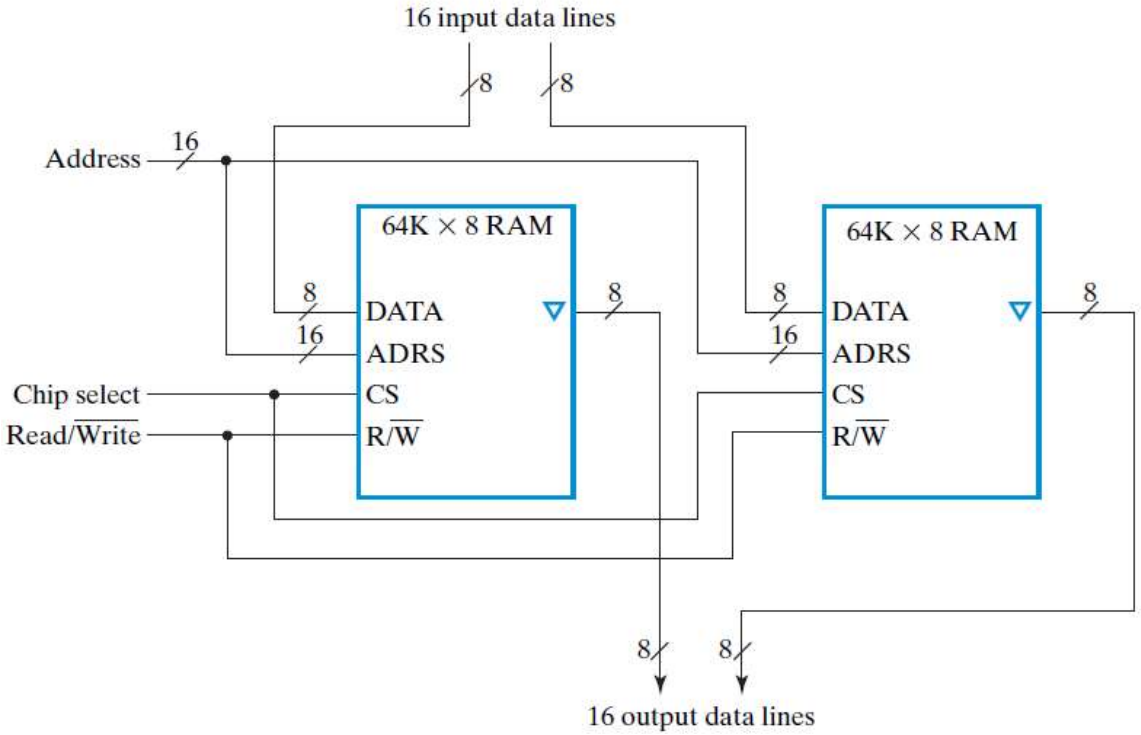
Making Larger Memories

- Constructing a 256K x 8 RAM with four 64K x 8 RAM chips.
- 256K-word memory requires an 18-bit address. The least 16 least significant bits are applied to the address of all four chips. The tow most significant bits are applied to a 2x4 decoder.



Making Wider Memories

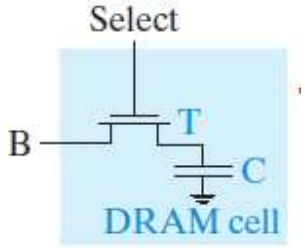
- To construct wider memories from narrow ones, we tie the address and control lines in parallel and keep the data lines separate.
- For example, two 64Kx8 chips can form a 64Kx16 memory.

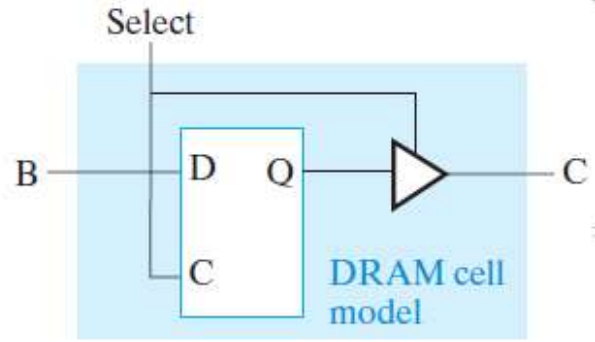


Dynamic RAM (DRAM)

- Because of its ability to provide high storage capacity at low cost, DRAM dominates the high capacity memory applications.
- Basic Principle: Storage of information on capacitors.
- DRAM Cell consits of a capacitor C and a transistor T.
  - The capacitor is used to store electrical charge.

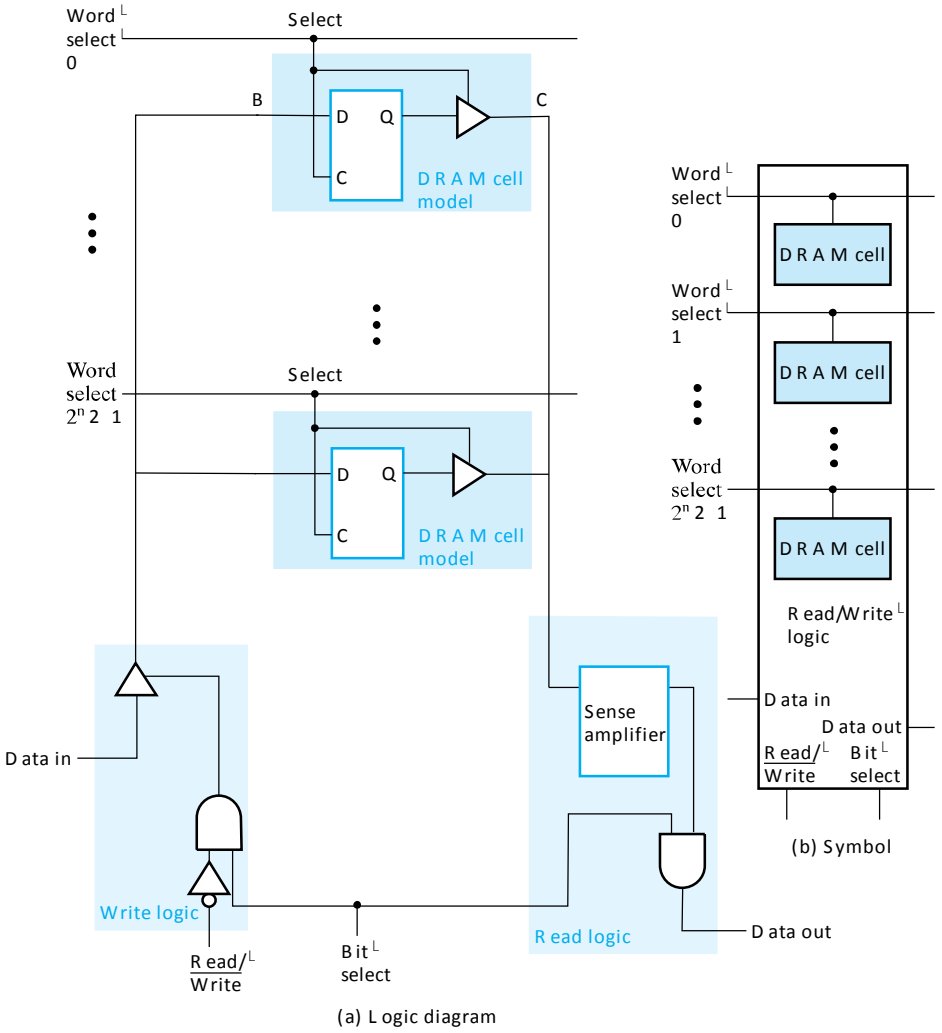
    Sufficient charge → Logic 1, Insufficint charge → Logic 0
  - Use of transistor as "switch" to:
    - Store charge
    - Charge or discharge



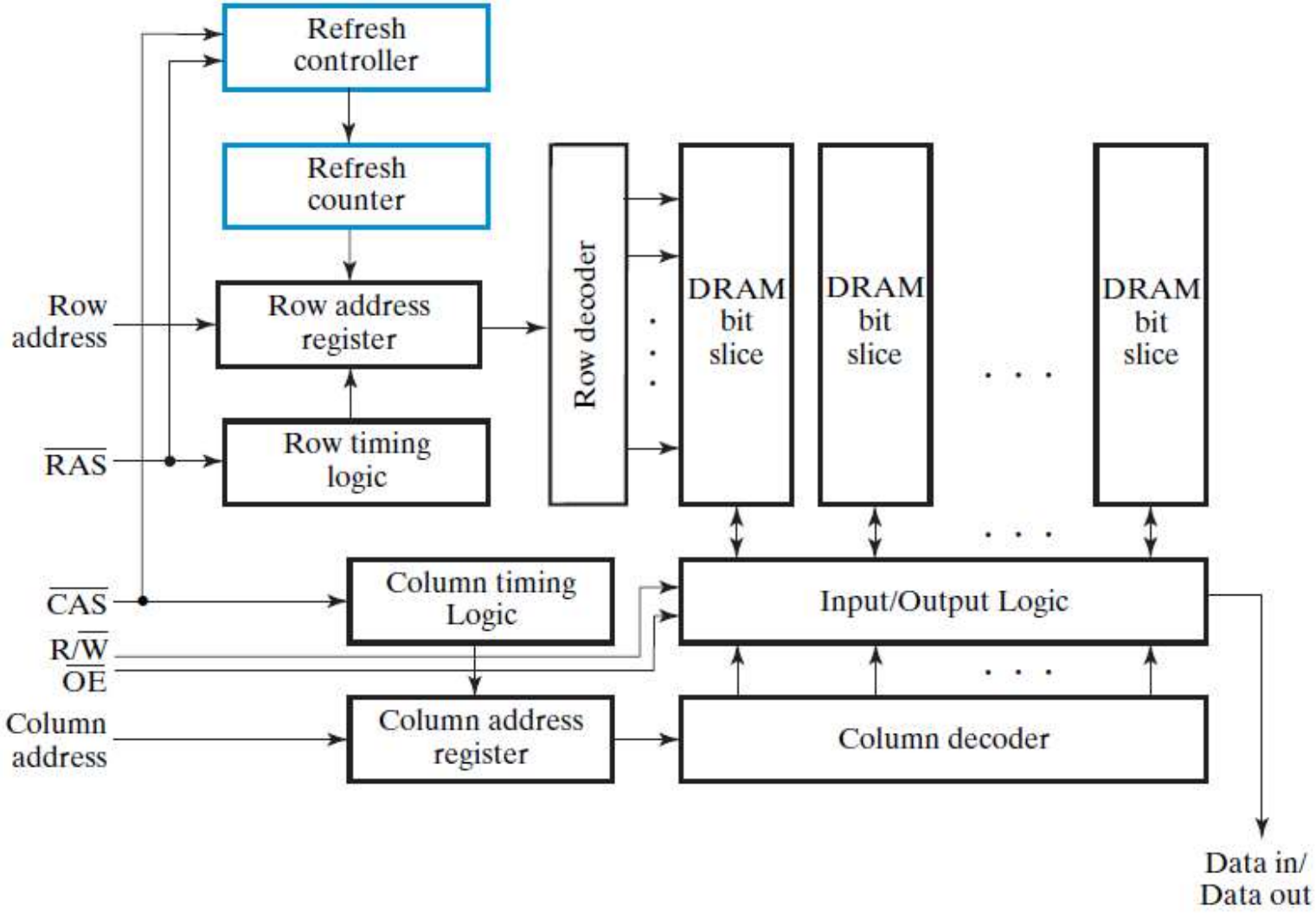


Dynamic RAM - Bit Slice

- C is driven by 3-state drivers
- Sense amplifier is used to change the small voltage change on C into H or L
- In the electronics, B, C, and the sense amplifier output are connected to make destructive read into non-destructive read



Dynamic RAM



Block Diagram of a DRAM Including Refresh Logic

DRAM Types

Type	Abbreviation	Description
Fast Page Mode DRAM	FPM DRAM	Takes advantage of the fact that, when a row is accessed, all of the row values are available to be read out. By changing the column address, data from different addresses can be read out without reapplying the row address and waiting for the delay associated with reading out the row cells to pass if the row portion of the addresses match.
Extended Data Output DRAM	EDO DRAM	Extends the length of time that the DRAM holds the data values on its output, permitting the CPU to perform other tasks during the access since it knows the data will still be available.
Synchronous DRAM	SDRAM	Operates with a clock rather than being asynchronous. This permits a tighter interaction between memory and CPU, since the CPU knows exactly when the data will be available. SDRAM also takes advantage of the row value availability and divides memory into distinct banks, permitting overlapped accesses.
Double Data Rate Synchronous DRAM	DDR SDRAM	The same as SDRAM except that data output is provided on both the negative and the positive clock edges.
Rambus DRAM	RDRAM	A proprietary technology that provides very high memory access rates using a relatively narrow bus.
Error-Correcting Code	ECC	May be applied to most of the DRAM types above to correct single bit data errors and often detect double errors.