Chapter 2

Memory Basics

Memory definitions

- Memory is a major component of a digital computer and is present in a large proportion of all digital systems.
- Memory is a collection of binary storage cells together with associated circuits needed to transfer information into and out of the cells.
- Random-access memory (RAM) stores data temporarily, and read-only memory (ROM) stores data permanently.

Memory Organization

- Organized as an indexed array of memory cells.
- Value of the index for each word is the memory address.
- A Logic circuits, e.g., Decoder, Chip select,…,etc, are required
- Often organized to fit the needs of a particular computer architecture (capacity, and Performance).

Memory Performance

- Access Time (Response time):

  How quickly the memory can respond to a request (R/W).

The max time for applying the address to the appearance of data (READ), or the completion of storing (WRITE).

- Memory cycle time:

  The minimum period between two successive requests

Random-access memory (RAM)

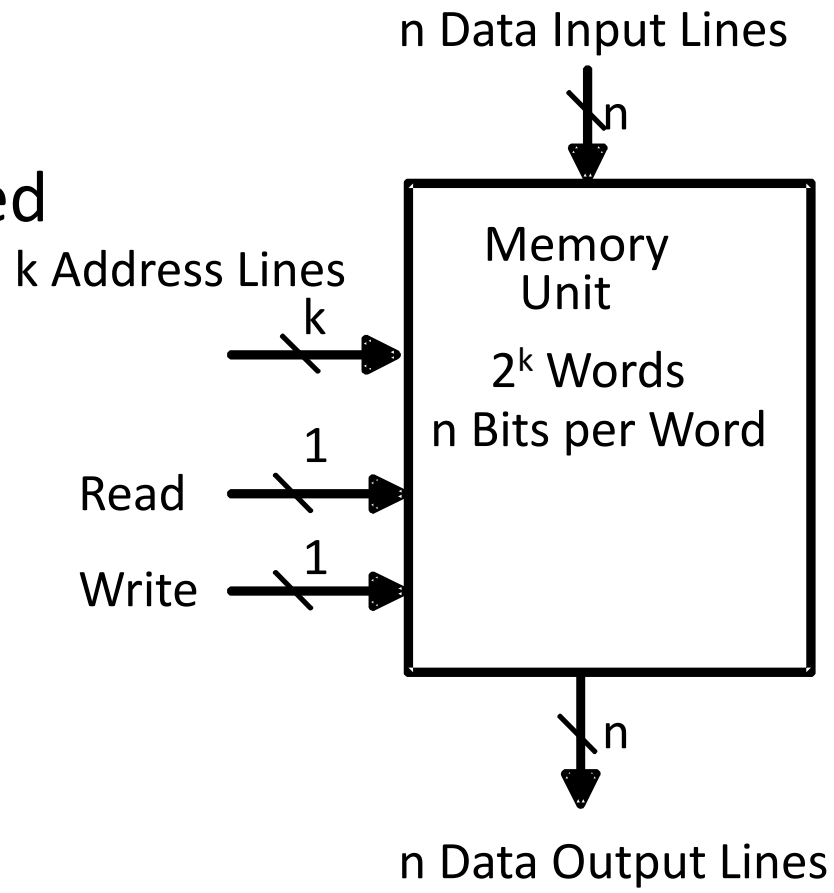
- Binary information is stored in memory in groups of bits, each group of which is called a *word*.
- A word is an entity of bits that moves in and out of memory as a unit.
- A group of eight bits is called a *byte*.
- Most computer memories use words that are multiples of eight bits in length.
- Communication between a memory and its environment is achieved through data input and output lines, address selection lines, and control lines that specify the direction of transfer of information.

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.
  - ***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

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.
- 2 M x 16 RAM: K= 21
  n= 16



Example

- Consider a memory with a capacity of 1K x 16 words . Find K and n?
- Solution:

  Since Memory of 1k words = 2^{10}

  Therefore, K address lines = 10

  Since word = 16 bits long

  Therefore, n data lines = 16

Ex: k = 24, n = 64, memory capacity = 16 M x 8B= 128MB

Basic Memory Operations (continued)

- **Read Memory**

The steps that must be taken for a read are as follows:

1. Apply the binary address of the desired word to the address lines.

2. Activate the Read input.

- **Write Memory**

The steps that must be taken for a write are as follows:

1. Apply the binary address of the desired word to the address lines.

2. Apply the data bits that must be stored in memory to the data input lines.

3. Activate the Write input.

TABLE 1
Control Inputs to a Memory Chip

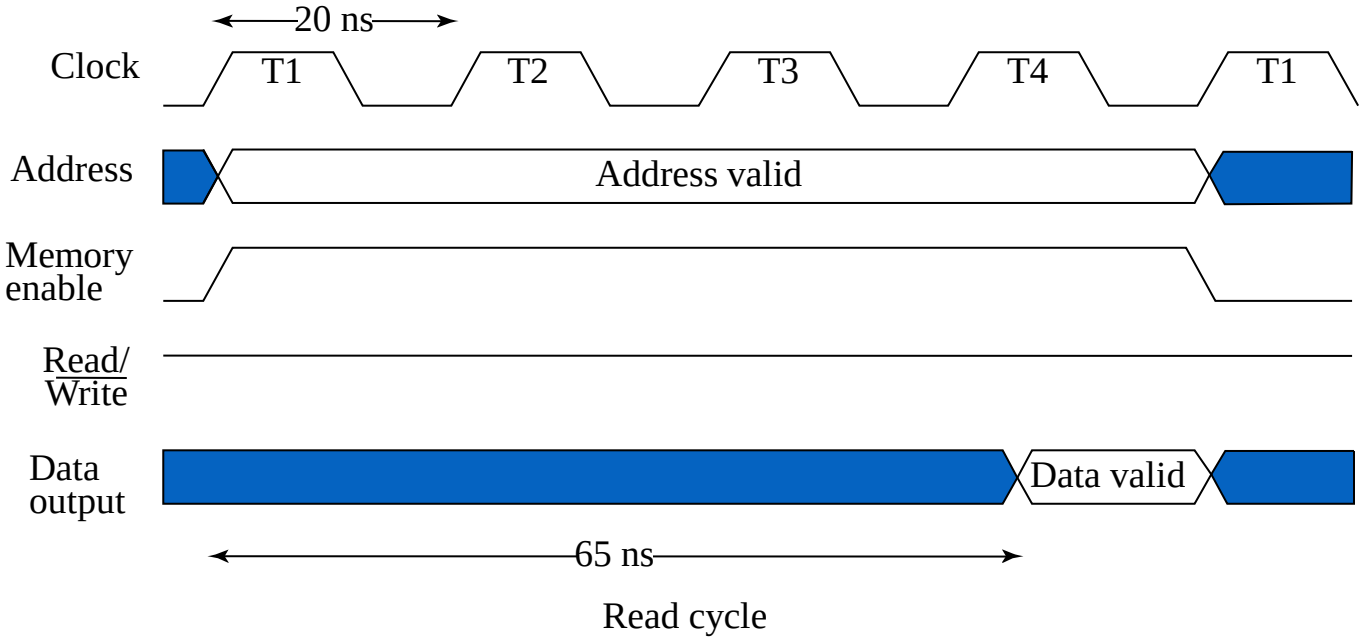
Chip Select CS	Read/$\overline{\text{Write}}$ R/$\overline{\text{W}}$	Memory operation
0	X	None
1	0	Write to selected word
1	1	Read from selected word

Example

- Assume that a CPU operates with a clock frequency of 50 MHz, giving a period of 20 ns for one clock pulse.
- Suppose now that the CPU communicates with a memory with an access time of 65 ns and a write cycle time of 75 ns.

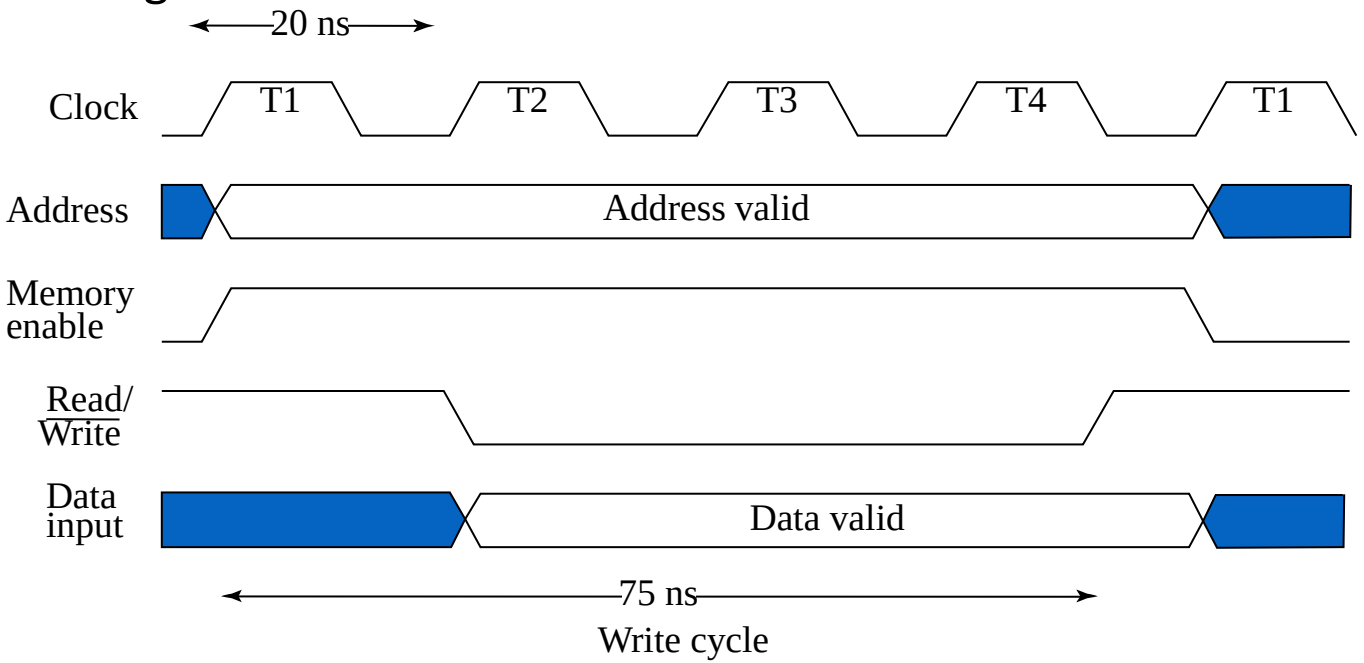
Timing Waveforms

- Most basic memories are asynchronous
  - Storage in latches or storage of electrical charge
  - No clock
- Controlled by control inputs and address
- Timing of signal changes and data observation is critical to the operation
- Read timing:



Memory Operation Timing

- Write timing:

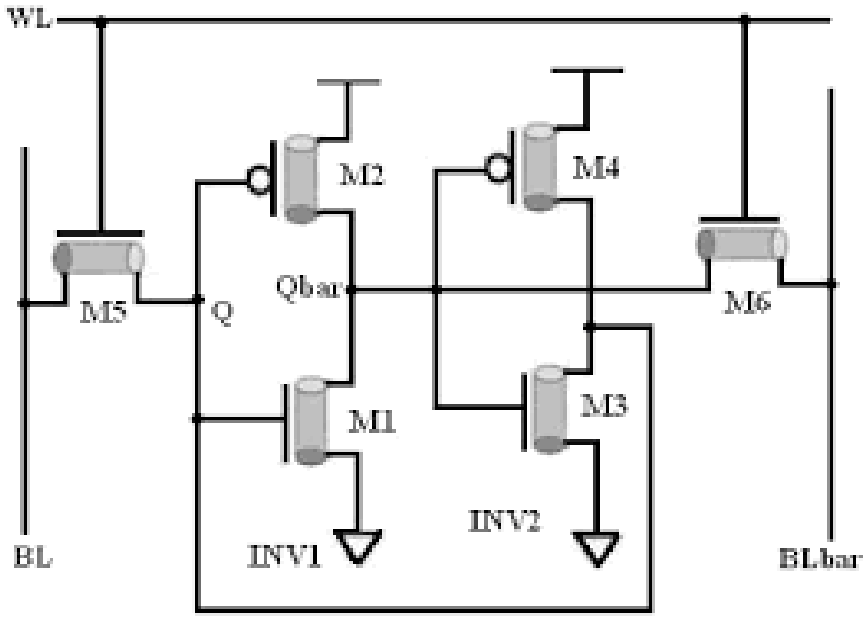


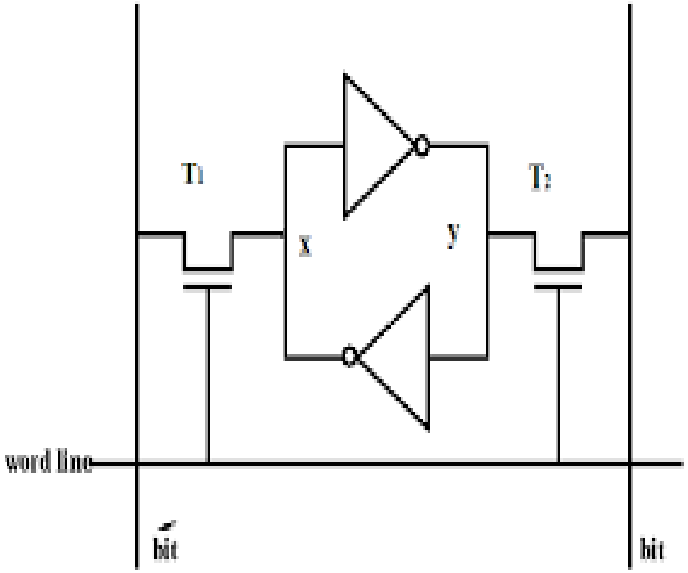
RAM Integrated Circuits

- 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 (SRAM)

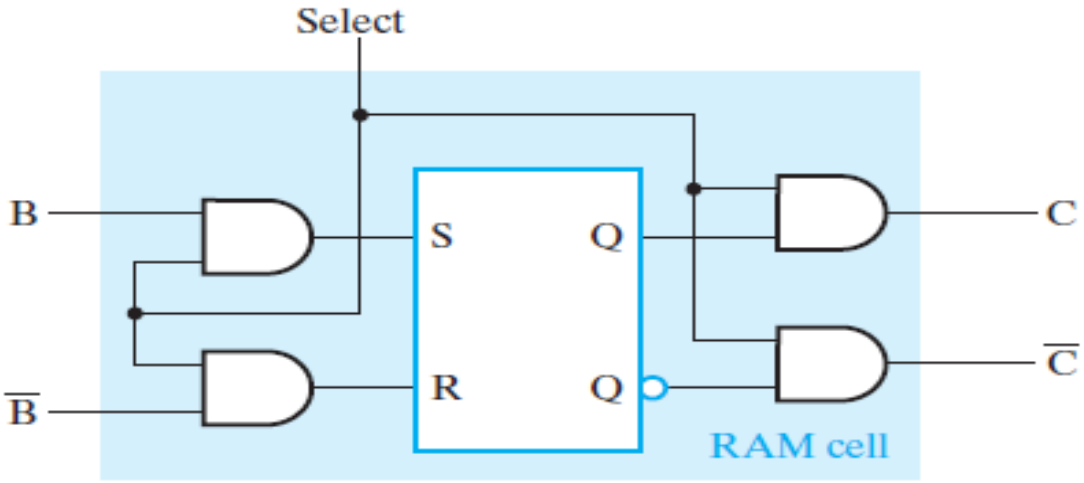
Basic SRAM Electronic Cell





SRAM Model

- As indicated earlier, memory consists of RAM chips plus additional logic. The following figure shows the logic model of the RAM cell. The storage part of the cell is modeled by an *SR* latch. The inputs to the latch are enabled by a Select signal.



Cont.,

- For Select equal to 0, the stored content is held.
- For Select equal to 1, the stored content is determined by the values on *B* and *B complement*.
- The outputs from the latch are gated by Select to produce cell outputs *C* and its *Complement*.
- For Select equal to 0, both *C* and *C complement* are 0, and for Select equal to 1, *C* is the stored value and *C complement* is its complement.

SRAM Bit Slice Model

- To obtain simplified static RAM diagrams, we interconnect a set of RAM cells and read and write circuits to form a *RAM bit slice* that contains all of the circuitry associated with a single bit position of a set of RAM words.
- The loading of a cell latch is now controlled by a Word Select input.
- If this is 0, then both *S* and *R* are 0, and the cell latch contents remain unchanged.
- If the Word Select input is 1, then the value to be loaded into the latch is controlled by two signals *B* and *B complement* from the Write Logic.

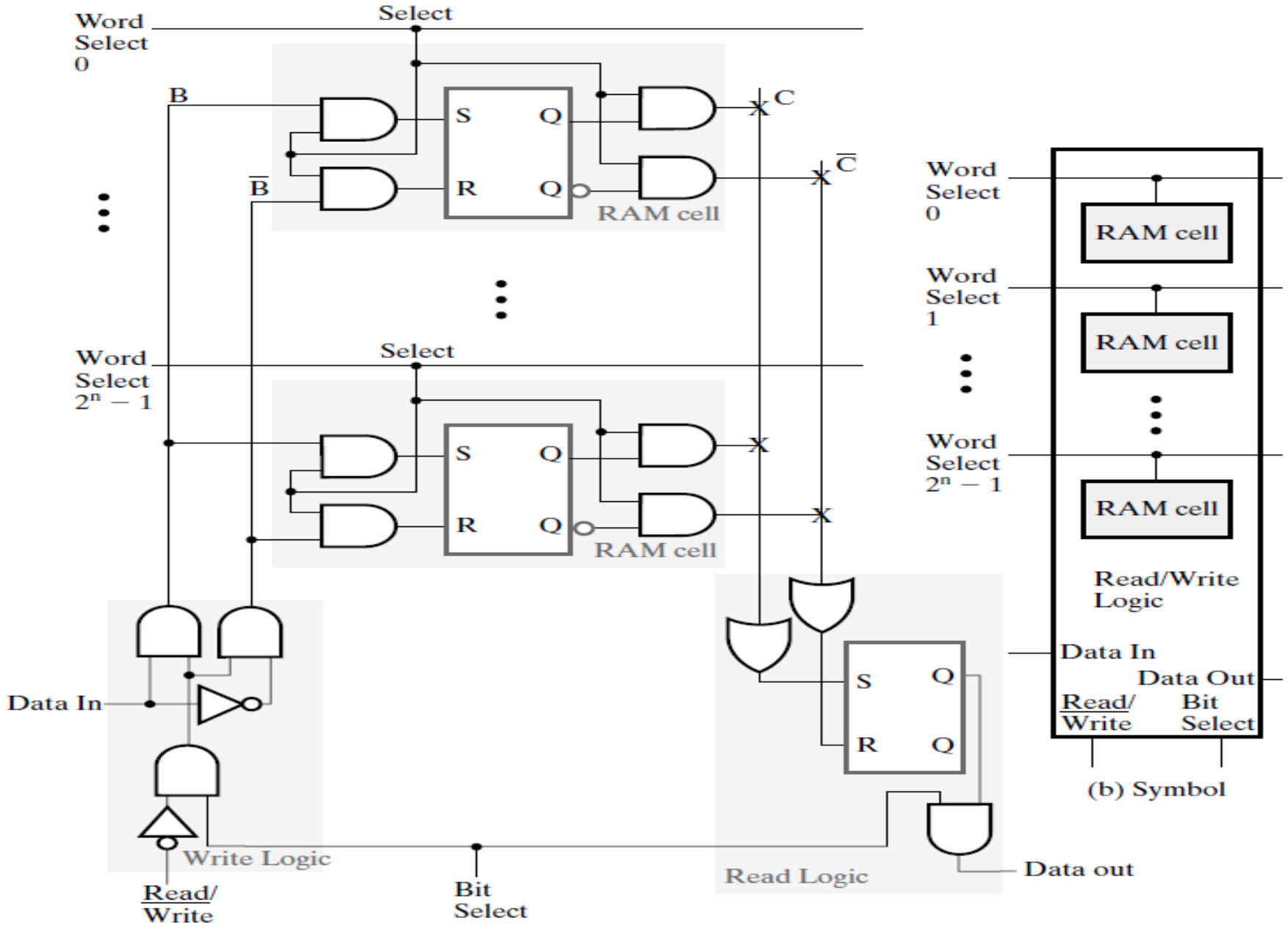
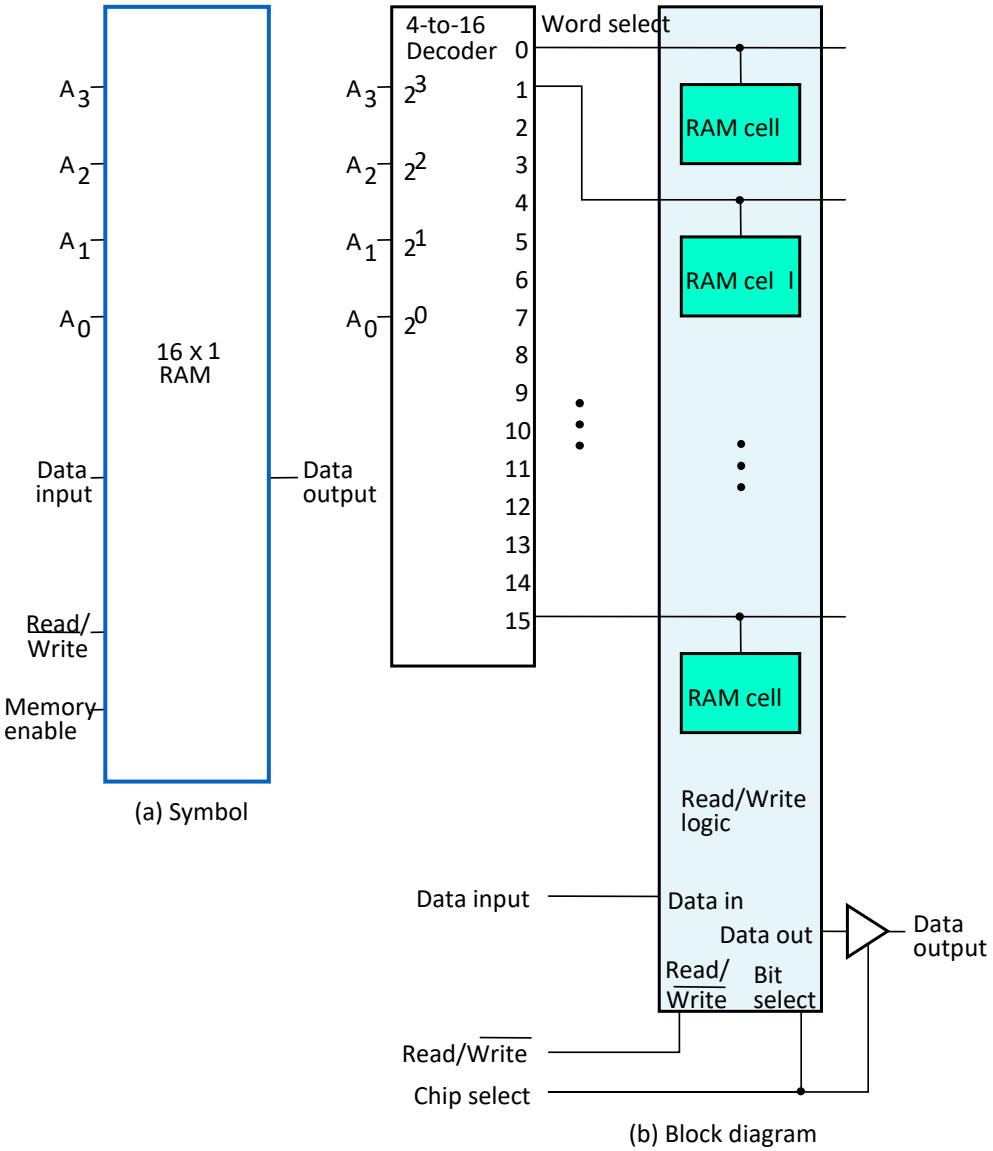


FIGURE 5
RAM Bit Slice Model

2^k Word × 1-Bit RAM IC

- To build a RAM IC from a RAM slice, we need:
  - Decoder :decodes the k address lines to 2^k word select lines
  - A Tri-state buffer : on the data output permits RAM ICs to be combined into a RAM with 2^k x n words



SRAM Design

1-Straightforward Design

- Inside a RAM chip, the decoder with *k* inputs and 2^k outputs requires 2^k AND gates with *k* inputs per gate if a straightforward design approach is used.
- In addition, if the number of words is large, and all bits for one bit position in the word are contained in a single RAM bit slice, the number of RAM cells sharing the read and write circuits is also large.
- Which causes the access and write cycle times of the RAM to become long.

2- Coincident Selection Design

- Since the Memory arrays can be very large => The decoder size and fan outs can be reduced 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 (vertical decoder)
  - Bit select becomes Column select (Horizontal decoder)
- I.e., two *k*/2-input decoders are used instead of one *k*-input decoder.

Procedure

1- Find physical address lines = K

2-Find the Square Root of the total number of memory bits = $\sqrt{total\ number\ of\ bits}$

To make the memory square:

1- The row decoder takes its k/2 lines from the MSB of the address lines

2- The rest of the physical address lines are given to the column decoder.

Example

- Design a 16 × 1 RAM Using a 4 × 4 RAM Cell Array ?
- Solution:
- Since, memory has 16 memory locations
- Therefore, Physical address lines K = 4
- $\sqrt{total\ number\ of\ bits} = \sqrt{16} = 2^2$ => K / 2 = 2
- Therefore, the row decoder takes its two MSBs,
- And , the column decoder takes its two LSBs.

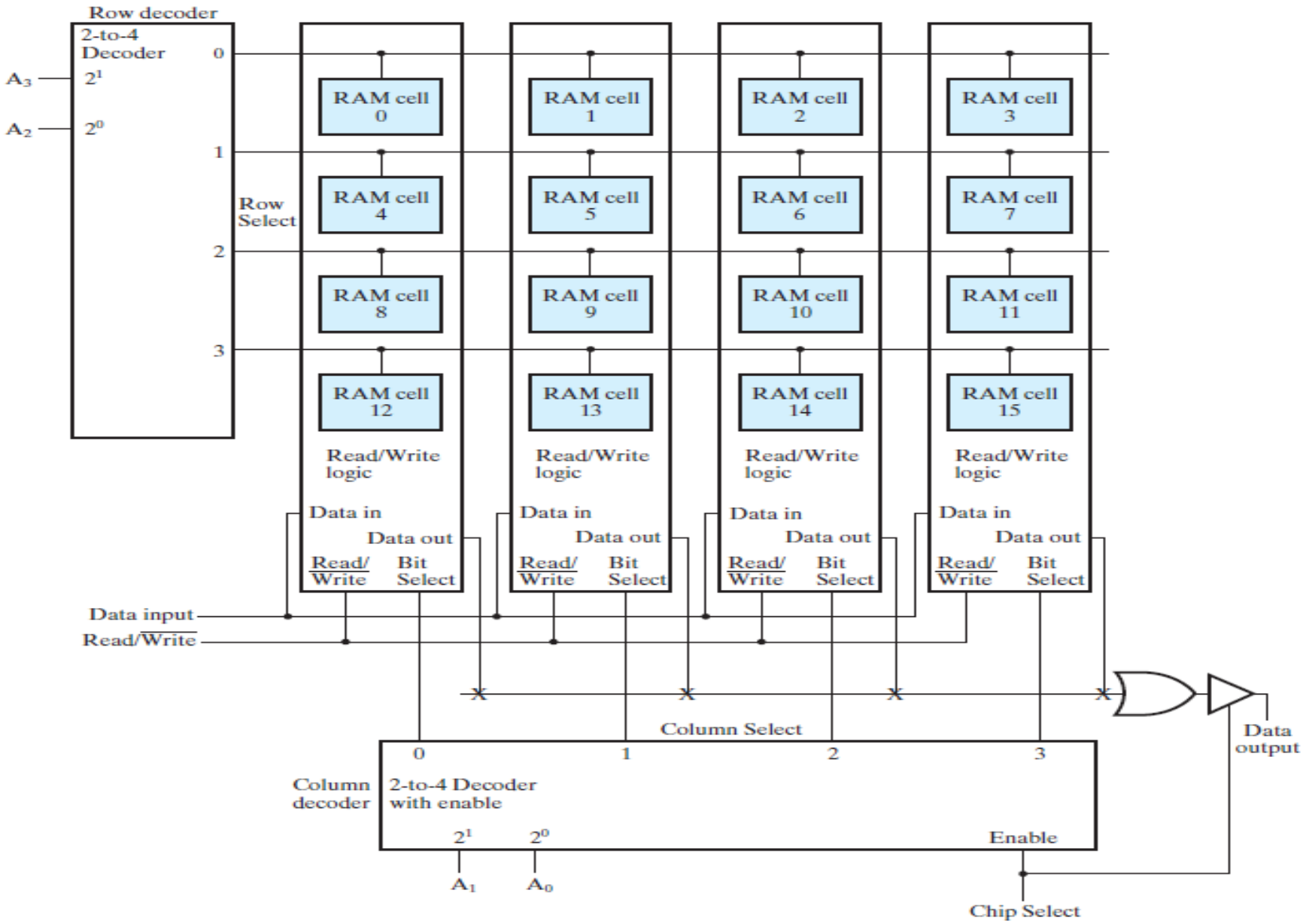


FIGURE 7-7
Diagram of a 16 × 1 RAM Using a 4 × 4 RAM Cell Array

Example

- Design 32k x 8 RAM using coincident selection ?
- Solution:
- Since, memory has 32K memory locations = 2^{15}
- Therefore, Physical address lines K = 15
- $\sqrt{total\ number\ of\ bits}$= $\sqrt{2^{15}\ x\ 8}$ = 2^{9}

  Therefore, K / 2 = 9
- Therefore, the row decoder takes its 9 MSBs,
- And , the column decoder takes the rest from the physical address = 15-9 = 6 LSBs.
- To make the memory array square, each output of the column decoder was multiplied by 8. i.e., select 8 bits.

Assignment

1. Explain the construction & operation of sense amplifier in SRAM?

2. The following memories are specified by the number of words times the number of bits per word. How many address lines and input–output data lines are needed in each case?

(a) 48K × 8, (b) 512K × 32, (c) 64M × 64, and

(d) 2G × 1.

Cont.,

3. *A 64K × 16 RAM chip uses coincident decoding by splitting the internal decoder into row select and column select. Assuming that the RAM cell array is square, what is the size of each decoder, and how many AND gates are required for decoding an address?

ARRAY OF SRAM ICs

- If the memory unit needed 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 memory.
- The capacity of the memory depends on two parameters: the number of words and the number of bits per word.

Procedure

- An increase in the number of words requires that we increase the address length.
- Every bit added to the length of the address doubles the number of words in memory.
- An increase in the number of bits per word requires that we increase the number of data input and output lines, but the address length remains the same.

Cont.,

- To construct the memory with address lines k1 and data lines n1, using memory chip with address lines K2 and data lines n2:

  1- number of chips / column = $2^{k1} / 2^{k2} = 2^{k1-k2}$

  So, we have to add a new decoder with address lines = k1 – k2, for CS

  2- number of columns = n1 / n2

Example

- Construct a 256k x 8 RAM using 64k x 8 RAM chips
- Solution:
- K1 = 18, n1 = 8
- K2 = 16, n2 = 8
- Therefore:

1- number of chips / column = $2^{18}/\ 2^{16} = 2^2 = 4$

2- number of columns = n1 / n2 = 8/8=1

3- Decoder 2 x 4

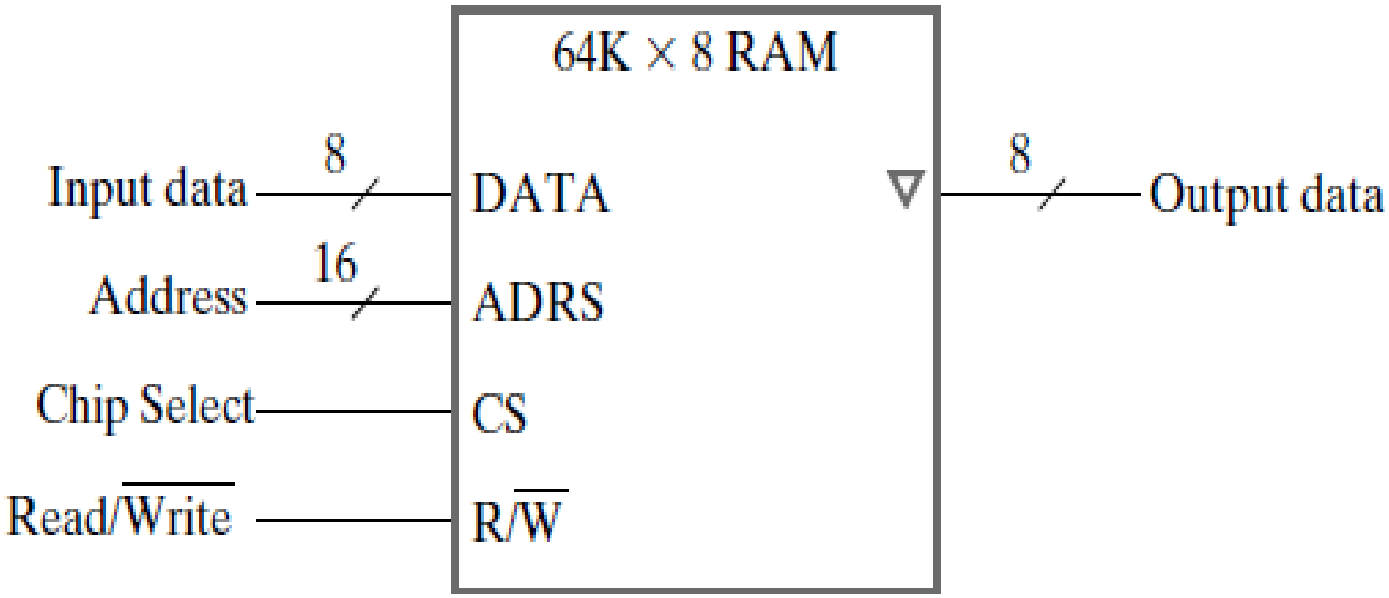


FIGURE 9
Symbol for a 64K × 8 RAM Chip

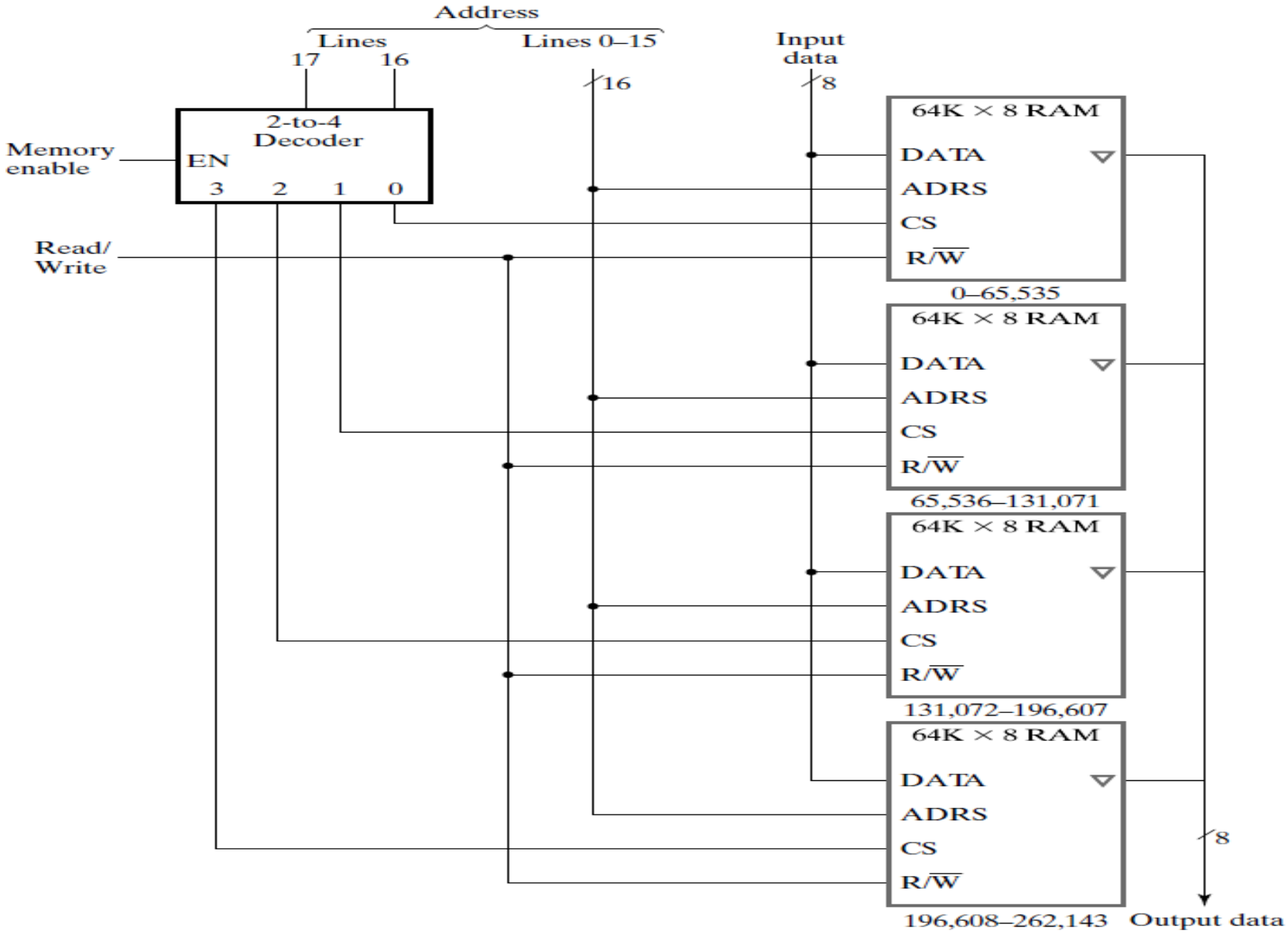
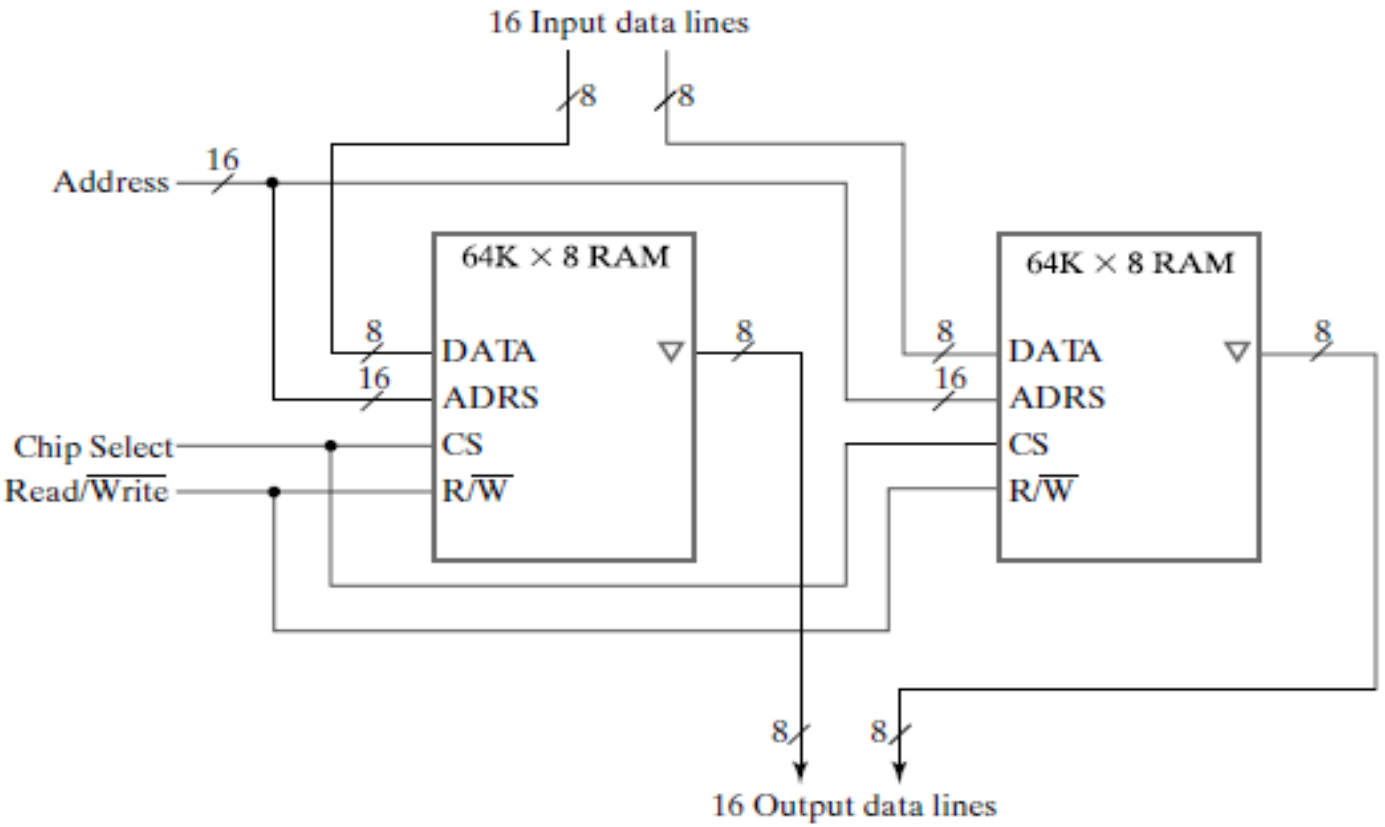


FIGURE 10
Block Diagram of a 256K × 8 RAM

Example

- Construct a 64K x 16 RAM using 64k x 8 RAM.



Sheet

1- Using the 64K × 8 RAM chip plus a decoder, construct the block diagram for a 512K × 16 RAM.

2- ***(a)** How many 128K × 16 RAM chips are needed to provide a memory capacity of 2 MB = 1 M x 16 RAM?

(b) How many address lines are required to access 2 MB? How many of these lines are connected to the address inputs of all chips?