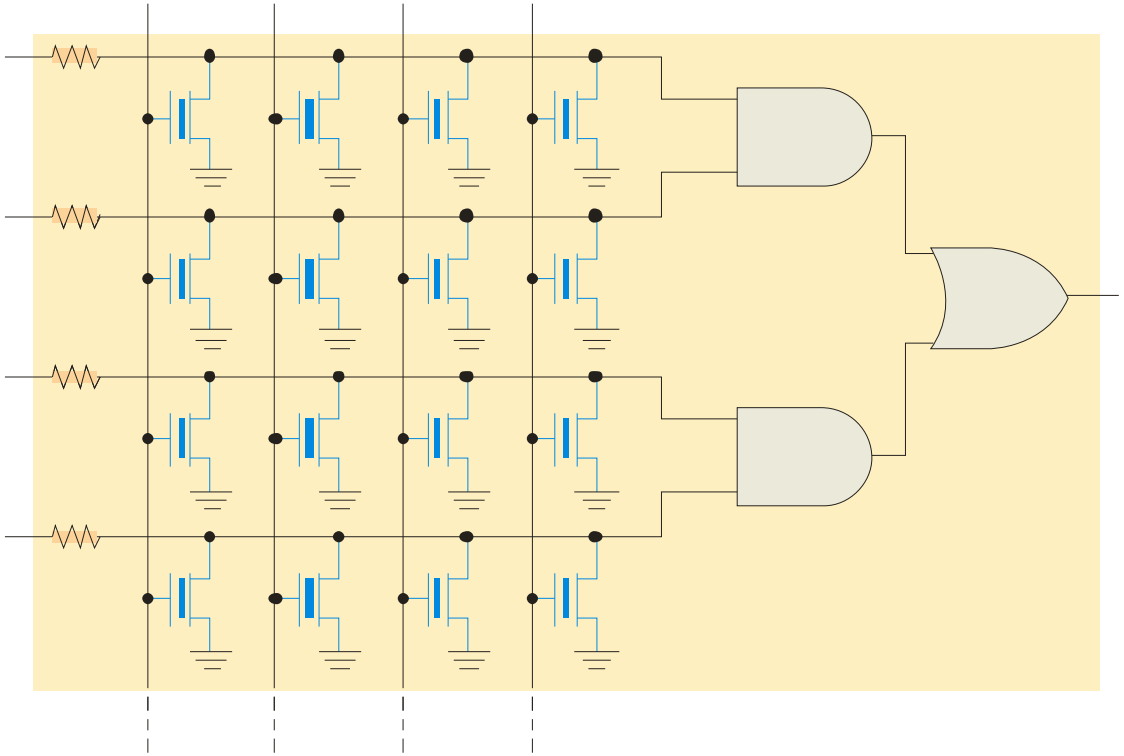
Digital Electronics

Memory & Storage



Memory Units

Memories store data in units from one to eight bits. The most common unit is the **byte**, which by definition is 8 bits.

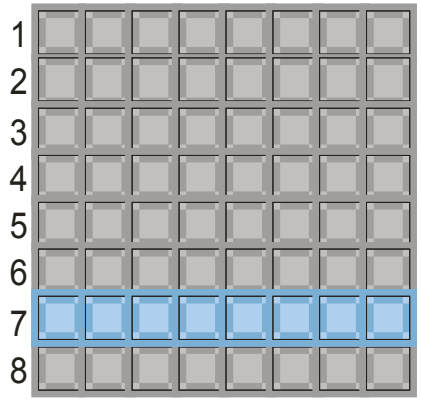
Computer memories are organized into multiples of bytes called words. Generally, a word is defined as the number of bits handled as one entity by a computer. By this definition, a word is equal to the internal register size (usually 16, 32, or 64 bits).

For historical reasons, assembly language defines a word as exactly two bytes. In assembly language, a 32 bit entity is called a double-word and 64 bits is defined as a quad-word.

Memory Units

The location of a unit of data in a memory is called the **address**. In PCs, a byte is the smallest unit of data that can be accessed.

In a 2-dimensional array, a byte is accessed by supplying a row number. For example the blue byte is located in row 7.



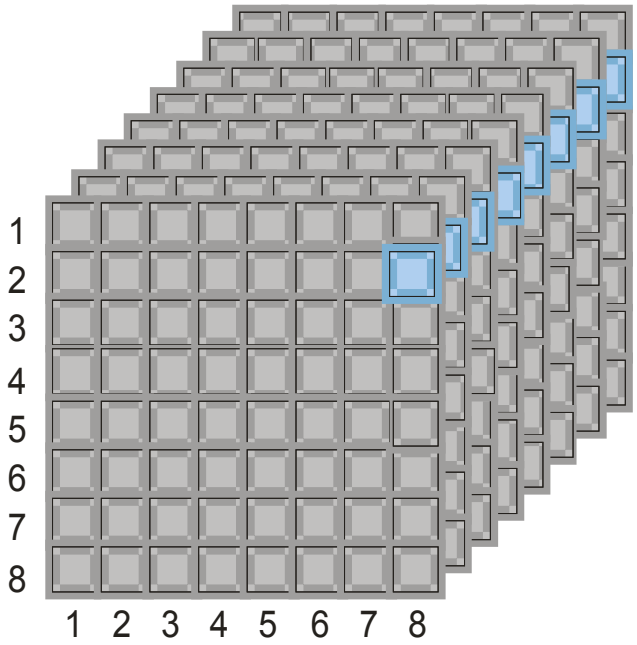
Memory Addressing

A 3-dimensional array is arranged as rows and columns. Each byte has a unique row and column address.

Question

a) How many bytes are shown?

b) What is the location of the blue byte?



Answers

a) 64 B

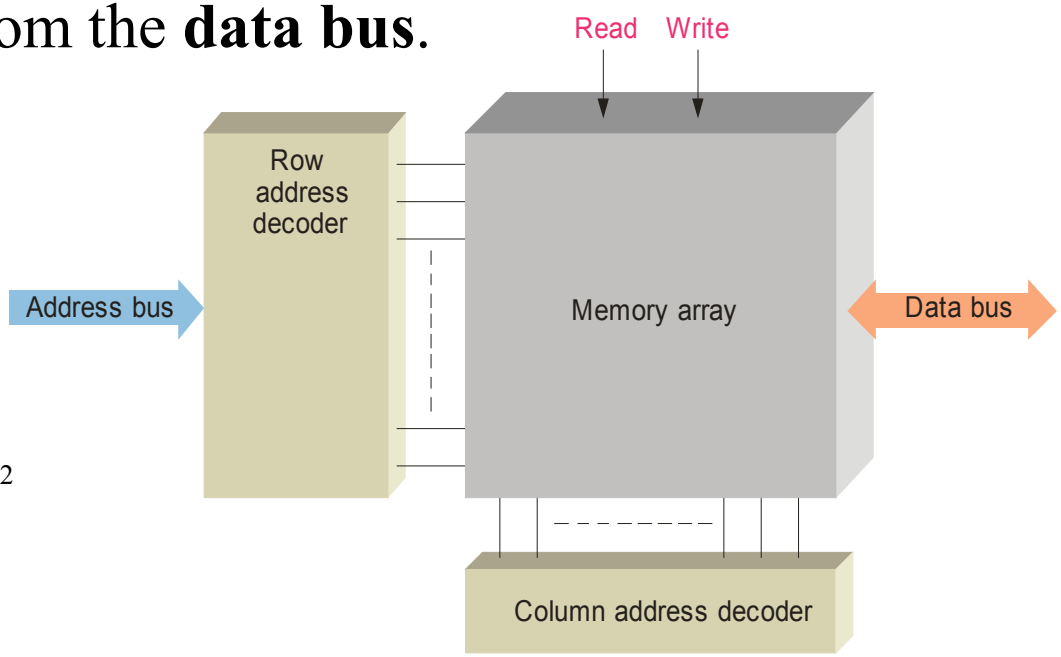
b) Row 2, column 8

This example is (of course) only for illustration. Typical computer memories have 256 MB or more of capacity.

Memory Addressing

In order to read or write to a specific memory location, a binary code is placed on the **address bus**. Internal decoders decode the address to determine the specific location. Data is then moved to or from the **data bus**.

The address bus is a group of conductors with a common function. Its size determines the number of locations that can be accessed. A 32 bit address bus can access 2^{32} locations, which is approximately 4G.



Memory Addressing

In addition to the address bus and data bus, semiconductor memories have read and write control signals and chip select signals. Depending on the type of memory, other signals may be required.

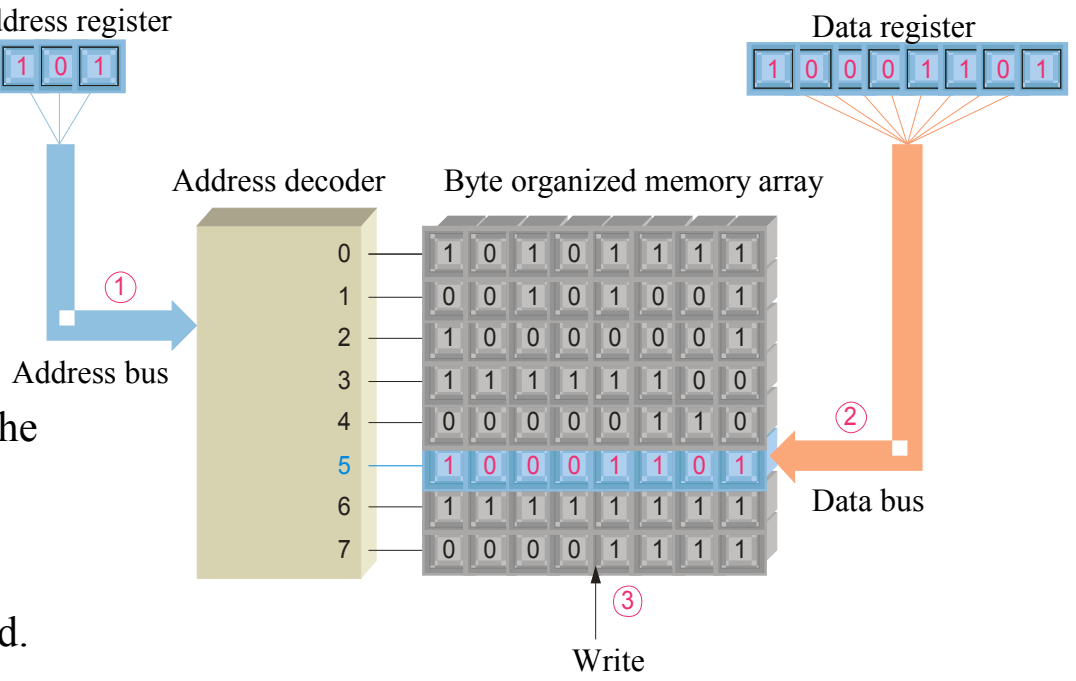
Read Enable ($\overline{RE}$) and **Write Enable** ($\overline{WE}$) signals are sent from the CPU to memory to control data transfer to or from memory.

Chip Select ($\overline{CS}$) or **Chip Enable** ($\overline{CE}$) is used as part of address decoding. All other inputs are ignored if the Chip Select is not active.

Output Enable ($\overline{OE}$) is active during a read operation, otherwise it is inactive. It connects the memory to the data bus.

Read and Write Operations

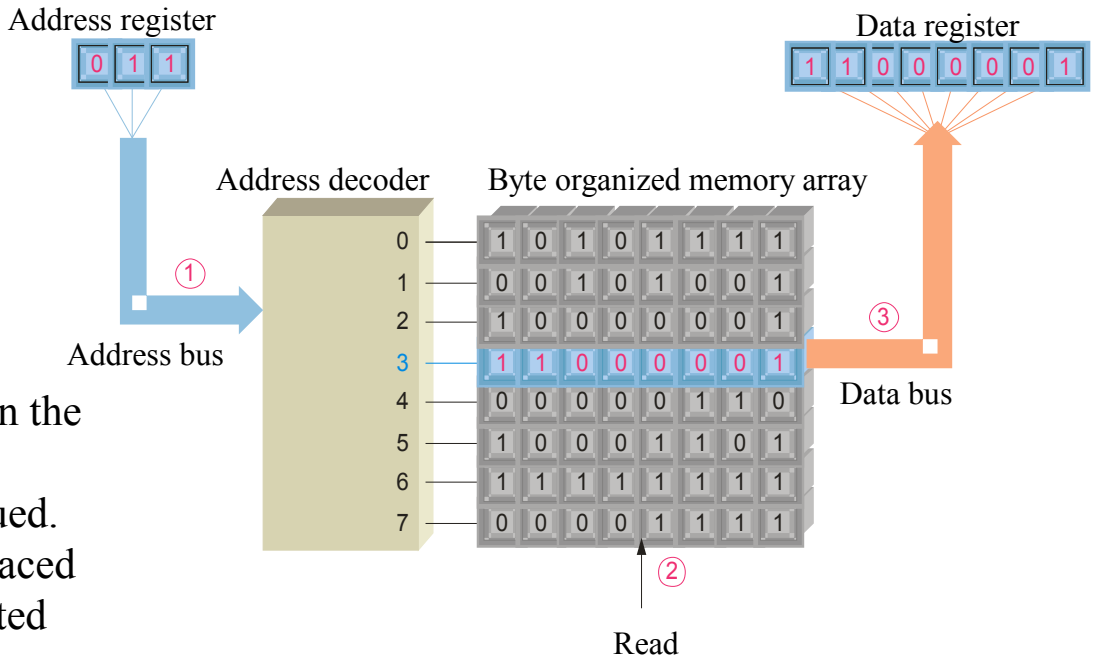
The two main memory operations are called **read** and **write**. A simplified write operation is shown in which new data overwrites the original data. Data moves *to* the memory.



1. The address is placed on the address bus.
2. Data is placed on the data bus.
3. A write command is issued.

Read and Write Operations

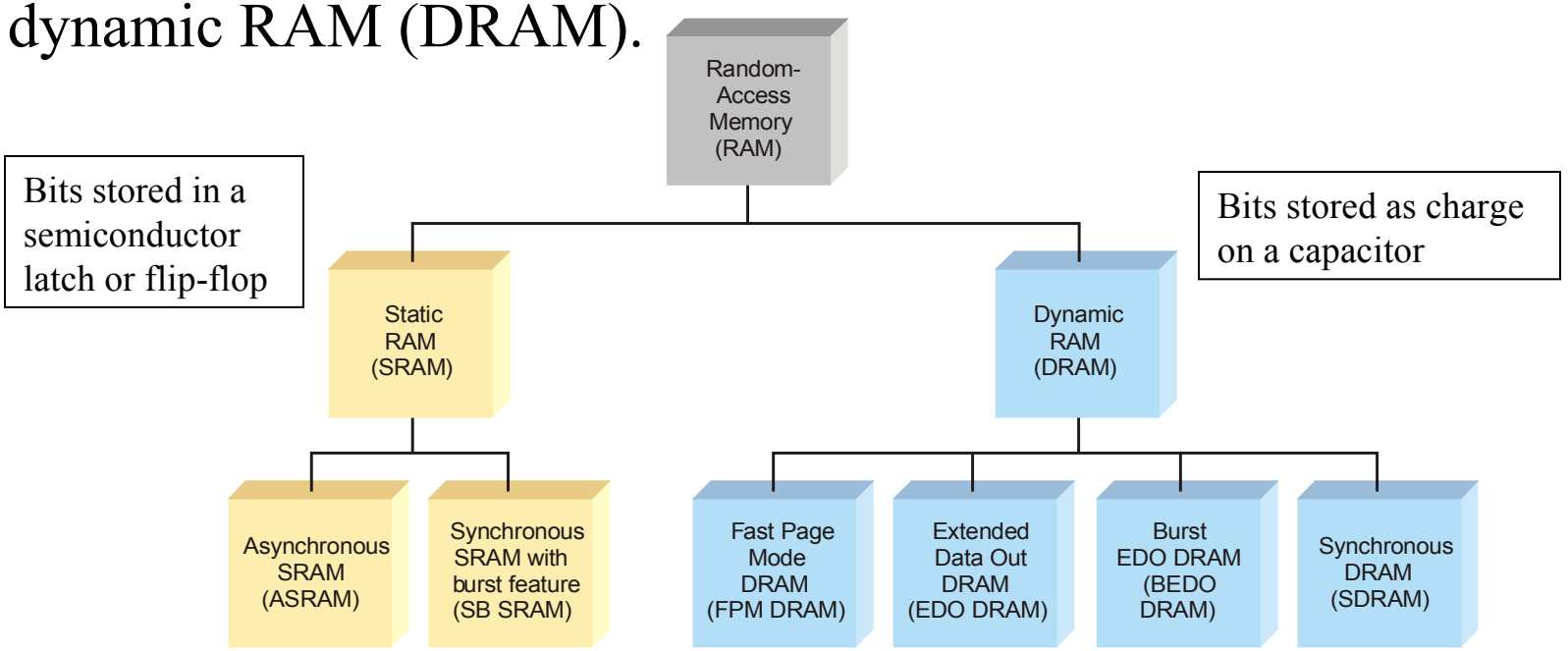
The read operation is actually a "copy" operation, as the original data is not changed. The data bus is a "two-way" path; data moves *from* the memory during a read operation.



1. The address is placed on the address bus.
2. A read command is issued.
3. A copy of the data is placed in the data bus and shifted into the data register.

Random Access Memory

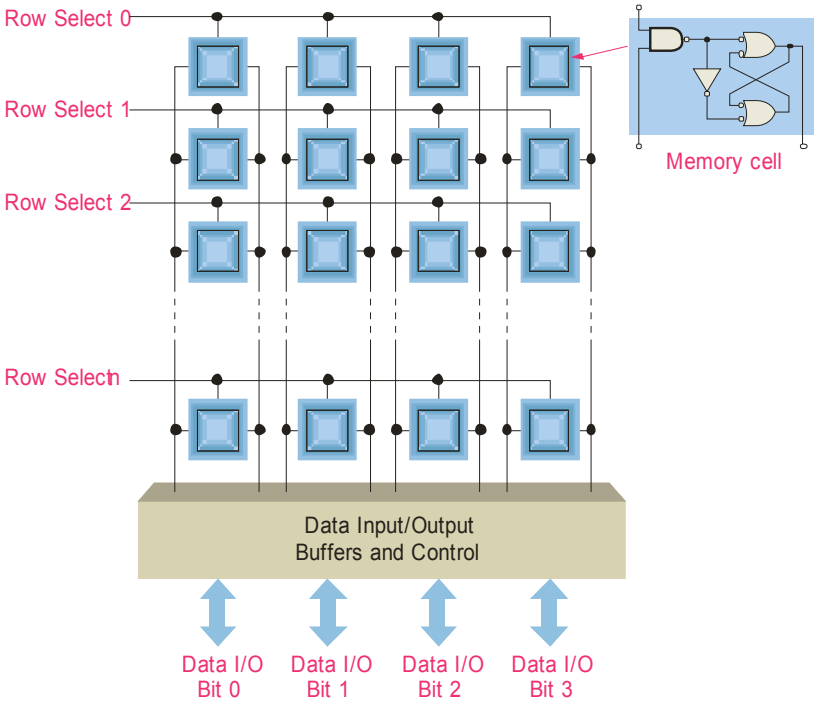
RAM is for temporary data storage. It is read/write memory and can store data only when power is applied, hence it is *volatile*. Two categories are static RAM (SRAM) and dynamic RAM (DRAM).



Static RAM

SRAM uses semiconductor latch memory cells. The cells are organized into an array of rows and columns.

SRAM is faster than DRAM but is more complex, takes up more space, and is more expensive. SRAMs are available in many configurations – a typical large SRAM is organized as 512 k X 8 bits.



Asynchronous Static RAM

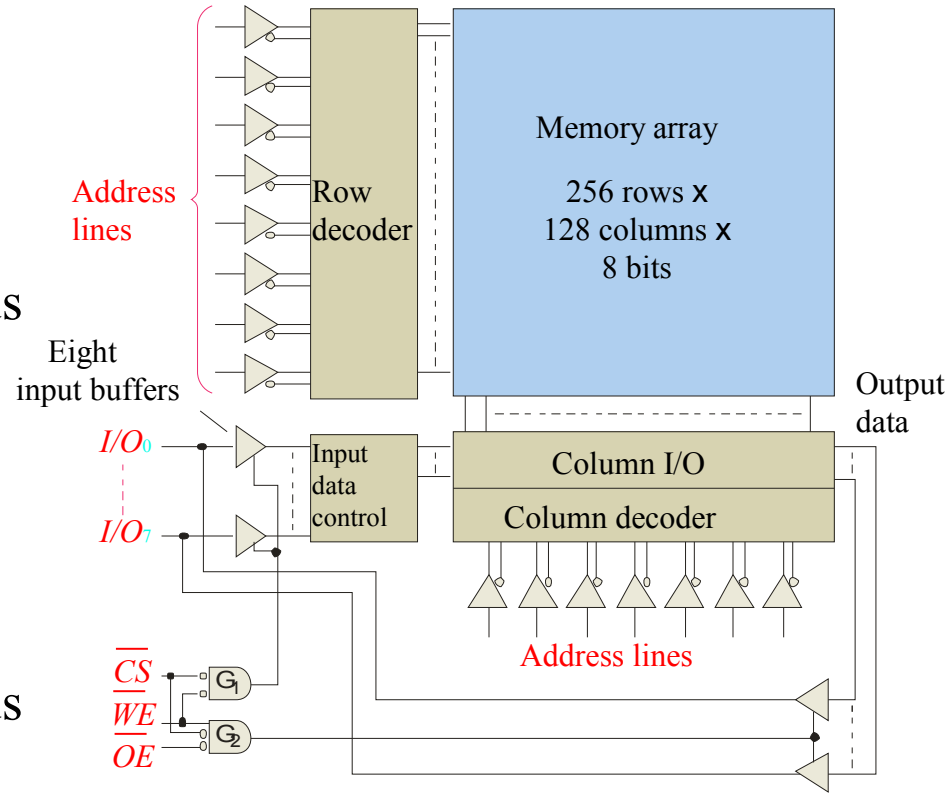
The basic organization of an asynchronous SRAM is shown.

Read cycle sequence:

- A valid address is put on the address bus
- Chip select is LOW
- Output enable is LOW
- Data is placed on the data bus

Write cycle sequence:

- A valid address is put on the address bus
- Chip select is LOW
- Write enable is LOW
- Data is placed on the data bus

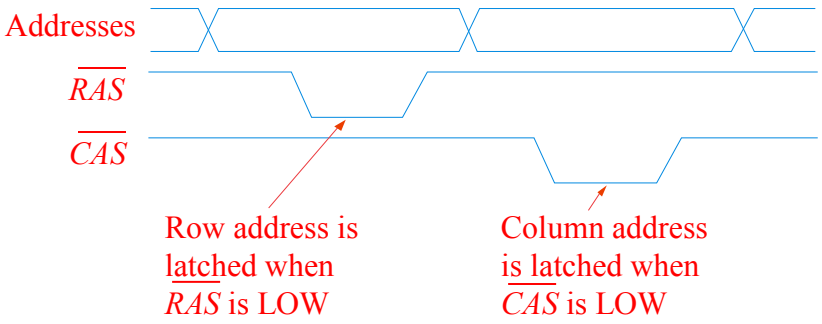


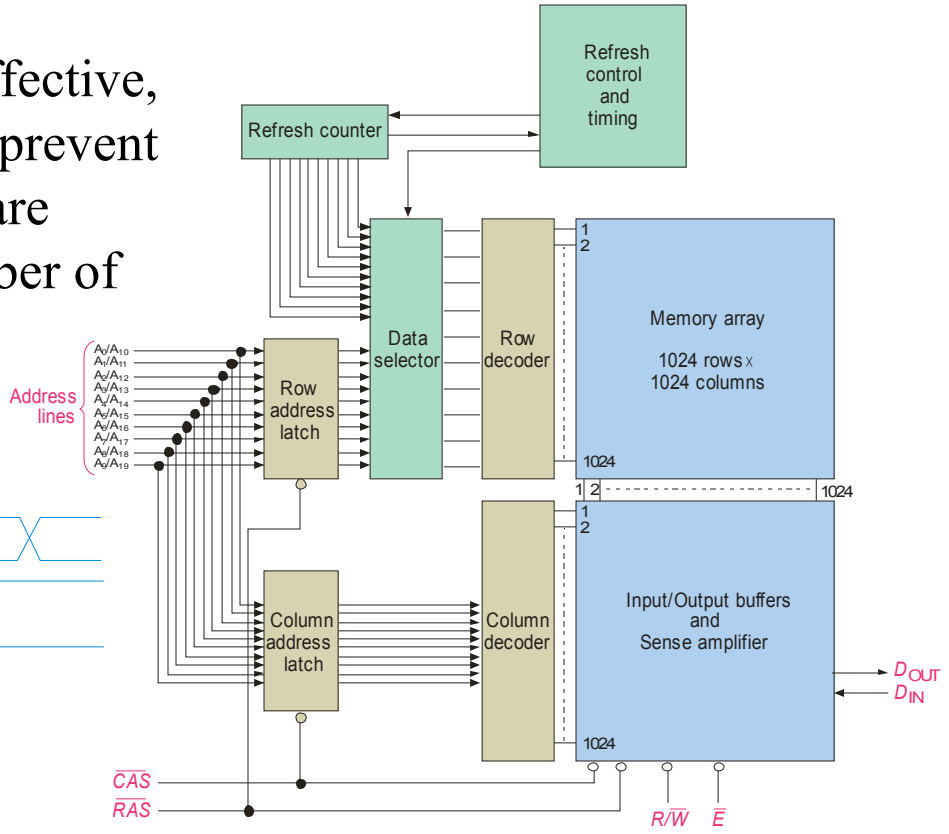
Dynamic RAM (DRAM)

Dynamic RAMs (DRAMs) store data bits as a charge on a capacitor.

DRAMs are simple and cost effective, but require refresh circuitry to prevent losing data. The address lines are multiplexed to reduce the number of address lines.

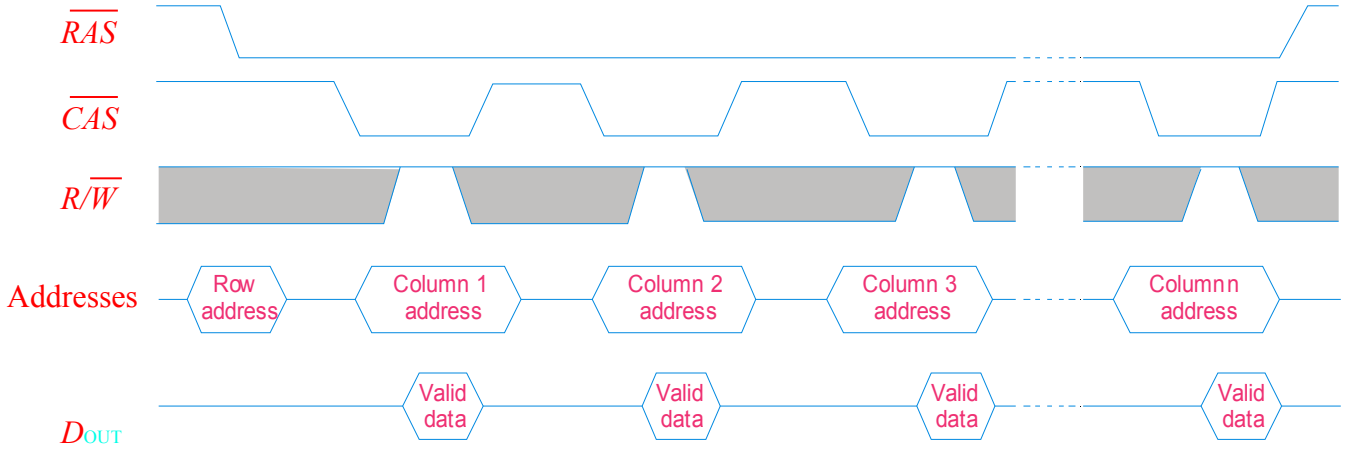
Multiplexed address lines:





Dynamic RAM (DRAM)

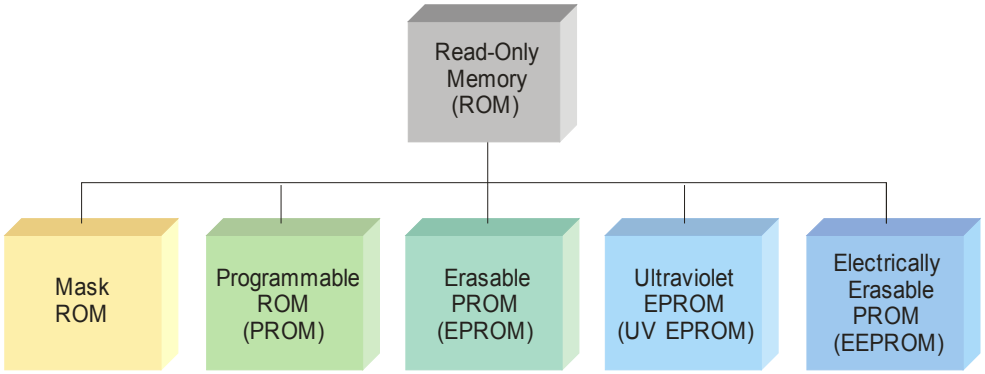
A feature with some DRAMs is fast page mode. Fast page mode allows successive read or write operations from a series of columns address that are all on the same row.



Other types of DRAMs have been developed to speed access and make the processor more efficient. These include EDO DRAMs, BEDO DRAMs and SDRAMs, as described in the text.

Read-Only Memory (ROM)

The ROM family is all considered non-volatile, because it retains data with power removed. It includes various members that can be either permanent memory or erasable.

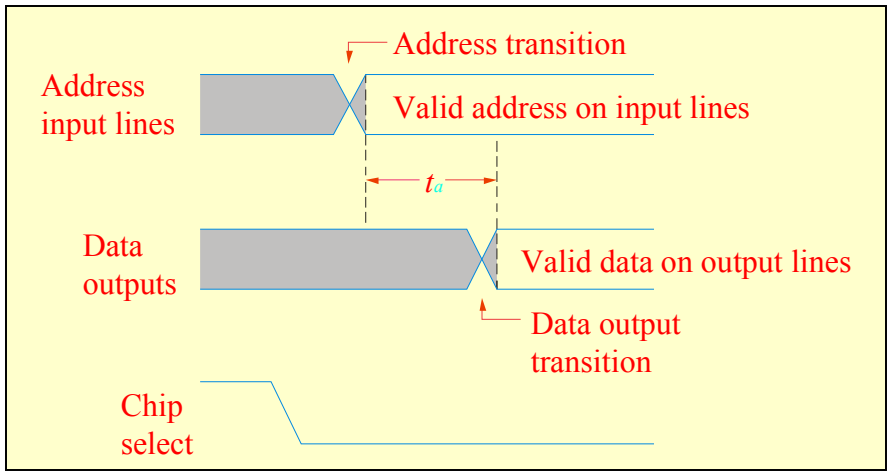


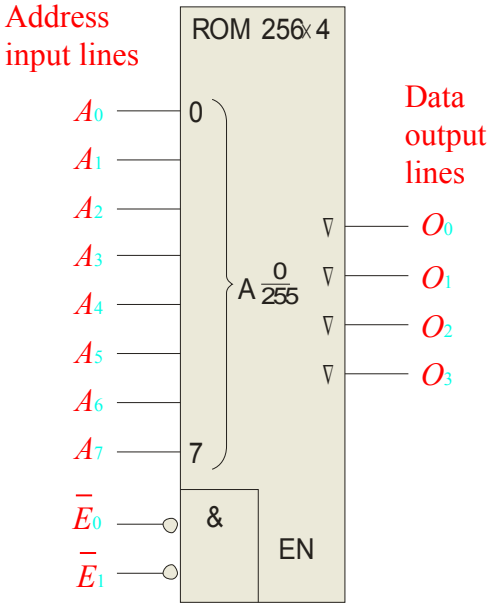
ROMs are used to store data that is never (or rarely) changed such as system initialization files. ROMs are *non-volatile*, meaning they retain the data when power is removed, although some ROMs can be reprogrammed using specialized equipment.

Read-Only Memory (ROM)

A ROM symbol is shown with typical inputs and outputs. The triangles on the outputs indicate it is a tri-stated device.

To read a value from the ROM, an address is placed on the address bus, the chip is enabled, and a short time later (called the access time), data appears on the data bus.

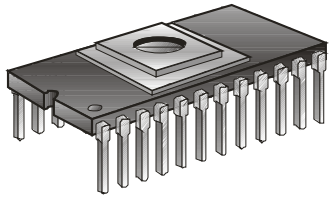


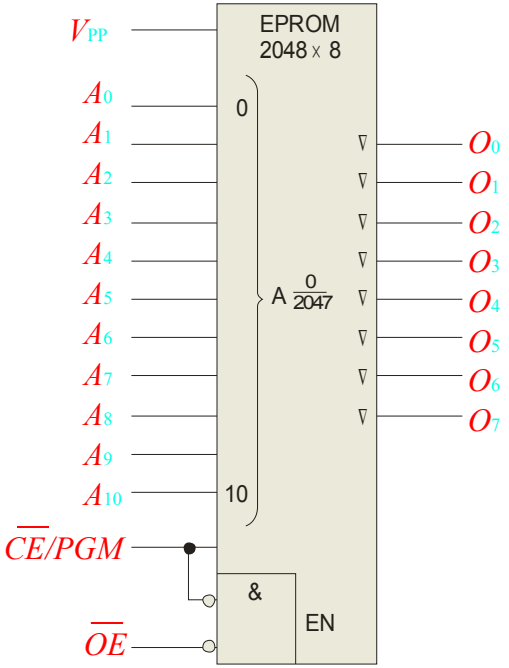


PROMs, EPROMs and EEPROMs

PROMs are programmable ROM, in which a fused link is burned open during the programming process. Once the PROM is programmed, it cannot be reversed.

An EPROM is an erasable PROM and can be erased by exposure to UV light through a window. To program it, a high voltage is applied to V_{PP} and $\overline{OE}$ is brought LOW.



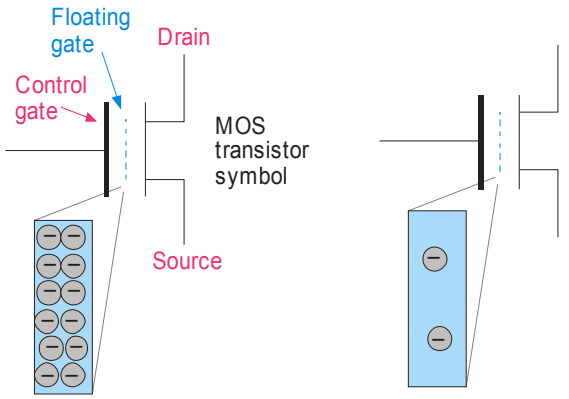


Another type of erasable PROM is the EEPROM, which can be erased and programmed with electrical pulses.

Flash Memory

Flash memories are high density read/write memories that are nonvolatile. They have the ability to retain charge for years with no applied power.

Flash memory uses a MOS transistor with a floating gate as the basic storage cell. The floating gate can store charge (logic 0) when a positive voltage is applied to the control gate. With little or no charge, the cell stores a logic 1.



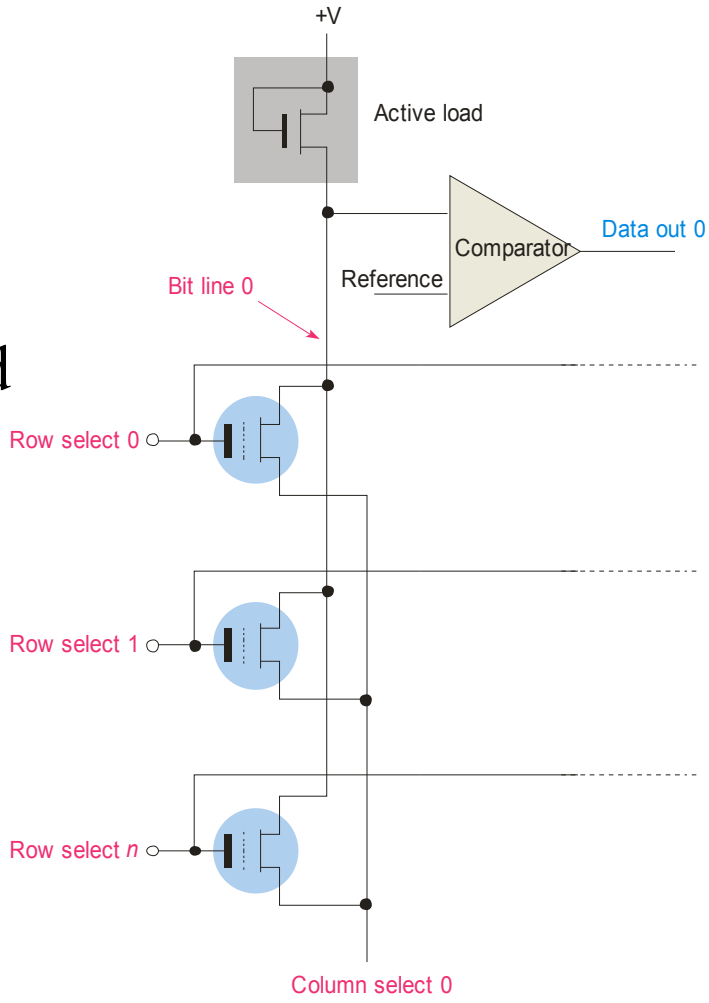
logic 0 is stored logic 1 is stored

The flash memory cell can be read by applying a positive voltage to the control gate. If the cell is storing a 1, the positive voltage is sufficient to turn on the transistor; if it is storing a 0, the transistor is off.

Flash Memory

Flash memories arranged in arrays with an active load. For simplicity, only one column is shown. When a specific row and column is selected during a read operation, the active load has current.

One drawback to flash memory is that once a bit has been set to 0, it can be reset to a 1 only by erasing an entire block of memory. Another limitation is that flash memory has a large but finite number of read/write cycles.

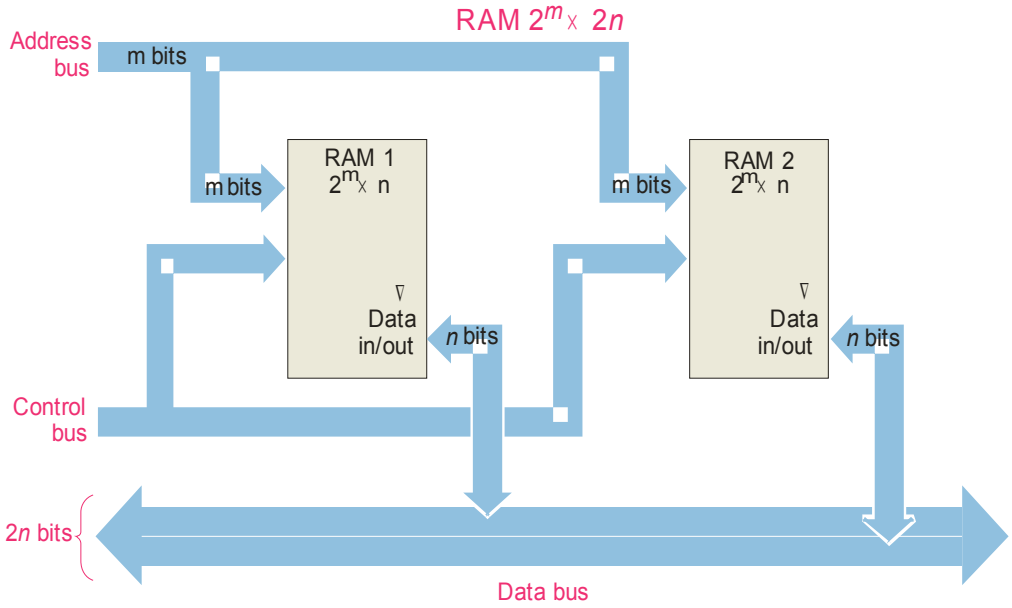


Memory Expansion

Memory can be expanded in either word size or word capacity or both.

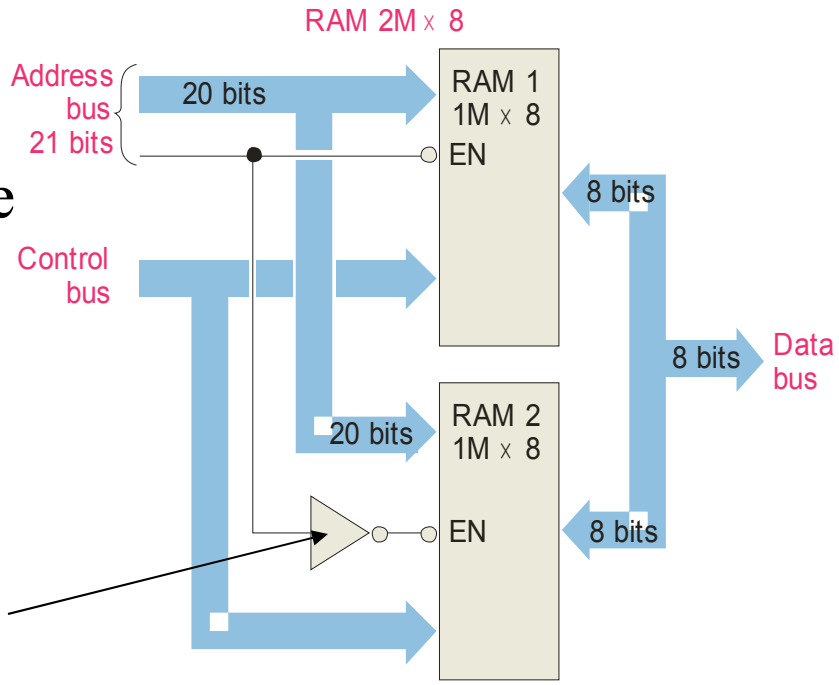
To expand word size:

Notice that the data bus size is larger, but the number of address is the same.



Memory Expansion

To expand word capacity, you need to add an address line as shown in this example



Notice that the data bus size does not change.

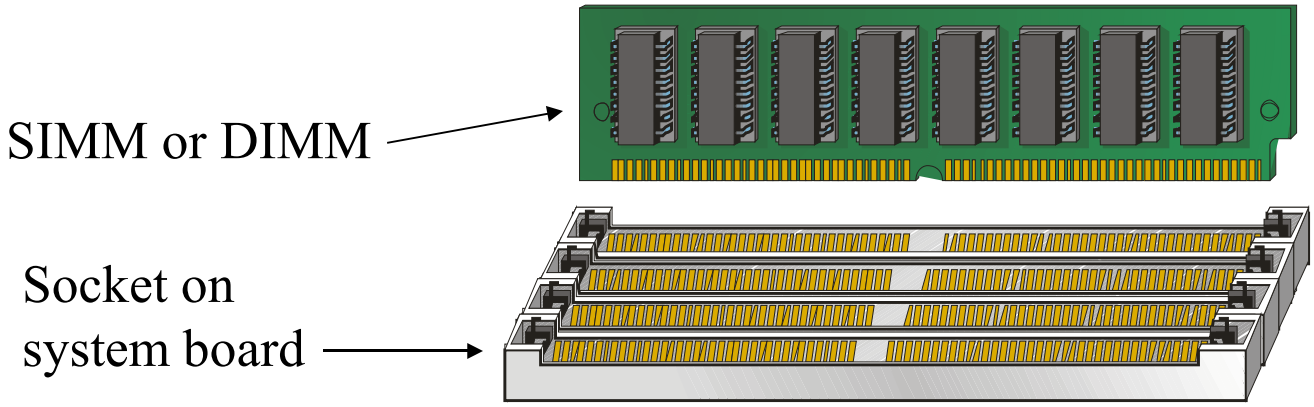
Question

What is the purpose of the inverter?

Answer Only one of the ICs is enabled at any time depending on the logic on the added address line.

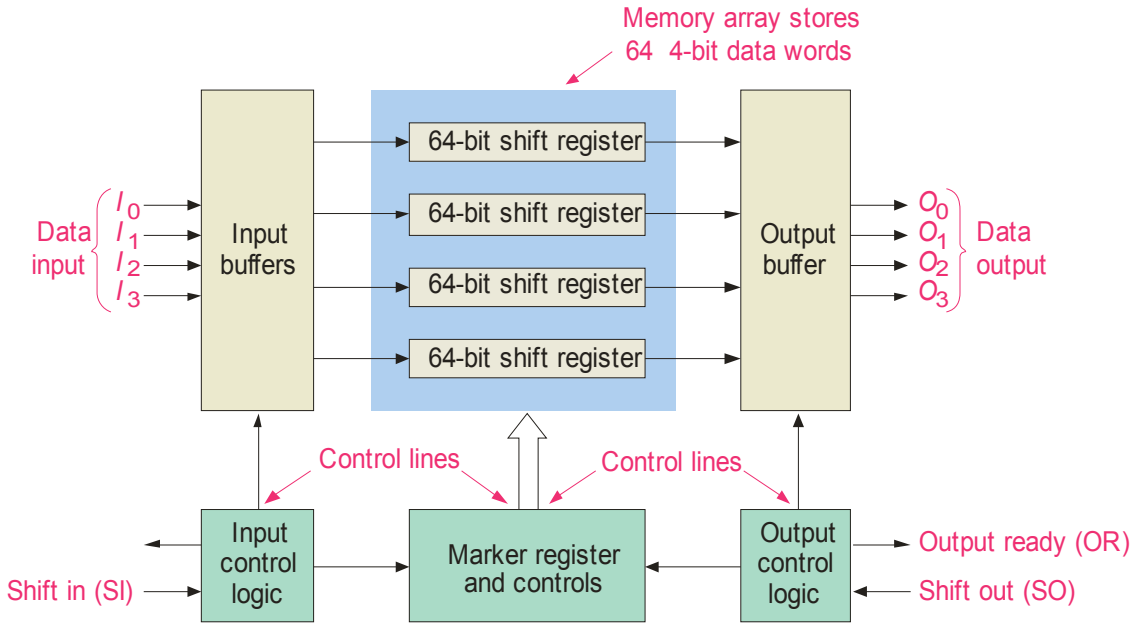
SIMMs and DIMMs

SIMMs (single in-line memory modules) and DIMMs (dual in-line memory modules) are plug-in circuit boards containing the ICs and I/O brought out on edge connectors. SIMMs have a 32-bit data path with I/O on only one side whereas DIMMs have a 64-bit data path with I/O on both sides of the board.



FIFO Memory

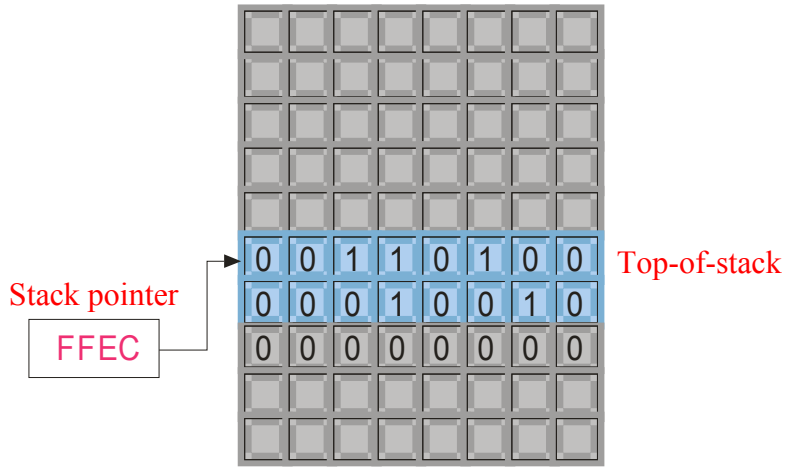
FIFO means first in-first out. This type of memory is basically an arrangement of shift registers. It is used in applications where two systems communicate at different rates.



LIFO Memory

LIFO means last in-first out. In microprocessors, a portion of RAM is devoted to this type of memory, which is called the **stack**. Stacks are very useful for temporary storage of internal registers, so that the processor can be interrupted but can easily return to a given task.

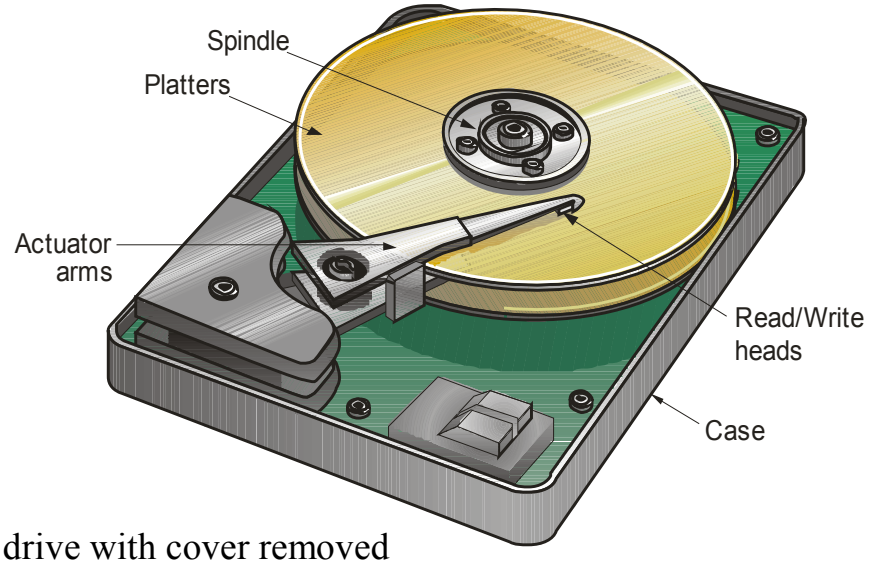
A special register, called the stack pointer, keeps track of the location that data was last stored on the stack. This will be the next data to be taken from the stack when needed.



Magnetic Hard Drive

The magnetic hard drive is the backbone of computer mass storage and is applied to other devices such as digital video recorders. Capacities of hard drives have increased exponentially, with 1 TB (1 trillion bytes!) drives available today.

Platters are arranged in tracks (circular shapes) and sectors (pie shaped). Files are listed in a File Allocation Table, (FAT) that keeps track of file names, locations, size, and more.



Hard drive with cover removed

Optical Storage

The compact disk (CD) uses a laser to burn tiny *pits* into the media. Surrounding the pits are flat areas called *lands*. The CD can be read using a low-power IR laser that detects the difference between pits and lands.

Binary data is encoded with a special method called negative non-return to zero encoding. A change from a pit to a land or a land to a pit represents a binary one, whereas no change represents a zero. A standard 120 mm CD can hold approximately 700 MB of data.

Selected Key Terms

Address	The location of a given storage cell or group of cells in memory.
Capacity	The total number of data units (bits, nibbles, bytes, words) that a memory can store.
SRAM	Static random access memory; a type of volatile read/write semiconductor memory.
DRAM	Dynamic random access memory; a type of read/write memory that uses capacitors as the storage elements and is a volatile read/write memory.
PROM	Programmable read-only memory; type of semiconductor memory.

Selected Key Terms

EPROM	Erasable programmable read-only memory; a type of semiconductor memory device that typically uses ultraviolet light to erase data.
Flash memory	A nonvolatile read/write random access semiconductor memory in which data are stored as charge on a floating gate of a certain type of FET.
FIFO	First in-first out memory.
LIFO	Last in-first out memory
Hard disk	A magnetic storage device; typically a stack of two or more rigid disks enclosed in a sealed housing.

1. Static RAM is

    a. nonvolatile read only memory

    b. nonvolatile read/write memory

    c. volatile read only memory

    d. volatile read/write memory

2. A nonvolatile memory is one that

    a. requires a clock

    b. must be refreshed regularly

    c. retains data without power applied

    d. all of the above

3. The advantage of dynamic RAM over static RAM is that

a. it is much faster

b. it does not require refreshing

c. it is simpler and cheaper

d. all of the above

4. The first step in a read or write operation for a random access memory is to

a. place a valid address on the address bus

b. enable the memory

c. send or obtain the data

d. start a refresh cycle

5. The output enable signal ($\overline{OE}$) on a RAM is active

   a. only during a write operation

   b. only during a read operation

   c. both of the above

   d. none of the above

6. When data is read from RAM, the memory location is

   a. cleared after the read operation

   b. set to all 1's after the read operation

   c. unchanged

   d. destroyed

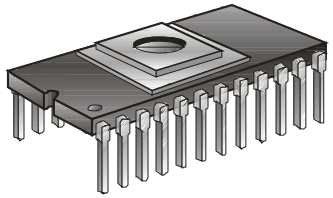
7. An EPROM has a window to allow UV light to enter under certain conditions. The purpose of this is to

a. refresh the data

b. read the data

c. program the IC

d. erase the data



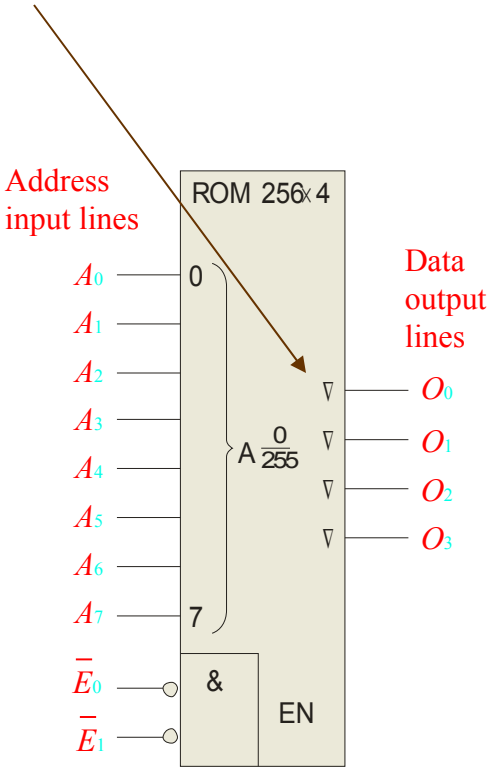
8. The small triangles on the logic diagram indicate that these outputs are

a. not used

b. tri-stated

c. inverted

d. grounded



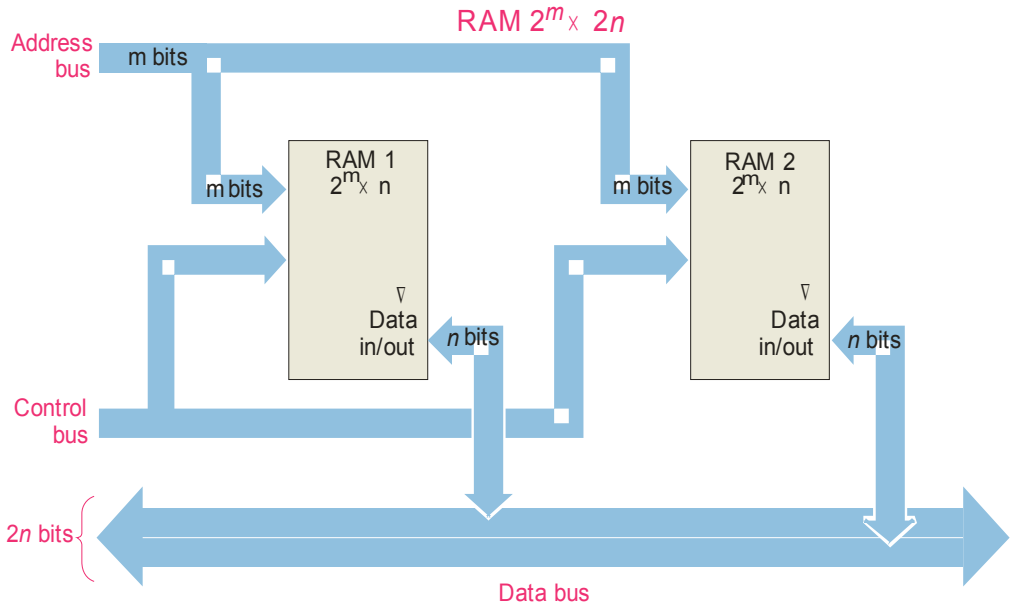
9. Using two ICs as shown will expand

a. the word size

b. the number of words available

c. both of the above

d. none of the above



10. On a hard drive. information about file names, locations, and file size are kept in a special location called the

a. file location list

b. file allocation table

c. disk directory

d. stack