

Fundamentos de diseño lógico y de computadoras

Tercera Edición

M. MORRIS MANO

California State University, Los Angeles

CHARLES R. KIME

University of Wisconsin, Madison

Traducción

José Antonio Herrera Camacho

Profesor Titular de Escuela Universitaria
Universidad Politécnica de Madrid

Martina Eckert

Dra. Ingeniería en Telecomunicación
Universidad Politécnica de Madrid

Beatriz Valcuende Lozano

Ingeniera Técnica en Telefonía y Transmisión de Datos
Universidad Politécnica de Madrid

Revisión técnica

José Antonio Herrera Camacho

Profesor Titular de Escuela Universitaria
Universidad Politécnica de Madrid



CAPÍTULO

9

MEMORIAS

La memoria es el componente más grande en una computadora digital y está presente en un gran porcentaje de los sistemas digitales. Las memorias de acceso aleatorio (RAM) almacenan datos temporalmente, y las memorias de sólo lectura (ROM) almacenan datos permanentemente. Una memoria ROM pertenece a un tipo de componentes llamados dispositivos lógicos programables (PLDs, del inglés *Programmable Logic Devices*) que utilizan la información almacenada para definir circuitos lógicos.

Nuestro estudio de las memorias RAM comienza viéndolas como un modelo con entradas, salidas y la temporización de sus señales. Usaremos, por tanto, modelos lógicos equivalentes para comprender el funcionamiento interno de las memorias RAM de los circuitos integrados. Se estudian las memorias RAM estáticas y dinámicas. También se estudian los distintos tipos de memorias RAM dinámicas usadas para el movimiento de datos a altas velocidades entre la CPU y la memoria. Finalmente juntaremos diversos chips de memoria RAM para construir la memoria de un sistema.

En algunos de los capítulos anteriores se utilizaron extensamente estos conceptos referentes a la computadora genérica al principio de Capítulo 1. En este capítulo, por primera vez vamos a ser más precisos y señalaremos los usos específicos de las memorias y sus componentes asociados. Empezando con el procesador, la caché interna es, básicamente, una memoria RAM muy rápida. Fuera de la CPU, la caché externa también es, básicamente, una memoria RAM muy rápida. El subsistema de memoria RAM, como su nombre indica, es un tipo de memoria. En la zona de entrada/salida, encontramos esencialmente memoria para almacenar información de la imagen de la pantalla en la tarjeta de vídeo. La memoria RAM aparece en la caché de disco en la tarjeta controladora del disco, acelerando los accesos a éste. Aparte del papel principal que tiene el subsistema de memoria RAM para almacenar datos y programas, encontramos memoria aplicada de varias formas en la mayoría de los subsistemas de una computadora genérica.

9-1 DEFINICIONES

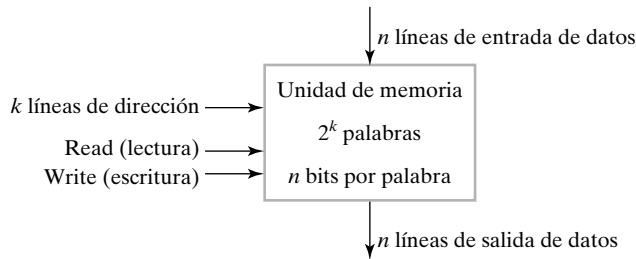
En los sistemas digitales, una memoria es una colección de celdas capaces de almacenar información binaria. Aparte de estas celdas, una memoria contiene circuitos electrónicos para almacenar y recuperar la información. Como se indicó en la explicación de la computadora genérica, la memoria se usa en diferentes partes de las computadoras modernas, proporcionando almacenamiento temporal o permanente para grandes cantidades de información binaria. Para que esta información sea procesada, se envía de la memoria al hardware de procesamiento, que está formado por registros y lógica combinacional. La información procesada se devuelve posteriormente a la misma memoria o a otra diferente. Los dispositivos de entrada y salida también interactúan con la memoria. La información de un dispositivo de entrada se coloca en la memoria, de forma que puede usarse en su procesamiento. La información procesada de salida se coloca en la memoria, y de allí se manda a un dispositivo de salida.

Se utilizan dos tipos de memoria en las diversas partes de una computadora: la *memoria de accesos aleatorio* (RAM, del inglés *random-access memory*) y la *memoria de sólo lectura* (ROM, del inglés *read-only memory*). La memoria RAM permite almacenar nueva información que estará disponible para su uso posteriormente. Al proceso de almacenamiento de nueva información en la memoria se le llama operación de *escritura* en la memoria. Al proceso de transferir la información almacenada en la memoria se le llama operación de *lectura* de la memoria. La memoria RAM puede realizar ambas operaciones, mientras que la memoria ROM, presentada en el Capítulo 3, sólo puede realizar operaciones de lectura. El tamaño de la memoria RAM puede variar entre cientos y millones de bits.

9-2 MEMORIA DE ACCESO ALEATORIO

Una memoria es una colección de celdas de almacenamiento binario junto con circuitos adicionales necesarios para transferir información de o desde un lugar determinado, con el mismo tiempo de acceso independientemente de dónde esté localizada, de aquí el nombre de *memoria de acceso aleatorio*. Por contra, la *memoria serie*, tal y como se produce en un disco magnético o en una unidad de cinta, necesita más o menos tiempo dependiendo de dónde este el dato deseado, puesto que depende de su localización física en el disco o en la cinta.

La información binaria se almacena en la memoria en grupos de bits, a cada grupo de bits se le llama *palabra* (del término inglés *word*). Una palabra es una entidad de bits que se mueve a dentro y a fuera de la memoria como una unidad, un grupo de unos y ceros que representan un número, una instrucción, uno o más caracteres alfanuméricos o cualquier otra información codificada. A un grupo de ocho bits se le llama *byte*. La mayoría de las computadoras usan palabras que son múltiplo de ocho bits. Así, una palabra de 16 bits contiene dos bytes y una palabra de 32 bits está formada por cuatro bytes. La capacidad de una unidad de memoria se expresa como el número total de bytes que puede almacenar. La comunicación entre la memoria y su entorno se consigue mediante: líneas de entradas y de salidas de datos, líneas de selección de dirección y las líneas de control que especifican la dirección de la transferencia de la información. En la Figura 9-1 se muestra un diagrama de bloques de una memoria. Las n líneas de entrada proporcionan la información a almacenar en la memoria y las n líneas de salida aportan la información que sale de la memoria. Las k líneas de dirección especifican la palabra escogida de entre las muchas disponibles. Las dos entradas de control especifican el sentido de la transferencia deseada: la entrada de escritura (Write) provoca que el dato binario se transfiera dentro la memoria, y la entrada de lectura (Read) hace que el dato binario se transfiera a fuera de la memoria.



□ **FIGURA 9-1**
Diagrama de bloques de una memoria

La unidad de memoria se caracteriza por el número de palabras que contiene y por el número de bits en cada palabra. Las líneas de dirección seleccionan una palabra concreta. A cada palabra de la memoria se le asigna un número de identificación llamada *dirección* (en término inglés *address*). El rango de direcciones va desde 0 hasta $2^k - 1$, donde k es el número de líneas de dirección. La selección de una determinada palabra dentro de la memoria se hace aplicando la dirección en binario de k bits a las líneas de dirección. Un decodificador toma esta dirección y abre los caminos necesarios para seleccionar la palabra especificada. La memoria de una computadora puede tener muchos tamaños. Es habitual referirse al número de palabras (o bytes) mediante los prefijos K (kilo), M (mega) o G (giga). K es igual a 2^{10} , M es igual a 2^{20} y G es igual a 2^{30} . De esta forma, $64\text{ K} = 2^{16}$, $2\text{ M} = 2^{21}$ y $4\text{ G} = 2^{32}$.

Considere, por ejemplo, una memoria con una capacidad de 1 K palabras de 16 bits cada una. Puesto que $1\text{ K} = 1024 = 2^{10}$, y 16 bits constituyen dos bytes, podemos decir que la memoria puede albergar 2048 o 2 K bytes. La Figura 9-2 muestra el posible contenido de las tres primeras y tres últimas palabras de la memoria de este tamaño. Cada palabra contiene 16 bits que se pueden dividir en dos bytes. Las palabras se reconocen por su dirección en decimal, desde 0 hasta 1023. Una dirección equivalente en binario tiene 10 bits. La primera dirección se especifica con 10 ceros y la última con 10 unos. Esto es así porque 1023 en binario es igual a 111111111. Una palabra de memoria se selecciona mediante su dirección en binario. Cuando una palabra se lee o se escribe, la memoria funciona con todos los 16 bits como una única unidad.

Direcciones de memoria

<u>Binario</u>	<u>Decimal</u>	Contenido de memoria
000000000	0	10110101 01011100
000000001	1	10101011 10001001
000000010	2	00001101 01000110
⋮	⋮	⋮
⋮	⋮	⋮
⋮	⋮	⋮
⋮	⋮	⋮
111111101	1021	10011101 00010101
111111110	1022	00001101 00011110
111111111	1023	11011110 00100100

□ **FIGURA 9-2**
Contenido de una memoria de 1024×16

La memoria de $1\text{ K} \times 16$ de la figura tiene 10 bits en las direcciones y 16 bits en cada palabra. Si tuviésemos una memoria de $64\text{ K} \times 10$, sería necesario incluir 16 bits en las direcciones y cada palabra tendría 10 bits. El número de bits necesarios en las direcciones depende del número total de palabras que pueden ser almacenadas y es independiente del número de bits en cada palabra. El número de bits en la dirección de una palabra se determina mediante la relación $2^k \geq m$, donde m es el número total de palabras y k es el número mínimo de bits de direcciones que satisface la relación.

Operaciones de lectura y escritura

Las dos operaciones que puede efectuar una memoria de acceso aleatoria son la escritura (*write*) y la lectura (*read*). Una *escritura* es una transferencia al interior de la memoria de un nuevo dato para ser almacenado. Una *lectura* es una transferencia de una copia de una palabra almacenada al exterior de la memoria. Una señal de escritura (Write) especifica la operación de entrada, y una señal de lectura (Read) determina la operación de salida. Aceptando una de estas señales de control, los circuitos internos de la memoria permiten realizar la función deseada.

Los pasos que se deben realizar para realizar una operación de escritura son los siguientes:

1. Aplicar la dirección binaria de la palabra elegida a las líneas de dirección.
2. Aplicar los bits de datos que se deben almacenar en la memoria a las líneas de entrada de datos.
3. Activar la entrada de escritura (Write).

La unidad de memoria tomará los bits de las líneas de entrada de datos y los almacenará en la palabra especificada en las líneas de direcciones.

Los pasos que se deben seguir para realizar una operación de lectura son los siguientes:

1. Aplicar la dirección binaria de la palabra elegida a las líneas de dirección.
2. Activar la entrada de lectura (Read)

La memoria tomará los bits de la palabra que ha sido seleccionada por la dirección y los llevará a las líneas de salida de datos. El contenido de la palabra seleccionada no se cambiará con su lectura.

La memoria RAM se construye con circuitos integrados (chips) más circuitos lógicos adicionales. Habitualmente, los chips de memoria RAM tienen dos entradas de control para las operaciones de lectura y escritura en una configuración diferente a la descrita anteriormente. En lugar de tener dos entradas separadas para la lectura y la escritura, la mayoría de circuitos tiene, al menos, una entrada de selección de chip (*Chip Select*) que selecciona el chip que se va a leer o a escribir, y una entrada de lectura/escritura ($\text{Read}/\overline{\text{Write}}$) que determina la operación a realizar. Las operaciones de la memoria que resultan de esta configuración de las señales de control se muestran en la Tabla 9-1.

La selección del chip (en adelante *Chip Select*) se usa para habilitar uno o varios chips que forman la RAM y que contienen la palabra a la que se quiere acceder. Si el *Chip Select* se activa, la entrada $\text{Read}/\overline{\text{Write}}$ determina la operación a realizar. Mientras la señal de *Chip Select* accede a los chips, también se proporciona una señal que accede a toda la memoria. Llamaremos a esta señal *Habilitación de Memoria (Memory Enable)*.

□ **TABLA 9-1**
Entradas de control de un chip de memoria

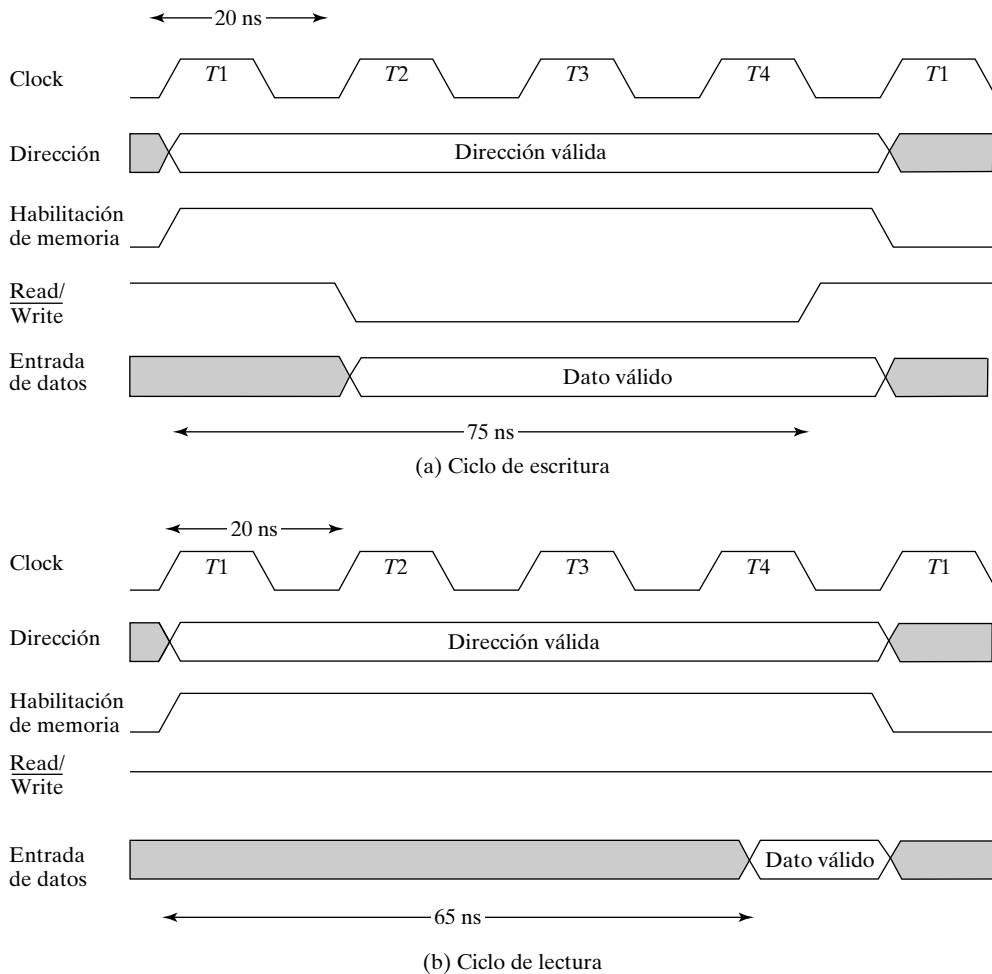
Chip select CS	Read/ $\overline{\text{Write}}$ R/\overline{W}	Operación de la memoria
0	×	Ninguna
1	0	Escriba la palabra seleccionada
1	1	Lee de la palabra seleccionada

Temporización de las formas de onda

El funcionamiento de la unidad de memoria se controla mediante un dispositivo externo, como por ejemplo, una CPU. La CPU está sincronizada con su señal de reloj. Sin embargo, la memoria no emplea este reloj. En cambio, sus operaciones de lectura y escritura se temporizan mediante cambios en sus entradas de control. El *tiempo de acceso* de una operación de lectura es el tiempo máximo que transcurre desde la aplicación de la dirección hasta que aparece la información en la salida de datos. De forma similar, el *tiempo del ciclo de escritura* es el tiempo máximo que transcurre desde que se pone la dirección hasta completar todas las operaciones internas que necesita la memoria para almacenar una palabra. Las escrituras en la memoria se pueden llevar a cabo una detrás de otra en intervalos de tiempo. La CPU debe proporcionar a la memoria señales de control de tal forma que se sincronizan sus propias operaciones internas, sincronizadas con el reloj, con las operaciones de lectura y escritura de la memoria. Esto significa que el tiempo de acceso y el tiempo del ciclo de escritura de la memoria deben estar relacionados con la CPU con un periodo igual a un número fijo de ciclos de reloj de la CPU.

Supongamos, como ejemplo, que una CPU funciona con un reloj de 50 MHz de frecuencia, que tiene un periodo de reloj de 20 ns ($1 \text{ ns} = 10^{-9} \text{ s}$). Supongamos ahora que la CPU se comunica con una memoria con un tiempo de acceso de 65 ns y un tiempo ciclo de escritura de 75 ns. El número de ciclos de reloj necesarios para una petición de memoria es un valor entero mayor o igual que el valor más grande del tiempo de acceso y del tiempo del ciclo de escritura, dividido por el periodo del reloj. Como el periodo del reloj es de 20 ns y el tiempo mayor entre el tiempo de acceso y el tiempo del ciclo de escritura es 75 ns, será necesario emplear, al menos, cuatro ciclo de reloj para cada petición a la memoria.

La temporización del ciclo de memoria se muestra en la Figura 9-3, para una CPU funcionando a 50 MHz y una memoria con un ciclo de escritura de 75 ns y un tiempo de acceso de 65 ns. El ciclo de lectura de la parte (a) muestra 4 pulsos $T1$, $T2$, $T3$ y $T4$ con un periodo de 20 ns. En una operación de escritura, la CPU debe proporcionar la dirección y el dato de entrada a la memoria. Se aplica la dirección y la habilitación de memoria se cambia a 1 en el flanco de subida del pulso $T1$. El dato, que es necesario algo más tarde, se aplica en el flanco de subida de $T2$. Las dos líneas que se cruzan una con otra en las forma de ondas de las direcciones y de los datos indica un posible cambio en el valor de estos buses. Las zonas sombreadas representan valores sin especificar. Un cambio de la señal $\text{Read}/\overline{\text{Write}}$ a 0 en el flanco positivo de $T2$ indica la operación de escritura. Para evitar la destrucción de los datos, en otras palabras de la memoria, es importante que este cambio ocurra después de que las señales en el bus de direcciones estén estables y con el valor de la dirección deseada. De lo contrario, una o más palabras pueden ser direccionadas momentáneamente y ser sobrescritas con diferentes datos. La señal debe



□ FIGURA 9-3
Formas de onda del ciclo de memoria

la habilitación de memoria para poder completar la operación de escritura. Finalmente, la dirección y el dato deben permanecer estables un poco después de que la señal *Read/Write* cambie a 1 de nuevo para evitar destruir los datos en otras palabras de la memoria. Después del cuarto pulso de reloj, la operación de escritura ha terminado con 5 ns de sobra, y la CPU puede poner la dirección y las señales para hacer otra petición de memoria en el siguiente pulso *T1*.

El ciclo de lectura de la Figura 9-3(b) tiene una dirección para la memoria que proporciona la CPU. La CPU pone la dirección, cambia la habilitación de memoria a 1 y la señal *Read/Write* a 1 para seleccionar la operación de lectura, todo esto en el flanco de subida de *T1*. La memoria coloca el dato de la palabra seleccionada, mediante la dirección en el bus de salida de datos, dentro de los 65 ns a partir de que se pone la dirección y se activa la habilitación de la memoria. Luego, la CPU transfiere el dato a uno de sus registros internos en el flanco de subida del siguiente pulso *T1*, en el que también se puede cambiar la dirección y las señales de control para la siguiente petición de memoria.

Características de las memorias

Las memorias en los circuitos integrados pueden ser estáticas o dinámicas. Las Memorias *Estáticas* (SRAM, del inglés *Static RAM*) están formadas con *latches* internos que almacenan la información binaria. La información permanece almacenada correctamente mientras que la memoria RAM esté alimentada. Las Memorias *Dinámicas* (DRAM, del inglés *Dynamic RAM*) almacenan la información binaria mediante cargas eléctricas en condensadores. Estos condensadores se fabrican dentro del chip utilizando transistores MOS de canal n. La carga almacenada en el condensador tiende a descargarse con el tiempo, por lo que los condensadores deben ser recargados periódicamente mediante el *refresco* de la memoria DRAM. Esto se hace cíclicamente en todas las palabras cada pocos milisegundos, leyendo y rescribiéndolos para restablecer la carga perdida. Las memorias DRAM de los chips son de bajo consumo y de gran capacidad de almacenamiento pero las memorias SRAM son más fáciles de usar y tienen unos ciclos de lectura y escritura más cortos, y además, no necesitan un ciclo de refresco.

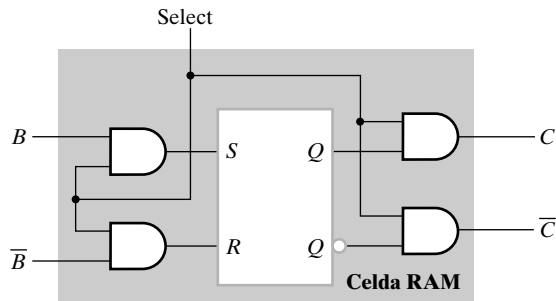
Las unidades de memorias que pierden la información cuando se apaga la alimentación se dicen que son *volátiles*. Las RAM de los circuitos integrados, tanto estáticas como dinámicas, pertenecen a esta categoría puesto que sus celdas necesitan una fuente de alimentación externa para mantener la información almacenada. Por el contrario, una *memoria no volátil*, como son los discos magnéticos, mantienen su información después de desconectar la alimentación. Esto es así porque los datos almacenados en soporte magnético se representan mediante la orientación de la magnetización, que no se pierde al desconectar la fuente de alimentación. Las memorias ROM son también memorias no volátiles, como se vio en la Sección 3-9.

9-3 MEMORIAS INTEGRADAS SRAM

Como se indicó anteriormente, las memorias están compuestas por circuitos integrados con memoria RAM más lógica adicional. Primero veremos la estructura interna de una memoria integrada RAM y luego estudiaremos las combinaciones de los chips de memoria RAM y la lógica adicional usada para construir una memoria. La estructura interna de un chip de memoria RAM de m palabras de n bits por palabra está compuesta de un array de $m \cdot n$ celdas de almacenamiento binario y una circuitería asociada. El circuito se construye con un decodificador para seleccionar la palabra que se va a leer o a escribir, circuitos de lectura, circuitos de escritura y lógica de salida. La celda de una memoria RAM es la célula de almacenamiento básica usada en un chip de una memoria RAM, que se diseña típicamente como un circuito electrónico en lugar de un circuito lógico. Sin embargo, es posible y conveniente modelar la memoria RAM de un chip usando un modelo lógico.

Usaremos un chip de memoria RAM estática como base para nuestra explicación. Primero presentamos la lógica de una célula RAM que almacena un solo bit y posteriormente usamos la célula jerárquicamente para describir un chip de memoria RAM. La Figura 9-4 muestra el modelo lógico de una celda de memoria RAM. La parte de almacenamiento de la celda se modela con un latch *SR*. Las entradas del latch se habilitan con la señal de selección, *Select*. Para *Select* igual a 0, el valor almacenado se retiene. Para *Select* igual a 1, el valor almacenado se determina mediante los valores de B y \bar{B} . Las salidas del latch se habilitan con *Select*, usando una puerta AND, para generar las salidas de la celda C y \bar{C} . Para *Select* igual 0, ambas salidas son 0 y para *Select* igual a 1, C tiene el valor almacenado y \bar{C} tiene su complemento.

Para obtener un diagrama interconectamos un conjunto de celdas de la memoria RAM y los circuitos de lectura y escritura para construir una tira de un bit de una memoria RAM que



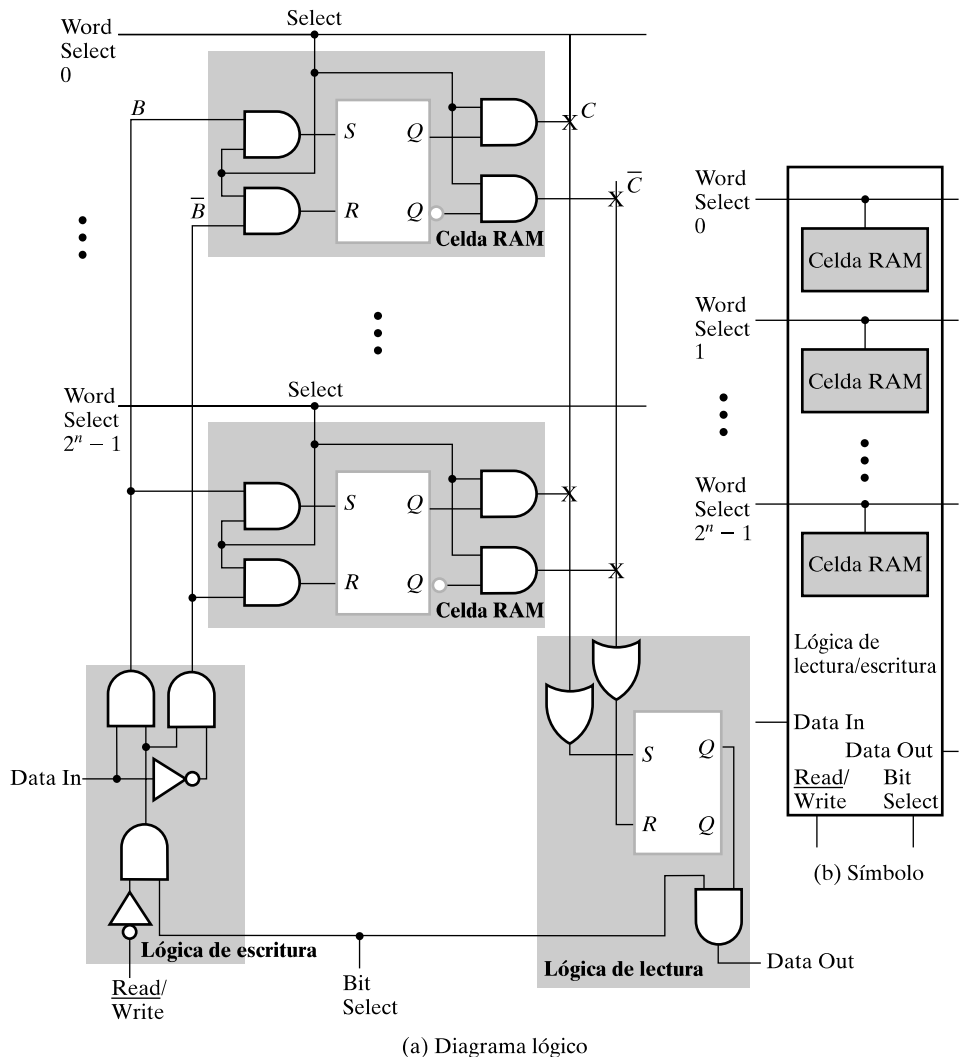
□ FIGURA 9-4
Celda de una memoria RAM estática

contiene todos los circuitos asociados con una posición de un bit de un conjunto de palabras de la memoria RAM. En la Figura 9-5 se muestra el diagrama lógico de una tira de un bit de memoria RAM. La parte del modelo que representa a cada celda de la memoria RAM se resalta en azul. La carga del latch de la célula se controla ahora con una entrada de selección de palabra, *Word Select*. Si ésta es 0, *S* y *R* son 0 y el contenido del latch de la célula permanece sin variar. Si la entrada *Word Select* es 1, entonces el valor a cargar en el latch se controla con las dos señales *B* y que parten de la lógica de escritura. Para que cualquiera de estas señales sea 1 y puedan cambiar el valor almacenado, *Read/Write* debe ser 0 y la señal *Bit Select* debe ser 1. Entonces, el dato de entrada, *Data_In*, y su complemento se aplica a *B* y \bar{B} , respectivamente, para poner a uno o a cero el latch de la celda de la memoria RAM seleccionada. Si *Data_In* es 1, el *latch* se pone a 1 y si es 0 se pone a 0, completando la operación de escritura.

Solamente se escribe una palabra cada vez. Es decir, sólo una línea *Word Select* es 1 y todas las demás son 0. Así, sólo la celda de la memoria RAM conectada a *B* y \bar{B} se escribe. La señal *Word Select* también controla la lectura de las celdas de la memoria RAM usando la lógica compartida de la escritura. Si *Word Select* es 0, el valor almacenado en el *latch SR* se bloquea mediante las puertas AND para que no alcance a las dos puertas OR de la lógica de lectura. Pero si *Word Select* es 1, el valor almacenado pasa a través de las puertas OR y se captura en el *latch SR* de la lógica de lectura. Si la señal *Bit Select* también es 1, el valor capturado también aparece en la línea *Data Out* de la tira de un bit de la memoria RAM. Véase que, para el diseño de esta lógica de lectura en particular, la lectura ocurre independientemente del valor de la señal *Read/Write*.

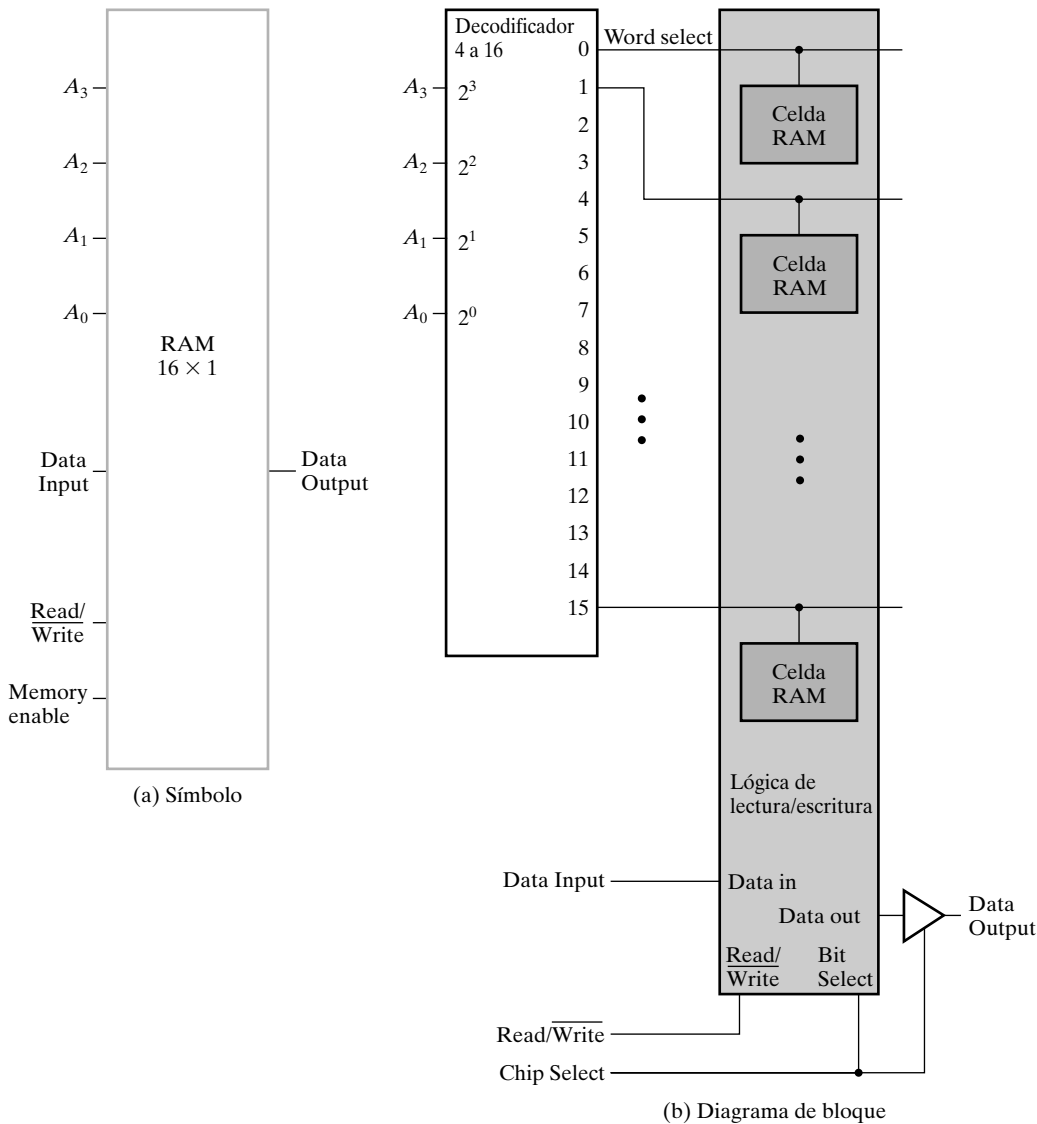
El símbolo de la tira de un bit de una memoria RAM, que se da en la Figura 9-5(b), se utiliza para representar la estructura interna de un chip de memoria RAM. Cada línea de selección se extiende más allá de la tira de un bit de forma que cuando varias tiras de un bit se colocan una junto a otra, se conectan las correspondientes líneas de selección. Las otras señales de la parte baja del símbolo se pueden conectar de diversas formas dependiendo de la estructura del chip de memoria RAM.

En la Figura 9-6 se muestra el símbolo y el diagrama de bloques de un chip de una memoria RAM de 16×1 . Ambos tienen 4 entradas de direcciones para las 16 palabras de un bit almacenadas en la memoria RAM. La habilitación del chip, *Chip Select*, a nivel del chip se corresponde con la habilitación de memoria, *Memory Enable*, a nivel de la memoria RAM formada por varios chips. La estructura interna de la RAM está compuesta por una tira de un bit de memoria RAM que tiene 16 celdas. Como hay 16 líneas de selección de palabra que controlar, de forma que una y solo una tienen un 1 lógico en un instante dado, se utiliza un decodificador de 4 a 16 líneas para descodificar los cuatro bits de dirección a los 16 bits de la palabra de selección.



□ FIGURA 9-5
Modelo de una tira de un bit de la memoria RAM

La única lógica adicional en la figura es un triángulo con una entrada normal, una salida normal y una segunda entrada debajo del símbolo. Este símbolo es un buffer triestado que permite la construcción de un multiplexor con un número arbitrario de entradas. Las salidas triestado se conectan juntas y se controlan adecuadamente usando las entradas de Chip Select. Con el uso de los buffers tri-estado en las salidas de la memoria RAM, dichas salidas pueden unirse para sacar la palabra del chip cuando éste se lee de las líneas de salidas conectadas a las salidas de la RAM. Las señales de habilitación anteriormente comentadas se corresponden con las entradas de Chip Select de los chips de la memoria RAM. Para leer una palabra un chip de memoria RAM concreto, el valor de Chip Select para ese chip debe ser 1 y para los chips restantes, que están conectados a las mismas líneas de salida, el Chip Select deben ser 0. Estas combinaciones que contienen un único 1 se pueden obtener de un decodificador.



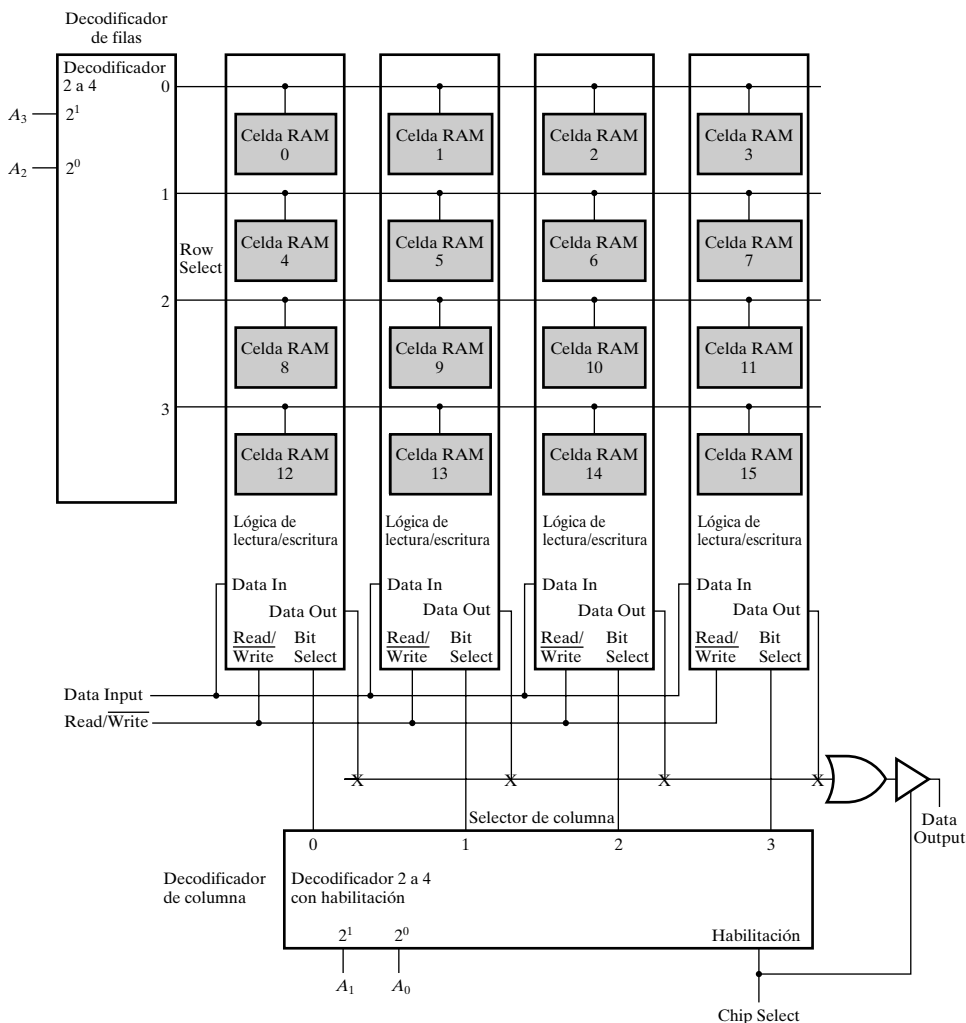
□ FIGURA 9-6
Chip de memoria RAM de 16 palabras de 1 bit.

Selección combinada

Dentro de un chip de memoria RAM, el decodificador de k entradas y 2^k salidas necesita 2^k puertas AND con k entradas por puerta si se utiliza un método de diseño directo. Además, si el número de palabras es grande, y todos los bits correspondientes a una posición de una palabra está en una única tira de un bit de una memoria RAM, el número de celdas RAM compartiendo los circuitos de lectura y escritura también es grande. Las características eléctricas resultantes de estas dos situaciones hacen que los tiempos de acceso de lectura y escritura de la memoria RAM sean largos, lo que cual es indeseable.

El número total de puertas del decodificador, el número de entradas por puerta y el número de celdas de memoria RAM por tira de un bit pueden reducirse empleando dos decodificadores empleando un método de *selección combinada*. Una posible configuración es utilizar dos decodificadores de $k/2$ entradas en lugar de uno de k entradas. Un decodificador controla las líneas de selección de palabra y el otro controla las líneas de selección de bit. El resultado es un método de selección con una matriz bidimensional. Si el chip de memoria RAM tiene m palabras de 1 bit por palabra, el método selecciona la célula de la memoria que está en la intersección de la fila de selección de palabra y la columna de selección de bit. Como *Word Select* no selecciona estrictamente palabras, su nombre se cambia a *Row Select* (selección de fila). A las salidas del decodificador añadido, que selecciona una o más tiras de bits se les llamará *Column Select* (selección de columna).

En la Figura 9-7 se muestra el método de selección combinada para un chip de memoria RAM. El chip esta formado por cuatro tiras de bits con cuatro bits en cada una y tiene un



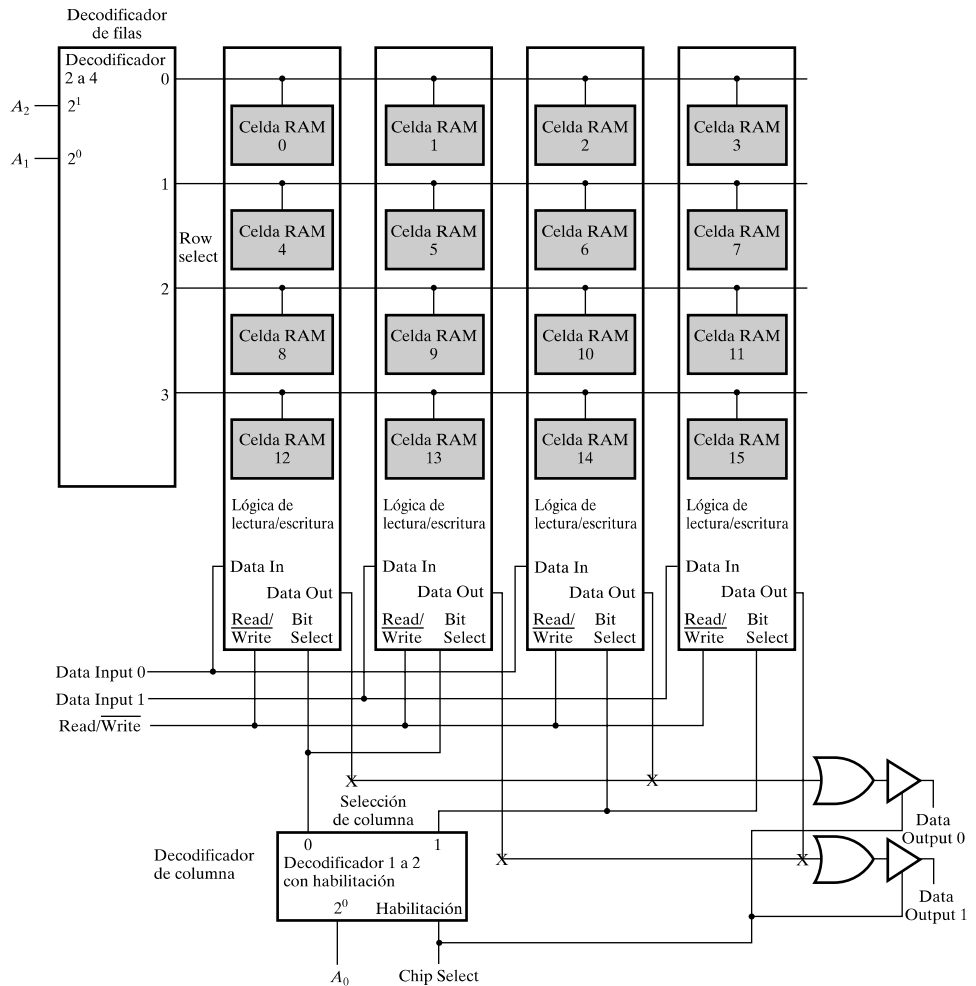
□ FIGURA 9-7

Diagrama de una memoria RAM de 16×1 utilizando un array de 4×4 celdas de memoria RAM

total de 16 celdas de memoria RAM en un array de dos dimensiones. Los dos bits más significativos del bus de direcciones van al decodificador de 2 a 4 líneas de las filas para seleccionar una de las cuatro filas del array. Las dos líneas menos significativas del bus de direcciones van al decodificador de 2 a 4 líneas de las columnas para seleccionar una de las cuatro columnas (tiras de un bit) del array. El decodificador de las columnas se habilita con la entrada de Chip Select. Cuando el Chip Select es 0, todas las salidas del decodificador están a 0 y no se selecciona ninguna de las celdas. Esto evita la escritura en cualquier celda del array de la memoria RAM. Cuando Chip Select es 1, se accede a un solo bit de la memoria. Por ejemplo, para la dirección 1001, los dos primeros bits de direcciones se decodifican para seleccionar la fila 10 (2_{10}) del array de celdas de la memoria. Los otros dos bits de direcciones se decodifican para seleccionar la columna 01 (1_{10}) del array. La celda de la memoria a la que se accede, en la fila 2 y columna 1 del array, es la celda 9 ($10_2 01_2$). Una vez seleccionada la fila y la columna, la entrada Read/Write determina la operación a realizar en la memoria. Durante la operación de lectura (Read/Write = 1), el bit seleccionado de la columna seleccionada pasa por la puerta OR al buffer triestado. Nótese que la puerta se dibuja de acuerdo con el array lógico presentado en la Figura 3-22. Como el buffer se habilita con la señal Chip Select, el valor leído aparece en la salida de datos, Data Output. En la operación de escritura (Read/Write = 0), el bit disponible en la línea de entrada de datos, Data Input, el bit disponible en la línea de entrada de datos, Data Input, se transfiere a la celda seleccionada de la memoria. El resto de las celdas no seleccionadas de la memoria están deshabilitadas y sus valores almacenados permanecen sin cambiar.

El mismo array de celdas se usa en la Figura 9-8 para construir una memoria RAM de 8×2 (ocho palabras de 2 bits). La decodificación de las filas no cambia con respecto de la Figura 9-7; los únicos cambios están en las columnas y en la lógica de salida. Al tener solo tres bits de direcciones y utilizar dos de ellos para el decodificador de filas, el decodificador de las columnas utiliza como entradas al bit de direcciones restante y la entrada Chip Select para generar dos líneas de selección de columna. Puesto que se quiere leer o escribir dos bits simultáneamente, las líneas de selección de columna van conectadas a los pares de tiras de un bit adyacentes. Las dos líneas de entrada, Data Input 0 y Data Input 1, van cada una a diferentes bits de cada uno de los pares. Finalmente, los bits correspondientes de cada par comparten la salida de las puertas OR y de los buffers triestado, dando lugar a las líneas de salida Data Output 0 y Data Output 1. La forma de operar de esta estructura se puede ilustrar mediante la aplicación de la dirección 3 (011_2). Los primeros dos bits de la dirección, 01, seleccionan la fila 1 del array. El último bit, 1, selecciona la columna 1, compuesta por las tiras 2 (10_2) y 3 (11_2) de un bit. De esta forma, la palabra a escribir o a leer está en las celdas 6 y 7 de la memoria ($011 0_2$ y $011 1_2$), que contienen a los bits 0 y 1, respectivamente, de la palabra 3.

Podemos demostrar el ahorro de la selección combinada considerando una memoria RAM estática de un tamaño real, $32 \text{ K} \times 8$. Esta memoria contiene un total de 256 K bits. Haciendo el número de columnas del array igual al de filas, calculamos la raíz cuadrada de 256 K, que da un resultado de $512 = 2^9$. De esta forma, los primeros nueve bits de dirección se conectan al decodificador de filas y los seis restantes al decodificador de columnas. Sin selección combinada, el único decodificador tendría 15 entradas y 32.768 salidas. Con selección combinada, hay un decodificador de 9 a 512 líneas y uno de 6 a 64 líneas. El número de puertas en el diseño con un solo decodificador sería de 32 800. En el caso de los dos decodificadores combinados, el número de puertas es 608, reduciéndose el número de puertas por un factor mayor que 50. Además, aunque parezca que hay 64 veces tantos circuitos de lectura/escritura, la selección de columna se puede hacer entre las celdas de la memoria RAM y los circuitos de lectura/escritura, así que sólo se necesitan los ocho circuitos originales. Debido al reducido número de celdas de memoria conectadas a cada circuito de lectura/escritura, los tiempos de acceso del chip también se mejoran.



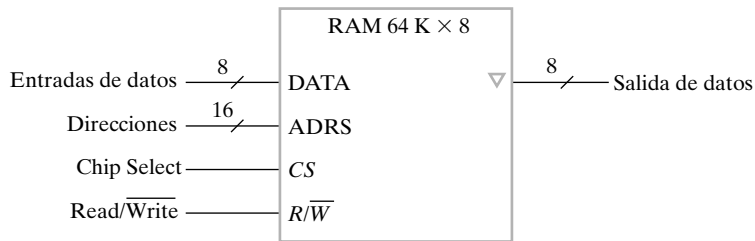
□ FIGURA 9-8

Diagrama de bloques de una memoria RAM de 8×21 utilizando un array de 4×4 celdas de memoria RAM

9-4 ARRAY DE CIRCUITOS INTEGRADOS DE MEMORIA SRAM

Los circuitos integrados de memoria RAM están disponibles en diversos tamaños. Si la unidad de memoria necesaria en una aplicación determinada es mayor que la capacidad de un chip, es necesario combinar un número de circuitos integrados en un array para construir la memoria con el tamaño requerido. La capacidad de la memoria depende de dos parámetros: el número de palabras y el número de bits por palabra. Un incremento en el número de palabras obliga a que incrementemos la longitud de la dirección. Cada bit añadido a la longitud de la dirección dobla el número de palabras en la memoria. Un incremento en el número de bits por palabra obliga a que incrementemos el número de líneas de entrada y salida de datos pero la longitud de la dirección permanece constante.

Para mostrar un array de circuitos integrados de memoria RAM vamos a utilizar un chip de memoria RAM, usando una representación condensada sus entradas y salidas, según se muestra en la Figura 9-9. La capacidad del chip es de 64 K palabras de 8 bits cada una. El integrado



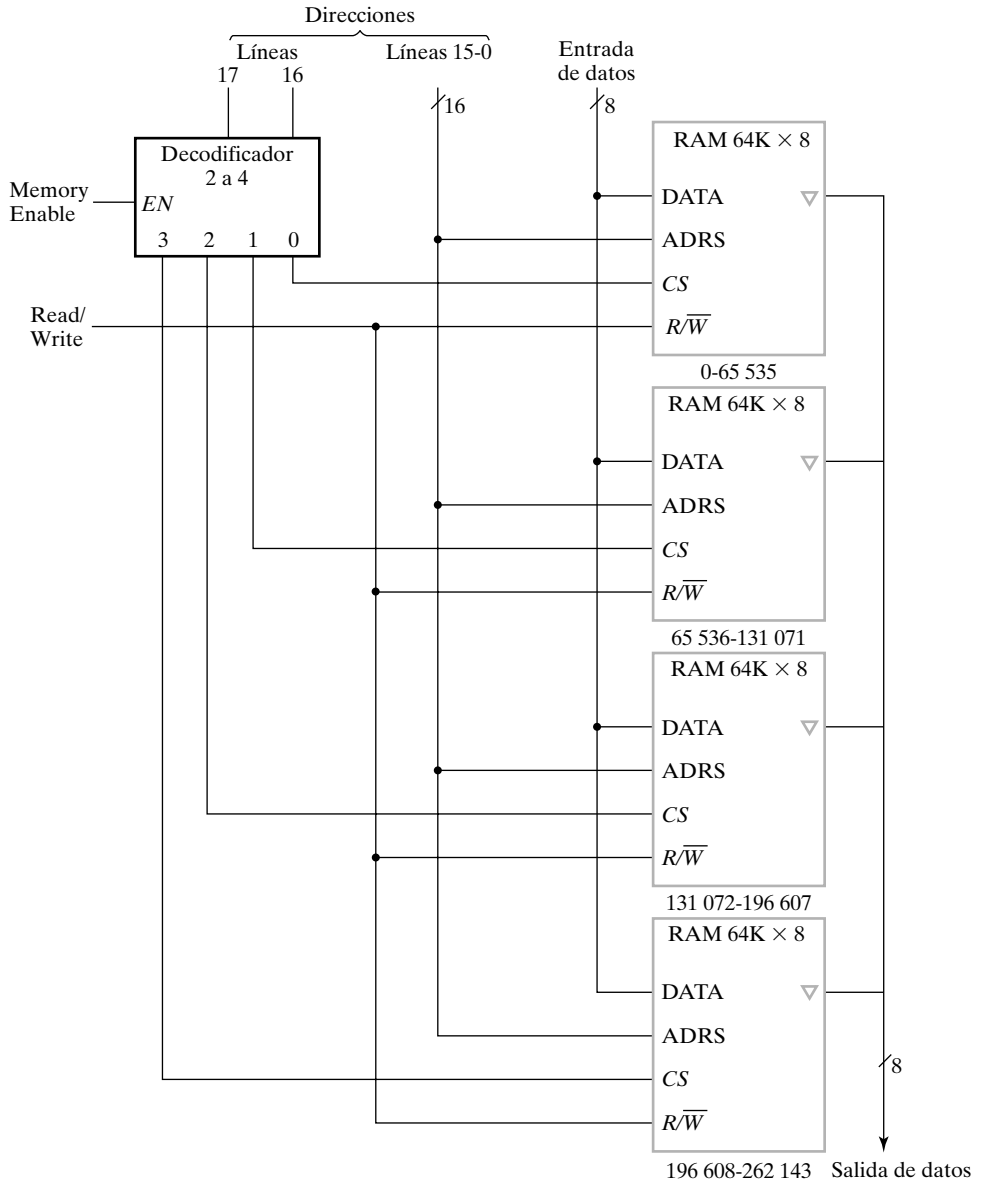
□ FIGURA 9-9
Símbolo de una memoria RAM de 64×8

necesita, por tanto, 16 líneas de direcciones, 8 líneas para la entrada y 8 líneas para la salida. En lugar de usar 16 líneas para las direcciones y 8 para la entrada y la salida, cada conjunto de líneas se mostrará en el diagrama de bloques como una línea simple. Cada línea está cruzada por una línea inclinada con un número que indica el número de líneas representadas en cada bus. La entrada CS (Chip Select) selecciona a un chip concreto de la memoria RAM y la entrada R/\overline{W} (Read/ \overline{Write}) especifica la operación de lectura o escritura cuando el chip ha sido seleccionado. El triángulo pequeño a la salida es el símbolo estándar para representar las salidas tri-estado. La entrada CS de la memoria RAM controla el comportamiento de las líneas de salida. Cuando $CS = 0$, el chip no ha sido seleccionado y todas las líneas de salida están en estado de alta impedancia. Cuando $CS = 1$, las líneas de salida de datos llevan los ocho bits de la palabra seleccionada.

Suponga que queremos incrementar el número de palabras en la memoria usando dos o más chips de memoria RAM. Puesto que por cada bit que se añade a la dirección se dobla el número binario que se puede formar, la forma natural de incrementar el número de palabras es por un factor de dos. Por ejemplo, dos chips de memoria doblarán el número de palabras y se añadirá un bit más para componer la dirección. Cuatro chips de memoria multiplican el número de palabras por cuatro y se añaden dos bits más para componer la dirección.

Considere la posibilidad de construir una memoria RAM de $256 K \times 8$ con cuatro chips de memoria RAM de $64 K \times 8$, como se muestra en la Figura 9-10. Las ocho líneas de datos llegan a todos los chips. Las salidas tri-estado se pueden conectar para formar un bus de salida de datos común. Este tipo de conexión de salida sólo es posible con salidas tri-estado. En cualquier instante, solamente se activará una entrada de selección de un chip, mientras que los restantes estarán deshabilitados. Las ocho salidas del chip seleccionado tendrán unos y ceros y las salidas de los otros tres estarán en estado de alta impedancia, presentándose sólo como circuitos abiertos a las señales de salida del circuito seleccionado.

La memoria de 256 K palabras necesita un bus de direcciones de 18 bits. Los 16 bits menos significativos se conectan a las entradas de direcciones de los cuatro chips. Los dos bits más significativos se llevan a las entradas de un decodificador de 2 a 4 líneas. Las cuatro salidas del decodificador se aplican a las entradas CS de los cuatro chips. La memoria se deshabilita cuando la entrada EN del decodificador, *Memory enable*, es igual a 0. Las cuatro salidas del decodificador son 0 y ningún chip está siendo seleccionado. Cuando el decodificador está habilitado, los bits de direcciones 17 y 16 determinan cuál de los cuatro chips ha sido seleccionado. Si estos bits son igual a 00, el chip de memoria seleccionado es el primero. El resto de los 16 bits de direcciones seleccionan entonces una palabra dentro del chip en el rango de 0 a 65 535. Las siguientes 65 535 palabras se seleccionan del segundo chip de memoria con una dirección de 18 bits que empieza con 01 seguido de los 16 bits restantes de las líneas comunes del bus de direcciones. El rango de direcciones para cada chip se enumera en decimal debajo de su símbolo.

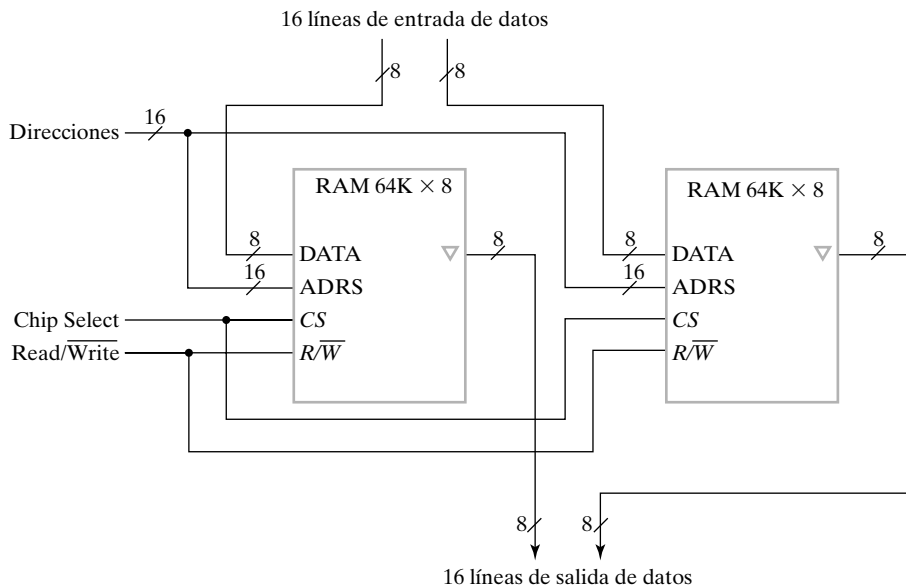


□ FIGURA 9-10

Diagrama de bloques de una memoria RAM de 256 K × 8

También es posible combinar dos chips para formar una memoria compuesta que contenga el mismo número de palabras pero con el doble de bits en cada palabra. En la Figura 9-11 se muestra la interconexión de dos chips para formar una memoria de 64 K × 16. Las 16 líneas de entrada y de salida se dividen entre los chips. Ambos reciben los 16 bits de direcciones y las entradas comunes de control CS y R/\bar{W} .

Las dos técnicas que se acaban de describir pueden combinarse para montar un array de chips idénticos para formar una memoria de gran capacidad. La memoria compuesta tendrá un número de bits por palabra que será múltiplo del número de bits por palabra de cada chip.



□ FIGURA 9-11
Diagrama de bloques de una memoria RAM de 64 K × 16

El número total de palabras se incrementará por un factor de dos veces la capacidad de palabras de un chip. Se necesita utilizar un decodificador externo para seleccionar a los chips individualmente según los bits adicionales de la memoria compuesta. Para reducir el número de pines del encapsulado del chip, muchos circuitos integrados tienen terminales comunes para la entrada y salida de datos. En este caso se dice que los terminales son bidireccionales, que quiere decir que para la operación de lectura éstos actúan como salidas y para la operación de escritura funcionan como entradas. Las líneas bidireccionales se construyen con buffers triestado, que ya se explicaron en la Sección 2-8. El uso de señales bidireccionales necesitan el control de los buffers triestado mediante las señales Chip Select y Read/Write.

9-5 CIRCUITOS INTEGRADOS DE MEMORIA DRAM

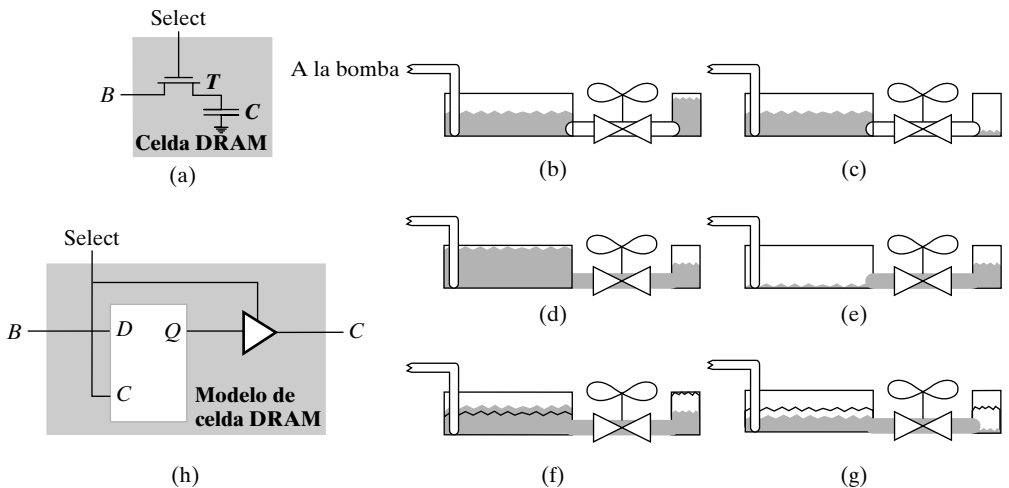
Las memorias RAM dinámicas (DRAM) dominan las aplicaciones con memorias de alta capacidad, incluyendo la memoria principal de las computadoras, debido a que proporcionan alta capacidad a bajo coste. Lógicamente, en muchos aspectos la memoria DRAM es similar a la memoria SRAM. Sin embargo, debido a los circuitos electrónicos utilizados para fabricar las celdas de almacenamiento, su diseño electrónico es considerablemente más desafiante. Además, como el nombre «dinámico» implica, el almacenamiento de la información es inherentemente temporal. Como consecuencia, la información debe «refrescarse» periódicamente para imitar el comportamiento del almacenamiento estático. Esta necesidad de refresco es la principal diferencia lógica en el comportamiento de una memoria DRAM en comparación con la SRAM. Exploraremos esta diferencia lógica examinando la celda de una memoria RAM dinámica, la lógica requerida para realizar la operación de refresco y el impacto de esta operación en el funcionamiento de la memoria del sistema.

Celda DRAM

En la Figura 9-12(a) se muestra la celda de una memoria RAM dinámica. Consiste en un condensador C y un transistor T . El condensador se usa para almacenar carga eléctrica. Si hay suficiente carga almacenada en el condensador, se puede considerar que hay almacenado un 1 lógico. Si la carga almacenada en el condensador es insuficiente, se puede considerar que hay almacenado un 0 lógico. El transistor actúa de forma muy parecida a un conmutador, de la misma forma que la puerta de transmisión presentada en el Capítulo 2. Cuando el conmutador está «abierto», la carga del condensador permanece fija aproximadamente, en otras palabras, está almacenada. Pero cuando el conmutador está «cerrado», la carga puede fluir a dentro y a fuera del condensador a través de la línea externa (B). Este flujo de carga permite escribir en la celda un 1 o un 0 y ser leído.

Para comprender las operaciones de lectura y escritura de la celda utilizaremos una analogía hidráulica con agua en lugar de carga, con un pequeño depósito en lugar del condensador y una válvula en lugar de un transistor. La línea B tiene una gran capacidad y se representa por un depósito grande y una bomba que puede llenar y vaciar el depósito rápidamente. En las Figuras 9-12(b) y 9-12(c) se representa esta analogía con la válvula cerrada. Nótese que en un caso el depósito pequeño está lleno, representando un 1 y en el otro caso está vacío, representando un 0. Suponga que se va a escribir un 1 en la celda. La válvula se abre y la bomba llena el depósito grande. El agua fluye a través de la válvula llenando el depósito pequeño, como se muestra en la Figura 9-12(d). Luego se cierra la válvula dejando el depósito pequeño lleno, lo cual representa un 1. Se puede escribir un 0 utilizando el mismo tipo de procedimiento con la excepción de que la bomba vacía el depósito grande, como se muestra en la Figura 9-12(e).

Suponga ahora que queremos leer un valor almacenado y ese valor es 1, que se corresponde con un depósito lleno. Con el depósito grande a un nivel conocido intermedio, se abre la válvula. Puesto que el depósito pequeño está lleno, el agua fluye del depósito pequeño al grande, incrementando su nivel de agua ligeramente, como se muestra en la Figura 9-12(f). Este incremento en el nivel se observa como la lectura de un 1 del almacenamiento del depósito. Corres-



□ FIGURA 9-12

Celda de memoria RAM dinámica, analogía hidráulica del funcionamiento de la celda y modelo de la celda

pondientemente, si el depósito de almacenamiento está inicialmente vacío, habrá un leve decremento en el nivel del depósito grande de la Figura 9-12(g), el cual se observa como la lectura de un 0 del depósito de almacenamiento.

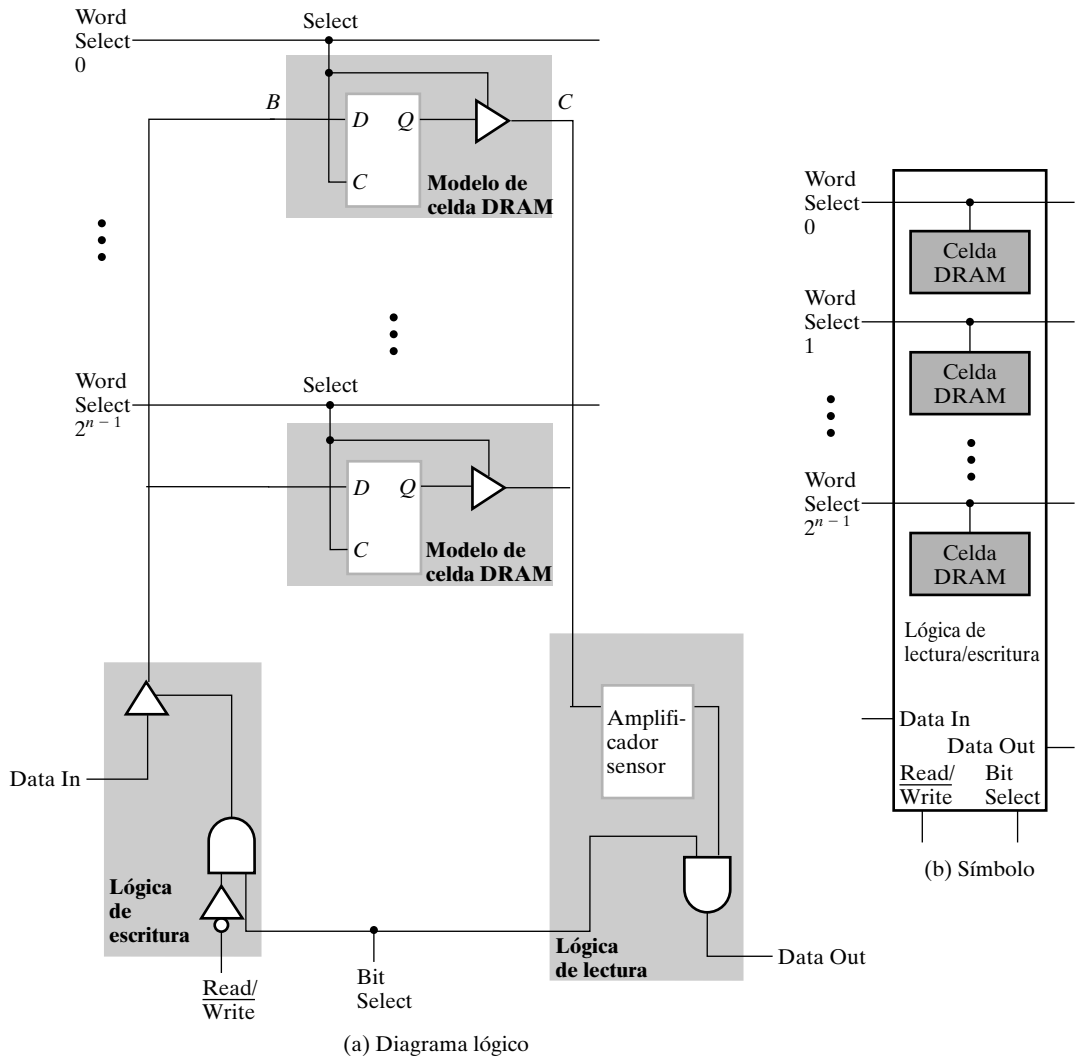
En la operación de lectura descrita, las Figuras 9-12(f) y 9-12(g) muestran que, independientemente del valor inicial almacenado en el depósito de almacenamiento, contiene ahora un valor intermedio que no provoca cambio suficiente en el nivel del depósito externo para permitir observar un 0 o un 1. Así, la operación de lectura ha destruido el valor almacenado; a esto le llamaremos *lectura destructiva*. Para permitir una lectura del valor original almacenado en un futuro, debemos *restaurar* dicho valor (es decir, devolver el depósito de almacenamiento a su nivel original). Para llevar a cabo el restablecimiento del 1 observado, el depósito grande se llena con la bomba y el depósito pequeño se llena a través de la válvula abierta. Para llevar al cabo el restablecimiento de un 0 almacenado que se ha observado, el depósito grande se vacía con la bomba y el depósito pequeño se desagua a través de la válvula abierta.

En las celdas de almacenamiento actuales hay otros caminos para el flujo de cargas. Estos caminos son análogos a pequeñas fugas en un depósito de almacenamiento. Debido a estas fugas, un depósito pequeño podrá desaguarse ocasionalmente hasta un punto en el que el incremento de nivel del depósito grande en una lectura puede no ser visto como un incremento. De hecho, si el depósito pequeño está a menos de la mitad cuando se lee, es posible que se pueda observar un decremento en el nivel del depósito grande. Para compensar estas fugas, el depósito pequeño, que almacena un 1, debe llenarse periódicamente. A esto se llama refresco del contenido de la celda. Cada celda de almacenamiento debe refrescarse antes de que su nivel haya caído a un punto tal que el valor almacenado no se pueda observar correctamente.

El funcionamiento de la memoria DRAM se ha explicado mediante una analogía hidráulica. Igual que hicimos con la memoria SRAM, empleamos un modelo lógico para la celda. El modelo se muestra en la Figura 9-12(h) es un latch tipo *D*. La entrada *C* al latch tipo *D* es *Select* y la entrada *D* del latch es *B*. Para modelar la salida de la celda de la memoria DRAM usamos un buffer triestado con la señal *Select* como entrada de control y *C* como su salida. En el circuito electrónico original de la celda de memoria DRAM de la Figura 9-12(a), *B* y *C* son la misma señal pero en el modelo lógico están separadas. Esto es necesario hacerlo en el proceso de modelado para evitar conectar las salidas de las puertas.

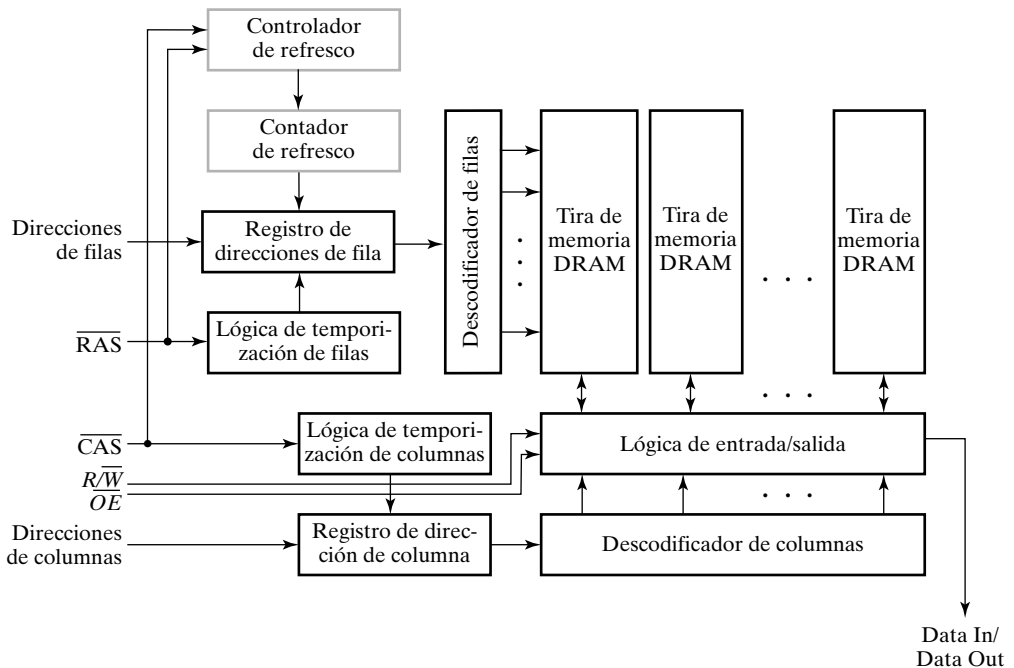
Tira de un bit de memoria DRAM

Usando el modelo lógico de la celda de memoria DRAM, vamos a construir el modelo de una tira de un bit de una memoria DRAM, como se muestra en la Figura 9-13. Este modelo es similar al de la tira de un bit de la memoria DRAM de la Figura 9-5. Es evidente que, aparte de la estructura de la celda, las dos tiras de un bit de memoria RAM son similares desde el punto de vista lógico. Sin embargo, desde el punto de vista del coste por bit, hay bastantes diferencias. La celda de una memoria DRAM está formada por un condensador más un transistor. La celda de memoria SRAM está formada típicamente por seis transistores, dando a la celda una complejidad tres veces superior, aproximadamente, a la de la memoria DRAM. Por tanto, el número de celdas de memoria SRAM, en un chip de un tamaño dado, es menor que un tercio del número de celdas en una memoria DRAM. El coste por bit de una memoria DRAM es menor que un tercio del coste por bit de una memoria SRAM, lo cual justifica el uso de memorias DRAM en memorias de gran capacidad.



□ FIGURA 9-13
Modelo de la tira de un bit de una memoria DRAM

Queda por discutir el refresco del contenido de la memoria DRAM. Antes de eso, necesitamos desarrollar la estructura típica usada para manejar el direccionamiento de las memorias DRAM. Puesto que muchos chips de memoria DRAM se usan para formar una memoria DRAM, queremos reducir el tamaño físico de los chips de la DRAM. Las memorias DRAM de gran capacidad necesitan 20 o más bits de dirección, lo cual supone 20 pines de dirección en cada chip. Para reducir el número de pines, las direcciones de la memoria DRAM se aplican vía serie en dos partes, la primera para la dirección de las filas y la segunda para la dirección de las columnas. Esto se puede hacer puesto que la dirección de las filas, que realiza la selección de fila, en realidad se necesita un tiempo antes que la dirección de las columnas, que es la que saca los datos para la lectura de la fila seleccionada. Para mantener las direcciones de la fila durante el ciclo de lectura o escritura, se almacena en un registro, según se muestra en la Figura 9-14. Las direcciones de las columnas también se almacenan en un registro. La señal de carga para el

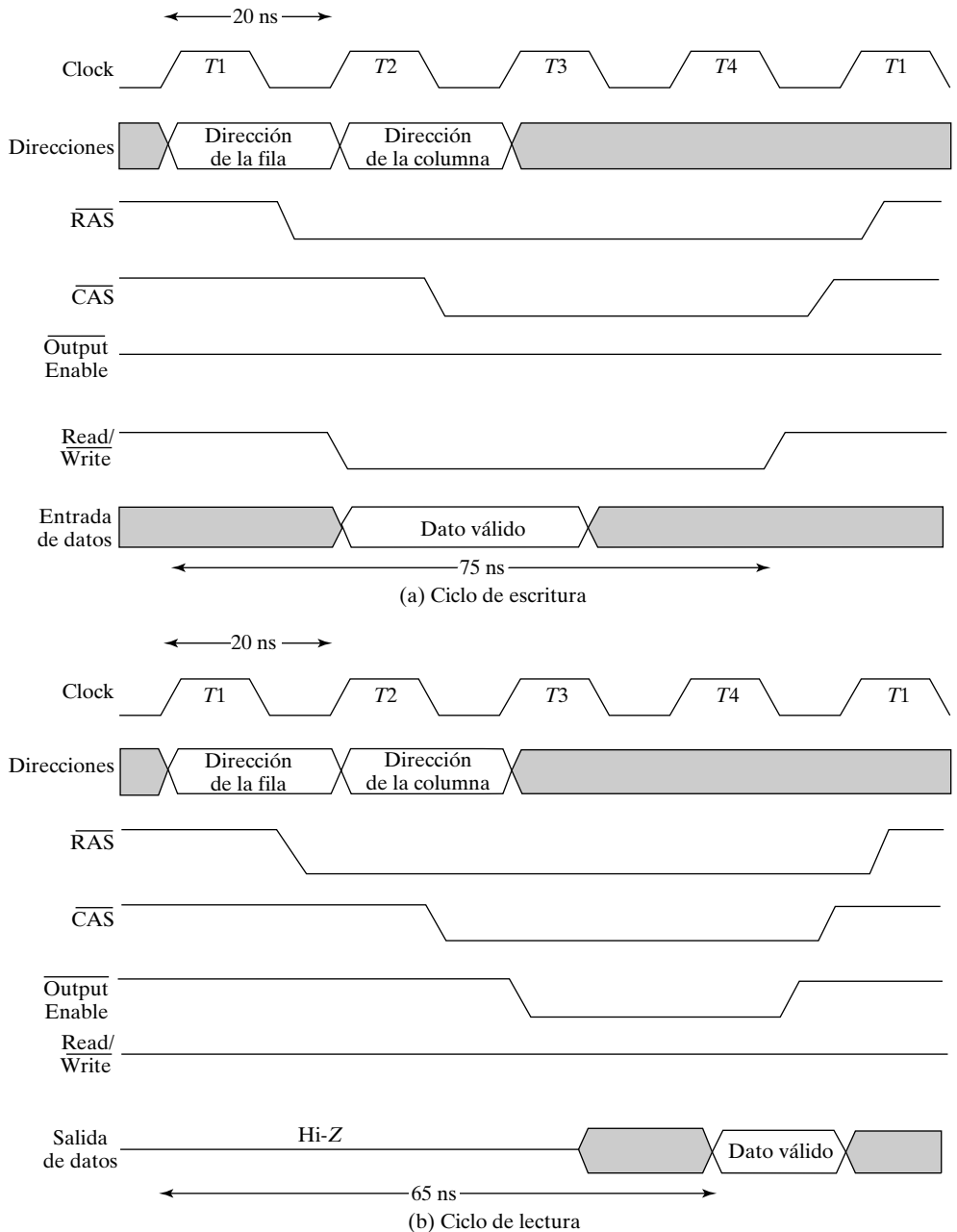


□ FIGURA 9-14
Diagrama de bloques de una memoria DRAM incluida su lógica de refresco

registro de direcciones de las filas es \overline{RAS} (del inglés *Row Adress Strobe*) y para el registro de direcciones de las columnas es \overline{CAS} (del inglés *Column Adress Strobe*). Además de estas señales de control, los chips de memoria DRAM también tienen las señales de control R/\overline{W} (Read/Write) y \overline{OE} (del inglés *Output Enable*). Véase que en este diseño usa señales activas a nivel bajo.

En la Figura 9-15(a) aparece la temporización de las señales para las operaciones de escritura y lectura. La dirección de la fila se aplica a las entradas de dirección, y la señal \overline{RAS} cambia de 1 a 0, cargando la dirección de la fila en el registro de direcciones de las filas. Esta dirección se aplica al decodificador de las direcciones de las filas y selecciona una fila de celdas de la memoria DRAM. Mientras tanto, se aplican la dirección de la columna y, luego, la señal \overline{CAS} cambia de 1 a 0, cargando la dirección de la columna en el registro de direcciones de las columnas. Esta dirección se aplica al decodificador de direcciones de las columnas, seleccionando un conjunto de columnas del array de la memoria RAM, de igual tamaño que el número de bit de los datos. La entrada de un dato, con $\text{Read/Write} = 0$, se aplica en un intervalo de tiempo similar al de la dirección de la columna. Los bits de datos se llevan a un conjunto de líneas de bits seleccionadas por el decodificador de direcciones de columnas, para llevar estos valores a las celdas de la memoria de la columna seleccionada, escribiendo los datos nuevos en las celdas. Cuando \overline{CAS} y \overline{RAS} pasan a valer 1, el ciclo de escritura se ha completado y las celdas de la memoria almacenan los nuevos datos escritos. Véase que el dato almacenado en el resto de las celdas de la fila direccionada se ha restaurado.

La temporización de las señales en el ciclo de lectura es similar, como se muestra en la Figura 9-15(b). La temporización de las direcciones es el mismo. Sin embargo, no se aplican datos y Read/Write es 1 en lugar de 0. Los valores del dato en las celdas de la memoria DRAM de la fila seleccionada se aplican a las líneas y se detectan por amplificadores sensores. El decodificador de direcciones de las columnas selecciona los valores que van a ser enviados a la



□ FIGURA 9-15
Temporización para las operaciones de lectura y escritura de una memoria DRAM

salida de datos, habilitada por \overline{OE} . Durante la operación de lectura, todos los valores de la fila seleccionada se restauran.

Para proporcionar el refresco está la lógica adicional en el diagrama de bloques de la Figura 9-14 (rectángulos coloreados). Hay un contador de refresco y un controlador de refresco. El contador de refresco se usa para proporcionar la dirección de la fila de las celdas de la me-

moria DRAM a refrescar. Esto es fundamental en los modos de refresco que necesita la dirección a ser proporcionada desde el interior del chip de la memoria DRAM. El contador de refresco avanza en cada ciclo de refresco. Según el número de bits del contador, cuando alcanza el valor $2^n - 1$, donde n es el número de filas del array de la memoria, pasa a 0 en el siguiente ciclo de refresco. Las formas estándar en la que el refresco se efectúa y el tipo de refresco correspondiente son los siguientes:

1. **RAS only refresh.** La dirección de una columna se coloca en las líneas de dirección y la señal RAS se pone a 0. En este caso, la dirección a refrescar debe proceder del exterior del chip de memoria DRAM, típicamente de un circuito integrado llamado controlador de DRAM.
2. **CAS before RAS refresh.** La señal CAS cambia de 1 a 0 y seguida de un cambio de 1 a 0 en la señal RAS. Se pueden llevar a cabo ciclos de refresco adicionales cambiando la señal RAS sin cambiar la señal CAS. En este caso, las direcciones de refresco proceden del contador de refresco, que se incrementa después del refresco para cada ciclo.
3. **Hidden refresh.** Siguiendo una escritura o lectura normal, la señal CAS se deja a 0 y la señal RAS se activa cíclicamente, efectuando un refresco del tipo *CAS before RAS refresh*. Durante un *hidden refresh*, la salida del dato de la anterior lectura permanece válido. De esta forma, el refresco permanece oculto. Desafortunadamente, el tiempo empleado por el *hidden refresh* es significativo, de forma que se retrasa la siguiente operación de lectura y escritura.

En todos los casos, tenga en cuenta que el inicio del refresco se controla externamente usando las señales *RAS* y *CAS*. Cada fila de un chip de memoria DRAM necesita un refresco dentro de un tiempo de refresco máximo especificado, típicamente en el rango de 16 a 64 milisegundos (ms). Los refrescos pueden realizarse en puntos espaciados uniformemente, llamándose entonces refresco en modo distribuido. Alternativamente, todos los refrescos se realizan uno después de otro, llamándose refresco en modo ráfaga. Por ejemplo, una memoria DRAM de $4\text{ M} \times 4$ tiene un tiempo de refresco de 64 ms y tiene 4096 filas para refrescar. La cantidad de tiempo para realizar un único refresco es de 60 ns y el intervalo de refresco en modo distribuido es $64\text{ ms}/4096 = 15,6\text{ microsegundos } (\mu\text{s})$. El tiempo total de refresco de 0.25 ms se saca del intervalo de 64 ms del intervalo de refresco. Durante los ciclos de refresco no se pueden efectuar operaciones de lectura ni de escritura en la memoria DRAM. Puesto que la ráfaga de refresco podría parar la operación de la computadora durante un periodo bastante largo, el modo de refresco más utilizado es el distribuido.

9-6 TIPOS DE MEMORIA DRAM

En las dos últimas décadas la capacidad y la velocidad de las memorias DRAM se han incrementado significativamente. La demanda de mayor velocidad ha dado lugar a la evolución de varios tipos de memoria DRAM. En la Tabla 9-2 se enumeran algunos tipos de memoria DRAM junto a una breve descripción. De los tipos de memoria enumerados, los dos primeros tipos han sido reemplazados en el mercado por memorias SDRAM y RDRAM más avanzadas. Al existir una presentación sobre códigos de corrección de errores (ECC) para memorias en la página web del libro, nos centraremos en las memorias DRAM síncronas, memorias DRAM de doble velocidad (*double data rate synchronous DRAM*) y memorias DRAM Rambus[®]. Antes de considerar estos tres tipos de memoria DRAM se presentaran brevemente algunos conceptos básicos.

□ **TABLA 9-2**
Tipos de DRAM

Tipo	Abreviatura	Descripción
Fast Page Mode DRAM	FPM DRAM	Toma la ventaja del hecho de que, cuando se accede a una fila, todos sus valores están disponibles para leerse. Cambiando la dirección de la columna se pueden leer los datos de diferentes direcciones sin necesidad de volver a poner la dirección de la fila y sin tener que esperar por el retardo asociado con la lectura de la fila de celdas si la porción de la fila coincide con las direcciones
Extended Data Output DRAM	EDO DRAM	Aumenta el tiempo que la DRAM mantiene los valores de los datos en su salida, permitiendo que la CPU realice otras tareas durante el acceso puesto que sabe que el dato estará todavía disponible.
Synchronous DRAM	SDRAM	Funciona con un reloj en lugar de operar asincrónicamente. Esto permite una más estrecha interacción entre la memoria y la CPU ya que la CPU conoce exactamente cuando los datos estarán disponibles. La memoria SDRAM usa la ventaja de la disponibilidad del valor de la fila y divide la memoria en distintos bancos, permitiendo el acceso solapado.
Double Data Rate Synchronous DRAM	DDR SDRAM	Igual que la memoria SDRAM excepto que la salida de los datos se proporciona tanto en el flanco de subida como en el de bajada.
Rambus [©] DRAM	RDRAM	Una tecnología propietaria que proporciona una muy alta velocidad de acceso usando un bus relativamente estrecho.
Error-Correcting Code	ECC	Se puede aplicar a la mayoría de los anteriores tipos de memorias DRAM para corregir errores en los datos de un bit y detectar los errores de dos bits.

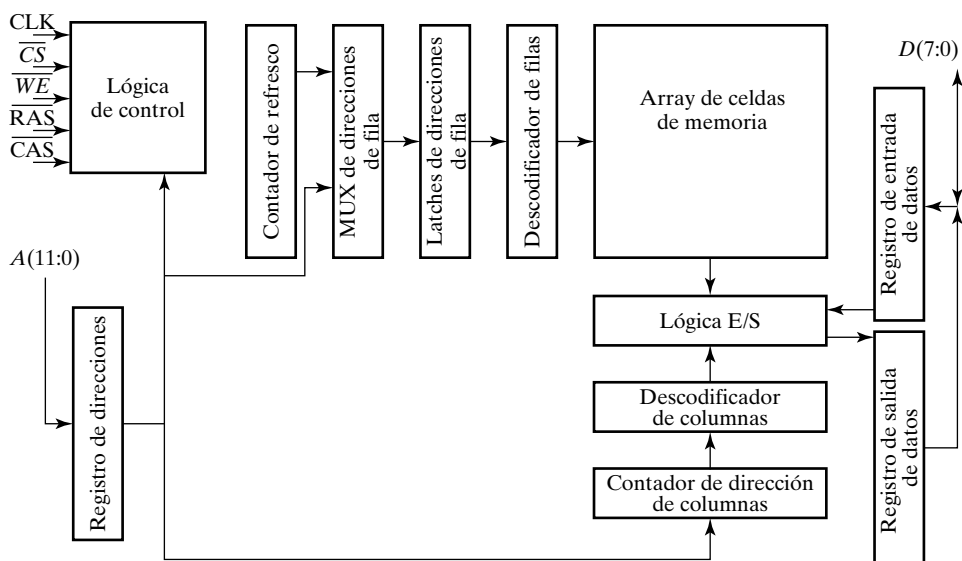
Primero, hemos de indicar que estos tres tipos de memoria DRAM funcionan bien debido al particular entorno en el que trabajan. En los modernos sistemas de alta velocidad, el procesador interactúa con la memoria DRAM dentro de una memoria jerárquica. La mayoría de las instrucciones y datos para el procesador se extraen de los dos niveles más bajos de la jerarquía, las cachés de primer y segundo nivel (L1 y L2 respectivamente). Estas memorias son, en comparación, más pequeña que las estructuras de memoria basadas en memorias SRAM, que se tratan con detalle en el capítulo 14. En nuestro estudio, la cuestión clave es que la mayoría de las lecturas de la DRAM no se realizan directamente por la CPU, en su lugar, se realizan lecturas

iniciales que llevan los datos y las instrucciones a estas cachés. Las lecturas se efectúan sobre un bloque de direcciones consecutivas cuya información se lleva a las cachés. Por ejemplo, realizar una lectura de 16 bytes de las direcciones 000000 a la 00000F. A esto se le llama ráfaga de lectura. En ráfagas de lectura el parámetro importante es la velocidad efectiva de lectura de bytes, que depende de las lecturas de las direcciones consecutivas, más que el tiempo de acceso. Según este parámetro, los tres tipos de memorias presentadas proporcionan un funcionamiento muy rápido.

Segundo, la efectividad de estos tres tipos de memoria DRAM depende del principio fundamental involucrado en el funcionamiento de las memorias DRAM, la lectura de todos y cada uno de los bits de una fila en cada operación de lectura. Este principio implica que todos los bits de la fila están disponibles después de una lectura usando esa fila si estos deben accederse. Con estos dos conceptos en mente, se puede presentar la memoria DRAM síncrona (SDRAM).

Memoria síncrona DRAM (SDRAM)

El uso de transferencias sincronizadas con el reloj diferencia a la memoria SDRAM de la DRAM convencional. En la Figura 9-16 aparece el diagrama de bloques de un circuito de memoria SDRAM de 16 megabytes. Las entradas y las salidas difieren en algo de las del diagrama de bloques de la memoria DRAM de la Figura 9-14 con la excepción de la presencia del reloj para sincronizar las operaciones. Internamente, hay ciertas diferencias. Puesto que la memoria SDRAM tiene una apariencia externa síncrona, tiene registros en las entradas de dirección y en las entradas y salidas de datos. Además se le ha incorporado un contador de direcciones de columna, que es la clave del funcionamiento de la memoria SDRAM. La lógica de control puede parecer similar pero, en este caso, es mucho más complicada ya que tiene una palabra de control de modo que se puede cargar del bus de direcciones. Suponiendo una memoria de 16 MB, el array de celdas de la memoria contiene 134 217 728 bits y es casi cuadrada, con



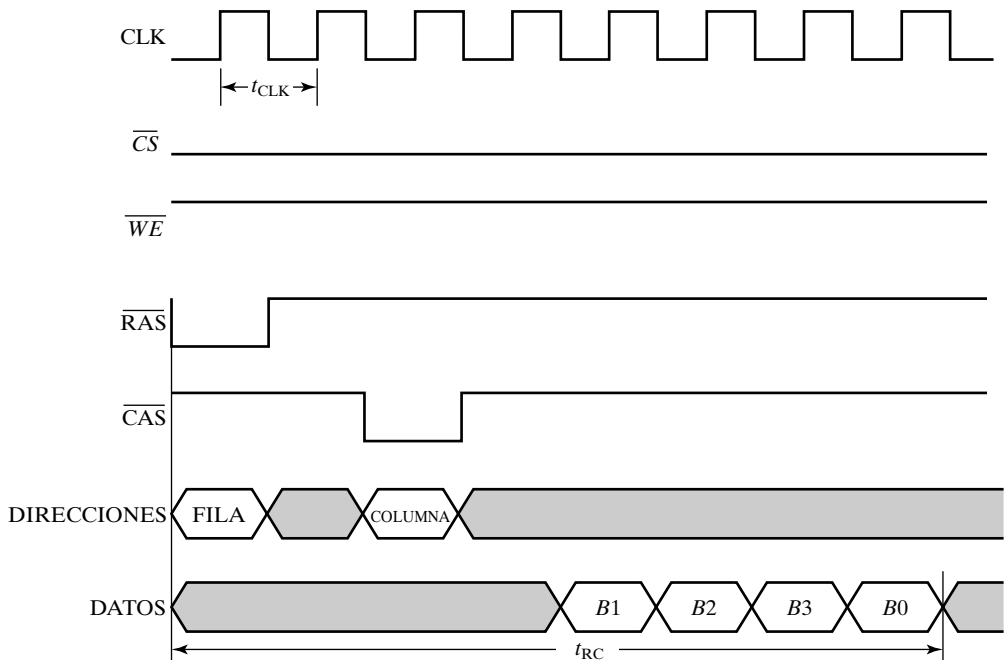
□ FIGURA 9-16

Diagrama de bloques de un memoria SDRAM de 16 MB

8 192 filas y 16 384 columnas. Tiene 13 bits de direcciones para las filas. Como tiene 8 bits por byte, el número de columnas direccionables es de 16 384 dividido por 8, es decir, 2048, es decir, hacen falta 11 bits para direccionar las columnas. Véase que 13 más 11 es igual a 24, que es el número correcto de bits de direcciones para una memoria de 16 MB.

Como en la DRAM normal, la memoria SDRAM aplica primero las direcciones de las filas seguidas de las direcciones de las columnas. Sin embargo, la temporización es algo diferente y se emplean algunos nuevos términos. Antes de realizar una operación de lectura de una determinada fila, todos los 2048 bytes de la fila especificada por la dirección se leen internamente y se almacenan en la lógica de entrada/salida. Este paso necesita internamente algunos ciclos de reloj. A continuación, la lectura se realiza con la dirección de la columna aplicada. Después de un retardo adicional de algunos ciclos de reloj, los bytes de datos empiezan a aparecer en la salida, uno por cada ciclo de reloj. El número de bytes que aparecen, la longitud de la ráfaga, ha sido cambiado cargando la palabra de control en la lógica de control desde la entrada de direcciones.

En la Figura 9.17 se muestra la temporización de un ciclo de lectura de una ráfaga de longitud igual a 4. La lectura comienza con la aplicación de la dirección de la fila y la habilitación de la dirección de la fila (*row address strobe*, RAS), que hace que la dirección de la fila sea capturada en el registro de dirección y que comience la lectura de la fila. Durante los siguientes dos ciclos de reloj tiene lugar la lectura de la fila. Durante el tercer ciclo de reloj se pone la dirección de las columnas y su habilitación (*column address strobe*, CAS), con la dirección de la columna capturada en el registro de direcciones y la lectura del primer byte iniciada. El byte ya está disponible para leerse de la memoria SDRAM en el flanco positivo dos ciclos después. El segundo, tercer y cuarto byte están disponibles para su lectura en los siguientes flancos de reloj. En la Figura 9-17 se puede observar que los bytes se presentan en el orden 1, 2, 3 y 0. Esto es así porque, en la identificación de la dirección de la columna, el byte que necesita inmediata-



□ FIGURA 9-17

Diagrama de tiempos de una memoria SDRAM

mente la CPU, los dos últimos bits son 01. Los siguientes bytes que aparecen dependen de la cuenta ascendente del módulo la longitud de la ráfaga que realiza el contador de direcciones de las columnas, dando direcciones que terminan en 01, 10, 11 y 00, permaneciendo el resto de bits fijos.

Es interesante comparar la velocidad de lectura de los bytes de una memoria SDRAM y de una memoria básica DRAM. Suponiendo que el tiempo del ciclo de lectura, t_{RC} , para una memoria DRAM básica es de 60 ns y que el periodo de reloj, t_{CLK} , es de 7.5 ns. La velocidad de lectura de una memoria DRAM básica es de un byte cada 60 ns o 16,67 MB/s. Para la SDRAM de la Figura 9-17, se necesitan 8.0 ciclos de reloj, es decir, 60 ns, para leer 4 bytes, dando una velocidad de 66.67 MB/s. Si la ráfaga es de ocho en lugar de cuatro, se necesita un tiempo de ciclo de lectura de 90 ns, dando una velocidad de 88.89 MB/s. Finalmente, si la ráfaga comprende los 2048 bytes de la fila de la memoria SDRAM, el tiempo del ciclo de lectura es $60 + (2048 - 4) \times 7.5 = 15.390$ ns, dando una tasa de transferencia de 133.07 MB/s, cercano al límite de un byte por ciclo de reloj de 7.5 ns.

Memoria SDRAM de doble tasa de transferencia de datos (DDR SDRAM)

El segundo tipo de memoria DRAM es la SDRAM de doble tasa de transferencia de datos (*double data rate* SDRAM, DDR SDRAM) supera el límite anteriormente presentado sin tener que decrementar el periodo de reloj. En cambio puede proporcionar dos bytes de datos por cada ciclo de reloj usando tanto el flanco de subida como el flanco de bajada. En la Figura 9-17 se leen cuatro bytes, uno por cada ciclo de reloj. Usando ambos flancos de reloj, se pueden transferir ocho bytes en el mismo tiempo de ciclo de lectura, t_{RC} . Para un periodo de reloj de 7.5 ns, la tasa de transferencia dobla la del ejemplo dando lugar a 266.24 MB/s.

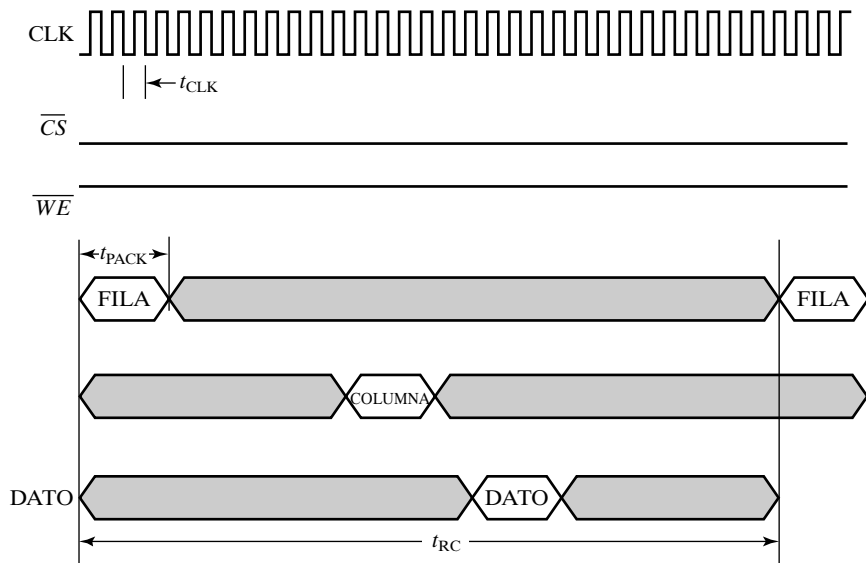
Se pueden aplicar otras técnicas básicas para incrementar la tasa de transferencia de bytes. Por ejemplo, en lugar de tener un dato de un solo byte se puede tener un chip de memoria SDRAM con una longitud del dato de entrada/salida de cuatro bytes (32 bits). Esto proporciona una tasa límite de transferencia de 1066 GB/s con un reloj con un periodo de 7.5 ns. Con 8 bytes se alcanzaría una tasa de 2130 GB/s.

La tasa de transferencia conseguida en los ejemplos son los límites superiores. Si un acceso se ha de realizar usando diferentes filas de la RAM, el retardo desde la aplicación de pulso RAS a la lectura del primer byte de dato es significativo y lleva a rendimientos bastante por debajo del límite. Esto se puede compensar parcialmente dividiendo la memoria en varios bancos donde cada uno realiza la lectura de la columna independientemente. Con tal de que la dirección de la fila y del banco estén disponibles con bastante antelación, la lectura de la fila se puede realizar sobre uno o más bancos mientras el dato de la columna activa se está aún transfiriendo. Cuando la lectura de la columna de la fila activa se completa, el dato puede estar potencialmente disponible inmediatamente de otros bancos, permitiendo un flujo ininterrumpido de datos de la memoria. Esto permite conseguir unas tasas de transferencia cercanas al límite. Sin embargo, debido al hecho de que puede ocurrir que se accedan secuencialmente a varias filas del mismo, esta tasa máxima nunca se alcanza.

Memoria RAMBUS[®] DRAM (RDRAM)

El último tipo de memoria DRAM que se va a presentar es la memoria RAMBUS DRAM. Los circuitos integrados RDRAM se han diseñado para ser integrados en la memoria de un sistema

que usa un bus basado en paquetes para la interacción de los circuitos de memoria RDRAM y el bus de memoria con el procesador. Los principales componentes del bus son: una ruta de 3 bits para las direcciones de las filas, una ruta de 5 bits para las direcciones de las columnas y una ruta de 16 o 18 bits para los datos. El bus es síncrono y efectúa transferencias en ambos flancos de reloj, como en la memoria DDR SDRAM. La información en las tres rutas anteriormente presentadas se transfiere en paquetes durante cuatro ciclos de reloj, es decir, se realizan 8 transferencias de paquetes. El número de bits por paquete de cada una de las rutas es de 24 bits para el paquete de direcciones de las filas, 40 bits para los paquetes de direcciones de las columnas y 128 o 144 bits para los paquetes de datos. El paquete más grande incluye 16 bits de paridad para realizar códigos de corrección de errores. Los circuitos RDRAM emplean el concepto de varios bancos de memoria, como se mencionó anteriormente, para proporcionar la posibilidad de realizar accesos concurrentes de diferentes direcciones de filas. La memoria RDRAM utiliza la técnica habitual de fila activada en la que se lee la fila direccionada de la memoria. A partir de esta fila de datos, la dirección de la columna se usa para seleccionar las parejas de bytes en el orden en que deben ser transmitidos en el paquete. En la Figura 9-18 se muestra una temporización típica de un acceso de lectura de una memoria RDARM. Como consecuencia del sofisticado diseño electrónico del sistema de memoria RAMBUS, consideramos un periodo de reloj de 1.875 ns. Así, el tiempo de transmisión de un paquete es $t_{\text{PACK}} = 4 \times 1.875 = 7.5$ ns. El tiempo del ciclo de acceso para un solo paquete de datos de 8 parejas de bytes o 16 bytes es de 266.67 MB/s. Si se accede a cuatro paquetes de un byte de la misma fila, la tasa se incrementa a 1.067 GB/s. Para leer todo el contenido de una fila de una memoria RDRAM de 2048 bytes, el tiempo de acceso se incrementa en $60 + (2048 - 64) \times 1.875/4 = 990$ ns o a la tasa límite de transferencia de bytes de $2048/(990 \times 10^{-9}) = 2.069$ MB/s, aproximándose al límite ideal de de $4/1.875$ ns o 2.133 GB/s.



□ FIGURA 9-18
Diagrama de tiempos de una RDRAM de 16 MB

9-7 ARRAYS DE CIRCUITOS INTEGRADOS DE MEMORIAS DINÁMICAS RAM

Muchos de los principios de diseños utilizados en la Sección 9-4 para hacer *arrays* de memorias SRAM se aplican al diseño de *arrays* DRAM. Sin embargo existen diversos requerimientos diferentes para controlar y direccionar los *arrays* de las memorias DRAM. Estos requerimientos se llevan a cabo mediante un controlador de memorias DRAM. Las funciones realizadas por dicho controlador son las siguientes:

1. realiza la separación de las direcciones en direcciones de las filas y las direcciones de las columnas, proporcionando la temporización necesaria,
2. proporciona las señales \overline{RAS} y \overline{CAS} con su temporización correcta para las operaciones de lectura, escritura y refresco.
3. lleva a cabo las operaciones de refresco en los intervalos apropiados, y
4. proporciona las señales de status al resto del sistema (por ejemplo, indica cuándo la memoria está ocupada realizando su intervalo de refresco).

El controlador de la memoria DRAM es un circuito secuencial complejo sincronizado con el reloj exterior de la CPU para su funcionamiento.

9-8 RESUMEN DEL CAPÍTULO

Las memorias son de dos tipos: memorias de acceso aleatorio (RAM) y memorias de sólo lectura (ROM). En ambos tipos aplicamos una dirección para leer o escribir un dato. Las operaciones de lectura y escritura han de cumplir ciertos pasos con sus parámetros temporales, incluyendo el tiempo de acceso y el tiempo del ciclo de escritura. Las memorias pueden ser estáticas o dinámicas y volátiles o no volátiles. Internamente, un circuito de una memoria RAM está formado por un array de celdas RAM, decodificadores, circuitos de escritura, circuitos de lectura y circuitos de salida. La combinación de estos circuitos de escritura, lectura y las celdas RAM asociadas se pueden modelar lógicamente como una tira de memoria RAM de un bit. Las tiras de memoria RAM de un bit se pueden combinar para formar *arrays* de celdas de dos dimensiones que, junto con los decodificadores y los circuitos de salida, forman la base para un chip de memoria RAM. Los circuitos de salida utilizan *buffers* tri-estado para facilitar la conexión de un array de chips de memoria con una lógica adicional reducida. En las memorias DRAM, debido a la necesidad de refresco, se debe utilizar circuitería adicional para llevarlo a cabo así como para utilizar *arrays* de chips. Como consecuencia de la necesidad de tener memorias con accesos más rápidos se han desarrollado nuevos tipos de memorias DRAM. Los tipos más recientes de memorias de alta velocidad DRAM emplean interfaces síncronas que usan un reloj para el control de los accesos de la memoria.



Los códigos de detección y corrección de errores, basados frecuentemente en Códigos Hamming, se usan para detectar o corregir errores de los datos almacenados en las memorias RAM. Existe material de la primera edición que cubre este tipo de códigos disponible en la dirección de Internet: <http://www.librosite.net/Mano>.



El materia que cubre las memorias en VHDL y Verilog está disponible en la dirección de Internet: <http://www.librosite.net/Mano>.

REFERENCIAS

1. WESTE, N. H. E., and ESHRAGHIAN, K.: *Principles of CMOS VLSI Design: A Systems Perspective*, 2nd ed. Reading, MA: Addison-Wesley, 1993.
2. Micron Technology, Inc. *Micron 256 Mb: x4, x8, x16 SDRAM*. www.micron.com, 2002.
3. Micron Technology, Inc. *Micron 64 Mb: x32 DDR SDRAM*. www.micron.com, 2001.
4. SOBELMAN, M.: «Rambus Technology Basics», *Rambus Developer Forum*. Rambus, Inc., October 2001.
5. Rambus, Inc. *Rambus Direct RDRAM 128/144-Mbit (256x16/18x32s) - Preliminary Information*, Documento DL0059 Versión 1.11.

PROBLEMAS



El signo (+) indica problemas más avanzados y el asterisco (*) indica que hay una solución disponible en la dirección de Internet: <http://www.librosite.net/Mano>.

- 9-1. *Las siguientes memorias se especifican mediante el número de palabras y el número de bits por palabra ¿Cuántas líneas de dirección y líneas de entrada/salida de datos se necesitan en cada caso?
 - (a) $16\text{ K} \times 8$
 - (b) $256\text{ K} \times 16$
 - (c) $64\text{ M} \times 32$
 - (d) $2\text{ G} \times 8$.
- 9-2. Indique el número de bits almacenados en las memorias enumeradas en el Problema 9-1.
- 9-3. *La palabra número $(835)_{10}$ de la memoria mostrada en la Figura 9-2 contiene el dato binario equivalente a $(15\ 103)_{10}$. Indique la dirección de 10 bits y el contenido de 16 bits de la palabra.
- 9-4. Un chip de memoria RAM de $64\text{ K} \times 16$ utiliza una descodificación coincidente dividiendo el decodificador interno en selección de fila y selección de columna.
 - (a) Suponiendo que el array de celdas RAM es cuadrado ¿Cuál es el tamaño de cada decodificador y cuántas puertas AND se necesitan para la descodificación de una dirección?
 - (b) Determine que las líneas de selección de fila y de columna están habilitadas cuando la dirección de entrada es el valor binario equivalente de $(32\ 000)_{10}$.
- 9-5. Suponga que el decodificador más grande que se puede usar en un chip de memoria de $m \times 1$ RAM tiene 13 entradas de dirección y que se emplea descodificación coincidente. Para construir chips de memoria RAM que contengan más de m palabras de 1 bit, se incluyen *arrays* múltiples de celdas RAM, cada una con su decodificador y circuitos de lectura/escritura.
 - (a) Con la restricción dada del decodificador ¿cuántos *arrays* de celdas RAM se necesitan para construir un chip de memoria RAM de $512\text{ M} \times 1$?
 - (b) Muestre qué decodificador es necesario emplear para seleccionar entre los diferentes *arrays* de memoria RAM de la memoria y sus conexiones a los bits de direcciones y los decodificadores de las columnas.

- 9-6.** Una memoria DRAM tiene 14 pines de dirección y su dirección de columna es de un bit. ¿Cuántas direcciones en total tiene la memoria DRAM?
- 9-7.** Una memoria DRAM de 256 Mb, que utiliza datos de 4 bits, tiene igual longitud de direcciones de filas y de columnas. ¿Cuántos pines de dirección tiene dicha memoria?
- 9-8.** Una memoria DRAM tiene un intervalo de refresco de 128 ms y 4096 filas. ¿Cuál es la duración del intervalo entre refresco en el modo de refresco distribuido? ¿Cuál es el número mínimo de pines de dirección de la memoria DRAM?
- 9-9.** *(a) ¿Cuántos chips de memoria RAM de 128×16 se necesitan para conseguir una memoria de 1 M bytes de capacidad?
(b) ¿Cuántas líneas de dirección se necesitan para direccionar 1 M bytes? ¿Cuántas de estas líneas están conectadas a las entradas de dirección de todos los chips?
(c) ¿Cuántas líneas se deben descodificar para generar el *chip select*? Especifique el tamaño del decodificador.
- 9-10.** Usando el chip de memoria RAM de capacidad $64 \text{ K} \times 8$ de la Figura 9-9 más un decodificador, construya el diagrama de bloques de una memoria RAM de $512 \text{ K} \times 16$.
- 9-11.** Explique cómo utiliza la memoria SDRAM la ventaja del *array* de dos dimensiones para conseguir altas tasas de transferencia de datos.
- 9-12.** Explique cómo se consigue doblar la tasa de transferencia de una memoria DDRAM en comparación con una memoria SDRAM.