

第5章 存储器系统

【学习目标】

本章首先以半导体存储器为对象,在讨论存储器及其基本电路、基础知识的基础上,讨论存储芯片及其与 CPU 之间的连接和扩充问题;然后,介绍内存的技术发展以及外部存储器;最后,简要介绍存储器系统的分层结构。

【学习要求】

- ◆ 存储器的分类、组成及功能。着重理解行选与列选对 1 位信息的读出。
- ◆ 重点掌握位扩充与地址扩充技术。
- ◆ 理解存储器与 CPU 的连接方法。
- ◆ 着重理解内存技术的发展。
- ◆ 理解存储器系统的分层结构。

5.1 存储器的分类与组成

计算机的存储器可分为两大类:一类为内部存储器,简称内存或主存,其基本存储元件多以半导体材料制造;另一类为外部存储器,简称外存,多以磁性材料或光学材料制造。

5.1.1 半导体存储器的分类

半导体存储器的分类如图 5-1 所示。按使用的功能可分为两大类:随机存取存储器(random access memory, RAM)和只读存储器(read only memory, ROM)。

RAM 按工艺又可分为双极型 RAM 和 MOS RAM 两类,而 MOS RAM 又可分为静态(static)和动态(dynamic)RAM 两种。双极型 RAM 的特点是存取速度快,但集成度低,功耗大,主要用于速度要求高的位片式微机中;静态 MOS RAM 的集成度高于双极型 RAM,而功耗低于双极型 RAM;动态 RAM 比静态 RAM 具有更高的集成度,靠电路中的栅极电容存储信息。由于电容器上的电荷会泄漏,因此,它需要定时进行刷新。

只读存储器 ROM 按工艺也可分为双极型和 MOS 型,但一般根据信息写入的方式不同,而分为不可编程掩膜式 ROM,可编程 ROM(PROM)和可擦除、可再编程 ROM(包括紫外线擦除 EPROM 与电子擦除 E²PROM 以及 Flash ROM)等几种。

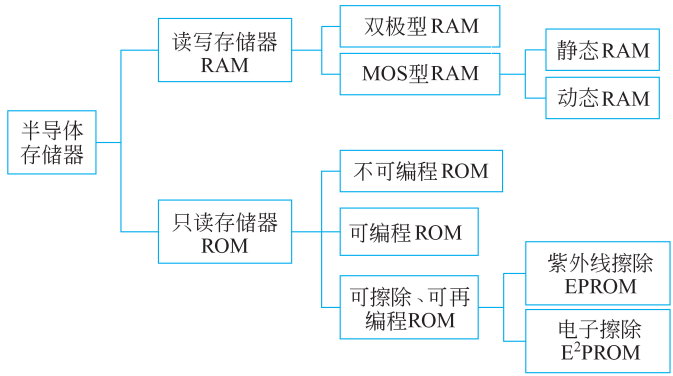


图 5-1 半导体存储器的分类

5.1.2 半导体存储器的组成

半导体存储器的组成框图如图 5-2 所示。它一般由存储体、地址选择电路、输入输出电路和控制电路组成。

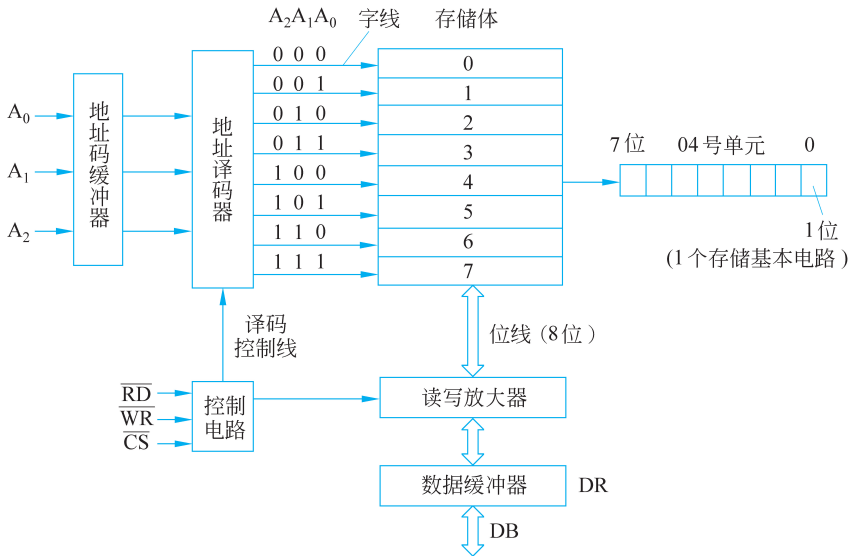


图 5-2 半导体存储器组成框图

1. 存储体

存储体是存储 1 或 0 信息的电路实体,它由许多个存储单元组成,每个存储单元赋予一个编号,称为地址单元号。而每个存储单元由若干相同的位组成,每个位需要一个存储元件。对存储容量为 1K(1024 个单元)×8 位的存储体,其总的存储位数为 1024×8

位=8192 位。

存储器的地址用一组二进制数表示,其地址线的位数 n 与存储单元的数量 N 之间的关系为 $2^n = N$ 。

地址线数与存储单元数的关系如表 5-1 所示。

表 5-1 地址线数与存储单元数的关系

地址线数 n	3	4	...	8	9	10	11	12	13	14	15	16
存储单元数 $N=2^n$	8	16	...	256	512	1024	2048	4096	8192	16 384	32 764	65 536
存储容量/B	8	16	...	256	512	1K	2K	4K	8K	16K	32K	64K

2. 地址选择电路

地址选择电路包括地址码缓冲器、地址译码器等。

地址译码器用来对地址码译码。设其输入端的地址线根数为 n ,输出线数为 N ,则它分别对应 2^n 个不同的地址码,作为对存储体地址单元的选择线。这些输出的选择线又称字线。

地址译码方式有以下两种。

(1) 单译码方式:或称字结构,其全部地址码只用一个地址译码器电路译码,译码输出的字选择线直接选中与输入地址码对应的存储单元。如图 5-2 所示,有 A_2 、 A_1 、 A_0 3 根输入地址线,经过地址译码器输出 8 种不同编号的字线:000、001、010、011、100、101、110、111。这 8 条字线分别对应 8 个不同的地址单元。这种单译码方式需要的选择线数较多,只适用于容量较小的存储器。

(2) 双译码方式:或称重合译码,双译码方式存储器结构如图 5-3 所示。它将地址码分为 X 与 Y 两部分,用两个译码电路分别译码。X 向译码又称行译码,其输出线称行选择线,它选中存储矩阵中一行的所有存储单元。Y 向译码又称列译码,其输出线称列选择线,它选中存储矩阵中一列的所有存储单元。只有 X 向和 Y 向的选择线同时选中的那一位存储单

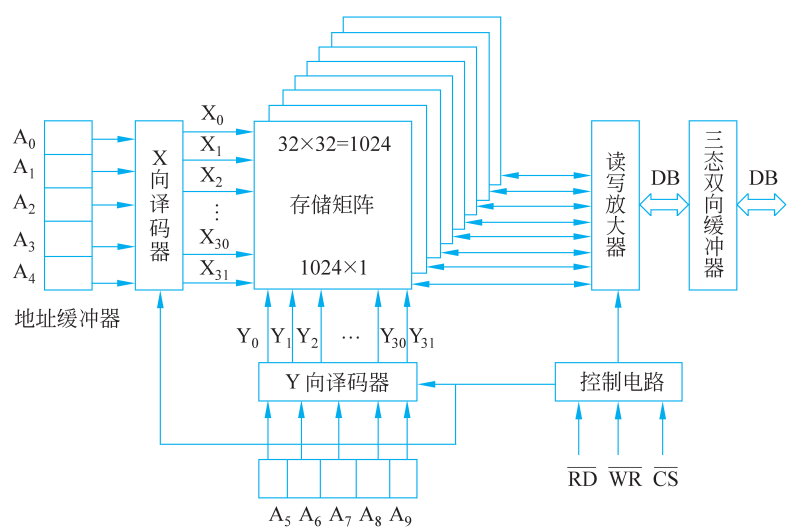


图 5-3 双译码存储器结构

元,才能进行读或写操作。由图可见,具有 1024 个基本单元电路的存储体排列成 32×32 的矩阵,它的 X 向和 Y 向译码器各有 32 根译码输出线,共 64 根。若采用单译码方式,则有 1024 根译码输出线。显然,双译码方式所需要的选择线数目较少,也简化了存储器的结构,故它适用于大容量的存储器。

3. 读写电路与控制电路

读写电路包括读写放大器、数据缓冲器(三态双向缓冲器)等,它是数据信息输入和输出的通道。

外界对存储器的控制信号有读信号($\overline{RD}$)、写信号($\overline{WR}$)和片选信号($\overline{CS}$)等,通过控制电路以控制存储器的读或写操作以及片选。只有片选信号处于有效状态,存储器才能与外界交换信息。

5.2 随机存取存储器

随机存取存储器(RAM)既可以读出,也可以写入。读出时并不损坏原来存储的内容,只有写入时才修改原来所存储的内容。断电后,存储内容立即消失,即具有易失性。它用于保存各种处理器需要使用的数据,可以加快计算机的运算速度。RAM 可分为静态(static RAM, SRAM)和动态(dynamic RAM, DRAM)两种。常用静态内存(SRAM)作为系统的高速缓存(通常用于一级缓存和二级缓存),而平常所提到的内存指的是动态内存,即 DRAM。

5.2.1 静态随机存取存储器

1. SRAM 基本存储电路

SRAM 的基本存储电路,是由 6 个 MOS(金属氧化物半导体)管组成的 RS 触发器,如图 5-4 所示。

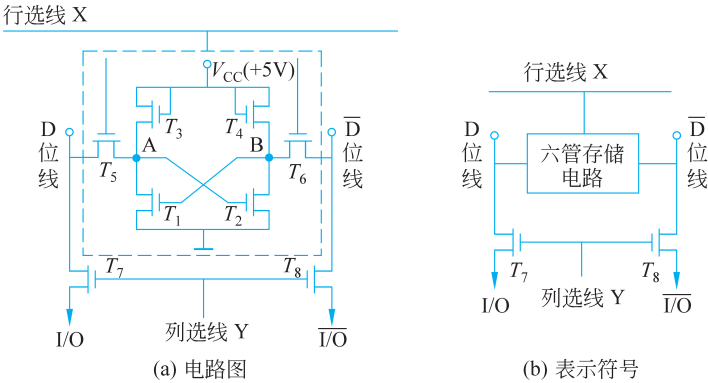


图 5-4 六管静态存储电路

在图 5-4 中, T_3 、 T_4 为负载管, T_1 、 T_2 交叉耦合组成了一个 RS 触发器, 具有两个稳定状态。在 A 点(相当于 Q 端)与 B 点(相当于 $\bar{Q}$ 端)可以分别寄存信息 1 和 0。 T_5 、 T_6 为行向选通门, 受行选线上的电平控制。 T_7 、 T_8 为列向选通门, 受列选线上的电平控制。由此, 组成了双译码方式。当行选线与列选线上的信号都为高电平时, 则分别将 T_5 、 T_6 与 T_7 、 T_8 导通, 使 A、B 两点的信息经 D 与 $\bar{D}$ 两点分别送至输入输出电路的 I/O 线及 $\bar{I}/\bar{Q}$ 线上, 从而存储器某单元位线上的信息同存储器外部的数据线相通。这时, 就可以对该单元位线上的信息进行读写操作。

写入时, 被写入的信息从 I/O 和 $\bar{I}/\bar{Q}$ 线输入。如写 1 时, 使 I/O 线为高电平, $\bar{I}/\bar{Q}$ 线为低电平, 经 T_7 、 T_5 与 T_8 、 T_6 分别加至 A 端和 B 端, 使 T_1 截止而 T_2 导通, 于是 A 端为高电平, 触发器为存 1 的稳态; 反之亦然。

读出时, 只要电路被选中, T_5 、 T_6 与 T_7 、 T_8 导通, 则 A 端与 B 端的电位就会送到 I/O 及 $\bar{I}/\bar{Q}$ 线上。若原存的信息为 1, 则 I/O 线上为 1, $\bar{I}/\bar{Q}$ 线上为 0; 反之亦然。读出信息时, 触发器的状态不受影响, 故为非破坏性读出。

2. SRAM 的组成

SRAM 的结构组成原理图, 如图 5-5 所示。存储体是一个由 $64 \times 64 = 4096$ 个六管静态存储电路组成的存储矩阵。在存储矩阵中, X 地址译码器输出端提供 $X_0 \sim X_{63}$ 计 64 根行选择线, 而每一行选择线接在同一行中的 64 个存储电路的行选端, 故行选择线能同时为该行 64 个行选端提供行选择信号。Y 地址译码器输出端提供 $Y_0 \sim Y_{63}$ 计 64 根列选择线, 而同一列中的 64 个存储电路共用同一位线, 故由列选择线可以同时控制它们与输入输出电路(I/O 电路)连通。显然, 只有行、列均被选中的某个单元存储电路(即 1 位), 在其 X 向选通门与 Y 向选通门同时被打开时, 才能进行信息的读出和写入操作。

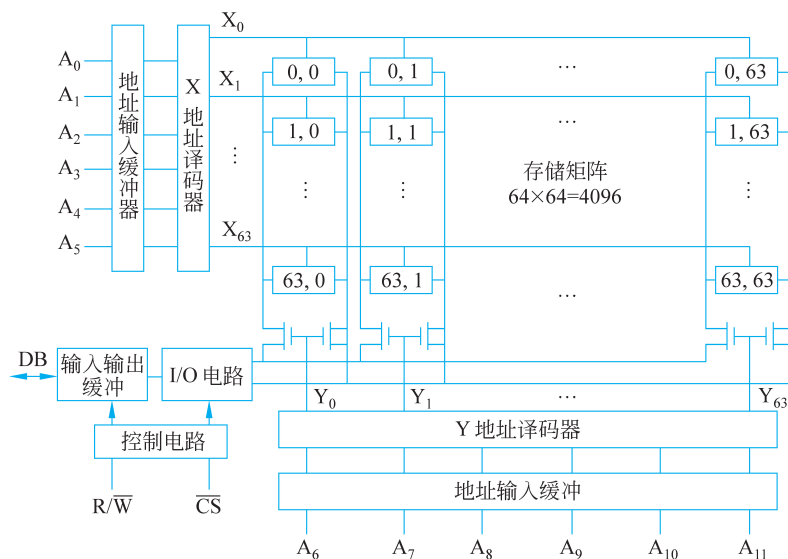


图 5-5 SRAM 结构组成原理图

图 5-5 中的存储体是容量为 $4K \times 1$ 位的存储器, 因此, 它仅有一个 I/O 电路, 用于存取

各存储单元中的 1 位信息。如果要组成字长为 4 位或 8 位的存储器,则每次存取时,同时应有 4 个或 8 个单元存储电路与外界交换信息。因此,在这种存储器中,要将列的列向选通门控制端引出线按 4 位或 8 位来分组,使每根列选择线能控制一组的列向门同时打开;相应地,I/O 电路也应有 4 个或 8 个。每一组的同一位共用一个 I/O 电路。这样,当存储体的某个存储单元在一次存取操作中,被地址译码器输出端的有效输出电平选中时,则该单元内的 4 位或 8 位信息被一次读写完毕。

必须指出,在图 5-5 中所示的存储体如果是 $4K \times 1$ 位的存储矩阵,则在读写操作时每次只能存取 1 位信息。如果是 8 个 $4K \times 1$ 位的存储矩阵,则在读写操作时每次才能存取 8 位信息,这时的存储容量为 $4K \times 8$ 位。通常,一个 RAM 芯片的存储容量是有限的,需要用若干片才能构成一个实用的存储器。这样,地址不同的存储单元,可能处于不同的芯片中,因此,在选中地址时,应先选择其所属的芯片。对于每块芯片,都有一个片选控制端($\overline{CS}$),只有当片选端加上有效信号时,才能对该芯片进行读或写操作。一般来说,片选信号由地址码的高位译码(通过译码器输出端)产生。

3. SRAM 的读写过程

SRAM 的读写过程参见图 5-5。

1) SRAM 读出过程

① 地址码 $A_0 \sim A_{11}$ 加到 RAM 芯片的地址输入端,经 X 与 Y 地址译码器译码,产生行选与列选信号,选中某一存储单元,该单元中存储的代码,经一定时间,出现在 I/O 电路的输入端。I/O 电路对读出的信号进行放大、整形,送至输出缓冲寄存器。缓冲寄存器一般具有三态控制功能,没有开门控制信号,所存数据还不能送到数据总线(DB)上。

② 在送上地址码的同时,还要送上读写控制信号($R/\overline{W}$ 或 $\overline{RD}$ 、 $\overline{WR}$)和片选信号($\overline{CS}$)。读出时,使 $R/\overline{W}=1$, $\overline{CS}=0$,这时,输出缓冲寄存器的三态门被打开,所存信息送至 DB 上,于是,存储单元中的信息被读出。

2) SRAM 写入过程

① 同 SRAM 读出过程①,先选中相应的存储单元,使其可以进行写操作。

② 将要写入的数据放在 DB 上。

③ 加上片选信号 $\overline{CS}=0$ 及写入信号 $R/\overline{W}=0$ 。这两个有效控制信号打开三态门使 DB 上的数据进入输入电路,送到存储单元的位线上,从而写入该存储单元。

4. SRAM 芯片举例

常用的 SRAM 芯片有 Intel 6116、6264、62256、628128、628512、6281024 等。

例如,Intel 6116 是一个 $2K \times 8$ 位的 CMOS SRAM 芯片,属双列直插式、24 条引脚封装。它的存储容量为 $2K \times 8$ 位,其引脚图及内部结构框图如图 5-6 所示。

Intel 6116 芯片内部的存储体是一个由 $128 \times 128 = 16\,384$ 个静态存储电路组成的存储矩阵。 $A_0 \sim A_{10}$ 11 根地址线供对其进行行地址、列地址译码,以便对 $2^{11} = 2048$ 个存储单元进行选址。每当选中的一个存储单元,将从该存储单元中同时读或写 8 位二进制信息,故 Intel 6116 有 8 根数据输入输出线 $I/O_0 \sim I/O_7$ 。Intel 6116 存储矩阵内部基本存储电路上的信息,正是通过 I/O 控制电路和数据输入输出缓冲器与 CPU 的数据总线连通的。数据

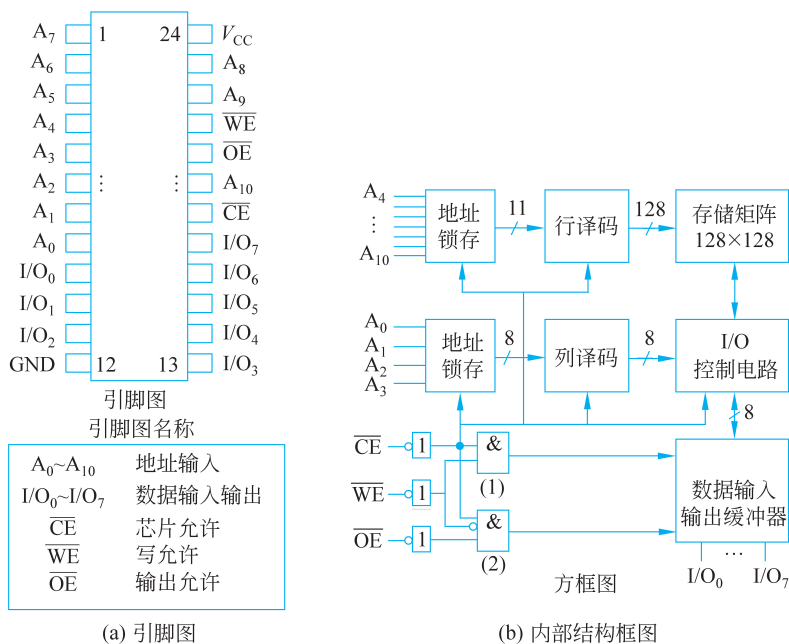


图 5-6 Intel 6116 芯片的引脚图及内部结构框图

的读出或写入将由片选允许信号 $\overline{CE}$ 、写允许信号 $\overline{WE}$ 以及数据输出允许信号 $\overline{OE}$ 一起控制。当 $\overline{CE}$ 有效而 $\overline{WE}$ 为低电平时, 1 门导通, 使数据输入缓冲器打开, 信息由 I/O₀~I/O₇ 写入被选中的存储单元; 当 $\overline{CE}$ 与 $\overline{OE}$ 同时有效而 $\overline{WE}$ 为高电平时, 2 门导通, 使数据输出缓冲器打开, CPU 从被选中的存储单元由 I/O₀~I/O₇ 读出信息送往数据总线。无论是写入或读出, 一次都是读写 8 位二进制信息。

Intel 6264 芯片的结构及工作原理与 Intel 6116 相似, 是一个存储容量为 8K×8 位的 CMOS SRAM 芯片, 其外部引脚如图 5-7 所示。它有 28 条引脚, 包括 13 根地址线 (A₁₂~A₀)、8 根双向数据线 (D₇~D₀) 以及 4 根控制线 (片选信号线 $\overline{CS_1}$ 、 $\overline{CS_2}$ 、输出允许信号 $\overline{OE}$ 与写允许信号), 另外, 还有 3 根其他信号线 (+5V 电源端 V_{CC}、接地端 GND、空端 NC)。这些引脚的功能及其用法是很容易理解的, 不再赘述。

需要补充的是, Intel 6264 芯片有两个片选端 $\overline{CS_1}$ 与 $\overline{CS_2}$, 在 CPU 选择 Intel 6264 芯片时, 必须使其两个片选信号 $\overline{CS_1}$ 与 $\overline{CS_2}$ 同时有效才行。事实上, 一个微机系统的内存空间通常是由若干块存储器芯片组成的, 各个存储器芯片究竟映射到内存空间的哪一段地址区间, 是由高位地址信号决定的。系统中的一组高位地址信号和控制信号通过译码器译码可产生对应的一组片选信号, 但每次只有一个特定的高位地址会将某个存储器芯片映射到所需要的地址范围上。

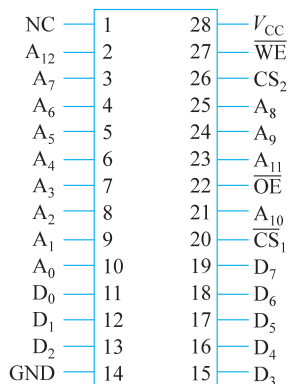


图 5-7 Intel 6264 SRAM 外部引脚图

5.2.2 动态随机存取存储器

动态随机存取存储器(DRAM)芯片是以 MOS 管栅极电容是否充有电荷来存储信息的,其基本单元电路一般由四管、三管和单管组成,以三管和单管较为常用。由于它所需要的管子较少,故可以扩大每片存储器芯片的容量,并且其功耗较低,所以在微机系统中,大多数采用 DRAM 芯片。

1. 动态基本存储电路

下面重点介绍常用的三管和单管两种基本存储电路。

1) 三管动态基本存储电路

三管动态基本存储电路如图 5-8 所示,它由 T_1 、 T_2 、 T_3 3 个管子和两条字选择线(读、写选择线),以及两条数据线(读、写数据线)组成。

T_1 是写数控制管; T_2 是存储管,用它的栅极电容 C_g 存储信息; T_3 是读数控制管; T_4 是一列基本存储电路上共同的预充电管,以控制对输出电容 C_D 的预充电。

写入操作时,写选择线上为高电平, T_1 导通。待写入的信息由写数据线通过 T_1 加到 T_2 管的栅极上,对栅极电容 C_g 充电。若写入 1,则 C_g 上充有电荷;若写入 0,则 C_g 上无电荷。写操作结束后, T_1 截止,信息被保存在电容 C_g 上。

读出操作时,先在 T_4 管栅极加上预充电脉冲,使 T_4 管导通,读数据线因有寄生电容 C_D 而预充到 1(V_{DD})。然后使读选择线为高电平, T_3 管导通。若 T_2 管栅极电容 C_g 上已存有 1 信息,则 T_2 管导通。这时,读数据线上的预充电荷将通过 T_3 、 T_2 而泄放,于是,读数据线上为 0。若 T_2 管栅极电容上所存为 0 信息, T_2 管不导通,则读数据线上为 1。因此,经过读操作,在读数据线上可以读出与原存储相反的信息。若再经过读出放大器反相后,就可以得到原存储信息了。

对于三管动态基本存储电路,即使电源不掉电, C_g 的电荷也会在几毫秒之内逐渐泄漏掉,而丢失原存 1 信息。为此,必须每隔 1~3ms 定时对 C_g 充电,以保持原存信息不变,此即动态存储器的刷新(或称再生)。

刷新要有刷新电路,若周期性地读出信息,但不往外输出(这由读信号 $\overline{RD}$ 为高电平来保证),经三态门(由刷新信号 $\overline{RFSH}$ 为低电平时使其导通)反相,再写入 C_g ,就可实现刷新。

2) 单管动态基本存储电路

单管动态基本存储电路如图 5-9 所示,它由 T_1 管和寄生电容 C_s 组成。

写入时,使字选线上为高电平, T_1 管导通,待写入的信息由位线 D(数据线)存入 C_s 。

读出时,同样使字选线上为高电平, T_1 管导通,则存储在 C_s 上的信息通过 T_1 管送到 D 线上,再通过放大,即可得到存储信息。

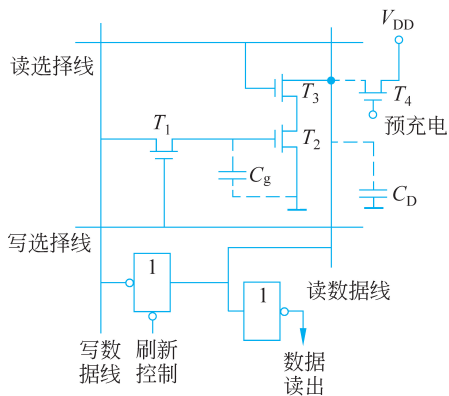


图 5-8 三管动态基本存储电路

为了节省面积,电容 C_s 不可能做得很大,一般使 $C_s < C_D$ 。这样,读出 1 和 0 时电平差别不大,故需要鉴别能力高的读出放大器。此外, C_s 上的信息被读出后,其已存的电压由 0.2V 下降为 0.1V。这是一个破坏性读出,要保持原存信息,读出后必须重写。因此,使用单管电路,其外围电路比较复杂。但由于使用管子最少,4KB 以上容量较大的 RAM,大多采用单管电路。

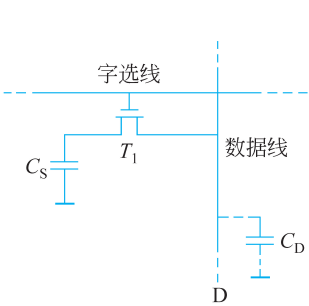


图 5-9 单管动态基本存储电路

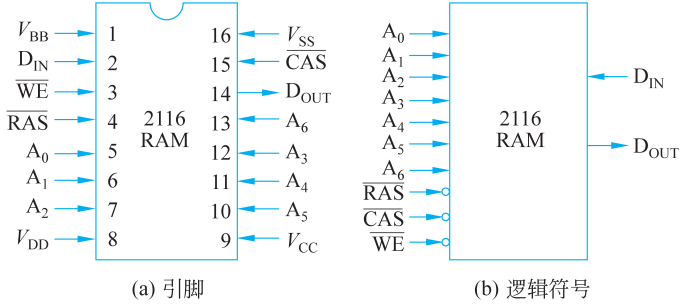


图 5-10 Intel 2116 引脚及逻辑符号

2. 动态 RAM 芯片举例

Intel 2116 单管动态 RAM 芯片的引脚和逻辑符号如图 5-10 所示。

Intel 2116 的引脚名称如表 5-2 所示。

表 5-2 Intel 2116 的引脚名称

引 脚	名 称	引 脚	名 称
$A_0 \sim A_6$	地址输入	$\overline{WE}$	写(或读)允许
$\overline{CAS}$	列地址选通	V_{BB}	电源(-5V)
$\overline{RAS}$	行地址选通	V_{CC}	电源(+5V)
D_{IN}	数据输入	V_{DD}	电源(+12V)
D_{OUT}	数据输出	V_{SS}	地

Intel 2116 芯片的存储容量为 $16K \times 1$ 位(简写 $16K \times 1$)，需用 14 条地址输入线，但 2116 芯片只有 16 条引脚。由于受封装引线的限制,只用了 $A_0 \sim A_6$ 7 条地址输入线,数据线只有 1 条(1 位),而且数据输入(D_{IN})和输出(D_{OUT})端是分开的,它们有各自的锁存器。写允许信号 $\overline{WE}$ 为低电平时允许写入,为高电平时可以读出,如表 5-2 所示,它需要 3 种电源。

Intel 2116 芯片的内部结构如图 5-11 所示。

为了解决用 7 条地址输入线传送 14 位地址码的矛盾,2116 芯片采用地址线分时复用技术,用 $A_0 \sim A_6$ 7 根地址线分两次将 14 位地址按行、列两部分分别引入芯片,即先把 7 位行地址 $A_0 \sim A_6$ 在行地址选通信号 $\overline{RAS}$ 有效时,通过 2116 芯片的 $A_0 \sim A_6$ 地址输入线送至行地址锁存器,而后把 7 位列地址 $A_7 \sim A_{13}$ 在列地址选通信号 $\overline{CAS}$ 有效时,通过 2116 芯片的 $A_0 \sim A_6$ 地址输入线送至列地址锁存器,从而实现了 14 位地址码的传送。

7 位行地址码经行译码器译码后,某一行的 128 个基本存储电路都被选中,而列译码器只选通 128 个基本存储电路中的一个(即 1 位),经列放大器放大后,在定时控制发生器及写

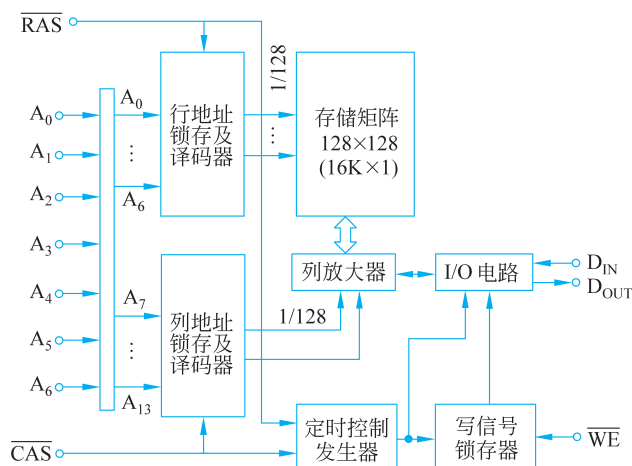


图 5-11 Intel 2116 芯片内部结构框图

信号锁存器的控制下送至 I/O 电路。

Intel 2116 芯片没有片选信号 $\overline{CS}$, 它的行地址选通信号 $\overline{RAS}$ 兼作片选信号, 且在整个读、写周期中均处于有效状态, 这是与其他芯片的不同之处。

此外, 地址输入线 $A_0 \sim A_6$ 还用作刷新地址的输入端, 刷新地址由 CPU 内部的刷新寄存器 R 提供。

与 Intel 2116 芯片类似的还有 2164、3764、4164 等 DRAM 芯片。

综上所述, 动态基本存储电路所需管子的数目比静态的要少, 提高了集成度, 降低了成本, 存取速度快。但由于要刷新, 需要增加刷新电路, 外围控制电路比较复杂。静态 RAM 尽管集成度低一些, 但静态基本存储电路工作较稳定, 也不需要刷新, 所以外围控制电路比较简单。究竟选用哪种 RAM, 要综合比较各方面的因素决定。

5.3 只读存储器

只读存储器(read only memory, ROM)只能读出原有的内容, 不能由用户再写入新内容。原来存储的内容是采用掩膜技术由厂家一次性写入的, 并永久保存下来。它一般用来存放专用的固定程序和数据, 不会因断电而丢失。

5.3.1 只读存储器存储信息的原理和组成

ROM 的存储元件如图 5-12 所示, 它可以看作一个单向导通的开关电路。当字线上加有选中信号时, 如果电子开关 S 是断开的, 位线 D 上将输出信息 1; 如果 S 是接通的, 则位线 D 经 T_1 接地, 将输出信息 0。

ROM 的组成结构与 RAM 相似, 一般也是由地址译码电路、存储矩阵、读出电路及控制电路等组成。图 5-13 是有 16 个存储单元、字长为 1 位的 ROM 示意图。16 个存储单元, 地址码应为 4 位, 因采用复合译码方式, 其行地址译码和列地址译码各占 2 位地址码。对某一