



微 处 理 机 基 础

(中 册)

福建省闽东无线电厂

第六章 半导体存储器

§ 1. 半导体存储器及其分类简介

什么是存储器，存储器是存储指令和数据的地方，存储器有两种基本形式，普通微处理机中最常用的一种存储器是随机存取存储器，RAM。RAM存储器允许使用者写入和读出任何存储单元的信息。另一种是只读存储器 ROM，ROM只允许从存储单元中读出数据，但不允许新的数据写入该存储单元，ROM是不能用来操作数据的，但是用它来存储常用的程序却是非常方便的。

在微处理器出现前半导体存储器已经开始应用，当今半导体存储器的种类，型号及规格繁多。怎样合理选择是设计中的一个重要方面。

按功能可分为读写存储器及只读存储器。

读写存储器：存放各种现场输入，输出数据，中间计算结果及作为堆栈用。只读存储器：存放各种固定的程序及常数。（例如采用微型系统中的应用程序，通用微型机系统中的汇编程序，外设管理程序，字符转换表）。

1.1 读写存储器

以 Random-Access Memory (RAM) 来命名，本意是指对存储器中任一地址单元的读（写）操作所需时间是相等的。由下图可见，RAM又分静态 RAM 和动态 RAM 两种。而按工艺分可分为 PMOS, NMOS, CMOS 等。（P沟道技术制成 PMOS 晶体管，N沟道以及互

补金属氧化物技术，P：位密度高，速度慢，N：速度快速度2倍于P沟，C：介于P，N两者，功耗低）。

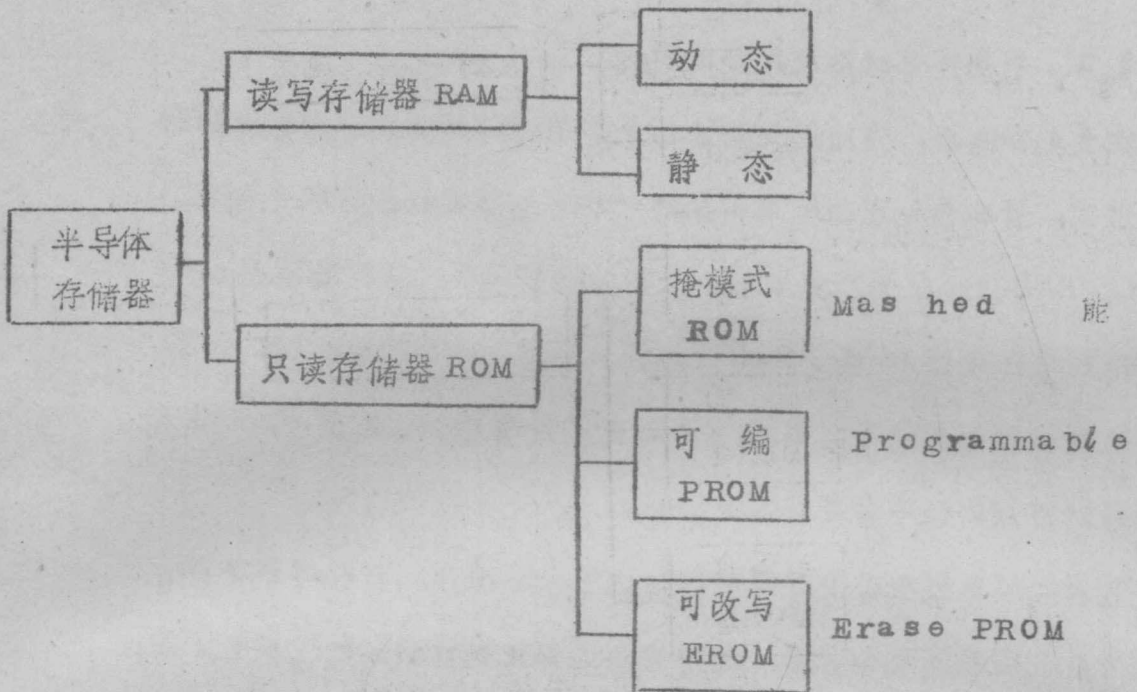


图1 . MOS存储器分类图

(1) 静态 RAM存储器单元一般是由双稳态触发电路构成。因此只要不停电就可以保存其中的信息。

(2) 动态 RAM是用 MOS电路中的栅极电容来存放信息，而这些电容器上的电荷，总会通过一些漏电阻（虽然MOS电路的漏电阻很大）消失，所以为了补充信息电荷的消失，动态 RAM隔一定时间必须重新写入一次原信息，这样的操作称为动态 RAM的再生操作。

(3) 两者比较，静态 RAM在一个触发器里存储一位信息。它的内容只要有电源，就经常保持稳定。动态 RAM是一位信息作为一个电荷来存储的。基本单元比静态 RAM触发器小，因而密度就高得多，另基

本单元的几何形状比较简单，因而速度较高，（64K位，300~500 ns），其缺点是存储电容器里的电荷会泄漏，可在几毫秒内失去大部分电荷，必须每隔1~2ms更新一次，先读出再写入，这样必须配置更新逻辑，并且干扰处理器的执行时间，对小型系统，通常采用静态RAM。

1.2 只读存储器

Read Only Memory (ROM) 来命名，由于只读不写，因而电路比起RAM来要简单得多，一般一个芯片内集成的位(bit)要比RAM大得多，在价格及芯片面积上都比RAM占有优势，因而微型系统尽量压缩RAM容量，尽可能多使用ROM，但ROM的特点就只能读，要改变其中的信息是非常困难的，只读存储器又分三类：

(1) 掩膜型ROM：在生产过程中一道掩膜工艺决定了其中的信息，而一旦生产完毕就不可能再改变信息，这适合于程序成熟且生产数量大（如：上万片）的产品。由于大量生产，生产周期以及投产后不能改，有错就必须整个交换，因此出现了其它型式的ROM。

(2) 可编程 PROM：是由用户，在专用设备下写入用户自己所需要的信息，但只能写入一次，它适合于小批量产品。

(3) 可改写 EPROM：在特定的条件下，花费较长时间进行的，改写次数可以任意多次，但由于特定条件下改写，因而应用时还是一只读存储器。

（以上几种只读存储器按工艺分，制作上有三种主要技术即双极技术，金属—氧化物—半导体 (MOS) 技术，金属氧化物氧化物半导

体)。

§ 2 . 随机存储器 (RAM)

2.1 对 RAM的要求

在微处理机中 RAM 用来组成主存储器, 对 RAM 的主要要求有:

- (1) 存取速度: 作为主存, 根据 CPU 的速度一般希望取数时间 $100 \sim 500 \text{ ns}$ 范围。
- (2) 体积小: 主存容量大, 它体积常在整个计算机中占有很高比例。
- (3) 功耗低: 减少每位存储器的功率消耗, 才能有效的减少大容量存储器的耗电功率, 尤其停电时存储信息的保存必须采用备用电池降低功耗更为重要。
- (4) 价格低: 存储器的成本在微型计算机中占有相当高的比例。
- (5) 可靠性高: 一种器件在系统中的用量越大则对其可靠性也就要求越高, 现存储单元是存储器的核心组成部分, 一个 1024 位 MOS RAM 芯片中就有 1024 个重复的单元电路, 所以存储单元的面积, 功耗以及稳定性对整个 MOS RAM 有决定性影响。在高集成度的 MOS RAM 工业产品在一个芯片上已经能达到 $16\text{K} \sim 64\text{K}$ 位。高速的已达到 $100 \mu\text{s}$ 以下, 低功耗的达到 1 mW/bit 以下, 与双极性的 RAM 相比, 它具有集成高, 功耗低, 制作方便, 价格便宜等优点。

动态 RAM 电路, 由于密度较高, 因而比静态 RAM 便宜。然而动态 RAM 要求附加一个更新电路, 至少在微处理器 (CPU) 内纳入更新

电路以前必须如此。对于小型系统如果不打算在日后使扩大成较大的存储系统的话，通常用静态 RAM 是最经济的。

静态 RAM 用双稳态触发器存储数据，由于数据锁存在这些触发器中，所以无须重写电路。这就是说只要电源不断数据就能无限期的保存而无需重写。

现在许多不同种类的静态 RAM，小些的静态 RAM 是用双极性 TTL 工艺，大部分大的 RAM 是用 MOS 工艺。最流行的静态 RAM 容量为 1024 位。这样大小的 RAM 包装在单块 IC 中，有 16 到 34 个引出端，在内部 1024 存储单元可以排列 128 字，每个字有 8 位，因此一个 RAM 为 512 字节之微处理机系统要求有 4 块这样的 IC。这种类型 IC 要求用 8 个引出端作为数据线。鉴于这种情况，经常采用 24 个引出端。

另外一种较为流行的排列为 1024×1 RAM。这种电路图每片只有一个数据线，8 字节就需要 8 个这样的片子。

2.2 静态 RAM 存储单元

图 2 表示了一个 MOS 型静态 RAM 的基本存储单元。这个单元只能存储一个数据。一块单元 IC 上可以有 1024 个这种存储单元以及地址译码器和总线驱动器。存储单元本身包含六个 MOS 晶体管，MOS 晶体管采用简化符号。假定这些晶体管是 N—沟道增强 MOS FET (MOS 场效应晶体管) MOS FET 通常是关断或者不通，然而这些 N—沟道设备，它们能够依靠在栅极加正电压的办法，使其导通，也就是说，当栅极接近地电位时，晶体管断开，呈现出一个非常高的阻抗。但是，当栅极为高 (接近 5 V) 时晶体管导通，呈现较低的阻抗。记住这几点，那就很

容易了解该单元的动作情况。

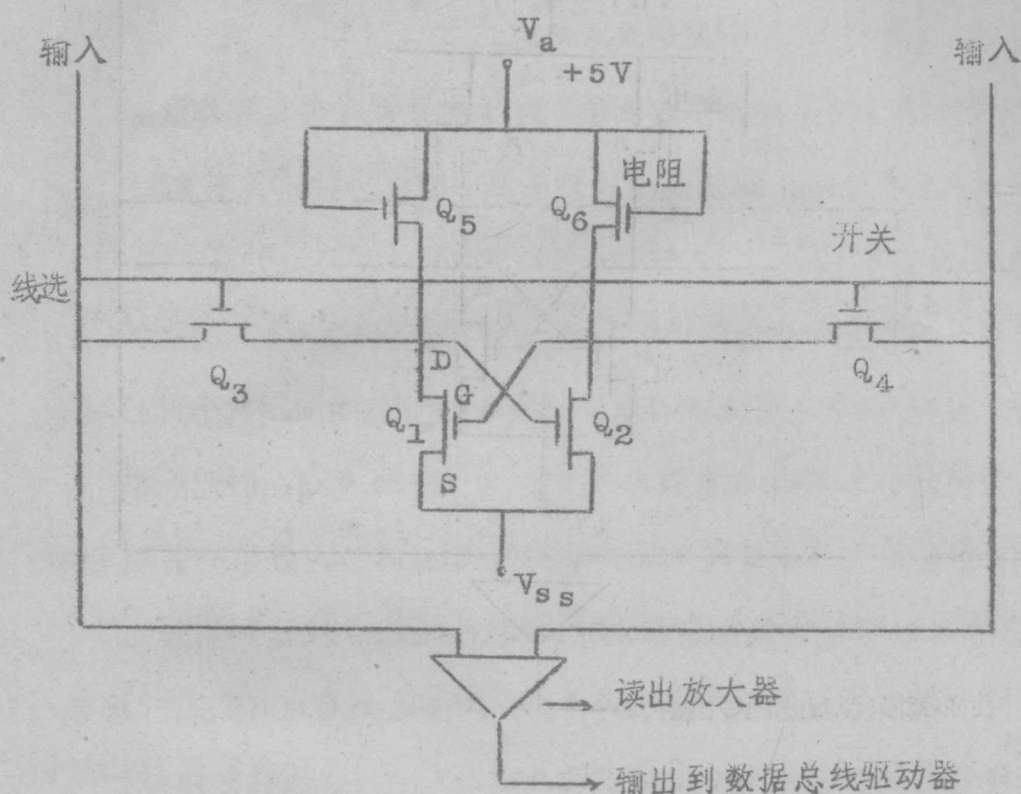


图2 . 静态RAM存储单元

晶体管 Q_1 及 Q_2 交叉耦合，所以它们组成了双稳态锁存器或触发器。位的数据存储在锁存器中。 Q_5 及 Q_6 的作用分别相当于 Q_1 及 Q_2 之电阻。用 MOS FET 来代替具体的电阻，因为这样可以减少片子的空间。 Q_3 和 Q_4 之作用相当于开关，当写入操作时，将输入数据连到锁存器中。同样，当读出操作时，将锁存器中的数据连到输出读出放大器，字选线连到 Q_3 及 Q_4 之栅极。字选线为逻辑 1 时， Q_3 及 Q_4 接通，字选线为 0 时， Q_3 及 Q_4 断开。RAM 存储单元，必须将数据写进去或从中读出来。

(1) 写入操作：将一位数据写入这种单元之前，先得讨论定义“1”和“0”。假定当 Q_2 通， Q_1 断时锁存器内存有二进制的 1。当然当 Q_1 通 Q_2 断时锁存器内存有二进制的 0。当起动电源时，触发器存 1 还是存 0，是不清楚的。我们向单元写入数据是通过控制输入， $\overline{\text{输入}}$ 及字选线来实现的。为了存二进制 1，我们将输入线为 1（正电压）， $\overline{\text{输入}}$ 线通常与输入线呈反状态，所以是逻辑 0（= 0 V），通常向字选线选加瞬时的逻辑 1（正电压）此时就可以将二进制存入。

正脉冲加在字选线后将使开关 Q_3 及 Q_4 导通，因此输入线上的正电压通过 Q_3 加到 Q_4 之基极（栅极）上，致使 Q_2 导通，当 Q_2 导通时，漏电压降低，降低后的电压传到 Q_1 栅极，使 Q_1 断开。当 Q_1 断开，它的漏电压变高，这个增加了的电压传到 Q_2 的栅极上，使 Q_2 保持在这个导通状态。

当字选线转为逻辑 0 时， Q_3 及 Q_4 断开，然而， Q_2 的导通使得 Q_1 继续断开，反之 Q_1 高，漏电压使得 Q_2 继续导通。因此二进制的“1”被锁存在触发器上。它将一直保存，直到电源断开或要写入二进制为止。

以后我们能够写 0，如果我们将逻辑 0 加到输入线，将逻辑 1 加到 $\overline{\text{输入}}$ 线以及将脉冲加到字选线的话，当字选线变高， Q_4 导通，将逻辑 1 从输入加到 Q_1 之栅极，使得 Q_1 离开截止区，同时 Q_3 导通，将逻辑 0 加到 Q_2 的栅极上，使 Q_2 趋向截止。此时触发器锁存了相反的状态，这个状态表示了二进制的 0。当用 R/W 线时（未画出），读出放大器及输出线（见图下部）被禁止。

(2) 读出操作: 输入及输入线三态线, 它们可以被 R/W 线使能或禁止。当 R/W 为高时 (未画出) 则输入及输入线有效地被脱开。在这周期中通过加脉冲在字选线上, 能够从存储单元读出数据, 假定现在单元中存的是逻辑 1, 这意味着 Q_1 截止, Q_2 导通, 因此, 当 Q_2 的漏电压为低时, Q_1 的漏电压为高 (逻辑 1)。

当字选线转为高时, Q_3 导通, 将 Q_1 漏上的高电压连到读放大器之左边的输入端, 与此同时, Q_4 导通, 将 Q_2 漏上的低电压连到读出放大器之右边的输入端。读出放大器译出逻辑 1 并相应地将输出线置 1。如果在加字线脉冲时, 触发器为逻辑 0, 读放右边输入端接收高电压, 左边输入端接收低电压, 读放译出逻辑 0。

2.3 128×8 RAM

图 3 展示 31 个由 1024 个单元组成的 128×8 RAM, 每个方块代表一个刚刚讲过的六个晶体管。

字选线在左边的输入端, 只画了 4 个, 实际上是 128 个——每个代表一个字。

字选线 00 连到 8 个存储单元中的每一个 (在图的上面横过), 在实际系统中, 这 8 个存储单元可以组成 8 位字节, 我们称它为 0000₁₆ 单元。

输入线画在图的顶部, 为简化计, 8 条线中画了 4 条。注意, 输入线是经过反相的, 因此“输入非”能够加到每个单元上。虽然详细的没有画, 但是输入及输入线是三态线, 所以不写操作时, 它们是有效地被脱开的。

输出线画在图的低部，当写操作时，它们被禁止；在读操作时，使能。至每个输出线上均有一个读放大器。

记住图并未将 RAM 画全，仅仅画了存储器矩阵及读出放大器。将 RAM 转入工作状态负载要附加电路，其中之一就是地址译码器。

存储器阵列是按 128 字节排列的。一个地址译码器能从 128×8 的存储单元中选出任意一个单元。因此，要求有 128 选 1 译码器，从 CPU 向译码器输入 7 根地址线。我们知道 7 位能够确定出 128 个不同的地址。地址译码器内 128 个 7 输入端的与门组成。图 4 中只画了三个门；第一，二两个及最后一个。当地址线 A_0 —— A_6 为低时，字选线 00 为高，注意七个反相器产生 A_0 —— A_6 。它们的“非”是输入到顶上面的与门上的。如果 A_0 —— A_6 全为低，则 A_0 —— A_6 全为高，因此与门输出变高。由此可见当低位地址为 00000000_2 时，字选线 00 被选中。

当地址为 10000001_2 时，字选线 01 起作用。下 125_{10} 个门没画出。当 A_0 —— A_6 全为高时，字选线 127_{10} 被选中，即地址为 11111111_2 时，字选线 127_{10} 起作用。

大部分 RAM 并没有单独的输入及输出线，代表它们的是数据线，该数据线不是作为输入线用就是作为输出线用。这是可能实现的，因为 CPU 不能同时作读出及写入操作。

图 5—A 是个简化的佈置，数据线在左边。依靠在写入线上加高信号的办法，使三态输入缓冲器使能。该写入线由 R/W 信号及片选 (CE) 信号所控制。当 CPU 向存储器写数据时，写入线为高，输入缓冲器使

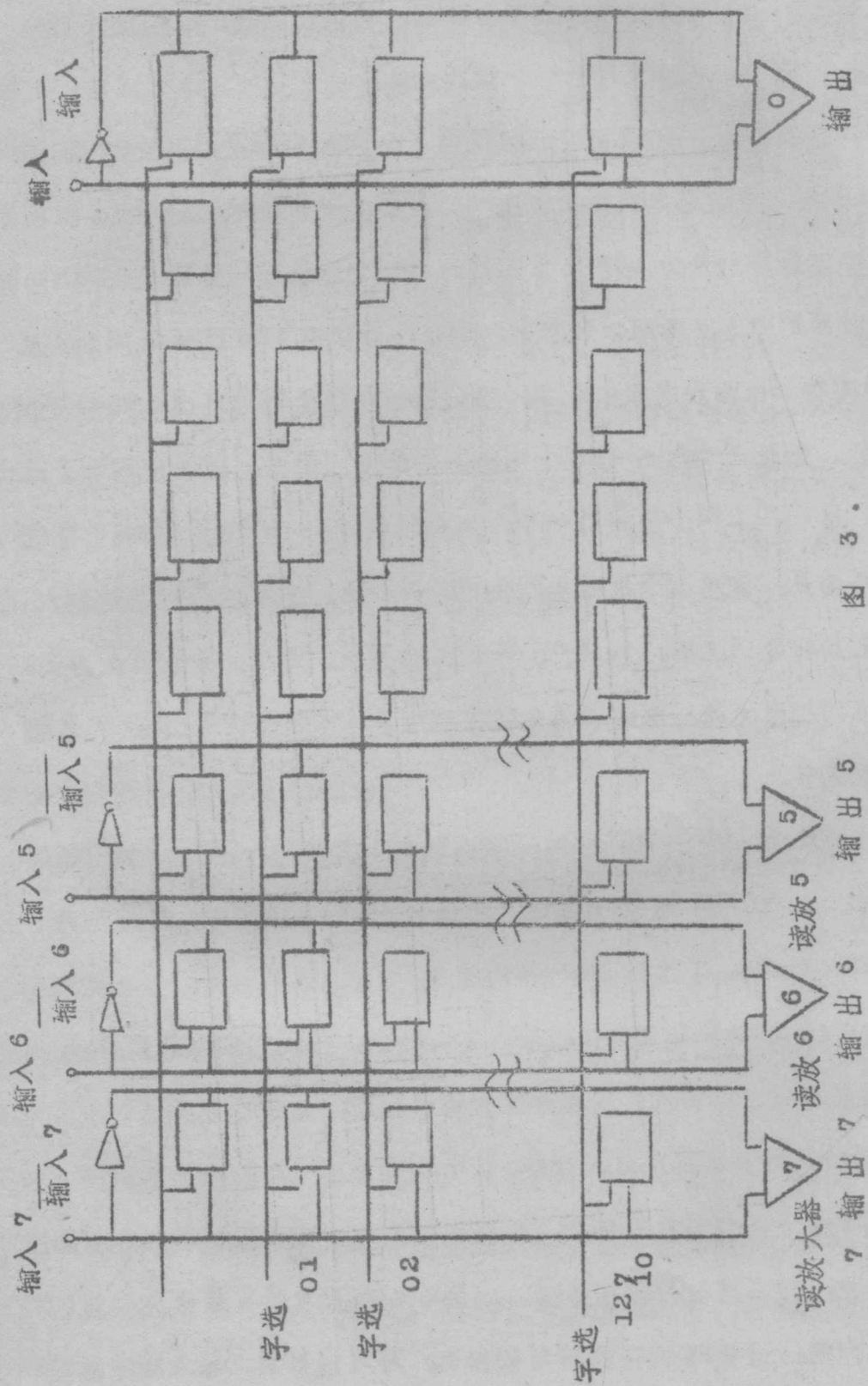


图 3.

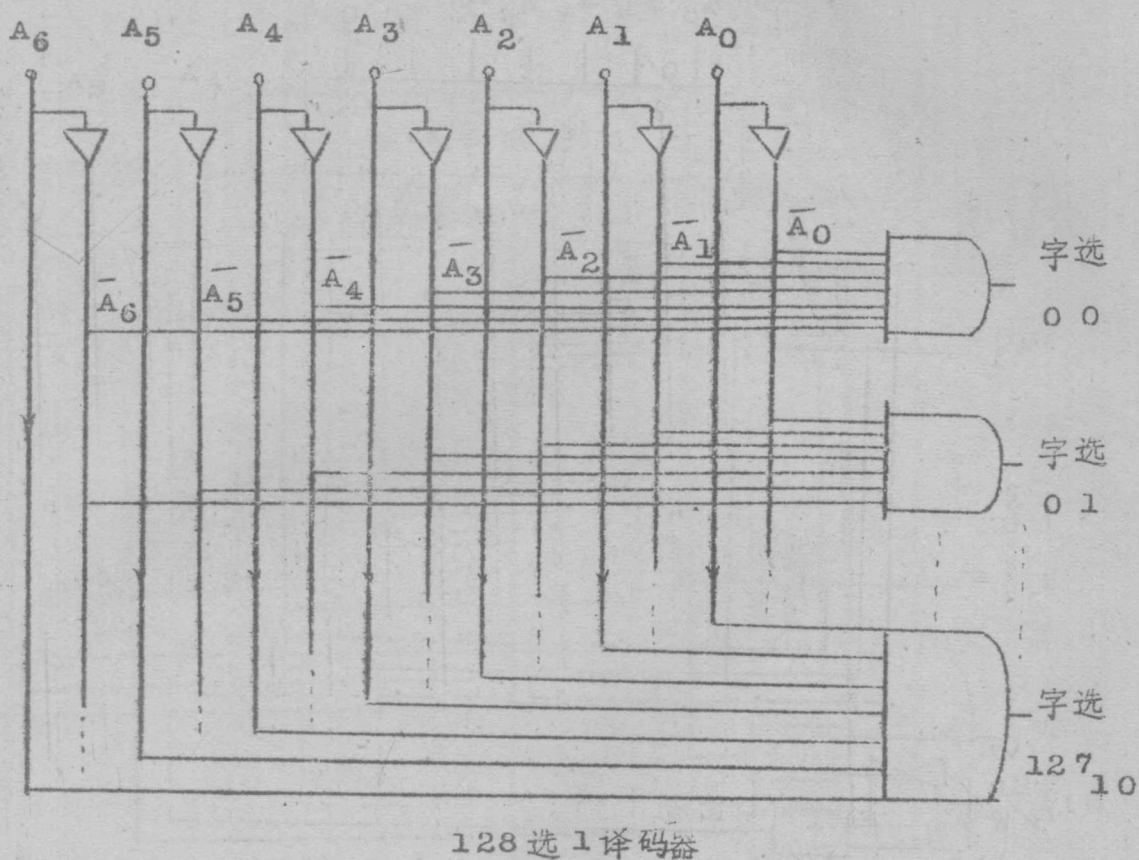


图 4

能，允许数据写入被选中的地址中去，在这个读出线为低信号的周期中，输出缓冲器被禁止。

当数据从 RAM 读出时，读出线转为高，写入线转为低。输入缓冲区禁止，输出缓冲区使能，于是选中地址上的数据被读出并放到数据总线上。

由图 5—A 可见，读出信号及写入信号被片使能 (CE) 信号及 R/W 线所控制。当这特定的存储器片子被选中时，CE 输入线变高。如 CE

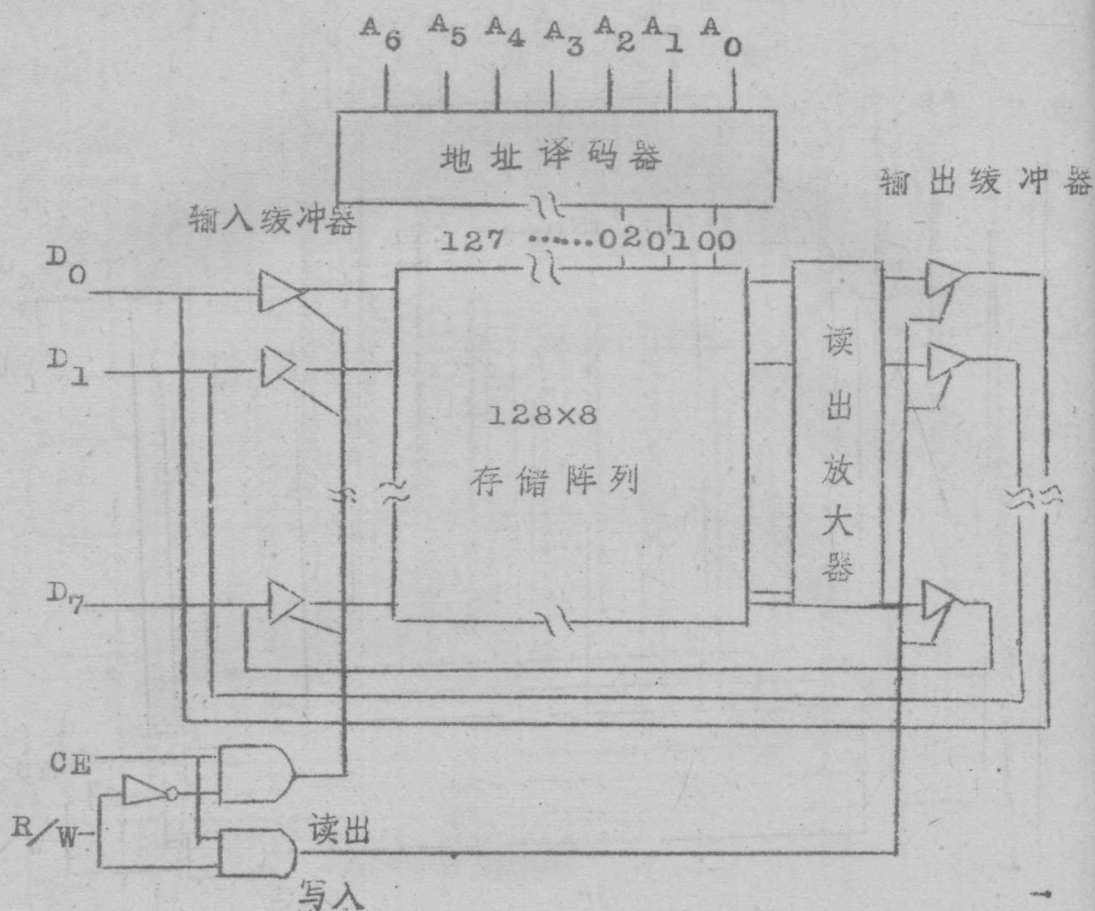


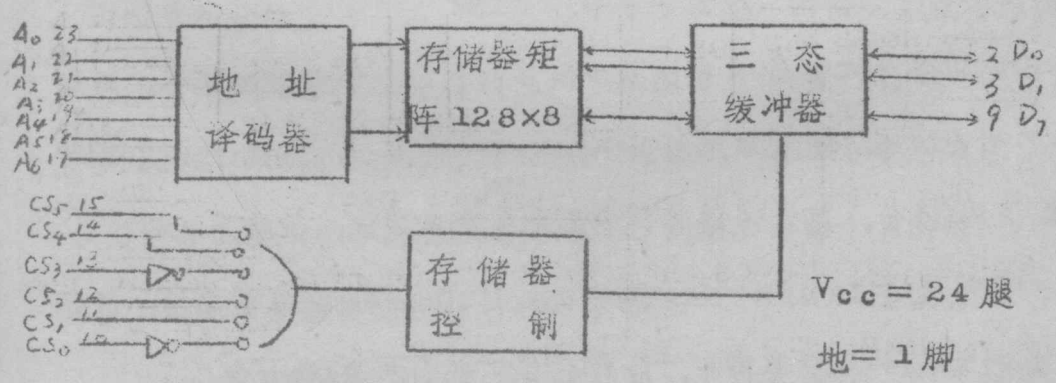
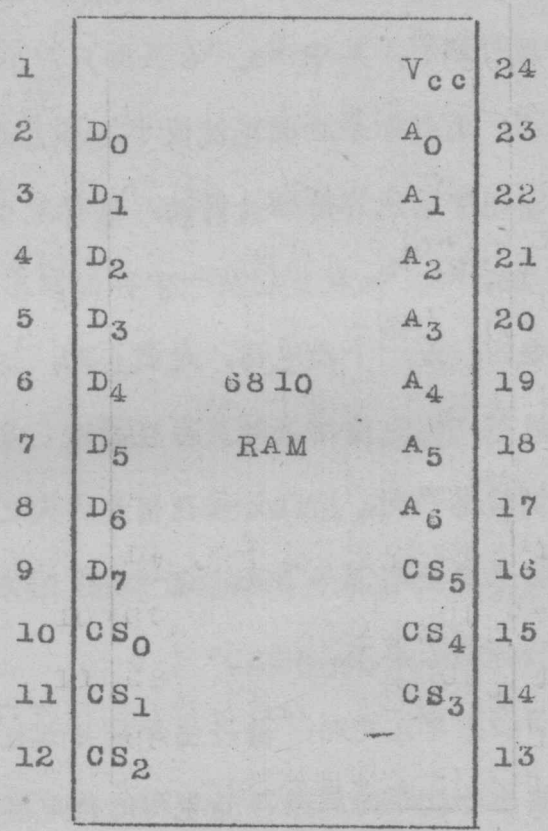
图5—1. 有双向数据线的 RAM

线为低，则门1及2被禁止。至使读出及与入信号两者均变低，输入、输出缓冲器均被禁止。其效果是，这块片子与数据总线脱离。

然而当CE为高，则R/W线控制读出及与入信号。该R/W线是连到CPU的R/W线。我们知道当CPU将数据与入RAM时，R/W线为低，当CPU将数据从RAM读出时，R/W线为高。当R/W为高时，则门2输出变高，使输出缓冲器使能。门1输出保持低，使输入缓冲器禁止。CPU处于读出状态。

当 R/W 变低, 111 输出变高, 112 输出变低。RAM 处于写入状态。有些 RAM 只有一根单独的片使能线。在许多 RAM 中片使能线标为 CE, 其意是当使能为低时, 该片被选中, 有一些 RAM 则有几个片使能 (CE) 或片选 (CS) 线。

6810 是一片 128x8
RAM



16 读出 / 写入 R/W

图 5—B. 6810 128x8 RAM

2.4 动态 MOS 单元

从上面静态 MOS 单元可看出为了存储信息“1”和“0”只要 Q_1 Q_2 两个管子，一个开放，一个截止，并把它们用门管 Q_3 、 Q_4 与外部隔断开就可以了，其中 Q_5 、 Q_6 实际只起不断补充存储管栅电容电荷泄漏的作用，显然如果泄漏电流很小，那么即没有直流电流来补充，栅电容的存储电荷仍能保持相当时间，这样就可以省略负载管 Q_5 、 Q_6 ，只要 Q_1 、 Q_2 、 Q_3 、 Q_4 就能组成一个存储单元。 Q_2 、 Q_3 栅电容一个有电荷是导通的，另一个没电荷，是截止的，这代表一种状态，反之代表另一种状态。但是由于实际上存在漏电（通过 PN 结等）栅电荷逐渐减少，而使信息丢失，所以必须在信息丢失之前，对它进行周期性地充电，才能把信息安全地保存下来，这一过程叫做再生，这种再生操作的存储单元，叫做动态单元。

四管动态单元电路，除再生外读写过程与静态 MOS 单元基本一致。实际上读出一次即向原来存储电荷的栅电容充电一次。所以对这种四管动态 MOS 单元的再生是方便的，只要对该单元进行一次读出再写入操作就行。

四管动态与六管 MOS 相比减少很贵面积的两个负载晶体管，同时由静态改为动态，因此单元面积和功耗都大大减少，适用于大规模的 MOS RAM 器件，四管还保持了六管单元比较稳定，速度可作得快等优点，因此现在 1KMOS RAM 部分 4K MOS RAM 器件被采用，但和其它动态 RAM 一样需要再生，应用起来不如静态 RAM 方便。

一般动态 MOS 保持信息时间为 2ms 左右，所以在 2ms 以内必须

把存储器存储的每一个地址单元进行一次再生操作,当然实现再生操作须要的控制电路较复杂,不过由于动态存储器芯片比静态存储器芯片在集成度、功耗、价格等方面都优势,所以在大容量存储体上应用还是有利的。

2.5 RAM与CPU之间的联接

在微处理机系统中,CPU对存储器进行读写操作,首先要给读写单元地址。(通过地址总线)及各种控制信号(以及写入时的数据),接着存储器才能进行读写操作。

RAM存储器芯片具有一位、四位、八位等不同的结构,例如1K位存储器芯片,可能具有 1024×1 位、 256×4 位、 128×8 位等不同结构,因此8位微处理器相联时,要采用位并联或地址串联的方法达到存储体所需要的容量及位数。用8个 1024×1 组成 1024×8 存储体。

图给出了 1024×1 位组成的 1024×8 位存储体与CPU地址总线,数据总线相连接的方式。

如图6所示由于每片有1024个可寻址单元,因而在地址容量上已满足存储体的要求,但每片只有一个地址单元中的一位,因而在位数上要用八片并联组成一个地址单元的八位数,由CPU来的地址总线 $A_0 \sim A_9$ 送往八片中的每一片的地址输入端,而每一片的输入,输出端(I/O)分别联到数据总线的各位。

(RAM中以其它控制信号图中略)

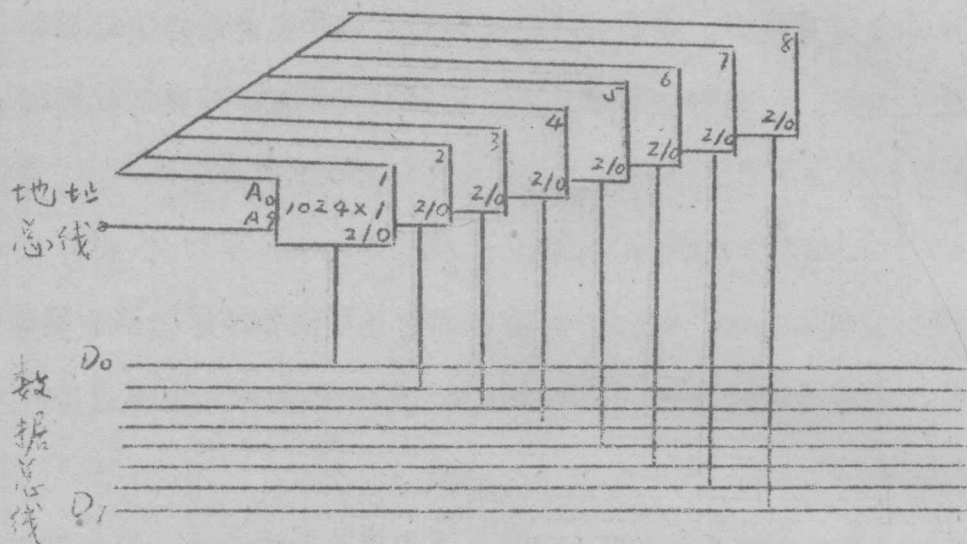


图 6 .

2. 256×4 位组成 1024×8 存储体

由于每片有四个输出端，即每个可寻址单元同时有四位数据可读出（写入）所以八位微处理器的数据总线相联时，有 2 片在位上并联即可组成一个可寻址单元的八位数。由于每个 RAM 芯片只有 256 个可寻址单元，因而在组成 1024 个地址的存储体时，必须由四片实行地址串联，地址串联的实现方法是用高位地址（在这个例子中是以 A_8 、 A_9 ）经过译码输出控制 RAM 芯片的选片输入端。这样如图中的四组（12, 34, 56, 78）只有一组是被选中而能进行读写。例如在读操作时其它三组未被选中的芯片，它们的 I/O 端处于浮动状态（三态电路里呈高阻）数据总线的电平只取决于被选中的一组的 I/O 端输出电平。

由上面两例可知，组成一定容量及位数的存储器可以由不同结构的芯片采用地址串联，位数并联的不同方案实现，在进行方案比较时，除