

第5章 微机的存储器

学习目标

存储器是微机的重要组成部分之一,它的种类很多,各种存储器存储信息的媒体、存储原理和方法也各不相同。本章主要以在微机中广泛应用的半导体存储器为对象,在讨论存储器及其基本电路、基本知识的基础上,着重研究存储芯片及其与 CPU 之间的连接与扩充问题。此外,简要介绍了内存条技术的发展、外存储器和存储器系统分层结构等。

学习要求

- 了解存储器的分类、组成及功能。着重理解行选与列选对 1 位信息的读出。
- 重点掌握位数扩充与地址扩充技术。
- 熟练掌握存储器与 CPU 的连接方法。
- 了解内存条技术的发展现状。
- 了解两种不同接口类型的硬盘。
- 着重理解存储器系统的分层结构。

5.1 存储器的分类与组成

存储器用来存储计算机的指令、数据和各种信息。按存储器和 CPU 的连接方式不同,它可分为内存储器(主要由半导体存储器构成)和外存储器(主要由磁、光存储材料构成)。通过 CPU 的外部总线直接与 CPU 相连的存储器称为内存储器(简称内存或主存)。CPU 要通过 I/O 接口电路才能访问的存储器称为外存储器(简称外存)。

5.1.1 半导体存储器的分类

半导体存储器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 在程序执行过程中,每个存储单元的内容根据程序的要求既可随时读出,又可随时写入,故可称读/写存储器。RAM 所存储的信息在断开电源时会立即消失,是一种易失性存储器。RAM 主要用来存放用户程序、原始数据、中间结果,也用来与外存交换信息和用于堆栈等。

ROM 在程序执行过程中,对每个存储单元的原存信息,只能读出,不能写入。ROM 在断开电源时,所存储的信息不会丢失。ROM 常用来存储固定的程序,如监控程序、汇编程序以及 BIOS 程序等。

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。

只读存储器 ROM 按工艺也可分为双极型 ROM 和 MOS 型 ROM,但一般根据信息写入的方式不同,而分为掩膜式 ROM、可编程 PROM 和可擦除、可再编程 EPROM 以及 Flash ROM(快擦除 ROM 或闪速存储器)等几种。其中,Flash ROM 是一种新型的非易失性存储器,它可以用软件在较高电压下进行改写或在线写入,一旦写入即相对固定。由于 Flash ROM 的单片容量大,易于修改,所以目前广泛用于主板的 ROM BIOS 及其他 ROM 中。

5.1.2 半导体存储器的组成

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,对每个存储单元要赋予一个编号,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。

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 之间的关系是: $2^n = N$ 。

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 N 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。

地址译码方式有单译码方式和双译码方式两种。

1) 单译码方式(或称字结构)

单译码方式的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。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)

双译码方式是将地址码分为 X 与 Y 两部分,用两个译码电路分别译码。X 向译码又

称为行译码,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 向译码又称为列译码,其输出线称列选择线,它选中一列的所有单元。只有 X 向和 Y 向的选择线同时选中的那一位存储单元,才能进行读或写操作。由于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

5.2.1 静态随机存取存储器

1. 静态 RAM 基本存储电路

静态 RAM 的基本存储电路,是由 6 个 MOS 管组成的 RS 触发器。每一个触发器就构成存储体的一位。

2. 静态 RAM 的组成

静态 RAM 的结构一般由存储体、译码电路和控制电路组成。

存储体由存储矩阵构成。在存储矩阵中,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,才能进行读出信息和写入信息的操作。

存储体中的每一位,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 I/O 电路也应有 4 个或 8 个。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 位信息将被一次读/写完毕。

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。一般地,片选信号由地址码的高位译码(通过译码器输出端)产生。

3. 静态 RAM 的读/写过程

1) 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将被打开,所存信息送至 DB 上。于是,存储单元中的信息被读出。

2) 写入过程

(1) 同上述读出过程(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。

需要着重理解,无论是存储器读还是存储器写,都必须保证能从存储器中读到或写入到存储器中一个稳定的数据。在读/写信号有效期间,数据线和地址线也必须是稳定的。

4. 静态 RAM 芯片举例

常用的静态 RAM 即 SRAM 芯片有 2114、2142、6116、6264 等。

本课程选用了 Intel 6116 和 6264。6116 芯片是一个 $2K \times 8$ 位的 CMOS SRAM 芯片,属双列直插式、24 引脚封装,它的存储容量为 $2K \times 8$ 位。6264 芯片的结构及工作原理与 6116 相似,它是一个存储容量为 $8K \times 8$ 位的 CMOS SRAM 芯片,其引脚有 28 条。

5.2.2 动态随机存取存储器

动态 RAM 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,大多采用 DRAM 芯片。

1. 动态基本存储电路

这里重点介绍常用的三管和单管这两种基本存储电路。

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 管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,信息被保存在电容 C_g 上。

读出操作时,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,必须每隔 $1\sim 3\text{ms}$ 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)。

刷新要由刷新电路来实现。

2) 单管动态基本存储电路

单管动态基本存储电路仅由一个 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,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。由于单管动态基本存储电路使用的管子少,4KB 以上容量较大的 RAM,大多采用单管电路。

2. 动态 RAM 芯片举例

Intel 2116 是单管动态 RAM 芯片。它的存储容量为 $16\text{K}\times 1$ 位。本来这需用 14 条地址输入线,但由于受封装引线的限制,2116 只有 16 条引脚,其中只有 7 条地址输入线。

为了解决用 7 条地址输入线传送 14 位地址码的矛盾,2116 采用地址线分时复用技术。

2116 芯片没有片选信号 $\overline{\text{CS}}$,它的行地址选通信号 $\overline{\text{RAS}}$ 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。

5.3 只读存储器

5.3.1 只读存储器存储信息的原理和组成

ROM 的存储元件可以看作一个单向导通的开关电路。

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部分组成。

5.3.2 只读存储器的分类

1. 不可编程掩膜式 MOS 只读存储器

不可编程掩膜式 MOS ROM 又称为固定存储器。它是由器件制造厂家根据用户事

先编好的机器码程序,把 0、1 信息存储在掩膜图形中而制成的 ROM 芯片。这种芯片制成以后,它的存储矩阵中每个 MOS 管所存储的信息 0 或 1 被固定下来,不能再改变,而只能读出。如果要修改其内容,只有重新制作。

2. 可编程只读存储器

为了克服上述掩膜式 MOS ROM 芯片不能修改内容的缺点,设计了一种可编程序的只读存储器(programmable ROM, PROM),用户在使用前可以根据自己的需要编制 ROM 中的程序。

3. 可擦除、可再编程的只读存储器

PROM 芯片虽然可供用户进行一次修改程序,但仍有局限。为了便于研究工作,试验各种 ROM 程序方案,研制出了一种可擦除、可再编程的 ROM,即 EPROM(erasable PROM)。在 EPROM 芯片出厂时,它是未编程的。若 EPROM 中写入的信息有错或不需时,可利用专用的紫外线灯对准芯片上的石英窗口照射 15~20min,即可擦除原写入的信息。

近几年来,采用金属—氮—氧化物—硅(MNOS)工艺生产的 MNOS 型 PROM,是一种利用电来改写的可编程只读存储器,即 EEPROM(或称 E²PROM)。用这种方法改写数万次,只需要 0.1~0.6s,信息存储时间可达十余年之久,这给需要经常修改程序和参数的应用领域带来极大的方便。但是,EEPROM 有存取时间较慢、完成改写程序需要较复杂的设备等缺点,现在正在迅速发展高密度、高存取速度的 EEPROM 技术和 Flash Memory(闪存)技术。

5.3.3 EPROM/E²PROM 常用芯片举例

1. Intel 2716 芯片

2716 EPROM 芯片的容量为 2K×8 位,采用 NMOS 工艺和双列直插式封装,有 24 条引脚。

2716 芯片的工作方式有:读出方式,未选中、待机、编程输入、校验编程内容与禁止编程。

与 2716 芯片属于同一类的常用 EPROM 芯片还有 2732、2764、27128、27256、27512 及 271024 等,它们的内部结构与外部引脚分配基本相同,主要是存储容量逐次成倍递增为 4K×8 位、8K×8 位、16K×8 位、32K×8 位、64K×8 位及 128K×8 位等。

2. Intel 2732 芯片

2732 EPROM 芯片的容量为 4K×8 位,采用 HN MOS-E(高速 NMOS 硅栅)工艺制造和双列直插式封装。

2732 EPROM 芯片也是 24 条引脚,与 2716 芯片相似,只是将其 21 引脚由 V_{PP} 改为 A_{11} ,20 引脚由 $\overline{CS}$ 改为 $\overline{OE}$ 与 V_{PP} 共用,而将 18 引脚由 PD/PGM 改为 $\overline{CE}$,其他引脚功能不变。

2732 EPROM 的工作方式也有读出、输出禁止、待机、编程输入、编程禁止和读标识码等多种。

与 2732 芯片属于同一类的常用 EPROM 芯片还有 2764、27128、27256、27512 及 271024 等,它们的内部结构与外部引脚分配基本相同,主要是存储容量逐次成倍递增为 $4K \times 8$ 位、 $8K \times 8$ 位、 $16K \times 8$ 位、 $32K \times 8$ 位、 $64K \times 8$ 位及 $128K \times 8$ 位等。

3. E²PROM 芯片

常用的 E²PROM 芯片有 2816/2816A、2817/2817A/2864A 等。其中,以 2864A 的 $8K \times 8$ 位的容量为最大,它与 6264 兼容。其主要特点是能像 SRAM 芯片一样在写之前自动擦除原内容。但它并不能像 RAM 芯片那样随机读写,而只能有条件地写入,即只有当一个字节或一页数据编程写入结束后,才可以写入下一个字节或下一页数据。在 E²PROM 的应用中,若需读某一个单元的内容,只要执行一条存储器读指令,即可读出;若需对其内容重新编程,可在线直接用字节写入或页写入方式写入。

5.4 存储器的连接

存储器的连接主要解决两个问题:一是如何用容量较小、字长较短的芯片,组成微机系统所需的存储器;二是存储器与 CPU 的连接方法与应注意的问题。

5.4.1 存储器芯片的扩充

1. 位数的扩充

一块实际的存储芯片,其存储单元的位数(即字长)通常与实际内存单元的字长并不相等。例如,SRAM 芯片 2114 为 $1K \times 4$ 位,DRAM 芯片 2164 为 $64K \times 1$ 位。显然,要用这些芯片来构成实际上按字节组织的内存空间,就需要进行位的扩充,以满足字长的要求。

用 1 位或 4 位的存储器芯片构成 8 位字长的存储器,可采用位并联的方法。例如,可以用两片 $4K \times 4$ 位的存储器芯片经过位扩充构成 4KB 的存储器。

2. 地址的扩充

地址的扩充即存储容量的扩充(又称字扩充)。当扩充存储容量时,采用地址串联的方法。这时,要用到地址译码电路,以其输入的地址码来区分高位地址,而以其输出端的控制线来对具有相同低位地址的几片存储器芯片进行片选。

地址译码电路是一种可以将地址码翻译成相应控制信号的电路。有 2-4 译码器、3-8 译码器等。

当需要同时位扩充与字扩充时,可以将上述两种方法结合起来使用。例如,当用 $16\text{K} \times 1$ 位的芯片组成 $64\text{K} \times 8$ 位的存储器时,共需用 32 片 $16\text{K} \times 1$ 位的芯片。先用位线并联方法将每 8 片组成一组 $16\text{K} \times 8$ 位存储器,再用字扩充方法,选用 2-4 译码器组成的译码电路,组成 4 组 $16\text{K} \times 8$ 位共 $64\text{K} \times 8$ 位存储器。

5.4.2 存储器与 CPU 的连接

1. 只读存储器与 8086 CPU 的连接

ROM、PROM 或 EPROM 芯片都可以与 8086 系统总线连接,以组成程序存储器。例如,Intel 2716、2732、2764 和 27128 这一类 EPROM 芯片,由于它们属于以 1 字节宽度输出组织的,因此,在连接到 8086 系统时,为了存储 16 位指令字,要使用两片这类芯片并联组成一组。

注意: 在由两片 2732 组成的 4KB 程序存储器中,两片的片选端 $\overline{\text{CE}}$ 一起连到译码器的片选信号 $\overline{\text{CS}}$ 。

2. 静态 RAM 与 8086 CPU 的连接

一般当微机系统的存储器容量少于 16KB 时,宜采用静态 RAM 芯片,因为大多数动态 RAM 芯片都是以 $16\text{K} \times 1$ 位或 $64\text{K} \times 1$ 位来组织的,并且动态 RAM 芯片还要求动态刷新电路,这种附加的支持电路会增加存储器的成本。

8086 CPU 无论是在最小模式或最大模式下,都可以寻址 1MB 的存储单元,存储器均按字节编址。

注意: 在由两片 6116 组成的 2KB 数据存储器中,两片的片选端 $\overline{\text{CE}}$ 分别连到系统的 A_0 、 $\overline{\text{BHE}}$ 端。

3. EPROM、静态 RAM 与 8086 CPU 连接的实例

图 5.1 给出了 8086 单处理器系统典型结构连接的实例,它就是一个实际单板机的电路图。深入理解本系统的原理和电路,就能综合掌握和灵活运用有关存储器与 CPU 的连接知识。

图 5.1 中,8086 接成最小工作方式($\text{MN}/\overline{\text{MX}}$ 引脚置逻辑高电平)。当机器复位时,8086 将执行 FFFF0H 单元的指令。

本系统具有 32KB 的 EPROM 区,使用了 8 片 2732($4\text{K} \times 8$ 位)EPROM 芯片,分别以 $\text{U}_{32} \sim \text{U}_{39}$ 表示。

本系统还具有 16KB 的 RAM,使用了 8 片 6116($2\text{K} \times 8$ 位)静态 RAM 芯片,分别以 $\text{U}_{24} \sim \text{U}_{31}$ 表示。

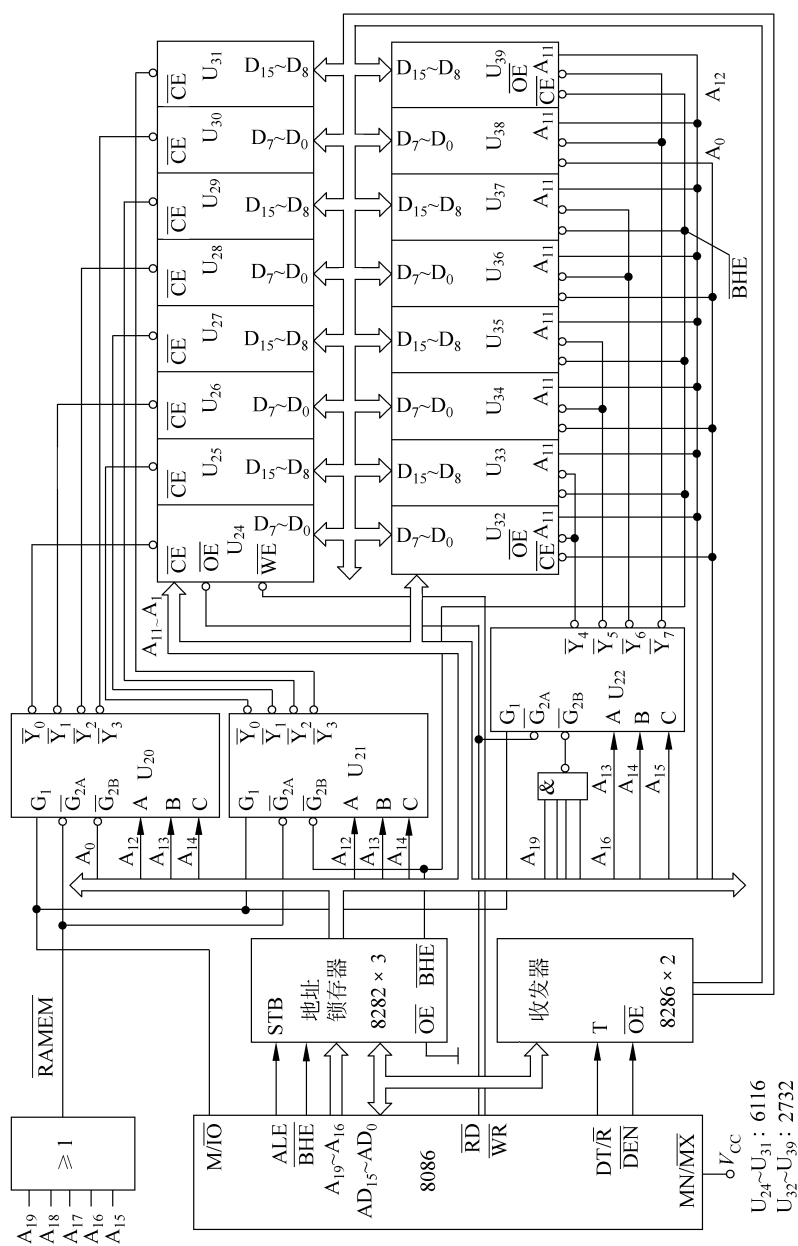


图 5.1 8086 单处理器系统典型结构连接实例

5.5 内存条技术的发展

内存历来都是系统中最大的性能瓶颈之一，特别是在 PC 技术发展的初期，PC 所使用的内存是一块块的集成电路芯片(IC)，并且将其焊接在主板上，这给后期维护与维修都带来许多麻烦。

随着 PC 技术的发展,PC 设计人员首次在 80286 主板上推出了模块化的条装内存,使每一条上集成了多块内存 IC,并在主板上也设计了相应的内存插槽,这样的内存条就大大地方便了安装与拆卸,内存的维修与升级也变得非常简单。此后,内存条从规格、技术、总线带宽等不断更新换代,内存的性能瓶颈问题得到较大改善。

1. SIMM 内存

最初(1982 年)出现在 80286 主板上的内存条,采用的是单边接触内存模组(single inline memory modules,SIMM)接口,容量为 30 线、256KB,由 8 片数据位和 1 片校验位组成 1 个存储区块(bank),因此,一般见到的 30 线 SIMM 都是 4 条一起使用。

1988 年到 1990 年,PC 技术进入 32 位的 386 和 486 时代,推出了 72 线 SIMM 内存,它支持 32 位快速页模式内存,内存带宽得以大幅提升。72 线 SIMM 内存单条容量一般为 512KB 至 2MB,要求两条同时使用。

注意,72 线的 SIMM 内存引进了一个快速页面动态内存(FP DRAM),在 386 时代很流行。

2. EDO DRAM 内存

外扩充数据模式动态存储器(extended data out DRAM,EDO DRAM)是 1991 年到 1995 年之间盛行的内存条,主要应用在当时的 486 以及早期的 Pentium 计算机中。

随着 EDO DRAM 在成本和容量上的突破,加上制作工艺的飞速发展,当时单条 EDO DRAM 内存的容量为 4MB~16MB。

3. SDRAM 时代

自 Intel Celeron 系列、AMD K6 处理器以及相关的主板芯片组推出后,EDO DRAM 内存性能再也无法满足需要,于是内存又开始进入 SDRAM 时代。

第一代 SDRAM 内存为 PC66 规范,之后有 PC100、PC133、PC150 等规范,其频率从早期的 66MHz,发展到 100MHz、133MHz 等。由于 SDRAM 的带宽为 64 位,正好对应 CPU 的 64 位数据总线宽度,因此,它只需要一条内存便可工作,便捷性进一步提高。

4. Rambus DRAM 内存

硬件技术竞争的特点是频率竞争,由于 CPU 主频的不断提升,Intel 公司在推出高频 Pentium III 以及 Pentium 4 CPU 的同时,与 Rambus 联合推出了 Rambus DRAM 内存(简称 RDRAM 内存)。与 SDRAM 不同的是,它采用了新一代高速简单内存架构,基于一种类 RISC(精简指令集计算机)理论,可以减少数据的复杂性,使得整个系统性能得到提高。

5. DDR 时代

双倍速率 SDRAM(Double Data Rate SDRAM,DDR SDRAM)简称 DDR,它实际上是 SDRAM 的升级版本,采用在时钟信号的上升沿和下降沿都可以传输数据,因而时钟率可以加倍提高,传输速率和带宽也相应提高。

DDR SDRAM 内存有 184 个引脚,引脚部分有一个缺口,其作用是在安装内存条时