



南京理工大学
NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

第 3 章

存储层次与系统

主讲：张功萱

©第1版 2023.08 张功萱

2024-10-10

1

非易失性存储器

第3.3节

1. 只读存储器的作用是什么？
2. 闪烁存储器的原理是什么？
3. 新型的又有哪些结构构成？

3.3 非易失性存储器

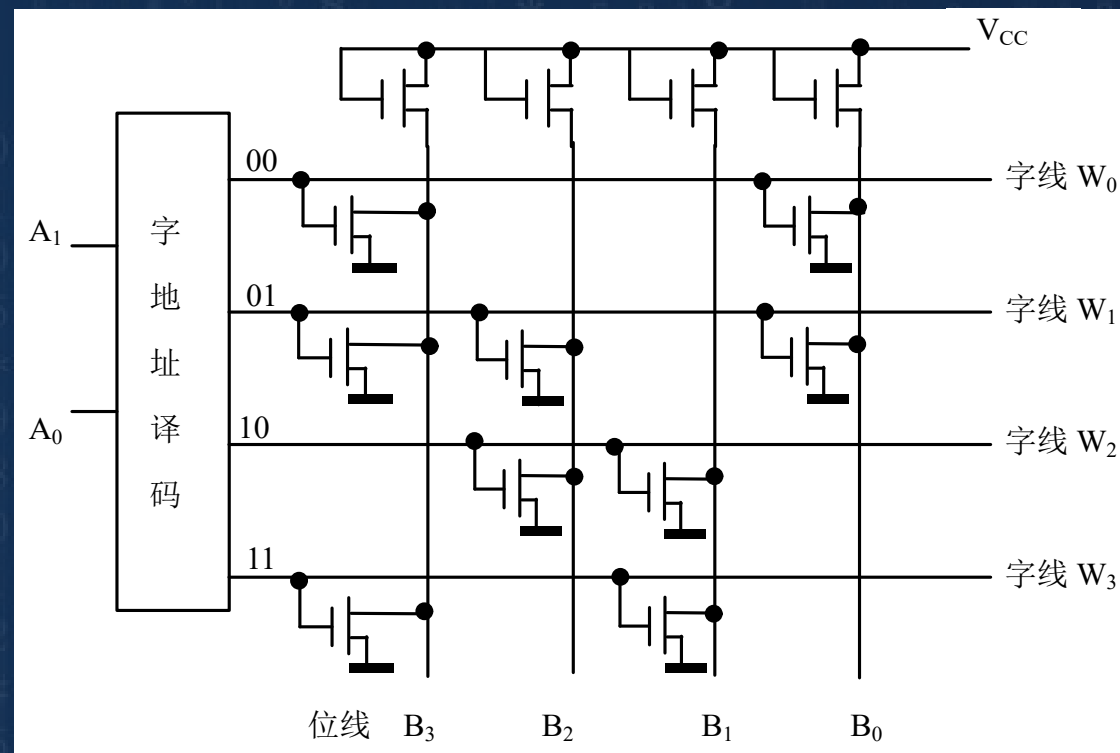
- 非易失性存储器（Non-Volatile Memory, NVM）是指当关机后，存储器中的内容不会随之消失的计算机存储器。
- 在非易失性存储器中，根据存储器中的内容是否能在计算机工作时随时改写为标准，可分为三大类：只读存储器（Read-Only Memory, ROM）、闪存存储器（Flash memory）和新型的非易失性存储器。

3.3.1 只读存储器

- 只读存储器属于非易失性存储器，在系统断电以后，其中所存储的内容不会丢失。
- 半导体只读存储器常作为主存的一部分，用于存放一些固定的程序，如监控程序、启动程序、磁盘引导程序等。只要一接通电源，这些程序就能自动运行。
- 只读存储器还可以用作控制存储器、函数发生器、代码转换器等。在输入、输出设备中，常用ROM存放字符、汉字等的点阵图形信息。

1. 掩膜ROM

- 掩膜ROM中的内容是由半导体存储芯片制造厂家在制做芯片时，直接写入芯片中的，即掩膜ROM中的内容用户不可编程。右图为4位掩膜ROM。
- 主要优点：价格便宜。
- 主要缺点：如果一批掩膜ROM中的某个代码或数据输入有误，则整批掩膜ROM都将作废。



2. 可编程ROM (PROM)

- 一次可编程只读存储器 (Programmable ROM, PROM) 的内部由行列式的熔丝构成。在出厂时，存储的内容为全“1”或全“0”。
- 用户可以通过ROM编程器，利用电流将PROM中某些单元的熔丝烧断，进行数据的写入，以实现对其“编程”的目的。
- PROM与掩膜ROM的区别在于，用户可以在PROM芯片制造完成后，根据需要写入数据。

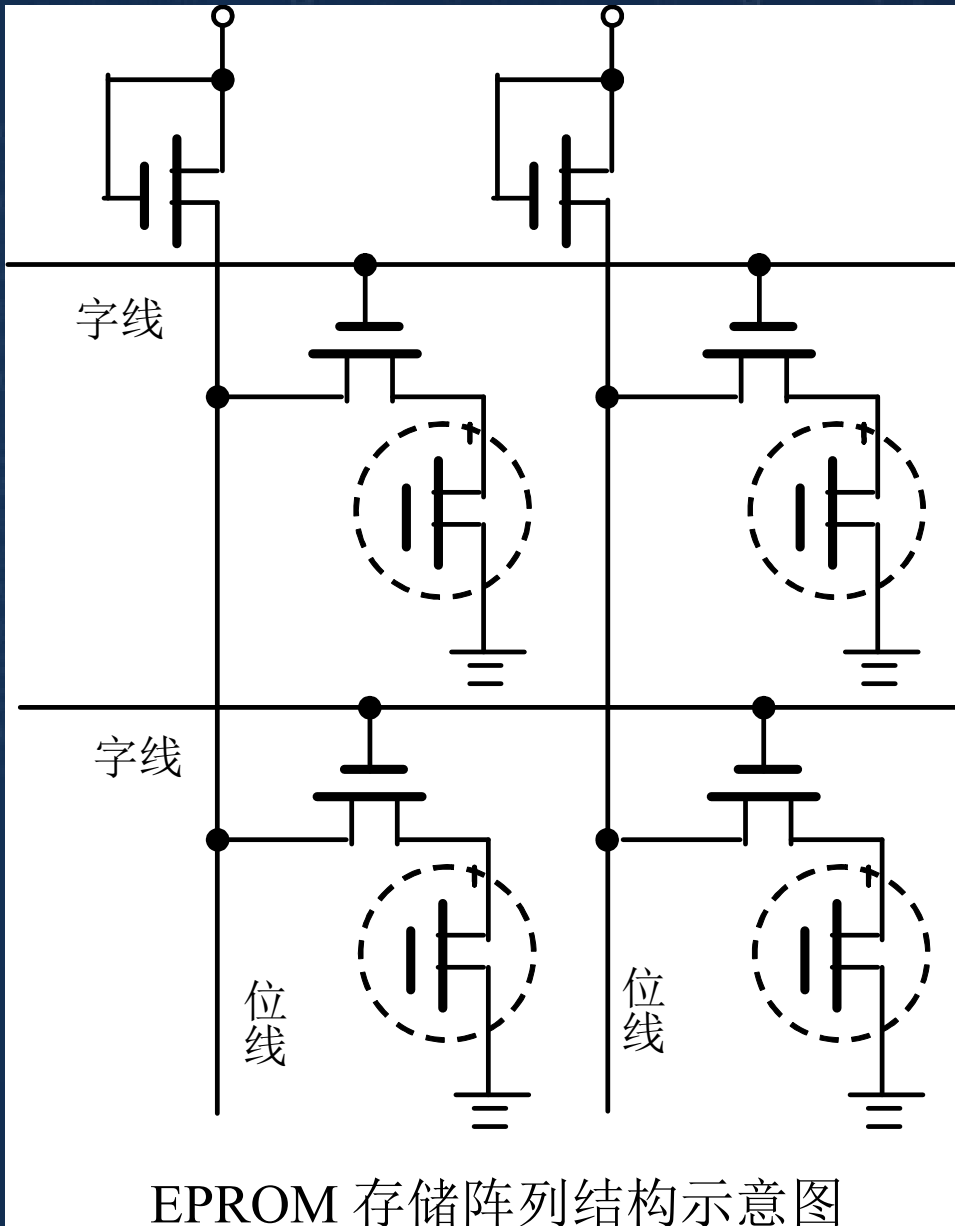
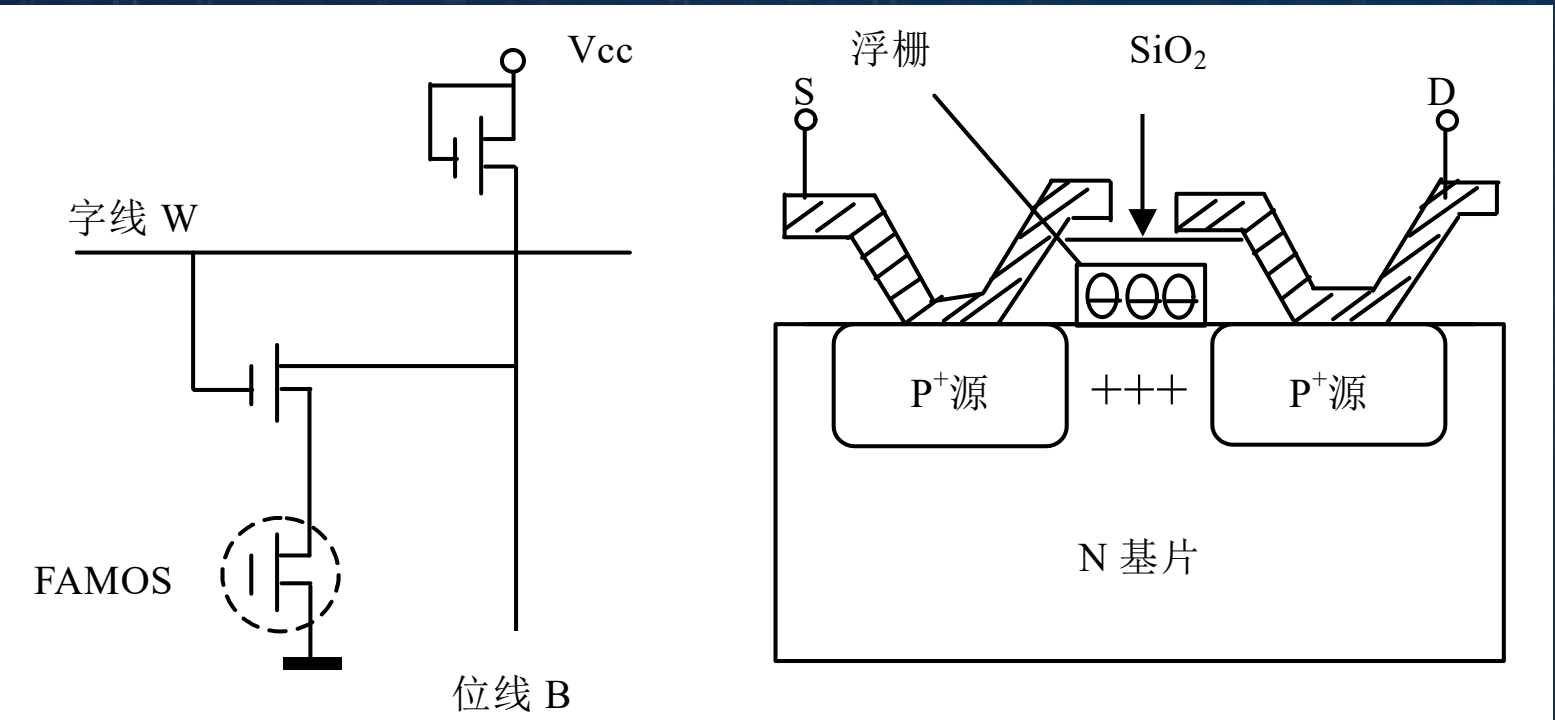


熔丝断:为“0”

熔丝未断:为“1”

3. 紫外线擦除PROM (EPROM)

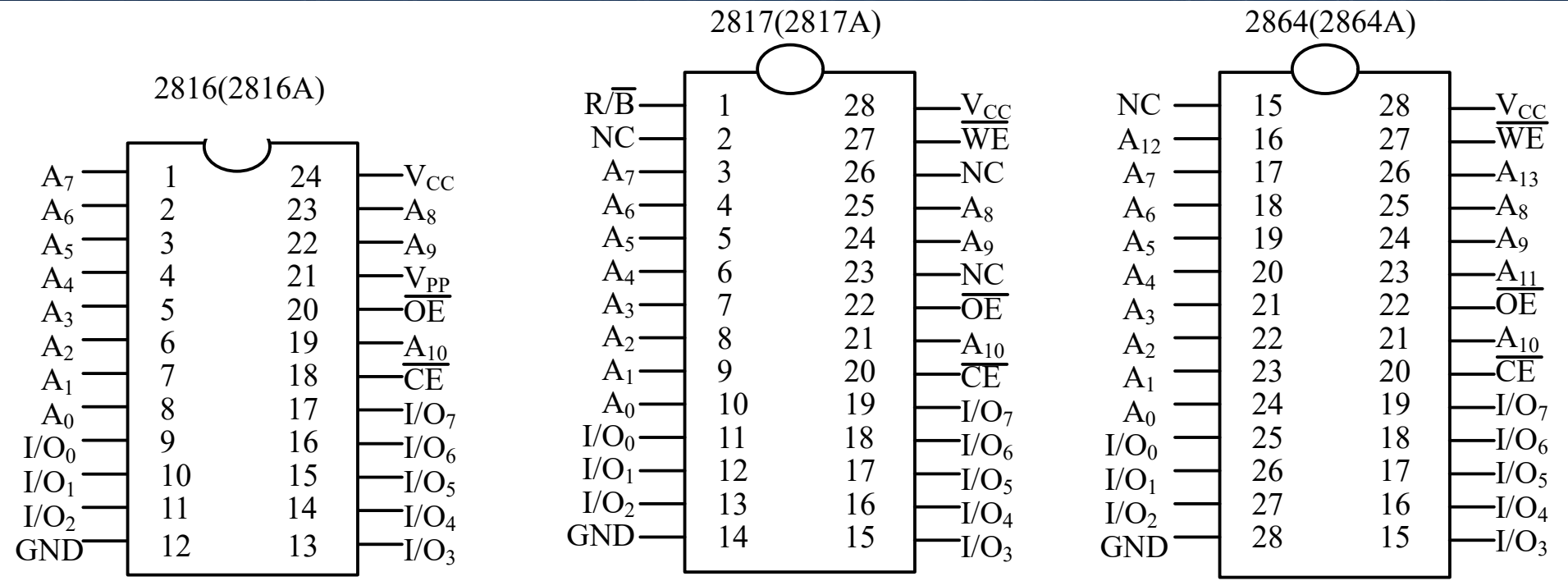
- EPROM 是用紫外线实现擦除的PROM (Erasable Programmable ROM, EPROM)。
- EPROM芯片上有一窗口用于接收紫外线，通过紫外线照射擦除其内容。
- 已写入EPROM中的信息可以被修改，并且可被擦除、编程上千次。
- EPROM的问题是：
 - 需要紫外设备
 - 擦除芯片的内容耗时较长，为分钟级。



EPROM 存储阵列结构示意图

4. 电擦除PROM (EEPROM)

- 电擦除的PROM (Electrically Erasable Programmable ROM, EEPROM、E²PROM)



E²PROM的特点：

- (1) 可用电擦除芯片中原有信息，实现瞬间擦除。
- (2) 用户可以有选择地擦除具体字节单元的内容，而不像EPROM那样只能整片擦除。
- (3) 系统设计者可在电路板上设置对E²PROM进行擦除和编程的电路，用户可直接在电路板上对E²PROM芯片进行擦除和编程，不需要额外的擦除和编程设备。

5. 闪速存储器

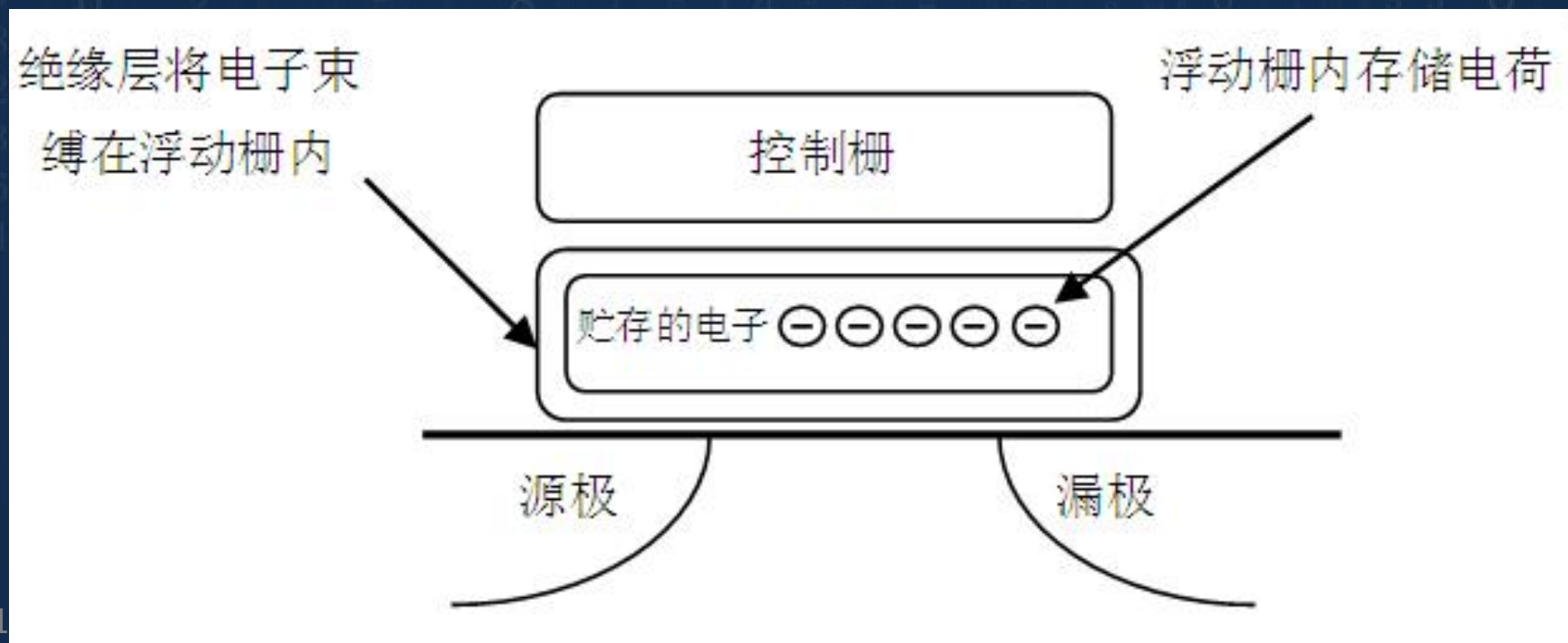
- 闪速存储器（Flash Memory，闪存）是取代传统的 EPROM 和 EEPROM 的高密度非易失性存储器。
- 闪存是用电擦除的，也被称为闪烁电擦除可编程ROM。
- 与EEPROM相比较，闪速存储器具有存储密度高、成本低、读写速度快等特点。

3.3.2 闪速存储器

- 闪速存储器(Flash Memory)是一类非易失性存储器NVM(Non-Volatile Memory)，即使在供电电源关闭后仍能保持片内信息。
- 相对传统的EEPROM芯片，闪存芯片可以快速地擦写。
- 由于闪存中不需要存储电容，故其集成度更高，且制造成本低于DRAM。
- 闪存使用方便，既具有SRAM读写的灵活性和较快的访问速度，又具有ROM在断电后可不丢失信息的特点，所以闪速存储器技术发展非常迅速。

1. 闪存存储器的基本原理

- 闪存以单晶体管作为二进制信号的存储单元。每个存储单元在晶体管的控制栅与漏极 / 源极之间设置浮动栅，用于存储信息。



- 当负电子在控制栅的作用下被注入到浮动栅中时，单晶体管的存储状态就由“1”变成“0”。
- 当负电子从浮动栅中移走后，存储状态就由“0”变成“1”。
- 包覆在浮动栅表面的绝缘体可将内部的电子“困住”，达到保存数据的目的。
- 写入数据时，须将浮动栅中的负电子全部移走，令目标存储区域都处于“1”状态，因此只有遇到数据“0”时才会发生写入动作。

闪存的电荷生成与存储方法

- (1) 通道热电子编程 (CHE)
- CHE (Channel Hot Electron)方法通过对控制栅施加高电压，使传导电子在电场的作用下突破绝缘体的屏障进入到浮动栅内部，反之亦然，以此来完成写入或擦除操作。
- (2) 隧道效应法 (FN)
- FN(Fowler-Nordheim)方法通过直接在绝缘层两侧施加高电压形成高强度电场，帮助电子穿越氧化层通道进出浮动栅，以实现写入或擦除操作。

2. 闪存存储器的特点

- (1) 固有的非易失性
 - 可在不加电的情况下长期保持存储的信息。
- (2) 廉价的高密度
 - 和SRAM及DRAM相比，相同存储容量的闪存存储器具有更低的成本。
- (3) 可直接执行
 - 闪存中存储的应用程序可以直接在闪存内运行，不必把代码读到系统RAM中。
- (4) 固态性能
 - 可用作固态大容量存储器，如固态硬盘。

3. 闪速存储器的分类

- 1) NOR型闪存
- NOR闪存工作时同时使用CHE和FN两种方法。CHE方法用于数据写入，支持单字节或单字编程；FN方法用于擦除。
- NOR型闪存不能单字节擦除，必须以块为单位或对整片区域执行擦除操作，其擦除和编程操作所花费的时间很长。
- NOR型闪存带有SRAM接口，可以方便地存取其内部的每一个字节，因此应用程序可以直接在闪存内运行，不必再把代码读到系统RAM中。
- NOR型闪存容量较小，价格较贵，较适合频繁随机读写的场合，如作为手机内存。

2) NAND型闪存

- NAND闪存工作时采用FN法写入和擦除，擦除动作很快。
- NAND闪存的存储密度较高，与NOR闪存相比，生产成本更低，也更容易生产大容量的芯片。
- NAND闪存理论擦写次数约为100万次，寿命上比NOR闪存要占优势。
- NAND闪存的缺陷在于容量越大，寻址时间就越长，并且不支持代码本地执行。
- NAND型闪存主要用于存储资料，常被应用于如数码照相机、数码摄像机及闪存卡等数码产品。

NOR型与NAND型闪存的区别

- ①**NOR**型闪存更像单片机闪存内存，有独立的地址线 and 数据线，但价格比较贵，容量比较小；而**NAND**型更像硬盘，地址线 and 数据线是共用的**I/O**线，类似硬盘的所有信息都通过一条硬盘线传送一般，而且**NAND**型与**NOR**型闪存相比，成本要低一些，而容量大得多。
- ②**NOR**型闪存比较适合频繁随机读写的场合，通常用于存储程序代码并直接在闪存内运行，手机就是使用**NOR**型闪存的大户，所以手机的“内存”容量通常不大；**NAND**型闪存主要用来存储资料，我们常用的闪存产品，如闪存盘、数码存储卡都是用**NAND**型闪存。

3.3.3 新型的非易失性存储器

- 新型的非易失性存储器通常包括铁电存储器（Ferroelectric Random Access Memory, FRAM）、相变存储器（Phase Change Random Access Memory, PRAM）、磁性存储器（Magnetic Random Access Memory, MRAM）和阻变存储器（Resistive Random Access Memory, RRAM）等。
- 相比其它非易失性存储技术，RRAM是高速存储器。
- 下面重点介绍RRAM芯片的重要电子元件——忆阻器（Memristor）。

*忆阻器的基本原理

- 忆阻器：是表示磁通与电荷关系的电路器件。其中，忆阻的阻值是由流经它的电荷来确定的。因此，通过测定忆阻的阻值，便可知道流经它的电荷量，从而具有记忆电荷的作用。简单来说，忆阻器是一种有记忆功能的非线性电阻，通过控制电流的变化可改变其阻值。
- 例如，如果把高阻值定义为“1”，低阻值定义为“0”，则这种电阻就可以实现存储数据的功能。

*忆阻器的基本原理

- 忆阻器是由两个金属电极夹着的氧化钽层构成的双端与双层交叉开关结构的半导体。其中一层氧化钽掺杂了氧空位，成为一个半导体，而相邻一层不掺任何东西，保持绝缘体的自然属性，通过检测交叉开关两端电极的阻性，就能判断RRAM的“开”或者“关”状态。

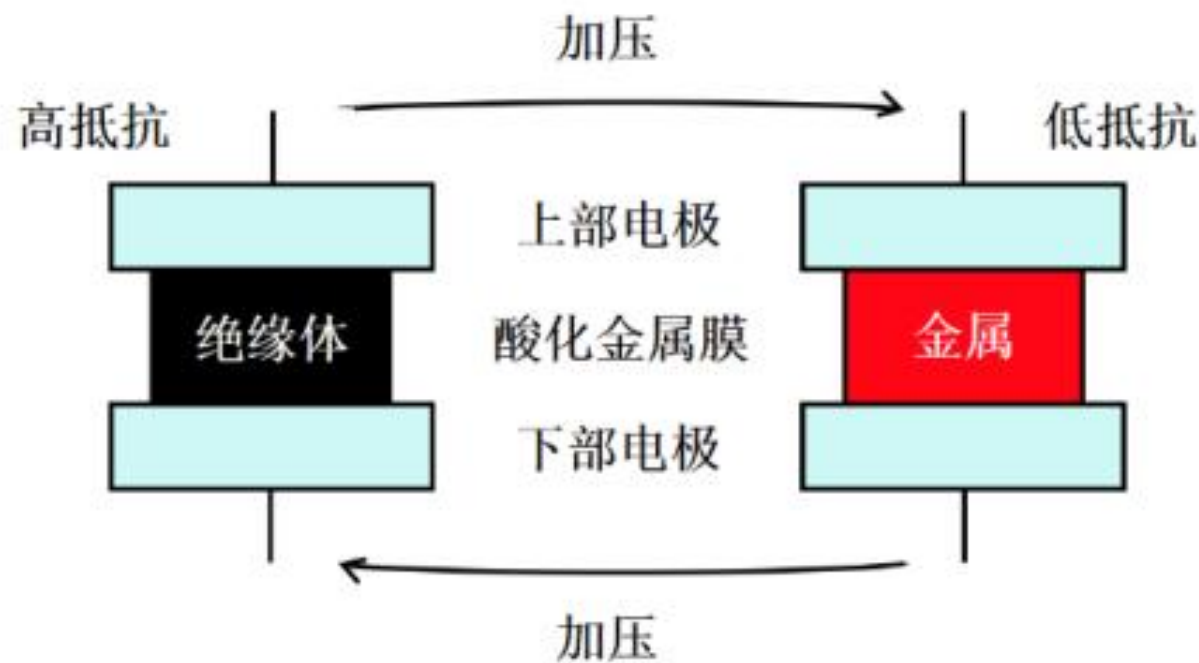


图 3-18 忆阻器的基本原理

并行存储器

第3.4节

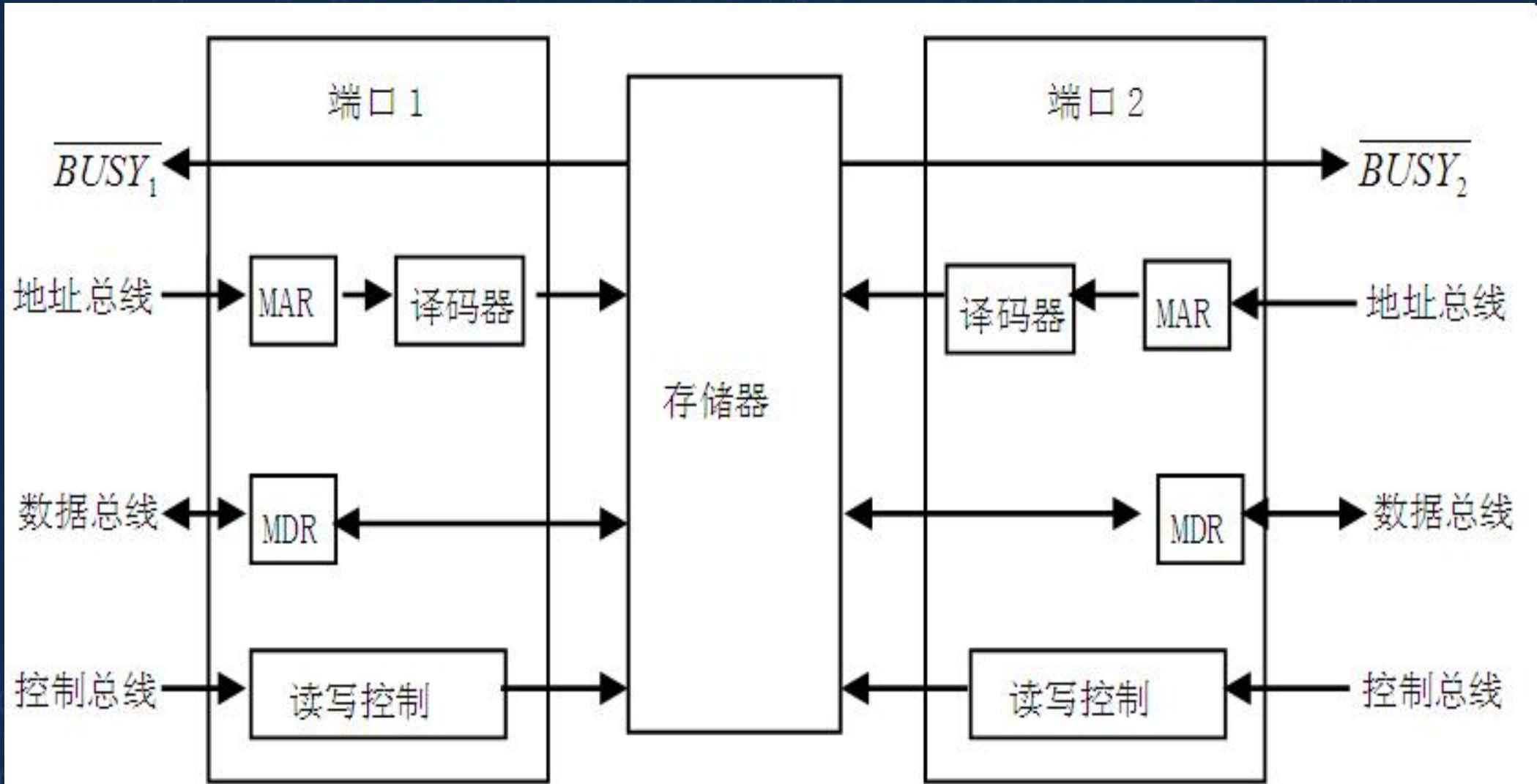
1. 提高存储器速度途径有哪些？
2. 双端口存储器的作用是什么？
3. 并行存储器提速有哪些方法？
4. 什么是联想(或相联)存储器？

- 虽然存储器系统的速度随着计算机技术的不断发展也在不断提高，但始终跟不上CPU速度的提高，成为了限制系统速度的瓶颈。
- 为了解决存储器与CPU的速度匹配问题，通常采用的方法：
 - (1) 采用更高速的主存储器，或加长存储器的字长。
 - (2) 采用并行操作的双端口存储器。
 - (3) 采用并行存储器。
 - (4) 在CPU和主存储器之间插入高速缓冲存储器（Cache）。

3.4.1 双端口存储器

- **单端口存储器：**
- 单端口存储器每次只接收一个地址，访问一个编址单元，从中读取或存入一个字节或一个字。
- **双端口存储器：**
双端口存储器具有两个彼此独立的读/写口，每个读/写口都有一套独立的地址寄存器和译码电路，可以并行地独立工作。两个读/写口可以按各自接收的地址，同时读出或写入，或一个写入而另一个读出。与两个独立的存储器不同，两套读/写口的访存空间相同，可以访问同一区间、同一单元。

双端口存储器



双端口存储器的读写规定

- 双端口存储器的每个端口都可独立对存储器进行读写，就像是两个存储器在同时工作，可实现并行存储操作。
- 当送达两个端口的访存地址不同时，在两个端口上进行读写操作不会发生冲突。
- 当送达两个端口的访存地址是存储器同一单元的地址时，会发生读写冲突。
- 为解决读写冲突问题，双端口存储器芯片设置了 BUSY BUSY标志。只有BUSY为“1”的端口才可以进行读写操作。

双端口存储器的常见应用场合

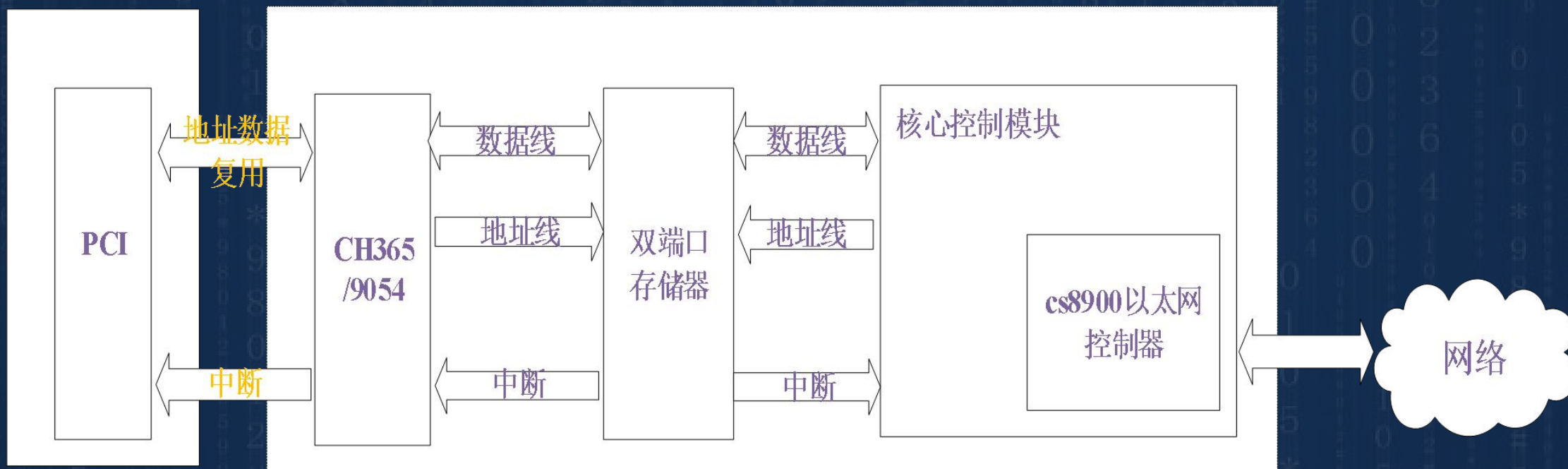
- (1) 在运算器中采用双端口存储芯片作为通用寄存器组，能快速提供双操作数，或快速实现寄存器间传送。（如AM2901芯片）
- (2) 让双端口存储器的一个读/写口面向CPU，通过专门的存储总线(或称局部总线)连接CPU与主存，使CPU能快速访问主存；另一个读 / 写口面向外围设备或输入输出处理机IOP，通过共享的系统总线连接，这种连接方式具有较大的信息吞吐量。（通过RAM“缓冲桥接”）
- (3) 在多机系统中采用双端口存储器甚至多端口存储器，作为各CPU的共享存储器。实现多CPU之间的通信。

项目实例

- PC处理器+PCI控制系统+双端口缓冲+ARM处理器+网络接口

宿主PC

网络安全处理板卡



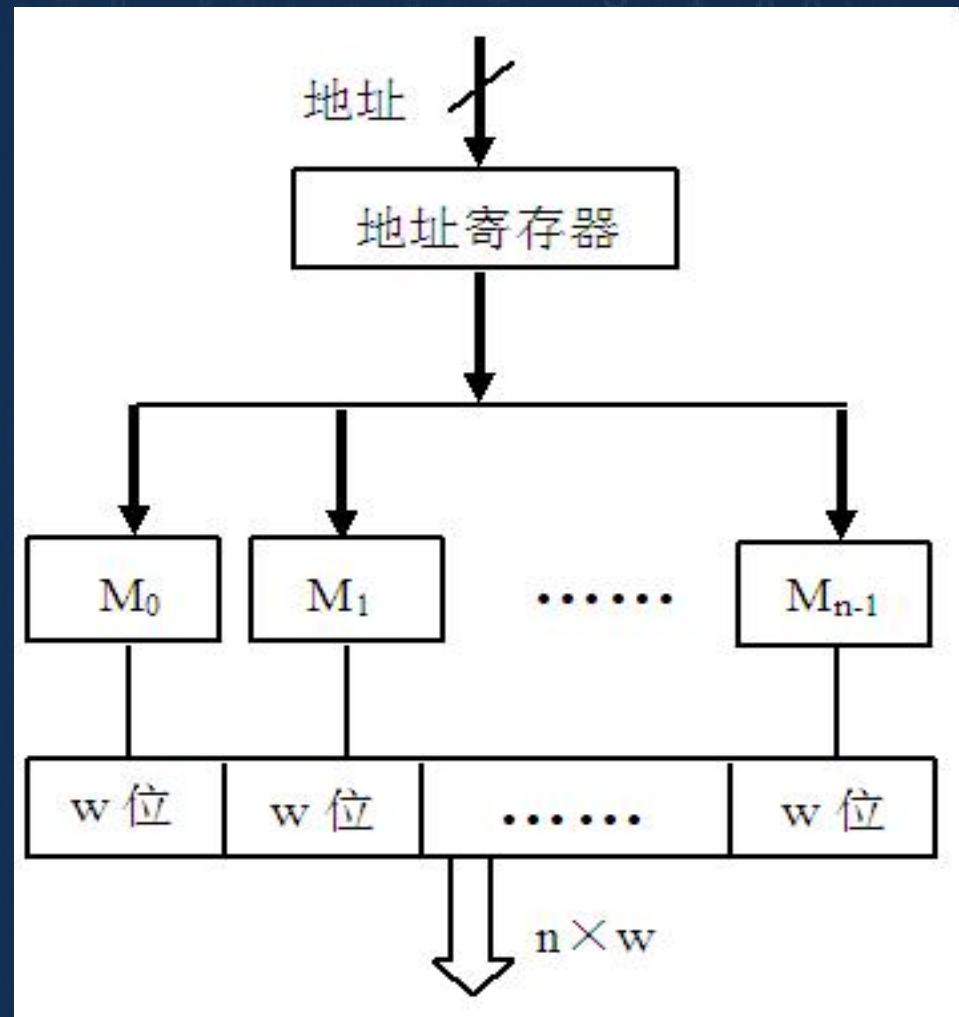
- 其中，双端口缓冲存储器是保证双核系统正常高速通信的桥梁，因此在辅核**Linux**内核中，需要对双端口存储器数据进行读/写，必须开发双端口存储器的管理模块程序，以完成对双端口存储器中存贮资源资源的访问。

3.4.2 并行存储器

- 为提高系统速度，在高速的大型计算机中普遍采用并行主存系统。
- 并行主存系统：
- 在一个存储周期内可并行存取多字的存储系统。
- 利用并行主存系统可以提高整个存储器系统的吞吐率（数据传送率），解决CPU与主存间的速度匹配问题。

1. 单体多字并行主存系统

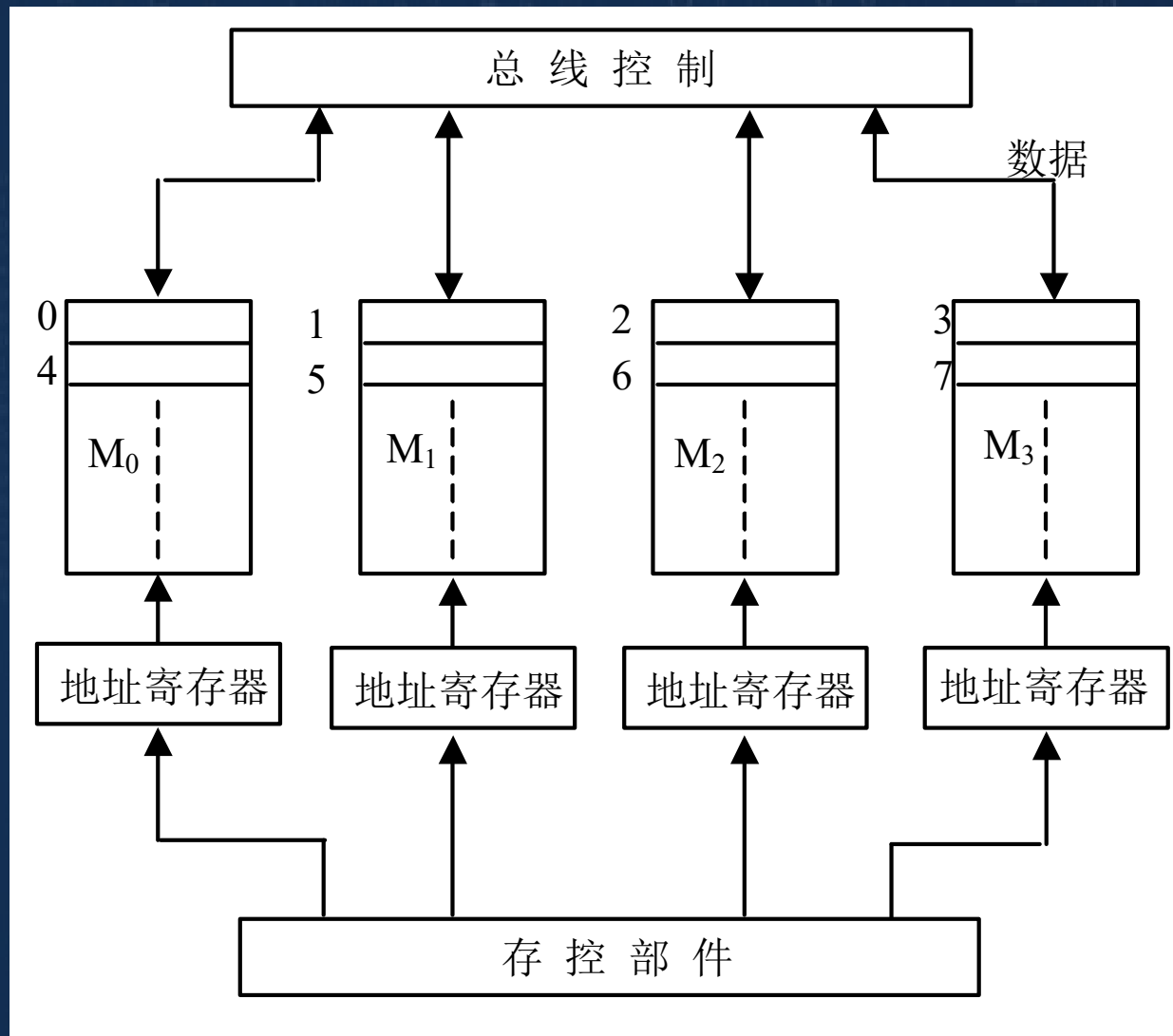
- 单体多字并行主存系统中，多个并行存储器共用一套地址寄存器，按同一地址码并行地访问各自的对应单元。



- 例如 n 个存储器顺序排列 n 个字，每个字有 w 位。假定送入的地址码为 A ，则 n 个存储器同时访问各自的 A 号单元。即将这 n 个存储器视作一个大存储器，每个编址对应于 n 字 $\times w$ 位，因而称为单体多字方式。
- 单体多字并行主存系统适用于向量运算一类的特定环境。
- 例如进行矩阵运算时，就适于采用单体多字并行存取的方式对操作数和运算结果进行读写。

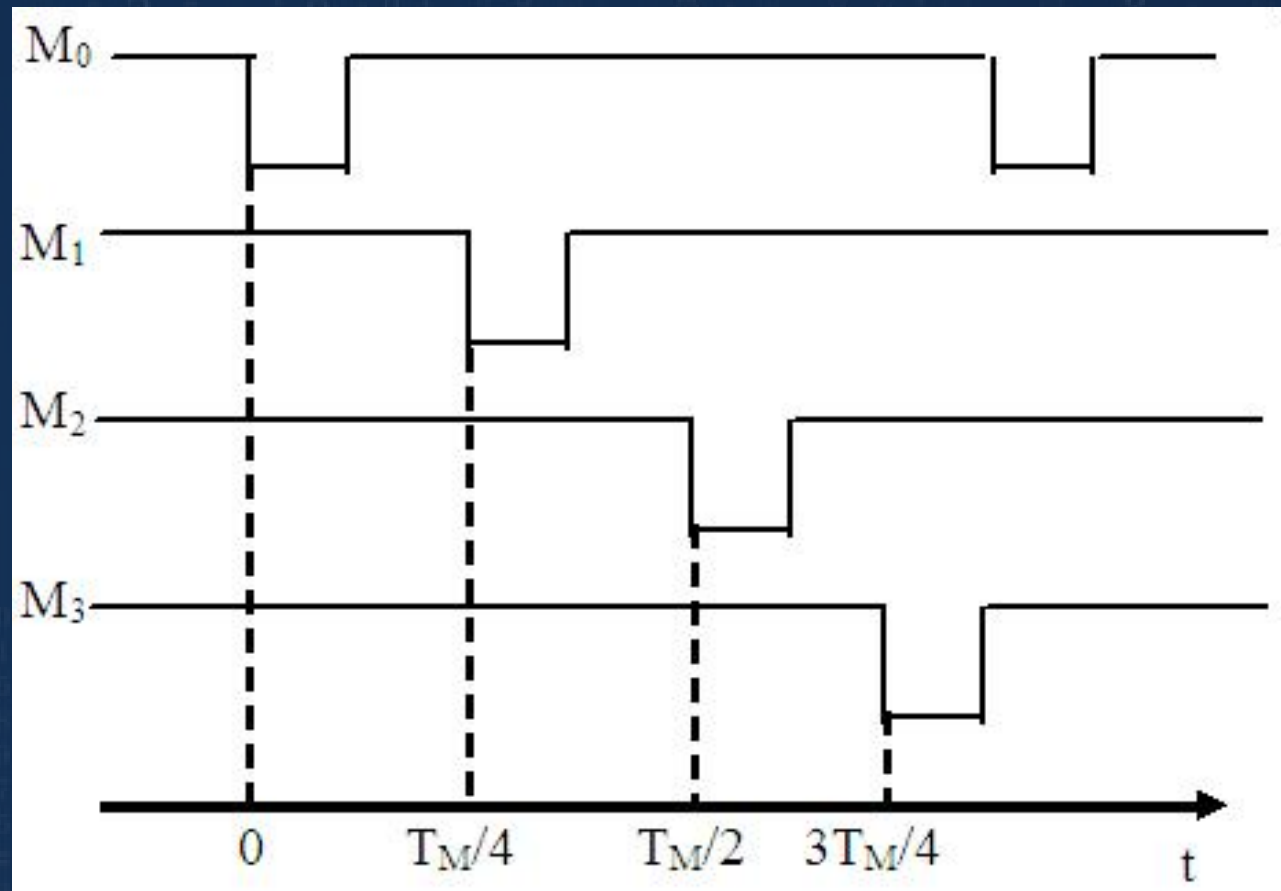
2. 多体交叉存取方式的并行主存系统

- 多体交叉存储器使用 n 个容量相同的存储器，或称为 n 个存储体，它们具有自己的地址寄存器、数据线、时序，可以独立编址的同时工作。



- 各存储体的编址大多采用交叉编址方式，即将一套统一的编址，按序号交叉地分配给各个存储体。
- 例如由四个存储体组成的多体交叉存储器中，各存储体的地址编址序列是：
 - M_0 : 0, 4, 8, 12, ...,
 - M_1 : 1, 5, 9, 13, ...,
 - M_2 : 2, 6, 10, 14, ...,
 - M_3 : 3, 7, 11, 15,
- 一段连续的程序或数据，可交叉地存放在 n 个存储体中，整个并行主存以 n 为模交叉存取。

- 在多体交叉存储器中，各存储体采取分时访问的时序。
- 例：设多体交叉存储器中有4个存储体，模等于4，各体分时启动读/写，每经过四分之一存取周期启动一个存储体。每个存取周期可访存四次。



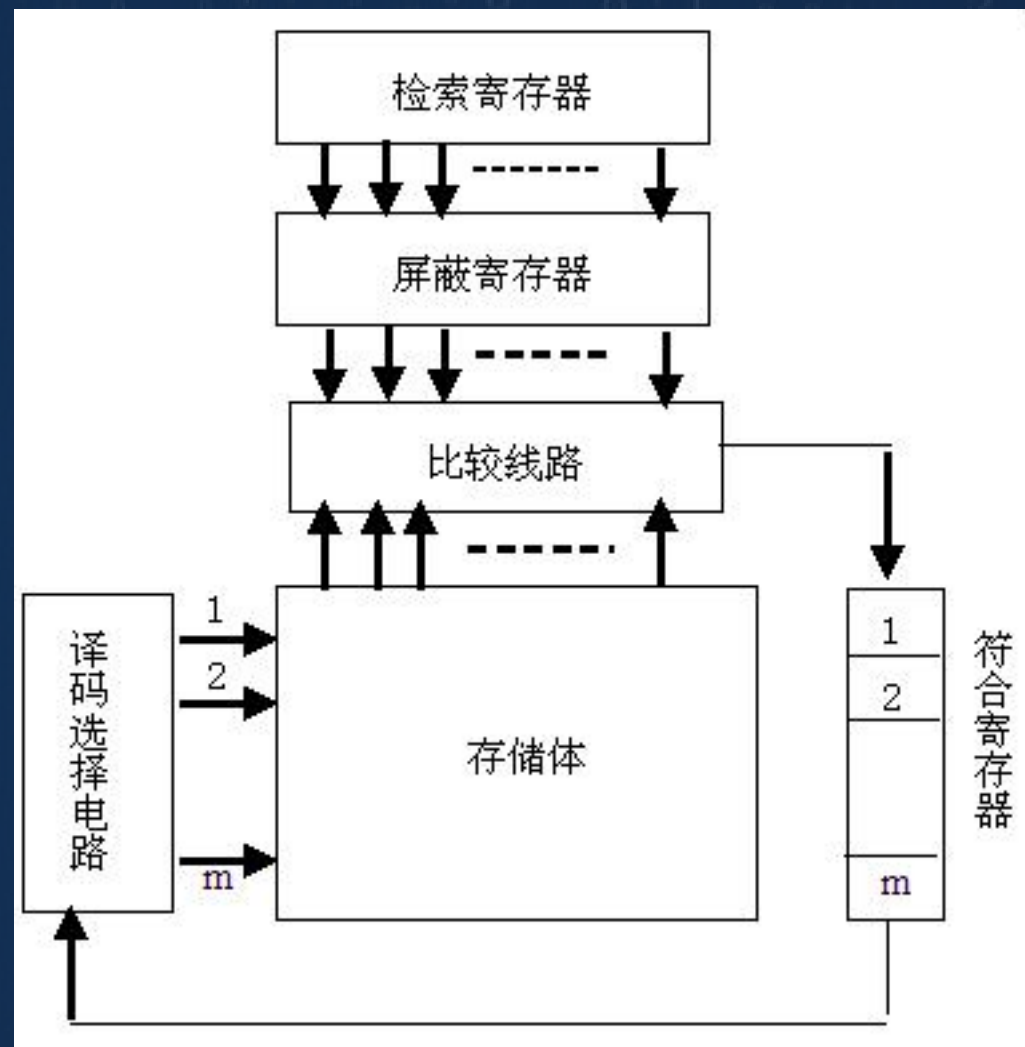
- 对每一个存储体来说，从CPU给出访存命令直到读出信息仍然使用了一个存取周期时间。
- 对CPU来说，它可以在一个存取周期内连续访问 n 个模块，各模块的读写过程并行进行。
- 在多体交叉存储器中，每个存储体都有自己的读写控制电路、地址寄存器和数据寄存器，各自以等同的方式与CPU传送信息。
- 在理想情况下，如果程序段或数据块都是连续地在主存中存取，那么将大大提高主存的访问速度。

- 采取多体交叉存取方式，需要一套存储器控制逻辑，简称为存控部件。
- 当CPU或其他设备发出访存请求时，存控部件按优先排队决定是否响应请求。响应后按交叉编址关系决定该地址应访问哪个存储体，并查询该存储体的“忙”触发器是否为1。若为1，表示该存储体正在进行读/写操作，需等待；若该存储体已完成一次读/写，则将“忙”触发器置0，然后可响应新的访存请求。当存储体完成读/写操作时，将发出一个回答信号。

- 多体交叉存取方式**适合支持流水线的处理方式**，而流水处理方式已是CPU中一种典型技术，因此，多体交叉存储结构是高速大型计算机的典型主存结构。

3.4.3 相联存储器

- **相联存储器** (Associative Memory)
- 又称为联想存储器。相联存储器是根据所存信息的全部特征或部分特征进行存取的，是一种按内容寻址的存储器。
- 如图是相联存储器的逻辑结构图。



- 相联存储器由存储体、检索寄存器、屏蔽寄存器、符合寄存器、比较线路、数据寄存器以及控制线路组成。
- 检索寄存器：用于存放数据本身或数据的特征标志部分（检索项）。
- 屏蔽寄存器：存放屏蔽字代码，用于屏蔽掉不需要进行符合比较的无效位。
- 符合寄存器：用于保存比较结果。
- 检索寄存器和屏蔽寄存器的位数与存储体中存储单元的位数（ n 位）相等，符合寄存器的位数则跟存储单元数（ m ）相等，即符合寄存器的每一位对应于存储体中的一个存储单元。

- 比较线路：
- 用于把需要查找的检索项同时和相联存储阵列中的每一个存储单元的相应部分进行逻辑比较。
- 若存储单元的内容与检索项完全相同，就把与该存储单元对应的符合寄存器的相应位置“1”，表示该字就是所要查找的字。然后利用符合寄存器中的符合信号去控制各个查找到的存储单元的读/写操作。

- 由于相联存储器要求每个基本存储单元都具有比较功能，因此设计复杂，成本较高，容量不大。
- 实用的联想存储器，一般除有按内容访问能力外，还有按地址访问能力。故仍保留有地址寄存器、译码电路和读出寄存器。
- 联想存储器的每个基本单元除了有存储能力和相等比较功能外，还可实现 \neq ， $<$ ， $>$ ， \leq ， \geq ，MAX，MIN，BETWEEN，NEXT HIGHER，NEXT LOWER等比较功能。