

4.1 概述

一、存储器分类

1. 按存储介质分类

(1) 半导体存储器	TTL、MOS	易失
(2) 磁表面存储器	磁头、载磁体	非易失
(3) 磁芯存储器	硬磁材料、环状元件	
(4) 光盘存储器	激光、磁光材料	



2. 按存取方式分类

(1) 存取时间与物理地址无关（随机访问）

- 随机存储器 在程序的执行过程中 可 读 可 写
- 只读存储器 在程序的执行过程中 只 读

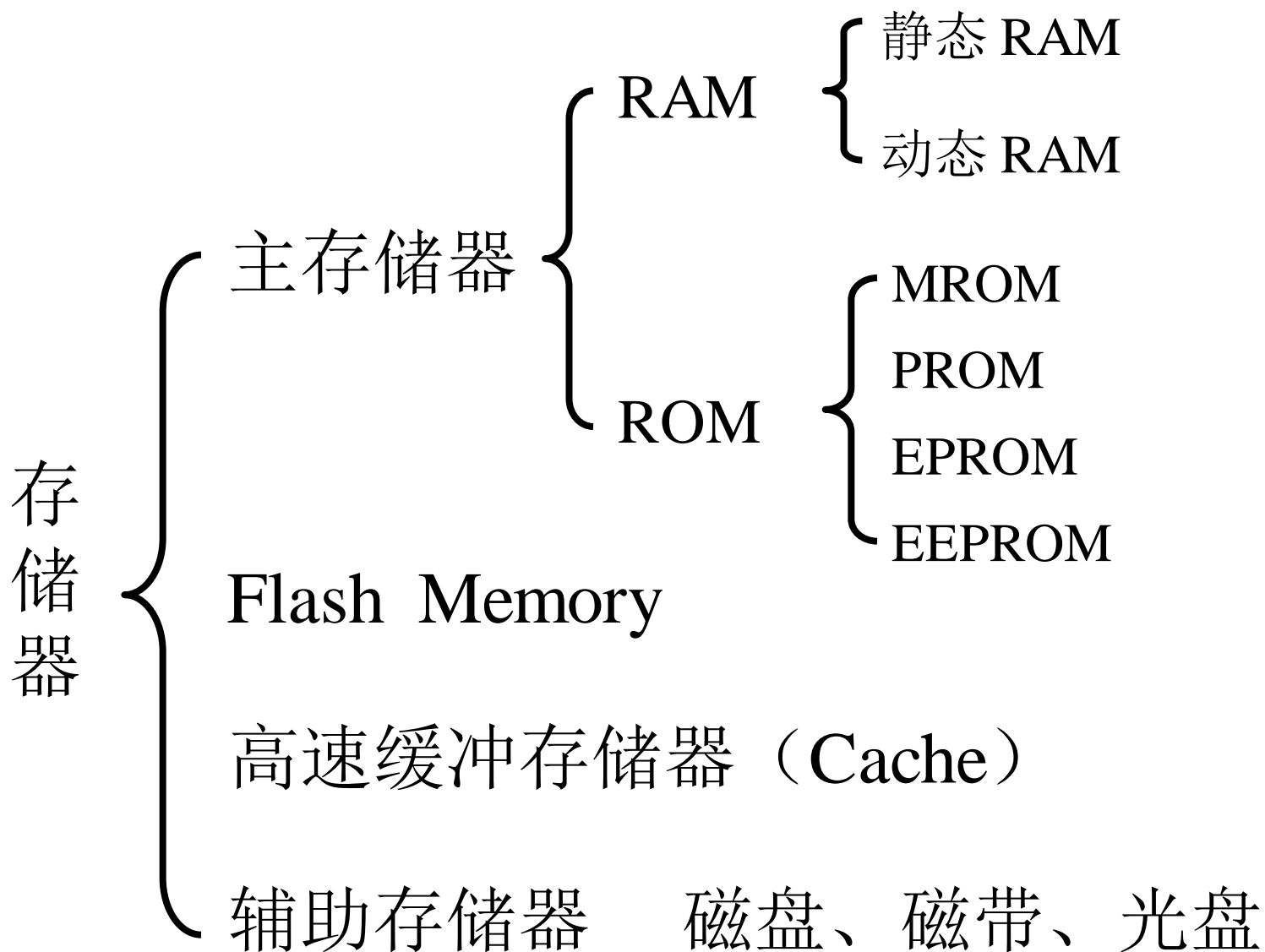
(2) 存取时间与物理地址有关（串行访问）

- 顺序存取存储器 磁带
- 直接存取存储器 磁盘



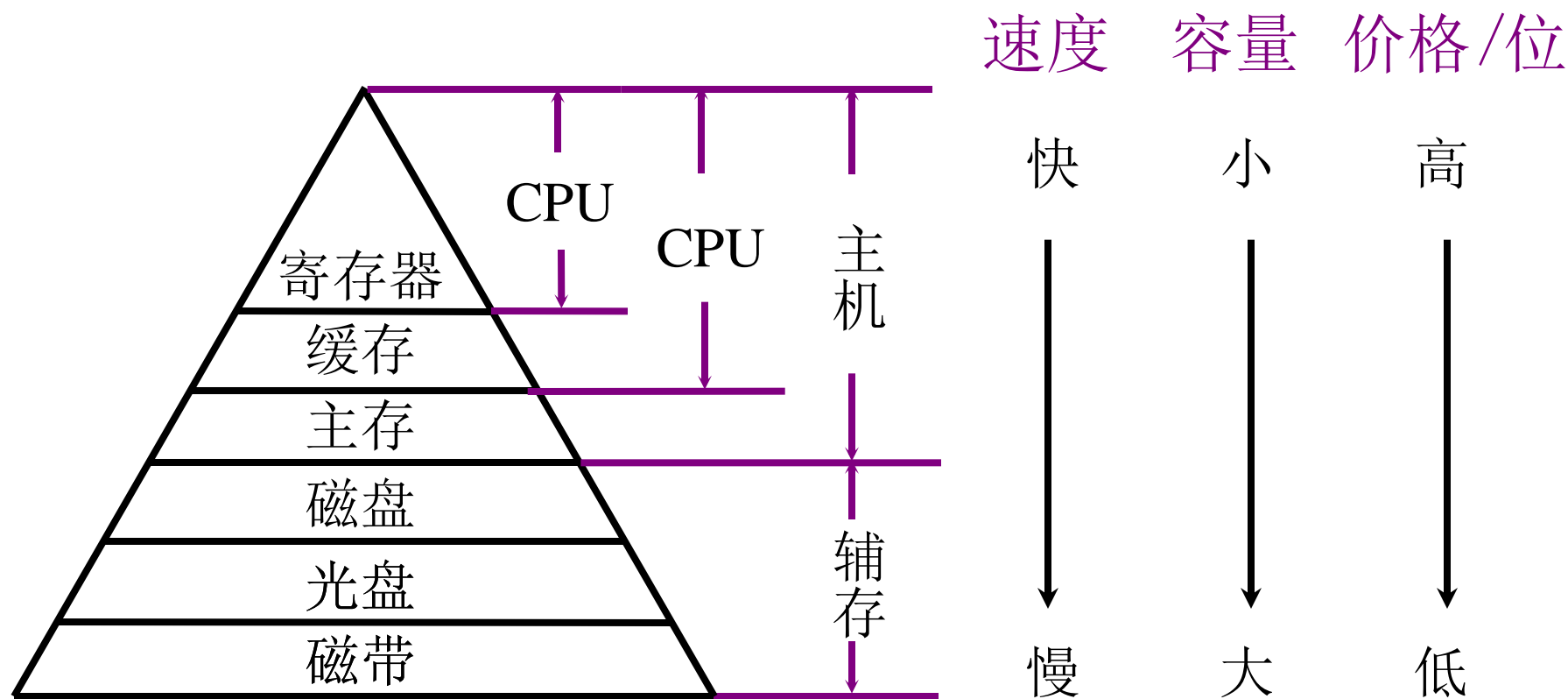
3. 按在计算机中的作用分类

4.1

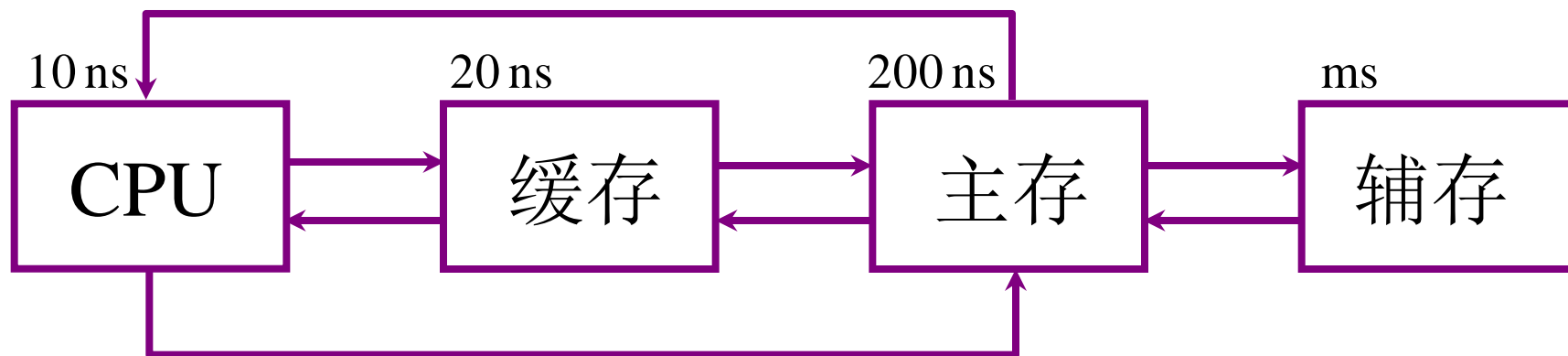


二、存储器的层次结构

1. 存储器三个主要特性的关系



2. 缓存—主存层次和主存—辅存层次 4.1



(速度) (容量)
缓存—主存 主存—辅存

主存储器

虚拟存储器

实地址

虚地址

物理地址

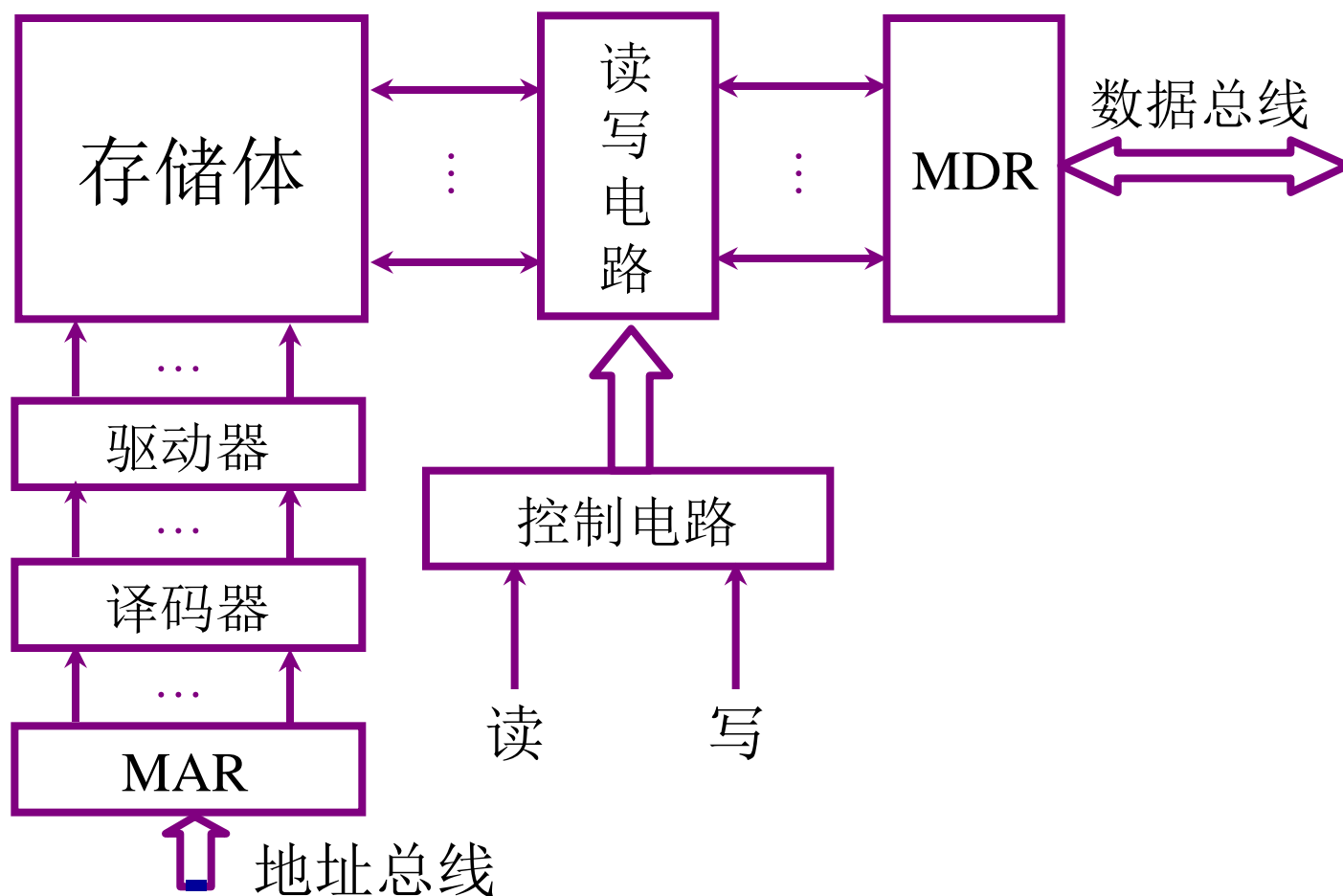
逻辑地址



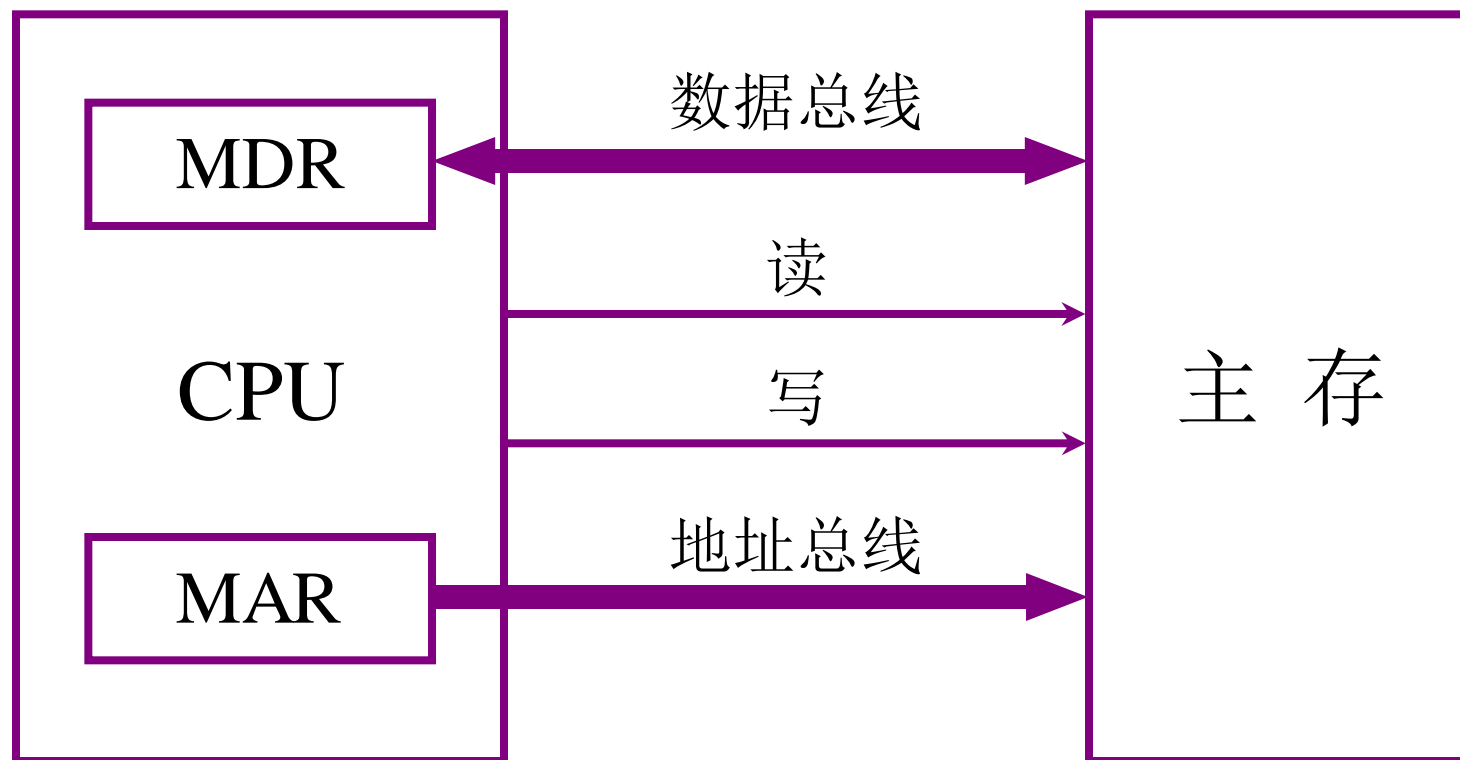
4.2 主存储器

一、概述

1. 主存的基本组成



2. 主存和 CPU 的联系



3. 主存中存储单元地址的分配

高位字节 地址为字地址

低位字节 地址为字地址

字地址	字节地址			
0	0	1	2	3
4	4	5	6	7
8	8	9	10	11

字地址	字节地址	
0	1	0
2	3	2
4	5	4

设地址线 24 根

按 字节 寻址 $2^{24} = 16 \text{ M}$

若字长为 16 位

按 字 寻址 8 M

若字长为 32 位

按 字 寻址 4 M

4. 主存的技术指标

(1) 存储容量 主存 存放二进制代码的总位数

(2) 存储速度

- 存取时间 存储器的 访问时间

读出时间 写入时间

- 存取周期 连续两次独立的存储器操作

（读或写）所需的 最小间隔时间

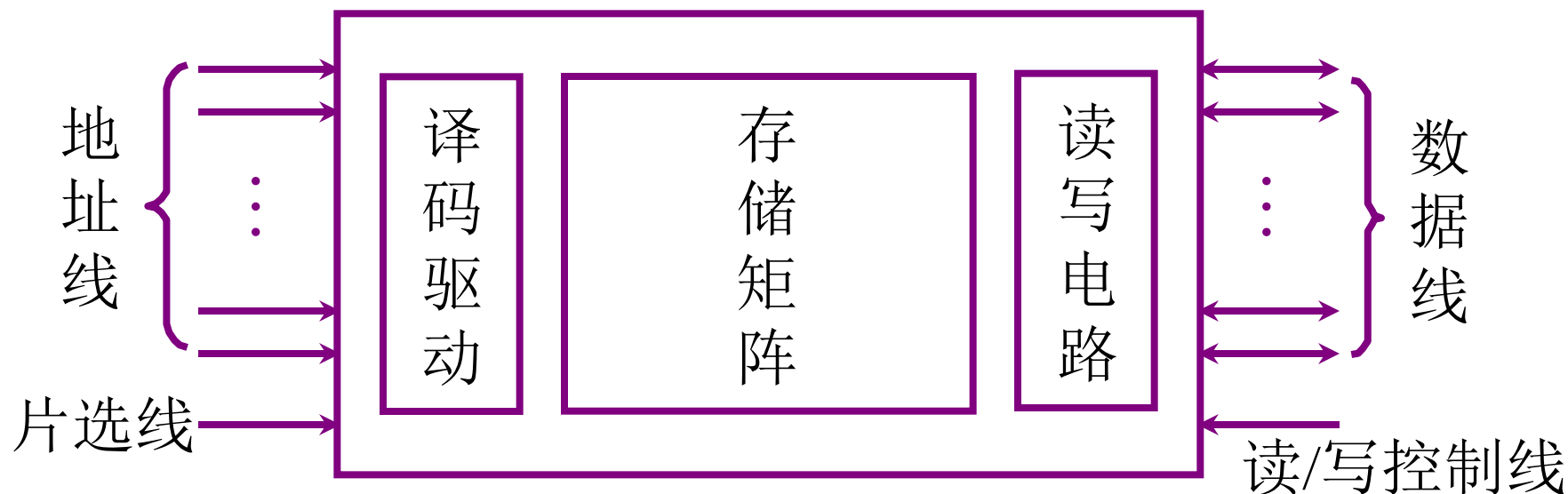
读周期 写周期

(3) 存储器的带宽 位/秒



二、半导体存储芯片简介

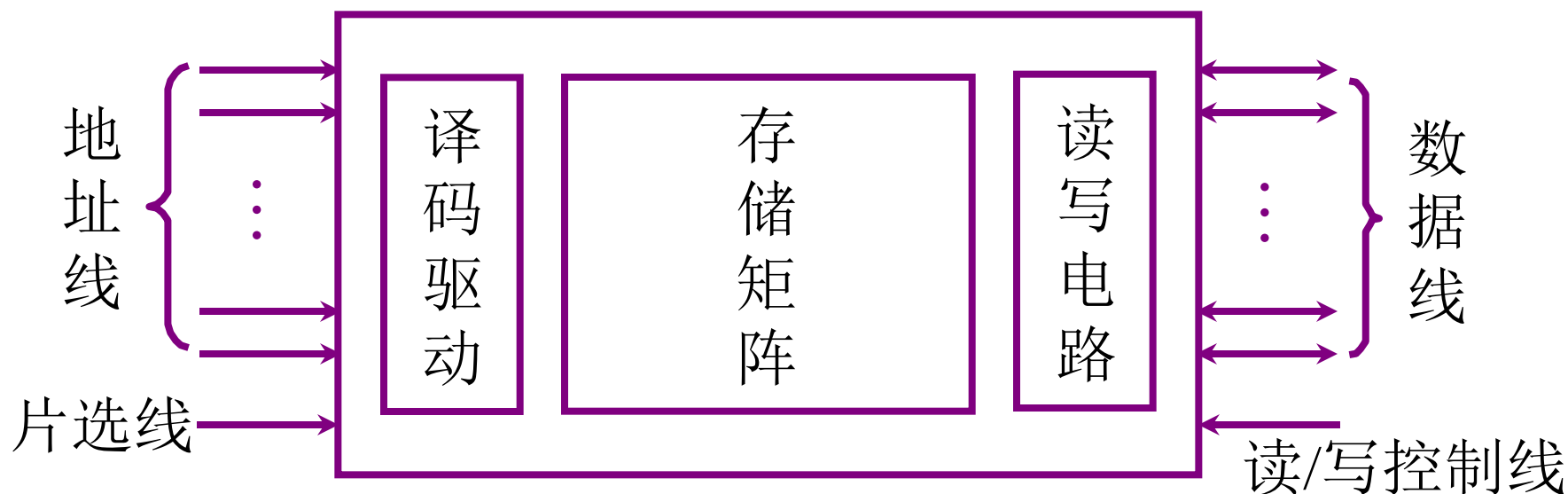
1. 半导体存储芯片的基本结构



地址线（单向）	数据线（双向）	芯片容量
10	4	1K×4位
14	1	16K×1位
13	8	8K×8位

二、半导体存储芯片简介

1. 半导体存储芯片的基本结构



片选线 $\overline{\text{CS}}$ $\overline{\text{CE}}$

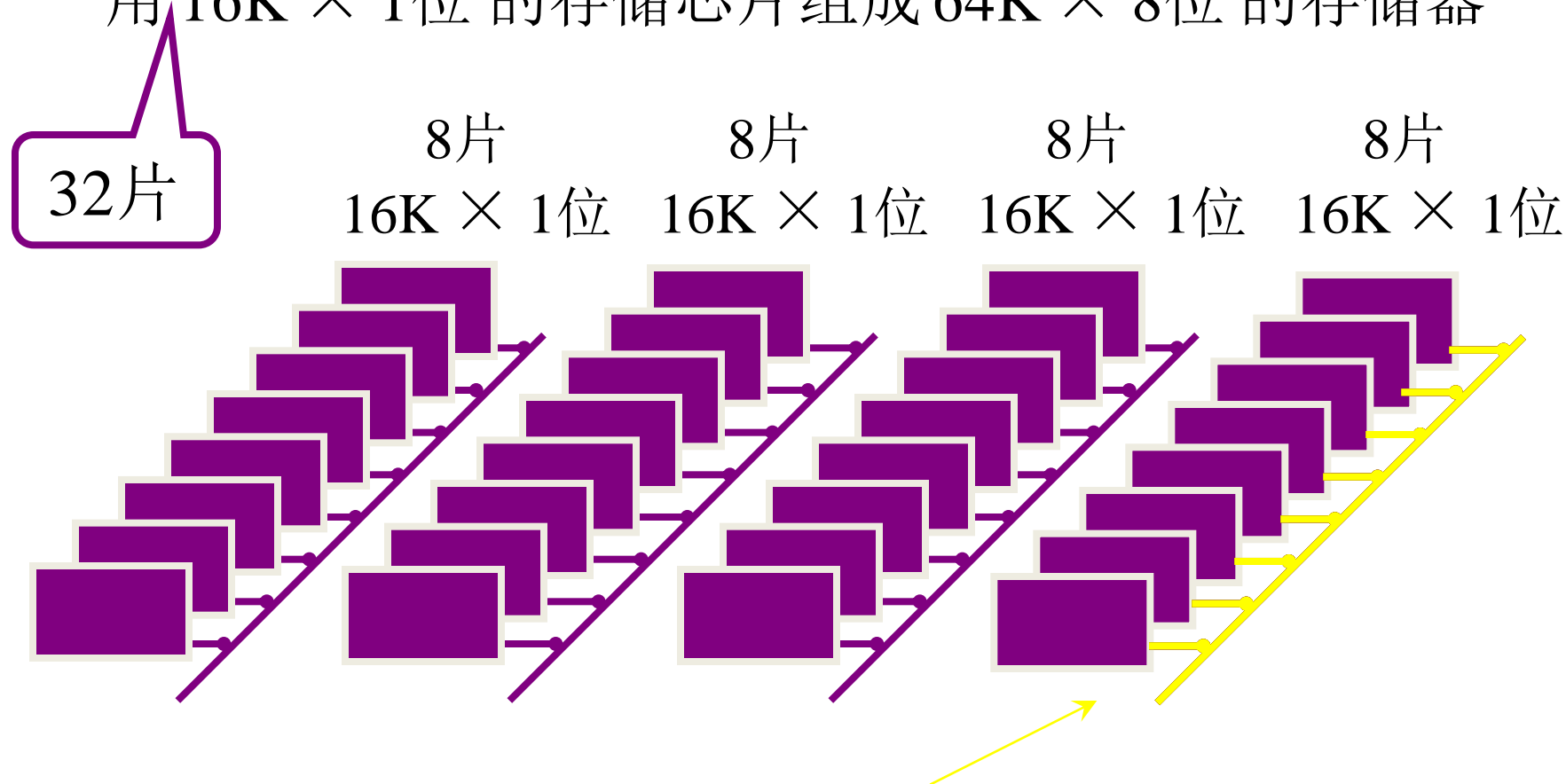
读/写控制线 $\overline{\text{WE}}$ (低电平写 高电平读)

$\overline{\text{OE}}$ (允许读) $\overline{\text{WE}}$ (允许写)



存储芯片片选线的作用

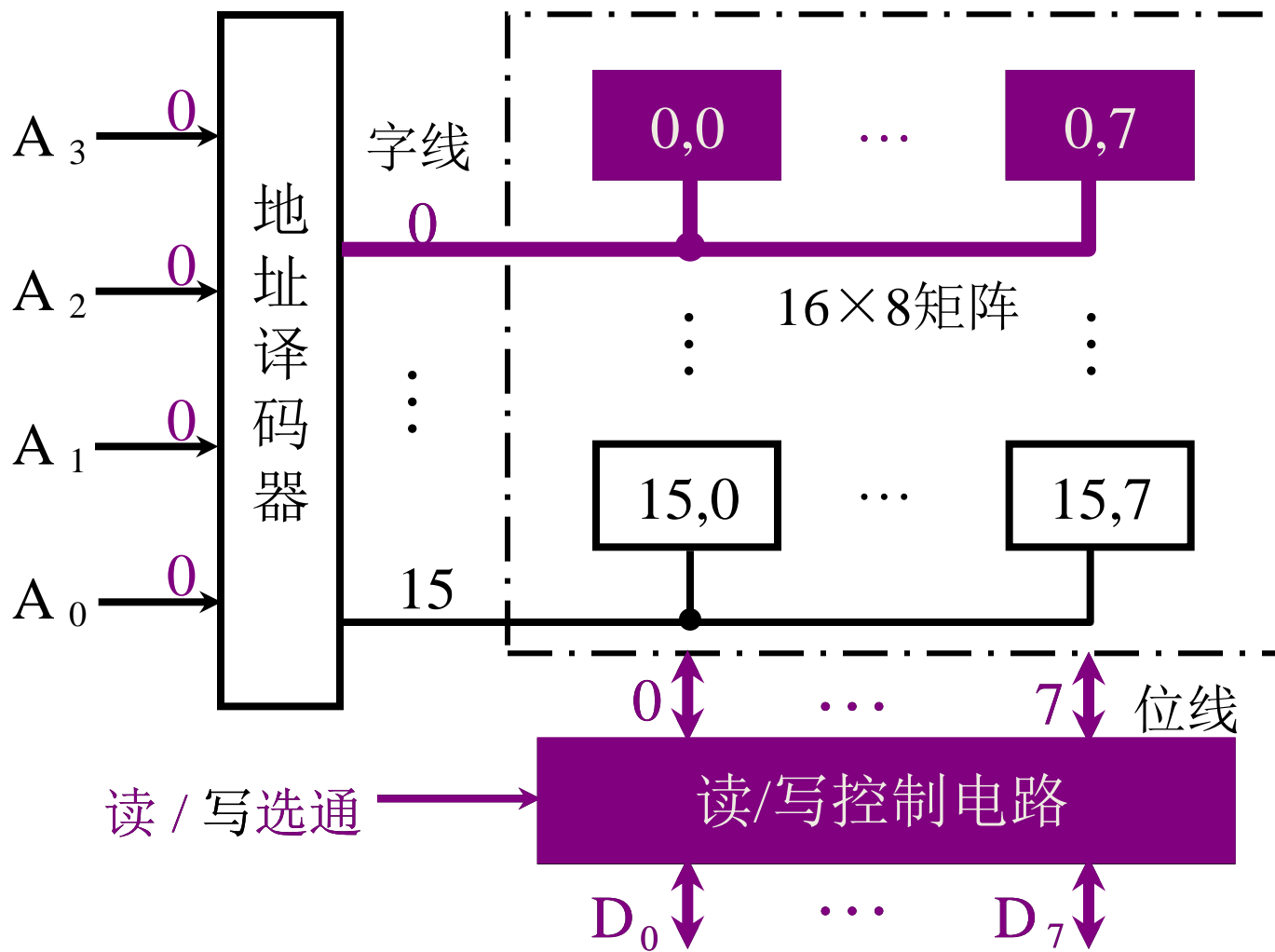
用 $16\text{K} \times 1$ 位的存储芯片组成 $64\text{K} \times 8$ 位的存储器



当地址为 65 535 时，此 8 片的片选有效

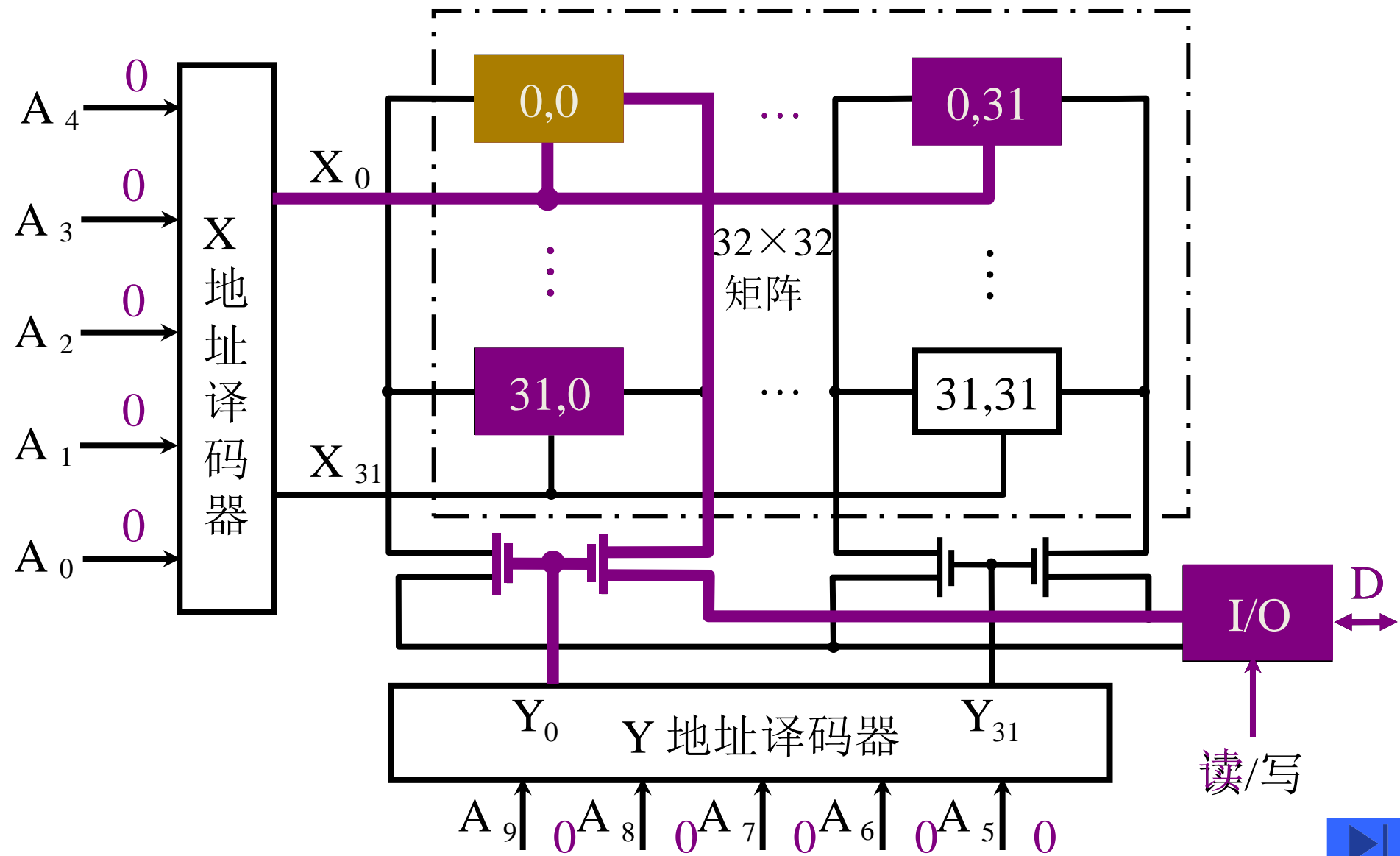
2. 半导体存储芯片的译码驱动方式 4.2

(1) 线选法



(2) 重合法

4.2

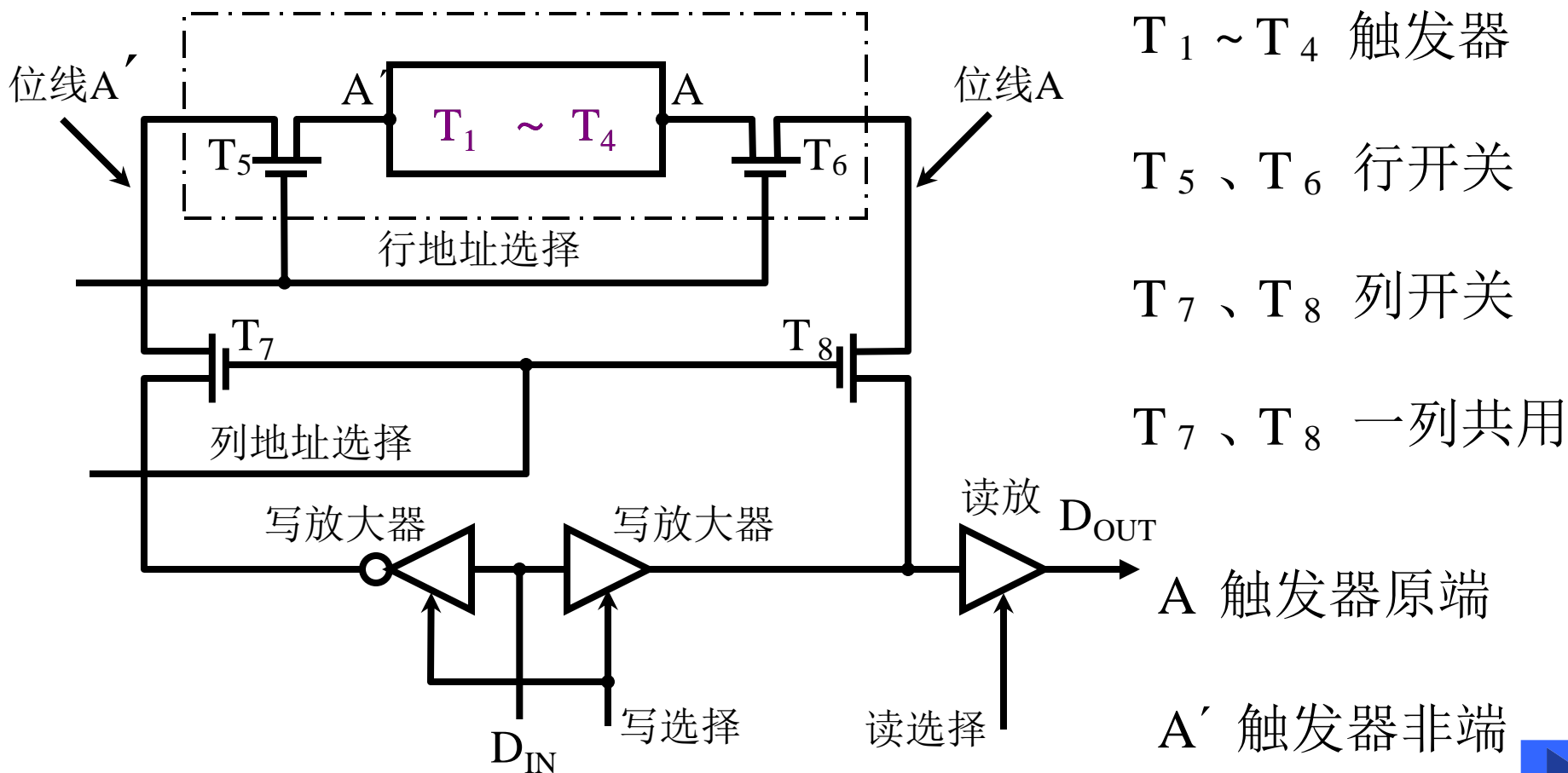


三、随机存取存储器 (RAM)

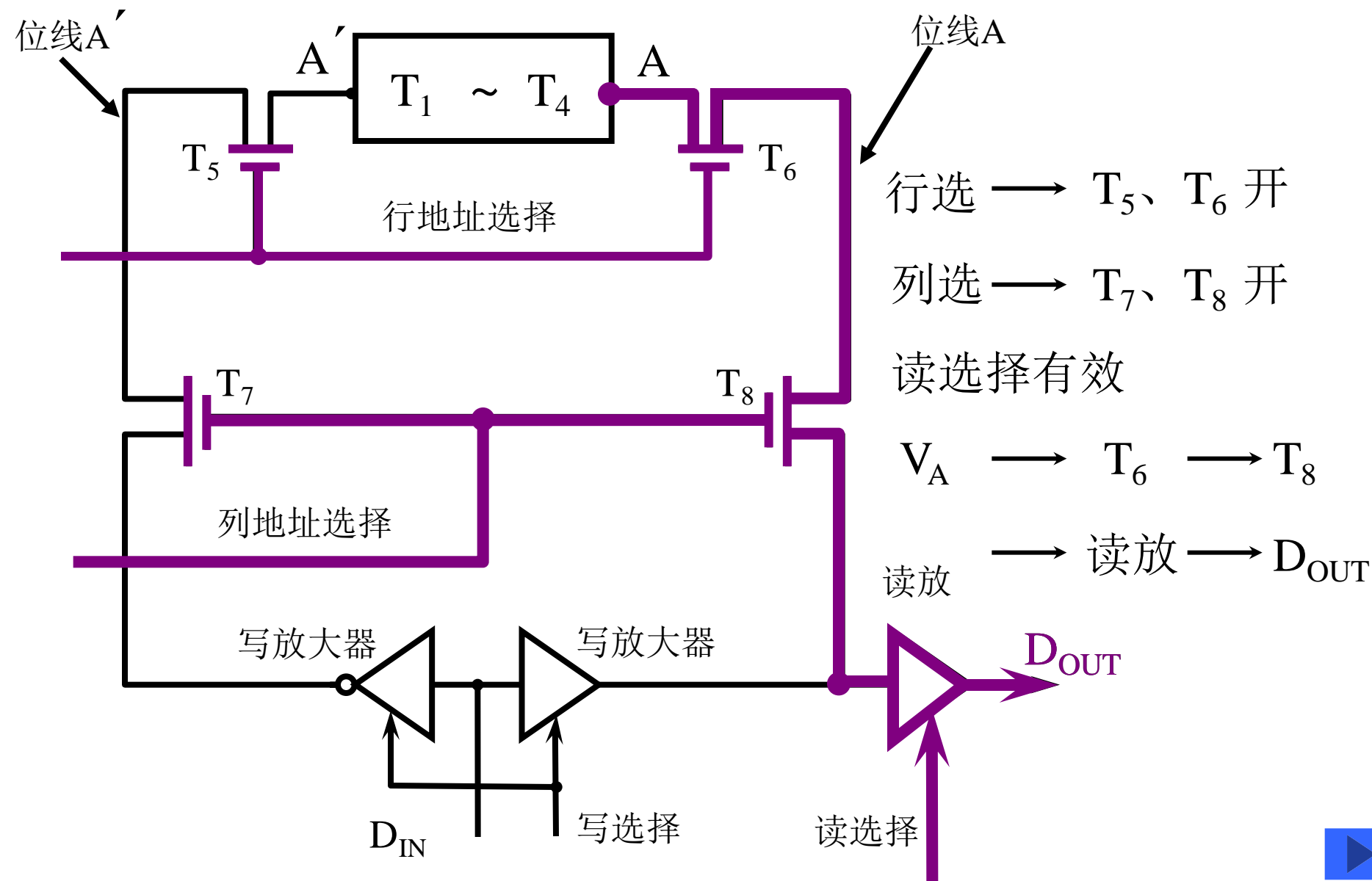
4.2

1. 静态 RAM (SRAM)

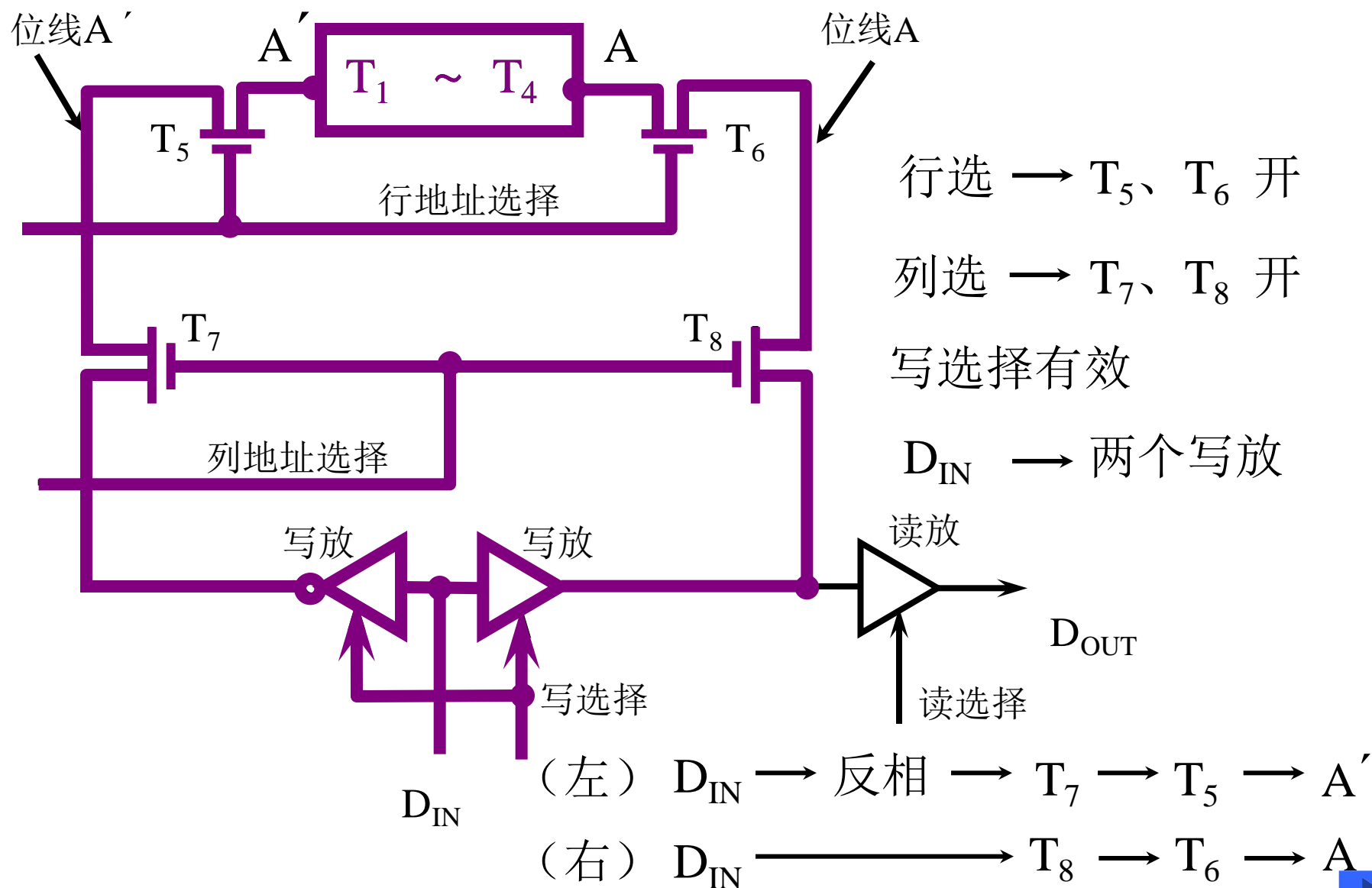
(1) 静态 RAM 基本电路



① 静态 RAM 基本电路的 读 操作

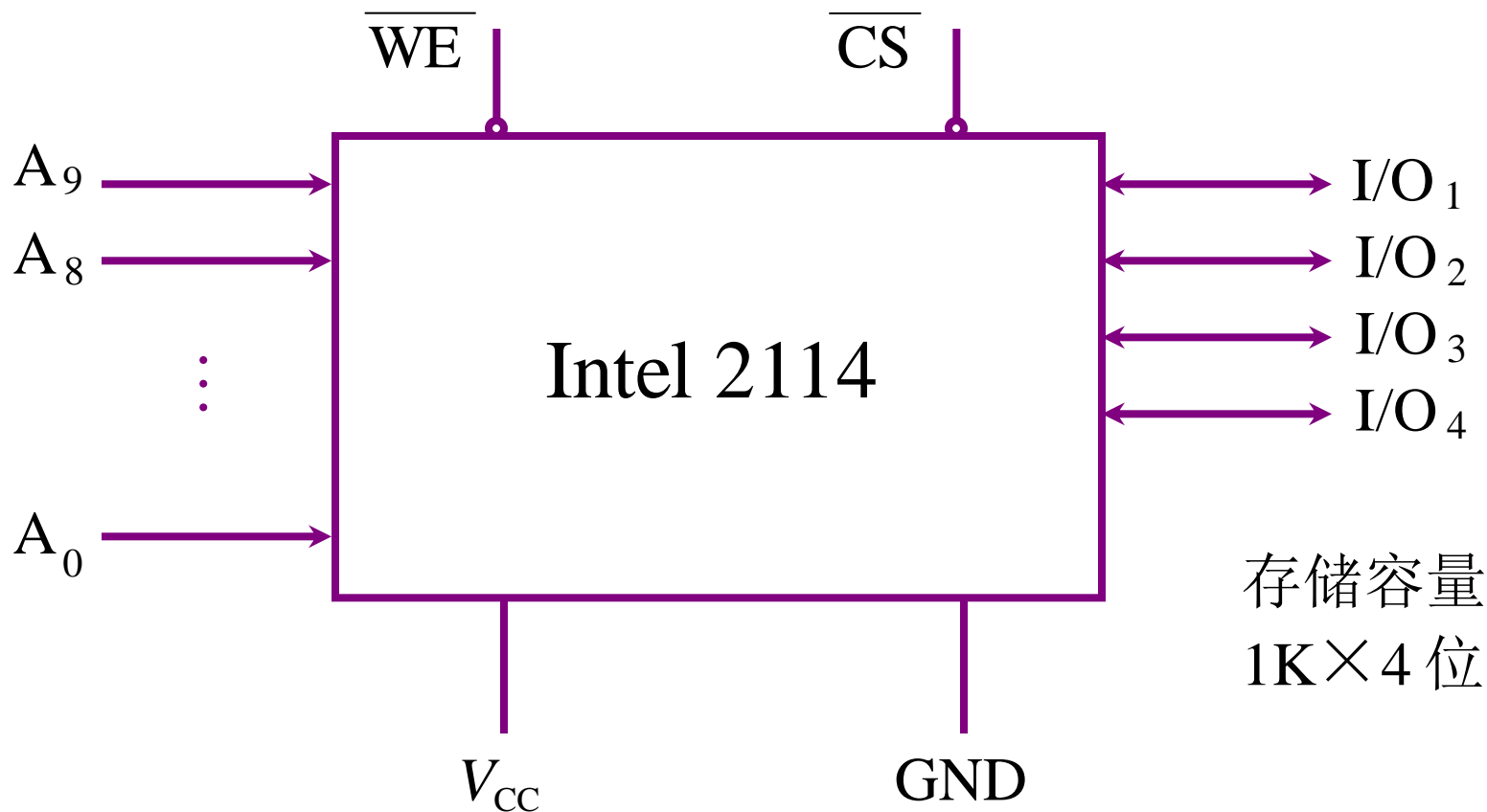


② 静态 RAM 基本电路的 写 操作

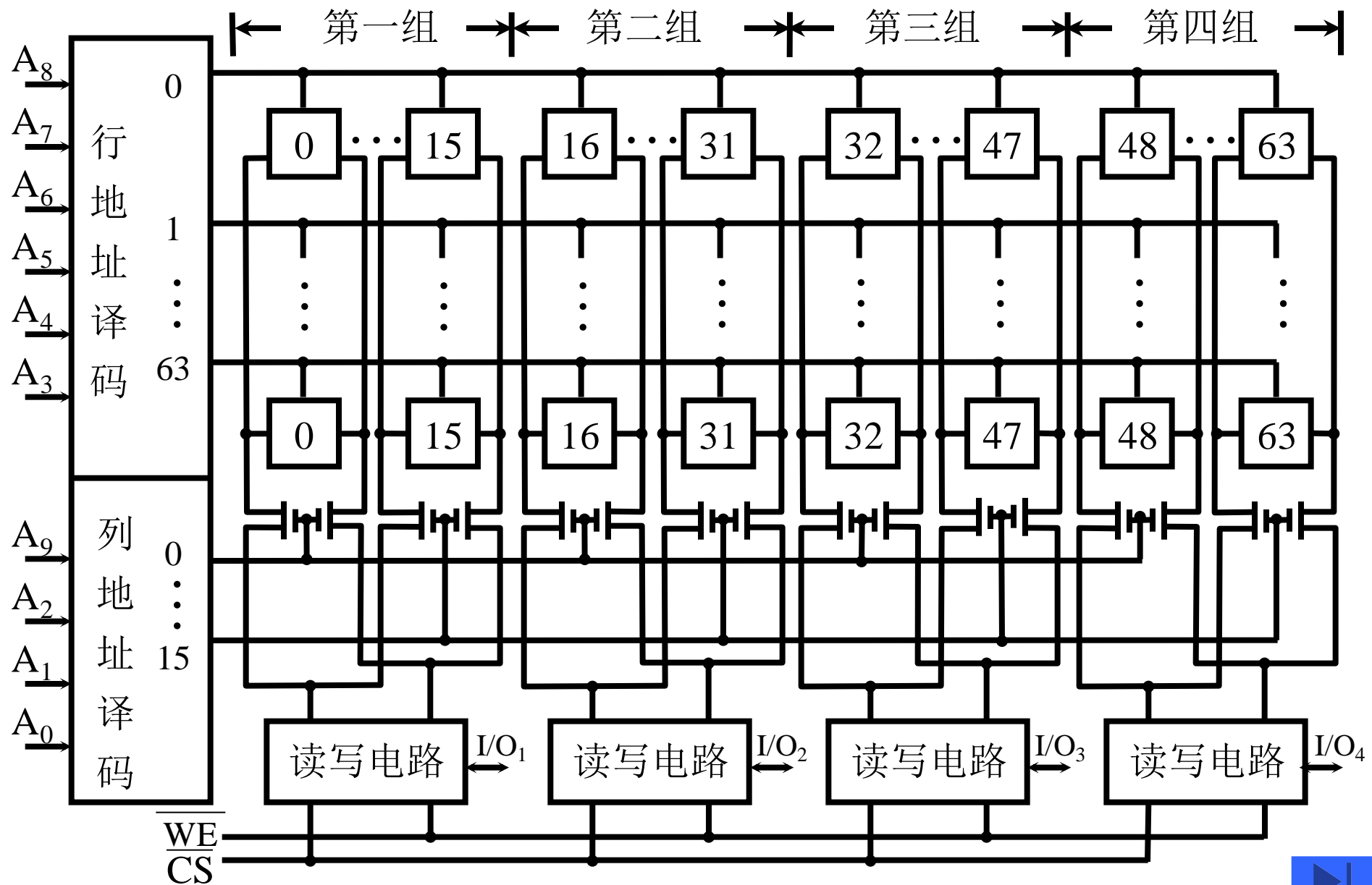


(2) 静态 RAM 芯片举例

① Intel 2114 外特性



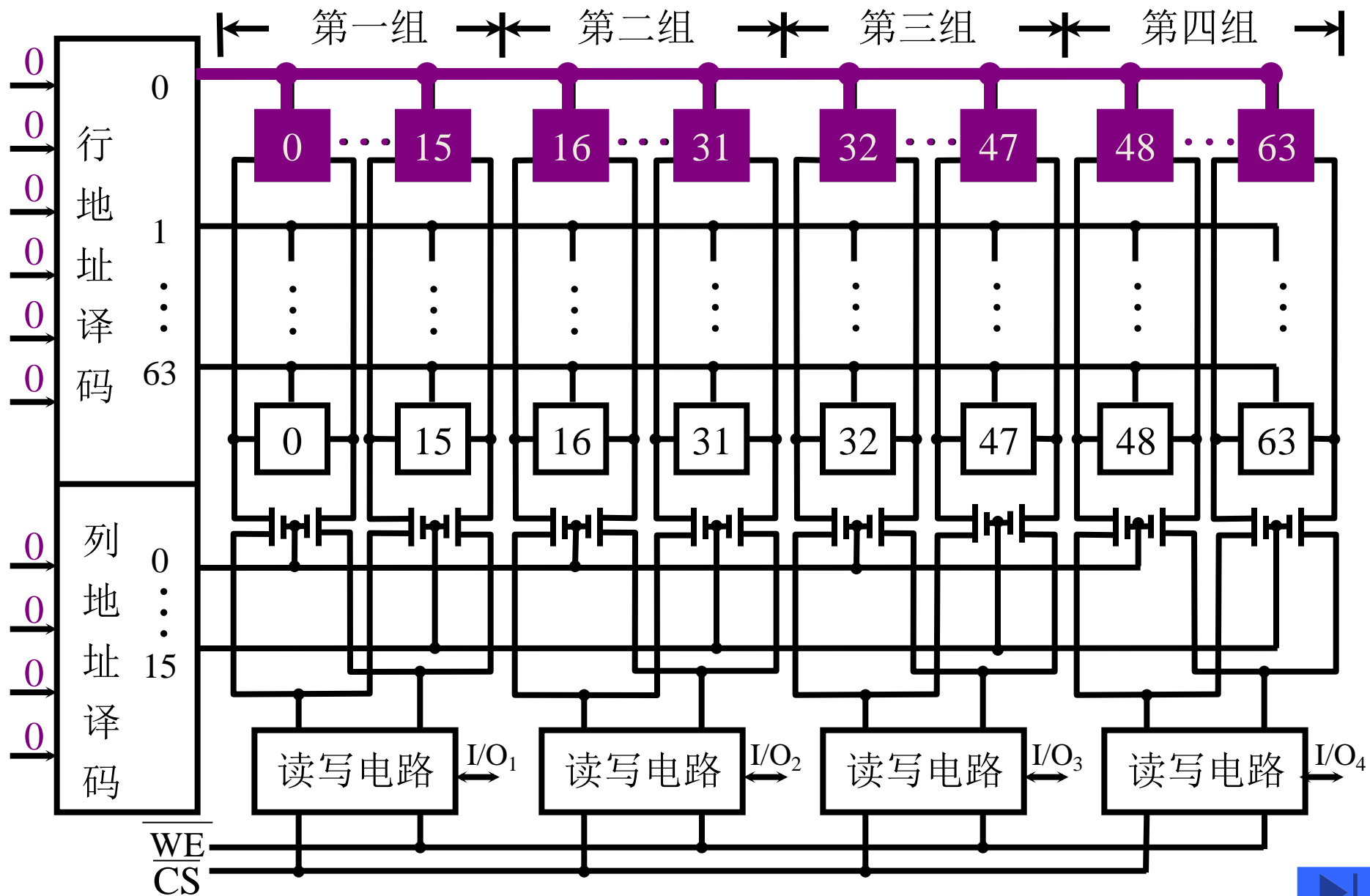
② Intel 2114 RAM 矩阵 (64×64) 读 4.2



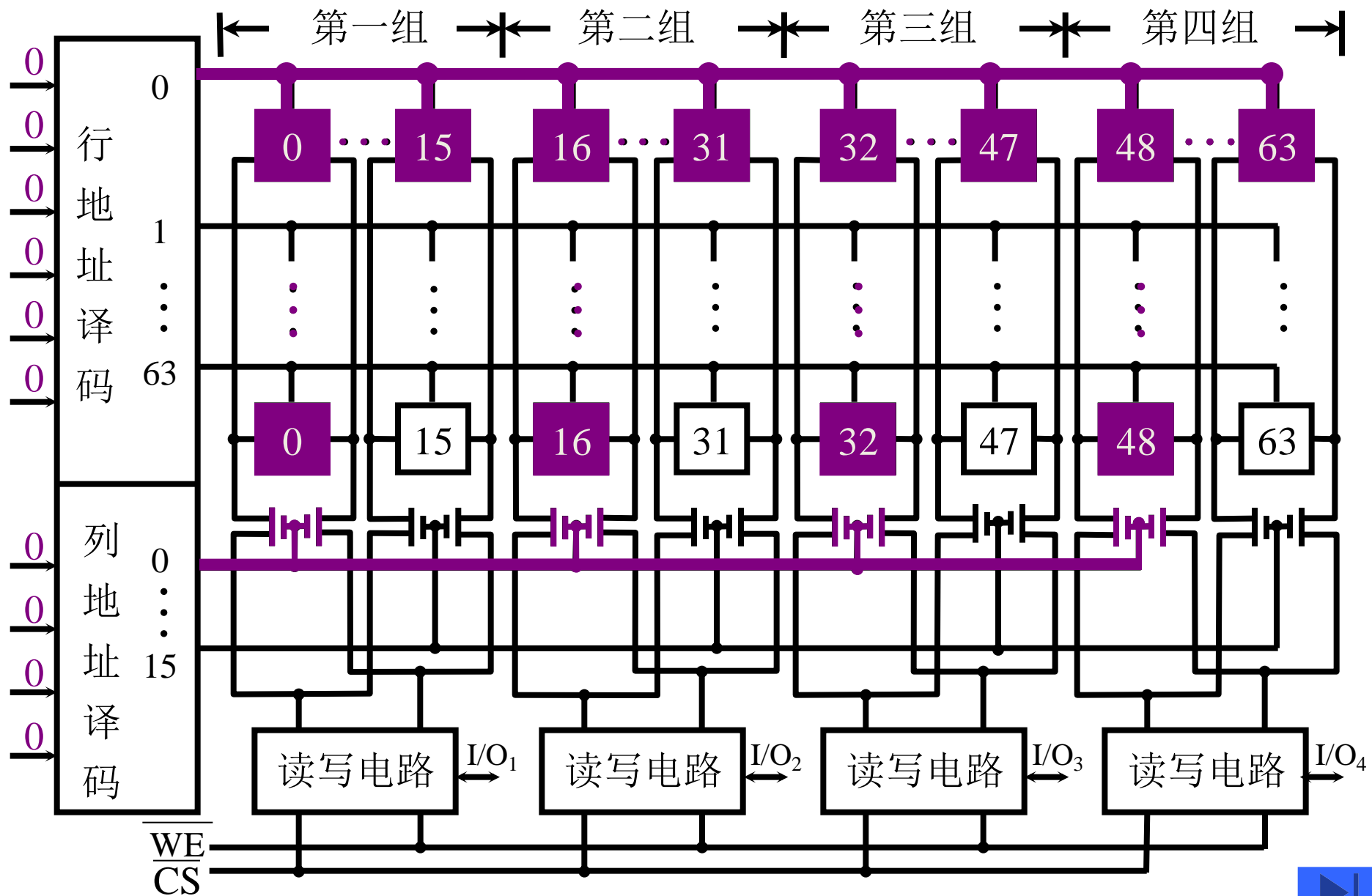
4.2



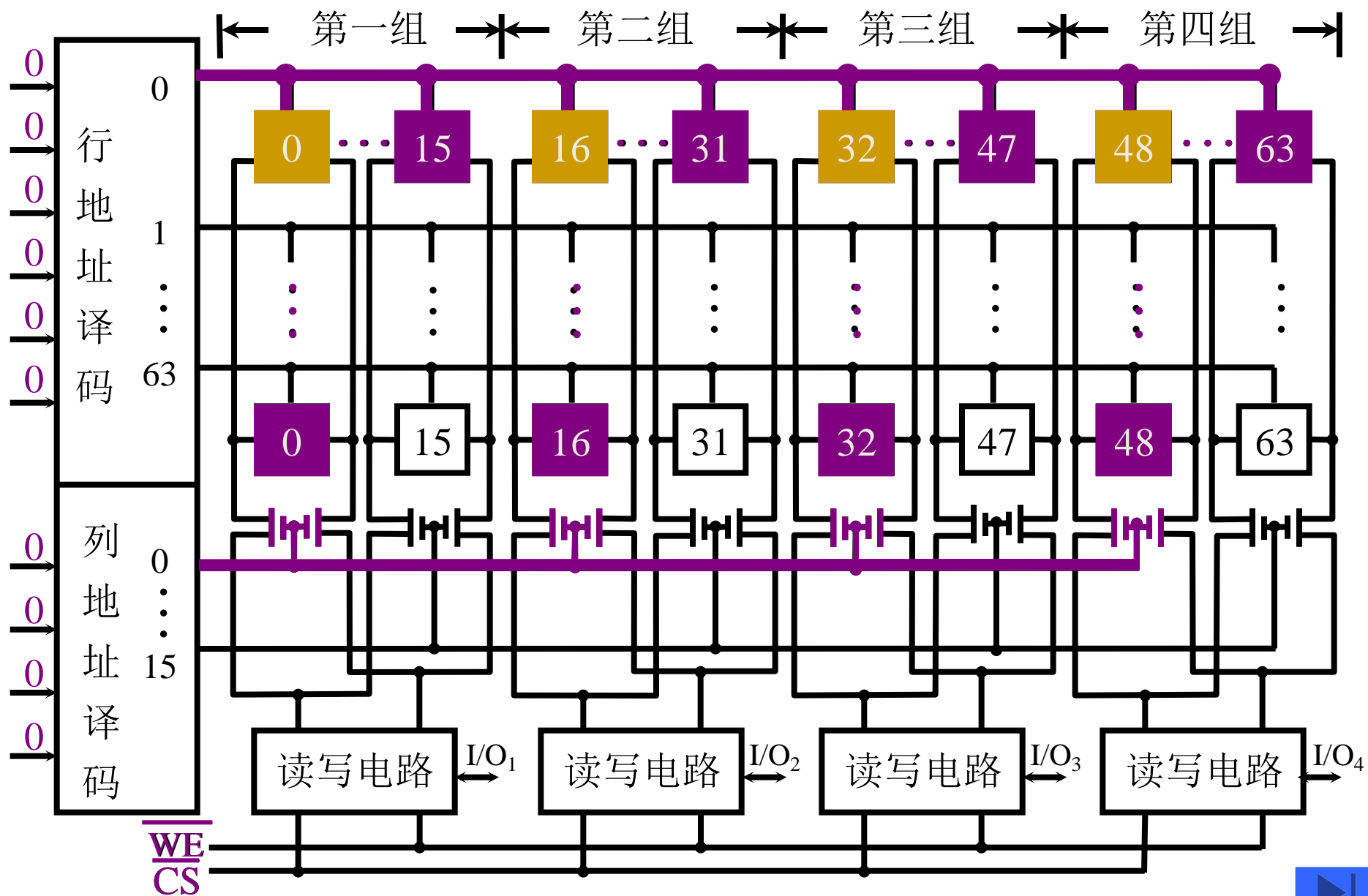
② Intel 2114 RAM 矩阵 (64×64) 读 4.2



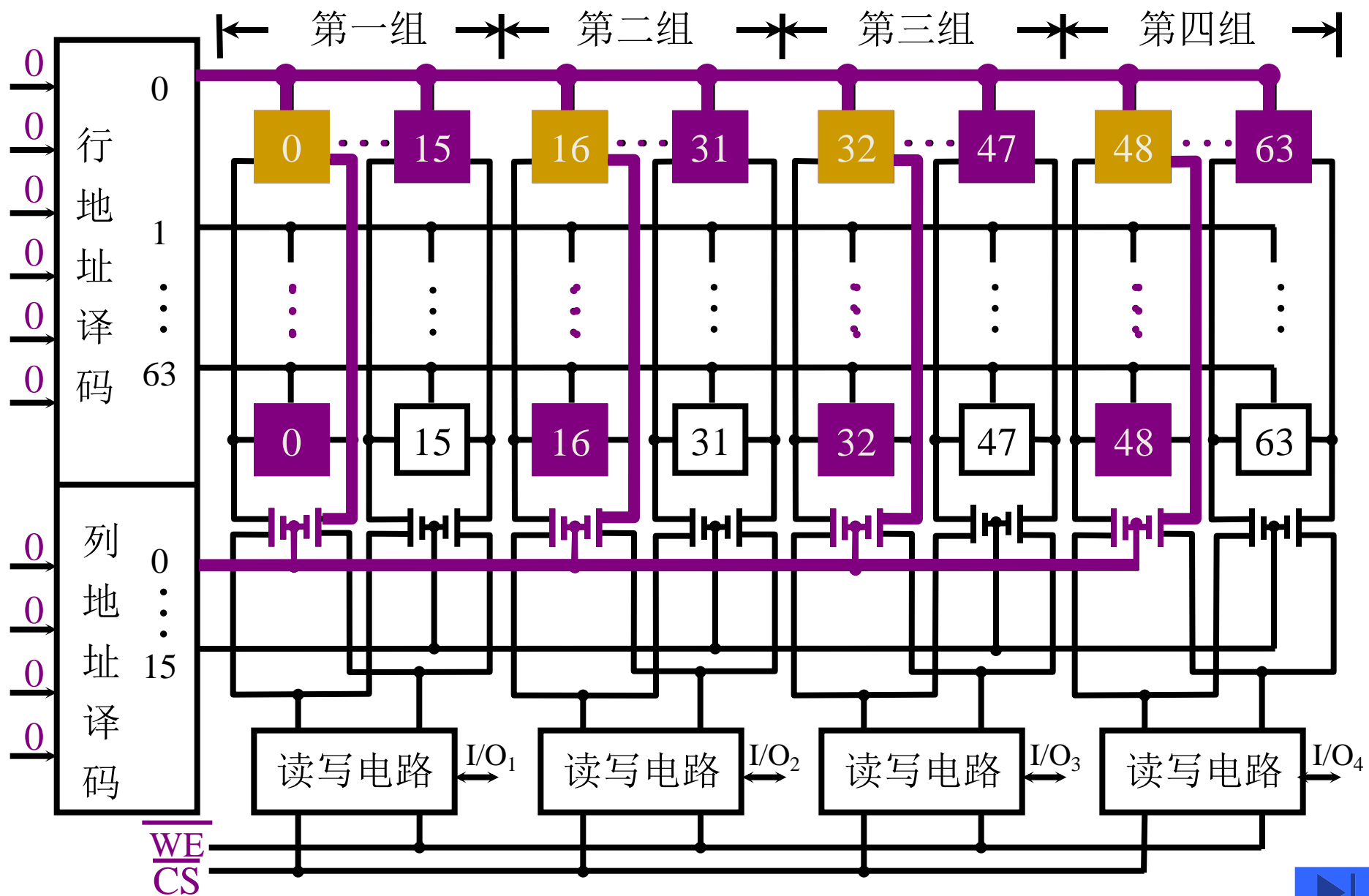
② Intel 2114 RAM 矩阵 (64×64) 读 4.2



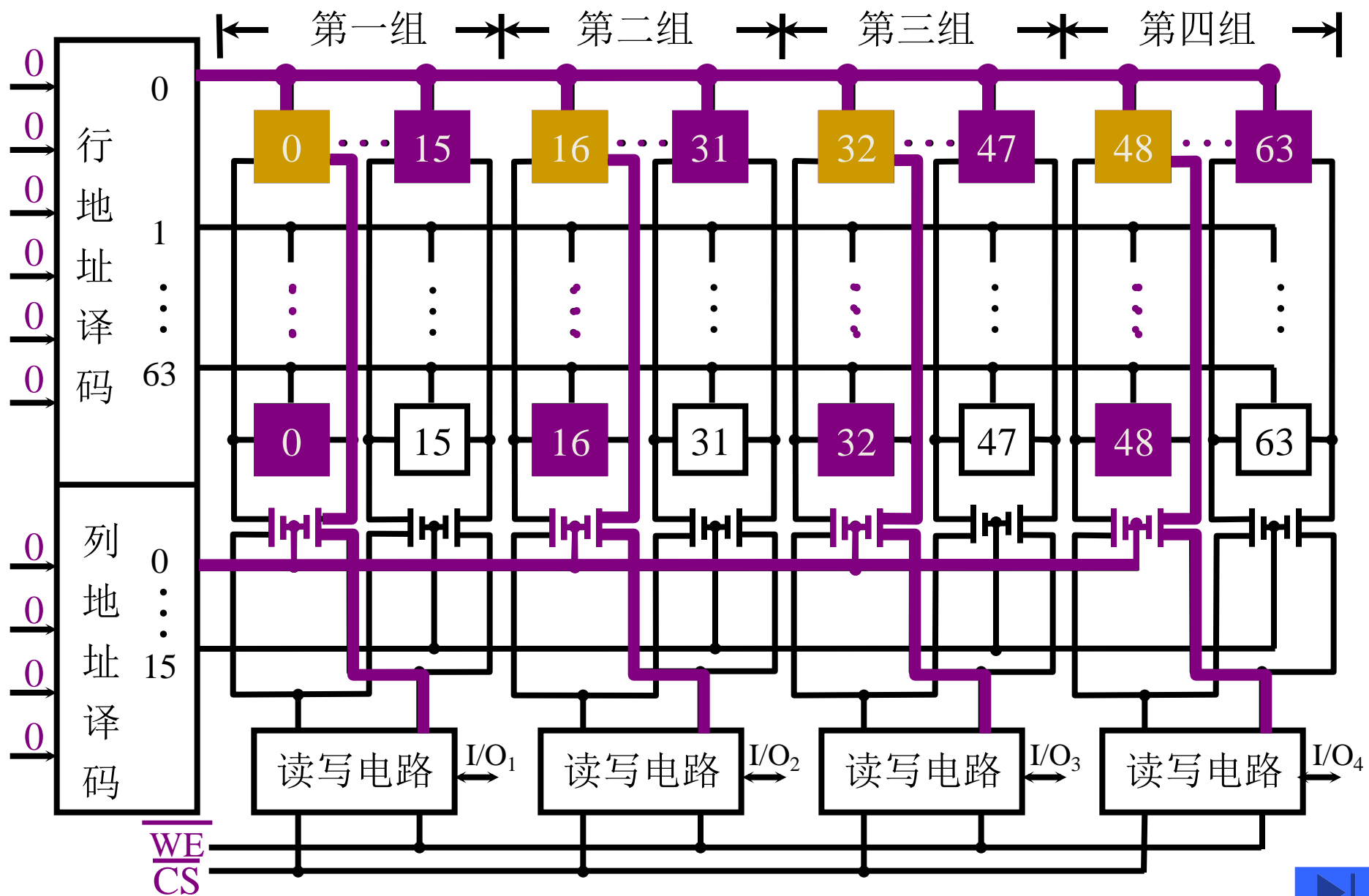
② Intel 2114 RAM 矩阵 (64 × 64) 读 4.2



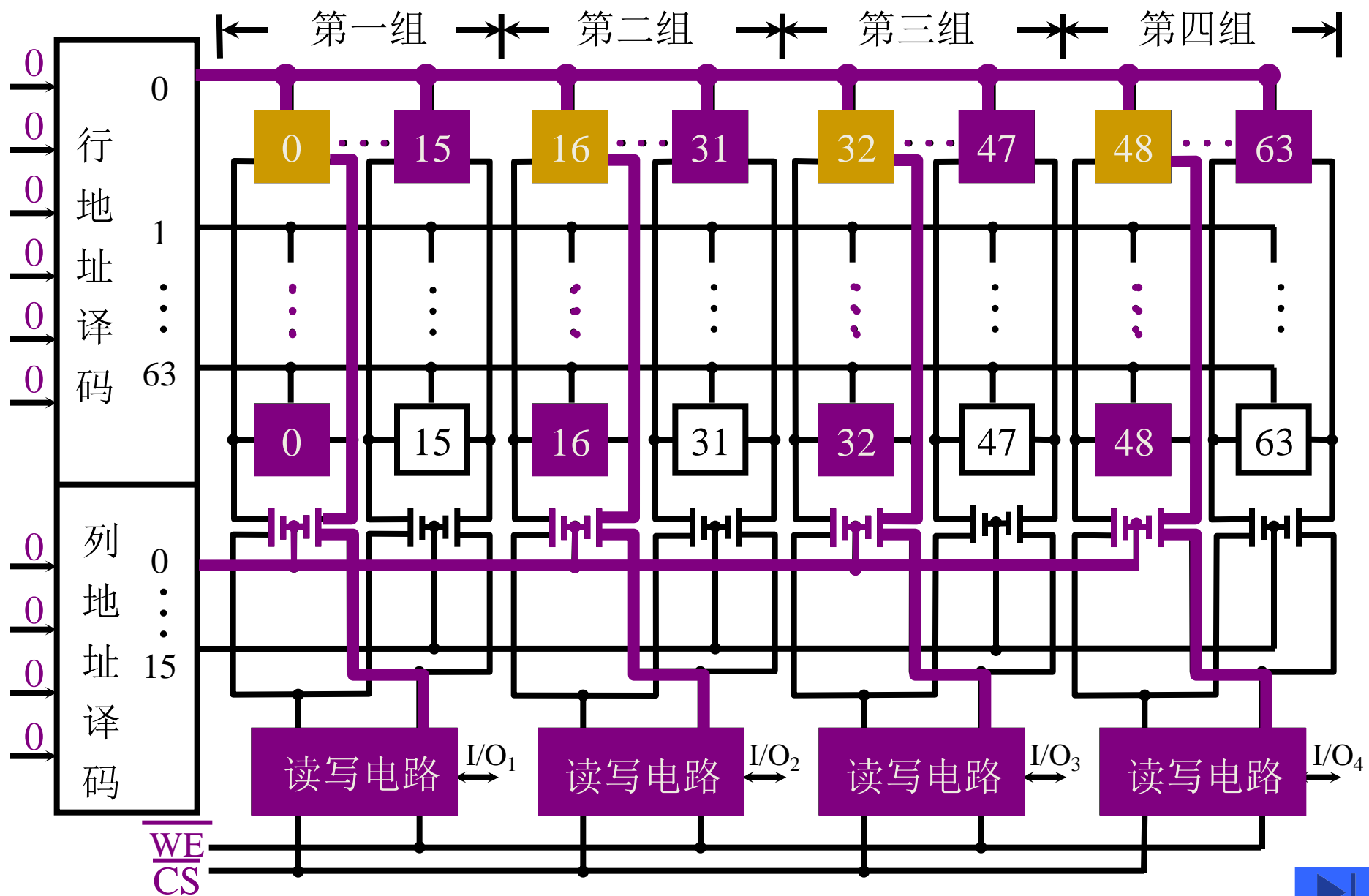
② Intel 2114 RAM 矩阵 (64 × 64) 读 4.2



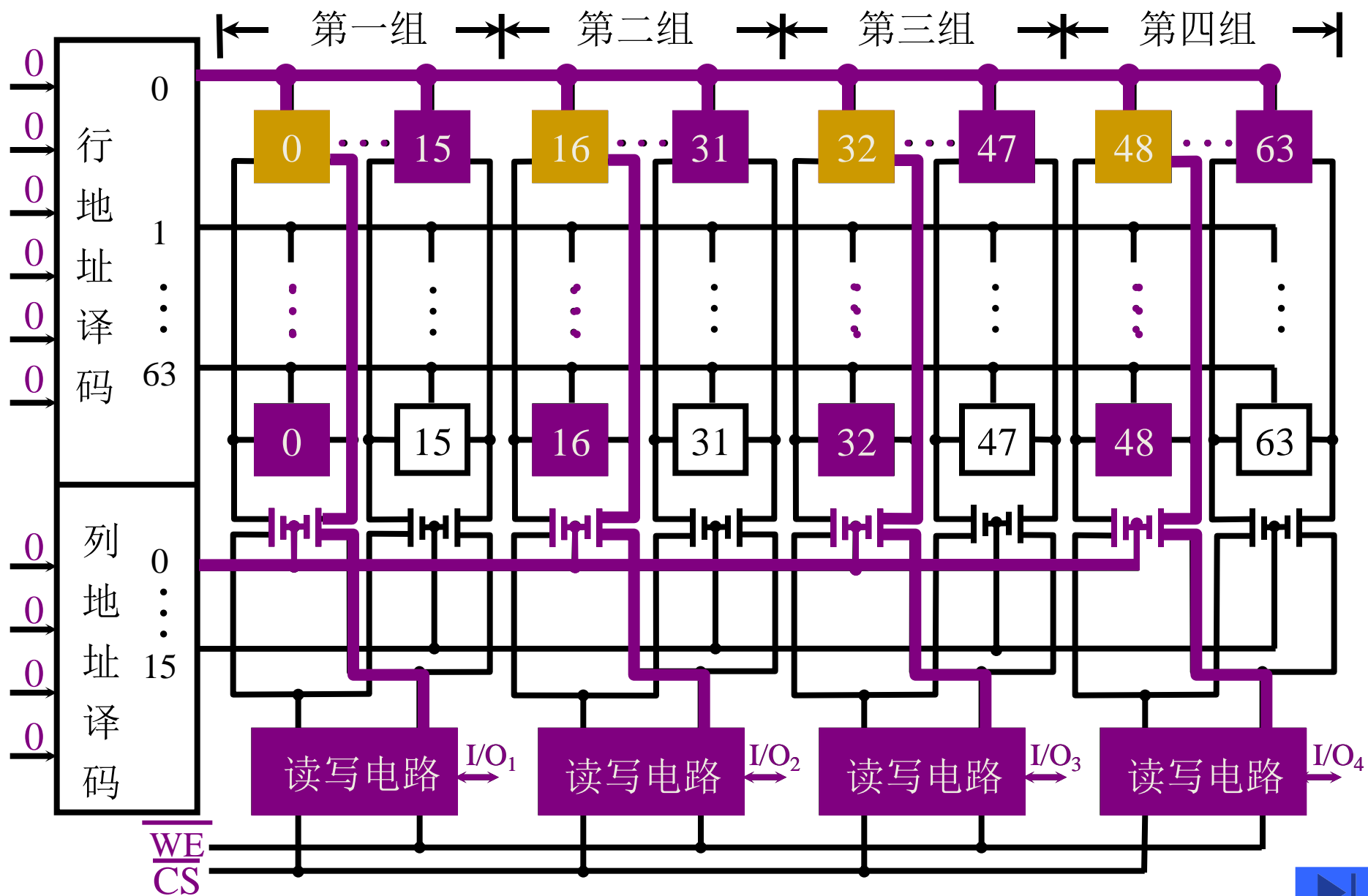
② Intel 2114 RAM 矩阵 (64 × 64) 读 4.2



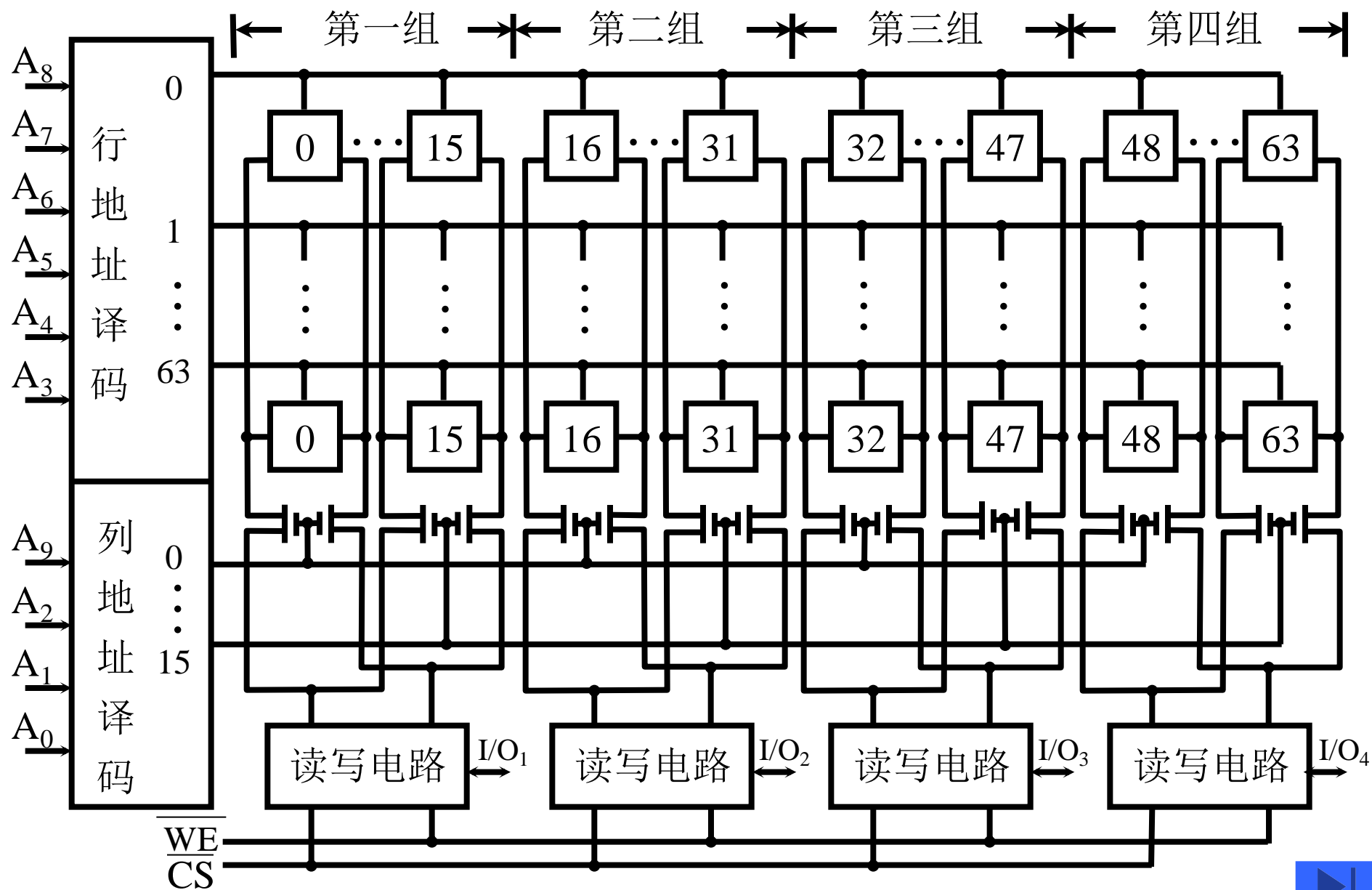
② Intel 2114 RAM 矩阵 (64 × 64) 读 4.2



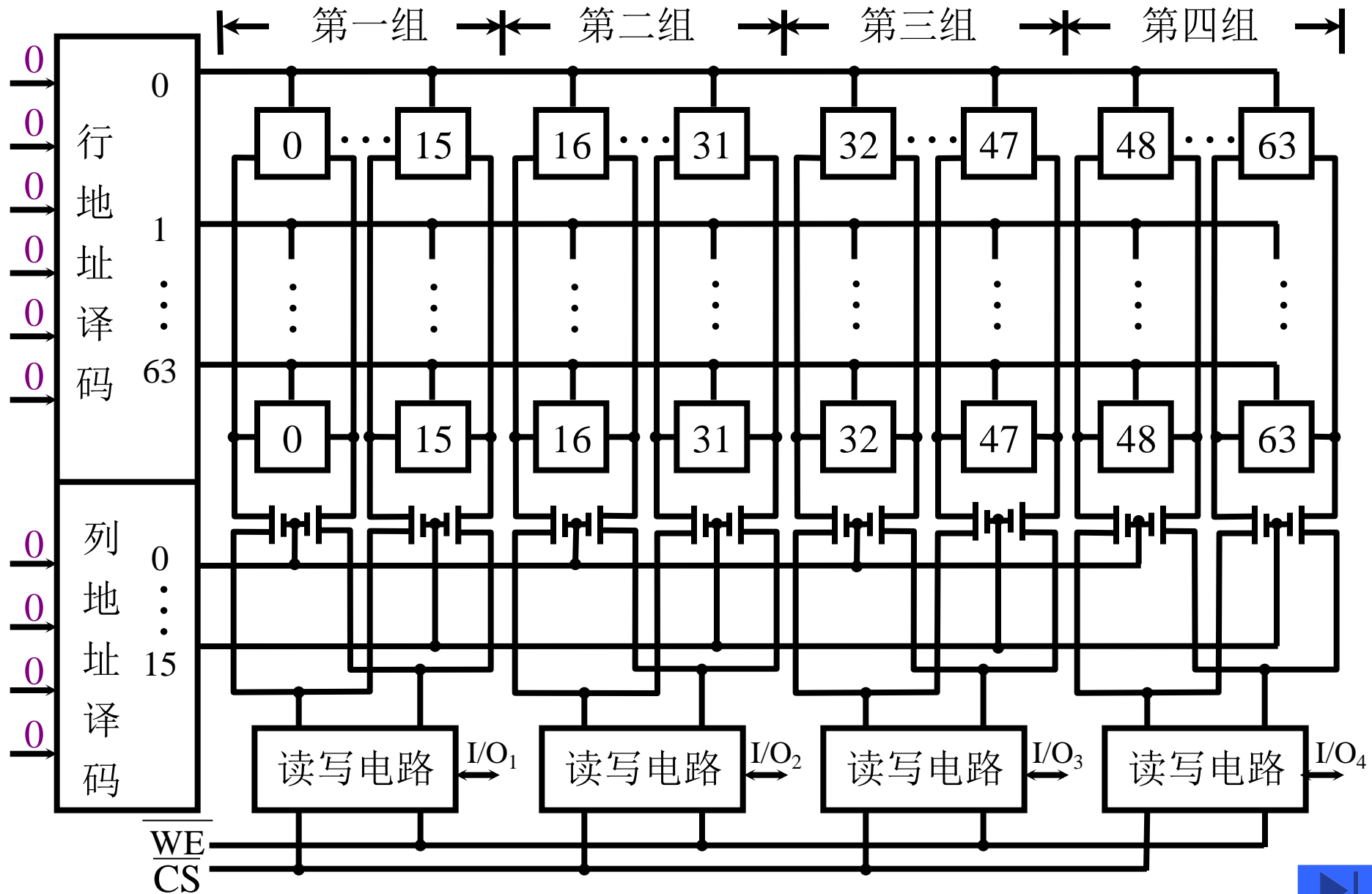
② Intel 2114 RAM 矩阵 (64×64) 读 4.2



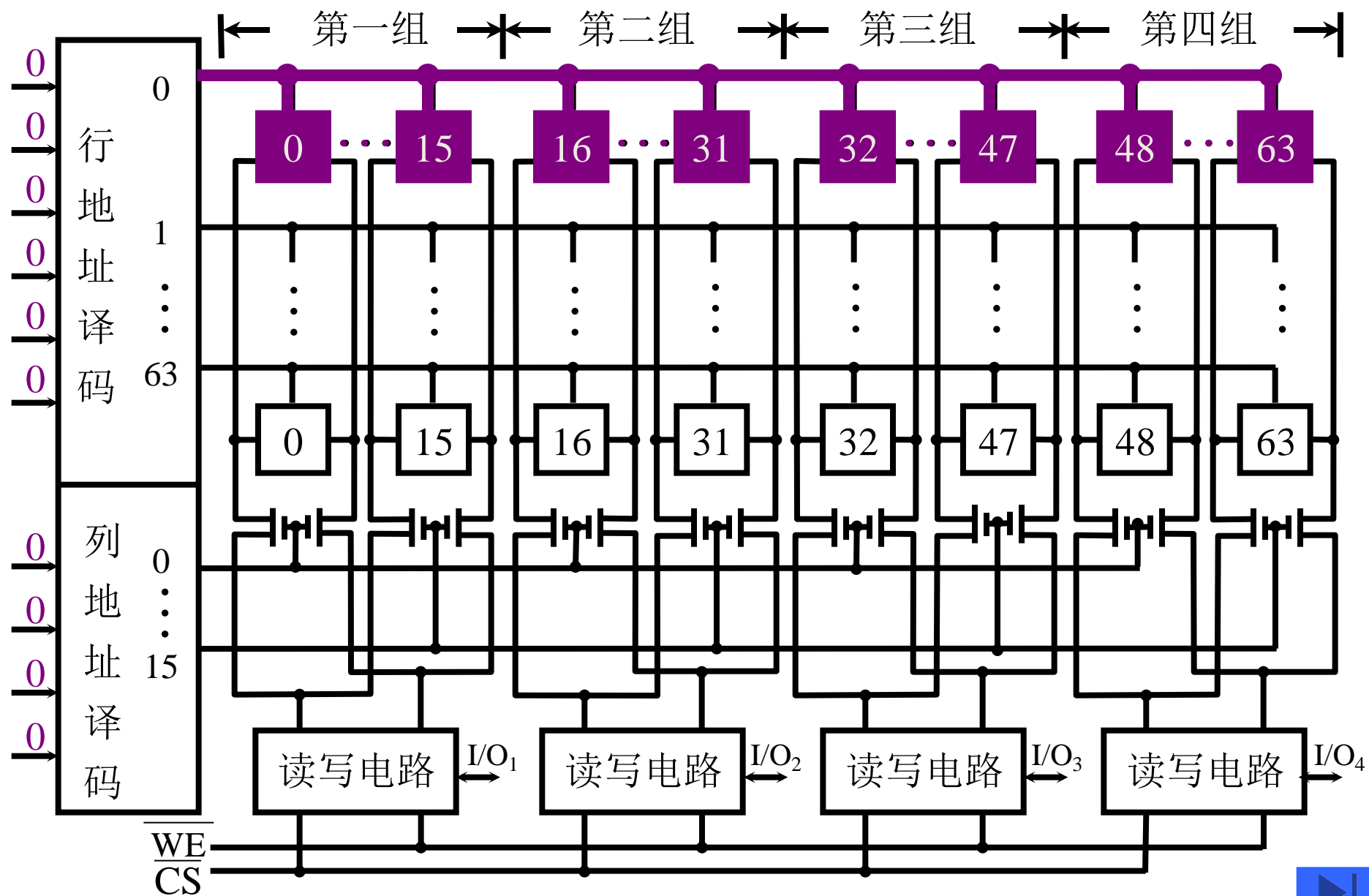
③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2



③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2



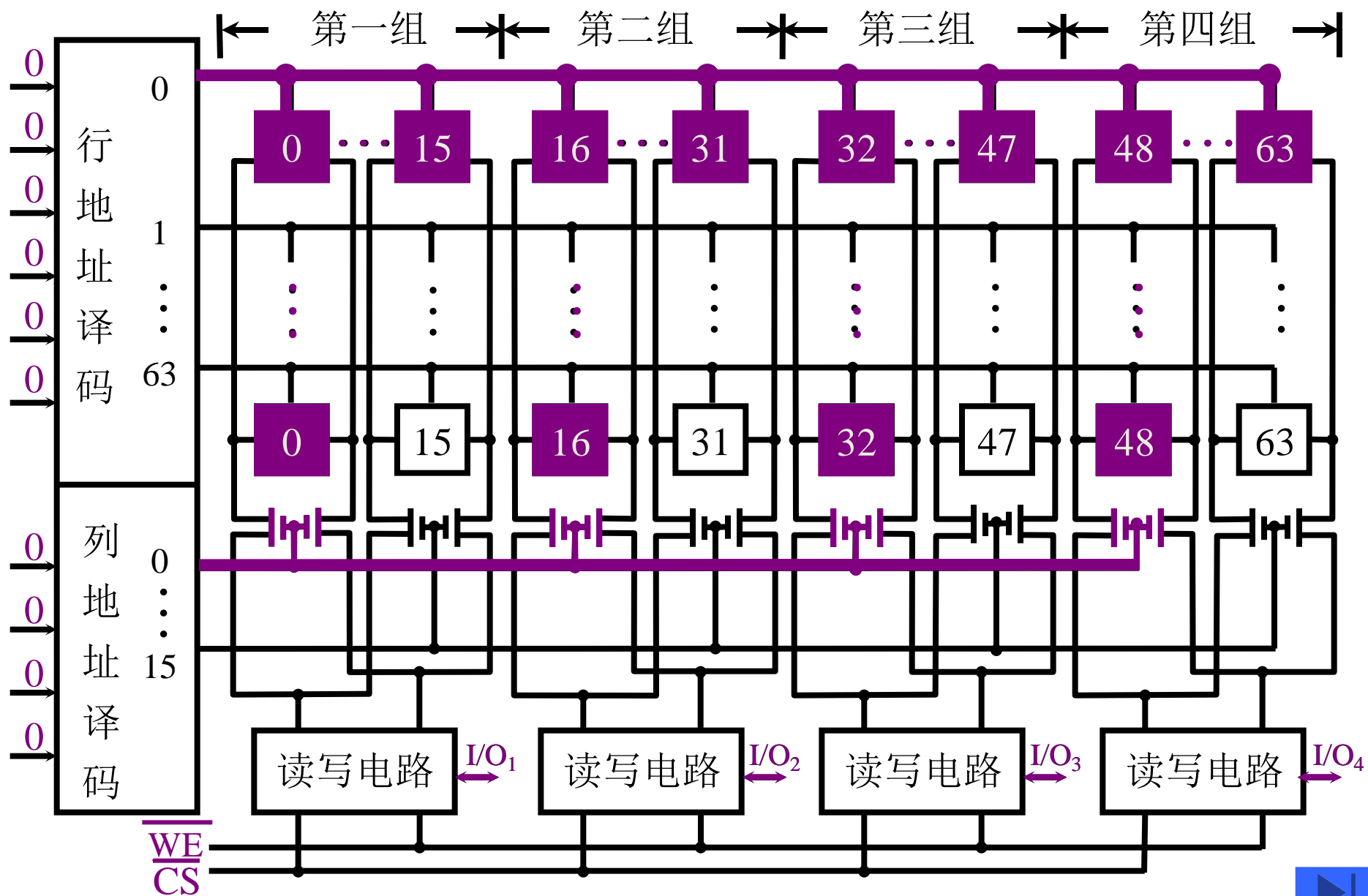
③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2



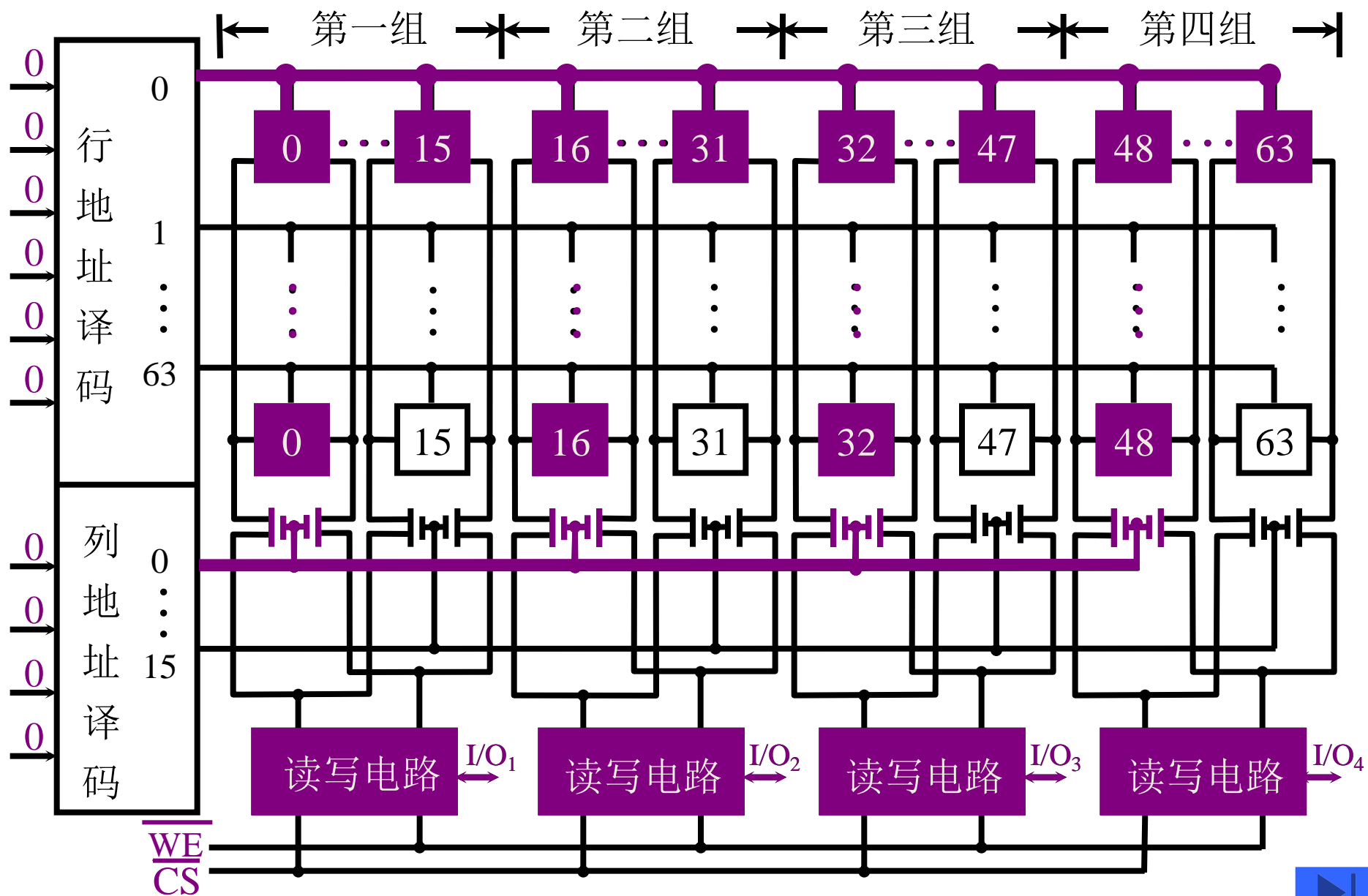
4.2



③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2



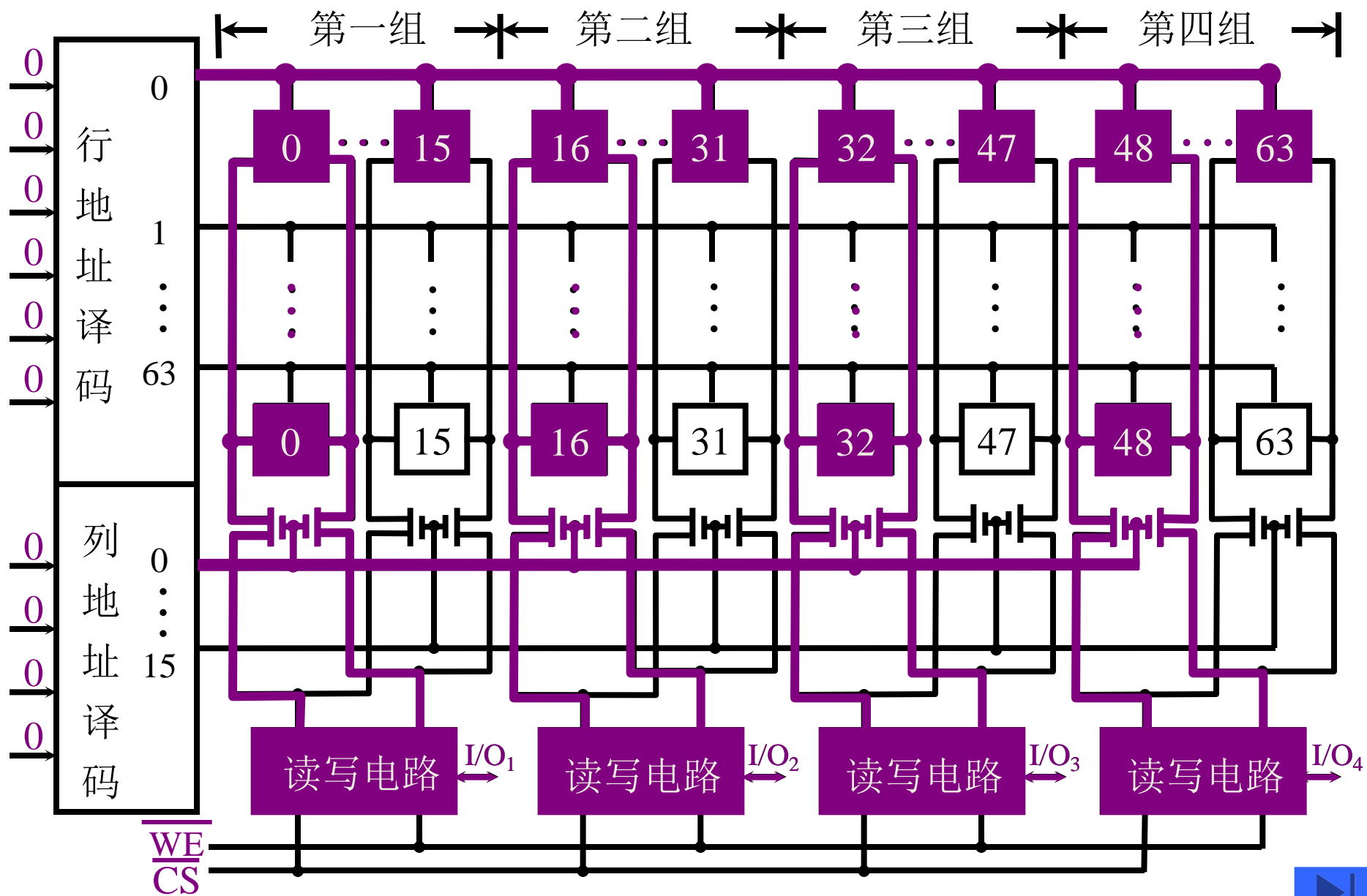
③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2



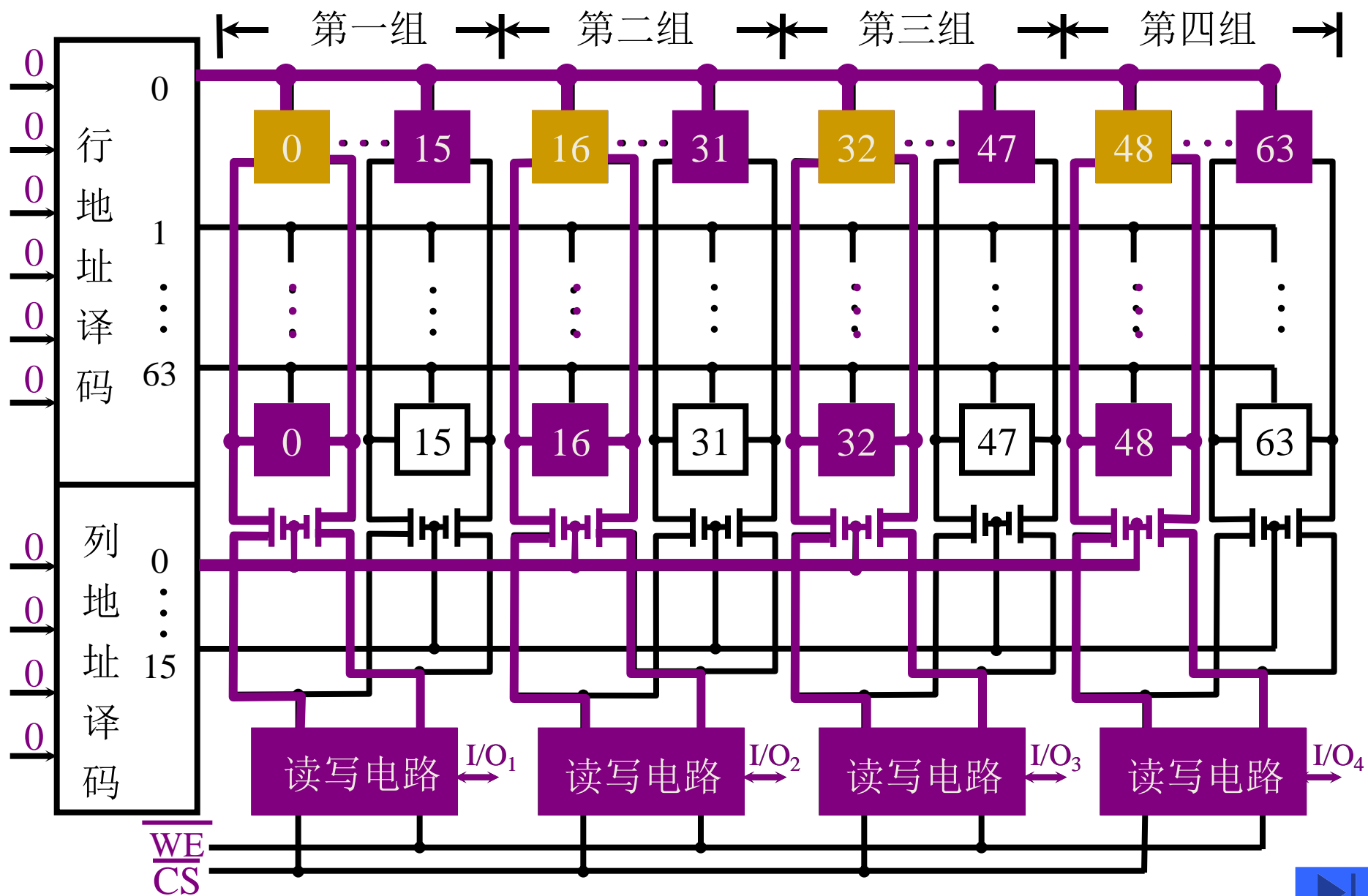
4.2



③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2

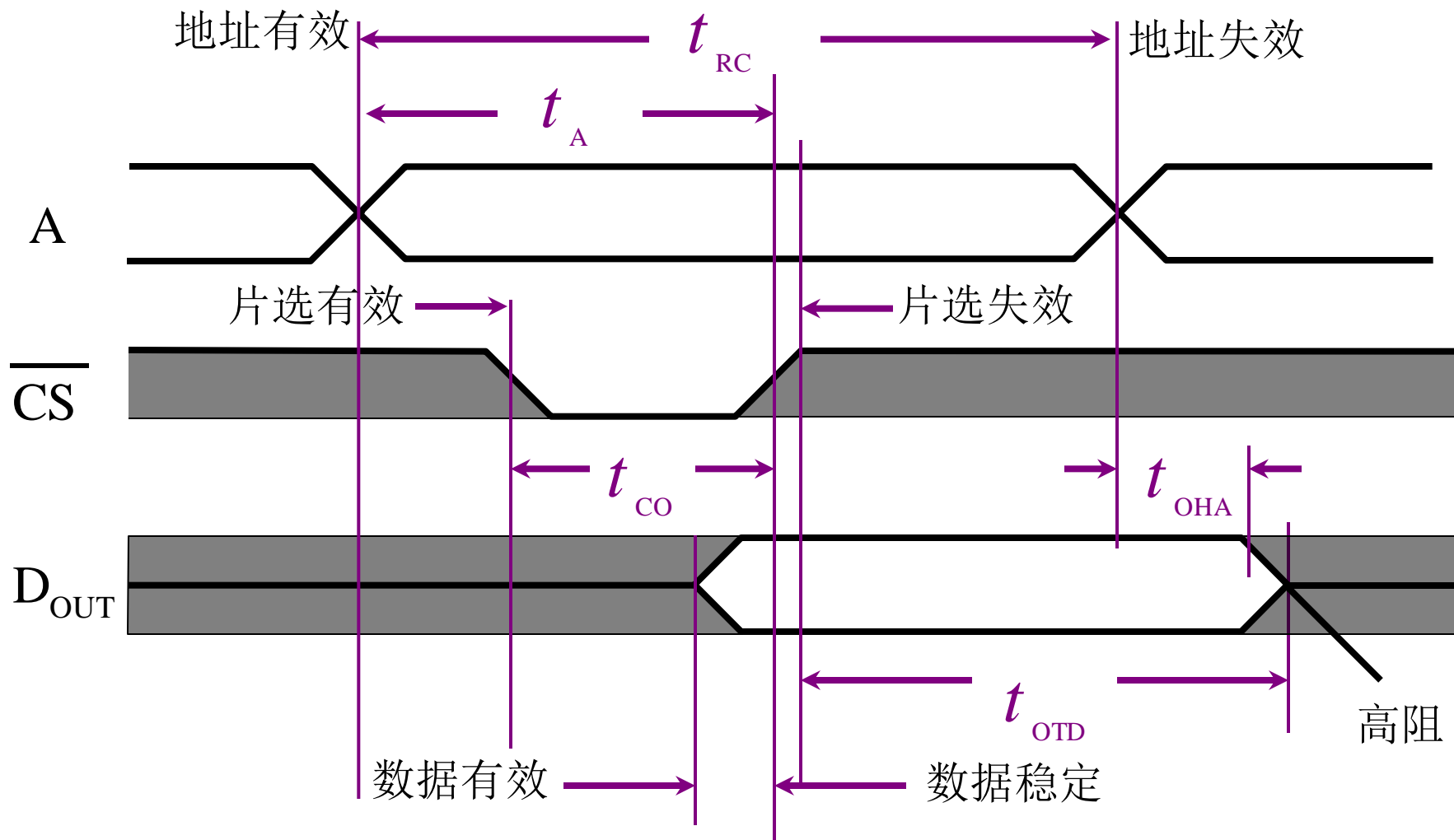


③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2



(3) 静态 RAM 读 时序

4.2

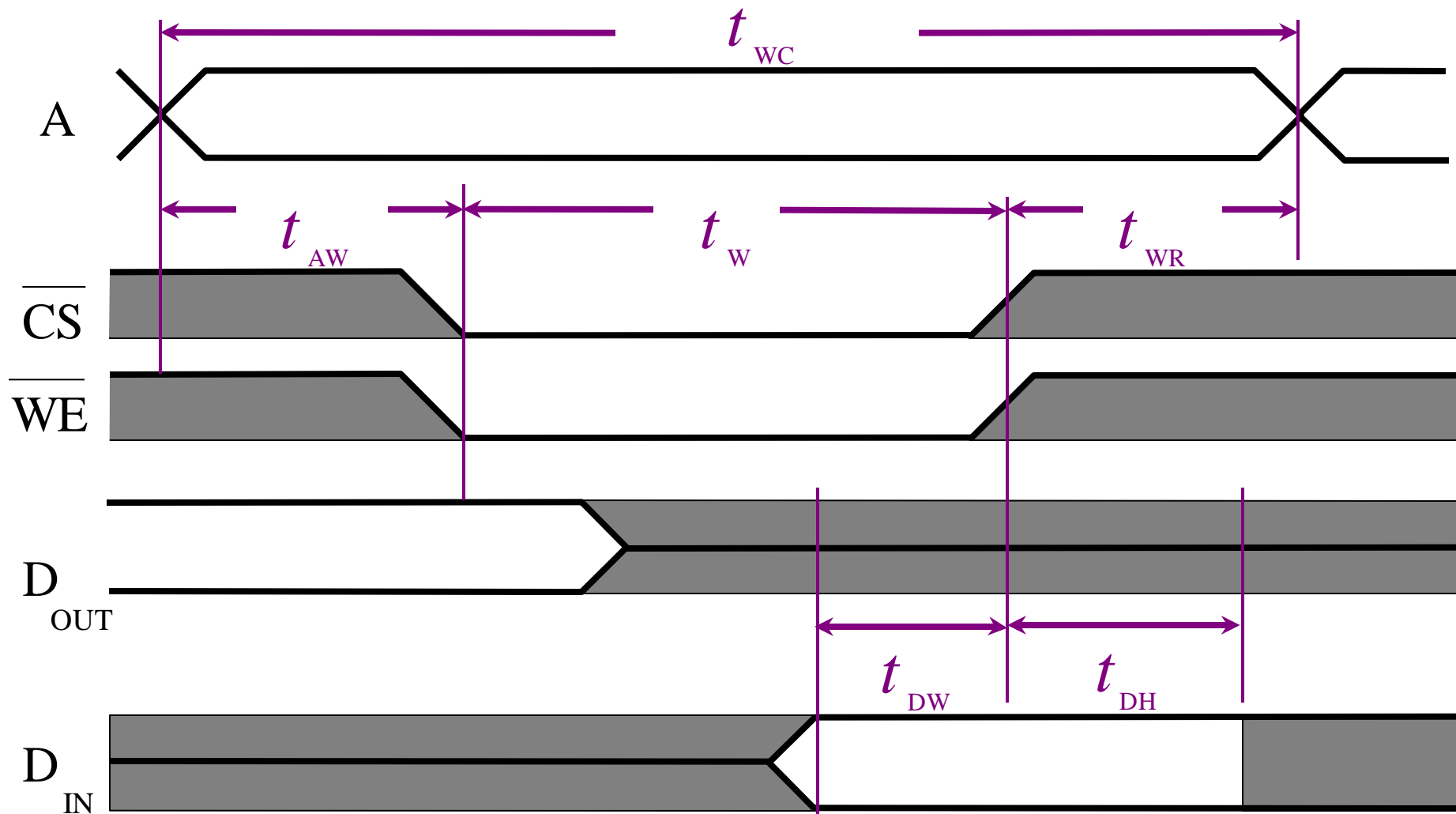


t_{OHA} 地址失效后的 数据维持时间



(4) 静态 RAM (2114) 写时序

4.2

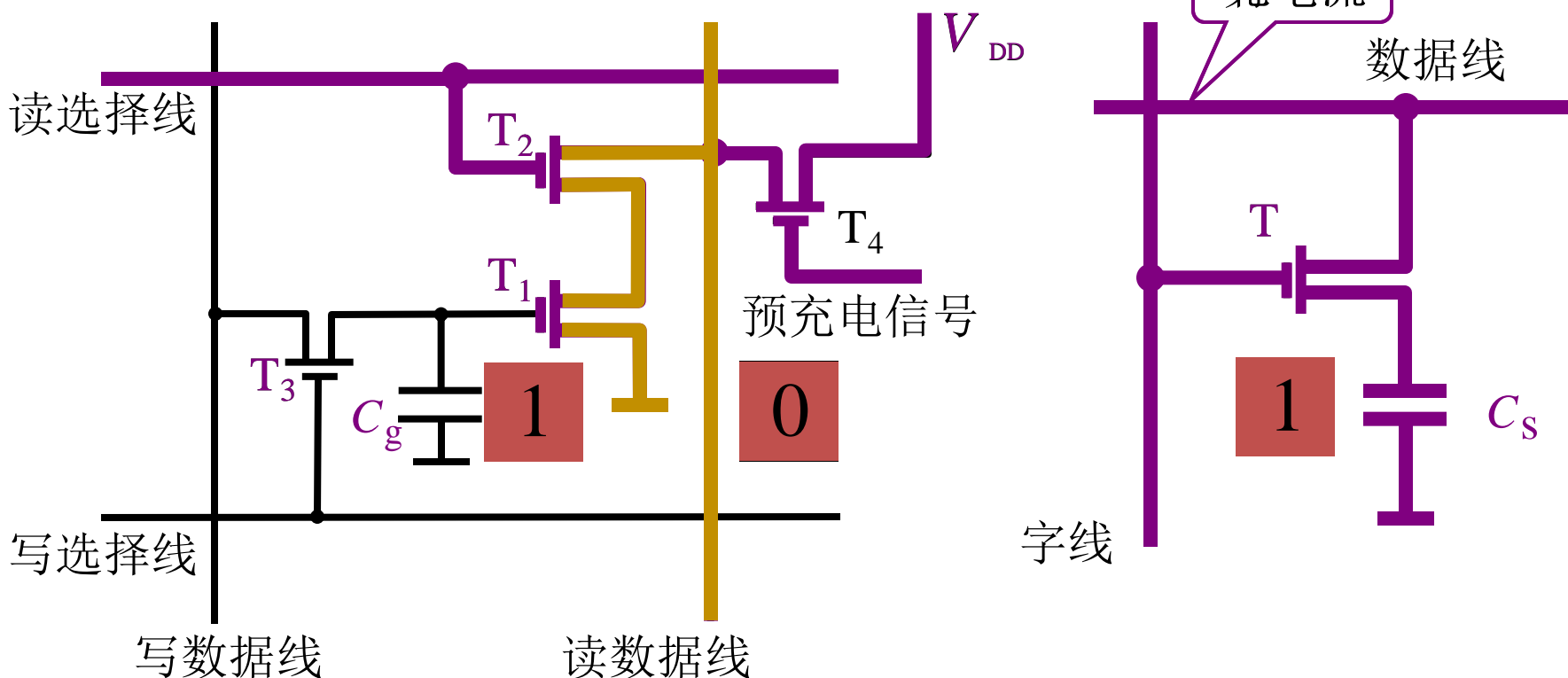


t_{DH} $\overline{\text{WE}}$ 失效后的数据维持时间



2. 动态 RAM (DRAM)

(1) 动态 RAM 基本单元电路



读出与原存信息相反

写入与输入信息相同

读出时数据线有电流为 “1”

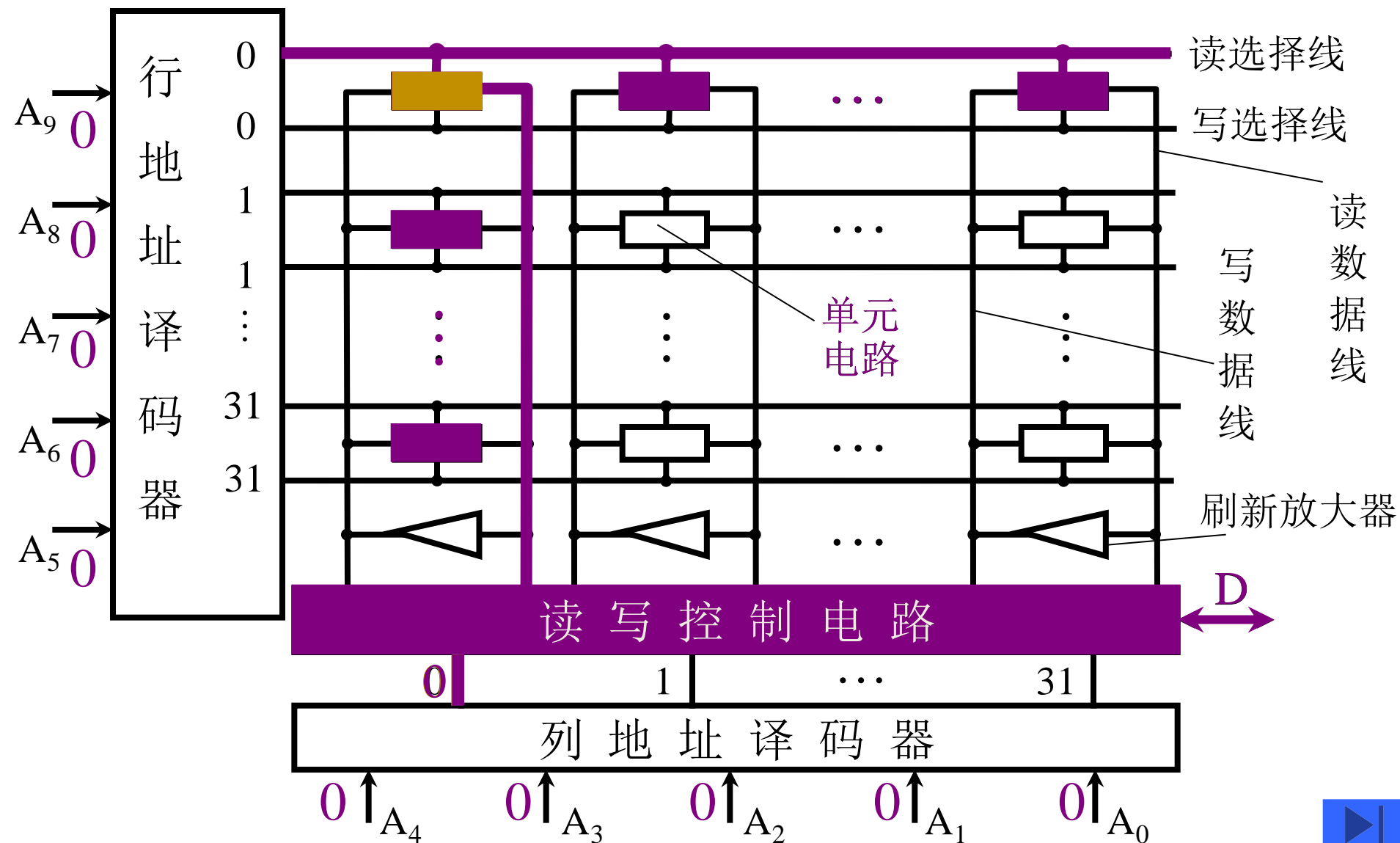
写入时 C_s 充电为 “1” 放电为 “0”



(2) 动态 RAM 芯片举例

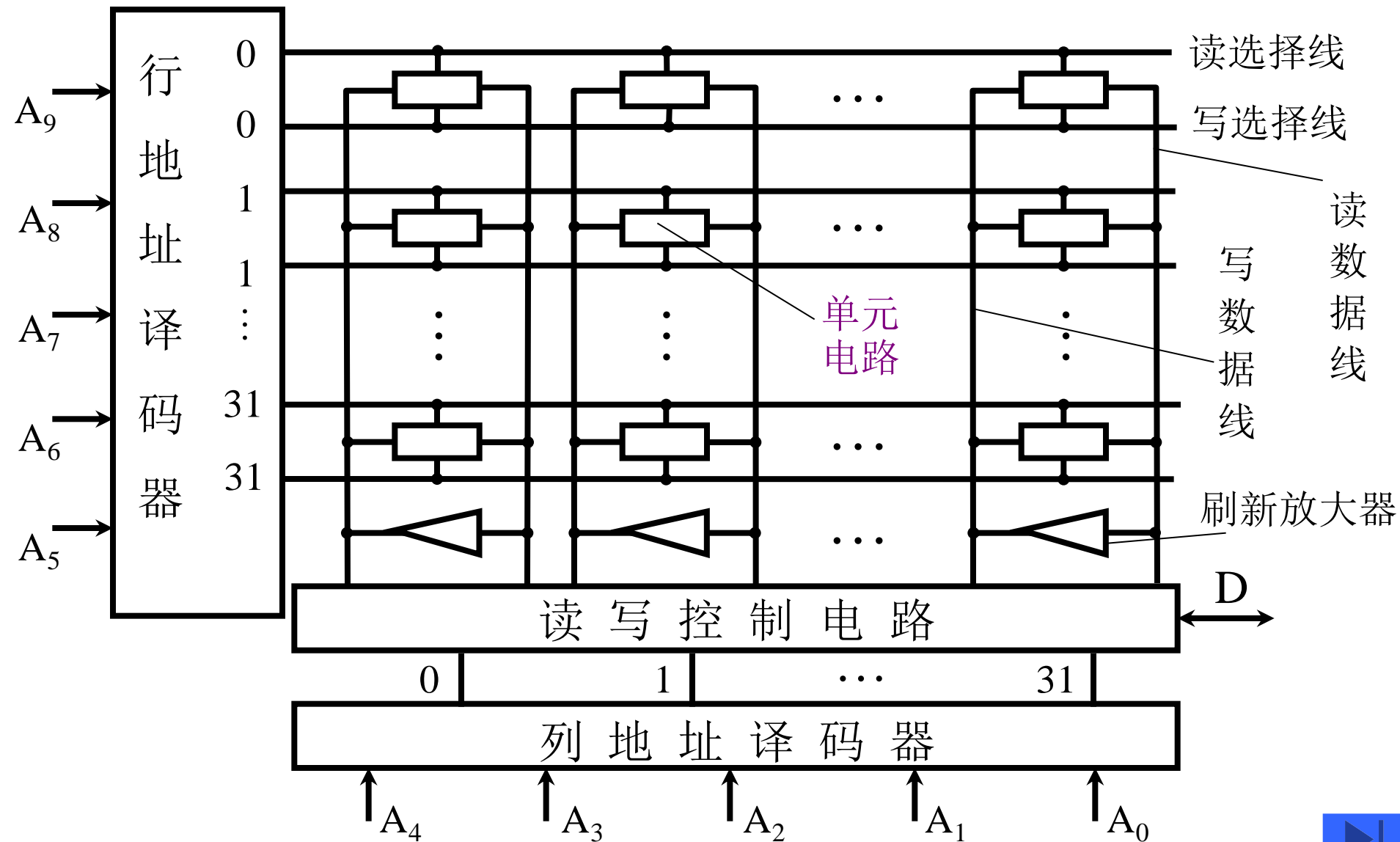
4.2

① 三管动态 RAM 芯片 (Intel 1103) 读



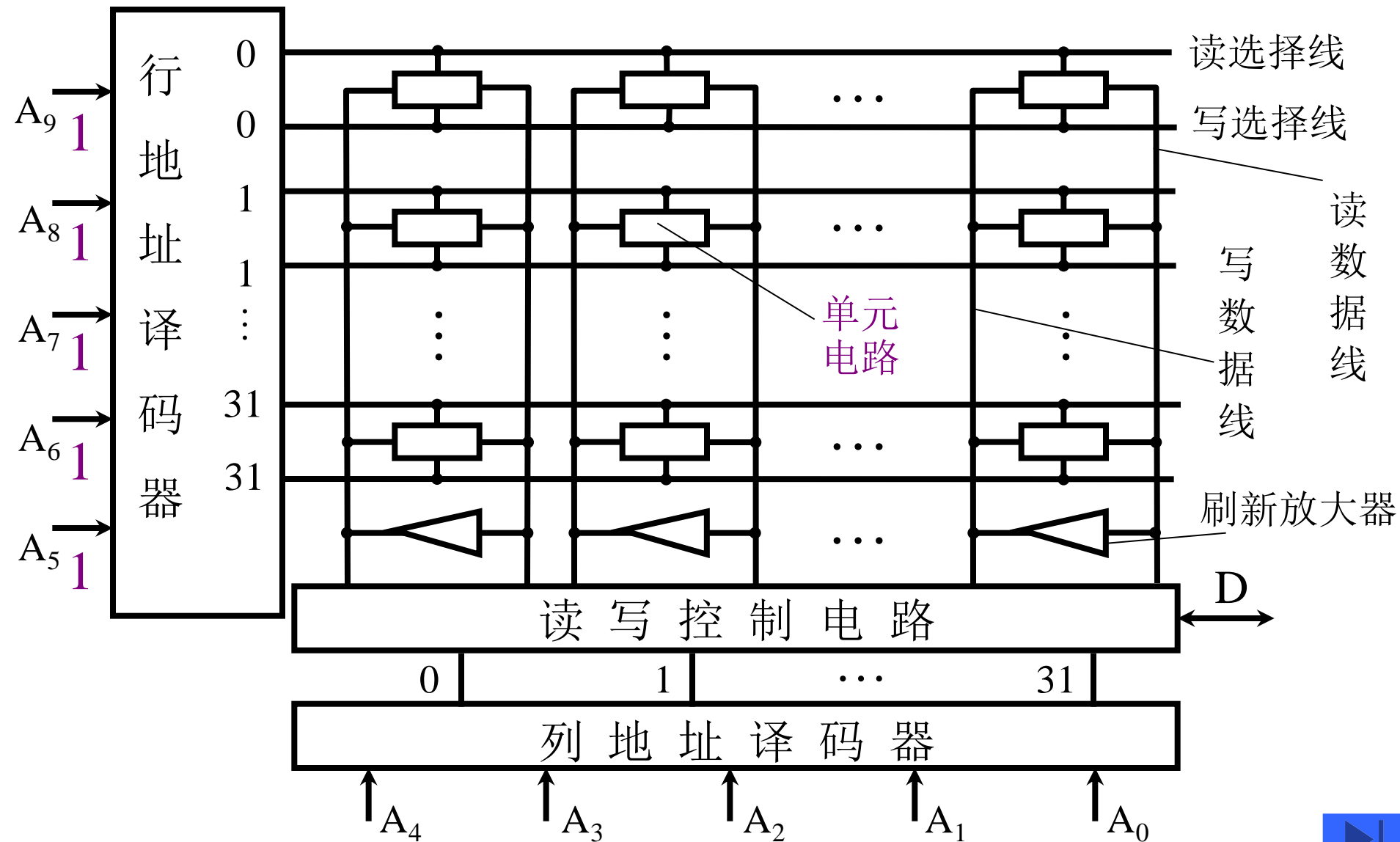
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2



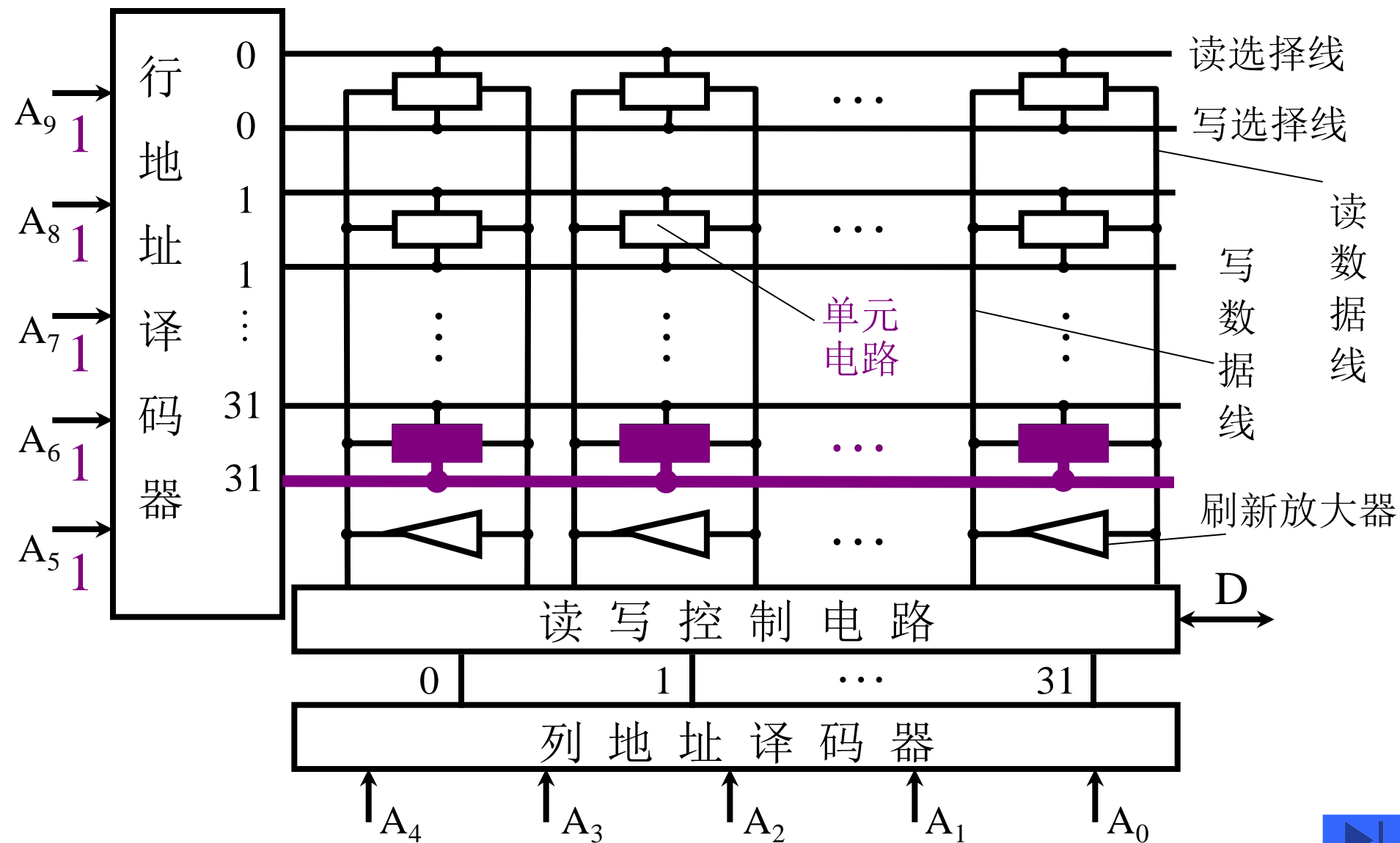
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2



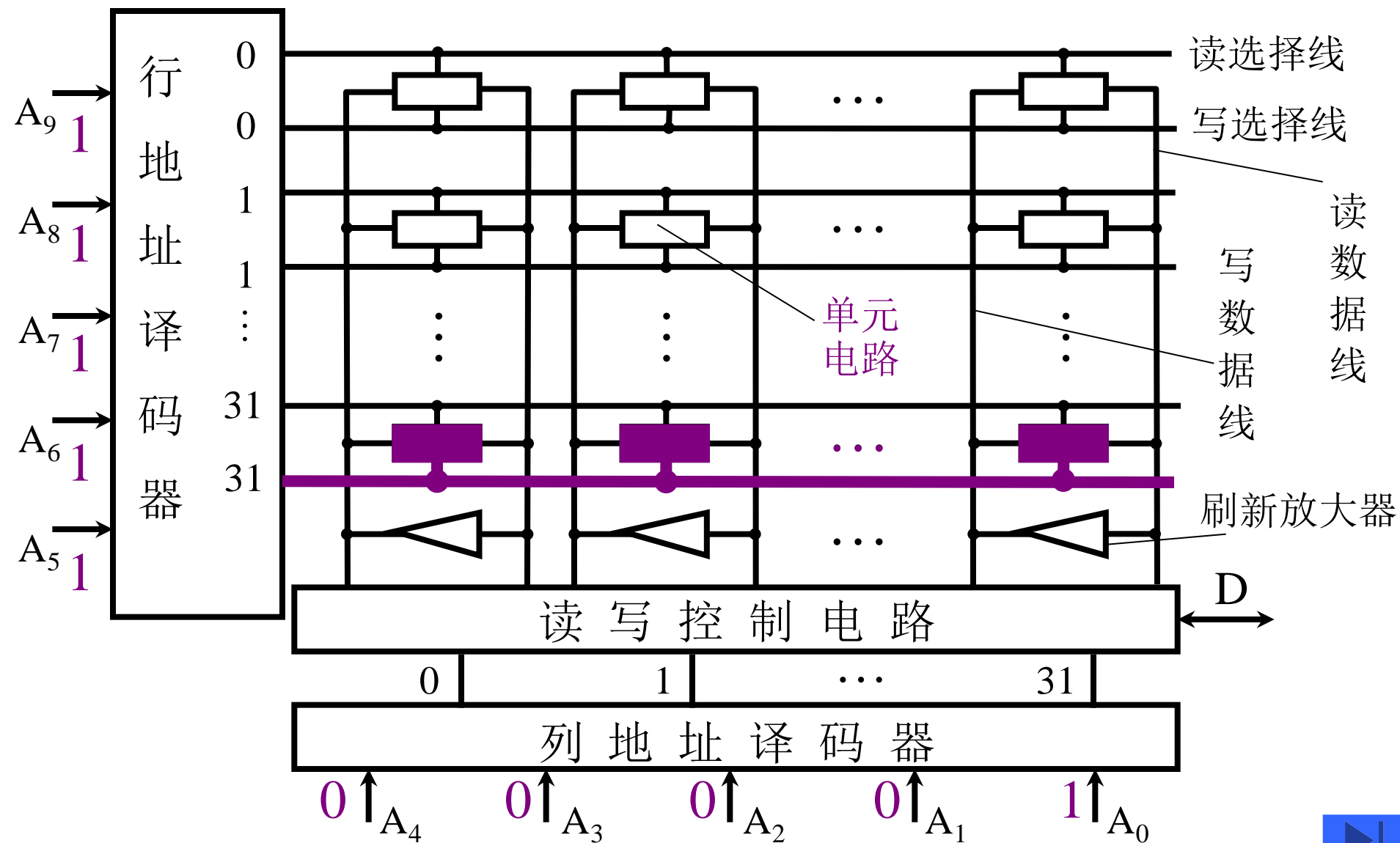
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2



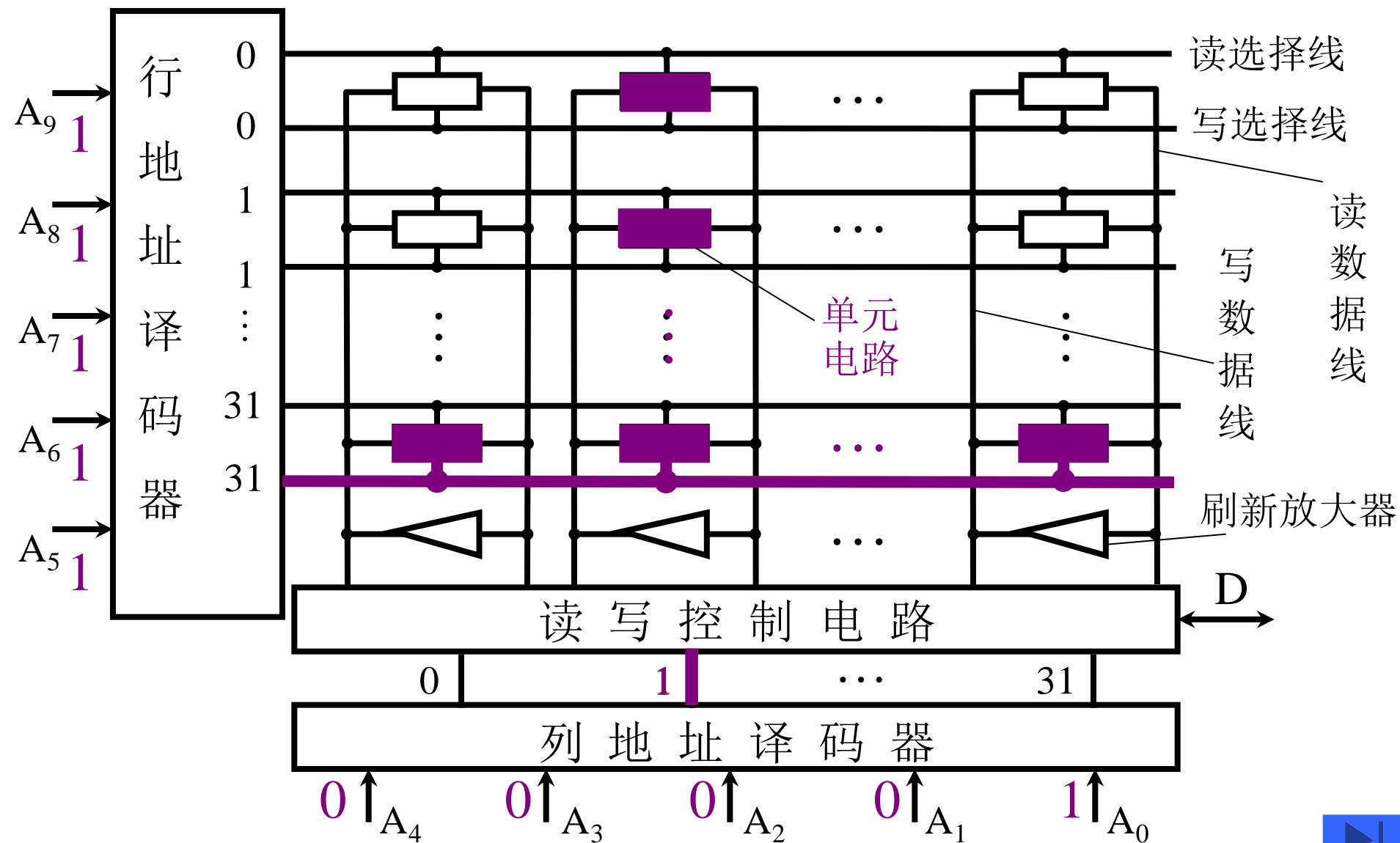
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2



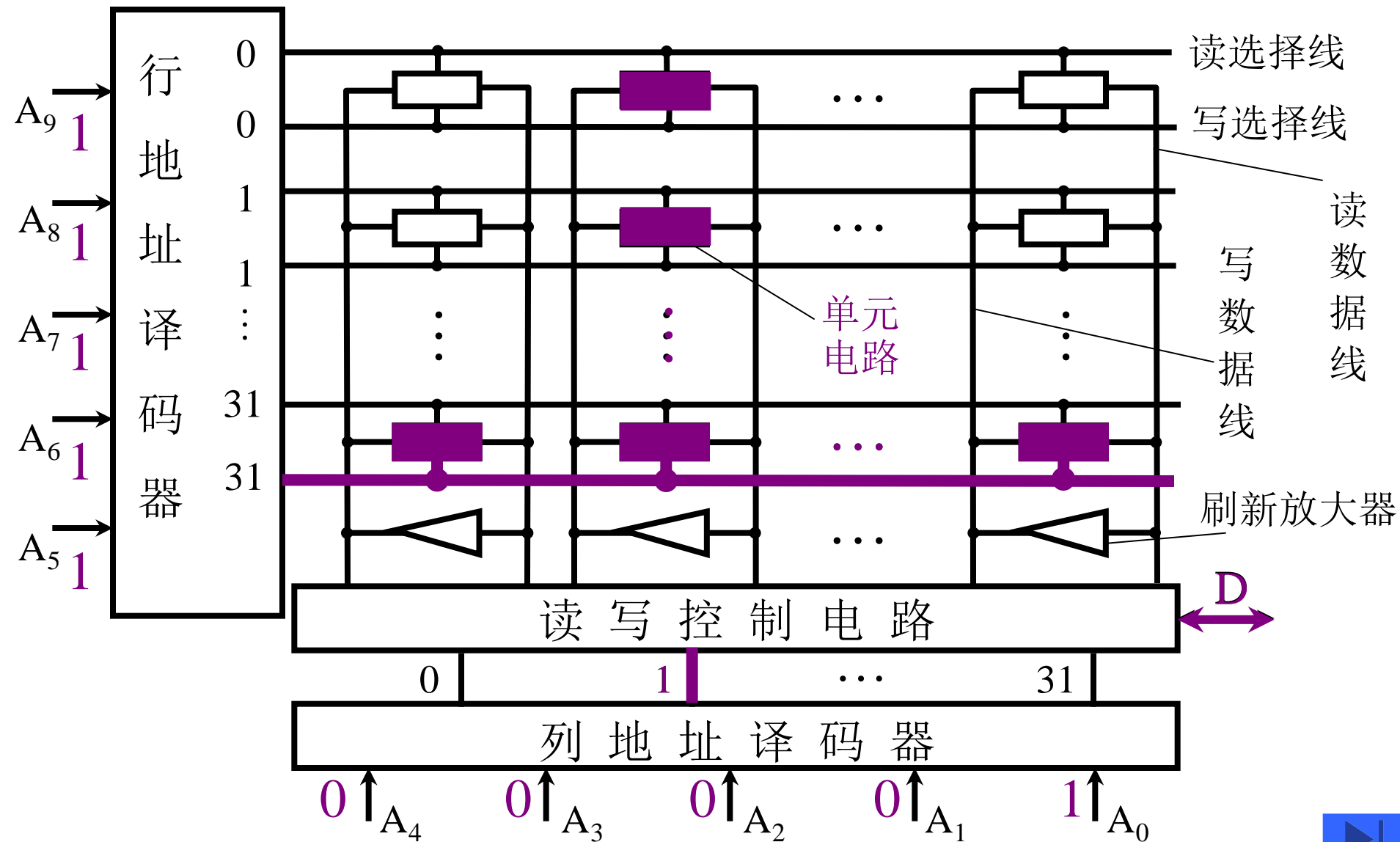
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2



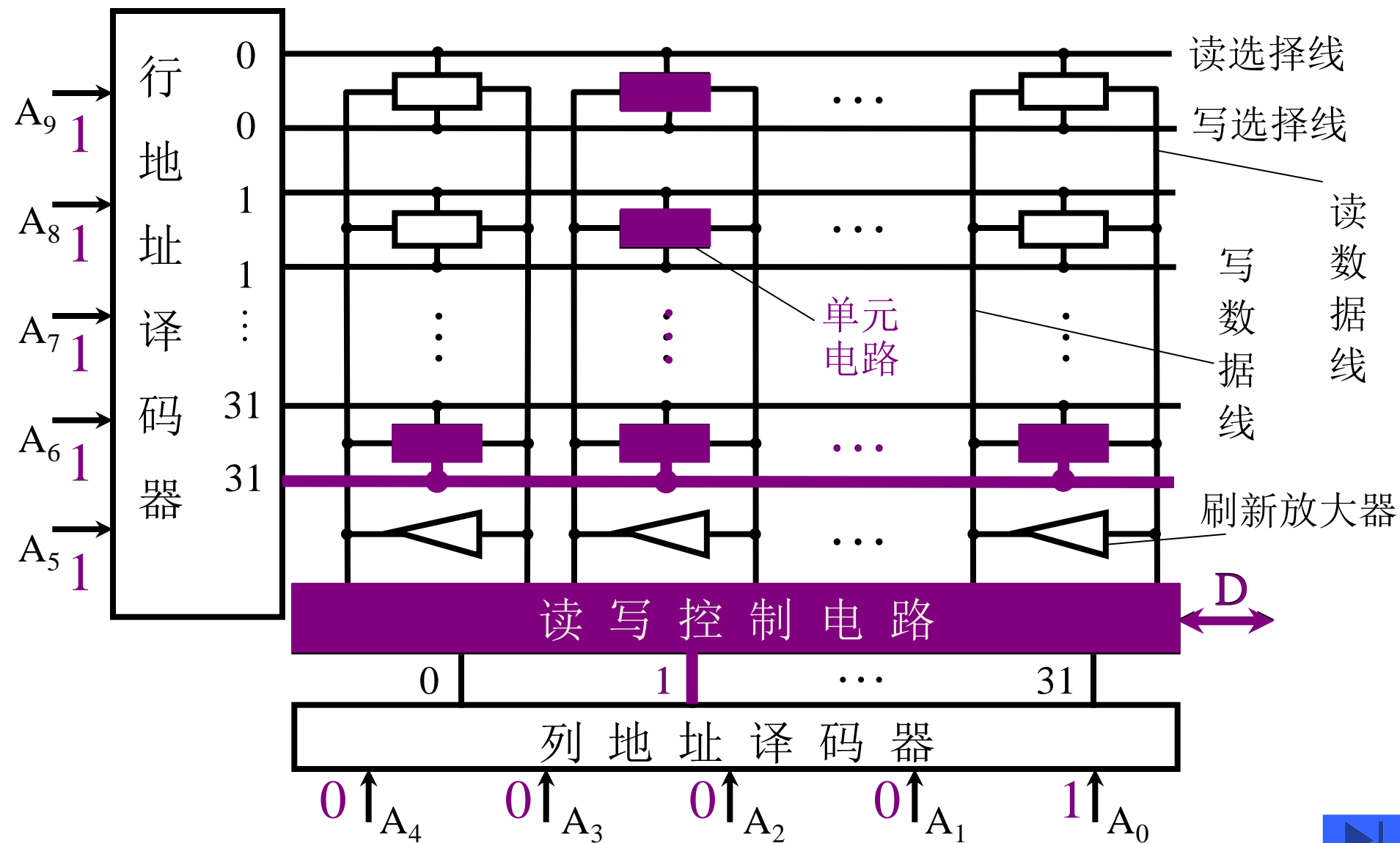
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2



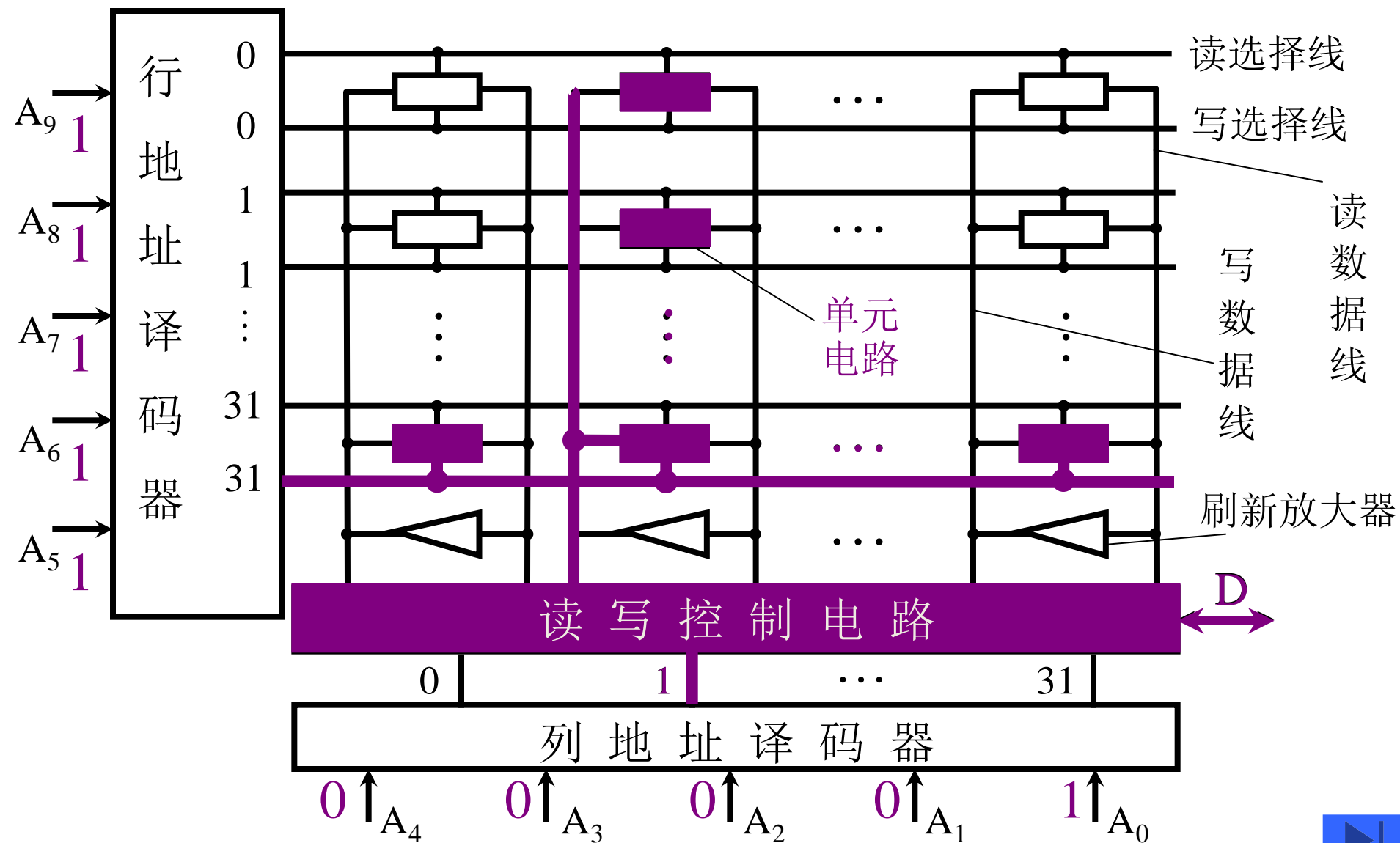
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2



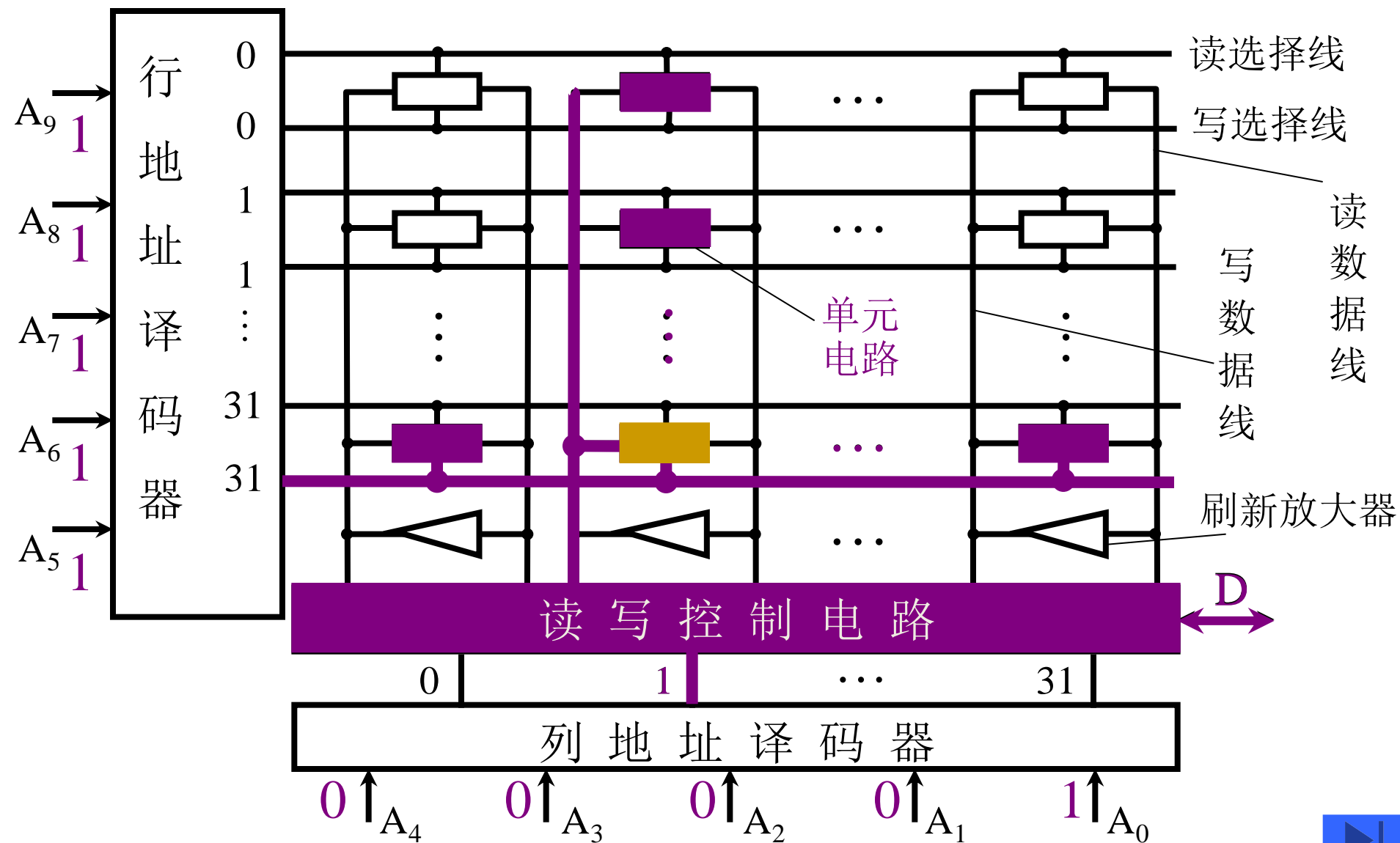
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2

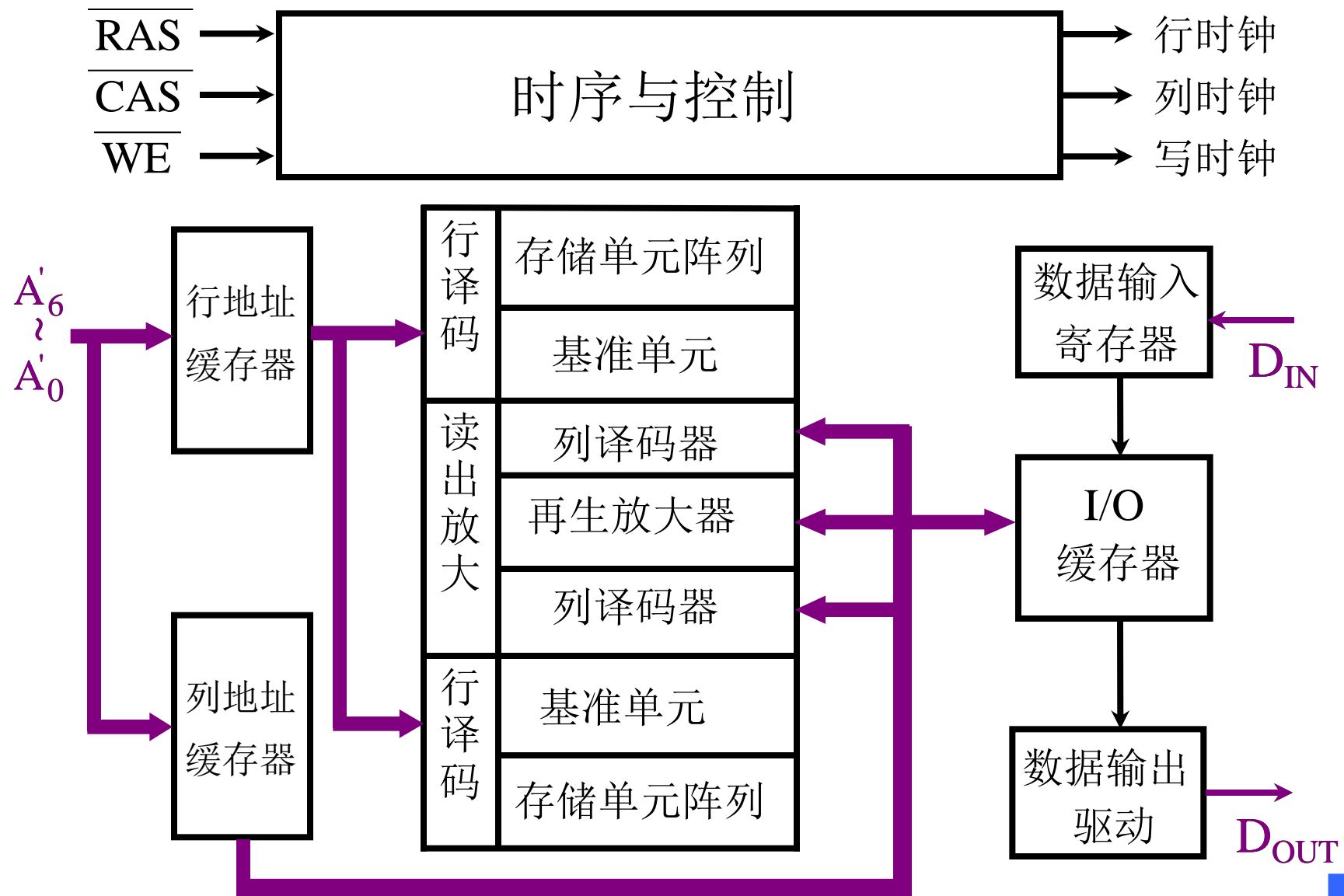


② 三管动态 RAM 芯片 (Intel 1103) 写

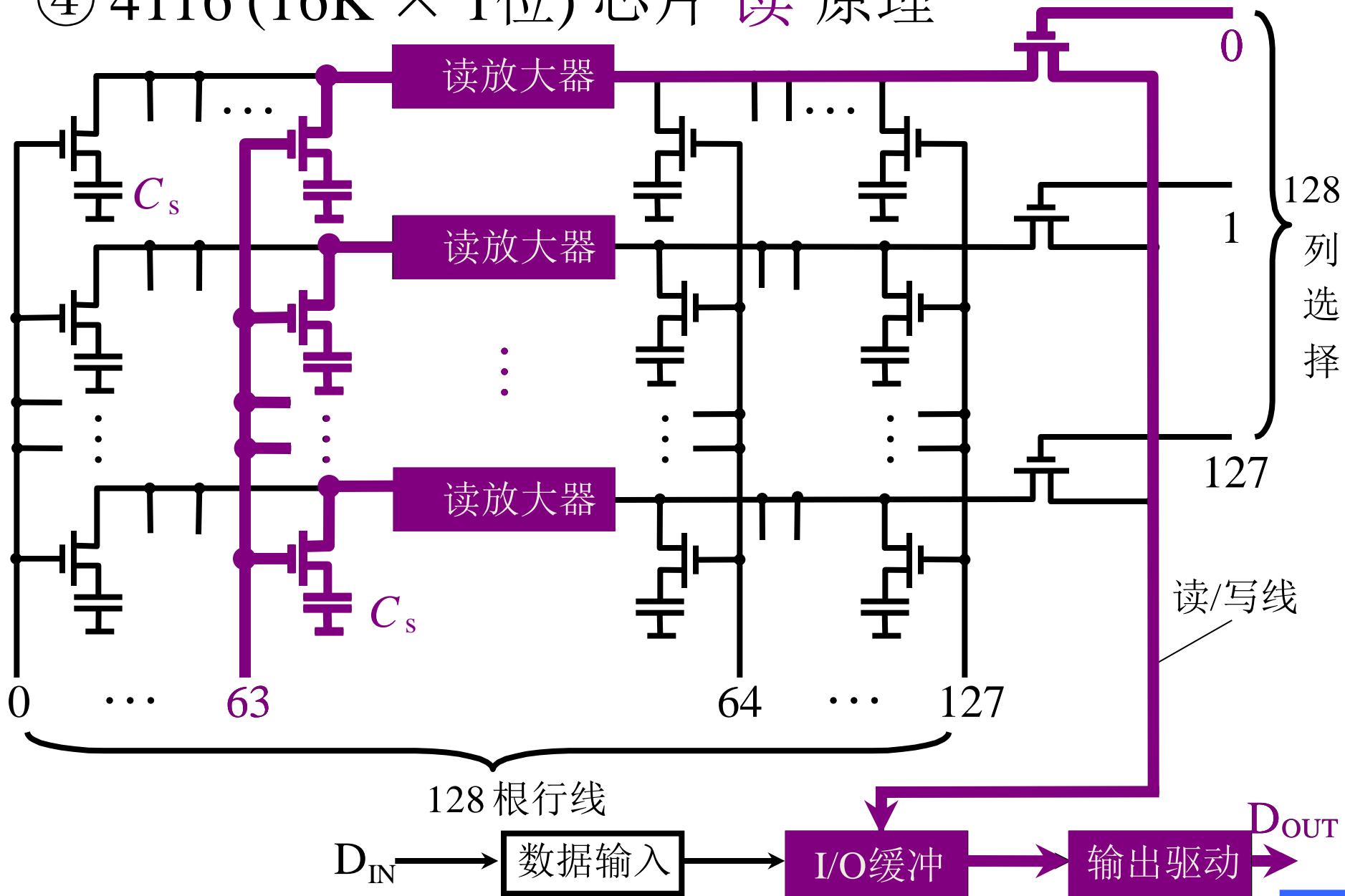
4.2



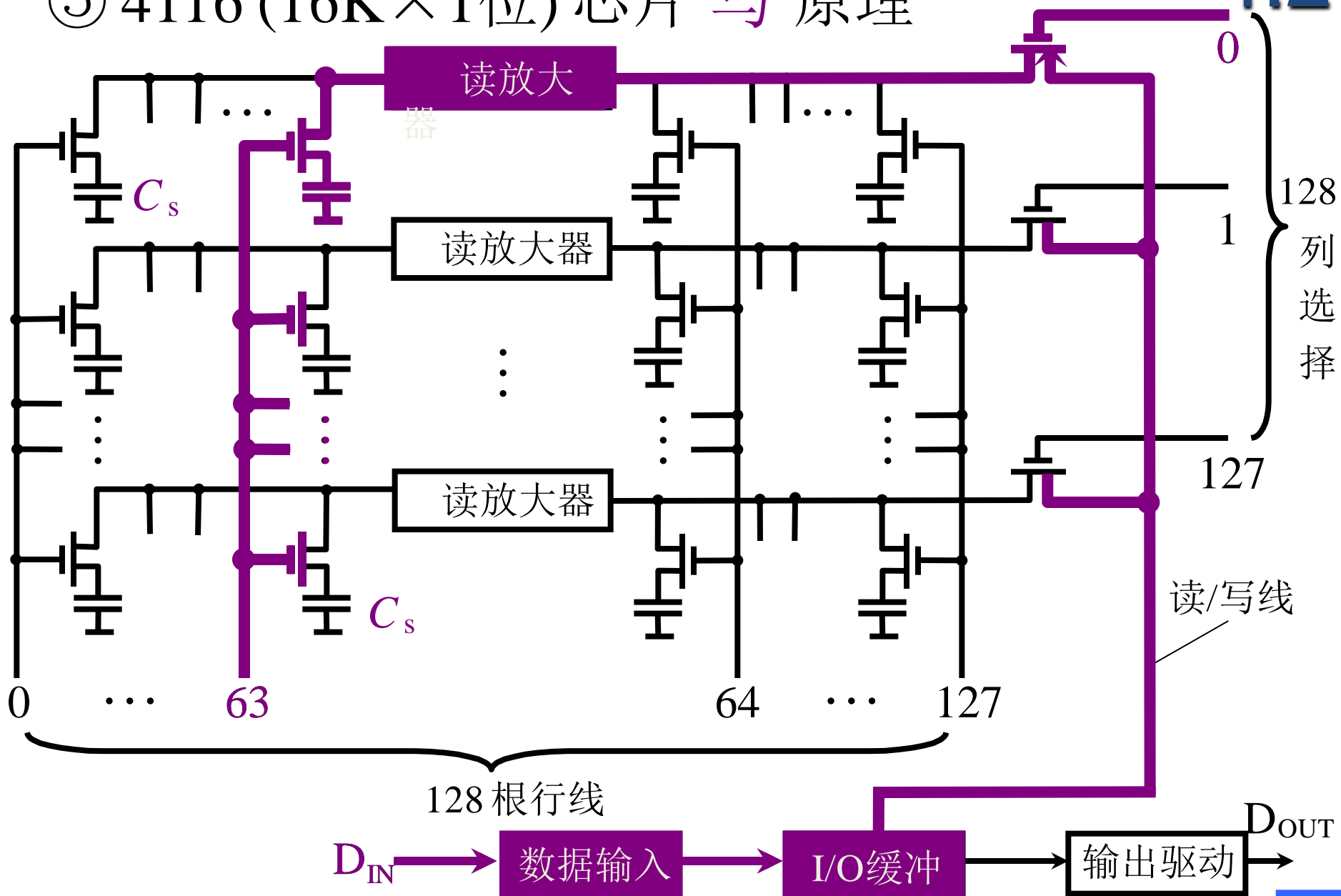
③ 单管动态 RAM 4116 (16K × 1位) 外特性 4.2



④ 4116 (16K × 1位) 芯片 读 原理



⑤ 4116 (16K×1位) 芯片 写 原理



(3) 动态 RAM 时序

行、列地址分开传送

读时序

行地址 $\overline{\text{RAS}}$ 有效

写允许 $\overline{\text{WE}}$ 有效(高)

列地址 $\overline{\text{CAS}}$ 有效

数据 D_{OUT} 有效

写时序

行地址 $\overline{\text{RAS}}$ 有效

写允许 $\overline{\text{WE}}$ 有效(低)

数据 D_{IN} 有效

列地址 $\overline{\text{CAS}}$ 有效

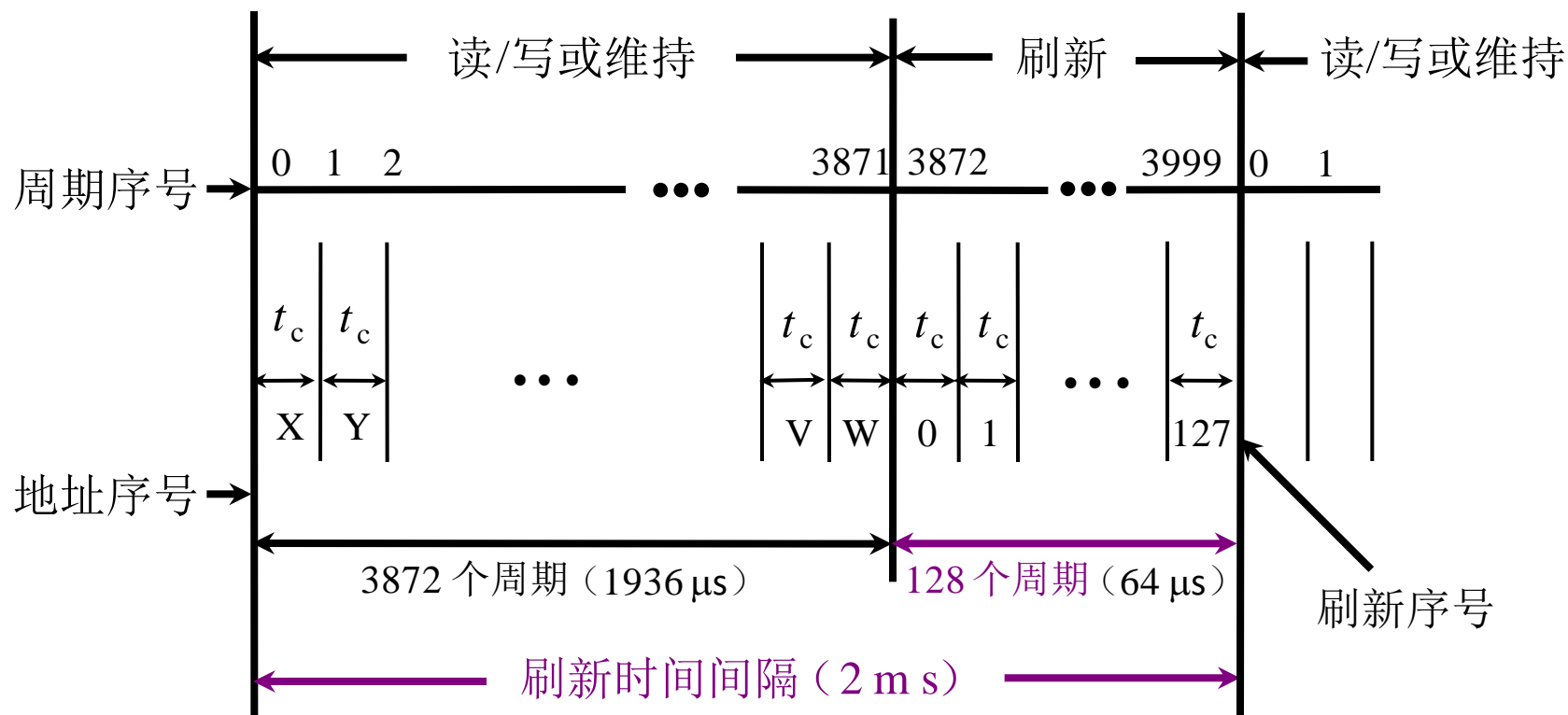


(4) 动态 RAM 刷新

4.2

刷新与行地址有关

① 集中刷新 （存取周期为 $0.5\ \mu\text{s}$ ）以 128×128 矩阵为例



“死区” 为 $0.5\ \mu\text{s} \times 128 = 64\ \mu\text{s}$

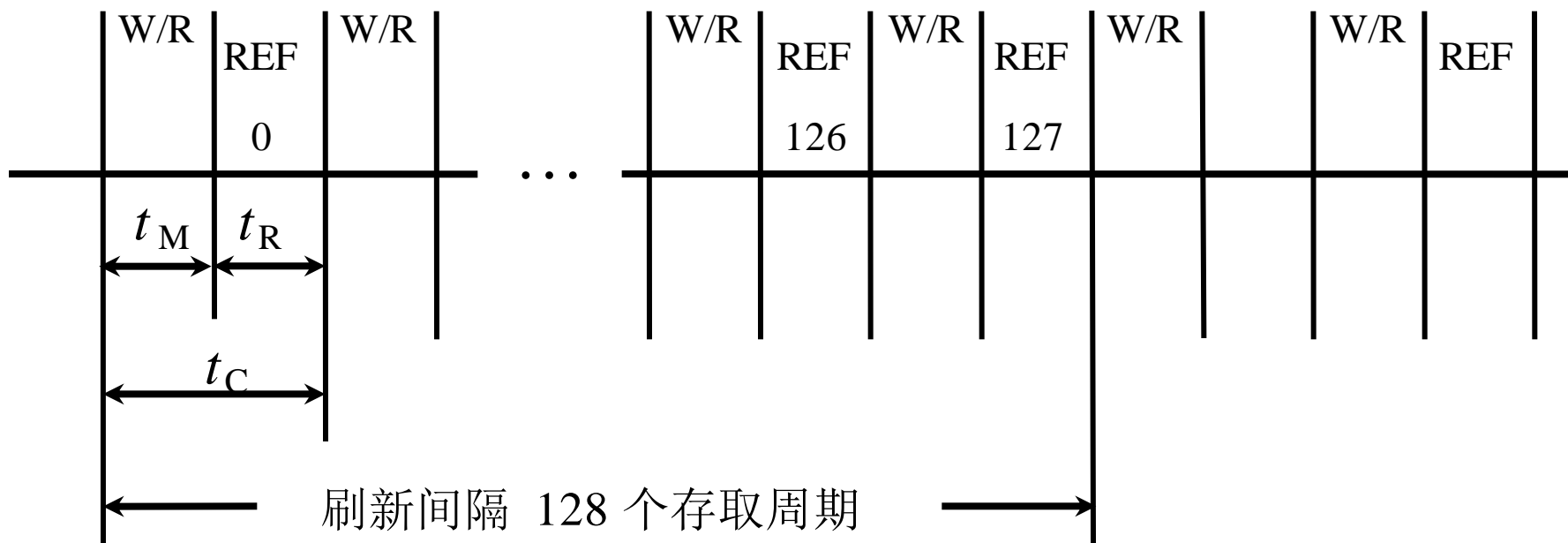
“死时间率” 为 $128/4\ 000 \times 100\% = 3.2\%$



② 分散刷新（存取周期为 $1\mu\text{s}$ ）

4.2

以 128×128 矩阵为例



$$t_C = t_M + t_R$$

无“死区”

↓ ↓
读写 刷新

(存取周期为 $0.5\mu\text{s} + 0.5\mu\text{s}$)



③ 分散刷新与集中刷新相结合（异步刷新）^{4.2}

对于 128×128 的存储芯片（存取周期为 $0.5 \mu\text{s}$ ）

若每隔 $15.6 \mu\text{s}$ 刷新一行

每行每隔 2 ms 刷新一次

“死区”为 $0.5 \mu\text{s}$

将刷新安排在指令译码阶段，不会出现“死区”



3. 动态 RAM 和静态 RAM 的比较

	主存 DRAM	SRAM 缓存
存储原理	电容	触发器
集成度	高	低
芯片引脚	少	多
功耗	小	大
价格	低	高
速度	慢	快
刷新	有	无



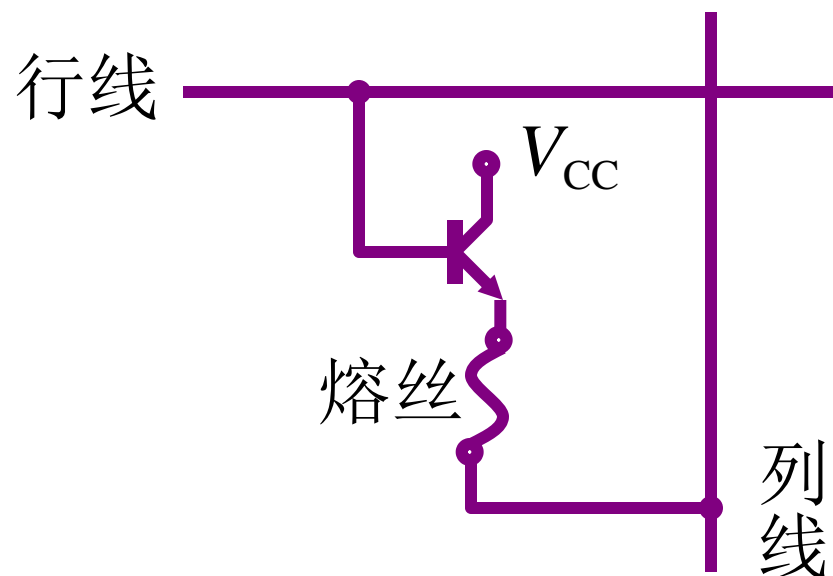
四、只读存储器（ROM）

1. 掩模 ROM (MROM)

行列选择线交叉处有 MOS 管为 “1”

行列选择线交叉处无 MOS 管为 “0”

2. PROM (一次性编程)



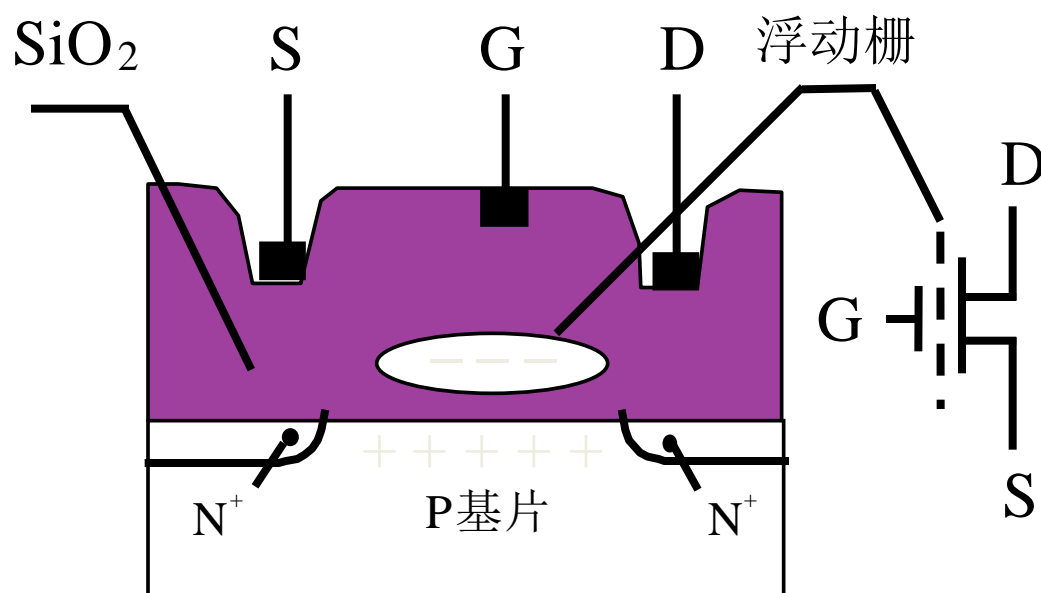
熔丝断 为 “0”

熔丝未断 为 “1”



3. EPROM (多次性编程)

(1) N型沟道浮动栅 MOS 电路



G 栅极

S 源

D 漏

紫外线全部擦洗

D 端加正电压

形成浮动栅

S 与 D 不导通为 “0”

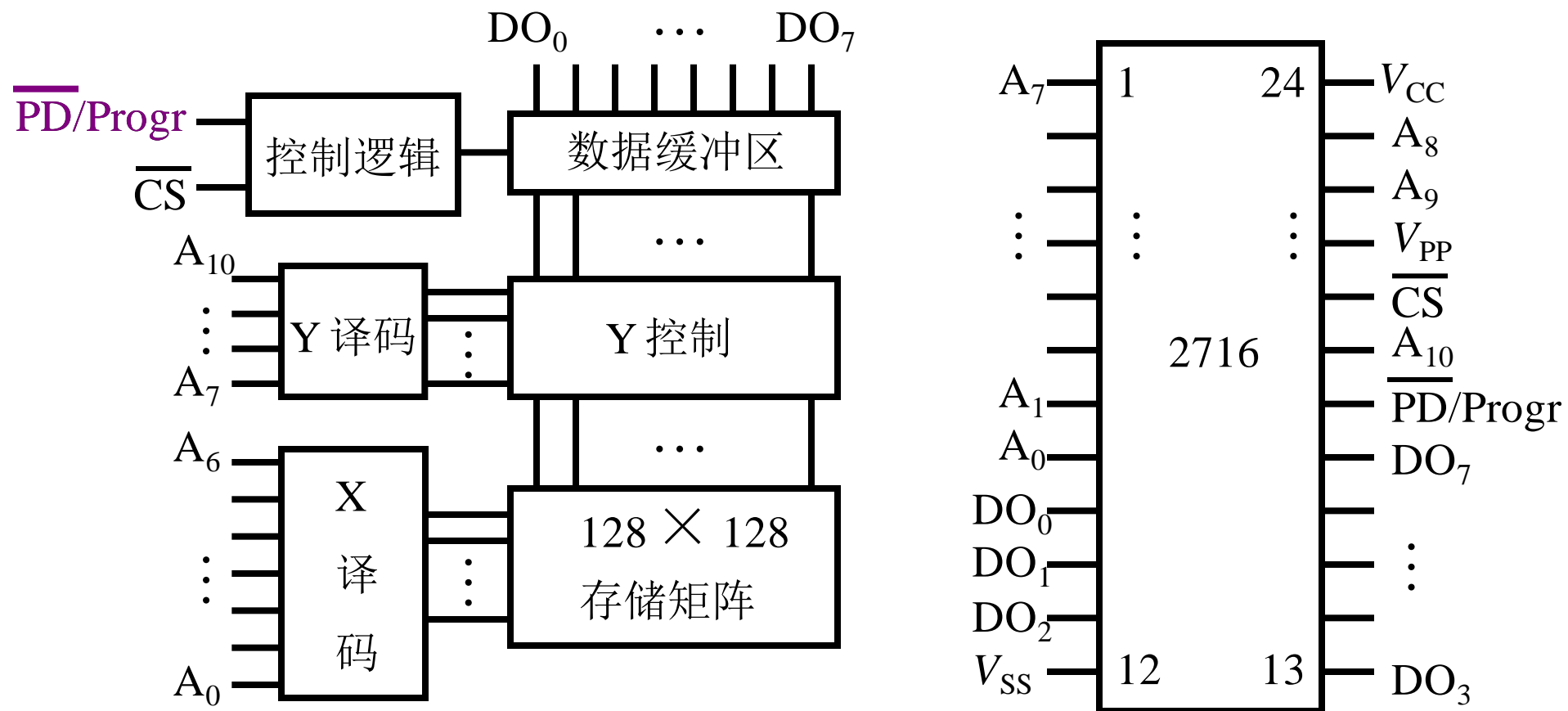
D 端不加正电压

不形成浮动栅

S 与 D 导通为 “1”



(2) 2716 EPROM 的逻辑图和引脚



$\overline{\text{PD/Progr}}$ 功率下降 / 编程输入端 读出时为低电平



4. EEPROM (多次性编程)

电可擦写

局部擦写

全部擦写

5. Flash Memory (闪速型存储器)

EPROM 价格便宜 集成度高

EEPROM 电可擦写重写

比 EEPROM快 具备 RAM 功能



五、存储器与 CPU 的连接

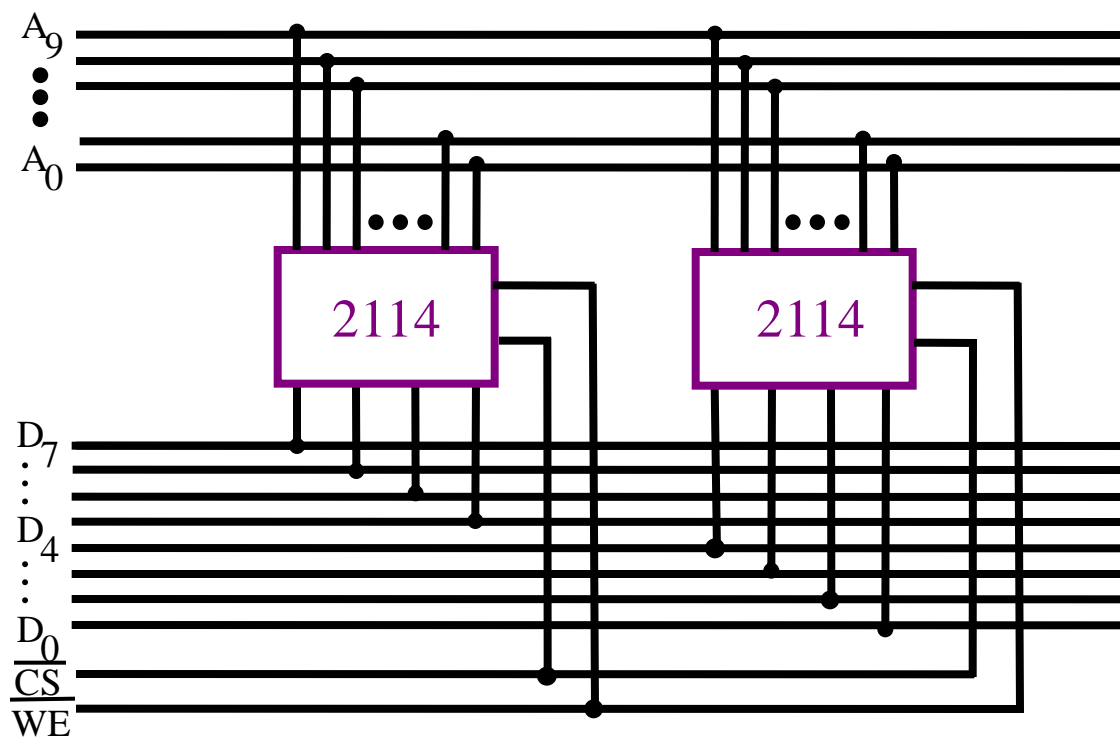
1. 存储器容量的扩展

(1) 位扩展（增加存储字长）

用 2 片 $1\text{K} \times 4$ 位 存储芯片组成 $1\text{K} \times 8$ 位的存储器

10根地址线

8根数据线

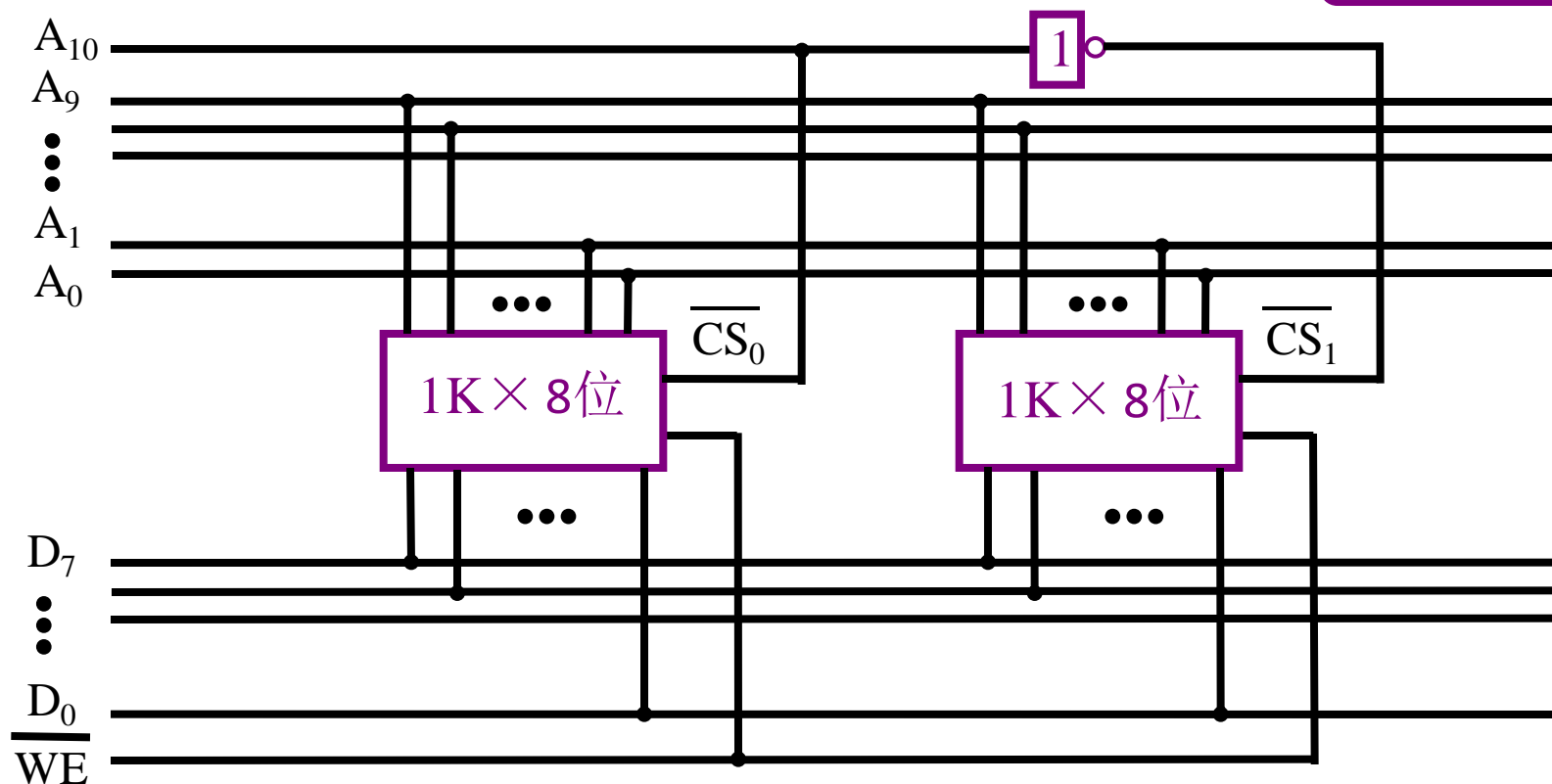


(2) 字扩展（增加存储字的数量）

11根地址线

用 2 片 $1\text{K} \times 8$ 位 存储芯片组成 $2\text{K} \times 8$ 位的存储器

8根数据线



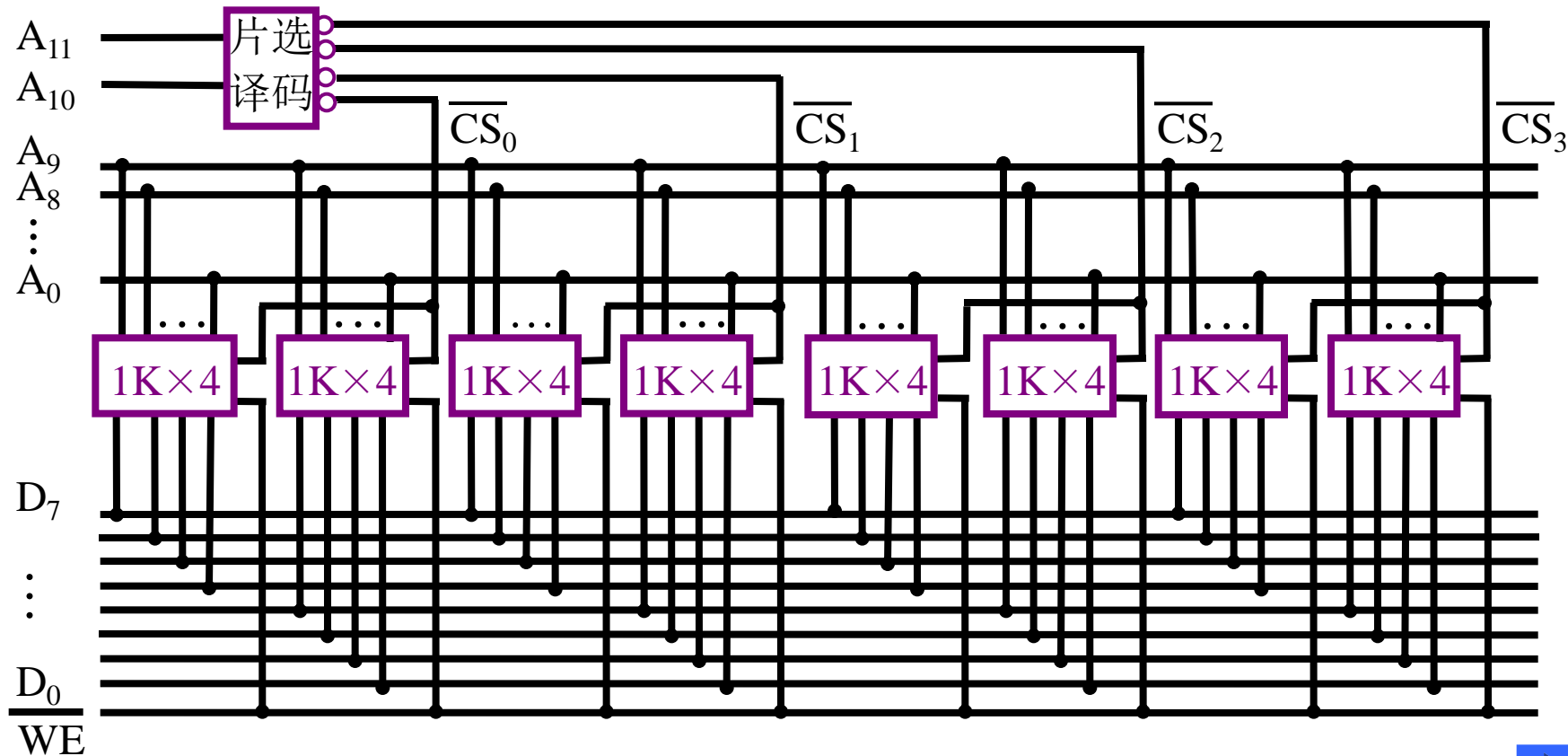
(3) 字、位扩展

4.2

用 8 片 $1\text{K} \times 4$ 位 存储芯片组成 $4\text{K} \times 8$ 位的存储器

12根地址线

8根数据线



2. 存储器与 CPU 的连接

- (1) 地址线的连接
- (2) 数据线的连接
- (3) 读/写命令线的连接
- (4) 片选线的连接
- (5) 合理选择存储芯片
- (6) 其他 时序、负载



例4.1 解:

4.2

(1) 写出对应的二进制地址码

A_{15}	A_{14}	A_{13}	A_{11}	A_{10}	...	A_7	...	A_4	A_3	...	A_0	
0	1	1	0	0	0	0	0	0	0	0	0	} ROM 2K×8位
⋮												
0	1	1	0	0	1	1	1	1	1	1	1	
0	1	1	0	1	0	0	0	0	0	0	0	} 1K×8位 RAM
⋮												
0	1	1	0	1	0	1	1	1	1	1	1	

1片 2K×8位

ROM

2K×8位

1K×8位

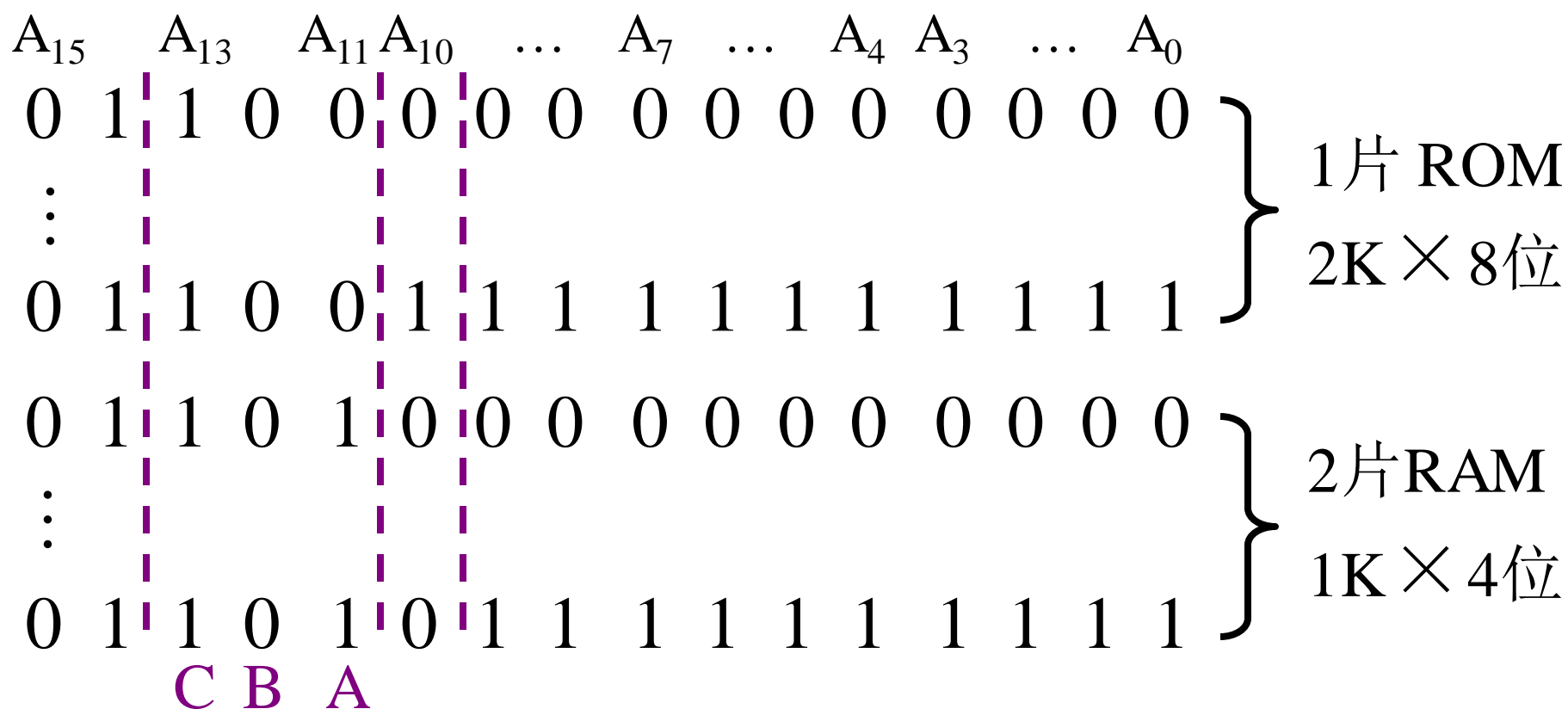
RAM

2片 1K×4位

(2) 确定芯片的数量及类型



(3) 分配地址线



$A_{10} \sim A_0$ 接 $2K \times 8$ 位 ROM 的地址线

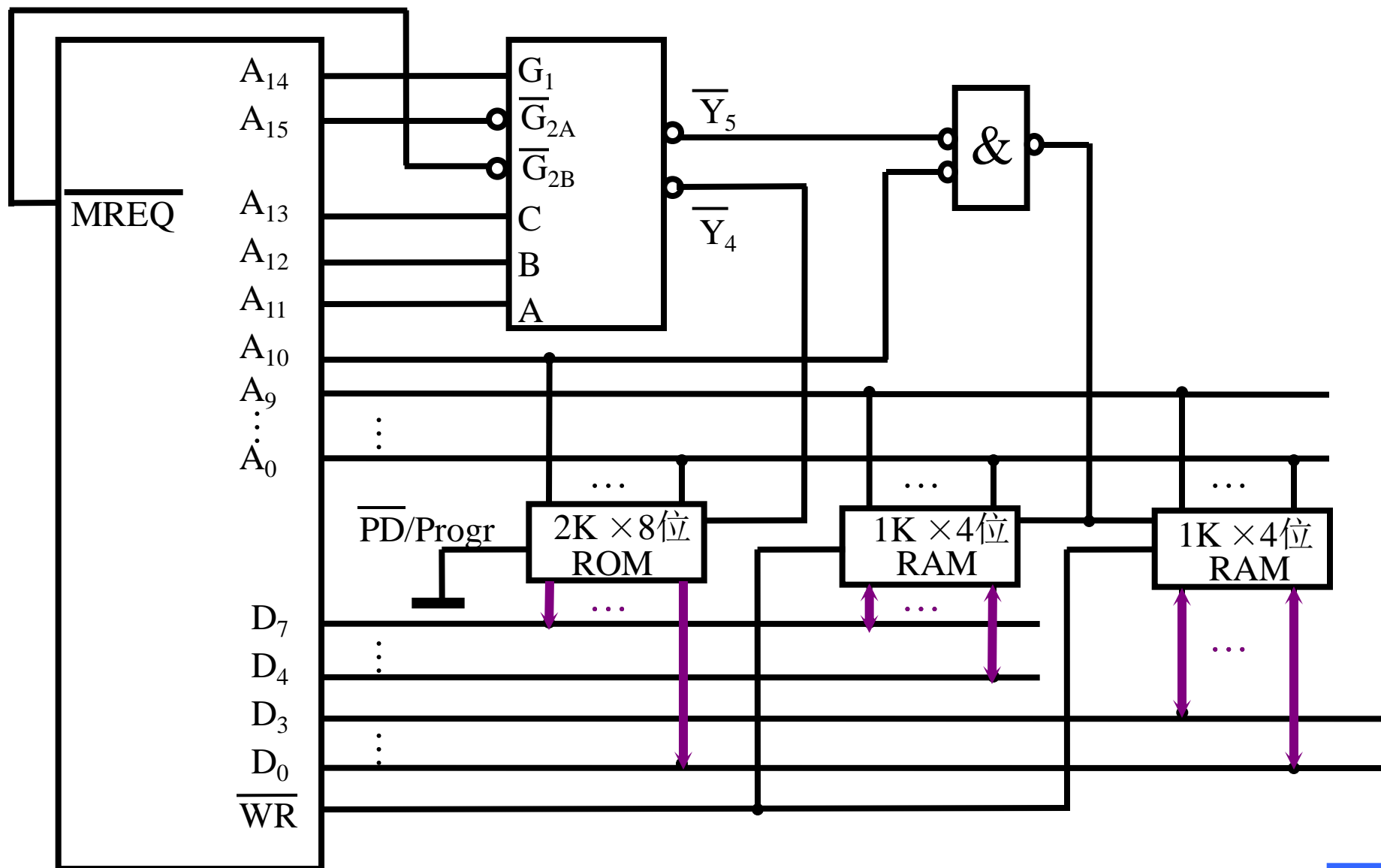
$A_9 \sim A_0$ 接 $1K \times 4$ 位 RAM 的地址线

(4) 确定片选信号



例 4.1 CPU 与存储器的连接图

4.2



例4.2 假设同前，要求最小 4K 为系统程序区，相邻 8K 为用户程序区。

(1) 写出对应的二进制地址码

(2) 确定芯片的数量及类型

1片 4K × 8位 ROM 2片 4K × 8位 RAM

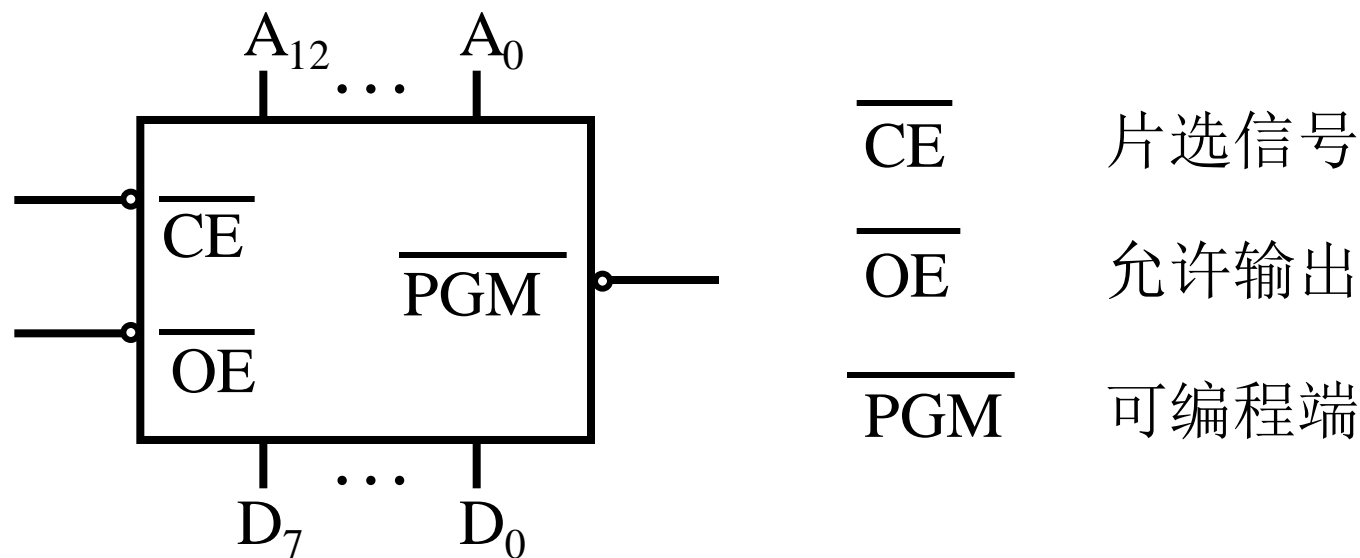
(3) 分配地址线

$A_{11} \sim A_0$ 接 ROM 和 RAM 的地址线

(4) 确定片选信号



例 4.3 设 CPU 有 20 根地址线, 8 根数据线。
并用 $\overline{\text{IO/M}}$ 作访存控制信号。 $\overline{\text{RD}}$ 为读命令,
 $\overline{\text{WR}}$ 为写命令。现有 2764 EPROM (8K × 8位),
外特性如下:



用 138 译码器及其他门电路 (门电路自定) 画出 CPU 和 2764 的连接图。要求地址为 F0000H~FFFFFFH, 并写出每片 2764 的地址范围。

六、存储器的校验

4.2

1. 编码的最小距离

任意两组合法代码之间 二进制位数 的 最少差异
编码的纠错、检错能力与编码的最小距离有关

$$L - 1 = D + C \quad (D \geq C)$$

L —— 编码的最小距离 $L = 3$

D —— 检测错误的位数 具有 一位 纠错能力

C —— 纠正错误的位数

汉明码是具有一位纠错能力的编码



2. 汉明码的组成

组成汉明码的三要素

汉明码的组成需增添 ? 位检测位

$$2^k \geq n + k + 1$$

检测位的位置 ?

$$2^i \ (i = 0, 1, 2, 3, \dots)$$

检测位的取值 ?

检测位的取值与该位所在的检测“小组”中承担的奇偶校验任务有关



各检测位 C_i 所承担的检测小组为

C_1 检测的 g_1 小组包含第 1, 3, 5, 7, 9, 11, ...

C_2 检测的 g_2 小组包含第 2, 3, 6, 7, 10, 11, ...

C_4 检测的 g_3 小组包含第 4, 5, 6, 7, 12, 13, ...

C_8 检测的 g_4 小组包含第 8, 9, 10, 11, 12, 13, 14, 15, 24, ...

g_i 小组独占第 2^{i-1} 位

g_i 和 g_j 小组共同占第 $2^{i-1} + 2^{j-1}$ 位

g_i 、 g_j 和 g_l 小组共同占第 $2^{i-1} + 2^{j-1} + 2^{l-1}$ 位



例4.4 求 0101 按 “偶校验” 配置的汉明码

解： $\because n = 4$

根据 $2^k \geq n + k + 1$

得 $k = 3$

汉明码排序如下：

二进制序号	1	2	3	4	5	6	7
名称	C_1	C_2	0	C_4	1	0	1
	0	1		0			

\therefore 0101 的汉明码为 0100101



练习1 按配偶原则配置 0011 的汉明码 4.2

解： $\because n = 4$ 根据 $2^k \geq n + k + 1$

取 $k = 3$

二进制序号	1	2	3	4	5	6	7
名称	C_1	C_2	0	C_4	0	1	1
	1	0		0			

$$C_1 = 3 \oplus 5 \oplus 7 = 1$$

$$C_2 = 3 \oplus 6 \oplus 7 = 0$$

$$C_4 = 5 \oplus 6 \oplus 7 = 0$$

\therefore 0011 的汉明码为 1000011



3. 汉明码的纠错过程

形成新的检测位 P_i ，其位数与增添的检测位有关，如增添 3 位（ $k=3$ ），新的检测位为 $P_4 P_2 P_1$ 。

以 $k=3$ 为例， P_i 的取值为

$$P_1 = \overset{C_1}{1} \oplus 3 \oplus 5 \oplus 7$$

$$P_2 = \overset{C_2}{2} \oplus 3 \oplus 6 \oplus 7$$

$$P_4 = \overset{C_4}{4} \oplus 5 \oplus 6 \oplus 7$$

对于按“偶校验”配置的汉明码
不出错时 $P_1=0, P_2=0, P_4=0$



例4.5 已知接收到的汉明码为 0100111

(按配偶原则配置) 试问要求传送的信息是什么?

解: 纠错过程如下

$$P_1 = 1 \oplus 3 \oplus 5 \oplus 7 = 0 \quad \text{无错}$$

$$P_2 = 2 \oplus \underset{\checkmark}{3} \oplus \boxed{6} \oplus \underset{\checkmark}{7} = 1 \quad \text{有错}$$

$$P_4 = 4 \oplus \underset{\checkmark}{5} \oplus \boxed{6} \oplus \underset{\checkmark}{7} = 1 \quad \text{有错}$$

$$\therefore P_4 P_2 P_1 = 110$$

第 6 位出错, 可纠正为 0100101,
故要求传送的信息为 0101。



练习2 写出按偶校验配置的汉明码

0101101 的纠错过程

$$P_4 = 4 \oplus 5 \oplus 6 \oplus 7 = 1$$

$$P_2 = 2 \oplus 3 \oplus 6 \oplus 7 = 0$$

$$P_1 = 1 \oplus 3 \oplus 5 \oplus 7 = 0$$

$\therefore P_4 P_2 P_1 = 100$ 第4位错，可不纠

练习3 按配奇原则配置0011的汉明码

配奇的汉明码为0101011



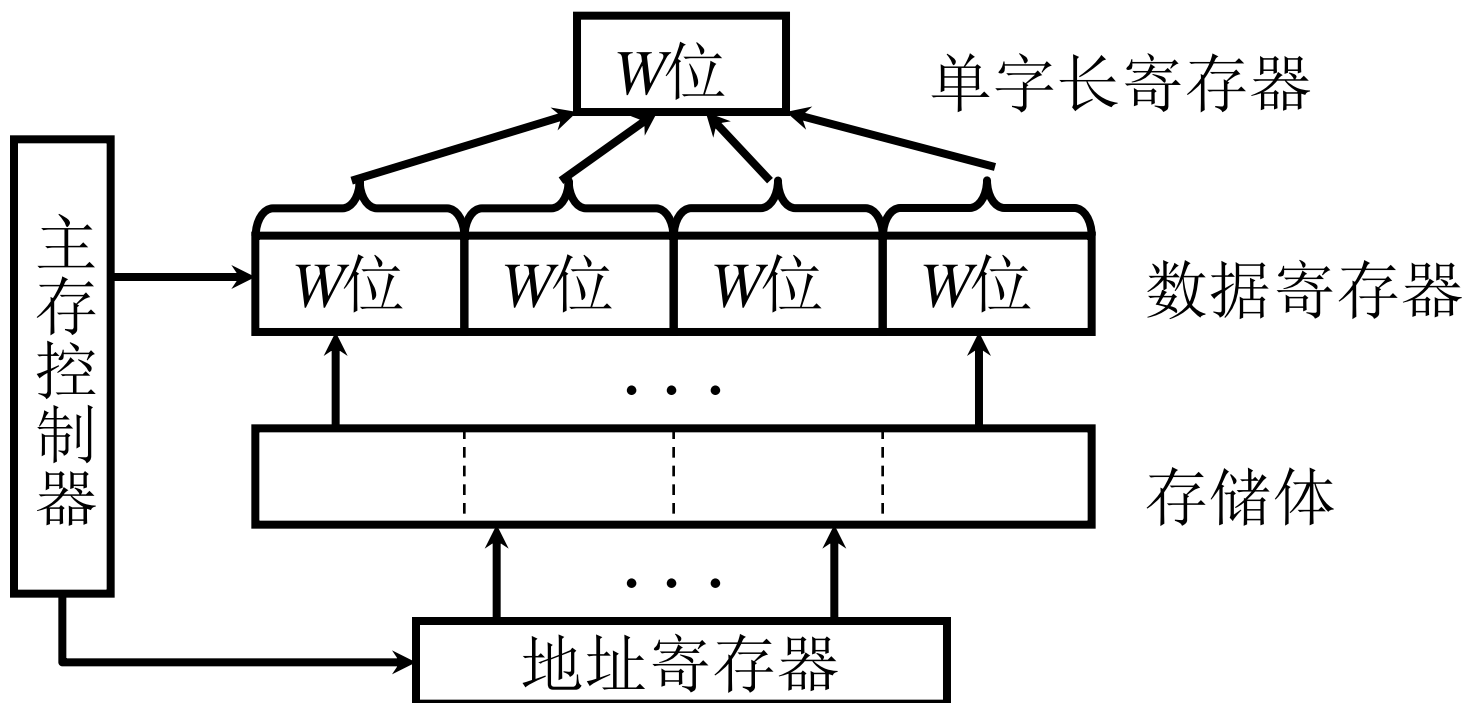
七、提高访存速度的措施

4.2

- 采用高速器件
- 采用层次结构 Cache – 主存
- 调整主存结构

1. 单体多字系统

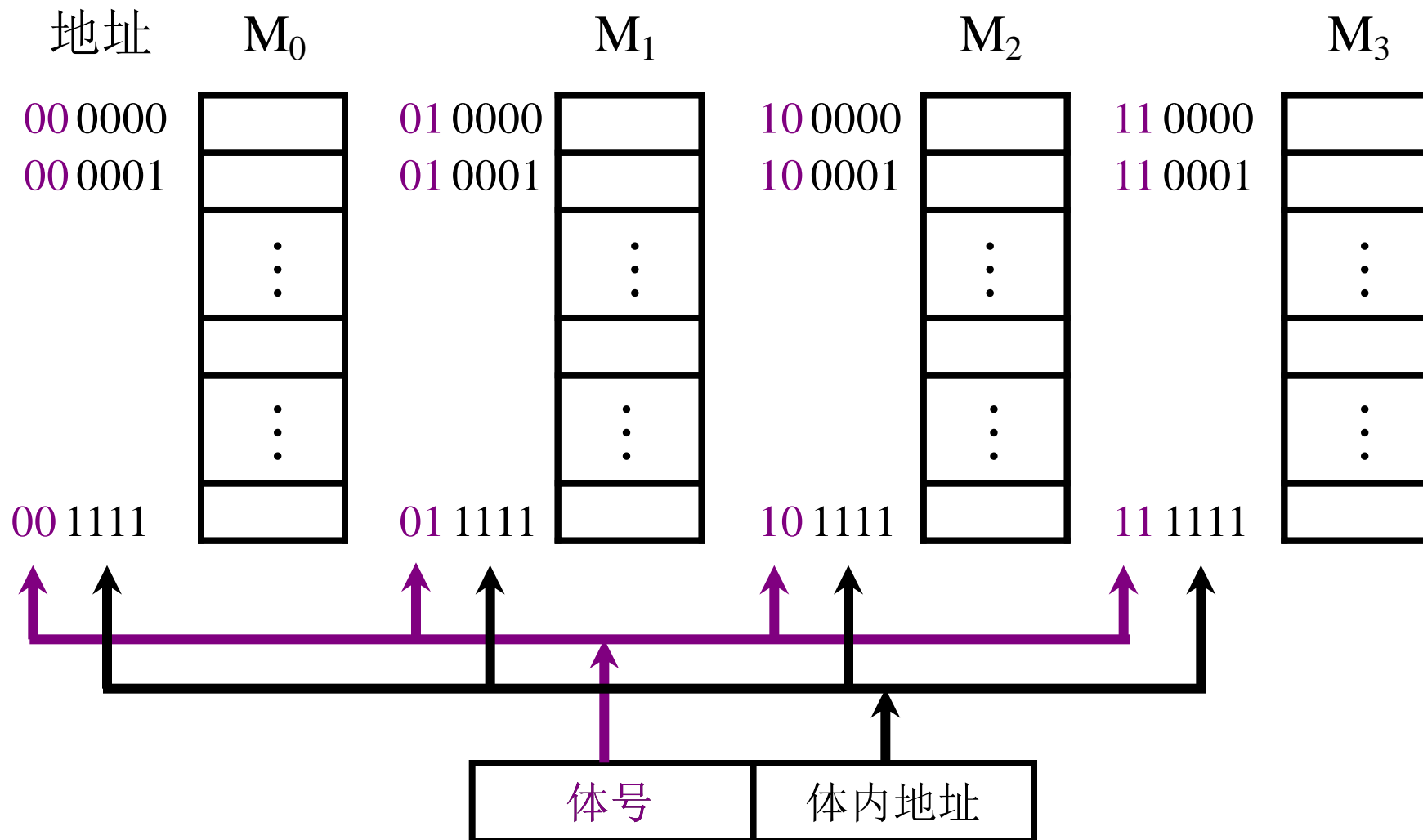
增加存储器的带宽



2. 多体并行系统

4.2

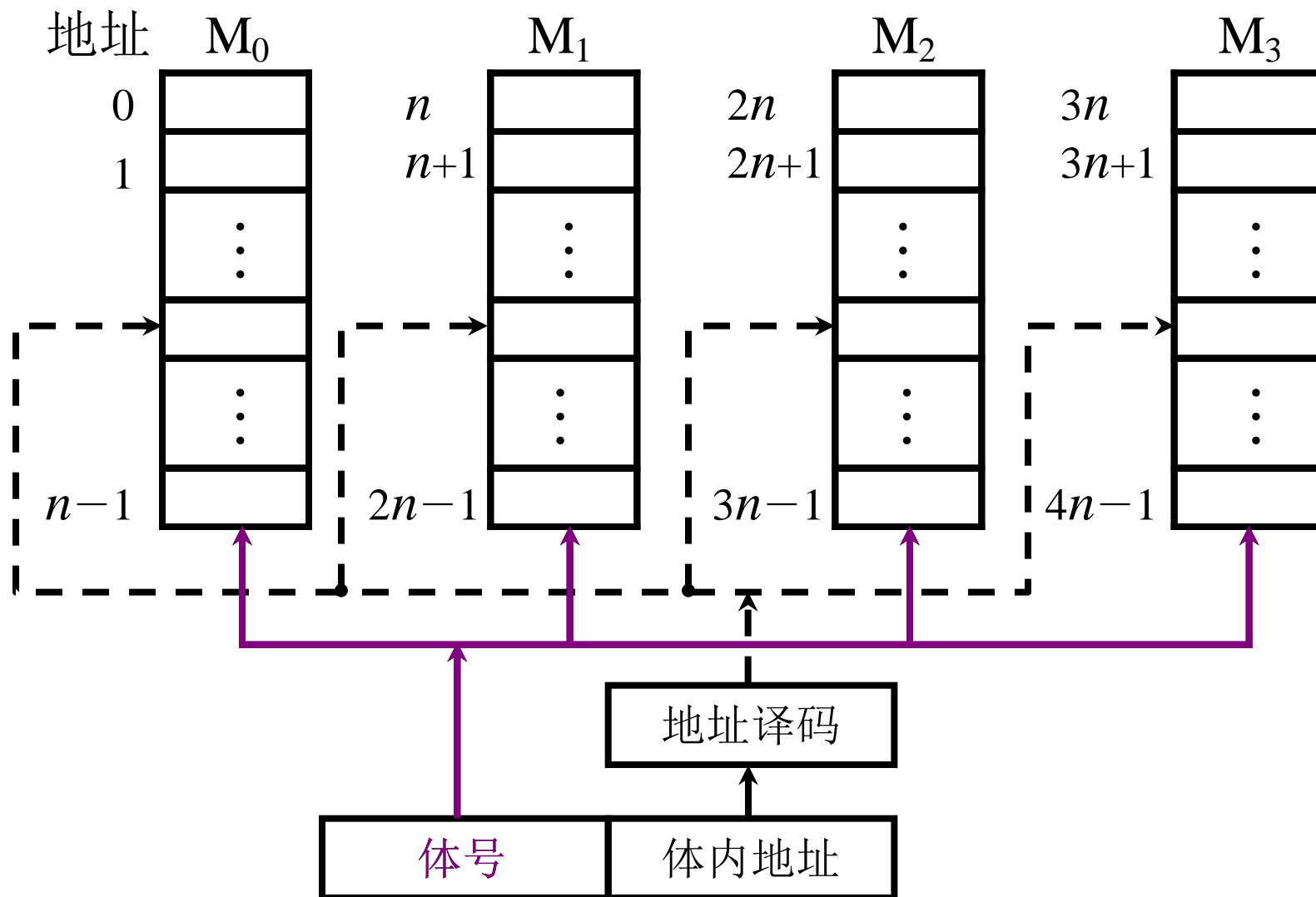
(1) 高位交叉 顺序编址



(1) 高位交叉

各个体并行工作

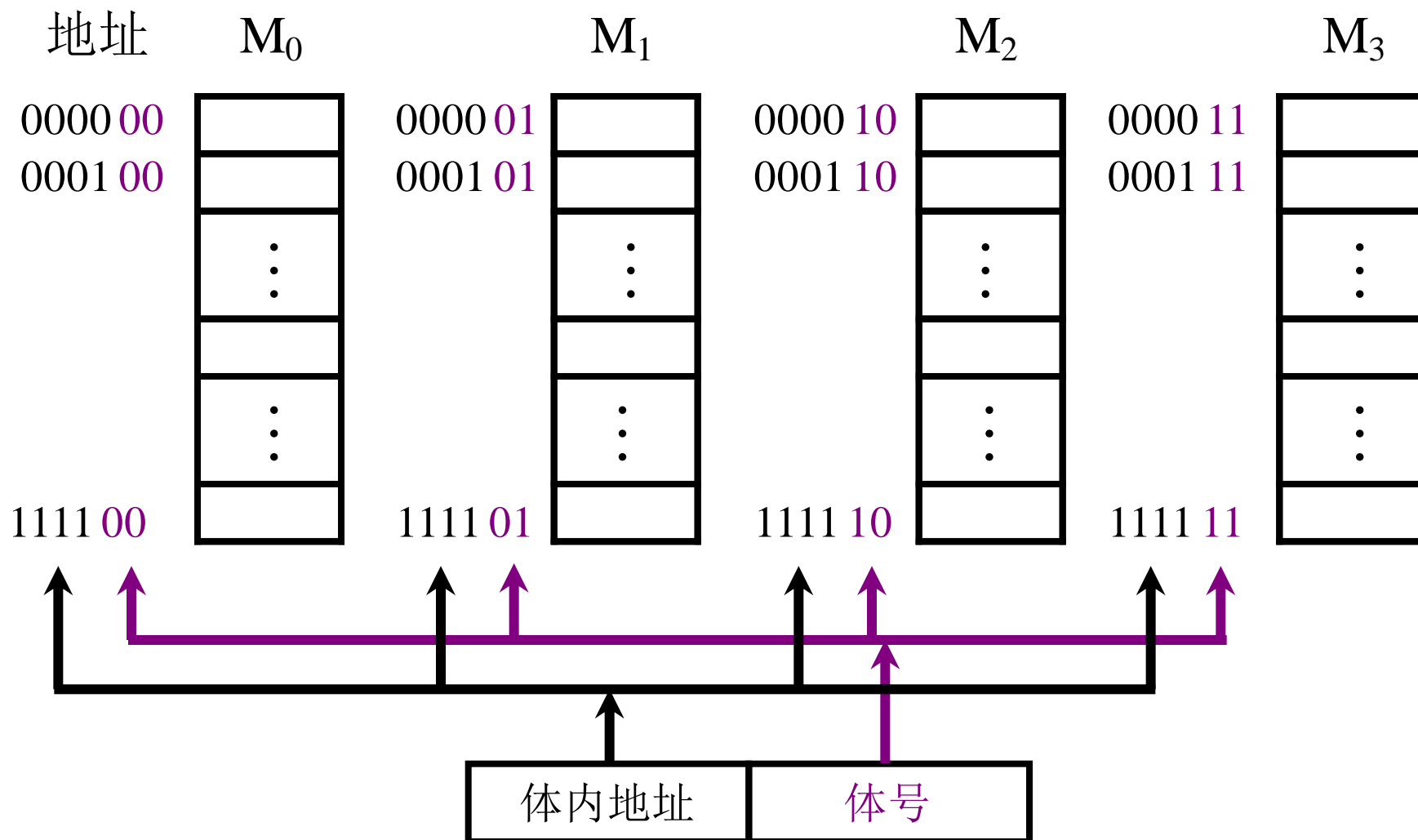
4.2



(2) 低位交叉

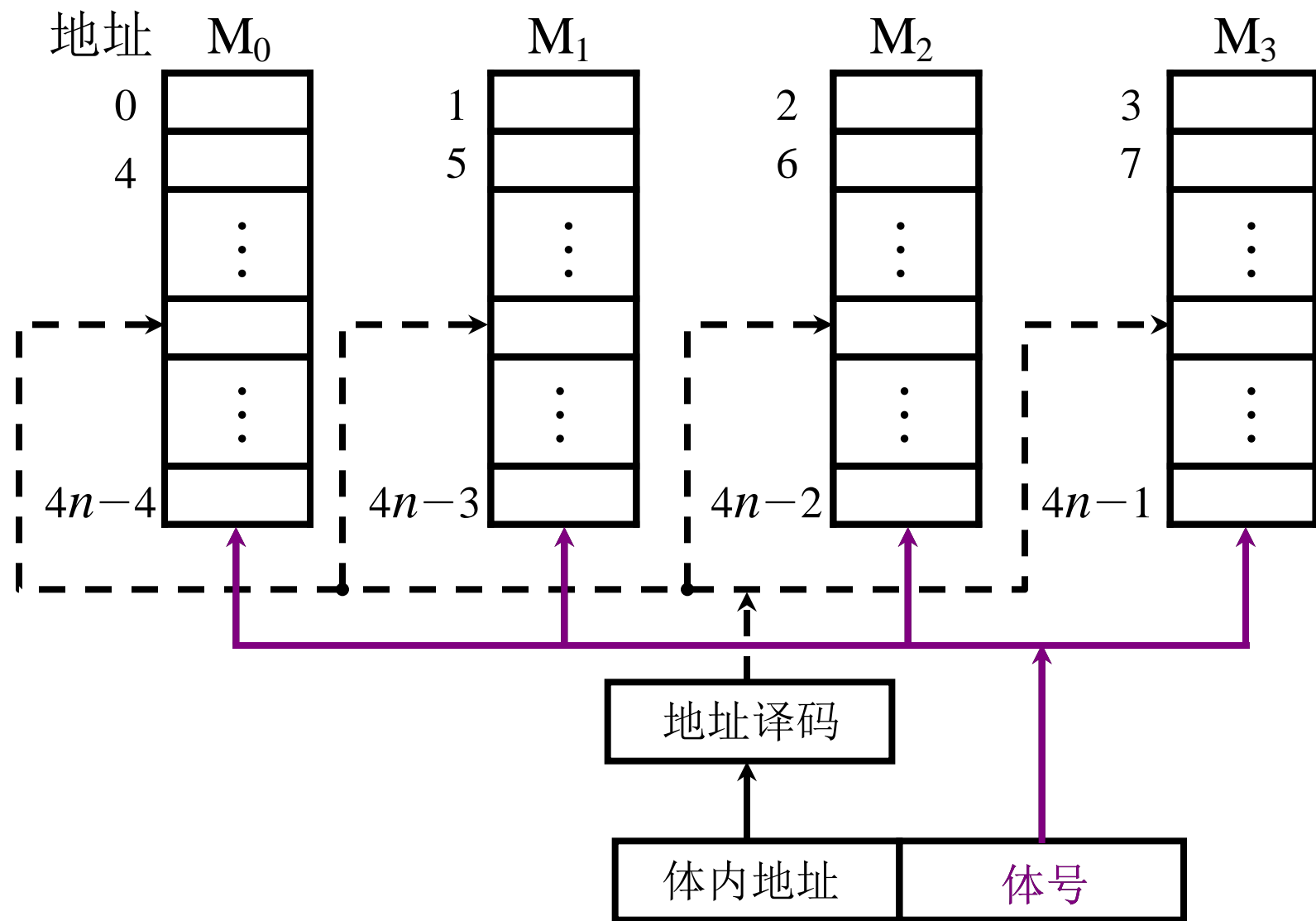
各个体轮流编址

4.2



(2) 低位交叉 各个体轮流编址

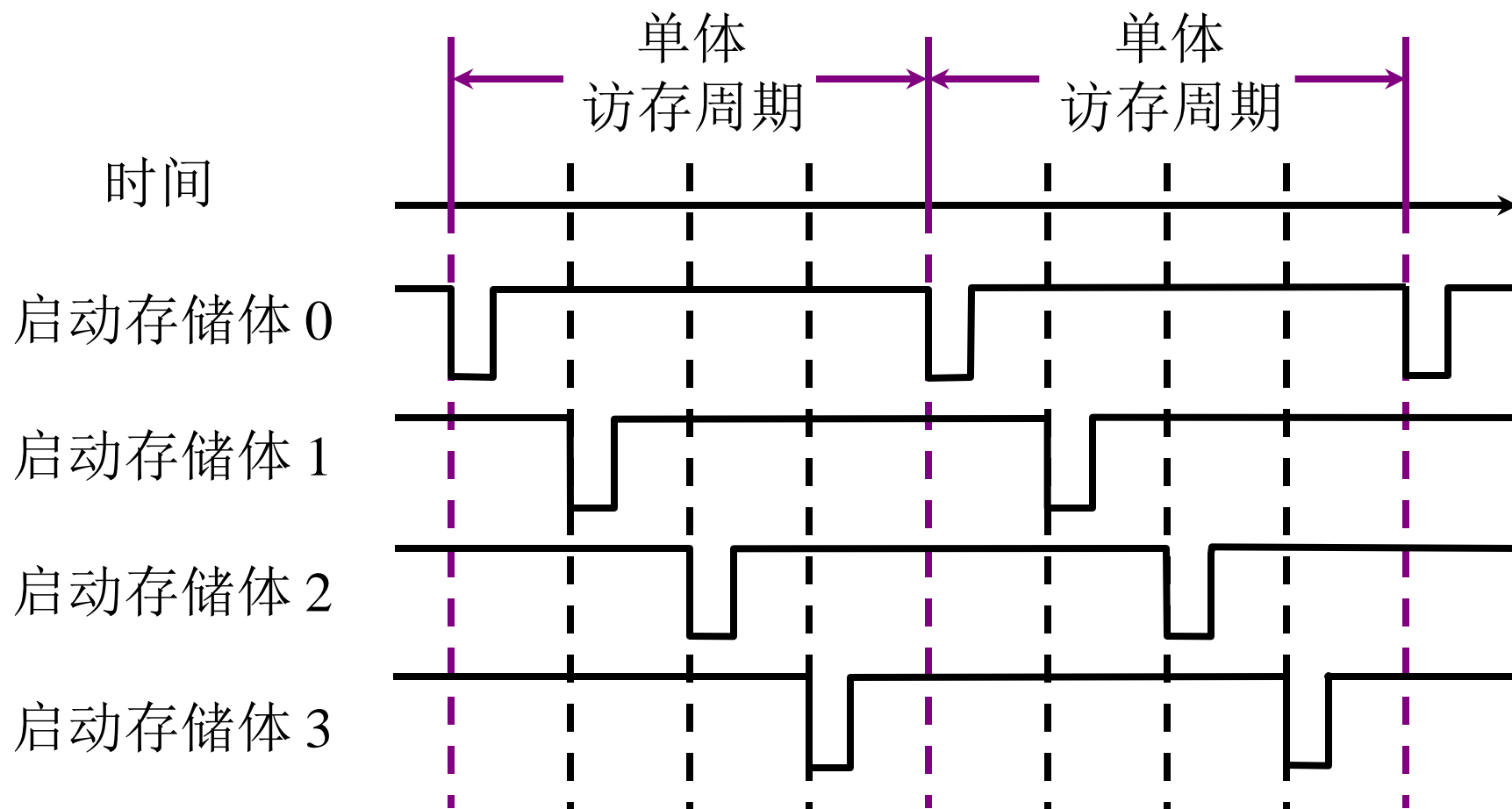
4.2



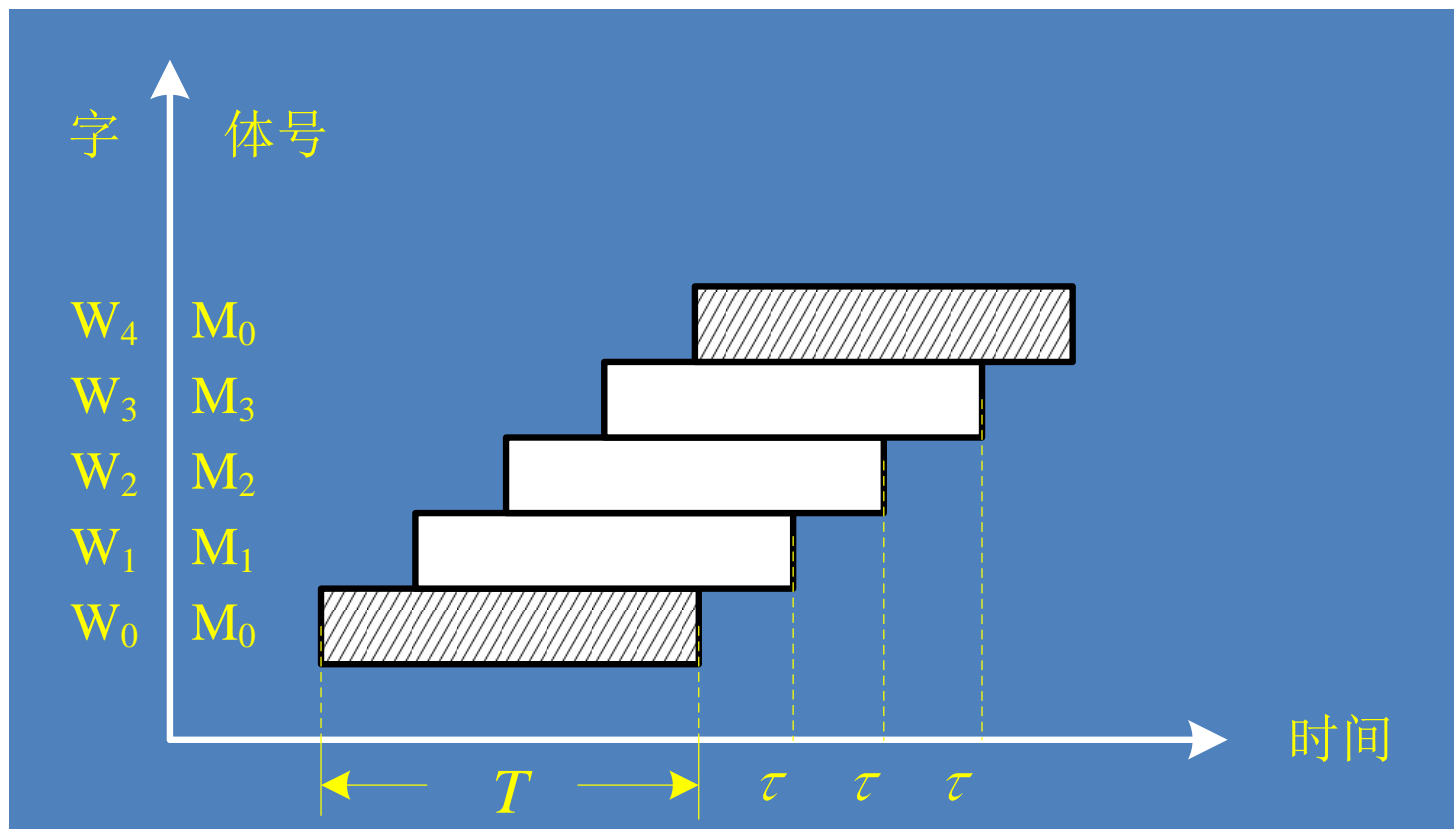
低位交叉的特点

4.2

在不改变存取周期的前提下，增加存储器的带宽

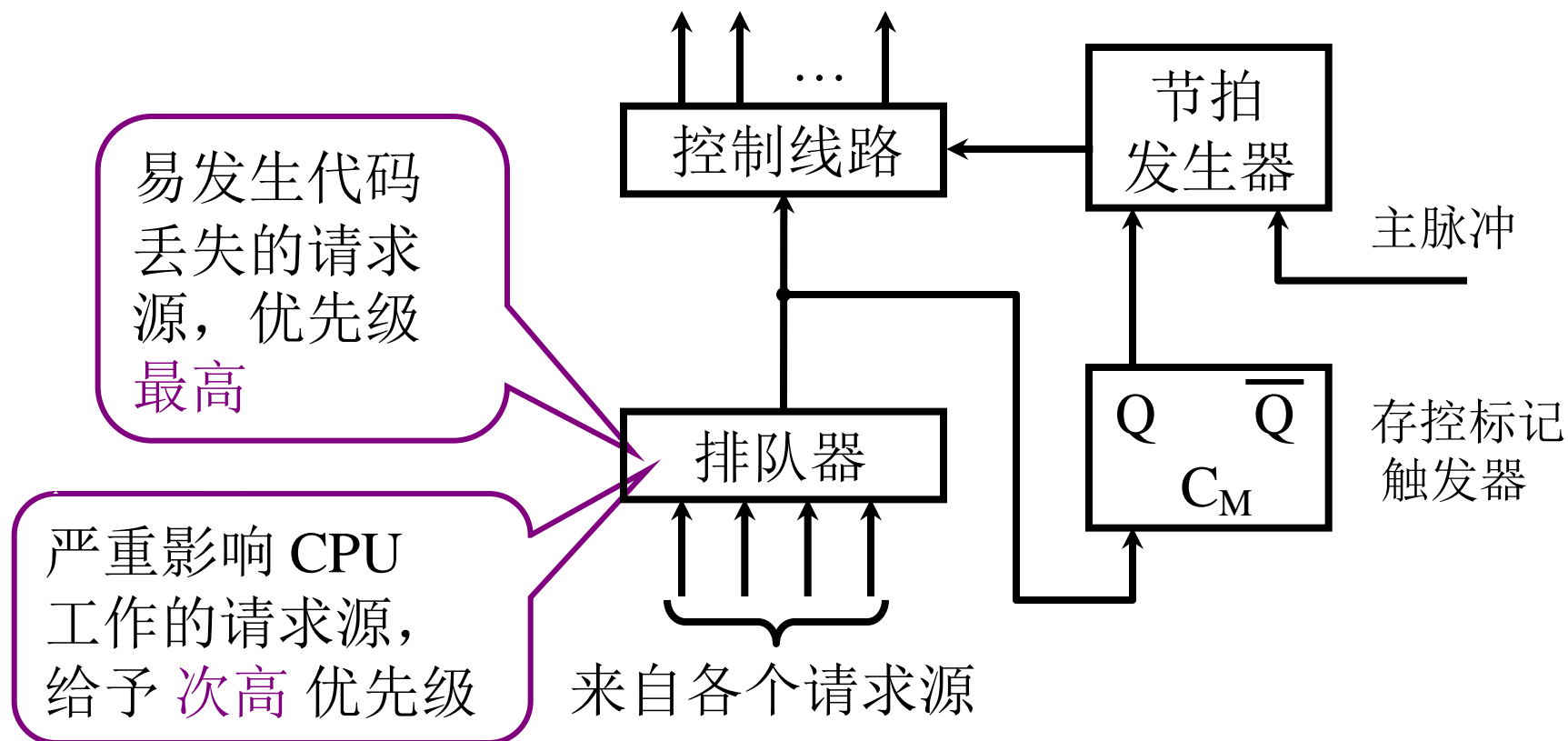


设四体低位交叉存储器，存取周期为 T ，总线传输周期为 τ ，为实现流水线方式存取，应满足 $T = 4\tau$ 。



连续读取 4 个字所需的时间为 $T + (4 - 1)\tau$

(3) 存储器控制部件（简称存控）



3.高性能存储芯片

(1) SDRAM (同步 DRAM)

在系统时钟的控制下进行读出和写入

CPU 无须等待

(2) RDRAM

由 Rambus 开发，主要解决 存储器带宽 问题

(3) 带 Cache 的 DRAM

在 DRAM 的芯片内 集成 了一个由 SRAM 组成的

Cache，有利于 猝发式读取



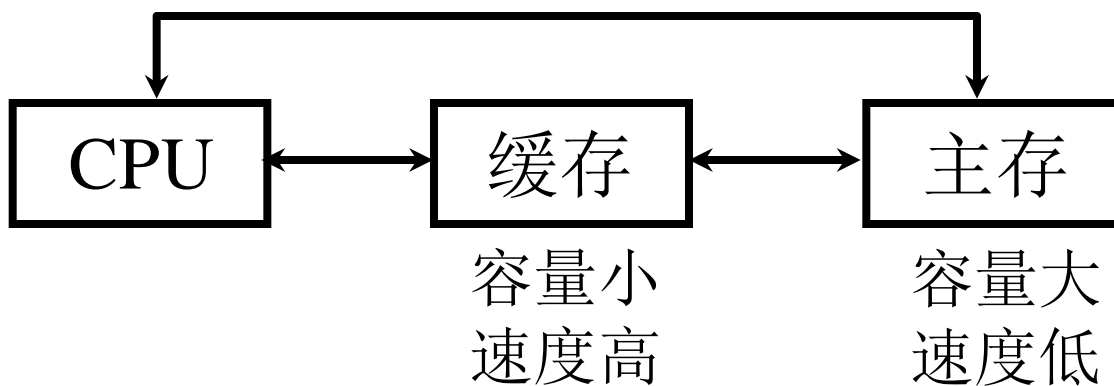
4.3 高速缓冲存储器

一、概述

1. 问题的提出

避免 CPU “空等” 现象

CPU 和主存（DRAM）的速度差异



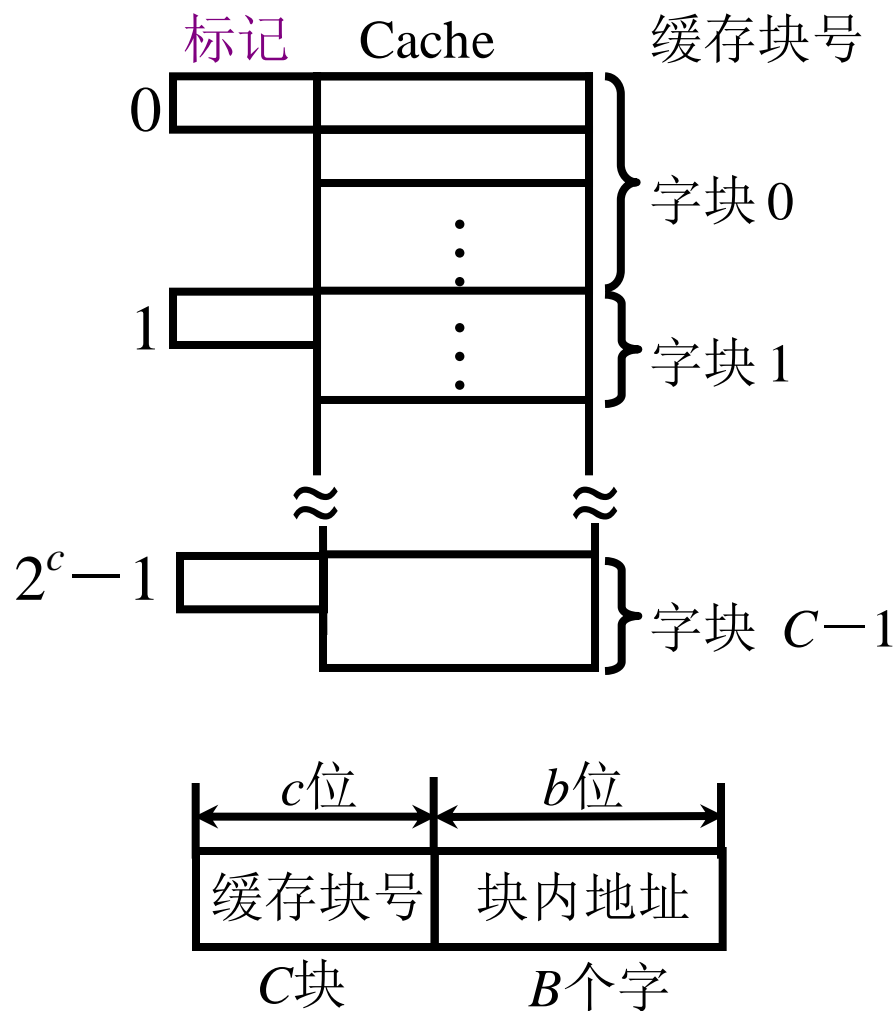
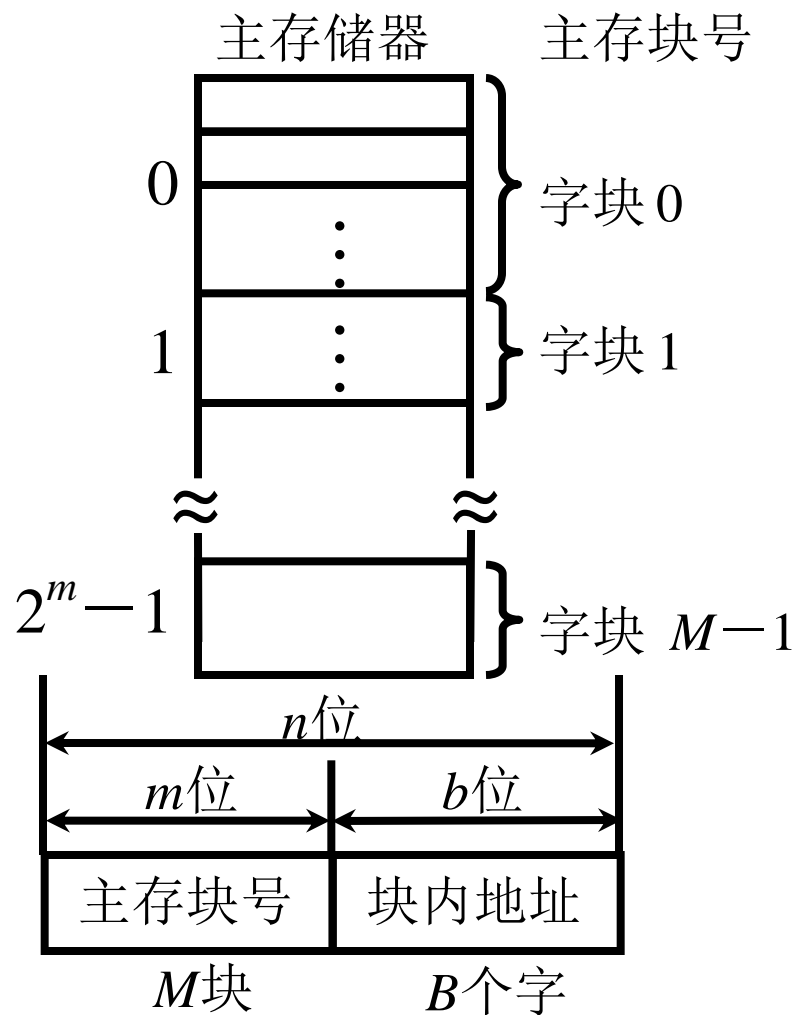
程序访问的局部性原理



2. Cache 的工作原理

4.3

(1) 主存和缓存的编址



主存和缓存按块存储

块的大小相同

B 为块长



(2) 命中与未命中

缓存共有 C 块

主存共有 M 块 $M \gg C$

命中 主存块 调入 缓存

主存块与缓存块 建立 了对应关系

用 标记记录 与某缓存块建立了对应关系的 主存块号

未命中 主存块 未调入 缓存

主存块与缓存块 未建立 对应关系



(3) Cache 的命中率

CPU 欲访问的信息在 Cache 中的 **比率**

命中率 与 Cache 的 **容量** 与 **块长** 有关

一般每块可取 4 ~ 8 个字

块长取一个存取周期内从主存调出的信息长度

CRAY_1	16体交叉	块长取 16 个存储字
--------	-------	-------------

IBM 370/168	4体交叉	块长取 4 个存储字
-------------	------	------------

(64位 \times 4 = 256位)



(4) Cache –主存系统的效率

效率 e 与 命中率 有关

$$e = \frac{\text{访问 Cache 的时间}}{\text{平均访问时间}} \times 100\%$$

设 Cache 命中率为 h ，访问 Cache 的时间为 t_c ，
访问 主存 的时间为 t_m

$$\text{则 } e = \frac{t_c}{h \times t_c + (1-h) \times t_m} \times 100\%$$



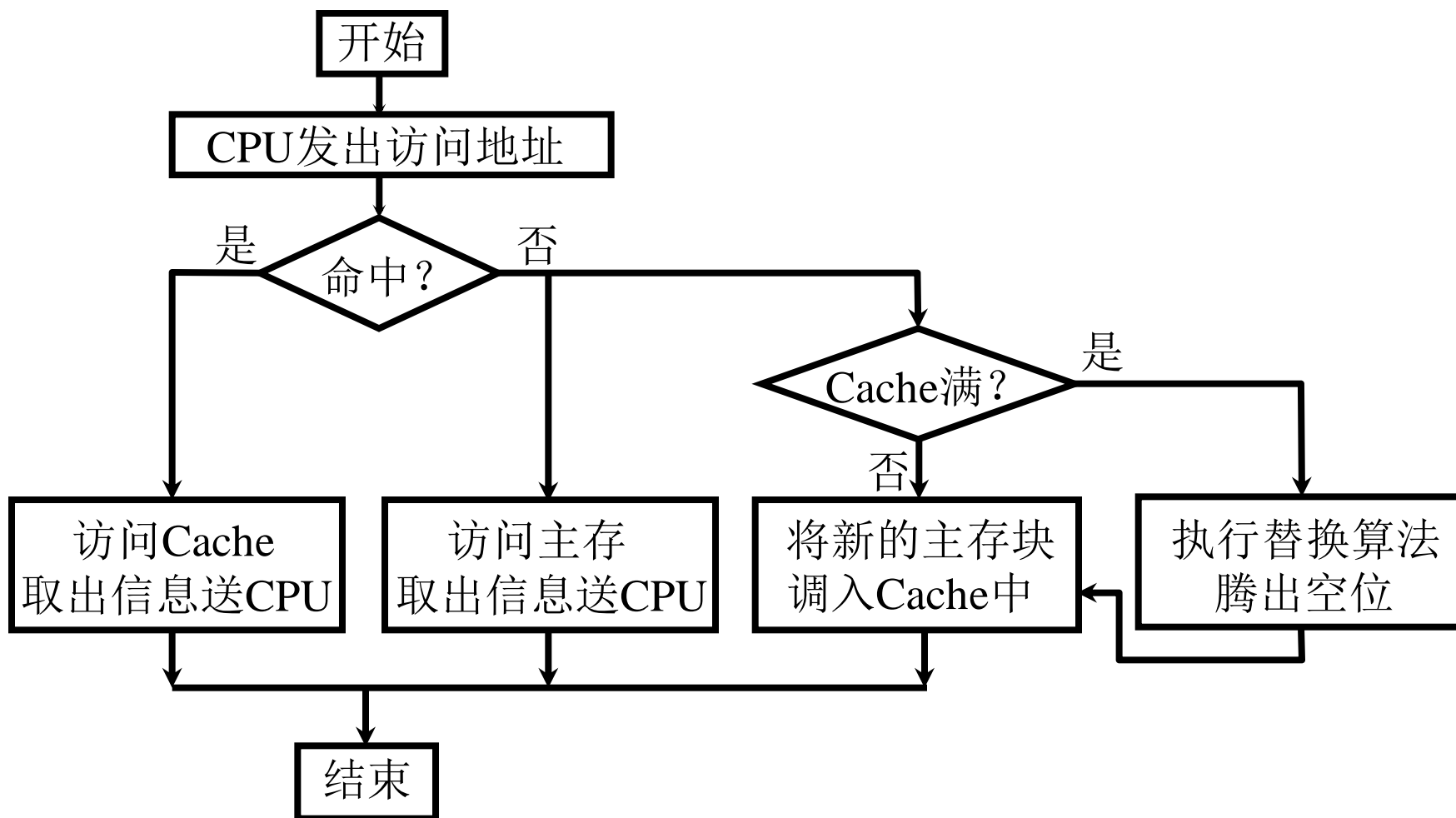
4.3



4. Cache 的 读写 操作

读

4.3



4. Cache 的 读写 操作

4.3

写 Cache 和主存的一致性

- 写直达法 (Write – through)

写操作时数据既写入Cache又写入主存

写操作时间就是访问主存的时间，读操作时不涉及对主存的写操作，更新策略比较容易实现

- 写回法 (Write – back)

写操作时只把数据写入 Cache 而不写入主存

当 Cache 数据被替换出去时才写回主存

写操作时间就是访问 Cache 的时间，

读操作 Cache 失效发生数据替换时，

被替换的块需写回主存，增加了 Cache 的复杂性



5. Cache 的改进

(1) 增加 Cache 的级数

片载（片内）Cache

片外 Cache

(2) 统一缓存和分立缓存

指令 Cache 数据 Cache

与主存结构有关

与指令执行的控制方式有关 是否流水

Pentium	8K 指令 Cache	8K 数据 Cache
---------	-------------	-------------

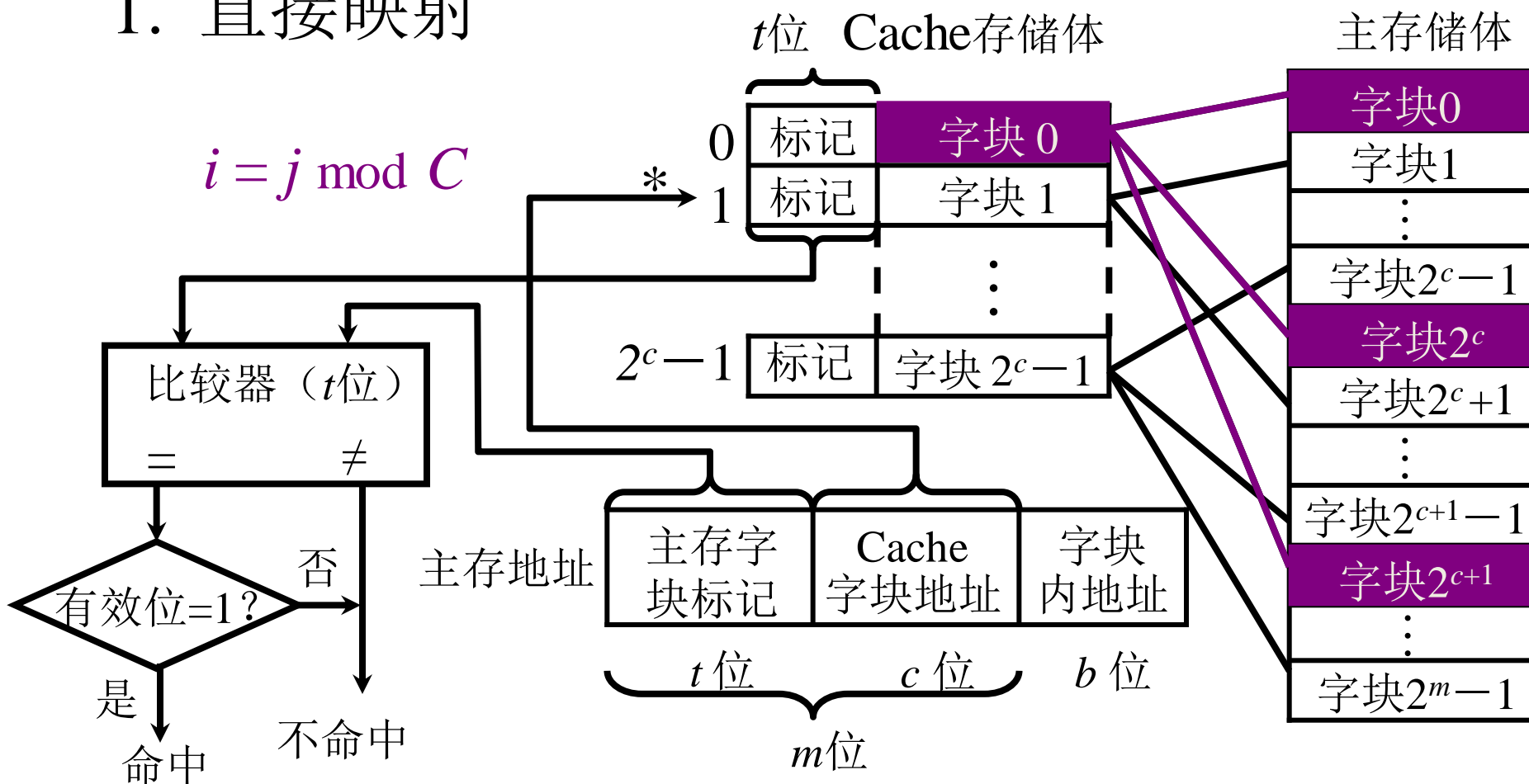
PowerPC620	32K 指令 Cache	32K 数据 Cache
------------	--------------	--------------



二、Cache – 主存的地址映射

4.3

1. 直接映射

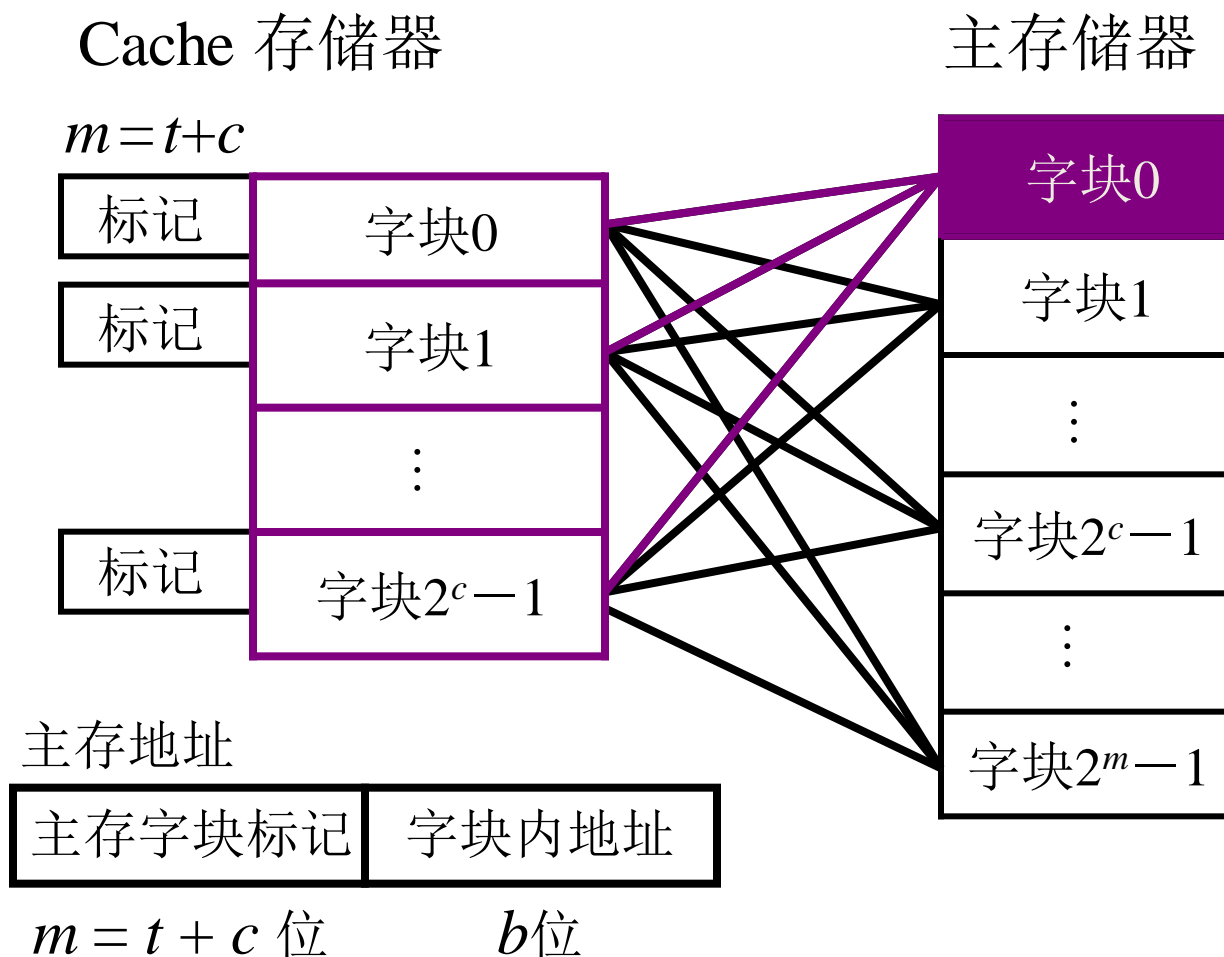


每个缓存块 i 可以和若干个主存块对应

每个主存块 j 只能和一个缓存块对应



2. 全相联映射



主存 中的 任一块 可以映射到 缓存 中的 任一块

3. 组相联映射

主存储器

组 Cache 共 Q 组，每组内两块 ($r=1$)

0	标记	字块 0	标记	字块 1
1	标记	字块 2	标记	字块 3
	⋮	⋮	⋮	⋮
$2^{c-r}-1$	标记	字块 2^c-2	标记	字块 2^c-1

主存地址

主存字块标记	组地址	字块内地址
$s = t + r$ 位	$q = c - r$ 位	b 位
m 位		

字块0
字块1
⋮
字块 $2^{c-r}-1$
字块 2^{c-r}
字块 $2^{c-r}+1$
⋮
字块 2^{c-r+1}
⋮
字块 2^m-1

$$i = j \bmod Q$$

直接组相联映射

某一主存块 j 按模 Q 映射到 缓存 的第 i 组中的 任一块



映射方法特点

- 全相联映像方式
 - 优点：冲突概率小，Cache的利用高。
 - 缺点：比较器难实现，需要一个访问速度很快代价高的相联存储器
 - 应用场合：
 - 适用于小容量的Cache

- 直接相联映像方式
 - 优点：比较电路少 m 倍线路，所以硬件实现简单，Cache地址为主存地址的低几位，不需变换。
 - 缺点：冲突概率高（抖动）
 - 应用场合
 - 适合大容量Cache

- 组相联映像方式
 - 比全相联容易实现，冲突低
 - $v=1$ ，则为直接相联映射方式
 - $u=1$ ，则为全相联映射方式
 - v 的取值一般比较小，一般是2的幂，称之为 v 路组相联cache.

举例

- 例1：假设主存容量为512KB，cache容量为4K，每个字块为16个字，每个字32位。
 - cache地址有多少位？可容纳多少块？
 - 主存地址多少位？可容纳多少块？
 - 在直接映射方式下，主存的第几块映射到cache的第5块（设起始字块为第1块）
 - 画出直接映射方式下主存地址字段中各段的位数。

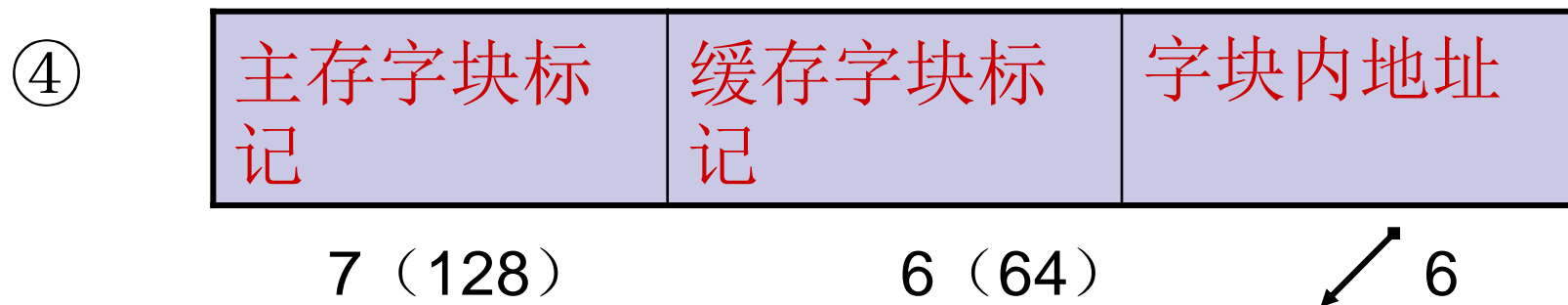
解：

①cache地址12位，可容纳64块。

②主存地址19位，共8192块。

③在直接映射方式下，由于cache共有64块，主存共有8192块，所以主存的5，64+5， $2 \times 64 + 5$ ，...， $2^{13} - 64 + 5$ 块能映射到cache的第5块中。以64位模， $8192/64=128$

$$i = j \bmod C$$



注：容量以Byte为单位，因为每个字32位，所以折算成字节，就等于地址范围是原来的4倍，等于 $16 \times 4 = 64 = 2^6$

- 例2：假设主存容量为 $512K \times 16$ 位，cache容量为 4096×16 位，块长为4个16位的字，访存地址为字地址
 - 在直接映射方式下，设计主存的地址格式
 - 在全相联映射方式下，设计主存的地址格式
 - 在二路组相联映射方式下，设计主存的地址格式。
 - 若主存容量为 $512K \times 32$ 位，块长不变，在四路组相联映射方式下，设计主存的地址格式。

解:

主存字块标记	缓存字块标记	字块内地址
7 (128)	10 (64)	2

(a) 直接映射方式主存地址格式

主存字块标记17	字块内地址2
----------	--------

(b) 全相联映射方式主存地址格式

主存字块标记	组地址	字块内地址
8	9	2

(c) 2路组相联映射方式主存地址格式

512K × 32 =

1024K × 16

20位

主存字块标记	组地址	字块内地址
10	8	2

(d) 4路组相联映射方式主存地址格式

- 例3：设某机主存容量为16MB，cache的容量为8KB。每字块有8个字，每字32位，设计一个四路组相联的cache组织。
 - 画出主存地址字段中各段的位数
 - 设cache初态为空，CPU依次从主存第0，1，2...,99号单元读出100个字（主存一次读出一个字），并重复此序读10次，问命中率是多少？
 - 若cache 的速度是主存的5倍，试问有cache 和无cache相比，速度提高多少？

解:

主存字块标记	组地址	字块内地址
13	6	5

(1) Cache容量为 $8KB=2^{13}B$, 字块大小为 2^5B , 得cache共有 2^8 。

根据四路组相联映射, $2^r=4$, 得 $r=2$. 则 $q=6$ 位。

(2) 由于每个字块有8个字, 初态cache为空, 因此CPU读第0号单元时, 未命中, 必须访问主存, 同时将该字所在主存块调入cache第0组中任一块, 接着CPU读1~7块单元时均命中。同理, CPU读第8, 16,...96号时均未命中。可见CPU在连续读100个字中共有13次未命中, 而后9次循环读100个字全部命中, 命中率为:

$$\frac{100 \times 10 - 13}{100 \times 10} = 0.987$$

组相联映像把cache分成Q组，每组R块，并有以下关系：

$$i = j \bmod Q$$

其中，i为缓存的组号，j为主存的块号。某一主存块按照模Q映射到缓存的第i组内。

本题中：Q=64组，组内4块。

这样地址单元0~100分别对应于字块0~12，如右表。

按照映射关系都映射到缓存的第0~12组中4块中的任一块。未命中，一次调入1块（即8个字）。

字块0	字块1	字块2	字块3
-----	-----	-----	-----

字块0	0~7
字块1	8~15
字块2	16~23
字块3	24~31
字块4	32~39
字块5	40~47
字块6	48~55
字块7	56~63
字块8	64~71
字块9	72~79
字块10	80~87
字块11	88~95
字块12	96~103

(3) 设主存周期为 $5t$,cache为 t 。没有cache的访问时间为 $5t \times 1000$, 有cache的访问时间为 $t(1000-13) + 5t \times 13$, 则有无的速度提高倍数为

$$\frac{5t \times 1000}{t(1000-13) + 5t \times 13} - 1 \approx 3.75$$

三、替换算法

1. 先进先出（FIFO）算法

2. 近期最少使用（LRU）算法

小结

成本与活

直接 某一主存块只能固定映射到某一缓存块

全相联 某一主存块能映射到任一缓存块

组相联 某一主存块只能映射到某一缓存组中的任一块



- 例子： 设cache有1、2、3、4共4个块，a、b、c、d等为主存中的块,访问顺序一次如下： a、b、c、d、b、b、c、c、d、d、a,下次若要再访问e块。问，采用LFU和LRU算法替换结果是不是相同？

	LFU（最不经常使用）					LRU（近期最少使用）				
	说明	1块	2块	3块	4块	说明	1块	2块	3块	4块
a	a进入	1	0	0	0	a进入	0	1	1	1
b	b进入	1	1	0	0	b进入	1	0	2	2
c	c进入	1	1	1	0	c进入	2	1	0	3
d	d进入	1	1	1	1	d进入	3	2	1	0
b	命中	1	2	1	1	命中	4	0	2	1
b	命中	1	3	1	1	命中	5	0	3	2
c	命中	1	3	2	1	命中	6	1	0	3
c	命中	1	3	3	1	命中	7	2	0	4
d	命中	1	3	3	2	命中	8	3	1	0
d	命中	1	3	3	3	命中	9	4	2	0
a	命中	2	3	3	3	命中	0	5	3	1
e	替换a	1	0	0	0	替换b	1	0	4	2

4.4 辅助存储器

一、概述

1. 特点 不直接与 CPU 交换信息

2. 磁表面存储器的技术指标

(1) 记录密度 道密度 D_t 位密度 D_b

(2) 存储容量 $C = n \times k \times s$

(3) 平均寻址时间 寻道时间 + 等待时间

辅存的速度 $\left\{ \begin{array}{l} \text{寻址时间} \\ \text{磁头读写时间} \end{array} \right.$

(4) 数据传输率 $D_r = D_b \times V$

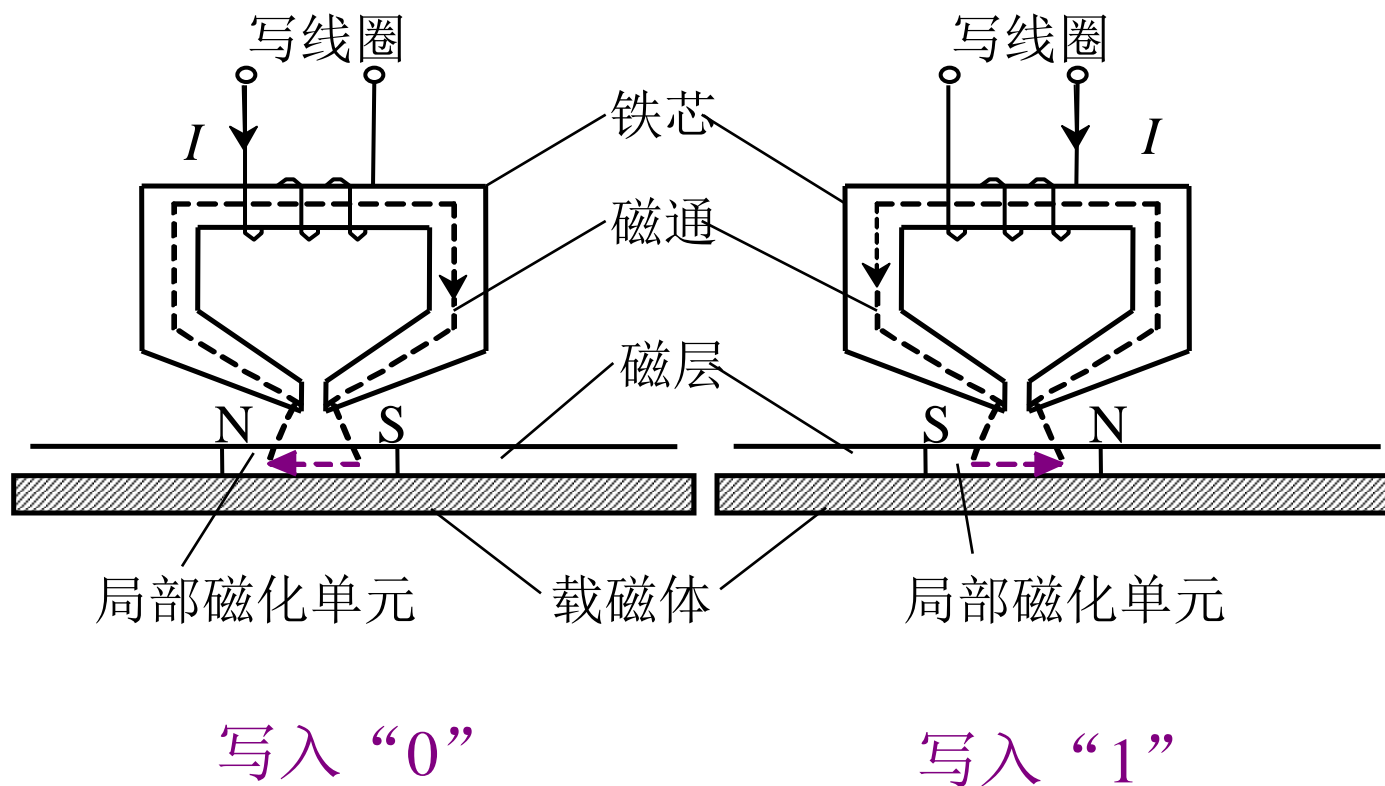
(5) 误码率 出错信息位数与读出信息的总位数之比



二、磁记录原理和记录方式

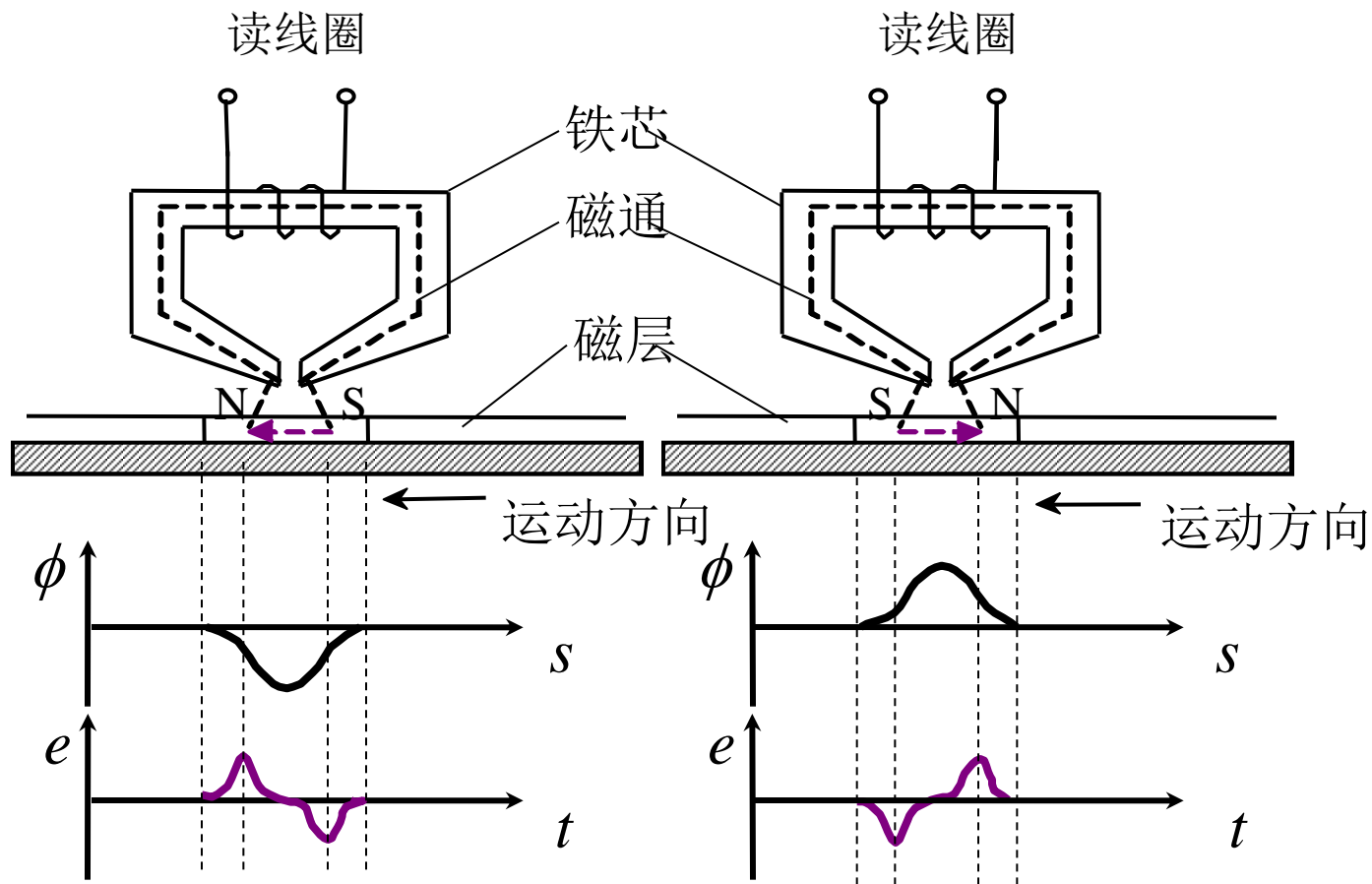
1. 磁记录原理

写



1. 磁记录原理

读



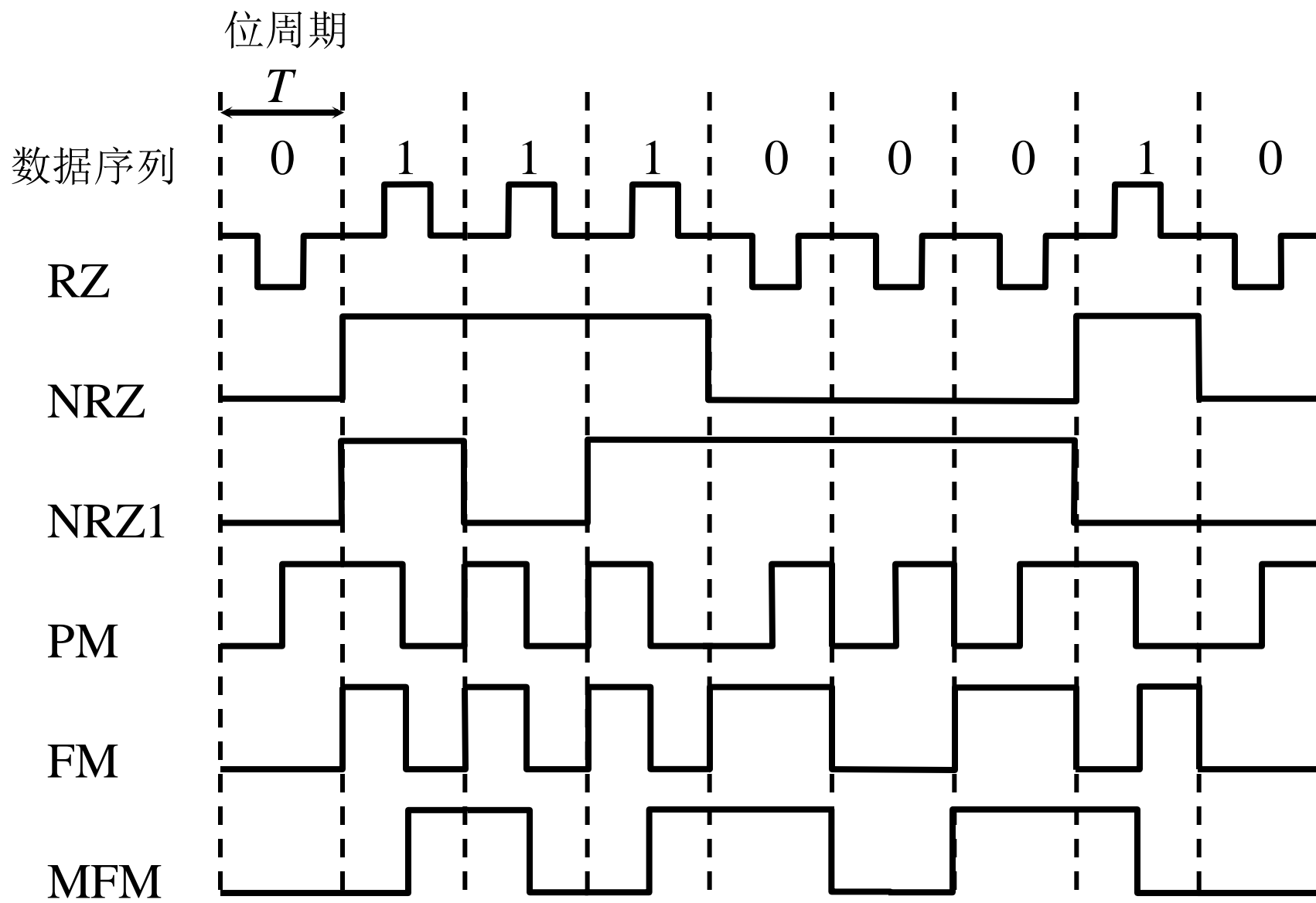
读出“0”

读出“1”

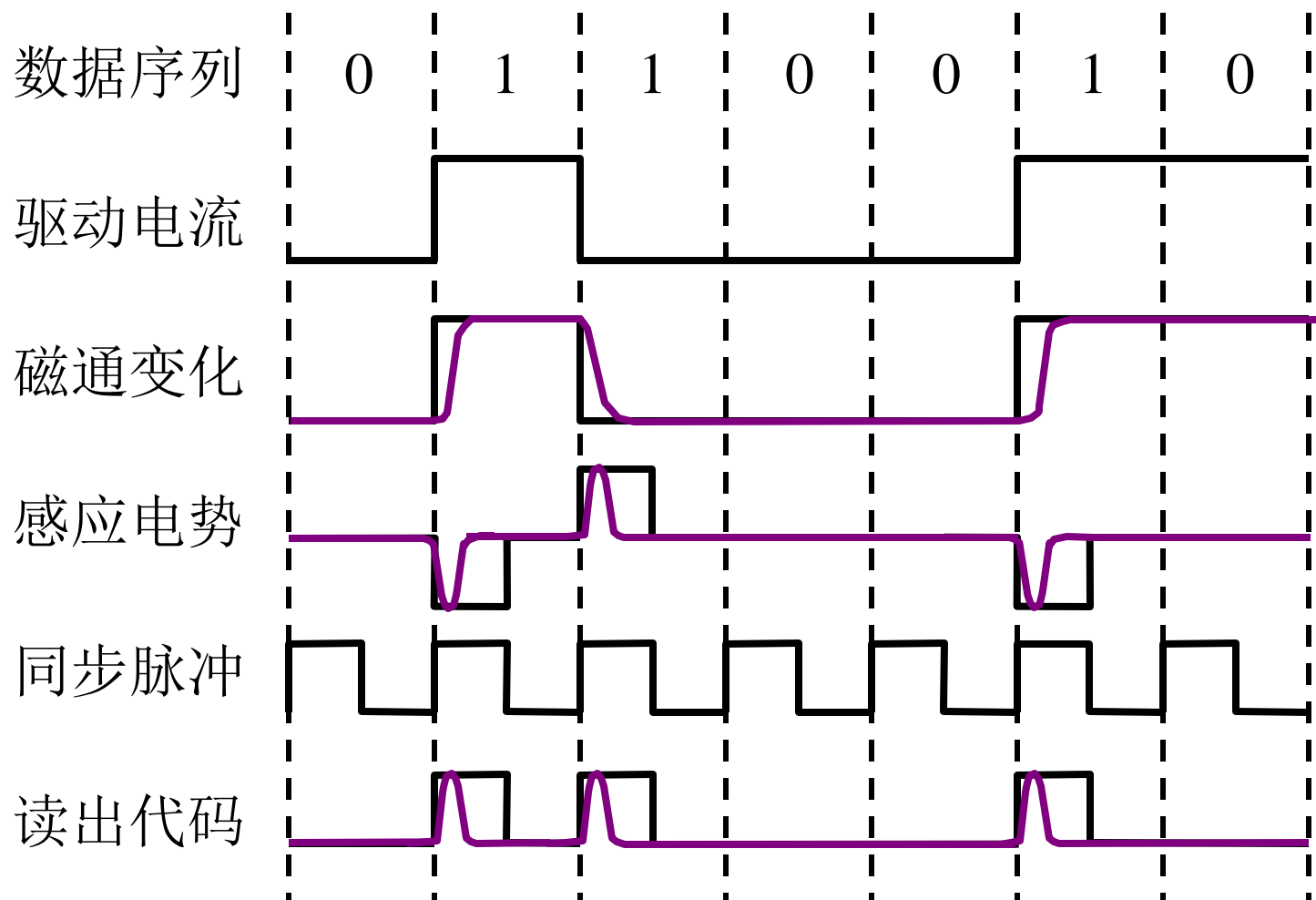


2. 磁表面存储器的记录方式

4.4



例 NRZ1 的读出代码波形



三、硬磁盘存储器

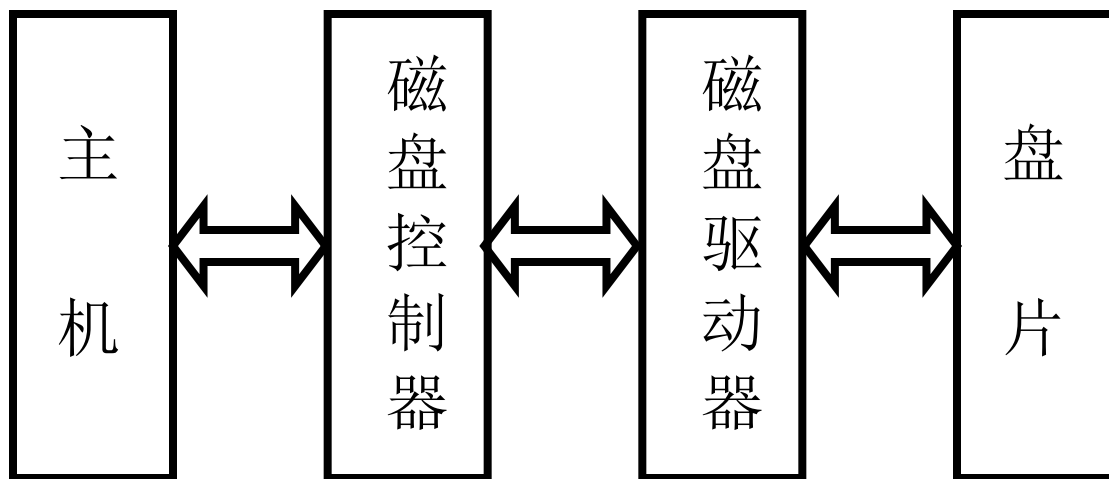
4.4

1. 硬磁盘存储器的类型

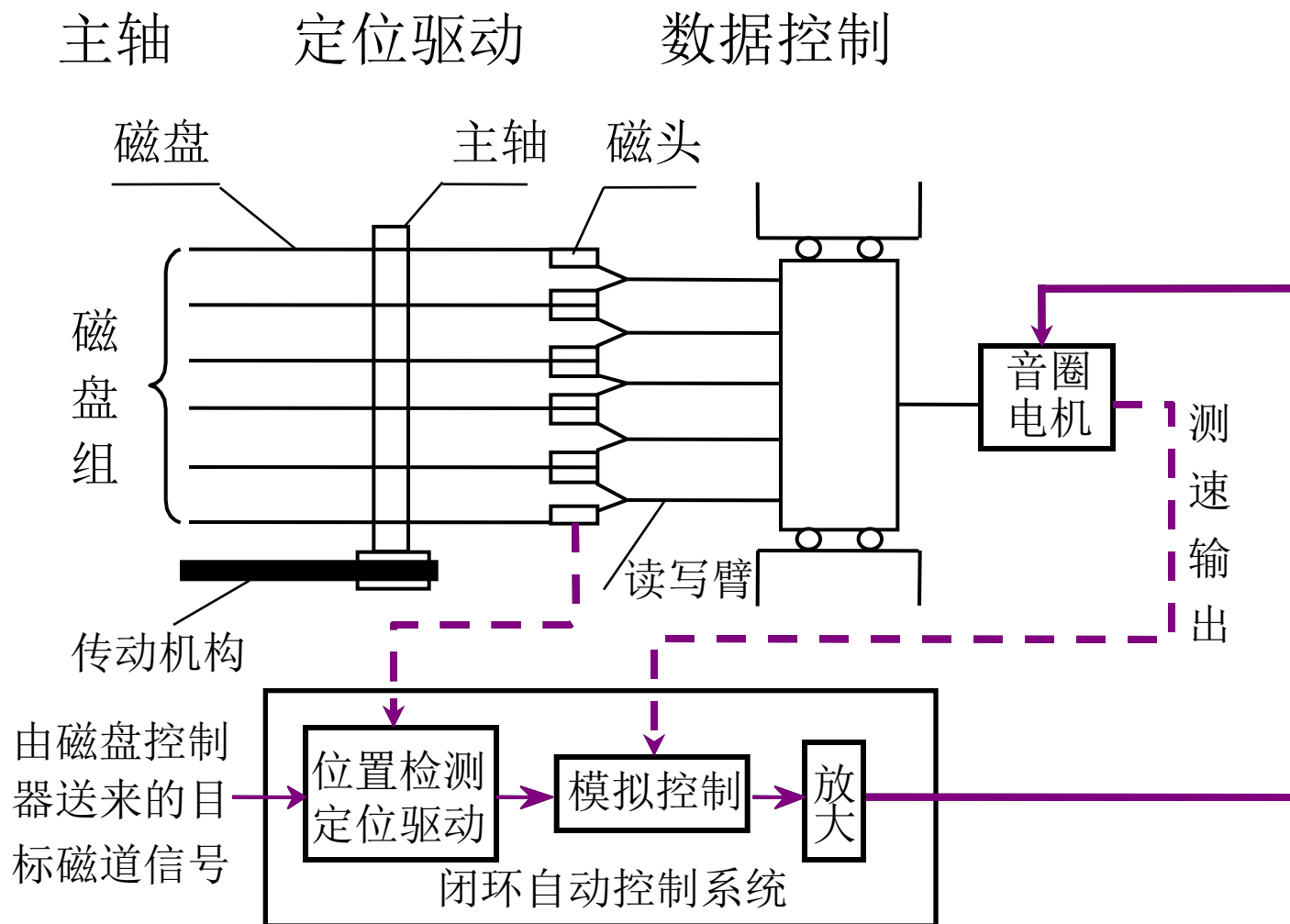
(1) 固定磁头和移动磁头

(2) 可换盘和固定盘

2. 硬磁盘存储器结构



(1) 磁盘驱动器



(2) 磁盘控制器

- 接收主机发来的命令，转换成磁盘驱动器的控制命令
- 实现主机和驱动器之间的数据格式转换
- 控制磁盘驱动器读写

磁盘控制器 是

主机与磁盘驱动器之间的 接口 { 对主机 通过总线
对硬盘 (设备)

(3) 盘片

由硬质铝合金材料制成

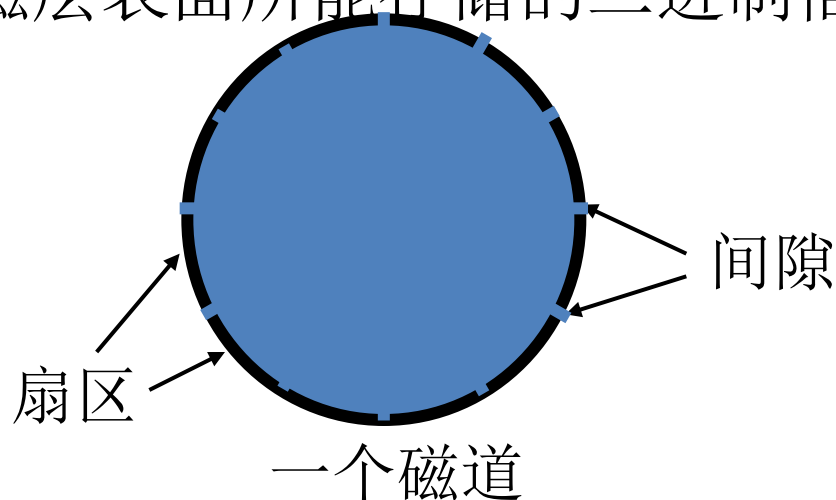


磁表面存储器的主要技术指标

- **静态属性**指的是与容量有关而与速度无关的参数，如存储密度、存储容量等。
- **动态属性**指的是磁盘读写数据时的访问速度，如寻址时间、数据传输率等

1. 存储密度(记录密度)

- 单位长度或面积的磁层表面所能存储的二进制信息量
- 道密度 D_t
- 道距 P
- 线密度 D_b



磁表面存储器的主要技术指标（续）

(1) 道密度Dt (横向密度): 沿磁盘半径方向上单位长度的磁道数目

TPI (Tracks Per Inch), 道/英寸

TPM (Tracks Per Millimetre), 道/毫米

(2) 道距P: 相邻两条磁道中心线之间的距离

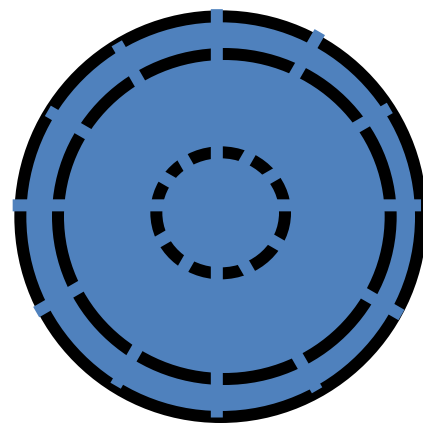
$P = 1 / Dt$ (英寸或毫米)

(3) 线密度Db (位密度) (纵向密度):

单位长度磁道所能记录的二进制信息位数

bpi (bits per inch) 位/英寸

bpm (bits per millimetre) 位/毫米



	$f \bullet t$	f —数码的脉冲频率(位/秒)
对磁盘:	$Db = \frac{f \bullet t}{\pi \bullet d_{min}}$	t —磁盘每转所需时间(秒)
		d_{min} —磁盘的内圈直径

磁表面存储器的主要技术指标（续）

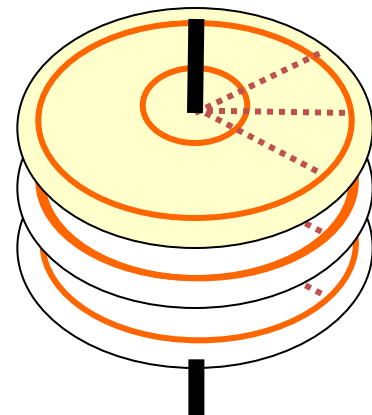
2. 存储容量

- 非格式化存储容量

- 可全部利用的磁化单元数 (物理容量)
- $\text{记录面数} \times \text{每面磁道数} \times \text{内圈磁道周长} \times \text{位密度}$

- 格式化存储容量

- 用户真正可以使用的容量
- $\text{记录面} \times \text{柱面/面} \times \text{扇区/道} \times \text{字节/扇区}$
- 例如, $4 \times 306 \times 17 \times 512$
- $= 10653696 \text{ 字节} = 10 \text{ MB}$

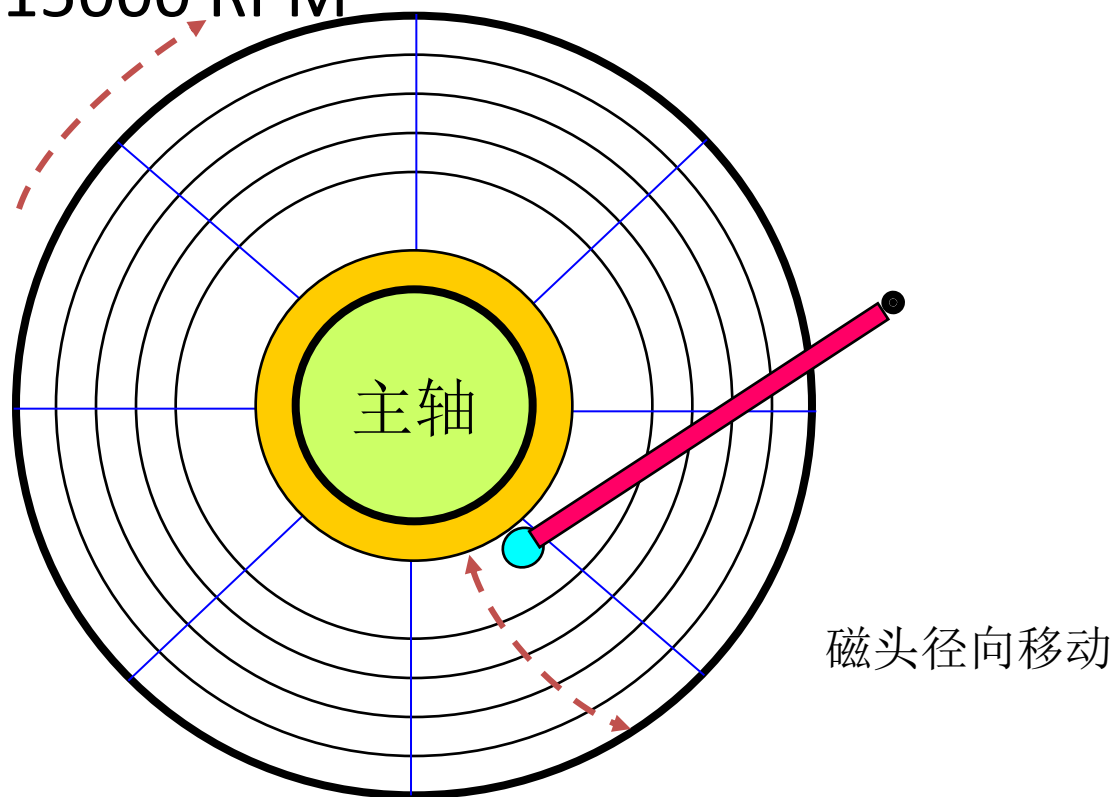


磁表面存储器的主要技术指标（续）

3. 旋转速度

— 5400~15000 RPM

盘片恒速旋转



磁表面存储器的主要技术指标（续）

4. 寻址时间：磁头接到读/写命令，从原来的位置移动到指定位置的时间

$$T_a = T_{s_a} + T_{w_a} = \frac{ts_{\max} + ts_{\min}}{2} + \frac{tw_{\max} + tw_{\min}}{2}$$

- T_{s_a} ：平均寻道时间—将磁头移到所需磁道上的时间，也称定位时间
- T_{w_a} ：平均等待时间—磁头等待读/写的区段旋转到它的下方所需等待时间

磁表面存储器的主要技术指标（续）

5. 数据传输率 D_r ：单位时间内向主机传送数据的位数或字节数（位/秒，字节/秒）

$$D_r = D \cdot V$$

D — 记录密度 （位密度 ； 位密度 * 磁道数 ）

V — 速度 （线速度 ； 走带速度 ）

6. 误码率：衡量磁表面存储器出错概率

$$\text{误码率} = \frac{\text{出错信息位数}}{\text{总信息量}}$$

7. 价格：设备价格

$$\text{位价格} = \text{设备价格} / \text{容量}$$

计算举例

例：设有一个盘面直径为18inch的磁盘组, 20记录面, 每面5inch用于记录信息, 记录密度为100TPI和1000bpi, 转速2400RPM, 道间移动时间为0.2ms, 计算:

每一记录面的磁道数 T : $T=5*100=500$ 道

最内圈磁道的周长 L :

该盘组的存储容量(非格式化容量) C : $L=\pi(18-2*5)=25.12$ inch

磁盘旋转一周的时间 t :

$$C=1000*25.12*500*20=251.2*10^7 \text{ bits}$$

数据传输率 D_r :

$$t = \frac{1}{2400} * 60 = 0.025 \text{ s}$$

平均寻址时间 T_a :

$$D_r = \frac{25.12*10^3 \text{ bits}}{0.025 \text{ s}} = 0.1256 \text{ MB/s}$$

$$T_a = \frac{0+0.2*499}{2} + \frac{0+25}{2} = 62 \text{ ms}$$

四、软磁盘存储器

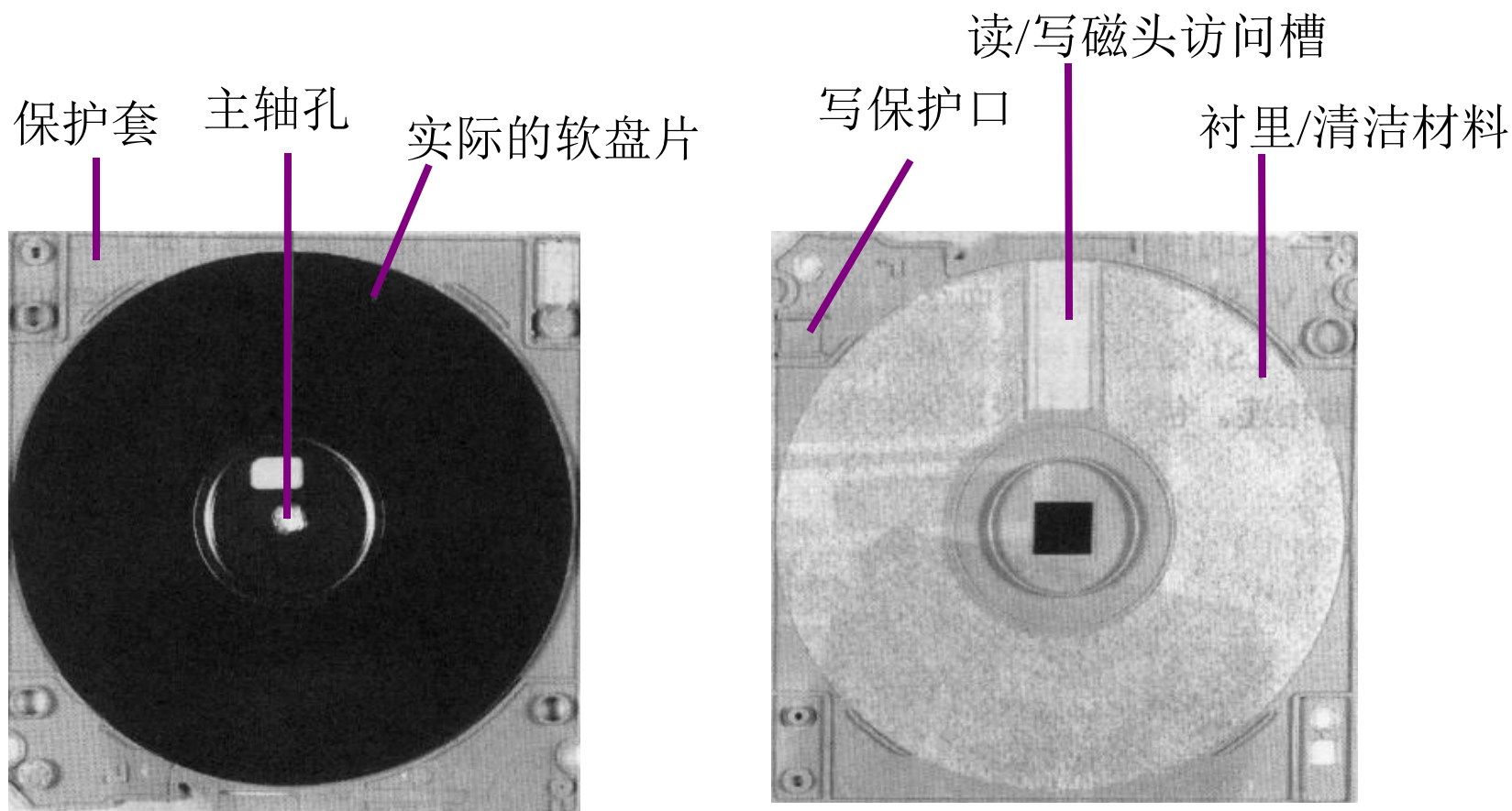
1. 概述

	硬盘	软盘
速度	高	低
磁头	固定、活动 浮动	活动 接触盘片
盘片	固定盘、盘组 大部分不可换	可换盘片
价格	高	低
环境	苛刻	



2. 软盘片

由聚酯薄膜制成



五、光盘存储器

1. 概述

采用光存储技术

利用激光写入和读出

{ 第一代光存储技术
第二代光存储技术

采用非磁性介质 不可擦写

采用磁性介质 可擦写

2. 光盘的存储原理

只读型和只写一次型 热作用（物理或化学变化）

可擦写光盘 热磁效应



2.1.5 数据校验码（续）

2. 循环冗余校验码 —— CRC码



也称 (n, k) 码, $(n=k+r)$

- 模2运算

模2加减：不考虑进位和借位（异或）

模2乘：不进位

模2除：不借位（首位为1，商1，首位为0，商0）

数据校验码（续）

2. 循环冗余校验码 —— CRC码



● CRC编码方法

(1) 已知信息:

$$M(x) = m_{k-1}x^{k-1} + m_{k-2}x^{k-2} + \dots + m_1x + m_0$$

其中, $m_i=0$ 或 1

(2) 选取生成多项式: $G(x) = x^r + g_{r-1}x^{r-1} + \dots + g_1x + 1$

其中, $g_i=0$ 或 1 , 选取条件: $n \leq 2^r - 1$

(3) $M(x)$ 左移 r 位: $M(x) \cdot x^r$

(4) 求余数 $R(x)$: $M(x) \cdot x^r / G(x)$

(5) CRC编码: $M(x) \cdot x^r + R(x)$

2.1.5 数据校验码（续）

CRC码举例：已知 $M(x)=1100=x^3+x^2$ ，

$G(x)=x^3+x+1=1011$ ，求CRC码

$$\begin{array}{r} 1110 \\ 1011 \overline{) 1100000} \\ \underline{1011} \\ 1110 \\ \underline{1011} \\ 1010 \\ \underline{1011} \\ 0010 \\ \underline{0000} \\ 010 \end{array}$$

CRC码为：

1100010

2.1.5 数据校验码（续）

- 生成多项式 $G(x)$ 满足条件:
 - 生成多项式的最高位和最低位必须为1
 - 任何一位发生错误都应使余数不为零
 - 不同位发生错误应当使余数不同
 - 对余数继续作模2除,应使余数循环
- CRC校验中广泛应用的标准多项式有4种, 即:
 - $\text{CRC-CCITT} = X^{16} + X^{12} + X^5 + 1$
 - $\text{CRC-12} = X^{12} + X^{11} + X^3 + X^2 + X + 1$
 - $\text{CRC-16(ANSI)} = X^{16} + X^{15} + X^2 + 1$
 - $\text{CRC-32} = X^{32} + X^{26} + X^{23} + X^{22} + X^{16} + X^{12} + X^{11} + X^{10} + X^8 + X^7 + X^5 + X^4 + X + 1$

2.1.5 数据校验码（续）

●CRC的译码与纠错

将接收到的信息用 $G(x)$ 去除，若余数为0，则无错，若余数不为0，则某一位出错，不同位数出错余数不同（由生成多项式决定）

	A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	A ₇	余数	出错位
正确	1	1	0	0	0	1	0	000	无
错误	1	1	0	0	0	1	1	001	7
	1	1	0	0	0	0	0	010	6
	1	1	0	0	1	1	0	100	5
	1	1	0	1	0	1	0	011	4
	1	1	1	0	0	1	0	110	3
	1	0	0	0	0	1	0	111	2
	0	1	0	0	0	1	0	101	1

$G=1011$ 时