

第四章 存储器

4.1 概述

4.2 主存储器

4.3 高速缓冲存储器

4.4 辅助存储器

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

(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

非
易
失

2. 按存取方式分类

(1) 存取时间与物理地址无关（随机访问）

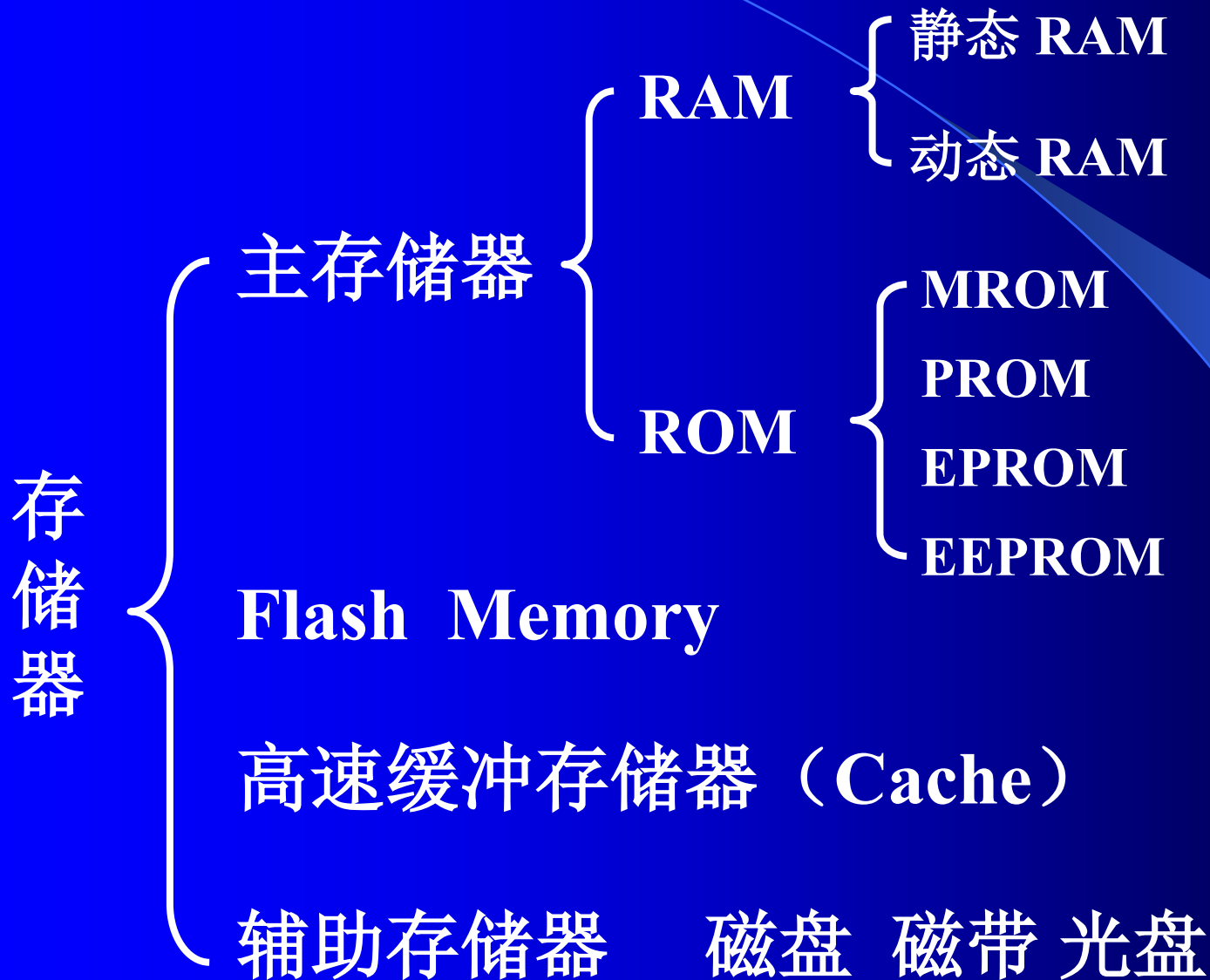
- 随机存储器 在程序的执行过程中 可读 可 写
- 只读存储器 在程序的执行过程中 只读

(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

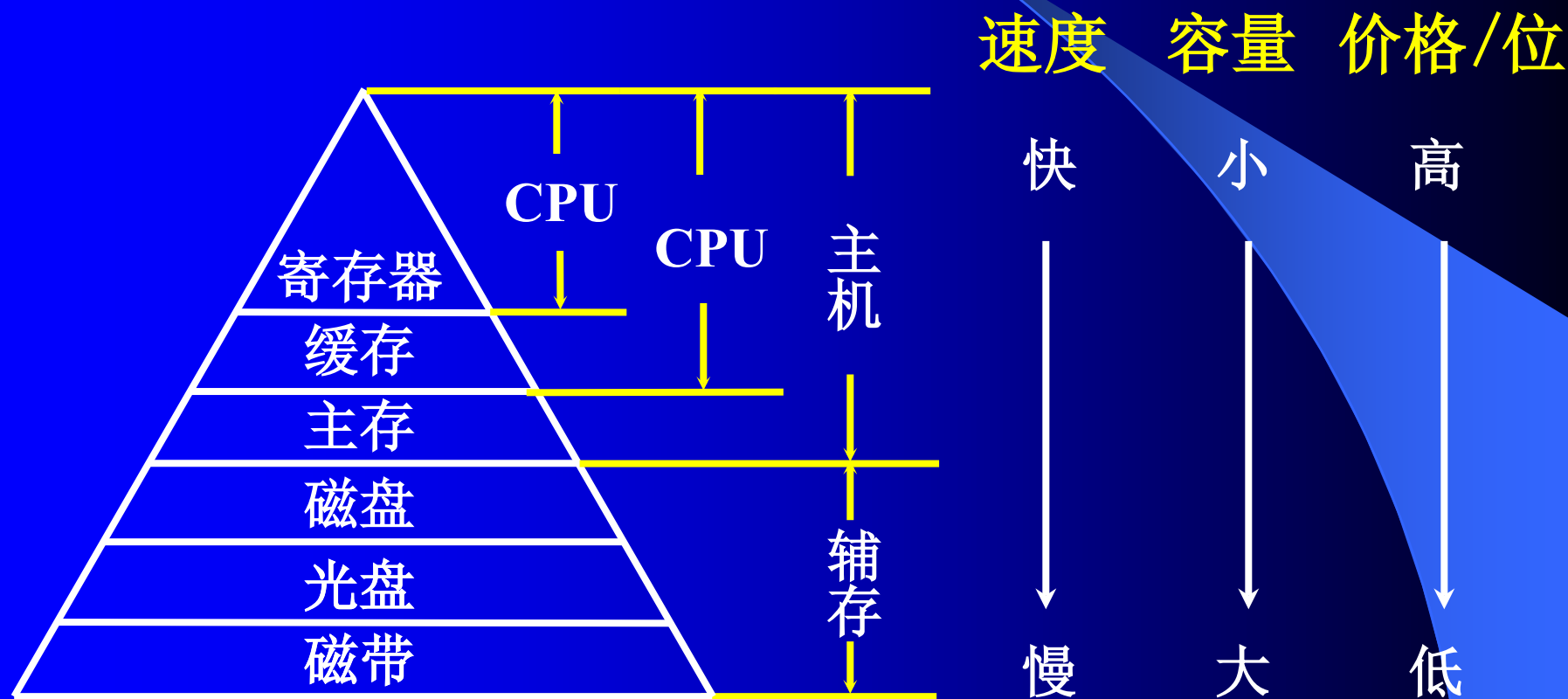
3. 按在计算机中的作用分类

4.1

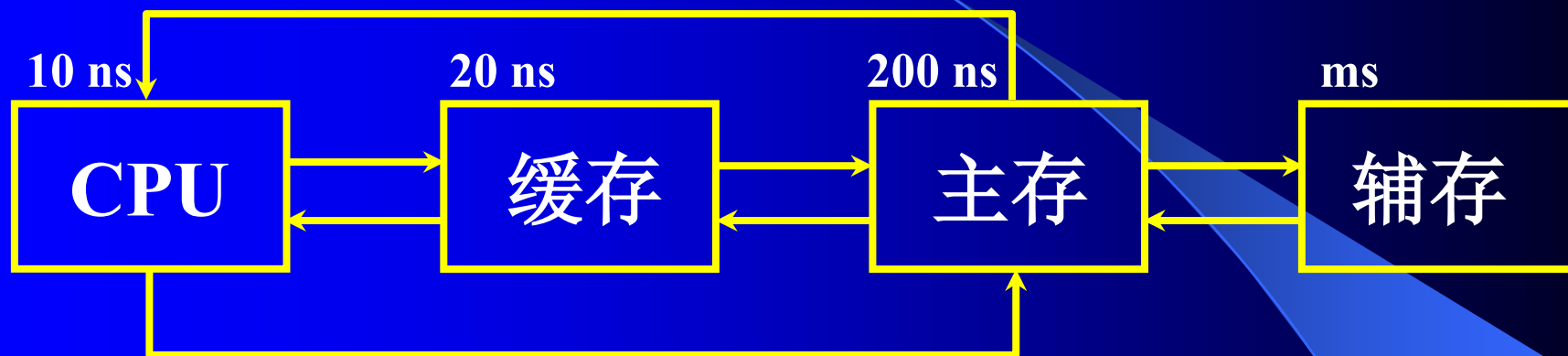


二、存储器的层次结构

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

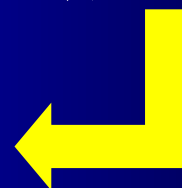
虚拟存储器

实地址

虚地址

物理地址

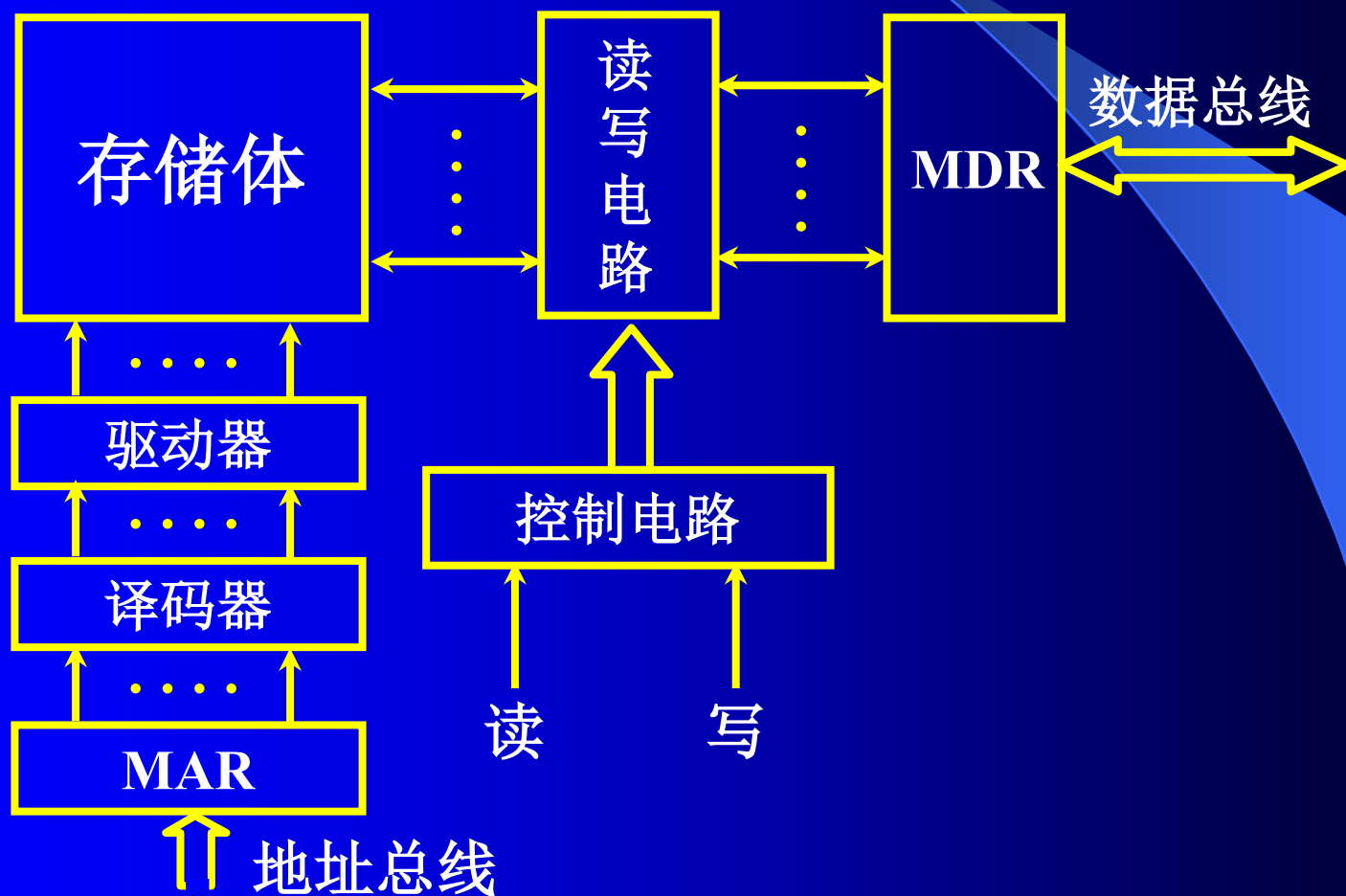
逻辑地址



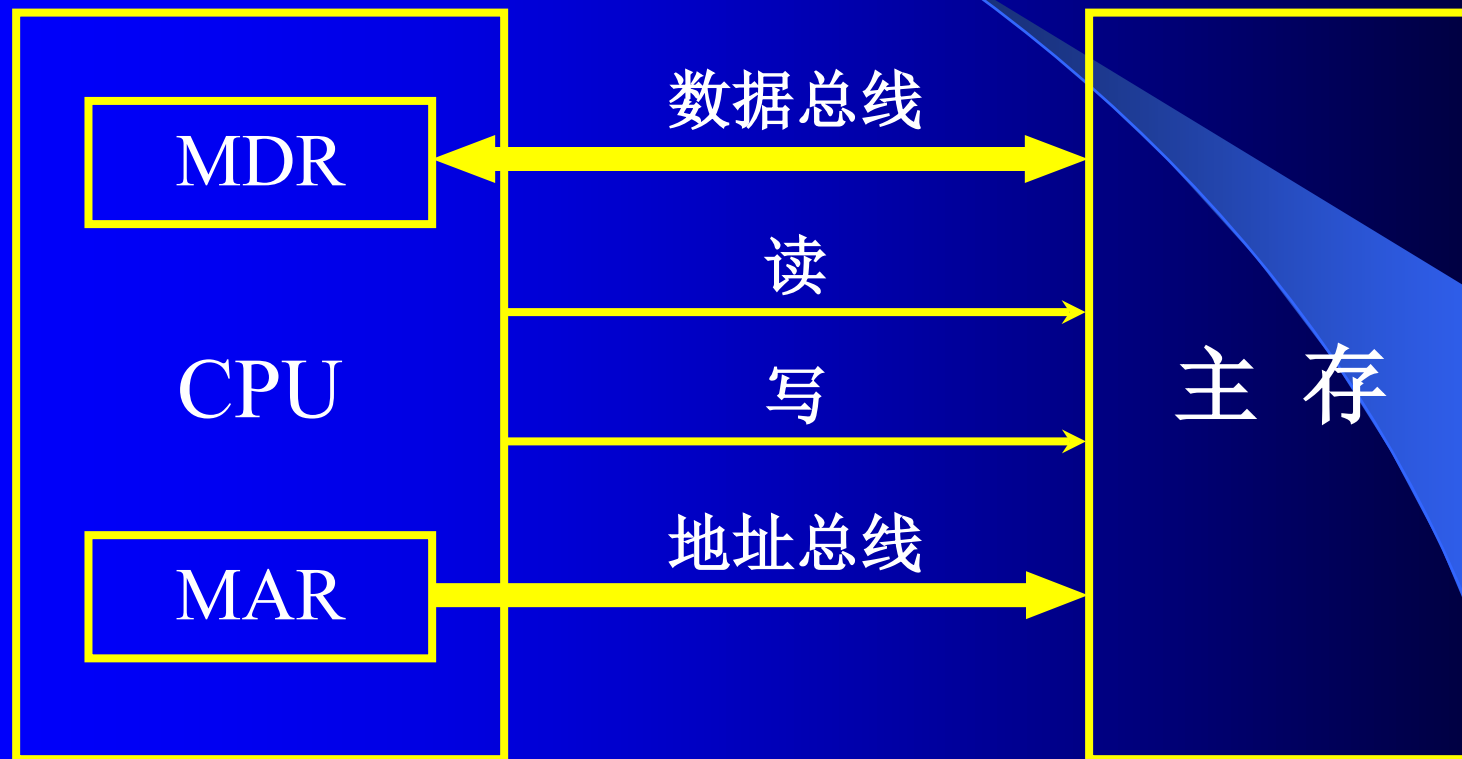
4.2 主存储器

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

若字长为 16 位

若字长为 32 位

按 字节 寻址 $2^{24} = 16 \text{ M}$

按 字 寻址 8 M

按 字 寻址 4 M

4. 主存的技术指标

4.2

(1) 存储容量 主存 存放二进制代码的总数量

(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

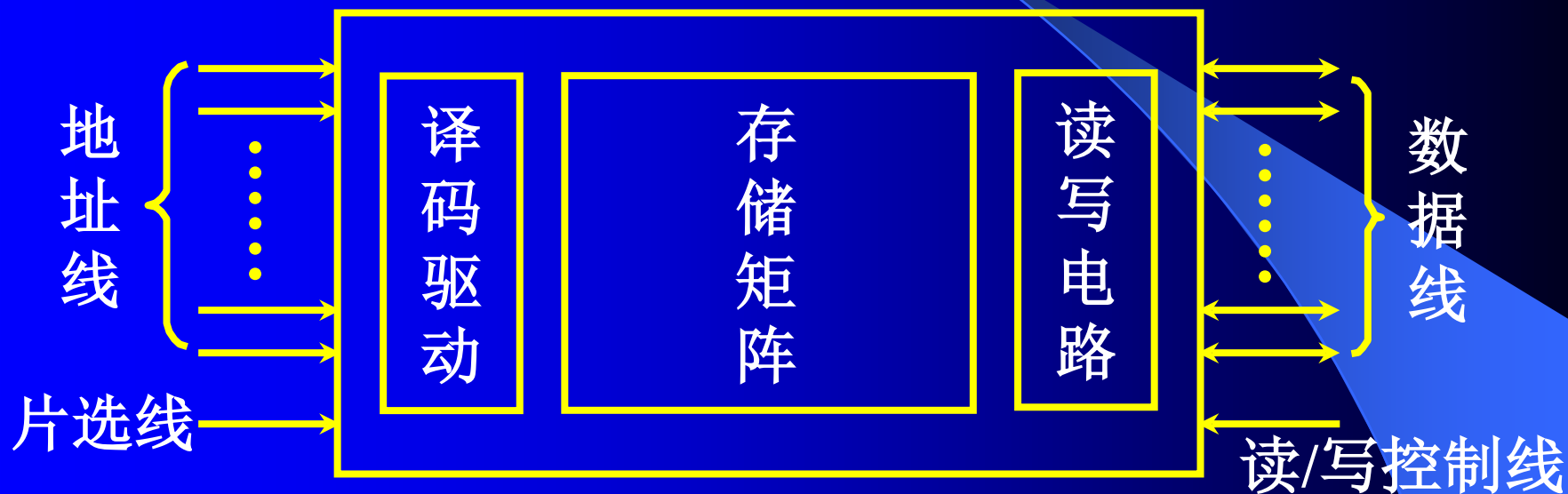
(读或写) 所需的 最小间隔时间

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

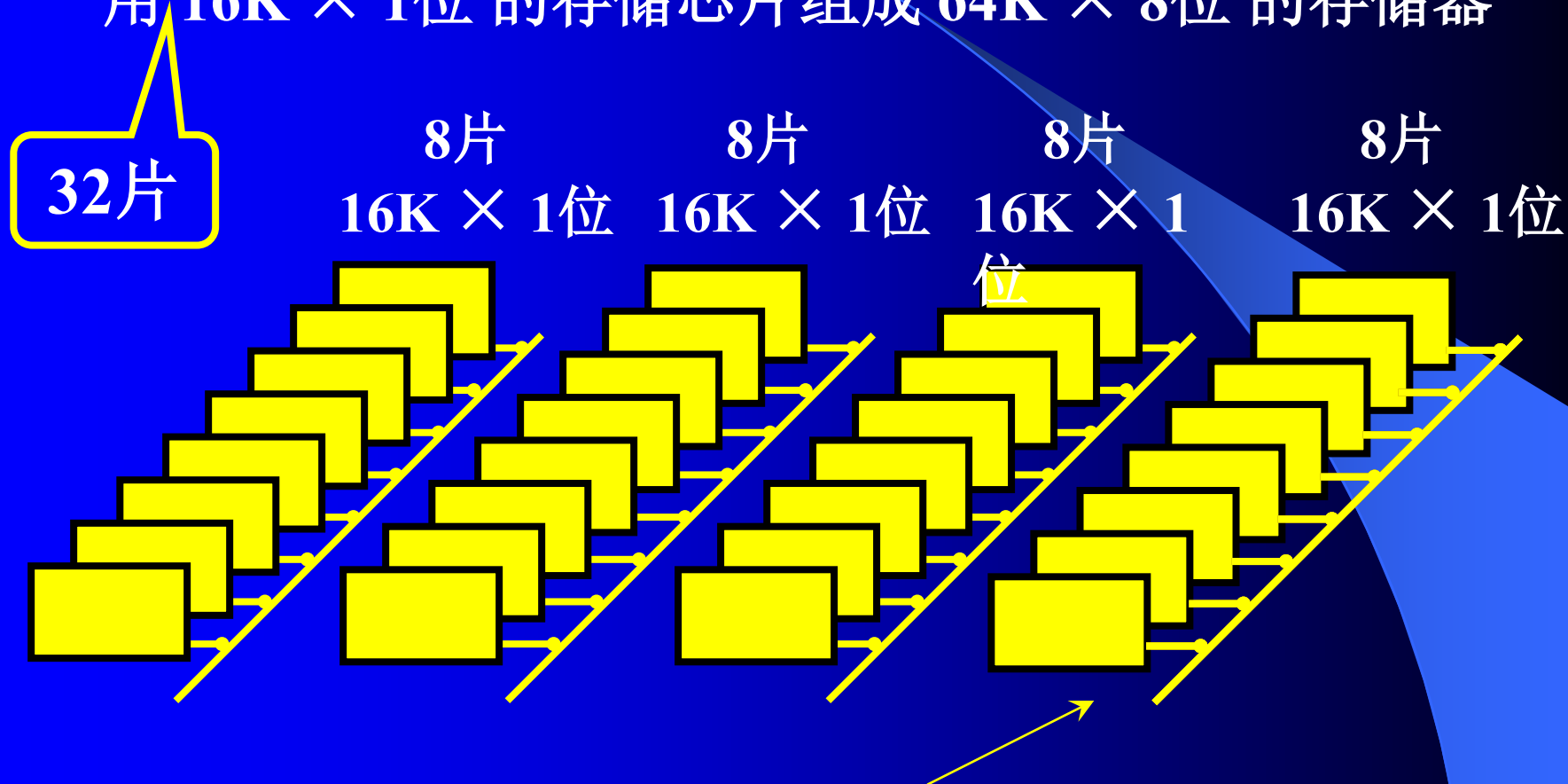
1. 半导体存储芯片的基本结构



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

存储芯片片选线的作用

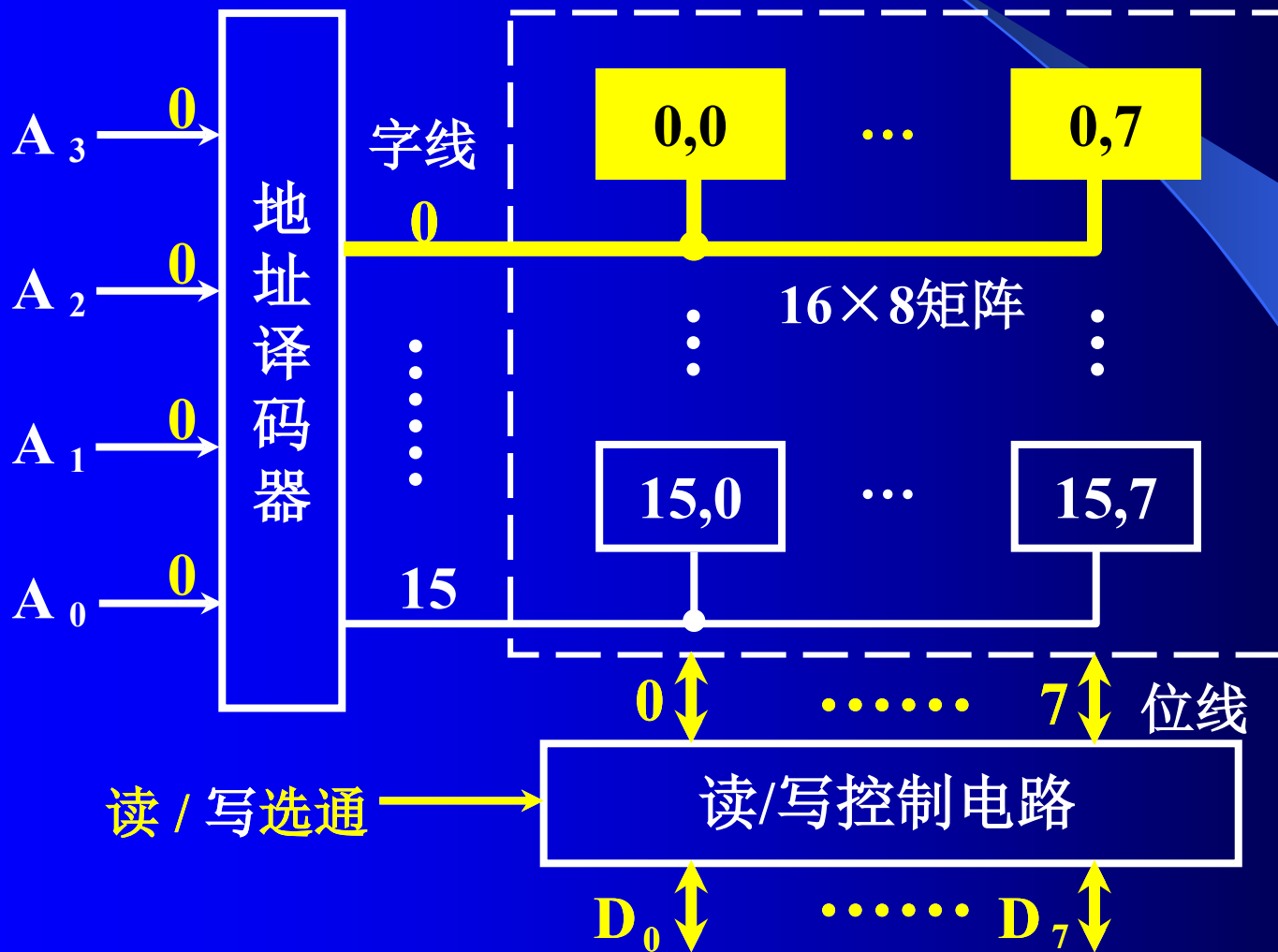
用 $16\text{K} \times 1$ 位的存储芯片组成 $64\text{K} \times 8$ 位的存储器



当地址为 65 535 时，此 8 片的片选有效

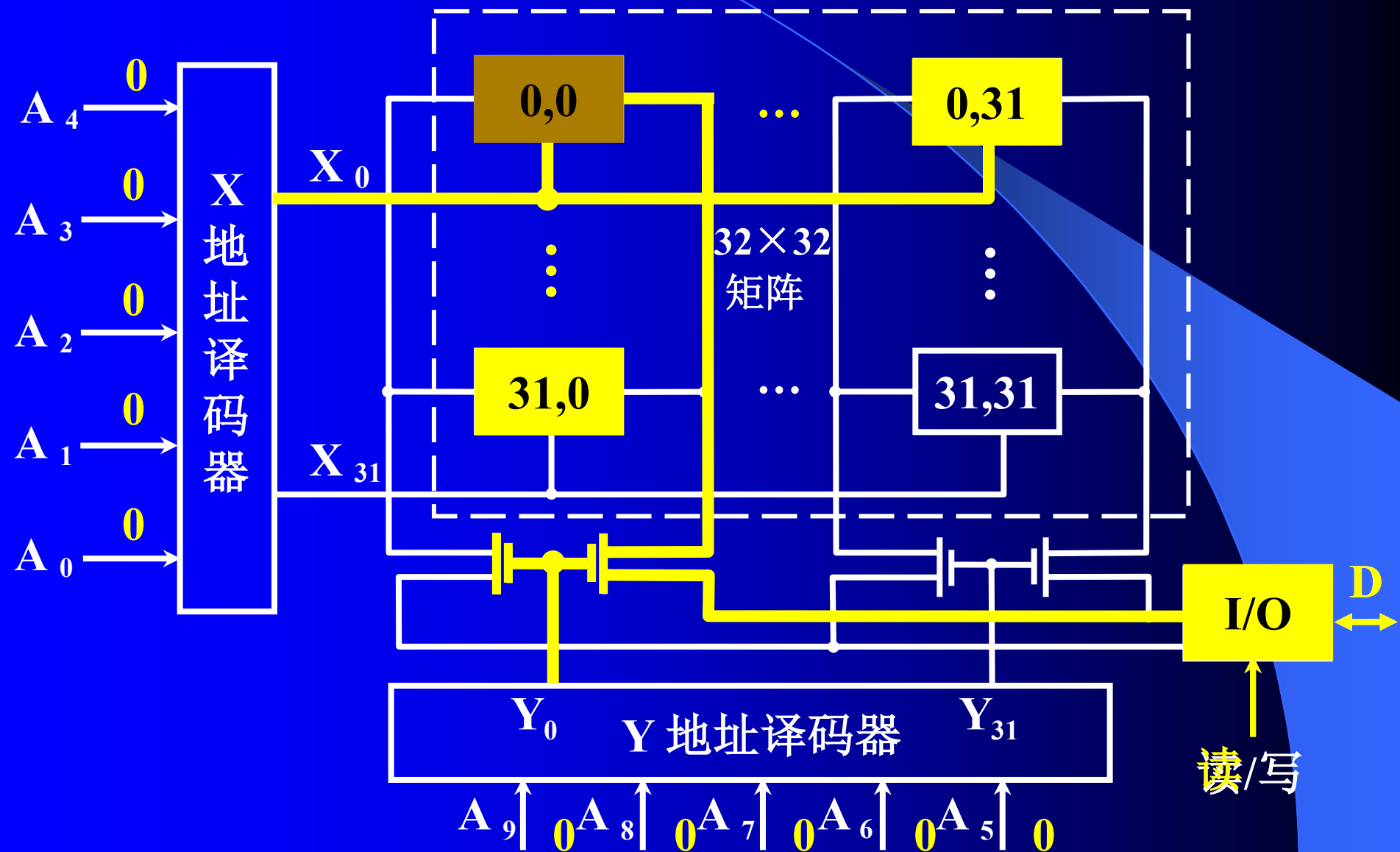
2. 半导体存储芯片的译码驱动方式 4.2

(1) 线选法



(2) 重合法

4.2

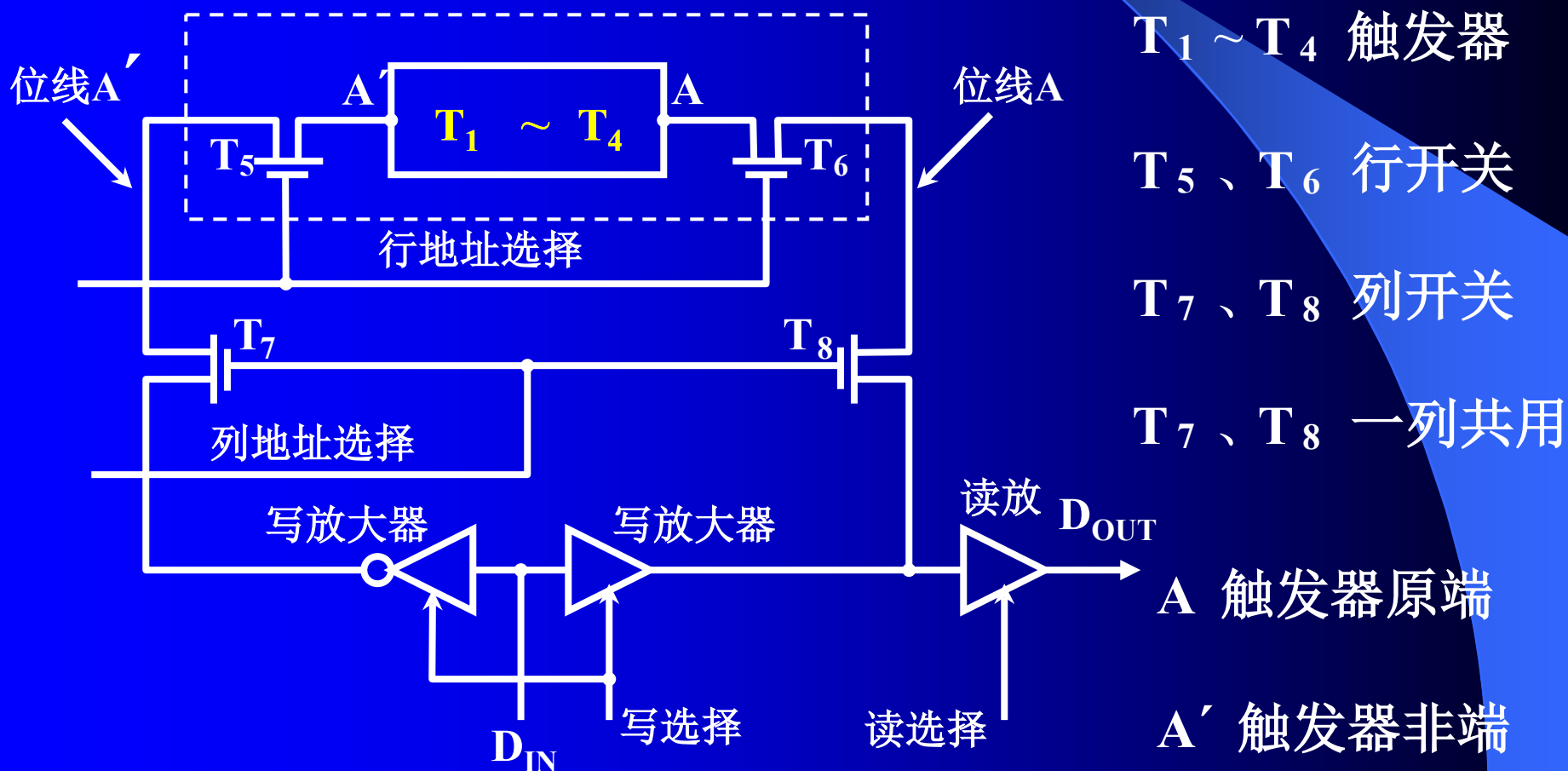


三、随机存取存储器 (RAM)

4.2

1. 静态 RAM (SRAM)

(1) 静态 RAM 基本电路



T₁ ~ T₄ 触发器

T₅、T₆ 行开关

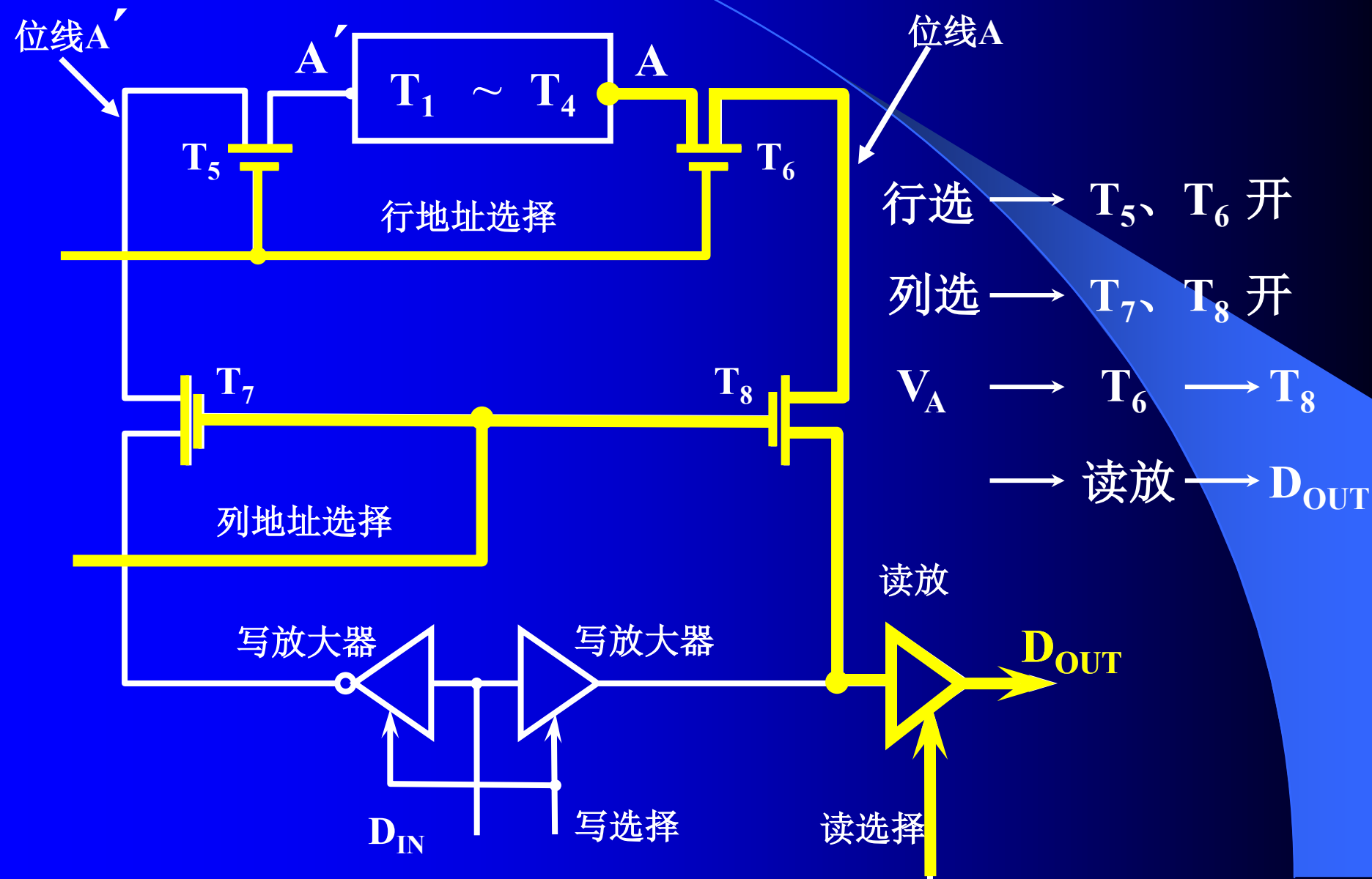
T₇、T₈ 列开关

T₇、T₈ 一列共用

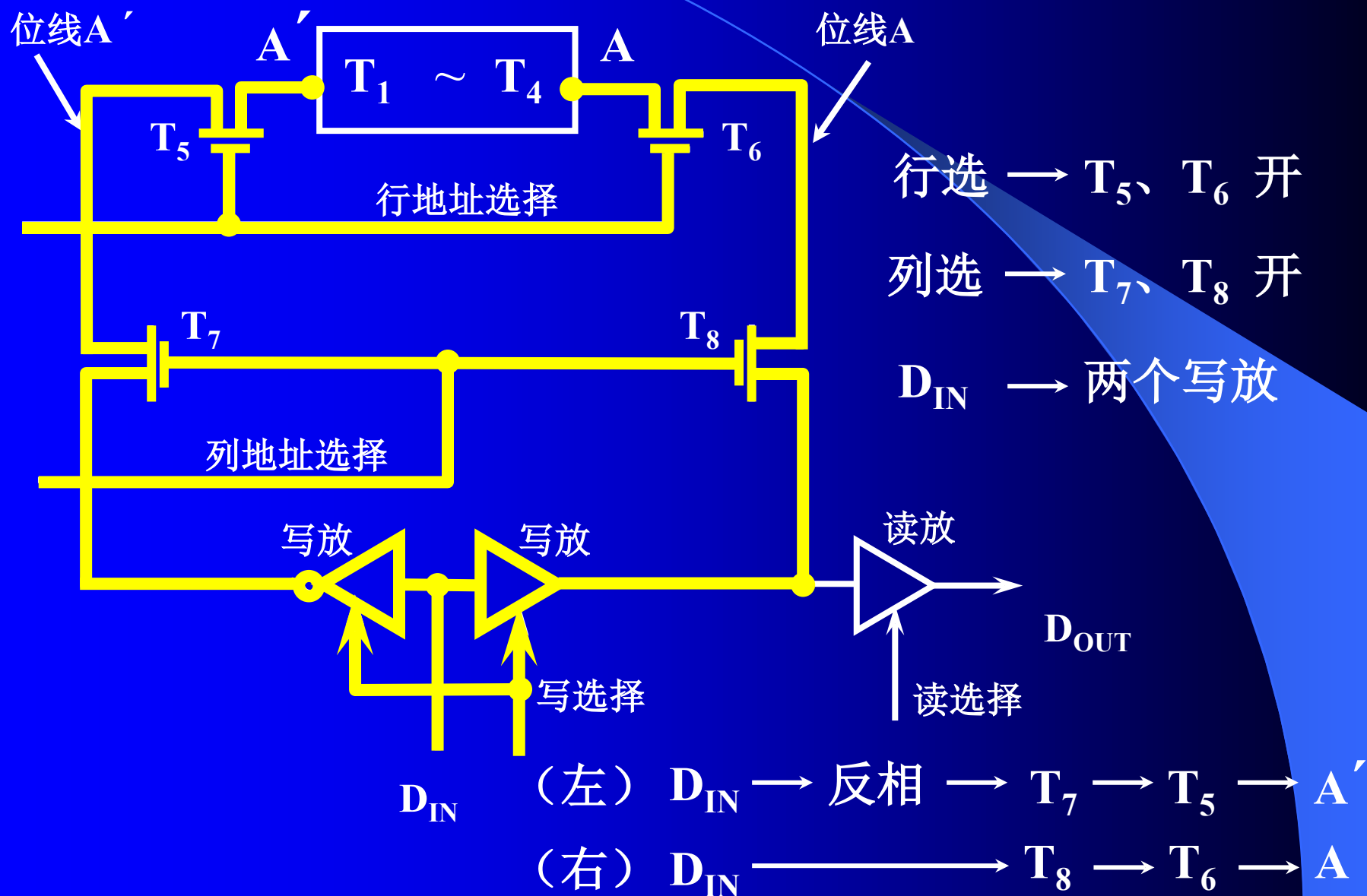
A 触发器原端

A' 触发器非端

① 静态 RAM 基本电路的 读 操作

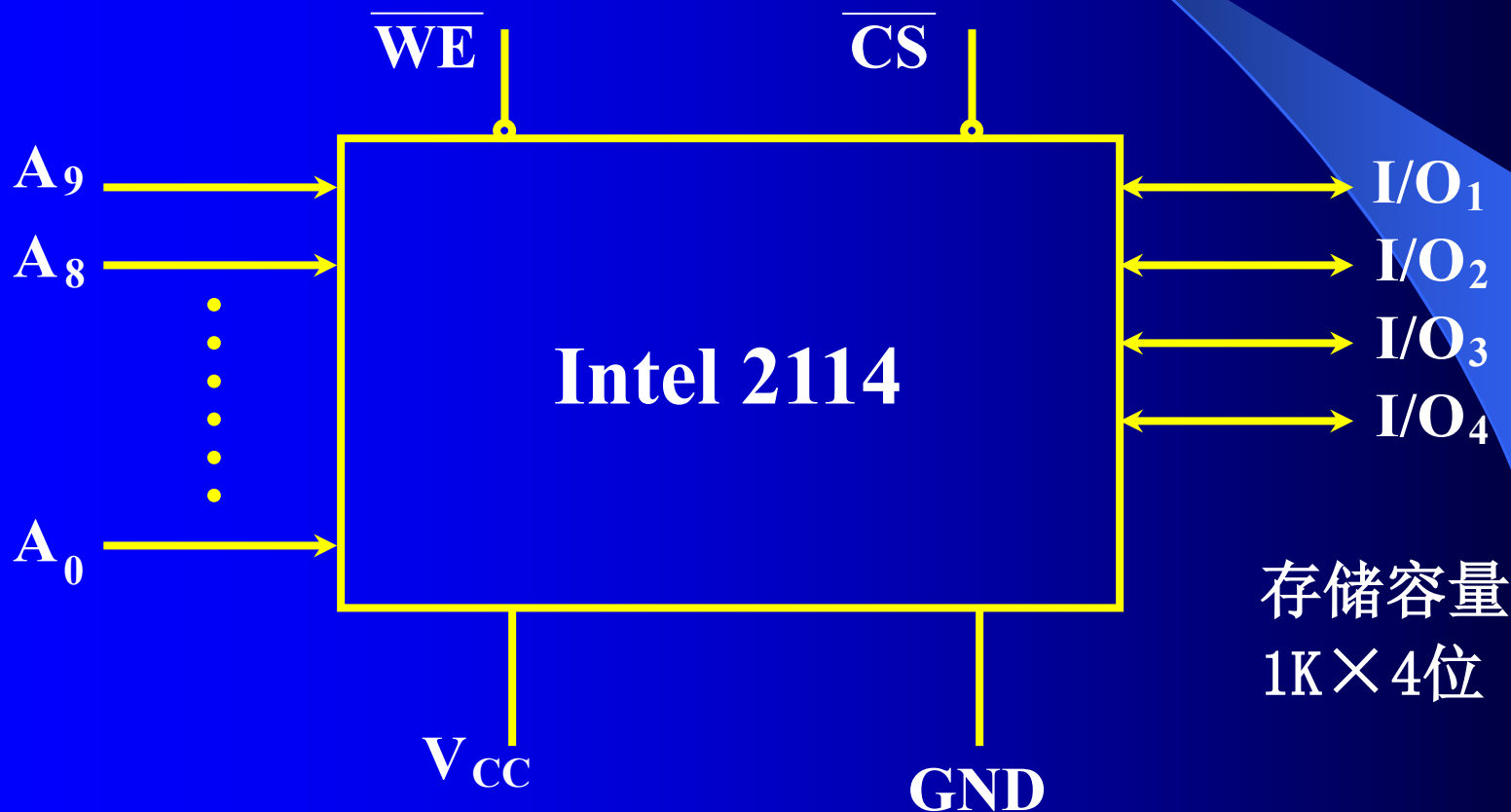


② 静态 RAM 基本电路的 写 操作

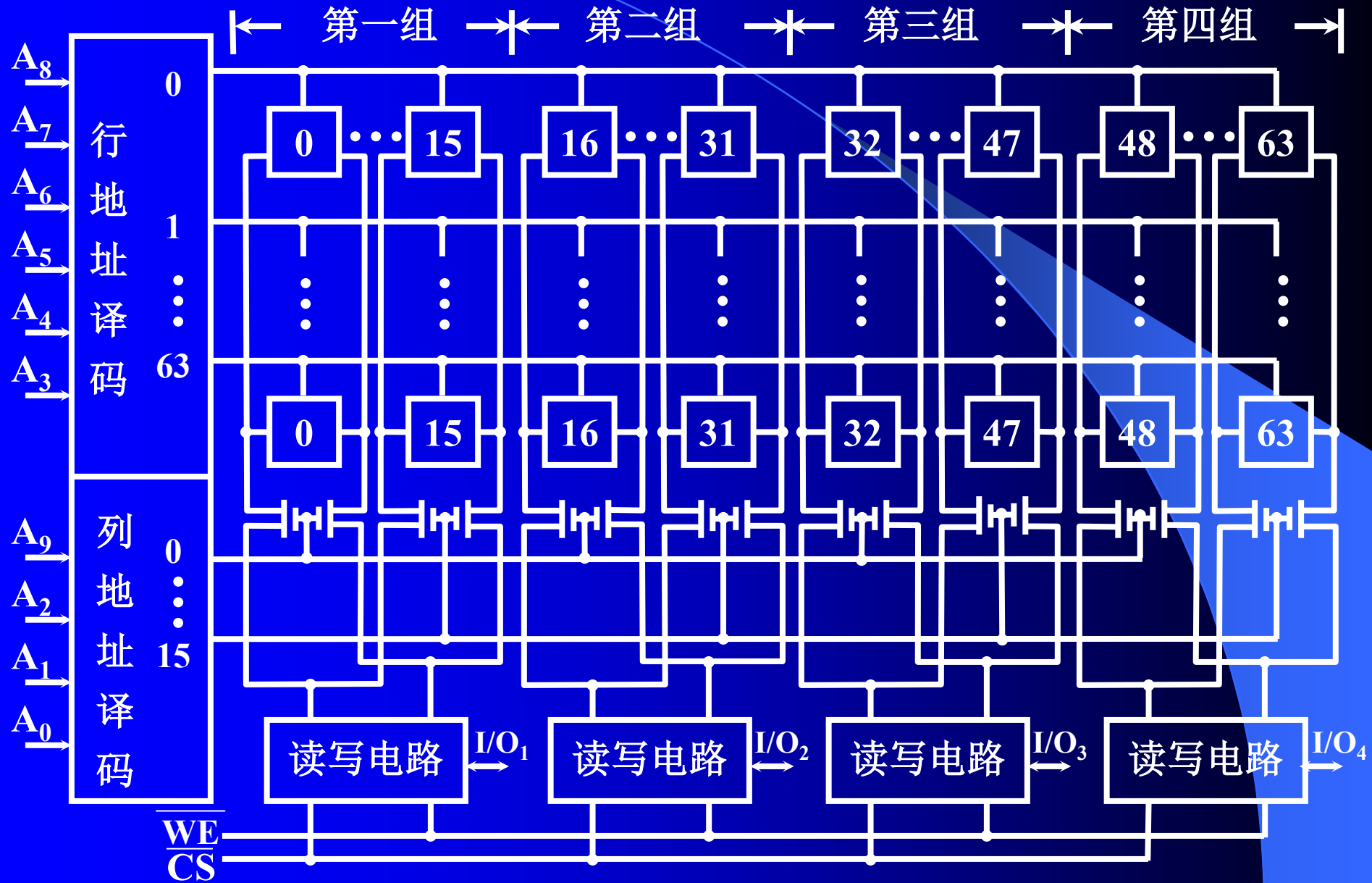


(2) 静态 RAM 芯片举例

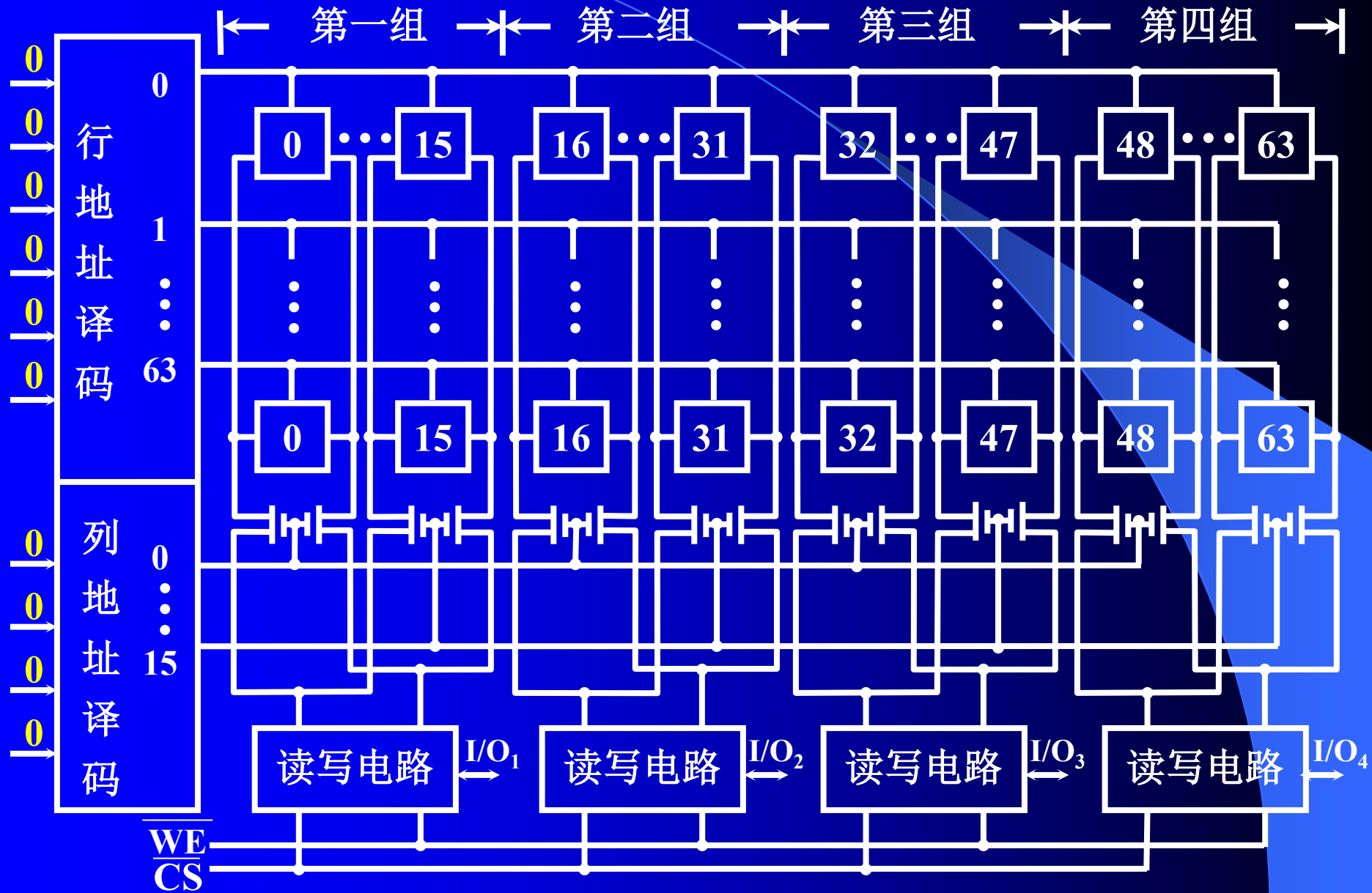
① Intel 2114 外特性



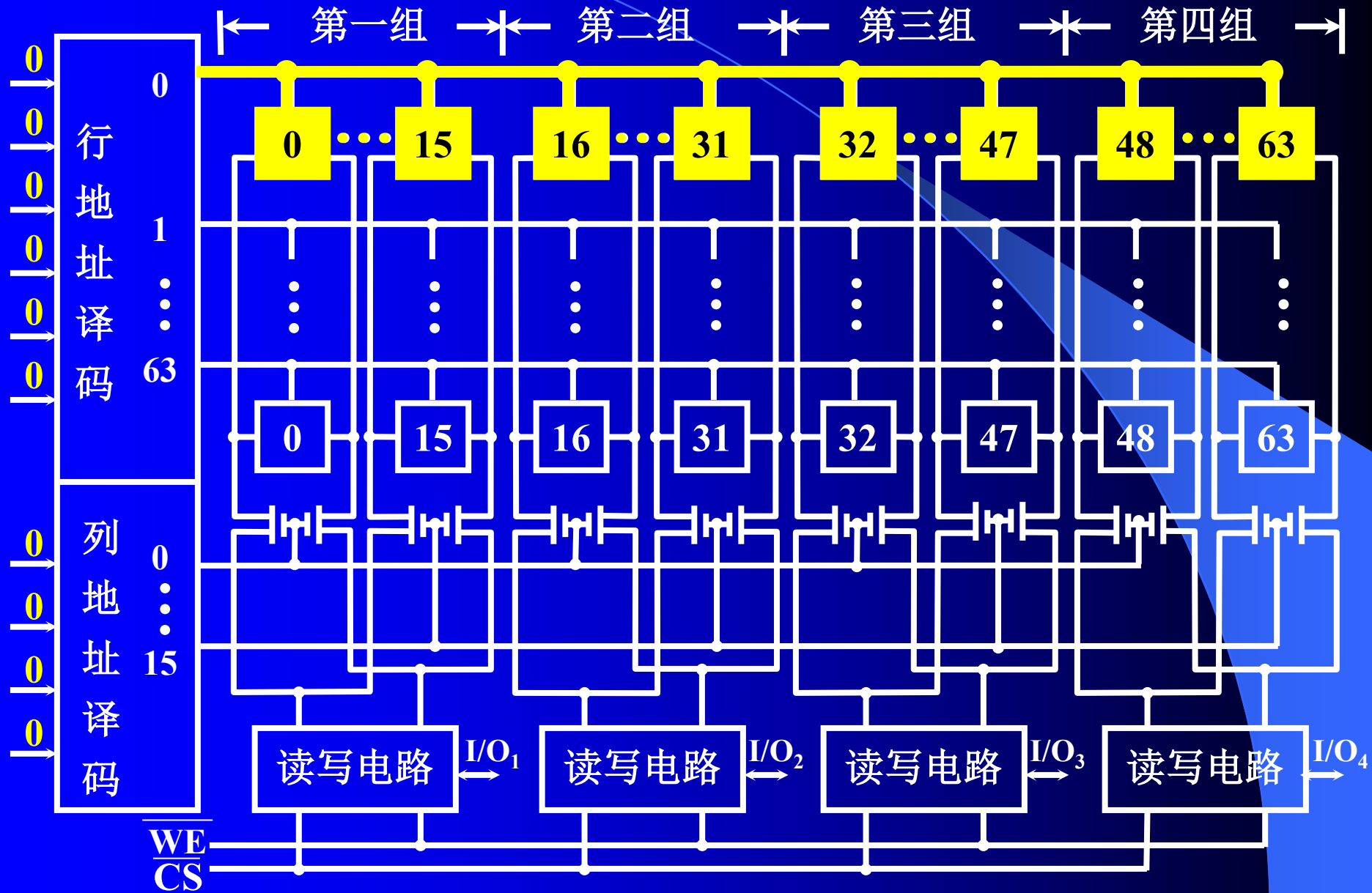
② Intel 2114 RAM 矩阵 (64 × 64) 读 4.2



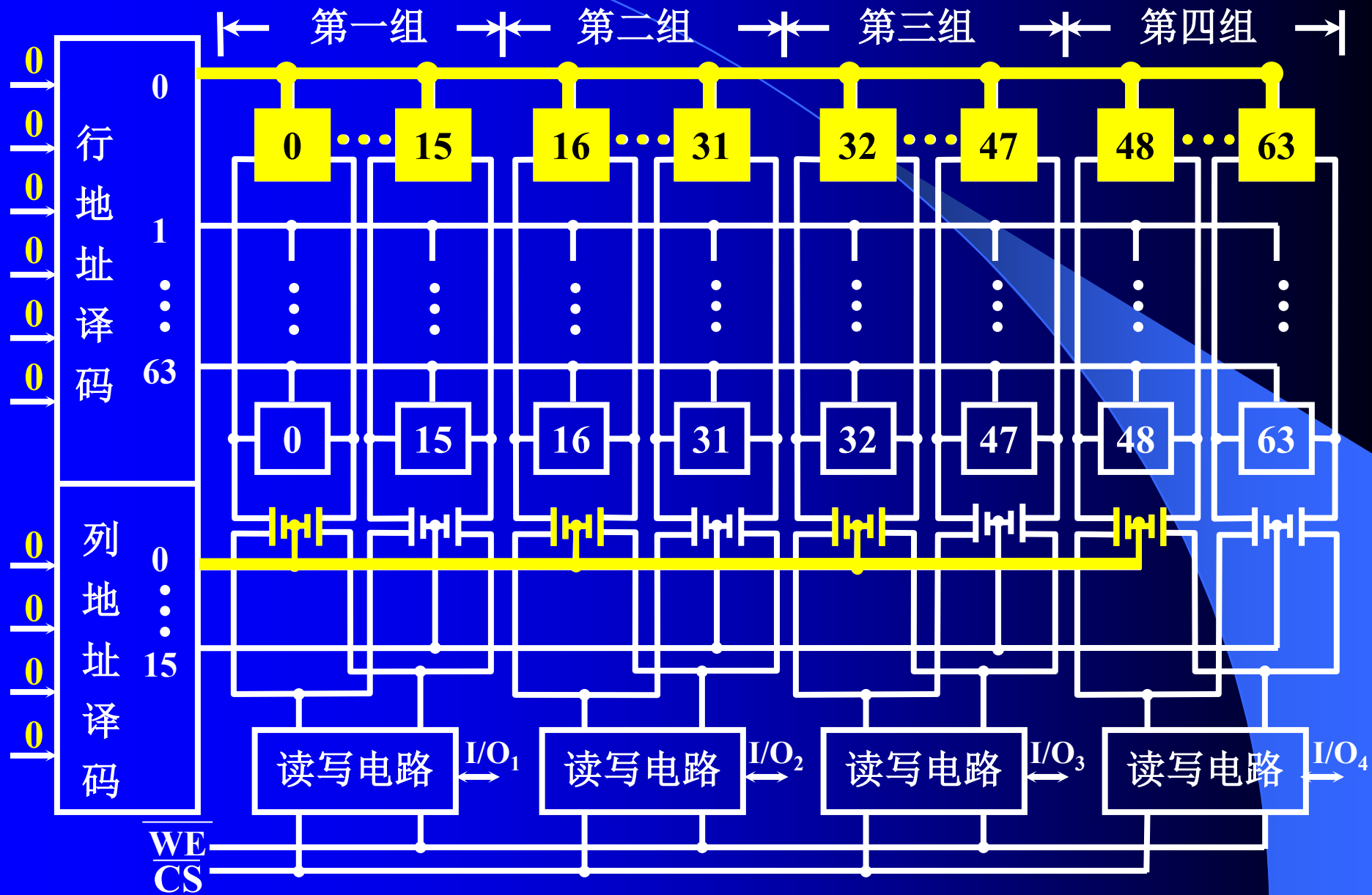
② Intel 2114 RAM 矩阵 (64×64) 读 4.2



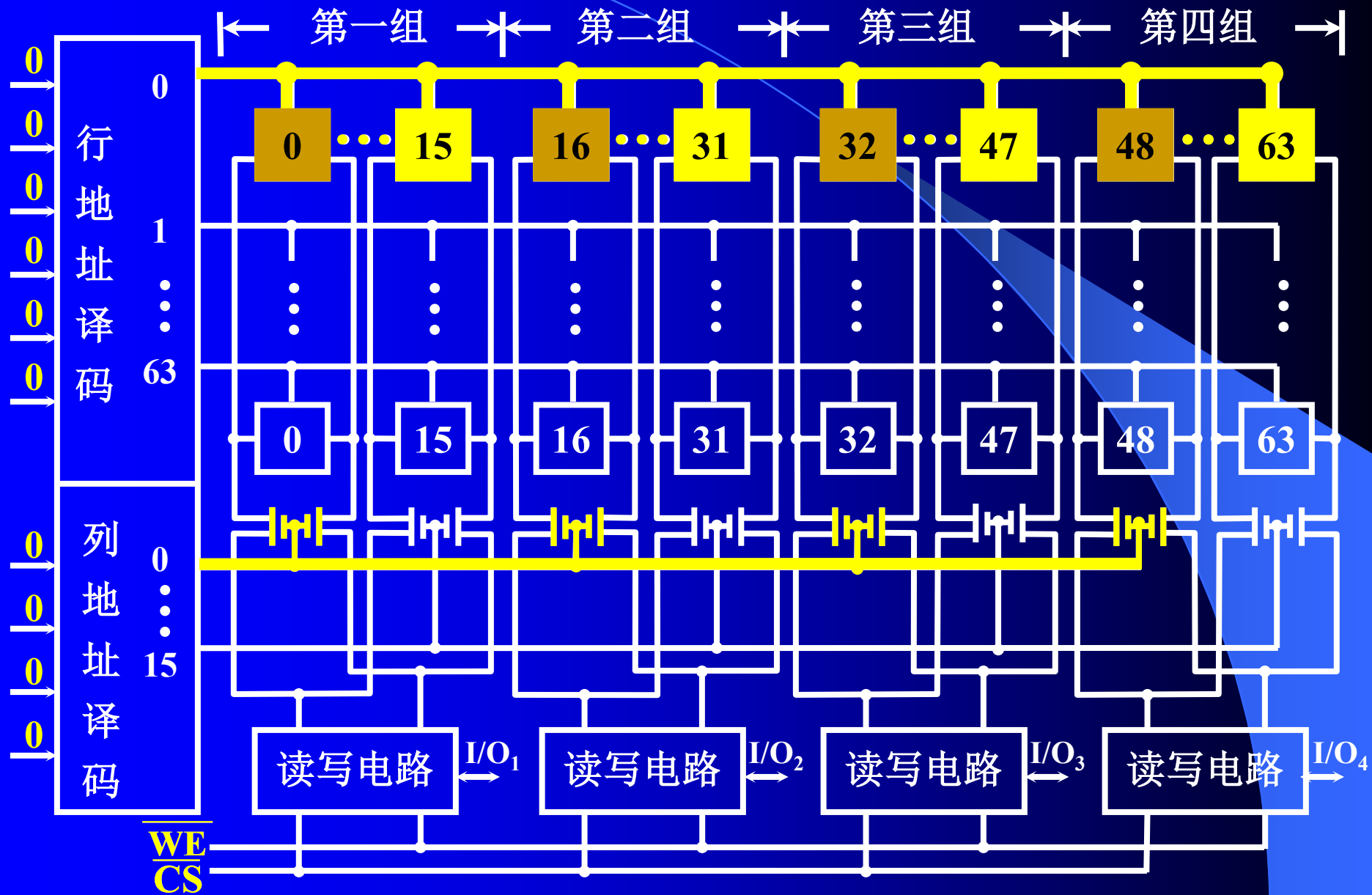
② Intel 2114 RAM 矩阵 (64×64) 读 4.2



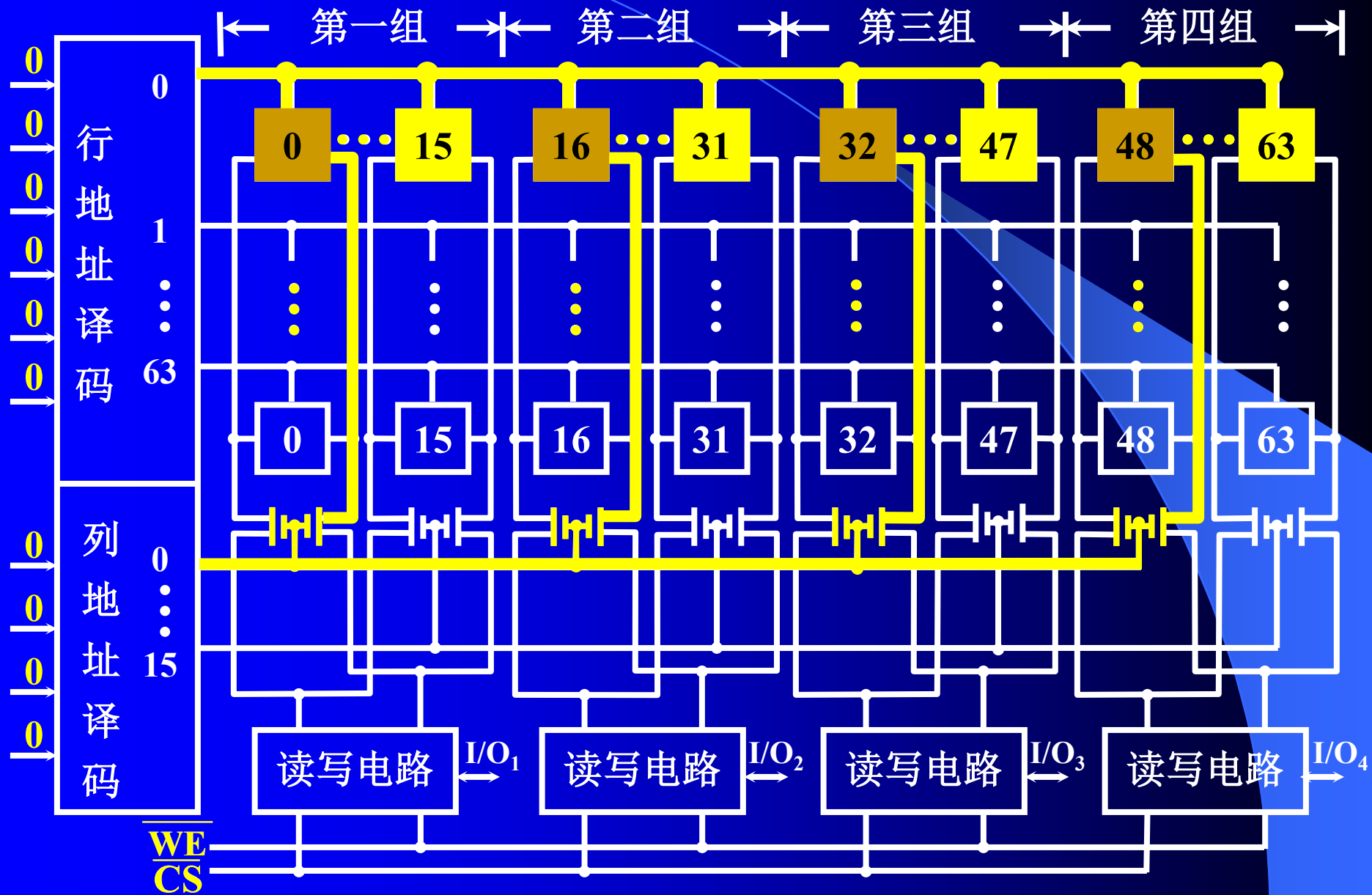
② Intel 2114 RAM 矩阵 (64×64) 读 4.2



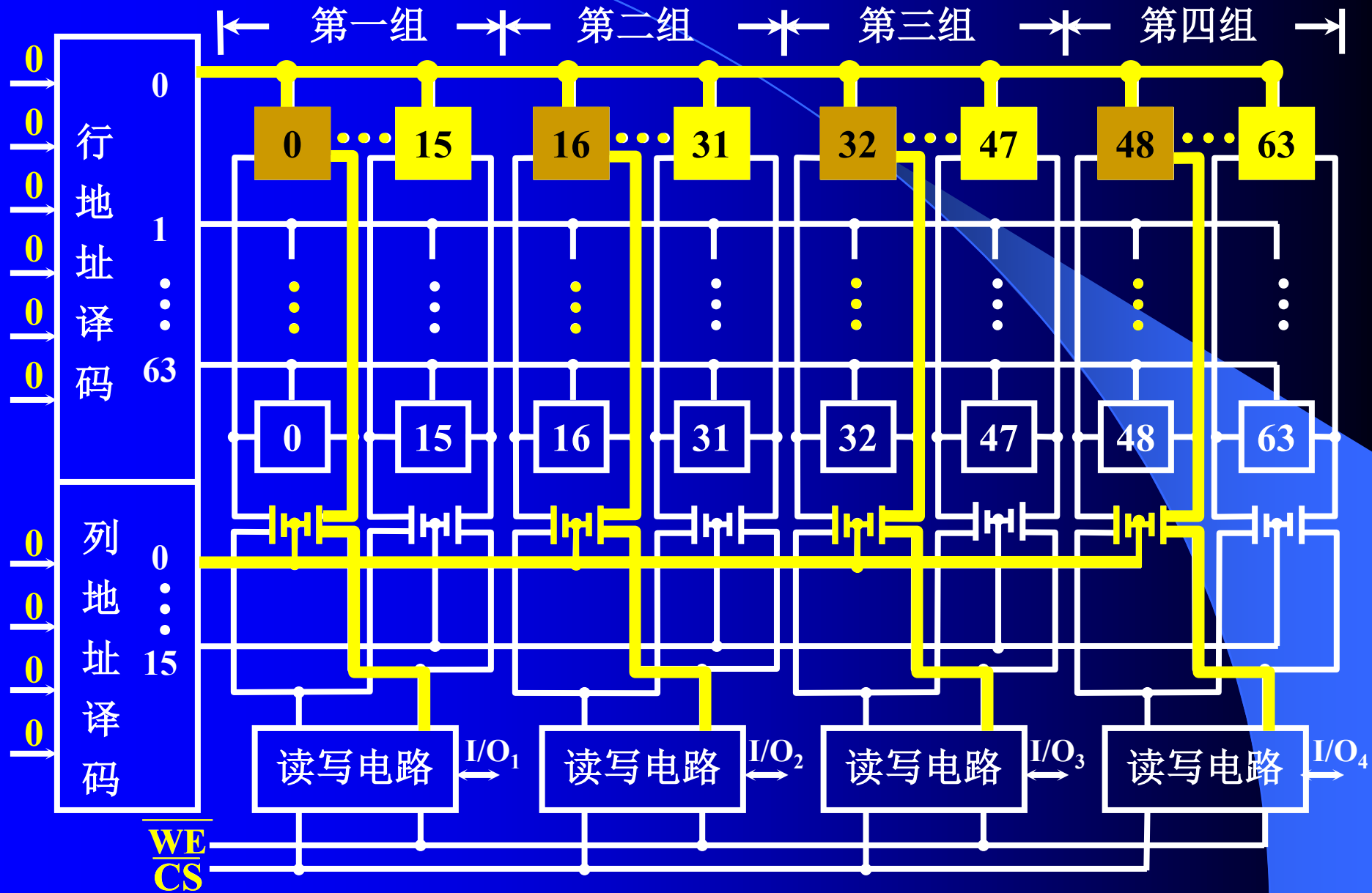
② Intel 2114 RAM 矩阵 (64×64) 读 4.2



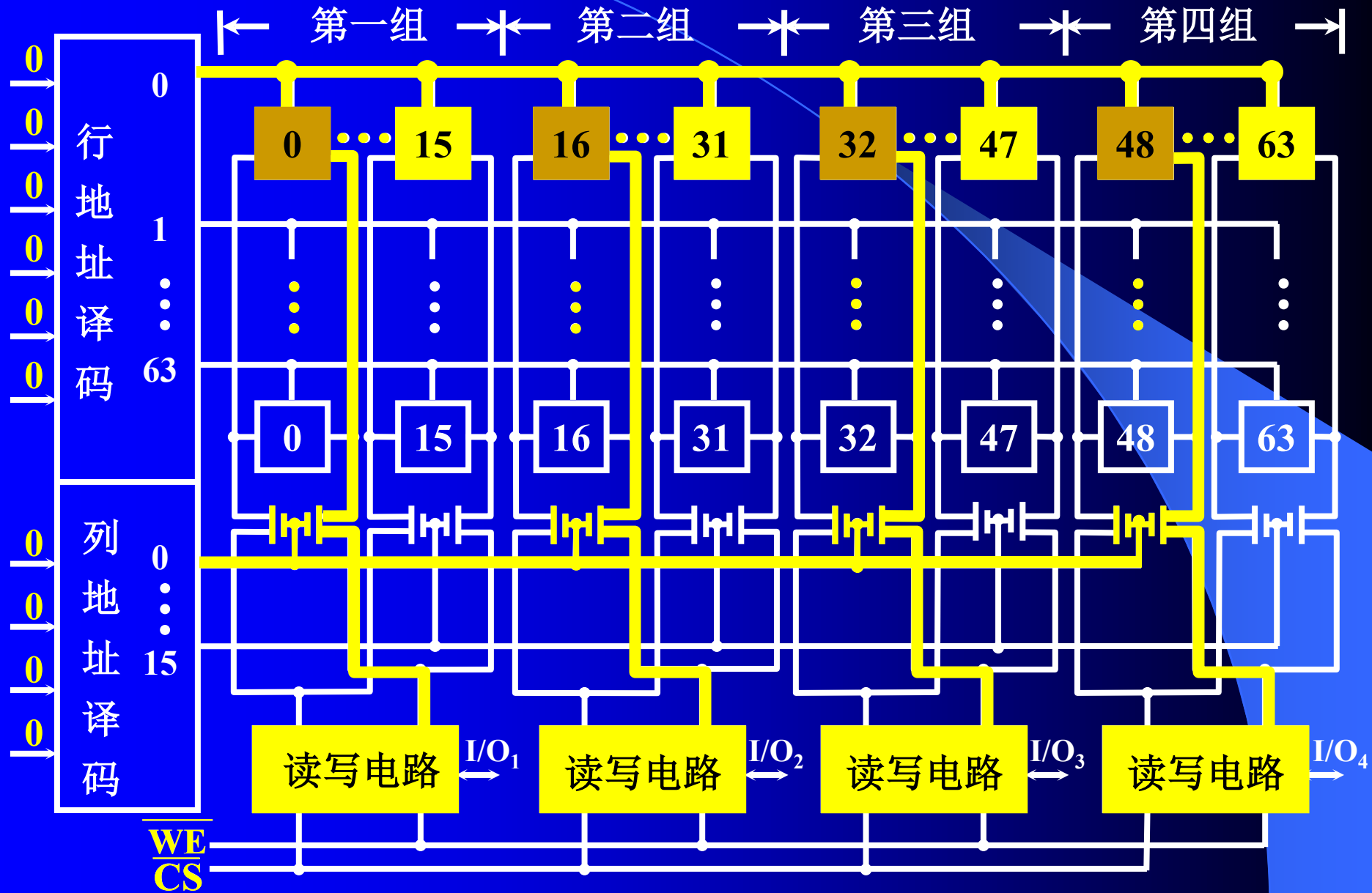
② Intel 2114 RAM 矩阵 (64×64) 读 4.2



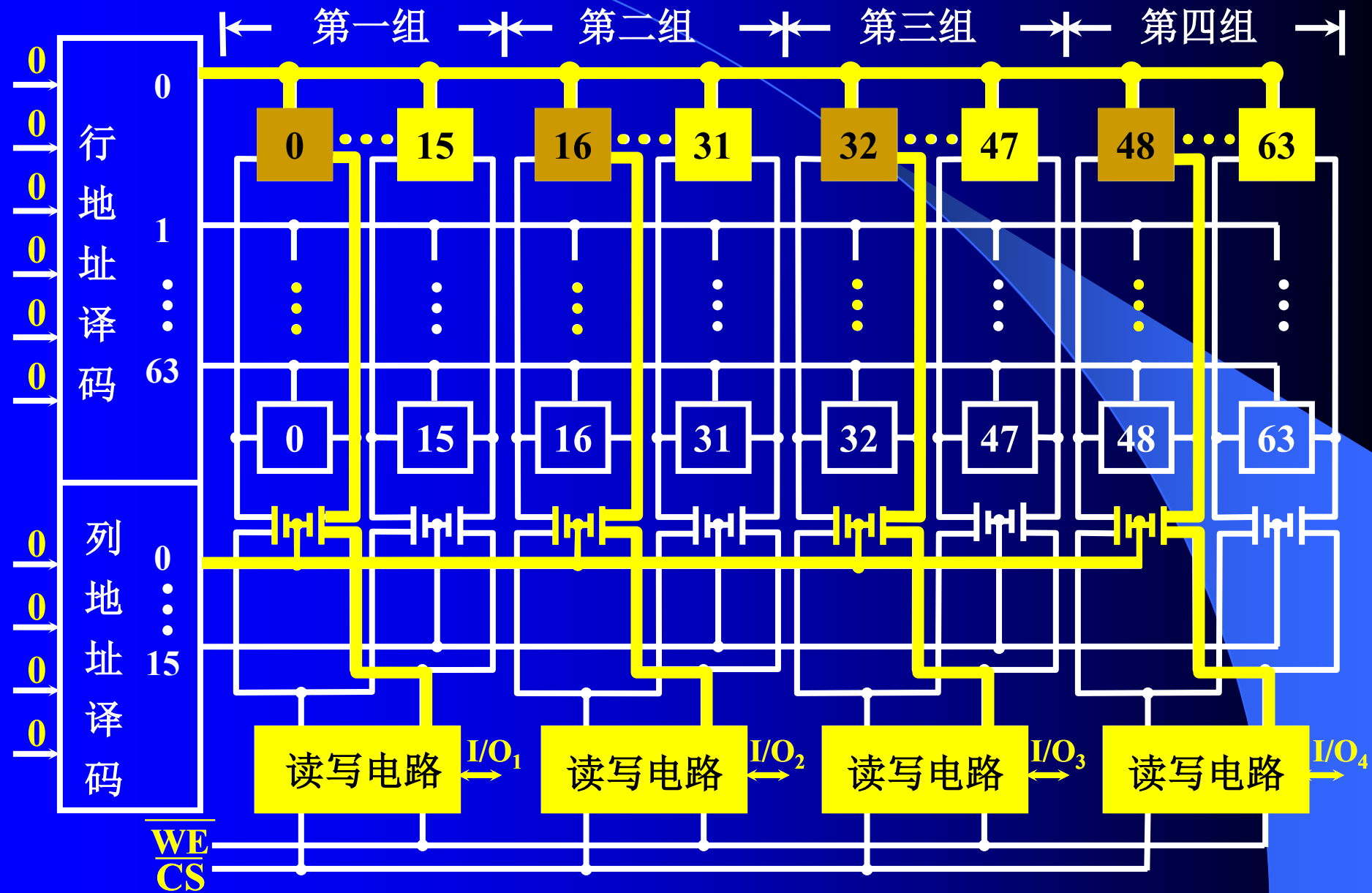
② Intel 2114 RAM 矩阵 (64×64) 读 4.2



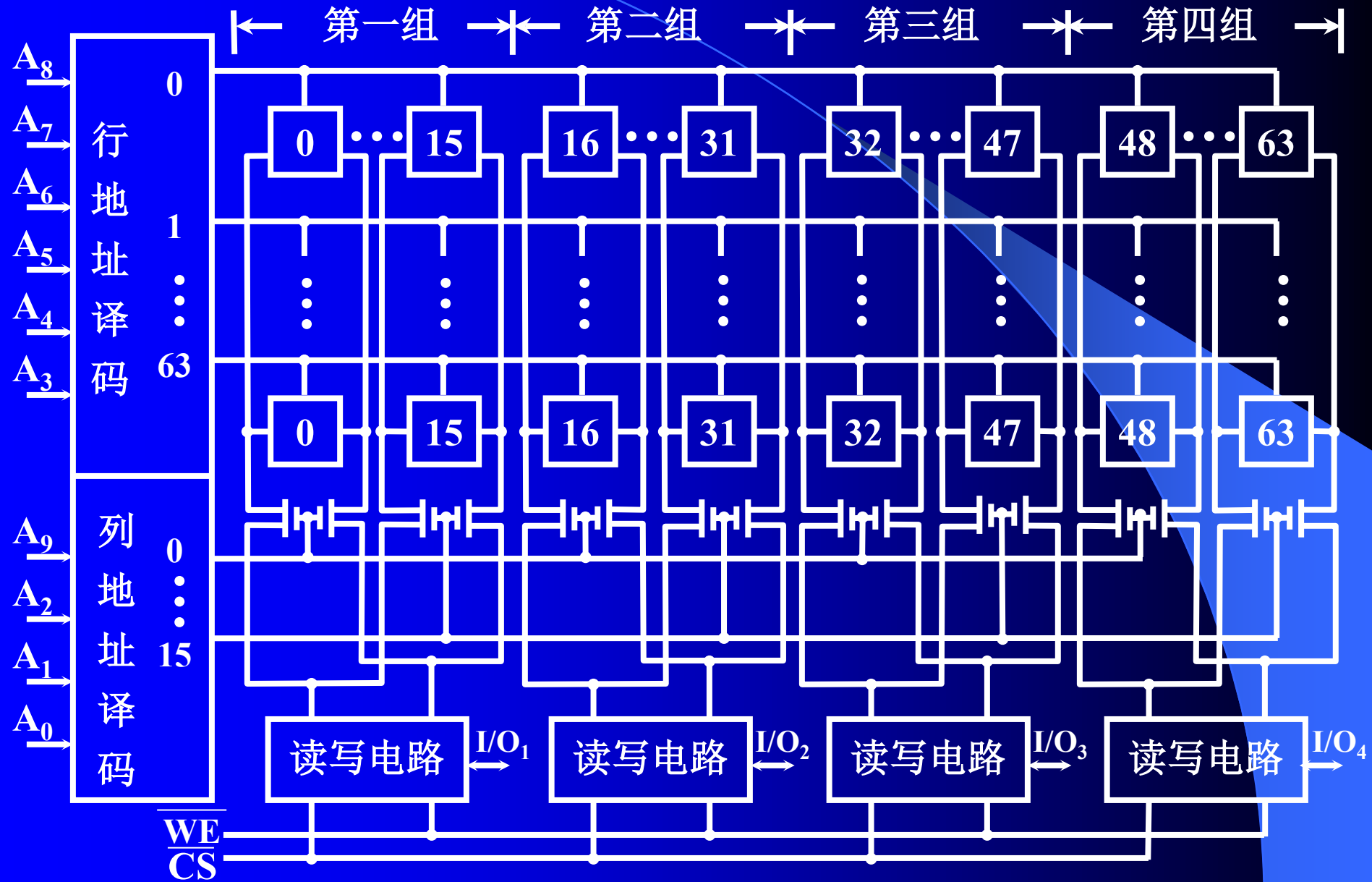
② Intel 2114 RAM 矩阵 (64×64) 读 4.2



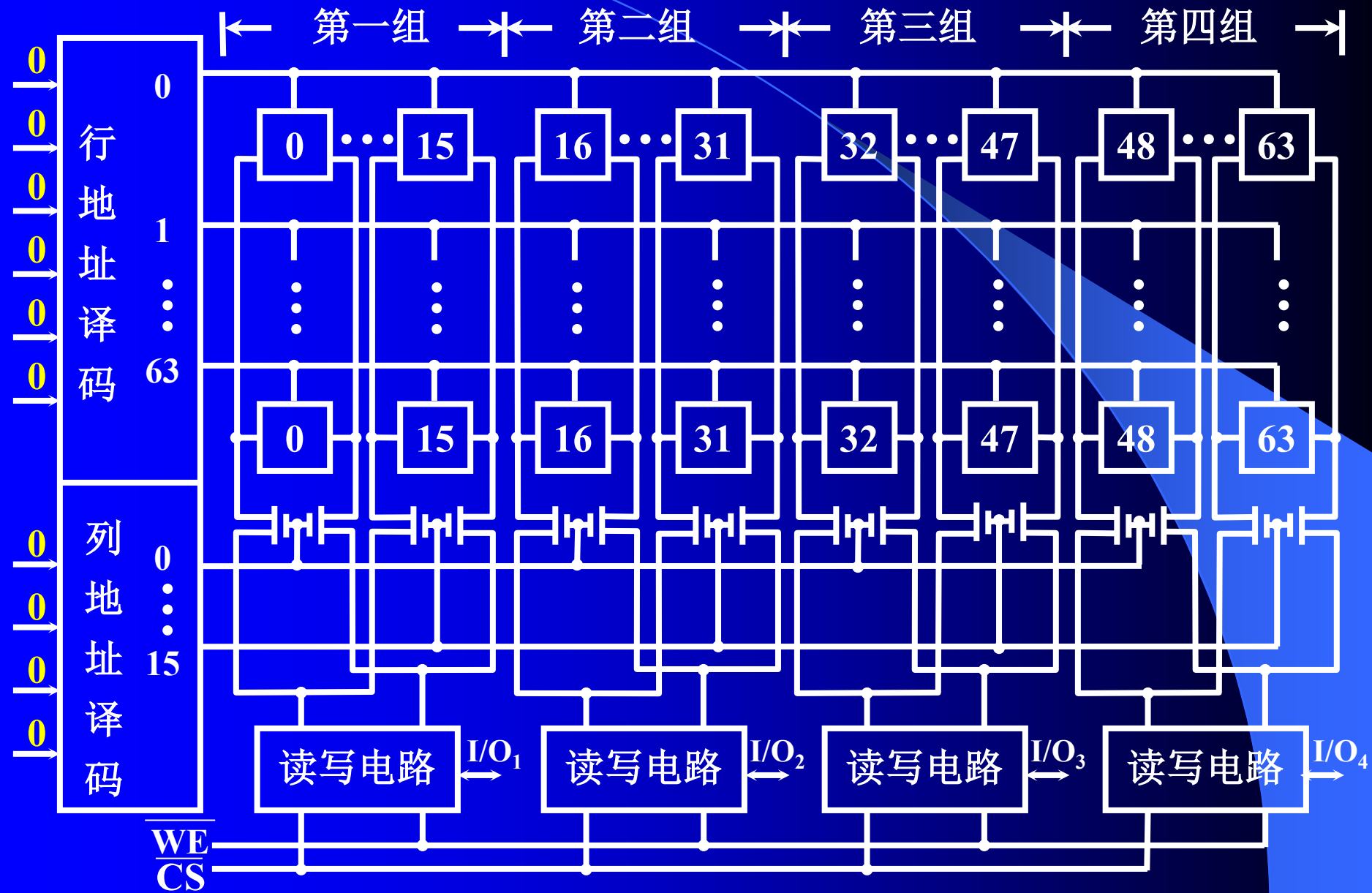
② Intel 2114 RAM 矩阵 (64×64) 读 4.2



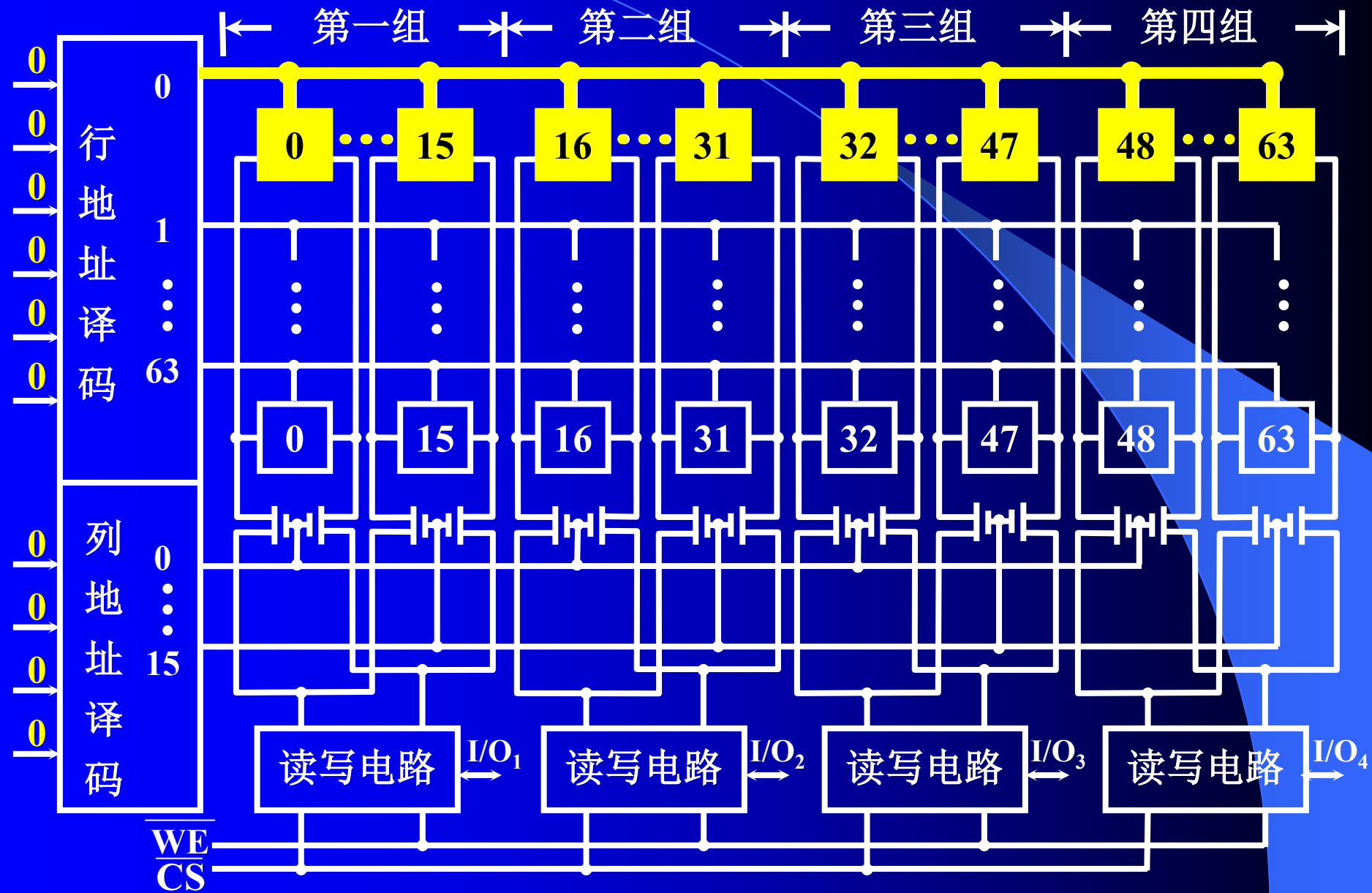
③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2



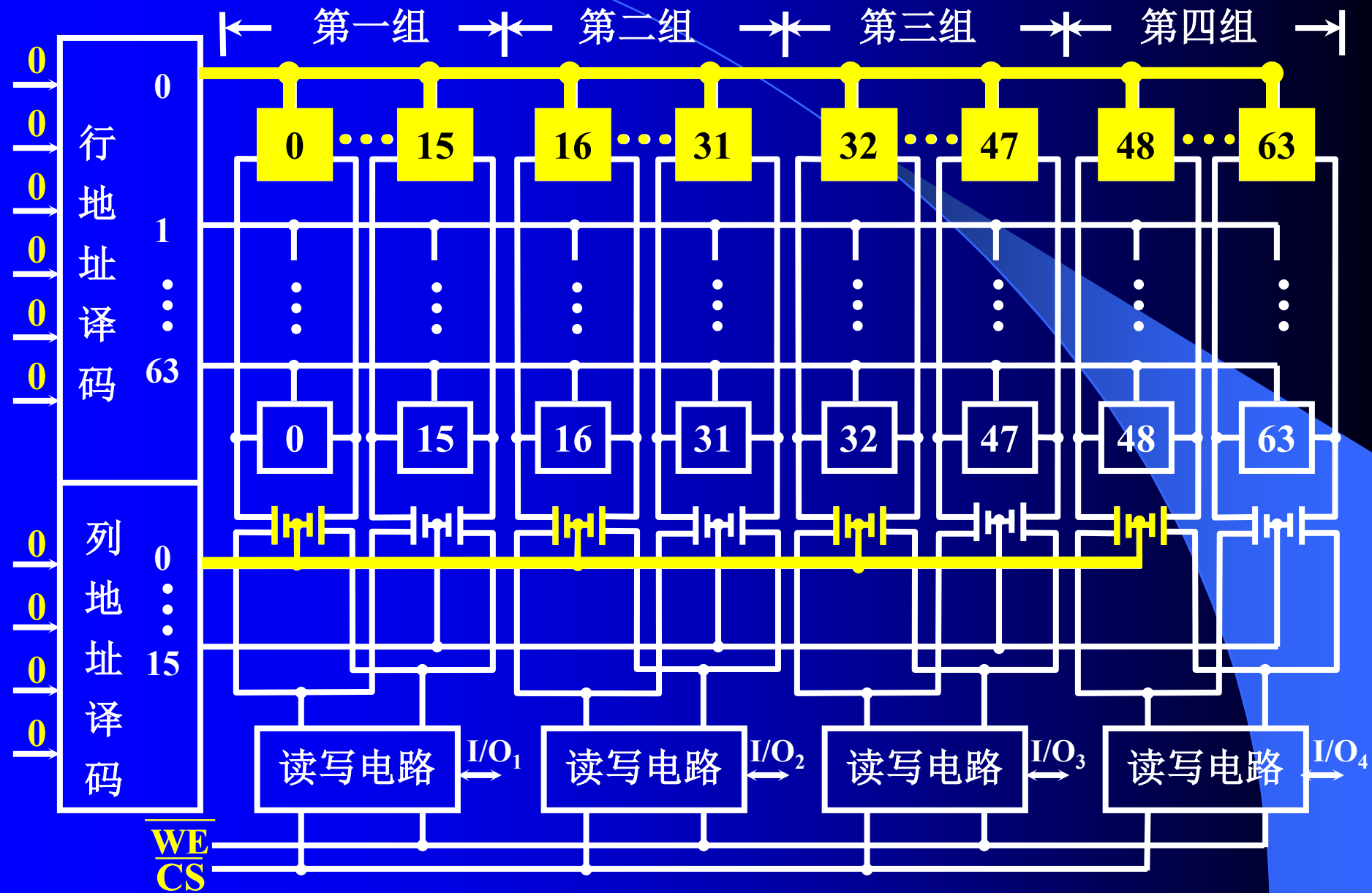
③ Intel 2114 RAM 矩阵 (64×64) 写 4.2



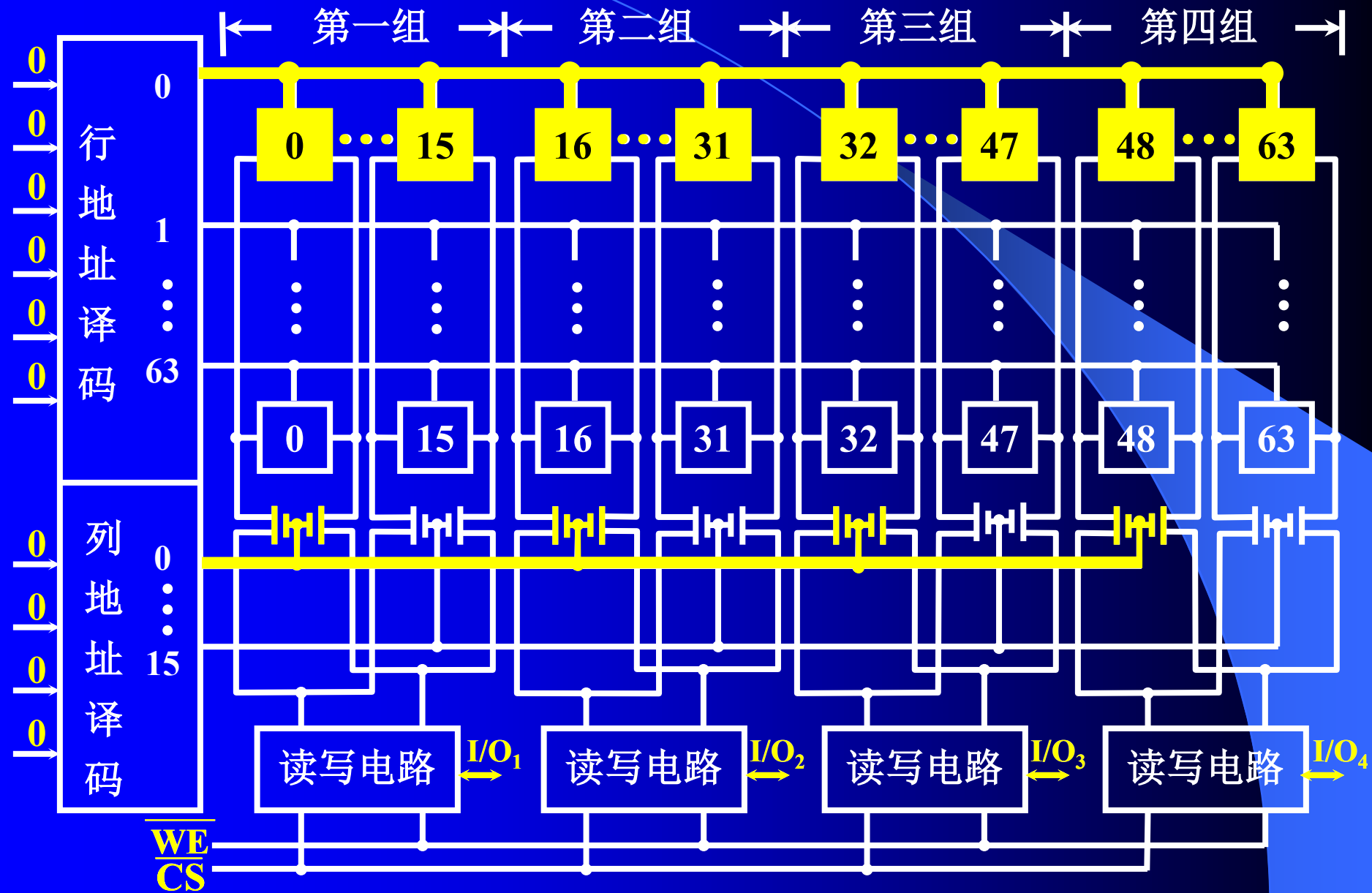
③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2



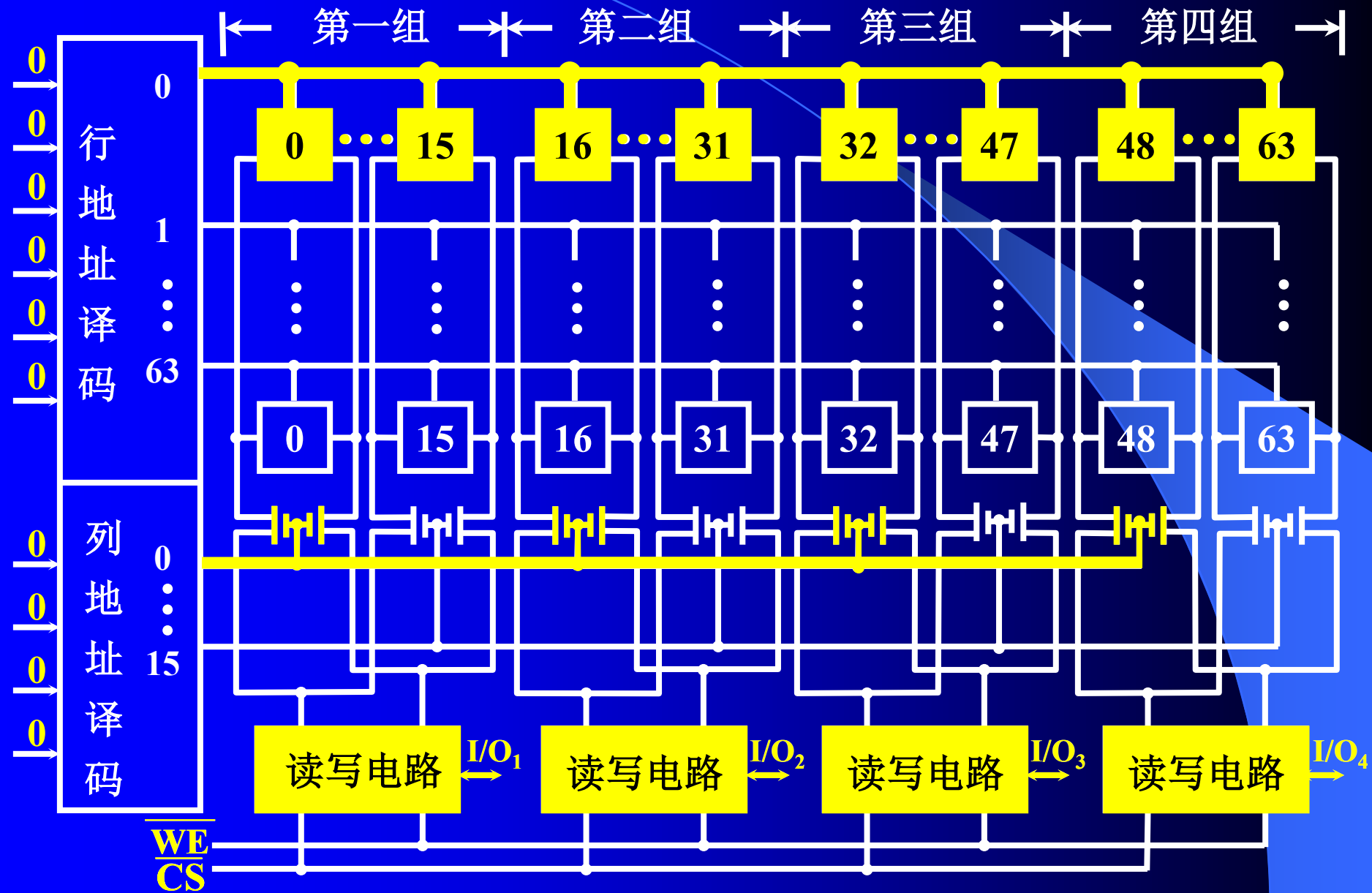
③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2



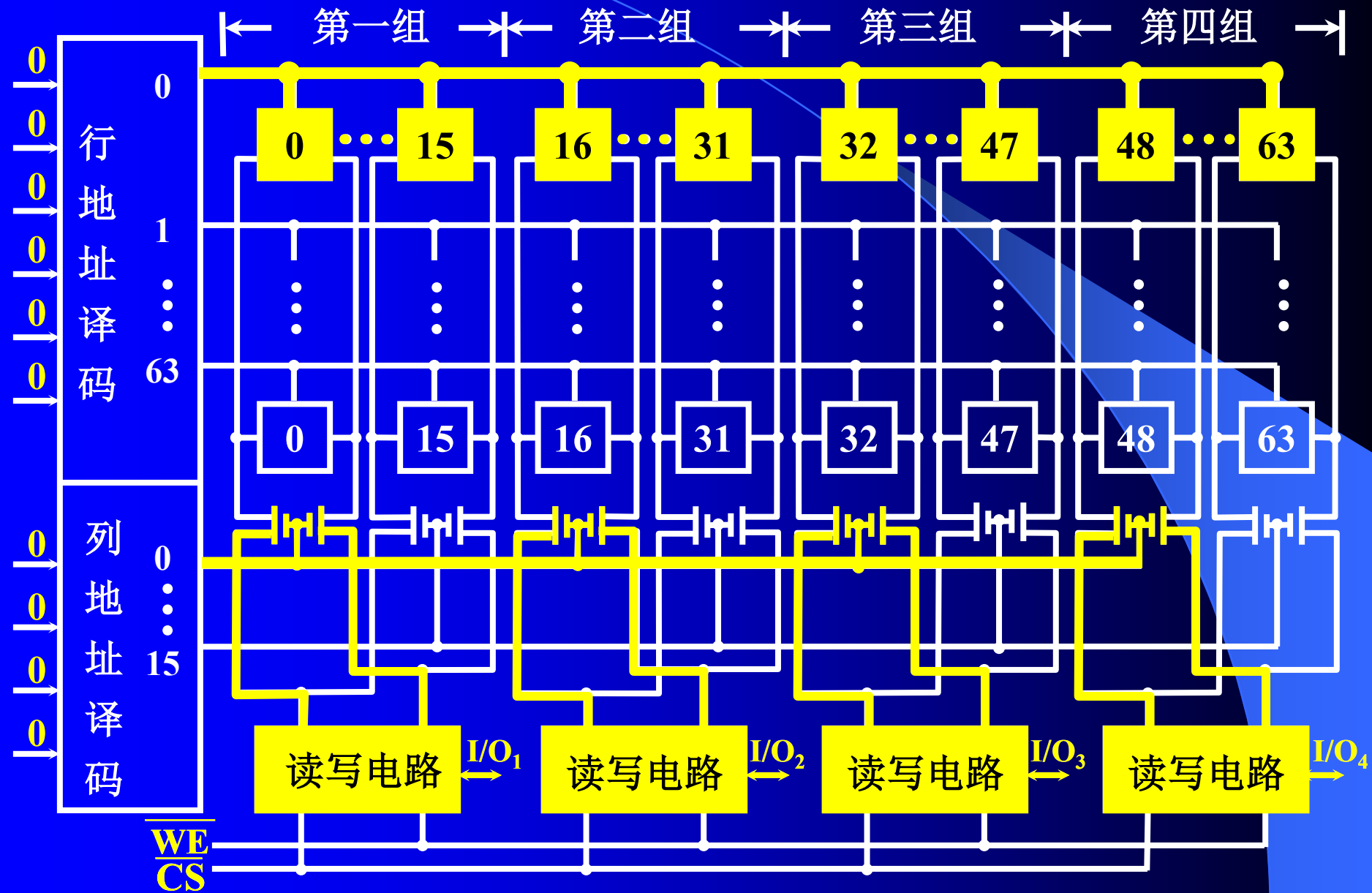
③ Intel 2114 RAM 矩阵 (64×64) 写 4.2



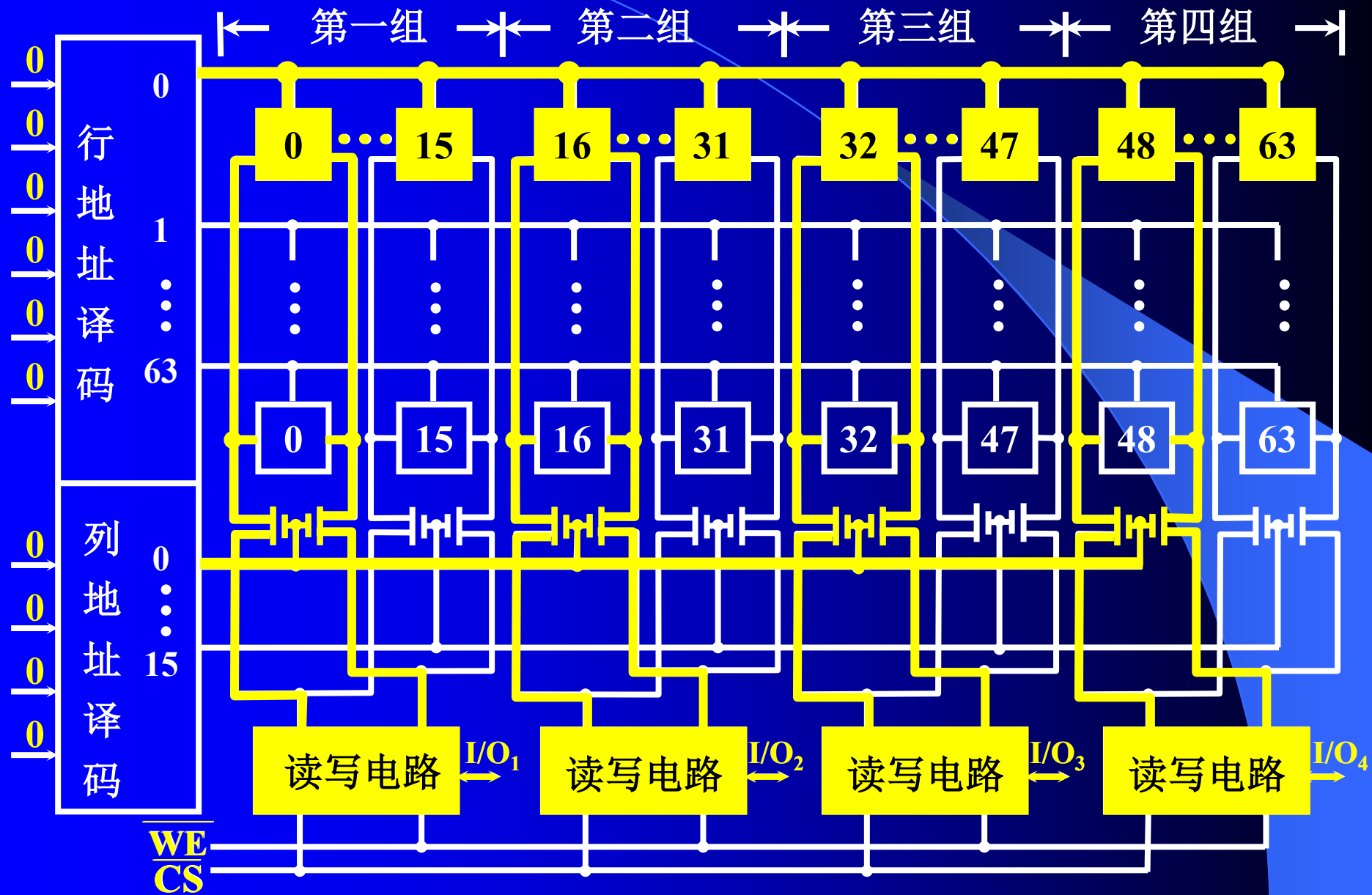
③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2



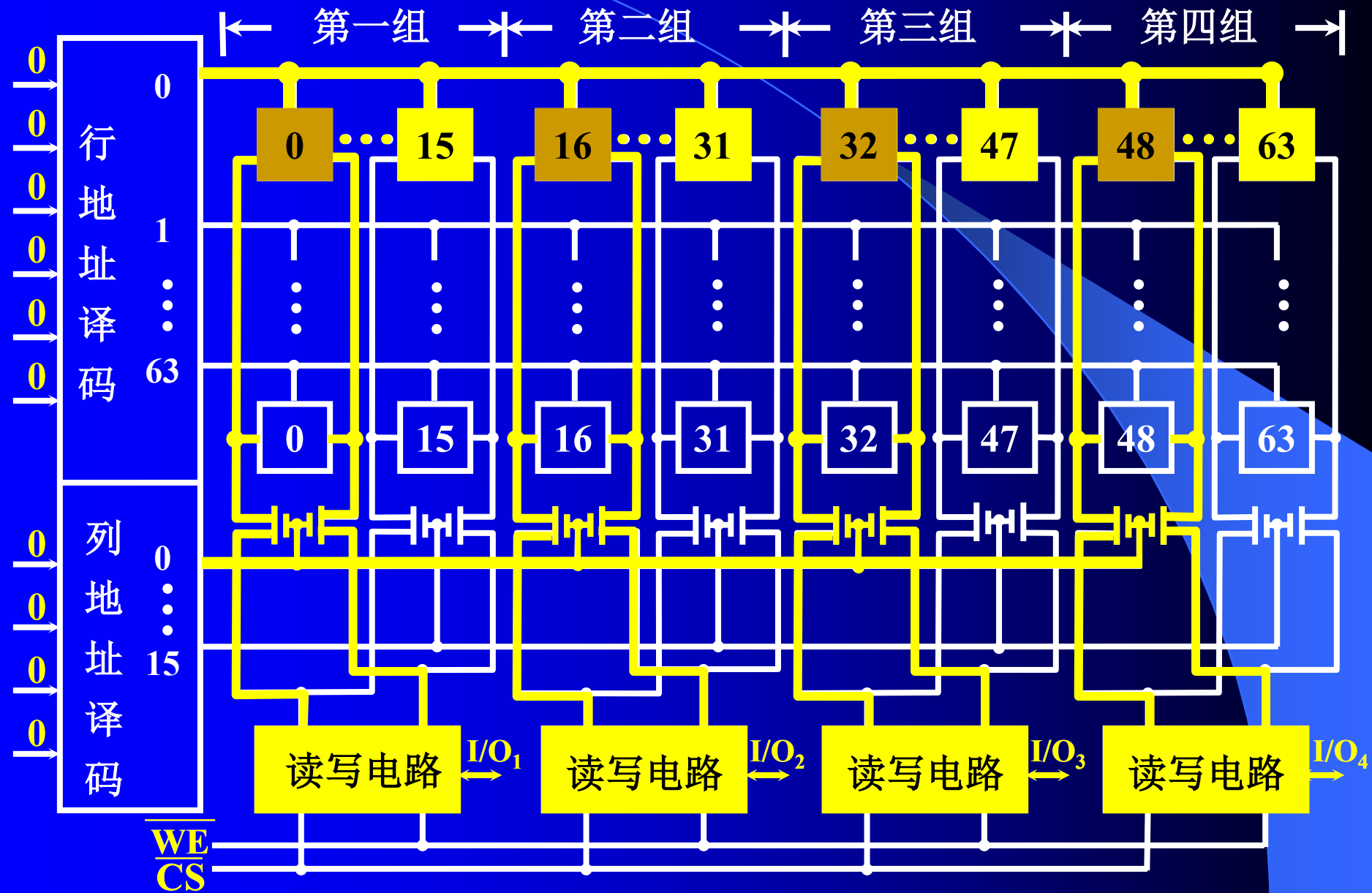
③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2



③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2

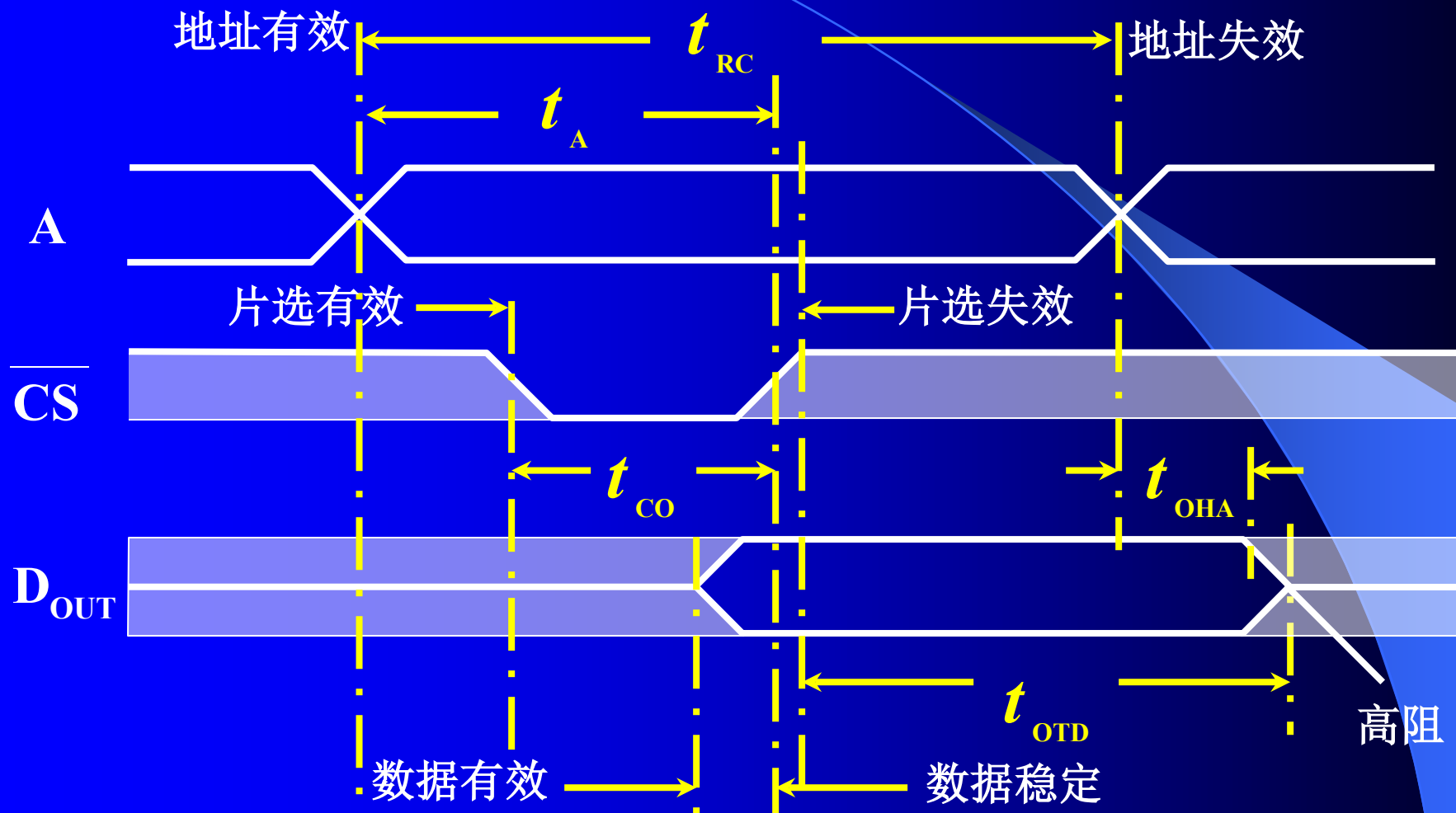


③ Intel 2114 RAM 矩阵 (64 × 64) 写 4.2



(3) 静态 RAM 读 时序

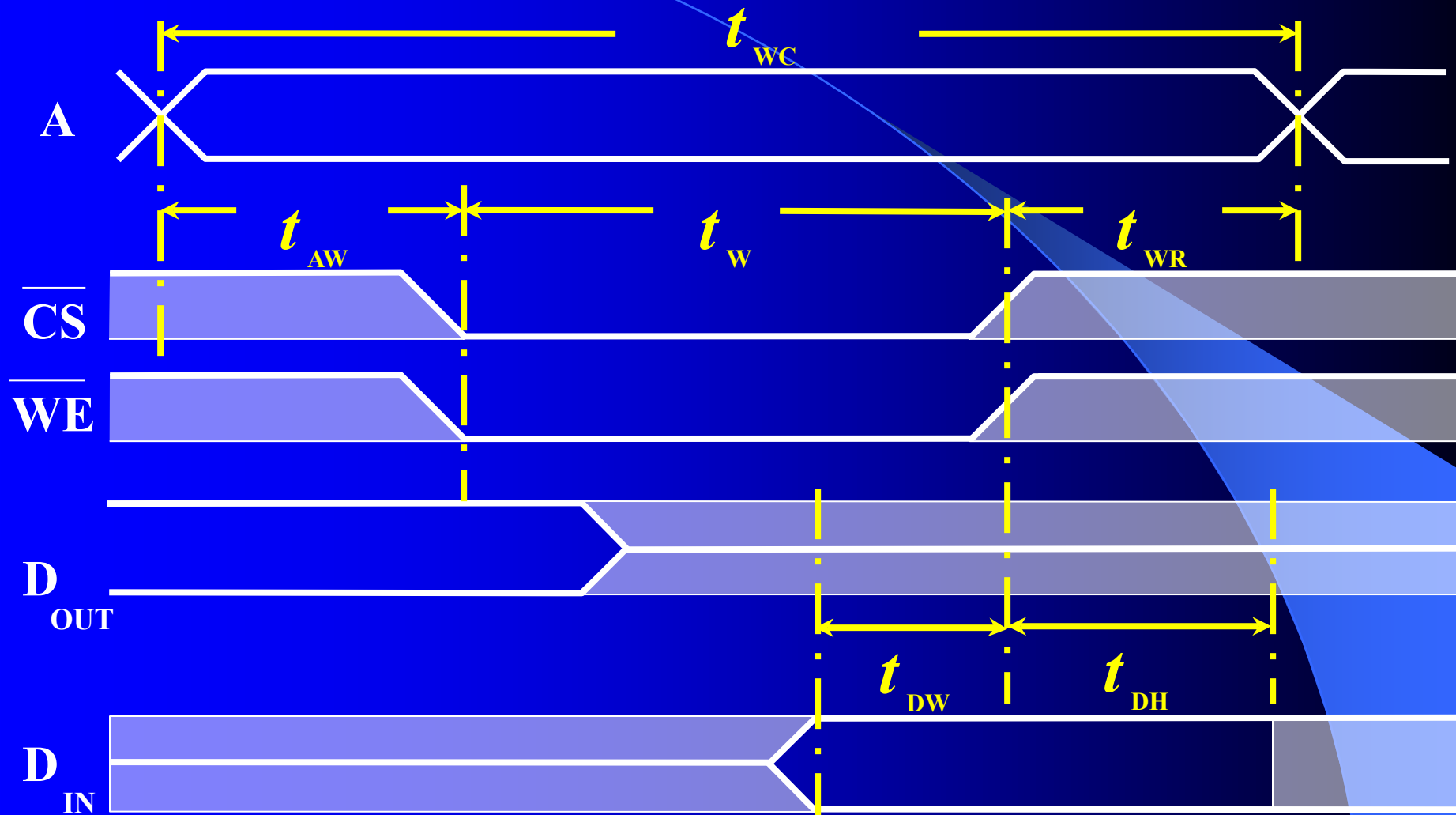
4.2



t_{OHA} 地址失效后的数据维持时间

(4) 静态 RAM (2114) 写时序

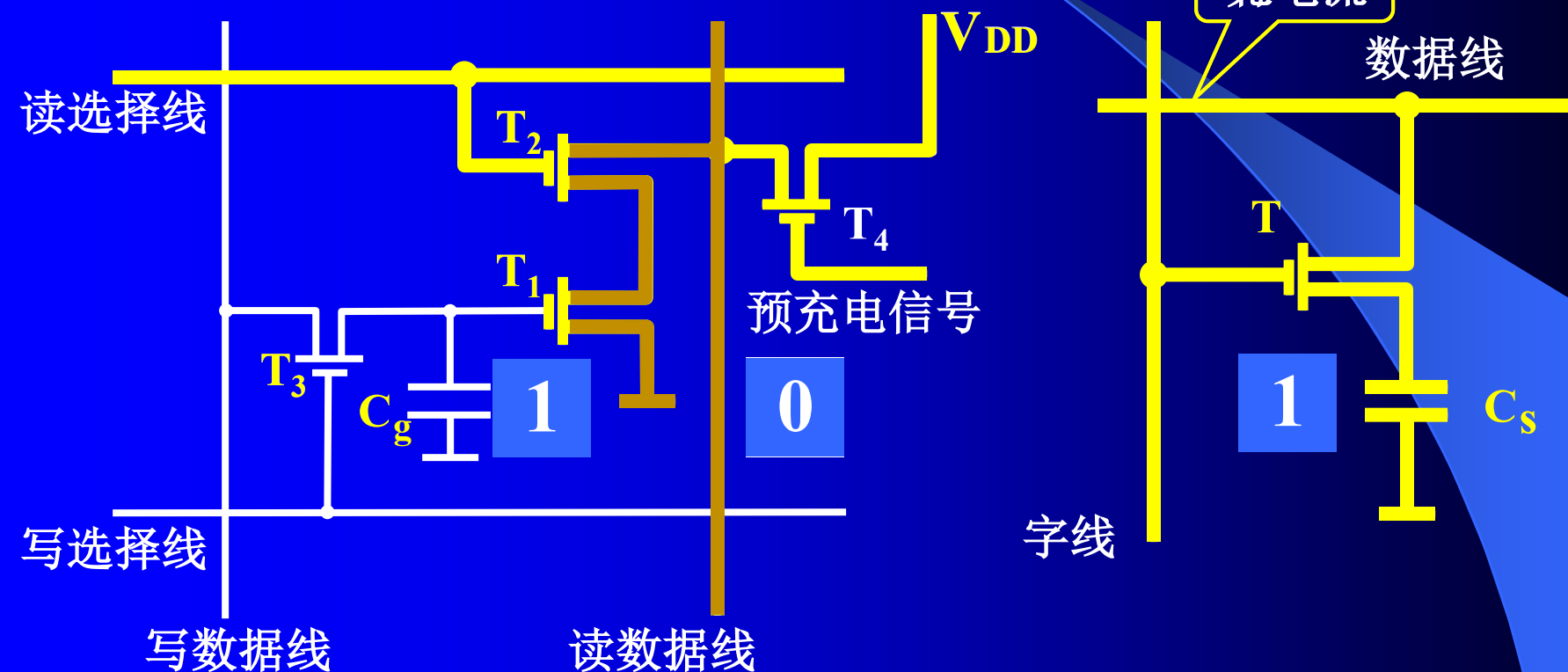
4.2



t_{DH} \overline{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

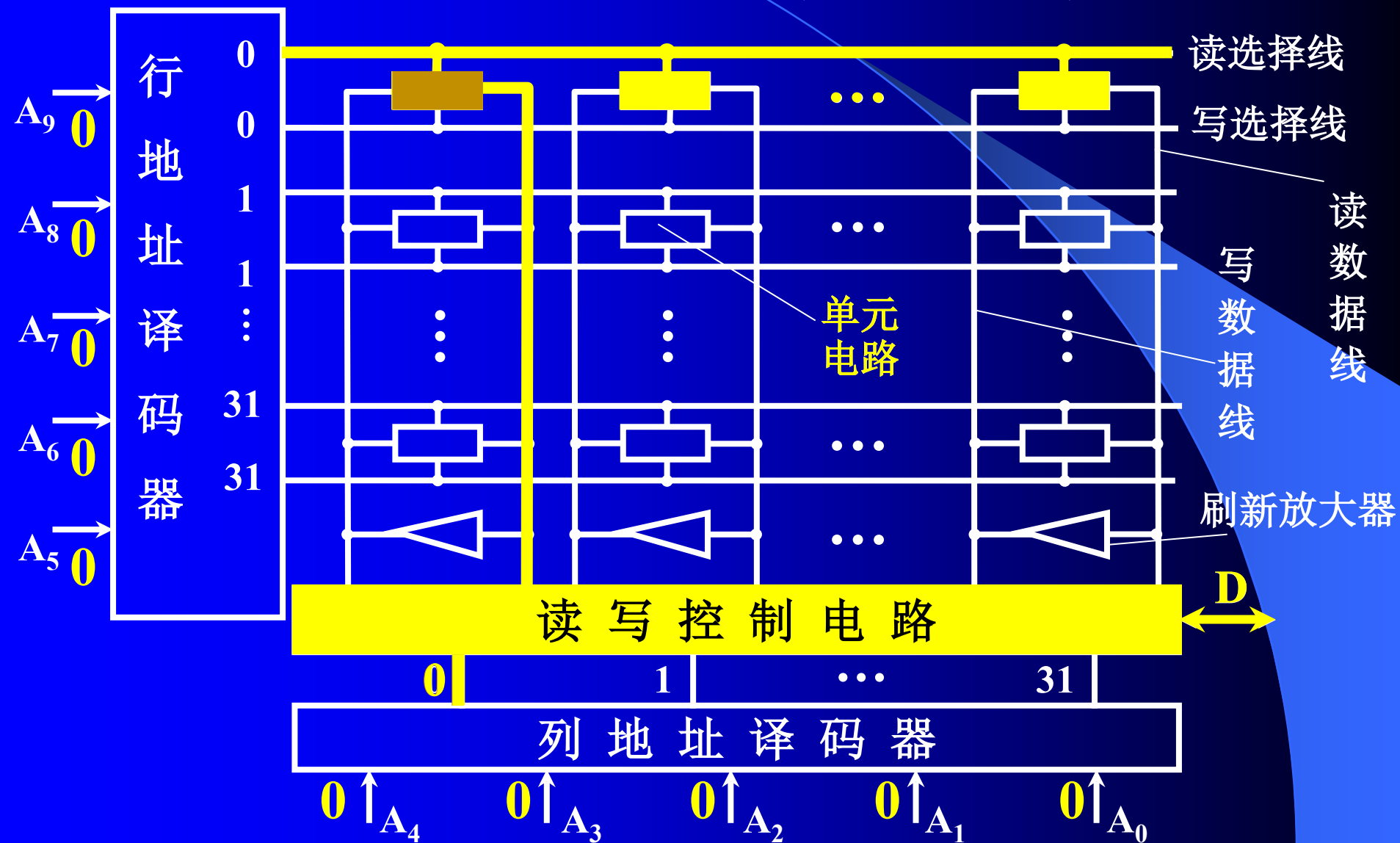
读出时数据线有电流为“1”

写入时 C_s 充电为“1”放电为“0”

(2) 动态 RAM 芯片举例

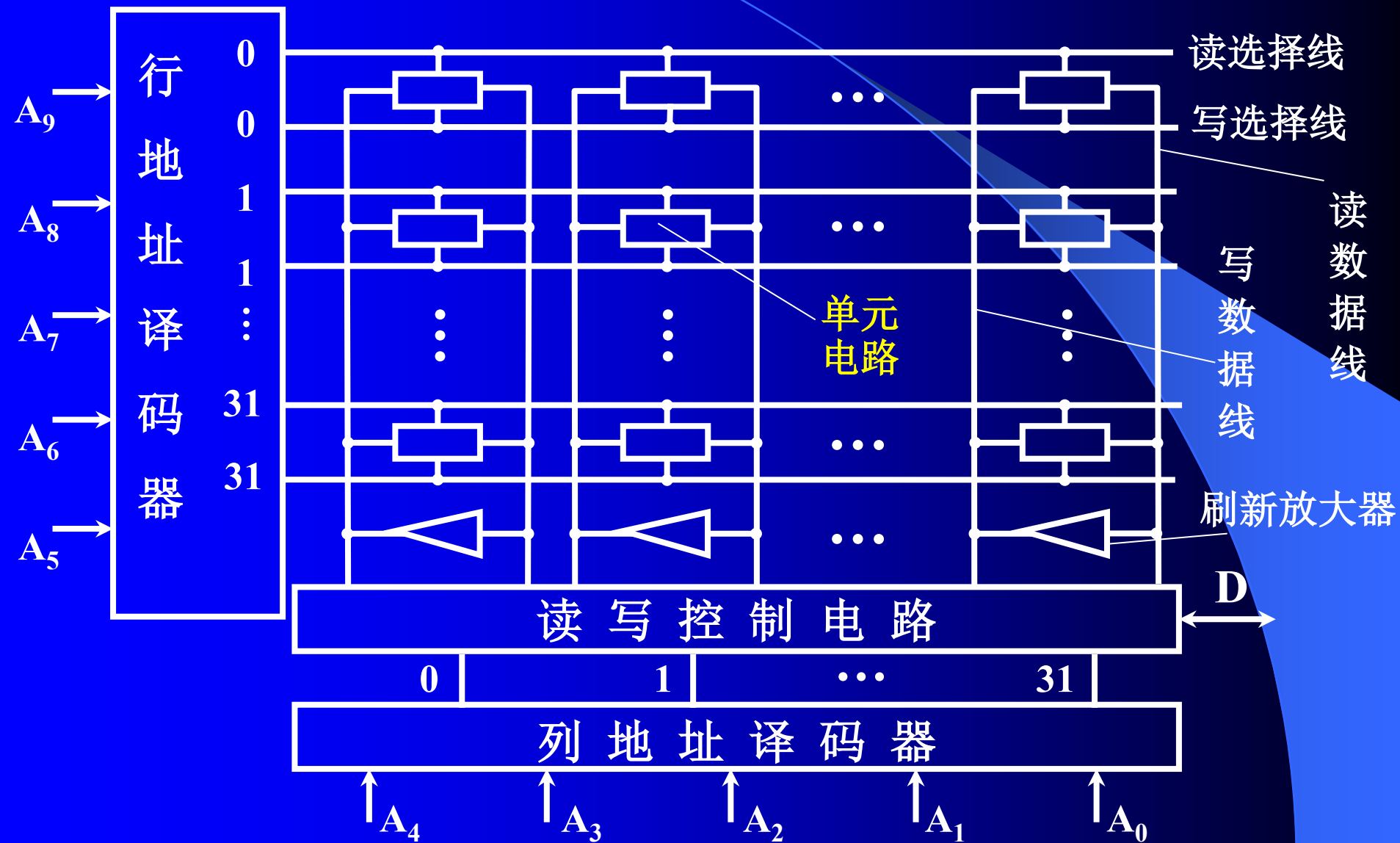
4.2

① 三管动态 RAM 芯片 (Intel 1103) 读



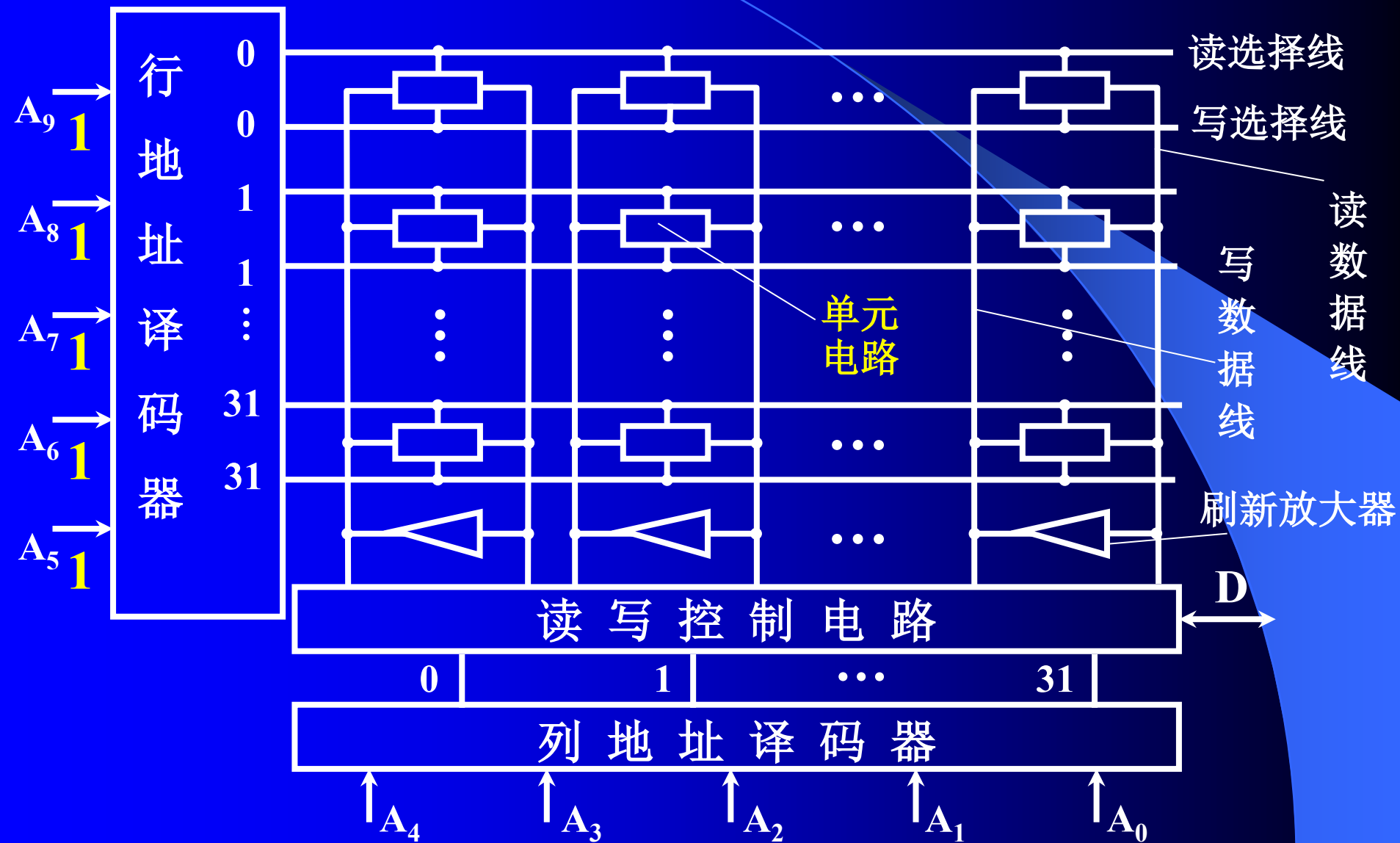
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2



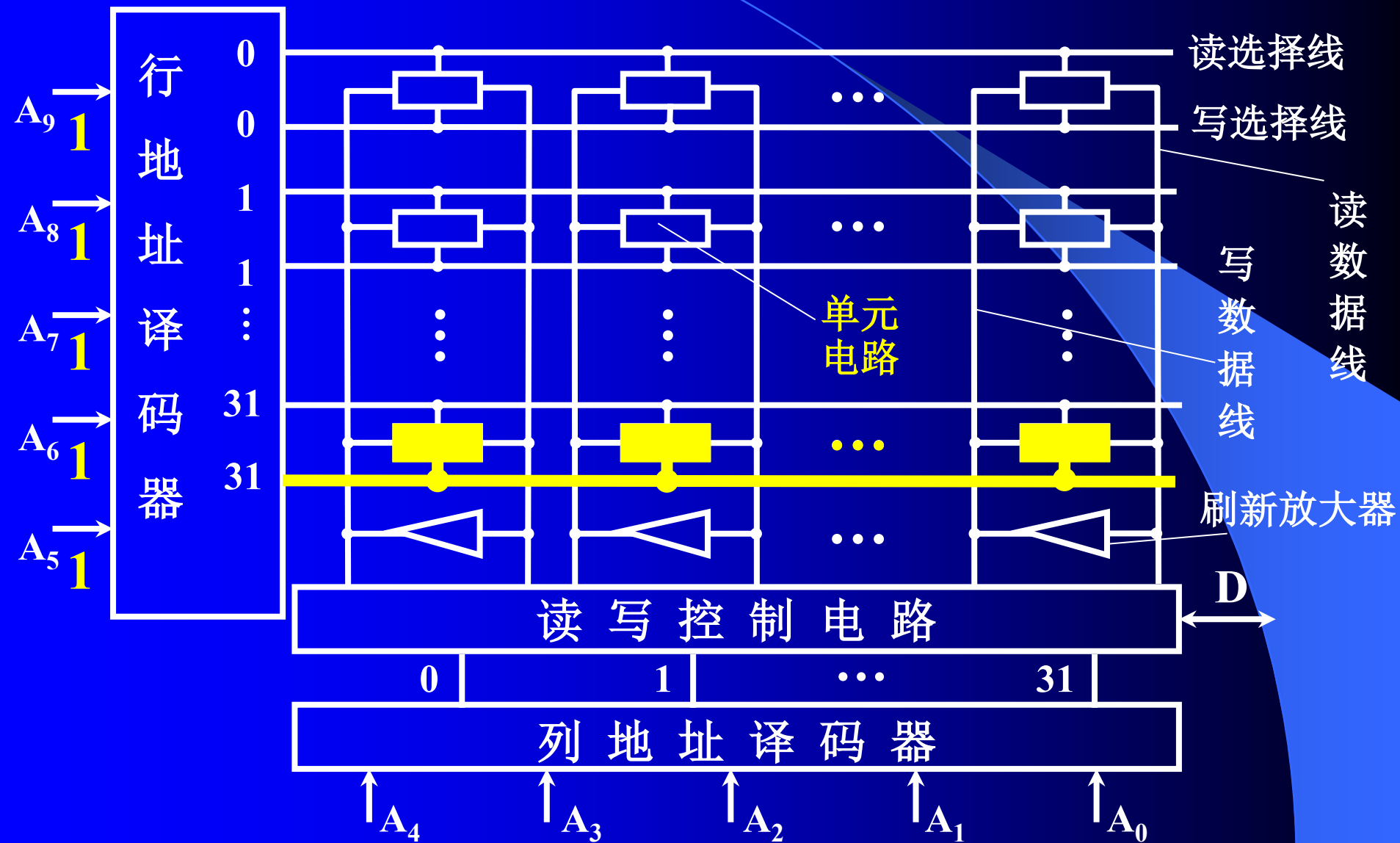
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2



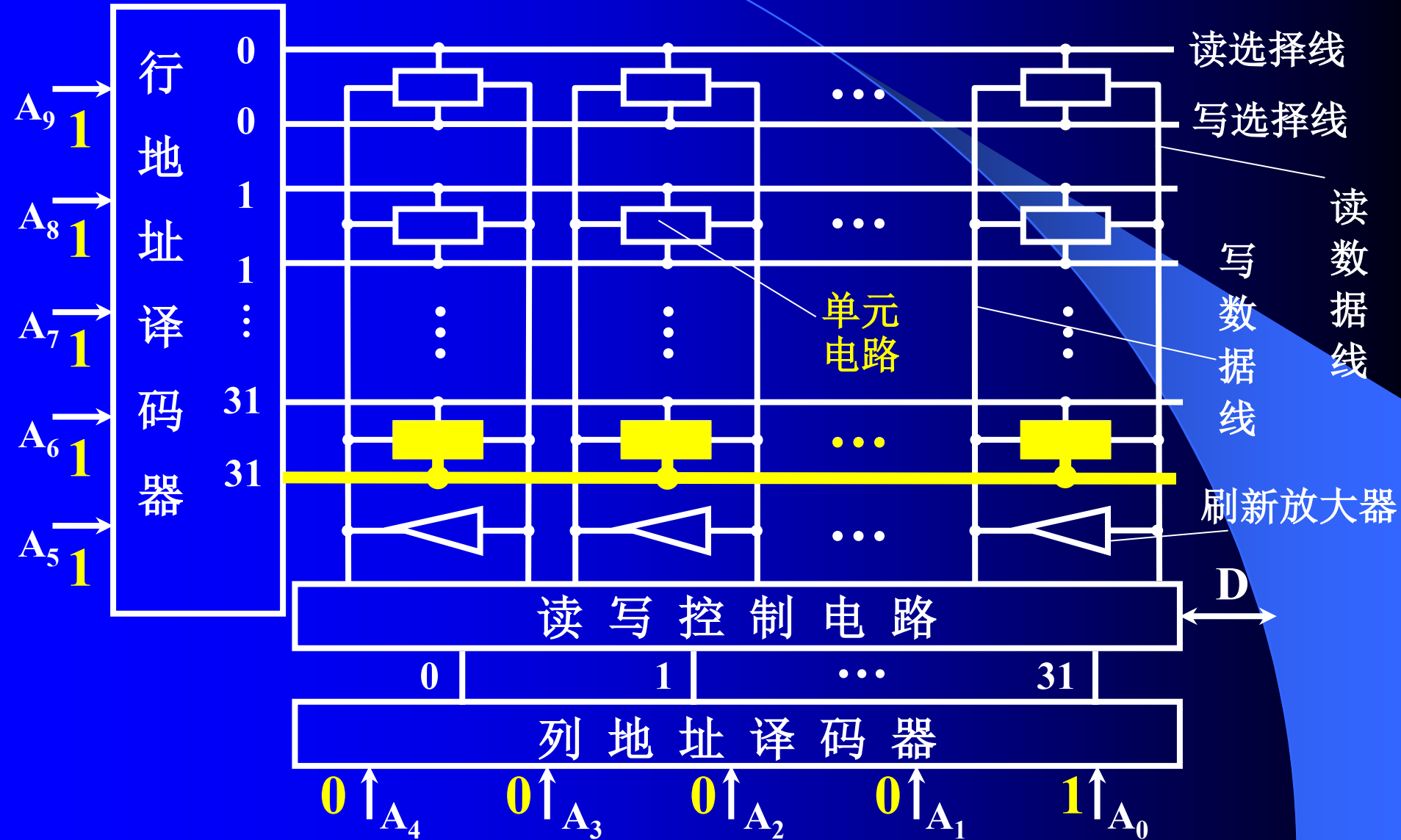
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2



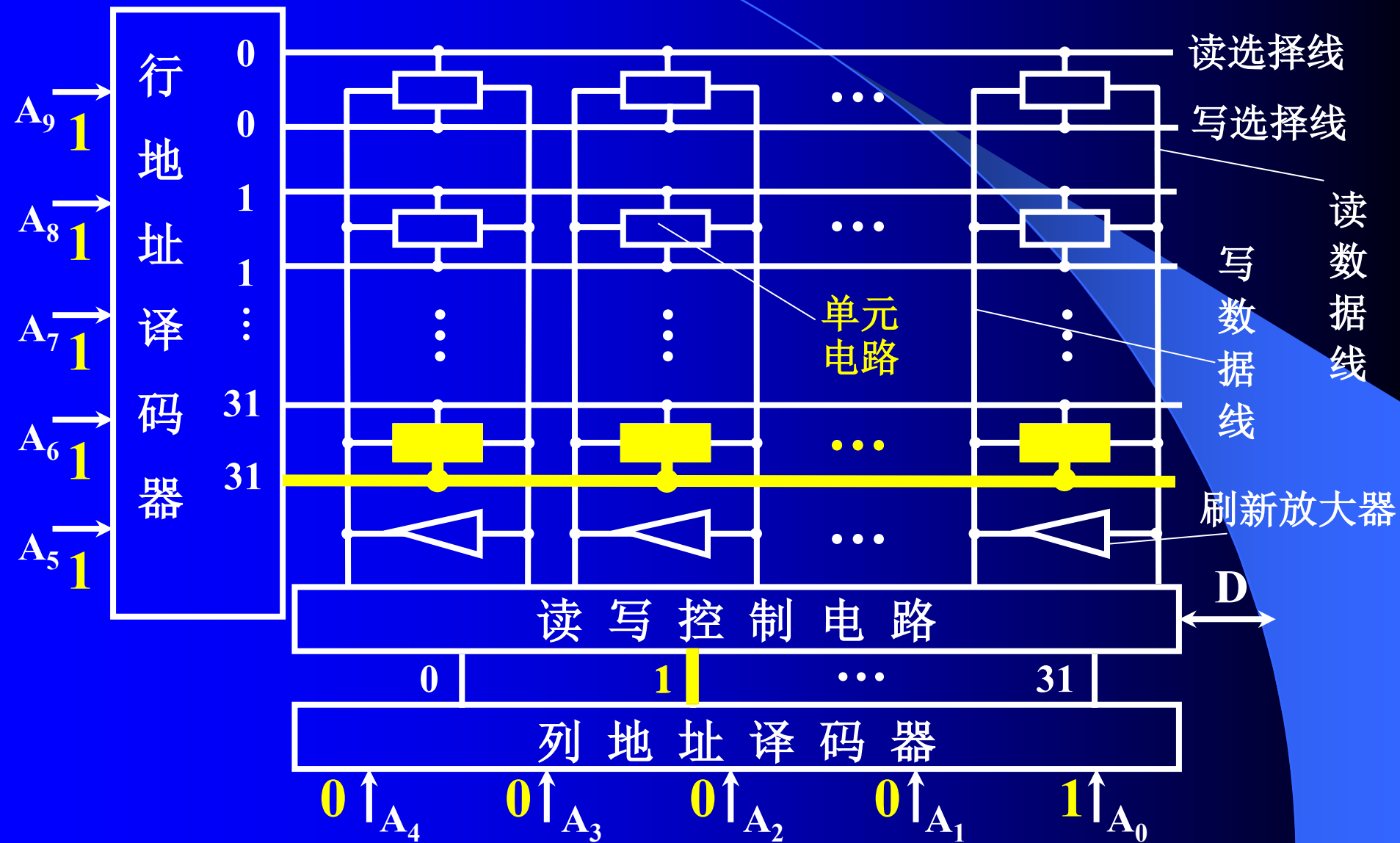
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2



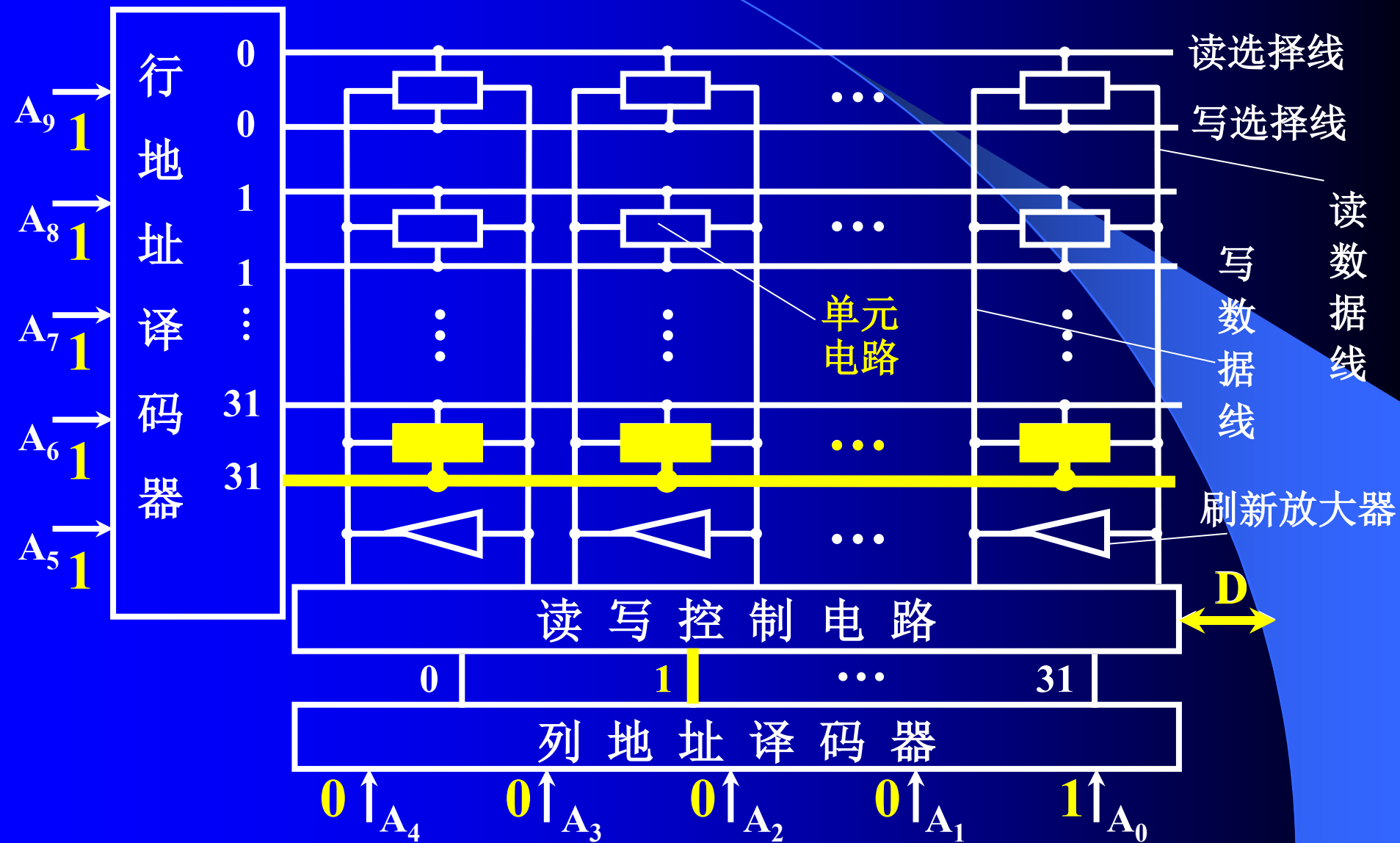
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2



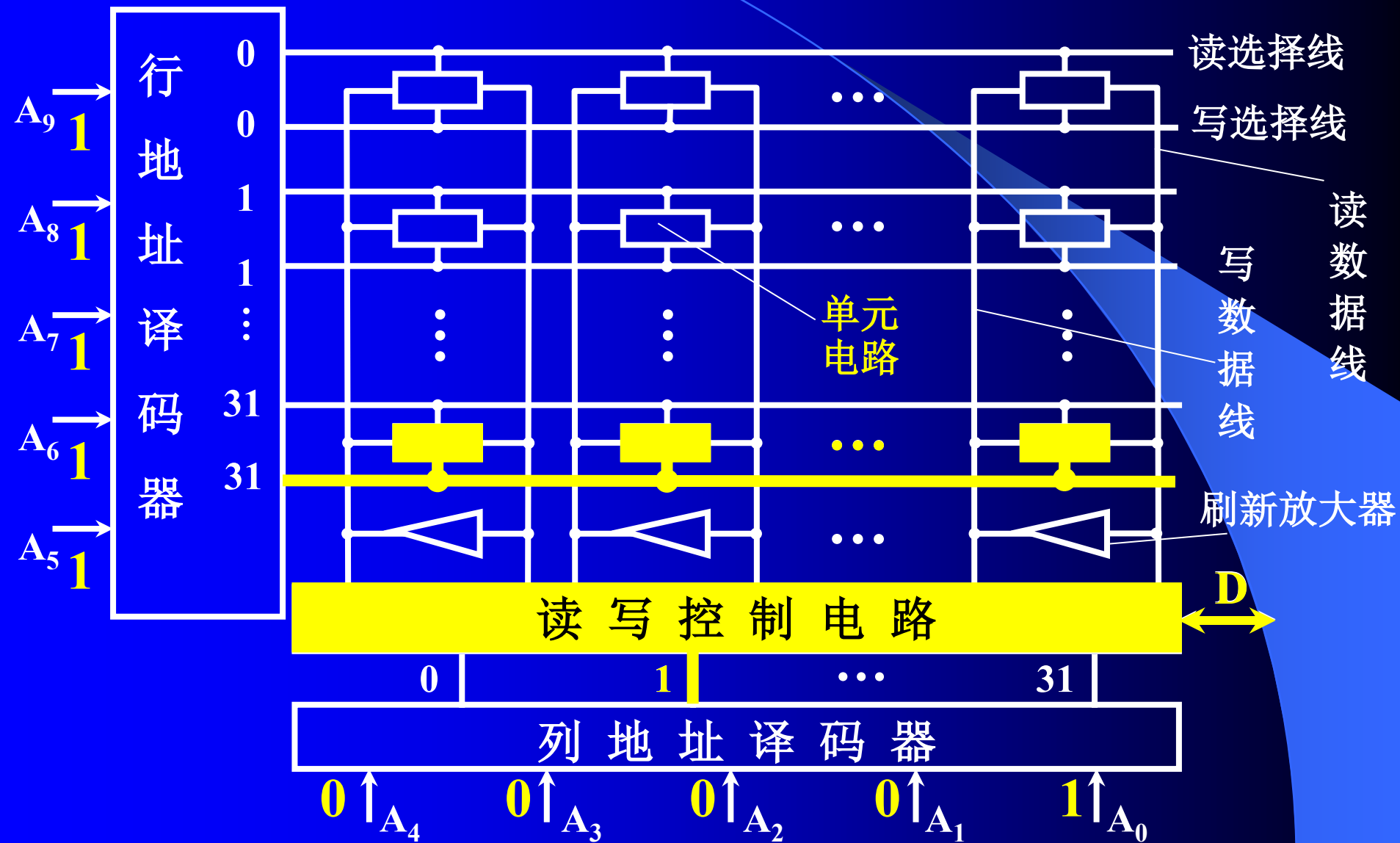
② 三管动态 RAM 芯片 (Intel 1103) 写

4.2

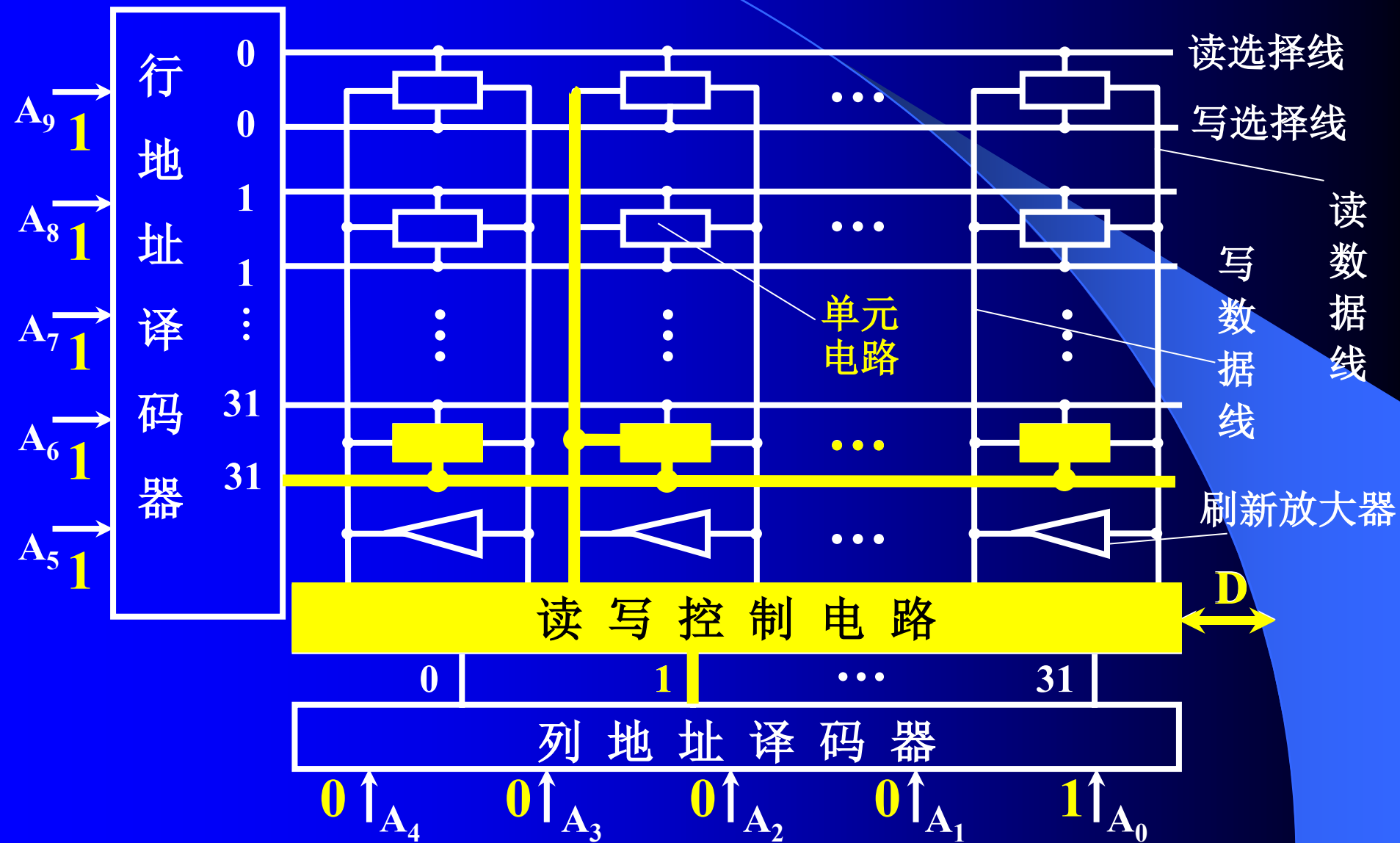


② 三管动态 RAM 芯片 (Intel 1103) 写

4.2

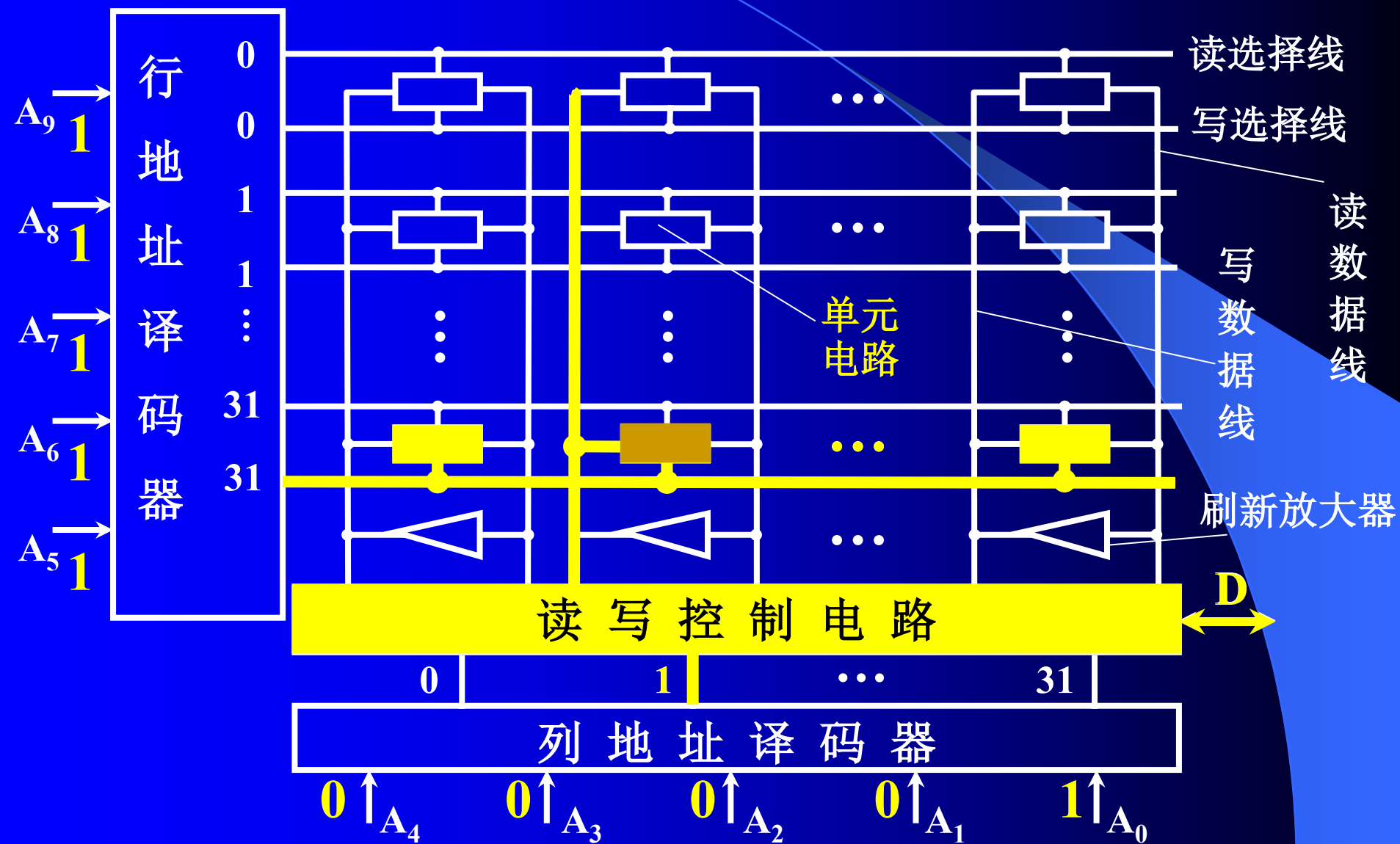


② 三管动态 RAM 芯片 (Intel 1103) 写 4.2

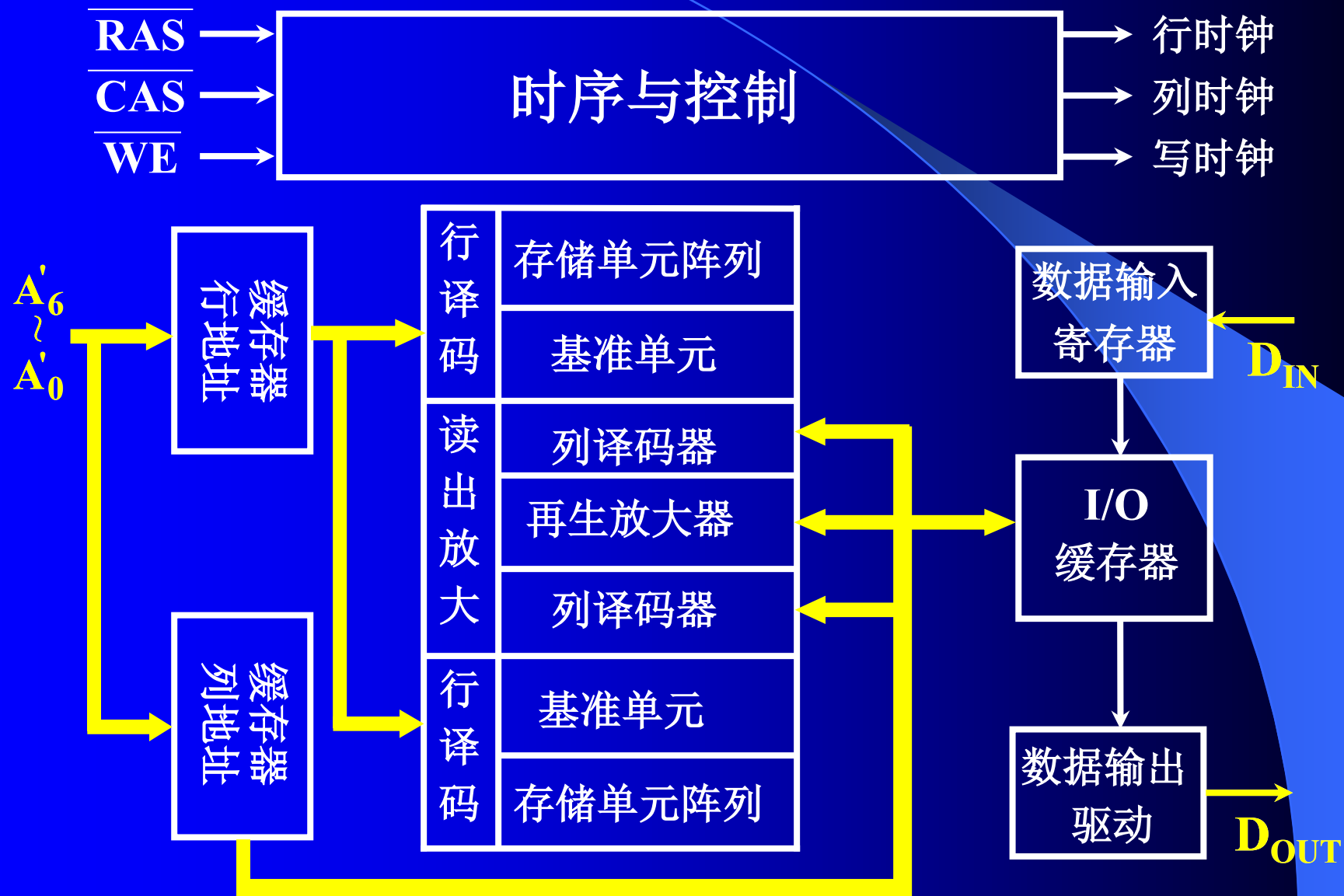


② 三管动态 RAM 芯片 (Intel 1103) 写

4.2

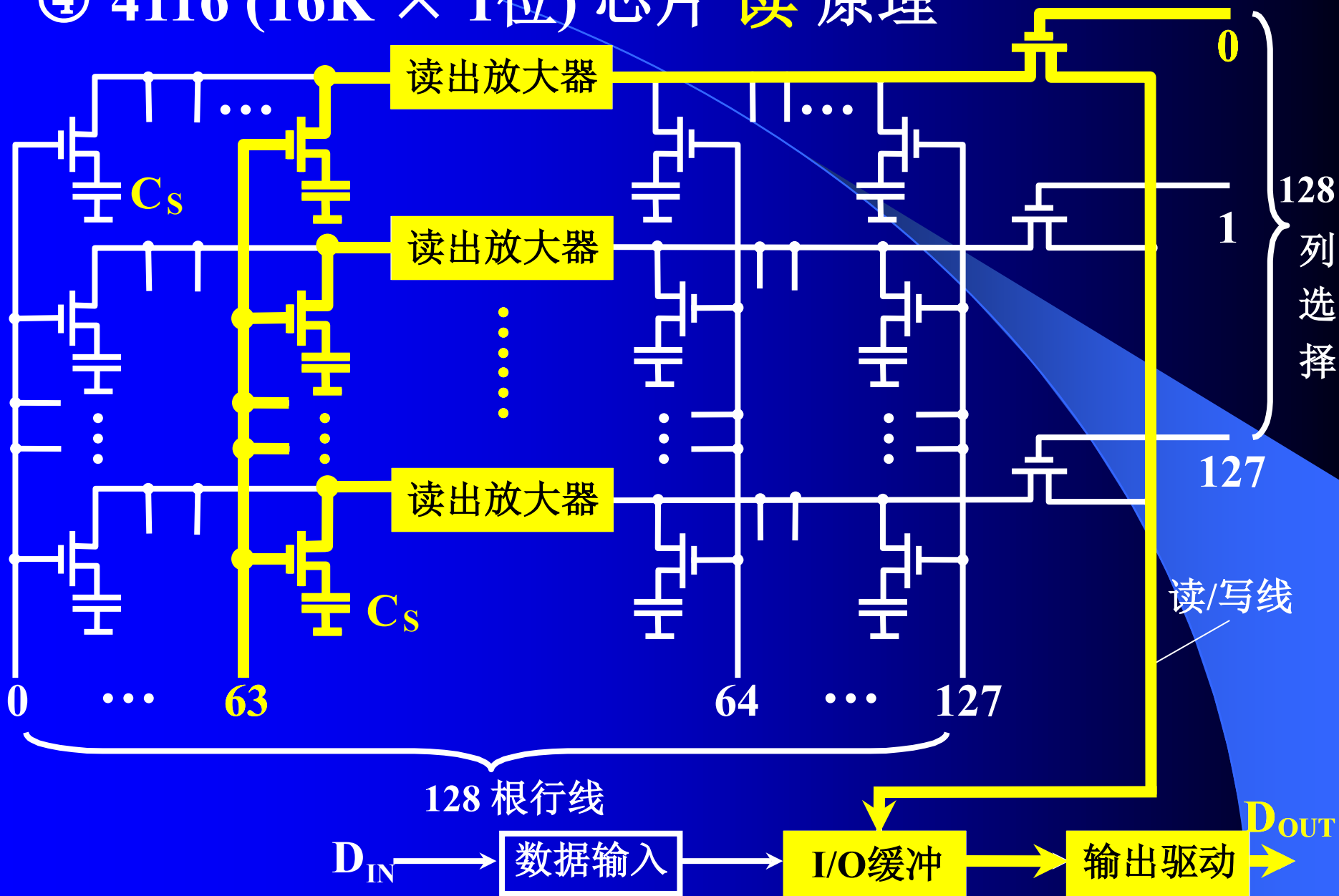


③ 单管动态 RAM 4116 (16K × 1位) 外特性 4.2

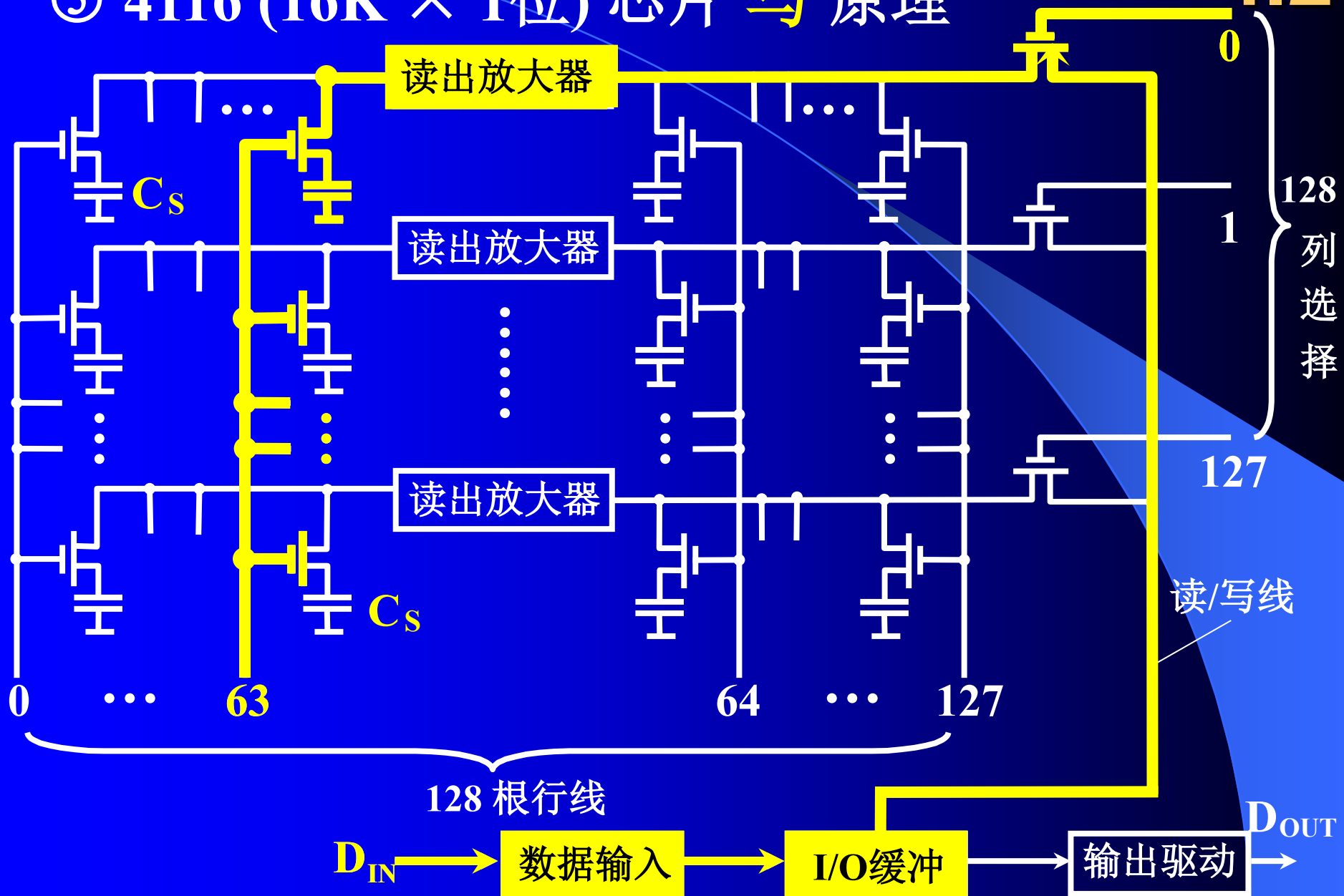


④ 4116 (16K × 1位) 芯片 读 原理

4.2



⑤ 4116 (16K × 1位) 芯片 写原理 4.2



(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

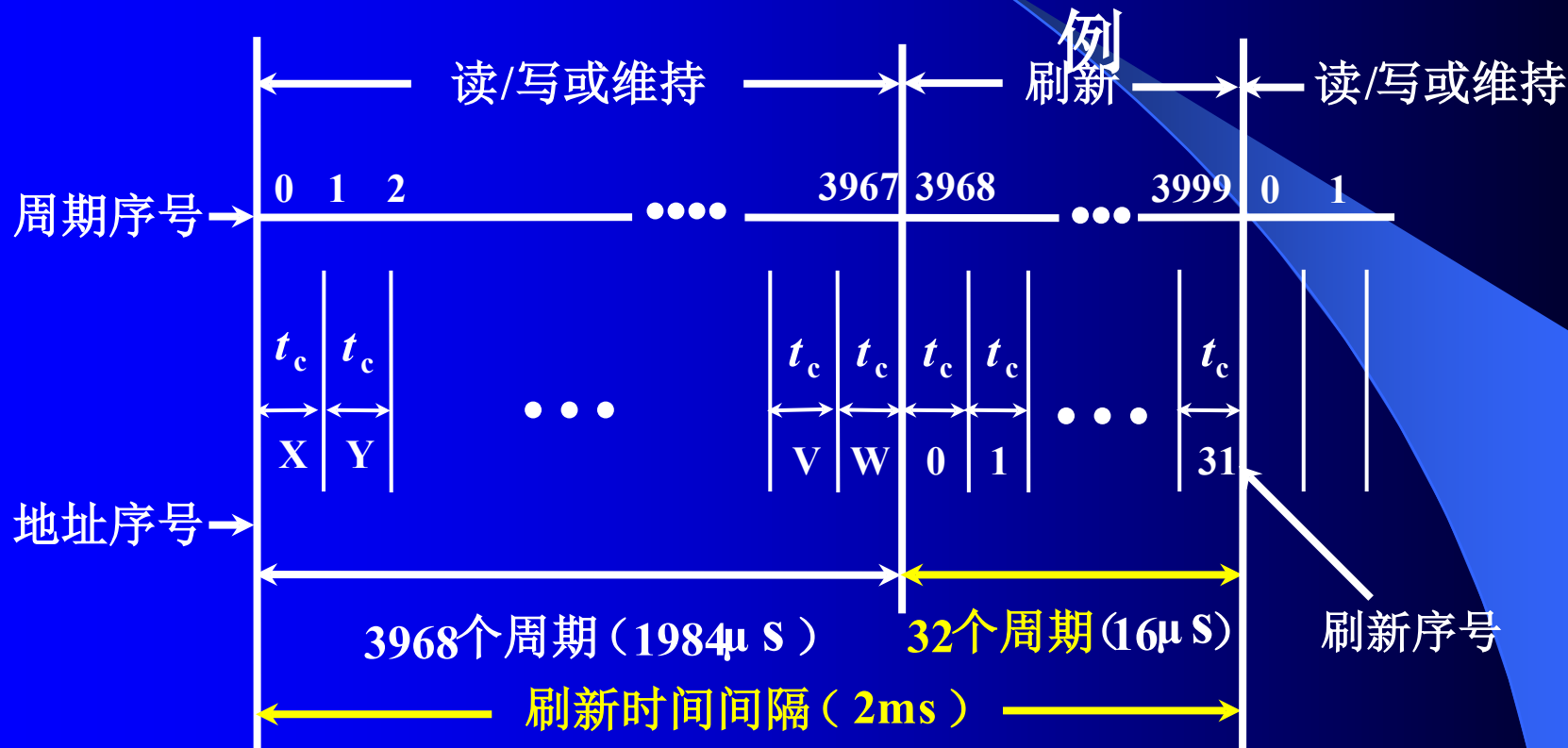
行地址 $\overline{\text{RAS}}$ 有效
写允许 $\overline{\text{WE}}$ 有效(低)
数据 D_{IN} 有效
列地址 $\overline{\text{CAS}}$ 有效

(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 32×32 矩阵为例



“死区” 为 $0.5\mu\text{s} \times 32 = 16\mu\text{s}$

“死时间率” 为 $32/4000 \times 100\% = 0.8\%$

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

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

若每隔 2 ms 集中刷新一次 “死区” 为 $64\mu\text{s}$

若每隔 $15.6\mu\text{s}$ 刷新一行

而且每行每隔 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

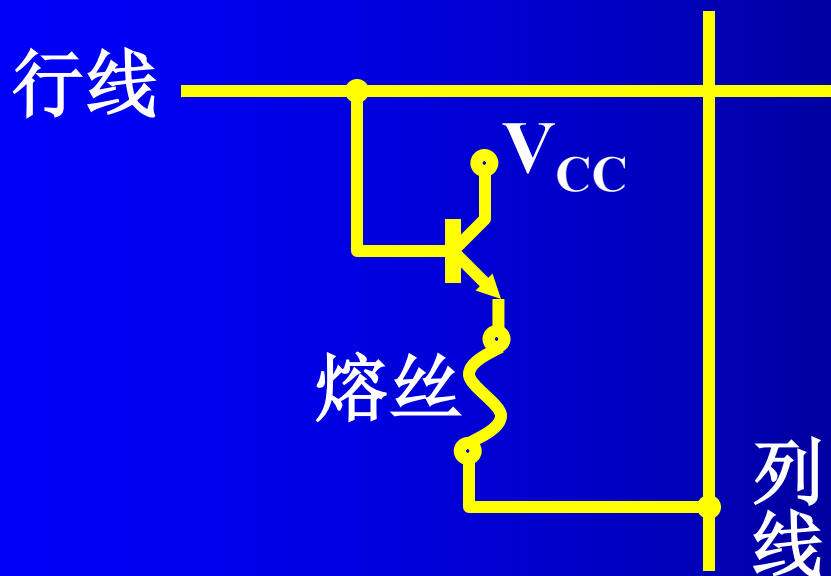
四、只读存储器 (ROM)

1. 掩膜 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)

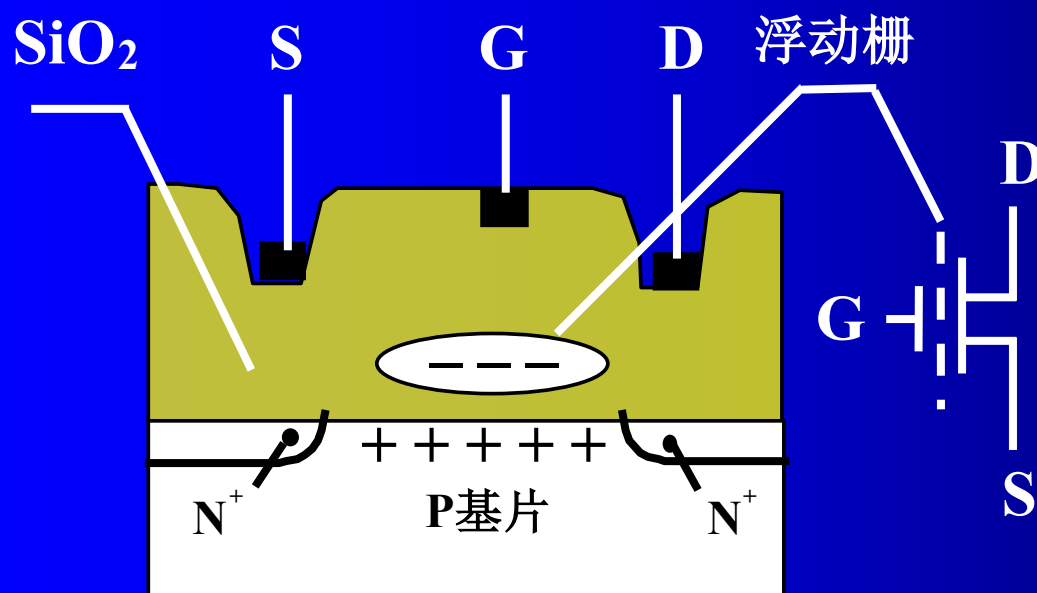


熔丝断 为 “0”

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”

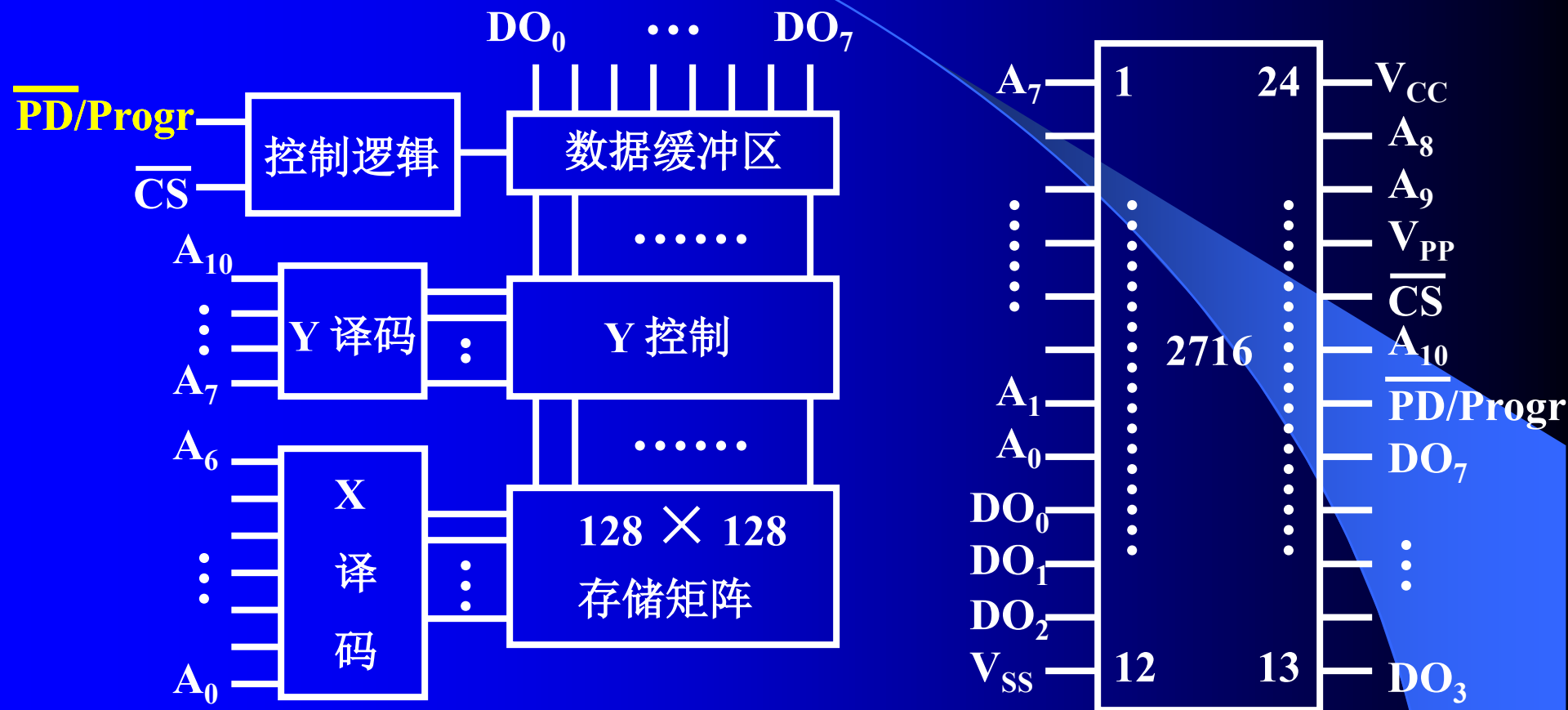
D 端不加正电压

不形成浮动栅

S 与 D 导通为 “1”

(2) 2716 EPROM 的逻辑图和引脚

4.2



$\overline{\text{PD/Progr}}$ 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 (快擦型存储器)

EPROM 价格便宜 集成度高

EEPROM 电可擦写重写

比 E²PROM快 具备 RAM 功能

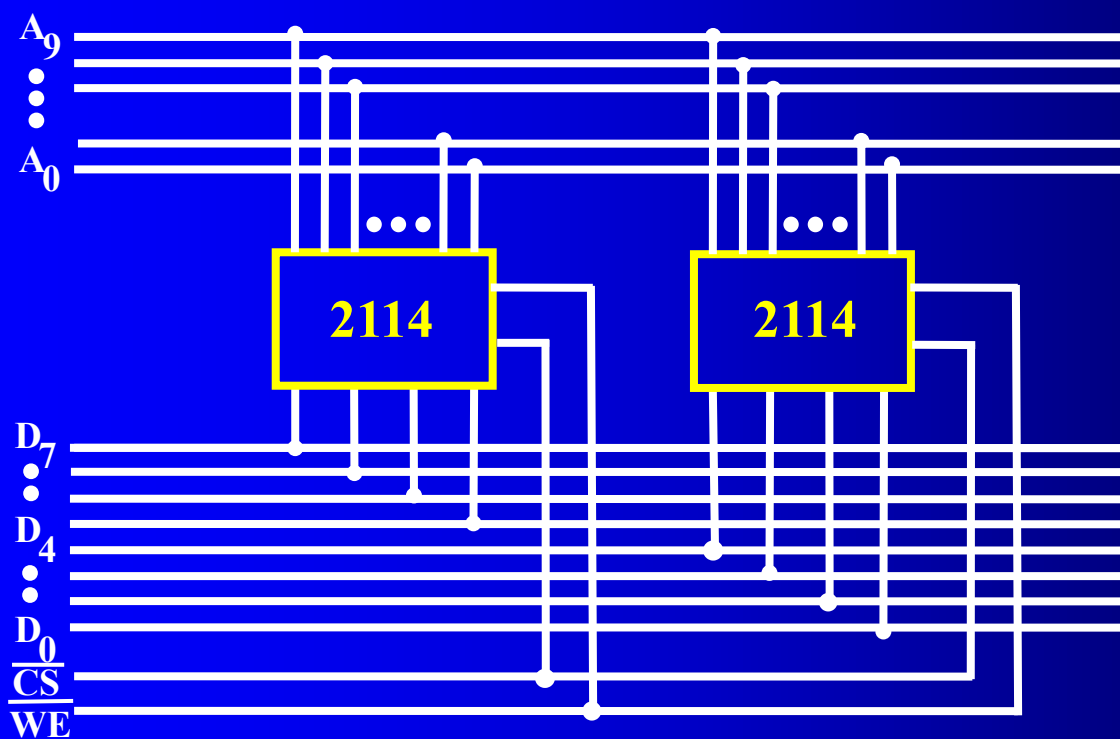
五、存储器与 CPU 的连接

4.2

1. 存储器容量的扩展

(1) 位扩展（增加存储字长）

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



10根地址线

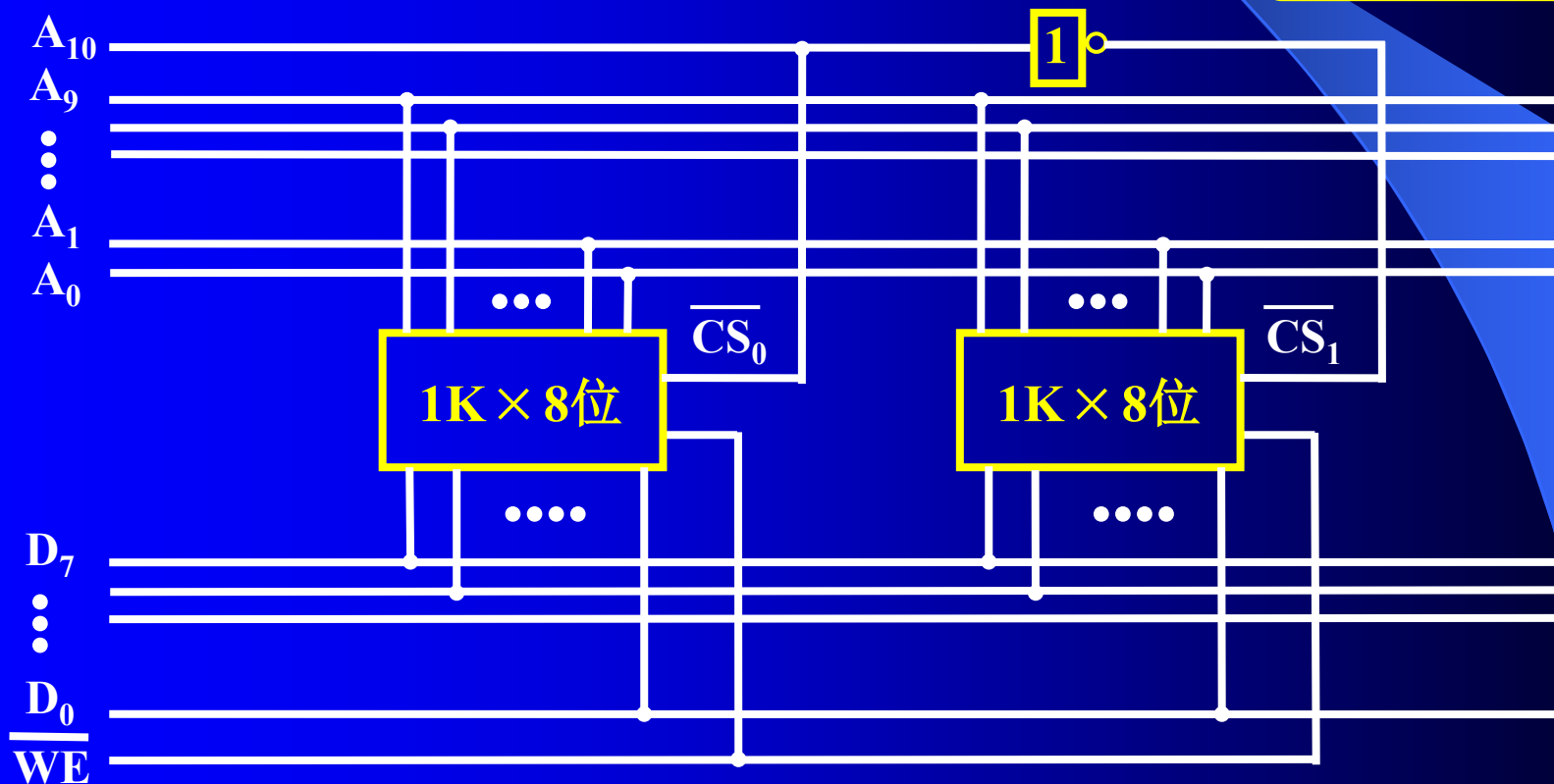
8根数据线

(2) 字扩展（增加存储字的数量）

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

11根地址线

8根数据线



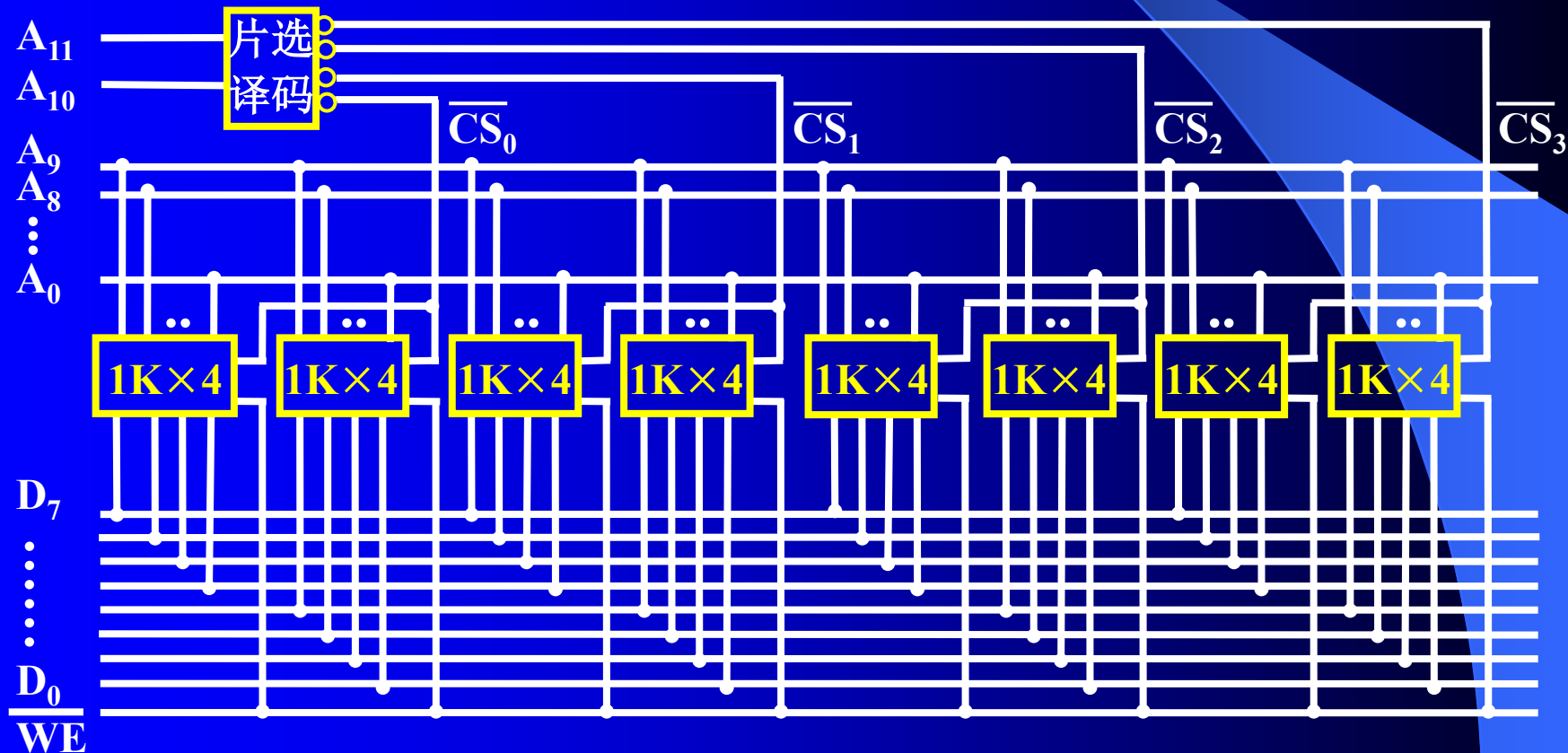
(3) 字、位扩展

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/写线的连接
- (4) 片选线的连接
- (5) 合理选用芯片
- (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

$A_{15}A_{14}A_{13} \quad A_{11}A_{10} \quad \dots \quad A_7 \quad \dots \quad A_4 A_3 \quad \dots \quad A_0$

0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0

⋮

0 1 1 0 0 1 1 1 1 1 1 1 1 1 1 1

0 1 1 0 1 0 0 0 0 0 0 0 0 0 0 0

⋮

0 1 1 0 1 0 1 1 1 1 1 1 1 1 1 1

1片 $2K \times 8$

位

ROM

$2K \times 8$ 位

$1K \times 8$ 位

RAM

2片 $1K \times 4$ 位

(2) 确定芯片的数量及类型

(3) 分配地址线

A_{15}	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	1片 ROM 2K × 8位
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	
0	1	1	0	0	1	1	1	1	1	1	
0	1	1	0	1	0	0	0	0	0	0	2片 RAM 1K × 4位
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	
0	1	1	0	1	0	1	1	1	1	1	

C B A

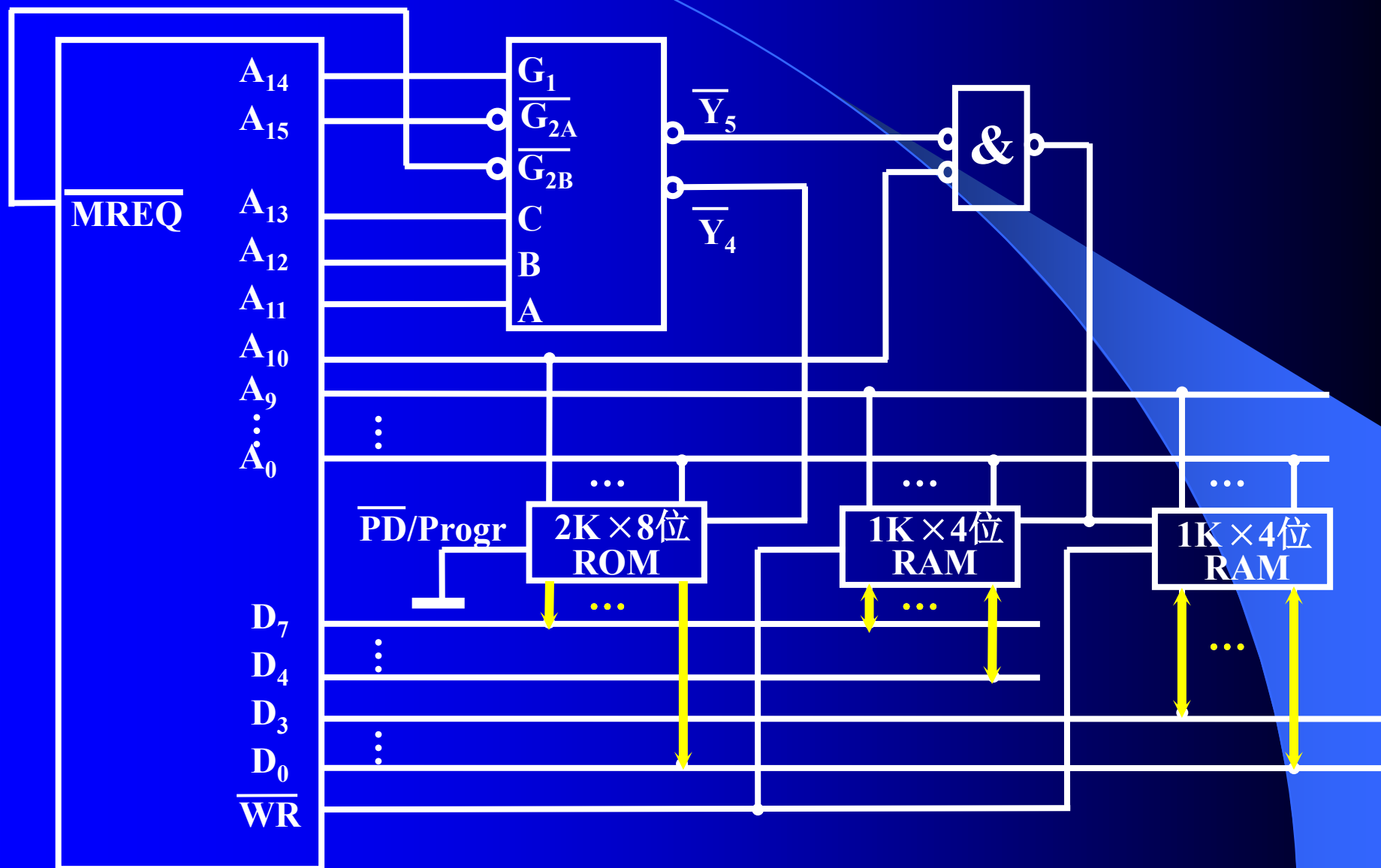
$A_{10} \sim A_0$ 接 2K × 8位 ROM 的地址线

$A_9 \sim A_0$ 接 1K × 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

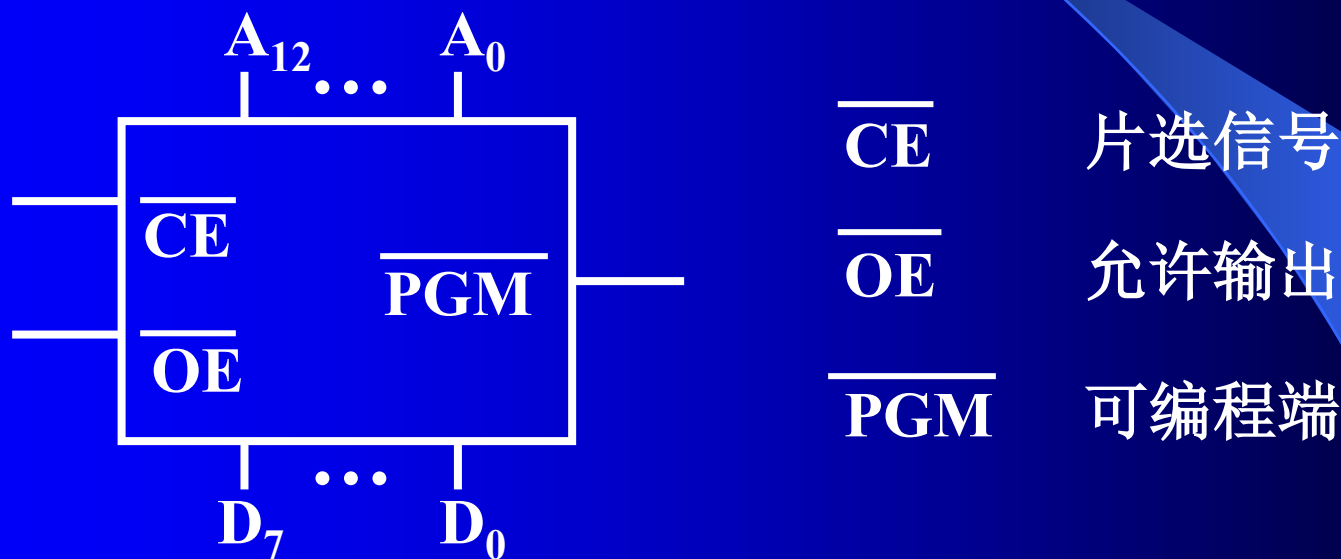
1片 4K×8位 ROM 2片 4K×8位 RAM

(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。 **4.2**
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~FFFFFH，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

海明码是具有一位纠错能力的编码

2. 海明码的组成

组成海明码的三要素

海明码的组成需增添 ? 位检测位

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

检测位的位置 ?

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

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”配置的海明码

解：∵ $n = 4$

根据 $2^k \geq n + k + 1$

得 $k = 3$

海明码排序如下：

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			

∴ 0101 的海明码为 **0100101**

练习1 按配偶原则配置 0011 的海明码 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 的海明码为 1000011

3. 海明码的纠错过程

4.2

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”配置的海明码
不出错时 $P_1=0, P_2=0, P_4=0$

例4.5 已知接收到的海明码为 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海明码

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 的海明码

配奇的海明码为 0101011

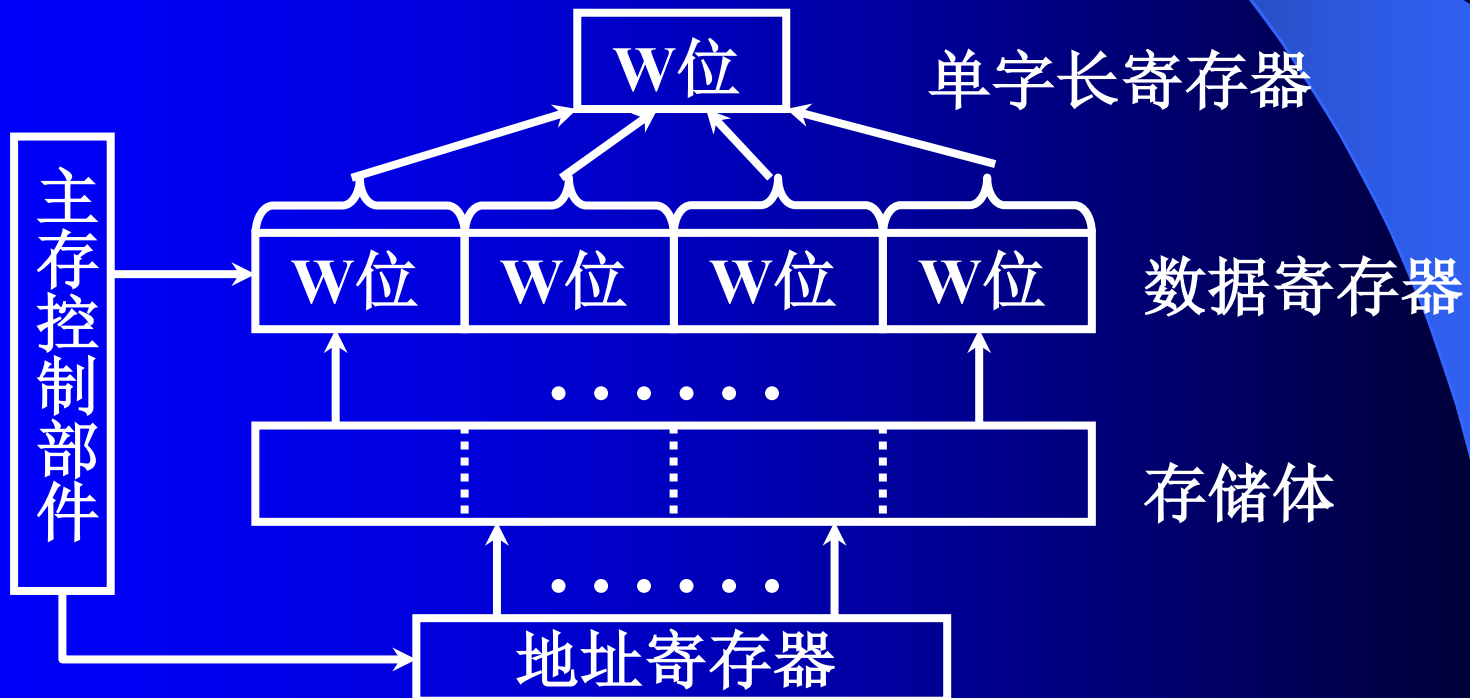
七、提高访存速度的措施

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

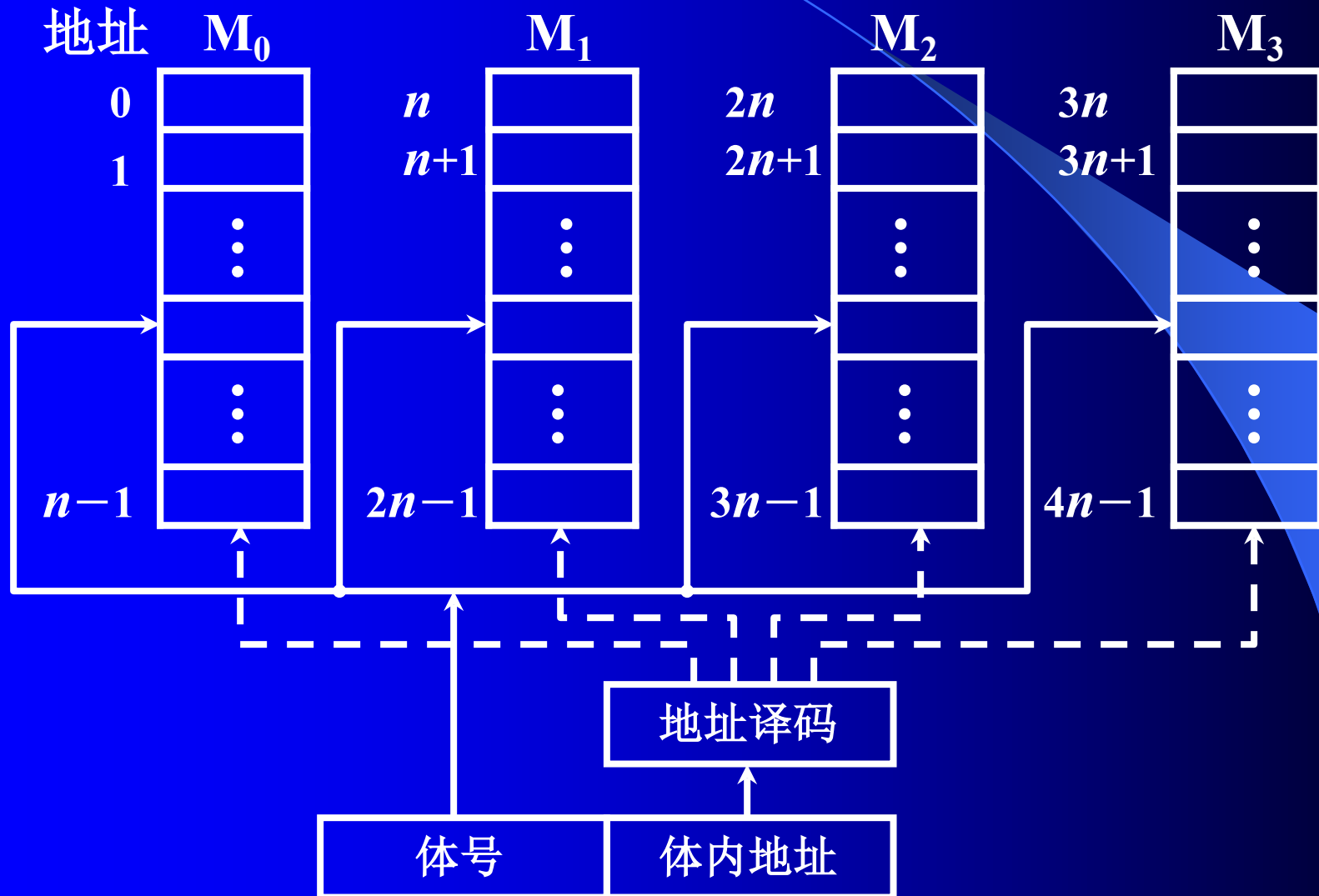


2. 多体并行系统

4.2

(1) 高位交叉

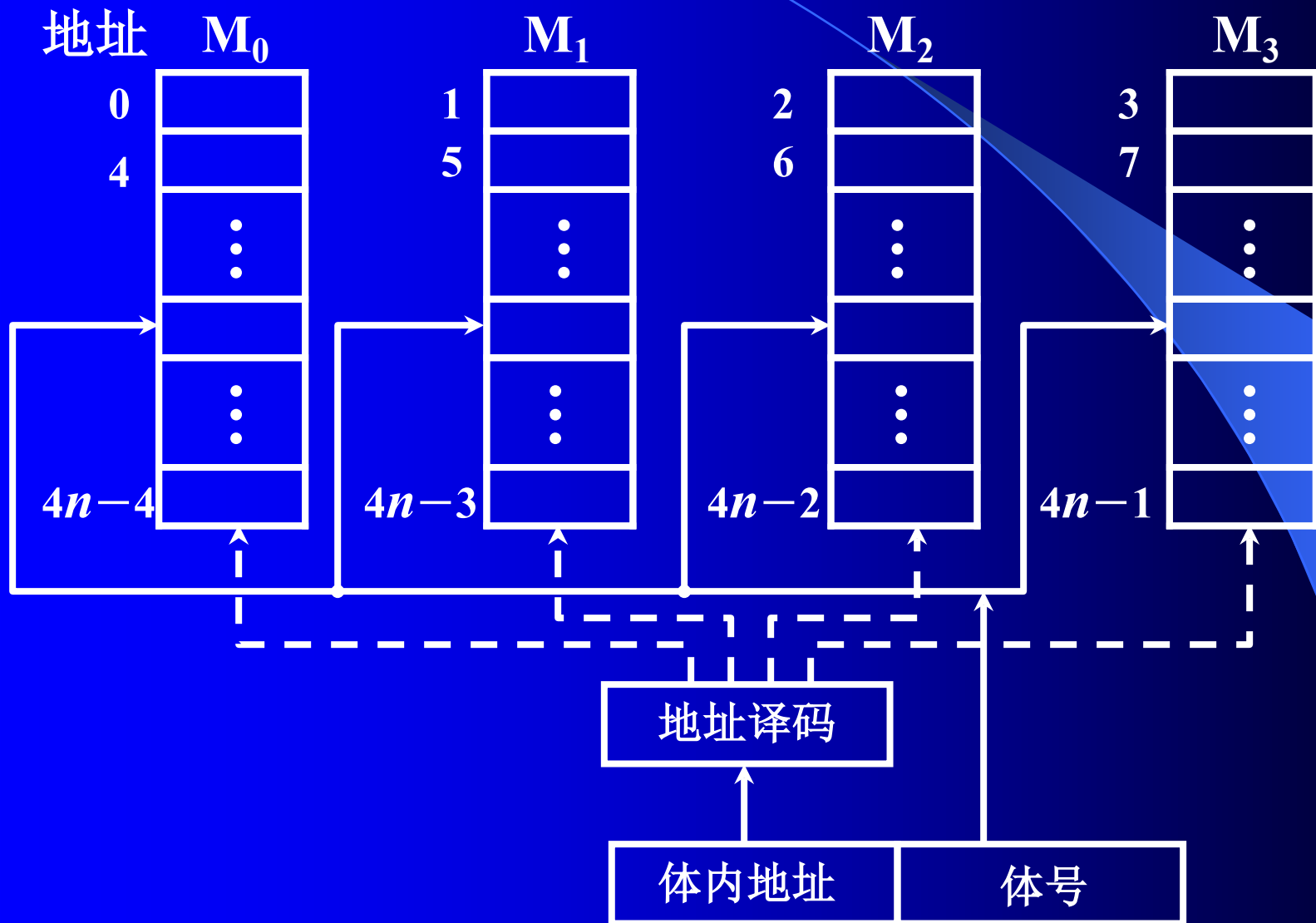
各个体并行工作



(2) 低位交叉

各个体轮流编址

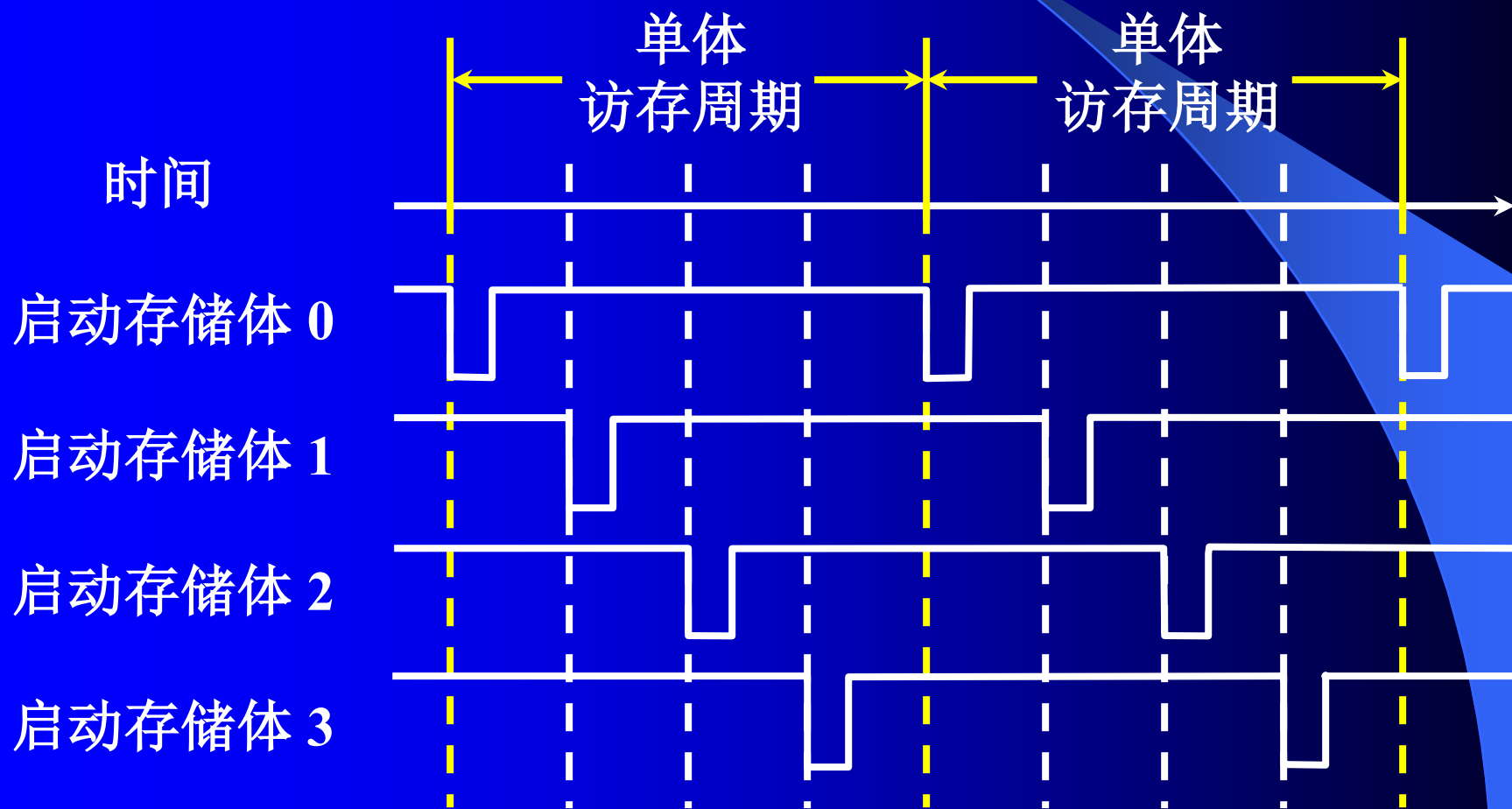
4.2



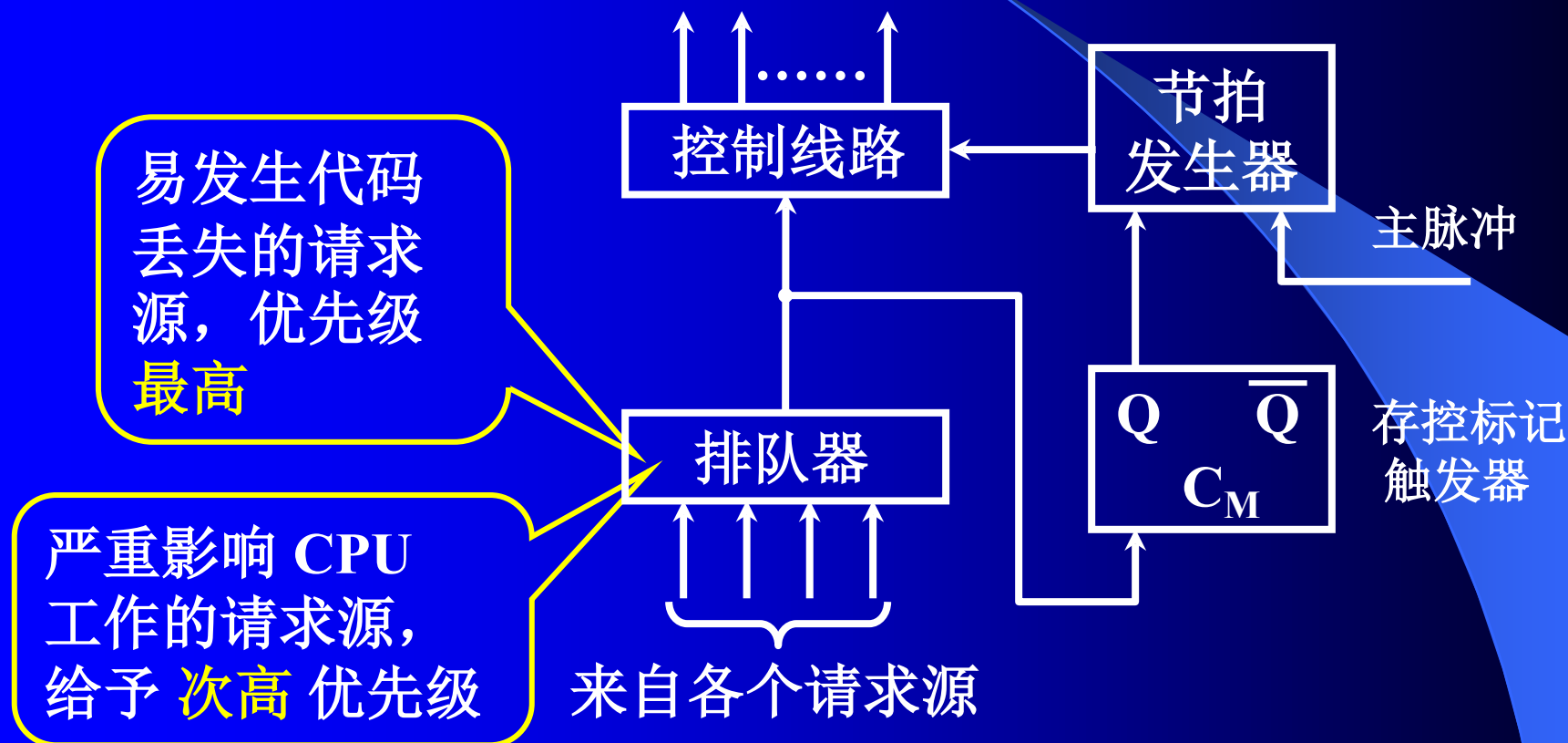
低位交叉的特点

4.2

在不改变存取周期的前提下，增加存储器的带宽



(3) 存储器控制部件（简称存控）



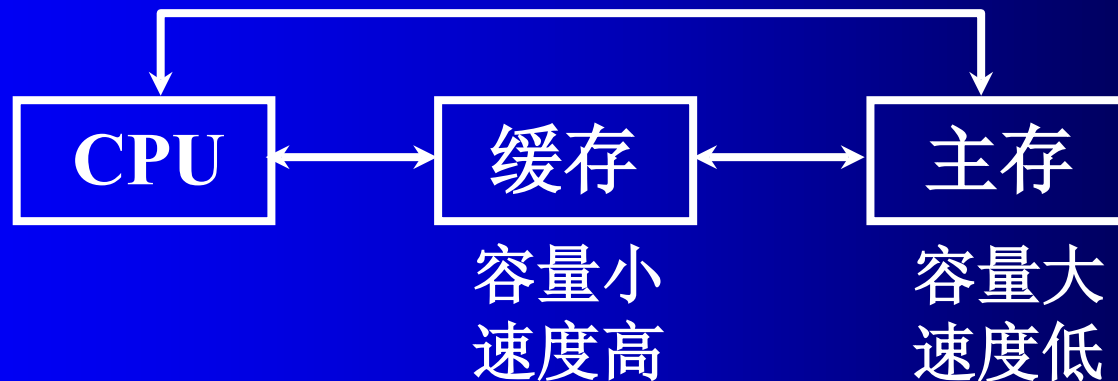
4.3 高速缓冲存储器

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

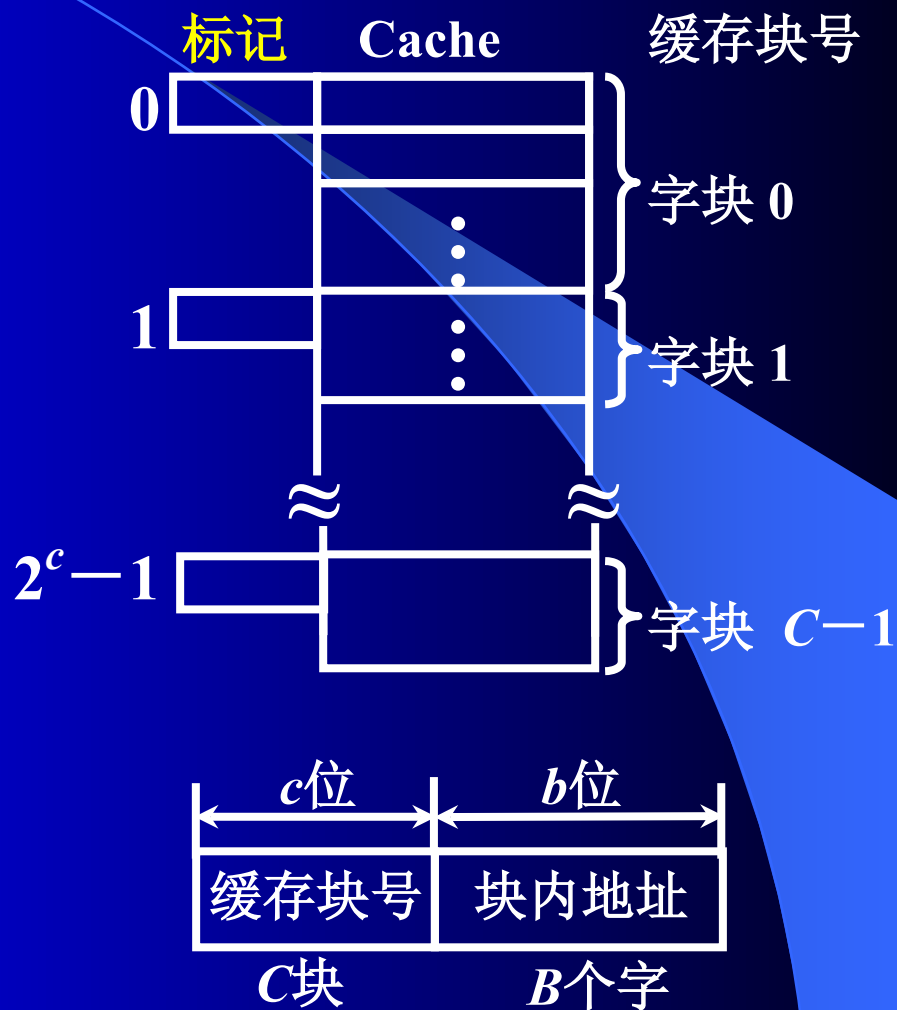
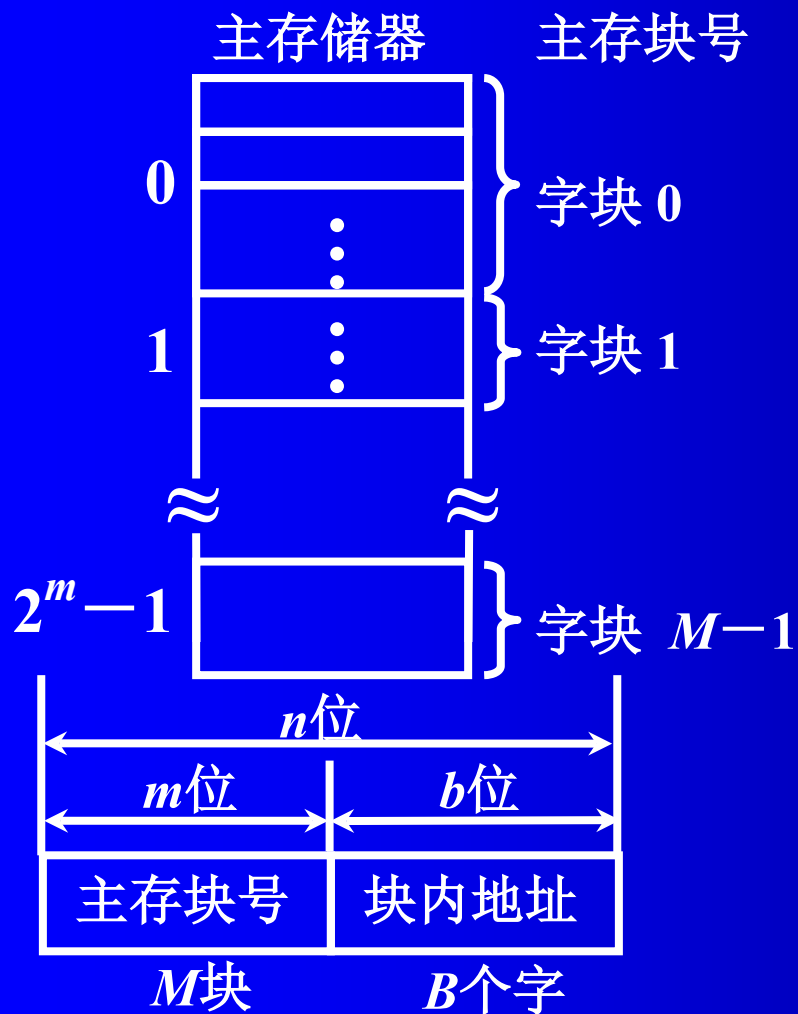


程序访问的局部性原理

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

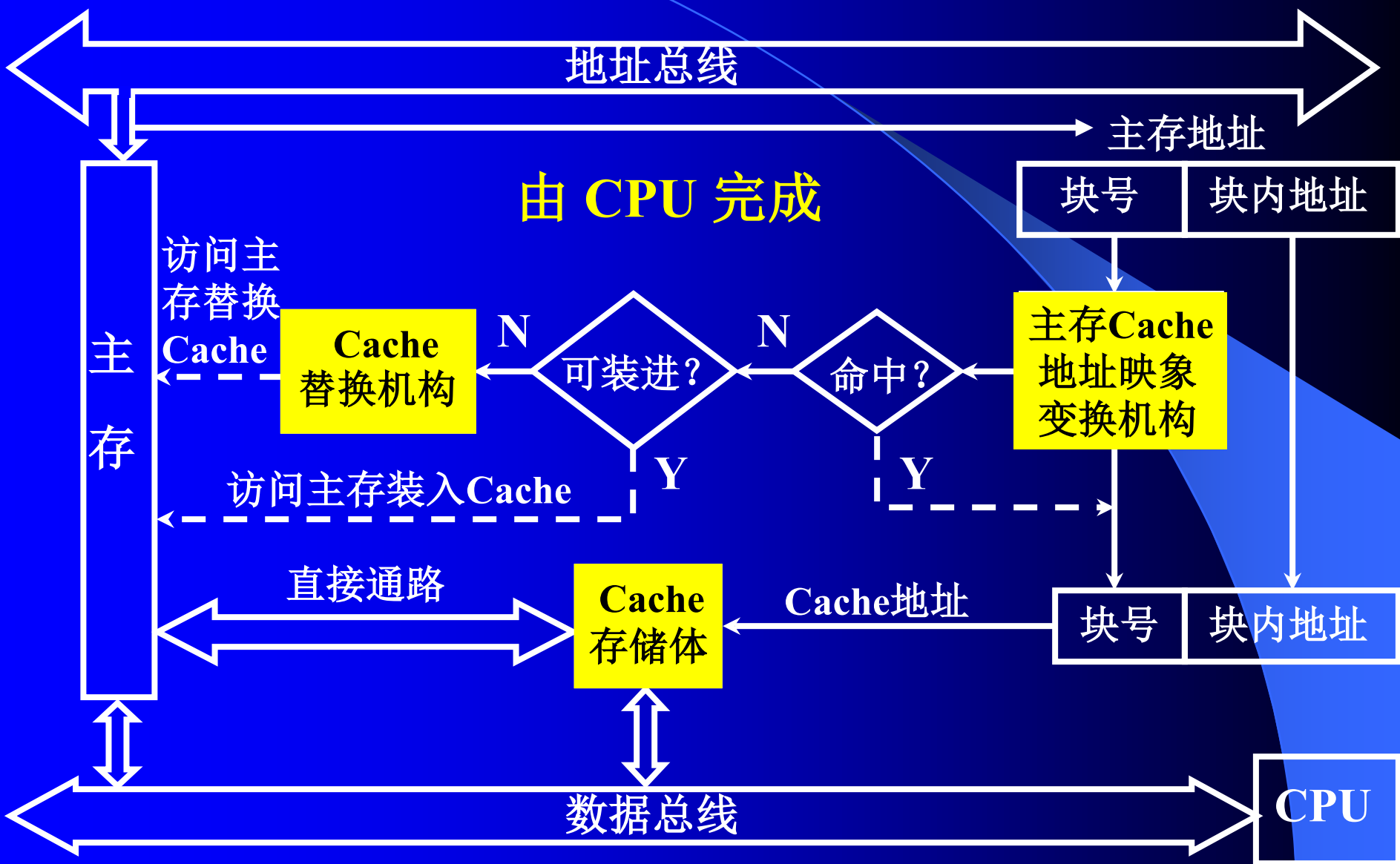
一般每块可取 4 至 8 个字

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)

3. Cache 的基本结构

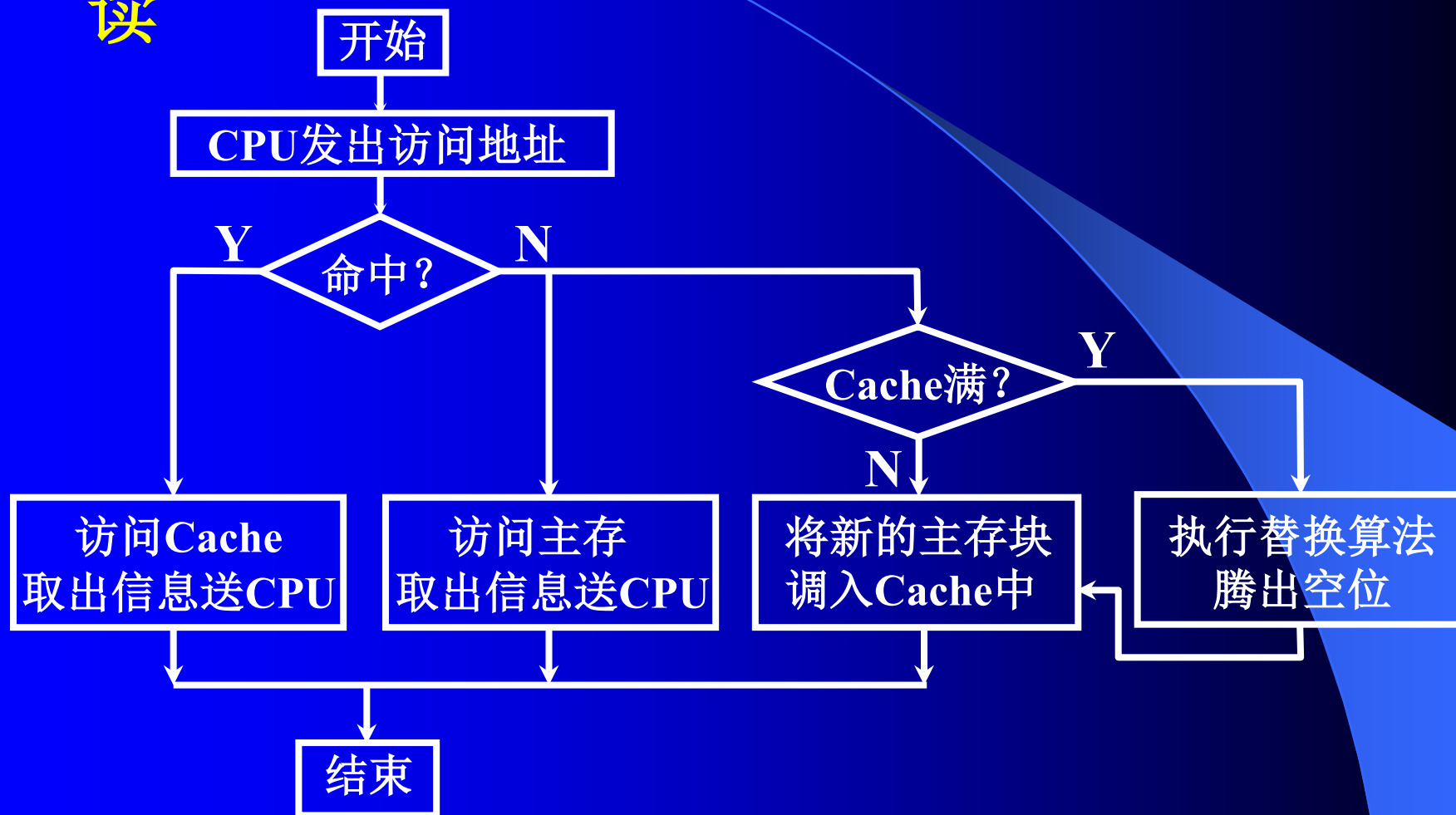
4.3



4. Cache 的读写操作

4.3

读



写

Cache 和主存的一致性

5. Cache 的改进

4.3

(1) 增加 Cache 的级数

片载（片内）Cache

片外 Cache

(2) 统一缓存和分开缓存

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
---------	-------------	-------------

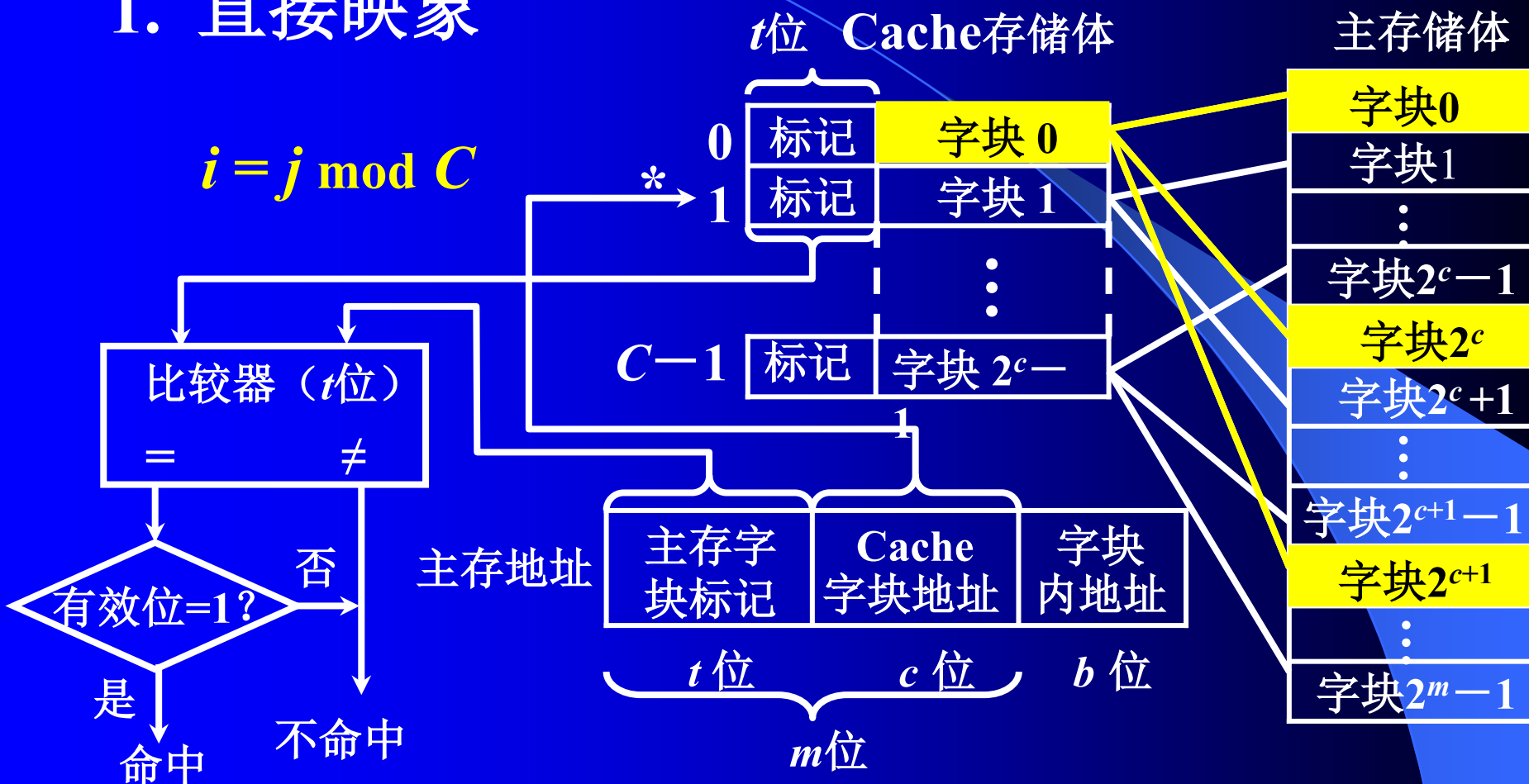
PowerPC620	32K 指令 Cache	32K 数据 Cache
------------	--------------	--------------

二、Cache — 主存的地址映象

4.3

1. 直接映象

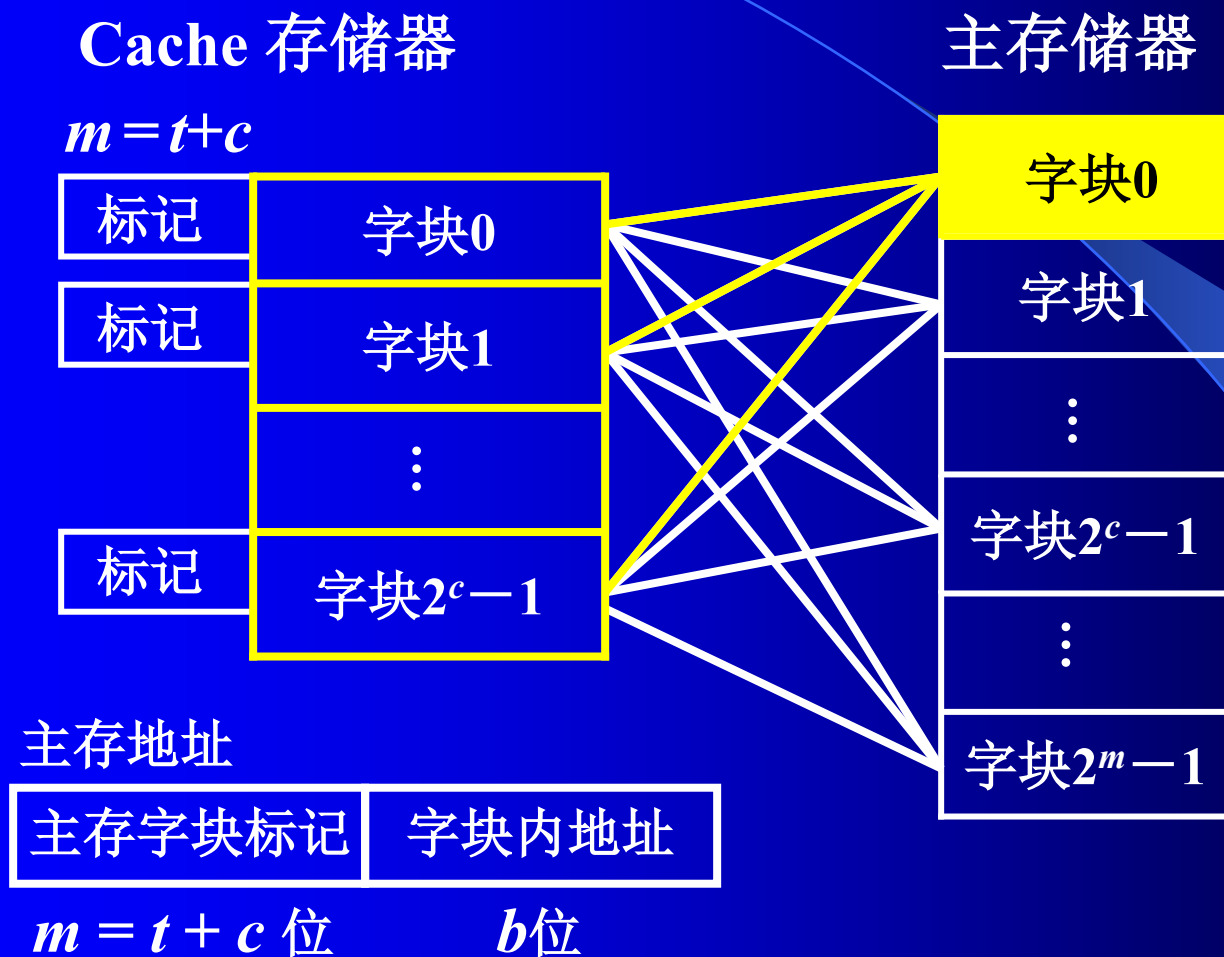
$$i = j \bmod C$$



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. 全相联映象

4.3



主存 中的 任一块 可以映象到 缓存 中的 任一块

3. 组相联映像

4.3

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

$$i = j \bmod Q$$

直接映像

某一主存块 j 按模 Q 映射到 缓存 的第 i 组中的 任一块

三、替换算法

1. 先进先出（FIFO）算法

2. 近期最少使用法（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 某一主存块能映射到某一缓存组中的任一块

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 { 寻址时间
磁头读写时间

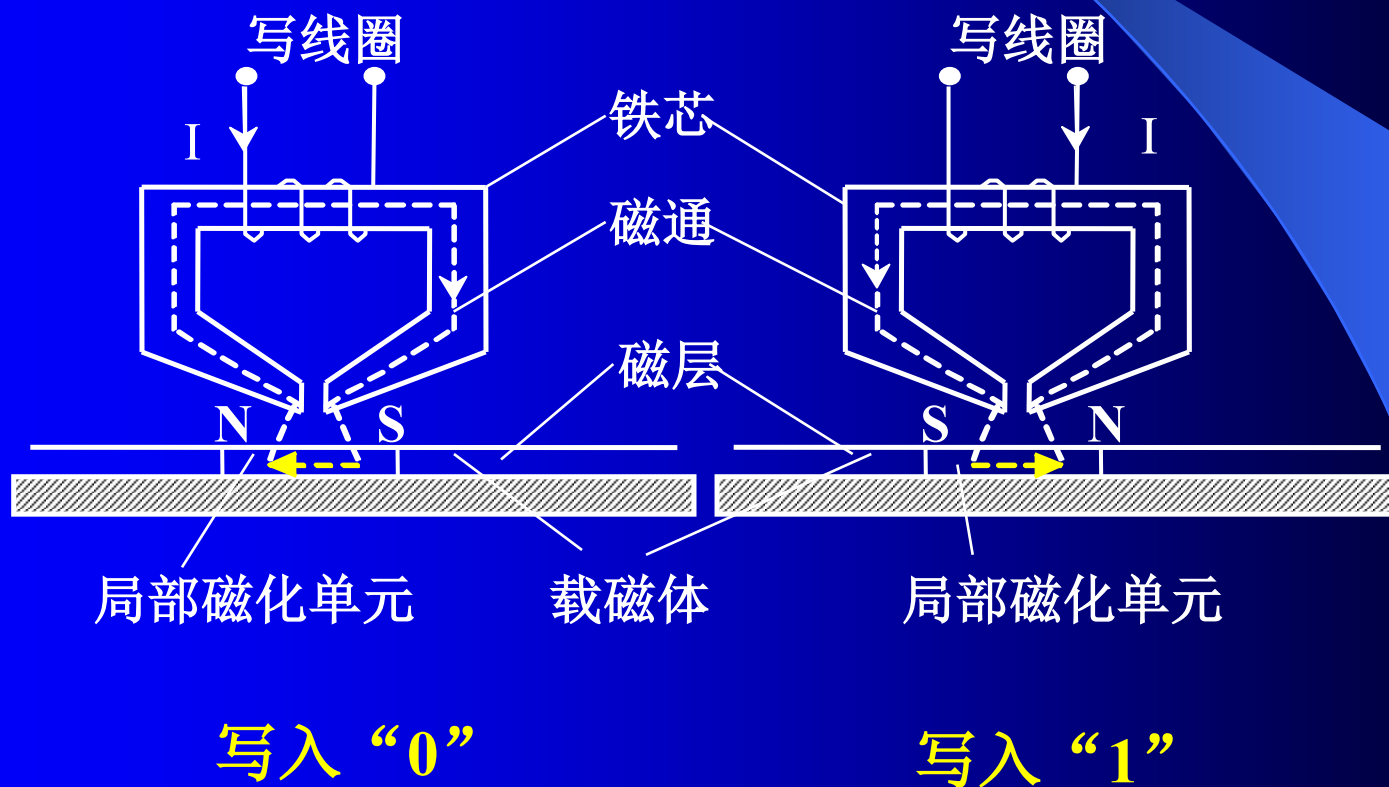
(4) 数据传输率 $D_r = D \times V$

(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

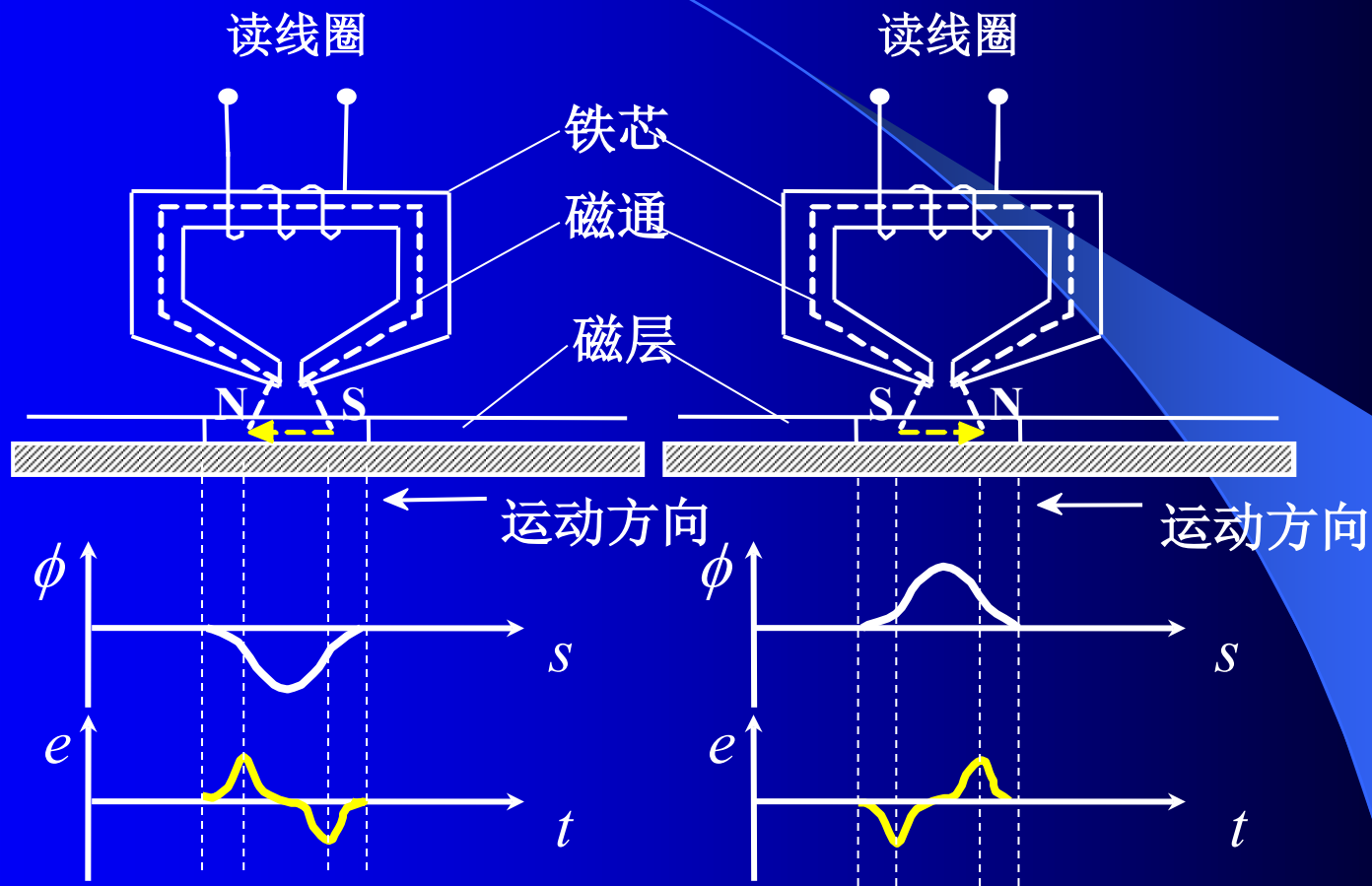
写



磁记录原理

4.4

读

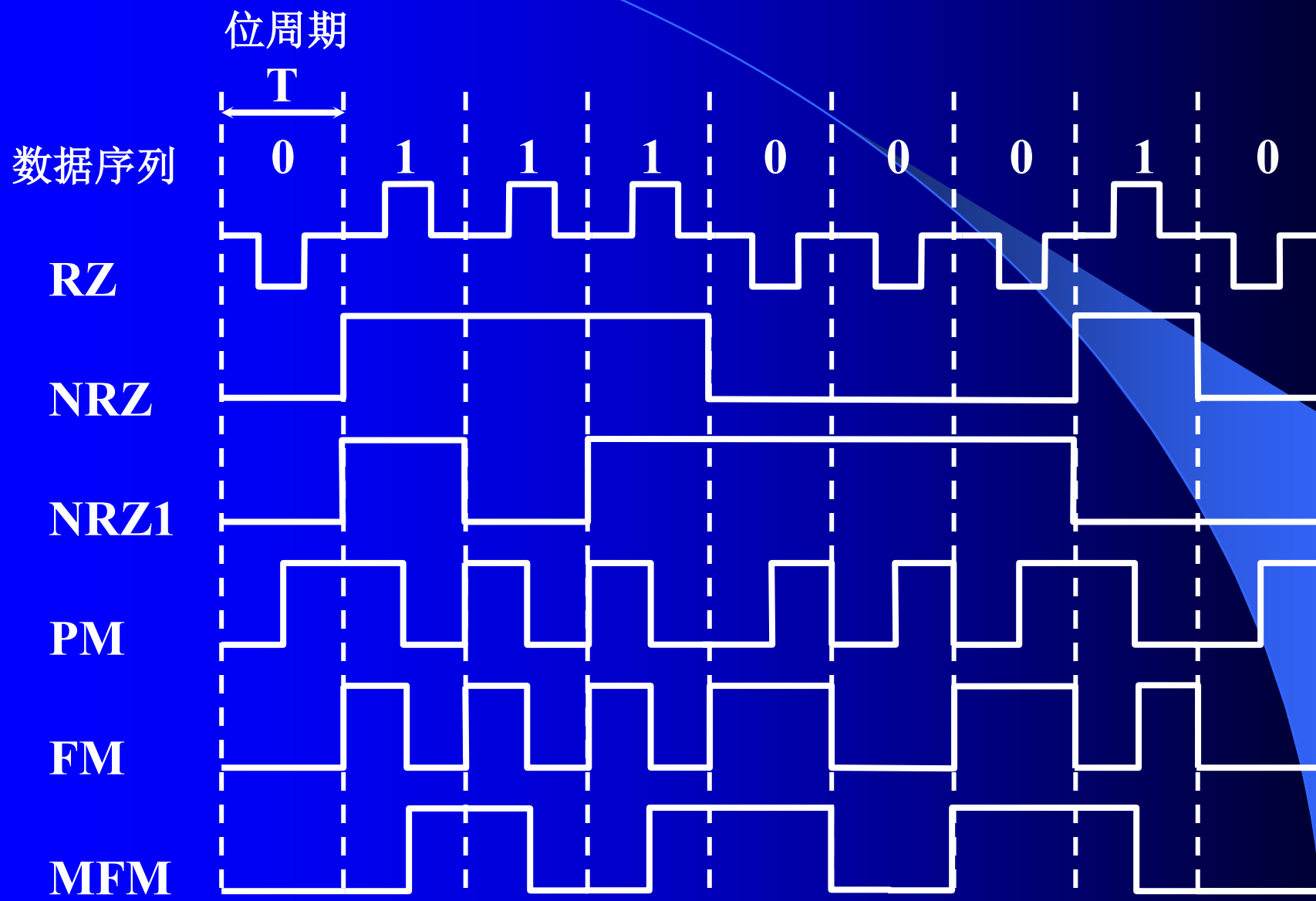


读出“0”

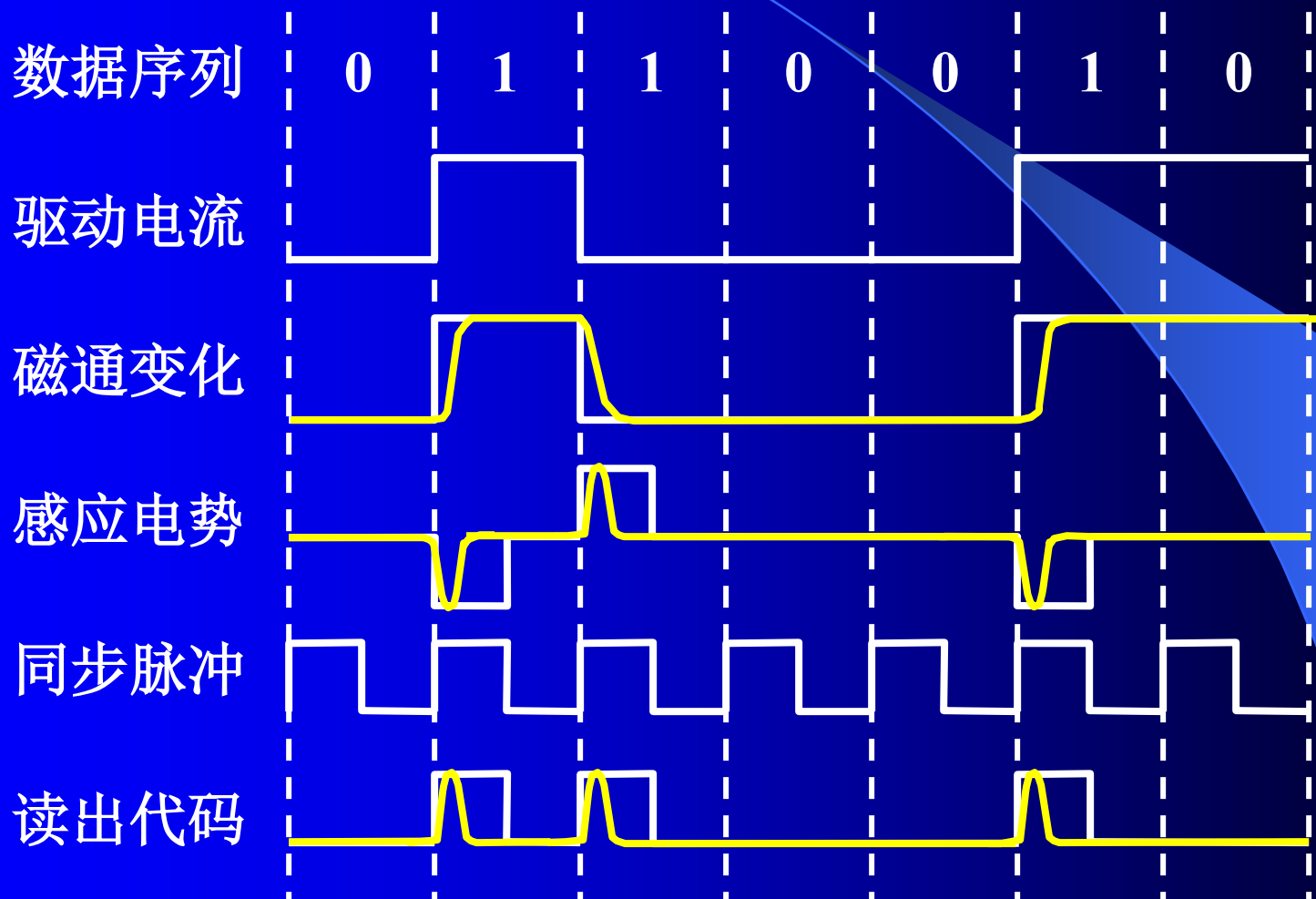
读出“1”

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

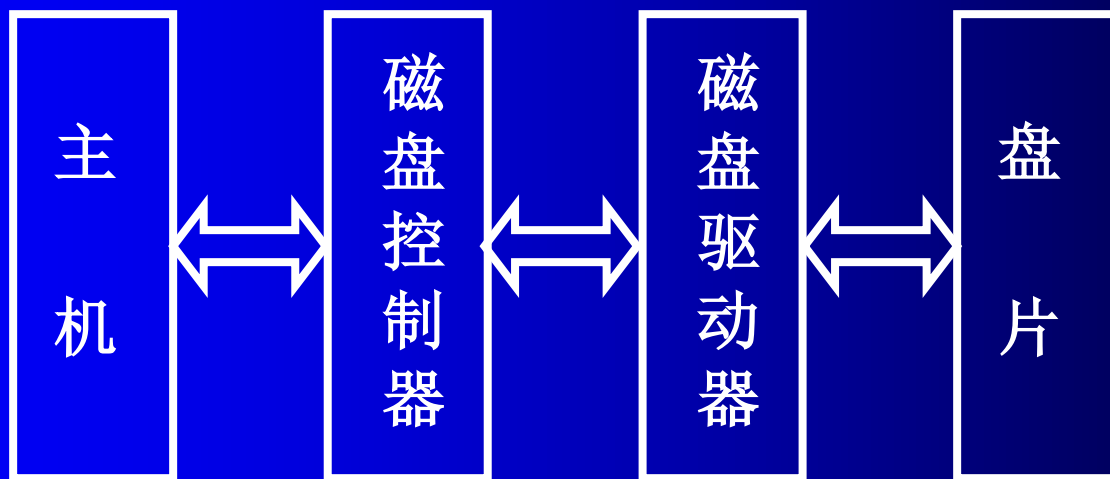
4.4

1. 硬磁盘存储器的类型

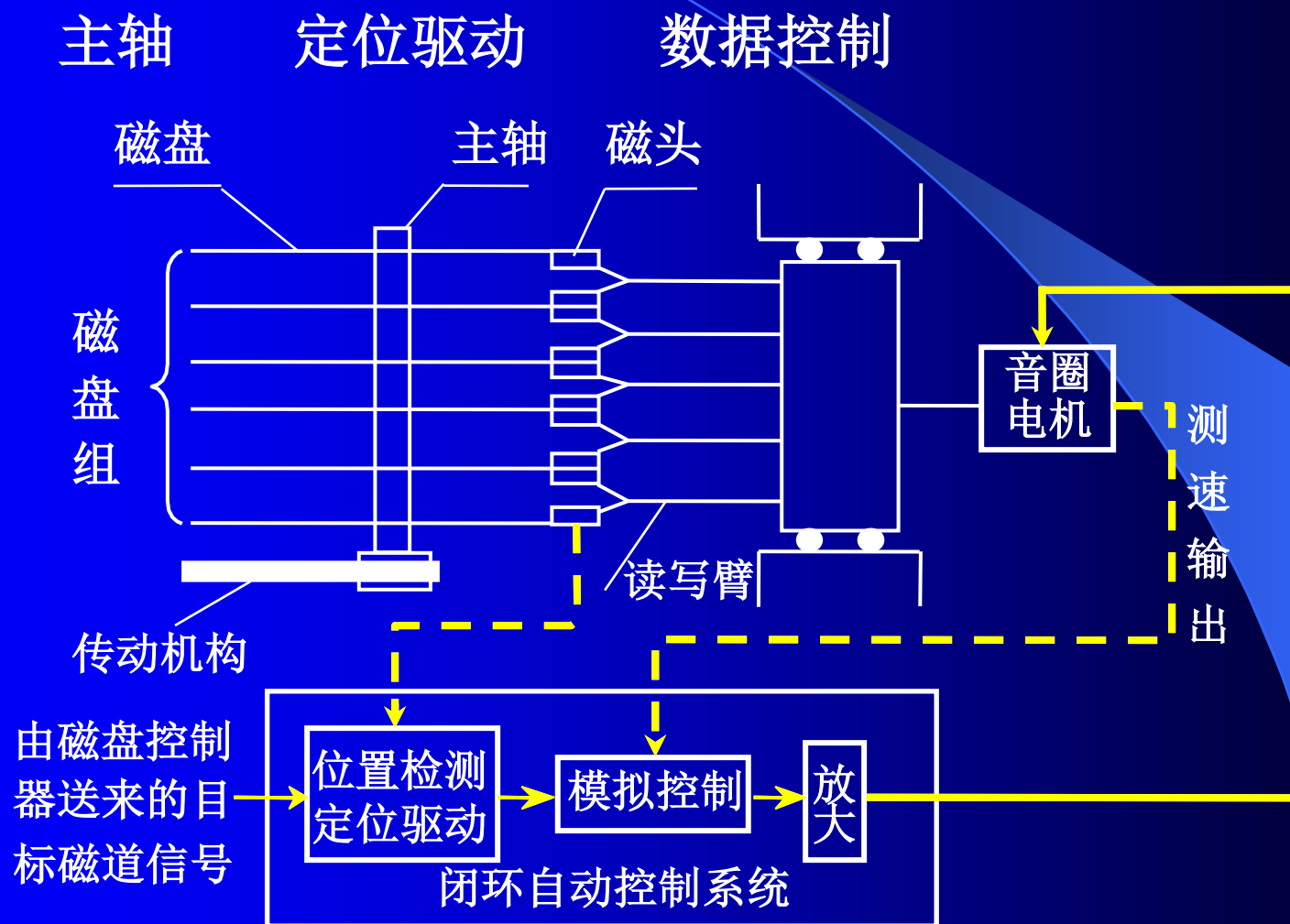
(1) 固定磁头和移动磁头

(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

- 接受主机发来的命令，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

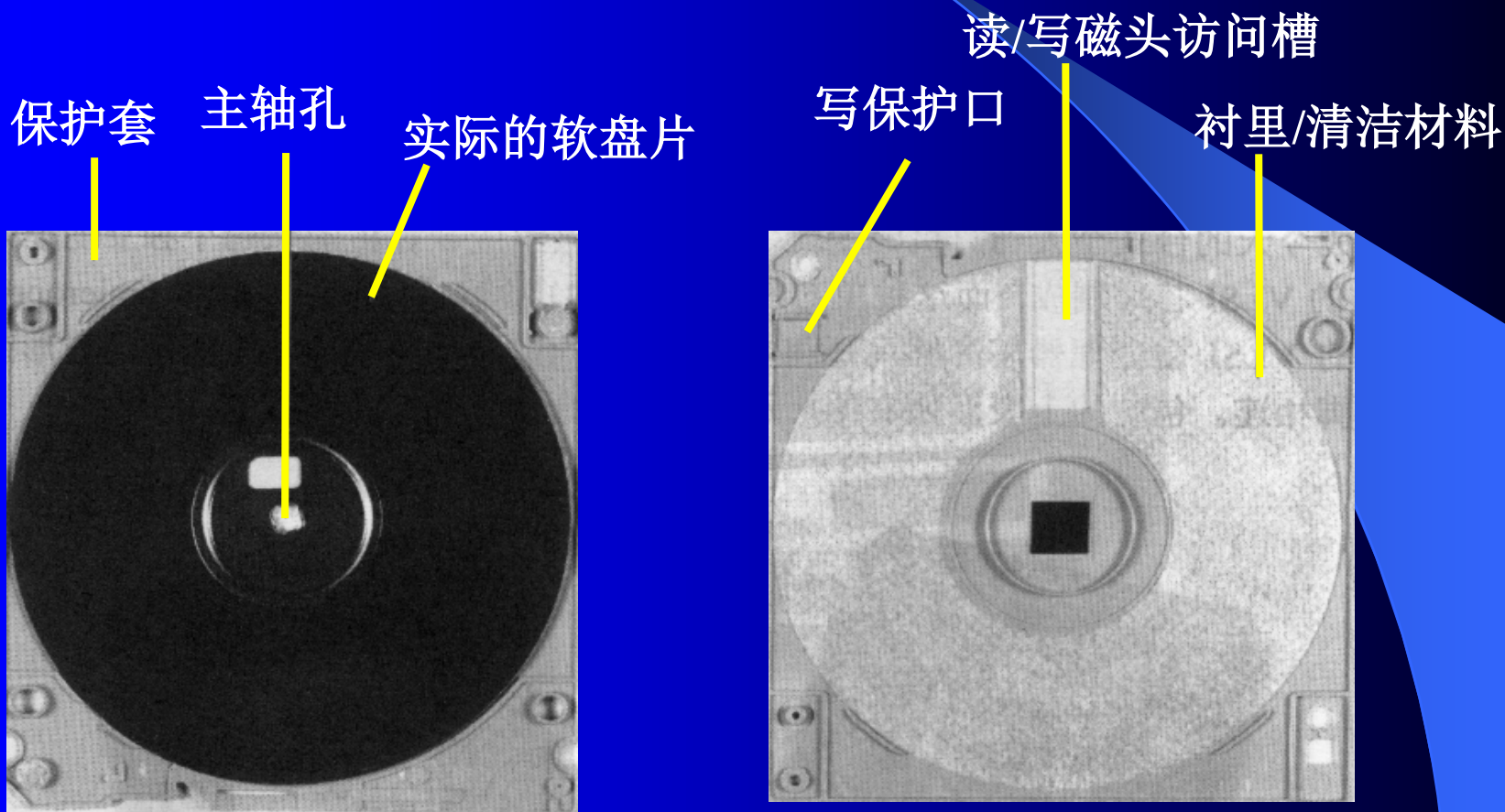
四、软磁盘存储器

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

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