第4章 存储器

> 课件来源：《计算机组成原理》课件第4章（Slide 94–205，共 112 页），本章为全书重点大章。一句话考试地位：本章是期末计算题与设计题的主产地——容量扩展与地址译码、Cache 映射与命中率、汉明码、DRAM 刷新四大计算考点几乎必考，概念对比题（SRAM/DRAM、ROM 家族、三种映射）覆盖面极广。

0. 本章考点清单

考点	常见题型	重要度
容量扩展（位/字/字位）与 CPU 连接、地址范围与片选译码	计算、综合设计	★★★
Cache 三种地址映射（直接/全相联/组相联）与地址字段划分	计算、填空	★★★
汉明码的组成与纠错（偶/奇校验）	计算	★★★
DRAM 刷新三方式（集中/分散/异步）、死区与死时间率	计算、简答	★★★
Cache 命中率、平均访问时间、系统效率	计算	★★★
SRAM/DRAM 存储原理、读写过程与对比	简答、选择	★★
ROM 家族（MROM/PROM/EPROM/EEPROM/Flash）分类与擦写方式	选择、填空	★★
存储器分类与层次结构（缓存–主存、主存–辅存/虚拟存储）	选择、填空	★★
芯片容量与地址线/数据线根数、末地址计算	填空、计算	★★
替换算法 FIFO/LRU 工作过程	模拟、计算	★★
多体并行（高位/低位交叉）与带宽、流水线时间	简答、计算	★★
主存技术指标（容量、存取时间/存取周期、带宽）	填空	★★
写直达/写回、Cache 结构改进	简答、选择	★★
SRAM 读/写时序参数（tRC、tA、tWC、tW 等）	填空	★
磁记录方式（RZ/NRZ/NRZ1/PM/FM/MFM）与 NRZ1 读出	波形分析	★
磁盘技术指标（容量、传输率、平均寻址时间）	计算	★
硬盘/软盘/光盘结构与特点	填空、选择	★

1. 存储器分类与层次结构（课件 p.95–99）

1.1 三种分类方式

按存储介质分（课件 p.95）：

介质	典型元件/材料	易失性（课件标注）
半导体存储器	TTL、MOS	易失
磁表面存储器	磁头、载磁体	非易失
磁芯存储器	硬磁材料、环状元件	非易失
光盘存储器	激光、磁光材料	非易失

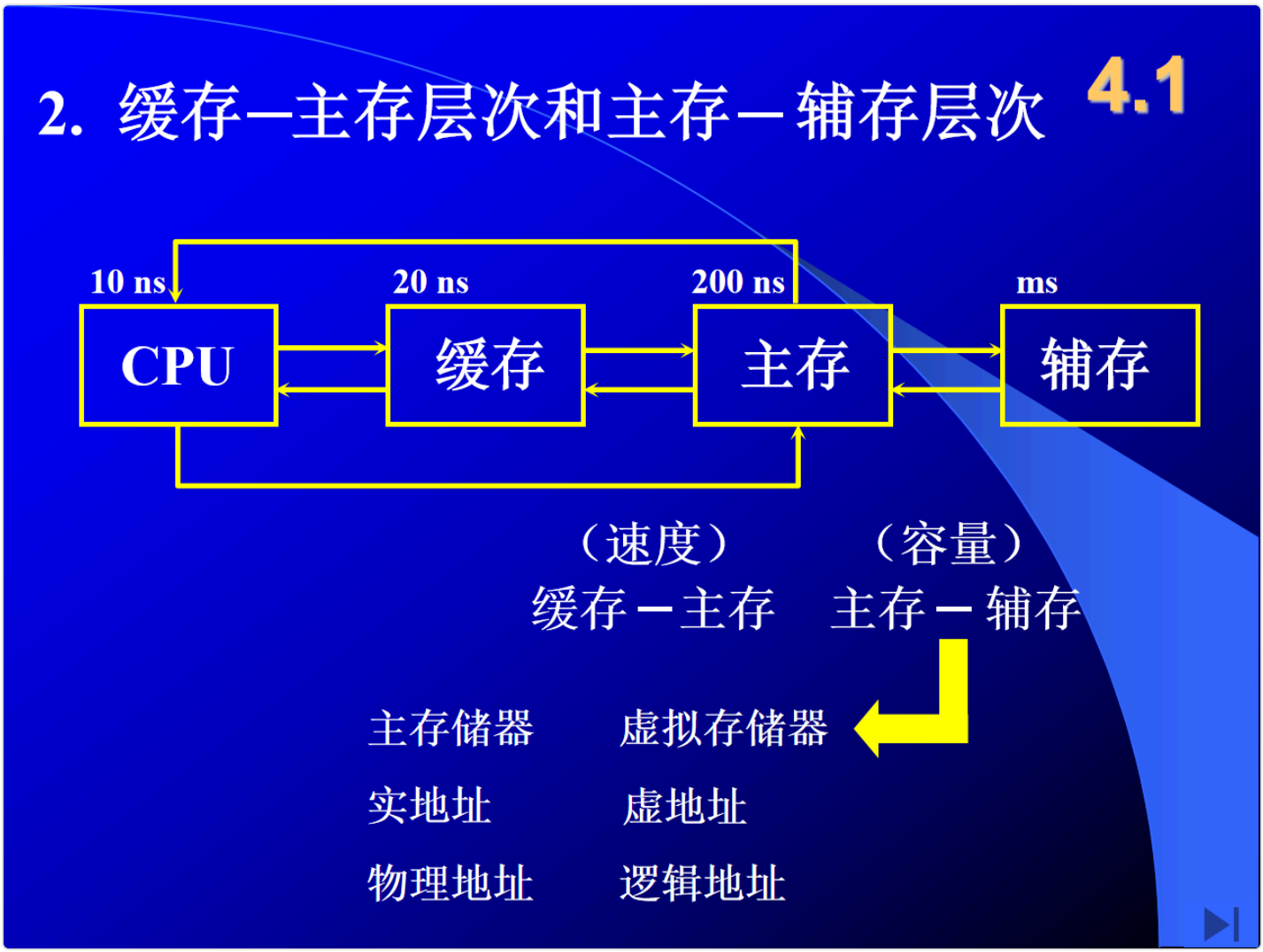
注意：课件把半导体存储器标为"易失"，指的是 RAM；半导体中的 ROM、Flash 断电不丢信息（见 p.97、p.155），做题时"半导体=易失"不能绝对化。

按存取方式分（课件 p.96）： - 存取时间与物理地址**无关**（随机访问）：随机存储器 RAM（程序执行过程中可读可写）、只读存储器 ROM（程序执行过程中只读）。 - 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）。

按在计算机中的作用分（课件 p.97）：主存储器（RAM：静态 RAM、动态 RAM；ROM：MROM、PROM、EPROM、EEPROM）、辅助存储器（磁盘、磁带、光盘）、高速缓冲存储器 Cache、Flash Memory。

1.2 层次结构

存储器三个主要特性：**速度、容量、价格/位**。从寄存器→缓存→主存→磁盘→光盘→磁带：速度由快到慢、容量由小到大、每位价格由高到低（课件 p.98）。



- **缓存–主存层次**：解决**速度**矛盾。课件给出典型速度：CPU 10 ns、缓存 20 ns、主存 200 ns、辅存 ms 级（课件 p.99）。
- **主存–辅存层次**：解决**容量**矛盾，构成**虚拟存储器**；程序员使用**虚地址（逻辑地址）**，实际访存用**实地址（物理地址）**（课件 p.99）。
- 易错警示：两个层次的目的不要记反——"缓存–主存=速度，主存–辅存=容量"。虚拟存储器在本课件中只讲到概念与地址名称（p.99），页式/段式等细节课件未展开，按课件范围复习。

2. 主存组成、地址分配与技术指标（课件 p.100–103）

2.1 基本组成与 CPU 联系

主存由**存储体、译码器、驱动器、MAR（地址寄存器）、MDR（数据寄存器）、读写电路、控制电路**组成（课件 p.100）。CPU 通过**地址总线**（单向，CPU→主存）送地址到 MAR，通过**数据总线**（双向）经 MDR 读写数据，读/写命令由控制电路执行（课件 p.101）。

2.2 存储单元地址的分配（课件 p.102）

设地址线 24 根： - 按**字节**寻址：2^{24} = 16 M 个字节； - 字长 16 位按**字**寻址：16 M ÷ 2 = 8 M 个字； - 字长 32 位按**字**寻址：16 M ÷ 4 = 4 M 个字。

字节地址与字地址的对应关系有两种约定（课件 p.102）： - **高位字节地址为字地址**（大端）：字的高字节存放在低字节地址，字地址即高字节地址； - **低位字节地址为字地址**（小端）：字的低字节存放在低字节地址，字地址即低字节地址。

例：16 位数据 1234H 存于字节地址 0000H、0001H：大端时 (0000H)=12H、(0001H)=34H；小端时 (0000H)=34H、(0001H)=12H。

2.3 主存技术指标（课件 p.103）

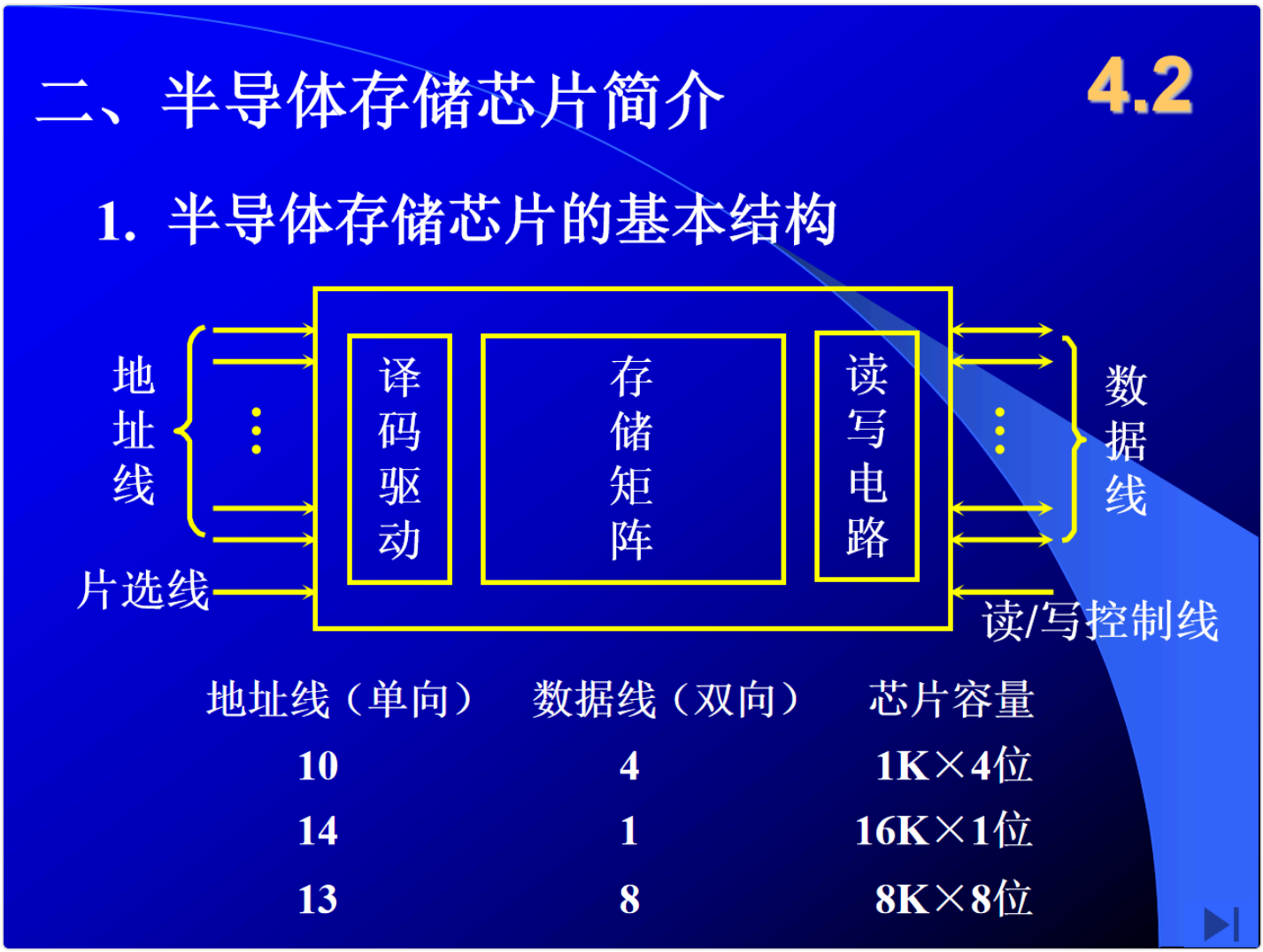
指标	含义
存储容量	主存存放二进制代码的**总位数**（也可用"字数×字长"或字节数表示）
存取时间（访问时间）	读出时间 / 写入时间，启动一次存储器操作到完成的时间
存取周期	连续两次**独立**的存储器操作（读或写）所需的**最小间隔时间**；分读周期、写周期
存储器带宽	单位时间传送的信息量，单位**位/秒**

易错警示：存取周期 ≠ 存取时间。存取周期包含操作后的恢复时间，故存取周期 ≥ 存取时间（补充常识）；带宽 = 每次存取的数据位数 ÷ 存取周期（补充常识）。

3. 半导体存储芯片结构与译码驱动（课件 p.104–108）

3.1 基本结构与引脚

芯片内部三大块：**译码驱动、存储矩阵、读写电路**；对外有**地址线（单向）、数据线（双向）、片选线、读/写控制线**（课件 p.104）。



芯片容量 = 2^(地址线根数) × 数据线根数（位）（课件 p.104）：

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

控制线（课件 p.105）：**$\overline{CS}$ / $\overline{CE}$ 片选**（低电平选中芯片）；**$\overline{WE}$ 读/写控制**（低电平写、高电平读）；**$\overline{OE}$ 允许输出（允许读）**。

3.2 片选线的作用（课件 p.106）

用 16K×1 位芯片组成 64K×8 位存储器：共 **32 片** = 4 组 × 8 片（每组 8 片位扩展凑 8 位数据，4 组字扩展凑 64K）。地址为 65 535（FFFFH）时，**第 4 组那 8 片**的片选有效。片选信号由高位地址译码产生，决定"哪一组芯片工作"。

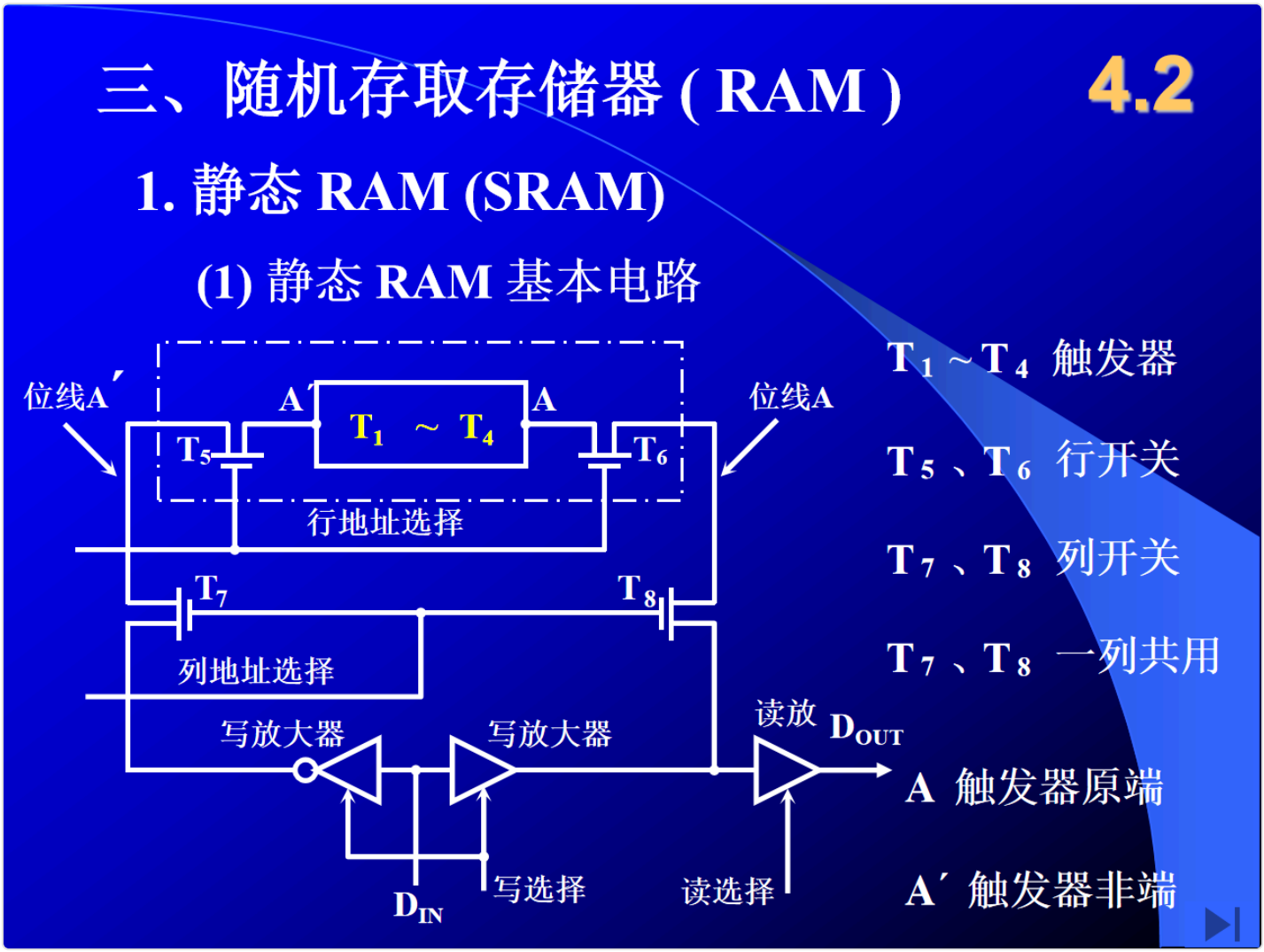
3.3 译码驱动方式

- **线选法**（课件 p.107）：一个译码器直接产生全部字线。如 16×8 矩阵用 4 根地址线译出 16 根字线，每根字线选中一行（一个字）。结构简单，但字数多时译码输出线太多。
- **重合法**（课件 p.108）：X、Y 两个译码器分别选行、列，交叉选中一个单元。如 32×32 矩阵用 10 根地址线分成 5+5，X、Y 各译出 32 根选择线（共 64 根），而不是 1024 根字线。大容量芯片普遍采用重合法。

易错警示：地址线根数由**字数**决定（$\log_2$字数），与数据位数无关；16K×1 是 14 根地址线、1 根数据线，不要按总位数 16K 算成别的值。

4. 静态 RAM（SRAM）（课件 p.109–132）

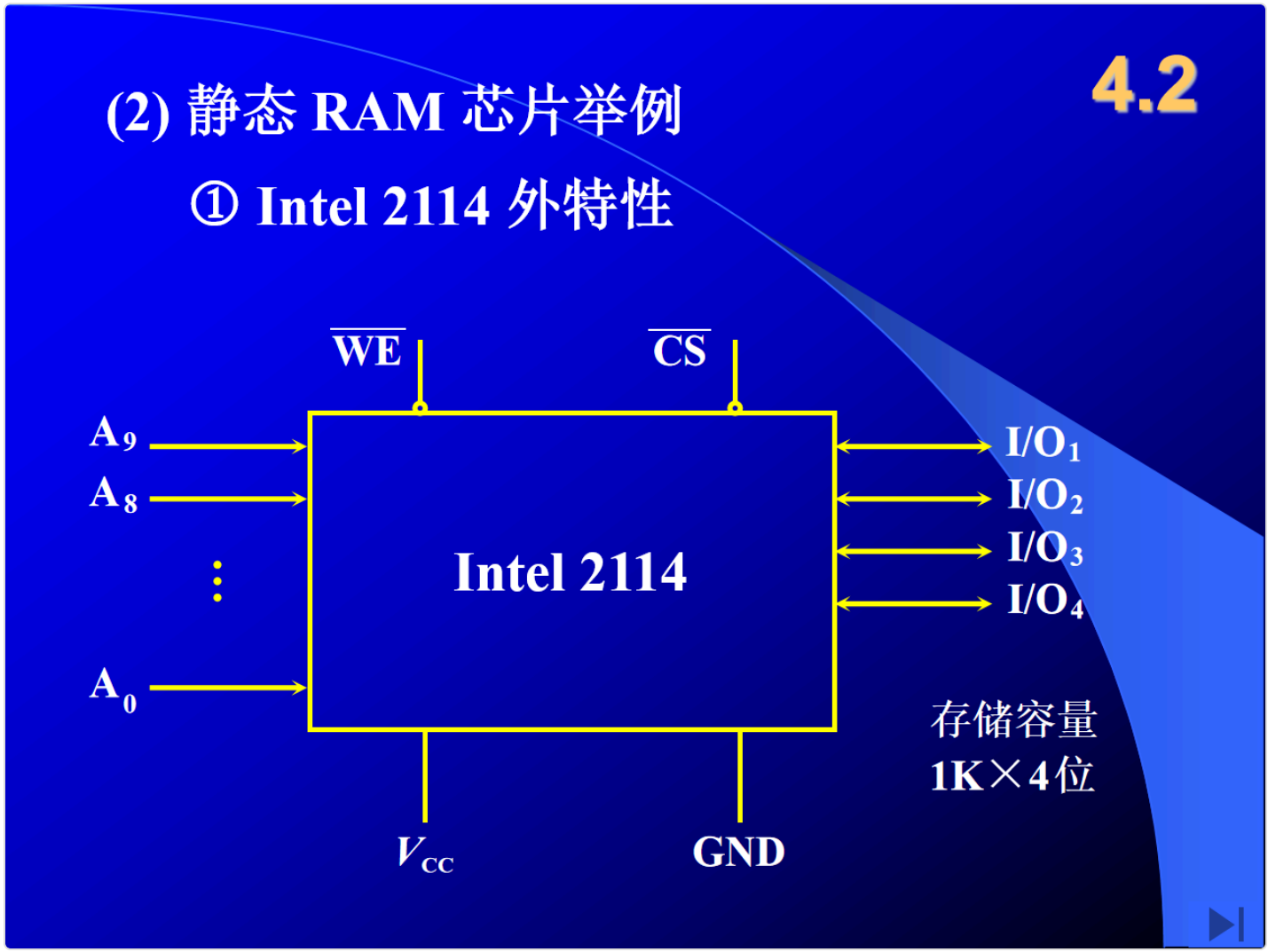
4.1 基本电路（六管单元）



- **T1～T4 构成触发器**（双稳态，存 0/1）；A 为触发器原端、A′ 为非端（课件 p.109）。
- **T5、T6 为行开关**，受行地址选择控制；**T7、T8 为列开关**，受列地址选择控制，**一列共用**（课件 p.109）。
- 位线 A、A′ 经列开关接读放（DOUT）与两个写放（DIN）。

读操作（课件 p.110）：行选→T5、T6 开；列选→T7、T8 开；读选择有效→触发器状态经位线、读放送出 DOUT。读出不破坏原状态。 **写操作**（课件 p.111）：行选、列选同上；写选择有效→DIN 经两个写放（其中一个反相）分别强制位线 A′、A，把触发器置成所需状态。

4.2 Intel 2114 芯片



- 容量 **1K×4 位**；地址线 **A9～A0**（10 根），数据线 **I/O1～I/O4**（4 根，双向），控制 **$\overline{WE}$、$\overline{CS}$**，电源 VCC、GND（课件 p.112）。
- 内部 **64×64 矩阵**：行地址 A8～A3（6 位）译出 64 根行线；列地址 A9、A2～A0（4 位）译出 16 根列选择线，每根列线同时选中 4 位（分属 I/O1～I/O4 四组），16×4=64 列（课件 p.113–130 读/写过程动画）。

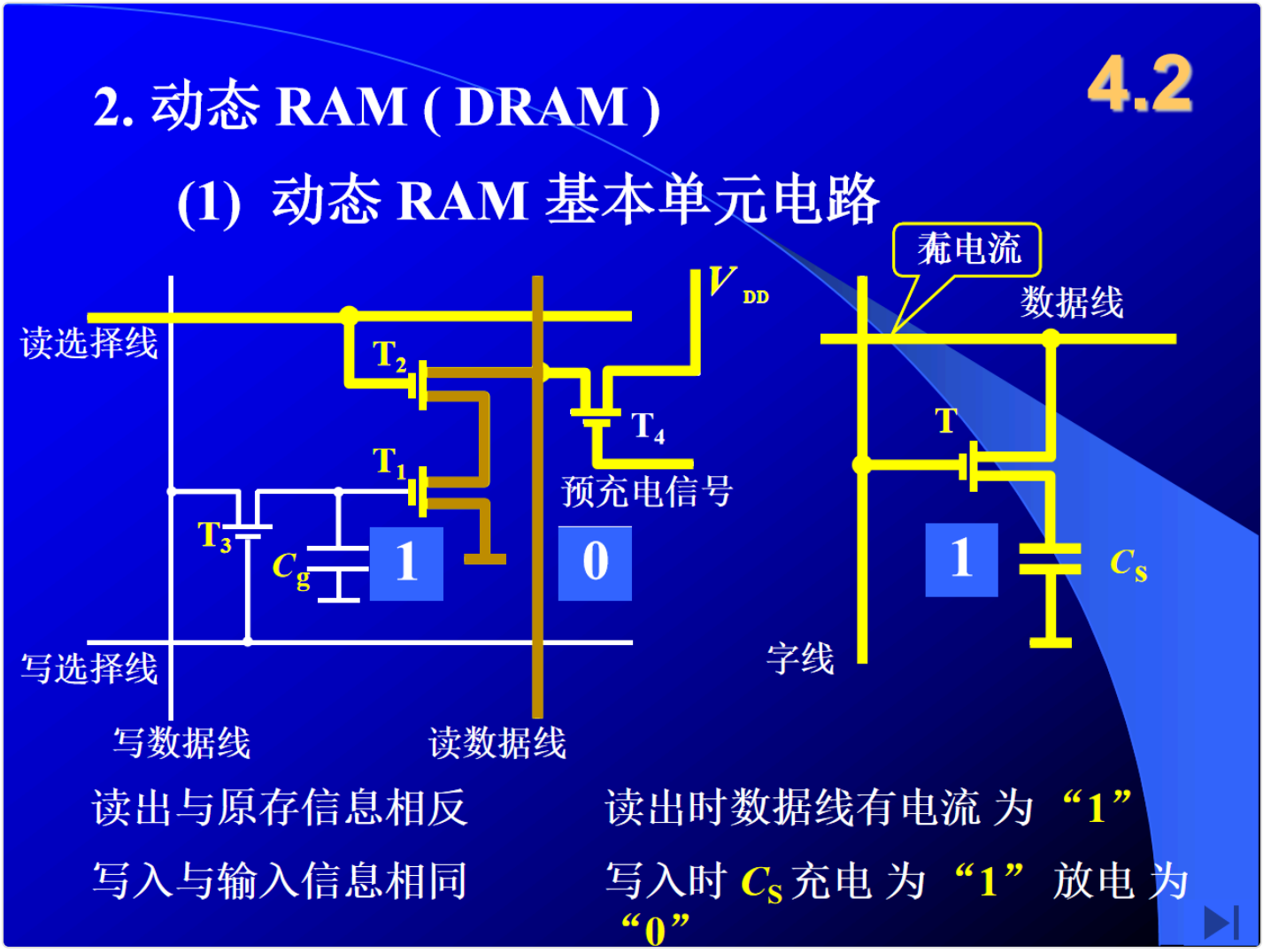
4.3 读/写时序（课件 p.131–132）

参数	含义
tRC 读周期	地址有效 → 下一次地址有效
tA 读时间	地址有效 → 数据稳定
tCO	片选有效 → 数据稳定
tOTD	片选失效 → 输出高阻
tOHA	地址失效后的数据维持时间
tWC 写周期	地址有效 → 下一次地址有效
tW 写时间	写命令 $\overline{WE}$ 的有效时间
tAW	地址有效 → 片选有效的滞后时间
tWR	片选失效 → 下一次地址有效
tDW	数据稳定 → $\overline{WE}$ 失效
tDH	$\overline{WE}$ 失效后的数据维持时间

易错警示：tA 从**地址有效**算起，tCO 从**片选有效**算起，两者起点不同；写操作要求数据在 $\overline{WE}$ 失效前已稳定（tDW）、失效后仍维持（tDH）。

5. 动态 RAM（DRAM）与刷新（课件 p.133–151）

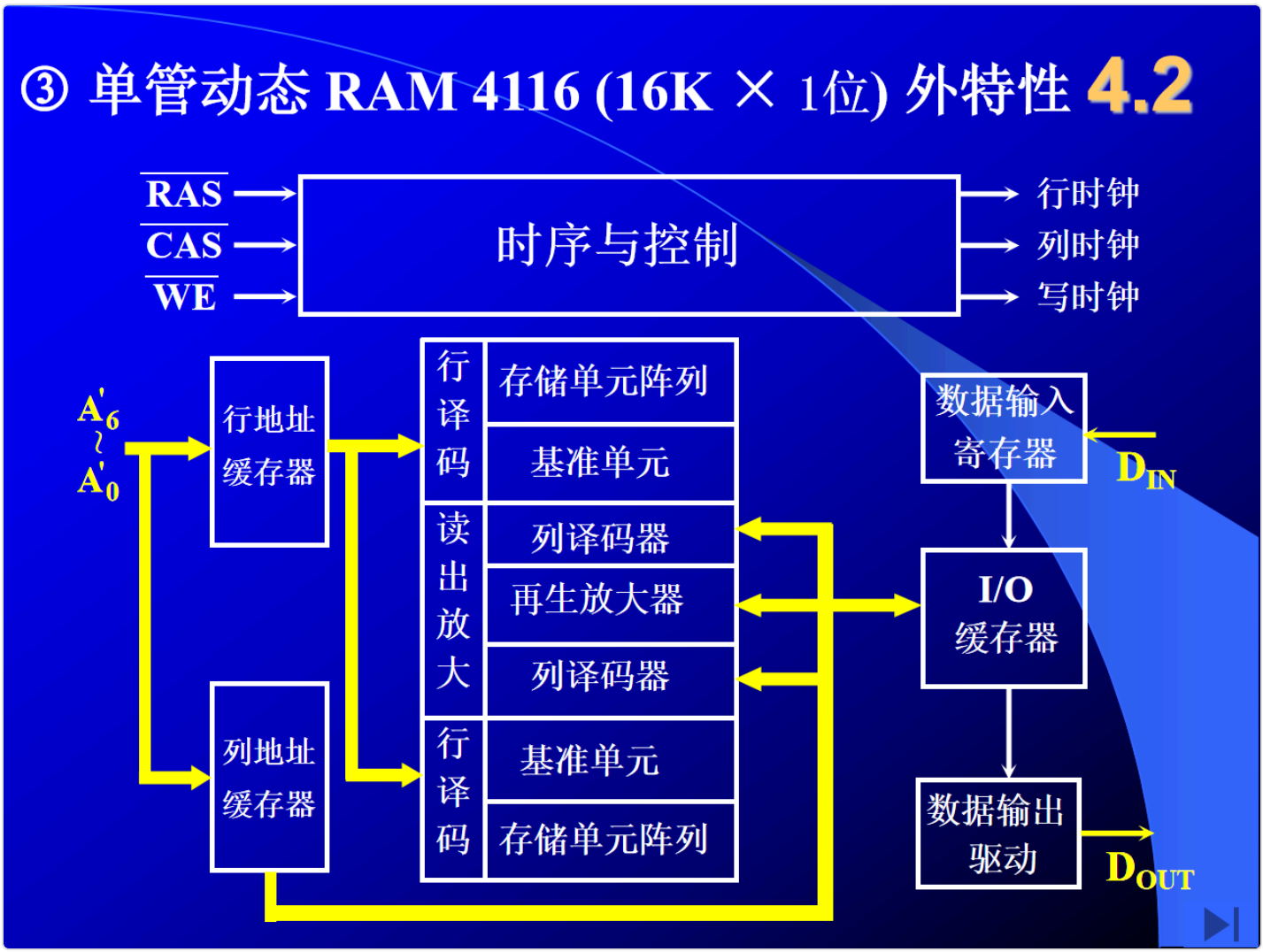
5.1 基本单元电路



- **三管单元**（课件 p.133 左）：靠栅极电容 Cg 存电荷；**读出与原存信息相反**，**写入与输入信息相同**；有读选择线、写选择线、读/写数据线、预充电信号。
- **单管单元**（课件 p.133 右）：一个管 T + 电容 Cs；**写入时 Cs 充电为"1"、放电为"0"**；**读出时数据线有电流为"1"**。单管读出的电荷分享会破坏原存信息，需再生放大器恢复（补充常识，课件 p.144 给出再生放大器）。

5.2 芯片举例

- **Intel 1103（三管）**：32×32 矩阵（1K×1），行地址 A9～A5、列地址 A4～A0，带刷新放大器，读/写分别走读选择线、写选择线（课件 p.134–143）。
- **Intel 4116（单管，16K×1 位）**：



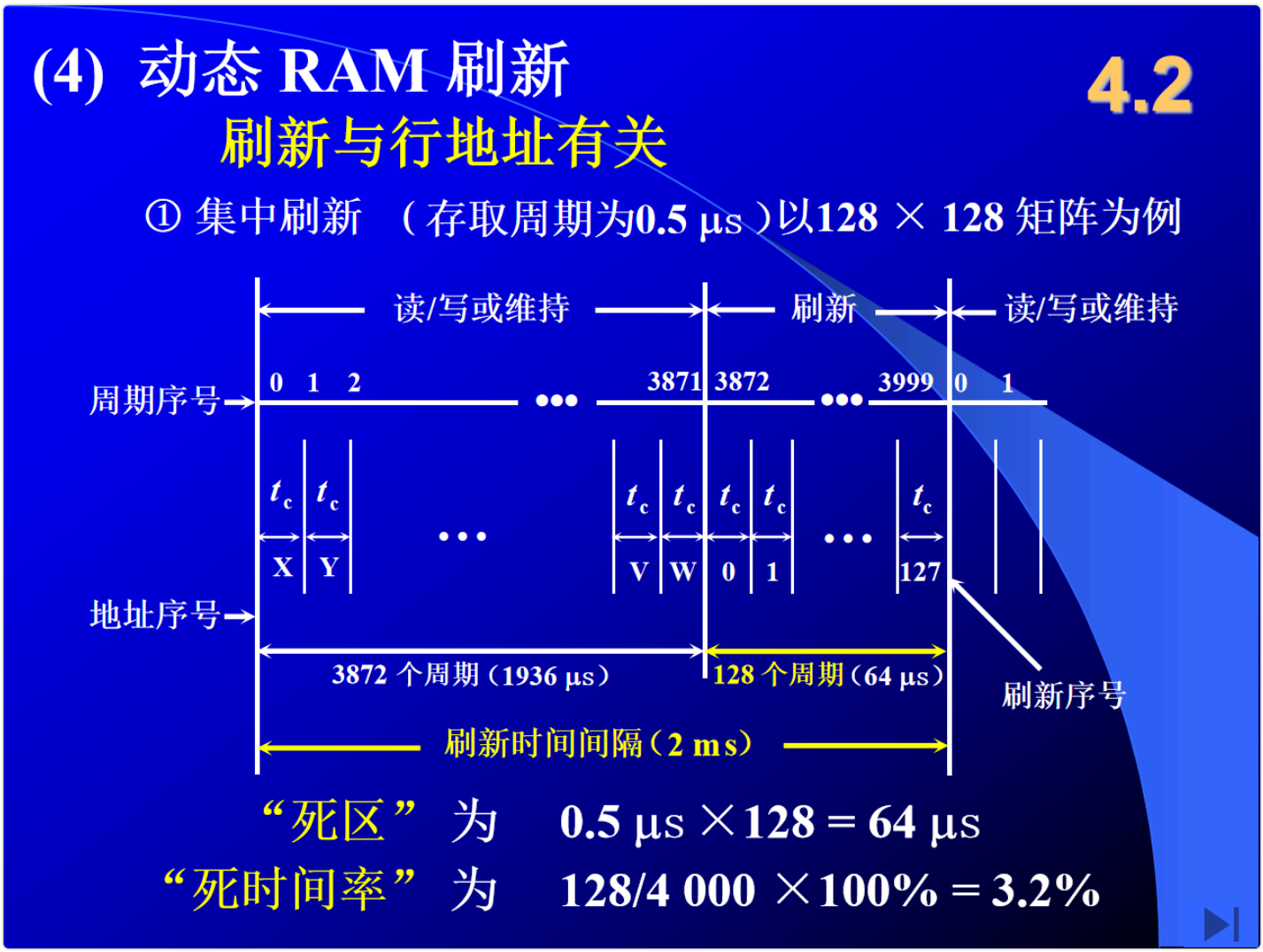
只有 **7 根地址引脚 A′6～A′0**，14 位地址**分两次**送入：行地址经 RA$\overline{S}$ 锁入行地址缓存器，列地址经 CA$\overline{S}$ 锁入列地址缓存器（地址复用，省引脚）；内部 **128×128 矩阵**，含基准单元、再生放大器、读出放大；控制线 RA$\overline{S}$、CA$\overline{S}$、W$\overline{E}$；数据 DIN、DOUT 分开（课件 p.144–146）。

5.3 DRAM 时序（课件 p.147）

行、列地址分开传送： - 读时序：行地址→RA$\overline{S}$ 有效；W$\overline{E}$ 有效（**高**）；列地址→CA$\overline{S}$ 有效；DOUT 有效。 - 写时序：行地址→RA$\overline{S}$ 有效；W$\overline{E}$ 有效（**低**）；DIN 有效；列地址→CA$\overline{S}$ 有效。

5.4 DRAM 刷新（重点计算考点）

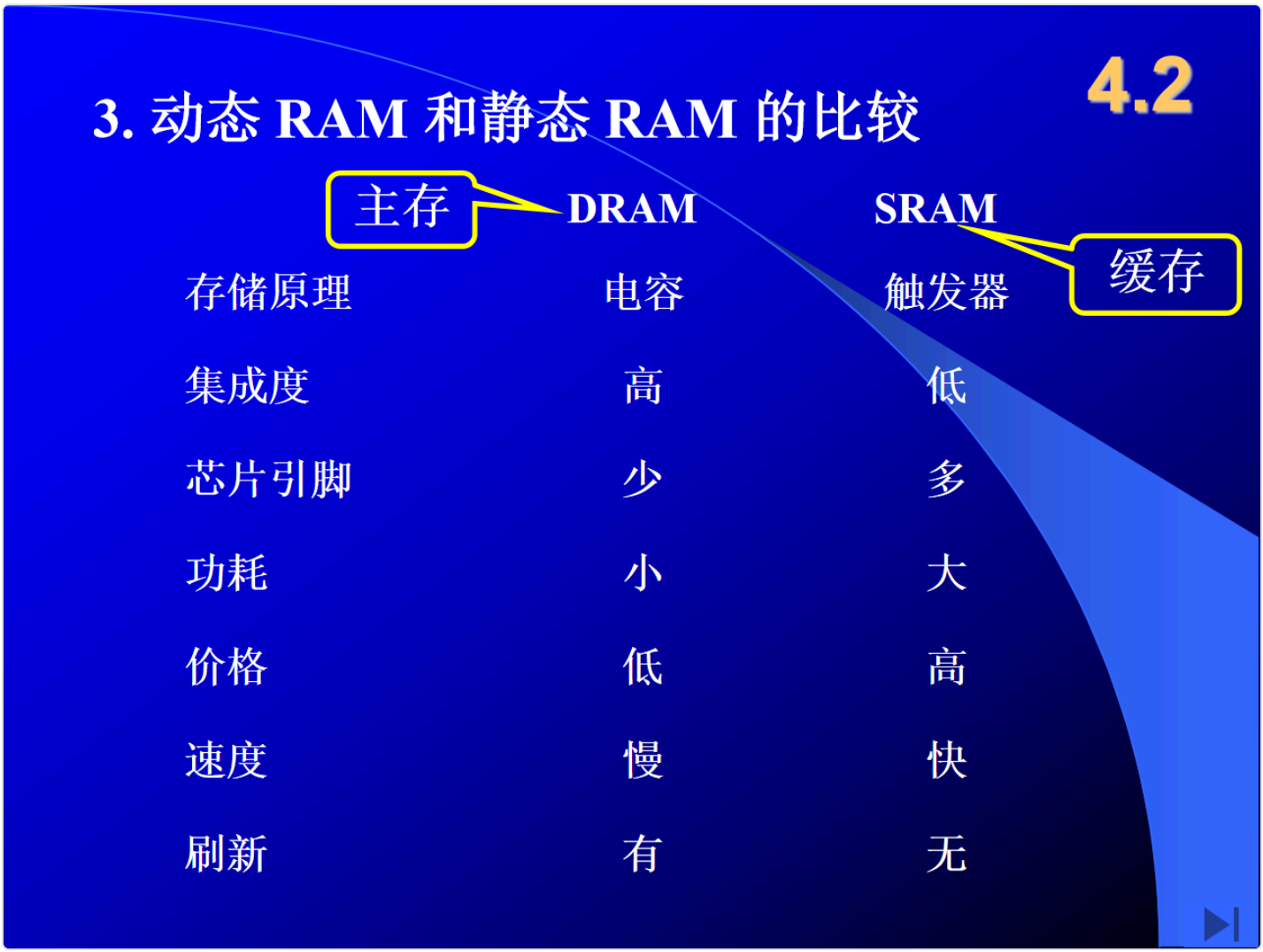
刷新与行地址有关——按行刷新，不需要列地址（课件 p.148）。以 128×128 矩阵、存取周期 0.5 μs、刷新间隔 2 ms 为例：



方式	做法	关键数据（课件 p.148–150）	优缺点
集中刷新	2 ms 内前段读写、末尾集中 128 个周期逐行刷新	2 ms÷0.5 μs=4000 周期；读写 3872 周期（1936 μs）+刷新 128 周期（64 μs）；**死区=0.5 μs×128=64 μs**；**死时间率=128/4000×100%=3.2%**	读写期间不受干扰，但存在"死区"
分散刷新	每个存取周期=读写 0.5 μs+刷新 0.5 μs，即 tC=tM+tR=1 μs，每周期刷一行	128 个周期（128 μs）刷完全部行；**无死区**	存取周期加长一倍，速度下降
异步刷新（分散+集中结合）	每隔 2 ms÷128≈**15.6 μs** 刷新一行，每行每隔 2 ms 刷新一次	**死区仅 0.5 μs**；可把刷新安排在指令译码阶段，不出现死区	兼顾两者，实际常用

易错警示：① 刷新间隔 2 ms 指**每一行**都要在 2 ms 内被刷到，不是整片只刷一次；② 死时间率的分母是 2 ms 内的**总周期数**（4000），不是读写周期数；③ 分散刷新"无死区"的代价是存取周期变 1 μs。

5.5 DRAM 与 SRAM 比较（课件 p.151）



比较项	DRAM	SRAM
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
用途	主存	缓存

易错警示：两者**都易失**（断电丢信息），"SRAM 不用刷新"不等于"断电不丢"。

6. 只读存储器（ROM）（课件 p.152–155）

类型	编程/擦除方式	存储原理要点（课件页）
MROM 掩模 ROM	厂家掩模，不可改	行列交叉处**有 MOS 管为"1"、无 MOS 管为"0"**（p.152）
PROM	一次性编程	**熔丝未断为"1"、熔丝断为"0"**（p.152）
EPROM	多次编程，**紫外线全部擦洗**	N 型沟道浮动栅 MOS：D 端加正电压形成浮动栅→S、D 不导通为"0"；不加→导通为"1"（p.153）
EEPROM	多次编程，**电可擦写**，可局部/全部擦写	具备 RAM 式的电擦写功能（p.155）
Flash Memory	电擦写，**比 EEPROM 快**	闪速型存储器（p.155）

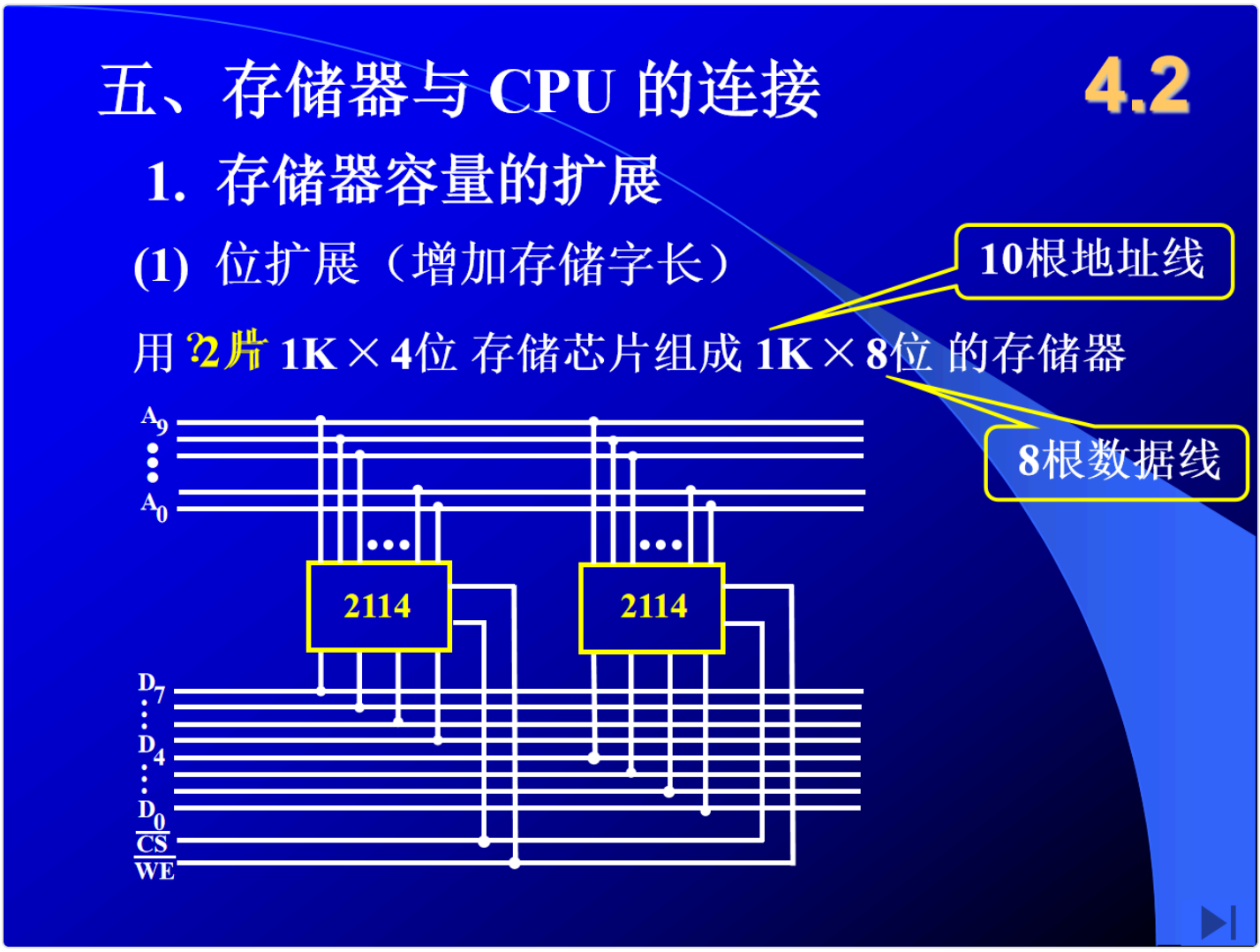
2716 EPROM（课件 p.154）：2K×8 位；地址 A10～A0（11 根），数据 DO7～DO0；$\overline{CS}$ 片选；**PD/Progr 功率下降/编程输入端，读出时为低电平**；VPP 编程电源、VCC、VSS；24 引脚。课件另给出 EPROM 价格便宜、集成度高的特点（p.155）。

易错警示：擦除方式配对不要错——EPROM=紫外线**整片**擦；EEPROM=电擦、可**局部**；PROM 只能写**一次**；Flash 比 EEPROM 快。

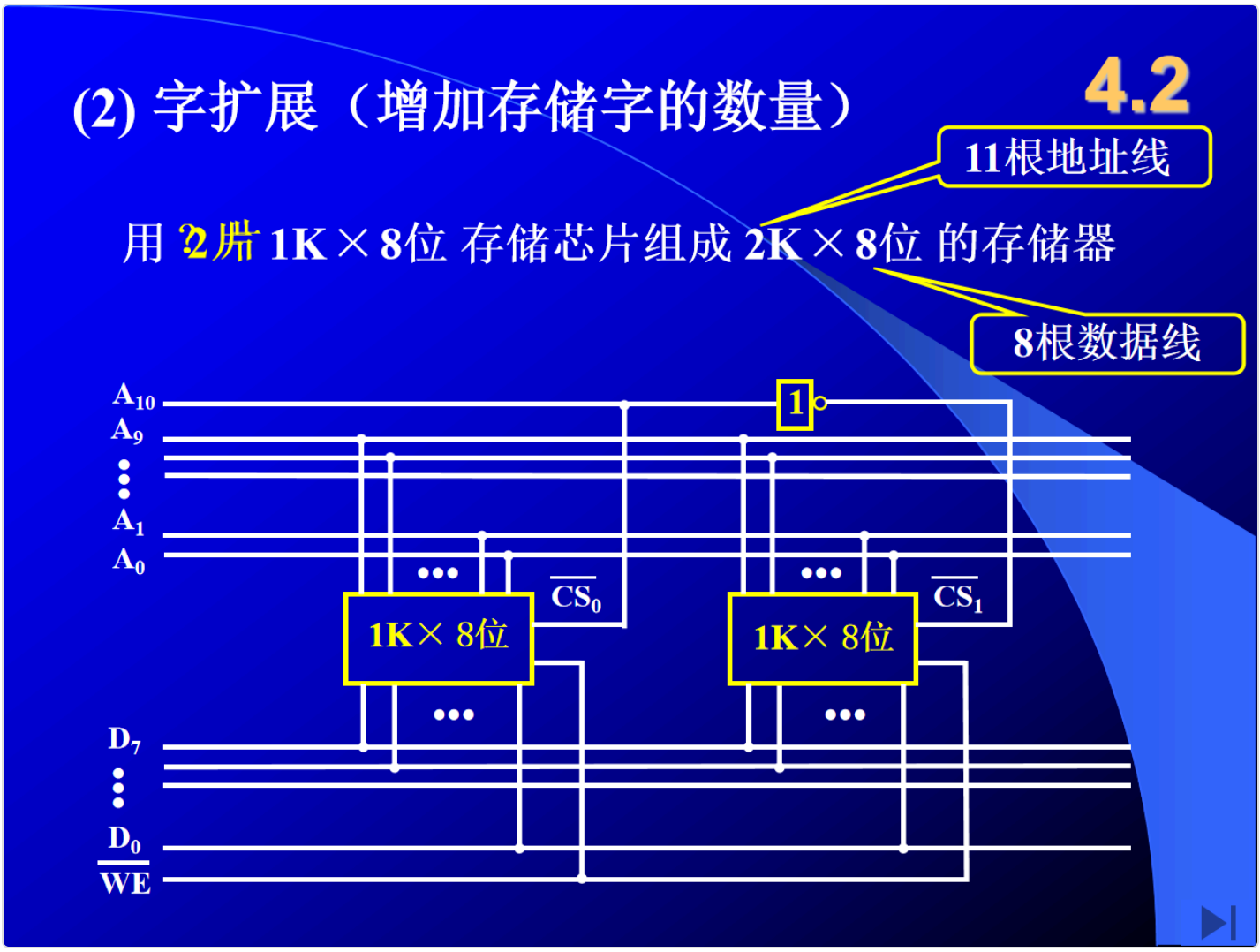
7. 存储器容量扩展与 CPU 连接（课件 p.156–164，核心设计考点）

7.1 三种扩展方式

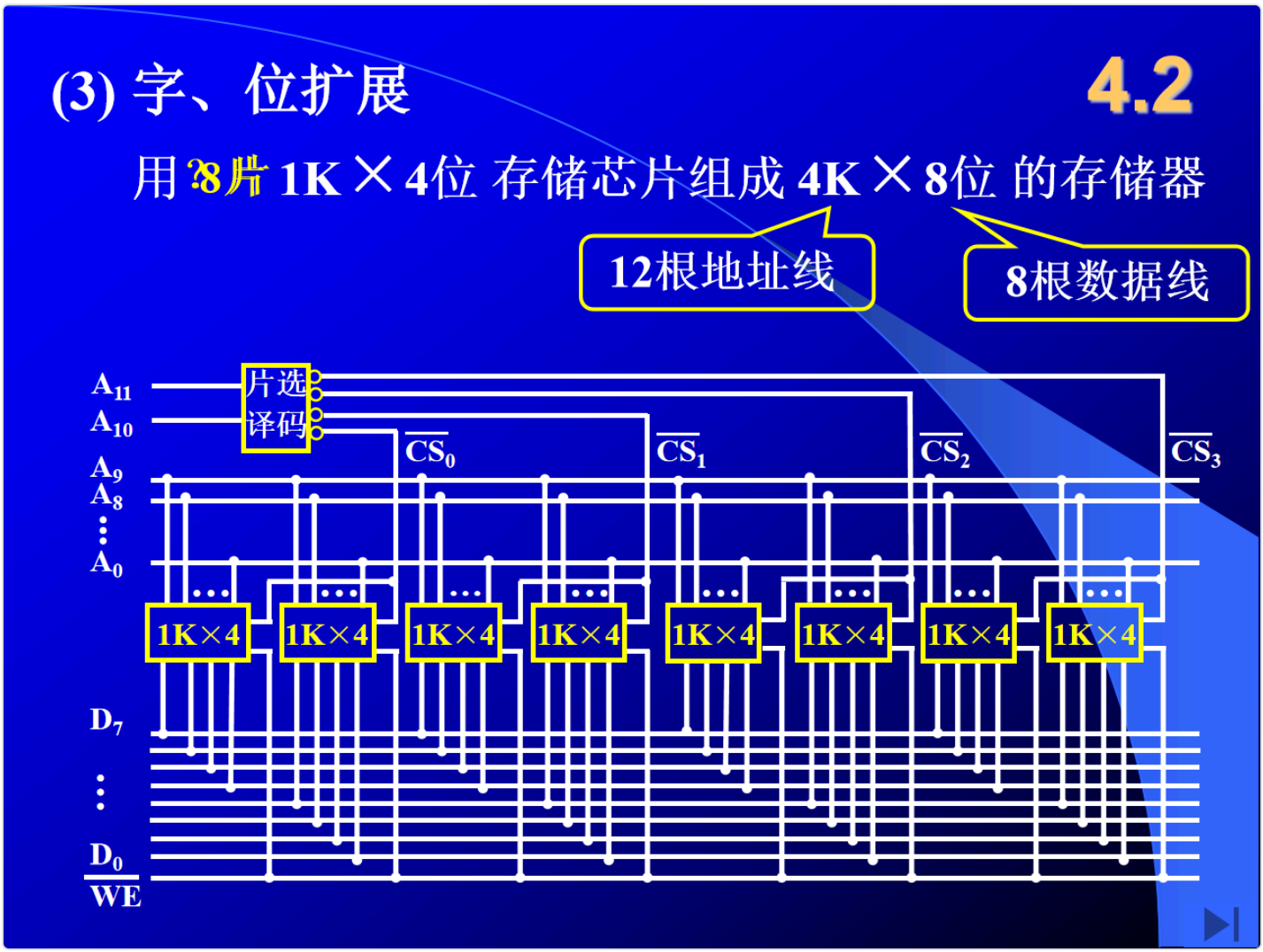
位扩展（增加存储字长）：用 2 片 1K×4 位（2114）组成 1K×8 位。地址线 10 根（A9～A0）**并联**到两片，$\overline{CS}$、$\overline{WE}$ 并联；数据线分开：一片接 D7～D4、另一片接 D3～D0（课件 p.156）。



字扩展（增加存储字的数量）：用 2 片 1K×8 位组成 2K×8 位。数据线 8 根、$\overline{WE}$ 并联；片内地址 A9～A0 并联；**最高位 A10 作片选**：$\overline{CS}0$ = A10（A10=0 选第 0 片），$\overline{CS}1$ = A10 经反相器（A10=1 选第 1 片）（课件 p.157）。



字、位扩展：用 8 片 1K×4 位组成 4K×8 位 = 4 组 × 2 片。12 根地址线：A9～A0 并联到所有芯片；**A11、A10 经片选译码（2:4）产生 $\overline{CS}0$～$\overline{CS}3$** 选组；每组内 2 片位扩展分接高/低 4 位数据；8 根数据线（课件 p.158）。



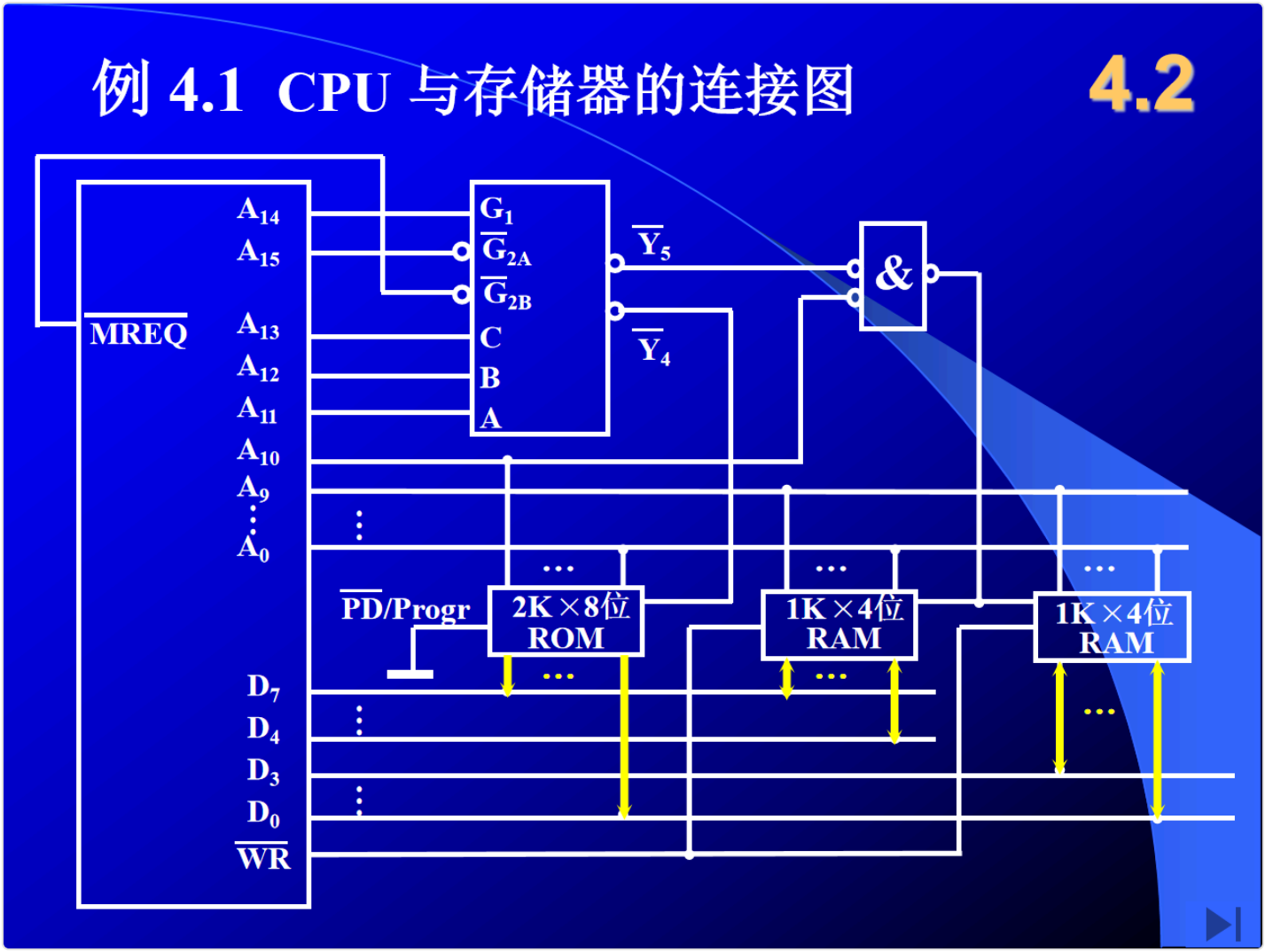
chip 数通用公式：总片数 = (目标字数÷芯片字数) × (目标位数÷芯片位数)。

7.2 与 CPU 连接的六件事（课件 p.159）

(1) 地址线的连接（低位并联进芯片、高位译码作片选）；(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（系统程序区用 ROM、用户程序区用 RAM）；(6) 其他：时序、负载。

7.3 例4.1 连接设计（课件 p.160–162）

地址分配：系统程序区 **6000H～67FFH**（2K，ROM），用户程序区 **6800H～6BFFH**（1K，RAM）。



- 芯片：ROM 2K×8 位 **1 片**；RAM 1K×4 位 **2 片**（位扩展成 1K×8）。
- 地址线：A10～A0 接 ROM；A9～A0 接 RAM。
- 片选（138 译码器）：**C、B、A ← A13、A12、A11；G1 ← A14；G2A̅ ← A15；G2B̅ ← MREQ̅**。
- ROM：A15A14A13A12A11 = 0 1 1 0 0 → CBA=100 → **Y4̅ 接 ROM 片选**；
- RAM：A15A14A13A12A11 = 0 1 1 0 1 → CBA=101 → Y5̅ 有效；但 Y5̅ 覆盖 6800H～6FFFH（2K），需再限定 **A10=0** 才是 1K，故图中 Y5̅ 与 A10 经门电路（低电平有效的或逻辑）接两片 RAM 的片选。
- 数据线：ROM 接 D7～D0；两片 RAM 分接 D7～D4 与 D3～D0；WR̅ 接两片 RAM 的 WE̅；ROM 的 PD/Progr 接地（读出方式）。

7.4 例4.2、例4.3（课件 p.163–164）

- 例4.2：最小 4K 为系统程序区（0000H～0FFFH，1 片 4K×8 ROM），相邻 8K 为用户程序区（1000H～2FFFH，2 片 4K×8 RAM）；A11～A0 接各芯片地址线；片选由 A14A13A12 译码（Y0̅/Y1̅/Y2̅）。完整解答见习题册题3。
- 例4.3：CPU 20 根地址线、8 根数据线，IO/M̅ 作访存控制，RD̅/WR̅ 读/写命令；2764 EPROM（8K×8）；要求地址 **F0000H～FFFFFH**（64K）→ **8 片** 2764，用 138 译码。完整解答见习题册题4。

易错警示：① 末地址 = 首地址 + 容量 − 1（1K=400H、2K=800H、4K=1000H、8K=2000H）；② 片选译码用**高位**地址，片内地址用**低位**，接反则地址重叠或空洞；③ 138 的使能端要把"地址范围固定位"和访存控制信号（MRE$\overline{Q}$/IO/$\overline{M}$）都用上，否则地址不唯一。

8. 存储器的校验与汉明码（课件 p.165–172）

8.1 编码的最小距离（课件 p.165）

最小距离 L：任意两组合法代码之间二进制位数的最少差异。纠错/检错能力满足：

L − 1 = D + C（D ≥ C），D 为检测错误位数、C 为纠正错误位数。

汉明码 L=3，具有一位纠错能力（此时 D=1、C=1；若只检错可检 2 位）。

8.2 汉明码组成三要素（课件 p.166–167）

1. **检测位个数 k**：满足 $2^k \geq n + k + 1$（n 为信息位数）。
2. **检测位位置**：第 2^i 位（i=0,1,2,…），即第 1、2、4、8、… 位放 C1、C2、C4、C8。
3. **检测位取值**：由该位所在检测小组的**奇偶校验**任务决定（偶校验使组内 1 的个数为偶数，奇校验为奇数）。

检测小组（课件 p.167）： - C1 的 g1 组：第 1,3,5,7,9,11,… 位 - C2 的 g2 组：第 2,3,6,7,10,11,… 位 - C4 的 g3 组：第 4,5,6,7,12,13,… 位 - C8 的 g4 组：第 8,9,10,11,12,13,14,15,24,… 位

规律：gi 独占第 2^{i-1} 位；gi、gj 共同占第 $2^{i-1}+2^{j-1}$ 位；三个组共同占相应三位之和（课件 p.167）。即"位置编号的二进制表示中含哪一位 2 的幂，就归哪个组管"。

8.3 纠错过程（课件 p.170–172）

接收端按同样小组做偶校验，形成新检测位 **P1、P2、P4**（k=3 时）： - P1 = 第1⊕3⊕5⊕7 位；P2 = 第2⊕3⊕6⊕7 位；P4 = 第4⊕5⊕6⊕7 位。 - 无错时 P4P2P1=000；**P4P2P1 的二进制值 = 出错位的位置**。 - 若出错位是**检测位本身**（如第 4 位），不影响信息位，**可不纠**（课件 p.172 练习2）。

例：0101 偶校验汉明码 = **0100101**（课件 p.168 例4.4）；接收 0100111 → P4P2P1=110 → 第 6 位错 → 纠正为 0100101 → 信息 0101（课件 p.171 例4.5）。完整步骤见习题册。

易错警示：① $2^k \geq n+k+1$ 是"≥"，取满足的最小 k；② 检测位插在第 1、2、4、8 位，信息位顺次填其余位置，不要从第 1 位连排信息；③ 偶校验与奇校验配置结果不同（0011 偶=1000011、奇=0101011，课件 p.169、p.172）。

9. 提高访存速度的措施（课件 p.173–181）

9.1 单体多字系统（课件 p.173）

一个存储体一次读出**多个字**（如 4 个 W 位），送单字长寄存器逐个使用；配合高速器件、Cache–主存层次结构，**增加存储器带宽**。

9.2 多体并行系统（课件 p.174–179）

方式	编址	特点
高位交叉	地址**高位=体号**，体内顺序编址	各体可并行工作（不同请求源访不同体）；但连续地址集中在同一体，**不能流水线式连续访问**，利于扩容
低位交叉	地址**低位=体号**，各体轮流编址	连续地址分布在不同体，可**流水线并行存取**，在不改变存取周期的前提下增加带宽

低位交叉流水线计算（课件 p.179）：四体低位交叉，存取周期 T、总线传输周期 τ，实现流水线需 **T = 4τ**；连续读取 4 个字所需时间 = **T + (4−1)τ = T + 3τ**。（推广：m 体时 T=mτ，连续 m 个字时间 = T+(m−1)τ，补充常识。）

9.3 存储器控制部件（存控，课件 p.180）

含控制线路、排队器、节拍发生器、存控标记触发器。优先级原则：**易发生代码丢失的请求源优先级最高；严重影响 CPU 工作的请求源给予次高优先级**。

9.4 高性能存储芯片（课件 p.181）

- **SDRAM**（同步 DRAM）：在系统时钟控制下读写，CPU 无须等待；
- **RDRAM**：Rambus 开发，主要解决**存储器带宽**问题；
- **带 Cache 的 DRAM**：芯片内集成 SRAM 组成的 Cache，有利于**猝发式读取**。

10. 高速缓冲存储器 Cache（课件 p.182–194）

10.1 概述与工作原理

- 问题提出：CPU 与主存（DRAM）速度差异，避免 CPU"空等"；依据**程序访问的局部性原理**（课件 p.182；时间局部性/空间局部性的展开为补充常识）。
- 编址：主存与缓存**按块存储、块大小相同**（块长 B）。主存共 $M=2^m$ 块，主存地址 = 主存块号 m 位 + 块内地址 b 位；Cache 共 $C=2^c$ 块，Cache 地址 = 缓存块号 c 位 + 块内地址 b 位；**M ≫ C**；用**标记**记录与某缓存块建立对应关系的主存块号（课件 p.183–184）。
- **命中**：欲访问信息已在 Cache（主存块与缓存块建立了对应关系）；**未命中**：需访问主存并把该块调入 Cache（课件 p.184）。
- **命中率 h**：CPU 欲访问的信息在 Cache 中的比率（h = 命中次数÷总访问次数，补充常识）。命中率与 Cache **容量**和**块长**有关；一般每块取 **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 16 个存储字；IBM 370/168 4 体交叉→块长 4 个存储字（64 位×4=256 位）（课件 p.185）。

- **平均访问时间与效率**（课件 p.186）：设命中率 h、访问 Cache 时间 tc、访问主存时间 tm，则
- 平均访问时间 ta = h×tc + (1−h)×tm
- 效率 **e = tc ÷ [h×tc + (1−h)×tm] × 100%**（访问 Cache 的时间÷平均访问时间）。

10.2 基本结构与读写操作

结构三部分：**Cache 存储体、替换机构、主存–Cache 地址映射变换机构**（课件 p.187）。

> 课件疑点：p.187 图中把地址映射变换标注为"由CPU完成"，课件原文如此，存疑：通用结论是 Cache 的地址映射变换由**硬件**完成、对程序员透明（"由 CPU 完成"与 Cache 硬件管理的基本概念不符），答题以"硬件实现"为准。

读操作流程（课件 p.188）：CPU 发地址→命中？是→访问 Cache 取信息送 CPU；否→访问主存取信息送 CPU，并将新主存块调入 Cache→Cache 满？是→执行替换算法腾出空位；否→直接装入。

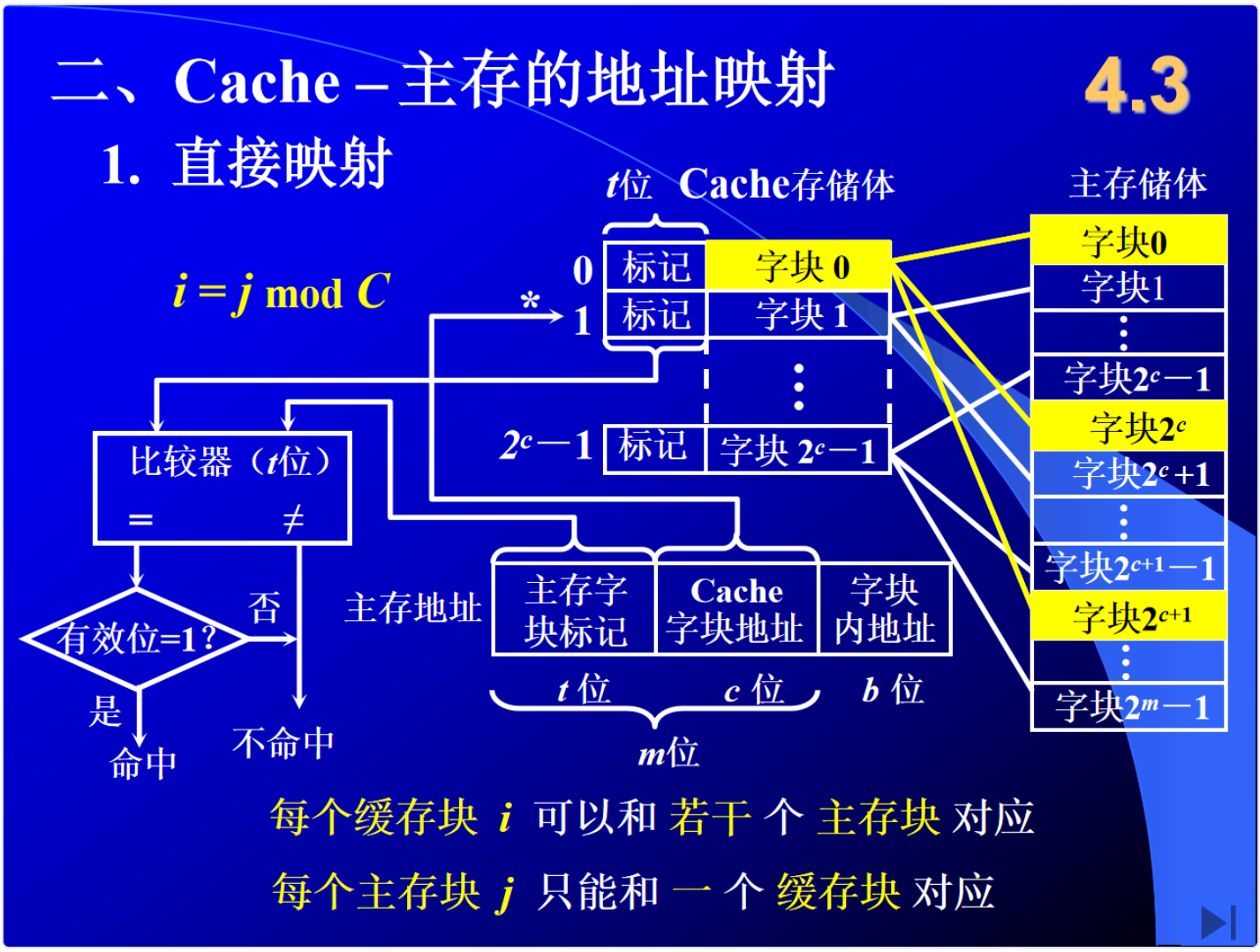
写操作（课件 p.189）：

策略	做法	写操作时间	特点
写直达 Write-through	数据既写 Cache 又写主存	访问主存的时间	一致性好、更新策略易实现
写回 Write-back	只写 Cache，被替换出去时才写回主存	访问 Cache 的时间	读失效替换时需写回，增加复杂性

Cache 改进（课件 p.190）：① 增加级数（片载/片内 Cache、片外 Cache）；② 统一缓存与分立缓存（指令 Cache、数据 Cache；与主存结构、指令执行控制方式/是否流水有关）。例：Pentium 8K 指令 Cache + 8K 数据 Cache；PowerPC620 32K + 32K。

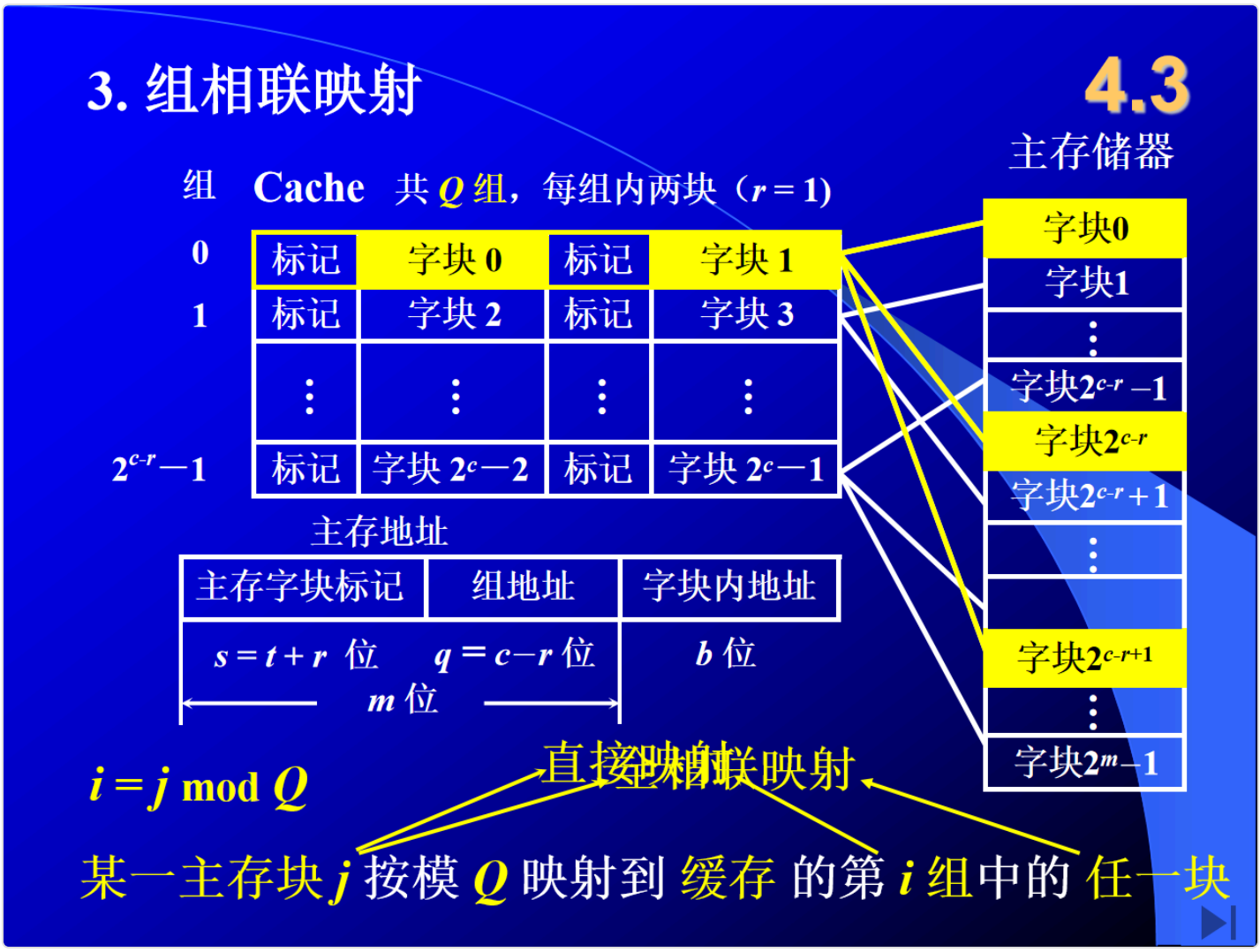
10.3 三种地址映射（核心考点）

直接映射（课件 p.191）：**i = j mod C**（j 为主存块号、i 为缓存块号、C 为缓存块数）。每个主存块只能对应**一个**缓存块；每个缓存块可对应**若干个**主存块。主存地址 = 主存字块标记 t 位 + Cache 字块地址 c 位 + 字块内地址 b 位（t = m − c）。命中判断：t 位比较器相等**且有效位=1**。



全相联映射（课件 p.192）：主存**任一块**可映射到缓存**任一块**。主存地址 = 主存字块标记（m = t+c 位，即整个主存块号）+ 字块内地址 b 位；需同时比较所有标记，**成本高**。

组相联映射（课件 p.193）：**i = j mod Q**（Q 为组数）。Cache 共 Q=2^(c−r) 组、每组 2^r 块（课件例 r=1，每组两块）；主存块按模 Q 映射到第 i 组中的**任一块**（组间直接、组内全相联）。主存地址 = 主存字块标记 s = t+r 位 + 组地址 q = c−r 位 + 字块内地址 b 位。



小结对比（课件 p.194）：

映射	对应关系	评价
直接	某一主存块只能固定映射到某一缓存块	不灵活（冲突多），硬件简单
全相联	某一主存块能映射到任一缓存块	灵活，成本高
组相联	某一主存块只能映射到某一缓存组中的任一块	折中方案

10.4 替换算法（课件 p.194）

- **FIFO 先进先出**：替换最早调入的块（补充常识：实现简单，未考虑使用频率）。
- **LRU 近期最少使用**：替换最近最久未被访问的块（补充常识：依据局部性，命中率通常更高）。
- 直接映射无需替换算法（位置固定）；替换发生在**全相联/组相联**且候选块已满时（补充常识）。

易错警示：① mod 的对象：直接映射 mod **C（块数）**、组相联 mod **Q（组数）**，不要 mod 每组块数；② 组相联标记位是 t+r 位而不是 t 位；③ 效率 e 的分子是 tc，不要写反；④ 块长并非越大越好（课件只给"与容量、块长有关"及典型取值，判断题按课件口径）。

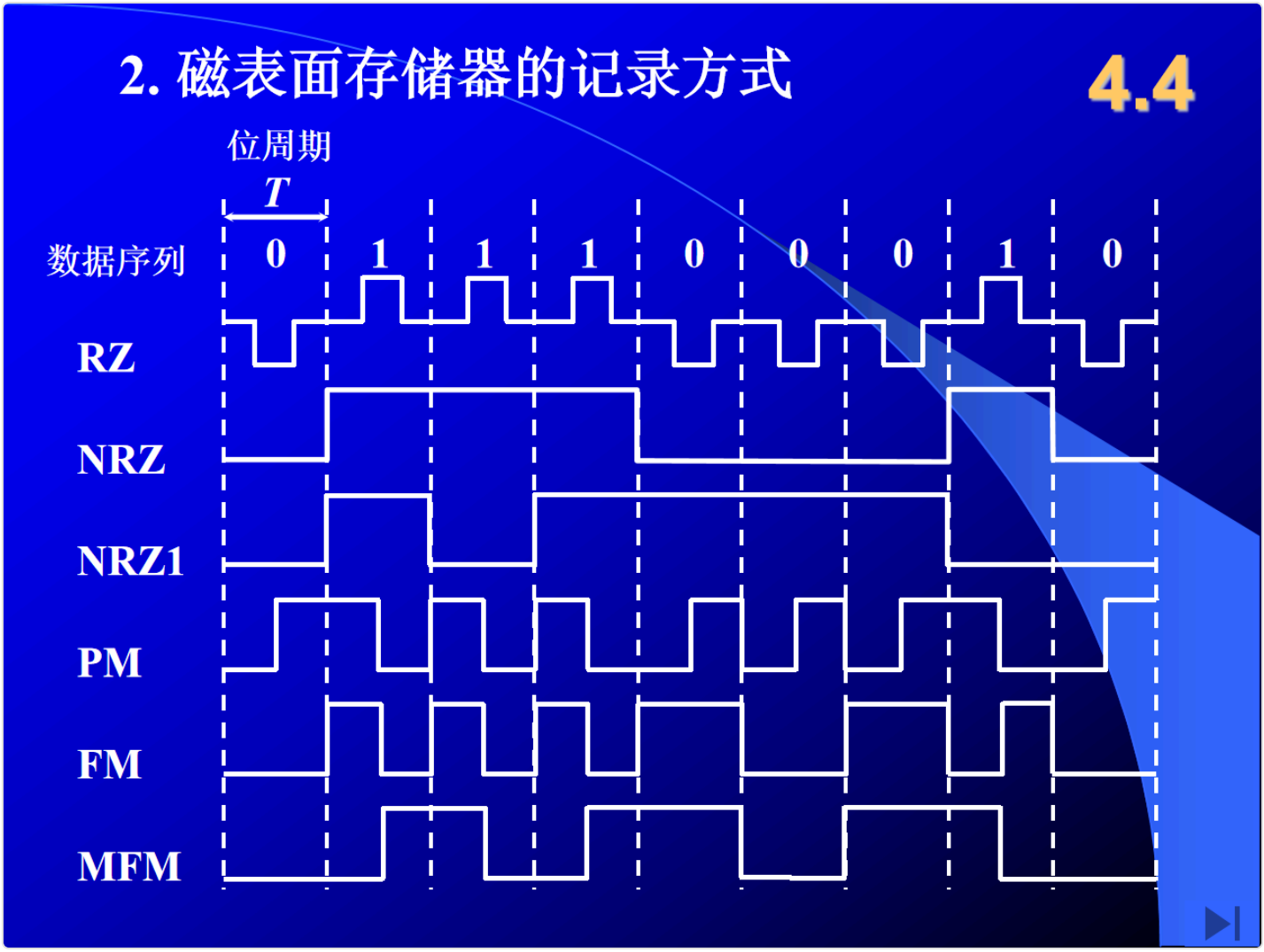
11. 辅助存储器（课件 p.195–205）

11.1 概述与技术指标（课件 p.195）

特点：**不直接与 CPU 交换信息**（经主存交换，补充常识）。磁表面存储器技术指标： - 记录密度：**道密度 Dt**（沿半径方向单位长度磁道数）、**位密度 Db**（磁道单位长度记录位数）（单位定义为补充常识）； - 存储容量：**C = n × k × s**（n 记录面数、k 每面磁道数、s 每道位数；符号定义为补充常识，课件只给公式）； - 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； - 数据传输率：**Dr = Db × V**（V 为磁表面相对磁头的线速度）； - 误码率：出错信息位数与读出信息总位数之比。

11.2 磁记录原理与记录方式（课件 p.196–199）

- **写**：写线圈电流→铁芯磁通→磁层局部磁化单元，磁化方向代表 0/1（课件 p.196）。
- **读**：磁层运动→磁通变化→读线圈感应电势（课件 p.197）。
- 六种记录方式（课件 p.198，数据序列 011100010 的波形）：



方式	规则
RZ 归零制	位周期前半脉冲：1 正脉冲、0 负脉冲，后半归零
NRZ 不归零制	电平表示 0/1（1 高、0 低），位周期内不归零
NRZ1 见1就翻	写 1 时电流在位周期起点**翻转**，写 0 保持
PM 调相制（曼彻斯特）	每位周期中心必跳变：1 高→低、0 低→高
FM 调频制	每位周期起点跳变（同步）；1 在中心再跳变一次，0 不跳
MFM 改进调频制	1 在中心跳变；0 中心不跳，仅当**前一位也是 0** 时在起点跳变

例：NRZ1 的读出代码波形（课件 p.199，数据序列 0110010）：驱动电流在每个"1"的起点翻转→磁通变化→感应电势在翻转处出脉冲；用**同步脉冲（位周期前半）窗口内有无感应电势脉冲**判读：有脉冲=1、无脉冲=0，读出代码 = 0110010，与写入一致。完整解答见习题册题11。

11.3 硬磁盘存储器（课件 p.200–202）

- 类型：固定磁头/移动磁头；可换盘/固定盘（课件 p.200）。
- 结构层次：**主机 — 磁盘控制器 — 磁盘驱动器 — 盘片**（课件 p.200）。
- 磁盘驱动器：**闭环自动控制系统**——控制器送来目标磁道信号→定位驱动→音圈电机带动读写臂；位置检测、测速输出反馈（课件 p.201）。
- 磁盘控制器：**主机与磁盘驱动器之间的接口**；接收主机命令转换成驱动器控制命令、实现主机与驱动器间数据格式转换、控制驱动器读写（课件 p.202）。
- 盘片：**硬质铝合金**材料制成（课件 p.202）。

11.4 软磁盘存储器（课件 p.203–204）

比较项	硬盘	软盘
速度	高	低
磁头	固定、活动	活动
磁头与盘片	浮动（不接触）	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	（课件未列，相对宽松，补充常识）

软盘片由**聚酯薄膜**制成；结构：保护套、主轴孔、**写保护口**、读/写磁头访问槽、衬里/清洁材料（课件 p.204）。

11.5 光盘存储器（课件 p.205）

采用**光存储技术**，利用**激光写入和读出**。 - 第一代光存储技术：**非磁性介质、不可擦写**（只读型和只写一次型），靠**热作用**（物理或化学变化）； - 第二代光存储技术：**磁性介质、可擦写**，靠**热磁效应**。

> 范围说明：本章课件未涉及 RAID 内容，复习以课件为准，不做扩展。

12. 易错点与高频陷阱汇总

1. 1K=1024=400H；**末地址 = 首地址 + 容量 − 1**，如 6000H 起 2K → 67FFH，漏减 1 必错。
2. 存取周期 ≥ 存取时间（周期含恢复时间），带宽按存取周期算而不是存取时间。
3. 芯片地址线根数 = $\log_2$**字数**：16K×1 是 14 根地址线、1 根数据线，与总位数无关。
4. SRAM 靠触发器、DRAM 靠电容；DRAM 要刷新、SRAM 不要，但**两者都易失**。
5. DRAM **按行刷新**；2 ms 是每行的刷新间隔上限，不是整片刷新一次的时间。
6. 集中刷新死区 = 行数×存取周期；死时间率分母是 2 ms 内总周期数；分散刷新无死区但周期加倍。
7. 单管 DRAM 读出是破坏性的（需再生放大）；三管 DRAM **读出与原存信息相反**。
8. ROM 擦写配对：PROM 一次、EPROM 紫外线整片、EEPROM 电擦可局部、Flash 比 EEPROM 快。
9. 位扩展：地址/片选并联、数据分高低位；字扩展：数据并联、**高位地址译码作片选**——两者别接反。
10. 138 使能端必须接入地址固定位与访存控制信号（MRE$\bar{Q}$/IO/$\bar{M}$），否则片选地址不唯一。
11. 汉明码检测位在第 1、2、4、8 位（2^i），不是顺序排；$2^k \geq n+k+1$ 取最小满足的 k。
12. 纠错指误码 P4P2P1 = 出错位置的二进制；错在检测位本身可不纠；偶/奇校验配置结果不同。
13. 缓存–主存层次解决**速度**、主存–辅存（虚拟存储）解决**容量**，目的别记反。
14. 直接映射 i = j mod **C**、组相联 i = j mod **Q**；组相联标记 s = t+r 位、组地址 q = c−r 位。
15. 平均访问时间 ta = h·tc+(1−h)·tm；效率 e = tc/ta，分子分母别颠倒。
16. 低位交叉才能流水线提带宽；高位交叉顺序编址、连续访问落在同一体，不能流水。
17. 写直达写时间=主存时间、写回写时间=Cache 时间（替换时才写回主存）。
18. 磁盘是**直接存取**（串行访问大类）、磁带是顺序存取，都不是随机访问。
19. NRZ1 是"见 1 翻转"（位周期起点），读出靠同步窗口内有无感应电势脉冲判 1/0。
20. 课件 p.187"由CPU完成"存疑：Cache 地址映射变换应为硬件完成、对程序员透明。

13. 典型例题精讲

例1（连接设计，课件 p.160–162 例4.1）

题：CPU 16 根地址线、8 根数据线，MRE$\bar{Q}$ 访存请求、W$\bar{R}$ 读写命令；系统程序区 6000H～67FFH、用户程序区 6800H～6BFFH；用 2K×8 ROM、1K×4 RAM 及 138 译码器设计连接。（题干由课件解答图整理）

解：① 容量核算：67FFH−6000H+1=800H=2K → ROM 2K×8 一片；6BFFH−6800H+1=400H=1K → RAM 1K×8=1K×4 两片（位扩展）。② 二进制地址：6000H=0110 0000 0000 0000，67FFH=0110 0111 1111 1111；6800H=0110 1000 0000 0000，6BFFH=0110 1011 1111 1111。可见 A15A14=01 固定，ROM 的 A13A12A11=100、RAM 的=101。③ 地址线：ROM 接 A10～A0，RAM 接 A9～A0。④ 片选：138 的

CBA=A13A12A11，G1=A14、G2$\overline{A}$=A15、G2$\overline{B}$=MRE$\overline{Q}$；Y4→ROM C$\overline{S}$；Y5与 A10 经门电路→两片 RAM C$\overline{S}$（Y5本身覆盖 6800H～6FFFH，须 A10=0 限定为 1K）。 ⑤ 数据与命令：ROM 接 D7～D0、PD/Progr 接地；RAM 分接 D7～D4、D3～D0，W$\overline{R}$→W$\overline{E}$。 **教学点**：先算容量定芯片，再写二进制定"固定位/译码位/片内位"，最后把固定位接使能、译码位接 CBA、片内位接芯片地址脚——三步法适用于所有连接设计题。

例2（汉明码组成+纠错，课件 p.168、p.171）

题：求 0101 的偶校验汉明码；若接收码为 0100111，求传送信息。 **解**：n=4，$2^k \geq 4+k+1$ → k=3（8≥8）。排位：1=C1、2=C2、3=0、4=C4、5=1、6=0、7=1。 C1=位3⊕5⊕7=0⊕1⊕1=0；C2=位3⊕6⊕7=0⊕0⊕1=1；C4=位5⊕6⊕7=1⊕0⊕1=0 → 汉明码 **0100101**。 接收 0100111：P1=0⊕0⊕1⊕1=0，P2=1⊕0⊕1⊕1=1，P4=0⊕1⊕1⊕1=1 → P4P2P1=110=6 → 第 6 位 1→0 → 0100101 → 取 3、5、6、7 位 = **0101**。 **教学点**：组成时"先挖坑（1/2/4 位）再填信息"，纠错时"指误码即位置"，两步都只用小组异或。

例3（刷新三方式计算，课件 p.148–150）

题：128×128 DRAM，存取周期 0.5 μs，刷新间隔 2 ms，求三种刷新的关键参数。 **解**：集中：总周期 2 ms÷0.5 μs=4000；死区=128×0.5 μs=64 μs；死时间率=128/4000=3.2%；读写占 3872 周期=1936 μs。分散：周期=0.5+0.5=1 μs，128 周期=128 μs 刷完一轮，无死区。异步：2 ms÷128≈15.6 μs 刷一行，死区 0.5 μs。 **教学点**：所有刷新计算都从"行数"和"2 ms 内总周期数"两个量出发。

例4（Cache 映射分解+效率，对应习题册题17、题18）

题：主存地址 20 位，Cache 16KB、块长 64B，直接映射；h=0.9、tc=20 ns、tm=200 ns。求地址字段与 ta、e。 **解**：$b=\log_2 64=6$；$C=16KB\div 64B=256=2^8$ → c=8；t=20−8−6=6。地址 12345H=0001 0010 0011 0100 0101 → 标记 000100、行号 10001101（141）、块内 000101（5）。ta=0.9×20+0.1×200=38 ns；e=20/38≈52.6%。 **教学点**：字段划分顺序"标记|行/组|块内"，从低位往高位切 b、c、t；效率公式分子恒为 tc。

附：本章速查表

容量与地址换算：1K=2^{10}=400H；2K=800H；4K=1000H；8K=2000H；16K=4000H；64K=10000H；1M=100000H。末地址=首地址+容量−1。

芯片速查：

芯片	容量	地址线	数据线	关键控制
2114（SRAM）	1K×4	10（A9～A0）	4（I/O1～I/O4）	$\overline{CS}$、$\overline{WE}$
1103（三管 DRAM）	1K×1（32×32）	10	1	读/写选择线、刷新放大器
4116（单管 DRAM）	16K×1（128×128）	7 复用（A′6～A′0）	DIN/DOUT 各 1	$\overline{RAS}$、$\overline{CAS}$、$\overline{WE}$
2716（EPROM）	2K×8	11（A10～A0）	8（DO7～DO0）	$\overline{CS}$、PD/Progr（读=低）、VPP
2764（EPROM）	8K×8	13（A12～A0）	8（D7～D0）	$\overline{CE}$、$\overline{OE}$、$\overline{PGM}$

扩展三方式：位扩展=地址/片选并联、数据分位；字扩展=数据并联、高位译码片选；字位扩展=先位后字、组选译码。片数=(字数比)×(位数比)。

刷新三方式（128 行、0.5 μs、2 ms）：集中→死区 64 μs、死时间率 3.2%；分散→周期 1 μs、无死区；异步→15.6 μs/行、死区 0.5 μs。

汉明码：$2^k \geq n+k+1$；检测位 1/2/4/8；g1={1,3,5,7,9,11…}、g2={2,3,6,7,10,11…}、g3={4,5,6,7,12,13…}、g4={8…15,24…}；P4P2P1=出错位；L−1=D+C（D≥C），汉明 L=3 纠 1 位。

Cache：ta=h·tc+(1−h)·tm；e=tc/ta×100%；直接 i=j mod C（t=m−c 位标记）；全相联标记 m 位；组相联 i=j mod Q，Q=2^(c−r)、每组 2^r 块、标记 t+r 位、组地址 c−r 位；块长一般 4～8 字；写直达/写回；FIFO/LRU。

多体交叉：高位=体号顺序编址（不流水）；低位=体号轮流编址（流水）；m 体 T=mτ，连续 m 字时间=T+(m−1)τ。

时序参数：读 tRC/tA/tCO/tOTD/tOHA；写 tWC/tW/tAW/tWR/tDW/tDH（含义见 §4.3 表）。

磁盘：C=n×k×s；Dr=Db×V；平均寻址=寻道+等待；误码率=出错位数/读出总位数。记录方式：RZ 归零、NRZ 电平、NRZ1 见1翻、PM 中心跳变、FM 起点+1中心、MFM 1中心/0仅前0时起点。

层次速度（课件 p.99）：CPU 10 ns、缓存 20 ns、主存 200 ns、辅存 ms 级。