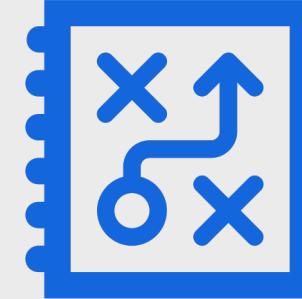


2. 数据的表示与存储

2.8 数据存储与主存储器



2. 数据的表示与存储

2.8 数据存储与主存储器

01 数据存储的概念

02 主存储器

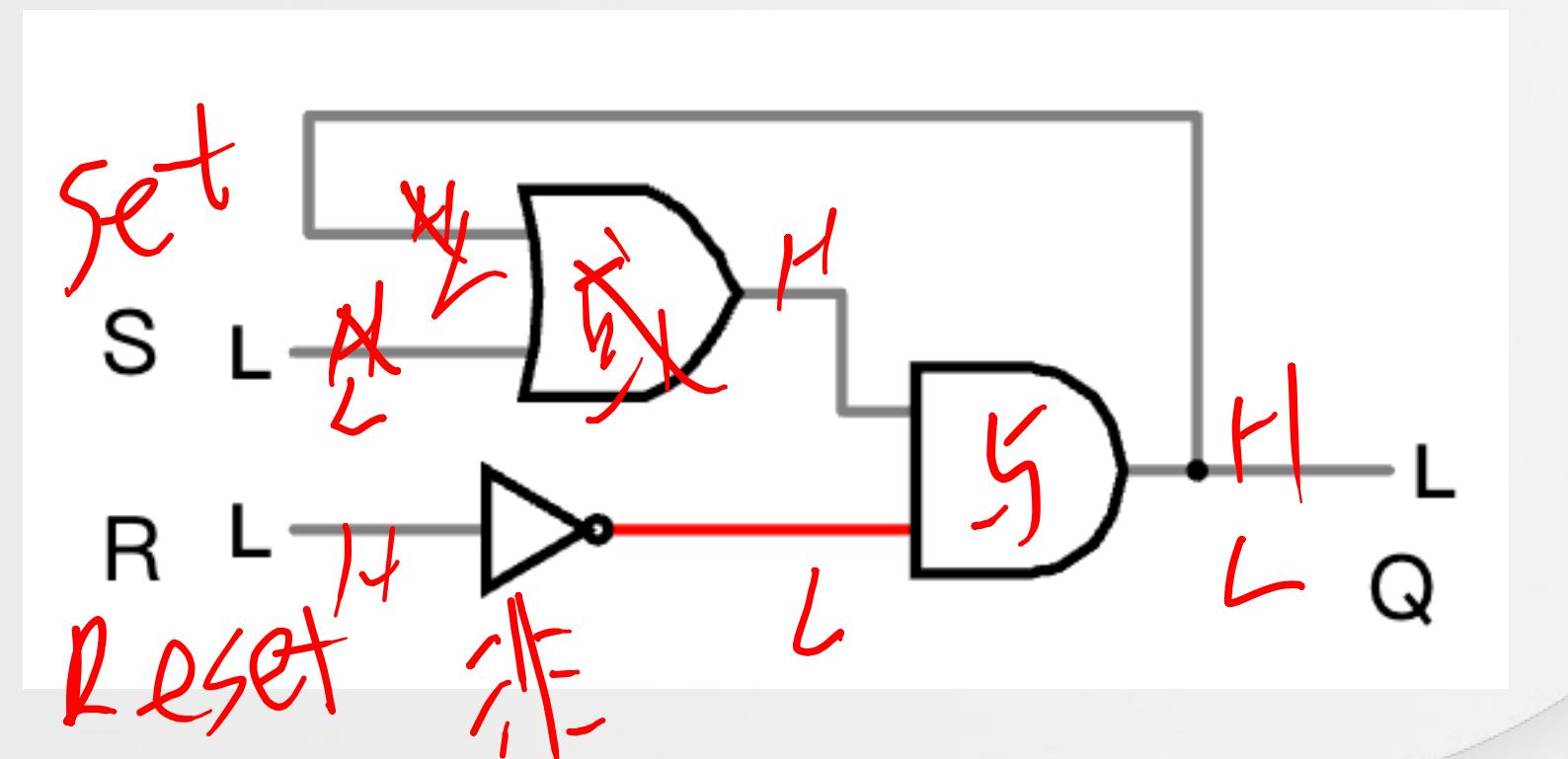
03 主存储器与CPU的连接

04 其他半导体存储器



2. 数据的表示与存储

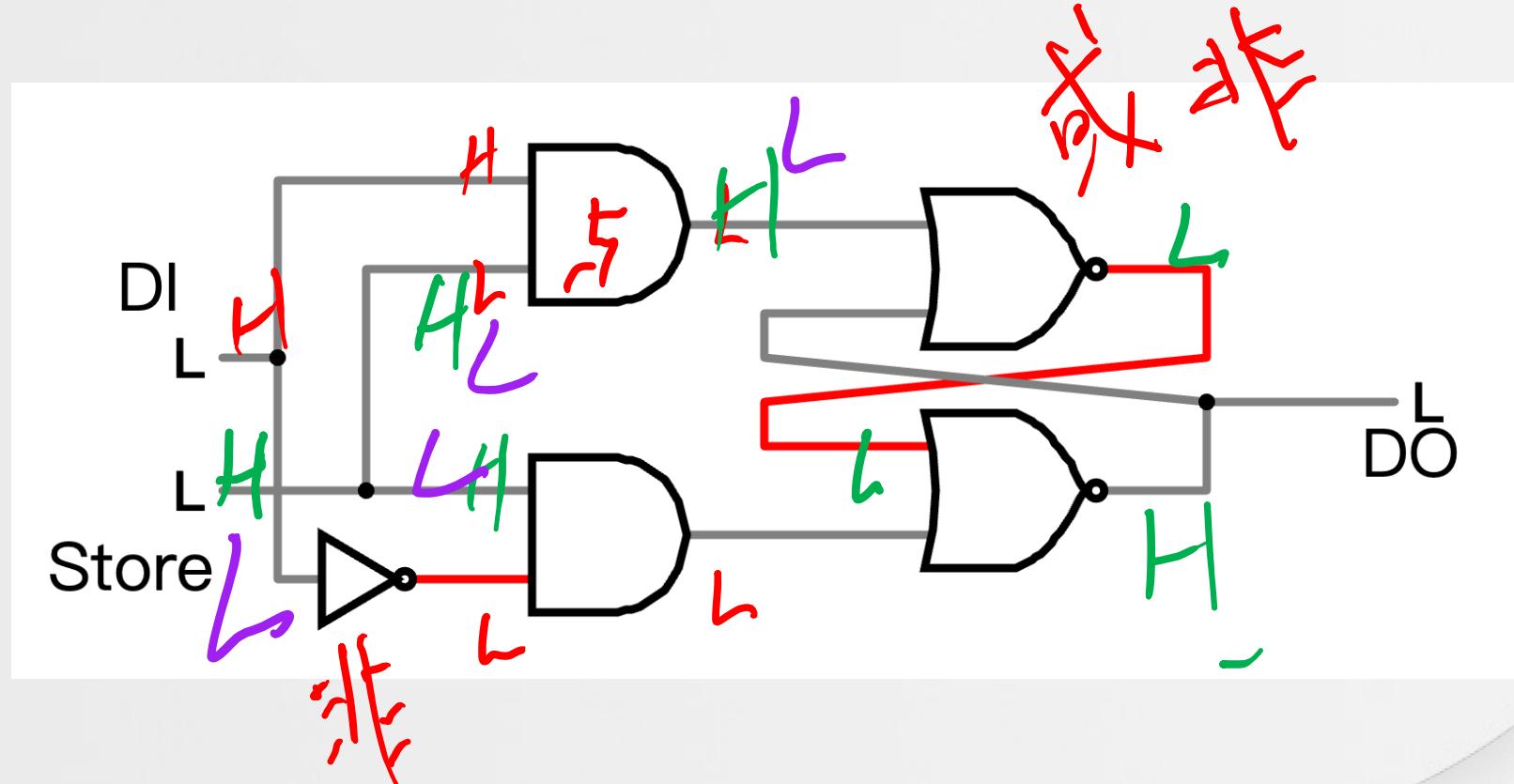
2.8.1 门电路构建数据存储





2. 数据的表示与存储

2.8.1 门电路构建数据存储



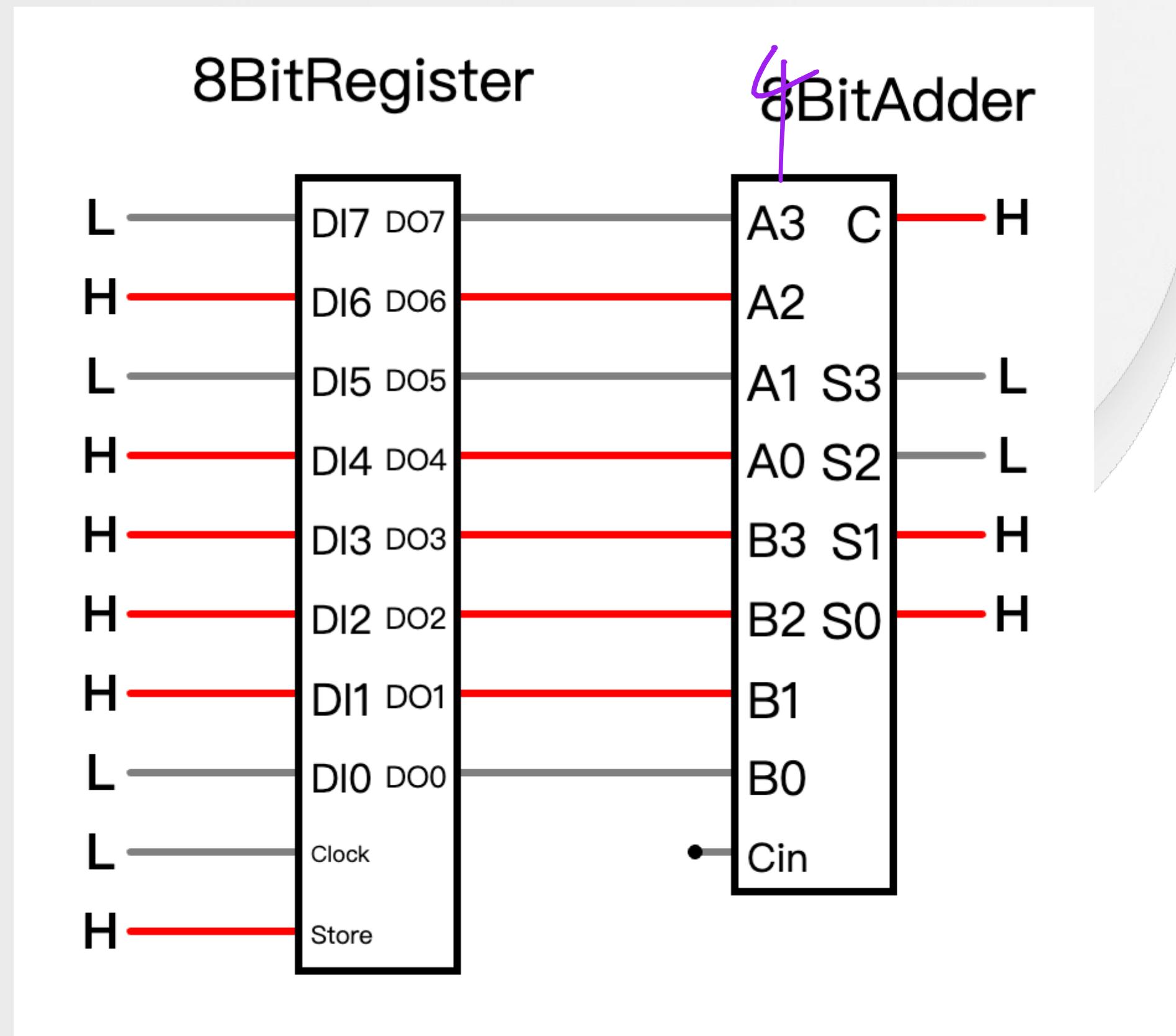
或非

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



2. 数据的表示与存储

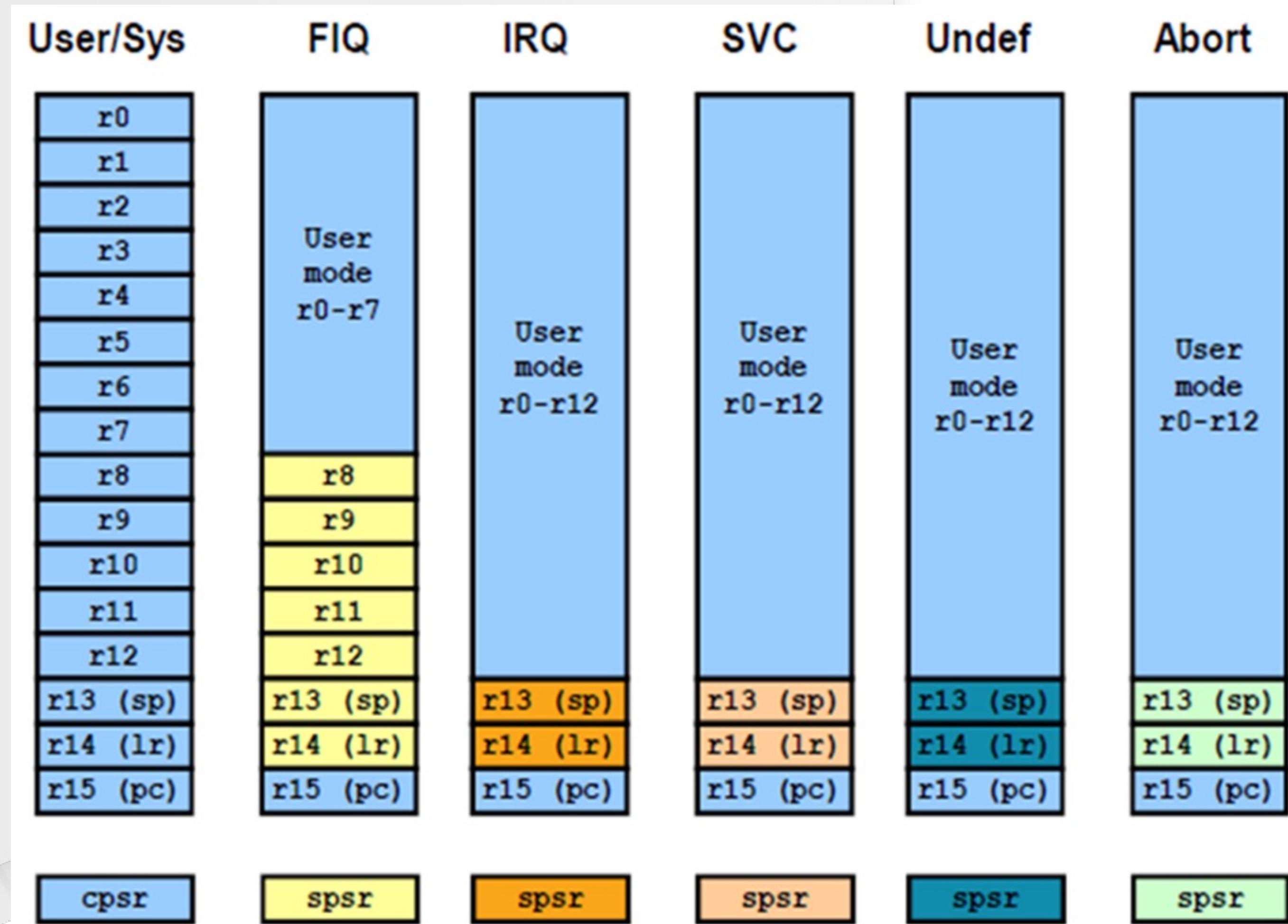
2.8.1 门电路构建数据存储





2. 数据的表示与存储

2.8.1 ARM Core Registers



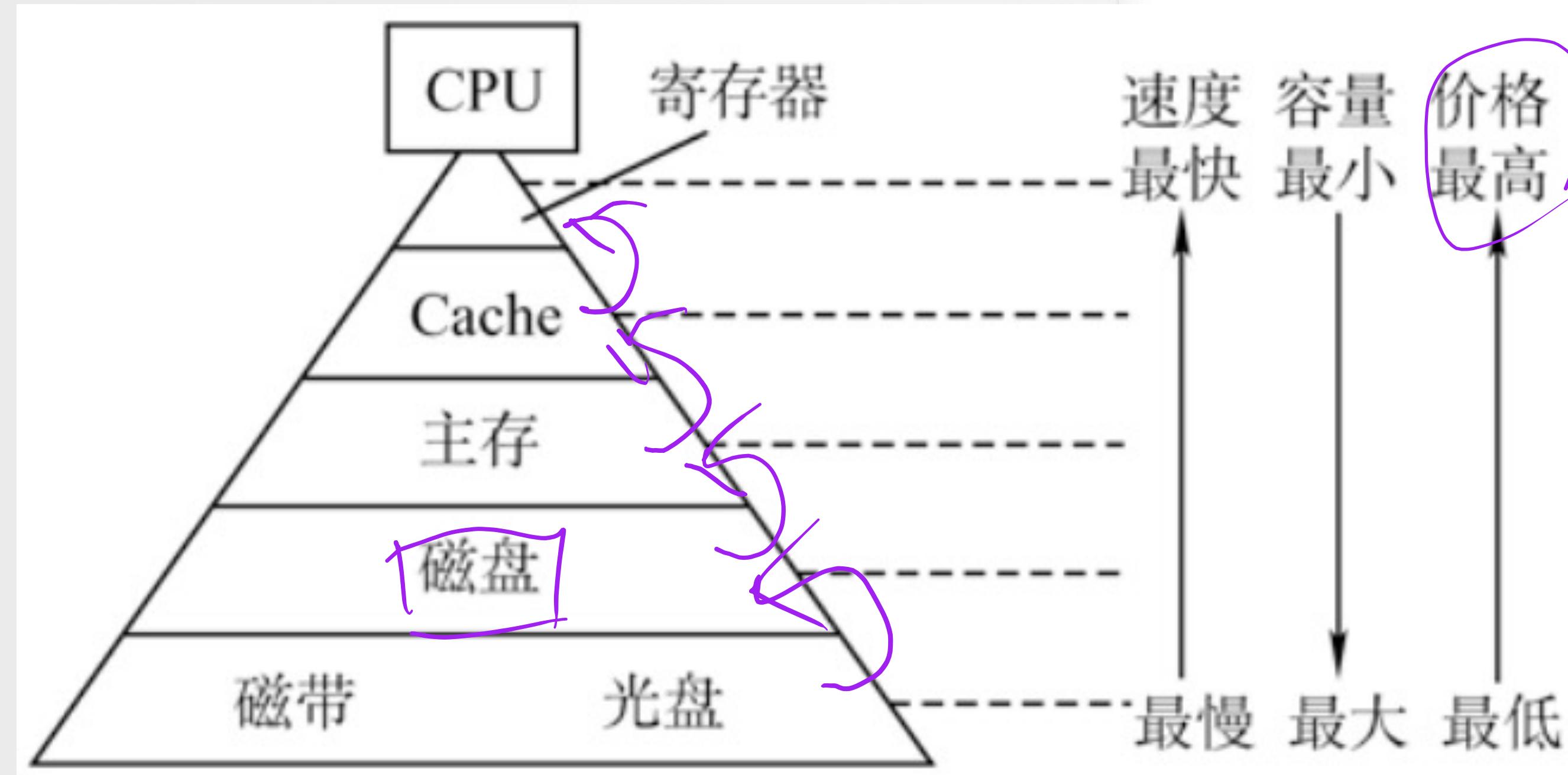
Note: System mode uses the User mode register set



2. 数据的表示与存储

2.8.1 Memory层次结构

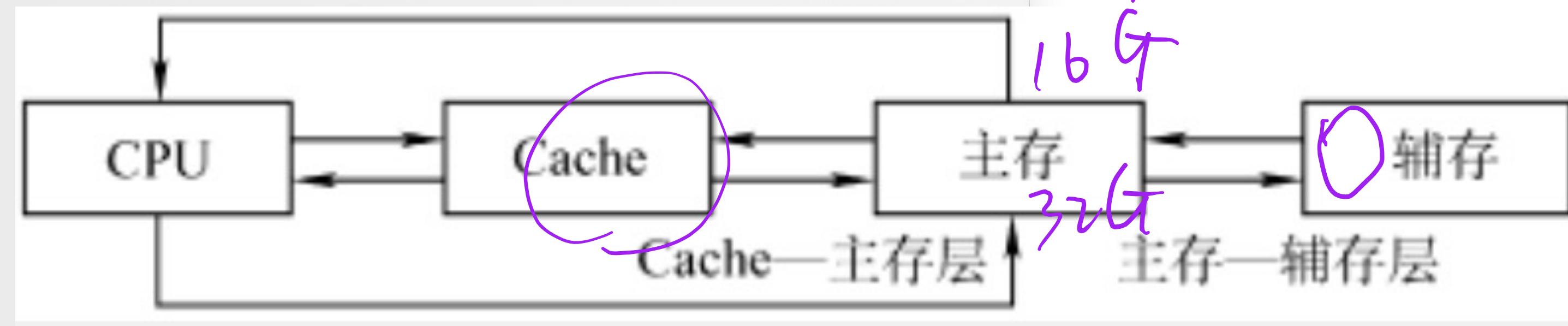
$\alpha = 10$





2. 数据的表示与存储

2.8.1 Memory层次结构



1、Cache-主存：主要对CPU加速，接近Cache，
寻址空间和价位接近主存。

这种存储结构的设计优化效果：速度快、容量大、位价低



2. 数据的表示与存储

2.8.1 存储器的性能指标

1、存储容量：存储字节数 × 字长 (如1Mx1B (8位), 1Mx4B (32位))，注意存储芯片容量标识(1Kx4, 16Kx8)

2、单位成本：每位价格 = 总成本/总容量。

$$1KB = 2^{10}B = 1024B = 2^{10} \times 8bit = 8192bit$$

$$1MB = 2^{10}KB = 2^{20}B$$

$$1GB = 2^{10}MB = 2^{20}KB = 2^{30}B$$

$$1TB = 2^{10}GB = 2^{20}MB = 2^{30}KB = 2^{40}B$$

字节

128b
8G, 100,
DDR4

48G 2条 1500
DDR5

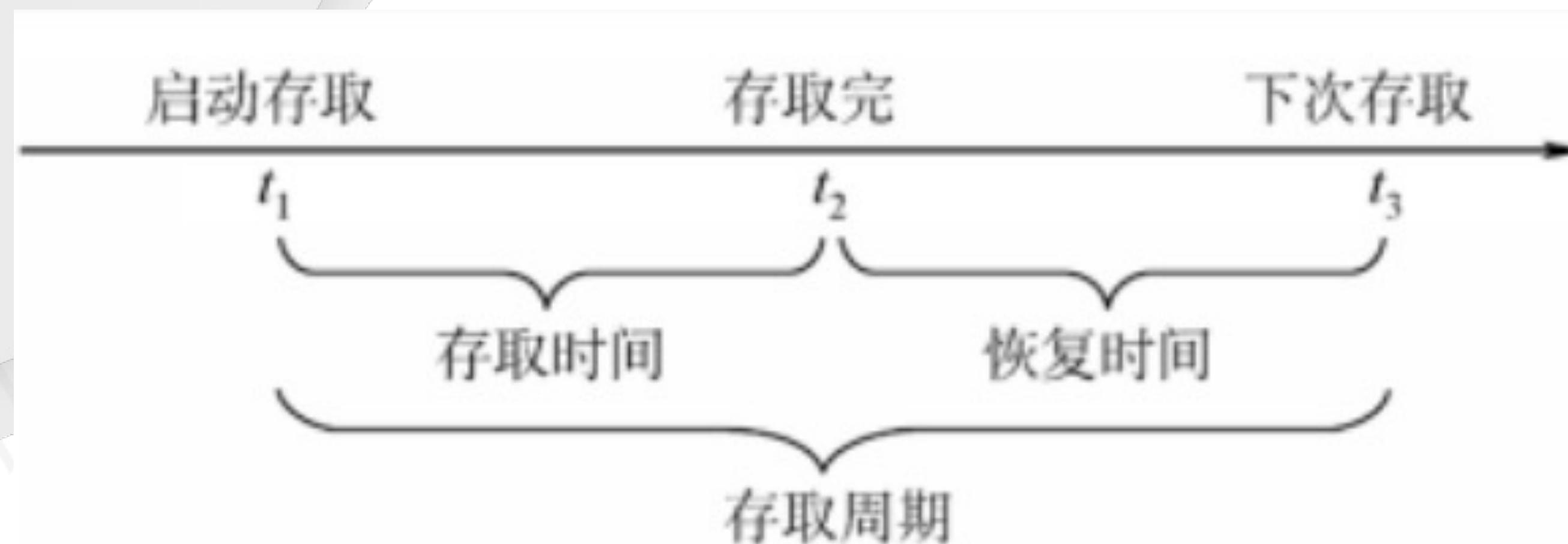


2. 数据的表示与存储

2.8.1 存储器的性能指标

3、存储速度：数据传输速度 = 数据的宽度/存取周期（或存储周期）

- ①存取时间Ta：启动一次存储操作到完成
- ②存取周期Tm：分读写周期和访问周期，两次完整的读或写全部
- ③主存带宽Bm：数据传输率，每秒从主存进出信息的最大数量





2. 数据的表示与存储

2.8.1 练习

32bit.

例1：设机器字长为32位，一个容量为16MB的存储器，CPU按半字寻址，其可寻址的单元数是（）。

- A 2^{24} B 2^{23} C 2^{22}

- D 2^{21}

$$16 \times 2^{20} = \frac{2^{28} \text{ byte}}{2 \text{ byte}} = 2^{23}$$



2. 数据的表示与存储

2.8.1 存储器分类

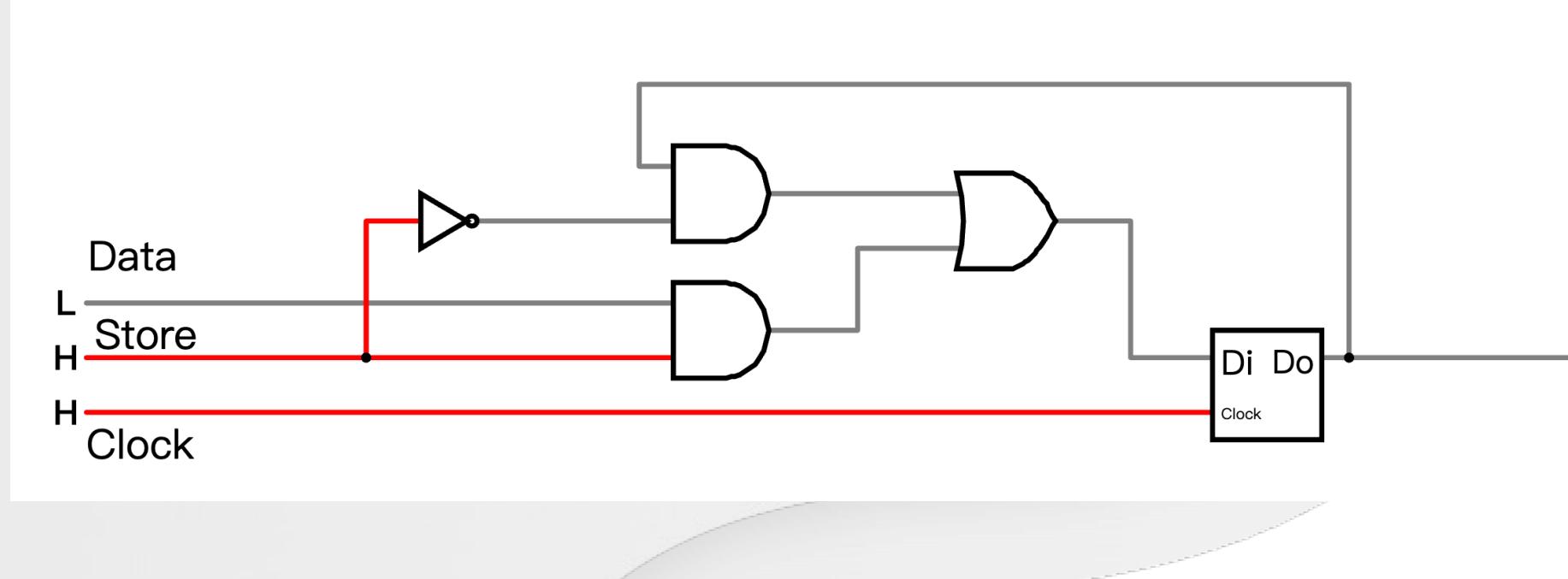
- 1、易失性存储器： RAM (DRAM) 、 Register、 Cache (SRAM)
- 2、非易失性存储器： ROM、 Flash、 EPROM、 EEPROM、 磁盘、 光存
储、 SSD、 U盘、 SD卡 *Nand*
- 3、只读存储器： ROM
- 4、串行访问存储器： 磁盘、 光盘、 磁带（顺序存取）
- 5、半导体随机存取存储器： SRAM、 DRAM、 Flash



2. 数据的表示与存储

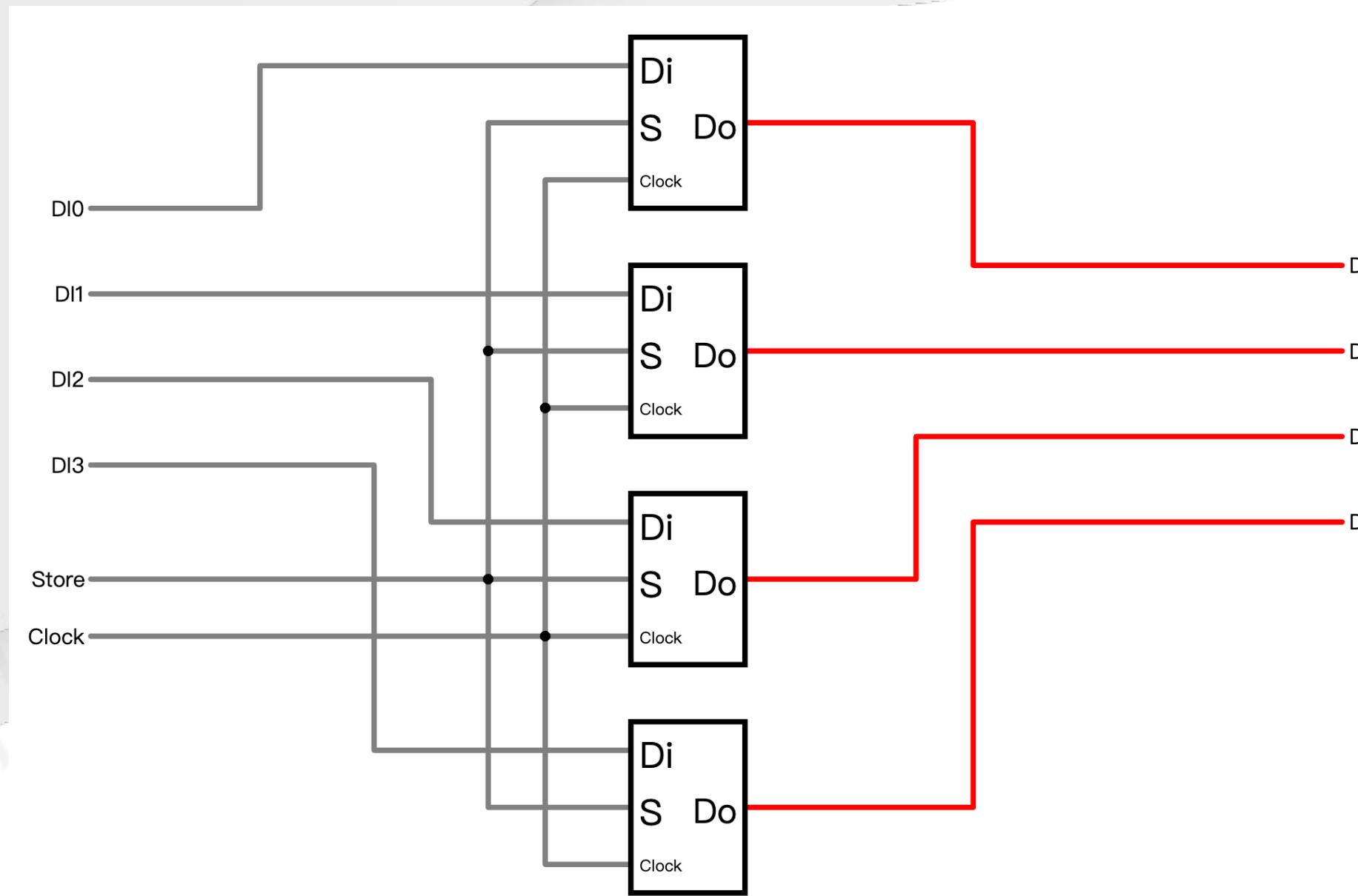
2.8.2 主存储器：SRAM

SRAM: Static Random Access Memory, 六MOS双稳态触发器。



特点：速度快、集成度低、功耗

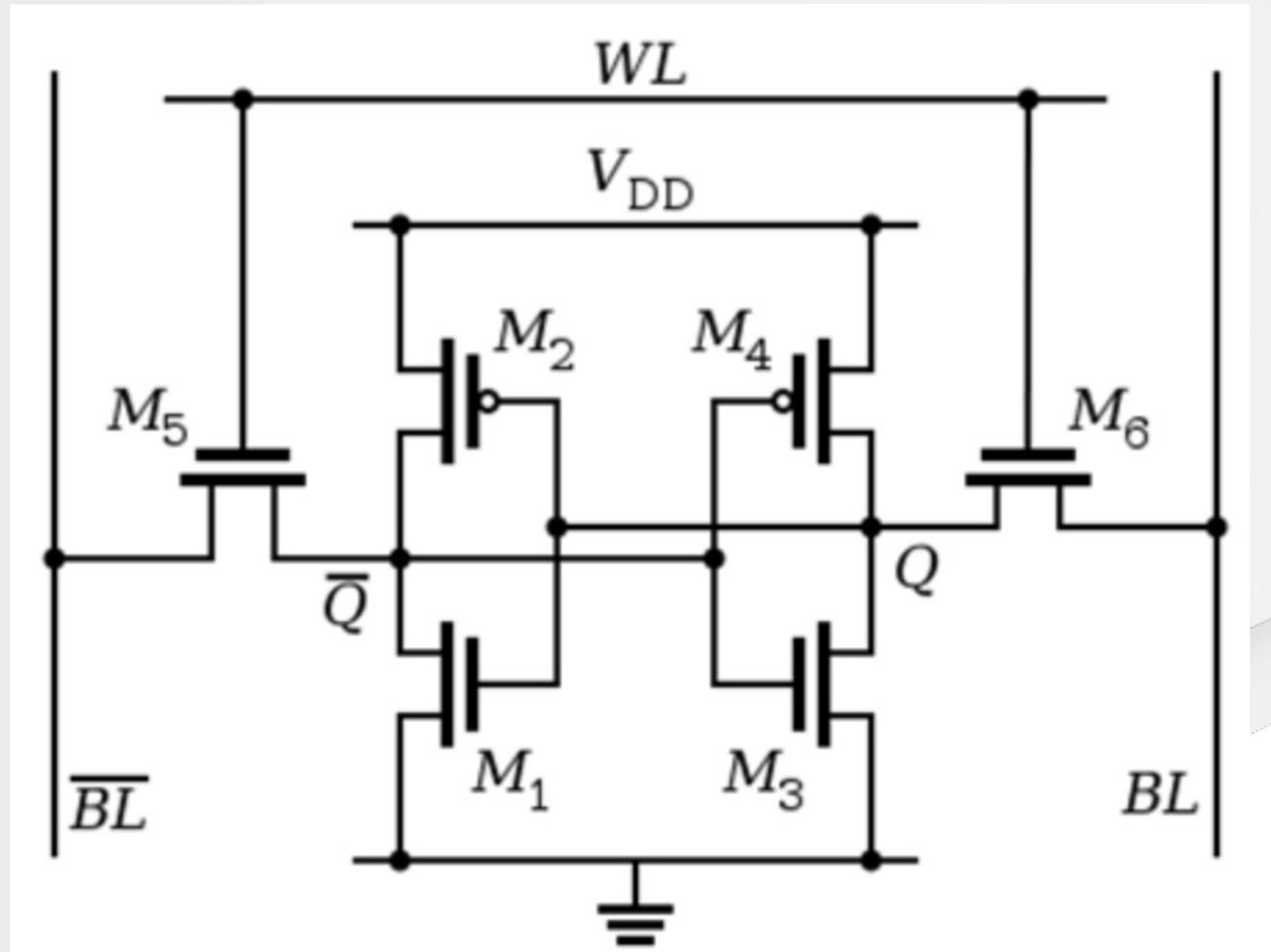
较大、价格贵、一般用于Cache





2. 数据的表示与存储

2.8.2 主存储器：SRAM

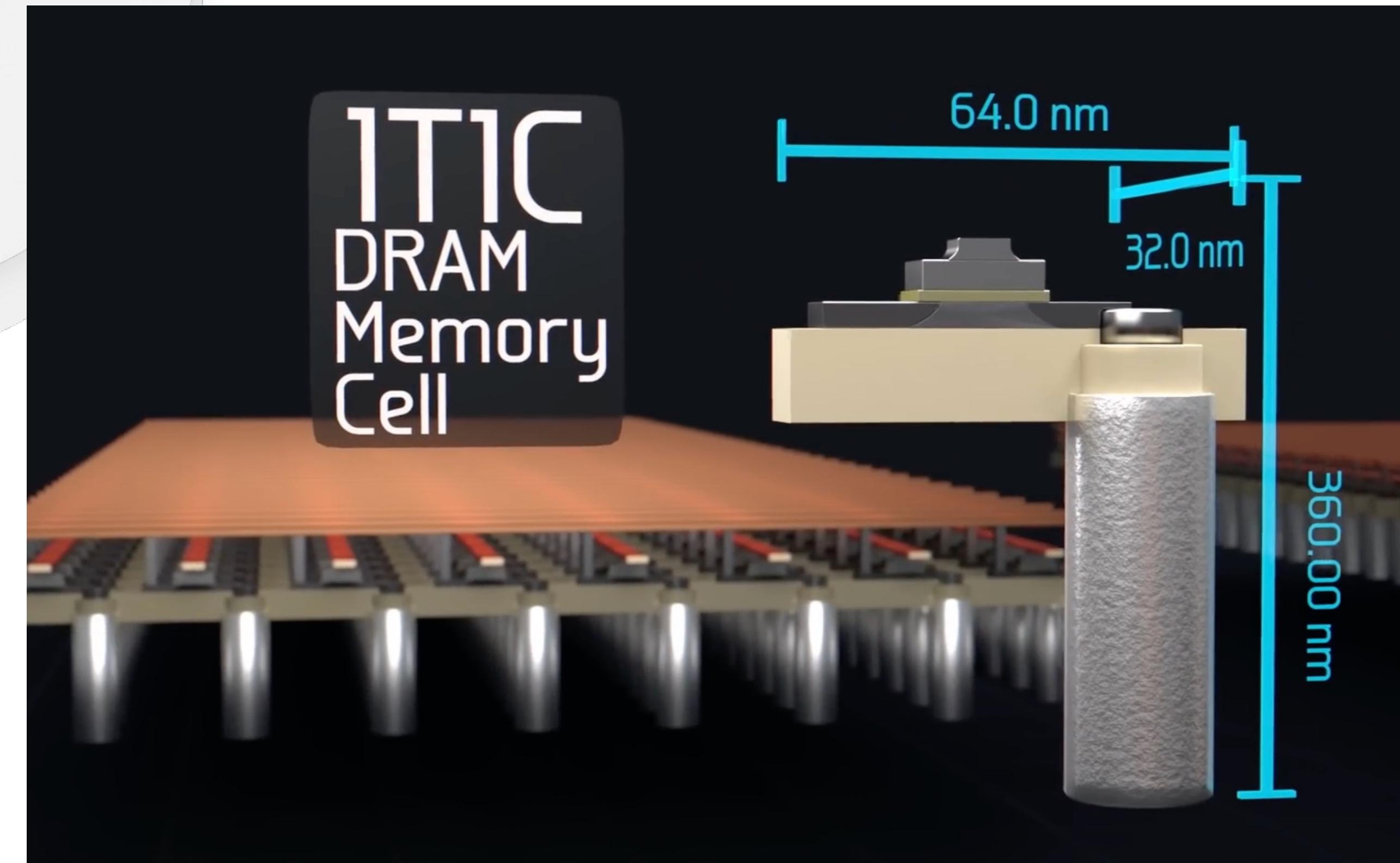
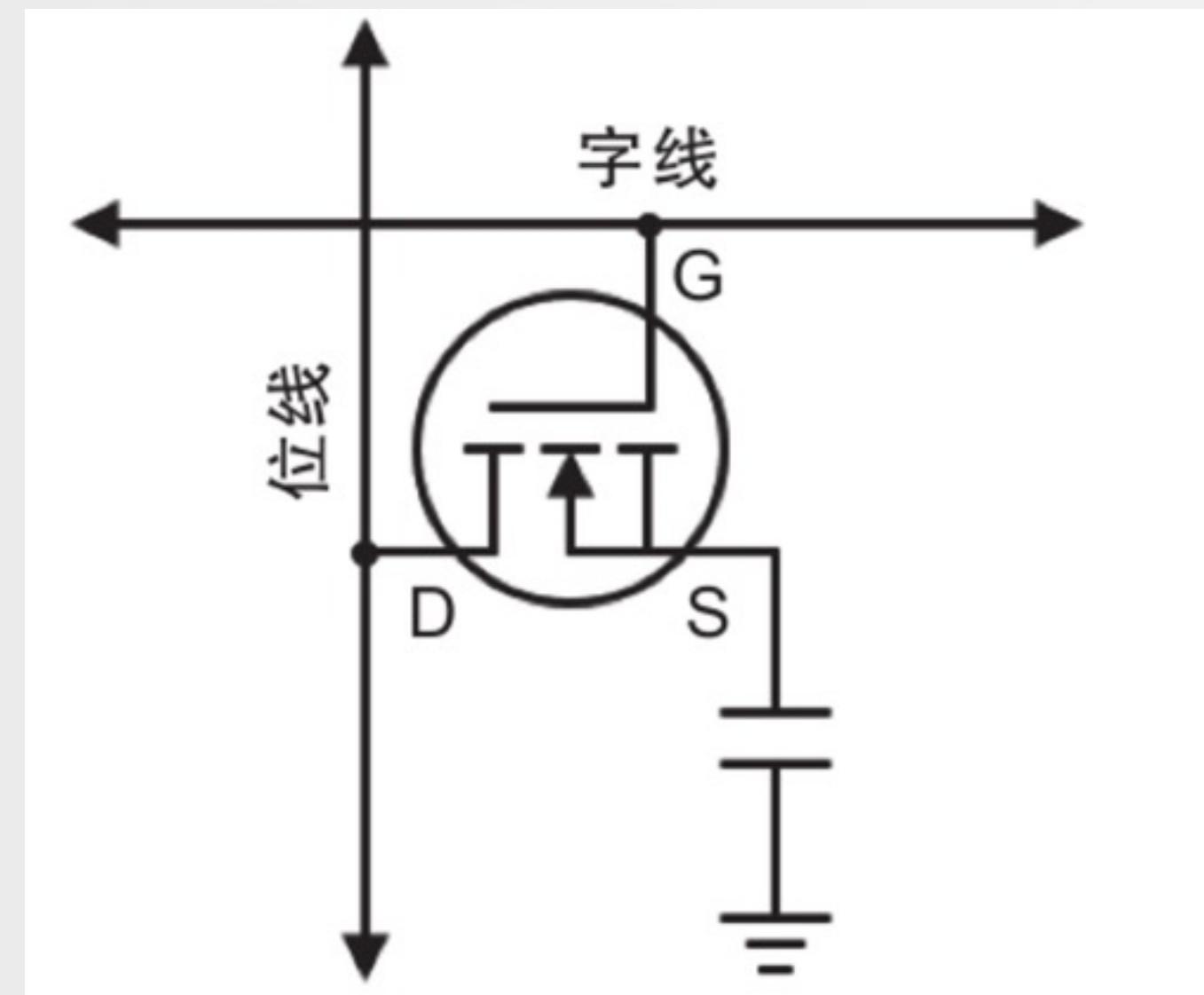




2. 数据的表示与存储

2.8.2 主存储器：DRAM

DRAM: Dynamic Random Access Memory, 单管型和多管型。



2. 数据的表示与存储

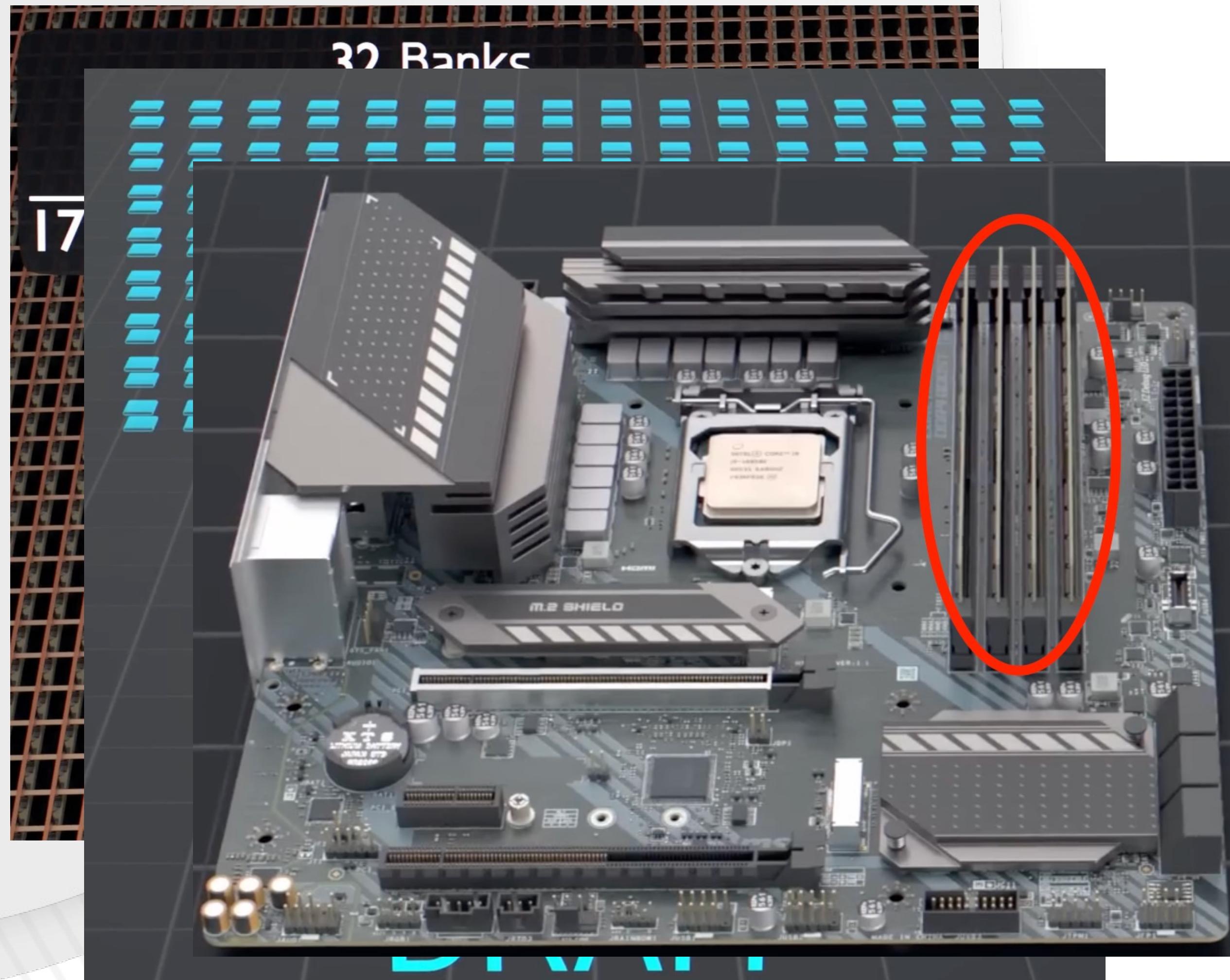
2.8.2 主存储器：DRAM





2. 数据的表示与存储

2.8.2 主存储器：DRAM



特点：集成度高、功耗低、价格低、
容量大、速度慢，一般用于主存。



2. 数据的表示与存储

2.8.2 DRAM的集中刷新



128x128的存储矩阵

w bit
存取周期 t_c : 0.5 μ s



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

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

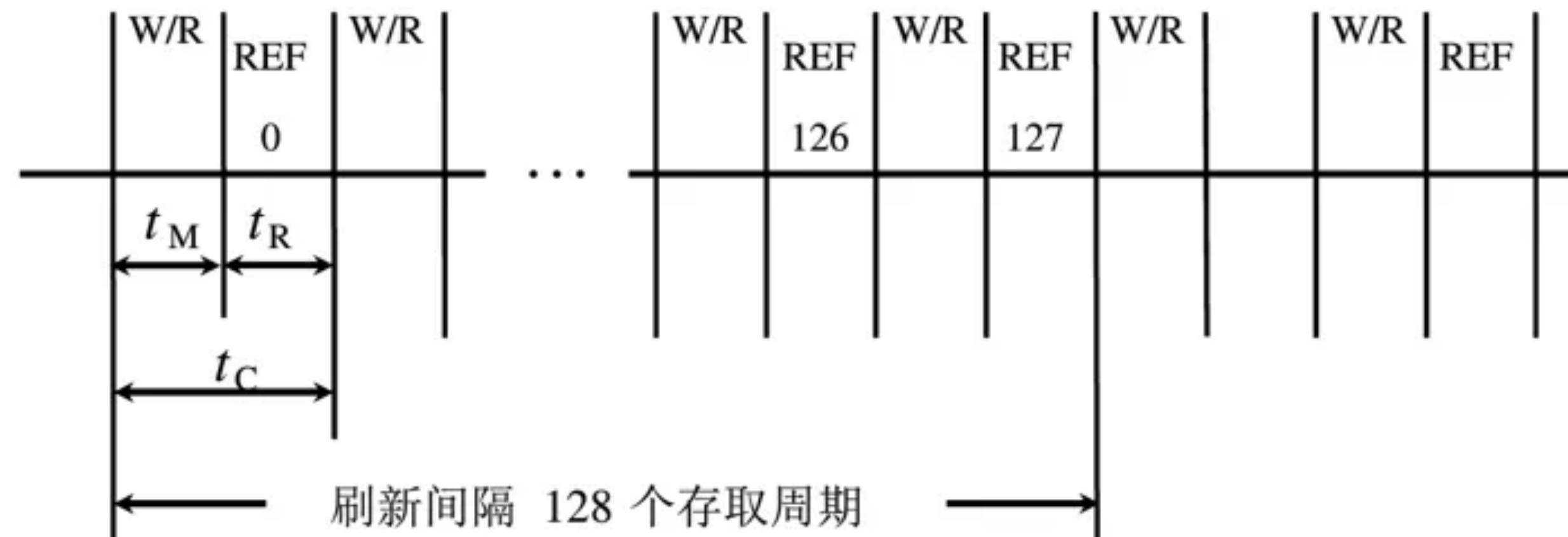
[@51CTO博客](https://blog.51cto.net)



2. 数据的表示与存储



2.8.2 DRAM的分散刷新



$$t_C = t_M + t_R$$

无“死区”

读写 刷新 (存取周期为 $0.5 \mu s + 0.5 \mu s$)

128x128的存储矩阵

读写周期tm: $0.5 \mu s$

<https://blog.csdn.net/@51CTO博客>

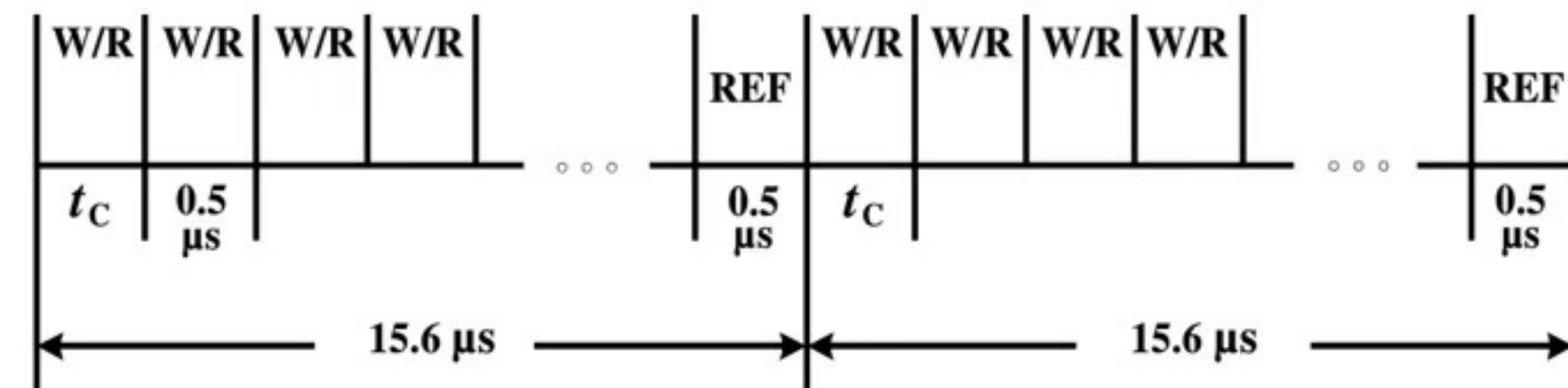


2. 数据的表示与存储

2.8.2 DRAM的异步刷新



128x128的存储矩阵



读写周期 t_m : 0.5μs

2000μs/128: 15.6μs

每行每隔 2 ms 刷新一次

“死区”为 0.5 μs



2. 数据的表示与存储

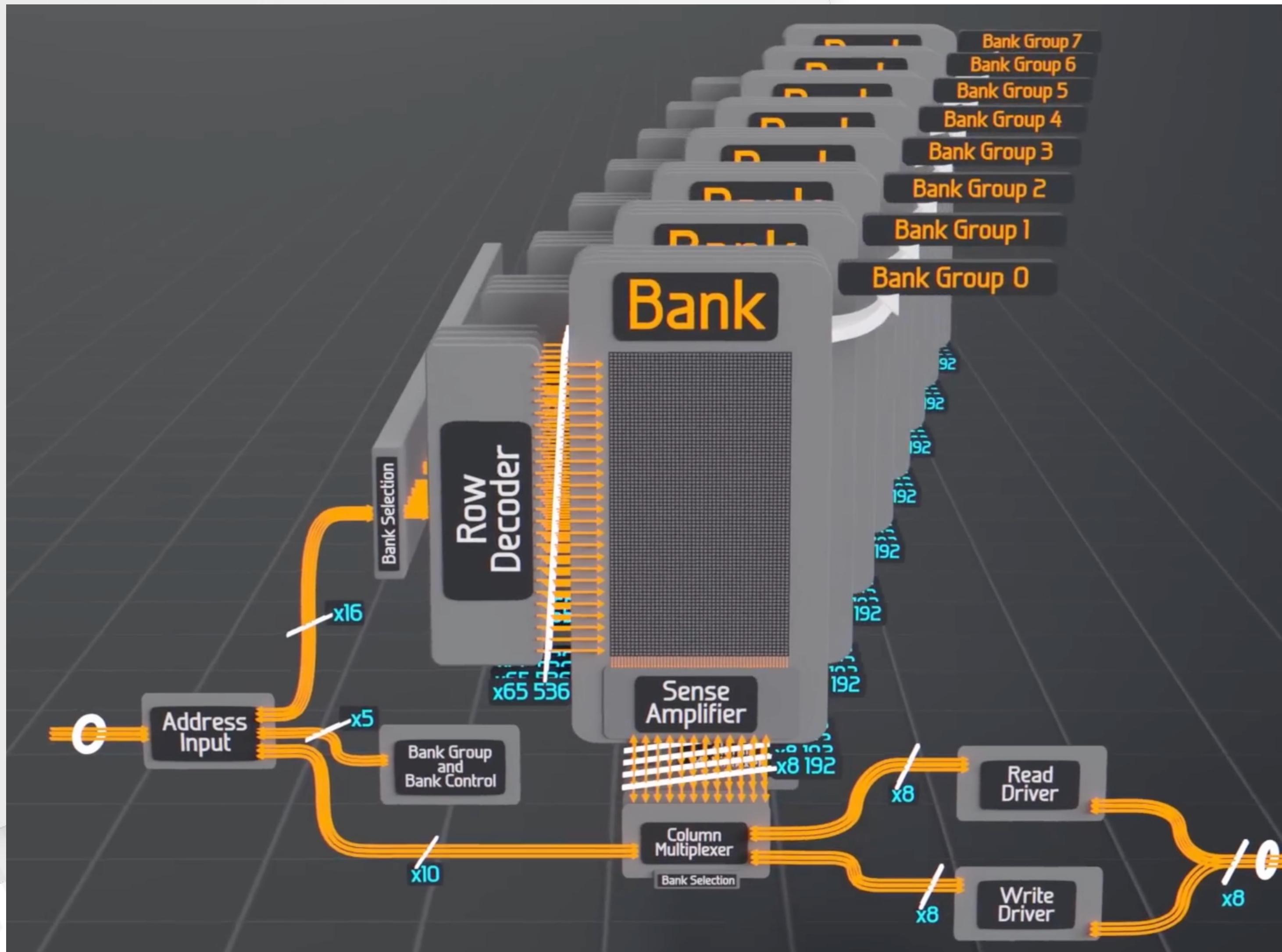
2.8.2 SRAM和DRAM

特点 \ 类型	SRAM	DRAM
存储信息	触发器	电容
破坏性读出	非	是
需要刷新	不要	需要
送行列地址	同时送	分两次送
运行速度	快	慢
集成度	低	高
发热量 (功耗)	大	小
存储成本	高	低
主要用途	高速缓存	主机内存



2. 数据的表示与存储

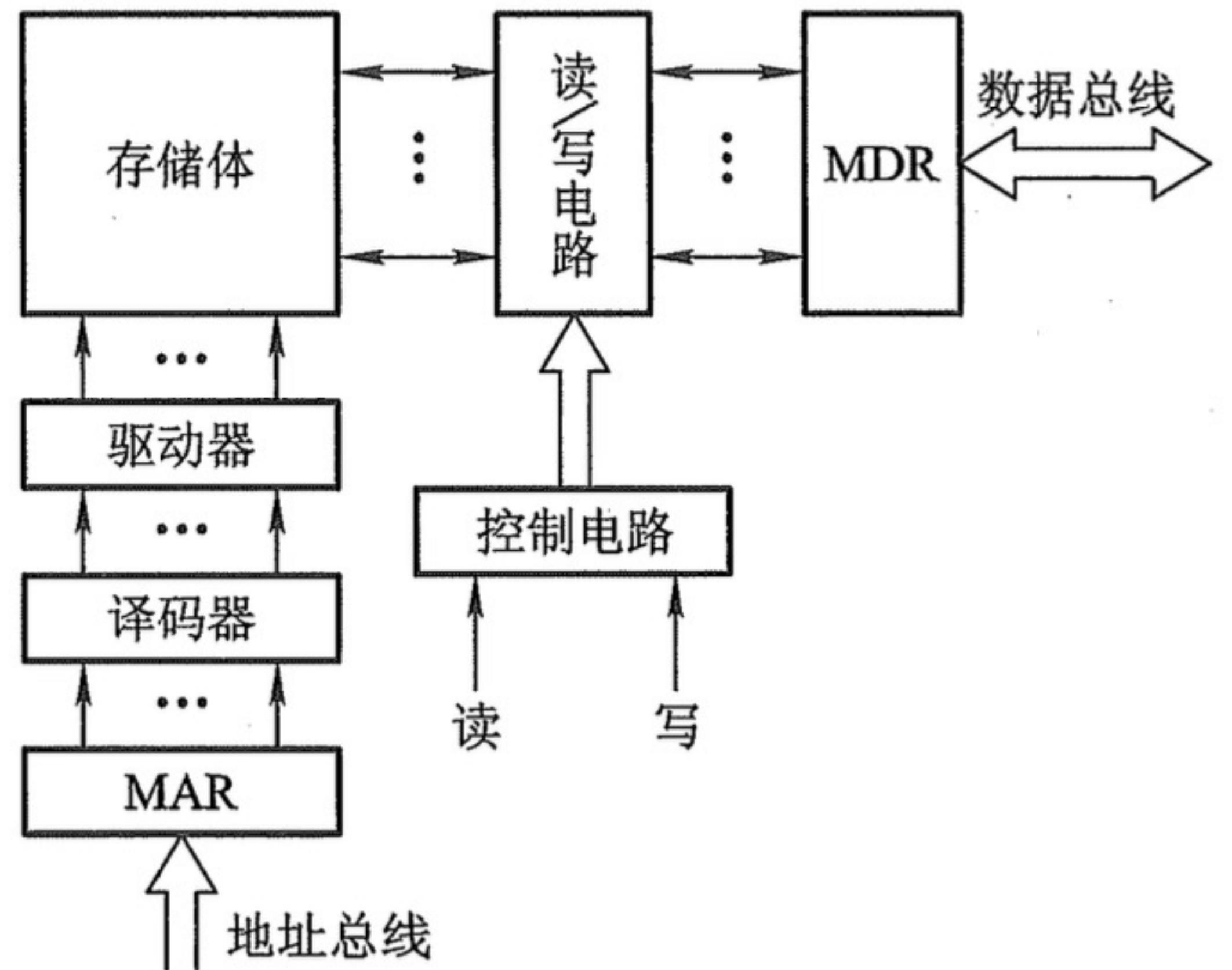
2.8.2 存储芯片的内部结构





2. 数据的表示与存储

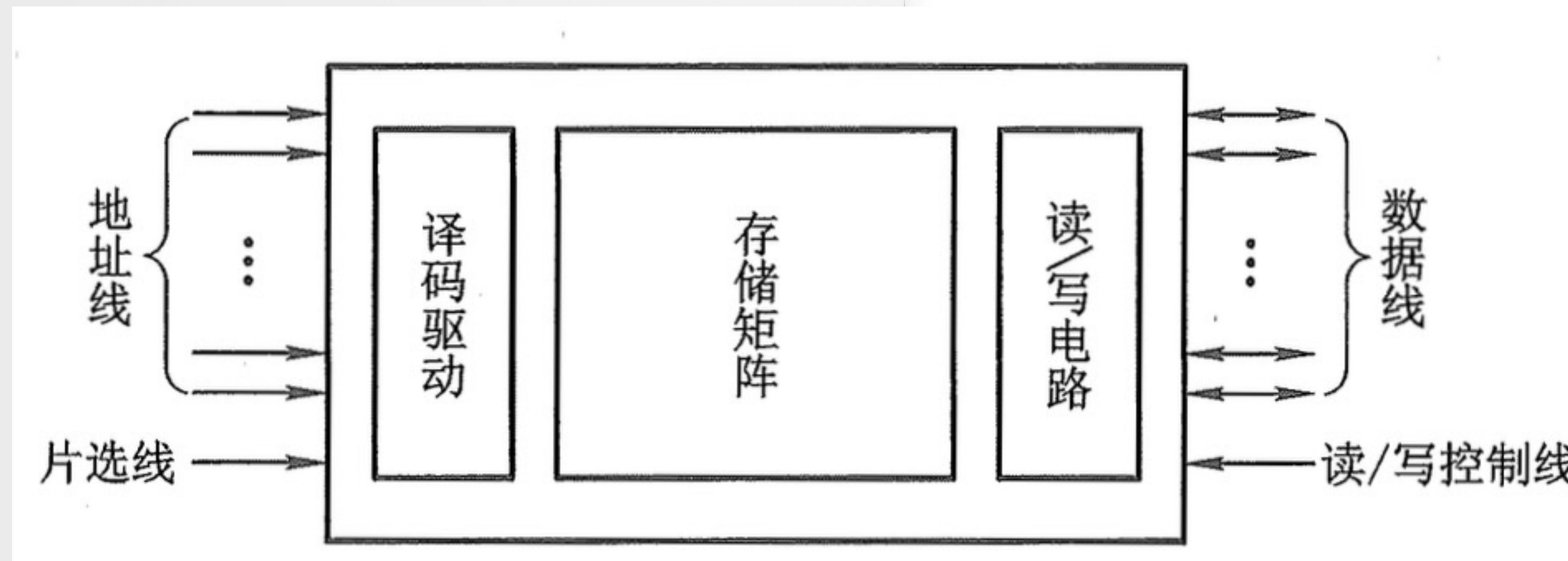
2.8.2 存储芯片的内部结构





2. 数据的表示与存储

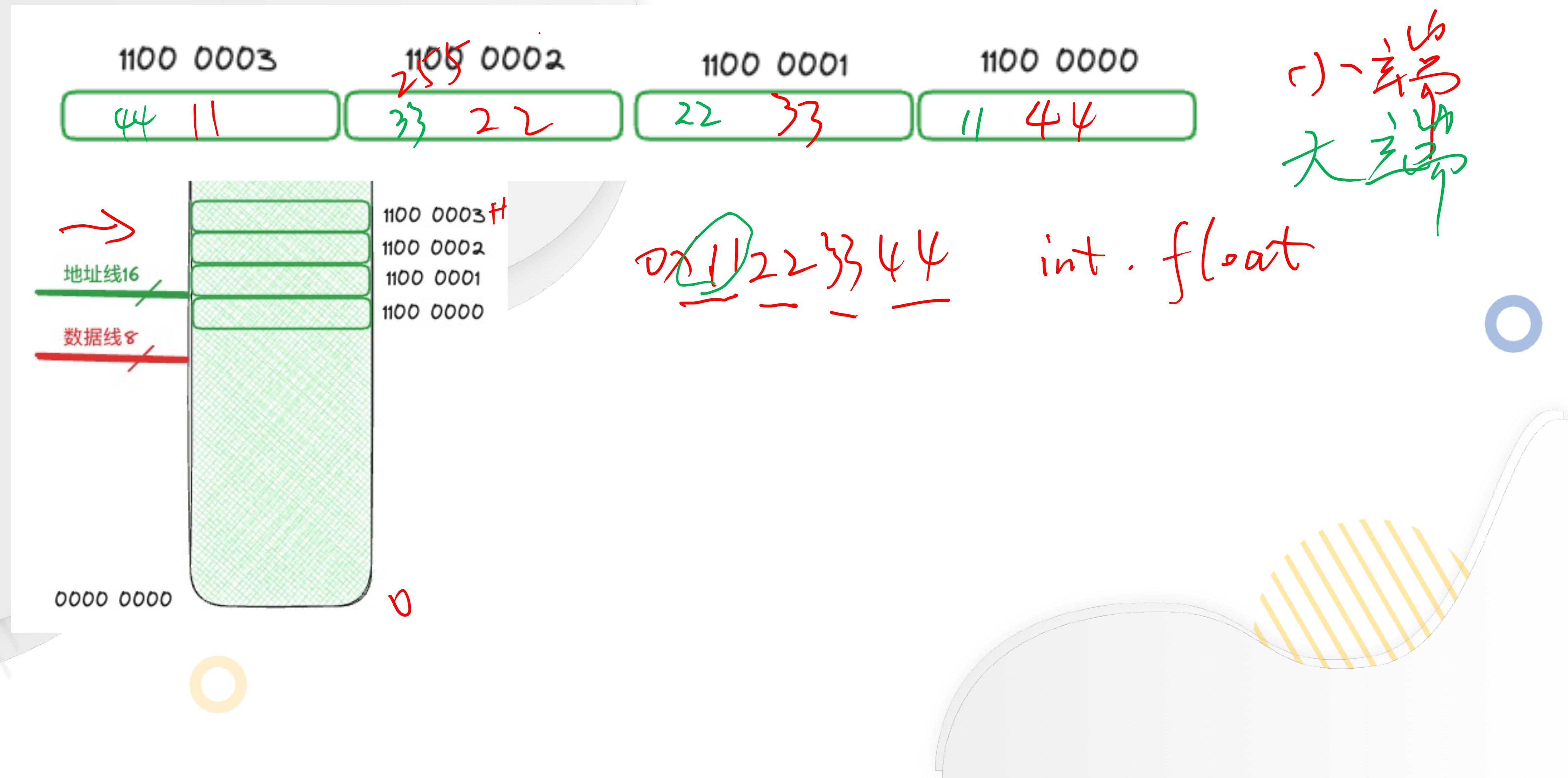
2.8.2 存储单元的地址线与数据线





2. 数据的表示与存储

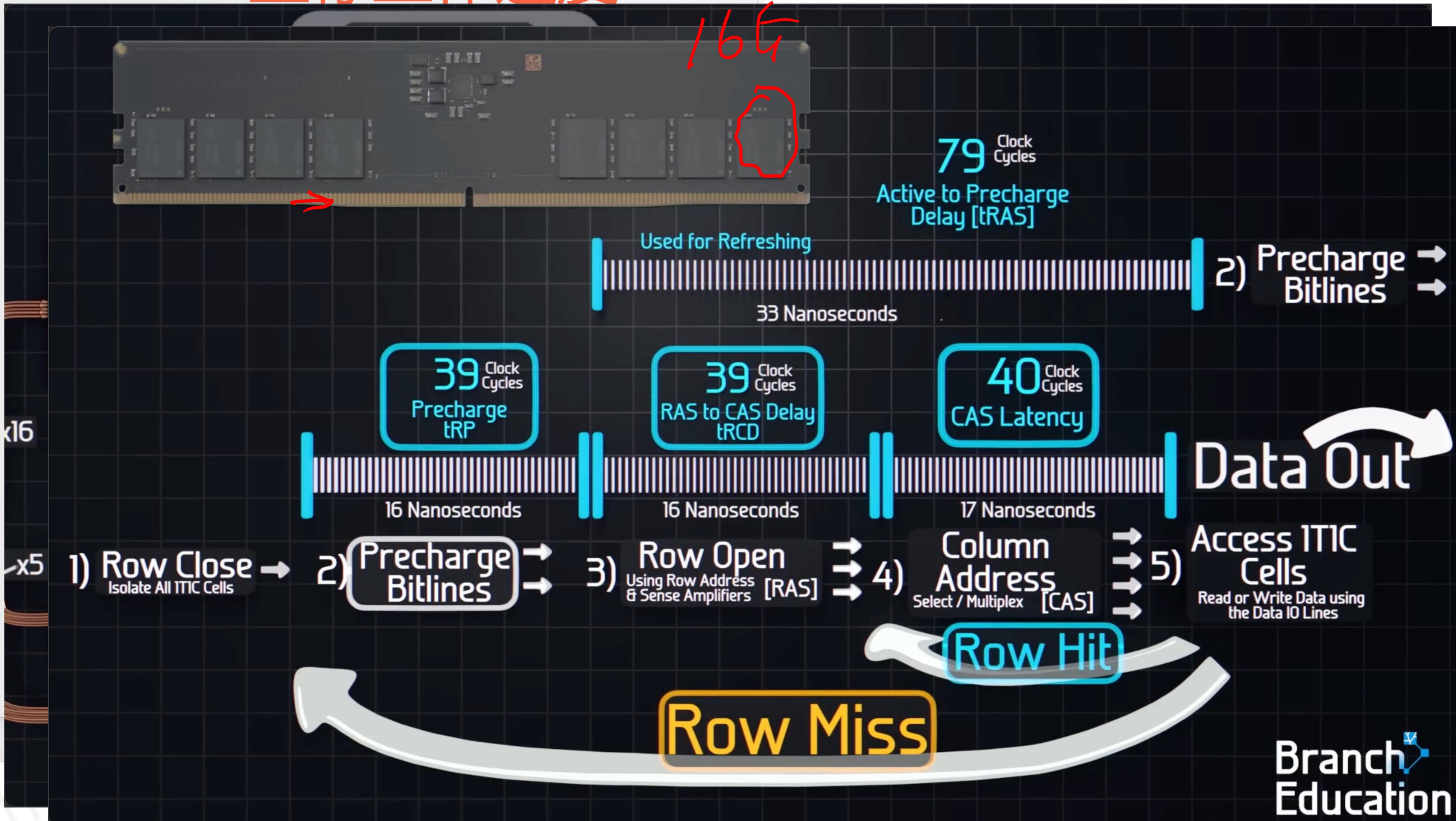
2.8.2 存储单元的地址及大小端





2. 数据的表示与存储

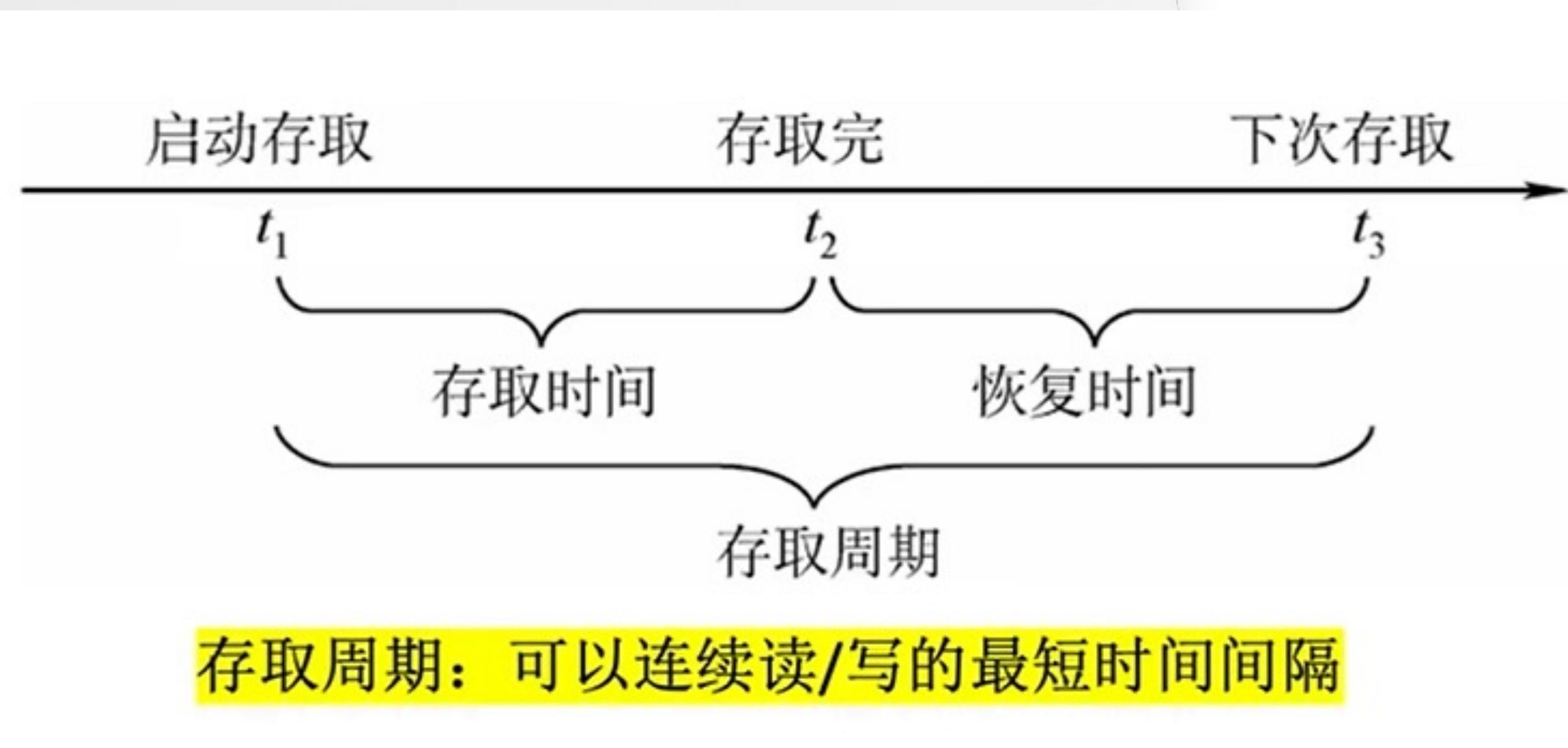
2.8.2 主存工作速度





2. 数据的表示与存储

2.8.2 CPU与主存速度匹配优化



$10 \sim 20 \text{ ns}$
DRAM, 200ns
 $\frac{1}{T_D}$

2. 数据的表示与存储

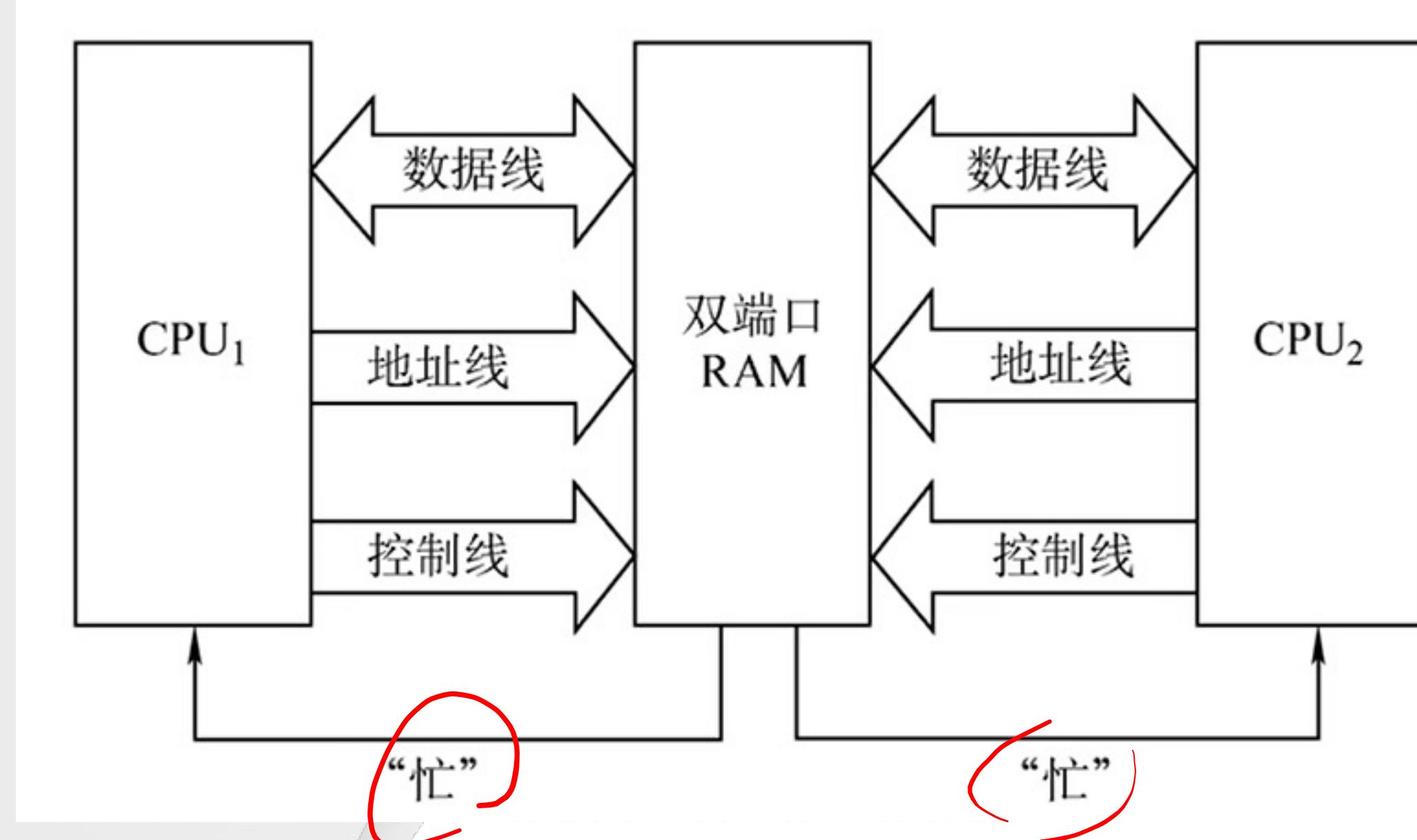
2.8.2 计算机处理速度优化方案

- 1、双口RAM
- 2、单体多字存储
- 3、多体并行存储
- 4、高速缓存Cache（多级）



2. 数据的表示与存储

2.8.2 优化方案1：双口RAM

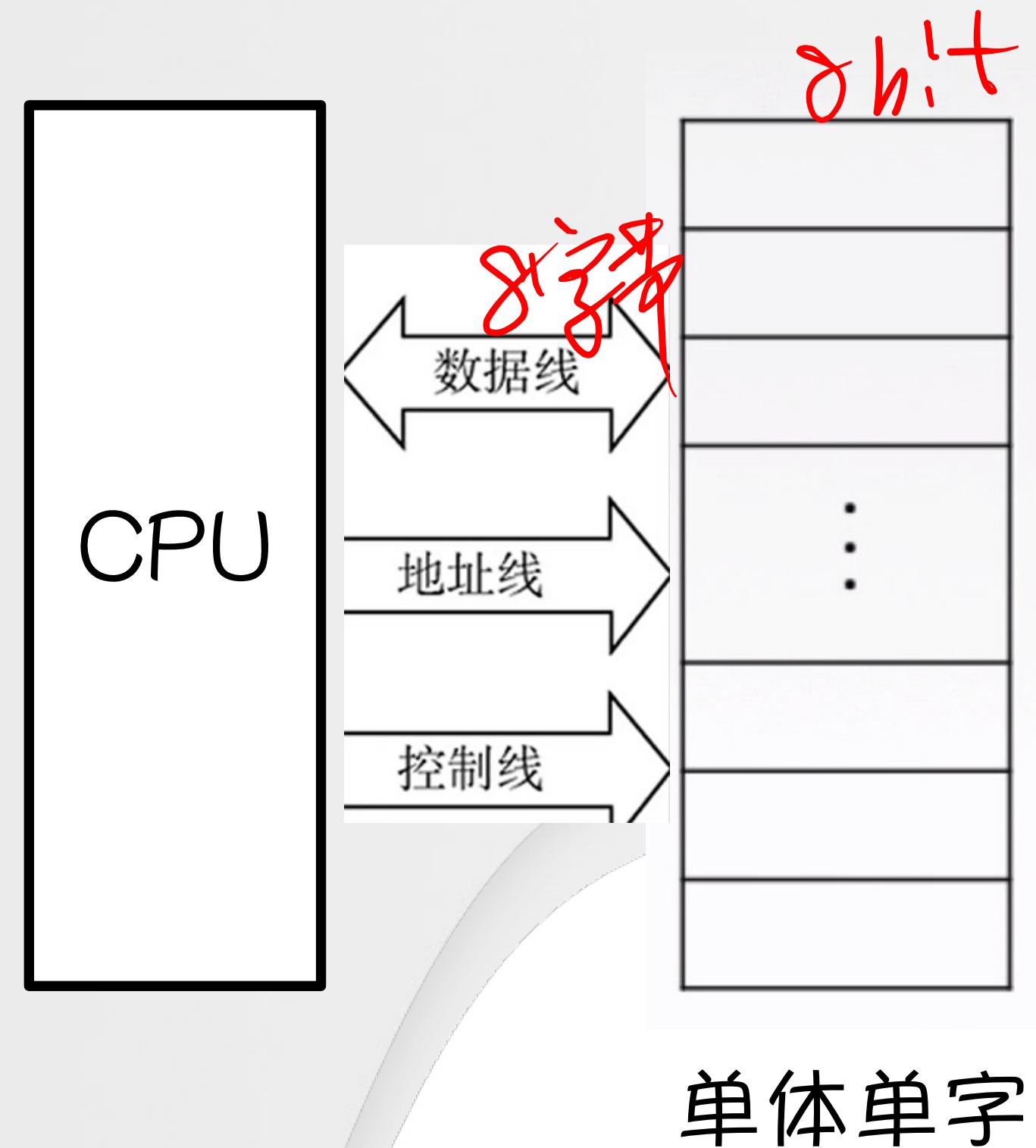


1. 两个端口同时对不同的地址单元存取数据 ✓
2. 两个端口同时对同一地址单元读出数据 ✓
3. 两个端口同时对同一地址单元写入数据 会
导致写入错误
4. 两个端口同时对同一地址，一个写入数据，
另一个读出数据 **会导致读出错误**



2. 数据的表示与存储

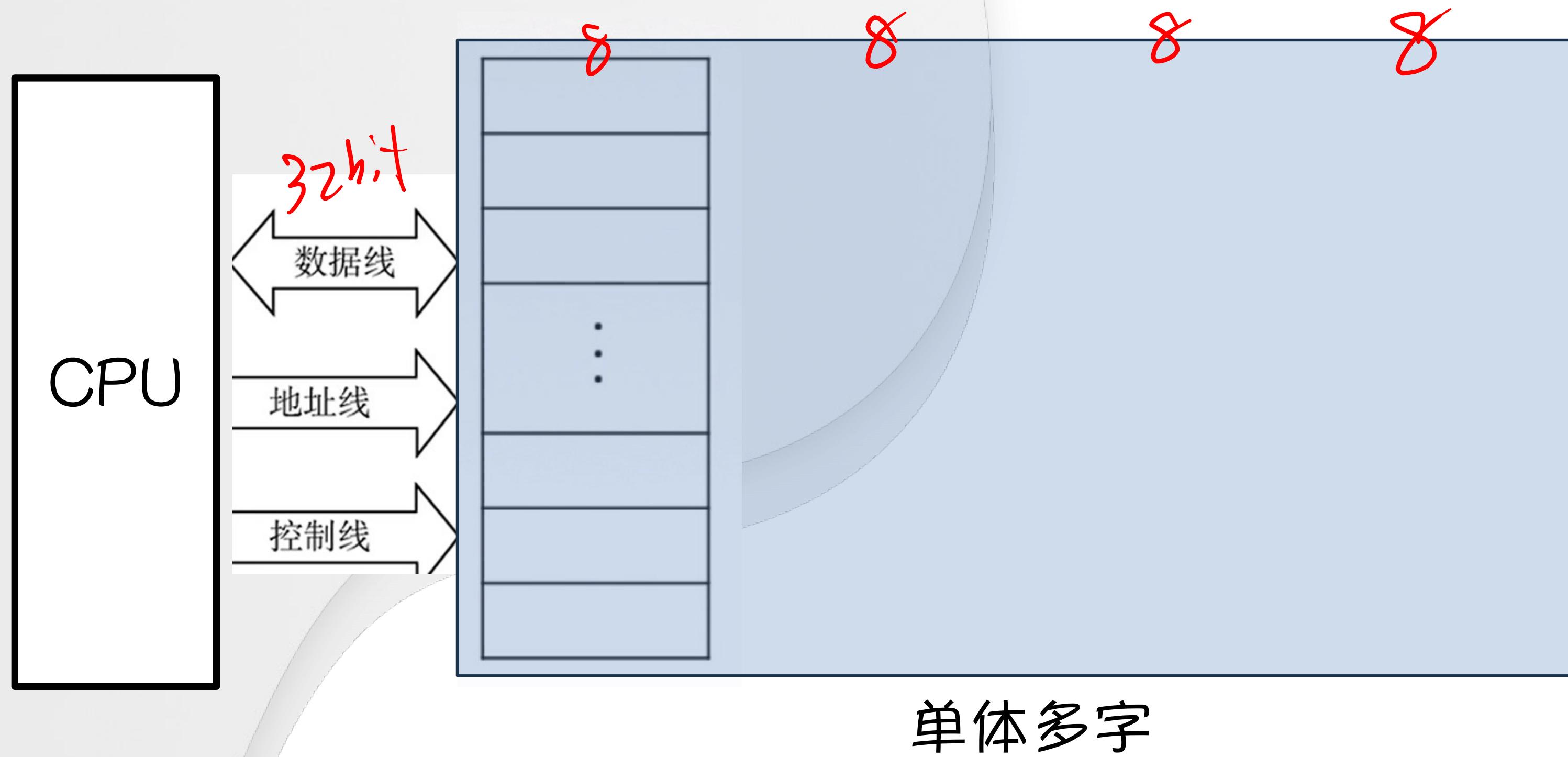
2.8.2 优化方案2：单体多字存储器





2. 数据的表示与存储

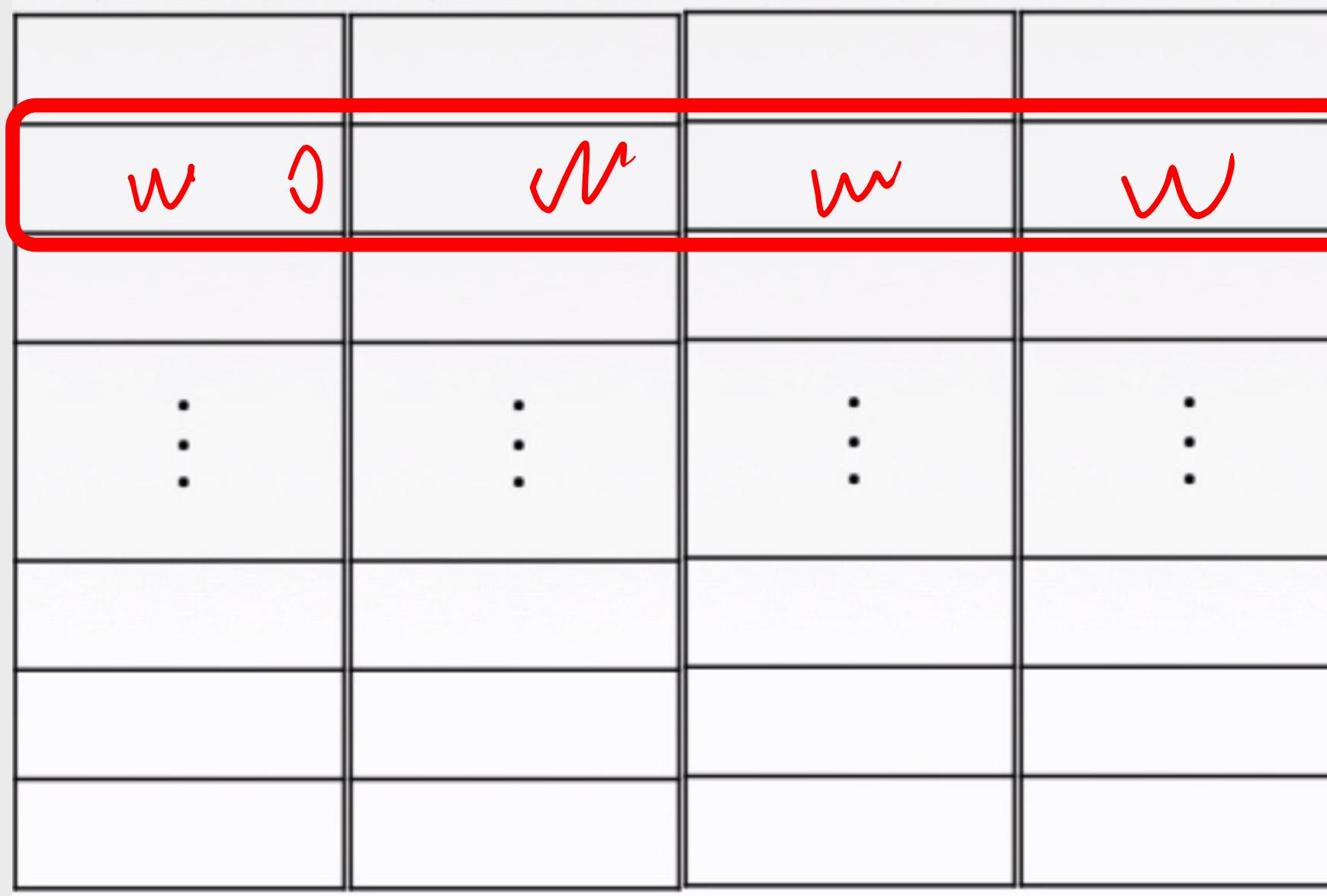
2.8.2 优化方案2：单体多字存储器





2. 数据的表示与存储

2.8.2 优化方案2：单体多字存储器



多字长数据寄存器

多字长数据寄存器

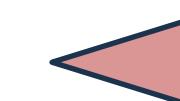


CPU

缺点：

- 1、凑齐了才能写
- 2、不一定一次读取的数据都有用

地址

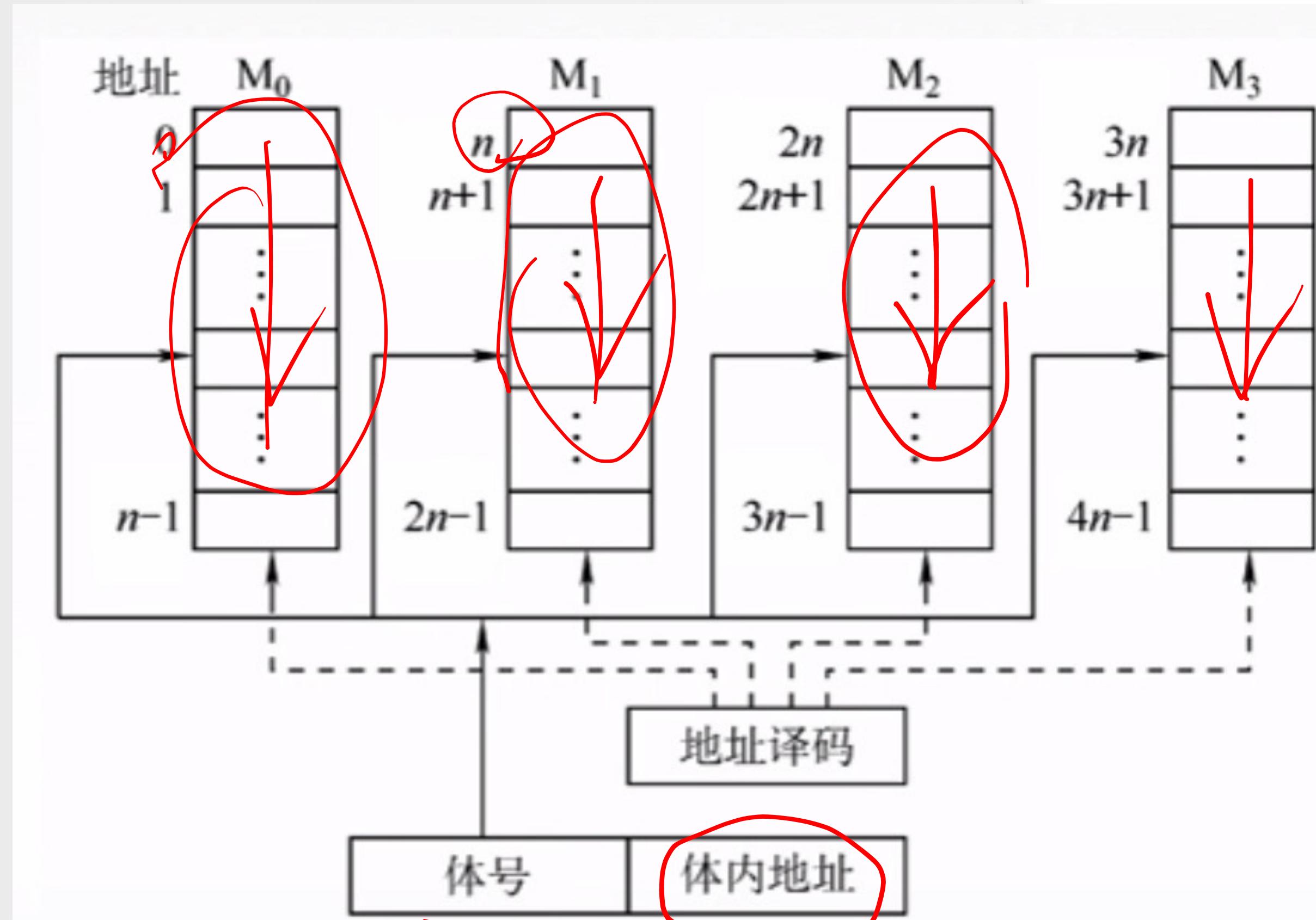


地址寄存器



2. 数据的表示与存储

2.8.2 优化方案3：多体并行之顺序存储

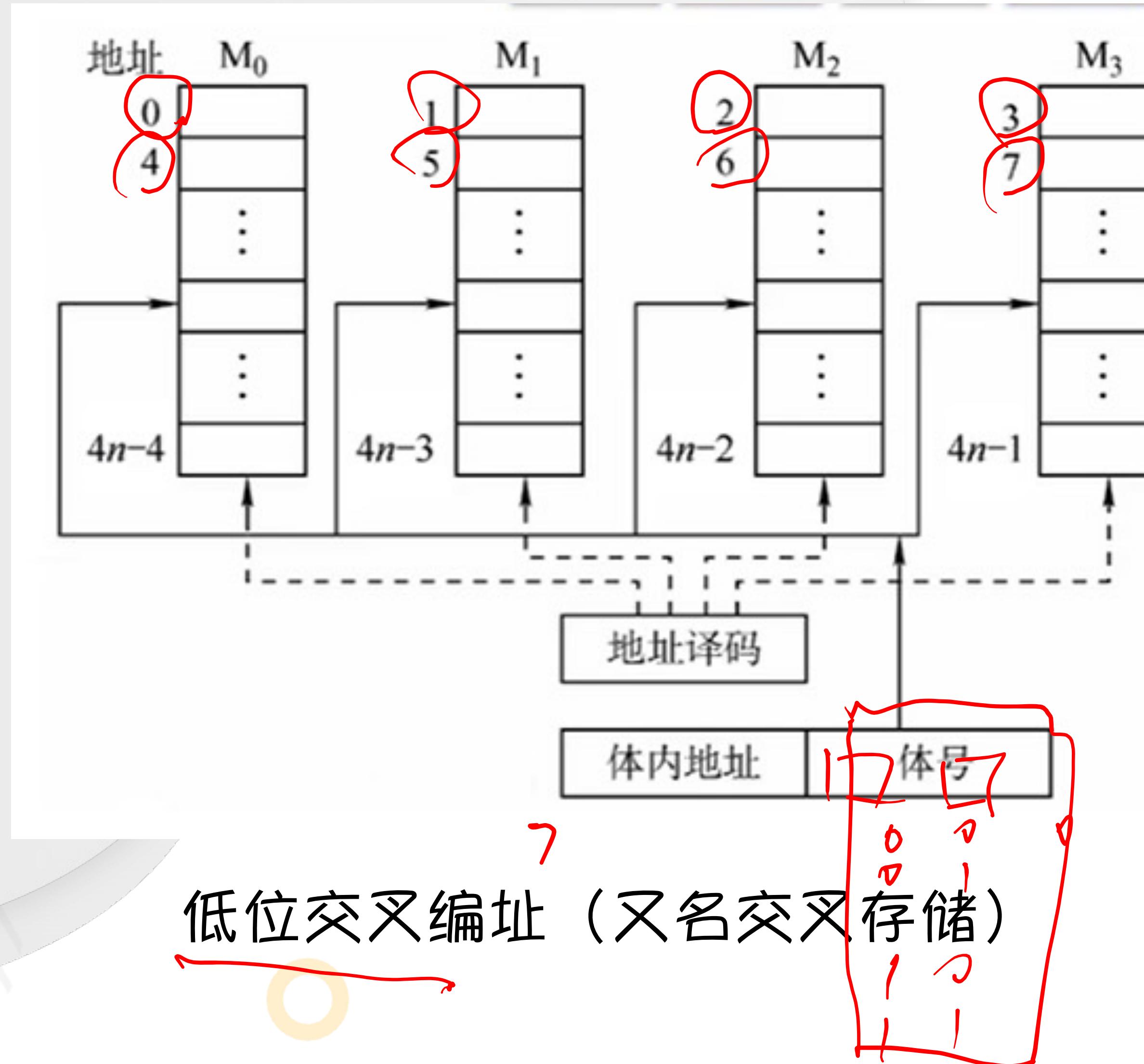


高位交叉编址 (又名顺序存储)



2. 数据的表示与存储

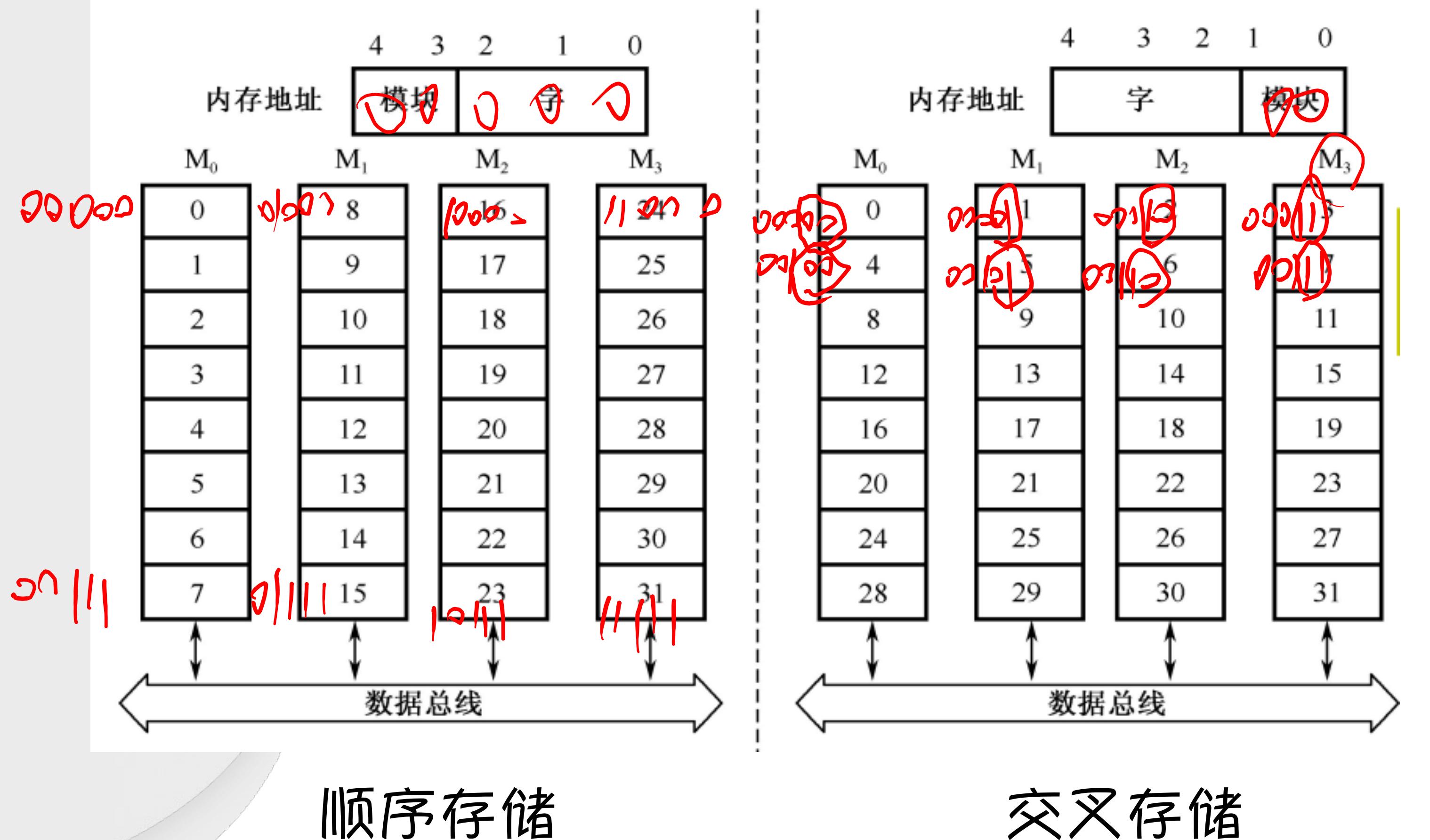
2.8.2 优化方案3：多体并行之交叉存储





2. 数据的表示与存储

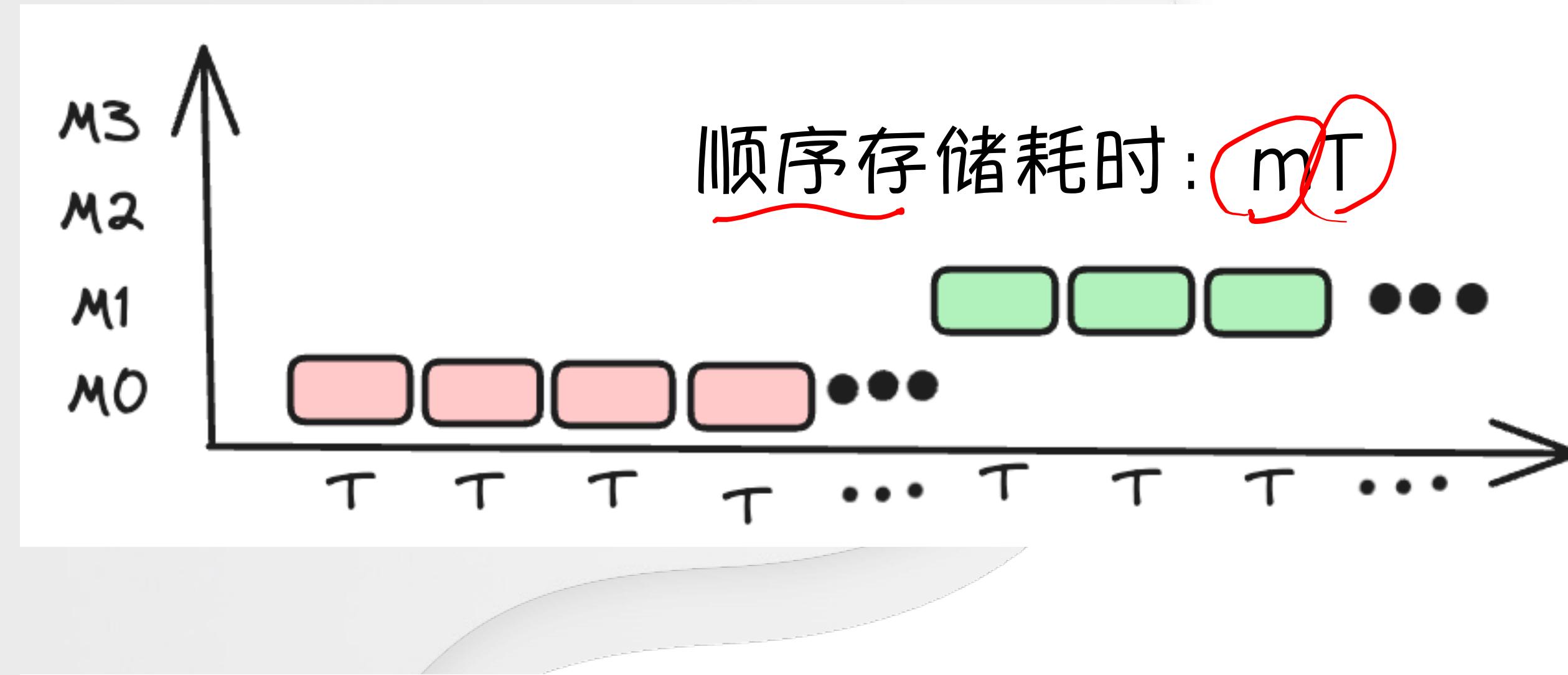
2.8.2 优化方案3：多体并行地址分布





2. 数据的表示与存储

2.8.2 多体并行耗时



说明:

模块存储周期: T

总线传送周期: r

前提条件: $m \geq T/r$

总结:

1、顺序存储的并行: 不同的请求源并行访问不同的存储体

2、交叉存储的并行: 同一请求源并行访问不同的存储体





2. 数据的表示与存储

2.8.2 练习

例2：设存储器容量为32个字，字长为64位，模块数m=4，分别采取顺序存储和交叉存储进行组织。存储周期T=200ns，数据总线宽度为64位，总线传输周期r=50ns。若连续读出4个字，两种组织的带宽是多少？

$$64 \times 4 = 256 \text{ bit}$$

$$t_{\text{读}} = 4 \times 200\text{ns} = 800\text{ns}$$

$$t_{\text{传}} = 200\text{ns} + (4-1)50\text{ns} = 350\text{ns}$$

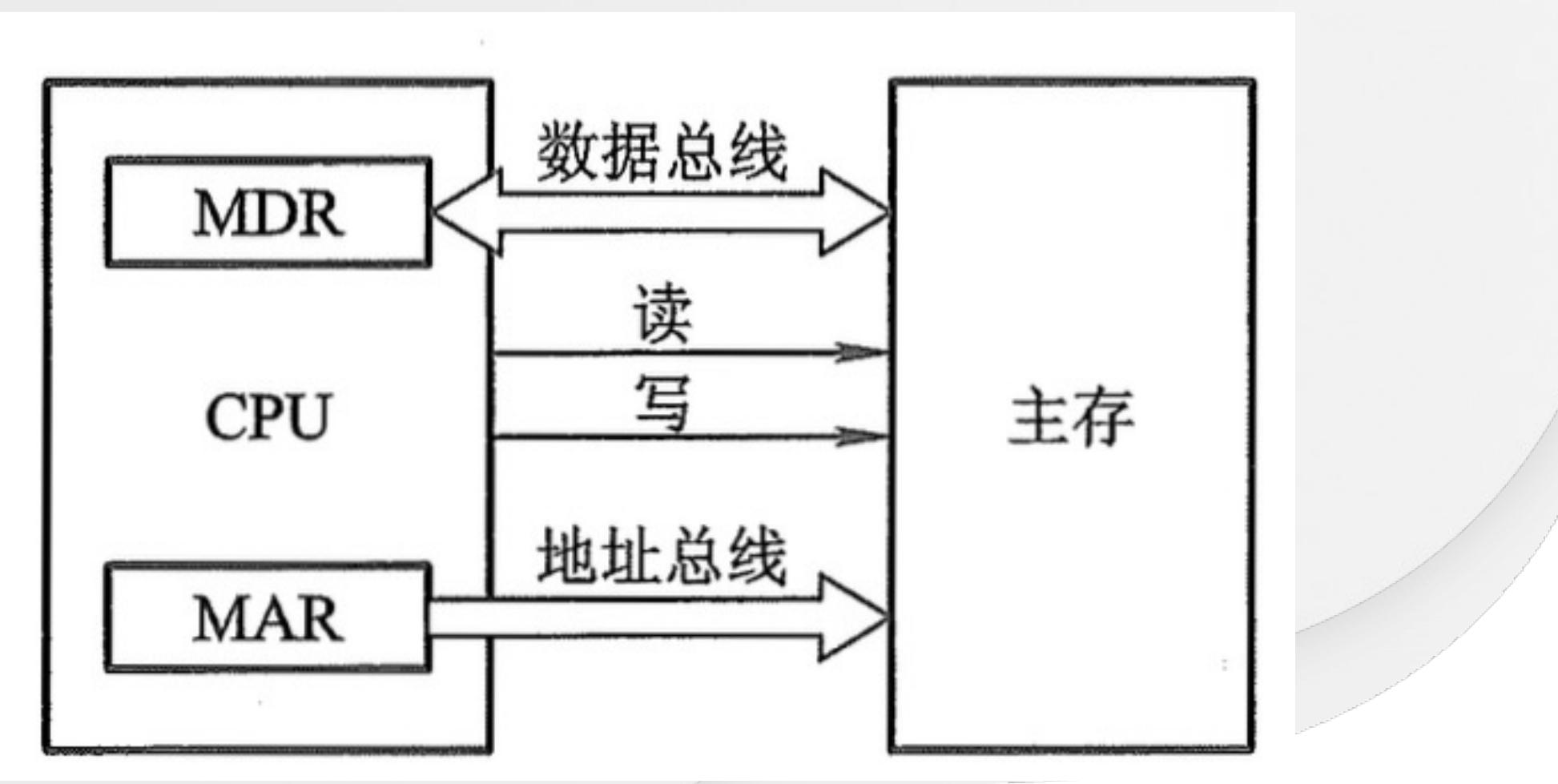
$$W_{\text{顺}} = 256 \text{ bit} / 800\text{ns} = 320 \text{ Mbit/s}$$

$$W_{\text{交}} = 256 \text{ bit} / 350\text{ns} = 731 \text{ Mbit/s}$$

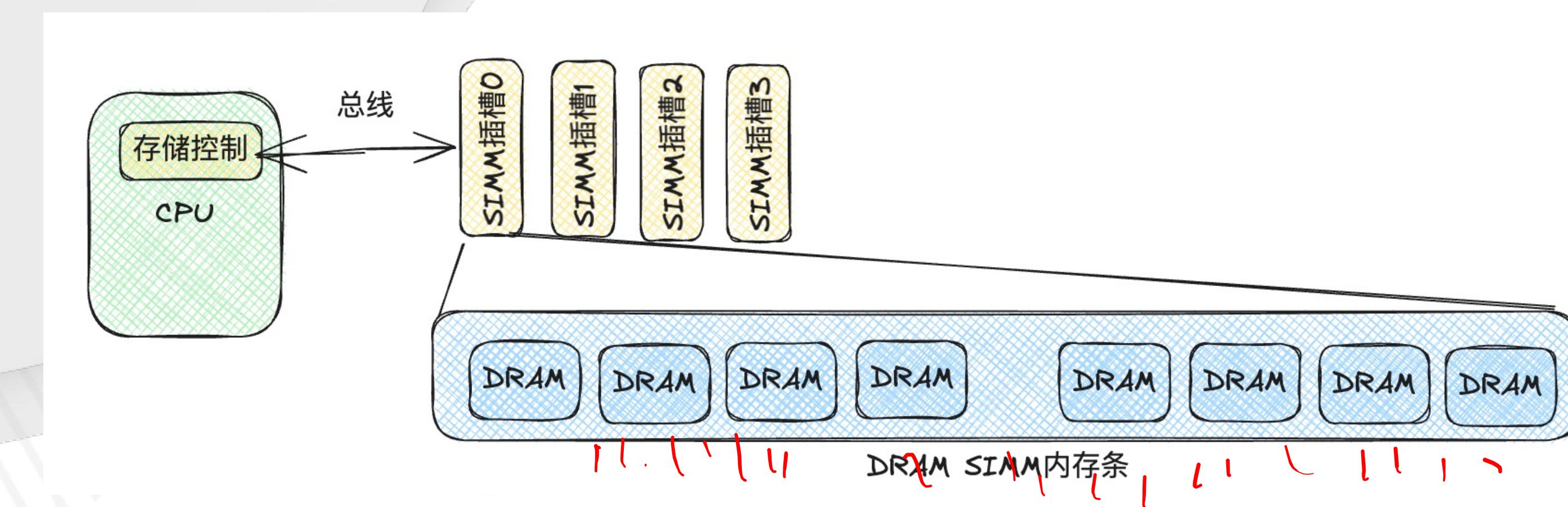


2. 数据的表示与存储

2.8.3 主存与CPU的连接



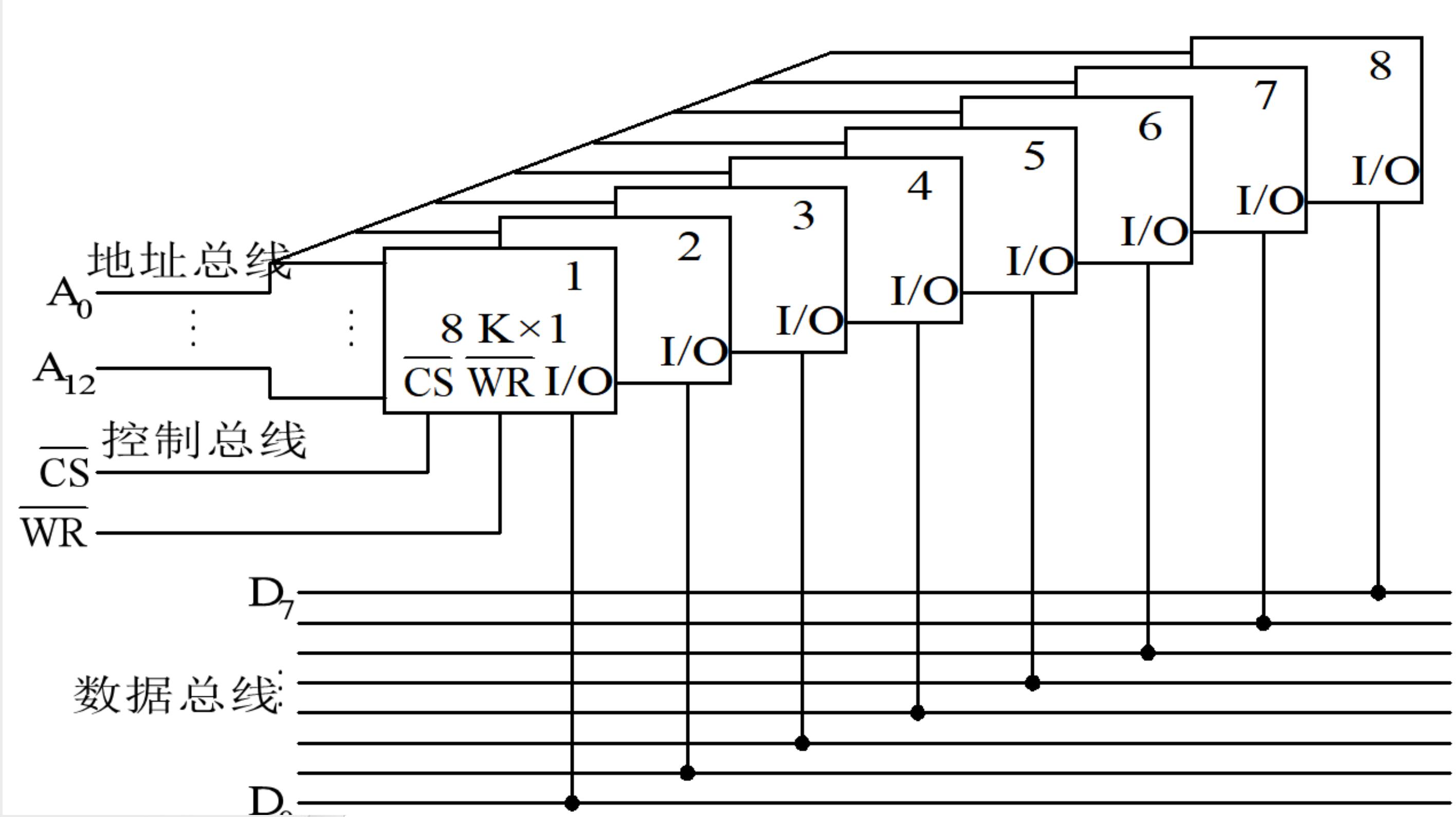
- 1、数据总线的位数与工作频率的乘积正比于数据传输率
- 2、地址总线的位数决定可寻址的最大内存空间
- 3、控制总线（读/写）指出总线周期的类型和本次输入输出完成时刻





2. 数据的表示与存储

2.8.3 主存扩展：位扩展

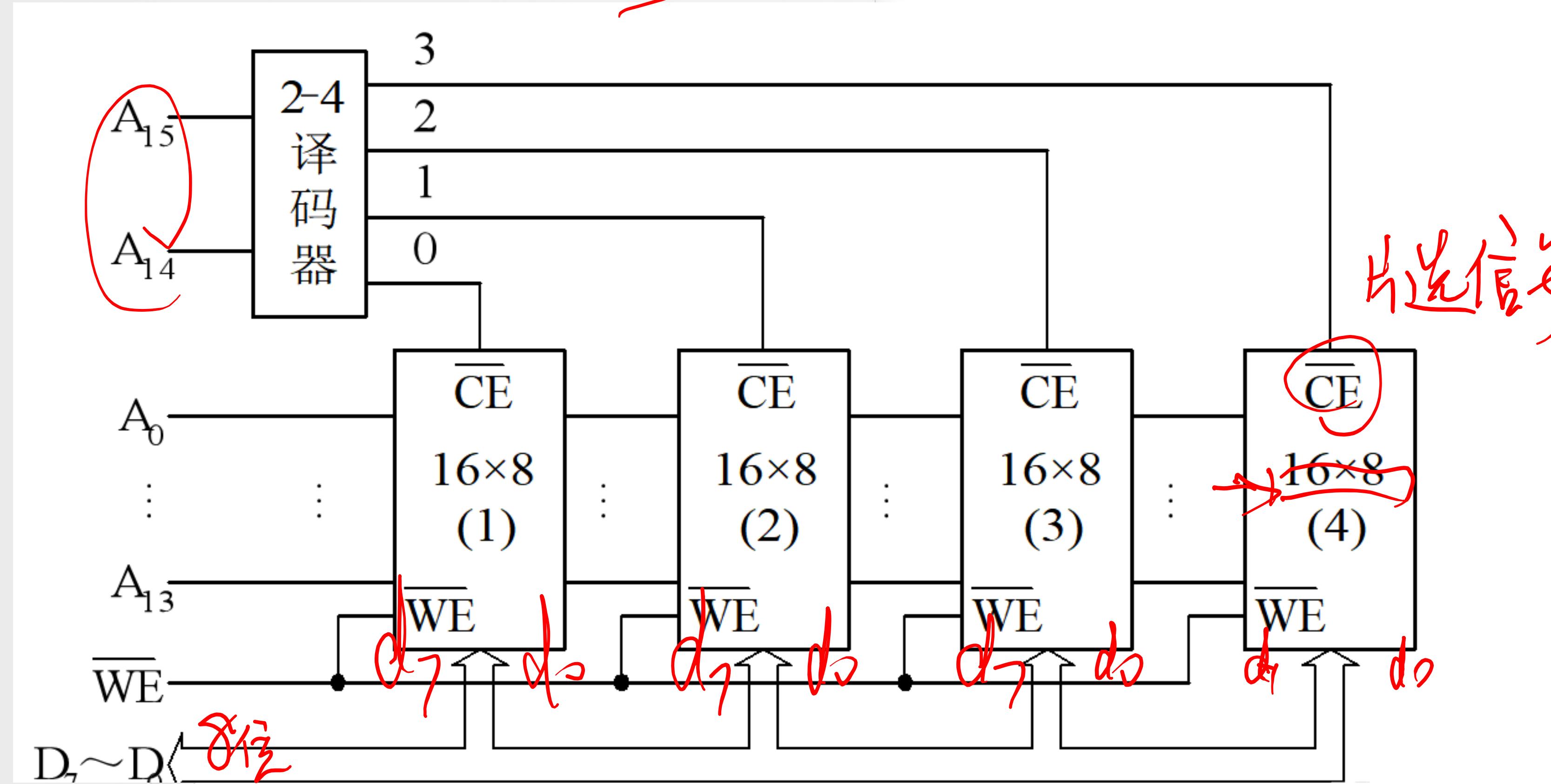


8片8Kx1位的RAM芯片组成8Kx8位。
各片地址线相同，数据线分开



2. 数据的表示与存储

2.8.3 主存扩展：字扩展

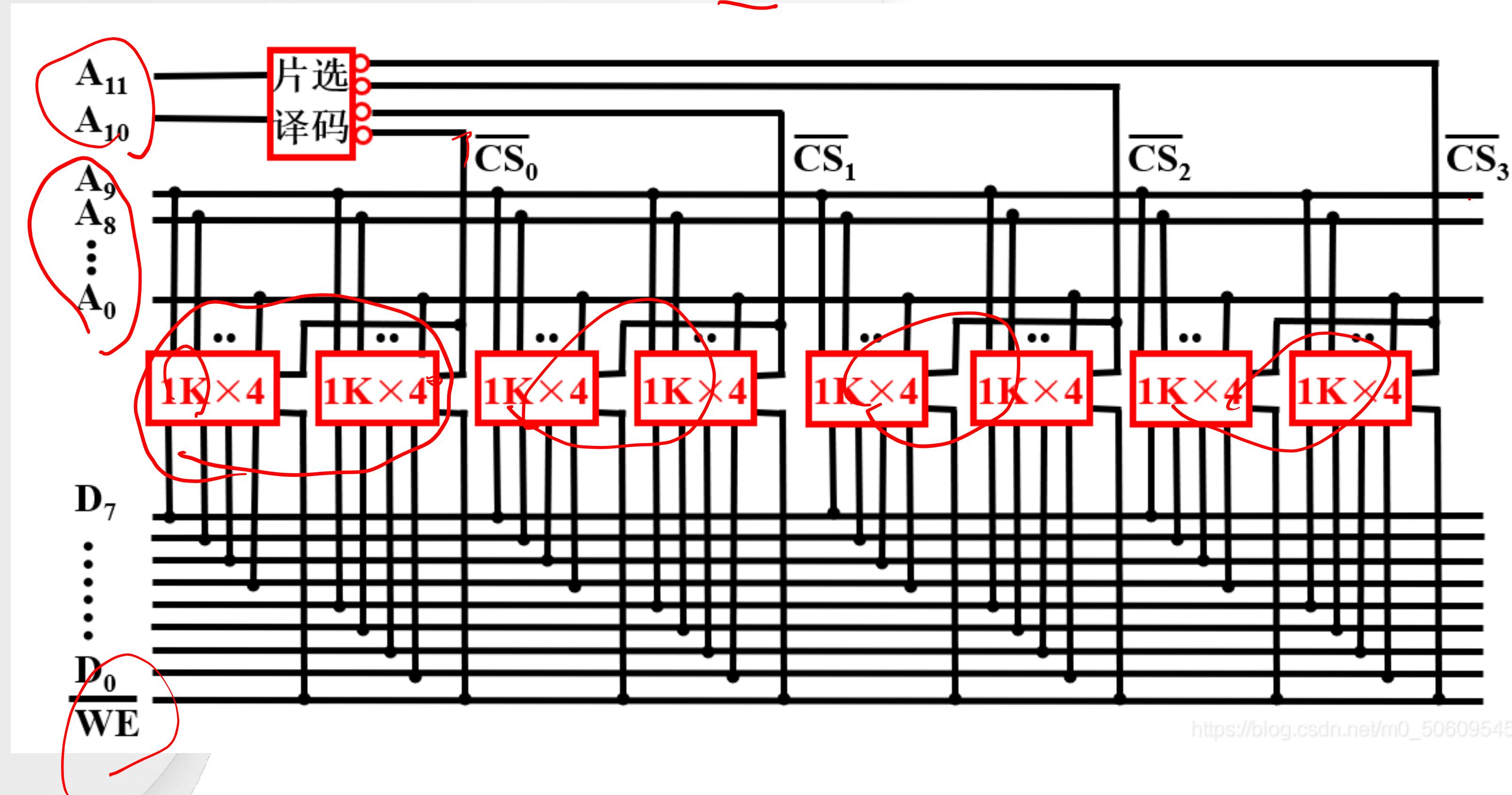


4片16Kx8位的RAM芯片组成64Kx8位。
各片数据线相同，地址线低位部分相同，
高位部分用来片选。



2. 数据的表示与存储

2.8.3 主存扩展：字位扩展



https://blog.csdn.net/m0_50609545

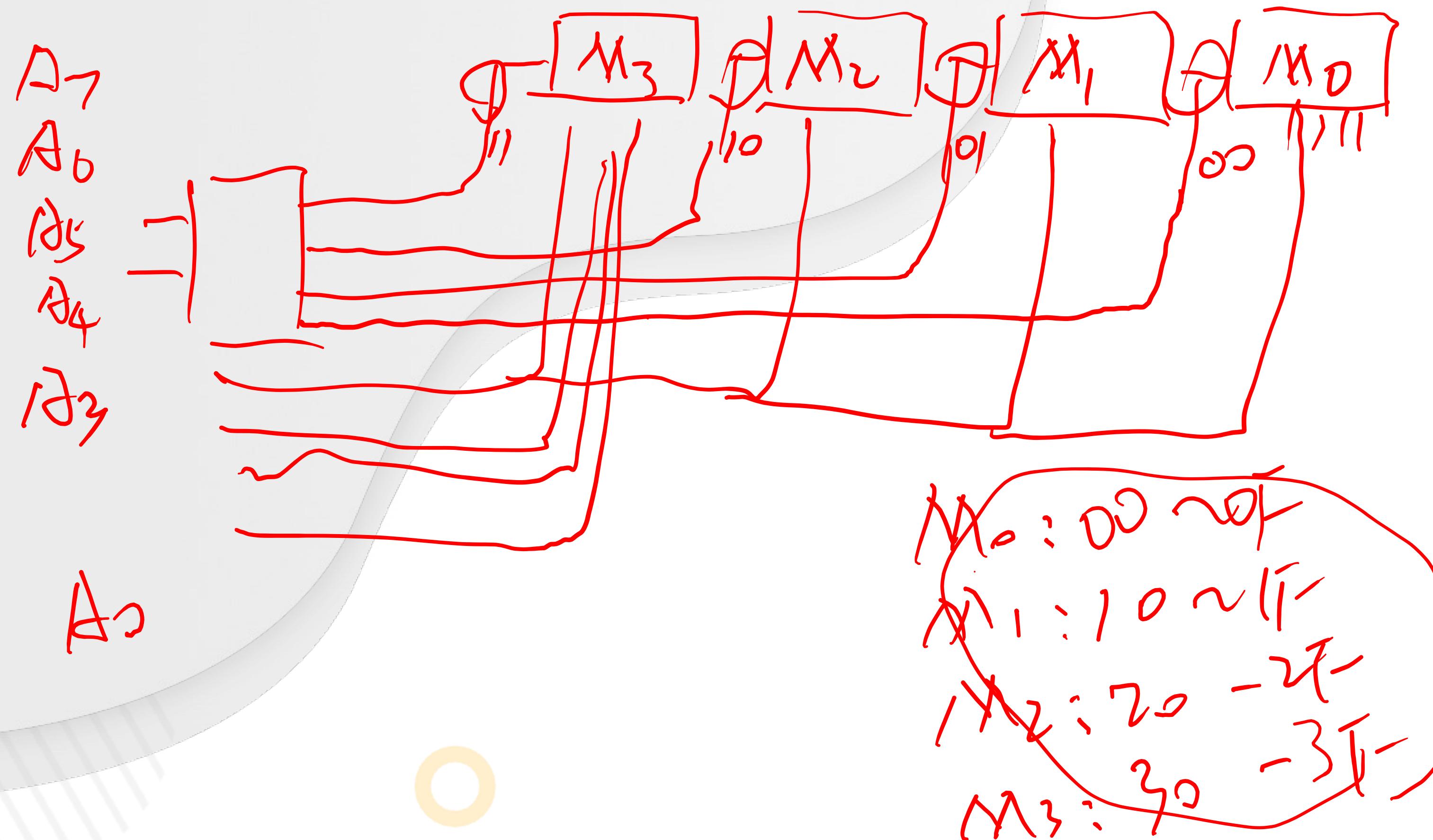
8片 $1K \times 4$ 位的RAM芯片组成 $4K \times 8$ 位。
两片数据线合并一组，地址线低位部分
相同，高位部分用来片选一组。



2. 数据的表示与存储

2.8.3 主存扩展：地址分配与片选

- 1、线选法：直接用地址线作为片选信号。浪费，空间不连续
- 2、译码片选法：扩展性强，地址连续



3 2 1 0
—— —— —— ——
1 0



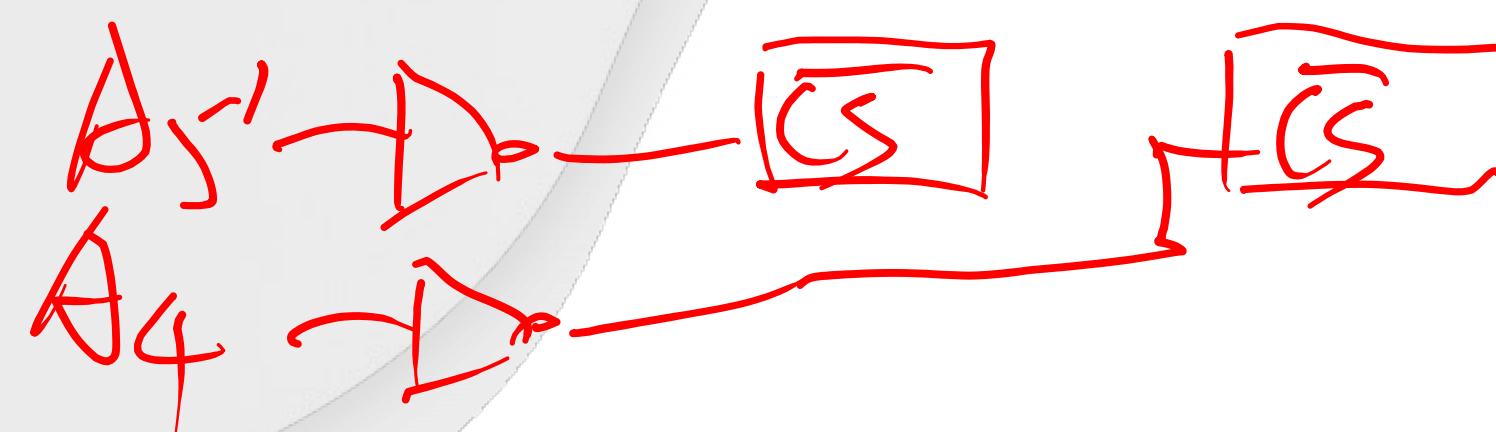
2. 数据的表示与存储

11

2.8.3 CPU和存储器的连接

host

- 1、合理选择存储芯片：RAM或ROM，数量合适，连线简单
- 2、地址线的连接：CPU送出的地址低位和芯片直连，高位通过译码片选，完成存储空间的扩展。
- 3、数据线的连接：根据CPU数据总线需求，选择位扩展，提升数据处理带宽。
- 4、读写命令线连接：CPU的读写信号直接连接存储芯片的RD和WE
- 5、片选线的选择：存储芯片的CS信号用来控制对应芯片是否有有效，主要配合地址扩展使用。有时CPU还有访存控制信号MREQ.





2. 数据的表示与存储

2.8.4 其他半导体存储器

- 1、掩膜型只读存储器MROM
- 2、可编程只读存储器PROM (使用熔丝)
- 3、可擦除可编程只读存储器EPROM (SAMOS管)
- 4、电可擦除可编程只读存储器EEPROM (FLOTOX管)
- 5、快擦除读写存储器 (Flash闪存，3和4的结合)

5.1、信息可读可写，读写都是以块为单位，块内如果有数据不能继续写，要先擦除后再写。所以写很慢。

512B
1024B
2048B x 16

5.2 断电后信息不丢失

5.3 存储单元由MOS管组成，也是半导体存储器

5.4 可随机访问存储单元，可作为外部分存储器

5.5 Flash根据结构不一样又分为Nor 和 Nand

SSD

1字



2. 数据的表示与存储

2.8 本节总结

1. 半导体结构的存储器大部分都是基于MOS管构建
2. SRAM基于6个MOS管组成的数据锁存器，用于寄存器、高速Cache
3. DRAM结构简单，可大规模封装，作为主存使用
4. SRAM一直保存着数据，DRAM需要刷新。 *2ms*
5. CPU连接存储器作为主存，可进行扩展
6. Flash存储器中的Nand应用普及（SSD硬盘）

欢迎参与学习

WELCOME FOR YOUR JOINING

船说：计算机基础