

计算机组织结构

7 内部存储器

刘博涵

2022年10月20日



南京大學
NANJING UNIVERSITY

教材对应章节



第7章 存储器分层体系结构



第5章 内部存储器

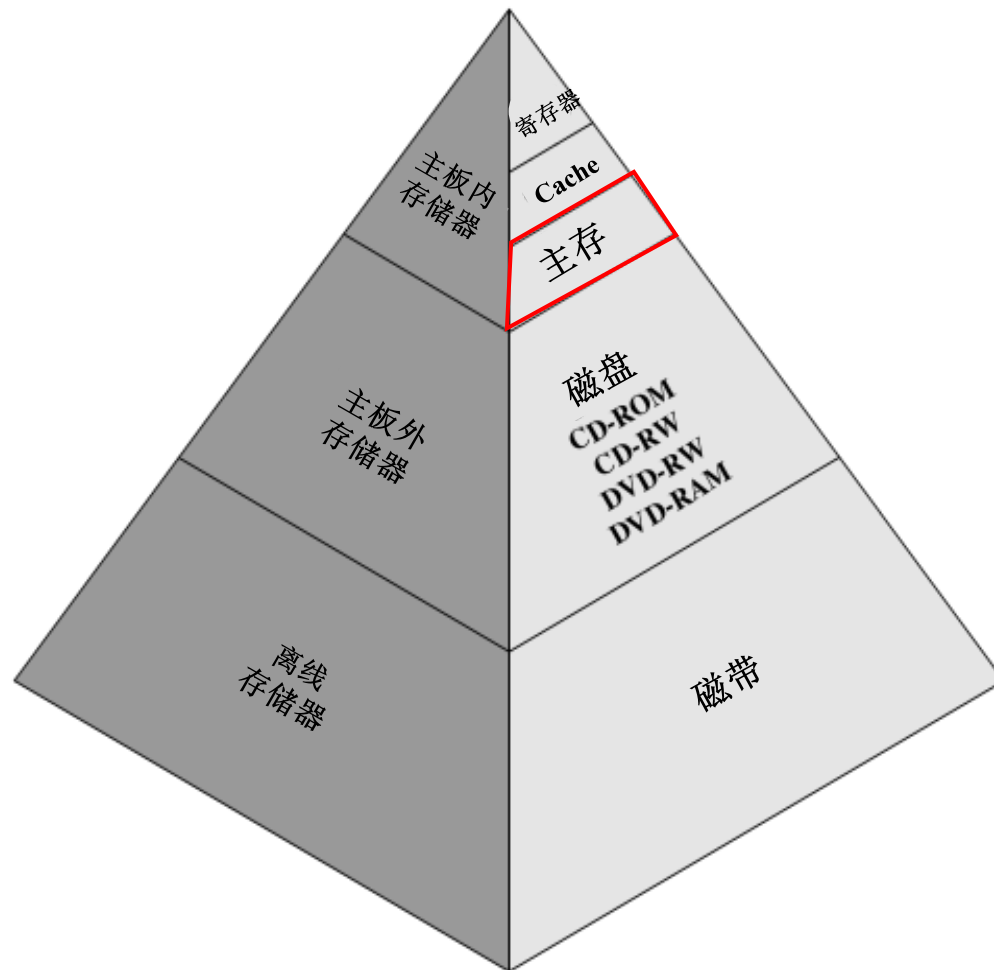
回顾：存储器

存储器 (Memory) 由一定数量的单元构成，每个单元可以被唯一标识，每个单元都有存储一个数值的能力

- **地址**：单元的唯一标识符（采用二进制）
- **地址空间**：可唯一标识的单元总数
- **寻址能力**：存储在每个单元中的信息的位数，即内存中能被单独识别并独立存放一个数据的最小内存空间
 - 大多数存储器是字节（8bit）寻址的，32位计算机的最大寻址空间为4GB

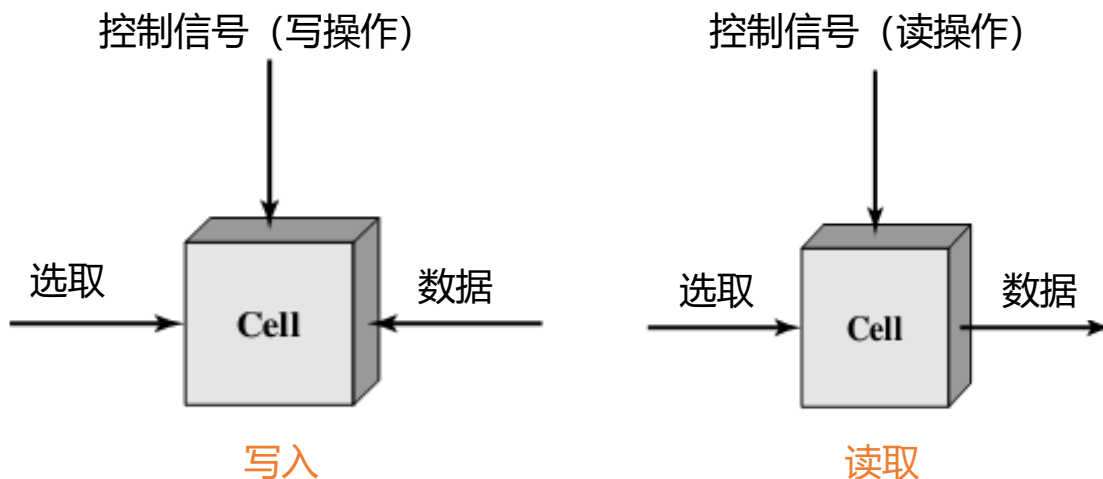


存储器层次结构



半导体存储器

- 用半导体芯片作主存储器是目前的主流做法
- **位元 (memory cell)**
 - 半导体存储器的基本元件，用于存储1位数据
 - 特性
 - 呈现两种稳态（或半稳态）：分别表示二进制的0和1
 - 它们能够至少被写入（write）数据一次：用来设置状态
 - 它们能够被读取（read）来获得状态信息
 - 操作



半导体存储器类型

存储器类型	种类	可擦除性	写机制	易失性
随机存取存储器（RAM）	读-写存储器	电可擦除，字节级	电	易失
只读存储器（ROM）	只读存储器	不可能	掩膜	非易失
可编程ROM（PROM）			电	
可擦除PROM（EPROM）	主要进行读操作的存储器	紫外线可擦除，芯片级		
电可擦除PROM（EEPROM）		电可擦除，字节级		
快闪存储器		电可擦除，块级		



随机存取存储器 (RAM)

- Random-Access Memory (RAM)

- 随机访问：对存储器中任意数据的访问所花费的时间与数据所在位置无关

- 特性

- 可以简单快速地进行读/写操作
 - 易失的 (Volatile)

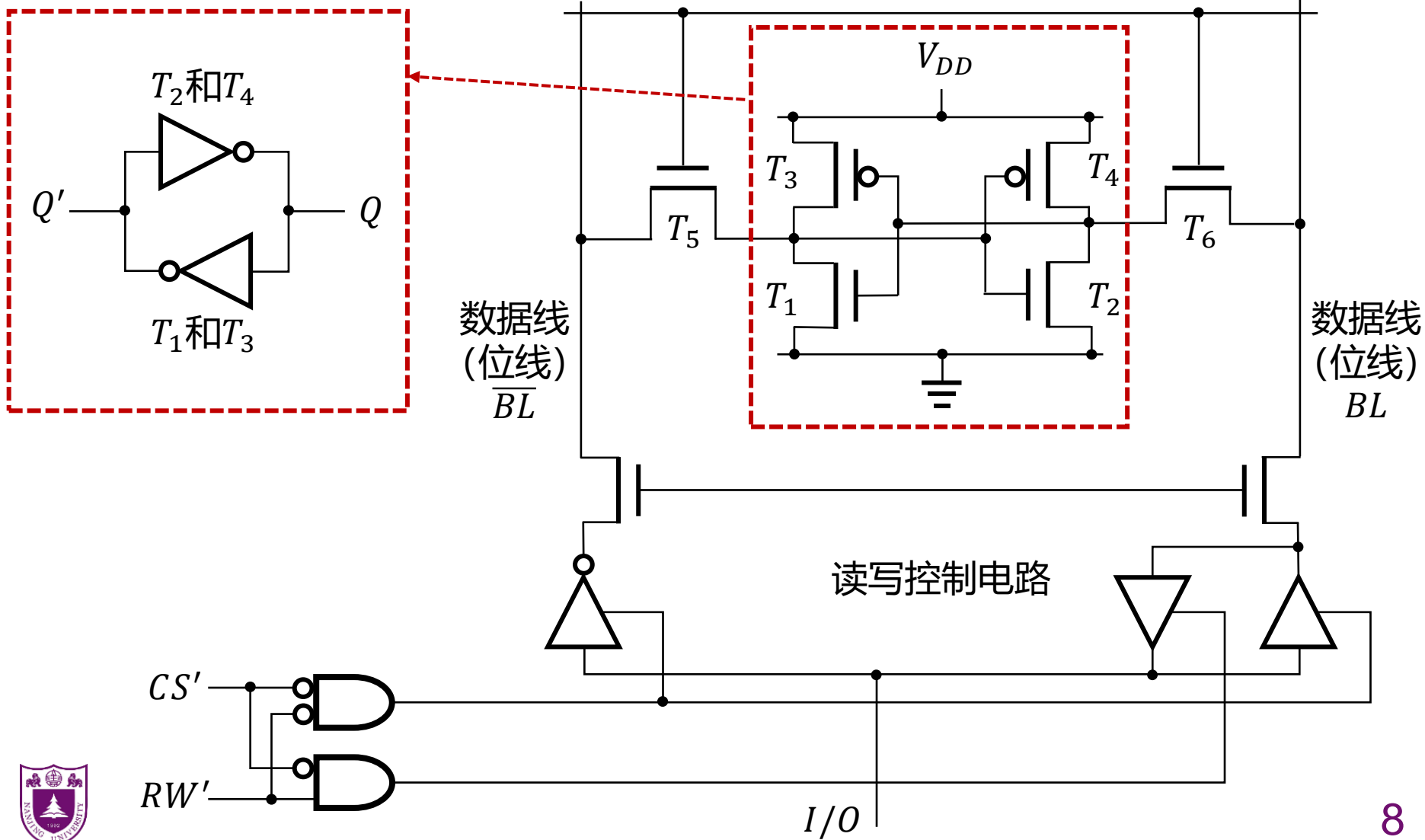
- 类型

- 静态RAM (SRAM) : Static RAM (1963年由Robert Norman在仙童发明)
 - 动态RAM (DRAM) : Dynamic RAM (1967年由Robert Dennard在IBM发明)

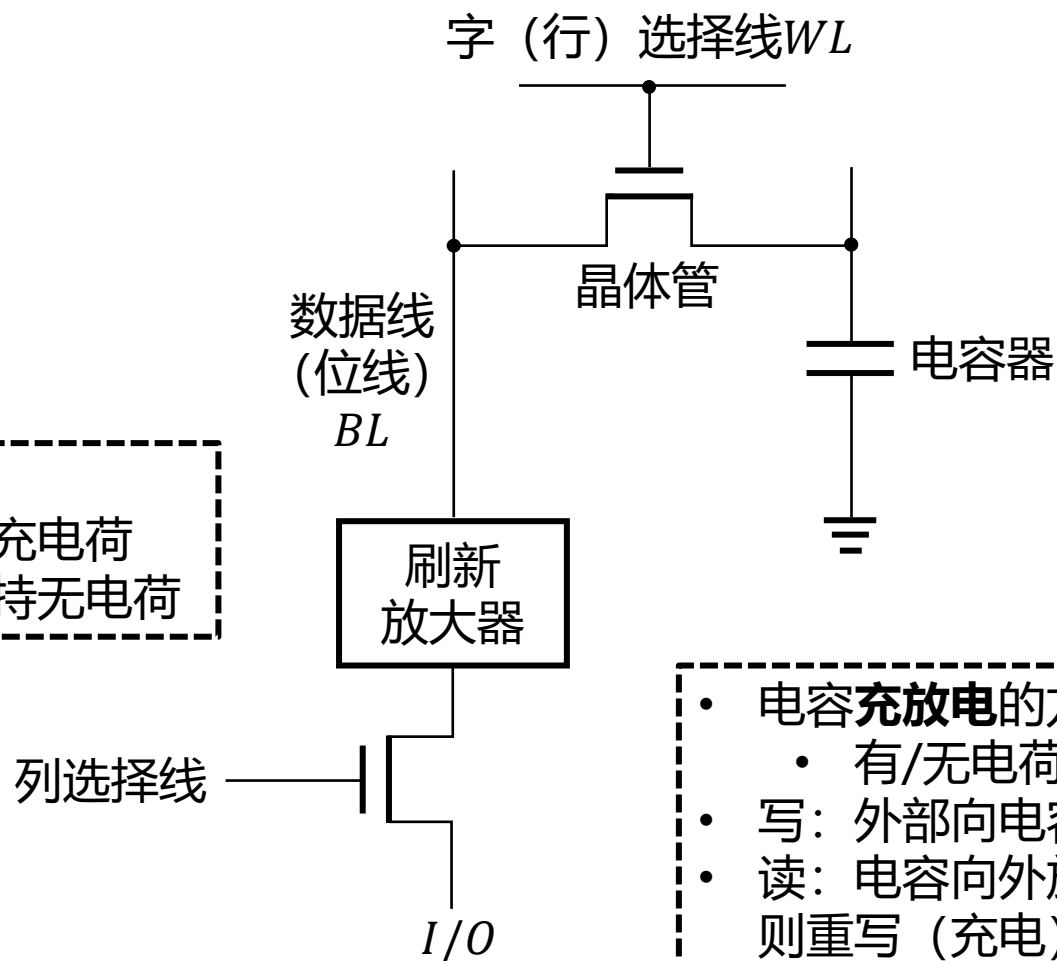


SRAM

加上输入构成RS触发器



DRAM



- **定期刷新**

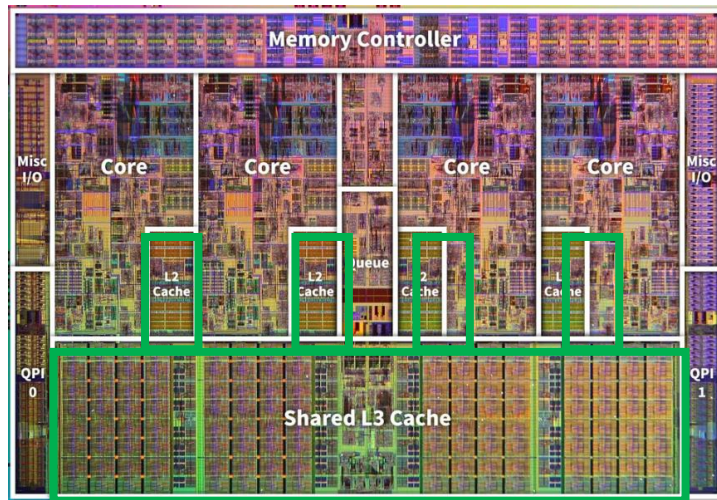
- 1则补充电荷
- 0则保持无电荷

- 电容**充放电**的方式存储数据
 - 有/无电荷: 1/0
- 写: 外部向电容充放电
- 读: 电容向外放电, 读后为1则重写 (充电)
- 电容存在**漏电效应**
 - 一段时间后电荷会流失, 导致数据丢失

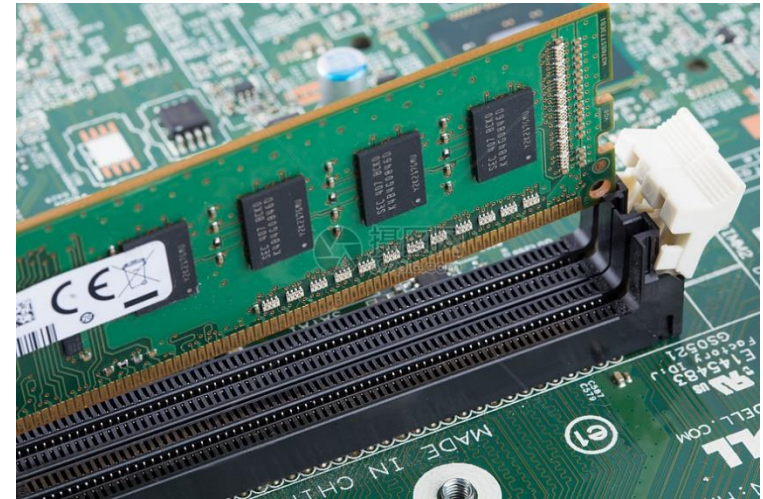


SRAM与DRAM的对比

	SRAM	DRAM
易失性	易失	易失
存储单元	双稳态触发器	电容
集成度	低	高
功耗	高	低
速度	快	慢
刷新	无	有
价格	\$1/MB	\$1/GB



CPU内部的SRAM (Cache L1~L3)



插在主板上的DRAM (内存条)

半导体存储器类型

存储器类型	种类	可擦除性	写机制	易失性
随机存取存储器 (RAM)	读-写存储器	电可擦除, 字节级	电	易失
只读存储器 (ROM)	只读存储器	不可能	掩膜	
可编程ROM (PROM)				
可擦除PROM (EPROM)	主要进行读操作的存储器	紫外线可擦除, 芯片级	电	非易失
电可擦除PROM (EEPROM)		电可擦除, 字节级		
快闪存储器		电可擦除, 块级		



只读存储器 (ROM)

- Read-only memory (ROM)
 - 一种可以长期保存信息的存储器，具有**断电后信息仍可继续保存**的特点，在正常工作时**只可读取数据，而不能写入数据**
- 特性
 - 非易失的：不要求供电来维持数据
 - 可读，但不能写入新数据
- 应用
 - 微程序设计，库子程序，系统程序，函数表
- 问题
 - 无出错处理机会：如果有一位出错，整批的ROM芯片只能报废
 - 用户无法写入数据：唯一的数据写入机会在出厂时完成



可编程ROM (PROM)

- **Programmable ROM (PROM)**

- 1956年由美籍华人电气工程科学家Wen Tsing Chow发明
- 为机载计算机而发明

- **特性**

- 非易失的
- 只能被写入一次
 - 写过程是用电信号执行
 - 需要特殊设备来完成写或“编程”过程

- **与ROM的对比**

- PROM提供了灵活性和方便性
- ROM在大批量生产领域仍具有吸引力



半导体存储器类型

存储器类型	种类	可擦除性	写机制	易失性
随机存取存储器 (RAM)	读-写存储器	电可擦除, 字节级	电	易失
只读存储器 (ROM)	只读存储器	不可能	掩膜	
可编程ROM (PROM)				
可擦除PROM (EPROM)	主要进行读操作的存储器	紫外线可擦除, 芯片级	电	非易失
电可擦除PROM (EEPROM)		电可擦除, 字节级		
快闪存储器		电可擦除, 块级		



主要进行读操作的存储器

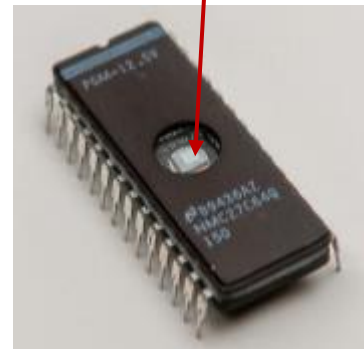
- Read-Mostly Memory
- 特性
 - 非易失的
 - 写操作与读操作相比，较为困难
- 应用
 - 读操作比写操作频繁得多的场景
- 类型
 - EPROM
 - EEPROM
 - Flash memory



光可擦除/可编程只读存储器 (EPROM)

- Erasable programmable read-only memory (EPROM)
 - 1971年由Intel的以色列电气工程科学家Dov Frohman发明
- 特性
 - 电写入 (1-→0)
 - 漏极和选择栅加高压 ($V_{PP}=12\sim24V$)，电子温度上升变为热电子，上升穿过氧化层，进入浮栅场效应管不导通，数据被写入0
 - 光擦除 (0-→1)
 - 擦除：在写操作前将封装芯片暴露在紫外线下
 - 照射使电子获得能量，逃出浮栅，场效应管导通
 - 所有的存储单元都变回相同的初始状态1
 - 每次擦除需要约20分钟（受玻璃影响）

紫外线照射



光可擦除/可编程只读存储器 (EPROM)



2022年Youtube视频下的一个回复:

"We **still use** EPROM chips in our **diecasting machines** due to the fact that **we need a system to run sequence without failure or alterations** they 're **locked inside UV proof** stainless steel flap inside a Allen and bradley serie 2 plc processing module the chip itself is windowless so can't be erased. the first programme was written in 1982 and is still running without interruption it cycled with 25 sequences per cycle for 40 years. @ **1000 cycles per day over 40 years** we estimate that' s **12,000,000 cycles without programme failure.**"

电可擦除/可编程只读存储器 (EEPROM)

- **Electrically erasable programmable read-only memory (EEPROM)**
 - 1978年由Intel的美籍希腊人，计算机科学家George Perlegos发明
 - 目前用于车载摄像头和传感器出厂设置、安全气囊的弹出记录以及需要长时间通电的工业自动化设备和服务器的数据记录系统等；在DDR SDRAM中也会使用EEPROM，用于保存内存的相关信息，计算机会在开机自检时会读取这些信息
- **特性**
 - 可以随时写入而不删除之前的内容；只更新寻址到的一个或多个字节
 - 电写入 (1- \rightarrow 0)
 - 源极和漏极接地，控制栅接高电平，电子进入浮栅
 - 写操作每字节需要大约5ms，而读一个字节大约0.5 μ s
 - 电擦除 (0- \rightarrow 1)
 - 源极和控制栅接地，漏极接高电平，浮栅中电子进入衬底
 - 每次擦除仅需4~6ms，可擦除百万次



快闪存储器

- **Flash Memory**

- 1984年日本电气工程科学家Fujio Masuoka发明了NOR Flash (Flash EEPROM), 可随机寻址, 即可用于内部存储。
- 1987年Fujio Masuoka又发明了NAND Flash (用于外部存储器)

- **特性**

- 电可擦除: 与EEPROM原理类似, 优于EPROM
- 可以在块级擦除, 不能在字节级擦除: 优于EPROM, 不如EEPROM
- 需要先擦除再写入

- **分类**

- **NOR Flash:** 数据线和地址线分开, 可以读任意字节
- **NAND Flash:** 数据线和地址线服用, 只能按页读取



各类ROM对比

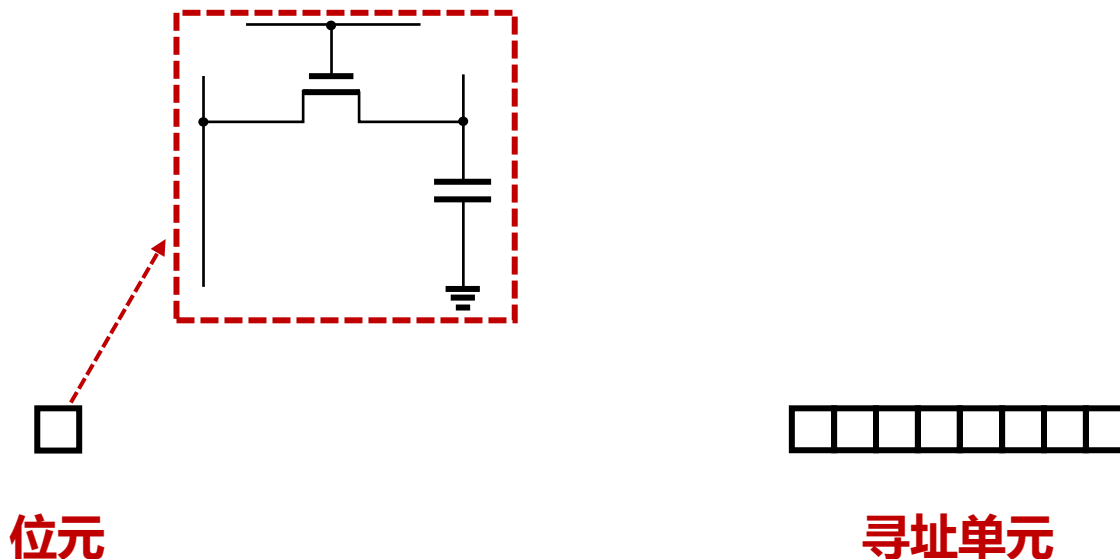
	ROM	PROM	EPROM	EEPROM	NOR Flash	NAND Flash
易失性	非易失	非易失	非易失	非易失	非易失	非易失
擦除方式	否	否	紫外线	电	电	电
擦除粒度	\	\	芯片级	字节级	块级	块级
擦除时间	\	\	20分钟	5ms	900ms	2ms
读取速度				5MB/s	100MB/s	20MB/s
写入速度	适合量产	慢 (1次)	慢	200B/s	0.5MB/s	5MB/s
寿命	几十年	几十年	几十年	>100万次	<10万次	<10万次
容量			MB级	KB级	MB级	GB级
价格	<PROM	<EPROM	<EEPROM	\$20/MB	\$0.01/MB	\$0.1/GB

技术在不断发展，仅供参考



从位元到主存：寻址单元

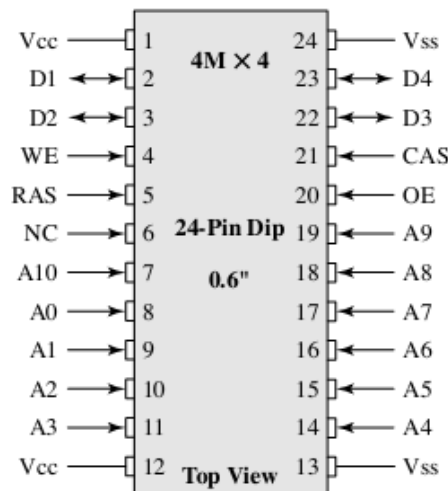
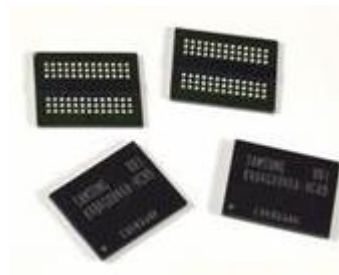
- 寻址单元 (Addressable unit) : 由若干相同地址的位元组成
- 寻址模式
 - 字节 (Byte) : 常用
 - 字 (Word)



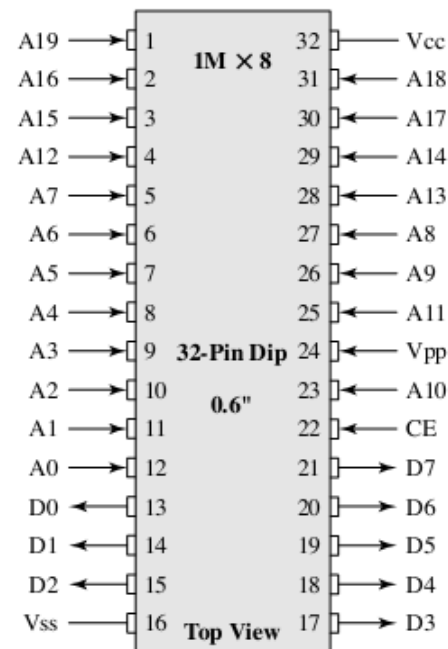
从位元到主存：芯片

• 芯片引脚

- Address: A0 – A19
- Data: D0 – D7
- Vcc: 电源
- Vss: 地线
- CE: 芯片允许引脚
- Vpp: 程序电压
- WE: 写允许
- OE: 读允许
- RAS: 行地址选通
- CAS: 列地址选通



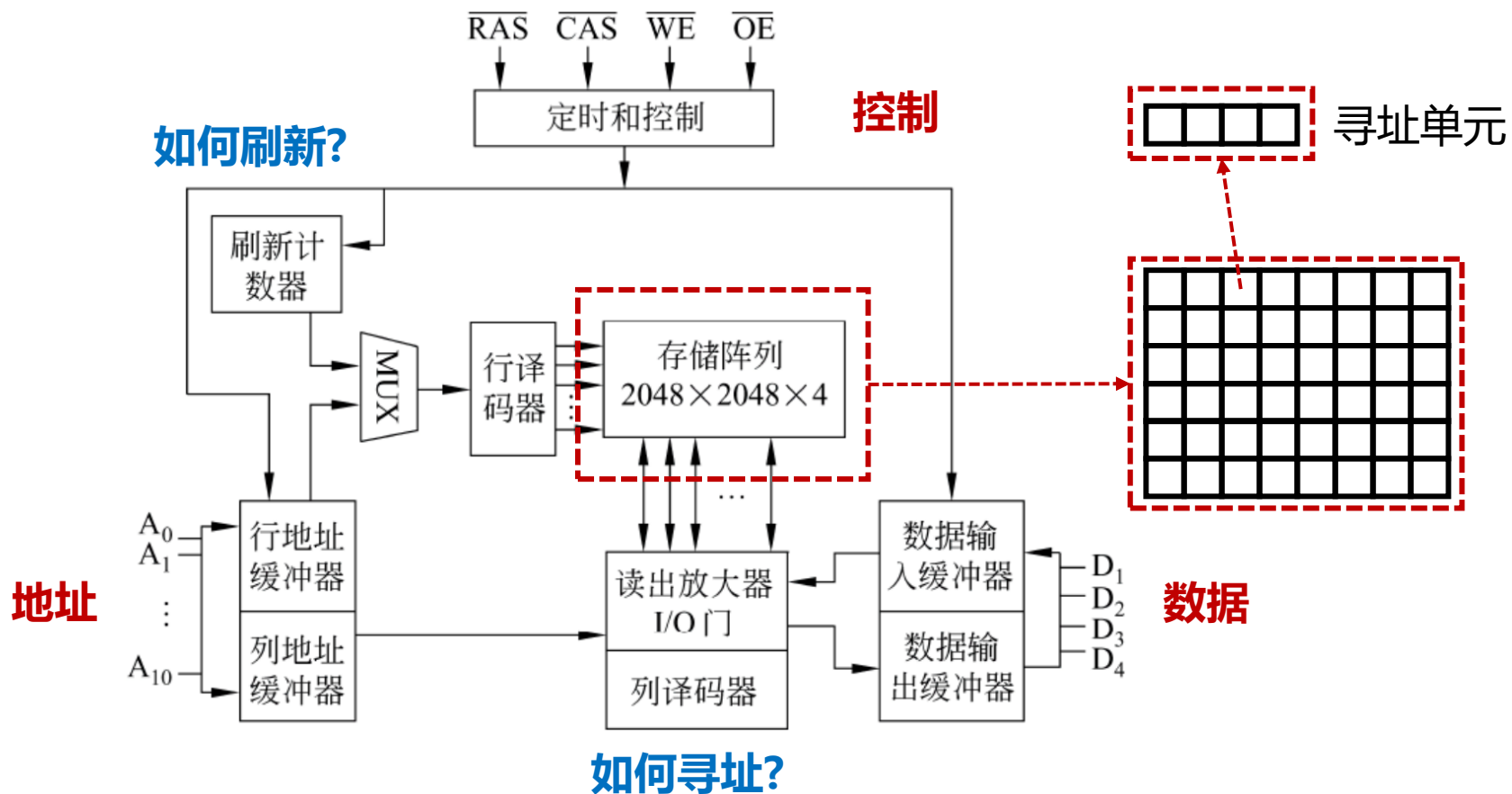
16-Mbit DRAM



8-Mbit EPROM

从位元到主存：存储阵列

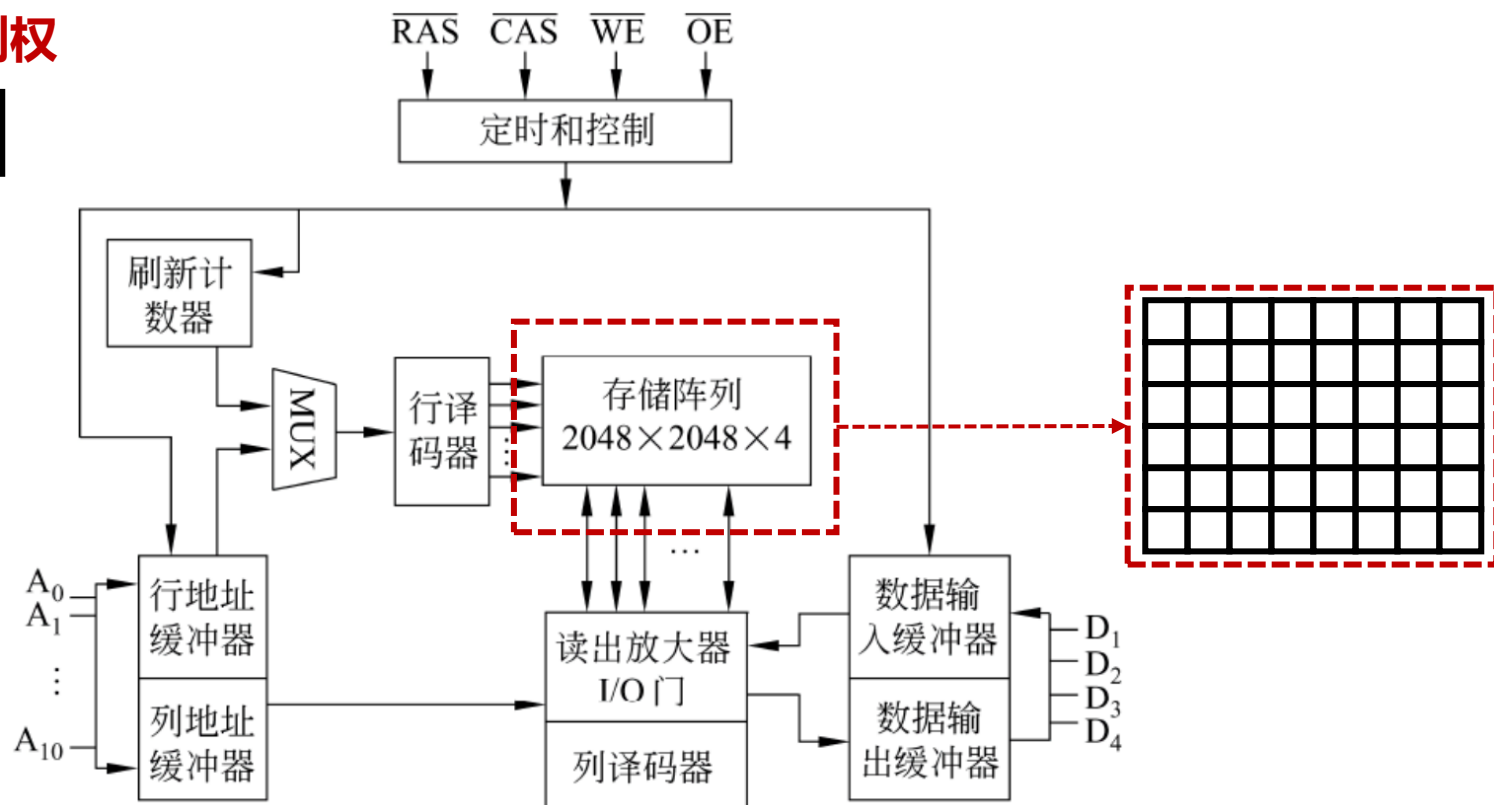
- 存储阵列 (Memory Array)：由大量寻址单元组成



如何寻址：获得总线控制权

1. 获得总线控制权

CPU

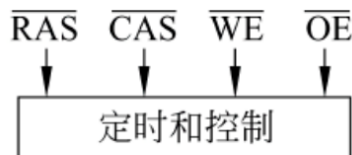


如何寻址：行访问

1. 获得总线控制权

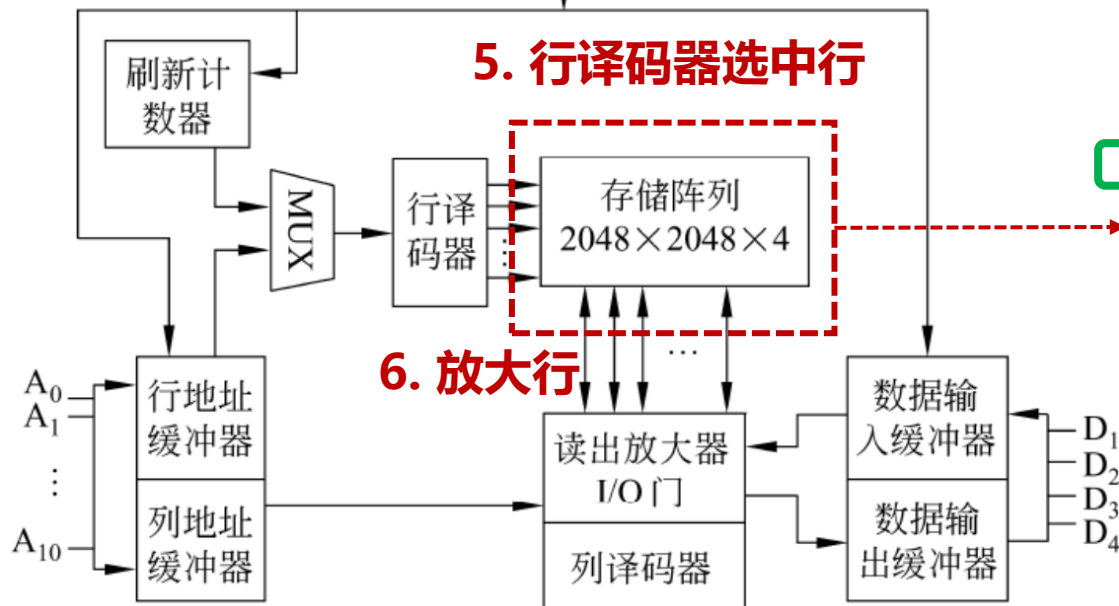


2. 发送地址



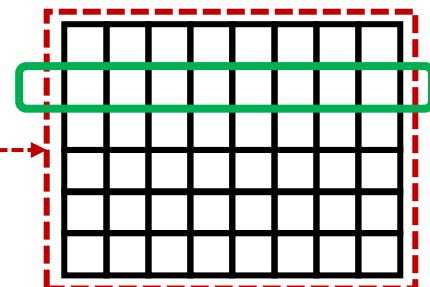
3. 分解为
行地址
和列地址

4. 发出行地址



5. 行译码器选中行

6. 放大行

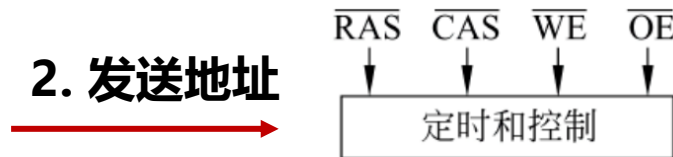


如何寻址：列访问

1. 获得总线控制权



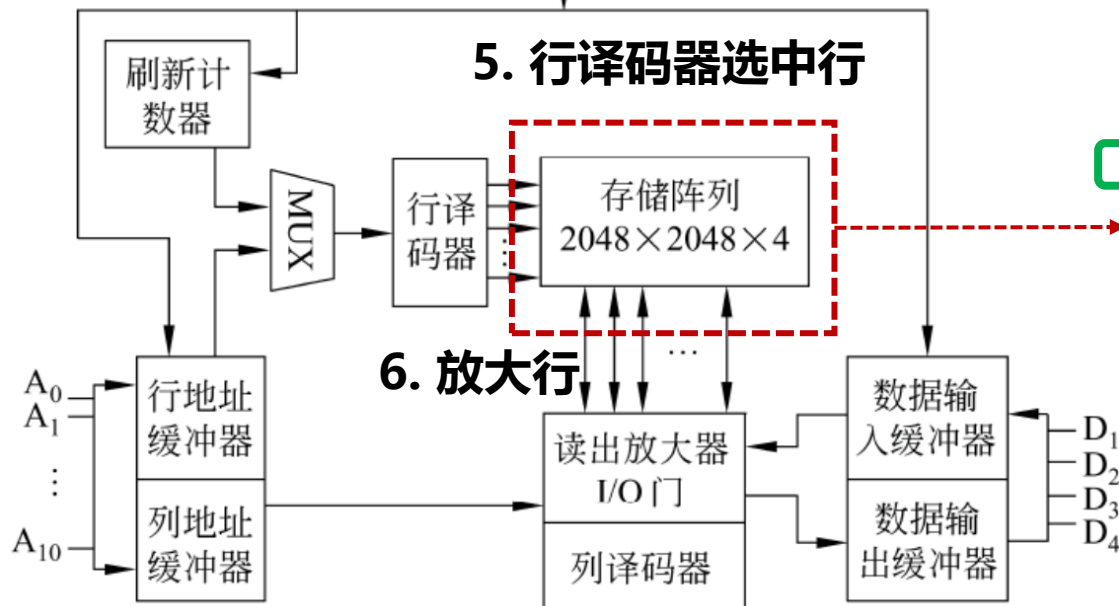
2. 发送地址



3. 分解为
行地址
和列地址

4. 发出行地址

7. 发出列地址



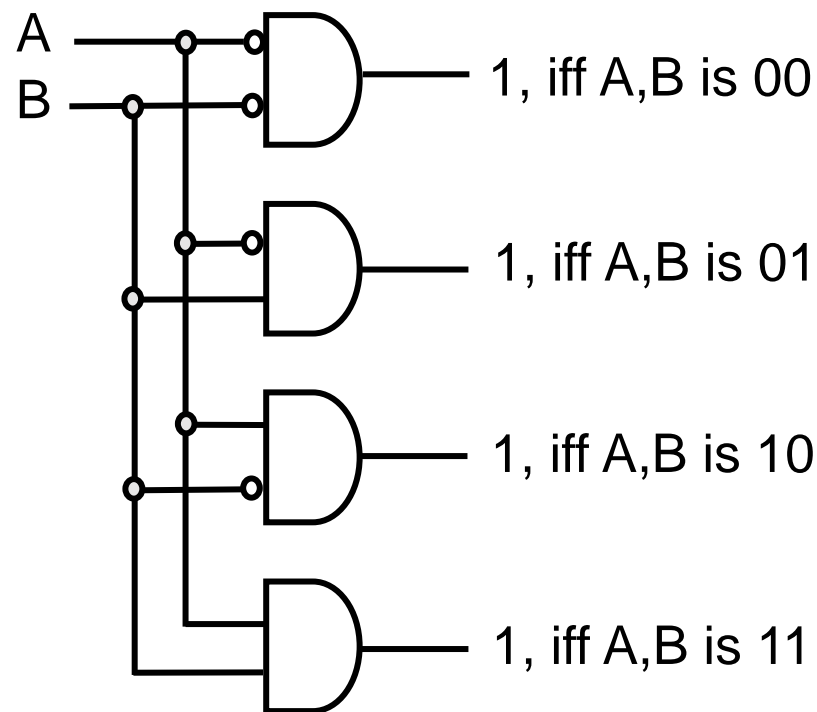
8. 选出对应列



如何寻址：地址译码器

- 地址译码器

- 一个 n 位译码器有 2^n 种输出.
- 当所有 n 个寻址位都满足条件时, 该输出为1
- 任何时候, 只有一个输出是1, 其他的都是0



2位译码器



如何刷新

约束：刷新会占用片选线、地址线、地址译码器

- **集中式刷新 (Centralized refresh)**

- 停止读写操作，并逐行刷新
- 刷新时无法操作内存（死区）

- **分散式刷新 (Decentralized refresh)**

- 在每个存储周期中，当读写操作完成时进行刷新
- 一次读写刷新一行，逐行刷新
- 会增加每个存储周期的时间

- **异步刷新 (Asynchronous refresh)**

- 每一行各自以固定间隔（小于最大刷新周期，毫秒级）刷新
- 将DRAM的刷新安排在CPU对指令的译码阶段，可有效避免死区
- 效率高：常用



高级的DRAM架构

- **问题：传统DRAM是异步的** **频率通常不超过66MHz**
 - 传统的DRAM芯片受到其内部架构和与处理器内存总线接口的限制
 - 处理器向内存提供地址和控制信号，表示内存中特定单元的一组数据应该被读出或写入DRAM
 - DRAM执行各种内部功能，如激活行和列地址线的高电容，读取数据，以及通过输出缓冲将数据输出，处理器只能等待这段延迟，即存取时间
 - 延时后，DRAM才写入或读取数据
- **解决：高级DRAM架构是同步的**
 - 同步DRAM (Synchronous DRAM, SDRAM) **频率通常不超过133MHz**
 - 双速率SDRAM (Double-Data-Rate SDRAM, DDR SDRAM / DDR)

DDR5频率可达4800MHz

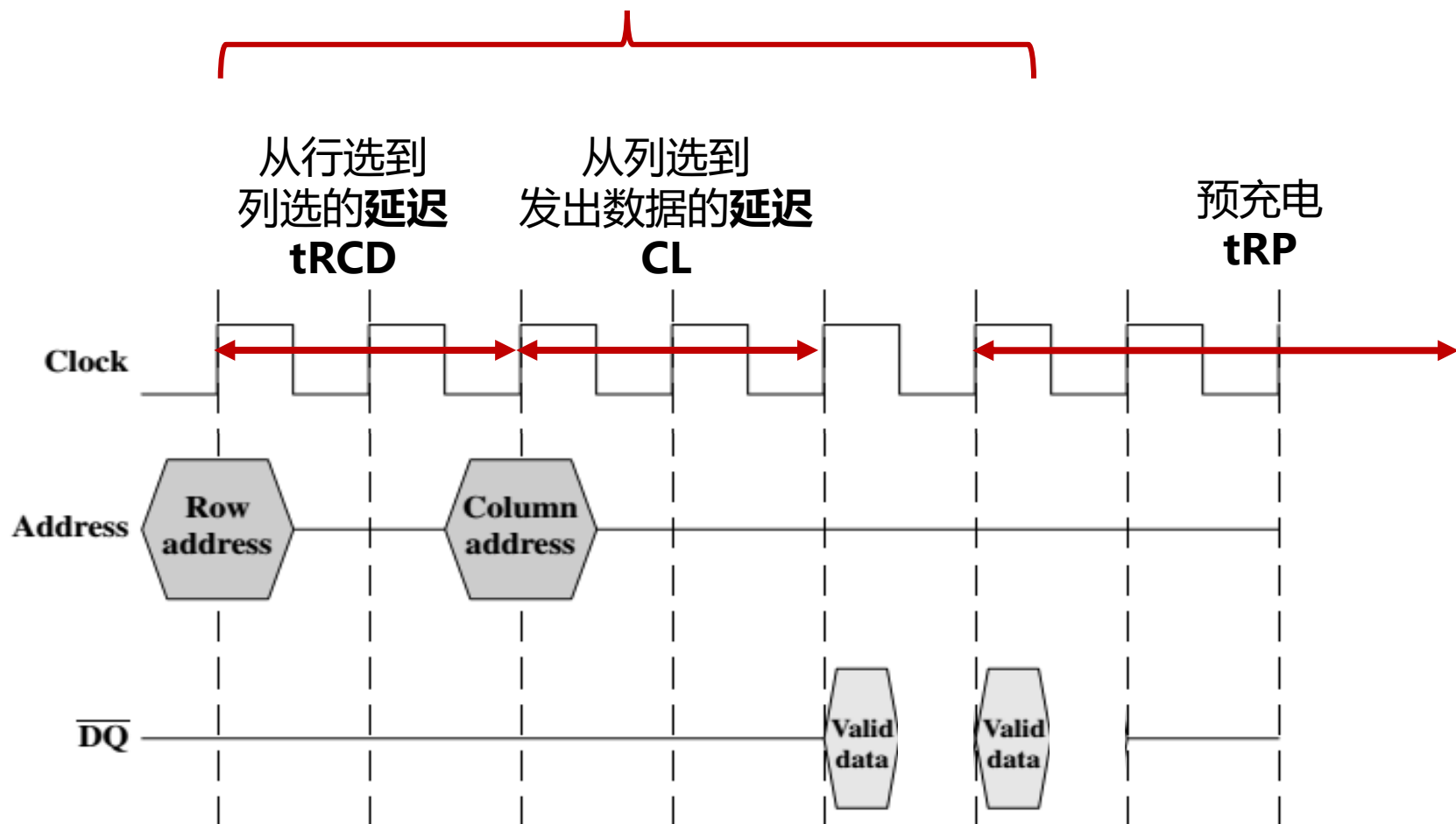


SDRAM

问题：电容充放电时间很难缩短

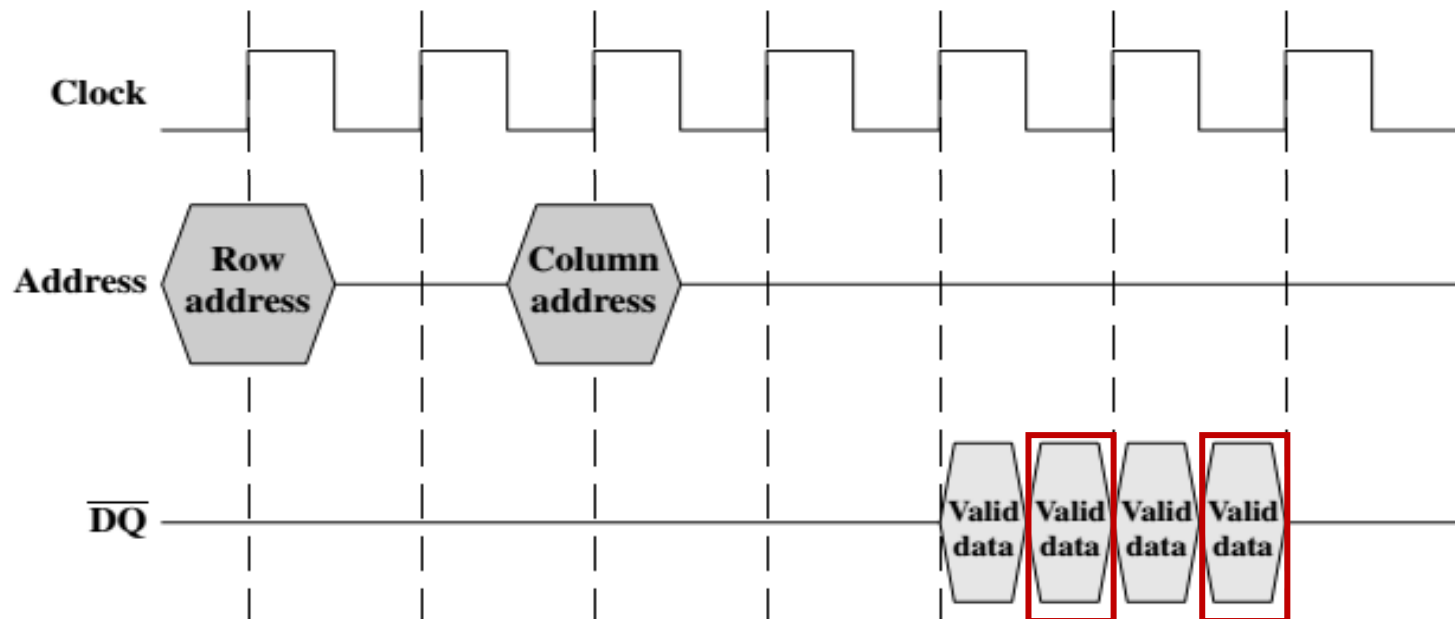
一次读数据操作

下一次



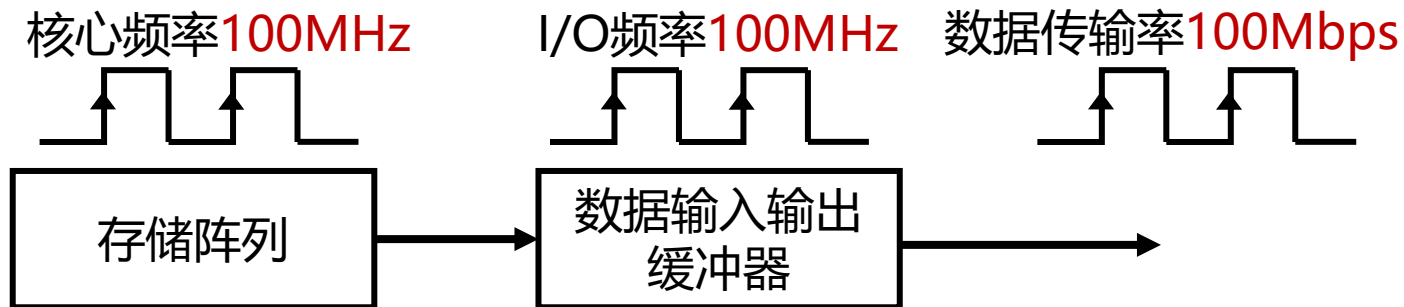
DDR SDRAM

- Double Data Rate: 每个时钟周期发送两次数据，一次在时钟脉冲的上升沿，一次在下降沿
- DDR → DDR2 → DDR3 → DDR4 → DDR5
 - 增加操作频率
 - 增加预取缓冲区

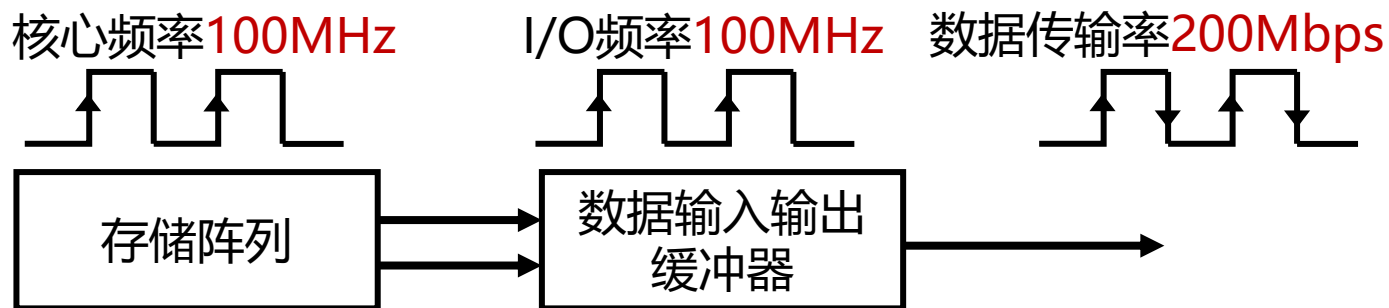


DDR SDRAM (续)

SDRAM



DDR1



DDR2



主要提升的是带宽



DDR SDRAM (续)

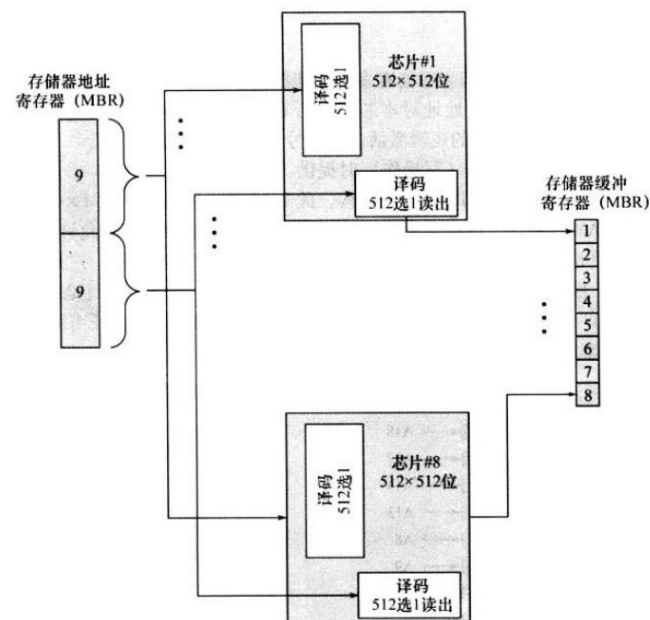
规格	核心频率	I/O频率	数据传输率	数据带宽
SDR-133	133MHz	133MHz	133MHz	1.06GB/s
DDR-266	133MHz	133MHz	266MHz	2.1GB/s
DDR-333	166MHz	166MHz	333MHz	2.7GB/s
DDR-400	200MHz	200MHz	400MHz	3.2GB/s
DDR2-533	133MHz	266MHz	533MHz	4.2GB/s
DDR2-667	166MHz	333MHz	667MHz	5.3GB/s
DDR2-800	200MHz	400MHz	800MHz	6.4GB/s
DDR3-1066	133MHz	533MHz	1066MHz	8.5GB/s
DDR3-1333	166MHz	667MHz	1333MHz	10.6GB/s
DDR3-1600	200MHz	800MHz	1600MHz	12.8GB/s
DDR4-2133	133MHz	1066MHz	2133MHz	17.06GB/s
DDR4-2666	166MHz	1333MHz	2666MHz	21.3GB/s
DDR4-3200	200MHz	1600MHz	3200MHz	25.6GB/s



从位元到主存：模块组织

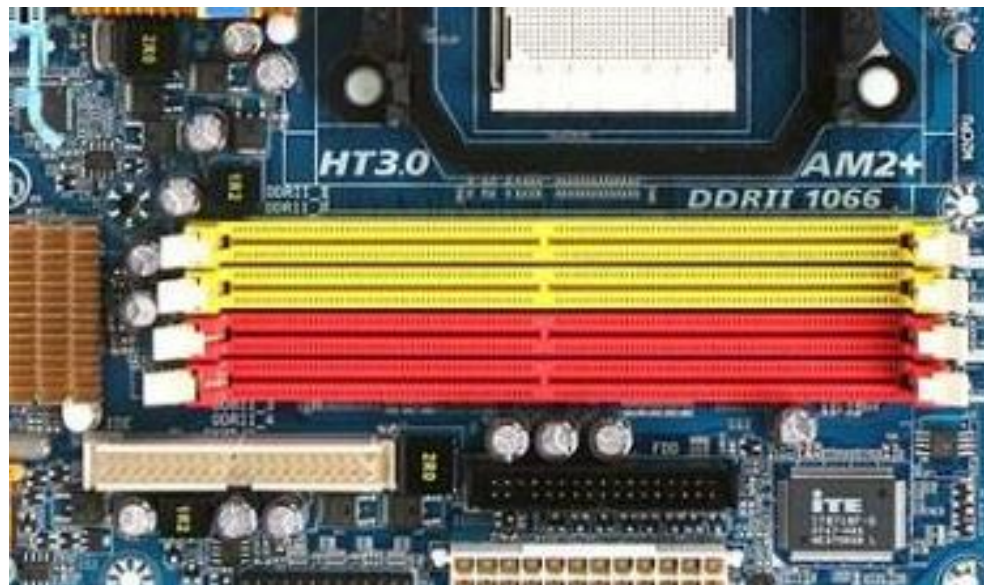
- 模块组织

- 位扩展：地址线不变，数据线增加
 - 使用 8 块 $4K \times 1$ bit 的芯片组成 $4K \times 8$ bit 的存储器
- 字扩展：地址线增加，数据线不变
 - 使用 4 个 $16K \times 8$ bit 的芯片组成 $64K \times 8$ bit 的存储器
- 字、位同时扩展：地址线增加，数据线增加
 - 使用 8 个 $16K \times 4$ bit 的芯片组成 $64K \times 8$ bit 的存储器



从位元到主存：主存

- 插槽
 - 组合多个存储模块



总结

- 半导体存储器
 - 读-写存储器：RAM
 - DRAM vs. SRAM
 - DRAM → SDRAM, DDR
 - 只读存储器：ROM, PROM
 - 主要进行读操作的存储器：EPROM, EEPROM, flash memory
- 从位元到主存
 - 位元 → 寻址单元 → 存储阵列 → 芯片 → 模块组织 → 主存



谢谢

bohanliu@nju.edu.cn



南京大學
NANJING UNIVERSITY