

第 3 章

CHAPTER

计算机组成原理习题

3.1 计算机系统概论

1. 什么是计算机系统、计算机硬件和计算机软件？硬件和软件哪个更重要？

【答】计算机系统指由计算机硬件系统和软件系统组成的综合体。

计算机硬件：指计算机中的电子线路和物理装置。

计算机软件：计算机运行所需的程序及相关资料。

硬件和软件在计算机系统中相互依存，缺一不可，因此同样重要。

2. 如何理解计算机的层次结构？

【答】计算机硬件、系统软件和应用软件构成了计算机系统的 3 个层次结构。

(1) 硬件系统是最内层的，它是整个计算机系统的基础和核心。

(2) 系统软件在硬件之外，为用户提供一个基本操作界面。

(3) 应用软件在最外层，为用户提供解决具体问题的应用系统界面。

通常将硬件系统之外的其余层称为虚拟机。各层次之间关系密切，上层是下层的扩展，下层是上层的基础，各层次的划分不是绝对的。

3. 说明高级语言、汇编语言和机器语言的差别及其联系。

【答】机器语言是计算机硬件能够直接识别的语言，汇编语言是机器语言的符号表示，高级语言是面向算法的语言。高级语言编写的程序（源程序）处于最高层，必须翻译成汇编语言，再由汇编程序汇编成机器语言（目标程序）之后才能被执行。

4. 如何理解计算机组成和计算机体系结构？

【答】计算机体系结构是指那些能够被程序员所见到的计算机系统的属性，如指令系统、数据类型、寻址技术组成及 I/O 机理等。计算机组成是指如何实现计算机体系结构所体现的属性，包含对程序员透明的硬件细节，如组成计算机系统的各个功能部件的结构和功能及相互连接方法等。

5. 冯·诺依曼计算机的特点是什么？

【答】冯·诺依曼计算机的特点是：

- (1) 计算机由运算器、控制器、存储器、输入设备、输出设备五大部件组成。
- (2) 指令和数据以同等地位存放于存储器内,并可以按地址访问。
- (3) 指令和数据均用二进制表示。
- (4) 指令由操作码、地址码两大部分组成,操作码用来表示操作的性质,地址码用来表示操作数在存储器中的位置。
- (5) 指令在存储器中顺序存放,通常自动顺序取出执行。
- (6) 机器以运算器为中心(原始冯·诺依曼机)。

6. 画出计算机硬件组成框图,说明各部件的作用及计算机系统的主要技术指标。

【答】计算机硬件组成框图如图 3.1 所示。

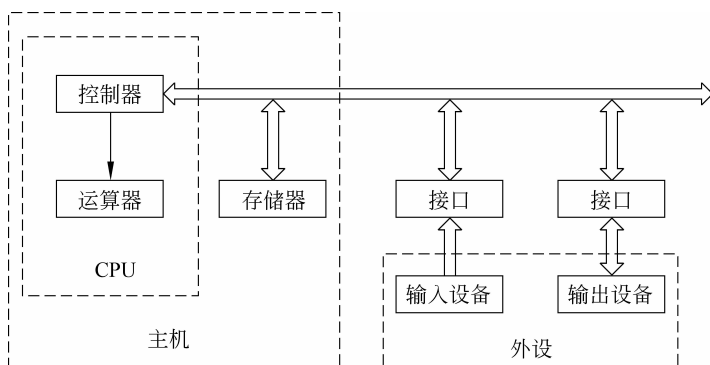


图 3.1 计算机硬件组成框图

各部件的作用如下：

控制器：整机的指挥中心,它使计算机的各个部件自动协调工作。

运算器：对数据信息进行处理的部分,用来进行算术运算和逻辑运算。

存储器：存放程序和数据,是计算机实现“存储程序控制”的基础。

输入设备：将人们熟悉的信息形式转换成计算机可以接受并识别的信息形式的设备。

输出设备：将计算机处理的结果(二进制信息)转换成人类或其他设备可以接收和识别的信息形式的设备。

计算机系统的主要技术指标有：

机器字长：指 CPU 一次能处理的数据的位数。通常与 CPU 的寄存器的位数有关,字长越长,数的表示范围越大,精度也越高。机器字长也会影响计算机的运算速度。

数据通路宽度：数据总线一次能并行传送的数据位数。

存储容量：指能存储信息的最大容量,通常以字节来衡量。一般包含主存容量和辅存容量。

运算速度：通常用 MIPS(每秒百万条指令)、MFLOPS(每秒百万次浮点运算)或 CPI(执行一条指令所需的时钟周期数)来衡量。CPU 执行时间是指 CPU 对特定程序的执行

时间。

主频：机器内部主时钟的运行频率，是衡量机器速度的重要参数。

吞吐量：指流入、处理和流出系统的信息速率。它主要取决于主存的存取周期。

响应时间：计算机系统对特定事件的响应时间，如实时响应外部中断的时间等。

7. 解释下列概念：

主机、CPU、主存、存储单元、存储元件、存储基元、存储元、存储字、存储字长、存储容量、机器字长、指令字长。

【答】主机：是计算机硬件的主体部分，由 CPU 和主存储器 MM 合成为主机。

CPU：中央处理器，是计算机硬件的核心部件，由运算器和控制器组成（早期的运算器和控制器不在同一芯片上，现在的 CPU 内除含有运算器和控制器外还集成了 Cache）。

主存：计算机中存放正在运行的程序和数据存储器，为计算机的主要工作存储器，可随机存取；由存储体、各种逻辑部件及控制电路组成。

存储单元：可存放一个机器字并具有特定存储地址的存储单位。

存储元件：存储一位二进制信息的物理元件，是存储器中最小的存储单位，又叫存储基元或存储元，不能单独存取。

存储字：一个存储单元所存二进制代码的逻辑单位。

存储字长：一个存储单元所存储的二进制代码的总位数。

存储容量：存储器中可存二进制代码的总量（通常主、辅存容量分开描述）。

机器字长：指 CPU 一次能处理的二进制数据的位数，通常与 CPU 的寄存器位数有关。

指令字长：机器指令中二进制代码的总位数。

8. 解释下列英文缩写的中文含义：

CPU、PC、IR、CU、ALU、ACC、MQ、X、MAR、MDR、I/O、MIPS、CPI、FLOPS、MFLOPS、GFLOPS、TFLOPS、PFLOPS、EFLOPS。

【答】CPU：Central Processing Unit，中央处理机（器），是计算机硬件的核心部件，主要由运算器和控制器组成。

PC：Program Counter，程序计数器，其功能是存放当前欲执行指令的地址，并可自动计数形成下一条指令地址。

IR：Instruction Register，指令寄存器，其功能是存放当前正在执行的指令。

CU：Control Unit，控制单元，为控制器的核心部件，功能是产生微操作命令序列。

ALU：Arithmetic Logic Unit，算术逻辑运算单元，为运算器的核心部件，其功能是进行算术、逻辑运算。

ACC：Accumulator，累加器，是运算器中既能存放运算前的操作数，又能存放运算结果的寄存器。

MQ：Multiplier-Quotient Register，乘商寄存器，乘法运算时存放乘数，除法运算时存放商的寄存器。

X：此字母没有专指的缩写含义，可以用作任一部件名，在此表示操作数寄存器，即运算器中工作寄存器之一，用来存放操作数。

MAR: Memory Address Register, 存储器地址寄存器, 在主存中用来存放欲访问的存储单元的地址。

MDR: Memory Data Register, 存储器数据缓冲寄存器, 在主存中用来存放从某单元读出或要写入某存储单元的数据。

I/O: Input/Output equipment, 输入/输出设备, 为输入设备和输出设备的总称, 用于计算机内部和外界信息的转换与传送。

MIPS: Million Instruction Per Second, 每秒执行百万条指令数, 为计算机运算速度指标的一种计量单位。

CPI: Cycle Per Instruction, 每条指令执行需要的时钟周期数。

FLOPS: Floating-point Operations Per Second, 每秒所执行的浮点运算次数。

MFLOPS: Million FLOPS; GFLOPS: Giga FLOPS; TFLOPS: Tera FLOPS; PFLOPS: Peta FLOPS; EFLOPS: Exa FLOPS。

9. 画出主机框图, 分别以存数指令“STA M”和加法指令“ADD M”(M 均为主存地址)为例, 按序标出完成该指令(包括取指令阶段)的信息流程。假设主存容量为 $256\text{M} \times 32$ 位, 在指令字长、存储字长、机器字长相等的条件下, 指出图 3.2 中各寄存器的位数。

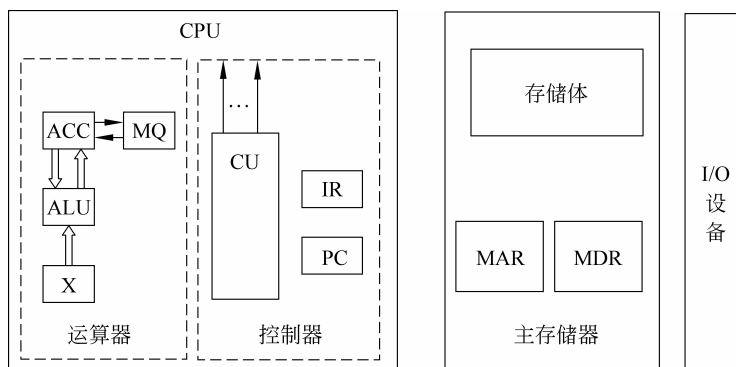


图 3.2 细化的计算机组成框图

【答】(1) STA M 指令: $\text{PC} \rightarrow \text{MAR}, \text{MAR} \rightarrow \text{MM}, \text{MM} \rightarrow \text{MDR}, \text{MDR} \rightarrow \text{IR}, \text{OP}(\text{IR}) \rightarrow \text{CU}, \text{AD}(\text{IR}) \rightarrow \text{MAR}, \text{ACC} \rightarrow \text{MDR}, \text{MAR} \rightarrow \text{MM}, \text{WR}$ 。

(2) ADD M 指令: $\text{PC} \rightarrow \text{MAR}, \text{MAR} \rightarrow \text{MM}, \text{MM} \rightarrow \text{MDR}, \text{MDR} \rightarrow \text{IR}, \text{OP}(\text{IR}) \rightarrow \text{CU}, \text{AD}(\text{IR}) \rightarrow \text{MAR}, \text{RD}, \text{MM} \rightarrow \text{MDR}, \text{MDR} \rightarrow \text{X}, \text{ADD}, \text{ALU} \rightarrow \text{ACC}, \text{ACC} \rightarrow \text{MDR}, \text{WR}$ 。

假设主存容量为 $256\text{M} \times 32$ 位, 在指令字长、存储字长、机器字长相等的条件下, ACC、X、IR、MDR 寄存器均为 32 位, PC 和 MAR 寄存器均为 28 位。

10. 指令和数据都存于存储器中, 计算机如何区分它们?

【答】计算机区分指令和数据有以下两种方法:

(1) 通过不同的时间段来区分指令和数据, 即在取指令阶段(或取微程序)取出的是指令, 在执行指令阶段(或相应微程序)取出的即为数据。

(2) 通过地址来源区分, 由 PC 提供存储单元地址的取出的是指令, 由指令地址码部分提供存储单元地址取出的是操作数。

3.2 计算机的发展及应用

1. 通常计算机的更新换代以什么为依据?

【答】主要以组成计算机基本电路的元器件为依据,如电子管、晶体管、集成电路等。

2. 举例说明专用计算机和通用计算机的区别。

【答】按照计算机的效率、速度、价格和运行的经济性和实用性,可以将计算机划分为通用计算机和专用计算机。通用计算机适应性强,但牺牲了效率、速度和经济性,而专用计算机是最有效、最经济和最快的计算机,但适应性很差,如个人计算机和计算器。

3. 摩尔定律是否永远生效?

【答】否。

3.3 系统总线

1. 什么是总线? 总线传输有何特点? 为了减轻总线负载,总线上的部件应具备什么特点?

【答】总线是一种能由多个部件分时共享的公共信息传送线路。

总线传输的特点是:某一时刻只允许有一个部件向总线发送信息,但多个部件可以同时从总线上接收相同的信息。

为了减轻总线负载,总线上的部件应通过三态驱动缓冲电路与总线连通。

2. 总线如何分类? 什么是系统总线? 系统总线又分为几类? 它们各有何作用? 是单向的还是双向的? 它们与机器字长、存储字长、存储单元有何关系?

【答】按照连接部件的不同,总线可以分为片内总线、系统总线和通信总线。

系统总线是连接 CPU、主存、I/O 各部件之间的信息传输线。

系统总线按照传输信息不同又分为地址线、数据线和控制线。地址线是单向的,其根数越多,寻址空间越大,即 CPU 能访问的存储单元的个数越多;数据线是双向的,其根数与存储字长相同,是机器字长的整数倍。

3. 为什么要设置总线判优控制? 常见的集中式总线控制有几种? 各有何特点? 哪种方式响应时间最快? 哪种方式对电路故障最敏感?

【答】总线判优控制可解决多个部件同时申请总线时的使用权分配问题。

常见的集中式总线控制有 3 种:链式查询、计数器定时查询、独立请求。

特点:链式查询方式连线简单,易于扩充,对电路故障最敏感;计数器定时查询方式优先级设置较灵活,对故障不敏感,连线及控制过程较复杂;独立请求方式速度最快,但硬件器件用量大,连线多,成本较高。

4. 解释下列概念:总线宽度、总线带宽、总线复用、总线的主设备(或主模块)、总线的从设备(或从模块)、总线的传输周期和总线的通信控制。

【答】总线宽度:通常指数据总线的根数。

总线带宽:总线的数据传输率,指单位时间内总线上传输数据的位数。

总线复用：指同一条信号线可以分时传输不同的信号。

总线的主设备(主模块)：指一次总线传输期间，拥有总线控制权的设备(模块)。

总线的从设备(从模块)：指一次总线传输期间，配合主设备完成数据传输的设备(模块)，它只能被动接受主设备发来的命令。

总线的传输周期：指总线完成一次完整而可靠的传输所需时间。

总线的通信控制：指总线传送过程中双方的时间配合方式。

5. 试比较同步通信和异步通信。

【答】同步通信：指由统一时钟控制的通信，控制方式简单，灵活性差，当系统中各部件工作速度差异较大时，总线工作效率明显下降。适合于速度差别不大的场合。

异步通信：指没有统一时钟控制的通信，部件间采用应答方式进行联系，控制方式较同步复杂，灵活性高，当系统中各部件工作速度差异较大时，有利于提高总线工作效率。

6. 为什么说半同步通信同时保留了同步通信和异步通信的特点？

【答】半同步通信既能像同步通信那样由统一时钟控制，又能像异步通信那样允许传输时间不一致，因此工作效率介于两者之间。

7. 分离式通信有何特点？主要用于什么系统？

【答】分离式通信的特点是：①各模块欲占用总线使用权都必须先提出申请；②在得到总线使用权后，主模块在选定的时间内向对方传送信息，采用同步方式传送，不再等待对方的回答信号；③各模块在准备数据的过程中都不占用总线，使总线可接收其他模块的请求；④总线被占用时都在做有效工作，或者通过它发送命令，或者通过它传送数据，不存在空闲等待时间，充分利用了总线的占用，从而实现了总线在多个主、从模块间进行信息交叉重叠并行传送。

分离式通信主要用于大型计算机系统。

8. 为什么要设置总线标准？你知道目前流行的总线标准有哪些？什么叫“Plug and play”？哪些总线有这一特点？

【答】总线标准的设置主要为了解决不同厂家各类模块化产品的兼容问题。

目前流行的总线标准有 ISA、EISA、PCI 等。

Plug and play：即插即用，EISA、PCI 等具有此功能。

9. 画一个具有双向传输功能的总线逻辑图。

【答】在总线的两端分别配置三态门，就可以使总线具有双向传输功能，如图 3.3 所示。

10. 设数据总线上接有 A、B、C、D 4 个寄存器，要求选用合适的 74 系列芯片，完成下列逻辑设计：

(1) 设计一个电路，在同一时间实现 $D \rightarrow A$ 、 $D \rightarrow B$ 和 $D \rightarrow C$ 寄存器间的传送。

(2) 设计一个电路，实现下列操作：

T_0 时刻完成 $D \rightarrow$ 总线； T_1 时刻完成总线 $\rightarrow A$ ；

T_2 时刻完成 $A \rightarrow$ 总线； T_3 时刻完成总线 $\rightarrow B$ 。

【答】(1) 由 T 打开三态门将 D 寄存器中的内容送至总线，由 CP 脉冲同时将总线上的数据打入到 A、B、C 寄存器中。T 和 CP 的时间关系如图 3.4 所示。

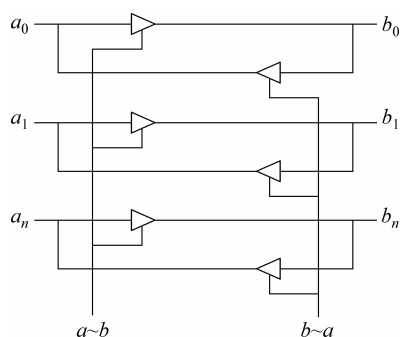
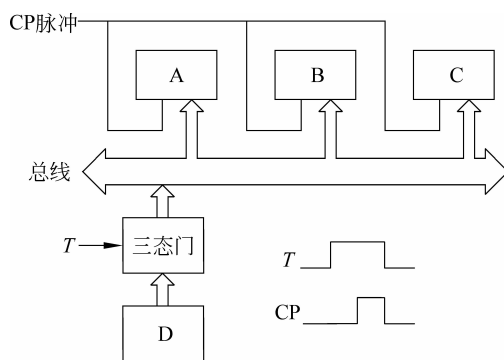
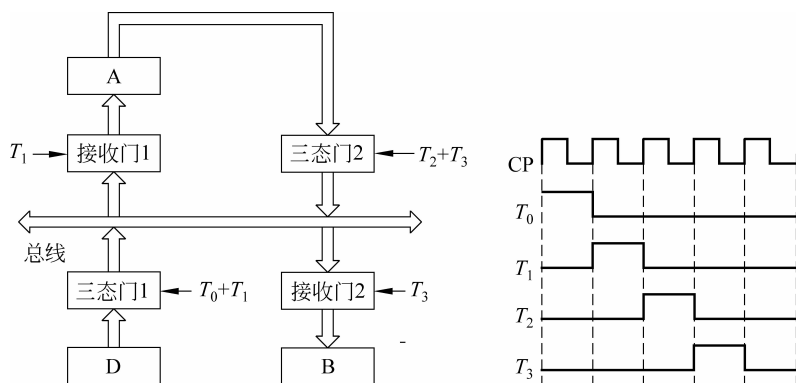


图 3.3 双向传输功能的总线逻辑电路

图 3.4 T 和 CP 的时间关系

(2) 三态门 1 受 $T_0 + T_1$ 控制，以确保 T_0 时刻 $D \rightarrow$ 总线，以及 T_1 时刻总线 $\rightarrow$ 接收门 1 $\rightarrow A$ 。三态门 2 受 $T_2 + T_3$ 控制，以确保 T_2 时刻 $A \rightarrow$ 总线，以及 T_3 时刻总线 $\rightarrow$ 接收门 2 $\rightarrow B$ 。 T_0 、 T_1 、 T_2 、 T_3 波形如图 3.5 所示。

图 3.5 T_0 、 T_1 、 T_2 、 T_3 波形

11. 什么是总线的数据传输率？它与哪些因素有关？

【答】 总线数据传输率即总线带宽，指单位时间内总线上传输数据的位数，通常用每秒传输信息的字节数来衡量。它与总线宽度和总线频率有关，总线宽度越宽，频率越快，数据传输率越高。

12. 设总线的时钟频率为 8MHz，一个总线周期等于一个时钟周期。如果一个总线周期中并行传送 16 位数据，试问总线的带宽是多少？

【答】 由于 $f = 8\text{MHz}$ ， $T = 1/f = 1/8\text{M}(\text{s})$ ，一个总线周期等于一个时钟周期
所以，总线带宽 $= 16 / (1/8\text{M}) = 128\text{Mbps}$ 。

13. 在一个 32 位的总线系统中，总线的时钟频率为 66MHz，假设总线最短传输周期为 4 个时钟周期，试计算总线的最大数据传输率。若想提高数据传输率，可采取什么措施？

【答】 总线传输周期 $= 4 \times 1/66\text{M}(\text{s})$

总线的最大数据传输率 $= 32 / (4/66\text{M}) = 528\text{Mbps}$ 。

若想提高数据传输率,可以提高总线时钟频率、增大总线宽度或者减少总线传输周期包含的时钟周期个数。

14. 在异步串行传送系统中,字符格式为:1个起始位、8个数据位、1个校验位、2个终止位。若要求每秒传送120个字符,试求传送的波特率和比特率。

【答】一帧包含: $1+8+1+2=12$ 位

故波特率为: $(1+8+1+2) \times 120 = 1440\text{bps}$

比特率为: $8 \times 120 = 960\text{bps}$

15. 某32位计算机,CPU主频为800MHz,Cache命中时的CPI为4,Cache块大小为32B;主存采用8体交叉存储方式,每个体的存储字长为32位、存储周期为40ns;存储器总线宽度为32位,总线时钟频率为200MHz,支持突发传送总线事务。每次读突发传送总线事务的过程包括送首地址和命令、存储器准备数据、传送数据。每次突发传送32B,传送地址或32位数据均需要一个总线时钟周期。请回答下列问题,要求给出理由或计算过程。

(1) CPU和总线的时钟周期各为多少?总线的带宽(即最大数据传输率)为多少?

(2) Cache缺失时,需要用几个读突发传送总线事务来完成一个主存块的读取?

(3) 存储器总线完成一次读突发传送总线事务所需的时间是多少?

(4) 若程序BP执行过程中共执行了100条指令,平均每条指令需进行1.2次访存,Cache缺失率为5%,不考虑替换等开销,则BP的CPU执行时间是多少?(2013年研究生)

【答】(1) CPU的时钟周期为: $1/800\text{MHz} = 1.25\text{ns}$ 。(1分)

总线的时钟周期为: $1/200\text{MHz} = 5\text{ns}$ 。(1分)

总线带宽为: $4\text{B} \times 200\text{MHz} = 800\text{MB/s}$ 或 $4\text{B}/5\text{ns} = 800\text{MB/s}$ 。(1分)

(2) Cache块大小是32B,因此Cache缺失时需要一个读突发传送总线事务读取一个主存块。(1分)

(3) 一次读突发传送总线事务包括一次地址传送和32B数据传送:用一个总线时钟周期传输地址;每隔 $40\text{ns}/8 = 5\text{ns}$ 启动一个体工作(各进行1次存取),第一个体读数据花费40ns,之后数据存取与数据传输重叠;用8个总线时钟周期传输数据。读突发传送总线事务时间: $5\text{ns} + 40\text{ns} + 8 \times 5\text{ns} = 85\text{ns}$ 。(2分)

(4) BP的CPU执行时间包括Cache命中时的指令执行时间和Cache缺失时带来的额外开销。命中时的指令执行时间: $100 \times 4 \times 1.25\text{ns} = 500\text{ns}$ 。(1分)指令执行过程中Cache缺失时的额外开销: $1.2 \times 100 \times 5\% \times 85\text{ns} = 510\text{ns}$ 。BP的CPU执行时间: $500\text{ns} + 510\text{ns} = 1010\text{ns}$ 。(2分)

3.4 存储器

1. 解释概念:主存、辅存、Cache、RAM、SRAM、DRAM、ROM、PROM、EPROM、EEPROM、CDROM、Flash Memory。

【答】主存:主存储器,用于存放正在执行的程序和数据。CPU可以直接进行随机读

写,访问速度较高。

辅存:辅助存储器,用于存放当前暂不执行的程序、数据和一些需要永久保存的信息。

Cache:高速缓冲存储器,介于CPU和主存之间以解决CPU和主存之间速度不匹配问题。

RAM:半导体随机存取存储器,主要用作计算机中的主存。

SRAM:静态半导体随机存取存储器。

DRAM:动态半导体随机存取存储器。

ROM:掩膜式半导体只读存储器。由芯片制造商在制造时写入内容,以后只能读出而不能写入。

PROM:可编程只读存储器,由用户根据需要确定写入内容,只能写入一次。

EPROM:紫外线擦写可编程只读存储器。需要修改内容时,先将其全部内容擦除,然后再编程。擦除依靠紫外线使浮动栅极上的电荷泄漏来实现。

EEPROM:电擦写可编程只读存储器。

CDROM:只读型光盘。

Flash Memory:闪存存储器,或称快擦型存储器。

2. 计算机中哪些部件可以用于存储信息?按速度、容量和价格/位排序说明。

【答】计算机中寄存器、Cache、主存、硬盘可以用于存储信息。

按速度由高至低排序为:寄存器、Cache、主存、硬盘。

按容量由小至大排序为:寄存器、Cache、主存、硬盘。

按价格/位由高至低排序为:寄存器、Cache、主存、硬盘。

3. 存储器的层次结构主要体现在什么地方?为什么要分这些层次?计算机如何管理这些层次?

【答】存储器的层次结构主要体现在Cache—主存和主存—辅存这两个存储层次上。

Cache—主存层次在存储系统中主要对CPU访存起加速作用,即从整体运行的效果分析,CPU访存速度加快,接近于Cache的速度,而寻址空间和位价却接近于主存。

主存—辅存层次在存储系统中主要起扩容作用,即从程序员的角度看,所使用的存储器的容量和位价接近于辅存,而速度接近于主存。

综合上述两个存储层次的作用,从整个存储系统来看,就达到了速度快、容量大、位价低的优化效果。

主存与Cache之间的信息调度功能全部由硬件自动完成。而主存与辅存层次的调度目前广泛采用虚拟存储技术实现,即将主存与辅存的一部分通过软硬结合的技术组成虚拟存储器,程序员可使用这个比主存实际空间(物理地址空间)大得多的虚拟地址空间(逻辑地址空间)编程,当程序运行时,再由软、硬件自动配合完成虚拟地址空间与主存实际物理空间的转换。因此,这两个层次上的调度或转换操作对于程序员来说都是透明的。

4. 说明存取周期和存取时间的区别。

【答】存取周期和存取时间的主要区别是:存取时间仅为完成一次操作的时间,而存取周期不仅包含操作时间,还包含操作后线路的恢复时间,即

存取周期 = 存取时间 + 恢复时间

5. 什么是存储器的带宽? 若存储器的数据总线宽度为 32 位, 存取周期为 200ns, 则存储器的带宽是多少?

【答】存储器的带宽指单位时间内从存储器进出信息的最大数量。

$$\begin{aligned}\text{存储器带宽} &= 1/200\text{ns} \times 32 \text{ 位} = 160\text{M 位/s} \\ &= 20\text{MB/s} = 5\text{M 字/s}\end{aligned}$$

注意: 字长 32 位, 不是 16 位。(注: $1\text{ns} = 10^{-9}\text{s}$)

6. 某机字长为 32 位, 其存储容量是 64KB, 按字编址它的寻址范围是多少? 若主存以字节编址, 试画出主存字地址和字节地址的分配情况。

【答】存储容量是 64KB 时, 按字节编址的寻址范围就是 64K。如按字编址, 其寻址范围为 $64\text{K}/(32/8) = 16\text{K}$, 主存字地址和字节地址的分配情况如图 3.6 所示。

7. 一个容量为 $16\text{K} \times 32$ 位的存储器, 其地址线和数据线的总和是多少? 当选用下列不同规格的存储芯片时, 各需要多少片?

$1\text{K} \times 4$ 位, $2\text{K} \times 8$ 位, $4\text{K} \times 4$ 位, $16\text{K} \times 1$ 位, $4\text{K} \times 8$ 位, $8\text{K} \times 8$ 位

【答】地址线和数据线的总和 = $14 + 32 = 46$ 根。

选择不同的芯片时, 各需要的片数为:

$1\text{K} \times 4$: $(16\text{K} \times 32)/(1\text{K} \times 4) = 16 \times 8 = 128$ 片

$2\text{K} \times 8$: $(16\text{K} \times 32)/(2\text{K} \times 8) = 8 \times 4 = 32$ 片

$4\text{K} \times 4$: $(16\text{K} \times 32)/(4\text{K} \times 4) = 4 \times 8 = 32$ 片

$16\text{K} \times 1$: $(16\text{K} \times 32)/(16\text{K} \times 1) = 1 \times 32 = 32$ 片

$4\text{K} \times 8$: $(16\text{K} \times 32)/(4\text{K} \times 8) = 4 \times 4 = 16$ 片

$8\text{K} \times 8$: $(16\text{K} \times 32)/(8\text{K} \times 8) = 2 \times 4 = 8$ 片

8. 什么叫刷新? 为什么要刷新? 说明刷新有几种方法。

【答】刷新: 指对 DRAM 定期进行的全部重写过程。

刷新原因: 因电容泄漏而引起的 DRAM 所存信息的衰减需要及时补充, 因此安排了定期刷新操作。

常用的刷新方法有 3 种: 集中式、分散式、异步式。

集中式: 在最大刷新间隔时间内, 集中安排一段时间进行刷新, 存在 CPU 访存死时间。

分散式: 在每个读/写周期之后插入一个刷新周期, 无 CPU 访存死时间。

异步式: 是集中式和分散式的折中。

讨论: 刷新与再生的异同?

共同点:

① 动作机制一样, 都是利用 DRAM 存储元破坏性读操作时的重写过程实现。

② 操作性质一样, 都是属于重写操作。

字节地址		字地址
0000H		0000H
0001H		
0002H		
0003H		
0004H		0001H
0005H		
0006H		
0007H		
0008H		0002H
0009H		

图 3.6 主存字地址和字节地址的分配示意图

区别:

① 解决的问题不一样。再生主要解决 DRAM 存储元破坏性读出时的信息重写问题;刷新主要解决长时间不访存时的信息衰减问题。

② 操作的时间不一样。再生紧跟在读操作之后,时间上是随机进行的;刷新以最大间隔时间为周期定时重复进行。

③ 动作单位不一样。再生以存储单元为单位,每次仅重写刚被读出的一个字的所有位;刷新以行为单位,每次重写整个存储器所有芯片内部存储矩阵的同一行。

④ 芯片内部 I/O 操作不一样。读出再生时芯片数据引脚上有读出数据输出;刷新时由于 CAS 信号无效,芯片数据引脚上无读出数据输出(唯 RAS 有效刷新,内部读)。

9. 半导体存储器芯片的译码驱动方式有几种?

【答】半导体存储器芯片的译码驱动方式有两种,即线选法和重合法。

线选法:地址译码信号只选中同一个字的所有位,结构简单,费器材。

重合法:地址分行、列两部分译码,行、列译码线的交叉点即为所选单元。这种方法通过行、列译码信号的重合来选址,也称矩阵译码。可大大节省器材用量,是最常用的译码驱动方式。

10. 一个 $8K \times 8$ 位的动态 RAM 芯片,其内部结构排列成 256×256 形式,存取周期为 $0.1\mu s$ 。试问采用集中刷新、分散刷新和异步刷新 3 种方式的刷新闻隔各为多少?

【答】采用分散刷新方式刷新闻隔为: $2ms$,其中刷新死时间为: $256 \times 0.1\mu s = 25.6\mu s$

采用分散刷新方式刷新闻隔为: $256 \times (0.1\mu s + 25.6\mu s) = 51.2\mu s$

采用异步刷新方式刷新闻隔为: $2ms$

11. 用 1024×4 位的存储芯片组成一个容量为 $64K \times 8$ 位的存储器,如果将 $64K$ 分成 4 个页面,每个页面分 16 组,共需多少片存储芯片?

【答】设采用 SRAM 芯片,则:

总片数 = $(64K \times 8 \text{ 位}) / (1024 \times 4 \text{ 位}) = 64 \times 2 = 128$ 片

题意分析:本题设计的存储器结构上分为总体、页面、组 3 级,因此画图时也应分 3 级画。首先应确定各级的容量:

页面容量 = 总容量 / 页面数 = $64K \times 8 / 4 = 16K \times 8$ 位,4 片 $16K \times 8$ 字串联成 $64K \times 8$ 位

组容量 = 页面容量 / 组数 = $16K \times 8 \text{ 位} / 16 = 1K \times 8$ 位,16 片 $1K \times 8$ 位字串联成 $16K \times 8$ 位

组内片数 = 组容量 / 片容量 = $1K \times 8 \text{ 位} / 1K \times 4 \text{ 位} = 2$ 片,两片 $1K \times 4$ 位芯片位并联成 $1K \times 8$ 位

12. 设有一个 $64K \times 8$ 位的 RAM 芯片,试问该芯片共有多少个基本单元电路(简称存储基元)?欲设计一种具有上述同样多存储基元的芯片,要求对芯片字长的选择应满足地址线和数据线的总和为最小,试确定这种芯片的地址线 and 数据线,并说明有几种解答。

【答】存储基元总数 = $64K \times 8 \text{ 位} = 512K \text{ 位} = 2^{19} \text{ 位}$;

思路:如要满足地址线和数据线总和最小,应尽量把存储元安排在字向,因为地址位

数和字数成 2 的幂的关系,可较好地压缩线数。

设地址线根数为 a ,数据线根数为 b ,则片容量为: $2^a \times b = 2^{19}$; $b = 2^{19-a}$ 。

若 $a=19, b=1$,总和 $=19+1=20$ 。

$a=18, b=2$,总和 $=18+2=20$ 。

$a=17, b=4$,总和 $=17+4=21$ 。

$a=16, b=8$,总和 $=16+8=24$ 。

.....

由上可以看出:芯片字数越少,芯片字长越长,引脚数越多。芯片字数减 1、芯片位数均按 2 的幂变化。

结论:如果满足地址线和数据线的总和为最小,这种芯片的引脚分配方案有两种:地址线=19 根,数据线=1 根;或地址线=18 根,数据线=2 根。

13. 某 8 位微型机地址码为 18 位,若使用 $4K \times 4$ 位的 RAM 芯片组成模块板结构的存储器,试问:

- (1) 该机所允许的最大主存空间是多少?
- (2) 若每个模块板为 $32K \times 8$ 位,共需几个模块板?
- (3) 每个模块板内共有几片 RAM 芯片?
- (4) 共有多少片 RAM?
- (5) CPU 如何选择各模块板?

【解】(1) 该机所允许的最大主存空间是: $2^{18} \times 8$ 位 $= 256K \times 8$ 位 $= 256KB$ 。

(2) 模块板总数 $= 256K \times 8 / 32K \times 8 = 8$ 块。

(3) 板内片数 $= 32K \times 8$ 位 $/ 4K \times 4$ 位 $= 8 \times 2 = 16$ 片。

(4) 总片数 $= 16$ 片 $\times 8 = 128$ 片。

(5) CPU 通过最高 3 位地址译码输出选择模板,次高 3 位地址译码输出选择芯片。地址格式分配如图 3.7 所示。

模板号(3位)	芯片号(3位)	片内地址(12位)
---------	---------	-----------

图 3.7 地址格式

14. 设 CPU 共有 16 根地址线,8 根数据线,并用 $\overline{MREQ}$ (低电平有效)作访存控制信号, $R/\overline{W}$ 作读写命令信号(高电平为读,低电平为写)。现有下列存储芯片: ROM($2K \times 8$ 位、 $4K \times 4$ 位、 $8K \times 8$ 位), RAM($1K \times 4$ 位、 $2K \times 8$ 位、 $4K \times 8$ 位),及 74138 译码器和其他门电路(门电路自定)。试从上述规格中选用合适的芯片,画出 CPU 和存储芯片的连接图。要求:

- (1) 最小 4K 地址为系统程序区,4096~16383 地址范围为用户程序区。
- (2) 指出选用的存储芯片类型及数量。
- (3) 详细画出片选逻辑。

【解】(1) 地址空间分配图:

系统程序区(ROM 共 4KB): 0000H~0FFFH

用户程序区(RAM 共 12KB): 1000H~3FFFH

(2) 选片：ROM：选择 4K×4 位芯片两片，位并联。

RAM：选择 4K×8 位芯片 3 片，字串联（RAM1 地址范围为：1000H～1FFFH，RAM2 地址范围为 2000H～2FFFH，RAM3 地址范围为 3000H～3FFFH）。

(3) 各芯片二进制地址分配如表 3.1 所列。

表 3.1 各芯片二进制地址分配表

	A ₁₅	A ₁₄	A ₁₃	A ₁₂	A ₁₁	A ₁₀	A ₉	A ₈	A ₇	A ₆	A ₅	A ₄	A ₃	A ₂	A ₁	A ₀
ROM1,2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1
RAM1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
RAM2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1
RAM3	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1

CPU 和存储器连接逻辑图及片选逻辑如图 3.8 所示。

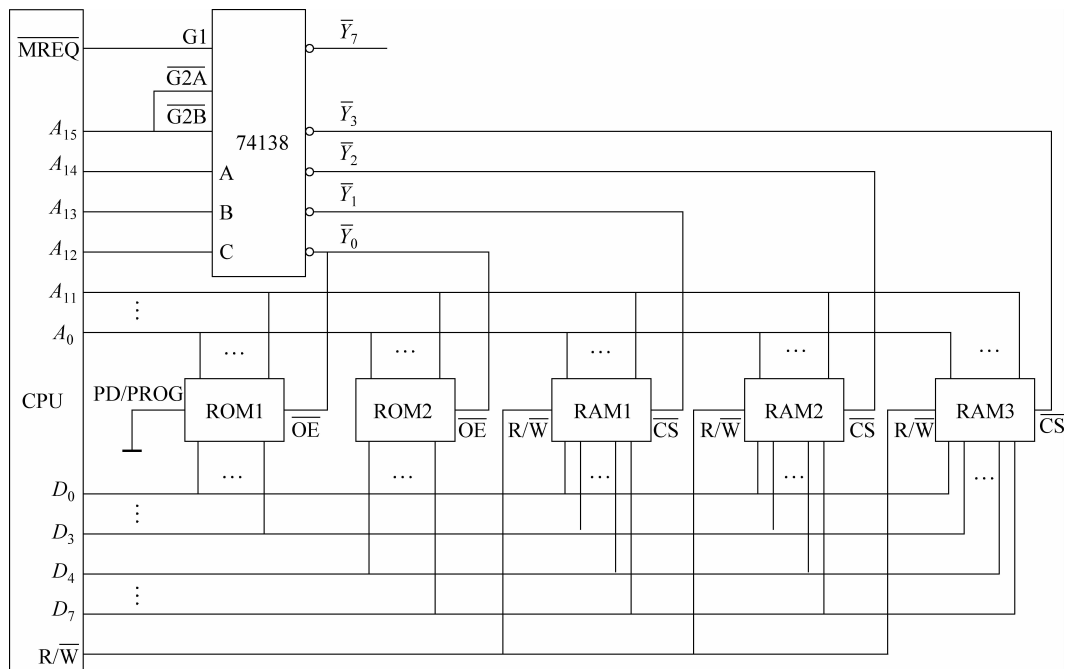


图 3.8 CPU 和存储器连接逻辑电路

15. CPU 假设同上题，现有 8 片 8K×8 位的 RAM 芯片与 CPU 相连，试回答：

(1) 用 74138 译码器画出 CPU 与存储芯片的连接图。

(2) 写出每片 RAM 的地址范围。

(3) 如果运行时发现不论往哪片 RAM 写入数据后,以 A000H 为起始地址的存储芯片都有与其相同的数据,分析故障原因。

(4) 根据(1)的连接图,若出现地址线 A_{13} 与 CPU 断线,并搭接到高电平上,将出现什么后果?

【解】(1) CPU 与存储器芯片连接逻辑电路如图 3.9 所示。

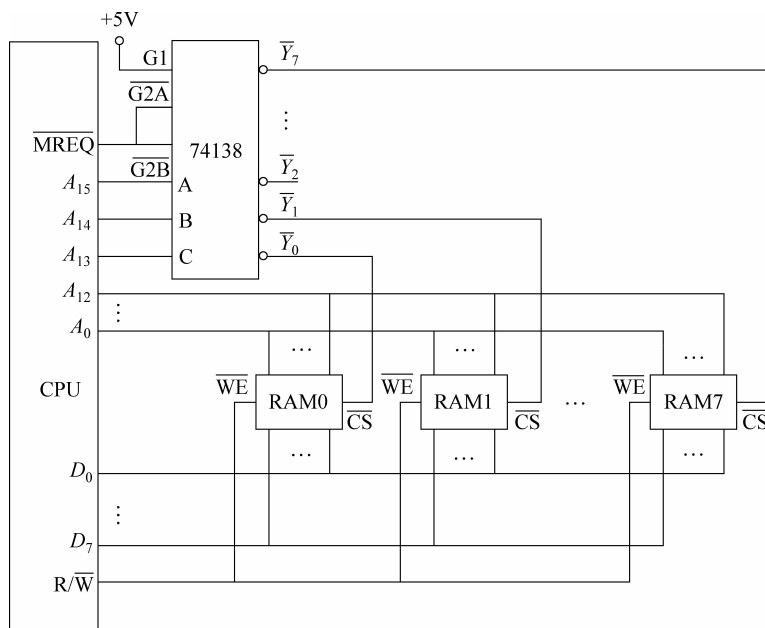


图 3.9 CPU 与存储芯片的连接

(2) 地址空间分配范围如下:

RAM0: 0000H~1FFFH	RAM1: 2000H~3FFFH
RAM2: 4000H~5FFFH	RAM3: 6000H~7FFFH
RAM4: 8000H~9FFFH	RAM5: A000H~BFFFH
RAM6: C000H~DFFFH	RAM7: E000H~FFFFH

(3) 如果运行时发现不论往哪片 RAM 写入数据后,以 A000H 为起始地址的存储芯片(RAM5)都有与其相同的数据,则根本的故障原因为:该存储芯片的片选输入端很可能总是处于低电平。假设芯片与译码器本身都是好的,可能的情况有:

- ① 该片的 $\overline{CS}$ 端与 $\overline{WE}$ 端错连或短路。
- ② 该片的 $\overline{CS}$ 端与 CPU 的 MREQ 端错连或短路。
- ③ 该片的 $\overline{CS}$ 端与地线错连或短路。

(4) 如果地址线 A_{13} 与 CPU 断线,并搭接到高电平上,将会出现 A_{13} 恒为“1”的情况。此时存储器只能寻址 $A_{13}=1$ 的地址空间(奇数片), $A_{13}=0$ 的另一半地址空间(偶数片)将永远访问不到。若对 $A_{13}=0$ 的地址空间(偶数片)进行访问,只能错误地访问到 $A_{13}=1$ 的对应空间(奇数片)中去。

16. 写出 1100、1101、1110、1111 对应的汉明码。

【解】有效信息均为 $n=4$ 位, 假设有效信息用 $b_4b_3b_2b_1$ 表示。

校验位位数 $k=3$ 位 ($2^k \geq n+k+1$)。

设校验位分别为 c_1, c_2, c_3 , 则汉明码共 $4+3=7$ 位, 即 $c_1c_2b_4c_3b_2b_1$

校验位在汉明码中分别处于第 1、2、4 位。

$$c_1 = b_4 \oplus b_3 \oplus b_1$$

$$c_2 = b_4 \oplus b_2 \oplus b_1$$

$$c_3 = b_3 \oplus b_2 \oplus b_1$$

当有效信息为 1100 时, $c_3c_2c_1=110$, 汉明码为 0111100。

当有效信息为 1101 时, $c_3c_2c_1=001$, 汉明码为 1010101。

当有效信息为 1110 时, $c_3c_2c_1=000$, 汉明码为 0010110。

当有效信息为 1111 时, $c_3c_2c_1=111$, 汉明码为 1111111。

17. 已知收到的汉明码(按配偶原则配置)为 1100100、1100111、1100000、1100001, 检查上述代码是否出错? 第几位出错?

【解】假设接收到的汉明码为: $c'_1c'_2b'_4c'_3b'_2b'_1$

纠错过程如下:

$$p_1 = c'_1 \oplus b'_4 \oplus b'_3 \oplus b'_1$$

$$p_2 = c'_2 \oplus b'_4 \oplus b'_2 \oplus b'_1$$

$$p_3 = c'_3 \oplus b'_3 \oplus b'_2 \oplus b'_1$$

如果收到的汉明码为 1100100, 则 $p_3p_2p_1=011$, 说明代码有错, 第 3 位(b'_4)出错, 有效信息为 1100。

如果收到的汉明码为 1100111, 则 $p_3p_2p_1=111$, 说明代码有错, 第 7 位(b'_1)出错, 有效信息为 0110。

如果收到的汉明码为 1100000, 则 $p_3p_2p_1=110$, 说明代码有错, 第 6 位(b'_2)出错, 有效信息为 0010。

如果收到的汉明码为 1100001, 则 $p_3p_2p_1=001$, 说明代码有错, 第 1 位(c'_1)出错, 有效信息为 0001。

18. 已经接收到下列汉明码, 分别写出它们所对应的欲传送代码。

(1) 1100000(按偶性配置)

(2) 1100010(按偶性配置)

(3) 1101001(按偶性配置)

(4) 0011001(按奇性配置)

(5) 1000000(按奇性配置)

(6) 1110001(按奇性配置)

【解】(一) 假设接收到的汉明码为 $C'_1C'_2B'_4C'_3B'_2B'_1$, 按偶性配置则:

$$P_1 = C'_1 \oplus B'_4 \oplus B'_3 \oplus B'_1$$

$$P_2 = C'_2 \oplus B'_4 \oplus B'_2 \oplus B'_1$$

$$P_3 = C'_3 \oplus B'_3 \oplus B'_2 \oplus B'_1$$

(1) 如接收到的汉明码为 1100000, 则

$$P_1 = 1 \oplus 0 \oplus 0 \oplus 0 = 1$$

$$P_2 = 1 \oplus 0 \oplus 0 \oplus 0 = 1$$

$$P_3 = 0 \oplus 0 \oplus 0 = 0$$

$P_3P_2P_1=011$,第3位出错,可纠正为1110000,故欲传送的信息为1000。

(2) 如接收到的汉明码为1100010,则

$$P_1 = 1 \oplus 0 \oplus 0 \oplus 0 = 1$$

$$P_2 = 1 \oplus 0 \oplus 1 \oplus 0 = 0$$

$$P_3 = 0 \oplus 0 \oplus 0 = 0$$

$P_3P_2P_1=001$,第1位出错,可纠正为0100010,故欲传送的信息为0010。

(3) 如接收到的汉明码为1101001,则

$$P_1 = 1 \oplus 0 \oplus 0 \oplus 1 = 0$$

$$P_2 = 1 \oplus 0 \oplus 0 \oplus 1 = 0$$

$$P_3 = 1 \oplus 0 \oplus 1 = 0$$

$P_3P_2P_1=000$,传送无错,故欲传送的信息为0001。

(二) 假设接收到的汉明码为 $C'_1C'_2B'_4C'_3B'_3B'_2B'_1$,按奇性配置则:

$$P_1 = C'_1 \oplus B'_4 \oplus B'_3 \oplus B'_1 \oplus 1$$

$$P_2 = C'_2 \oplus B'_4 \oplus B'_2 \oplus B'_1 \oplus 1$$

$$P_3 = C'_3 \oplus B'_3 \oplus B'_1 \oplus 1$$

(4) 如接收到的汉明码为0011001,则

$$P_1 = 0 \oplus 1 \oplus 0 \oplus 1 \oplus 1 = 1$$

$$P_2 = 0 \oplus 1 \oplus 0 \oplus 1 \oplus 1 = 1$$

$$P_3 = 1 \oplus 0 \oplus 1 \oplus 1 = 1$$

$P_3P_2P_1=111$,第7位出错,可纠正为0011000,故欲传送的信息为1000。

(5) 如接收到的汉明码为1000000,则

$$P_1 = 1 \oplus 0 \oplus 0 \oplus 0 \oplus 1 = 0$$

$$P_2 = 0 \oplus 1 \oplus 0 \oplus 0 \oplus 1 = 0$$

$$P_3 = 0 \oplus 0 \oplus 0 \oplus 1 = 1$$

$P_3P_2P_1=100$,第4位出错,可纠正为1001000,故欲传送的信息为0000。

(6) 如接收到的汉明码为1110001,则

$$P_1 = 1 \oplus 1 \oplus 0 \oplus 1 \oplus 1 = 0$$

$$P_2 = 1 \oplus 1 \oplus 0 \oplus 1 \oplus 1 = 0$$

$$P_3 = 0 \oplus 0 \oplus 1 \oplus 1 = 0$$

$P_3P_2P_1=000$,传送无错,故欲传送的信息为1001。

19. 欲传送的二进制代码为1001101,用奇校验来确定其对应的汉明码,若在第6位出错,说明纠错过程。

【解】欲传送的二进制代码为1001101,有效信息位数为 $n=7$ 位,则汉明校验的校验位为 k 位,则: $2k \geq n + k + 1, k = 4$,进行奇校验设校验位为 $C_1C_2C_3C_4$,汉明码为 $C_1C_2B_7C_3B_6B_5B_4C_4B_3B_2B_1$ 。

$$C_1 = 1 \oplus B_7 \oplus B_6 \oplus B_4 \oplus B_3 \oplus B_1 = 1 \oplus 1 \oplus 0 \oplus 1 \oplus 1 \oplus 1 = 1$$

$$C_2 = 1 \oplus B_7 \oplus B_5 \oplus B_4 \oplus B_2 \oplus B_1 = 1 \oplus 1 \oplus 0 \oplus 1 \oplus 0 \oplus 1 = 0$$

$$C_3 = 1 \oplus B_6 \oplus B_5 \oplus B_4 = 1 \oplus 0 \oplus 0 \oplus 1 = 0$$

$$C_4 = 1 \oplus B_3 \oplus B_2 \oplus B_1 = 1 \oplus 1 \oplus 0 \oplus 1 = 1$$

故传送的汉明码为 10100011101, 若第 6 位 (B_5) 出错, 即接收的码字为 10100111101,

则

$$P_1 = 1 \oplus C'_1 \oplus B'_7 \oplus B'_6 \oplus B'_4 \oplus B'_3 \oplus B'_1 = 1 \oplus 1 \oplus 1 \oplus 0 \oplus 1 \oplus 1 \oplus 1 = 0$$

$$P_2 = 1 \oplus C'_2 \oplus B'_7 \oplus B'_5 \oplus B'_4 \oplus B'_2 \oplus B'_1 = 1 \oplus 0 \oplus 1 \oplus 1 \oplus 1 \oplus 0 \oplus 1 = 1$$

$$P_3 = 1 \oplus C'_3 \oplus B'_6 \oplus B'_5 \oplus B'_4 = 1 \oplus 0 \oplus 0 \oplus 1 \oplus 1 = 1$$

$$P_4 = 1 \oplus C'_4 \oplus B'_3 \oplus B'_2 \oplus B'_1 = 1 \oplus 1 \oplus 1 \oplus 0 \oplus 1 = 0$$

$P_4 P_3 P_2 P_1 = 0110$ 说明第 6 位出错, 对第 6 位取反即完成纠错。

20. 为什么在汉明码纠错过程中, 新的检测位 $P_4 P_2 P_1$ 的状态即指出了编码中错误的信息位?

【解】 汉明码属于分组奇偶校验, $P_4 P_2 P_1 = 000$, 说明接收方生成的校验位和收到的校验位相同; 否则, 不同说明出错。由于分组时校验位只参加一组奇偶校验, 有效信息参加至少两组奇偶校验, 如果校验位出错, $P_4 P_2 P_1$ 的某一位将为 1, 刚好对应位号 4、2、1; 如果有效信息出错, 将引起 $P_4 P_2 P_1$ 中至少两位为 1, 如 B_1 出错, 将使 $P_4 P_1$ 均为 1, $P_2 = 0$, $P_4 P_2 P_1 = 101$ 。

21. 某机字长 16 位, 常规的存储空间为 64K 字, 若想不改用其他高速的存储芯片, 而使访存速度提高到 8 倍, 可采取什么措施? 画图说明。

【解】 若想不改用高速存储芯片, 而使访存速度提高到 8 倍, 可采取 8 体交叉存取技术, 8 体交叉访问时序如图 3.10 所示。

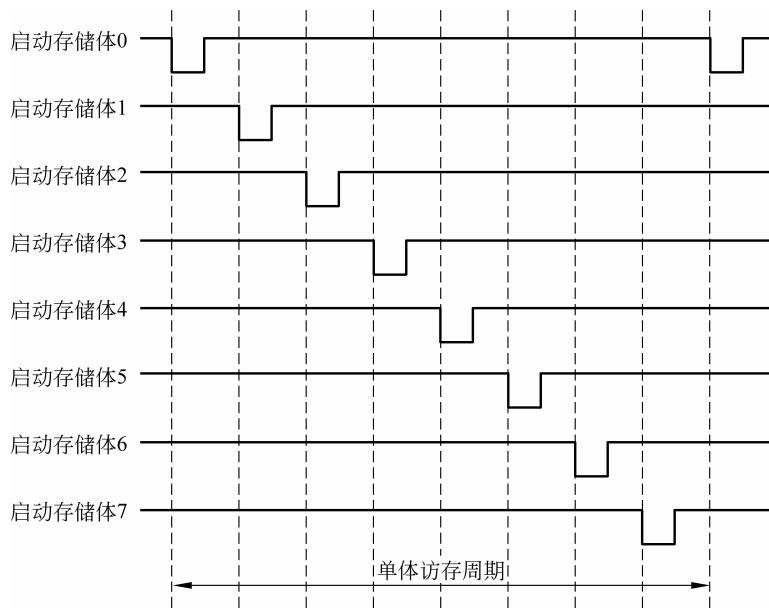


图 3.10 8 体交叉访问时序图

22. 设 CPU 共有 16 根地址线, 8 根数据线, 并用 $M/\overline{IO}$ 作为访问存储器或 I/O 的控制信号(高电平为访存, 低电平为访 I/O), $\overline{WR}$ (低电平有效) 为写命令, $\overline{RD}$ (低电平有效) 为读命令。设计一个容量为 64KB 的采用低位交叉编址的 8 体并行结构存储器。现有图 3.11 所示的存储器芯片和 138 译码器。

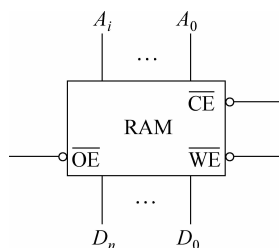


图 3.11 存储器芯片和 138 译码器示意图

画出 CPU 和存储器芯片(芯片容量自定)的连接图, 并写出图中每个存储芯片的地址范围(用十六进制数表示)。

【解】 8 体低位交叉并行存储器的每个存储体容量为 $64KB/8 = 8KB$, 因此应选择 8KB RAM 芯片, 芯片地址线 12 根 ($A_0 \sim A_{12}$), 数据线 8 根 ($D_0 \sim D_7$), 用 138 译码器进行存储体的选择。设计如图 3.12 所示。

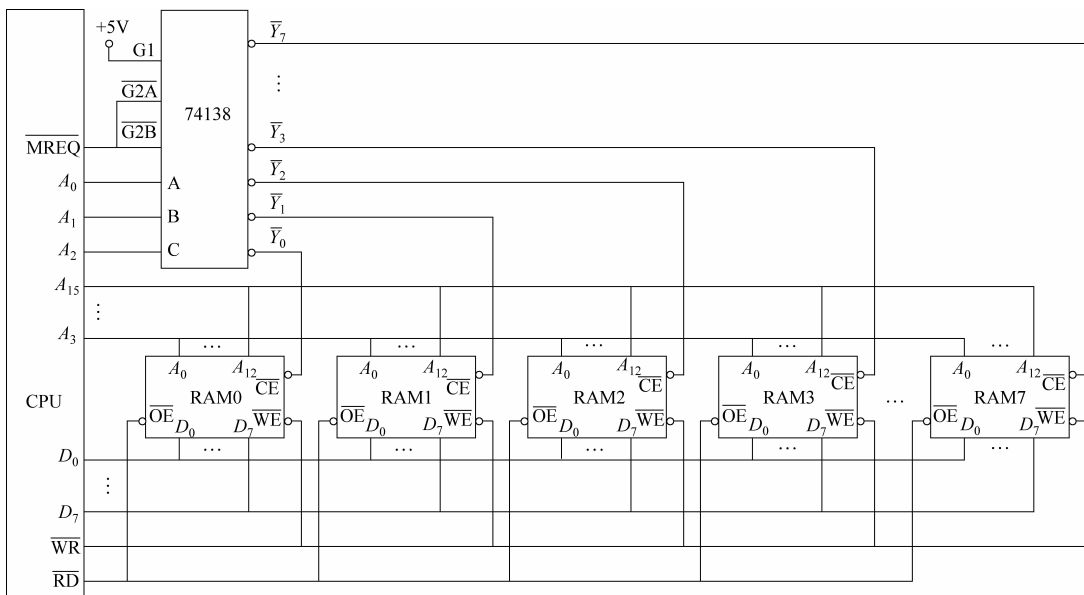


图 3.12 CPU 和存储器芯片连接

23. 一个 4 体低位交叉的存储器, 假设存储周期为 T , CPU 每隔 $T/4$ 启动一个存储体, 试问依次访问 64 个字需多少个存取周期?

【解】 4 体低位交叉的存储器的总线传输周期为 τ , $\tau = T/4$, 依次访问 64 个字所需时间为

$$t = T + (64 - 1)\tau = T + 63T/4 = 16.75T$$

24. 什么是“程序访问的局部性”? 存储系统中哪一级采用了程序访问的局部性原理?

【解】 程序运行的局部性原理: 指在一小段时间内, 最近被访问过的程序和数据很可

能再次被访问;在空间上,这些被访问的程序和数据往往集中在一小片存储区;在访问顺序上,指令顺序执行比转移执行的可能性大($\approx 5:1$)。存储系统中 Cache—主存层次和主存—辅存层次均采用了程序访问的局部性原理。

25. 计算机中设置 Cache 的作用是什么?能否将 Cache 的容量扩大最后取代主存?为什么?

【解】计算机中设置 Cache 的作用是解决 CPU 和主存速度不匹配问题。

不能将 Cache 的容量扩大取代主存,原因是:①Cache 容量越大成本越高,难以满足人们追求低价格的要求;②如果取消主存,当 CPU 访问 Cache 失败时,需要将辅存的内容调入 Cache 再由 CPU 访问,造成 CPU 等待时间太长,损失更大。

26. Cache 做在 CPU 芯片内有什么好处?将指令 Cache 和数据 Cache 分开又有什么好处?

【解】Cache 做在 CPU 芯片内主要有下面几个好处:

(1) 可提高外部总线的利用率。因为 Cache 在 CPU 芯片内,CPU 访问 Cache 时不必占用外部总线。

(2) Cache 不占用外部总线就意味着外部总线可更多地支持 I/O 设备与主存的信息传输,增强了系统的整体效率。

(3) 可提高存取速度。因为 Cache 与 CPU 之间的数据通路大大缩短,故存取速度得以提高。

将指令 Cache 和数据 Cache 分开有以下好处:

(1) 可支持超前控制和流水线控制,有利于这类控制方式下指令预取操作的完成。

(2) 指令 Cache 可用 ROM 实现,以提高指令存取的可靠性。

(3) 数据 Cache 对不同数据类型的支持更为灵活,既可支持整数(如 32 位),也可支持浮点数据(如 64 位)。

补充: Cache 结构改进的第 3 个措施是分级实现,如二级缓存结构,即在片内 Cache(L1)和主存之间再设一个片外 Cache(L2),片外缓存既可以弥补片内缓存容量不够大的缺点,又可在主存与片内缓存间起到平滑速度差的作用,加速片内缓存的调入调出速度。

27. 设主存容量为 256K 字,Cache 容量为 2K 字,块长为 4。

(1) 设计 Cache 地址格式,Cache 中可装入多少块数据?

(2) 在直接映射方式下,设计主存地址格式。

(3) 在 4 路组相联映射方式下,设计主存地址格式。

(4) 在全相联映射方式下,设计主存地址格式。

(5) 若存储字长为 32 位,存储器按字节寻址,写出上述 3 种映射方式下主存的地址格式。

【解】(1) Cache 容量为 2K 字,块长为 4,Cache 共有 $2K/4=2^{11}/2^2=2^9=512$ 块,Cache 字地址 9 位,字块内地址为 2 位,因此,Cache 地址格式设计如图 3.13 所示。

(2) 主存容量为 256K 字= 2^{18} 字,主存地址共 18 位,共分 $256K/4=2^{16}$ 块,主存字块标记为 $18-9-2=7$ 位。

直接映射方式下主存地址格式如图 3.14 所示。

Cache字块地址(9位)	字块内地址(2位)
---------------	-----------

图 3.13 Cache 地址格式示意图

主存字块标记(7位)	Cache字块地址(9位)	字块内地址(2位)
------------	---------------	-----------

图 3.14 直接映射方式下主存地址格式示意图

(3) 根据 4 路组相联的条件, 一组内共有 4 块, 得 Cache 共分为 $512/4=128=2^7$ 组, 主存字块标记为 $18-7-2=9$ 位, 主存地址格式设计如图 3.15 所示。

主存字块标记(9位)	组地址(7位)	字块内地址(2位)
------------	---------	-----------

图 3.15 4 路组相联主存地址格式示意图

(4) 在全相联映射方式下, 主存字块标记为 $18-2=16$ 位, 其地址格式如图 3.16 所示。

主存字块标记(16位)	字块内地址(2位)
-------------	-----------

图 3.16 全相联映射方式地址格式示意图

(5) 若存储字长为 32 位, 存储器按字节寻址, 则主存容量为 $256\text{K} \times 32/4=2^{21}\text{B}$, Cache 容量为 $2\text{K} \times 32/4=2^{14}\text{B}$, 块长为 $4 \times 32/4=32\text{B}=2^5\text{B}$, 字块内地址为 5 位。在直接映射方式下, 主存字块标记为 $21-9-5=7$ 位, 主存地址格式如图 3.17 所示。

主存字块标记(7位)	Cache字块地址(9位)	字块内地址(5位)
------------	---------------	-----------

图 3.17 32 位字长的直接映射方式主存地址格式示意图

在 4 路组相联映射方式下, 主存字块标记为 $21-7-5=9$ 位, 主存地址格式如图 3.18 所示。

主存字块标记(9位)	组地址(7位)	字块内地址(5位)
------------	---------	-----------

图 3.18 32 位字长的 4 路组相联映射方式主存地址格式示意图

在全相联映射方式下, 主存字块标记为 $21-5=16$ 位, 主存地址格式如图 3.19 所示。

主存字块标记(16位)	字块内地址(5位)
-------------	-----------

图 3.19 32 位字长的全相联映射方式主存地址格式示意图

28. 假设 CPU 执行某段程序时共访问 Cache 命中 4800 次, 访问主存 200 次, 已知 Cache 的存取周期为 30ns, 主存的存取周期为 150ns, 求 Cache 的命中率以及 Cache—主存系统的平均访问时间和效率, 试问该系统的性能提高了多少倍?

【解】 Cache 被访问命中率为: $4800/(4800+200)=24/25=96\%$

则 Cache—主存系统的平均访问时间为: $t_a=0.96 \times 30\text{ns} + (1-0.96) \times 150\text{ns} =$