

专业课

计算机

计算机体系结构

袁礼

华图网校

版权所有 盗版必究

目录

计算机体系结构	1
1. 计算机体系结构的基础知识	1
2. 数据标识、寻址方式与指令系统	5
3. 存储系统	10
4. 总线、中断与输入输出系统	12
5. 标量处理机	15
6. 向量处理机	19

计算机体系结构

- 计算机体系结构的基础知识
- 数据标识、寻址方式与指令系统
- 存储系统
- 总线、中断与输入输出系统
- 标量和流水处理机
- 向量处理机

1. 计算机体系结构的基础知识

一、计算机系统的多级层次结构



(1) 软件和硬件在逻辑功能上具有等效性

具有相同功能的计算机系统, 其软、硬件功能分配比例可以在很宽的范围内变化。



（2）计算机体系结构

1964 年 G.M.Amdahl 在介绍 IBM360 系统时提出：计算机体系结构是从程序员所看到的计算机属性，即程序员编写出能在机器上正确运行的程序所必须了解的概念性结构和功能特性。

系统结构是对计算机系统中各级界面的划分、定义及其上下功能的分配。

系统结构设计主要研究界面的属性的透明性的取舍。

计算机体系结构指的是传统机器级的系统结构。

计算机体系结构研究的是软、硬件之间的功能分配以及对传统机器级界面的确定。

二、计算机组成、实现

计算机组成指的是计算机体系结构的逻辑实现，包括机器级内部的数据流、控制流的组成以及逻辑设计等。

它着眼于机器级内各事件的排序方式与控制机构、各部件的功能及各部件间的联系。

计算机实现指的是计算机组成的物理实现，包括处理机、主存等部件的物理结构，器件的集成度和速度，器件、模块、插件、底板的划分与连接，专用器件的设计，微组装技术，信号传输，电源、冷却及整机装配技术等。

它着眼于器件技术和微组装技术，其中，器件技术在实现技术中起着主导作用。

（1）计算机体系结构、组成、实现的关系

三者互不相同，但又相互影响。

计算机体系结构学科实际上包括了系统结构和组成两个方面的内容。

研究的是软、硬件的功能分配以及如何更好、更合理地实现分配给硬件的功能。

代表从程序设计者、计算机设计者两个不同角度看到的计算机体系结构。

（2）软硬件取舍的基本原则

系统有高的性能价格比。

尽可能不要过多或不合理地限制各种组成、实现技术的采用。

进一步缩短高级语言与机器语言、操作系统与计算机体系结构之间的语义差距。

计算机系统的性能评测

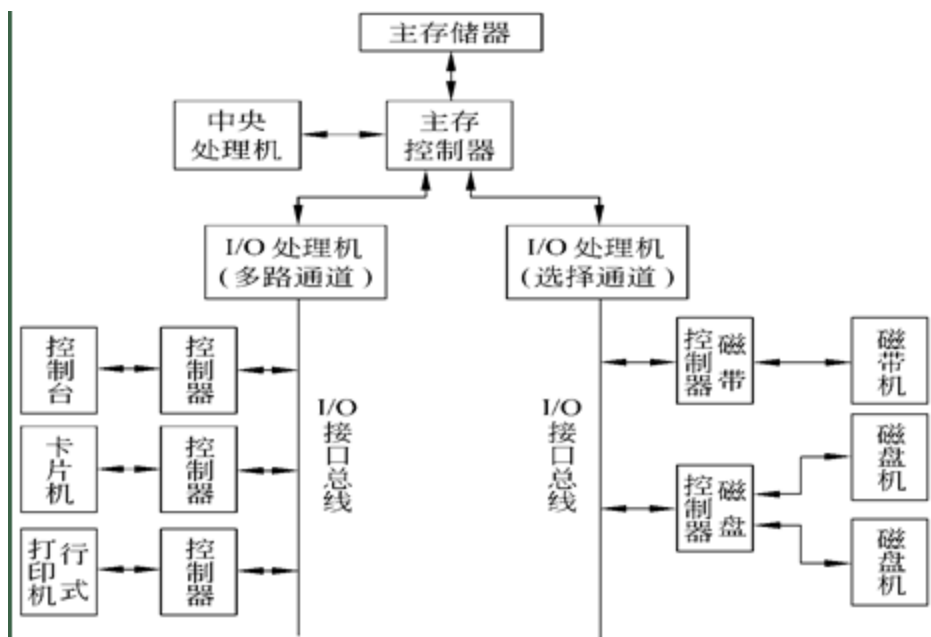
峰值性能—理想情况下计算机系统可获得的最高理论性能值。

持续性能—实际性能，有算术平均、调和平均和几何平均三种表示。

MIPS—每秒百万条指令数。

MFLOPS—每秒百万次浮点运算。

IBM370 系列



三、计算机系统的定量设计原理

Huffman 压缩原理：加速处理高概率事件远比加速处理低概率事件对性能提高要显著。

Admal 定律：系统加速比定义为系统改进后的性能与未改进时的性能的比值。
Admal 定律表明了性能提高量的递减规律。

程序的局部性原理：时间局部性、空间局部性。

计算机系统的设计思路

- 由上往下一适合于专用机设计。
- 由下往上一60-70 年代以前的通用机设计思想。
- 由中间往两边一合理的软、硬件功能分配。

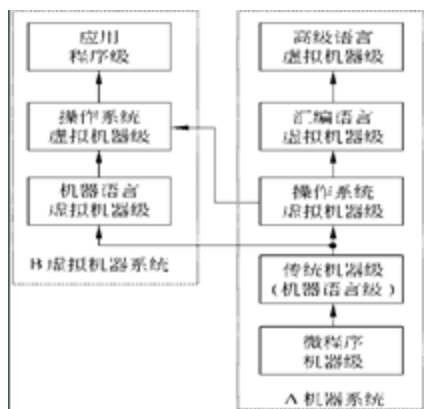
四、软件、应用、器件对系统结构的影响

软件的可移植性。

实现软件移植的基本技术：统一高级语言、采用系列机、模拟与仿真。

向上（下）兼容、向前（后）兼容。

系列机软件必须保证向后兼容，力争向上兼容。



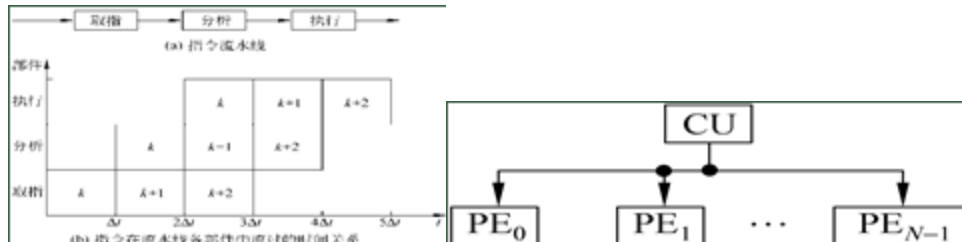
五、系统结构中的并行性发展

并行性 — 可以同时进行运算或操作的特性，包括同时性、并发性两重含义。

并行性的等级：

1. 从执行程序的角度：指令内部，指令之间，任务或进程之间，作业或程序之间。
2. 从处理数据的角度：位串字串，位并字串，位片串字并，全并行。
3. 从信息加工步骤与阶段的角度：存储器操作并行，处理器操作步骤并行，处理器操作并行，指令、任务、作业并行。

并行性开发的途径：时间重叠，资源重复，资源共享。



3T 性能目标 — 1TFLOPS 计算能力、1TB 主存容量、1TBps I/O 带宽。

并行处理计算机结构：流水线计算机、阵列处理机、多处理机、数据流计算机。

多机系统：多处理机系统、多计算机系统。

多处理机系统：同构性多处理机系统、异构型多处理机系统。

多机系统的耦合度：反映多机系统中各机器之间物理连接的紧密度和交叉作用能力的强弱，可分为最低耦合、松散耦合、紧密耦合。

六、弗林分类法

1966 年，弗林提出按指令流和数据流的多倍性分类。

多倍性是指在系统性能瓶颈部件上处于同一执行阶段的指令或数据的最大可能个数。分为 SISD（单指令流单数据流）、SIMD（单指令流多数据流）、MISD（多指令流单数据流）、MIMD 四大类（多指令流多数据流）。

【练习】计算机系统多级层次中，从下层到上层，各级相对顺序正确的应当是(B)

- A.汇编语言机器级---操作系统机器级---高级语言机器级
- B.微程序机器级---传统机器语言机器级---汇编语言机器级
- C.传统机器语言机器级---高级语言机器级---汇编语言机器级
- D.汇编语言机器级---应用语言机器级---高级语言机器级

2. 数据标识、寻址方式与指令系统

一、数据表示

指计算机硬件能够直接识别，可以被指令系统直接调用的那些数据类型。(由硬件实现的数据类型)

同数据结构的关系：数据结构和数据表示是软、硬件的界面。

同运算器及运算指令的关系：决定了运算器的结构及运算指令的设置。

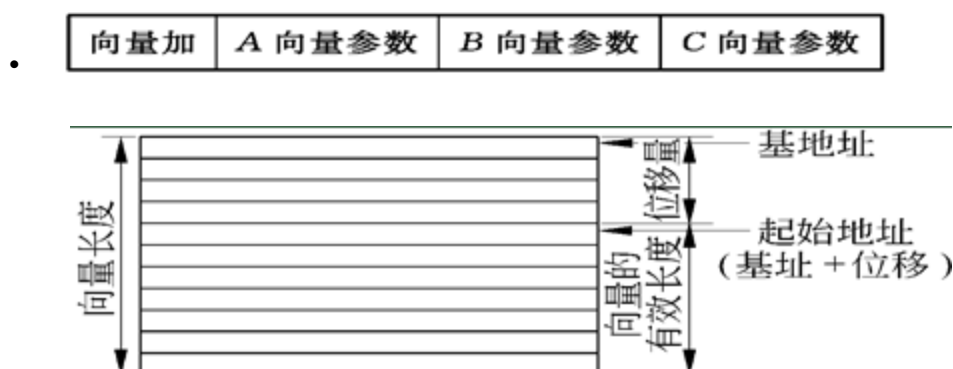
简单数据表示：定点数（小数点固定）、二一十进制数、变址操作、串操作等。

二、确定数据表示的原则

- 1.缩短程序的运行时间
- 2.减少 CPU 与主存储器之间的通信量
- 3.这种数据表示的通用性和利用率

①向量数组数据表示

- 有向量数据表示的处理机就是向量处理机。引入向量数据表示的主要目的是为了便于实现向量运算的并行性。



②堆栈数据表示

有堆栈数据表示的机器就称为堆栈机器。

堆栈机器与通用寄存器型机器的主要区别在于堆栈指令、堆栈存储的位置等。

引入数据表示的原则：效率、通用性、利用率。



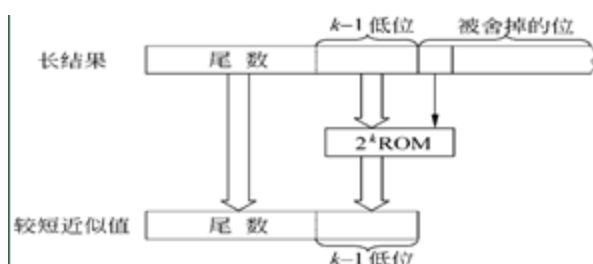
③浮点数表示

浮点数表示的值： r 阶值， X 尾数值；

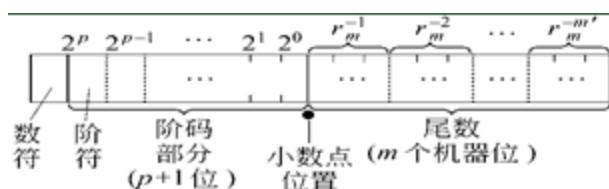
阶码决定了数的范围，尾数决定了数的精度；

当尾数右移一个 r 进制数位时，为保持数不变，阶码增 1；

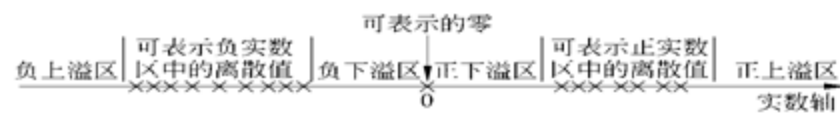
浮点数尾数的下溢处理方法：截断法、舍入法、恒置 1 法、查表舍入法。



浮点数的一般格式



浮点数的特性：只能表示出数轴上分散于正、负两个区间上的离散值；



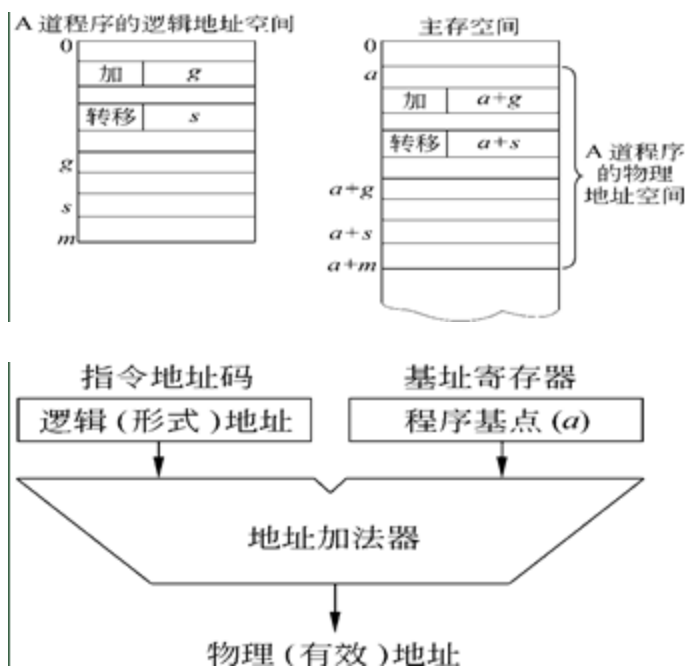
三、寻址方式

指指令按什么方式寻找（或访问）所需的操作数。

指明方式：占用操作码的某些位；在地址码部分设置专门的寻址方式位字段。

程序的定位技术：逻辑地址、物理地址。

逻辑地址空间至物理地址空间的变换：静态重定位、动态重定位。

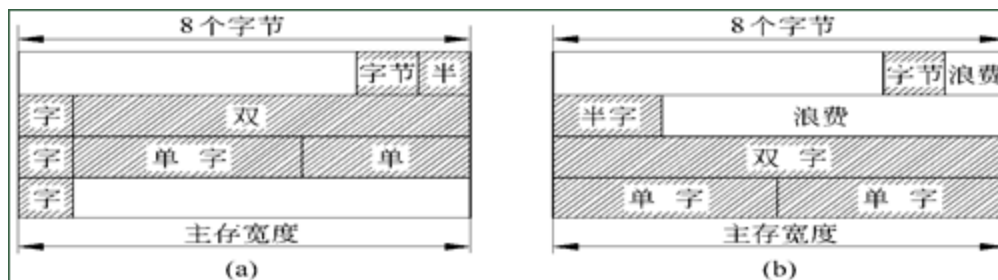


按整数边界存储:

字长

主存宽度 (数据通路宽度) (IBM370 为 64 位)

最小寻址单位: 最小数据访问块大小



四、指令系统的设计和优化

指令系统的设计主要包括功能、格式方面的设计。

指令操作码的优化: 信息源的熵、平均码长、信息冗余量、Huffman 编码。

(1) 指令系统的功能设计

- 完整性是指应该具备的基本指令种类, 通用计算机必须有 5 类基本指令:
 - 数据传送类指令、运算类指令、程序控制指令、输入输出指令、处理机控制和调试指令
- 规整性包括对称性和均匀性
 - 对称性: 所有寄存器同等对待, 操作码的设置等都要对称
 - 均匀性: 不同的数据类型、字长、存储设备、操作种类要设置相同的指令

- 高效率：指令的执行速度要快；指令的使用频度要高；各类指令之间要有一定的比例
- 兼容性：在同一系列机内指令系统不变(可以适当增加)

(2) 计算机指令集结构分类

- 根据五个因素对计算机指令集结构进行分类：
- 在 CPU 中操作数的存储方法；
- 指令中显式表示的操作数个数；
- 操作数的寻址方式；
- 指令集所提供的操作类型；
- 操作数的类型和大小。

(3) 指令集结构分为

- 堆栈型指令集结构
 - 指令短小。不能随机访问堆栈，从而很难生成有效代码；同时，由于堆栈是瓶颈，所以很难被高效地实现。
- 累加器型指令集结构
 - 减少了机器的内部状态；指令短小。由于累加器是唯一的暂存器，这种机器的存储器通信开销最大。
- 通用寄存器型指令集结构
 - 易于生成高效的目标代码。所有操作数均需命名，且要显式表示，因而指令比较长。现代大多数机器均采用通用寄存器型指令集结构，
 - 原因一是寄存器和 CPU 内部其他存储单元一样，要比存储器快；其次是对编译器而言，可以更加容易、有效地分配和使用寄存器。

(4) 指令系统的优化设计

指令系统的优化设计有两个截然相反的方向：

- 1. 复杂指令系统计算机 CISC
 - (Complex Instruction Set Computer)
 - 增强指令功能，设置功能复杂的指令
 - 面向目标代码、高级语言和操作系统
 - 用一条指令代替一串指令
- 2. 精简指令系统计算机 RISC
 - (Reduced Instruction Set Computer)

- 只保留功能简单的指令
- 功能较复杂的指令用子程序来实现

(5) 指令的组成



- 一般的指令主要由两部分组成：
 - 操作码和地址码
- 地址码通常包括三部分内容：
 - 地址：地址码、立即数、寄存器、变址寄存器
 - 地址的附加信息：偏移量、块长度、跳距
 - 寻址方式：直接寻址、间接寻址、立即数寻址、变址寻址、相对寻址、寄存器寻址
- 操作码主要包括两部分内容
 - 操作种类：加、减、乘、除、数据传送、移位、转移、输入输出、程序控制、处理机控制等
 - 操作数描述：
 - 数据的类型：定点数、浮点数、复数、字符、字符串、逻辑数、向量
 - 进位制：2 进制、10 进制、16 进制
 - 数据字长：字、半字、双字、字节

【练习】浮点数尾数基值取大，可使浮点数（C）

- A.运算过程中数的精度损失降低
- B.数在数轴上的分布变密
- C.可表示数的范围增大
- D.可表示数的个数增多

【练习】SIMD 是指（B）

- A、单指令流单数据流
- B、单指令流多数据流
- C、多指令流单数据流
- D、多指令流多数据流

【练习】RISC 思想主要是基于什么样的目的?(A)

- A. 减少指令的平均执行周期数

- B. 减少指令的复杂程度
- C. 减少硬件的复杂程度
- D. 便于编译器编写

3. 存储系统

一、存储系统的性能指标

- 1. 存储容量
- 虚拟存储技术为用户设计了一个虚拟地址空间，这个虚拟地址空间比主存的实际地址空间要大得多，并采用像主存一样的随机访问方式。
- 2. 存储系统带
- 存储器被连续访问时能提供的数据传输速率称为存储器的最大带宽。

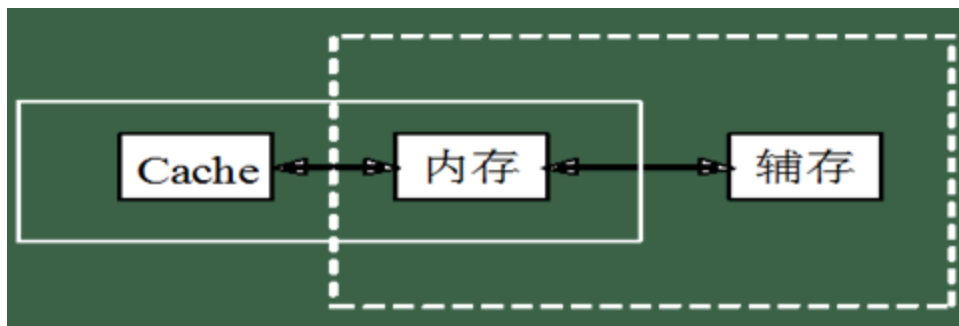
二、并行存储器

- 并行存储器（Parallel Memory）是指在一个存储器访问周期能并行访问到多个存储字的存储器，能有效地提高存储器的带宽。
- 并行存储器主要有 2 种，一种是单体多字并行存储器，另一种是低位交叉编址多体并行存储器。

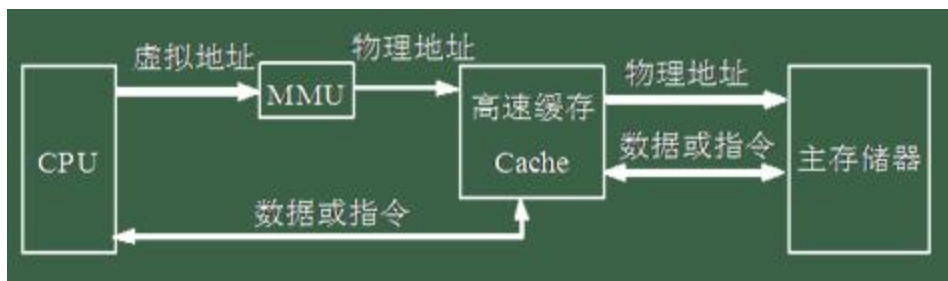
三、虚拟存储器

- 在虚拟存储技术中，把程序经编译生成的访存地址称为虚拟地址（或称为虚地址），由虚地址表示的存储空间称为虚空间。程序代码运行时，必须先把虚地址转换成主存物理地址（或称为主存实地址），才能按实地址访问主存。虚地址与实地址之间对应关系的规则称为地址映像（Address Mapping）。程序在运行时，虚拟存储系统按照某种地址映像把虚地址转换成实地址称为地址变换（Address Translation）。地址变换对应用程序员是透明的，由存储系统自动完成。
- 如果按虚地址要访问的数据不在主存中（未命中），则称为发生访问失效，那么就需要访问磁盘存储器。这时，先要把虚地址变换成磁盘存储器物理地址，称之为外部地址变换，然后才能访问磁盘存储器，将要访问的数据块调入主存。外部地址变换更多地依靠软件来实现。
- 当有新的数据块要调入主存时，如果主存中已经没有空闲的位置，则称为发生实存冲突，那么就要采用某种替换算法来确定新数据块调入主存的位置。

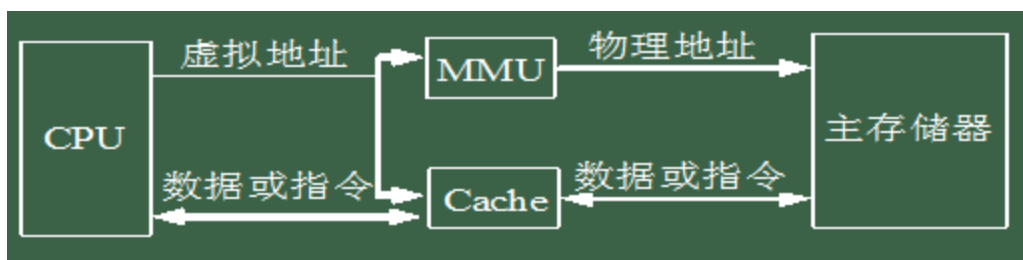
四、三级存储系统



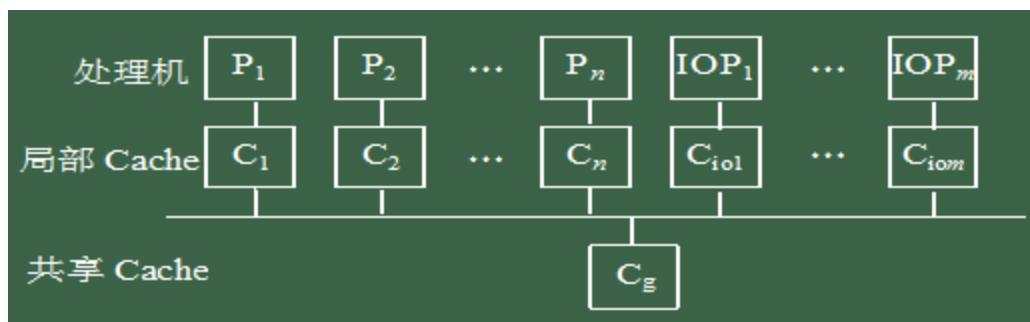
- 虚拟存储器（Virtual Memory）是针对主存容量不能满足要求而提出的。
- Cache 存储器（Cache Memory）是针对主存速度不能满足要求而提出的。
- 从程序员来看，不管存储系统的组织多么复杂，都只看到一个存储器，这个存储器采用与主存相同的按地址随机访问的方式工作，它的等效访问速度接近于 Cache 的访问速度，等效存储容量是虚拟地址空间的容量。
- Cache、主存、磁盘这 3 个物理存储器组织成一个三级存储系统的方式可以分别构成“Cache-主存”和“主存-磁盘”两个存储系统，也可以构成一个“Cache-主存-磁盘”存储系统。
- 两个存储系统的组织方式
- 这种组织方式组织成“Cache-主存”和“主存-磁盘”两个独立的存储系统。这种三级存储系统的结构也称为物理地址 Cache 存储系统。



- 一个存储系统的组织方式
- 这种组织方式是把 Cache、主存和磁盘 3 个存储器组织成一个“Cache-主存-磁盘”存储系统。这种三级存储系统的结构也称为虚拟地址 Cache 存储系统。Cache 能够直接接受虚拟地址的访问，找出 CPU 所需要的数据或指令。



- 全 Cache 存储系统的组织方式
- 全 Cache 存储系统的组织方式中，没有主存储器，只用 Cache 和磁盘两种存储器构成“Cache-磁盘”存储系统。



【练习】与虚拟存储器的等效访问速度无关的是（D）。

- A. 访存页地址流
- B. 页面替换算法
- C. 主存的容量
- D. 辅存的容量

4. 总线、中断与输入输出系统

一、输入输出系统概述

经历了 3 个阶段对应着 3 种方式：

- 程序控制 I/O（程序查询、中断驱动）
- 直接存储器访问（DMA）
- I/O 处理机方式（通道、外围处理机 PPU）

输入输出设备分外存和传输设备两大类。

总线分类

按信息传送方向：半双工、全双工；

按位置：芯片级、板级（局部总线）、系统级；

按用法：专用总线、非专用总线。

（一）基本输入输出方式

对于工作速度、工作方式和性质不同的外围设备，基本输入输出方式有如下三种。

1 程序控制输入输出方式

又称为状态驱动输入输出方式、应答输入输出方式、查询输入输出方式、条件驱动输入输出方式等

程序控制输入输出方式的 4 个特点：

- (1) 何时、 对何设备进行输入或输出操作受 CPU 控制
- (2) CPU 要通过指令对设备进行测试才能知道设备的工作状态。空闲、准备就绪、正在忙碌等
- (3) 数据的输入和输出都要经过 CPU
- (4) 用于连接低速外围设备，如终端、打印机等

(二) 中断输入输出方式

中断输入输出方式的定义如下：

当出现来自系统外部，机器内部，甚至处理机本身的任何例外的，或者虽然是事先安排的，但出现在现行程序的什么地方是事先不知道的事件时，

CPU 暂停执行现行程序，转去处理这些事件，等处理完成后再返回来继续执行原先的程序。

中断输入输出方式的特点：

- (1) CPU 与外围设备能够并行工作
- (2) 能够处理例外事件。例如，电源掉电、非法指令、地址越界、数据溢出、数据校验错、页面失效等
- (3) 数据的输入和输出都要经过 CPU
- (4) 灵活性好
- (5) 用于连接低速外围设备

(三) 直接存储器访问(DMA)方式

直接存储器访问方式又称为 DMA(Direct Memory Access)方式，这种输入输出方式主要用来连接高速外围设备。

- 如，磁盘存储器，磁带存储器等

DMA 方式具有如下特点：

- (1) 外围设备访问请求直接发往主存储器
- (2) 不需要 CPU 进行保存现场和恢复现场
- (3) DMA 控制器中，需设置数据寄存器、设备状态或控制寄存器、主存地址寄存器、设备地址寄存器和数据交换个数计数器
- (4) 在 DMA 方式开始和结束时，需要处理机进行管理
- (5) 在 DMA 方式中，数据的传送过程不需要 CPU 的干预

二、总线的控制方式

集中式控制

串行链接

定时查询

独立请求

分布式控制

(1) 总线的通信技术

同步通信

异步通信：

单向控制通信（源控、目控）

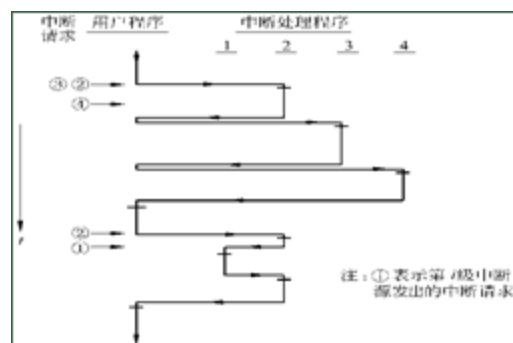
双向控制通信（非互锁、互锁）

三、中断系统

中断的分类和分级。

中断的响应优先级、处理次序：

响应优先级由硬件排队器固定，处理次序可通过屏蔽字灵活改变。



中断处理 程序级别	中断级屏蔽位			
	1级	2级	3级	4级
第1级	0	0	0	0
第2级	1	0	0	0
第3级	1	1	0	0
第4级	1	1	1	0

(1) 中断系统的软硬功能分配

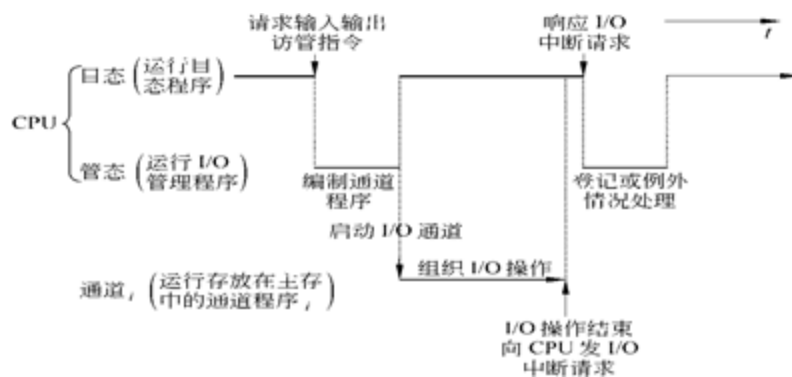
指中断处理程序软件和中断响应硬件的功能分配。

中断系统的功能包括中断请求的保存和清除、优先级的确定、中断断点及现场的保存、对中断请求的分析和处理以及终端返回等。

四、通道处理机

通道处理机是 IBM 首先提出来的一种 I/O 处理机方式。

通道处理机的输入输出过程：



(1) 通道流量分析

通道流量—单位时间内通道传送的字节数。极限流量、实际最大流量。

I/O 系统的极限流量、实际最大流量—所有通道流量之和。

极限流量 \geq 实际最大流量

通道的响应优先级

【练习】若某个计算机系统中 I/O 地址统一编址，访问内存单元和 I/O 设备是靠 B 来区分的。

- A.数据总线上输出的数据
- B.不同的地址代码
- C.内存与 I/O 设备使用不同的地址总线
- D.不同的指令

5、标量处理机

一、标量处理机

标量处理机--只有标量数据表示和标量指令系统的处理机

提高指令执行速度(处理机设计主要任务之一)的主要途径

- 提高处理机的工作主频
- 采用更好的算法和设计更好的功能部件
- 采用指令级并行技术

三种指令级并行处理机

- 流水线处理机和超流水线(Super-pipelining)处理机
- 超标量(Superscalar)处理机
- 超长指令字(VLIW: Very Long Instruction Word)处理机

三种基本技术

- 先行控制技术
- 流水线技术
- 相关性分析技术
- 动态调度技术

二、数据相关

数据相关：在执行本条指令的过程中

- 如果用到的指令、操作数、变址量等正好是前面指令的执行结果，
- 则必须等待前面的指令执行完成，
- 并把结果写到主存或寄存器中之后，
- 本条指令才能开始执行。这种相关称为数据相关。

数据相关有四种：指令相关、主存操作数相关、通用寄存器相关和变址相关

解决数据相关的方法有两种：推后分析法和设置专用路径法。

指令相关

发生指令相关的情况：

- n: STORE R1, n+1
- n+1:

满足关系： 结果地址(n)=指令地址(n+1)

当第 n 条指令还没有把执行结果写到主存之前，取出的第 n+1 条指令显然是错误的。

控制相关

控制相关：由条件分支指令、转子程序指令、中断等引起的相关，因为程序的执行方向可能被改变而引起的相关

- 可能改变程序执行方向的指令通常有
 无条件转移
- 一般条件转移
- 复合条件转移
- 子程序调用
- 中断

数据相关称为局部相关，影响到的仅仅是本条指令附近的少数几条指令

控制相关称为全局相关，影响的范围要大得多，它可能引起程序执行方向的改变

三、转移预测技术

在程序执行过程中，转移预测的方向不能改变的称为静态分支预测。

在程序执行过程中能够动态地改变转移预测方向的称为动态分支预测。

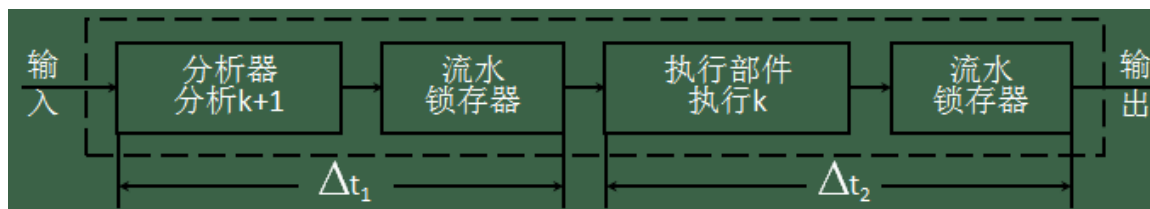
四、流水线处理机

两个方面来开发处理机的并行性

- 空间并行性
 - 设置多个独立的操作部件
 - 多操作部件处理机、超标量处理机
- 时间并行性
 - 采用流水线技术。
 - 不增加或只增加少量硬件就能使运算速度提高几倍
 - 流水线处理机、超流水线处理机

(1) 流水线工作原理

从重叠到流水线



流水线的每一个阶段称为流水节拍、流水步、流水步骤、流水阶段、流水线阶段、功能段、流水流水段、流水段、流水级等等。

在每一个流水段的末尾或开头必须设置一个寄存器

- 称为流水寄存器、流水锁存器、流水闸门寄存器等
- 会增加指令的执行时间。
- 锁存器作用：保存分析或执行结果

为了简化，在一般流水线中不画出流水锁存器

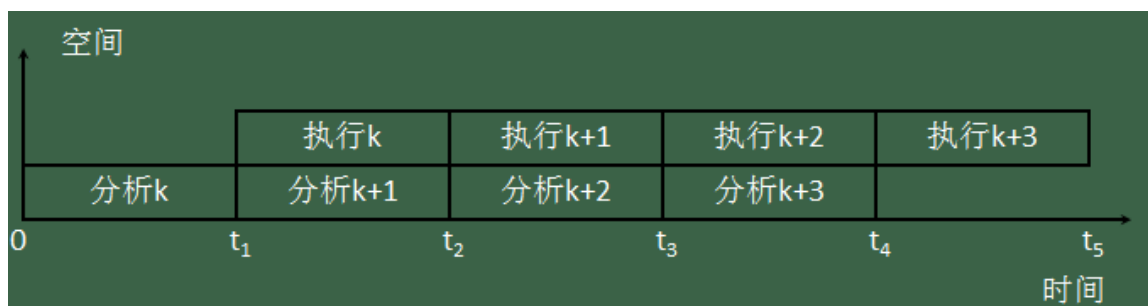
指令分析与执行并行

- 输出端每 Δt 输出一条指令

• 时空图

流水线的表示方法通常有三种：连接图、时空图和预约表。

一条简单流水线的时空图，如下所示



(2) 单功能流水线与多功能流水线

单功能流水线：

- 只能完成一种固定功能的流水线
- Cray-1 计算机中有 12 条；YH-1 计算机有 18 条；Pentium 有一条 5 段的定点和一条 8 段的浮点流水线；PentiumIII 有三条指令流水线，其中两条定点指令流水线，一条浮点指令流水线。

多功能流水线：

- 流水线的各段通过不同连接实现不同功能
- Texas 公司的 ASC 计算机中的 8 段流水线，能够实现：定点加减法、定点乘法、浮点加法、浮点乘法、逻辑运算、移位操作、数据转换、向量运算等。

(3) 静态流水线与动态流水线

静态流水线：

- 同一段时间内，多功能流水线中的各个功能段只能按照一种固定的方式连接，实现一种固定的功能。
- 只有连续出现同一种运算时，流水线的效率才能得到充分的发挥。

动态流水线：

- 在同一段时间内，多功能流水线中的各段可以按照不同的方式连接，同时执行多种功能。

(4) 线性流水线的调度

线性流水：每一任务在流水中每一段流且仅流过一次

- 调度简单：每一时钟周期注入一个任务
- 流水线充满后，每一时钟周期输出一个任务

(5) 非线性流水线的调度

非线性流水：存在反馈回路

- 流水冲突（功能性部件冲突）

- 反馈回路→任务可能多次流过同一功能段
- 每一时钟周期注入一个任务→多个任务竞争同一功能段
- 冲突解决方法：延迟输入新任务
- 非线性流水线调度
 - 间隔周期变化：找出一个最小的循环周期
 - 按照这周期向流水线输入新任务
 - 流水线的各个功能段都不会发生冲突
 - 流水线的吞吐率和效率最高

【练习】利用时间重叠概念实现并行处理的是（ A ）

- A.流水处理
- B.多处理机
- C.并行（阵列）处理机
- D.相联处理机

【练习】静态流水线是指（ C ）

- A.只有一种功能的流水线
- B.功能不能改变的流水线
- C.同时只能完成一种功能的多功能流水线
- D.可同时执行多种功能的流水线

6. 向量处理机

一、什么是向量处理

从标量到向量

- 如一个简单的 C 语言程序
for (i = 10; i <= 1010; i++)
c[i] = a[i] + b[i+5];
- 在向量处理机上，可以只用一条指令：
C(10:1010)=A(10:1010) + B(15 :1015)
- 一条向量指令可处理 N 个或 N 对操作数
- 在标量处理机上用 10 多条机器指令，其中有 8 条指令要循环 1000 次。
- 采用多寄存器结构的两地址指令编写程序。
- 存储器采用字节编址方式，字长为 32 位

$$\begin{pmatrix} A1 \\ A2 \\ A3 \\ A4 \end{pmatrix} + \begin{pmatrix} B1 \\ B2 \\ B3 \\ B4 \end{pmatrix} \Rightarrow \begin{pmatrix} C1 \\ C2 \\ C3 \\ C4 \end{pmatrix}$$

二、向量处理方式

要根据向量运算的特点和向量处理机的类型选择向量的处理方式。

有三种处理方式：

- 1. 横向处理方式，又称为水平处理方式，横向加工方式等
 - 向量计算是按行的方式从左至右横向地进行。
- 2. 纵向处理方式，又称为垂直处理方式，纵向加工方式等
 - 向量计算是按列的方式自上而下纵向地进行。
- 3. 纵横处理方式，又称为分组处理方式，纵横向加工方式等
 - 横向处理和纵向处理相结合的方式。

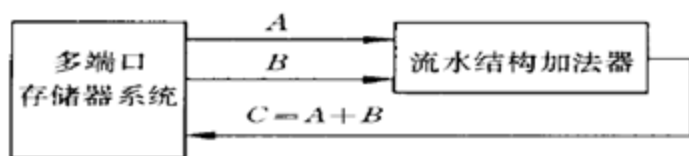
向量横向处理是向量的处理方式，但不是向量的流水处理方式，而向量纵向处理和分组建横处理是向量的处理方式，也是向量的流水处理方式

三、向量处理机的结构

向量处理机基本思想

- 两个向量的对应向量分别运算，产生结果向量
 - $C=A+B$, A 、 B 、 C 均为 n 维向量

可用流水线实现



- 存储器每个时钟周期提供 A 和 B 的一个元素到相应数据通路
- 每个时钟周期产生一个结果

向量处理机存储系统设计难点

- 提供运算器连续不断的数据流、接收来自运算器连续的结果
- 一个时钟周期内完成读两个操作数和写一个运算结果
- 一般随机存储器一个时钟周期只能完成一个操作

四、向量流水处理机的结构

- 向量处理机的指令系统：
- 分类： $V = OP\ V$ 、 $S = OP\ V$ 、 $V = V\ OP\ V$ 、 $V = S\ OP\ V$ 等。
- 其中 V 代表向量， S 代表标量， OP 代表操作。
- 另外还有：比较、压缩、归并、传送等特殊操作的向量指令。
- 指令格式：



五、向量链接技术(chaining)

- 结果寄存器可能成为后继指令的操作数寄存器
- 两条有数据相关的向量指令并行执行，这种技术称为两条流水线的链接技术。

六、向量循环或分段开采技术

当向量的长度大于向量寄存器的长度时，必须把长向量分成长度固定的段，采用循环结构处理这个长向量，这种技术称为向量循环开采技术，也称为向量分段开采技术。

【练习】为了提高向量处理机的性能，会采用一些特殊的技术，下面哪一项不是提高向量处理机性能而采用的技术?(C)

- A. 链接技术
- B. 向量递归技术
- C. Cache 缓存技术
- D. 稀疏矩阵的处理技术

■ 华图网校介绍

华图网校（V.HUATU.COM）于2007年3月由华图教育投资创立，是华图教育旗下的远程教育高端品牌。她专注于公职培训，目前拥有遍及全国各地500万注册用户，已成为公职类考生学习提高的专业门户网站。

华图网校是教育部中国远程教育理事单位。她拥有全球最尖端高清录播互动技术和国际领先的网络课程设计思想，融汇华图教育十余年公职辅导模块教学法，凭借强大师资力量与教学资源、利用教育与互联网的完美结合，真正为考生带来“乐享品质”的学习体验，通过“高效学习”成就品质人生。

华图网校课程丰富多元，涵盖公务员、事业单位、招警、法院、检察院、军转干、选调生、村官、政法干警、三支一扶、乡镇公务员、党政公选等热门考试、晋升及选拔。同时，华图网校坚持以人为本的原则，不断吸引清华、北大等高端人才加入经营管理，优化课程学习平台，提升用户体验，探索网络教育新技术和教学思想，力争为考生提供高效、个性、互动、智能的高品质课程和服务。

华图网校将秉承“以教育推动社会进步”的使命，加快网站国际化进程，打造全球一流的网络学习平台。

我们的使命：以教育推动社会进步

我们的愿景：德聚最优秀人才，仁就基业长青的教育机构

我们的价值观：诚信为根、质量为本、知难而进、开拓创新。

- 咨询电话：400-678-1009
- 听课网址：v.huatu.com（华图网校）