

# 计算机 体系结构



**计算机体系结构**是计算机相关专业本科生必修的专业课。也称为**计算机系统结构**。

本课程先修的课程：**计算机组成原理**。



本课程是一门从计算机组成和结构的角度上学习和领会计算机系统的课程。

计算机系统是由软件和硬件组合的一个复杂的综合体。本课程主要研究如何对计算机系统软件和硬件的功能进行更合理的分配；研究如何更好、更合理地实现分配给硬件的那部分功能，使系统有尽可能高的性能价格比。



本课程学习的内容是计算机系统结构设计、硬件设计、高层次应用系统开发和系统软件开发所必须了解和掌握的基本知识。

通过本课程的学习，能进一步树立计算机系统的整体概念，熟悉有关计算机系统结构的概念、原理，了解常用的基本结构，领会结构设计思想和方法，提高分析和解决问题的能力。



## 本课程的地位和重要性

1. 计算机系统结构是计算机科学与技术一级学科中的三个二级学科之一。

2. 计算机系统结构是国家同等学历申请硕士学位考试科目之一。

3. 硕士研究生入学考试“计算机组成原理”部分的试题中有体系结构方面的内容。



## 教材与参考书

1. 蒋本珊、马忠梅、郑宏. 计算机体系结构简明教程. 北京: 清华大学出版社, **2015**
2. 郑纬民、汤志忠. 计算机系统结构. 北京: 清华大学出版社, **2001.2**
3. 李学干. 计算机系统的体系结构 (第**5**版). 西安: 西安电子科技大学出版社, **2011**
4. **John L. Hennessy**等. 计算机系统结构——量化研究方法. 北京: 机械工业出版社 (英文版), 电子工业出版社 (中文版), **2004**
5. 张晨曦. 计算机系统结构教程. 北京: 清华大学出版社, **2009**
6. 李学干. 《计算机系统结构》学习指导与题解. 西安: 西安电子科技大学出版社, **2001**



# 第1章 计算机系统结构的基本概念

1.1 计算机系统的多级层次结构

1.2 计算机系统结构、组成与实现

1.3 软硬取舍与计算机系统的设计思路

1.4 计算机设计的量化准则

1.5 对系统结构的影响因素

1.6 系统结构中的并行性

1.7 计算机系统分类



## 1.1 计算机系统的多级层次结构

- 计算机系统 = 软件 + 硬件 / 固件
- 可以从多个角度考察计算机系统的结构
- 一种观点：从使用语言的角度，可以将计算机系统按功能划分为多级层次结构





## 1.1 计算机系统的多级层次结构

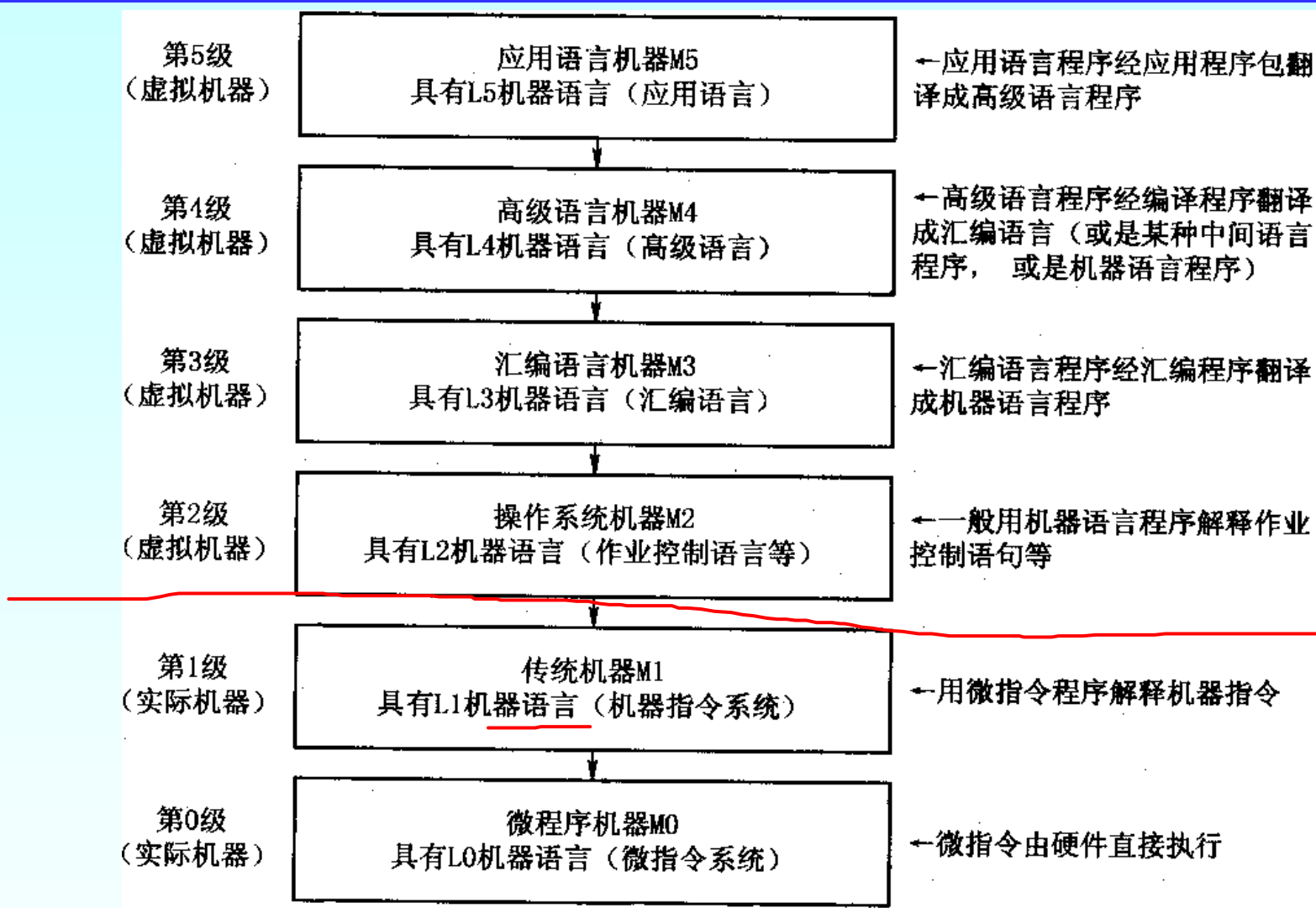


图 1.1 计算机系统的多级层次结构  
北京理工大学计算机学院



## 1.1 计算机系统的多级层次结构

### 从学科领域来划分

- 第**0**和第**1**级属于计算机组成与系统结构
- 第**2**至第**4**级是系统软件
- 第**5**级是应用软件

它们之间仍有交叉

- 第**1**级涉及汇编语言程序设计的内容
- 第**2**级与计算机系统结构密切相关

在特殊的计算机系统中，有些级别可能不存在。



## 1.1 计算机系统的多级层次结构

### 虚拟机概念

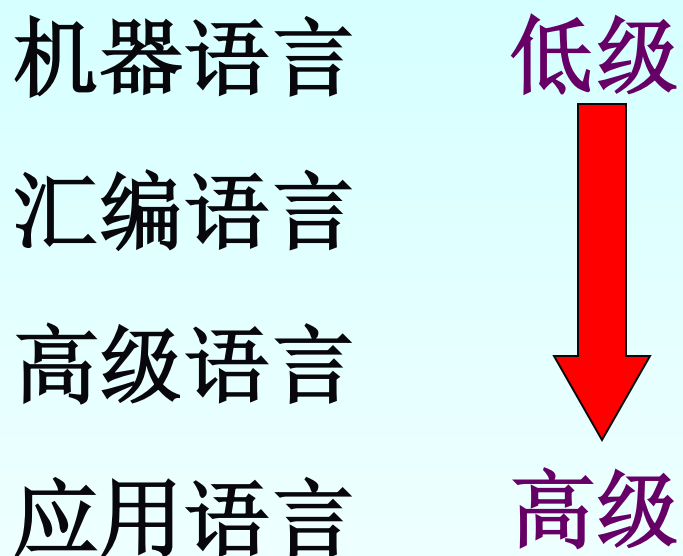
在计算机系统的多层次结构中，除第**0**、**1**级外，上面四级均为虚拟机。

虚拟计算机是指这个计算机只对该级的观察者存在。对某一层次的观察者来说，它只能是通过该层次的语言来了解和使用计算机，至于下层是如何工作和实现的就不必关心了。简而言之，**虚拟机即是由软件实现的机器。**



## 1.1 计算机系统的多级层次结构

从使用语言的角度上，将计算机系统看成按功能划分的多级层次结构。



特点：后者以前者为基础；比前者功能强、使用更方便。



## 1.1 计算机系统的多级层次结构

计算机只能直接识别和执行机器语言。

汇编语言是一种符号式程序设计语言。

可以想象在使用机器指令的实际机器上出现了用汇编语言作为机器语言的“**虚拟**”机器。



## 1.1 计算机系统的多级层次结构

**翻译(Translation)**：先用转换程序将高一级机器级上的程序整个地变换成低一级机器级上可运行的等效程序，然后再在低一级机器级上去实现的技术。（**先翻译后执行**）

例：英语翻译

**解释(Interpretation)**：在低一级机器级上用它的一串语句或指令来等效高一级机器上的一条语句或指令的功能，通过对高一级机器语言程序中的每条语句或指令逐条解释来实现的技术。（**边解释边执行**）

例：解释一件事



## 1.1 计算机系统的多级层次结构

翻译和解释是语言实现的两种基本技术。一般来说，解释执行比翻译花的时间多，但占用存储空间较少。

在多层次结构中，通常第**1**、**2**级是用解释方法实现的，而第**3**级或更高级则用翻译方法实现。



## 1.1 计算机系统的多级层次结构

计算机系统结构主要研究软硬件功能分配和对软硬件界面的确定。

计算机系统由软件、硬件和固件组成，它们在功能上是同等的。

同一种功能可以用硬件实现，也可以用软件或固件实现。





## 1.1 计算机系统的多级层次结构

**固件(Firmware)**是指那些存储在能永久保存信息的器件（如**ROM**）中的程序，是**具有软件功能的硬件**。固件的性能指标介于硬件与软件之间，吸收了软、硬件各自的优点，其执行速度快于软件，灵活性优于硬件，是软、硬件结合的产物，计算机功能的固件化将成为计算机发展中的一个趋势。



## 1.1 计算机系统的多级层次结构

**软件和硬件实现在逻辑功能上等效。**

计算机系统结构设计者的主要任务就是要确定软硬件的分界；软件、硬件和固件的功能分配。

### 软件与硬件实现的特点

硬件实现：速度快、成本高；灵活性差、占用内存少。

软件实现：速度低、复制费用低；灵活性好、占用内存多。



## 1.1 计算机系统的多级层次结构

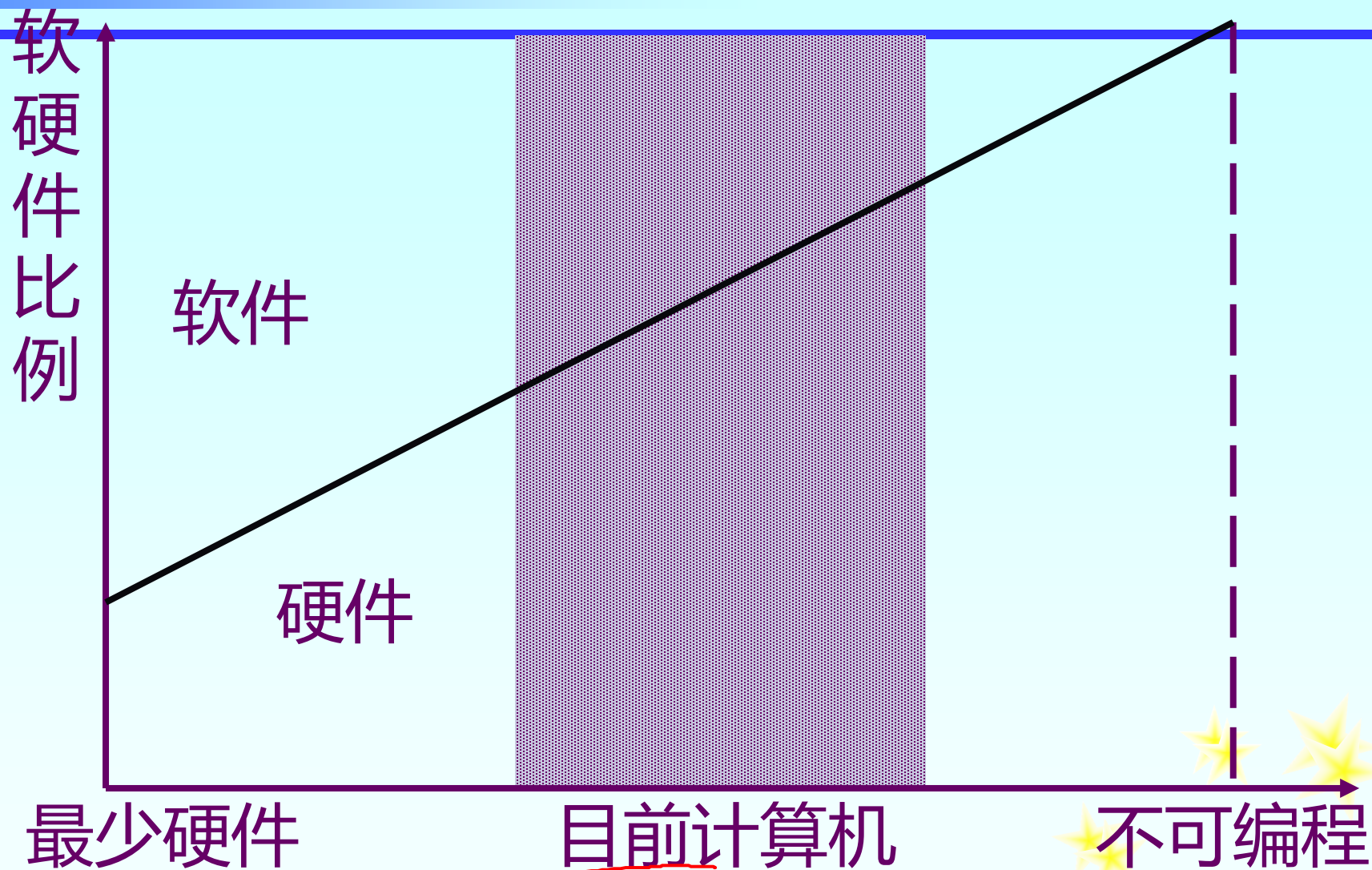


图 1.2 计算机系统的软、硬件功能分配



## 1.2 计算机系统结构、组成与实现

### 1.2.1 系统结构、组成与实现

我们这里所称的计算机系统结构或计算机体系结构(Computer Architecture) 指的是层次结构中传统机器级的系统结构，其界面之上的功能包括操作系统级、汇编语言级、高级语言级和应用语言级中所有软件的功能。界面之下的功能包括所有硬件和固件的功能。



## 1.2 计算机系统结构、组成与实现

从计算机系统的层次结构定义：

计算机系统结构是对计算机系统中各级界面的划分、定义及其上下的功能分配。



## 1.2 计算机系统结构、组成与实现

### 计算机系统结构定义

**Amdahl**于**1964**年在推出**IBM360**系列计算机时提出：**程序员所看到的一个计算机系统的属性，即概念性结构和功能特性。**

所谓概念性结构与功能特性，实际上就是计算机系统的外特性。

**程序员：**机器语言、操作系统、汇编语言、编译程序的程序员。

**看到的：**编写出能够在机器上正确运行的程序所必须了解到的东西。



## 1.2 计算机系统结构、组成与实现

对于这一定义，计算机界是有争议的，主要争议点基于这样一个事实，即由于计算机系统是包括软、硬件乃至固件资源的较复杂系统，因此处于不同级别的使用者（各级程序员）所看到的计算机具有不同的属性。例如，用高级语言编程的程序员，可以把**IBM PC**与**RS6000**两种机器看成是同一属性的机器；但对使用汇编语言编程的程序员来说，**IBM PC**与**RS6000**是两种截然不同的机器。



## 1.2 计算机系统结构、组成与实现

事实上，**Amdahl**提出的系统结构定义中的程序员指的是传统机器语言程序员。他们所看到的计算机属性，是硬件子系统的概念性结构和功能特性。包括：

(1)硬件能直接识别和处理的数据类型和格式等的**数据表示**；

(2)最小可寻址单位、寻址种类、地址计算等的**寻址方式**；

(3)通用/专用寄存器的设置、数量、字长、使用约定等的**寄存器组织**；





## 1.2 计算机系统结构、组成与实现

**(4)**二进制或汇编级指令的操作类型、格式、排序方式、控制机构等的**指令系统**；

**(5)**内存的最小编址单位、编址方式、容量、最大可编址空间等的**存储系统组织**；

**(6)**中断的分类与分级、中断处理程序功能及入口地址等的**中断机构**；

**(7)**系统机器级的**管态和用户态的定义和切换**；

**(8)**输入输出设备的连接、使用方式、流量、操作结束、出错指示等的**机器级I/O结构**；

**(9)**系统各部分的**信息保护方式和保护机构**。



## 1.2 计算机系统结构、组成与实现

计算机组成(Computer Organization)指的是计算机系统结构的逻辑实现，包括机器级内的数据流和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。计算机组成设计要解决的问题是在所希望达到的性能和价格下，怎样最佳、最合理地把各种设备和部件组织成计算机，以实现所确定的系统结构。



## 1.2 计算机系统结构、组成与实现

计算机组成设计要确定的方面一般应包括：

- (1)数据通路宽度
- (2)专用部件的设置
- (3)各种操作对部件的共享程度
- (4)功能部件的并行度
- (5)控制机构的组成方式
- (6)缓冲和排队技术
- (7)预估、预判技术



## 1.2 计算机系统结构、组成与实现

计算机实现是指计算机组成的物理实现。它主要着眼于器件技术和微组装技术。

计算机组成和实现都属于计算机系统的内特性，这些特性对程序员来说是透明的（即程序员是看不到的）。



## 1.2 计算机系统结构、组成与实现

**透明性概念：本来存在的事物或属性，从某种角度看似乎不存在。**

这与日常生活中的“透明”的含义正好相反。日常生活中的“透明”是要公开，让大家看得到，而计算机中的“透明”，则是指看不到的意思。

所谓透明实际上就是指那些不属于自己管的部分。对于计算机系统结构而言，前述内容都是不透明的；而全部由硬件实现的，或是在机器语言、汇编语言编程中不会出现和不需要了解的部分都是透明的。



## 1.2 计算机系统结构、组成与实现

例如：浮点数表示、乘法指令，对高级语言程序员、应用程序员透明，对汇编语言程序员、机器语言程序员不透明。

再例如：数据总线宽度、微程序对汇编语言程序员、机器语言程序员透明，对硬件设计者、计算机维修人员不透明。



## 1.2 计算机系统结构、组成与实现

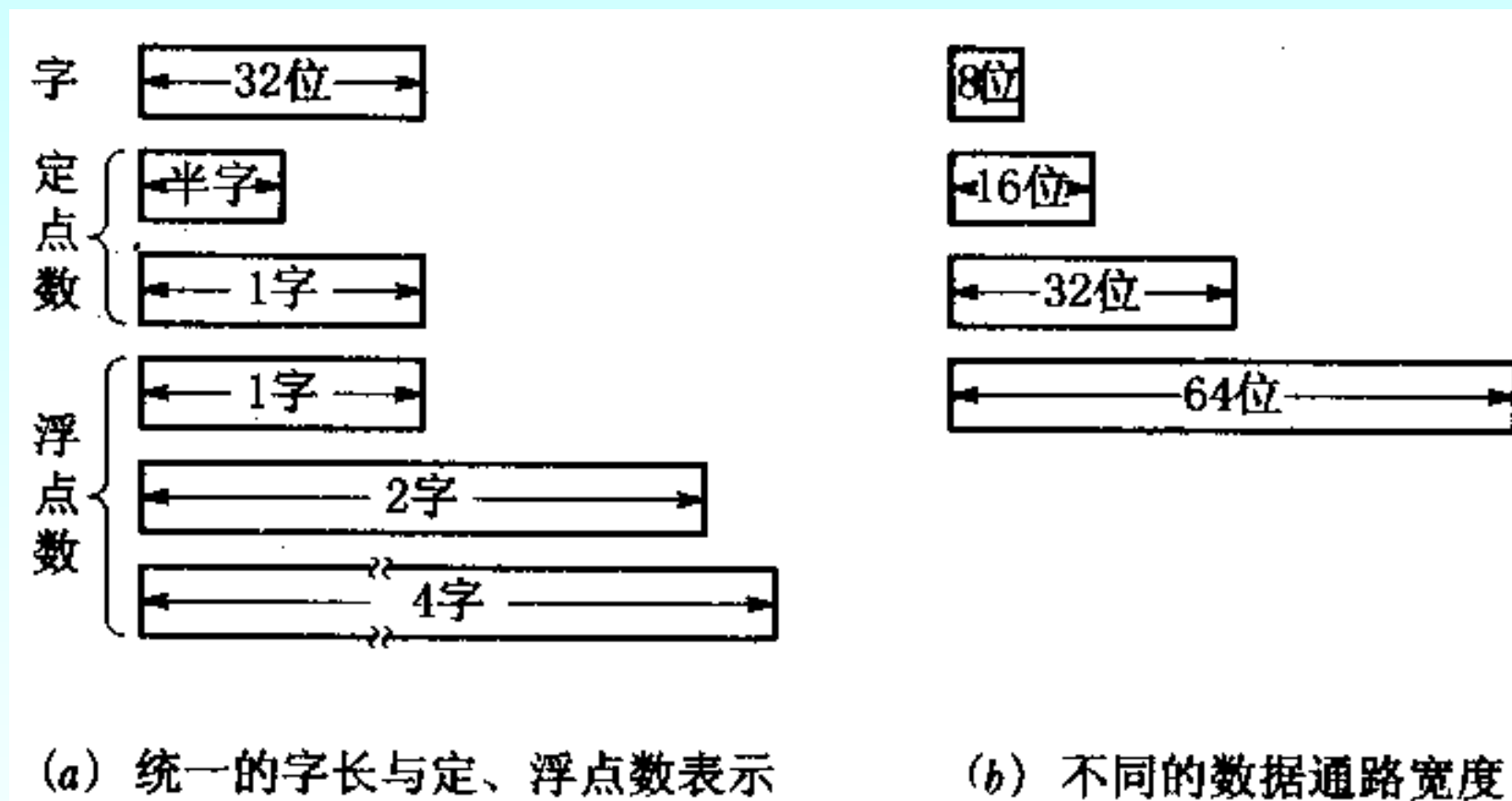


图 IBM 370系列机字长、数的表示和数据通路宽度



## 1.2 计算机系统结构、组成与实现

计算机系统结构、计算机组成和计算机实现是三个不同的概念。同一系统结构可因速度要求不同采用不同组成，一种组成可以采用多种不同的实现。

区分计算机系统结构与计算机组成这两个概念是十分重要的。

指令系统的确定属于计算机系统结构；而指令的实现，如取指令、分析指令、取操作数、运算、送结果等的操作安排和排序属于计算机组成。





## 1.2 计算机系统结构、组成与实现

确定指令系统中是否要设置乘法指令属于计算机系统结构；乘法指令是采用专门的高速乘法器实现，还是靠用加法器和移位器经时序信号控制其相加和移位来实现属于计算机组成。

主存容量与编址方式（按位、按字节还是字编址等）的确定属于计算机系统结构；而为达到性能价格要求，主存速度应选多快，采用何种逻辑结构等则属于计算机组成。



## 1.2 计算机系统结构、组成与实现

计算机制造商可能会向用户提供一系列系统结构相同的计算机，而它们的组成却有相当大的差别，即使是同一系列不同型号的机器，其价格和性能也是有极大差异的。

例如，有**3**台计算机：

**1.没有Cache;**

**2.有单级CPU片外Cache;**

**3. 第三台计算机既有CPU片内Cache，又有片外Cache。**



## 1.2 计算机系统结构、组成与实现

这**3**台计算机具有不同的组成，但它们的系统结构可能是相同的。因此，只知其结构，不知其组成，就选不好性能价格比最合适的机器。此外，一种机器的系统结构可能维持许多年，但机器的组成却会随着计算机技术的发展而不断变化。

如果两台计算机具有不同的计算机组成和相同的计算机系统结构，那么在其中一台计算机上编译后的目标程序，拿到另一台计算机上也能运行，但两者的运行时间可能不同。



## 1.2 计算机系统结构、组成与实现

### 1.2.2 计算机系统结构、组成和实现的相互关系

- 1、系统结构要考虑组成和实现的发展，不要有过多或不合理的限制；**
- 2、组成要考虑系统结构和实现。  
决定于系统结构，受限于实现；**
- 3、组成与实现不是被动的。  
折中权衡；**
- 4、实现是物质基础。**



## 1.2 计算机系统结构、组成与实现

“计算机体系结构”学科=系统结构+组成

研究软硬件功能分配，最佳、最合理地实现分配给硬件的功能。

分为：

从程序设计者看——机器级界面

从计算机设计者看——分配给硬件的功能



## 1.3 软硬取舍与计算机系统的设计思路

### 1.3.1 软硬取舍的基本原则

确定软、硬件功能分配的第一个基本原则是，在现有硬件和器件条件下，系统要有高的性能价格比。

- 性能是一个综合指标，主要包括：
  - 实现费用
  - 速度
  - 其它性能等



## 1.3 软硬取舍与计算机系统的设计思路

- 实现费用 = 设计费用( $D_s + D_h$ ) + 重复生产费( $M_s + M_h$ )
- 硬件设计费用( $D_h$ ) = 软件设计费用( $D_s$ ) \* 100
- 硬件重复生产费( $M_h$ ) = 软件重复生产费( $M_s$ ) \* 100
- 软件设计费用( $D_s$ ) = 软件重复生产费( $M_s$ ) \* 10000

**R**为软件重复出现次数（占用内存、占用介质）

**C**为该功能在软件实现时需重新设计的次数

当台数为**V**时，每台的硬件费用和软件费用



## 1.3 软硬取舍与计算机系统的设计思路

$$D_h/V + M_h < C \times D_s / V + R \times M_s$$

$$100 D_s / V + 100 M_s < C \times D_s / V + R \times M_s$$

$$10^6 / V + 100 < 10^4 \times C / V + R$$

**结论1:** 当R很大时，即经常使用的基本功能适宜用硬件实现。

**结论2:** 当V很大时，即生产台数很多时适宜用硬件实现。





## 1.3 软硬取舍与计算机系统的设计思路

确定软、硬件功能分配的第二个基本原则是，要考虑到准备采用和可能采用的组成技术，使它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。

确定软、硬件功能分配的第三个基本原则是，不能仅从“硬”的角度去考虑如何便于应用组成技术的成果和发挥器件技术的进展，还应从“软”的角度把为编译和操作系统的实现，以至高级语言程序的设计提供更多更好的硬件支持放在首位。



## 1.3 软硬取舍与计算机系统的设计思路

### 1.3.2 计算机系统的设计思路

- 出发点：多级层次结构
- 三种设计思路：

方法1：由上向下（**Top-Down**）

方法2：由下向上（**Bottom-Up**）

方法3：中间开始（**Middle-Out**）



## 1.3 软硬取舍与计算机系统的设计思路

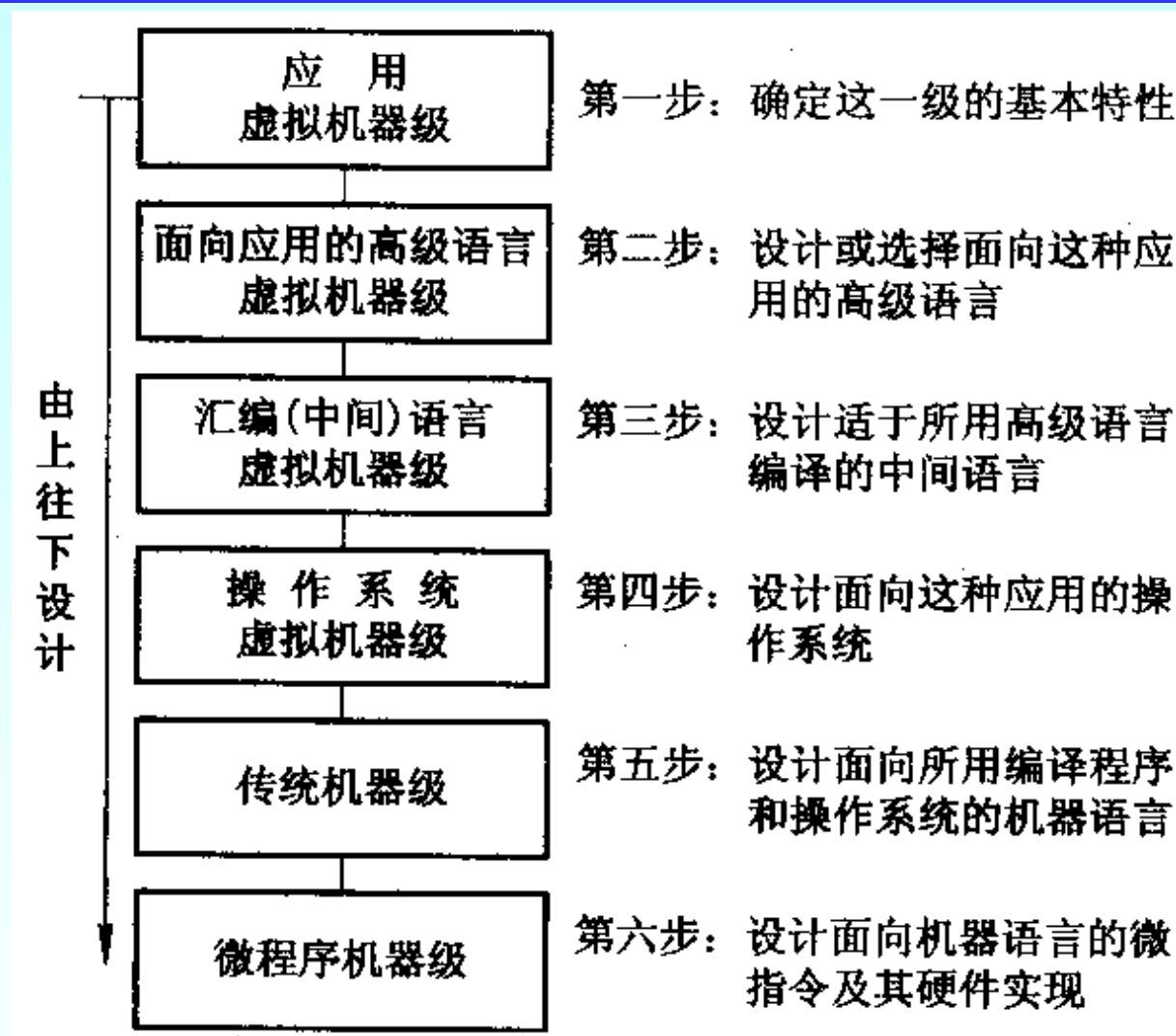


图 1.5 计算机系统“由上往下”设计的方法



## 1.3 软硬取舍与计算机系统的设计思路

### 由上往下设计

**特点：**从应用开始，逐级往下。

**优点：**运行效率高，软硬分配合理，适用于专用机的设计。

**缺点：**适应性差，周期长。

**解决方法：**不完全优化，不专门设计机器级“选型”。



## 1.3 软硬取舍与计算机系统的设计思路

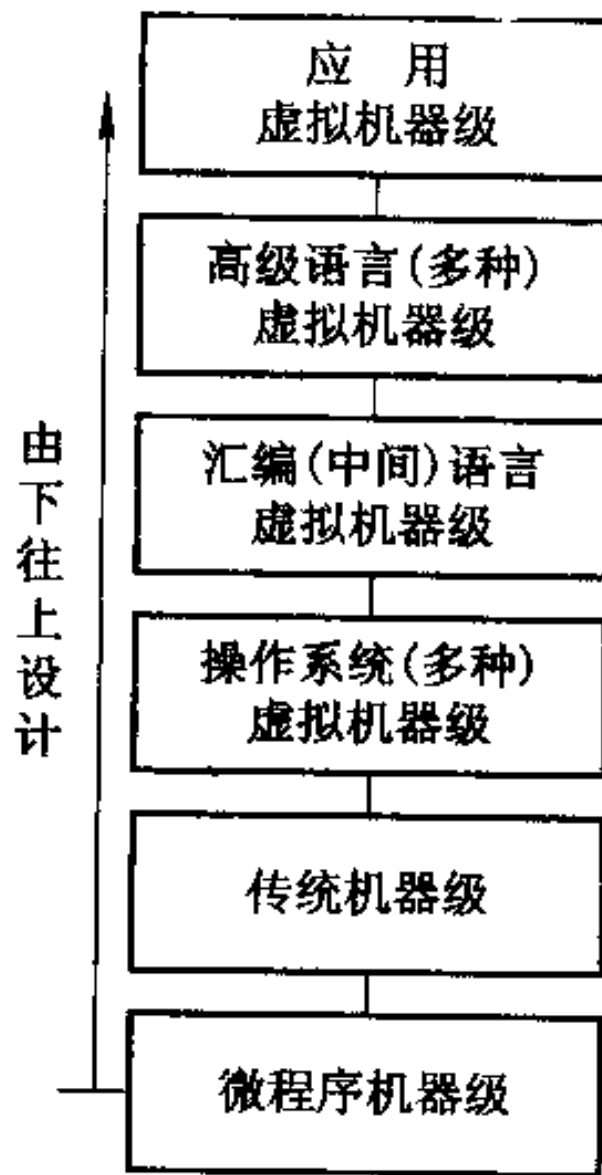


图 1.6 计算机体系结构设计的方法



## 1.3 软硬取舍与计算机系统的设计思路

### 由下往上设计

**特点：**根据器件等情况研制硬件，根据要求配置软件。

**优点：**可设计通用计算机。

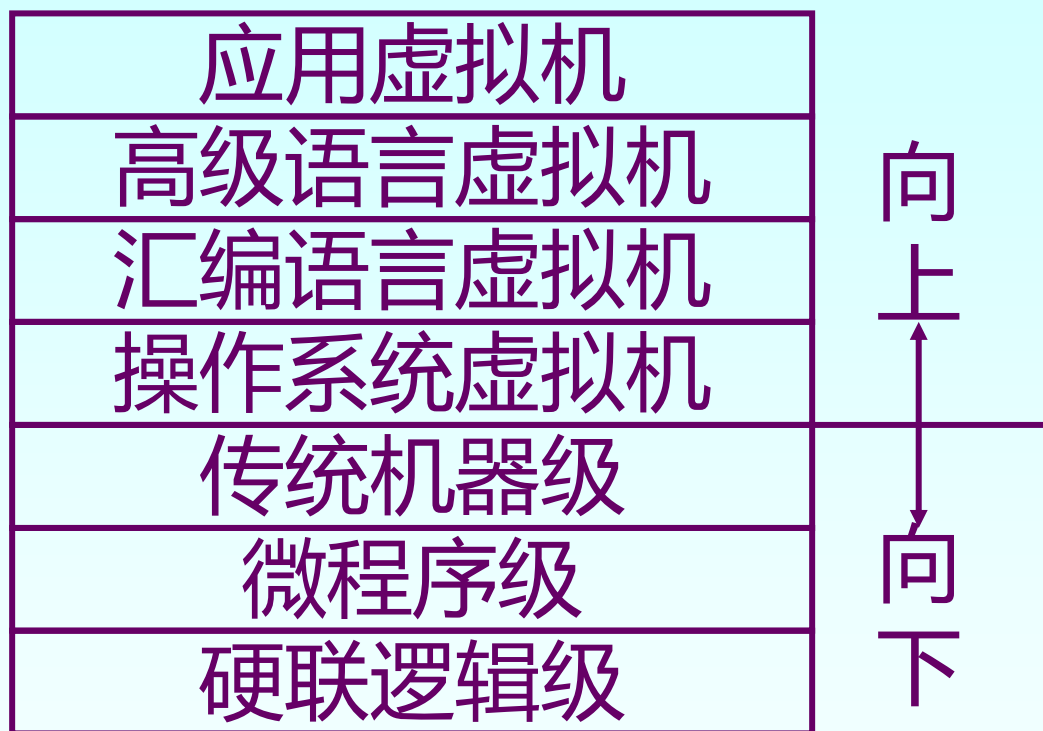
**缺点：**软硬脱节、分离；硬件无法改变，某些性能指标是虚假的。

——**很少使用**——



## 1.3 软硬取舍与计算机系统的设计思路

## 由中间开始



## 1.3 软硬取舍与计算机系统的设计思路

### 由中间开始设计

**特点：**从软硬界面开始，同时进行软硬件设计。

**优点：**软硬件功能分配比较合理，缩短了研制周期，有利于硬件和软件设计人员之间的交流协调，解决软硬设计分离和脱节的问题。

**“中间”**指的是层次结构中的软硬交界面，目前多数是在传统机器级与操作系统机器级之间。





## 1.3 软硬取舍与计算机系统的设计思路

计算方向的变化导致三个不同的计算机市场：

**1、桌面电脑：性价比**

**2、服务器：可用性、可扩展性、有效的吞吐量**

**3、嵌入式计算机：实时需求、最小化存储器需求、最小化功耗需求**



## 1.3 软硬取舍与计算机系统的设计思路

四种飞速发展的实现技术对现代计算机实现的影响深远：

1. 集成电路逻辑技术
2. 半导体**DRAM**（动态随机访问存储器）
3. 磁盘存储技术
4. 网络技术

