

计算机组成原理

傅颖勋

北方工业大学计算机科学与技术系

邮箱: fuyx@ncut.edu.cn

电话: Mobile:13488717987

学习指南

- 本课程的性质
 - 专业核心基础课程
 - 计算机硬件基础课程
 - 考研硬件基础课程（四门之一，30%）
- 本课程在专业体系中的地位
 - 本课程在计算机科学与技术学科中处于承上启下的地位，要求先修计算机导论，数字逻辑，程序设计基础
- 本课程的主要任务

通过学习本课程，掌握计算机各大部件的原理、逻辑实现及其互连构成整机系统的技术。为后继课程的学习打好基础

主要内容

- 第一章 计算机系统概述
- 第二章 运算方法与运算器
- 第三章 指令系统
- 第四章 中央处理器
- 第五章 主存储器与存储系统
- 第六章 辅助存储器
- 第七章 输入输出设备与系统

学习要点

- 从复杂的计算机系统抽象出主要模型
- 各部分构成原理(重点控制器、运算器、存储器)
- 学习后希望掌握高级语言程序怎么转换成机器语言程序，硬件如何执行这些程序
- 通过硬件设计怎样提高计算机系统的性能
- 什么是并行处理？(结合多核系统发展状况引出并行系统的重要性)
- 能够理解计算机体系结构与高级语言、编译系统、操作系统的关系

学习指南

- 学习目标

- 具有对计算机硬件系统的**认知能力**；
- 在认知能力的基础上，学会**运算部件、存储部件、指令系统、控制单元**以及整机硬件系统的**设计方法**；
- 在认知和设计能力的基础上，为具备一定的创新意识和创新能力打下基础

- 课程安排

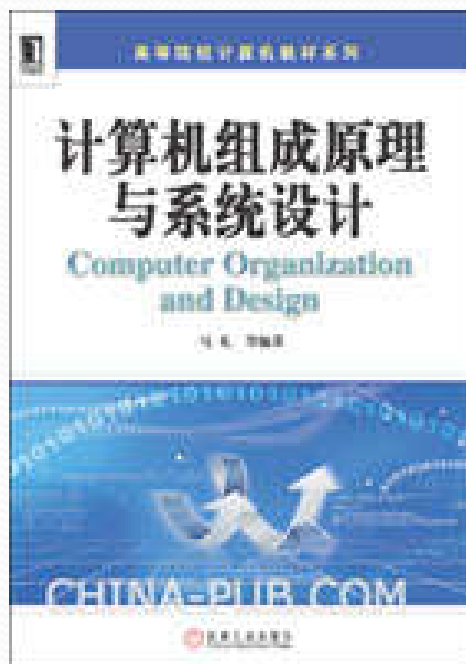
- **64学时=48(理论教学)+16(实验)**
- **考核:平时成绩(40%)+期末考试(60%)**
- **平时:出勤(0.3)、作业(0.2)、期中(0.2)、实验(0.3)**

参考书籍



- 唐朔飞.计算机组成原理(第二版) 高等教育出版社.2010
- 蒋本珊,计算机组成原理,清华大学出版社. 2013
- 白中英.计算机组成原理.科学出版社.2013

参考书籍



- 马礼.计算机组成原理与系统设计.机械工业出版社.2011
- David Patterson, John Hennessy. 王党辉，康继昌等译.计算机组成与设计/硬件软件接口，机械工业出版社，2015
- Randal E. Bryant. 龚奕利,贺莲译.深入理解计算机系统，机械工业出版社,2016

第一章 计算机系统概述

- 1.1 电子数字计算机与存储程序控制
- 1.2 计算机的类型和应用
- 1.3 计算机系统的硬件组成
- 1.4 计算机系统的层次结构
- 1.5 计算机的工作过程和主要性能指标

在本章中我们将从存储程序的概念入手，讨论电子数字计算机的**基本组成与工作原理**，使大家对于计算机系统先有一个简单的整体概念，为今后深入讨论各个部件打下基础。

电子数字计算机

- 电子计算机是一种不需要人工直接干预，能够自动、高速、准确地处理和存储各种信息的电子设备。
- 可以有非电子计算机如：量子计算机。
- 可以有非数字计算机如：模拟计算机。
- 世界上第一台电子数字计算机是1946年2月15日问世的ENIAC。(宾夕法尼亚大学)
 - 莫齐利(1907-1980): 首席顾问
 - 艾克特(1919-1995): 首席工程师
 - ENIAC的设计开始于1943年, 一直使用到1955年。

第一台电子计算机—ENIAC

ENIAC的特点:

- 采用十进制
- 20 个10位的累加器
- 用开关手动编程
- 18,000个电子管
- 重30 吨
- 占地170平方米
- 耗电140 KW
- 5,000次/秒加法运算



超级计算机 (2016.11)

Rank	Name	Site	Country	Year	Rmax	Rpeak	Nmax	Power	Mflops/Watt	Processor	System Model
1	Sunway TaihuLight	National Supercomputing Center in Wuxi	China	2016	93014594	125435904	12288000	15371	6051.3	Sunway SW26010 260C 1.45GHz	Sunway MPP
2	Tianhe-2 (MilkyWay-2)	National Super Computer Center in Guangzhou	China	2013	33862700	54902400	9960000	17808	1901.54	Intel Xeon E5-2692v2 12C 2.2GHz	TH-IVB-FEP Cluster
3	Titan	DOE/SC/Oak Ridge National Laboratory	United States	2012	17590000	27112550	0	8209	2142.77	Opteron 6274 16C 2.2GHz	Cray XK7
4	Sequoia	DOE/NNSA/LLNL	United States	2011	17173224	20132659.2	0	7890	2176.58	Power BQC 16C 1.6GHz	BlueGene/Q
5	Cori	DOE/SC/LBNL/NERSC	United States	2016	14014700	27880653	6984960	3939	3557.93	Intel Xeon Phi 7250 68C 1.4GHz	Cray XC40
6	Oakforest-PACS	Joint Center for Advanced High Performance Computing	Japan	2016	13554600	24913459	9938880	2718.7	4985.69	Intel Xeon Phi 7250 68C 1.4GHz	PRIMERGY CX1640 M1
7	Trubuty	RIKEN Advanced Institute for Computational Science (AICS)	Japan	2011	10510000	11280384	11870208	12659.89	830.18	SPARC64 VIIIfx 8C 2GHz	K computer
8	Piz Daint	Swiss National Supercomputing Centre (CSCS)	Switzerland	2016	9779000	15987968	2488320	1312	7453.51	Xeon E5-2690v3 12C 2.6GHz	Cray XC50
9	Mira	DOE/SC/Argonne National Laboratory	United States	2012	8586612	10066330	0	3945	2176.58	Power BQC 16C 1.6GHz	BlueGene/Q
10	Trinity	DOE/NNSA/LLNL/SNL	United States	2015	8100900	11078861	8847936	4232.63	1913.92	Xeon E5-2698v3 16C 2.3GHz	Cray XC40

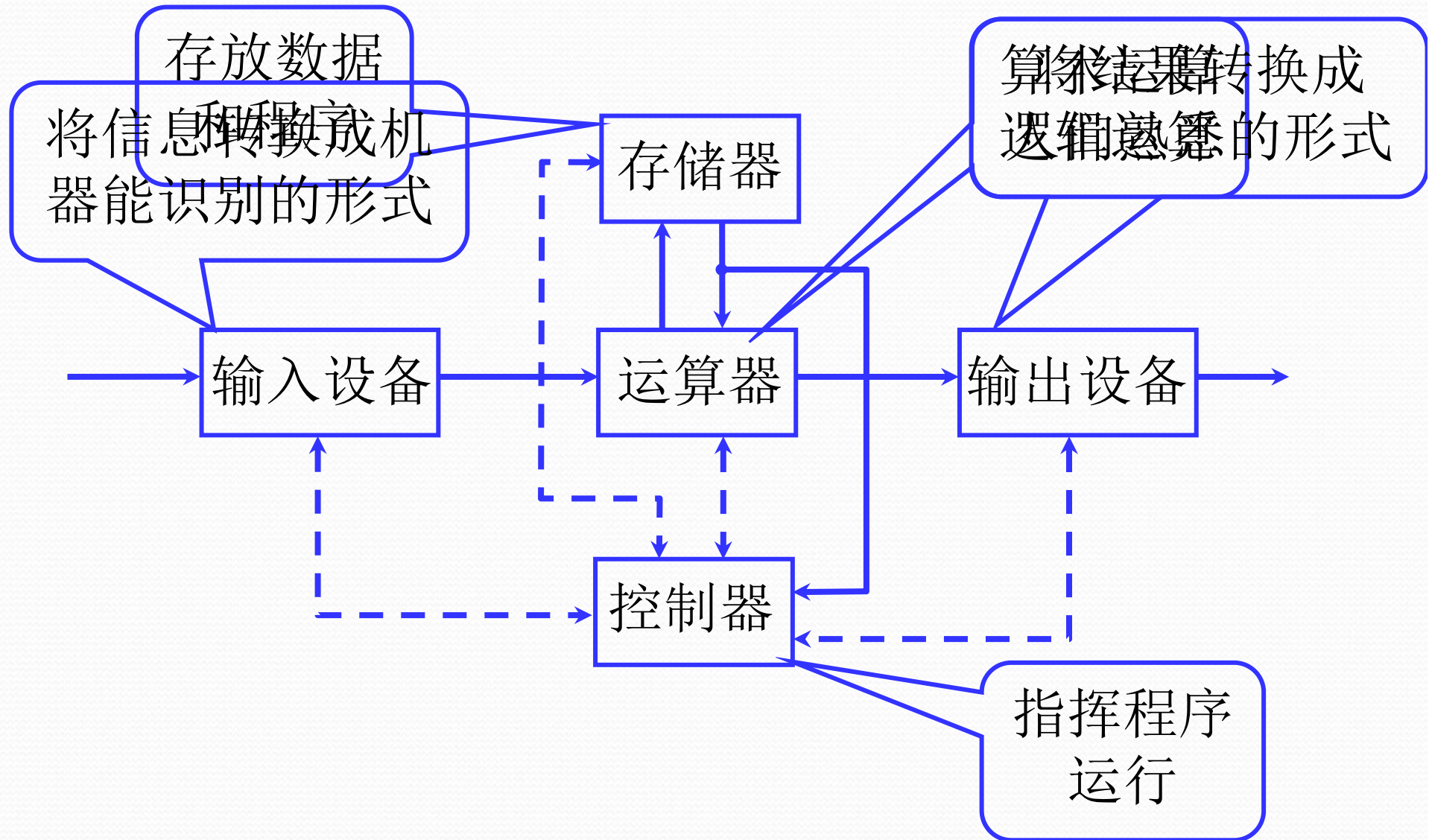
一些理念（思考）

- Abstraction (Layers of Representation/Interpretation)
- Moore's Law (Designing through trends)
- Principle of Locality (Memory Hierarchy)
- Parallelism & Amdahl's law (which limits it)
- Dependability via Redundancy
- 思考问题
 - 1) 计算机由哪些部分组成？
 - 2) 各部分之间什么关系？
 - 3) 为什么能够自动运行程序？

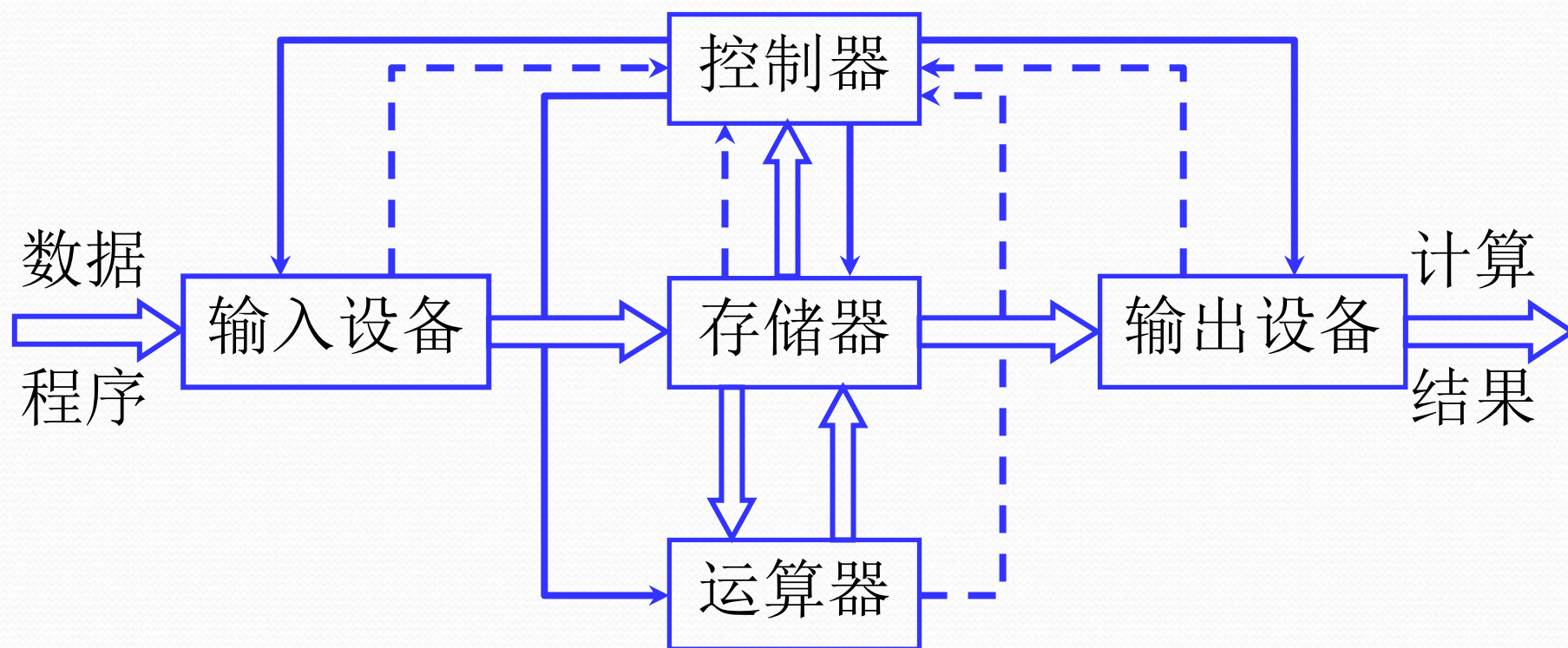
存储程序概念

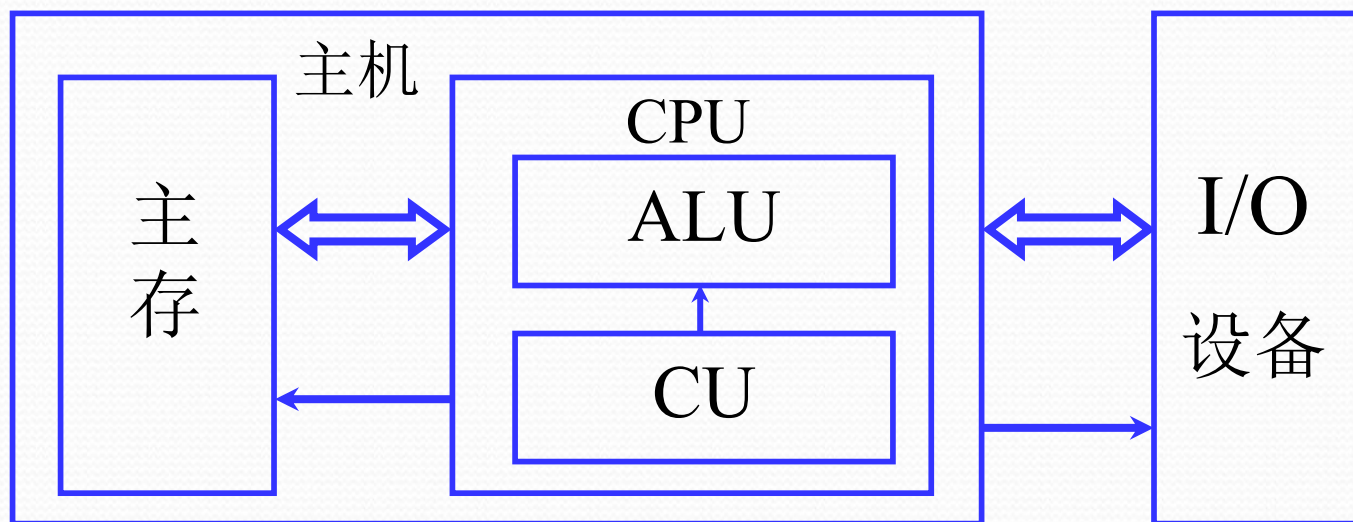
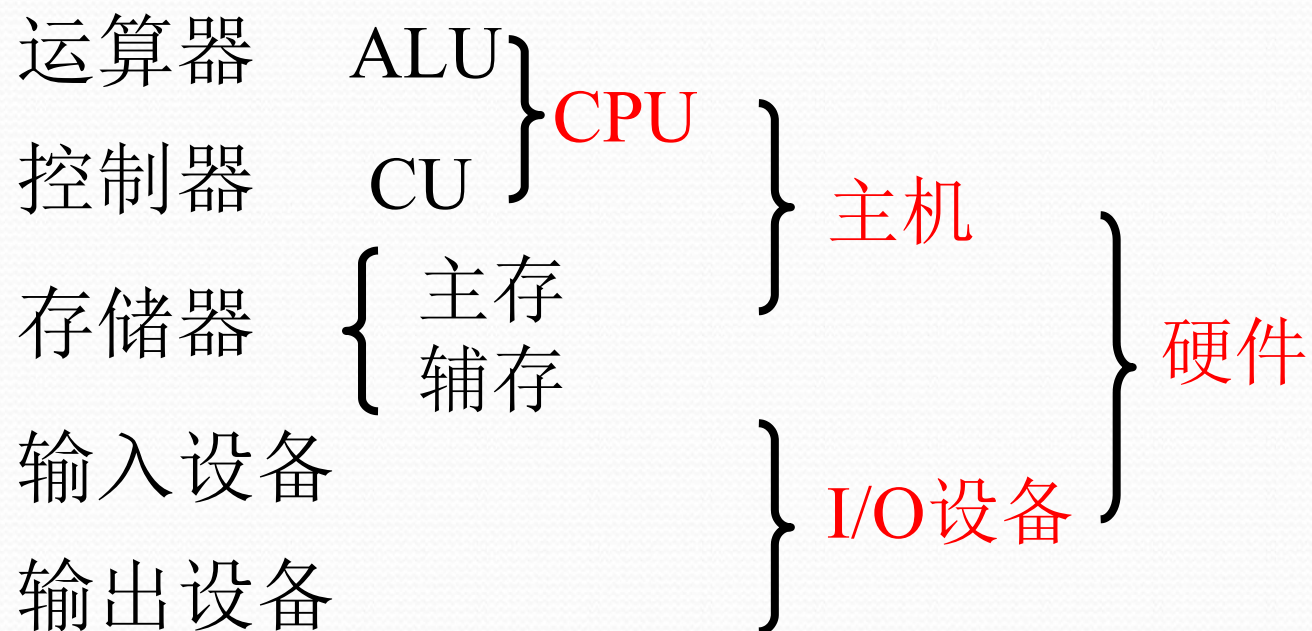
- 冯·诺依曼等人在1946年6月提出存储程序概念：
 - 计算机（指硬件）应由运算器、存储器、控制器、输入设备和输出设备五大基本部件组成；
 - 计算机内部采用二进制来表示指令和数据；
 - 将编好的程序和原始数据事先存入存储器中，然后再启动计算机工作，使计算机在不需要人工干预的情况下，自动、高速地从存储器中取出指令加以执行。
 - 目前绝大多数计算机仍建立在存储程序概念的基础上，称冯·诺依曼型计算机。也出现了一些突破冯·诺依曼结构的计算机，统称非冯结构计算机，如：数据驱动的数据流计算机、需求驱动的归约计算机和模式匹配驱动的智能计算机等。

冯·诺依曼计算机硬件框图



以存储器为中心的硬件框图







第一章 计算机系统概述

- 1.1 电子数字计算机与存储程序控制
- 1.2 计算机的类型和应用
- 1.3 计算机系统的硬件组成
- 1.4 计算机系统的层次结构
- 1.5 计算机的工作过程和主要性能指标

计算机的类型与应用

- 国内分类
 - 小型机、中型机
 - 大型机、巨型机
- 美国分类
 - **PMD**、桌面计算机
 - 服务器、集群/仓库级、嵌入式
- 计算机应用
 - 科学计算、数据处理、过程控制
 - 计算机辅助设计与制造、人工智能

系列机

- 系列机是指一个厂家生产的，具有相同的体系结构，但具有不同组成和实现的一系列不同型号的机器。
- 系列机应在指令系统、数据格式、字符编码、中断系统、控制方式、输入/输出操作方式等方面保持统一，从而保证软件的兼容性。
- 软件兼容
 - 向上兼容
 - 向下兼容 
 - 向前兼容
 - 向后兼容 

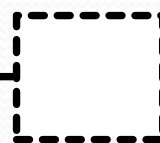
机器档次



向上
兼容

向下
兼容

当前机器



向前
兼容

向后
兼容



时间



第一章 计算机系统概述

- 1.1 电子数字计算机与存储程序控制
- 1.2 计算机的类型和应用
- 1.3 计算机系统的硬件组成
- 1.4 计算机系统的层次结构
- 1.5 计算机的工作过程和主要性能指标

输入输出设备

- 输入设备

- 输入设备的任务是把人们编好的程序和原始数据送到计算机中去，并且将它们转换成计算机内部所能识别和接受的信息方式。常用的有键盘、鼠标、扫描仪等。

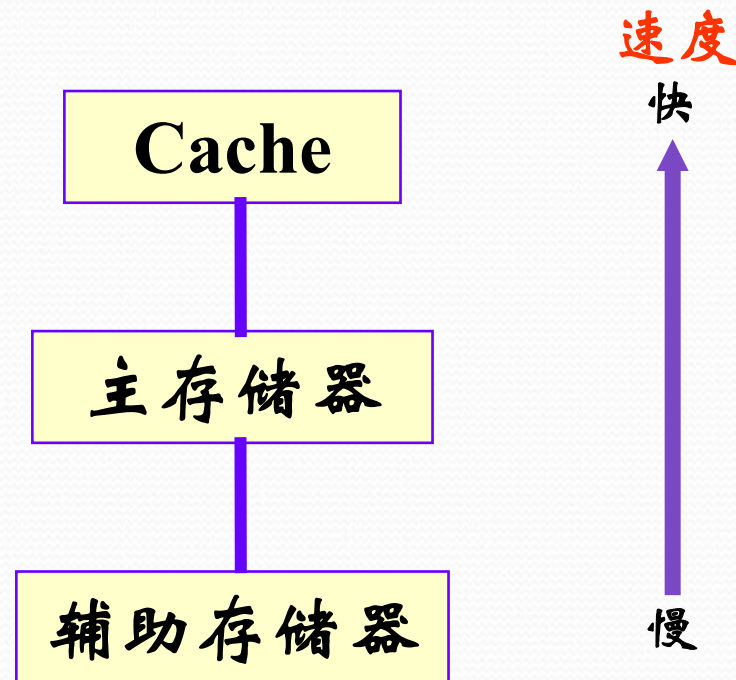
- 输出设备

- 输出设备的任务是将计算机的处理结果以人或其他设备所能接受的形式送出计算机。常用的有显示器、打印机、绘图仪等。

存储器

- 存储器是用来存放程序和数据部件，它是一个记忆装置，是计算机能够实现“存储程序控制”的基础。

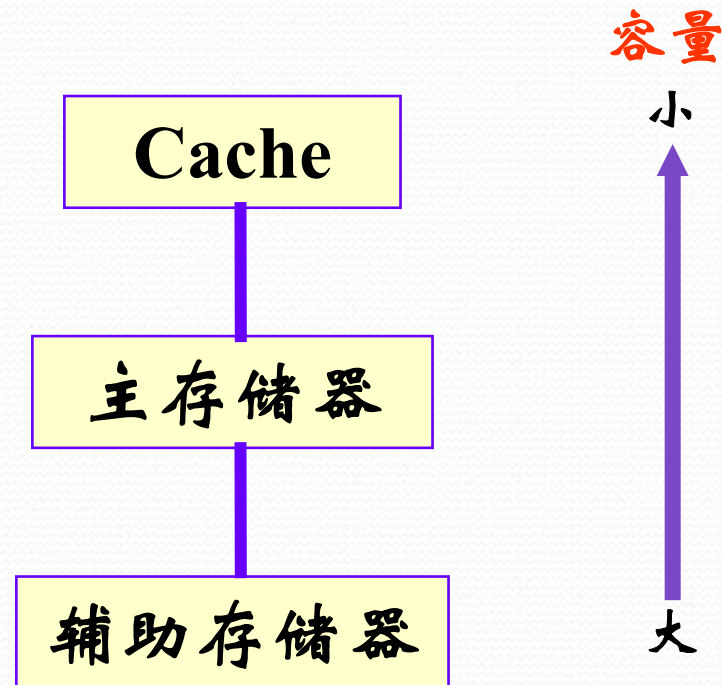
三级存储系统



存储器

- 存储器是用来存放程序和数据部件，它是一个记忆装置，是计算机能够实现“存储程序控制”的基础。

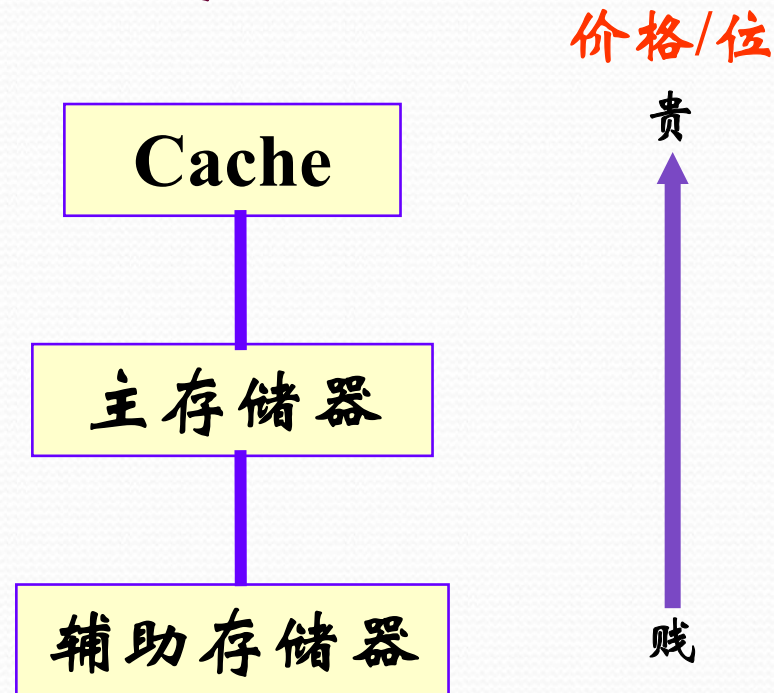
三级存储系统



存储器

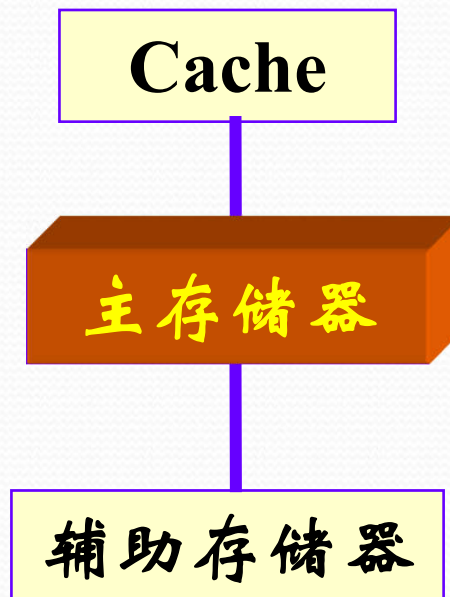
- 存储器是用来存放程序和数据部件，它是一个记忆装置，是计算机能够实现“存储程序控制”的基础。

三级存储系统



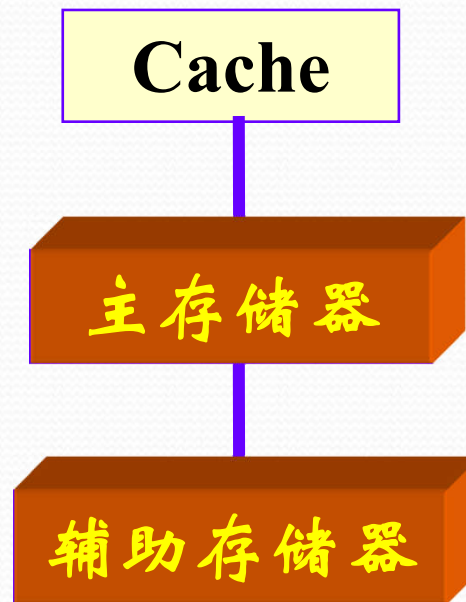
主存储器

- 可由CPU直接访问，用来存放当前正在执行的程序和数据。



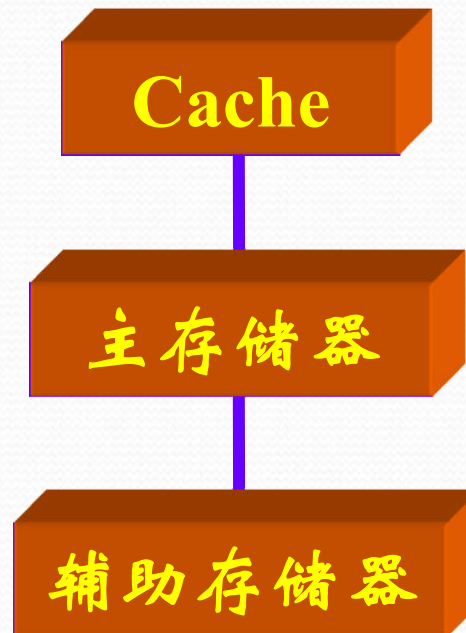
辅助存储器

- 设置在主机外部，CPU不能直接访问，用来存放暂时不参与运行的程序和数据，需要时再传送到主存。



高速缓冲存储器

- CPU可以直接访问，用来存放当前正在执行的程序中的活跃部分，以便快速地向CPU提供指令和数据。



运算器与控制器

- 运算器

- 运算器是对信息进行处理和运算的部件，经常进行的运算是算术运算和逻辑运算，因此运算器的核心是算术逻辑运算部件ALU。
- 运算器中有若干个寄存器（累加寄存器、暂存器等）。

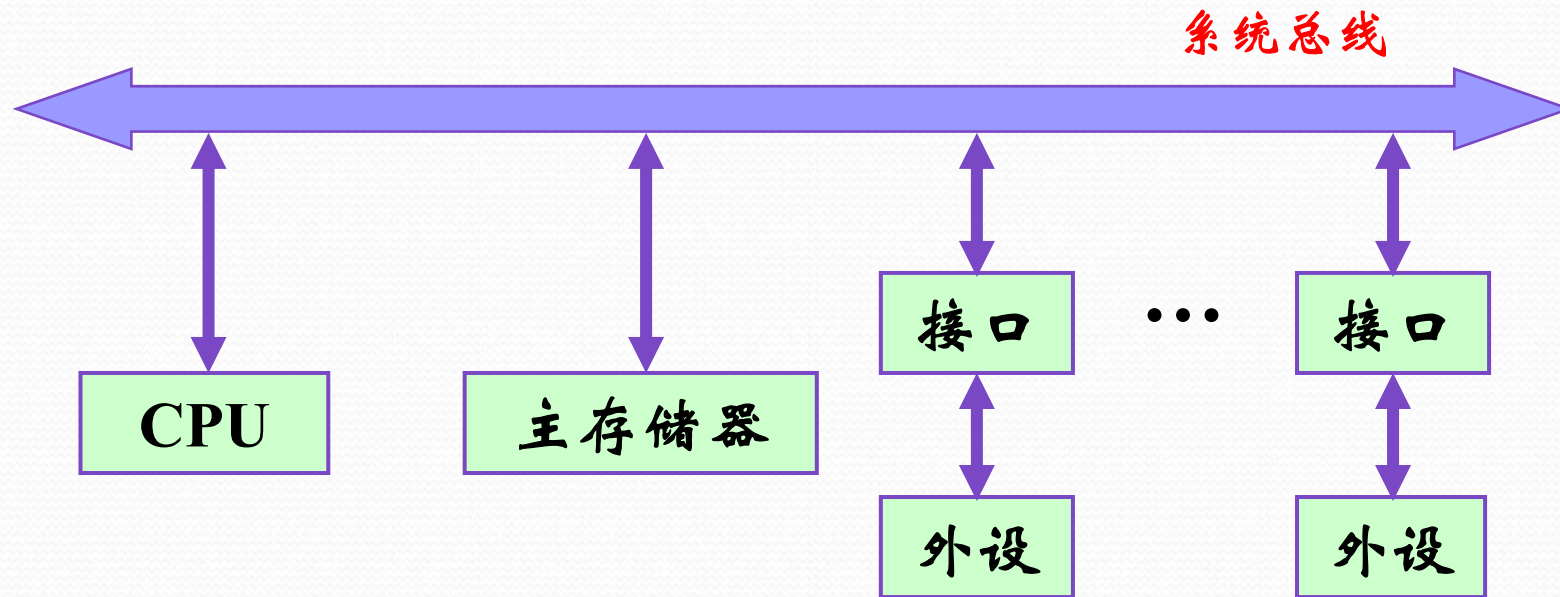
- 控制器

- 控制器是整个计算机的指挥中心。
- 控制器中主要包括时序控制信号形成部件和一些专用的寄存器。

计算机总线结构

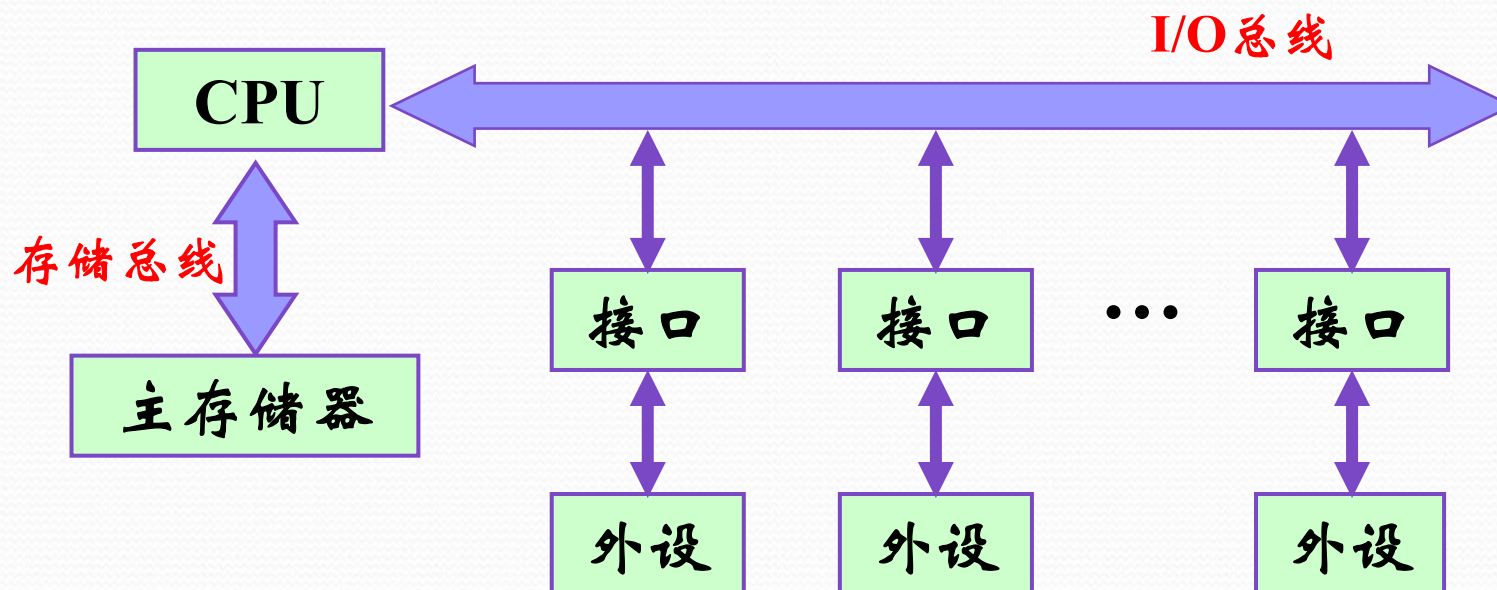
- 将各大基本部件，按某种方式连接起来就构成了计算机的硬件系统。
- 目前许多计算机(主要指小、微型计算机)的各大基本部件之间是用总线(Bus)连接起来的。
- 总线是一组能为多个部件服务的公共信息传送线路，它能分时地发送与接收各部件的信息，具有共享性和分时性等特点
- 小型、微型机的设计目标是以较小的硬件代价组成具有较强功能的系统，总线结构正好能满足这一要求。

单总线结构



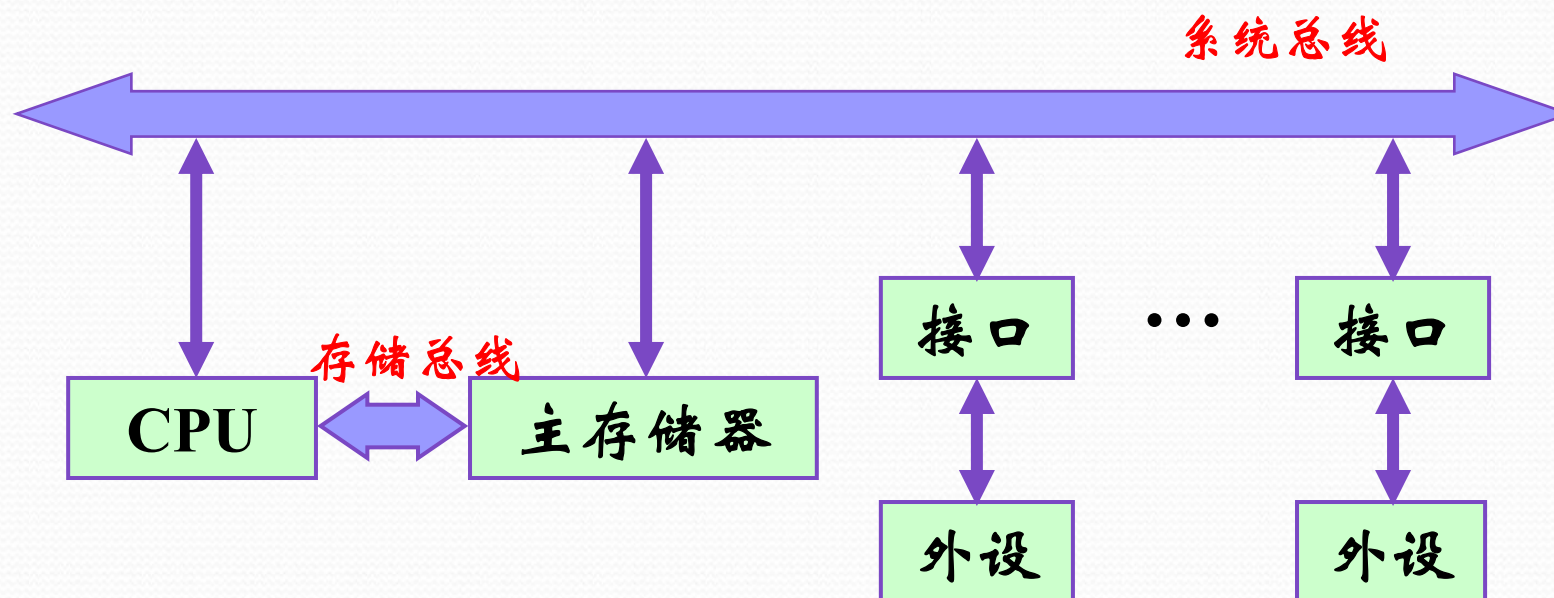
提高了CPU的工作效率，外设连接灵活。但信息传送的吞吐率受到限制，控制逻辑比专用存储复杂。

以CPU为中心的双总线结构



结构简单。但外设与主存间必须通过CPU进行信息交换，降低了CPU的工作效率。

面向存储器的双总线结构

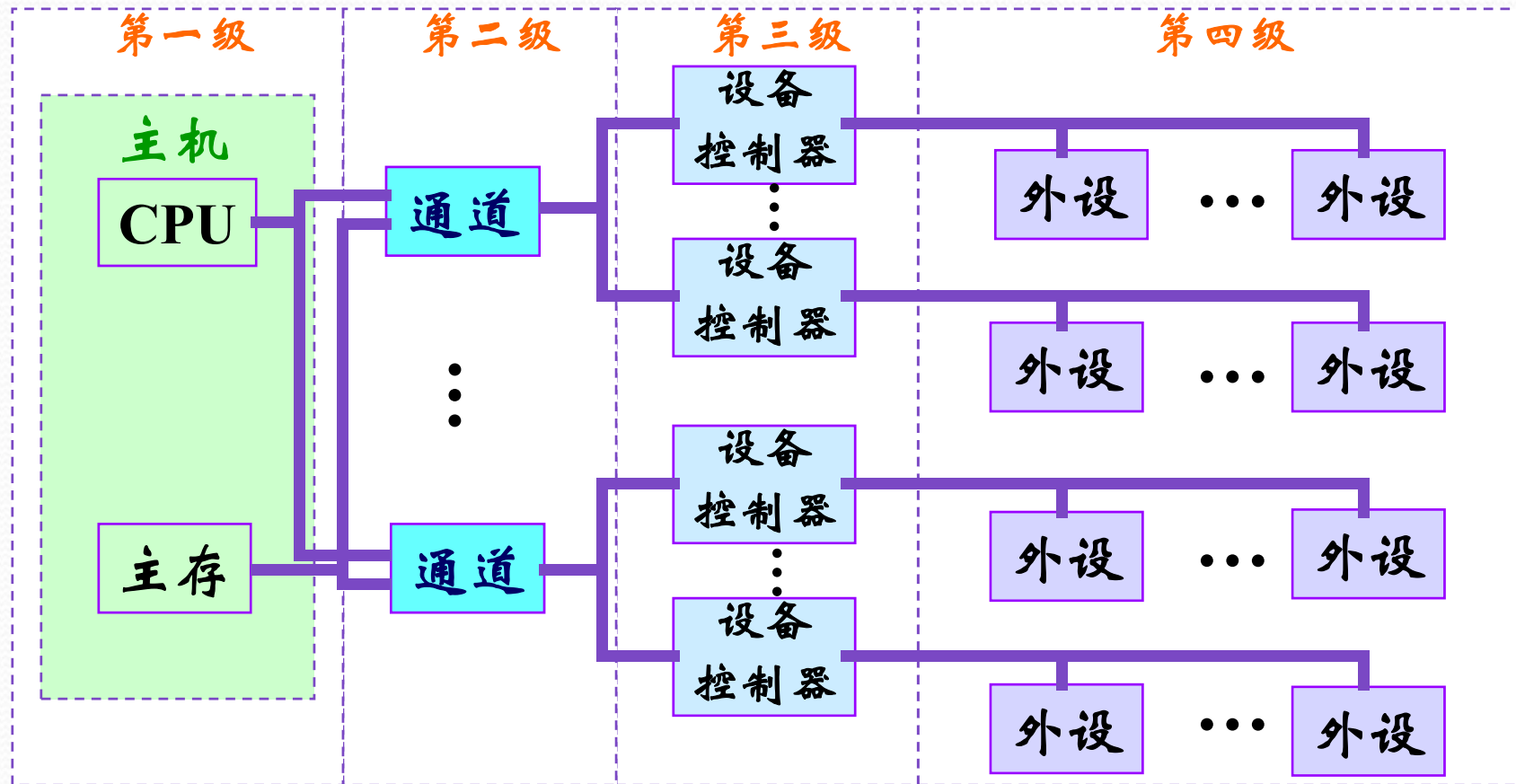


保留了单总线结构的优点，而存取速度大大提高。但硬件的代价随之增加。

大、中型机的典型结构

- 大、中型计算机系统的设计目标更着重于系统功能的扩大与效率的提高。
- 主机可以连接多个通道，每个通道可以接一台或几台设备控制器，每个设备控制器又可接一台或几台外部设备，这样整个系统就可以连接很多的外部设备。
- 通道是承担I/O操作管理的主要部件，能使CPU的数据处理和外部设备交换信息这两项操作同时进行。

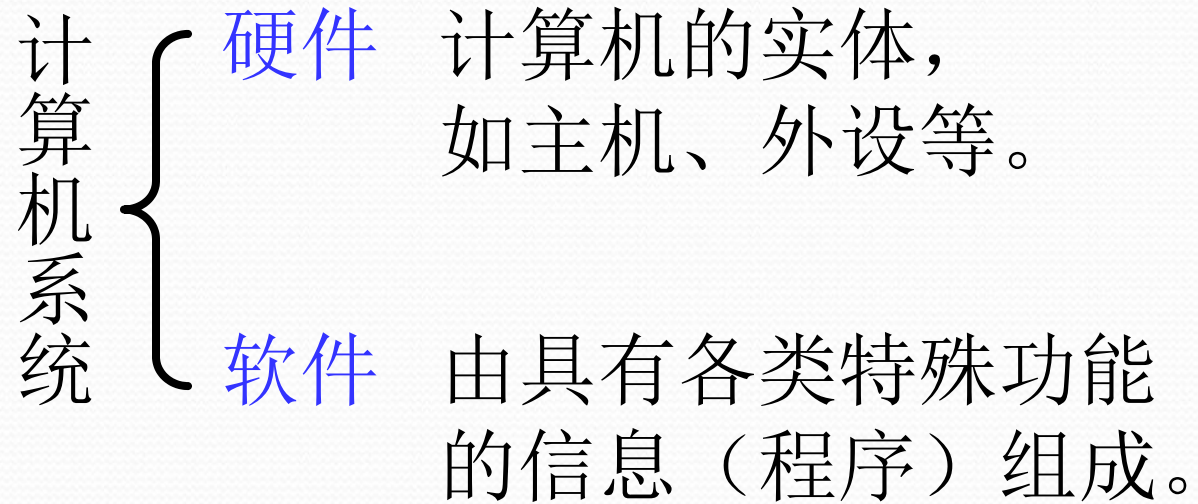
大、中型机的典型结构



第一章 计算机系统概述

- 1.1 电子数字计算机与存储程序控制
- 1.2 计算机的类型和应用
- 1.3 计算机系统的硬件组成
- 1.4 计算机系统的层次结构
- 1.5 计算机的工作过程和主要性能指标

计算机系统



硬件与软件

- 计算机系统=硬件系统+软件系统
 - 硬件是计算机系统的物质基础，软件是计算机系统的灵魂。硬件和软件是相辅相成的，不可分割的整体。
 - 当前计算机的硬件和软件正朝着互相渗透，互相融合的方向发展，在计算机系统中没有一条明确的硬件与软件的分界线。硬件和软件之间的界面是浮动的，对于程序设计人员来说，硬件和软件在逻辑上是等价的。
 - 硬件软化：原来由硬件实现的操作改由软件来实现。它可以增强系统的功能和适应性。
 - 软件硬化：原来由软件实现的操作改由硬件来实现。它可以显著降低软件在时间上的开销。

固件

- 固件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。

计算机软件

软件

系统软件 用来管理整个计算机系统

语言处理程序

操作系统

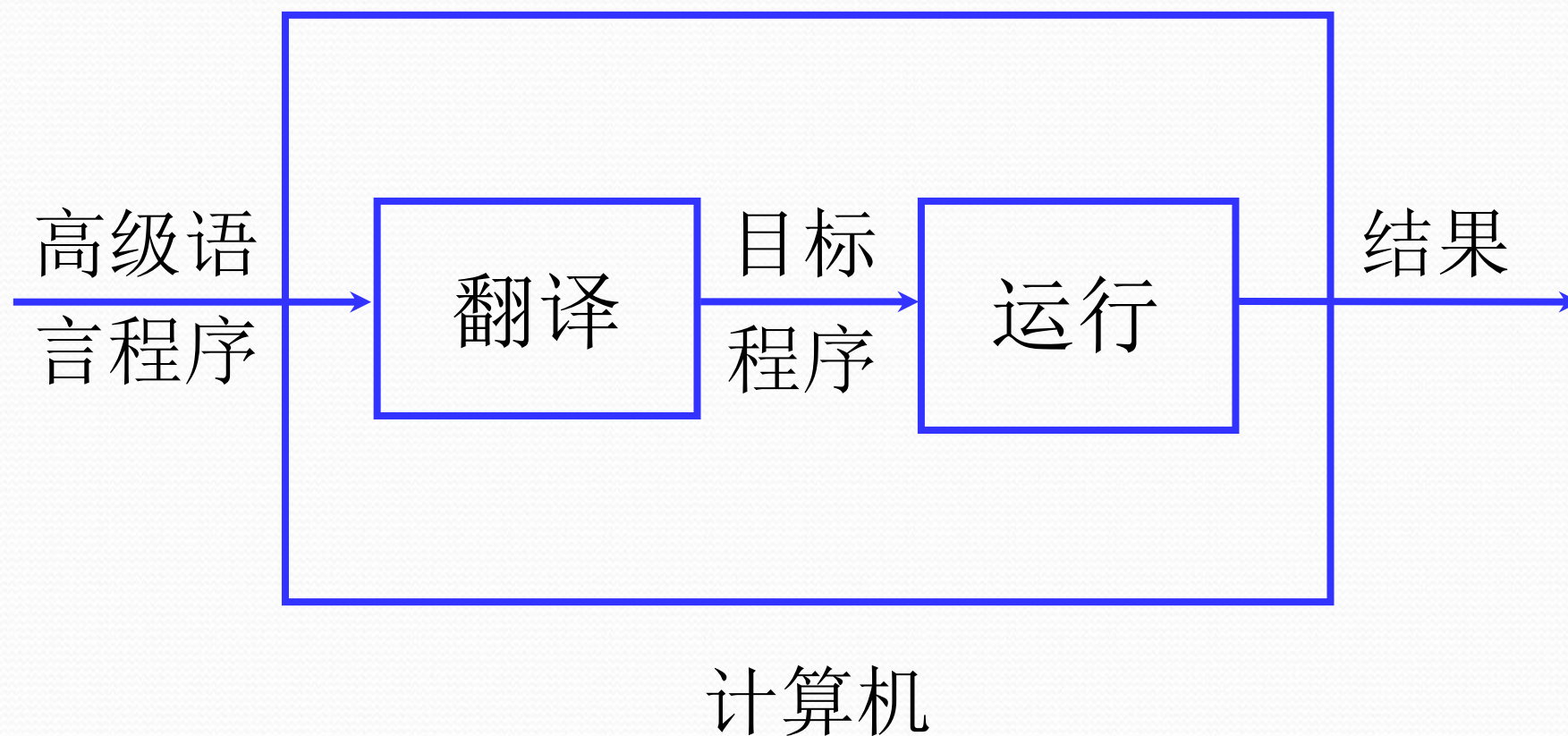
服务性程序

数据库管理系统

网络软件

应用软件 按任务需要编制成的各种程序

计算机解题过程



计算机解题示例

High-level language 高级语言 (C)

```
void swap(int a[],int k)
{
    int temp;
    temp=a[k];
    a[k]=a[k+1];
    a[k+1]=temp;
}
```

compiler

Assemble language(x86)

```
void swap(int a[],int k)
{
    push    ebp
    mov     ebp,esp
    sub     esp,0CCh
    push    ebx
    push    esi
    push    edi
    lea     edi,[ebp-0CCh]
    mov     ecx,33h
    mov     eax,0CCCCCCCCh
    rep stos dword ptr es:[edi]
    int temp;
    temp=a[k];
    mov     eax,dword ptr [k]
    mov     ecx,dword ptr [a]
    mov     edx,dword ptr [ecx+eax*4]
    mov     dword ptr [temp],edx
    a[k]=a[k+1];
    mov     eax,dword ptr [k]
    mov     ecx,dword ptr [a]
    mov     edx,dword ptr [k]
    mov     esi,dword ptr [a]
    mov     edx,dword ptr [esi+edx*4+4]
    mov     dword ptr [ecx+eax*4],edx
    a[k+1]=temp;
    mov     eax,dword ptr [k]
    mov     ecx,dword ptr [a]
    mov     edx,dword ptr [temp]
    mov     dword ptr [ecx+eax*4+4],edx
}
pop     edi
pop     esi
pop     ebx
mov     esp,ebp
pop     ebp
ret
```

assembler

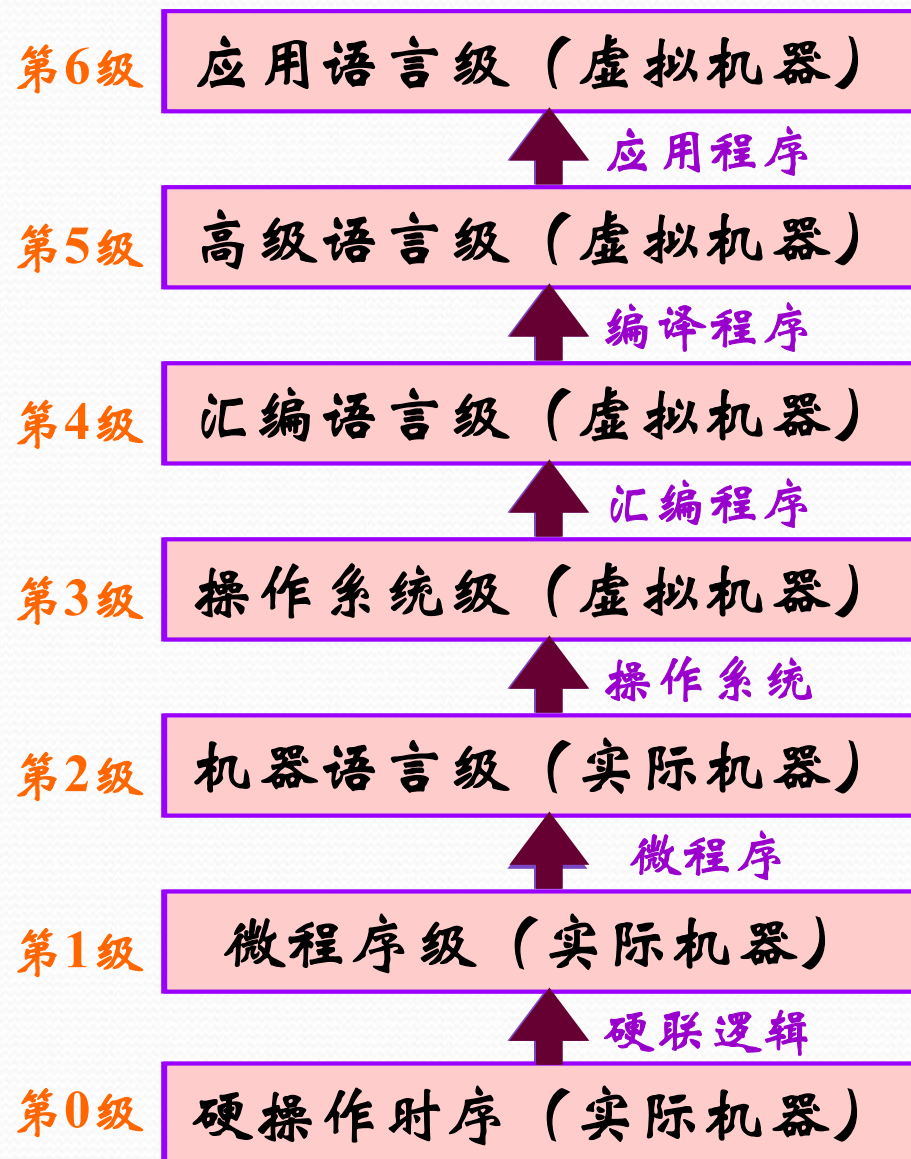
Hardware representation (x86)

```
00411390
00411391
00411393
00411399
0041139A
0041139B
0041139C
004113A2
004113A7
004113AC
004113AE
004113B1
004113B4
004113B7
004113BA
004113BD
004113C0
004113C3
004113C6
004113CA
004113CD
004113D0
004113D3
004113D6
004113DA
004113DB
004113DC
004113DD
004113DF
004113E0
```


计算机层次结构

现代计算机系统是一个硬件与软件组成的综合体，我们可以把它看成是按功能划分的多级层次结构。

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，他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，虚拟计算机即是由软件实现的机器。



计算机层次结构

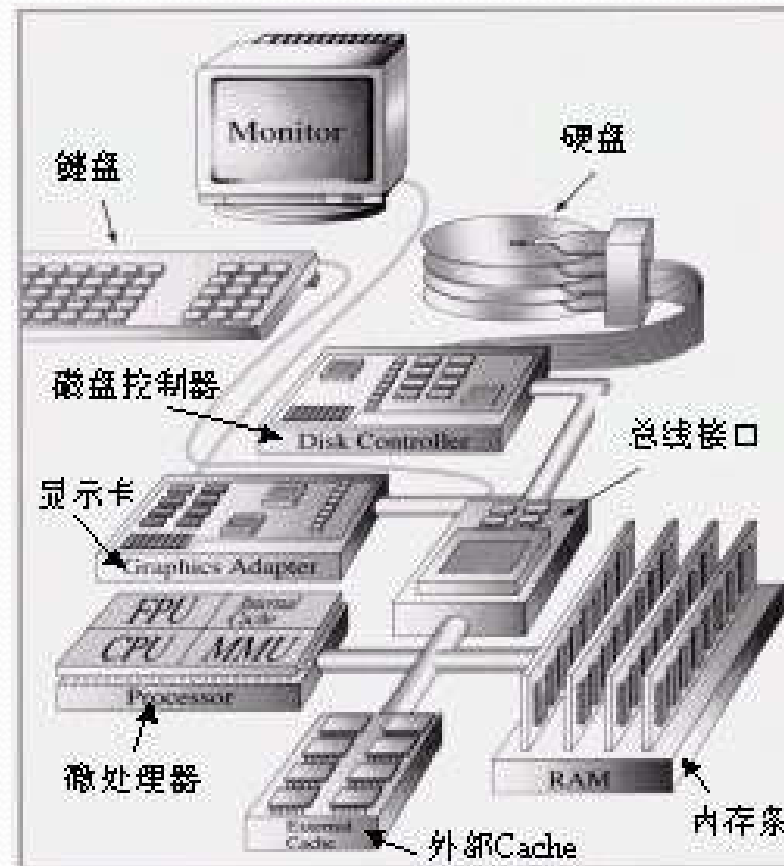


一般用户
观察到的
计算机硬
件系统



计算机层次结构

系统集成专业用户观察到的计算机硬件系统

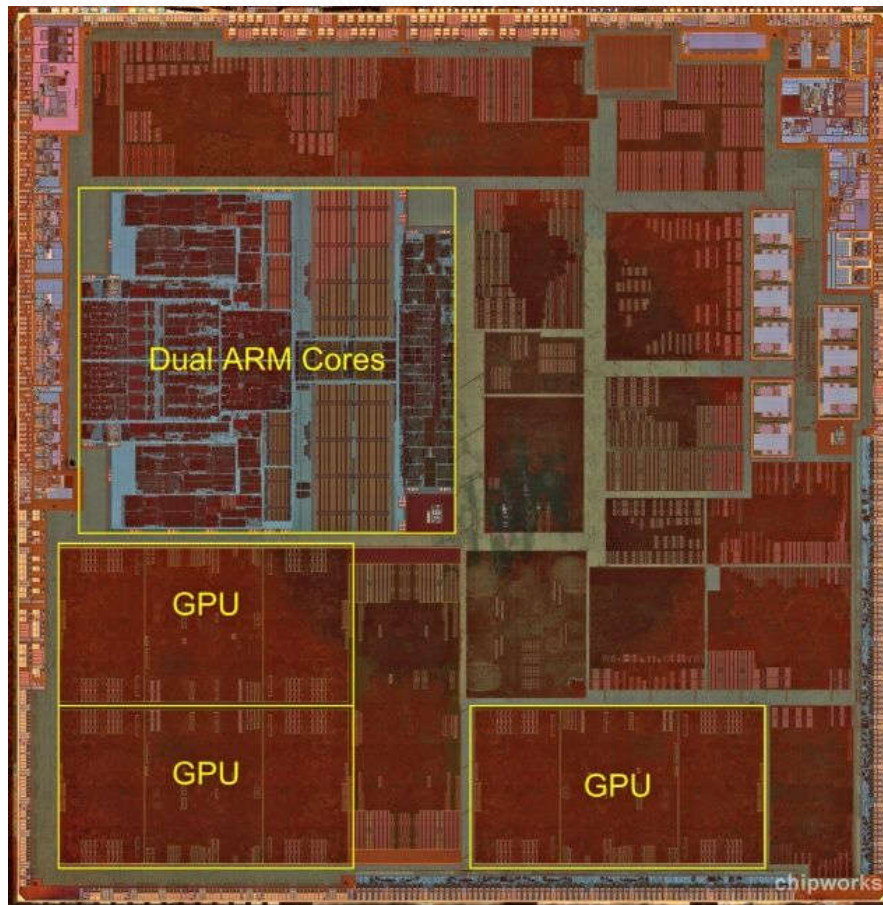


Inside the Processor (CPU)

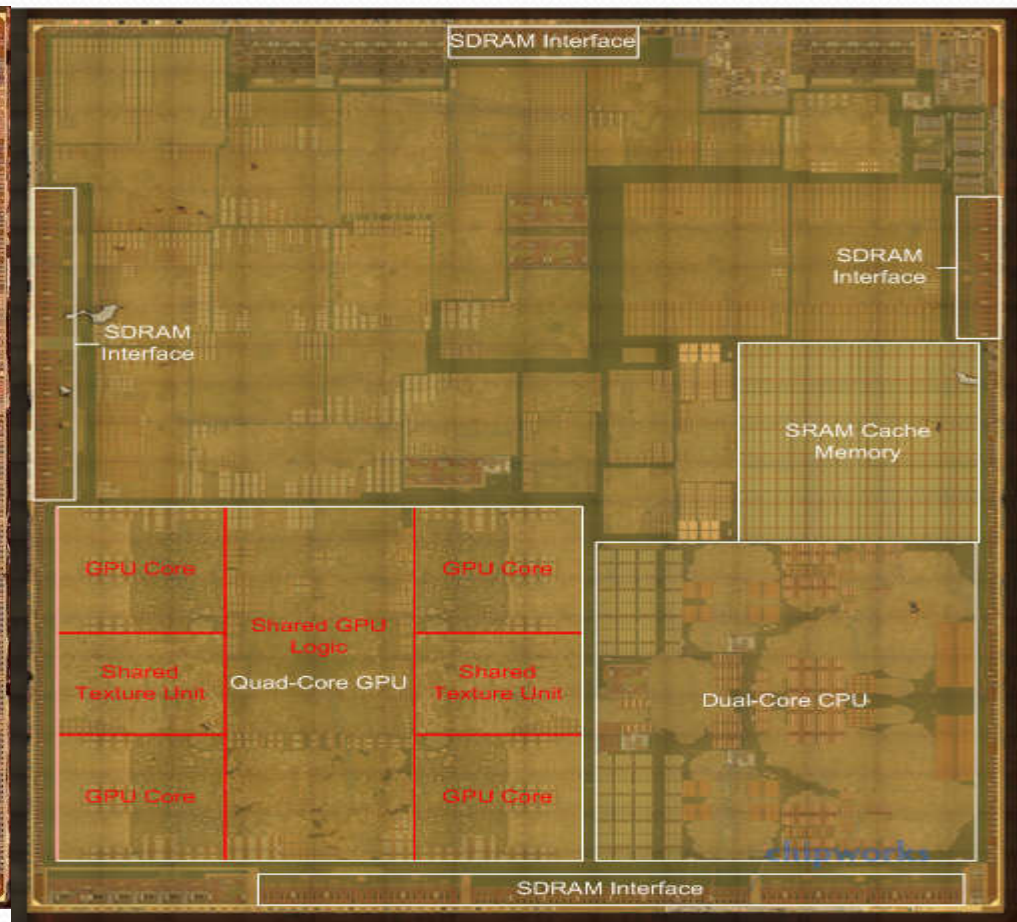
- Datapath: performs operations on data
- Control: sequences datapath, memory, ...
- Cache memory
 - Small fast SRAM memory for immediate access to data

Inside the Processor

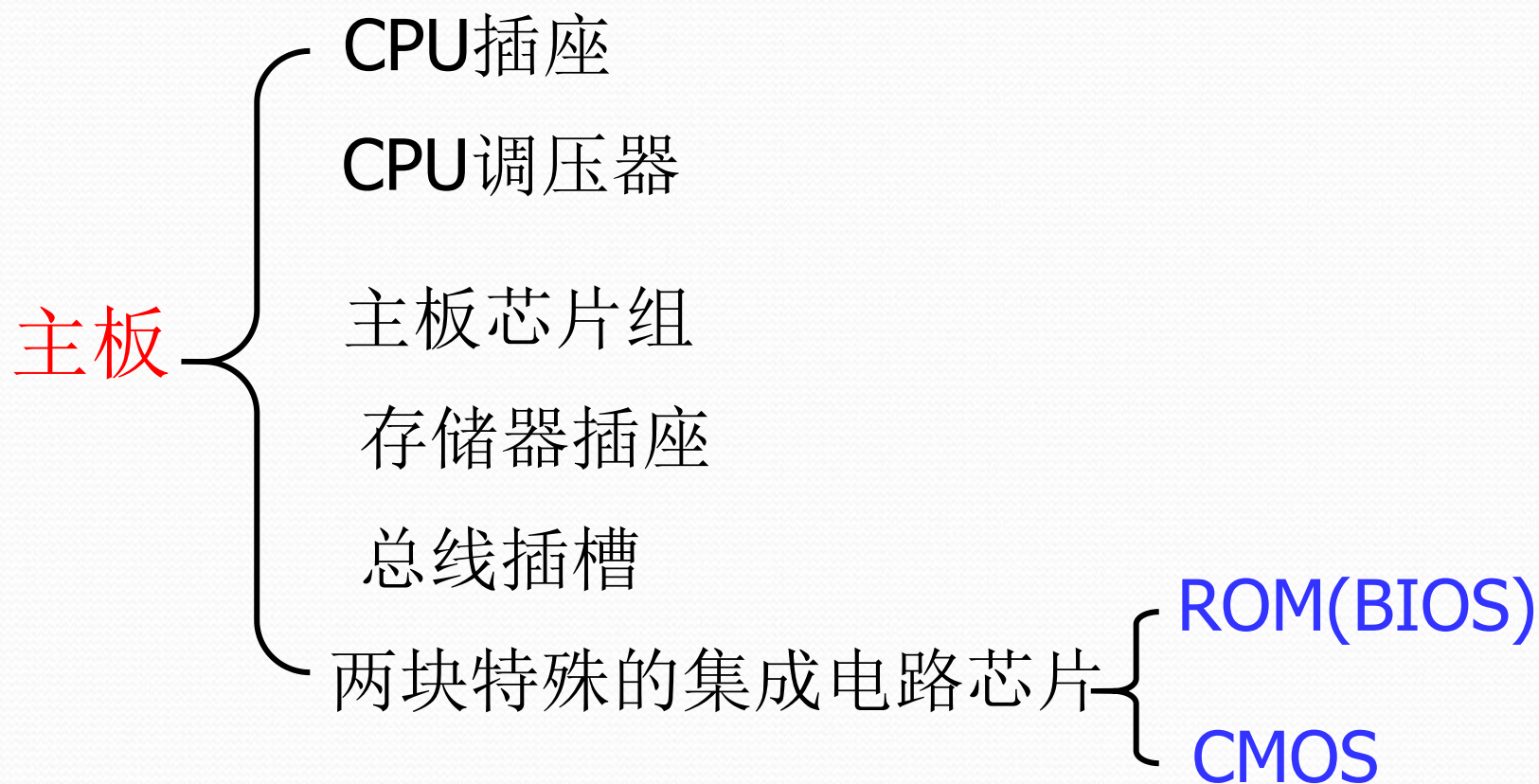
- Apple A7



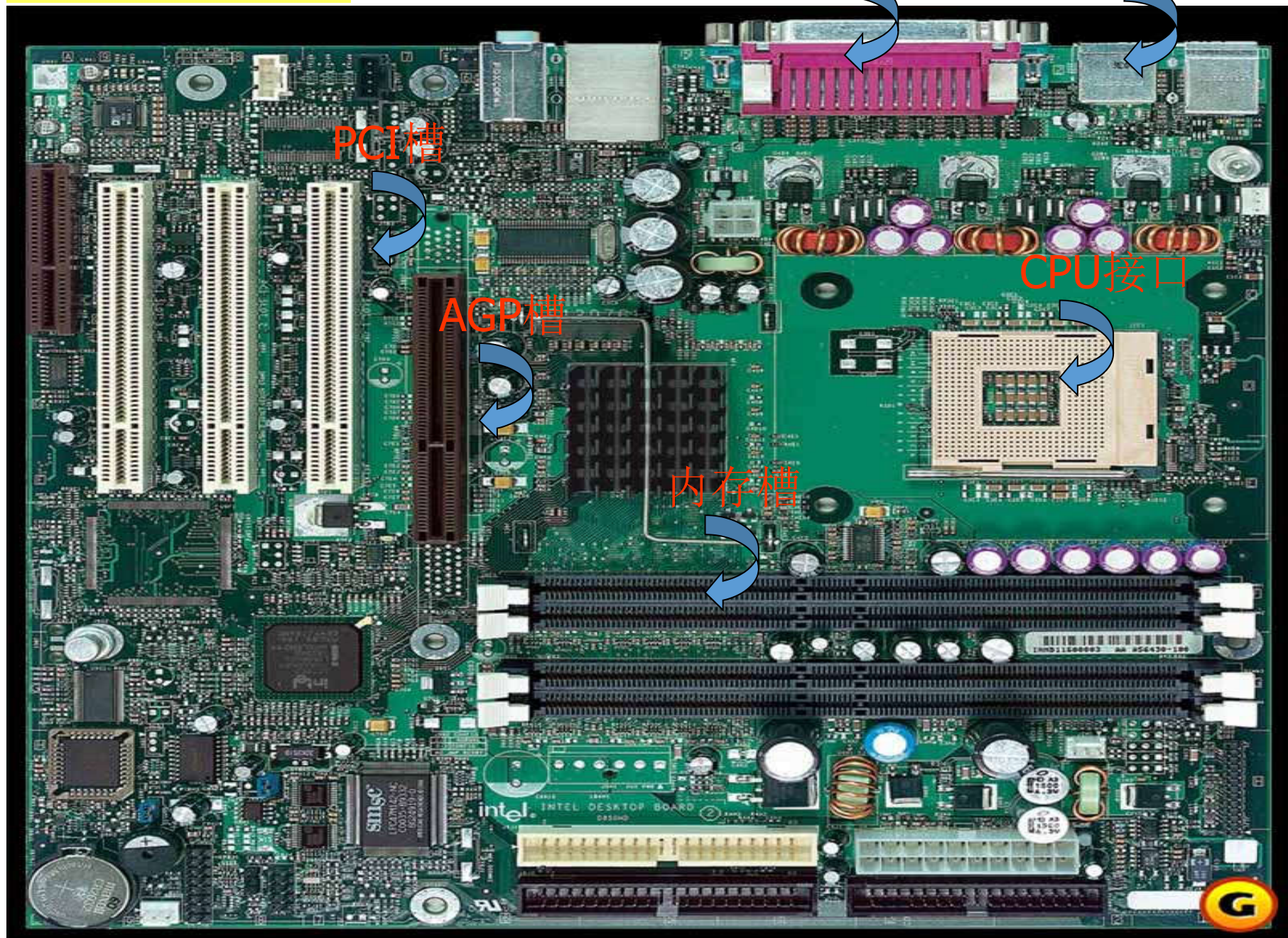
- Apple A8



主板、芯片组与BIOS

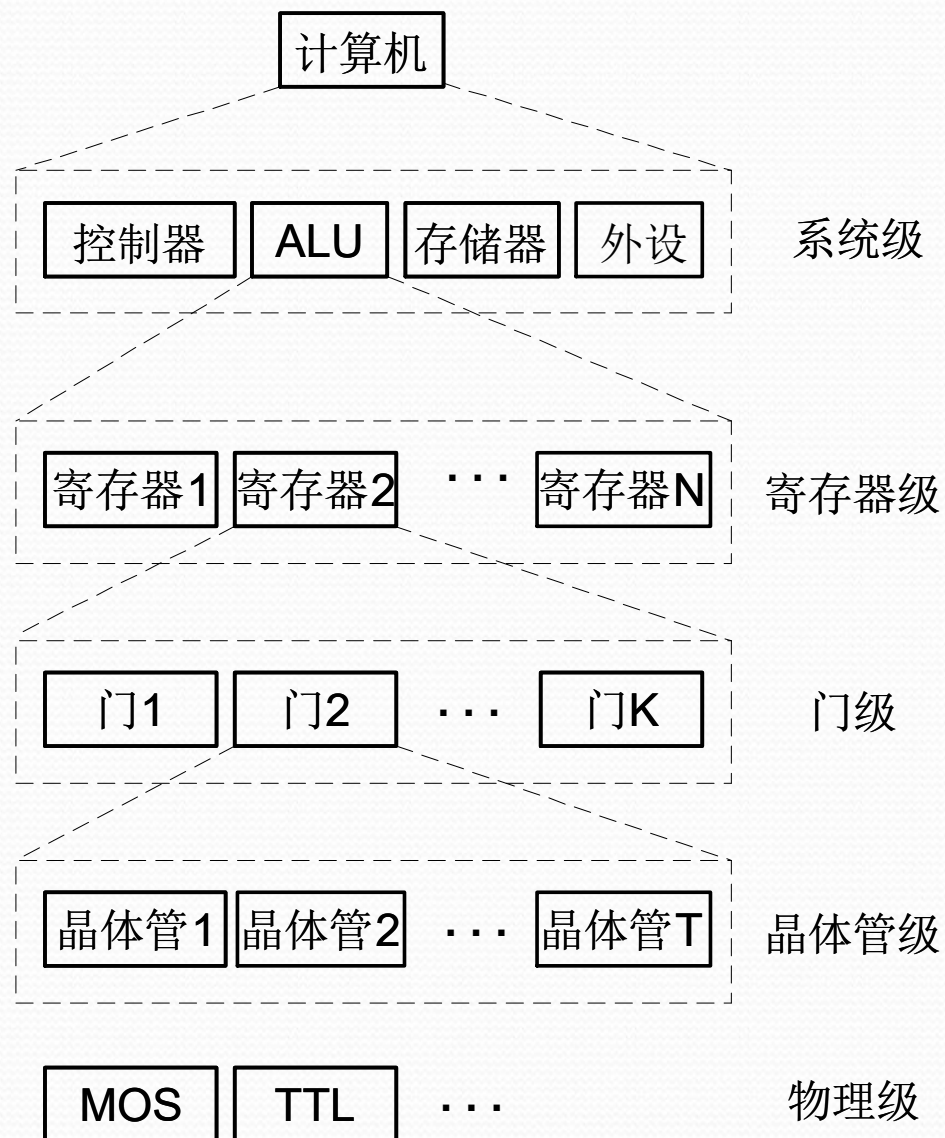


系统主板



计算机层次结构

计算机设计者
观察到的计算
机硬件系统



第一章 计算机系统概述

- 1.1 电子数字计算机与存储程序控制
- 1.2 计算机的类型和应用
- 1.3 计算机系统的硬件组成
- 1.4 计算机系统的层次结构
- 1.5 计算机的工作过程和主要性能指标

计算机的工作过程

例：设变量a、b、c分别存储在存储单元中(存储地址为005、006、007)，试编程计算 $a+b-c$ 的值，并将其存储在008地址单元中。以此程序执行为例说明计算机内部执行过程。

解：Load 005

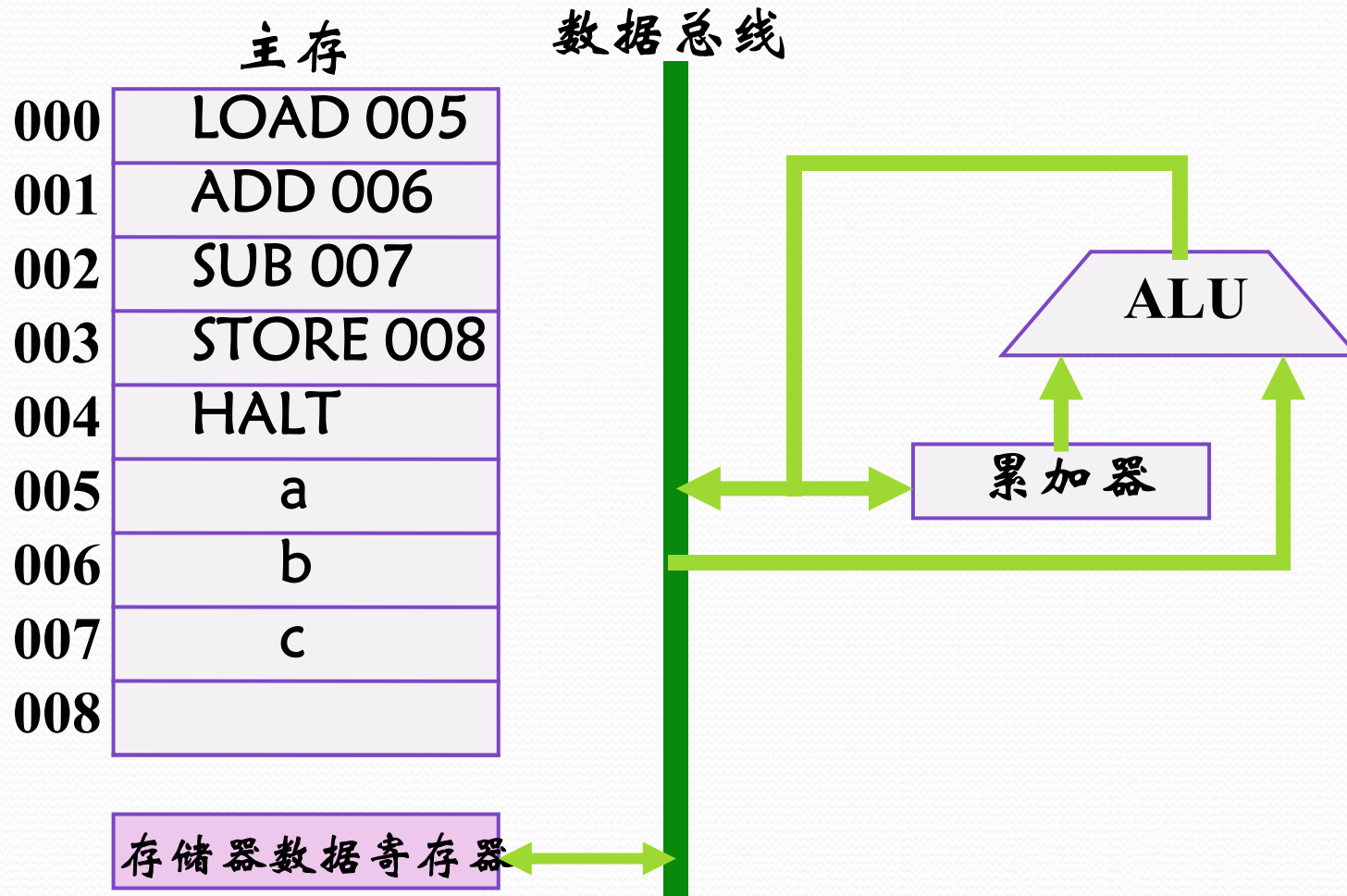
Add 006

Sub 007

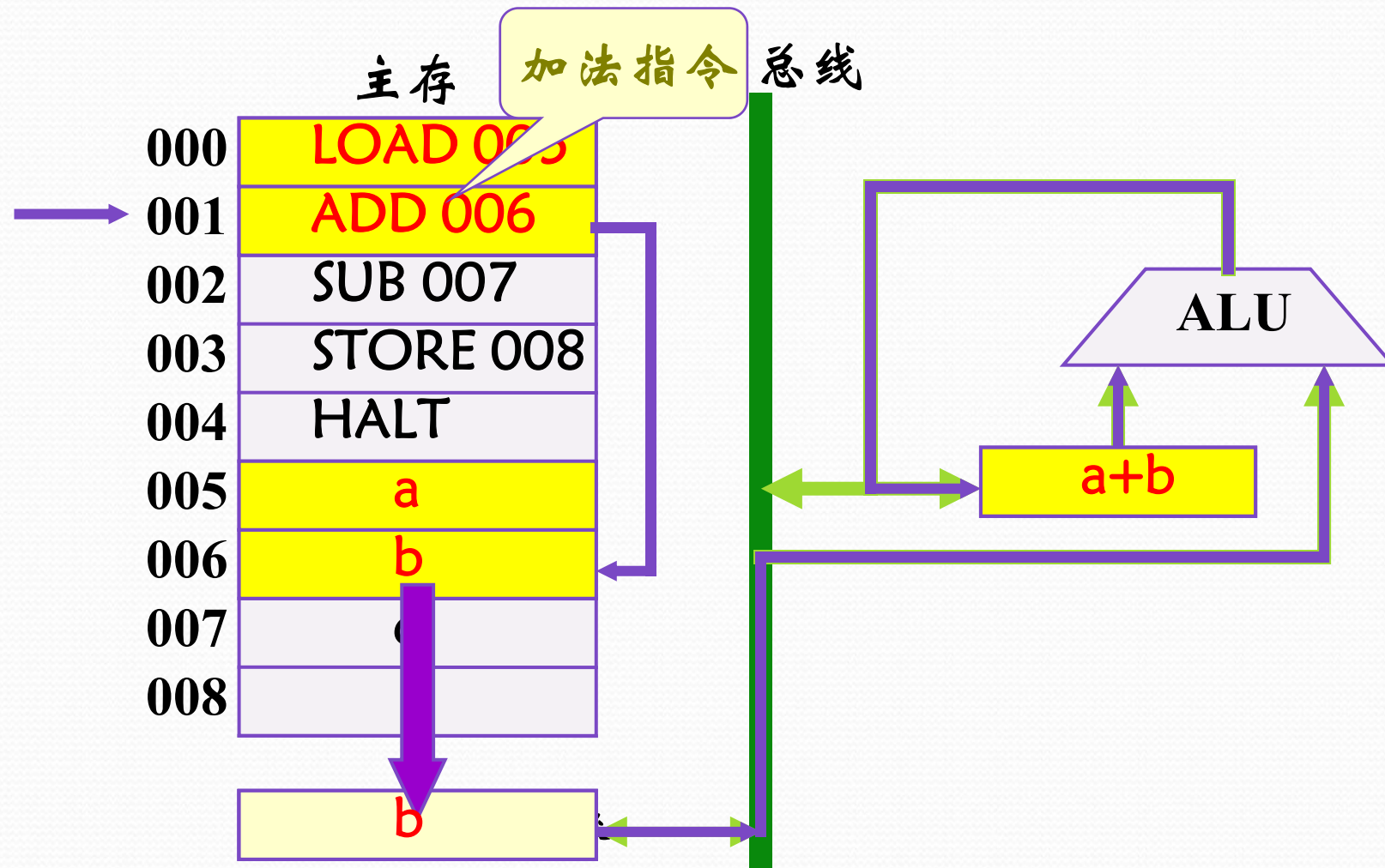
Store 008

HLT

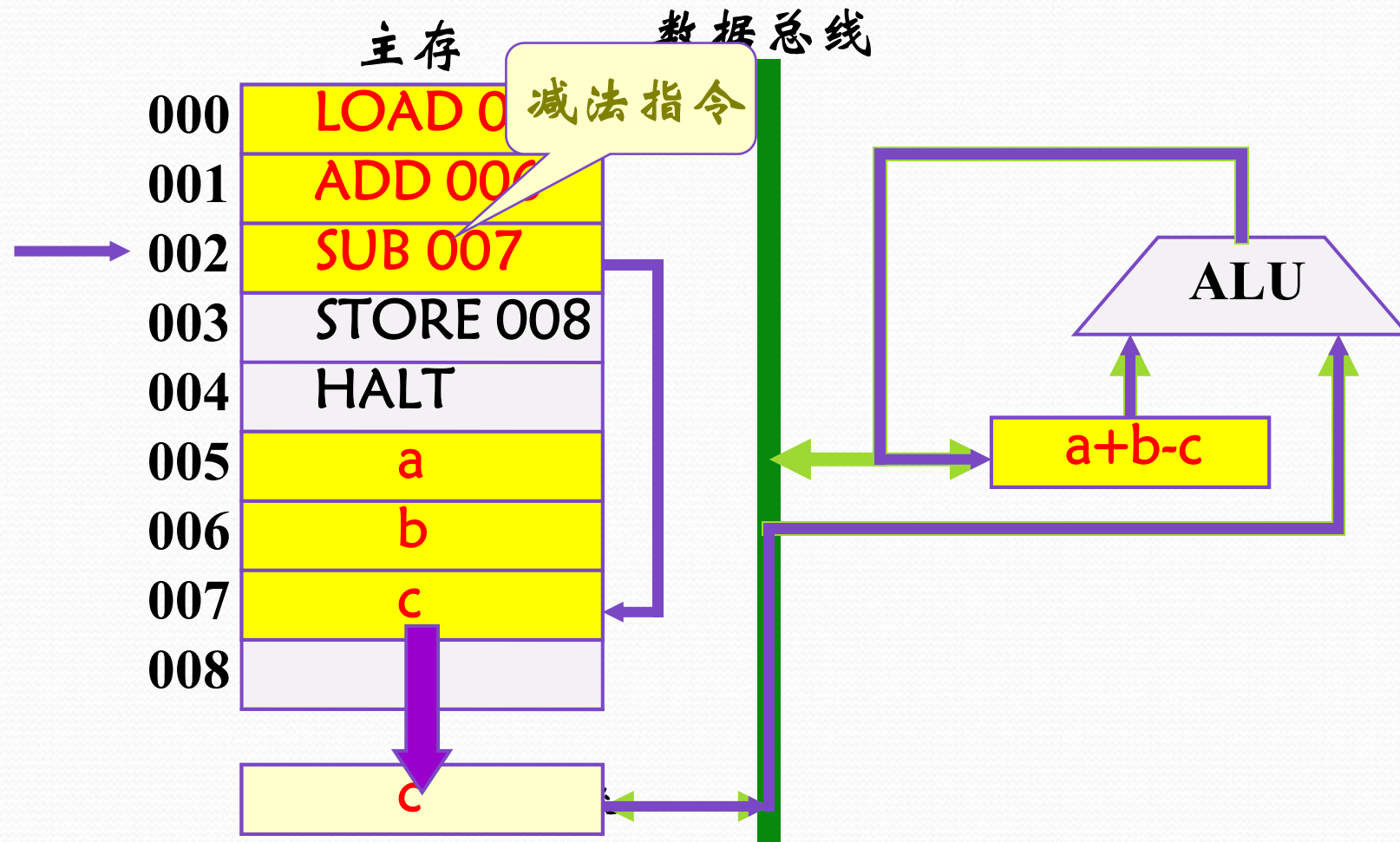
计算机的工作过程



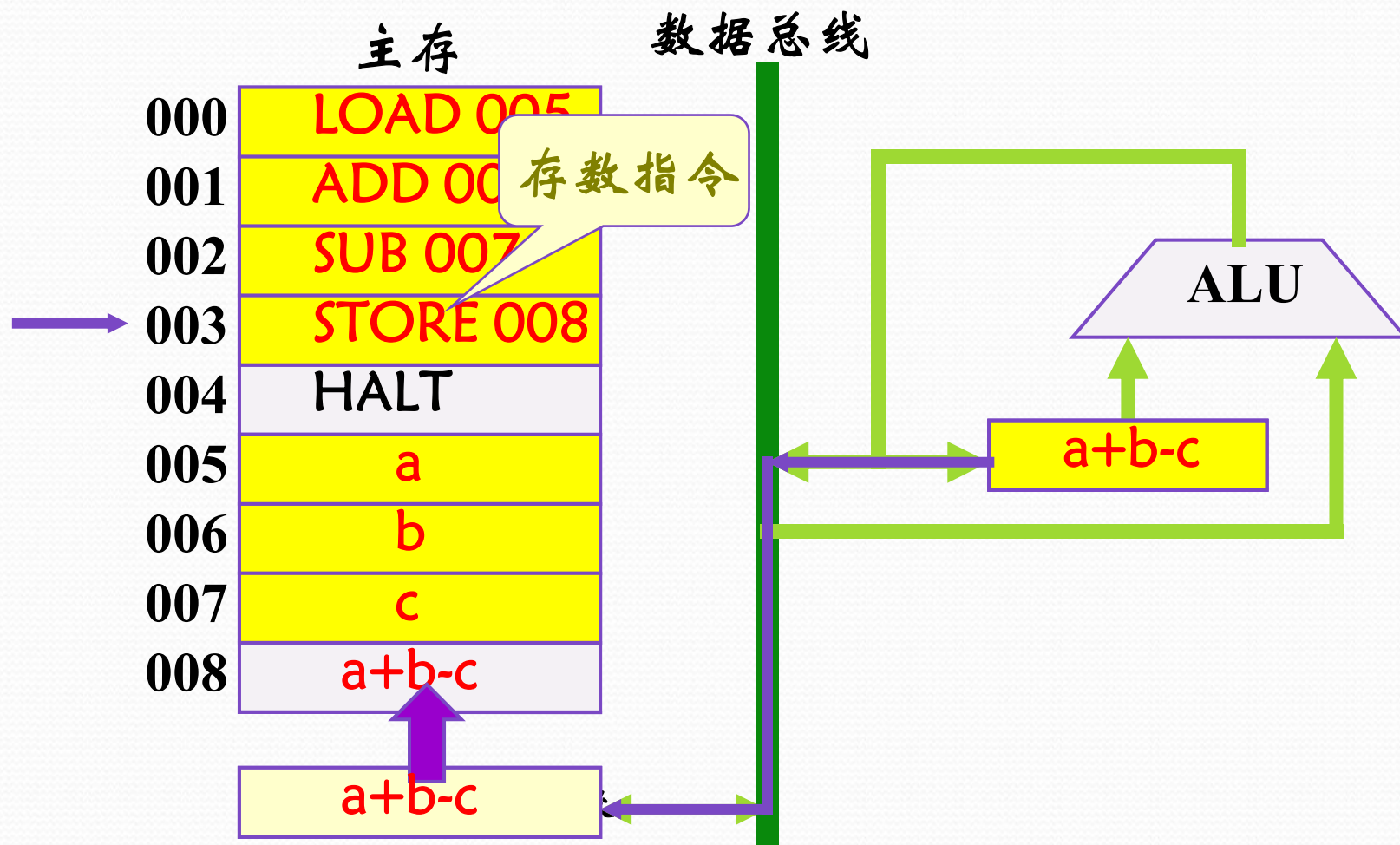
计算机的工作过程



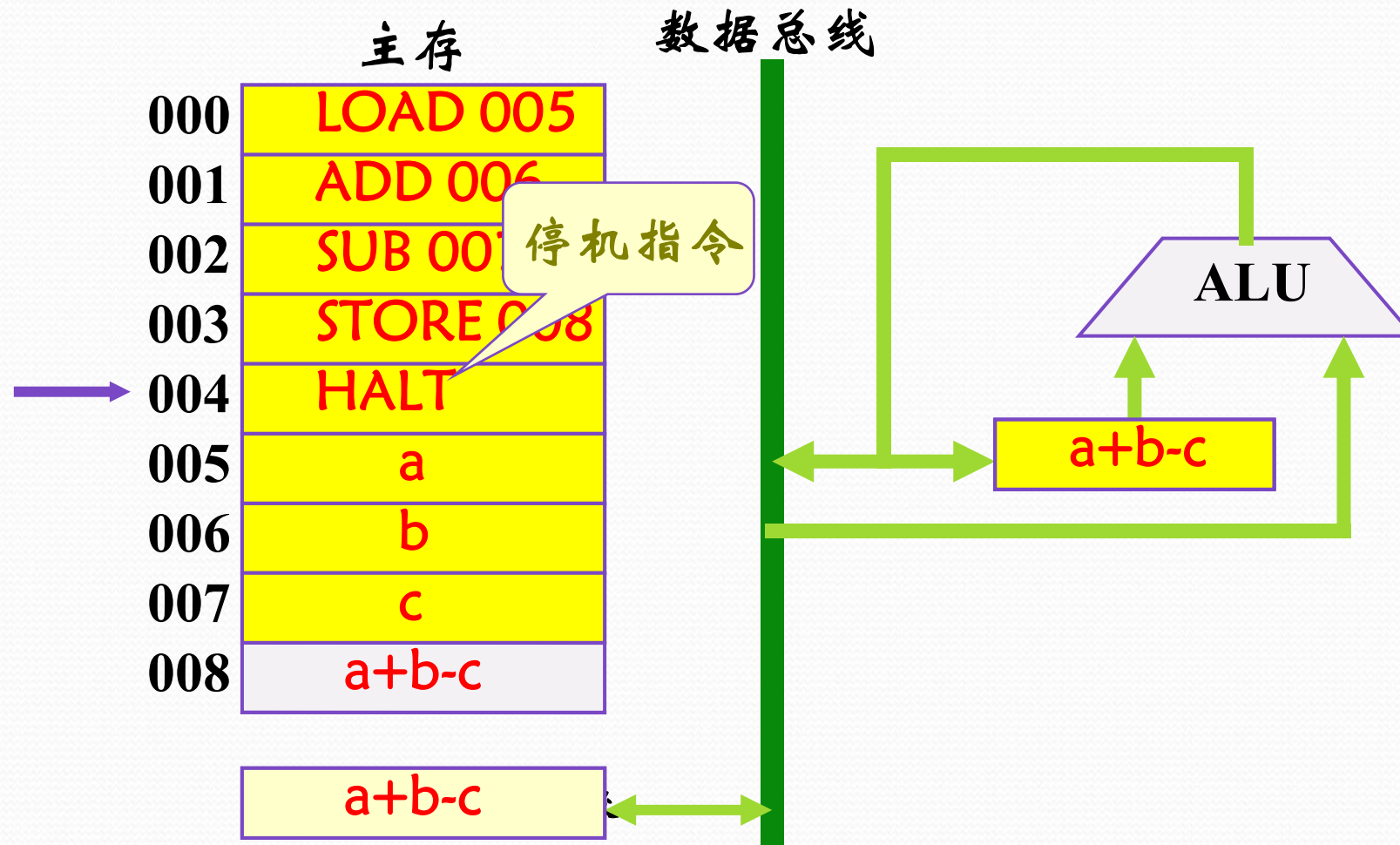
计算机的工作过程



计算机的工作过程



计算机的工作过程



基本字长

- 基本字长是指参与运算的数的基本位数，它是由加法器、寄存器、数据总线的位数决定的。
- 在计算机中为了更灵活地表达和处理信息，许多计算机又以字节(Byte)为基本单位，一个字节等于8位二进制位(bit)。
- 不同的计算机，字(Word)可以不相同，但对于系列机来说，在同一系列中，字却是固定的，如80X86系列中，一个字等于16位；IBM303X系列中，一个字等于32位。

数据通路宽度

- 数据总线一次所能并行传送的位数，称为数据通路宽度。它影响到信息的传送能力，从而影响计算机的有效处理速度。CPU内部的数据通路宽度一般等于基本字长，而外部数据通路宽度则取决于系统总线。
- 内、外数据通路宽度相等的CPU有：Intel 8086、80286、80486等；
- 外部<内部的CPU有：8088、80386SX；
- 外部>内部的CPU有：Pentium等。

主存容量

- 一个主存储器所能存储的全部信息量称为主存容量。
衡量主存容量单位有两种：
 - ① 字节数。这类计算机称为字节编址的计算机。每1024个字节称为1K字节（ $2^{10}=1K$ ），每1024K字节称为1M字节（ $2^{20}=1M$ ），每1024M字节称为1G字节（ $2^{30}=1G$ ）。
 - ② 字数×字长。这类计算机称为字编址的计算机。如：
4096×16表示存储器有4096个存储单元，每个存储单元字长为16位。
- 计算机的主存容量越大，存放的信息就越多，处理问题的能力就越强。

常用的容量单位

十进制术语	缩写	数值	二进制术语	缩写	数值	数值差别
Kilobyte	KB	10^3	Kibibyte	KiB	2^{10}	2%
Megabyte	MB	10^6	Mebibyte	MiB	2^{20}	5%
Gigabyte	GB	10^9	Gibibyte	GiB	2^{30}	7%
Terabyte	TB	10^{12}	Tebibyte	TiB	2^{40}	10%
Petabyte	PB	10^{15}	Pebibyte	PiB	2^{50}	13%
Exabyte	EB	10^{18}	Exbibyte	EiB	2^{60}	15%
Zettabyte	ZB	10^{21}	Zebibyte	ZiB	2^{70}	18%
Yottabyte	YB	10^{24}	Yobibyte	YiB	2^{80}	21%

运算速度

- 常用MIPS和MFLOPS来衡量运算速度。
- MIPS表示每秒百万条指令，多用于衡量标量机。
- MFLOPS每秒表示百万次浮点运算，MFLOPS比较适用于衡量向量机。
- 主频/时钟周期:主时钟的频率。
- CPU执行时间:表示CPU执行一段程序所占用的时间

运算速度

- CPU执行时间：表示CPU执行一段程序所占用的时间，可用下式计算：

$$\text{CPU时钟周期数} \times \text{CPU时钟周期}$$

- CPI：每条指令的周期数，即执行一条指令所需要的平均时钟周期数。
- $\text{CPI} = \text{执行某段程序所需的CPU时钟周期数} \div \text{指令条数}$
- $\text{CPU执行时间} = \text{CPI} \times \text{IC} \times \text{CPU时钟周期}$
 $= \text{CPI} \times \text{IC} / \text{CPU时钟频率}$

运算速度

练习：某计算机的CPU主频为500MHz，CPI为5(即执行每条指令平均需要5个时钟周期)。假定某外设的数据传输率为0.5MB/s，采用中断方式与主机进行数据传送，以32位为传输单位，对应的中断服务程序包含18条指令，中断服务的其他开销相当于2条指令的执行时间。请回答下列问题，要求给出计算过程。在中断方式下，CPU用于该外设I/O的时间占整个CPU时间的百分比是多少？

解：

中断一次所用的时间 $t = 4B/r = 4B/0.5(MB/s) = 8\mu s$

CPU时间： $t_{CPU} = \text{指令条数} \times \text{CPI} / \text{主频} = (18+2) \times 5 / 500M = 0.2\mu s$

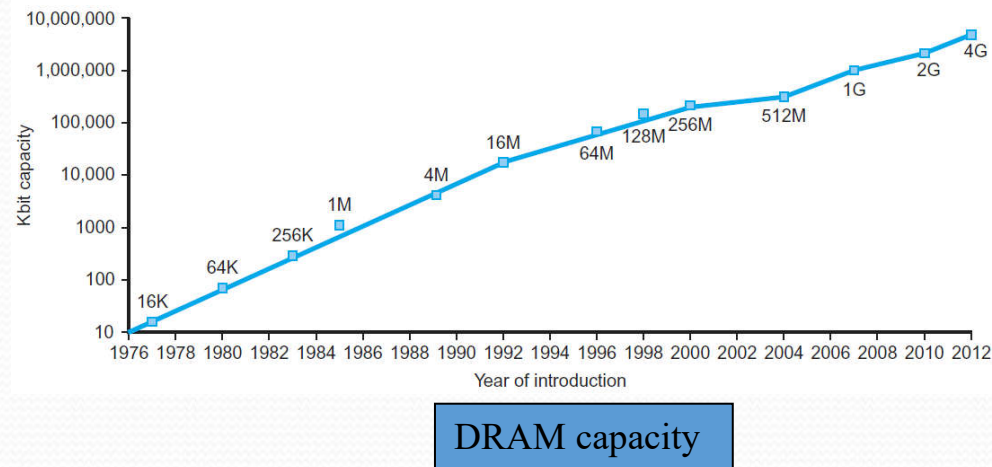
比例： $P = t_{CPU}/t = 0.2/8 = 2.5\%$

工作频率

- 工作频率是衡量计算机速度的重要参数。
- 工作频率又可以细分为内频和外频。
 - 内频，就是CPU的内部工作频率，内频也可称为内部时钟。
 - 外频，即外部工作频率，也就是主板的总线速度或系统时钟。
 - 早期，CPU的内频就等于外频。
 - 目前，CPU的内频越来越高，外频与内频不再是一比一的同步关系，从而出现了所谓的内部倍频技术，导致了“倍频”的出现。
- 内频、外频和倍频三者之间的关系是：
 - $\text{内频} = \text{外频} \times \text{倍频}$

Technology Trends

- Electronics technology continues to evolve
 - Increased capacity and performance
 - Reduced cost



Year	Technology	Relative performance/cost
1951	Vacuum tube	1
1965	Transistor	35
1975	Integrated circuit (IC)	900
1995	Very large scale IC (VLSI)	2,400,000
2013	Ultra large scale IC	250,000,000,000

第一章 小结

1.1 电子数字计算机与存储程序控制

- 存储程序概念
- 主机

1.2 计算机的类型和应用

- 系列机概念
- 软件兼容

1.3 计算机系统的硬件组成

- 计算机的五大基本部件
- 总线

第一章 小结

1.4 计算机系统的层次结构

- 计算机系统
- 硬件和软件的关系
- 固件的概念
- 虚拟机概念

1.5 计算机的工作过程和主要性能指标

- 主要性能指标

基本字长、数据通路宽度、主存容量

作业:

P19: 1, 2, 4, 6, 8

谢谢！