



# 计算机组成原理

北京邮电大学 计算机学院

易秋萍

yiqiuping@bupt.edu.cn





# 课前调查1

- 你现在使用什么设备进入课堂的
  - A 台式机
  - B 笔记本
  - C 平板电脑
  - D 手机





# 课程计划

总 学 时	64 学时
课堂教学	48 学时
实      验	16 学时
作      业	6-8次
考      试	期中考试+期末考试（闭卷）
成      绩	硬件实验20%、期中考试10%、平时作业10%、期末考试60%
答      疑	课间+课后





# 课程内容

- 计算机概述
  - ◆ 性能指标、层次结构
- 数据表示与运算器
  - ◆ 数据的表示、运算器
  - ◆ 浮点运算方法
- 存储器
  - ◆ 存储器系统、存储器控制器
  - ◆ Cache、虚拟存储器
- 指令系统
  - ◆ 指令集架构、寻址方式
  - ◆ 典型指令
- 中央处理器
  - CPU模型机、指令执行过程、微程序控制器
    - ◆ 硬布线控制器
- 总线
  - ◆ 总线分类、总线仲裁
  - ◆ 总线定时
- 外围设备
  - ◆ 磁盘、显示
- 输入输出系统
  - ◆ 程序查询、中断、DMA
  - ◆ 通道方式





# 实验内容

1. 运算器组成
  2. 双端口存储器
  3. 数据通路
  4. 微程序控制器
  5. CPU组成与机器指令的执行
  6. 中断原理
- 实验设备： TEC-8





# 课程目的

- 了解计算机的组成
  - ◆ 构成计算机的五大部件
- 掌握计算机的运行机制
  - ◆ 计算机是如何执行程序
- 理解计算机中处理器、存储器、输入/输出设备等部件的工作原理
- 建立软件与硬件之间的关联、层次化结构计算机系统的概念
- 通过本课程的理论和实践学习，掌握单处理机系统的组成原理和实现方法，建立起完整的计算机系统概念，初步具备硬件的设计与调试能力
- 成为计算机“大牛”





# 课程目的

## ■ David A. [Patterson@CS. UCB](mailto:Patterson@CS.UCB)

- ◆ professor of computer science at the University of California, Berkeley since 1976
- ◆ noted for his pioneering contributions to RISC design, and by leading the Berkeley RISC project
- ◆ becoming a distinguished software engineer at Google, 2016

“软件系统性能取决于软件设计者对系统硬件的理解程度。操作系统设计者、更多的软件工程师需要有较强的计算机组织与设计的背景知识”

程序执行结果不仅取决于算法和程序设计，而且取决于其运行时环境和计算机体系结构！

不明白





# 学习方法

- 课堂听课
- 课前预习
- 课后消化所学内容
- 阅读参考资料
- 完成习题
- 完成实验及实验报告

专业课学习的本质是：理解原理，  
而不是刷题！





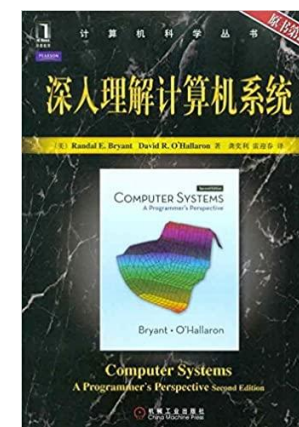
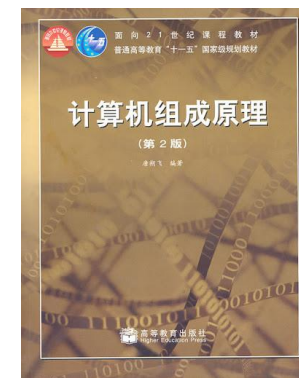
# 教材及参考书

## 教材

- 《计算机组成原理》（第6版）白中英等编 科学出版社

## 参考书

- 《计算机组成与设计》（第5版）David A. Patterson、John L. Hennessy著，机械工业出版社
- 《计算机组成原理》唐朔飞主编，高等教育出版社
- 《深入理解计算机系统》（美）Randal E. Bryant 等第2、3、4、6、9章





# 处理器内核高危漏洞

2018年  
热点事件

学习完本课程后，可以  
基本理解该漏洞的原理！

- **Meltdown（熔断）、Spectre（幽灵）**
- 国家信息安全漏洞共享平台（**CNVD**）对其评级为“高危”
- 攻击者利用这两组漏洞，在一定条件下可绕过内存访问的安全隔离机制获取密钥等隐私信息
- 近20年的**Intel、AMD、高通**厂家和其它**ARM**处理器均受到影响，影响面甚广
- 攻击原理：
  - ◆ 现代处理器采用乱序执行（**Out-of-Order Execution**）和推测执行（**Speculative Execution**）加速程序执行，这两种执行在遇到异常或分支预测错误时，**CPU**会丢弃之前执行结果，将 **CPU**状态恢复到乱序执行或预测执行前的正确状态，然后选择对应正确的指令继续执行，但此时并没有恢复缓存的内容，黑客可通过旁道攻击（**Side Channel Attack**）获取到缓存的敏感内容

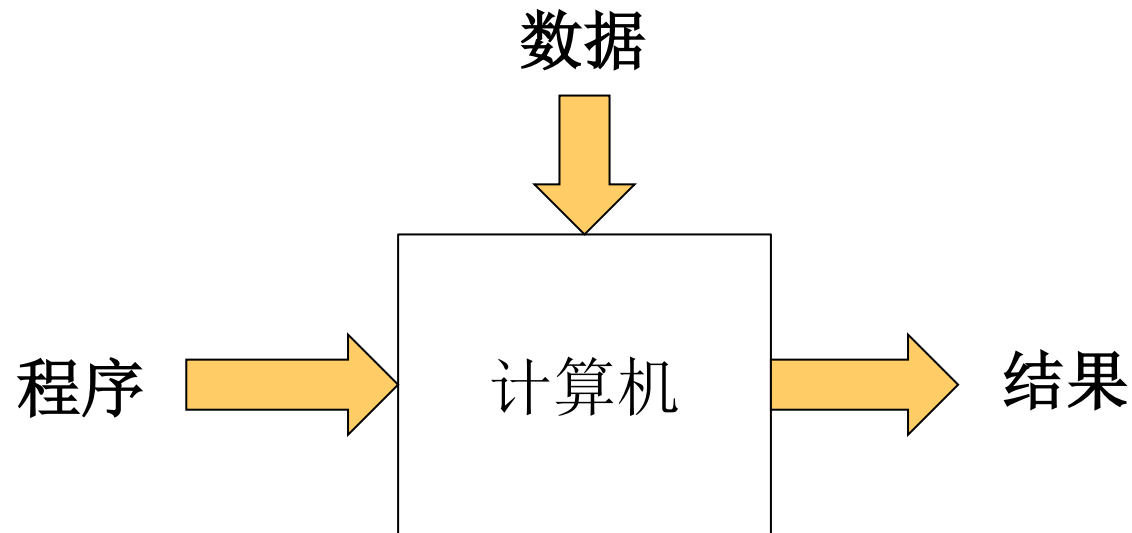


# 第1章 计算机概述

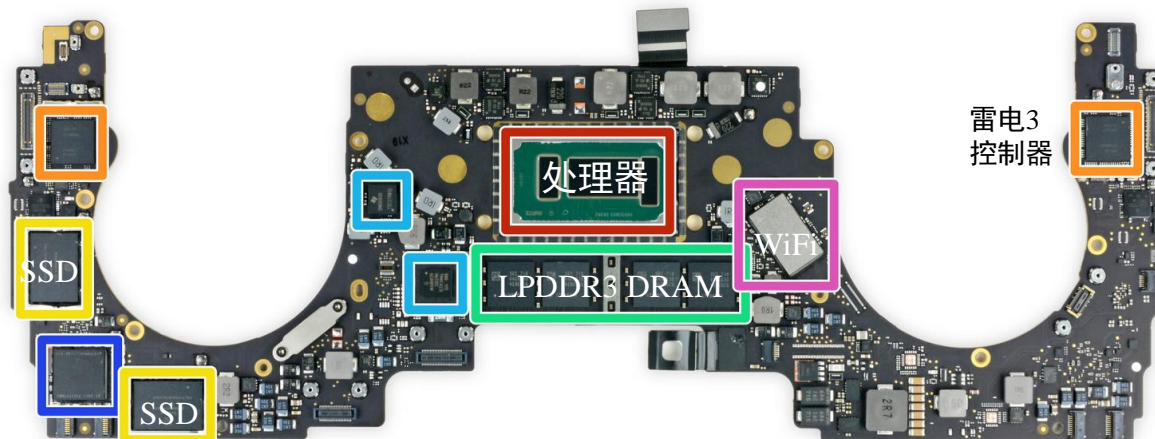
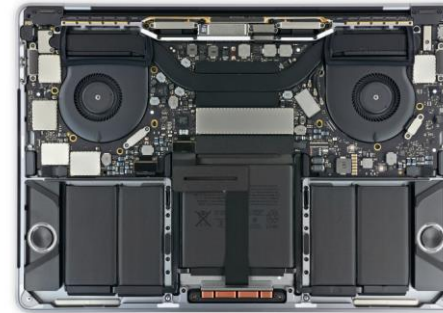




# 计算机在干啥？

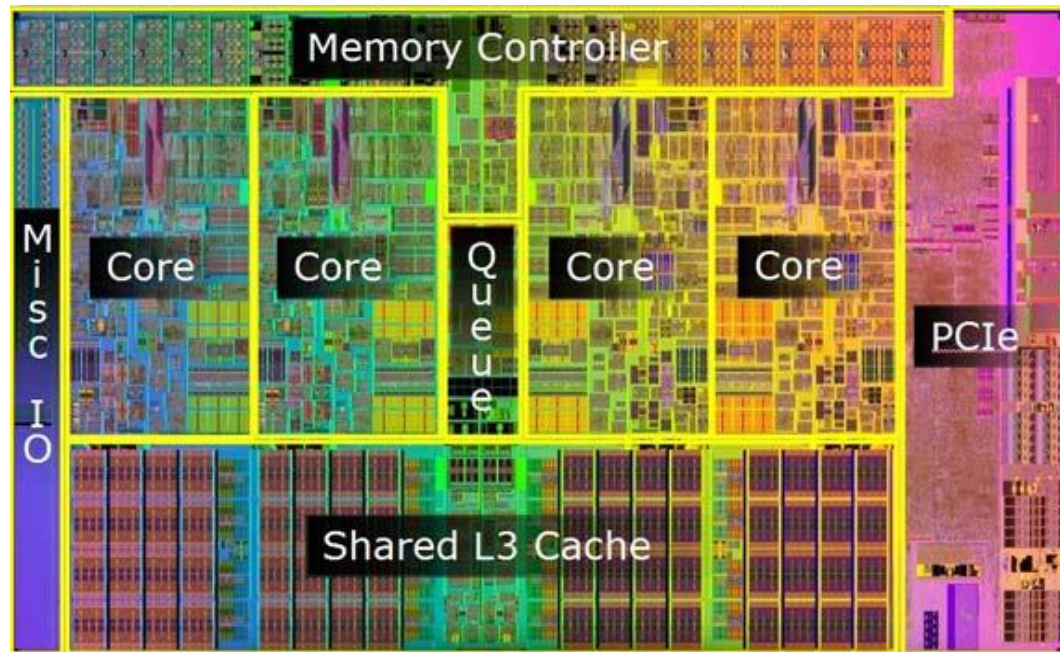


# 计算机里都有啥？





# 处理器内部视图



Core i7 Mobile Processor Die Map



# 1.1 计算机的分类和应用

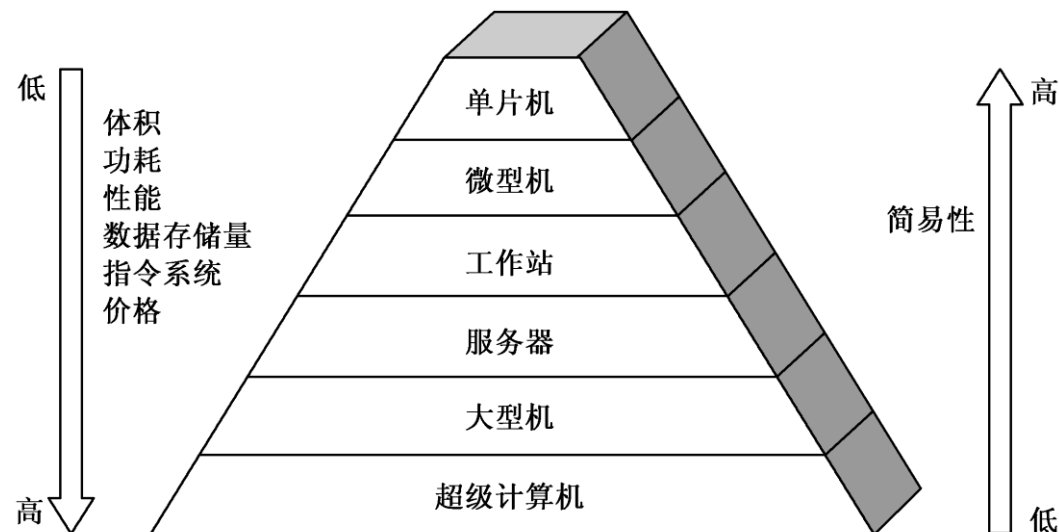






# 通用计算机分类

超级计算机、大型机、服务器、工作站、微型机、单片机







# 面向应用的分类

## ■ 通用计算机

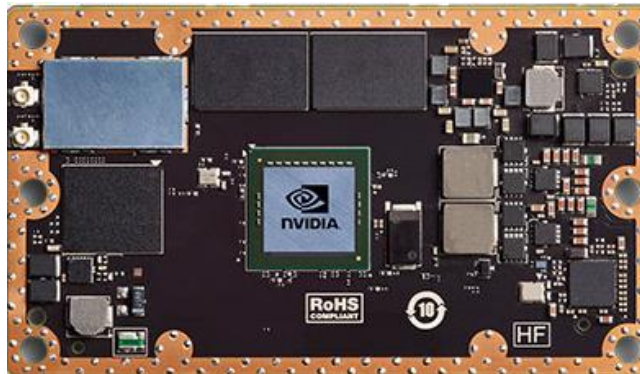
具有计算机的标准形态，通过安装不同的应用软件，以相似的外观呈现并应用在各行各业，例如：**你的PC机**

## ■ 专用计算机/嵌入式（Embedded）计算机

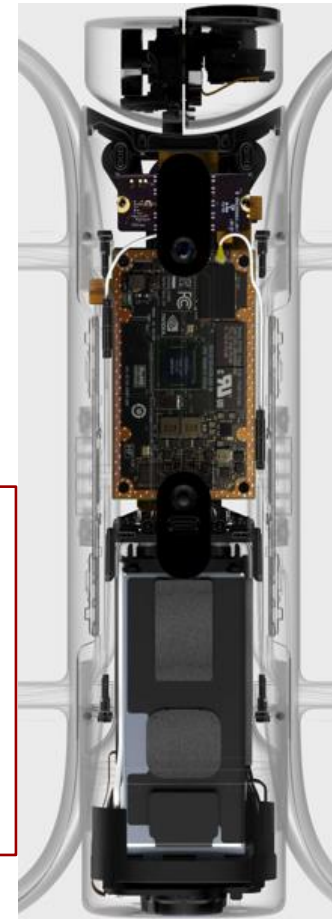
安装或嵌入到交通工具、仪器仪表、控制系统、通信设备和家电产品等中的模块化计算机，例如：



## e.g. Autonomous Drone With AI and CV



嵌入式系统:  
4核64bit ARM  
256核GPU  
4GBRAM  
64GB  
12x cameras  
WiFi  
...





# 课前调查

- 课堂直播效果
  - A 画面不卡顿、声音清晰
  - B 画面卡顿
  - C 画面卡顿，家里网络质量差
  - D 效果差，只能用手机上网





# 计算机的应用领域

- 科学计算：密码破译、天气预报、地质勘探、卫星轨道计算
- 信息处理：信息数据处理、搜索引擎、文字处理、事务处理、信息管理
- 实时控制：程控交换机、路由器、自动驾驶、测控系统
- 人工智能：自然语言处理、模式识别、机器学习
- 计算机辅助设计：电路设计、模拟仿真
- 娱乐和游戏



# 1.2 计算机发展简史





# 计算机的演进

- 第0代 机械时代 (BC-1940)
- 第1代 电子管时代 (1946-1957)  
耗电高, 体积大, 定点计算, 机器语言, 汇编语言
- 第2代 晶体管时代 (1958-1964)  
变集中处理为分级处理, 浮点运算、高级语言
- 第3代 集成电路时代 (1965-1971)  
存储容量大, 运算速度快, 几十至几百万次/秒
- 第4代 超大规模集成电路时代 (1972—现在)
- 第5代 量子计算机、光计算机、生物计算机?

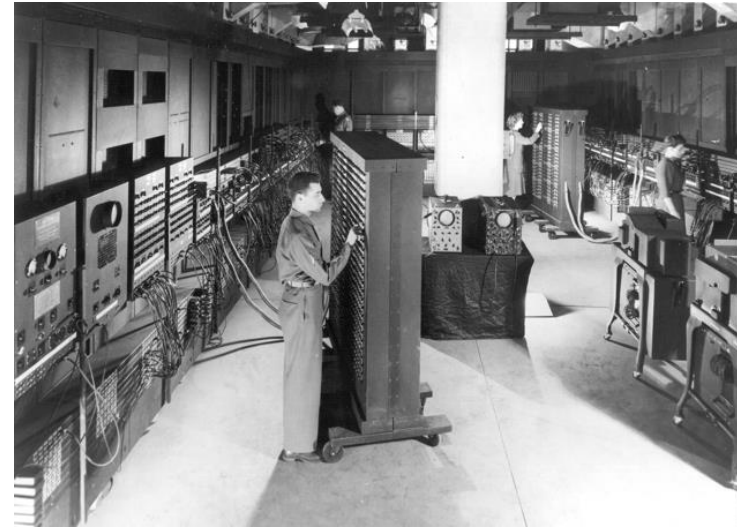


# 世界上第一台计算机ENIAC

1946美国宾夕法尼亚大学

**Mauchly** 和 **Eckert** 设计  
美国国防部用它来进行弹道计算

- 18000多个电子管
- 耗电功率150千瓦
- 170平方米，重30吨
- 5000次加法/秒
- 使用十进制数
- 20个累加寄存器，每个存放10位有符号十进制数
- 通过设置6000个开关和拔插电缆来实现编程
- 采用手动编程，通过设置开关和插拔电缆实现。



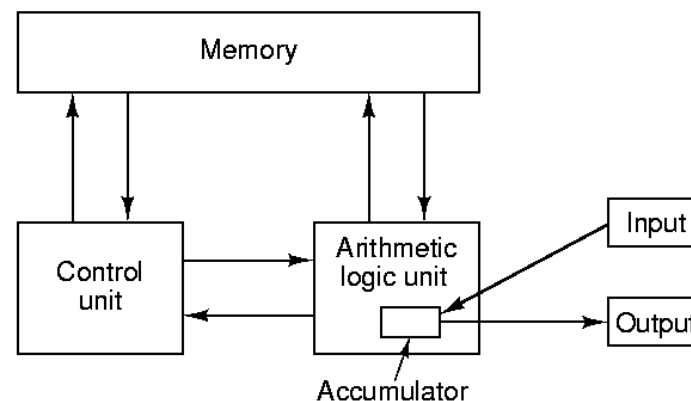
Electronic Numerical Integrate And Calculator





# IAS 计算机

- 1945 年冯·诺依曼 / 图灵同时提出“存储程序”思想，于 1946 年开始设计“存储程序”计算机，被称为 IAS 计算机。
- 也称为 Von Neumann 机。冯·诺依曼结构计算机
- 现代计算机体系结构的鼻祖：五大功能部件，是后来通用计算机的原型。
- 将程序和数据存储在同一个存储器中
- 采用二进制数据：简化了存储器设计
- 字长40位，存储器地址12位，指令长度20位
- 只实现整数运算：任何有能力的数学家都能在头脑中记住小数点的位置







# 计算机技术飞速发展

- 1946年 → 2019年，计算机处理速度：
  - ◆ 5000次加法/秒 → 14.86亿亿次/秒（超级计算机完成）
- 比尔盖茨曾在一个演讲中说到：假如GM的技术能象计算机技术那样发展，我们现在应该能用25美元买到一辆一加仑汽油跑1000英里的汽车
- GM反驳道：如果GM发展的技术象微软那样，那么现在开的汽车会有以下特点：
  - ◆ 你的汽车可能毫无道理地每天抛锚两次
  - ◆ 每条公路上重新画线，你都得买辆新的汽车
  - ◆ 有时候你的车在高速公路上莫名其妙地熄火，你必需Accept，然后Restart
  - ◆ 安全气囊弹出前将询问“ARE YOU SURE?” 要求您加以确认。
  - ◆ 有时候你的车在左转弯时突然熄火，而且无法Restart，所以你必需Reinstall发动机

.....



# 对比



- **1976 Cray1**
  - ◆ 250 M Ops/second
  - ◆ ~170,000 chips
  - ◆ 0.5B transistors
  - ◆ 5,000 kg, 115 KW
  - ◆ \$9M
  - ◆ 80 manufactured



- **2017 iPhone X**
  - ◆ > 10 B Ops/second
  - ◆ 16 chips
  - ◆ 4.3B transistors (CPU only)
  - ◆ 174 g, < 5 W
  - ◆ \$999
  - ◆ ~3 million sold in first 3 days



# 全球超级计算机排行（2021年11月最新榜单）

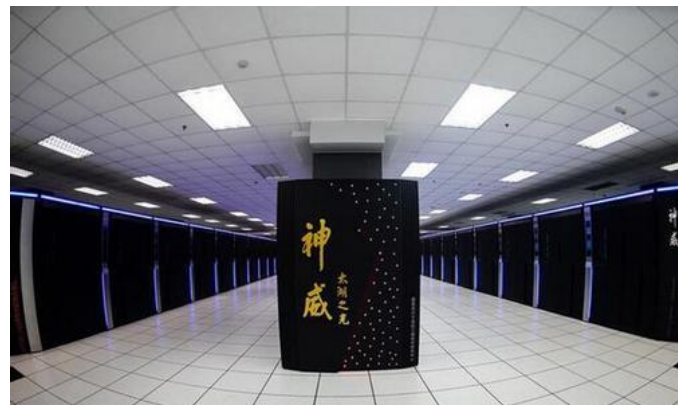
Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	<b>Supercomputer Fugaku</b> - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442,010.0	537,212.0	29,899
2	<b>Summit</b> - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148,600.0	200,794.9	10,096
3	<b>Sierra</b> - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
4	<b>Sunway TaihuLight</b> - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
5	<b>Perlmutter</b> - HPE Cray EX235n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Slingshot-10, HPE DOE/SC/LBNL/NERSC United States	761,856	70,870.0	93,750.0	2,589





# 2021年11月Top500排名第4

- 中国：神威·太湖之光
- 制造商：国家并行计算机工程技术研究中心
- 地点：国家超级计算无锡中心
- 实测性能：93,014.6 TFlop/s（9.3亿亿次/秒）
- 处理器：SW26010 260C 1.45GHz
- 处理器核数：10,649,600
- 内存容量：1,310,720GB
- OS：Sunway RaiseOS
- 互联方式：Sunway
- 功耗：15371kW





# IBM Blue Gene/L

$(2 \text{ processors/chip}) \cdot (2 \text{ chips/compute card}) \cdot (16 \text{ compute cards/node board}) \cdot (32 \text{ node boards/tower}) \cdot (64 \text{ tower}) = \underline{128k = 131072 (0.7 \text{ GHz PowerPC 440) processors (64k nodes)}}$

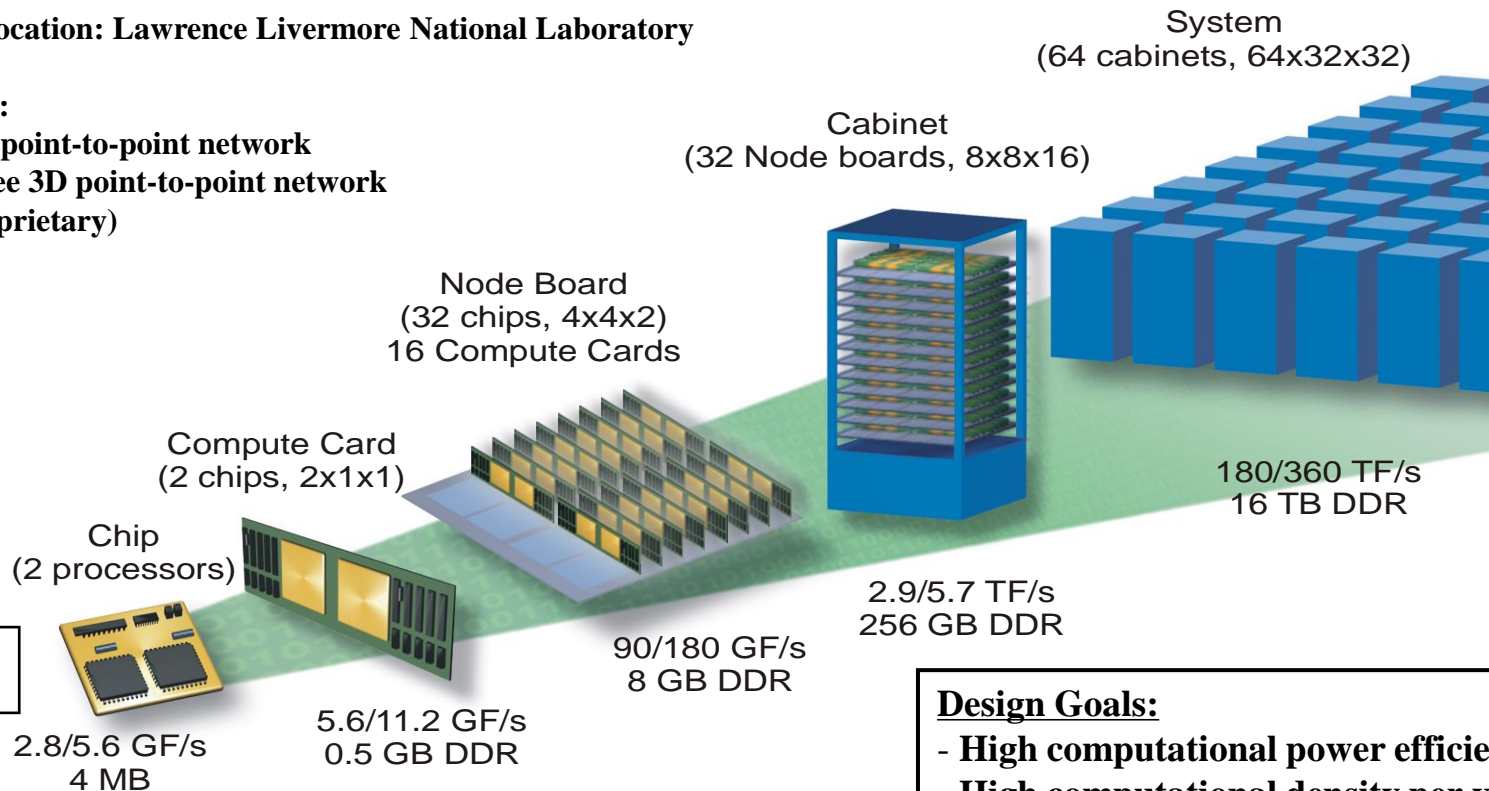
System Location: Lawrence Livermore National Laboratory

## Networks:

3D Torus point-to-point network

Global tree 3D point-to-point network

(both proprietary)



## Design Goals:

- High computational power efficiency
- High computational density per volume

## LINPACK Performance:

280,600 GFLOPS = 280.6 TeraFLOPS = 0.2806 Peta FLOP

## Top Peak FP Performance:

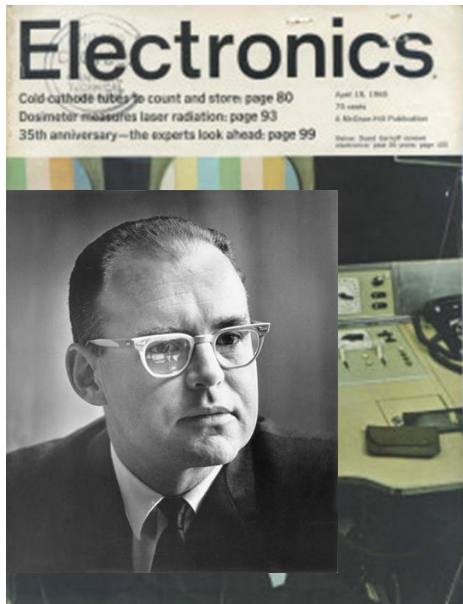
Now about 367,000 GFLOPS = 367 TeraFLOPS = 0.367 Peta FLOP





# 硬件定律

- 信号传输速度赶不上光速
- 内存赶不上CPU速度
- 软件赶不上硬件
- Moore's Law



April 19, 1965

## Cramming more components onto integrated circuits

With unit cost falling as the number of components per circuit rises, by 1975 economics may dictate squeezing as many as 65,000 components on a single silicon chip

By Gordon E. Moore

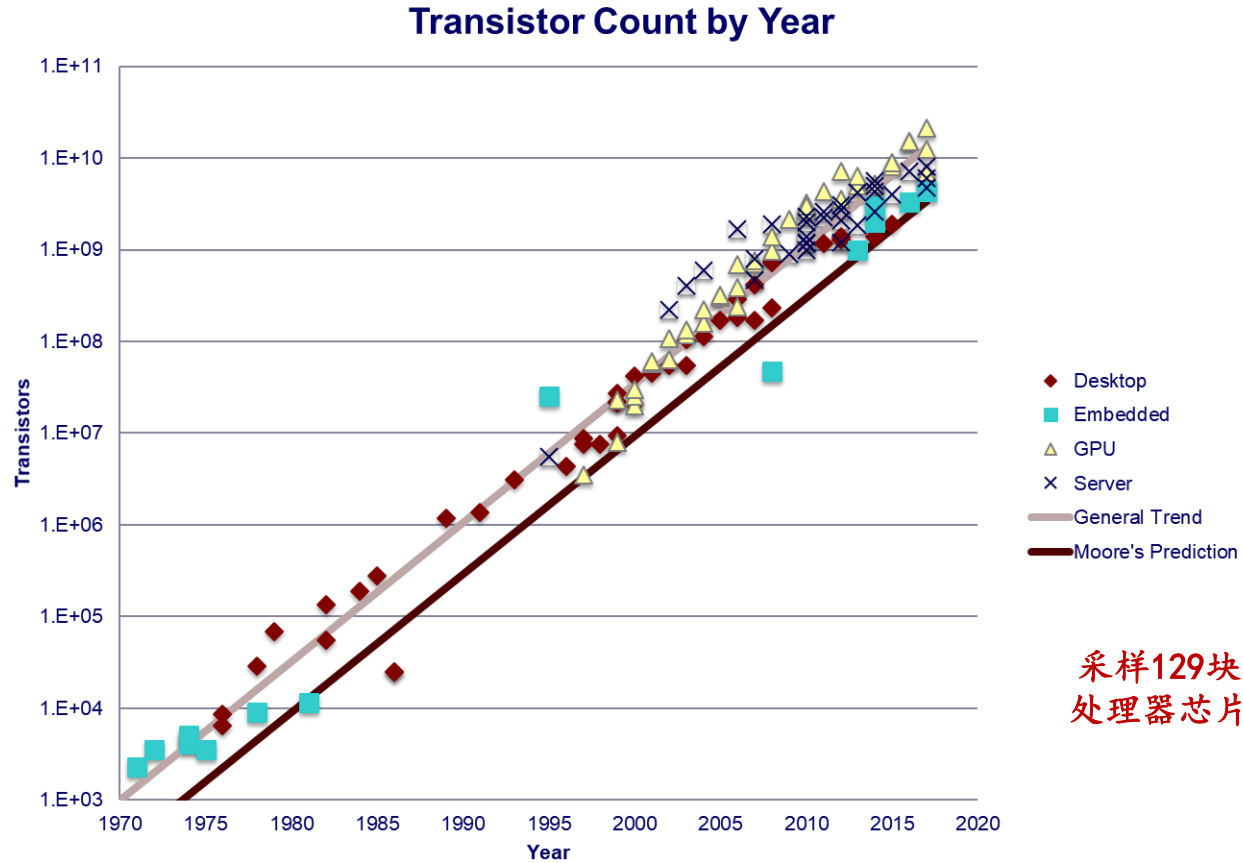
Director, Research and Development Laboratories, Fairchild Semiconductor division of Fairchild Camera and Instrument Corp.

- 摩尔定律：集成电路芯片中晶体管数量大体上每2年翻一番，或者价格下降一半





# Moore's Law: 50年





# 计算机的性能指标

容 量

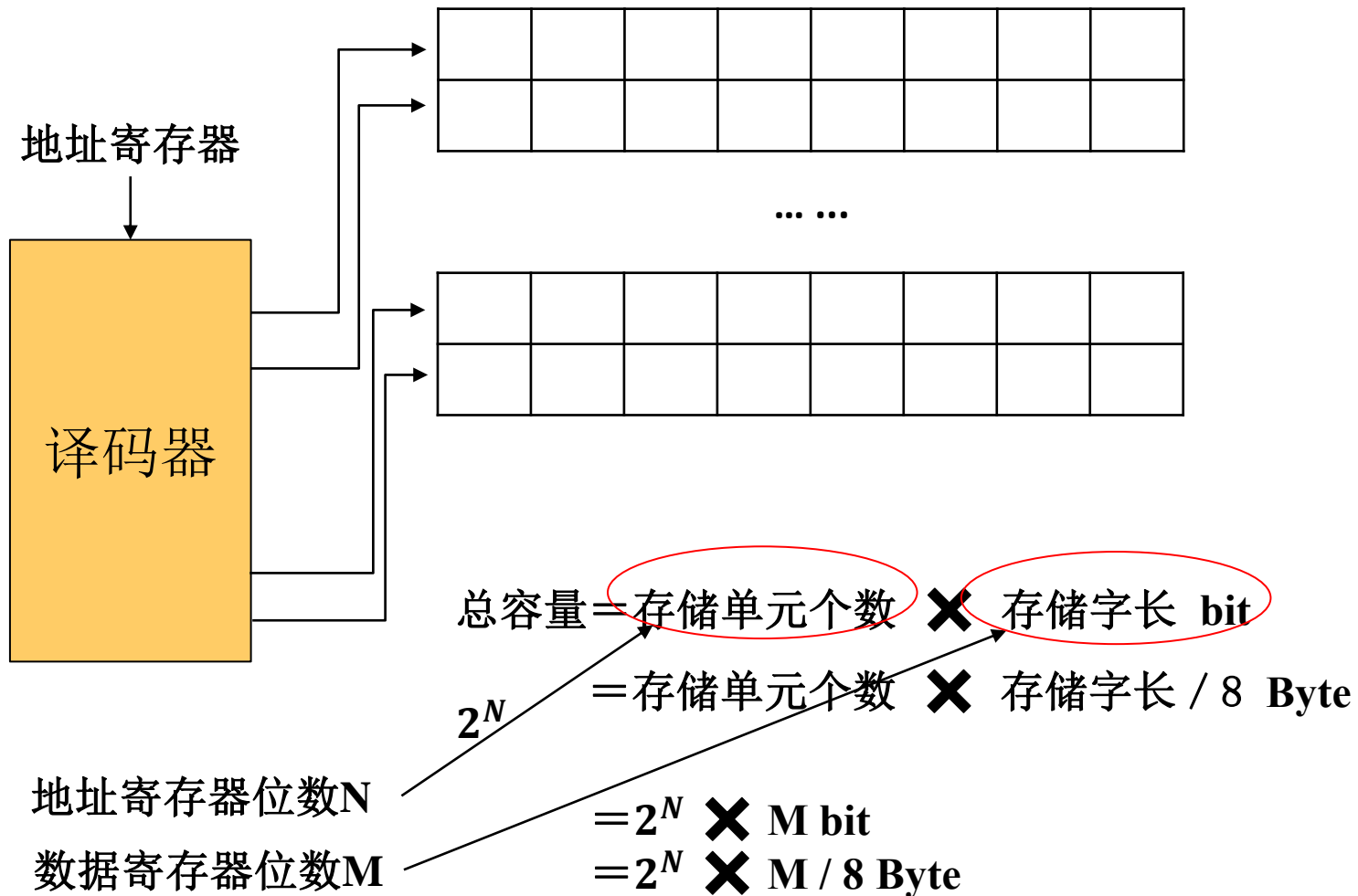
速 度







# 计算机性能指标（容量）





# 计算机性能指标（容量）

■  $n$ 个二进制位能表示多少种不同的状态？

1个二进制位	0, 1	$2^1$
2个二进制位	00, 01, 10, 11	$2^2$
3个二进制位	000, 001, 010, 011, 100, 101, 110, 111	$2^3$
...	...	
$n$ 个二进制位		$2^n$

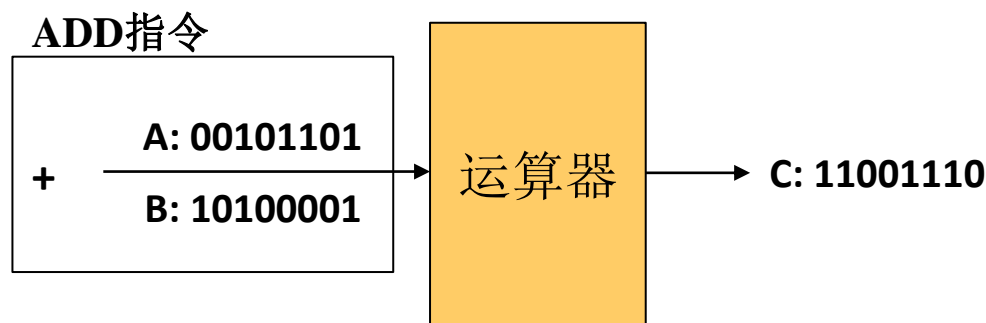
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2	4	8	16	32	64	128	256	512	1024	2048	4096	8192	16384	32768	65536

$2^{10}$ : K    $2^{20}$ : M    $2^{30}$ : G    $2^{40}$ : T

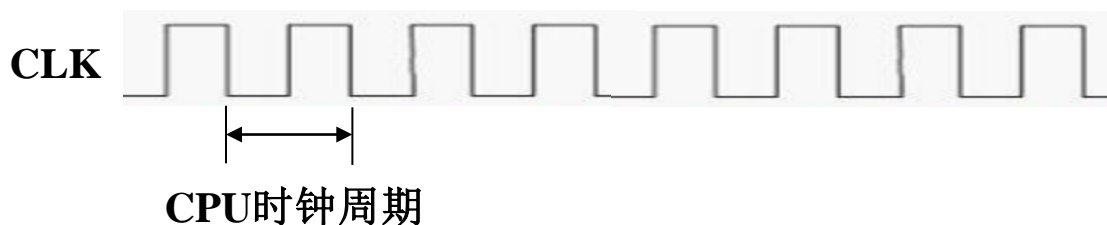




# 计算机性能指标（速度）



机器字长：指处理机运算器中一次能够完成二进制数运算的位数。  
（一般等于内部寄存器的位数）



$$\text{CPU时钟频率（主频）} = \frac{1}{\text{CPU时钟周期}}$$

CPU的工作节拍受主时钟控制，主时钟不断产生固定频率的时钟，主时钟的频率称为CPU的主频。

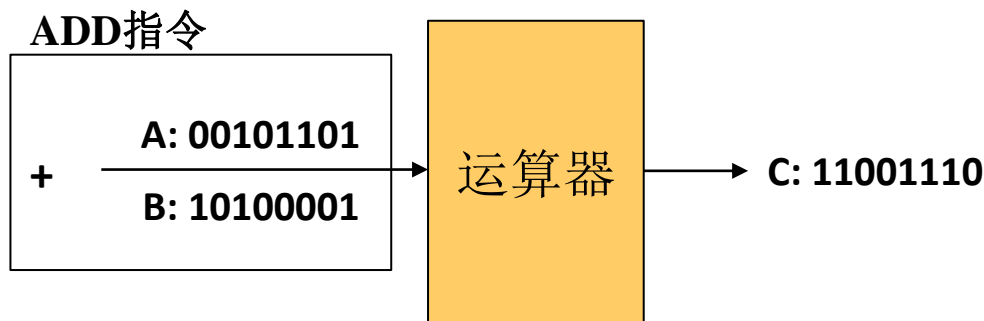
**CPI（Clock Cycles Per Instruction）：** 每条指令执行所需要的时钟周期数，或每条指令执行所需的平均时钟周期数

$$\text{指令耗时} = \text{CPI} \times \text{CPU时钟周期}$$

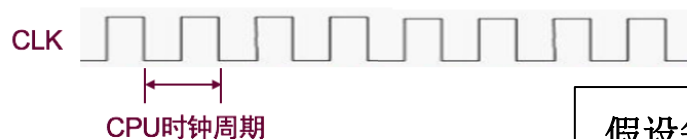




# 计算机性能指标（速度）



机器字长：指处理机运算器中一次能够完成二进制数运算的位数。  
（一般等于内部寄存器的位数）



$$\text{CPU时钟频率（主频）} = \frac{1}{\text{CPU时钟周期}}$$

CPI（Clock Cycles Per Instruction）：每条指令数，或每条指令执行所需的平均时钟周期数

$$\text{指令耗时} = \text{CPI} \times \text{CPU时钟周期}$$

假设每次只能处理一位，那完成8位运算，编程需要8条指令：

整个程序执行时间 = 指令1执行时间 + ... + 指令8执行时间

$$= (\text{指令1的CPI} + \dots + \text{指令8的CPI}) \times \text{CPU时钟周期}$$

$$= (\text{平均指令CPI} \times \text{指令条数}) \times \text{CPU时钟周期}$$

$$= (\text{平均指令CPI} \times \text{指令条数}) / \text{CPU时钟频率}$$





# 计算机性能指标（速度）

- MIPS: 每秒百万指令数

$$\text{MIPS} = (\text{指令总数} \div \text{程序执行时间}) \times 10^{-6}$$

- MFLOPS: 每秒**百万**次浮点操作次数

$$\text{MFLOPS} = (\text{程序中浮点运算次数} \div \text{程序执行时间}) \times 10^{-6}$$

- GFLOPS: 每秒**十亿**次浮点操作次数

$$\text{GFLOPS} = (\text{程序中浮点运算次数} \div \text{程序执行时间}) \times 10^{-9}$$

- TFLOPS: 每秒**万亿**次浮点操作次数

$$\text{TFLOPS} = (\text{程序中浮点运算次数} \div \text{程序执行时间}) \times 10^{-12}$$





# 计算机性能指标（速度）

- 总线宽度：数据总线**一次**所能并行传送信息的位数
- 存储器带宽：**单位时间内**从存储器读出的字节数，一般用字节数/秒表示。
- 吞吐量：表征一台计算机在**某一时间间隔**内能够处理的信息量，单位是字节/秒（B/S）
- 响应时间：指从用户向计算机发送一个请求，到系统对该请求作出响应并获得它所需要的结果的等待时间
  - ◆ 包括CPU时间与等待时间
- 利用率：在给定的时间间隔内系统被实际使用的时间所占的比率，用百分比表示





# 例 题

某计算机主频为1.2GHz，其指令分为A、B、C、D四类，对应的CPI依次为2、3、4、5，我们可通过运行一基准程序来测定该机的执行速度，已知在这段基准程序中A、B、C、D这四类指令所占比例依次为50%、20%、10%、20%，请问该机的MIPS数是多少？

解:设该基准程序为P，指令条数为n，其平均CPI为：

$$\text{CPI} = (0.5n \times 2 + 0.2n \times 3 + 0.1n \times 4 + 0.2n \times 5) / n = 3$$

故该机的MIPS数为：

$$1.2 \times 10^9 / 3 \times 10^{-6} = 400 \text{ MIPS}$$



# 1.2 计算机的硬件

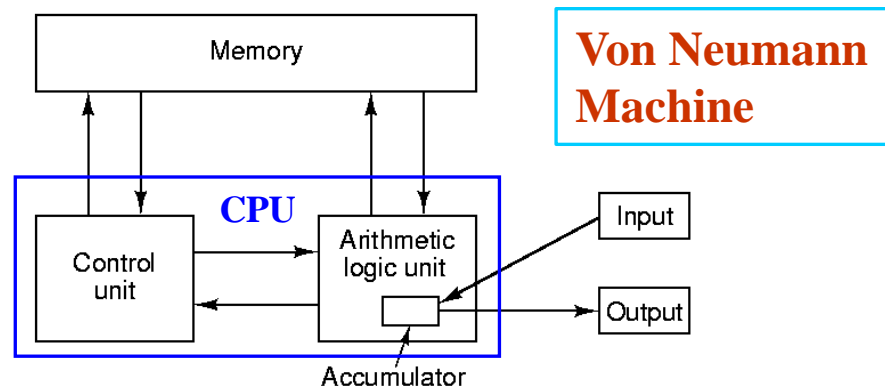






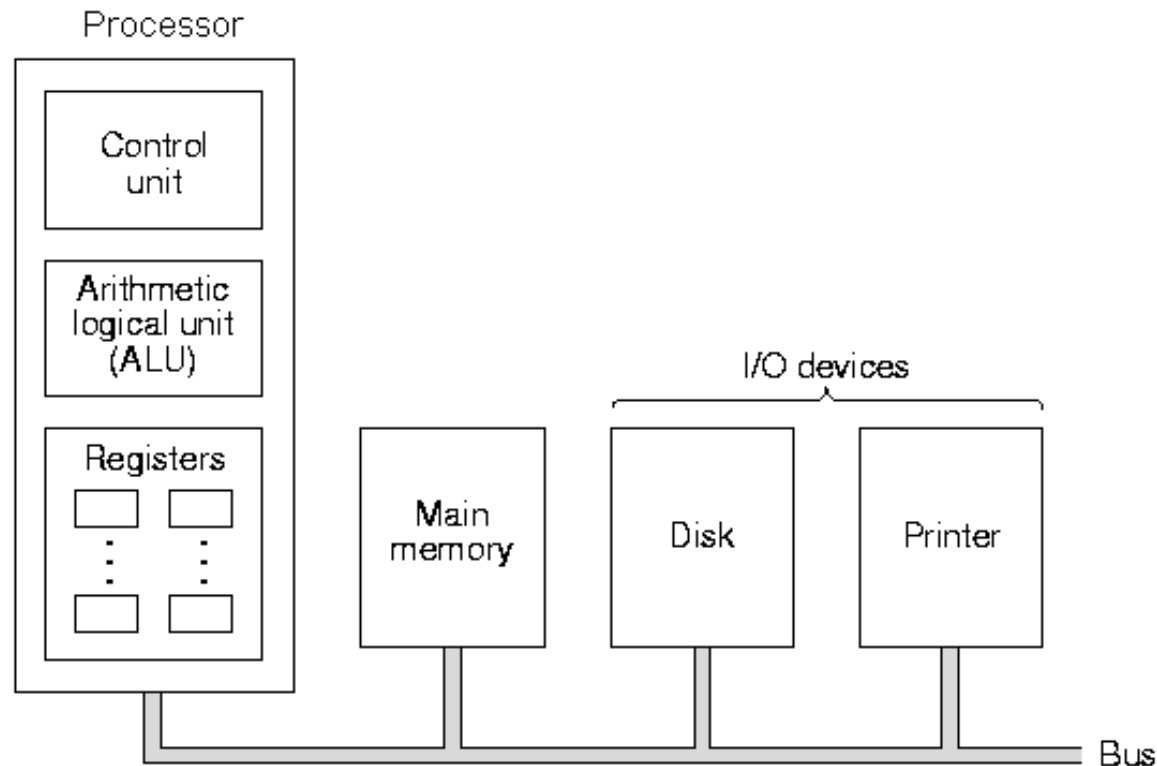
# 计算机组成

- 计算机由运算器、控制器、存储器、输入设备和输出设备五大部件组成
  - ◆ 存储器：存放程序和数据
  - ◆ 控制器：根据取得的指令向其他部件发出控制信号，完成指令规定操作
  - ◆ 运算器：完成算术和逻辑运算操作，也称为数据通路
  - ◆ 输入/输出设备：完成人与计算机的相互通信
- 在机器内部，指令和数据均以二进制码表示
  - ◆ 指令由操作码和地址码组成
- 机器以运算器为中心，数据传送都经过运算器
- 采用存储程序的方式，编制好的程序和数据存放在同一存储器中，计算机自动完成逐条取出指令和执行指令的操作。故称之为**存储程序式计算机**



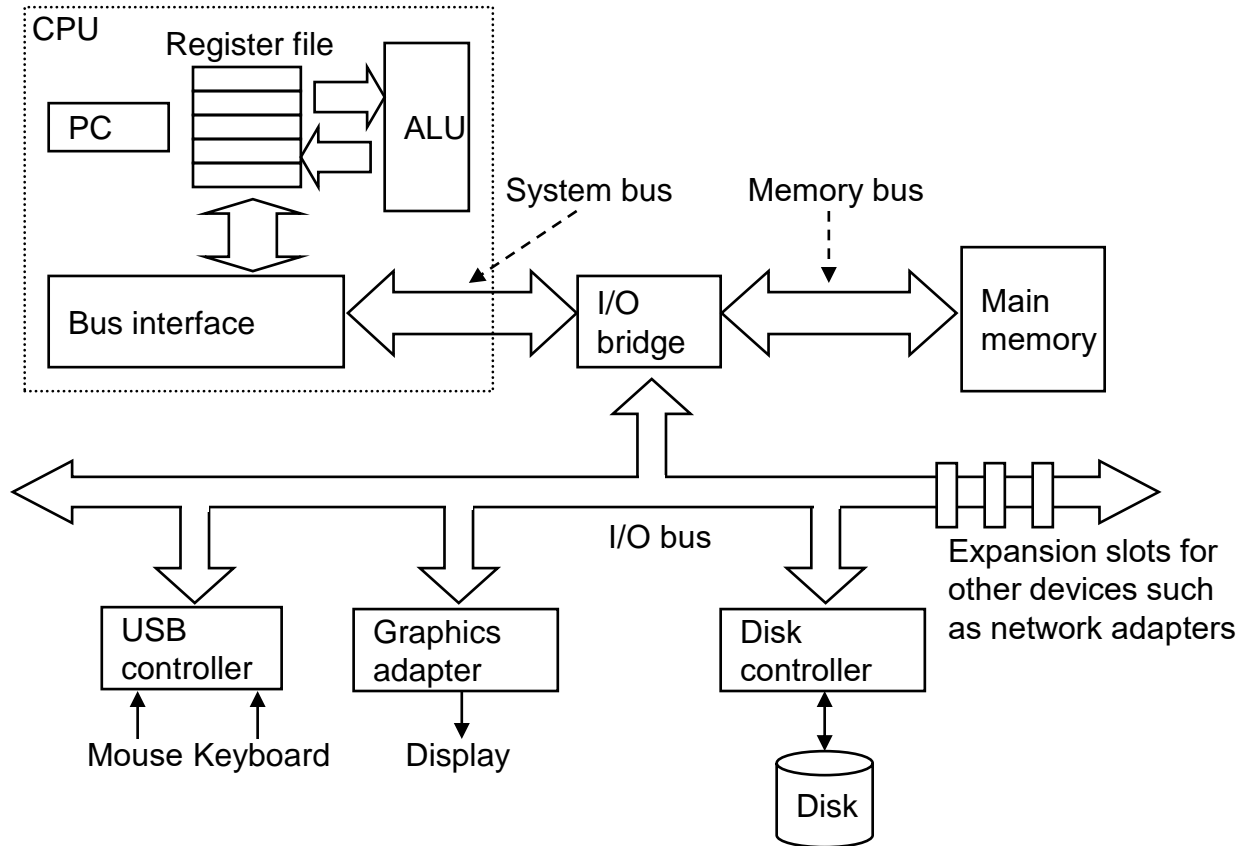


# 基于总线的计算机



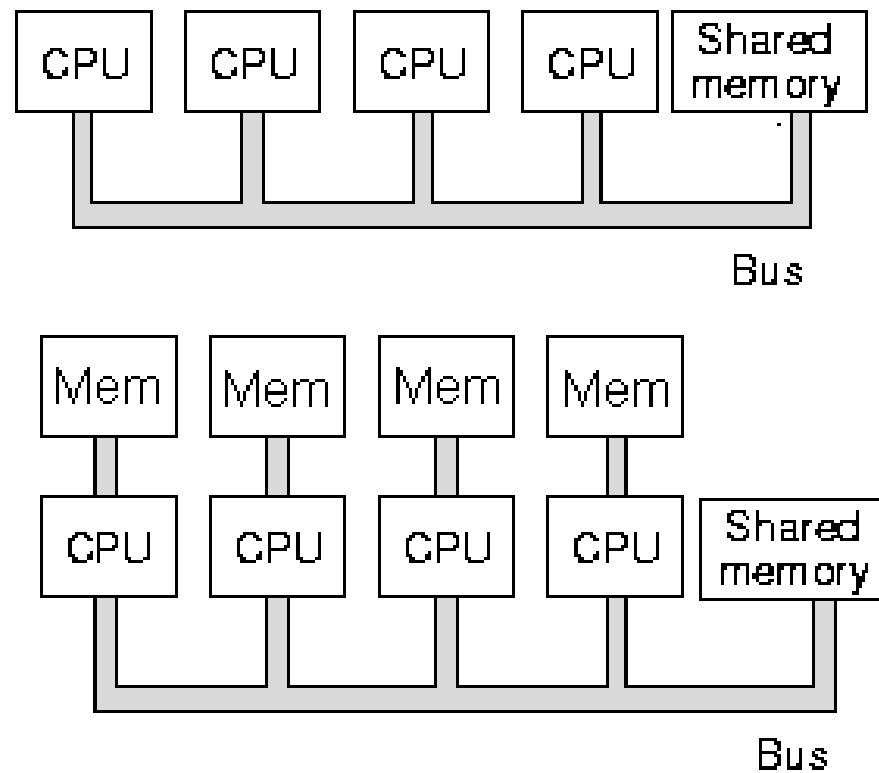


# 典型计算机的硬件组成





# 多处理器

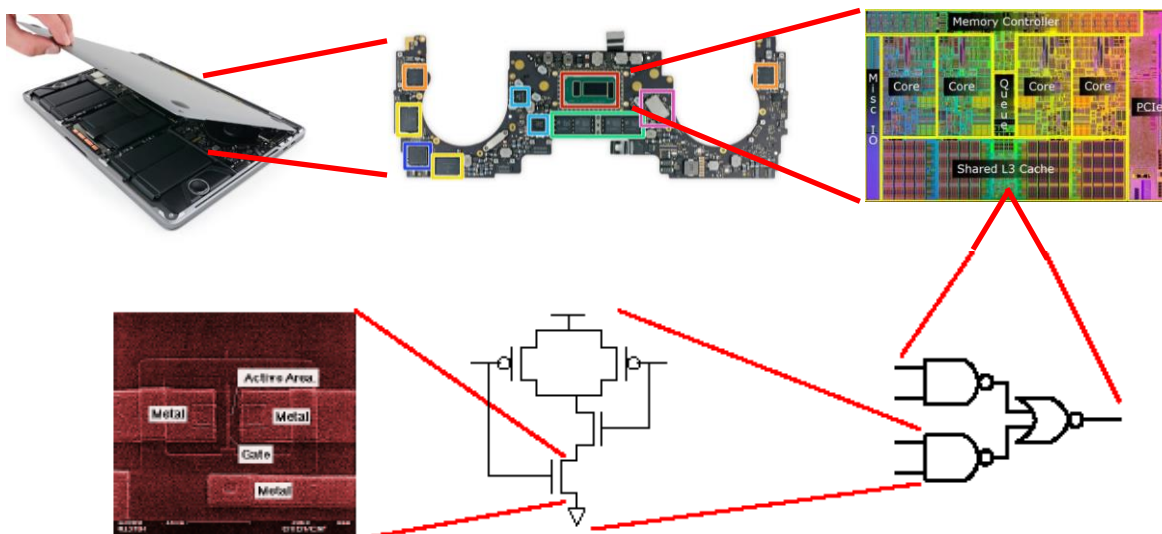


# 计算机硬件层剖析

计算机系统

计算机结构

处理器



设备制造

电路 (VLSI) 设计

逻辑设计

# 1.4 计算机的软件





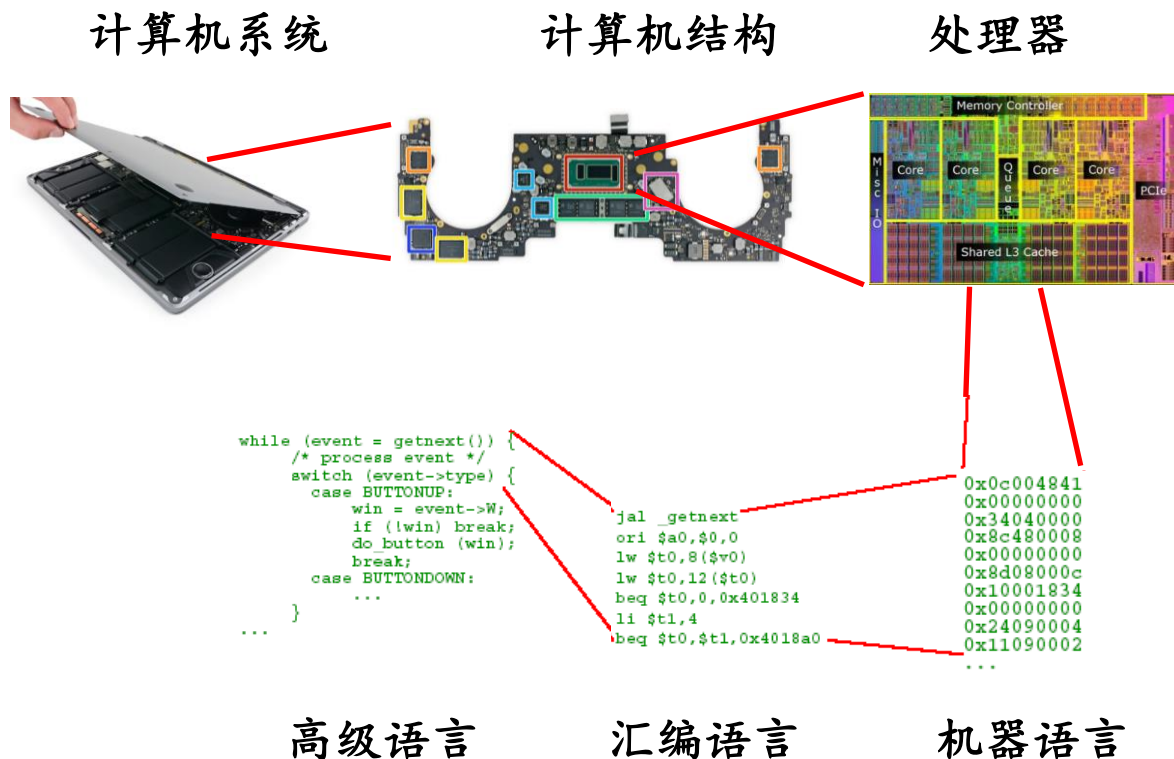
# 计算机软件分类

- 系统程序：管理整个操作系统，使系统资源得到合理调度
  - ◆ 操作系统
  - ◆ 调试程序、诊断程序
  - ◆ 汇编器、编译器、解释程序
  - ◆ 数据库管理系统
  
- 应用程序：完成用户的特定任务，使用系统软件提供的资源接口
  - ◆ EDA设计软件
  - ◆ 数据处理软件
  - ◆ 社交娱乐
  - ◆ 办公软件
  - ◆ ... ..





# 计算机软件层剖析







# C语言程序→汇编语言程序

sample.c

```
#include <stdio.h>
void sum(int x,int y,int* z){
    *z=x+y;
}

int main(void){
    int result;
    sum(100,200,&result);
    printf("100 + 200 = %d\n",result);
    return 0;
}
```

sample.s

```
sum:
    addl    %edi, %esi
    movl    %esi, (%rdx)
    ret
main:
    subq    $24, %rsp
    movq    %fs:40, %rax
    movq    %rax, 8(%rsp)
    xorl    %eax, %eax
    leaq    4(%rsp), %rdx
    movl    $200, %esi
    movl    $100, %edi
    call    sum
    movl    4(%rsp), %edx
    movl    $.LC0, %esi
    movl    $1, %edi
    movl    $0, %eax
    call    __printf_chk
    ...
```

64位汇编

编译命令 `gcc -Og -S sample.c`





# 机器码

00000000004005bb <main>:

4005bb:	48 83 ec 18	sub	\$0x18,%rsp
4005bf:	64 48 8b 04 25 28 00	mov	%fs:0x28,%rax
4005c6:	00 00		
4005c8:	48 89 44 24 08	mov	%rax,0x8(%rsp)
4005cd:	31 c0	xor	%eax,%eax
4005cf:	48 8d 54 24 04	lea	0x4(%rsp),%rdx
4005d4:	be c8 00 00 00	mov	\$0xc8,%esi
4005d9:	bf 64 00 00 00	mov	\$0x64,%edi
4005de:	e8 d3 ff ff ff	callq	4005b6 <sum>
4005e3:	8b 54 24 04	mov	0x4(%rsp),%edx
4005e7:	be a4 06 40 00	mov	\$0x4006a4,%esi
4005ec:	bf 01 00 00 00	mov	\$0x1,%edi
4005f1:	b8 00 00 00 00	mov	\$0x0,%eax
4005f6:	e8 a5 fe ff ff	callq	4004a0 <__printf_chk@plt>
...			
400610:	b8 00 00 00 00	mov	\$0x0,%eax
400615:	48 83 c4 18	add	\$0x18,%rsp
400619:	c3	retq	



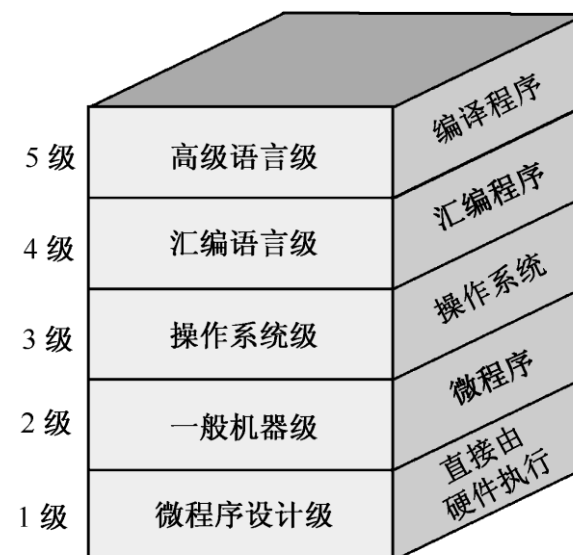
# 1.2 计算机系统的层次结构





# 计算机系统的层次结构

- 第1级为微程序（或硬布线）级。是一个实际机器层，由硬件直接执行微指令。
- 第2级为一般机器级。也称为机器语言级，它由微程序解释机器指令系统。这一级也是硬件级
- 第3级为操作系统级。用于对计算机系统的软硬件资源进行管理和调度。也称为混合机
- 第4级为汇编语言级。通过汇编器将汇编程序翻译成机器指令
- 第5级为高级语言级。用编译程序翻译成第4级的汇编语言程序





# n层计算机系统层次结构

第n层

虚拟机Mn, 其机器语言为Ln

⋮

第4层

虚拟机M4, 其机器语言为L4

第3层

虚拟机M3, 其机器语言为L3

第2层

虚拟机M2, 其机器语言为L2

第1层

实际计算机M1, 其机器语言为L1

用Ln编写的程序需经运行在更低级别的机器

上的解释器解释执行, 或翻译成更低级别机器

上语言后执行  
用L3编写的程序需经运行在M2或M1上的解释器解释执行, 或翻译成L2或L1语言后执行

用L2编写的程序需经运行在M1上的解释器解释执行或翻译成L1语言后执行

用L1编写的程序可直接由电子电路执行





# 软件与硬件的逻辑等价性

- 硬件 – 能实现高速的算术逻辑运算功能，但难以实现较复杂的功能或实现的代价太高
- 软件 – 易于实现各种复杂的算术逻辑运算功能，但是频繁的访存操作制约了处理速度
- 从理论上讲，任何软件算法可以用软件实现，也可以用硬件实现，即：软件和硬件在逻辑上等价
- 对于某一具体功能，采用硬件方案还是软件方案，取决于设计目标、技术水平、器件价格、速度、可靠性、易维护性和变更周期等因素

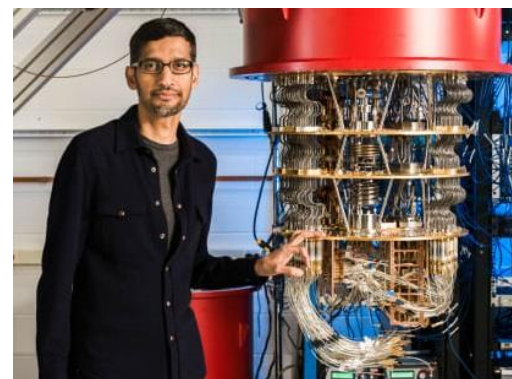
	器件价格	速度	易维护性	变更周期
硬件	高	快	低	长
软件	低	慢	高	短





# 发展趋势

- 以前由软件实现的功能，现在更多地直接由硬件实现
- 计算机系统的软、硬件界限已经变得模糊了
- 众核处理器
- 并行计算、分布式处理技术
- 寻找硅之外的突破
  - ◆ 光计算机
  - ◆ 量子计算机Quantum Computer
    - 量子比特Qubit
    - 谷歌量子计算机Sycamore，54位量子比特
  - ◆ 生物计算机
    - 生化反应机理生物计算机的DNA计算模型
    - 天然高集成度





# 本章小结

- 计算机分类
- 冯诺依曼结构计算机
- 计算机性能指标
- 计算机组成
  - ◆ 软件
  - ◆ 硬件
- 计算机系统层次结构
- 软硬件逻辑等价性

