

第1章 绪 论

计算思维 (Computational Thinking) 是指运用计算机科学的基础概念进行问题求解、系统设计及人类行为理解等涵盖计算机科学之广度的一系列思维活动。计算思维的核心是抽象 (Abstraction) 和自动化 (Automation)。其中,“抽象”是指用符号 (例如数字) 来表示客观事物,“自动化”是指让计算机解决问题而不需要人的干预。

——周以真 (Jeannette M. Wing) 教授, 美国卡内基·梅隆大学

1.1 计算机的发展历程

利用工具来放大大脑力和体力是人类具有智慧的象征,也是人类不懈追求的目标。如果说机械的发明是扩展了人手的功能,交通工具的使用是扩大了人腿的功能,望远镜和显微镜极大地开阔了人们的视野,那么计算工具的发明与利用就是扩展和提高了人脑的功能。

在电子计算机出现以前,人类就曾发明和使用了许多辅助计算工具。例如在远古时期,人们用石头、木棒、刻痕或结绳来延长自己的记忆。在公元前 7 世纪,算筹开始用于辅助计算。公元 8 世纪,我们的祖先又发明了算盘,并提出了基于算盘的珠算。

算盘及珠算的发明是人类计算工具史上的一大飞跃,它的科学性和实用性经受住了长期实践的考验。时至今日,珠算仍在被使用,这是中华民族对人类文明的一大贡献。15 世纪后,算盘及珠算传至日本,并影响欧洲,激励了各国对计算工具的研究。

在电子计算器或电子计算机普及之前,计算尺也是工程技术人员广泛使用的一种计算工具。计算尺是在木质的尺子形状的材料上,印上各种刻度和数字标记,通过拉动中间可移动的部分,找出不同位置上刻度 (数字标记) 的对应关系,来完成一次计算过程。

在 1642 年,年仅 19 岁的法国科学家帕斯卡 (Blaise Pascal, 1623—1662 年) 发明了第一台机械计算器 Pascaline。这台手摇转动的齿轮进位式计算机是帕斯卡为他担任税收官员的父亲设计的,能够完成 6 位数字的加减法运算。Pascaline 的最大贡献就是解决了自动进位这个**关键问题,体现了计算思维的核心概念——自动化。**

德国科学家莱布尼茨 (Baron Gottfried Wilhelm von Leibnitz, 1646—1716 年) 对 Pascaline 进行了改进,于 1673 年研制出具有加、减、乘、除功能的手摇式机械计算器,并提出了“可以用机械代替人进行烦琐重复的计算工作”这一重要思想。遗憾的是,由于当时技术水平的限制,还不能提供大量廉价而精密的机械零件,使得两个世纪后才出现商品化的手摇计算机。

上述计算工具的特点是:机器要由人按照一定的步骤来操作,每一步运算都要由操作者供给操作数,决定进行什么样的操作并安排计算结果。人们常说的“算盘珠子不拨不动”就是这类计算工具的生动写照。计算工程中频繁的人工干预极大地限制了计算速度的提高,并将人束缚于机器之上。人们期待着一种能够实现自动运算的计算工具,将人从计算过程中解放出来。

最早研究自动计算工具的科学家是英国剑桥大学教授巴贝奇 (Charles Babbage, 1792—1871 年)。1822 年,他首先设计制造出“差分机 (Difference Engine)”。**“差分机”是为计算航**

海数据表而设计的，只能运行一个算法，即用多项式计算有限差分。有趣的是它的数据输出方法是用钢锥将结果刻在铜板上，这可以看作一次性写存储介质（如穿孔卡片和光盘）的雏形。

虽然“差分机”运行得很好，但巴贝奇还是对它只能完成一种算法很不满意。1833年，巴贝奇开始了研制“差分机”的更新换代产品，并将其命名为“分析机（Analytical Engine）”。按照巴贝奇的设计，“分析机”由“存储部分”、“计算部分”、“输入与控制部分”和“输出部分”4部分组成。“存储部分”为1000排、每排50个的齿轮阵列。用这些齿轮的不同位置来表示不同十进制数的数据。

当时用于提花机上编织复杂图案的穿孔卡片略加修改后被用来向“分析机”输入计算步骤和数据。**巴贝奇的这个妙想就属于计算思维的核心概念——抽象。**

“分析机”设计蓝图的最大亮点是，将自动计算装置在组成上划分成存储部件、计算部件、输入部件（卡片穿孔设备）、控制部件（穿孔卡片及其阅读设备）、输出部件。

“分析机”的另一个亮点是通用性——从穿孔卡片上读取指令，依据指令进行运算。这样人们就可以通过在穿孔卡片上编制不同的指令来在同一台计算机上完成不同的运算。

尽管由于种种原因，“分析机”没有能够最终研制出来，但它奠定了现代计算机的基本组成结构和主要功能。因此，巴贝奇被誉为“计算机之父”。

1854年，英国数学家布尔（George Boole）出版了《布尔代数》，为计算机采用二进制来进行信息的表示与运算奠定了理论基础。19世纪末20世纪初，一批基于上述研究成果的手摇计算机、电动计算机和卡片式计算机被相继发明出来，如美国国际商业机器 IBM（International Business Machine）公司推出的“插销继电器计算机”。这些辅助计算工具，为人们解决烦琐的数据处理问题提供了很大的帮助。

但是，这些辅助计算工具还是存在诸多不足，难以满足科技发展对计算能力的需求。

第一，它们的计算速度慢，只有大约3000次运算/小时。例如，要准确预报24小时的天天气情况，大约需要200多万次的运算。若使用上述辅助计算工具，要费时5个多月，这就失去预报的意义了。在地球物理勘探中，测量数据的处理往往需要上亿次的计算，上述辅助计算工具更是无法胜任。

第二，运行可靠性差，容易出现计算错误。这是因为在使用这些辅助计算工具时，操作者要参与整个计算过程，人的主观因素直接影响了计算的正确性。

第三，这些辅助计算工具只具有计算功能，而人们在生产实践、科学研究和社会生活中大量需要的是机械控制、信息处理等工作，上述辅助计算工具在这些任务面前就无能为力了。

1937年，致力于研究数学机械化的英国数学家图灵（Alan Mathison Turing，1912—1954年）在《关于可计算的数及其对判定问题的应用》学术论文中提出了一个被后人称为“图灵机（Turing Machine）”的计算模型（参见附录A），这就是现代计算机的理论模型。

鉴于图灵对计算科学的杰出贡献，美国计算机协会（Associative of Computing Machine，ACM）于1966年设立了“图灵奖（A.M. Turing Award）”，以纪念这位杰出的科学巨匠。至今已有50多位科学家荣获了“图灵奖”，这些获奖者的工作代表了计算科学在各个时期最重要的成果，影响着计算科学发展的方向（参见附录B）。第一位华裔“图灵奖”获得者是美国普林斯顿大学的姚期智（Andrew Chi-Chih Yao）博士，他荣获了2000年度的“图灵奖”。

1938年，德国大学生朱斯（Konrad Zuse）成功地制造出第一台二进制计算机Z-1。此后，他又继续研制Z系列计算机，其中Z-3型计算机是世界上第一台通用程序控制机电式计算机，

它的开关元件为继电器，采用浮点计数法和带数字存储地址的指令形式。

1944年，美国哈佛大学的研究生艾肯(H. Aiken)成功地研制了一台机电式计算机 Mark-I。

至此，计算机走过两条技术道路：一条是机械式(Mechanical)，另一条是机电式(Electro-mechanical)。后来，电子(Electrical)计算机都从这两条道路上汲取了很多思想。

1946年2月14日，世界上第一台通用电子计算机 ENIAC (Electronic Numerical Integrator And Computer) 诞生了。这是人类文明史上的一个重要里程碑。从此，电子计算机把人类从繁重的脑力计算和烦琐的数据处理工作中解放出来，使人们能够将更多的时间和精力投入到具有创造性的工作中去。电子计算机的发明是20世纪最杰出的科学成就之一。

从第一台电子计算机的诞生算起，至今仅60多年的时间，但是就在这短短的60多年的时间里，电子计算机得到了迅速的发展和普及。目前，电子计算机已经深入到我们社会生活的每一个角落，它不仅改变了人类工作和学习的方式，而且还改变着人类的观念和思维。

电子计算机的发展经历了四代，目前正在向着第五代计算机发展。

(1) 第一代电子计算机(1946—1958年)是电子管(Vacuum Tube)计算机。它的特征是采用电子管作为逻辑元件，能够处理的数据类型只有定点数，用机器语言或汇编语言来编制程序。第一代电子计算机的应用仅仅局限于科学计算。

第一台电子计算机 ENIAC 是由美国宾夕法尼亚大学莫尔学院的物理教授莫克利(J. W. Mauchley)和工程师埃克特(J. P. Eckert)领导的科研小组研制成功的。它使用了约18800只电子管和1500个继电器，几十万枚电阻和电容，体积为 460m^3 ，自重30t，功耗为140kW，占地面积约为 170m^2 。计算速度约5000次加减运算/秒。然而，ENIAC在1946年2月正式试算时就创造了奇迹：用短于炮弹实际飞行的时间，求出了16英寸海军炮的弹道。

1946年，莫克利和埃克特开始研究“离散变量自动电子计算机 EDVAC (Electronic Discrete Variable Automatic Computer)”，但是最终没有研制出来。与此同时，ENIAC项目的顾问、美籍匈牙利科学家冯·诺依曼(von Neumann)与他的同事在普林斯顿大学高等研究院 IAS (Institute of Advanced Study) 设计他们自己的 EDVAC (IAS 机)。为了实现自动连续地工作，冯·诺依曼在 IAS 机的设计中提出了“存储程序(Stored Program)”的思想。这个设计思想对后来计算机的发展产生了深远的影响。后人称采用该机结构的计算机为“冯·诺依曼计算机”，目前绝大多数电子计算机都是“冯·诺依曼计算机”。因此，冯·诺依曼被誉为“现代计算机之父”。

1946年，莫克利和埃克特离开莫尔学院，创办了电子控制公司 ECC (Electric Control Corp.)，这是世界上第一家计算机公司。1947年，ECC 开始研制 UNIVAC-I (UNIVersal Automatic Computer-I) 计算机。这台计算机首次采用磁带(Magnetic Tape)作为外存储器，采用奇偶校验和双重运算线路来提高系统可靠性。但是，UNIVAC-I 的重要意义更在于它是第一款批量生产的计算机，是计算机产业的起点。UNIVAC-I 当时的售价为25万美元，共生产了48套系统。从此，计算机的研究不再只是学术机构的技术行为，产业和市场需求越来越成为计算机技术发展的主要动力。

1954年，第一个高级程序设计语言 Fortran 问世了。Fortran 语言的问世一举奠定了高级程序设计语言在程序设计中的地位，并成为世界上应用最广泛、最有生命力的高级程序设计语言。Fortran 语言的设计者巴克斯(John Backus)也因此荣获了1977年度的图灵奖。

(2) 第二代电子计算机(1958—1965年)是晶体管(Transistor)计算机。它的特征是采用晶

体管代替电子管作为逻辑元件；用磁芯（Magnetic Core）作为主存储器；采用磁带、磁鼓（Magnetic Drums）、纸带（Paper Tape）、卡片穿孔机和阅读机（Card Punch and Readers）作为输入/输出设备；在软件方面有了很大的发展，相继出现了 Algol、COBOL 等一系列高级程序设计语言。其中 Algol 60 语言的诞生更是标志着对计算机程序设计语言的研究正式成为一门专门的学科，它的设计者诺尔（Peter Naur）荣获了 2005 年度的“图灵奖”。

为了提高计算机系统的利用率、吞吐率，人们开发了管理程序（Monitor）和批处理系统（Batch System），这就是操作系统（Operating System，OS）的雏形。

除了科学计算之外，第二代电子计算机开始应用于数据处理和工业过程控制。

在第二代电子计算机中，代表性的机器有 IBM 公司生产的 36 位计算机 IBM 7094 和美国数字设备公司 DEC（Digital Equipment Company）生产的 18 位计算机 PDP-1。与第一代电子计算机相比，第二代电子计算机具有体积小、重量轻、耗电低、可靠性高等优点，计算速度可达到几万到几十万次运算/秒。

（3）第三代电子计算机（1965—1970 年）是集成电路（Integrated Circuits，IC）计算机。其特征是采用集成电路代替分立的晶体管元件；半导体存储器逐渐取代磁芯存储器；控制单元设计开始采用微程序控制技术。在软件方面，操作系统日益成熟和功能逐渐强化，也是第三代电子计算机一个显著的特点。多道程序、并行处理、多处理机、虚拟存储器（Virtual Memory）、系列计算机（Family of Computers）以及图形用户界面 GUI（Graphical User Interface）等技术的提出，大大推动了计算机科学与技术的发展。

具有代表性的第三代计算机有 IBM 公司的大型计算机（Mainframe）System 360 系列，美国控制数据公司 CDC（Control Data Corporation）的超级计算机（Supercomputer）CDC-6600、CDC-6700。它们的计算速度可达几百万次每秒，甚至几千万次每秒。

20 世纪 60 年代中期，计算机技术出现了一个引人注目的新方向——低成本的小型计算机（Mini-Computer）。当 32 位计算机大行其道时，美国 DEC 公司推出了比 PDP-1 更便宜的 8 位的小型计算机 PDP-8。出人意料地是，这款计算机在市场上深受欢迎。其后，DEC 公司又推出了 16 位的小型计算机 PDP-11。从此，小型计算机以其成本低廉、适用面广、性能价格比高，成为计算机市场的重要角色。

（4）第四代电子计算机（1971 年以后）的特征是采用大规模集成电路和半导体存储器，UNIX 操作系统逐渐成为主流。在这一时代，计算机的性能有了快速提高，由美国人西蒙·克雷（Seymour Cray）创办的克雷（CRAY）公司于 1976 年推出了世界上首台计算速度超过 1 亿次/秒的超级计算机 Cray-1。

第四代计算机的另外一个重要代表是微处理器（Microprocessor）与微型计算机（Microcomputer）/个人计算机 PC（Personal Computer），它们推动了电子计算机的普及。

第三代以后的电子计算机从本质上使用的都是集成电路，只不过是集成度越来越高，所以有人就将集成电路的集成度作为划分第三代以后电子计算机代次的依据。把由集成度为 1~10 个等效逻辑门的小规模集成电路 SSI（Small Scale Integration）和集成度为 10~100 个等效逻辑门的中规模集成电路 MSI（Medium Scale Integration）构成的电子计算机称为第三代电子计算机；把由集成度为 100~10000 个等效逻辑门的大规模集成电路 LSI（Large Scale Integration）和集成度为 10000 个以上等效逻辑门的超大规模集成电路 VLSI（Very Large Scale Integration）构成的电子计算机称为第四代电子计算机。

(5) 那么,什么是第五代电子计算机呢?

沿用按集成度划分的思路,有人提出在超大规模集成电路 VLSI 量纲中,进一步将由集成度为 1 万~100 万个等效逻辑门的超大规模集成电路 VLSI 构成的电子计算机称为第四代电子计算机;将集成度为 100 万~1 亿个等效逻辑门的集成电路定义为巨大规模集成电路 ULSI (Ultra Large Scale Integration),称基于 ULSI 的电子计算机为第五代电子计算机。

但是这种“第五代电子计算机”的定义并没有得到广泛的认同,更多的人认为第五代电子计算机应该是具有广泛知识、能推理、会学习的智能计算机。理想的智能计算机拥有由各种类型专家系统组成的知识库,具有理解、联想、推理、学习、判断和决策的能力。智能计算机能够理解人类的自然语言,能直接接收语言、文字、图形或图像等输入信息,在经过相应处理后,利用知识库中的知识和规则进行推理,从而使问题得到解决。在解决问题的同时,智能计算机的知识库也将自动更新或补充。

人类一直在努力地进行对自身智能的研究。1949 年,美国科学家维纳(N. Wiener)采用开关网络来模拟动物神经系统。20 世纪 80 年代,该思想发展成为所谓的“神经网络计算机”。但是这种计算机与人的大脑毫无相似之处,所以它只是在有限的应用领域(例如非程序化和自适应的数据处理)收到了一些效果。美国科学家西蒙(Herbert A. Simon)在 20 世纪 50 年代提出了“物理符号系统假说 PSSH (Physics Symbol System Hypothesis)”:人的大脑是一个物理符号系统,计算机也是一个物理符号系统,所以可以用计算机来模拟人的大脑。按照 PSSH,人们研制出了棋艺水平接近国际大师的国际象棋机器人,例如 1998 年由 IBM 公司研制的“深蓝(Deep Blue)”就战胜了国际象棋世界冠军卡斯帕罗夫。但是“深蓝”的智能仍然是有限的,它与人类的较量还是负多胜少,它的胜利主要得益于计算机处理能力的提高。

IBM 公司研制“深蓝”的团队是由出生于中国台湾省的许峰雄博士领导的,许博士也因此被誉为“深蓝之父”。

日本曾于 20 世纪 80 年代初提出了一个“第五代计算机系统 FGCS (Fifth Generation Computer System)”计划,并于 1982 年 4 月成立了“新一代计算机技术研究所(Institute for New Generation Computer Technology)”。该研究所当时制订了一个为期 10 年(1982—1991 年)的“智能计算机研究计划”,但是该计划并没有取得预期的成果。

在 1992 年,日本又提出了一个“RWC (Real World Computing)”计划,后来还研制出当时世界上计算速度最快的计算机——地球模拟器(Earth Simulator)。美国自然也不会甘居人后,也提出了对“更新一代计算机”的研究设想。所谓“更新一代计算机”将不再只是采用传统的电子器件,而是更多地采用光电子器件、超导器件、生物电子器件、量子器件。

2011 年 2 月 16 日,IBM 公司推出的人工智能计算机“华生(Watson)”在美国著名的益智节目“Jeopardy (危险边缘)!”中,战胜了两位前冠军参赛者,以绝对优势赢得了历史上第一次人机智力问答比赛的胜利,这标志着智能计算机达到了一个新的水平。

在长达 4 年的研发过程中,共有来自 IBM 美国研究院、中国研究院、日本研究院、以色列研究院的 30 位研究员参与了“华生”的研发。

在第五代计算机的研究中,中国也应该占有一席之地,有志于计算机事业的中国青年,应该积极投身到这一领域中去。

1.2 冯·诺依曼计算机模型

在 ENIAC 设计与研制的过程中，他们的设计者曾向冯·诺依曼进行过咨询。ENIAC 投入使用后，冯·诺依曼曾亲自到现场参观，对 ENIAC 表现出强烈的兴趣。在对 ENIAC 存在的不足（ENIAC 是专用计算机，它的功能是由电路连线来决定的，改变功能需要人为地改变电路连线）进行深入思考的基础上，冯·诺依曼等人于 1946 年 6 月发表了一篇旨在构建一台通用计算机的技术报告《关于电子计算装置逻辑结构初探》。这份报告提出了基于“存储程序”控制的 EDVAC 的设计方案。

“存储程序”的思想是：计算机的用途和硬件完全分离。硬件采用固定性逻辑，提供某些固定不变的功能。通过编制不同的程序（Program）来满足不同用户对计算机的应用需求。

依照这个思想，在计算机上求解一个问题，需要将求解该问题的过程分解成一系列简单、有序的计算步骤，一个步骤由计算机提供的一条计算机指令（Instruction）完成。然后将这些有序的计算步骤一一对应计算机能够识别并可执行的指令——汇总在一起，就形成所谓的程序，并存储在计算机中。计算机通过逐条、顺序执行程序中的指令来完成问题的求解。“存储程序”思想体现了计算思维的核心概念——自动化。

根据《关于电子计算装置逻辑结构初探》设计的“冯·诺依曼计算机”的特点如下。

（1）具备五大功能：数据存储、操作判断与控制、数据处理、数据输入与输出。对应五个功能部件：存储器（Memory）、控制单元 CU（Control Unit）、算术逻辑运算单元 ALU（Arithmetic Logic Unit）、输入单元（Input Unit）和输出单元（Output Unit），如图 1-1 所示。图中，有两股信息在流动。一股是控制流（操作命令），它从 CU 发出，分散流向各部件；另一股是数据流（包括指令和地址），它在 CU 的控制下，从一个部件流到另一个部件。

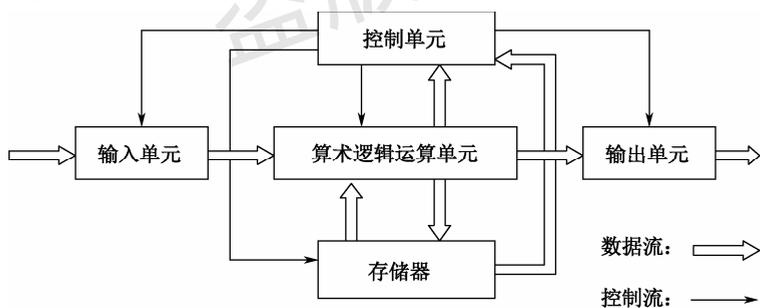


图 1-1 冯·诺依曼计算机结构

（2）存储器由一组一维排列、线性编址的存储单元组成，每个存储单元的位数是相等且固定的，存储单元按地址访问。这是最简单、最易于实现的信息存储与查找方案，就如同一排平房，房间大小都是一样的，按照房间号查找房间，房间号逐间递增。

（3）“程序”由一条一条的指令有序排列而成，而指令由操作码和地址码两部分组成。操作码规定了该指令的操作类型（功能），地址码指示存储操作数和运算结果的存储单元地址。操作数的数据类型由操作码来规定，操作数可能是定点数、浮点数、双精度浮点数、十进制数、逻辑数、字符或字符串等。

（4）指令和数据均采用二进制表示，并以二进制形式进行运算。二进制的计算规则是最简

单的，加法仅有 4 种： $0+0=0$ ， $0+1=1$ ， $1+0=1$ ， $1+1=10$ 。

把纷繁复杂的信息抽象成 0/1，这是计算机的根，是计算机的哲学，是计算思维的核心概念“抽象”的具体体现——信息符号化，符号数字化。

(5) 为了简化计算机的控制与组成，程序（指令）与数据是同等地、不加区分地存储在同一个存储器中。但可以从时间和空间上，将它们区分开。在取指周期，从存储器流向控制器的是指令；在执行周期，存储器与运算器交换的是数据。

(6) 为了实现“逐条、顺序执行程序中指令”，冯·诺依曼提出了一个极易实现的解决方案：设置一个“程序计数器 PC (Program Counter)”来指示下一条将要执行的指令的地址。在一般情况下，每执行完一条指令，PC 就自动加 1，指向下一条指令的存储单元。

当然，为了赋予计算机更多的“灵性”，PC 的值也可以通过执行特殊的指令来修改，从而达到改变指令执行顺序的目的。这样，执行指令的顺序就不受限于指令存储的顺序了。

不过，尽管冯·诺依曼非常聪明，冯·诺依曼计算机的结构还是存在一些问题，后来人们对它进行了改进和发展。

例如，由于以 ALU（算术逻辑运算单元）为中心，输入/输出单元与存储器之间的数据传递都要经由 ALU，这使得 ALU 无法专注于运算，低速的输入/输出和高速的运算不得不相互等待，串行工作。因此，很快“冯·诺依曼计算机”就被改进成以存储器为中心（如图 1-2 所示）。这样，输入/输出设备就可以与运算器并行工作，输入设备也可与输出设备并行工作，提高了设备的效率和利用率。同时这使得计算机五个功能单元的互连更加简单。

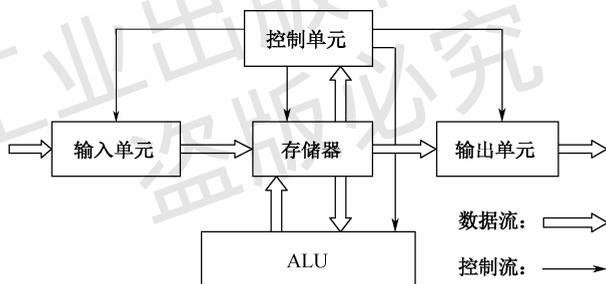


图 1-2 现代冯·诺依曼计算机结构图

时至今日，冯·诺依曼计算机结构经历近 70 年，依然占据统治地位。这是为什么呢？

说到底，冯·诺依曼的最大贡献就是确定了计算机五个部件的互连结构。在考虑到 ALU 既需要从存储器中读取数据又需要将运算结果写回到存储器，请读者自行设计一个计算机五个部件的互连结构。看看能否比图 1-2 更简单。

抽象地看，“冯·诺依曼计算机结构”就是“以一个部件为中心，实现五个部件互连的星形结构”。这样实现的“五个部件互连”是最简单的、连接链路是最少的。由于外围的四个部件只能与中心部件进行数据通信，而且有些数据通路还是单向的，所以冯·诺依曼计算机所需要的数据传送功能也是最少的、最简单的。

综上所述，“冯·诺依曼计算机”可以概括为：用最简单、最易于实现的思想（二进制及其运算规则）来表示数据并实现运算，用最简单的互连结构来组成一台计算机。

这不仅体现了计算思维，还很好地体现了“简单就是美”的工程哲学理念！

1.3 计算机的组成结构

1.3.1 计算机的基本组成

1. 概述

从组成的角度看，计算机由 CU、ALU、存储器、输入单元和输出单元组成。在具体实现时，通常将 CU 和 ALU 集成在一起，构成“处理单元 PU (Processing Unit)”。处理单元也被称为“处理器 (Processor)”。

一台计算机通常只拥有一个 PU，而这个 PU 又是计算机的核心部件，所以这样的处理单元又称为“中央处理单元或中央处理器 CPU (Central Processing Unit)”。事实上，一台计算机可以拥有多个 PU，这样的计算机称为“并行计算机 (Parallel Computer)”或“高性能计算机 (High Performance Computer)”。**追求更高的计算性能，是科学研究的永恒主题。**

2. 硬件 (Hardware)

从功能的实现载体看，计算机由硬件和软件 (Software) 两部分组成。

硬件指的是构成计算机的物理实体，例如处理器、存储器、输入/输出控制器等芯片及其集成这些芯片的印制电路板——主板 (Mainboard)。其中，主板上的存储器是可以被处理器直接访问的，被称为“主存储器 (Main Memory, 简称主存)”。其他硬件还包括软盘驱动器、硬盘驱动器、光盘驱动器及连接它们的线缆。上述硬件一般都安装在一个机箱内部的机架上，这个机箱及其中的硬件被统称为主机。主机之外的硬件设备主要是各种输入/输出设备 (Input/Output Devices)。

输入设备用以将程序与数据、各种人机交互命令转换成电信号，并在控制器的指挥下，按一定的地址存储在各种存储介质中。传统的输入设备有光电机、键盘、鼠标、触摸屏等。近年来，随着声图文智能人机接口技术的发展，麦克、摄像头、手写板逐渐成为受人欢迎的输入设备。

输出设备在控制器的指挥下将计算结果或计算机中存储的信息，转换成人们能够识别的形式呈现给用户。常见的输出设备有显示器、打印机、绘图仪、音箱等。

计算机系统中还有一个重要的硬件就是辅助存储器 (Secondary Memory, 简称辅存)。

辅存是计算机系统不可或缺的存储介质。在计算机系统关机 (或断电) 后，存储在主存储器中的信息将消失。这样的存储器被称为“易失性存储器 (Volatile Memory)”。所以需要长久保存的程序和数据必须存储在“非易失的存储器 (Non-volatile Memory)”——辅存里。

常见的辅存有：硬盘 (Hard Disk, HD)、软盘 (Floppy Disk, FD)、光盘 (Optical Disk Memory, ODM)、磁带等。辅存的容量比主存的容量要大得多，每位的平均价格也要低得多，但是它的访问速度却明显慢于主存。所以，为了保证速度很快的处理器有较高的工作效率，计算机的设计者规定处理器只与速度较快的主存交换信息，而不直接访问辅存，辅存中的信息要装入到主存后才能供处理器使用。

由于辅存位于机箱外部，故又称为“外存 (External Memory)”。相应地，位于机箱内部的主存也称为“内存 (Internal Memory)”。

在有些场合，除主板及其上芯片之外的硬件器件称为外部设备 (简称外设)。外设包括：输入设备、输出设备、辅存。

因为计算机中所有的操作和命令都有一定的时间顺序，所以主板上需要设置一个定时部件。定时部件有时钟 CLK (Clock) 和时序信号发生器 TSG (Timing Signal Generator)。

3. PU 的组成

PU 是计算机系统的核心，由 ALU、CU 以及一些暂存单元——寄存器 (Register) 组成。

ALU 能完成“加”、“减”、“乘”、“除”等算术运算和“与”、“或”、“非”、“异或”等逻辑运算。但它并不是只靠一套电路来“智慧地”完成各种运算，而是用不同的电路来完成不同的运算，即由加法器来完成算术运算，由逻辑运算器来完成逻辑运算，由移位器 (Shifter) 来完成移位运算，由求补器 (Complementer) 来完成“取反/求补码”运算。

CU 是 PU 的指挥机构，由程序计数器 PC、存放当前指令的指令寄存器 IR (Instruction Register)、解释指令的指令译码器 ID (Instruction Decoder)、发出各种命令信号的控制信号发生器 CSG (Control Signal Generator) 及相应的控制逻辑组成。CU 依据指令译码器产生的一系列操作命令/信号来指挥、协调 PU 乃至计算机系统中各个部件的工作。

PU 内部的寄存器有若干个存放数据的数据寄存器、若干个存放操作数地址的地址寄存器和一个存放各种“标志 (Flag)”的标志寄存器 FR (Flag Register)。

有些文献将这些寄存器称为“寄存器文件 (Register File)”或“寄存器堆”。在有的计算机中，寄存器既可以存数据，也可以存地址。这样的寄存器称为通用寄存器 GPR (General Purpose Register)。

最常用的数据寄存器是用于存放加法的一个操作数及运算结果的累加寄存器 ACC (Accumulator)。同样是一串“0101”，如果存储在数据寄存器中，则计算机将其理解为一个数；如果存储在地址寄存器中，则计算机将其理解为一个操作数的地址。

引入“标志”来表示/区分计算过程的各种状态是计算思维的一个具体体现。例如，当加法运算的结果为零时，“零标志 ZF (Zero Flag)”被置为 1，否则为 0。运算结果的符号被复制到符号标志 SF (Sign Flag) 中。SF 为 1，表示运算结果为负数，否则为正数。当运算结果中“1”的个数为零或偶数时，“奇偶标志 PF (Parity Flag)”被置为 1，否则为 0。类似的还有“进位标志 CF (Carry Flag)”和“溢出标志 OF (Overflow Flag)”等。这些标志分别占据标志寄存器中的不同位置，后继的指令就可以根据 ZF、SF、PF 的值来选择不同的操作。

访问主存是 PU 经常执行的操作。为了实现这个操作，PU 内部设置了专门存放访存地址的寄存器 MAR (Memory Access Register)、专门存放与主存交换数据的寄存器 MDR (Memory Data Register)。有的文献称 MDR 为 MBR (Memory Buffer Register)。

在用户看来，只要把主存地址送入 MAR，启动读命令，在一个访存周期内，目标数据就会从主存被读入到 MDR 中；或者只要把主存地址送入 MAR 并把目标数据送入 MDR，启动写命令，在一个访存周期内，目标数据就会从 MDR 被写到主存中。

图 1-3 显示了计算机系统硬件的基本组成。

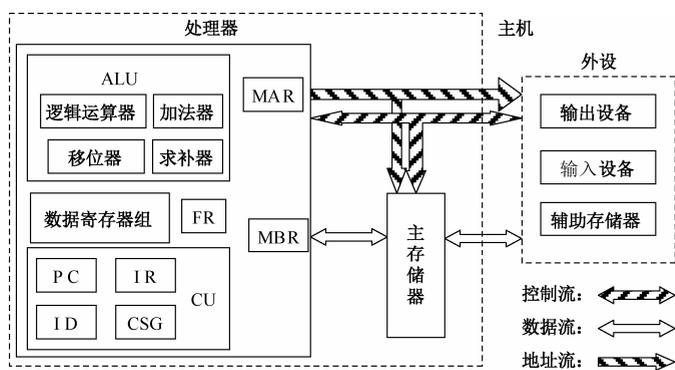


图 1-3 计算机系统硬件的基本组成

4. 软件（Software）与软件/硬件的等价性原理

广义上，软件是“计算机程序、过程、规则及与这些程序、过程、规则有关的文档，以及从属于计算机系统运行的数据。”狭义上，软件指的是发挥电子计算机功能的各种程序及相应数据。

通常，软件分为系统软件和应用软件。也可以更详细地分为：系统软件、应用软件、支持软件、测试与维护软件。

系统软件是指构成一个计算机系统所必需的基本软件，与具体用户无关。常见的系统软件有：操作系统和数据库管理系统。在互联网时代，某些网络基础软件（如网络浏览器、电子邮件等）也可以算是系统软件。

应用软件是由用户根据各自的应用需要而安装的、解决专用领域特殊问题的软件，如 AutoCAD、3ds Max、Protel、Mathlab 等。

支持软件指用于帮助和支持软件开发的软件，如汇编程序、高级语言编译器、文本编辑器、设计工具软件、项目管理软件、配置管理软件等。

测试与维护软件指用于软件故障诊断、错误隔离、系统调试及检测系统可靠性的软件。

除去软件和硬件外，还有一个概念就是“固件（Firmware）”。对于那些不再需要改动而且经常被调用的软件，为了使其有更快的执行速度，可以将其存储在访问速度较快的、具有非易失性的只读存储器 ROM（Read Only Memory）芯片中。这相当于将软件“固化”在硬件（ROM 芯片）中。如果需要修改或升级软件，只需要更换一块 ROM 芯片即可。

这种吸收软件、硬件各自优点，性能介于软件和硬件之间（执行速度快于软件，灵活性优于硬件），以硬件形式出现的软件，被称为“固件”。

在器件成本不断下降的今天，固件越来越多地被采用以提高计算机的性能。

事实上，计算机是面向算法的机器，而程序/软件是某个算法的实现。一种算法可以由硬件或固件来实现，也可以由软件来实现。例如数组运算，在微机上由软件来实现，而在大型计算机或超级计算机上则是由硬件来实现，即硬件和软件在逻辑功能上是完全等价的。软件的功能在原理上可以由硬件来实现，硬件的功能在原理上也可以由软件来实现，这称为“软件/硬件的等价性原理（The Principle of Equivalence of Hardware and Software）”。

1.3.2 计算机体系结构、计算机组成与计算机实现

1. 计算机体系结构与系列机

“计算机体系结构（Computer Architecture）”是计算机学科中最重要的概念之一。但这一概念并不是同计算机或计算机科学与技术学科一起诞生的，它是在第一台电子计算机诞生近 20 年后，由安达尔（G. M. Amdahl）等人于 1964 年在设计 IBM System 360 时提出的。

在 20 世纪 50 年代，计算机系统的设计大多是逐个进行的。由于设计者的不同，计算机一般具有自己的、与众不同的指令系统，程序员按照其特定的汇编符号编写汇编程序并在特定的机器上运行。当时每个用户拥有的计算机数目极少，计算机的昂贵也使得用户很少更换计算机。用户只关心自己计算机的功能与性能，而不提其他需求。

从 20 世纪 50 年代末期开始，随着计算机的日益普及，需要编写的程序越来越多，程序也越来越大，程序的开发成本越来越高、开发周期越来越长。这时，人们就希望别的部门开发的程序能够拿到自己部门的机器上运行，为原来机器开发的程序能够不加修改地在新买的机器上

运行。因此，为了保证优秀的软件能够长期使用，保护用户的投资，降低软件开发人员的重复工作量，使用户或软件开发人员能够把更多的资金或精力投入到新的软件上去，程序可移植性 (Portability) 的概念被提了出来。所谓可移植性是指在一台计算机上能够运行的程序，不加修改或只需少许加工就可以在另外一台计算机上正确运行，并给出相同的结果。不具备可移植性的程序，就不能在其他计算机上运行，不能被其他用户共享。

这时，计算机的发展进入到第二代阶段，计算机的应用领域也从传统的科学计算领域，逐渐扩大到数据处理、过程控制和事务管理等领域，计算机生产厂商不断增多，计算机的型号五花八门，既有面向科学计算的超级计算机，又有面向大规模数据处理和事务管理的大型计算机，还有低端用户的中、小型机。

因此，产品多样性和软件难于移植的矛盾，就摆在计算机设计师的面前。那时，IBM 公司投资 50 亿美元计划开发一个大型计算机系统，总设计师安达尔雄心勃勃，决心一举解决这个矛盾。

首先，安达尔计划让新的计算机系统功能齐全、无所不能。用户只要购买了一台新机器，就可以满足他对计算机所有的需求。为此，这个新系统被命名为 System 360 (S/360)，寓意它能够满足全方位 (360°) 的应用。

其次，安达尔希望 S/360 长盛不衰，10 年、20 年，甚至上百年后仍然受到用户的喜爱。用户一旦购买了 S/360，就没必要购买别的厂商的机器。由于用户始终使用 S/360，也就不用于程序难于移植到其他机器上而发愁了。

只要肯投入，付得起高价钱，实现第一个计划并不是一件难事。但由于新技术的不断提出、新材料的不断涌现、计算机厂商之间的竞争日益激烈，计算机系统更新换代的速度越来越快，计算机系统的生命周期越来越短。在这种情况下，如何实现安达尔的第二个计划呢？

通过对用户和市场进行深入的分析，安达尔定下了解决问题的哲学：一切为用户着想，而计算机系统最直接的用户就是程序员。

对于程序员而言，他们最大的期望就是一劳永逸，即一旦掌握某一机器的属性及其编程方法，就永远只对这一机器编程，不愿意再学习其他机器的属性及其编程方法。但同时他们又想拥有性能不断提升的新机器来充分施展其软件的功能。

这样，不变的机器属性及其编程方法与对不断提升的机器性能的期望构成了一对矛盾。

为此，安达尔提出了一个新的概念——“计算机体系结构”来解决这一矛盾，并将其定义为程序员所看到的机器的属性，即机器的概念性结构和功能特性。在当时，程序员主要指汇编程序员。

安达尔认为：计算机的设计/制造者可以利用新技术、新材料来设计制造性能更高的新机器，但只要新机器保持原机器的体系结构，程序员就无须学习新的机器属性及机器语言，对原有的软件自然不用做任何修改就能在新机器上运行。

安达尔还考虑到：用户对性能的需求是多方面的，但是对软件可移植性的要求是一致的，所以可以通过生产具有相同体系结构但不同档次的机器来同时满足用户这两方面的要求。

可见，安达尔真是全心全意为用户着想。

安达尔将同一厂商生产的具有相同体系结构的机器定义为系列计算机 (简称系列机)。系列机可以是不同年代生产的机器，也可以是相同年代生产但不同档次的机器。安达尔设计的 IBM S/360 就是计算机历史上的第一个系列机。

若将软件兼容 (Software Compatibility) 定义为同一个程序可以不加修改地在具有相同体

系结构的各档机器上正确运行，唯一区别仅仅在于运行时间长短不同，则安达尔就是以软件兼容（统一机器语言）的方式来解决程序的可移植性问题的。

通过提出“计算机体系结构”概念，安达尔解决了 IBM S/360 的软件可移植性问题。

计算机体系结构和系列机这两个概念的提出，一举打破了 20 世纪 50 年代计算机“手工作坊式”的设计生产方式。使得已有的软件资源得到了充分的利用，有效地减少了开发软件的投资与工作量。很好地解决了软件开发环境要求相对稳定，而新技术、新材料又不断涌现之间的矛盾。使得程序员能够把更多的时间和精力投入到完善软件功能、提高软件质量的工作中，并为积累和重用软件，提高软件生产率奠定了基础。同时，计算机厂商又可以不断地采用新技术、新材料，为市场提供性能更高、价格更低的新机器，来更好地满足用户不断增长的应用需求。计算机产业从此走向了供需两旺、蓬勃发展的新时期。

其后，IBM 公司在 20 世纪 70 年代后期和 90 年代又先后推出了与 System/360 兼容的 System/370 系列和 System/390 系列。

系列机的概念一经提出，各个著名计算机厂商纷纷效仿，先后分别提出了各自不同的系列机产品。

DEC 公司在 PDP-11 系列之后，又推出了 VAX-11 (Virtual Address eXtension of PDP11) 系列计算机。CDC 公司的 6600、7600、CYBER 系列超级计算机也曾一度非常显赫。CRAY 计算机公司一直研制并生产着 CRAY 系列超级计算机。英特尔公司从 1980 年开始推出的 Intel 80x86 系列微处理器更是创下了产量最大的系列机的纪录。

所以，计算机体系结构和系列机概念的提出是计算机发展史上的一个重要的里程碑。“计算机体系结构”这一概念至今仍在计算机科学与技术领域中处于重要的地位，按“系列机”的思想来设计计算机仍是所有计算机厂商所必须遵循的原则。

同时，这一概念也影响了软件开发，导致软件工程领域出现了一个新的概念——软件体系结构。欲了解这一概念及相关技术，请学习“软件体系结构”课程或相关书籍。

2. 计算机组成 (Computer Organization) 与计算机实现 (Computer Implementation)

伴随计算机体系结构概念的提出，两个与之相关联的概念也相继提出。它们是：计算机组成和计算机实现。

计算机组成是计算机体系结构的逻辑实现。一种计算机体系结构可以有多种不同的计算机组成，例如实现加法，不同的计算机可以选择不同的逻辑电路。关于计算机组成的具体内容将在后继课程“计算机组成原理”中介绍。计算机组成设计的任务是按照所要求的性能价格比，最佳、最合理地用各种器件和部件组成计算机，以实现设计规定的体系结构。

计算机实现是计算机组成的物理实现。一种计算机组成也可以有多种不同的计算机实现。计算机实现的设计内容包括：中央处理器、主存储器等部件的物理结构、封装技术和制造工艺；信号传输的频率、电气属性；电源、冷却、微组装技术和整机装配技术等。

计算机实现是计算机体系结构和组成的基础。先进的计算机实现技术，尤其是器件技术的发展，一直是推动计算机体系结构和组成发展的最活跃的因素。著名的摩尔定律 (Moore's Law) 揭示的就是这一现象。

3. 计算机体系结构、计算机组成与计算机实现三者的关系

根据“软件/硬件的等价性原理”，计算机的功能可以由硬件来实现，也可以由软件来实现。对于一台具体的计算机系统，它所具有的功能到底由硬件承担多少，由软件承担多少，这是计算机体系结构设计所要解决的首要问题。

软硬件功能的分界面就是计算机的全部指令的集合——指令集 (Instruction Set)。加大软件所承担的比例,可以减少硬件成本,但是解题时间会延长。随着大规模集成电路的迅速发展,硬件成本不断下降,加大硬件所承担的比例,或将软件固化,成为当今计算机技术发展的重要趋势。

在计算机体系结构设计工作——确定指令集——完成之后,就可以开展计算机组成设计工作,即确定指令的逻辑实现方案,如取指令、译码、取操作数、运算、写回结果等具体操作的定义、划分及其排序方式等。此后,就应该考虑这些操作的具体实现电路、器件的设计与装配等。这些就属于计算机实现的设计工作。

例如,主存的最小编址单位、CPU 所能够访问的最大主存容量以及所采用的寻址方式的选择与确定属于计算机体系结构,而主存所采用的逻辑结构及管理策略属于计算机组成,主存的器件设计、电气性能则属于计算机实现。

对于系列机,它们都具有相同的体系结构。但是,由于系列机中不同型号的机器所面对的目标用户不同,这些目标用户对机器的性能、价格的要求不同,所以需要采用不同的计算机组成或实现技术来实现相同的体系结构。另外,新推出的系列机也会采用新的组成或实现技术来提高其性能或降低其成本。

例如,低档的计算机系统采用的是单总线结构,而较高档的计算机系统则采用的双总线结构甚至三总线结构来提高系统的通信能力或系统的可靠性。这些都属于相同的计算机体系结构,为了不同的设计目标,采用了不同的计算机组成。

同一型号的处理器的可以采用不同的主频来实现,这就属于相同的计算机组成采用了不同的计算机实现,以满足不同用户在价格或性能上的不同要求。

1.3.3 计算机系统的层次结构

计算机系统是一个由硬件和软件组成的复杂系统。为了实现高性能、低成本、适于大批量生产、使用方便、易于维护的设计目标,计算机系统的分析、设计和制造就逐渐地从一门艺术转变成一项工程,即将工程的思想、观点、方法、技术、工具等应用于计算机系统的分析、设计和制造。实践已经证明,只有把一门艺术转变成一项工程,才能做到其产品生产的“多快好省”。软件开发技术,从编制程序发展到软件工程,也说明了相同的道理。

工程的一个最重要的思想就是引入“分工与协作”。因此,计算机系统的设计和制造就需要划分成不同的模块。在定义好各模块之间的接口后,不同的设计人员就可以同时开始计算机系统不同模块的设计工作。这样就可以调动尽可能多的人力资源来加快设计的进程。同时,这种做法还使得计算机系统中某一模块的实现技术出现新的改进时,设计人员在遵循原定的接口的前提下,可以对这一模块进行重新设计与实现,并将新的模块直接替换旧的模块,而不会影响到系统中其他模块的正常工作。

目前最主流的计算机系统的划分方法就是从计算机程序设计语言的角度,按计算机系统所提供的功能将其划分成一个“层次结构 (Hierarchy)”,如图 1-4 所示。

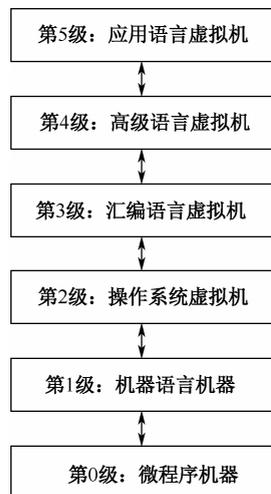


图 1-4 计算机系统的层次结构

第1级机器是机器语言机器，即由CPU、主存、I/O等组成的物理上的计算机。它只能运行以二进制形式表示的计算机程序。这样的计算机使用起来是很不方便的。

为此，人们在它上面安装了操作系统，从而得到第2级机器——操作系统虚拟机。操作系统是最重要的系统软件之一，它负责管理和调度整个计算机的资源。通过学习后继课程“操作系统”，读者将了解、掌握操作系统，甚至可以自行研发新的操作系统。

之所以称为虚拟机（Virtual Machine），就是因为用户在使用计算机时，看到的就是软件的界面，他/她并不了解也不必了解计算机内部的结构及工作原理。软件的界面就是计算机所具有功能的具体体现。虚拟机的实现有解释（Interpretation）和翻译（Translation）两种途径。在计算机专业内，“翻译”常称为“编译（Compile）”。

所谓解释是指在执行某一层机器的源程序时，其中的指令/语句是逐条地、实时替换成用下一层机器语言编写的等效程序段，然后立即在下一层机器上运行。虚拟机（解释程序）是边解释边执行，不保留目标代码。下次执行时，需要重新解释。

所谓编译是指某一层机器的源程序，在运行之前，先一次性完整地转换成用下一层机器语言编写的程序，然后在下一层机器上运行。虚拟机（编译程序 Compiler，也叫编译器）是先编译后执行。下次执行时，不需要重新编译。

根据需要了解计算机组成与结构程度的不同，可以把计算机用户分为：系统程序员（System Programmer）、应用程序员（Application Programmer）、最终用户（End User）和系统管理员（System Administrator）。

最终用户不需了解任何的计算机组成与体系结构，他只需了解应用软件（如网络浏览器、游戏、财务软件）的功能和使用即可，大多数计算机用户属于最终用户。

最终用户面对的是应用语言虚拟机，即最高级虚拟机——第5级机器，使用应用语言——如操作数据库管理软件的SQL语言、操作AutoCAD软件的绘图命令——来操作的软件。在用户看来，使用计算机就是使用应用语言。今天，应用语言已经变成软件界面上的操作“选单（Menu）”或“按钮（Button）”。

应用软件是应用程序员使用高级语言（High Level Language, HLL）（如Fortran、C/C++、Java等）编写的。应用程序员只需有一些基本的计算机组成与体系结构的概念即可，而这些概念只是作为他学习、掌握高级语言的基础。应用程序员面对的是高级语言虚拟机（第4级机器）。

不过，诸如C/C++、Java这样的高级程序设计语言，大量使用操作系统提供的应用程序接口API（Application Program Interface），这也是应用程序员需要了解的。

实际上，高级语言开发的应用软件并不能直接在计算机上运行，高级语言源程序必须由翻译程序（编译器 Compiler）或解释程序（也叫解释器 Interpreter）转换成目标计算机的汇编语言（Assembly Language）源程序，汇编语言源程序还要经过“汇编程序（Assembler）”汇编成机器指令才能在计算机上运行。

编写编译器、解释器以及汇编程序的程序员就属于系统程序员。此外，现代计算机都安装有操作系统和数据库，开发操作系统或数据库的程序员也属于系统程序员。系统程序员必须深入理解计算机组成与体系结构，才能开发出系统软件。而系统软件的质量依赖于开发它的系统程序员对计算机组成与体系结构的理解程度。

编译器/解释器就是高级语言虚拟机，汇编程序就是汇编语言虚拟机。系统程序员面对的是汇编语言虚拟机、操作系统虚拟机或者机器语言机器。

对于多任务、多用户的计算机系统，需要引入系统管理员来管理用户账户、分配计算机资源，安装、配置和维护系统的软件和硬件。他主要面向操作系统虚拟机，需要深入了解操作系统的功能与特性，需要理解计算机组成与体系结构的基本概念和基本原理。

由于编译器、解释器和汇编程序都运行在物理计算机（第1级机器）上、服从操作系统的调度，所以汇编语言虚拟机就是第3级机器，操作系统虚拟机是第2级机器。

如果处理器的控制器是以微程序方式实现的（控制器有微程序、硬连线两种实现方式，将在第4章中介绍），则第1级机器以下还有一层：第0级微程序机器。如果存在这一级机器，计算机的设计者就可以通过改变指令的微程序，轻松地改变计算机的指令集。

在上述计算机的层次结构中，上一层机器向下一层机器发出操作命令，下一层机器执行操作命令，并向上一层机器返回操作结果。至于下一层机器是如何完成这些操作，它的内部结构如何，上一层机器并不关心。

因此，面向不同层次机器的程序员所看到的计算机的属性是不同的。例如，面向机器语言计算机的程序员看到的计算机属性就是机器语言计算机的属性；面向高级语言虚拟机的程序员看到的计算机属性就是该高级语言的语法、语义和语用，即高级语言所具有的功能特性。高级语言虚拟机的属性与机器语言计算机的属性是无关系的，即同一个高级语言虚拟机的属性可以建立在不同的机器语言计算机之上，所以面向高级语言虚拟机的程序员可以对机器语言计算机的属性“视而不见”，或者他根本就不想关心。

这又引出了计算机与软件工程领域另外一个重要的概念——透明性（Transparency）。

计算机系统中客观上存在的事物或属性，从某个角度去看好像不存在，这种现象被称为“透明性”。

在计算机系统的层次结构中，底层机器的属性，对上一层机器的程序员是透明的。

例如，计算一个用高级语言表示的算术表达式，在CPU内部是通过执行多条指令来完成的，这些指令需要访问寄存器或内存单元，多次使用运算器。这些客观存在的事物和操作对于高级语言程序员来说是看不见的，他/她也根本不想去看，不关心。所以低层机器的属性，对于高级机器的程序员就是透明的。

类似的例子还有，计算机用户很轻松地使用操作系统，而不需关心和了解操作系统的设计与实现。可见，“透明性”的提出促进了社会分工，提高了劳动生产率。

计算机体系结构设计，可理解为：决定哪些事物对汇编程序员透明，哪些事物对汇编程序员不透明。

【例 1-1】（2010 年硕士研究生入学统一考试计算机专业基础综合考试试题）

下列存储器中，汇编语言程序员可见的是_____。

- A. 存储器地址寄存器（MAR） B. 程序计数器（PC）
C. 存储器数据寄存器（MDR） D. 指令寄存器（IR）

答：根据计算机体系结构与透明性的概念，PC对汇编语言程序员是可见的，因为程序员需要依据PC的值计算相对转移的偏移量。而MAR、MDR和IR属于计算机组成的范畴，对汇编语言程序员是透明的，即不可见的。故选择B。

1.4 计算机系统的分类

1.4.1 综述

根据数据表示原理的不同，电子计算机分为模拟式和数字式。

模拟式电子计算机所处理的电信号在时间上是连续的，称为模拟电信号。用电信号的电位大小来模拟数值，不同的电位值对应不同的数值。模拟计算机的处理过程均由模拟电路来实现，处理速度快，但是电路结构复杂，处理的精度低，抗干扰能力差，目前已很少使用。

数字式电子计算机所处理的电信号在时间上是离散的，称为数字量。如在电子线路中，用电平的“高/低”或脉冲的“有/无”来表示数值“1/0”。这样就可以用一组触发器的输出电平或一串脉冲来表示一个二进制的数值。不同的电平组合就能够表示不同的数值，增加组合位数就能增大数的表示范围和精度。

数字式电子计算机将信息数字化，具有许多独特的好处。例如，数字化信息便于利用各种存储器和寄存器保存，使计算机可以具有极大的存储量，从而可以对海量数据信息进行处理。数字信息还可以用来表示各种物理量和逻辑变量，以及文字、符号、图像等，因而计算机除了可以进行数值计算外，还可以用于逻辑判断、文字编辑、图像处理等。由于目前广泛使用的计算机都是数字式电子计算机，所以为了简便起见，人们就将数字式电子计算机直接称为电子计算机，简称计算机。

将数字技术和模拟技术相结合而实现的计算机，称为模拟、数字混合式计算机。

依据性能的高低，可以将计算机分为超级计算机、大型计算机、中型计算机、小型计算机、工作站（Workstation）和微型计算机。超级计算机是指其所处年代性能最强的计算机。

按照结构集成的不同方式，微型计算机分为单片机（Single Chip Microcomputer）和单板机（Single Board Microcomputer）。

单片机是指组成计算机系统的处理器、存储器、输入/输出接口、定时器和计数器都集成在一个芯片上。单片机主要应用于工业控制领域。

单板机是指组成计算机系统的处理器、存储器、输入/输出模块是以分立芯片的形式集成在一个印制电路板（主板）上。常见的各类台式机、服务器、笔记本计算机都属于单板机。

计算机按其设计目的又可分为专用计算机和通用计算机。

专用计算机能以最佳的性能、最高的设备利用率来满足特定用户的使用需求，但是它的适应性差。一旦用户的使用需求发生变化，专用计算机的性能和效率都会明显下降。

相比之下，通用计算机的适应性就要好得多，它的应用领域较广，但是对于具体的应用领域，其性能和效率都很难达到最佳。

根据计算机的用途不同，计算机分为个人计算机、工业控制计算机（简称工控机）、军用计算机（也叫加固型计算机或抗恶劣环境计算机）和嵌入式计算机（Embedded Computer）。目前，嵌入式计算机的发展十分迅速。手机、汽车、家电、机器人里都有嵌入式计算机。

在一些安全攸关（Safety-critical）领域，如航空/航天或铁路调度，还需要使用容错计算机（Fault-tolerant Computer）或可信计算机（Trusted Computer）。

1.4.2 弗林分类法

计算机的工作过程是执行一串指令，来对一组数据进行处理。据此，弗林（M. J. Flynn）

于 1966 年提出按照指令流和数据流的多倍性对计算机进行分类。所谓指令流 IS (Instruction Stream) 是指机器执行的指令序列, 指令被译码后形成控制流 CS (Control Stream), 数据流 DS (Data Stream) 是指由控制流调用的数据序列 (包括输入数据和中间结果), 多倍性 (Multiplicity) 是指在系统最受限制的元件上, 同时处于同一执行阶段的指令流或数据流的最大可能个数。

按照指令流和数据流分别具有的多倍性, 弗林将计算机分为以下四类:

- (1) 单指令流单数据流 SISD (Single Instruction Stream, Single Data Stream)
- (2) 单指令流多数据流 SIMD (Single Instruction Stream, Multiple Data Stream)
- (3) 多指令流单数据流 MISD (Multiple Instruction Stream, Single Data Stream)
- (4) 多指令流多数据流 MIMD (Multiple Instruction Stream, Multiple Data Stream)

图 1-5 画出了这四类计算机的基本结构 (没有包括 I/O 设备)

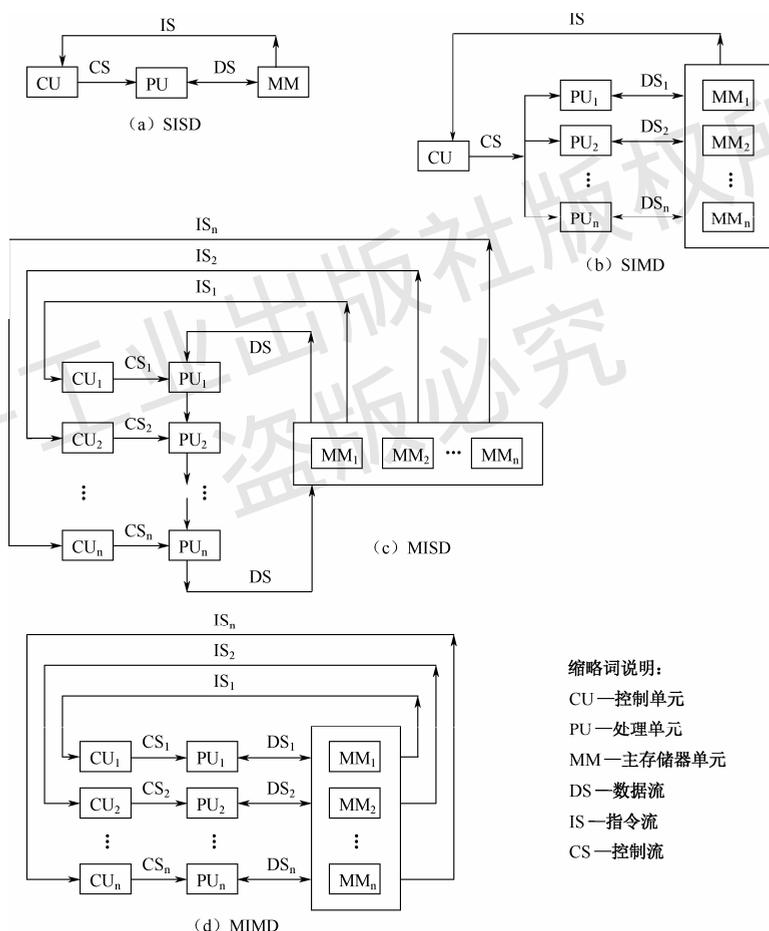


图 1-5 弗林分类法的各类计算机的基本结构

SISD 计算机就是传统的计算机, 它只设有一个控制单元, 该控制单元一次只对一条指令进行译码并且只对一个操作单元分配数据。尽管这类计算机也可以设置有多个并行工作的存储器或者操作单元。但只要是一次只对一个操作单元分配数据, 它就属于 SISD 计算机。

SIMD 计算机设有一个控制单元和 n 个操作单元, 该控制单元一次只对一条指令进行译码,

但是这条指令同时向 n 个操作单元分配数据。

MISD 计算机设有 n 个控制单元和 n 个操作单元, 这些控制单元分别对 n 条指令进行译码, 这 n 条指令在 n 个操作单元上对同一数据流及其中间结果分别进行不同的处理, 一个处理单元的输出作为另外一个处理单元的输入。这类计算机目前没有实用价值, 很少见。

MIMD 计算机是指各种多处理器系统, 它设有 n 个控制单元、 n 个操作单元以及 n 个局部存储器。这 n 个局部存储器构成一个共享主存空间。

如果访问任何一个局部存储器中任意一个存储单元所花费的时间是一样的, 则这样的多处理器系统称为“一致存储器访问 (Uniform Memory Access, UMA)”的多处理器系统, 否则称为“非一致存储器访问 (Non-Uniform Memory Access, NUMA)”的多处理器系统。

如果局部存储器只能由与其对应的控制单元和操作单元访问, 一个控制单元和操作单元只有通过基于计算机网络的“消息传递 (Message Passing)”才能获得其他控制单元和操作单元对应的局部存储器中的数据, 这样的 MIMD 计算机称为多计算机系统。

1.5 计算机的性能评价指标

评价计算机系统性能的指标非常多。与计算机主机有关的评价指标包括: 基本字长、主存储器容量、处理速度、存储器的存取周期、处理器主频、软件兼容性等。与计算机系统有关的评价指标包括: 系统软件的配置、吞吐率、响应时间、辅助存储器容量、外围设备的配置、RASIS 特性、可扩展性等。

1. 基本字长

基本字长是指处理器中的算术逻辑单元所输入的操作数的二进制位数, 也是处理器内部标准的数据寄存器所包含的二进制位数。

计算机处理的一个字 (Word, W) 通常是由若干个字节 (Byte, B) 组成的。一般情况下, 一个字节包括 8 个二进制位 (这与早期 IBM 公司大量生产以 8 位为一个字节的存储器芯片有关), 所以字长基本上都是 8 的整数倍。早期微型计算机的基本字长有 8 位、16 位, 目前主要是 32 位, 64 位的微处理器也已出现在市场上。

基本字长决定了计算机中数据表示的范围与精度, 所以它是评价计算机性能最重要的指标。字长越大, 计算机所能表示与处理的数据的范围越大、精度越高, 但是计算机 (处理器) 的价格也就越高。所以在一些低端的数字设备或工业控制计算机中仍然大量使用字长为 16 位、甚至 8 位的微处理器。

2. 主存储器容量

由于处理器只能访问主存储器, 所以主存储器容量对计算机性能的高低起着重要的影响。主存储器容量常用主存储器的存储单元个数乘以存储单元宽度来表示, 如 1024×16 表示主存储器有 1024 个单元, 每个单元的宽度是 16 位的。计算机系统主存储器的最大存储单元个数取决于处理器地址总线的线数/宽度。

目前, 微型计算机都采用以字节作为最小的编址单位, 即每个存储单元都是 8 位的, 所以表示主存储器容量时, 存储单元宽度可用 B (字节的第一个英文字母) 表示。同时, 在表示存储单元个数时, 为了简便起见, 将 $1024 (2^{10})$ 记为 1K (Kilo), $1024K (2^{20})$ 记为 1M (Mega), $1024M (2^{30})$ 记为 1G (Giga), $1024G (2^{40})$ 记为 1T (Tera), $1024T (2^{50})$ 记为 1P (Peta), $1024P (2^{60})$ 记为 1E (Exascale), $1024E (2^{70})$ 记为 1Z (Zetta), $1024Z (2^{80})$ 记为 1Y (Yotta)。

现代微型计算机的主存储器容量一般从几 GB 到几十 GB, 超级计算机的主存储器容量可达 1TB 以上。

计算机系统的实际主存储器容量的确定主要考虑的是它对计算机价格的影响。增加主存储器容量有助于提高解题速度, 但是计算机系统的价格也会相应升高。一般来说, 如果继续增大主存储器容量已经不能明显地提高解题速度, 则不宜继续增大主存储器容量。

3. 处理速度

计算机处理速度是用户最为关心的性能指标。目前常用的指标有: 百万条指令每秒 MIPS (Million Instructions Per Second)、百万次浮点操作每秒 MFLOPS (Million Floating Point Operation Per Second) 和每条指令的平均时钟周期 CPI (Cycles Per Instruction)。

MIPS 源于美国人吉伯森 (Gibson) 提出的以指令的平均执行时间来评价处理器性能的观点。指令的平均执行时间就是各种不同指令执行时间的某种加权平均。一种很自然的权就是指令的使用频率, 吉伯森和弗林分别在 IBM 7090 和 IBM S/360 上统计出的指令使用频率如表 1-1 所示。指令平均执行时间的倒数就是每秒平均执行的指令条数 (以 MIPS 为单位), 目前常以 MIPS 作为评价处理器综合性能的指标。

表 1-1 指令使用频率 (单位: %)

指令类型	吉伯森	弗林
存/取	31.2	45.1
变址	18	
转移	16.6	27.5
比较	3.8	10.8
定点加/减	6.1	7.6
定点乘	0.6	
定点除	0.2	
浮点加/减	6.9	3.2
浮点乘	3.8	
浮点除	1.5	
移位/逻辑运算	6.0	4.5
其他	5.3	1.3

$CPI = \text{一个程序的 CPU 时钟周期数} / \text{该程序的指令条数}$ 。
它是一个衡量计算机组成设计优劣、计算机性能高低的有效指标, 当今微处理器的 CPI 普遍小于 1。

以上指标都是针对计算机的综合性能而制定的。在用户决定购买或使用哪种机器最适合他们的应用需求时, 往往是运行一些具有代表性的典型应用程序来做出判断的, 这样的典型应用程序被称为“基准程序 (Benchmark)”。

可选择的基准程序很多。不同的基准测试程序侧重的目标是不同的: 有的测试定点运算能力, 有的测试浮点运算能力, 有的测试网络通信能力, 等等。

目前国际上流行的基准测试程序有: 测试整数、浮点数与三角函数等运算性能的 Whetstone 和测试整数与逻辑运算性能的 Dhrystone, SPEC (Standard Performance Evaluation Cooperation) 系列, 测试方程组求解能力的 LinPACK 等。

另外, 一台计算机运行不同的基准程序可能会得到不同的测试结果, 其中测试到的最高速度被称为“峰值速度”或“峰值性能 (Peak Performance)”。

4. 主频

处理器的工作是在主时钟的控制下进行的, 主时钟的频率称为处理器的主频。主频的倒数称为时钟周期。

执行一个程序所需的处理器时间可用“该程序的指令条数 \times CPI \times 时钟周期”来估算。

可见, 提高主频有助于缩短程序的执行时间。原先微处理器的主频在几兆赫 (MHz) 到几百兆赫之间, 随着器件技术的迅速发展, 目前主流微处理器的主频已经达到上千兆赫兹 (GHz)。但事实上, 处理器性能的提高并不能与主频的提高一起线性增长。相反, 主频的提高却带来了功耗增加、产生热量高等一系列问题。

5. 存储器的存取周期

对存储器进行一次完整的读/写操作所需的全部时间, 也是连续对存储器进行存 (写) / 取

(读)的最小时间间隔,称为存储器的存取周期。半导体存储器的存取周期通常在十几到上百纳秒(10^{-9} s, ns)之间,磁盘的存取周期一般在10毫秒(10^{-3} s, ms)以上。

6. 功耗

随着主频和片内晶体管数量的不断提高,微处理器的功耗也不断升高,现代通用微处理器功耗的峰值已经超过100W。随着节能意识的不断增强,人们越来越重视降低微处理器的功耗。例如IBM公司的POWER6处理器是其上一代POWER5处理器的2倍,但运行和散热所消耗的电能耗却基本相同。

事实上,超过150W的功耗,无论是目前芯片的封装还是主板的供电能力,都难以为继。在移动计算领域,功耗更是压倒一切的性能指标。

7. 软件兼容性

软件兼容可分为向上(下)兼容和向前(后)兼容。

“向上(下)兼容”是指为某档机器编制的软件,不加修改就可以正确运行在比它更高(低)档的机器上;“向前(后)兼容”是指为某个时期投入市场的某种型号机器编制的软件,不加修改就可以正确运行在比它早(晚)投入市场的相同型号机器上。

一个计算机系列,从开始设计到最后被淘汰,往往要持续10年以上的时间。在这段时间内,计算机体系结构不可能是一成不变的,它总是要在满足用户新的需求、改正原有的设计失误中不断更新变化。但是这种变化不能是随意的,它必须满足软件兼容性的要求,即在同一系列中的所有机型之间必须做到“向后兼容”(为某个年代投放市场的机器编制的程序应能够不加修改地正确运行于在它之后投放市场的机器上),力争做到“向上兼容”(为某档机器编制的程序应能够不加修改地正确运行于档次比它高的机器上),对于“向下兼容”或“向前兼容”不做要求。

8. 系统软件的配置

常见的系统软件有操作系统、数据库系统、文本编辑器、高级语言程序开发环境、互联网浏览器等。不同的系统软件,其性能也不同,价格差别也很大。

9. 吞吐率(Throughput)、响应时间(Response Time)与周转时间(Turnaround Time)

吞吐率是指计算机系统在单位时间内完成的任务数,响应时间是指从用户输入命令或数据到获得第一个结果的时间间隔,周转时间是指从提交作业到作业完成的时间间隔。

计算机系统结构的设计目标是追求高性能、低成本。但同样一台计算机,从不同角度,会得到不同的性能评价结论。例如,从用户的角度看,响应时间短的计算机好;从计算机系统管理员的角度看,系统吞吐率高的计算机好。

10. 辅助存储器容量

常用的辅助存储器有硬盘、磁带、光盘。辅助存储器容量决定了计算机系统所能够存储的信息总量。辅助存储器的组成形式有:单一的硬盘、硬盘阵列、磁带库、光盘。单一硬盘的容量可达几百GB、甚至数TB,而磁带库的容量则在几千TB以上。

11. RASIS 特性

可靠性(Reliability)、可用性(Availability)、可服务性/可维护性(Serviceability)、完整性(Integrity)和安全性(Security)统称为RASIS特性。

其中,可靠性用“平均无故障时间MTTF(Mean Time To Failure)”或“平均故障间隔时间MTBF(Mean Time Between Failure)”来衡量,可服务性/可维护性用“平均修复时间MTTR(Mean Time To Repair)”来衡量,可用性 $A=MTBF/(MTBF+MTTR)$ 。(信息)完整性是指信息

在计算机系统中，不应被未经授权者修改和破坏，始终保证数据的一致性。（信息）安全性是指通过采取某些安全保密措施，计算机系统中信息不会泄露给未经授权者。

12. 可扩展性（Scalability）

可扩展性也叫可伸缩性。如果一个计算机系统在保证软件兼容性的同时，不仅可以向上扩展（增加资源）来提供更高的性能和更强的功能，还能够通过向下收缩（减少资源）来降低价格，则称这个计算机系统具有可扩展性。

13. 外设的配置

为了拓展计算机系统的功能，需要为它配置相应的外设。通常，计算机系统要尽可能满足用户配备不同类型、不同数量外设的需求。所以，主机与外设的接口应该设计成可扩展的。

以上讨论的性能指标可以称为绝对的性能指标，即不需参照的性能指标。还有一些性能指标是相对的，即需要对应一定的参照，例如加速比（Speedup）。加速比是指解决某个问题，在原有的计算条件（算法、程序或者硬件平台）下所花费的时间与在新的计算条件下所花费的时间的比值。

【例 1-2】MBR 的位数取决于____，MAR 的位数取决于____。

- I. 机器字长 II. 指令字长 III. 存储字长 IV. 主存地址空间大小
A. I、II B. II、III C. II、IV D. III、IV

答：MBR 的位数取决于存储字长，MAR 的位数取决于主存地址空间大小，主存容量常指实际的主存储器容量，它一般小于主存地址空间大小。故选 D。

【例 1-3】（2011 年硕士研究生入学统一考试计算机专业基础综合考试试题之选择题）

下列选项中，描述浮点数操作速度指标的是_____。

- A. MIPS B. CPI C. IPC D. MFLOPS

答：只有选项 D 带“F（意味着浮点数）”，故选 D。

【例 1-4】下列关于计算机性能的说法，正确的是_____。

- A. 指令条数少的代码序列执行时间一定短。
B. 同一程序在时钟频率不同的系列机上运行，时钟频率提高的倍数等于执行速度提高的倍数。
C. 执行不同程序，测得的同一台计算机的 CPI 可能不同。
D. 执行不同程序，测得的同一条机器指令的 CPI 可能不同。

答：功能复杂的指令要比功能简单的指令执行时间长，所以指令条数少的代码序列执行时间不一定短；程序执行过程中，需要访存和 I/O，处理器时钟频率的提高不能直接导致访存和 I/O 时间的缩短；执行不同程序（包含的指令类型和指令条数不同），测得的同一条机器指令的 CPI 原则上是相同的，但同一台计算机的 CPI 可能不同，故选 C。

【例 1-5】某计算机有 I_1 、 I_2 、 I_3 和 I_4 四条指令，其 CPI 分别为 1、3、4 和 5。某程序先被编译成目标代码 A，A 包含这四条指令的条数分别是 3、6、9 和 2。采用优化编译后，该程序得到的目标代码为 B，B 包含这四条指令的条数分别是 10、5、5 和 2。问：哪个目标代码包含的指令条数少？哪个目标代码的执行时间短？A 和 B 的 CPI 分别是多少？

答：A 包含的指令条数为： $3+6+9+2=20$ 。B 包含的指令条数为： $10+5+5+2=22$ 。

所以，A 包含的指令条数少。

A 的时钟周期数为： $3 \times 1 + 6 \times 3 + 9 \times 4 + 2 \times 5 = 67$ ；

B 的时钟周期数为： $10 \times 1 + 5 \times 3 + 5 \times 4 + 2 \times 5 = 55$ 。

可见，B 的时钟周期数少，执行时间短。

程序的 CPI = 程序总的时钟周期数 / 程序包含的指令条数。

所以，A 的 CPI = $67/20 = 3.35$ ，B 的 CPI = $55/22 = 2.5$ 。

【例 1-6】已知某程序编译得到的目标代码 A 包含四类指令 I_1 、 I_2 、 I_3 和 I_4 ，其 CPI 分别为 2、3、4 和 5，它们在目标代码中所占比例分别为 40%、20%、30%、10%。采用优化编译后，该程序得到的目标代码为 B，B 中 I_3 的指令条数减少了 20%，其他指令的条数没有变化。请问：

(1) A 和 B 的 CPI 分别是多少？

(2) 设机器的主频为 1GHz，基于 A 和 B 测得的机器 MIPS 分别是多少？

答：优化编译后，目标代码中各类指令所占比例如下。

$$I_3 : [30 \times (1 - 20\%)] / \{ [30 \times (1 - 20\%)] + 40 + 20 + 10 \} \times 100\% = 25.53\%$$

$$I_1 : 40 / (40 + 20 + 24 + 10) \times 100\% = 42.55\%$$

$$I_2 : 20 / (40 + 20 + 24 + 10) \times 100\% = 21.28\%$$

$$I_4 : 10 / (40 + 20 + 24 + 10) \times 100\% = 10.64\%$$

(1) A 的 CPI = $2 \times 40\% + 3 \times 20\% + 4 \times 30\% + 5 \times 10\% = 3.1$ 。

$$B \text{ 的 CPI} = 2 \times 42.55\% + 3 \times 21.28\% + 4 \times 25.53\% + 5 \times 10.64\% = 3.04。$$

(2) 基于 A 测得的机器 MIPS = $1G/3.1 = 1000M/3.1 = 322.58 \text{ MIPS}$ 。

$$\text{基于 B 测得的机器 MIPS} = 1G/3.04 = 328.95 \text{ MIPS}。$$

【例 1-7】某计算机有甲、乙、丙三类指令，其 CPI 分别为 1、2、5。编译器使用不同的优化编译技术对某应用程序进行编译，得到两个功能等价但指令序列不同的目标代码 A 和 B。已知 A 中甲类指令有 5 条，乙类指令有 3 条，丙类指令有 1 条；B 中甲类指令有 3 条，乙类指令有 2 条，丙类指令有 2 条。问：在理想情况下，哪个目标代码运行的时间短？

答：A 的执行时间 = 5 条甲类指令 \times CPI_甲 + 3 条乙类指令 \times CPI_乙 + 1 条丙类指令 \times CPI_丙
 $= 5 \times 1 + 3 \times 2 + 1 \times 5 = 16$ 个时钟周期；

B 的执行时间 = 3 条甲类指令 \times CPI_甲 + 2 条乙类指令 \times CPI_乙 + 2 条丙类指令 \times CPI_丙
 $= 3 \times 1 + 2 \times 2 + 2 \times 5 = 17$ 个时钟周期。

所以，目标代码 A 运行的时间短。

【例 1-8】(2013 年硕士研究生入学统一考试计算机专业基础综合考试试题)

某计算机主频为 1.2GHz，其指令分为 4 类，它们在基准程序中所占比例及 CPI 如表 1-2 所示。

表 1-2 例 1-8 数据表

指令类型	所占比例	CPI
A	50%	2
B	20%	3
C	10%	4
D	20%	5

该机的 MIPS 是_____。

A . 100

B . 200

C . 400

D . 600

答：该计算机指令集的 CPI = $2 \times 50\% + 3 \times 20\% + 4 \times 10\% + 5 \times 20\% = 3$ 。

该机的运算速度=主频/CPI=1.2G/3=400MIPS。故选 C。

【例 1-9】(2012 年硕士研究生入学统一考试计算机专业基础综合考试试题)

假定基准程序 A 在某计算机上的运行时间为 100s, 其中 90s 为 CPU 时间, 其余为 I/O 时间。若 CPU 速度提高 50%, I/O 速度不变, 则运行程序 A 所耗费的时间是_____。

A . 55s

B . 60s

C . 65s

D . 70s

答: 执行时间=CPU 时间+I/O 时间=90/(1+50%)+10=90/1.5+10=60+10=70s。故选 D。

1.6 微处理器与微型计算机

1.6.1 微处理器与微型计算机的产生与发展

微处理器的发展历史, 主要就是美国英特尔 (Intel) 公司的发展历史。

英特尔公司 1968 年 7 月 18 日由戈登·摩尔 (Gordon Moore) 集成电路的发明人鲍勃·诺伊斯 (Bob Noyce) 和旧金山风险投资人洛克 (Arthur Rock) 共同创立。英特尔公司创立不久, 安迪·格鲁夫 (Andy Grove) 也加入到创业者的行列中。英特尔公司开始只生产半导体存储芯片, 这种存储芯片是当时流行的磁芯存储器的替代产品。

真正成就英特尔公司辉煌的产品——微处理器的开发始于 1969 年。当时, 一家名为 Basicom 的日本公司要求英特尔公司为一款新型袖珍计算器设计生产一个简单的专用处理芯片, 摩尔就将这个任务交给了工程师霍夫 (Marcian Hoff), 并希望他设计成标准化的产品, 以便卖给更多的用户。霍夫很快实现了摩尔的目标, 设计出来的产品不仅可以应用于袖珍计算器, 还可以用来控制交通信号灯以及家用电器。

发展到这个地步, 对于英特尔公司来说, 已不只是技术上的突破, 而是商业上的一大拓展。于是, 摩尔请更资深的工程师费根 (Federico Faggin) 接手领导设计标准化的 4 位处理芯片。1971 年 11 月, 英特尔公司推出了编号为 4004 的处理芯片, 其名中的第一个“4”表示 4 位的字长, 最后一个“4”表示它是该公司推出的第 4 款专用芯片。4004 采用“P 型金属氧化物半导体 (P-Metal-Oxide-Semiconductor, PMOS)”工艺, 含有 2300 个晶体管, 时钟频率为 108kHz, 寻址空间为 640B, 售价为 200 美元。这就是世界上第一个微处理器。

费根还同时开发出另外三款芯片: 4001、4002、4003, 分别是随机存取存储器 RAM (Random Access Memory)、只读存储器 ROM 和寄存器。这四颗芯片合起来就可以组成一台微型计算机。由于微处理器采用大规模集成电路技术, 所以微型计算机属于第四代计算机。

自 Intel 4004 问世以来, 微处理器和微型计算机如雨后春笋般蓬勃发展起来, 逐渐占领了计算机市场, 有力地推动了计算机的普及。目前, 各种类型的微处理器和微型计算机在国民经济的各个领域以及人们的日常生活中得到了广泛的应用。微处理器和微型计算机之所以发展如此迅速, 其主要原因就是其性能价格比在各类计算机系统中处于领先地位。另外, 微型计算机维护方便、小巧灵活的特点, 也受到了人们的青睐。

至今, 微处理器和微型计算机的发展经历了五代。

第一代 (1971—1972 年): 4 位的微处理器和微型计算机, 例如以 Intel 4004 为处理器加上一片 320 位的 RAM、一片 256 字节的 ROM 和一片 10 位的寄存器, 通过总线连接在一起就构成了世界上第一台的微型计算机 MCS-4。可使用机器语言或汇编语言进行编程, 能够进行十进

制的算术运算。作为第一代微型计算机虽然不够完善,但由于价格较低,所以一经问世就赢得了市场。于是,英特尔公司对它进行了改进,正式生产了通用的4位微处理器4040。这些微处理器让英特尔公司获得了巨大的成功。

第二代(1972—1977年):8位的微处理器和微型计算机。1972年4月,英特尔公司推出了第一个8位微处理器8008,它含有3500个晶体管,时钟频率为108kHz,寻址空间为16千字节(Kilo-Byte,KB)。1974年,集成了4900个晶体管、时钟频率为2MHz、寻址空间为64KB的Intel 8080问世,随即以8080为中央处理器的微型计算机Altair问世。

1975年,年仅19岁的哈佛大学大二学生比尔·盖茨(Bill Gates)创立了面向微型计算机软件开发的微软(Microsoft)公司,开始了他“让每一个家庭,每一张桌子上都有一台计算机”的事业。微软公司开发的第一个软件就是为Altair编写了BASIC语言的解释程序。

1974年,摩托罗拉(Motorola)公司也推出了集成了6800个晶体管的8位的微处理器M6800。

4004的设计者费根离开英特尔公司,成立了齐洛格(Zilog)公司。1976年,齐洛格公司推出了集成有10000个晶体管的Z-80。Z-80完全兼容8080,而性能明显优于8080。所以Z-80迅速走红,成为当时最受欢迎的微处理器。

之后,英特尔公司和摩托罗拉公司又分别推出了性能更高的Intel 8085和M6809。从1973年起,主流微处理器(如Intel 8080、M6800、Z-80)开始采用开关速度更快的“N型金属氧化物半导体(N-Metal-Oxide-Semiconductor, NMOS)”工艺。尽管8080与8008具有相同的指令集,但是性能却提高了近10倍。

从第二代微处理器来看,英特尔公司的8080与8085、摩托罗拉公司的M6800和齐洛格公司的Z-80构成三足鼎立之势,垄断了市场。

第二代微处理器的特点是软件日益丰富,除了汇编程序外,通常还配有BASIC、FORTRAN及PL/M等高级语言的解释程序和编译程序。面向微型计算机的操作系统也被开发出来。价格低廉、使用方便的软件极大地推动了微型计算机的普及。

第三代(1978—1983年):16位的微处理器和微型计算机。1978年6月,英特尔公司发布了第一个16位微处理器8086,它含有29000个晶体管,时钟频率为4MHz,数据总线为16位,地址总线为20位。1979年,8086的变型产品8088问世,8088与8086的不同主要体现在数据总线降为8位。同年,摩托罗拉公司推出了集成有68000个晶体管的M68000。齐洛格公司也相继推出了集成有37500个晶体管的Z-8000。Intel 8086/8088、M68000和Z-8000都是早期16位的微处理器的典型代表,它们的主频为4~8MHz,平均指令执行时间为0.5 μ s。

20世纪80年代以后,英特尔公司又推出了性能更高的16位微处理器80286,摩托罗拉公司推出了M68010。后期的16位微处理器主频超过10MHz,平均指令执行时间为0.2 μ s,集成度超过100000个晶体管/片。

这时,面向微型计算机的操作系统、数据库系统日趋完善,各种高级语言的解释程序和编译程序也相继开发出来,微处理器还被用来构成多处理器系统。微型计算机从性能上开始超过小型计算机,在市场上成为小型计算机的竞争对手。由于价格低廉、实时性能优异,微处理器还被广泛应用于实时数据处理和工业控制等领域。

第四代(1984年以后):32位的微处理器和微型计算机。1984年,摩托罗拉公司率先推出了首个32位的微处理器M68020。1985年,英特尔公司发布了它的第一个32位微处理器

80386, 它含有 45 万个晶体管, 内部设置了一级高速缓冲存储器 (Cache)。这一时期的微处理器的时钟频率一般在 20 ~ 40MHz, 平均指令执行时间为 0.1 μ s。1989 年, 英特尔公司发布了它的第二个 32 位微处理器 80486, 将第四代微处理器的性能又大大提高一步。

英特尔公司的第三个 32 位微处理器是在 1993 年推出的。当时, 为了防止竞争对手跟风搭车, 也因为数字无法登记注册为商标, 所以英特尔公司决定放弃以 586 来命名新的微处理器, 而采用了一个响亮的名字“Pentium”。“Pent”在拉丁文中表示“第五”的意思, “ium”是结尾音, 可使整个词听起来很响亮。Pentium 的时钟频率为 150MHz, 平均指令执行时间为 0.05 μ s。随后, Intel 公司又推出了 Pentium 的增强型产品 Pentium Pro 和 Pentium MMX。

1997 年, 英特尔公司发布了它的第四个 32 位微处理器 Pentium II。1999 年, Pentium III 问世。2000 年, Pentium 4 投入市场。

第五代 (1992 年以后): 64 位的微处理器和微型计算机。1992 年, 美国 DEC 公司率先推出了首个 64 位的微处理器 Alpha 21064。其名意味着: 21 世纪的 64 位微处理器。最新的 Alpha 微处理器为 21364。

2001 年, 英特尔公司和惠普 (Hewlett-Packard, HP) 公司联合推出了基于 IA-64 体系结构的 64 位的微处理器——安腾 (Itanium)。2002 年推出安腾-2, 2003 年推出改进的安腾-2。

1995 年, IBM 公司和摩托罗拉公司联合发布了它们的第一款 64 位的微处理器 PowerPC 620。2007 年 7 月, IBM 公司推出新一代 64 位微处理器——Power 6。

多年以来, 单个微处理器芯片内部的晶体管数目一直在以每 18 ~ 24 个月增加 1 倍的速度增长着。这个规律是由摩尔于 1965 年发现并预测的, 故称为“摩尔定律 (Moore's Law)”。摩尔于 1975 年对摩尔定律做了一次修改, 最终确定为“每隔 18 个月, 微处理器内部的晶体管数目增加一倍, 同时计算性能翻一番, 而价格保持不变”。

微处理器按照摩尔定律发展了 30 多年。进入 21 世纪, 微处理器的发展趋势已经变缓, 由原先的 18 个月一代变为约 40 个月一代。

1.6.2 多核微处理器

进入 21 世纪, 处理器领域可谓是风起云涌, 似乎在一夜之间, 处理器就进入了多核时代。

首先是 IBM 公司于 2001 年推出了集成了 2 个 Power 3 核的 Power 4 双核处理器。然后在 2004 年, IBM 公司、AMD 公司和 SUN 公司分别推出双核处理器 Power 5、Opteron 和 UltraSPARC IV。

英特尔公司不甘落后, 在 2005 年 10 月推出了双核处理器 Paxville DP, 随后在 2006 年 5 月发布了升级版的双核处理器 Dempsey, 7 月又发布了基于 Core 微架构的 Woodcrest 处理器, 接着在 12 月推出了四核处理器 Clovertown。

多核处理器 (Multi-core Processor), 也叫片上多处理器 CMP (Chip Multiprocessor), 是指在一个芯片中集成两个或多个处理内核的处理器。相对而言, 传统的处理器被称为单核 (Single-core) 处理器。多核处理器中的内核首先是一个完整的处理单元, 能够独立执行指令; 其次, 它们往往具有设计简单、功耗低的特点。

在多核处理器运行过程中, 操作系统将每个执行内核作为一个独立的逻辑处理器。通过在并行执行的内核之间分配任务, 多核处理器可在特定的时钟周期内执行更多的任务。

在多核处理器发展之前, 商业化处理器一直致力于单核处理器的发展。但应用对处理器性

能需求的增长速度远远超过处理器的发展速度。单核处理器越来越难以满足要求，其局限性也日渐突出。

单核处理器的局限性主要表现在如下三个方面。

(1) 仅靠提高频率的办法，难以实现性能的突破。当处理器的工作频率提高到 4GHz 时，几乎接近目前集成电路制造工艺的极限。

(2) 单一处理器内部器件的增加，导致以下两个方面的后果：一是不断增加的芯片面积提高了生产成本；二是设计和验证所花费的时间变得更长。处理器的性价比已经令人难以接受，速度稍快的处理器的价格要高很多。

(3) 功耗与散热问题日渐突出。目前，通用处理器的峰值功耗已经高达上百瓦，例如 AMD 的 Opteron 是 90W，英特尔的安腾-2 已经超过 100W。相应地，主板上向处理器供电的电流接近 100A，与发动汽车时蓄电池供出的电流差不多。功耗增加的原因主要是处理器工作频率不断上升。

随着功耗的上升，超快单核芯片的冷却代价也越来越高，它要求采用更大的散热器和更有力的风扇，以降低其工作温度。否则，过高的温度将导致处理器的性能和稳定性下降。

所以，英特尔和 AMD 采用多核技术的真正原因，不是因为多核技术是一种突然出现的一种优秀创意，而是因为利用多核技术既可以继续提高处理器性能，又可以暂时避开功耗和散热难题。因此，要了解多核处理器，首先要从并行性 (Parallelism) 谈起。

所谓并行性是指问题或者任务中具有可以同时进行计算或操作的特性。它包括同时性 (Simultaneity) 和并发性 (Concurrency) 两重含义。同时性是指两个或者多个事件在同一时刻发生，并发性是指两个或者多个事件在同一时间间隔内发生。

能够实现并行性的计算机被称为并行计算机。依照弗林分类法给出的四类计算机模型，能够用来构建并行计算机的只有 SIMD 型和 MIMD 型计算机。

根据 SIMD 型构建的并行计算机一般包括以下三种。

(1) 向量处理器 (Vector Processor)

向量处理器采用的是时间重叠技术 (流水线) 来实现并行处理。具体来说就是将一个功能单元划分成若干个不同的子单元。按照一定的顺序，这些子单元组成一个处理流水线，分别完成这个功能的部分处理任务。这样，这些子单元就可以在同一时刻，对多个操作数进行并行处理。Cray-1、Convex C1 和国产的 YH-1 等就属于向量处理器。

(2) 阵列处理器 (Array Processor)

阵列处理器由大量的相对简单的处理单元组成，这些处理单元可以在同一时刻，对多个操作数或操作数的位片进行并行处理。美国 Burroughs 公司生产的 ILLIAC-IV、英国 ICL 公司生产的 DAP 和国产的 150-AP 等就属于阵列处理器。

(3) 相联处理器 (Associative Processor)

相联处理器是以相联存储器为核心，配上必要的中央处理部件、指令存储器、控制器和 I/O 接口，构成的以“存储器操作并行”为特征的计算机。

相联存储器是一种按照内容访问、具有信息处理功能的存储器。它可以按所给内容的部分或全部特征，同时对存储器中所有存储单元进行并行访问。这里的访问，不仅仅是读/写存储单元，还包括对各个存储单元的有关字段内容并行地进行比较、符合 (判断是否相等、是否大于/大于或等于/小于/小于或等于、是否在区间内/区间外以及是否是最大值/最小值/次大值/次小

值等)、分解等处理。最后将内容与该特征相符的所有存储单元在一次存储器访问中全部都找出来。所以,相联处理器是基于存储器操作并行的 SIMD 型并行计算机。

根据 MIMD 型构建的并行计算机一般包括以下三种。

(1) 共享存储的多处理器系统 (Shared Memory Multi-Processors)

共享存储多处理器系统是指多个处理器通过互连网络共享一个统一的主存储器空间,并通过这个主存储器空间来实现处理器之间的协调。这个统一的主存储器空间可以是一个庞大的存储器单元,但在更多情况下是由多个存储器模块组成的。

在共享存储多处理器系统中,不同的处理器可以执行相同或不同的指令流,可以直接访问到所有的主存储器单元,处理器之间的通信是通过共享主存储器来实现的。

组成共享存储多处理器系统的处理器可以是普通的标量处理器(如 SGI 的 Power Challenge),也可以是高性能的向量处理器(如 Cray 的 Cray X-MP、Cray Y-MP, Fujitsu 的 VPP500, NEC 的 SX-3, Hitachi 的 S-3800, Convex 的 C3, 国产的 YH-2 等)。

共享存储多处理器系统存在的一个主要问题是可扩展性差。当处理器规模增加时,由于处理器需要同时访问共享全局变量而导致的主存储器竞争急剧增加,反而使并行计算效率下降。因此,共享存储多处理器系统主要适合于中小规模的并行处理问题。

所以从理论上讲,多核处理器就是一种共享存储的多处理器系统。

(2) 分布存储的多计算机系统 (Distributed Memory Multi-Computers)

在分布存储多计算机系统中,各个计算节点就是一个独立的计算机系统,即计算节点是由一个独立的处理器及其私有的主存储器空间组成。计算节点通过互连网络相互连接从而构成了一个并行计算系统。由于每台处理器只能访问到局部的存储器,对其他存储器的访问必须以消息传递的方式实现,故称之为分布存储多计算机系统。

分布存储多计算机系统克服了共享存储多处理器系统的缺点,具有良好的可扩展性,是目前实现超大规模科学与工程计算最重要的工具。

分布存储多计算机系统目前主要有两种实现形式:一种是通过专用的高性能互连网络来实现各个计算节点的互连,另一种是采用常规的计算机网络系统来实现各个计算节点的互连。前者一般被称为紧耦合的多计算机系统,后者则被称为工作站机群系统。

(3) 多线程处理器/计算机系统。

由于指令之间存在着相关性(如分支指令、后一条指令需要前一条指令的运算结果、两条指令同时访问一个存储单元等),所以在一个正在执行的程序[也叫一个进程(Process)]中找到大量可以并行执行的指令是困难的。

为此,计算机科学工作者引入了线程(Thread)的概念,所谓线程是指程序内部的一段功能/操作相对独立的指令序列。线程内部的指令是顺序执行的,但是不同线程是可以并行执行的。这就是线程级并行(Thread Level Parallelism, TLP),支持 TLP 的处理器就是多线程(Multiple Thread, MT)处理器。

在多线程的意义下,普通的程序被认为是单线程的。当然,普通的程序也可以通过特殊的编译器编译成多线程的程序。

同样在多线程的意义下,处理器可以分为:单线程处理器和多线程处理器。这两种处理器都可以分别成为多核处理器的内核。

目前主流的多线程处理器主要有两种:多线程芯片 CMT (Chip Multithreading) 与同时多

线程处理器 SMT (Simultaneous Multithreading)。

其中，CMT 芯片是 SUN 公司的发明，它是片上多处理器 CMP 和多线程 MT 技术的综合。在一个 CMT 芯片中，具有多个普通的（单线程）处理器，这些处理器同时工作，分别执行不同线程的指令。不过，运行在 CMT 芯片上的程序必须是多线程的程序。

而 SMT 处理器是具有多个功能部件（如多条指令流水线）的单核处理器，它允许在一个时钟周期内，在不同的功能部件上，执行来自不同线程的不同指令。这有利于提高一个核内多个功能部件的利用率。

也就是说，被鱼贯发射到同一条指令流水线上的指令可以来自不同的线程。

这样，SMT 处理器可以同时运行多个用户的程序，每个用户都感觉在独立使用一个普通的单核处理器（逻辑处理器）。这些逻辑处理器通过资源动态分割共享一个物理处理器。运行在 SMT 处理器上的程序可以是多线程的程序，也可以是普通的单线程的程序。

综上所述，SMT 的处理器资源利用率比 CMP 的处理器资源利用率高。特别是在线程级并行不够时，CMP 处理器会存在资源浪费，而在 SMT 处理器中，这个问题不太突出。另外，SMT 处理器的灵活性要优于 CMP。

相比 SMT，CMT 的最大优势体现在模块化设计的简洁性上。不仅复制简单，设计也非常容易，指令调度更加简单，同时也不存在 SMT 中多个线程对共享资源的争用。因此，当应用程序的线程级并行性较高时，CMT 性能一般要优于 SMT。另外，随着集成电路工艺的发展，芯片上的线延迟会超过 CMOS 门延迟占主导地位。在克服线延迟方面，短连线分布式的 CMP 芯片要比长连线集中式的 SMT 处理器更容易提高芯片的运行频率，从而在一定程度上起到性能优化的效果。

目前，多核处理器的主要发展趋势：一是增加核数，二是将通用处理器和协处理器（GPU 或流处理器）集成在一起形成异构多核。有人称核数多于 8 的多核处理器为“众核(Many_core) 处理器”。但是相对于多核处理器中的核，众核处理器一般选用的是更加简单的核。IBM 公司的 CELL 就是一款异构多核处理器。

多核处理器的时代已经到来，但是，构建基于多核处理器的计算机，还面临着计算机整体架构、I/O、操作系统、应用软件等多方面的挑战。例如，多核处理器在实际应用中并没有带来性能的明显提升，其原因是程序并没有提供足够多的线程让多核处理器中的每一个核都“奔腾”起来。

可见，计算机硬件技术应该与软件技术协同发展。在关注多核处理器发展的同时，应该研究与其相适应的“多核程序设计”模式。有兴趣的同学可以选修“多核程序设计”或者“并行程序设计”课。

1.7 中国计算机事业的发展历程

新中国计算机科学与技术的发展始于 1956 年国务院制定的《1956—1967 年科学技术发展远景规划纲要》，其中的第 41 项为“计算技术的建立”。为了落实这一规划纲要，中国科学院数学所所长华罗庚教授组织闵乃大、夏培肃、王传英成立了新中国第一个电子计算机科研小组，并于 1956 年 6 月 19 日成立了以华罗庚为主任的“计算技术研究所筹备委员会”，这就是今日中国科学院计算技术研究所的前身。该委员会成立后，立即开展了国产 103 机和 104 机的研制。

这两台机器都属于电子管计算机（第一代计算机），其中 103 机是仿制前苏联的 M—3 小型数字电子计算机。103 机于 1958 年 8 月 1 日仿制成功，结束了我国没有计算机的历史。

104 机是我国科学家在前苏联专家的帮助下，自行研制的大型通用数字电子计算机。该机使用了 4200 个真空管和 4000 多只二极管，每秒钟可以完成 1 万次浮点数运算，功耗为 70kW。它的磁芯体内存只有 2048 字，字长 39 位。它还有两台磁鼓和一台 1 英寸宽的磁带机作为外存。104 机于 1959 年 4 月 30 日清晨调通，并算出第一个课题：“五一”节的天气预报。104 机于 1959 年 10 月 1 日宣布完成，为共和国“十年大庆”献上了一份厚礼。

1956 年，清华大学和哈尔滨工业大学分别成立了计算机专业，揭开了我国计算机教育事业的序幕。1958 年，哈尔滨工业大学和东北大学（原名东北工学院）分别研制成功国产第一台数字式电子计算机和模拟电子计算机。

1960 年 4 月，夏培肃教授领导的科研小组研制出我国第一台自行设计电子计算机——107 机，这是一台小型通用计算机。1964 年，我国第一台自行设计的大型通用数字电子管计算机 119 机在中国科学院计算技术研究所研制成功，平均浮点运算速度为 5 万次每秒。

从 20 世纪 60 年代开始，我国的计算机科学工作者自行设计并成功研制出了一批晶体管计算机（第二代计算机），如哈尔滨军事工程学院研制的 441B 机，计算技术研究所研制的 109 乙机和 109 丙机，华北计算技术研究所研制的 108 机、108 乙机、121 机和 320 机等。

在这期间，我国自主研制的存储设备也取得了显著的成果。1964 年，天津大学许镇宇教授主持研制成功“立式双浮动磁鼓存储器（48 磁头浮动，锥形磁鼓也浮动）”。1966 年，哈尔滨工业大学的陈光熙教授以挤压方法代替传统的冲压成型方法，生产出外径为 0.5mm 的磁芯，进而成功研制了超小型磁芯存储器。

进入 20 世纪 70 年代，我国的集成电路计算机（第三代计算机）也相继研制成功。1971 年，中国科学院计算技术研究所和华北计算技术研究所分别研制出 111 计算机和 112 计算机。1973 年，北京有线电厂和北京大学合作研制出 150 计算机，华东计算所研制出 655 计算机，其中 655 机是我国首台具有百万次浮点计算能力的机器。

中国软件领域奠基人之一、中国科学院院士、北京大学教授杨芙清先生主持研制了中国第一台百万次集成电路计算机 150 机操作系统和第一个全部用高级语言编写的操作系统，为中国操作系统的早期发展做出了重大贡献。中国最早的两位计算机软件博士生导师之一、培养出我国第一位软件学博士的南京大学教授徐家福先生于 1965 年主持研制出中国第一个 ALGOL 编译系统。1977 年，他倡导并主持研发系统程序设计语言，在 655 机上设计并实现了基于自编译语言的软件自动产生系统 NDHD。为表彰他们对中国计算机事业的创建、开拓和发展做出的卓越贡献，中国计算机学会（CCF）将 2011 年度 CCF 杰出贡献奖授予这两位先生。

1976 年，国产 DJS-100 系列小型多功能计算机投入批量生产（DJS 分别是电子、计算机、数字式三个词的汉语拼音字头，这是国家计算机产品系列的型号名），与美国 NOVA 小型机兼容的 DJS-130 生产了上千台，该机曾装备在郑州铁路局的调度系统中。

1977 年 200 万次大型集成电路计算机 DJS-183 问世，与 IBM System 360/370 兼容的 DJS-200 系列计算机于 1979 年研制成功。

1981 年，北京大学的王选教授带领的团队成功研制了我国自行设计的计算机——激光汉字编辑排版系统，彻底改造了我国沿用上百年的铅字印刷技术。国产激光照排系统使我国传统出版印刷行业仅用了短短数年时间，从铅字排版直接跨越到激光照排，走完了西方用几十年才

完成的技术改造道路,被公认为自毕昇发明活字印刷术后中国印刷技术的第二次革命。为了纪念王选的成就,CCF于2006年将“中国计算机学会创新奖”更名为“CCF王选奖”,以奖励那些为中国计算机事业做出重大贡献的人士/成果。铁道科学研究院电子计算技术研究所朱建生等研制的“中国铁路客票发售和预订系统V5.0”获得当年的二等奖。2007年,北京中科辅龙计算机技术股份有限公司唐卫清研究员等人完成的“PDSOFT计算机辅助工厂设计系统”获得一等奖。中国科学院计算技术研究所的徐志伟研究员撰写的科普著作《电脑启示录》(上、中、下三篇)获得二等奖。

1983年,中科院计算技术研究所研制了我国第一台大型向量计算机——757机,它的计算速度达到1000万次/秒。同年,国防科技大学研制出我国第一台运算速度达到1亿次/秒的超级计算机“银河(YH)-1”,标志着我国的计算机技术达到了世界先进水平。

我国的微型计算机的研究开始于1974年。DJS-050系列、DJS-060系列是我国最早的微型计算机,它们属于8位微型计算机,其中DJS-050系列与Intel 8080兼容、DJS-060系列与Motorola M6800兼容。1981年,骊山微电子公司成功研制了16位微型处理器LS77。1984年,电子工业部和电子科技大学(原名为成都电讯工程学院)分别成功研制了与IBM PC/XT兼容的“长城0520系列”和与Apple机兼容的“紫金-II系列”16位微型计算机。

从1988年5月到1989年9月,金山公司的求伯君独自开发了WPS1.0,填补了我国中文字处理软件的空白。

1992年,国防科技大学成功研制了峰值速度达到4亿次浮点运算/秒的通用超级计算机“银河-2”。1997年,峰值速度为130亿次浮点运算/秒的“银河-3”问世。

自20世纪80年代以来,中国科学院计算技术研究所、国防科技大学、江南计算技术研究所、联想公司等单位先后研制成功了“曙光1000”、“曙光2000”、“曙光3000”、“曙光4000”、“银河-4”、“神州/神威”系列大型/超级计算机、联想深腾1800和6800等。

1989年,西安交通大学成功研制了面向人工智能语言LISP的LISP-M1智能计算机。

2008年6月24日,由中科院计算所国家智能计算机研究开发中心、曙光信息产业(北京)有限公司、上海超级计算中心联合研制并由曙光公司定型制造的集群超级计算机——曙光5000A系统问世,它的峰值速度达到230亿次浮点运算/秒。这标志着我国成为世界上第二个可以研发生产超百万亿次超级计算机的国家。

在2008年11月17日公布的全球高性能计算机500强(TOP500)排行榜中,曙光5000A位列世界超级计算机第十位,是我国高性能计算机的最好成绩(曙光4000也曾经以11万亿/秒的速度位列世界超级计算机第十位)。

2009年10月29日,我国首台千万亿次超级计算机系统——“天河一号”由中国国防科学技术大学研制成功。该计算机峰值性能为1206万亿次双精度浮点数操作,LINPACK实测性能为560.3万亿次,排列“TOP500”第四位。使我国成为继美国之后世界上第二个能够研制千万亿次超级计算机的国家。

天河一号采用6144个Intel通用多核处理器和5120个AMD图形加速处理器,内存容量为98TB,点对点通信带宽为40Gbps,磁盘总容量达到1PB,综合技术水平进入世界前列。

同时,具有我国自主知识产权的高性能微处理器的研发也取得了明显的进展。中国科学院计算技术研究所研制的“龙芯”系列微处理器、北京大学研制的“众志”微处理器、清华大学研制的“THUMP”微处理器和西北工业大学研制的“龙腾R1”微处理器等,已被应用到我国

的科技和经济工作中。

在 2013 年 11 月 18 日公布的 TOP500 强排行榜中,国防科学技术大学研制的“天河二号”以峰值计算速度为 5.49 亿亿次/秒、持续计算速度为 3.39 亿亿次/秒的双精度浮点运算速度再次夺得全球第一。“天河二号”服务阵列采用了由国防科学技术大学研制的新一代“飞腾-1500”CPU,这是当前国内主频最高的自主高性能通用 CPU。

1.8 计算机的特点及应用

1.8.1 计算机的特点

1. 能在程序的控制下自动连续地工作

目前,市场上能见到的计算机采用的都是“冯·诺依曼”体系结构,而“冯·诺依曼”体系结构最核心的特征就是“存储程序”,即控制计算机工作的指令被事先编制好并按执行顺序排列形成所谓的“程序”,程序被存储在计算机控制单元能够直接访问的存储器中。一旦计算机启动,控制单元就顺序地访问存储器,读出指令,逐条解释执行指令。除了执行“输入指令”外,不再需要人们的干预。这是电子计算机的一个基本特点。

2. 运算速度快、计算精度高

计算机采用电子器件作为处理单元,所以运算速度快。随着器件技术的不断发展,计算机的处理速度从几千次每秒,发展到几万次每秒、几百万次每秒、几亿次每秒。目前,计算机的计算速度已达到 1000 万亿次每秒量级。

计算机的计算精度主要取决于机器的字长。只要增加字长,提高计算精度在理论上几乎没有限制。目前,最新的微处理器的字长为 64 位,借助软件还能实现双倍字长、多倍字长的运算,从而达到更高的精度。

3. 具有逻辑判断能力

由于计算机的工作原理是基于逻辑的,所以它还具有逻辑判断能力。例如,计算机能够判断一个数是正数还是负数,能够判断一个数是大于、等于,还是小于另外一个数。有了逻辑判断能力,计算机在工作时,就能对运算结果的性质进行判断,并根据判断结果选择下一步的处理方法。有了逻辑判断能力,计算机除了能够进行数值运算外,还能够对各种信息进行逻辑运算,使得计算机能够模拟人类智能。

4. 通用性强

现代计算机都是数字式的,用数字逻辑部件来处理数字信号,即处理功能逻辑化,这就使得计算机具有统一的逻辑基础。不管是什么信息,只要数字化,计算机就能对其进行算术运算或逻辑运算。

对于不同的问题,计算机只是执行的程序不同而已,计算机的内部组成与结构是不变的,所以说计算机的通用性极强。只要为计算机编制好了应用程序,同一台计算机就能够应用于不同的领域,解决不同的问题。目前,计算机除了在科学计算领域外,还被广泛地应用于信息检索、图像处理、语音处理、工业控制等领域,渗透到社会生活的方方面面。

5. 具有很强的“记忆”功能

计算机系统通常都设置有容量很大的存储器,能够长期保存用户的程序和数据。在需要

时,计算机可以迅速地将存储的内容读出使用或对其进行更改。这是计算机的又一个基本特点。计算机的许多功能都是在此特点的基础上发展起来的。计算机能够存储程序,所以才能够自动地连续工作;计算机存储容量大,可存储的应用程序多,存储的信息量大,计算机的用途才广;通过存储不同的程序来实现不同的功能使计算机具有通用性;计算机访问存储器的速度快,才能支持计算机运算器满负荷、高速的工作。

1.8.2 计算机的应用

1. 科学与工程计算（或数值计算）

科学与工程计算是指利用计算机来完成科学研究和工程设计中所需的数学计算,这是计算机最早的、最重要的应用领域。由于计算机具有计算速度快、计算精度高、存储容量大、能够连续运算的特点,所以它不仅提高了计算的效率,还解决了一些原先靠人工无法解决的科学计算问题。

例如,在建筑设计中为了确定构件的尺寸,需要求解由弹性力学导出的一系列复杂方程。过去由于计算工具和计算方法的限制,只能通过简化,得到这些方程的近似解。而电子计算机的使用,不但能够求解这些方程,还引起了计算方法上的突破,导致了“有限元法”的诞生。

2. 数据处理（或信息处理）

数据处理是指利用计算机来对在生产组织、企业管理、市场分析、情报检索等过程中出现的大量数据进行收集、存储、归纳、分类、整理、检索、统计、分析、列表、图形化等处理。这些处理的特点是算法比较简单,一般只需要进行简单的算术运算和逻辑运算,但是数据量极大。在引入电子计算机以前,这些繁重而单调的数据处理工作是由人工完成的,极易出错。而采用计算机来进行数据处理,不仅速度快,而且质量高。

数据处理从简单到复杂已经历了三个发展阶段。

第一阶段:电子数据处理(Electronic Data Processing, EDP)。它是以文件为对象,实现文件内部的单一数据类型的简单数据处理。

第二阶段:管理信息系统(Management Information System, MIS)。它是以数据库为对象,实现一个部门的多种数据类型的综合数据处理。

第三阶段:决策支持系统(Decision Support System, DSS)。它是以数据库、模型库和方法库为基础,实现多种数据类型、多种数据来源的高层次数据处理,帮助管理者、决策者制定正确、有效的企业运营策略。

3. 过程控制

过程控制是指利用计算机及时采集、检测工业生产过程中的状态参数,按照相应的标准或最优化的目标,迅速对控制对象进行自动调节或控制,也称计算机控制。计算机控制的引入,不仅改善工人的劳动条件,降低了工人的劳动强度,提高控制的及时性和精确性,还通过提高工厂的自动化水平来提高劳动生产率和产品质量,降低产品成本,缩短生产周期。目前,计算机过程控制已广泛地应用在冶金、化工、船舶等领域。

有一类重要的过程控制称为实时控制。所谓实时控制是指要求响应速度极快的一类过程控制。例如,核电站中一些关键设备的临界故障处理,航天/航空中的运动控制等。

4. 辅助技术（或计算机辅助设计与制造）

利用计算机来辅助和改善人们的工作,是当今计算机一个重要的应用。

对汽车、飞机、船舶、建筑、大规模集成电路、新型药物的设计，既追求设计目标优化，又希望缩短设计周期，还要减轻设计者的劳动强度。随着计算机技术的发展与普及，一个新的学科方向——计算机辅助设计 CAD(Computer Aided Design)应运而生。波音公司设计的 B-777 就全部是在计算机上完成的，号称“无纸设计”。

在 CAD 技术的发展过程中，一些新的技术分支不断派生出来，如计算机辅助制造 CAM(Computer Aided Manufacture)、计算机辅助测试 CAT(Computer Aided Test)、计算机集成制造系统 CIMS(Computer Integrated Manufacturing System)等。

计算机辅助技术还可以应用在教育活动中，比如让计算机来演示教学内容、协助教师组卷与判卷、代替教师回答学生提问等。这就是计算机辅助教学 CAI(Computer Aided Instruction)。今天，基于计算机和网络技术的“慕课(Massive Open Online Courses, MOOCs)”正在引发一场高等教育的变革。

5. 智能模拟

智能模拟是用计算机软/硬件系统来模拟人类的某些智能行为，如感知、思维、推理、学习、理解和问题求解等，它是在计算机科学、控制论、仿生学和心理学等学科的基础上发展起来的一门交叉学科。它的核心是人工智能(Artificial Intelligence, AI)，包括：专家系统、模式识别、问题求解、定理证明、机器翻译、机器学习、人工生命、自然语言理解等。AI 始终是计算机科学与技术领域的一个重要的研究方向。

AI 的一个重要成果就是机器人(Robot)，机器人的视觉、听觉、触觉、决策等，都需要人工智能理论与技术的支持。目前，国际上已研发出大量的工业机器人、家用机器人、会下棋的机器人、会踢足球的机器人、会跳舞的机器人，等等。可以相信，随着 AI 理论与技术的不断发展，未来的机器人将会在我们的生产和生活中发挥越来越大的作用。

6. 网络通信

计算机技术与现代通信技术相结合构成了计算机网络(Computer Network)。计算机网络，特别是因特网(Internet)，不仅解决了一个单位、一个城市、一个国家乃至全世界的计算机之间的相互通信、软硬件资源共享等问题，还大大改变了人们的工作习惯和思维习惯，改变着人类的文化。

计算机网络的飞速发展还催生了一些新的学科和应用领域，如物联网、电子商务、数字图书馆、网络存储、海量信息检索，等等。

7. 图形图像处理

将计算机应用到图形图像处理导致出现了两个新的学科分支——计算机图形学(Computer Graphic, CG)和图像处理(Image Processing)，以及两个新的产业方向——游戏制作与动漫制作。

目前，计算机已广泛应用在影视制作、电子游戏、社会安全监控、虚拟现实等中。其中，虚拟现实(Virtual Reality, VR)是利用计算机系统来生成一个极具真实感的模拟环境。通过虚拟现实，人们可以进行训练、娱乐、学习等活动。现在开发的虚拟现实系统已被用来培训驾驶员、宇航员、技术工人，还有人研究基于虚拟现实的旅游。虚拟现实的发展及应用必将深刻地影响着人们的思维方式和生活方式。

8. 物联网(The Internet of Things, IOT)

“物联网”是在“互联网”的基础上，将其用户端延伸和扩展到任何物品与物品之间，进行信息交换和通信的一种网络概念。其定义是：通过射频识别(Radio Frequency Identification,

RFID)、红外感应器、全球定位系统(GPS)等传感设备,按约定的协议,把任何物品与互联网相连接,进行信息交换,以实现智能化识别、定位、跟踪、监控和管理的一种网络。

最早提出物联网的设想和概念的是比尔·盖茨。在1995年出版的《未来之路》中,盖茨描述了未来的住房:当你走进去时,所遇到的第一件事是有一根电子别针夹住你的衣服,这根别针把你和房子里的各种电子服务接通了……凭你戴的电子别针,房子会知道你是谁,你在哪儿,房子将用这一信息尽量满足甚至预见你的需求……当你沿大厅的路走时,你可能不会注意到前面的光渐渐变强,身后的光正在消失……

早期,这样的网络被称为传感网或感知网,中科院于1999年就启动了“感知中国”项目,开展针对传感网的研究。

2005年,国际电信联盟ITU正式提出“物联网”的概念。物联网是互联网发展的延伸,互联网是物联网的基础。物联网的发展又将极大地促进互联网的发展。建设物联网的目的是实现人、物理世界与信息世界的完全融合。物联网的特性如下。

(1)全面感知。利用RFID、传感器及其他各种感知设备采集对象信息,人们可通过物联网随时随地(anywhere, anytime)地掌握所关心物品(anything)的精确位置及其周围环境。

(2)可靠的传送。利用以太网、无线网、移动网将感知的信息进行实时、可靠的传送。

(3)智能化控制。对物体实现智能化的控制和管理,达到了人与物、物与物的沟通或互动。

例如,上海浦东国际机场和上海世博会安装了无锡传感网中心的安全防护设备。该设备由10万个微小传感器组成,散布在墙头、墙角及路面上。传感器能根据声音、图像、震动等信息分析判断,爬上墙的究竟是人还是猫狗,防止人员的翻越、恐怖袭击等。

通过在汽车和汽车点火钥匙上植入感应器,物联网还可应用在交通安全中。当喝了酒的司机掏出汽车钥匙时,钥匙能通过气味感应器察觉到一股酒气,立即用无线信号通知汽车“不要发动”,汽车会意地“罢工”,并“命令”司机的手机给其亲友发短信,通知他们司机所在的位置,请亲友们来处理,杜绝了“酒后驾车”。

物联网也是一种层次结构。早期的物联网分为感知层、网络层和应用层。最新的观点认为物联网分为四层,从下往上依次为感知层、网络层、管理服务层、应用层。

感知层是实现物联网全面感知的基础。由RFID标签(又叫电子标签)和读写器、摄像头、GPS、传感器和传感器网络等组成,感知和识别物体,采集和捕获物体的标识、运动和位置等信息。

网络层由各种有线或无线的通信网络与计算机互联网有机地结合而成,实现信息的传输和交换,网络的接入与会聚。作为物联网的基础设施,网络层的运行必须是可靠的。

管理服务层通过中间件平台,完成数据存储与处理、数据挖掘与智能决策,屏蔽掉感知层和网络层的工作细节,为应用程序员提供一个统一而简洁的信息服务接口。

应用层是将物联网技术与行业技术相结合,实现广泛的智能化应用。例如,灾害预报与防治,精细化农业,智能交通,智能物流,智能医疗,智慧城市等。

在智能交通中应用物联网的一个例子是“路桥不停车收费(ETC)系统”:安装电子标签和IC卡的车辆在通过路桥时,系统会自动识别车型、车号,计算通行费,并直接从IC卡上扣除费用。交易完成后,电动栏杆升起,放行车辆。

提出物联网的概念,无疑是一个创新。那么,创新来源何处呢?答案很简单——**提出自己的需求,发现别人的需求,全心全意为人民服务!**美国苹果公司的成功也有力地证明了这个道理。

当前,跨越商业和 IT 行业的服务计算,如软件即服务 SaaS (Software as a Service)、基础设施即服务 IaaS (Infrastructure as a Service)、平台即服务 PaaS (Platform as a Service) 等,已成为 IT 行业的热门研发领域。请读者查阅相关资料,了解各种 XaaS (X as a Service) 的概念与技术。

总之,科学技术以人为本。计算机替代人们完成相当一部分重复性的脑力劳动,使人们有更多的时间进行创造性的工作。随着计算机性能的不不断提高、应用的日益普及,新的学科分支或应用领域还将不断地涌现出来。

1.9 计算机的发展

1.9.1 计算机发展的动力

电子计算机的诞生,不仅仅是计算工具的一次革命,更重要的是它作为科学技术发展史上的一个里程碑,引发了一场新的技术革命——信息革命。信息革命把人类社会生产力提高到一个新的水平,使信息资源成为当今世纪最重要的资源,成为世界各国争夺的第一目标。目前,世界各国都投入一大批优秀人才,把最先进的技术应用于计算机科学与技术领域的研究。在竞争中,计算机科学与技术得到了飞速的发展。

计算机发展的动力主要来源于应用的需求与技术的推动两个方面。

1. 应用的需求

应用的需求可分为求解疑难问题和改善人机接口两个方面。

20 世纪 90 年代,美国提出了科学与工程领域的 7 个重大挑战问题 (Grand Challenges)。

(1) 量子化学、统计力学、相对论物理。

(2) 天体物理学。

(3) 计算流体动力学和湍流。

(4) 材料设计与超导性。

(5) 生物学、药理学、基因测序、遗传工程、蛋白质折叠、酶活力、细胞建模。

(6) 医药、人体器官与骨骼的建模。

(7) 全球天气和环境建模。

这些重大挑战问题像“海绵吸水”一样吸干了目前任何高性能微处理器和超级计算机所能提供的计算能力。

和谐的人机接口一直是计算机科学工作者努力的目标。手写输入、语言输入、图形/图像的输入/输出等就是这方面的研究成果。

2. 技术的推动

技术的推动可分为软件的推动和硬件的推动两个方面。

软件技术的推动则表现在以下几个方面。

(1) 新的应用问题和应用领域的出现。

(2) 软件工程的发展,降低了产品价格,缩短了软件开发周期,保护用户的投资。

(3) 不断地消除软件系统中的“瓶颈”,使软件中的各个模块平衡工作。

(4) 市场需求的刺激,产品的不断细化。

(5) 其他学科研究成果的应用。

(6) 人们对自身“智能”的探索。

硬件技术的推动来源于器件的进步和体系结构与组成技术的进步。其中，器件的进步一直是推动计算机科学与技术不断发展、计算机性能不断提高的最直接的因素。

从第一台电子计算机诞生至今只有 60 多年的时间，但是电子计算机却经历了真空管、晶体管分立元件、小规模集成电路、大规模集成电路、超大规模集成电路五代。元器件的进步还使得计算机系统的设计与应用环境发生了深刻的变化，元器件的小型化使计算机的存储容量大幅增加、计算速度急剧提高。目前，超级计算机的运算速度已超过 1 亿亿次/秒，主存储器容量超过 1 万亿字节。

但是研究表明，半导体器件技术的发展已接近其物理极限。例如，Pentium 4 的主频已达到 2GHz，集成电路线宽已是 $0.18\mu\text{m}$ ，虽然预计未来还可以生产出主频达到 10GHz、集成电路线宽低于 $0.1\mu\text{m}$ 的微处理器，但是再进一步发展，恐怕就要另辟蹊径了。

不过，计算机的发展前途还是一片光明的。因为计算机科学工作者最善于学习和利用其他学科/领域的新成果、新技术。这也是电子计算机 60 多年来取得惊人发展的原因。

一个很有前途的方向就是发展光学计算机。光学计算机用不同波长的光代表不同的数据，用光束代替电子进行计算和存储，以大量的透镜、棱镜和反射镜来实现数据的传输。光学计算机具有电子计算机无法比拟的许多优势。首先在速度方面，电子的传播速度只能达到 593km/s ，而光子的速度却是 $30\times 10^4\text{km/s}$ 。其次在超并行性、抗干扰性和容错性方面，光学计算机都优于电子计算机。另外，光还具有光路间可交叉，也可与电子信号交叉而不互相干扰的特点。

1990 年，贝尔实验室宣布成功研制世界上第一台光学计算机，尚待解决的困难是没有与其匹配的存储器件。目前，光学计算机有全光型和光电混合型。

也有一部分学者认为：由于硅半导体在工艺上已经成熟，是最经济的器件，还有潜力可挖，在今后相当长的时间内仍将是计算机的主流器件，所以把提高电子计算机性能的希望寄托于超导现象的研究。

此外，以 DNA 计算机为代表的生物化学计算机的研究也取得了突破性的进展。研究表明：DNA 上含有大量的遗传密码，可通过生物化学反应进行遗传信息的传递，这是生命现象的基本特征之一。DNA 计算机的原理如下。

(1) 利用 DNA 存储遗传密码的原理存储数据。有关科学家认为：DNA 的存储容量非常巨大， 1m^3 的 DNA 溶液可以存储 1×10^{20} 位的信息，相当于目前世界上所有计算机存储容量的总和。

(2) DNA 单元之间可以在某种酶的作用下瞬间完成生物化学反应，而且数以万亿计的 DNA 单元能够同时/并行地进行生化操作，以此就可以实现人们需要的计算。乐观的科学家认为：DNA 计算机的计算速度极快，它在几天内完成的运算量就相当于目前世界上有史以来所有计算机完成的运算量的总和。

(3) 反应前的基因代码可以作为计算的输入，反应后的基因代码可以作为计算的输出。生物化学计算机是人类一直梦寐以求的愿望，它可以实现现有计算机无法真正实现的模糊推理和神经网络计算，是智能计算机乃至“人造大脑”最有希望的突破口。

1994 年 11 月，在美国《Science（科学）》杂志上刊登了南加利福尼亚大学伦纳德·阿德拉曼（L. Adleman）博士利用 DNA 溶液成功求解出多个城市间最短回路问题的论文，这是第

一个 DNA 计算机的报道。2001 年，以色列魏茨曼 (Weizmann) 研究所成功研制了一台全自动运行的 DNA 计算机。该机可以在生物体内存储和处理信息，具有普通计算机的大部分功能，标志着 DNA 计算机的研制向实用化迈出了一大步。2002 年，日本开发出全球第一台能够投入商业应用的 DNA 计算机，该机可进行基因诊断。

量子计算机是目前研究的另一个热点，它以处于量子状态的原子作为处理与存储的对象，利用原子的量子特性来实现计算和信息传输。

1.9.2 计算机的发展趋势

现代科学技术的水平和发展速度告诉我们，无论是在元器件方面，还是在体系结构方面，电子计算机的发展前景都是十分乐观的。随着信息时代的到来，计算机的地位日益突出、作用越来越大、发展日新月异。就目前而言，计算机科学与技术的研发重点主要集中在“高”、“开”、“多”、“智”、“网”和“虚”六个方面。

(1) 高：高速度。提高计算机的速度是计算机科学与技术研究与开展永恒的主题。10 年前，人们追求的“3T”计算机 (1TFlops 的运算速度，1TB 的主存储器容量，1TB/s 的 I/O 带宽) 已基本实现，现在的目标是“3P”计算机 (1P=1024T)。

我国的“天河二号”就属于一台“3P”计算机，它包含 16000 个运算节点，每节点配备两颗时钟频率为 2.2GHz 的 Xeon E5 12 核心的 CPU (英特尔 Ivy Bridge 微架构)、三个时钟频率为 1.1GHz 的 Xeon Phi 57 核心的协处理器 (基于英特尔众核架构)。累计 32000 颗 Xeon E5 主处理器和 48000 个 Xeon Phi 协处理器，共 312 万个计算核心。峰值计算速度达到 5.49PFlops (PetaFlops, Flops 是英文 Floating-point Operations per Second 的缩写，Peta 指拍，1PetaFlops 等于千万亿 (拍，即 10^{15}) 次/秒的浮点运算)、持续计算速度为 3.39 PFlops，在 2013 年 11 月位列世界第一。

“天河二号”的每个节点拥有 88GB 内存，内存总容量为 1.408PB。外存为 12.4PB 的硬盘阵列 (相当于存储每册 10 万字的图书 600 亿册)。整个系统由 170 个机柜组成，包括 125 个计算机柜、8 个服务机柜、13 个通信机柜和 24 个存储机柜，占地面积为 720m²。

(2) 开：开放 (Open)，指计算机系统具有开放性，不同厂商、不同时期的产品能够无障碍地相互操作、相互通信。这使得计算机系统的灵活性和通用性得到极大的增强。

(3) 多：多媒体 (Multimedia)，指未来的计算机将能够通过声、图、文等多种媒介来与人类交互。现在，人们已经可以通过语音、手势、图像来控制计算机；未来，人们还可以通过表情、语气、眼神来操作计算机。

(4) 智：指人工智能，让计算机具有智能是计算机科学工作者不懈的目标。经过多方面的努力，计算机在知识表示、机器学习、定理证明、专家系统、博弈等多方面已经取得了长足的进展。首届中国最高科学技术奖的两位获奖者之一吴文俊院士的主要研究成果就是几何定理机器证明，医学专家系统已有效地解除了患者的病痛。

目前，微软亚洲研究院的周明研究员研究的“计算机对对联”，厦门大学的周昌乐教授研究的“计算机作诗”都是人工智能领域的新的、有趣的研究方向。

(5) 网：指计算机网络。自进入 20 世纪 90 年代以来，计算机网络 (特别是因特网) 得到了迅猛的发展。“计算机就是网络，网络就是计算机”的观点已被人们接受。

目前，计算机网络的基本工作模式是客户机 (Client) /服务器 (Server) 模式 (简称 C/S

模式)。在 C/S 模式下,网络上的计算机根据所担任的角色不同,被分为客户机或服务器。客户机的任务是提出请求,接收结果,它本身不做太多的计算;服务器的任务是接受请求,进行处理,并返回结果,即提供服务。

目前,大多数客户机都是通用计算机,这些计算机的特点是:配有高性能的微处理器,而网络功能却不是很强。这就导致高性能微处理器的计算功能没有得到充分的利用,而计算机的网络功能又满足不了用户的需求,显然很不经济。为此,在 Java 语言(一种与运行平台无关的程序设计语言)出现后,SUN 公司就率先提出了“网络计算机 NC(Network Computer)”的概念,NC 的特点是具有很强的网络功能,采用较简单的微处理器,配置较简单的操作系统,存储器系统的容量也没必要很大。由于 NC 的定位是客户机,所以它的工作模式主要是下载 Java 程序来执行。为此,NC 必须安装 Java 语言解释执行程序——JVM(Java Virtual Machine)。

由于 NC 机具有结构简单、价格低廉、功能时尚等特点,一经提出就受到广泛的欢迎。目前已有很多厂商推出了多款 NC 产品。例如,北大众志公司的基于“北大众志-863”CPU 系统芯片的网络计算机 NC-I30/I40、NC-A15/A25。

(6) 虚:虚拟化(Virtualization),指以用户/应用程序很容易获得的方式来表示计算机资源的过程,而不是根据这些资源的实现、地理位置或物理包装的专用方式来表示计算机资源,即提供计算能力、存储资源等的逻辑视图而不是物理视图。

虚拟化是计算机层次结构概念的延伸与发展,代表着计算机系统的发展方向——通过共享有限的资源,为用户提供更好的服务,而用户将在感觉不到具体计算部件、存储部件甚至计算机程序存在的情况下,享受计算机的服务。2008 年,谷歌公司(Google)和 IBM 公司提倡的“云计算”就是这方面的例子。

网络技术和虚拟化技术的发展,推动着微型计算机 PC 向着融合了 3C(Computer、Communication、Consumer)技术的掌上计算机和可穿戴计算机发展,进入“风(Phone)起云(云计算)涌”的“移动互联网(Mobile-Internet,MI)”时代——用户借助智能手机就能够完成原先需要使用 PC 才能完成的事情。例如,苹果公司推出的智能手机 iPhone 就结合了照相机、个人数码助理 PDA、媒体播放器 MP3/MP4 以及无线通信和无线上网的功能。而平板电脑 iPad 的功能定位,介于 iPhone 和笔记本电脑之间,除手机的功能外,还提供浏览互联网、收发电子邮件、观看电子书、播放音频或视频等功能。而在 iPhone 之后,联想公司推出了“乐 Phone(LePhone)”智能手机。主流的智能机操作系统有苹果公司的 iOS、谷歌公司的安卓(Android)、微软公司的 Windows Phone。最新市场调研数据显示,2014 年第一季度中国移动互联网市场规模达到 330.7 亿元,同比增长 78.5%,主要包括移动购物、移动增值、移动游戏和移动营销四个方面。

习题 1

1-1 第一台电子计算机_____于 1946 年 2 月在美国诞生,它是由_____和_____领导的科研小组研制成功的。_____被誉为“计算机之父”,_____被誉为“现代计算机之父”。

1-2 为了纪念帕斯卡在发展计算机方面的杰出贡献,1971 年,瑞士计算机科学家沃思(Niklaus Wirth)将其发明的一个程序设计语言命名为 Pascal。Pascal 语言在 20 世纪七八十年代

被广泛应用于大学的计算机教育。Pascal 语言的特点是什么？

1-3 第一个高级程序设计语言是 _____，_____ 语言的诞生标志着对计算机程序设计语言的研究正式成为一门专门的学科。你喜欢的程序设计语言是 _____。

1-4 为了编制控制“分析机”运行的穿孔卡片上的指令，巴贝奇雇用了英国著名诗人拜伦（Byron）爵士的女儿爱达（Ada Augusta Lovelace）来完成这一任务。从某种意义上说，爱达是世界上首位程序员。20 世纪 80 年代，美国国防部将其研发的一种通用程序设计语言命名为 Ada，以纪念这位计算机事业的先驱。请查阅相关资料，说明 Ada 语言的提出背景及其特点。

1-5 现代计算机的理论模型是由_____提出的，这个模型被称为_____。

1-6 冯·诺依曼计算机在体系结构上的特点是什么？

1-7 冯·诺依曼计算机采用二进制来表示指令与数据，并采用二进制来运算。二进制的起源可追溯到中国古代的八卦图。莱布尼茨曾对它进行过深入的研究，并将研究成果应用到了他研制的计算机上。1842 年，在巴贝奇研制“分析机”的过程中，他的助手爱达女士也曾建议用二进制存储代替十进制存储。那么，迄今为止，计算机中的所有信息仍是以二进制方式表示的原因是_____。

A．运算规则简单

B．物理器件的限制

C．运算速度快

D．软件兼容

1-8 解释下列概念：计算机系统、硬件、软件、存储程序、主机、虚拟机、主存储器、辅助存储器、透明性、吞吐率、响应时间、软件兼容性、可伸缩性、C/S 模式、计算机体系结构、计算机组成、计算机实现。

1-9 请在计算机硬件系统框图（如图 1-6 所示）中，根据数据流和控制流的走向，填写各个部件的名称。

1-10 尽管冯·诺依曼在 1946 年就完成了 EDVAC 的设计，但可惜的是直到 1952 年才完成 EDVAC。而此时，英国剑桥大学的 Maurice Vincent Wilkes 在 1946 年 5 月，阅读了冯·诺依曼关于 EDVAC 的设计方案，相信“存储程序”是计算机发展的正确方向。之后，Wilkes 详细了解了 ENIAC 并参加了 EDVAC 设计的讨论；1946 年的 10 月，Wilkes 开始 EDSAC 的设计与实现工作，并于 1949 年 5 月成功完成了 EDSAC（Electronic Delay Storage Automatic Calculator）的建造。所以 EDSAC 是事实上的第一台冯·诺依曼结构的电子计算机。为此，Wilkes 于 1967 年荣获图灵奖。你对此有何感想？

1-11 查阅相关资料，说明最近三届图灵奖获得者的名字及其成果。

1-12 事实上，在第二次世界大战期间，英国政府为了破译德军的密码，委托图灵设计制造了一台名为 COLOSSUS 的电子计算机。该机于 1943 年投入运行。但由于英国政府将其列为最高机密，所以该机没有对外发布。在第二次世界大战结束后，也没有对该机进行进一步的发展。英国人认为 COLOSSUS 才是世界上第一台电子计算机。英国失去计算机技术领先地位的原因是什么？

1-13 按照采用的逻辑器件分，电子计算机的发展经历了哪几代？每一代的特征是什么？

1-14 “超级计算机之父”西蒙·克雷一生致力于研制世界上最快的计算机，由他主持研

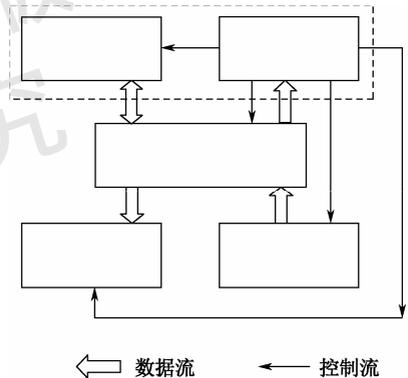


图 1-6 习题 1-9 的框图

制完成的 CDC6600、CDC7600、Cray-1、Cray-2、Cray-3 一直是当时世界上最快的计算机。现代的多核处理器、精简指令系统计算机 RISC (Reduced Instruction Set Computer) 都可以从 CDC6600 中找到启发。请查阅相关资料，了解西蒙·克雷的生平事迹。

1-15 计算机软件分为：_____、_____、_____和_____。

1-16 将功能固定的软件存储在 ROM 中，称之为_____。

1-17 简述控制器 CU 的功能和基本组成。

1-18 早期微处理器中仅表示程序运算状态的标志位寄存器已经无法满足面向多用户、多任务的操作系统的要求，所以在现代位处理器中，标志位寄存器已经发展成为“程序状态字 PSW (Program Status Word)”寄存器。请查阅文献资料，了解奔腾处理器 PSW 寄存器的内容。

1-19 画图说明计算机系统硬件的基本组成以及计算机系统的层次结构。

1-20 目前，物联网的层次结构有两种分法。第一种分为四层，从下往上依次为感知层、网络层、管理服务层、应用层。第二种分为三层：感知层、网络层、应用层。请评价这两种划分方法。如果你要设计、开发一个物联网系统，你倾向于选择哪一种？

1-21 程序计数器 PC 是冯·诺依曼计算机在组成上最明显的特征之一。请说明 PC 的功能。多年来，计算机科学家一直致力于研制不带程序计数器的计算机——非冯·诺依曼计算机的研究，相继研制出数据流计算机、面向函数程序设计的归约机、智能推理机、模拟人脑神经元结构的连接机等。请查阅文献，了解非冯·诺依曼计算机的研究进展。

1-22 什么叫程序的可移植性？实现程序的可移植性有哪些技术途径？

1-23 什么叫系列机？列举出你所知道的系列机。

1-24 系列机的设计目的是_____。第一个系列机是_____。

1-25 除了系列机外，实现程序可移植的另一个技术途径是统一高级语言。统一高级语言是采用高级语言虚拟机来屏蔽掉物理机器的属性，Ada 语言和 Java 语言的提出都带有类似的目的。用统一高级语言实现程序可移植可行吗？为什么？

1-26 简述微处理器和微型计算机发展的五个阶段及各个阶段的特征。

1-27 查阅相关资料，说明英特尔公司最新一代微处理器的研发进展。

1-28 在个人计算机领域，另外一家著名的计算机公司是苹果公司 (Apple Inc.)，它的注册商标是“苹果”，所以它生产的计算机被称为“苹果计算机 (Apple Computer)”，俗称为“苹果电脑”。苹果公司是在 1976 年 4 月 1 日由斯蒂夫·乔布斯 (Steve Jobs) 和斯蒂夫·沃兹尼亚克 (Steve Wozniak) 创立的。在当年开发并销售个人电脑 Apple I。1977 年发售最早的个人电脑 Apple II。1984 年推出革命性的 Macintosh 电脑。2001 年推出 iPod 数字音乐随身听。2003 年推出最早的 64 位个人电脑 Apple PowerMac G5。至今，苹果电脑仍然在个人计算机市场占据重要的位置。请查阅相关文献，评价苹果电脑的优点。

1-29 简述摩尔定律的内容。摩尔定律是商业规律还是科学规律？

1-30 _____是中国电子计算机事业的奠基人。目前，国内研发计算机系统的单位有_____、_____、_____等，有影响的人物有_____、_____、_____等。

1-31 超级计算机指的是一个特定的时间点上运算速度最快的计算机。请登录“国际最快的 500 台计算机网站 www.top500.org”和“国内最快的 100 台计算机网站 www.samss.org.cn”，了解目前国外和国内最快的电子计算机的名称及其处理速度。

1-32 近 60 年来，计算机迅速发展的主要原因是什么？

1-33 随着计算机性能的不断提高,计算模拟已成为继理论分析、实验验证之后的第三者科学研究手段,例如在航空、航天和船舶领域出现了“数字风洞”和“数字水洞”。查阅相关资料,说明用计算机模拟自然现象的例子。

1-34 对计算机进行分类是计算机科学领域的基础工作,目前已经提出了很多种分类的方法。例如,1972年,美籍华人冯泽云(Tse-yun Feng)提出用最大并行度(Degree of Parallelism),即在单位时间内计算机能够处理的最大二进制位数,来对计算机进行分类。1977年,Handler提出根据计算机指令执行的并行度和流水线来对计算机进行分类。1978年,Kuck提出用指令流、执行流及其多倍性对计算机进行分类。请读者查阅文献了解这些计算机分类法,不妨也尝试提出自己的分类法。

1-35 美国政府商务部还使用过“数据处理速率 PDR (Processing Data Rate)”来评价处理器的性能,并以此作为控制高性能计算机产品出口的标准。数据处理速率 PDR 等于每条指令传送的平均字节数乘以处理速率,单位是百万字节/秒。

每条指令传送的平均字节数 = 定点指令的字节数 + $0.4 \times$ 定点操作数长度 + $0.15 \times$ 浮点操作数长度。

处理速率 = $1/[0.85 \times$ 定点加法指令的执行时间(μs) + $0.09 \times$ 浮点加法指令的执行时间(μs) + $0.06 \times$ 浮点乘法指令的执行时间(μs)]。

美国政府商务部在 1980 年 10 月,规定出口到中国的计算机系统的数据处理速率 PDR 不得超过 32,当时 DEC 公司的 VAX 11/782 计算机系统的 PDR 为 101。到 1981 年 12 月,这一标准被改成了不得超过 64。你知道为什么吗?请你谈谈对美国政府限制对华出口高技术的看法。

1-36 IBM 公司还曾提出用每秒所完成操作的数目来表示处理器的性能,其单位是 KOPS (Kilo-Operations Per Second)。请你也提出一个能够表示处理器性能的技术指标。

1-37 在计算机性能指标中,_____决定了计算机中数据表示的范围与精度。系统管理员最关心的性能指标是_____,用户最关心的性能指标是_____。

1-38 计算机性能的统计方法:_____、_____和_____。

1-39 程序在速度为 100MIPS 的计算机上运行所花的时间一定长于在速度为 500MIPS 的计算机上运行所花的时间。对吗?

1-40 什么叫软件兼容?对系列机而言,软件兼容的基本要求是什么?

1-41 软件和硬件在_____上是等价的,在性能上是不等效的。这称为软/硬件的等价性原理。在计算机系统的层次结构中,位于硬件以上的所有层次统称为_____。

1-42 一家著名的美国计算机公司及其出品的微处理器也以 MIPS 命名,不过它的 MIPS 是 Microprocessor without Interlocked Pipeline Staged 的缩写。该公司推出的 MIPS R4000 是最早的 64 位微处理器之一。请查阅相关资料,了解该公司的最新产品。

1-43 在研制 ENIAC 的同时,爱肯在哈佛大学也开始了新型电子计算装置 Mark- 的研究,并相继开发了 Mark-、Mark- 和 Mark-。在 Mark- 和 Mark- 中,爱肯将存储系统设计成两个分离的指令存储器和数据存储器。当时,人们就把这种分离的存储器结构称为哈佛结构(Harvard Architecture)。相应地,称冯·诺依曼设计的单一的存储器结构为普林斯顿结构(Princeton Architecture)。请比较两种存储器结构的优缺点。

1-44 开发并行性的方法主要有:时间重叠、资源重复和资源共享。时间重叠是将相邻处

理过程在时间上错开,轮流重叠地使用同一套硬件的各部分,如生产流水线,这是投入少,收益较高的方法;资源重复是通过重复设置硬件资源来提高性能或可靠性,如阵列机、多核处理器和容错计算机,这是高投入、高收益的方法;资源共享是利用软件的方法让多个用户按照一定的顺序轮流使用同一套资源,这是着眼于提高资源利用率的方法。分布式系统(Distributed System)就是基于资源共享而提出的。请查阅文献,了解容错计算机和分布式系统的基本原理。

1-45 多核处理器是一个重大的技术创新。请分析这个创新的产生背景、要解决什么问题、是如何解决这些问题的。谈谈你对这个创新的想法。

1-46 机器学习(Machine Learning)是研究计算机怎样模拟或实现人类的学习行为,以获取新的知识或技能,重新组织已有的知识结构使之不断改善自身的性能。它是人工智能的核心,是使计算机具有智能的根本途径。已故的哈尔滨工业大学计算机系的洪家荣教授当年在这方面的研究取得了世界瞩目的成果。斯坦福大学公开课中就有一门课程是“机器学习”。有兴趣了解的读者可以学习这门课程,了解机器学习最新的研究成果。

1-47 1950年,计算机逻辑和AI研究的奠基人图灵在论文《电脑能思考吗?》中提出了著名的“图灵测试”——在不知情的条件下,通过某种特殊的方式,一个人与一台机器相互问答,假若在相当长的时间内这个人分辨不出与之交流的对象是人还是机器,那么这台机器便可认为是能思考的。谈谈你对这个问题的看法。

1-48 利用工具来增强脑力和体力是人类具有智慧的象征,也是人类不懈追求的目标。计算机/软件的科学工作者的重要使命是研究和开发计算工具。这既包括硬的计算工具(如计算机),也包括软的计算工具(如程序设计语言和操作系统)。具有里程碑意义的程序设计语言的设计者几乎都获得了图灵奖,例如ALGOL 60语言的设计者Perlis(1966年首届图灵奖)和Naur(2005年图灵奖),第一个ALGOL 60编译器的实现者Dijkstra(1972年图灵奖),Fortran语言的设计者Backus(1977年图灵奖),APL语言的设计者Iversior(1979年图灵奖),OCCAM语言的设计参与者Hoare(1980年图灵奖),C语言的设计者Ritchie(1983年图灵奖),Pascal语言的设计者Wirth(1984年图灵奖),ML语言的设计者Milner(1991年图灵奖),SIMULA 67语言的设计者Dahl和Nygaard(2001年图灵奖),Smalltalk语言的设计者Kay(2003年图灵奖),CLU语言和Argus语言语言的设计者Liskov(女,2008年图灵奖)等。中国的计算机/软件科学工作者在这方面也卓有建树。其中较著名的有南京大学徐家福教授、中国科学院计算技术研究所仲萃豪研究员、北京大学杨芙清教授等于1978年设计的XCY语言(系统程序设计语言)和中国科学院软件研究所唐稚松院士于1979年起设计的一个广谱的结构程序语言XYZ语言(Xiliehua Yuyan Zu,系列化语言族)。XCY语言规模适度,简明易用,主要用于书写系统程序。DJS-240机的操作系统完全用XCY语言书写。XCY语言本身的编译程序也是用XCY语言书写的。XYZ语言将时序逻辑理论与软件工程技术有机地结合起来,提高了软件开发的自动化水平和生产率。其中,时序逻辑语言XYZ/E是世界上第一个可执行的时序逻辑语言。请查阅资料,了解上述程序设计语言的特性,并提出你对程序设计语言的新要求。

1-49 随着环境保护的呼声日益高涨,“绿色计算机”的概念被提了出来,所谓绿色计算机是指符合环保概念的计算机主机和相关产品(含显示器、打印机等外设),具有省电、低噪声、低污染、低辐射、材料可回收及符合人体工程学特性等特点。衡量一台计算机是否绿色的一个指标就是效能比——每消耗1W电力能提供多少百万个浮点操作(MFlops/W),该值越大

越好。目前,即便是关注性能的超级计算机也重视效能比,我国“天河一号”的能效比是 431.7 MFlops/W,“天河二号”的能效比是 1.935GFlops/W。请登录 <http://www.green500.org/>, 查找世界上效能比最高的计算机。并试着提出一个衡量计算机是否绿色的新指标。

1-50 有学者提出,在信息技术高度发展的今天,人类进入了“大数据 (Big Data)”时代。所谓“大数据”是指蕴藏着特定知识或信息,而又无法依靠现有计算平台在合理时间内获取、管理、处理、分析的巨量资讯数据。大数据具有 4V 特点: Volume (量大)、Velocity (激增)、Variety (纷杂)、Veracity (客观)。请查阅相关资料,了解目前“大数据”领域的研究与应用状况,并谈谈你对“大数据”的看法。

1-51 在 2008 年开始的国际金融危机的背景下,以制造和硬件为重点的 IT 公司都遭遇重创,但是 IBM 公司 2008 年的全年收入为 1036 亿美元,较去年增长 5%,利润增长超过了 18%。IBM 公司 2008 年财报显示硬件部门的收入下降,而最大的利润增长贡献部门来自软件和服务部门。IBM 之所以有效避开国际金融危机的冲击,而且保持利润的高增长,完全得益于几年前就开始酝酿的战略调整。

进入 2008 年,IBM 公司开始把自己的业务范围从“商业机器”延伸到了“医疗、食品、能源、淡水”等领域。2008 年年底,IBM 公司宣布:2009 年 IBM 将在全球推出全新战略理念“智慧的地球 (A Smarter Planet)”。

“智慧的地球”的基本要素是:更透彻的感应和度量,更全面的互联互通,更深入的智能洞察。IBM 公司称世界不仅在变得更小、扁平,而且将变得更“智慧”。

如果把 IBM 公司从专注“硬件”变身为“软件+硬件”视为第一次产业转型,从“软件+硬件”变成“软件+硬件+服务”视为第二次产业转型,那么这一次将是 IBM 的第三次产业转型。IBM 中国研究院院长李实恭说:“IBM 将关注人从出生到结束的一切活动,包含此过程中的教育、娱乐、生活等的民生需要,IBM 变得更有‘生命感’了。”

进入 21 世纪以来,人类已经意识到人类真正需要的是一个环境友好的、以人为本的、和谐的地球。在这样的理念下,IT 可以做什么?IT 人应该怎么做?

1-52 我国计算机学者徐志伟研究员于 2008 年年底提出了“为人民计算 (Computing for the Masses)”的概念。所谓“为人民计算”是指以全民普及为目标的计算机技术的研究与应用。事实上,普及计算能力、为大众服务一直是计算机科学技术和产业发展的目标。个人电脑的发明是“为人民计算”发展历程中的一个里程碑。请举例说明代表“为人民计算”理念的计算机技术。你还希望“为人民计算”为你做什么?

1-53 我国计算机学者张尧学院士于 2004 年年底提出了“透明计算”的概念,即通过网络,把存储、运算、管理进行逻辑或物理分离,将应用和系统软件以及硬件分开,实现个人在任意地点的“不知不觉、用户可控”的统一体验,有简化用户使用、降低成本等好处。实现“透明计算”后,用户可以在不感知“操作系统”存在的情况(操作系统透明化)下,自由选择和使用各种软件(服务)。2012 年 10 月,中南大学成立了“中南大学——英特尔透明计算实验室”。透明计算的核心思想是什么?请查阅相关文献,了解“透明计算”的研究进展。

1-54 拟态章鱼是自然界最为奇妙的“伪装大师”,它能扭曲身体和触手,改变颜色,模仿至少 15 种动物的外表和行为。受拟态章鱼的启发,中国工程院院士邬江兴带领科研团队融合仿生学、认知科学和现代信息技术,提出拟态计算新理论,并于 2013 年 9 月成功研制出世界首台结构动态可变的拟态计算机。通常的计算机“结构固定不变、靠软件编程计算”,而拟

态计算机的结构动态可变，“靠变结构、软硬件结合计算”。针对用户不同的应用需求，拟态计算机可通过改变自身结构提高效能。测试表明，拟态计算机典型应用的能效，比一般计算机可提升十几倍甚至上百倍。请查找资料，撰写高效能计算机研究进展综述，并设想未来计算机的发展趋势和新的应用领域。

1-55 据中国互联网络信息中心 CNNIC 统计，截至 2011 年 12 月底，中国网民达到 5.13 亿人，手机网民达到 3.56 亿人，占全国网民总数的 69.4%，手机网民人数已超过电脑网民人数。手机的功能也日益多样化，给中国的企业和年轻人参与世界 IT 技术的竞争提供了机会。例如，2013 年 6 月，在第 13 届挑战杯大学生课外学术科技作品竞赛（黑龙江赛区）上，东北农业大学的学生展示了一款用手机检测西瓜成熟度的作品。用户用手机接收手拍西瓜发出的声音，手机内的软件分析出声音的波长和频率，以此判断西瓜的成熟度。你对未来的手机功能有什么期待？

1-56 “早上好，你是吉林大学今天第 361 个起床的人！你击败了全校其他 67832 人。”这是吉林大学学生王祯晗早上收到的早起签到回复。吉大学生张天译、徐昊天、张宗达为实现“中国梦”而独立研发的励志正能量传播平台——“还睡呀”微信平台于 2013 年 4 月 15 日正式上线。只要发送数字“1”到“还睡呀”微信平台，就能知道自己是全校第几位起床的人。为了鼓励大家早起，排在起床排行榜前边的同学还可以领到免费豆浆作为奖励。

“很多大学生都有睡懒觉的习惯，美好的光阴往往就这么蹉跎掉了，我们想通过这样一种喜闻乐见的方式——起床签到排行榜，把赖床的同学从床上‘薅’起来。”张天译说，“中国梦不是喊口号，需要我们踏踏实实地从每一天、每一分钟做起。如果每个同学都能早起一分钟，实现‘中国梦’就近了很多步。”

“还睡呀”微信平台不仅有早起签到功能，还有一分钟讲堂和实用信息等基本功能。一分钟讲堂主要包含由学校知名教授讲述的知识、优秀学子的经验及小故事等。实用信息则包含了每日校内通知及订餐电话等生活信息。

“让更多的同学充满正能量地生活是我们共同的追求”，张天译和他的研发团队表示，接下来他们会根据同学们的需求，不断地丰富该微信平台的功

你了解手机 APP 吗？你开发过 APP 吗？若开发过，请介绍它的功能与架构。

1-57 “钥匙！”只需对着储物柜喊一声，柜子就会乖乖地把钥匙送到你面前。这种好像神话般的事情，已经在郑州大学的大学生手中变成了现实。智能储物柜的发明者之一、郑州大学学生潘世豪说，智能储物柜借鉴物联网理论，能把物品分类分层放好。当你需要柜子里的任何物品时，只要对着柜子上的语音识别系统喊一声，柜子就会自动把它送到你面前。你对未来物联网技术改善我们的生活有什么期待？