永不停顿和可伸缩的
64位RISC计算机技术概论

Non-Stop and Scalable
64-bit RISC Computer Technologies

王德安　张报昌　著

原子能出版社

图书在版编目(CIP)数据

永不停顿和可伸缩的64位RISC计算机技术概论/王德安,张报昌著.—北京:原子能出版社,2000.8

ISBN 7-5022-2203-0

Ⅰ.永… Ⅱ.①王… ②张… Ⅲ.计算机,RISC-系统结构 Ⅳ.TP338.6

中国版本书馆CIP数据核字(2000)第3801号

内容提要

本书针对Internet电子商务和技术计算应用对高可用性和可伸缩性的并行计算机系统的迫切需要,介绍具有永不停顿可用性和高可伸缩性的计算机系统。由于64位RISC处理器目前已经成为装备高端计算机系统的主流技术,本书将主要介绍基于64位RISC处理器的计算机系统。

本书包括7章,分为三部分:第一部分(第1章)以Internet和电子商务时代对计算机系统的全面需求为背景,介绍计算机技术的发展趋势和设计要求以及评估现代计算机系统的各种指标;第二部分(第2-4章)介绍设计和建造具有高可用性和可伸缩性的并行计算机系统的主要技术,包括处理器技术、内存系统技术和系统互联技术;第三部分(第5—7章)介绍UMA、NUMA、MPP和集群等体系结构的并行系统新发展以及各种体系结构系统的代表性机型。

本书取材新颖,力图反映成稿(2000年6月)前现代计算机技术的最新发展。本书可以作为了解和研究64位RISC计算机系统技术的参考书,也可用作进行计算机设备选型及评估工作的参考资料。

原子能出版社出版 发行

责任编辑:孔玥

社址:北京市海淀区阜成路43号邮政编码:100037

北京市朝阳区科普印刷厂印刷 新华书店经销

开本850×1168 mm 1/32 印张12.25 字数329千字

2000年8月北京第1版 2000年8月北京第1次印刷

印数:1—3000

定价:28.00元

序　言

Internet 电子商务爆炸性的发展和信息技术在人类探索未知世界的征程中应用日益广泛，要求计算机系统具有更高的性能、永不停顿的高可用性以及能够容纳几代技术和长期发展计划的可伸缩性。这些需求推动计算机技术和产品都取得了史无前例的飞跃发展。

自 Alpha 在 1992 年开创了 64 位计算先河以来，目前各厂商基于 RISC 处理器的产品都已先后进入了 64 位时代。商品化的 64 位 RISC 处理器正在成为推动信息时代发展的主要引擎之一。新的产品层出不穷、日新月异，不仅反映了市场竞争的激烈，也反映了 64 位 RISC 技术的巨大生命力。

本书取材新颖，力图反映基于 64 位 RISC 处理器的现代计算机技术发展趋势和最新成果。对新一代 Alpha 处理器、Near UMA 体系结构 AlphaServer 服务器、提供单一系统映象的 TruCluster 集群以及把 Alpha 超级计算能力与 Linux 开放性完美地结合在一起的超级计算机等方面都以翔实的资料予以了介绍。我很高兴地指出这些新的产品和成果正在服务于世界上大多数最著名的 Internet 和电子商务网站，支持世界上最

多的关键任务应用，完成了许多规模最大、最复杂的计算任务，包括为人类最近取得的划时代成就——完成人体基因代码草图作出的重大贡献。

俞新昌博士

二零零零年七月四日

目　录

第一章 绪 论

1.1 计算机技术的发展

1.1.1 计算机技术的分代

第一台计算机问世已经半个世纪了，在这期间计算机技术已经经历了五次更新换代。

表 1-1 中，前三代的每一代持续大约 10 年，第四代的时间跨度约为 15 年，10 年前进入了第五代。可以看到，换代的标志主要有两个：第一个标志是计算机的器件。从第一代到第五代计算机，器件发生了根本的变化：从电子管、晶体管发展到集成电路，而集成电路又经小规模、中规模、大规模、非常大规模(VLSI)等阶段发展到超大规模(ULSI)阶段。

器件的更新，其速度、功能、可靠性的不断提高和成本的不断降低，是计算机发展的物质基础。因此，器件的换代是计算机换代的最突出的标志。第二个标志是系统体系结构的特点。系统体系结构的不断改进，许多重要概念的不断提出并且得到实现，推动计算机技术向更高的层次发展。从早期的变址寄存器、通用寄存器、程序中断和 I/O 通道等概念，到虚拟存储器、Cache 存储器、微程序设计、系列机、基于总线的多 CPU 系统、向量处理机等概念，发展到 64 位 RISC 处

表 1-1 计算机技术时代划分简表

	第一代计算机	第二代计算机	第三代计算机	第四代计算机	第五代计算机
时期	1945－1955	1955－1964	1965－1974	1974－1991	1991－现在
器件	电子管和继电器存储器	晶体管和磁芯存储器	中、小规模集成电路	LSI 和 VLSI 微处理器以及半导体存储器	ULSI 微处理器和存储器芯片
互联设备	绝缘导线	印刷电路	多层印刷电路	专用互联网络	通用互联网络
代表性的系统技术	单 CPU 和程序控制 I/O	变址寄存器、浮点运算、多路存储器和 I/O 处理器	微程序控制、流水线、高速缓存和先行处理器	RISC 处理器、多处理器系统、向量超级计算机	可伸缩的并行计算机、集群系统、Internet
软件系统技术		批处理监控程序	多道程序设计和分时操作系统	并行处理的多处理操作系统、编译程序	面向大规模并行处理软件系统技术
程序语言和编程工具	机器语言和汇编语言	高级语言、编译程序和子程序库	高级语言、编译程序和子程序库	用于并行处理和分布式处理的软件工具和环境	Java、微内核、多线程、WWW
代表性系统	ENIAC IBM 701 PricetonIAS	IBM 7030 CDC 1604 Digital PDP－8	Digital PDP－11 IBM 360/370 CDC 6600	IBM PC VAX 780 Cray X/MP	IBM RS/6000 and RS/6000 SP SGI Origin 2000 Compaq AlphaServer and TruCluster

理器、基于 MPP、NUMA、集群等体系结构的可伸缩并行处理系统，计算机系统技术也取得了突飞猛进的发展。如果有人用 ULSI 器件，按照早期的系统结构生产一台计算机，人们也决不会认为它是一

台现代的计算机。因为,早期的计算机和现代计算机的系统结构是根本不可同日而语的。因此,系统结构方面的特点也是计算机换代的重要标志。

第五代计算机在系统结构方面的最大特征是并行处理。研究结果表明:近年来计算机器件的延迟时间仅仅缩短到原来的十几分之一,而计算机系统的处理能力却提高了几百倍,这主要是通过在计算机各个层次广泛使用并行技术来实现的。因此,为了满足 Internet 时代对计算机系统的全面需求,单靠提高器件性能是不够的,必须在器件性能和系统结构两方面双管齐下方能充分发挥器件潜力、提高计算机系统性能。当前,提高计算机性能、满足高可伸缩性和高可用性要求的最主要的系统技术是实现并行处理。

1.1.2 基于 64 位 RISC 处理器计算机系统的地位

自从 1945 年第一台计算机问世以来,在计算机技术的许多重要变革中,最有意义的变革也许是从复杂指令集(CISC)过渡到精简指令集(RISC)体系结构。多年来,处理器的体系结构都朝着复杂化方向发展:指令系统越来越复杂和庞大,人们倾向于设置更多的指令、寻址方式、专用寄存器和功能单元,使得处理器结构越来越复杂,成本越来越高。RISC 体系结构表现出对传统发展趋势的彻底背离,这对当时的计算机体系结构的常规思路是一个很大的突破。RISC 处理器结构简单,便于利用 VLSI 以至 ULSI 工艺制造,也便于在处理器内部实现并行处理。RISC 技术奠定了现代计算机发展基础。

计算机技术的另一重大发展是高端系统和应用从 32 位向 64 位时代过渡。这一过渡首先是从 RISC 处理器开始的。众所周知,处理器"字长"是处理器性能指标主要量度之一。处理器字长与计算机其他性能指标(如内存最大容量、文件的最大长度、数据在计算机内部的传输速度、计算机处理速度和精度等等)也有十分密切的关系。

字长是计算机系统体系结构、操作系统结构和应用软件设计的基础，也是决定计算机系统综合性能的基础。64 位 RISC 微处理器为基于它的计算机系统提供无与伦比的性能潜力和应用前景。

64 位计算机系统并不能够自动发挥其潜力，基于 64 位处理器的计算机系统必须从硬件设计和集成、操作系统、中间件和应用软件开发等方面全面跟上，才能真正发挥 64 位处理器的潜力、提供具有高综合性能的系统。最早问世的 64 位 RISC 处理器是 Alpha 21064 和 MIPS R4000。但是，Alpha 率先完成了从 32 位过渡到 64 位各项技术任务，包括 64 位微处理器设计和生产、各种体系结构、各个档次 64 位计算机系统的设计和生产、64 位操作系统、中间件和应用软件的开发、64 位系统集成和解决方案、技术服务和支持，在此基础上建立了成熟的 64 位技术，创造了大量成功的应用实例、丰富的应用经验和应用成果，充分展示了 64 位技术的优越性，开创了 64 位 RISC 技术应用时代。

此后，HP，IBM，SUN 等厂商也都先后推出了自己的 64 位处理器。到 21 世纪初，32 位芯片的“霸主”Intel 也将推出 64 位 IA-64 体系结构的处理器。这标志着计算机高端技术全面进入了 64 位的时代。

目前，64 位 RISC 处理器已经完全商品化批量生产，其性能已经接近甚至超过 IBM 主机和 Cray 超级计算机中使用的专门设计的处理器，而价格却便宜得多。因此，当前计算机工业的一个十分明显的发展趋势是利用商品化的 64 位 RISC 处理器和并行体系结构来设计和生产高端的服务器和工作站，甚至超级计算机。IBM 的 RS/6000 SP和 SGI 的 Origin 2800 等超级计算机都基于商品化 64 位 RISC 微处理器，康柏最近也推出了基于 Alpha 21264 的超级计算机产品系列 AlphaServer SC。与此相反，基于专门设计微处理器的向量计算机日趋衰落，正在失去其原有的阵地。事实上，在 TOP 500 中占据重要地位的 Cray T3E/T3D 所使用的也是 Alpha 微处理器。

IBM 和 SGI 为美国核科技应用提供的“蓝色太平洋”和“蓝色山脉”等著名的超级计算机、以及康柏为法国原子能委员会生产的欧洲最大超级计算机，都是基于 64 位 RISC 处理器的并行计算机系统。

目前各主要厂商都有自己的 64 位 RISC 处理器产品，并且不断地用主频更高、速度更快的处理器装备自己的服务器和工作站系列。在此基础上，提供各种的集群系统以及基于 CC-NUMA 或 MPP 体系结构的超级计算机系统。基于 64 位 RISC 处理器计算机系统已经应用到几乎一切领域，积累了十分丰富的应用软件。单在基于 Alpha 处理器平台上的 64 位应用软件就已经超过了 12000 个。RISC/UNIX 服务器的市场规模也已经超过每年 500 亿美元。

虽然基于 64 位 RISC 处理器的计算机系统不可能像基于 Wintel 工业标准的 PC 机和服务器那样普及，也不可能取代 PC 机。但在高端应用中，基于 64 位 RISC 处理器的计算机系统在当前和可以预见的未来都将占据中心位置，从入口级的服务器到超级计算机都采用 64 位 RISC 处理器。特别是在高端应用中，商品化的 64 位 RISC 处理器产品不仅已经成为装备服务器和工作站的主流芯片，而且越来越多地取代传统的专门设计的向量处理器，成为装备超级计算机的主流芯片。这已经成为超级计算机技术重要的新发展趋势之一。无论是通过 Web 查询万里以外的信息、在电子商务网点上购物、收听气象预报、预订机票和客房、观看影片“坦泰尼克号”中的数字特技，都有基于 64 位 RISC 处理器的计算机系统为此服务。

因此，本书将围绕基于 64 位 RISC 处理器的计算机系统来介绍计算机技术的发展和应用。

1.2 Internet 和电子商务对计算机系统的需求

Internet 和电子商务应用爆炸性的发展对计算机技术提出了越来越高的要求，其中主要有：

丰富的系统资源容量 要求提供支持更广泛类型、更大规模应用的系统资源容量,包括速度更快和数量更多的处理器、更大的内存和磁盘容量、提供更高的系统带宽和I/O及网络联接能力等。

高可伸缩性 要求提供全面和经济的可伸缩性,一方面允许用户以最低的入口价格购买本档次的产品,另一方面又提供具有最高的扩展上限,对现有设备运行干扰最小,效率最高的可伸缩性。

高可用性 要求计算机系统提供支持全球性Internet和电子商务应用所需的永不停顿的高可用性。

1.2.1 计算机系统的资源容量

计算机系统资源一般是指CPU性能和数量、主存储器容量、I/O和网络联接能力等。当今的世界已经进入了信息时代,全球性的Internet和电子商务应用、人类探索和预测未知世界的高性能技术计算应用对计算机系统的资源容量提出了越来越高的要求。这些要求不是凭空而来、而是由这些应用的规模和实际需求决定的:

根据IDC 2000年3月的报告,目前全球已有Internet用户1.96亿个,到2003年将达到5亿个(据国内的统计,我国的Internet用户已经超过1千万个),能够访问Internet的设备将超过7.5亿台。如此大规模的应用,要求计算机系统能够接待大量的同时用户,并且以最快的速度完成用户的查询请求。例如,著名的Internet门户网站Alta Vista使用几十台运行Tru64 UNIX的AlphaServer,每天检索全球27万多个Web服务器上4 500万个Web页中120亿个字节的信息。每天有2 000多万人次使用这一服务,检索响应时间一般小于1 s。

根据IDC同一报告,目前全球电子商务的营业额为2 690亿美元,到2004年将达到25 000亿美元。电子商务从根本上改变了商业的模式:传统商务每年也许仅仅处理几千个订单,可以容忍几个星

期的延迟;而电子商务每年需要处理几百万笔交易,必须把全球价值链(GVC)、客户关系管理(CRM)和知识管理(KM)等核心应用集成在一起,实现所谓"零延迟"(Zero Latency)。目前客户要求所访问的电子商务以最快的速度响应,延迟时间超过 50 s,客户往往就将失去耐心,转向别的网站。

这一切都要求计算机系统具有最大的资源容量,不仅在Internet和电子商务应用中,在帮助人们发展科技、探索未知世界的高性能技术计算应用中,也要求最大的计算机系统资源。90 年代初,为了满足许多大型应用的需要,人们设想了"3T 的系统资源目标",即生产出具有 1TFLOPS 计算能力、1TB 内存容量和 1TB/s I/O 带宽的计算

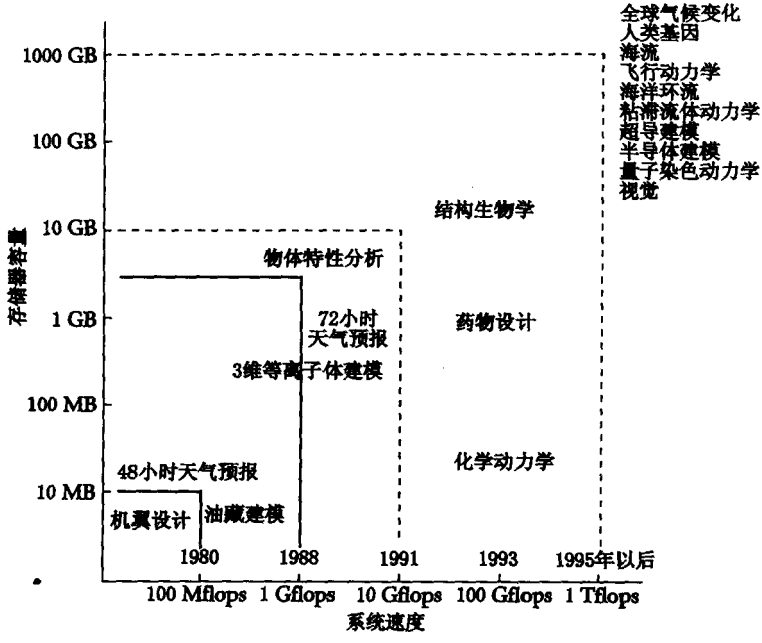

图 1-1 大型课题对计算机系统资源容量的要求

机系统。图 1－1 表示一些大型课题对计算机系统资源容量的要求。

事实上,"3T 系统资源目标"还是不能满足一些高科技项目和全球性电子商务应用的需求。例如,1994 年,美国能源部(DOE)启动了加速战略计算协作计划(ASCI)。ASCI 是一个投资高达 10 亿美元的 10 年计划,目标是尽可能地使用商品化市售硬件和软件成批地生产出速度达到多个 TFLOPS 的超级计算机,并应用于模拟长期储存对核武器影响、生物技术、医疗和新药研究、长期气象预报、飞机和汽车设计、工业过程改进和环境保护等领域。ASCI 根据美国核科学技术等领域对计算机系统的需求,提出了比"3T"更高的发展目标,如表 1-2 所示。

表 1-2 ASCI 提供平衡的可伸缩计算环境的发展规划

属性	1996 年	1997 年	2000 年	2003 年
应用性能(倍数)	1		1000	100 000
计算速度峰值(GFLOPS)	100	1000	10,000	100 000
内存容量(TB)	0.05	0.5	5	50
磁盘容量(TB)	0.1-1	1-10	10-100	100-1000
归档的存储容量(PB)	0.13	1.3	13	130
I/O 速度(GB/s)	5	50	500	5000
网络速度(GB/s)	0.13	1.3	13	130

除了对超级计算机外,许多大规模的企业级应用也对服务器系统的资源容量提出了很高的要求。表 1-3 列出各厂商高端服务器目前能够提供的系统资源容量。