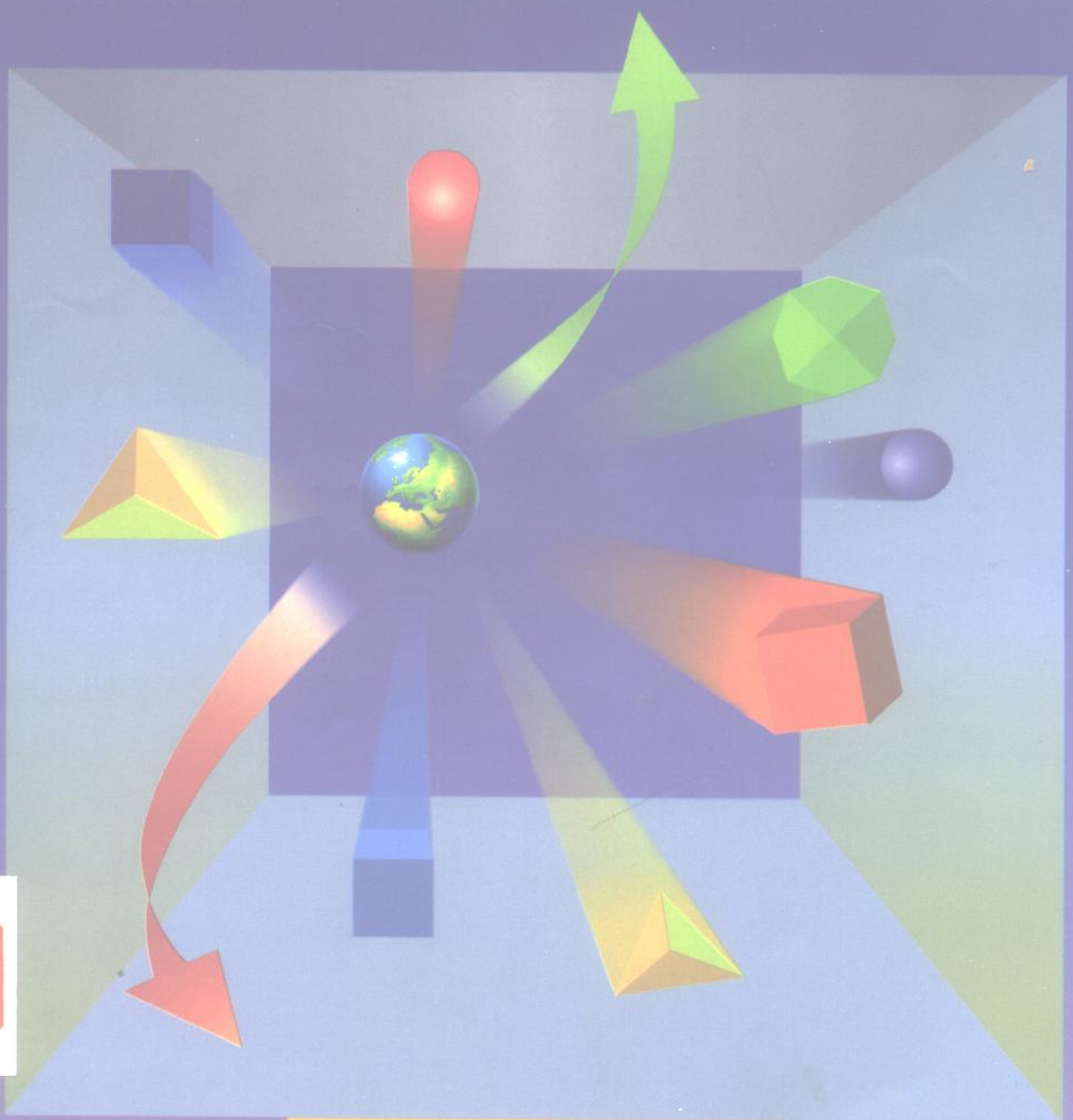
Pentium Ⅱ/Ⅲ

体系结构及扩展技术

刘清森　马鸣锦　吴灏　曾光裕　赵秋霞　田家斌　编著



国防工业出版社

Pentium Ⅱ/Ⅲ体系结构及扩展技术

刘清森　马鸣锦　吴　灏
曾光裕　赵秋霞　田家斌　编著

国防工业出版社
·北京·

图书在版编目(CIP)数据

Pentium Ⅱ/Ⅲ体系结构及扩展技术/刘清森等编著.
—北京:国防工业出版社,2000.7
ISBN 7-118-02270-5

Ⅰ.P… Ⅱ.刘… Ⅲ.电子计算机-中央处理机,
Pentium Ⅱ/Ⅲ-系统结构 Ⅳ.TP332

中国版本图书馆 CIP 数据核字(2000)第 05282 号

国防工业出版社出版发行
(北京市海淀区紫竹院南路 23 号)
(邮政编码 100044)
三河市腾飞胶印厂印刷
新华书店经售
*
开本 787×1092 1/16 印张 21¾ 496 千字
2000 年 7 月第 1 版 2000 年 7 月北京第 1 次印刷
印数:1—3500 册 定价:32.00 元

(本书如有印装错误,我社负责调换)

前　言

并行处理技术、多媒体计算机技术与超大规模集成电路技术和微电子组装技术相结合，促进了微处理器和微计算机向超级化方向发展。Pentium Ⅱ和 Pentium Ⅲ处理器就是超级性能处理器的一类。Pentium Ⅱ/Ⅲ处理器内部采用超标量、超流水线微结构，做到每个处理器时钟周期可以并发多条指令；采用双芯片设计，使得两级高速缓存与处理器紧耦合，不仅提高了高速缓存的命中率，而且能支持处理器全速工作；新引入的 MMX 技术和流式 SIMD 扩展技术，大大增强了对通信和多媒体的处理能力；新颖的处理器总线结构，能在不增加外部逻辑的情况下，建立起对称式的多处理器计算机系统。

Pentium Ⅱ/Ⅲ是 Intel 体系结构(IA)系列中最新一代处理器，既引进了大量先进技术，又在目的代码一级上与 IA 系列家族保持向上兼容。因此，从 IBM PC 微机以来所开发的系统软件和应用软件，均可以不经修改而在新系统中正确运行。这就是 1997 年推出 Pentium Ⅱ系统和 1999 年推出 Pentium Ⅲ系统后能很快推广的主要原因。不过，这些软件虽然可以在 Pentium Ⅱ/Ⅲ系统上运行，但运行这类软件不能充分发挥处理器的功能，因此，在 Pentium Ⅱ/Ⅲ系统上，开发出体系结构所支持的并经过代码优化的新软件，是应用开发的一项重要任务，而从事 Pentium Ⅱ/Ⅲ系统的高级应用开发，往往需要了解处理器在体系结构上的特点、性能，使用方法等，编写本书就是为了满足这方面的需求。

本书是作为高校教学的参考书而编写的，按照高等学校计算机学科课程改革的思路来组织内容。我们的设想是：如果课程设置偏重于计算机应用的话，则学完计算机组织与结构(或者计算机原理)课程之后再选修本课程，可在已有的计算机和微机知识基础之上，增加高速缓存、虚拟存储器和并行处理的基本知识，就能够理解新一代奔腾处理器的工作原理；如果课程设置偏重于计算机系统设计的话，则可在系统结构课程之后选修本课程，作为一般的系统结构知识的补充。本书还可以作为计算机专业硕士研究生的“高性能微机系统”、“高级计算机系统结构”及“多媒体计算机技术”等课程的参考用书。

全书共 11 章及 5 个附录，可分为四部分。第一部分为第 1、2 章，基于 IA 体系结构之上介绍 P6 系列处理器的基本组成及系统结构特点。第二部分为第 3～7 章，介绍 Pentium 系列的存储管理、存储保护、中断及异常处理和多任务管理，此外还介绍了虚拟 8086 任务的基本概念。第一、二部分作为基本内容，是一门课程中的应授部分，其余为选授部分。第三部分为第 8、9 章，主要讨论多处理器管理问题。第四部分为第 10、11 章，这部分介绍了 MMX 技术和流式 SIMD 扩展技术，为了帮助读者理解这类新技术的应用，特举一些例子作说明。5 个附录为整型指令、浮点指令、系统指令、MMX 指令集和流式 SIMD 扩展指令集。在第 10 章和第 11 章中简要地介绍了两类多媒体与通信扩展的指令系统，如果要深入了解指令的功能和用法，还必须参看附录 D 及附录 E。

本书第 1 章由马鸣锦编写，第 2、7 章由曾光裕编写，第 3、4 章由赵秋霞编写，第 5、6 章由吴灏编写，第 8～11 章和附录 A～E 由刘清森和田家斌编写。全书在相互审阅的基

础上由刘清森定稿。田家斌制作了书中全部图表。

全书的内容是在 Intel 公司网上发布的新资料、我们的讲课资料和研究生论文的基础上整理而成。感谢解放军信息工程大学计算机科学与技术系的领导和同事们给予的支持和帮助。

这是国内第一本介绍 Pentium Ⅱ/Ⅲ系统的书，资料收集分析不完善，加上作者知识水平有限，书中有错误和不足之处，敬请读者批评指正。

编　者

此为试读，需要完整PDF请访问：www.ertongbook.com

目　录

此为试读，需要完整PDF请访问：www.ertongbook.com

第 1 章　P6 系列处理器概况

P6 系列处理器中两种典型的处理器 Pentium Ⅱ和 Pentium Ⅲ，今天已经是家喻户晓。从微处理器的发展历程来看，Pentium Ⅱ/Ⅲ处理器仍属于微处理器一类，或者称之为高性能微处理器，同样，以它们为主处理器所构成的计算机系统也属于微型计算机一类。但是，从体系结构的观点来看，今天的微处理器和微型计算机系统已经超越了微型系统的界线，进入了大型计算机甚至巨型系统的行列，使得现在的微型计算机与超级小型机和大型机之间的界线已愈来愈模糊了。同样，大型计算机的体系结构和系统组成也发生了很大的变化，它们已经不是过去的分立系统的概念，而是使用一种或多种微处理器，由多个微处理器芯片构成其主 CPU 系统，所以，从另一角度来说，今天的大型计算机系统又是传统观念上的微型计算机系统。

Pentium Ⅱ/Ⅲ处理器芯片的执行速度已经超过了每秒钟 10 亿条指令（或者每秒 10 亿次微操作）。它广泛采用 RISC（精简指令系统计算机）技术，实现了超标量、超流水线结构，强调硬件设计与软件技术相结合，成为一种超级微处理器。Pentium Ⅱ/Ⅲ处理器还引入了支持通信与多媒体数据处理的扩展技术，显著地增强了该类微机系统的处理能力。

综上所述，P6 系列处理器与传统观念的微处理器有很大的差别，因此本书不使用传统的“微型计算机原理”课程的讲述方法来介绍 Pentium Ⅱ/Ⅲ处理器，而着重从体系结构及应用的角度介绍它们。

由于在讨论处理器体系结构的过程中，经常要涉及系统组成的硬件知识，因此本章将对一些基本的硬件知识进行概述性的介绍，以有利于对本书整体内容的理解。本章首先回顾 IA 体系结构的发展过程，从 IA 系列处理器的发展中看出 P6 系列处理器的特点；接着从处理器微结构的角度来讨论其超标量、超流水线的组织，以及处理器内部实现“动态运行”的方法和效果；高速缓存是支持处理器高速运行以及支持多处理器结构的一个重要部件，本章 1.3 节简要介绍 P6 处理器高速缓存结构及其一致性协议，为了便于初学者阅读，稍微揉入了一些对高速缓存一般性硬件工作原理的叙述；本章最后一节介绍 P6 系列处理器的外部总线结构及总线操作规程，从这种与众不同的总线操作方法中可以看出，P6 系列处理器是如何有效地支持对称式多处理器的建立的。

1.1　Intel 体系结构的发展

1.1.1　Intel 体系结构发展的历史回顾

回顾 Intel 体系结构（IA）的发展历程，可以追溯到 30 年前，1969 年 Intel 设计的第一个微处理器是 4 位微处理器——4004。70 年代初设计出 8 位微处理器——8080/8085。然而，开创 IA 系列的第一个处理器实际上是 8086，于 1978 年推出；紧接着是 8088，这是

用于较小型系统的一种 8086 低价版本微处理器。为 8086/8088 处理器所开发的目标代码程序,至今仍可在 IA 系列中的各种处理器上运行。

8086 具有 16 位内部寄存器和 16 位外部数据总线,有 20 位地址,可直接寻址 1MB 的地址空间。8088 除了外部数据总线为 8 位之外,其余与 8086 相同。8086/8088 引入了存储器段管理机制,但只是在"实方式"下分段。4 个段寄存器,每个含有当前段 20 位基地址的高 16 位,以 16 位为段内偏移量可寻址的段内地址空间达 64KB。分段可在整个 1MB 的地址空间上实现。

Intel286 处理器将"保护方式"引入 IA。这种新方式用段寄存器的内容作为选择符或者作为指向描述符表的指针,去索引描述符表中的一个段描述符或者指向描述符表。段描述符提供 24 位基地址,最大的物理存储器空间达 16MB。保护方式提供了基于段切换的虚拟存储器管理,还提供了多种保护机制。这些保护机制包括段限检查,只读/只执行段的选择,以及设置 4 级特权级以避免操作系统代码受应用程序或用户程序的影响。此外,硬件任务切换和局部描述符表的实施,使得操作系统能够防止各应用程序或用户程序之间的相互影响。

Intel386 处理器把 32 位寄存器引入 IA 结构,32 位寄存器可作为操作数用于运算和寻址。每一个 32 位寄存器的低半部分保留着 16 位寄存器的特性以提供向上的兼容性。当在这种 32 位机器上执行 8086/8088 的程序时,可选择在虚拟 8086 方式下运行,以获得更高的效率。Intel386 有 32 位外部地址总线,支持 32 位寻址,提供 4GB 地址空间,也使每个段最大可以达到 4GB。386 在原指令系统的基础上增加了 32 位的操作数并扩展了寻址方式,还提供了新的指令。Intel386 处理器把分页技术引入 IA 结构,页面大小固定为 4KB,为管理虚拟存储器提供了一种机制。当把段定义为 4GB 时,用分页技术就可以在结构上形成一种保护的"平面模型"(只用一个 32 位地址元素便可访问到整个地址空间)的寻址系统,386 一类主机的 UNIX 操作系统就是这样实现的。

IA 结构一直在目标代码一级上维持向后兼容,以保护用户在软件方面的巨大投资,与此同时,在每一代产品中也使用当时最先进的微处理器结构及硅制造技术,尽可能制造出最快的、具有最强功能的处理器。Intel386 处理器是 IA 结构中第一个含有多个并行部件(或称为单元)的处理器。它有 6 个并行部件:总线接口单元、指令代码预取单元、指令译码单元、执行单元、段管理部件和页管理部件。

Intel486 处理器增加了并行处理的能力,它把 Intel386 处理器的指令译码单元和执行单元扩展成 5 个流水线级,每一级都与其他级并行操作,在不同的执行级上可同时运行 5 条指令。在一个时钟内,每一级完成一条指令中属于它的那部分工作,这样,Intel486 处理器最快可以每个 CPU 时钟执行一条指令。Intel486 片内增加了一个 8KB 的第一级(L1)高速缓存(cache),使得每时钟执行一条指令的几率增加。比如,在执行一条存储器访问指令时,如果欲访问的操作数就在 L1 高速缓存中,那么就能在一个时钟内完成。Intel486 也是第一个将浮点数值运算单元集成在片内的 IA CPU。另外它还增加了新的引脚、控制位和相应的指令用以支持更复杂、功能更强的系统(支持 L2 高速缓存和多处理器)。

Intel486 后期产品,引入了支持节电的系统管理设计,以专门用于电池供电的笔记本 PC。这种新设计采用新的系统管理方式(SMM),由它自己专用的中断引脚触发,故在把

复杂的系统管理功能(例如,PC 内系统时钟和系统电源管理)加入到系统中去的时候,对主机操作系统和应用程序是透明的。停止时钟和自动暂停是两种低功耗电源管理方式,前者是 CPU 本身以低时钟速率运行,后者是更省电的关机。

Intel Pentium(奔腾)处理器增加了一条执行流水线,实现了超标量结构。两条流水线称为 U 和 V 流水线,每时钟可执行两条指令。片内的 L1 高速缓存比 486 增加了一倍,8KB 用于代码,8KB 用于数据,数据高速缓存使用 MESI 协议,支持高效的回写方式,也支持 Intel486 处理器所采用的通写方式。Pentium 处理器增加了具有片内分支表的分支预测功能,使循环指令的执行效率提高。另外还作了一些扩展和改善,使虚拟 8086 方式更加有效;使页面大小可以为 4KB 和 4MB。主要的寄存器仍然是 32 位,但内部的数据通路增加到 128 位和 256 位,用来加快内部数据的传送;可进行猝发传送的外部数据总线宽度增加到 64 位。设置了新的引脚和一种特殊的模式,以支持无需附加连接逻辑的双处理器系统。增加了先进的可编程中断控制器(APIC),用以支持多处理器系统。

Intel Pentium Pro 处理器是 IA 结构中 P6 系列处理器的第一代产品。Pentium Pro 处理器已具有三路超标量结构,这意味着每个 CPU 时钟能够执行三条指令,显然比 Pentium 具有更强的并行性。Pentium Pro 处理器在超标量的履行过程中采用动态执行技术(微数据流分析、乱序执行、先进的分支预测和猜测执行)。3 个指令译码单元并行工作,将目标代码翻译成更小的操作,称为"微操作"(micro-op)。这些微操作再送入指令池中,当微操作之间的相关性不妨碍独立执行时,可由 5 个并行执行部件(两个整型部件、两个浮点部件和一个存储器接口部件)乱序执行,然后退出单元按照原程序的顺序(包括分支)撤下已完成的微操作。

Pentium Pro 处理器通过内部的 cache 进一步增强其能力:与 Pentium 一样也有两个片内 8KB L1 cache,另外,256KB 的 L2 cache(第二级高速缓存)芯片与 CPU 芯片封装在同一模块里,使用专用的 64 位"全时钟速度"总线紧耦合到 CPU。它的 64 位外部数据总线是面向事务运行的,即按分立的请求和响应来处理每次访问,在等待响应时,允许多个请求存在。这些用于数据访问的并行特性连同并行执行能力一起,可提供一种"无阻塞"的结构,在这种结构里,处理器得到充分的利用,处理性能显著增强。Pentium Pro 处理器地址总线扩展为 36 位,地址空间最大为 64GB。

P6 系列第二代产品是 Pentium Ⅱ处理器。它在 Pentium Pro 处理器先进结构的基础上增加了对 MMX(多媒体扩展技术)的支持。这种支持包括 4 种新数据类型、8 个 64 位 MMX 寄存器、57 条 MMX 指令,使得处理器对具有大数据量特点的多媒体数据的运算和处理能力大大提高。Pentium Ⅱ处理器把 L1 数据 cache 和 L1 指令 cache 都扩充为 16KB,L2 cache 的容量可以为 256KB、512KB、1MB 或 2MB。

IA 结构中的最新处理器是 Pentium Ⅲ处理器,它是基于 Pentium Pro 和 Pentium Ⅱ处理器结构上的 P6 系列第三代产品。Pentium Ⅲ处理器引入了流式 SIMD 扩展技术,包括增加了 70 条新指令、8 个 128 位寄存器和紧缩单精度浮点数据类型,以支持多媒体与通信的处理。

1.1.2 IA 性能的提高

早在 60 年代的中期,Intel 的委员会主席戈登·摩尔曾推断过一条原则或"定律":硅

集成电路微处理器的计算能力和复杂程度(粗略地说,即 CPU 芯片的晶体管数目)每 1～2 年提高一倍,而 CPU 芯片的价格降低一半。这个推断已在过去的 30 多年中得到证实。这条定律实际上是对计算机革命的诠释,在这场革命中,IA 起了促进作用。

表 1-1 示出了在 IA 发展历史上,IA 处理器的性能和晶体管数目的增长情况(正如摩尔定律所预言的那样),也汇总了关键特性的变革情况。

表 1-1 各时期处理器性能及关键特性

Intel 处理器	产品日期	性能 MIP①	CPU 最高频率	晶体管数	主 CPU 寄存器规模②	外部数据总线规模②	外部最大地址空间	CPU 封装内的 cache③
8086	1978	0.8	8MHz	29K	16	16	1MB	无
Intel286	1982	2.7	12.5MHz	134K	16	16	16MB	无
Intel386DX	1985	6.0	20MHz	275K	32	32	4GB	无
Intel486DX	1989	20	25MHz	1.2M	32	32	4GB	8KB L1
Pentium	1993	100	60MHz	3.1M	32	64	4GB	16KB L1
Pentium Pro	1995	440	200MHz	5.5M	32	64	64GB	16KB L1; 256KB 或 512KB L2
Pentium II	1997	466	266MHz	7M	32	64	64GB	32KB L1; 256KB 或 512KB L2
Pentium III	1999	1000	500MHz	8.2M	32GP 128SIMD-FP	64	64GB	32KB L1; 512KB L2

注:

① 此处的性能仍用 Dhrystone 的 MIPS(每秒钟百万条指令)表示,尽管不再认为 MIPS 是一种最佳的 CPU 性能测试方法,但它还是估测 IA 处理器的唯一基准。此处的 MIPS 值是最高频率下的值。

② 主 CPU 寄存器的规模和外部数据总线的规模以位给出。另外,在所有的 CPU 里都有 8 位和 16 位的数据寄存器;Intel486 以上处理器中的浮点部件里有 8 个 80 位的寄存器;对每一种处理器,内部的数据路径宽度是外部数据总线的 2～4 倍。

③ 除了表中所列到的通用 cache 以外,各处理器还有小型专用的 cache,情况如下:Intel286 中每个段寄存器有 6 字节的描述符 cache;Intel386 中每个段寄存器有 8 字节的描述符 cache,还有一个 32 项、4 路组相联的转换后援缓冲器(TLB)cache,存放最近所用到的页的地址转换信息;Intel486 有与 Intel386 相同的 cache;Intel Pentium 和 Pentium Pro 处理器除了有描述符 cache,还有两个 TLB cache,每个 8KB L1 cache 配一个 TLB cache,分别进行 L1 指令 cache 和 L1 数据 cache 的虚拟地址到物理地址的快速转换;Pentium II 和 Pentium III 处理器的 cache 结构与 Pentium Pro 处理器一样。

1.1.3 IA 浮点部件发展的历史回顾

IA 系列的第一个处理器 8086 推出后,Intel 就为 8086/8088 设计了相配套的浮点运算处理器 8087。8087 是专门用硬件实现数值运算的处理器,大大加快了计算机数值运算的速度。其指令系统支持 16/32/64 位整数、32/64/80 位实数、18 位压缩 BCD 数的加减乘除运算,还支持求平方根、三角函数、双曲函数和指数函数的运算,采用早期的 IEEE754 二进制浮点运算标准。为简化设计,8087 不具有独立访问存储器的能力,它与 8086/8088 紧耦合连接,共用处理器总线,构成主从关系,由主处理器 8086/8088 替它取指,在 8086/8088 协助下读/写内存操作数,可与 8086/8088 并行执行各自的指令,所以常称之为“数值协处理器(MCP)”,或“数值处理器扩展(NPX)”。

Intel286 推出后，Intel 为它设计了相配套的 80287NPX。80287 的内部寄存器、数据类型和指令与 8087 兼容。从硬件上看，80286 与 80287 仍为主协结构，但 80287 与 80286 的耦合关系及并行工作机理与 8087 有所不同。80287 与 80286 仅共用部分处理器总线，80287 在系统中的地位好像是挂在处理器总线上的一个 I/O 芯片，完全没有访问存储器的能力。80286 取到 287 的指令后，通过固定的 I/O 端口交付于 80287，80287 执行指令需读/写内存操作数时，80286 就相当于一个 DMA 控制器，在内存与 80287 之间传送数据，这样，就不要求 80287 像 8087 一样与主处理器以同一时钟同步工作，简化了设计，而且这一切都由硬件实现，对软件透明。

80287 可以在 80286 的多任务、多用户系统中运行，80286 的全部保护功能能支持使用数值处理器扩展(NPX)的多任务。

Intel80387DX 和 SX 协处理器是 Intel 的第三代数值运算处理器。它们完整地执行了 IEEE754 标准。增加了新的三角函数指令，使用了新的设计和工艺以提高时钟频率并减少每条指令所需要的时钟周期数。

从 Intel80486 处理器开始，浮点运算部件 FPU 集成于 CPU 芯片内部，Intel80486 内的 FPU 与 Intel80387DX 数值协处理器等价，并遵循 IEEE754 标准和后来推出的 IEEE854 标准。FPU 放置于 CPU 芯片内部，使引线缩短，内部数据总线加宽，因此处理速度比 80387 提高 3～5 倍。

Pentium 处理器的 FPU 在 Intel486 处理器的 FPU 基础上全面重新设计，同时仍保持与 IEEE754 和 854 标准相容。对于通用操作(浮点加、乘和加载)，Pentium FPU 的性能比 Intel486 FPU 至少要高 3 倍。当使用指令调度和流水执行时，很多应用程序可以达到 Intel486 FPU5 倍以上的速度性能。

1.2 P6 系列处理器的微结构

1995 年，Intel 公布了 P6 系列的第一代处理器 Pentium Pro。P6 系列处理器芯片结构设计的主要目标是，在仍使用与 Pentium 相同的 0.6μm、4 层金属 BICMOS(双 CMOS)制造工艺的条件下，要使处理器的性能明显超过 Pentium 处理器，这意味着高性能的获得只有通过对微结构的重大改进才能达到。

1.2.1 流水线的动态执行机构

P6 系列处理器的微结构设计为三路超标量、超流水线结构，这表明 P6 处理器使用了更高性能的并行处理技术，每个时钟平均可以译码、分配和执行三条指令。为了提高指令的通过量，P6 系列处理器使用一种已去除了级间相关性的 12 级流水线，以支持指令的乱序执行(动态执行)。

图 1-1 所示的是流水线的基本思路。流水线分为 4 个处理单元(取指/译码单元、分配/执行单元、退出单元和指令池)，通过总线接口单元向处理单元供给指令和数据。

为了保证能稳定地向指令执行流水线供给指令和数据，P6 系列处理器微结构引入 L1 和 L2 两级高速缓存。L1 高速缓存与流水线紧密相连。静态 RAM 构成的 L2 高速缓存通过全时钟速度的 64 位专用 cache 总线与处理器内核耦合。

此为试读，需要完整PDF请访问：www.ertongbook.com