

新一代信息通信技术书系·互联网专辑

网络处理器 原理与技术

张宏科 苏伟 武勇 编著

WANGLUO

CHULIQI YUANLI

YU JISHU



北京邮电大学出版社
www.buptpress.com

TP393.02
10

新一代信息通信技术书系·互联网专辑

网络处理器原理与技术

张宏科 苏伟 武勇 编著

北方工业大学图书馆



00565337

北京邮电大学出版社

·北京·

内 容 简 介

本书首先从基本原理和具体技术的角度对网络处理器进行了概要介绍,然后主要以 Intel 公司的第二代网络处理器 IXP2XXX 为例,分硬件结构、软件流程、编程指令、应用实例、软件开发工具的使用等方面对网络处理器进行了详细的分析。

本书取材新颖,内容丰富,实用性强,可作为初学者了解网络处理器基本知识的入门指南,也可作为利用网络处理器进行项目研发的工程人员的参考手册,还可作为通信专业的研究生和高年级本科生学习网络处理器知识的教材。

图书在版编目(CIP)数据

网络处理器原理与技术/张宏科,苏伟,武勇编著. —北京:北京邮电大学出版社,2004
ISBN 7-5635-0953-4

I. 网... II. ①张...②苏...③武... III. 计算机网络—微处理器 IV. TP393.02

中国版本图书馆 CIP 数据核字 (2004) 第 107810 号

书 名: 网络处理器原理与技术

编 著: 张宏科 苏伟 武勇

责任编辑: 马莹娜

出版发行: 北京邮电大学出版社

社 址: 北京市海淀区西土城路 10 号(100876)

电话传真: 010-62282185(发行部) 010-62283578(FAX)

E-mail: publish@bupt.edu.cn

经 销: 各地新华书店

印 刷: 北京源海印刷有限责任公司

开 本: 787 mm×1 092 mm 1/16

印 张: 20.25

字 数: 425 千字

印 数: 1—5 000 册

版 次: 2004 年 11 月第 1 版 2004 年 11 月第 1 次印刷

ISBN 7-5635-0953-4 / TP·128

定 价: 29.80 元

· 如有印装质量问题,请与北京邮电大学出版社发行部联系 ·

前 言

进入 21 世纪,网络发展越来越明显地呈现出宽带化和综合化的趋势。首先,随着网络传输技术的发展,特别是光纤传输技术的成熟,网络传输速度已从 155 Mbit/s、622 Mbit/s 等中低端速率迅速攀升到 2.5 Gbit/s、10 Gbit/s,甚至 40 Gbit/s 等中高端速率。其次,由于三网合一趋势的日益增强,以及人们需求的快速增长,要求网络能够同时承载语音、数据、多媒体以及其他形式的网络业务。

从整体上来看,网络可分为终端、传输、交换(网络处理)三大部分。从技术层面上来看,随着智能、高速终端的发展,高速传输技术(特别是光纤传输技术)的成熟,交换日益成为决定网络宽带化和综合化的瓶颈因素。再从业务驱动的角度来看,随着多种网络业务的综合化以及新型业务的不断出现,要求网络处理设备不仅能够进行快速的路由交换,而且还能对不同网络业务进行复杂、深度的智能化处理。

传统的网络处理设备基本可分为两大类,一类是基于通用处理器来构建的,另一类是基于专用处理器(ASIC)来实现的。通用处理器的优点是功能灵活,能够处理多种网络业务,通过更新升级网络处理软件即可适应网络业务的更新和升级;其缺点是处理速度比较慢,只能应用在中低端速率场合(如低端路由器)。针对通用处理器处理速度慢的不足,为满足高速的网络处理要求,人们设计了专用处理器来进行网络业务处理,将某些网络处理功能固化为特殊的硬件结构,以实现高速的网络处理(如高速的 2 层以太网交换机)。专用处理器的缺点是网络处理功能单一,灵活性不好,为提供新的网络处理功能以适应新的网络业务,需要重新设计硬件结构,因此开发周期长,开发成本高,市场存活时间短。

为了更好地适应网络宽带化和综合化的趋势,弥补通用处理器和专用处理器在网络处理方面的不足,同时也结合了通用处理器与专用处理器在网络处理方面的优势,人们设计了网络处理器(Network Processor, NP)。网络处理器有效地综合了通用处理器和专用处理器各自的优势,在网络处理功能和性能之间取得了较好的平衡,已被设备制造商和网络运营商认定为新一代网络设备的核心,其价值也逐步为市场所承认。在网络处理器中,专门为网络处理而进行优化设计的硬件结构提供了高速强大的网络处理平台,以适应网络处理宽带化的趋势,完全可编程的网络处理模式提供了灵活多样的网络处理功能,以满足网络业务综合化的要求。

网络处理器的出现不仅意味着技术的进步,还代表着理念的更新,更是事物发展的逻辑结果和必然趋势。通用处理器强调了网络处理功能的灵活性而忽略了性能的强大性,

专用处理器则正好相反。这两种单极化的思路都只强调了事物的某一方面,而事物总是两面性的,只有有效地兼顾了矛盾的两方面,才能获得最优的整体效果。网络处理器正是基于这样的理念而设计的。就功能的灵活性而言,网络处理器比不上通用处理器,就性能的强大性而言,也不如专用处理器。但网络的宽带化和综合化是相伴而生、相辅相成的,网络业务的处理需求也是多方面的,因此,从最优化理论的角度来看,网络处理器比通用处理器和专用处理器更具有优势,更能适应当前和今后网络发展的趋势。

本书主要以 Intel 公司的第二代网络处理器 IXP2XXX 为例,从网络处理器的基本原理和具体技术的角度,分硬件结构、软件流程、编程指令、应用实例等方面对网络处理器进行介绍。本书可作为初学者了解网络处理器基本知识的入门指南,也可作为利用网络处理器进行项目研发的工程人员的参考手册。

本书的章节安排如下:第 1 章 绪论,概要介绍网络处理器的基本概念、发展历史、研究现状。第 2 章 Intel IXA 简介,介绍 Intel IXA(Intel 因特网交换体系架构)的基本概念和结构组成。第 3 章 IXP2400 硬件体系结构,以 IXP2400 为例详细介绍 Intel 第二代网络处理器 IXP2XXX 的硬件结构。第 4 章 基于 IXP2XXX 的网络处理,介绍网络处理器进行网络处理的软件流程,以及相应的特殊的硬件结构和软件机制。第 5 章 IXP2XXX 编程参考,介绍 IXP2400 的编程指令语法及其用法。第 6 章 IXP2XXX 的应用,介绍基于 IXP2XXX 进行项目开发的一般过程以及其中的重要问题。附录 开发工具 Workbench 的使用,介绍 IXP2XXX 软件开发工具的使用。

本书的内容框架和结构组织由张宏科老师统筹规划,另外张老师还对书稿进行了最后的修改、把关,第 1、2、6 章由苏伟编写,第 3、4 章以及附录由武勇编写,第 5 章由林宗平整理,另外林宗平还负责了全书插图和表格的绘制。除此之外,实验室的张思东老师、秦雅娟老师、王江林老师、郜帅老师、刘颖老师以及研究生宋坤、罗华、龙宇祥、杨冬、候成杰、刘枫也对本书的编写提供了有益的帮助,在此一并致谢。

在本书的编写过程中,Intel 公司给予了大力的支持,提供了关于 IXP2XXX 网络处理器的硬件指南、软件手册、编程指南、参考设计等多方面详细的资料。本书正是在对这些资料进行归纳整理的基础上编写而成的。另外,本书还特别参考了 Bill Carlson 先生的《Intel Internet Exchange Architecture and Application——A Practical Guide to IXP2XXX Network Processors》一书。在全书的出版过程中,还得到了 Intel(中国)公司大学计划项目部经理曹捷先生以及 Intel 工程师刘英伟先生的大力支持和热心帮助,在此特表谢意。

北京邮电大学出版社对本书的出版给予了大力支持,在此致以深深的谢意。

由于时间仓促,编者水平和所掌握的资料有限,书中错误之处在所难免,恳请读者批评指正。

编 者

2004 年 9 月于北京交通大学

目 录

第 1 章 绪论	1
1.1 网络处理器的基本概念	1
1.1.1 网络处理器的产生背景	1
1.1.2 网络处理器的定义	1
1.1.3 网络处理器的基本组成	2
1.2 网络处理器的发展状况	3
1.3 Intel 网络处理器	4
第 2 章 Intel IXA 简介	6
2.1 Intel IXA 概述	6
2.2 为什么要使用 IXP2XXX 网络处理器	8
2.3 Intel IXP2XXX 简介	8
2.4 Intel IXA 可移植性框架	12
2.5 基于 IXP2XXX 的典型系统结构	14
2.6 典型的数据包处理流程	15
第 3 章 IXP2400 硬件体系结构	18
3.1 IXP2400 硬件结构概述	18
3.1.1 IXP2400 硬件结构组成	18
3.1.2 IXP2400 硬件结构特征	20
3.2 微引擎	20
3.2.1 概述	20
3.2.2 微引擎的组成部分	21
3.2.3 线程	25
3.2.4 事件信号	26
3.2.5 数据通道寄存器的寻址模式	27
3.3 DRAM 控制器	29

3.3.1 概述	29
3.3.2 DRAM 控制器的特征	30
3.3.3 DRAM 的读/写操作	30
3.4 SRAM 控制器	32
3.4.1 概述	32
3.4.2 SRAM 控制器支持的操作	33
3.5 SHaC 单元	40
3.5.1 概述	40
3.5.2 Scratchpad 操作	41
3.5.3 CAP 操作	43
3.5.4 哈希运算	43
3.6 MSF 的硬件结构	44
3.6.1 概述	44
3.6.2 介质总线接口	46
3.6.3 UTOPIA 协议模式	47
3.6.4 POS-PHY 协议模式	49
3.6.5 CSIX 协议模式	51
3.6.6 MSF 接收	52
3.6.7 MSF 发送	58
3.7 Intel Xscale Core	64
3.8 PCI 控制器	64
3.9 Chassis 单元	65
3.9.1 IXP2400 中的并行处理机制	65
3.9.2 Chassis 的结构	66
3.10 IXP2400 的外围设备	67
3.11 基于 IXP2400 的系统构架方案	69
3.11.1 Dual Ingress/Egress 构架方案	69
3.11.2 Single Ingress/Egress 构架方案	71
3.11.3 Chaining 构架方案	71
3.11.4 一个基于 IXP2400 的系统构架实例	72
第 4 章 基于 IXP2XXX 的网络处理	74
4.1 基于 IXP2XXX 的网络处理的基本知识	74
4.1.1 典型的硬件平台	74

4.1.2	典型的软件结构	75
4.1.3	基本的数据包处理流程	77
4.1.4	数据包缓存	83
4.2	基于 IXP2XXX 的并行处理	85
4.2.1	通用处理器的网络处理缺陷	85
4.2.2	多处理和多线程	85
4.2.3	串行数据流处理的问题	88
4.2.4	分布式缓存	90
4.2.5	小结	94
4.3	IXP2XXX 中数据结构的管理和操作	95
4.3.1	存储器的使用	95
4.3.2	对于 Linked List 的管理	96
4.3.3	SRAM Q-Array	101
4.3.4	Ring Buffer 数据结构	108
4.3.5	小结	110
4.4	微引擎编程模型	111
4.4.1	Mutual Exclusion、Packet Ordering 问题	111
4.4.2	Hyper Task Chaining 模型	112
4.4.3	Pool of Threads 模型	117
4.4.4	HTC 与 POTs 的比较	129
4.5	数据包处理流程的性能分析	130
4.5.1	性能分析中的重要参数	131
4.5.2	IXP2XXX 系统的硬件资源	131
4.5.3	任务/资源利用率的分析	132
4.5.4	性能分析举例	133
4.5.5	性能分析小结	137
第 5 章	IXP2XXX 编程参考	139
5.1	数据包处理软件概述	139
5.2	源文件	140
5.2.1	指令	140
5.2.2	伪指令	140
5.2.3	注释	140
5.3	模块化结构	141

5.4 汇编过程	141
5.5 汇编预处理	142
5.6 常量表达式	143
5.7 运算符	144
5.8 函数	144
5.9 参数、标记列表.....	146
5.10 寄存器和信号.....	146
5.10.1 寄存器.....	146
5.10.2 信号	153
5.10.3 地址运算符.....	156
5.11 IXP2XXX 的伪指令	157
5.12 IXP2XXX 的指令集	168
5.12.1 指令概述.....	168
5.12.2 IXP2XXX 中的 I/O 指令	171
5.12.3 IXP2XXX 的指令详解	172
第 6 章 IXP2XXX 的应用	216
6.1 基于 IXP2400 的 IPv6 路由器设计	216
6.1.1 硬件概述	217
6.1.2 软件概述	218
6.1.3 数据处理流程	218
6.1.4 Dispatch Loop	220
6.1.5 系统性能分析	221
6.1.6 系统资源分配	222
6.1.7 软件模块详细设计	223
6.2 基于 IXP2400 的 DiffServ 系统的设计	264
6.2.1 硬件概述	265
6.2.2 软件概述	266
附录 开发工具 Workbench 的使用	286
参考文献.....	311

1.1 网络处理器的基本概念

1.1.1 网络处理器的产生背景

在过去的十几年中,网络设备的研制开发和因特网的发展相互促进。早期的网络设备(例如网桥和路由器)主要由通用 CPU 和专用标准设备组成,功能依赖于软件设计,因此容易扩展,便于适应网络中出现的新协议和新业务。这种设计的缺点是处理能力低,无法适应网络传输技术的发展和无法满足附加性能的需求。

随着网络规模和接口速度的增加,通用 CPU 对某些网络处理任务尤其是数据接收、交换转发等方面的处理性能难以满足要求,这就促使网络设备厂家开发出了基于 ASIC 的简单且具有固定功能的设备,如第二层以太网交换机。这种 ASIC 设计方案能有效提高网络设备的处理性能和速度,却降低了灵活性。后来在 ASIC 基础上又出现了交换芯片构建方案,如第二层以太网“单片交换”设备。这种方案优化了硬件间的连接,提高了指令执行效率,但仍没有克服 ASIC 缺乏灵活性的固有缺点。

当今网络规模和性能迅速增长,新业务接连不断出现,这就要求网络设备具有线速和智能的处理能力。网络处理器便是适应这种需要的一种新兴、有效的统一解决方案。

1.1.2 网络处理器的定义

网络处理器(NP, Network Processor)是新一代用来执行数据处理和转发的高速可编程处理器。与传统的处理器不同,它的设计采用了全新的理念,使其既有 ASIC(Application Specific Integrated Circuit, 专用集成电路)的高速处理能力,又有完全的可编程特性。由于 NP 在网络数据处理方面的明显优势,它必将成为高速网络设备支持业务管理、安全与网络监控、QoS 等网络功能必不可少的元件。因此,可以说网络处理器代表了未来网络设备设计的发展方向。

从功能上讲,网络处理器主要是完成数据处理和转发任务。例如,对数据分段或重组,对数据帧进行识别,实施流量控制,保证服务质量,进行报文过滤等。

网络处理器可以广泛地应用在因特网的核心层、边沿层和局域网/企业网中。在核心网中,网络处理器可以用于实现核心路由器;在边沿网中,它可以应用于 ISP 接入设备和数据中心,提供对新型业务和网络汇接的支持;在局域网中,则可以用于用户管理和接入控制,实施安全与网络监控,以及 Intranet 防火墙等。

1.1.3 网络处理器的基本组成

网络处理器的组成可以从硬件和软件两方面来描述。

事实上,正是网络处理器在硬件结构和软件组成方面的特殊性,尤其是硬件结构方面采取了与传统处理器不同的策略,才使网络处理器具有上述的各种优点。图 1.1 简要描述了网络处理器的硬件结构。从图中可以看出,网络处理器主要包括两个功能模块,即网络处理单元和专用的智能协处理单元。其中,智能协处理单元是网络处理器的核心,它一般需要嵌入式操作系统的支持,用于对网络处理器单元和其他硬件单元进行控制;通过运行操作系统之上的路由协议软件包,完成路由信息的接收、处理和发送,生成并维护路由表等。网络处理单元(一般有多个)通常采用多线程结构,可以完成高速、大容量的智能数据处理功能,如数据包的收发、包头处理和路由查询等。网络处理器的这种硬件结构,大大提高了其处理速度:多个处理器并行工作的结构可以提高网络处理器的处理能力;处理器内部的硬件多线程结构可以降低数据处理时访问存储器所需的平均时间,提高芯片的效率;实时性要求不同的任务交由不同的处理器来完成,从一定程度上可以减少数据包的处理延迟。从系统灵活性的角度看,智能协处理单元是可编程的,网络处理单元一般由多个 RISC 芯核组成,而 RISC 芯核也是可编程的。这就充分保证了网络处理器的灵活性。

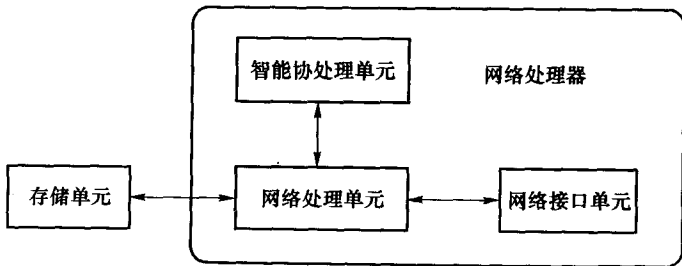


图 1.1 网络处理器的硬件结构

从上面的介绍不难看出,网络处理器的高速处理能力主要是通过一个芯片内集成多个微处理器以及一个微处理器包含多个硬件线程实现的,当然也采用了其他一些硬件加速技术;而网络处理器的灵活性则主要是通过智能协处理单元与网络处理单元的完全可编程能力实现的。

网络处理器的软件包括:板级支持包(BSP, Board Support Packet)、嵌入式操作系统、

路由协议软件包和微代码(或标准语言代码)等4部分。其中前三者运行在智能协处理单元上。板级支持包记录着智能协处理单元需要管理的硬件信息以及它们的主要配置信息;嵌入式操作系统是路由协议或其他应用程序运行的基础;智能协处理单元通过运行路由协议软件包可以生成并维护路由表;微代码运行在网络处理单元上,用于对数据进行处理和转发。

1.2 网络处理器的发展状况

网络处理器既有 ASIC 的处理速度,同时又克服了 ASIC 灵活性差的缺点,具有良好的可编程能力,这两个特点决定了网络处理器具有强大的生命力。从 1999 年到现在,已有包括 AMCC、Intel、Hifn、Motorola 等数家大公司相继推出了自己的网络处理器产品。下面简要介绍一下几家大公司的比较典型的网络处理器产品。

AMCC: AMCC 收购了 MMC 公司的网络处理器生产线,因而也就继承了 MMC 公司网络处理器产品丰富的特点。AMCC 公司的网络处理器产品主要有:10 Gbit/s 的 nP7510、5 Gbit/s 的 nP3700/nP3710、2.5 Gbit/s 的 nP7120/nP7250。由于 Cisco 等大型 OEM 厂商的支持,AMCC 近年来在网络处理器市场上一直处于领先地位,占据了最大的市场份额。

Intel: Intel 公司的网络处理器主要有 IXP4XX 系列、IXP1XXX 系统、IXP2XXX 系列。围绕 IXP2XXX 网络处理器产品,Intel 还开发了 Intel IXA(Intel 因特网体系架构),为基于 Intel 网络处理器的应用开发提供了完整的软硬件解决方案。由于雄厚的技术实力,广阔的传统处理器市场,以及有力的第三方支持,在过去几年中 Intel 已经后来居上,成为网络处理器市场上为数不多的有力的竞争者之一,再假以时日,很有可能优势扩大,成为网络处理器行业的优胜者。本书主要以 IXP2400 为主线,结合 Intel IXA,介绍网络处理器的原理与技术方面的知识。

Hifn: Hifn 公司于 2003 年底收购了 IBM 公司的 Rainier 结构。此后,Hifn 在 IBM 公司网络处理器技术的基础上,基于 Rainier 结构研制出了自己的网络处理器产品 5NP4G。5NP4G 中集成了 16 个 Picoprocessor(网络处理器单元)、1 个 Power PC 智能协处理器单元以及多个硬件加速单元。在 5NP4G 中,2 个 Picoprocessor 共享 1 个硬件加速单元,可取得 2128 MIPS 的计算性能,提供 OC-48 的线速处理能力。5NP4G 可应用于核心路由器、边缘路由器、负载均衡设备、防火墙等场合。

C-Port(Motorola): C-Port 公司的网络处理器产品有 C-3x、C-5x 等。其中,C-5x 处理器集成了 16 个 RISC 芯核,32 个串接数据处理器,可以执行第七层以下的分组分类作业,速度为 5 Gbps。它具有强大的处理能力、标准的编程接口和通用的编程模式(采用标准 C 或 C++)。尤其在进行复杂的数据处理方面有很大的优势。

1.3 Intel 网络处理器

Intel 公司在网络处理器的开发方面走在了同行的前列。该公司开发的网络处理器产品有 IXP12XXX 系列第一代网络处理器,以及以 IXP2400、IXP2800、IXP2850 为代表的第二代网络处理器。

IXP1200 是 Intel 公司推出的第一代网络处理器的典型产品,主要面向 OC-3 到 OC-12数据率的网络应用。它们可以用于:多服务交换、宽带接入平台、协议转换、2/3 层转发、QoS、防火墙、VPN(Virtual Private Network,虚拟专用网)、IDS(Intrusion Detection System,入侵检测系统)和负载均衡等。从硬件结构上讲,IXP1200 包括 6 个完全可编程多线程微引擎和一个 Strong ARM 智能协处理器。这种硬件结构可以最大限度地提高性能,提供综合的网络管理功能;另外,利用 ARM 结构的兼容性可以使编程很容易,支持各种开发工具。该系列产品的开发工具是 IXA SDK 2.0。IXP12XX 系列产品包括 4 种类型:IXP1200、IXP1240、IXP1250 和 IXP1250 温度扩展型,它们的区别如表 1.1 所示。

表 1.1 IXP1200、IXP1240、IXP1250 以及 IXP1250 温度扩展型的比较

网络处理器	IXP1200	IXP1240	IXP1250	IXP1250 温度扩展型
芯片主频	166 MHz、200 MHz 和 232 MHz	166 MHz、200 MHz 和 232 MHz	166 MHz、200 MHz 和 232 MHz	166 MHz
CRC 支持	—	✓	✓	✓
ECC 支持	—	—	✓	✓
工作温度	0~70 ℃	0~70 ℃	0~70 ℃	-45~85 ℃

注: CRC(Cyclic Redundancy Check)是循环冗余码校验, ECC 是纠错码。

随着网络技术的发展应用需求的增长,以 IXP1200 为代表的第一代网络处理器在功能上已经不能很好地满足日益复杂的网络处理功能要求(如: IPSec、VPN、QoS 等);在性能上, IXP1200 只能满足 OC-12 网络线速处理要求,但对于更高的速度则显得力不从心。为了很好地适应当前和今后网络宽带化和业务综合化的趋势, Intel 公司推出了 IXP2XXX 第二代网络处理器。相对于 IXP1200, IXP2XXX 能处理 OC-12~OC-192 速率的网络数据,很好的适应了网络宽带化的趋势;另外, IXP2XXX 充分考虑了网络数据的复杂多样性,采取相应的硬件结构和软件算法有效地进行网络数据处理,很好地适应了网络业务综合化的趋势。

IXP2XXX 是 Intel 第二代网络处理器的统称,包括 IXP2400、IXP2800 和 IXP2850 共 3 款网络处理器产品。其中, IXP2400 主要是面向 OC-12 到 OC-48/2.5 Gbit/s 的网络接入和边缘网的应用。它可以应用于: WAN 多服务交换、DSLAM、2.5 Gbit/s 和 3 Gbit/s 的基站控制器和网关、4~7 层交换、VoIP(Voice over IP,基于 IP 网的语音业务)网关、多服

务接入平台、高端路由器和 VPN 等。从硬件结构上讲,它包括 8 个完全可编程多线程微引擎(主频可以达到 600 MHz),每秒可完成超过 50 亿次的操作,实现 2.5 Gbit/s 的包转发速率。IXP2800 主要是面向 OC-48 到 OC-192/10 Gbit/s 的边缘网和核心网的应用。它可以应用于:骨干网的路由与交换、无线设备、IPSec 和 VPN、10 Gbit/s 的企业交换和路由等。从硬件结构上讲,它里面有 16 个完全可编程多线程微引擎(主频可以达到 1.4 GHz),每秒可完成超过 250 亿次的操作,实现 10 Gbit/s 的包转发速率。IXP2850 在 IXP2800 的基础之上增加了 2 个 crypto 单元,用于加、解密,可达到 10 Gbit/s 的加、解密速度,可用于 IPSec,VPN,QoS Router 等场合。

就目前的情况来看,Intel 公司开发的第二代网络处理器在众多的网络处理器产品中是比较有代表性的,也是比较有优势的。因此,在本书的后续章节中,我们以 Intel 的第二代网络处理器 IXP2400 为基点对 Intel IXA 的结构组成和技术原理进行分析,进而对网络处理器的原理和技术进行介绍。

随着网络本身的发展和业务需求的增长,开发者对新一代的网络处理器提出了更高的要求,主要有如下4个方面:

(1) 灵活性:要求新一代网络处理器能够胜任多种不同的网络应用,满足日益复杂多样的网络业务需求,适应网络业务综合化的趋势;

(2) 高性能:要求新一代网络处理器能以较高的线速(Line Rate)对网络数据进行处理,适应网络宽带化的趋势;

(3) 可扩展性:要求新一代网络处理器在硬件结构上具有可扩展性,能够适应不同速率等级的网络应用;

(4) 可移植性:要求基于新一代网络处理器开发的软件具有良好的可移植性,以节省开发成本、缩短开发周期。

为了满足上述要求,Intel公司推出了Intel IXA(Intel Internet Exchange Architecture,英特尔因特网体系架构),它包括两个方面:在硬件平台方面,Intel公司开发了IXP2XXX第二代网络处理器(包括IXP2400、IXP2800、IXP2850 3款网络处理器产品);在软件平台方面,Intel开发了所谓的IXA Software Portability Framework(IXA软件可移植性框架)。在上述软硬件平台的配合支持下,Intel公司开发的第二代网络处理器IXP2XXX很好地满足了上述要求。

2.1 Intel IXA 概述

如图2.1所示,IXA中有如下3个重要组成部分:

(1) Microengine(ME,微引擎):微引擎是IXP2XXX网络处理器的核心部件,在其他硬件功能单元的支持下,执行快通道(Fast Path,也称为数据面:Data Plane)上的数据包处理任务;

(2) Xscale Core: Xscale Core是IXP2XXX网络处理器的高层控制、管理单元,是与IXP1200中Strong ARM兼容的一种微处理器。Xscale上面运行VxWorks, Embedded Linux等实时操作系统(RTOS),对整个IXP2XXX系统进行控制和管理;

(3) IXA Portability Framework(IXA可移植性框架):IXA Portability Framework为

基于 IXP2XXX 的软件开发提供应用编程接口 (API) 和硬件抽象层 (HAL, Hardware Abstraction Layer), 使得运行在 Xscale Core 和 ME 上的软件具有良好的可移植性和可重用性。

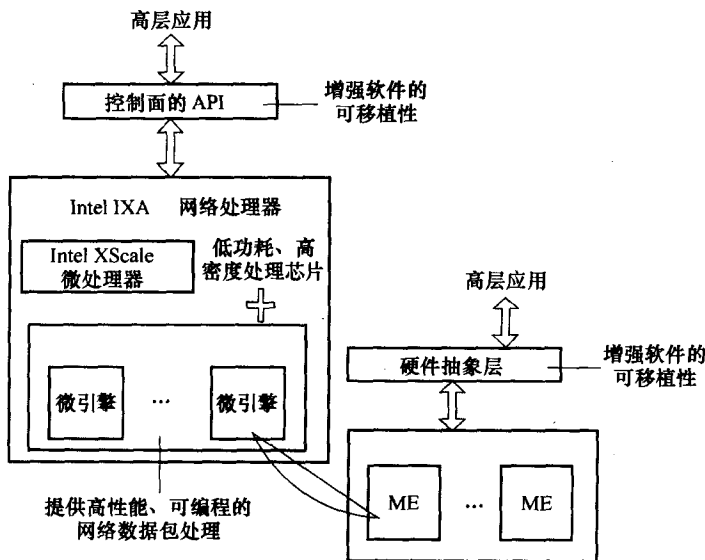


图 2.1 Intel IXA 模型示意图

如上文所述,为了满足对新一代网络处理器提出的 4 方面要求,Intel 开发了 IXA,下面具体论述 IXA 是如何满足这些要求的:

(1) 灵活性:在 IXA 的 IXP2XXX 中,微引擎完全可编程,通过采用不同的编程模型以及相应的协议、算法,在低层硬件设备的支持下,实现许多复杂的网络处理功能,具有较好的功能灵活性;

(2) 可移植性:IXA Portability Framework 提供了 API 和 HAL,使得基于 IXA Portability Framework 开发的软件与底层硬件实现松耦合,具有良好的可移植性及重用性;

(3) 高性能:在 IXA 的 IXP2XXX 中,通过提高 ME、Xscale Core 的时钟频率,采用新技术减小存储器访问时延(Memory Latency),使得 IXP2XXX 能线速处理高达 OC-192 的网络数据,具备了网络宽带化所要求的高性能;

(4) 可扩展性:通过提高 ME、Xscale Core 的时钟频率,以及采用新的微引擎结构 MEV2(第二代微引擎结构),使得 IXP2XXX 网络处理器具有很好的可扩展性,不仅能够处理高达 OC-192 的高端速率的网络数据,还能向下兼容,有效地处理 OC-3、OC-12、OC-48 等中、低端速率的网络数据。

2.2 为什么要使用 IXP2XXX 网络处理器

如上所述,由于技术的发展和需求的增长,开发者对网络处理器提出了更高的要求,但是在第一代网络处理器(如 IXP1200)中,存储器访问时延过大,无法满足高速网络数据的处理要求,因此 IXP1200 只能应用在 OC-12 以及更低速率的场合,对于 OC-48、OC-192 等中高端速率情形就无能为力了。通用处理器的时钟频率较高,能以较快的速度进行数据处理,但是通用处理器不能有效地解决网络处理中经常遇到的串行数据流处理问题(SSPP, Serial Stream Processing Problem),所以不适用于网络数据的处理。在 IXP2XXX 中,通过提高 ME、Xscale Core 的时钟频率以提高处理速度,采取 Multi-Master(多主控设备)和 Multi-Target(多从属设备)的分布式处理结构、Multi-Processing(多微引擎)并行处理结构、Multi-Threading(多线程交换)技术、Distributed Cache(分布式缓存)结构以及其他专用的硬件结构和特殊的软件技术增强并行处理能力,减小存储器访问时延,有效地解决 SSPP,提高串行处理效率,较好地满足了开发者对于网络处理的要求。具体而言,SSPP 可分为两个方面:

(1) 相关情形下的 SSPP:当对前后相关的数据包进行处理时,经常需要访问共同的数据结构。在此情形下,为了避免数据冲突从而导致数据完整性被破坏,需要依次访问该共同的数据结构,串行处理这些前后相关的数据包,从而导致 SSPP。

(2) 不相关情形下的 SSPP:当进行不相关的数据包的处理时,也有可能需要访问共同的数据结构。在此情形下,同样需要依次访问该共同的数据结构,串行处理这些前后不相关的数据包,从而导致 SSPP。

上述两种情形下的 SSPP,导致只能对数据包进行串行处理,使得网络处理器的并行处理优势得不到充分的发挥。因此,为了高速、有效地进行网络处理,必须采用特殊的软硬件机制有效地解决 SSPP。

另外,第一代网络处理器是在假设网络数据包的前后不相关的基础上进行处理而取得相应的性能指标的,但是实际情况不总是简单明了的,网络数据包在某些场合(特别是高速场合)下具有相当的相关性,因此第一代网络处理器仅适用于 OC-12 或更低速率的场合,对于更高速率的应用,第一代网络处理器是不适合的。而第二代网络处理器 IXP2XXX 则充分而全面地考虑了网络数据流的统计特性,并采取了相应的硬件加速设备和软件技术来解决复杂多变的网络数据流所带来的问题,在 OC-48~OC-192 的高速场合也能取得较好的性能指标,较好地适应了网络宽带化的趋势。

2.3 Intel IXP2XXX 简介

IXP2XXX 是 Intel 第二代网络处理器的统称,包括 IXP2400、IXP2800、IXP2850 3 款