第 36 卷 第 3 期 计 算 机 工 程 2010 年 2 月
Vol.36 No.3 Computer Engineering February 2010
·工程应用技术与实现· 文章编号：1000—3428(2010)03—0243—03 文献标识码：A 中图分类号：TP303

网络处理器数据处理内核的设计与实现

李 苗，王叶辉，周彩宝
(华东计算技术研究所，上海 200233)

摘 要：通用处理器和专用芯片 ASIC 的数据处理能力无法满足日益增长的网络带宽和各种复杂网络协议的要求，针对该问题，研究网络处理器的系统结构，讨论网络处理器中数据处理内核的设计实现，提出一种可编程的精简指令集计算机(RISC)处理器微结构，对其进行现场可编程门阵列(FPGA)原型验证，结果证明了该设计方案的有效性。

关键词：网络处理器；数据处理内核；流水线

Design and Realization of Data Processing Core in Network Processor

LI Miao, WANG Ye-hui, ZHOU Cai-bao
(East China Institute of Computer Technology, Shanghai 200233)

【Abstract】Data processing capability of General Purpose Processor(GPP) and Application Specific Integrated Circuit(ASIC) can not meet the requirements of increasing network bandwidth and complex protocols, so that this paper researches the architecture of Network Processor(NP), discusses the design and realization of data processor core in network processor, and proposes a programmable Reduced Instruction Set Computer(RISC) micro-architecture of processor. It demonstrates the feasibility of design scheme through Field Programmable Gate Array(FPGA) prototype.

【Key words】Network Processor(NP); data processing core; pipeline

由于网络用户和数据流量呈指数级增长，因此要求网络处理的带宽更宽、速度更快。同时，网络上出现了更为复杂的协议，要求网络处理报文的能力更强、更灵活。通用处理器(General Purpose Processor, GPP)的成本低，但是性能不足以处理高速网络流量，专用芯片 ASIC(Application Specific Integrated Circuit)的设计灵活性差，这就要求由处理性能更高的网络处理器(Network Processor, NP)实现数据报文的交换转发。网络处理器是面向网络应用的专用指令处理器(Application Specific Instruction Processor, ASIP)，其内部是由若干数据处理内核和协处理器组成的并行结构，能够对报文进行并行的深度处理，从而实现硬件加速。网络处理器提供专用的微指令系统，用户可以根据不同的网络应用，开发不同的微处理程序，以灵活的软件体系提高硬件的处理性能[1]。

1 网络处理器的硬件体系结构

网络处理器通常采用多内核并行处理器的体系结构，控制平面和数据平面通过一组通信协议进行通信，相互配合完成网络节点的处理任务。目前典型的网络处理器产品有 Intel 公司的 IXP2XXX 系列、IBM 的 PowerNP 等。

1.1 网络处理器的组成

网络处理器一般集成了一个片内通用处理器核、多个数据报文处理内核、多个专用的协处理器以及各种接口模块[2]。网络处理器的结构如图 1 所示。其中，通用处理器核处理非实时管理任务，包括操作系统的运行、模块的初始化和异常数据报文的处理等；数据处理内核负责实时、高速的数据报文处理，包括分组头有效性检验、路由表查找和 TTL 修改等；专用协处理器完成报文处理中执行频度较高、处理较复杂的功能，如 CRC 校验、Hash 运算；接口模块则包括 SRAM 和 DRAM 存储器接口、网络媒介接口和 PCI 接口等。

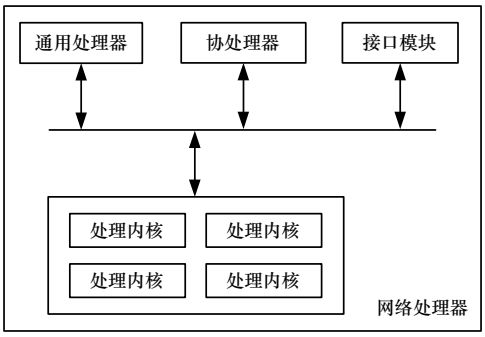

图 1 网络处理器结构

1.2 网络处理器中的并行机制

为了达到千兆甚至万兆的数据报文线速处理能力，网络处理器大多采用多层次并行机制，包括处理内核之间、处理内核与协处理器或接口模块之间、处理内核内部的并行。

多个处理内核的配置方法有串行和并行 2 种。在串行配置中，多个处理内核串成一条流水线，每个内核负责报文处理的一部分。在并行配置中，每个处理内核的功能相同，并行地对不同的数据报文进行处理。

处理内核与协处理器或接口模块之间的通信方式一般有同步和异步 2 种。在同步方式下，处理内核与协处理器或接口模块是串行工作的，协处理器或接口模块执行完指令之后，处理内核才执行下一条指令。而在异步方式下，处理内核不必等待协处理器或接口模块结果就可以执行下一条指令。

作者简介：李 苗(1979－)，女，工程师，主研方向：计算机体系结构，数字系统设计；王叶辉，工程师；周彩宝，研究员

收稿日期：2009-06-02 **E-mail：**limiao111@sohu.com

处理内核可以采用不同的微结构，如精简指令集计算机(Reduced Instruction Set Computer, RISC)、超长指令字 VLIW。RISC 技术是当前应用最广泛的技术，其中，单标量结构每周期发射一条指令，结构简单，而超标量结构一次可以向执行部件发射多条指令，提高指令执行的并行度。VLIW 结构将多条不相干的指令构成一组，同时发射执行，由于编译器确保同组的指令可以并行执行，因此降低了硬件复杂度。

由于网络处理器通常将接收到的数据报文存储在外部存储器中，或者将报文交给协处理器处理，因此为了隐藏访问外部存储器和协处理器的延时，处理内核一般采用多线程机制[3]，实现多个任务的并发处理。理想情况下的四线程切换如图 2 所示。

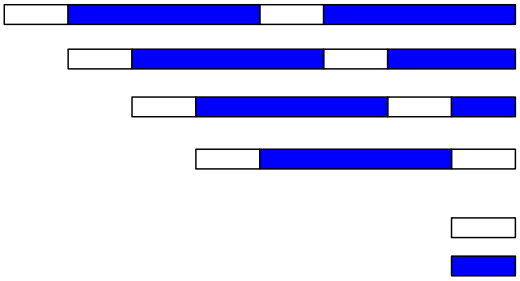

图 2　线程切换示意图

2　数据处理内核的设计

本文设计的网络处理器包括通用处理器核、4 个数据处理内核、Hash 运算协处理器、DRAM 和 SRAM 控制器和支持 SPI3 协议的网络媒介接口等。数据处理内核在硬件上支持 4 个线程，可隐藏访问外部存储器和协处理器的延迟，提高系统的性能。本文重点讨论数据处理内核的硬件设计与实现。

2.1　数据处理内核的结构

本文设计的数据处理内核是可编程的 RISC 处理器核，其内核结构如图 3 所示。

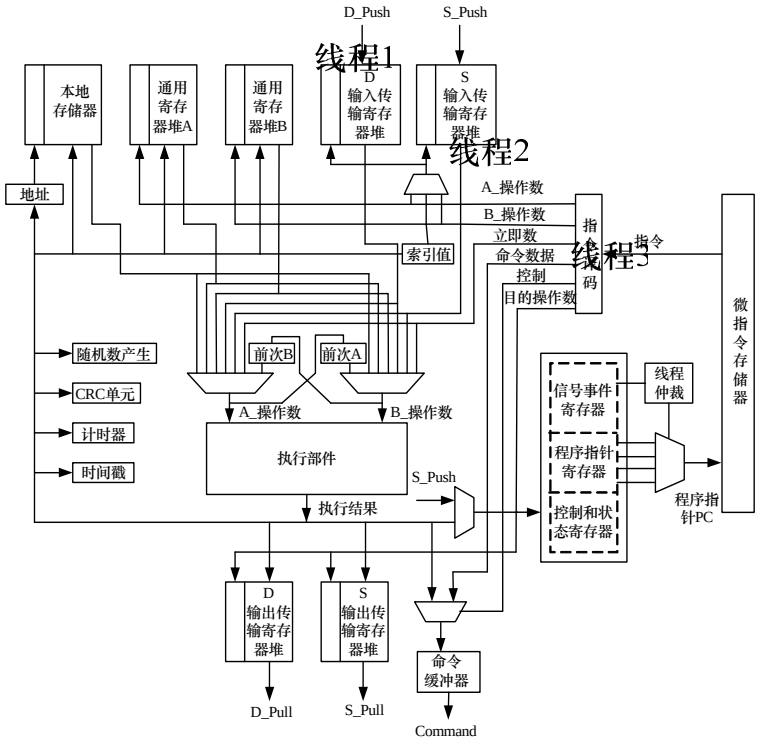

图 3　数据处理内核结构

数据处理内核的主要部件有：

(1)微指令存储器：通用处理器核将报文处理程序下载到该存储器中，内部程序指针控制指令的单周期读取。

(2)指令译码部件：将读取的微指令翻译成内部的数据和控制信息。

(3)数据通路寄存器：用于存放指令源、目的操作数，包括本地存储器、通用寄存器堆、DRAM 输入传输寄存器堆、SRAM 输入传输寄存器堆和命令缓冲器。

(4)执行部件：根据指令译码的结果，完成源操作数的不同操作。

(5)控制和状态寄存器：程序员配置处理内核或者读取处理内核的状态。

(6)5 套分立的总线：包括 D_push, S_push, D_pull, S_pull 和 Command 总线。

2.2　数据处理内核的指令系统

本文设计的指令系统参考了国际主流网络处理器的指令系统，主要包括通用指令、分支跳转指令、线程切换指令和外设访问指令 4 类。

通用指令实现移位、加减法、逻辑、寻找第 1 个“1”、乘法等操作。针对网络数据处理的特点，还有用于快速查找的 CAM 指令、用于校验数据的 CRC 指令和用于数据对准的字节对齐指令。

分支跳转指令可以无条件转移、根据位、字节、线程号或者信号量的比较结果转移、根据条件码或者指定标号转移。为了减少因指令转移而浪费的时钟周期，可以在转移指令的延迟槽(slot)中插入若干不相关的指令。

线程切换指令可以实现处理内核中 4 个硬件线程的切换。当前线程执行线程切换指令后，就交出执行控制权，进入睡眠状态，当指定的信号量到来时会被唤醒。

处理内核不处理外设访问指令，而是将该指令通过总线发送给外设(指网络处理器内部的协处理器、存储控制器和介质访问控制器)，同时将线程切换，外设完成操作后会向处理内核返回信号量，依据信号量寄存器值唤醒线程。

2.3　流水线

为了使各个流水级路径延迟基本均衡，本设计中的处理内核采用单标量流水线结构，流水线划分为 6 级，如图 4 所示。

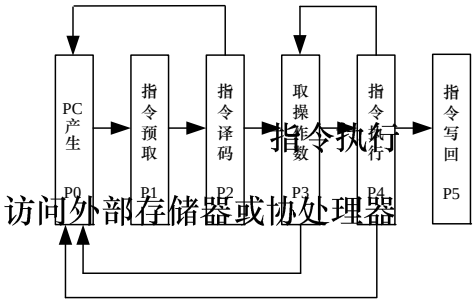

图 4　数据处理内核流水线

本设计中处理内核的 6 个功能模块分别对应于 6 级流水线：

(1)PC 产生模块 P0

该模块产生指令预取的程序指针 PC 值。处理内核正常执行时，PC 值每周期加 1，遇

到 P2 决策的分支和线程切换指令、P3 决策的条件码分支指令、P4 决策的条件分支和跳转指令时，PC 值为相应的跳转目标地址。如果指令处于分支指令的延迟槽中并且被分支指令中止(abort)，PC 值保持不变。

(2)指令预取模块 P1

该模块包含了一个 4 096 字的微指令存储器，通用处理器核可以通过一组寄存器对微指令存储器进行指令的装载。处理内核工作时，从微指令存储器中读取指令，如果没有被分支指令中止，就将预取的指令寄存输出到 P2，否则，输出空指令。

(3)指令译码模块 P2

该模块对 P1 预取的指令进行译码，从指令中解析出源操作数信息、目的操作数信息和操作控制信息，输出到 P3。如果解析出分支指令，则通知 P0 产生新的 PC 值。如果解析出线程切换指令，通知 P0 产生新的 PC 值，将 PC 值和信号量等保存在相关寄存器中，并用轮转算法仲裁出下一个要执行的线程号。

(4)取操作数模块 P3

该模块使用的存储资源有输入传输寄存器堆、通用寄存器堆和本地存储器，根据 P2 级产生的源操作数信息从存储资源中取出 32 位的 A 和 B 操作数。处理 P3 需要决策的条件码转移指令，通知 P0 产生新的 PC 值。如果检测出当前指令的源操作数是前一条指令的目的操作数，即遇到写后读(RAW)相关，则选择 P4 旁路到 P3 的数据作为源操作数。此外，通用处理器核可以通过网络处理器内部的访问代理模块从 S_Push 总线向控制寄存器写数据，DRAM 和 SRAM 控制器可以分别通过 D_Push 和 S_Push 总线向输入传输寄存器堆写数据。

(5)指令执行模块 P4

该模块实现数据内核的数据通路，完成指令的执行，向 P5 输出执行结果，产生条件码(包括负、零、溢出和进位标志)。负责分支跳转地址和外设访问指令中外设地址的产生，并通知 P0 产生新的 PC 值。

(6)指令写回模块 P5

该模块将执行完成的指令结果写回输出传输寄存器堆、通用寄存器堆和本地存储器中，组装外设访问命令的控制信号，以命令的方式存入命令 FIFO，通过 Command 总线向外设传送命令。此外，通用处理器核可以通过网络处理器内部的访问代理模块从 S_Pull 总线读取控制和状态寄存器的数据，DRAM 和 SRAM 控制器可以分别通过 D_Pull 和 S_Pull 总线从输出传输寄存器堆读取数据。

3 设计验证

本文用 Verilog 语言对数据处理内核进行了寄存器传输级的描述，采用软平台把通用处理器核、协处理器以及外围接口集成在一起，同时编写覆盖所有指令的测试程序，经过功能验证，指令执行的结果以及程序处理的能力都达到设计要求。此外，选用了网络处理器的测试基准工程(IPv4 转发工程)作为应用实例，在该网络处理器上成功运行了 IPv4 转发工程。

之后将设计代码下载到 Xilinx 公司的 Virtex4 芯片进行现场可编程门阵列(Field Programmable Gate Array, FPGA)原型验证。由于网络处理器的规模比较大，因此采用 2 片 Virtex4 芯片级联，一片中装载 4 个数据处理内核、网络媒介接口、总线仲裁器等设计，由于 Virtex4 芯片中内嵌通用处理器核 PowerPC，因此还需要装载总线转接桥；另一片中装载了 DRAM 和 SRAM 控制器、Hash 运算协处理器、寄存器访问代理等设计。FPGA 原型运行的频率为 50 MHz，将 IPv4 转发程序下载到处理内核的微指令存储器中，通过转接芯片实现 SPI3 接口到以太网 MII 接口的转换，最终可以实现 4 个千兆以太网端口之间 IPv4 数据报文的转发。

4 结束语

本文主要研究了网络处理器的硬件体系结构，重点讨论了数据处理内核的设计实现，提出了一种可编程的 RISC 微处理器结构，并且采用 6 级流水线实现。将 4 个数据处理内核、DRAM 和 SRAM 控制器、Hash 运算协处理器、SPI3 媒介接口等模块集成在一起，采用 IPv4 数据报文转发工程对原型系统进行了验证。软平台功能验证和 FPGA 原型验证的结果证实了该方案的可行性。

参考文献

[1] 石晶林，程　胜，孙江明. 网络处理器原理、设计与应用[M]. 北京: 清华大学出版社, 2003.

[2] 李　诚，李华伟. 网络处理器中处理单元的设计与实现[J]. 计算机工程, 2007, 33(2): 253-254.

[3] 张宏科，苏　伟，武　勇. 网络处理器原理与技术[M]. 北京: 北京邮电大学出版社, 2004.

编辑　张　帆

（上接第 242 页）

参考文献

[1] 宫长荣，潘建斌，宋朝鹏. 我国烟叶烘烤设备的演变与研究进展[J]. 烟草科技, 2005, (11): 34-37.

[2] 贺桂芳. 基于 SHT11 的温湿度无线测控系统设计[J]. 微计算机信息, 2007, 23(23): 307-308.

[3] 周文举. PC 串口与多个单片机红外无线通信的实现[J]. 工业控制计算机, 2004, 17(7): 29-31.

[4] Santos M S M, Freire R C S, da Silva J F. Wireless Data Acquisition System for Remote Care of Newly Born Prematures[C]//Proc. of IEEE International Workshop on Medical Measurement and Applications. [S. l.]: IEEE Press, 2006: 28-32.

[5] 龚永坚，盛法生，陈　霓. 基于无线传输的轮胎气压监测系统设计[J]. 农业机械学报, 2005, 36(6): 79-81.

编辑　张　帆