

并行处理的 IPv6 控制路由器快路径结构实现

王 光

(西安文理学院, 陕西 西安 710065)

摘 要: 提出了一种适用于 IPv6 路由器的快路径结构, 采用 NoC 结构作为并行处理的基本通信设施, 利用其可扩展性及高通信效率的特点, 设计实现了包括预处理、输入队列、数据包处理和片上网络四部分的快路径结构, 从而实现了分布式 IP 地址查找与并行 IP 包交换相融合的并行 IP 报文处理. 测试表明该结构解决了 IP 路由查找方面的可扩展性差、查找延时大等瓶颈问题, 而且在一个 NoC 结构之上同时实现了并行 IP 包交换, 极大提高交换效率, 节省硬件资源.

关键词: IPv6 路由器; 片上网络; 并行处理

中图分类号: TN402

文献标识码: A

文章编号: 1000-7180(2014)10-0134-04

A Fast-Path Architecture for IPv6 Control Router

WANG Guang

(Xi'an University of Arts and Science, Xi'an 710065, China)

Abstract: A fast-path architecture of IPv6 router suitable for parallel processing is proposed based on the NoC structure. The fast-path architecture of IPv6 router is designed and implemented using NoC structure. Thereby, a parallel processing structure of IP packets combining the distributed IP address lookup and the parallel IP packet exchange is implemented. It is convenient to replace the search algorithm and to expand the hardware scale, and it can improve the overall performance of the router effectively.

Key words: IPv6 router; Network-on-Chip; Parallel processing

1 引言

随着网络技术和网络应用的不断进步, 人们所需要的数据传输数量越来越大. 尤其是 IPTV、移动语音、P2P 等近些年快速增长, 直接导致 Internet 骨干网带宽承受着前所未有的压力. 以 IPv4 协议为核心技术的互联网正面临越来越严峻的技术挑战^[1]. 为了解决 IPv4 所存在的一些问题和不足, IETF 组织(Internet Engineering Task Force)设计并发布了面向下一代互联网数据包传送格式 IPv6 协议, 从而在路由、自动配置等方面进行了改进.

IP 路由器(Router)是一种“专用”计算机, 是计算机网络的核 心设备, 在网络通信中完成路由

(Routing)和转发(Forwarding)任务^[2-3].

尽管 IPv6 最终会完全取代 IPv4 在互连网上占据统治地位, 但 IPv4 和 IPv6 两个协议必然在很长时期内共存. 目前应用的 IP 路由器主要是基于 IPv4 的, 而已有的 IPv4 路由器体系结构难以解决路由器所面临的容量、扩展性和功耗等问题. 所以, 迫切需要新一代可持续发展的超大容量路由器.

本文引入 Mesh 技术, 提出了一种新的基于 IPV6 路由器设计.

2 关键技术

(1)路由查找算法

收稿日期: 2014-01-20; 修回日期: 2014-03-12

基金项目: 西安市科技计划项目(CXY1352WL35)

IP 地址查找问题就一直是学术界和工业界研究的重要课题,由此也产生了众多被广泛应用于互联网的路由器的经典算法^[4-5].当前关于 IP 地址查找的研究重点逐步转移到利用并行计算技术来提高查找效率,即并行 IP 路由查找.主要思路是借助 N 个查找引擎协同工作,从而实现结构和性能的扩展.

(2) 交换结构

近年来,基于输入输出队列的单级交换结构是大多数商业千兆比特级路由器采用的结构^[6].但这种交换结构面临存储器带宽瓶颈、吞吐率保证、QoS (Quality of Service) 保证以及可扩展性等系列问题不能很好解决,必须从高性能交换体系结构上寻求创新和突破.

(3) Mesh 技术

Mesh 是片上网络 (Network-on-Chip, NoC) 的一种基本形式,可以解决以总线结构为通信基础的 SoC 技术所面临的性能、功耗、延迟和可靠性等问题^[7-8].NoC 借鉴通信网络和并行计算的思想,采用包交换和路由的方法代替传统的总线,实现处理单元与通信结构分离^[9].

3 解决方案

适用于并行处理的 IPv6 控制路由快路径结构如图 1 所示,主要包含预处理模块、数据包处理模块以及片上网络等部分.四路并行的数据通路是由相同的模块实例化而来.另外还包括系统的寄存器管理部分.

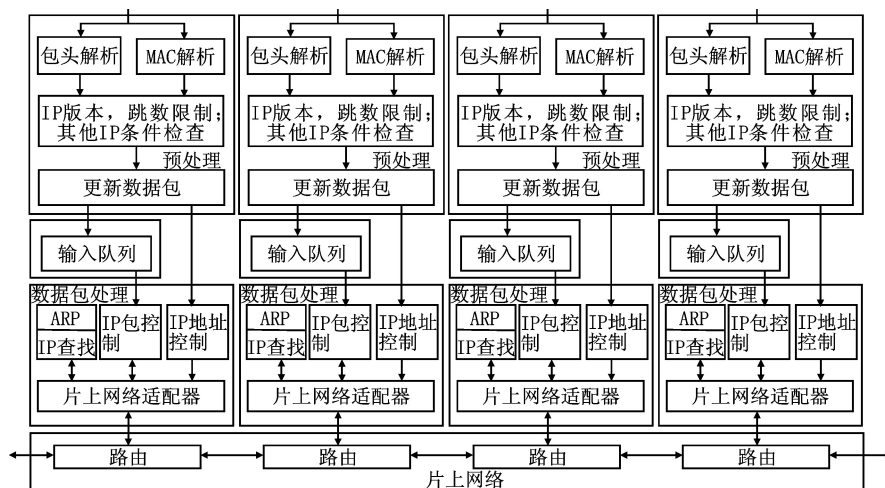


图 1 IPv6 快路径结构框图

路由器的性能主要取决于路由器快路径设计的优劣.它是整个设计中最核心的模块.该结构的特点在于采用了片上网络的思想来并行交换数据包,可以同时完成 IP 报文的交换与分布式 IP 地址查找.

本设计是在 MAC 层完成数据链路层到网络层数据转换的基础上进行的.系统数据字宽度为 64 位,MAC 层需要将携带有 IP 报文的 MAC 帧转换为数据宽度为 64 位的数据格式.另外在模块间进行数据传输时加上了与数据信号相匹配的控制信号与握手信号.模块发送端的数据信号如表 1 所示,信号之间的时序关系如图 3 所示.

表 1 模块间数据传输信号

信号名称	输入/输出	信号描述
Data	输出	64 位的数据信号
Ctrl	输出	8 位的控制信号,用来指示当前数据信号的内容
Wr	输出	写信号,将当前 data 和 ctrl 信号写入下一级模块,高有效.
Rdy	输入	下级模块向本级模块发送的指示信号,高有效.

Data 信号与 Ctrl 信号作为要传输到下一级的数据,在 Wr 为高电平时写入下一级模块. Rdy 信号为高电平时表示下一级模块可以接收数据,因此,在发送端需要先检查 Rdy 信号是否为高电平,只有当 Rdy 信号为高电平时发送端才将数据输出.

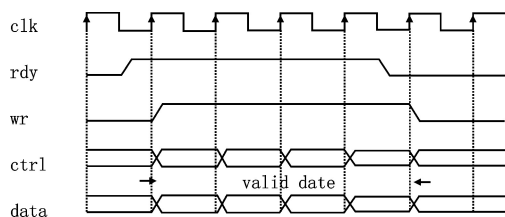


图2 模块间数据传输信号时序图

4 验证

4.1 验证平台结构

验证平台分为软件和硬件两部分。软件部分主要根据路由表信息完成 IP 数据包头和查找结果的生成、验证平台的配置、功能检查及性能统计和待验证设计(如路由器交换设计、转发表设计等)管理。硬件部分对待验证设计进行输入数据控制、输出管理和检查待验证系统功能及统计性能信息。整个平台的软硬件两部分通过 PCI 接口相连接,结构如图4所示。

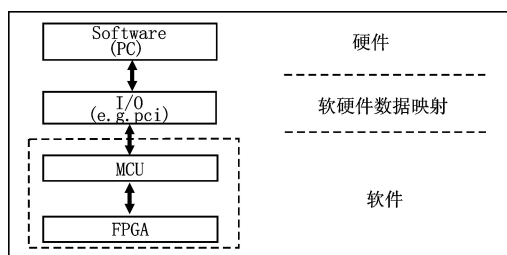


图3 路由器验证平台结构

4.2 系统测试方案

系统测试采用 Xilinx ML605 开发板。该开发板所搭载的 FPGA 芯片型号为 Virtex-6 XC6VLX240T-1FFG1156。使用 Xilinx ISE 12.4 进行 FPGA 综合,时钟频率达到 151 MHz。因为在该 FPGA 中,DDR3 接口支持的最大时钟频率为 400 MHz,所以在进行 FPGA 测试时最大支持 100 MHz 的时钟频率。

测试的数据主要包括:

- (1)数据包的转发是否正确;
- (2)发送不同长度数据包与不同数据包注入速率的情况下系统的丢包率;
- (3)发送不同长度数据包与不同数据包注入速率的情况下系统的延迟。

测试时采用线速 1—25 Gbit/s 流量输入;DIX Ethernet V2 标准规定输入数据包长在 46~1 500B

之间;对四个 NoC 节点和输出端口都采用均匀分布的流量模型。

4.3 系统测试结果

经测试,路由器的 IP 地址查找和数据转发功能正确。

路由器在固定包长(1 kB)情况下丢包率与线速的关系如图4所示。

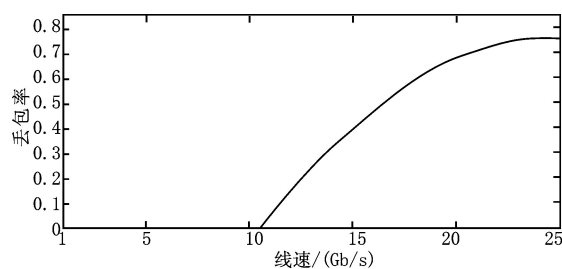


图4 丢包率与线速的关系

从图4可以看出,系统的输入线速在低于10.3 Gb/s 时丢包率非常小;当超过 10.3 Gb/s 时由于查找延迟与 NoC 传输的延迟导致系统交换的能力趋于饱和,丢包率迅速增大,并且随着线速的进一步增加,丢包率也迅速增加。

在固定线速下(10 Gbit/s)丢包率与数据包长的关系如图5所示。数据包的包长与平均输出延时的关系如图6所示。

可见,系统正常工作时丢包率与数据包的长度相关,数据包越长,占用存储队列越多,且占用交换资源与输出端口时间也越长,因此丢包率也就越大。

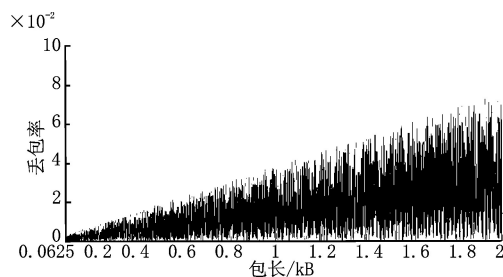


图5 丢包率与包长的关系

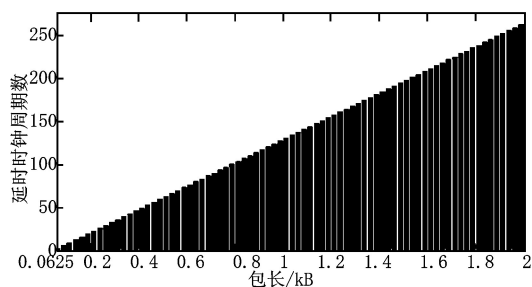


图6 数据包平均输出延时与包长的关系

5 结束语

本文提出一种适用于并行处理的 IPv6 控制路由器快路径结构. 该结构采用 NoC 结构作为并行处理的基本通信设施, 利用其可扩展性及高通信效率的特点, 设计实现了包括预处理、输入队列、数据包处理和片上网络四部分, 从而实现了分布式 IP 地址查找与并行 IP 包交换相融合的并行 IP 报文处理. 经测试, 该结构一方面解决了 IP 路由查找方面的可扩展性差、查找延时大等瓶颈问题, 另一方面, 在一个 NoC 结构之上同时实现了并行 IP 包交换, 极大提高了交换效率, 节省了硬件资源.

下一步将研究快速高效且资源消耗低的 IPv6 硬件查找算法, 以替代传统的 IPv4 硬件查找算法.

参考文献:

- [1] Chao H J, Liu B. High performance switches and routers[M]. Boston: Wiley-Interscience, 2007.
- [2] Li Q, Jinmei T, Shima K. IPv6 core protocols implementation the morgan kaufmann series in networking [M]. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. 2006.
- [3] Baker, F. Requirements for IP version 4 routers. RFC 1812. [2014-01-20]. <https://tools.ietf.org/html/rfc1812>. June 1995.
- [4] Knuth D E, The art of computer programming[M]. Sorting and Searching. Reading, MA: Addison-Wesley, 1973.
- [5] Lampson B, Srinivasan V, Varghese G. IP lookups using multiway and multicolumn search [J]. IEEE/ACM Transactions on Networking (TON), 1999, 7 (3): 324-334.
- [6] Minkenberg C. Current issues in packet switch design [J]. ACM SigComm Comput Comm Rev Jan, 2003, 33(1):1-6.
- [7] W. Dally and B. Towles. Route packets, not wires: on-chip interconnection networks[C]// Proc the Design Automation Conference, Las Vegas, NV: IEEE, 2001:684-689.
- [8] De G Micheli, Benini L. Eds networks on chips: technology and tools[M]. San Mateo, CA: Morgan Kaufmann, 2006.
- [9] Hu J, Marculescu R. Energy and performance-aware mapping for regular NoC architectures [J]. IEEE Trans on Computer-Aided Design of Integrated Circuits and Systems, 2005, 24(4):551-562.
- [6] Lin Hsuan-Tien, Lin Chih-Jen, A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods[R]. Taipei, Taiwan, 2003.
- [7] 谢承旺. 不同种类支持向量机算法的比较研究[J]. 小型微型计算机系统, 2008, 1(1):98-108.
- [8] 李敏, 李俊. 基于人类视觉系统特性的图像质量评价算法[J]. 科技通报, 2013, 29(2): 160-162.
- [9] Lowe D. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91-121.
- [10] 马建红, 姬莉霞. 基于免疫多 Agent 的网络监控系统模型研究[J]. 计算机仿真, 2013, 30(5): 213-216.

作者简介:

王 光 男, (1973—), 副教授. 研究方向为计算机系统结构.

作者简介:

徐 甜 女, (1968—), 硕士研究生, 副教授. 主要研究为图像处理、数据挖掘.

刘凌霞 女, (1977—), 硕士, 副教授. 研究方向为图像处理.