

基于多核的 IPv4/IPv6 过渡技术研究是实现

杨志义, 李晓燕

(西北工业大学 计算机学院, 西安 710072)

(shuixian00@126.com)

摘要:针对目前 IPv4/IPv6 过渡时期存在的问题,结合目前主要过渡技术的不足,提出基于多核的 IPv6 隧道技术,设计了多核下 IPv6 隧道模型。利用多核流技术、锁技术和多核负载分担方法,加速转发处理、保护多核临界资源和提高系统整体性能。仿真实验结果表明,IPv6 隧道技术基于多核比基于单核的实现,在转发与处理性能上成倍提高,能更好地解决过渡时期存在的问题与适应未来的发展。

关键词:多核;IPv6 过渡技术;IPv6 隧道技术

中图分类号:TP393.01 **文献标志码:**A

Study and implementation of IPv4/ IPv6 transition technology based on multi-core

YANG Zhi-yi, LI Xiao-yan

(School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an Shaanxi 710072, China)

Abstract: According to the problem in the transitional period of IPv4/ IPv6 and the shortages of the mainly techniques during this period, this paper proposed the IPv6 tunnel technology based on multi-core, designed its model, used the multi-core flow technology, the lock-technique and the multi-core load-share model to speed up transmitting, protected the critical resources and improved the whole performance. Finally, the IPv6 tunnel technology based on multi-core is proved to have better transition and service performance than that based on single-core.

Key words: multi-core; IPv6 transition technology; IPv6 tunnel technology

随着互联网技术的不断发展,IPv4 在地址空间、QoS 服务等方面日益不能满足需求,IPv6 以其巨大的地址空间、良好的 QoS 保证、便利的网络管理以及良好的移动性支持等多项技术优势可以彻底解决 IPv4 所存在问题^[1]。然而现有大部分骨干网络应用都建立在 IPv4 协议上,要实现到 IPv6 的全面转换是一个长期渐进的过程,因此 IPv4 和 IPv6 并存现象将会持续很长的一段时间^[2],于是就出现了 IPv4 海洋与 IPv6 孤岛之间以及 IPv6 孤岛之间的通信问题,如何在尽可能减少过渡成本^[4]的前提下渐进地、无伤害地、高效地解决过渡期节点的通信问题是目前研究的重点。

1 IPv4/IPv6 过渡阶段主流技术及其不足

1.1 IPv4 向 IPv6 过渡的三个阶段

根据 IPv4 与 IPv6 网络的发展,可以将过渡时期分为三个阶段。IPv4 网络占主导地位的初始阶段;IPv6 得到较大规模的应用和发展,并通过 IPv4 网连接实现通信的 IPv6 骨干网络的共存阶段;IPv6 网络占主导地位阶段。

1.2 IPv4/IPv6 过渡时期的三种技术

为解决 IPv4/IPv6 过渡时期的问题,实现 IPv4 向 IPv6 的高效无缝转换,NGTrans 和 IETF 都对此问题进行了研究且已经提出了多种 IPv4/IPv6 过渡技术。基本过渡技术中成熟的技术包括双协议栈技术、隧道技术和协议转换三种^[2],用于解决不同过渡阶段、不同环境的通信问题。

1) 双协议栈。双协议栈技术是一台设备上同时启用 IPv4/ IPv6 协议栈^[3]。它是所有过渡技术的基础,三个阶段均适用,主要用于小型企业网络,易于部署,可同时提供 IPv4 应用和 IPv6 应用。

2) 隧道技术。隧道技术是一种封装技术,它是利用一种网络协议来透明传输另一种网络协议的虚拟的点对点的连接。主要用于第一、三阶段,且它只要求 IPv4/IPv6 网络的边缘设备支持双协议栈,可以更好地利用现有 IPv4 网络,减少投资^[3]。目前 6_Bone 采用的就是隧道技术。

3) 协议转换技术。协议转换技术是通过 IPv4 内部地址向 IPv6 全局地址或 IPv6 内部地址向 IPv4 全局地址^[2]的转换来实现 IPv4/IPv6 网络互通的。目前使用广泛的是 NAT_PT。只需支持 NAT_PT 的设备和 IPv6 DNS 解析,无需双协议栈的支持,且终端系统不需要升级,部署比较简单。

1.3 当前主流技术在应用中的不足

由于双协议栈要求路由器同时支持 IPv4/IPv6 协议,对于大型网络不利于升级与扩展,不能解决 IPv4 地址短缺问题^[3];协议转换技术存在单点失效问题,且在 IPv4 和 IPv6 报文之间进行转换也影响转发效率,而其他如在应用层进行验证、加密的应用几乎无法利用这一技术实现^[2],因此隧道技术被广泛采纳。

2 基于多核的 IPv6 隧道技术及设计方案

随着需求的不断扩大,网络承载的业务类型也更为复杂,单核处理器的处理速度和性能不能满足应用的需要,使得隧道强大功能的发挥受限于 CPU 处理能力。尽管有少部分采用 ASIC 或 NP 作处理器,但因 ASIC 扩展性差,NP 软件开发复杂且成本较高,也不是很理想的选择^[5]。多核以其良好的可扩展性、可重用性、软件开发简单和性价比高,已成为高性能微处理器体系结构的发展趋势和诸多工业应用的最佳选择。鉴于以上原因,给出了基于多核的 IPv6 隧道技术(6in4)

收稿日期:2008-09-26;修回日期:2008-11-25。

作者简介:杨志义(1952-),男,陕西西安人,教授,主要研究方向:网络化嵌入式计算、分布式计算; 李晓燕(1983-),女,河南三门峡人,硕士研究生,主要研究方向:测控技术、嵌入式计算。

的设计方案与实现模型,如图1所示。

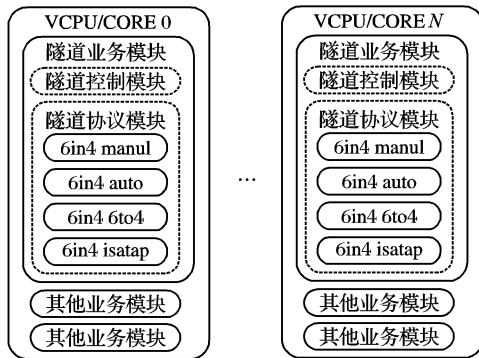


图1 多核隧道模块模型

2.1 采用 6in4 隧道技术的四种隧道模式

1) 手工配置隧道(6 in 4 manual)。该模式的隧道端口地址需要在隧道两端的双协议路由器或双协议主机上手动配置。隧道的源接口地址与目的接口地址必须是全球唯一的 IPv4 地址且具有有效的路由。

2) 自动配置隧道(6 in 4 auto)。该隧道模式不用配置隧道的目的端口地址,但站点/主机必须拥有全球唯一的 IPv4 地址且用 IPv4 兼容的 IPv6 地址(0::[IPv4 地址]/96)作为其地址。隧道目的端口的 IPv4 地址根据通过其的报文的 IPv4 兼容 IPv6 地址直接产生。

3) 6to4 隧道(6 in 4 6 to 4)。该隧道模式不用配置隧道的目的端口地址。它采用前缀为 2002::/16 的特殊 IPv6 地址。拥有全球唯一 IPv4 地址的站点或主机可拥有相应 IPv6 地址(2002:[IPv4 地址]::/48)。隧道目的端口的 IPv4 地址根据通过其的报文的 IPv6 地址直接产生。

4) 6over4 隧道(6 in 4 isatap)。该隧道模式不用配置隧道的目的端口地址,与 6to4 不同的是,6over4 利用 IPv4 的多播机制来实现虚拟链路,它采用前缀 ::0:5EFE:[IPv4 地址]/96 的 IPv6 地址,隧道目的端口的 IPv4 地址通过报文的 IPv6 地址直接产生。

2.2 设计方案

对于多核方式下隧道的设计与实现,需要以多核处理方式为前提,尽最大可能发挥多核的优势,提高隧道业务处理的性能。

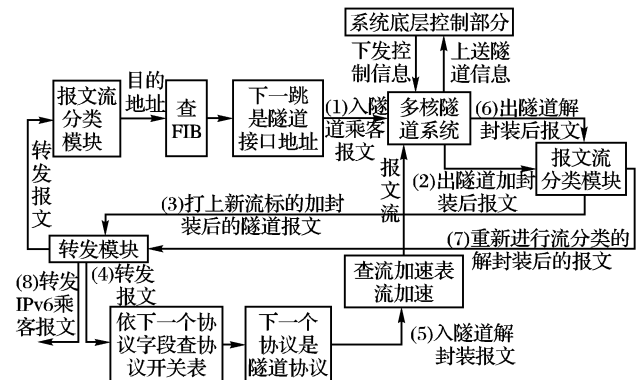
采用多核负载分担模型与单一系统映像机制。每个 VCPU/CORE 上均映射系统中相同的隧道模块代码,抽象为多个处理单元,如图1所示,实现各 VCPU/CORE 并行化处理的同时也便于实现负载分担,提高系统整体性能。

多核流技术是本文提高转发处理性能的关键。多核的引入使得报文能正确高效地到达每个 VCPU/CORE,并接受业务处理。单核报文处理流程是基于业务匹配策略,每个包都要通过查 FIB 走完整的处理流程。基于多核流的实现只要一条流的首包走完整业务处理流程并建立与业务相关的流加速表项(流 ID 与隧道的对应关系),后续属于同一条流的其他包由流分类模块分配到同一 VCPU/CORE,利用首包的成果(已经建立流加速表项)不必查 FIB,直接进入相应隧道处理并转发,提高处理速度的同时也避免了分配到不同核处理时的核间通信开销。除了硬件与软件流分类模块的流分类(相同五元组特征的报文分配相同流 ID 属于同一条流)外,各 VCPU/CORE 除业务处理外也具有流分类功能,它只是为报文打上流标签(流 ID)不进行报文分发,避免流分类模块的性能瓶颈。

多核方式下隧道接口信息与统计信息等临界资源的保护。针对不同的临界资源选用不同的锁保护技术,尽量缩小锁的作用空间和时间。对于读操作多的统计信息采用读写锁与资源分散存储,对于写操作占有性要求高的流加速表采用自旋锁,对于不能在同一个指令周期完成的统计计数运算操作使用原子操作以保证数据操作的一致性,对于读/写操作不在同一层次的隧道哈希表采用延迟删除机制,以确保写入新数据时不会影响到正在引用旧数据的业务处理的中断。

为确保良好的封装性,将隧道模块定义为一个抽象的接口,包括隧道协议模块和隧道管理模块。隧道协议模块又抽象为包括以上四种不同隧道模式对应的隧道业务,由隧道管理模块来控制隧道协议模块的工作,在初始化时由其向隧道依赖的其他周边模块注册隧道协议模块的处理函数以及创建和初始化隧道控制块,同时也向隧道协议模块提供查询和获得统计信息的统一接口,隧道协议模块与其他模块不直接相关,它只进行单纯的业务处理与报文转发,其上的各子业务功能相互独立,并行处理业务。

基于流的隧道转发。报文由转发进入隧道之前,先经流分类模块分配流 ID,然后依据报文的目的地址查 FIB,若下一跳地址是隧道地址,则进入隧道系统依隧道类型执行不同的隧道加封装,因加封装后报文的五元组特征发生变化,所以要重新进行流分类为其分配新的流 ID 并入 IP 队列进行二次转发。报文出隧道时,依据报文头中的下一个协议字段查协议开关表(定义注册在转发模块的隧道协议类型数组),按照不同的隧道协议调用相应隧道解封封装函数解封封装,解封封装后也要重新进行流分类和二次转发,如图2所示。



注:在(5)处,若是首包则走完整处理流程并建立相关流加速表项;若是同一条流的其他包,则利用流加速表项直接进行业务处理。

图2 多核隧道原理

隧道系统的稳定性与可靠性。设置隧道软状态检查,当隧道出现异常时软状态位置位,同时隧道依据乘客报文构造 ICMP6 差错报文发往源端。设置分片检测,依据 RFC2893 中 MTU 发现与分片思想实现对大报文转发的支持。以上三个功能均采用命令行方式注册。

采用主板定时轮询与消息事件驱动的方式来实现多核分布式下的板间同步。主板定时查询消息事件,若资源有更新则向其他各接口板与备板发 IPC 消息同步最新资源。

多核下统计计数的问题。采用资源分散管理,在每个隧道接口所在板分配统计信息资源,统计数为各隧道控制块中统计数的叠加,既对资源实施了保护又确保各隧道统计的独立性。

3 隧道仿真模型及评测数据与分析

3.1 多核隧道系统仿真模型

多核隧道系统采用混合模型,如图3所示,既克服了流水

线模型的可扩展性差及资源利用率低的缺点,同时解决了并行模型对于某些处理开销大且设计复杂的问题,可以针对不同特性的处理选择不同的处理方式。

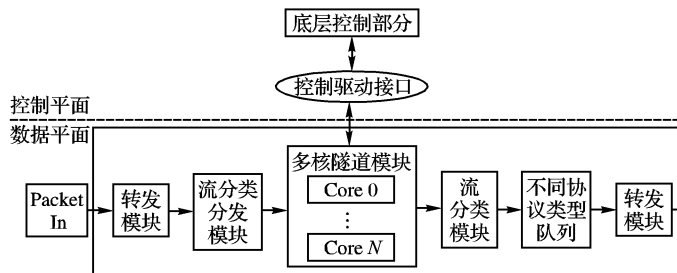


图3 多核隧道系统仿真模型

3.2 测试数据与分析评价

依据上述系统模型与实现原理,利用图4组网进行测试,在PC机A与PC机B之间用SMB打双向包流量,以0丢包时的吞吐量作为转发流量,SMB测试时间为60s,数据精度要求0.1%,其中SMB带宽1000Mbps,多核CPU有8核32线程(VCPU),型号为RMIXLR732。获得数据见表1。

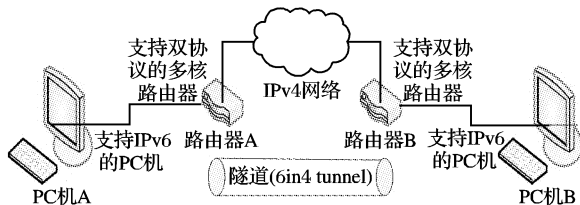


图4 隧道测试组网图

表1 6in4隧道单核与多核下吞吐量(双向)

测试包长/B	吞吐量/pps	
	多核方式	单核方式
128	1332875.840	295113.774
256	842108.243	292792.652
512	450099.859	292739.887
1024	234412.428	233781.754
1280	188251.212	188189.372
1400	173580.546	172962.593

依表1中6in4隧道单核与多核方式下测试数据,得两者的数据折线比较图,如图5所示。

从表1与图5可见,相对于单核方式,多核下隧道对小报文与大报文在转发处理性能上成倍提高,这表明多核方式下

IPv6隧道技术能更高效地解决过渡时期问题。同时采用多核来实现隧道技术更易于进行业务功能的扩展,适应未来网络对高效与业务复杂的要求。

4 结语

多核处理器具有高性能、设计和验证周期短、易扩展、通信延迟低等优点,在推动PC安全性和虚拟技术方面起到关键作用^[6]。IPv6彻底解决了地址空间耗尽和路由表爆炸等问题,协议的设计使路由器处理报文更加简便^[1]。

本文采用多核与隧道技术相结合的方案,能满足目前高速网络处理的同时适应未来网络的发展方向,更好地解决IPv4/IPv6过渡问题。

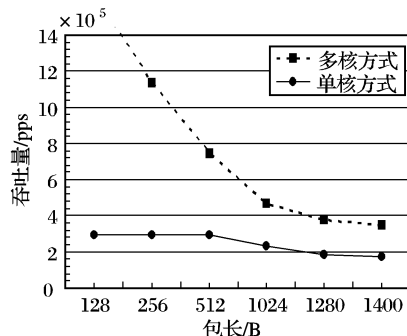


图5 6in4隧道单核与多核方式下测试数据折线图

参考文献:

- [1] DESMEULES R. Cisco ipv6 网络实现技术[M]. 王玲芳, 张宇, 李颖, 等译. 北京: 人民邮电出版社, 2004.
- [2] 王忠培, 周健, 李勇. 基于隧道和 NAT-PT 相结合的 IPv6 过渡方案[J]. 微型计算机, 2008, 24(1/3): 109-111.
- [3] 沈庆伟, 张霖. 基于隧道的 IPv4/IPv6 过渡技术分析[J]. 计算机技术与发展, 2007, 17(5): 171-176.
- [4] 杜慧军. 基于双协议栈的 6to4 隧道技术的应用[J]. 广东技术师范学院学报, 2007(12): 16-20.
- [5] 杭州华三通信技术有限公司. 多核分布式 Net Stream 技术白皮书[Z]. 杭州华三通信技术有限公司, 2007: 6-8.
- [6] 吴继雁. 多核技术与发展趋势[J]. 哈尔滨轴承, 2007, 28(2): 57-58.
- [7] 戴彬, 钱德沛, 刘铁, 等. IPv4/IPv6 协议过渡机制的实验研究[J]. 微电子学与计算机, 2004, 21(9): 64-67.

(上接第 677 页)

从图3和4中可以看出,在出现场景转换的位置, JM 算法中剧烈增加的编码字节数在新算法中得到了很好的控制。说明当图像序列出现场景转换时,和 JM 算法相比,新算法可以更加有效地进行码率控制。

4 结语

本文基于 H.264 视频编码标准的码率控制方案,提出了一种改进的码率控制算法。实验结果表明,改进后的新算法不仅能在各种信道传输中对码率的控制更加准确,而且能在相同的码率下得到比原有算法更高的 PSNR 值,较为明显地提高了解码图像的质量。

参考文献:

- [1] ISO/IEC JTC1/SC29/WG11 MPEG93/457, MPEG-2 Video Test Model, Version 5.0[S]. Genevese: ISO, 1993.

- [2] ISO/IEC JTC1/SC29/WG11 W1796, MPEG-4 Video Verification Model, Version 8.0[S]. Genevese: ISO, 1999.
- [3] ITU-T TMN8, Video Codec Test Model, Near-Term, Version 8.0[S]. Genevese: ISO, 1997.
- [4] WIEGAND T, SULLIVAN G, BJONTEGAARD G. Overview of the H.264/AVC video coding standard[J]. IEEE Transaction on Circuit and Systems for Video Technology, 2003, 13(7): 560-576.
- [5] 葛晓妮, 王宇, 郝重阳, 等. H.264/AVC 率失真优化(RDO)策略研究[J]. 无线通信技术, 2006, 15(2): 14-22.
- [6] 蒋建国, 袁炜, 王德宝. H.264 码率控制算法分析[J]. 电视技术, 2007(5): 10-22.
- [7] 王友炳, 蔡灿辉. 一种 H.264/AVC 码率控制的改进算法[J]. 计算机工程与应用, 2008, 44(11): 199-202.
- [8] KIM M-J, KIM K-H, HONG M-C. Adaptive rate control in frame-layer for real-time H.264/AVC[J]. Advanced Communication Technology, 2008, 3(2): 1875-1880.